This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2020.3004482, IEEE Transactions on Information Theory

Weakly Secure Symmetric Multilevel Diversity Coding

Tao Guo, *Member, IEEE*, Chao Tian, *Senior Member, IEEE*, Tie Liu, *Senior Member, IEEE*, and Raymond W. Yeung, *Fellow, IEEE*

***Abstract*—Multilevel diversity coding is a classical coding model where multiple mutually independent information messages are encoded, such that different reliability requirements can be afforded to different messages. It is well known that *superposition coding*, namely separately encoding the independent messages, is optimal for symmetric multilevel diversity coding (SMDC) (Yeung-Zhang 1999). In the current paper, we consider weakly secure SMDC where security constraints are injected on each individual message, and provide a complete characterization of the conditions under which superposition coding is sum-rate optimal. Two joint coding strategies, which lead to rate savings compared to superposition coding, are proposed, where some coding components for one message can be used as the encryption key for another. By applying different variants of Han's inequality, we show that the lack of opportunity to apply these two coding strategies directly implies the optimality of superposition coding. It is further shown that under a set of particular security constraints, one of the proposed joint coding strategies can be used to construct a code that achieves the optimal rate region.**

I. Introduction

Symmetric multilevel diversity coding (SMDC) was introduced by Roche *et al.* [1] for applications in distributed data storage and robust network communication. Albanese *et al.* [2] independently studied the problem of *priority encoding transmission* (PET), which shares the same mathematical model as SMDC. In a symmetric L-level diversity coding system, there are L independent messages $(M_1, M_2, \dots, M_L)$, where the importance of messages decreases with the subscript l. The messages are encoded by L encoders. There are totally $2^L - 1$ decoders, each of which has access to the outputs of a distinct subset of the encoders. A decoder which can access any α encoders, called a Level-α decoder, is required to reconstruct the first α most important messages. The system is symmetric in the sense that the reconstruction requirement of a decoder depends on the set of encoders it can access only via its cardinality.

The work of C. Tian was supported in part by the National Science Foundation under Grant CCF-18-32309 and CCF-18-16546. This paper was presented in part at 2019 IEEE Information Theory Workshop (ITW).

T. Guo was with the Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX, USA. He is now with the Department of Electrical and Computer Engineering, the University of California, Los Angeles, CA, USA. (e-mail: guotao@ucla.edu)

C. Tian and T. Liu are with the Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX, USA. (e-mail: chao.tian@tamu.edu, tieliu@tamu.edu)

R. W. Yeung is with the Institute of Network Coding and the Department of Information Engineering, The Chinese University of Hong Kong, Hong Kong (e-mail: whyeung@ie.cuhk.edu.hk)

It was shown [1], [3] that separately encoding these independent messages, referred to as *superposition coding*, is optimal in terms of achieving the entire rate region. The characterization of the coding rate region therein involves implicit and uncountably many inequalities, and an explicit characterization of the coding rate region was recently obtained [4]. The problem has also been extended and generalized, e.g., to allow node regeneration [5] and to allow asymmetric decoders [6]. Li *et al.* [7] studied the multilevel diversity coding problem with at most 3 sources and 4 encoders in a systematic way and obtained the exact rate region of each of the over 7,000 instances with the aid of computation.

The SMDC problem with a *strong* security guarantee was considered by Balasubramanian et al. [8] and Jiang et al. [9]. In this setting, a security threshold N is given, and the first N messages are degenerate. For the remaining $L - N$ messages M_α, $\alpha = N+1, N+2, \cdots, L$, in addition to the standard multilevel reconstruction requirement, it is also required that all these messages need to be kept perfectly *jointly* secure if no more than N encoders are accessible by an eavesdropper. Despite the additional security constraints, it was shown that superposition coding remains to be optimal in terms of both the sum rate [8] and the entire rate region [9].

In this paper we consider a *weakly* secure setting of the classical SMDC problem, where the security level of each message is specified by a separate security parameter N_α. More specifically, for any $\alpha = 1, 2, \dots, L$, we require the message M_α to be kept perfectly secure if the outputs of no more than N_α encoders are accessible by an eavesdropper. Such a security requirement is "weak" in the sense that the eavesdropper is only prevented from obtaining any information about the *individual* messages. By comparison, the security requirement of [8]–[10] is strong in that it prevents the eavesdropper to obtain any information about the *entire* set of messages. The notion of weak security has been considered in various network coding settings [11]–[14] and also channel coding perspectives [15]–[18] in the literature and is generally considered to be more practical for protecting individual messages. For example, when the messages are video sequences, the user should not obtain information about any individual video segment, but obtaining the binary XOR of two video sequences may not be an issue since it will not lead to a meaningfully decodable video sequence. Moreover, a protocol with a weak security constraint can potentially be implemented more efficiently in practical settings and may not require encryption keys. Note that the notion of "weak/strong security" here is different from the asymptotic notion of weak/strong security in [19]–[22], wherein asymptotic weak security requires vanishing of the information leakage rate and the corresponding strong security requires the vanishing of leaked information content. Another notion of "weak security" is defined in [23] which requires the eavesdropper to be unable to obtain any meaningful information about the source.

On the one hand, the notion of weak security has significantly enriched the collection of secure SMDC problems: Unlike the strongly secure setting where a single security parameter is set for all the messages, for the weakly secure setting, a different security parameter can be set for each message. On the other hand, the notion of weak security has also cast the optimality of superposition coding in much greater doubt, as requiring the messages to be protected only *marginally* (instead of jointly) significantly opens up the set of feasible coding strategies. The main goals of this paper are: 1) to understand under what configurations of the security parameters $(N_1, N_2, \dots, N_L)$ superposition coding remains to be optimal; and 2) to identify optimal coding strategies when superposition coding is suboptimal.

The main message of this paper is that the optimality of superposition coding depends *critically* on the security parameters $(N_1, N_2, \dots, N_L)$. More specifically, we consider a natural joint coding strategy that encodes a pair of messages together by using one of the messages as part of the secret key for securing the other. We term this coding strategy *pairwise encoding*, and Sections IV-A and IV-B discuss two scenarios for which pairwise encoding is possible. The main results of the paper are:

1) We show that superposition coding can achieve the minimum sum rate whenever pairwise encoding is not possible between any two messages. This immediately leads to a necessary and sufficient condition on the security parameters $(N_1, N_2, \cdots, N_L)$ for superposition coding to be optimal in terms of minimizing the sum rate.
2) We consider a special class, referred to as *differential-constant secure SMDC* (DS-SMDC), for which the more important messages are maximally protected ($N_\alpha = \alpha - 1$) and the less important messages are not protected at all ($N_\alpha = 0$), and show that a simple extension of the pairwise encoding strategy (from a pair of messages to a pair of groups of messages and hence termed as *group pairwise encoding*) can achieve the entire rate region.

Note that the min-cut capacity for multicasting a single source is achievable using linear network codes [24]–[26]. It was shown in [23] that the min-cut bound can also be achieved for a single-source secure network coding model, where the security measure is similar to the weak security notion we used in this work. However, the min-cut bound may not be achievable for general multi-source network coding problems (even without any security measure), e.g., the example illustrated by Fig. 21.3 in [27]. In particular, the min-cut bound is not achievable for the secure SMDC problem here.

The rest of the paper is organized as follows. We first formulate the problem and state some preliminary results in Section II. In Section III, we state the main results, i) a precise classification of the cases where superposition is sum-rate optimal; ii) the optimal rate region for DS-SMDC.

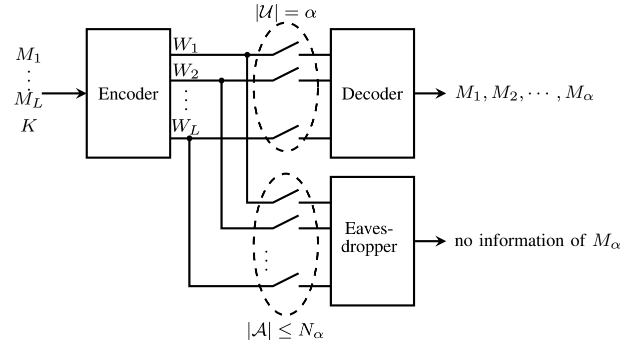

Fig. 1: The Weakly Secure SMDC Model

In Sections IV and V, we describe the pairwise encoding strategies that reduce coding rates and prove the optimality of superposition under the conditions in i). Section VI is devoted to the proof of the optimal rate region for DS-SMDC. We conclude the paper in Section VII. Some technical proofs can be found in the appendices.

II. Problem Formulation and Preliminaries

A. Problem Formulation

Let $\mathcal{L} \triangleq \{1, 2, \cdots, L\}$, where $L \geq 2$. Let $M_1, M_2, \cdots, M_L$ be a collection of L mutually independent messages uniformly distributed over the direct product of certain finite sets. For simplicity, we assume the message set to be $\mathbb{F}_{p^{m_1}} \times \mathbb{F}_{p^{m_2}} \times \cdots \times \mathbb{F}_{p^{m_L}}$, where $\mathbb{F}_{p^{m_1}}$ is a finite field of order p^{m_1} and p itself can be an integer power of some prime number. We may also regard M_α ($\alpha \in \mathcal{L}$) as $M_\alpha = (M_\alpha^1, M_\alpha^2, \cdots, M_\alpha^{m_\alpha})$ where $M_\alpha^i \in \mathbb{F}_p$ for $i = 1, 2, \cdots, m_\alpha$.

The weakly secure SMDC problem is depicted in Fig. 1. There are L encoders, indexed by $\mathcal{L}$, each of which can access all the L information messages. There are also $2^L - 1$ decoders. For each $\mathcal{U} \subseteq \mathcal{L}$ such that $\mathcal{U} \neq \emptyset$, Decoder-$\mathcal{U}$ can access the outputs of the subset of encoders indexed by $\mathcal{U}$. For $\alpha \in \mathcal{L}$ and any $\mathcal{U}$ such that $|\mathcal{U}| = \alpha$, Decoder-$\mathcal{U}$ can completely recover the first α messages $M_1, M_2, \cdots, M_\alpha$. In addition, there is an eavesdropper who has access to the outputs of a subsets $\mathcal{A}$ of encoders. Let $\boldsymbol{N} = (N_1, N_2, \cdots, N_L)$ be L non-negative integers, where $N_\alpha < \alpha$ for $\alpha \in \mathcal{L}$. Weak security requires that each individual message M_α should be kept perfectly secure from the eavesdropper if $|\mathcal{A}| \leq N_\alpha$.

Let $\mathcal{K}$ be the key space. An $(m_1, m_2, \cdots, m_L, R_1, R_2, \cdots, R_L)$ code is formally defined by the encoding functions

$$E_l : \prod_{i=1}^{L} \mathbb{F}_{p^{m_i}} \times \mathcal{K} \rightarrow \mathbb{F}_{p^{R_l}}, \text{ for } l \in \mathcal{L} \tag{1}$$

and decoding functions

$$D_\mathcal{U} : \prod_{l \in \mathcal{U}} \mathbb{F}_{p^{R_l}} \rightarrow \prod_{i=1}^{|\mathcal{U}|} \mathbb{F}_{p^{m_i}}, \text{ for } \mathcal{U} \subseteq \mathcal{L} \text{ and } \mathcal{U} \neq \emptyset. \tag{2}$$

Denote the shared key as K (accessible to all the encoders), which is uniformly distributed in the key space $\mathcal{K}$. Let $W_l = E_l(M_1, M_2, \cdots, M_L, K)$ be the output of Encoder-l

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2020.3004482, IEEE Transactions on Information Theory

and $W_{\mathcal{U}} = (W_l : l \in \mathcal{U})$ for $\mathcal{U} \subseteq \mathcal{L}$. Define the normalized message rates $\mathsf{m}_l \triangleq m_l / \sum_{l=1}^{L} m_l$, from which it follows that $\sum_l \mathsf{m}_l = 1$. A normalized non-negative rate tuple $\mathbf{R} \triangleq (\mathsf{R}_1, \mathsf{R}_2, \cdots, \mathsf{R}_L)$ is *achievable* for the normalized message rates $(\mathsf{m}_1, \ldots, \mathsf{m}_L)$, if for any $\epsilon > 0$, there exist an integer a and an $(a\mathsf{m}_1, a\mathsf{m}_2, \cdots, a\mathsf{m}_L, R_1, R_2, \cdots, R_L)$ code such that

$$\textbf{perfect reconstruction: } D_{\mathcal{U}}(W_{\mathcal{U}}) = (M_1, M_2, \cdots, M_{|\mathcal{U}|}), \quad \forall\, \mathcal{U} \subseteq \mathcal{L} \text{ s.t. } \mathcal{U} \neq \emptyset, \tag{3}$$

$$\textbf{perfect secure: } H(M_\alpha | W_{\mathcal{A}}) = H(M_\alpha), \quad \forall\, \alpha \in \mathcal{L} \text{ and } \mathcal{A} \subseteq \mathcal{L} \text{ s.t. } |\mathcal{A}| \leq N_\alpha, \tag{4}$$

and

$$\textbf{coding rate: } \mathsf{R}_l + \epsilon \geq a^{-1} R_l, \qquad l \in \mathcal{L}. \tag{5}$$

The optimal coding rate region $\mathcal{R}$ is defined as the collection of all achievable rate tuples.

Remark 1. Here each message M_α can be essentially represented in $m_\alpha \log_2 p$ bits, and each codeword W_l can be represented in $R_l \log_2 p$ bits. Thus R_l can be viewed as the coding rate of encoder E_l, when the definition of the entropy function uses logarithm of base p, which will be adopted from here on. The quantity R_l is then essentially the normalized R_l.

The minimum achievable normalized sum rate is defined as $\mathsf{R}^*_{\text{sum}} \triangleq \min \sum_{l=1}^{L} \mathsf{R}_l$, and one of our main results is a necessary and sufficient condition for superposition coding to be sum-rate optimal. We also study an important case where $\boldsymbol{N}$ is given by

$$N_\alpha = \begin{cases} \alpha - 1, & \text{for } 1 \leq \alpha \leq r \\ 0, & \text{for } r+1 \leq \alpha \leq L, \end{cases} \tag{6}$$

for certain parameters (L, r), where $r \geq 1$. We refer to this system as the (L, r) differential-constant secure SMDC (DS-SMDC), where the more important messages (i.e., small α values) are maximally secure ($N_\alpha = \alpha - 1$) and the less important messages do not have any security guarantee at all ($N_\alpha = 0$). For the protected messages ($1 \leq \alpha \leq r$), while the security constraint N_α grows with the reconstruction requirement α, the difference between N_α and α remains to be a constant equal to 1. We refer to this feature as “differential-constant secure”, in contrast to the “level-constant secure” guarantee [8], [9] which requires $N_\alpha = N$ for all $\alpha > N$. Denote the optimal coding rate region of the (L, r) DS-SMDC problem by $\mathcal{R}_{L,r}$, which is the collection of all achievable normalized rate tuples. For $r = 1$, the problem reduces to the classical SMDC.

B. An Achievable Rate Region via Superposition Coding

Let M be a message encoded by n encoders. For any $0 \leq c < k \leq n$, the (c, k, n) ramp secret sharing problem [28], also known as the secure symmetrical single-level diversity coding (S-SSDC) problem in [8], requires that the outputs from any subset of no more than c encoders provide no information about the message, and the outputs from any subset of k encoders can completely recover the message. The optimal rate region for this problem can be found in [8], [29], as stated in the following lemma.

Lemma 1. *The optimal rate region of the (c, k, n) ramp secret sharing problem is the collection of rate tuples $(R_1, R_2, \cdots, R_n)$ such that*

$$\sum_{l \in \mathcal{B}} R_l \geq H(M), \ \forall \mathcal{B} \subseteq \{1, 2, \cdots, n\}, |\mathcal{B}| = k - c. \tag{7}$$

Remark 2. If $k = c + 1$, the (c, k, n) ramp secret sharing problem reduces to the (k, n) threshold secret sharing problem and the rate region reduces accordingly.

In light of this result, a natural coding scheme (i.e., superposition coding) for the weakly secure SMDC problem formulated above is to separately encode each message M_α using an (N_α, α, L) ramp secret sharing code as shown in Fig. 2. The rate region induced by superposition coding provides an inner bound $\mathcal{R}_{\text{sup}}$ for $\mathcal{R}$, and by Lemma 1, it can be written as the set of non-negative rate tuples $\mathbf{R} = (\mathsf{R}_1, \mathsf{R}_2, \cdots, \mathsf{R}_L)$ such that

$$\mathsf{R}_l = \sum_{\alpha=1}^{L} r_l^\alpha, \ \text{ for } l \in \mathcal{L} \tag{8}$$

for some $r_l^\alpha \geq 0, \ l, \alpha \in \mathcal{L}$, satisfying

$$\sum_{l \in \mathcal{B}} r_l^\alpha \geq \mathsf{m}_\alpha, \ \text{ for } \mathcal{B} \subseteq \mathcal{L} \text{ s.t. } |\mathcal{B}| = \alpha - N_\alpha. \tag{9}$$

The induced sum rate provides an upper bound $\bar{\mathsf{R}}_{\text{sum}}$ for $\mathsf{R}^*_{\text{sum}}$, and can be written simply as,

$$\bar{\mathsf{R}}_{\text{sum}} \triangleq \sum_{\alpha=1}^{L} \frac{L\mathsf{m}_\alpha}{\alpha - N_\alpha}. \tag{10}$$

C. Properties of MDS Code for Secret Sharing

In this section, we describe in some details two (n, k) maximum distance separable (MDS) codes for ramp secret sharing that achieve the minimum sum rate in Lemma 1, and provide important properties that are instrumental to the joint coding strategy we later propose.

Let $M = (U_1, U_2, \cdots, U_{k-c})$ be a length-$(k - c)$ message where each symbol is chosen uniformly and independently from the finite field $\mathbb{F}_p$. Let $Z_1, Z_2, \cdots, Z_c$ be independent random keys chosen uniformly from the same finite field $\mathbb{F}_p$. For $i = 1, 2, \cdots, k$, define the following length-k vectors:

$$f_i = [\underbrace{0 \ \cdots \ 0}_{i-1} \ 1 \ 0 \ \cdots \ 0]^T. \tag{11}$$

Let $g_1, g_2, \cdots, g_n$ be length-k vectors with entries from $\mathbb{F}_p$ such that any k vectors $\{h_{j_1}, h_{j_2}, \cdots, h_{j_k}\}$ chosen from the set $\{f_1, f_2, \cdots, f_k, \ g_1, g_2, \cdots, g_n\}$ satisfy the full rank condition over $\mathbb{F}_p$, i.e.,

$$\text{rank}\,[h_{j_1} \ h_{j_2} \ \cdots \ h_{j_k}] = k. \tag{12}$$

It can be shown that as long as $p \geq n + k$, there exist such vectors $g_1, g_2, \cdots, g_n$, e.g., it can be chosen as the columns

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2020.3004482, IEEE Transactions on Information Theory

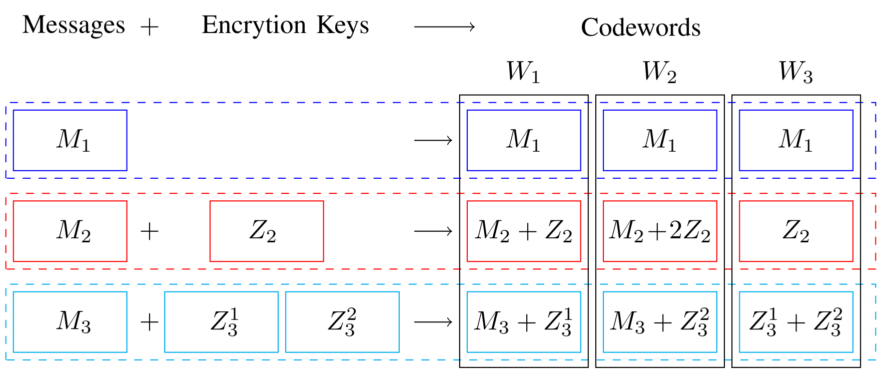

Fig. 2: The superposition coding scheme for $(3, 3)$ DS-SMDC with $(m_1, m_2, m_3) = (1, 1, 1)$, $(N_1, N_2, N_3) = (0, 1, 2)$, and $p = 3$.

from a Cauchy matrix. The generator matrices of the two MDS codes of interest are given, respectively, as

$$G^{(1)} = [f_{k-c+1} \ \cdots \ f_k \ g_1 \ g_2 \ \cdots \ g_{n-c}], \tag{13}$$
$$G^{(2)} = [g_1 \ g_2 \ \cdots \ g_n]. \tag{14}$$

Then the codewords of two MDS codes are, respectively,

$$[Y_1, Y_2, \cdots, Y_n] = \big[U_1 \ \cdots \ U_{k-c} \ \ Z_1 \ \cdots \ Z_c\big]G^{(1)}, \tag{15}$$
$$[Y_1, Y_2, \cdots, Y_n] = \big[U_1 \ \cdots \ U_{k-c} \ \ Z_1 \ \cdots \ Z_c\big]G^{(2)}. \tag{16}$$

We shall refer these two codes as MDS-A and MDS-B, respectively. By the definition of $f_{k-c+1}, \cdots, f_k$ in (11), MDS-A has the random keys explicitly as part of the coded message,

$$[Y_1, Y_2, \cdots, Y_c] = \big[Z_1 \ Z_2 \ \cdots \ Z_c\big]. \tag{17}$$

It is obvious that for both codes, M and $Z_1, Z_2, \cdots, Z_c$ can be perfectly recovered from any k coded symbols.

Since all the coded symbols are linear combinations of the messages and the random keys that are uniformly distributed, we have the following lemma.

Lemma 2. *Any k coded symbols of MDS-A and MDS-B are uniformly distributed over $\mathbb{F}_{p^k}$.*

The main difference between the two codes, which is the most relevant to this work, is given in the following two lemmas.

Lemma 3. *For any integer t such that $c \le t \le k$, let $\mathcal{E} \subseteq \{1, 2, \cdots, n\}$ where $|\mathcal{E}| = t$, and $\mathcal{A} \subseteq \{1, 2, \cdots, k-c\}$ where $|\mathcal{A}| = k - t$. The codewords of MDS-A has the following property:*

$$I(Y_{\mathcal{E}}; U_{\mathcal{A}}) = 0, \tag{18}$$

where $Y_{\mathcal{E}} \triangleq \{Y_i : i \in \mathcal{E}\}$ and $U_{\mathcal{A}} \triangleq \{U_i : i \in \mathcal{A}\}$.

Proof. We consider the following chain of equality

$$
\begin{aligned}
&I(Y_{\mathcal{E}}; U_{\mathcal{A}}) \\
=&H(Y_{\mathcal{E}}) - H(Y_{\mathcal{E}}|U_{\mathcal{A}}) && (19)\\
=&H(Y_{\mathcal{E}}) - H(Y_{\mathcal{E}}|U_{\mathcal{A}}) + H(Y_{\mathcal{E}}|U_{\mathcal{A}}U_{\bar{\mathcal{A}}}Z_1Z_2\cdots Z_c) && (20)\\
=&H(Y_{\mathcal{E}}) - H(U_{\bar{\mathcal{A}}}Z_1Z_2\cdots Z_c|U_{\mathcal{A}}) && (21)\\
&\ + H(U_{\bar{\mathcal{A}}}Z_1Z_2\cdots Z_c|Y_{\mathcal{E}}U_{\mathcal{A}}) && (22)\\
=&H(Y_{\mathcal{E}}) - H(U_{\bar{\mathcal{A}}}Z_1Z_2\cdots Z_c) && (23)\\
=&\frac{t}{k-c}H(M) - \frac{t-c+c}{k-c}H(M) && (24)\\
=&0, && (25)
\end{aligned}
$$

where (20) follows from (15), and both (23) and (24) follow from the full rank condition in (12) and the uniform and mutually independent distribution of the messages and the encryption key. □

Remark 3. For $t = c$, Lemma 3 reduces to the stated security constraint of parameter c; on the other hand, for $t > c$ (but $t \le k$), any t coded symbols reveal no information about any subset of $k - t$ message symbols.

Lemma 4. *For any integer t such that $0 \le t \le k$, let $\mathcal{E} \subseteq \{1, 2, \cdots, n\}$ where $|\mathcal{E}| = t$, and $\mathcal{A}_1 \subseteq \{1, 2, \cdots, k-c\}$, and $\mathcal{A}_2 \subseteq \{1, 2, \cdots, c\}$ where $|\mathcal{A}_1| + |\mathcal{A}_2| = k-t$. The codewords of MDS-B has the following property:*

$$I(Y_{\mathcal{E}}; U_{\mathcal{A}_1}, Z_{\mathcal{A}_2}) = 0, \tag{26}$$

where $Y_{\mathcal{E}} \triangleq \{Y_i : i \in \mathcal{E}\}$, $U_{\mathcal{A}_1} \triangleq \{U_i : i \in \mathcal{A}_1\}$, and $Z_{\mathcal{A}_2} \triangleq \{Z_i : i \in \mathcal{A}_2\}$.

Proof. This is direct from the full-rank condition in (12) and the uniform and mutually independent distribution of the messages and the encryption key. □

From the above two lemmas, in contrast to MDS-A, MDS-B has the additional advantage that part of the keys can also be made secure against some t eavesdroppers. This property becomes important to us in the sequel.

III. Main Results

A. Sum-rate Optimality Conditions of Superposition

The main question we seek to answer here is under what condition the equality $\mathsf{R}^*_{\mathsf{sum}} = \bar{\mathsf{R}}_{\mathsf{sum}}$ will hold, and the following theorem provides the exact answer to this question.

Theorem 1. $\mathsf{R}^*_{sum} = \bar{\mathsf{R}}_{sum}$, *if and only if for any* $\alpha < \beta \in \mathcal{L}$ *where* $\mathsf{m}_\alpha, \mathsf{m}_\beta > 0$, *we have*

$$\text{either } N_\alpha < \alpha \leq N_\beta < \beta, \text{ or } N_\alpha = N_\beta = 0. \tag{27}$$

Remark 4. If all L messages are non-degenerate, i.e., all the message entropies are non-zero, the condition in (27) is equivalent to that there exists a $T_s \in \{1, 2, \cdots, L\}$ such that for all $\alpha \in \mathcal{L}$,

$$N_\alpha = \begin{cases} 0, & \text{for } \alpha \leq T_s \\ \alpha - 1, & \text{for } \alpha > T_s. \end{cases} \tag{28}$$

If we do not assume non-degeneration, then the following necessary condition for optimality can be induced from (27): There exists a $T_s \in \{1, 2, \cdots, L\}$ such that for any $\alpha \in \mathcal{L}$ satisfying $m_\alpha > 0$,

$$\begin{cases} N_\alpha = 0, & \text{for } \alpha \leq T_s \\ N_\alpha > 0, & \text{for } \alpha > T_s. \end{cases} \tag{29}$$

Remark 5. The following are two examples that superposition coding is optimal in terms of achieving the entire rate region and thus Theorem 1 reduces correctly.

- If the threshold in (28) is $T_s = L$, the security constraints are given as
$$N_\alpha = 0, \text{ for all } \alpha \in \mathcal{L}, \tag{30}$$
then the problem reduces to the classical SMDC problem without security constraints, where superposition is known to be optimal [3].
- If the threshold in (28) is $T_s = 1$, the security constraint becomes
$$N_\alpha = \alpha - 1, \text{ for all } \alpha \in \mathcal{L}, \tag{31}$$
and the problem reduces to the special case of DS-SMDC for $r = L$ in Section III-B.

The following definition will be used in the sequel.

Definition 1. For any $\alpha < \beta \in \mathcal{L}$, we define two conditions.

$$\text{Condition 1}: \ N_\alpha < N_\beta < \alpha; \tag{32}$$
$$\text{Condition 2}: \ N_\beta \leq N_\alpha \ \& \ N_\alpha > 0. \tag{33}$$

Theorem 1 can be alternatively written in the following form, by taking the complement of the conditions in (27).

Theorem 1'. $\mathsf{R}^*_{sum} < \bar{\mathsf{R}}_{sum}$, *if and only if there exist* $\alpha < \beta \in \mathcal{L}$ *where* $\mathsf{m}_\alpha, \mathsf{m}_\beta > 0$ *such that either Condition 1 in* (32) *or Condition 2 in* (33) *holds.*

We prove Theorem 1 in two parts. In Section IV, we show that superposition is suboptimal under the security constraints in (32) or (33), by providing joint coding strategies that can reduce coding rates. In Section V, the optimality of superposition coding is established by proving that the sum rate is lower bounded by $\bar{\mathsf{R}}_{\mathsf{sum}}$ in (10).

Remark 6. Superposition coding is optimal for classical SMDC where there is no security constraints, i.e., suboptimality only happens when there is a security constraint. In view of the suboptimality in Sections IV-A and IV-B, we see intuitively that joint encoding helps only when some message can perform as the secret key of another message.

B. Rate Region of DS-SMDC

When superposition is not optimal, it is generally hard to characterize the coding rate region or even the minimum sum rate, since it is difficult to find the optimal code structures. In this section, we study the (L, r) DS-SMDC problem for which we fully characterize the optimal rate region. The pairwise coding strategy in Section IV-B can be generalized to a multi-message regime, and we obtain a *group pairwise* coding scheme that achieves the entire rate region of the DS-SMDC problem.

We first present an example that motivates the general group pairwise coding scheme.

Example 1. Let $L = 4, (m_1, m_2, m_3, m_4) = (1, 1, 1, 4)$, and $p = 11$. The security constraint for the $(4, 3)$ DS-SMDC problem should be $(N_1, N_2, N_3, N_4) = (0, 1, 2, 0)$. We can follow a naive strategy as illustrated in (34): use generator matrices G_2 and G_3 generated from MDS-B to encode M_2 and M_3 separately with encryption keys Z_2 and Z_3^1, Z_3^2; equally partition M_4 into four pieces $M_4^1, M_4^2, M_4^3, M_4^4$.

$$\begin{aligned} W_1 &= (M_2 + Z_2, \quad M_3 + 2Z_3^1 + 9Z_3^2, W_4^1), \\ W_2 &= (M_2 + 2Z_2, \ 9M_3 + 8Z_3^1 + 6Z_3^2, W_4^2), \\ W_3 &= (M_2 + 3Z_2, 6M_3 + 10Z_3^1 + 7Z_3^2, W_4^3), \\ W_4 &= (M_2 + 4Z_2, \ 7M_3 + 9Z_3^1 + 7Z_3^2, W_4^4); \end{aligned} \tag{34}$$

The first part of the group pairwise coding scheme is simply to use M_4, specifically M_4^1, M_4^2, M_4^3, to replace Z_2 and Z_3^1, Z_3^2 as secret keys to encrypt M_2 and M_3, as given in (35).

$$\begin{aligned} W_1' &= (M_2 + M_4^1, \quad M_3 + 2M_4^2 + 9M_4^3 \quad), \\ W_2' &= (M_2 + 2M_4^1, \ 9M_3 + 8M_4^2 + 6M_4^3 \quad), \\ W_3' &= (M_2 + 3M_4^1, 6M_3 + 10M_4^2 + 7M_4^3 \quad), \\ W_4' &= (M_2 + 4M_4^1, \ 7M_3 + 9M_4^2 + 7M_4^3, W_4^4). \end{aligned} \tag{35}$$

The second part of the group pairwise coding scheme simply encodes M_4^4 as part of the fourth coded message. Since M_4^1, M_4^2, M_4^3 does not need to be separately encoded, rate saving is obtained compared to the naive version. The reconstruction and security requirements of M_2 and M_3 are immediate from the MDS-B code. The reconstruction requirement of M_4 is straightforward since M_4^1, M_4^2, M_4^3 is recovered with any three coded symbols.

Coding scheme for general parameters:
The group pairwise coding scheme is illustrated in Fig. 3. For each $\alpha \in \{1, 2, \cdots, r\}$, we will use an (α, L)-threshold secret sharing scheme to encode M_α and use the last $L - r$ messages $M_{r+1}, \cdots, M_L$ as keys. It is proved in [28] that the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2020.3004482, IEEE Transactions on Information Theory

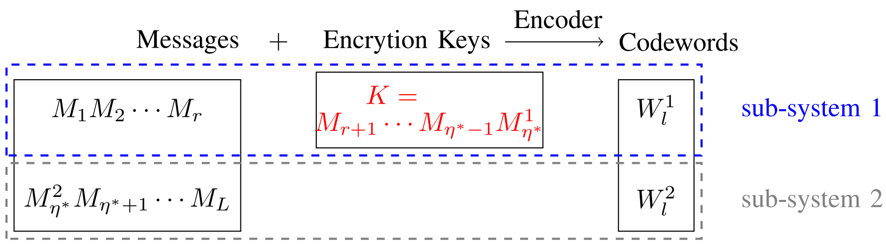

Fig. 3: The group pairwise coding scheme.

minimum key size for M_α is $(\alpha-1)m_\alpha$. Thus the total size of keys needed is

$$|\mathcal{K}| = \sum_{\alpha=1}^{r} (\alpha-1) m_\alpha. \tag{36}$$

For notational simplicity, we define an auxiliary message[1] M_{L+1}, which is independent with other messages and uniformly distributed over $\mathbb{F}_{p^{m_{L+1}}}$ with

$$m_{L+1} = \left[\sum_{\alpha=1}^{r} (\alpha-1) m_\alpha - \sum_{\alpha=r+1}^{L} m_\alpha \right]^+, \tag{37}$$

where for any $x \in \mathbb{R}$, $[x]^+ \triangleq \max\{0, x\}$. It is easy to check that

$$\sum_{\alpha=r+1}^{L+1} m_\alpha \geq \sum_{\alpha=1}^{r} (\alpha-1) m_\alpha. \tag{38}$$

Thus, there exists a unique $\eta^* \in \{r+1, r+2, \cdots, L+1\}$ such that

$$\sum_{\alpha=r+1}^{\eta^*-1} m_\alpha < \sum_{\alpha=1}^{r} (\alpha-1) m_\alpha \leq \sum_{\alpha=r+1}^{\eta^*} m_\alpha. \tag{39}$$

The parameter η^* determines which messages of $M_{r+1}, M_{r+2}, \cdots, M_{L+1}$ will be used as the encryption keys. In light of the definition of η^* in (39), denote the first $\frac{\sum_{\alpha=1}^{r}(\alpha-1)m_\alpha - \sum_{\alpha=r+1}^{\eta^*-1} m_\alpha}{m_{\eta^*}}$ fraction of M_{η^*} by $M^1_{\eta^*}$, and the rest by $M^2_{\eta^*}$. Then we use the messages $(M_{r+1}, M_{r+2}, \cdots, M_{\eta^*-1}, M^1_{\eta^*})$ to replace the keys of $M_1, \cdots, M_r$. The messages $M^2_{\eta^*}, M_{\eta^*+1}, M_{\eta^*+2}, \cdots, M_L$ are separately encoded in the same way as in classical SMDC.

Next, we verify the reconstruction and security constraints.

Reconstruction: By the code construction in Section II-C, the reconstruction requirements of all messages $(M_1, M_2, \cdots, M_r)$, $(M_{r+1}, M_{r+2}, \cdots, M_{\eta^*-1}, M^1_{\eta^*})$, and $(M^2_{\eta^*}, M_{\eta+1}, \cdots, M_L)$ are satisfied immediately.

Security: The security constraints of $(M_1, M_2, \cdots, M_r)$ is straightforward, and there is no security constraint for $(M_{r+1}, M_{r+2}, \cdots, M_L)$.

Remark 7. The first r messages $M_1, M_2, \cdots, M_r$ are encoded separately, and the last r messages $M_{r+1}, M_{r+2}, \cdots, M_L$ are also encoded separately. The reason why we call the coding scheme "group pairwise" is that joint encoding are only performed between the two groups of messages

$$\{M_1, M_2, \cdots, M_r\} \text{ and } \{M_{r+1}, M_{r+2}, \cdots, M_{\eta^*}\}. \tag{40}$$

The group pairwise coding scheme can also be interpreted as superposition coding of the messages $M_1, M_2, \cdots, M_r$, $M^*_{r+1}, \cdots, M^*_L$, where the independent pseudo-messages $M^*_\alpha (r+1 \leq \alpha \leq L)$ are defined by the message size m^*_α as

$$\mathsf{m}^*_\alpha = \begin{cases} 0, & \text{for } r+1 \leq \alpha \leq \eta^*-1 \\ \sum_{j=r+1}^{\eta^*} \mathsf{m}_j - \sum_{j=1}^{r} (j-1)\mathsf{m}_j, & \text{for } \alpha = \eta^* \\ \mathsf{m}_\alpha, & \text{for } \eta^*+1 \leq \alpha \leq L. \end{cases} \tag{41}$$

Then the coding rate region $\mathcal{R}^{L,r}_{\text{gp}}$ induced by group pairwise coding is the set of $\mathbf{R} \geq \mathbf{0}$ such that

$$\mathsf{R}_l = \sum_{\alpha=1}^{L} r^\alpha_l, \text{ for } l \in \mathcal{L}, \tag{42}$$

where $r^\alpha_l \geq 0$ and

$$r^\alpha_l \geq \mathsf{m}_\alpha, \text{ for } 1 \leq \alpha \leq r, \tag{43}$$

$$\sum_{l \in \mathcal{B}} r^\alpha_l \geq \mathsf{m}^*_\alpha, \text{ for all } \mathcal{B} \subseteq \mathcal{L} \text{ s.t. } |\mathcal{B}| = \alpha, \; r+1 \leq \alpha \leq L. \tag{44}$$

Our main result on DS-SMDC is the following theorem.

Theorem 2. *$\mathcal{R}_{L,r} = \mathcal{R}^{L,r}_{gp}$.*

Proof. The achievability is immediate from the group pairwise coding scheme. The converse is proved through a sophisticated iteration of information inequalities, which can be found in Section VI. □

Remark 8. From the group pairwise code design and the converse proof in Section VI, we see that both the group pairwise coding scheme and the converse are compatible with $r=1$ and $r=L$. Nevertheless, in order to emphasize the specificity of the case $r=L$ and to distinguish superposition and group pairwise joint coding, we discuss the optimality for $r=L$ separately in the following.

1) Optimality of Superposition Coding for (L,L)-DS-SMDC: For $r=L$, all the messages are protected. We separately encode the L independent messages, where each M_α is encoded using an (α, L) threshold secret sharing scheme. The induced superposition rate region $\mathcal{R}^L_{\text{sup}}$ can be obtained from (8) and (9) by letting $\alpha - N_\alpha = 1$ for all

[1]We use the auxiliary message M_{L+1} to perform as encryption keys for the first r messages if the messages $M_{r+1}, \cdots, M_L$ are not enough. Thus, M_{L+1} is non-vanishing (i.e., $m_{L=1} > 0$), only when the total key size needed is strictly larger than the total size of messages $M_{r+1}, \cdots, M_L$.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2020.3004482, IEEE Transactions on Information Theory

$\alpha \in \mathcal{L}$. To be specific, $\mathcal{R}^L_{\text{sup}}$ is the set of nonnegative rate tuples $\mathbf{R}$ such that

$$\mathsf{R}_l = \sum_{\alpha=1}^{L} r_l^\alpha, \text{ for } l \in \mathcal{L} \tag{45}$$

where $r_l^\alpha \geq 0$, and

$$r_l^\alpha \geq \mathsf{m}_\alpha, \text{ for } 1 \leq \alpha \leq L. \tag{46}$$

It is easy to eliminate r_l^α $(l, \alpha \in \mathcal{L})$ and obtain the following equivalent characterization of the superposition region,

$$\mathcal{R}^L_{\text{sup}} = \{\mathbf{R} : \mathsf{R}_l \geq \sum_{\alpha=1}^{L} \mathsf{m}_\alpha, \text{ for all } l \in \mathcal{L}\}. \tag{47}$$

The following corollary of Theorem 2 states that superposition coding is optimal for the (L, L) DS-SMDC problem.

Corollary 2.1. $\mathcal{R}_{L,L} = \mathcal{R}^L_{sup}$.

Proof. The proof of the converse part is straightforward, so we omit the details and derive the conclusion directly from Theorem 2. It is easily seen by comparing (42)-(44) and (45)-(46) that $\mathcal{R}^{L,r}_{\text{gp}}$ reduces to $\mathcal{R}^L_{\text{sup}}$ for $r = L$. Thus, by Theorem 2, we have $\mathcal{R}_{L,L} = \mathcal{R}^{L,r}_{\text{gp}} = \mathcal{R}^L_{\text{sup}}$. □

IV. Achievability of Theorem 1: Joint Coding Strategies

In order to prove the necessity part of Theorem 1, we instead prove the sufficiency part of Theorem 1', in the two separate cases given in (32) and (33).

A. Low Security Level at Higher Diversity Level

In this section, we provide a joint coding strategy for the case that Condition 1 in (32) holds which provides rate saving, compared to superposition coding. We first discuss a motivating example to illustrate the key insight on how such rate saving is obtained.

Example 2. Let $L = 3, (\alpha, \beta) = (2, 3), (m_2, m_3) = (2, 2), (N_2, N_3) = (0, 1)$, and $p = 5$. Let Z_3 be an independent random key uniformly chosen from $\mathbb{F}_p$. Let the two messages be encoded with generator matrices constructed using MDS-A, which induce the coded symbols as shown in Table I(a) through superposition. The important insight is that the coded message of M_2 can be used as the secret key to encode M_3, which reduces the coding rate. More precisely, we replace Z_3 by $Y_2^1 = Z_2^1 + Z_2^2$ to serve as the key for M_3. The coded symbols for this joint coding strategy are shown in Table I(b). By comparing the two tables, it is seen that the sum rate is reduced since the coded symbol Z_3 is eliminated.

The reconstruction requirements of both M_2 and M_3 are straightforward. There is no security requirement on M_2. For M_3, it is seen that any one coded symbol W_l reveals no information about M_3. For instance, eavesdropping W_2 gives

TABLE I: Coding strategy for Example 2

	W_1	W_2	W_3
$\alpha = 2$	$Y_2^1 = M_2^1 + M_2^2$	$2M_2^1 + M_2^2$	$M_2^1 + 2M_2^2$
$\beta = 3$	Z_3	$M_3^1 + 2M_3^2 + Z_3$	$2M_3^1 + M_3^2 + Z_3$

(a) Superposition coding strategy

	W_1	W_2	W_3
$\alpha = 2$	$Y_2^1 = M_2^1 + M_2^2$	$2M_2^1 + M_2^2$	$M_2^1 + 2M_2^2$
$\beta = 3$		$M_3^1 + 2M_3^2 + Y_2^1$	$2M_3^1 + M_3^2 + Y_2^1$

(b) Joint coding strategy

TABLE II: Coding strategy to replace encryption keys for M_β

	W_1	W_2	$\cdots$	W_θ	$W_{\theta+1}$	$\cdots$	W_L
α	Y_α^1	Y_α^2	$\cdots$	Y_α^θ	$Y_\alpha^{\theta+1}$	$\cdots$	Y_α^L
β	Y_β^1	Y_β^2	$\cdots$	Y_β^θ	$Y_\beta^{\theta+1}$	$\cdots$	Y_β^L

(a) Superposition coding strategy

	W_1	W_2	$\cdots$	W_θ	$W_{\theta+1}$	$\cdots$	W_L
α	Y_α^1	Y_α^2	$\cdots$	Y_α^θ	$Y_\alpha^{\theta+1}$	$\cdots$	Y_α^L
β					$Y_\beta^{(\theta+1)*}$	$\cdots$	Y_β^{L*}

(b) Joint coding strategy

$$\begin{aligned}
&H(M_3|W_2) = H(M_3|M_3^1 + 2M_3^2 + Y_2^1, M_2^1 + Y_2^1) && (48)\\
&= H(M_3, M_3^1 + 2M_3^2 + Y_2^1|M_2^1 + Y_2^1) \\
&\quad - H(M_3^1 + 2M_3^2 + Y_2^1|M_2^1 + Y_2^1) && (49)\\
&= H(M_3, M_3^1 + 2M_3^2 + Y_2^1) - H(M_3^1 + 2M_3^2 + Y_2^1) && (50)\\
&= H(M_3|M_3^1 + 2M_3^2 + Y_2^1) && (51)\\
&= H(M_3), && (52)
\end{aligned}$$

where (50) follows from that M_2^1 is independent of $M_3, M_3^1 + 2M_3^2 + Y_2^1, Y_2^1$ and M_2^1 is independent of $M_3^1 + 2M_3^2 + Y_2^1, Y_2^1$.

Coding strategy for general parameters:
First encode separately M_α and M_β with generator matrices G_α and G_β using MDS-A in Section II-C. The coded symbols for superposition coding strategy are as given in Table II(a). The joint coding strategy we propose is then to replace the first $\theta = \min\{N_\beta, \alpha - N_\beta\}$ encryption key symbols $(Z_\beta^1\ Z_\beta^2\ \cdots\ Z_\beta^\theta)$ by the coded symbols $(Y_\alpha^1, Y_\alpha^2, \cdots, Y_\alpha^\theta)$. The parameter θ is strictly positive, which is implied by Condition 1 in (32). Denote the corresponding codewords for M_β thus obtained as $(Y_\beta^{1*}, Y_\beta^{2*}, \cdots, Y_\beta^{L*})$. The joint coding strategy of M_α and M_β is illustrated in Table II(b) and can be described as follows:

$$W_i = \begin{cases} Y_\alpha^i, & \text{for } 1 \leq i \leq \theta \\ [Y_\alpha^i, Y_\beta^{i*}], & \text{for } \theta < i \leq L. \end{cases} \tag{53}$$

By comparing Table II(a) and Table II(b), it can be seen that the coding rate is reduced compared to superposition coding because $(Y_\beta^1, Y_\beta^2, \cdots, Y_\beta^\theta)$ are removed from the codewords, while the rates for all the others are unchanged. Next, we verify the reconstruction and security constraints for the two messages.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2020.3004482, IEEE Transactions on Information Theory

Reconstruction: The verification of the reconstruction requirements of both M_α and M_β is straightforward.

Security: We consider the security requirements for the two levels separately.

1) Assume we can access N_α coded symbols $W_{\mathcal{B}}, |\mathcal{B}| = N_\alpha$. Partition $\mathcal{B}$ into $\mathcal{B}_1$ and $\mathcal{B}_2$ such that $\mathcal{B}_1 \subseteq \{1, 2, \cdots, \theta\}$ and $\mathcal{B}_2 \subseteq \{\theta+1, \cdots, L\}$. Notice that

$$
\begin{aligned}
& H(Y_\beta^{*\mathcal{B}_2} | M_\alpha, Y_\alpha^{\mathcal{B}_1} Y_\alpha^{\mathcal{B}_2}) \\
& \geq H(Y_\beta^{*\mathcal{B}_2} | M_\alpha, Y_\alpha^{1:\theta}, Y_\alpha^{\mathcal{B}_2}) && (54)\\
& = H(Y_\beta^{*\mathcal{B}_2} | Y_\alpha^{1:\theta}) && (55)\\
& = H(Y_\beta^{*\mathcal{B}_2}), && (56)
\end{aligned}
$$

where the second equality follows from the fact that conditioning does not increase entropy, and the last equality follows from Lemma 2 because

$$
\begin{aligned}
|\mathcal{B}_2| + \theta & \leq N_\alpha + \theta && (57)\\
& = N_\alpha + \min\{\alpha - N_\beta, N_\beta\} && (58)\\
& \leq \alpha && (59)\\
& < \beta, && (60)
\end{aligned}
$$

where the second inequality follows from $N_\alpha < N_\beta$ which is part of Condition 1 in (32). Since conditioning does not increase entropy, in light of (56), we obtain

$$
H(Y_\beta^{*\mathcal{B}_2} | M_\alpha, Y_\alpha^{\mathcal{B}_1} Y_\alpha^{\mathcal{B}_2}) = H(Y_\beta^{*\mathcal{B}_2}). \quad (61)
$$

It follows that

$$
\begin{aligned}
& I(W_{\mathcal{B}}; M_\alpha) \\
& = I(W_{\mathcal{B}_1} W_{\mathcal{B}_2}; M_\alpha) \\
& = I(Y_\alpha^{\mathcal{B}_1}\ Y_\alpha^{\mathcal{B}_2} Y_\beta^{*\mathcal{B}_2}; M_\alpha) && (62)\\
& = I(Y_\alpha^{\mathcal{B}_1} Y_\alpha^{\mathcal{B}_2}; M_\alpha) + I(Y_\beta^{*\mathcal{B}_2}; M_\alpha | Y_\alpha^{\mathcal{B}_1} Y_\alpha^{\mathcal{B}_2}) && (63)\\
& = I(Y_\beta^{*\mathcal{B}_2}; M_\alpha | Y_\alpha^{\mathcal{B}_1} Y_\alpha^{\mathcal{B}_2}) && (64)\\
& = 0, && (65)
\end{aligned}
$$

where the last but one equality follows from Lemma 3 and the fact that $|\mathcal{B}_1| + |\mathcal{B}_2| = N_\alpha$, and (65) follows from (61). Thus indeed $W_{\mathcal{B}}$ reveals nothing about M_α.

2) Assume we can access N_β coded symbols $W_{\mathcal{B}}, |\mathcal{B}| = N_\beta$. Partition $\mathcal{B}$ into $\mathcal{B}_1$ and $\mathcal{B}_2$ such that $\mathcal{B}_1 \subseteq \{1, 2, \cdots, \theta\}$ and $\mathcal{B}_2 \subseteq \{\theta+1, \cdots, L\}$. We first consider

$$
\begin{aligned}
& H(Y_\alpha^{\mathcal{B}_2} | Y_\alpha^{\mathcal{B}_1} Y_\beta^{*\mathcal{B}_2}) \\
& \geq H(Y_\alpha^{\mathcal{B}_2} | Y_\alpha^{\mathcal{B}_1} Y_\beta^{*\mathcal{B}_2} M_\beta) && (66)\\
& \geq H(Y_\alpha^{\mathcal{B}_2} | Y_\alpha^1 \cdots Y_\alpha^\theta, Z_\beta^{\theta+1} \cdots Z_\beta^{N_\beta}, M_\beta Y_\beta^{*\mathcal{B}_2}) && (67)\\
& = H(Y_\alpha^{\mathcal{B}_2} | Y_\alpha^1 \cdots Y_\alpha^\theta, Z_\beta^{\theta+1} \cdots Z_\beta^{N_\beta}, M_\beta) && (68)\\
& = H(Y_\alpha^{\mathcal{B}_2}, Y_\alpha^1 \cdots Y_\alpha^\theta | Z_\beta^{\theta+1} \cdots Z_\beta^{N_\beta}, M_\beta) \\
& - H(Y_\alpha^1 \cdots Y_\alpha^\theta | Z_\beta^{\theta+1} \cdots Z_\beta^{N_\beta}, M_\beta) && (69)\\
& = H(Y_\alpha^{\mathcal{B}_2}, Y_\alpha^1 \cdots Y_\alpha^\theta) - H(Y_\alpha^1 \cdots Y_\alpha^\theta) && (70)\\
& = H(Y_\alpha^{\mathcal{B}_2}), && (71)
\end{aligned}
$$

where both (66) and (67) follow from the fact that conditioning does not increase entropy, (68) follows from that $Y_\beta^{*\mathcal{B}_2}$ is a function of $(Y_\alpha^1 \cdots Y_\alpha^\theta, Z_\beta^{\theta+1} \cdots Z_\beta^{N_\beta}, M_\beta)$, (70) follows from that $(Z_\beta^{\theta+1} \cdots Z_\beta^{N_\beta}, M_\beta)$ are independent of $(Y_\alpha^{\mathcal{B}_2}, Y_\alpha^1 \cdots Y_\alpha^\theta)$, and the last equality follows from Lemma 2, since $|\mathcal{B}_2| + \theta \leq \alpha$ which is induced by $\theta \leq \alpha - N_\beta$. Since conditioning does not increase entropy, in light of (71), we obtain

$$
\begin{aligned}
& H(Y_\alpha^{\mathcal{B}_2} | Y_\alpha^{\mathcal{B}_1} Y_\beta^{*\mathcal{B}_2} M_\beta) \\
& = H(Y_\alpha^{\mathcal{B}_2} | Y_\alpha^{\mathcal{B}_1} Y_\beta^{*\mathcal{B}_2}) = H(Y_\alpha^{\mathcal{B}_2}). && (72)
\end{aligned}
$$

Then we have

$$
\begin{aligned}
& I(W_{\mathcal{B}}; M_\beta) = I(W_{\mathcal{B}_1} W_{\mathcal{B}_2}; M_\beta) && (73)\\
& = I(Y_\alpha^{\mathcal{B}_1}\ Y_\alpha^{\mathcal{B}_2} Y_\beta^{*\mathcal{B}_2}; M_\beta) && (74)\\
& = I(Y_\alpha^{\mathcal{B}_1} Y_\beta^{*\mathcal{B}_2}; M_\beta) + I(Y_\alpha^{\mathcal{B}_2}; M_\beta | Y_\alpha^{\mathcal{B}_1} Y_\beta^{*\mathcal{B}_2}) && (75)\\
& = I(Y_\alpha^{\mathcal{B}_2}; M_\beta | Y_\alpha^{\mathcal{B}_1} Y_\beta^{*\mathcal{B}_2}) && (76)\\
& = H(Y_\alpha^{\mathcal{B}_2} | Y_\alpha^{\mathcal{B}_1} Y_\beta^{*\mathcal{B}_2}) - H(Y_\alpha^{\mathcal{B}_2} | Y_\alpha^{\mathcal{B}_1} Y_\beta^{*\mathcal{B}_2} M_\beta) && (77)\\
& = H(Y_\alpha^{\mathcal{B}_2}) - H(Y_\alpha^{\mathcal{B}_2}) && (78)\\
& = 0, && (79)
\end{aligned}
$$

where (76) follows from Lemma 3 and the fact that $|\mathcal{B}_1| + |\mathcal{B}_2| = N_\beta$, and (78) follows from (72). Thus we obtain that $W_{\mathcal{B}}$ reveals nothing about M_β.

B. Reversed Security Level

We next provide a joint coding strategy for the case that Condition 2 in (33) holds.

Example 3. Let $L = 4, (\alpha, \beta) = (3, 4), (m_3, m_4) = (1, 1), (N_3, N_4) = (2, 1)$, and $p = 11$. We use generator matrix G_3 generated using MDS-B to encode M_3 separately with encryption keys Z_1, Z_2, as given in (80). The joint coding strategy is simply to use M_4 to replace Z_1 as secret keys to encrypt M_3, as given in (81).

$$
\begin{aligned}
& M_3 + 2Z_1 + 9Z_2,\ 9M_3 + 8Z_1 + 6Z_2, \\
& 6M_3 + 10Z_1 + 7Z_2,\ 7M_3 + 9Z_1 + 7Z_2; && (80)\\
\longrightarrow & M_3 + 2M_4 + 9Z_2,\ 9M_3 + 8M_4 + 6Z_2, \\
& 6M_3 + 10M_4 + 7Z_2,\ 7M_3 + 9M_4 + 7Z_2. && (81)
\end{aligned}
$$

Since M_4 does not need to be separately encoded, rate saving is obtained. The reconstruction and security requirements of M_3 are immediate. The reconstruction requirement of M_4 is straightforward since everything is recovered with any three coded symbols. The security requirement of M_4 can be easily seen that any one coded symbol reveals nothing about M_4.

Coding strategy for general parameters:

Next, we present the general coding strategy that M_β performs as secret keys for M_α so that we can reduce the coding rates. Let G_α be a generator matrix generated using MDS-B in Section II-C, which can be used to encode M_α separately with encryption keys $(Z_1, Z_2, \ldots, Z_{N_\alpha})$. The joint coding strategy is simply to use $\eta = \min\{N_\alpha, \alpha - N_\beta\}$ symbols of the message M_β (i.e., $M_\beta^1, M_\beta^2, \cdots, M_\beta^\eta$) to replace the encryption keys $(Z_1\ Z_2\ \cdots\ Z_\eta)$ for encrypting M_α. The parameter η is strictly positive, which is implied by Condition 2 in (33) as well as $\alpha > N_\alpha$. Denote the corresponding coded symbols for M_α after this replacement as $(Y_\alpha^{1*}, Y_\alpha^{2*}, \cdots, Y_\alpha^{L*})$. Since the η message symbols of M_β do not need to be separately encoded, rate saving is thus obtained. Next, we verify the reconstruction and security constraints.

Reconstruction: By the code construction in Section II-C, both the message M_α and the keys M_β can be losslessly recovered from any α coded symbols. Since $\alpha < \beta$, the reconstruction requirements of both M_α and M_β are satisfied immediately.

Security: The security constraint of M_α is straightforward, and thus let us consider M_β. For any $\mathcal{B} \subseteq \mathcal{L}$ such that $|\mathcal{B}| = N_\beta$, let $Y_\alpha^{*\mathcal{B}} = (Y_\alpha^{i*} : i \in \mathcal{B})$. By Lemma 4, we have

$$I(Y_\alpha^{*\mathcal{B}}; M_\beta^1, M_\beta^2, \cdots, M_\beta^\eta) = 0, \tag{82}$$

since $\eta \leq \alpha - N_\beta$.

V. Converse of Theorem 1

To show the optimality of Theorem 1, we only need to prove that under the condition in (27), the sum rate is lower bounded by (10), i.e.,

$$\sum_{l=1}^{L} \mathsf{R}_l \geq \sum_{\alpha=1}^{L} \frac{L\mathsf{m}_\alpha}{\alpha - N_\alpha}. \tag{83}$$

For any $\alpha \in \mathcal{L}$, let $\mathbb{B}_\alpha$ be the set of *disjoint subset* pairs $(\mathcal{B}_\alpha^1, \mathcal{B}_\alpha^2)$ such that $\mathcal{B}_\alpha^1, \mathcal{B}_\alpha^2 \subseteq \mathcal{L}$,

$$|\mathcal{B}_\alpha^1| = \alpha - N_\alpha \text{ and } |\mathcal{B}_\alpha^2| = N_\alpha. \tag{84}$$

For $\alpha \in \mathcal{L}$, let $M_{1:\alpha} \triangleq (M_1, M_2, \cdots, M_\alpha)$. Define μ_α by

$$\mu_\alpha = \frac{L}{\alpha - N_\alpha} \frac{1}{\binom{L}{N_\alpha}\binom{L-N_\alpha}{\alpha-N_\alpha}} \sum_{(\mathcal{B}_\alpha^1, \mathcal{B}_\alpha^2) \in \mathbb{B}_\alpha} H(W_{\mathcal{B}_\alpha^1} | W_{\mathcal{B}_\alpha^2} M_{1:\alpha}). \tag{85}$$

We need the following lemma to proceed.

Lemma 5. *Under the condition in* (27)*, for any $\alpha \in \mathcal{L}$, we have*

$$\sum_{l=1}^{L} H(W_l) \geq \sum_{j=1}^{\alpha} \frac{Lm_j}{j - N_j} + \mu_\alpha. \tag{86}$$

Proof. For $\alpha \leq T_s$, (86) is simply the inequality (27) in [1]. For $\alpha \geq T_s$, we prove the lemma by induction on α. Similar to the proof of Theorem 2 in [1] where Han's inequality plays a key role, we apply Han's inequality and its complementary conditioning version. The details of the proof can be found in Appendix B. □

For $\alpha = L$, in light of (86), we have

$$\sum_{l=1}^{L} R_l = \sum_{l=1}^{L} H(W_l) \geq \sum_{\alpha=1}^{L} \frac{Lm_\alpha}{\alpha - N_\alpha} + \mu_L \geq \sum_{\alpha=1}^{L} \frac{Lm_\alpha}{\alpha - N_\alpha}, \tag{87}$$

from which we can obtain, by normalization, the sum rate bound (83).

Remark 9. It is clear that superposition coding must induce $\mu_L = 0$ under the condition in (27). Since the messages are encoded separately, we can indeed verify that for any $\alpha \in \mathcal{L}$,

$$H(Y_\alpha^{\mathcal{B}_L^1} | Y_\alpha^{\mathcal{B}_L^2} M_\alpha) = 0, \tag{88}$$

where $Y_\alpha^1, Y_\alpha^2, \cdots, Y_\alpha^L$ are coded symbols of M_α and $Y_\alpha^{\mathcal{B}} \triangleq (Y_\alpha^i : i \in \mathcal{B})$ for any $\mathcal{B} \subseteq \mathcal{L}$. To see this, observe that if the weakly secure SMDC problem reduces to classical SMDC, (88) is true immediately. Otherwise, by (27), we have $N_L \geq N_\alpha$ for any $\alpha \in \mathcal{L}$. Since we use an (N_α, α, L) ramp secret sharing code to encode M_α, any α symbols from the set $\{M_\alpha^1, M_\alpha^2, \cdots, M_\alpha^{\alpha-N_\alpha}, Y_\alpha^1, Y_\alpha^2, \cdots, Y_\alpha^L\}$ can completely recover the whole set. Thus, $(Y_\alpha^{\mathcal{B}_L^2}, M_\alpha)$ provide complete information about $Y_\alpha^{\mathcal{B}_L^1}$, which verifies (88).

VI. Converse Proof of Theorem 2

Before proving Theorem 2, we introduce some terminologies and notations in [3]. Let $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \cdots, \lambda_L)$ and

$$\mathbb{R}_+^L = \{\boldsymbol{\lambda} : \ \boldsymbol{\lambda} \neq \mathbf{0} \text{ and } \lambda_i \in \mathbb{R}, \lambda_i \geq 0 \text{ for } i \in \mathcal{L}\}. \tag{89}$$

Let $\Omega_L^\alpha = \{\boldsymbol{v} \in \{0,1\}^L : |\boldsymbol{v}| = \alpha\}$, where $|\boldsymbol{v}|$ is the Hamming weight of a vector $\boldsymbol{v} = (v_1, v_2, \cdots, v_L)$. For any $\boldsymbol{v} \in \Omega_L^\alpha$, let $c_\alpha(\boldsymbol{v})$ be any nonnegative real number. For any $\boldsymbol{\lambda} \in \mathbb{R}_+^L$ and $\alpha \in \mathcal{L}$, let $f_\alpha(\boldsymbol{\lambda})$ be the optimal solution to the following optimization problem:

$$f_\alpha(\boldsymbol{\lambda}) \triangleq \quad \max \sum_{\boldsymbol{v} \in \Omega_L^\alpha} c_\alpha(\boldsymbol{v}) \tag{90}$$

$$\text{s.t.} \sum_{\boldsymbol{v} \in \Omega_L^\alpha} c_\alpha(\boldsymbol{v}) \cdot \boldsymbol{v} \leq \boldsymbol{\lambda} \tag{91}$$

$$c_\alpha(\boldsymbol{v}) \geq 0, \forall \boldsymbol{v} \in \Omega_L^\alpha. \tag{92}$$

A set $\{c_\alpha(\boldsymbol{v}) : \boldsymbol{v} \in \Omega_L^\alpha\}$ is called an α-*resolution* for $\boldsymbol{\lambda}$ if (91) and (92) are satisfied and it will be abbreviated as $\{c_\alpha(\boldsymbol{v})\}$ if there is no ambiguity. Furthermore, an α-resolution is called *optimal* if it achieves the optimal value $f_\alpha(\boldsymbol{\lambda})$. In the following proof, we will take advantage of some lemmas and theorems from [3] and [4], which are enclosed in Appendix A for convenience.

To prove the converse of Theorem 2, we follow the idea of Theorem 2 in [3], i.e., we provide an alternative characterization of the group pairwise region $\mathcal{R}_{\text{gp}}^{L,r}$. For simplicity, let $f_{L+1}(\boldsymbol{\lambda}) = 0$ for all $\boldsymbol{\lambda} \in \mathbb{R}_+^L$. For $\eta \in \{r+1, r+2, \cdots, L+1\}$, let

$$g_\eta(\boldsymbol{\lambda}) = \sum_{\alpha=1}^{r} f_1(\boldsymbol{\lambda})\mathsf{m}_\alpha + \sum_{\alpha=\eta+1}^{L} f_\alpha(\boldsymbol{\lambda})\mathsf{m}_\alpha + f_\eta(\boldsymbol{\lambda}) \left[\sum_{\alpha=r+1}^{\eta} \mathsf{m}_\alpha - \sum_{\alpha=1}^{r} (\alpha - 1)\mathsf{m}_\alpha\right]. \tag{93}$$

In particular, for $\eta = \eta^*$ which is defined by (39), we have

$$g_{\eta^*}(\boldsymbol{\lambda}) = \sum_{\alpha=1}^{r} f_1(\boldsymbol{\lambda})\mathsf{m}_\alpha + \sum_{\alpha=r+1}^{L} f_\alpha(\boldsymbol{\lambda})\mathsf{m}_\alpha^*, \tag{94}$$

where m_α^* is defined in (41). From the group pairwise coding scheme in Fig. 3, we have the following intuitions on the coding rates.

i. Superposition of $M_1, M_2, \cdots, M_r$ induces the rate

$$\sum_{l=1}^{L} \lambda_l \mathsf{R}_l = \sum_{\alpha=1}^{r} f_1(\boldsymbol{\lambda})\mathsf{m}_\alpha. \tag{95}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2020.3004482, IEEE Transactions on Information Theory

ii. The messages $M_{r+1}, M_{r+2}, \cdots, M_{\eta^*-1}, M^1_{\eta^*}$ perform as keys for $M_1, M_2, \cdots, M_r$. Thus, we do not need extra rates to encode them beyond the rate given in (95).
iii. The other messages $M^2_{\eta^*}, M_{\eta^*+1}, \cdots, M_L$ will be encoded in the same way as in classical SMDC, i.e., superposition coding. The coding rate is characterized in [3] using the technique of α-resolution, which is

$$\sum_{l=1}^{L} \lambda_l \mathsf{R}_l = f_{\eta^*}(\boldsymbol{\lambda}) \left[\sum_{\alpha=r+1}^{\eta^*} \mathsf{m}_\alpha - \sum_{\alpha=1}^{r} (\alpha-1)\mathsf{m}_\alpha \right] + \sum_{\alpha=\eta^*+1}^{L} f_\alpha(\boldsymbol{\lambda})\mathsf{m}_\alpha. \tag{96}$$

Summing up the rates in (95) and (96), we obtain $g_{\eta^*}(\boldsymbol{\lambda})$ which is the total rate of group pairwise coding. Let $\mathcal{R}^*_{L,r}$ be the set of all $\mathbf{R} \geq \mathbf{0}$ such that

$$\boldsymbol{\lambda} \cdot \mathbf{R} \geq g_{\eta^*}(\boldsymbol{\lambda}). \tag{97}$$

In particular, for $\boldsymbol{\lambda} = (100\cdots)$ and $\eta^* = L+1$, the constraint in (97) becomes the single rate bound

$$\mathsf{R}_l \geq \sum_{\alpha=1}^{r} \mathsf{m}_\alpha. \tag{98}$$

For $\boldsymbol{\lambda} = \mathbf{1}$, the constraint in (97) becomes the sum rate bound

$$\mathsf{R}^*_{\text{sum}} = \sum_{\alpha=1}^{r} (L-\alpha+1)\mathsf{m}_\alpha + \sum_{\alpha=r+1}^{\eta^*} \frac{L\mathsf{m}_\alpha}{\eta^*} + \sum_{\alpha=\eta^*+1}^{L} \frac{L\mathsf{m}_\alpha}{\alpha}. \tag{99}$$

For $\eta^* = r+1$, the constraint becomes

$$\boldsymbol{\lambda} \cdot \mathbf{R} \geq \sum_{\alpha=1}^{r} \left[f_1(\boldsymbol{\lambda}) - (\alpha-1) f_{r+1}(\boldsymbol{\lambda}) \right] \mathsf{m}_\alpha + \sum_{\alpha=r+1}^{L} f_\alpha(\boldsymbol{\lambda})\mathsf{m}_\alpha. \tag{100}$$

Inspired by the above intuitions on the group pairwise coding rates, we can alternatively characterize $\mathcal{R}^{L,r}_{\text{gp}}$ in another equivalent form, given in the following theorem.

Theorem 3. $\mathcal{R}^{L,r}_{gp} = \mathcal{R}^*_{L,r}$.

Proof. See Appendix C. □

To complete the converse proof of Theorem 2, in light of the fact $\mathcal{R}^{L,r}_{\text{gp}} \subseteq \mathcal{R}_{\text{L,r}}$ as well as Theorem 3, we now only need to show $\mathcal{R}_{\text{L,r}} \subseteq \mathcal{R}^*_{L,r}$, i.e., for any $\mathbf{R} \in \mathcal{R}_{\text{L,r}}$, the following inequality holds

$$\boldsymbol{\lambda} \cdot \mathbf{R} \geq g_{\eta^*}(\boldsymbol{\lambda}). \tag{101}$$

The following lemma provides an alternative representation of $g_{\eta^*}(\boldsymbol{\lambda})$.

Lemma 6. $\max_{\eta=r+1,\cdots,L+1} \{g_\eta(\boldsymbol{\lambda})\} = g_{\eta^*}(\boldsymbol{\lambda})$.

Proof. See Appendix D. □

By Lemma 6, it only remains to show that for any $\mathbf{R} \in \mathcal{R}_{\text{L,r}}$ and $\eta = r+1, \cdots, L+1$, the inequality $\boldsymbol{\lambda} \cdot \mathbf{R} \geq g_\eta(\boldsymbol{\lambda})$ holds. The converse for SMDC in [3] is proved using iterations to extract the entropies $H(M_1), H(M_2), \cdots, H(M_L)$ successively with coefficient $f_1(\boldsymbol{\lambda}), f_2(\boldsymbol{\lambda}), \cdots, f_L(\boldsymbol{\lambda})$ which have the same form of expression. In the secure setting here, the desired inequality $\boldsymbol{\lambda} \cdot \mathbf{R} \geq g_\eta(\boldsymbol{\lambda})$ will have two forms of coefficients, i.e., coefficients related to the secure messages and those related to the non-secure messages. The latter is the same as that in [3], but the former is different. For this reason, the iterations in the converse proof in [3] do not apply to the former, i.e., the secure messages. Therefore, we need to derive new iterations to extract the entropies of the secure messages, such that the r-th iteration can be connected with the iterations in [3]. Specifically, the main idea of proving $\boldsymbol{\lambda} \cdot \mathbf{R} \geq g_\eta(\boldsymbol{\lambda})$ is as follows:

i) we extract the entropies $H(M_1), H(M_2), \cdots, H(M_L)$ with proper coefficients in (93) from $\sum_{l=1}^{L} \lambda_l H(W_l)$ successively and iteratively;
ii) when extracting $H(M_\alpha)$ for $\alpha \in \{1, 2, \cdots, r\}$, we explicitly design the coefficients of each intermediate term in closed-form so that we can finally connect to the r-th iteration of the converse proof in [3];
iii) for $\alpha \geq r+1$, since there is no security constraints, we simply use the iterations in [3].

One of the main contributions of the converse proof compared with that in [3] is the new technique of explicitly designing the coefficients in closed-form in each iteration for the secure messages. In contrast, in each iteration of the non-secure messages which is simply the iteration in [3], the coefficients in the iteration do not have a closed-form.

Instead of formally proving this inequality here, we provide an example for $(L, r) = (4, 2)$ and $\eta = 3$ to illustrate the main idea, and relegate the formal proof to Appendix E. The connection between this example and the formal proof will be discussed in Remark 11, Remark 12, and Remark 13 in Appendix E. For different $i, j, k \in \{1, 2, 3, 4\}$, we first present two equalities that will be used in the example:

$$\begin{aligned} H(W_i|W_k M_1) &= H(M_2|W_k M_1) + H(W_i|W_k M_{1:2}) \\ &\quad - H(M_2|W_i W_k M_1) \\ &= H(M_2) + H(W_i|W_k M_{1:2}), \end{aligned} \tag{102}$$

$$\begin{aligned} H(W_i W_j W_k|M_{1:2}) &= H(W_i W_j W_k M_3|M_{1:2}) \\ &= H(M_3) + H(W_i W_j W_k M_{1:3}). \end{aligned} \tag{103}$$

Now we can write the following chain of inequalities without much difficulty:

$$\begin{aligned} & R_1 + R_2 + R_3 + R_4 \\ &= H(W_1) + H(W_2) + H(W_3) + H(W_4) \qquad (104) \\ &= 4H(M_1) + H(W_1|M_1) + H(W_2|M_1) \\ &\quad + H(W_3|M_1) + H(W_4|M_1) \qquad (105) \\ &= 4H(M_1) + \underbrace{\Big\{ 0H(W_1|M_1) + \frac{2}{3}H(W_2|M_1) + H(W_3|M_1) + H(W_4|M_1) \Big\}}_{\triangleq S_1} + \underbrace{\Big\{ H(W_1|M_1) + \frac{1}{3}H(W_2|M_1) + 0H(W_3|M_1) + 0H(W_4|M_1) \Big\}}_{\triangleq S_2} \qquad (106) \\ &\geq 4H(M_1) + \frac{8}{3}H(M_2) \\ &\quad + \Big\{ \frac{1}{3}H(W_2 W_3|W_1 M_{1:2}) + \frac{1}{3}H(W_2 W_4|W_1 M_{1:2}) \end{aligned}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2020.3004482, IEEE Transactions on Information Theory

$$
\begin{aligned}
&+\frac{1}{3}H(W_3W_4|W_1M_{1:2})+\frac{1}{3}H(W_3W_4|W_2M_{1:2})\Big\}\\
&+\Big\{H(W_1|M_{1:2})+\frac{1}{3}H(W_2|M_{1:2})\Big\} &(107)\\
&=4H(M_1)+\frac{8}{3}H(M_2)\\
&+\frac{1}{3}H(W_1W_2W_3|M_{1:2})+\frac{1}{3}H(W_1W_2W_4|M_{1:2})\\
&+\frac{1}{3}H(W_1W_3W_4|M_{1:2})+\frac{1}{3}H(W_2W_3W_4|M_{1:2}) &(108)\\
&\overset{(103)}{=}4H(M_1)+\frac{8}{3}H(M_2)+\frac{4}{3}H(M_3)\\
&+\frac{1}{3}H(W_1W_2W_3|M_{1:3})+\frac{1}{3}H(W_1W_2W_4|M_{1:3})\\
&+\frac{1}{3}H(W_1W_3W_4|M_{1:3})+\frac{1}{3}H(W_2W_3W_4|M_{1:3}) &(109)\\
&\geq 4H(M_1)+\frac{8}{3}H(M_2)+\frac{4}{3}H(M_3)+H(M_4) &(110)\\
&=4m_1+\frac{8}{3}m_2+\frac{4}{3}m_3+m_4, &(111)
\end{aligned}
$$

where (110) follows from the fact that $(\frac{1}{3},\frac{1}{3},\frac{1}{3},\frac{1}{3})$ is an optimal 3-resolution for $\boldsymbol{\lambda}=(1,1,1,1)$ (*cf.* (90)-(92)), and the nontrivial step from (106) to (107) can be derived as follows

$$
\begin{aligned}
S_1&=\frac{2}{3}H(W_2|M_1)+H(W_3|M_1)+H(W_4|M_1) &(112)\\
&=\Big[\frac{1}{3}H(W_2|M_1)+\frac{1}{3}H(W_3|M_1)\Big]\\
&+\Big[\frac{1}{3}H(W_2|M_1)+\frac{1}{3}H(W_4|M_1)\Big]\\
&+\Big[\frac{1}{3}H(W_3|M_1)+\frac{1}{3}H(W_4|M_1)\Big]\\
&+\Big[\frac{1}{3}H(W_3|M_1)+\frac{1}{3}H(W_4|M_1)\Big] &(113)\\
&\geq\Big[\frac{1}{3}H(W_2|W_1M_1)+\frac{1}{3}H(W_3|W_1M_1)\Big]\\
&+\Big[\frac{1}{3}H(W_2|W_1M_1)+\frac{1}{3}H(W_4|W_1M_1)\Big]\\
&+\Big[\frac{1}{3}H(W_3|W_1M_1)+\frac{1}{3}H(W_4|W_1M_1)\Big]\\
&+\Big[\frac{1}{3}H(W_3|W_2M_1)+\frac{1}{3}H(W_4|W_2M_1)\Big] &(114)\\
&\overset{(102)}{=}\frac{8}{3}H(M_2)+\Big[\frac{1}{3}H(W_2|W_1M_{1:2})+\frac{1}{3}H(W_3|W_1M_{1:2})\Big]\\
&+\Big[\frac{1}{3}H(W_2|W_1M_{1:2})+\frac{1}{3}H(W_4|W_1M_{1:2})\Big]\\
&+\Big[\frac{1}{3}H(W_3|W_1M_{1:2})+\frac{1}{3}H(W_4|W_1M_{1:2})\Big]\\
&+\Big[\frac{1}{3}H(W_3|W_2M_{1:2})+\frac{1}{3}H(W_4|W_2M_{1:2})\Big] &(115)\\
&\geq\frac{8}{3}H(M_2)\\
&+\Big\{\frac{1}{3}H(W_2W_3|W_1M_{1:2})+\frac{1}{3}H(W_2W_4|W_1M_{1:2})\\
&+\frac{1}{3}H(W_3W_4|W_1M_{1:2})+\frac{1}{3}H(W_3W_4|W_2M_{1:2})\Big\}_{\triangleq S_1'}, &(116)
\end{aligned}
$$

and

$$
\begin{aligned}
S_2&=H(W_1|M_1)+\frac{1}{3}H(W_2|M_1) &(117)\\
&\geq\Big\{H(W_1|M_{1:2})+\frac{1}{3}H(W_2|M_{1:2})\Big\}_{\triangleq S_2'}. &(118)
\end{aligned}
$$

The main ideas of the example are as follows:

1) The two terms S_1 and S_2 have a similar form in (106), but with different coefficient vectors $(0,\frac{2}{3},1,1)$ and $(1,\frac{1}{3},0,0)$, respectively, which are chosen strategically for this bound. The two terms are bounded in rather different manners. We extract $\frac{8}{3}H(M_2)$ from S_1 (with S_1' left) and use S_2 to convert S_1' from the form $H(W_iW_j|W_kM_{1:2})$ to the form $H(W_iW_jW_k|M_{1:2})$. This further generates the terms $H(W_iW_jW_k|M_{1:3})$ in (109), which ensures that the α-resolution technique can be applied subsequently.
2) When bounding S_1, we reorganize its coefficient vector $(0,\frac{2}{3},1,1)$ as given in (113) for two purposes: firstly, Han's inequality can be applied as in (116); secondly, $H(W_iW_j|W_kM_{1:2})$ in S_1' can be converted to $H(W_1W_jW_k|M_{1:2})$ using S_2';
3) The coefficient $(\frac{1}{3},\frac{1}{3},\frac{1}{3},\frac{1}{3})$ in (108)-(109) is an optimal 3-resolution for $(1,1,1,1)$. In the general proof, the α-resolution technique used in the converse proof for SMDC [3] will be invoked in a more systematic manner.

VII. Conclusion

We studied the weakly secure SMDC problem and characterized the condition that superposition coding is optimal in terms of achieving the minimum sum rate. It is generally difficult to design the optimal coding schemes and characterize the rate regions for those cases that superposition is suboptimal. In this paper, we consider a special case called differential-constant secure SMDC, for which the optimal rate region is characterized. A group pairwise coding scheme is shown to be optimal in terms of achieving the entire rate region.

The optimality condition is proved only for the minimum sum rate, we conjecture that it is also the optimality condition that superposition coding can achieve the entire rate region. This is currently under our investigation.

Appendix A
Some lemmas/theorems from [3], [4]

In the following lemmas and theorem, we assume $\boldsymbol{\lambda}\in\mathbb{R}_+^L$ (c.f. (89)) is ordered, i.e., $\lambda_1\geq\lambda_2\geq\cdots\geq\lambda_L$. Let $\{c(\boldsymbol{v})\}$ be an α-resolution for $\boldsymbol{\lambda}$ (c.f. (91),(92)) and $\tilde{\boldsymbol{\lambda}}=\sum_{\boldsymbol{v}\in\Omega_L^\alpha}c(\boldsymbol{v})\cdot\boldsymbol{v}$. An α-resolution is called *perfect* if the equality in (91) holds, i.e., $\sum_{\boldsymbol{v}\in\Omega_L^\alpha}c(\boldsymbol{v})\cdot\boldsymbol{v}=\boldsymbol{\lambda}$.

Lemma 2 in [3]: *Let $\{c(\boldsymbol{v})\}$ be an optimal α-resolution for $\boldsymbol{\lambda}$. Then there exists $0\leq l\leq\alpha-1$ such that $\lambda_i-\tilde{\lambda}_i>0$ if and only if $1\leq i\leq l$.*

Lemma 4 in [3]:

(i) $f_\alpha(\boldsymbol{\lambda})\leq\alpha^{-1}\sum_{i=1}^L\lambda_i$;

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2020.3004482, IEEE Transactions on Information Theory

(ii) $\sum_{\boldsymbol{v}\in\Omega_L^\alpha} c(\boldsymbol{v}) = \alpha^{-1}\sum_{i=1}^L \lambda_i$ *if and only if* $\{c(\boldsymbol{v})\}$ *is a perfect* α*-resolution for* $\boldsymbol{\lambda}$*. In this case,* $f_\alpha(\boldsymbol{\lambda}) = \alpha^{-1}\sum_{i=1}^L \lambda_i$.

Lemma 7 in [3]: *For* $\alpha \geq 2$*,* $\boldsymbol{\lambda}$ *has a perfect* α*-resolution if and only if* $\lambda_1 \leq \frac{\lambda_2+\cdots+\lambda_L}{\alpha-1}$.

Theorem 1 in [4]: $f_\alpha(\boldsymbol{\lambda}) = \min_{\beta\in\{0,1,\cdots,\alpha-1\}} \frac{1}{\alpha-\beta}\sum_{i=\beta+1}^L \lambda_i$.

Lemma 1 in [4]: *For* $\alpha \geq 2$*, if* $\lambda_1 \leq \frac{\lambda_2+\lambda_3+\cdots\lambda_L}{\alpha-1}$*, then* $f_\alpha(\boldsymbol{\lambda}) = \frac{1}{\alpha}\sum_{i=1}^L \lambda_i$.

For any permutation ω on $\{1,2,\cdots,L\}$, denote $\big(\lambda_{\omega(1)},\lambda_{\omega(2)},\cdots,\lambda_{\omega(L)}\big)$ by $\omega(\boldsymbol{\lambda})$.

Lemma 2 in [4]: $f_\alpha\big(\omega(\boldsymbol{\lambda})\big) = f_\alpha(\boldsymbol{\lambda})$ *for any* $\alpha\in\mathcal{L}$.

Lemma 5 in [4]: *Let* $\boldsymbol{\lambda}_1 = (\lambda_{1,1},\lambda_{1,2},\cdots,\lambda_{1,L})$ *and* $\boldsymbol{\lambda}_2 = (\lambda_{2,1},\lambda_{2,2},\cdots,\lambda_{2,L})$ *be two ordered vectors such that* $\lambda_{1,1} > \lambda_{2,1}$ *and* $\lambda_{1,i} = \lambda_{2,i}$ *for all* $2\leq i\leq L$*. For any* $\alpha_0\in\mathcal{L}$*, if* $f_{\alpha_0}(\boldsymbol{\lambda}_1) = f_{\alpha_0}(\boldsymbol{\lambda}_2)$*, then* $f_\alpha(\boldsymbol{\lambda}_1) = f_\alpha(\boldsymbol{\lambda}_2)$ *for all* $\alpha\geq\alpha_0$.

Let $\boldsymbol{\lambda}^{[1]}$ be the length-L vector with the first component being 1 and the rest being 0, i.e., $\boldsymbol{\lambda}^{[1]} = (1,0,0,\cdots,0)$.

Lemma 6 in [4]: *If* $\lambda_1 > \sum_{i=2}^L \lambda_i$*, let* $\boldsymbol{\lambda}' = \Big(\sum_{i=2}^L \lambda_i,\lambda_2,\lambda_3,\cdots,\lambda_L\Big)$*. Then for all* $\alpha\in\mathcal{L}$*,*

$$f_\alpha(\boldsymbol{\lambda}) = \left(\lambda_1 - \sum_{i=2}^L \lambda_i\right) f_\alpha\left(\boldsymbol{\lambda}^{[1]}\right) + f_\alpha(\boldsymbol{\lambda}').$$

Lemma 7 in [4]: *For any* $\eta\in\{1,2,\cdots,L-1\}$*,*

(i) *if* $\lambda_1 \leq \frac{1}{\eta}\sum_{i=2}^L \lambda_i$*, then* $f_\alpha(\boldsymbol{\lambda}) = \frac{1}{\alpha}\sum_{i=1}^L\lambda_i$ *for* $\alpha = 1,2,\cdots,\eta+1$*;*

(ii) *if* $\lambda_1 \geq \frac{1}{\eta}\sum_{i=2}^L \lambda_i$*, then* $f_\alpha(\boldsymbol{\lambda}) = f_{\alpha-1}(\lambda_2,\lambda_3,\cdots,\lambda_L)$ *for* $\alpha = \eta+1,\eta+2,\cdots,L$.

Appendix B
Proof of Lemma 5

When condition (27) is satisfied, there must exist a T_s as defined in (29). For $\alpha\leq T_s$, since $N_\alpha = 0$, we have

$$\mu_\alpha = \frac{L}{\alpha}\frac{1}{\binom{L}{\alpha}}\sum_{\mathcal{B}_\alpha^1\subseteq\mathcal{L}:|\mathcal{B}_\alpha^1|=\alpha} H(W_{\mathcal{B}_\alpha^1}|M_{1:\alpha}). \tag{119}$$

The claim in (86) is exactly inequality (27) in [1] which was proved by applying Han's inequality.

For any $\alpha\geq T_s$, we prove the claim by induction. Firstly, the claim is true for $\alpha = T_s$. Then we assume the claim is true for $\alpha=\zeta$ for some $\zeta\geq T_s$. We now show that it is true for $\alpha=\varphi$, which is the index of the next non-vanishing message following M_ζ. In light of (86), we only need to show that

$$\mu_\zeta \geq \frac{L}{\varphi-N_\varphi}H(M_\varphi)+\mu_\varphi. \tag{120}$$

Since now $\varphi>\zeta>0$, the first condition in (27) must hold, i.e.,

$$N_\zeta<\zeta\leq N_\varphi<\varphi. \tag{121}$$

For any $(\mathcal{B}_\varphi^1,\mathcal{B}_\varphi^2)\in\mathbb{B}_\varphi$ and $(\mathcal{B}_\zeta^1,\mathcal{B}_\zeta^2)\in\mathbb{B}_\zeta$ such that $\mathcal{B}_\zeta^2\subseteq\mathcal{B}_\varphi^2$, from the reconstruction and security constraints of M_φ, we obtain

$$\begin{aligned}
H(M_\varphi) &= H(M_\varphi|W_{\mathcal{B}_\varphi^2}) - H(M_\varphi|W_{\mathcal{B}_\varphi^1}W_{\mathcal{B}_\varphi^2}) && (122)\\
&= I(M_\varphi;W_{\mathcal{B}_\varphi^1}|W_{\mathcal{B}_\varphi^2}) && (123)\\
&= H(W_{\mathcal{B}_\varphi^1}|W_{\mathcal{B}_\varphi^2}) - H(W_{\mathcal{B}_\varphi^1}|W_{\mathcal{B}_\varphi^2}M_\varphi) && (124)\\
&= H(W_{\mathcal{B}_\varphi^1}|W_{\mathcal{B}_\varphi^2}M_{1:\zeta}) - H(W_{\mathcal{B}_\varphi^1}|W_{\mathcal{B}_\varphi^2}M_{1:\varphi}) && (125)
\end{aligned}$$

where (125) follows from the fact that $N_\varphi\geq\zeta$ and the reconstruction constraints of $M_1,M_2,\cdots,M_\zeta$. In the following, we prove the iteration of (120) in two different situations:

i. $\zeta-N_\zeta\leq\varphi-N_\varphi$;
ii. $\zeta-N_\zeta>\varphi-N_\varphi$.

Remark 10. It is easy to see by checking the following proof that the case of $\zeta-N_\zeta=\varphi-N_\varphi$ is compatible with both (i) and (ii).

Case i. $\underline{\zeta-N_\zeta\leq\varphi-N_\varphi}$: Consider the following,

$$\begin{aligned}
\mu_\zeta &= \frac{L}{\zeta-N_\zeta}\frac{1}{\binom{L}{N_\zeta}\binom{L-N_\zeta}{\zeta-N_\zeta}}\cdot\sum_{(\mathcal{B}_\zeta^1,\mathcal{B}_\zeta^2)\in\mathbb{B}_\zeta} H(W_{\mathcal{B}_\zeta^1}|W_{\mathcal{B}_\zeta^2}M_{1:\zeta})\\
&\geq \frac{L}{\zeta-N_\zeta}\frac{1}{\binom{L}{N_\zeta}\binom{L-N_\zeta}{\zeta-N_\zeta}}\cdot\\
&\qquad\sum_{(\mathcal{B}_\zeta^1,\mathcal{B}_\zeta^2)\in\mathbb{B}_\zeta}\ \sum_{\substack{\mathcal{V}\subseteq\mathcal{L}\setminus(\mathcal{B}_\zeta^1\cup\mathcal{B}_\zeta^2):\\|\mathcal{V}|=N_\varphi-N_\zeta}} \frac{1}{\binom{L-\zeta}{N_\varphi-N_\zeta}} H(W_{\mathcal{B}_\zeta^1}|W_{\mathcal{B}_\zeta^2\cup\mathcal{V}}M_{1:\zeta}) && (126)\\
&= \frac{L}{\zeta-N_\zeta}\frac{1}{\binom{L}{N_\zeta}\binom{L-N_\zeta}{\zeta-N_\zeta}}\cdot\\
&\qquad\sum_{\substack{\mathcal{B}_\zeta^1\subseteq\mathcal{L}:\\|\mathcal{B}_\zeta^1|=\zeta-N_\zeta}}\ \sum_{\substack{\mathcal{B}_\varphi^2\subseteq\mathcal{L}\setminus\mathcal{B}_\zeta^1:\\|\mathcal{B}_\varphi^2|=N_\varphi}} \frac{\binom{N_\varphi}{N_\zeta}}{\binom{L-\zeta}{N_\varphi-N_\zeta}} H(W_{\mathcal{B}_\zeta^1}|W_{\mathcal{B}_\varphi^2}M_{1:\zeta})\\
&= \frac{L}{\zeta-N_\zeta}\frac{\binom{L-N_\varphi-1}{\zeta-N_\zeta-1}}{\binom{L}{N_\zeta}\binom{L-N_\zeta}{\zeta-N_\zeta}}\frac{\binom{N_\varphi}{N_\zeta}}{\binom{L-\zeta}{N_\varphi-N_\zeta}}\cdot\\
&\qquad\sum_{\substack{\mathcal{B}_\varphi^2\subseteq\mathcal{L}:\\|\mathcal{B}_\varphi^2|=N_\varphi}}\ \sum_{\substack{\mathcal{B}_\zeta^1\subseteq\mathcal{L}\setminus\mathcal{B}_\varphi^2:\\|\mathcal{B}_\zeta^1|=\zeta-N_\zeta}} \frac{1}{\binom{L-N_\varphi-1}{\zeta-N_\zeta-1}} H(W_{\mathcal{B}_\zeta^1}|W_{\mathcal{B}_\varphi^2}M_{1:\zeta})\\
&\geq \frac{L}{\zeta-N_\zeta}\frac{\binom{L-N_\varphi-1}{\zeta-N_\zeta-1}}{\binom{L}{N_\zeta}\binom{L-N_\zeta}{\zeta-N_\zeta}}\frac{\binom{N_\varphi}{N_\zeta}}{\binom{L-\zeta}{N_\varphi-N_\zeta}}\cdot\\
&\qquad\sum_{(\mathcal{B}_\varphi^1,\mathcal{B}_\varphi^2)\in\mathbb{B}_\varphi} \frac{1}{\binom{L-N_\varphi-1}{\varphi-N_\varphi-1}} H(W_{\mathcal{B}_\varphi^1}|W_{\mathcal{B}_\varphi^2}M_{1:\zeta}) && (127)\\
&= \frac{L}{\varphi-N_\varphi}\frac{1}{\binom{L}{N_\varphi}\binom{L-N_\varphi}{\varphi-N_\varphi}}\cdot\sum_{(\mathcal{B}_\varphi^1,\mathcal{B}_\varphi^2)\in\mathbb{B}_\varphi} H(W_{\mathcal{B}_\varphi^1}|W_{\mathcal{B}_\varphi^2}M_{1:\zeta})\\
&= \frac{L}{\varphi-N_\varphi}\frac{1}{\binom{L}{N_\varphi}\binom{L-N_\varphi}{\varphi-N_\varphi}}\cdot\\
&\qquad\sum_{(\mathcal{B}_\varphi^1,\mathcal{B}_\varphi^2)\in\mathbb{B}_\varphi}\Big[H(M_\varphi)+H(W_{\mathcal{B}_\varphi^1}|W_{\mathcal{B}_\varphi^2}M_{1:\varphi})\Big] && (128)\\
&= \frac{L}{\varphi-N_\varphi}H(M_\varphi)+\mu_\varphi, && (129)
\end{aligned}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2020.3004482, IEEE Transactions on Information Theory

where (126) follows from the fact that conditioning does not increase entropy, (127) follows from Han's inequality and the assumption that $\zeta - N_\zeta \le \varphi - N_\varphi$, and (128) follows from (125).

Case ii. $\underline{\zeta - N_\zeta > \varphi - N_\varphi}$: We derive the iteration as follows,

$$\begin{aligned}
\mu_\zeta &= \frac{L}{\zeta - N_\zeta} \frac{1}{\binom{L}{N_\zeta}\binom{L-N_\zeta}{\zeta-N_\zeta}} \cdot \sum_{(\mathcal{B}^1_\zeta,\mathcal{B}^2_\zeta)\in\mathbb{B}_\zeta} H(W_{\mathcal{B}^1_\zeta}|W_{\mathcal{B}^2_\zeta} M_{1:\zeta}) \\
&\ge \frac{L}{\zeta - N_\zeta} \frac{1}{\binom{L}{N_\zeta}\binom{L-N_\zeta}{\zeta-N_\zeta}} \cdot \\
&\quad \sum_{(\mathcal{B}^1_\zeta,\mathcal{B}^2_\zeta)\in\mathbb{B}_\zeta} \sum_{\substack{\mathcal{V}\subseteq\mathcal{L}\setminus(\mathcal{B}^1_\zeta\cup\mathcal{B}^2_\zeta):\\ |\mathcal{V}|=\varphi-\zeta}} \frac{1}{\binom{L-\zeta}{\varphi-\zeta}} H(W_{\mathcal{B}^1_\zeta}|W_{\mathcal{B}^2_\zeta\cup\mathcal{V}} M_{1:\zeta}) \\
&= \frac{L}{\zeta - N_\zeta} \frac{1}{\binom{L}{N_\zeta}\binom{L-N_\zeta}{\zeta-N_\zeta}} \cdot \\
&\quad \sum_{\substack{\mathcal{B}^1_\zeta\subseteq\mathcal{L}:\\ |\mathcal{B}^1_\zeta|=\zeta-N_\zeta}} \sum_{\substack{\mathcal{V}'\subseteq\mathcal{L}\setminus\mathcal{B}^1_\zeta:\\ |\mathcal{V}'|=\varphi-(\zeta-N_\zeta)}} \frac{\binom{\varphi-(\zeta-N_\zeta)}{N_\zeta}}{\binom{L-\zeta}{\varphi-\zeta}} H(W_{\mathcal{B}^1_\zeta}|W_{\mathcal{V}'} M_{1:\zeta}) \\
&= \frac{L}{\zeta - N_\zeta} \frac{1}{\binom{L}{N_\zeta}\binom{L-N_\zeta}{\zeta-N_\zeta}} \frac{\binom{\varphi-(\zeta-N_\zeta)}{N_\zeta}}{\binom{L-\zeta}{\varphi-\zeta}} \cdot \\
&\quad \sum_{\substack{\mathcal{D}\subseteq\mathcal{L}:\\ |\mathcal{D}|=\varphi}} \sum_{\substack{\mathcal{B}^1_\zeta\subseteq\mathcal{D}:\\ |\mathcal{B}^1_\zeta|=\zeta-N_\zeta}} \sum_{\mathcal{V}'=\mathcal{D}\setminus\mathcal{B}^1_\zeta} H(W_{\mathcal{B}^1_\zeta}|W_{\mathcal{V}'} M_{1:\zeta}) \\
&= \frac{L}{\zeta - N_\zeta} \frac{1}{\binom{L}{N_\zeta}\binom{L-N_\zeta}{\zeta-N_\zeta}} \frac{\binom{\varphi-(\zeta-N_\zeta)}{N_\zeta}}{\binom{L-\zeta}{\varphi-\zeta}} \binom{\varphi}{\zeta - N_\zeta} (\zeta - N_\zeta) \cdot \\
&\quad \sum_{\substack{\mathcal{D}\subseteq\mathcal{L}:\\ |\mathcal{D}|=\varphi}} \sum_{\substack{\mathcal{B}^1_\zeta\subseteq\mathcal{D}:\\ |\mathcal{B}^1_\zeta|=\zeta-N_\zeta}} \sum_{\mathcal{B}^2_\varphi=\mathcal{D}\setminus\mathcal{B}^1_\zeta} \frac{1}{\binom{\varphi}{\zeta-N_\zeta}(\zeta - N_\zeta)} H(W_{\mathcal{B}^1_\zeta}|W_{\mathcal{V}'} M_{1:\zeta}) \\
&\ge \frac{L}{\zeta - N_\zeta} \frac{1}{\binom{L}{N_\zeta}\binom{L-N_\zeta}{\zeta-N_\zeta}} \frac{\binom{\varphi-(\zeta-N_\zeta)}{N_\zeta}}{\binom{L-\zeta}{\varphi-\zeta}} \binom{\varphi}{\zeta - N_\zeta} (\zeta - N_\zeta) \cdot \\
&\quad \sum_{\substack{\mathcal{D}\subseteq\mathcal{L}:\\ |\mathcal{D}|=\varphi}} \sum_{\substack{\mathcal{B}^1_\varphi\subseteq\mathcal{D}:\\ |\mathcal{B}^1_\varphi|=\varphi-N_\varphi}} \sum_{\mathcal{V}''=\mathcal{D}\setminus\mathcal{B}^1_\varphi} \frac{1/(\varphi - N_\varphi)}{\binom{\varphi}{\varphi-N_\varphi}} H(W_{\mathcal{B}^1_\varphi}|W_{\mathcal{V}''} M_{1:\zeta}) \qquad (130) \\
&= \frac{L}{\varphi - N_\varphi} \frac{1}{\binom{L}{N_\varphi}\binom{L-N_\varphi}{\varphi-N_\varphi}} \cdot \\
&\quad \sum_{\substack{\mathcal{D}\subseteq\mathcal{L}:\\ |\mathcal{D}|=\varphi}} \sum_{\substack{\mathcal{B}^1_\varphi\subseteq\mathcal{D}:\\ |\mathcal{B}^1_\varphi|=\varphi-N_\varphi}} \sum_{\mathcal{B}^2_\varphi=\mathcal{D}\setminus\mathcal{B}^1_\varphi} H(W_{\mathcal{B}^1_\varphi}|W_{\mathcal{B}^2_\varphi} M_{1:\zeta}) \\
&= \frac{L}{\varphi - N_\varphi} \frac{1}{\binom{L}{N_\varphi}\binom{L-N_\varphi}{\varphi-N_\varphi}} \cdot \sum_{(\mathcal{B}^1_\varphi,\mathcal{B}^2_\varphi)\in\mathbb{B}_\varphi} H(W_{\mathcal{B}^1_\varphi}|W_{\mathcal{B}^2_\varphi} M_{1:\zeta}) \\
&= \frac{L}{\varphi - N_\varphi} \frac{1}{\binom{L}{N_\varphi}\binom{L-N_\varphi}{\varphi-N_\varphi}} \cdot \\
&\quad \sum_{(\mathcal{B}^1_\varphi,\mathcal{B}^2_\varphi)\in\mathbb{B}_\varphi} \left[H(M_\varphi) + H(W_{\mathcal{B}^1_\varphi}|W_{\mathcal{B}^2_\varphi} M_{1:\varphi})\right] \qquad (131) \\
&= \frac{L}{\varphi - N_\varphi} H(M_\varphi) \\
&\quad + \frac{L}{\varphi - N_\varphi} \frac{1}{\binom{L}{N_\varphi}\binom{L-N_\varphi}{\varphi-N_\varphi}} \sum_{(\mathcal{B}^1_\varphi,\mathcal{B}^2_\varphi)\in\mathbb{B}_\varphi} H(W_{\mathcal{B}^1_\varphi}|W_{\mathcal{B}^2_\varphi} M_{1:\varphi}) \\
&= \frac{L}{\varphi - N_\varphi} H(M_\varphi) + \mu_\varphi, \qquad (132)
\end{aligned}$$

where (130) follows from Han's inequality (complementary conditioning version), and (131) follows from (125). This proves Lemma 5.

Appendix C
Proof of Theorem 3

Similar to Lemma 11 in [3], the theorem can be obtained by proving i) $\mathcal{R}^{L,r}_{\text{gp}} \subseteq \mathcal{R}^*_{L,r}$; ii) for any $\boldsymbol{\lambda} \in \mathbb{R}^L_+$, there exists $\mathbf{R} \in \mathcal{R}^{L,r}_{\text{gp}}$ such that $\boldsymbol{\lambda} \cdot \mathbf{R} = g_{\eta^*}(\boldsymbol{\lambda})$.

i) We first show that $\mathcal{R}^{L,r}_{\text{gp}} \subseteq \mathcal{R}^*_{L,r}$. For any $1 \le \alpha \le r$, let $\boldsymbol{r}^\alpha = (r^\alpha_1, r^\alpha_2, \cdots, r^\alpha_L)$. For any $\boldsymbol{\lambda} \in \mathbb{R}^L_+$ and $\mathbf{R} \in \mathcal{R}^{L,r}_{\text{gp}}$, we have from (43) that

$$\boldsymbol{\lambda} \cdot \boldsymbol{r}^\alpha \ge f_1(\boldsymbol{\lambda})\mathsf{m}_\alpha. \qquad (133)$$

For $r+1 \le \alpha \le L$, let $\{c_\alpha(\boldsymbol{v})\}$ be an optimal α-resolution for $\boldsymbol{\lambda}$, which implies that

$$\boldsymbol{\lambda} \ge \sum_{\boldsymbol{v}\in\Omega^\alpha_L} c_\alpha(\boldsymbol{v})\boldsymbol{v}. \qquad (134)$$

Then we have

$$\begin{aligned}
\boldsymbol{\lambda} \cdot \boldsymbol{r}^\alpha &\ge \left(\sum_{\boldsymbol{v}\in\Omega^\alpha_L} c_\alpha(\boldsymbol{v})\boldsymbol{v}\right) \cdot \boldsymbol{r}^\alpha && (135)\\
&= \sum_{\boldsymbol{v}\in\Omega^\alpha_L} \left(c_\alpha(\boldsymbol{v})(\boldsymbol{v} \cdot \boldsymbol{r}^\alpha)\right) && (136)\\
&\ge \sum_{\boldsymbol{v}\in\Omega^\alpha_L} \left(c_\alpha(\boldsymbol{v})\mathsf{m}^*_\alpha\right) && (137)\\
&= \left(\sum_{\boldsymbol{v}\in\Omega^\alpha_L} c_\alpha(\boldsymbol{v})\right) \mathsf{m}^*_\alpha && (138)\\
&= f_\alpha(\boldsymbol{\lambda})\mathsf{m}^*_\alpha && (139)
\end{aligned}$$

where (135) follows from (134), (137) follows from (44), and (139) follows from the optimality of $\{c_\alpha(\boldsymbol{v})\}$. Summing up (133) and (139) over α, we have

$$\begin{aligned}
\boldsymbol{\lambda} \cdot \mathbf{R} &\ge \sum_{\alpha=1}^{r} f_1(\boldsymbol{\lambda})\mathsf{m}_\alpha + \sum_{\alpha=r+1}^{L} f_\alpha(\boldsymbol{\lambda})\mathsf{m}^*_\alpha && (140)\\
&= g_{\eta^*}(\boldsymbol{\lambda}). && (141)
\end{aligned}$$

This implies $\mathbf{R} \in \mathcal{R}^*_{L,r}$ and thus $\mathcal{R}^{L,r}_{\text{gp}} \subseteq \mathcal{R}^*_{L,r}$.

ii) We now construct a rate tuple $\mathbf{R}$ for each $\boldsymbol{\lambda} \in \mathbb{R}^L_+$ such that $\mathbf{R} \in \mathcal{R}^{L,r}_{\text{gp}}$ and $\boldsymbol{\lambda} \cdot \mathbf{R} = g_{\eta^*}(\boldsymbol{\lambda})$. For $r+1 \le \alpha \le L$, let $\{c_\alpha(\boldsymbol{v})\}$ be an optimal α-resolution for $\boldsymbol{\lambda}$ and let

$$\tilde{\boldsymbol{\lambda}} = \sum_{\boldsymbol{v}\in\Omega^\alpha_L} c_\alpha(\boldsymbol{v}) \cdot \boldsymbol{v}. \qquad (142)$$

By Lemma 2 in [3], there exists $1 \le l_\alpha \le \alpha - 1$ such that $\lambda_i > \tilde{\lambda}_i$ if and only if $1 \le i \le l_\alpha$. Let $\mathsf{R}_l = \sum_{\alpha=1}^{L} r^\alpha_l$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2020.3004482, IEEE Transactions on Information Theory

for $l \in \mathcal{L}$. We construct $\mathbf{R}$ by designing the sub-rates r_l^α as follows.

a) For $1 \le \alpha \le r$, let
$$r_l^\alpha = \mathsf{m}_\alpha,\ \text{for all } 1 \le l \le L. \tag{143}$$

b) For $r+1 \le \alpha \le \eta^* - 1$, let
$$r_l^\alpha = 0,\ \text{for all } 1 \le l \le L. \tag{144}$$

c) For $\eta^* \le \alpha \le L$, let
$$r_l^\alpha = \begin{cases} 0, & \text{for } 1 \le l \le l_\alpha \\ \frac{\mathsf{m}_\alpha^*}{\alpha - l_\alpha}, & \text{for } l_\alpha + 1 \le l \le L. \end{cases} \tag{145}$$

We first verify that such a construction implies $\mathbf{R} \in \mathcal{R}_{\text{gp}}^{L,r}$.

a) For $1 \le \alpha \le r$, it is obvious that (43) is satisfied.

b) For $r+1 \le \alpha \le \eta^*-1$, since $\mathsf{m}_\alpha^* = 0$, (44) is satisfied.

c) For $\eta^* \le \alpha \le L$, consider any $\mathcal{B} \subseteq \mathcal{L}$ such that $|\mathcal{B}| = \alpha$. Let $\boldsymbol{e}_\alpha$ be an L-vector with the first l_α components being 0 and the last $L - l_\alpha$ components being 1. Let $\boldsymbol{v}_\mathcal{B} = (v_1, v_2, \cdots, v_L)$ be such that $v_i = 1$ if and only if $i \in \mathcal{B}$. Since $\sum_{i=1}^{l_\alpha} v_i \le l_\alpha$, we have $\boldsymbol{e}_\alpha \cdot \boldsymbol{v}_\mathcal{B} \ge \alpha - l_\alpha$. Thus,
$$\begin{aligned} \sum_{l\in\mathcal{B}} r_l^\alpha &= \left(\frac{\mathsf{m}_\alpha^*}{\alpha - l_\alpha}\boldsymbol{e}_\alpha\right)\cdot \boldsymbol{v}_\mathcal{B} && (146)\\ &= \frac{\mathsf{m}_\alpha^*}{\alpha - l_\alpha}(\boldsymbol{e}_\alpha \cdot \boldsymbol{v}_\mathcal{B}) && (147)\\ &\ge \frac{\mathsf{m}_\alpha^*}{\alpha - l_\alpha}(\alpha - l_\alpha) && (148)\\ &= \mathsf{m}_\alpha^*. && (149)\end{aligned}$$

Thus, $\mathbf{R} \in \mathcal{R}_{\text{gp}}^{L,r}$. Now it remains to show that $\boldsymbol{\lambda} \cdot \mathbf{R} = g_{\eta^*}(\boldsymbol{\lambda})$. We consider the following cases.

a) For $1 \le \alpha \le r$, it is easy to check that
$$\boldsymbol{\lambda} \cdot \boldsymbol{r}^\alpha = f_1(\boldsymbol{\lambda})\mathsf{m}_\alpha. \tag{150}$$

b) For $r+1 \le \alpha \le \eta^* - 1$, it is obvious that
$$\boldsymbol{\lambda} \cdot \boldsymbol{r}^\alpha = 0. \tag{151}$$

c) For $\eta^* \le \alpha \le L$, only the first l_α components of $\boldsymbol{\lambda} - \tilde{\boldsymbol{\lambda}}$ are nonzero. Thus, we have
$$\left(\boldsymbol{\lambda} - \sum_{\boldsymbol{v}\in\Omega_L^\alpha} c_\alpha(\boldsymbol{v})\boldsymbol{v}\right)\cdot \boldsymbol{r}^\alpha = 0, \tag{152}$$
which implies that
$$\boldsymbol{\lambda}\cdot\boldsymbol{r}^\alpha = \left(\sum_{\boldsymbol{v}\in\Omega_L^\alpha} c_\alpha(\boldsymbol{v})\boldsymbol{v}\right)\cdot\boldsymbol{r}^\alpha = \sum_{\boldsymbol{v}\in\Omega_L^\alpha}\big(c_\alpha(\boldsymbol{v})(\boldsymbol{v}\cdot\boldsymbol{r}^\alpha)\big). \tag{153}$$
By Lemma 2 in [3], for any $\boldsymbol{v} \in \Omega_L^\alpha$ such that $c_\alpha(\boldsymbol{v}) > 0$, the first l_α components are equal to 1, $(\alpha - l_\alpha)$ of the other $L - l_\alpha$ components are equal to 1, and the rest are equal to 0. On the other hand, the first l_α components of $\boldsymbol{r}^\alpha$ are equal to zero. Thus, for any $\boldsymbol{v} \in \Omega_L^\alpha$ such that $c_\alpha(\boldsymbol{v}) > 0$, we have
$$\boldsymbol{v}\cdot\boldsymbol{r}^\alpha = (\alpha - l_\alpha)\frac{\mathsf{m}_\alpha^*}{\alpha - l_\alpha} = \mathsf{m}_\alpha^*. \tag{154}$$
Then
$$\begin{aligned} \boldsymbol{\lambda}\cdot\boldsymbol{r}^\alpha &= \sum_{\boldsymbol{v}\in\Omega_L^\alpha}\big(c_\alpha(\boldsymbol{v})(\boldsymbol{v}\cdot\boldsymbol{r}^\alpha)\big) && (155)\\ &= \sum_{\boldsymbol{v}\in\Omega_L^\alpha} c_\alpha(\boldsymbol{v})\mathsf{m}_\alpha^* && (156)\\ &= \left(\sum_{\boldsymbol{v}\in\Omega_L^\alpha} c_\alpha(\boldsymbol{v})\right)\mathsf{m}_\alpha^* && (157)\\ &= f_\alpha(\boldsymbol{\lambda})\mathsf{m}_\alpha^*. && (158)\end{aligned}$$

Summing up (150), (151), and (158) over all $1 \le \alpha \le L$, we obtain $\boldsymbol{\lambda}\cdot\mathbf{R} = g_{\eta^*}(\boldsymbol{\lambda})$. Therefore, Theorem 3 is proved.

Appendix D
Proof of Lemma 6

We prove the lemma by proving (i) for $r+1 \le \eta^* \le L$, $\sum_{\alpha=1}^r(\alpha-1)\mathsf{m}_\alpha \le \sum_{\alpha=r+1}^{\eta^*}\mathsf{m}_\alpha$ is equivalent to $g_{\eta^*}(\boldsymbol{\lambda}) \ge g_{\eta^*+1}(\boldsymbol{\lambda}) \ge \cdots \ge g_{L+1}(\boldsymbol{\lambda})$; (ii) for $r+2 \le \eta^* \le L+1$, $\sum_{\alpha=r+1}^{\eta^*-1}\mathsf{m}_\alpha < \sum_{\alpha=1}^r(\alpha-1)\mathsf{m}_\alpha$ is equivalent to $g_{\eta^*}(\boldsymbol{\lambda}) > g_{\eta^*-1}(\boldsymbol{\lambda}) > \cdots > g_{r+1}(\boldsymbol{\lambda})$.

(i) For $\eta^* \le \eta \le L$, we have
$$g_\eta(\boldsymbol{\lambda}) \ge g_{\eta+1}(\boldsymbol{\lambda}) \tag{159}$$
$$\Updownarrow$$
$$\begin{aligned} &\sum_{\alpha=\eta+1}^{L} f_\alpha(\boldsymbol{\lambda})\mathsf{m}_\alpha + f_\eta(\boldsymbol{\lambda})\left[\sum_{\alpha=r+1}^{\eta}\mathsf{m}_\alpha - \sum_{\alpha=1}^{r}(\alpha-1)\mathsf{m}_\alpha\right]\\ &\ge \sum_{\alpha=\eta+2}^{L} f_\alpha(\boldsymbol{\lambda})\mathsf{m}_\alpha\\ &\quad + f_{\eta+1}(\boldsymbol{\lambda})\left[\sum_{\alpha=r+1}^{\eta+1}\mathsf{m}_\alpha - \sum_{\alpha=1}^{r}(\alpha-1)\mathsf{m}_\alpha\right] \end{aligned} \tag{160}$$
$$\Updownarrow$$
$$\begin{aligned} &f_{\eta+1}(\boldsymbol{\lambda})\mathsf{m}_{\eta+1} + f_\eta(\boldsymbol{\lambda})\left[\sum_{\alpha=r+1}^{\eta}\mathsf{m}_\alpha - \sum_{\alpha=1}^{r}(\alpha-1)\mathsf{m}_\alpha\right]\\ &\ge f_{\eta+1}(\boldsymbol{\lambda})\left[\sum_{\alpha=r+1}^{\eta+1}\mathsf{m}_\alpha - \sum_{\alpha=1}^{r}(\alpha-1)\mathsf{m}_\alpha\right] \end{aligned} \tag{161}$$
$$\Updownarrow$$
$$(f_\eta(\boldsymbol{\lambda}) - f_{\eta+1}(\boldsymbol{\lambda}))\left[\sum_{\alpha=r+1}^{\eta}\mathsf{m}_\alpha - \sum_{\alpha=1}^{r}(\alpha-1)\mathsf{m}_\alpha\right] \ge 0 \tag{162}$$
$$\Updownarrow$$
$$\sum_{\alpha=r+1}^{\eta}\mathsf{m}_\alpha \ge \sum_{\alpha=1}^{r}(\alpha-1)\mathsf{m}_\alpha. \tag{163}$$

Thus, we conclude that
$$g_{\eta^*}(\boldsymbol{\lambda}) \ge g_{\eta^*+1}(\boldsymbol{\lambda}) \ge \cdots \ge g_{L+1}(\boldsymbol{\lambda}) \tag{164}$$
is equivalent to
$$\sum_{\alpha=1}^{r}(\alpha-1)\mathsf{m}_\alpha \le \sum_{\alpha=r+1}^{\eta}\mathsf{m}_\alpha \text{ for all } \eta^* \le \eta \le L+1, \tag{165}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2020.3004482, IEEE Transactions on Information Theory

which is also equivalent to

$$\sum_{\alpha=1}^{r}(\alpha-1)\mathsf{m}_\alpha \leq \sum_{\alpha=r+1}^{\eta^*}\mathsf{m}_\alpha. \tag{166}$$

(ii) For $r+1 \leq \eta \leq \eta^*$, we have

$$g_\eta(\boldsymbol{\lambda}) > g_{\eta-1}(\boldsymbol{\lambda}) \tag{167}$$

$$\Updownarrow$$

$$\sum_{\alpha=\eta+1}^{L} f_\alpha(\boldsymbol{\lambda})\mathsf{m}_\alpha + f_\eta(\boldsymbol{\lambda})\left[\sum_{\alpha=r+1}^{\eta}\mathsf{m}_\alpha - \sum_{\alpha=1}^{r}(\alpha-1)\mathsf{m}_\alpha\right]$$
$$> \sum_{\alpha=\eta}^{L} f_\alpha(\boldsymbol{\lambda})\mathsf{m}_\alpha$$
$$+ f_{\eta-1}(\boldsymbol{\lambda})\left[\sum_{\alpha=r+1}^{\eta-1}\mathsf{m}_\alpha - \sum_{\alpha=1}^{r}(\alpha-1)\mathsf{m}_\alpha\right] \tag{168}$$

$$\Updownarrow$$

$$f_\eta(\boldsymbol{\lambda})\left[\sum_{\alpha=r+1}^{\eta}\mathsf{m}_\alpha - \sum_{\alpha=1}^{r}(\alpha-1)\mathsf{m}_\alpha\right]$$
$$> f_\eta(\boldsymbol{\lambda})\mathsf{m}_\eta + f_{\eta-1}(\boldsymbol{\lambda})\left[\sum_{\alpha=r+1}^{\eta-1}\mathsf{m}_\alpha - \sum_{\alpha=1}^{r}(\alpha-1)\mathsf{m}_\alpha\right] \tag{169}$$

$$\Updownarrow$$

$$(f_{\eta-1}(\boldsymbol{\lambda}) - f_\eta(\boldsymbol{\lambda}))\left[\sum_{\alpha=r+1}^{\eta-1}\mathsf{m}_\alpha - \sum_{\alpha=1}^{r}(\alpha-1)\mathsf{m}_\alpha\right] < 0 \tag{170}$$

$$\Updownarrow$$

$$\sum_{\alpha=r+1}^{\eta-1}\mathsf{m}_\alpha < \sum_{\alpha=1}^{r}(\alpha-1)\mathsf{m}_\alpha. \tag{171}$$

Thus, we conclude that

$$g_{\eta^*}(\boldsymbol{\lambda}) > g_{\eta^*-1}(\boldsymbol{\lambda}) > \cdots > g_{r+1}(\boldsymbol{\lambda}) \tag{172}$$

is equivalent to

$$\sum_{\alpha=r+1}^{\eta-1}\mathsf{m}_\alpha < \sum_{\alpha=1}^{r}(\alpha-1)\mathsf{m}_\alpha \text{ for all } r+1 \leq \eta \leq \eta^*, \tag{173}$$

which is also equivalent to

$$\sum_{\alpha=r+1}^{\eta^*-1}\mathsf{m}_\alpha < \sum_{\alpha=1}^{r}(\alpha-1)\mathsf{m}_\alpha. \tag{174}$$

Appendix E
Converse Proof of Theorem 2 (continuing)

In order to prove the inequality in (97), i.e., $\boldsymbol{\lambda}\cdot\mathbf{R} \geq g_{\eta^*}(\boldsymbol{\lambda})$, we first introduce some lemmas and important parameters that will be used. The connection between the example at the end of Section VI and the general converse proof here will be provided when the corresponding parameters are defined.

Similar to Lemma 6 in [4], the following lemma gives a sufficient condition of redundancy in the characterization of the rate region.

Lemma 7. *For any $\eta = r+1, r+2, \cdots, L+1$, the rate constraint $\boldsymbol{\lambda}\cdot\mathbf{R} \geq g_\eta(\boldsymbol{\lambda})$ is redundant in the characterization of $\mathcal{R}^*_{L,r}$ if*

$$\lambda_1 > \frac{\lambda_2+\lambda_3+\cdots+\lambda_L}{\eta-1}. \tag{175}$$

Proof. See Appendix F. □

For any $\eta \in \{r+1, r+2, \cdots, L+1\}$, $\boldsymbol{\lambda}$ is called an *η-considerable coefficient vector* if $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_L$ and

$$\lambda_1 \leq \frac{\lambda_2+\lambda_3+\cdots+\lambda_L}{\eta-1}. \tag{176}$$

Denote the set of all η-considerable coefficient vectors by $\mathbb{R}^L_\eta$. Then let

$$\mathbb{R}^L_{\text{con}} = \bigcup_{\eta=r+1}^{L+1}\mathbb{R}^L_\eta. \tag{177}$$

We have the following property on vectors in $\mathbb{R}^L_{\text{con}}$, for which a simple proof is given in Appendix G.

Lemma 8. *For $\eta = r+1, r+2, \cdots, L+1$ and $\boldsymbol{\lambda} \in \mathbb{R}^L_\eta$, we have $f_\eta(\boldsymbol{\lambda}) \geq \lambda_1$.*

By Lemma 7, we only need to prove $\boldsymbol{\lambda}\cdot\mathbf{R} \geq g_{\eta^*}(\boldsymbol{\lambda})$ for $\boldsymbol{\lambda} \in \mathbb{R}^L_{\text{con}}$. Thus, we assume $\boldsymbol{\lambda} \in \mathbb{R}^L_{\text{con}}$ in the sequel. From Theorem 1 in [4], we can verify that

$$\sum_{i=1}^{L}\lambda_i = f_1(\boldsymbol{\lambda}) \geq \eta f_\eta(\boldsymbol{\lambda}), \tag{178}$$

which implies that

$$\sum_{i=1}^{L}\lambda_i - (r-1)f_\eta(\boldsymbol{\lambda}) \geq [\eta-(r-1)]\, f_\eta(\boldsymbol{\lambda}) > 0. \tag{179}$$

Let $\xi_\alpha \in \mathcal{L}$ be the index of $\boldsymbol{\lambda}$ such that

$$\sum_{i=1}^{\xi_\alpha-1}\lambda_i < \alpha f_\eta(\boldsymbol{\lambda}) \leq \sum_{i=1}^{\xi_\alpha}\lambda_i. \tag{180}$$

For simplicity, let $\xi_0 = 1$. From Lemma 8, we can see that

$$f_\eta(\boldsymbol{\lambda}) \geq \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_L, \tag{181}$$

which implies

$$\xi_0 \leq \xi_1 < \xi_2 < \cdots < \xi_r. \tag{182}$$

and

$$\xi_i \geq i. \tag{183}$$

Due to (179), we can subtract $r-1$ of $f_\eta(\boldsymbol{\lambda})$ one by one from the sequence $\lambda_1, \lambda_2, \cdots, \lambda_L$. The subtraction process is illustrated in Fig. 4. For $\alpha = 1, 2, \cdots, r-1$, let $\boldsymbol{\gamma}^{(\alpha)} = \left(\gamma_1^{(\alpha)}, \gamma_2^{(\alpha)}, \cdots, \gamma_L^{(\alpha)}\right)$ be the α-th *subtraction* and $\boldsymbol{\lambda}^{(\alpha)} = \left(\lambda_1^{(\alpha)}, \lambda_2^{(\alpha)}, \cdots, \lambda_L^{(\alpha)}\right)$ be the α-th *residue* after the first α subtractions such that

$$\gamma_i^{(\alpha)} = \begin{cases} \sum_{i=1}^{\xi_{\alpha-1}}\lambda_i - (\alpha-1)f_\eta(\boldsymbol{\lambda}), & \text{if } i = \xi_{\alpha-1} \\ \alpha f_\eta(\boldsymbol{\lambda}) - \sum_{i=1}^{\xi_\alpha-1}\lambda_i, & \text{if } i = \xi_\alpha \\ \lambda_i, & \text{if } \xi_{\alpha-1} < i < \xi_\alpha \\ 0, & \text{if } i < \xi_{\alpha-1} \text{ or } i > \xi_\alpha. \end{cases} \tag{184}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2020.3004482, IEEE Transactions on Information Theory

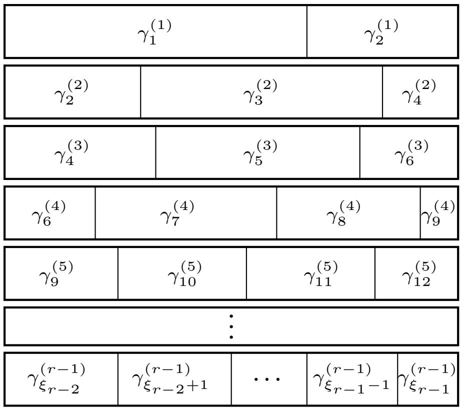

Fig. 4: Illustration of $\gamma_i^{(\alpha)}$

and $\lambda_i^{(\alpha)} = \lambda_i - \sum_{j=1}^{\alpha} \gamma_i^{(j)}$. Thus,

$$\lambda_i^{(\alpha)} = \begin{cases} 0, & \text{if } i < \xi_\alpha \\ \sum_{i=1}^{\xi_\alpha} \lambda_i - \alpha f_\eta(\boldsymbol{\lambda}), & \text{if } i = \xi_\alpha \\ \lambda_i, & \text{if } i > \xi_\alpha \end{cases} \tag{185}$$

It is easy to check that

$$\sum_{i=\xi_{\alpha-1}}^{\xi_\alpha} \gamma_i^{(\alpha)} = f_\eta(\boldsymbol{\lambda}) \tag{186}$$

and

$$\gamma_{\xi_\alpha}^{(\alpha)} + \gamma_{\xi_\alpha}^{(\alpha+1)} = \lambda_{\xi_\alpha}. \tag{187}$$

Remark 11. In the example at the end of Section VI, the subtraction and residue parameters are the coefficients in (106), which is $\boldsymbol{\lambda}^{(1)} = (0, \frac{2}{3}, 1, 1)$ and $\gamma^{(1)} = (1, \frac{1}{3}, 0, 0)$.

Let $\boldsymbol{\lambda}^{(r-1)} = \left(\lambda_{\xi_{r-1}}^{(r-1)}, \lambda_{\xi_{r-1}+1}^{(r-1)}, \cdots, \lambda_L^{(r-1)}\right)$. The following lemma will be used in the converse. The detailed proof of the lemma is given in Appendix H.

Lemma 9. $f_{\eta-(r-1)}\left(\boldsymbol{\lambda}^{(r-1)}\right) \geq f_\eta(\boldsymbol{\lambda})$.

By the definition of $f_{\eta-(r-1)}\left(\boldsymbol{\lambda}^{(r-1)}\right)$ in (90), the value of the objective function $\sum_{\boldsymbol{v}\in\Omega_{L-\xi_{r-1}+1}^{\eta-(r-1)}} c_{\eta-(r-1)}(\boldsymbol{v})$ lies in the range $\left[0, f_{\eta-(r-1)}\left(\boldsymbol{\lambda}^{(r-1)}\right)\right]$. The inequality in Lemma 9 implies $f_\eta(\boldsymbol{\lambda}) \in \left[0, f_{\eta-(r-1)}\left(\boldsymbol{\lambda}^{(r-1)}\right)\right]$. Thus, there exists an $[\eta-(r-1)]$-resolution $\left\{c_{\eta-(r-1)}(\boldsymbol{v}) : \boldsymbol{v} \in \Omega_{L-\xi_{r-1}+1}^{\eta-(r-1)}\right\}$ for $\boldsymbol{\lambda}^{(r-1)}$ such that

$$\sum_{\boldsymbol{v}\in\Omega_{L-\xi_{r-1}+1}^{\eta-(r-1)}} c_{\eta-(r-1)}(\boldsymbol{v}) = f_\eta(\boldsymbol{\lambda}). \tag{188}$$

For $\boldsymbol{v} \in \Omega_{L-\xi_{r-1}+1}^{\eta-(r-1)}$ such that $c_{\eta-(r-1)}(\boldsymbol{v}) > 0$, let $\boldsymbol{v} = (v_1, v_2, \cdots, v_{L-\xi_{r-1}+1})$ and

$$D_{\boldsymbol{v}} = \left\{i \in \{\xi_{r-1}, \xi_{r-1}+1, \cdots, L\} : v_{i-\xi_{r-1}+1} = 1\right\}. \tag{189}$$

Let $\mathcal{D} = \left\{D_{\boldsymbol{v}} : \boldsymbol{v} \in \Omega_{L-\xi_{r-1}+1}^{\eta-(r-1)},\ c_{\eta-(r-1)}(\boldsymbol{v}) > 0\right\}$ and $|\mathcal{D}| = b_1$. For simplicity, let $\mathcal{D} = \{D_1, D_2, \cdots, D_{b_1}\}$. For $k = \{1, 2, \cdots, b_1\}$, if $D_k = D_{\boldsymbol{v}}$ for some $\boldsymbol{v} \in \Omega_{L-\xi_{r-1}+1}^{\eta-(r-1)}$, let $c(D_k) = c_{\eta-(r-1)}(\boldsymbol{v})$. Then

$$\sum_{k=1}^{b_1} c(D_k) = f_\eta(\boldsymbol{\lambda}). \tag{190}$$

For $\alpha \in \{1, 2, \cdots, r-1\}$, let $A_\alpha = \{i_1, i_2, \cdots, i_{\alpha-1}\}$, where $i_j \in \{\xi_{j-1}, \xi_{j-1}+1, \cdots, \xi_j\}$ for $j \in \{1, 2, \cdots, \alpha-1\}$. Let $\mathcal{A}^{(\alpha)}$ be the collection of all A_α. For $A_\alpha \in \mathcal{A}^{(\alpha)}$, for notational simplicity, let

$$H_{A_\alpha} = \min_{j=1,2,\cdots,\alpha-1} \left\{ \sum_{k=\xi_{j-1}}^{i_j} \gamma_k^{(j)} \right\} \tag{191}$$

and

$$Q_{A_\alpha} = \max_{j=1,2,\cdots,\alpha-1} \left\{ \sum_{k=\xi_{j-1}}^{i_j-1} \gamma_k^{(j)} \right\}. \tag{192}$$

For each $i_\alpha \in \{\xi_{\alpha-1}, \xi_{\alpha-1}+1, \cdots, \xi_\alpha\}$, let

$$h_\alpha = \sum_{k=\xi_{j-1}}^{i_\alpha} \gamma_k^{(\alpha)} \tag{193}$$

and

$$q_\alpha = \sum_{k=\xi_{j-1}}^{i_\alpha-1} \gamma_k^{(\alpha)}. \tag{194}$$

Then for $\alpha \in \{1, 2, \cdots, r-1\}$, $i_\alpha \in \{\xi_{\alpha-1}, \xi_{\alpha-1}+1, \cdots, \xi_\alpha\}$, and $A_\alpha \in \mathcal{A}^{(\alpha)}$, define $\gamma_{i_\alpha}^{A_\alpha}$ by

$$\gamma_{i_\alpha}^{A_\alpha} \triangleq \left[\min\{h_\alpha, H_{A_\alpha}\} - \max\{q_\alpha, Q_{A_\alpha}\}\right]^+, \tag{195}$$

where for any $x \in \mathbb{R}$, $[x]^+ \triangleq \max\{0, x\}$ as defined after (37). For notational simplicity, we denote $\gamma_{i_\alpha}^{(\alpha)}$ and $\gamma_{i_\alpha}^{A_\alpha}$ by $\gamma_i^{(\alpha)}$ and $\gamma_i^{A_\alpha}$ respectively, where $i \in \{\xi_{\alpha-1}, \xi_{\alpha-1}+1, \cdots, \xi_\alpha\}$.

Let $\mathcal{A}_0^{(\alpha)}$ be the collection of A_α such that $\gamma_i^{A_\alpha} > 0$. We can verify that for any $i \in \{\xi_{\alpha-1}, \xi_{\alpha-1}+1, \cdots, \xi_\alpha\}$,

$$\sum_{A_\alpha \in \mathcal{A}_0^{(\alpha)}} \gamma_i^{A_\alpha} = \gamma_i^{(\alpha)}. \tag{196}$$

This means that $\gamma_i^{A_\alpha},\ A_\alpha \in \mathcal{A}^{(\alpha)}$ is a partition of $\gamma_i^{(\alpha)}$. This partition is the key idea of the converse proof in (225)-(227) that we recursively partition the coefficient of an entropy term into coefficients of entropies in a lower layer. For example, the coefficient of $H(W_1, W_2, W_3|W_4)$ is partitioned into coefficients of $H(W_1, W_2|W_3, W_4)$, $H(W_1, W_3|W_2, W_4)$, and $H(W_2, W_3|W_1, W_4)$.

For $\alpha = 1$, we can see that $\mathcal{A}_0^{(1)} = \{\emptyset\}$ and for $i \in \{1, 2, \cdots, \xi_1\}$,

$$\gamma_i^{A_1} = \gamma_i^{(1)}. \tag{197}$$

If there is an A_α such that $i \in A_\alpha$, then $i = \xi_{\alpha-1}$. In particular, for all A_α such that $\xi_{\alpha-1} \in A_\alpha$, we have $\gamma_{\xi_{\alpha-1}}^{A_\alpha} = 0$ since

$$\begin{aligned} \gamma_{\xi_\alpha}^{(\alpha+1)} &= \lambda_{\xi_\alpha} - \gamma_{\xi_\alpha}^{(\alpha)} \\ &\leq f_\eta(\boldsymbol{\lambda}) - \gamma_{\xi_\alpha}^{(\alpha)} \\ &= \sum_{i=\xi_{\alpha-1}}^{\xi_\alpha-1} \gamma_i^{(\alpha)}, \end{aligned} \tag{198}$$

where the inequality follows from Lemma 8. It is easy to check that for $i \in \{1, 2, \cdots, \xi_{\alpha-1} - 1\}$,

$$\sum_{k=\xi_{\alpha-1}}^{\xi_\alpha} \sum_{A_\alpha \in \mathcal{A}_0^{(\alpha)}:\ i \in A_\alpha} \gamma_k^{A_\alpha} = \lambda_i \tag{199}$$

and for $i = \xi_{\alpha-1}$,

$$\sum_{k=\xi_{\alpha-1}}^{\xi_\alpha} \sum_{A_\alpha \in \mathcal{A}_0^{(\alpha)}:\ \xi_{\alpha-1} \in A_\alpha} \gamma_k^{A_\alpha} = \gamma_{\xi_{\alpha-1}}^{(\alpha-1)}. \tag{200}$$

Thus,

$$\begin{aligned} &\sum_{k=\xi_{\alpha-1}}^{\xi_\alpha} \sum_{A_\alpha \in \mathcal{A}_0^{(\alpha)}:\ \xi_{\alpha-1} \in A_\alpha} \gamma_k^{A_\alpha} + \sum_{A_\alpha \in \mathcal{A}_0^{(\alpha)}} \gamma_{\xi_{\alpha-1}}^{A_\alpha} \\ &= \gamma_{\xi_{\alpha-1}}^{(\alpha-1)} + \gamma_{\xi_{\alpha-1}}^{(\alpha)} \\ &= \lambda_{\xi_{\alpha-1}}, \end{aligned} \tag{201}$$

where the first equality follows from (200) and (196), and the second equality follows from (187).

For any $k \in \{1, 2, \cdots, \alpha - 1\}$ and $A_\alpha \in \mathcal{A}_0^{(\alpha)}$, let $A_\alpha^k = \{i_1, i_2, \cdots, i_k\}$ be the set of the first k smallest elements in A_α. In particular, $A_\alpha^{\alpha-1} = A_\alpha$. Then the condition $\gamma_i^{A_\alpha} = \sum_{k=\xi_\alpha}^{\xi_{\alpha+1}} \gamma_k^{\{i\} \cup A_\alpha}$ implies that

$$\gamma_i^{A_\alpha} = \sum_{A_{\alpha+1}^{\alpha-1} \in \mathcal{A}_0^{(\alpha+1)}:\ A_{\alpha+1}^{\alpha-1} = A_\alpha} \gamma_j^{A_{\alpha+1}}. \tag{202}$$

For $i \in \mathcal{L}$ and $\alpha \in \{1, 2, \cdots, r-1\}$, we have

$$\lambda_i = \lambda_i^{(\alpha)} + \sum_{k=1}^{\alpha} \gamma_i^{(k)} = \lambda_i^{(\alpha)} + \sum_{k=1}^{\alpha} \sum_{A_k \in \mathcal{A}_0^{(k)}} \gamma_i^{A_k}. \tag{203}$$

In particular, for $\alpha = r - 1$,

$$\lambda_i = \lambda_i^{(r-1)} + \sum_{k=1}^{r-1} \sum_{A_k \in \mathcal{A}_0^{(k)}} \gamma_i^{A_k}. \tag{204}$$

Let $\mathcal{A}^{(r)} = \Big\{\{i\} \cup A_{r-1} : \gamma_i^{A_{r-1}} > 0 \text{ for } i \in \{\xi_{r-2}, \xi_{r-2} + 1, \cdots, \xi_{r-1}\} \text{ and } A_{r-1} \in \mathcal{A}_0^{(r-1)}\Big\}$. Denote the cardinality of $\mathcal{A}^{(r)}$ by b_2. For simplicity, let $\mathcal{A}^{(r)} = \{B_1, B_2, \cdots, B_{b_2}\}$. For $j \in \{1, 2, \cdots, b_2\}$, (198) implies that

$$|B_j| = r - 1. \tag{205}$$

Without loss of generality, let $B_j = \{i\} \cup A_{r-1}$ for some $i \in \{\xi_{r-2}, \xi_{r-2} + 1, \cdots, \xi_{r-1}\}$ and $A_{r-1} \in \mathcal{A}_0^{(r-1)}$. For $k \in \{1, 2, \cdots, r-1\}$, let

$$B_j^k = \begin{cases} A_{r-1}^k, & \text{if } 1 \le k \le r-2 \\ B_j, & \text{if } k = r-1. \end{cases} \tag{206}$$

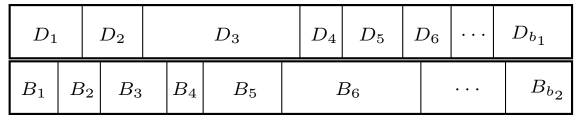

Fig. 5: a one-to-one mapping

Note that B_j^k is the set of the first k smallest elements in B_j. Let $\gamma(B_j) = \gamma_i^{A_{r-1}}$ which is the number of B_j. Then we have

$$\begin{aligned} \sum_{j=1}^{b_2} \gamma(B_j) &= \sum_{i=\xi_{r-2}}^{\xi_{r-1}} \sum_{A_{r-1} \in \mathcal{A}_0^{(r-1)}} \gamma_i^{A_{r-1}} \\ &= \sum_{i=\xi_{r-2}}^{\xi_{r-1}} \gamma_i^{(r-1)} \qquad (207) \\ &= f_\eta(\boldsymbol{\lambda}) \qquad (208) \\ &= \sum_{k=1}^{b_1} c(D_k), \qquad (209) \end{aligned}$$

where (207) follows from (196), (208) follows from (186), and (209) follows from (190). This implies that we have a one-to-one correspondence between $f_\eta(\boldsymbol{\lambda})$ of D_k's and $f_\eta(\boldsymbol{\lambda})$ of B_j's. The mapping defined by overlap in Fig. 5 is a simple one-to-one correspondence. The inequality in (198) ensures that the number of D_k's that contains ξ_{r-1} is less than or equal to the number of B_j's that don't contain ξ_{r-1}. Thus, there exists a correspondence such that $B_j \cap D_k = \emptyset$ if B_j and D_k have overlap in Fig. 5. Without loss of generality, assume the the mapping in Fig. 5 is such a correspondence. Let

$$\mathcal{O} = \big\{(j,k) : B_j \text{ and } D_k \text{ have overlap in Fig. 5}\big\}. \tag{210}$$

Then we have for all $(j,k) \in \mathcal{O}$ that

$$B_j \cap D_k = \emptyset \tag{211}$$

and

$$|B_j \cup D_k| = \eta. \tag{212}$$

For $k \in \{1, 2, \cdots, b_1\}$, let $s_k = \sum_{i=1}^k c(D_i)$. For $j \in \{1, 2, \cdots, b_2\}$, let $t_j = \sum_{i=1}^j \gamma(B_i)$. For $(j,k) \in \mathcal{O}$, let $c(B_j, D_k)$ be the length overlap of B_j and D_k in Fig. 5, which is equal to

$$c(B_j, D_k) = \begin{cases} \gamma(B_j), & \text{if } s_{k-1} \le t_{j-1} \le t_j \le s_k \\ s_k - t_{j-1}, & \text{if } s_{k-1} \le t_{j-1} \le s_k \le t_j \\ c(D_k), & \text{if } t_{j-1} \le s_{k-1} \le s_k \le t_j \\ t_j - s_{k-1}, & \text{if } t_{j-1} \le s_{k-1} \le t_j \le s_k \\ 0, & \text{otherwise.} \end{cases} \tag{213}$$

It is easy to check that for $k \in \{1, 2, \cdots, b_1\}$,

$$\sum_{j=1}^{b_2} c(B_j, D_k) = c(D_k) \tag{214}$$

and for $j \in \{1, 2, \cdots, b_2\}$,

$$\sum_{k=1}^{b_1} c(B_j, D_k) = \gamma(B_j). \tag{215}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2020.3004482, IEEE Transactions on Information Theory

Then we have

$$\sum_{(j,k)\in\mathcal{O}} c(B_j, D_k) = \sum_{k=1}^{b_1} c(D_k) = \sum_{j=1}^{b_2} \gamma(B_j) = f_\eta(\boldsymbol{\lambda}). \tag{216}$$

The following lemma states the relation between the coefficients $c(B_j, D_k)$ and $\boldsymbol{\lambda}$. The detailed proof of the lemma can be found in Appendix I.

Lemma 10. *For $i \in \mathcal{L}$, we have*

$$\sum_{(j,k)\in\mathcal{O}:\ i\in B_j\cup D_k} c(B_j, D_k) \le \lambda_i. \tag{217}$$

For any $\boldsymbol{v} \in \Omega_L^\eta$ and $\boldsymbol{v} = (v_1, v_2, \cdots, v_L)$, if $\{i : v_i = 1\} = B_j \cup D_k$ for some $(j,k) \in \mathcal{O}$, let $c_\eta(\boldsymbol{v}) = c(B_j, D_k)$. Otherwise, if there is no $(j,k) \in \mathcal{O}$ such that $\{i : v_i = 1\} = B_j \cup D_k$, let $c_\eta(\boldsymbol{v}) = 0$. Then by (212), (216), and Lemma 10, we can see that $\{c_\eta(\boldsymbol{v}) :\ \boldsymbol{v} \in \Omega_L^\eta\}$ is an optimal η-resolution for $\boldsymbol{\lambda}$.

For $i \in \{\xi_{r-1}, \xi_{r-1}+1, \cdots, L\}$ and $j \in \{1, 2, \cdots, b_2\}$, let

$$c(\{i\} \cup B_j) = \sum_{k\in\{1,2,\cdots,b_1\}:\ i\in D_k} c(B_j, D_k). \tag{218}$$

It is easy to check that

$$\sum_{j=1}^{b_2} c(\{i\} \cup B_j) = \sum_{(j,k)\in\mathcal{O}:\ i\in D_k} c(B_j, D_k) \le \lambda_i^{(r-1)} \tag{219}$$

and

$$\begin{aligned}\sum_{i=\xi_{r-1}}^{L} c(\{i\} \cup B_j) &= \sum_{i=\xi_{r-1}}^{L} \sum_{k\in\{1,2,\cdots,b_1\}:\ i\in D_k} c(B_j, D_k) \\ &= \sum_{k=1}^{b_1} c(B_j, D_k) &(220)\\ &= \gamma(B_j). &(221)\end{aligned}$$

Remark 12. In the example at the end of Section VI, the parameter $c(B_j, D_k)$ is the coefficients in (107)-(109), where for example, the coefficient $\frac{1}{3}$ of $\frac{1}{3}H(W_2W_3|W_1M_1M_2)$ in (107) and $\frac{1}{3}H(W_1W_2W_3|M_1M_2)$ in (108)-(109) is $c(\{1\}, \{2,3\})$. The fact that $\{c(B_j, D_k) : (j,k) \in \mathcal{O}\}$ is an optimal η-resolution ensures us to proceed after the η-th iteration in the converse proof. The parameter $c(\{i\} \cup B_j)$ is the coefficient in (113)-(115), where for example, $\frac{1}{3}$ of $\frac{1}{3}H(W_2|W_1M_{1:2})$ in (115) is $c(\{i\} \cup B_j)$ for $i = 2$ and $B_j = \{1\}$;

Before proving the converse, we introduce two important relations that will be repeated used in the proof. For $\alpha = 1, 2, \cdots, r-1$, and $i, j \in \mathcal{L}$, $\mathcal{B} \subseteq \mathcal{L}$ such that $|\mathcal{B}| = \alpha - 1$ and $i, j \notin \mathcal{B}$, we have

$$\begin{aligned}H(W_i|W_{\mathcal{B}}M_{1:\alpha}) &\ge H(W_i|W_jW_{\mathcal{B}}M_{1:\alpha}) \\ &= H(W_i|W_{\{j\}\cup\mathcal{B}}M_{1:\alpha}). &(222)\end{aligned}$$

and

$$\begin{aligned}&H(W_i|W_{\mathcal{B}}M_{1:\alpha-1}) \\ &= H(W_i|W_{\mathcal{B}}M_{1:\alpha-1}M_\alpha) + H(M_\alpha|W_{\mathcal{B}}M_{1:\alpha-1}) \\ &\quad - H(M_\alpha|W_iW_{\mathcal{B}}M_{1:\alpha-1}) \\ &= H(W_i|W_{\mathcal{B}}M_{1:\alpha}) + H(M_\alpha) &(223)\end{aligned}$$

For notational simplicity, let $\xi_{-1} = 0$. For $\alpha = 1, 2, \cdots, r-1$, let

$$\begin{aligned}I_\alpha \triangleq &\sum_{i=\xi_{\alpha-2}}^{\xi_{\alpha-1}} \sum_{A_{\alpha-1}\in\mathcal{A}_0^{(\alpha-1)}} \gamma_i^{A_{\alpha-1}} H(W_iW_{A_{\alpha-1}}|M_{1:\alpha}) \\ &+ \sum_{i=\xi_{r-1}}^{L} \sum_{j=1}^{b} c(\{i\} \cup B_j) H(W_i|W_{B_j^{\alpha-1}}M_{1:\alpha}) \\ &+ \sum_{i=1}^{L} \left(\sum_{k=\alpha}^{r-1} \sum_{A_k\in\mathcal{A}_0^{(k)}} \gamma_i^{A_k} H(W_i|W_{A_k^{\alpha-1}}M_{1:\alpha}) \right). \end{aligned} \tag{224}$$

We have the following lemma which provides an iteration that is useful in the sequel. The proof of the lemma can be found in Appendix J.

Lemma 11. $I_\alpha \ge I_{\alpha+1} + [f_1(\boldsymbol{\lambda}) - \alpha f_\eta(\boldsymbol{\lambda})]\, H(M_{\alpha+1})$ *for* $\alpha = 1, 2, \cdots, r-1$.

We prove the converse of DS-SMDC (i.e., $\boldsymbol{\lambda} \cdot \mathbf{R} \ge g_\eta(\boldsymbol{\lambda})$ for all $\eta = r+1, r+2, \cdots, L+1$) as follows.

$$\begin{aligned}\boldsymbol{\lambda} \cdot \boldsymbol{R} &= \lambda_1 H(W_1) + \lambda_2 H(W_2) + \cdots + \lambda_L H(W_L) \\ &= (\sum_{i=1}^{L} \lambda_i) H(M_1) + \sum_{i=1}^{L} \lambda_i H(W_i|M_1) &(225)\\ &= f_1(\boldsymbol{\lambda}) H(M_1) \\ &\quad + \sum_{i=1}^{L} \left(\lambda_i^{(r-1)} + \sum_{k=1}^{r-1} \sum_{A_k\in\mathcal{A}_0^{(k)}} \gamma_i^{A_k} \right) H(W_i|M_1) &(226)\\ &\ge \sum_{\alpha=1}^{r} [f_1(\boldsymbol{\lambda}) - (\alpha-1) f_\eta(\boldsymbol{\lambda})]\, H(M_\alpha) \\ &\quad + \sum_{i=\xi_{r-2}}^{\xi_{r-1}} \sum_{A_{r-1}\in\mathcal{A}_0^{(r-1)}} \gamma_i^{A_{r-1}} H(W_iW_{A_{r-1}}|M_{1:r}) \\ &\quad + \sum_{i=\xi_{r-1}}^{L} \sum_{j=1}^{b_2} c(\{i\} \cup B_j) H(W_i|W_{B_j^{r-1}}M_{1:r}) &(227)\\ &= \sum_{\alpha=1}^{r} [f_1(\boldsymbol{\lambda}) - (\alpha-1) f_\eta(\boldsymbol{\lambda})]\, H(M_\alpha) \\ &\quad + \sum_{j=1}^{b_2} \gamma(B_j) H(W_{B_j}|M_{1:r}) \\ &\quad + \sum_{i=\xi_{r-1}}^{L} \sum_{j=1}^{b_2} c(\{i\} \cup B_j) H(W_i|W_{B_j}M_{1:r}) &(228)\\ &\ge \sum_{\alpha=1}^{r} [f_1(\boldsymbol{\lambda}) - (\alpha-1) f_\eta(\boldsymbol{\lambda})]\, H(M_\alpha) \\ &\quad + \sum_{(j,k)\in\mathcal{O}} c(B_j, D_k) H(W_{B_j}|M_{1:r}) \\ &\quad + \sum_{(j,k)\in\mathcal{O}} c(B_j, D_k) H(W_{D_k}|W_{B_j}M_{1:r}) &(229)\end{aligned}$$

$$\geq \sum_{\alpha=1}^{r} [f_1(\boldsymbol{\lambda}) - (\alpha-1)f_\eta(\boldsymbol{\lambda})] H(M_\alpha) + \sum_{(j,k)\in\mathcal{O}} c(B_j, D_k) H(W_{D_k} W_{B_j} | M_{1:r}) \tag{230}$$

$$= \sum_{\alpha=1}^{r} [f_1(\boldsymbol{\lambda}) - (\alpha-1)f_\eta(\boldsymbol{\lambda})] H(M_\alpha) + \sum_{(j,k)\in\mathcal{O}} c(B_j, D_k) H(W_{B_j\cup D_k} M_{r+1}^{\eta} | M_{1:r}) \tag{231}$$

$$\geq \sum_{\alpha=1}^{r} [f_1(\boldsymbol{\lambda}) - (\alpha-1)f_\eta(\boldsymbol{\lambda})] H(M_\alpha) + f_\eta(\boldsymbol{\lambda}) \sum_{\alpha=r+1}^{\eta} H(M_\alpha) + \sum_{\boldsymbol{v}\in\Omega_L^\eta} c_\eta(\boldsymbol{v}) H(W_{\boldsymbol{v}} | M_{1:\eta}) \tag{232}$$

$$\geq \sum_{\alpha=1}^{r} [f_1(\boldsymbol{\lambda}) - (\alpha-1)f_\eta(\boldsymbol{\lambda})] H(M_\alpha) + f_\eta(\boldsymbol{\lambda}) \sum_{\alpha=r+1}^{\eta} H(M_\alpha) + \sum_{\alpha=\eta+1}^{L} f_\alpha(\boldsymbol{\lambda}) H(M_\alpha) \tag{233}$$

$$= \sum_{\alpha=1}^{r} (f_1(\boldsymbol{\lambda})) H(M_\alpha) + \sum_{\alpha=\eta+1}^{L} f_\alpha(\boldsymbol{\lambda}) H(M_\alpha) + f_\eta(\boldsymbol{\lambda}) \left[\sum_{\alpha=r+1}^{\eta} H(M_\alpha) - \sum_{\alpha=1}^{r} (\alpha-1) H(M_\alpha) \right]$$

$$= \sum_{\alpha=1}^{r} f_1(\boldsymbol{\lambda}) m_\alpha + \sum_{\alpha=\eta+1}^{L} f_\alpha(\boldsymbol{\lambda}) m_\alpha + f_\eta(\boldsymbol{\lambda}) \left[\sum_{\alpha=r+1}^{\eta} m_\alpha - \sum_{\alpha=1}^{r} (\alpha-1) m_\alpha \right], \tag{234}$$

where (226) follows from (204), (227) follows by applying Lemma 11 for $\alpha = 1, 2, \cdots, r-1$ successively, (228) follows from the definition of B_j, (232) follows from (216), (233) follows from the fact that $\{c_\eta(\boldsymbol{v}) : \boldsymbol{v} \in \Omega_L^\eta\}$ is an optimal η-resolution for $\boldsymbol{\lambda}$ and the iteration in the converse for SMDC in [3], and (229) follows from (215) and

$$\sum_{i=\xi_{r-1}}^{L} \sum_{j=1}^{b_2} c(\{i\} \cup B_j) H(W_i | W_{B_j} M_{1:r})$$

$$= \sum_{j=1}^{b_2} \left[\sum_{i=\xi_{r-1}}^{L} c(\{i\} \cup B_j) H(W_i | W_{B_j} M_{1:r}) \right] \tag{235}$$

$$= \sum_{j=1}^{b_2} \left[\sum_{i=\xi_{r-1}}^{L} \sum_{k\in\{1,2,\cdots,b_1\}:\ i\in D_k} c(B_j, D_k) H(W_i | W_{B_j} M_{1:r}) \right] \tag{236}$$

$$= \sum_{j=1}^{b_2} \left[\sum_{k=1}^{b_1} c(B_j, D_k) \left(\sum_{i\in D_k} H(W_i | W_{B_j} M_{1:r}) \right) \right] \tag{237}$$

$$\geq \sum_{j=1}^{b_2} \sum_{k=1}^{b_1} c(B_j, D_k) H(W_{D_k} | W_{B_j} M_{1:r}) \tag{238}$$

$$= \sum_{(j,k)\in\mathcal{O}} c(B_j, D_k) H(W_{D_k} | W_{B_j} M_{1:r}). \tag{239}$$

Dividing both sides of (234) by a, we obtain by the definition of $g_\eta(\boldsymbol{\lambda})$ in (93) that for any $\eta = r+1, r+2, \cdots, L+1$,

$$\sum_{l=1}^{L} \lambda_l (\mathsf{R}_l + \epsilon) \geq g_\eta(\boldsymbol{\lambda}). \tag{240}$$

Letting $\epsilon \to 0$, the inequality $\boldsymbol{\lambda} \cdot \mathbf{R} \geq g_\eta(\boldsymbol{\lambda})$ is proved.

Remark 13. The step-by-step correspondence between the general proof in (225)-(234) and the example in (105)-(110) is as follows:

- The iteration in (227) is the generalization of the step in (107);
- The transform of conditional entropies in (228)-(232) play the same role as (108);
- The application of the α-resolution technique in (233) is the generalization of that in (109).

Appendix F
Proof of Lemma 7

Let $\boldsymbol{\lambda}' = (\lambda_1', \lambda_2', \cdots, \lambda_L')$, where

$$\lambda_i' = \lambda_i, \text{ for all } i = 2, 3, \cdots, L \tag{241}$$

and

$$\lambda_1' = \frac{\lambda_2' + \lambda_3' + \cdots + \lambda_L'}{\eta - 1}. \tag{242}$$

By Lemma 7 in [4], (175) implies that

$$f_\eta(\boldsymbol{\lambda}) = f_{\eta-1}(\lambda_2, \lambda_3, \cdots, \lambda_L), \tag{243}$$

and similarly, from (242),

$$f_\eta(\boldsymbol{\lambda}') = f_{\eta-1}(\lambda_2, \lambda_3, \cdots, \lambda_L). \tag{244}$$

Thus, we have

$$f_\eta(\boldsymbol{\lambda}) = f_\eta(\boldsymbol{\lambda}'), \tag{245}$$

which by Lemma 5 in [4] implies that

$$f_\alpha(\boldsymbol{\lambda}) = f_\alpha(\boldsymbol{\lambda}'), \text{ for all } \eta \leq \alpha \leq L. \tag{246}$$

The rate constraint $\boldsymbol{\lambda}' \cdot \mathbf{R} \geq g_\eta(\boldsymbol{\lambda}')$ is the following,

$$\boldsymbol{\lambda}' \cdot \mathbf{R} \geq \sum_{\alpha=1}^{r} f_1(\boldsymbol{\lambda}') \mathsf{m}_\alpha + \sum_{\alpha=\eta+1}^{L} f_\alpha(\boldsymbol{\lambda}') \mathsf{m}_\alpha + f_\eta(\boldsymbol{\lambda}') \left[\sum_{\alpha=r+1}^{\eta} \mathsf{m}_\alpha - \sum_{\alpha=1}^{r} (\alpha-1) \mathsf{m}_\alpha \right]. \tag{247}$$

This implies

$$\begin{aligned}
\boldsymbol{\lambda}\cdot\mathbf{R} &= \boldsymbol{\lambda}'\cdot\mathbf{R}+(\lambda_1-\lambda_1')\mathsf{R}_1\\
&\geq \sum_{\alpha=1}^{r} f_1(\boldsymbol{\lambda}')\mathsf{m}_\alpha+\sum_{\alpha=\eta+1}^{L} f_\alpha(\boldsymbol{\lambda}')\mathsf{m}_\alpha\\
&\quad + f_\eta(\boldsymbol{\lambda}')\left[\sum_{\alpha=r+1}^{\eta}\mathsf{m}_\alpha-\sum_{\alpha=1}^{r}(\alpha-1)\mathsf{m}_\alpha\right]\\
&\quad +(\lambda_1-\lambda_1')\left(\sum_{\alpha=1}^{r}\mathsf{m}_\alpha\right) \qquad (248)\\
&= \sum_{\alpha=1}^{r} f_1(\boldsymbol{\lambda})\mathsf{m}_\alpha+\sum_{\alpha=\eta+1}^{L} f_\alpha(\boldsymbol{\lambda})\mathsf{m}_\alpha\\
&\quad + f_\eta(\boldsymbol{\lambda})\left[\sum_{\alpha=r+1}^{\eta}\mathsf{m}_\alpha-\sum_{\alpha=1}^{r}(\alpha-1)\mathsf{m}_\alpha\right], \qquad (249)
\end{aligned}$$

which is exactly the constraint $\boldsymbol{\lambda}\cdot\mathbf{R}\geq g_\eta(\boldsymbol{\lambda})$. This proves the lemma.

Appendix G
Proof of Lemma 8

For any $\boldsymbol{\lambda}\in\mathbb{R}^L_\eta$, we have

$$\lambda_1 \leq \frac{1}{\eta-1}\sum_{i=2}^{L}\lambda_i, \qquad (250)$$

which by Lemma 1 in [4] implies that

$$f_\eta(\boldsymbol{\lambda})=\frac{1}{\eta}\sum_{i=1}^{L}\lambda_i. \qquad (251)$$

It is easy to check that (250) is equivalent to

$$\lambda_1\leq\frac{1}{\eta}\sum_{i=1}^{L}\lambda_i. \qquad (252)$$

Thus, we have $f_\eta(\boldsymbol{\lambda})\geq\lambda_1$, which proves the lemma.

Appendix H
Proof of Lemma 9

For $\alpha=1,2,\cdots,r-1$, we have

$$\sum_{i=\xi_\alpha}^{L}\lambda_i^{(\alpha)}=\sum_{i=1}^{L}\lambda_i-\alpha f_\eta(\boldsymbol{\lambda})\geq(\eta-\alpha)f_\eta(\boldsymbol{\lambda}). \qquad (253)$$

where the inequality follows from (178) and the fact that $f_1(\boldsymbol{\lambda})=\sum_{i=1}^{L}\lambda_i$. In particular, for $\alpha=r-1$,

$$\sum_{i=\xi_{r-1}}^{L}\lambda_i^{(r-1)}\geq[\eta-(r-1)]\,f_\eta(\boldsymbol{\lambda}). \qquad (254)$$

Denote the ordered permutation of $\boldsymbol{\lambda}^{(r-1)}$ by $\tilde{\boldsymbol{\lambda}}^{(r-1)}=\left(\tilde{\lambda}^{(r-1)}_{\xi_{r-1}},\tilde{\lambda}^{(r-1)}_{\xi_{r-1}+1},\cdots,\tilde{\lambda}^{(r-1)}_{L}\right)$. Then from (181), we obtain

$$\frac{1}{\eta-(r-1)}\sum_{i=\xi_{r-1}}^{L}\tilde{\lambda}_i^{(r-1)}\geq f_\eta(\boldsymbol{\lambda})\geq\tilde{\lambda}^{(r-1)}_{\xi_{r-1}}, \qquad (255)$$

which implies

$$\tilde{\lambda}^{(r-1)}_{\xi_{r-1}}\leq\frac{1}{[\eta-(r-1)]-1}\sum_{i=\xi_{r-1}+1}^{L}\tilde{\lambda}_i^{(r-1)}. \qquad (256)$$

By Lemma 4 and Lemma 7 in [3], this implies that $\tilde{\boldsymbol{\lambda}}^{(r-1)}$ has a perfect $[\eta-(r-1)]$-resolution (c.f. Appendix A) and

$$f_{\eta-(r-1)}\left(\tilde{\boldsymbol{\lambda}}^{(r-1)}\right)=\frac{1}{\eta-(r-1)}\sum_{i=\xi_{r-1}}^{L}\tilde{\lambda}_i^{(r-1)}. \qquad (257)$$

From Lemma 2 in [4] and (255), this implies that

$$f_{\eta-(r-1)}\left(\boldsymbol{\lambda}^{(r-1)}\right)=f_{\eta-(r-1)}\left(\tilde{\boldsymbol{\lambda}}^{(r-1)}\right)\geq f_\eta(\boldsymbol{\lambda}). \qquad (258)$$

This proves the lemma.

Appendix I
Proof of Lemma 10

Consider the following five cases where the set $\mathcal{L}$ is partitioned into five subsets.

i. For $i\in\{1,2,\cdots,\xi_{r-2}-1\}$, we have

$$\begin{aligned}
\sum_{(j,k)\in\mathcal{O}:\ i\in B_j} c(B_j,D_k) &= \sum_{i\in B_j}\sum_{k=1}^{b_1} c(B_j,D_k)\\
&=\sum_{i\in B_j}\gamma(B_j) \qquad (259)\\
&=\sum_{k=\xi_{\alpha-1}}^{\xi_\alpha}\ \sum_{A_\alpha\in\mathcal{A}_0^{(\alpha)}:\ i\in A_\alpha}\gamma_k^{A_\alpha}\\
&=\lambda_i, \qquad (260)
\end{aligned}$$

where (259) follows from (215) and (260) follows from (199).

ii. For $i=\xi_{r-2}$, it follows from (201) that

$$\begin{aligned}
&\sum_{(j,k)\in\mathcal{O}:\ i\in B_j} c(B_j,D_k)\\
&=\sum_{k=\xi_{r-2}}^{\xi_{r-1}}\ \sum_{A_{r-1}\in\mathcal{A}_0^{(r-1)}:\ \xi_{r-2}\in A_{r-1}}\gamma_k^{A_{r-1}}\\
&\quad+\sum_{A_{r-1}\in\mathcal{A}_0^{(r-1)}}\gamma^{A_{r-1}}_{\xi_{r-2}}\\
&=\gamma^{(r-2)}_{\xi_{r-2}}+\gamma^{(r-1)}_{\xi_{r-2}}\\
&=\lambda_{\xi_{r-2}}. \qquad (261)
\end{aligned}$$

iii. For $i\in\{\xi_{r-2}+1,\xi_{r-2}+2,\cdots,\xi_{r-1}-1\}$, we have

$$\begin{aligned}
\sum_{(j,k)\in\mathcal{O}:\ i\in B_j} c(B_j,D_k) &= \sum_{A_{r-1}\in\mathcal{A}_0^{(r-1)}}\gamma_i^{A_{r-1}}\\
&=\gamma_i^{(r-1)} \qquad (262)\\
&=\lambda_i, \qquad (263)
\end{aligned}$$

where (262) follows from (196) and (263) follows from (184).

iv. For $i=\xi_{r-1}$,

$$\begin{aligned}
&\sum_{(j,k)\in\mathcal{O}:\ i\in B_j} c(B_j,D_k)\\
&=\sum_{A_{r-1}\in\mathcal{A}_0^{(r-1)}}\gamma_{\xi_{r-1}}^{A_{r-1}}+\sum_{\xi_{r-1}\in D_k}\sum_{j=1}^{b_2}c(B_j,D_k)\\
&=\sum_{A_{r-1}\in\mathcal{A}_0^{(r-1)}}\gamma_{\xi_{r-1}}^{A_{r-1}}+\sum_{\xi_{r-1}\in D_k}c(D_k) &(264)\\
&\le\gamma_{\xi_{r-1}}^{(r-1)}+\gamma_{\xi_{r-1}}^{(r)} &(265)\\
&=\lambda_{\xi_{r-1}}, &(266)
\end{aligned}$$

where (264) follows from (214), (265) follows from (196) and the fact that $\{c(D_k):k=1,2,\cdots,b_1\}$ is an $[\eta-(r-1)]$-resolution for $\boldsymbol{\lambda}^{(r-1)}$, and (266) follows from (184).

v. For $i\in\{\xi_{r-1}+1,\xi_{r-1}+2,\cdots,L\}$, we have

$$\begin{aligned}
\sum_{(j,k)\in\mathcal{O}:\ i\in D_k}c(B_j,D_k)&=\sum_{i\in D_k}\sum_{j=1}^{b_2}c(B_j,D_k)\\
&=\sum_{i\in D_k}c(D_k) &(267)\\
&\le\lambda_i, &(268)
\end{aligned}$$

where (267) follows from (214) and (268) follows from the fact that $\{c(D_k):k=1,2,\cdots,b_1\}$ is an $[\eta-(r-1)]$-resolution for $\boldsymbol{\lambda}^{(r-1)}$.

Appendix J
Proof of Lemma 11

For $\alpha=1,2,\cdots,r-1$, we have the following iteration,

$$\begin{aligned}
I_\alpha&=\sum_{i=\xi_{\alpha-2}}^{\xi_{\alpha-1}}\sum_{A_{\alpha-1}\in\mathcal{A}_0^{(\alpha-1)}}\gamma_i^{A_{\alpha-1}}H(W_iW_{A_{\alpha-1}}|M_{1:\alpha})\\
&\quad+\sum_{i=\xi_{r-1}}^{L}\sum_{j=1}^{b}c(\{i\}\cup B_j)H(W_i|W_{B_j^{\alpha-1}}M_{1:\alpha})\\
&\quad+\sum_{i=1}^{L}\left(\sum_{k=\alpha}^{r-1}\sum_{A_k\in\mathcal{A}_0^{(k)}}\gamma_i^{A_k}H(W_i|W_{A_k^{\alpha-1}}M_{1:\alpha})\right) &(269)\\
&=\sum_{i=\xi_{\alpha-2}}^{\xi_{\alpha-1}}\sum_{A_{\alpha-1}\in\mathcal{A}_0^{(\alpha-1)}}\gamma_i^{A_{\alpha-1}}H(W_iW_{A_{\alpha-1}}|M_{1:\alpha})\\
&\quad+\sum_{i=\xi_{\alpha-1}}^{\xi_\alpha}\sum_{A_\alpha\in\mathcal{A}_0^{(\alpha)}}\gamma_i^{A_\alpha}H(W_i|W_{A_\alpha}M_{1:\alpha})\\
&\quad+\sum_{i=\xi_{r-1}}^{L}\sum_{j=1}^{b}c(\{i\}\cup B_j)H(W_i|W_{B_j^{\alpha-1}}M_{1:\alpha})\\
&\quad+\sum_{i=1}^{L}\left(\sum_{k=\alpha+1}^{r-1}\sum_{A_k\in\mathcal{A}_0^{(k)}}\gamma_i^{A_k}H(W_i|W_{A_k^{\alpha-1}}M_{1:\alpha})\right) &(270)\\
&=\sum_{i=\xi_{\alpha-1}}^{\xi_\alpha}\sum_{A_\alpha\in\mathcal{A}_0^{(\alpha)}}\gamma_i^{A_\alpha}H(W_{A_\alpha}|M_{1:\alpha})\\
&\quad+\sum_{i=\xi_{\alpha-1}}^{\xi_\alpha}\sum_{A_\alpha\in\mathcal{A}_0^{(\alpha)}}\gamma_i^{A_\alpha}H(W_i|W_{A_\alpha}M_{1:\alpha})\\
&\quad+\sum_{i=\xi_{r-1}}^{L}\sum_{j=1}^{b}c(\{i\}\cup B_j)H(W_i|W_{B_j^{\alpha-1}}M_{1:\alpha})\\
&\quad+\sum_{i=1}^{L}\left(\sum_{k=\alpha+1}^{r-1}\sum_{A_k\in\mathcal{A}_0^{(k)}}\gamma_i^{A_k}H(W_i|W_{A_k^{\alpha-1}}M_{1:\alpha})\right) &(271)\\
&=\sum_{i=\xi_{\alpha-1}}^{\xi_\alpha}\sum_{A_\alpha\in\mathcal{A}_0^{(\alpha)}}\gamma_i^{A_\alpha}H(W_iW_{A_\alpha}|M_{1:\alpha})\\
&\quad+\sum_{i=\xi_{r-1}}^{L}\sum_{j=1}^{b}c(\{i\}\cup B_j)H(W_i|W_{B_j^{\alpha-1}}M_{1:\alpha})\\
&\quad+\sum_{i=1}^{L}\left(\sum_{k=\alpha+1}^{r-1}\sum_{A_k\in\mathcal{A}_0^{(k)}}\gamma_i^{A_k}H(W_i|W_{A_k^{\alpha-1}}M_{1:\alpha})\right) &(272)\\
&\ge\sum_{i=\xi_{\alpha-1}}^{\xi_\alpha}\sum_{A_\alpha\in\mathcal{A}_0^{(\alpha)}}\gamma_i^{A_\alpha}H(W_iW_{A_\alpha}|M_{1:\alpha+1})\\
&\quad+\sum_{i=\xi_{r-1}}^{L}\sum_{j=1}^{b}c(\{i\}\cup B_j)H(W_i|W_{B_j^{\alpha}}M_{1:\alpha})\\
&\quad+\sum_{i=1}^{L}\left(\sum_{k=\alpha+1}^{r-1}\sum_{A_k\in\mathcal{A}_0^{(k)}}\gamma_i^{A_k}H(W_i|W_{A_k^{\alpha}}M_{1:\alpha})\right) &(273)\\
&=\sum_{i=\xi_{\alpha-1}}^{\xi_\alpha}\sum_{A_\alpha\in\mathcal{A}_0^{(\alpha)}}\gamma_i^{A_\alpha}H(W_iW_{A_\alpha}|M_{1:\alpha+1})\\
&\quad+\left[f_1(\boldsymbol{\lambda})-\alpha f_\eta(\boldsymbol{\lambda})\right]H(M_{\alpha+1})\\
&\quad+\sum_{i=\xi_{r-1}}^{L}\sum_{j=1}^{b}c(\{i\}\cup B_j)H(W_i|W_{B_j^{\alpha}}M_{1:\alpha+1})\\
&\quad+\sum_{i=1}^{L}\left(\sum_{k=\alpha+1}^{r-1}\sum_{A_k\in\mathcal{A}_0^{(k)}}\gamma_i^{A_k}H(W_i|W_{A_k^{\alpha}}M_{1:\alpha+1})\right) &(274)\\
&=I_{\alpha+1}+\left[f_1(\boldsymbol{\lambda})-\alpha f_\eta(\boldsymbol{\lambda})\right]H(M_{\alpha+1}), &(275)
\end{aligned}$$

where (271) follows from

$$\begin{aligned}
&\sum_{i=\xi_{\alpha-2}}^{\xi_{\alpha-1}}\sum_{A_{\alpha-1}\in\mathcal{A}_0^{(\alpha-1)}}\gamma_i^{A_{\alpha-1}}H(W_iW_{A_{\alpha-1}}|M_{1:\alpha}) &(276)\\
&=\sum_{i=\xi_{\alpha-2}}^{\xi_{\alpha-1}}\sum_{A_{\alpha-1}\in\mathcal{A}_0^{(\alpha-1)}}\sum_{j=\xi_{\alpha-1}}^{\xi_\alpha}\gamma_j^{\{i\}\cup A_{\alpha-1}}H(W_{\{i\}\cup A_{\alpha-1}}|M_{1:\alpha}) &(277)
\end{aligned}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2020.3004482, IEEE Transactions on Information Theory

$$= \sum_{j=\xi_{\alpha-1}}^{\xi_\alpha} \sum_{i=\xi_{\alpha-2}}^{\xi_{\alpha-1}} \sum_{A_{\alpha-1}\in\mathcal{A}_0^{(\alpha-1)}} \gamma_j^{\{i\}\cup A_{\alpha-1}} H(W_{\{i\}\cup A_{\alpha-1}}|M_{1:\alpha}) \tag{278}$$

$$= \sum_{j=\xi_{\alpha-1}}^{\xi_\alpha} \sum_{A_\alpha\in\mathcal{A}_0^{(\alpha)}} \gamma_j^{A_\alpha} H(W_{A_\alpha}|M_{1:\alpha}), \tag{279}$$

(273) follows from the fact that conditioning does not increase entropy, and (274) follows from (223).

References

[1] J. R. Roche, R. W. Yeung, and K. P. Hau, "Symmetrical multilevel diversity coding," *IEEE Trans. Inf. Theory*, vol. 43, pp. 1059–1064, May 1997.

[2] A. Albanese, J. Blömer, J. Edmonds, M. Luby, and M. Sudan, "Priority encoding transmission," *IEEE Trans. Inf. Theory*, vol. 42, pp. 1737–1744, Nov. 1996.

[3] R. W. Yeung and Z. Zhang, "On symmetrical multilevel diversity coding," *IEEE Trans. Inf. Theory*, vol. 45, pp. 609–621, Mar. 1999.

[4] T. Guo and R. W. Yeung, "The explicit coding rate region of symmetric multilevel diversity coding," *IEEE Trans. Inf. Theory*, vol. 66, pp. 1053–1077, Feb. 2020.

[5] C. Tian and T. Liu, "Multilevel diversity coding with regeneration," *IEEE Trans. Inf. Theory*, vol. 62, pp. 4833–4847, Sep. 2016.

[6] S. Mohajer, C. Tian, and S. N. Diggavi, "Asymmetric multilevel diversity coding and asymmetric gaussian multiple descriptions," *IEEE Trans. Inf. Theory*, vol. 56, pp. 4367–4387, Sep. 2010.

[7] C. Li and J. W. S. Weber, "Multilevel diversity coding systems: Rate regions, codes, computation, & forbidden minors," *IEEE Trans. Inf. Theory*, vol. 63, pp. 230–251, Nov. 2016.

[8] A. Balasubramanian, H. D. Ly, S. Li, T. Liu, and S. L. Miller, "Secure symmetrical multilevel diversity coding," *IEEE Trans. Inf. Theory*, vol. 59, pp. 3572–3581, Jun. 2013.

[9] J. Jiang, N. Marukala, and T. Liu, "Symmetrical multilevel diversity coding and subset entropy inequalities," *IEEE Trans. Inf. Theory*, vol. 60, pp. 84–103, Jan. 2014.

[10] N. Cai and R. W. Yeung, "Secure network coding," in *IEEE International Symposium on Information Theory (ISIT)*, (Lausanne, Switzerland), Jun. 2002.

[11] T. Guo, C. Tian, T. Liu, and R. W. Yeung, "Weakly secure symmetric multilevel diversity coding," in *2019 IEEE Information Theory Workshop (ITW)*, (Visby, Gotland, Sweden), Aug. 2019.

[12] M. Yan and A. Sprintson, "Weakly secure network coding for wireless cooperative data exchange," in *IEEE Global Telecommunications Conference (GLOBECOM)*, (Kathmandu, Nepal), Dec. 2011.

[13] M. Yan and A. Sprintson, "Algorithms for weakly secure data exchange," in *Proc. NetCod 2013*, (Calgary, Alberta, Canada), Jun. 2013.

[14] M. Yan, A. Sprintson, and I. Zelenko, "Weakly secure data exchange with generalized reed solomon codes," in *IEEE International Symposium on Information Theory (ISIT)*, (Honolulu, HI, USA), Jun. 2014.

[15] Y. Chen, O. O. Koyluoglu, and A. J. H. Vinck, "On secure communication over the multiple access channel," in *International Symposium on Information Theory and Its Applications (ISITA)*, (Monterey, CA, USA), Oct. 2016.

[16] Y. Chen, O. O. Koyluoglu, and A. Sezgin, "Individual secrecy for broadcast channels with receiver side information," *IEEE Trans. Inf. Theory*, vol. 63, pp. 4687–4708, Jul. 2017.

[17] Y. Chen, O. O. Koyluoglu, and A. Sezgin, "Individual secrecy for the broadcast channel," *IEEE Trans. Inf. Theory*, vol. 63, pp. 5981–5999, Sep. 2017.

[18] A. S. Mansour, R. F. Schaefer, and H. Boche, "On the individual secrecy capacity regions of the general, degraded, and gaussian multi-receiver wiretap broadcast channel," *IEEE Trans. Inf. Forensics and Security*, vol. 11, pp. 2107–2122, Sep. 2016.

[19] E. Tekin and A. Yener, "The general gaussian multiple-access and two-way wiretap channels: Achievable rates and cooperative jamming," *IEEE Trans. Inf. Theory*, vol. 54, pp. 2735–2751, Jun. 2008.

[20] E. Tekin and A. Yener, "The gaussian multiple access wire-tap channel," *IEEE Trans. Inf. Theory*, vol. 54, pp. 5747–5755, Dec. 2008.

[21] U. Maurer and S. Wolf, "Information-theoretic key agreement: From weak to strong secrecy for free," in *Proceedings of EUROCRYPT 2000, Lecture Notes in Computer Science*, vol. 1807, pp. 351–368, Springer-Verlag, 2000.

[22] M. Nafea and A. Yener, "Generalizing multiple access wiretap and wiretap II channel models: Achievable rates and cost of strong secrecy," *IEEE Trans. Inf. Theory*, vol. 65, pp. 5125–5143, Aug. 2019.

[23] K. Bhattad and K. R. Narayanan, "Weakly secure network coding," in *Proc. NetCod 2005*, (Riva del Garda, Italy), Apr. 2005.

[24] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. Inf. Theory*, vol. 46, pp. 1204–1216, Jul. 2000.

[25] S.-Y. R. Li, R. W. Yeung, and N. Cai, "Linear network coding," *IEEE Trans. Inf. Theory*, vol. 49, pp. 371–381, Feb. 2003.

[26] N. Cai and R. W. Yeung, "Secure network coding on a wiretap network," *IEEE Trans. Inf. Theory*, vol. 57, pp. 424–435, Jan. 2011.

[27] R. W. Yeung, *Information Theory and Network Coding*. Springer, 2008.

[28] H. Yamamoto, "Secret sharing system using (k, L, n) threshold scheme," *IEICE Trans. Fund.(Jpn. Edition)*, vol. J68-A, Sep. 1985. (English Translation: Scripta Technica, Inc., Electronics and Commun. in Japan, Part I, vol. 69, pp. 4654, 1986).

[29] W.-A. Jackson and K. M. Martin, "A combinatorial interpretation of ramp schemes," *Australasian Journal of Combinatorics*, vol. 14, pp. 51–60, 1996.

Tao Guo (S'16–M'19) received his B.E. degree in Telecommunications Engineering from Xidian University in 2013, and the Ph.D. degree from the Department of Information Engineering, The Chinese University of Hong Kong in 2018. He was a Postdoctoral Research Associate at the Department of Electrical and Computer Engineering, Texas A&M University from 2018 to 2020. He is currently a Postdoctoral Scholar in the Department of Electrical and Computer Engineering at the University of California, Los Angeles. His research interests include information theory and it applications to security and privacy, multi-user source coding, and coding for distributed storage systems.

Chao Tian (S'00–M'05–SM'12) received the B.E. degree in Electronic Engineering from Tsinghua University, Beijing, China, in 2000 and the M.S. and Ph. D. degrees in Electrical and Computer Engineering from Cornell University, Ithaca, NY in 2003 and 2005, respectively. Dr. Tian was a postdoctoral researcher at Ecole Polytechnique Federale de Lausanne (EPFL) from 2005 to 2007, a member of technical staff–research at AT&T Labs–Research in New Jersey from 2007 to 2014, and an Associate Professor in the Department of Electrical Engineering and Computer Science at the University of Tennessee Knoxville from 2014 to 2017. He joined the Department of Electrical and Computer Engineering at Texas A&M University in 2017. His research interests include data storage systems, multi-user information theory, joint source-channel coding, signal processing, and compute algorithms.

Dr. Tian received the Liu Memorial Award at Cornell University in 2004, AT&T Key Contributor Award in 2010, 2011 and 2013. His authored and co-authored papers received the 2014 IEEE ComSoc DSTC Data Storage Best Paper Award and the 2017 IEEE Jack Keil Wolf ISIT Student Paper Award. He was an Associate Editor for THE IEEE SIGNAL PROCESSING LETTERS from 2012 to 2014, and is currently an Editor for THE IEEE TRANSACTIONS ON COMMUNICATIONS and an Associate Editor for THE IEEE TRANSACTIONS ON INFORMATION THEORY.

Tie Liu received his B.S. (1998) and M.S. (2000) degrees, both in Electrical Engineering, from Tsinghua University, Beijing, China and a second M.S. degree in Mathematics (2004) and a Ph.D. degree in Electrical and Computer Engineering (2006) from the University of Illinois at Urbana-Champaign. Since August 2006 he has been with Texas A&M University, where he is currently a Professor in the Department of Electrical and Computer Engineering. His primary research interest is in the area of information and statistical learning theory.

Dr. Liu received an M. E. Van Valkenburg Graduate Research Award (2006) from the University of Illinois at Urbana-Champaign, a CAREER Award (2009) from the National Science Foundation, and an Outstanding Professor Award from Texas A&M University (2018). He was a Technical Program Committee Co-Chair for the 2008 IEEE GLOBECOM, a General Co-Chair for the 2011 IEEE North American School of Information Theory, and an Associate Editor for Shannon Theory for the IEEE Transactions on Information Theory during 2014-2016.

Raymond W. Yeung (S'85-M'88-SM'92-F'03) was born in Hong Kong on June 3, 1962. He received the B.S., M.Eng., and Ph.D. degrees in electrical engineering from Cornell University, Ithaca, NY, in 1984, 1985, and 1988, respectively.

He was on leave at Ecole Nationale Supérieure des Télécommunications, Paris, France, during fall 1986. He was a Member of Technical Staff of AT&T Bell Laboratories from 1988 to 1991. Since 1991, he has been with The Chinese University of Hong Kong, where he is now Choh-Ming Li Professor of Information Engineering and Co-Director of Institute of Network Coding. He has held visiting positions at Cornell University, Nankai University, the University of Bielefeld, the University of Copenhagen, Tokyo Institute of Technology, Munich University of Technology, and Columbia University. He was a consultant in a project of Jet Propulsion Laboratory, Pasadena, CA, for salvaging the malfunctioning Galileo Spacecraft and a consultant for NEC, USA. His 25-bit synchronization marker was used onboard the Galileo Spacecraft for image synchronization.

His research interests include information theory and network coding. He is the author of the textbooks *A First Course in Information Theory* (Kluwer Academic/Plenum 2002) and its revision *Information Theory and Network Coding* (Springer 2008), which have been adopted by over 100 institutions around the world. This book has also been published in Chinese (Higher Education Press 2011, translation by Ning Cai *et al.*). He also co-authored with Shenghao Yang the monograph *BATS Codes: Theory and Applications* (Morgan & Claypool Publishers, 2017). In spring 2014, he gave the first MOOC on information theory that reached over 25,000 students.

Dr. Yeung was a member of the Board of Governors of the IEEE Information Theory Society from 1999 to 2001. He has served on the committees of a number of information theory symposiums and workshops. He was General Chair of the First and the Fourth Workshops on Network, Coding, and Applications (NetCod 2005 and 2008), a Technical Co-Chair for the 2006 IEEE International Symposium on Information Theory, a Technical Co-Chair for the 2006 IEEE Information Theory Workshop (Chengdu, China), and a General Co-Chair of the 2015 IEEE International Symposium on Information Theory. He currently serves as an Editor-at-Large of *Communications in Information and Systems*, an Editor of *Foundation and Trends in Communications and Information Theory* and of *Foundation and Trends in Networking*, and was an Associate Editor for Shannon Theory of the *IEEE Transactions on Information Theory* from 2003 to 2005. In 2011-12, he serves as a Distinguished Lecturer of the IEEE Information Theory Society.

He was a recipient of the Croucher Foundation Senior Research Fellowship for 2000/2001, the Best Paper Award (Communication Theory) of the 2004 International Conference on Communications, Circuits and System, the 2005 IEEE Information Theory Society Paper Award, the Friedrich Wilhelm Bessel Research Award of the Alexander von Humboldt Foundation in 2007, the 2016 IEEE Eric E. Sumner Award ("for pioneering contributions to the field of network coding"), and the 2018 ACM SIGMOBILE Test-of-Time Paper Award. In 2015, he was named (together with Zhen Zhang) an Outstanding Overseas Chinese Information Theorist by the China Information Theory Society. In 2019, his team won a Gold Medal with Congratulations of the Jury at the 47th International Exhibition of Inventions of Geneva for their invention "BATS: Enabling the Nervous System of Smart Cities." He is a Fellow of the IEEE, Hong Kong Academy of Engineering Sciences, and Hong Kong Institution of Engineers.