Degraded Broadcast Channel With Secrecy Outside a Bounded Range

Shaofeng Zou, *Member, IEEE*, Yingbin Liang, *Senior Member, IEEE*, Lifeng Lai, *Member, IEEE*, H. Vincent Poor, *Fellow, IEEE*, and Shlomo Shamai (Shitz), *Fellow, IEEE*

***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 $1, \ldots, k-2$. The secrecy capacity region is fully characterized. The achievable scheme designates one superposition layer to each message with binning employed for each layer. Joint embedded coding and 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 of 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.**

Manuscript received September 20, 2016; revised August 23, 2017; accepted December 22, 2017. Date of publication January 11, 2018; date of current version February 15, 2018. S. Zou was supported in part by a National Science Foundation CAREER Award under Grant CCF-1026565 and in part by the National Science Foundation under Grant CNS-1116932. Y. Liang was supported in part by a National Science Foundation CAREER Award under Grant CCF-1026565, in part by the National Science Foundation under Grant CNS-1116932, and in part by the National Science Foundation under Grant CCF-1801846. L. Lai was supported by a National Science Foundation CAREER Award under Grant CCF-17-60889 and in part by the National Science Foundation under Grant CCF-16-65037 and Grant ECCS-16-60140. H. V. Poor was supported by the National Science Foundation under Grant CNS-1702808 and Grant ECCS-1647198. S. Shamai was supported in part by the European Commission in the Framework of the Network of Excellence in Wireless Communications NEWCOM and in part by the European Union's Horizon 2020 Research and Innovation Programme under Grant 694630. This paper was presented in part at the 2015 IEEE Information Theory Workshop [1], the 2015 IEEE International Symposium on Information Theory [2], and the 2016 IEEE Information Theory Workshop [3].

S. Zou is with the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA (e-mail: szou3@illinois.edu).

Y. Liang is with the Department of Electrical and Computer Engineering, The Ohio State University, Columbus, OH 43220 USA (e-mail: liang.889@osu.edu).

L. Lai is with the Electrical and Computer Engineering, University of California, Davis, CA 95616 USA (e-mail: lflai@ucdavis.edu).

H. V. Poor is with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544 USA (e-mail: poor@princeton.edu).

S. Shamai is with the Department of Electrical Engineering, Technion-Israel Institute of Technology, Haifa 32000, Israel (e-mail: sshlomo@ee.technion.ac.il).

Communicated by M. Bloch, Associate Editor for Shannon Theory.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2018.2791995

***Index Terms*— Binning, broadcast channel, embedded coding, multi-user, secrecy capacity region.**

I. 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 in 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]–[9]).

The basic broadcast channel with a secrecy constraint was the wiretap channel initiated by Wyner [10], in which a transmitter has a message intended for a legitimate receiver and wishes to keep this message secure 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 [11].

These broadcast models were further generalized to the multi-receiver case in [12] and [13], in which a transmitter has a number of messages intended for a set of receivers, and all messages need to be secure from an eavesdropper. Another type of extension is the broadcast channel with layered decoding and layered secrecy [13]–[15], 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 secure from all receivers with worse channel quality. More specifically, a K-receiver broadcast channel is considered in [15] ($K = 3$ in [13] and [14]), in which a transmitter sends K messages to K receivers. The channel quality gradually degrades from receiver K to receiver 1. 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 ignorant of messages $W_{k+1}, \ldots, W_K$ for $1 \leq k \leq K-1$.

We note that for the model considered in [13]–[15], the additional message decoded by a better receiver needs to be kept secure from the receiver with only one level worse channel quality. Here, any message W_k should be decoded

by receiver k, and be kept secure from receiver $k-1$. Such a model is well defined when receivers k and $k-1$ have nonzero difference in channel quality so that nonzero secrecy rate can be achieved for W_k. However, such a model is not useful if the difference in channel quality between the adjacent receivers becomes asymptotically small (i.e., close to zero), because essentially no secrecy rate can be achieved under the secrecy requirement of the model. For example, consider a fading broadcast channel, in which the channel to each receiver is determined by a channel gain coefficient with amplitude h, where h is continuous (the larger h, the better the channel). Here, the channel gains between two adjacent receivers can be arbitrarily close, and hence zero secrecy rate can be achieved for a message required to be decoded by one receiver and secured from the other receiver.

In this paper, we are interested in a model in which any message decoded at a certain receiver is not required to be kept secure from the one-level-worse receiver, but kept secure from the m-level-worse ($m > 1$) receiver. Such a model is valid as long as the "m levels" create nonzero differences in channel quality between receivers. In the fading channel, such a model captures scenarios in which messages intended for receivers with $h > h_0$ be kept secure from receivers with $h < h_0 - \Delta$, i.e., the messages are not necessarily kept secure from receivers with channel quality between $h_0 - \Delta$ and h_0. Here, $\Delta > 0$ guarantees nonzero difference between receivers required to decode the messages and receivers required to be ignorant of the messages, so that nonzero secrecy rate can be achieved. We refer to such a secrecy requirement as *secrecy outside a bounded range*.

We note that although this paper focuses on the case with Δ corresponding to two levels of channel quality (as we describe below in more detail), the technical treatment here already contains all the necessary ingredients to design capacity-achieving secrecy schemes for the general case with secrecy outside arbitrary m levels of channel quality. We discuss this generalization in Section V. We also note that we recently applied/generalized this study to the fading channel in [16].

More formally, 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 satisfies 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 $1 \le k \le K$, 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 secure from the receiver $k-2$, which has two level worse channel quality, for $3 \le k \le K$. 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 of this study [1], [2]. For the

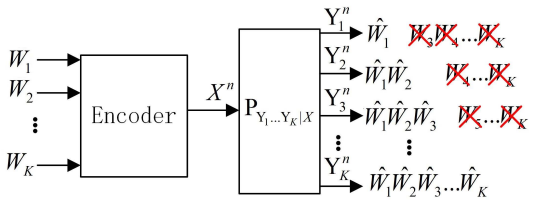

Fig. 1. The K-receiver broadcast channel with secrecy outside a bounded range.

three-receiver model, we show in [1] that superposition of messages and joint 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 as random sources to protect higher-layer messages. If the message at a certain layer (say layer k) is more than enough to protect the higher-layer messages, then such a message can also partially protect the message at layer k. Consequently, the protected message at layer k can be shared between layer k and its upper layer to enlarge the secrecy rate region.

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 sources of randomness to protect all higher-layer messages (from lower layer receivers). The design of joint embedded coding and binning is very complicated to handle. For example, consideration of whether to adopt 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) 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) Because we employ joint embedded coding and binning to secure multiple messages, the analysis of secrecy guarantee is much more involved than the cases with only one or two messages secured by 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 binning in each layer,

which avoids the complex consideration of whether or not it is necessary to employ 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 obtain the strong secrecy guarantee for the case with arbitrary K users, we generalize the arguments in [17]–[21] in which strong secrecy is obtained through channel resolvability. (4) Our development of the converse proof involves recursive construction of upper bounds on the rate of 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 II, we introduce our system model. In Section III, we present two example models with three receivers and four receivers, respectively, which motivate the design of the achievable scheme for the model with arbitrary K receivers. In Section IV, we present our main results for the model with arbitrary K receivers. In Section V, we discuss potential extensions of our results. Finally, in Section VI, we conclude our paper.

II. 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:

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

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

- K message sets: $W_k \in \mathcal{W}_k = \{1, \ldots, 2^{nR_k}\}$ for $k = 1, \ldots, K$, 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 $g_k^n : \mathcal{Y}_k^n \to (\mathcal{W}_1, \ldots, \mathcal{W}_k)$ that maps an output y_k^n to a message tuple $(\hat{w}_1, \ldots, \hat{w}_k)$ for $k = 1, \ldots, K$.

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

$$P_e^n = \Pr\left(\cup_{k=1}^K \{(W_1, \ldots, W_k) \neq g_k^n(Y_k^n)\}\right) \to 0, \tag{2}$$

and the secrecy metric at receiver k

$$I(W_k, \ldots, W_K; Y_{k-2}^n) \to 0, \tag{3}$$

for $k = 3, \ldots, K$. Here, we consider the strong secrecy metric instead of the weak secrecy metric as in [10] and [11], which requires the mutual information in (3) averaged over the block length n go to zero as n goes to infinity. The results in this paper may also be extended to an even stronger security notion, namely the semantic security [22], which enables quantifying the security of codes at finite lengths and is of practical importance in cryptography.

The asymptotically small error probability in (2) implies that each receiver k is able to decode messages $W_1, \ldots, W_k$, and the asymptotically small secrecy metric in (3) for each receiver k implies that $W_k, \ldots, W_K$ and Y_{k-2}^n are asymptotically independent, i.e., receiver k is kept ignorant of messages $W_{k+2}, \ldots, W_K$. Our goal is to characterize the *secrecy capacity region* which consists of all achievable rate tuples.

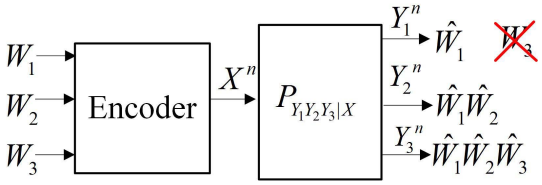

Fig. 2. The three-receiver broadcast channel with secrecy outside a bounded range.

III. Motivating Examples

In this section, we study two motivating examples 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 binning. 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.

A. Lessons Learned From $K = 3$

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 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.

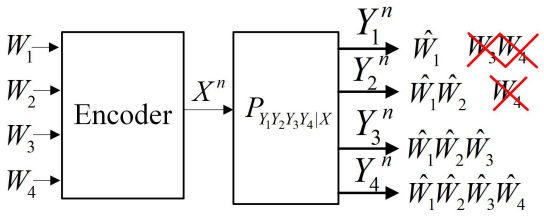

Fig. 3. The four-receiver broadcast channel with secrecy outside a bounded range.

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, binning is not necessary. If this is not sufficient to protect W_3, we apply 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 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 II. 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, $R_3 \le I(U_2; Y_2|U_1) - I(X; Y_1|U_1) + I(X; Y_3|U_2)$.

We note that a graphical representation of rate and equivocation quantities for the scalar Gaussian broadcast channel with secrecy outside a bounded range ($K = 3$) is presented in [23], which is based on the fundamental relationship between the mutual information and the minimum mean square error (MMSE) (I-MMSE approach [24]).

B. Lessons Learned From $K = 4$

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 secure from receiver 1, and the message W_4 is required to be secure 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 binning in layer 4 (if necessary) to protect W_4, and use the random message W_2 as embedded coding together with the 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

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

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 $R_4 \le I(X; Y_4|U_3)$ 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. 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 enlarged. In fact, such an enlarged region is 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 II. The secrecy capacity region consists of rate tuples (R_1, R_2, R_3, R_4) *satisfying*

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

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

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

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 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 binning in the fifth layer depends on whether the embedded coding in the third layer and (possibly) 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.

IV. 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 behind the design of the achievable scheme.

A. 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 II. The secrecy capacity region consists of rate tuples $(R_1, R_2, \ldots, R_K)$ satisfying

$$R_1 \le I(U_1; Y_1), \quad (9a)$$

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

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

for some $P_{U_1U_2\ldots U_K}$ such that the following Markov chain condition holds:

$$U_1 \to U_2 \to \cdots \to U_K \to Y_K \to \cdots \to Y_2 \to Y_1. \quad (10)$$

Proof: The proof of the achievability and the proof of converse are provided in Appendices A and C, 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 $1 \le k \le K$. 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 $3 \le l \le k \le K$. Furthermore, the bounds (9c) can be further written as

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

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 secure 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 $l-1 \le j \le k$ 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. However, the achievable schemes for the three-receiver model in Section III-A and the four-receiver model in Section III-B 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 binning in each layer, which avoids the complex consideration of whether or not it is necessary to employ binning for each layer. 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 IV-B.

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.

B. 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, 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.

1) Superposition, Binning, Embedded Coding: We design one layer of codebook for each message, i.e., layer k corresponds to W_k, for $1 \le k \le K$. To avoid the complex

consideration of whether to use binning, we employ 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 binning to provide randomness for secrecy.

2) 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 binning to protect $W_{k,2}$ and the higher layer messages from receiver Y_{k-1}, i.e., we require that $(W_{k,2}, W_{k+1,1}, W_{k+1,2}, \ldots, W_{K,1}, W_{K,2})$ be secure from receiver Y_{k-1}, for $2 \le k \le K-1$. 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.

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 secure 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:

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

where we use the convention that $\sum_{i=j}^{k} X_i = 0$ if $j > k$.

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 $3 \le k \le K-1$, $R_2 = R_{2,1}$ 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 $2 \le k \le K$.

3) 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 this 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 $3 \le k \le K$ 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:

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

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 B 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

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

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.

V. Extensions

In this paper, we have focused on the case with secrecy outside two levels of channel quality. In fact, such a case captures the essence of this type of model, and the design of the capacity-achieving secrecy schemes already consists of all the necessary ingredients to address the general case with secrecy outside arbitrary m levels of channel quality, i.e., the techniques of joint embedded coding and binning, rate splitting and sharing, and inductive Fourier-Motzkin elimination.

For $m > 2$ the rate splitting and sharing are more involved than for the case with $m = 2$. Each message W_k should be split into m submessages, $W_{k,1}, \ldots, W_{k,m}$. All the submessages in layers indexed from $k-m+1$ to $k-1$ but $\{W_{k-i,i+1}\}_{i=1}^{m-1}$ serve as embedded coding in addition to binning to protect $\{W_{k-i,i+1}\}_{i=1}^{m-1}$ and all higher-layer (with index no less than k)

submessages from receiver Y_{k-m}. Here, we index the submessages such that $W_{k,i}$ is secured from receiver $Y_{k-m+i-1}$, for $1 \leq i \leq m$ and $2 \leq k \leq K$. The upstream receiver Y_k can also decode all the submessages $\{W_{k-i,i+1}\}_{i=1}^{m-1}$. Hence, $\{W_{k-i,i+1}\}_{i=1}^{m-1}$ satisfy both the decoding and secrecy requirements for message W_k. Then the rates of $\{W_{k-i,i+1}\}_{i=1}^{m-1}$ can be counted towards the rate of W_k. We then define the rate sharing such that $R_k = \sum_{i=0}^{m-1} R_{k-i,i+1}$. Based on the above achievable scheme, we can obtain an achievable region in terms of $W_{k,j}$, for $1 \leq k \leq K$ and $1 \leq j \leq m$. We then project this region onto the rate space $(R_1, \ldots, R_K)$. This can be done by a similar but rather complicated inductive Fourier-Motzkin elimination.

The results here can be further generalized to models with continuously changing channel state parameters, e.g., Gaussian fading channel [25], and Gaussian multiple input multiple output (MIMO) channel [26]. For example, in our recent work [16], we study the fading channel with secrecy outside a bounded range. More specifically, we first quantify the continuous channel state with infinitely many discrete channel states, and then apply/generalize the techniques in this paper.

VI. 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 binning, and rate splitting and sharing. The combination of embedded coding and binning to achieve secrecy captures the fact that lower-layer messages can serve as embedded coding to protect higher-layer messages. 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.

This paper has focused on characterizing the information theoretic performance limits which are based on random coding arguments. It is of further interest to design practical coding schemes such as polar codes [27]–[33] and low density parity check (LDPC) codes [34] to achieve secrecy capacity.

Appendix A
Achievability Proof of Theorem 1

The achievability proof is based on superposition coding, embedded coding, 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). 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 [35, ch. 3], [36]. The achievable scheme is designed as follows:

A. 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 by $\prod_{i=1}^n P(u_{2,i}|u_{1,i})$. 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 $2^{n(R_{3,1}+R_{3,2}+R_{3,3})}$ i.i.d. u_3^n by $\prod_{i=1}^n P(u_{3,i}|u_{2,i})$. Partition these codewords into $2^{nR_{3,1}}$ bins, and further partition each bin into $2^{nR_{3,2}}$ sub-bins. Hence, there are $2^{nR_{3,3}}$ 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 $l_3 \in [1 : 2^{nR_{3,3}}]$ 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 $4 \leq k \leq K$, for each $u_{k-1}^n(w_1, \ldots, w_{k-1,1}, w_{k-1,2}, l_{k-1})$, generate $2^{n(R_{k,1}+R_{k,2}+R_{k,3})}$ i.i.d. u_k^n by $\prod_{i=1}^n P(u_{k,i}|u_{k-1,i})$. Partition these codewords into $2^{nR_{k,1}}$ bins, and further partition each bin into $2^{nR_{k,2}}$ sub-bins. Hence, there are $2^{nR_{k,3}}$ 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 $l_k \in [1 : 2^{nR_{k,3}}]$ 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.

B. 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 $l_k \in [1 : 2^{nR_{k,3}}]$ for $3 \leq k \leq K$, and sends

$$x^n(w_1, \ldots, w_{K,1}, w_{K,2}, l_3, \ldots, l_K).$$

C. 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

$$\left(u_1^n(\widehat{w}_1), u_2^n(\widehat{w}_1, \widehat{w}_{2,1}, \widehat{w}_{2,2}), y_2^n\right) \in T_\epsilon^n(P_{U_1U_2\ Y_2}).$$

  Otherwise, it declares an error.

- For $3 \leq k \leq K$, 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.

D. Analysis of Error Probability

By the law of large numbers and the packing lemma [37], 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:

$$\begin{aligned} R_1 &\leq I(U_1; Y_1), \\ R_{2,1} + R_{2,2} &\leq I(U_2; Y_2|U_1), \\ R_{k,1} + R_{k,2} + R_{k,3} &\leq I(U_k; Y_k|U_{k-1}), \quad \text{for } 3 \leq k \leq K. \end{aligned} \tag{14}$$

E. Analysis of Secrecy

We require that

$$(W_{k-1,2}, W_{k,1}, W_{k,2}, \ldots, W_{K,1}, W_{K,2})$$

be secure from receiver Y_{k-2}, for $3 \leq k \leq K$. It then suffices to show that as $n \to \infty$,

$$I\left(W_{k-1,2}\ , W_{k,1}, W_{k,2}, \ldots, W_{K,1}, W_{K,2}; Y_{k-2}^n \middle| \mathcal{C}\right) \to 0, \tag{15}$$

for $3 \leq k \leq K$, where $\mathcal{C}$ denotes a random codebook over the codebook ensemble. 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. For notational convenience, let $L_j^k = (L_j, \ldots, L_k)$, $l_j^k = (l_j, \ldots, l_k)$ for $3 \leq j \leq k \leq K$, and $\mathcal{M}_k = (W_{k-1,2}, W_{k,1}, W_{k,2}, \ldots, W_{K,1}, W_{K,2})$, for $3 \leq k \leq K$.

By the independence of the messages, i.e., $\mathcal{M}_k$ and $(W_1, \ldots, W_{k-2,1}, W_{k-2,2}, L_{k-2})$ are independent, and the fact that given $\mathcal{C}$, U_{k-2}^n is a deterministic function of $(W_1, \ldots, W_{k-2,1}, W_{k-2,2}, L_{k-2})$, it follows that U_{k-2}^n is independent of $\mathcal{M}_k$, and thus

$$\begin{aligned} I(\mathcal{M}_k; Y_{k-2}^n|\mathcal{C}) &= H(\mathcal{M}_k|\mathcal{C}) - H(\mathcal{M}_k|Y_{k-2}^n, \mathcal{C}) \\ &= H(\mathcal{M}_k|U_{k-2}^n, \mathcal{C}) - H(\mathcal{M}_k|Y_{k-2}^n, \mathcal{C}) \\ &\leq H(\mathcal{M}_k|U_{k-2}^n, \mathcal{C}) - H(\mathcal{M}_k|Y_{k-2}^n, U_{k-2}^n, \mathcal{C}) \\ &\leq I(\mathcal{M}_k; Y_{k-2}^n|U_{k-2}^n, \mathcal{C}). \end{aligned} \tag{16}$$

Connecting the idea of channel resolvability to secrecy [18]–[21], it follows that

$$\begin{aligned} &I(\mathcal{M}_k; Y_{k-2}^n|U_{k-2}^n, \mathcal{C}) \\ &\quad = \mathbb{E} \log \frac{P(\mathcal{M}_k, Y_{k-2}^n|U_{k-2}^n, \mathcal{C})}{P(\mathcal{M}_k|U_{k-2}^n, \mathcal{C}) P(Y_{k-2}^n|U_{k-2}^n, \mathcal{C})} \\ &\quad = \mathbb{E} \log \frac{P(Y_{k-2}^n|\mathcal{M}_k, U_{k-2}^n, \mathcal{C})}{P(Y_{k-2}^n|U_{k-2}^n, \mathcal{C})} \\ &\quad = \mathbb{E}\left[\log \frac{P(Y_{k-2}^n|\mathcal{M}_k, U_{k-2}^n, \mathcal{C})}{P(Y_{k-2}^n|U_{k-2}^n)} + \log \frac{P(Y_{k-2}^n|U_{k-2}^n)}{P(Y_{k-2}^n|U_{k-2}^n, \mathcal{C})}\right] \\ &\quad \leq \mathbb{E}\left[\log \frac{P(Y_{k-2}^n|\mathcal{M}_k, U_{k-2}^n, \mathcal{C})}{P(Y_{k-2}^n|U_{k-2}^n)}\right], \end{aligned} \tag{17}$$

where the last step is due to the fact that

$$\begin{aligned} &\mathbb{E}\left[\log \frac{P(Y_{k-2}^n|U_{k-2}^n)}{P(Y_{k-2}^n|U_{k-2}^n, \mathcal{C})}\right] \\ &\qquad = -\mathbb{E}\left[D(P_{Y_{k-2}^n|U_{k-2}^n, \mathcal{C}} \| P_{Y_{k-2}^n|U_{k-2}^n})\right] \leq 0, \end{aligned} \tag{18}$$

where

$$D(P||Q) = \sum_i P(i) \log \frac{P(i)}{Q(i)}$$

is the Kullback-Leibler divergence between two distributions P and Q.

Conditioned on a realization C of the random codebook $\mathcal{C}$, we obtain that

$$\begin{aligned} &P\left(y_{k-2}^n|w_{k-1,2}, w_{k,1}, w_{k,2}, \ldots, w_{K,1}, w_{K,2}, u_{k-2}^n, C\right) \\ &\quad = \sum_{w_{k-1,1}, l_{k-1}^K} \left[P\left(y_{k-2}^n, w_{k-1,1}, l_{k-1}^K|w_{k-1,2}, w_{k,1}, w_{k,2}, \ldots, w_{K,1}, w_{K,2}, u_{k-2}^n, C\right)\right] \\ &\quad = \frac{\sum_{w_{k-1,1}, l_{k-1}^K} \left[P\left(y_{k-2}^n|w_{k-1,1}, w_{k-1,2}, \ldots, w_{K,1}, w_{K,2}, l_{k-1}^K, u_{k-2}^n, C\right)\right]}{2^{n(R_{k-1,1} + R_{k-1,3} + \ldots + R_{K,3})}} \\ &\quad = \frac{\sum_{w_{k-1,1}, l_{k-1}^K} \left[P\left(y_{k-2}^n|u_K^n(w_1, \ldots, w_{k,1}, w_{k,2}, \ldots, w_{K,1}, w_{K,2}, l_3^K), C\right)\right]}{2^{n(R_{k-1,1} + R_{k-1,3} + \ldots + R_{K,3})}} \\ &\quad = \frac{\sum_{w_{k-1,1}, l_{k-1}^K} \left[P\left(y_{k-2}^n|u_K^n(w_1, \ldots, w_{k,1}, w_{k,2}, \ldots, w_{K,1}, w_{K,2}, l_3^K)\right)\right]}{2^{n(R_{k-1,1} + R_{k-1,3} + \ldots + R_{K,3})}}, \end{aligned} \tag{19}$$

where the last step is due to the Markov chain condition

$$\mathcal{C} \to U_K^n \to Y_{k-2}^n.$$

Due to the symmetry of the random codebook construction, when computing the expectation in (17), we can assume that all the indices except $(W_{k-1,1}, L_{k-1}^K)$ are fixed constants and equal to one. For notational convenience, we only include the indices $(W_{k-1,1}, L_{k-1}^K)$ and ignore all those fixed indices when labeling the codewords. For example, instead of $u_{k-2}^n(w_1, \ldots, w_{k-2,1}, w_{k-2,2}, l_{k-2})$ and $u_K^n(w_1, \ldots, w_{k,1}, w_{k,2}, \ldots, w_{K,1}, w_{K,2}, l_3^K)$, we use u_{k-2}^n and $u_K^n(w_{k-1,1}, l_{k-1}^K)$.

Following steps similar to those in [19], we obtain that

$$\begin{aligned}
&\mathbb{E}\left[\log\frac{P(Y_{k-2}^n|\mathcal{M}_k,U_{k-2}^n,\mathcal{C})}{P(Y_{k-2}^n|U_{k-2}^n)}\right]\\
&\overset{(a)}{=}\sum_C P(C)\sum_{y_{k-2}^n}P(y_{k-2}^n|1,\ldots,1,u_{k-2}^n,C)\\
&\quad\times\log\frac{P(y_{k-2}^n|1,\ldots,1,u_{k-2}^n,C)}{P(y_{k-2}^n|u_{k-2}^n)}\\
&=\sum_C P(C)\sum_{y_{k-2}^n}\frac{1}{2^{n(R_{k-1,1}+R_{k-1,3}+\ldots+R_{K,3})}}\\
&\quad\times\sum_{w_{k-1,1},l_{k-1}^K}P(y_{k-2}^n|u_K^n(w_{k-1,1}l_{k-1}^K))\\
&\qquad\sum_{\tilde{w}_{k-1,1},\tilde{l}_{k-1}^K}P(y_{k-2}^n|u_K^n(\tilde{w}_{k-1,1},\tilde{l}_{k-1}^K))\\
&\quad\times\log\frac{}{2^{n(R_{k-1,1}+R_{k-1,3}+\ldots+R_{K,3})}P(y_{k-2}^n|u_{k-2}^n)}\\
&\overset{(b)}{=}\sum_C P(u_{k-2}^n)\prod_{\hat{w}_{k-1,1},\hat{l}_{k-1}}\Bigg[P(u_{k-1}^n(\hat{w}_{k-1,1},\hat{l}_{k-1})|u_{k-2}^n)\\
&\quad\times\Bigg[\cdots\prod_{\hat{l}_K}P(u_K^n(\hat{w}_{k-1,1},\hat{l}_{k-1}^K)|u_{K-1}^n(\hat{w}_{k-1,1},\hat{l}_{k-1}^{K-1}))\Bigg]\Bigg]\\
&\quad\times\sum_{y_{k-2}^n}\frac{1}{2^{n(R_{k-1,1}+R_{k-1,3}+\ldots+R_{K,3})}}\\
&\quad\times\sum_{w_{k-1,1},l_{k-1}^K}P(y_{k-2}^n|u_K^n(w_{k-1,1}l_{k-1}^K))\\
&\qquad\sum_{\tilde{w}_{k-1,1},\tilde{l}_{k-1}^K}P(y_{k-2}^n|u_K^n(\tilde{w}_{k-1,1},\tilde{l}_{k-1}^K))\\
&\quad\times\log\frac{}{2^{n(R_{k-1,1}+R_{k-1,3}+\ldots+R_{K,3})}P(y_{k-2}^n|u_{k-2}^n)}\\
&\overset{(c)}{=}\frac{1}{2^{n(R_{k-1,1}+R_{k-1,3}+\ldots+R_{K,3})}}\sum_C\sum_{y_{k-2}^n}\sum_{w_{k-1,1},l_{k-1}^K}P(u_{k-2}^n)\\
&\quad\times\prod_{\hat{w}_{k-1,1},\hat{l}_{k-1}}\Bigg[P(u_{k-1}^n(\hat{w}_{k-1,1},\hat{l}_{k-1})|u_{k-2}^n)\\
&\quad\times\Bigg[\cdots\prod_{\hat{l}_K}P(u_K^n(\hat{w}_{k-1,1},\hat{l}_{k-1}^K)|u_{K-1}^n(\hat{w}_{k-1,1},\hat{l}_{k-1}^{K-1}))\Bigg]\Bigg]\\
&\quad\times P(y_{k-2}^n|u_K^n(w_{k-1,1}l_{k-1}^K))\\
&\quad\times\log\left(\frac{\sum_{\tilde{w}_{k-1,1},\tilde{l}_{k-1}^K}P(y_{k-2}^n|u_K^n(\tilde{w}_{k-1,1},\tilde{l}_{k-1}^K))}{2^{n(R_{k-1,1}+R_{k-1,3}+\ldots+R_{K,3})}P(y_{k-2}^n|u_{k-2}^n)}\right),
\end{aligned}\tag{20}$$

where in (a), by the symmetry of the random codebook construction, we let $U_{k-2}^n = u_{k-2}^n(1,\ldots,1)$, $\mathcal{M}_k = (1,\ldots,1)$; and in (b) and the following equations, C consists only of those codewords with all the indices except $(w_{k-1,1}, l_{k-1}^K)$ being one; (c) is obtained by reordering those summations.

From (20), we further obtain (21) on bottom of this page, where (a) follows by the concavity of the logarithm and Jensen's inequality applied to the expectation over all the codewords except

$$\left(u_{k-2}^n, u_{k-1}^n(w_{k-1},l_{k-1}),\ldots,u_K^n(w_{k-1},l_{k-1}^K)\right).$$

$$\begin{aligned}
&\mathbb{E}\left[\log\left(\frac{P(Y_{k-2}^n|\mathcal{M}_k,U_{k-2}^n,\mathcal{C})}{P(Y_{k-2}^n|U_{k-2}^n)}\right)\right]\\
&=\frac{1}{2^{n(R_{k-1,1}+R_{k-1,3}+\ldots+R_{K,3})}}\sum_{w_{k-1,1},l_{k-1}^K}\sum_{y_{k-2}^n}\sum_{u_{k-2}^n}\sum_{u_{k-1}^n(w_{k-1,1},l_{k-1})}\cdots\sum_{u_K^n(w_{k-1,1}l_{k-1}^K)}\\
&\quad P(y_{k-2}^n,u_K^n(w_{k-1,1}l_{k-1}^K),\ldots,u_{k-2}^n)\sum_{\substack{C\setminus\{u_{k-2}^n,\ldots,\\ u_K^n(w_{k-1,1}l_{k-1}^K)\}}}\prod_{\hat{w}_{k-1,1},\hat{l}_{k-1}}\Bigg[P(u_{k-1}^n(\hat{w}_{k-1,1},\hat{l}_{k-1})|u_{k-2}^n)\\
&\quad\times\Bigg[\cdots\prod_{\hat{l}_K}P(u_K^n(\hat{w}_{k-1,1},\hat{l}_{k-1}^K)|u_{K-1}^n(\hat{w}_{k-1,1},\hat{l}_{k-1}^{K-1}))\Bigg]\Bigg]\\
&\qquad\qquad\log\left(\frac{\sum_{\tilde{w}_{k-1,1},\tilde{l}_{k-1}^K}P(y_{k-2}^n|u_K^n(\tilde{w}_{k-1,1},\tilde{l}_{k-1}^K))}{2^{n(R_{k-1,1}+R_{k-1,3}+\ldots+R_{K,3})}P(y_{k-2}^n|u_{k-2}^n)}\right)\\
&\qquad\qquad\frac{}{P(u_{k-1}^n(w_{k-1},l_{k-1}),\ldots,w_K(w_{k-1},l_{k-1}^K)|u_{k-2}^n)}\\
&\overset{(a)}{\le}\frac{1}{2^{n(R_{k-1,1}+R_{k-1,3}+\ldots+R_{K,3})}}\sum_{w_{k-1,1},l_{k-1}^K}\sum_{y_{k-2}^n}\sum_{u_{k-2}^n}\sum_{u_{k-1}^n(w_{k-1,1},l_{k-1})}\cdots\sum_{u_K^n(w_{k-1,1}l_{k-1}^K)}\\
&\quad P(y_{k-2}^n,u_K^n(w_{k-1,1}l_{k-1}^K),\ldots,u_{k-2}^n)\log\left(\begin{array}{l}\displaystyle\sum_{(\tilde{w}_{k-1,1},\tilde{l}_{k-1}^K)\neq(w_{k-1,1},l_{k-1}^K)}\frac{\mathbb{E}\left[P(y_{k-2}^n|U_K^n(\tilde{w}_{k-1,1},\tilde{l}_{k-1}^K))\right]}{P(y_{k-2}^n|u_{k-2}^n)2^{n(R_{k-1,1}+R_{k-1,3}+\ldots+R_{K,3})}}\\ \quad+\frac{P(y_{k-2}^n|u_K^n(w_{k-1,1},l_{k-1}^K))}{2^{n(R_{k-1,1}+R_{k-1,3}+\ldots+R_{K,3})}P(y_{k-2}^n|u_{k-2}^n)}\end{array}\right)
\end{aligned}\tag{21}$$

We now consider the expectation in (21) for different values of $(\tilde{w}_{k-1,1}, \tilde{l}_{k-1}^K)$. We first define

$$\frac{P(y_{k-2}^n|u_K^n(w_{k-1,1}, l_{k-1}^K))}{2^{n(R_{k-1,1}+R_{k-1,3}+\ldots+R_{K,3})}P(y_{k-2}^n|u_{k-2}^n)} \triangleq A_K. \quad (22)$$

For $(\tilde{w}_{k-1,1}, \tilde{l}_{k-1}^{K-1}) = (w_{k-1,1}, l_{k-1}^{K-1})$ but $\tilde{l}_K \neq l_K$, we obtain

$$\begin{aligned}\sum_{\tilde{l}_K \neq l_K} & \frac{P(y_{k-2}^n|u_{K-1}^n(w_{k-1,1}, l_{k-1}^{K-1}))}{2^{n(R_{k-1,1}+R_{k-1,3}+\ldots+R_{K,3})}P(y_{k-2}^n|u_{k-2}^n)} \\ &\leq \frac{P(y_{k-2}^n|u_{K-1}^n(w_{k-1,1}, l_{k-1}^{K-1}))}{2^{n(R_{k-1,1}+R_{k-1,3}+\ldots+R_{K-1,3})}P(y_{k-2}^n|u_{k-2}^n)} \\ &\triangleq A_{K-1}. \end{aligned} \quad (23)$$

More generally, for any $k-1 \leq j \leq K-1$, for $(\tilde{w}_{k-1,1}, \tilde{l}_{k-1}^j) = (w_{k-1,1}, l_{k-1}^j)$ but $\tilde{l}_{j+1} \neq l_{j+1}$, we obtain

$$\begin{aligned}\sum_{\tilde{l}_{j+1}^K : \tilde{l}_{j+1} \neq l_{j+1}} & \frac{P(y_{k-2}^n|u_j^n(w_{k-1,1}, l_{k-1}^j))}{2^{n(R_{k-1,1}+R_{k-1,3}+\ldots+R_{K,3})}P(y_{k-2}^n|u_{k-2}^n)} \\ &\leq \frac{P(y_{k-2}^n|u_j^n(w_{k-1,1}, l_{k-1}^j))}{2^{n(R_{k-1,1}+R_{k-1,3}+\ldots+R_{j,3})}P(y_{k-2}^n|u_{k-2}^n)} \\ &\triangleq A_j. \end{aligned} \quad (24)$$

For $(\tilde{w}_{k-1,1}, \tilde{l}_{k-1}) \neq (w_{k-1,1}, l_{k-1})$, we obtain

$$\sum_{\substack{(\tilde{w}_{k-1,1}, \tilde{l}_{j+1}^K): \\ (\tilde{w}_{k-1,1}, \tilde{l}_{k-1}) \neq (w_{k-1,1}, l_{k-1})}} \frac{1}{2^{n(R_{k-1,1}+R_{k-1,3}+\ldots+R_{K,3})}} \leq 1. \quad (25)$$

Combining (22) (24) and (25) yields that the term within the log in (21) is upper bounded by $1 + \sum_{j=k-1}^K A_j$, which further implies that

$$\begin{aligned}\mathbb{E}\left[\log \frac{P(Y_{k-2}^n|\mathcal{M}_k, U_{k-2}^n, \mathcal{C})}{P(Y_{k-2}^n|U_{k-2}^n)}\right] &\leq \mathbb{E}\log\left(1 + \sum_{j=k-1}^K \frac{P(Y_{k-2}^n|U_j^n(W_{k-1,1}, L_{k-1}^j))}{\left[\begin{smallmatrix}P(Y_{k-2}^n|U_{k-2}^n) \\ \times 2^{n(R_{k-1,1}+R_{k-1,3}+\ldots+R_{j,3})}\end{smallmatrix}\right]}\right) \\ &\leq \sum_{j=k-1}^K \mathbb{E}\log\left(1 + \frac{P(Y_{k-2}^n|U_j^n(W_{k-1,1}, L_{k-1}^j))}{\left[\begin{smallmatrix}P(Y_{k-2}^n|U_{k-2}^n) \\ \times 2^{n(R_{k-1,1}+R_{k-1,3}+\ldots+R_{j,3})}\end{smallmatrix}\right]}\right). \end{aligned} \quad (26)$$

By the symmetry of the random codeword generation, we assume that $(W_{k-1,1}, L_{k-1}^K)$ are fixed, and thus in the following proof, we ignore these indices. For any $k-1 \leq j \leq K$, it then follows that

$$\begin{aligned}\mathbb{E}\log & \left(1 + \frac{P(Y_{k-2}^n|U_j^n)}{2^{n(R_{k-1,1}+R_{k-1,3}+\ldots+R_{j,3})}P(Y_{k-2}^n|U_{k-2}^n)}\right) \\ &= \sum_{(u_{k-2}^n, u_j^n, y_{k-2}^n) \in T_\epsilon^n(P_{U_{k-2}U_jY_{k-2}})} P(u_{k-2}^n, u_j^n, y_{k-2}^n) \\ &\quad \times \log\left(1 + \frac{P(y_{k-2}^n|u_j^n)}{2^{n(R_{k-1,1}+R_{k-1,3}+\ldots+R_{j,3})}P(y_{k-2}^n|u_{k-2}^n)}\right) \\ &\quad + \sum_{(u_{k-2}^n, u_j^n, y_{k-2}^n) \notin T_\epsilon^n(P_{U_{k-2}U_jY_{k-2}})} P(u_{k-2}^n, u_j^n, y_{k-2}^n) \\ &\quad \times \log\left(1 + \frac{P(y_{k-2}^n|u_j^n)}{2^{n(R_{k-1,1}+R_{k-1,3}+\ldots+R_{j,3})}P(y_{k-2}^n|u_{k-2}^n)}\right) \\ &\triangleq d_1 + d_2. \end{aligned} \quad (27)$$

Using the inequalities in [36, Appendix] and following steps similar to those in [19], we have

$$\begin{aligned} d_1 &\leq \sum_{(u_{k-2}^n, u_j^n, y_{k-2}^n) \in T_\epsilon^n(P_{U_{k-2}U_jY_{k-2}})} P(u_{k-2}^n, u_j^n, y_{k-2}^n) \\ &\quad \times \log\left(1 + \frac{2^{-n(1-\epsilon)H(Y_{k-2}|U_j)}}{2^{n(R_{k-1,1}+R_{k-1,3}+\ldots+R_{j,3})-n(1+\epsilon)H(Y_{k-2}|U_{k-2})}}\right) \\ &\leq \log\left(1 + \frac{2^{-n(1-\epsilon)H(Y_{k-2}|U_j)}}{2^{n(R_{k-1,1}+R_{k-1,3}+\ldots+R_{j,3})-n(1+\epsilon)H(Y_{k-2}|U_{k-2})}}\right) \end{aligned} \quad (28)$$

which vanishes as $n \to \infty$ if

$$\begin{aligned} R_{k-1,1} + R_{k-1,3} + \ldots + R_{j,3} \\ > I(U_j; Y_{k-2}|U_{k-2}) + 2\epsilon H(Y_{k-2}|U_{k-2}). \end{aligned} \quad (29)$$

To show $d_2 \to 0$ as $n \to \infty$, it follows that

$$\begin{aligned} d_2 &\leq \sum_{\substack{(u_{k-2}^n, u_j^n, y_{k-2}^n) \\ \notin T_\epsilon^n\left(P_{U_{k-2}U_jY_{k-2}}\right) \\ \left(u_{k-2}^n, u_j^n, y_{k-2}^n\right) \\ \in \mathrm{supp}\left(P_{U_{k-2}^n, U_j^n, Y_{k-2}^n}\right)}} P(u_{k-2}^n, u_j^n, y_{k-2}^n) \log\left(1 + \left(\frac{1}{\mu}\right)^n\right) \\ &\leq 2|\mathcal{U}_{k-2}||\mathcal{U}_j||\mathcal{Y}_{k-2}|e^{-\epsilon^2\phi n/3} n \log\left(1 + \frac{1}{\mu}\right) \\ &\to 0, \text{ as } n \to \infty. \end{aligned} \quad (30)$$

where $\mathrm{supp}(P_X)$ is defined to be the support of a distribution P_X, $|\mathcal{U}_{k-2}|$, $|\mathcal{U}_j|$ and $|\mathcal{Y}_{k-2}|$ are the support sizes of U_{k-2}, U_j and Y_{k-2}, respectively, and

$$\begin{aligned} \mu &= \min_{(u_{k-2}, y_{k-2}) \in \mathrm{supp}\left(P_{U_{k-2}Y_{k-2}}\right)} P(y_{k-2}|u_{k-2}), \\ \phi &= \min_{(u_{k-2}, u_j, y_{k-2}) \in \mathrm{supp}\left(P_{U_{k-2}U_jY_{k-2}}\right)} P(u_{k-2}u_j y_{k-2}). \end{aligned} \quad (31)$$

Therefore, if the following conditions are satisfied for $3 \leq k \leq K$ and $k-1 \leq j \leq K$:

$$R_{k-1,1} + R_{k-1,3} + \ldots + R_{j,3} > I(U_j; Y_{k-2}|U_{k-2}), \quad (32)$$

then, for $3 \leq k \leq K$,

$$I(\mathcal{M}_k; Y_{k-2}^n|U_{k-2}^n, \mathcal{C}) \to 0, \text{ as } n \to \infty. \quad (33)$$

Combining the bounds in (14) and (32), and by choosing $R_{k,1} + R_{k,2} + R_{k,3} = I(U_k; Y_k|U_{k-1})$, we conclude that

the rate tuple $(R_1, R_{2,1}, R_{2,2}, \ldots, R_{K,1}, R_{K,2})$ is achievable if

$$\begin{aligned} R_1 &\le I(U_1; Y_1), \\ R_{k,1} + R_{k,2} &\le I(U_k; Y_k|U_{k-1}), \text{ for } 2 \le k \le K, \\ R_{k-1,2} + \sum_{i=k}^{j} (R_{i,1} + R_{i,2}) &\le \sum_{i=k-1}^{j} I(U_i; Y_i|U_{i-1}) - I(U_j; Y_{k-2}|U_{k-2}), \\ &\text{for } 3 \le k \le K, \text{ and } k-1 \le j \le K. \end{aligned} \tag{34}$$

F. 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 secure from receiver Y_{k-2}, for $3 \le k \le K$. 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 $3 \le k \le K-1$, 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 $2 \le k \le K$ 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 B, and obtain the achievable region given in Theorem 1.

Appendix B
Inductive Fourier-Motzkin Elimination

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

$$R_1 \le I(U_1; Y_1), \tag{35a}$$

$$R_{k,1} + R_{k,2} \le I(U_k; Y_k|U_{k-1}), \text{ for } 2 \le k \le K, \tag{35b}$$

$$\begin{aligned} R_{l-1,2} + \sum_{i=l}^{j} (R_{i,1} + R_{i,2}) &\le \sum_{i=l-1}^{j} I(U_i; Y_i|U_{i-1}) - I(U_j; Y_{l-2}|U_{l-2}), \\ &\text{for } 3 \le l \le K, l-1 \le j \le K, \end{aligned} \tag{35c}$$

$$R_2 = R_{2,1}, \tag{35d}$$

$$R_k = R_{k-1,2} + R_{k,1}, \text{ for } 3 \le k \le K-1, \tag{35e}$$

$$R_K = R_{K-1,2} + R_{K,1} + R_{K,2}, \tag{35f}$$

where the bounds (35a), (35b) and (35c) correspond to the achievable region after rate splitting, which are expressed in terms of component rates, and the bounds (35d), (35e) and (35f) 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 $3 \le k \le K-1$ 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 $3 \le k \le K-1$, and $R_{K-1,2} + R_{K,1} + R_{K,2}$ with R_K, and we obtain the following region:

$$\begin{aligned} R_1 &\le I(U_1; Y_1), \\ R_2 + R_{2,2} &\le I(U_2; Y_2|U_1), \\ R_{k,1} + R_{k,2} &\le I(U_k; Y_k|U_{k-1}), \text{ for } 3 \le k \le K, \\ \sum_{i=l}^{j} R_i + R_{j,2} &\le \sum_{i=l-1}^{j} I(U_i; Y_i|U_{i-1}) - I(U_j; Y_{l-2}|U_{l-2}), \\ &\text{for } 3 \le l \le K, \quad l-1 \le j \le K-1, \\ \sum_{i=l}^{K} R_i &\le \sum_{i=l-1}^{K} I(U_i; Y_i|U_{i-1}) - I(U_K; Y_{l-2}|U_{l-2}), \\ &\text{for } 3 \le l \le K, \\ R_k &= R_{k-1,2} + R_{k,1}, \quad \text{for } 3 \le k \le K-1, \\ R_K &= R_{K-1,2} + R_{K,1} + R_{K,2}. \end{aligned} \tag{36}$$

To start, 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:

$$\begin{aligned} R_1 &\le I(U_1; Y_1), \\ R_2 + R_{2,2} &\le I(U_2; Y_2|U_1), \\ R_{3,1} + R_{3,2} &\le I(U_3; Y_3|U_2), \\ R_{2,2} &\le I(U_2; Y_2|U_1) - I(U_2; Y_1|U_1), \\ R_3 + R_{3,2} &\le \sum_{i=2}^{3} I(U_i; Y_i|U_{i-1}) - I(U_3; Y_1|U_1), \\ R_{3,2} &\le I(U_3; Y_3|U_2) - I(U_3; Y_2|U_2), \\ R_3 &= R_{2,2} + R_{3,1}. \end{aligned} \tag{37}$$

We then obtain the following inequalities after elimination:

$$\begin{aligned} R_1 &\le I(U_1; Y_1), \\ R_2 &\le I(U_2; Y_2|U_1), \\ \sum_{i=2}^{3} R_i + R_{3,2} &\le \sum_{i=2}^{3} I(U_i; Y_i|U_{i-1}), \\ R_3 + R_{3,2} &\le \sum_{i=2}^{3} I(U_i; Y_i|U_{i-1}) - I(U_3; Y_1|U_1), \\ R_{3,2} &\le I(U_3; Y_3|U_2) - I(U_3; Y_2|U_2), \end{aligned} \tag{38}$$

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:

$$\begin{aligned} R_{4,1} + R_{4,2} &\le I(U_4; Y_4|U_3), \\ \sum_{i=j}^{4} R_i + R_{4,2} &\le \sum_{i=j-1}^{4} I(U_i; Y_i|U_{i-1}) - I(U_4; Y_{j-2}|U_{j-2}), \\ &\text{for } 3 \le j \le 5, \\ R_4 &= R_{3,2} + R_{4,1}. \end{aligned} \tag{39}$$

We then obtain the following bounds after elimination:

$$\begin{aligned}
R_1 &\le I(U_1;Y_1),\\
\sum_{i=2}^{j} R_i &\le \sum_{i=2}^{j} I(U_i;Y_i|U_{i-1}), \text{ for } 2\le j\le 3,\\
\sum_{i=2}^{4} R_i + R_{4,2} &\le \sum_{i=2}^{4} I(U_i;Y_i|U_{i-1}),\\
\sum_{i=l}^{j} R_i &\le \sum_{i=l-1}^{j} I(U_i;Y_i|U_{i-1}) - I(U_j;Y_{l-2}|U_{l-2}),\\
&\qquad \text{for } 3\le l\le j\le 3,\\
\sum_{i=l}^{4} R_i + R_{4,2} &\le \sum_{i=l-1}^{4} I(U_i;Y_i|U_{i-1}) - I(U_4;Y_{l-2}|U_{l-2}),\\
&\qquad \text{for } 3\le l\le 5, \qquad (40)
\end{aligned}$$

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 $k=4$:

$$\begin{aligned}
R_1 &\le I(U_1;Y_1),\\
\sum_{i=2}^{j} R_i &\le \sum_{i=2}^{j} I(U_i;Y_i|U_{i-1}), \text{ for } 2\le j\le k-1,\\
\sum_{i=2}^{k} R_i + R_{k,2} &\le \sum_{i=2}^{k} I(U_i;Y_i|U_{i-1}),\\
\sum_{i=l}^{j} R_i &\le \sum_{i=l-1}^{j} I(U_i;Y_i|U_{i-1}) - I(U_j;Y_{l-2}|U_{l-2}),\\
&\qquad 3\le l\le j\le k-1,\\
\sum_{i=l}^{k} R_i + R_{k,2} &\le \sum_{i=l-1}^{k} I(U_i;Y_i|U_{i-1}) - I(U_k;Y_{l-2}|U_{l-2}),\\
&\qquad \text{for } 3\le l\le k+1. \qquad (41)
\end{aligned}$$

We next show that the region $\mathcal{R}_k$ takes the structure (41) 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 $k+1$:

$$\begin{aligned}
R_{k+1,1} + R_{k+1,2} &\le I(U_{k+1};Y_{k+1}|U_k),\\
\sum_{i=j}^{k+1} R_i + R_{k+1,2} &\le \sum_{i=j-1}^{k+1} I(U_i;Y_i|U_{i-1})\\
&\quad - I(U_{k+1};Y_{j-2}|U_{j-2}), \text{ for } 3\le j\le k+2,\\
R_{k+1} &= R_{k,2} + R_{k+1,1}. \qquad (42)
\end{aligned}$$

Then the resulting region, following standard steps of Fourier-Motzkin elimination to eliminate $R_{k,2}$ and $R_{k+1,1}$, equals (41) for $k+1$.

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

Appendix C
Converse Proof of Theorem 1

We note that the converse proof is based on the weak secrecy requirement, which is necessarily valid under the strong secrecy requirement. Such a converse proof also implies that the secrecy capacity region under the weak and strong secrecy requirements are the same.

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

$$H(W_k|Y_k^n) \le n\epsilon_n, \quad \text{for } 1\le k\le K, \qquad (43)$$

$$I(W_k,\ldots,W_K;Y_{k-2}^n) \le \epsilon_n \le n\epsilon_n, \quad \text{for } 3\le k\le K, \qquad (44)$$

both of which implies that for $3\le k\le K$,

$$I(W_k,\ldots,W_K;Y_{k-2}^n|W_1,\ldots,W_{k-2}) \le n\epsilon_n\ . \qquad (45)$$

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 $3\le k\le K$. 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 for $1\le i\le n$:

$$U_{1,i}\to\ldots\to U_{K,i}\to X_i\to Y_{K,i}\to\ldots\to Y_{1,i}. \qquad (46)$$

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:

$$\begin{aligned}
nR_1 &= H(W_1)\\
&= I(W_1;Y_1^n) + H(W_1|Y_1^n)\\
&\overset{(a)}{\le} I(W_1;Y_1^n) + n\epsilon_n\\
&= \sum_{i=1}^{n} I(W_1;Y_{1i}|Y_1^{i-1}) + n\epsilon_n\\
&\le \sum_{i=1}^{n} I(W_1,Y_1^{i-1};Y_{1i}) + n\epsilon_n\\
&= \sum_{i=1}^{n} I(U_{1,i};Y_{1,i}) + n\epsilon_n, \qquad (47)
\end{aligned}$$

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

We further bound the rate R_2 as follows:

$$\begin{aligned}
nR_2 &= H(W_2) = H(W_2|W_1)\\
&= I(W_2;Y_2^n|W_1) + H(W_2|Y_2^n,W_1)\\
&\overset{(a)}{\le} I(W_2;Y_2^n|W_1) + n\epsilon_n\\
&= \sum_{i=1}^{n} I(W_2;Y_{2,i}|W_1,Y_2^{i-1}) + n\epsilon_n\\
&\overset{(b)}{\le} \sum_{i=1}^{n} I(W_1,W_2,Y_2^{i-1};Y_{2,i}|W_1,Y_1^{i-1}) + n\epsilon_n\\
&= \sum_{i=1}^{n} I(U_{2,i};Y_{2,i}|U_{1,i}) + n\epsilon_n, \qquad (48)
\end{aligned}$$

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

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

$$\begin{aligned}
n\sum_{j=2}^{k} R_j &= H(W_2, \ldots, W_k) \\
&\overset{(a)}{=} H(W_2|W_1) + H(W_3|W_1, W_2) \\
&\quad + \ldots + H(W_k|W_1, \ldots, W_{k-1}) \\
&\overset{(b)}{\le} I(W_2; Y_2^n|W_1) + I(W_3; Y_3^n|W_1, W_2) \\
&\quad + \ldots + I(W_k; Y_k^n|W_1, \ldots, W_{k-1}) + n(k-1)\epsilon_n \\
&= \sum_{i=1}^{n} I(W_2; Y_{2,i}|W_1, Y_2^{i-1}) + I(W_3; Y_{3,i}|W_1, W_2, Y_3^{i-1}) \\
&\quad + \ldots + I(W_k; Y_{k,i}|W_1, \ldots, W_{k-1}, Y_k^{i-1}) + n(k-1)\epsilon_n \\
&= n(k-1)\epsilon_n + \sum_{i=1}^{n} \Bigg(I(W_2, Y_2^{i-1}; Y_{2,i}|W_1, Y_1^{i-1}) \\
&\quad - I(Y_2^{i-1}; Y_{2,i}|W_1, Y_1^{i-1}) \\
&\quad + I(W_3, Y_3^{i-1}, Y_{1,i+1}^n; Y_{3,i}|W_1, W_2, Y_2^{i-1}) \\
&\quad - I(Y_3^{i-1}; Y_{3,i}|W_1, W_2, Y_2^{i-1}) \\
&\quad - I(Y_{1,i+1}^n; Y_{3,i}|W_1, W_2, W_3, Y_3^{i-1}) \\
&\quad + \sum_{j=4}^{k} \Bigg(I\big(W_j, Y_j^{i-1}, Y_{j-2,i+1}^n; Y_{j,i}|W_1, \ldots, W_{j-1}, \\
&\quad Y_{j-3,i+1}^n, Y_{j-1}^{i-1}\big) + I(Y_{j-3,i+1}^n; Y_{j,i}|W_1, \ldots, W_{j-1}, Y_j^{i-1}) \\
&\quad - I(Y_j^{i-1}; Y_{j,i}|W_1, \ldots, W_{j-1}, Y_{j-3,i+1}^n, Y_{j-1}^{i-1}) \\
&\quad - I(Y_{j-2,i+1}^n; Y_{j,i}|W_1, \ldots, W_j, Y_j^{i-1}) \Bigg) \Bigg) \\
&\overset{(c)}{\le} n(k-1)\epsilon_n + \sum_{j=2}^{k}\sum_{i=1}^{n} I(U_{j,i}; Y_{j,i}|U_{j-1,i}), \qquad (49)
\end{aligned}$$

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

$$-I(Y_2^{i-1}; Y_{2,i}|W_1, Y_1^{i-1}) \le 0, \qquad (50)$$

$$-I(Y_3^{i-1}; Y_{3,i}|W_1, W_2, Y_2^{i-1}) \le 0, \qquad (51)$$

$$-I(Y_{k-2,i+1}^n; Y_{k,i}|W_1, \ldots, W_k, Y_k^{i-1}) \le 0, \qquad (52)$$

and the following inequalities:

$$\begin{aligned}
&-I(Y_{j-2,i+1}^n; Y_{j,i}|W_1, \ldots, W_j, Y_j^{i-1}) \\
&\quad + I(Y_{j-2,i+1}^n; Y_{j+1,i}|W_1, \ldots, W_j, Y_{j+1}^{i-1}) \\
&\quad - I(Y_{j+1}^{i-1}; Y_{j+1,i}|W_1, \ldots, W_j, Y_{j-2,i+1}^n, Y_j^{i-1}) \\
&\overset{(a)}{=} -I(Y_j^{i-1}; Y_{j-2,i}|W_1, \ldots, W_j, Y_{j-2,i+1}^n) \\
&\quad + I(Y_{j+1}^{i-1}; Y_{j-2,i}|W_1, \ldots, W_j, Y_{j-2,i+1}^n) \\
&\quad - I(Y_{j+1}^{i-1}; Y_{j+1,i}|W_1, \ldots, W_j, Y_{j-2,i+1}^n, Y_j^{i-1}) \\
&\overset{(b)}{=} I(Y_{j+1}^{i-1}; Y_{j-2,i}|W_1, \ldots, W_j, Y_{j-2,i+1}^n, Y_j^{i-1}) \\
&\quad - I(Y_{j+1}^{i-1}; Y_{j+1,i}|W_1, \ldots, W_j, Y_{j-2,i+1}^n, Y_j^{i-1}) \\
&\overset{(c)}{=} -I(Y_{j+1}^{i-1}; Y_{j+1,i}|W_1, \ldots, W_j, Y_{j-2,i+1}^n, Y_j^{i-1}, Y_{j-2,i}) \\
&\le 0, \qquad (53)
\end{aligned}$$

where (a) is due to Csiszár's sum identity property [11], 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 $3 \le l \le k \le K$, which correspond to the secrecy constraints:

$$\begin{aligned}
\sum_{j=l}^{k} nR_j &= H(W_l, \ldots, W_k) + H(W_{l-1}) - H(W_{l-1}) \\
&\overset{(a)}{\le} \sum_{j=l-1}^{k} H(W_j) - H(W_{l-1}) + n\epsilon_n \\
&\quad - I(W_l \ldots, W_k; Y_{l-2}^n|W_1, \ldots, W_{l-2}) \\
&\overset{(b)}{\le} \sum_{j=l-1}^{k} H(W_j) + n\epsilon_n \\
&\quad - I(W_{l-1} \ldots, W_k; Y_{l-2}^n|W_1, \ldots, W_{l-2}) \qquad (54)
\end{aligned}$$

where (a) is due to the secrecy requirement (45) and the independence of the messages, and (b) is due to the fact that

$$\begin{aligned}
&-H(W_{l-1}) - I(W_l \ldots, W_k; Y_{l-2}^n|W_1, \ldots, W_{l-2}) \\
&\quad = -H(W_{l-1}) - H(W_l \ldots, W_k|W_1, \ldots, W_{l-2}) \\
&\qquad + H(W_l \ldots, W_k|Y_{l-2}^n, W_1, \ldots, W_{l-2}) \\
&\quad = -H(W_{l-1} \ldots, W_k|W_1, \ldots, W_{l-2}) \\
&\qquad + H(W_l \ldots, W_k|Y_{l-2}^n, W_1, \ldots, W_{l-2}) \\
&\quad \le -H(W_{l-1} \ldots, W_k|W_1, \ldots, W_{l-2}) \\
&\qquad + H(W_{l-1}, W_l \ldots, W_k|Y_{l-2}^n, W_1, \ldots, W_{l-2}) \\
&\quad = -I(W_{l-1} \ldots, W_k; Y_{l-2}^n|W_1, \ldots, W_{l-2}). \qquad (55)
\end{aligned}$$

We next bound each term in (54) one by one. We first bound $H(W_j)$ for $l \le j \le k$ as shown in (56) on next page, where (a) is due to the independence of the messages and the Fano's inequality (45), (b) is due to Csiszár sum identity property, (c) is due to the degradedness condition (1) and the fact that

$$\begin{aligned}
&I(Y_j^{i-1}; Y_{j-3,i}|W_1, \ldots, W_{j-1}, Y_{j-3,i+1}^n) \\
&\quad = I(Y_j^{i-1}; Y_{j-3,i}|W_1, \ldots, W_{j-1}, Y_{j-3,i+1}^n, Y_{j-1}^{i-1}) \\
&\qquad + I(Y_{j-1}^{i-1}; Y_{j-3,i}|W_1, \ldots, W_{j-1}, Y_{j-3,i+1}^n), \qquad (57)
\end{aligned}$$

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

$$\begin{aligned}
&-I(Y_j^{i-1}; Y_{j,i}|W_1, \ldots, W_{j-1}, Y_{j-1}^{i-1}, Y_{j-3,i+1}^n) \\
&\qquad + I(Y_j^{i-1}; Y_{j-3,i}|W_1, \ldots, W_{j-1}, Y_{j-3,i+1}^n, Y_{j-1}^{i-1}) \\
&\quad = -I(Y_j^{i-1}; Y_{j,i}|W_1, \ldots, W_{j-1}, Y_{j-1}^{i-1}, Y_{j-3,i+1}^n, Y_{j-3,i}) \\
&\quad \le 0, \qquad (58)
\end{aligned}$$

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

Following the intermediate step in (56), $H(W_j)$ is also upper bounded as follows:

$$\begin{aligned}
H(W_j) \le n\epsilon_n + \sum_{i=1}^{n} \Big(&I(U_{j,i}; Y_{j,i}|U_{j-1,i}) \\
&- I(Y_j^{i-1}; Y_{j,i}|W_1, \ldots, W_{j-1}, Y_{j-1}^{i-1}, Y_{j-3,i+1}^n) \\
&+ I(Y_{j-3,i+1}^n; Y_{j,i}|W_1, \ldots, W_{j-1}, Y_j^{i-1}) \\
&- I(Y_{j-2,i+1}^n; Y_{j,i}|W_1, \ldots, W_j, Y_j^{i-1}) \Big). \qquad (59)
\end{aligned}$$

$$
\begin{aligned}
H(W_j) &\overset{(a)}{\leq} H(W_j|W_1,\ldots,W_{j-1}) + n\epsilon_n - H(W_j|Y_j^n, W_1,\ldots,W_{j-1}) \\
&= I(W_j; Y_j^n|W_1,\ldots,W_{j-1}) + n\epsilon_n \\
&= n\epsilon_n + \sum_{i=1}^{n} I(W_j; Y_{j,i}|W_1,\ldots,W_{j-1}, Y_j^{i-1}) \\
&= \sum_{i=1}^{n} \Big(I(W_j, Y_j^{i-1}, Y_{j-2,i+1}^n; Y_{j,i}|W_1,\ldots,W_{j-1}, Y_{j-1}^{i-1}, Y_{j-3,i+1}^n) - I(Y_j^{i-1}; Y_{j,i}|W_1,\ldots,W_{j-1}, Y_{j-1}^{i-1}, Y_{j-3,i+1}^n) \\
&\quad + I(Y_{j-3,i+1}^n; Y_{j,i}|W_1,\ldots,W_{j-1}, Y_j^{i-1}) - I(Y_{j-2,i+1}^n; Y_{j,i}|W_1,\ldots,W_j, Y_j^{i-1}) \Big) + n\epsilon_n \\
&= n\epsilon_n + \sum_{i=1}^{n} \Big(I(U_{j,i}; Y_{j,i}|U_{j-1,i}) - I(Y_j^{i-1}; Y_{j,i}|W_1,\ldots,W_{j-1}, Y_{j-1}^{i-1}, Y_{j-3,i+1}^n) \\
&\quad + I(Y_{j-3,i+1}^n; Y_{j,i}|W_1,\ldots,W_{j-1}, Y_j^{i-1}) - I(Y_{j-2,i+1}^n; Y_{j,i}|W_1,\ldots,W_j, Y_j^{i-1}) \Big) \\
&\overset{(b)}{=} n\epsilon_n + \sum_{i=1}^{n} \Big(I(U_{j,i}; Y_{j,i}|U_{j-1,i}) - I(Y_j^{i-1}; Y_{j,i}|W_1,\ldots,W_{j-1}, Y_{j-1}^{i-1}, Y_{j-3,i+1}^n) \\
&\quad + I(Y_j^{i-1}; Y_{j-3,i}|W_1,\ldots,W_{j-1}, Y_{j-3,i+1}^n) - I(Y_{j-2,i+1}^n; Y_{j,i}|W_1,\ldots,W_j, Y_j^{i-1}) \Big) \\
&\overset{(c)}{=} \sum_{i=1}^{n} \Big(I(U_{j,i}; Y_{j,i}|U_{j-1,i}) - I(Y_j^{i-1}; Y_{j,i}|W_1,\ldots,W_{j-1}, Y_{j-1}^{i-1}, Y_{j-3,i+1}^n) + I(Y_{j-1}^{i-1}; Y_{j-3,i}|W_1,\ldots,W_{j-1}, Y_{j-3,i+1}^n) \\
&\quad + I(Y_j^{i-1}; Y_{j-3,i}|W_1,\ldots,W_{j-1}, Y_{j-3,i+1}^n, Y_{j-1}^{i-1}) - I(Y_{j-2,i+1}^n; Y_{j,i}|W_1,\ldots,W_j, Y_j^{i-1}) \Big) + n\epsilon_n \\
&\overset{(d)}{\leq} n\epsilon_n + \sum_{i=1}^{n} \Big(I(U_{j,i}; Y_{j,i}|U_{j-1,i}) + I(Y_{j-1}^{i-1}; Y_{j-3,i}|W_1,\ldots,W_{j-1}, Y_{j-3,i+1}^n) - I(Y_{j-2,i+1}^n; Y_{j,i}|W_1,\ldots,W_j, Y_j^{i-1}) \Big) \\
&\overset{(e)}{=} n\epsilon_n + \sum_{i=1}^{n} \Big(I(U_{j,i}; Y_{j,i}|U_{j-1,i}) + I(Y_{j-3,i+1}^n; Y_{j-1,i}|W_1,\ldots,W_{j-1}, Y_{j-1}^{i-1}) - I(Y_{j-2,i+1}^n; Y_{j,i}|W_1,\ldots,W_j, Y_j^{i-1}) \Big)
\end{aligned} \tag{56}
$$

Hence, substituting (56) for $l \leq j \leq k$, and (59) for $j = l-1$ into the first term in (54), we obtain,

$$
\begin{aligned}
&\sum_{j=l-1}^{k} H(W_j) \\
&\quad \leq n(k-l+2)\epsilon_n + \sum_{i=1}^{n}\sum_{j=l-1}^{k} I(U_{j,i}; Y_{j,i}|U_{j-1,i}) \\
&\qquad + I(Y_{l-4,i+1}^n; Y_{l-1,i}|W_1,\ldots,W_{l-2}, Y_{l-1}^{i-1}) \\
&\qquad - I(Y_{l-1}^{i-1}; Y_{l-1,i}|W_1,\ldots,W_{l-2}, Y_{l-2}^{i-1}, Y_{l-4,i+1}^n) \\
&\qquad - I(Y_{k-2,i+1}^n; Y_{k,i}|W_1,\ldots,W_k, Y_k^{i-1}).
\end{aligned} \tag{60}
$$

We then bound the third term in (54) for $3 \leq l \leq k \leq K$ as shown in (61) on next page, where (a) is due to the following fact:

$$
\begin{aligned}
&\sum_{i=1}^{n} -I(Y_{l-2}^{i-1}; Y_{l-2,i}|W_1,\ldots,W_{l-2}, Y_{l-4,i+1}^n) \\
&\qquad + I(Y_{l-2,i+1}^n; Y_{l-2,i}|W_1,\ldots,W_{l-2}, Y_{l-4,i+1}^n) \\
&\quad = \sum_{i=1}^{n} H(Y_{l-2,i}|W_1,\ldots,W_{l-2}, Y_{l-2}^{i-1}, Y_{l-4,i+1}^n) \\
&\qquad - H(Y_{l-2,i}|W_1,\ldots,W_{l-2}, Y_{l-2}^{i-1}) \\
&\quad = \sum_{i=1}^{n} -I(Y_{l-4,i+1}^n; Y_{l-2,i}|W_1,\ldots,W_{l-2}, Y_{l-2}^{i-1}).
\end{aligned} \tag{62}
$$

Substituting (60) and (61) into (54), we obtain

$$
\begin{aligned}
&n\sum_{j=l}^{k} R_j \\
&\quad \leq \sum_{j=l-1}^{k} H(W_j) + n\epsilon_n \\
&\qquad - I(W_{l-1}\ldots, W_k; Y_{l-2}^n|W_1,\ldots,W_{l-2}) \\
&\quad \leq n(k-l+3)\epsilon_n + \sum_{i=1}^{n} \Bigg(\Big(\sum_{j=l-1}^{k} I(U_{j,i}; Y_{j,i}|U_{j-1,i}) \Big) \\
&\qquad - I(U_{k,i}; Y_{l-2,i}|U_{l-2,i}) \\
&\qquad + I(Y_{l-4,i+1}^n; Y_{l-1,i}|W_1,\ldots,W_{l-2}, Y_{l-1}^{i-1}) \\
&\qquad - I(Y_{l-1}^{i-1}; Y_{l-1,i}|W_1,\ldots,W_{l-2}, Y_{l-2}^{i-1}, Y_{l-4,i+1}^n) \\
&\qquad - I(Y_{k-2,i+1}^n; Y_{k,i}|W_1,\ldots,W_k, Y_k^{i-1}) \\
&\qquad - I(Y_{l-4,i+1}^n; Y_{l-2,i}|W_1,\ldots,W_{l-2}, Y_{l-2}^{i-1}) \\
&\qquad + I(Y_k^{i-1}, Y_{k-2,i+1}^n; Y_{l-2,i}|W_1,\ldots,W_k, Y_{l-2,i+1}^n) \Bigg), \\
&\quad \overset{(a)}{\leq} n(k-l+3)\epsilon_n + \sum_{i=1}^{n} \Bigg(\Big(\sum_{j=l-1}^{k} I(U_{j,i}; Y_{j,i}|U_{j-1,i}) \Big) \\
&\qquad - I(U_{k,i}; Y_{l-2,i}|U_{l-2,i}) \Bigg),
\end{aligned} \tag{63}
$$

$$
\begin{aligned}
&-I(W_{l-1}\ldots,W_k;Y_{l-2}^n|W_1,\ldots,W_{l-2})\\
&=\sum_{i=1}^{n}-I(W_{l-1}\ldots,W_k;Y_{l-2,i}|W_1,\ldots,W_{l-2},Y_{l-2,i+1}^n)\\
&=\sum_{i=1}^{n}-I(W_{l-1}\ldots,W_k,Y_k^{i-1};Y_{l-2,i}|W_1,\ldots,W_{l-2},Y_{l-2,i+1}^n)\\
&\quad+I(Y_k^{i-1};Y_{l-2,i}|W_1,\ldots,W_k,Y_{l-2,i+1}^n)\\
&=\sum_{i=1}^{n}-I(W_{l-1}\ldots,W_k,Y_k^{i-1},Y_{l-2,i+1}^n;Y_{l-2,i}|W_1,\ldots,W_{l-2},Y_{l-4,i+1}^n)\\
&\quad+I(Y_{l-2,i+1}^n;Y_{l-2,i}|W_1,\ldots,W_{l-2},Y_{l-4,i+1}^n)+I(Y_k^{i-1};Y_{l-2,i}|W_1,\ldots,W_k,Y_{l-2,i+1}^n)\\
&=\sum_{i=1}^{n}-I(W_{l-1}\ldots,W_k,Y_k^{i-1},Y_{k-2,i+1}^n;Y_{l-2,i}|W_1,\ldots,W_{l-2},Y_{l-2}^{i-1},Y_{l-4,i+1}^n)\\
&\quad+I(Y_{k-2,i+1}^n;Y_{l-2,i}|W_1,\ldots,W_k,Y_k^{i-1},Y_{l-2,i+1}^n)-I(Y_{l-2}^{i-1};Y_{l-2,i}|W_1,\ldots,W_{l-2},Y_{l-4,i+1}^n)\\
&\quad+I(Y_{l-2,i+1}^n;Y_{l-2,i}|W_1,\ldots,W_{l-2},Y_{l-4,i+1}^n)+I(Y_k^{i-1};Y_{l-2,i}|W_1,\ldots,W_k,Y_{l-2,i+1}^n)\\
&=\sum_{i=1}^{n}-I(U_{k,i};Y_{l-2,i}|U_{l-2,i})\\
&\quad+I(Y_{k-2,i+1}^n;Y_{l-2,i}|W_1,\ldots,W_k,Y_k^{i-1},Y_{l-2,i+1}^n)-I(Y_{l-2}^{i-1};Y_{l-2,i}|W_1,\ldots,W_{l-2},Y_{l-4,i+1}^n)\\
&\quad+I(Y_{l-2,i+1}^n;Y_{l-2,i}|W_1,\ldots,W_{l-2},Y_{l-4,i+1}^n)+I(Y_k^{i-1};Y_{l-2,i}|W_1,\ldots,W_k,Y_{l-2,i+1}^n)\\
&\overset{(a)}{=}\sum_{i=1}^{n}-I(U_{k,i};Y_{l-2,i}|U_{l-2,i})\\
&\quad-I(Y_{l-4,i+1}^n;Y_{l-2,i}|W_1,\ldots,W_{l-2},Y_{l-2}^{i-1})+I(Y_k^{i-1},Y_{k-2,i+1}^n;Y_{l-2,i}|W_1,\ldots,W_k,Y_{l-2,i+1}^n)
\end{aligned}
\tag{61}
$$

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

$$
\begin{aligned}
&\sum_{i=1}^{n}I(Y_k^{i-1},Y_{k-2,i+1}^n;Y_{l-2,i}|W_1,\ldots,W_k,Y_{l-2,i+1}^n)\\
&\qquad-I(Y_{k-2,i+1}^n;Y_{k,i}|W_1,\ldots,W_k,Y_k^{i-1})\\
&=\sum_{i=1}^{n}H(Y_{l-2,i}|W_1,\ldots,W_k,Y_{l-2,i+1}^n)\\
&\qquad-H(Y_{l-2,i}|W_1,\ldots,W_k,Y_k^{i-1},Y_{k-2,i+1}^n)\\
&\qquad-H(Y_{k,i}|W_1,\ldots,W_k,Y_k^{i-1})\\
&\qquad+H(Y_{k,i}|W_1,\ldots,W_k,Y_k^{i-1},Y_{k-2,i+1}^n)\\
&=H(Y_{l-2}^n|W_1,\ldots,W_k)-H(Y_k^n|W_1,\ldots,W_k)\\
&\qquad+\sum_{i=1}^{n}H(Y_{k,i}|W_1,\ldots,W_k,Y_k^{i-1},Y_{k-2,i+1}^n,Y_{l-2,i})\\
&=-H(Y_k^n|W_1,\ldots,W_k,Y_{l-2}^n)\\
&\qquad+\sum_{i=1}^{n}H(Y_{k,i}|W_1,\ldots,W_k,Y_k^{i-1},Y_{k-2,i+1}^n,Y_{l-2,i})\\
&=\sum_{i=1}^{n}-H(Y_{k,i}|W_1,\ldots,W_k,Y_{l-2}^n,Y_k^{i-1})\\
&\qquad+H(Y_{k,i}|W_1,\ldots,W_k,Y_k^{i-1},Y_{k-2,i+1}^n,Y_{l-2,i})\\
&\le 0,
\end{aligned}
\tag{64}
$$

where the last inequality is due to the Markov chain condition in (1). The second fact is shown as follows:

$$
\begin{aligned}
&\sum_{i=1}^{n}I(Y_{l-4,i+1}^n;Y_{l-1,i}|W_1,\ldots,W_{l-2},Y_{l-1}^{i-1})\\
&\qquad-I(Y_{l-1}^{i-1};Y_{l-1,i}|W_1,\ldots,W_{l-2},Y_{l-2}^{i-1},Y_{l-4,i+1}^n)\\
&\qquad-I(Y_{l-4,i+1}^n;Y_{l-2,i}|W_1,\ldots,W_{l-2},Y_{l-2}^{i-1})\\
&=\sum_{i=1}^{n}H(Y_{l-1,i}|W_1,\ldots,W_{l-2},Y_{l-1}^{i-1})\\
&\qquad-H(Y_{l-1,i}|W_1,\ldots,W_{l-2},Y_{l-1}^{i-1},Y_{l-4,i+1}^n)\\
&\qquad-H(Y_{l-1,i}|W_1,\ldots,W_{l-2},Y_{l-2}^{i-1},Y_{l-4,i+1}^n)\\
&\qquad+H(Y_{l-1,i}|W_1,\ldots,W_{l-2},Y_{l-4,i+1}^n,Y_{l-1}^{i-1})\\
&\qquad-H(Y_{l-2,i}|W_1,\ldots,W_{l-2},Y_{l-2}^{i-1})\\
&\qquad+H(Y_{l-2,i}|W_1,\ldots,W_{l-2},Y_{l-2}^{i-1},Y_{l-4,i+1}^n)\\
&=H(Y_{l-1}^n|W_1,\ldots,W_{l-2})-H(Y_{l-2}^n|W_1,\ldots,W_{l-2})\\
&\qquad-\sum_{i=1}^{n}H(Y_{l-1,i}|W_1,\ldots,W_{l-2},Y_{l-4,i+1}^n,Y_{l-2}^{i-1},Y_{l-2,i})\\
&=H(Y_{l-1}^n|W_1,\ldots,W_{l-2},Y_{l-2}^n)\\
&\qquad-\sum_{i=1}^{n}H(Y_{l-1,i}|W_1,\ldots,W_{l-2},Y_{l-4,i+1}^n,Y_{l-2}^{i-1},Y_{l-2,i})\\
&=\sum_{i=1}^{n}H(Y_{l-1,i}|W_1,\ldots,W_{l-2},Y_{l-2}^n,Y_{l-1}^{i-1})\\
&\qquad-H(Y_{l-1,i}|W_1,\ldots,W_{l-2},Y_{l-4,i+1}^n,Y_{l-2}^{i-1},Y_{l-2,i})\\
&\le 0.
\end{aligned}
\tag{65}
$$

Furthermore, based on (49), we bound $\sum_{j=2}^{K} R_j$ as follows:

$$\begin{aligned} n\sum_{j=2}^{K} R_j &\le n(k-1)\epsilon_n + \sum_{j=2}^{K}\sum_{i=1}^{n} I(U_{j,i}; Y_{j,i}|U_{j-1,i}) \\ &\le n(k-1)\epsilon_n + \sum_{j=2}^{K-1}\sum_{i=1}^{n} I(U_{j,i}; Y_{j,i}|U_{j-1,i}) \\ &\quad + \sum_{i=1}^{n} I(X_i; Y_{K,i}|U_{K-1,i}). \end{aligned} \tag{66}$$

Based on (63), we bound $\sum_{j=l}^{K} R_j$ as follows:

$$\begin{aligned} & n\sum_{j=l}^{K} R_j \\ &\le n(K-l+3)\epsilon_n + \sum_{i=1}^{n}\Bigg(\Bigg(\sum_{j=l-1}^{K} I(U_{j,i}; Y_{j,i}|U_{j-1,i})\Bigg) \\ &\quad - I(U_{K,i}; Y_{l-2,i}|U_{l-2,i})\Bigg) \\ &= n(K-l+3)\epsilon_n + \sum_{i=1}^{n}\Bigg(\Bigg(\sum_{j=l-1}^{K-1} I(U_{j,i}; Y_{j,i}|U_{j-1,i})\Bigg) \\ &\quad + I(U_{K,i}; Y_{K,i}|U_{K-1,i}) - I(U_{K,i}; Y_{l-2,i}|U_{l-2,i})\Bigg) \\ &\overset{(a)}{\le} n(K-l+3)\epsilon_n + \sum_{i=1}^{n}\Bigg(\Bigg(\sum_{j=l-1}^{K-1} I(U_{j,i}; Y_{j,i}|U_{j-1,i})\Bigg) \\ &\quad + I(X_i; Y_{K,i}|U_{K-1,i}) - I(X_i; Y_{l-2,i}|U_{l-2,i})\Bigg), \end{aligned} \tag{67}$$

where (a) is due to the Markov chain condition (46).

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 $X \triangleq (Q, X_Q)$.

ACKNOWLEDGEMENT

The authors would like to thank the associate editor Dr. Matthieu Bloch for his constructive suggestions to obtain strong secrecy. They would also like to thank the anonymous reviewers for their insightful comments, which helped to significantly improve the presentation of the paper.

REFERENCES

[1] S. Zou, Y. Liang, L. Lai, and S. Shamai (Shitz), "Degraded broadcast channel: Secrecy outside of a bounded range," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Jerusalem, Israel, Apr. 2015, pp. 1–5.

[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. Inf. Theory (ISIT)*, Hong Kong, Jun. 2015, pp. 1357–1361.

[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 Inf. Theory Workshop (ITW)*, Cambridge, U.K., Sep. 2016, pp. 31–35.

[4] Y. Liang, H. V. Poor, and S. Shamai (Shitz), "Information theoretic security," *Found. Trends Commun. Inf. Theory*, vol. 5, nos. 4–5, pp. 355–580, 2009.

[5] M. Bloch and J. Barros, *Physical-Layer Security: From Information Theory to Security Engineering*. New York, NY, USA: Cambridge Univ. Press, 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," *Proc. IEEE*, vol. 103, no. 10, pp. 1841–1856, Sep. 2015.

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

[8] H. V. Poor and R. F. Schaefer, "Wireless physical layer security," *Proc. Nat. Acad. Sci. USA*, vol. 114, no. 1, pp. 19–26, Jan. 2017.

[9] A. Hyadi, Z. Rezki, and M.-S. Alouini, "An overview of physical layer security in wireless communication systems with CSIT uncertainty," *IEEE Access*, vol. 4, pp. 6121–6132, 2016.

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

[11] I. Csiszár and J. Korner, "Broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. IT-24, no. 3, pp. 339–348, May 1978.

[12] E. Ekrem and S. Ulukus, "Secrecy capacity of a class of broadcast channels with an eavesdropper," *EURASIP J. Wireless Commun. Netw.*, vol. 2009, pp. 1:1–1:29, Mar. 2009.

[13] R. Liu, T. Liu, H. V. Poor, and S. Shamai (Shitz), "A vector generalization of Costa's entropy-power inequality with applications," *IEEE Trans. Inf. Theory*, vol. 56, no. 4, pp. 1865–1879, Apr. 2010.

[14] E. Ekrem and S. Ulukus, "Degraded compound multi-receiver wiretap channels," *IEEE Trans. Inf. Theory*, vol. 58, no. 9, pp. 5681–5698, Sep. 2012.

[15] S. Zou, Y. Liang, L. Lai, and S. Shamai (Shitz), "An information theoretic approach to secret sharing," *IEEE Trans. Inf. Theory*, vol. 61, no. 6, pp. 3121–3136, Apr. 2015.

[16] S. Zou, Y. Liang, and S. Shamai (Shitz), "Gaussian fading channel with secrecy outside a bounded range," in *Proc. IEEE Conf. Commun. Netw. Secur. (CNS)*, Oct. 2017, pp. 545–549.

[17] T. S. Han and S. Verdú, "Approximation theory of output statistics," *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 752–772, May 993.

[18] M. R. Bloch and J. N. Laneman, "Strong secrecy from channel resolvability," *IEEE Trans. Inf. Theory*, vol. 59, no. 12, pp. 8077–8098, Dec. 2013.

[19] J. Hou and G. Kramer, "Informational divergence approximations to product distributions," in *Proc. Can. Workshop Inf. Theory (CWIT)*, 2013, pp. 76–81.

[20] J. Hou and G. Kramer, "Effective secrecy: Reliability, confusion and stealth," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, 2014, pp. 601–605.

[21] M. Hayashi, "General nonasymptotic and asymptotic formulas in channel resolvability and identification capacity and their application to the wiretap channel," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1562–1575, Apr. 2006.

[22] A. Thangaraj, "Coding for wiretap channels: Channel resolvability and semantic security," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Nov. 2014, pp. 232–236.

[23] R. Bustin, R. F. Schaefer, H. V. Poor, and S. Shamai (Shitz), "An I-MMSE based graphical representation of rate and equivocation for the Gaussian broadcast channel," in *Proc. Conf. Commun. Netw. Secur. (CNS)*, Sep. 2015, pp. 53–58.

[24] D. Guo, S. Shamai (Shitz), and S. Verdú, "Mutual information and minimum mean-square error in Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1261–1282, Apr. 2005.

[25] Y. Liang, L. Lai, H. V. Poor, and S. Shamai (Shitz), "A broadcast approach for fading wiretap channels," *IEEE Trans. Inf. Theory*, vol. 60, no. 2, pp. 842–858, Feb. 2014.

[26] S. Shamai (Shitz) and A. Steiner, "A broadcast approach for a single-user slowly fading MIMO channel," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2617–2635, Oct. 2003.

[27] O. O. Koyluoglu and H. El Gamal, "Polar coding for secure transmission and key agreement," *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 5, pp. 1472–1483, Oct. 2012.

[28] J. del Olmo and J. R. Fonollosa. (Jul. 2016). "Strong secrecy on a class of degraded broadcast channels using polar codes." [Online]. Available: https://arxiv.org/abs/1607.07815

[29] H. Mahdavifar and A. Vardy, "Achieving the secrecy capacity of wiretap channels using polar codes," *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 6428–6443, Oct. 2011.

[30] R. A. Chou and M. R. Bloch, "Polar coding for the broadcast channel with confidential messages: A random binning analogy," *IEEE Trans. Inf. Theory*, vol. 62, no. 5, pp. 2410–2429, May 2016.

[31] T. C. Gulcu and A. Barg, "Achieving secrecy capacity of the wiretap channel and broadcast channel with a confidential component," *IEEE Trans. Inf. Theory*, vol. 63, no. 2, pp. 1311–1324, Feb. 2017.
[32] Y.-P. Wei and S. Ulukus, "Polar coding for the general wiretap channel with extensions to multiuser scenarios," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 2, pp. 278–291, Feb. 2016.
[33] M. Bloch, M. Hayashi, and A. Thangaraj, "Error-control coding for physical-layer secrecy," *Proc. IEEE*, vol. 103, no. 10, pp. 1725–1746, Oct. 2015.
[34] D. Klinc, J. Ha, S. W. McLaughlin, J. Barros, and B.-J. Kwak, "LDPC codes for the Gaussian wiretap channel," *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 3, pp. 532–540, Sep. 2011.
[35] J. L. Massey, *Applied Digital Information Theory*. Zürich, Switzerland: ETH Zurich, 1980.
[36] A. Orlitsky and J. R. Roche, "Coding for computing," *IEEE Trans. Inf. Theory*, vol. 47, no. 3, pp. 903–917, Mar. 2001.
[37] A. El Gamal and Y.-H. Kim, *Network Information Theory*. New York, NY, USA: Cambridge University Press, 2012.

Shaofeng Zou (S'14–M'16) received the Ph.D. degree in Electrical and Computer Engineering from Syracuse University in 2016. He received the B.E. degree (with honors) from Shanghai Jiao Tong University, Shanghai, China, in 2011. Since July 2016, he has been a postdoctoral research associate at University of Illinois at Urbana-Champaign. Dr. Zou's research interests include information theory, machine learning, and statistical signal processing. He received the National Scholarship from the Ministry of Education of China in 2008. He also received the award of the outstanding graduate of Shanghai in 2011.

Yingbin Liang (S'01–M'05–SM'16) received the Ph.D. degree in Electrical Engineering from the University of Illinois at Urbana-Champaign in 2005. In 2005–2007, she was working as a postdoctoral research associate at Princeton University. In 2008–2009, she was an assistant professor at University of Hawaii. In 2010–2017, she was an assistant and then an associate professor at Syracuse University. Since August 2017, she has been an associate professor at the Department of Electrical and Computer Engineering at the Ohio State University. Dr. Liang's research interests include machine learning, statistical signal processing, optimization, information theory, and wireless communication and networks.

Dr. Liang was a Vodafone Fellow at the University of Illinois at Urbana-Champaign during 2003-2005, and received the Vodafone-U.S. Foundation Fellows Initiative Research Merit Award in 2005. She also received the M. E. Van Valkenburg Graduate Research Award from the ECE department, University of Illinois at Urbana-Champaign, in 2005. In 2009, she received the National Science Foundation CAREER Award, and the State of Hawaii Governor Innovation Award. In 2014, she received EURASIP Best Paper Award for the *EURASIP Journal on Wireless Communications and Networking*. She served as an Associate Editor for the Shannon Theory of the IEEE TRANSACTIONS ON INFORMATION THEORY during 2013–2015.

Lifeng Lai (M'07) received the B.E. and M. E. degrees from Zhejiang University, Hangzhou, China in 2001 and 2004 respectively, and the PhD degree from The Ohio State University at Columbus, OH, in 2007. He was a postdoctoral research associate at Princeton University from 2007 to 2009, an assistant professor at University of Arkansas, Little Rock from 2009 to 2012, and an assistant professor at Worcester Polytechnic Institute. Since 2016, he has been an associate professor at University of California, Davis. Dr. Lais research interests include information theory, stochastic signal processing and their applications in wireless communications, security and other related areas.

Dr. Lai was a Distinguished University Fellow of the Ohio State University from 2004 to 2007. He is a co-recipient of the Best Paper Award from IEEE Global Communications Conference (Globecom) in 2008, the Best Paper Award from IEEE Conference on Communications (ICC) in 2011 and the Best Paper Award from IEEE Smart Grid Communications (SmartGridComm) in 2012. He received the National Science Foundation CAREER Award in 2011, and Northrop Young Researcher Award in 2012. He served as a Guest Editor for IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, Special Issue on Signal Processing Techniques for Wireless Physical Layer Security. He is currently serving as an Editor for IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, and an Associate Editor for IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY.

H. Vincent Poor (S'72–M'77–SM'82–F'87) received the Ph.D. degree in electrical engineering and computer science from Princeton University in 1977. From 1977 until 1990, he was on the faculty of the University of Illinois at Urbana-Champaign. Since 1990 he has been on the faculty at Princeton, where he is the Michael Henry Strater University Professor of Electrical Engineering. During 2006 to 2016, he served as Dean of Princetons School of Engineering and Applied Science. He has also held visiting appointments at several other institutions, most recently at Berkeley and Cambridge. His research interests are in the areas of information theory and signal processing, and their applications in wireless networks, energy systems and related fields. Among his publications in these areas is the recent book *Information Theoretic Security and Privacy of Information Systems* (Cambridge University Press, 2017).

Dr. Poor is a member of the National Academy of Engineering and the National Academy of Sciences, and is a foreign member of the Chinese Academy of Sciences, the Royal Society, and other national and international academies. In 1990, he served as President of the IEEE Information Theory Society, in 2004–07 as the Editor-in-Chief of these TRANSACTIONS, and in 2009 as General Co-chair of the IEEE INTERNATIONAL SYMPOSIUM ON INFORMATION THEORY, held in Seoul, South Korea. Recent recognition of his work includes the 2017 IEEE Alexander Graham Bell Medal, Honorary Professorships from Peking University and Tsinghua University, both conferred in 2017, and a D.Sc. *honoris causa* from Syracuse University awarded in 2017.

Shlomo Shamai (Shitz) (S'80–M'82–SM'88–F'94) received the B.Sc., M.Sc., and Ph.D. degrees in electrical engineering from the Technion—Israel Institute of Technology, in 1975, 1981 and 1986 respectively.

During 1975–1985 he was with the Communications Research Labs, in the capacity of a Senior Research Engineer. Since 1986 he is with the Department of Electrical Engineering, Technion—Israel Institute of Technology, where he is now a Technion Distinguished Professor, and holds the William Fondiller Chair of Telecommunications. His research interests encompasses a wide spectrum of topics in information theory and statistical communications.

Dr. Shamai (Shitz) is an IEEE Fellow, an URSI Fellow, a member of the Israeli Academy of Sciences and Humanities and a foreign member of the US National Academy of Engineering. He is the recipient of the 2011 Claude E. Shannon Award, the 2014 Rothschild Prize in Mathematics/Computer Sciences and Engineering and the 2017 IEEE Richard W. Hamming Medal.

He has been awarded the 1999 van der Pol Gold Medal of the Union Radio Scientifique Internationale (URSI), and is a co-recipient of the 2000 IEEE Donald G. Fink Prize Paper Award, the 2003, and the 2004 joint IT/COM societies paper award, the 2007 IEEE Information Theory Society Paper Award, the 2009 and 2015 European Commission FP7, Network of Excellence in Wireless COMmunications (NEWCOM++, NEWCOM#) Best Paper Awards, the 2010 Thomson Reuters Award for International Excellence in Scientific Research, the 2014 EURASIP Best Paper Award (for the *EURASIP Journal on Wireless Communications and Networking*), and the 2015 IEEE Communications Society Best Tutorial Paper Award. He is also the recipient of 1985 Alon Grant for distinguished young scientists and the 2000 Technion Henry Taub Prize for Excellence in Research. He has served as Associate Editor for the Shannon Theory of the IEEE TRANSACTIONS ON INFORMATION THEORY, and has also served twice on the Board of Governors of the Information Theory Society. He has also served on the Executive Editorial Board of the IEEE TRANSACTIONS ON INFORMATION THEORY and on the IEEE Information Theory Society Nominations and Appointments Committee.