Hindawi Publishing Corporation
EURASIP Journal on Wireless Communications and Networking
Volume 2009, Article ID 824235, 29 pages
doi:10.1155/2009/824235

*Research Article*

# Secrecy Capacity of a Class of Broadcast Channels with an Eavesdropper

**Ersen Ekrem and Sennur Ulukus**

*Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742, USA*

Correspondence should be addressed to Sennur Ulukus, ulukus@umd.edu

Received 1 December 2008; Accepted 21 June 2009

Recommended by Shlomo Shamai (Shitz)

We study the security of communication between a single transmitter and many receivers in the presence of an eavesdropper for several special classes of broadcast channels. As the first model, we consider the degraded multireceiver wiretap channel where the legitimate receivers exhibit a degradedness order while the eavesdropper is more noisy with respect to all legitimate receivers. We establish the secrecy capacity region of this channel model. Secondly, we consider the parallel multireceiver wiretap channel with a less noisiness order in each subchannel, where this order is not necessarily the same for all subchannels, and hence the overall channel does not exhibit a less noisiness order. We establish the common message secrecy capacity and sum secrecy capacity of this channel. Thirdly, we study a class of parallel multireceiver wiretap channels with two subchannels, two users and an eavesdropper. For channels in this class, in the first (resp., second) subchannel, the second (resp., first) receiver is degraded with respect to the first (resp., second) receiver, while the eavesdropper is degraded with respect to both legitimate receivers in both subchannels. We determine the secrecy capacity region of this channel, and discuss its extensions to arbitrary numbers of users and subchannels. Finally, we focus on a variant of this previous channel model where the transmitter can use only one of the subchannels at any time. We characterize the secrecy capacity region of this channel as well.

Copyright © 2009 E. Ekrem and S. Ulukus. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## 1. Introduction

Information theoretic secrecy was initiated by Wyner in his seminal work [1], where he introduced the wiretap channel and established the capacity-equivocation region of the *degraded* wiretap channel. Later, his result was generalized to arbitrary, *not necessarily degraded*, wiretap channels by Csiszar and Korner [2]. Recently, many multiuser channel models have been considered from a secrecy point of view [3–22]. One basic extension of the wiretap channel to the multiuser environment is *secure broadcasting to many users* in the presence of an eavesdropper. In the most general form of this problem (see Figure 1), one transmitter wants to have confidential communication with an arbitrary number of users in a broadcast channel, while this communication is being eavesdropped by an external entity. Our goal is to understand the theoretical limits of secure broadcasting, that is, largest simultaneously achievable secure rates. Characterizing the secrecy capacity region of this channel model in its most general form is difficult, because the version of this problem without any secrecy constraints, is the broadcast channel with an arbitrary number of receivers, whose capacity region is open. Consequently, to have progress in understanding the limits of secure broadcasting, we resort to studying several special classes of channels, with increasing generality. The approach of studying special channel structures was also followed in the existing literature on secure broadcasting [8, 9].

The work in [9] first considers an arbitrary wiretap channel with two legitimate receivers and one eavesdropper, and provides an inner bound for achievable rates when each user wishes to receive an independent message. Secondly, [9] focuses on the degraded wiretap channel with two receivers and one eavesdropper, where there is a degradedness order among the receivers, and the eavesdropper is degraded with respect to both users (see Figure 2 for a more general version of the problem that we study). For this setting, the work in [9] finds the secrecy capacity region. This result is

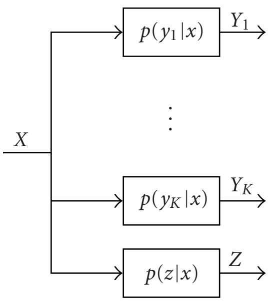

Figure 1: Secure broadcasting to many users in the presence of an eavesdropper.

concurrently and independently obtained in this work as a special case, see Corollary 2, which is also published in a conference version in [23].

Another relevant work on secure broadcasting is [8] which considers secure broadcasting to $K$ users using $M$ subchannels (see Figure 3) for two different scenarios. In the first scenario, the transmitter wants to convey only a common confidential message to all users, and in the second scenario, the transmitter wants to send independent messages to all users. For both scenarios, the work in [8] considers a subclass of parallel multireceiver wiretap channels, where in any given subchannel, there is a degradation order such that each receiver's observation (except the best one) is a degraded version of some other receiver's observation, and this degradation order is not necessarily the same for all subchannels. For the first scenario, the work in [8] finds the common message secrecy capacity for this subclass. For the second scenario, where each user wishes to receive an independent message, [8] finds the sum secrecy capacity for this subclass of channels.

In this paper, our approach will be two-fold: first, we will identify more general channel models than considered in [8, 9] and generalize the results in [8, 9] to those channel models, and secondly, we will consider somewhat more specialized channel models than in [8] and provide more comprehensive results. More precisely, our contributions in this paper are as follows.

(1) We consider the degraded multireceiver wiretap channel with an arbitrary number of users and one eavesdropper, where users are arranged according to a degradedness order, and each user has a less noisy channel with respect to the eavesdropper, see Figure 2. We find the secrecy capacity region when each user receives both an independent message and a common confidential message. Since degradedness implies less noisiness [2], this channel model contains the subclass of channel models where in addition to the degradedness order users exhibit, the eavesdropper is degraded with respect to all users. Consequently, our result can be specialized to the degraded multireceiver wiretap channel with an arbitrary number of users and a degraded eavesdropper, see Corollary 2 and also [23]. The two-user version of the degraded multireceiver wiretap channel was studied and the capacity region was found independently and concurrently in [9].

(2) We then focus on a class of parallel multireceiver wiretap channels with an arbitrary number of legitimate receivers and an eavesdropper, see Figure 3, where in each subchannel, for any given user, either the user's channel is less noisy with respect to the eavesdropper's channel, or vice versa. We establish the common message secrecy capacity of this channel, which is a generalization of the corresponding capacity result in [8] to a broader class of channels. Secondly, we study the scenario where each legitimate receiver wishes to receive an independent message for another subclass of parallel multireceiver wiretap channels. For channels belonging to this subclass, in each subchannel, there is a less noisiness order which is not necessarily the same for all subchannels. Consequently, this ordered class of channels is a subset of the class for which we establish the common message secrecy capacity. We find the sum secrecy capacity for this class, which is again a generalization of the corresponding result in [8] to a broader class of channels.

(3) We also investigate a class of parallel multireceiver wiretap channels with two subchannels, two users, and one eavesdropper, see Figure 4. For the channels in this class, there is a specific degradation order in each subchannel such that in the first (resp., second) subchannel the second (resp., first) user is degraded with respect to the first (resp., second) user, while the eavesdropper is degraded with respect to both users in both subchannels. This is the model of [8] for $K = 2$ users and $M = 2$ subchannels. This model is more restrictive compared to the one mentioned in the previous item. Our motivation to study this more special class is to provide a stronger and more comprehensive result. In particular, for this class, we determine the entire secrecy capacity region when each user receives both an independent message and a common message. In contrast, the work in [8] gives the common message secrecy capacity (when only a common message is transmitted) and sum secrecy capacity (when only independent messages are transmitted) of this class. We discuss the generalization of this result to arbitrary numbers of users and subchannels.

(4) We finally consider a variant of the previous channel model. In this model, we again have a parallel multireceiver wiretap channel with two subchannels, two users, and one eavesdropper, and the degradation order in each subchannel is exactly the same as in the previous item. However, in this case, the input and output alphabets of one subchannel are nonintersecting with the input and output alphabets of the other subchannel. Moreover, we can use only one of these subchannels at any time. We determine the secrecy capacity region of this channel when the transmitter sends both an independent message to each receiver and a common message to both receivers.

It is clear that all of the channel models we consider exhibit some kind of an ordered structure, where this ordered structure is in the form of degradedness in some channel models, and it is in the form of less noisiness in others. This common ordered structure in all channel models we considered implies that our achievability schemes and converse proofs use some common techniques. In particular, for achievability, we use stochastic encoding [2] in conjunction with superposition coding [24]; and for the

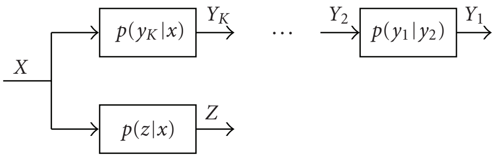

Figure 2: The degraded multireceiver wiretap channel with a more noisy eavesdropper.

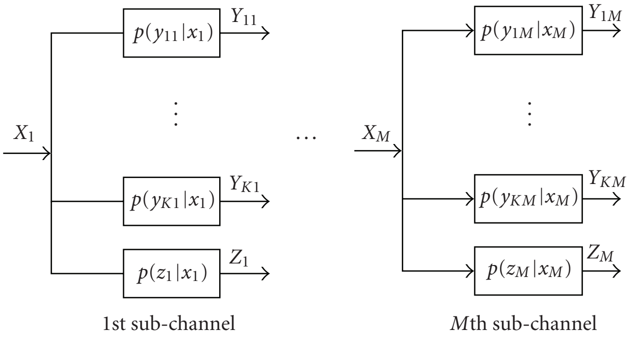

Figure 3: The parallel multireceiver wiretap channel.

converse proofs, we use outer bounding techniques in [1, 2], more specifically, the Csiszar-Korner identity, [2, Lemma 7].

## 2. Degraded Multireceiver Wiretap Channels

We first consider the generalization of Wyner's degraded wiretap channel to the case with many legitimate receivers. In particular, the channel consists of a transmitter with an input alphabet $x \in \mathcal{X}$, $K$ legitimate receivers with output alphabets $y_k \in \mathcal{Y}_k$, $k = 1, \ldots, K$, and an eavesdropper with output alphabet $z \in \mathcal{Z}$. The transmitter sends a confidential message to each user, say $w_k \in \mathcal{W}_k$ to the $k$th user, in addition to a common message, $w_0 \in \mathcal{W}_0$, which is to be delivered to all users. All messages are to be kept secret from the eavesdropper. The channel is assumed to be memoryless with a transition probability $p(y_1, y_2, \ldots, y_K, z \mid x)$.

In this section, we consider a special class of these channels, see Figure 2, where users exhibit a certain degradation order, that is, their channel outputs satisfy the following Markov chain:

$$X \longrightarrow Y_K \longrightarrow \cdots \longrightarrow Y_1 \tag{1}$$

and each user has a less noisy channel with respect to the eavesdropper, that is, we have

$$I(U; Y_k) > I(U; Z) \tag{2}$$

for every $U$ such that $U \rightarrow X \rightarrow (Y_k, Z)$. In fact, since a degradation order exists among the users, it is sufficient to say that user 1 has a less noisy channel with respect to the eavesdropper to guarantee that all users do. Hereafter, we call this channel *the degraded multireceiver wiretap channel with a more noisy eavesdropper*. We note that this channel model contains the degraded multireceiver wiretap channel which is defined through the Markov chain:

$$X \longrightarrow Y_K \longrightarrow \cdots \longrightarrow Y_1 \longrightarrow Z \tag{3}$$

because the Markov chain in (3) implies the less noisiness condition in (2).

A $(2^{nR_0}, 2^{nR_1}, \ldots, 2^{nR_K}, n)$ code for this channel consists of $K+1$ message sets, $\mathcal{W}_k = \{1, \ldots, 2^{nR_k}\}$, $k = 0, 1, \ldots, K$, an encoder $f : \mathcal{W}_0 \times \cdots \times \mathcal{W}_K \rightarrow \mathcal{X}^n$, $K$ decoders, one at each legitimate receiver, $g_k : \mathcal{Y}_k^n \rightarrow \mathcal{W}_0 \times \mathcal{W}_k$, $k = 1, \ldots, K$. The probability of error is defined as $P_e^n = \max_{k=1,\ldots,K} \Pr[g_k(Y_k^n) \neq (W_0, W_k)]$. A rate tuple $(R_0, R_1, \ldots, R_K)$ is said to be achievable if there exists a code with $\lim_{n\rightarrow\infty} P_e^n = 0$ and

$$\lim_{n\rightarrow\infty} \frac{1}{n} H(\mathcal{S}(W) \mid Z^n) \geq \sum_{k\in\mathcal{S}(W)} R_k, \quad \forall \mathcal{S}(W), \tag{4}$$

where $\mathcal{S}(W)$ denotes any subset of $\{W_0, W_1, \ldots, W_K\}$. Hence, we consider only perfect secrecy rates. The secrecy capacity region is defined as the closure of all achievable rate tuples.

The secrecy capacity region of the degraded multireceiver wiretap channel with a more noisy eavesdropper is given by the following theorem whose proof is provided in Appendix A.

**Theorem 1.** *The secrecy capacity region of the degraded multireceiver wiretap channel with a more noisy eavesdropper is given by the union of the rate tuples $(R_0, R_1, \ldots, R_K)$ satisfying*

$$R_0 + \sum_{k=1}^{\ell} R_k \leq \sum_{k=1}^{\ell} I(U_k; Y_k \mid U_{k-1}) - I(U_\ell; Z), \quad \ell = 1, \ldots, K, \tag{5}$$

*where $U_0 = \phi$, $U_K = X$, and the union is over all probability distributions of the form*

$$p(u_1)p(u_2 \mid u_1) \cdots p(u_{K-1} \mid u_{K-2})p(x \mid u_{K-1}). \tag{6}$$

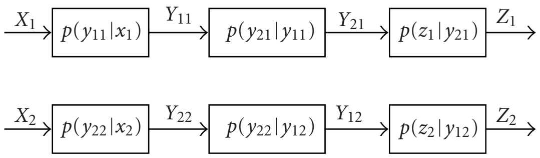

Figure 4: The parallel degraded multireceiver wiretap channel.

*Remark 1.* Theorem 1 implies that a modified version of superposition coding can achieve the boundary of the capacity region. The difference between the superposition coding scheme used to achieve (5) and the standard one in [24], which is used to achieve the capacity region of the degraded broadcast channel, is that the former uses stochastic encoding in each layer of the code to associate each message with many codewords. This controlled amount of redundancy prevents the eavesdropper from being able to decode the message.

As stated earlier, the degraded multireceiver wiretap channel with a more noisy eavesdropper contains the degraded multireceiver wiretap channel which requires the eavesdropper to be degraded with respect to all users as stated in (3). Thus, we can specialize our result in Theorem 1 to the degraded multireceiver wiretap channel as given in the following corollary.

**Corollary 2.** *The secrecy capacity region of the degraded multireceiver wiretap channel is given by the union of the rate tuples* $(R_0, R_1, \ldots, R_K)$ *satisfying*

$$R_0 + \sum_{k=1}^{\ell} R_k \leq \sum_{k=1}^{\ell} I(U_k; Y_k \mid U_{k-1}, Z), \qquad \ell = 1, \ldots, K, \tag{7}$$

*where* $U_0 = \phi, U_K = X$, *and the union is over all probability distributions of the form*

$$p(u_1)p(u_2 \mid u_1) \cdots p(u_{K-1} \mid u_{K-2})p(x \mid u_{K-1}). \tag{8}$$

The proof of this corollary can be carried out from Theorem 1 by noting the following identity:

$$I(U_\ell; Z) = \sum_{k=1}^{\ell} I(U_k; Z \mid U_{k-1}), \tag{9}$$

and the following Markov chains:

$$U_{k-1} \longrightarrow U_k \longrightarrow Y_k \longrightarrow Z, \quad k = 1, \ldots, K. \tag{10}$$

We acknowledge an independent and concurrent work regarding the degraded multireceiver wiretap channel. The work in [9] considers the two-user case and establishes the secrecy capacity region as well.

So far we have determined the entire secrecy capacity region of the degraded multireceiver wiretap channel with a more noisy eavesdropper. This class of channels requires a certain degradation order among the legitimate receivers which may be viewed as being too restrictive from a practical point of view. Our goal is to consider progressively more general channel models. Toward that goal, in the following section, we consider channel models where the users are not ordered in a degradedness or noisiness order. However, the concepts of degradedness and noisiness are essential in proving capacity results. In the following section, we will consider multireceiver broadcast channels which are composed of independent subchannels. We will assume some noisiness properties in these subchannels in order to derive certain capacity results. However, even though the subchannels will have certain noisiness properties, the overall broadcast channel will not have any degradedness or noisiness properties.

## 3. Parallel Multireceiver Wiretap Channels

Here, we investigate the parallel multireceiver wiretap channel where the transmitter communicates with $K$ legitimate receivers using $M$ independent subchannels in the presence of an eavesdropper, see Figure 3. The channel transition probability of a parallel multireceiver wiretap channel is

$$\begin{aligned} &p\left(\{y_{1m}, \ldots, y_{Km}, z_m\}_{m=1}^{M} \mid \{x_m\}_{m=1}^{M}\right) \\ &\quad = \prod_{m=1}^{M} p(y_{1m}, \ldots, y_{Km}, z_m \mid x_m), \end{aligned} \tag{11}$$

where $x_m \in \mathcal{X}_m$ is the input in the $m$th subchannel where $\mathcal{X}_m$ is the corresponding channel input alphabet, $y_{km} \in \mathcal{Y}_{km}$ (resp., $z_m \in \mathcal{Z}_m$) is the output in the $k$th user's (resp., eavesdropper's) $m$th subchannel where $\mathcal{Y}_{km}$ (resp., $\mathcal{Z}_m$) is the $k$th user's (resp., eavesdropper's) $m$th subchannel output alphabet.

We note that the parallel multireceiver wiretap channel can be regarded as an extension of the parallel wiretap channel [21, 22] to the case of multiple legitimate users. Though the work in [21, 22] establishes the secrecy capacity of the parallel wiretap channel for the most general case, for the parallel multireceiver wiretap channel, obtaining the secrecy capacity region for the most general case seems to be intractable for now. Thus, in this section, we investigate special classes of parallel multireceiver wiretap channels. These channel models contain the class of channel models studied in [8] as a special case. Similar to [8], our emphasis will be on the common message secrecy capacity and the sum secrecy capacity.

### 3.1. The Common Message Secrecy Capacity.

We first consider the simplest possible scenario where the transmitter sends a common confidential message to all users. Despite its simplicity, the secrecy capacity of a common confidential message (hereafter will be called the common message secrecy capacity) in a general broadcast channel is unknown.

The common message secrecy capacity for a special class of parallel multireceiver wiretap channels was studied in [8]. In this class of parallel multireceiver wiretap channels [8], each subchannel exhibits a certain degradation order which is not necessarily the same for all subchannels, that is, the following Markov chain is satisfied:

$$X_l \longrightarrow Y_{\pi_l(1)} \longrightarrow Y_{\pi_l(2)} \longrightarrow \cdots \longrightarrow Y_{\pi_l(K+1)} \tag{12}$$

in the $l$th subchannel, where $(Y_{\pi_l(1)}, Y_{\pi_l(2)}, \ldots, Y_{\pi_l(K+1)})$ is a permutation of $(Y_{1l}, \ldots, Y_{Kl}, Z_l)$. Hereafter, we call this channel the parallel degraded multireceiver wiretap channel.( In [8], these channels are called *reversely degraded* parallel channels. Here, we call them parallel degraded multireceiver wiretap channels to be consistent with the terminology used in the rest of the paper.) Although [8] established the common message secrecy capacity for this class of channels, in fact, their result is valid for the broader class in which we have either

$$X_l \longrightarrow Y_{kl} \longrightarrow Z_l \tag{13}$$

or

$$X_l \longrightarrow Z_l \longrightarrow Y_{kl} \tag{14}$$

valid for every $X_l$ and for any $(k, l)$ pair where $k \in \{1, \ldots, K\}$, $l \in \{1, \ldots, M\}$. Thus, it is sufficient to have a degradedness order between each user and the eavesdropper in any subchannel instead of the long Markov chain between all users and the eavesdropper as in (12).

Here, we focus on a broader class of channels where in each subchannel, for any given user, either the user's channel is less noisy than the eavesdropper's channel or vice versa. More formally, we have either

$$I(U; Y_{kl}) > I(U; Z_l) \tag{15}$$

or

$$I(U; Y_{kl}) < I(U; Z_l) \tag{16}$$

for all $U \rightarrow X_l \rightarrow (Y_{kl}, Z)$ and any $(k, l)$ pair where $k \in \{1, \ldots, K\}$, $l \in \{1, \ldots, M\}$. Hereafter, we call this channel *the parallel multireceiver wiretap channel with a more noisy eavesdropper*. Since the Markov chain in (12) implies either (15) or (16), the parallel multireceiver wiretap channel with a more noisy eavesdropper contains the parallel degraded multireceiver wiretap channel studied in [8].

A $(2^{nR}, n)$ code for this channel consists of a message set, $\mathcal{W}_0 = \{1, \ldots, 2^{nR}\}$, an encoder, $f : \mathcal{W}_0 \rightarrow \mathcal{X}_1^n \times \cdots \times \mathcal{X}_M^n$, $K$ decoders, one at each legitimate receiver $g_k : \mathcal{Y}_{k1}^n \times \cdots \times \mathcal{Y}_{kM}^n \rightarrow \mathcal{W}_0, k = 1, \ldots, K$. The probability of error is defined as $P_e^n = \max_{k=1,\ldots,K} \Pr[\widehat{W}_{k0} \neq W_0]$, where $\widehat{W}_{k0}$ is the $k$th user's decoder output. The secrecy of the common message is measured through the equivocation rate which is defined as $(1/n)H(W_0 \mid Z_1^n, \ldots, Z_M^n)$. A common message secrecy rate, $R$, is said to be achievable if there exists a code such that $\lim_{n\rightarrow\infty} P_e^n = 0$, and

$$\lim_{n\rightarrow\infty} \frac{1}{n} H(W_0 \mid Z_1^n, \ldots, Z_M^n) \geq R. \tag{17}$$

The common message secrecy capacity is the supremum of all achievable secrecy rates.

The common message secrecy capacity of the parallel multireceiver wiretap channel with a more noisy eavesdropper is stated in the following theorem whose proof is given in Appendix B.

**Theorem 3.** *The common message secrecy capacity, $C_0$, of the parallel multireceiver wiretap channel with a more noisy eavesdropper is given by*

$$C_0 = \max \min_{k=1,\ldots,K} \sum_{l=1}^{M} [I(X_l; Y_{kl}) - I(X_l; Z_l)]^+, \tag{18}$$

*where the maximization is over all distributions of the form $p(x_1, \ldots, x_M) = \prod_{l=1}^{M} p(x_l)$.*

*Remark 2.* Theorem 3 implies that we should not use the subchannels in which there is no user that has a less noisy channel than the eavesdropper. Moreover, Theorem 3 shows that the use of independent inputs in each subchannel is sufficient to achieve the capacity, that is, inducing correlation between channel inputs of subchannels cannot provide any improvement. This is similar to the results of [25, 26] in the sense that the work in [25, 26] established the optimality of the use of independent inputs in each subchannel for the product of two degraded broadcast channels.

As stated earlier, the parallel multireceiver wiretap channel with a more noisy eavesdropper encompasses the parallel degraded multireceiver wiretap channel studied in [8]. Hence, we can specialize Theorem 3 to recover the common message secrecy capacity of the parallel degraded multireceiver wiretap channel established in [8]. This is stated in the following corollary whose proof can be carried out from Theorem 3 by noting the Markov chain $X_l \rightarrow Y_{kl} \rightarrow Z_l$, for all $(k, l)$.

**Corollary 4.** *The common message secrecy capacity of the parallel degraded multireceiver wiretap channel is given by*

$$C_0 = \max \min_{k=1,\ldots,K} \sum_{l=1}^{M} I(X_l; Y_{kl} \mid Z_l), \tag{19}$$

*where the maximization is over all distributions of the form $p(x_1, \ldots, x_M) = \prod_{l=1}^{M} p(x_l)$.*

### *3.2. The Sum Secrecy Capacity.*

We now consider the scenario where the transmitter sends an independent confidential message to each legitimate receiver, and focus on the sum secrecy capacity. We consider a class of parallel multireceiver wiretap channels where the legitimate receivers and the eavesdropper exhibit a certain less noisiness order in each subchannel. These less noisiness orders are not necessarily the same for all subchannels. Therefore, the overall channel does not have a less noisiness order. In the $l$th subchannel, for all $U \rightarrow X_l \rightarrow (Y_{1l}, \ldots, Y_{Kl}, Z_l)$, we have

$$I(U; Y_{\pi_l(1)}) > I(U; Y_{\pi_l(2)}) > \cdots > I(U; Y_{\pi_l(K+1)}), \tag{20}$$

where $(Y_{\pi_l(1)}, Y_{\pi_l(2)}, \ldots, Y_{\pi_l(K+1)})$ is a permutation of $(Y_{1l}, \ldots, Y_{Kl}, Z_l)$. We call this channel *the parallel multireceiver wiretap channel with a less noisiness order in each subchannel.* We note that this class of channels is a subset of the parallel multireceiver wiretap channel with a more noisy eavesdropper studied in Section 3.1, because of the additional ordering imposed between users' subchannels. We also note that the class of parallel degraded multireceiver wiretap channels with a degradedness order in each subchannel studied in [8] is not only a subset of parallel multireceiver wiretap channels with a more noisy eavesdropper studied in Section 3.1 but also a subset of parallel multireceiver wiretap channels with a less noisiness order in each subchannel studied in this section.

A $(2^{nR_1}, \ldots, 2^{nR_K}, n)$ code for this channel consists of $K$ message sets, $\mathcal{W}_k = \{1, \ldots, 2^{nR_k}\}$, $k = 1, \ldots, K$, an encoder, $f : \mathcal{W}_1 \times \cdots \times \mathcal{W}_K \rightarrow \mathcal{X}_1^n \times \cdots \times \mathcal{X}_M^n$, $K$ decoders, one at each legitimate receiver $g_k : \mathcal{Y}_{k1}^n \times \cdots \times \mathcal{Y}_{kM}^n \rightarrow \mathcal{W}_k, k = 1, \ldots, K$. The probability of error is defined as $P_e^n = \max_{k=1,\ldots,K} \Pr[\widehat{W}_k \neq W_k]$, where $\widehat{W}_k$ is the $k$th user's decoder output. The secrecy is measured through the equivocation rate which is defined as $(1/n)H(W_1, \ldots, W_K \mid Z_1^n, \ldots, Z_M^n)$. A sum secrecy rate, $R_s$, is said to be achievable if there exists a code such that $\lim_{n\rightarrow\infty} P_e^n = 0$, and

$$\lim_{n\rightarrow\infty} \frac{1}{n} H(W_1, \ldots, W_K \mid Z_1^n, \ldots, Z_M^n) \geq R_s. \tag{21}$$

The sum secrecy capacity is defined to be the supremum of all achievable sum secrecy rates.

The sum secrecy capacity for the class of parallel multireceiver wiretap channels with a less noisiness order in each subchannel studied in this section is stated in the following theorem whose proof is given in Appendix C.

**Theorem 5.** *The sum secrecy capacity of the parallel multireceiver wiretap channel with a less noisiness order in each subchannel is given by*

$$\max \sum_{l=1}^{M} \left[ I\left(X_l; Y_{\rho(l)l}\right) - I(X_l; Z_l) \right]^+, \tag{22}$$

*where the maximization is over all input distributions of the form $p(x_1, \ldots, x_M) = \prod_{l=1}^{M} p(x_l)$ and $\rho(l)$ denotes the index of the strongest user in the lth subchannel in the sense that*

$$I(U; Y_{kl}) \leq I\left(U; Y_{\rho(l)l}\right) \tag{23}$$

*for all $U \rightarrow X_l \rightarrow (Y_{1l}, \ldots, Y_{Kl}, Z_l)$ and any $k \in \{1, \ldots, K\}$.*

*Remark 3.* Theorem 5 implies that the sum secrecy capacity is achieved by sending information only to the strongest user in each subchannel. As in Theorem 3, here also, the use of independent inputs for each subchannel is capacity-achieving, which is again reminiscent of the result in [25, 26] about the optimality of the use of independent inputs in each subchannel for the product of two degraded broadcast channels.

As mentioned earlier, since the class of parallel multireceiver wiretap channels with a less noisiness order in each subchannel contains the class of parallel degraded multireceiver wiretap channels studied in [8], Theorem 5 can be specialized to give the sum secrecy capacity of the latter class of channels as well. This result was originally obtained in [8]. This is stated in the following corollary. Since the proof of this corollary is similar to the proof of Corollary 4, we omit its proof.

**Corollary 6.** *The sum secrecy capacity of the parallel degraded multireceiver wiretap channel is given by*

$$\max \sum_{l=1}^{M} I\left(X_l; Y_{\rho(l)l} \mid Z_l\right), \tag{24}$$

*where the maximization is over all input distributions of the form $p(x_1, \ldots, x_M) = \prod_{l=1}^{M} p(x_l)$ and $\rho(l)$ denotes the index of the strongest user in the lth subchannel in the sense that*

$$X_l \longrightarrow Y_{\rho(l)l} \longrightarrow Y_{kl} \tag{25}$$

*for all input distributions on $X_l$ and any $k \in \{1, \ldots, K\}$.*

So far, we have considered special classes of parallel multireceiver wiretap channels for specific scenarios and obtained results similar to [8], only for broader classes of channels. In particular, in Section 3.1, we focused on the transmission of a common message, whereas in Section 3.2, we focused on the sum secrecy capacity when only independent messages are transmitted to all users. In the subsequent sections, we will specialize our channel model, but we will develop stronger and more comprehensive results. In particular, we will let the transmitter send both common and independent messages, and we will characterize the entire secrecy capacity region.

## 4. Parallel Degraded Multireceiver Wiretap Channels

We consider a special class of parallel degraded multireceiver wiretap channels with two subchannels, two users, and one eavesdropper. We consider the most general scenario where each user receives both an independent message and a common message. All messages are to be kept secret from the eavesdropper.

For the special class of parallel degraded multireceiver wiretap channels in consideration, there is a specific degradation order in each subchannel. In particular, we have the following Markov chain:

$$X_1 \longrightarrow Y_{11} \longrightarrow Y_{21} \longrightarrow Z_1 \tag{26}$$

in the first subchannel, and the following Markov chain:

$$X_2 \longrightarrow Y_{22} \longrightarrow Y_{12} \longrightarrow Z_2 \tag{27}$$

in the second subchannel. Consequently, although in each subchannel, one user is degraded with respect to the other one, this does not hold for the overall channel, and the overall channel is not degraded for any user. The corresponding channel transition probability is

$$\begin{aligned} &p(y_{11} \mid x_1)p(y_{21} \mid y_{11}) \\ &\quad \times p\left(z_1 \mid y_{21}\right)p(y_{22} \mid x_2)p\left(y_{12} \mid y_{22}\right)p(z_2 \mid y_{12}). \end{aligned} \tag{28}$$

If we ignore the eavesdropper by setting $Z_1 = Z_2 = \phi$, this channel model reduces to the broadcast channel that was studied in [25, 26].

A $(2^{nR_0}, 2^{nR_1}, 2^{nR_2}, n)$ code for this channel consists of three message sets, $\mathcal{W}_0 = \{1, \dots, 2^{nR_0}\}$, $\mathcal{W}_j = \{1, \dots, 2^{nR_j}\}$, $j = 1, 2$, one encoder $f : \mathcal{W}_0 \times \mathcal{W}_1 \times \mathcal{W}_2 \rightarrow \mathcal{X}_1^n \times \mathcal{X}_2^n$, two decoders one at each legitimate receiver $g_j : \mathcal{Y}_{j1}^n \times \mathcal{Y}_{j2}^n \rightarrow \mathcal{W}_0 \times \mathcal{W}_j$, $j = 1, 2$. The probability of error is defined as $P_e^n = \max_{j=1,2} \Pr[g_j(Y_{j1}^n, Y_{j2}^n) \neq (W_0, W_j)]$. A rate tuple $(R_0, R_1, R_2)$ is said to be achievable if there exists a code such that $\lim_{n\rightarrow\infty} P_e^n = 0$ and

$$\lim_{n\rightarrow\infty} \frac{1}{n} H(\mathcal{S}(W) \mid Z_1^n, Z_2^n) \geq \sum_{k \in \mathcal{S}(W)} R_k, \quad \forall \mathcal{S}(W), \tag{29}$$

where $\mathcal{S}(W)$ denotes any subset of $\{W_0, W_1, W_2\}$. The secrecy capacity region is the closure of all achievable secrecy rate tuples.

The secrecy capacity region of this parallel degraded multireceiver wiretap channel is characterized by the following theorem whose proof is given in Appendix D.1.

**Theorem 7.** *The secrecy capacity region of the parallel degraded multireceiver wiretap channel defined by* (28) *is the union of the rate tuples* $(R_0, R_1, R_2)$ *satisfying*

$$\begin{aligned} R_0 &\leq I(U_1; Y_{11} \mid Z_1) + I(U_2; Y_{12} \mid Z_2), \\ R_0 &\leq I(U_1; Y_{21} \mid Z_1) + I(U_2; Y_{22} \mid Z_2), \\ R_0 + R_1 &\leq I(X_1; Y_{11} \mid Z_1) + I(U_2; Y_{12} \mid Z_2), \\ R_0 + R_2 &\leq I(X_2; Y_{22} \mid Z_2) + I(U_1; Y_{21} \mid Z_1), \\ R_0 + R_1 + R_2 &\leq I(X_1; Y_{11} \mid Z_1) + I(U_2; Y_{12} \mid Z_2) \\ &\qquad + I(X_2; Y_{22} \mid U_2, Z_2), \\ R_0 + R_1 + R_2 &\leq I(X_2; Y_{22} \mid Z_2) + I(U_1; Y_{21} \mid Z_1) \\ &\qquad + I(X_1; Y_{11} \mid U_1, Z_1), \end{aligned} \tag{30}$$

*where the union is over all distributions of the form* $p(u_1, u_2, x_1, x_2) = p(u_1, x_1)p(u_2, x_2)$.

*Remark 4.* If we let the encoder use an arbitrary joint distribution $p(u_1, x_1, u_2, x_2)$ instead of the ones that satisfy $p(u_1, x_1, u_2, x_2) = p(u_1, x_1)p(u_2, x_2)$, this would not enlarge the region given in Theorem 7, because all rate expressions in Theorem 7 depend on either $p(u_1, x_1)$ or $p(u_2, x_2)$ but not on the joint distribution $p(u_1, u_2, x_1, x_2)$.

*Remark 5.* The capacity-achieving scheme uses either superposition coding in both subchannels or superposition coding in one of the subchannels, and a dedicated transmission in the other one. We again note that this superposition coding is different from the standard one [24] in the sense that it associates each message with many codewords by using stochastic encoding at each layer of the code due to secrecy concerns.

*Remark 6.* If we set $Z_1 = Z_2 = \phi$, we recover the capacity region of the underlying broadcast channel [26].

*Remark 7.* If we disable one of the subchannels, say the first one, by setting $Y_{11} = Y_{21} = Z_1 = \phi$, the parallel degraded multireceiver wiretap channel of this section reduces to the degraded multireceiver wiretap channel of Section 2. The corresponding secrecy capacity region is then given by the union of the rate tuples $(R_0, R_1, R_2)$ satisfying

$$\begin{aligned} R_0 + R_1 &\leq I(U_2; Y_{12} \mid Z_2) \\ R_0 + R_1 + R_2 &\leq I(X_2; Y_{22} \mid U_2, Z_2) + I(U_2; Y_{12} \mid Z_2), \end{aligned} \tag{31}$$

where the union is over all $p(u_2, x_2)$. This region can be obtained through either Corollary 2 or Theorem 7 (by setting $Y_{11} = Y_{21} = Z_1 = \phi$ and eliminating redundant bounds) implying the consistency of the results.

Next, we consider the scenario where the transmitter does not send a common message, and find the secrecy capacity region.

**Corollary 8.** *The secrecy capacity region of the parallel degraded multireceiver wiretap channel defined by* (28) *with no common message is given by the union of the rate pairs* $(R_1, R_2)$ *satisfying*

$$\begin{aligned} R_1 &\leq I(X_1; Y_{11} \mid Z_1) + I(U_2; Y_{12} \mid Z_2), \\ R_2 &\leq I(X_2; Y_{22} \mid Z_2) + I(U_1; Y_{21} \mid Z_1), \\ R_1 + R_2 &\leq I(X_1; Y_{11} \mid Z_1) + I(U_2; Y_{12} \mid Z_2) \\ &\qquad + I(X_2; Y_{22} \mid U_2, Z_2), \\ R_1 + R_2 &\leq I(X_2; Y_{22} \mid Z_2) + I(U_1; Y_{21} \mid Z_1) \\ &\qquad + I(X_1; Y_{11} \mid U_1, Z_1), \end{aligned} \tag{32}$$

*where the union is over all distributions of the form* $p(u_1)p(u_2)p(x_1 \mid u_1)p(x_2 \mid u_2)$.

*Proof.* Since the common message rate can be exchanged with any user's independent message rate, we set $R_0 = \alpha + \beta, R_1' = R_1 + \alpha, R_2' = R_2 + \beta$, where $\alpha, \beta \geq 0$. Plugging these expressions into the rates in Theorem 7 and using Fourier-Moztkin elimination, we get the region given in the corollary. □

*Remark 8.* If we disable the eavesdropper by setting $Z_{11} = Z_{22} = \phi$, we recover the capacity region of the underlying broadcast channel without a common message, which was found originally in [25].

At this point, one may ask whether the results of this section can be extended to arbitrary numbers of users and parallel subchannels. Once we have Theorem 7, the extension of the results to an arbitrary number of parallel subchannels is rather straightforward. Let us consider the parallel degraded multireceiver wiretap channel with $M$ subchannels, and in each subchannel, we have either the following Markov chain:

$$X_l \longrightarrow Y_{1l} \longrightarrow Y_{2l} \longrightarrow Z_l, \tag{33}$$

or this Markov chain:

$$X_l \longrightarrow Y_{2l} \longrightarrow Y_{1l} \longrightarrow Z_l \tag{34}$$

for any $l \in \{1, \ldots, M\}$. We define the set of indices $\mathcal{S}_1$ (resp., $\mathcal{S}_2$) as those where for every $l \in \mathcal{S}_1$ (resp., $l \in \mathcal{S}_2$), the Markov chain in (33) (resp., in (34)) is satisfied. Then, using Theorem 7, we obtain the secrecy capacity region of the channel with two users and $M$ subchannels as given in the following theorem which is proved in Appendix D.2.

**Theorem 9.** *The secrecy capacity region of the parallel degraded multireceiver wiretap channel with $M$ subchannels, where each subchannel satisfies either* (33) *or* (34), *is given by the union of the rate tuples $(R_0, R_1, R_2)$ satisfying*

$$\begin{aligned}
R_0 &\leq \sum_{l=1}^{M} I(U_l; Y_{1l} \mid Z_l), \\
R_0 &\leq \sum_{l=1}^{M} I(U_l; Y_{2l} \mid Z_l), \\
R_0 + R_1 &\leq \sum_{l\in\mathcal{S}_1} I(X_l; Y_{1l} \mid Z_l) + \sum_{l\in\mathcal{S}_2} I(U_l; Y_{1l} \mid Z_l), \\
R_0 + R_2 &\leq \sum_{l\in\mathcal{S}_2} I(X_l; Y_{2l} \mid Z_l) + \sum_{l\in\mathcal{S}_1} I(U_l; Y_{2l} \mid Z_l), \\
R_0 + R_1 + R_2 &\leq \sum_{l\in\mathcal{S}_1} I(X_l; Y_{1l} \mid Z_l) + \sum_{l\in\mathcal{S}_2} I(U_l; Y_{1l} \mid Z_l) \\
&\quad + \sum_{l\in\mathcal{S}_2} I(X_l; Y_{2l} \mid U_l, Z_l), \\
R_0 + R_1 + R_2 &\leq \sum_{l\in\mathcal{S}_2} I(X_l; Y_{2l} \mid Z_l) + \sum_{l\in\mathcal{S}_1} I(U_l; Y_{2l} \mid Z_l) \\
&\quad + \sum_{l\in\mathcal{S}_1} I(X_l; Y_{1l} \mid U_l, Z_l),
\end{aligned} \tag{35}$$

*where the union is over all distributions of the form $\prod_{l=1}^{M} p(u_l, x_l)$.*

We are now left with the question whether these results can be generalized to an arbitrary number of users. If we consider the parallel degraded multireceiver wiretap channel with more than two subchannels and an arbitrary number of users, the secrecy capacity region for the scenario where each user receives a common message in addition to an independent message does not seem to be characterizable. Our intuition comes from the fact that, as of now, the capacity region of the corresponding broadcast channel without secrecy constraints is unknown [27]. However, if we consider the scenario where each user receives only an independent message, that is, there is no common message, then the secrecy capacity region may be found, because the capacity region of the corresponding broadcast channel without secrecy constraints can be established [27], although there is no explicit expression for it in literature. We expect this particular generalization to be rather straightforward, and do not pursue it here.

## 5. Sum of Degraded Multireceiver Wiretap Channels

We now consider a different multireceiver wiretap channel which can be viewed as a sum of two degraded multireceiver wiretap channels with two users and one eavesdropper. In this channel model, the transmitter has two nonintersecting input alphabets, that is, $\mathcal{X}_1, \mathcal{X}_2$ with $\mathcal{X}_1 \cap \mathcal{X}_2 = \emptyset$, and each receiver has two nonintersecting alphabets, that is, $\mathcal{Y}_{j1}, \mathcal{Y}_{j2}$ with $\mathcal{Y}_{j1} \cap \mathcal{Y}_{j2} = \emptyset$ for the $j$th user, $j = 1, 2$, and $\mathcal{Z}_1, \mathcal{Z}_2$ with $\mathcal{Z}_1 \cap \mathcal{Z}_2 = \emptyset$ for the eavesdropper. The channel is again memoryless with transition probability

$$\begin{aligned}
&p(y_1, y_2, z \mid x) \\
&= \begin{cases}
p(y_{11} \mid x_1) p(y_{21} \mid y_{11}) p(z_1 \mid y_{21}) \\
\qquad \text{if } (x, y_1, y_2, z) \in \mathcal{X}_1 \times \mathcal{Y}_{11} \times \mathcal{Y}_{21} \times \mathcal{Z}_1, \\
p(y_{22} \mid x_2) p(y_{12} \mid y_{22}) p(z_2 \mid y_{12}) \\
\qquad \text{if } (x, y_1, y_2, z) \in \mathcal{X}_2 \times \mathcal{Y}_{12} \times \mathcal{Y}_{22} \times \mathcal{Z}_2, \\
0 \qquad \text{otherwise,}
\end{cases}
\end{aligned} \tag{36}$$

where $x \in \mathcal{X} = \mathcal{X}_1 \cup \mathcal{X}_2$, $y_j \in \mathcal{Y}_j = \mathcal{Y}_{j1} \cup \mathcal{Y}_{j2}, j = 1, 2$ and $z \in \mathcal{Z} = \mathcal{Z}_1 \cup \mathcal{Z}_2$. Thus, if the transmitter chooses to use its first alphabet, that is, $\mathcal{X}_1$, the second user (resp. eavesdropper) receives a degraded version of user 1's (resp., user 2's) observation. However, if the transmitter uses its second alphabet, that is, $\mathcal{X}_2$, the first user (resp. eavesdropper) receives a degraded version of user 2's (resp. user 1's) observation. Consequently, the overall channel is not degraded from any user's perspective, however, it is degraded from eavesdropper's perspective.

A $(2^{nR_0}, 2^{nR_1}, 2^{nR_2}, n)$ code for this channel consists of three message sets, $\mathcal{W}_0 = \{1, \ldots, 2^{nR_0}\}$, $\mathcal{W}_j = \{1, \ldots, 2^{nR_j}\}$, $j = 1, 2$, one encoder $f : \mathcal{W}_0 \times \mathcal{W}_1 \times \mathcal{W}_2 \rightarrow \mathcal{X}^n$ and two decoders, one at each legitimate receiver, $g_j : \mathcal{Y}_j^n \rightarrow \mathcal{W}_0 \times \mathcal{W}_j$, $j = 1, 2$. The probability of error is defined as $P_e^n = \max_{j=1,2} \Pr[g_j(Y_j^n) \neq (W_0, W_j)]$. A rate tuple $(R_0, R_1, R_2)$ is said to be achievable if there exists a code with $\lim_{n\to\infty} P_e^n = 0$ and

$$\lim_{n\to\infty} \frac{1}{n} H(\mathcal{S}(W) \mid Z^n) \geq \sum_{j\in\mathcal{S}(W)} R_j, \quad \forall \mathcal{S}(W), \tag{37}$$

where $\mathcal{S}(W)$ denotes any subset of $\{W_0, W_1, W_2\}$. The secrecy capacity region is the closure of all achievable secrecy rate tuples.

The secrecy capacity region of this channel is given in the following theorem which is proved in Appendix E.

**Theorem 10.** *The secrecy capacity region of the sum of two degraded multireceiver wiretap channels is given by the union of the rate tuples* $(R_0, R_1, R_2)$ *satisfying*

$$\begin{aligned}
R_0 &\le \alpha I(U_1; Y_{11} \mid Z_1) + \overline{\alpha} I(U_2; Y_{12} \mid Z_2),\\
R_0 &\le \alpha I(U_1; Y_{21} \mid Z_1) + \overline{\alpha} I(U_2; Y_{22} \mid Z_2),\\
R_0 + R_1 &\le \alpha I(X_1; Y_{11} \mid Z_1) + \overline{\alpha} I(U_2; Y_{12} \mid Z_2),\\
R_0 + R_2 &\le \alpha I(U_1; Y_{21} \mid Z_1) + \overline{\alpha} I(X_2; Y_{22} \mid Z_2),\\
R_0 + R_1 + R_2 &\le \alpha I(X_1; Y_{11} \mid Z_1) + \overline{\alpha} I(U_2; Y_{12} \mid Z_2)\\
&\quad + \overline{\alpha} I(X_2; Y_{22} \mid U_2, Z_2),\\
R_0 + R_1 + R_2 &\le \alpha I(U_1; Y_{21} \mid Z_1) + \alpha I(X_1; Y_{11} \mid U_1, Z_1)\\
&\quad + \overline{\alpha} I(X_2; Y_{22} \mid Z_2),
\end{aligned} \tag{38}$$

*where the union is over all* $\alpha \in [0,1]$ *and distributions of the form* $p(u_1, u_2, x_1, x_2) = p(u_1, x_1) p(u_2, x_2)$.

*Remark 9.* This channel model is similar to the parallel degraded multireceiver wiretap channel of the previous section in the sense that it can be viewed to consist of two parallel subchannels, however, now the transmitter cannot use both subchannels simultaneously. Instead, it should invoke a time-sharing approach between these two so-called parallel subchannels ($\alpha$ reflects this concern). Moreover, superposition coding scheme again achieves the boundary of the secrecy capacity region, however, it differs from the standard one [24] in the sense that it needs to be modified to incorporate secrecy constraints, that is, it needs to use stochastic encoding to associate each message with multiple codewords.

*Remark 10.* An interesting point about the secrecy capacity region is that if we drop the secrecy constraints by setting $Z_1 = Z_2 = \phi$, we are unable to recover the capacity region of the corresponding broadcast channel that was found in [26]. After setting $Z_1 = Z_2 = \phi$, we note that each expression in Theorem 10 and its counterpart describing the capacity region [26] differ by exactly $h(\alpha)$. The reason for this is as follows. Here, $\alpha$ not only denotes the time-sharing variable but also carries an additional information, that is, the change of the channel that is in use is part of the information transmission. However, since the eavesdropper can also decode these messages, the term $h(\alpha)$, which is the amount of information that can be transmitted via changes of the channel in use, disappears in the secrecy capacity region.

## 6. Conclusions

In this paper, we studied secure broadcasting to many users in the presence of an eavesdropper. Characterizing the secrecy capacity region of this channel in its most general form seems to be intractable for now, since the version of this problem without any secrecy constraints is the broadcast channel with an arbitrary number of receivers, whose capacity region is open. Consequently, we took the approach of considering special classes of channels. In particular, we considered degraded multireceiver wiretap channels, parallel multireceiver wiretap channels with a more noisy eavesdropper, parallel multireceiver wiretap channels with less noisiness orderings in each subchannel, and parallel degraded multireceiver wiretap channels. For each channel model, we obtained either partial characterization of the secrecy capacity region or the entire region.

## Appendices

## A. Proof of Theorem 1

First, we show achievability, then provide the converse.

*A.1. Achievability.* Fix the probability distribution as

$$p(u_1)p(u_2 \mid u_1) \cdots p(u_{K-1} \mid u_{K-2}) p(x \mid u_{K-1}). \tag{A.1}$$

*Codebook Generation.*

(i) Generate $2^{n(R_0+R_1+\widetilde{R}_1)}$ length-$n$ sequences $\mathbf{u}_1$ through $p(\mathbf{u}_1) = \prod_{i=1}^{n} p(u_{1,i})$ and index them as $\mathbf{u}_1(w_0, w_1, \widetilde{w}_1)$ where $w_0 \in \{1, \dots, 2^{nR_0}\}$, $w_1 \in \{1, \dots, 2^{nR_1}\}$ and $\widetilde{w}_1 \in \{1, \dots, 2^{n\widetilde{R}_1}\}$.

(ii) For each $\mathbf{u}_{j-1}$, where $j = 2, \dots, K-1$, generate $2^{n(R_j+\widetilde{R}_j)}$ length-$n$ sequences $\mathbf{u}_j$ through $p(\mathbf{u}_j|\mathbf{u}_{j-1}) = \prod_{i=1}^{n} p(u_{j,i} \mid u_{j-1,i})$ and index them as $\mathbf{u}_j(w_0, w_1, \dots, w_j, \widetilde{w}_1, \dots, \widetilde{w}_j)$, where $w_j \in \{1, \dots, 2^{nR_j}\}$ and $\widetilde{w}_j \in \{1, \dots, 2^{n\widetilde{R}_j}\}$.

(iii) Finally, for each $\mathbf{u}_{K-1}$, generate $2^{n(R_K+\widetilde{R}_K)}$ length-$n$ sequences $\mathbf{x}$ through $p(\mathbf{x} \mid \mathbf{u}_{K-1}) = \prod_{i=1}^{n} p(x_i \mid u_{K,i})$ and index them as $\mathbf{x}(w_0, w_1, \dots, w_K, \widetilde{w}_1, \dots, \widetilde{w}_K)$ where $w_K \in \{1, \dots, 2^{nR_K}\}$ and $\widetilde{w}_K \in \{1, \dots, 2^{n\widetilde{R}_K}\}$.

(iv) Furthermore, we set

$$\widetilde{R}_i = I(U_i; Z \mid U_{i-1}), \quad i = 1, \dots, K, \tag{A.2}$$

where $U_0 = \phi$ and $U_K = X$.

*Encoding.* Assume the messages to be transmitted are $(w_0, w_1, \dots, w_K)$. Then, the encoder randomly picks a set $(\widetilde{w}_1, \dots, \widetilde{w}_K)$ and sends $\mathbf{x}(w_0, w_1, \dots, w_K, \widetilde{w}_1, \dots, \widetilde{w}_K)$.

*Decoding.* It is straightforward to see that if the following conditions are satisfied:

$$\begin{aligned}
R_0 + R_1 + \widetilde{R}_1 &\le I(U_1; Y_1),\\
R_j + \widetilde{R}_j &\le I\left(U_j; Y_j \mid U_{j-1}\right), \quad j = 2, \dots, K-1,\\
R_K + \widetilde{R}_K &\le I(X; Y_K \mid U_{K-1}),
\end{aligned} \tag{A.3}$$

then all users can decode both the common message and the independent message directed to itself with vanishingly small error probability. Moreover, since the channel is degraded, each user, say the $j$th one, can decode all of the independent messages intended for the users whose channels are degraded with respect to the $j$th user's channel. Thus, these degraded users' rates can be exploited to increase the $j$th user's rate which leads to the following achievable region:

$$R_0 + \sum_{j=1}^{\ell} R_j + \sum_{j=1}^{\ell} \tilde{R}_j \leq \sum_{j=1}^{\ell} I\left(U_j; Y_j \mid U_{j-1}\right), \quad \ell = 1, \ldots, K, \tag{A.4}$$

where $U_0 = \phi$ and $U_K = X$. Moreover, after eliminating $\{\tilde{R}_j\}_{j=1}^{K}$, (A.4) can be expressed as

$$R_0 + \sum_{j=1}^{\ell} R_j \leq \sum_{j=1}^{\ell} I\left(U_j; Y_j \mid U_{j-1}\right) - I(U_\ell; Z), \quad \ell = 1, \ldots, K, \tag{A.5}$$

where we used the fact that

$$\sum_{j=1}^{\ell} \tilde{R}_j = \sum_{j=1}^{\ell} I\left(U_j; Z \mid U_{j-1}\right) = I(U_1, \ldots, U_\ell; Z) = I(U_\ell; Z), \tag{A.6}$$

where the second and the third equalities are due to the following Markov chain:

$$U_1 \longrightarrow \cdots \longrightarrow U_{K-1} \longrightarrow X \longrightarrow Z. \tag{A.7}$$

*Equivocation Calculation.* We now calculate the equivocation of the code described above. To that end, we first introduce the following lemma which states that a code satisfying the sum rate secrecy constraint fulfills all other secrecy constraints.

**Lemma 11.** *If the sum rate secrecy constraint is satisfied, that is,*

$$\frac{1}{n} H(W_0, W_1, \ldots, W_K \mid Z^n) \geq \sum_{j=0}^{K} R_j - \epsilon_n, \tag{A.8}$$

*then all other secrecy constraints are satisfied as well, that is,*

$$\frac{1}{n} H(\mathcal{S}(W) \mid Z^n) \geq \sum_{j \in \mathcal{S}(W)} R_j - \epsilon_n, \tag{A.9}$$

*where $\mathcal{S}(W)$ denotes any subset of $\{W_0, W_1, \ldots, W_K\}$.*

*Proof.* The proof of this lemma is as follows.

$$\begin{aligned}
&\frac{1}{n} H(\mathcal{S}(W) \mid Z^n) \\
&\quad = \frac{1}{n} H(\mathcal{S}(W), \mathcal{S}^c(W) \mid Z^n) - \frac{1}{n} H(\mathcal{S}^c(W) \mid \mathcal{S}(W), Z^n) && \text{(A.10)} \\
&\quad \geq \sum_{j=0}^{K} R_j - \epsilon_n - \frac{1}{n} H(\mathcal{S}^c(W) \mid \mathcal{S}(W), Z^n) && \text{(A.11)} \\
&\quad = \sum_{j \in \mathcal{S}(W)} R_j - \epsilon_n + \sum_{j \in \mathcal{S}^c(W)} R_j - \frac{1}{n} H(\mathcal{S}^c(W) \mid \mathcal{S}(W), Z^n) && \text{(A.12)} \\
&\quad = \sum_{j \in \mathcal{S}(W)} R_j - \epsilon_n + \frac{1}{n} H(\mathcal{S}^c(W)) - \frac{1}{n} H(\mathcal{S}^c(W) \mid \mathcal{S}(W), Z^n) && \text{(A.13)} \\
&\quad \geq \sum_{j \in \mathcal{S}(W)} R_j - \epsilon_n, && \text{(A.14)}
\end{aligned}$$

where (A.11) is due to the fact that we assumed that sum rate secrecy constraint (A.8) is satisfied and (A.13) follows from

$$\sum_{j \in \mathcal{S}^c(W)} R_j = \frac{1}{n} H(\mathcal{S}^c(W)), \tag{A.15}$$

which is a consequence of the fact that message sets are uniformly and independently distributed. □

Hence, it is sufficient to check whether coding scheme presented satisfies the sum rate secrecy constraint.

$$\begin{aligned}
&H(W_0, W_1, \ldots, W_K \mid Z^n) \\
&\quad = H(W_0, W_1, \ldots, W_K, Z^n) - H(Z^n) && \text{(A.16)} \\
&\quad = H(U_1^n, \ldots, U_{K-1}^n, X^n, W_0, W_1, \ldots, W_K, Z^n) - H(Z^n) \\
&\qquad - H(U_1^n, \ldots, U_{K-1}^n, X^n \mid W_0, W_1, \ldots, W_K, Z^n) && \text{(A.17)} \\
&\quad = H(U_1^n, \ldots, U_{K-1}^n, X^n) \\
&\qquad + H(W_0, W_1, \ldots, W_K, Z^n \mid U_1^n, \ldots, U_{K-1}^n, X^n) - H(Z^n) \\
&\qquad - H(U_1^n, \ldots, U_{K-1}^n, X^n \mid W_0, W_1, \ldots, W_K, Z^n) && \text{(A.18)} \\
&\quad \geq H(U_1^n, \ldots, U_{K-1}^n, X^n) - I(U_1^n, \ldots, U_{K-1}^n, X^n; Z^n) \\
&\qquad - H(U_1^n, \ldots, U_{K-1}^n, X^n \mid W_0, W_1, \ldots, W_K, Z^n), && \text{(A.19)}
\end{aligned}$$

where each term will be treated separately. Since given $U_k^n = u_k^n$, $U_{k+1}^n$ can take $2^{n(R_{k+1}+\tilde{R}_{k+1})}$ values uniformly, the first term is

$$\begin{aligned} &H(U_1^n,\dots,U_{K-1}^n,X^n)\\ &\quad = H(U_1^n)+\sum_{k=2}^{K-1}H\left(U_k^n \mid U_{k-1}^n\right)+H(X^n \mid U_{K-1}^n) \end{aligned}\tag{A.20}$$

$$= nR_0 + n\sum_{k=1}^{K}R_k + n\sum_{k=1}^{K}\widetilde{R}_k, \tag{A.21}$$

where the first equality follows from the following Markov chain:

$$U_1^n \longrightarrow U_2^n \longrightarrow \cdots \longrightarrow U_{K-1}^n \longrightarrow X^n. \tag{A.22}$$

The second term in (A.19) is

$$\begin{aligned} &I(U_1^n,\dots,U_{K-1}^n,X^n;Z^n)\\ &\quad = I(X^n;Z^n)+I(U_1^n,U_2^n,\dots,U_{K-1}^n;Z^n \mid X^n) \end{aligned}\tag{A.23}$$

$$= I(X^n;Z^n) \tag{A.24}$$

$$\le nI(X;Z)+\gamma_n, \tag{A.25}$$

where (A.24) follows from the Markov chain in (A.22) and (A.25) can be shown by following the approach devised in [1]. We now bound the third term in (A.19). To that end, assume that the eavesdropper tries to decode $(U_1^n,\dots,U_{K-1}^n,X^n)$ using the side information $(W_0,W_1,\dots,W_K)$ which is equivalent to decoding $(\widetilde{W}_1,\dots,\widetilde{W}_K)$. Since $\widetilde{R}_j$s are selected to ensure that the eavesdropper can decode them successively, see (A.2), then using Fano's lemma, we have

$$H(U_1^n,\dots,U_{K-1}^n,X^n \mid W_0,W_1,\dots,W_K,Z^n) \le \epsilon_n. \tag{A.26}$$

Thus, using (A.21), (A.25), and (A.26) in (A.19), we get

$$\begin{aligned} &H(W_0,W_1,\dots,W_K \mid Z^n)\\ &\quad \ge n\sum_{j=0}^{K}R_j + n\sum_{j=1}^{K}\widetilde{R}_j - nI(X;Z)-\epsilon_n \end{aligned}\tag{A.27}$$

$$= n\sum_{j=0}^{K}R_j - \epsilon_n - \gamma_n, \tag{A.28}$$

where (A.28) follows from the following, see (A.2) and (A.6),

$$\sum_{j=1}^{K}\widetilde{R}_j = I(X;Z). \tag{A.29}$$

*A.2. Converse.* First let us define the following auxiliary random variables:

$$U_{k,i} = W_0W_1\cdots W_kY_{k+1}^{i-1}Z_{i+1}^n, \quad k=1,\dots,K-1, \tag{A.30}$$

which satisfy the following Markov chain:

$$\begin{aligned} &U_{1,i} \longrightarrow U_{2,i} \longrightarrow \cdots \longrightarrow U_{K-1,i}\\ &\quad \longrightarrow X_i \longrightarrow (Z_i,Y_{K,i},\dots,Y_{1,i}). \end{aligned}\tag{A.31}$$

To provide a converse, we will show

$$\frac{1}{n}H(W_0,W_1,\dots,W_\ell \mid Z^n) \le \sum_{k=1}^{\ell}I(U_k;Y_k \mid U_{k-1}) - I(U_\ell;Z), \quad \ell=1,\dots,K, \tag{A.32}$$

where $U_0=\phi$, $U_K=X$. We show this in three steps. First, let us write down

$$\begin{aligned} H(W_0,W_1,\dots,W_\ell \mid Z^n) &= H(W_0,W_1 \mid Z^n)\\ &\quad + \sum_{k=2}^{\ell}H(W_k \mid W_0,W_1,\dots,W_{k-1},Z^n). \end{aligned}\tag{A.33}$$

The first term on the right-hand side of (A.33) is bounded as follows:

$$\begin{aligned} &H(W_0,W_1 \mid Z^n)\\ &\quad \le I(W_0,W_1;Y_1^n)-I(W_0,W_1;Z^n)+\epsilon_n \end{aligned}\tag{A.34}$$

$$\begin{aligned} &\le \sum_{i=1}^{n}I\left(W_0,W_1;Y_{1,i} \mid Y_1^{i-1},Z_{i+1}^n\right)\\ &\quad - I\left(W_0,W_1;Z_i \mid Y_1^{i-1},Z_{i+1}^n\right)+\epsilon_n \end{aligned}\tag{A.35}$$

$$\begin{aligned} &\le \sum_{i=1}^{n}I\left(W_0,W_1;Y_{1,i} \mid Y_1^{i-1},Z_{i+1}^n\right)\\ &\quad - I\left(W_0,W_1;Z_i \mid Y_1^{i-1},Z_{i+1}^n\right)\\ &\quad + I\left(Y_1^{i-1},Z_{i+1}^n;Y_{1,i}\right)-I\left(Y_1^{i-1},Z_{i+1}^n;Z_i\right)+\epsilon_n \end{aligned}\tag{A.36}$$

$$\begin{aligned} &= \sum_{i=1}^{n}I\left(W_0,W_1,Y_1^{i-1},Z_{i+1}^n;Y_{1,i}\right)\\ &\quad - I\left(W_0,W_1,Y_1^{i-1},Z_{i+1}^n;Z_i\right)+\epsilon_n \end{aligned}\tag{A.37}$$

$$\begin{aligned} &\le \sum_{i=1}^{n}I\left(W_0,W_1,Y_1^{i-1},Z_{i+1}^n;Y_{1,i}\right)\\ &\quad - I\left(W_0,W_1,Y_1^{i-1},Z_{i+1}^n;Z_i\right)\\ &\quad + I\left(Y_2^{i-1};Y_{1,i} \mid W_0,W_1,Y_1^{i-1},Z_{i+1}^n\right)\\ &\quad - I\left(Y_2^{i-1};Z_i \mid W_0,W_1,Y_1^{i-1},Z_{i+1}^n\right)+\epsilon_n \end{aligned}\tag{A.38}$$

$$\begin{aligned} &= \sum_{i=1}^{n}I\left(W_0,W_1,Y_1^{i-1},Z_{i+1}^n,Y_2^{i-1};Y_{1,i}\right)\\ &\quad - I\left(W_0,W_1,Y_1^{i-1},Z_{i+1}^n,Y_2^{i-1};Z_i\right)+\epsilon_n \end{aligned}\tag{A.39}$$

$$\begin{aligned}
&= \sum_{i=1}^{n} I\left(W_0, W_1, Z_{i+1}^n, Y_2^{i-1}; Y_{1,i}\right) \\
&\quad - I\left(W_0, W_1, Z_{i+1}^n, Y_2^{i-1}; Z_i\right)
\end{aligned} \tag{A.40}$$

$$\begin{aligned}
&\quad + I\left(Y_1^{i-1}; Y_{1,i} \mid W_0, W_1, Z_{i+1}^n, Y_2^{i-1}\right) \\
&\quad - I\left(Y_1^{i-1}; Z_i \mid W_0, W_1, Z_{i+1}^n, Y_2^{i-1}\right) + \epsilon_n
\end{aligned} \tag{A.41}$$

$$\begin{aligned}
&= \sum_{i=1}^{n} I\left(W_0, W_1, Z_{i+1}^n, Y_2^{i-1}; Y_{1,i}\right) \\
&\quad - I\left(W_0, W_1, Z_{i+1}^n, Y_2^{i-1}; Z_i\right) + \epsilon_n
\end{aligned} \tag{A.42}$$

$$= \sum_{i=1}^{n} I\left(U_{1,i}; Y_{1,i}\right) - I\left(U_{1,i}; Z_i\right) + \epsilon_n, \tag{A.43}$$

where (A.34) follows from Fano's lemma, (A.35) is obtained using Csiszar-Korner identity (see [2, Lemma 7]), and (A.36) is due to the fact that

$$I\left(Y_1^{i-1}, Z_{i+1}^n; Y_{1,i}\right) - I\left(Y_1^{i-1}, Z_{i+1}^n; Z_i\right) > 0, \tag{A.44}$$

which follows from the fact that each user's channel is less noisy with respect to the eavesdropper. Similarly, (A.38) follows from the fact that

$$\begin{aligned}
&I\left(Y_2^{i-1}; Y_{1,i} \mid W_0, W_1, Y_1^{i-1}, Z_{i+1}^n\right) \\
&\qquad - I\left(Y_2^{i-1}; Z_i \mid W_0, W_1, Y_1^{i-1}, Z_{i+1}^n\right) > 0,
\end{aligned} \tag{A.45}$$

which is a consequence of the fact that each user's channel is less noisy with respect to the eavesdropper's channel. Finally, (A.42) is due to the following Markov chain:

$$Y_1^{i-1} \longrightarrow Y_2^{i-1} \longrightarrow \left(W_0, W_1, Z_{i+1}^n, Y_{1,i}, Z_i\right), \tag{A.46}$$

which is a consequence of the fact that the legitimate receivers exhibit a degradation order.

We now bound the terms of the summation in (A.33) for $2 \le k \le K-1$. Let us use the shorthand notation, $\widetilde{W}_{k-1} = (W_0, W_1, \dots, W_{k-1})$, then

$$\begin{aligned}
&H\left(W_k \mid \widetilde{W}_{k-1}, Z^n\right) \\
&\quad \le I\left(W_k; Y_k^n \mid \widetilde{W}_{k-1}\right) - I\left(W_k; Z^n \mid \widetilde{W}_{k-1}\right) + \epsilon_n
\end{aligned} \tag{A.47}$$

$$\begin{aligned}
&\quad \le \sum_{i=1}^{n} I\left(W_k; Y_{k,i} \mid \widetilde{W}_{k-1}, Y_k^{i-1}, Z_{i+1}^n\right) \\
&\qquad - I\left(W_k; Z_i \mid \widetilde{W}_{k-1}, Y_k^{i-1}, Z_{i+1}^n\right) + \epsilon_n
\end{aligned} \tag{A.48}$$

$$\begin{aligned}
&\le \sum_{i=1}^{n} I\left(W_k; Y_{k,i} \mid \widetilde{W}_{k-1}, Y_k^{i-1}, Z_{i+1}^n\right) \\
&\quad - I\left(W_k; Z_i \mid \widetilde{W}_{k-1}, Y_k^{i-1}, Z_{i+1}^n\right) \\
&\quad + I\left(Y_{k+1}^{i-1}; Y_{k,i} \mid \widetilde{W}_{k-1}, Y_k^{i-1}, Z_{i+1}^n, W_k\right) \\
&\quad - I\left(Y_{k+1}^{i-1}; Z_i \mid \widetilde{W}_{k-1}, Y_k^{i-1}, Z_{i+1}^n, W_k\right) + \epsilon_n
\end{aligned} \tag{A.49}$$

$$\begin{aligned}
&= \sum_{i=1}^{n} I\left(W_k, Y_{k+1}^{i-1}; Y_{k,i} \mid \widetilde{W}_{k-1}, Y_k^{i-1}, Z_{i+1}^n\right) \\
&\quad - I\left(W_k, Y_{k+1}^{i-1}; Z_i \mid \widetilde{W}_{k-1}, Y_k^{i-1}, Z_{i+1}^n\right) + \epsilon_n
\end{aligned} \tag{A.50}$$

$$\begin{aligned}
&= \sum_{i=1}^{n} I\left(U_{k,i}; Y_{k,i} \mid U_{k-1,i}\right) \\
&\quad - I\left(U_{k,i}; Z_i \mid U_{k-1,i}\right) + \epsilon_n
\end{aligned} \tag{A.51}$$

where (A.47) follows from Fano's lemma, (A.48) is obtained through Csiszar-Korner identity, and (A.49) is a consequence of the fact that

$$\begin{aligned}
&I\left(Y_{k+1}^{i-1}; Y_{k,i} \mid \widetilde{W}_{k-1}, Y_k^{i-1}, Z_{i+1}^n, W_k\right) \\
&\qquad - I\left(Y_{k+1}^{i-1}; Z_i \mid \widetilde{W}_{k-1}, Y_k^{i-1}, Z_{i+1}^n, W_k\right) > 0,
\end{aligned} \tag{A.52}$$

which follows from the fact that each user's channel is less noisy with respect to the eavesdropper's channel. Finally, we bound the following term where we again use the shorthand notation $\widetilde{W}_{K-1} = (W_0, W_1, \dots, W_{K-1})$,

$$\begin{aligned}
&H\left(W_K \mid \widetilde{W}_{K-1}, Z^n\right) \\
&\quad \le I\left(W_K; Y_K^n \mid \widetilde{W}_{K-1}\right) - I\left(W_K; Z^n \mid \widetilde{W}_{K-1}\right) + \epsilon_n
\end{aligned} \tag{A.53}$$

$$\begin{aligned}
&\le \sum_{i=1}^{n} I\left(W_K; Y_{K,i} \mid \widetilde{W}_{K-1}, Y_K^{i-1}, Z_{i+1}^n\right) \\
&\quad - I\left(W_K; Z_i \mid \widetilde{W}_{K-1}, Y_K^{i-1}, Z_{i+1}^n\right) + \epsilon_n
\end{aligned} \tag{A.54}$$

$$\begin{aligned}
&\le \sum_{i=1}^{n} I\left(W_K; Y_{K,i} \mid \widetilde{W}_{K-1}, Y_K^{i-1}, Z_{i+1}^n\right) \\
&\quad - I\left(W_K; Z_i \mid \widetilde{W}_{K-1}, Y_K^{i-1}, Z_{i+1}^n\right) \\
&\quad + I\left(X_i; Y_{K,i} \mid \widetilde{W}_{K-1}, Y_K^{i-1}, Z_{i+1}^n, W_K\right) \\
&\quad - I\left(X_i; Z_i \mid \widetilde{W}_{K-1}, Y_K^{i-1}, Z_{i+1}^n, W_K\right) + \epsilon_n
\end{aligned} \tag{A.55}$$

$$\begin{aligned}
&= \sum_{i=1}^{n} I\left(W_K, X_i; Y_{K,i} \mid \widetilde{W}_{K-1}, Y_K^{i-1}, Z_{i+1}^n\right) \\
&\quad - I\left(W_K, X_i; Z_i \mid \widetilde{W}_{K-1}, Y_K^{i-1}, Z_{i+1}^n\right) + \epsilon_n
\end{aligned} \tag{A.56}$$

$$\begin{aligned}
&= \sum_{i=1}^{n} I\left(X_i; Y_{K,i} \mid \widetilde{W}_{K-1}, Y_K^{i-1}, Z_{i+1}^n\right) \\
&\quad + I\left(W_K; Y_{K,i} \mid \widetilde{W}_{K-1}, Y_K^{i-1}, Z_{i+1}^n, X_i\right) \\
&\quad - I\left(X_i; Z_i \mid \widetilde{W}_{K-1}, Y_K^{i-1}, Z_{i+1}^n\right) \\
&\quad - I\left(W_K; Z_i \mid \widetilde{W}_{K-1}, Y_K^{i-1}, Z_{i+1}^n, X_i\right) + \epsilon_n
\end{aligned} \tag{A.57}$$

$$\begin{aligned}
&= \sum_{i=1}^{n} I\left(X_i; Y_{K,i} \mid \widetilde{W}_{K-1}, Y_K^{i-1}, Z_{i+1}^n\right) \\
&\quad - I\left(X_i; Z_i \mid \widetilde{W}_{K-1}, Y_K^{i-1}, Z_{i+1}^n\right) + \epsilon_n
\end{aligned} \tag{A.58}$$

$$\begin{aligned}
&= \sum_{i=1}^{n} I(X_i; Y_{K,i} \mid U_{K-1,i}) \\
&\quad - I(X_i; Z_i \mid U_{K-1,i}) + \epsilon_n,
\end{aligned} \tag{A.59}$$

where (A.53) follows from Fano's lemma, (A.54) is obtained by using Csiszar-Korner identity, and (A.55) follows from the fact that

$$\begin{aligned}
&I\left(X_i; Y_{K,i} \mid \widetilde{W}_{K-1}, Y_K^{i-1}, Z_{i+1}^n, W_K\right) \\
&\quad - I\left(X_i; Z_i \mid \widetilde{W}_{K-1}, Y_K^{i-1}, Z_{i+1}^n, W_K\right) > 0,
\end{aligned} \tag{A.60}$$

which is due to the fact that each user's channel is less noisy with respect to the eavesdropper and (A.58) is due to the Markov chain:

$$(Y_{K,i}, Z_i) \longrightarrow X_i \longrightarrow \left(W_0, W_1, \ldots, W_K, Y_K^{i-1}, Z_{i+1}^n\right), \tag{A.61}$$

which follows from the fact that the channel is memoryless. Finally, plugging (A.43), (A.51), and (A.59) into (A.33), we get

$$\begin{aligned}
H(W_0, W_1, \ldots, W_\ell \mid Z^n) &\le n \sum_{k=1}^{\ell} I(U_k; Y_k \mid U_{k-1}) \\
&\quad - nI(U_\ell; Z), \ell = 1, \ldots, K,
\end{aligned} \tag{A.62}$$

where $U_0 = \phi$ and $U_K = X$, and this concludes the converse.

## B. Proof of Theorem 3

Achievability of these rates follows from [8, Proposition 2]. We provide the converse. First let us define the following random variables:

$$\begin{aligned}
Z^n &= (Z_1^n, \ldots, Z_M^n), \\
Y_k^n &= \left(Y_{k1}^n, \ldots, Y_{kM}^n\right), \\
Z_{i+1}^n &= \left(Z_{1,i+1}^n, \ldots, Z_{M,i+1}^n\right), \\
Y_k^{i-1} &= \left(Y_{k1}^{i-1} \ldots, Y_{kM}^{i-1}\right), \\
Y_k(i) &= (Y_{k1}(i), \ldots, Y_{kM}(i)), \\
Z(i) &= (Z_1(i), \ldots, Z_M(i)),
\end{aligned} \tag{B.1}$$

where $Y_{kl}^{i-1} = (Y_{kl}(1), \ldots, Y_{kl}(i-1))$, $Z_{l,i+1}^n = (Z_l(i+1), \ldots, Z_l(n))$. Start with the definition

$$H(W_0 \mid Z^n) = H(W_0) - I(W_0; Z^n) \tag{B.2}$$

$$\le I\left(W_0; Y_k^n\right) - I(W_0; Z^n) + \epsilon_n \tag{B.3}$$

$$\begin{aligned}
&= \sum_{i=1}^{n} I\left(W_0; Y_k(i) \mid Y_k^{i-1}\right) \\
&\quad - I(W_0; Z(i) \mid Z_{i+1}^n) + \epsilon_n
\end{aligned} \tag{B.4}$$

$$\begin{aligned}
&= \sum_{i=1}^{n} I\left(W_0, Z_{i+1}^n; Y_k(i) \mid Y_k^{i-1}\right) \\
&\quad - I\left(Z_{i+1}^n; Y_k(i) \mid Y_k^{i-1}, W_0\right) \\
&\quad - I\left(W_0, Y_k^{i-1}; Z(i) \mid Z_{i+1}^n\right) \\
&\quad + I\left(Y_k^{i-1}; Z(i) \mid Z_{i+1}^n, W_0\right) + \epsilon_n
\end{aligned} \tag{B.5}$$

$$\begin{aligned}
&= \sum_{i=1}^{n} I\left(W_0, Z_{i+1}^n; Y_k(i) \mid Y_k^{i-1}\right) \\
&\quad - I\left(W_0, Y_k^{i-1}; Z(i) \mid Z_{i+1}^n\right) + \epsilon_n
\end{aligned} \tag{B.6}$$

$$\begin{aligned}
&= \sum_{i=1}^{n} I\left(W_0; Y_k(i) \mid Y_k^{i-1}, Z_{i+1}^n\right) \\
&\quad + I\left(Z_{i+1}^n; Y_k(i) \mid Y_k^{i-1}\right) \\
&\quad - I\left(W_0; Z(i) \mid Z_{i+1}^n, Y_k^{i-1}\right) \\
&\quad - I\left(Y_k^{i-1}; Z(i) \mid Z_{i+1}^n\right) + \epsilon_n
\end{aligned} \tag{B.7}$$

$$\begin{aligned}
&= \sum_{i=1}^{n} I\left(W_0; Y_k(i) \mid Y_k^{i-1}, Z_{i+1}^n\right) \\
&\quad - I\left(W_0; Z(i) \mid Z_{i+1}^n, Y_k^{i-1}\right) + \epsilon_n,
\end{aligned} \tag{B.8}$$

where (B.6) and (B.8) are due the following identities:

$$\begin{aligned}
\sum_{i=1}^{n} I\left(Z_{i+1}^n; Y_k(i) \mid Y_k^{i-1}, W_0\right) &= \sum_{i=1}^{n} I\left(Y_k^{i-1}; Z(i) \mid Z_{i+1}^n, W_0\right), \\
\sum_{i=1}^{n} I\left(Z_{i+1}^n; Y_k(i) \mid Y_k^{i-1}\right) &= \sum_{i=1}^{n} I\left(Y_k^{i-1}; Z(i) \mid Z_{i+1}^n\right),
\end{aligned} \tag{B.9}$$

respectively, which are due to [2, Lemma 7]. Now, we will bound each summand in (B.8) separately. First, define the following variables:

$$\begin{aligned}
U_{k,i} &= \left(Z_{i+1}^n, Y_k^{i-1}\right), \\
\widetilde{Y}_k^{l-1}(i) &= (Y_{k1}(i), \ldots, Y_{k(l-1)}(i)), \\
\widetilde{Z}_{l+1}^M(i) &= (Z_{l+1}(i), \ldots, Z_M(i)).
\end{aligned} \tag{B.10}$$

Hence, the summand in (B.8) can be written as follows:

$$I\Big(W_0; Y_k(i) \mid Y_k^{i-1}, Z_{i+1}^n\Big) - I\Big(W_0; Z(i) \mid Z_{i+1}^n, Y_k^{i-1}\Big) \tag{B.11}$$

$$= I(W_0; Y_k(i) \mid U_{k,i}) - I(W_0; Z(i) \mid U_{k,i}) \tag{B.12}$$

$$\begin{aligned} &= I(W_0; Y_{k1}(i), \dots, Y_{kM}(i) \mid U_{k,i}) \\ &\quad - I(W_0; Z_1(i), \dots, Z_M(i) \mid U_{k,i}) \end{aligned} \tag{B.13}$$

$$\begin{aligned} &= \sum_{l=1}^{M} I\Big(W_0; Y_{kl}(i) \mid U_{k,i}, \widetilde{Y}_k^{l-1}(i)\Big) \\ &\quad - I\Big(W_0; Z_l(i) \mid U_{k,i}, \widetilde{Z}_{l+1}^M(i)\Big) \end{aligned} \tag{B.14}$$

$$\begin{aligned} &= \sum_{l=1}^{M} I\Big(W_0, \widetilde{Z}_{l+1}^M(i); Y_{kl}(i) \mid U_{k,i}, \widetilde{Y}_k^{l-1}(i)\Big) \\ &\quad - I\Big(\widetilde{Z}_{l+1}^M(i); Y_{kl}(i) \mid U_{k,i}, \widetilde{Y}_k^{l-1}(i), W_0\Big) \\ &\quad - I\Big(W_0, \widetilde{Y}_k^{l-1}(i); Z_l(i) \mid U_{k,i}, \widetilde{Z}_{l+1}^M(i)\Big) \\ &\quad + I\Big(\widetilde{Y}_k^{l-1}(i); Z_l(i) \mid U_{k,i}, \widetilde{Z}_{l+1}^M(i), W_0\Big) \end{aligned} \tag{B.15}$$

$$\begin{aligned} &= \sum_{l=1}^{M} I\Big(W_0, \widetilde{Z}_{l+1}^M(i); Y_{kl}(i) \mid U_{k,i}, \widetilde{Y}_k^{l-1}(i)\Big) \\ &\quad - I\Big(W_0, \widetilde{Y}_k^{l-1}(i); Z_l(i) \mid U_{k,i}, \widetilde{Z}_{l+1}^M(i)\Big) \end{aligned} \tag{B.16}$$

$$\begin{aligned} &= \sum_{l=1}^{M} I\Big(\widetilde{Z}_{l+1}^M(i); Y_{kl}(i) \mid U_{k,i}, \widetilde{Y}_k^{l-1}(i)\Big) \\ &\quad + I\Big(W_0; Y_{kl}(i) \mid U_{k,i}, \widetilde{Y}_k^{l-1}(i), \widetilde{Z}_{l+1}^M(i)\Big) \\ &\quad - I\Big(\widetilde{Y}_k^{l-1}(i); Z_l(i) \mid U_{k,i}, \widetilde{Z}_{l+1}^M(i)\Big) \\ &\quad - I\Big(W_0; Z_l(i) \mid U_{k,i}, \widetilde{Z}_{l+1}^M(i), \widetilde{Y}_k^{l-1}(i)\Big) \end{aligned} \tag{B.17}$$

$$\begin{aligned} &= \sum_{l=1}^{M} I\Big(W_0; Y_{kl}(i) \mid U_{k,i}, \widetilde{Y}_k^{l-1}(i), \widetilde{Z}_{l+1}^M(i)\Big) \\ &\quad - I\Big(W_0; Z_l(i) \ \mid U_{k,i}, \widetilde{Z}_{l+1}^M(i), \widetilde{Y}_k^{l-1}(i)\Big), \end{aligned} \tag{B.18}$$

where (B.16) and (B.18) follow from the following identities:

$$\begin{aligned} &\sum_{l=1}^{M} I\Big(\widetilde{Z}_{l+1}^M(i); Y_{kl}(i) \mid U_{k,i}, \widetilde{Y}_k^{l-1}(i), W_0\Big) \\ &\qquad = \sum_{l=1}^{M} I\Big(\widetilde{Y}_k^{l-1}(i); Z_l(i) \mid U_{k,i}, \widetilde{Z}_{l+1}^M(i), W_0\Big) \\ &\sum_{l=1}^{M} I\Big(\widetilde{Z}_{l+1}^M(i); Y_{kl}(i) \mid U_{k,i}, \widetilde{Y}_k^{l-1}(i)\Big) \\ &\qquad = \sum_{l=1}^{M} I\Big(\widetilde{Y}_k^{l-1}(i); Z_l(i) \ \mid U_{k,i}, \widetilde{Z}_{l+1}^M(i)\Big), \end{aligned} \tag{B.19}$$

respectively, which are again due to [2, Lemma 7]. Now, define the set of subchannels, say $\mathcal{S}(k)$, in which the $k$th user is less noisy with respect to the eavesdropper. Thus, the summands in (B.18) for $l \notin \mathcal{S}(k)$ are negative and by dropping them, we can bound (B.18) as follows:

$$\begin{aligned} &I\Big(W_0; Y_k(i) \mid Y_k^{i-1}, Z_{i+1}^n\Big) - I\Big(W_0; Z(i) \mid Z_{i+1}^n, Y_k^{i-1}\Big) \\ &\quad \le \sum_{l \in \mathcal{S}(k)} I\Big(W_0; Y_{kl}(i) \mid U_{k,i}, \widetilde{Y}_k^{l-1}(i), \widetilde{Z}_{l+1}^M(i)\Big) \\ &\qquad - I\Big(W_0; Z_l(i) \ \mid U_{k,i}, \widetilde{Z}_{l+1}^M(i), \widetilde{Y}_k^{l-1}(i)\Big). \end{aligned} \tag{B.20}$$

Moreover, for $l \in \mathcal{S}(k)$, we have

$$\begin{aligned} &I\Big(U_{k,i}, \widetilde{Y}_k^{l-1}(i), \widetilde{Z}_{l+1}^M(i); Y_{kl}(i)\Big) \\ &\quad - I\Big(U_{k,i}, \widetilde{Y}_k^{l-1}(i), \widetilde{Z}_{l+1}^M(i); Z_l(i)\Big) \ge 0 \end{aligned} \tag{B.21}$$

$$\begin{aligned} &I\Big(X_l(i); Y_{kl}(i) \mid U_{k,i}, \widetilde{Y}_k^{l-1}(i), \widetilde{Z}_{l+1}^M(i), W_0\Big) \\ &\quad - I\Big(X_l(i); Z_l(i) \mid U_{k,i}, \widetilde{Z}_{l+1}^M(i), \widetilde{Y}_k^{l-1}(i), W_0\Big) \ge 0, \end{aligned} \tag{B.22}$$

where both are due to the fact that for $l \in \mathcal{S}(k)$, in this subchannel the $k$th user is less noisy with respect to the eavesdropper. Therefore, adding (B.21) and (B.22) to each summand in (B.20), we get the following bound:

$$\begin{aligned} &I\Big(W_0; Y_k(i) \mid Y_k^{i-1}, Z_{i+1}^n\Big) - I\Big(W_0; Z(i) \mid Z_{i+1}^n, Y_k^{i-1}\Big) \\ &\quad \le \sum_{l \in \mathcal{S}(k)} I\Big(X_l(i), W_0, U_{k,i}, \widetilde{Y}_k^{l-1}(i), \widetilde{Z}_{l+1}^M(i); Y_{kl}(i)\Big) \\ &\qquad - I\Big(X_l(i), W_0, U_{k,i}, \widetilde{Y}_k^{l-1}(i), \widetilde{Z}_{l+1}^M(i); Z_l(i)\Big) \end{aligned} \tag{B.23}$$

$$= \sum_{l \in \mathcal{S}(k)} I(X_l(i); Y_{kl}(i)) - I(X_l(i); Z_l(i)), \tag{B.24}$$

where an equality follows from the following Markov chain:

$$\Big(W_0, U_{k,i}, \widetilde{Y}_k^{l-1}(i), \widetilde{Z}_{l+1}^M(i)\Big) \longrightarrow X_l(i) \longrightarrow (Y_{kl}(i), Z_l(i)), \tag{B.25}$$

which is a consequence of the facts that channel is memoryless and subchannels are independent. Finally, using (B.24) in (B.8), we get

$$\begin{aligned} H(W_0 \mid Z^n) &\le \sum_{i=1}^{n} \sum_{l \in \mathcal{S}(k)} I(X_l(i); Y_{kl}(i)) - I(X_l(i); Z_l(i)) + \epsilon_n \\ &\le n \sum_{l \in \mathcal{S}(k)} I(X_l; Y_{kl}) - I(X_l; Z_l) + \epsilon_n \\ &= n \sum_{l=1}^{M} [I(X_l; Y_{kl}) - I(X_l; Z_l)]^+ + \epsilon_n, \end{aligned} \tag{B.26}$$

which completes the proof.

## C. Proof of Theorem 5

Achievability of Theorem 5 is a consequence of the achievability result for wiretap channels in [2]. We provide the converse proof here. We first define the function $\rho(l)$ which denotes the index of the strongest user in the $l$th subchannel in the sense that

$$I(U;Y_{kl}) \leq I\left(U;Y_{\rho(l)l}\right) \tag{C.1}$$

for all $U \rightarrow X_l \rightarrow (Y_{1l},\ldots,Y_{Kl},Z_l)$ and any $k \in \{1,\ldots,K\}$. Moreover, we define the following shorthand notations:

$$\begin{aligned}
\widetilde{Y}_l^n &= Y_{\rho(l)l}^n, \quad l = 1,\ldots,M,\\
\widetilde{Y}^n &= \left(\widetilde{Y}_1^n,\ldots,\widetilde{Y}_M^n\right),\\
Y_k^n &= \left(Y_{k1}^n,\ldots,Y_{kM}^n\right), \quad k = 1,\ldots,K,\\
Z^n &= (Z_1^n,\ldots,Z_M^n),\\
Y_k^{i-1} &= \left(Y_{k1}^{i-1},\ldots,Y_{kM}^{i-1}\right), \quad k = 1,\ldots,K,\\
Z^{i-1} &= \left(Z_1^{i-1},\ldots,Z_M^{i-1}\right),\\
\widetilde{Y}_{i+1}^n &= \left(\widetilde{Y}_{1,i+1}^n,\ldots,\widetilde{Y}_{M,i+1}^n\right),\\
Y_k^{l-1}(i) &= (Y_{k1}(i),\ldots,Y_{k,l-1}(i)), \quad l = 1,\ldots,M,\\
Z^{l-1}(i) &= (Z_1(i),\ldots,Z_{l-1}(i)), \quad l = 1,\ldots,M,\\
\widetilde{Y}_{l+1}^M(i) &= \left(\widetilde{Y}_{l+1}(i),\ldots,\widetilde{Y}_M(i)\right), \quad l = 1,\ldots,M.
\end{aligned} \tag{C.2}$$

We first introduce the following lemma.

**Lemma 12.** *For the parallel multireceiver wiretap channel with less noisiness order, one has*

$$I\left(W_k;Y_k^n\right) \leq I\left(W_k;\widetilde{Y}^n\right), \quad k = 1,\ldots,K. \tag{C.3}$$

*Proof.* Consecutive uses of Csiszar-Korner identity [2], as in Appendix B, yield

$$\begin{aligned}
&I\left(W_k;Y_k^n\right) - I\left(W_k;\widetilde{Y}^n\right)\\
&= \sum_{i=1}^n\sum_{l=1}^M \Big[I\left(W_k;Y_{kl}(i) \mid Y_k^{i-1},\widetilde{Y}_{i+1}^n,Y_k^{l-1}(i),\widetilde{Y}_{l+1}^M(i)\right)\\
&\qquad -I\left(W_k;\widetilde{Y}_l(i) \mid Y_k^{i-1},\widetilde{Y}_{i+1}^n,Y_k^{l-1}(i),\widetilde{Y}_{l+1}^M(i)\right)\Big],
\end{aligned} \tag{C.4}$$

where each of the summand is negative, that is, we have

$$\begin{aligned}
&I\left(W_k;Y_{kl}(i) \mid Y_k^{i-1},\widetilde{Y}_{i+1}^n,Y_k^{l-1}(i),\widetilde{Y}_{l+1}^M(i)\right)\\
&\quad - I\left(W_k;\widetilde{Y}_l(i) \mid Y_k^{i-1},\widetilde{Y}_{i+1}^n,Y_k^{l-1}(i),\widetilde{Y}_{l+1}^M(i)\right) \leq 0
\end{aligned} \tag{C.5}$$

because $\widetilde{Y}_l(i)$ is the observation of the strongest user in the $l$th subchannel, that is, its channel is less noisy with respect to all other users in the $l$th subchannel. This concludes the proof of the lemma. □

This lemma implies that

$$H\left(W_k \mid \widetilde{Y}^n\right) \leq H\left(W_k \mid Y_k^n\right) \leq \epsilon_n, \tag{C.6}$$

where the second inequality is due to Fano's lemma. Using (C.6), we get

$$H\left(W_1,\ldots,W_K \mid \widetilde{Y}^n\right) \leq \sum_{k=1}^K H\left(W_k \mid \widetilde{Y}^n\right) \leq K\epsilon_n, \tag{C.7}$$

where the first inequality follows from the fact that conditioning cannot increase entropy.

We now start the converse proof:

$$\begin{aligned}
&H(W_1,\ldots,W_K \mid Z^n)\\
&\leq I\left(W_1,\ldots,W_K;\widetilde{Y}^n\right) - I(W_1,\ldots,W_K;Z^n) + K\epsilon_n
\end{aligned} \tag{C.8}$$

$$\begin{aligned}
&= \sum_{i=1}^n\sum_{l=1}^M \Big[I\left(W_1,\ldots,W_K;\widetilde{Y}_l(i) \mid Z^{i-1},\widetilde{Y}_{i+1}^n,Z^{l-1}(i),\widetilde{Y}_{l+1}^M(i)\right)\\
&\qquad - I\left(W_1,\ldots,W_K;Z_l(i) \mid Z^{i-1},\widetilde{Y}_{i+1}^n,\right.\\
&\qquad\qquad \left.Z^{l-1}(i),\widetilde{Y}_{l+1}^M(i)\right)\Big] + K\epsilon_n,
\end{aligned} \tag{C.9}$$

where (C.8) is a consequence of (C.7) and (C.9) is obtained via consecutive uses of the Csiszar-Korner identity [2] as we did in Appendix B. We define the set of indices $\mathcal{S}$ such that for all $l \in \mathcal{S}$, the strongest user in the $l$th subchannel has a less noisy channel with respect to the eavesdropper, that is, we have

$$I\left(U;\widetilde{Y}_l(i)\right) \geq I(U;Z_l(i)) \tag{C.10}$$

for all $U \rightarrow X_l(i) \rightarrow (\widetilde{Y}_l(i),Z_l(i))$ and any $l \in \mathcal{S}$. Thus, we can further bound (C.9) as follows:

$$\begin{aligned}
&H(W_1,\ldots,W_K \mid Z^n)\\
&\leq \sum_{i=1}^n\sum_{l\in\mathcal{S}} \Big[I\left(W_1,\ldots,W_K;\widetilde{Y}_l(i) \mid Z^{i-1},\widetilde{Y}_{i+1}^n,Z^{l-1}(i),\widetilde{Y}_{l+1}^M(i)\right)\\
&\qquad -I\left(W_1,\ldots,W_K;Z_l(i) \mid Z^{i-1},\widetilde{Y}_{i+1}^n,Z^{l-1}(i),\widetilde{Y}_{l+1}^M(i)\right)\Big]\\
&\quad + K\epsilon_n
\end{aligned} \tag{C.11}$$

$$\begin{aligned}
&\leq \sum_{i=1}^n\sum_{l\in\mathcal{S}} \Big[I\left(W_1,\ldots,W_K,Z^{i-1},\widetilde{Y}_{i+1}^n,Z^{l-1}(i),\widetilde{Y}_{l+1}^M(i);\widetilde{Y}_l(i)\right)\\
&\qquad -I\left(W_1,\ldots,W_K,Z^{i-1},\widetilde{Y}_{i+1}^n,Z^{l-1}(i),\widetilde{Y}_{l+1}^M(i);Z_l(i)\right)\Big]\\
&\quad + K\epsilon_n
\end{aligned} \tag{C.12}$$

$$\begin{aligned}&\leq \sum_{i=1}^{n}\sum_{l\in\mathcal{S}}\Big[I\Big(X_l(i),W_1,\ldots,W_K,Z^{i-1},\widetilde{Y}_{i+1}^{n},Z^{l-1}(i),\\&\qquad\widetilde{Y}_{l+1}^{M}(i);\widetilde{Y}_l(i)\Big)-I(X_l(i),W_1,\ldots,W_K,\\&\qquad Z^{i-1},\widetilde{Y}_{i+1}^{n},Z^{l-1}(i),\widetilde{Y}_{l+1}^{M}(i);Z_l(i)\Big)\Big]+K\epsilon_n\end{aligned}\tag{C.13}$$

$$=\sum_{i=1}^{n}\sum_{l\in\mathcal{S}}\Big[I\Big(X_l(i);\widetilde{Y}_l(i)\Big)-I(X_l(i);Z_l(i))\Big]+K\epsilon_n,\tag{C.14}$$

where (C.11) is obtained by dropping the negative terms, (C.12)-(C.13) are due to the following inequalities:

$$\begin{aligned}&I\Big(Z^{i-1},\widetilde{Y}_{i+1}^{n},Z^{l-1}(i),\widetilde{Y}_{l+1}^{M}(i);\widetilde{Y}_l(i)\Big)\\&\quad\geq I\Big(Z^{i-1},\widetilde{Y}_{i+1}^{n},Z^{l-1}(i),\widetilde{Y}_{l+1}^{M}(i);Z_l(i)\Big),\\&I\Big(X_l(i);\widetilde{Y}_l(i)\mid W_1,\ldots,W_K,Z^{i-1},\widetilde{Y}_{i+1}^{n},Z^{l-1}(i),\widetilde{Y}_{l+1}^{M}(i)\Big)\\&\quad\geq I\Big(X_l(i);Z_l(i)\mid W_1,\ldots,W_K,Z^{i-1},\widetilde{Y}_{i+1}^{n},Z^{l-1}(i),\widetilde{Y}_{l+1}^{M}(i)\Big),\end{aligned}\tag{C.15}$$

which come from the fact that for any $l\in\mathcal{S}$, the strongest user in the $l$th subchannel has a less noisy channel with respect to the eavesdropper. Finally, we get (C.14) using the following Markov chain:

$$\begin{aligned}&\Big(W_1,\ldots,W_K,Z^{i-1},\widetilde{Y}_{i+1}^{n},Z^{l-1}(i),\widetilde{Y}_{l+1}^{M}(i)\Big)\\&\quad\longrightarrow X_l(i)\longrightarrow\Big(\widetilde{Y}_l,Z_l(i)\Big),\end{aligned}\tag{C.16}$$

which is a consequence of the facts that channel is memoryless, and the subchannels are independent.

## D. Proofs of Theorems 7 and 9

*D.1. Proof of Theorem 7.* We prove Theorem 7 in two parts, first achievability and then converse. Throughout the proof, we use the shorthand notations $Y_1^n=(Y_{11}^n,Y_{12}^n)$, $Y_2^n=(Y_{21}^n,Y_{22}^n)$, $Z_1^n=(Z_1^n,Z_2^n)$.

*D.1.1. Achievability.* To show the achievability of the region given by (30), first we need to note that the boundary of this region can be decomposed into three surfaces as follows [26].

(i) First surface:

$$\begin{aligned}&R_0\leq I(U_2;Y_{12}\mid Z_2)\\&R_2\leq I(X_2;Y_{22}\mid U_2,Z_2)\\&R_0+R_1\leq I(X_1;Y_{11}\mid Z_1)+I(U_2;Y_{12}\mid Z_2),\quad U_1=\phi.\end{aligned}\tag{D.1}$$

(ii) Second surface:

$$\begin{aligned}&R_0\leq I(U_1;Y_{21}\mid Z_1)\\&R_1\leq I(X_1;Y_{11}\mid U_1,Z_1)\\&R_0+R_2\leq I(X_2;Y_{22}\mid Z_2)+I(U_1;Y_{21}\mid Z_1),\quad U_2=\phi.\end{aligned}\tag{D.2}$$

(iii) Third surface:

$$\begin{aligned}&R_0\leq I(U_1;Y_{11}\mid Z_1)+I(U_2;Y_{12}\mid Z_2)\\&R_0\leq I(U_1;Y_{21}\mid Z_1)+I(U_2;Y_{22}\mid Z_2)\\&R_1\leq I(X_1;Y_{11}\mid U_1,Z_1)\\&R_2\leq I(X_2;Y_{22}\mid U_2,Z_2).\end{aligned}\tag{D.3}$$

We now show the achievability of these regions separately. Start with the first region.

**Proposition 13.** *The region defined by* (D.1) *is achievable.*

*Proof.* Fix the probability distribution

$$p(x_1)p(u_2)p(x_2\mid u_2)p(y_1,y_2,z\mid x).\tag{D.4}$$

*Codebook Generation.*

(i) Split the private message rate of user 1 as $R_1=R_{11}+R_{12}$.

(ii) Generate $2^{n(R_{11}+\widetilde{R}_{11})}$ length-$n$ sequences $\mathbf{x}_1$ through $p(\mathbf{x}_1)=\prod_{i=1}^{n}p(x_{1,i})$ and index them as $\mathbf{x}_1(w_{11},\widetilde{w}_{11})$ where $w_{11}\in\{1,\ldots,2^{nR_{11}}\}$ and $\widetilde{w}_{11}\in\{1,\ldots,2^{n\widetilde{R}_{11}}\}$.

(iii) Generate $2^{n(R_0+R_{12}+\widetilde{R}_{12})}$ length-$n$ sequences $\mathbf{u}_2$ through $p(\mathbf{u}_2)=\prod_{i=1}^{n}p(u_{2,i})$ and index them as $\mathbf{u}_2(w_0,w_{12},\widetilde{w}_{12})$ where $w_0\in\{1,\ldots,2^{nR_0}\}$, $w_{12}\in\{1,\ldots,2^{nR_{12}}\}$ and $\widetilde{w}_{12}\in\{1,\ldots,2^{n\widetilde{R}_{12}}\}$.

(iv) For each $\mathbf{u}_2$, generate $2^{n(R_2+\widetilde{R}_2)}$ length-$n$ sequences $\mathbf{x}_2$ through $p(\mathbf{x}_2)=\prod_{i=1}^{n}p(x_{2,i}\mid u_{2,i})$ and index them as $\mathbf{x}_2(w_0,\widetilde{w}_{02},w_2,\widetilde{w}_2)$ where $w_2\in\{1,\ldots,2^{nR_2}\}$, $\widetilde{w}_2\in\{1,\ldots,2^{n\widetilde{R}_2}\}$.

(v) Furthermore, set the confusion message rates as follows:

$$\begin{aligned}&\widetilde{R}_{11}=I(X_1;Z_1),\\&\widetilde{R}_{12}=I(U_2;Z_2),\\&\widetilde{R}_2=I(X_2;Z_2\mid U_2).\end{aligned}\tag{D.5}$$

*Encoding.* If $(w_0,w_{11},w_{12},w_2)$ is the message to be transmitted, then the receiver randomly picks $(\widetilde{w}_{11},\widetilde{w}_{12},\widetilde{w}_2)$ and sends the corresponding codewords through each channel.

*Decoding.* It is straightforward to see that if the following conditions are satisfied, then both users can decode the messages directed to themselves with vanishingly small error probability.

$$\begin{aligned} &R_0 + \widetilde{R}_{12} + R_{12} \le I(U_2; Y_{12}), \\ &R_{11} + \widetilde{R}_{11} \le I(X_1; Y_{11}), \\ &R_2 + \widetilde{R}_2 \le I(X_2; Y_{22} \mid U_2). \end{aligned} \tag{D.6}$$

After eliminating $R_{11}$ and $R_{12}$ and plugging the values of $\widetilde{R}_{11}, \widetilde{R}_{12}, \widetilde{R}_2$, we can reach the following conditions:

$$\begin{aligned} &R_0 \le I(U_2; Y_{12} \mid Z_2), \\ &R_2 \le I(X_2; Y_{22} \mid U_2, Z_2), \\ &R_0 + R_1 \le I(X_1; Y_{11} \mid Z_1) + I(U_2; Y_{12} \mid Z_2), \end{aligned} \tag{D.7}$$

where we used the degradedness of the channel. Thus, we only need to show that this coding scheme satisfies the secrecy constraints.

*Equivocation Computation.* As shown previously in Lemma 11 of Appendix A, checking the sum rate secrecy condition is sufficient:

$$\begin{aligned} &H(W_0, W_1, W_2 \mid Z^n) \\ &\quad = H(W_0, W_1, W_2, Z^n) - H(Z^n) \\ &\quad = H(W_0, W_1, W_2, U_2^n, X_2^n, X_1^n, Z^n) \\ &\qquad - H(U_2^n, X_2^n, X_1^n \mid W_0, W_1, W_2, Z^n) - H(Z^n) \end{aligned} \tag{D.8}$$

$$\begin{aligned} &\quad = H(U_2^n, X_2^n, X_1^n) \\ &\qquad + H(W_0, W_1, W_2, Z^n \mid U_2^n, X_2^n, X_1^n) - H(Z^n) \\ &\qquad - H(U_2^n, X_2^n, X_1^n \mid W_0, W_1, W_2, Z^n) \end{aligned} \tag{D.9}$$

$$\begin{aligned} &\quad \ge H(U_2^n, X_2^n, X_1^n) + H(Z^n \mid U_2^n, X_2^n, X_1^n) \\ &\qquad - H(Z^n) - H(U_2^n, X_2^n, X_1^n \mid W_0, W_1, W_2, Z^n). \end{aligned} \tag{D.10}$$

We treat each term in (D.10) separately. The first term in (D.10) is

$$\begin{aligned} &H(U_2^n, X_2^n, X_1^n) \\ &\quad = H(U_2^n, X_2^n) + H(X_1^n) \\ &\quad = n\Big(R_0 + R_{11} + R_2 + R_{12} + \widetilde{R}_{11} + \widetilde{R}_{12} + \widetilde{R}_2\Big), \end{aligned} \tag{D.11}$$

where the first equality is due to the independence of $(U_2^n, X_2^n)$ and $X_1^n$, and the second equality is due the fact that both messages and confusion codewords are uniformly distributed. The second and the third terms in (D.10) are

$$\begin{aligned} &H(Z^n) - H(Z^n \mid U_2^n, X_2^n, X_1^n) \\ &\quad = H(Z_1^n, Z_2^n) - H(Z^n \mid U_2^n, X_2^n, X_1^n) \end{aligned} \tag{D.12}$$

$$\le H(Z_1^n) + H(Z_2^n) - H(Z_1^n, Z_2^n \mid U_2^n, X_2^n, X_1^n) \tag{D.13}$$

$$= H(Z_1^n) + H(Z_2^n) - H(Z_1^n, Z_2^n \mid X_2^n, X_1^n) \tag{D.14}$$

$$= H(Z_1^n) + H(Z_2^n) - H(Z_1^n \mid X_1^n) - H(Z_2^n \mid X_2^n) \tag{D.15}$$

$$= I(X_1^n; Z_1^n) + I(X_2^n; Z_2^n) \tag{D.16}$$

$$\le nI(X_1; Z_1) + nI(X_2; Z_2) + \gamma_{1,n} + \gamma_{2,n}, \tag{D.17}$$

where the equalities in (D.14) and (D.15) are due to the following Markov chains:

$$\begin{aligned} &U_2^n \longrightarrow X_2^n \longrightarrow (X_1^n, Z_1^n, Z_2^n), \\ &Z_2^n \longrightarrow X_2^n \longrightarrow X_1^n \longrightarrow Z_1^n, \end{aligned} \tag{D.18}$$

respectively, and the last inequality in (D.17) can be shown using the technique devised in [1]. To bound the last term in (D.10), assume that the eavesdropper tries to decode $(U_2^n, X_2^n, X_1^n)$ using the side information $W_0, W_1, W_2$ and its observation. Since the rates of the confusion codewords are selected such that the eavesdropper can decode them given $W_0 = w_0, W_1 = w_1, W_2 = w_2$ (see (D.5)), using Fano's lemma, we get

$$H(U_2^n, X_2^n, X_1^n \mid W_0, W_1, W_2, Z^n) \le \epsilon_n \tag{D.19}$$

for the third term in (D.10). Plugging (D.11), (D.17), and (D.19) into (D.10), we get

$$H(W_0, W_1, W_2 \mid Z^n) \ge n(R_0 + R_1 + R_2) - \epsilon_n - \gamma_{1,n} - \gamma_{2,n}, \tag{D.20}$$

which completes the proof. □

Achievability of the region defined by (D.2) follows due to symmetry. We now show the achievability of the region defined by (D.3).

**Proposition 14.** *The region described by* (D.3) *is achievable.*

*Proof.* Fix the probability distribution as follows:

$$p(u_1)p(x_1 \mid u_1)p(u_2)p(x_2 \mid u_2)p(y_1, y_2, z \mid x). \tag{D.21}$$

*Codebook Generation.*

(i) Generate $2^{n(R_0+\widetilde{R}_{01})}$ length-$n$ sequences $\mathbf{u}_1$ through $p(\mathbf{u}_1) = \prod_{i=1}^n p(u_{1,i})$ and index them as $\mathbf{u}_1(w_0, \widetilde{w}_{01})$ where $w_0 \in \{1, \ldots, 2^{nR_0}\}$, $\widetilde{w}_{01} \in \{1, \ldots, 2^{n\widetilde{R}_{01}}\}$.

(ii) For each $\mathbf{u}_1$, generate $2^{n(R_1+\widetilde{R}_1)}$ length-$n$ sequences $\mathbf{x}_1$ through $p(\mathbf{x}_1) = \prod_{i=1}^n p(x_{1,i} \mid u_{1,i})$ and index them as $\mathbf{x}_1(w_0, \widetilde{w}_{01}, w_1, \widetilde{w}_1)$ where $w_1 \in \{1, \ldots, 2^{nR_1}\}$, $\widetilde{w}_1 \in \{1, \ldots, 2^{n\widetilde{R}_1}\}$.

(iii) Generate $2^{n(R_0+\widetilde{R}_{02})}$ length-$n$ sequences $\mathbf{u}_2$ through $p(\mathbf{u}_2) = \prod_{i=1}^{n} p(u_{2,i})$ and index them as $\mathbf{u}_2(w_0, \widetilde{w}_{02})$ where $w_0 \in \{1, \dots, 2^{nR_0}\}$, $\widetilde{w}_{02} \in \{1, \dots, 2^{n\widetilde{R}_{02}}\}$.

(iv) For each $\mathbf{u}_2$, generate $2^{n(R_2+\widetilde{R}_2)}$ length-$n$ sequences $\mathbf{x}_2$ through $p(\mathbf{x}_2) = \prod_{i=1}^{n} p(x_{2,i} \mid u_{2,i})$ and index them as $\mathbf{x}_2(w_0, \widetilde{w}_{02}, w_2, \widetilde{w}_2)$ where $w_2 \in \{1, \dots, 2^{nR_2}\}$, $\widetilde{w}_2 \in \{1, \dots, 2^{n\widetilde{R}_2}\}$.

(v) Moreover, set the rates of confusion messages as follows:

$$\begin{aligned} \widetilde{R}_{01} &= I(U_1; Z_1), \\ \widetilde{R}_{02} &= I(U_2; Z_2), \\ \widetilde{R}_1 &= I(X_1; Z_1 \mid U_1), \\ \widetilde{R}_2 &= I(X_2; Z_2 \mid U_2). \end{aligned} \tag{D.22}$$

*Encoding.* Assume that the messages to be transmitted are $(w_0, w_1, w_2)$. Then, after randomly picking the tuple $(\widetilde{w}_{01}, \widetilde{w}_{02}, \widetilde{w}_1, \widetilde{w}_2)$, corresponding codewords are sent.

*Decoding.* Users decode $w_0$ using their both observations. If $w_0$ is the only message that satisfies

$$\begin{aligned} E_{i1}^{w_0} &= \{\exists \widetilde{w}_{01} : (\mathbf{u}_1(w_0, \widetilde{w}_{01}), \mathbf{y}_{i1}) \in A_\epsilon^n\} \\ E_{i2}^{w_0} &= \{\exists \widetilde{w}_{02} : (\mathbf{u}_2(w_0, \widetilde{w}_{02}), \mathbf{y}_{i2}) \in A_\epsilon^n\} \end{aligned} \tag{D.23}$$

simultaneously for user $i$, $w_0$ is declared to be transmitted. Assume $w_0 = 1$ is transmitted. The error probability for user $i$ can be bounded as

$$\Pr(E_i) \le \Pr\Big(\big(E_{i1}^1, E_{i2}^1\big)^c\Big) + \sum_{j=2}^{2^{nR_0}} \Pr\Big(E_{i1}^j, E_{i2}^j\Big), \tag{D.24}$$

using the union bound. Let us consider the following:

$$\begin{aligned} \Pr\Big(E_{i1}^j\Big) &= \Pr(\exists \widetilde{w}_{01} : (\mathbf{u}_1(j, \widetilde{w}_{01}), \mathbf{y}_{i1}) \in A_\epsilon^n) \\ &\le \sum_{\forall \widetilde{w}_{01}} \Pr((\mathbf{u}_1(j, \widetilde{w}_{01}), \mathbf{y}_{i1}) \in A_\epsilon^n) \\ &\le 2^{n\widetilde{R}_{01}} 2^{-n(I(U_1;Y_{i1})-\epsilon_n)} \\ &= 2^{n(\widetilde{R}_{01}-I(U_1;Y_{i1})+\epsilon_n)}. \end{aligned} \tag{D.25}$$

Similarly, we have

$$\Pr\Big(E_{i2}^j\Big) \le 2^{n(\widetilde{R}_{02}-I(U_2;Y_{i2})+\epsilon_n)}. \tag{D.26}$$

Thus, the probability of declaring that the $j$th message was transmitted can be bounded as

$$\begin{aligned} &\Pr\Big(E_{i1}^j, E_{i2}^j\Big) \\ &\quad = \Pr\Big(E_{i1}^j\Big) \times \Pr\Big(E_{i2}^j\Big) \\ &\quad \le 2^{n(\widetilde{R}_{01}-I(U_1;Y_{i1})+\epsilon_n)} \times 2^{n(\widetilde{R}_{02}-I(U_2;Y_{i2})+\epsilon_n)} \\ &\quad = 2^{n(\widetilde{R}_{01}-I(U_1;Y_{i1})+\widetilde{R}_{02}-I(U_2;Y_{i2})+2\epsilon_n)}, \end{aligned} \tag{D.27}$$

where the first equality is due to the independence of subchannels and codebooks used for each channel. Therefore, error probability can be bounded as

$$\begin{aligned} &\Pr(E_i) \\ &\quad \le \epsilon_n + \sum_{j=2}^{2^{nR_0}} 2^{n(\widetilde{R}_{01}-I(U_1;Y_{i1})+\widetilde{R}_{02}-I(U_2;Y_{i2})+2\epsilon_n)} \\ &\quad = \epsilon_n + 2^{n(R_0+\widetilde{R}_{01}-I(U_1;Y_{i1})+\widetilde{R}_{02}-I(U_2;Y_{i2})+2\epsilon_n)} \end{aligned} \tag{D.28}$$

which vanishes if the following are satisfied:

$$R_0 + \widetilde{R}_{01} + \widetilde{R}_{02} \le I(U_1; Y_{i1}) + I(U_2; Y_{i2}), \quad i = 1, 2. \tag{D.29}$$

After decoding the common message, both users decode their private messages if the rates satisfy

$$R_1 + \widetilde{R}_1 \le I(X_1; Y_{11} \mid U_1), \tag{D.30}$$

$$R_2 + \widetilde{R}_2 \le I(X_2; Y_{22} \mid U_2). \tag{D.31}$$

After plugging the values of $\widetilde{R}_{01}, \widetilde{R}_{02}, \widetilde{R}_1, \widetilde{R}_2$ given by (D.22) into (D.29)–(D.31), one can recover the region described by (D.3) using the degradedness of the channel.

*Equivocation Calculation.* It is sufficient to check the sum rate constraint:

$$H(W_0, W_1, W_2 \mid Z^n) = H(W_0, W_1, W_2, Z^n) - H(Z^n) \tag{D.32}$$

$$\begin{aligned} &= H(U_1^n, U_2^n, X_1^n, X_2^n, W_0, W_1, W_2, Z^n) \\ &\quad - H(U_1^n, U_2^n, X_1^n, X_2^n \mid W_0, W_1, W_2, Z^n) - H(Z^n) \end{aligned} \tag{D.33}$$

$$\begin{aligned} &= H(U_1^n, U_2^n, X_1^n, X_2^n) \\ &\quad + H(W_0, W_1, W_2, Z^n \mid U_1^n, U_2^n, X_1^n, X_2^n) \\ &\quad - H(Z^n) - H(U_1^n, U_2^n, X_1^n, X_2^n \mid W_0, W_1, W_2, Z^n) \end{aligned} \tag{D.34}$$

$$\begin{aligned} &\ge H(U_1^n, U_2^n, X_1^n, X_2^n) + H(Z^n \mid U_1^n, U_2^n, X_1^n, X_2^n) - H(Z^n) \\ &\quad - H(U_1^n, U_2^n, X_1^n, X_2^n \mid \ W_0, W_1, W_2, Z^n), \end{aligned} \tag{D.35}$$

where each term will be treated separately. The first term is

$$\begin{aligned} &H(U_1^n, U_2^n, X_1^n, X_2^n) \\ &\quad = H(U_1^n, U_2^n) + H(X_1^n \mid U_1^n, U_2^n) + H(X_2^n \mid U_1^n, U_2^n) \end{aligned} \tag{D.36}$$

$$= n\Big(R_0 + R_1 + R_2 + \widetilde{R}_{01} + \widetilde{R}_{02} + \widetilde{R}_1 + \widetilde{R}_2\Big), \tag{D.37}$$

where we first use the fact that $X_1^n$ and $X_2^n$ are independent given $(U_1^n, U_2^n)$, and secondly, we use the fact that messages are uniformly distributed. The second and third terms of (D.35) are

$$\begin{aligned} & H(Z^n) - H(Z^n \mid U_1^n, U_2^n, X_1^n, X_2^n) \\ & \quad = H(Z_1^n, Z_2^n) - H(Z_1^n \mid X_1^n) - H(Z_2^n \mid X_2^n) \end{aligned} \tag{D.38}$$

$$\le H(Z_1^n) + H(Z_2^n) - H(Z_1^n \mid X_1^n) - H(Z_2^n \mid X_2^n) \tag{D.39}$$

$$= I(X_1^n; Z_1^n) + I(X_2^n; Z_2^n) \tag{D.40}$$

$$\le nI(X_1; Z_1) + nI(X_2; Z_2) + \gamma_{1,n} + \gamma_{2,n}, \tag{D.41}$$

where the first equality is due to the independence of the subchannels. We now consider the last term of (D.35) for which assume that eavesdropper tries to decode $(U_1^n, U_2^n, X_1^n, X_2^n)$ using the side information $(W_0, W_1, W_2)$ and its observation. Since the rates of the confusion messages are selected to ensure that the eavesdropper can decode $(U_1^n, U_2^n, X_1^n, X_2^n)$ given $(W_0 = w_0, W_1 = w_1, W_2 = w_2)$ (see (D.22)), using Fano's lemma we have

$$H(U_1^n, U_2^n, X_1^n, X_2^n \mid \ W_0, W_1, W_2, Z^n) \le \epsilon_n. \tag{D.42}$$

Plugging (D.37), (D.41), and (D.42) into (D.35), we have

$$H(W_0, W_1, W_2 \ \mid Z^n) \ge n(R_0 + R_1 + R_2) - \epsilon_n - \gamma_{1,n} - \gamma_{2,n}, \tag{D.43}$$

which concludes the proof. □

*D.1.2. Converse.* First let us define the following auxiliary random variables:

$$\begin{aligned} U_{1,i} &= W_0 W_2 Y_{12}^n Y_{11}^{i-1} Z_{1,i+1}^n, \\ U_{2,i} &= W_0 W_1 Y_{21}^n Y_{22}^{i-1} Z_{2,i+1}^n, \end{aligned} \tag{D.44}$$

which satisfy the following Markov chains:

$$\begin{aligned} U_{1,i} &\longrightarrow X_{1,i} \longrightarrow (Y_{11,i}, Y_{21,i}, Z_{1,i}), \\ U_{2,i} &\longrightarrow X_{2,i} \longrightarrow (Y_{12,i}, Y_{22,i}, Z_{2,i}). \end{aligned} \tag{D.45}$$

We remark that although $U_{1,i}$ and $U_{2,i}$ are correlated, at the end of the proof, it will turn out that selection of them as independent will yield the same region. We start with the common message rate:

$$H(W_0 \mid Z^n) = H(W_0) - I(W_0; Z^n) \tag{D.46}$$

$$\le I(W_0; Y_1^n) - I(W_0; Z^n) + \epsilon_n \tag{D.47}$$

$$= I(W_0; Y_1^n \mid Z^n) + \epsilon_n \tag{D.48}$$

$$= I(W_0; Y_{12}^n \mid Z^n) + I(W_0; Y_{11}^n \mid Y_{12}^n, Z^n) + \epsilon_n \tag{D.49}$$

$$\begin{aligned} & \le I(W_0, W_1; Y_{12}^n \mid Z^n) \\ & \quad + I(W_0, W_2; Y_{11}^n \mid \ Y_{12}^n, Z^n) + \epsilon_n, \end{aligned} \tag{D.50}$$

where (D.47) is due to Fano's lemma, the equality in (D.48) is due to the fact that the eavesdropper's channel is degraded with respect to the first user's channel. We bound each term in (D.50) separately. First term is

$$I(W_0, W_1; Y_{12}^n \mid Z^n)$$

$$= \sum_{i=1}^n I\Big(W_0, W_1; Y_{12,i} \mid Y_{12}^{i-1}, Z_1^n, Z_2^n\Big) \tag{D.51}$$

$$\begin{aligned} &= \sum_{i=1}^n H\Big(Y_{12,i} \mid Y_{12}^{i-1}, Z_1^n, Z_2^n\Big) \\ &\quad - H\Big(Y_{12,i} \mid Y_{12}^{i-1}, Z_1^n, Z_2^n, W_0, W_1\Big) \end{aligned} \tag{D.52}$$

$$\begin{aligned} &\le \sum_{i=1}^n H(Y_{12,i} \mid Z_{2,i}) \\ &\quad - H\Big(Y_{12,i} \mid Y_{12}^{i-1}, Z_1^n, Z_2^n, W_0, W_1, Y_{21}^n, Y_{22}^{i-1}\Big) \end{aligned} \tag{D.53}$$

$$\begin{aligned} &= \sum_{i=1}^n H(Y_{12,i} \mid Z_{2,i}) \\ &\quad - H\Big(Y_{12,i} \mid W_0, W_1, Y_{21}^n, Y_{22}^{i-1}, Z_{2,i+1}^n, Z_{2,i}\Big) \end{aligned} \tag{D.54}$$

$$= \sum_{i=1}^n I(U_{2,i}; Y_{12,i} \mid Z_{2,i}), \tag{D.55}$$

where (D.53) follows from the fact that conditioning cannot increase entropy and the equality in (D.54) is due to the following Markov chains:

$$\begin{aligned} Z_1^n &\longrightarrow Y_{21}^n \longrightarrow (W_0, W_1, Y_{22}^n, Z_2^n, Y_{12}^n), \\ Y_{12}^{i-1} Z_2^{i-1} &\longrightarrow Y_{22}^{i-1} \longrightarrow \Big(W_0, W_1, Y_{21}^n, Y_{12,i}, Z_{2,i}^n, Z_1^n\Big), \end{aligned} \tag{D.56}$$

both of which are due to the fact that subchannels are independent, memoryless, and degraded. We now consider the second term in (D.50),

$$I(W_0, W_2; Y_{11}^n \mid Y_{12}^n, Z^n)$$

$$= \sum_{i=1}^n I\Big(W_0, W_2; Y_{11,i} \mid Y_{12}^n, Z_1^n, Z_2^n, Y_{11}^{i-1}\Big) \tag{D.57}$$

$$= \sum_{i=1}^n I\Big(W_0, W_2; Y_{11,i} \mid Y_{12}^n, Y_{11}^{i-1}, Z_{1,i+1}^n, Z_{1,i}\Big) \tag{D.58}$$

$$\le \sum_{i=1}^n I\Big(W_0, W_2, Y_{12}^n, Y_{11}^{i-1}, Z_{1,i+1}^n; Y_{11,i} \mid Z_{1,i}\Big) \tag{D.59}$$

$$= \sum_{i=1}^n I(U_{1,i}; Y_{11,i} \mid Z_{1,i}), \tag{D.60}$$

where (D.58) follows from the following Markov chains:

$$\begin{aligned} Z_2^n &\longrightarrow Y_{12}^n \longrightarrow \Big(W_0, W_2, Y_{11}^{i-1}, Z_1^n, Y_{11,i}\Big), \\ Z_1^{i-1} &\longrightarrow Y_{11}^{i-1} \longrightarrow \Big(W_0, W_2, Y_{12}^n, Z_{1,i+1}^n, Z_{1,i}, Y_{11,i}\Big), \end{aligned} \tag{D.61}$$

both of which are due to the fact that subchannels are independent, memoryless, and degraded. Plugging (D.55) and (D.60) into (D.50), we get the following outer bound on the common rate.

$$\begin{aligned} &H(W_0 \mid Z^n) \\ &\quad \le \sum_{i=1}^{n} I(U_{2,i};Y_{12,i} \mid Z_{2,i}) + \sum_{i=1}^{n} I(U_{1,i};Y_{11,i} \mid Z_{1,i}) + \epsilon_n. \end{aligned} \tag{D.62}$$

Using the same analysis on the second user, we can obtain the following outer bound on the common rate as well.

$$\begin{aligned} &H(W_0 \mid Z^n) \\ &\quad \le \sum_{i=1}^{n} I(U_{2,i};Y_{22,i} \mid Z_{2,i}) + \sum_{i=1}^{n} I(U_{1,i};Y_{21,i} \mid Z_{1,i}) + \epsilon_n. \end{aligned} \tag{D.63}$$

We now bound the sum of independent and common message rates for each user,

$$H(W_0, W_1 \mid Z^n) \le I(W_0, W_1; Y_1^n) - I(W_0, W_1; Z^n) + \epsilon_n \tag{D.64}$$

$$= I(W_0, W_1; Y_1^n \mid Z^n) + \epsilon_n \tag{D.65}$$

$$= I(W_0, W_1; Y_{11}^n, Y_{12}^n \mid Z^n) + \epsilon_n \tag{D.66}$$

$$\begin{aligned} &= I(W_0, W_1; Y_{12}^n \mid Z^n) \\ &\quad + I(W_0, W_1; Y_{11}^n \mid Y_{12}^n, Z^n) + \epsilon_n, \end{aligned} \tag{D.67}$$

where (D.64) is due to Fano's lemma, (D.65) is due to the fact that the eavesdropper's channel is degraded with respect to the first user's channel. Using (D.55), the first term in (D.67) can be bounded as

$$I(W_0, W_1; Y_{12}^n \mid Z^n) \le \sum_{i=1}^{n} I(U_{2,i}; Y_{12,i} \mid Z_{2,i}). \tag{D.68}$$

Thus, we only need to bound the second term of (D.67):

$$I(W_0, W_1; Y_{11}^n \mid Y_{12}^n, Z^n)$$

$$\begin{aligned} &= H(Y_{11}^n \mid Y_{12}^n, Z_1^n, Z_2^n) \\ &\quad - H(Y_{11}^n \mid Y_{12}^n, Z_1^n, Z_2^n, W_0, W_1) \end{aligned} \tag{D.69}$$

$$\begin{aligned} &\le H(Y_{11}^n \mid Z_1^n) \\ &\quad - H(Y_{11}^n \mid Y_{12}^n, Z_1^n, Z_2^n, W_0, W_1, X_1^n) \end{aligned} \tag{D.70}$$

$$= H(Y_{11}^n \mid Z_1^n) - H(Y_{11}^n \mid Z_1^n, X_1^n) \tag{D.71}$$

$$= I(X_1^n; Y_{11}^n \mid Z_1^n) \tag{D.72}$$

$$\le \sum_{i=1}^{n} H(Y_{11,i} \mid Z_{1,i}) - H\left(Y_{11,i} \mid Z_1^n, X_1^n, Y_{11}^{i-1}\right) \tag{D.73}$$

$$= \sum_{i=1}^{n} H(Y_{11,i} \mid Z_{1,i}) - H(Y_{11,i} \mid Z_{1,i}, X_{1,i}) \tag{D.74}$$

$$= \sum_{i=1}^{n} I(X_{1,i}; Y_{11,i} \mid Z_{1,i}), \tag{D.75}$$

where (D.70) is due to the fact that conditioning cannot increase entropy, (D.71) is due to the following Markov chain:

$$(Y_{11}^n, Z_1^n) \longrightarrow X_1^n \longrightarrow (Y_{12}^n, Z_2^n, W_0, W_1), \tag{D.76}$$

and (D.73) follows from the fact that conditioning cannot increase entropy. Finally, (D.74) is due to the fact that each subchannel is memoryless. Hence, plugging (D.68) and (D.75) into (D.67), we get the following outer bound:

$$\begin{aligned} H(W_0, W_1 \mid Z^n) &\le \sum_{i=1}^{n} I(X_{1,i}; Y_{11,i} \mid Z_{1,i}) \\ &\quad + \sum_{i=1}^{n} I(U_{2,i}; Y_{12,i} \mid Z_{2,i}) + \epsilon_n. \end{aligned} \tag{D.77}$$

Similarly, for the second user, we can get the following outer bound:

$$\begin{aligned} H(W_0, W_2 \mid Z^n) &\le \sum_{i=1}^{n} I(X_{2,i}; Y_{22,i} \mid Z_{2,i}) \\ &\quad + \sum_{i=1}^{n} I(U_{1,i}; Y_{21,i} \mid Z_{1,i}) + \epsilon_n. \end{aligned} \tag{D.78}$$

We now bound the sum rates to conclude the converse:

$$\begin{aligned} &H(W_0, W_1, W_2 \mid Z^n) \\ &\quad = H(W_0, W_1, W_2) - I(W_0, W_1, W_2; Z^n) \end{aligned} \tag{D.79}$$

$$\begin{aligned} &\le I(W_0, W_1; Y_1^n) + I(W_2; Y_2^n \mid W_0, W_1) \\ &\quad - I(W_0, W_1, W_2; Z^n) + \epsilon_n \end{aligned} \tag{D.80}$$

$$= I(W_0, W_1; Y_1^n \mid Z^n) + I(W_2; Y_2^n \mid W_0, W_1, Z^n) + \epsilon_n \tag{D.81}$$

$$\begin{aligned} &= I(W_0, W_1; Y_{12}^n \mid Z^n) + I(W_0, W_1; Y_{11}^n \mid Z^n, Y_{12}^n) \\ &\quad + I(W_2; Y_{21}^n \mid W_0, W_1, Z^n) \\ &\quad + I(W_2; Y_{22}^n \mid W_0, W_1, Z^n, Y_{21}^n) + \epsilon_n \end{aligned} \tag{D.82}$$

$$\begin{aligned} &= I(W_0, W_1, Y_{21}^n; Y_{12}^n \mid Z^n) - I(Y_{21}^n; Y_{12}^n \mid W_0, W_1, Z^n) \\ &\quad + I(W_0, W_1; Y_{11}^n \mid Z^n, Y_{12}^n) + I(W_2; Y_{21}^n \mid W_0, W_1, Z^n) \\ &\quad + I(W_2; Y_{22}^n \mid W_0, W_1, Z^n, Y_{21}^n) + \epsilon_n \end{aligned} \tag{D.83}$$

$$= S_1 - S_2 + S_3 + S_4 + S_5, \tag{D.84}$$

where (D.80) follows from Fano's lemma, (D.81) is due to the fact that the eavesdropper's channel is degraded with respect to both users' channels, (D.83) is obtained by adding and subtracting $S_2$ from the first term of (D.82). Now, we proceed as follows:

$$\begin{aligned} S_4 - S_2 &= I(W_2; Y_{21}^n \mid W_0, W_1, Z^n) \\ &\quad - I(Y_{21}^n; Y_{12}^n \mid W_0, W_1, Z^n) \end{aligned} \tag{D.85}$$

$$\begin{aligned} &\le I(W_2, Y_{12}^n; Y_{21}^n \mid W_0, W_1, Z^n) \\ &\quad - I(Y_{21}^n; Y_{12}^n \mid W_0, W_1, Z^n) \end{aligned} \tag{D.86}$$

$$= I(W_2; Y_{21}^n \mid W_0, W_1, Z^n, Y_{12}^n). \tag{D.87}$$

Adding $S_3$ to (D.87), we get

$$\begin{aligned} S_3 + S_4 - S_2 &\le I(W_0, W_1; Y_{11}^n \mid Z^n, Y_{12}^n) \\ &\quad + I(W_2; Y_{21}^n \mid W_0, W_1, Z^n, Y_{12}^n) \end{aligned} \tag{D.88}$$

$$\begin{aligned} &\le I(W_0, W_1; Y_{11}^n \mid Z^n, Y_{12}^n) \\ &\quad + I(W_2; Y_{11}^n, Y_{21}^n \mid W_0, W_1, Z^n, Y_{12}^n) \end{aligned} \tag{D.89}$$

$$\begin{aligned} &= I(W_0, W_1; Y_{11}^n \mid Z^n, Y_{12}^n) \\ &\quad + I(W_2; Y_{11}^n \mid W_0, W_1, Z^n, Y_{12}^n) \\ &\quad + I(W_2; Y_{21}^n \mid W_0, W_1, Z^n, Y_{12}^n, Y_{11}^n) \end{aligned} \tag{D.90}$$

$$\begin{aligned} &= I(W_0, W_1, W_2; Y_{11}^n \mid Z^n, Y_{12}^n) \\ &\quad + I(W_2; Y_{21}^n \mid W_0, W_1, Z^n, Y_{12}^n, Y_{11}^n), \end{aligned} \tag{D.91}$$

where the second term is zero as we show in what follows:

$$\begin{aligned} &I(W_2; Y_{21}^n \mid W_0, W_1, Z^n, Y_{12}^n, Y_{11}^n) \\ &\quad = H(W_2 \mid W_0, W_1, Z_1^n, Z_2^n, Y_{12}^n, Y_{11}^n) \\ &\qquad - H(W_2 \mid W_0, W_1, Z_1^n, Z_2^n, Y_{12}^n, Y_{11}^n, Y_{21}^n) \\ &\quad = H(W_2 \mid W_0, W_1, Y_{12}^n, Y_{11}^n) \\ &\qquad - H(W_2 \mid W_0, W_1, Y_{12}^n, Y_{11}^n) = 0, \end{aligned} \tag{D.92}$$

where we used the following Markov chain:

$$(W_0, W_1, W_2) \longrightarrow (Y_{11}^n, Y_{12}^n) \longrightarrow (Y_{21}^n, Z_1^n, Z_2^n), \tag{D.93}$$

which is a consequence of the degradation orders that subchannels exhibit. Thus, (D.91) can be expressed as

$$S_3 + S_4 - S_2 \le I(W_0, W_1, W_2; Y_{11}^n \mid Z^n, Y_{12}^n) \tag{D.94}$$

$$= I(W_0, W_1, W_2; Y_{11}^n \mid Z_1^n, Y_{12}^n) \tag{D.95}$$

$$\le I(X_1^n, W_0, W_1, W_2; Y_{11}^n \mid Z_1^n, Y_{12}^n) \tag{D.96}$$

$$\begin{aligned} &= I(X_1^n; Y_{11}^n \mid Z_1^n, Y_{12}^n) \\ &\quad + I(W_0, W_1, W_2; Y_{11}^n \mid Z_1^n, Y_{12}^n, X_1^n), \end{aligned} \tag{D.97}$$

where (D.95) follows from the following Markov chain:

$$Z_2^n \longrightarrow Y_{12}^n \longrightarrow (W_0, W_1, W_2, Y_{11}^n, Z_1^n), \tag{D.98}$$

which is due to the degradedness of the channel. Moreover, the second term in (D.97) is zero as we show in what follows:

$$\begin{aligned} &I(W_0, W_1, W_2; Y_{11}^n \mid Z_1^n, Y_{12}^n, X_1^n) \\ &\quad = H(W_0, W_1, W_2 \mid Z_1^n, Y_{12}^n, X_1^n) \\ &\qquad - H(W_0, W_1, W_2 \mid Z_1^n, Y_{12}^n, X_1^n, Y_{11}^n) \end{aligned} \tag{D.99}$$

$$\begin{aligned} &= H(W_0, W_1, W_2 \mid Y_{12}^n, X_1^n) \\ &\quad - H(W_0, W_1, W_2 \mid Y_{12}^n, X_1^n) = 0, \end{aligned} \tag{D.100}$$

where (D.100) follows from the following Markov chain:

$$(Y_{11}^n, Z_1^n) \longrightarrow X_1^n \longrightarrow (W_0, W_1, W_2, Y_{12}^n). \tag{D.101}$$

Thus, (D.97) turns out to be

$$S_3 + S_4 - S_2 \le I(X_1^n; Y_{11}^n \mid Z_1^n, Y_{12}^n). \tag{D.102}$$

which can be further bounded as follows:

$$\begin{aligned} S_3 + S_4 - S_2 &\le H(Y_{11}^n \mid Z_1^n, Y_{12}^n) \\ &\quad - H(Y_{11}^n \mid Z_1^n, Y_{12}^n, X_1^n) \end{aligned} \tag{D.103}$$

$$\begin{aligned} &\le H(Y_{11}^n \mid Z_1^n) \\ &\quad - H(Y_{11}^n \mid Z_1^n, Y_{12}^n, X_1^n) \end{aligned} \tag{D.104}$$

$$= H(Y_{11}^n \mid Z_1^n) - H(Y_{11}^n \mid Z_1^n, X_1^n) \tag{D.105}$$

$$\le \sum_{i=1}^{n} I(X_{1,i}; Y_{11,i} \mid Z_{1,i}), \tag{D.106}$$

where (D.104) is due to the fact that conditioning cannot increase entropy, (D.105) is due to the following Markov chain:

$$(Y_{11}^n, Z_1^n) \longrightarrow X_1^n \longrightarrow Y_{12}^n. \tag{D.107}$$

Finally, (D.106) is due to our previous result in (D.75). We keep bounding terms in (D.84):

$$S_5 = I(W_2; Y_{22}^n \mid W_0, W_1, Y_{21}^n, Z_1^n, Z_2^n) \tag{D.108}$$

$$= I(W_2; Y_{22}^n \mid W_0, W_1, Y_{21}^n, Z_2^n) \tag{D.109}$$

$$= \sum_{i=1}^{n} I\left(W_2; Y_{22,i} \mid W_0, W_1, Y_{21}^n, Z_2^n, Y_{22}^{i-1}\right) \tag{D.110}$$

$$= \sum_{i=1}^{n} I\left(W_2; Y_{22,i} \mid W_0, W_1, Y_{21}^n, Z_{2,i+1}^n, Y_{22}^{i-1}, Z_{2,i}\right) \tag{D.111}$$

$$= \sum_{i=1}^{n} I(W_2; Y_{22,i} \mid U_{2,i}, Z_{2,i}) \tag{D.112}$$

$$\begin{aligned} &\le \sum_{i=1}^{n} H(Y_{22,i} \mid U_{2,i}, Z_{2,i}) \\ &\quad - H(Y_{22,i} \mid U_{2,i}, Z_{2,i}, W_2, X_{2,i}) \end{aligned} \tag{D.113}$$

$$\le \sum_{i=1}^{n} I(X_{2,i}; Y_{22,i} \mid U_{2,i}, Z_{2,i}), \tag{D.114}$$

where (D.109) and (D.111) are due to the following Markov chains:

$$\begin{aligned} &Z_1^n \longrightarrow Y_{21}^n \longrightarrow (W_0, W_1, W_2, Y_{22}^n, Z_2^n), \\ &Z_2^{i-1} \longrightarrow Y_{22}^{i-1} \longrightarrow \left(W_0, W_1, W_2, Y_{21}^n, Z_{2,i}^n, Y_{22,i}\right), \end{aligned} \tag{D.115}$$

respectively, (D.113) follows from that conditioning cannot increase entropy and (D.114) is due to the following Markov chain:

$$(Y_{22,i}, Z_{2,i}) \longrightarrow X_{2,i} \longrightarrow (W_2, U_{2,i}), \tag{D.116}$$

which is a consequence of the fact that each subchannel is memoryless. Thus, we only need to bound $S_1$ in (D.84) to reach the outer bound for the sum secrecy rate:

$$\begin{aligned} S_1 &= I(W_0, W_1, Y_{21}^n; Y_{12}^n \mid Z^n) && \text{(D.117)} \\ &= \sum_{i=1}^n I\left(W_0, W_1, Y_{21}^n; Y_{12,i} \mid Z_1^n, Z_2^n, Y_{12}^{i-1}\right) && \text{(D.118)} \\ &\leq \sum_{i=1}^n H(Y_{12,i} \mid Z_{2,i}) \\ &\quad - H\left(Y_{12,i} \mid Z_1^n, Z_2^n, Y_{12}^{i-1}, W_0, W_1, Y_{21}^n, Y_{22}^{i-1}\right) && \text{(D.119)} \\ &= \sum_{i=1}^n H(Y_{12,i} \mid Z_{2,i}) \\ &\quad - H\left(Y_{12,i} \mid Z_2^n, Y_{12}^{i-1}, W_0, W_1, Y_{21}^n, Y_{22}^{i-1}\right) && \text{(D.120)} \\ &= \sum_{i=1}^n H(Y_{12,i} \mid Z_{2,i}) \\ &\quad - H\left(Y_{12,i} \mid W_0, W_1, Y_{21}^n, Y_{22}^{i-1}, Z_{2,i+1}^n, Z_{2,i}\right) && \text{(D.121)} \\ &= \sum_{i=1}^n I(U_{2,i}; Y_{12,i} \mid Z_{2,i}), && \text{(D.122)} \end{aligned}$$

where (D.119) is due to the fact that conditioning cannot increase entropy, (D.120) and (D.121) follow from the following Markov chains:

$$\begin{aligned} &Z_1^n \longrightarrow Y_{21}^n \longrightarrow \left(W_0, W_1, Y_{22}^{i-1}, Y_{12}^n, Z_2^n\right), \\ &\left(Y_{12}^{i-1}, Z_2^{i-1}\right) \longrightarrow Y_{22}^{i-1} \longrightarrow \left(W_0, W_1, W_2, Y_{21}^n, Z_{2,i}^n, Y_{12,i}\right), \end{aligned} \tag{D.123}$$

respectively. Thus, plugging (D.106), (D.114), and (D.122) into (D.84), we get the following outer bound on the sum secrecy rate:

$$\begin{aligned} &H(W_0, W_1, W_2 \mid Z^n) \\ &\quad \leq \sum_{i=1}^n I(X_{1,i}; Y_{11,i} \mid Z_{1,i}) + I(X_{2,i}; Y_{22,i} \mid U_{2,i}, Z_{2,i}) \\ &\qquad + I(U_{2,i}; Y_{12,i} \mid Z_{2,i}) + \epsilon_n. \end{aligned} \tag{D.124}$$

Following similar steps, we can also get the following one:

$$\begin{aligned} &H(W_0, W_1, W_2 \mid Z^n) \\ &\quad \leq \sum_{i=1}^n I(X_{2,i}; Y_{22,i} \mid Z_{2,i}) + I(X_{1,i}; Y_{11,i} \mid U_{1,i}, Z_{1,i}) \\ &\qquad + I(U_{1,i}; Y_{21,i} \mid Z_{1,i}) + \epsilon_n. \end{aligned} \tag{D.125}$$

So far, we derived outer bounds, (D.62), (D.63), (D.77), (D.78), (D.124), (D.125), on the capacity region which match the achievable region provided. The only difference can be on the joint distribution that they need to satisfy. However, the outer bounds depend on either $p(u_1, x_1)$ or $p(u_2, x_2)$ but not on the joint distribution $p(u_1, u_2, x_1, x_2)$. Hence, for the outer bound, it is sufficient to consider the joint distributions having the form $p(u_1, u_2, x_1, x_2) = p(u_1, x_1)p(u_2, x_2)$. Thus, the outer bounds derived and the achievable region coincide yielding the capacity region.

### *D.2. Proof of Theorem 9.*

*D.2.1. Achievability.* To show the achievability of the region given in Theorem 9, we use Theorem 7. First, we group subchannels into two sets $\mathcal{S}_j, j = 1, 2$, where $\mathcal{S}_j, j = 1, 2$, contains the subchannels in which user $j$ has the best observation. In other words, we have the Markov chain:

$$X_l \longrightarrow Y_{1l} \longrightarrow Y_{2l} \longrightarrow Z_l, \tag{D.126}$$

for $l \in \mathcal{S}_1$, and we have this Markov chain:

$$X_l \longrightarrow Y_{2l} \longrightarrow Y_{1l} \longrightarrow Z_l \tag{D.127}$$

for $l \in \mathcal{S}_2$.

We replace $U_j$ with $\{U_l\}_{l\in\mathcal{S}_j}$, $X_j$ with $\{X_l\}_{l\in\mathcal{S}_j}$, $Y_{j1}$ with $\{Y_{jl}\}_{l\in\mathcal{S}_1}$, $Y_{j2}$ with $\{Y_{jl}\}_{l\in\mathcal{S}_2}$, and $Z_j$ with $\{Z_l\}_{l\in\mathcal{S}_j}$, $j = 1, 2$, in Theorem 7. Moreover, if we select the pairs $\{(U_l, X_l)\}_{l=1}^M$ to be mutually independent, we get the following joint distribution:

$$p\left(\{u_l, x_l, y_{1l}, y_{2l}, z_l\}_{l=1}^M\right) = \prod_{l=1}^M p(u_l, x_l) p(y_{1l}, y_{2l}, z_l \mid x_l), \tag{D.128}$$

which implies that random variable tuples $\{(u_l, x_l, y_{1l}, y_{2l}, z_l)\}_{l=1}^M$ are mutually independent. Using this fact, one can reach the expressions given in Theorem 9.

*D.2.2. Converse.* For the converse part, we again use the proof of Theorem 7. First, without loss of generality, we assume $\mathcal{S}_1 = \{1, \ldots, L_1\}$, and $\mathcal{S}_2 = \{L_1 + 1, \ldots, M\}$. We define the following auxiliary random variables:

$$\begin{aligned} U_{1,i} &= W_0 W_2 Y_{1[L_1+1:M]}^n Y_{1[1:L_1]}^{i-1} Z_{[1:L_1],i+1}^n, \\ U_{2,i} &= W_0 W_1 Y_{2[1:L_1]}^n Y_{2[L_1+1:M]}^{i-1} Z_{[L_1+1:M],i+1}^n, \end{aligned} \tag{D.129}$$

which satisfy the Markov chains:

$$
\begin{aligned}
U_{1,i} &\longrightarrow X_{l,i} \longrightarrow (Y_{1l,i}, Y_{2l,i}, Z_{l,i}), \quad l = 1, \ldots, L_1, \\
U_{2,i} &\longrightarrow X_{l,i} \longrightarrow (Y_{1l,i}, Y_{2l,i}, Z_{l,i}), \quad l = L_1 + 1, \ldots, M.
\end{aligned} \tag{D.130}
$$

Using the analysis carried out for the proof of Theorem 7, we get

$$
\begin{aligned}
nR_0 \leq{}& \sum_{i=1}^{n} I(U_{1,i}; Y_{1[1:L_1],i} \mid Z_{[1:L_1],i}) \\
&+ \sum_{i=1}^{n} I(U_{2,i}; Y_{1[L_1+1:M],i} \mid Z_{[L_1+1:M],i}) + \epsilon_n,
\end{aligned} \tag{D.131}
$$

where each term will be treated separately. The first term can be bounded as follows:

$$
\begin{aligned}
& I(U_{1,i}; Y_{1[1:L_1],i} \mid Z_{[1:L_1],i}) \\
&\quad = \sum_{l=1}^{L_1} I(U_{1,i}; Y_{1l,i} \mid Y_{1[1:l-1],i}, Z_{[1:L_1],i}) && \text{(D.132)} \\
&\quad = \sum_{l=1}^{L_1} I(U_{1,i}; Y_{1l,i} \mid Y_{1[1:l-1],i}, Z_{[l:L_1],i}) && \text{(D.133)} \\
&\quad \leq \sum_{l=1}^{L_1} I(U_{1,i}, Y_{1[1:l-1],i}, Z_{[l+1:L_1],i}; Y_{1l,i} \mid Z_{l,i}), && \text{(D.134)}
\end{aligned}
$$

where (D.133) follows from the Markov chain:

$$
Z_{[1:l-1],i} \longrightarrow Y_{1[1:l-1],i} \longrightarrow (U_{1,i}, Y_{1l,i}, Z_{[l:L_1],i}), \tag{D.135}
$$

which is due to the degradedness of the subchannels. To this end, we define the following auxiliary random variables:

$$
V_{l,i} = Y_{1[1:l-1],i} Z_{[l+1:L_1],i} U_{1,i}, \quad l = 1, \ldots, L_1, \tag{D.136}
$$

which satisfy the Markov chains:

$$
V_{l,i} \longrightarrow X_{l,i} \longrightarrow (Y_{1l,i}, Y_{2l,i}, Z_{l,i}), \quad l = 1, \ldots, L_1. \tag{D.137}
$$

Thus, using these new auxiliary random variables in (D.134), we get

$$
I(U_{1,i}; Y_{1[1:L_1],i} \mid Z_{[1:L_1],i}) \leq \sum_{l=1}^{L_1} I(V_{l,i}; Y_{1l,i} \mid Z_{l,i}). \tag{D.138}
$$

We now bound the second term in (D.131) as follows:

$$
\begin{aligned}
& I(U_{2,i}; Y_{1[L_1+1:M],i} \mid Z_{[L_1+1:M],i}) \\
&\quad = \sum_{l=L_1+1}^{M} I(U_{2,i}; Y_{1l,i} \mid Z_{[L_1+1:M],i}, Y_{1[L_1+1:l-1],i}) && \text{(D.139)} \\
&\quad = \sum_{l=L_1+1}^{M} I(U_{2,i}; Y_{1l,i} \mid Z_{[l:M],i}, Y_{1[L_1+1:l-1],i}) && \text{(D.140)} \\
&\quad \leq \sum_{l=L_1+1}^{M} H(Y_{1l,i} \mid Z_{l,i}) \\
&\qquad - H(Y_{1l,i} \mid Z_{[l:M],i}, Y_{1[L_1+1:l-1],i}, U_{2,i}) && \text{(D.141)} \\
&\quad \leq \sum_{l=L_1+1}^{M} H(Y_{1l,i} \mid Z_{l,i}) \\
&\qquad - H(Y_{1l,i} \mid Z_{[l:M],i}, Y_{1[L_1+1:l-1],i}, U_{2,i}, Y_{2[L_1+1:l-1],i}) && \text{(D.142)} \\
&\quad = \sum_{l=L_1+1}^{M} H(Y_{1l,i} \mid Z_{l,i}) \\
&\qquad - H(Y_{1l,i} \mid Z_{[l:M],i}, U_{2,i}, Y_{2[L_1+1:l-1],i}) && \text{(D.143)} \\
&\quad = \sum_{l=L_1+1}^{M} I(Z_{[l+1:M],i}, U_{2,i}, Y_{2[L_1+1:l-1],i}; Y_{1l,i} \mid Z_{l,i}), && \text{(D.144)}
\end{aligned}
$$

where (D.140) follows from the Markov chain:

$$
Z_{[L_1+1:l-1],i} \longrightarrow Y_{1[L_1+1:l-1],i} \longrightarrow (U_{2,i}, Z_{[l:M],i}, Y_{1l,i}), \tag{D.145}
$$

which is a consequence of the degradedness of the subchannels, (D.141) and (D.142) follow from the fact that conditioning cannot increase entropy, and (D.143) is due to the Markov chain:

$$
Y_{1[L_1+1:l-1],i} \longrightarrow Y_{2[L_1+1:l-1],i} \longrightarrow (U_{2,i}, Z_{[l:M],i}, Y_{1l,i}), \tag{D.146}
$$

which is again a consequence of the degradedness of the subchannels. To this end, we define the following auxiliary random variables:

$$
V_{l,i} = Y_{2[L_1+1:l-1],i} Z_{[l+1:M],i} U_{2,i}, \quad l = L_1 + 1, \ldots, M, \tag{D.147}
$$

which satisfy the Markov chains:

$$
V_{l,i} \longrightarrow X_{l,i} \longrightarrow (Y_{1l,i}, Y_{2l,i}, Z_{l,i}), \quad l = L_1 + 1, \ldots, M, \tag{D.148}
$$

Thus, using these new auxiliary random variables in (D.144), we get

$$
I(U_{2,i}; Y_{1[L_1+1:M],i} \mid Z_{[L_1+1:M],i}) \leq \sum_{l=L_1+1}^{M} I(V_{l,i}; Y_{1l,i} \mid Z_{l,i}). \tag{D.149}
$$

Finally, using (D.138) and (D.149) in (D.131), we obtain

$$nR_0 \le \sum_{i=1}^{n}\sum_{l=1}^{M} I(V_{l,i}; Y_{1l,i} \mid Z_{l,i}) + \epsilon_n. \tag{D.150}$$

Due to symmetry, we also have

$$nR_0 \le \sum_{i=1}^{n}\sum_{l=1}^{M} I(V_{l,i}; Y_{2l,i} \mid Z_{l,i}) + \epsilon_n. \tag{D.151}$$

We now bound the sum of common and independent message rates. Using the converse proof of Theorem 7, we get

$$\begin{aligned} n(R_0+R_1) \le{}& \sum_{i=1}^{n} I(X_{[1:L_1],i}; Y_{1[1:L_1],i} \mid Z_{[1:L_1],i}) \\ &+ \sum_{i=1}^{n} I(U_{2,i}; Y_{1[L_1+1:M],i} \mid Z_{[L_1+M],i}) + \epsilon_n, \end{aligned} \tag{D.152}$$

where, for the second term we already obtained, an outer bound given in (D.149). We now bound the first term:

$$\begin{aligned} & I(X_{[1:L_1],i}; Y_{1[1:L_1],i} \mid Z_{[1:L_1],i}) \\ &= \sum_{l=1}^{L_1} I(X_{[1:L_1],i}; Y_{1l,i} \mid Z_{[1:L_1],i}, Y_{1[1:l-1],i}) && \text{(D.153)} \\ &\le \sum_{l=1}^{L_1} H(Y_{1l,i} \mid Z_{l,i}) \\ &\quad - H(Y_{1l,i} \mid Z_{[1:L_1],i}, Y_{1[1:l-1],i}, X_{[1:L_1],i}) && \text{(D.154)} \\ &= \sum_{l=1}^{L_1} H(Y_{1l,i} \mid Z_{l,i}) - H(Y_{1l,i} \mid Z_{l,i}, X_{l,i}) && \text{(D.155)} \\ &= \sum_{l=1}^{L_1} I(X_{l,i}; Y_{1l,i} \mid Z_{l,i}), && \text{(D.156)} \end{aligned}$$

where (D.154) follows from the fact that conditioning cannot increase entropy, and (D.155) is due to the following Markov chain:

$$\begin{aligned} & (Y_{1l,i}, Z_{l,i}) \longrightarrow X_{l,i} \\ &\quad \longrightarrow (X_{[1:l-1],i}, X_{[l+1:L_1],i}, Y_{1[1:l-1],i}Z_{[1:l-1],i}, Z_{[l+1:L_1],i}), \end{aligned} \tag{D.157}$$

which follows from the facts that channel is memoryless and subchannels are independent. Thus, plugging (D.149) and (D.156) into (D.152), we obtain

$$\begin{aligned} n(R_0+R_1) \le{}& \sum_{i=1}^{n}\sum_{l\in\mathcal{S}_1} I(X_{l,i}; Y_{1l,i} \mid Z_{l,i}) \\ &+ \sum_{i=1}^{n}\sum_{l\in\mathcal{S}_2} I(V_{l,i}; Y_{1l,i} \mid Z_{l,i}) + \epsilon_n. \end{aligned} \tag{D.158}$$

Due to symmetry, we also have

$$\begin{aligned} n(R_0+R_2) \le{}& \sum_{i=1}^{n}\sum_{l\in\mathcal{S}_2} I(X_{l,i}; Y_{2l,i} \mid Z_{l,i}) \\ &+ \sum_{i=1}^{n}\sum_{l\in\mathcal{S}_1} I(V_{l,i}; Y_{2l,i} \mid Z_{l,i}) + \epsilon_n. \end{aligned} \tag{D.159}$$

We now bound the sum secrecy rate. We first borrow the following outer bound from the converse proof of Theorem 7:

$$\begin{aligned} & n(R_0+R_1+R_2) \\ &\quad \le \sum_{i=1}^{n} I(X_{[1:L_1],i}; Y_{1[1:L_1],i} \mid Z_{[1:L_1],i}) && \text{(D.160)} \\ &\quad + \sum_{i=1}^{n} I(X_{[L_1+1:M],i}; Y_{2[L_1+1:M],i} \mid U_{2,i}, Z_{[L_1+1:M],i}) \\ &\quad + \sum_{i=1}^{n} I(U_{2,i}; Y_{1[L_1+1:M],i} \mid Z_{[L_1+1:M],i}), && \text{(D.161)} \end{aligned}$$

where, for the first and third terms, we already obtained outer bounds given in (D.156) and (D.149), respectively. We now bound the second term as follows:

$$\begin{aligned} & I(X_{[L_1+1:M],i}; Y_{2[L_1+1:M],i} \mid U_{2,i}, Z_{[L_1+1:M],i}) \\ &= \sum_{l=L_1+1}^{M} I(X_{[L_1+1:M],i}; Y_{2l,i} \mid U_{2,i}, Z_{[L_1+1:M],i}, Y_{2[L_1+1:l-1],i}) && \text{(D.162)} \\ &= \sum_{l=L_1+1}^{M} I(X_{[L_1+1:M],i}; Y_{2l,i} \mid U_{2,i}, Z_{[l:M],i}, Y_{2[L_1+1:l-1],i}) && \text{(D.163)} \\ &= \sum_{l=L_1+1}^{M} I(X_{[L_1+1:M],i}; Y_{2l,i} \mid V_{l,i}, Z_{l,i}) && \text{(D.164)} \\ &= \sum_{l=L_1+1}^{M} H(Y_{2l,i} \mid V_{l,i}, Z_{l,i}) \\ &\quad - H(Y_{2l,i} \mid V_{l,i}, Z_{l,i}, X_{[L_1+1:M],i}) && \text{(D.165)} \\ &= \sum_{l=L_1+1}^{M} H(Y_{2l,i} \mid V_{l,i}, Z_{l,i}) \\ &\quad - H(Y_{2l,i} \mid V_{l,i}, Z_{l,i}, X_{l,i}) && \text{(D.166)} \\ &= \sum_{l=L_1+1}^{M} I(X_{l,i}; Y_{2l,i} \mid V_{l,i}, Z_{l,i}), && \text{(D.167)} \end{aligned}$$

where (D.163) follows from the Markov chain:

$$Z_{[L_1+1:l-1],i} \longrightarrow Y_{2[L_1+1:l-1],i} \longrightarrow U_{2,i}, Z_{[l:M],i}, X_{[L_1+1:M],i}, Y_{2l,i}, \tag{D.168}$$

which is a consequence of the degradedness of the subchannels, (D.164) is obtained via using the definition of $V_{2,i}$ given in (D.147), and (D.166) follows from the Markov chain:

$$(Z_{l,i}, Y_{2l,i}) \longrightarrow X_{l,i} \longrightarrow \left(V_{l,i}, X_{[L_1+1:l-1],i}, X_{[l+1:M]}\right), \quad \text{(D.169)}$$

which is due to the facts that channel is memoryless and subchannels are independent. Thus, plugging (D.149), (D.156), and (D.167) into (D.161), we get

$$\begin{aligned} n(R_0+R_1+R_2) \leq & \sum_{i=1}^{n}\sum_{l\in\mathcal{S}_1} I(X_{l,i};Y_{1l,i} \mid Z_{l,i}) \\ & + \sum_{i=1}^{n}\sum_{l\in\mathcal{S}_2} I(X_{l,i};Y_{2l,i} \mid V_{l,i},Z_{l,i}) \quad \text{(D.170)} \\ & + \sum_{i=1}^{n}\sum_{l\in\mathcal{S}_2} I(V_{l,i};Y_{1l,i} \mid Z_{l,i}) + \epsilon_n. \end{aligned}$$

Due to symmetry, we also have

$$\begin{aligned} n(R_0+R_1+R_2) \leq & \sum_{i=1}^{n}\sum_{l\in\mathcal{S}_2} I(X_{l,i};Y_{2l,i} \mid Z_{l,i}) \\ & + \sum_{i=1}^{n}\sum_{l\in\mathcal{S}_1} I(X_{l,i};Y_{1l,i} \mid V_{l,i},Z_{l,i}) \quad \text{(D.171)} \\ & + \sum_{i=1}^{n}\sum_{l\in\mathcal{S}_1} I(V_{l,i};Y_{2l,i} \mid Z_{l,i}) + \epsilon_n. \end{aligned}$$

Finally, we note that all outer bounds depend on the distributions $p(v_{l,i}, x_{l,i}, y_{1l,i}, y_{2l,i}, z_{l,i}) = p(v_{l,i}, x_{l,i}) p(y_{1l,i}, y_{2l,i}, z_{l,i} \mid x_{l,i})$ but not on any joint distributions of the tuples $(v_{l,i}, x_{l,i}, y_{1l,i}, y_{2l,i}, z_{l,i})$ implying that selection of the pairs $(v_{l,i}, x_{l,i})$ to be mutually independent is optimum.

## E. Proof of Theorem 10

We prove Theorem 10 in two parts; first, we show achievability, and then we prove the converse.

*E.1. Achievability.* Similar to what we have done to show the achievability of Theorem 7, we first note that boundary of the capacity region can be decomposed into three surfaces [26].

(i) First surface:

$$\begin{aligned} & R_0 \leq \overline{\alpha} I(U_2;Y_{12} \mid Z_2) \\ & R_2 \leq \overline{\alpha} I(X_2;Y_{22} \mid U_2,Z_2) \\ & R_0+R_1 \leq \alpha I(X_1;Y_{11} \mid Z_1) + \overline{\alpha} I(U_2;Y_{12} \mid Z_2), \quad U_1 = \phi. \end{aligned} \quad \text{(E.1)}$$

(ii) Second surface:

$$\begin{aligned} & R_0 \leq \alpha I(U_1;Y_{21} \mid Z_1) \\ & R_1 \leq \alpha I(X_1;Y_{11} \mid U_1,Z_1) \\ & R_0+R_2 \leq \alpha I(U_1;Y_{21} \mid Z_1) + \overline{\alpha} I(X_2;Y_{22} \mid Z_2), \quad U_2 = \phi. \end{aligned} \quad \text{(E.2)}$$

(iii) Third surface:

$$\begin{aligned} & R_1 \leq \alpha I(X_1;Y_{11} \mid U_1,Z_1) \\ & R_2 \leq \overline{\alpha} I(X_2;Y_{22} \mid U_2,Z_2) \\ & R_0 \leq \alpha I(U_1;Y_{11} \mid Z_1) + \overline{\alpha} I(U_2;Y_{12} \mid Z_2) \\ & R_0 \leq \alpha I(U_1;Y_{21} \mid Z_1) + \overline{\alpha} I(U_2;Y_{22} \mid Z_2). \end{aligned} \quad \text{(E.3)}$$

To show the achievability of each surface, we first introduce a codebook structure.

*Codebook Structure.* Fix the probability distribution as

$$p(u_1,x_1)p(u_2,x_2)p(y_1,y_2,z \mid x). \quad \text{(E.4)}$$

(i) Generate $2^{n(R_{01}+R_{11}+\widetilde{R}_{11})}$ length-$n_1$ sequences $\mathbf{u}_1$ through $p(\mathbf{u}_1) = \prod_{i=1}^{n_1} p(u_{1,i})$ and index them as $\mathbf{u}_1(w_{01}, w_{11}, \widetilde{w}_{11})$ where $w_{01} \in \{1,\ldots,2^{nR_{01}}\}$, $w_{11} \in \{1,\ldots,2^{nR_{11}}\}$ and $\widetilde{w}_{11} \in \{1,\ldots,2^{n\widetilde{R}_{11}}\}$.

(ii) For each $\mathbf{u}_1$, generate $2^{n(R_{12}+\widetilde{R}_{12})}$ length-$n_1$ sequences $\mathbf{x}_1$ through $p(\mathbf{x}_1) = \prod_{i=1}^{n_1} p(x_{1,i} \mid u_{1,i})$ and index them as $\mathbf{x}_1(w_{01}, w_{11}, \widetilde{w}_{11}, w_{12}, \widetilde{w}_{12})$ where $w_{12} \in \{1,\ldots,2^{nR_{12}}\}$, $\widetilde{w}_{12} \in \{1,\ldots,2^{n\widetilde{R}_{12}}\}$.

(iii) Generate $2^{n(R_{02}+R_{21}+\widetilde{R}_{21})}$ length-$(n-n_1)$ sequences $\mathbf{u}_2$ through $p(\mathbf{u}_2) = \prod_{i=1}^{n-n_1} p(u_{2,i})$ and index them as $\mathbf{u}_2(w_{02}, w_{21}, \widetilde{w}_{21})$ where $w_{02} \in \{1,\ldots,2^{nR_{02}}\}$, $w_{21} \in \{1,\ldots,2^{nR_{21}}\}$ and $\widetilde{w}_{21} \in \{1,\ldots,2^{n\widetilde{R}_{21}}\}$.

(iv) For each $\mathbf{u}_2$, generate $2^{n(R_{22}+\widetilde{R}_{22})}$ length-$(n-n_1)$ sequences $\mathbf{x}_2$ through $p(\mathbf{x}_2) = \prod_{i=1}^{n-n_1} p(x_{2,i} \mid u_{2,i})$ and index them as $\mathbf{x}_2(w_{02}, w_{21}, \widetilde{w}_{21}, w_{22}, \widetilde{w}_{22})$ where $w_{22} \in \{1,\ldots,2^{nR_{22}}\}$, $\widetilde{w}_{22} \in \{1,\ldots,2^{n\widetilde{R}_{22}}\}$.

(v) We remark that this codebook uses first channel $n_1$ times and the other one $(n-n_1)$ times. We define

$$\alpha = \frac{n_1}{n} \quad \text{(E.5)}$$

and $\overline{\alpha} = 1-\alpha$.

(vi) Furthermore, we set

$$\widetilde{R}_{11} = \alpha I(U_1;Z_1), \quad \text{(E.6)}$$

$$\widetilde{R}_{12} = \alpha I(X_1;Z_1 \mid U_1), \quad \text{(E.7)}$$

$$\widetilde{R}_{21} = \overline{\alpha} I(U_2;Z_2), \quad \text{(E.8)}$$

$$\widetilde{R}_{22} = \overline{\alpha} I(X_2;Z_2 \mid U_2), \quad \text{(E.9)}$$

$$R_1 = R_{11}+R_{12}, \quad \text{(E.10)}$$

$$R_2 = R_{21}+R_{22}. \quad \text{(E.11)}$$

*Encoding.* When the transmitted messages are $(w_{01}, w_{02}, w_{11}, w_{12}, w_{21}, w_{22})$, we randomly pick $(\widetilde{w}_{11}, \widetilde{w}_{12}, \widetilde{w}_{21}, \widetilde{w}_{22})$ and send corresponding codewords.

*Decoding.* Using this codebook structure, we can show that all three surfaces which determine the boundary of the capacity region are achievable. For example, if we set $U_1 = \phi$ (that implies $R_{01} = R_{11} = \widetilde{R}_{11} = 0$) and $R_{21} = 0$, then we achieve the following rates with vanishingly small error probability:

$$\begin{aligned} R_1 &\leq \alpha I(X_1; Y_{11} \mid Z_1), \\ R_0 &\leq \overline{\alpha} I(U_2; Y_{12} \mid Z_2), \\ R_2 &\leq \overline{\alpha} I(X_2; Y_{22} \mid U_2, Z_2). \end{aligned} \tag{E.12}$$

Exchanging common message rate with user 1's independent message rate, one can obtain the first surface. Second surface follows from symmetry. For the third surface, we first set $R_{11} = R_{21} = 0$. Moreover, we send common message in its entirety, that is, we do not use a rate splitting for the common message, hence we set $R_{01} = R_{02} = R_0$, $w_{01} = w_{02} = w_0$. In this case, each user, say the $j$th one, decodes the common message by looking for a unique $w_0$ which satisfies

$$\begin{aligned} E_{j1}^{w_0} &= \left\{ \exists \widetilde{w}_{01} : \left( \mathbf{u}_1(w_0, \widetilde{w}_{01}), \mathbf{y}_{j1} \right) \in A_\epsilon^n \right\}, \\ E_{j2}^{w_0} &= \left\{ \exists \widetilde{w}_{02} : \left( \mathbf{u}_2(w_0, \widetilde{w}_{02}), \mathbf{y}_{j2} \right) \in A_\epsilon^n \right\}. \end{aligned} \tag{E.13}$$

Following the analysis carried out in (D.24)-(D.29), the sufficient conditions for the common message to be decodable by both users can be found as

$$R_0 \leq \alpha I\left( U_1; Y_{j1} \mid Z_1 \right) + \overline{\alpha} I\left( U_2; Y_{j2} \mid Z_2 \right), \quad j = 1, 2. \tag{E.14}$$

After decoding the common message, each user can decode its independent message if

$$\begin{aligned} R_1 &\leq \alpha I(X_1; Y_{11} \mid U_1, Z_1), \\ R_2 &\leq \overline{\alpha} I(X_2; Y_{22} \mid U_2, Z_2). \end{aligned} \tag{E.15}$$

Thus, the third surface can be achieved with vanishingly small error probability. As of now, we showed that all rates in the so-called capacity region are achievable with vanishingly small error probability, however we did not claim anything about the secrecy conditions which will be considered next.

*Equivocation Calculation.* To complete the achievability part of the proof, we need to show that this codebook structure also satisfies the secrecy conditions. For that purpose, it is sufficient to consider the sum rate secrecy condition:

$$\begin{aligned} & H\left( W_0, W_1, W_2 \mid Z_1^{n_1}, Z_2^{n-n_1} \right) \\ & \quad = H\left( W_0, W_1, W_2, Z_1^{n_1}, Z_2^{n-n_1} \right) - H\left( Z_1^{n_1}, Z_2^{n-n_1} \right) \end{aligned} \tag{E.16}$$

$$\begin{aligned} & = H\left( W_0, W_1, W_2, U_1^{n_1}, U_2^{n-n_1}, X_1^{n_1}, X_2^{n-n_1}, Z_1^{n_1}, Z_2^{n-n_1} \right) \\ & \quad - H\left( Z_1^{n_1}, Z_2^{n-n_1} \right) \\ & \quad - H\left( U_1^{n_1}, U_2^{n-n_1}, X_1^{n_1}, X_2^{n-n_1} \mid W_0, W_1, W_2, Z_1^{n_1}, Z_2^{n-n_1} \right) \end{aligned} \tag{E.17}$$

$$\begin{aligned} & = H\left( U_1^{n_1}, U_2^{n-n_1}, X_1^{n_1}, X_2^{n-n_1} \right) \\ & \quad + H\left( W_0, W_1, W_2, Z_1^{n_1}, Z_2^{n-n_1} \mid U_1^{n_1}, U_2^{n-n_1}, X_1^{n_1}, X_2^{n-n_1} \right) \\ & \quad - H\left( Z_1^{n_1}, Z_2^{n-n_1} \right) \\ & \quad - H\left( U_1^{n_1}, U_2^{n-n_1}, X_1^{n_1}, X_2^{n-n_1} \mid W_0, W_1, W_2, Z_1^{n_1}, Z_2^{n-n_1} \right) \end{aligned} \tag{E.18}$$

$$\begin{aligned} & \geq H\left( U_1^{n_1}, U_2^{n-n_1}, X_1^{n_1}, X_2^{n-n_1} \right) \\ & \quad + H\left( Z_1^{n_1}, Z_2^{n-n_1} \mid U_1^{n_1}, U_2^{n-n_1}, X_1^{n_1}, X_2^{n-n_1} \right) \\ & \quad - H\left( Z_1^{n_1}, Z_2^{n-n_1} \right) \\ & \quad - H\left( U_1^{n_1}, U_2^{n-n_1}, X_1^{n_1}, X_2^{n-n_1} \mid W_0, W_1, W_2, Z_1^{n_1}, Z_2^{n-n_1} \right), \end{aligned} \tag{E.19}$$

where each term will be treated separately. The first term is

$$\begin{aligned} & H\left( U_1^{n_1}, U_2^{n-n_1}, X_1^{n_1}, X_2^{n-n_1} \right) \\ & \quad = H\left( U_1^{n_1}, U_2^{n-n_1} \right) + H(X_1^{n_1} \mid U_1^{n_1}) + H\left( X_2^{n-n_1} \mid U_2^{n-n_1} \right) \end{aligned} \tag{E.20}$$

$$\begin{aligned} & = n\left( R_0 + R_{11} + \widetilde{R}_{11} + R_{21} + \widetilde{R}_{21} \right) \\ & \quad + n\left( R_{12} + \widetilde{R}_{12} \right) + n\left( R_{22} + \widetilde{R}_{22} \right) \end{aligned} \tag{E.21}$$

$$\begin{aligned} & = n(R_0 + R_1 + R_2) + n_1 I(X_1; Z_1) \\ & \quad + (n - n_1) I(X_2; Z_2), \end{aligned} \tag{E.22}$$

where the first equality is due to the Markov chain

$$X_1^{n_1} \longrightarrow U_1^{n_1} \longrightarrow U_2^{n-n_1} \longrightarrow X_2^{n-n_1}. \tag{E.23}$$

The equality in (E.21) is due to the fact that $(U_1^{n_1}, U_2^{n-n_1})$ can take $2^{n(R_0+R_{11}+\widetilde{R}_{11}+R_{21}+\widetilde{R}_{21})}$ values uniformly, and given $U_1^{n_1}$ (resp., $U_2^{n-n_1}$), $X_1^{n_1}$ (resp., $X_2^{n-n_1}$) can take $2^{n(R_{12}+\widetilde{R}_{12})}$ (resp., $2^{n(R_{22}+\widetilde{R}_{22})}$) values with equal probability. To reach (E.22), we use the definitions in (E.6)–(E.11). We consider the second and third terms in (E.19):

$$H\left(Z_1^{n_1}, Z_2^{n-n_1}\right) - H\left(Z_1^{n_1}, Z_2^{n-n_1} \mid U_1^{n_1}, U_2^{n-n_1}, X_1^{n_1}, X_2^{n-n_1}\right)$$

$$\leq H(Z_1^{n_1}) + H\left(Z_2^{n-n_1}\right) - H\left(Z_1^{n_1}, Z_2^{n-n_1} \mid U_1^{n_1}, U_2^{n-n_1}, X_1^{n_1}, X_2^{n-n_1}\right) \quad \text{(E.24)}$$

$$= H(Z_1^{n_1}) + H\left(Z_2^{n-n_1}\right) - H(Z_1^{n_1} \mid X_1^{n_1}) + H\left(Z_2^{n-n_1} \mid X_2^{n-n_1}\right) \quad \text{(E.25)}$$

$$= I(X_1^{n_1}; Z_1^{n_1}) + I\left(X_2^{n-n_1}; Z_2^{n-n_1}\right) \quad \text{(E.26)}$$

$$\leq n_1 I(X_1; Z_1) + (n - n_1) I(X_2; Z_2) + \gamma_{1,n} + \gamma_{2,n}, \quad \text{(E.27)}$$

where (E.24) is due to the fact that conditioning cannot increase entropy, (E.25) follows from the Markov chain:

$$Z_1^{n_1} \longrightarrow X_1^{n_1} \longrightarrow U_1^{n_1} \longrightarrow U_2^{n-n_1} \longrightarrow X_2^{n-n_1} \longrightarrow Z_2^{n-n_1} \quad \text{(E.28)}$$

and (E.27) can be shown using the technique devised in [1]. We bound the fourth term of (E.19). To this end, assume that the eavesdropper tries to decode $(U_1^{n_1}, X_1^{n_1}, U_2^{n-n_1}, X_2^{n-n_1})$ given side information $(W_0 = w_0, W_1 = w_1, W_2 = w_2)$. Since the confusion message rates are selected as given in (E.6)-(E.9), the eavesdropper can decode them as long as this side information is available. Consequently, the use of Fano's lemma yields

$$H\left(U_1^{n_1}, U_2^{n-n_1}, X_1^{n_1}, X_2^{n-n_1} \mid W_0, W_1, W_2, Z_1^{n_1}, Z_2^{n-n_1}\right) < \epsilon_n. \quad \text{(E.29)}$$

Finally, plugging (E.22),(E.27), and (E.29) into (E.19), we get

$$H\left(W_0, W_1, W_2 \mid Z_1^{n_1}, Z_2^{n-n_1}\right) \geq n(R_0 + R_1 + R_2) - \epsilon_n - \gamma_{1,n} - \gamma_{2,n}, \quad \text{(E.30)}$$

which completes the achievability part of the proof.

*E.2. Converse.* First, let us define the following auxiliary random variables:

$$\begin{aligned} U_{1,i} &= W_0 W_2 Y_{12}^{n-n_1} Y_{11}^{i-1} Z_{1,i+1}^{n_1}, \quad i = 1, \dots, n_1, \\ U_{2,i} &= W_0 W_1 Y_{21}^{n_1} Y_{22}^{i-1} Z_{2,i+1}^{n-n_1}, \quad i = 1, \dots, n - n_1, \end{aligned} \quad \text{(E.31)}$$

where we assume that first channel is used $n_1$ times. We again define

$$\alpha = \frac{n_1}{n}. \quad \text{(E.32)}$$

We note that the auxiliary random variables, $U_{1,i}, U_{2,i}$, satisfy the Markov chains:

$$\begin{aligned} U_{1,i} &\longrightarrow X_{1,i} \longrightarrow (Y_{11,i}, Y_{21,i}, Z_{1,i}), \\ U_{2,i} &\longrightarrow X_{2,i} \longrightarrow (Y_{21,i}, Y_{22,i}, Z_{2,i}). \end{aligned} \quad \text{(E.33)}$$

Similar to the converse of Theorem 7, here again, $U_{1,i}$ and $U_{2,i}$ can be arbitrarily correlated. However, at the end of converse, it will be clear that selection of them as independent would yield the same region. Start with the common message rate:

$$H\left(W_0 \mid Z_1^{n_1}, Z_2^{n-n_1}\right) \quad \text{(E.34)}$$

$$\leq I\left(W_0; Y_{11}^{n_1}, Y_{12}^{n-n_1}\right) - I\left(W_0; Z_1^{n_1}, Z_2^{n-n_1}\right) + \epsilon_n \quad \text{(E.35)}$$

$$= I\left(W_0; Y_{11}^{n_1}, Y_{12}^{n-n_1} \mid Z_1^{n_1}, Z_2^{n-n_1}\right) + \epsilon_n \quad \text{(E.36)}$$

$$= I\left(W_0; Y_{12}^{n-n_1} \mid Z_1^{n_1}, Z_2^{n-n_1}\right) + I\left(W_0; Y_{11}^{n_1} \mid Z_1^{n_1}, Z_2^{n-n_1}, Y_{12}^{n-n_1}\right) + \epsilon_n \quad \text{(E.37)}$$

$$\leq I\left(W_0, W_1; Y_{12}^{n-n_1} \mid Z_1^{n_1}, Z_2^{n-n_1}\right) + I\left(W_0, W_2; Y_{11}^{n_1} \mid Z_1^{n_1}, Z_2^{n-n_1}, Y_{12}^{n-n_1}\right) + \epsilon_n, \quad \text{(E.38)}$$

where (E.35) is due to Fano's lemma, (E.36) is due to the fact that the eavesdropper's channel is degraded with respect to the first user's channel. Once we obtain (E.38), using the analysis carried out in the proof of Theorem 7, we can obtain the following bounds:

$$I\left(W_0, W_1; Y_{12}^{n-n_1} \mid Z_1^{n_1}, Z_2^{n-n_1}\right) \leq \sum_{i=1}^{n-n_1} I(U_{2,i}; Y_{12,i} \mid Z_{2,i}), \quad \text{(E.39)}$$

$$I\left(W_0, W_2; Y_{11}^{n_1} \mid Z_1^{n_1}, Z_2^{n-n_1}, Y_{12}^{n-n_1}\right) \leq \sum_{i=1}^{n_1} I(U_{1,i}; Y_{11,i} \mid Z_{1,i}), \quad \text{(E.40)}$$

where (E.39) (resp., (E.40)) can be derived following the lines from (D.51) (resp., (D.57)) to (D.55) (resp., (D.60)). Thus, we have

$$H\left(W_0 \mid Z_1^{n_1}, Z_2^{n-n_1}\right) \leq \sum_{i=1}^{n-n_1} I(U_{2,i}; Y_{12,i} \mid Z_{2,i}) + \sum_{i=1}^{n_1} I(U_{1,i}; Y_{11,i} \mid Z_{1,i}) + \epsilon_n, \quad \text{(E.41)}$$

and similarly, we can get

$$H\left(W_0 \mid Z_1^{n_1}, Z_2^{n-n_1}\right) \leq \sum_{i=1}^{n-n_1} I(U_{2,i}; Y_{22,i} \mid Z_{2,i}) + \sum_{i=1}^{n_1} I(U_{1,i}; Y_{21,i} \mid Z_{1,i}) + \epsilon_n. \quad \text{(E.42)}$$

We now consider the sum of common and independent message rates:

$$\begin{aligned}
&H\left(W_0, W_1 \mid Z_1^{n_1}, Z_2^{n-n_1}\right)\\
&\quad\le I\left(W_0, W_1; Y_{11}^{n_1}, Y_{12}^{n-n_1}\right) - I\left(W_0, W_1; Z_1^{n_1}, Z_2^{n-n_1}\right) + \epsilon_n && \text{(E.43)}\\
&\quad= I\left(W_0, W_1; Y_{11}^{n_1}, Y_{12}^{n-n_1} \mid Z_1^{n_1}, Z_2^{n-n_1}\right) + \epsilon_n && \text{(E.44)}\\
&\quad= I\left(W_0, W_1; Y_{12}^{n-n_1} \mid Z_1^{n_1}, Z_2^{n-n_1}\right)\\
&\qquad + I\left(W_0, W_1; Y_{11}^{n_1} \mid Z_1^{n_1}, Z_2^{n-n_1}, Y_{12}^{n-n_1}\right) + \epsilon_n, && \text{(E.45)}
\end{aligned}$$

where (E.43) is due to Fano's lemma, (E.44) follows from the fact that the eavesdropper's channel is degraded with respect to the first user's channel. The first term of (E.45) is already bounded in (E.39). The second term can be bounded as

$$I\left(W_0, W_1; Y_{11}^{n_1} \mid Z_1^{n_1}, Z_2^{n-n_1}, Y_{12}^{n-n_1}\right) \le \sum_{i=1}^{n_1} I(X_{1,i}; Y_{11,i} \mid Z_{1,i}), \qquad \text{(E.46)}$$

which can be obtained following the lines from (D.69) to (D.75). Hence, plugging (E.39) and (E.46) into (E.45), we get

$$\begin{aligned}
&H\left(W_0, W_1 \mid Z_1^{n_1}, Z_2^{n-n_1}\right)\\
&\quad\le \sum_{i=1}^{n-n_1} I(U_{2,i}; Y_{12,i} \mid Z_{2,i}) + \sum_{i=1}^{n_1} I(X_{1,i}; Y_{11,i} \mid Z_{1,i}) + \epsilon_n. && \text{(E.47)}
\end{aligned}$$

Similarly, we can obtain

$$\begin{aligned}
&H\left(W_0, W_2 \mid Z_1^{n_1}, Z_2^{n-n_1}\right)\\
&\quad\le \sum_{i=1}^{n-n_1} I(X_{2,i}; Y_{22,i} \mid Z_{2,i}) + \sum_{i=1}^{n_1} I(U_{1,i}; Y_{21,i} \mid Z_{1,i}) + \epsilon_n. && \text{(E.48)}
\end{aligned}$$

Finally, we derive the outer bounds for the sum secrecy rate:

$$\begin{aligned}
&H\left(W_0, W_1, W_2 \mid Z_1^{n_1}, Z_2^{n-n_1}\right)\\
&\quad\le I\left(W_0, W_1; Y_{11}^{n_1}, Y_{12}^{n-n_1}\right) + I\left(W_2; Y_{21}^{n_1}, Y_{22}^{n-n_1} \mid W_0, W_1\right)\\
&\qquad - I\left(W_0, W_1, W_2; Z_1^{n_1}, Z_2^{n-n_1}\right) + \epsilon_n && \text{(E.49)}\\
&\quad= I\left(W_0, W_1; Y_{11}^{n_1}, Y_{12}^{n-n_1} \mid Z_1^{n_1}, Z_2^{n-n_1}\right)\\
&\qquad + I\left(W_2; Y_{21}^{n_1}, Y_{22}^{n-n_1} \mid W_0, W_1, Z_1^{n_1}, Z_2^{n-n_1}\right) + \epsilon_n && \text{(E.50)}\\
&\quad= I\left(W_0, W_1; Y_{12}^{n-n_1} \mid Z_1^{n_1}, Z_2^{n-n_1}\right)\\
&\qquad + I\left(W_0, W_1; Y_{11}^{n_1} \mid Z_1^{n_1}, Z_2^{n-n_1}, Y_{12}^{n-n_1}\right)\\
&\qquad + I\left(W_2; Y_{21}^{n_1} \mid W_0, W_1, Z_1^{n_1}, Z_2^{n-n_1}\right)\\
&\qquad + I\left(W_2; Y_{22}^{n-n_1} \mid W_0, W_1, Z_1^{n_1}, Z_2^{n-n_1}, Y_{21}^{n_1}\right) + \epsilon_n && \text{(E.51)}\\
&\quad= I\left(W_0, W_1, Y_{21}^{n_1}; Y_{12}^{n-n_1} \mid Z_1^{n_1}, Z_2^{n-n_1}\right)\\
&\qquad - I\left(Y_{21}^{n_1}; Y_{12}^{n-n_1} \mid Z_1^{n_1}, Z_2^{n-n_1}, W_0, W_1\right)\\
&\qquad + I\left(W_0, W_1; Y_{11}^{n_1} \mid Z_1^{n_1}, Z_2^{n-n_1}, Y_{12}^{n-n_1}\right)\\
&\qquad + I\left(W_2; Y_{21}^{n_1} \mid W_0, W_1, Z_1^{n_1}, Z_2^{n-n_1}\right)\\
&\qquad + I\left(W_2; Y_{22}^{n-n_1} \mid W_0, W_1, Z_1^{n_1}, Z_2^{n-n_1}, Y_{21}^{n_1}\right) + \epsilon_n && \text{(E.52)}\\
&\quad= S_1 - S_2 + S_3 + S_4 + S_5 + \epsilon_n, && \text{(E.53)}
\end{aligned}$$

where in (E.49), we used Fano's lemma and (E.50) follows from the fact that the eavesdropper's channel is degraded with respect to both users' channels. We can again use the analysis carried out in the converse proof of Theorem 7 to bound (E.53). For example, following lines from (D.85) to (D.106), we can obtain

$$S_4 + S_3 - S_2 \le \sum_{i=1}^{n_1} I(X_{1,i}; Y_{11,i} \mid Z_{1,i}). \qquad \text{(E.54)}$$

Similarly, if we follow the analysis from (D.108) to (D.114), we can get

$$S_5 \le \sum_{i=1}^{n-n_1} I(X_{2,i}; Y_{22,i} \mid U_{2,i}, Z_{2,i}), \qquad \text{(E.55)}$$

and if we follow the lines from (D.117) to (D.122), we can get

$$S_1 \le \sum_{i=1}^{n-n_1} I(U_{2,i}; Y_{12,i} \mid Z_{2,i}). \qquad \text{(E.56)}$$

Thus, plugging (E.54), (E.55), and (E.56) into (E.53), we get

$$\begin{aligned}
&H\left(W_0, W_1, W_2 \mid Z_1^{n_1}, Z_2^{n-n_1}\right)\\
&\quad\le \sum_{i=1}^{n_1} I(X_{1,i}; Y_{11,i} \mid Z_{1,i})\\
&\qquad + \sum_{i=1}^{n-n_1} I(U_{2,i}; Y_{12,i} \mid Z_{2,i})\\
&\qquad + \sum_{i=1}^{n-n_1} I(X_{2,i}; Y_{22,i} \mid U_{2,i}, Z_{2,i}) + \epsilon_n. && \text{(E.57)}
\end{aligned}$$

Similarly, it can be shown that

$$H\Big(W_0, W_1, W_2 \mid Z_1^{n_1}, Z_2^{n-n_1}\Big) \leq \sum_{i=1}^{n_1} I(U_{1,i}, Y_{21,i} \mid Z_{1,i}) + \sum_{i=1}^{n_1} I(X_{1,i}; Y_{11,i} \mid U_{1,i}, Z_{1,i}) + \sum_{i=1}^{n-n_1} I(X_{2,i}; Y_{22,i} \mid \ Z_{2,i}). \tag{E.58}$$

So far, we derived outer bounds on the secrecy capacity region which match the achievable region. Hence, to claim that this is indeed the capacity region, we need to show that computing the outer bounds over all distributions of the form $p(u_1, x_1)p(u_2, x_2)$ yields the same region which we would obtain by computing over all $p(u_1, u_2, x_1, x_2)$. Since all the expressions involved in the outer bounds depend on either $p(u_1, x_1)$ or $p(u_2, x_2)$ but not on the joint distribution $p(u_1, u_2, x_1, x_2)$, this argument follows, establishing the secrecy capacity region.

## Acknowledgments

This work was supported by NSF Grants CCF 04-47613, CCF 05-14846, CNS 07-16311, and CCF 07-29127, and was presented in part at the 42nd Asilomar Conference on Signals, Systems and Computers, Pacific Grove, Calif, USA, October 2008 [23].

## References

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

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

[3] E. Tekin and A. Yener, "The Gaussian multiple access wire-tap channel," *IEEE Transactions on Information Theory*, vol. 54, no. 12, pp. 5747–5755, 2008.

[4] E. Tekin and A. Yener, "The general Gaussian multiple access and two-way wire-tap channels: achievable rates and cooperative jamming," *IEEE Transactions on Information Theory*, vol. 54, no. 6, pp. 2735–2751, 2008.

[5] E. Ekrem and S. Ulukus, "On the secrecy of multiple access wiretap channel," in *Proceedings of the 46th Annual Allerton Conference on Communication, Control, and Computing*, pp. 1014–1021, September 2008.

[6] R. Liu, I. Maric, P. Spasojevic, and R. D. Yates, "Discrete memoryless interference and broadcast channels with confidential messages: secrecy rate regions," *IEEE Transactions on Information Theory*, vol. 54, no. 6, pp. 2493–2507, 2008.

[7] R. Liu and H. V. Poor, "Secrecy capacity region of a multi-antenna Gaussian broadcast channel with confidential messages," *IEEE Transactions on Information Theory*, vol. 55, no. 3, pp. 1235–1249, 2009.

[8] A. Khisti, A. Tchamkerten, and G. W. Wornell, "Secure broadcasting over fading channels," *IEEE Transactions on Information Theory*, vol. 54, no. 6, pp. 2453–2469, 2008.

[9] G. Bagherikaram, A. S. Motahari, and A. K. Khandani, "The secrecy rate region of the broadcast channel," in *Proceedings of the Allerton Conference on Communications, Control and Computing*,, July 2008.

[10] Y. Oohama, "Relay channels with confidential messages," http://arxiv.org/abs/cs/0611125.

[11] L. Lai and H. El Gamal, "The relay-eavesdropper channel: cooperation for secrecy," *IEEE Transactions on Information Theory*, vol. 54, no. 9, pp. 4005–4019, 2008.

[12] M. Yuksel and E. Erkip, "The relay channel with a wire-tapper," in *Proceedings of the 41st Annual Conference on Information Sciences and Systems*, March 2007.

[13] X. He and A. Yener, "Cooperation with an untrusted relay: a secrecy perspective," submitted to *IEEE Transactions on Information Theory*.

[14] X. He and A. Yener, "On the equivocation region of relay channels with orthogonal components," in *Proceedings of the Conference Record Asilomar Conference on Signals, Systems and Computers*, pp. 883–887, November 2007.

[15] E. Ekrem and S. Ulukus, "Effects of cooperation on the secrecy of multiple access channels with generalized feedback," in *Proceedings of the Conference on Information Sciences and Systems*, March 2008.

[16] E. Ekrem and S. Ulukus, "Secrecy in cooperative relay broadcast channels," submitted to *IEEE Transactions on Information Theory*.

[17] M. Bloch and A. Thangaraj, "Confidential messages to a cooperative relay," in *Proceedings of the IEEE Information Theory Workshop*, May 2008.

[18] Y. Liang and H. V. Poor, "Generalized multiple access channels with confidential messages," *IEEE Transactions on Information Theory*, vol. 54, no. 3, pp. 976–1002, 2008.

[19] R. Liu, I. Maric, R. D. Yates, and P. Spasojevic, "The discrete memoryless multiple access channel with confidential messages," in *Proceedings of the IEEE International Symposium on Information Theory*, July 2006.

[20] X. Tang, R. Liu, P. Spasojevic, and H. V. Poor, "Multiple access channels with generalized feedback and confidential messages," in *Proceedings of the IEEE Information Theory Workshop on Frontiers in Coding Theory*, September 2007.

[21] Y. Liang, H. V. Poor, and S. Shamai, "Secure communication over fading channels," *IEEE Transactions on Information Theory*, vol. 54, no. 6, pp. 2470–2492, 2008.

[22] Z. Li, R. Yates, and W. Trappe, "Secrecy capacity of independent parallel channels," in *Proceedings of the 44th Annual Allerton Conference on Communication, Control, and Computing*, pp. 841–848, September 2006.

[23] E. Ekrem and S. Ulukus, "On secure broadcasting," in *Proceedings of the Asilomar Conference on Signals, Systems and Computers*, October 2008.

[24] T. Cover and J. Thomas, *Elements of Information Theory*, John Wiley & Sons, 2nd edition, 2006.

[25] G. S. Poltyrev, "Capacity for a sum of certain broadcast channels," *Problemy Peredachi Informatsii*, vol. 15, no. 2, pp. 40–44, 1979.

[26] A. El Gamal, "Capacity of the product and sum of two unmatched broadcast channels," *Problemy Peredachi Informatsii*, vol. 16, no. 1, pp. 3–23, 1980.

[27] A. J. Goldsmith and M. Effros, "The capacity region of broadcast channels with intersymbol interference and colored Gaussian noise," *IEEE Transactions on Information Theory*, vol. 47, no. 1, pp. 219–240, 2001.