

An Information Theoretic Approach to Secret Sharing

Shaofeng Zou, *Student Member, IEEE*, Yingbin Liang, *Member, IEEE*, Lifeng Lai, *Member, IEEE*, and Shlomo Shamai (Shitz), *Fellow, IEEE*

Abstract—A novel information theoretic approach is proposed to solve the secret sharing problem, in which a dealer distributes one or multiple secrets among a set of participants in such a manner that for each secret only qualified sets of users can recover this secret by pooling their shares together while nonqualified sets of users obtain no information about the secret even if they pool their shares together. While existing secret sharing systems (implicitly) assume that communications between the dealer and participants are noiseless, this paper takes a more practical assumption that the dealer delivers shares to the participants via a noisy broadcast channel. Thus, in contrast to the existing solutions that are mainly based on number theoretic tools, an information theoretic approach is proposed, which exploits the channel randomness during delivery of shares as additional resources to achieve secret sharing requirements. In this way, secret sharing problems can be reformulated as equivalent secure communication problems via wiretap channel models, and can hence be solved by employing the powerful information theoretic security techniques. This approach is first developed for the classic secret sharing problem, in which only one secret is to be shared. This classic problem is shown to be equivalent to a communication problem over a compound wiretap channel. Thus, the lower and upper bounds on the secrecy capacity of the compound channel provide the corresponding bounds on the secret sharing rate, and the secrecy scheme designed for the compound channel provides the secret sharing schemes. The power of the approach is further demonstrated by a more general layered multisecret sharing problem, which is shown to be equivalent to the degraded broadcast multiple-input multiple-output (MIMO) channel with layered decoding and secrecy constraints. The secrecy capacity region for the degraded

MIMO broadcast channel is characterized, which provides the secret sharing capacity region. Furthermore, the secure encoding scheme that achieves the secrecy capacity region provides an information theoretic scheme for sharing the secrets.

Index Terms—Broadcast channel, layered secrecy, MIMO channel, secrecy capacity region, secret sharing, wiretap channel.

I. INTRODUCTION

IN the classic secret sharing problem, a dealer intends to distribute a secret among a set of participants such that only qualified sets of participants can correctly recover the secret by pooling their shares together, while the non-qualified set of participants obtain no information about the secret even if they pool their shares together. There are rich applications of secret sharing including construction of protocols and algorithms for secure multiparty computations [4], [5], Byzantine agreement [6], threshold cryptography [7], access control [8], attribute-based encryption [9], and generalized oblivious transfer [10]. The existing solutions for the secret sharing problems are mainly based on the number theoretic tools, in which contents of the shares that the dealer delivers to the participants are specially designed in order to guarantee the secret sharing requirements. While such approaches work well for simple secret sharing problems, they are not readily extendable to more complicated problems, in which qualified and non-qualified sets become more complicated, and/or multiple secrets are simultaneously shared.

While in existing secret sharing systems, it is implicitly assumed that communications between the dealer and participants are noiseless, in this paper we take a more practical assumption that the dealer delivers shares to the participants via a noisy broadcast channel. Thus, we propose a novel information theoretic approach to solving secret sharing problems, which exploits the channel randomness during delivery of shares from dealers to participants as additional resources to achieve secret sharing requirements. In this way, secret sharing problems can be equivalently reformulated into secure communication problems via wiretap channel models studied in information theory [11], [12] (see [13], [14] for more references of these studies). More importantly, such an approach is general enough to incorporate complex secret sharing requirements with multiple secrets and various structures of qualified and non-qualified sets into wiretap models. These secure communication problems can then be solved by employing powerful information theoretic security techniques, which thus enable the design of secret sharing strategies. Furthermore, such an approach also allows

Manuscript received March 18, 2014; revised February 17, 2015; accepted March 30, 2015. Date of publication April 10, 2015; date of current version May 15, 2015. S. Zou and Y. Liang were supported by the National Science Foundation (NSF) CAREER Award under Grant CCF-10-26565 and by NSF under Grants CCF-10-26566 and CNS-11-16932. L. Lai was supported by NSF CAREER Award under Grant CCF-13-18980 and by NSF under Grant CNS-13-21223. S. Shamai was supported by the Israel Science Foundation, and by the European Commission in the framework of the Network of Excellence in Wireless Communications NEWCOM++. This paper was presented in part at the IEEE International Symposium on Information Theory (ISIT) in 2011 [1], in part at the IEEE 14th Workshop on Signal Processing Advances in Wireless Communications (SPAWC) in 2013 [2], and in part at the IEEE International Symposium on Information Theory (ISIT) in 2014 [3].

S. Zou and Y. Liang are with the Department of Electrical Engineering and Computer Science, Syracuse University, Syracuse, NY 13244 USA (e-mail: szou02@syr.edu; yliang06@syr.edu).

L. Lai is with the Department of Electrical and Computer Engineering, Worcester Polytechnic Institute, Worcester, MA 01609 USA (e-mail: llai@wpi.edu).

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

Communicated by T. Liu, 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.2015.2421905

one to characterize the secret sharing capacity region based on the information theoretic characterization of the secrecy capacity region of physical layer wiretap models.

We first illustrate the basic idea of our approach using the classic secret sharing problem, in which the dealer wishes to distribute one secret to qualified sets of participants specified by an arbitrary access structure. We propose to achieve the secret sharing property via broadcasting the secret message from the dealer to all participants. To design a secret sharing scheme, we construct an equivalent compound wiretap channel [15] by creating one virtual legitimate receiver for each qualified set (with the receiver's output including channel outputs from all participants in the qualified set) and creating one virtual eavesdropper for each non-qualified set (with the eavesdropper's output including channel outputs from all participants in the non-qualified set). Thus, a secure communication scheme for the equivalent compound wiretap channel guarantees that the qualified sets of participants can decode the secret while the non-qualified sets of participants obtain a negligible amount of information about the secret. By applying the results for the compound wiretap channel in [15], we obtain a secret sharing scheme and characterize lower and upper bounds on the secret sharing rate. We also characterize the secret sharing capacity for some secret sharing scenarios.

We then demonstrate the power of our approach via a more complicated multi-secret sharing problem. Although we here solve an example problem, our goal is to demonstrate that this approach can be applied to more general secret sharing problems, not limited to the one that we present in this paper. We consider the following multi-secret sharing problem (see Fig. 4), in which multiple secrets are intended for different corresponding qualified sets to recover in a layered fashion. More specifically, a dealer equipped with multiple antennas wishes to distribute K secrets to K participants by broadcasting via multiple antennas over a wireless channel. It is required that participant 1 recover the first secret, and as one more participant joins the group to share its output, one more secret should be recovered by the group, and this new secret should be kept secure from the smaller groups. Hence, if the first k of the K participants share their channel outputs, they can recover the first k secrets, while all the remaining secrets are kept confidential to the group of the first k participants.

The above problem involves sharing multiple secrets in a layered fashion, and can be very challenging to solve using the traditional number theoretic tools. To solve this problem, following the approach developed above, we design a virtual receiver for each qualified set (i.e., sharing group) of receivers, and thus convert this secret sharing problem to a communication problem over the degraded Gaussian multiple-input multiple-output (MIMO) broadcast channel with layered decoding and secrecy constraints. The requirements of the secret sharing problem is exactly mapped into the layered decoding and secrecy requirements for the communication problem. More specifically, in the communication model (see Fig. 3 for an illustration), the transmitter wishes to transmit K messages to K (virtual) receivers. The channel

outputs at receivers naturally satisfy the degradedness condition, i.e., from receiver K to receiver 1, the quality of their channels gets worse gradually. It is required that receiver k decodes one more message than receiver $k - 1$ for $k = 2, \dots, K$, and this additional message should be kept secure from all receivers with worse channel outputs, i.e., with lower indices. Design of secure communication schemes and characterization of the secrecy capacity region for such a channel model readily provide secret sharing schemes and the corresponding secret sharing rate regions.

Characterizing the secrecy capacity region of the degraded Gaussian MIMO broadcast channel with layered decoding and secrecy constraints is a challenging information theoretic problem. Previously, there have been a number of broadcast models with layered decoding and secrecy requirements proposed and studied. In particular, [16] studied a model (model 1) with two legitimate receivers and one eavesdropper. It is required that one message be decoded at both receivers and kept secure from the eavesdropper, and that the second message be decoded at one receiver and kept secure from the other receiver and the eavesdropper. [16] studied one more model (model 2), in which the second message does not need to be kept secure from the other receiver. Both models were further generalized in [17] in that each receiver and the eavesdropper in the above model was replaced by a group of nodes. In [16] and [17], the secrecy capacity region was established for the MIMO Gaussian channels.

The model that we study in this paper generalizes model 1 in [16] to K receivers. We characterize the secrecy capacity region for the discrete memoryless channel and the single-input-single-output (SISO) and MIMO Gaussian channels with layered decoding and secrecy constraints. Towards this end, the challenges lie in both the achievability and converse proofs due to the layered secrecy constraints on more than two receivers. Our achievable scheme is based on the stochastic encoding (i.e., binning) and heterogeneous superposition. The major challenge of achievability arises in the analysis of leakage rates, which is much more involved than the cases with two secure messages. Our contribution here lies in novel generalization of the analysis of the leakage rate provided in [18] for one secure message to multiple secure messages. On the other hand, due to the layered secrecy constraints, outer bounds on the secrecy rates should be developed in certain recursive structures for three or more consecutive layers of receivers. Consequently, techniques used in [16] and [17] for two layers cannot be readily applied here, although some properties on matrix manipulations are useful in our proof for the MIMO channel. Our main technical development in the converse proof lies in the construction of a series of covariance matrices representing input resources for layered messages such that the secrecy rates can be upper bounded as the desired recursive forms in terms of these covariance matrices.

Due to the equivalence of the multi-secret sharing problem via the multiple-input-single-output (MISO) broadcast channel and the secure communication problem over the degraded MIMO broadcast channel, we hence establish the secret

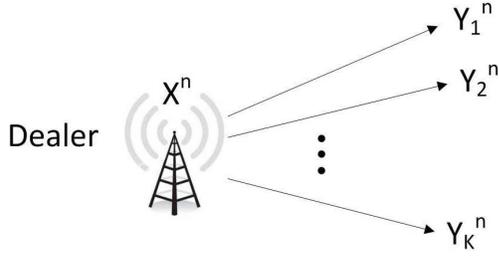


Fig. 1. The noisy broadcast channel.

sharing capacity region. Furthermore, the secure encoding scheme that achieves the secret capacity region provides an information theoretic scheme for sharing multiple secrets.

The rest of the paper is organized as follows. In Section II, we introduce the classical secret sharing problem and its connection to an [17], [19] information theoretic secrecy model. We then use this secret sharing problem to illustrate our information theoretic approach for secret sharing. As further demonstration of our approach, in Sections III and IV, we study an information theoretic model of the degraded broadcast channel with layered decoding and secrecy constraints, and apply the results obtained to study a layered multi-secret sharing problem. Finally, in Section V, we conclude our paper with further remarks.

II. SECRET SHARING AND INFORMATION THEORETIC SECRECY

In this section, we first introduce the secret sharing problem, and then connect this problem to a model studied in information theoretic secrecy.

A. Model of Secret Sharing

We consider the following secret sharing problem. Suppose the system consists of a dealer and a set of participants $\mathcal{P} = \{1, 2, \dots, K\}$. The dealer has a secret W (taken from a set \mathcal{W}) for the K participants to share. We define an access structure \mathcal{A} , which contains all subsets of \mathcal{P} that are required to recover the secret. Each set $A \in \mathcal{A}$ is called a qualified set. We assume that the access structure considered in this paper is monotone [20], that is if $A \in \mathcal{A}$ and $A \subseteq A_1$, then $A_1 \in \mathcal{A}$. For the secret sharing scheme, we require that if the users in any qualified set $A \in \mathcal{A}$ gather their observations together, then they can recover the secret with a negligible error probability. We define a non-access structure \mathcal{B} such that for any set $B \notin \mathcal{B}$, we require that even if users in the set B gather their observations together, they obtain negligible information about the secret message. In many applications, $\mathcal{B} = \mathcal{A}^C$.

In the existing secret sharing schemes, the communications between the dealer and participants are assumed to be noiseless as the classic secret sharing problem does not involve channel. In this paper, we assume that the dealer and the participants are connected by a noisy broadcast channel, as shown in Figure 1. If the dealer transmits X^n , participant k receives Y_k^n , and the relationship among the input and outputs is characterized by the transition probability distribution

$$P_{Y_1^n \dots Y_K^n | X^n}(y_1^n \dots y_K^n | x^n) = \prod_{k=1}^K \prod_{i=1}^n P_{Y_k | X}(y_k(i) | x(i)), \quad (1)$$

where $x(i)$ is taken from a finite set \mathcal{X} , and $y_k(i)$ is taken from a finite set \mathcal{Y}_k for $k = 1, \dots, K$.

Definition 1: A $(2^{nR}, n)$ code for secret sharing over the noisy broadcast channel consists of the following:

- 1) a secret set: $\mathcal{W} = \{1, 2, \dots, 2^{nR}\}$ with the secret W uniformly distributed over \mathcal{W} ;
- 2) an encoder $f: \mathcal{W} \rightarrow \mathcal{X}^n$ mapping each secret message $w \in \mathcal{W}$ to a codeword $x^n \in \mathcal{X}^n$;
- 3) a decoder g_A for each qualified set $A \in \mathcal{A}$: $\{\mathcal{Y}_k^n : k \in A\} \rightarrow \mathcal{W}$.

The average block error probability for set A is

$$P_{e,A}^n = \frac{1}{2^{nR}} \sum_{w=1}^{2^{nR}} \Pr\{\hat{w}^A \neq w | w \text{ was sent}\}, \quad (2)$$

where \hat{w}^A is the secret recovered by the decoder g_A .

Definition 2: A secret sharing rate R is said to be achievable if there exists a sequence of $(2^{nR}, n)$ codes such that for any $\epsilon > 0$ the following two conditions are satisfied:

- 1) $\forall A \in \mathcal{A}$, we have

$$P_{e,A}^n \leq \epsilon; \quad (3)$$

- 2) $\forall B \in \mathcal{B}$, we have

$$\frac{1}{n} I(W; \{Y_k^n : k \in B\}) \leq \epsilon. \quad (4)$$

The above condition (3) requires that the decoding error probability at any qualified set should be small, while condition (4) requires that any non-qualified set gains negligible amount of information about the secret even if users in this set share their observations together.

The *secret sharing capacity* is defined to be the maximal achievable secret sharing rate.

B. Connection to Wiretap Channels

In the remaining part of this section, we study the above secret sharing problem with non-access structure $\mathcal{B} = \mathcal{A}^C$. Our main idea is to connect this problem to a communication problem over an equivalent compound wiretap channel [15] (i.e., the Wyner's wiretap channel model with multiple legitimate receivers and multiple eavesdroppers). Thus, secure coding design for the equivalent compound wiretap channel can be applied to achieve secret sharing.

More specifically, for each set $A \in \mathcal{A}$, we construct a virtual legitimate receiver V_A such that the observation at the virtual receiver is $Y_A = \{Y_k : k \in A\}$. Clearly, this receiver is not an actual node, but an identity representing that the users in set A share their outputs. For an access structure \mathcal{A} , we will construct $|\mathcal{A}|$ virtual legitimate receivers, representing that in each of these sharing scenarios, the secret message is required to be recovered. In addition, for each set $B \in \mathcal{B}$, we construct a virtual wiretapper $V_{E,B}$ with the observation $Y_B = \{Y_k : k \in B\}$. These virtual eavesdroppers represent that in each of these sharing scenarios, the message should be kept secure. We note that these virtual eavesdroppers are also not actual devices, but the identities representing that nodes in non-qualified sets share their outputs.

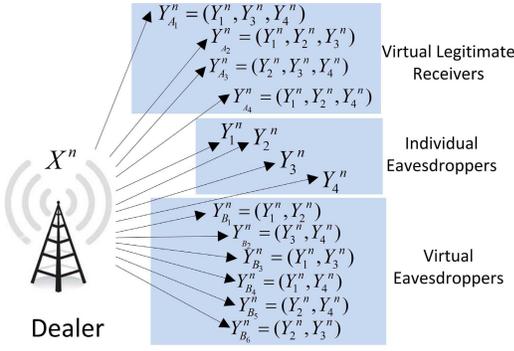


Fig. 2. An equivalent compound wiretap channel for a secret sharing with four participants.

We note that if $A_1 \subset A_2$, then $X \rightarrow Y_{A_2} \rightarrow Y_{A_1}$. In this case, if we design a code such that the users in the qualified set A_1 can decode the secret, then the users in the qualified set A_2 can also decode the secret. Hence, it is not necessary to construct a virtual legitimate receiver for the qualified set A_2 . In this way, the number of virtual legitimate receivers can be reduced in the constructed equivalent compound wiretap channel. Similarly, if $B_1 \subset B_2$, then it is not necessary to construct a virtual wiretapper for the set B_1 , since if the message is kept secure from the set B_2 , then it is also kept secret from the set B_1 . Hence, we can also reduce the number of constructed virtual wiretappers.

Figure 2 illustrates an equivalent compound wiretap channel for a secret sharing system with four participants. Any three participants are required to recover the secret, and hence the access structure \mathcal{A} includes four qualified sets $\{1, 2, 3\}$, $\{1, 2, 4\}$, $\{1, 3, 4\}$, $\{2, 3, 4\}$. We note that the set $\{1, 2, 3, 4\}$ is not included as a virtual legitimate receiver due to the reason mentioned above. Furthermore, any two participants should not recover the secret, and hence six virtual eavesdroppers are created corresponding to these non-qualified sets $\{1, 2\}$, $\{1, 3\}$, $\{1, 4\}$, $\{2, 3\}$, $\{2, 4\}$, $\{3, 4\}$.

Secure coding schemes for the compound wiretap channel have been proposed in [15], which guarantee that all legitimate receivers recover the message and none of the wiretappers obtain any information about the message. By applying the approach developed in [15] to the equivalent compound wiretap channel that we construct for the secret sharing problem, we obtain corresponding schemes such that each virtual legitimate receiver can decode the message W while each virtual wiretapper has negligible information about the message W , which are exactly the conditions required by secret sharing. The following bounds on the secret sharing capacity follows from the bounds on the secrecy capacity in [15].

Theorem 1: The following secret sharing rate is achievable via the broadcast channel described in (1)

$$R_N = \max_{P_{UX}} \left[\min_A I(U; Y_A) - \max_B I(U; Y_B) \right], \quad (5)$$

where U is an auxiliary random variable that satisfies the Markov chain relationship:

$$U \rightarrow X \rightarrow (Y_1, \dots, Y_K). \quad (6)$$

Furthermore, the following rate is an upper-bound on the secret sharing capacity

$$R = \min_{A, B} \max_{P_{UX}} [I(U; Y_A) - I(U; Y_B)]. \quad (7)$$

In particular, the main difference of the inner and outer bounds (5) and (7) lies in that the bound (5) is characterized as a max-min optimization problem, but the bound (7) is characterized as a min-max optimization problem. In general, for the two bounds to match, the objective function needs to satisfy the saddle-point property. Although the conditions can be derived via cardinality bound on U , parametrization of P_{UX} , and expression of mutual information terms, such conditions have very involved forms, and do not provide immediate insights.

However, there are some interesting special cases of the secret sharing problem, which naturally correspond to degraded compound wiretap channels. For these cases, the lower bound (5) matches with the upper bound (7), and hence the secrecy sharing capacity can be characterized.

For example, we consider the scenario, in which the secret is decodable only if all participants share their outputs, and is kept secret from any subset of participants. The secret sharing capacity is given as follows.

Corollary 1: Consider the K -participant secret sharing problem over the noisy channel $P_{Y_1, \dots, Y_K | X}$. Suppose \mathcal{A} contains a single set $A = \{1, \dots, K\}$. The secret sharing capacity is given by

$$C_s = \max_{P_X} \min_B [I(X; Y_1, \dots, Y_K) - I(X; Y_B)]. \quad (8)$$

where the set B can be any strict subset of $A = \{1, \dots, K\}$.

For the two-participant secret sharing problem, the secret sharing capacity is given by

$$C_{two} = \max_{P_X} [I(X; Y_1, Y_2) - \max\{I(X; Y_1), I(X; Y_2)\}]. \quad (9)$$

Another example for which we can fully characterize the secret sharing capacity is the secret sharing problem over Gaussian broadcast channel, in which

$$Y_k = X + Z_k \quad \text{for } k = 1, \dots, K. \quad (10)$$

where Z_k , for $k = 1, \dots, K$ are independent Gaussian noise variables with mean zero and variance N_k . The dealer has an average power constraint:

$$\frac{1}{n} \sum_{i=1}^n X^2(i) \leq P. \quad (11)$$

For the Gaussian channel example, one can obtain an analytical form of the achievable secret sharing rate for an arbitrary access structure \mathcal{A} by setting U in (5) to be Gaussian random variable jointly distributed with X . However, the obtained rate takes a complex form and may not provide any insight. A simpler and often-encountered scenario is the so-called (t, K) -secret sharing, in which any t or more users can recover the secret by sharing their observations, while any $t - 1$ users obtain only negligible information from their joint observations. This secret sharing problem can be reformulated into the secure communication problem over a

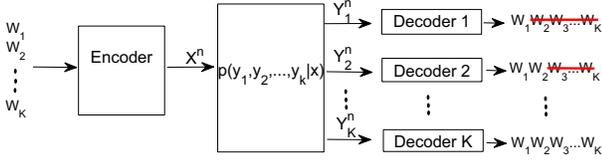


Fig. 3. The broadcast channel with layered decoding and secrecy.

compound wiretap channel, which consists of one transmitter with single antenna, multiple virtual legitimate receivers with each having t antennas (i.e., each virtual receiver corresponds to a group of t users), and multiple virtual eavesdroppers with each having $t - 1$ antennas (i.e., each virtual eavesdropper corresponds to a group of $t - 1$ users). For such a single-input multiple-output (SIMO) compound wiretap channel, the secrecy capacity region can be obtained because the same Gaussian input maximizes the secrecy capacity for each pair of legitimate receiver and eavesdropper, and hence achieves the secrecy capacity of the compound channel. Therefore, for (t, K) -secret sharing, we obtain the secret sharing capacity given below.

Corollary 2: For the (t, K) -secret sharing problem, the secret sharing capacity is given by

$$C_{t,K} = \min_{A_t, A_{t-1}} \log \left(\frac{1 + \sum_{l \in A_t} P/N_l}{1 + \sum_{l \in A_{t-1}} P/N_l} \right). \quad (12)$$

where A_t can be any subset of $\{1, \dots, K\}$ with t indices, and A_{t-1} can be any subset of $\{1, \dots, K\}$ with $t - 1$ indices.

For the special case when the channels to all receivers are symmetric, i.e., $N_k = 1$ for all $k = 1, \dots, K$, the secret sharing capacity is given by

$$C_{t,K} = \log \left(\frac{1 + tP}{1 + (t-1)P} \right).$$

Remark 1: It is interesting to note that for the special case with symmetric channels, the secret sharing capacity depends only on t but not on K which is the total number of participants.

III. BROADCAST CHANNEL WITH LAYERED DECODING AND SECRECY

In order to study a more general secret sharing problem in which simultaneously sharing multiple secrets is required as we introduce in Section IV, we need to study a broadcast wiretap model with layered decoding and secrecy. In this section, we first introduce the system model for this channel, and then we provide our characterization of the secrecy capacity region for this channel. These results will then be applied to the layered multiple secrets sharing problem in Section IV.

A. System Model

In this section, we consider the model of the degraded broadcast channel with layered decoding and secrecy constraints (see Fig. 3), in which a transmitter transmits to K receivers. The channel transition probability function is given by $P_{Y_1 \dots Y_K | X}$, in which $X \in \mathcal{X}$ is the channel input and

$Y_k \in \mathcal{Y}_k$ is the channel output of receiver k for $k = 1, \dots, K$. It is assumed that the receivers have degraded outputs, i.e., Y_1, \dots, Y_K satisfy the following Markov chain condition, i.e., the degradedness condition:

$$X \rightarrow Y_K \rightarrow Y_{K-1} \rightarrow \dots \rightarrow Y_2 \rightarrow Y_1. \quad (13)$$

Hence, the quality of channels gradually degrades from receiver K to receiver 1. The transmitter has K messages W_1, \dots, W_K intended for the K receivers. The system is required to satisfy the following layered decoding and secrecy constraints. For $k = 1, \dots, K$, receiver k needs to decode the messages W_1, \dots, W_k , and to be kept ignorant of messages W_{k+1}, \dots, W_K (see Fig. 3 for an illustration). We note that the model presented here is well motivated by application of secret sharing problems to be discussed in Section IV.

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

- K message sets: $W_k \in \mathcal{W}_k = \{1, \dots, 2^{nR_k}\}$ for $k = 1, \dots, K$, which are independent from each other and each message is uniformly distributed over the corresponding message set;
- an (possibly stochastic) encoder $f^n: \mathcal{W}_1 \times \dots \times \mathcal{W}_K \rightarrow \mathcal{X}^n$;
- K decoders $g_k^n: \mathcal{Y}_k^n \rightarrow (\mathcal{W}_1 \times \dots \times \mathcal{W}_k)$ for $k = 1, \dots, K$.

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

$$P_e^n = \Pr \left(\bigcup_{k=1}^K \{(W_1, \dots, W_k) \neq g_k^n(Y_k^n)\} \right) \quad (14)$$

and the leakage rate at each receiver k for $k = 1, \dots, K$

$$\frac{1}{n} I(W_{k+1}, \dots, W_K; Y_k^n) \quad (15)$$

approach zero as n goes to infinity.

Here, the asymptotically small error probability as in (14) implies that each receiver k is able to decode messages W_1, \dots, W_k , and asymptotically small leakage rate as in (15) for each receiver k implies that receiver k is kept ignorant of messages W_{k+1}, \dots, W_K . Our goal is to characterize the secrecy capacity region that consists of all achievable rate tuples.

We also consider the K -receiver degraded Gaussian broadcast channel, in which

$$Y_k = X + Z_k, \quad k = 1, \dots, K, \quad (16)$$

where Z_k is a zero mean Gaussian noise variable with variance N_k at receiver k . We assume that $0 < N_K < N_{K-1} < \dots < N_1$. The transmitter has an average power constraint P . Since the decoding and secrecy requirements in (14) and (15) only depend on the marginal conditional distribution of the channel output at each receiver given the channel input, not on the joint distribution of the channel outputs given the channel input, changing the correlation of those noise variables has no effects on the secrecy capacity region. Hence, we can adjust the correlation of those noise variables such that the channel outputs at each receiver and the channel input satisfy the same Markov chain as shown in (13).

We further consider the K -receiver degraded Gaussian MIMO broadcast channel. The received signal at receiver k for one channel use is given by

$$\mathbf{Y}_k = \mathbf{X} + \mathbf{Z}_k, \quad k = 1, \dots, K, \quad (17)$$

where the channel input \mathbf{X} , the channel output \mathbf{Y}_k and the noise \mathbf{Z}_k are r -dimensional vectors. Furthermore, the noise variables \mathbf{Z}_k are zero-mean Gaussian random vectors with covariance matrices $\mathbf{\Sigma}_k$ for $k = 1, \dots, K$ that satisfy the following order:

$$\mathbf{0} < \mathbf{\Sigma}_K \preceq \mathbf{\Sigma}_{K-1} \preceq \dots \preceq \mathbf{\Sigma}_1. \quad (18)$$

The channel input \mathbf{X} is subject to a covariance constraint

$$E[\mathbf{X}\mathbf{X}^\top] \preceq \mathbf{S} \quad (19)$$

where $\mathbf{S} \succ \mathbf{0}$. Since the secrecy capacity region does not depend on the correlation across the channel outputs, we can adjust the correlation between the noise vectors such that the channel inputs and channel outputs satisfy the following Markov chain:

$$\mathbf{X} \rightarrow \mathbf{Y}_K \rightarrow \mathbf{Y}_{K-1} \rightarrow \dots \rightarrow \mathbf{Y}_2 \rightarrow \mathbf{Y}_1. \quad (20)$$

Hence, the quality of channels gradually degrades from receiver K to receiver 1.

B. Characterization of Secrecy Capacity Region

In this subsection, we characterize the secrecy capacity region for the model presented in Section III-A. The following theorem characterized the secrecy capacity region of the discrete memoryless channel.

Theorem 1: The secrecy capacity region of the degraded broadcast channel with layered decoding and secrecy constraints as described in Section III-A contains rate tuples (R_1, \dots, R_K) satisfying

$$\begin{aligned} R_1 &\leq I(U_1; Y_1), \\ R_k &\leq I(U_k; Y_k | U_{k-1}) - I(U_k; Y_{k-1} | U_{k-1}), \\ &\quad \text{for } k = 2, \dots, K-1, \\ R_K &\leq I(X; Y_K | U_{K-1}) - I(X; Y_{K-1} | U_{K-1}), \end{aligned} \quad (21)$$

for some $P_{U_1 U_2 \dots U_{K-1} X}$ such that the following Markov chain holds

$$U_1 \rightarrow U_2 \rightarrow \dots \rightarrow U_{K-1} \rightarrow X \rightarrow Y_K \rightarrow \dots \rightarrow Y_1. \quad (22)$$

Remark 2: By setting $R_1 = 0$ and $K = 3$, Theorem 1 reduces to the results for scenario 2 in [17] with each group having a single user for the model in [21], and for the example in [16].

Proof: The proof of the achievability and the proof of converse are provided in Appendices A and C, respectively. \square

We here briefly introduce the idea of the achievable scheme, which is based on the stochastic encoding (i.e., random binning) and superposition coding. For each message, we design one layer of codebook. This codebook contains codewords that are divided into a number of bins, where the bin number contains the information of the corresponding message. The receivers that are required to decode the message

can tell which bin the codeword is in with a small probability of error, while other receivers (i.e., those with worse channel quality) are kept ignorant of this message. These layers of codebooks are superposed together via superposition coding. The major challenge of the achievability proof arises in the analysis of leakage rates, which is much more involved than the cases with two secure messages studied in [16] and [17]. In our proof, we develop novel generalization of the analysis provided in [18] for the case with one secure message to multiple secure messages. More details are provided in Appendix A.

We next characterize the secrecy capacity region of the degraded Gaussian broadcast channel with layered decoding and secrecy constraints. We note that although the Gaussian channel is a special case of the MIMO channel, we present the result for the Gaussian channel here, because the result for the Gaussian channel is simpler, and hence it is easier to follow the converse proof for this case. This helps the understanding of the more complicated proof of the converse for the MIMO channel.

Theorem 2: The secrecy capacity region of a K -user Gaussian broadcast channel with layered decoding and secrecy constraints as described in Section III-A contains rate tuples (R_1, R_2, \dots, R_K) satisfying

$$\begin{aligned} R_1 &\leq \frac{1}{2} \log \left(\frac{N_1 + \sum_{j=1}^K P_j}{N_1 + \sum_{j=2}^K P_j} \right) \\ R_k &\leq \frac{1}{2} \log \left(\frac{N_k + \sum_{j=k}^K P_j}{N_k + \sum_{j=k+1}^K P_j} \right) \\ &\quad - \frac{1}{2} \log \left(\frac{N_{k-1} + \sum_{j=k}^K P_j}{N_{k-1} + \sum_{j=k+1}^K P_j} \right), \quad \text{for } 2 \leq k \leq K, \end{aligned} \quad (23)$$

for some nonnegative variables P_1, P_2, \dots, P_K such that $\sum_{k=1}^K P_k \leq P$.

Proof: The achievability is based on Theorem 1 by setting (U_1, \dots, U_K, X) to be jointly Gaussian distributed random variables with $U_k \sim \mathcal{N}(0, \sum_{j=1}^k P_j)$.

The converse proof is given in Appendix D. \square

We now characterize the secrecy capacity region for the degraded Gaussian MIMO channel with layered decoding and secrecy constraints in the following theorem.

Theorem 3: The secrecy capacity region of the degraded Gaussian MIMO Broadcast channel with layered decoding and secrecy constraints as described in Section III-A contains rate tuples (R_1, \dots, R_K) satisfying the following inequalities:

$$\begin{aligned} R_1 &\leq \frac{1}{2} \log \frac{|\mathbf{\Sigma}_1 + \mathbf{S}|}{|\mathbf{\Sigma}_1 + \mathbf{S}_1|} \\ R_k &\leq \frac{1}{2} \log \frac{|\mathbf{\Sigma}_k + \mathbf{S}_{k-1}|}{|\mathbf{\Sigma}_k + \mathbf{S}_k|} - \frac{1}{2} \log \frac{|\mathbf{\Sigma}_{k-1} + \mathbf{S}_{k-1}|}{|\mathbf{\Sigma}_{k-1} + \mathbf{S}_k|}, \\ &\quad \text{for } 2 \leq k \leq K-1, \\ R_K &\leq \frac{1}{2} \log \frac{|\mathbf{\Sigma}_K + \mathbf{S}_{K-1}|}{|\mathbf{\Sigma}_K|} - \frac{1}{2} \log \frac{|\mathbf{\Sigma}_{K-1} + \mathbf{S}_{K-1}|}{|\mathbf{\Sigma}_{K-1}|}, \end{aligned} \quad (24)$$

for some $\mathbf{0} \preceq \mathbf{S}_{K-1} \preceq \mathbf{S}_{K-2} \preceq \dots \preceq \mathbf{S}_2 \preceq \mathbf{S}_1 \preceq \mathbf{S}$.

We note that if the system has only a single transmit antenna and receive antenna, then the secrecy capacity region in Theorem 3 reduces to that in Theorem 2.

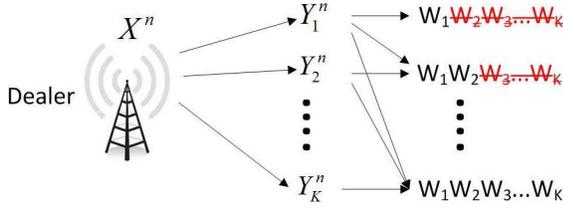


Fig. 4. The model of secret sharing via a broadcast channel.

We further note that due to the layered secrecy constraints, the major challenge in the converse proof for the secrecy capacity region lies in development of upper bounds in certain recursive structures for three or more consecutive layers of receivers. Our contribution here lies in the construction of a sequence of covariance matrices representing input resources for layered messages such that the secrecy rates can be upper bounded as the desired recursive forms in terms of these covariance matrices. More details are provided in Appendix E.

Proof: The achievability of the region (24) follows by choosing the auxiliary random variables $\mathbf{U}_1, \dots, \mathbf{U}_{K-1}, \mathbf{X}$ in (21) to be jointly Gaussian distributed and satisfy the following Markov chain condition:

$$\mathbf{U}_1 \rightarrow \mathbf{U}_2 \rightarrow \dots \rightarrow \mathbf{U}_{K-1} \rightarrow \mathbf{X}, \quad (25)$$

where the covariance of \mathbf{U}_k is set to be $\mathbf{S} - \mathbf{S}_k$ for $k = 1, \dots, K-1$, and the covariance of \mathbf{X} is set to be \mathbf{S} .

The proof of converse is given in Appendix E. \square

IV. APPLICATION TO SHARING MULTIPLE SECRETS

In this section, we apply our result in Section III-B for the degraded MIMO channel to studying the following problem of sharing multiple secrets. Here, a dealer wishes to share K secrets W_1, W_2, \dots, W_K with K participants via a broadcast channel (see Fig. 4). The channel input sent by the dealer is denoted by \mathbf{X} and the channel output received at participant k is denoted by Y_k for $k = 1, \dots, K$. It is required that participant 1 decodes W_1 , and participants 1 and 2 decode W_1 and W_2 by sharing their outputs (Y_1, Y_2) , but W_2 should be kept secure from participant 1. Such requirements extend to k participants for $k = 1, \dots, K$ in the sense that participants 1 to k can recover the first k messages W_1, \dots, W_k by sharing their outputs (Y_1, \dots, Y_k) , but the new message W_k should be secure from the first $k-1$ participants. Hence, as one more participant joins the group, one more secret can be recovered, and this new secret is secure from (and hence cannot be recovered by) a smaller group. The goal is to characterize the secret sharing capacity region, which contains all possible achievable rate tuples (R_1, R_2, \dots, R_K) for K secrets.

This secret sharing problem involves sharing multiple secrets in a layered fashion, and is challenging to solve using the classical approach based on number theory. Here, we solve this problem by constructing an equivalent Gaussian MIMO broadcast model as described in Section III-A.

We assume that the dealer communicates to the participants via a Gaussian MISO broadcast channel corrupted by additive Gaussian noise variables. The dealer has K antennas and each

receiver has one antenna. The relationship of the channel input from the dealer and the channel outputs at all participants is given by

$$\begin{pmatrix} Y_1 \\ \vdots \\ Y_K \end{pmatrix} = \mathbf{H} \begin{pmatrix} X_1 \\ \vdots \\ X_K \end{pmatrix} + \begin{pmatrix} Z_1 \\ \vdots \\ Z_K \end{pmatrix} \quad (26)$$

where \mathbf{H} is the $K \times K$ channel matrix, which is assumed to be invertible, (Y_1, \dots, Y_K) are channel outputs at the K participants, (X_1, \dots, X_K) are the channel inputs from the K antennas of the dealer, and (Z_1, \dots, Z_K) is a random Gaussian vector with the covariance matrix Σ with each entry $\Sigma_{ij} = E[Z_i Z_j] = \sigma_{ij}^2$. We assume that the dealer's input is subject to a resource constraint, $E[\mathbf{X}\mathbf{X}^T] \leq \mathbf{S}$.

We note that it is reasonable to assume that \mathbf{H} is invertible in order to guarantee that each participant's output contains new information compared to other participants so that new secret can be recovered when this participant joins a group.

We reformulate the above secret sharing model into a degraded MIMO broadcast communication system by designing a virtual receiver for each sharing group of participants. More specifically, we design a virtual receiver V_k for the group of the first k participants, i.e., $\mathbf{Y}_{V_k} = (Y_1, \dots, Y_k)$, for $1 \leq k \leq K$. For technical convenience, we add $K-k$ specially designed outputs $\tilde{Y}_{k+1}, \dots, \tilde{Y}_K$ to V_k so that it contains K components, i.e., the virtual receiver V_k has K antennas. The channel outputs at those K antennas are given by,

$$\mathbf{Y}_{V_k} = \begin{pmatrix} Y_1 \\ \vdots \\ Y_k \\ \tilde{Y}_{k+1} \\ \vdots \\ \tilde{Y}_K \end{pmatrix} = \mathbf{H} \begin{pmatrix} X_1 \\ \vdots \\ X_K \end{pmatrix} + \begin{pmatrix} Z_1 \\ \vdots \\ Z_k \\ Z_{k+1} + t\tilde{Z}_{k+1} \\ \vdots \\ Z_K + t\tilde{Z}_K \end{pmatrix} \quad (27)$$

where \tilde{Z}_k , $2 \leq k \leq K$, is random Gaussian noise variables with mean zero and variance $\tilde{\sigma}_{kk}^2 > 0$, and \tilde{Z}_k is independent from all other random variables. Here, t is a large enough constant (i.e., $t \rightarrow \infty$), so that $\tilde{Y}_{k+1}, \dots, \tilde{Y}_K$ are fully corrupted by the noise. We define a new random Gaussian vector $\mathbf{Z}_{V_k} = (Z_1, \dots, Z_k, Z_{k+1} + t\tilde{Z}_{k+1}, \dots, Z_K + t\tilde{Z}_K)^T$ and rewrite (27) as

$$\mathbf{Y}_{V_k} = \mathbf{H}\mathbf{X} + \mathbf{Z}_{V_k}, \quad \text{for } k = 1, \dots, K. \quad (28)$$

Since the channel matrix \mathbf{H} is invertible, we have

$$\mathbf{H}^{-1}\mathbf{Y}_{V_k} = \mathbf{X} + \mathbf{H}^{-1}\mathbf{Z}_{V_k}. \quad (29)$$

By treating $\mathbf{H}^{-1}\mathbf{Y}_{V_k}$ as the new channel output \mathbf{Y}'_{V_k} at virtual receiver V_k , and define a new random Gaussian noise vector $\mathbf{Z}'_{V_k} = \mathbf{H}^{-1}\mathbf{Z}_{V_k}$, we have

$$\mathbf{Y}'_{V_k} = \mathbf{X} + \mathbf{Z}'_{V_k}, \quad (30)$$

which is equivalent to the model in (28).

We now state a lemma that provides the order of the covariance matrices of \mathbf{Z}'_{V_k} , denoted by Σ'_{V_k} , for $1 \leq k \leq K$.

Lemma 1: Let \mathbf{Z}'_{V_k} , $1 \leq k \leq K$, be random Gaussian vectors defined as above. The covariance matrices of \mathbf{Z}'_{V_k} satisfy the following ordering property:

$$\Sigma'_{V_1} \succeq \Sigma'_{V_2} \succeq \dots \succeq \Sigma'_{V_K}. \quad (31)$$

Proof: For any $1 \leq k \leq K - 1$,

$$\mathbf{Z}'_{V_k} = \mathbf{Z}'_{V_{k+1}} + \mathbf{H}^{-1}(0, \dots, 0, t\tilde{\mathbf{Z}}_{k+1}, 0, \dots, 0)^T, \quad (32)$$

where $\mathbf{H}^{-1}(0, \dots, 0, t\tilde{\mathbf{Z}}_{k+1}, 0, \dots, 0)^T$ is a random Gaussian vector, hence the covariance matrices of \mathbf{Z}'_{V_k} satisfy such an order in (31). \square

Therefore, by designing virtual receivers, we reformulate the problem of secret sharing via the MISO broadcast channel into the problem of secure communication over the degraded MIMO broadcast channel described in Section III-A. It can also be seen that the requirements of the secret sharing problem is equivalent to the layered decoding and secrecy requirements for the communication problem. It is also due to the secret sharing requirements and Lemma 1, the degradedness condition in the equivalent MIMO channel naturally holds. Thus, the secret sharing capacity region equals the secrecy capacity region of the degraded MIMO broadcast channel. Thus, applying Theorem 3 we obtain the following secret sharing capacity region.

Corollary 3: The capacity region for the secret sharing problem described above contains rate tuples (R_1, R_2, \dots, R_K) satisfying

$$\begin{aligned} R_1 &\leq \frac{1}{2} \log \frac{|\Sigma'_{V_1} + \mathbf{S}|}{|\Sigma'_{V_1} + \mathbf{S}_1|} \\ R_k &\leq \lim_{t \rightarrow \infty} \frac{1}{2} \log \frac{|\Sigma'_{V_k} + \mathbf{S}_{k-1}|}{|\Sigma'_{V_k} + \mathbf{S}_k|} - \frac{1}{2} \log \frac{|\Sigma'_{V_{k-1}} + \mathbf{S}_{k-1}|}{|\Sigma'_{V_{k-1}} + \mathbf{S}_k|}, \\ &\quad \text{for } 2 \leq k \leq K - 1, \\ R_K &\leq \lim_{t \rightarrow \infty} \frac{1}{2} \log \frac{|\Sigma'_{V_K} + \mathbf{S}_{K-1}|}{|\Sigma'_{V_K}|} - \frac{1}{2} \log \frac{|\Sigma'_{V_{K-1}} + \mathbf{S}_{K-1}|}{|\Sigma'_{V_{K-1}}|}, \end{aligned} \quad (33)$$

for some $\mathbf{0} \leq \mathbf{S}_{K-1} \leq \mathbf{S}_{K-2} \leq \dots \leq \mathbf{S}_2 \leq \mathbf{S}_1 \leq \mathbf{S}$.

V. CONCLUSION

In this paper, we have proposed a novel approach based on information theoretic secrecy to solving secret sharing problems. The basic idea is to reformulate the secret sharing problem into a secure communication problem, and then apply techniques for the latter case to solving the secret sharing problem. In order to illustrate the basic idea, we have first studied the classic problem of sharing one secret among a set of participants, and provided a solution by reformulating the secret sharing system into an equivalent compound wiretap channel with multiple legitimate receivers and multiple eavesdroppers. We have then demonstrated the power of our approach by solving a more complicated problem of sharing multiple secrets with layered sharing requirements. We have characterized the secret sharing capacity region by

reformulating the problem into the problem of the degraded broadcast MIMO channel with layered decoding and secrecy, for which we have characterized the secrecy capacity region. Our approach can be generally applicable to solving various secret sharing problems, which can be difficult using traditional number theoretic tools. For example, various multiple-secret sharing problems can be reformulated into secure communication problems with multiple confidential messages via compound MIMO broadcast channels, and hence the information theoretic techniques and results developed in existing literature, see [22]–[24], can be applied to solving these secret sharing problems.

APPENDIX A

ACHIEVABILITY PROOF OF THEOREM 1

The achievability proof is based on stochastic encoding and superposition coding. We use random codes and fix a joint probability distribution $P_{U_1 \dots U_{K-1} X}$ satisfying the Markov chain condition given in (22). Let $T_\epsilon^n(P_{U_1 \dots U_{K-1} X Y_1 \dots Y_K})$ denote the strongly jointly ϵ -typical set based on the fixed distribution.

Random Codebook Generation: In the following achievability proof, for notational convenience, we write X as U_K , i.e., $P_{U_1 \dots U_{K-1} X} = P_{U_1 \dots U_K}$.

- 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 $u_{k-1}^n(w_1, w_2, l_2, \dots, w_{k-1}, l_{k-1})$, $k = 2, \dots, K$, generate $2^{n\tilde{R}_k}$ i.i.d. sequences u_k^n with distribution $\prod_{i=1}^n p(u_{k,i} | u_{k-1,i})$. Partition these sequences into 2^{nR_k} bins, each with $2^{n(\tilde{R}_k - R_k)}$ sequences. We use $w_k \in [1 : 2^{nR_k}]$ to denote the bin index, and $l_k \in [1 : 2^{n(\tilde{R}_k - R_k)}]$ to denote the index within each bin. Hence each u_k^n is indexed by $(w_1, w_2, l_2, \dots, w_k, l_k)$.

The chosen codebook is revealed to the transmitter and all receivers.

Encoding: To send a message tuple (w_1, w_2, \dots, w_K) , for each $2 \leq k \leq K$, the encoder randomly generate $l_k \in [1 : 2^{n(\tilde{R}_k - R_k)}]$ based on a uniform distribution. The transmitter then sends $u_K^n(w_1, w_2, l_2, \dots, w_K, l_K)$.

Decoding: For $k = 1, \dots, K$, receiver k claims that $(\hat{w}_1, \dots, \hat{w}_k)$ is sent, if there exists a unique tuple $(\hat{w}_1, \hat{w}_2, \hat{l}_2, \dots, \hat{w}_k, \hat{l}_k)$ such that $(u_1^n(\hat{w}_1), u_2^n(\hat{w}_1, \hat{w}_2, \hat{l}_2), \dots, u_k^n(\hat{w}_1, \hat{w}_2, \hat{l}_2, \dots, \hat{w}_k, \hat{l}_k), y_k^n) \in T_\epsilon^n(P_{U_1 \dots U_k Y_k})$. Otherwise, it declares an error.

Analysis of Error Probability: By the law of large numbers and the packing lemma [18], it can be shown that if the following inequalities are satisfied, receiver k (for $k = 1, \dots, K$) can decode messages w_1, w_2, \dots, w_k with a vanishing error probability:

$$\begin{aligned} R_1 &\leq I(U_1; Y_1), \\ \tilde{R}_k &\leq I(U_k; Y_k | U_{k-1}), \quad \text{for } 2 \leq k \leq K. \end{aligned} \quad (34)$$

Analysis of Leakage Rate: We first compute an average of the leakage rate over the random codebook ensemble as follows. For convenience, we let $W^k = (W_1, \dots, W_k)$,

$W_{k+1}^K = (W_{k+1}, \dots, W_K)$, and $L^K = (L_1, \dots, L_K)$. We note that l_k appeared above is realization of the random variable L_k here.

$$\begin{aligned}
 & I(W_{k+1}^K; Y_k^n | \mathcal{C}) \\
 & \stackrel{(a)}{=} I(W^K, L^K; Y_k^n | \mathcal{C}) - I(W^k, L^K; Y_k^n | W_{k+1}^K, \mathcal{C}) \\
 & \stackrel{(b)}{\leq} I(W^K, L^K; Y_k^n | \mathcal{C}) - I(W^k, L^K; Y_k^n | W_{k+1}^K, \mathcal{C}) \\
 & \stackrel{(c)}{\leq} I(U_K^n; Y_k^n | \mathcal{C}) - I(W^k, L^K; Y_k^n | W_{k+1}^K, \mathcal{C}) \\
 & = I(U_K^n; Y_k^n | \mathcal{C}) - H(W^k, L^K | W_{k+1}^K, \mathcal{C}) \\
 & \quad + H(W^k, L^K | Y_k^n, W_{k+1}^K, \mathcal{C}), \tag{35}
 \end{aligned}$$

where step (a) is due to the independence of W^k and W_{k+1}^K , (b) follows from Fano's inequality, step (c) follows from the Markov chain $(W^K, L^K) \rightarrow (U_K^n, \mathcal{C}) \rightarrow Y_k^n$.

We bound the above three terms one by one. For the first term, we have

$$\begin{aligned}
 & I(U_K^n; Y_k^n | \mathcal{C}) \\
 & \stackrel{(a)}{=} I(U_k^n, U_K^n; Y_k^n | \mathcal{C}) \\
 & = I(U_k^n; Y_k^n | \mathcal{C}) + I(U_K^n; Y_k^n | U_k^n, \mathcal{C}) \\
 & \leq H(U_k^n | \mathcal{C}) + I(U_K^n; Y_k^n | U_k^n, \mathcal{C}) \\
 & \leq \sum_{j=1}^k \tilde{R}_j + H(Y_k^n | U_k^n, \mathcal{C}) - H(Y_k^n | U_K^n, U_k^n, \mathcal{C}) \\
 & = n \sum_{j=1}^k \tilde{R}_j + \sum_{j=1}^n H(Y_{k,j} | U_k^n, Y_k^{j-1}, \mathcal{C}) \\
 & \quad - \sum_{j=1}^n H(Y_{k,j} | U_K^n, U_k^n, Y_k^{j-1}, \mathcal{C}) \\
 & \stackrel{(b)}{\leq} n \sum_{j=1}^k \tilde{R}_j + \sum_{j=1}^n H(Y_{k,j} | U_{k,j}) - \sum_{j=1}^n H(Y_{k,j} | U_{K,j}) \\
 & = n \sum_{j=1}^k \tilde{R}_j + nH(Y_k | U_k) - nH(Y_k | U_K) \\
 & = n \sum_{j=1}^k \tilde{R}_j + nI(U_K; Y_k | U_k), \tag{36}
 \end{aligned}$$

where (a) follows from the Markov chain $U_k^n \rightarrow U_K^n \rightarrow Y_k^n$, (b) follows from the fact that $H(Y_{k,j} | U_k^n, Y_k^{j-1}, \mathcal{C}) \leq H(Y_{k,j} | U_{k,j})$ and from the Markov chain $(U_k^n, U_K^{j-1}, U_{K,j+1}^{j-1}, \mathcal{C}) \rightarrow U_{k,j} \rightarrow Y_{k,j}$.

For the second term, due to the independence of W_1, \dots, W_K and L_1, \dots, L_K , we have

$$H(W^k, L^K | W_{k+1}^K, \mathcal{C}) = \sum_{j=1}^k n\tilde{R}_j + \sum_{j=k+1}^K n(\tilde{R}_j - R_j). \tag{37}$$

We now bound the last term as follows.

$$\begin{aligned}
 & H(W^k, L^K | Y_k^n, W_{k+1}^K, \mathcal{C}) \\
 & = H(W^k | Y_k^n, W_{k+1}^K, \mathcal{C}) + H(L^K | Y_k^n, W^K, \mathcal{C}) \\
 & \stackrel{(a)}{\leq} H(L_{k+1}^K | Y_k^n, W^K, L^k, \mathcal{C}) + 2n\epsilon_n \\
 & = \sum_{j=k+1}^K H(L_j | Y_k^n, W^K, L^{j-1}, \mathcal{C}) + 2n\epsilon_n \\
 & \stackrel{(b)}{=} \sum_{j=k+1}^K H(L_j | Y_k^n, W^K, L^{j-1}, U_{j-1}^n, \mathcal{C}) + 2n\epsilon_n \\
 & \leq \sum_{j=k+1}^K H(L_j | Y_k^n, U_{j-1}^n, W_j) + 2n\epsilon_n \\
 & \stackrel{(c)}{\leq} \sum_{j=k+1}^K n(\tilde{R}_j - R_j - I(U_j; Y_k | U_{j-1})) + n\epsilon'_n \\
 & \stackrel{(d)}{=} \sum_{j=k+1}^K n(\tilde{R}_j - R_j) - I(U_K; Y_k | U_k) + n\epsilon'_n, \tag{38}
 \end{aligned}$$

where (a) follows from the chain rule and Fano's inequality, (b) follows from the fact that U_{j-1}^n is a function of $(\mathcal{C}, W^{j-1}, L^{j-1})$, and (c) follows from to Lemma 2 with the condition that $\tilde{R}_j - R_j \geq I(U_j; Y_k | U_{j-1})$, and (d) follows from the Markov chain $U_1 \rightarrow U_2 \rightarrow \dots \rightarrow U_K \rightarrow Y_k$.

Lemma 2: If $\tilde{R}_j - R_j \geq I(U_j; Y_k | U_{j-1})$ for $k+1 \leq j \leq K$, then

$$\frac{1}{n} H(L_j | Y_k^n, U_{j-1}^n, W_j) \leq \tilde{R}_j - R_j - I(U_j; Y_k | U_{j-1}) + \epsilon'_n.$$

Proof: See Appendix B. \square

Combining the analysis of the three terms together, we conclude that as $n \rightarrow \infty$ for $1 \leq k \leq K-1$, $\frac{1}{n} I(W_{k+1}^K; Y_k^n | \mathcal{C}) \rightarrow 0$, if

$$\tilde{R}_k - R_k \geq I(U_k; Y_k | U_{k-1}), \quad \text{for } 2 \leq k \leq K. \tag{39}$$

It is also clear that the sum of the error probability and the leakage rates averaged over the codebook ensemble converges to zero as $n \rightarrow \infty$. Hence, there exists one codebook such that the error probability and the leakage rate converge to zero as $n \rightarrow \infty$.

Combining the bounds in (34) and (39), we obtain that the rate tuple (R_1, \dots, R_K) is achievable if

$$\begin{aligned}
 R_1 & \leq I(U_1; Y_1), \\
 R_k & \leq I(U_k; Y_k | U_{k-1}) - I(U_k; Y_{k-1} | U_{k-1}), \\
 & \quad \text{for } 2 \leq k \leq K. \tag{40}
 \end{aligned}$$

APPENDIX B PROOF OF LEMMA 2

We first bound $\frac{1}{n} H(L_j | Y_k^n, U_{j-1}^n, w_j)$ for any w_j , and hence, $\frac{1}{n} H(L_j | Y_k^n, U_{j-1}^n, W_j)$ is bounded.

Fix $L_j = l_j$ and a joint typical sequence $(u_{j-1}^n, y_k^n) \in T_\epsilon^{(n)}(U_{j-1}, Y_k)$. We define

$$\begin{aligned}
 & N(w_j, l_j, u_{j-1}^n, y_k^n) \\
 & := |\{\tilde{l}_j \neq l_j : (U_j^n(w_j, \tilde{l}_j), u_{j-1}^n, y_k^n) \in T_\epsilon^{(n)}\}|. \tag{41}
 \end{aligned}$$

In fact, $N(w_j, l_j, u_{j-1}^n, y_k^n)$ can be viewed as a binomial distributed random variable, with $2^{n(\tilde{R}_j - R_j)} - 1$ Bernoulli distributed random variables, each taking the value 1 with probability

$$2^{-nI(U_j; Y_k | U_{j-1}) - n\delta_n(\epsilon)} \leq p \leq 2^{-nI(U_j; Y_k | U_{j-1}) + n\delta_n(\epsilon)} \quad (42)$$

It can be shown that the expectation and variance of N satisfy the following inequalities:

$$\begin{aligned} 2^{n(\tilde{R}_j - R_j) - nI(U_j; Y_k | U_{j-1}) - n\delta_n(\epsilon) - n\epsilon_n} \\ \leq E(N(w_j, l_j, u_{j-1}^n, y_k^n)) \\ \leq 2^{n(\tilde{R}_j - R_j) - nI(U_j; Y_k | U_{j-1}) + n\delta_n(\epsilon) - n\epsilon_n}, \end{aligned} \quad (43)$$

$$\begin{aligned} \text{Var}(N(w_j, l_j, u_{j-1}^n, y_k^n)) \\ \leq 2^{n(\tilde{R}_j - R_j) - nI(U_j; Y_k | U_{j-1}) + n\delta_n(\epsilon) - n\epsilon_n}, \end{aligned} \quad (44)$$

where $\delta_n(\epsilon), \epsilon_n \rightarrow 0$ as $n \rightarrow \infty$.

We next define the random event,

$$\begin{aligned} \varepsilon(w_j, l_j, u_{j-1}^n, y_k^n) := \{N(w_j, l_j, u_{j-1}^n, y_k^n) \\ \geq 2^{n(\tilde{R}_j - R_j - I(U_j; Y_k | U_{j-1}) + \delta_n(\epsilon) - \epsilon_n/2) + 1}\}. \end{aligned} \quad (45)$$

Using Chebyshev's inequality, we obtain

$$\begin{aligned} P(\varepsilon(w_j, l_j, u_{j-1}^n, y_k^n)) \\ = P(N(w_j, l_j, u_{j-1}^n, y_k^n) \\ \geq 2^{n(\tilde{R}_j - R_j - I(U_j; Y_k | U_{j-1}) + \delta_n(\epsilon) - \epsilon_n/2) + 1}) \\ \leq P(N(w_j, l_j, u_{j-1}^n, y_k^n) \geq E(N(w_j, l_j, u_{j-1}^n, y_k^n)) \\ + 2^{n(\tilde{R}_j - R_j - I(U_j; Y_k | U_{j-1}) + \delta_n(\epsilon) - \epsilon_n/2)}) \\ \leq (|N(w_j, l_j, u_{j-1}^n, y_k^n) - E(N(w_j, l_j, u_{j-1}^n, y_k^n))| \\ \geq 2^{n(\tilde{R}_j - R_j - I(U_j; Y_k | U_{j-1}) + \delta_n(\epsilon) - \epsilon_n/2)}) \\ \leq \frac{\text{Var}(N(w_j, l_j, u_{j-1}^n, y_k^n))}{2^{2n(\tilde{R}_j - R_j - I(U_j; Y_k | U_{j-1}) + \delta_n(\epsilon) - \epsilon_n/2)}} \\ \leq \frac{1}{2^{2n(\tilde{R}_j - R_j - I(U_j; Y_k | U_{j-1}) + \delta_n(\epsilon))}} \end{aligned} \quad (46)$$

which goes to zero as $n \rightarrow \infty$ if $\tilde{R}_j - R_j \geq I(U_j; Y_k | U_{j-1})$. This implies that

$$P(\varepsilon(w_j, l_j, u_{j-1}^n, y_k^n)) \rightarrow 0$$

as $n \rightarrow \infty$.

For each message w_j , we define the following random variable and event:

$$\begin{aligned} N(w_j) := |\{\tilde{l}_j : (U_j^n(w_j, \tilde{l}_j), Y_k^n, U_{j-1}^n) \in T_\epsilon^{(n)}, \tilde{l}_j \neq L_j\}| \\ \varepsilon(w_j) := \{N(w_j) \geq 2^{n(\tilde{R}_j - R_j - I(U_j; Y_k | U_{j-1}) + \delta_n(\epsilon) - \epsilon_n/2) + 1}\} \end{aligned}$$

Finally, define the indicator random variable $E(w_j) := 0$ if $(U_j^n(w_j, L_j), Y_k^n, U_{j-1}^n) \in T_\epsilon^{(n)}$ and $\varepsilon(w_j)^c$ occurs; and $E(w_j) := 1$, otherwise. Therefore, we have

$$\begin{aligned} P(E(w_j) = 1) \leq P((U_j^n(w_j, L_j), U_{j-1}^n, Y_k^n) \notin T_\epsilon^{(n)}) \\ + P(\varepsilon(w_j) | (U_{j-1}^n, Y_k^n) \in T_\epsilon^{(n)}). \end{aligned} \quad (47)$$

It is clear that the first term in (47) goes to zero as $n \rightarrow \infty$. For the second term in (47), we have

$$\begin{aligned} P(\varepsilon(w_j) | (U_{j-1}^n, Y_k^n) \in T_\epsilon^{(n)}) \\ \leq \sum_{(u_{j-1}^n, y_k^n) \in T_\epsilon^{(n)}} P(u_{j-1}^n, y_k^n) P(\varepsilon(w_j) | u_{j-1}^n, y_k^n) \\ = \sum_{(u_{j-1}^n, y_k^n) \in T_\epsilon^{(n)}} \sum_{l_j} \left(P(u_{j-1}^n, y_k^n) P(l_j | u_{j-1}^n, y_k^n) \right. \\ \left. \cdot P(\varepsilon(w_j) | u_{j-1}^n, y_k^n, l_j) \right) \\ \rightarrow 0, \quad \text{if } \tilde{R}_j - R_j \geq I(U_j; Y_k | U_{j-1}). \end{aligned} \quad (48)$$

Therefore,

$$\begin{aligned} H(L_j | w_j, U_{j-1}^n, Y_k^n) \\ \leq H(L_j, E(w_j) | w_j, U_{j-1}^n, Y_k^n) \\ \leq H(E(w_j)) + H(L_j | w_j, U_{j-1}^n, Y_k^n, E(w_j)) \\ \leq 1 + P(E(w_j) = 1) H(L_j | w_j, Y_k^n, U_{j-1}^n, E(w_j) = 1) \\ + H(L_j | w_j, Y_k^n, U_{j-1}^n, E(w_j) = 0) \\ \leq 1 + P(E(w_j) = 1) n(\tilde{R}_j - R_j) \\ + \log 2^{n(\tilde{R}_j - R_j - I(U_j; Y_k | U_{j-1}) + \delta(\epsilon) - \epsilon/2) + 1} \\ = 1 + n(\tilde{R}_j - R_j) P(E(w_j) = 1) \\ + n(\tilde{R}_j - R_j - I(U_j; Y_k | U_{j-1}) + \delta(\epsilon) - \epsilon/2) + 1 \end{aligned} \quad (49)$$

Following from (49), we obtain,

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} H(L_j | w_j, U_{j-1}^n, Y_k^n) \\ \leq \tilde{R}_j - R_j - I(U_j; Y_k | U_{j-1}) + \delta'(\epsilon), \end{aligned} \quad (50)$$

where $\delta'(\epsilon) \rightarrow 0$ as $n \rightarrow \infty$. This concludes the proof.

APPENDIX C

CONVERSE PROOF OF THEOREM 1

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

$$H(W_k | Y_k^n) \leq n\epsilon_n, \quad \text{for } 1 \leq k \leq K, \quad (51)$$

and,

$$\begin{aligned} \frac{1}{n} I(W_{k+1}, \dots, W_K; Y_k^n | W_1, \dots, W_k) \leq \epsilon_n, \\ \text{for } 1 \leq k \leq K-1. \end{aligned} \quad (52)$$

We let $Y_k^{i-1} := (Y_{k,1}, \dots, Y_{k,i-1})$, and $Y_{k,i+1}^n := (Y_{k,i+1}, \dots, Y_{k,n})$. We set $U_{k,i} := \{W_1, \dots, W_k, Y_k^{i-1}, Y_{k-1,i+1}^n\}$ for $k = 1, \dots, K$ where $Y_0^n = \phi$. Due to the degradedness condition (13), it can be verified that $(U_{1,i}, \dots, U_{K-1,i}, X_i)$ satisfy the following Markov chain condition:

$$\begin{aligned} U_{1,i} \rightarrow U_{2,i} \rightarrow \dots \rightarrow U_{K-1,i} \rightarrow X_i \rightarrow Y_{K,i} \rightarrow \dots \rightarrow Y_{1,i}, \\ \text{for } 1 \leq i \leq n. \end{aligned} \quad (53)$$

$$\begin{aligned}
 nR_k &= H(W_k) = H(W_k|W^{k-1}) \\
 &= I(W_k; Y_k^n|W^{k-1}) + H(W_k|W^{k-1}, Y_k^n) \\
 &\stackrel{(a)}{\leq} I(W_k; Y_k^n|W^{k-1}) + n\epsilon_n \\
 &\stackrel{(b)}{\leq} I(W_k; Y_k^n|W^{k-1}) + 2n\epsilon_n - I(W_k; Y_{k-1}^n|W^{k-1}) \\
 &= \sum_{i=1}^n I(W_k; Y_{k,i}|W^{k-1}, Y_k^{i-1}) + 2n\epsilon_n - \sum_{i=1}^n I(W_k; Y_{k-1,i}|W^{k-1}, Y_{k-1,i+1}^n) \\
 &= \sum_{i=1}^n \left[I(W_k, Y_{k-1,i+1}^n; Y_{k,i}|W^{k-1}, Y_k^{i-1}) - I(W_k, Y_k^{i-1}; Y_{k-1,i}|W^{k-1}, Y_{k-1,i+1}^n) \right. \\
 &\quad \left. - I(Y_{k-1,i+1}^n; Y_{k,i}|W^k, Y_k^{i-1}) + I(Y_k^{i-1}; Y_{k-1,i}|W^k, Y_{k-1,i+1}^n) \right] + 2n\epsilon_n \\
 &\stackrel{(c)}{=} \sum_{i=1}^n \left[I(Y_{k-1,i+1}^n; Y_{k,i}|W^{k-1}, Y_k^{i-1}) + I(W_k; Y_{k,i}|W^{k-1}, Y_k^{i-1}, Y_{k-1,i+1}^n) \right. \\
 &\quad \left. - I(Y_k^{i-1}; Y_{k-1,i}|W^{k-1}, Y_{k-1,i+1}^n) - I(W_k; Y_{k-1,i}|W^{k-1}, Y_k^{i-1}, Y_{k-1,i+1}^n) \right] + 2n\epsilon_n \\
 &\stackrel{(d)}{=} \sum_{i=1}^n \left[I(W_k; Y_{k,i}|W^{k-1}, Y_k^{i-1}, Y_{k-1,i+1}^n) - I(W_k; Y_{k-1,i}|W^{k-1}, Y_k^{i-1}, Y_{k-1,i+1}^n) \right] + 2n\epsilon_n \\
 &= \sum_{i=1}^n \left[I(W_k, Y_k^{i-1}, Y_{k-1,i+1}^n; Y_{k,i}|W^{k-1}) - I(Y_k^{i-1}, Y_{k-1,i+1}^n; Y_{k,i}|W^{k-1}) \right. \\
 &\quad \left. - I(W_k, Y_k^{i-1}, Y_{k-1,i+1}^n; Y_{k-1,i}|W^{k-1}) + I(Y_k^{i-1}, Y_{k-1,i+1}^n; Y_{k-1,i}|W^{k-1}) \right] + 2n\epsilon_n \\
 &= \sum_{i=1}^n \left[I(W_k, Y_k^{i-1}, Y_{k-1,i+1}^n, Y_{k-1}^{i-1}, Y_{k-2,i+1}^n; Y_{k,i}|W^{k-1}) \right. \\
 &\quad \left. - I(W_k, Y_k^{i-1}, Y_{k-1,i+1}^n, Y_{k-1}^{i-1}, Y_{k-2,i+1}^n; Y_{k-1,i}|W^{k-1}) - I(Y_k^{i-1}, Y_{k-1,i+1}^n; Y_{k,i}|W^{k-1}) \right. \\
 &\quad \left. + I(Y_k^{i-1}, Y_{k-1,i+1}^n; Y_{k-1,i}|W^{k-1}) \right] + 2n\epsilon_n \\
 &= \sum_{i=1}^n \left[I(W_k, Y_k^{i-1}, Y_{k-1,i+1}^n; Y_{k,i}|W^{k-1}, Y_{k-1}^{i-1}, Y_{k-2,i+1}^n) \right. \\
 &\quad \left. - I(W_k, Y_k^{i-1}, Y_{k-1,i+1}^n; Y_{k-1,i}|W^{k-1}, Y_{k-1}^{i-1}, Y_{k-2,i+1}^n) + I(Y_{k-1}^{i-1}, Y_{k-2,i+1}^n; Y_{k,i}|W^{k-1}) \right. \\
 &\quad \left. - I(Y_{k-1}^{i-1}, Y_{k-2,i+1}^n; Y_{k-1,i}|W^{k-1}) - I(Y_k^{i-1}, Y_{k-1,i+1}^n; Y_{k,i}|W^{k-1}) \right. \\
 &\quad \left. + I(Y_k^{i-1}, Y_{k-1,i+1}^n; Y_{k-1,i}|W^{k-1}) \right] + 2n\epsilon_n \\
 &\stackrel{(e)}{=} \sum_{i=1}^n \left[I(W_k, Y_k^{i-1}, Y_{k-1,i+1}^n; Y_{k,i}|W^{k-1}, Y_{k-1}^{i-1}, Y_{k-2,i+1}^n) \right. \\
 &\quad \left. - I(W_k, Y_k^{i-1}, Y_{k-1,i+1}^n; Y_{k-1,i}|W^{k-1}, Y_{k-1}^{i-1}, Y_{k-2,i+1}^n) - I(Y_k^{i-1}, Y_{k-1,i+1}^n; Y_{k,i}|W^{k-1}, Y_{k-1}^{i-1}, Y_{k-2,i+1}^n) \right. \\
 &\quad \left. + I(Y_{k-1}^{i-1}, Y_{k-1,i+1}^n; Y_{k-1,i}|W^{k-1}, Y_{k-1}^{i-1}, Y_{k-2,i+1}^n) \right] + 2n\epsilon_n \\
 &\stackrel{(f)}{\leq} \sum_{i=1}^n \left[I(W_k, Y_k^{i-1}, Y_{k-1,i+1}^n; Y_{k,i}|W^{k-1}, Y_{k-1}^{i-1}, Y_{k-2,i+1}^n) \right. \\
 &\quad \left. - I(W_k, Y_k^{i-1}, Y_{k-1,i+1}^n; Y_{k-1,i}|W^{k-1}, Y_{k-1}^{i-1}, Y_{k-2,i+1}^n) \right] + 2n\epsilon_n \\
 &= \sum_{i=1}^n \left[I(U_{k,i}; Y_{k,i}|U_{k-1,i}) - I(U_{k,i}; Y_{k-1,i}|U_{k-1,i}) \right] + 2n\epsilon_n. \tag{54}
 \end{aligned}$$

We first bound the rate R_1 . Since there is no secrecy constraint for W_1 , following the standard steps, we obtain the following bound:

$$\begin{aligned}
nR_1 &= H(W_1) = I(W_1; Y_1^n) + H(W_1|Y_1^n) \\
&\leq I(W_1; Y_1^n) + n\epsilon_n \\
&= \sum_{i=1}^n I(W_1; Y_{1i}|Y_1^{i-1}) + n\epsilon_n \\
&\leq \sum_{i=1}^n I(W_1, Y_1^{i-1}; Y_{1i}) + n\epsilon_n \\
&= \sum_{i=1}^n I(U_{1,i}; Y_{1,i}) + n\epsilon_n. \tag{55}
\end{aligned}$$

For the message W_k , $2 \leq k \leq K$, we derive the bound as shown in (54) in the previous page, where (a) follows from Fano's inequality, (b) follows from (52), i.e., the secrecy constraint, (c) and (d) follow from the sum identity property in [12, Lemma 7], and (e) and (f) follows from the degradedness condition (13).

For $k = K$, based on (54), we obtain

$$\begin{aligned}
nR_K &\leq \sum_{i=1}^n \left[I(U_{K,i}; Y_{K,i}|U_{K-1,i}) - I(U_{K,i}; Y_{K-1,i}|U_{K-1,i}) \right] \\
&\quad + 2n\epsilon_n \\
&= \sum_{i=1}^n \left[I(U_{K,i}, X_i; Y_{K,i}|U_{K-1,i}) \right. \\
&\quad \left. - I(U_{K,i}, X_i; Y_{K-1,i}|U_{K-1,i}) - I(X_i; Y_{K,i}|U_{K,i}) \right. \\
&\quad \left. + I(X_i; Y_{K-1,i}|U_{K,i}) \right] + 2n\epsilon_n \\
&\leq \sum_{i=1}^n \left[I(X_i; Y_{K,i}|U_{K-1,i}) - I(X_i; Y_{K-1,i}|U_{K-1,i}) \right] \\
&\quad + 2n\epsilon_n, \tag{56}
\end{aligned}$$

where the last step follows from (53) and (13). The proof of the converse is completed by defining a uniformly distributed random variable $Q \in \{1, \dots, n\}$, and setting $U_k \triangleq (Q, U_{k,Q})$, $Y_k \triangleq Y_{k,Q}$, for $k \in [1 : K]$, and $X \triangleq (Q, X_Q)$.

APPENDIX D

PROOF OF CONVERSE OF FOR THEOREM 2

We continue the bounds proved for the discrete memoryless channel in their single-letter forms. We first bound R_1 as follows:

$$\begin{aligned}
R_1 &\leq I(U_1; Y_1) + \epsilon_n \\
&\leq \frac{1}{2} \log 2\pi e(P + N_1) - h(Y_1|U_1) + \epsilon_n, \tag{57}
\end{aligned}$$

where $h(Y_1|U_1)$ will be bounded later.

Following from (54), we have,

$$\begin{aligned}
R_2 &\leq I(U_2; Y_2|U_1) - I(U_2; Y_1|U_1) + 2\epsilon_n \\
&= h(Y_2|U_1) - h(Y_1|U_1) - (h(Y_2|U_1, U_2) \\
&\quad - h(Y_1|U_1, U_2)) + 2\epsilon_n \tag{58}
\end{aligned}$$

It can be shown as [17, eq. (140)] that

$$\frac{1}{2} \log \frac{N_2}{N_1} \leq h(Y_2|U_1) - h(Y_1|U_1) \leq \frac{1}{2} \log \frac{P + N_2}{P + N_1}. \tag{59}$$

Hence, there must exist an α_1 , $0 \leq \alpha_1 \leq 1$ such that

$$h(Y_2|U_1) - h(Y_1|U_1) = \frac{1}{2} \log \frac{(1 - \alpha_1)P + N_2}{(1 - \alpha_1)P + N_1}. \tag{60}$$

Due to the degradedness condition, it is clear that $-I(U_2; Y_2|U_1) + I(U_2; Y_1|U_1) \leq 0$, which implies

$$h(Y_2|U_1, U_2) - h(Y_1|U_1, U_2) \leq h(Y_2|U_1) - h(Y_1|U_1). \tag{61}$$

Hence, there must exist an α_2 , $0 \leq \alpha_2 \leq 1 - \alpha_1$, such that

$$h(Y_2|U_1, U_2) - h(Y_1|U_1, U_2) = \frac{1}{2} \log \frac{(1 - \alpha_1 - \alpha_2)P + N_2}{(1 - \alpha_1 - \alpha_2)P + N_1}. \tag{62}$$

Substituting (60) and (62) into (58), we obtain

$$\begin{aligned}
R_2 &\leq \frac{1}{2} \log \left(1 + \frac{\alpha_2 P}{(1 - \alpha_1 - \alpha_2)P + N_2} \right) \\
&\quad - \frac{1}{2} \log \left(1 + \frac{\alpha_2 P}{(1 - \alpha_1 - \alpha_2)P + N_1} \right). \tag{63}
\end{aligned}$$

In order to bound $h(Y_1|U_1)$, we note that Y_1 can be written as $Y_2 + Z'$, where Z' is Gaussian with variance $N_1 - N_2$. By the entropy power inequality [25], we have

$$2^{2h(Y_1|U_1)} \geq 2^{2h(Y_2|U_1)} + 2^{2h(Z'|U_1)}. \tag{64}$$

Using (60) and (64), we obtain

$$H(Y_1|U_1) \geq \frac{1}{2} \log(2\pi e((1 - \alpha_1)P + N_1)). \tag{65}$$

Substituting (65) into (57), we obtain

$$R_1 \leq \frac{1}{2} \log \left(1 + \frac{\alpha_1 P}{(1 - \alpha_1)P + N_1} \right). \tag{66}$$

We next bound R_3 and have

$$\begin{aligned}
R_3 &\leq h(Y_3|U_1, U_2) - h(Y_2|U_1, U_2) \\
&\quad - (h(Y_3|U_1, U_2, U_3) - h(Y_2|U_1, U_2, U_3)) + 2\epsilon_n \tag{67}
\end{aligned}$$

We let $\sqrt{t} = \sqrt{\frac{N_2 - N_3}{N_1 - N_3}}$, and write $Z_2 = Z_3 + \sqrt{t}\tilde{Z}$, where \tilde{Z} is Gaussian with variance $N_1 - N_3$. By the entropy power inequality in [26],

$$\begin{aligned}
2^{2h(Y_2|U_1, U_2)} &= 2^{2h(Y_3 + \sqrt{t}\tilde{Z}|U_1, U_2)} \\
&\geq (1 - t)2^{2h(Y_3|U_1, U_2)} + t2^{2h(Y_1|U_1, U_2)}, \tag{68}
\end{aligned}$$

which, jointly with (62), implies that

$$\begin{aligned}
h(Y_3|U_1, U_2) - h(Y_2|U_1, U_2) \\
\leq \frac{1}{2} \log \frac{(1 - \alpha_1 - \alpha_2)P + N_3}{(1 - \alpha_1 - \alpha_2)P + N_2}. \tag{69}
\end{aligned}$$

Using the same argument for obtaining (62), we can show that there exists $0 < \alpha_3 < 1 - \alpha_1 - \alpha_2$, such that

$$\begin{aligned}
h(Y_3|U_1, U_2, U_3) - h(Y_2|U_1, U_2, U_3) \\
= \frac{1}{2} \log \frac{(1 - \alpha_1 - \alpha_2 - \alpha_3)P + N_3}{(1 - \alpha_1 - \alpha_2 - \alpha_3)P + N_2}. \tag{70}
\end{aligned}$$

Substituting (69) and (70) into (67), we obtain the desired bound on R_3 .

It can be seen that the technique for bounding R_3 can be extended to bound R_4, \dots, R_K , and we hence complete the proof by noting that $P_k = \alpha_k P$ for $k = 1, \dots, K$.

APPENDIX E PROOF OF CONVERSE FOR THEOREM 3

In this proof, we first introduce some necessary definitions and useful lemmas in the previous studies [17], [19]. We then present our main proof.

A. Preliminaries

Definition 1 [17]: Let (\mathbf{U}, \mathbf{X}) be an arbitrarily correlated length- n random vector pair with well defined densities. The conditional Fisher information matrix of \mathbf{X} given \mathbf{U} is defined as

$$\mathbf{J}(\mathbf{X}|\mathbf{U}) = E[\rho(\mathbf{X}|\mathbf{U})\rho(\mathbf{X}|\mathbf{U})^T] \quad (71)$$

where the expectation is taken over the joint density $f(\mathbf{u}, \mathbf{x})$, and the conditional score function $\rho(\mathbf{x}|\mathbf{u})$ is given by

$$\begin{aligned} \rho(\mathbf{x}|\mathbf{u}) &= \nabla \log f(\mathbf{x}|\mathbf{u}) \\ &= \left[\frac{\partial \log f_U(\mathbf{x}|\mathbf{u})}{\partial x_1} \dots \frac{\partial \log f_U(\mathbf{x}|\mathbf{u})}{\partial x_n} \right]^T. \end{aligned} \quad (72)$$

Lemma 3 [17, Th. 11]: Let $(\mathbf{Z}_1, \mathbf{Z}_2, \mathbf{Z}_3, \mathbf{Z}_4)$ be Gaussian random vectors with covariance matrices $\Sigma_1, \Sigma_2, \Sigma_3, \Sigma_4$, respectively, where

$$\Sigma_4 \leq \Sigma_3 \leq \Sigma_2 \leq \Sigma_1. \quad (73)$$

Let (\mathbf{U}, \mathbf{X}) be an arbitrarily dependent random vector pair, which is independent of the Gaussian random vectors $(\mathbf{Z}_1, \mathbf{Z}_2, \mathbf{Z}_3, \mathbf{Z}_4)$, and the second moment of \mathbf{X} be constrained as $E[\mathbf{X}\mathbf{X}^T] \leq \mathbf{S}$. Then, for any feasible (\mathbf{U}, \mathbf{X}) , for any $\Sigma_1, \Sigma_2, \Sigma_3, \Sigma_4$ satisfying the order in (73), there exists a positive semidefinite matrix \mathbf{K}^* such that $\mathbf{K}^* \leq \mathbf{S}$, and

$$h(\mathbf{X} + \mathbf{Z}_2|\mathbf{U}) - h(\mathbf{X} + \mathbf{Z}_3|\mathbf{U}) = \frac{1}{2} \log \frac{|\mathbf{K}^* + \Sigma_2|}{|\mathbf{K}^* + \Sigma_3|}, \quad (74)$$

and

$$h(\mathbf{X} + \mathbf{Z}_1|\mathbf{U}) - h(\mathbf{X} + \mathbf{Z}_3|\mathbf{U}) \leq \frac{1}{2} \log \frac{|\mathbf{K}^* + \Sigma_1|}{|\mathbf{K}^* + \Sigma_3|}, \quad (75)$$

$$h(\mathbf{X} + \mathbf{Z}_3|\mathbf{U}) - h(\mathbf{X} + \mathbf{Z}_4|\mathbf{U}) \geq \frac{1}{2} \log \frac{|\mathbf{K}^* + \Sigma_3|}{|\mathbf{K}^* + \Sigma_4|}. \quad (76)$$

Lemma 4 [19]: Let (\mathbf{U}, \mathbf{X}) be an arbitrarily correlated random vector pair, and the second moment of \mathbf{X} is constrained as $E[\mathbf{X}\mathbf{X}^T] \leq \mathbf{S}$. Let $\mathbf{Z}_1, \mathbf{Z}_2$ be Gaussian random vectors that are independent from (\mathbf{U}, \mathbf{X}) , and have mean zero and covariance matrices Σ_1, Σ_2 , respectively, where $\Sigma_1 \geq \Sigma_2$. Then $h(\mathbf{X} + \mathbf{Z}_1|\mathbf{U}) - h(\mathbf{X} + \mathbf{Z}_2|\mathbf{U})$ is upper and lower bounded as follows,

$$\begin{aligned} &\frac{1}{2} \log \frac{|\mathbf{J}(\mathbf{X} + \mathbf{Z}_1|\mathbf{U})^{-1}|}{|\mathbf{J}(\mathbf{X} + \mathbf{Z}_1|\mathbf{U})^{-1} + \Sigma_2 - \Sigma_1|} \\ &\leq h(\mathbf{X} + \mathbf{Z}_1|\mathbf{U}) - h(\mathbf{X} + \mathbf{Z}_2|\mathbf{U}) \\ &\leq \frac{1}{2} \log \frac{|\mathbf{J}(\mathbf{X} + \mathbf{Z}_2|\mathbf{U})^{-1} + \Sigma_1 - \Sigma_2|}{|\mathbf{J}(\mathbf{X} + \mathbf{Z}_2|\mathbf{U})^{-1}|}. \end{aligned} \quad (77)$$

Proof: The proof of the unconditioned version of Lemma 4 is given in [19, Sec. V-B]. The proof can be generalized to the conditioned version by applying mathematical tools given in [19, Sec. V-D]. \square

Lemma 5 [19, Lemma 17]: Let $(\mathbf{V}, \mathbf{U}, \mathbf{X})$ be n -dimensional random vectors with well-defined densities. Moreover, assume that the partial derivatives of $f(\mathbf{u}|\mathbf{v}, \mathbf{x})$ with respect to x_i , $i = 1, \dots, n$, exist and satisfy

$$\max_{1 \leq i \leq n} \left| \frac{\partial f(\mathbf{u}|\mathbf{v}, \mathbf{x})}{\partial x_i} \right| \leq g(\mathbf{u}), \quad (78)$$

for some integrable function $g(\mathbf{u})$. If $(\mathbf{V}, \mathbf{U}, \mathbf{X})$ satisfy the Markov chain $\mathbf{V} \rightarrow \mathbf{U} \rightarrow \mathbf{X}$, then

$$\mathbf{J}(\mathbf{X}|\mathbf{U}) \geq \mathbf{J}(\mathbf{X}|\mathbf{V}). \quad (79)$$

Lemma 6 [19, Lemma 10]: Consider the function

$$r(t) = \frac{1}{2} \log \frac{|\mathbf{A} + \mathbf{B} + t\Delta|}{|\mathbf{A} + t\Delta|}, \quad 0 \leq t \leq 1. \quad (80)$$

where $\mathbf{A}, \mathbf{B}, \Delta$ are real symmetric matrices, and $\mathbf{A} \succ \mathbf{0}, \mathbf{B} \geq \mathbf{0}, \Delta \geq \mathbf{0}$. Then $r(t)$ is continuous and monotonically decreasing with respect to t .

Lemma 7 [19]: Suppose (\mathbf{U}, \mathbf{X}) is a random vector pair with arbitrary joint distribution and the second order moment of \mathbf{X} satisfies $E(\mathbf{X}\mathbf{X}^T) \leq \mathbf{S}$. Let \mathbf{Z} be a random Gaussian vector that is independent from \mathbf{U} and \mathbf{X} and has mean zero and covariance Σ . Then we have

$$0 \leq \mathbf{J}(\mathbf{X} + \mathbf{Z}|\mathbf{U})^{-1} - \Sigma \leq \mathbf{S}. \quad (81)$$

Lemma 8 [19]: Suppose (\mathbf{U}, \mathbf{X}) is a random vector pair with arbitrary joint distribution and the second order moment of \mathbf{X} satisfies $E(\mathbf{X}\mathbf{X}^T) \leq \mathbf{S}$. Let $\mathbf{Z}_1, \mathbf{Z}_2$ be random Gaussian vectors that are independent from \mathbf{U} and \mathbf{X} and have mean zero and covariance matrices $\Sigma_1 \geq \Sigma_2$. Then we have

$$\mathbf{J}(\mathbf{X} + \mathbf{Z}_1|\mathbf{U})^{-1} + \Sigma_2 - \Sigma_1 - \mathbf{J}(\mathbf{X} + \mathbf{Z}_2|\mathbf{U})^{-1} \geq \mathbf{0}. \quad (82)$$

The proof of Lemma 8 follows the arguments in [19, Proof of Lemma 6] for the unconditional case, but using [19, Corollary 4] for the conditional case.

B. Main Proof

Following the converse proof of Theorem 1 in Appendix C, we have the inequalities as follows:

$$R_1 \leq I(\mathbf{U}_1; \mathbf{Y}_1), \quad (83)$$

$$R_k \leq I(\mathbf{U}_k; \mathbf{Y}_k|\mathbf{U}_{k-1}) - I(\mathbf{U}_k; \mathbf{Y}_{k-1}|\mathbf{U}_{k-1}), \quad (84)$$

$$\text{for } 2 \leq k \leq K,$$

where the random variables satisfy the Markov chain condition in (22).

We first derive the bounds on R_2 and R_3 in order to show that the bounding techniques can be extended to prove the bounds on R_4, \dots, R_K . We then derive the bound on R_1 .

To bound R_2 , we start with (83), and have

$$\begin{aligned} R_2 &\leq I(\mathbf{U}_2; \mathbf{Y}_2|\mathbf{U}_1) - I(\mathbf{U}_2; \mathbf{Y}_1|\mathbf{U}_1) \\ &\stackrel{(a)}{=} h(\mathbf{Y}_2|\mathbf{U}_1) - h(\mathbf{Y}_2|\mathbf{U}_2) - (h(\mathbf{Y}_1|\mathbf{U}_1) - h(\mathbf{Y}_1|\mathbf{U}_2)) \\ &= (h(\mathbf{Y}_2|\mathbf{U}_1) - h(\mathbf{Y}_1|\mathbf{U}_1)) - (h(\mathbf{Y}_2|\mathbf{U}_2) - h(\mathbf{Y}_1|\mathbf{U}_2)), \end{aligned} \quad (85)$$

where (a) follows from the Markov chain condition in (22).

Following from Lemma 4, we obtain the following upper and lower bounds on $h(\mathbf{Y}_2|\mathbf{U}_1) - h(\mathbf{Y}_1|\mathbf{U}_1)$.

$$\begin{aligned} \frac{1}{2} \log \frac{|\mathbf{J}(\mathbf{X} + \mathbf{Z}_2|\mathbf{U}_1)^{-1}|}{|\mathbf{J}(\mathbf{X} + \mathbf{Z}_2|\mathbf{U}_1)^{-1} + \Sigma_1 - \Sigma_2|} \\ \leq h(\mathbf{Y}_2|\mathbf{U}_1) - h(\mathbf{Y}_1|\mathbf{U}_1) \\ \leq \frac{1}{2} \log \frac{|\mathbf{J}(\mathbf{X} + \mathbf{Z}_1|\mathbf{U}_1)^{-1} + \Sigma_2 - \Sigma_1|}{|\mathbf{J}(\mathbf{X} + \mathbf{Z}_1|\mathbf{U}_1)^{-1}|}. \end{aligned} \quad (86)$$

Define

$$\begin{aligned} \mathbf{A} &= \mathbf{J}(\mathbf{X} + \mathbf{Z}_2|\mathbf{U}_1)^{-1}, \\ \mathbf{B} &= \Sigma_1 - \Sigma_2, \\ \Delta &= \mathbf{J}(\mathbf{X} + \mathbf{Z}_1|\mathbf{U}_1)^{-1} + \Sigma_2 - \Sigma_1 - \mathbf{J}(\mathbf{X} + \mathbf{Z}_2|\mathbf{U}_1)^{-1}, \end{aligned}$$

and

$$r(t) = \frac{1}{2} \log \frac{|\mathbf{A} + \mathbf{B} + t\Delta|}{|\mathbf{A} + t\Delta|}.$$

Therefore, (86) can be rewritten into,

$$-r(0) \leq h(\mathbf{Y}_2|\mathbf{U}_1) - h(\mathbf{Y}_1|\mathbf{U}_1) \leq -r(1). \quad (87)$$

It can be verified that $\mathbf{A} > 0$, $\mathbf{B} \geq 0$, and $\Delta \geq 0$. In particular, $\Delta \geq 0$ is due to Lemma 8.

Following from Lemma 6, $r(t)$ is a continuous and monotonically decreasing function in t . Hence, (87) implies that there must exist a constant t_1 with $0 \leq t_1 \leq 1$, such that

$$h(\mathbf{Y}_2|\mathbf{U}_1) - h(\mathbf{Y}_1|\mathbf{U}_1) = -r(t_1). \quad (88)$$

We define

$$\begin{aligned} \mathbf{S}_1 &:= \mathbf{A} + t_1\Delta - \Sigma_2 = \mathbf{J}(\mathbf{X} + \mathbf{Z}_2|\mathbf{U}_1)^{-1} \\ &+ t_1(\mathbf{J}(\mathbf{X} + \mathbf{Z}_1|\mathbf{U}_1)^{-1} + \Sigma_2 - \Sigma_1 - \mathbf{J}(\mathbf{X} + \mathbf{Z}_2|\mathbf{U}_1)^{-1}) - \Sigma_2. \end{aligned} \quad (89)$$

Therefore,

$$h(\mathbf{Y}_2|\mathbf{U}_1) - h(\mathbf{Y}_1|\mathbf{U}_1) = -r(t_1) = \frac{1}{2} \log \frac{|\mathbf{S}_1 + \Sigma_2|}{|\mathbf{S}_1 + \Sigma_1|}. \quad (90)$$

It can be seen that \mathbf{S}_1 satisfies $\mathbf{A} - \Sigma_2 \leq \mathbf{S}_1 \leq \mathbf{A} + \Delta - \Sigma_2$. Following Lemma 7, we have

$$\begin{aligned} \mathbf{0} &\leq \mathbf{J}(\mathbf{X} + \mathbf{Z}_2|\mathbf{U}_1)^{-1} - \Sigma_2 = \mathbf{A} - \Sigma_2 \leq \mathbf{S}_1 \\ &\leq \mathbf{A} + \Delta - \Sigma_2 = \mathbf{J}(\mathbf{X} + \mathbf{Z}_1|\mathbf{U}_1)^{-1} - \Sigma_1 \leq \mathbf{S}, \end{aligned} \quad (91)$$

which implies

$$\mathbf{0} \leq \mathbf{S}_1 \leq \mathbf{S}. \quad (92)$$

We next study the term $h(\mathbf{Y}_2|\mathbf{U}_2) - h(\mathbf{Y}_1|\mathbf{U}_2)$. Due to the Markov chain condition (22), it is clear that $-I(\mathbf{U}_2; \mathbf{Y}_2|\mathbf{U}_1) + I(\mathbf{U}_2; \mathbf{Y}_1|\mathbf{U}_1) \leq 0$, which implies that

$$\begin{aligned} h(\mathbf{Y}_2|\mathbf{U}_2) - h(\mathbf{Y}_1|\mathbf{U}_2) &\leq h(\mathbf{Y}_2|\mathbf{U}_1) - h(\mathbf{Y}_1|\mathbf{U}_1) \\ &= \frac{1}{2} \log \frac{|\mathbf{S}_1 + \Sigma_2|}{|\mathbf{S}_1 + \Sigma_1|}. \end{aligned} \quad (93)$$

Applying Lemma 4, we obtain

$$\begin{aligned} \frac{1}{2} \log \frac{|\mathbf{J}(\mathbf{X} + \mathbf{Z}_2|\mathbf{U}_2)^{-1}|}{|\mathbf{J}(\mathbf{X} + \mathbf{Z}_2|\mathbf{U}_2)^{-1} + \Sigma_1 - \Sigma_2|} \\ \leq h(\mathbf{Y}_2|\mathbf{U}_2) - h(\mathbf{Y}_1|\mathbf{U}_2) \\ \leq \frac{1}{2} \log \frac{|\mathbf{J}(\mathbf{X} + \mathbf{Z}_2|\mathbf{U}_2)^{-1} + \Sigma_2 - \Sigma_1|}{|\mathbf{J}(\mathbf{X} + \mathbf{Z}_1|\mathbf{U}_2)^{-1}|}. \end{aligned} \quad (94)$$

Combining (93) and (94), we have

$$\begin{aligned} \frac{1}{2} \log \frac{|\mathbf{J}(\mathbf{X} + \mathbf{Z}_2|\mathbf{U}_2)^{-1}|}{|\mathbf{J}(\mathbf{X} + \mathbf{Z}_2|\mathbf{U}_2)^{-1} + \Sigma_1 - \Sigma_2|} \\ \leq h(\mathbf{Y}_2|\mathbf{U}_2) - h(\mathbf{Y}_1|\mathbf{U}_2) \leq \frac{1}{2} \log \frac{|\mathbf{S}_1 + \Sigma_2|}{|\mathbf{S}_1 + \Sigma_1|}. \end{aligned} \quad (95)$$

We now consider the function

$$r(t) = \frac{1}{2} \log \frac{|\mathbf{A} + \mathbf{B} + t\Delta|}{|\mathbf{A} + t\Delta|} \quad (96)$$

with \mathbf{A} , \mathbf{B} and Δ being redefined as,

$$\begin{aligned} \mathbf{A} &= \mathbf{J}(\mathbf{X} + \mathbf{Z}_2|\mathbf{U}_2)^{-1} \\ \mathbf{B} &= \Sigma_1 - \Sigma_2 \\ \Delta &= \mathbf{S}_1 + \Sigma_2 - \mathbf{J}(\mathbf{X} + \mathbf{Z}_2|\mathbf{U}_2)^{-1}, \end{aligned} \quad (97)$$

where $\mathbf{A} > 0$, $\mathbf{B} \geq 0$, and $\Delta \geq 0$. In order to show $\Delta \geq \mathbf{0}$, we show that $\mathbf{S}_1 \geq \mathbf{J}(\mathbf{X} + \mathbf{Z}_2|\mathbf{U}_2)^{-1} - \Sigma_2$. Using Lemma 5, we have

$$\mathbf{J}(\mathbf{X} + \mathbf{Z}_2|\mathbf{U}_2) \geq \mathbf{J}(\mathbf{X} + \mathbf{Z}_2|\mathbf{U}_1). \quad (98)$$

Hence,

$$\mathbf{J}(\mathbf{X} + \mathbf{Z}_2|\mathbf{U}_1)^{-1} \geq \mathbf{J}(\mathbf{X} + \mathbf{Z}_2|\mathbf{U}_2)^{-1}. \quad (99)$$

Since $\mathbf{S}_1 \geq \mathbf{J}(\mathbf{X} + \mathbf{Z}_2|\mathbf{U}_1)^{-1} - \Sigma_2$, we have

$$\mathbf{S}_1 \geq \mathbf{J}(\mathbf{X} + \mathbf{Z}_2|\mathbf{U}_2)^{-1} - \Sigma_2. \quad (100)$$

Thus, (95) can be rewritten as

$$-r(0) \leq h(\mathbf{Y}_2|\mathbf{U}_2) - h(\mathbf{Y}_1|\mathbf{U}_2) \leq -r(1). \quad (101)$$

Since the function $r(t)$ is monotone and continuous, there exists a constant t_2 with $0 \leq t_2 \leq 1$ such that $h(\mathbf{Y}_2|\mathbf{U}_2) - h(\mathbf{Y}_1|\mathbf{U}_2) = -r(t_2)$. Let $\mathbf{S}_2 = \mathbf{A} + t_2\Delta - \Sigma_2$, and obtain

$$h(\mathbf{Y}_2|\mathbf{U}_2) - h(\mathbf{Y}_1|\mathbf{U}_2) = -r(t_2) = \frac{1}{2} \log \frac{|\mathbf{S}_2 + \Sigma_2|}{|\mathbf{S}_2 + \Sigma_1|} \quad (102)$$

It can be seen that

$$\mathbf{J}(\mathbf{X} + \mathbf{Z}_2|\mathbf{U}_2)^{-1} - \Sigma_2 = \mathbf{A} - \Sigma_2 \leq \mathbf{S}_2 \leq \mathbf{A} + \Delta - \Sigma_2 = \mathbf{S}_1. \quad (103)$$

Therefore, combining (90) and (102), we obtain

$$\begin{aligned} R_2 &\leq \frac{1}{2} \log \frac{|\mathbf{S}_1 + \Sigma_2|}{|\mathbf{S}_1 + \Sigma_1|} - \frac{1}{2} \log \frac{|\mathbf{S}_2 + \Sigma_2|}{|\mathbf{S}_2 + \Sigma_1|} \\ &= \frac{1}{2} \log \frac{|\mathbf{S}_1 + \Sigma_2|}{|\mathbf{S}_2 + \Sigma_2|} - \frac{1}{2} \log \frac{|\mathbf{S}_1 + \Sigma_1|}{|\mathbf{S}_2 + \Sigma_1|}. \end{aligned} \quad (104)$$

We next derive an upper bound on R_3 , which is a necessary step to show that the proof techniques can be iteratively extended to bound R_4, \dots, R_K . Following from (83), we have

$$R_3 \leq h(\mathbf{Y}_3|\mathbf{U}_2) - h(\mathbf{Y}_2|\mathbf{U}_2) - (h(\mathbf{Y}_3|\mathbf{U}_3) - h(\mathbf{Y}_2|\mathbf{U}_3)). \quad (105)$$

Using Lemma 3 and (102), we obtain

$$h(\mathbf{Y}_3|\mathbf{U}_2) - h(\mathbf{Y}_2|\mathbf{U}_2) \leq \frac{1}{2} \log \frac{|\mathbf{S}_2 + \Sigma_3|}{|\mathbf{S}_2 + \Sigma_2|}. \quad (106)$$

Similarly to (93), due to the Markov chain condition (22), we have

$$h(\mathbf{Y}_3|\mathbf{U}_3) - h(\mathbf{Y}_2|\mathbf{U}_3) \leq h(\mathbf{Y}_3|\mathbf{U}_2) - h(\mathbf{Y}_2|\mathbf{U}_2). \quad (107)$$

Using Lemma 4 and (106) and (107), we have,

$$\begin{aligned} & \frac{1}{2} \log \frac{|\mathbf{J}(\mathbf{X} + \mathbf{Z}_3|\mathbf{U}_3)^{-1}|}{|\mathbf{J}(\mathbf{X} + \mathbf{Z}_3|\mathbf{U}_3)^{-1} + \Sigma_2 - \Sigma_3|} \\ & \leq h(\mathbf{Y}_3|\mathbf{U}_3) - h(\mathbf{Y}_2|\mathbf{U}_3) \leq \frac{1}{2} \log \frac{|\mathbf{S}_2 + \Sigma_3|}{|\mathbf{S}_2 + \Sigma_2|}. \end{aligned} \quad (108)$$

It can be shown that $\mathbf{S}_2 \geq \mathbf{J}(\mathbf{X} + \mathbf{Z}_3|\mathbf{U}_3)^{-1} - \Sigma_3$ by using Lemma 5 and Lemma 8. Then following the similar arguments that yield (102), we can show that there exists a \mathbf{S}_3 , such that $\mathbf{0} \leq \mathbf{S}_3 \leq \mathbf{S}_2 \leq \mathbf{S}_1 \leq \mathbf{S}$ and

$$h(\mathbf{Y}_3|\mathbf{U}_3) - h(\mathbf{Y}_2|\mathbf{U}_3) = \frac{1}{2} \log \frac{|\mathbf{S}_3 + \Sigma_3|}{|\mathbf{S}_3 + \Sigma_2|}. \quad (109)$$

Therefore, substituting (106) and (109) into (105), we obtain

$$\begin{aligned} R_3 & \leq \frac{1}{2} \log \frac{|\mathbf{S}_2 + \Sigma_3|}{|\mathbf{S}_2 + \Sigma_2|} - \frac{1}{2} \log \frac{|\mathbf{S}_3 + \Sigma_3|}{|\mathbf{S}_3 + \Sigma_2|} \\ & = \frac{1}{2} \log \frac{|\mathbf{S}_2 + \Sigma_3|}{|\mathbf{S}_3 + \Sigma_3|} - \frac{1}{2} \log \frac{|\mathbf{S}_2 + \Sigma_2|}{|\mathbf{S}_3 + \Sigma_2|}. \end{aligned} \quad (110)$$

Using techniques similar to those for bounding R_2 and R_3 , we can derive the desired bounds on R_4, \dots, R_K iteratively.

Finally, we bound the rate R_1 . We introduce a virtual receiver $\mathbf{Y}_0 = \mathbf{X} + \mathbf{Z}_0$, where \mathbf{Z}_0 is a Gaussian vector with the covariance matrix of $\Sigma_0 = t\Sigma_1$ with $t \geq 1$. Hence, $\Sigma_0 \geq \Sigma_1$. Following from (90) and Lemma 3, we have,

$$h(\mathbf{Y}_0|\mathbf{U}_1) - h(\mathbf{Y}_1|\mathbf{U}_1) \leq \frac{1}{2} \log \frac{|\mathbf{S}_1 + \Sigma_0|}{|\mathbf{S}_1 + \Sigma_1|}, \quad (111)$$

for any $t \geq 1$. On the other hand, we have

$$\begin{aligned} \frac{1}{2} \log(2\pi e)^r |\Sigma_0| & = h(\mathbf{Z}_0) \leq h(\mathbf{Y}_0|\mathbf{U}_1) \leq h(\mathbf{Y}_0) \\ & \leq \frac{1}{2} \log(2\pi e)^r |\mathbf{S} + \Sigma_0|, \end{aligned} \quad (112)$$

which implies that

$$\begin{aligned} \frac{1}{2} \log \frac{|\Sigma_0|}{|\mathbf{S}_1 + \Sigma_0|} & \leq h(\mathbf{Y}_0|\mathbf{U}_1) - \frac{1}{2} \log(2\pi e)^r |\mathbf{S}_1 + \Sigma_0| \\ & \leq \frac{1}{2} \log \frac{|\mathbf{S} + \Sigma_0|}{|\mathbf{S}_1 + \Sigma_0|}. \end{aligned} \quad (113)$$

As $t \rightarrow \infty$, $\frac{1}{2} \log \frac{|\Sigma_0|}{|\mathbf{S}_1 + \Sigma_0|} \rightarrow 0$ and $\frac{1}{2} \log \frac{|\mathbf{S} + \Sigma_0|}{|\mathbf{S}_1 + \Sigma_0|} \rightarrow 0$. Hence, $h(\mathbf{Y}_0|\mathbf{U}_1) - \frac{1}{2} \log(2\pi e)^r |\mathbf{S}_1 + \Sigma_0| \rightarrow 0$ as $t \rightarrow \infty$. Since (111) holds for any $t \geq 1$, we have $h(\mathbf{Y}_1|\mathbf{U}_1) \geq \frac{1}{2} \log(2\pi e)^r |\mathbf{S}_1 + \Sigma_1|$.

Following from (83),

$$\begin{aligned} R_1 & \leq I(\mathbf{U}_1; \mathbf{Y}_1) \\ & = h(\mathbf{Y}_1) - h(\mathbf{Y}_1|\mathbf{U}_1) \\ & \leq \frac{1}{2} \log(2\pi e)^r |\mathbf{S} + \Sigma_1| - \frac{1}{2} \log(2\pi e)^r |\mathbf{S}_1 + \Sigma_1| \\ & = \frac{1}{2} \log \frac{|\mathbf{S} + \Sigma_1|}{|\mathbf{S}_1 + \Sigma_1|}, \end{aligned} \quad (114)$$

which completes the proof.

REFERENCES

- [1] L. Lai, Y. Liang, W. Du, and S. Shamai (Shitz), "Secret sharing via noisy broadcast channels," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, St. Petersburg, Russia, Jul./Aug. 2011, pp. 1955–1959.
- [2] S. Zou, Y. Liang, L. Lai, and S. Shamai (Shitz), "Layered decoding and secrecy over degraded broadcast channels," in *Proc. IEEE 14th Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jun. 2013, pp. 679–683.
- [3] S. Zou, Y. Liang, L. Lai, and S. Shamai (Shitz), "Layered secure broadcasting over MIMO channels and application in secret sharing," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun./Jul. 2014, pp. 396–400.
- [4] A. C. Yao, "Protocols for secure computations," in *Proc. 23rd Annu. Symp. Found. Comput. Sci.*, 1982, pp. 160–164.
- [5] A. C.-C. Yao, "How to generate and exchange secrets," in *Proc. 27th Annu. Symp. Found. Comput. Sci.*, 1986, pp. 162–167.
- [6] M. O. Rabin, "Randomized Byzantine generals," in *Proc. 24th Annu. Symp. Found. Comput. Sci.*, Nov. 1983, pp. 403–409.
- [7] Y. Desmedt and Y. Frankel, "Shared generation of authenticators and signatures," in *Advances in Cryptology—CRYPTO* (Lecture Notes in Computer Science), vol. 576. Berlin, Germany: Springer-Verlag, 1992, pp. 457–469.
- [8] M. Naor and A. Wool, "Access control and signatures via quorum secret sharing," in *Proc. 3rd ACM Conf. Comput. Commun. Secur. (CCS)*, New York, NY, USA, 1996, pp. 157–168.
- [9] V. Goyal, O. Pandey, A. Sahai, and B. Waters, "Attribute-based encryption for fine-grained access control of encrypted data," in *Proc. 13th ACM Conf. Comput. Commun. Secur. (CCS)*, New York, NY, USA, 2006, pp. 89–98.
- [10] B. Shankar, K. Srinathan, and C. P. Rangan, "Alternative protocols for generalized oblivious transfer," in *Proc. 9th Int. Conf. Distrib. Comput. Netw. (ICDCN)*, Berlin, Germany, 2008, pp. 304–309.
- [11] A. D. Wyner, "The wire-tap channel," *Bell Syst. Tech. J.*, vol. 54, no. 8, pp. 1355–1387, Oct. 1975.
- [12] I. Csiszár and J. Körner, "Broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. 24, no. 3, pp. 339–348, May 1978.
- [13] Y. Liang, H. V. Poor, and S. Shamai (Shitz), "Information theoretic security," in *Foundations and Trends in Communications and Information Theory*, vol. 5. Hanover, MA, USA: NOW Publishers, 2008.
- [14] M. Bloch and J. Barros, *Physical-Layer Security: From Information Theory to Security Engineering*. Cambridge, U.K.: Cambridge Univ. Press, 2011.
- [15] Y. Liang, G. Kramer, H. V. Poor, and S. Shamai (Shitz), "Compound wiretap channels," *EURASIP J. Wireless Commun. Netw.*, vol. 2009, Mar. 2009, Art. ID 142374.
- [16] 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.
- [17] E. Ekrem and S. Ulukus, "Degraded compound multi-receiver wiretap channels," *IEEE Trans. Inf. Theory*, vol. 58, no. 9, pp. 5681–5698, Sep. 2012.
- [18] A. El Gamal and Y.-H. Kim, *Network Information Theory*. New York, NY, USA: Cambridge Univ. Press, 2012.
- [19] E. Ekrem and S. Ulukus, "The secrecy capacity region of the Gaussian MIMO multi-receiver wiretap channel," *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 2083–2114, Apr. 2011.
- [20] C. Blundo, A. De Santis, A. Giorgio Gaggia, and U. Vaccaro, "New bounds on the information rate of secret sharing schemes," *IEEE Trans. Inf. Theory*, vol. 41, no. 2, pp. 549–554, Mar. 1995.
- [21] G. Bagherikaram, A. S. Motahari, and A. K. Khandani, "Secure broadcasting: The secrecy rate region," in *Proc. 46th Annu. Allerton Conf. Commun. Control Comput.*, Sep. 2008, pp. 834–841.
- [22] H. Weingarten, T. Liu, S. Shamai (Shitz), Y. Steinberg, and P. Viswanath, "The capacity region of the degraded multiple-input multiple-output compound broadcast channel," *IEEE Trans. Inf. Theory*, vol. 55, no. 11, pp. 5011–5023, Nov. 2009.
- [23] M. Kobayashi, Y. Liang, S. Shamai (Shitz), and M. Debbah, "On the compound MIMO broadcast channels with confidential messages," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun./Jul. 2009, pp. 1283–1287.
- [24] M. A. Maddah-Ali, "On the degrees of freedom of the compound MISO broadcast channels with finite states," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2010, pp. 2273–2277.
- [25] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York, NY, USA: Wiley, 1991.
- [26] M. H. M. Costa, "A new entropy power inequality," *IEEE Trans. Inf. Theory*, vol. 31, no. 6, pp. 751–760, Nov. 1985.

Shaofeng Zou (S'14) received the B.E. degree (with honors) from Shanghai Jiao Tong University, Shanghai, China in 2011. Since September 2011, he has been a PhD student at Syracuse University. His research interests are on information theory, nonparametric detection and machine learning. He received the Excellent Student Scholarship from Shanghai Jiaotong University in 2008, 2009 and 2010, and received the National scholarship from the Ministry of Education of China in 2008. He also received the National Encouragement Scholarship from the Ministry of Education of China in 2009 and 2010.

Yingbin Liang (S'01–M'05) 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 the Department of Electrical Engineering at the University of Hawaii. Since December 2009, she has been on the faculty at Syracuse University, where she is an associate professor. Dr. Liang's research interests include information theory, wireless communications and networks, and machine learning.

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. More recently, her paper received the 2014 EURASIP Best Paper Award for the *EURASIP Journal on Wireless Communications and Networking*. She is currently serving as an Associate Editor for the Shannon Theory of the IEEE TRANSACTIONS ON INFORMATION THEORY.

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, and was an assistant professor at University of Arkansas, Little Rock from 2009 to 2012. Since Aug. 2012, he has been an assistant professor at Worcester Polytechnic Institute. Dr. Lai's 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.

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, 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 and the 2014 Rothschild Prize in Mathematics/Computer Sciences and Engineering.

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 European Commission FP7, Network of Excellence in Wireless COMMUNICATIONS (NEWCOM++) Best Paper Award, and 2014 EURASIP Best Paper Award (for the *EURASIP Journal on Wireless Communications and Networking*). He is also the recipient of the 2010 Thomson Reuters Award for International Excellence in Scientific Research and is listed in the 2014 Thomson Reuters "The World's Most Influential Scientific Minds." 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 on the Executive Editorial Board of the IEEE TRANSACTIONS ON INFORMATION THEORY and has also served as a Shannon Theory Associate Editor for this journal. He has served twice on the Board of Governors of the Information Theory Society.