

Secure Multiplex Coding with Dependent and Non-Uniform Multiple Messages

Masahito Hayashi, *Senior Member, IEEE*, Ryutaroh Matsumoto, *Member, IEEE*,

Abstract—The secure multiplex coding (SMC) is a technique to remove rate loss in the coding for wire-tap channels and broadcast channels with confidential messages caused by the inclusion of random bits into transmitted signals. SMC replaces the random bits by other meaningful secret messages, and a collection of secret messages serves as the random bits to hide the rest of messages. In the previous researches, multiple secret messages were assumed to have independent and uniform distributions, which is difficult to be ensured in practice. We remove this restrictive assumption by a generalization of the channel resolvability technique.

We also give practical construction techniques for SMC by using an arbitrary given error-correcting code as an ingredient, and channel-universal coding of SMC. By using the same principle as the channel-universal SMC, we give coding for the broadcast channel with confidential messages universal to both channel and source distributions.

Index Terms—broadcast channel with confidential messages, information theoretic security, multiuser information theory, universal coding, the secure multiplex coding

I. INTRODUCTION

A. Overview

Recently, the security of personal information is demanded much more. The wire-tap model is a typical secure message transmission model with the presence of an eavesdropper. Specially, there are the legitimate sender called Alice, the legitimate receiver called Bob, and the eavesdropper Eve. There is also a noisy broadcast channel from Alice to Bob and Eve. Alice wants to send secret messages reliably to Bob and secretly from Eve. This problem was first formulated by Wyner [35]. Csiszár and Körner generalized Wyner’s original problem to include common messages from Alice to both Bob and Eve, and determined the optimal information rate tuples of the secret message and the common message, and the information

leakage rate of the secret message to Eve, which is measured by the conditional entropy of the secret message given Eve’s received signal [9]. They called their generalized problem as the broadcast channel with confidential messages, hereafter abbreviated as BCC. The secrecy of messages over the wire-tap channel and the BCC is realized by including meaningless random variable, which is called the dummy message, into Alice’s transmitted signal. This decreases the information rate.

In order to get rid of this information rate loss, Yamamoto et al. [22] proposed the secure multiplex coding, hereafter abbreviated as SMC, as a generalization of the wire-tap channel coding. The SMC can be used, for example, in the following case. When a company treats a collection of personal information, it is required to keep the secrecy of the respective personal information. However, it may not be required to keep the secrecy of the relation among several personal information. For example, when all of personal information are subject to the uniform distribution of the same length bit sequence, the secrecy of their exclusive OR may not be required. Consider the case when the sender Alice sends the collection of T persons’ personal information S_1, \dots, S_T via the channel partially leaked to Eve. It is required that the receiver Bob can decode all of S_1, \dots, S_T , and that Eve cannot obtain any information of the respective personal information. In order to keep the secrecy of the message S_i from Eve, Yamamoto et al. [22] proposed to use the remaining information $S_1, \dots, S_{i-1}, S_{i+1}, \dots, S_T$ as the dummy message for the message S_i . Then, they realized the secrecy of the message S_i without loss of the information rate. This type of coding problem is called the SMC. It is known that the application of the channel resolvability [13] yields the security of the wire-tap channel model [15]. Hence, employing this method, Yamamoto et al. [22] proved the security of SMC.

On the other hand, since $S_1, \dots, S_{i-1}, S_{i+1}, \dots, S_T$ are personal information, they are not necessarily uniform random bits and might be dependent, while the existing papers [27], [22] assumed their uniformity and independence. Such assumption is difficult to be ensured in practice. Unfortunately, the application of the original channel resolvability can prove the security only when the messages $S_1, \dots, S_{i-1}, S_{i+1}, \dots, S_T$ are conditionally uniform and independent of S_i because it treats the approximation of the channel output distribution with the uniform input random variable. One may consider that the compressed data satisfies that assumption so that the removal of that assumption is not needed. However, as is shown in [14], [16], the compressed data is not uniform in the sense of the variational distance nor the divergence. That is, the uniformity assumption does not hold for such compressed data. Hence, the removal of the assumption is essential for non-uniform

This research was partially supported by the MEXT Grant-in-Aid for Young Scientists (A) No. 20686026, (B) No. 22760267, Grant-in-Aid for Scientific Research (A) No. 23246071, and the ImPACT Program of Council for Science, Technology and Innovation (Cabinet Office, Government of Japan), the Villum Foundation through their VELUX Visiting Professor Programme 2011–2012. The Centre for Quantum Technologies is funded by the Singapore Ministry of Education and the National Research Foundation as part of the Research Centres of Excellence programme. This paper was presented in part at 2010 IEEE International Symposium on Information Theory, Austin, Texas, USA, June, 2010 [18], in part at 2011 IEEE International Symposium on Information Theory, Saint Petersburg, Russia, August 2011 [27], in part at 49th Annual Allerton Conference, University of Illinois at Urbana-Champaign, IL, USA, September 2011 [19], and in part at 50th Annual Allerton Conference, University of Illinois at Urbana-Champaign, IL, USA, October 2012 [20].

M. Hayashi is with Graduate School of Mathematics, Nagoya University, Furocho, Chikusaku, Nagoya, 464-8602, Japan, and Centre for Quantum Technologies, National University of Singapore, 3 Science Drive 2, Singapore 117542. (e-mail: masahito@math.nagoya-u.ac.jp)

R. Matsumoto is with Department of Communications and Computer Engineering, Tokyo Institute of Technology, 152-8550 Japan

information source.

The reader might also conceive that this problem could be solved by a straightforward combination of the coding for intrinsic randomness [33] and that for the original secure multiplex coding [22], [27]. We emphasize that this is false. We cannot recover the original secret messages from a codeword generated by an intrinsic randomness encoder, and a new technique must be deployed to remove the independence and uniform assumption on the multiple secret messages. One of the main contributions of this paper is to remove that assumption. In order to treat the non-uniform and dependent case, we need a generalization of the channel resolvability. Hence, this paper also studies a generalization of the channel resolvability problem [13], [15].

Even after we solve the above problem by a generalization of the channel resolvability problem, the security of S_i depends on the randomness and the dependence of the remaining messages $S_1, \dots, S_{i-1}, S_{i+1}, \dots, S_T$ on S_i . This dependence causes another difficulty in the asymptotic formulation of SMC. That is, we need to characterize the randomness and the dependence in the asymptotic setting. For this purpose, we introduce several kinds of asymptotic conditional uniformity conditions and study their properties. In addition to this, for the case when the channel is unknown, we also treat universal coding for the secure multiplex coding [22]. Further, as a byproduct, we obtain source-channel universal coding for the broadcast channel with confidential messages [9]. We divide the introductory section to six subsections.

Finally, we should explain the assumptions for our probability spaces. In the main body, we assume that all of probability spaces are finite sets. However, our result can be extended to the case of measurable spaces except for the contents in Sections VIII-A, XII, and XIII. This generalization contains the case of continuous sets. In Appendix D, we summarize how to generalize our results to the case of measurable spaces. As a byproduct, we show the strong security for the Gaussian channel.

B. Generalization of the Channel Resolvability

For a given channel W with input alphabet \mathcal{X} and output alphabet \mathcal{Y} , and given information source X on \mathcal{X} , Han and Verdú [13] considered to find a coding $f : \mathcal{A} \rightarrow \mathcal{X}$ and a random variable A such that the distributions of $W(f(A))$ is close to $W(X)$ with respect to the variational distance or the normalized divergence, and evaluated the minimum resolution of A to make the variational distance or the normalized divergence asymptotically zero. In their problem formulation, one can choose the randomness A used to simulate the channel output distribution.

In this paper, we shall consider the situation in which we are given a channel W , an information source X , and randomness A and asked to find coding $f : \mathcal{A} \rightarrow \mathcal{X}$ such that $W(f(A))$ is as close as possible to $W(X)$ with respect to unnormalized divergence. We shall study how close $W(f(A))$ can be to $W(X)$ in Theorems 14 and 17 in Section VI. Hence, this problem can be regarded as a generalization of channel resolvability because this problem contains the original channel resolvability as a special case in the above sense.

C. Asymptotic Conditional Uniformity

In Subsection VIII-A, in order to characterize the randomness and the dependence of the messages $S_1, \dots, S_{i-1}, S_{i+1}, \dots, S_T$ on the other message S_i asymptotically, we introduce three asymptotic conditional uniformity conditions. Then, we can characterize what a conditional distribution of the messages $S_1, \dots, S_{i-1}, S_{i+1}, \dots, S_T$ has a similar performance to the conditionally uniform distribution when we apply SMC. We summarize the relations among those conditions as Theorem 29. In particular, in Appendix C, we show that two introduced asymptotic conditional uniformity conditions are equivalent. Hence, we essentially have two different conditional uniformity conditions, namely, the weaker and the stronger asymptotic conditional uniformity conditions.

In Subsection VIII-B, we give sufficient conditions for the Slepian-Wolf compression so that the compressed data satisfies these asymptotic conditional uniformity conditions. For the stationary ergodic sources, we show the existence of a sequence of Slepian-Wolf codes whose compressed data satisfies the weaker asymptotic conditional uniformity conditions (Theorem 30 and Remark 31). Also for the i.i.d. sources, we show the existence of a sequence of Slepian-Wolf codes whose compressed data satisfies the stronger asymptotic conditional uniformity conditions (Theorem 32 and Remark 33).

D. Secure Multiplex Coding

Here, we explain the detail of our contributions to SMC. As is explained above, we have to realize the security of S_i when the remaining messages $S_1, \dots, S_{i-1}, S_{i+1}, \dots, S_T$ are not uniform and are dependent on the message S_i . In order to solve this problem, we employ our generalized channel resolvability coding in Theorems 14 and 17. Then, we can construct coding for a wire-tap channel that can ensure the secrecy of message against the eavesdropper Eve when the dummy message used by the encoder is non-uniform and statistically dependent on the secret message that has to be kept secret from Eve. We apply our generalized channel resolvability coding to the above SMC case. Hence, we can remove the independence and uniform assumption on the multiple secret messages while the original paper [22] by Yamamoto et al. and the previous paper [27] by the present authors assumed the independence and the uniformity of the multiple secret messages.

Indeed, Yamamoto et al. [22] treated only the secrecy of each message S_i , and did not evaluate the information leakage of multiple messages S_{i_1}, \dots, S_{i_n} to Eve, and the present authors analyzed such information leakage in [27]. The present authors also generalized coding in [27] so that Alice's encoder can support the common message S_0 to both Bob and Eve. The present authors also characterized the achievable information leakage rate in [27]. Those enhancements are retained in this paper.

In Section VII, we shall give two code constructions for SMC. The first construction given in Subsection VII-B is a simple application of channel resolvability coding in Theorem 14. Although it achieves the capacity region when there is no

common message, it is insufficient to fully prove the capacity region. In Subsection VII-C, to overcome this defect, we propose the second construction given in Theorem 17, which is based on another type of the channel resolvability coding. By using these constructions, we shall evaluate the decoding error probability and the mutual information to Eve in Section VII in single-shot setting in the sense of [34].

In Section IX we formulate the capacity region of SMC, analyze the asymptotic performance of two constructions, and prove that the second construction achieves the capacity region of SMC. The capacity region is defined based on the weaker asymptotic conditional uniformity condition given in Definition 36. In Section X, we shall prove that the mutual information to Eve converges to zero when the normalized mutual information to Eve converges to zero under the stronger asymptotic conditional uniformity given in Definition 28. The convergence is so-called the strong security [28]. In Subsection X-B, we also derive the exponent of the mutual information to Eve. The relation between our results and the paper [22] is explained as (145).

Section XI addresses a more practical issue. In Theorem 22 of Section VII, we show that we can have an upper bound of mutual information between multiple secret messages and Eve's received signal, by attaching randomly chosen group homomorphisms satisfying Condition 15 to *any* given error-correcting code for channels with single sender and single receiver or the broadcast channel with degraded message sets [23]. However, the upper bound in Theorem 22 becomes difficult to be computed when the error-correcting code is not given by the standard random coding in information theory. In Section XI, we shall construct more practical codes by combining the construction of Section VII with an arbitrary given error-correcting code. Under these codes, we shall give two upper bounds on the leaked mutual information that can be computed easily in practice. Section XI gives enhancement of our earlier proceeding paper [18].

E. Universal Coding

Universal coding is construction of encoder and decoder that do not use the statistical knowledge on the underlying information system (usually channel and/or source) [8]. In Section XII we shall give a construction of SMC universal to channel. The basic idea in Section XII is to combine the construction in Section VII with the universal coding using constant-type codes for the broadcast channel with degraded messages sets (BCD) in [24], while in Sections VII–X the superposition random coding in [23] is used as their error-correcting mechanism. The exponent given in Section XII is better than that given in our earlier proceeding paper [19].

Channel-universal coding for BCC had not been studied before [19], and coding for BCC can be regarded as a special case of SMC while Muramatsu et al. [29] treat channel-universal coding for wire-tap channel independently of [19]. In Section XII and [19] we consider SMC universal to channel, but its universality to the source is not considered. In Section XIII we give a coding for BCC universal to both channel and source. Its channel-universality is realized by the same

principle as Section XII and [19]. The exponent given in Section XII is also greater than that given in our earlier proceeding paper [19].

In Section XIV, we compare the exponent of leaked information given in Sections XII and XIII and that given in Subsection X-B. As a result, we show that the exponent in Sections XII and XIII is greater than one of exponents in Subsection X-B, which is the same as that in [19]. We also derive the equality condition.

F. Organization of This Paper

The outline of this paper is given as follows. First, we prepare notations used in this paper in Section II. Second, we prepare information quantities and their properties used in this paper in Section III. Then, we review the formulation and existing results of BCC in Subsection IV-A. We give its reformulation for the dependent and non-uniform messages case in Subsection IV-B. This new formulation is essential in the later discussion for SMC with dependent and non-uniform multiple messages. In Subsection V-A, we review the formulation and existing results of BCD as a special case of BCC, which will be used for our codes of SMC. In Subsection V-B, we review Körner and Sgarro [24]'s result for universal code for BCD, which will be used for our construction of universal codes for SMC and BCC. In Section VI, we proceed to generalization of channel resolvability, which is a key idea of the paper and is used for codes of SMC and universal codes for SMC and BCC. Section VII introduces SMC with the single-shot setting. Section VIII introduces three asymptotic conditional uniformity conditions. Based on these conditions, Sections IX–XI treats SMC with the asymptotic setting, as is explained in Subsection I-D. In Section XII, combining the discussion of Subsections V-A and VII-D, we propose universal coding for SMC by using Körner and Sgarro [24]'s universal coding for BCD. In Section XIII, we propose source-channel universal coding for BCC. Appendices are devoted for several additionally required discussions for asymptotic conditional uniformity conditions. This paper contains two types of descriptions for each topics, i.e., the single-shot description [34] and the n -fold description. Formulations and many coding theorems are given with the single-shot description. The definitions of capacity regions are given in the n -fold description.

II. NOTATION IN THIS PAPER

\mathcal{X} denotes the channel input alphabet and \mathcal{Y} (resp. \mathcal{Z}) denotes the channel output alphabet to Bob (resp. Eve). We assume that \mathcal{X} , \mathcal{Y} , and \mathcal{Z} are finite unless otherwise stated. We denote the conditional probability of the channel to Bob and Eve by $P_{YZ|X}$. Then, taking the marginal distribution, we denote the conditional probability of the channel to Bob (resp. Eve) by $P_{Y|X}$ (resp. $P_{Z|X}$). Also, we denote the distribution of the random variable X by P_X .

We denote the uniform distribution on Ω by $P_{\text{mix},\Omega}$. When Ω is a subset of $\mathcal{X} \times \mathcal{Y}$, $P_{\text{mix},\Omega}$ is a joint distribution for the random variables X and Y . We denote the marginal distribution of $P_{\text{mix},\Omega}$ for the random variable X and the random variable Y

by $P_{X,\text{mix},\Omega}$ and $P_{Y,\text{mix},\Omega}$, respectively. Further, the conditional distribution on the random variable X conditioned to the other random variable Y is denoted by $P_{X|Y,\text{mix},\Omega}$, i.e.,

$$P_{X|Y,\text{mix},\Omega}(x|y) = P_{X|Y=y,\text{mix},\Omega}(x) := \frac{P_{\text{mix},\Omega}(x,y)}{P_{Y,\text{mix},\Omega}(y)} \quad (1)$$

for $x \in \mathcal{X}$ and $y \in \mathcal{Y}$. We denote the support of the distribution P_X by $\text{supp}(P_X)$. Given a joint distribution P_{XY} , we define the distribution $P_{X|Y=y}$ on \mathcal{X} by $P_{X|Y=y}(x) := P_{X|Y}(x|y)$. When we need to treat another distribution of the same random variables X and Y , we denote it by Q_{XY} . This is because it is crucial to consider several distributions on the same probability space in this paper¹. In this case, we denote the marginal distribution over \mathcal{X} by Q_X , and the conditional distribution by $Q_{X|Y}$. We also define the distribution $Q_{X|Y=y}$ on \mathcal{X} by $Q_{X|Y=y}(x) := Q_{X|Y}(x|y)$.

When we have to treat more than two distributions on \mathcal{X} , \mathcal{Y} , and \mathcal{Z} , the above notation is not useful. In this case, we consider the set $\mathcal{P}(\mathcal{X})$ of probability distributions on \mathcal{X} or the set $\mathcal{W}(\mathcal{X}, \mathcal{Y})$ of conditional probability distributions from \mathcal{X} to \mathcal{Y} , which are mathematically equivalent to probability transition matrices. When the output alphabet of the channel is given as a product set $\mathcal{Y} \times \mathcal{Z}$, the alphabet is written by $\mathcal{W}(\mathcal{X}, \mathcal{Y} \times \mathcal{Z})$. For any probability transition matrix $W \in \mathcal{W}(\mathcal{X}, \mathcal{Y} \times \mathcal{Z})$, W_x expresses the output distribution when the input X is x . When we focus on the random variable Y , we use the notation $W_x^Y(y) := \sum_{z \in \mathcal{Z}} W_x(y, z)$.

In the following, we treat an arbitrary probability transition matrix $W \in \mathcal{W}(\mathcal{X}, \mathcal{Y})$. Given a subset $\Omega \subset \mathcal{X}$, we define the restriction $W|_{\Omega} \in \mathcal{W}(\Omega, \mathcal{Y})$ by $W|_{\Omega}(y|x) = W(y|x)$ for $x \in \Omega$ and $y \in \mathcal{Y}$. We often employ another probability transition matrix Ξ from \mathcal{V} to \mathcal{X} . We define the probability transition matrix from \mathcal{V} to \mathcal{Y} by $W \circ \Xi_v(y) := \sum_{x \in \mathcal{X}} W_x(y) \Xi_v(x)$ for $v \in \mathcal{V}$ and $y \in \mathcal{Y}$. When a probability distribution P on \mathcal{X} is given, we define the distribution on \mathcal{Y} by $W \circ P(y) := \sum_{x \in \mathcal{X}} W_x(y) P(x)$ for $y \in \mathcal{Y}$. When we need the joint distribution on $\mathcal{X} \times \mathcal{Y}$, we use the notation $W \times P(x, y) := W_x(y) P(x)$ for $x \in \mathcal{X}$ and $y \in \mathcal{Y}$ as [6]. Similarly, when a distribution P_{XV} on $\mathcal{X} \times \mathcal{V}$ is given, we use the notation $W \times P_{XV}(v, x, y) := W_x(y) P_{XV}(x, v)$ for $v \in \mathcal{V}$, $x \in \mathcal{X}$, and $y \in \mathcal{Y}$.

When a function $f : \mathcal{V} \rightarrow \mathcal{X}$ is given and a random variable V taking the values in \mathcal{V} obeys the distribution P_V , we can define the random variable $f(V)$ taking the values in \mathcal{X} . The random variable $f(V)$ takes the value x with probability $\sum_{v \in f^{-1}(x)} P_V(v)$. We also use the same symbol $f : \mathcal{V} \rightarrow \mathcal{X}$ to denote the probability transition matrix from \mathcal{V} to \mathcal{X} , in which, the output value is deterministically determined by the input. Then, $W \circ f$ is a stochastic mapping \mathcal{V} to \mathcal{Y} , and we have

$$(W \circ f)(y|v) = W(y|f(v)) \quad (2)$$

¹Recently, the meta converse theorem was introduced for the channel coding in [48], [50]. In the meta converse theorem, it is the key point to optimize the choice of the distribution on the output alphabet and we usually denote the distribution different from the marginal distribution by Q [49], [50]. Also, another recent paper [51] adopts this notation for optimizing the distribution. This kind notation becomes more popular, recently.

for $v \in \mathcal{V}$ and $y \in \mathcal{Y}$. Given a probability transition matrix $W' \in \mathcal{W}(\mathcal{U}, \mathcal{V})$, we define $f \circ W' \in \mathcal{W}(\mathcal{U}, \mathcal{X})$ by

$$(f \circ W')(x|u) := \sum_{v \in f^{-1}(x)} W'(v|u) \quad (3)$$

for $x \in \mathcal{X}$ and $u \in \mathcal{U}$. As a special case, given a distribution Q on \mathcal{V} , $f \circ Q$ is defined as a distribution on \mathcal{X} in the following way.

$$(f \circ Q)(x) := \sum_{v \in f^{-1}(x)} Q(v). \quad (4)$$

Remember that W_x denotes the output distribution on the output alphabet \mathcal{Y} with input x . Then, W_X is the random variable taking its values on the output distributions on \mathcal{Y} . Given a real valued function g of distributions on \mathcal{Y} , we regard $g(W_X)$ as a random variable taking the value $g(W_x)$ with the probability $P_X(x)$. Hence, we obtain

$$\mathbf{E}_X g(W_X) = \sum_x P_X(x) g(W_x),$$

where \mathbf{E}_X denotes the expectation concerning X .

Given two random variables X and Y , for a real valued function h on $\mathcal{X} \times \mathcal{Y}$, we regard $\mathbf{E}_{X|Y} h(X, Y)$ as a random variable taking the value $\mathbf{E}_{X|Y=y} h(X, y)$ with the probability $P_Y(y)$. In order to identify an information quantity, e.g., mutual information $I(X; Y)$ and the Shannon entropy $H(X)$, we sometimes need to specify the distribution P of interest. In such a case, we use the notations $I(X; Y)[P]$ and $H(X)[P]$ for identifying what distribution is considered.

Further, in this paper, we discuss our codes and their performances in the single-shot setting[34] when their descriptions do not require their asymptotic discussions. However, in several parts, we need to treat n -fold memoryless extensions when we discuss their asymptotic performances. Hence, we need to prepare the notations for n -fold independent and identical distributions and n -fold memoryless extensions of given channels. For a given probability distributions Q and P_X of the random variable X on \mathcal{X} , we denote their n -fold independent and identical distributions by Q^n and P_X^n .

When we consider the random variables on \mathcal{X}^n , even if they do not obey the independent and identical distributions, we denote the random variables by X^n and denote their distributions by P_{X^n} . However, when we consider a general sequence of random variables those take values not in the product sets \mathcal{X}^n but in general sets \mathcal{X}_n , we denote the random variables by X_n and denote their distributions by P_{X_n} . Similarly, for a given probability transition matrices W and $P_{Y|X}$ from \mathcal{X} to \mathcal{Y} , we denote their n -fold memoryless extensions by W^n and $P_{Y|X}^n$.

We also denote the set of positive real numbers by \mathbf{R}^+ , and denote the set of non-negative real numbers by $\mathbf{R}_{\geq 0}$.

III. INFORMATION QUANTITIES

In this paper, to evaluate the secrecy and the decoding error probabilities, we employ several information quantities. For distributions P_A on \mathcal{A} and P_{AB} on $\mathcal{A} \times \mathcal{B}$, we define Rényi

entropy and conditional Rényi entropy

$$H_{1+\rho}(A) := -\frac{1}{\rho} \log \sum_a P_A(a)^{1+\rho}$$

$$H_{1+\rho}(A|B) := -\frac{1}{\rho} \log \sum_{a,b} P_B(b)P_{A|B=b}(a)^{1+\rho}.$$

$H_1(A)$ and $H_1(A|B)$ are defined to be $H(A)$ and $H(A|B)$. Then, we have several important properties for Rényi entropy and conditional Rényi entropy. Since $\rho \mapsto \rho H_{1+\rho}(A)$, $\rho \mapsto \rho H_{1+\rho}(A|B)$ are concave and $\lim_{\rho \rightarrow 0} \rho H_{1+\rho}(A) = \lim_{\rho \rightarrow 0} \rho H_{1+\rho}(A|B) = 0$, we have

$$H_{1+\rho'}(A) \leq H_{1+\rho}(A), \quad H_{1+\rho'}(A|B) \leq H_{1+\rho}(A|B) \quad (5)$$

for $0 \leq \rho \leq \rho'$.

Similarly, as is shown in [17], we have the following proposition for the function

$$\psi(\rho|Q||P) := \log \sum_a Q(a)^{1+\rho} P(a)^{-\rho}. \quad (6)$$

Proposition 1: [17] The function $\psi(\rho|Q||P)$ satisfies the following properties:

- (1) $\rho \mapsto \psi(\rho|Q||P)$ is convex.
- (2) $\psi(0|Q||P) = 0$.
- (3) $\frac{d}{d\rho} \psi(\rho|Q||P)|_{\rho=0} = D(Q||P)$.
- (4) The relations

$$D(Q||P) := \sum_a P(a) \log \frac{P(a)}{Q(a)} = \lim_{\rho \rightarrow +0} \frac{\psi(\rho|Q||P)}{\rho} \leq \frac{\psi(\rho|Q||P)}{\rho} \quad (7)$$

hold for $0 < \rho^2$.

For a given channel W from \mathcal{X} to \mathcal{Y} , we define the function [17]:

$$\psi(\rho|W, P_X) := \log \sum_x P_X(x) e^{\psi(\rho|W_x||W \circ P_X)}. \quad (8)$$

When the channel is written as $P_{Z|L}$, $\psi(\rho|W, P)$ can be rewritten as follows.

$$\psi(\rho|P_{Z|L}, P_L) = \log \sum_z \sum_\ell P_L(\ell) P_{Z|L}(z|\ell)^{1+\rho} P_Z(z)^{-\rho}. \quad (9)$$

This quantity is extended as

$$\psi(\rho|P_{Z|V}, P_{V|U}, P_U) := \log \sum_u P_U(u) \sum_v P_{V|U}(v|u) \sum_z P_{Z|V}(z|v)^{1+\rho} P_{Z|U}(z|u)^{-\rho}. \quad (10)$$

for conditional distributions $P_{Z|V}$, $P_{V|U}$ and a distribution P_U . Also, we introduce the following functions as in [17].

$$E_0(\rho|P_{Z|L}, P_L) := \log \sum_z \left(\sum_\ell P_L(\ell) (P_{Z|L}(z|\ell))^{1/(1-\rho)} \right)^{1-\rho}, \quad (11)$$

$$E_0(\rho|P_{Z|V}, P_{V|U}, P_U) := \log \sum_u P_U(u) \sum_z \left(\sum_v P_{V|U}(v|u) (P_{Z|V}(z|v))^{1/(1-\rho)} \right)^{1-\rho}. \quad (12)$$

²Item (4) was not directly given in [17]. However, it can be shown by the combination of other items.

Observe that E_0 is essentially Gallager's function E_0 [12]. As can be easily shown, these quantities satisfy the additivity as follows [17], [12].

$$\psi(\rho|P_{Z|L}^n, P_L^n) = n\psi(\rho|P_{Z|L}, P_L) \quad (13)$$

$$\psi(\rho|P_{Z|V}^n, P_{V|U}^n, P_U^n) = n\psi(\rho|P_{Z|V}, P_{V|U}, P_U) \quad (14)$$

$$E_0(\rho|P_{Z|L}^n, P_L^n) = nE_0(\rho|P_{Z|L}, P_L) \quad (15)$$

$$E_0(\rho|P_{Z|V}^n, P_{V|U}^n, P_U^n) = nE_0(\rho|P_{Z|V}, P_{V|U}, P_U) \quad (16)$$

Then, we have the following proposition.

Proposition 2: [12], [17] We have the following five items for fixed $0 < \rho < 1$ and fixed conditional distribution $P_{Z|L}$.

- (1) The function $\rho \mapsto E_0(\rho|P_{Z|L}, P_L)$ is convex for a given distribution P_L [12].
- (2) $\exp(E_0(\rho|P_{Z|L}, P_L))$ is concave with respect to P_L [17, Lemma 1].
- (3) The relation $\psi(\rho|P_{Z|L}, P_L) \leq E_0(\rho|P_{Z|L}, P_L)$, i.e.,

$$\exp(\psi(\rho|P_{Z|L}, P_L)) \leq \exp(E_0(\rho|P_{Z|L}, P_L)) \quad (17)$$

holds for any distribution P_L of L [17, (16)].

- (4) The relation

$$\lim_{\rho \rightarrow 0} \frac{\psi(\rho|P_{Z|L}, P_L)}{\rho} = \lim_{\rho \rightarrow 0} \frac{E_0(\rho|P_{Z|L}, P_L)}{\rho} = I(Z; L) \quad (18)$$

holds for a distribution P_L [17, Section III][12].

Lemma 3: When two distributions Q_L and P_L of L satisfy $P_L(\ell) \leq C_1 Q_L(\ell)$ for any ℓ with given constants $C_1 \geq 1$ and $0 < \rho < 1$, we have

$$\exp(E_0(\rho|P_{Z|L}, P_L)) \leq C_1 \exp(E_0(\rho|P_{Z|L}, Q_L)). \quad (19)$$

Proof: (19) can be shown as follows.

$$\begin{aligned} \exp(E_0(\rho|P_{Z|L}, P_L)) &= \sum_z \left(\sum_\ell P_L(\ell) (P_{Z|L}(z|\ell))^{1/(1-\rho)} \right)^{1-\rho} \\ &\leq \sum_z \left(\sum_\ell C_1 Q_L(\ell) (P_{Z|L}(z|\ell))^{1/(1-\rho)} \right)^{1-\rho} \\ &\leq C_1^{1-\rho} \sum_z \left(\sum_\ell Q_L(\ell) (P_{Z|L}(z|\ell))^{1/(1-\rho)} \right)^{1-\rho} \\ &= C_1^{1-\rho} \exp(E_0(\rho|P_{Z|L}, Q_L)) \leq C_1 \exp(E_0(\rho|P_{Z|L}, Q_L)). \end{aligned}$$

As a generalization of Item (4) of Proposition 2, we have the following lemma.

Lemma 4: The relation

$$\lim_{\rho \rightarrow 0} \frac{\psi(\rho|P_{Z|V}, P_{V|U}, P_U)}{\rho} = \lim_{\rho \rightarrow 0} \frac{E_0(\rho|P_{Z|V}, P_{V|U}, P_U)}{\rho} = I(Z; V|U) \quad (20)$$

holds for a distribution P_U , and conditional distributions $P_{Z|V}$ and $P_{V|U}$.

Proof: Due to (18), we have

$$\begin{aligned} e^{\psi(\rho|P_{Z|V}, P_{V|U}, P_U)} &= \sum_u P_U(u) 1 + \rho I(Z; V|U = u) + o(\rho) \\ &= 1 + \rho I(Z; V|U) + o(\rho). \end{aligned}$$

Taking the logarithm, we obtain $\lim_{\rho \rightarrow 0} \frac{\psi(\rho P_{ZV}, P_{V|U}, P_U)}{E_0(\rho P_{ZV}, P_{V|U}, P_U)} = I(Z; V|U)$. Similarly, we can show $\lim_{\rho \rightarrow 0} \frac{E_0(\rho P_{ZV}, P_{V|U}, P_U)}{\rho} = I(Z; V|U)$. ■

Considering the Legendre transforms, we define

$$\tilde{E}^\psi(R, P_{Z,V,U}) := \max_{0 \leq \rho \leq 1} \rho R - \psi(\rho P_{Z|V}, P_{V|U}, P_U), \quad (21)$$

$$\tilde{E}^{E_0}(R, P_{Z,V,U}) := \max_{0 \leq \rho \leq 1} \rho R - E_0(\rho P_{Z|V}, P_{V|U}, P_U). \quad (22)$$

Taking the maximum, we define

$$\begin{aligned} E_{0,\max}(\rho P_{Z|V}) &:= \max_{P_V} E_0(\rho P_{Z|V}, P_V) \\ &= \log \max_{P_V} \sum_z \left(\sum_v P_V(v) P_{Z|V}(z|v)^{\frac{1}{1-\rho}} \right)^{1-\rho} \\ &= \max_{P_{V,U}} E_0(\rho P_{Z|V}, P_{V|U}, P_U). \end{aligned} \quad (23)$$

Lemma 5: The function $\rho \mapsto E_{0,\max}(\rho P_{Z|V})$ is convex.

Proof: Given convex functions $x \mapsto f_i(x)$, the function $x \mapsto \max_i f_i(x)$ is also convex. Hence, the item (1) of Proposition 2 yields the desired argument. ■

Next, for $\bar{W}^Z \in \mathcal{W}(\mathcal{V}, \mathcal{Z})$, we consider a different information quantity \tilde{E}^I :

$$\begin{aligned} \tilde{E}^I(R, \bar{W}^Z \times Q_{VU}) \\ := \min_{W^Z \in \mathcal{W}(\mathcal{U} \times \mathcal{V}, \mathcal{Z})} \left(D(W^Z \| \bar{W}^Z | Q_{VU}) \right. \\ \left. + [R - I(V; Z|U)[W^Z \times Q_{VU}]_+ \right). \end{aligned} \quad (24)$$

Due to Item (3) of Proposition 2, we have

$$\tilde{E}^\psi(R, \bar{W}^Z \times Q_{VU}) \geq \tilde{E}^{E_0}(R, \bar{W}^Z \times Q_{VU}). \quad (25)$$

In this paper, we will derive the following relations:

$$\tilde{E}^I(R, \bar{W}^Z \times Q_{VU}) \geq \tilde{E}^{E_0}(R, \bar{W}^Z \times Q_{VU}) \quad (26)$$

and

$$\begin{aligned} \min_{Q_V} \tilde{E}^I(R, \bar{W}^Z \times Q_V) &= \min_{Q_V} \tilde{E}^{E_0}(R, \bar{W}^Z \times Q_V) \\ &= \max_{\rho \in [0,1]} \rho R - E_0(\rho \bar{W}^Z) \end{aligned} \quad (27)$$

as Theorems 67 and 80 in Section XIV, respectively.

Similar to \tilde{E}^I , we introduce the following quantities for $W^Y \in \mathcal{W}(\mathcal{V}, \mathcal{Y})$ and $W^Z \in \mathcal{W}(\mathcal{V}, \mathcal{Z})$

$$\begin{aligned} \hat{E}^b(R_p, R_c, \tilde{W}^Y \times Q_{VU}) \\ := \min \left([I(VU; Y)[\tilde{W}^Y \times Q_{U,V}] - R_p - R_c]_+, \right. \\ \left. [I(V; Y|U)[\tilde{W}^Y \times Q_{U,V}] - R_p]_+ \right), \end{aligned} \quad (28)$$

$$\begin{aligned} \tilde{E}^b(R_p, R_c, W^Y \times Q_{VU}) \\ := \min_{\tilde{W}^Y \in \mathcal{W}(\mathcal{U} \times \mathcal{V}, \mathcal{Y})} D(\tilde{W}^Y \| W^Y | Q_{VU}) + \hat{E}^b(R_p, R_c, \tilde{W}^Y \times Q_{VU}), \end{aligned} \quad (29)$$

$$\begin{aligned} \tilde{E}^e(R_c, W^Z \times Q_U) \\ := \min_{\tilde{W}^Z \in \mathcal{W}(\mathcal{U} \times \mathcal{V}, \mathcal{Z})} D(\tilde{W}^Z \| W^Z | Q_{VU}) + [I(U; Z)[\tilde{W}^Z \times Q_{VU}] - R_c]_+, \end{aligned} \quad (30)$$

where $D(\tilde{W}^Y \| W^Y | Q_{VU})$ is defined for $\tilde{W}^Y, W^Y \in \mathcal{W}(\mathcal{V}, \mathcal{Y})$ as

$$D(\tilde{W}^Y \| W^Y | Q_{VU}) := \sum_{u,v} Q_{VU}(u, v) D(\tilde{W}_{u,v}^Y \| W_v^Y). \quad (31)$$

In the above definition, W^Y and W^Z are treated as elements of $\mathcal{W}(\mathcal{U} \times \mathcal{V}, \mathcal{Y})$ and $\mathcal{W}(\mathcal{U} \times \mathcal{V}, \mathcal{Z})$, respectively.

IV. BROADCAST CHANNELS WITH CONFIDENTIAL MESSAGES

A. Review of Existing Results

First, we give a formulation of broadcast channels with confidential messages with single shot setting[34]. Let Alice, Bob, and Eve be as defined in Section I. \mathcal{X} denotes the channel input alphabet and \mathcal{Y} (resp. \mathcal{Z}) denotes the channel output alphabet to Bob (resp. Eve). We assume that \mathcal{X} , \mathcal{Y} , and \mathcal{Z} are finite unless otherwise stated.

We denote the conditional probability of the channel to Bob (resp. Eve) by $P_{Y|X}$ (resp. $P_{Z|X}$). The purpose of broadcast channels with confidential messages is the following. (1) Alice reliably sends the common message E to Bob and Eve. (2) Alice confidentially and reliably sends the secret message S to Bob. Here, we denote the sets of the common messages and the secret messages by \mathcal{E} and \mathcal{S} . Our code is given by Alice's stochastic encoder φ_a from $\mathcal{S} \times \mathcal{E}$ to \mathcal{X} , Bob's deterministic decoder $\varphi_b : \mathcal{Y} \rightarrow \mathcal{S} \times \mathcal{E}$ and Eve's deterministic decoder $\varphi_e : \mathcal{Z} \rightarrow \mathcal{E}$. The triple $\varphi = (\varphi_a, \varphi_b, \varphi_e)$ is called a code for broadcast channels with confidential messages. Then, when the common message E and the secret message S obey the distribution $P_{S,E}$, the performance is evaluated by the following quantities. (1) The sizes of the sets of the common messages and the secret messages, i.e., $|\mathcal{E}|$ and $|\mathcal{S}|$. (2) Bob's decoding error probability $P_b[P_{Y|X}, \varphi, P_{S,E}]$, which is the probability $\Pr\{(S, E) \neq \varphi_b(Y)\}$ under the distribution $(P_{Y|X} \circ \varphi_a) \times P_{S,E}$. (3) Eve's decoding error probability $P_e[P_{Y|X}, \varphi, P_{S,E}]$, which is the probability $\Pr\{E \neq \varphi_e(Z)\}$ under the distribution $(P_{Z|X} \circ \varphi_a) \times P_{S,E}$. (4) Eve's uncertainty $H(S|Z)[P_{Z|X}, \varphi_a, P_{S,E}]$, which is the conditional entropy $H(S|Z)$ under the distribution $(P_{Z|X} \circ \varphi_a) \times P_{S,E}$. Since these quantities are functions of the channel and the code, such dependencies are denoted by the symbol $[P_{Y|X}, \varphi, P_{S,E}]$ in the above notation. Instead of $H(S|Z)[P_{Z|X}, \varphi_a, P_{S,E}]$, we sometimes treat (5) leaked information $I(S; Z)[P_{Z|X}, \varphi_a, P_{S,E}]$, which is the mutual information $I(S; Z)$ under the distribution $(P_{Z|X} \circ \varphi_a) \times P_{S,E}$.

We sometimes need to evaluate the error probability when S and/or E is fixed. In such a case, we denote it by $P_b[P_{Y|X}, \varphi, P_{E|S=s}]$, $P_b[P_{Y|X}, \varphi, S = s, E = e]$, and $P_e[P_{Y|X}, \varphi, P_{S|E=e}]$.

Now, we review the asymptotic formulation of broadcast channels with confidential messages with the n -fold discrete memoryless extension when both of the common messages and the secret messages are subject to uniform distributions. The set \mathcal{S}_n denotes the set of the confidential message and \mathcal{E}_n does the set of the common message when the block coding of length n is used. We shall define the achievability of a rate triple (R_1, R_e, R_0) , where R_0 and R_1 are the rates of the common and confidential messages, and R_e is the entropy rate conditioned with Eve's random variable for the

confidential message. For the notational convenience, we fix the base of logarithm, including one used in entropy and mutual information, to the base of natural logarithm.

Definition 6: [9] The rate triple (R_1, R_e, R_0) is said to be *achievable* for the information leakage rate criterion if the following condition holds. The size of the sets of the common and confidential messages are $|\mathcal{E}_n| = e^{nR_0}$ and $|\mathcal{S}_n| = e^{nR_1}$. The common and confidential messages are subject to the uniform and independent distribution on \mathcal{S}_n and \mathcal{E}_n . There exists a sequence of the codes $\varphi_n = (\varphi_{a,n}, \varphi_{b,n}, \varphi_{e,n})$, i.e., Alice's stochastic encoder $\varphi_{a,n}$ from $\mathcal{S}_n \times \mathcal{E}_n$ to \mathcal{X}^n , Bob's deterministic decoder $\varphi_{b,n} : \mathcal{Y}^n \rightarrow \mathcal{S}_n \times \mathcal{E}_n$ and Eve's deterministic decoder $\varphi_{e,n} : \mathcal{Z}^n \rightarrow \mathcal{E}_n$ such that

$$\begin{aligned} \lim_{n \rightarrow \infty} P_b[P_{Y|X}^n, \varphi_n, P_{\text{mix}, \mathcal{S}_n, \mathcal{E}_n}] &= 0 \\ \lim_{n \rightarrow \infty} P_e[P_{Z|X}^n, \varphi_n, P_{\text{mix}, \mathcal{S}_n, \mathcal{E}_n}] &= 0 \\ \liminf_{n \rightarrow \infty} \frac{H(S_n | Z^n)[P_{Y|X}^n, \varphi_{a,n}, P_{\text{mix}, \mathcal{S}_n, \mathcal{E}_n}]}{n} &\geq R_e. \end{aligned}$$

The capacity region with the information leakage rate criterion of the BCC is the closure of the achievable rate triples for the information leakage rate criterion.

Theorem 7: [9] The capacity region with the information leakage rate criterion of the BCC is given by the set of R_0, R_1 and R_e such that there exists a Markov chain $U \rightarrow V \rightarrow X \rightarrow YZ$ and

$$\begin{aligned} R_1 + R_0 &\leq I(V; Y|U) + \min[I(U; Y), I(U; Z)], \\ R_0 &\leq \min[I(U; Y), I(U; Z)], \\ R_e &\leq I(V; Y|U) - I(V; Z|U), \\ R_e &\leq R_1. \end{aligned}$$

As described in [25], U can be regarded as the common message, V the combination of the common and the confidential messages, and X the transmitted signal.

In this paper, we treat the source-channel universal coding for BCC, in which, we guarantee the security independently of the choice of the source distribution. While the lower bound of the above conditional entropy $H(S_n | Z^n)[P_{Y|X}^n, \varphi_{a,n}, P_{\mathcal{S}_n, \mathcal{E}_n}]$ depends on the the source distribution $P_{\mathcal{S}_n, \mathcal{E}_n}$, we can find an upper bound of mutual information that does not depend on the source distribution, as is shown in Section XIII. As a preparation for the above source-channel universal coding for BCC, we propose another type of capacity region for the uniform and independent distributed case while the non-uniform and dependent case will be treated latter.

Definition 8: The rate triple (R_1, R_l, R_0) is said to be *achievable* for the leaked information criterion if the following conditions hold. In this notation, R_1, R_l , and R_0 denote the rates of the confidential message, the leaked information, and the common message, respectively. The size of the sets of the common and confidential messages are $|\mathcal{E}_n| = e^{nR_0}$ and $|\mathcal{S}_n| = e^{nR_1}$, and the common and confidential messages are subject to the uniform and independent distribution on \mathcal{S}_n and \mathcal{E}_n . There exists a sequence of the codes $\varphi_n = (\varphi_{a,n}, \varphi_{b,n}, \varphi_{e,n})$, i.e., Alice's stochastic encoder $\varphi_{a,n}$ from $\mathcal{S}_n \times \mathcal{E}_n$ to \mathcal{X}^n , Bob's deterministic decoder $\varphi_{b,n} : \mathcal{Y}^n \rightarrow \mathcal{S}_n \times \mathcal{E}_n$ and Eve's

deterministic decoder $\varphi_{e,n} : \mathcal{Z}^n \rightarrow \mathcal{E}_n$ such that

$$\begin{aligned} \lim_{n \rightarrow \infty} P_b[P_{Y|X}^n, \varphi_n, P_{\text{mix}, \mathcal{S}_n, \mathcal{E}_n}] &= 0 \\ \lim_{n \rightarrow \infty} P_e[P_{Z|X}^n, \varphi_n, P_{\text{mix}, \mathcal{S}_n, \mathcal{E}_n}] &= 0 \\ \limsup_{n \rightarrow \infty} \frac{I(S_n; Z^n)[P_{Y|X}^n, \varphi_{a,n}, P_{\text{mix}, \mathcal{S}_n, \mathcal{E}_n}]}{n} &\leq R_l. \end{aligned}$$

The capacity region with the leaked information criterion of the BCC is the closure of the achievable rate triples.

The capacity region with the leaked information criterion of the BCC is characterized as a corollary of Theorem 7.

Corollary 9: The capacity region with the leaked information criterion of the BCC is given by the set of R_0, R_1 and R_l , such that there exists a Markov chain $U \rightarrow V \rightarrow X \rightarrow YZ$ and

$$\begin{aligned} R_1 + R_0 &\leq I(V; Y|U) + \min[I(U; Y), I(U; Z)], \\ R_0 &\leq \min[I(U; Y), I(U; Z)], \\ R_l &\geq R_1 - [I(V; Y|U) - I(V; Z|U)]_+, \end{aligned}$$

where $[x]_+ := \max(x, 0)$. That is, when $R_1 + R_0 < I(V; Y|U) + \min[I(U; Y), I(U; Z)]$ and $R_0 < \min[I(U; Y), I(U; Z)]$, there exists a sequence of the codes $\varphi_n = (\varphi_{a,n}, \varphi_{b,n}, \varphi_{e,n})$, i.e., Alice's stochastic encoder $\varphi_{a,n}$ from $\mathcal{S}_n \times \mathcal{E}_n$ to \mathcal{X}^n , Bob's deterministic decoder $\varphi_{b,n} : \mathcal{Y}^n \rightarrow \mathcal{S}_n \times \mathcal{E}_n$ and Eve's deterministic decoder $\varphi_{e,n} : \mathcal{Z}^n \rightarrow \mathcal{E}_n$ such that

$$\begin{aligned} \lim_{n \rightarrow \infty} P_b[P_{Y|X}^n, \varphi_n, P_{\text{mix}, \mathcal{S}_n, \mathcal{E}_n}] &= 0 \\ \lim_{n \rightarrow \infty} P_e[P_{Z|X}^n, \varphi_n, P_{\text{mix}, \mathcal{S}_n, \mathcal{E}_n}] &= 0 \end{aligned}$$

and

$$\begin{aligned} \limsup_{n \rightarrow \infty} \frac{I(S_n; Z^n)[P_{Y|X}^n, \varphi_{a,n}, P_{\text{mix}, \mathcal{S}_n, \mathcal{E}_n}]}{n} \\ \leq R_1 - [I(V; Y|U) - I(V; Z|U)]_+. \end{aligned}$$

B. Our Approach to BCC

Next, we consider the BCC with the single-shot setting when the common and confidential messages do not obey the uniform and independent distributions on \mathcal{S} and \mathcal{E} , i.e., the confidential message S may have a correlation with the common messages E . When the confidential message S is independent of the common messages E ,

$$\begin{aligned} I(S; Z) &\leq I(S; ZE) = I(S; Z|E) + I(S; E) = I(S; Z|E), \\ I(S; Z) &= H(S) - H(S|Z) \geq H(S|E) - H(S|Z) \\ &= H(S|E) - (H(S|ZE) + I(S; E|Z)) = I(S; Z|E) - I(S; E|Z) \\ &\geq I(S; Z|E) - H(E|Z) \geq I(S; Z|E) - H(E|\varphi_e(Z)). \end{aligned}$$

When the error probability goes to zero, Fano's inequality guarantees that $H(E|Z)$ goes to zero. Hence, $I(S; Z)$ and $I(S; Z|E)$ have the same asymptotic behaviors. So, even if we replace $I(S; Z)$ by $I(S; Z|E)$ in Definition 8, we obtain the same capacity region. However, when the confidential message S is dependent on the common messages E , $I(S; Z)$ and $I(S; Z|E)$ have the different asymptotic behavior as follows. Since

$$\begin{aligned} I(S; Z) &= I(S; ZE) - I(S; E|Z) \\ &\geq I(S; E) - H(E|Z) \geq I(S; E) - H(E|\varphi_e(Z)), \end{aligned}$$

$I(S; Z)$ is asymptotically lower bounded by $I(S; E)$ when the error probability goes to zero. That is, when the mutual information $I(S; E)$ is positive, the mutual information $I(S; Z)$ cannot go to zero because Eve can infer the secret message from the common message. Thus, it is not suitable to treat the mutual information $I(S; Z)$ as leaked information from Z . Hence, we adopt the conditional mutual information $I(S; Z|E)$ as leaked information from Z .

Remark 10: Csiszár and Körner [9] treated BCC with non-uniform information source. However, their formulation was different from our formulation in the following point. In their formulation, they fixed a correlated non-uniform distribution $P_{S,E}$ on $\mathcal{S} \times \mathcal{E}$ and assumed that the information source S_n and E_n obey its n -fold independent and identical distribution $P_{S,E}^n$. In addition to this, their code depends on the distribution $P_{S,E}$. However, in our formulation, we do not assume the independent and identical distributed condition for the distribution P_{S_n, E_n} of the information source S_n and E_n . This is because information source is not given as an independent and identical distribution or known, in general. Hence, we study a universal code independent of the distribution P_{S_n, E_n} of sources in Section XIII. Thus, our code is useful for a realistic case.

V. BROADCAST CHANNELS WITH DEGRADED MESSAGE SETS

A. Capacity Region

Next, we review the broadcast channel with degraded message sets (abbreviated as BCD) considered by Körner and Marton [23] in the single-shot setting. If we set $R_e = 0$ in the BCC, the secrecy requirement is removed from BCC, and the coding problem is equivalent to BCD. In this problem, we treat the private message S_p taking values in \mathcal{S}_p and the common message S_c taking values in \mathcal{S}_c .

Corollary 11: [23] The capacity region of the BCD is given by the pair of the rate R_c of common message and the rate R_p of private message such that there exists a Markov chain $U \rightarrow V = X \rightarrow YZ$ and

$$\begin{aligned} R_c &\leq \min[I(U; Y), I(U; Z)], \\ R_c + R_p &\leq I(V; Y|U) + \min[I(U; Y), I(U; Z)]. \end{aligned}$$

Note that the statement of our Corollary 11 is the same as [9, Corollary 5] and different from [23]. However, as is stated in [9, Remark 5], the equivalence between the two statements can be easily shown by some algebra.

Here, we only consider a sequence of codes that achieves the rate pair (R_c, R_p) satisfying

$$R_c < \min[I(U; Y), I(U; Z)], \quad R_p < I(V; Y|U). \quad (32)$$

For a given Markov chain $U \rightarrow V = X \rightarrow YZ$, we construct an ensemble of codes by the following random coding with the single-shot setting, which is mathematically equivalent to the construction by Kaspi and Merhav [21].

Code Ensemble 1 (Kaspi and Merhav [21, Section II]):

³ For an arbitrary element $s_c \in \mathcal{S}_c$, $\Phi_c(s_c)$ is the random variable taking values in \mathcal{U} and is subject to the distribution P_U , and is independent of $\Phi_c(s'_c)$ with $s'_c \neq s_c \in \mathcal{S}_c$. For an arbitrary element $s_p \in \mathcal{S}_p$, $\Phi_p(s_c, s_p)$ is the random variable taking values in \mathcal{V} , is independent of $\Phi_p(s'_c, s'_p)$ with $s'_c \neq s_c$, and depends on the random variable $\Phi_c(s_c)$. Under the condition $\Phi_c(s_c) = u$, the random variable $\Phi_p(s_c, s_p)$ is subject to the distribution $P_{V|U=u}$ and is conditionally independent of $\Phi_p(s_c, s'_p)$ with $s'_p \neq s_p$. Bob's decoder Φ_b and Eve's decoder Φ_e are defined as the maximum likelihood decoders. The quartet $(\Phi_p, \Phi_c, \Phi_b, \Phi_e)$ is abbreviated as Φ .

Here, the all values of the random variables $\{\Phi_c(s_c)\}_{s_c}$ and $\{\Phi_p(s_c, s_p)\}_{s_c, s_p}$ are disclosed to all players prior to the real communication because these random variables decides our code.

Lemma 12: [21, Theorem 1 and Section IV] The above ensemble of codes Φ satisfies the following inequalities.

$$\begin{aligned} E_{\Phi} P_b[P_{Y|V}, \Phi] &\leq |\mathcal{S}_p|^{\rho} e^{E_0(-\rho|P_{Y|V}, P_{V|U}, P_U)} \\ &\quad + (|\mathcal{S}_c| |\mathcal{S}_p|)^{\rho} e^{E_0(-\rho|P_{Y|U, V}, P_{U, V})} \end{aligned} \quad (33)$$

$$E_{\Phi} P_e[P_{Z|V}, \Phi] \leq |\mathcal{S}_c|^{\rho} e^{E_0(-\rho|P_{Z|U}, P_U)}, \quad (34)$$

where $E_0(-\rho|P_{Z|U}, P_U)$ and $E_0(-\rho|P_{Y|V}, P_{V|U}, P_U)$ are defined in (11) and (12).

Here, we should remark that Inequalities (33) and (34) hold for any distribution over the messages because the proof by [21] does not make any assumption for the distribution over the messages.

Due to Lemma 12, Markov inequality guarantees that

$$\begin{aligned} \Pr \Omega_1 &< \frac{1}{2}, \quad \Pr \Omega_2 < \frac{1}{2} \\ \Omega_1 &:= \left\{ P_b[P_{Y|V}, \Phi, P_{\text{mix}, \mathcal{S}_p, \mathcal{S}_c}] > 2|\mathcal{S}_p|^{\rho} e^{E_0(-\rho|P_{Y|V}, P_{V|U}, P_U)} \right. \\ &\quad \left. + 2(|\mathcal{S}_c| |\mathcal{S}_p|)^{\rho} e^{E_0(-\rho|P_{Y|U, V}, P_{U, V})} \right\} \\ \Omega_2 &:= \{ P_e[P_{Z|V}, \Phi, P_{\text{mix}, \mathcal{S}_p, \mathcal{S}_c}] > 2|\mathcal{S}_c|^{\rho} e^{E_0(-\rho|P_{Z|U}, P_U)} \}. \end{aligned}$$

Since $\Pr(\Omega_1 \cup \Omega_2) < 1$, we have $\Pr(\Omega_1^c \cap \Omega_2^c) > 0$. That is, for an arbitrary distribution P_{S_p, S_c} over the messages, there exists a code φ such that

$$\begin{aligned} P_b[P_{Y|V}, \varphi, P_{S_p, S_c}] &\leq 2|\mathcal{S}_p|^{\rho} e^{E_0(-\rho|P_{Y|V}, P_{V|U}, P_U)} \\ &\quad + 2(|\mathcal{S}_c| |\mathcal{S}_p|)^{\rho} e^{E_0(-\rho|P_{Y|U, V}, P_{U, V})} \end{aligned} \quad (35)$$

$$P_e[P_{Z|V}, \varphi, P_{S_p, S_c}] \leq 2|\mathcal{S}_c|^{\rho} e^{E_0(-\rho|P_{Z|U}, P_U)}. \quad (36)$$

Now, we apply the above inequalities to the n -fold discrete memoryless extension. Then, for an arbitrary distribution $P_{S_p, n, S_c, n}$ over the messages, there exists a sequence of codes φ_n with the rate of common message R_c and the rate of private

³A code ensemble and a code construction play a distinguished role in this paper because they give a procedure to make our codes. Hence, we give them serial numbers that are separate from other environments, Theorems, Lemmas, and Remarks. Although both of a code ensemble and a code construction give a procedure for our code, the procedure by a code ensemble is less practical, and that by a code construction is more practical. To clarify this difference, we assigned one of two environments to them dependently of their properties. Code constructions will be given in Section XI after code ensembles are presented in the previous sections.

message R_p of length n such that

$$P_b[P_{Y|V}^n, \varphi_n, P_{S_{p,n}, S_{c,n}}] \leq 2e^{n(\rho R_p + E_0(-\rho|P_{Y|V}, P_{V|U}, P_U))} + 2e^{n(\rho(R_p + R_c) + E_0(-\rho|P_{Y|U,V}, P_{U,V}))} \quad (37)$$

$$P_e[P_{Z|V}^n, \varphi_n, P_{S_{p,n}, S_{c,n}}] \leq 2e^{n(\rho R_c + E_0(-\rho|P_{Z|U}, P_U))}. \quad (38)$$

The above values go to zero under the condition (32), because the condition (32) guarantees that both exponents are positive with sufficiently small $\rho > 0$.

Indeed, Kaspi and Merhav [21] derived a better bound than (34) by employing four parameters even in the single-shot setting. The bound (34) can be seen as a special case of Kaspi and Merhav [21]'s bound. Since the bound (34) can derive the capacity region of SMC, we only use the bound (34) for simplicity.

B. Universal Code for BCD

Körner and Sgarro [24] provided the code that attains the above rate region universally for source and channel in the following sense.

Theorem 13: [24] For an arbitrary real number $\epsilon > 0$, there exists an integer N satisfying the following. For an arbitrary integer $n \geq N$, a given joint type Q_{VU} of length n on the sets $\mathcal{V} \times \mathcal{U}$, and rates R_p and R_c , there exists a code φ_n with the rates R_p and R_c such that

$$P_b[W^n, \varphi_n, S_{p,n} = s_{p,n}, S_{c,n} = s_{c,n}] \leq \exp(-n[\tilde{E}^b(R_p, R_c, W^Y \times Q_{U,V}) - \epsilon]), \quad (39)$$

$$P_e[W^n, \varphi_n, S_{p,n} = s_{p,n}, S_{c,n} = s_{c,n}] \leq \exp(-n[\tilde{E}^e(R_c, W^Z \times Q_{U,V}) - \epsilon]) \quad (40)$$

for any $s_{p,n} \in S_{p,n}$, $s_{c,n} \in S_{c,n}$ and any $W \in \mathcal{W}(\mathcal{V}, \mathcal{Y} \times \mathcal{Z})$, where the exponents $\tilde{E}^b(R_p, R_c, W^Y \times Q_{U,V})$ and $\tilde{E}^e(R_c, W^Z \times Q_{U,V})$ are defined in (29) and (30), respectively.

VI. GENERAL CHANNEL RESOLVABILITY

In the wire-tap channel model, when the dummy message obeys the uniform distribution, channel resolvability [13] can be used for guaranteeing the security [15]. In this paper, we consider the security of SMC with non-uniform and dependent secret messages. For the analysis of this case, we have to consider the secrecy when the dummy message does not necessarily obey the uniform distribution. Hence, the security evaluation [15] based on the original channel resolvability cannot be extended to the security of SMC with non-uniform and dependent secret messages. Thus, we need a generalization of channel resolvability. In this section, we propose a generalization of channel resolvability in the single-shot setting.

First, we fix a channel W from the alphabet \mathcal{X} to the alphabet \mathcal{Y} . For a fixed distribution P_X on \mathcal{X} , we focus on an encoder Λ from the message set \mathcal{A} to the alphabet \mathcal{X} . The purpose of the encoder Λ is approximation of the average output distribution $W \circ P_X$ by the output distribution with input $\Lambda(A)$. The original channel resolvability [13] treats the minimum asymptotic rate of $|\mathcal{A}|$ such that the output distribution $W \circ \Lambda \circ P_{\text{mix}, \mathcal{A}}$ can approximate the average output

distribution $W \circ P_X$ with a suitable choice of Λ in the sense that the variational distance goes to zero. In the single-shot setting, the problem can be converted to the following way: How well the given average output distribution $W \circ P_X$ can be approximated by the output distribution $W \circ \Lambda \circ P_{\text{mix}, \mathcal{A}}$ when the cardinality $|\mathcal{A}|$ is less than a given amount. In this paper, we consider this approximation problem when the message A does not obey the uniform distribution $P_{\text{mix}, \mathcal{A}}$. Since our problem can be regarded as a generalization of channel resolvability, it is called general channel resolvability, which is essential for the secure multiplex coding with common messages with dependent and non-uniform secret messages.

Now, we apply the random coding on the alphabet A with the probability distribution P_A . For an arbitrary $a \in \mathcal{A}$, $\Lambda(a)$ is the random variable subject to the distribution P_X on \mathcal{X} . For $a \neq a' \in \mathcal{A}$, $\Lambda(a)$ is independent of $\Lambda(a')$. Then, the random encoder $\Lambda := \{\Lambda(a)\}_{a \in \mathcal{A}}$ gives the map from \mathcal{A} to \mathcal{X} as $a \mapsto \Lambda(a)$.

Then, we have the following theorem:

Theorem 14 (General channel resolvability): For $\rho \in (0, 1]$, we have

$$\mathbf{E}_\Lambda e^{\rho D(W \circ \Lambda \circ P_A \| W \circ P_X)} \leq \mathbf{E}_\Lambda e^{\psi(\rho | W \circ \Lambda \circ P_A \| W \circ P_X)} \leq 1 + e^{-\rho H_{1+\rho}(A)} e^{\psi(\rho | W, P_X)}.$$

By applying Jensen inequality to the function $x \mapsto e^x$, Theorem 14 yields

$$\mathbf{E}_\Lambda D(W \circ \Lambda \circ P_A \| W \circ P_X) \leq \frac{1}{\rho} \log \mathbf{E}_\Lambda e^{\rho D(W \circ \Lambda \circ P_A \| W \circ P_X)} \leq \frac{1}{\rho} \log(1 + e^{-\rho H_{1+\rho}(A)} e^{\psi(\rho | W, P_X)}),$$

which is non-uniform generalization of [15, Lemma 2]. This theorem will be used for the proof of Theorem 20.

Proof: Due to (7), we have

$$\rho D(W \circ \Lambda \circ P_A \| W \circ P_X) \leq \psi(\rho | W \circ \Lambda \circ P_A \| W \circ P_X).$$

The average of $e^{\psi(\rho|W \circ \Lambda \circ P_A \| W \circ P_X)}$ is evaluated as

$$\begin{aligned}
& \mathbf{E}_\Lambda e^{\psi(\rho|W \circ \Lambda \circ P_A \| W \circ P_X)} \\
&= \mathbf{E}_\Lambda \sum_y \left(\sum_a P_A(a) W_{\Lambda(a)}(y) \right)^{1+\rho} (W \circ P_X)(y)^{-\rho} \\
&= \mathbf{E}_\Lambda \sum_y \left(\sum_a P_A(a) W_{\Lambda(a)}(y) \right) \left(\sum_{a'} P_A(a') W_{\Lambda(a')}(y) \right)^\rho (W \circ P_X)(y)^{-\rho} \\
&= \sum_y \sum_a \left(\mathbf{E}_{\Lambda(a)} P_A(a) W_{\Lambda(a)}(y) \mathbf{E}_{\Lambda \setminus \Lambda(a)} \left(P_A(a) W_{\Lambda(a)}(y) \right. \right. \\
&\quad \left. \left. + \sum_{a' \neq a} P_A(a') W_{\Lambda(a')}(y) \right)^\rho (W \circ P_X)(y)^{-\rho} \right) \\
&\leq \sum_y \sum_a \left(\mathbf{E}_{\Lambda(a)} P_A(a) W_{\Lambda(a)}(y) \left(P_A(a) W_{\Lambda(a)}(y) \right. \right. \\
&\quad \left. \left. + \mathbf{E}_{\Lambda \setminus \Lambda(a)} \sum_{a' \neq a} P_A(a') W_{\Lambda(a')}(y) \right)^\rho (W \circ P_X)(y)^{-\rho} \right) \quad (41)
\end{aligned}$$

$$\begin{aligned}
&= \sum_y \sum_a \left(\mathbf{E}_{\Lambda(a)} P_A(a) W_{\Lambda(a)}(y) \left(P_A(a) W_{\Lambda(a)}(y) \right. \right. \\
&\quad \left. \left. + \sum_{a' \neq a} P_A(a') (W \circ P_X)(y) \right)^\rho (W \circ P_X)(y)^{-\rho} \right) \\
&\leq \sum_y \sum_a \left(\mathbf{E}_{\Lambda(a)} P_A(a) W_{\Lambda(a)}(y) \left(P_A(a) W_{\Lambda(a)}(y) + (W \circ P_X)(y) \right)^\rho \right. \\
&\quad \left. \cdot (W \circ P_X)(y)^{-\rho} \right) \quad (42)
\end{aligned}$$

$$\begin{aligned}
&\leq \sum_y \sum_a \mathbf{E}_{\Lambda(a)} P_A(a) W_{\Lambda(a)}(y) \\
&\quad \left(P_A(a)^\rho W_{\Lambda(a)}(y)^\rho + (W \circ P_X)(y)^\rho \right) (W \circ P_X)(y)^{-\rho} \quad (43) \\
&= \sum_y \sum_a \mathbf{E}_{\Lambda(a)} P_A(a) W_{\Lambda(a)}(y) \left(1 + P_A(a)^\rho W_{\Lambda(a)}(y)^\rho (W \circ P_X)(y)^{-\rho} \right) \\
&= 1 + \sum_y \sum_a \mathbf{E}_{\Lambda(a)} P_A(a)^{1+\rho} W_{\Lambda(a)}(y)^{1+\rho} (W \circ P_X)(y)^{-\rho} \\
&= 1 + \sum_a P_A(a)^{1+\rho} \sum_y \sum_x P_X(x) W_x(y)^{1+\rho} (W \circ P_X)(y)^{-\rho} \\
&= 1 + \left(\sum_a P_A(a)^{1+\rho} \right) e^{\psi(\rho|W, P_X)}.
\end{aligned}$$

In the above derivation, (41) follows from the concavity of $x \mapsto x^\rho$, (42) follows from $\sum_{a' \neq a} P_A(a') \leq 1$, (43) follows from the inequality $(x+y)^\rho \leq x^\rho + y^\rho$. ■

Next, in order to reduce the complexity of encoding, we consider the case when \mathcal{X} and \mathcal{A} are Abelian groups. We introduce the following condition for the ensemble for injective homomorphisms F from \mathcal{A} to \mathcal{X} .

Condition 15: Let F be a random variable that takes its values on injective⁴ homomorphisms from \mathcal{A} to \mathcal{X} . For arbitrary elements $x \neq 0 \in \mathcal{X}$ and $a \neq 0 \in \mathcal{A}$, the relation $F(a) = x$ holds with probability at most $\frac{1}{|\mathcal{X}|-1}$.

When \mathcal{X} and \mathcal{A} are vector spaces over a finite field \mathbb{F}_q , the set of all injective homomorphisms from \mathcal{A} to \mathcal{X} satisfies Condition 15.

Remark 16: When \mathcal{X} and \mathcal{A} have the same Abelian group structure as the vector space over a finite field \mathbb{F}_2 with the same dimension k , these can be regarded as the finite filed \mathbb{F}_{2^k} . For $y \in \mathbb{F}_{2^k}$, the homomorphism f_y from \mathcal{A} to \mathcal{X} from \mathcal{A}

to \mathcal{X} is defined by the multiplication as $f_y : x \rightarrow xy$. Then, as mentioned in [44, Remark 9], when the random variable Y chosen in \mathbb{F}_{2^k} subject to the uniform distribution, the function-valued random variable f_Y satisfies Condition 15. To realize the function-valued random variable f_Y , we need to choose a finite filed \mathbb{F}_{2^k} with efficient multiplication. Constructions of such a finite filed \mathbb{F}_{2^k} are given in [45, Appendix D], [46, Section 7.3.1].

We choose another random variable G in \mathcal{X} that obeys the uniform distribution on \mathcal{X} and is independent of the choice of F . Then, we define a map $\Lambda_{F,G}(a) := F(a) + G$ and have the following theorem:

Theorem 17 (Algebraic channel resolvability): Under the above choice, we obtain

$$\begin{aligned}
& \mathbf{E}_{F,G} e^{\rho D(W \circ \Lambda_{F,G} \circ P_A \| W \circ P_{\text{mix},\mathcal{X}})} \leq \mathbf{E}_{F,G} e^{\psi(\rho|W \circ \Lambda_{F,G} \circ P_A \| W \circ P_{\text{mix},\mathcal{X}})} \\
& \leq 1 + e^{-\rho H_{1+\rho}(A)} e^{\psi(\rho|W, P_{\text{mix},\mathcal{X}})}. \quad (44)
\end{aligned}$$

This theorem will be used for the proof of Lemma 21, which is essential for the proof of Theorem 22.

Proof: We introduce the random variable $Z_a := \Lambda_{F,G}(a) = F(a) + G$. The random variable Z_a is independent of the choice of F . For $a' \in \mathcal{A}$, $\Lambda_{F,G}(a') = F(a' - a) + Z_a$. Since $(|\mathcal{X}| - 1) \mathbf{E}_{F|Z_a} W_{\Lambda_{F,G}(a)}(y) = (|\mathcal{X}| - 1) \mathbf{E}_F W_{F(a' - a) + Z_a}(y) \leq \sum_x W_x(y) = |\mathcal{X}| W \circ P_{\text{mix},\mathcal{X}}(y)$ for $a \in \mathcal{A}$ and $y \in \mathcal{Y}$, we obtain $\mathbf{E}_{F|Z_a} W_{\Lambda_{F,G}(a)}(y) \leq \frac{|\mathcal{X}|}{|\mathcal{X}|-1} W \circ P_{\text{mix},\mathcal{X}}(y)$ for $a \in \mathcal{A}$ and $y \in \mathcal{Y}$. Further, since F is injective, we have $|\mathcal{A}| \leq |\mathcal{X}|$, which implies $\sum_a P_A(a)^2 \geq \frac{1}{|\mathcal{A}|} \geq \frac{1}{|\mathcal{X}|}$. Hence, since $x \mapsto x^\rho$ is concave, we obtain

$$\sum_a P_A(a) \left(\frac{1 - P_A(a)}{1 - 1/|\mathcal{X}|} \right)^\rho \leq \left(\frac{1 - \sum_a P_A(a)^2}{1 - 1/|\mathcal{X}|} \right)^\rho \leq \left(\frac{1 - 1/|\mathcal{X}|}{1 - 1/|\mathcal{X}|} \right)^\rho = 1. \quad (45)$$

Our proof of Theorem 14 can be applied to our proof of Theorem 17 by replacing $\Lambda(a)$, $\Lambda \setminus \Lambda(a)$, and P_X by Z_a , $F|Z_a$

⁴The condition of injectivity is not necessarily for Theorem 17. However, the injectivity for F will be needed in the discussion in Subsection XI-C. Hence, to avoid to make so many conditions, we assume the injectivity, here.

and $P_{\text{mix},\mathcal{X}}$. Then, we obtain

$$\begin{aligned} & \mathbf{E}_{F,G} e^{\psi(\rho|W \circ \Lambda_{F,G} \circ P_A \| W \circ P_{\text{mix},\mathcal{X}})} \\ & \leq \sum_y \sum_a \left(\mathbf{E}_{Z_a} P_A(a) W_{\Lambda_{F,G}(a)}(y) \left(P_A(a) W_{\Lambda_{F,G}(a)}(y) \right. \right. \\ & \quad \left. \left. + \mathbf{E}_{F|Z_a} \sum_{a' \neq a} P_A(a') W_{\Lambda_{F,G}(a')}(y) \right)^{\rho} W \circ P_{\text{mix},\mathcal{X}}(y)^{-\rho} \right) \end{aligned} \quad (46)$$

$$\begin{aligned} & \leq \sum_y \sum_a \left(\mathbf{E}_{Z_a} P_A(a) W_{Z_a}(y) \left(P_A(a) W_{Z_a}(y) \right. \right. \\ & \quad \left. \left. + \frac{|\mathcal{X}|}{|\mathcal{X}|-1} \sum_{a' \neq a} P_A(a') W \circ P_{\text{mix},\mathcal{X}}(y) \right)^{\rho} W \circ P_{\text{mix},\mathcal{X}}(y)^{-\rho} \right) \end{aligned} \quad (47)$$

$$\begin{aligned} & \leq \sum_y \sum_a \left(\mathbf{E}_{Z_a} P_A(a) W_{Z_a}(y) \left(P_A(a) W_{Z_a}(y) \right. \right. \\ & \quad \left. \left. + \frac{1 - P_A(a)}{1 - 1/|\mathcal{X}|} W \circ P_{\text{mix},\mathcal{X}}(y) \right)^{\rho} W \circ P_{\text{mix},\mathcal{X}}(y)^{-\rho} \right) \end{aligned} \quad (48)$$

$$\begin{aligned} & \leq \sum_y \sum_a \left(\mathbf{E}_{Z_a} P_A(a) W_{Z_a}(y) \left(P_A(a)^{\rho} W_{Z_a}(y)^{\rho} \right. \right. \\ & \quad \left. \left. + \left(\frac{1 - P_A(a)}{1 - 1/|\mathcal{X}|} \right)^{\rho} W \circ P_{\text{mix},\mathcal{X}}(y)^{\rho} \right) W \circ P_{\text{mix},\mathcal{X}}(y)^{-\rho} \right) \end{aligned} \quad (49)$$

$$\begin{aligned} & = \sum_y \sum_a \left(\mathbf{E}_{Z_a} P_A(a) W_{Z_a}(y) \left(\left(\frac{1 - P_A(a)}{1 - 1/|\mathcal{X}|} \right)^{\rho} \right. \right. \\ & \quad \left. \left. + P_A(a)^{\rho} W_{Z_a}(y)^{\rho} W \circ P_{\text{mix},\mathcal{X}}(y)^{-\rho} \right) \right) \\ & = \sum_a P_A(a) \left(\frac{1 - P_A(a)}{1 - 1/|\mathcal{X}|} \right)^{\rho} \\ & \quad + \sum_y \sum_a \mathbf{E}_{Z_a} P_A(a)^{1+\rho} W_{Z_a}(y)^{1+\rho} W \circ P_{\text{mix},\mathcal{X}}(y)^{-\rho} \\ & = \sum_a P_A(a) \left(\frac{1 - P_A(a)}{1 - 1/|\mathcal{X}|} \right)^{\rho} \\ & \quad + \sum_a P_A(a)^{1+\rho} \sum_y \sum_x P_X(x) W_x(y)^{1+\rho} W \circ P_{\text{mix},\mathcal{X}}(y)^{-\rho} \\ & \leq 1 + \left(\sum_a P_A(a)^{1+\rho} \right) e^{\psi(\rho|W, P_{\text{mix},\mathcal{X}})}. \end{aligned} \quad (50)$$

In the above derivation, (46) follows in the same way as (41), (47) follows from Condition 15, (48) follows from $\sum_{a' \neq a} P_A(a') \leq 1$, (49) follows from the inequality $(x+y)^{\rho} \leq x^{\rho} + y^{\rho}$. The final inequality follows from (45). ■

In the following, we assume that the input alphabet \mathcal{X} is an Abelian group, and an action of \mathcal{X} on the output alphabet \mathcal{Y} is given as $x \cdot y$ for $x \in \mathcal{X}$ and $y \in \mathcal{Y}$. A channel W from \mathcal{X} to \mathcal{Y} is regular in the sense of Delsarte-Piret [10], if there is a probability distribution P_Y such that

$$W_x(y) = P_Y(x \cdot y).$$

Since a regular channel W satisfies

$$D(W \circ \Lambda_{F,g} \circ P_A \| W \circ P_{\text{mix},\mathcal{X}}) = D(W \circ \Lambda_{F,g'} \circ P_A \| W \circ P_{\text{mix},\mathcal{X}})$$

for any $g, g' \in \mathcal{X}$, we obtain the following corollary. This corollary implies that we do not need the additional random variable G in the regular channel case.

Corollary 18: When the channel W is a regular channel given by a distribution P_Y on \mathcal{Y} , we obtain

$$\begin{aligned} & \mathbf{E}_F e^{\rho D(W \circ \Lambda_{F,g} \circ P_A \| W \circ P_{\text{mix},\mathcal{X}})} \leq \mathbf{E}_F e^{\psi(\rho|W \circ \Lambda_{F,g} \circ P_A \| W \circ P_{\text{mix},\mathcal{X}})} \\ & \leq 1 + e^{-\rho H_{1+\rho}(A)} e^{\psi(\rho|W, P_{\text{mix},\mathcal{X}})} = 1 + e^{-\rho H_{1+\rho}(A)} e^{\psi(\rho|P_Y \| \bar{P}_Y)} \end{aligned} \quad (51)$$

for any $g \in \mathcal{X}$, where $\bar{P}_Y(y) := \sum_x P_{\text{mix},\mathcal{X}}(x) P_Y(x \cdot y)$.

Proof: Due to Theorem 14, it is enough to show $\psi(\rho|W, P_{\text{mix},\mathcal{X}}) = \psi(\rho|P_Y \| \bar{P}_Y)$. Since $\bar{P}_Y(y) = W \circ P_{\text{mix},\mathcal{X}}(y) = W \circ P_{\text{mix},\mathcal{X}}(x \cdot y)$, we have

$$\begin{aligned} e^{\psi(\rho|W, P_{\text{mix},\mathcal{X}})} & = \sum_x P_{\text{mix},\mathcal{X}}(x) \sum_y P_Y(x \cdot y)^{1+\rho} \bar{P}_Y(y)^{-\rho} \\ & = \sum_x P_{\text{mix},\mathcal{X}}(x) \sum_y P_Y(y)^{1+\rho} \bar{P}_Y(x^{-1} \cdot y)^{-\rho} \\ & = \sum_x P_{\text{mix},\mathcal{X}}(x) \sum_y P_Y(y)^{1+\rho} \bar{P}_Y(y)^{-\rho} \\ & = \sum_y P_Y(y)^{1+\rho} \bar{P}_Y(y)^{-\rho} = e^{\psi(\rho|P_Y \| \bar{P}_Y)}. \end{aligned}$$

VII. SECURE MULTIPLEX CODING WITH COMMON MESSAGES: SINGLE-SHOT SETTING

In this section, we give the formulation of the secure multiplex coding with common messages. After the formulation, we give two kinds of random construction of codes for the secure multiplex coding with common messages and evaluate their performance in the single-shot setting.

A. Formulation and Preparation

In the secure multiplex coding with common messages, Alice sends the common message S_0 to Bob and Eve, and T secret messages S_1, \dots, S_T to Bob. We do not necessarily assume the uniformity nor independence for the distributions of messages S_0, S_1, \dots, S_T . Hence, there might exist statistical correlations among messages S_0, S_1, \dots, S_T . Even in this scenario, Alice and Bob can use $S_1, \dots, S_{i-1}, S_{i+1}, \dots, S_T$ as random bits making S_i ambiguous to Eve. When we focus on $S_{\mathcal{I}} := (S_i; i \in \mathcal{I})$ for a non-empty proper subset $\mathcal{I} (\neq \emptyset) \subseteq \{1, \dots, T\}$, the remaining information $S_{\mathcal{I}^c}$ serves as random bits making $S_{\mathcal{I}}$ ambiguous to Eve. The messages S_0, S_1, \dots, S_T are assumed to belong to the sets $\mathcal{S}_0, \mathcal{S}_1, \dots, \mathcal{S}_T$. The set $\mathcal{S}_1 \times \dots \times \mathcal{S}_T$ of all secret messages is denoted by \mathcal{S} . In order to explain the SMC model without S_0 , we consider the following example. Consider the case when S_1, \dots, S_T are personal information for T persons. That is, S_i corresponds to the personal information of the i -th person. Assume that it is required only to keep the secrecy of the respective personal information S_1, \dots, S_T from the third party. The secrecy of the relation among respective personal informations is not required. For example, when S_1, \dots, S_T are the uniform random bits with the same size, the secrecy of the sum $S_1 \oplus \dots \oplus S_T$ is not required, where \oplus is exclusive OR. In order to treat this secrecy problem, we give a formulation of the SMC model as follows.

The purpose of the coding in the SMC model is to reliably send the messages S_0, S_1, \dots, S_T to Bob, and to make $S_{\mathcal{I}}$ ambiguous to Eve by using the remaining information $S_{\mathcal{I}^c}$ for several non-empty proper subsets $\mathcal{I} \subseteq \{1, \dots, T\}$. Our code is given by Alice's stochastic encoder φ_a from $\mathcal{S} \times \mathcal{S}_0$ to \mathcal{X} , Bob's deterministic decoder $\varphi_b : \mathcal{Y} \rightarrow \mathcal{S} \times \mathcal{S}_0$ and Eve's deterministic decoder $\varphi_e : \mathcal{Z} \rightarrow \mathcal{S}_0$. The triple $\varphi = (\varphi_a, \varphi_b, \varphi_e)$

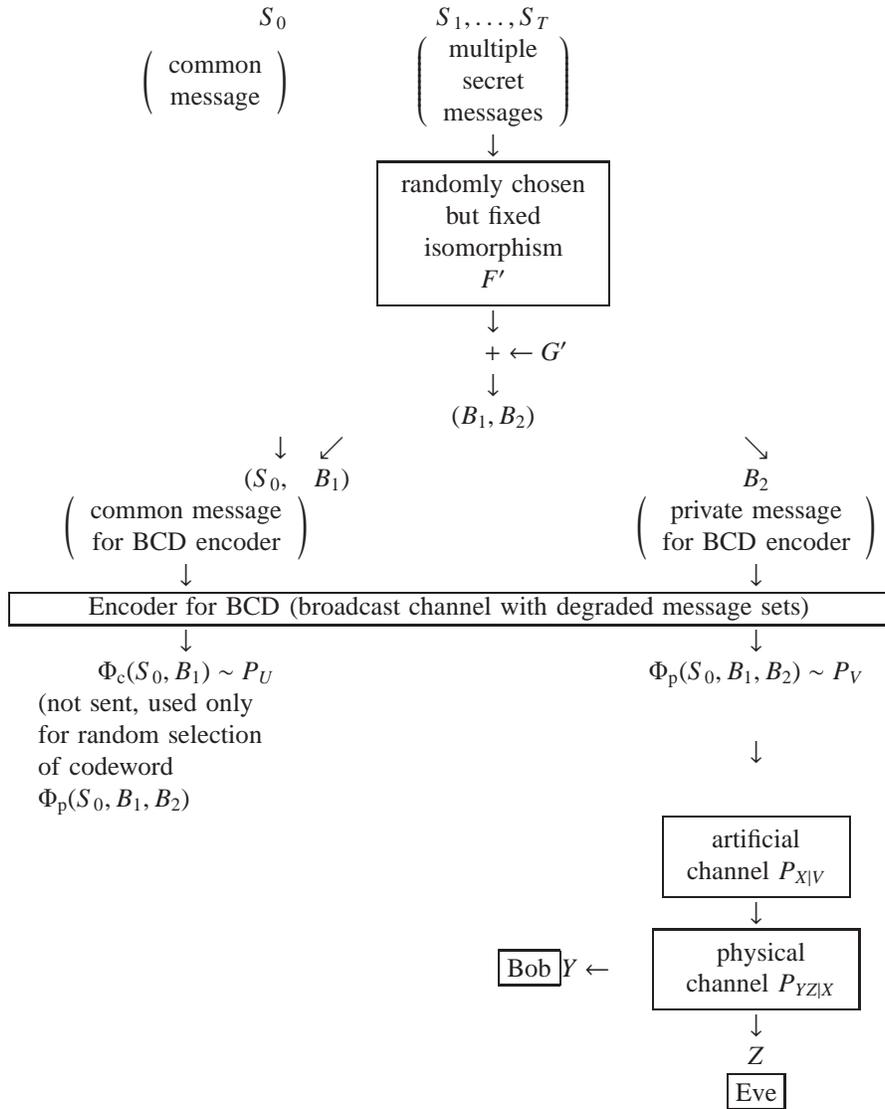


Fig. 1. Communication structure used in Sections VII–XII

is called a code for the secure multiplex coding with common messages. Then, the performance is evaluated by the following quantities: (1) The sizes of the sets of the common messages and all of the secret messages, i.e., $|S_0|, |S_1|, \dots, |S_T|$. (2) Bob’s decoding error probability $P_b[P_{Y|X}, \varphi, P_{S_T}]$, which is the probability $\Pr\{(S_0, S_1, \dots, S_T) \neq \varphi_b(Y)\}$ under the distribution $(P_{Y|X} \circ \varphi_a) \times P_{S_T}$ with $\mathcal{T} := \{0, \dots, T\}$. (3) Eve’s decoding error probability $P_e[P_{Z|X}, \varphi, P_{S_T}]$, which is the probability $\Pr\{S_0 \neq \varphi_e(Z)\}$ under the distribution $(P_{Z|X} \circ \varphi_a) \times P_{S_T}$. (4) Leaked information $I(S_I; Z|S_0)[P_{Z|X}, \varphi_a, P_{S_T}]$ for non-empty proper subset $I \subseteq \{1, \dots, T\}$, which is the mutual information $I(S_I; Z|S_0)$ under the distribution $(P_{Z|X} \circ \varphi_a) \times P_{S_T}$. Instead of $I(S_I; Z|S_0)[P_{Z|X}, \varphi_a, P_{S_T}]$, other researchers sometimes treat (5) Eve’s uncertainty $H(S_I|Z, S_0)[P_{Z|X}, \varphi_a, P_{S_T}]$, which is the conditional entropy $H(S_I|Z, S_0)$ under the distribution $(P_{Z|X} \circ \varphi_a) \times P_{S_T}$. However, when we treat the universality of our code, leaked information $I(S_I; Z|S_0)[P_{Z|X}, \varphi_a, P_{S_T}]$ is used as criterion for performance of our code. That is, we adopt leaked information $I(S_I; Z|S_0)[P_{Z|X}, \varphi_a, P_{S_T}]$ rather than

Eve’s uncertainty $H(S_I|Z, S_0)[P_{Z|X}, \varphi_a, P_{S_T}]$.

In the above formulation, we treat the leaked information $I(S_I; Z|S_0)[P_{Z|X}, \varphi_a, P_{S_T}]$ for several non-empty proper subsets $I \subseteq \{1, \dots, T\}$. Depending on the situation, we decide which non-empty proper subset I is considered. Hence, in that case, we can fix a family \mathbf{J} of non-empty proper subsets I of $\{1, \dots, T\}$ for which we discuss the leaked information $I(S_I; Z|S_0)[P_{Z|X}, \varphi_a, P_{S_T}]$. For example, in the case of the above personal information, we consider the subsets $\{1\}, \{2\}, \dots, \{T\}$. Hence, we choose \mathbf{J} as $\mathbf{J} := \{\{1\}, \{2\}, \dots, \{T\}\}$. When we do not specify the family \mathbf{J} , we treat the leaked information $I(S_I; Z|S_0)[P_{Z|X}, \varphi_a, P_{S_T}]$ for all non-empty proper subsets I of $\{1, \dots, T\}$.

This model can be regarded as a generalization of the wire-tap model in the following way. When there is no common messages and $T = 2$, there exist only two messages S_1 and S_2 in the secure multiplex coding. In the wire-tap channel model, S_1 corresponds to the message to be secretly sent to Bob, and S_2 does to the dummy message making S_1

ambiguous to Eve. As a special case of our code, a wire-tap code is given by Alice's stochastic encoder φ_a from $\mathcal{S}_1 \times \mathcal{S}_2$ to \mathcal{X} and Bob's deterministic decoder $\varphi_b : \mathcal{Y} \rightarrow \mathcal{S}_1$. Then, the performance is evaluated by the following quantities. (1) The size of the secret message $|\mathcal{S}_1|$. (2) Bob's decoding error probability $P_b[P_{Y|X}, \varphi, P_{S_{1,2}}]$. (4) Leaked information $I(S_1; Z)[P_{Z|X}, \varphi_a, P_{S_{1,2}}]$.

In order to guarantee that the leaked information is small, we employ the method of generalized channel resolvability given in Section VI. In order to employ this method, we have to use the random coding method to construct a code φ . In this section, we propose two kinds of random construction for our code. For a simple application of Theorem 14, which is a simple generalization of channel resolvability, we propose the first construction in Subsection VII-B. When there is no common message, this construction achieves the capacity region, as is mentioned in Remark 39. However, it cannot fully achieve the capacity region that will be defined in Section IX-B when there exists a common message S_0 .

To resolve this defect, in Subsection VII-C, we propose the second construction, which attains the capacity region. This construction has two steps. In the first step, similar to the BCD encoder, we use the superposition random coding. In the second step, as illustrated in Fig. 1, we split the confidential message into the private message B_2 and a part B_1 of the common message encoded by the BCD encoder. The coding scheme for BCC in [9] uses this kind of message splitting. The average leaked information under this kind of construction is evaluated by Theorem 17, which is an algebraic version of channel resolvability. However, when there is no common message, the first construction realizes a better exponential decreasing rate for leaked information than the second construction.

When we fix a code φ , we obtain the following observations. Any distribution \tilde{P}_Z on \mathcal{Z} and any non-empty proper subset $\mathcal{I} \subseteq \{1, \dots, T\}$ satisfy

$$\begin{aligned} & \rho I(S_{\mathcal{I}}; Z|S_0)[P_{Z|V}, \varphi, P_{S_{\mathcal{T}}}] \\ &= \rho \sum_{s_0} P_{S_0}(s_0) I(S_{\mathcal{I}}; Z|S_0 = s_0)[P_{Z|V}, \varphi, P_{S_{\mathcal{T}}}] \\ &= \rho \sum_{s_0} P_{S_0}(s_0) D(P_{Z, S_{\mathcal{I}}|S_0=s_0, \varphi} \| P_{Z|S_0=s_0, \varphi} \times P_{S_{\mathcal{I}}|S_0=s_0, \varphi}) \\ &\leq \rho \sum_{s_0} P_{S_0}(s_0) D(P_{Z, S_{\mathcal{I}}|S_0=s_0, \varphi} \| \tilde{P}_Z \times P_{S_{\mathcal{I}}|S_0=s_0, \varphi}) \quad (52) \\ &= \sum_{s_0} P_{S_0}(s_0) \sum_{S_{\mathcal{I}}} P_{S_{\mathcal{I}}|S_0}(S_{\mathcal{I}}|s_0) \rho D(P_{Z|S_{\mathcal{I}}=S_{\mathcal{I}}, S_0=s_0, \varphi} \| \tilde{P}_Z), \quad (53) \end{aligned}$$

where (52) follows from the following general inequality

$$D(P_{X,Y} \| P_X \times P_Y) \leq D(P_{X,Y} \| Q_X \times P_Y) \quad (54)$$

for any distribution Q_X over \mathcal{X} . Due to (7), we have

$$\rho D(P_{Z|S_{\mathcal{I}}=S_{\mathcal{I}}, S_0=s_0, \varphi} \| \tilde{P}_Z) \leq \psi(\rho | P_{Z|S_{\mathcal{I}}=S_{\mathcal{I}}, S_0=s_0, \varphi} \| \tilde{P}_Z). \quad (55)$$

Thus, combining Jensen inequality and the above observations, we obtain the following lemma.

Lemma 19: Any distribution \tilde{P}_Z on \mathcal{Z} and any non-empty proper subset $\mathcal{I} \subseteq \{1, \dots, T\}$ satisfy

$$\begin{aligned} e^{\rho I(S_{\mathcal{I}}; Z|S_0)[P_{Z|V}, \varphi, P_{S_{\mathcal{T}}}]} &\leq e^{\sum_{s_0} P_{S_0}(s_0) \sum_{S_{\mathcal{I}}} P_{S_{\mathcal{I}}|S_0}(S_{\mathcal{I}}|s_0) \rho D(P_{Z|S_{\mathcal{I}}=S_{\mathcal{I}}, S_0=s_0, \varphi} \| \tilde{P}_Z)} \\ &\leq \sum_{s_0} P_{S_0}(s_0) \sum_{S_{\mathcal{I}}} P_{S_{\mathcal{I}}|S_0}(S_{\mathcal{I}}|s_0) e^{\rho D(P_{Z|S_{\mathcal{I}}=S_{\mathcal{I}}, S_0=s_0, \varphi} \| \tilde{P}_Z)} \quad (56) \end{aligned}$$

$$\leq \sum_{s_0} P_{S_0}(s_0) \sum_{S_{\mathcal{I}}} P_{S_{\mathcal{I}}|S_0}(S_{\mathcal{I}}|s_0) e^{\psi(\rho | P_{Z|S_{\mathcal{I}}=S_{\mathcal{I}}, S_0=s_0, \varphi} \| \tilde{P}_Z)}. \quad (57)$$

B. First Construction

Now, we introduce the first kind of random coding for SMC.

Code Ensemble 2: For a given Markov chain $U \rightarrow V \rightarrow X \rightarrow YZ$, we give the random coding Φ_c and Φ_p in the same way as Code Ensemble 1 with $\mathcal{S}_c = \mathcal{S}_0$ and $\mathcal{S}_p = \mathcal{S}_1 \times \dots \times \mathcal{S}_T$. Similar to the case of BCD, Bob's decoder Φ_b and Eve's decoder Φ_e are defined as the maximum likelihood decoders. Hence, our code is written by the quartet $(\Phi_c, \Phi_p, \Phi_b, \Phi_e)$.

As a special case of Code Ensemble 2, a wire-tap code is given as the case when $T = 2$ and we do not have the random variables S_0 . The averaged performance of the above code is evaluated by the following theorem. Indeed, we cannot derive the capacity region from the following theorem. However, the following theorem has an advantage when the conditional mutual information goes to zero. As is explained in Section X, the following theorem yields a better bound for the exponential decreasing rate of the conditional mutual information than Theorem 22 in a specific case.

Theorem 20: The above ensemble of codes $\Phi = (\Phi_c, \Phi_p, \Phi_b, \Phi_e)$ satisfies the following inequalities.

$$\begin{aligned} & \mathbf{E}_{\Phi} \exp(\rho I(S_{\mathcal{I}}; Z|S_0)[P_{Z|V}, \Phi, P_{S_{\mathcal{T}}}]) \\ &\leq 1 + e^{-\rho H_{1+\rho}(S_{\mathcal{I}^c}|S_{\mathcal{I}}, S_0) + \psi(\rho | P_{Z|V}, P_{V|U}, P_U)}, \quad (58) \end{aligned}$$

$$\begin{aligned} & \mathbf{E}_{\Phi} P_b[P_{Y|V}, \Phi, P_{S_{\mathcal{T}}}] \\ &\leq |\mathcal{S}|^{\rho} e^{E_0(-\rho | P_{Y|V}, P_{V|U}, P_U)} + (|\mathcal{S}_0| |\mathcal{S}|)^{\rho} e^{E_0(-\rho | P_{Y|U,V}, P_U)}, \quad (59) \end{aligned}$$

$$\mathbf{E}_{\Phi} P_e[P_{Z|V}, \Phi, P_{S_{\mathcal{T}}}] \leq |\mathcal{S}_0|^{\rho} e^{E_0(-\rho | P_{Z|U}, P_U)}. \quad (60)$$

Theorem 20 yields the following observation. Applying Jensen's inequality to the convex function $x \mapsto e^x$, we obtain

$$\begin{aligned} & \mathbf{E}_{\Phi} \rho I(S_{\mathcal{I}}; Z|S_0)[P_{Z|V}, \Phi, P_{S_{\mathcal{T}}}] \\ &\leq \log(1 + e^{-\rho H_{1+\rho}(S_{\mathcal{I}^c}|S_{\mathcal{I}}, S_0) + \psi(\rho | P_{Z|V}, P_{V|U}, P_U)}) \\ &\leq e^{-\rho H_{1+\rho}(S_{\mathcal{I}^c}|S_{\mathcal{I}}, S_0) + \psi(\rho | P_{Z|V}, P_{V|U}, P_U)}. \quad (61) \end{aligned}$$

The number of non-empty proper subsets $\mathcal{I} \subseteq \{1, \dots, T\}$ is $2^T - 2$. Similar to (35) and (36), since $2(2^T - 2) + 2 = 2^{T+1} - 2 < 2^{T+1}$, Markov inequality guarantees that there exists a code φ

such that

$$\begin{aligned} & \exp(\rho I(S_I; Z|S_0)[P_{Z|V}, \varphi, P_{S_T}]) \\ & \leq 2^{T+1} (1 + e^{-\rho H_{1+\rho}(S_{I^c}|S_I, S_0) + \psi(\rho|P_{Z|V}, P_{V|U}, P_U)}) \\ & \leq 2^{T+2} e^{[-\rho H_{1+\rho}(S_{I^c}|S_I, S_0) + \psi(\rho|P_{Z|V}, P_{V|U}, P_U)]_+}, \end{aligned} \quad (62)$$

$$\begin{aligned} & \rho I(S_I; Z|S_0)[P_{Z|V}, \varphi, P_{S_T}] \\ & \leq 2^{T+1} e^{-\rho H_{1+\rho}(S_{I^c}|S_I, S_0) + \psi(\rho|P_{Z|V}, P_{V|U}, P_U)}, \end{aligned} \quad (63)$$

$$\begin{aligned} & P_b[P_{Y|V}, \varphi, P_{S_T}] \\ & \leq 2^{T+1} |\mathcal{S}|^\rho e^{E_0(-\rho|P_{Y|V}, P_{V|U}, P_U)} + 2^{T+1} |\mathcal{S}_0|^\rho e^{E_0(-\rho|P_{Y|U}, P_U)}, \end{aligned} \quad (64)$$

$$\begin{aligned} & P_e[P_{Z|V}, \varphi, P_{S_T}] \\ & \leq 2^{T+1} |\mathcal{S}_0|^\rho e^{E_0(-\rho|P_{Z|U}, P_U)}. \end{aligned} \quad (65)$$

Taking the logarithm in (62), we obtain

$$\begin{aligned} & I(S_I; Z|S_0)[P_{Z|V}, \Phi, P_{S_T}] \\ & \leq (T+2) \frac{\log 2}{\rho} + \left[\frac{1}{\rho} \psi(\rho|P_{Z|V}, P_{V|U}, P_U) - H_{1+\rho}(S_{I^c}|S_I, S_0) \right]_+. \end{aligned} \quad (66)$$

Proof of Theorem 20:

Inequalities (59) and (60) can be shown by Lemma 12. The remaining inequality (58) can be shown as follows.

$$\begin{aligned} & \mathbf{E}_\Phi e^{\rho I(S_I; Z|S_0, \Phi)} \\ & \stackrel{(a)}{\leq} \mathbf{E}_\Phi \sum_{s_0} P_{S_0}(s_0) \sum_{S_I} P_{S_I|S_0}(S_I|s_0) e^{\psi(\rho|P_{Z|S_I=S_I, S_0=s_0, \Phi} \| P_{Z|U=\Phi_c(s_0)})} \\ & = \sum_{s_0} P_{S_0}(s_0) \sum_{S_I} P_{S_I|S_0}(S_I|s_0) \\ & \quad \cdot \mathbf{E}_{\Phi_c} \mathbf{E}_{\Phi_p|\Phi_c} e^{\psi(\rho|P_{Z|S_I=S_I, S_0=s_0, \Phi} \| P_{Z|U=\Phi_c(s_0)})} \\ & \stackrel{(b)}{\leq} \sum_{s_0} P_{S_0}(s_0) \sum_{S_I} P_{S_I|S_0}(S_I|s_0) \\ & \quad \cdot \mathbf{E}_{\Phi_c} (1 + e^{-\rho H_{1+\rho}(S_{I^c}|S_I, S_0=s_0)}) e^{\psi(\rho|P_{Z|V}, P_{V|U}=\Phi_c(s_0))} \\ & = \sum_{s_0} P_{S_0}(s_0) \sum_{S_I} P_{S_I|S_0}(S_I|s_0) \\ & \quad \cdot (1 + e^{-\rho H_{1+\rho}(S_{I^c}|S_I, S_0=s_0)}) e^{\psi(\rho|P_{Z|V}, P_{V|U}, P_U)} \\ & = 1 + e^{-\rho H_{1+\rho}(S_{I^c}|S_I, S_0)} e^{\psi(\rho|P_{Z|V}, P_{V|U}, P_U)}, \end{aligned}$$

(a) follows from application of (57) to the case with $\tilde{P}_Z = P_{Z|U=\Phi_c(s_0)}$, and (b) follows from Theorem 14. \blacksquare

C. Second Construction

Next, we give the second kind of random coding for SMC as follows.

Code Ensemble 3: First Step: For a given Markov chain $U \rightarrow V \rightarrow X \rightarrow YZ$, we introduce two random variables B_1 and B_2 that take values in Abelian groups \mathcal{B}_1 and \mathcal{B}_2 and are subject to the uniform distributions. The pair of random variables (B_1, B_2) is used for sending the all of secret messages in $\mathcal{S}_1 \times \dots \times \mathcal{S}_T$. Assuming that $\mathcal{S}_1 \times \dots \times \mathcal{S}_T$ has an Abelian group structure, we give the random coding Φ_c and Φ_p in the same way as Code Ensemble 1 with $\mathcal{S}_c = \mathcal{S}_0 \times \mathcal{B}_1$ and $\mathcal{S}_p = \mathcal{B}_2$.

Second Step: We choose an ensemble satisfying Condition 15 of isomorphisms F' from $\mathcal{S}_1 \times \dots \times \mathcal{S}_T$ to $\mathcal{B}_1 \times \mathcal{B}_2$ as Abelian groups. We choose the random variable $G' \in \mathcal{B}_1 \times \mathcal{B}_2$ that

obeys the uniform distribution on $\mathcal{B}_1 \times \mathcal{B}_2$ and is independent of the choice of F' and anything else. Then, we define a map $\Lambda_{F', G'}(s) := F'(s) + G'$. Combining the above codes, we construct the code $\Phi_a = \Phi_p \circ \Lambda_{F', G'} : \mathcal{S}_0 \times \mathcal{S}_1 \times \dots \times \mathcal{S}_T \rightarrow \mathcal{V}$ as $(s_0, s_1, \dots, s_T) \mapsto \Phi_p(s_0, \Lambda_{F', G'}(s_1, \dots, s_T))$. Similar to the case of BCD, Bob's decoder Φ_b and Eve's decoder Φ_e are defined as the maximum likelihood decoders. Hence, our code is written by the triple (Φ_a, Φ_b, Φ_e) . The structure of encoder is illustrated in Fig. 1.

As a special case of Code Ensemble 3, a wire-tap code is given as the case when $T = 2$ and we do not have the random variables S_0 . For a fixed code φ_p , $P_{Z|S_0=s_0, \Phi_p=\varphi_p}$ denotes the average output distribution of the channel of the transmitted codeword $\varphi_p(s_0, B_1, B_2)$ averaged over B_1, B_2 . In order to evaluate the averaged performance of the above code (Φ_a, Φ_b, Φ_e) , we prepare the following lemma.

Lemma 21: When the code Φ_p is fixed to φ_p in the BCD part, we have the following average performance.

$$\begin{aligned} & \mathbf{E}_{F', G'} \exp(\rho I(S_I; Z|S_0)[P_{Z|V}, \varphi_p \circ \Lambda_{F', G'}, P_{S_T}]) \\ & \leq \mathbf{E}_{F', G'} \sum_{s_0} P_{S_0}(s_0) \sum_{S_I} P_{S_I|S_0}(S_I|s_0) \\ & \quad \cdot e^{\rho D(P_{Z|S_I=S_I, S_0=s_0, \Phi_p=\varphi_p} \| P_{Z|S_0=s_0, \Phi_p=\varphi_p})} \\ & \leq 1 + \sum_{s_0} P_{S_0}(s_0) \sum_{S_I} P_{S_I|S_0}(S_I|s_0) e^{-\rho H_{1+\rho}(S_{I^c}|S_I, S_0=s_0)} \\ & \quad \cdot e^{\psi(\rho|P_{Z|B_1, B_2, S_0=s_0, \Phi_p=\varphi_p}, P_{\text{mix}, \mathcal{B}_1, \mathcal{B}_2})}. \end{aligned} \quad (67)$$

Further, when $P_{Z|V}$ is a regular channel and the map $\varphi_p|_{S_0=s_0} : (b_1, b_2) \mapsto \varphi_p(b_1, b_2, s_0)$ is a homomorphism from an Abelian group $\mathcal{B}_1 \times \mathcal{B}_2$ to an Abelian group \mathcal{V} for any $s_0 \in \mathcal{S}_0$, the inequalities (67) hold even when G' is a constant g' .

Lemma 21 will be applied for the evaluation of the performance of Code Ensemble 3. However, it will be also used for the evaluation of the performance of another type of codes without common messages based on a specific error correcting code in Section XI. Hence, Lemma 21 addresses the case when the map $\varphi_p|_{S_0=s_0}$ is a homomorphism.

Lemma 21 yields the following observation. Applying Jensen's inequality for the convex function $x \mapsto e^x$ and the inequality $\log(1+x) \leq x$, we obtain

$$\begin{aligned} & \mathbf{E}_{F', G'} \rho I(S_I; Z|S_0)[P_{Z|V}, \varphi_p \circ \Lambda_{F', G'}, P_{S_T}] \\ & \leq \log \left(1 + \sum_{s_0} P_{S_0}(s_0) \sum_{S_I} P_{S_I|S_0}(S_I|s_0) e^{-\rho H_{1+\rho}(S_{I^c}|S_I, S_0=s_0)} \right. \\ & \quad \left. \cdot e^{\psi(\rho|P_{Z|B_1, B_2, S_0=s_0, \Phi_p=\varphi_p}, P_{\text{mix}, \mathcal{B}_1, \mathcal{B}_2})} \right) \\ & \leq \sum_{s_0} P_{S_0}(s_0) \sum_{S_I} P_{S_I|S_0}(S_I|s_0) e^{-\rho H_{1+\rho}(S_{I^c}|S_I, S_0=s_0)} \\ & \quad \cdot e^{\psi(\rho|P_{Z|B_1, B_2, S_0=s_0, \Phi_p=\varphi_p}, P_{\text{mix}, \mathcal{B}_1, \mathcal{B}_2})}. \end{aligned} \quad (68)$$

Proof: Applying (56) and (57) to the case when $\tilde{P}_Z = \tilde{P}_{Z|S_0=s_0, \Phi_p=\varphi_p}$, we obtain

$$\begin{aligned} & \mathbf{E}_{F', G'} e^{\rho I(S_I; Z|S_0)[P_{Z|V}, \varphi_p \circ \Lambda_{F', G'}, P_{S_T}]} \\ & \leq \mathbf{E}_{F', G'} \sum_{s_0} P_{S_0}(s_0) \sum_{s_I} P_{S_I|S_0}(s_I|s_0) \\ & \quad \cdot e^{\rho D(P_{Z|S_I=s_I, S_0=s_0, \Phi_p=\varphi_p} \| P_{Z|S_0=s_0, \Phi_p=\varphi_p})} \\ & \leq \mathbf{E}_{F', G' | \Phi_p=\varphi_p} \sum_{s_0} P_{S_0}(s_0) \sum_{s_I} P_{S_I|S_0}(s_I|s_0) \\ & \quad \cdot e^{\psi(\rho) P_{Z|S_I=s_I, S_0=s_0, \Phi_p=\varphi_p} \| P_{Z|S_0=s_0, \Phi_p=\varphi_p}}. \end{aligned} \quad (69)$$

For a fixed s_I , we apply Theorem 17 to the case when \mathcal{A} is S_{I^c} , \mathcal{X} is $\mathcal{B}_1 \times \mathcal{B}_2$, G is $G' + F'(s_I, 0)$, which is independent of F' , and F is the map $s_{I^c} \mapsto F'(0, s_{I^c})$ that satisfies Condition 15. Then, $\Lambda_{F', G'}(s_I, s_{I^c}) = F'(s_I, s_{I^c}) + G' = F'(0, s_{I^c}) + Z_{s_I}$. Thus, we obtain

$$\begin{aligned} & \mathbf{E}_{F', G'} e^{\psi(\rho) P_{Z|S_I=s_I, S_0=s_0, \Phi_p=\varphi_p} \| \tilde{P}_{Z|S_0=s_0, \Phi_p=\varphi_p}} \\ & \leq 1 + e^{-\rho H_{1+\rho}(S_{I^c}|S_I=s_I, S_0=s_0)} e^{\psi(\rho) P_{Z|\mathcal{B}_1, \mathcal{B}_2, S_0, \Phi_p=\varphi_p} \cdot P_{\text{mix}, \mathcal{B}_1, \mathcal{B}_2}}. \end{aligned} \quad (70)$$

Thus, we obtain (67).

Further, when $P_{Z|V}$ is a regular channel and the map $\varphi_p|_{S_0=s_0} : (b_1, b_2) \mapsto \varphi_p(b_1, b_2, s_0)$ is a homomorphism from an Abelian group $\mathcal{B}_1 \times \mathcal{B}_2$ to an Abelian group \mathcal{V} for any $s_0 \in \mathcal{S}_0$, the channel $P_{Z|V} \circ \varphi_p|_{S_0=s_0}$ is a regular channel from $\mathcal{B}_1 \times \mathcal{B}_2$ to \mathcal{V} . Hence, due to Corollary 18, the inequalities (67) hold even when G' is a constant g' . ■

Using the above lemma, we obtain the following theorem, which gives the averaged performance of the above code (Φ_a, Φ_b, Φ_e) . By using this theorem, we will give the capacity region in Subsection IX-B.

Theorem 22: Assume that the code $\Phi = (\Phi_a, \Phi_b, \Phi_e)$ is the ensemble given in Code Ensemble 3. Then, the inequalities

$$\begin{aligned} & \mathbf{E}_{\Phi_a} \exp(\rho I(S_I; Z|S_0)[P_{Z|V}, \Phi_a, P_{S_T}]) \\ & \leq \mathbf{E}_{\Phi_a} \sum_{s_0} P_{S_0}(s_0) \sum_{s_I} P_{S_I|S_0}(s_I|s_0) e^{\rho D(P_{Z|S_I=s_I, S_0=s_0, \Phi_a} \| P_{Z|S_0=s_0, \Phi_p})} \\ & \leq 1 + |\mathcal{B}_1|^\rho e^{-\rho H_{1+\rho}(S_{I^c}|S_I, S_0)+E_0(\rho|P_{Z|V}, P_{V|U}, P_U)}, \end{aligned} \quad (71)$$

and

$$\begin{aligned} \mathbf{E}_{\Phi} P_b[P_{Y|V}, \Phi, P_{S_T}] & \leq |\mathcal{B}_2|^\rho e^{E_0(-\rho|P_{Y|V}, P_{V|U}, P_U)} \\ & \quad + (|\mathcal{S}_0||\mathcal{S}|)^\rho e^{E_0(-\rho|P_{Y|UV}, P_{UV})} \end{aligned} \quad (72)$$

$$\mathbf{E}_{\Phi} P_e[P_{Z|V}, \Phi, P_{S_T}] \leq |\mathcal{S}_0|^\rho e^{E_0(-\rho|P_{Z|U}, P_U)}. \quad (73)$$

hold.

Theorem 22 yields the following observation. Applying Jensen's inequality to the convex function $x \mapsto e^x$, we obtain

$$\begin{aligned} & \mathbf{E}_{\Phi_a} \rho I(S_I; Z|S_0)[P_{Z|V}, \Phi_a, P_{S_T}] \\ & \leq \log(1 + |\mathcal{B}_1|^\rho e^{-\rho H_{1+\rho}(S_{I^c}|S_I, S_0)+E_0(\rho|P_{Z|V}, P_{V|U}, P_U)}) \\ & \leq |\mathcal{B}_1|^\rho e^{-\rho H_{1+\rho}(S_{I^c}|S_I, S_0)+E_0(\rho|P_{Z|V}, P_{V|U}, P_U)}. \end{aligned} \quad (74)$$

Here, we choose ρ_0 as

$$\begin{aligned} \rho_0 := \operatorname{argmin}_{\rho \in [0,1]} & \left[\log |\mathcal{B}_1| + \frac{1}{\rho} E_0(\rho|P_{Z|V}, P_{V|U}, P_U) \right. \\ & \left. - H_{1+\rho}(S_{I^c}|S_I, S_0) \right]_+ + (T+2) \frac{\log 2}{\rho}. \end{aligned} \quad (75)$$

Then, Similar to (35) and (36), since $2(2^T - 2) + 2 = 2^{T+1} - 2 < 2^{T+1}$, Markov inequality guarantees that there exists a code $\varphi = (\varphi_a, \varphi_b, \varphi_e)$ such that

$$\begin{aligned} & \exp(\rho_0 I(S_I; Z|S_0)[P_{Z|V}, \varphi_a, P_{S_T}]) \\ & \leq 2^{T+1} (1 + |\mathcal{B}_1|^{\rho_0} e^{-\rho_0 H_{1+\rho_0}(S_{I^c}|S_I, S_0)+E_0(\rho_0|P_{Z|V}, P_{V|U}, P_U)}) \\ & \leq 2^{T+2} e^{[\rho_0 \log |\mathcal{B}_1| - \rho_0 H_{1+\rho_0}(S_{I^c}|S_I, S_0)+E_0(\rho_0|P_{Z|V}, P_{V|U}, P_U), P_{S_T}]_+}, \end{aligned} \quad (76)$$

$$\begin{aligned} & I(S_I; Z|S_0)[P_{Z|V}, \varphi_a, P_{S_T}] \\ & \leq \min_{0 \leq \rho \leq 1} \frac{2^{T+1}}{\rho} |\mathcal{B}_1|^\rho e^{-\rho H_{1+\rho}(S_{I^c}|S_I, S_0)+E_0(\rho|P_{Z|V}, P_{V|U}, P_U)}, \end{aligned} \quad (77)$$

$$\begin{aligned} & P_b[P_{Y|V}, \varphi, P_{S_T}] \\ & \leq 2^{T+1} \min_{0 \leq \rho \leq 1} (|\mathcal{B}_2|^\rho e^{E_0(-\rho|P_{Y|V}, P_{V|U}, P_U)} + (|\mathcal{S}_0||\mathcal{S}|)^\rho e^{E_0(-\rho|P_{Y|UV}, P_{UV})}), \end{aligned} \quad (78)$$

$$\begin{aligned} & P_e[P_{Z|V}, \varphi, P_{S_T}] \\ & \leq 2^{T+1} \min_{0 \leq \rho \leq 1} |\mathcal{S}_0|^\rho e^{E_0(-\rho|P_{Z|U}, P_U)} \end{aligned} \quad (79)$$

for any non-empty proper subset $I \subseteq \{1, \dots, T\}$. Taking the logarithm in (76), we obtain

$$\begin{aligned} & I(S_I; Z|S_0)[P_{Z|V}, \Phi_a, P_{S_T}] \\ & \leq \left[\log |\mathcal{B}_1| + \frac{1}{\rho_0} E_0(\rho_0|P_{Z|V}, P_{V|U}, P_U) - H_{1+\rho_0}(S_{I^c}|S_I, S_0) \right]_+ \\ & \quad + (T+2) \frac{\log 2}{\rho_0} \\ & = \min_{\rho \in [0,1]} \left[\log |\mathcal{B}_1| + \frac{1}{\rho} E_0(\rho|P_{Z|V}, P_{V|U}, P_U) - H_{1+\rho}(S_{I^c}|S_I, S_0) \right]_+ \\ & \quad + (T+2) \frac{\log 2}{\rho}. \end{aligned} \quad (80)$$

Proof of Theorem 22: We show (71). Using (17), we obtain

$$\begin{aligned} & \mathbf{E}_{\Phi_p, \Phi_c} e^{\psi(\rho) P_{Z|B_1, B_2, S_0=s_0, \Phi_p, P_{\text{mix}, B_1, B_2}}} \\ & \leq \mathbf{E}_{\Phi_p, \Phi_c} e^{E_0(\rho) P_{Z|B_1, B_2, S_0=s_0, \Phi_p, P_{\text{mix}, B_1, B_2}}} \end{aligned} \quad (81)$$

$$\begin{aligned} & = \mathbf{E}_{\Phi_p, \Phi_c} \sum_z \left(\sum_{b_1, b_2} P_{B_1, B_2}(b_1, b_2) P_{Z|B_1, B_2, S_0=s_0, \Phi_p}(z|b_1, b_2)^{\frac{1}{1-\rho}} \right)^{1-\rho} \\ & = \mathbf{E}_{\Phi_p, \Phi_c} \sum_z \left(\sum_{b_1, b_2} \frac{1}{|\mathcal{B}_1||\mathcal{B}_2|} P_{Z|V}(z|\Phi_p(s_0, b_1, b_2))^{\frac{1}{1-\rho}} \right)^{1-\rho} \end{aligned}$$

$$\leq \mathbf{E}_{\Phi_p, \Phi_c} \sum_z \sum_{b_1} \left(\sum_{b_2} \frac{1}{|\mathcal{B}_1||\mathcal{B}_2|} P_{Z|V}(z|\Phi_p(s_0, b_1, b_2))^{\frac{1}{1-\rho}} \right)^{1-\rho} \quad (82)$$

$$\begin{aligned} & = \mathbf{E}_{\Phi_p, \Phi_c} \sum_z \sum_{b_1} \frac{|\mathcal{B}_1|^\rho}{|\mathcal{B}_1|} \left(\sum_{b_2} \frac{1}{|\mathcal{B}_2|} P_{Z|V}(z|\Phi_p(s_0, b_1, b_2))^{\frac{1}{1-\rho}} \right)^{1-\rho} \\ & \quad (83) \end{aligned}$$

$$\begin{aligned} & \leq \mathbf{E}_{\Phi_c} \sum_z \sum_{b_1} \frac{|\mathcal{B}_1|^\rho}{|\mathcal{B}_1|} \left(\sum_{b_2} \frac{1}{|\mathcal{B}_2|} \mathbf{E}_{\Phi_p|\Phi_c} P_{Z|V}(z|\Phi_p(s_0, b_1, b_2))^{\frac{1}{1-\rho}} \right)^{1-\rho} \\ & \quad (84) \end{aligned}$$

$$\begin{aligned} & = \sum_z \sum_{b_1} \frac{|\mathcal{B}_1|^\rho}{|\mathcal{B}_1|} \mathbf{E}_{\Phi_c} \left(\sum_{b_2} \frac{1}{|\mathcal{B}_2|} \sum_v P_{V|U}(v|\Phi_c(s_0, b_1)) P_{Z|V}(z|v)^{\frac{1}{1-\rho}} \right)^{1-\rho} \\ & \quad (85) \end{aligned}$$

$$\begin{aligned} & = \sum_z \sum_{b_1} \frac{|\mathcal{B}_1|^\rho}{|\mathcal{B}_1|} \mathbf{E}_{\Phi_c} \left(\sum_v P_{V|U}(v|\Phi_c(s_0, b_1)) P_{Z|V}(z|v)^{\frac{1}{1-\rho}} \right)^{1-\rho} \\ & = \sum_z \sum_{b_1} \frac{|\mathcal{B}_1|^\rho}{|\mathcal{B}_1|} \sum_u P_U(u) \left(\sum_v P_{V|U}(v|u) P_{Z|V}(z|v)^{\frac{1}{1-\rho}} \right)^{1-\rho} \end{aligned}$$

$$\begin{aligned} & = \sum_z |\mathcal{B}_1|^\rho \sum_u P_U(u) \left(\sum_v P_{V|U}(v|u) P_{Z|V}(z|v)^{\frac{1}{1-\rho}} \right)^{1-\rho} \\ & = |\mathcal{B}_1|^\rho e^{E_0(\rho) P_{Z|V}, P_{V|U}, P_U}, \end{aligned} \quad (86)$$

where (81), (82) (84), and (85) follow from (17), the inequality $(x+y)^{1-\rho} \leq x^{1-\rho} + y^{1-\rho}$, the concavity of $x \mapsto x^{1-\rho}$, and the definition of the ensemble of the code Φ_p , respectively.

Summarizing the above discussion, we obtain

$$\begin{aligned} & \mathbf{E}_{\Phi_c} e^{\rho I(S_I; Z|S_0) [P_{Z|V}, \Phi_c, P_{S_I}]} \\ & \leq \mathbf{E}_{\Phi_c} \sum_{s_0} P_{S_0}(s_0) \sum_{S_I} P_{S_I|S_0}(S_I|s_0) e^{\rho D(P_{Z|B_1, B_2, S_0=s_0, \Phi_p} \| \bar{P}_{Z|S_0=s_0, \Phi_p})} \end{aligned} \quad (87)$$

$$\begin{aligned} & = \mathbf{E}_{\Phi_p} \mathbf{E}_{F', G'|\Phi_p} \sum_{s_0} P_{S_0}(s_0) \sum_{S_I} P_{S_I|S_0}(S_I|s_0) \\ & \quad \cdot e^{\rho D(P_{Z|B_1, B_2, S_0=s_0, \Phi_p} \| \bar{P}_{Z|S_0=s_0, \Phi_p})} \\ & \leq \sum_{s_0} P_{S_0}(s_0) \sum_{S_I} P_{S_I|S_0}(S_I|s_0) \\ & \quad \cdot \mathbf{E}_{\Phi_p} (1 + e^{-\rho H_{1+\rho}(S_I^c|S_I=S_I, S_0=s_0)} e^{\psi(\rho) P_{Z|B_1, B_2, S_0=s_0, \Phi_p}, P_{B_1, B_2}}) \end{aligned} \quad (88)$$

$$\begin{aligned} & \leq \sum_{s_0} P_{S_0}(s_0) \sum_{S_I} P_{S_I|S_0}(S_I|s_0) \\ & \quad \cdot (1 + e^{-\rho H_{1+\rho}(S_I^c|S_I=S_I, S_0=s_0)} |\mathcal{B}_1|^\rho e^{E_0(\rho) P_{Z|V}, P_{V|U}, P_U}) \\ & = 1 + e^{-\rho H_{1+\rho}(S_I^c|S_I, S_0)} |\mathcal{B}_1|^\rho e^{E_0(\rho) P_{Z|V}, P_{V|U}, P_U}, \end{aligned} \quad (89)$$

where (87), (88), and (89) follow from (56), the second inequality in Lemma 21, and (86), respectively. Then, we obtain (71).

Further, (72) and (73) follow from Lemma 12. \blacksquare

D. Group Symmetry

Next, when the channel has a nice property with respect to group action, we treat the upper bound of the leaked information with a fixed BCD code φ_p . That is, we discuss the upper bound given in Lemma 21 under an assumption for group action, which will be given latter. The following analysis is required for evaluation of universal coding in Sections XII and XIII and a practical code construction in Subsection XI-B.

For simplicity, we first discuss the case with no common message, i.e., $|\mathcal{S}_0| = 1$ and $|\mathcal{B}_1| = 1$. Assume that a group \mathcal{G} acts on \mathcal{V} and \mathcal{Z} . The action of $g \in \mathcal{G}$ is written as $g \cdot v$ and $g \cdot z$ for $v \in \mathcal{V}$ and $z \in \mathcal{Z}$. Then, due to Eqs. (2), (3), and (4), we have

$$\begin{aligned} (g^{-1} \circ P_{Z|V} \circ g)(z|v) & = P_{Z|V}(g \cdot z|g \cdot v) \\ (g^{-1} \circ P_V)(v) & = P_V(g \cdot v). \end{aligned}$$

Then, the set \mathcal{V} can be divided to orbits $\{\mathcal{V}_o\}_{o \in O}$ by the action of \mathcal{G} . The set O of indexes of the orbits is called the orbit space. Given a code φ_p as an injective map from \mathcal{B}_2 to \mathcal{V} , Recall that we denote the uniform distribution on the image $\text{Im } \varphi_p$ by $P_{\text{mix}, \text{Im } \varphi_p}$, and we define the distribution $P_{\varphi_p}(o) := |\text{Im } \varphi_p \cap \mathcal{V}_o| / |\text{Im } \varphi_p|$ on the orbit space O and the distribution \bar{P}_{φ_p} on \mathcal{V} by $\bar{P}_{\varphi_p}(v) := \frac{P_{\varphi_p}(o)}{|\mathcal{V}_o|}$ when the element v belongs to the subset \mathcal{V}_o . Then, we obtain the following lemma.

Lemma 23: When the relation $g^{-1} \circ P_{Z|V} \circ g = P_{Z|V}$ holds for any $g \in \mathcal{G}$, $v \in \mathcal{Z}$, and $v \in \mathcal{V}$,

$$\begin{aligned} \psi(\rho) P_{Z|B_2, \Phi_p=\varphi_p, P_{\text{mix}, B_2}} & = \psi(\rho) P_{Z|V}, P_{\text{mix}, \text{Im } \varphi_p} \\ \leq E_0(\rho) P_{Z|V}, P_{\text{mix}, \text{Im } \varphi_p} & \leq E_0(\rho) P_{Z|V}, \bar{P}_{\varphi_p}. \end{aligned} \quad (90)$$

In particular, when the image $\text{Im } \varphi_p$ is included in one orbit \mathcal{V}_o , \bar{P}_{φ_p} is the uniform distribution on the orbit \mathcal{V}_o .

Proof: Since $e^{E_0(\rho) [g^{-1} \circ P_{Z|V} \circ g, g^{-1} \circ P_{\text{mix}, \varphi_p}]} = e^{E_0(\rho) [P_{Z|V}, g^{-1} \circ P_{\text{mix}, \varphi_p}]}$, we have

$$\begin{aligned} & e^{\psi(\rho) P_{Z|V}, P_{\text{mix}, \text{Im } \varphi_p}} \leq e^{E_0(\rho) P_{Z|V}, P_{\text{mix}, \text{Im } \varphi_p}} \\ & = \sum_{g \in \mathcal{G}} \frac{1}{|\mathcal{G}|} e^{E_0(\rho) [g^{-1} \circ P_{Z|V} \circ g, g^{-1} \circ P_{\text{mix}, \text{Im } \varphi_p}]} \\ & = \sum_{g \in \mathcal{G}} \frac{1}{|\mathcal{G}|} e^{E_0(\rho) [P_{Z|V}, g^{-1} \circ P_{\text{mix}, \text{Im } \varphi_p}]} \\ & \leq e^{E_0(\rho) [P_{Z|V}, \sum_{g \in \mathcal{G}} \frac{1}{|\mathcal{G}|} g^{-1} \circ P_{\text{mix}, \text{Im } \varphi_p}]} = e^{E_0(\rho) P_{Z|V}, \bar{P}_{\varphi_p}}. \end{aligned} \quad (91)$$

Next, we consider the general case. Assume that a group \mathcal{G} acts on \mathcal{U} , \mathcal{V} , and \mathcal{Z} . The code pair code (φ_c, φ_p) is a map from $\mathcal{S}_0 \times \mathcal{B}_1 \times \mathcal{B}_2$ to $\mathcal{U} \times \mathcal{V}$. For a given $s_0 \in \mathcal{S}_0$, we define the maps $\varphi_c|_{S_0=s_0}$ and $(\varphi_c, \varphi_p)|_{S_0=s_0}$ by

$$\begin{aligned} \varphi_c|_{S_0=s_0}(b_1) & := \varphi_c(s_0, b_1) \in \mathcal{U} \\ (\varphi_c, \varphi_p)|_{S_0=s_0}(b_1, b_2) & := (\varphi_c(s_0, b_1), \varphi_p(s_0, b_1, b_2)) \in \mathcal{U} \times \mathcal{V}. \end{aligned}$$

For simplicity, we assume that the image of $(\varphi_c, \varphi_p)|_{S_0=s_0}$ is included in one orbit in $\mathcal{U} \times \mathcal{V}$, which is denoted by $(\mathcal{V} \times \mathcal{U})_o$. Hence, the image of $\varphi_c|_{S_0=s_0}$ is included in one orbit in \mathcal{U} , which is denoted by \mathcal{U}_o .

Lemma 24: Assume that the image of $(\varphi_c, \varphi_p)|_{S_0=s_0}$ is included in a orbit $(\mathcal{V} \times \mathcal{U})_o$ in $\mathcal{U} \times \mathcal{V}$. When the relation $g^{-1} \circ P_{Z|V} \circ g = P_{Z|V}$ holds for any $g \in \mathcal{G}$, the relation

$$e^{\psi(\rho|P_{Z|B_1, B_2, S_0=s_0, \Phi_p=\varphi_p}, P_{\text{mix}, B_1, B_2})} \leq |\mathcal{B}_1|^\rho e^{E_0(\rho|P_{Z|V}, P_{V|U, \text{mix}, (\mathcal{V} \times \mathcal{U})_o}, P_{\text{mix}, \mathcal{U}_o})} \quad (92)$$

holds for any $s_0 \in S_0$.

Proof: For a given $u \in \mathcal{U}_o$, we define the stabilizer of u by $\mathcal{H}_u := \{g \in \mathcal{G} | g \cdot u = u\}$, which is a subgroup of \mathcal{G} . For arbitrary $u \in \mathcal{U}_o$, we define the two subsets $\mathcal{V}'_u, \mathcal{V}_u \subset \mathcal{V}$ by $\{u\} \times \mathcal{V}'_u = \text{Im}(\varphi_c, \varphi_p)|_{S_0=s_0} \cap (\{u\} \times \mathcal{V})$ and $\{u\} \times \mathcal{V}_u = (\mathcal{V} \times \mathcal{U})_o \cap (\{u\} \times \mathcal{V})$. Then, we obtain the relations

$$P_{V|U=u, \text{mix}, \text{Im}(\varphi_c, \varphi_p)|_{S_0=s_0}} = P_{V|\text{mix}, \mathcal{V}'_u} \quad (93)$$

$$P_{V|U=u, \text{mix}, (\mathcal{V} \times \mathcal{U})_o} = P_{V|\text{mix}, \mathcal{V}_u}. \quad (94)$$

For the definitions of the left hand sides, see (1). We can also show that

$$\cup_{g \in \mathcal{H}_u} \{g \cdot v | v \in \mathcal{V}'_u\} = \mathcal{V}_u.$$

Since $g^{-1} \circ P_{V|U=g \cdot u, \text{mix}, (\mathcal{V} \times \mathcal{U})_o} = P_{V|U=u, \text{mix}, (\mathcal{V} \times \mathcal{U})_o}$, the condition $g^{-1} \circ P_{Z|V} \circ g = P_{Z|V}$ implies that

$$e^{E_0(\rho|g^{-1} \circ P_{Z|V} \circ g, g^{-1} \circ P_{V|U=g \cdot u, \text{mix}, (\mathcal{V} \times \mathcal{U})_o})} = e^{E_0(\rho|P_{Z|V}, P_{V|U=u, \text{mix}, (\mathcal{V} \times \mathcal{U})_o})}. \quad (95)$$

We obtain the following relations. In the following derivation, (96) and (98) follow from (83) and (95), respectively. Applying Lemma 23 to the case of $\mathcal{G} = \mathcal{H}_u$, we obtain the inequality (97) from (93) and (94).

$$\begin{aligned} & e^{\psi(\rho|P_{Z|B_1, B_2, S_0=s_0, \Phi_p=\varphi_p}, P_{\text{mix}, B_1, B_2})} \\ & \leq \sum_z \sum_{b_1} \frac{|\mathcal{B}_1|^\rho}{|\mathcal{B}_1|} \left(\sum_{b_2} \frac{1}{|\mathcal{B}_2|} P_{Z|V}(z | \varphi_p(s_0, b_1, b_2)) \right)^{\frac{1}{1-\rho}} \quad (96) \\ & = |\mathcal{B}_1|^\rho \sum_z \sum_u P_{U, \text{mix}, \text{Im} \varphi_c|_{S_0=s_0}}(u) \\ & \quad \cdot \left[\sum_v P_{V|U=u, \text{mix}, \text{Im}(\varphi_c, \varphi_p)|_{S_0=s_0}}(v) P_{Z|V}(z|v) \right]^{\frac{1}{1-\rho}} \\ & = |\mathcal{B}_1|^\rho \sum_u P_{U, \text{mix}, \text{Im} \varphi_c|_{S_0=s_0}}(u) e^{E_0(\rho|P_{Z|V}, P_{V|U=u, \text{mix}, \text{Im}(\varphi_c, \varphi_p)|_{S_0=s_0})} \\ & \leq |\mathcal{B}_1|^\rho \sum_u P_{U, \text{mix}, \text{Im} \varphi_c|_{S_0=s_0}}(u) e^{E_0(\rho|P_{Z|V}, P_{V|U=u, \text{mix}, (\mathcal{V} \times \mathcal{U})_o})} \quad (97) \\ & = |\mathcal{B}_1|^\rho \sum_{g \in \mathcal{G}} \frac{1}{|\mathcal{G}|} \sum_u P_{U, \text{mix}, \text{Im} \varphi_c|_{S_0=s_0}}(g \cdot u) e^{E_0(\rho|P_{Z|V}, P_{V|U=u, \text{mix}, (\mathcal{V} \times \mathcal{U})_o})} \\ & = |\mathcal{B}_1|^\rho \sum_u P_{U, \text{mix}, (\mathcal{V} \times \mathcal{U})_o}(u) e^{E_0(\rho|P_{Z|V}, P_{V|U=u, \text{mix}, (\mathcal{V} \times \mathcal{U})_o})} \\ & = |\mathcal{B}_1|^\rho e^{E_0(\rho|P_{Z|V}, P_{V|U=u, \text{mix}, (\mathcal{V} \times \mathcal{U})_o}, P_{U, \text{mix}, \mathcal{U}_o})}. \end{aligned} \quad (98)$$

Remark 25: Section VII deals with the security when a channel $P_{Z|V}$ from \mathcal{V} to \mathcal{Z} is given. The discussion of Section VII can be extended to the case with a channel $P_{Z|VU}$ from $\mathcal{V} \times \mathcal{U}$ to \mathcal{Z} . In this case, $\psi(\rho|P_{Z|V}, P_{V|U}, P_U)$ and

$E_0(\rho|P_{Z|V}, P_{V|U}, P_U)$ are modified to

$$\begin{aligned} & \psi(\rho|P_{Z|V, U}, P_{V|U}, P_U) \\ & := \log \sum_u P_U(u) \sum_v P_{V|U}(v|u) \sum_z P_{Z|V, U}(z|v, u)^{1+\rho} P_{Z|U}(z|u)^{-\rho} \\ & E_0(\rho|P_{Z|V, U}, P_{V|U}, P_U) \\ & := \log \sum_u P_U(u) \sum_z \left(\sum_v P_{V|U}(v|u) P_{Z|V, U}(z|v, u)^{1/(1-\rho)} \right)^{1-\rho}. \end{aligned}$$

All of the discussions in this section are still valid even if we replace $P_{Z|V}(z|v)$ by $P_{Z|V, U}(z|v, u)$ with the above modification. These extensions to the channel $P_{Z|VU}$ will be used in Section XII as a mathematical tool for our proof.

VIII. ASYMPTOTIC CONDITIONAL UNIFORMITY

A. Three Kinds of Asymptotic Conditional Uniformity Conditions

In SMC, we use the message S_{I^c} as a dummy message. The secrecy of the message S_I depends on the conditional entropy of the dummy message S_{I^c} given S_I . Then, it is not easy to treat the asymptotic performance without fixing the conditional entropy rate of the dummy message S_{I^c} . Hence, we need to characterize the randomness of the dummy message S_{I^c} under the condition with respect to S_I in the asymptotic setting. In order to treat the capacity region and the strong security, we introduce several kinds of asymptotic conditional uniformity conditions for a general sequence of source distributions $P_{S_{T,n}}$ on the message sets $\mathcal{S}_{i,n}$ for $i = 0, 1, \dots, T$ satisfying the relations $|\mathcal{S}_{i,n}| := e^{nR_i}$ for $i = 0, 1, \dots, T$.

Definition 26: The sequence of distributions $P_{S_{T,n}}$ of the dummy message $S_{I^c, n}$ is called *weak asymptotically conditionally uniform (WACU)* for a non-empty proper subset $\mathcal{I}(\neq \emptyset) \subseteq \{1, \dots, T\}$ when

$$\lim_{n \rightarrow \infty} \frac{1}{n} H(S_{I^c, n} | S_{\mathcal{I}, n}, S_{0, n}) = \sum_{i \in \mathcal{I}^c} R_i. \quad (99)$$

Definition 27: The sequence of distributions $P_{S_{T,n}}$ of the dummy message $S_{I^c, n}$ is called *semi-weak asymptotically conditionally uniform (SWACU)* for a non-empty proper subset $\mathcal{I}(\neq \emptyset) \subseteq \{1, \dots, T\}$ when the relation

$$\lim_{n \rightarrow \infty} \frac{1}{n} H_{1+\frac{\delta}{n}}(S_{I^c, n} | S_{\mathcal{I}, n}, S_{0, n}) = \sum_{i \in \mathcal{I}^c} R_i \quad (100)$$

holds for any $\delta > 0$.

Definition 28: Fix an arbitrary fixed real number $\epsilon \geq 0$. The sequence of distributions $P_{S_{T,n}}$ of the dummy message $S_{I^c, n}$ is called ϵ -*strong asymptotically conditionally uniform (ϵ -SACU)* for a non-empty proper subset $\mathcal{I}(\neq \emptyset) \subseteq \{1, \dots, T\}$ when the relation

$$\underline{H}_{\log}(\mathcal{I}^c) \geq \sum_{i \in \mathcal{I}^c} (R_i - \epsilon), \quad (101)$$

where

$$\underline{H}_{\log}(\mathcal{I}^c) := \lim_{\delta \rightarrow \infty} \liminf_{n \rightarrow \infty} \frac{1}{n} H_{1+\frac{\delta \log n}{n}}(S_{I^c, n} | S_{\mathcal{I}, n}, S_{0, n}). \quad (102)$$

Since $\rho - 1$ behaves as $\delta \frac{\log n}{n}$ in (102), we use the subscript \log in (102). In the case of $\epsilon = 0$, it is simply called *strong*

asymptotically conditionally uniform (SACU) for a non-empty proper subset $\mathcal{I} (\neq \emptyset) \subseteq \{1, \dots, T\}$. In this case, the condition (101) is equivalent with

$$H_{\log}(\mathcal{I}^c) = \sum_{i \in \mathcal{I}^c} R_i \quad (103)$$

because the opposite inequality holds due to the cardinalities of respective message sets.

In particular, when the sequence of distributions $P_{S_{\mathcal{T},n}}$ of the dummy message $S_{\mathcal{I}^c,n}$ is WACU for any non-empty proper subset $\mathcal{I} \subseteq \{1, \dots, T\}$, it is simply called WACU. We sometimes fix a family \mathbf{J} of non-empty proper subsets \mathcal{I} of $\{1, \dots, T\}$, and treat only non-empty proper subsets $\mathcal{I} \in \mathbf{J}$. In this case, we call the sequence of distributions $P_{S_{\mathcal{T},n}}$ WACU for a family \mathbf{J} when it is WACU for any non-empty proper subset $\mathcal{I} \in \mathbf{J}$. We also apply these conventions to SWACU, SACU, and ϵ -SACU. The relations among the above conditions are summarized as follows.

Theorem 29: The following relations hold.

$$\begin{array}{ccc} \text{SACU} & \Rightarrow & \text{SWACU} \Leftrightarrow \text{WACU} \\ \downarrow & & \\ \epsilon\text{-SACU} & & \end{array}$$

Proof: The equivalence between SWACU and WACU will be shown as Lemma 93 in Appendix C. Other relations are trivial from their definitions. ■

In fact, as is shown in Subsection VIII-B, even if the original information does not satisfy the WACU condition (99) or the SACU condition (103) with $\epsilon = 0$, if we apply Slepian-Wolf data compression [30] to the original sources so that the total compressed rate of the whole data attains the entropy rate of the whole sources, the compressed data satisfies the WACU condition (99) and/or the SACU condition (103). Similarly, as is shown in Subsection VIII-B, even if the original information does not satisfy the ϵ -SACU condition (101), if we apply Slepian-Wolf data compression [30] to the original sources so that the error probability goes to zero exponentially and the difference between the entropy rate of the whole system and the total compressed rate is less than ϵ , the compressed data satisfies the ϵ -SACU condition (101).

B. Asymptotic Conditional Uniformity Conditions and Slepian-Wolf Data Compression

In Subsection X-A, we have introduced several asymptotic conditional uniformity conditions. In this subsection, we clarify which kind of data compressed by Slepian-Wolf compression satisfies asymptotic conditional uniformity conditions. For this purpose, we assume that the random variables $S_{\mathcal{T}}^n = (S_0^n, S_1^n, \dots, S_T^n)$ are subject to the n -fold stationary ergodic joint distribution $P_{S_{\mathcal{T}}}^n$ over $S_0^n \times S_1^n \times \dots \times S_T^n$. The symbols $H(S_0, \dots, S_T)$, $H(S_{\mathcal{I}})$, and $H(S_0, S_{\mathcal{I}})$ describe the entropy rates of the respective random variables for any non-empty proper subset $\mathcal{I} \subseteq \{1, \dots, T\}$. The following theorem treats the WACU condition for the compressed data.

Theorem 30: We choose the asymptotic compression rates R_0, \dots, R_T such that $\sum_{i=0}^T R_i = H(S_0, \dots, S_T)$ and $\sum_{i \in \mathcal{I}} R_i \leq H(S_{\mathcal{I}})$, $R_0 + \sum_{i \in \mathcal{I}} R_i \leq H(S_0, S_{\mathcal{I}})$ for any non-empty proper

subset $\mathcal{I} \subseteq \{1, \dots, T\}$. Choose a sequence m_n such that $\frac{m_n}{n} \rightarrow 1$.

Let $\varphi_i^n : S_i^{m_n} \rightarrow \{1, \dots, \lceil e^{nR_i} \rceil\}$ be Slepian-Wolf encoders and $\hat{\varphi}^n : \{1, \dots, \lceil e^{nR_0} \rceil\} \times \dots \times \{1, \dots, \lceil e^{nR_T} \rceil\} \rightarrow S_0^{m_n} \times \dots \times S_T^{m_n}$ be its Slepian-Wolf decoder for any positive integer n such that

$$\varepsilon(\varphi^n, \hat{\varphi}^n) := \Pr\{(S_0^{m_n}, \dots, S_T^{m_n}) \neq \hat{\varphi}^n(\varphi_0^n(S_0^{m_n}), \dots, \varphi_T^n(S_T^{m_n}))\} \rightarrow 0, \quad (104)$$

where $\varphi^n = (\varphi_0^n, \dots, \varphi_T^n)$. Then, we have

$$\lim_{n \rightarrow \infty} \frac{1}{n} H((\varphi_i^n(S_i^{m_n}))_{i \in \mathcal{I}^c} | (\varphi_i^n(S_i^{m_n}))_{i \in \mathcal{I}}, \varphi_0^n(S_0^{m_n})) = \sum_{i \in \mathcal{I}^c} R_i \quad (105)$$

for any non-empty proper subset $\mathcal{I} \subseteq \{1, \dots, T\}$. That is, the compressed data satisfies the WACU condition (99).

Remark 31: Theorem 30 gives only a sufficient condition (104) for the compressed data satisfying the WACU condition. For construction of the compressed data satisfying the WACU condition, it is needed to clarify the existence of a code whose the compressed data satisfying the condition (104).

In the single terminal Markovian case, under the condition $\frac{m_n}{n} \rightarrow 1$, the second order asymptotic analysis in [16, Section VII] guarantees that there exists sequence of the pairs of an encoder and a decoder satisfying (104) if and only if $\frac{n-m_n}{\sqrt{n}} \rightarrow \infty$. The extension to the Slepian-Wolf coding has been done with the i.i.d. case [32]. For the boundary of the attainable rate region of Slepian-Wolf data compression in the stationary ergodic case [5], we can show the existence of the pair of an encoder and a decoder satisfying (104) with a suitable choice of the sequence m_n under the condition $\frac{m_n}{n} \rightarrow 1$ in the following way⁵.

Choose the rates $R_i + \delta$ for any $\delta > 0$. Let $\varphi_{i,\delta}^n : S_i^n \rightarrow \{1, \dots, \lceil e^{nR_i(1+\delta)} \rceil\}$ be Slepian-Wolf encoders and $\hat{\varphi}_\delta^n : \{1, \dots, \lceil e^{nR_0(1+\delta)} \rceil\} \times \dots \times \{1, \dots, \lceil e^{nR_T(1+\delta)} \rceil\} \rightarrow S_0^n \times \dots \times S_T^n$ be its Slepian-Wolf decoder such that $\varepsilon(\varphi_\delta^n, \hat{\varphi}_\delta^n) \rightarrow 0$ with $\varphi_\delta^n := (\varphi_{0,\delta}^n, \dots, \varphi_{T,\delta}^n)$. For an arbitrary integer l , we choose an integer n_l such that the inequality $\varepsilon(\varphi_{1/l}^n, \hat{\varphi}_{1/l}^n) \leq \frac{1}{l}$ holds for any $n \geq n_l$. We define m_n to be $m_n := \lfloor \frac{n}{1+1/l} \rfloor$, where we choose l such that $n_l \leq n < n_{l+1}$. Here, we can choose the integer l for any positive integer n . The construction guarantees that $R_i(1+1/l)(m_n+1) \geq R_i n \geq R_i(1+1/l)m_n$. We define the pair of an encoder and a decoder $(\varphi^n, \hat{\varphi}^n)$ to be $(\varphi_{1/l}^{m_n}, \hat{\varphi}_{1/l}^{m_n})$. That is, φ_i^n is chosen to be $\varphi_{i,1/l}^{m_n}$. Our choices guarantee that $\frac{m_n}{n} \cong \frac{1}{1+1/l} \rightarrow 1$, and $\varepsilon(\varphi^n, \hat{\varphi}^n) = \varepsilon(\varphi_{1/l}^{m_n}, \hat{\varphi}_{1/l}^{m_n}) \leq 1/l \rightarrow 0$. In this construction, the encoder φ_i^n is a map from $S_i^{m_n}$ to $\{1, \dots, \lceil e^{m_n R_i(1+1/l)} \rceil\} \subset \{1, \dots, \lceil e^{nR_i} \rceil\}$ because $R_i n \geq m_n R_i(1+1/l)$. Hence, the pair of an encoder and a decoder $(\varphi^n, \hat{\varphi}^n)$ satisfies the assumption of Theorem 30.

Proof of Theorem 30: Assume that the code $\varphi^n = (\varphi_0^n, \dots, \varphi_T^n)$ satisfies (104). Since the stationary ergodic source satisfies the strong converse property for the data compression, due to folklore source coding theorem [14, Theorem 3.1], the

⁵The following discussion does not require any property for source distribution. That is, it can be extended to Slepian-Wolf data compression for the general information source [42] in the sense of Han-Verdú[13].

code φ^n satisfies

$$\lim_{n \rightarrow \infty} \frac{1}{n} H(\varphi_0^n(S_0^{m_n}), \dots, \varphi_T^n(S_T^{m_n})) = \sum_{i=0}^T R_i.$$

Since $\frac{1}{n} H((\varphi_i^n(S_i^{m_n}))_{i \in \mathcal{I}^c} | (\varphi_i^n(S_i^{m_n}))_{i \in \mathcal{I}}, \varphi_0^n(S_0^{m_n})) \leq \sum_{i \in \mathcal{I}^c} R_i$ and $\frac{1}{n} H((\varphi_i^n(S_i^{m_n}))_{i \in \mathcal{I}}, \varphi_0^n(S_0^{m_n})) \leq R_0 + \sum_{i \in \mathcal{I}} R_i$, we obtain (105). ■

In Subsection X-A, we have introduced the ϵ -strong asymptotic conditional uniformity (101) as another kind of asymptotic conditional uniformity. The following theorem shows the ϵ -strong asymptotic conditional uniformity for the compressed data.

Theorem 32: We fix a sequence m_n such that $\frac{m_n}{n} \rightarrow 1$. We also fix an arbitrary $\epsilon \geq 0$ and an arbitrary non-empty proper subset $\mathcal{I} \subseteq \{1, \dots, T\}$. Then, we choose the asymptotic compression rates R_0, \dots, R_T such that $\sum_{i=0}^T R_i = H(S_0, \dots, S_T) + \epsilon$ and

$$\sum_{i \in \mathcal{I}} R_i \leq H(S_{\mathcal{I}}), \quad R_0 + \sum_{i \in \mathcal{I}} R_i \leq H(S_0, S_{\mathcal{I}}). \quad (106)$$

We choose a Slepian-Wolf encoder $\varphi^n = (\varphi_0^n, \dots, \varphi_T^n)$ and a Slepian-Wolf decoder $\hat{\varphi}^n$ as a map $\varphi_i^n : \mathcal{S}_i^{m_n} \rightarrow \{1, \dots, \lceil e^{nR_i} \rceil\}$ and a map $\hat{\varphi}^n : \{1, \dots, \lceil e^{nR_0} \rceil\} \times \dots \times \{1, \dots, \lceil e^{nR_T} \rceil\} \rightarrow \mathcal{S}_0^{m_n} \times \dots \times \mathcal{S}_T^{m_n}$. When the decoding error probability $\varepsilon(\varphi^n, \hat{\varphi}^n)$ satisfies that

$$\varepsilon(\varphi^n, \hat{\varphi}^n) p(n) \rightarrow 0 \quad (107)$$

for any polynomial $p(n)$, the relation

$$\liminf_{n \rightarrow \infty} \frac{1}{n} H_{1+\rho_n}((\varphi_i^n(S_i^n))_{i \in \mathcal{I}^c} | (\varphi_i^n(S_i^n))_{i \in \mathcal{I}}, \varphi_0^n(S_0^n)) \geq \left(\sum_{i \in \mathcal{I}^c} R_i \right) - \epsilon \geq \sum_{i \in \mathcal{I}^c} (R_i - \epsilon) \quad (108)$$

holds with $\rho_n = \frac{\delta \log n}{n}$ for any $\delta > 0$. That is, the compressed data $(\varphi_0^n(S_0^n), \dots, \varphi_T^n(S_T^n))$ satisfies the ϵ -SACU condition (101) for the non-empty proper subset $\mathcal{I} \subseteq \{1, \dots, T\}$. In particular, in the case of $\epsilon = 0$, the compressed data $(\varphi_0^n(S_0^n), \dots, \varphi_T^n(S_T^n))$ satisfies the SACU condition for the non-empty proper subset $\mathcal{I} \subseteq \{1, \dots, T\}$.

Hence, if the relation (106) holds for any non-empty proper subset $\mathcal{I} \subseteq \{1, \dots, T\}$, the compressed data $(\varphi_0^n(S_0^n), \dots, \varphi_T^n(S_T^n))$ satisfies the ϵ -SACU condition (101).

Remark 33: Theorem 32 gives only a sufficient condition (107) for the compressed data satisfying the ϵ -SACU condition (101). Hence, it is necessary to clarify the existence of a code whose compressed data satisfying the condition (107).

In the i.i.d. case, for an arbitrary $\epsilon > 0$ and an arbitrary sequence m_n satisfying $\lim_{n \rightarrow \infty} \frac{m_n}{n} = 1$, there exists a sequence of Slepian-Wolf codes $(\varphi^n, \hat{\varphi}^n)$ with any rate tuples given in Theorem 32 such that the decoding error probability $\varepsilon(\varphi^n, \hat{\varphi}^n)$ goes to zero exponentially with respect to n [39]. That is, there exists a Slepian-Wolf code satisfying the condition (107) in Theorem 32. However, it is not so easy to give a required code in the case of $\epsilon = 0$. In Appendix B, we give such a code when $m_n := \frac{n}{1 + \frac{\epsilon}{t}}$ with $t > 1/2$ and $\infty > c > 0$.

C. Proof of Theorem 32

For the proof of Theorem 32, we prepare the following lemma for treating the relation between the conditional Rényi entropy of the compressed data and the decoding error probability. The following lemma treats the single terminal data compression for a random variable S on a set \mathcal{S} in the single-shot setting.

Lemma 34: Any encoder $\varphi : \mathcal{S} \rightarrow \{1, \dots, M\}$ and any decoder $\hat{\varphi} : \{1, \dots, M\} \rightarrow \mathcal{S}$ for a random variable S satisfy

$$e^{-\rho H_{1+\rho}(S)} \leq e^{-\rho H_{1+\rho}(\varphi(S))} \leq 2^\rho e^{-\rho H_{1+\rho}(S)} + 2^\rho \varepsilon(\varphi, \hat{\varphi})^{1+\rho}, \quad (109)$$

where $\varepsilon(\varphi, \hat{\varphi})$ is the decoding error probability $\Pr\{S \neq \hat{\varphi}(\varphi(S))\}$.

Proof: First, we show the first inequality. Using the inequality $x^{1+\rho} + y^{1+\rho} \leq (x+y)^{1+\rho}$ for $x, y \geq 0$, we obtain

$$\left(\sum_{s \in \varphi^{-1}(i)} P_S(s) \right)^{1+\rho} \geq \sum_{s \in \varphi^{-1}(i)} P_S(s)^{1+\rho}$$

for any $i = 1, \dots, M$. Hence,

$$\begin{aligned} e^{-\rho H_{1+\rho}(\varphi(S))} &= \sum_{i=1}^M \left(\sum_{s \in \varphi^{-1}(i)} P_S(s) \right)^{1+\rho} \\ &\geq \sum_{i=1}^M \sum_{s \in \varphi^{-1}(i)} P_S(s)^{1+\rho} = \sum_s P_S(s)^{1+\rho} = e^{-\rho H_{1+\rho}(S)}, \end{aligned}$$

which implies the first inequality of (109).

Next, we show the second inequality of (109). Given an arbitrary element i in the codebook, we have two cases: (1) The element $s_i := \hat{\varphi}(i)$ belongs to $\varphi^{-1}(i)$, i.e., there exists exact one element $s_i \in \varphi^{-1}(i)$ such that $\hat{\varphi}(\varphi(s_i)) = s_i$. (2) There exists no element $s_i \in \varphi^{-1}(i)$ such that $\hat{\varphi}(\varphi(s_i)) = s_i$. In case (1),

$$\begin{aligned} \left(\sum_{s \in \varphi^{-1}(i)} P_S(s) \right)^{1+\rho} &= \left(P_S(s_i) + \sum_{s \in \varphi^{-1}(i): \hat{\varphi}(\varphi(s)) \neq s} P_S(s) \right)^{1+\rho} \\ &= 2^{1+\rho} \left(\frac{1}{2} P_S(s_i) + \frac{1}{2} \sum_{s \in \varphi^{-1}(i): \hat{\varphi}(\varphi(s)) \neq s} P_S(s) \right)^{1+\rho} \\ &\leq 2^{1+\rho} \left(\frac{1}{2} P_S(s_i)^{1+\rho} + \frac{1}{2} \left(\sum_{s \in \varphi^{-1}(i): \hat{\varphi}(\varphi(s)) \neq s} P_S(s) \right)^{1+\rho} \right) \\ &= 2^\rho P_S(s_i)^{1+\rho} + 2^\rho \left(\sum_{s \in \varphi^{-1}(i): \hat{\varphi}(\varphi(s)) \neq s} P_S(s) \right)^{1+\rho}. \end{aligned}$$

In case (2),

$$\left(\sum_{s \in \varphi^{-1}(i)} P_S(s) \right)^{1+\rho} = \left(\sum_{s \in \varphi^{-1}(i): \hat{\varphi}(\varphi(s)) \neq s} P_S(s) \right)^{1+\rho}.$$

Hence, we obtain

$$\begin{aligned} e^{-\rho H_{1+\rho}(\varphi(S))} &= \sum_i \left(\sum_{s \in \varphi^{-1}(i)} P_S(s) \right)^{1+\rho} \\ &\leq 2^\rho \sum_i P_S(s_i)^{1+\rho} + 2^\rho \sum_i \left(\sum_{s \in \varphi^{-1}(i): \hat{\varphi}(\varphi(s)) \neq s} P_S(s) \right)^{1+\rho} \\ &\leq 2^\rho \sum_s P_S(s)^{1+\rho} + 2^\rho \left(\sum_i \sum_{s \in \varphi^{-1}(i): \hat{\varphi}(\varphi(s)) \neq s} P_S(s) \right)^{1+\rho} \quad (110) \\ &= 2^\rho \sum_s P_S(s)^{1+\rho} + 2^\rho \left(\sum_{s: \hat{\varphi}(\varphi(s)) \neq s} P_S(s) \right)^{1+\rho} \\ &= 2^\rho e^{-\rho H_{1+\rho}(S)} + 2^\rho \varepsilon(\varphi, \hat{\varphi})^{1+\rho}, \end{aligned}$$

where (110) follow from the inequality $x^{1+\rho} + y^{1+\rho} \leq (x+y)^{1+\rho}$ for $x, y \geq 0$. Hence, we obtain the second inequality. ■

Then, we obtain the following corollary of Lemma 34. The following corollary treats the single terminal data compression for a general sequence of random variables S_n .

Corollary 35: Let φ^n be an encoder and $\hat{\varphi}^n$ be a decoder for a general sequence of random variables S_n . When the decoding error probabilities $\varepsilon(\varphi^n, \hat{\varphi}^n)$ and the sequence $\{\rho_n\}$ of positive real numbers satisfy

$$\lim_{n \rightarrow \infty} \varepsilon(\varphi^n, \hat{\varphi}^n)^{1+\rho_n} e^{\rho_n H_{1+\rho_n}(S_n)} = 0, \quad (111)$$

we have

$$\lim_{n \rightarrow \infty} \frac{1}{n} H_{1+\rho_n}(\varphi^n(S_n)) = \lim_{n \rightarrow \infty} \frac{1}{n} H_{1+\rho_n}(S_n). \quad (112)$$

Proof of Corollary 35: The inequality $\lim_{n \rightarrow \infty} \frac{1}{n} H_{1+\rho_n}(\varphi^n(S_n)) \leq \lim_{n \rightarrow \infty} \frac{1}{n} H_{1+\rho_n}(S_n)$ follows from the first inequality (109). We show only the inequality $\lim_{n \rightarrow \infty} \frac{1}{n} H_{1+\rho_n}(\varphi^n(S_n)) \geq \lim_{n \rightarrow \infty} \frac{1}{n} H_{1+\rho_n}(S_n)$. Using the second inequality in (109), we have

$$\begin{aligned} & \lim_{n \rightarrow \infty} \frac{1}{n} H_{1+\rho_n}(\varphi^n(S_n)) = \lim_{n \rightarrow \infty} \frac{-1}{n \rho_n} \log e^{-\rho_n H_{1+\rho_n}(\varphi^n(S_n))} \\ & \geq \lim_{n \rightarrow \infty} \frac{-1}{n \rho_n} \log(2^{\rho_n} e^{-\rho_n H_{1+\rho_n}(S_n)} + 2^{\rho_n} \varepsilon(\varphi^n, \hat{\varphi}^n)^{1+\rho_n}) \\ & = \lim_{n \rightarrow \infty} \frac{-1}{n \rho_n} \log(2^{\rho_n} e^{-\rho_n H_{1+\rho_n}(S_n)}) \\ & = \lim_{n \rightarrow \infty} \frac{1}{n} (H_{1+\rho_n}(S_n) - \log 2) = \lim_{n \rightarrow \infty} \frac{1}{n} H_{1+\rho_n}(S_n), \end{aligned} \quad (113)$$

where (113) follows from the assumption (111). ■

Now, we show Theorem 32.

Proof of Theorem 32: For the proof of Theorem 32, we choose ρ'_n so that $\rho'_n(1 - \rho'_n) = \rho_n$. Since $\lim_{n \rightarrow \infty} \frac{m_n}{n} = 1$ and $\rho \geq \rho'_n$ for all n , we have

$$\begin{aligned} H_{1+\rho}(S_0, \dots, S_T) & \leq \liminf_{n \rightarrow \infty} \frac{1}{n} H_{1+\rho'_n}(S_0^{m_n}, \dots, S_T^{m_n}) \\ & \leq \limsup_{n \rightarrow \infty} \frac{1}{n} H_{1+\rho'_n}(S_0^{m_n}, \dots, S_T^{m_n}) \leq H(S_0, \dots, S_T). \end{aligned}$$

Since $\rho'_n \rightarrow 0$ and $\lim_{\rho \rightarrow +0} H_{1+\rho}(S_0, \dots, S_T) = H(S_0, \dots, S_T)$,

$$\lim_{n \rightarrow \infty} \frac{1}{n} H_{1+\rho'_n}(S_0^{m_n}, \dots, S_T^{m_n}) = H(S_0, \dots, S_T). \quad (114)$$

Since ρ'_n behaves as $\frac{\delta \log n}{n}$, due to the relation (114), the quantity $e^{\rho'_n H_{1+\rho'_n}(S_0^{m_n}, \dots, S_T^{m_n})}$ behaves as $e^{\delta(\log n)H(S_0, \dots, S_T)} = n^{\delta H(S_0, \dots, S_T)}$. Since $\varepsilon(\varphi^n, \hat{\varphi}^n)^{1+\rho'_n} \leq \varepsilon(\varphi^n, \hat{\varphi}^n)$, the condition (107) guarantees the condition (111). Hence, Corollary 35 guarantees that

$$\lim_{n \rightarrow \infty} \frac{1}{n} H_{1+\rho'_n}(\varphi_0^n(S_0^{m_n}), \dots, \varphi_T^n(S_T^{m_n})) = \left(\sum_{i=0}^T R_i \right) - \epsilon.$$

Since $\log |\varphi_0^n(S_0^{m_n}) \times \prod_{i \in \mathcal{I}} \varphi_i^n(S_i^{m_n})| = n(R_0 + \sum_{i \in \mathcal{I}} R_i)$, Corollary 87 in Appendix A implies (108). ■

IX. SECURE MULTIPLEX CODING WITH COMMON MESSAGES: ASYMPTOTIC PERFORMANCE

In this section, we treat the asymptotic performance for the secure multiplex coding with common messages when the channel is given as the n -fold discrete memoryless channel of a given broadcast channel $P_{YZ|X}$. First, we treat what performance can be achieved by using Code Ensemble 3 and Theorem 22 in Subsection VII-C without any assumption for the distribution of sources. In the next step, we define the capacity region under the asymptotic uniformity of information sources. In SMC, this restriction for the sources is essential for our definition of the capacity region. After this definition, we concretely give the capacity region.

A. General Sequence of Information Sources

First, we treat the secure multiplex coding with common messages with general sequence of information sources. For a given set of rates $(R_i)_{i=0}^T$, we give a general sequence of source distributions $P_{S_{i,n}}$ on the message sets $\mathcal{S}_{i,n}$ for $i = 0, 1, \dots, T$ satisfying the relations $|\mathcal{S}_{i,n}| := e^{nR_i}$ for $i = 0, 1, \dots, T$. For a given Markov chains $U \rightarrow V \rightarrow X \rightarrow YZ$, we give an asymptotic code construction in the following way.

Code Construction 4: Let φ_n be a code given in Code Ensemble 2 in Subsection VII-B satisfying (66), (63), (64), and (65) of length n with $|\mathcal{S}_{i,n}| := e^{nR_i}$ for $i = 0, 1, \dots, T$ and a given Markov chain $U \rightarrow V \rightarrow X$.

The performance of the code φ_n of Code Construction 4 is characterized as follows. The conditions (64) and (65) guarantee (115) and (116) given as follows.

$$\begin{aligned} & \liminf_{n \rightarrow \infty} \frac{-1}{n} \log P_b[P_{Y|V}^n, \Phi_n, P_{S_{\mathcal{I},n}}] \\ & \geq -\rho \sum_{i=1}^T R_i - \max[E_0(-\rho|P_{Y|V}, P_{V|U}, P_U), E_0(-\rho|P_{Y|U,V}, P_{V,U})], \end{aligned} \quad (115)$$

$$\liminf_{n \rightarrow \infty} \frac{-1}{n} \log P_e[P_{Z|V}^n, \Phi_n, P_{S_{\mathcal{I},n}}] \geq -\rho R_0 - E_0(-\rho|P_{Z|U}, P_U) \quad (116)$$

with any $\rho \in (0, 1]$. Further, due to (66), the leaked information for $S_{\mathcal{I},n}$ can be evaluated as

$$\begin{aligned} & \frac{1}{n} I(S_{\mathcal{I},n}; Z^n | S_{0,n}) [P_{Z|V}^n, \varphi_{a,n}, P_{S_{\mathcal{I},n}}] \\ & \leq \left[\frac{1}{\rho} \psi(\rho | P_{Z|V}, P_{V|U}, P_U) - \frac{1}{n} H_{1+\rho}(S_{\mathcal{I}^c, n} | S_{\mathcal{I},n}, S_{0,n}) \right]_+ \\ & \quad + (T+2) \frac{\log 2}{n\rho}. \end{aligned}$$

We substitute $\rho = a/n$ with an arbitrary real $a > 0$ and take the limits $n \rightarrow \infty$. Then, (20) of Lemma 4 leads the inequality

$$\begin{aligned} & \limsup_{n \rightarrow \infty} \frac{1}{n} I(S_{\mathcal{I},n}; Z^n | S_{0,n}) [P_{Z|V}^n, \varphi_{a,n}, P_{S_{\mathcal{I},n}}] \\ & \leq \left[I(V; Z|U) - \liminf_{n \rightarrow \infty} \frac{1}{n} H_{1+a/n}(S_{\mathcal{I}^c, n} | S_{\mathcal{I},n}, S_{0,n}) \right]_+ + (T+2) \frac{\log 2}{a}. \end{aligned}$$

Taking the limits $a \rightarrow \infty$, we obtain

$$\limsup_{n \rightarrow \infty} \frac{1}{n} I(S_{I,n}; Z^n | S_{0,n}) [P_{Z|V}^n, \varphi_{a,n}, P_{S_{T,n}}] \\ \leq \left[I(V; Z|U) - \lim_{a \rightarrow \infty} \liminf_{n \rightarrow \infty} \frac{1}{n} H_{1+a/n}(S_{I^c,n} | S_{I,n}, S_{0,n}) \right]_+ \quad (117)$$

So, the asymptotic performance of our code given in Code Construction 4 is characterized in (115), (116), and (117).

In Code Construction 4, the parameter R_0 is chosen to be R_c in BCD. However, to realize the capacity region of SMC, we need to choose the parameter R_0 to be a smaller value than R_c in BCD in general. To realize such a choice, we introduce another code construction by using Code Ensemble 3 in Subsection VII-C. As is explained in Remark 39, such a construction is crucial for achieving the capacity region in general although Code Construction 4 achieves the capacity region with no common message.

Code Construction 5: For a given set of rates $(R_i)_{i=0}^T$, we introduce other parameters R_p and R_c satisfying

$$R_c + R_p = \sum_{i=0}^T R_i, \quad R_c \geq R_0. \quad (118)$$

In the following, we denote the set of $((R_i)_{i=0}^T, R_p, R_c)$ satisfying the above condition by \mathcal{R}_T . In order to apply Code Ensemble 3 in Subsection VII-C, we fix Abelian groups $\mathcal{B}_{1,n}$ and $\mathcal{B}_{2,n}$ satisfying $|\mathcal{B}_{1,n}| := e^{n(R_c - R_0)}$ and $|\mathcal{B}_{2,n}| := e^{nR_p}$. Applying Code Ensemble 3 and Theorem 22 to the n -fold discrete memoryless extension $U^n \rightarrow V^n \rightarrow X^n \rightarrow Y^n Z^n$ of the above Markov chain and the Abelian groups $\mathcal{B}_{1,n}$ and $\mathcal{B}_{2,n}$, we find the code $\varphi_n = (\varphi_{a,n}, \varphi_{b,n}, \varphi_{e,n})$ with the message sets $\mathcal{S}_{i,n}$ for $i = 0, 1, \dots, T$ satisfying (76), (77), (78), and (79).

The performance of the code φ_n of Code Construction 5 is characterized as follows. The relations (78) and (79) guarantee that

$$\liminf_{n \rightarrow \infty} \frac{-1}{n} \log P_b[P_{Y|V}^n, \varphi_n, P_{S_{T,n}}] \\ \geq \min \left[-\rho R_p - E_0(-\rho |P_{Y|V}, P_{V|U}, P_U), \right. \\ \left. -\rho(R_p + R_c) - E_0(-\rho |P_{Y|U,V}, P_{V,U}) \right], \quad (119)$$

$$\liminf_{n \rightarrow \infty} \frac{-1}{n} \log P_e[P_{Z|V}^n, \varphi_n, P_{S_{T,n}}] \geq -\rho R_c - E_0(-\rho |P_{Z|U}, P_U) \quad (120)$$

for any $\rho \in (0, 1]$. Hence, due to (18) and (20), above both exponents (119) and (120) are positive, i.e., both error probabilities go to zero exponentially when

$$R_p < I(Y; V|U), \quad R_p + R_c < I(Y; VU) = I(Y; U) + I(Y; V|U), \\ R_c < I(Z; U),$$

which are satisfied when

$$R_c < \min[I(Y; U), I(Z; U)], \quad R_p < I(Y; V|U). \quad (121)$$

Further, due to (80), the leaked information for $S_{I,n}$ can be evaluated as

$$\frac{1}{n} I(S_{I,n}; Z^n | S_{0,n}) [P_{Z|V}^n, \varphi_{a,n}, P_{S_{T,n}}] \\ \leq \left[[R_c - R_0]_+ + \frac{1}{\rho} E_0(\rho |P_{Z|V}, P_{V|U}, P_U) - \frac{1}{n} H_{1+\rho}(S_{I^c,n} | S_{I,n}, S_{0,n}) \right]_+ \\ + (T+2) \frac{\log 2}{n\rho}.$$

Similar to (117), we obtain

$$\limsup_{n \rightarrow \infty} \frac{1}{n} I(S_{I,n}; Z^n | S_{0,n}) [P_{Z|V}^n, \varphi_{a,n}, P_{S_{T,n}}] \\ \leq \left[(R_c - R_0) + I(V; Z|U) \right. \\ \left. - \lim_{a \rightarrow \infty} \liminf_{n \rightarrow \infty} \frac{1}{n} H_{1+a/n}(S_{I^c,n} | S_{I,n}, S_{0,n}) \right]_+. \quad (122)$$

So, the asymptotic performance of our code in Code Construction 5 is characterized in (119), (120), and (122).

B. Capacity Region

Next, in order to characterize the limit of the asymptotic performance of the secure multiplex coding with common messages, we define the capacity region based on the WACU condition (99). For this purpose, we treat the transmission rate tuple $(R_i)_{i=0,\dots,T} = (R_0, R_1, \dots, R_T)$ and the information leakage rate tuple $(R_{l,I})_{\emptyset \neq I \subseteq \{1,\dots,T\}}$, where I takes every non-empty proper subset of $\{1, \dots, T\}$. The latter describes the rates of the leaked information for the message $S_{I,n}$. Combining both tuples, we call $((R_i)_{i=0,\dots,T}, (R_{l,I})_{\emptyset \neq I \subseteq \{1,\dots,T\}})$ the rate tuple.

Definition 36: The rate tuple $((R_i)_{i=0,\dots,T}, (R_{l,I})_{\emptyset \neq I \subseteq \{1,\dots,T\}})$ is said to be *achievable* for the secure multiplex coding with T secret messages for the channel $P_{YZ|X}$ if there exist a sequence of codes $\varphi_n = (\varphi_{a,n}, \varphi_{b,n}, \varphi_{e,n})$, i.e., Alice's stochastic encoder $\varphi_{a,n}$ from $\mathcal{S}_{0,n} \times \mathcal{S}_{1,n} \times \dots \times \mathcal{S}_{T,n}$ to \mathcal{X}^n , Bob's deterministic decoder $\varphi_{b,n} : \mathcal{Y}^n \rightarrow \mathcal{S}_{0,n} \times \mathcal{S}_{1,n} \times \dots \times \mathcal{S}_{T,n}$ and Eve's deterministic decoder $\varphi_{e,n} : \mathcal{Z}^n \rightarrow \mathcal{S}_{0,n}$ satisfying the following conditions: (1) The i -th secret message set $\mathcal{S}_{i,n}$ has cardinality e^{nR_i} for $i = 1, \dots, T$, and the common message set $\mathcal{S}_{0,n}$ has cardinality e^{nR_0} . (2) When a sequence of joint distributions $P_{S_{T,n}}$ on the message sets $\mathcal{S}_{i,n}$ for $T = 0, 1, \dots, T$ satisfies the WACU condition (99) for a non-empty proper subset $I (\neq \emptyset) \subseteq \{1, \dots, T\}$, the relations

$$\lim_{n \rightarrow \infty} P_b[P_{Y|X}^n, \varphi_n, P_{S_{T,n}}] = 0 \quad (123)$$

$$\lim_{n \rightarrow \infty} P_e[P_{Z|X}^n, \varphi_n, P_{S_{T,n}}] = 0 \quad (124)$$

$$\limsup_{n \rightarrow \infty} I(S_{I,n}; Z^n | S_0) [P_{Z|X}^n, \varphi_{a,n}, P_{S_{T,n}}] \leq R_{l,I} \quad (125)$$

hold. The capacity region \mathcal{C} of the secure multiplex coding is the closure of the achievable rate tuples $((R_i)_{i=0,\dots,T}, (R_{l,I})_{\emptyset \neq I \subseteq \{1,\dots,T\}})$.

Theorem 37: The capacity region of the secure multiplex coding with common messages is given by the set of rate tuples $((R_i)_{i=0,\dots,T}, (R_{l,I})_{\emptyset \neq I \subseteq \{1,\dots,T\}})$ such that there exist a Markov

chain $U \rightarrow V \rightarrow X \rightarrow YZ$ and

$$\begin{aligned} R_0 &\leq \min[I(U; Y), I(U; Z)], \\ \sum_{i=0}^T R_i &\leq I(V; Y|U) + \min[I(U; Y), I(U; Z)] \\ R_{l, \mathcal{I}} &\geq \sum_{i \in \mathcal{I}} R_i - [I(V; Y|U) - I(V; Z|U)]_+ \end{aligned} \quad (126)$$

for any non-empty proper subset $\mathcal{I} \subsetneq \{1, \dots, T\}$.

Now, we define the capacity region C_{nc} of the secure multiplex coding with no common messages as the set of rate tuples $((R_i)_{i=1, \dots, T}, (R_{l, \mathcal{I}})_{\emptyset \neq \mathcal{I} \subsetneq \{1, \dots, T\}})$ satisfying $(0, (R_i)_{i=1, \dots, T}, (R_{l, \mathcal{I}})_{\emptyset \neq \mathcal{I} \subsetneq \{1, \dots, T\}}) \in C$. As a corollary, the case with no common message is characterized as follows.

Corollary 38: C_{nc} is given as the set of rate tuples $((R_i)_{i=1, \dots, T}, (R_{l, \mathcal{I}})_{\emptyset \neq \mathcal{I} \subsetneq \{1, \dots, T\}})$ such that there exist a Markov chain $V \rightarrow X \rightarrow YZ$ and

$$\begin{aligned} \sum_{i=1}^T R_i &\leq I(V; Y) \\ R_{l, \mathcal{I}} &\geq \sum_{i \in \mathcal{I}} R_i - [I(V; Y) - I(V; Z)]_+ \end{aligned} \quad (127)$$

for any non-empty proper subset $\mathcal{I} \subsetneq \{1, \dots, T\}$.

Proof of Theorem 37: The converse part of this coding theorem follows from that for Corollary 9 with the uniform distribution on the whole message sets. The direct part can be shown by Lemma 41. That is, for a rate tuple $((R_i)_{i=1, \dots, T}, (R_{l, \mathcal{I}})_{\emptyset \neq \mathcal{I} \subsetneq \{1, \dots, T\}})$ given in (126) and an arbitrary small real number $\varepsilon > 0$, the rate tuple $((R_i - \frac{\varepsilon}{T})_{i=1, \dots, T}, (R_{l, \mathcal{I}})_{\emptyset \neq \mathcal{I} \subsetneq \{1, \dots, T\}})$ can be achieved by Lemma 41 when the $T + 1$ -th message S_{T+1} is used as the dummy message subject to the uniform distribution and its rate R_{T+1} is chosen to be $\max(I(V; Y|U) - \sum_{i=0}^T R_i - \frac{\varepsilon}{T}, 0)$. ■

Remark 39: As is mentioned in Proof of Theorem 37, to derive the capacity region, we employ Lemma 41, which is based on Code Construction 5 instead of Code Construction 4 because the case $\sum_{i=1}^T R_i > I(V; Y|U)$ requires Code Construction 5. This is the reason why we introduce Code Construction 5 as well as Code Construction 4. When $\sum_{i=1}^T R_i \leq I(V; Y|U)$, the rate tuple $((R_i)_{i=1, \dots, T}, (R_{l, \mathcal{I}})_{\emptyset \neq \mathcal{I} \subsetneq \{1, \dots, T\}})$ given in (126) can be approximately achieved by Lemma 40, which is based on Code Construction 4. That is, the rate tuple $((R_i - \frac{\varepsilon}{T})_{i=1, \dots, T}, (R_{l, \mathcal{I}})_{\emptyset \neq \mathcal{I} \subsetneq \{1, \dots, T\}})$ can be achieved by Lemma 40 when the $T + 1$ -th message S_{T+1} is used as the dummy message subject to the uniform distribution and its rate R_{T+1} is chosen to be $\max(I(V; Y|U) - \sum_{i=0}^T (R_i - \frac{\varepsilon}{T}) - \varepsilon, 0)$. Then, Code Construction 4 gives only the special rate tuple in the capacity region.

When there is no common message, it is enough to attain the region given in Corollary 38. Hence, it is sufficient to consider the case with $R_0 = 0$, which implies that $\sum_{i=1}^T R_i \leq I(V; Y|U)$. That is, if we need to show only Corollary 38, it is enough to use Lemma 40, which is based on Code Construction 4 instead of Code Construction 5.

Lemma 40: Choose a sufficiently small real number $\varepsilon > 0$

and $(R_i)_{i=0}^{T+1}$ for $i = 0, 1, \dots, T, T + 1$ satisfying

$$R_0 < \min[I(U; Y), I(U; Z)], \quad (128)$$

$$\sum_{i=1}^{T+1} R_i < I(V; Y|U) \leq \left(\sum_{i=1}^{T+1} R_i \right) + \varepsilon. \quad (129)$$

Then, the code φ_n given by Code Construction 4 satisfies

$$\lim_{n \rightarrow \infty} P_b[P_{Y|V}^n, \varphi_n, P_{S_{T,n}} \times P_{S_{T+1,n}}] = 0 \quad (130)$$

$$\lim_{n \rightarrow \infty} P_e[P_{Z|V}^n, \varphi_n, P_{S_{T,n}} \times P_{S_{T+1,n}}] = 0 \quad (131)$$

and

$$\begin{aligned} &\limsup_{n \rightarrow \infty} \frac{1}{n} I(S_{\mathcal{I}, n}; Z_n | S_{0, n}) [P_{Z|V}^n, \varphi_n, P_{S_{T,n}} \times P_{S_{T+1,n}}] \\ &\leq \sum_{i \in \mathcal{I}} R_i - [I(V; Y|U) - I(V; Z|U)]_+ + \varepsilon \end{aligned} \quad (132)$$

when the sequence of the joint distributions $P_{S_{T,n}}$ of information source satisfies the WACU condition (99) for any non-empty proper subset $\mathcal{I} \subsetneq \{1, \dots, T\}$ and $P_{S_{T+1,n}}$ is the uniform distribution.

Lemma 41: Choose a sufficiently small real number $\varepsilon > 0$ and $(R_i)_{i=0}^{T+1}$ for $i = 0, 1, \dots, T, T + 1$ satisfying

$$R_0 < \min[I(U; Y), I(U; Z)], \quad (133)$$

$$I(V; Y|U) \leq \left(\sum_{i=0}^{T+1} R_i \right) + \varepsilon < I(V; Y|U) + \min[I(U; Y), I(U; Z)]. \quad (134)$$

Then, the code φ_n given by Code Construction 5 with the choices

$$R_p := I(V; Y|U) - \varepsilon \text{ and } R_c := \sum_{i=0}^{T+1} R_i - R_p \quad (135)$$

satisfies (130), (131), and (132) when the sequence of the joint distributions $P_{S_{T,n}}$ of information source satisfies the WACU condition (99) for any non-empty proper subset $\mathcal{I} \subsetneq \{1, \dots, T\}$ and $P_{S_{T+1,n}}$ is the uniform distribution.

Proof of Lemma 40: Since the conditions (128) and (129) guarantee the conditions (121), we obtain (130) and (131). We need to show only (132). Assume that $I(V; Y|U) \leq I(V; Z|U)$. Since $|S_{\mathcal{I}, n}| = e^{n \sum_{i \in \mathcal{I}} R_i}$, we obtain $\frac{1}{n} I(S_{\mathcal{I}, n}; Z_n | S_{0, n}) [P_{Z|V}^n, \varphi_n, P_{S_{T,n}} \times P_{S_{T+1,n}}] \leq \sum_{i \in \mathcal{I}} R_i$, which implies (132). Hence, it is enough to consider the case $I(V; Y|U) > I(V; Z|U)$. Since, as is shown in Lemma 93 in Appendix C, the equivalence between the SWACU condition (100) and the WACU condition (99) holds, we obtain

$$\lim_{a \rightarrow \infty} \lim_{n \rightarrow \infty} \frac{1}{n} H_{1+a/n}(S_{\mathcal{I}, n} | S_{\mathcal{I}, n}, S_{0, n}) = \sum_{i \in \mathcal{I}} R_i. \quad (136)$$

The relations (117) and (136) yield

$$\begin{aligned}
& \limsup_{n \rightarrow \infty} \frac{1}{n} I(S_{\mathcal{I},n}; Z_n | S_{0,n}) [P_{Z|V}^n, \varphi_{a,n}, P_{S_{\mathcal{T},n}} \times P_{S_{\mathcal{T}+1,n}}] \\
& \leq I(V; Z|U) - \sum_{i \in \mathcal{I}^c} R_i \\
& = - \sum_{i=1}^{T+1} R_i + I(V; Z|U) + \sum_{i \in \mathcal{I}} R_i \\
& \leq \epsilon - I(V; Y|U) + I(V; Z|U) + \sum_{i \in \mathcal{I}} R_i, \tag{137}
\end{aligned}$$

which implies (132) \blacksquare

Proof of Lemma 41: Since the conditions (133), (134), and (135) guarantee the conditions (121), we obtain (130) and (131). We need to show only (132). When $I(V; Y|U) \leq I(V; Z|U)$, we can show (132) by the same way as Lemma 40. Hence, it is enough to consider the case $I(V; Y|U) > I(V; Z|U)$. By the same way as Lemma 40, the relations (122) and (136) yield

$$\begin{aligned}
& \limsup_{n \rightarrow \infty} \frac{1}{n} I(S_{\mathcal{I},n}; Z_n | S_{0,n}) [P_{Z|V}^n, \varphi_{a,n}, P_{S_{\mathcal{T},n}} \times P_{S_{\mathcal{T}+1,n}}] \\
& \leq (R_c - R_0) + I(V; Z|U) - \sum_{i \in \mathcal{I}^c} R_i \\
& = R_c - \sum_{i=0}^{T+1} R_i + I(V; Z|U) + \sum_{i \in \mathcal{I}} R_i \\
& = -R_p + I(V; Z|U) + \sum_{i \in \mathcal{I}} R_i. \tag{138}
\end{aligned}$$

Therefore, since $R_p = I(V; Y|U) - \epsilon$, (138) implies (132) when $I(V; Y|U) > I(V; Z|U)$. \blacksquare

X. SECURE MULTIPLEX CODING WITH COMMON MESSAGES: STRONG SECURITY

A. Strong Security

In this section, we treat the strong security. A sequence of codes φ_n is called *strongly secure* for a subset $\mathcal{I} \subseteq \{1, \dots, T\}$ and a sequence of distributions $P_{S_{\mathcal{T},n}}$ when the relation

$$\lim_{n \rightarrow \infty} I(S_{\mathcal{I},n}; Z_n | S_{0,n}) [P_{Z|X}^n, \varphi_n, P_{S_{\mathcal{T},n}}] = 0 \tag{139}$$

holds. Now, we fix a family \mathbf{J} of non-empty proper subsets \mathcal{I} of $\{1, \dots, T\}$, and consider only the security of the messages $S_{\mathcal{I},n}$ for all $\mathcal{I} \in \mathbf{J}$.

Theorem 42: Assume that the transmission rate tuple $(R_i)_{i=0, \dots, T} = (R_0, R_1, \dots, R_T)$ belongs to the inner of the capacity region with $R_{\mathcal{I}, \mathcal{I}} = 0$ for any subset $\mathcal{I} \in \mathbf{J}$, i.e., there exist an information leakage rate tuple $(R_{\mathcal{I}, \mathcal{I}})_{\emptyset \neq \mathcal{I} \in \mathbf{J}}$ such that

$$((R_i)_{i=0, \dots, T}, (0)_{\mathcal{I} \in \mathbf{J}}, (R_{\mathcal{I}, \mathcal{I}})_{\emptyset \neq \mathcal{I} \in \mathbf{J}}) \in \text{inn}(C), \tag{140}$$

where $\text{inn}(C)$ denotes the inner of the set C . Then, there exists a Markov chain $U \rightarrow V \rightarrow X$ such that

$$\epsilon := \min_{\mathcal{I} \in \mathbf{J}} \frac{I(V; Y|U) - I(V; Z|U) - \sum_{i \in \mathcal{I}} R_i}{|\mathcal{I}^c|} > 0, \tag{141}$$

$$R_0 < \min[I(U; Y), I(U; Z)],$$

$$\sum_{i=0}^T R_i < I(V; Y|U) + \min[I(U; Y), I(U; Z)].$$

Next, we choose $R_{T+1} := \max(I(V; Y|U) - \sum_{i=0}^T R_i, 0)$ and a small real $\epsilon' > 0$ such that $\epsilon' < \frac{\epsilon}{2}$, $\epsilon' < I(V; Y|U) + \min[I(U; Y), I(U; Z)] - \sum_{i=0}^{T+1} R_i$. The code φ_n given by Code Construction 5 with the choices $R_p := I(V; Y|U) - \epsilon'$ and $R_c := \sum_{i=0}^{T+1} R_i - R_p$ satisfies (130), (131), and the strong security

$$\lim_{n \rightarrow \infty} I(S_{\mathcal{I},n}; Z_n | S_{0,n}) [P_{Z|V}^n, \varphi_n, P_{S_{\mathcal{T},n}}] = 0 \tag{142}$$

for any subset $\mathcal{I} \in \mathbf{J}$ when the sequence of distributions $P_{S_{\mathcal{T},n}}$ satisfies the $(\epsilon - 2\epsilon')$ -SACU condition (101) for the subset \mathcal{I} .

Thanks to Theorem 42, the strong security holds at all inner points of the capacity region C with $R_{\mathcal{I}, \mathcal{I}} = 0$ for any subset $\mathcal{I} \in \mathbf{J}$ under the ϵ -SACU condition (101) for any subset $\mathcal{I} \in \mathbf{J}$.

Here, we address the relation with the paper [22]. When there is no common message, the paper [22] defined the region $\mathcal{R}_{\text{sto}}^I$ as follows.

Definition 43: The region $\mathcal{R}_{\text{sto}}^I$ is the closure of the set of the rate tuples $(R_i)_{i=1, \dots, T}$ satisfying the following. There exist a sequence of codes $\varphi_n = (\varphi_{a,n}, \varphi_{b,n}, \varphi_{e,n})$, i.e., Alice's stochastic encoder $\varphi_{a,n}$ from $\mathcal{S}_{1,n} \times \dots \times \mathcal{S}_{T,n}$ to \mathcal{X}^n , Bob's deterministic decoder $\varphi_{b,n} : \mathcal{Y}^n \rightarrow \mathcal{S}_{1,n} \times \mathcal{S}_{1,n} \times \dots \times \mathcal{S}_{T,n}$ satisfying the following conditions: (1) The i -th secret message set $\mathcal{S}_{i,n}$ has cardinality e^{nR_i} for $i = 1, \dots, T$, (2) When the message obeys the uniform distribution, the relations (123) and

$$\limsup_{n \rightarrow \infty} I(S_{t,n}; Z^n | S_0) [P_{Z|X}^n, \varphi_{a,n}, P_{S_{\mathcal{T},n}} \times P_{S_{\mathcal{T}+1,n}}] = 0 \tag{143}$$

hold for $t = 1, \dots, T$.

On the other hand, we define the region $\tilde{\mathcal{R}}_{\text{sto}}^I$ as the set of rate tuples $(R_i)_{i=1, \dots, T}$ such that there exists a Markov chain $V \rightarrow X \rightarrow YZ$ and

$$\sum_{i=1}^T R_i \leq I(V; Y), \quad R_t \leq [I(V; Y) - I(V; Z)]_+ \tag{144}$$

for $t = 1, \dots, T$. Then, Theorem 42 and Corollary 38 guarantee the relation

$$\mathcal{R}_{\text{sto}}^I = \tilde{\mathcal{R}}_{\text{sto}}^I, \tag{145}$$

which is the same as the result by the paper [22, (138)]. Here, Corollary 38 implies $\mathcal{R}_{\text{sto}}^I \subset \tilde{\mathcal{R}}_{\text{sto}}^I$ and Theorem 42 does $\mathcal{R}_{\text{sto}}^I \supset \text{inn}(\tilde{\mathcal{R}}_{\text{sto}}^I)$. Since $\mathcal{R}_{\text{sto}}^I$ and $\tilde{\mathcal{R}}_{\text{sto}}^I$ are the closed sets, we obtain (145).

In order to show Theorem 42, we prepare the following lemma.

Lemma 44: We fix a subset $\mathcal{I} \subseteq \{1, \dots, T\}$. Assume that the transmission rate tuple $(R_i)_{i=0, \dots, T}$, the sequence of distributions $P_{S_{\mathcal{T},n}}$, and a Markov chain $U \rightarrow V \rightarrow X$ satisfy that

$$\begin{aligned}
\delta' & := \frac{1}{2} \left(\underline{H}_{\log}(\mathcal{I}^c) \right. \\
& \quad \left. - \left(\sum_{i=1}^T R_i - I(V; Y|U) + I(V; Z|U) \right) \right) > 0, \tag{146}
\end{aligned}$$

$$R_0 < \min[I(U; Y), I(U; Z)],$$

$$\sum_{i=0}^T R_i < I(V; Y|U) + \min[I(U; Y), I(U; Z)].$$

When we choose $R_{T+1} := \max(I(V; Y|U) - \sum_{i=0}^T R_i, 0)$ and a small real $\epsilon' > 0$ such that $\epsilon' \leq \delta'$ and $\epsilon' < I(V; Y|U) +$

$\min[I(U; Y), I(U; Z)] - \sum_{i=0}^{T+1} R_i$, the code φ_n given by Code Construction 5 with the choices $R_p := I(V; Y|U) - \epsilon'$ and $R_c := \sum_{i=0}^{T+1} R_i - R_p$ satisfies (130), (131), and the strong security

$$\lim_{n \rightarrow \infty} I(S_{\mathcal{I},n}; Z_n | S_{0,n}) [P_{Z|V}^n, \varphi_n, P_{S_{\mathcal{T},n}} \times P_{S_{T+1,n}}] = 0. \quad (147)$$

Proof of Theorem 42: First, we fix an arbitrary subset $\mathcal{I} \in \mathbf{J}$. Hence,

$$\begin{aligned} & \sum_{i \in \mathcal{I}^c} (R_i - (\epsilon - 2\epsilon')) - \left(\sum_{i=1}^{T+1} R_i - I(V; Y|U) + I(V; Z|U) \right) \\ & \geq \left(\sum_{i \in \mathcal{I}^c} R_i \right) - |\mathcal{I}^c|(\epsilon - 2\epsilon') - \left(\sum_{i=1}^{T+1} R_i - I(V; Y|U) + I(V; Z|U) \right) \\ & = I(V; Y|U) - I(V; Z|U) - \sum_{i \in \mathcal{I}} R_i - |\mathcal{I}^c|(\epsilon - 2\epsilon') \\ & \geq |\mathcal{I}^c|\epsilon - |\mathcal{I}^c|(\epsilon - 2\epsilon') = 2|\mathcal{I}^c|\epsilon' \geq 2\epsilon'. \end{aligned}$$

Thus, since the sequence of distributions $P_{S_{\mathcal{T},n}}$ satisfies the $\epsilon - 2\epsilon'$ -SACU condition (101) for the subset \mathcal{I} ,

$$\begin{aligned} \delta' & := \frac{1}{2} \left(\underline{H}_{\log}(\mathcal{I}^c) \right. \\ & \quad \left. - \left(\sum_{i=1}^{T+1} R_i - I(V; Y|U) + I(V; Z|U) \right) \right) \\ & \geq \frac{1}{2} \left(\sum_{i \in \mathcal{I}^c} (R_i - (\epsilon - 2\epsilon')) - \left(\sum_{i=1}^{T+1} R_i - I(V; Y|U) + I(V; Z|U) \right) \right) \\ & \geq \epsilon'. \end{aligned}$$

Hence, any real number $\epsilon' > 0$ given in Theorem 42 satisfies the condition for $\epsilon' > 0$ in Lemma 44. Thus, applying Lemma 44, we obtain (142) for the subset \mathcal{I} . Since the subset \mathcal{I} is an arbitrary element of \mathbf{J} , we obtain Theorem 42. \blacksquare

Proof of Lemma 44: Since $\epsilon' > 0$, we have the second condition of (121). Due to the choice of $\epsilon' > 0$,

$$\begin{aligned} 0 & = I(V; Y|U) - \epsilon' - R_p \\ & > I(V; Y|U) - \left(I(V; Y|U) + \min[I(U; Y), I(U; Z)] - \sum_{i=0}^{T+1} R_i \right) \\ & \quad - R_p \\ & = \sum_{i=0}^{T+1} R_i - \min[I(U; Y), I(U; Z)] - R_p \\ & = R_c - \min[I(U; Y), I(U; Z)], \end{aligned}$$

which implies the first condition of (121). Hence, we obtain (130) and (131).

Next, we define

$$\begin{aligned} \rho_n & := \frac{2 \log n}{n \delta'}, \\ C_n & := \left(-\rho_n n (R_c - R_0) + \rho_n H_{1+\rho_n}(S_{\mathcal{I}^c,n} | S_{\mathcal{I},n}, S_{0,n}) \right. \\ & \quad \left. - n E_0(\rho_n | P_{Z|V}, P_{V|U}, P_U) \right). \end{aligned}$$

The condition (146) and $\epsilon' \leq \delta'$ imply that

$$\begin{aligned} & \liminf_{n \rightarrow \infty} \frac{C_n}{n \rho_n} \\ & = \liminf_{n \rightarrow \infty} \frac{1}{n} H_{1+\rho_n}(S_{\mathcal{I}^c,n} | S_{\mathcal{I},n}, S_{0,n}) - \sum_{i=1}^{T+1} R_i + R_p - I(V; Z|U) \\ & \geq \underline{H}_{\log}(\mathcal{I}^c) - \sum_{i=1}^{T+1} R_i + I(V; Y|U) - \delta' - I(V; Z|U) \\ & = \frac{1}{2} \left(\underline{H}_{\log}(\mathcal{I}^c) - \sum_{i=1}^{T+1} R_i + I(V; Y|U) - I(V; Z|U) \right) \\ & = \delta' > 0. \end{aligned} \quad (148)$$

That is, we can choose a sufficiently large integer N such that

$$\frac{C_n}{n \rho_n} \geq \frac{\delta'}{2} \quad (149)$$

for $n \geq N$. Due to (77), the leaked information for $S_{\mathcal{I},n}$ can be evaluated as

$$I(S_{\mathcal{I},n}; Z_n | S_{0,n}) [P_{Z|V}^n, \varphi_n, P_{S_{\mathcal{T},n}}] \leq \frac{2^{T+2}}{\rho_n} e^{-C_n}.$$

Since (149) implies that

$$\begin{aligned} & -\log\left(\frac{2^{T+2}}{\rho_n} e^{-C_n}\right) = -(T+2) \log 2 + C_n + \log \rho_n \\ & \geq -(T+2) \log 2 + \frac{\delta'}{2} n \rho_n + \log \rho_n \\ & = -(T+2) \log 2 + \log \log n - \log \frac{\delta'}{2} \rightarrow \infty, \end{aligned}$$

we obtain (147). \blacksquare

B. Exponential Decreasing Rate

In this subsection, we treat the exponential decreasing rate of leaked information. In this subsection, we assume that the $T+1$ -th message $S_{T+1,n}$ is subject to the uniform distribution. We simplify $P_{S_{\mathcal{T},n}} \times P_{S_{T+1,n}}$ by $P_{S_{\mathcal{T},n}}$. For a subset $\mathcal{I} \subseteq \{1, \dots, T\}$, we denote the complementary set in $\{1, \dots, T\}$ by \mathcal{I}^c and simplify the set $\mathcal{I}^c \cup \{T+1\}$ to $\mathcal{I}^{c,*}$. Unfortunately, the ϵ -SACU condition (101) is not sufficient for deriving a good exponential decreasing rate of leaked information. Hence, in this subsection, given a sequence of distributions $P_{S_{\mathcal{T},n}}$, we introduce the following quantity

$$\underline{H}_{1+\rho}(\mathcal{I}^{c,*}) := \liminf_{n \rightarrow \infty} \frac{1}{n} H_{1+\rho}(S_{\mathcal{I}^{c,*},n} | S_{\mathcal{I},n}, S_{0,n}) \quad (150)$$

for any subset $\mathcal{I} \subseteq \{1, \dots, T\}$ and any $\rho \in (0, 1]$.

Theorem 45: For given $(R_i)_{i=0}^T$, we choose R_p and R_c as follows.

$$R_c \geq R_0, \quad R_c + R_p = \sum_{i=0}^{T+1} R_i.$$

We fix a real number $\epsilon > 0$. We choose a code φ_n given by Code Construction 5 with the above choices R_p and R_c and a given Markov chain $U \rightarrow V \rightarrow X$. When the sequence of distributions $P_{S_{\mathcal{T},n}}$ satisfies the ϵ -SACU condition (101) for a

non-empty proper subset $\mathcal{I} (\neq \emptyset) \subseteq \{1, \dots, T\}$, the sequence of codes φ_n satisfies (119), (120), and

$$\begin{aligned} & \liminf_{n \rightarrow \infty} \frac{-1}{n} \log I(S_{\mathcal{I},n}; Z_n | S_{0,n}) [P_{Z|V}^n, \Phi_n, P_{S_{\mathcal{I},n}}] \\ & \geq \sup_{0 < \rho < 1} \rho (\underline{H}_{1+\rho}(\mathcal{I}^{c,*}) - R_c + R_0) - E_0(\rho | P_{Z|V}, P_{V|U}, P_U). \end{aligned} \quad (151)$$

In particular, when the distribution $P_{S_{\mathcal{I},n}}$ is uniform, we obtain

$$\begin{aligned} & \liminf_{n \rightarrow \infty} \frac{-1}{n} \log I(S_{\mathcal{I},n}; Z_n | S_{0,n}) [P_{Z|V}^n, \Phi_n, P_{S_{\mathcal{I},n}}] \\ & \geq \tilde{E}^{E_0}(R_p - \sum_{i \in \mathcal{I}} R_i, P_{Z,V,U}), \end{aligned} \quad (152)$$

where $\tilde{E}^{E_0}(R, P_{Z,V,U})$ is defined in (22).

Theorem 45 yields the following observation. When $R_p - \epsilon - \sum_{i \in \mathcal{I}} R_i > I(V; Z|U)$ and $\underline{H}_{1+\rho}(\mathcal{I}^c) \geq (\sum_{i \in \mathcal{I}^c} R_i) - \epsilon$ holds with a small $\rho > 0$, the exponent (151) is positive, i.e., the leaked information goes to zero exponentially. In particular, when

$$\sum_{i=1}^{T+1} R_i < I(V; Y|U), \quad R_0 < \min[I(U; Y), I(U; Z)], \quad (153)$$

we can choose R_p and R_c by

$$R_p := \sum_{i=1}^{T+1} R_i, \quad R_c := R_0. \quad (154)$$

Then, the inequalities (119) and (120) can be simplified to (115) and (116). Then, the both decoding error probabilities goes zero exponentially. Further, the inequality (151) can be simplified to

$$\begin{aligned} & \liminf_{n \rightarrow \infty} \frac{-1}{n} \log I(S_{\mathcal{I},n}; Z_n | S_{0,n}) [P_{Z|V}^n, \Phi_n, P_{S_{\mathcal{I},n}}] \\ & \geq \sup_{0 < \rho < 1} \rho \underline{H}_{1+\rho}(\mathcal{I}^{c,*}) - E_0(\rho | P_{Z|V}, P_{V|U}, P_U). \end{aligned} \quad (155)$$

Further, in the case of (153) and (154), when the WACU condition holds for \mathcal{I} , the inequality (122) can be simplified to

$$\begin{aligned} & \limsup_{n \rightarrow \infty} \frac{1}{n} I(S_{\mathcal{I},n}; Z_n | S_{0,n}) [P_{Z|V}^n, \Phi_n, P_{S_{\mathcal{I},n}}] \\ & \leq R_c - R_0 + I(V; Z|U) - \sum_{i \in \mathcal{I}^{c,*}} R_i = I(V; Z|U) - \sum_{i \in \mathcal{I}^{c,*}} R_i. \end{aligned} \quad (156)$$

Proof of Theorem 45: In Subsection IX-A, we have already shown (119) and (120). Hence, we need to only show (151). Due to (77), the leaked information for $S_{\mathcal{I},n}$ can be evaluated as

$$\begin{aligned} & I(S_{\mathcal{I},n}; Z_n | S_{0,n}) [P_{Z|V}^n, \varphi_n, P_{S_{\mathcal{I},n}}] \\ & \leq \frac{2^{T+2}}{\rho} e^{\rho n(R_c - R_0) - \rho H_{1+\rho}(S_{\mathcal{I}^{c,*},n} | S_{\mathcal{I},n}, S_{0,n}) + n E_0(\rho | P_{Z|V}, P_{V|U}, P_U)}. \end{aligned}$$

Hence,

$$\begin{aligned} & \liminf_{n \rightarrow \infty} \frac{-1}{n} \log I(S_{\mathcal{I},n}; Z_n | S_{0,n}) \\ & \geq \rho \liminf_{n \rightarrow \infty} \frac{1}{n} H_{1+\rho}(S_{\mathcal{I}^{c,*},n} | S_{\mathcal{I},n}, S_{0,n}) \\ & \quad - \rho(R_c - R_0) - E_0(\rho | P_{Z|V}, P_{V|U}, P_U) \\ & \geq \rho (\underline{H}_{1+\rho}(\mathcal{I}^{c,*}) - R_c + R_0) - E_0(\rho | P_{Z|V}, P_{V|U}, P_U). \end{aligned}$$

Taking the supremum for $\rho \in [0, 1]$, we obtain (151). \blacksquare

When the condition (153) holds, the exponent (155) can be improved by using Theorem 20 with Code Construction 4 in the following way.

Theorem 46: We fix a real number $\epsilon \geq 0$. Let φ_n be a code given in Code Construction 4 in Subsection IX-A. The sequence of codes φ_n satisfies (115), (116), (156), and

$$\begin{aligned} & \liminf_{n \rightarrow \infty} \frac{-1}{n} \log I(S_{\mathcal{I},n}; Z_n | S_{0,n}) [P_{Z|V}^n, \Phi_n, P_{S_{\mathcal{I},n}}] \\ & \geq \max_{0 \leq \rho \leq 1} \rho \underline{H}_{1+\rho}(\mathcal{I}^{c,*}) - \psi(\rho | P_{Z|V}, P_{V|U}, P_U). \end{aligned} \quad (157)$$

In particular, when the distribution $P_{S_{\mathcal{I},n}}$ is uniform, we obtain

$$\begin{aligned} & \liminf_{n \rightarrow \infty} \frac{-1}{n} \log I(S_{\mathcal{I},n}; Z_n | S_{0,n}) [P_{Z|V}^n, \Phi_n, P_{S_{\mathcal{I},n}}] \\ & \geq \tilde{E}^\psi(\sum_{i \in \mathcal{I}^{c,*}} R_i, P_{Z,V,U}), \end{aligned}$$

where $\tilde{E}^\psi(R, P_{Z,V,U})$ is defined in (21).

Now, we compare Theorems 45 and 46. Since the RHS of (157) is larger than the RHS of (155) due to (17), Theorem 46 is better than Theorem 45 when the relation (153) holds. Otherwise, the error exponent of (115) and/or (116) is not positive. That is, Theorem 46 cannot yield a reliable communication. In summary, Theorem 45 has a wider applicability than Theorem 46. In the special case (153), Theorem 46 is better than Theorem 45.

Proof: Relations (115) and (116) have been shown in Subsection IX-A. Due to the ϵ -SACU condition, (117) guarantees (156). Using (63) and the ϵ -SACU condition, we obtain

$$\begin{aligned} & I(S_{\mathcal{I},n}; Z_n | S_{0,n}) [P_{Z|V}^n, \Phi_n, P_{S_{\mathcal{I},n}}] \\ & \leq \frac{2^{T+2}}{\rho} e^{-\rho H_{1+\rho}(S_{\mathcal{I}^{c,*},n} | S_{\mathcal{I},n}, S_{0,n}) + n \psi(\rho | P_{Z|V}, P_{V|U}, P_U)}. \end{aligned}$$

Then,

$$\begin{aligned} & \liminf_{n \rightarrow \infty} \frac{-1}{n} \log I(S_{\mathcal{I},n}; Z_n | S_{0,n}) [P_{Z|V}^n, \Phi_n, P_{S_{\mathcal{I},n}}] \\ & \geq \rho \underline{H}_{1+\rho}(\mathcal{I}^{c,*}) - \psi(\rho | P_{Z|V}, P_{V|U}, P_U). \end{aligned} \quad (158)$$

Hence, we obtain (157). \blacksquare

When the above discussion is applied to the wire-tap channel model, we obtain an extension of existing results to the case of the asymptotic uniform dummy message. That is, we consider the case with no common messages and $T = 2$ when S_1 corresponds to the message to be secretly sent to Bob, and S_2 does to the dummy message making S_1 ambiguous to Eve. For a given rate R_1 of secret message and a given rate R_2 of dummy message, the RHS of (115) coincides with the Gallager exponents, the RHS of (155) coincides with the RHS of (59) in [15], and the RHS of (157) coincides with the exponents of the RHS of (15) in [17].

XI. PRACTICAL CODE CONSTRUCTION

In Section XI, we consider how we can construct practically usable encoder and decoder for the secure multiplex coding. When the channel has additive structure, the paper [17, Section V] constructed a code for wire-tap channel code from an ordinary linear error correcting code, and the paper [22, Section

VI] did a secure multiple code without common message from an ordinary linear error correcting code. Here, we construct a secure multiple code with/without common message when the channel does not necessarily have additive structure and the message does not necessarily obey the uniform distribution. We shall show how to convert an ordinary error correcting code without secrecy consideration to a code for the secure multiplex coding. In this section, we treat practical code construction in the single-shot setting unless otherwise stated.

It is a common practice to assume the uniform distribution of messages when one evaluates the decoding error probability, and decoding error probabilities with non-uniform message distributions are rarely considered in practice. Thus, we always assume the uniform message distribution because this assumption is necessary for the analysis of the decoding error probability. However, this assumption is unnecessary for that of the leaked information to Eve. The analysis of this section holds for general channels with finite alphabets except for Lemma 50. Only Lemma 50 assumes the regularity of the channel.

A. First Practical Code Construction: First Type Evaluation

We construct a code for the secure multiplex coding based on a given code φ_p for BCD with the common message in \mathcal{S}_c and the private message in \mathcal{S}_p . We assume that encoding and decoding of φ_p can be efficiently executed. We shall attach F' and G' in the second step of Code Ensemble 3 to φ_p so that the resulting code for SMC enables efficient encoding and decoding. This type of construction is much more practical than Code Ensemble 3 because Code Ensemble 3 uses the random coding for the error correcting code φ_p , which does not enable efficient encoding nor decoding. To use the code with F' and G' attached, we have to evaluate decoding error probability and the amount of information leaked to Eve. The former is less than or equal to that of the underlying error correcting code φ_p , and the average of the latter over the ensemble of F' and G' can be evaluated by Lemma 21 with a fixed error correcting code φ_p . In our code, we employ a dummy message to realize the secrecy of message when the leaked information is very close to the mutual information with the normal receiver and the number of T is fixed. Now, we present a code construction.

Code Construction 6: First, in order to apply Lemma 21, we divide the common message set \mathcal{S}_c of the BCD code φ_p to $\mathcal{S}_0 \times \mathcal{B}_1$, and denote the private message set \mathcal{S}_p of φ_p by \mathcal{B}_2 . That is, the code φ_p is regarded as a map from $\mathcal{S}_0 \times \mathcal{B}_1 \times \mathcal{B}_2$ to \mathcal{X} . Then, based on the code φ_p , assuming the Abelian group structures in \mathcal{B}_1 and \mathcal{B}_2 , we choose an ensemble of isomorphisms⁶ F' from $\mathcal{S}_1 \times \dots \times \mathcal{S}_{T+1}$ to $\mathcal{B}_1 \times \mathcal{B}_2$ as Abelian groups satisfying Condition 15 while we do not assume any algebraic assumption for the code φ_p . In this scenario, \mathcal{S}_0 is common message, $\mathcal{S}_1, \dots, \mathcal{S}_T$ are secret messages, and \mathcal{S}_{T+1} is the dummy randomness whose secrecy is not required. We choose the random variable $G' \in \mathcal{B}_1 \times \mathcal{B}_2$ that obeys the uniform distribution on $\mathcal{B}_1 \times \mathcal{B}_2$ and

⁶Remark 16 discusses an efficient realization of an ensemble of isomorphisms F satisfying Condition 15.

is independent of the choice of F' and anything else. Then, by defining a map $\Lambda_{F',G'}(s) := F'(s) + G'$, we obtain our encoder $\varphi_p \circ \Lambda_{F',G'}(s_0, s_1, \dots, s_{T+1}) = \varphi_p(s_0, \Lambda_{F',G'}(s_1, \dots, s_{T+1}))$. The decoder is constructed by applying the inverse $\Lambda_{F',G'}^{-1}(b_1, b_2) = F'^{-1}((b_1, b_2) - G')$ to the decoded message of the code φ_p .

The average of the leaked information of the above constructed code is evaluated as follows.

Lemma 47: For a subset $\mathcal{I} \subseteq \{1, \dots, T\}$, the quantity $E_{0,\max}(\rho|P_{Z|V})$ defined in (23) satisfies

$$\begin{aligned} & \mathbf{E}_{F',G'} I(\mathcal{S}_{\mathcal{I}}; Z|S_0)[P_{Z|V}, \varphi_p \circ \Lambda_{F',G'}, P_{\mathcal{S}_{\mathcal{T}}}] \\ & \leq \frac{e^{E_{0,\max}(\rho|P_{Z|V}) - \rho H_{1+\rho}(\mathcal{S}_{\mathcal{I}^c} | \mathcal{S}_{\mathcal{I}}, S_0)}}{\rho}. \end{aligned} \quad (159)$$

Proof: Applying Lemma 21, we obtain

$$\begin{aligned} & \mathbf{E}_{F',G'} \exp(\rho I(\mathcal{S}_{\mathcal{I}}; Z|S_0)[P_{Z|V}, \varphi_p \circ \Lambda_{F',G'}, P_{\mathcal{S}_{\mathcal{T}}}]) \\ & \leq 1 + \sum_{s_0} P_{S_0}(s_0) \sum_{\mathcal{S}_{\mathcal{I}}} P_{\mathcal{S}_{\mathcal{I}}|S_0}(\mathcal{S}_{\mathcal{I}}|s_0) e^{-\rho H_{1+\rho}(\mathcal{S}_{\mathcal{I}^c} | \mathcal{S}_{\mathcal{I}}, S_0 = s_0)} \\ & \quad \cdot e^{\psi(\rho|P_{Z|B_1, B_2, S_0 = s_0, \varphi_p}, P_{\text{mix}, B_1, B_2})}. \end{aligned} \quad (160)$$

Since

$$\begin{aligned} & e^{\psi(\rho|P_{Z|B_1, B_2, \varphi_p, S_0}, P_{\text{mix}, B_1, B_2})} \leq e^{E_0(\rho|P_{Z|B_1, B_2, \varphi_p, S_0}, P_{\text{mix}, B_1, B_2})} \\ & = \sum_z \left(\sum_{b_1, b_2} \frac{1}{|\mathcal{B}_1| |\mathcal{B}_2|} P_{Z|V}(z|\varphi_p(s_0, b_1, b_2)) \right)^{\frac{1}{1-\rho}} \\ & \quad \sum_{\mathcal{S}_{\mathcal{I}}} P_{\mathcal{S}_{\mathcal{I}}|S_0}(\mathcal{S}_{\mathcal{I}}|s_0) e^{-\rho H_{1+\rho}(\mathcal{S}_{\mathcal{I}^c} | \mathcal{S}_{\mathcal{I}}, S_0 = s_0)} = e^{-\rho H_{1+\rho}(\mathcal{S}_{\mathcal{I}^c} | \mathcal{S}_{\mathcal{I}}, S_0 = s_0)}, \end{aligned}$$

we obtain

$$\begin{aligned} & \mathbf{E}_{F',G'} \exp(\rho I(\mathcal{S}_{\mathcal{I}}; Z|S_0)[P_{Z|V}, \varphi_p \circ \Lambda_{F',G'}, P_{\mathcal{S}_{\mathcal{T}}}]) \\ & \leq 1 + \sum_{s_0} P_{S_0}(s_0) e^{-\rho H_{1+\rho}(\mathcal{S}_{\mathcal{I}^c} | \mathcal{S}_{\mathcal{I}}, S_0 = s_0)} \\ & \quad \cdot \sum_z \left(\sum_{b_1, b_2} \frac{1}{|\mathcal{B}_1| |\mathcal{B}_2|} P_{Z|V}(z|\varphi_p(s_0, b_1, b_2)) \right)^{\frac{1}{1-\rho}}. \end{aligned} \quad (161)$$

It can be simplified as follows.

$$\begin{aligned} & \sum_z \left(\sum_{b_1, b_2} \frac{1}{|\mathcal{B}_1| |\mathcal{B}_2|} P_{Z|V}(z|\varphi_p(s_0, b_1, b_2)) \right)^{\frac{1}{1-\rho}} \\ & \leq \max_{P_V} \sum_z \left(\sum_v P_V(v) P_{Z|V}(z|v) \right)^{\frac{1}{1-\rho}} \\ & = \max_{P_V} e^{E_0(\rho|P_{Z|V}, P_V)} = e^{E_{0,\max}(\rho|P_{Z|V})}. \end{aligned}$$

That is, using the relation $\sum_{s_0} P_{S_0}(s_0) e^{-\rho H_{1+\rho}(\mathcal{S}_{\mathcal{I}^c} | \mathcal{S}_{\mathcal{I}}, S_0 = s_0)} = e^{-\rho H_{1+\rho}(\mathcal{S}_{\mathcal{I}^c} | \mathcal{S}_{\mathcal{I}}, S_0)}$, we have

$$\begin{aligned} & \mathbf{E}_{F',G'} \exp(\rho I(\mathcal{S}_{\mathcal{I}}; Z|S_0)[P_{Z|V}, \varphi_p \circ \Lambda_{F',G'}, P_{\mathcal{S}_{\mathcal{T}}}]) \\ & \leq 1 + e^{-\rho H_{1+\rho}(\mathcal{S}_{\mathcal{I}^c} | \mathcal{S}_{\mathcal{I}}, S_0)} e^{E_{0,\max}(\rho|P_{Z|V})}. \end{aligned} \quad (162)$$

Combining the Jensen inequality for $x \mapsto e^x$, we obtain the desired upper bound (159). ■

The logarithm of the RHS of (159) has the following property.

Lemma 48: The functions $\rho \mapsto E_0(\rho|P_{Z|V}) - \rho H_{1+\rho}(\mathcal{S}_{\mathcal{I}^c} | \mathcal{S}_{\mathcal{I}}, S_0) - \log \rho$ and $\rho \mapsto E_{0,\max}(\rho|P_{Z|V}) - \rho H_{1+\rho}(\mathcal{S}_{\mathcal{I}^c} | \mathcal{S}_{\mathcal{I}}, S_0) - \log \rho$ are convex.

Proof: The function $\rho \mapsto E_0(\rho|\overline{W}^Z, Q_V)$ is convex [12]. Also the function $\rho \mapsto \rho H_{1+\rho}(\mathcal{S}_{\mathcal{I}^c} | \mathcal{S}_{\mathcal{I}}, S_0)$ is concave. Hence,

$E_0(\rho|P_{Z|V}, Q_V) - \rho H_{1+\rho}(S_{I^c}|S_I, S_0) - \log \rho$ is convex. Similarly, due to Lemma 5, the function $\rho \mapsto E_{0,\max}(\rho|P_{Z|V}) - \rho H_{1+\rho}(S_{I^c}|S_I, S_0) - \log \rho$ is convex. ■

As is explained latter, the bound $e^{E_{0,\max}(\rho|P_{Z|V})}$ is computable in the discrete memoryless case. On the other hand, the error probabilities can be upper bounded by the average error probabilities of the code φ_p .

Next, we determine the necessary amount of dummy randomness so that the amounts of leaked information is below specified levels. Suppose that we are given arbitrary error-correcting code φ_p for the broadcast channel $P_{Y|Z|V}$. The code φ_p can be, for example, an LDPC code [40] or a Turbo code [41] when there is no common message. Then, we assume that S_{T+1} obeys the uniform distribution on its alphabet \mathcal{S}_{T+1} and is statistically independent of all other random variables. As a corollary to Lemma 47, we have:

Lemma 49: For $\mathcal{I} \subset \{1, \dots, T\}$, we have

$$\frac{\mathbf{E}_{F',G'} I(S_{\mathcal{I}}; Z|S_0)[P_{Z|V}, \varphi_p \circ \Lambda_{F',G'}, P_{S_{\mathcal{T}}}] e^{E_{0,\max}(\rho|P_{Z|V}) - \rho(\log|\mathcal{S}_{T+1}| + H_{1+\rho}(S_{I^c}|S_I, S_0))}}{\rho} \leq \epsilon_{\mathcal{I}}. \quad (163)$$

By using Eq. (163), from φ_p we can construct a code for the secure multiplex coding as follows. For each proper nonempty set $\mathcal{I} \subseteq \{1, \dots, T\}$, $\epsilon_{\mathcal{I}}$ denotes the maximum acceptable information leakage for $I(S_{\mathcal{I}}; Z)$. Denote by ϵ_2 the maximum acceptable probability for a chosen F', G' not making $I(S_{\mathcal{I}}; Z|S_0)$ below $\epsilon_{\mathcal{I}}$ for some \mathcal{I} .

Adjust the size $|\mathcal{S}_{T+1}|$ of the dummy randomness so that

$$\epsilon_{\mathcal{I}} := \frac{2^T}{\epsilon_2} \left(\inf_{\rho \in (0,1)} \frac{e^{E_{0,\max}(\rho|P_{Z|V}) - \rho(\log|\mathcal{S}_{T+1}| + H_{1+\rho}(S_{I^c}|S_I, S_0))}}{\rho} \right).$$

Then, due to (163), we obtain

$$\mathbf{E}_{F',G'} I(S_{\mathcal{I}}; Z|S_0)[P_{Z|V}, \varphi_p \circ \Lambda_{F',G'}, P_{S_{\mathcal{T}}}] \leq \epsilon_2 \epsilon_{\mathcal{I}} / 2^T$$

Then, by the Markov inequality the probability of choosing F' and G' making $I(S_{\mathcal{I}}; Z|S_0) \leq \epsilon_{\mathcal{I}}$ simultaneously for all $\mathcal{I} \subseteq \{1, \dots, T\}$ is $\geq 1 - \epsilon_2$.

When the channel is a regular channel in the sense of Delsarte-Piret [10], the value $E_{0,\max}(\rho|P_{Z|V})$ can be calculated as follows:

Lemma 50: When the channel $P_{Z|V}$ is regular in the sense of Delsarte-Piret [10],

$$E_{0,\max}(\rho|P_{Z|V}) = E_0(\rho|P_{Z|V}, P_{\text{mix},\mathcal{V}}). \quad (164)$$

Further, when the code φ_p is a homomorphism as Abelian group, the inequality

$$\frac{\mathbf{E}_{F',G'} I(S_{\mathcal{I}}; Z|S_0)[P_{Z|V}, \varphi_p \circ \Lambda_{F',G'}, P_{S_{\mathcal{T}}}] e^{E_0(\rho|P_{Z|V}, P_{\text{mix},\mathcal{V}}) - \rho(\log|\mathcal{S}_{T+1}| + H_{1+\rho}(S_{I^c}|S_I, S_0))}}{\rho} \leq \epsilon_{\mathcal{I}} \quad (165)$$

holds for any $g' \in G'$.

Thanks to Lemma 50, in the regular case, when the code φ_p is a homomorphism as Abelian group, the above procedure for the construction of our code (Code Construction 6) can be simplified to the following way. It is enough to choose F' and to fix G' to be 0, and we can replace $E_{0,\max}(\rho|P_{Z|V})$ by $E_0(\rho|P_{Z|V}, P_{\text{mix},\mathcal{V}})$. That is, it is enough to calculate

$\inf_{\rho \in (0,1)} E_0(\rho|P_{Z|V}, P_{\text{mix},\mathcal{V}}) - \rho(\log|\mathcal{S}_{T+1}| + H_{1+\rho}(S_{I^c}|S_I, S_0)) - \log \rho$. Due to Lemma 48, $E_0(\rho|P_{Z|V}, P_{\text{mix},\mathcal{V}}) - \rho(\log|\mathcal{S}_{T+1}| + H_{1+\rho}(S_{I^c}|S_I, S_0)) - \log \rho$ is convex with respect to ρ , and the infimum is computable by the bisection method [4, Algorithm 4.1].

Proof of Lemma 50: First, we choose P'_V such that

$$E_{0,\max}(\rho|P_{Z|V}) = E_0(\rho|P_{Z|V}, P'_V). \quad (166)$$

Define P'_{V,v_0} for $v_0 \in \mathcal{V}$ by

$$P'_{V,v_0}(v) = P'_V(v + v_0).$$

Then,

$$e^{E_0(\rho|P_{Z|V}, P'_V)} = e^{E_0(\rho|P_{Z|V}, P'_{V,v_0})}. \quad (167)$$

Hence, we obtain

$$\begin{aligned} e^{E_{0,\max}(\rho|P_{Z|V})} &\stackrel{(a)}{=} e^{E_0(\rho|P_{Z|V}, P'_V)} \stackrel{(b)}{=} \sum_{v_0 \in \mathcal{V}} \frac{1}{|\mathcal{V}|} e^{E_0(\rho|P_{Z|V}, P'_{V,v_0})} \\ &\stackrel{(c)}{\leq} e^{E_0(\rho|P_{Z|V}, \sum_{v_0 \in \mathcal{V}} \frac{1}{|\mathcal{V}|} P'_{V,v_0})} = e^{E_0(\rho|P_{Z|V}, P_{\text{mix},\mathcal{V}})} \stackrel{(d)}{\leq} e^{E_{0,\max}(\rho|P_{Z|V})}, \end{aligned}$$

where (a), (b), (c), and (d) follow from (166), (167), the concavity of $P_V \mapsto e^{E_0(\rho|P_{Z|V}, P_V)}$ (Item (2) of Proposition 2), and the definition (23) of $E_{0,\max}(\rho|P_{Z|V})$, respectively. Thus, we have (164).

Next, we show (165). When the code φ_p is a homomorphism as Abelian group, as is mentioned in Lemma 21, we have $\mathbf{E}_{F',G'=g'} I(S_{\mathcal{I}}; Z|S_0)[P_{Z|V}, \varphi_p \circ \Lambda_{F',g'}, P_{S_{\mathcal{T}}}] = \mathbf{E}_{F',G'} I(S_{\mathcal{I}}; Z|S_0)[P_{Z|V}, \varphi_p \circ \Lambda_{F',g'}, P_{S_{\mathcal{T}}}]$. Hence, combining (163), we obtain (165). ■

When the channel is given as the n -fold discrete memoryless extension $P_{Z|V}^n$ of $P_{Z|V}$, $E_{0,\max}(\rho|P_{Z|V}^n)$ has the following characterization. Using [1], we obtain

$$\max_{P_{V^n}} \sum_{z^n} \left(\sum_{v^n} P_{V^n}(v^n) P_{Z^n|V^n}(z^n|v^n)^{\frac{1}{1-\rho}} \right)^{1-\rho} = e^{nE_{0,\max}(\rho|P_{Z|V})}.$$

Thus, we can apply the above discussion to the n -fold memoryless case by replacing $E_{0,\max}(\rho|P_{Z|V})$ and $P_{Z|V}$ by $nE_{0,\max}(\rho|P_{Z|V})$ and $P_{Z|V}^n$. That is, it is enough to calculate $\inf_{\rho \in (0,1)} nE_{0,\max}(\rho|P_{Z|V}) - \rho(\log|\mathcal{S}_{T+1}| + H_{1+\rho}(S_{I^c}|S_I, S_0)) - \log \rho$. Since, as is mentioned in Proposition 2, $Q_V \mapsto e^{E_0(\rho|\overline{W}^Z, Q_V)}$ is concave and $x \mapsto \log x$ is monotone increasing and concave, $Q_V \mapsto E_0(\rho|\overline{W}^Z, Q_V)$ is concave. Hence, $E_{0,\max}(\rho|P_{Z|V}, Q_V) = \max_{Q_V} E_0(\rho|P_{Z|V}, Q_V)$ can be easily computed. Due to Lemma 48, $nE_{0,\max}(\rho|P_{Z|V}) - \rho(\log|\mathcal{S}_{T+1}| + H_{1+\rho}(S_{I^c}|S_I, S_0)) - \log \rho$ is convex concerning with respect to ρ , the infimum is computable by the bisection method [4, Algorithm 4.1]. Therefore, we can calculate the minimum size $|\mathcal{S}_{T+1}|$ satisfying that $nE_{0,\max}(\rho|P_{Z|V}) - \rho(\log|\mathcal{S}_{T+1}| + H_{1+\rho}(S_{I^c}|S_I, S_0)) - \log \rho$ is smaller than a specified level for all of $\mathcal{I} \subseteq \{1, \dots, T\}$.

B. First Practical Construction: Second Type Evaluation

In the above discussion, we have to consider the maximum value $E_{0,\max}(\rho|P_{Z|V})$. However, when there is no common message and the channel $P_{Z|V}$ is not regular, one can improve the bound (159) in the n -fold memoryless case under the same

code construction (Code Construction 6) as the following way. In the following, we treat the n -fold memoryless extension $P_{Z|V}^n$. Given an encoder $\varphi_p : \mathcal{B}_2 \rightarrow \mathcal{V}^n$, we define the weight distribution P_{φ_p} over the set $T_n(\mathcal{V})$ of types of length n of the set \mathcal{V} by

$$P_{\varphi_p}(Q_V) := \frac{|\{v^n \in \text{Im } \varphi_p | \text{the type of } v^n \text{ is } Q_V.\}|}{|\text{Im } \varphi_p|} \quad (168)$$

for $Q_V \in T_n(\mathcal{V})$. Using the above weight distribution P_{φ_p} , we define the distribution

$$\bar{P}_{\varphi_p}(v^n) := \frac{P_{\varphi_p}(Q_V)}{|T_n(Q_V)|}$$

for $v^n \in \mathcal{V}^n$, where Q_V is the type of v^n and

$$T_n(Q_V) := \{v^n \in \mathcal{V}^n | \text{the type of } v^n \text{ is } Q_V.\}.$$

We construct our code by the same way as Subsection XI-A. We apply Lemma 23 to the case when \mathcal{G} is the n -th permutation group, \mathcal{V} is \mathcal{V}^n , and $P_{Z|V}$ is $P_{Z|V}^n$. Then,

$$e^{\psi(\rho|P_{Z|V}^n, P_{\text{mix}, \mathcal{B}_2})} \leq e^{E_0(\rho|P_{Z|V}^n, \bar{P}_{\varphi_p})}.$$

Hence, combining (160), we obtain

$$\begin{aligned} & \mathbf{E}_{F', G'} \exp(\rho I(S_I; Z) [P_{Z|V}^n, \varphi_p \circ \Lambda_{F', G'}, P_{S_T}]) \\ & \leq 1 + e^{E_0(\rho|P_{Z|V}^n, \bar{P}_{\varphi_p}) - \rho(\log |S_{T+1}| + H_{1+\rho}(S_{T^c} | S_I))}. \end{aligned}$$

Since e^x is convex, we obtain

$$\begin{aligned} & \mathbf{E}_{F', G'} I(S_I; Z) [P_{Z|V}^n, \varphi_p \circ \Lambda_{F', G'}, P_{S_T}] \\ & \leq \frac{e^{E_0(\rho|P_{Z|V}^n, \bar{P}_{\varphi_p}) - \rho(\log |S_{T+1}| + H_{1+\rho}(S_{T^c} | S_I))}}{\rho}. \end{aligned}$$

However, it is not easy to calculate the weight distribution P_{φ_p} for a given code φ_p , but it is possible to give an upper bound for each $P_{\varphi_p}(Q_V)$ in some special cases. For example, the upper bound in the case of binary BCH codes is discussed in [31]. We assume that another distribution Q_{φ_p} over the set $T_n(\mathcal{V})$ and a constant C_1 satisfy

$$C_1 Q_{\varphi_p}(Q_V) \geq P_{\varphi_p}(Q_V)$$

for any $Q_V \in T_n(\mathcal{V})$. Similar to \bar{P}_{φ_p} , we define the distribution \bar{Q}_{φ_p} by

$$\bar{Q}_{\varphi_p}(v^n) := \frac{Q_{\varphi_p}(Q_V)}{|T_n(Q_V)|}$$

for $v^n \in \mathcal{V}^n$, where Q_V is the type of v^n . Hence, Proposition 2 yields

$$e^{E_0(\rho|P_{Z|V}^n, \bar{P}_{\varphi_p})} \leq C_1 e^{E_0(\rho|P_{Z|V}^n, \bar{Q}_{\varphi_p})}.$$

Therefore, we obtain

$$\begin{aligned} & \mathbf{E}_{F', G'} I(S_I; Z) [P_{Z|V}^n, \varphi_p \circ \Lambda_{F', G'}, P_{S_T}] \\ & \leq C_1 \frac{e^{E_0(\rho|P_{Z|V}^n, \bar{Q}_{\varphi_p}) - \rho(\log |S_{T+1}| + H_{1+\rho}(S_{T^c} | S_I))}}{\rho}. \end{aligned} \quad (169)$$

When C_1 is sufficiently small and \bar{Q}_{φ_p} does not give the maximum $E_{0, \max}(\rho|P_{Z|V}^n)$, the RHS of (169) is smaller than the RHS of (159). Similar to the regular case of Subsection XI-A, we can calculate $\inf_{\rho \in (0, 1)} E_0(\rho|P_{Z|V}^n, \bar{Q}_{\varphi_p}) - \rho(\log |S_{T+1}| +$

$H_{1+\rho}(S_{T^c} | S_I, S_0)) - \log \rho + \log C_1$ by the bisection method [4, Algorithm 4.1]. Therefore, in the above case, the method in this subsection improves that in Subsection XI-A.

C. Second Practical Construction

In the previous construction, when the channel is not a regular channel, we have to use an upper bound (159), which is larger than $\frac{e^{E_0(\rho|P_{Z|V}^n, P_{\text{mix}, \mathcal{V}}) - \rho H_{1+\rho}(S_{T^c, *}|S_I, S_0)}}{\rho}$. In order to use a smaller upper bound $\frac{e^{E_0(\rho|P_{Z|V}^n, P_{\text{mix}, \mathcal{V}}) - \rho H_{1+\rho}(S_{T^c, *}|S_I, S_0)}}{\rho}$ even for a non-regular channel, we introduce another practical construction when there is no common message.

Assume that \mathcal{V} has an Abelian group structure. Now, we give a code ensemble from an arbitrary Abelian group \mathcal{B} and an arbitrary encoder $\varphi : \mathcal{B}_2 \rightarrow \mathcal{V}$ satisfying that the map φ is an injective homomorphism. In particular, when \mathcal{B}_2 and \mathcal{V} are vector spaces over the finite field \mathbb{F}_2 , the map φ can be given as a linear code, such as an LDPC code [40] or a Turbo code [41]. However, we do not necessarily need to assume any algebraic structure in the channel $P_{Z, Y|V}$, for now. We stress that in Code Ensemble 7 we use single encoder φ , while in Code Construction 8 we use multiple encoders with the same code length and different information rates.

Code Ensemble 7: We modify the random code given in Lemma 21 as follows. We choose an ensemble of isomorphisms F' from $S_1 \times \dots \times S_{T+1}$ to \mathcal{B}_2 satisfying Condition 15. We choose the random variable $G'' \in \mathcal{V}$ that obeys the uniform distribution on \mathcal{V} statistically independent of the choice of F' . Then, we define the encoder $\tilde{\Lambda}_{F', G''}(s) := (\varphi \circ F')(s) + G''$. The decoder is given by $\hat{\Lambda}_{F', G''}(v) = F'^{-1}(\hat{\varphi}(v - G''))$ by using the decoder $\hat{\varphi}$ of φ .

This code ensemble can be understood in the following way. We define the random variable H in the quotient group $\mathcal{V}/\varphi(\mathcal{B}_2)$ that obeys the uniform distribution. Let $\{y_h\}$ be the set of coset representatives. Let G' be the random variable subject to the uniform distribution on \mathcal{B}_2 . Then, G'' is given as $\varphi(G') + y_H$. That is, the encoder and the decoder can be given as follows. $\tilde{\Lambda}_{F', G', H}(s) := (\varphi \circ F')(s) + G' + y_H$ and $\hat{\Lambda}_{F', G', H}(v) := F'^{-1}(\hat{\varphi}(v - G' - y_H))$.

In Code Ensemble 7, the random variable H corresponds to the choice of the codebook for error correction. Let ε_H be the decoding error probability when we use H as the codebook and the message obeys the uniform distribution. Hence, we consider that ε_H expresses the decoding error probability when we use H as the codebook in the following code construction.

For Code Ensemble 7, we have the following lemma:

Lemma 51: The inequality

$$\begin{aligned} & \mathbf{E}_{F', G', H} e^{\rho I(S_I; Z) [P_{Z|V}, \tilde{\Lambda}_{F', G', H}, P_{S_T}]} \\ & \leq 1 + e^{-\rho H_{1+\rho}(S_{T^c, *}|S_I)} e^{E_0(\rho|P_{Z|V}, P_{\text{mix}, \mathcal{V}})} \end{aligned} \quad (170)$$

holds for each subset $I \subseteq \{1, \dots, T\}$. Thus, applying Jensen inequality to $x \mapsto e^x$, we have

$$\begin{aligned} & \mathbf{E}_{F', G', H} I(S_I; Z) [P_{Z|V}, \tilde{\Lambda}_{F', G', H}, P_{S_T}] \\ & \leq \frac{e^{E_0(\rho|P_{Z|V}, P_{\text{mix}, \mathcal{V}}) - \rho H_{1+\rho}(S_{T^c, *}|S_I)}}{\rho}. \end{aligned} \quad (171)$$

Proof: We apply (161) to the case when $|\mathcal{S}_0| = 1$, $\mathcal{S}_0 = \{s_0\}$, $|\mathcal{B}_1| = 1$, $\mathcal{B}_1 = \{b_1\}$, and the map φ_p is given as $\varphi_p(s_0, b_1, b_2) = \varphi(b_2) + y_h$ for any $b_2 \in \mathcal{B}_2$. Then, we obtain

$$\begin{aligned} & \mathbf{E}_{F', G', H'} e^{\rho I(S_I; Z) [P_{Z|V}, \tilde{\Lambda}_{F', G', H'}, P_{S_T}]} \\ & \leq 1 + e^{-\rho H_{1+\rho}(S_{I^c, *}|S_I)} \sum_z \left(\sum_{b_2} \frac{1}{|\mathcal{B}_2|} P_{Z|V}(z | \varphi(b_2) + y_h)^{\frac{1}{1-\rho}} \right)^{1-\rho}. \end{aligned}$$

Hence, we obtain

$$\begin{aligned} & \mathbf{E}_{F', G', H'} e^{\rho I(S_I; Z) [P_{Z|V}, \tilde{\Lambda}_{F', G', H'}, P_{S_T}]} \\ & = \mathbf{E}_H \mathbf{E}_{F', G', H'} e^{\rho I(S_I; Z) [P_{Z|V}, \tilde{\Lambda}_{F', G', H'}, P_{S_T}]} \\ & \leq 1 + e^{-\rho H_{1+\rho}(S_{I^c, *}|S_I)} \mathbf{E}_H \sum_z \left(\sum_{b_2} \frac{1}{|\mathcal{B}_2|} P_{Z|V}(z | \varphi(b_2) + y_H)^{\frac{1}{1-\rho}} \right)^{1-\rho} \\ & \leq 1 + e^{-\rho H_{1+\rho}(S_{I^c, *}|S_I)} \sum_z \left(\mathbf{E}_H \sum_{b_2} \frac{1}{|\mathcal{B}_2|} P_{Z|V}(z | \varphi(b_2) + y_H)^{\frac{1}{1-\rho}} \right)^{1-\rho} \\ & = 1 + e^{-\rho H_{1+\rho}(S_{I^c, *}|S_I)} e^{E_0(\rho | P_{Z|V}, P_{\text{mix}, \gamma^n})}, \end{aligned}$$

which implies (170). \blacksquare

In order to construct a code for the secure multiplex coding (with no common message), we define the notations as follows. Let ϵ_I be the maximum acceptable information leakage for $I(S_I; Z)$ for each $I \subseteq \{1, \dots, T\}$. Let ϵ_b be the maximum acceptable error probability. Let ϵ_2 be the maximum acceptable probability a chosen F', G'' not making $I(S_I; Z)$ below ϵ_I . These parameters ϵ_b , ϵ_I , and ϵ_2 are the requirements for our code construction.

Code Construction 8: In this construction, in contrast to Subsections XI-A and XI-B we assume that we are given multiple error-correcting codes with the same code length n and different information rates. Using (171), we construct a code for the secure multiplex coding (with no common message) as follows:

- 1) We choose a suitable Abelian group \mathcal{B}_2 , a suitable code φ , a suitable sacrifice bit length (the size of T -th message), and a suitable real value $\epsilon_1 \in (0, 1)$ satisfying that

$$\epsilon_b \geq \frac{\mathbf{E}_H \epsilon_H}{\epsilon_1} \quad (172)$$

$$\epsilon_I \geq 2^T \min_{\rho \in (0, 1)} \frac{e^{E_0(\rho | P_{Z|V}, P_{\text{mix}, \gamma^n}) - \rho H_{1+\rho}(S_{I^c, *}|S_I)}}{\rho \epsilon_2 (1 - \epsilon_1)}. \quad (173)$$

- 2) We choose H randomly. Then, we check that ϵ_H is less than ϵ_b . If not, we choose another H . We repeat this process until it is successful. We denote the final choice of H by H' . Thanks to Markov inequality and (172), the successful probability for one trial is at least $1 - \epsilon_1$.
- 3) We choose F' and G' randomly. Then, we obtain the pair of the encoder $\tilde{\Lambda}_{F', G', H'}(s) := (\varphi \circ F')(s) + G' + y_{H'}$ and the decoder $\hat{\Lambda}_{F', G', H'}(v) := F'^{-1}(\hat{\varphi}(v - G' - y_{H'}))$.

Theorem 52: Under the above construction, the inequality

$$I(S_I; Z) [P_{Z|V}, \tilde{\Lambda}_{F', G', H'}, P_{S_T}] \leq \epsilon_I \quad (174)$$

holds for all subsets $I \subseteq \{1, \dots, T\}$ with at least with probability $1 - \epsilon_2$.

Proof: Markov inequality guarantees that $\Pr\{\epsilon_H \leq \epsilon_b\} \geq 1 - \epsilon_1$. Hence, we obtain

$$\begin{aligned} & \mathbf{E}_{F', G', H'} I(S_I; Z) [P_{Z|V}, \tilde{\Lambda}_{F', G', H'}, P_{S_T}] \\ & = \mathbf{E}_{F', G', H' | \epsilon_H \leq \epsilon_b} I(S_I; Z) [P_{Z|V}, \tilde{\Lambda}_{F', G', H'}, P_{S_T}] \\ & \leq \frac{\Pr\{\epsilon_H \leq \epsilon_b\}}{\Pr\{\epsilon_H \leq \epsilon_b\}} \mathbf{E}_{F', G', H' | \epsilon_H \leq \epsilon_b} I(S_I; Z) [P_{Z|V}, \tilde{\Lambda}_{F', G', H'}, P_{S_T}] \\ & \quad + \frac{\Pr\{\epsilon_H > \epsilon_b\}}{\Pr\{\epsilon_H \leq \epsilon_b\}} \mathbf{E}_{F', G', H' | \epsilon_H > \epsilon_b} I(S_I; Z) [P_{Z|V}, \tilde{\Lambda}_{F', G', H'}, P_{S_T}] \\ & = \frac{1}{\Pr\{\epsilon_H \leq \epsilon_b\}} \mathbf{E}_{F', G', H'} I(S_I; Z) [P_{Z|V}, \tilde{\Lambda}_{F', G', H'}, P_{S_T}] \\ & \leq \frac{1}{1 - \epsilon_1} \mathbf{E}_{F', G', H'} I(S_I; Z) [P_{Z|V}, \tilde{\Lambda}_{F', G', H'}, P_{S_T}] \\ & \leq \epsilon_2 \epsilon_I / 2^T \end{aligned}$$

for every I , where $\mathbf{E}_{F', G', H' | \epsilon_H \leq \epsilon_b}$ denotes the expectation under the condition $\epsilon_H \leq \epsilon_b$. The final inequality follows from (171). Since the above choice of F', G' and H' is restricted to the set $\{(f', g', h') | \epsilon_h \leq \epsilon_b\}$, due to Markov inequality, the probability of choosing F', G' and H' making (174) simultaneously for all $I \subseteq \{1, \dots, T\}$ is not less than $1 - \epsilon_2$. \blacksquare

Further, when the channel is given as the n -fold discrete memoryless extension $P_{Z|V}^n$ of $P_{Z|V}$, the quantity $E_0(\rho | P_{Z|V}^n, P_{\text{mix}, \gamma^n})$ is simplified to $n E_0(\rho | P_{Z|V}, P_{\text{mix}, \gamma^n})$. Hence, similar to the regular case of Subsection XI-A, we can calculate the right hand side of (173) by the bisection method [4, Algorithm 4.1].

XII. CHANNEL-UNIVERSAL CODING FOR SECURE MULTIPLEX CODING WITH COMMON MESSAGES

In order to treat universal coding for the multiplex coding with common messages, we introduce the universally attainable exponents of the multiplex coding with common messages in the n -fold discrete memoryless setting by adjusting the original definition for the BCD given by Körner and Sgarro [24]. Similar to Subsection X-B, in this section, we employ $T + 1$ -th message S_{T+1} as a dummy message subject to the uniform distribution, and assume that the $T + 1$ -th message $S_{T+1, n}$ is subject to the uniform distribution. We simplify $P_{S_{T, n}} \times P_{S_{T+1, n}}$ by $P_{S_{T, n}}$. For a subset $I \subseteq \{1, \dots, T\}$, we denote the complementary set in $\{1, \dots, T\}$ by I^c and simplify the set $I^c \cup \{T + 1\}$ to $I^{c, *}$.

In order to treat universal coding for secure multiplex coding with common messages, we focus on $2^{T+1} - 2$ functions to express the evaluations of the exponential decreasing rates of decoding error probabilities and the asymptotic evaluations of leaked information. For describing bounds of the exponential decreasing rates of both decoding error probabilities, we need two functions. For treating the asymptotic evaluations of leaked information, we need $2^{T+1} - 4$ functions because the number of non-empty proper subsets $I (\neq \emptyset) \subseteq \{1, \dots, T\}$ is $2^T - 2$ and we treat the exponential decreasing rates and the information leakage rates of leaked information for respective non-empty proper subsets $I (\neq \emptyset) \subseteq \{1, \dots, T\}$. Then, we need to treat $2^{T+1} - 2$ functions. Since we do not assume the uniformity, we cannot describe our bounds of the exponential decreasing rate and the information leakage

rate of leaked information as functions of the rate tuples $(R_p, R_c, (R_i)_{i=0,1,\dots,T,T+1})$. In the following discussion, we treat our bound of the exponential decreasing rate of leaked information for a non-empty proper subset $\mathcal{I}(\neq \emptyset) \subseteq \{1, \dots, T\}$ as a function of $\underline{H}_2(\mathcal{I}^{c,*})$, R_c , and R_0 as well as the channel W . Similarly, we treat our bound of the information leakage rate of leaked information for a non-empty proper subset $\mathcal{I}(\neq \emptyset) \subseteq \{1, \dots, T\}$ as a function of $\underline{H}_{\log}(\mathcal{I}^{c,*})$, R_c , and R_0 as well as the channel W . Our bounds of the exponential decreasing rates of both decoding error probabilities are described as functions of R_p , R_c , and the channel W . Hence, the outcomes of the above $2^{T+1} - 2$ functions are decided by $2^{T+1} - 1$ real numbers R_p , R_c , R_0 , and $(\underline{H}_2(\mathcal{I}^{c,*}), \underline{H}_{\log}(\mathcal{I}^{c,*}))_{\mathcal{I}(\neq \emptyset) \subseteq \{1,\dots,T\}}$ as well as the channel W .

Definition 53: A set of functions $(E^b, E^e, (E_+^{\mathcal{I}}, E_-^{\mathcal{I}})_{\mathcal{I} \subseteq \{1,\dots,T\}})$ from $\mathbf{R}_{\geq 0}^{2^{T+1}-1} \times \mathcal{W}(\mathcal{X}, \mathcal{Y} \times \mathcal{Z})$ to $\mathbf{R}_{\geq 0}^{2^{T+1}-2}$ is said to be a universally attainable set of exponents and information leakage rate for the family $\mathcal{W}(\mathcal{X}, \mathcal{Y} \times \mathcal{Z})$ if for any $\epsilon > 0$ and any rate tuples $(R_p, R_c, (R_i)_{i=0,1,\dots,T})$, there exist a sufficiently large integer N and a sequence of codes φ_n of length n satisfying the following conditions: (1) The i -th secret message set $\mathcal{S}_{i,n}$ of the code φ_n has cardinality e^{nR_i} for $i = 1, \dots, T$, and the common message sets $\mathcal{S}_{0,n}$ has cardinality e^{nR_0} . (2) Any sequence of joint distributions $P_{S_{\mathcal{T},n}}$ for all of the i -th secret $S_{i,n}$ on $\mathcal{S}_{i,n}$ and the common message $S_{0,n}$ on $\mathcal{S}_{0,n}$ satisfies the inequalities

$$P_b[W^n, \varphi_n, P_{S_{\mathcal{T}+\infty,n}}] \leq \exp(-n[E^b(R_p, R_c, R_0, W) - \epsilon]), \quad (175)$$

$$P_e[W^n, \varphi_n, P_{S_{\mathcal{T}+\infty,n}}] \leq \exp(-n[E^e(R_p, R_c, R_0, W) - \epsilon]), \quad (176)$$

and

$$\liminf_{n \rightarrow \infty} \frac{-1}{n} \log I(S_{\mathcal{I},n}; Z^n | S_{0,n})[W^n, \varphi_n, P_{S_{\mathcal{T}+\infty,n}}] \geq E_+^{\mathcal{I}}(R_p, R_c, R_0, (\underline{H}_2(\mathcal{I}^{c,*}), \underline{H}_{\log}(\mathcal{I}^{c,*}))_{\mathcal{I}(\neq \emptyset) \subseteq \{1,\dots,T\}}, W), \quad (177)$$

$$\limsup_{n \rightarrow \infty} \frac{1}{n} I(S_{\mathcal{I},n}; Z^n | S_{0,n})[W^n, \varphi_n, P_{S_{\mathcal{T}+\infty,n}}] \leq E_-^{\mathcal{I}}(R_p, R_c, R_0, (\underline{H}_2(\mathcal{I}^{c,*}), \underline{H}_{\log}(\mathcal{I}^{c,*}))_{\mathcal{I}(\neq \emptyset) \subseteq \{1,\dots,T\}}, W), \quad (178)$$

hold for any channel $W \in \mathcal{W}(\mathcal{X}, \mathcal{Y} \times \mathcal{Z})$, any non-empty proper subset $\mathcal{I}(\neq \emptyset) \subseteq \{1, \dots, T\}$, and any $n \geq N$. Here, $E^b(R_p, R_c, R_0, (\underline{H}_2(\mathcal{I}^{c,*}), \underline{H}_{\log}(\mathcal{I}^{c,*}))_{\mathcal{I}(\neq \emptyset) \subseteq \{1,\dots,T\}}, W)$ and $E^e(R_p, R_c, R_0, (\underline{H}_2(\mathcal{I}^{c,*}), \underline{H}_{\log}(\mathcal{I}^{c,*}))_{\mathcal{I}(\neq \emptyset) \subseteq \{1,\dots,T\}}, W)$ are abbreviated to $E^b(R_p, R_c, R_0, W)$ and $E^e(R_p, R_c, R_0, W)$ because they do not depend on

$$(\underline{H}_2(\mathcal{I}^{c,*}), \underline{H}_{\log}(\mathcal{I}^{c,*}))_{\mathcal{I}(\neq \emptyset) \subseteq \{1,\dots,T\}}.$$

For the reason why we employ the limiting forms in (177) and (178), see Remark 60. Note that we do not consider here the universality for source while Körner and Sgarro [24] show the universality for source as well as that for channel, as reviewed in Theorem 13 of this paper. In order to guarantee the secrecy for $\mathcal{S}_{\mathcal{I},n}$, we need sufficient randomness of $\mathcal{S}_{\mathcal{I}^c,n}$. That is, the secrecy of $\mathcal{S}_{\mathcal{I},n}$ depends on $\underline{H}_2(\mathcal{I}^c)$ and $\underline{H}_{\log}(\mathcal{I}^c)$, which depends on the source distribution. Hence, it is impossible to show the universality for source in SMC.

We fix a distribution Q_{VU} on $\mathcal{U} \times \mathcal{V}$ and a channel $\Xi : \mathcal{V} \rightarrow \mathcal{X}$. Then, we present a universally attainable set of exponents and leaked information rate in terms of Q_{VU} and Ξ in the following way. Given a broadcast $W : \mathcal{X} \rightarrow \mathcal{Y} \times \mathcal{Z}$ and the real

numbers $(R_p, R_c, R_0, (\underline{H}_2(\mathcal{I}^{c,*}), \underline{H}_{\log}(\mathcal{I}^{c,*}))_{\mathcal{I}(\neq \emptyset) \subseteq \{1,\dots,T\}})$, the tuple of exponents and information leakage rate are given as

$$\begin{aligned} E^b &= E^b(R_p, R_c, R_0, W) \\ &:= \tilde{E}^b(R_p, R_c, (W^{\mathcal{Y}} \circ \Xi) \times Q_{VU}), \end{aligned} \quad (179)$$

$$\begin{aligned} E^e &= E^e(R_p, R_c, R_0, W) \\ &:= \tilde{E}^e(R_c, (W^{\mathcal{Z}} \circ \Xi) \times Q_{VU}), \end{aligned} \quad (180)$$

$$\begin{aligned} E_+^{\mathcal{I}} &= E_+^{\mathcal{I}}(R_p, R_c, R_0, (\underline{H}_2(\mathcal{I}^{c,*}), \underline{H}_{\log}(\mathcal{I}^{c,*}))_{\mathcal{I}(\neq \emptyset) \subseteq \{1,\dots,T\}}, W) \\ &:= \tilde{E}^{\mathcal{I}}(\underline{H}_2(\mathcal{I}^{c,*}) - R_c + R_0, (W^{\mathcal{Z}} \circ \Xi) \times Q_{VU}), \end{aligned} \quad (181)$$

$$\begin{aligned} E_-^{\mathcal{I}} &= E_-^{\mathcal{I}}(R_p, R_c, R_0, (\underline{H}_2(\mathcal{I}^{c,*}), \underline{H}_{\log}(\mathcal{I}^{c,*}))_{\mathcal{I}(\neq \emptyset) \subseteq \{1,\dots,T\}}, W) \\ &:= I(V; Z|U)[(W^{\mathcal{Z}} \circ \Xi) \times Q_{VU}] - \underline{H}_{\log}(\mathcal{I}^{c,*}) + R_c - R_0 \end{aligned} \quad (182)$$

for a non-empty proper subset $\mathcal{I}(\neq \emptyset) \subseteq \{1, \dots, T\}$, where \tilde{E}^b , \tilde{E}^e , \tilde{E}^{E_0} , and $\tilde{E}^{\mathcal{I}}$ are given by (29), (30), (22), and (24), respectively.

Hence, our quadruple of exponents and information leakage rate depends on Q_{VU} and Ξ .

Theorem 54 (Extension of [24, Theorem 1, part (a)]):

Eqs. (179)–(182) are universally attainable rates of exponents and information leakage rate in the sense of Definition 53.

Proof: In the proof, since we treat the channel $W^{\mathcal{Z}} \circ \Xi : \mathcal{V} \rightarrow \mathcal{Z}$, we abbreviate it as $\overline{W}^{\mathcal{Z}}$. First, we give the outline of our proof. We shall modify the constant composition code used by Körner and Sgarro [24]. We do not evaluate the decoding error probability, because that of our code is not larger than that given in [24]. Observe that our exponents in Eqs. (179) and (180) are the same as [24] with the channel $\overline{W}^{\mathcal{Z}} = W^{\mathcal{Z}} \circ \Xi$. We shall evaluate only the mutual information. For this purpose, we prepare general notations and properties of type and conditional type in Step (1). Next, in Steps (2) and (3), we prepare several notations and properties of type and conditional type that are specific to our proof. In Step (4), we apply the random coding and evaluate the leaked information when the channel is given by the conditional types. Then, we choose a code whose leaked information is evaluated for all conditional types and whose error is evaluated for all discrete memoryless channels. In Step (5), we evaluate the leaked information under the above chosen code for all discrete memoryless channels.

Step (1): Preparation of general notations and properties of type and conditional type:

For the following construction of our code, we prepare general notations for types. These notations will be used also in the next section. For a given type Q_U of length n on a set \mathcal{U} , we define the set $T_n(Q_U)$ as

$$T_n(Q_U) := \{u^n \in \mathcal{U}^n \mid \text{the type of } u^n \text{ is } Q_U\}.$$

Hence, for a given type Q_{VU} of length n on a set $\mathcal{V} \times \mathcal{U}$, the set $T_n(Q_{VU})$ is written as

$$T_n(Q_{VU}) = \{(v^n, u^n) \in \mathcal{V}^n \times \mathcal{U}^n \mid \text{the type of } (v^n, u^n) \text{ is } Q_{VU}\}.$$

The marginal distribution Q_U over \mathcal{U} of the type Q_{VU} of length n on the set $\mathcal{V} \times \mathcal{U}$ is a type of length n on the set \mathcal{U} . Given a type Q_V of length n on the set \mathcal{V} , we define the set of

conditional types on the set \mathcal{V} with respect to Q_V as

$$\begin{aligned} & \mathcal{T}_{n,\mathcal{V}}(Q_U) \\ & := \{\text{probability transition matrix } W \text{ from } \mathcal{U} \text{ to } \mathcal{V} \\ & \quad | W \times Q_U \text{ is a type of length } n \text{ on a set } \mathcal{V} \times \mathcal{U}\}. \end{aligned}$$

The cardinality $|\mathcal{T}_{n,\mathcal{V}}(Q_U)|$ is upper bounded as [8]

$$|\mathcal{T}_{n,\mathcal{V}}(Q_U)| \leq (n+1)^{|\mathcal{V} \times \mathcal{U}|}. \quad (183)$$

In particular, given a type Q_{VU} of length n on the set $\mathcal{V} \times \mathcal{U}$, we define the conditional type $Q_{V|U}$ such that $Q_{VU} = Q_{V|U} \times Q_U$. We also define the set $T_n(Q_{V|U})_{U^n=U^n}$ as

$$T_n(Q_{V|U})_{U^n=U^n} := \{v^n \in \mathcal{V}^n \mid \text{the type of } (v^n, U^n) \text{ is } Q_{VU}\}.$$

We denote the uniform distribution $P_{\text{mix},T_n(Q_U)}$ on $T_n(Q_U)$ by $\Upsilon_n(Q_U)$. Then, for a given type Q_{VU} of length n on a set $\mathcal{V} \times \mathcal{U}$, $\Upsilon_n(Q_{VU})$ represents the uniform distribution $P_{\text{mix},T_n(Q_{VU})}$ on $T_n(Q_{VU})$. Further, for an arbitrary $W \in \mathcal{T}_{n,\mathcal{V}}(Q_U)$, $\Upsilon_n(W \times Q_U)$ represents the uniform distribution on $T_n(W \times Q_U)$. Then, we define the probability transition matrix $\Upsilon_n(W)$ from \mathcal{V}^n to \mathcal{U}^n such that $\Upsilon_n(W) \times \Upsilon_n(Q_U) = \Upsilon_n(W \times Q_U)$.

When $P_{V^n U^n}$ is a distribution over $\mathcal{V}^n \times \mathcal{U}^n$ and invariant under the permutation of the indices, the distribution $P_{V^n U^n}$ can be written as

$$P_{V^n U^n} = \sum_{Q_{VU}} \lambda_{P_{V^n U^n}}(Q_{VU}) \Upsilon_n(Q_{VU}) \quad (184)$$

with non-negative constants $\lambda(Q_{VU})$. In particular, the independent and identical distribution P_V^n of P_V can be written as

$$P_V^n = \sum_{Q_V} \lambda_{P_V^n}(Q_V) \Upsilon_n(Q_V) \quad (185)$$

with

$$\lambda_{P_V^n}(Q_V) = P_V^n(T_n(Q_V)) \leq e^{-nD(Q_V \| P_V)}. \quad (186)$$

When the marginal distribution over \mathcal{U}^n of $P_{V^n U^n}$ can be written as $P_{\text{mix},T_n(Q_U)} = \Upsilon_n(Q_U)$ with a type Q_U on the set \mathcal{U} , we have

$$\begin{aligned} P_{V^n U^n} &= \sum_{Q_{V|U} \in \mathcal{T}_{n,\mathcal{V}}(Q_U)} \lambda_{P_{V^n U^n}}(Q_{V|U} \times Q_U) \Upsilon_n(Q_{V|U} \times Q_U) \\ &= \sum_{Q_{V|U} \in \mathcal{T}_{n,\mathcal{V}}(Q_U)} \lambda_{P_{V^n U^n}}(Q_{V|U} \times Q_U) (\Upsilon_n(Q_{V|U}) \times \Upsilon_n(Q_U)) \\ &= \left(\sum_{Q_{V|U} \in \mathcal{T}_{n,\mathcal{V}}(Q_U)} \lambda_{P_{V^n U^n}}(Q_{V|U} \times Q_U) \Upsilon_n(Q_{V|U}) \right) \times \Upsilon_n(Q_U). \end{aligned} \quad (187)$$

We define the channel $P_{V^n|U^n}$ by $P_{V^n|U^n} = P_{V^n U^n} \times \Upsilon_n(Q_U)$ and the real number $\lambda_{P_{V^n|U^n}}(Q_{V|U}) := \lambda_{P_{V^n U^n}}(Q_{V|U} \times Q_U)$ for $Q_{V|U} \in \mathcal{T}_{n,\mathcal{V}}(Q_U)$. Then, we obtain

$$P_{V^n|U^n} = \sum_{Q_{V|U} \in \mathcal{T}_{n,\mathcal{V}}(Q_U)} \lambda_{P_{V^n|U^n}}(Q_{V|U}) \Upsilon_n(Q_{V|U}). \quad (188)$$

Now, we consider the n -fold discrete memoryless channel $P_{V^n|U^n}$. For a given type Q_U on the set \mathcal{U} , we apply the relation

(187) to the joint distribution $P_{V^n|U^n} \times \Upsilon_n(Q_U)$. Then, (188) implies that

$$P_{V^n|U^n}^{P_{V^n|U^n}} = \sum_{Q_{V|U} \in \mathcal{T}_{n,\mathcal{V}}(Q_U)} \lambda_{P_{V^n|U^n}}(Q_{V|U}) \Upsilon_n(Q_{V|U}). \quad (189)$$

Choosing $u^n \in T_n(Q_U)$, we have

$$\Upsilon_n(Q_{V|U})(T_n(Q_{V|U})_{U^n=U^n} | U^n = u^n) = \begin{cases} 1 & \text{if } Q_{V|U} = Q_{V|U} \\ 0 & \text{otherwise.} \end{cases} \quad (190)$$

Combining (189) and (190), we obtain

$$\begin{aligned} & \lambda_{P_{V^n|U^n}}(Q_{V|U}) \\ &= P_{V^n|U^n}^{P_{V^n|U^n}}(T_n(Q_{V|U})_{U^n=U^n} | U^n = u^n) \\ &= \prod_{u \in \mathcal{U}} (P_{V|U=u})^{n Q_U(u)} (T_{n_u}(Q_{V|U=u})) \\ &\leq e^{-\sum_{u \in \mathcal{U}} n Q_U(u) D(Q_{V|U=u} \| P_{V|U=u})} \\ &= e^{-nD(Q_{V|U} \| P_{V|U} | Q_U)}, \end{aligned} \quad (191)$$

$$= e^{-nD(Q_{V|U} \| P_{V|U} | Q_U)}, \quad (192)$$

where (191) follows from (186).

Step (2): Preparation of notations and properties of conditional types based on a joint type on $\mathcal{U} \times \mathcal{V}$:

In this step, we prepare several important properties based on a type of length n on the set $\mathcal{U} \times \mathcal{V} \times \mathcal{Z}$. Now, we focus on a conditional type $W^Z \in \mathcal{T}_{n,\mathcal{Z}}(Q_{VU})$, which gives a type $W^Z \times Q_{VU}$ of length n on the set $\mathcal{U} \times \mathcal{V} \times \mathcal{Z}$. Note that in order to make a type of length n on the set $\mathcal{U} \times \mathcal{V} \times \mathcal{Z}$, we need to choose W^Z not from $\mathcal{T}_{n,\mathcal{Z}}(Q_V)$ but from $\mathcal{T}_{n,\mathcal{Z}}(Q_{VU})$. Now, we treat the channel \overline{W}^Z as a channel from $\mathcal{V} \times \mathcal{U}$ to \mathcal{Z} while the output distribution of the channel \overline{W}^Z does not depend on the choice of $u \in \mathcal{U}$. In our code $\varphi_{a,n}$, the random variable $V^n U^n$ takes values in the subset $T_n(Q_{VU})$. Hence, it is sufficient to treat the channel whose input alphabet is the subset $T_n(Q_{VU})$ of $\mathcal{V}^n \times \mathcal{U}^n$. Based on (189), we make a convex decomposition

$$\overline{W}^{Z,n} |_{T_n(Q_{VU})} = \sum_{W^Z \in \mathcal{T}_{n,\mathcal{Z}}(Q_{VU})} \lambda_{n,T}(W^Z) \Upsilon_n(W^Z), \quad (193)$$

with non-negative constants $\lambda_{n,T}(W^Z)$. Then, due to (192), we have

$$\lambda_{n,T}(W^Z) \leq e^{-nD(W^Z \| \overline{W}^Z | Q_{VU})}. \quad (194)$$

For an arbitrary code $\varphi_{a,n}$, the joint convexity of the conditional relative entropy yields that

$$\begin{aligned} & I(S_{I,n}; Z^n | S_{0,n}) [\overline{W}^{Z,n}, \varphi_{a,n}, P_{S_{\mathcal{T}+\infty,n}}] \\ & \leq \sum_{W^Z \in \mathcal{T}_{n,\mathcal{Z}}(Q_{VU})} \lambda_{n,T}(W^Z) I(S_{I,n}; Z^n | S_{0,n}) [\Upsilon_n(W^Z), \varphi_{a,n}, P_{S_{\mathcal{T}+\infty,n}}]. \end{aligned} \quad (195)$$

Next, in order to treat each channel $\Upsilon_n(W^Z)$, we fix a conditional type $W^Z \in \mathcal{T}_{n,\mathcal{Z}}(Q_{VU})$ and study the properties of the channel $\Upsilon_n(W^Z)$. Under the joint type $Q_{ZVU} := W^Z \times Q_{VU}$, we define the numbers

$$\begin{aligned} N(U) &:= |T_n(Q_U)|, & N(UZ) &:= |T_n((W^Z \circ Q_{V|U}) \times Q_U)|, \\ N(VU) &:= |T_n(Q_{VU})|, & N(VUZ) &:= |T_n(W^Z \times Q_{VU})|, \end{aligned}$$

and

$$\begin{aligned} N(Z|U) &:= N(UZ)/N(U), & N(V|UZ) &:= N(VUZ)/N(UZ), \\ N(V|U) &:= N(VU)/N(U), & N(Z|VU) &:= N(VUZ)/N(VU). \end{aligned}$$

Then, due to [8], we have

$$|\mathcal{T}_{n,Z}(Q_U)|^{-1} e^{nH(Z|U)[W^Z \times Q_{VU}]} \leq N(Z|U) \leq e^{nH(Z|U)[W^Z \times Q_{VU}]} \quad (196)$$

$$|\mathcal{T}_{n,Z}(Q_{VU})|^{-1} e^{nH(Z|VU)[W^Z \times Q_{VU}]} \leq N(Z|VU) \leq e^{nH(Z|VU)[W^Z \times Q_{VU}]} \quad (197)$$

Then, we obtain the following lemma.

Lemma 55: Any conditional type $W^Z \in \mathcal{T}_{n,Z}(Q_{VU})$ satisfies

$$\begin{aligned} &E_0(\rho | \Upsilon_n(W^Z), P_{V^n|U^n, \text{mix}, T_n(Q_{VU})}, P_{\text{mix}, T_n(Q_{VU})}) \\ &= \rho \log \frac{N(Z|U)}{N(Z|VU)} \end{aligned} \quad (198)$$

$$= \rho I(V; Z|U) [\Upsilon_n(W^Z) \times P_{\text{mix}, T_n(Q_{VU})}] \quad (199)$$

$$\leq n\rho I(V; Z|U)[W^Z \times Q_{VU}] + \rho \log |\mathcal{T}_{n,Z}(Q_{VU})| \quad (200)$$

for any $\rho \in (0, 1)$. Here $P_{V^n|U^n, \text{mix}, T_n(Q_{VU})}$ is defined as a special case of Eq.(1).

Proof: Under the joint type $Q_{ZVU} := W^Z \times Q_{VU}$, since $\Upsilon_n(W^Z) = P_{Z^n|V^n U^n, \text{mix}, T_n(Q_{ZVU})}$, we obtain

$$\begin{aligned} &e^{E_0(\rho | \Upsilon_n(W^Z), P_{V^n|U^n, \text{mix}, T_n(Q_{VU})}, P_{\text{mix}, T_n(Q_{VU})})} \\ &= e^{E_0(\rho | P_{Z^n|V^n U^n, \text{mix}, T_n(Q_{ZVU})}, P_{V^n|U^n, \text{mix}, T_n(Q_{VU})}, P_{\text{mix}, T_n(Q_{VU})})} \\ &= \sum_{u^n \in T_n(Q_U)} \frac{1}{N(U)} \sum_{z^n \in T_n(Q_{ZU}) | U^n = u^n} \left(\sum_{v \in T_n(Q_{VZU}) | z^n U^n = (z^n, u^n)} P_{V^n|U^n, \text{mix}, T_n(Q_{VU})}(v^n | u^n) \right. \\ &\quad \left. \cdot (P_{Z^n|V^n U^n, \text{mix}, T_n(Q_{ZVU})}(z^n | v^n, u^n))^{\frac{1}{1-\rho}} \right)^{1-\rho} \\ &= \sum_{u^n \in T_n(Q_U)} \frac{1}{N(U)} \sum_{z^n \in T_n(Q_{ZU}) | U^n = u^n} \left(\sum_{v \in T_n(Q_{VZU}) | z^n U^n = (z^n, u^n)} \frac{1}{N(V|U)} \left(\frac{1}{N(Z|VU)} \right)^{\frac{1}{1-\rho}} \right)^{1-\rho} \\ &= N(U) \frac{1}{N(U)} N(Z|U) (N(V|UZ)) \frac{1}{N(V|U)} \left(\frac{1}{N(Z|VU)} \right)^{\frac{1}{1-\rho}})^{1-\rho} \\ &= \frac{N(ZU)^\rho N(VU)^\rho}{N(VUZ)^\rho N(U)^\rho} = \frac{N(Z|U)^\rho}{N(Z|VU)^\rho}, \end{aligned}$$

which implies (198). Since

$$\begin{aligned} &\log N(Z|U) - \log N(Z|VU) \\ &= H(Z|U) [\Upsilon_n(W^Z) \times P_{\text{mix}, T_n(Q_{VU})}] \\ &\quad - H(Z|VU) [\Upsilon_n(W^Z) \times P_{\text{mix}, T_n(Q_{VU})}] \\ &= I(V; Z|U) [\Upsilon_n(W^Z) \times P_{\text{mix}, T_n(Q_{VU})}], \end{aligned}$$

we obtain (199). Combining (196) and (197), we obtain (200). \blacksquare

Step (3): Preparation of notations and properties concerning conditional types based on a type on \mathcal{V} :

In this step, we focus only on a convex decomposition different from (193). For a given type Q_V of length n on a

set \mathcal{V} , we focus on the set

$$\mathcal{W}_{n,Z}(Q_V) := \{\Upsilon_n(W^Z) | W^Z \in \mathcal{T}_{n,Z}(Q_V)\}.$$

In our code $\varphi_{a,n}$, the random variable V^n takes values in the subset $T_n(Q_V)$. Hence, if we focus on the set \mathcal{V}^n as inputs, it is sufficient to treat the channel whose input alphabet is the subset $T_n(Q_V)$ of \mathcal{V}^n . Then, due to (189), we have another type of convex combination:

$$\overline{W}^{Z,n} |_{T_n(Q_V)} = \sum_{\Theta_n \in \mathcal{W}_{n,Z}(Q_V)} \lambda_{n,W}(\Theta_n) \Theta_n, \quad (201)$$

where $\lambda_{n,W}(\Theta_n)$ is a non-negative constant. Then, for an arbitrary code $\varphi_{a,n}$, the joint convexity of the conditional relative entropy yields that

$$\begin{aligned} &I(S_{I,n}; Z^n | S_{0,n}) [\overline{W}^{Z,n}, \varphi_{a,n}, P_{S_{\mathcal{T}+\infty,n}}] \\ &\leq \sum_{\Theta_n \in \mathcal{W}_{n,Z}(Q_V)} \lambda_{n,W}(\Theta_n) I(S_{I,n}; Z^n | S_{0,n}) [\Theta_n, \varphi_{a,n}, P_{S_{\mathcal{T}+\infty,n}}]. \end{aligned} \quad (202)$$

Next, we introduce the quantity

$$\begin{aligned} &\varepsilon_{n,\rho,I}(W^{Z^n}, Q_{V^n, U^n}) \\ &:= \exp\left(n\rho(R_c - R_0) - \rho H_{1+\rho}(S_{I^{c,*},n} | S_{I,n}, S_{0,n}) \right. \\ &\quad \left. + E_0(\rho | W^{Z^n}, Q_{V^n|U^n}, Q_{U^n})\right) \end{aligned} \quad (203)$$

for any channel W^{Z^n} from \mathcal{V}^n to \mathcal{Z}^n and any distribution $Q_{V^n U^n}$ on $\mathcal{V}^n \times \mathcal{U}^n$.

Then, we have the following lemma.

Lemma 56: Any joint type Q_{VU} of length n on a set $\mathcal{V} \times \mathcal{U}$ and any channel $\Theta_n \in \mathcal{W}_{n,Z}(Q_V)$ satisfy

$$\begin{aligned} &\exp(E_0(\rho | \overline{W}^{Z,n}, P_{V^n|U^n, \text{mix}, T_n(Q_{VU})}, P_{\text{mix}, T_n(Q_{VU})})) \\ &\leq (n+1)^{|\mathcal{U}|^2 |\mathcal{V}|} \exp(E_0(\rho | \overline{W}^{Z,n}, Q_{V|U}^n, Q_U^n)), \end{aligned} \quad (204)$$

$$\begin{aligned} &\lambda_{n,W}(\Theta_n) \varepsilon_{n,\rho}(\Theta_n, P_{\text{mix}, T_n(Q_{VU})}) \\ &\leq (n+1)^{|\mathcal{U}|^2 |\mathcal{V}|} \varepsilon_{n,\rho,I}(\overline{W}^{Z,n}, Q_{V,U}). \end{aligned} \quad (205)$$

We have

$$\begin{aligned} &\limsup_{n \rightarrow \infty} \frac{1}{n\rho_n} \log \varepsilon_{n,\rho_n,I}(\overline{W}^{Z,n}, Q_{V,U}^n) \\ &\leq I(V; Z|U) [\overline{W}^Z \times Q_{VU}] - \underline{H}_{\log}(I^{c,*}) + R_c - R_0 = E_-^I. \end{aligned} \quad (206)$$

with $\rho_n = \frac{\delta \log n}{n}$ for any $\delta > 0$. Further, when $S_{I^{c,*},n}$ is the uniform random number and independent of $S_{I,n}$ and $S_{0,n}$, we have

$$\varepsilon_{n,\rho,I}(\overline{W}^{Z,n}, Q_{V,U}^n) = \varepsilon_{1,\rho,I}(\overline{W}^Z, Q_{V,U})^n \quad (207)$$

and

$$\lim_{\rho \rightarrow 0} \frac{[\log \varepsilon_{1,\rho,I}(\overline{W}^Z, Q_{V,U})]_+}{\rho} = I(V; Z|U) - R_p + \sum_{i \in \mathcal{I}} R_i. \quad (208)$$

The convergence in (208) is uniform.

Proof: First, we show (204). For arbitrary $u \in \mathcal{U}$ and $v \in \mathcal{V}$, the distribution $P_{\text{mix}, T_n(Q_{VU})}$ satisfies

$$P_{V^n|U^n, \text{mix}, T_n(Q_{VU})}(v|u) \leq (n+1)^{|\mathcal{U} \times \mathcal{V}|} Q_{V|U}^n(v|u) \quad (209)$$

by [8, Lemma 2.5, Chapter 1], and

$$P_{\text{mix},T_n(Q_U)}(u) \leq (n+1)^{|\mathcal{U}|} Q_U^n(u), \quad (210)$$

by [8, Lemma 2.3, Chapter 1]. Then, due to the relation (209), and (210), Lemma 3 with $C_1 = (n+1)^{|\mathcal{U}|^2|\mathcal{V}|}$ yields the relation (204).

Next, we show (205). We can also show that

$$\begin{aligned} & \lambda_{n,W}(\Theta_n) e^{E_0(\rho|\Theta_n, P_{V^n|U^n, \text{mix}, T_n(Q_{VU})}, P_{\text{mix}, T_n(Q_U)})} \\ &= \sum_u P_{\text{mix}, T_n(Q_U)}(u) \sum_z \left(\sum_v P_{V^n|U^n, \text{mix}, T_n(Q_{VU})}(v|u) \right. \\ & \quad \left. \cdot \left(\lambda_{n,W}(\Theta_n) \Theta_n(z|v) \right)^{\frac{1}{1-\rho}} \right)^{1-\rho} \\ & \leq \sum_u P_{\text{mix}, T_n(Q_U)}(u) \sum_z \left(\sum_v P_{V^n|U^n, \text{mix}, T_n(Q_{VU})}(v|u) \right. \\ & \quad \left. \cdot \left(\sum_{\Theta'_n \in \mathcal{W}_{n,Z}(Q_V)} \lambda_{n,W}(\Theta'_n) \Theta'_n(z|v) \right)^{\frac{1}{1-\rho}} \right)^{1-\rho} \\ & = e^{E_0(\rho|\overline{W}^Z, P_{V^n|U^n, \text{mix}, T_n(Q_{VU})}, P_{\text{mix}, T_n(Q_U)})}. \end{aligned} \quad (211)$$

Combining (204) and (211), we obtain

$$\begin{aligned} & (n+1)^{|\mathcal{U}|^2|\mathcal{V}|} e^{E_0(\rho|\overline{W}^Z, Q_{VU}^n, Q_U^n)} \\ & \geq \lambda_{n,W}(\Theta_n) e^{E_0(\rho|\Theta_n, P_{V^n|U^n, \text{mix}, T_n(Q_{VU})}, P_{\text{mix}, T_n(Q_U)})}. \end{aligned} \quad (212)$$

Due to the definition of $\varepsilon_{n,\rho}(W^Z, Q_{V^n}, U^n)$, the relation (212) is equivalent with the relation (205).

By using (16), the relation (206) can be shown as follows.

$$\begin{aligned} & \limsup_{n \rightarrow \infty} \frac{1}{n\rho_n} \log \varepsilon_{n,\rho_n, I}(\overline{W}^Z, Q_{V,U}^n) \\ &= \limsup_{n \rightarrow \infty} \left[(R_c - R_0) - \frac{1}{n} H_{1+\frac{\delta \log n}{n}}(S_{I^{c,*},n}|S_{I,n}, S_{0,n}) \right. \\ & \quad \left. + \frac{1}{\rho_n} E_0(\rho_n|\overline{W}^Z, Q_{VU}, Q_U) \right] \\ & \leq R_c - R_0 - \underline{H}_{\log}(I^{c,*}) + I(V; Z|U) = E_-^I. \end{aligned}$$

The relations (207) and (208) are trivial. \blacksquare

Step (4): Evaluation of the leaked information when the channel is given by the uniform distribution on a fixed conditional type:

Recall the fixed code $\varphi_{p,n}$ for BCD given in Theorem 13. The message sets of the code $\varphi_{p,n}$ are $\mathcal{S}_{0,n} \times \mathcal{B}_{1,n}$ and $\mathcal{B}_{2,n}$ with $|\mathcal{B}_{1,n}| = e^{n(R_c - R_0)}$ and $|\mathcal{B}_{2,n}| = e^{nR_p}$. We attach the other random coding $\Lambda_{F,G,n}$ for message $S_{1,n}, \dots, S_{T,n}$ given as Second Step of Code Ensemble 3 in Subsection VII-C to the code $\varphi_{p,n}$. That is, the encoder is given by $\Phi_{a,n} = (\varphi_{p,n}, \Lambda_{F,G,n})$. In the following, Bob's decoder $\Phi_{b,n}$ and Eve's decoder $\Phi_{e,n}$ are given as the maximum mutual information decoder. We treat the ensemble of codes $\Phi_n := (\Phi_{a,n}, \Phi_{b,n}, \Phi_{e,n})$.

First, related to the decomposition (193), we focus on a fixed arbitrary element $W^Z \in \mathcal{T}_{n,Z}(Q_{VU})$. We recall the discussion in Subsection VII-D. As is mentioned in Remark 25, the discussion in Section VII can be applied the channel W^Z , whose output distribution depends on the element of \mathcal{U} as well as the element of \mathcal{V} . Then, we apply Lemma 24 to the case when $P_{Z|V} = W^Z$, \mathcal{G} is the n -th permutation group, $(\mathcal{U} \times \mathcal{V})_o$ is

$T_n(Q_{UV})$, and $P_{V|U}$ is $\Upsilon_n(W^Z)$. Note that the n -th permutation group acts on $T_n(Q_{UV})$ transitively. We obtain

$$\begin{aligned} & e^{\psi(\rho|P_{Z^n|B_1, B_2, S_{0,n}} \cdot P_{\text{mix}, B_1, B_2})} \\ &= e^{\psi(\rho|\Upsilon_n(W^Z), P_{V^n|U^n, \text{mix}, \text{Im } \varphi_p}, P_{U, \text{mix}, \text{Im } \varphi_p})} \\ & \leq e^{n\rho(R_c - R_0) + E_0(\rho|\Upsilon_n(W^Z), P_{V^n|U^n, \text{mix}, T_n(Q_{VU})}, P_{\text{mix}, T_n(Q_U)})}. \end{aligned}$$

Combining Lemma 21 and the above inequality, we obtain

$$\begin{aligned} & \mathbf{E}_{\Phi_{a,n}} \exp(\rho I(S_{I,n}; Z^n | S_{0,n}) [\Upsilon_n(W^Z), \Phi_{a,n}, P_{S_{T+\infty,n}}]) \\ & \leq 1 + e^{n\rho(R_c - R_0) - \rho H_{1+\rho}(S_{I^{c,*},n}|S_{I,n}, S_{0,n})} e^{E_0(\rho|\Upsilon_n(W^Z), P_{V^n|U^n, \text{mix}, T_n(Q_{VU})}, P_{\text{mix}, T_n(Q_U)})}. \end{aligned} \quad (213)$$

Hence, we obtain the following relations. In the following derivation, the first inequality follows from the convexity of $x \mapsto e^x$. The third inequality follows from (200).

$$\begin{aligned} & \exp(\rho \mathbf{E}_{\Phi_{a,n}} I(S_{I,n}; Z^n | S_{0,n}) [\Upsilon_n(W^Z), \Phi_{a,n}, P_{S_{T+\infty,n}}]) \\ & \leq \mathbf{E}_{\Phi_{a,n}} \exp(\rho I(S_{I,n}; Z^n | S_{0,n}) [\Upsilon_n(W^Z), \Phi_{a,n}, P_{S_{T+\infty,n}}]) \\ & \leq 1 + e^{n\rho(R_c - R_0) - \rho H_{1+\rho}(S_{I^{c,*},n}|S_{I,n}, S_{0,n})} e^{E_0(\rho|\Upsilon_n(W^Z), P_{V^n|U^n, \text{mix}, T_n(Q_{VU})}, P_{\text{mix}, T_n(Q_U)})} \\ & \leq 1 + |\mathcal{T}_{n,Z}(Q_{VU})|^\rho e^{n\rho(R_c - R_0) - \rho H_{1+\rho}(S_{I^{c,*},n}|S_{I,n}, S_{0,n})} e^{n\rho I(V; Z|U)[W^Z \times Q_{VU}]} \end{aligned}$$

for any $\rho \in (0, 1)$. Taking the limit $\rho \rightarrow 1 - 0$, we have

$$\begin{aligned} & \exp(\mathbf{E}_{\Phi_{a,n}} I(S_{I,n}; Z^n | S_{0,n}) [\Upsilon_n(W^Z), \Phi_{a,n}, P_{S_{T+\infty,n}}]) \\ & \leq 1 + |\mathcal{T}_{n,Z}(Q_{VU})| e^{n(R_c - R_0) - H_2(S_{I^{c,*},n}|S_{I,n}, S_{0,n})} e^{nI(V; Z|U)[W^Z \times Q_{VU}]}. \end{aligned} \quad (214)$$

Since $\log(1+x) \leq x$, taking the logarithm in (214), we have

$$\begin{aligned} & \mathbf{E}_{\Phi_{a,n}} I(S_{I,n}; Z^n | S_{0,n}) [\Upsilon_n(W^Z), \Phi_{a,n}, P_{S_{T+\infty,n}}] \\ & \leq \log(1 + |\mathcal{T}_{n,Z}(Q_{VU})| e^{n(R_c - R_0) - H_2(S_{I^{c,*},n}|S_{I,n}, S_{0,n})}) e^{nI(V; Z|U)[W^Z \times Q_{VU}]} \\ & \leq |\mathcal{T}_{n,Z}(Q_{VU})| e^{n(R_c - R_0) - H_2(S_{I^{c,*},n}|S_{I,n}, S_{0,n})} e^{nI(V; Z|U)[W^Z \times Q_{VU}]}. \end{aligned}$$

Since $\log|\mathcal{Z}^n| = n \log|\mathcal{Z}| \leq |\mathcal{T}_{n,Z}(Q_{VU})|$, we have

$$\mathbf{E}_{\Phi_{a,n}} I(S_{I,n}; Z^n | S_{0,n}) [\Upsilon_n(W^Z), \Phi_{a,n}, P_{S_{T+\infty,n}}] \leq |\mathcal{T}_{n,Z}(Q_{VU})|. \quad (215)$$

Hence,

$$\begin{aligned} & \mathbf{E}_{\Phi_{a,n}} I(S_{I,n}; Z^n | S_{0,n}) [\Upsilon_n(W^Z), \Phi_{a,n}, P_{S_{T+\infty,n}}] \\ & \leq |\mathcal{T}_{n,Z}(Q_{VU})| e^{-[H_2(S_{I^{c,*},n}|S_{I,n}, S_{0,n}) - n(R_c - R_0) + I(V; Z|U)[W^Z \times Q_{VU}]]}. \end{aligned} \quad (216)$$

Next, related to the decomposition (201), we focus on a fixed arbitrary $\Theta_n \in \mathcal{W}_{n,Z}(Q_V)$. Similar to (213), Lemmas 21 and 24 yield that

$$\begin{aligned} & \mathbf{E}_{\Phi_{a,n}} \exp(\rho I(S_{I,n}; Z^n | S_{0,n}) [\Theta_n, \Phi_{a,n}, P_{S_{T+\infty,n}}]) \\ & \leq 1 + e^{n\rho(R_c - R_0) - \rho H_{1+\rho}(S_{I^{c,*},n}|S_{I,n}, S_{0,n})} e^{E_0(\rho|\Theta_n, P_{V^n|U^n, \text{mix}, T_n(Q_{VU})}, P_{\text{mix}, T_n(Q_U)})} \\ & = 1 + \varepsilon_{n,\rho, I}(\Theta_n, P_{\text{mix}, T_n(Q_{VU})}). \end{aligned} \quad (217)$$

Observe that we have shown that the averages over $\Phi_{a,n}$ of $\exp(\rho I(S_{I,n}; Z^n | S_{0,n}) [\Upsilon_n(W^Z), \Phi_{a,n}, P_{S_{T+\infty,n}}])$ and $I(S_{I,n}; Z^n | S_{0,n}) [\Theta_n, \Phi_{a,n}, P_{S_{T+\infty,n}}]$ are smaller than (216) and (217), respectively.

Choosing $p_1(n) := 2^T (|\mathcal{T}_{n,Z}(Q_{VU})| + |\mathcal{W}_{n,Z}(Q_V)|) + 1$, thanks to the Markov inequality in the same as (35) and (36), given a

fixed $\rho \in (0, 1)$, we can see that there exists at least one code φ_n such that the relations

$$\begin{aligned} & I(S_{I,n}; Z^n | S_{0,n}) [\Upsilon_n(W^Z), \varphi_{a,n}, P_{S_{\mathcal{T}+\infty,n}}] \\ & \leq p_1(n) \mathbf{E}_{\Phi_{a,n}} I(S_{I,n}; Z^n | S_{0,n}) [\Upsilon_n(W^Z), \Phi_{a,n}, P_{S_{\mathcal{T}+\infty,n}}] \\ & \leq p_1(n) |\mathcal{T}_{n,Z}(Q_{VU})| e^{n(R_c - R_0) - H_2(S_{I^{c,*},n} | S_{I,n}, S_{0,n})} e^{nI(V; Z|U)[W^Z \times Q_{VU}]} \end{aligned} \quad (218)$$

$$\begin{aligned} & \exp(\rho I(S_{I,n}; Z^n | S_{0,n}) [\Theta_n, \varphi_{a,n}, P_{S_{\mathcal{T}+\infty,n}}]) \\ & \leq p_1(n) \mathbf{E}_{\Phi_{a,n}} \exp(\rho I(S_{I,n}; Z^n | S_{0,n}) [\Theta_n, \Phi_{a,n}, P_{S_{\mathcal{T}+\infty,n}}]) \\ & \leq p_1(n) (1 + \varepsilon_{n,\rho,I}(\Theta_n, P_{\text{mix},T_n(Q_{VU})})). \end{aligned} \quad (219)$$

hold for any $W^Z \in \mathcal{T}_{n,Z}(Q_{VU})$ and $\Theta_n \in \mathcal{W}_{n,Z}(Q_V)$.

Step (5): Evaluation of the leaked information when the channel is given by discrete memoryless channel:

Using (218), we obtain

$$\begin{aligned} & I(S_{I,n}; Z^n | S_{0,n}) [\overline{W}^{Z,n}, \varphi_{a,n}, P_{S_{\mathcal{T}+\infty,n}}] \\ & \leq \sum_{W^Z \in \mathcal{T}_{n,Z}(Q_{VU})} \lambda_{n,T}(W^Z) I(S_{I,n}; Z^n | S_{0,n}) [\Upsilon_n(W^Z), \varphi_{a,n}, P_{S_{\mathcal{T}+\infty,n}}] \end{aligned} \quad (220)$$

$$\begin{aligned} & \leq \sum_{W^Z \in \mathcal{T}_{n,Z}(Q_{VU})} \left[\lambda_{n,T}(W^Z) p_1(n) |\mathcal{T}_{n,Z}(Q_{VU})| \right. \\ & \quad \left. \cdot e^{-[H_2(S_{I^{c,*},n} | S_{I,n}, S_{0,n}) - n(R_c - R_0) + I(V; Z|U)[W^Z \times Q_{VU}]]_+} \right] \end{aligned} \quad (221)$$

$$\begin{aligned} & \leq \sum_{W^Z \in \mathcal{T}_{n,Z}(Q_{VU})} \left[p_1(n) |\mathcal{T}_{n,Z}(Q_{VU})| \right. \\ & \quad \left. \cdot e^{-nD(W^Z | \overline{W}^Z | Q_{VU}) - [H_2(S_{I^{c,*},n} | S_{I,n}, S_{0,n}) - n(R_c - R_0) + I(V; Z|U)[W^Z \times Q_{VU}]]_+} \right] \end{aligned} \quad (222)$$

$$\leq \sum_{W^Z \in \mathcal{T}_{n,Z}(Q_{VU})} p_1(n) |\mathcal{T}_{n,Z}(Q_{VU})| e^{-K_n(\overline{W}^Z, Q_{VU}, R_c, R_0 | S)} \quad (223)$$

$$= p_1(n) |\mathcal{T}_{n,Z}(Q_{VU})|^2 e^{-K_n(\overline{W}^Z, Q_{VU}, R_c, R_0 | S)}, \quad (224)$$

where $K_n(\overline{W}^Z, Q_{VU}, R_c, R_0 | S)$ is defined as

$$\begin{aligned} & K_n(\overline{W}^Z, Q_{VU}, R_c, R_0 | S) \\ & := \min_{W^Z} \left[nD(W^Z | \overline{W}^Z | Q_{VU}) + [H_2(S_{I^{c,*},n} | S_{I,n}, S_{0,n}) \right. \\ & \quad \left. - n(R_c - R_0 + I(V; Z|U)[W^Z \times Q_{VU}]]_+ \right], \end{aligned}$$

and (220), (221), and (222) follow from (195), (218), and (194), respectively.

Hence,

$$\begin{aligned} & \liminf_{n \rightarrow \infty} \frac{-1}{n} \log I(S_{I,n}; Z^n | S_{0,n}) [\overline{W}^{Z,n}, \varphi_{a,n}, P_{S_{\mathcal{T}+\infty,n}}] \\ & \geq \liminf_{n \rightarrow \infty} \frac{1}{n} \min_{W^Z} \left[nD(W^Z | \overline{W}^Z | Q_{VU}) + [H_2(S_{I^{c,*},n} | S_{I,n}, S_{0,n}) \right. \\ & \quad \left. - n(R_c - R_0 + I(V; Z|U)[W^Z \times Q_{VU}]]_+ \right] \\ & = \min_{W^Z} \left[D(W^Z | \overline{W}^Z | Q_{VU}) \right. \\ & \quad \left. + [H_2(I^{c,*}) - R_c + R_0 - I(V; Z|U)[W^Z \times Q_{VU}]]_+ \right] \\ & = E_+^I \end{aligned} \quad (225)$$

Next, defining

$$p_2(n) := p_1(n)(n+1)^{|\mathcal{U}^{[2]V}|} |\mathcal{W}_{n,Z}(Q_V)|, \quad (226)$$

we obtain the following inequalities, in which, the first, second, and third inequalities follow from the convexity of function $x \mapsto \exp(x)$ and (202), (219), and (205), respectively. The final equation follows from (226).

$$\begin{aligned} & \exp(\rho I(S_{I,n}; Z^n | S_{0,n}) [\overline{W}^{Z,n}, \varphi_{a,n}, P_{S_{\mathcal{T}+\infty,n}}]) \\ & \leq \sum_{\Theta_n \in \mathcal{W}_{n,Z}(Q_V)} \lambda_{n,W}(\Theta_n) \exp(\rho I(S_{I,n}; Z^n | S_{0,n}) [\overline{W}_n, \varphi_{a,n}, P_{S_{\mathcal{T}+\infty,n}}]) \\ & \leq \sum_{\Theta_n \in \mathcal{W}_{n,Z}(Q_V)} \lambda_{n,W}(\Theta_n) p_1(n) (1 + \varepsilon_{n,\rho,I}(\Theta_n, P_{\text{mix},T_n(Q_{VU})})) \\ & \leq \sum_{\Theta_n \in \mathcal{W}_{n,Z}(Q_V)} p_1(n)(n+1)^{|\mathcal{U}^{[2]V}|} (1 + \varepsilon_{n,\rho,I}(\overline{W}^{Z,n}, Q_{V,U})) \\ & = p_1(n) |\mathcal{W}_{n,Z}(Q_V)| (n+1)^{|\mathcal{U}^{[2]V}|} (1 + \varepsilon_{n,\rho,I}(\overline{W}^{Z,n}, Q_{V,U})) \\ & = p_2(n) (1 + \varepsilon_{n,\rho,I}(\overline{W}^{Z,n}, Q_{V,U})). \end{aligned} \quad (227)$$

Taking the logarithm, we have

$$\begin{aligned} & I(S_{I,n}; Z^n | S_{0,n}) [\overline{W}^{Z,n}, \varphi_{a,n}, P_{S_{\mathcal{T}+\infty,n}}] \\ & \leq \frac{\log p_2(n) (1 + \varepsilon_{n,\rho,I}(\overline{W}^{Z,n}, Q_{V,U}))}{\rho} \\ & \leq \frac{\log(2p_2(n))}{\rho} + \frac{[\log \varepsilon_{n,\rho,I}(\overline{W}^{Z,n}, Q_{V,U})]_+}{\rho}. \end{aligned} \quad (228)$$

Now, we have

$$\lim_{n \rightarrow \infty} \frac{\log(2p_2(n))}{n \cdot \frac{\delta \log n}{n}} = \lim_{n \rightarrow \infty} \frac{\log(2p_2(n))}{\delta \log n} = \frac{\deg(p_2)}{\delta}, \quad (229)$$

where $\deg(p_2)$ is the degree of the polynomial p_2 . Due to (206) in Lemma 56, (228), and (229), choosing $\rho_n = \frac{\delta \log n}{n}$, we obtain

$$\limsup_{n \rightarrow \infty} \frac{1}{n} I(S_{I,n}; Z^n | S_{0,n}) [\overline{W}^{Z,n}, \varphi_{a,n}, P_{S_{\mathcal{T}+\infty,n}}] \leq \frac{\deg(p_2)}{\delta} + E_-^I.$$

Since $\delta > 0$ is arbitrary, we have

$$\limsup_{n \rightarrow \infty} \frac{1}{n} I(S_{I,n}; Z^n | S_{0,n}) [\overline{W}^{Z,n}, \varphi_{a,n}, P_{S_{\mathcal{T}+\infty,n}}] \leq E_-^I. \quad (230)$$

Therefore, using (225) and (230), we can see that (E^b, E^e, E_+^I, E_-^I) is a universally attainable quadruple of exponents in the sense of Definition 53. ■

Remark 57: One might consider that if we apply the random coding of Theorem 20 to the uniform distribution $P_{\text{mix},T_n(Q_{VU})}$, we obtain a better exponent. However, this method yields the same exponent because $\psi(\rho | \Upsilon_n(W^Z), P_{V^n|U^n, \text{mix}, T_n(Q_{VU})}, P_{\text{mix}, T_n(Q_{VU})})$ is the same as $E_0(\rho | \Upsilon_n(W^Z), P_{V^n|U^n, \text{mix}, T_n(Q_{VU})}, P_{\text{mix}, T_n(Q_{VU})})$, which is shown as

$$\begin{aligned} & e^{\psi(\rho | \Upsilon_n(W^Z), P_{V^n|U^n, \text{mix}, T_n(Q_{VU})}, P_{\text{mix}, T_n(Q_{VU})})} \\ & = \sum_{u \in \mathcal{T}_n(Q_U)} \frac{1}{N(U)} \sum_{v \in \mathcal{T}_n(Q_{V|U=u})} \left[\frac{1}{N(V|U)} \sum_{z \in \mathcal{T}_n(Q_{Z|VU=(u,v)})} \left(\frac{1}{N(Z|VU)} \right)^{1+\rho} \left(\frac{1}{N(Z|U)} \right)^{-\rho} \right] \\ & = \frac{N(Z|U)^\rho}{N(Z|VU)^\rho}. \end{aligned}$$

XIII. SOURCE-CHANNEL UNIVERSAL CODING FOR BCC

Now, we introduce the concept of ‘‘source-channel universal code for BCC’’ for the n -fold discrete memoryless extension of a discrete channel. In a realistic setting, we do not have statistical knowledge of the sources and the channel, precisely. In order to treat such a case, we have to make a code whose performance is guaranteed independently of the statistical properties of the sources and the channel. Such a kind of universality is called source-channel universality, and studied for the case of BCD [24]. For the case of wire-tap channel, the source universality is divided into two parts. One is the source universality for decoding error probability and the other is that for the leaked information. The paper [26] studied the latter part. Although the transmission rates are characterized by the pair (R_0, R_1) , in order to make a code achieving the capacity region of BCC, we employ other two parameters R_c and R_p that satisfy $R_0 \leq R_c$ and $R_0 + R_1 \leq R_c + R_p$. Hence, in the following definition of a universally attainable quadruple of exponents and leaked information rate, we focus on the set $\mathbf{R}_{\text{BCC}}^4 := \{(R_p, R_c, R_0, R_1) \in (\mathbf{R}^+)^4 | R_0 \leq R_c, R_0 + R_1 \leq R_c + R_p\}$.

Definition 58: A set of functions (E^b, E^e, E_+, E_-) from $\mathbf{R}_{\text{BCC}}^4 \times \mathcal{W}(\mathcal{X}, \mathcal{Y} \times \mathcal{Z})$ to $\mathbf{R}_{\geq 0}^4$ is said to be a universally attainable quadruple of exponents and leaked information rate for the family of channels $\mathcal{W}(\mathcal{X}, \mathcal{Y} \times \mathcal{Z})$ and for sources if for $\epsilon > 0$ and $(R_p, R_c, R_0, R_1) \in \mathbf{R}_{\text{BCC}}^4$, there exist a sufficiently large integer N and a sequence of codes Φ_n of length n satisfying the following conditions. (1) The confidential message set \mathcal{S}_n of the code Φ_n has cardinality e^{nR_1} and the common message set \mathcal{E}_n of the code Φ_n has cardinality e^{nR_0} . (2) The inequalities

$$P_b[W^n, \Phi_n, P_{\mathcal{S}_n, \mathcal{E}_n}] \leq \exp(-n[E^b(R_p, R_c, R_0, R_1, W) - \epsilon]), \quad (231)$$

$$P_e[W^n, \Phi_n, P_{\mathcal{S}_n, \mathcal{E}_n}] \leq \exp(-n[E^e(R_p, R_c, R_0, R_1, W) - \epsilon]), \quad (232)$$

and

$$\begin{aligned} & I(\mathcal{S}_n; Z^n | \mathcal{E}_n)[W^n, \Phi_n, P_{\mathcal{S}_n, \mathcal{E}_n}] \\ & \leq \max \left[\exp(-n[E_+^l(R_p, R_c, R_0, R_1, W) - \epsilon]), \right. \\ & \quad \left. n[E_-^l(R_p, R_c, R_0, R_1, W) + \epsilon] \right] \quad (233) \end{aligned}$$

hold for any sequence of joint distributions $P_{\mathcal{S}_n, \mathcal{E}_n}$ for the confidential message \mathcal{S}_n on \mathcal{S}_n and the common message \mathcal{E}_n on \mathcal{E}_n , and the n -th memoryless extension W^n of any channel $W \in \mathcal{W}(\mathcal{X}, \mathcal{Y} \times \mathcal{Z})$ and $n \geq N$.

Then, given a distribution Q_{VU} on $\mathcal{U} \times \mathcal{V}$ and a channel (probability transition matrix) $\Xi : \mathcal{V} \rightarrow \mathcal{X}$, we present a universally attainable quadruple of exponents and leaked information rate as follows. Given rates $(R_p, R_c, R_0, R_1) \in (\mathbf{R}^+)^4$ and a broadcast $W \in \mathcal{W}(\mathcal{X}, \mathcal{Y} \times \mathcal{Z})$, the quadruple E^b, E^e, E_+, E_-

and E_-^l are given as

$$E^b = E^b(R_p, R_c, R_0, R_1, W) := \tilde{E}^b(R_p, R_c, (W \circ \Xi) \times Q_{VU}), \quad (234)$$

$$E^e = E^e(R_p, R_c, R_0, R_1, W) := \tilde{E}^e(R_c, (W \circ \Xi) \circ Q_{VU}), \quad (235)$$

$$E_+^l = E_+^l(R_p, R_c, R_0, R_1, W) := \tilde{E}^l(R_p - R_1, (W \circ \Xi) \times Q_{VU}), \quad (236)$$

$$E_-^l = E_-^l(R_p, R_c, R_0, R_1, W) := I(V; Z|U) - R_p + R_1. \quad (237)$$

Theorem 59 (Extension of [24, Theorem 1, part (a)]):

Eqs. (234)–(237) are source-channel universally attainable rates of exponents and information leakage rate in the sense of Definition 58.

Therefore, our source-channel universal code attaining Eqs. (234)–(237) depends on R_p, R_c , the distribution Q_{VU} on $\mathcal{U} \times \mathcal{V}$, and the channel $\Xi : \mathcal{V} \rightarrow \mathcal{X}$.

We prove Theorem 59 by expurgating the messages in the code given in Theorem 54. The outline of the proof is as follows: First, in Step (1), similar to Theorem 54, we evaluate the leaked information when the channel is given by the conditional types and the source obeys the uniform distribution. Then, for a given code in Step (1), we expurgate the common message E_n in Step (2) and the secret message \mathcal{S}_n in Step (3). We evaluate the leaked information of the expurgated code for an arbitrary source distribution and an arbitrary conditional type in Step (4). Based on this evaluation, we evaluate the leaked information of the expurgated code for an arbitrary source distribution and an arbitrary discrete memoryless channel in Step (5).

In the following proof, we assume that the secret message \mathcal{S}_n and the common message E_n obey the uniform distributions on \mathcal{S}_n and \mathcal{E}_n . However, expurgations \mathcal{S}'_n and E'_n of the secret message \mathcal{S}_n and the common message E_n are allowed to obey arbitrary distributions.

Step (1): Evaluation of the leaked information when the channel is given as the uniform distribution on a fixed conditional type:

Recall the fixed code $\varphi_{p,n}$ for BCD given in Theorem 13. The code $\varphi_{p,n}$ has the private message set $\mathcal{S}_{0,n} \times \mathcal{B}_{1,n}$ and the common message set $\mathcal{B}_{2,n}$. We attach the random coding $\Lambda_{F,G,n}$ for message $S_{1,n}, \dots, S_{T,n}$ given as Second Step of Code Ensemble 3 in Subsection VII-C to the code $\varphi_{p,n}$ when $T = 2$, $S_{1,n} = \mathcal{S}_n$, $S_{0,n} = E_n$, and $S_{2,n}$ is the random number subject to the uniform distribution, which is used as the dummy for making \mathcal{S}_n secret for Eve. The uniformity of the distribution guarantees that

$$H_{1+\rho}(S_{2,n} | S_{1,n}, S_{0,n}) = n(R_c + R_p - R_1 - R_2) \quad (238)$$

for any $\rho \in (0, 1]$. Then, the encoder is given by $\Phi_{a,n} = (\varphi_{p,n}, \Lambda_{F,G,n})$. In the following, Bob’s decoder $\Phi_{b,n}$ and Eve’s decoder $\Phi_{e,n}$ are given as the maximum mutual information decoder. We treat the ensemble of codes $\Phi_n := (\Phi_{a,n}, \Phi_{b,n}, \Phi_{e,n})$.

For an arbitrary $\Theta_n \in \mathcal{W}_{n,\mathcal{Z}}(Q_V)$ and an arbitrary $\rho \in (0, 1)$, the combination of Lemmas 21 and 24 yields that

$$\begin{aligned} & \mathbf{E}_{\Phi_{a,n}} \sum_e P_{E_n}(e) \sum_s P_{S_n|E_n}(s|e) \\ & \quad \cdot \exp(\rho D(P_{Z^n|S_n=s, E_n=e, \Phi_{a,n}} \| P_{Z^n|E_n=e, \Phi_{a,n}})[\Theta_n]) \\ & \leq 1 + e^{n\rho(R_1 - R_p)} e^{E_0(\Theta_n, P_{V^n|U^n, \text{mix}, T_n(Q_VU)}, P_{\text{mix}, T_n(Q_V)})} \\ & = 1 + \varepsilon_{n,\rho,\{1\}}(\Theta_n, P_{\text{mix}, T_n(Q_VU)}), \end{aligned} \quad (239)$$

where $D(P_{Z^n|S_n=s, E_n=e, \varphi_{a,n}} \| P_{Z^n|E_n=e, \varphi_{a,n}})[\Theta_n]$ denotes the relative entropy $D(P_{Z^n|S_n=s, E_n=e, \varphi_{a,n}} \| P_{Z^n|E_n=e, \varphi_{a,n}})$ when the channel is $\Theta_n \in \mathcal{W}_{n,\mathcal{Z}}(Q_V)$.

The relations (238) and (216) with $T = 2$ yield

$$\begin{aligned} & \mathbf{E}_{\Phi_{a,n}} I(S_{I,n}; Z^n | S_{0,n})[\Upsilon_n(W^Z), \Phi_{a,n}, P_{S_{T,n}}] \\ & \leq |\mathcal{T}_{n,\mathcal{Z}}(Q_VU)| e^{-n[R_p - R_1 - I(V; Z|U)[W^Z \times Q_VU]]_+}. \end{aligned} \quad (240)$$

Thanks to the Markov inequality in the same way as (35) and (36), given a fixed $\rho \in (0, 1)$, due to (239) and (240), we can see that there exists at least one code $\varphi_{a,n}$ such that the relations

$$\begin{aligned} & I(S_{I,n}; Z^n | S_{0,n})[\Upsilon_n(W^Z), \varphi_{a,n}, P_{S_{T,n}}] \\ & \leq p_1(n) |\mathcal{T}_{n,\mathcal{Z}}(Q_VU)| e^{-n[R_p - R_1 - I(V; Z|U)[W^Z \times Q_VU]]_+}, \quad (241) \\ & \sum_e P_{E_n}(e) \sum_s P_{S_n|E_n}(s|e) \\ & \quad \cdot \exp(\rho D(P_{Z^n|S_n=s, E_n=e, \varphi_{a,n}} \| P_{Z^n|E_n=e, \varphi_{a,n}})[\Theta_n]) \\ & \leq p_1(n) (1 + \varepsilon_{n,\rho,\{1\}}(\Theta_n, P_{\text{mix}, T_n(Q_VU)})) \end{aligned} \quad (242)$$

hold for any $W^Z \in \mathcal{T}_{n,\mathcal{Z}}(Q_VU)$ and $\Theta_n \in \mathcal{W}_{n,\mathcal{Z}}(Q_V)$.

Step (2): Expurgation for common message E_n :

We choose $p_3(n) := 2p_1(n)$. When e is randomly chosen from \mathcal{E}_n subject to the uniform distribution, the element e satisfies all of the following conditions at least with probability of $1 - p_1(n)/p_3(n) = \frac{1}{2}$. The relations

$$\begin{aligned} & \sum_s P_{S_n|E_n}(s|e) \exp(\rho D(P_{Z^n|S_n=s, E_n=e, \varphi_{a,n}} \| P_{Z^n|E_n=e, \varphi_{a,n}})[\Theta_n]) \\ & \leq p_1(n) p_3(n) (1 + \varepsilon_{n,\rho,\{1\}}(\Theta_n, P_{\text{mix}, T_n(Q_VU)})), \\ & \sum_s P_{S_n|E_n}(s|e) D(P_{Z^n|S_n=s, E_n=e, \varphi_{a,n}} \| P_{Z^n|E_n=e, \varphi_{a,n}})[\Upsilon_n(W^Z)] \\ & = I(S_n; Z^n)[\Upsilon_n(W^Z), \varphi_{a,n}, P_{\text{mix}, S_n|E_n=e}] \\ & \leq p_1(n) p_3(n) |\mathcal{T}_{n,\mathcal{Z}}(Q_VU)| e^{-n[R_p - R_1 - I(V; Z|U)[W^Z \times Q_VU]]_+} \end{aligned} \quad (243)$$

hold for any elements $W^Z \in \mathcal{T}_{n,\mathcal{Z}}(Q_VU)$ and $\Theta_n \in \mathcal{W}_{n,\mathcal{Z}}(Q_V)$, and $n \geq N$. Thus, there exist $|\mathcal{E}_n|/2$ elements $e \in \mathcal{E}_n$ satisfies the above conditions. So, we denote the set of such elements by \mathcal{E}'_n .

Step (3): Expurgation for secret message S_n :

Then, when s is randomly chosen from \mathcal{S}_n subject to the uniform distribution, the element s satisfies all of the following conditions at least with probability of $1 - p_1(n)/p_3(n) \geq \frac{1}{2}$: The relations

$$\begin{aligned} & \exp(\rho D(P_{Z^n|S_n=s, E_n=e', \varphi_{a,n}} \| P_{Z^n|E_n=e', \varphi_{a,n}})[\Theta_n]) \\ & \leq p_1(n) p_3(n)^2 (1 + \varepsilon_{n,\rho,\{1\}}(\Theta_n, P_{\text{mix}, T_n(Q_VU)})), \end{aligned} \quad (244)$$

$$\begin{aligned} & D(P_{Z^n|S_n=s, E_n=e', \varphi_{a,n}} \| P_{Z^n|E_n=e', \varphi_{a,n}})[\Upsilon_n(W^Z)] \\ & \leq p_1(n) p_3(n)^2 |\mathcal{T}_{n,\mathcal{Z}}(Q_VU)| e^{-n[R_p - R_1 - I(V; Z|U)[W^Z \times Q_VU]]_+} \end{aligned} \quad (245)$$

hold for any elements $e' \in \mathcal{E}'_n$, $W^Z \in \mathcal{T}_{n,\mathcal{Z}}(Q_VU)$, $\Theta_n \in \mathcal{W}_{n,\mathcal{Z}}(Q_V)$, and $n \geq N$. Thus, there exist $|\mathcal{S}_n|/2$ elements $s \in \mathcal{S}_n$ satisfies the above conditions. So, we denote the set of such elements by \mathcal{S}'_n .

Step (4): Universal code that works for all sources when the channel is given as the uniform distribution on a fixed conditional type:

In the following discussion, $P_{S'_n, E'_n}$ is an arbitrary joint distribution of the random variables S'_n and E'_n on $\mathcal{S}'_n \times \mathcal{E}'_n$. For a given $e \in \mathcal{E}'_n$, we consider two kinds of marginal distributions of Z^n as follows.

$$\begin{aligned} P_{Z^n|E'_n=e, \varphi_{a,n}} & = \sum_{s \in \mathcal{S}'_n} P_{S'_n}(s) P_{Z^n|S'_n=s, E'_n=e, \varphi_{a,n}} \\ P'_{Z^n|E'_n=e, \varphi_{a,n}} & := \sum_{s' \in \mathcal{S}'_n} P_{S'_n|E'_n}(s'|e) P_{Z^n|S'_n=s', E'_n=e, \varphi_{a,n}}. \end{aligned}$$

The former marginal distribution is discussed in Steps (1), (2), and (3). Hence, using (54) and (245), we obtain

$$\begin{aligned} & I(S'_n; Z^n | E'_n)[\Upsilon_n(W^Z), \varphi_{a,n}, P_{S'_n, E'_n}] \\ & = \sum_{e \in \mathcal{E}'_n} P_{E'_n}(e) D(P_{Z^n, S'_n|E'_n=e, \varphi_{a,n}} \| P'_{Z^n|E'_n=e, \varphi_{a,n}} \times P_{S'_n|E'_n=e})[\Upsilon_n(W^Z)] \\ & \leq \sum_{e \in \mathcal{E}'_n} P_{E'_n}(e) D(P_{Z^n, S'_n|E'_n=e, \varphi_{a,n}} \| P_{Z^n|E'_n=e, \varphi_{a,n}} \times P_{S'_n|E'_n=e})[\Upsilon_n(W^Z)] \\ & = \sum_{e \in \mathcal{E}'_n} P_{E'_n}(e) \sum_{s \in \mathcal{S}'_n} [P_{S'_n|E'_n}(s|e) \\ & \quad \cdot D(P_{Z^n|S'_n=s, E'_n=e, \varphi_{a,n}} \| P_{Z^n|E'_n=e, \varphi_{a,n}})[\Upsilon_n(W^Z)]] \\ & \leq p_1(n) p_3(n)^2 |\mathcal{T}_{n,\mathcal{Z}}(Q_VU)| e^{-n[R_p - R_1 - I(V; Z|U)[W^Z \times Q_VU]]_+}, \end{aligned} \quad (246)$$

for any elements $W^Z \in \mathcal{T}_{n,\mathcal{Z}}(Q_VU)$, $\Theta_n \in \mathcal{W}_{n,\mathcal{Z}}(Q_V)$, and $n \geq N$. Similarly, using the convexity of $x \mapsto e^x$, (54), (244), and (245), we obtain

$$\begin{aligned} & e^{\rho I(S'_n; Z^n | E'_n)[\Theta_n, \varphi_{a,n}, P_{S'_n, E'_n}]} \\ & \leq \sum_{e \in \mathcal{E}'_n} P_{E'_n}(e) e^{\rho D(P_{Z^n, S'_n|E'_n=e, \varphi_{a,n}} \| P'_{Z^n|E'_n=e, \varphi_{a,n}} \times P_{S'_n|E'_n=e})[\Theta_n]} \\ & \leq \sum_{e \in \mathcal{E}'_n} P_{E'_n}(e) e^{\rho D(P_{Z^n, S'_n|E'_n=e, \varphi_{a,n}} \| P_{Z^n|E'_n=e, \varphi_{a,n}} \times P_{S'_n|E'_n=e})[\Theta_n]} \\ & \leq \sum_{e \in \mathcal{E}'_n} P_{E'_n}(e) \sum_{s \in \mathcal{S}'_n} P_{S'_n|E'_n}(s|e) e^{\rho D(P_{Z^n|S'_n=s, E'_n=e, \varphi_{a,n}} \| P_{Z^n|E'_n=e, \varphi_{a,n}})[\Theta_n]} \\ & \leq p_1(n) p_3(n)^2 (1 + \varepsilon_{n,\rho,\{1\}}(\Theta_n, P_{\text{mix}, T_n(Q_VU)})) \end{aligned} \quad (247)$$

for any elements $W^Z \in \mathcal{T}_{n,\mathcal{Z}}(Q_VU)$, $\Theta_n \in \mathcal{W}_{n,\mathcal{Z}}(Q_V)$, and $n \geq N$.

Step (5): Evaluation of leaked information for all sources and all discrete memoryless channels:

Similar to (224) and (227), defining $p_4(n) := p_1(n) p_3(n)^2 |\mathcal{T}_{n,\mathcal{Z}}(Q_VU)|^2$ and $p_5(n) := p_2(n) p_3(n)^2$ and using (246) and (247), we obtain

$$I(S'_n; Z^n | E'_n)[\overline{W}^{Z,n}, \varphi_{a,n}, P_{S'_n, E'_n}] \leq p_4(n) e^{-nE_+^l(R_p, R_c, R_0, R_1, W)}, \quad (248)$$

and

$$\begin{aligned} & \exp(\rho I(S'_n; Z^n | E'_n) [\overline{W}^{Z,n}, \varphi_{a,n}, P_{S'_n, E'_n}]) \\ & \leq p_5(n) (1 + \varepsilon_{n,\rho, \{1\}}(\overline{W}^{Z,n}, Q_{V,U}^n)) \\ & = p_5(n) (1 + \varepsilon_{1,\rho, \{1\}}(\overline{W}^Z, Q_{V,U})^n) \end{aligned} \quad (249)$$

for any sequence of joint distributions $P_{S'_n, E'_n}$ and $n \geq N$.

Using (248), for an arbitrary $\epsilon > 0$, we can choose an integer N_1 such that

$$\begin{aligned} & \log I(S'_n; Z^n | E'_n) [\overline{W}^{Z,n}, \varphi_{a,n}, P_{S'_n, E'_n}] \\ & \leq -n(E_+^l(R_p, R_c, R_0, R_1, W) - \epsilon) \end{aligned} \quad (250)$$

for $n \geq N_1$. Due to (249), we obtain

$$\begin{aligned} & \frac{1}{n} I(S'_n; Z^n | E'_n) [\overline{W}^{Z,n}, \varphi_{a,n}, P_{S'_n, E'_n}] \\ & \leq \frac{\log p_5(n) + \log(1 + \varepsilon_{1,\rho, \{1\}}(\overline{W}^Z, Q_{V,U})^n)}{n\rho} \\ & \leq \frac{\log p_5(n) + \log 2 + \log \varepsilon_{1,\rho, \{1\}}(\overline{W}^Z, Q_{V,U})^n}{n\rho} \\ & \leq \frac{\log 2p_5(n)}{n\rho} + \frac{\log \varepsilon_{1,\rho, \{1\}}(\overline{W}^Z, Q_{V,U})}{\rho}. \end{aligned} \quad (251)$$

When $\rho = \frac{1}{\sqrt{n}}$, as is mentioned in Lemma 56, the RHS of (251) converges $E_-^l(R_p, R_c, R_0, R_1, W)$ uniformly. Hence, for an arbitrary $\epsilon > 0$, we can choose an integer N_2 such that

$$\begin{aligned} & I(S'_n; Z^n | E'_n) [\overline{W}^{Z,n}, \varphi_{a,n}, P_{S'_n, E'_n}] \\ & \leq n(E_-^l(R_p, R_c, R_0, R_1, W) + \epsilon) \end{aligned} \quad (252)$$

for $n \geq N_2$.

Therefore, since the original code $\varphi_{p,n}$ satisfies (39) and (40), using (250) and (252), we can see that (E^b, E^e, E_+^l, E_-^l) is a universally attainable quadruple of exponents in the sense of Definition 58. ■

Remark 60: In this section, we treat the leaked information asymptotically as (233). However, in Section XII, we have treated it non-asymptotically as (177) and (178). The difference is caused by the condition for the sequence of joint distributions $P_{S_{T,n}}$. In Section XII, we do not assume the uniformity. However, in this section, we can use uniform distribution of $S_{2,n}$. Hence, we can calculate the relative Rényi entropy as (238) non-asymptotically.

Remark 61: Here, we remark the relation with the discussion for secure multiplex coding in [22, Section IV-D]. The preceding paper [22] showed the existence of the code φ_n satisfying that

$$\max_s D(P_{Z^n | S_1 = s_1, \varphi_n} \| P_{Z^n, \varphi_n}) \rightarrow 0 \quad (253)$$

when there is no common message E_n and the random variables S_1, \dots, S_T obey the uniform distribution. However, to show the source universality for leaked information in secure multiplex coding we need to evaluate the above value when the random variables S_1, \dots, S_T do not necessarily obey the uniform distribution. In this section, we show the source universality for leaked information for S_1 by assuming the uniformity of the other random variable S_2 . Although this

method brings us the source universality for BCC, it cannot derive the source universality for secure multiplex coding.

XIV. COMPARISON OF EXPONENTS OF LEAKED INFORMATION

In this section, we compare the exponent of leaked information given in Sections XII and XIII and the exponents of leaked information given in Subsection X-B when the source distribution $P_{S_{T,n}}$ is uniform. First, in Subsection XIV-A, we compare the exponent given in Sections XII and XIII with the above mentioned exponent. Then, we clarify that the exponent in Sections XII and XIII is greater than one of exponents in Subsection X-B, which is the same as that in [19]. Next, in Subsection XIV-B, we give equality conditions between two exponents. In the remaining subsections, we give proofs of Lemmas used in Subsections XIV-A and XIV-B.

A. Comparison between Two Exponents $\tilde{E}^l(R, \overline{W}^Z \times Q_{V,U})$ and $\tilde{E}^{E_0}(R, \overline{W}^Z \times Q_{V,U})$

First, we characterize the exponent $\tilde{E}^{E_0}(R, \overline{W}^Z \times Q_{V,U}) = \sup_{\rho \in (0,1)} \rho R - E_0(\rho | \overline{W}^Z, Q_{V,U}, Q_U)$, which describes the exponent of leaked information when R is $R_p - \sum_{i \in \mathcal{I}} R_i$ and the source distribution $P_{S_{T,n}}$ is uniform, as is shown in Subsection X-B. The exponent can be attained by the code constructed in the second construction (Subsection VII-C). Since $E_0(\rho | \overline{W}^Z, Q_{V,U}, Q_U)$ is convex with respect to ρ [12], $F_\rho(Q_{V|U}, Q_U) := \frac{d}{d\rho} E_0(\rho | \overline{W}^Z, Q_{V|U}, Q_U)$ is monotonically increasing with respect to ρ . As limits, we define

$$F_1(Q_{V|U}, Q_U) := \lim_{\rho \rightarrow 1-0} F_\rho(Q_{V|U}, Q_U) \quad (254)$$

$$E_0(1 | \overline{W}^Z, Q_{V|U}, Q_U) := \lim_{\rho \rightarrow 1-0} E_0(\rho | \overline{W}^Z, Q_{V|U}, Q_U). \quad (255)$$

In particular, when $Q_{V,U}$ equal $Q_V \times Q_U$, $\tilde{E}^l(R, \overline{W}^Z \times Q_{V,U})$, $\tilde{E}^{E_0}(R, \overline{W}^Z \times Q_{V,U})$, and the above values depend only on Q_V . Then, $\tilde{E}^l(R, \overline{W}^Z \times Q_{V,U})$, $\tilde{E}^{E_0}(R, \overline{W}^Z \times Q_{V,U})$, $E_0(1 | \overline{W}^Z, Q_{V|U}, Q_U)$, $F_1(Q_{V|U}, Q_U)$, and $F_\rho(Q_{V|U}, Q_U)$ are simplified to $\tilde{E}^l(R, \overline{W}^Z \times Q_V)$, $\tilde{E}^{E_0}(R, \overline{W}^Z \times Q_V)$, $E_0(1 | \overline{W}^Z, Q_V)$, $F_1(Q_V)$, and $F_\rho(Q_V)$. Then, we obtain the following lemma.

Lemma 62: (1) Case of $R < F_1(Q_{V|U}, Q_U)$. There uniquely exists $\rho \in (0, 1)$ such that $R = F_\rho(Q_{V|U}, Q_U)$. Then, the exponent $\tilde{E}^{E_0}(R, \overline{W}^Z \times Q_{V,U})$ can be characterized as

$$\tilde{E}^{E_0}(R, \overline{W}^Z \times Q_{V,U}) = \rho R - E_0(\rho | \overline{W}^Z, Q_{V|U}, Q_U). \quad (256)$$

(2) Case of $R \geq F_1(Q_{V|U}, Q_U)$. The exponent $\tilde{E}^{E_0}(R, \overline{W}^Z \times Q_{V,U})$ can be characterized as

$$\tilde{E}^{E_0}(R, \overline{W}^Z \times Q_{V,U}) = R - E_0(1 | \overline{W}^Z, Q_{V|U}, Q_U). \quad (257)$$

The quantities appearing in Lemma 62 can be characterized by Lemma 63, which is displayed in the wide space in the next page.

The proof of Lemma 63 will be given in Subsection XIV-D. For a detail analysis for the exponent $\tilde{E}^{E_0}(R, \overline{W}^Z \times Q_{V,U})$, we

Lemma 63: The quantities $F_\rho(Q_{V|U}, Q_U)$, $F_1(Q_{V|U}, Q_U)$, and $E_0(1|\overline{W}^Z, Q_{V|U}, Q_U)$ are calculated as

$$F_\rho(Q_{V|U}, Q_U) = \frac{\sum_u Q_U(u) \sum_z (\sum_v \frac{1}{1-\rho} (\log \overline{W}^Z(z|v)) Q_{V|U}(v|u) \overline{W}^Z(z|v)^{\frac{1}{1-\rho}}) (\sum_v Q_{V|U}(v|u) \overline{W}^Z(z|v)^{\frac{1}{1-\rho}})^{-\rho}}{\sum_u Q_U(u) \sum_z (\sum_v Q_{V|U}(v|u) \overline{W}^Z(z|v)^{\frac{1}{1-\rho}})^{1-\rho}} - \frac{\sum_u Q_U(u) \sum_z \log(\sum_v Q_{V|U}(v|u) \overline{W}^Z(z|v)^{\frac{1}{1-\rho}}) (\sum_v Q_{V|U}(v|u) \overline{W}^Z(z|v)^{\frac{1}{1-\rho}})^{1-\rho}}{\sum_u Q_U(u) \sum_z (\sum_v Q_{V|U}(v|u) \overline{W}^Z(z|v)^{\frac{1}{1-\rho}})^{1-\rho}}. \quad (258)$$

$$F_1(Q_{V|U}, Q_U) = - \frac{\sum_u Q_U(u) \sum_z \log(\sum_{v \in \mathcal{V}_z} Q_{V|U}(v|u)) \max_{v'} \overline{W}^Z(z|v')}{\sum_z \max_{v'} \overline{W}^Z(z|v')} \quad (259)$$

$$E_0(1|\overline{W}^Z, Q_{V|U}, Q_U) = \log \sum_u Q_U(u) \sum_z \max_{v \in \text{supp}(Q_{V|U=u})} \overline{W}^Z(z|v). \quad (260)$$

In particular, $F_\rho(Q_V)$, $F_1(Q_V)$, and $E_0(1|\overline{W}^Z, Q_V)$ are simplified to

$$F_\rho(Q_V) = \frac{\sum_z (\sum_v \frac{1}{1-\rho} (\log \overline{W}^Z(z|v)) Q_V(v) \overline{W}^Z(z|v)^{\frac{1}{1-\rho}}) (\sum_v Q_V(v) \overline{W}^Z(z|v)^{\frac{1}{1-\rho}})^{-\rho}}{\sum_z (\sum_v Q_V(v) \overline{W}^Z(z|v)^{\frac{1}{1-\rho}})^{1-\rho}} - \frac{\sum_z \log(\sum_v Q_V(v) \overline{W}^Z(z|v)^{\frac{1}{1-\rho}}) (\sum_v Q_V(v) \overline{W}^Z(z|v)^{\frac{1}{1-\rho}})^{1-\rho}}{\sum_z (\sum_v Q_V(v) \overline{W}^Z(z|v)^{\frac{1}{1-\rho}})^{1-\rho}}. \quad (261)$$

$$= \frac{\sum_{z,v} (\frac{1}{1-\rho} (\log \overline{W}^Z(z|v)) - \log(\sum_{v''} Q_V(v'') \overline{W}^Z(z|v'')^{\frac{1}{1-\rho}})) (Q_V(v) \overline{W}^Z(z|v)^{\frac{1}{1-\rho}}) (\sum_{v'} Q_V(v') \overline{W}^Z(z|v')^{\frac{1}{1-\rho}})^{-\rho}}{\sum_z (\sum_v Q_V(v) \overline{W}^Z(z|v)^{\frac{1}{1-\rho}})^{1-\rho}}$$

$$F_1(Q_V) = - \frac{\sum_z \log(\sum_{v \in \mathcal{V}_z} Q_V(v)) \max_{v'} \overline{W}^Z(z|v')}{\sum_u Q_U(u) \sum_z \max_{v'} \overline{W}^Z(z|v')} \quad (262)$$

$$E_0(1|\overline{W}^Z, Q_V) = \log \sum_z \max_{v \in \text{supp}(Q_V)} \overline{W}^Z(z|v). \quad (263)$$

Further, the map $Q_V \mapsto F_1(Q_V)$ is concave.

define

$$F_\rho := \frac{d}{d\rho} E_{0,\max}(\rho|\overline{W}^Z), \quad F_1 := \lim_{\rho \rightarrow 1-0} F_\rho, \quad (264)$$

$$\mathcal{K} := \{(z, v) \in \mathcal{Z} \times \mathcal{V} | \overline{W}^Z(z|v) = \max_{v'} \overline{W}^Z(z|v')\}$$

$$\mathcal{Z}_v := \{z \in \mathcal{Z} | (z, v) \in \mathcal{K}\}, \quad \mathcal{V}_z := \{v \in \mathcal{V} | (z, v) \in \mathcal{K}\}. \quad (265)$$

Due to the compactness of the set $\mathcal{P}(\mathcal{U})$, we have

$$\lim_{\rho \rightarrow 1-0} \max_{Q'_V} E_0(1|\overline{W}^Z, Q'_V) = \max_{Q'_V} \lim_{\rho \rightarrow 1-0} E_0(1|\overline{W}^Z, Q'_V).$$

Hence, we obtain the following lemma for characterization of the quantity $E_{0,\max}(1|\overline{W}^Z)$ defined in (23).

Lemma 64: We have

$$E_{0,\max}(1|\overline{W}^Z) = \log \sum_z \max_v \overline{W}^Z(z|v) = \lim_{\rho \rightarrow 1-0} E_{0,\max}(\rho|\overline{W}^Z). \quad (266)$$

Then, we have the following characterization for a special case of Case (2) of Lemma 62.

Lemma 65: Assume that $\cup_{v \in \text{supp}(Q_u)} \mathcal{Z}_v = \mathcal{Z}$ for any $u \in \text{supp}(Q_U)$. When $R \geq F_1(Q_{V|U}, Q_U)$, we have

$$E_{0,\max}(1|\overline{W}^Z) = E_0(1|\overline{W}^Z, Q_{V|U}, Q_U) \quad (267)$$

and

$$\tilde{E}^{E_0}(R, \overline{W}^Z \times Q_{VU}) = R - E_{0,\max}(1|\overline{W}^Z). \quad (268)$$

The proof of Lemma 65 will be given in Subsection XIV-E.

For comparison between two exponential decreasing rates $\tilde{E}^{E_0}(R, \overline{W}^Z \times Q_{VU})$ and $\tilde{E}^l(R, \overline{W}^Z \times Q_{VU})$, we prepare the following lemma.

Lemma 66: Any channel $\overline{W}^Z \in \mathcal{W}(\mathcal{V}, \mathcal{Z})$ satisfies

$$\min_{w^Z \in \mathcal{W}(\mathcal{U} \times \mathcal{V}, \mathcal{Z})} D(W^Z \| \overline{W}^Z | Q_{VU}) - \rho I(V; Z|U)[W^Z \times Q_{VU}] \geq -E_0(\rho|\overline{W}^Z, Q_{V|U}, Q_U) \quad (269)$$

for any $\rho \in (0, 1)$.

The proof of Lemma 66 will be given in Subsection XIV-I. Since the inequalities

$$\begin{aligned} & \tilde{E}^l(R, \overline{W}^Z \times Q_{VU}) \\ &= \min_{w^Z \in \mathcal{W}(\mathcal{U} \times \mathcal{V}, \mathcal{Z})} D(W^Z \| \overline{W}^Z | Q_{VU}) + [R - I(V; Z|U)][W^Z \times Q_{VU}]_+ \\ &\geq \min_{w^Z \in \mathcal{W}(\mathcal{U} \times \mathcal{V}, \mathcal{Z})} D(W^Z \| \overline{W}^Z | Q_{VU}) + \rho [R - I(V; Z|U)][W^Z \times Q_{VU}]_+ \\ &\geq \min_{w^Z \in \mathcal{W}(\mathcal{U} \times \mathcal{V}, \mathcal{Z})} D(W^Z \| \overline{W}^Z | Q_{VU}) + \rho (R - I(V; Z|U))[W^Z \times Q_{VU}] \end{aligned}$$

hold for any $\rho \in (0, 1)$, we obtain the following theorem, which is (26).

Theorem 67:

$$\begin{aligned} & \tilde{E}^l(R, \overline{W}^Z \times Q_{VU}) \\ & \geq \sup_{\rho \in (0,1)} \rho R - E_0(\rho | \overline{W}^Z, Q_{V|U}, Q_U) = \tilde{E}^{E_0}(R, \overline{W}^Z \times Q_{VU}). \end{aligned} \quad (270)$$

B. Equality Conditions of (270)

In this subsection, we derive equality conditions of (270). For this purpose, we prepare two lemmas.

Lemma 68: For a fixed $\rho \in (0, 1)$, the following three conditions for a distribution Q_V are equivalent.

(i) The following value does not depend on $v \in \mathcal{V}$.

$$\sum_z \overline{W}^Z(z|v)^{\frac{1}{1-\rho}} \left(\sum_{v'} Q_V(v') \overline{W}^Z(z|v')^{\frac{1}{1-\rho}} \right)^{-\rho}$$

(ii) The following relation holds.

$$E_0(\rho | \overline{W}^Z, Q_V) = E_{0, \max}(\rho | \overline{W}^Z) = \max_{Q'_V} E_0(\rho | \overline{W}^Z, Q'_V). \quad (271)$$

(iii) The following relations hold for any $v \in \mathcal{V}$.

$$\begin{aligned} & \sum_z \overline{W}^Z(z|v)^{\frac{1}{1-\rho}} \left(\sum_{v'} Q_V(v') \overline{W}^Z(z|v')^{\frac{1}{1-\rho}} \right)^{-\rho} \\ & = \max_{Q'_V} \sum_z \left(\sum_{v'} Q'_V(v') \overline{W}^Z(z|v')^{\frac{1}{1-\rho}} \right)^{1-\rho} \\ & = \max_{Q'_V} e^{E_0(\rho | \overline{W}^Z, Q'_V)} = e^{E_{0, \max}(\rho | \overline{W}^Z)}. \end{aligned}$$

The proof of Lemma 68 will be given in Subsection XIV-F.

Lemma 69: The following three conditions for a distribution Q_V are equivalent.

(i) The following value does not depend on $v \in \mathcal{V}$.

$$\sum_{z \in \mathcal{Z}_v} \frac{\max_{v' \in \mathcal{V}} \overline{W}^Z(z|v')}{\sum_{v'' \in \mathcal{V}_z} Q_V(v'')} = \sum_{z \in \mathcal{Z}_v} \frac{\overline{W}^Z(z|v)}{\sum_{v'' \in \mathcal{V}_z} Q_V(v'')}.$$

(ii) The following relation holds.

$$F_1(Q_V) = \min_{Q'_V} F_1(Q'_V).$$

(iii) The following relations hold for any $v \in \mathcal{V}$.

$$\begin{aligned} & \sum_{z \in \mathcal{Z}_v} \frac{\max_{v' \in \mathcal{V}} \overline{W}^Z(z|v')}{\sum_{v'' \in \mathcal{V}_z} Q_V(v'')} = \sum_{z \in \mathcal{Z}_v} \frac{\overline{W}^Z(z|v)}{\sum_{v'' \in \mathcal{V}_z} Q_V(v'')} \\ & = \sum_z \max_{v'} \overline{W}^Z(z|v'). \end{aligned} \quad (272)$$

The proof of Lemma 68 will be given in Subsection XIV-G.

Then, we introduce two conditions for a distribution Q_V .

Condition 70: Given a fixed $\rho \in (0, 1)$, the distribution Q_V satisfies the condition given in Lemma 68

Condition 71: The distribution Q_V satisfies the condition given in Lemma 69

Since Condition 70 depends on ρ , we describe it by ‘‘Condition 70 with ρ ’’ when we need to clarify the dependence on ρ .

Lemma 72: When distribution Q_V and Q'_V satisfy Condition 70 with ρ , the relation $\sum_v Q_V(v) \overline{W}^Z(z|v)^{\frac{1}{1-\rho}} =$

$\sum_v Q'_V(v) \overline{W}^Z(z|v)^{\frac{1}{1-\rho}}$ holds for any $z \in \mathcal{Z}$. That is the value $\sum_v Q_V(v) \overline{W}^Z(z|v)^{\frac{1}{1-\rho}}$ does not depend on the choice of Q_V as long as the distribution Q_V satisfies Condition 70 with ρ . The proof of Lemma 72 will be given in Subsection XIV-F.

Lemma 73: When distribution Q_V and Q'_V satisfy Condition 71 with ρ , the relation $\sum_{v'' \in \mathcal{V}_z} Q_V(v'') = \sum_{v'' \in \mathcal{V}_z} Q'_V(v'')$ holds for any $z \in \mathcal{Z}$. That is the value $\sum_v Q_V(v) \overline{W}^Z(z|v)^{\frac{1}{1-\rho}}$ does not depend on the choice of Q_V as long as the distribution Q_V satisfies Condition 71.

The proof of Lemma 73 will be given in Subsection XIV-G. Hence, we can define the transition matrices $W^{Z,\rho}$ and $W^{Z,1}$ from \mathcal{V} to \mathcal{Z} by

$$\begin{aligned} W^{Z,\rho}(z|v) & := \frac{\overline{W}^Z(z|v)^{\frac{1}{1-\rho}} \left(\sum_{v'} Q_{V,\rho}(v') \overline{W}^Z(z|v')^{\frac{1}{1-\rho}} \right)^{-\rho}}{\sum_z \overline{W}^Z(z|v)^{\frac{1}{1-\rho}} \left(\sum_{v'} Q_{V,\rho}(v') \overline{W}^Z(z|v')^{\frac{1}{1-\rho}} \right)^{-\rho}}, \\ W^{Z,1}(z|v) & := \begin{cases} \frac{\overline{W}^Z(z|v)}{\sum_{v'' \in \mathcal{V}_z} Q_{V,1}(v'') \sum_{z'} \max_{v'} \overline{W}^Z(z'|v')} & z \in \mathcal{Z}_v \\ 0 & z \in \mathcal{Z}_v^c, \end{cases} \end{aligned}$$

where the distributions $Q_{V,\rho}$ and $Q_{V,1}$ satisfy Condition 70 with ρ and Condition 71, respectively. These definitions do not depend on the choices of $Q_{V,\rho}$ and $Q_{V,1}$.

Lemma 74: When $Q_{V,\rho}$ satisfies Condition 70 with ρ , we have

$$F_\rho = F_\rho(Q_{V,\rho}) = I(V; Z)[W^{Z,\rho} \times Q_{V,\rho}] \quad (273)$$

$$D(W^{Z,\rho} \| \overline{W}^Z | Q_{V,\rho}) = \rho F_\rho - E_{0, \max}(\rho | \overline{W}^Z). \quad (274)$$

The proof of Lemma 74 will be given in Subsection XIV-F.

Lemma 75: When $Q_{V,1}$ satisfies Condition 71, we have

$$F_1 = F_1(Q_{V,1}) = I(V; Z)[W^{Z,1} \times Q_{V,1}] \quad (275)$$

$$D(W^{Z,1} \| \overline{W}^Z | Q_{V,1}) = F_1 - E_{0, \max}(1 | \overline{W}^Z). \quad (276)$$

The proof of Lemma 75 will be given in Subsection XIV-G.

Lemma 76: For any $\rho \in (0, 1)$, we choose the distribution $Q_{V,\rho}$ satisfying Condition 70 with ρ . We choose a sequence ρ_n such that $\rho_n \rightarrow 0$ as $n \rightarrow \infty$ and the limit distribution $\lim_{n \rightarrow \infty} Q_{V,\rho_n}$ exists. (Since the set of distributions over \mathcal{V} is compact, such a sequence ρ_n exists.) Then, the limit distribution $\lim_{n \rightarrow \infty} Q_{V,\rho_n}$ satisfies Condition 71.

The proof of Lemma 76 will be given in Subsection XIV-H.

Then, using the above lemmas, we can characterize equality conditions of (270) for the case $Q_{UV} = Q_U \times Q_V$ in the following way.

Theorem 77: (1) Case of $R < F_1$. We choose $\rho \in (0, 1)$ such that $R = F_\rho$. When $Q_{V,\rho}$ satisfies Condition 70 with ρ , the relations

$$\begin{aligned} & \min_{Q_V} \tilde{E}^l(R, \overline{W}^Z \times Q_V) = \min_{Q_V} \tilde{E}^{E_0}(R, \overline{W}^Z \times Q_V) \\ & = \tilde{E}^l(R, \overline{W}^Z \times Q_{V,\rho}) = \tilde{E}^{E_0}(R, \overline{W}^Z \times Q_{V,\rho}) = \rho R - E_{0, \max}(\rho | \overline{W}^Z) \end{aligned} \quad (277)$$

hold, which implies the equality in (270).

(2) Case of $R \geq F_1$. When $Q_{V,1}$ satisfies Condition 71, the relations

$$\begin{aligned} & \min_{Q_V} \tilde{E}^l(R, \overline{W}^Z \times Q_V) = \min_{Q_V} \tilde{E}^{E_0}(R, \overline{W}^Z \times Q_V) \\ & = \tilde{E}^l(R, \overline{W}^Z \times Q_{V,1}) = \tilde{E}^{E_0}(R, \overline{W}^Z \times Q_{V,1}) = R - E_{0, \max}(1 | \overline{W}^Z) \end{aligned} \quad (278)$$

hold, which implies the equality in (270).

Combining the discussions in both cases in Theorem 77, we obtain

$$\begin{aligned} \min_{Q_V} \tilde{E}^l(R, \overline{W}^Z \times Q_V) &= \min_{Q_V} \tilde{E}^{E_0}(R, \overline{W}^Z \times Q_V) \\ &= \max_{\rho \in [0,1]} \rho R - E_{0,\max}(\rho|\overline{W}^Z), \end{aligned} \quad (279)$$

which is (27).

Proof of Theorem 77: First, we show (277). Since $I(V; Z)[W^{Z,\rho} \times Q_{V,\rho}] = F_\rho = R$ follows from (273), we have

$$\begin{aligned} &\tilde{E}^l(R, \overline{W}^Z \times Q_{V,\rho}) \\ &\stackrel{(a)}{\leq} D(W^{Z,\rho} \| \overline{W}^Z | Q_{V,\rho}) + [R - I(V; Z)[W^{Z,\rho} \times Q_{V,\rho}]]_+ \\ &\stackrel{(b)}{=} \rho F_\rho - E_{0,\max}(\rho|\overline{W}^Z) \stackrel{(c)}{=} \rho R - E_0(\rho|\overline{W}^Z, Q_{V,\rho}) \\ &\stackrel{(d)}{=} \tilde{E}^{E_0}(R, \overline{W}^Z \times Q_{V,\rho}), \end{aligned} \quad (280)$$

where (a), (b), (c), and (d) follow from the Definition (24) of $\tilde{E}^l(R, \overline{W}^Z \times Q_{V,\rho})$, (274), (271), and Item (1) of Lemma 62, respectively.

Any distribution Q_V satisfies

$$\rho R - E_{0,\max}(\rho|\overline{W}^Z) \leq \rho R - E_0(\rho|\overline{W}^Z, Q_V) \leq \tilde{E}^{E_0}(R, \overline{W}^Z \times Q_V),$$

which implies

$$\rho R - E_{0,\max}(\rho|\overline{W}^Z) \leq \min_{Q_V} \tilde{E}^{E_0}(R, \overline{W}^Z \times Q_V). \quad (281)$$

Combining the above relations and we obtain

$$\begin{aligned} &\tilde{E}^l(R, \overline{W}^Z \times Q_{V,\rho}) \stackrel{(a)}{\leq} \rho R - E_{0,\max}(\rho|\overline{W}^Z) \\ &\stackrel{(b)}{\leq} \min_{Q_V} \tilde{E}^{E_0}(R, \overline{W}^Z \times Q_V) \stackrel{(c)}{\leq} \min_{Q_V} \tilde{E}^l(R, \overline{W}^Z \times Q_V), \end{aligned} \quad (282)$$

where (a), (b), and (c) follow from (280), (281), and Theorem 67, respectively. Hence, the combination of (282) and (d) of (280) leads (277).

Next, we show (278). The relations (275) and (276) imply

$$\begin{aligned} &\tilde{E}^l(R, \overline{W}^Z \times Q_{V,1}) \\ &\leq D(W^{Z,1} \| \overline{W}^Z | Q_{V,1}) + [R - I(V; Z)[W^{Z,1} \times Q_{V,1}]]_+ \\ &= F_1 - E_{0,\max}(1|\overline{W}^Z) + [R - F_1]_+ \\ &= F_1 - E_{0,\max}(1|\overline{W}^Z) + R - F_1 = R - E_{0,\max}(1|\overline{W}^Z) \\ &= R - E_0(1|\overline{W}^Z, Q_{V,1}) = \tilde{E}^{E_0}(R, \overline{W}^Z \times Q_{V,1}). \end{aligned}$$

Any distribution Q_V satisfies

$$R - E_{0,\max}(1|\overline{W}^Z) \leq R - E_0(1|\overline{W}^Z, Q_V) \leq \tilde{E}^{E_0}(R, \overline{W}^Z \times Q_V),$$

which implies

$$R - E_{0,\max}(1|\overline{W}^Z) \leq \min_{Q_V} \tilde{E}^{E_0}(R, \overline{W}^Z \times Q_V).$$

Combining the above relations and Lemma 67, we obtain

$$\begin{aligned} &\tilde{E}^l(R, \overline{W}^Z \times Q_{V,\rho}) \leq R - E_{0,\max}(1|\overline{W}^Z) = \tilde{E}^{E_0}(R, \overline{W}^Z \times Q_{V,\rho}) \\ &\leq \min_{Q_V} \tilde{E}^{E_0}(R, \overline{W}^Z \times Q_V) \leq \min_{Q_V} \tilde{E}^l(R, \overline{W}^Z \times Q_V), \end{aligned}$$

which implies (278). \blacksquare

For the general case, we prepare the generalizations of Lemmas 74 and 75. The following lemmas follow from Lemmas 74 and 75.

Lemma 78: When $Q_{V|U=u}$ satisfies Condition 70 with ρ , for any $u \in \text{supp}(Q_U)$,

$$\begin{aligned} F_\rho &= F_\rho(Q_{V|U}, Q_U) = I(V; Z|U)[W^{Z,\rho} \times Q_{VU}] \\ D(W^{Z,\rho} \| \overline{W}^Z | Q_{VU}) &= F_\rho - E_{0,\max}(\rho|\overline{W}^Z). \end{aligned}$$

Lemma 79: When $Q_{V|U=u}$ satisfies Condition 71 for any $u \in \text{supp}(Q_U)$,

$$\begin{aligned} F_1 &= F_1(Q_{V|U}, Q_U) = I(V; Z|U)[W^{Z,1} \times Q_{VU}] \\ D(W^{Z,1} \| \overline{W}^Z | Q_{VU}) &= F_1 - E_{0,\max}(1|\overline{W}^Z). \end{aligned}$$

Then, we can characterize equality conditions for (270) in the general case. That is, similar to Theorem 77, using Lemmas 78 and 79, we can show the following theorem.

Theorem 80: (1) Case of $R < F_1$. We choose $\rho \in (0, 1)$ such that $R = F_\rho$. When $Q_{V|U=u}$ satisfies Condition 70 with ρ for any $u \in \text{supp}(Q_U)$, the relations

$$\begin{aligned} &\min_{Q'_{VU}} \tilde{E}^l(R, \overline{W}^Z \times Q'_{VU}) = \min_{Q'_V} \tilde{E}^l(R, \overline{W}^Z \times Q'_V) \\ &= \min_{Q'_{VU}} \tilde{E}^{E_0}(R, \overline{W}^Z \times Q'_{VU}) = \min_{Q'_V} \tilde{E}^{E_0}(R, \overline{W}^Z \times Q'_V) \\ &= \tilde{E}^l(R, \overline{W}^Z \times Q_{VU}) = \tilde{E}^{E_0}(R, \overline{W}^Z \times Q_{VU}) = \rho R - E_{0,\max}(\rho|\overline{W}^Z) \end{aligned} \quad (283)$$

hold, which implies the equality in (270).

(2) Case of $R \geq F_1$. When $Q_{V|U=u}$ satisfies Condition 71 for any $u \in \text{supp}(Q_U)$, the relations

$$\begin{aligned} &\min_{Q'_{VU}} \tilde{E}^l(R, \overline{W}^Z \times Q'_{VU}) = \min_{Q'_V} \tilde{E}^l(R, \overline{W}^Z \times Q'_V) \\ &= \min_{Q'_{VU}} \tilde{E}^{E_0}(R, \overline{W}^Z \times Q'_{VU}) = \min_{Q'_V} \tilde{E}^{E_0}(R, \overline{W}^Z \times Q'_V) \\ &= \tilde{E}^l(R, \overline{W}^Z \times Q_{VU}) = \tilde{E}^{E_0}(R, \overline{W}^Z \times Q_{VU}) = R - E_{0,\max}(1|\overline{W}^Z) \end{aligned} \quad (284)$$

hold, which implies the equality in (270).

Then, we obtain the following two corollaries.

Corollary 81: When the channel W^Z is regular and Q_V is the uniform distribution, the equality in (270) holds.

Proof: When the channel W^Z is regular, the uniform distribution over \mathcal{V} satisfies Condition 70 with ρ . Hence, when Q_V is the uniform distribution, the equality in (270) holds. \blacksquare

Corollary 82: When $R = F_\rho$ and $Q_{V|U=u}$ satisfies Condition 71 for any $u \in \text{supp}(Q_U)$, we have

$$\begin{aligned} &\tilde{E}^l(R, \overline{W}^Z \times Q_{VU}) = \tilde{E}^{E_0}(R, \overline{W}^Z \times Q_{VU}) \\ &\leq \tilde{E}^\psi(R, \overline{W}^Z \times Q_{VU}). \end{aligned}$$

In the above case of Corollary 82, the exponent $\tilde{E}^l(R, \overline{W}^Z \times Q_{VU})$ cannot improve the exponent $\tilde{E}^\psi(R, \overline{W}^Z \times Q_{VU})$, which is the exponent of the code constructed in the first construction (Subsection VII-B) and is given in Subsection X-B. However, the relation between $\tilde{E}^l(R, \overline{W}^Z \times Q_{VU})$ and $\tilde{E}^\psi(R, \overline{W}^Z \times Q_{VU})$ remains unknown up to now.

C. Examples

In this subsection, we numerically compare

$$\begin{aligned} & \tilde{E}^l(R, \bar{W}^Z \times Q_V) \\ &= \min_{W^Z \in \mathcal{W}(V, Z)} D(W^Z | \bar{W}^Z | Q_V) + [R - I(V; Z)[W^Z \times Q_V]]_+ \end{aligned}$$

and

$$\begin{aligned} \tilde{E}^{E_0}(R, \bar{W}^Z \times Q_V) &= \max_{0 \leq \rho \leq 1} \rho R - E_0(\rho | \bar{W}^Z, Q_V) \\ \tilde{E}^\psi(R, \bar{W}^Z \times Q_V) &= \max_{0 \leq \rho \leq 1} \rho R - \psi(\rho | \bar{W}^Z, Q_V) \end{aligned}$$

in the following two examples.

Example 83: In this example, we address the channel given by a 2×2 general transition matrix. Consider the case when $\mathcal{Z} = \mathcal{V} = \{1, 2\}$. Define the transition matrix \bar{W}^Z by

$$\bar{W}^Z := \begin{pmatrix} 1-p & q \\ p & 1-q \end{pmatrix} \quad (285)$$

with $p > q \in (0, 1/2)$. When $Q_V(1) = 1/2$ and $Q_V(2) = 1/2$, we have

$$\begin{aligned} & E_0(\rho | \bar{W}^Z, Q_V) \\ &= \log\left(\left(\frac{1}{2}(1-p)^{\frac{1}{1-\rho}} + \frac{1}{2}q^{\frac{1}{1-\rho}}\right)^{1-\rho} + \left(\frac{1}{2}p^{\frac{1}{1-\rho}} + \frac{1}{2}(1-q)^{\frac{1}{1-\rho}}\right)^{1-\rho}\right), \end{aligned} \quad (286)$$

$$\begin{aligned} & \psi(\rho | \bar{W}^Z, Q_V) \\ &= \log\left(\frac{1}{2}(1-p)^{1+\rho}\left(\frac{1-p+q}{2}\right)^{-\rho} + \frac{1}{2}p^{1+\rho}\left(\frac{1-q+p}{2}\right)^{-\rho}\right. \\ & \quad \left. + \frac{1}{2}q^{1+\rho}\left(\frac{1-p+q}{2}\right)^{-\rho} + \frac{1}{2}(1-q)^{1+\rho}\left(\frac{1-q+p}{2}\right)^{-\rho}\right). \end{aligned} \quad (287)$$

Fig. 2 suggests that $\tilde{E}^\psi(R, \bar{W}^Z \times Q_V)$ is larger than $\tilde{E}^l(R, \bar{W}^Z \times Q_V)$. In Fig. 3, we numerically calculate $\arg\max_{0 \leq \rho \leq 1} \rho R - E_0(\rho | \bar{W}^Z, Q_V)$ and $\arg\max_{0 \leq \rho \leq 1} \rho R - \psi(\rho | \bar{W}^Z, Q_V)$ which realize $\tilde{E}^{E_0}(R, \bar{W}^Z \times Q_V)$ and $\tilde{E}^\psi(R, \bar{W}^Z \times Q_V)$, respectively.

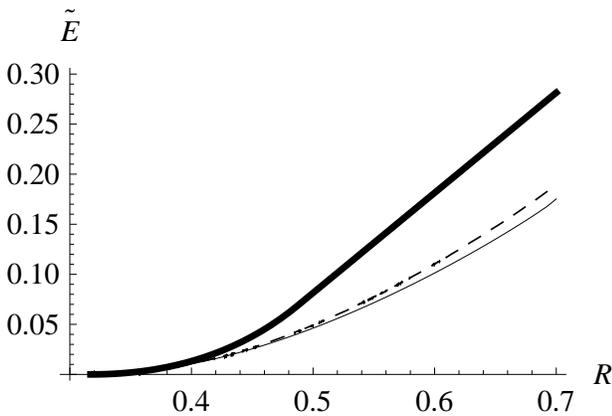


Fig. 2. Lower bounds of exponent in Example 83 with $p = 0.01$ and $q = 0.3$. In this case, $I(V; Z)[\bar{W}^Z \times Q_V] = 0.317054$. Thick line, Dashed line, and Normal line plot $\tilde{E}^\psi(R, \bar{W}^Z \times Q_V)$, $\tilde{E}^l(R, \bar{W}^Z \times Q_V)$, and $\tilde{E}^{E_0}(R, \bar{W}^Z \times Q_V)$ as functions of R from $R = 0.317054$ to $R = \log 2 = 0.693147$ with the origin $(0.3, 0)$.

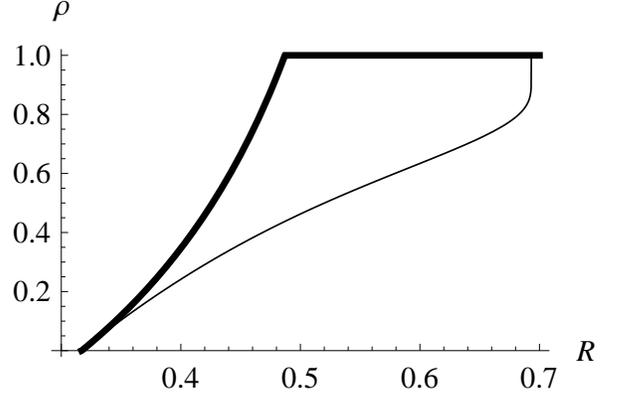


Fig. 3. Relation between R and ρ realizing the optimal value in Example 83 with $p = 0.01$ and $q = 0.3$. Thick line expresses $\arg\max_{0 \leq \rho \leq 1} \rho R - \psi(\rho | \bar{W}^Z, Q_V)$, which realizes $\tilde{E}^\psi(R, \bar{W}^Z \times Q_V)$. Normal line expresses $\arg\max_{0 \leq \rho \leq 1} \rho R - E_0(\rho | \bar{W}^Z, Q_V)$, which realizes $\tilde{E}^{E_0}(R, \bar{W}^Z \times Q_V)$. There is no graph corresponding to $\tilde{E}^l(R, \bar{W}^Z \times Q_V)$ because $\tilde{E}^l(R, \bar{W}^Z \times Q_V)$ is not given as maximization with respect to ρ . The origin is $(0.3, 0)$.

Example 84: In this example, we consider the case when states satisfying Conditions 70 and 71 are not unique. Consider the case when $\mathcal{Z} = \mathcal{V} = \{1, 2, 3, 4\}$. Define the transition matrix \bar{W}^Z by

$$\bar{W}^Z := \begin{pmatrix} \frac{1}{2}-p & p & \frac{1}{2}-p & p \\ p & \frac{1}{2}-p & p & \frac{1}{2}-p \\ \frac{1}{2}-p & p & p & \frac{1}{2}-p \\ p & \frac{1}{2}-p & \frac{1}{2}-p & p \end{pmatrix} \quad (288)$$

with $p \in (0, 1/4)$. When $Q_V(1) = q$, $Q_V(2) = q$, $Q_V(3) = \frac{1}{2} - q$, and $Q_V(4) = \frac{1}{2} - q$, we have

$$\begin{aligned} & \sum_z \bar{W}^Z(z|v)^{\frac{1}{1-\rho}} \left(\sum_{v'} Q_V(v') \bar{W}^Z(z|v')^{\frac{1}{1-\rho}} \right)^{-\rho} \\ &= 4 \left(\frac{1}{2} \left(\frac{1}{2} - p \right)^{\frac{1}{1-\rho}} + \frac{1}{2} p^{\frac{1}{1-\rho}} \right)^{1-\rho} = 2^{1+\rho} \left(\left(\frac{1}{2} - p \right)^{\frac{1}{1-\rho}} + p^{\frac{1}{1-\rho}} \right)^{1-\rho}. \end{aligned} \quad (289)$$

for all $v \in \mathcal{V}$, which implies Condition 70. Hence,

$$\begin{aligned} & E_{0, \max}(\rho | \bar{W}^Z) = E_0(\rho | \bar{W}^Z, Q_V) \\ &= (1 + \rho) \log 2 + (1 - \rho) \log \left(\left(\frac{1}{2} - p \right)^{\frac{1}{1-\rho}} + p^{\frac{1}{1-\rho}} \right), \end{aligned} \quad (290)$$

$$\begin{aligned} & F_\rho = F_\rho(Q_V) \\ &= \log 2 - \log \left(\left(\frac{1}{2} - p \right)^{\frac{1}{1-\rho}} + p^{\frac{1}{1-\rho}} \right) \\ & \quad + \frac{1}{1-\rho} \frac{\left(\frac{1}{2} - p \right)^{\frac{1}{1-\rho}} \log \left(\frac{1}{2} - p \right) + p^{\frac{1}{1-\rho}} \log p}{\left(\frac{1}{2} - p \right)^{\frac{1}{1-\rho}} + p^{\frac{1}{1-\rho}}}, \end{aligned} \quad (291)$$

$$\psi(\rho | \bar{W}^Z, Q_V) = (2\rho + 1) \log 2 + \log \left(\left(\frac{1}{2} - p \right)^{1+\rho} + p^{1+\rho} \right). \quad (292)$$

Next, we check Condition 71. For this purpose, we check Condition (i) in Lemma 69 by treating \mathcal{V}_z given in (265). Since $\mathcal{V}_1 = \{1, 3\}$, $\mathcal{V}_2 = \{2, 4\}$, $\mathcal{V}_3 = \{1, 4\}$, and $\mathcal{V}_4 = \{2, 3\}$,

in the above choice of Q_V , we have $\sum_{v'' \in \mathcal{V}_z} Q_V(v'') = \frac{1}{2}$, which implies

$$\sum_{z \in \mathcal{Z}_V} \frac{\max_{v' \in \mathcal{V}} \bar{W}^Z(z|v')}{\sum_{v'' \in \mathcal{V}_z} Q_V(v'')} = 2 \frac{\frac{1}{2} - p}{\frac{1}{2}} = 4(\frac{1}{2} - p) \quad (293)$$

for all $v \in \mathcal{V}$. Thus, Condition 71 holds. Hence,

$$E_{0,\max}(1|\bar{W}^Z) = \log 4(\frac{1}{2} - p) \quad (294)$$

$$F_1 = \log 2. \quad (295)$$

Further, Theorem 80 guarantees that $\tilde{E}^{E_0}(R, \bar{W}^Z \times Q_V) = \tilde{E}^l(R, \bar{W}^Z \times Q_V)$. So, we numerically compare only $\tilde{E}^\psi(R, \bar{W}^Z \times Q_V)$ and $\tilde{E}^{E_0}(R, \bar{W}^Z \times Q_V)$ in Fig. 4. Since $\tilde{E}^{E_0}(R, \bar{W}^Z \times Q_V)$ attains the minimum value due to Theorem 80, $\tilde{E}^{E_0}(R, \bar{W}^Z \times Q_V)$ does not depend on q . Further, $\tilde{E}^\psi(R, \bar{W}^Z \times Q_V)$ also does not depend on q due to the form of $\tilde{E}^\psi(R, \bar{W}^Z \times Q_V)$. Similar to Fig. 3, Fig. 5 suggests that the parameter ρ realizing $\tilde{E}^{E_0}(R, \bar{W}^Z \times Q_V)$ has a behavior different from the parameter ρ realizing $\tilde{E}^\psi(R, \bar{W}^Z \times Q_V)$.

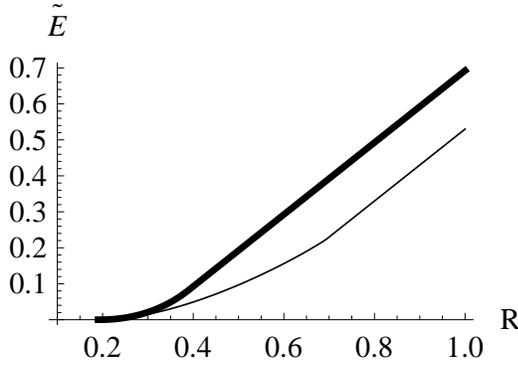


Fig. 4. Lower bounds of exponent in Example 84 with $p = 0.1$. In this case, $I(V; Z|\bar{W}^Z \times Q_V) = 0.192745$. Thick line and Normal line express $\tilde{E}^\psi(R, \bar{W}^Z \times Q_V)$ and $\tilde{E}^{E_0}(R, \bar{W}^Z \times Q_V) = \tilde{E}^l(R, \bar{W}^Z \times Q_V)$ as functions of R from $R = 0.192745$ to $R = 1.0$ with the origin $(0, 1, 0)$. Thick line is straight when $R \geq 0.4$ because $\operatorname{argmax}_{0 \leq \rho \leq 1} \rho R - \psi(\rho|\bar{W}^Z, Q_V)$ is 1 when $R \geq 0.4$, as in Fig 5. Normal line is straight when $R \geq 0.7$ because $\operatorname{argmax}_{0 \leq \rho \leq 1} \rho R - E_0(\rho|\bar{W}^Z, Q_V)$ is 1 when $R \geq 0.7$, as in Fig 5.

D. Proof of Lemma 63

Proof: We can show (258) and (260) by direct calculations. Now, we show (260). In general, when $b_i > 0$ and $a_1 = a_2 = \dots = a_l > a_{l+1} > 0$ for $i = l+1, \dots, k$, the relation

$$\begin{aligned} & \lim_{\rho \rightarrow 1-0} \left(\sum_{i=1}^k b_i a_i^{\frac{1}{1-\rho}} \right)^{1-\rho} \\ &= \lim_{\rho \rightarrow 1-0} \left(\left(\sum_{i=1}^l b_i a_1^{\frac{1}{1-\rho}} \right)^{1-\rho} \left(1 + \sum_{i=l+1}^k \frac{b_i}{\sum_{i=1}^l b_i} \frac{a_i^{\frac{1}{1-\rho}}}{a_1^{\frac{1}{1-\rho}}} \right)^{1-\rho} \right) \\ &= \lim_{\rho \rightarrow 1-0} \left(\left(\sum_{i=1}^l b_i a_1^{\frac{1}{1-\rho}} \right)^{1-\rho} \right) = a_1 \end{aligned} \quad (296)$$

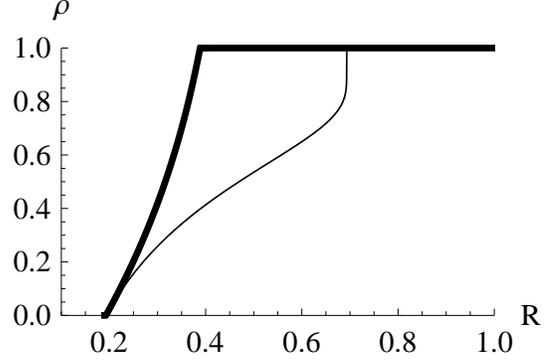


Fig. 5. Relation between R and ρ realizing the optimal value in Example 84 with $p = 0.1$. Normal line expresses $\operatorname{argmax}_{0 \leq \rho \leq 1} \rho R - E_0(\rho|\bar{W}^Z, Q_V)$, which realizes $\tilde{E}^{E_0}(R, \bar{W}^Z \times Q_V)$. Thick line expresses $\operatorname{argmax}_{0 \leq \rho \leq 1} \rho R - \psi(\rho|\bar{W}^Z, Q_V)$, which realizes $\tilde{E}^\psi(R, \bar{W}^Z \times Q_V)$. There is no graph corresponding to $\tilde{E}^l(R, \bar{W}^Z \times Q_V)$ because $\tilde{E}^l(R, \bar{W}^Z \times Q_V)$ is not given as maximization with respect to ρ . The origin is $(0, 1, 0)$.

holds. That is, the difference $(\sum_{i=1}^k b_i a_i^{\frac{1}{1-\rho}})^{1-\rho} - ((\sum_{i=1}^l b_i a_1^{\frac{1}{1-\rho}})^{1-\rho})$ behaves as $O(\exp(-\frac{a}{1-\rho}))$ with a constant a . Applying the above general discussion, we have

$$\begin{aligned} & \lim_{\rho \rightarrow 1-0} \sum_u Q_U(u) \sum_z \left[\sum_v Q_{V|U}(v|u) \bar{W}^Z(z|v)^{\frac{1}{1-\rho}} \right]^{1-\rho} \\ &= \lim_{\rho \rightarrow 1-0} \sum_u Q_U(u) \sum_z \left[\sum_{v \in \mathcal{V}_z(Q_{V|U=u})} Q_{V|U}(v|u) \right. \\ & \quad \left. \cdot \left(\max_{v \in \operatorname{supp}(Q_{V|U=u})} \bar{W}^Z(z|v) \right)^{\frac{1}{1-\rho}} \right]^{1-\rho} \\ &= \lim_{\rho \rightarrow 1-0} \sum_u Q_U(u) \sum_z \left[\left(\sum_{v \in \mathcal{V}_z(Q_{V|U=u})} Q_{V|U}(v|u) \right)^{1-\rho} \right. \\ & \quad \left. \cdot \left(\max_{v \in \operatorname{supp}(Q_{V|U=u})} \bar{W}^Z(z|v) \right) \right] \\ &= \sum_u Q_U(u) \sum_z \left(\max_{v \in \operatorname{supp}(Q_{V|U=u})} \bar{W}^Z(z|v) \right). \end{aligned}$$

where $\mathcal{V}_z(Q_{V|U=u}) := \{v \in \operatorname{supp}(Q_{V|U=u}) | \max_{v \in \operatorname{supp}(Q_{V|U=u})} \bar{W}^Z(z|v)\}$. Hence, we obtain (260).

Further, since $x \mapsto -\log x$ is concave, the map $Q_V \mapsto F_1(Q_V)$ is concave. The remaining task is the proof of the equation (259), will be shown in the wide space style in the next page. ■

E. Proof of Lemma 65

Proof: Due to (260), we have

$$\begin{aligned} E_{0,\max}(1|\bar{W}^Z) &= \max_{Q'_{VV}} \lim_{\rho \rightarrow 1-0} E_0(\rho|\bar{W}^Z, Q'_{V|U}, Q'_U) \\ &= \max_{Q'_{VV}} \log \sum_u Q_U(u) \sum_z \max_{v \in \operatorname{supp}(Q_{V|U=u})} \bar{W}^Z(z|v) \\ &= \log \sum_z \max_v \bar{W}^Z(z|v), \end{aligned}$$

Proof of (259): We have

$$\begin{aligned} & \frac{d}{d\rho} E_0(\rho|\overline{W}^Z, Q_{V|U}, Q_U) \\ &= \frac{\sum_u Q_U(u) \sum_z (\sum_v \frac{1}{1-\rho} (\log \overline{W}^Z(z|v)) Q_{V|U}(v|u) \overline{W}^Z(z|v)^{\frac{1}{1-\rho}}) (\sum_v Q_{V|U}(v|u) \overline{W}^Z(z|v)^{\frac{1}{1-\rho}})^{-\rho}}{\sum_u Q_U(u) \sum_z (\sum_v Q_{V|U}(v|u) \overline{W}^Z(z|v)^{\frac{1}{1-\rho}})^{1-\rho}} \\ & \quad - \frac{\sum_u Q_U(u) \sum_z \log(\sum_v Q_{V|U}(v|u) \overline{W}^Z(z|v)^{\frac{1}{1-\rho}}) (\sum_v Q_{V|U}(v|u) \overline{W}^Z(z|v)^{\frac{1}{1-\rho}})^{1-\rho}}{\sum_u Q_U(u) \sum_z (\sum_v Q_{V|U}(v|u) \overline{W}^Z(z|v)^{\frac{1}{1-\rho}})^{1-\rho}}. \end{aligned}$$

When ρ approaches 1, $\sum_v Q_{V|U}(v|u) \overline{W}^Z(z|v)^{\frac{1}{1-\rho}}$ approaches $(\sum_{v \in \mathcal{V}_z} Q_{V|U}(v|u)) (\max_{v'} \overline{W}^Z(z|v'))^{\frac{1}{1-\rho}}$. Hence,

$$\begin{aligned} & \lim_{\rho \rightarrow 1-0} \frac{d}{d\rho} E_0(\rho|\overline{W}^Z, Q_{V|U}, Q_U) \\ &= \lim_{\rho \rightarrow 1-0} \left(\frac{\sum_u Q_U(u) \sum_z (\frac{1}{1-\rho} \log \max_{v'} \overline{W}^Z(z|v')) (\sum_{v \in \mathcal{V}_z} Q_{V|U}(v|u))^{1-\rho} \max_{v'} \overline{W}^Z(z|v')}{\sum_u Q_U(u) \sum_z (\sum_{v \in \mathcal{V}_z} Q_{V|U}(v|u))^{1-\rho} \max_{v'} \overline{W}^Z(z|v')} \right. \\ & \quad \left. - \frac{\sum_u Q_U(u) \sum_z (\frac{1}{1-\rho} \log \max_{v'} \overline{W}^Z(z|v') + \log(\sum_{v \in \mathcal{V}_z} Q_{V|U}(v|u))) (\sum_{v \in \mathcal{V}_z} Q_{V|U}(v|u))^{1-\rho} \max_{v'} \overline{W}^Z(z|v')}{\sum_u Q_U(u) \sum_z (\sum_{v \in \mathcal{V}_z} Q_{V|U}(v|u))^{1-\rho} \max_{v'} \overline{W}^Z(z|v')} \right) \\ &= \lim_{\rho \rightarrow 1-0} \frac{-\sum_u Q_U(u) \sum_z \log(\sum_{v \in \mathcal{V}_z} Q_{V|U}(v|u)) (\sum_{v \in \mathcal{V}_z} Q_{V|U}(v|u))^{1-\rho} \max_{v'} \overline{W}^Z(z|v')}{\sum_u Q_U(u) \sum_z (\sum_{v \in \mathcal{V}_z} Q_{V|U}(v|u))^{1-\rho} \max_{v'} \overline{W}^Z(z|v')} \\ &= \lim_{\rho \rightarrow 1-0} \frac{-\sum_u Q_U(u) \sum_z \log(\sum_{v \in \mathcal{V}_z} Q_{V|U}(v|u)) \max_{v'} \overline{W}^Z(z|v')}{\sum_u Q_U(u) \sum_z \max_{v'} \overline{W}^Z(z|v')}, \end{aligned} \tag{297}$$

which implies (259). \blacksquare

which implies (266).

Assume that the support of $Q_{V|U=u}$ contains $\{v \in \mathcal{V} | \min_z \frac{\max_{v'} \overline{W}^Z(z|v')}{\overline{W}^Z(z|v)} = 1\}$ for any $u \in \text{supp}(Q_U)$. Due to (260), we have

$$E_0(1|\overline{W}^Z, Q_{V|U}, Q_U) = \log \sum_z \max_v \overline{W}^Z(z|v). \tag{298}$$

Combining (266), we obtain (267). Hence, as a special case of (257), we obtain (268). \blacksquare

F. Proofs of Lemmas 68, 72, and 74

Lemma 85: Let f be a concave C^1 function from \mathbf{R}^d to \mathbf{R} and $\mathcal{P}(d)$ be the subset $\{(x_1, \dots, x_d) \in \mathbf{R}^d | x_i \geq 0, \sum_{i=1}^d x_i = 1\}$. The following two conditions for $x = (x_1, \dots, x_d) \in \mathcal{P}(d)$ are equivalent.

(i)

$$f(x) = \max_{x' \in \mathcal{P}(d)} f(x'). \tag{299}$$

(ii) The following relation holds for any $i \neq j$.

$$\frac{\partial}{\partial x^i} f(x) = \frac{\partial}{\partial x^j} f(x). \tag{300}$$

Proof of Lemma 85: We choose variable $y = (y_1, \dots, y_{d-1}) \in \mathbf{R}^{d-1}$, and define a function $\tilde{f}(y) := f(y_1, \dots, y_{d-1}, 1 - \sum_{i=1}^{d-1} y_i)$. Due to the concavity, the condition (i) holds if and only if $\frac{\partial}{\partial y_i} \tilde{f}(y) = 0$ for $i = 1, \dots, d-1$. This condition is equivalent

to the condition (ii) because $\frac{\partial}{\partial y_i} \tilde{f}(y) = \frac{\partial}{\partial x_i} f(y_1, \dots, y_{d-1}, 1 - \sum_{i=1}^{d-1} y_i) - \frac{\partial}{\partial x_d} f(y_1, \dots, y_{d-1}, 1 - \sum_{i=1}^{d-1} y_i)$. \blacksquare

Proof of Lemma 68: In order to apply Lemma 85, we regard all of probabilities $Q_V(v)$ as independent parameters by removing the constraint $\sum_v Q_V(v) = 1$. The partial derivatives are calculated as

$$\begin{aligned} & \frac{\partial}{\partial Q_V(v)} \sum_z (\sum_{v'} Q_V(v') \overline{W}^Z(z|v')^{\frac{1}{1-\rho}})^{1-\rho} \\ &= \sum_z (1-\rho) (\sum_{v'} Q_V(v') \overline{W}^Z(z|v')^{\frac{1}{1-\rho}})^{-\rho} \overline{W}^Z(z|v)^{\frac{1}{1-\rho}}. \end{aligned}$$

Hence, Lemma 85 guarantees the equivalence between (i) and (ii). Condition (iii) trivially implies Condition (i).

The remaining task is showing Condition (i) implies Condition (iii). Assume Condition (i). Since $\sum_z \overline{W}^Z(z|v)^{\frac{1}{1-\rho}} (\sum_{v'} Q_V(v') \overline{W}^Z(z|v')^{\frac{1}{1-\rho}})^{-\rho}$ does not depend on v and Condition (ii) holds,

$$\begin{aligned} & \sum_z \overline{W}^Z(z|v)^{\frac{1}{1-\rho}} (\sum_{v'} Q_V(v') \overline{W}^Z(z|v')^{\frac{1}{1-\rho}})^{-\rho} \\ &= \sum_v Q_V(v) \sum_z \overline{W}^Z(z|v)^{\frac{1}{1-\rho}} (\sum_{v'} Q_V(v') \overline{W}^Z(z|v')^{\frac{1}{1-\rho}})^{-\rho} \\ &= \sum_z (\sum_v Q_V(v) \overline{W}^Z(z|v)^{\frac{1}{1-\rho}})^{1-\rho} = e^{E_0(\rho|\overline{W}^Z, Q_V)} \\ &= \max_{Q'_V} e^{E_0(\rho|\overline{W}^Z, Q'_V)} = e^{E_0, \max(\rho|\overline{W}^Z)}. \end{aligned}$$

Proof of Lemma 72: Assume that

$$\sum_v Q_V(v) \bar{W}^Z(z|v)^{\frac{1}{1-\rho}} \neq \sum_v Q'_V(v) \bar{W}^Z(z|v)^{\frac{1}{1-\rho}} \quad (301)$$

for any $z \in \mathcal{Z}$. Due to the strict concavity of $x \mapsto x^{1-\rho}$, we have

$$\begin{aligned} & \frac{1}{2} \left(\sum_v Q_V(v) \bar{W}^Z(z|v)^{\frac{1}{1-\rho}} \right)^{1-\rho} + \frac{1}{2} \left(\sum_v Q'_V(v) \bar{W}^Z(z|v)^{\frac{1}{1-\rho}} \right)^{1-\rho} \\ & < \left(\sum_v \left(\frac{1}{2} Q_V(v) + \frac{1}{2} Q'_V(v) \right) \bar{W}^Z(z|v)^{\frac{1}{1-\rho}} \right)^{1-\rho}. \end{aligned} \quad (302)$$

Hence,

$$\begin{aligned} & \frac{1}{2} \sum_z \left(\sum_v Q_V(v) \bar{W}^Z(z|v)^{\frac{1}{1-\rho}} \right)^{1-\rho} + \frac{1}{2} \sum_z \left(\sum_v Q'_V(v) \bar{W}^Z(z|v)^{\frac{1}{1-\rho}} \right)^{1-\rho} \\ & < \sum_z \left(\sum_v \left(\frac{1}{2} Q_V(v) + \frac{1}{2} Q'_V(v) \right) \bar{W}^Z(z|v)^{\frac{1}{1-\rho}} \right)^{1-\rho}. \end{aligned} \quad (303)$$

However, Lemma 68 guarantees that

$$\begin{aligned} \sum_z \left(\sum_v Q_V(v) \bar{W}^Z(z|v)^{\frac{1}{1-\rho}} \right)^{1-\rho} &= \sum_z \left(\sum_v Q'_V(v) \bar{W}^Z(z|v)^{\frac{1}{1-\rho}} \right)^{1-\rho} \\ &= \max_{Q'_V} e^{E_0(\rho | \bar{W}^Z, Q'_V)}. \end{aligned} \quad (304)$$

Since (303) contradicts (304), we obtain the desired argument. \blacksquare

Proof of Lemma 74: As

$$W^{Z,\rho} \circ Q_{V,\rho}(z) = \frac{(\sum_v Q_{V,\rho}(v) \bar{W}^Z(z|v)^{\frac{1}{1-\rho}})^{1-\rho}}{\sum_z (\sum_v Q_{V,\rho}(v) \bar{W}^Z(z|v)^{\frac{1}{1-\rho}})^{1-\rho}},$$

we can calculate the mutual information $I(V; Z)[W^{Z,\rho} \times Q_{V,\rho}]$ as

$$\begin{aligned} & I(V; Z)[W^{Z,\rho} \times Q_{V,\rho}] \\ &= \sum_{v,z} \frac{Q_{V,\rho}(v) \bar{W}^Z(z|v)^{\frac{1}{1-\rho}} (\sum_v Q_{V,\rho}(v) \bar{W}^Z(z|v)^{\frac{1}{1-\rho}})^{-\rho}}{\sum_z \bar{W}^Z(z|v)^{\frac{1}{1-\rho}} (\sum_v Q_{V,\rho}(v) \bar{W}^Z(z|v)^{\frac{1}{1-\rho}})^{-\rho}} \\ & \quad \cdot \left[\log \left[\bar{W}^Z(z|v)^{\frac{1}{1-\rho}} (\sum_v Q_{V,\rho}(v) \bar{W}^Z(z|v)^{\frac{1}{1-\rho}})^{-\rho} \right] \right. \\ & \quad \left. - \log \left[\left(\sum_v Q_{V,\rho}(v) \bar{W}^Z(z|v)^{\frac{1}{1-\rho}} \right)^{1-\rho} \right] \right] \\ &= \sum_{v,z} \frac{Q_{V,\rho}(v) \bar{W}^Z(z|v)^{\frac{1}{1-\rho}} (\sum_v Q_{V,\rho}(v) \bar{W}^Z(z|v)^{\frac{1}{1-\rho}})^{-\rho}}{\sum_z \bar{W}^Z(z|v)^{\frac{1}{1-\rho}} (\sum_v Q_{V,\rho}(v) \bar{W}^Z(z|v)^{\frac{1}{1-\rho}})^{-\rho}} \\ & \quad \cdot \left[\frac{1}{1-\rho} \log \bar{W}^Z(z|v) - \log \left[\sum_v Q_{V,\rho}(v) \bar{W}^Z(z|v)^{\frac{1}{1-\rho}} \right] \right] \\ &= F_\rho(Q_{V,\rho}), \end{aligned} \quad (305)$$

where the final equation follows from (261). We obtain the second equation of (273).

Since the constraint (i) in Lemma 68 for $Q_{V,\rho}$ is differentiable with respect to ρ , for a given $\rho_0 \in (0, 1)$, we can choose

■ $Q_{V,\rho}$ such that the map $\rho \mapsto Q_{V,\rho}$ is differentiable at least in an enough small neighborhood of ρ_0 . Since

$$\frac{d}{d\rho} E_0(\rho_0 | \bar{W}^Z, Q_{V,\rho})|_{\rho=\rho_0} = 0, \quad (306)$$

we have

$$\begin{aligned} & F_{\rho_0} = \frac{d}{d\rho} E_0(\rho | \bar{W}^Z, Q_{V,\rho})|_{\rho=\rho_0} \\ &= \frac{d}{d\rho} E_0(\rho | \bar{W}^Z, Q_{V,\rho_0})|_{\rho=\rho_0} + \frac{d}{d\rho} E_0(\rho_0 | \bar{W}^Z, Q_{V,\rho})|_{\rho=\rho_0} \\ &= \frac{d}{d\rho} E_0(\rho | \bar{W}^Z, Q_{V,\rho_0})|_{\rho=\rho_0} = F_{\rho_0}(Q_{V,\rho_0}). \end{aligned} \quad (307)$$

Hence, we obtain the first equation of (273).

The conditional divergence $D(W^Z \| \bar{W}^Z | Q_{V,\rho})$ is calculated to

$$\begin{aligned} & D(W^{V,\rho} \| \bar{W}^Z | Q_{V,\rho}) \\ &= \sum_{v,z} \frac{Q_{V,\rho}(v) \bar{W}^Z(z|v)^{\frac{1}{1-\rho}} (\sum_v Q_{V,\rho}(v) \bar{W}^Z(z|v)^{\frac{1}{1-\rho}})^{-\rho}}{\sum_z \bar{W}^Z(z|v)^{\frac{1}{1-\rho}} (\sum_{v'} Q_{V,\rho}(v') \bar{W}^Z(z|v')^{\frac{1}{1-\rho}})^{-\rho}} \\ & \quad \cdot \left(\log \left[\bar{W}^Z(z|v)^{\frac{1}{1-\rho}} (\sum_v Q_{V,\rho}(v) \bar{W}^Z(z|v)^{\frac{1}{1-\rho}})^{-\rho} \right] - \log \bar{W}^Z(z|v) \right) \\ & \quad - \sum_v Q_{V,\rho}(v) \log \left[\sum_z \bar{W}^Z(z|v)^{\frac{1}{1-\rho}} (\sum_{v'} Q_{V,\rho}(v') \bar{W}^Z(z|v')^{\frac{1}{1-\rho}})^{-\rho} \right] \\ &= \sum_{v,z} \frac{Q_{V,\rho}(v) \bar{W}^Z(z|v)^{\frac{1}{1-\rho}} (\sum_v Q_{V,\rho}(v) \bar{W}^Z(z|v)^{\frac{1}{1-\rho}})^{-\rho}}{\sum_z \bar{W}^Z(z|v)^{\frac{1}{1-\rho}} (\sum_v Q_{V,\rho}(v) \bar{W}^Z(z|v)^{\frac{1}{1-\rho}})^{-\rho}} \\ & \quad \cdot \left(\frac{\rho}{1-\rho} \log \bar{W}^Z(z|v) - \rho \log \left[\sum_v Q_{V,\rho}(v) \bar{W}^Z(z|v)^{\frac{1}{1-\rho}} \right] \right) \\ & \quad - \sum_v Q_{V,\rho}(v) \log \left[\sum_z \bar{W}^Z(z|v)^{\frac{1}{1-\rho}} (\sum_{v'} Q_{V,\rho}(v') \bar{W}^Z(z|v')^{\frac{1}{1-\rho}})^{-\rho} \right] \\ &= \rho F_\rho(Q_{V,\rho}) - \sum_v Q_{V,\rho}(v) \log \left[\sum_z (\sum_{v'} Q_{V,\rho}(v') \bar{W}^Z(z|v')^{\frac{1}{1-\rho}})^{1-\rho} \right] \\ &= \rho F_\rho - E(\rho | \bar{W}^Z, Q_{V,\rho}). \end{aligned}$$

We obtain (274). \blacksquare

G. Proofs of Lemmas 69, 73, and 75

Proof of Lemma 69: In order to apply Lemma 85, we regard all of probabilities $Q_V(v)$ as independent parameters by removing the constraint $\sum_v Q_V(v) = 1$. The partial derivatives are calculated as

$$\begin{aligned} & \frac{\partial}{\partial Q_V(v)} - \frac{\sum_z \log(\sum_{v' \in \mathcal{V}_z} Q_V(v')) \max_{v'} \bar{W}^Z(z|v')}{\sum_z \max_{v'} \bar{W}^Z(z|v')} \\ &= - \sum_{z \in \mathcal{Z}_v} \frac{\max_{v' \in \mathcal{V}} \bar{W}^Z(z|v')}{\sum_{v'' \in \mathcal{V}_z} Q_V(v'')}. \end{aligned}$$

Hence, Lemma 85 guarantees the equivalence between (i) and (ii). Condition (iii) trivially implies Condition (i).

The remaining task is showing Condition (i) implies Condition (iii). Assume Condition (i). Since

$\sum_z \bar{W}^Z(z|v)^{\frac{1}{1-\rho}} (\sum_{v'} Q_V(v') \bar{W}^Z(z|v')^{\frac{1}{1-\rho}})^{-\rho}$ does not depend on v and Condition (ii) holds, we have

$$\begin{aligned} & \sum_{z \in \mathcal{Z}_v} \frac{\bar{W}^Z(z|v)}{\sum_{v'' \in \mathcal{V}_z} Q_V(v'')} = \sum_{z \in \mathcal{Z}_v} \frac{\max_{v' \in \mathcal{V}} \bar{W}^Z(z|v')}{\sum_{v'' \in \mathcal{V}_z} Q_V(v'')} \\ &= \sum_v Q_V(v) \sum_{z \in \mathcal{Z}_v} \frac{\max_{v' \in \mathcal{V}} \bar{W}^Z(z|v')}{\sum_{v'' \in \mathcal{V}_z} Q_V(v'')} \\ &= \sum_{(z,v) \in \mathcal{K}} Q_V(v) \frac{\max_{v' \in \mathcal{V}} \bar{W}^Z(z|v')}{\sum_{v'' \in \mathcal{V}_z} Q_V(v'')} \\ &= \sum_z \sum_{v \in \mathcal{V}_z} Q_V(v) \frac{\max_{v' \in \mathcal{V}} \bar{W}^Z(z|v')}{\sum_{v'' \in \mathcal{V}_z} Q_V(v'')} = \sum_z \max_{v'} \bar{W}^Z(z|v'). \end{aligned}$$

Proof of Lemma 73: We focus on the function $\{\sum_{v'' \in \mathcal{V}_z} Q_V(v'')\}_z \mapsto -\frac{\sum_z \log(\sum_{v \in \mathcal{V}_z} Q_V(v)) \max_{v'} \bar{W}^Z(z|v')}{\sum_z \max_{v'} \bar{W}^Z(z|v')}$, which is strictly concave. Hence, when there exists an element $z \in \mathcal{Z}$ such that $\sum_{v'' \in \mathcal{V}_z} Q_V(v'') \neq \sum_{v'' \in \mathcal{V}_z} Q'_V(v'')$ for two distributions Q_V and Q'_V , the convex combination $\frac{Q_V + Q'_V}{2}$ gives a strictly greater value for the above function, which contradicts (ii) of Lemma 69. Hence, $\sum_{v'' \in \mathcal{V}_z} Q_V(v'') = \sum_{v'' \in \mathcal{V}_z} Q'_V(v'')$ for all $z \in \mathcal{Z}$. ■

Proof of Lemma 75: Since

$$W^{Z,1} \times Q_{V,1}(v, z) = \begin{cases} \frac{Q_{V,1}(v) \bar{W}^Z(z|v)}{\sum_{v'' \in \mathcal{V}_z} Q_{V,1}(v'') \sum_{z'} \max_{v'} \bar{W}^Z(z'|v')} & z \in \mathcal{Z}_v \\ 0 & z \in \mathcal{Z}_v^c, \end{cases} \quad (308)$$

the mutual information $I(V; Z)[W^{Z,1} \times Q_{V,1}]$ is calculated as

$$\begin{aligned} I(V; Z)[W^{Z,1} \times Q_{V,1}] &= -\frac{\sum_z \log(\sum_{v \in \mathcal{V}_z} Q_{V,1}(v)) \max_{v'} \bar{W}^Z(z|v')}{\sum_z \max_{v'} \bar{W}^Z(z|v')} \\ &= F_1(Q_{V,1}), \end{aligned} \quad (309)$$

where the final equation follows from (262). Hence, we obtain the second equation in (275). The first equation in (275) follows from the limit $\rho \rightarrow 1 - 0$ at (307).

When Q_V satisfies Condition 71,

$$\begin{aligned} & D(W^{Z,1} \| \bar{W}^Z | Q_V) \\ &= -\sum_{z,v} W^{Z,1} \times Q_{V,1}(v, z) \log \left[\sum_{v'' \in \mathcal{V}_z} Q_V(v'') \sum_{z'} \max_{v'} \bar{W}^Z(z'|v') \right] \\ &= -\log \left[\sum_{z'} \max_{v'} \bar{W}^Z(z'|v') \right] \\ &\quad - \sum_z \log \left[\sum_{v'' \in \mathcal{V}_z} Q_V(v'') \right] W^{Z,1} \circ Q_V(z) \\ &= -\log \left[\sum_{z'} \max_{v'} \bar{W}^Z(z'|v') \right] \\ &\quad - \frac{\sum_z \log \left[\sum_{v \in \mathcal{V}_z} Q_V(v) \right] \max_{v'} \bar{W}^Z(z|v')}{\sum_z \max_{v'} \bar{W}^Z(z|v')} \\ &= F_1 - E_{0, \max}(1 | \bar{W}^Z), \end{aligned}$$

which implies (276). ■

H. Proof of Lemma 76

Proof of Lemma 76: Due to Condition 70 with ρ , we can choose a constant C_ρ in the following way: the relation

$$C_\rho = \sum_z \bar{W}^Z(z|v)^{\frac{1}{1-\rho}} \left(\sum_{v'} Q_{V,\rho}(v') \bar{W}^Z(z|v')^{\frac{1}{1-\rho}} \right)^{-\rho} \quad (310)$$

holds for all v . Due to the general relation as (296), we have

$$\begin{aligned} C &:= \lim_{\rho \rightarrow 1-0} C_\rho \\ &= \lim_{\rho \rightarrow 1-0} \sum_z \bar{W}^Z(z|v)^{\frac{1}{1-\rho}} \left(\sum_{v'} Q_{V,\rho}(v') \bar{W}^Z(z|v')^{\frac{1}{1-\rho}} \right)^{-\rho} \\ &= \lim_{\rho \rightarrow 1-0} \sum_{z \in \mathcal{Z}_v} \left(\sum_{v'' \in \mathcal{V}_z} Q_{V,\rho}(v'') \right)^{-\rho} \max_{v'} \bar{W}^Z(z|v') \\ &= \sum_{z \in \mathcal{Z}_v} \frac{\max_{v'} \bar{W}^Z(z|v')}{\sum_{v'' \in \mathcal{V}_z} (\lim_{n \rightarrow \infty} Q_{V,\rho_n}(v''))}. \end{aligned}$$

Since C does not depend on v , the distribution $\lim_{n \rightarrow \infty} Q_{V,\rho_n}$ satisfies Condition 71. ■

I. Proof of Lemma 66

We show the inequality in (269). First, we obtain the inequality (314), which is displayed in the wide space in the next page.

Since $\frac{1}{1-\rho} + \frac{-\rho}{1-\rho} = 1$, the reverse Hölder inequality yields that

$$\begin{aligned} & \sum_z \left(\sum_v Q_{V|U}(v|u) \bar{W}^Z(z|v)^{\frac{1}{1-\rho}} \right) \tilde{Q}_Z(z)^{\frac{-\rho}{1-\rho}} \\ & \geq \left(\sum_z \left(\sum_v Q_{V|U}(v|u) \bar{W}^Z(z|v)^{\frac{1}{1-\rho}} \right)^{1-\rho} \right)^{\frac{1}{1-\rho}} \left(\sum_z \left(\tilde{Q}_Z(z)^{\frac{-\rho}{1-\rho}} \right)^{-\frac{1-\rho}{\rho}} \right)^{\frac{-\rho}{1-\rho}} \\ & \geq \min_{\tilde{Q}_Z \in \mathcal{P}(\mathcal{Z})} \left(\sum_z \left(\sum_v Q_{V|U}(v|u) \bar{W}^Z(z|v)^{\frac{1}{1-\rho}} \right)^{1-\rho} \right)^{\frac{1}{1-\rho}} \left(\sum_z \tilde{Q}_Z(z)^{\frac{-\rho}{1-\rho}} \right)^{\frac{-\rho}{1-\rho}} \\ & = \left(\sum_z \left(\sum_v Q_{V|U}(v|u) \bar{W}^Z(z|v)^{\frac{1}{1-\rho}} \right)^{1-\rho} \right)^{\frac{1}{1-\rho}}. \end{aligned}$$

The equality holds only when $(\sum_v Q_{V|U}(v|u) \bar{W}^Z(z|v)^{\frac{1}{1-\rho}})^{1-\rho} = C \tilde{Q}_Z(z)$ with a constant C . Hence,

$$\begin{aligned} & \min_{\tilde{Q}_Z \in \mathcal{P}(\mathcal{Z})} \sum_z \left(\sum_v Q_{V|U}(v|u) \bar{W}^Z(z|v)^{\frac{1}{1-\rho}} \right) \tilde{Q}_Z(z)^{\frac{-\rho}{1-\rho}} \\ & = \left(\sum_z \left(\sum_v Q_{V|U}(v|u) \bar{W}^Z(z|v)^{\frac{1}{1-\rho}} \right)^{1-\rho} \right)^{\frac{1}{1-\rho}}. \end{aligned}$$

Thus,

$$\begin{aligned} & -(1-\rho) \sum_u Q_U(u) \log \left[\min_{\tilde{Q}_Z \in \mathcal{P}(\mathcal{Z})} \sum_z \left(\sum_v Q_{V|U}(v|u) \bar{W}^Z(z|v)^{\frac{1}{1-\rho}} \right) \tilde{Q}_Z(z)^{\frac{-\rho}{1-\rho}} \right] \\ &= -(1-\rho) \sum_u Q_U(u) \log \left[\left(\sum_z \left(\sum_v Q_{V|U}(v|u) \bar{W}^Z(z|v)^{\frac{1}{1-\rho}} \right)^{1-\rho} \right)^{\frac{1}{1-\rho}} \right] \\ &= -\sum_u Q_U(u) \log \left(\sum_z \left(\sum_v Q_{V|U}(v|u) \bar{W}^Z(z|v)^{\frac{1}{1-\rho}} \right)^{1-\rho} \right) \\ & \geq -\log \sum_u Q_U(u) \left(\sum_z \left(\sum_v Q_{V|U}(v|u) \bar{W}^Z(z|v)^{\frac{1}{1-\rho}} \right)^{1-\rho} \right) \\ &= -E_0(\rho | \bar{W}^Z, Q_{V|U}, Q_U), \end{aligned} \quad (315)$$

$$= -E_0(\rho | \bar{W}^Z, Q_{V|U}, Q_U), \quad (316)$$

$$\begin{aligned}
& \min_{W^Z \in \mathcal{W}(U \times V, Z)} D(W^Z \| \bar{W}^Z | Q_{VU}) - \rho I(V; Z|U)[W^Z \times Q_{VU}] \\
&= \min_{W^Z \in \mathcal{W}(U \times V, Z)} \left(\sum_u Q_U(u) \left(\sum_v Q_{V|U}(v|u) \sum_z W^Z(z|u, v) \log \frac{W^Z(z|u, v)}{\bar{W}^Z(z|v)} \right. \right. \\
&\quad \left. \left. - \rho \min_{\tilde{Q} \in \mathcal{P}(Z)} \sum_v Q_{V|U}(v|u) \sum_z W^Z(z|u, v) \log \frac{W^Z(z|u, v)}{\tilde{Q}(z)} \right) \right) \\
&= \min_{W^Z \in \mathcal{W}(U \times V, Z)} \max_{\tilde{W}^Z \in \mathcal{W}(U, Z)} \sum_u Q_U(u) \sum_v Q_{V|U}(v|u) \left(\sum_z W^Z(z|u, v) \log \frac{W^Z(z|u, v)}{\bar{W}^Z(z|v)} - \rho \sum_z W^Z(z|u, v) \log \frac{W^Z(z|u, v)}{\tilde{W}^Z(z|u)} \right) \\
&= \min_{W^Z \in \mathcal{W}(U \times V, Z)} \max_{\tilde{W}^Z \in \mathcal{W}(U, Z)} \sum_u Q_U(u) \sum_v Q_{V|U}(v|u) \sum_z W^Z(z|u, v) \log \frac{W^Z(z|u, v)^{1-\rho} \tilde{W}^Z(z|u)^\rho}{\bar{W}^Z(z|v)} \\
&= \max_{\tilde{W}^Z \in \mathcal{W}(U, Z)} \min_{W^Z \in \mathcal{W}(U \times V, Z)} \sum_u Q_U(u) \sum_v Q_{V|U}(v|u) \sum_z W^Z(z|u, v) \log \frac{W^Z(z|u, v)^{1-\rho} \tilde{W}^Z(z|u)^\rho}{\bar{W}^Z(z|v)} \tag{311} \\
&= (1-\rho) \max_{\tilde{W}^Z \in \mathcal{W}(U, Z)} \sum_u Q_U(u) \sum_v Q_{V|U}(v|u) \min_{\tilde{P}_Z \in \mathcal{P}(Z)} \sum_z \tilde{P}_Z(z) \log \frac{\tilde{P}_Z(z)}{\bar{W}^Z(z|v)^{\frac{1}{1-\rho}} \tilde{W}^Z(z|u)^{\frac{\rho}{1-\rho}}} \\
&= -(1-\rho) \min_{\tilde{W}^Z \in \mathcal{W}(U, Z)} \sum_u Q_U(u) \sum_v Q_{V|U}(v|u) \log \sum_z \bar{W}^Z(z|v)^{\frac{1}{1-\rho}} \tilde{W}^Z(z|u)^{\frac{\rho}{1-\rho}} \\
&\geq -(1-\rho) \min_{\tilde{W}^Z \in \mathcal{W}(U, Z)} \sum_u Q_U(u) \log \sum_v Q_{V|U}(v|u) \sum_z \bar{W}^Z(z|v)^{\frac{1}{1-\rho}} \tilde{W}^Z(z|u)^{\frac{\rho}{1-\rho}} \tag{313} \\
&= -(1-\rho) \sum_u Q_U(u) \log \min_{\tilde{Q}_Z \in \mathcal{P}(Z)} \sum_z \left(\sum_v Q_{V|U}(v|u) \bar{W}^Z(z|v)^{\frac{1}{1-\rho}} \tilde{Q}_Z(z)^{\frac{\rho}{1-\rho}} \right). \tag{314}
\end{aligned}$$

The above derivation can be shown in the following way. The equality (311) follows from the minimax theorem [11, Chap. IV Prop. 2.3] because the function is concave for \tilde{W}^Z and is convex for W^Z . The equality (312) holds because the minimum is attained with $\tilde{P}_Z(z) = \bar{W}^Z(z|v)^{\frac{1}{1-\rho}} \tilde{W}^Z(z|u)^{\frac{\rho}{1-\rho}} / \sum_z \bar{W}^Z(z|v)^{\frac{1}{1-\rho}} \tilde{W}^Z(z|u)^{\frac{\rho}{1-\rho}}$. The inequality (313) follows from the concavity of $x \mapsto \log x$.

where (315) follows from the concavity of $x \mapsto \log x$. The combination of (314) and (316) yields (269).

The equality in (313) holds if and only if for an arbitrary fixed u , $\sum_z \bar{W}^Z(z|v)^{\frac{1}{1-\rho}} \tilde{W}^Z(z|u)^{\frac{\rho}{1-\rho}}$ does not depend on v with $\tilde{W}^Z(z|u) = (\sum_v Q_{V|U}(v|u) \bar{W}^Z(z|v)^{\frac{1}{1-\rho}})^{1-\rho} / \sum_z (\sum_v Q_{V|U}(v|u) \bar{W}^Z(z|v)^{\frac{1}{1-\rho}})^{1-\rho}$, i.e., the quantity $\sum_z \bar{W}^Z(z|v)^{\frac{1}{1-\rho}} (\sum_v Q_{V|U}(v|u) \bar{W}^Z(z|v)^{\frac{1}{1-\rho}})^{-\rho}$ does not depend on v for an arbitrary fixed u . The condition holds when $Q_{V|U=u}$ is $\text{argmin}_{Q_V} E_0(\rho | \bar{W}^Z, Q_V)$ because of Lemma 68. Further, the equality in (315) holds in this case. Hence, when $Q_{V|U=u}$ is $\text{argmin}_{Q_V} E_0(\rho | \bar{W}^Z, Q_V)$, the equality holds in the inequality (269).

XV. CONCLUSION

In order to treat the secure multiplex coding with dependent and non-uniform multiple messages and common messages, we have generalized resolvability to the case when input random variable is subject to a non-uniform distribution. Two kinds of generalization have been given. The first one (Theorem 14) is a simple extension of Han-Verdú's channel resolvability coding [13] with the non-uniform inputs. The second one (Theorem 17) uses randomly chosen affine mapping satisfying Condition 15 with the non-uniform inputs.

We have constructed two kinds of codes for the above type of SMC. Similar to BCC in [9], the second construction has two steps. In the first step, similar to the BCD encoder, we

apply superposition random coding. In the second step, as is illustrated in Fig. 1, we split the confidential message into the private message B_2 and a part B_1 of the common message encoded by the BCD encoder. Employing the second type of channel resolvability, we have derived a non-asymptotic formula for the average leaked information under this kind of code construction. On the other hand, in the first construction, the confidential message is simply sent as the private message encoded by the BCD encoder. Hence, it has only one step. Employing the first type of channel resolvability, we have derived a non-asymptotic formula for the average leaked information under this kind of code construction.

For asymptotic treatment for the non-uniform and dependent sources, we have introduced three kinds of asymptotic conditional uniformity conditions. Then, we have clarified the relation among three conditions, especially, that two of them are equivalent. Further, we have shown that these conditions can be satisfied by data compressed by Slepian-Wolf compression, in the respective senses. Extending the above formula for the second construction to the asymptotic case, we have derived the capacity region of SMC defined in our general setting, in which, the message is allowed to be dependent and non-uniform while it has to satisfy the weaker asymptotic conditional uniformity condition. We have shown the strong security when the the leaked information rate is zero and the message satisfies the stronger asymptotic conditional uniformity condition. Using the both formulas, we have also

derived the exponential decreasing rate of leaked information. While the first formula gives an upper bound in any case, the second one gives a better upper bound in some specific cases.

We have also given two kinds of practical constructions for SMC by using ordinary linear codes. Following our constructions, we can make a code satisfying a required security level. Further, we have given a universal code for SMC, which does not depend on the channel. Extending this result, we have derived a source-channel universal code for BCC, which does not depend on the channel or the source distribution.

ACKNOWLEDGMENT

RM would like to thank Prof. H. Yamamoto to teach him the secure multiplex coding. The authors are grateful to Prof. Alexander Vardy for pointing out the importance for the non-independent case for the multiple secret messages. The authors are grateful to Dr. Shun Watanabe for informing the references [36], [37], [38], [39]. They also would like to express their appreciation to the referee of this paper for his/her helpful comments. A part of this research was done during RM's stay at the Institute of Network Coding, the Chinese University of Hong Kong, and Department of Mathematical Sciences, Aalborg University. He greatly appreciates the hospitality by Prof. R. Yeung and Prof. O. Geil.

This research was partially supported by the MEXT Grant-in-Aid for Young Scientists (A) No. 20686026 and (B) No. 22760267, Grant-in-Aid for Scientific Research (A) No. 23246071, and the ImPACT Program of Council for Science, Technology and Innovation (Cabinet Office, Government of Japan). The Center for Quantum Technologies is funded by the Singapore Ministry of Education and the National Research Foundation as part of the Research Centres of Excellence programme.

APPENDIX A

INEQUALITY BETWEEN RÉNYI ENTROPY AND CONDITIONAL RÉNYI ENTROPY

In this appendix, we derive a useful inequality between Rényi entropy and conditional Rényi entropy, which was used in Subsection VIII-B. For this purpose, we prepare the following lemma.

Lemma 86: Any two distributions P_{XY} and Q_{XY} over $\mathcal{X} \times \mathcal{Y}$ satisfy

$$\psi(\rho|P_{X,Y}\|Q_{X,Y}) \geq \frac{1}{1-\rho} \psi(\rho(1-\rho)|P_{X,Y}\|Q_{Y|X} \times P_X) \quad (317)$$

for $\rho > 0$, where P_X is the marginal distribution of $P_{X,Y}$ on \mathcal{X} , and $Q_{Y|X}$ is the conditional distribution of $Q_{X,Y}$ on \mathcal{Y} conditioned with X .

When $Q_{X,Y}$ is the uniform distribution, $\frac{1}{\rho} \psi(\rho|P_{X,Y}\|Q_{X,Y}) = \log(|\mathcal{X}||\mathcal{Y}|) - H_{1+\rho}(X, Y)$ and $\frac{1}{\rho(1-\rho)} \psi(\rho(1-\rho)|P_{X,Y}\|Q_{Y|X} \times P_X) = \log|\mathcal{Y}| - H_{1+\rho(1-\rho)}(Y|X)$, which implies the following corollary of the above lemma as an inequality between Rényi entropy and conditional Rényi entropy.

Corollary 87: For $\rho > 0$, arbitrary random variables X and Y over \mathcal{X} and \mathcal{Y} satisfy

$$\log(|\mathcal{X}||\mathcal{Y}|) - H_{1+\rho}(X, Y) \geq \log|\mathcal{Y}| - H_{1+\rho(1-\rho)}(Y|X), \quad (318)$$

which implies

$$\log|\mathcal{X}| + H_{1+\rho(1-\rho)}(Y|X) \geq H_{1+\rho}(X, Y). \quad (319)$$

Proof of Lemma 86: Applying Hölder inequality $\sum_x P_X(x)|A(x)B(x)| \leq (\sum_x P_X(x)|A(x)|^{\frac{1}{1-\rho}})^{1-\rho} (\sum_x P_X(x)|B(x)|^{\frac{1}{\rho}})^{\rho}$, to the case $A(x) = P_X(x)^{\rho} Q_X(x)^{-\rho} (\sum_y P_{Y|X}(y|x)^{1+\rho(1-\rho)} Q_{Y|X}(y|x)^{-\rho(1-\rho)})^{\frac{1}{1-\rho}}$ and $B(x) = P_X(x)^{-\rho} Q_X(x)^{\rho}$, we obtain the following. In the following derivation, we employ the above Hölder inequality in (321), and the Jensen inequality for the convex function $x \mapsto x^{\frac{1}{1-\rho}}$ in (320), (322), and (323).

$$\begin{aligned} & e^{\frac{1}{1-\rho} \psi(\rho(1-\rho)|P_{X,Y}\|Q_{Y|X} \times P_X)} \\ &= \left(\sum_x P_X(x) \sum_y P_{Y|X}(y|x)^{1+\rho(1-\rho)} Q_{Y|X}(y|x)^{-\rho(1-\rho)} \right)^{\frac{1}{1-\rho}} \\ &\leq \sum_x P_X(x) \left(\sum_y P_{Y|X}(y|x)^{1+\rho(1-\rho)} Q_{Y|X}(y|x)^{-\rho(1-\rho)} \right)^{\frac{1}{1-\rho}} \quad (320) \\ &= \sum_x P_X(x) \left[(P_X(x)^{\rho} Q_X(x)^{-\rho} \right. \\ &\quad \cdot \left. \sum_y (P_{Y|X}(y|x)^{1+\rho(1-\rho)} Q_{Y|X}(y|x)^{-\rho(1-\rho)})^{\frac{1}{1-\rho}} (P_X(x)^{-\rho} Q_X(x)^{\rho}) \right] \\ &\leq \left[\sum_x P_X(x) P_X(x)^{\frac{\rho}{1-\rho}} Q_X(x)^{-\frac{\rho}{1-\rho}} \right. \\ &\quad \cdot \left. \left(\sum_y P_{Y|X}(y|x)^{1+\rho(1-\rho)} Q_{Y|X}(y|x)^{-\rho(1-\rho)} \right)^{\frac{1}{(1-\rho)^2}} \right]^{1-\rho} \\ &\quad \cdot \left(\sum_x P_X(x) P_X(x)^{-1} Q_X(x)^{\rho} \right)^{\rho} \quad (321) \\ &= \left[\sum_x P_X(x) P_X(x)^{\frac{\rho}{1-\rho}} Q_X(x)^{-\frac{\rho}{1-\rho}} \right. \\ &\quad \cdot \left. \left(\sum_y P_{Y|X}(y|x) (P_{Y|X}(y|x)^{\rho(1-\rho)} Q_{Y|X}(y|x)^{-\rho(1-\rho)}) \right)^{\frac{1}{(1-\rho)^2}} \right]^{1-\rho} \cdot 1^{\rho} \\ &\leq \sum_x P_X(x) P_X(x)^{\rho} Q_X(x)^{-\rho} \left[\right. \\ &\quad \left. \sum_y P_{Y|X}(y|x) \left(P_{Y|X}(y|x)^{\rho(1-\rho)} Q_{Y|X}(y|x)^{-\rho(1-\rho)} \right) \right]^{\frac{1}{1-\rho}} \quad (322) \end{aligned}$$

$$\begin{aligned} &\leq \sum_x P_X(x) P_X(x)^{\rho} Q_X(x)^{-\rho} \left[\right. \\ &\quad \left. \sum_y P_{Y|X}(y|x) \left(P_{Y|X}(y|x)^{\rho(1-\rho)} Q_{Y|X}(y|x)^{-\rho(1-\rho)} \right)^{\frac{1}{1-\rho}} \right] \quad (323) \\ &= \sum_x P_X(x) P_X(x)^{\rho} Q_X(x)^{-\rho} \left[\right. \\ &\quad \left. \sum_y P_{Y|X}(y|x) \left(P_{Y|X}(y|x)^{\rho} Q_{Y|X}(y|x)^{-\rho} \right) \right] \\ &= \sum_{x,y} P_{X,Y}(x, y)^{1+\rho} Q_{X,Y}(x, y)^{-\rho} = e^{\psi(\rho|P_{X,Y}\|Q_{X,Y})}. \quad (324) \end{aligned}$$

■

APPENDIX B

EXISTENCE OF CODE REQUIRED IN THEOREM 32 WITH $\epsilon = 0$

In this appendix, we show the existence of Slepian-Wolf data compression code satisfying the condition (107) required in Theorem 32 with $\epsilon = 0$ in the two-terminal and i.i.d. case. For this purpose, we assume that the random variables (S_1^n, S_2^n) are subject to the n -fold i.i.d. distribution of a given non-uniform joint distribution of S_1 and S_2 . For this purpose, we recall the definition of achievable rate pair for Slepian-Wolf compression.

Definition 88: A rate pair (R_1, R_2) is called *achievable* when there exists a sequence of encoders $\varphi^n = (\varphi_1^n, \varphi_2^n)$ ($\varphi_i^n : S_i^n \rightarrow \{1, \dots, \lceil e^{nR_i} \rceil\}$) and decoders $\hat{\varphi}^n$ ($\hat{\varphi}^n : \{1, \dots, \lceil e^{nR_1} \rceil\} \times \{1, \dots, \lceil e^{nR_2} \rceil\} \rightarrow S_1^n \times S_2^n$) such that the decoding error probability $\varepsilon(\varphi^n, \hat{\varphi}^n)$ satisfies

$$\lim_{n \rightarrow \infty} \varepsilon(\varphi^n, \hat{\varphi}^n) = 0. \quad (325)$$

Then, we prepare the following lemma.

Lemma 89: Let (R_1, R_2) be a pair of achievable rates for Slepian-Wolf compression satisfying $R_1 + R_2 = H(S_1, S_2)$. When the compression rate pair $(R_{1,n}, R_{2,n})$ behaves as $R_{1,n} = R_1 + \frac{c_1}{n^t}$ and $R_{2,n} = R_2 + \frac{c_2}{n^t}$ with $0 < t < 1/2$ and $c_1 > 0, c_2 > 0$, there exists a sequence of Slepian-Wolf codes $(\varphi^n, \hat{\varphi}^n) = ((\varphi_1^n, \varphi_2^n), \hat{\varphi}^n)$ for any positive integer n such that φ_i^n is a map from S_i^n to $\{1, \dots, \lceil e^{nR_{i,n}} \rceil\}$ for $i = 1, 2$ and the decoding error probability $\varepsilon(\varphi^n, \hat{\varphi}^n)$ satisfies

$$\begin{aligned} & \liminf_{n \rightarrow \infty} -n^{2t-1} \log \varepsilon(\varphi^n, \hat{\varphi}^n) \\ & \geq \min \left(\lambda \frac{c_1^2}{2V(S_1)}, \lambda \frac{c_2^2}{2V(S_2|S_1)}, \right. \\ & \quad \left. (1-\lambda) \frac{c_2^2}{2V(S_2)}, (1-\lambda) \frac{c_1^2}{2V(S_1|S_2)} \right), \quad (326) \end{aligned}$$

where $V(S_2|S_1) := \sum_{s_1, s_2} P_{S_1, S_2}(s_1, s_2) (\log P_{S_2|S_1}(s_2|s_1) - H(S_2|S_1))^2$ and $\lambda \in [0, 1]$ is the real number satisfying that

$$(R_1, R_2) = \lambda(H(S_1), H(S_2|S_1)) + (1-\lambda)(H(S_1|S_2), H(S_2)). \quad (327)$$

Further, when $R_1 = H(S_1)$ and $R_2 = H(S_2|S_1)$ and the compression rates $(R_{1,n}, R_{2,n})$ behaves as $R_{1,n} = H(S_1) + \frac{c_1}{n^t}$ and $R_{2,n} = H(S_2|S_1) + \frac{c_2}{n^t}$ with $0 < t < 1/2$ and $c_1 > 0, c_2 > 0$, there exists a sequence of Slepian-Wolf codes $(\varphi^n, \hat{\varphi}^n)$ such that the decoding error probability $\varepsilon(\varphi^n, \hat{\varphi}^n)$ satisfies

$$\liminf_{n \rightarrow \infty} -n^{2t-1} \log \varepsilon(\varphi^n, \hat{\varphi}^n) \geq \min \left(\frac{c_1^2}{2V(S_1)}, \frac{c_2^2}{2V(S_2|S_1)} \right). \quad (328)$$

We will prove Lemma 89 after preparing several lemmas. Using Lemma 89, we make a Slepian-Wolf compression whose compressed data satisfies the SACU condition. Let (R_1, R_2) be a pair of achievable rates for Slepian-Wolf compression satisfying $R_1 + R_2 = H(S_1, S_2)$. Then, let $\varphi^n = (\varphi_1^n, \varphi_2^n)$ and $\hat{\varphi}^n$ be the Slepian-Wolf encoders and the Slepian-Wolf decoder given in Lemma 89 with the case of $c_1 = R_1 c$ and $c_2 = R_2 c$. We choose the integer $m_n := \lfloor \frac{n}{1+\frac{c}{n}} \rfloor = \lfloor \frac{R_1 n}{R_1 + R_1 \frac{c}{n}} \rfloor = \lfloor \frac{R_2 n}{R_2 + R_2 \frac{c}{n}} \rfloor = \lfloor \frac{R_1 n}{R_{1,n}} \rfloor = \lfloor \frac{R_2 n}{R_{2,n}} \rfloor$ for $0 < t < \frac{1}{2}$ and $c > 0$. Then, we obtain the Slepian-Wolf encoders

$\varphi_i^{m_n} : S_i^{m_n} \rightarrow \{1, \dots, \lceil e^{nR_i} \rceil\}$ and the Slepian-Wolf decoder $\hat{\varphi}^{m_n} : \{1, \dots, \lceil e^{nR_1} \rceil\} \times \{1, \dots, \lceil e^{nR_2} \rceil\} \rightarrow S_1^{m_n} \times S_2^{m_n}$. Using the code, we define the Slepian-Wolf encoders $\varphi_{i,u}^n : S_i^{m_n} \rightarrow \{1, \dots, \lceil e^{nR_i} \rceil\}$ and the Slepian-Wolf decoder $\hat{\varphi}_u^n : \{1, \dots, \lceil e^{nR_1} \rceil\} \times \{1, \dots, \lceil e^{nR_2} \rceil\} \rightarrow S_1^{m_n} \times S_2^{m_n}$ by

$$\varphi_{i,u}^n(s^{m_n}) := \varphi_i^{m_n}(s^{m_n}) \quad (329)$$

$$\hat{\varphi}_u^n(x_1, x_2) := \hat{\varphi}^{m_n}(x_1, x_2). \quad (330)$$

Then, due to Lemma 89, since $m_n(R_1 + R_1 \frac{c}{n}) = nR_1$ and $m_n(R_2 + R_2 \frac{c}{n}) = nR_2$, the code $((\varphi_{1,u}^n, \varphi_{2,u}^n), \hat{\varphi}_u^n)$ satisfies the condition (107) in Theorem 32 with $\epsilon = 0$. Theorem 32 guarantees that the compressed data satisfies the SACU condition.

Now, in order to show Lemma 89, we prepare several lemmas.

Lemma 90 ([36], [37], [38]): For a given compression rate $R_2 > 0$, there exists a pair of the encoder φ^n and the decoder $\hat{\varphi}^n$ of the random variable S_2^n with the side information S_1^n such that the decoding error probability $\varepsilon(\varphi^n, \hat{\varphi}^n)$ satisfies

$$\varepsilon(\varphi^n, \hat{\varphi}^n) \leq e^{-n(\rho R_2 - E_0(-\rho|S_2|S_1))} \quad (331)$$

for any $\rho \in (0, 1]$, where

$$E_0(\rho|S_2|S_1) := \log \sum_{s_1} \left(\sum_{s_2} P_{S_1, S_2}(s_1, s_2)^{\frac{1}{1-\rho}} \right)^{1-\rho}. \quad (332)$$

Note that when there is no side information, we have

$$E_0(-\rho|S_2) = \rho H_{\frac{1}{1-\rho}}(S_2). \quad (333)$$

Lemma 91: The quantity $E_0(-\rho|S_2|S_1)$ has the expansion

$$E_0(-\rho|S_2|S_1) = \rho H(S_2|S_1) + \frac{\rho^2}{2} V(S_2|S_1) \quad (334)$$

with small ρ . In particular, the quantity $\rho H_{\frac{1}{1-\rho}}(S_1)$ has the expansion

$$\rho H_{\frac{1}{1-\rho}}(S_1) = \rho H(S_1) + \frac{\rho^2}{2} V(S_1) \quad (335)$$

with small ρ and $V(S_1) := \sum_{s_1} P_{S_1}(s_1) (\log P_{S_1}(s_1) - H(S_1))^2$.

Proof: Take the Taylor expansion of $e^{E_0(\rho|S_2|S_1)}$ as

$$\begin{aligned} & e^{E_0(-\rho|S_2|S_1)} \\ & = 1 + \rho H(S_2|S_1) \\ & \quad + \frac{\rho^2}{2} \sum_{s_1, s_2} P_{S_1, S_2}(s_1, s_2) (\log P_{S_2|S_1}(s_2|s_1))^2 + o(\rho^2). \quad (336) \end{aligned}$$

Taking the logarithm, we obtain (334). \blacksquare

Lemma 92: Let (R_1, R_2) belong to the Slepian-Wolf compression region of (S_1^n, S_2^n) . We choose the rates R'_1, R'_2, R''_1 , and R''_2 and the real number $\lambda \in [0, 1]$ such that

$$(R_1, R_2) = \lambda(R'_1, R'_2) + (1-\lambda)(R''_1, R''_2). \quad (337)$$

Then, there exists a pair of the Slepian-Wolf encoder φ^n and the decoder $\hat{\varphi}^n$ such that the decoding error probability $\varepsilon(\varphi^n, \hat{\varphi}^n)$ satisfies

$$\begin{aligned} & \varepsilon(\varphi^n, \hat{\varphi}^n) \\ & \leq \inf_{\rho \in (0, 1]} e^{-\lambda n(\rho R'_1 - \rho H_{\frac{1}{1-\rho}}(S_1))} + \inf_{\rho \in (0, 1]} e^{-\lambda n(\rho R'_2 - E_0(-\rho|S_2|S_1))} \\ & \quad + \inf_{\rho \in (0, 1]} e^{-(1-\lambda)n(\rho R''_1 - E_0(-\rho|S_1|S_2))} + \inf_{\rho \in (0, 1]} e^{-(1-\lambda)n(\rho R''_2 - \rho H_{\frac{1}{1-\rho}}(S_2))}, \quad (338) \end{aligned}$$

Also, there exists a pair of the Slepian-Wolf encoder φ^n and the decoder $\hat{\varphi}^n$ such that the decoding error probability $\varepsilon(\varphi^n, \hat{\varphi}^n)$ satisfies

$$\varepsilon(\varphi^n, \hat{\varphi}^n) \leq \inf_{\rho \in (0,1]} e^{-n(\rho R_1 - \rho H_{\frac{1}{1-\rho}}(S_1))} + \inf_{\rho \in (0,1]} e^{-n(\rho R_2 - E_0(-\rho|S_2|S_1))}, \quad (339)$$

Proof: First, we show the existence of a sequence of codes satisfying (339). We apply the usual data compression for S_2^n , and the data compression given in Lemma 90 for S_1^n . The decoder is given by combination of the respective decoders. Since the decoding error probability is bounded by the sum of the decoding error probabilities of S_1^n and S_2^n , we obtain (339).

Next, we show the existence of a sequence of codes satisfying (338). We divide n symbols into two parts, λn symbols and $(1-\lambda)n$ symbols. We apply the construction given in the previous paragraph with the rates (R'_1, R'_2) to the first part, and apply the same construction with the rates (R''_1, R''_2) to the second part. Due to Lemma 90, the decoding error probability of the first part is less than $\inf_{\rho \in (0,1]} e^{-\lambda n(\rho R'_1 - \rho H_{\frac{1}{1-\rho}}(S_1))} + \inf_{\rho \in (0,1]} e^{-\lambda n(\rho R'_2 - E_0(-\rho|S_2|S_1))}$, and the decoding error probability of the second part is less than $\inf_{\rho \in (0,1]} e^{-(1-\lambda)n(\rho R''_1 - E_0(-\rho|S_1|S_2))} + \inf_{\rho \in (0,1]} e^{-(1-\lambda)n(\rho R''_2 - \rho H_{\frac{1}{1-\rho}}(S_2))}$. Then, we obtain (338). ■

Proof of Lemma 89: First, we consider the case when $R_1 = H(S_1)$ and $R_2 = H(S_2|S_1)$. Since $R_{1,n} = H(S_1) + \frac{c_1}{n}$ and $R_{2,n} := H(S_2|S_1) + \frac{c_2}{n}$, we can show that

$$\lim_{n \rightarrow \infty} -n^{2t-1} \log \inf_{\rho \in (0,1]} e^{-n(\rho R_{1,n} - \rho H_{\frac{1}{1-\rho}}(S_1))} = \frac{c_1^2}{2V(S_1)} \quad (340)$$

$$\lim_{n \rightarrow \infty} -n^{2t-1} \log \inf_{\rho \in (0,1]} e^{-n(\rho R_{2,n} - E_0(-\rho|S_2|S_1))} = \frac{c_2^2}{2V(S_2|S_1)}. \quad (341)$$

Since the proof of (340) is similar to those of (341), we show only (340). When ρ is sufficiently small, due to Lemma 91, we have

$$\begin{aligned} \rho R_{1,n} - \rho H_{\frac{1}{1-\rho}}(S_1) &\cong \rho \frac{c_1}{n^t} - \frac{\rho^2}{2} V(S_1) \\ &= -\frac{V(S_1)}{2} \left(\rho - \frac{c_1}{V(S_1)n^t} \right)^2 + \frac{c_1^2}{2V(S_1)n^{2t}}. \end{aligned} \quad (342)$$

Hence, $\inf_{\rho \in (0,1]} e^{-n(\rho R'_{1,n} - \rho H_{\frac{1}{1-\rho}}(S_1))} \cong e^{-n \frac{c_1^2}{2V(S_1)n^{2t}}}$, which implies (340). Then, we apply the evaluation (339) for the decoding error probability in Lemma 92 to the case when R_1, R_2 are $R_{1,n}, R_{2,n}$. Combining the relations (340) and (341), we obtain (328).

Next, we show the general case. We choose $R'_{1,n} := H(S_1) + \frac{c_1}{n}$, $R'_{2,n} := H(S_2|S_1) + \frac{c_2}{n}$, $R''_{1,n} := H(S_1|S_2) + \frac{c_1}{n}$, $R''_{2,n} := H(S_2) + \frac{c_2}{n}$. Then, we obtain

$$(R_{1,n}, R_{2,n}) = \lambda(R'_{1,n}, R'_{2,n}) + (1-\lambda)(R''_{1,n}, R''_{2,n}). \quad (343)$$

Then, similar to (340) and (341), we can show that

$$\lim_{n \rightarrow \infty} -n^{2t-1} \log \inf_{\rho \in (0,1]} e^{-\lambda n(\rho R'_{1,n} - \rho H_{\frac{1}{1-\rho}}(S_1))} = \lambda \frac{c_1^2}{2V(S_1)} \quad (344)$$

$$\lim_{n \rightarrow \infty} -n^{2t-1} \log \inf_{\rho \in (0,1]} e^{-\lambda n(\rho R'_{2,n} - E_0(-\rho|S_2|S_1))} = \lambda \frac{c_2^2}{2V(S_2|S_1)} \quad (345)$$

$$\lim_{n \rightarrow \infty} -n^{2t-1} \log \inf_{\rho \in (0,1]} e^{-(1-\lambda)n(\rho R''_{1,n} - E_0(-\rho|S_1|S_2))} = (1-\lambda) \frac{c_1^2}{2V(S_2)} \quad (346)$$

$$\lim_{n \rightarrow \infty} -n^{2t-1} \log \inf_{\rho \in (0,1]} e^{-(1-\lambda)n(\rho R''_{2,n} - \rho H_{\frac{1}{1-\rho}}(S_2))} = (1-\lambda) \frac{c_1^2}{2V(S_1|S_2)}. \quad (347)$$

We apply the evaluation (338) for the decoding error probability in Lemma 92 to the case when R'_1, R'_2, R''_1, R''_2 are $R'_{1,n}, R'_{2,n}, R''_{1,n}, R''_{2,n}$. Combining the relations (344), (345), (346) and (347), we obtain (326). ■

APPENDIX C

EQUIVALENCE BETWEEN THE SWACU CONDITION AND THE WACU CONDITION

In Subsection VIII-A, we have introduced three asymptotic conditional uniformity conditions. The aim of this appendix is to show the equivalence between the SWACU condition and the WACU condition, which was used in our proof of Theorem 37.

Lemma 93: Let A_n be a random variable on the set \mathcal{A}_n with the cardinality e^{nR} and B_n be another random variable for any positive inter n . Then, the relation

$$\lim_{n \rightarrow \infty} \frac{1}{n} H(A_n|B_n) = R \quad (348)$$

holds, if and only if

$$\lim_{n \rightarrow \infty} \frac{1}{n} H_{1+\alpha/n}(A_n|B_n) = R \quad (349)$$

for any $\alpha > 0$.

Lemma 93 will be shown after Lemma 94, which is used in the proof of Lemma 93. Thanks to Lemma 93, we can replace the WACU condition (99) by the SWACU condition (100). Indeed, in order to apply our results in Section VII to the proof of Theorem 37, we need evaluation conditional Rényi entropy instead of conditional entropy, as is discussed around (122). Lemma 93 provides the evaluation of conditional Rényi entropy (349) from the evaluation of conditional entropy (348). Hence, Lemma 93 is useful for the application of our results in Section VII to the asymptotic setting.

Lemma 94: Let A be a random variable on the set \mathcal{A} with the cardinality M and B be another random variable. For arbitrary $\epsilon_1 > 0$ and $1 \geq \epsilon_2 > 0$, we define the subset of joint distributions for A and B as

$$\mathcal{P}_{\epsilon_1, \epsilon_2, M}^{A|B} := \{P_{A,B}|P_{A,B}\{(a,b)|-\log P_{A|B}(a|b) \leq \log M - \epsilon_1\} \leq \epsilon_2\}. \quad (350)$$

Then,

$$\max_{P_{A,B} \in \mathcal{P}_{\epsilon_1, \epsilon_2, M}^{AB}} H(A|B) \leq \log M - \epsilon_2(e^{-\epsilon_1} - 1 + \epsilon_1) \quad (351)$$

$$\min_{P_{A,B} \in \mathcal{P}_{\epsilon_1, \epsilon_2, M}^{AB}} H_{1+\rho}(A|B) \geq -\frac{1}{\rho} \log\left((1 - \epsilon_2) \frac{e^{\rho\epsilon_1}}{M^\rho} + \epsilon_2\right). \quad (352)$$

Here, since the region $\mathcal{P}_{\epsilon_1, \epsilon_2, M}^{AB}$ is compact, the above maximum and the above minimum exist.

Proof of Lemma 94: For an arbitrary integer k , we define the set

$$\mathcal{P}_{\epsilon_1, \epsilon_2, M, k}^A := \left\{ P_A \left| \begin{array}{l} P_A\{a | -\log P_A(a) \leq \log M - \epsilon_1\} \leq \epsilon_2, \\ \left| \{a | -\log P_A(a) \leq \log M - \epsilon_1\} \right| = k \end{array} \right. \right\}$$

$$\mathcal{P}_{\epsilon_1, \epsilon_2, M}^A := \{P_A | P_A\{a | -\log P_A(a) \leq \log M - \epsilon_1\} \leq \epsilon_2\},$$

and define the function

$$f(x) := \epsilon_2(\log x - \log \epsilon_2) + (1 - \epsilon_2)(\log(M - x) - \log(1 - \epsilon_2))$$

for $\epsilon_2 \in (0, 1)$. The set $\mathcal{P}_{\epsilon_1, \epsilon_2, M, k}^A$ is a non-empty set only when the integer k belongs to $[0, \epsilon_2 M e^{-\epsilon_1}]$. Under the above choice of k , we have

$$\max_{P_A \in \mathcal{P}_{\epsilon_1, \epsilon_2, M, k}^A} H(A) = f(k)$$

and

$$\max_{P_A \in \mathcal{P}_{\epsilon_1, \epsilon_2, M}^A} H(A) = \max_{k \in [0, \epsilon_2 M e^{-\epsilon_1}]} f(k),$$

where k is restricted to an integer in the maximum. Taking the derivative, we have

$$f'(x) = \frac{\epsilon_2}{x} - \frac{1 - \epsilon_2}{M - x},$$

which is positive when $x < M\epsilon_2$. Hence,

$$\begin{aligned} & \max_{P_A \in \mathcal{P}_{\epsilon_1, \epsilon_2, M}^A} H(A) \\ & \leq f(\epsilon_2 M e^{-\epsilon_1}) \\ & = \epsilon_2(\log M - \epsilon_1) + (1 - \epsilon_2)(\log M + \log(1 - \epsilon_2 e^{-\epsilon_1}) - \log(1 - \epsilon_2)) \\ & = \log M - \epsilon_2 \epsilon_1 + (1 - \epsilon_2) \log \left[1 + \frac{\epsilon_2(1 - e^{-\epsilon_1})}{1 - \epsilon_2} \right] \\ & \leq \log M - \epsilon_2 \epsilon_1 + (1 - \epsilon_2) \frac{\epsilon_2(1 - e^{-\epsilon_1})}{1 - \epsilon_2} \\ & = \log M - \epsilon_2(e^{-\epsilon_1} - 1 + \epsilon_1). \end{aligned}$$

Since $\log M - \epsilon_2(e^{-\epsilon_1} - 1 + \epsilon_1)$ is an affine function of ϵ_2 , we obtain (351).

On the other hand, using the set $\Omega := \{a | -\log P_A(a) \leq \log M - \epsilon_1\}$, we have

$$\begin{aligned} \max_{P_A \in \mathcal{P}_{\epsilon_1, \epsilon_2, M}^A} e^{-\rho H_{1+\rho}(A)} & = \sum_{a \in \Omega^c} (P_A(a))^{1+\rho} + \sum_{a \in \Omega} (P_A(a))^{1+\rho} \\ & \leq (1 - \epsilon_2) \frac{e^{\rho\epsilon_1}}{M^\rho} + \epsilon_2^{1+\rho} \leq (1 - \epsilon_2) \frac{e^{\rho\epsilon_1}}{M^\rho} + \epsilon_2. \end{aligned}$$

Since $(1 - \epsilon_2) \frac{e^{\rho\epsilon_1}}{M^\rho} + \epsilon_2$ is a linear function of ϵ_2 , we obtain

$$\max_{P_{A|B} \in \mathcal{P}_{\epsilon_1, \epsilon_2, M}^{AB}} e^{-\rho H_{1+\rho}(A|B)} \leq (1 - \epsilon_2) \frac{e^{\rho\epsilon_1}}{M^\rho} + \epsilon_2,$$

which implies (352). \blacksquare

Proof of Lemma 93: Since (349) implies (348), we only show (349) from (348). For an arbitrary small number $\epsilon > 0$, we define the probability

$$\delta_n := P_{A^n, B^n}\{(a, b) | -\frac{1}{n} \log P_{A^n|B^n}(a|b) \leq R - \epsilon\}.$$

Applying Eq. (351) of Lemma 94 to the case when $\epsilon_1 = n\epsilon$ and $\epsilon_2 = \delta_n$, we obtain

$$H(A_n|B_n) \leq nR - \delta_n(e^{-n\epsilon} - 1 + n\epsilon).$$

That is,

$$\delta_n \leq \frac{R - \frac{1}{n}H(A_n|B_n)}{\frac{e^{-n\epsilon} - 1}{n} + \epsilon}. \quad (353)$$

Thus, $\lim_{n \rightarrow \infty} \delta_n = 0$. Hence, Eq. (352) of Lemma 94 guarantees that

$$H_{1+\alpha/n}(A_n|B_n) \geq -\frac{n}{\alpha} \log((1 - \delta_n)e^{\alpha(\epsilon - R)} + \delta_n). \quad (354)$$

Thus,

$$\begin{aligned} \liminf_{n \rightarrow \infty} \frac{1}{n} H_{1+\alpha/n}(A_n|B_n) & \geq \liminf_{n \rightarrow \infty} -\frac{1}{\alpha} \log((1 - \delta_n)e^{\alpha(\epsilon - R)} + \delta_n) \\ & = R - \epsilon. \end{aligned}$$

Since $\epsilon > 0$ is arbitrary,

$$\liminf_{n \rightarrow \infty} \frac{1}{n} H_{1+\alpha/n}(A_n|B_n) \geq R.$$

Since the cardinality of \mathcal{A}_n is e^{nR} , we have $\frac{1}{n} H_{1+\alpha/n}(A_n|B_n) \leq R$. Hence,

$$\lim_{n \rightarrow \infty} \frac{1}{n} H_{1+\alpha/n}(A_n|B_n) = R.$$

Combining relation (5), we obtain the desired argument. \blacksquare

APPENDIX D

EXTENSION TO GENERAL MEASURABLE SPACES

A. Information quantities

Our results has been obtained based on discrete sets, i.e., sets with countable elements. Here, we explain how our results are extended to the case of measurable spaces, which contain continuous sets. Firstly, we state the assumptions used in Appendix D. As before, \mathcal{X} is the input alphabet of the channel and \mathcal{Z} is the output alphabet to Eve. In general, a channel from \mathcal{X} to \mathcal{Z} is described as a collection of conditional probability measures $\mu_{\mathcal{Z}|X=x}$ on \mathcal{Z} for all inputs $x \in \mathcal{X}$, and $\mu_{\mathcal{Z}|X=x}$ might not have a probability density for some $x \in \mathcal{X}$. In this appendix, however, we assume that there exists a finite measure $\nu_{\mathcal{Z}}$ on \mathcal{Z} such that for all $x \in \mathcal{X}$, $\mu_{\mathcal{Z}|X=x}$ is absolutely continuous with respect $\nu_{\mathcal{Z}}$. In the following $P_{\mathcal{Z}|X}(\cdot|x)$ denotes the Radon-Nikodym derivative $d\mu_{\mathcal{Z}|X=x}/d\nu_{\mathcal{Z}}$. We also make the same assumption on the channel from Alice to Bob.

In addition, as before, we consider probability measures η on $\mathcal{U} \times \mathcal{V} \times \mathcal{X}$. We assume that there exist finite measures $\nu_{\mathcal{U}}$ on \mathcal{U} , $\nu_{\mathcal{V}}$ on \mathcal{V} and $\nu_{\mathcal{X}}$ on \mathcal{X} such that η is absolutely continuous with respect to the product measure $\nu_{\mathcal{U}} \times \nu_{\mathcal{V}} \times \nu_{\mathcal{X}}$.

Under this assumption we can denote by P_{UVX} the Radon-Nikodym derivative $d\eta/d(\nu_U \times \nu_V \times \nu_X)$, and marginal probability densities P_U , etc. and conditional probability densities $P_{V|U}$, etc. can be computed from P_{UVX} . In the following, dv , dz , etc. denote $d\nu_V$, $d\nu_Z$, etc. assumed above.

Firstly, we give the definition of the information quantities in the general measurable case. Although $E_0(\rho|P_{Z|V}, P_V)$ and $E_0(\rho|P_{Z|V}, P_{V|U}, P_U)$ are defined for distributions P_V and P_U and conditional distributions $P_{Z|V}$ and $P_{V|U}$ with discrete sets in (11), they can be defined as follows even when \mathcal{Z} , \mathcal{V} , and \mathcal{U} are measurable spaces in the sense of [47, Theorem 32.2]. Then, we define

$$\begin{aligned} & E_0(\rho|P_{Z|V}, P_V) \\ & := \log \int_{\mathcal{Z}} dz \left(\int_{\mathcal{V}} dv P_V(v) (P_{Z|V}(z|v))^{1/(1-\rho)} \right)^{1-\rho}, \quad (355) \\ & E_0(\rho|P_{Z|V}, P_{V|U}, P_U) \\ & := \log \int_{\mathcal{U}} du \int_{\mathcal{Z}} dz \left(\int_{\mathcal{V}} dv P_{V|U}(v|u) (P_{Z|V}(z|v))^{1/(1-\rho)} \right)^{1-\rho}. \end{aligned}$$

The above definition formally depends on the choices of the measures dz , du , dv . But in the next paragraph we will explain the above values are independent of the choice of measures dz , du , dv .

Now, suppose that we choose other measures dz' , du' , dv' so that the measures dz' , du' , dv' and the original measures dz , du , dv are absolutely continuous with respect to each other, respectively. As is shown in the left hand side of [43, p.7740], even when these information quantities are defined with the measures dz' , du' , dv' , these information quantities have the same values as those defined with the original measures dz , du , dv . So, these information quantities do not depend on the choice of the measures dz , du , dv whenever the measures and the original measures are absolutely continuous with respect to each other.

When Q and P are probability density functions on a measurable space \mathcal{Z} with respect to a common finite measure dz , $\psi(\rho|Q||P)$ is defined as

$$\psi(\rho|Q||P) := \log \int_{\mathcal{Z}} dz Q(z)^{1+\rho} P(z)^{-\rho}.$$

Further, $\psi(\rho|P_{Z|V}, P_V)$ and $\psi(\rho|P_{Z|V}, P_{V|U}, P_U)$ are defined as follows.

$$\begin{aligned} & \psi(\rho|P_{Z|V}, P_V, P_U) \\ & = \log \int_{\mathcal{V}} dv P_V(v) \int_{\mathcal{Z}} dz P_{Z|V}(z|v)^{1+\rho} P_Z(z)^{-\rho}, \quad (356) \\ & \psi(\rho|P_{Z|V}, P_{V|U}, P_U) \\ & = \log \int_{\mathcal{U}} du P_U(u) \int_{\mathcal{V}} dv P_{V|U}(v|u) \int_{\mathcal{Z}} dz P_{Z|V}(z|v)^{1+\rho} P_{Z|U}(z|u)^{-\rho}. \quad (357) \end{aligned}$$

Similar to the information quantities $E_0(\rho|P_{Z|V}, P_V)$ and $E_0(\rho|P_{Z|V}, P_{V|U}, P_U)$, we can show that the information quantities $\psi(\rho|P_{Z|V}, P_V, P_U)$ and $\psi(\rho|P_{Z|V}, P_{V|U}, P_U)$ do not depend on the choice of the measures dz , du , dv whenever the measures and the original measures are absolutely continuous with respect to each other.

The above quantities can be defined for a channel. When the input and output systems \mathcal{Z} and \mathcal{V} are measurable spaces, a channel W is defined as a set of probability density functions $\{W_v\}_{v \in \mathcal{V}}$ on \mathcal{Z} . That is, substituting W into a conditional probability density function $P_{Z|V}$ as $P_{Z|V}(z|v) = W_v(z)$, we define the above information quantities for the channel W . So, when the channels W^Z and W^Y satisfy the above conditions, the code construction and security evaluation given in the next subsection work well. Note that the above generalization works well even when \mathcal{V} is a finite set because a finite set is also a measurable space.

B. Code construction and security evaluation

Under the above extension, our results can be extended as follows. Firstly, we focus on Theorem 14. Assume that W is a channel from a measurable space \mathcal{X} to a measurable space \mathcal{Y} and that A is a discrete random variable on a finite set \mathcal{A} subject to the distribution P_A . Theorem 14 holds even under this assumption, whose proof can be done by replacing \sum_x and \sum_y by $\int_{\mathcal{X}} dx$ and $\int_{\mathcal{Y}} dy$. Theorem 17 and Corollary 18 also hold with a slightly different extension. Assume that W is a channel from a finite-dimensional vector space \mathcal{X} over \mathbb{F}_q to a measurable space \mathcal{Y} and that A is a discrete random variable on a finite-dimensional vector space \mathcal{A} over \mathbb{F}_q subject to P_A . Then, Theorem 17 and Corollary 18 hold even under this assumption, whose proof can be done by replacing \sum_y by $\int_{\mathcal{Y}} dy$.

Now, we consider the extension of Code Ensemble 1. Assume that $\mathcal{X} = \mathcal{V}$, \mathcal{Y} , \mathcal{Z} , and \mathcal{U} are measurable, and that the private and common messages S_p and S_c take values in finite sets. Then, we can apply Code Ensemble 1 to the above situation. Hence, Lemma 12 holds even under this assumption because the proof by Kaspi and Merhav [21, Section II] is still valid under this assumption.

Next, we proceed to the extension of Code Ensemble 2. Assume that \mathcal{X} , \mathcal{Y} , \mathcal{Z} , \mathcal{V} , and \mathcal{U} are measurable, and that all messages S_0, S_1, \dots, S_T take values in finite sets. Then, we can apply Code Ensemble 2 to the above situation. Hence, Theorem 20 holds even under this assumption because (57) holds under this assumption.

Then, we extend the contents of Section VII. We consider the extension of Code Ensemble 3. Assume that \mathcal{X} , \mathcal{Y} , \mathcal{Z} , \mathcal{V} , and \mathcal{U} are measurable, and that \mathcal{B}_1 and \mathcal{B}_2 are finite Abelian groups. In this case, all messages S_0, S_1, \dots, S_T take values in finite sets. Then, we can apply Code Ensemble 3 to the above situation. First, notice that Theorem 12 still holds in the above situation. Hence, Lemma 21 and Theorem 22 hold even under this assumption, whose proof can be done by applying the extension of Theorems 12 and 17. Lemma 24 holds with a slightly different extension. That is, Lemma 24 holds when the sets \mathcal{U} and \mathcal{V} are finite set, i.e., only the set \mathcal{Z} is allowed to be a general measurable space. This is because we need to consider the cardinalities of the subsets in \mathcal{U} and \mathcal{V} . Since the contents of Sections V and VI are extended to the case of measurable spaces in the above way, the contents of Sections VIII and IX also can be extended to the case of measurable spaces in the same way.

In Section XI, we have proposed several types of practical code constructions. Code Constructions 6 and 7 can be applied to the channel $P_{Z|V}$ from a measurable space \mathcal{V} to a measurable space \mathcal{Z} . In these constructions, since the code φ_p is given, we can restrict the set \mathcal{V} to the finite subset given as the image of the map φ_p . Hence, we can apply Lemma 24 with the above extension in this context.

When the above discussion is applied to the wire-tap channel model, we obtain an extension of existing results to the case of the asymptotic uniform dummy message. That is, we consider the case with no common messages and $T = 2$ when S_1 corresponds to the message to be secretly sent to Bob, and S_2 does to the dummy message making S_1 ambiguous to Eve. For a given rate R_1 of secret message and a given rate R_2 of dummy message, the RHS of (115) coincides with the Gallager exponents, the RHS of (155) coincides with the RHS of (59) in [15], and the RHS of (157) coincides with the exponents of the RHS of (15) in [17].

C. Gaussian case

Finally, when the channel $P_{Y|Z|X}$ is a degraded Gaussian channel as (358), we demonstrate how the strong security can be shown for the wire-tap channel, which is given as the case with no common messages and $T = 2$ when S_1 corresponds to the message S to be secretly sent to Bob, and S_2 does to the dummy message A making S ambiguous to Eve. Assume that \mathcal{X} , \mathcal{Y} , and \mathcal{Z} are the set of real numbers. So, we choose the measures dx , dy , and dz to be the Lebesgue measure. Then, we assume that the conditional probability density functions corresponding to the channels are

$$P_{Y|X}(y|x) := \frac{1}{\sqrt{2\pi v_1}} e^{-\frac{(y-x)^2}{2v_1}}, \quad P_{Z|X}(z|x) := \frac{1}{\sqrt{2\pi v_2}} e^{-\frac{(z-x)^2}{2v_2}}, \quad (358)$$

where $v_2 > v_1$. Since the channel is degraded, we do not need to introduce random variables U and V . Now, we choose the probability density function P_X to be $P_X(x) = \frac{1}{\sqrt{2\pi v_3}} e^{-\frac{x^2}{2v_3}}$. Then,

$$E_0(\rho|P_{Z|X}, P_X) = \frac{\rho}{2} \log\left(1 + \frac{v_3}{(1-\rho)v_2}\right), \quad (359)$$

$$\begin{aligned} \psi(\rho|P_{Z|X=x}, P_Z) &= \frac{(1+\rho)\rho}{2(v_2 + (1+\rho)v_3)} x^2 - \frac{\rho}{2} \log v_2 \\ &\quad + \frac{1+\rho}{2} \log(v_2 + v_3) - \frac{1}{2} \log(v_2 + (1+\rho)v_3), \end{aligned} \quad (360)$$

$$\begin{aligned} \psi(\rho|P_{Z|X}, P_X) &= \frac{1+\rho}{2} \log(v_2 + v_3) \\ &\quad - \frac{1}{2} \log(v_2 + (1-\rho^2)v_3) - \frac{\rho}{2} \log v_2 \\ &= \frac{\rho}{2} \log\left(1 + \frac{v_3}{v_2}\right) - \frac{1}{2} \log\left(1 - \frac{v_3}{v_2+v_3}\rho^2\right). \end{aligned} \quad (361)$$

Hereafter, we denote the average leaked information under our code Φ by $I(S; E)[\Phi]$. Assume that we use the Gaussian channel $P_{Y|Z|X}$ n times, and that the rates of secret message S and dummy message A are R_1 and R_2 , respectively. When the

dummy message A has the Rényi entropy $H_{1+\rho}(A)$, Theorem 20 guarantees that

$$\mathbf{E}_\Phi[e^{\rho I(S; E)}] \leq 1 + e^{-\rho H_{1+\rho} + n(\frac{\rho}{2} \log(1 + \frac{v_3}{v_2}) - \frac{1}{2} \log(1 - \frac{v_3}{v_2+v_3}\rho^2))} \quad (362)$$

i.e.,

$$\mathbf{E}_\Phi[I(S; E)] \leq \frac{1}{\rho} e^{-\rho H_{1+\rho} + n(\frac{\rho}{2} \log(1 + \frac{v_3}{v_2}) - \frac{1}{2} \log(1 - \frac{v_3}{v_2+v_3}\rho^2))} \quad (363)$$

for $\rho \in (0, 1]$. Since there is no common messages, the cardinality of \mathcal{B}_1 is 1 in Code Ensemble 3. Theorem 22 guarantees that

$$\mathbf{E}_\Phi[e^{\rho I(S; E)[\Phi]}] \leq 1 + e^{-\rho H_{1+\rho}(A) + n\frac{\rho}{2} \log(1 + \frac{v_3}{(1-\rho)v_2})}, \quad (364)$$

i.e.,

$$\mathbf{E}_\Phi[I(S; E)] \leq \frac{1}{\rho} e^{-\rho H_{1+\rho}(A) + n\frac{\rho}{2} \log(1 + \frac{v_3}{(1-\rho)v_2})} \quad (365)$$

for $\rho \in (0, 1]$. When the dummy message A is uniform, (365) and (363) are simplified as follows

$$\mathbf{E}_\Phi[I(S; E)] \leq \frac{1}{\rho} e^{-n(\rho R_2 - (\frac{\rho}{2} \log(1 + \frac{v_3}{v_2}) - \frac{1}{2} \log(1 - \frac{v_3}{v_2+v_3}\rho^2)))}. \quad (366)$$

$$\mathbf{E}_\Phi[I(S; E)] \leq \frac{1}{\rho} e^{-n(\rho R_2 - \frac{\rho}{2} \log(1 + \frac{v_3}{(1-\rho)v_2}))}. \quad (367)$$

Since $\lim_{\rho \rightarrow 0} \frac{1}{\rho} (\frac{\rho}{2} \log(1 + \frac{v_3}{(1-\rho)v_2})) = \lim_{\rho \rightarrow 0} \frac{1}{\rho} (\frac{\rho}{2} \log(1 + \frac{v_3}{v_2}) - \frac{1}{2} \log(1 - \frac{v_3}{v_2+v_3}\rho^2)) = \frac{1}{2} \log(1 + \frac{v_3}{v_2})$, both (366) and (367) yield the strong security when $R_2 > \frac{1}{2} \log(1 + \frac{v_3}{v_2})$.

REFERENCES

- [1] S. Arimoto, "On the converse to the coding theorem for discrete memoryless channels," *IEEE Trans. Inform. Theory*, vol. 19, no. 3, pp. 357–359, May 1973.
- [2] P. Bergmans, "Random coding theorem for broadcast channels with degraded components", *IEEE Trans. Inform. Theory*, vol. 19, no. 2, pp. 197–207, 1973.
- [3] M. R. Bloch, "Achieving secrecy: Capacity vs. resolvability," in *Proc. ISIT 2011*, Saint-Petersburg, Russia, Aug. 2011, pp. 633–637.
- [4] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
- [5] T. Cover, "A proof of the data compression theorem of Slepian and Wolf for ergodic sources", *IEEE Trans. Inform. Theory*, vol. 21, no. 2, pp. 226–228, 1975.
- [6] I. Csiszár, "The Method of Types," *IEEE Trans. Inform. Theory*, vol. 44, no. 6, pp. 2505–2523, 1998.
- [7] —, "Almost independence and secrecy capacity," *Problems of Information Transmission*, vol. 32, no. 1, pp. 40–47, 1996.
- [8] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Budapest, Hungary: Akadémiai Kiadó, 1981.
- [9] —, "Broadcast channels with confidential messages," *IEEE Trans. Inform. Theory*, vol. 24, no. 3, pp. 339–348, May 1978.
- [10] P. Delsarte and P. M. Piret, "Algebraic constructions of Shannon codes for regular channels," *IEEE Trans. Inform. Theory*, vol. 28, no. 4, pp. 593–599, Jul. 1982.
- [11] I. Ekeland, R. Témán, *Convex Analysis and Variational Problems*, (North-Holland, Amsterdam, 1976); (SIAM, Philadelphia, 1999).
- [12] R. G. Gallager, *Information Theory and Reliable Communication*. New York: John Wiley & Sons, 1968.
- [13] T. S. Han and S. Verdú, "Approximation theory of output statistics," *IEEE Trans. Inform. Theory*, vol. 39, no. 3, pp. 752–772, May 1993.
- [14] T. S. Han, "Folklore in source coding: Information-spectrum approach," *IEEE Trans. Inform. Theory*, vol. 51, no. 2, pp. 747–753, Feb. 2005.
- [15] M. Hayashi, "General non-asymptotic and asymptotic formulas in channel resolvability and identification capacity and its application to wire-tap channel," *IEEE Trans. Inform. Theory*, vol. 52, no. 4, pp. 1562–1575, Apr. 2006.

- [16] —, “Second-Order Asymptotics in Fixed-Length Source Coding and Intrinsic Randomness,” *IEEE Trans. Inform. Theory*, vol. 54, 4619–4637, 2008.
- [17] —, “Exponential decreasing rate of leaked information in universal random privacy amplification,” *IEEE Trans. Inform. Theory*, vol. 57, no. 6, pp. 3989–4001, Jun. 2011.
- [18] M. Hayashi and R. Matsumoto, “Construction of wiretap codes from ordinary channel codes,” in *Proc. 2010 IEEE ISIT*, Austin, Texas, USA, Jun. 2010, pp. 2538–2542.
- [19] —, “Universally attainable error and information exponents, and equivocation rate for the broadcast channels with confidential messages,” in *Proc. 49th Annual Allerton Conf.*, Allerton House, Monticello, IL, USA, 2011, pp. 439–444, arXiv:1104.4285.
- [20] —, “Secure Multiplex Coding with Dependent and Non-Uniform Multiple Messages,” in *Proc. 50th Annual Allerton Conf.* Allerton House, Monticello, IL, USA, 2012, pp. 954–959.
- [21] Y. Kaspi and N. Merhav, “Error exponents for broadcast channels with degraded message sets,” *IEEE Trans. Inform. Theory*, vol. 57, no. 1, pp. 101–123, Jan. 2011.
- [22] D. Kobayashi, H. Yamamoto, T. Ogawa, “Secure multiplex coding attaining channel capacity in wiretap channels,” *IEEE Trans. Inform. Theory*, vol. 59, no. 12, pp. 8131–8143, Dec. 2013.
- [23] J. Körner and K. Marton, “General broadcast channels with degraded message sets,” *IEEE Trans. Inform. Theory*, vol. 23, no. 1, pp. 60–64, Jan. 1977.
- [24] J. Körner and A. Sgarro, “Universally attainable error exponents for broadcast channels with degraded message sets,” *IEEE Trans. Inform. Theory*, vol. 26, no. 6, pp. 670–679, Nov. 1980.
- [25] Y. Liang, H. V. Poor, and S. Shamai (Shitz), *Information Theoretic Security*. Hanover, MA, USA: NOW Publishers, 2009.
- [26] M. Bellare, S. Tessaro, and A. Vardy, “Semantic security for the wiretap channel,” *CRYPTO, LNCS*, vol. 7417, pp. 294–311, 2012.
- [27] R. Matsumoto and M. Hayashi, “Secure multiplex coding with a common message,” in *Proc. 2011 IEEE ISIT*, Saint-Petersburg, Russia, Jul. 2011, pp. 1931–1935, arXiv:1101.4036.
- [28] U. M. Maurer, “The strong secret key rate of discrete random triples,” in *Communications and Cryptography – Two Sides of One Tapestry*, R. E. Blahut et al., Eds. Kluwer Academic Publishers, 1994, pp. 271–285.
- [29] J. Muramatsu and S. Miyake, “Construction of Codes for the Wiretap Channel and the Secret Key Agreement From Correlated Source Outputs Based on the Hash Property,” *IEEE Trans. Inform. Theory*, vol. 58, no. 2, pp. 671–692, 2012.
- [30] D. Slepian and J. Wolf, “Noiseless coding of correlated information sources,” *IEEE Trans. Inform. Theory*, vol. 19, pp. 471–480, July 1973.
- [31] T. Kasami, *Weight distribution of Bose-Chaudhuri-Hocquenghem codes*, Defense Technical Information Center, 1966; R.C Bose, and T.A DOWLING (Eds.), *Combinatorial Mathematics and Its Applications*, Univ. of North Carolina Press, Chapel Hill (1969), pp. 335–357.
- [32] V. Y. F. Tan and O. Kosut, “The Dispersion of Slepian-Wolf Coding,” in *Proc. 2012 IEEE ISIT*, Cambridge, MA, USA, Jul., 2012, pp. 915–919.
- [33] S. Vembu and S. Verdú, “Generating random bits from an arbitrary source: Fundamental limits,” *IEEE Trans. Inform. Theory*, vol. 41, no. 5, pp. 1322–1332, 1995.
- [34] S. Verdú, “Non-Asymptotic Achievability Bounds in Multiuser Information Theory,” *Proc. 50th Allerton Conf.*, 2012, pp. 1–8.
- [35] A. D. Wyner, “The wire-tap channel,” *Bell System Tech. J.*, vol. 54, no. 8, pp. 1355–1387, Oct. 1975.
- [36] J. Chen, D.-k. He, A. Jagmohan, L. A. Lastras-Montano, and E.-h. Yang, “On the Linear Codebook-Level Duality Between Slepian-Wolf Coding and Channel Coding,” *IEEE Trans. Inform. Theory*, vol. 55, pp. 5575 (2009).
- [37] H. Yagi, “Finite Blocklength Bounds for Multiple Access Channels with Correlated Sources,” *ISITA2012* 377–381 (2012)
- [38] R. G. Gallager, “Source coding with side information and universal coding,” presented at the IEEE Int. Symp. Inform. Theory, Ronneby, Sweden, July 1976.
- [39] I. Csiszár and J. Körner, “Graph Decomposition: A New Key to Coding Theorems,” *IEEE Trans. Inform. Theory*, vol. 27, no. 1, pp. 5–12 (1981).
- [40] T. Richardson and R. Urbanke, *Modern Coding Theory*, Cambridge University Press, 2008.
- [41] L. Hanzo et al., *Turbo Coding, Turbo Equalisation and Space-Time Coding*, Wiley-IEEE Press, 2011.
- [42] S. Miyake and F. Kanaya, “Coding theorems on correlated general sources,” *IEICE Trans. Fundamentals*, vol. E78-A, no. 9, 1063–1070 (1995).
- [43] M. Hayashi, “Tight exponential analysis of universally composable privacy amplification and its applications,” *IEEE Trans. Inform. Theory*, vol. 59, no. 11, pp. 7728–7746 (2013).
- [44] R. Matsumoto, and M. Hayashi, “Universal Strongly Secure Network Coding with Dependent and Non-Uniform Messages,” arXiv:1111.4174 (2011)
- [45] M. Hayashi, and T. Tsurumaru, “More Efficient Privacy Amplification with Less Random Seeds via Dual Universal Hash Function.” arXiv:1311.5322 (2013); Accepted for publication in *IEEE Trans. Inform. Theory*.
- [46] G. Van Assche, *Quantum Cryptography and Secret-Key Distillation*, Cambridge University Press, 2006.
- [47] P. Billingsley, *Probability and Measure*, Wiley, 2012.
- [48] H. Nagaoka, “Strong Converse Theorems in Quantum Information Theory,” In *Proc. ERATO Workshop on Quantum Information Science 2001*, page 33, (2001).
- [49] M. Hayashi, “Information Spectrum Approach to Second-Order Coding Rate in Channel Coding,” *IEEE Trans. Inform. Theory*, vol. 55, no. 11, 4947–4966, 2009.
- [50] Y. Polyanskiy, H.V. Poor, and S. Verdú, “Channel coding rate in the finite blocklength regime,” *IEEE Trans. Inform. Theory*, vol. 56, no. 5, 2307–2359, 2010.
- [51] M. Tomamichel, and M. Hayashi, “Operational Interpretation of Rényi Information Measures via Composite Hypothesis Testing Against Product and Markov Distributions,” arXiv:1511.04874 (2015).

Masahito Hayashi (M’06–SM’13) was born in Japan in 1971. He received the B.S. degree from the Faculty of Sciences in Kyoto University, Japan, in 1994 and the M.S. and Ph.D. degrees in Mathematics from Kyoto University, Japan, in 1996 and 1999, respectively.

He worked in Kyoto University as a Research Fellow of the Japan Society of the Promotion of Science (JSPS) from 1998 to 2000, and worked in the Laboratory for Mathematical Neuroscience, Brain Science Institute, RIKEN from 2000 to 2003, and worked in ERATO Quantum Computation and Information Project, Japan Science and Technology Agency (JST) as the Research Head from 2000 to 2006. He also worked in the Superrobust Computation Project Information Science and Technology Strategic Core (21st Century COE by MEXT) Graduate School of Information Science and Technology, The University of Tokyo as Adjunct Associate Professor from 2004 to 2007. In 2006, he published the book “Quantum Information: An Introduction” from Springer. He worked in the Graduate School of Information Sciences, Tohoku University as Associate Professor from 2007 to 2012. In 2012, he joined the Graduate School of Mathematics, Nagoya University as Professor. He also worked in Centre for Quantum Technologies, National University of Singapore as Visiting Research Associate Professor from 2009 to 2012 and as Visiting Research Professor from 2012 to now. In 2011, he received Information Theory Society Paper Award (2011) for Information-Spectrum Approach to Second-Order Coding Rate in Channel Coding. In 2016, he received the Japan Academy Medal from the Japan Academy and the JSPS Prize from Japan Society for the Promotion of Science.

He is on the Editorial Board of *International Journal of Quantum Information* and *International Journal On Advances in Security*. His research interests include classical and quantum information theory and classical and quantum statistical inference.

Ryutaroh Matsumoto (M’00) was born in Nagoya, Japan, on November 29, 1973. He received the B.E. degree in computer science, the M.E. degree in information processing, and the Ph.D. degree in electrical and electronic engineering, all from Tokyo Institute of Technology, Japan, in 1996, 1998 and 2001, respectively. He was an Assistant Professor from 2001 to 2004, and has been an Associate Professor since 2004 in the Department of Communications and Computer Engineering, Tokyo Institute of Technology. He also served as a Velux visiting professor for the Department of Mathematical Sciences, Aalborg University, Denmark during 2011–2012 and 2014. His research interests include error-correcting codes, quantum information theory, information theoretic security, and communication theory. Dr. Matsumoto received the Young Engineer Award from IEICE and the Ericsson Young Scientist Award from Ericsson Japan in 2001. He received the Best Paper Awards from IEICE in 2001, 2008, 2011 and 2014.