Degraded Broadcast Channel with Secrecy Outside a Bounded Range

Shaofeng Zou, Yingbin Liang, Lifeng Lai, H. Vincent Poor and Shlomo Shamai (Shitz)

Abstract

The K-receiver degraded broadcast channel with secrecy outside a bounded range is studied, in which a transmitter sends K messages to K receivers, and the channel quality gradually degrades from receiver K to receiver 1. Each receiver k is required to decode message $W_1, \ldots, W_k$, for $1 \leq k \leq K$, and to be kept ignorant of $W_{k+2}, \ldots, W_K$, for $k = 1, \ldots, K-2$. Thus, each message W_k is kept secure from receivers with at least two-level worse channel quality, i.e., receivers 1, ..., $k-2$. The secrecy capacity region is fully characterized. The achievable scheme designates one superposition layer to each message with random binning employed for each layer. Joint embedded coding and random binning are employed to protect all upper-layer messages from lower-layer receivers. Furthermore, the scheme allows adjacent layers to share rates so that part of the rate of each message can be shared with its immediate upper-layer message to enlarge the rate region. More importantly, an induction approach is developed to perform Fourier-Motzkin elimination of $2K$ variables from the order K^2 bounds to obtain a close-form achievable rate region. An outer bound is developed that matches the achievable rate region, whose proof involves recursive construction of the rate bounds and exploits the intuition gained from the achievable scheme.

Key words: Broadcast channel, embedded coding, random binning, rate splitting and sharing, secrecy outside a bounded range, secrecy capacity region

---

The material in this paper was presented in part at the IEEE Information Theory Workshop (ITW), Jerusalem, Israel, April 2015 [1], at the IEEE International Symposium on Information Theory (ISIT), Hong Kong, China, June 2015 [2], and at the IEEE Information Theory Workshop (ITW), Cambridge, UK, September, 2016 [3].

The work of S. Zou and Y. Liang was supported by a National Science Foundation CAREER Award under Grant CCF-10-26565 and by the National Science Foundation under Grant CCF-16-18127 and CNS-11-16932. The work of L. Lai was supported by a National Science Foundation CAREER Award under Grant CCF-13-18980 and the National Science Foundation under Grant CNS-13-21223. The work of H. V. Poor was supported by the National Science Foundation under Grant CMMI-1435778. The work of S. Shamai (Shitz) was supported by the Israel Science Foundation (ISF), and the European Commission in the framework of the Network of Excellence in Wireless COMmunications NEWCOM#.

Shaofeng Zou is with the Department of Electrical and Computer Engineering and Coordinated Science Laboratory, University of Illinois at Urbana Champaign, Urbana, IL 61801 USA (email: szou3@illinois.edu). Yingbin Liang is with the Department of Electrical Engineering and Computer Science, Syracuse University, Syracuse, NY 13244 USA (email: yliang06@syr.edu). Lifeng Lai is with the Electrical and Computer Engineering, University of California, Davis, CA 95616 USA (email: lflai@ucdavis.edu). H. Vincent Poor is with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544 USA (email: poor@princeton.edu). Shlomo Shamai (Shitz) is with the Department of Electrical Engineering, Technion-Israel Institute of Technology, Technion City, Haifa 32000 Israel (email: sshlomo@ee.technion.ac.il).

1 Introduction

The broadcast channel models an important type of scenarios in which the transmitter's signal can simultaneously reach multiple receivers, and it has been widely used to model wireless communications. Within the communication range of the transmitter, some receivers are intended while some are non-intended or even eavesdroppers from which the messages should be kept secure. Due to this broadcast nature of wireless communications, security has arisen as an important issue. Various broadcast channel models with different transmission reliability constraints (i.e., legitimate receivers should decode messages destined for them) and different secrecy constraints (i.e., eavesdroppers should be kept ignorant of messages) have been intensively studied (see recent surveys [4–7]).

The basic broadcast channel with the secrecy constraint was the wiretap channel initiated by Wyner in [8], in which a transmitter has a message intended for a legitimate receiver and wishes to keep this message secured from an eavesdropper. Csiszár and Körner further generalized this model to the case with one more common message intended for both the legitimate receiver and the eavesdropper in [9].

These broadcast models were further generalized to the multi-receiver case in [10] and [11], in which a transmitter has a number of messages intended for a set of receivers, and all messages need to be secured from an eavesdropper. Another class of extension is degraded broadcast channel with layered decoding and layered secrecy [11–13], in which the transmitter has a number of messages intended for a set of receivers and as the channel quality of a receiver gets one level better, one more message is required to be decoded, and this message is required to be secured from all receivers with worse channel quality.

We note that for the model with layered decoding and layered secrecy, the additional message decoded by a better receiver needs to be kept secure from the receiver with only one level worse channel quality (layered secrecy, zero secrecy range). Although such a model is feasible for broadcast channels with discrete states (i.e., quality of receivers can be captured by discrete channel states), it cannot capture the scenarios with receivers' channel quality varying continuously. For such a case, it is more reasonable to require the message to be secured from the receivers with a certain amount of worse channel quality, instead of being secured from the receiver with one level worse channel quality, which is not even well defined for continuous channel quality. To be more explicit, we use an example to illustrate the motivation of such a model. Consider a degraded broadcast channel with infinite number of receivers, in which h denotes the amplitude of the channel gain (the larger h, the better the channel). In this case, it is impossible to require that the message intended for receivers with $h \geq h_0$ to be secured from receivers with $h < h_0$, because no positive secrecy rate can be achieved. Instead, it is more natural to require that the messages intended for receivers with $h \geq h_0$ to be secured from receivers with $h \leq h_0 - \Delta$, where $\Delta > 0$. We refer to such a

secrecy requirement as *secrecy outside a bounded range*.

In this paper, we consider the K-receiver degraded broadcast channel with secrecy outside a bounded range (see Fig. 1), in which a transmitter sends K messages to K receivers. The channel is assumed to satisfy the degradedness condition, i.e., the channel quality gradually degrades from receiver K to receiver 1. Furthermore, receiver k is required to decode the first k messages, $W_1, \ldots, W_k$, for $1 \leq k \leq K$, and to be kept secure of $W_{k+2}, \ldots, W_K$ for $k = 1, \ldots, K-2$. Each message W_k is required to be secured from the receiver $k-2$, which has two level worse channel quality, for $3 \leq k \leq K$. In this way, the secrecy is required outside a range of two level channel quality.

The main result of this paper lies in the complete characterization of the secrecy capacity region for the K-receiver degraded broadcast channel with secrecy outside a bounded range. To understand the challenges of the problem and the novelty of the paper, we first describe special cases, namely three-receiver and four-receiver models, studied by the authors in earlier conference versions [1, 2]. For three-receiver model, we show in [1] that superposition of messages and joint random binning and embedded coding (using lower layer messages to protect higher layer messages) achieves the secrecy capacity. However, in [2] we show that a natural generalization of such a scheme does not provide the capacity region for the four-receiver model. A novel rate splitting and sharing scheme was proposed in [2], which is shown to be critical to further enlarge the achievable region and establish the secrecy capacity region for the four-receiver model. The idea is to first use lower-layer messages to serve as random sources to protect high-layer messages, and if the lower-layer messages are more than enough to protect high-layers messages, then further share remaining rate of lower-layer messages with upper-layer messages, as such part of lower layer messages can satisfy the same secrecy constraints as high-layer messages.

Further generalization of the capacity characterization for the above four-receiver model to the arbitrary K-user case becomes very challenging due to the following reasons. (1) Based on the understanding in the four-receiver model, each message as well as the random bin number at each layer can potentially serve as random source to protect all higher-layer messages (from lower layer receivers). The design of joint embedded coding and random binning is very complicated to handle. For example, consideration of whether to adopt random binning at layer k depends on whether embedded coding of layer $k-1$ is sufficient to protect W_k from receiver $k-2$, and whether embedded coding of layer $k-2$ and (possible) random binning in layer $k-1$ are sufficient to protect W_{k-1} and W_k from receiver $k-3$, and so on. Incorporating all these considerations into the design of an achievable scheme is not feasible for arbitrary K-user model. (2) Due to rate splitting and sharing across adjacent layers, the rate region is expressed in terms of individual rate components. A typical technique to convert the rate region in terms of the (total) rate for each message is Fourier-Motzkin elimination. However, for the arbitrary K-user model, a large number

of rate variables (more specifically, $2K$) should be eliminated from the order of K^2 rate bounds. Such procedure is not analytically tractable in general. (3) Due to the reason that we employ joint embedded coding and random binning to secure multiple messages, the analysis of leakage rates is much more involved than the cases with only one or two messages secured by random binning.

Despite the challenges mentioned above, in this paper, we fully characterize the secrecy capacity region for the K-receiver model with secrecy outside a bounded range. Our solution of the problem includes the following new ingredients. (1) Our achievable scheme employs random binning in each layer, which avoids the complex consideration of whether or not it is necessary to employ random binning for each layer. We also make an important observation that rate sharing only between adjacent layers is sufficient. This observation is critical to keep the obtained rate region simple enough for further manipulation. (2) We design an induction algorithm to perform Fourier-Motzkin elimination. Instead of directly eliminating $2K$ variables from the order of K^2 rate bounds, we eliminate a pair of variables at a time. We then further show that the region after each elimination step possesses a common structure by induction. (3) In order to analyze the leakage rate for arbitrary K-user case, we generalize the analysis of the leakage rate provided in [14] for one confidential message using random binning to multiple confidential messages using joint embedded coding and random binning. (4) Our development of the converse proof involves recursive construction of rate upper bound on each message such that proper terms cancel out across adjacent messages, and manipulation of the terms by exploiting intuition in achievable schemes.

The remainder of this paper is organized as follows. In Section 2, we introduce our system model. In Section 3, we present two example models with three receivers and four receivers, respectively, which motivate the design of the achievable scheme for arbitrary K-receiver model. In Section 4, we present our main results for the model with arbitrary K receivers. Finally, in Section 5, we conclude our paper.

2 Channel Model

In this paper, we consider a K-receiver degraded broadcast channel model with secrecy outside a bounded range (see Fig. 1). A transmitter sends information to K receivers through a discrete memoryless channel. The channel transition probability function is $P_{Y_1 \cdots Y_K|X}$, where $X \in \mathcal{X}$ denotes the channel input, and $Y_k \in \mathcal{Y}_k$ denotes the channel output at receiver k, for $1 \leq k \leq K$. The channel is assumed to be degraded, i.e., the following Markov chain condition holds:

$$X \to Y_K \to Y_{K-1} \to \cdots \to Y_1. \tag{1}$$

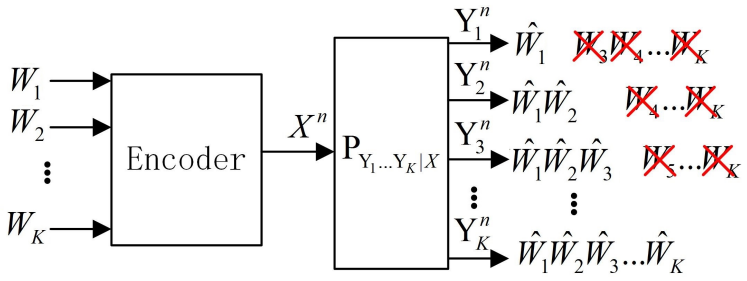

Figure 1: The K-receiver broadcast channel with secrecy outside a bounded range

Hence, the channel quality gradually degrades from receiver K to receiver 1. There are in total K messages $W_1, W_2, \ldots, W_K$ intended for K receivers with the following decoding and secrecy requirements. Receiver k is required to decode messages $W_1, W_2, \ldots, W_k$, for $k = 1, 2, \ldots, K$, and to be kept secure of $W_{k+2}, \ldots, W_K$, for $k = 1, \ldots, K-2$ (see Fig. 1).

A $(2^{nR_1}, \cdots, 2^{nR_K}, n)$ code for the channel consists of

- K message sets: $W_k \in \mathcal{W}_k = \{1, \cdots, 2^{nR_k}\}$ for $k = 1, \cdots, K$, which are independent from each other and each message is uniformly distributed over the corresponding message set;
- A (possibly stochastic) encoder f^n: $\mathcal{W}_1 \times \cdots \times \mathcal{W}_K \to \mathcal{X}^n$ that maps a message tuple to an input x^n;
- K decoders $g_k^n : \mathcal{Y}_k^n \to (\mathcal{W}_1, \cdots, \mathcal{W}_k)$ that maps an output y_k^n to a message tuple $(\hat{w}_1, \ldots, \hat{w}_k)$ for $k = 1, \cdots, K$.

A rate tuple $(R_1, \cdots, R_K)$ is said to be *achievable*, if there exists a sequence of $(2^{nR_1}, \cdots, 2^{nR_K}, n)$ codes such that both the average error probability

$$P_e^n = \Pr\left(\cup_{k=1}^K \{(W_1, \cdots, W_k) \neq g_k^n(Y_k^n)\}\right) \tag{2}$$

and the leakage rate at each receiver k for $k = 3, \ldots, K$

$$\frac{1}{n} I(W_k, \cdots, W_K; Y_{k-2}^n) \tag{3}$$

approach zero as n goes to infinity.

Here, the asymptotically small error probability as in (2) implies that each receiver k is able to decode messages $W_1, \ldots, W_k$, and asymptotically small leakage rate as in (3) for each receiver k implies that receiver k is kept ignorant of messages $W_{k+2}, \ldots, W_K$. Our goal is to characterize the *secrecy capacity region* that consists of all achievable rate tuples.

3 Motivating Examples

In this section, we study two simple cases with $K = 3$ and $K = 4$. The purpose is to motivate the development of the optimal achievable scheme for the case with arbitrary K receivers step by step. More specifically, we study the example with three receivers to introduce the technique of joint embedded coding and random binning. And we study the example with four receivers to introduce the technique of rate splitting and sharing. These schemes turn out to be necessary to achieve the secrecy capacity region for the case with arbitrary K receivers.

3.1 Lessons Learned from $K = 3$

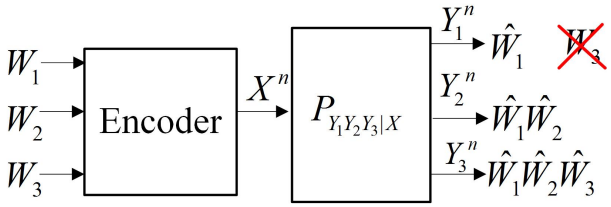

Figure 2: The three-receiver broadcast channel with secrecy outside a bounded range

We start with the case in which there are three receivers (see Fig. 2). In this case, receiver 1 is required to decode W_1, receiver 2 is required to decode W_1, W_2, and receiver 3 is required to decode W_1, W_2, W_3. The system is also required to satisfy the secrecy constraint that the message W_3 is kept secure from receiver 1.

For such a model, a natural idea is to design superposition coding for encoding three messages W_1, W_2, W_3 into three layers, and then apply random binning in the third layer to protect W_3 from receiver 1. However, such a scheme is suboptimal because it ignores an important fact that the random message W_2, which is not required to be decoded by receiver 1, can provide additional randomness to protect W_3 from receiver 1. This is referred to as *embedded coding*. In fact, if such a random source of W_2 is sufficient to protect W_3 from receiver 1, random binning is not necessary. If this is not sufficient to protect W_3, we apply random binning in the third layer to further protect W_3 from receiver 1. The novelty of such an achievable scheme lies in exploiting the superposition layer of W_2 as embedded coding in addition to the random binning scheme to protect W_3. Such a scheme turns out to achieve the secrecy capacity region as characterized in the following proposition.

Proposition 1. *Consider the three-receiver degraded broadcast channel with secrecy outside a bounded range as described in Section 2. The secrecy capacity region contains rate tuples*

(R_1, R_2, R_3) *satisfying*

$$\begin{aligned} R_1 &\le I(U_1; Y_1), \\ R_2 &\le I(U_2; Y_2|U_1), \\ R_3 &\le \min\{0, I(U_2; Y_2|U_1) - I(X; Y_1|U_1)\} + I(X; Y_3|U_2) \end{aligned} \tag{4}$$

for some $P_{U_1U_2X}$ *such that the following Markov chain condition holds*

$$U_1 \to U_2 \to X \to Y_3 \to Y_2 \to Y_1. \tag{5}$$

Proof. The proof can be found in [1]. □

The idea of the achievable scheme is also reflected in the expression of the capacity region in (4). The two bounds in "min" are corresponding to the two cases with the second layer of W_2 being sufficient and insufficient to protect W_3, respectively. If $I(U_2; Y_2|U_1) > I(X; Y_1|U_1)$, the randomness of W_2 is sufficient to exhaust receiver 1's decoding capability, and hence is good enough for protecting W_3. Thus, in this case, no binning is required in layer 3, and $R_3 \le I(X; Y_3|U_2)$. On the other hand, if $I(U_2; Y_2|U_1) \le I(X; Y_1|U_1)$, binning is required at layer 3 to protect W_3 in addition to randomness of W_2, and hence, $R_3 \le I(U_2; Y_2|U_1) - I(X; Y_1|U_1) + I(X; Y_3|U_2)$.

3.2 Lessons Learned from $K = 4$

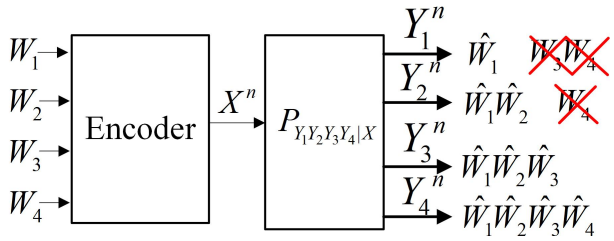

Figure 3: The four-receiver broadcast channel with secrecy outside a bounded range

In this subsection, we study the model with four receivers (see Fig. 3). In this model, receiver k is required to decode messages $W_1, \ldots, W_k$, for $1 \le k \le 4$. Furthermore, the message W_3 is required to be secured from receiver 1, and the message W_4 is required to be secured from receivers 1 and 2.

Although this four-receiver model seems to be a straightforward generalization of the three-receiver model, our exploration turns out to show that the achievable scheme for the

three-receiver model is not sufficient to establish the secrecy capacity region for the four-receiver model. In order to understand this, we note that a direct generalization of the achievable scheme for the three-model involves first applying superposition coding to encode the four messages, and then use the random message W_3 as embedded coding together with the random binning in layer 4 (if necessary) to protect W_4, and use the random message W_2 as embedded coding together with the random binning in layer 3 and layer 4 (if necessary) to protect W_3 and W_4. Such a scheme then yields an achievable region with rate tuples (R_1, R_2, R_3, R_4) satisfying

$$\begin{aligned}
R_1 &\leq I(U_1; Y_1), \\
R_2 &\leq I(U_2; Y_2; U_1), \\
R_3 &\leq I(U_3; Y_3|U_2) + \min\Big(0, I(U_2; Y_2|U_1) - I(U_3; Y_1|U_1)\Big), \\
R_4 &\leq I(X; Y_4|U_3), \\
R_4 &\leq I(X; Y_4|U_3) + I(U_3; Y_3|U_2) - I(X; Y_2|U_2), \\
R_3 + R_4 &\leq I(U_3; Y_3|U_2) + I(X; Y_4|U_3) + I(U_2; Y_2|U_1) - I(X; Y_1|U_1),
\end{aligned} \tag{6}$$

for some $P_{U_1U_2U_3X}$ satisfying the Markov chain condition $U_1 \to U_2 \to U_3 \to X \to Y_4 \to \cdots \to Y_1$. It turns out to be very difficult to develop the converse proof for the bound $R_4 \leq I(X; Y_4|U_3)$ in the above region. Thus, the optimality of the region (6) cannot be guaranteed.

The major novelty in our scheme for this four-receiver model lies in the design of rate splitting and sharing, which helps enlarge the achievable region and thus establish the secrecy capacity region. In order to further enlarge the above region, we incorporate the newly designed ingredient of rate splitting and sharing. More specifically, if W_3 is sufficient to protect W_4, we further split W_3 into two parts, i.e., $W_{3,1}$ and $W_{3,2}$, such that $W_{3,1}$ serves as a random source to secure both $W_{3,2}$ and W_4 from receiver 2. Thus,$W_{3,2}$ satisfies both the decoding and secrecy requirements for W_4, and hence the rate of $W_{3,2}$ can be counted towards the rate of either W_3 or W_4. In this way, the achievable region can be potentially enlarged. In fact, such an enlarged region can be shown to be the secrecy capacity region as characterized in the following proposition.

Proposition 2. *Consider the four-receiver degraded broadcast channel with secrecy outside a bounded range as described in Section 2. The secrecy capacity region consists of rate tuples*

(R_1, R_2, R_3, R_4) *satisfying*

$$\begin{aligned}
R_1 &\le I(U_1; Y_1), \\
R_2 &\le I(U_2; Y_2|U_1), \\
R_3 &\le I(U_3; Y_3|U_2) + \min\Big(0, I(U_2; Y_2|U_1) - I(U_3; Y_1|U_1)\Big), \\
R_4 &\le I(X; Y_4|U_3) + I(U_3; Y_3|U_2) - I(X; Y_2|U_2), \\
R_3 + R_4 &\le I(U_3; Y_3|U_2) + I(X; Y_4|U_3) + \min\Big(0, I(U_2; Y_2|U_1) - I(X; Y_1|U_1)\Big), \qquad (7)
\end{aligned}$$

for some $P_{U_1U_2U_3X}$ *such that the following Markov chain condition holds*

$$U_1 \to U_2 \to U_3 \to X \to Y_4 \to Y_3 \to Y_2 \to Y_1. \qquad (8)$$

Proof. The proof can be found in [2]. □

We note that by using rate splitting and sharing, the bound $R_4 \le I(X; Y_4|U_3)$ in the region (6) is replaced by the bound $R_3 + R_4 \le I(U_3; Y_3|U_2) + I(X; Y_4|U_3)$ in the region (7). Clearly, the region (7) is larger than the region (6) (for a given distribution of auxiliary random variables). Furthermore, the converse proof for the new bound on $R_3 + R_4$ in (7) can be derived, and thus establishes the region (7) as the secrecy capacity region.

Moreover, although we learn useful coding ingredients from the three-receiver and four-receiver cases, direct generalization to arbitrary K-receiver model still gives rise to an analytically intractable achievable scheme. More specifically, the consideration of whether or not to use random binning in the higher layers and whether or not to split and share the rates will be complex. For example, when $K = 5$, whether to use random binning in the fifth layer depends on whether the embedded coding in the third layer and (possibly) random binning in the fourth layer are sufficient to protect W_4, W_5 from receiver 3 and whether the embedded coding in the fourth layer is sufficient to protect W_5 from receiver 3. Such considerations become intractable when K is large. Thus, the major design issue for the arbitrary K-receiver case is to develop an achievable scheme that effectively incorporates the necessary coding ingredients as well as yielding a tractable rate region for analysis. This is the focus of the following section.

4 Main Results

In this section, we first present our main result of characterization of the secrecy capacity region for the K-receiver model, and then describe the idea of the design of the achievable scheme.

4.1 Secrecy Capacity Region

The following theorem states our main result. For simplicity of notation, we define $U_K = X$.

Theorem 1. *Consider the K-receiver degraded broadcast channel with secrecy outside a bounded range as described in Section 2. The secrecy capacity region consists of rate tuples $(R_1, R_2, \ldots, R_K)$ satisfying*

$$R_1 \leq I(U_1; Y_1), \tag{9a}$$

$$\sum_{j=2}^{k} R_j \leq \sum_{j=2}^{k} I(U_j; Y_j|U_{j-1}), \quad \text{for } 2 \leq k \leq K, \tag{9b}$$

$$\sum_{j=l}^{k} R_j \leq \left(\sum_{j=l-1}^{k} I(U_j; Y_j|U_{j-1}) \right) - I(U_k; Y_{l-2}|U_{l-2}), \quad \text{for } 3 \leq l \leq k \leq K, \tag{9c}$$

for some $P_{U_1 U_2 \ldots U_K}$ such that the following Markov chain condition holds:

$$U_1 \to U_2 \to \cdots \to U_K \to Y_K \to \cdots \to Y_2 \to Y_1. \tag{10}$$

Proof. The proof of the achievability and the proof of converse are provided in Appendices A and D, respectively. □

In the above capacity region, the bounds (9a) and (9b) are due to the decoding requirements, i.e., receiver k should decode messages $W_1, \ldots, W_k$, for $1 \leq k \leq K$. The sum rate bounds (9b) are due to the rate sharing scheme we design. The bounds (9c) are due to the secrecy requirements, i.e., messages $W_l, \ldots, W_k$ need to be kept secure from receiver $l-2$ for $3 \leq l \leq k \leq K$. Furthermore, the bounds (9c) can be further written as

$$\sum_{j=l}^{k} R_i \leq \sum_{j=l-1}^{k} \Big(I(U_j; Y_j|U_{j-1}) - I(U_j; Y_{l-2}|U_{j-1}) \Big),$$

which has clear intuitive interpretation. The term $I(U_j; Y_j|U_{j-1}) - I(U_j; Y_{l-2}|U_{j-1})$ is corresponding to the rate in layer j that can be secured from receiver $l-2$ given the knowledge of layer $j-1$. Those rates $I(U_j; Y_j|U_{j-1}) - I(U_j; Y_{l-2}|U_{j-1})$ for $l-1 \leq j \leq k$ can all be counted towards $\sum_{j=l}^{k} R_j$ in accordance to the secrecy requirement of keeping $W_l, \ldots, W_k$ secured from receiver $l-2$.

If we set $K = 3$ and $K = 4$, the region in Theorem 1 reduces to equivalent but different forms from the regions in Proposition 1 and Proposition 2. The equivalence is justified by the converse proofs. In fact, the achievable schemes for the three-receiver model in Section 3.1 and the four-receiver model in Section 3.2 cannot be easily generalized to the arbitrary K-receiver model.

Our design of the achievable scheme for the general arbitrary K-receiver model is different from those for the three-receiver and four receiver models, and includes the following new ingredients. The scheme employs random binning in each layer, which avoids the complex consideration of whether or not it is necessary to employ random binning for each layer. And the rate sharing scheme is limited only between adjacent layers which captures the essence of the problem and helps simplify the obtained rate region. Furthermore, we design an induction algorithm to perform Fourier-Motzkin elimination, which makes the problem of eliminating $2K$ variables from the order of K^2 bounds analytically tractable. These ideas are described in more detail in Subsection 4.2.

The converse for the achievable region can be developed. The bounds (9a) and (9b) can be derived following standard steps. However, the proof for the bounds (9c) is more involved and requires careful recursive construction of the terms such that proper terms cancel out across adjacent messages.

4.2 Achievable Scheme

In this subsection, we introduce the idea of the achievable scheme for the arbitrary K-receiver model, which is based on superposition coding, random binning, embedded coding, and rate splitting and sharing. We also sketch the novel induction idea to analyze Fourier-Motzkin elimination to characterize the achievable region.

Superposition, binning, embedded coding. We design one layer of codebook for each message, i.e., layer k corresponds to W_k, for $1 \leq k \leq K$. To avoid the complex consideration of whether to use random binning, we employ random binning in each layer. We divide the codewords in each layer into a number of bins, where the bin number contains the information of the corresponding message. We use joint embedded coding and random binning to provide randomness for secrecy.

Rate splitting and sharing. We design rate splitting and sharing to enlarge the achievable region. More specifically, within the k-th layer, we split the message W_k into two parts $W_{k,1}$ and $W_{k,2}$. The message $W_{k,1}$ serves as embedded coding which is a random source in addition to the random binning to protect $W_{k,2}$ and the higher layer messages from receiver Y_{k-1}, i.e., we require that $(W_{k,2}, W_{k+1,1}, W_{k+1,2}, \ldots, W_{K,1}, W_{K,2})$ are secured from receiver Y_{k-1}, for $2 \leq k \leq K-1$. Furthermore, the upstream receiver Y_{k+1} can also decode $W_{k,2}$ because Y_{k+1} has a better channel quality than Y_k. Thus, the message $W_{k,2}$ satisfies both the decoding and secrecy requirements for message W_{k+1}, and hence, the rate of $W_{k,2}$ can be counted towards the rate of either W_k or W_{k+1}. By such a rate sharing strategy, the achievable region is enlarged.

We note that the rate can only be shared between adjacent receivers, which is an important

observation of this problem, and is critical to reduce the complexity of the design of the rate splitting and sharing strategy. More specifically, the rate of $W_{k,2}$ cannot be counted towards the rates of $W_{k+2}, \ldots, W_K$, because $W_{k+2}, \ldots, W_K$ are required to be secured not only from receiver Y_{k-1} but also from Y_k that are required to decode $W_{k,2}$.

Based on the above achievable scheme, we obtain the following achievable region:

$$\begin{aligned} R_1 &\leq I(U_1; Y_1), \\ R_{k,1} + R_{k,2} &\leq I(U_k; Y_k | U_{k-1}), \text{ for } 2 \leq k \leq K, \\ R_{k-1,2} + \sum_{i=k}^{j} (R_{i,1} + R_{i,2}) &\leq \sum_{i=k-1}^{j} I(U_i; Y_i | U_{i-1}) - I(U_j; Y_{k-2} | U_{k-2}), \\ &\quad \text{for } 3 \leq k \leq K, k-1 \leq j \leq K. \end{aligned} \tag{11}$$

The above region are expressed in terms of component rates due to rate splitting. In order to express the above region in terms of total rate for each message, we introduce the technique of rate sharing. We define $R_k = R_{k-1,2} + R_{k,1}$ for $3 \leq k \leq K-1$, $R_2 = R_{2,1}$ and $R_K = R_{K-1,2} + R_{K,1} + R_{K,2}$. We then wish to project the region (11) onto the rate space $(R_1, \ldots, R_K)$. This can be done by adding the above rate definitions to the achievable region (11) and then perform the Fourier-Motzkin elimination to eliminate $R_{k,1}$ and $R_{k,2}$ for $2 \leq k \leq K$.

Fourier-Motzkin elimination via induction. The total number of bounds in the achievable region (11) is on the order of K^2 with $2K$ variables to be eliminated. Directly applying Fourier-Motzkin elimination is not analytically tractable. In order to solve such a problem, we design the following induction algorithm to perform Fourier Motzkin elimination. We eliminate the rate pairs $R_{k-1,2}$ and $R_{k,1}$ for $3 \leq k \leq K$ one at each step, and wish to show that the region $\mathcal{R}_k$ after eliminating $R_{k-1,2}$ and $R_{k,1}$ possesses the following

structure:

$$
\begin{aligned}
R_1 &\le I(U_1;Y_1),\\
\sum_{i=2}^{j} R_i &\le \sum_{i=2}^{j} I(U_i;Y_i|U_{i-1}),\ \text{for } 2\le j\le k-1,\\
\sum_{i=2}^{k} R_i + R_{k,2} &\le \sum_{i=2}^{k} I(U_i;Y_i|U_{i-1}),\\
\sum_{i=l}^{j} R_i &\le \sum_{i=l-1}^{j} I(U_i;Y_i|U_{i-1}) - I(U_j;Y_{l-2}|U_{l-2}),\\
&\qquad \text{for } 3\le l\le j\le k-1,\\
\sum_{i=l}^{k} R_i + R_{k,2} &\le \sum_{i=l-1}^{k} I(U_i;Y_i|U_{i-1}) - I(U_k;Y_{l-2}|U_{l-2}),\\
&\qquad \text{for } 3\le l\le k+1. \qquad (12)
\end{aligned}
$$

Such a claim can be easily verified for the case when $k = 3, 4, 5$. If such a claim holds for $\mathcal{R}_k$, we then are able to show (see Appendix C for detailed proof) that the region $\mathcal{R}_{k+1}$ after eliminating $R_{k,2}$ and $R_{k+1,1}$ possesses the same structure given by

$$
\begin{aligned}
R_1 &\le I(U_1;Y_1), \qquad (13)\\
\sum_{i=2}^{j} R_i &\le \sum_{i=2}^{j} I(U_i;Y_i|U_{i-1}),\ \text{for } 2\le j\le k, \qquad (14)\\
\sum_{i=2}^{k+1} R_i + R_{k+1,2} &\le \sum_{i=2}^{k+1} I(U_i;Y_i|U_{i-1}), \qquad (15)\\
\sum_{i=l}^{j} R_i &\le \sum_{i=l-1}^{j} I(U_i;Y_i|U_{i-1}) - I(U_j;Y_{l-2}|U_{l-2}), \qquad (16)\\
&\qquad \text{for } 3\le l\le j\le k, \qquad (17)\\
\sum_{i=l}^{k+1} R_i + R_{k+1,2} &\le \sum_{i=l-1}^{k+1} I(U_i;Y_i|U_{i-1}) - I(U_{k+1};Y_{l-2}|U_{l-2}), \qquad (18)\\
&\qquad \text{for } 3\le l\le k+2. \qquad (19)
\end{aligned}
$$

The last step is to eliminate $(R_{K-1,2}, R_{K,1}, R_{K,2})$. Thus, the above induction algorithm and arguments yield the achievable region in Theorem 1.

5 Conclusion

In this paper, we have studied the K-receiver degraded broadcast channel with secrecy outside a bounded range. We have proposed a novel achievable scheme based on superposition coding, joint embedded coding and random binning, and rate splitting and sharing. The combination of embedded coding and random binning to achieve secrecy captures the fact that lower-layer message can serve as embedded coding to protect higher-layer messages. And rate splitting and sharing are critical to enlarge the achievable region for which the converse proof can be established. Moreover, our design exploits an important property that the rate sharing should be only between adjacent receivers, which significantly reduces the complexity of the achievable scheme. We have further proposed a novel induction algorithm to perform Fourier-Motzkin elimination on the achievable region with $2K$ variables to be eliminated from the order of K^2 bounds. We have also constructed a converse proof, which involves careful recursive construction of rate bounds, and exploits the intuition gained from embedded coding in the achievable scheme. By the converse proof, we have demonstrated the optimality of our achievable scheme and established the secrecy capacity region.

Several directions are interesting to explore in the future. The current results can be further generalized to the model with continuously changing channel state parameters, e.g., Gaussian fading channel [15], and Gaussian multiple input multiple output (MIMO) channel [16]. This paper has focused on characterizing the information theoretic performance limits which is based on random coding arguments. It is of further interest to design practical coding schemes such as low density parity check (LDPC) codes [17] and polar codes [18, 19] to achieve the secrecy capacity region.

Appendix

A Achievability Proof of Theorem 1

The achievability proof is based on superposition coding, embedded coding, random binning, rate splitting and sharing. We use random codes and fix a distribution $P_{U_1U_2\ldots,U_{K-1}X}P_{Y_1\ldots Y_K|X}$ satisfying the Markov chain condition in (10). We let $T_\epsilon^n(P_{U_1\ldots U_{K-1}XY_1\ldots Y_K})$ denote the strongly jointly ϵ-typical set based on the fixed distribution. We design the achievable scheme as follows:

Random codebook generation: For simplicity, we define $U_K = X$ in the following proof, i.e., $P_{U_1\cdots U_{K-1}X} = P_{U_1\cdots U_K}$.

- Generate 2^{nR_1} independent and identically distributed (i.i.d.) u_1^n with distribution $\prod_{i=1}^n P(u_{1,i})$. Index these codewords as $u_1^n(w_1)$, $w_1 \in [1, 2^{nR_1}]$.
- For each $u_1^n(w_1)$, generate $2^{n(R_{2,1}+R_{2,2})}$ i.i.d. u_2^n with distribution $\prod_{i=1}^n P(u_{2,i}|u_{1,i})$. Partition these codewords into $2^{nR_{2,2}}$ bins. Index these codewords as $u_2^n(w_1, w_{2,1}, w_{2,2})$, $w_{2,1} \in [1, 2^{nR_{2,1}}]$, $w_{2,2} \in [1, 2^{nR_{2,2}}]$.
- For each $u_2^n(w_1, w_{2,1}, w_{2,2})$, generate $2^{n\widetilde{R}_3}$ i.i.d. u_3^n with distribution $\prod_{i=1}^n P(u_{3,i}|u_{2,i})$. Partition these codewords into $2^{nR_{3,1}}$ bins, and further partition each bin into $2^{nR_{3,2}}$ sub-bins. Hence, there are $2^{n(\widetilde{R}_3 - R_{3,1} - R_{3,2})}$ u_3^n in each sub-bin. We use $w_{3,1} \in [1 : 2^{nR_{3,1}}]$ to denote the bin number, $w_{3,2} \in [1 : 2^{nR_{3,2}}]$ to denote the sub-bin number, and $l_3 \in [1 : 2^{n(\widetilde{R}_3 - R_{3,1} - R_{3,2})}]$ to denote the index within the bin. Hence, each u_3^n is indexed by $(w_1, w_{2,1}, w_{2,2}, w_{3,1}, w_{3,2}, l_3)$.
- For $4 \le k \le K$, for each $u_{k-1}^n(w_1, \ldots, w_{k-1,1}, w_{k-1,2}, l_{k-1})$, generate $2^{n\widetilde{R}_k}$ i.i.d. u_k^n with distribution $\prod_{i=1}^n P(u_{k,i}|u_{k-1,i})$. Partition these codewords into $2^{nR_{k,1}}$ bins, and further partition each bin into $2^{nR_{k,2}}$ sub-bins. Hence, there are $2^{n(\widetilde{R}_k - R_{k,1} - R_{k,2})}$ u_k^n in each sub-bin. We use $w_{k,1} \in [1 : 2^{nR_{k,1}}]$ to denote the bin number, $w_{k,2} \in [1 : 2^{nR_{k,2}}]$ to denote the sub-bin number, and $l_k \in [1 : 2^{n(\widetilde{R}_k - R_{k,1} - R_{k,2})}]$ to denote the index within the bin. Hence, each u_k^n is indexed by $(w_1, \ldots, w_{k-1,1}, w_{k-1,2}, l_{k-1}, w_{k,1}, w_{k,2}, l_k)$.

The codebook is revealed to both the transmitter and the receivers.

Encoding:

To send a message tuple $(w_1, w_{2,1}, w_{2,2}, \ldots, w_{K,1}, w_{K,2})$, the transmitter randomly and uniformly generates $l_k \in [1 : 2^{n(\widetilde{R}_k - R_{k,1} - R_{k,2})}]$ for $3 \le k \le K$, and sends $x^n(w_1, \ldots, w_{K,1}, w_{K,2}, l_3, \ldots, l_K)$.

Decoding:

- Receiver 1 claims that $\widehat{w}_1$ is sent, if there exists a unique $\widehat{w}_1$ such that
$$\left(u_1^n(\widehat{w}_1), y_1^n\right) \in T_\epsilon^n(P_{U_1Y_1}).$$
Otherwise, it declares an error.
- Receiver 2 claims that $(\widehat{w}_1, \widehat{w}_{2,1}, \widehat{w}_{2,2})$ is sent, if there exists a unique tuple $(\widehat{w}_1, \widehat{w}_{2,1}, \widehat{w}_{2,2})$ such that
$$\left(u_1^n(\widehat{w}_1), u_2^n(\widehat{w}_1, \widehat{w}_{2,1}, \widehat{w}_{2,2}), y_2^n\right) \in T_\epsilon^n(P_{U_1U_2Y_2}).$$
Otherwise, it declares an error.
- For $3 \le k \le K$, receiver k claims that $(\widehat{w}_1, \ldots, \widehat{w}_{k,1}, \widehat{w}_{k,2})$ is sent, if there exists a unique tuple $(\widehat{w}_1, \ldots, \widehat{w}_{k,1}, \widehat{w}_{k,2}, \widehat{l}_3, \ldots, \widehat{l}_k)$ such that
$$\left(u_1^n(\widehat{w}_1), \ldots, u_k^n(\widehat{w}_1, \ldots, \widehat{w}_{k,1}, \widehat{w}_{k,2}, \widehat{l}_3, \ldots, \widehat{l}_k), y_k^n\right) \in T_\epsilon^n(P_{U_1 \cdots U_k Y_k}).$$
Otherwise, it declares an error.

Analysis of error probability: By the law of large numbers and packing lemma, it can be shown that receiver k decodes the message $(w_1, \ldots, w_{k,1}, w_{k,2})$ for $2 \leq k \leq K$ and receiver 1 decodes the message w_1 with asymptotically small error probabilities if the following inequalities are satisfied:

$$\begin{aligned} R_1 &\leq I(U_1; Y_1), \\ R_{2,1} + R_{2,2} &\leq I(U_2; Y_2|U_1), \\ \widetilde{R}_k &\leq I(U_k; Y_k|U_{k-1}), \text{ for } 3 \leq k \leq K. \end{aligned} \tag{20}$$

Analysis of leakage rate: We require that $W_{k-1,2}, W_{k,1}, W_{k,2}, \ldots, W_{K,1}, W_{K,2}$ be secured from receiver Y_{k-2} for $3 \leq k \leq K$. It suffices to show that the average of the leakage rate over the random codebook ensemble converges to zero, as $n \to \infty$, i.e.,

$$\frac{1}{n} I\left(W_{k-1,2}, W_{k,1}, W_{k,2}, \ldots, W_{K,1}, W_{K,2}; Y_{k-2}^n|\mathcal{C}\right) \to 0, \text{ as } n \to \infty, \tag{21}$$

for $3 \leq k \leq K$. This implies the existence of one codebook that guarantees secrecy.

We note that l_k in random codebook generation is a realization of the random variable L_k here. And we let $L_3^K = (L_3, \ldots, L_K)$. We generalize the analysis in [14] for one secure message via random binning to multiple secure messages via joint embedded coding and random binning. For any $3 \leq k \leq K$, we derive the following bound:

$$\begin{aligned} & I(W_{k-1,2}, W_{k,1}, W_{k,2}, \ldots, W_{K,1}, W_{K,2}; Y_{k-2}^n|\mathcal{C}) \\ & \overset{(a)}{=} I(W_1, W_{2,1}, W_{2,2}, \ldots, W_{K,1}, W_{K,2}, L_3^K; Y_{k-2}^n|\mathcal{C}) \\ & \quad - I(W_1, W_{2,1}, W_{2,2}, \ldots, W_{k-1,1}, L_3^K; Y_{k-2}^n|W_{k-1,2}, \ldots, W_{K,2}, \mathcal{C}) \\ & \overset{(b)}{\leq} I(U_K^n; Y_{k-2}^n|\mathcal{C}) - I(W_1, W_{2,1}, W_{2,2}, \ldots, W_{k-1,1}, L_3^K; Y_{k-2}^n|W_{k-1,2}, \ldots, W_{K,2}, \mathcal{C}) \\ & = I(U_K^n; Y_{k-2}^n|\mathcal{C}) - H(W_1, W_{2,1}, W_{2,2}, \ldots, W_{k-1,1}, L_3^K|W_{k-1,2}, \ldots, W_{K,2}, \mathcal{C}) \\ & \quad + H(W_1, W_{2,1}, W_{2,2}, \ldots, W_{k-1,1}, L_3^K|W_{k-1,2}, \ldots, W_{K,2}, Y_{k-2}^n, \mathcal{C}), \end{aligned} \tag{22}$$

where (a) follows from the chain rule of mutual information, and (b) is due to the Markov chain condition $(W_1, \ldots, W_{K,1}, W_{K,2}, L_3^K) \to (U_K^n, \mathcal{C}) \to Y_{k-2}^n$.

We then bound the three terms in (22) one by one. We bound the first term as follows:

$$
\begin{aligned}
&I(U_K^n; Y_{k-2}^n|\mathcal{C}) \\
&\overset{(a)}{=} I(U_{k-2}^n, U_K^n; Y_{k-2}^n|\mathcal{C}) \\
&= I(U_{k-2}^n; Y_{k-2}^n|\mathcal{C}) + I(U_K^n; Y_{k-2}^n|U_{k-2}^n, \mathcal{C}) \\
&\leq H(U_{k-2}^n|\mathcal{C}) + I(U_K^n; Y_{k-2}^n|U_{k-2}^n, \mathcal{C}) \\
&= nR_1 + nR_{2,1} + nR_{2,2} + n\sum_{i=3}^{k-2} \widetilde{R}_i + H(Y_{k-2}^n|U_{k-2}^n, \mathcal{C}) - H(Y_{k-2}^n|U_{k-2}^n, U_K^n, \mathcal{C}) \\
&= nR_1 + nR_{2,1} + nR_{2,2} + n\sum_{i=3}^{k-2} \widetilde{R}_i + \sum_{i=1}^{n} H(Y_{k-2,i}|Y_{k-2,i+1}^n, U_{k-2}^n, \mathcal{C}) - H(Y_{k-2,i}|U_K^n, Y_{k-2,i+1}^n, \mathcal{C}) \\
&\overset{(b)}{\leq} nR_1 + nR_{2,1} + nR_{2,2} + n\sum_{i=3}^{k-2} \widetilde{R}_i + \sum_{i=1}^{n} H(Y_{k-2,i}|U_{k-2,i}) - H(Y_{k-2,i}|U_{K,i}) \\
&= nR_1 + nR_{2,1} + nR_{2,2} + n\sum_{i=3}^{k-2} \widetilde{R}_i + nH(Y_{k-2}|U_{k-2}) - nH(Y_{k-2}|U_K) \\
&= nR_1 + nR_{2,1} + nR_{2,2} + n\sum_{i=3}^{k-2} \widetilde{R}_i + nI(U_K; Y_{k-2}|U_{k-2}) \qquad (23)
\end{aligned}
$$

where (a) is due to the Markov chain condition $U_{k-2}^n \to U_K^n \to Y_{k-2}^n$, and (b) is due to the fact that $H(Y_{k-2,i}|Y_{k-2,i+1}^n, U_{k-2}^n, \mathcal{C}) \leq H(Y_{k-2,i}|U_{k-2,i})$ and the Markov chain condition $(U_K^{i-1}, U_{K,i+1}^n, Y_{k-2,i+1}^n, \mathcal{C}) \to U_{K,i} \to Y_{k-2,i}$.

For the second term, due to the independence of the messages $W_1, W_{2,1}, W_{2,2}, \ldots, W_{K,1}, W_{K,2}$ and $L_3, \ldots, L_K$, we have

$$
\begin{aligned}
&- H(W_1, W_{2,1}, W_{2,2}, \ldots, W_{k-1,1}, L_3^K|W_{k-1,2}, \ldots, W_{K,2}, \mathcal{C}) \\
&= -n\bigg(R_1 + R_{2,1} + R_{2,2} + \sum_{i=3}^{k-2} \widetilde{R}_i + \widetilde{R}_{k-1} - R_{k-1,2} + \sum_{i=k}^{K} (\widetilde{R}_i - R_{i,1} - R_{i,2})\bigg). \qquad (24)
\end{aligned}
$$

We bound the third term as follows:

$$
\begin{aligned}
&H(W_1, W_{2,1}, W_{2,2}, \ldots, W_{k-1,1}, L_3^K | W_{k-1,2}, \ldots, W_{K,2}, Y_{k-2}^n, \mathcal{C}) \\
&= H(W_1, W_{2,1}, W_{2,2}, \ldots, W_{k-2,1}, W_{k-2,2}, L_3^{k-2} | W_{k-1,2}, \ldots, W_{K,2}, Y_{k-2}^n, \mathcal{C}) \\
&\quad + H(W_{k-1,1}, L_{k-1}^K | W_1, W_{2,1}, W_{2,2}, \ldots, W_{k-2,2}, W_{k-1,2} \ldots, W_{K,2}, L_3^{k-2}, Y_{k-2}^n, \mathcal{C}) \\
&\overset{(a)}{\leq} n\epsilon_n + H(W_{k-1,1}, L_{k-1}^K | W_1, W_{2,1}, W_{2,2}, \ldots, W_{k-2,2}, W_{k-1,2} \ldots, W_{K,2}, L_3^{k-2}, Y_{k-2}^n, \mathcal{C}) \\
&\overset{(b)}{=} n\epsilon_n + H(W_{k-1,1}, L_{k-1}^K | U_{k-2}^n, W_1, W_{2,1}, W_{2,2}, \ldots, W_{k-2,2}, W_{k-1,2} \ldots, W_{K,2}, L_3^{k-2}, Y_{k-2}^n, \mathcal{C}) \\
&\leq n\epsilon_n + H(W_{k-1,1}, L_{k-1}^K | U_{k-2}^n, W_{k-1,2} \ldots, W_{K,2}, Y_{k-2}^n, \mathcal{C}) \\
&\overset{(c)}{\leq} n\left(\widetilde{R}_{k-1} - R_{k-1,2} + \sum_{i=k}^{K} (\widetilde{R}_i - R_{i,1} - R_{i,2}) - I(U_K; Y_{k-2} | U_{k-2}) + 2\epsilon_n \right) \qquad (25)
\end{aligned}
$$

where (a) is due to Fano's inequality [20], (b) is due to the fact that U_{k-2}^n is a function of $(\mathcal{C}, W_1, \ldots, W_{k-2,2}, L_3^{k-2})$, and (c) follows from Lemma 1 given below if the conditions (26) are satisfied.

Lemma 1. *For $3 \leq k \leq K$, if the following conditions are satisfied:*

$$
\begin{aligned}
\widetilde{R}_{k-1} - R_{k-1,2} + \sum_{i=k}^{j} \widetilde{R}_i - R_{i,1} - R_{i,2} \geq I(U_j; Y_{k-2} | U_{k-2}), \\
\text{for } k-1 \leq j \leq K, \qquad (26)
\end{aligned}
$$

then the following inequality holds:

$$
\begin{aligned}
&H(W_{k-1,1}, L_{k-1}^K | W_{k-1,2} \ldots, W_{K,2}, U_{k-2}^n, Y_{k-2}^n, \mathcal{C}) \\
&\leq n\left(\widetilde{R}_{k-1} - R_{k-1,2} + \sum_{i=k}^{K} (\widetilde{R}_i - R_{i,1} - R_{i,2}) - I(U_K; Y_{k-2} | U_{k-2}) + \epsilon_n \right). \qquad (27)
\end{aligned}
$$

Proof. See Appendix B. □

Combining (23), (24) and (25), we conclude that if (26) holds, then

$$
\begin{aligned}
\frac{1}{n} I(W_{k-1,2}, W_{k,1}, W_{k,2}, \ldots, W_{K,1}, W_{K,2}; Y_{k-2}^n | \mathcal{C}) \to 0, \text{ as } n \to \infty, \\
\text{for } 3 \leq k \leq K. \qquad (28)
\end{aligned}
$$

Combining the bounds in (20) and (26), we conclude that the rate tuple $(R_1, R_{2,1},$

$R_{2,2}, \ldots, R_{K,1}, R_{K,2})$ is achievable if

$$\begin{aligned} R_1 &\leq I(U_1; Y_1), \\ R_{k,1} + R_{k,2} &\leq I(U_k; Y_k|U_{k-1}), \text{ for } 2 \leq k \leq K, \\ R_{k-1,2} + \sum_{i=k}^{j}(R_{i,1} + R_{i,2}) &\leq \sum_{i=k-1}^{j} I(U_i; Y_i|U_{i-1}) - I(U_j; Y_{k-2}|U_{k-2}), \\ &\qquad \text{for } 3 \leq k \leq K, \text{ and } k-1 \leq j \leq K. \end{aligned} \tag{29}$$

Rate Sharing: We note that our achievable scheme guarantees $W_{k-1,2}, W_{k,1}, W_{k,2}, \ldots, W_{K,1}$, $W_{K,2}$ to be secured from receiver Y_{k-2}, for $3 \leq k \leq K$. Furthermore, due to the degradedness condition, $W_{k-1,2}$ can be decoded by receiver Y_k. Thus, $W_{k-1,2}$ satisfies both the decoding and secrecy requirements as W_k. Hence, the rate of $W_{k-1,2}$ can be counted towards either R_{k-1} or R_k. Based on such an understanding, we design the following rate sharing scheme. We define $R_2 = R_{2,1}$, $R_k = R_{k-1,2} + R_{k,1}$ for $3 \leq k \leq K-1$, and $R_K = R_{K-1,2} + R_{K,1} + R_{K,2}$, and include these equations to the above achievable region. We then perform Fourier-Motzkin elimination to eliminate $R_{k,1}, R_{k,2}$ for $2 \leq k \leq K$ and obtain a closed-form achievable rate region. Such a process involves eliminating $2K-2$ variables from the order of K^2 bounds, which is intractable for arbitrary K. We propose an inductive Fourier Motzkin elimination approach as shown in Appendix C, and obtain the achievable region given in Theorem 1.

B Proof of Lemma 1

We bound $H(W_{k-1,1}, L_{k-1}^K | w_{k-1,2}, \ldots, w_{K,2}, U_{k-2}^n, Y_{k-2}^n, \mathcal{C})$ for each given tuple $(w_{k-1,2}, \ldots, w_{K,2})$, and hence, $H(W_{k-1,1}, L_{k-1}^K | W_{k-1,2} \ldots, W_{K,2}, U_{k-2}^n, Y_{k-2}^n, \mathcal{C})$ is bounded.

We fix $(W_{k-1,1}, L_{k-1}, \ldots, L_K) = (w_{k-1,1}, l_{k-1}, \ldots, l_K)$ and a jointly typical sequence $(u_{k-2}^n, Y_{k-2}^n) \in T_\epsilon^{(n)}(U_{k-2}, Y_{k-2})$. We define the following random variable:

$$\begin{aligned} N(w_{k-1,1}, l_{k-1}, \ldots, l_K, u_{k-2}^n, y_{k-2}^n) = |\{&(\widetilde{w}_{k-1,1}, \widetilde{l}_{k-1}, \ldots, \widetilde{l}_K) \neq (w_{k-1,1}, l_{k-1}, \ldots, l_K) \\ &: (U_K^n(\widetilde{w}_{k-1,1}, \widetilde{l}_{k-1}, \ldots, \widetilde{l}_K, w_{k-1,2} \ldots, w_{K,2}), u_{k-2}^n, y_{k-2}^n) \in T_\epsilon^{(n)}(U_K, U_{k-2}, Y_{k-2})\}|. \end{aligned} \tag{30}$$

It can be shown that the expectation of $N(w_{k-1,1}, l_{k-1}, \ldots, l_K, u_{k-2}^n, y_{k-2}^n)$ satisfies the fol-

lowing inequalities:

$$
\begin{aligned}
&\mathbb{E}[N(w_{k-1,1}, l_{k-1}, \ldots, l_K, u_{k-2}^n, y_{k-2}^n)] \\
&\geq 2^{n\left(\widetilde{R}_{k-1}-R_{k-1,2}+\sum_{i=k}^{K}(\widetilde{R}_i-R_{i,1}-R_{i,2})-I(U_K;Y_{k-2}|U_{k-2})-\delta_n(\epsilon)-\epsilon_n\right)} \\
&\quad + 2^{n\left(\sum_{i=k}^{K}(\widetilde{R}_i-R_{i,1}-R_{i,2})-I(U_K;Y_{k-2}|U_{k-1})-\delta_n(\epsilon)-\epsilon_n\right)} \\
&\quad + 2^{n\left(\sum_{i=k+1}^{K}(\widetilde{R}_i-R_{i,1}-R_{i,2})-I(U_K;Y_{k-2}|U_k)-\delta_n(\epsilon)-\epsilon_n\right)} \\
&\quad + \cdots\cdots + 2^{n\left(\widetilde{R}_K-R_{K,1}-R_{K,2}-I(U_K;Y_{k-2}|U_{K-1})-\delta_n(\epsilon)-\epsilon_n\right)} \\
&= 2^{n\left(\widetilde{R}_{k-1}-R_{k-1,2}+\sum_{i=k}^{K}(\widetilde{R}_i-R_{i,1}-R_{i,2})-I(U_K;Y_{k-2}|U_{k-2})-\delta_n(\epsilon)-\epsilon_n\right)} \\
&\quad \Bigg(1 + 2^{-n(\widetilde{R}_{k-1}-R_{k-1,2}-I(U_{k-1};Y_{k-2}|U_{k-2}))} + 2^{-n(\widetilde{R}_{k-1}-R_{k-1,2}+\widetilde{R}_k-R_{k,1}-R_{k,2}-I(U_k;Y_{k-2}|U_{k-2}))} \\
&\quad + 2^{-n(\widetilde{R}_{k-1}-R_{k-1,2}+\sum_{i=k}^{k+1}\widetilde{R}_i-R_{i,1}-R_{i,2}-I(U_{k+1};Y_{k-2}|U_{k-2}))} \\
&\quad + \cdots\cdots + 2^{-n(\widetilde{R}_{k-1}-R_{k-1,2}+\sum_{i=k}^{K-1}\widetilde{R}_i-R_{i,1}-R_{i,2}-I(U_{K-1};Y_{k-2}|U_{k-2}))}\Bigg)
\end{aligned} \tag{31}
$$

and,

$$
\begin{aligned}
&\mathbb{E}[N(w_{k-1,1}, l_{k-1}, \ldots, l_K, u_{k-2}^n, y_{k-2}^n)] \\
&\leq 2^{n\left(\widetilde{R}_{k-1}-R_{k-1,2}+\sum_{i=k}^{K}(\widetilde{R}_i-R_{i,1}-R_{i,2})-I(U_K;Y_{k-2}|U_{k-2})+\delta_n(\epsilon)-\epsilon_n\right)} \\
&\quad + 2^{n\left(\sum_{i=k}^{K}(\widetilde{R}_i-R_{i,1}-R_{i,2})-I(U_K;Y_{k-2}|U_{k-1})+\delta_n(\epsilon)-\epsilon_n\right)} \\
&\quad + 2^{n\left(\sum_{i=k+1}^{K}(\widetilde{R}_i-R_{i,1}-R_{i,2})-I(U_K;Y_{k-2}|U_k)+\delta_n(\epsilon)-\epsilon_n\right)} \\
&\quad + \cdots\cdots + 2^{n\left(\widetilde{R}_K-R_{K,1}-R_{K,2}-I(U_K;Y_{k-2}|U_{K-1})+\delta_n(\epsilon)-\epsilon_n\right)} \\
&= 2^{n\left(\widetilde{R}_{k-1}-R_{k-1,2}+\sum_{i=k}^{K}(\widetilde{R}_i-R_{i,1}-R_{i,2})-I(U_K;Y_{k-2}|U_{k-2})+\delta_n(\epsilon)-\epsilon_n\right)} \\
&\quad \Bigg(1 + 2^{-n(\widetilde{R}_{k-1}-R_{k-1,2}-I(U_{k-1};Y_{k-2}|U_{k-2}))} + 2^{-n(\widetilde{R}_{k-1}-R_{k-1,2}+\widetilde{R}_k-R_{k,1}-R_{k,2}-I(U_k;Y_{k-2}|U_{k-2}))} \\
&\quad + 2^{-n(\widetilde{R}_{k-1}-R_{k-1,2}+\sum_{i=k}^{k+1}\widetilde{R}_i-R_{i,1}-R_{i,2}-I(U_{k+1};Y_{k-2}|U_{k-2}))} \\
&\quad + \cdots\cdots + 2^{-n(\widetilde{R}_{k-1}-R_{k-1,2}+\sum_{i=k}^{K-1}\widetilde{R}_i-R_{i,1}-R_{i,2}-I(U_{K-1};Y_{k-2}|U_{k-2}))}\Bigg).
\end{aligned} \tag{32}
$$

Hence, if

$$
\widetilde{R}_{k-1} - R_{k-1,2} + \sum_{i=k}^{j}\widetilde{R}_i - R_{i,1} - R_{i,2} \geq I(U_j; Y_{k-2}|U_{k-2}), \text{ for } k-1 \leq j \leq K-1, \tag{33}
$$

then, $\epsilon'_n \to 0$ as $n \to \infty$, where ϵ'_n is defined as follows:

$$2^{n\epsilon'_n} = \Bigg(1 + 2^{-n(\widetilde{R}_{k-1} - R_{k-1,2} - I(U_{k-1};Y_{k-2}|U_{k-2}))} + 2^{-n(\widetilde{R}_{k-1} - R_{k-1,2} + \widetilde{R}_k - R_{k,1} - R_{k,2} - I(U_k;Y_{k-2}|U_{k-2}))}$$
$$+ 2^{-n(\widetilde{R}_{k-1} - R_{k-1,2} + \sum_{i=k}^{k+1} \widetilde{R}_i - R_{i,1} - R_{i,2} - I(U_{k+1};Y_{k-2}|U_{k-2}))}$$
$$+ \cdots\cdots + 2^{-n(\widetilde{R}_{k-1} - R_{k-1,2} + \sum_{i=k}^{K-1} \widetilde{R}_i - R_{i,1} - R_{i,2} - I(U_{K-1};Y_{k-2}|U_{k-2}))}\Bigg). \tag{34}$$

Hence, we have

$$\begin{aligned} &2^{n\left(\widetilde{R}_{k-1} - R_{k-1,2} + \sum_{i=k}^{K}(\widetilde{R}_i - R_{i,1} - R_{i,2}) - I(U_K;Y_{k-2}|U_{k-2}) - \delta_n(\epsilon) - \epsilon_n + \epsilon'_n\right)} \\ &\quad \le \mathbb{E}(N(w_{k-1,1}, l_{k-1}, \ldots, l_K, u_{k-2}^n, y_{k-2}^n)) \\ &\quad \le 2^{n\left(\widetilde{R}_{k-1} - R_{k-1,2} + \sum_{i=k}^{K}(\widetilde{R}_i - R_{i,1} - R_{i,2}) - I(U_K;Y_{k-2}|U_{k-2}) + \delta_n(\epsilon) - \epsilon_n + \epsilon'_n\right)}. \end{aligned} \tag{35}$$

Due to the fact that $N(w_{k-1,1}, l_{k-1}, \ldots, l_K, u_{k-2}^n, y_{k-2}^n)$ is a sum of Bernoulli distributed random variables, one can show that

$$\begin{aligned} &\mathrm{Var}(N(w_{k-1,1}, l_{k-1}, \ldots, l_K, u_{k-2}^n, y_{k-2}^n)) \\ &\le \mathbb{E}[N(w_{k-1,1}, l_{k-1}, \ldots, l_K, u_{k-2}^n, y_{k-2}^n)] \\ &\le 2^{n\left(\widetilde{R}_{k-1} - R_{k-1,2} + \sum_{i=k}^{K}(\widetilde{R}_i - R_{i,1} - R_{i,2}) - I(U_K;Y_{k-2}|U_{k-2}) + \delta_n(\epsilon) - \epsilon_n + \epsilon'_n\right)}. \end{aligned} \tag{36}$$

We next define the random event,

$$\begin{aligned} \varepsilon(w_{k-1,1}, l_{k-1}, \ldots, l_K, u_{k-2}^n, y_{k-2}^n) := \Bigg\{ &N(w_{k-1,1}, l_{k-1}, \ldots, l_K, u_{k-2}^n, y_{k-2}^n) \\ &\ge 2^{n\left(\widetilde{R}_{k-1} - R_{k-1,2} + \sum_{i=k}^{K}(\widetilde{R}_i - R_{i,1} - R_{i,2}) - I(U_K;Y_{k-2}|U_{k-2}) + \delta_n(\epsilon) - \frac{\epsilon_n}{2} + \epsilon'_n\right) + 1}\Bigg\}. \end{aligned} \tag{37}$$

Using Chebyshev's inequality, we can bound the probability of this random event as follows:

$$\begin{aligned} &P(\varepsilon(w_{k-1,1}, l_{k-1}, \ldots, l_K, u_{k-2}^n, y_{k-2}^n)) \\ &= P\Big(N(w_{k-1,1}, l_{k-1}, \ldots, l_K, u_{k-2}^n, y_{k-2}^n) \\ &\qquad \ge 2^{n\left(\widetilde{R}_{k-1} - R_{k-1,2} + \sum_{i=k}^{K}(\widetilde{R}_i - R_{i,1} - R_{i,2}) - I(U_K;Y_{k-2}|U_{k-2}) + \delta_n(\epsilon) - \frac{\epsilon_n}{2} + \epsilon'_n\right) + 1}\Big) \\ &\le P\Big(N(w_{k-1,1}, l_{k-1}, \ldots, l_K, u_{k-2}^n, y_{k-2}^n) \ge \mathbb{E}[N(w_{k-1,1}, l_{k-1}, \ldots, l_K, u_{k-2}^n, y_{k-2}^n)] \\ &\qquad + 2^{n\left(\widetilde{R}_{k-1} - R_{k-1,2} + \sum_{i=k}^{K}(\widetilde{R}_i - R_{i,1} - R_{i,2}) - I(U_K;Y_{k-2}|U_{k-2}) + \delta_n(\epsilon) - \frac{\epsilon_n}{2} + \epsilon'_n\right)}\Big) \\ &\le P\Big(|N(w_{k-1,1}, l_{k-1}, \ldots, l_K, u_{k-2}^n, y_{k-2}^n) - \mathbb{E}[N(w_{k-1,1}, l_{k-1}, \ldots, l_K, u_{k-2}^n, y_{k-2}^n)]| \\ &\qquad \ge 2^{n\left(\widetilde{R}_{k-1} - R_{k-1,2} + \sum_{i=k}^{K}(\widetilde{R}_i - R_{i,1} - R_{i,2}) - I(U_K;Y_{k-2}|U_{k-2}) + \delta_n(\epsilon) - \frac{\epsilon_n}{2} + \epsilon'_n\right)}\Big) \\ &\le \frac{\mathrm{Var}(N(w_{k-1,1}, l_{k-1}, \ldots, l_K, u_{k-2}^n, y_{k-2}^n))}{2^{2n\left(\widetilde{R}_{k-1} - R_{k-1,2} + \sum_{i=k}^{K}(\widetilde{R}_i - R_{i,1} - R_{i,2}) - I(U_K;Y_{k-2}|U_{k-2}) + \delta_n(\epsilon) - \frac{\epsilon_n}{2} + \epsilon'_n\right)}}, \end{aligned} \tag{38}$$

which converges to zero if

$$\widetilde{R}_{k-1} - R_{k-1,2} + \sum_{i=k}^{K}(\widetilde{R}_i - R_{i,1} - R_{i,2}) \geq I(U_K; Y_{k-1}|U_{k-2}). \tag{39}$$

For each $(w_{k-1,2} \ldots, w_{K,2})$, we define the following random variable:

$$\begin{aligned} N(w_{k-1,2} \ldots, w_{K,2}) := |\{(\widetilde{w}_{k-1,1}, \widetilde{l}_{k-1}, \ldots, \widetilde{l}_K) : (U_K^n(\widetilde{w}_{k-1,1}, w_{k-1,2}, \ldots, W_{K,2}, \widetilde{l}_{k-1}, \ldots, \widetilde{l}_K), \\ Y_{k-2}^n, U_{k-2}^n) \in T_\epsilon^{(n)}(U_{k-2}, U_K, Y_{k-2}), (\widetilde{w}_{k-1,1}, \widetilde{l}_{k-1}, \ldots, \widetilde{l}_K) \neq (w_{k-1,1}, l_{k-1}, \ldots, l_K)\}|, \end{aligned} \tag{40}$$

and define the following event:

$$\begin{aligned} \varepsilon(w_{k-1,2} \ldots, w_{K,2}) := \{N(w_{k-1,2}, \ldots, w_{K,2}) \\ \geq 2^{n\left(\widetilde{R}_{k-1} - R_{k-1,2} + \sum_{i=k}^{K}(\widetilde{R}_i - R_{i,1} - R_{i,2}) - I(U_K; Y_{k-2}|U_{k-2}) + \delta_n(\epsilon) - \frac{\epsilon_n}{2} + \epsilon_n'\right) + 1}\}. \end{aligned} \tag{41}$$

We further define the random variable $E(w_{k-1,2}, w_{k,1}, \ldots, w_{K,2}) := 0$ if $(U_K^n(w_{k-1,2}, w_{k,1}, \ldots, w_{K,2}, W_{k-1,1}, L_{k-1}, \ldots, L_K), Y_{k-2}^n, U_{k-2}^n) \in T_\epsilon^{(n)}(U_{k-2}, U_K, Y_{k-2})$ and the event $\varepsilon(w_{k-1,2} \ldots, w_{K,2})^c$ occurs; otherwise, $E(w_{k-1,2}, w_{k,1}, \ldots, w_{K,2}) := 1$. Therefore, we have

$$\begin{aligned} &P(E(w_{k-1,2}, w_{k,1}, \ldots, w_{K,2}) = 1) \\ &\leq P\big((U_K^n(w_{k-1,2}, w_{k,1}, \ldots, w_{K,2}, W_{k-1,1}, L_{k-1}, \ldots, L_K), Y_{k-2}^n, U_{k-2}^n) \notin T_\epsilon^{(n)}\big) \\ &\quad + P(\varepsilon(w_{k-1,2} \ldots, w_{K,2})). \end{aligned} \tag{42}$$

It is clear that the first term in (42) goes to zero as $n \to \infty$. For the second term in (42), we have

$$\begin{aligned} &P(\varepsilon(w_{k-1,2} \ldots, w_{K,2})) \\ &\leq \sum_{(u_{k-2}^n, y_{k-2}^n) \in T_\epsilon^{(n)}} P(u_{k-2}^n, y_{k-2}^n) P(\varepsilon(w_{k-1,2} \ldots, w_{K,2})|u_{k-2}^n, y_{k-2}^n) + P((U_{k-2}^n, Y_{k-2}^n) \notin T_\epsilon^{(n)}) \\ &= \sum_{(u_{k-2}^n, y_{k-2}^n) \in T_\epsilon^{(n)}} \sum_{w_{k-1,1}, l_{k-1}, \ldots, l_K} P(u_{k-2}^n, y_{k-2}^n) P(w_{k-1,1}, l_{k-1}, \ldots, l_K|u_{k-2}^n, y_{k-2}^n) \\ &\quad \cdot P(\varepsilon(w_{k-1,2} \ldots, w_{K,2})|u_{k-2}^n, y_{k-2}^n, w_{k-1,1}, l_{k-1}, \ldots, l_K) + P((U_{k-2}^n, Y_{k-2}^n) \notin T_\epsilon^{(n)}) \\ &\to 0, \text{ as } n \to \infty \text{ if } \widetilde{R}_{k-1} - R_{k-1,2} + \sum_{i=k}^{K} \widetilde{R}_i - R_{i,1} - R_{i,2} \geq I(U_K; Y_{k-2}|U_{k-2}). \end{aligned} \tag{43}$$

Hence, $P(E(w_{k-1,2}, w_{k,1}, \ldots, w_{K,2}) = 1) \to 0$ as $n \to \infty$, if

$$\widetilde{R}_{k-1} - R_{k-1,2} + \sum_{i=k}^{K} \widetilde{R}_i - R_{i,1} - R_{i,2} \geq I(U_K; Y_{k-2}|U_{k-2}). \tag{44}$$

Therefore, we derive the following bound:

$$
\begin{aligned}
&H(W_{k-1,1}, L_{k-1}, \ldots, L_K | Y_{k-2}^n, U_{k-2}^n, W_{k-1,2}, \ldots, W_{K,2}) \\
&\leq H(E(w_{k-1,2}, w_{k,1}, \ldots, w_{K,2}), W_{k-1,1}, L_{k-1}, \ldots, L_K | Y_{k-2}^n, U_{k-2}^n, W_{k-1,2}, \ldots, W_{K,2}) \\
&= H(W_{k-1,1}, L_{k-1}, \ldots, L_K | Y_{k-2}^n, U_{k-2}^n, W_{k-1,2}, \ldots, W_{K,2}, E(w_{k-1,2}, w_{k,1}, \ldots, w_{K,2})) \\
&\quad + H(E(w_{k-1,2}, w_{k,1}, \ldots, w_{K,2})) \\
&\leq P(E(w_{k-1,2}, w_{k,1}, \ldots, w_{K,2}) = 1) H\big(W_{k-1,1}, L_{k-1}, \ldots, L_K | Y_{k-2}^n, U_{k-2}^n, W_{k-1,2}, \ldots, W_{K,2}, \\
&\quad E(w_{k-1,2}, w_{k,1}, \ldots, w_{K,2}) = 1\big) + H\big(W_{k-1,1}, L_{k-1}, \ldots, L_K | Y_{k-2}^n, U_{k-2}^n, W_{k-1,2}, \ldots, W_{K,2}, \\
&\quad E(w_{k-1,2}, w_{k,1}, \ldots, w_{K,2}) = 0\big) + 1 \\
&\leq P(E(w_{k-1,2}, w_{k,1}, \ldots, w_{K,2}) = 1) H\big(W_{k-1,1}, L_{k-1}, \ldots, L_K | Y_{k-2}^n, U_{k-2}^n, W_{k-1,2}, \ldots, W_{K,2}, \\
&\quad E(w_{k-1,2}, w_{k,1}, \ldots, w_{K,2}) = 1\big) + n\Bigg(\widetilde{R}_{k-1} - R_{k-1,2} + \sum_{i=k}^{K} (\widetilde{R}_i - R_{i,1} - R_{i,2}) \\
&\quad - I(U_K; Y_{k-2} | U_{k-2}) + \delta_n(\epsilon) - \frac{\epsilon_n}{2} + \epsilon_n' \Bigg) + 1 + 1. \qquad (45)
\end{aligned}
$$

Combining (33), (44) and (45), we obtain

$$
\begin{aligned}
&\lim_{n\to\infty} \frac{1}{n} H(W_{k-1,1}, L_{k-1}, \ldots, L_K | Y_{k-2}^n, U_{k-2}^n, W_{k-1,2}, \ldots, W_{K,2}) \\
&\qquad \leq \widetilde{R}_{k-1} - R_{k-1,2} + \sum_{i=k}^{K} (\widetilde{R}_i - R_{i,1} - R_{i,2}) - I(U_K; Y_{k-2} | U_{k-2}), \qquad (46)
\end{aligned}
$$

which concludes the proof.

C Induction Algorithm for Fourier-Motzkin Elimination

As we have shown in Appendix A, we need to eliminate $R_{k,1}, R_{k,2}$ for $2 \leq k \leq K$ in the following region:

$$R_1 \leq I(U_1;Y_1), \tag{47a}$$

$$R_{k,1} + R_{k,2} \leq I(U_k;Y_k|U_{k-1}), \text{ for } 2 \leq k \leq K, \tag{47b}$$

$$R_{l-1,2} + \sum_{i=l}^{j}(R_{i,1}+R_{i,2}) \leq \sum_{i=l-1}^{j} I(U_i;Y_i|U_{i-1}) - I(U_j;Y_{l-2}|U_{l-2}), \tag{47c}$$

$$\text{for } 3 \leq l \leq K, l-1 \leq j \leq K,$$

$$R_2 = R_{2,1}, \tag{47d}$$

$$R_k = R_{k-1,2} + R_{k,1}, \text{ for } 3 \leq k \leq K-1, \tag{47e}$$

$$R_K = R_{K-1,2} + R_{K,1} + R_{K,2}, \tag{47f}$$

where the bounds (47a), (47b) and (47c) correspond to the achievable region after rate splitting, which are expressed in terms of component rates, and the bounds (47d), (47e) and (47f) are corresponding to the rate sharing strategy.

It can be seen that the total number of bounds in the above region is on the order of K^2 over which $2K-2$ variables need to be eliminated. Directly applying Fourier-Motzkin elimination is not analytically tractable. We design an inductive algorithm, in which we eliminate the rate pairs $(R_{k-1,2}, R_{k,1})$ for $3 \leq k \leq K-1$ one at each step, and finally eliminate $(R_{K-1,2}, R_{K,1}, R_{K,2})$. We first replace $R_{2,1}$ with R_2, $R_{k-1,2}+R_{k,1}$ with R_k for $3 \leq k \leq K-1$, and $R_{K-1,2}+R_{K,1}+R_{K,2}$ with R_K, and we obtain the following region:

$$\begin{aligned}
R_1 &\leq I(U_1;Y_1),\\
R_2 + R_{2,2} &\leq I(U_2;Y_2|U_1)\\
R_{k,1} + R_{k,2} &\leq I(U_k;Y_k|U_{k-1}), \text{ for } 3 \leq k \leq K,\\
\sum_{i=l}^{j} R_i + R_{j,2} &\leq \sum_{i=l-1}^{j} I(U_i;Y_i|U_{i-1}) - I(U_j;Y_{l-2}|U_{l-2}),\\
&\qquad \text{for } 3 \leq l \leq K, l-1 \leq j \leq K-1,\\
\sum_{i=l}^{K} R_i &\leq \sum_{i=l-1}^{K} I(U_i;Y_i|U_{i-1}) - I(U_K;Y_{l-2}|U_{l-2}),\\
&\qquad \text{for } 3 \leq l \leq K,\\
R_k &= R_{k-1,2} + R_{k,1}, \text{ for } 3 \leq k \leq K-1,\\
R_K &= R_{K-1,2} + R_{K,1} + R_{K,2}.
\end{aligned} \tag{48}$$

To start the elimination process, we first eliminate $(R_{2,2}, R_{3,1})$ from the inequalities given below, corresponding to the decoding and secrecy requirements of receiver 1 to receiver 3:

$$\begin{aligned}
R_1 &\leq I(U_1;Y_1),\\
R_2 + R_{2,2} &\leq I(U_2;Y_2|U_1),\\
R_{3,1} + R_{3,2} &\leq I(U_3;Y_3|U_2),\\
R_{2,2} &\leq I(U_2;Y_2|U_1) - I(U_2;Y_1|U_1),\\
R_3 + R_{3,2} &\leq \sum_{i=2}^{3} I(U_i;Y_i|U_{i-1}) - I(U_3;Y_1|U_1).\\
R_{3,2} &\leq I(U_3;Y_3|U_2) - I(U_3;Y_2|U_2),\\
R_3 &= R_{2,2} + R_{3,1}.
\end{aligned} \tag{49}$$

We then obtain the following inequalities after elimination:

$$\begin{aligned}
R_1 &\leq I(U_1;Y_1),\\
R_2 &\leq I(U_2;Y_2|U_1),\\
\sum_{i=2}^{3} R_i + R_{3,2} &\leq \sum_{i=2}^{3} I(U_i;Y_i|U_{i-1}),\\
R_3 + R_{3,2} &\leq \sum_{i=2}^{3} I(U_i;Y_i|U_{i-1}) - I(U_3;Y_1|U_1),\\
R_{3,2} &\leq I(U_3;Y_3|U_2) - I(U_3;Y_2|U_2),
\end{aligned} \tag{50}$$

which we denote as $\mathcal{R}_3$.

We then eliminate $(R_{3,2}, R_{4,1})$ from the inequalities in $\mathcal{R}_3$ and the inequalities given below, which together are corresponding to the decoding and secrecy requirements of receiver 1 to receiver 4:

$$\begin{aligned}
R_{4,1} + R_{4,2} &\leq I(U_4;Y_4|U_3),\\
\sum_{i=j}^{4} R_i + R_{4,2} &\leq \sum_{i=j-1}^{4} I(U_i;Y_i|U_{i-1}) - I(U_4;Y_{j-2}|U_{j-2}), \text{ for } 3 \leq j \leq 5\\
R_4 &= R_{3,2} + R_{4,1}.
\end{aligned} \tag{51}$$

We then obtain the following bounds after elimination:

$$\begin{aligned}
R_1 &\le I(U_1;Y_1),\\
\sum_{i=2}^{j} R_i &\le \sum_{i=2}^{j} I(U_i;Y_i|U_{i-1}),\ \text{for } 2\le j\le 3,\\
\sum_{i=2}^{4} R_i + R_{4,2} &\le \sum_{i=2}^{4} I(U_i;Y_i|U_{i-1}),\\
\sum_{i=l}^{j} R_i &\le \sum_{i=l-1}^{j} I(U_i;Y_i|U_{i-1}) - I(U_j;Y_{l-2}|U_{l-2}),\\
&\qquad \text{for } 3\le l\le j\le 3,\\
\sum_{i=l}^{4} R_i + R_{4,2} &\le \sum_{i=l-1}^{4} I(U_i;Y_i|U_{i-1}) - I(U_4;Y_{l-2}|U_{l-2}),\\
&\qquad \text{for } 3\le l\le 5,
\end{aligned} \tag{52}$$

which we denote as $\mathcal{R}_4$.

As we observe, the region $\mathcal{R}_3$ and $\mathcal{R}_4$ conform to the following structure for $k=3$ and $k=4$:

$$\begin{aligned}
R_1 &\le I(U_1;Y_1),\\
\sum_{i=2}^{j} R_i &\le \sum_{i=2}^{j} I(U_i;Y_i|U_{i-1}),\ \text{for } 2\le j\le k-1,\\
\sum_{i=2}^{k} R_i + R_{k,2} &\le \sum_{i=2}^{k} I(U_i;Y_i|U_{i-1}),\\
\sum_{i=l}^{j} R_i &\le \sum_{i=l-1}^{j} I(U_i;Y_i|U_{i-1}) - I(U_j;Y_{l-2}|U_{l-2}),\\
&\qquad \text{for } 3\le l\le j\le k-1,\\
\sum_{i=l}^{k} R_i + R_{k,2} &\le \sum_{i=l-1}^{k} I(U_i;Y_i|U_{i-1}) - I(U_k;Y_{l-2}|U_{l-2}),\\
&\qquad \text{for } 3\le l\le k+1.
\end{aligned} \tag{53}$$

We next show that the region $\mathcal{R}_k$ takes the structure (53) for any $3\le k\le K-1$ using induction. We have verified such a claim for $k=3,4$. If such a claim holds for $\mathcal{R}_k$, we eliminate $R_{k,2}$ and $R_{k+1,1}$ from the inequalities in $\mathcal{R}_k$ and the inequalities given below, which

together are corresponding to the decoding and secrecy requirements of receiver 1 to receiver $k+1$:

$$\begin{aligned}
&R_{k+1,1} + R_{k+1,2} \leq I(U_{k+1}; Y_{k+1}|U_k), \\
&\sum_{i=j}^{k+1} R_i + R_{k+1,2} \leq \sum_{i=j-1}^{k+1} I(U_i; Y_i|U_{i-1}) - I(U_{k+1}; Y_{j-2}|U_{j-2}), \text{ for } 3 \leq j \leq k+2 \\
&R_{k+1} = R_{k,2} + R_{k+1,1}. \qquad (54)
\end{aligned}$$

Then the resulting region, following standard steps of Fourier-Motzkin elimination to eliminate $R_{k,2}$ and $R_{k+1,1}$, equals (53) for $k+1$.

Finally, we eliminate $(R_{K-1,2}, R_{K,1}, R_{K,2})$, and obtain the achievable region in Theorem 1.

D Converse Proof of Theorem 1

By Fano's inequality and the secrecy requirements, we have the following inequalities:

$$H(W_k|Y_k^n) \leq n\epsilon_n, \quad \text{for } 1 \leq k \leq K, \qquad (55)$$

$$I(W_k, \ldots, W_K; Y_{k-2}^n) \leq n\epsilon_n, \quad \text{for } 3 \leq k \leq K, \qquad (56)$$

both of which implies that

$$I(W_k, \ldots, W_K; Y_{k-2}^n|W_1, \ldots, W_{k-2}) \leq n\epsilon_n, \quad \text{for } 3 \leq k \leq K. \qquad (57)$$

We denote $Y_k^{i-1} := (Y_{k,1}, \ldots, Y_{k,i-1})$, and $Y_{k,i+1}^n := (Y_{k,i+1}, \ldots, Y_{k,n})$. We set $U_{1,i} := (W_1, Y_1^{i-1})$, $U_{2,i} := (W_1, W_2, Y_2^{i-1})$, $U_{k,i} := (W_1, \ldots, W_k, Y_k^{i-1}, Y_{k-2,i+1}^n)$, for $3 \leq k \leq K$. And we note that $Y_0^n = Y_{-1}^n = \Phi$. Due to the degradedness condition, it can be verified that $(U_{1,i}, U_{2,i}, \ldots, U_{K-1,i}, U_{K,i}, X_i)$ satisfy the following Markov chain condition:

$$U_{1,i} \to U_{2,i} \to \ldots \to U_{K,i} \to X_i \to Y_{K,i} \to \ldots \to Y_{1,i}, \text{ for } 1 \leq i \leq n. \qquad (58)$$

We first bound the rate R_1. Since W_1 is only required to be decoded by receiver Y_1, we obtain the following bound:

$$\begin{aligned}
nR_1 &= H(W_1) = I(W_1; Y_1^n) + H(W_1|Y_1^n) \\
&\overset{(a)}{\leq} I(W_1; Y_1^n) + n\epsilon_n = \sum_{i=1}^{n} I(W_1; Y_{1i}|Y_1^{i-1}) + n\epsilon_n \\
&\leq \sum_{i=1}^{n} I(W_1, Y_1^{i-1}; Y_{1i}) + n\epsilon_n = \sum_{i=1}^{n} I(U_{1,i}; Y_{1,i}) + n\epsilon_n, \qquad (59)
\end{aligned}$$

where (a) is due to Fano's inequality.

We further bound the rate R_2 as follows:

$$\begin{aligned} nR_2 = H(W_2) = H(W_2|W_1) &= I(W_2; Y_2^n|W_1) + H(W_2|Y_2^n, W_1) \\ &\overset{(a)}{\leq} I(W_2; Y_2^n|W_1) + n\epsilon_n \\ &= \sum_{i=1}^{n} I(W_2; Y_{2,i}|W_1, Y_2^{i-1}) + n\epsilon_n \\ &\overset{(b)}{\leq} \sum_{i=1}^{n} I(W_1, W_2, Y_2^{i-1}; Y_{2,i}|W_1, Y_1^{i-1}) + n\epsilon_n \\ &= \sum_{i=1}^{n} I(U_{2,i}; Y_{2,i}|U_{1,i}) + n\epsilon_n, \end{aligned} \tag{60}$$

where (a) is due to Fano's inequality, and (b) is due to the Markov chain condition $Y_1^{i-1} \rightarrow Y_2^{i-1} \rightarrow (W_1, W_2, Y_{2,i})$.

We then bound the sum rate bounds on $\sum_{i=2}^{k} R_i$, for $3 \leq k \leq K$:

$$
\begin{aligned}
n\sum_{j=2}^{k} R_j &= H(W_2, \ldots, W_k) \\
&\overset{(a)}{=} H(W_2|W_1) + H(W_3|W_1, W_2) + \ldots + H(W_k|W_1, \ldots, W_{k-1}) \\
&\overset{(b)}{\leq} I(W_2; Y_2^n|W_1) + I(W_3; Y_3^n|W_1, W_2) + \ldots + I(W_k; Y_k^n|W_1, \ldots, W_{k-1}) + n(k-1)\epsilon_n \\
&= \sum_{i=1}^{n} I(W_2; Y_{2,i}|W_1, Y_2^{i-1}) + I(W_3; Y_{3,i}|W_1, W_2, Y_3^{i-1}) \\
&\qquad + \ldots + I(W_k; Y_{k,i}|W_1, \ldots, W_{k-1}, Y_k^{i-1}) + n(k-1)\epsilon_n \\
&= n(k-1)\epsilon_n + \sum_{i=1}^{n} \Bigg(I(W_2, Y_2^{i-1}; Y_{2,i}|W_1, Y_1^{i-1}) - I(Y_2^{i-1}; Y_{2,i}|W_1, Y_1^{i-1}) \\
&\quad + I(W_3, Y_3^{i-1}, Y_{1,i+1}^n; Y_{3,i}|W_1, W_2, Y_2^{i-1}) - I(Y_3^{i-1}; Y_{3,i}|W_1, W_2, Y_2^{i-1}) \\
&\quad - I(Y_{1,i+1}^n; Y_{3,i}|W_1, W_2, W_3, Y_3^{i-1}) \\
&\quad + \sum_{j=4}^{k} \Bigg(I(W_j, Y_j^{i-1}, Y_{j-2,i+1}^n; Y_{j,i}|W_1, \ldots, W_{j-1}, Y_{j-3,i+1}^n, Y_{j-1}^{i-1}) \\
&\quad + I(Y_{j-3,i+1}^n; Y_{j,i}|W_1, \ldots, W_{j-1}, Y_j^{i-1}) \\
&\quad - I(Y_j^{i-1}; Y_{j,i}|W_1, \ldots, W_{j-1}, Y_{j-3,i+1}^n, Y_{j-1}^{i-1}) - I(Y_{j-2,i+1}^n; Y_{j,i}|W_1, \ldots, W_j, Y_j^{i-1}) \Bigg) \Bigg) \\
&\overset{(c)}{\leq} n(k-1)\epsilon_n + \sum_{j=2}^{k}\sum_{i=1}^{n} I(U_{j,i}; Y_{j,i}|U_{j-1,i}),
\end{aligned} \tag{61}
$$

where (a) is due to the independence between the messages $(W_1, \ldots, W_k)$, (b) is due to Fano's inequality, and (c) is due to the facts that $-I(Y_2^{i-1}; Y_{2,i}|W_1, Y_1^{i-1}) \leq 0$, $-I(Y_3^{i-1}; Y_{3,i}|W_1, W_2, Y_2^{i-1}) \leq 0$, $-I(Y_{k-2,i+1}^n; Y_{k,i}|W_1, \ldots, W_k, Y_k^{i-1}) \leq 0$ and the following inequalities:

$$
\begin{aligned}
&-I(Y_{j-2,i+1}^n; Y_{j,i}|W_1, \ldots, W_j, Y_j^{i-1}) + I(Y_{j-2,i+1}^n; Y_{j+1,i}|W_1, \ldots, W_j, Y_{j+1}^{i-1}) \\
&-I(Y_{j+1}^{i-1}; Y_{j+1,i}|W_1, \ldots, W_j, Y_{j-2,i+1}^n, Y_j^{i-1}) \\
&\quad \overset{(a)}{=} -I(Y_j^{i-1}; Y_{j-2,i}|W_1, \ldots, W_j, Y_{j-2,i+1}^n) + I(Y_{j+1}^{i-1}; Y_{j-2,i}|W_1, \ldots, W_j, Y_{j-2,i+1}^n) \\
&\qquad - I(Y_{j+1}^{i-1}; Y_{j+1,i}|W_1, \ldots, W_j, Y_{j-2,i+1}^n, Y_j^{i-1}) \\
&\quad \overset{(b)}{=} I(Y_{j+1}^{i-1}; Y_{j-2,i}|W_1, \ldots, W_j, Y_{j-2,i+1}^n, Y_j^{i-1}) - I(Y_{j+1}^{i-1}; Y_{j+1,i}|W_1, \ldots, W_j, Y_{j-2,i+1}^n, Y_j^{i-1}) \\
&\quad \overset{(c)}{=} -I(Y_{j+1}^{i-1}; Y_{j+1,i}|W_1, \ldots, W_j, Y_{j-2,i+1}^n, Y_j^{i-1}, Y_{j-2,i}) \\
&\quad \leq 0,
\end{aligned} \tag{62}
$$

where (a) is due to Csiszár's sum identity property [9], and (b) and (c) are due to the degradedness condition (1).

We next bound the sum rate bounds on $\sum_{j=l}^{k} R_j$, for $3 \leq l \leq k \leq K$, which correspond to the secrecy constraints:

$$\begin{aligned} n\sum_{j=l}^{k} R_j &= H(W_l, \ldots, W_k) + H(W_{l-1}) - H(W_{l-1}) \\ &\overset{(a)}{\leq} \sum_{j=l-1}^{k} H(W_j) - H(W_{l-1}) + n\epsilon_n - I(W_l \ldots, W_k; Y_{l-2}^n | W_1, \ldots, W_{l-2}) \\ &\overset{(b)}{\leq} \sum_{j=l-1}^{k} H(W_j) + n\epsilon_n - I(W_{l-1} \ldots, W_k; Y_{l-2}^n | W_1, \ldots, W_{l-2}) \end{aligned} \tag{63}$$

where (a) is due to the secrecy requirement (57) and the independence of the messages, and (b) is due to the following fact:

$$\begin{aligned} &-H(W_{l-1}) - I(W_l \ldots, W_k; Y_{l-2}^n | W_1, \ldots, W_{l-2}) \\ &\quad = -H(W_{l-1}) - H(W_l \ldots, W_k | W_1, \ldots, W_{l-2}) + H(W_l \ldots, W_k | Y_{l-2}^n, W_1, \ldots, W_{l-2}) \\ &\quad \overset{(a)}{=} -H(W_{l-1} \ldots, W_k | W_1, \ldots, W_{l-2}) + H(W_l \ldots, W_k | Y_{l-2}^n, W_1, \ldots, W_{l-2}) \\ &\quad \leq -H(W_{l-1} \ldots, W_k | W_1, \ldots, W_{l-2}) + H(W_{l-1}, W_l \ldots, W_k | Y_{l-2}^n, W_1, \ldots, W_{l-2}) \\ &\quad = -I(W_{l-1} \ldots, W_k; Y_{l-2}^n | W_1, \ldots, W_{l-2}), \end{aligned} \tag{64}$$

where (a) is due to the independence of the messages.

We next bound each term in (63) one by one. We first bound $H(W_j)$ for $l \leq j \leq k$ as

follows:

$$
\begin{aligned}
H(W_j) &\overset{(a)}{\leq} H(W_j|W_1,\ldots,W_{j-1}) + n\epsilon_n - H(W_j|Y_j^n, W_1,\ldots,W_{j-1}) \\
&= I(W_j; Y_j^n|W_1,\ldots,W_{j-1}) + n\epsilon_n \\
&= n\epsilon_n + \sum_{i=1}^{n} I(W_j; Y_{j,i}|W_1,\ldots,W_{j-1}, Y_j^{i-1}) \\
&= n\epsilon_n + \sum_{i=1}^{n} \Big(I(W_j, Y_j^{i-1}, Y_{j-2,i+1}^n; Y_{j,i}|W_1,\ldots,W_{j-1}, Y_{j-1}^{i-1}, Y_{j-3,i+1}^n) \\
&\quad - I(Y_j^{i-1}; Y_{j,i}|W_1,\ldots,W_{j-1}, Y_{j-1}^{i-1}, Y_{j-3,i+1}^n) + I(Y_{j-3,i+1}^n; Y_{j,i}|W_1,\ldots,W_{j-1}, Y_j^{i-1}) \\
&\quad - I(Y_{j-2,i+1}^n; Y_{j,i}|W_1,\ldots,W_j, Y_j^{i-1}) \Big) \\
&= n\epsilon_n + \sum_{i=1}^{n} \Big(I(U_{j,i}; Y_{j,i}|U_{j-1,i}) - I(Y_j^{i-1}; Y_{j,i}|W_1,\ldots,W_{j-1}, Y_{j-1}^{i-1}, Y_{j-3,i+1}^n) \\
&\quad + I(Y_{j-3,i+1}^n; Y_{j,i}|W_1,\ldots,W_{j-1}, Y_j^{i-1}) - I(Y_{j-2,i+1}^n; Y_{j,i}|W_1,\ldots,W_j, Y_j^{i-1}) \Big) \\
&\overset{(b)}{=} n\epsilon_n + \sum_{i=1}^{n} \Big(I(U_{j,i}; Y_{j,i}|U_{j-1,i}) - I(Y_j^{i-1}; Y_{j,i}|W_1,\ldots,W_{j-1}, Y_{j-1}^{i-1}, Y_{j-3,i+1}^n) \\
&\quad + I(Y_j^{i-1}; Y_{j-3,i}|W_1,\ldots,W_{j-1}, Y_{j-3,i+1}^n) - I(Y_{j-2,i+1}^n; Y_{j,i}|W_1,\ldots,W_j, Y_j^{i-1}) \Big) \\
&\overset{(c)}{=} n\epsilon_n + \sum_{i=1}^{n} \Big(I(U_{j,i}; Y_{j,i}|U_{j-1,i}) - I(Y_j^{i-1}; Y_{j,i}|W_1,\ldots,W_{j-1}, Y_{j-1}^{i-1}, Y_{j-3,i+1}^n) \\
&\quad + I(Y_j^{i-1}; Y_{j-3,i}|W_1,\ldots,W_{j-1}, Y_{j-3,i+1}^n, Y_{j-1}^{i-1}) + I(Y_{j-1}^{i-1}; Y_{j-3,i}|W_1,\ldots,W_{j-1}, Y_{j-3,i+1}^n) \\
&\quad - I(Y_{j-2,i+1}^n; Y_{j,i}|W_1,\ldots,W_j, Y_j^{i-1}) \Big) \\
&\overset{(d)}{\leq} n\epsilon_n + \sum_{i=1}^{n} \Big(I(U_{j,i}; Y_{j,i}|U_{j-1,i}) + I(Y_{j-1}^{i-1}; Y_{j-3,i}|W_1,\ldots,W_{j-1}, Y_{j-3,i+1}^n) \\
&\quad - I(Y_{j-2,i+1}^n; Y_{j,i}|W_1,\ldots,W_j, Y_j^{i-1}) \Big) \\
&\overset{(e)}{=} n\epsilon_n + \sum_{i=1}^{n} \Big(I(U_{j,i}; Y_{j,i}|U_{j-1,i}) + I(Y_{j-3,i+1}^n; Y_{j-1,i}|W_1,\ldots,W_{j-1}, Y_{j-1}^{i-1}) \\
&\quad - I(Y_{j-2,i+1}^n; Y_{j,i}|W_1,\ldots,W_j, Y_j^{i-1}) \Big) \qquad (65)
\end{aligned}
$$

where (a) is due to the independence of the messages and the Fano's inequality (57), (b) is due to Csiszár sum identity property, (c) is due to the degradedness condition (1) and the

fact that

$$I(Y_j^{i-1};Y_{j-3,i}|W_1,\ldots,W_{j-1},Y_{j-3,i+1}^n)$$
$$= I(Y_j^{i-1};Y_{j-3,i}|W_1,\ldots,W_{j-1},Y_{j-3,i+1}^n,Y_{j-1}^{i-1}) + I(Y_{j-1}^{i-1};Y_{j-3,i}|W_1,\ldots,W_{j-1},Y_{j-3,i+1}^n), \tag{66}$$

the inequality (d) is due to the degradedness condition (1) and the fact that

$$\begin{aligned}&-I(Y_j^{i-1};Y_{j,i}|W_1,\ldots,W_{j-1},Y_{j-1}^{i-1},Y_{j-3,i+1}^n) + I(Y_j^{i-1};Y_{j-3,i}|W_1,\ldots,W_{j-1},Y_{j-3,i+1}^n,Y_{j-1}^{i-1})\\ &\quad= -I(Y_j^{i-1};Y_{j,i}|W_1,\ldots,W_{j-1},Y_{j-1}^{i-1},Y_{j-3,i+1}^n,Y_{j-3,i})\\ &\quad\le 0,\end{aligned} \tag{67}$$

and (e) is due to Csiszár's sum identity property.

Following the intermediate step in (65), $H(W_j)$ is also upper bounded as follows:

$$\begin{aligned}H(W_j)\le n\epsilon_n + \sum_{i=1}^{n}\Big(&I(U_{j,i};Y_{j,i}|U_{j-1,i}) - I(Y_j^{i-1};Y_{j,i}|W_1,\ldots,W_{j-1},Y_{j-1}^{i-1},Y_{j-3,i+1}^n)\\ &+ I(Y_{j-3,i+1}^n;Y_{j,i}|W_1,\ldots,W_{j-1},Y_j^{i-1}) - I(Y_{j-2,i+1}^n;Y_{j,i}|W_1,\ldots,W_j,Y_j^{i-1})\Big).\end{aligned} \tag{68}$$

Hence, substituting (65) for $l\le j\le k$, and (68) for $j=l-1$ into the first term in (63), we obtain,

$$\begin{aligned}&\sum_{j=l-1}^{k} H(W_j)\\ &\quad\le n(k-l+2)\epsilon_n + \sum_{i=1}^{n}\sum_{j=l-1}^{k} I(U_{j,i};Y_{j,i}|U_{j-1,i})\\ &\qquad+ I(Y_{l-4,i+1}^n;Y_{l-1,i}|W_1,\ldots,W_{l-2},Y_{l-1}^{i-1}) - I(Y_{l-1}^{i-1};Y_{l-1,i}|W_1,\ldots,W_{l-2},Y_{l-2}^{i-1},Y_{l-4,i+1}^n)\\ &\qquad- I(Y_{k-2,i+1}^n;Y_{k,i}|W_1,\ldots,W_k,Y_k^{i-1}).\end{aligned} \tag{69}$$

We then bound the third term in (63) for $3 \leq l \leq k \leq K$ as follows:

$$
\begin{aligned}
&-I(W_{l-1}\ldots,W_k;Y_{l-2}^n|W_1,\ldots,W_{l-2}) \\
&\quad=\sum_{i=1}^{n} -I(W_{l-1}\ldots,W_k;Y_{l-2,i}|W_1,\ldots,W_{l-2},Y_{l-2,i+1}^n) \\
&\quad=\sum_{i=1}^{n} -I(W_{l-1}\ldots,W_k,Y_k^{i-1};Y_{l-2,i}|W_1,\ldots,W_{l-2},Y_{l-2,i+1}^n) \\
&\qquad+I(Y_k^{i-1};Y_{l-2,i}|W_1,\ldots,W_k,Y_{l-2,i+1}^n) \\
&\quad=\sum_{i=1}^{n} -I(W_{l-1}\ldots,W_k,Y_k^{i-1},Y_{l-2,i+1}^n;Y_{l-2,i}|W_1,\ldots,W_{l-2},Y_{l-4,i+1}^n) \\
&\qquad+I(Y_{l-2,i+1}^n;Y_{l-2,i}|W_1,\ldots,W_{l-2},Y_{l-4,i+1}^n)+I(Y_k^{i-1};Y_{l-2,i}|W_1,\ldots,W_k,Y_{l-2,i+1}^n) \\
&\quad=\sum_{i=1}^{n} -I(W_{l-1}\ldots,W_k,Y_k^{i-1},Y_{k-2,i+1}^n;Y_{l-2,i}|W_1,\ldots,W_{l-2},Y_{l-2}^{i-1},Y_{l-4,i+1}^n) \\
&\qquad+I(Y_{k-2,i+1}^n;Y_{l-2,i}|W_1,\ldots,W_k,Y_k^{i-1},Y_{l-2,i+1}^n)-I(Y_{l-2}^{i-1};Y_{l-2,i}|W_1,\ldots,W_{l-2},Y_{l-4,i+1}^n) \\
&\qquad+I(Y_{l-2,i+1}^n;Y_{l-2,i}|W_1,\ldots,W_{l-2},Y_{l-4,i+1}^n)+I(Y_k^{i-1};Y_{l-2,i}|W_1,\ldots,W_k,Y_{l-2,i+1}^n) \\
&\quad=\sum_{i=1}^{n} -I(U_{k,i};Y_{l-2,i}|U_{l-2,i}) \\
&\qquad+I(Y_{k-2,i+1}^n;Y_{l-2,i}|W_1,\ldots,W_k,Y_k^{i-1},Y_{l-2,i+1}^n)-I(Y_{l-2}^{i-1};Y_{l-2,i}|W_1,\ldots,W_{l-2},Y_{l-4,i+1}^n) \\
&\qquad+I(Y_{l-2,i+1}^n;Y_{l-2,i}|W_1,\ldots,W_{l-2},Y_{l-4,i+1}^n)+I(Y_k^{i-1};Y_{l-2,i}|W_1,\ldots,W_k,Y_{l-2,i+1}^n) \\
&\quad\overset{(a)}{=}\sum_{i=1}^{n} -I(U_{k,i};Y_{l-2,i}|U_{l-2,i}) \\
&\qquad-I(Y_{l-4,i+1}^n;Y_{l-2,i}|W_1,\ldots,W_{l-2},Y_{l-2}^{i-1})+I(Y_k^{i-1},Y_{k-2,i+1}^n;Y_{l-2,i}|W_1,\ldots,W_k,Y_{l-2,i+1}^n),
\end{aligned}
\tag{70}
$$

where (a) is due to the following fact:

$$
\begin{aligned}
&\sum_{i=1}^{n} -I(Y_{l-2}^{i-1};Y_{l-2,i}|W_1,\ldots,W_{l-2},Y_{l-4,i+1}^n)+I(Y_{l-2,i+1}^n;Y_{l-2,i}|W_1,\ldots,W_{l-2},Y_{l-4,i+1}^n) \\
&\quad=\sum_{i=1}^{n} -H(Y_{l-2,i}|W_1,\ldots,W_{l-2},Y_{l-4,i+1}^n)+H(Y_{l-2,i}|W_1,\ldots,W_{l-2},Y_{l-2}^{i-1},Y_{l-4,i+1}^n) \\
&\qquad+H(Y_{l-2,i}|W_1,\ldots,W_{l-2},Y_{l-4,i+1}^n)-H(Y_{l-2,i}|W_1,\ldots,W_{l-2},Y_{l-2,i+1}^n) \\
&\quad=\sum_{i=1}^{n} H(Y_{l-2,i}|W_1,\ldots,W_{l-2},Y_{l-2}^{i-1},Y_{l-4,i+1}^n)-H(Y_{l-2,i}|W_1,\ldots,W_{l-2},Y_{l-2}^{i-1}) \\
&\quad=\sum_{i=1}^{n} -I(Y_{l-4,i+1}^n;Y_{l-2,i}|W_1,\ldots,W_{l-2},Y_{l-2}^{i-1}).
\end{aligned}
\tag{71}
$$

Substituting (69) and (70) into (63), we obtain

$$
\begin{aligned}
&n\sum_{j=l}^{k}R_j \\
&\quad\le \sum_{j=l-1}^{k} H(W_j)+n\epsilon_n - I(W_{l-1}\ldots,W_k;Y_{l-2}^n|W_1,\ldots,W_{l-2}) \\
&\quad\le n(k-l+3)\epsilon_n+\sum_{i=1}^{n}\Bigg(\Bigg(\sum_{j=l-1}^{k} I(U_{j,i};Y_{j,i}|U_{j-1,i})\Bigg) - I(U_{k,i};Y_{l-2,i}|U_{l-2,i}) \\
&\qquad + I(Y_{l-4,i+1}^n;Y_{l-1,i}|W_1,\ldots,W_{l-2},Y_{l-1}^{i-1}) - I(Y_{l-1}^{i-1};Y_{l-1,i}|W_1,\ldots,W_{l-2},Y_{l-2}^{i-1},Y_{l-4,i+1}^n) \\
&\qquad - I(Y_{k-2,i+1}^n;Y_{k,i}|W_1,\ldots,W_k,Y_k^{i-1}) \\
&\qquad - I(Y_{l-4,i+1}^n;Y_{l-2,i}|W_1,\ldots,W_{l-2},Y_{l-2}^{i-1}) + I(Y_k^{i-1},Y_{k-2,i+1}^n;Y_{l-2,i}|W_1,\ldots,W_k,Y_{l-2,i+1}^n)\Bigg), \\
&\quad\overset{(a)}{\le} n(k-l+3)\epsilon_n+\sum_{i=1}^{n}\Bigg(\Bigg(\sum_{j=l-1}^{k} I(U_{j,i};Y_{j,i}|U_{j-1,i})\Bigg) - I(U_{k,i};Y_{l-2,i}|U_{l-2,i})\Bigg), \qquad (72)
\end{aligned}
$$

where (a) is due to the following two facts. The first fact is shown as follows:

$$
\begin{aligned}
&\sum_{i=1}^{n} I(Y_k^{i-1},Y_{k-2,i+1}^n;Y_{l-2,i}|W_1,\ldots,W_k,Y_{l-2,i+1}^n) - I(Y_{k-2,i+1}^n;Y_{k,i}|W_1,\ldots,W_k,Y_k^{i-1}) \\
&\quad= \sum_{i=1}^{n} H(Y_{l-2,i}|W_1,\ldots,W_k,Y_{l-2,i+1}^n) - H(Y_{l-2,i}|W_1,\ldots,W_k,Y_k^{i-1},Y_{k-2,i+1}^n) \\
&\qquad - H(Y_{k,i}|W_1,\ldots,W_k,Y_k^{i-1}) + H(Y_{k,i}|W_1,\ldots,W_k,Y_k^{i-1},Y_{k-2,i+1}^n) \\
&\quad= H(Y_{l-2}^n|W_1,\ldots,W_k) - H(Y_k^n|W_1,\ldots,W_k) \\
&\qquad + \sum_{i=1}^{n} H(Y_{k,i}|W_1,\ldots,W_k,Y_k^{i-1},Y_{k-2,i+1}^n,Y_{l-2,i}) \\
&\quad= -H(Y_k^n|W_1,\ldots,W_k,Y_{l-2}^n) + \sum_{i=1}^{n} H(Y_{k,i}|W_1,\ldots,W_k,Y_k^{i-1},Y_{k-2,i+1}^n,Y_{l-2,i}) \\
&\quad= \sum_{i=1}^{n} -H(Y_{k,i}|W_1,\ldots,W_k,Y_{l-2}^n,Y_k^{i-1}) + H(Y_{k,i}|W_1,\ldots,W_k,Y_k^{i-1},Y_{k-2,i+1}^n,Y_{l-2,i}) \\
&\quad\le 0. \qquad (73)
\end{aligned}
$$

The second fact is shown as follows:

$$\begin{aligned}
&\sum_{i=1}^{n} I(Y_{l-4,i+1}^{n}; Y_{l-1,i}|W_1,\ldots,W_{l-2},Y_{l-1}^{i-1}) - I(Y_{l-1}^{i-1}; Y_{l-1,i}|W_1,\ldots,W_{l-2},Y_{l-2}^{i-1},Y_{l-4,i+1}^{n}) \\
&\quad - I(Y_{l-4,i+1}^{n}; Y_{l-2,i}|W_1,\ldots,W_{l-2},Y_{l-2}^{i-1}) \\
&\quad = \sum_{i=1}^{n} H(Y_{l-1,i}|W_1,\ldots,W_{l-2},Y_{l-1}^{i-1}) - H(Y_{l-1,i}|W_1,\ldots,W_{l-2},Y_{l-1}^{i-1},Y_{l-4,i+1}^{n}) \\
&\quad - H(Y_{l-1,i}|W_1,\ldots,W_{l-2},Y_{l-2}^{i-1},Y_{l-4,i+1}^{n}) + H(Y_{l-1,i}|W_1,\ldots,W_{l-2},Y_{l-4,i+1}^{n},Y_{l-1}^{i-1}) \\
&\quad - H(Y_{l-2,i}|W_1,\ldots,W_{l-2},Y_{l-2}^{i-1}) + H(Y_{l-2,i}|W_1,\ldots,W_{l-2},Y_{l-2}^{i-1},Y_{l-4,i+1}^{n}) \\
&\quad = H(Y_{l-1}^{n}|W_1,\ldots,W_{l-2}) - H(Y_{l-2}^{n}|W_1,\ldots,W_{l-2}) \\
&\qquad + \sum_{i=1}^{n} -H(Y_{l-1,i}|W_1,\ldots,W_{l-2},Y_{l-4,i+1}^{n},Y_{l-2}^{i-1},Y_{l-2,i}) \\
&\quad = H(Y_{l-1}^{n}|W_1,\ldots,W_{l-2},Y_{l-2}^{n}) - \sum_{i=1}^{n} H(Y_{l-1,i}|W_1,\ldots,W_{l-2},Y_{l-4,i+1}^{n},Y_{l-2}^{i-1},Y_{l-2,i}) \\
&\quad = \sum_{i=1}^{n} H(Y_{l-1,i}|W_1,\ldots,W_{l-2},Y_{l-2}^{n},Y_{l-1}^{i-1}) - H(Y_{l-1,i}|W_1,\ldots,W_{l-2},Y_{l-4,i+1}^{n},Y_{l-2}^{i-1},Y_{l-2,i}) \\
&\quad \leq 0. \qquad (74)
\end{aligned}$$

Furthermore, based on (61), we bound $\sum_{j=2}^{K} R_j$ as follows:

$$\begin{aligned}
n\sum_{j=2}^{K} R_j &\leq n(k-1)\epsilon_n + \sum_{j=2}^{K}\sum_{i=1}^{n} I(U_{j,i}; Y_{j,i}|U_{j-1,i}) \\
&\leq n(k-1)\epsilon_n + \sum_{j=2}^{K-1}\sum_{i=1}^{n} I(U_{j,i}; Y_{j,i}|U_{j-1,i}) + \sum_{i=1}^{n} I(X_i; Y_{K,i}|U_{K-1,i}). \qquad (75)
\end{aligned}$$

And based on (72), we bound $\sum_{j=l}^{K} R_j$ as follows:

$$\begin{aligned} n\sum_{j=l}^{K} R_j &\le n(K-l+3)\epsilon_n + \sum_{i=1}^{n}\left(\left(\sum_{j=l-1}^{K} I(U_{j,i};Y_{j,i}|U_{j-1,i})\right) - I(U_{K,i};Y_{l-2,i}|U_{l-2,i})\right) \\ &= n(K-l+3)\epsilon_n + \sum_{i=1}^{n}\left(\left(\sum_{j=l-1}^{K-1} I(U_{j,i};Y_{j,i}|U_{j-1,i})\right)\right. \\ &\quad \left. + I(U_{K,i};Y_{K,i}|U_{K-1,i}) - I(U_{K,i};Y_{l-2,i}|U_{l-2,i})\right) \\ &\overset{(a)}{\le} n(K-l+3)\epsilon_n + \sum_{i=1}^{n}\left(\left(\sum_{j=l-1}^{K-1} I(U_{j,i};Y_{j,i}|U_{j-1,i})\right)\right. \\ &\quad \left. + I(X_i;Y_{K,i}|U_{K-1,i}) - I(X_i;Y_{l-2,i}|U_{l-2,i})\right), \end{aligned} \tag{76}$$

where (a) is due to the Markov chain condition (58).

The proof of the converse is then completed by defining a uniformly distributed random variable $Q \in \{1,\ldots,n\}$, and setting $U_k \triangleq (Q, U_{k,Q})$, $Y_k \triangleq Y_{k,Q}$, for $k \in [1:K]$, and $X \triangleq (Q, X_Q)$.

References

[1] S. Zou, Y. Liang, L. Lai, and S. Shamai (Shitz). Degraded broadcast channel: Secrecy outside of a bounded range. In *Proc. IEEE Information Theory Workshop (ITW)*, Jerusalem, Israel, April 2015.

[2] S. Zou, Y. Liang, L. Lai, and S. Shamai (Shitz). Rate splitting and sharing for degraded broadcast channel with secrecy outside a bounded range. In *Proc. IEEE Int. Symp. Information Theory (ISIT)*, pages 1357–1361, Hong Kong, China, June 2015.

[3] S. Zou, Y. Liang, L. Lai, H. V. Poor, and S. Shamai (Shitz). K-user degraded broadcast channel with secrecy outside a bounded range. In *Proc. IEEE Information Theory Workshop (ITW)*, Cambridge, UK, September 2016.

[4] Y. Liang, H. V. Poor, and S. Shamai (Shitz). Information theoretic security. *Foundations and Trends in Communications and Information Theory*, 5(4-5):355–580, Now Publishers, Hanover, MA, USA, 2008.

[5] M. Bloch and J. Barros. *Physical-Layer Security: From Information Theory to Security Engineering*. Cambridge University Press, New York, USA, 2011.

[6] S. Zou, Y. Liang, L. Lai, H. V. Poor, and S. Shamai (Shitz). Broadcast networks with layered decoding and layered secrecy: Theory and applications. *Proceedings of the IEEE*, 103(10):1841–1856, Sept 2015.

[7] M. Baldi and S. Tomasin. *Physical and Data-Link Security Techniques for Future Communication Systems*. Springer, Switzerland, 2016.

[8] A. D. Wyner. The wire-tap channel. *Bell Syst. Tech. J.*, 54(8):1355–1387, October 1975.

[9] I. Csiszár and J. Körner. Broadcast channels with confidential messages. *IEEE Trans. Inform. Theory*, 24(3):339–348, May 1978.

[10] E. Ekrem and S. Ulukus. Secrecy capacity of a class of broadcast channels with an eavesdropper. *EURASIP J. Wirel. Commun. Netw.*, 2009:1:1–1:29, March 2009.

[11] R. Liu, T. Liu, H. V. Poor, and S. Shamai (Shitz). A vector generalization of Costa's entropy-power inequality with applications. *IEEE Trans. Inform. Theory*, 56(4):1865 –1879, April 2010.

[12] E. Ekrem and S. Ulukus. Degraded compound multi-receiver wiretap channels. *IEEE Trans. Inform. Theory*, 58(9):5681–5698, September 2012.

[13] S. Zou, Y. Liang, L. Lai, and S. Shamai. An information theoretic approach to secret sharing. *IEEE Trans. Inform. Theory*, 61(6):3121–3136, April 2015.

[14] A. El Gamal and Y.-H. Kim. *Network Information Theory*. Cambridge University Press, New York, 2012.

[15] Y. Liang, L. Lai, H. V. Poor, and S. Shamai (Shitz). A broadcast approach for fading wiretap channels. *IEEE Trans. Inform. Theory*, 60(2):842–858, Feb 2014.

[16] S. Shamai (Shitz) and A. Steiner. A broadcast approach for a single-user slowly fading MIMO channel. *IEEE Trans. Inform. Theory*, 49(10):2617–2635, October 2003.

[17] V. Rathi, M. Andersson, R. Thobaben, J. Kliewer, and M. Skoglund. Performance analysis and design of two edge-type LDPC codes for the BEC wiretap channel. *IEEE Trans. Inform. Theory*, 59(2):1048–1064, October 2013.

[18] O. O. Koyluoglu and H. El Gamal. Polar coding for secure transmission and key agreement. *IEEE Transactions on Information Forensics and Security*, 7(5):1472–1483, Oct 2012.

[19] J. del Olmo and J. R. Fonollosa. Strong secrecy on a class of degraded broadcast channels using polar codes. *ArXiv e-prints 1607.07815*, July 2016.

[20] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. Wiley-Interscience, New York, NY, USA, 1991.