Distributed Source Coding for Satellite Communications

Raymond W. Yeung, *Senior Member, IEEE*, and Zhen Zhang, *Senior Member, IEEE*

***Abstract*—Inspired by mobile satellite communications systems, we consider a source coding system which consists of multiple sources, multiple encoders, and multiple decoders. Each encoder has access to a certain subset of the sources, each decoder has access to certain subset of the encoders, and each decoder reconstructs a certain subset of the sources almost perfectly. The connectivity between the sources and the encoders, the connectivity between the encoders and the decoders, and the reconstruction requirements for the decoders are all arbitrary. Our goal is to characterize the admissible coding rate region. Despite the generality of the problem, we have developed an approach which enables us to study all cases on the same footing. We obtain inner and outer bounds of the admissible coding rate region in terms of Γ_N^* and $\overline{\Gamma}_N^*$, respectively, which are fundamental regions in the entropy space recently defined by Yeung. So far, there has not been a full characterization of Γ_N^*, so these bounds cannot be evaluated explicitly except for some special cases. Nevertheless, we obtain an alternative outer bound which can be evaluated explicitly. We show that this bound is tight for all the special cases for which the admissible coding rate region is known. The model we study in this paper is more general than all previously reported models on multilevel diversity coding, and the tools we use are new in multiuser information theory.**

***Index Terms*— Diversity coding, multiterminal source coding, multiuser information theory, satellite communication.**

I. Introduction

A mobile satellite communication system, like Motorola's Iridium™ System and Qualcomm's GlobalStar™ System, provides telephone and data services to mobile users within a certain geographical range through satellite links. For example, the Iridium™ System covers users anywhere in the world, while the GlobalStar™ System covers users between ±70° latitude. At any time, a mobile user is covered by one or more satellites. Through satellite links, the message from a mobile user is transmitted to a certain set of mobile users within the system.

In a generic mobile satellite communication system, transmitters and receivers are not necessarily colocated. In principle, a transmitter can transmit information to all the satellites within the line of sight simultaneously; a satellite can combine, encode, and broadcast the information it receives from all the transmitters it covers; and a receiver can combine and decode the information it receives from all the satellites within the line of sight.

We are motivated to study a new problem called the *distributed source coding* problem. A distributed source coding system consists of multiple sources, multiple encoders, and multiple decoders. Each encoder has access to a certain subset of the sources, each decoder has access to a certain subset of the encoders, and each decoder reconstructs a certain subset of the sources almost perfectly. The connectivity between the sources and the encoders, the connectivity between the encoders and the decoders, and the reconstruction requirements for the decoders are all arbitrary.

The sources, the encoders, and the decoders in a distributed coding system correspond to the transmitters, the satellites, and the receivers in a mobile satellite communication system, respectively. The set of encoders connected to a source refers to the set of satellites within the line of sight of a transmitter, while the set of decoders connected to an encoder refers to the set of receivers covered by a satellite.

Throughout the paper, we will use a boldfaced letter to denote a vector. The ith component of a vector $\boldsymbol{x}$ is denoted by x_i unless otherwise specified. For a random variable X, we will use $\mathcal{X}$ to denote the alphabet set of X, and x to denote a generic outcome of X. For a set A, we will use $\overline{A}$ to denote the closure of A. We will use $P\{\cdot\}$ to denote the probability of an event, and $H(\cdot)$ to denote the entropy of a set of random variables in base 2.

Let us now present the formal description of the problem. A distributed source coding system consists of the following elements:

1) $\mathcal{S}$, the index set of the information sources;
2) $\mathcal{E}$, the index set of the encoders;
3) $\mathcal{D}$, the index set of the decoders;
4) $\mathcal{A} \subset \mathcal{S} \times \mathcal{E}$, the set of connections between the sources and the encoders;
5) $\mathcal{B} \subset \mathcal{E} \times \mathcal{D}$, the set of connections between the encoders and the decoders; and
6) $F_m \subset 2^{\mathcal{S}} \backslash \phi$, $m \in \mathcal{D}$, which specify the reconstruction requirements of the decoders.

The jth source is denoted by $S_j = \{X_{jk}\}_{k=1}^{\infty}$, $j \in \mathcal{S}$. We assume that S_j, $j \in \mathcal{S}$ are independent, and X_{jk}, $k = 1, 2, \cdots$ are independent and identically distributed (i.i.d.) copies of a generic random variable X_j, where $|\mathcal{X}_j| < \infty$. The lth encoder

Manuscript received December 12, 1996; revised November 4, 1998. The work of R. W. Yeung was supported in part by the Research Grant Council of Hong Kong under Earmarked Grant CUHK 332/96E. The work of Z. Zhang was supported in part by the National Science Foundation under Grant NCR-9508282.

R. W. Yeung is with Department of Information Engineering, The Chinese University of Hong Kong, N.T., Hong Kong, China (e-mail: whyeung@ie.cuhk.edu.hk).

Z. Zhang is with the Communication Sciences Institute, Department of Electrical Engineering–Systems, University of Southern California, Los Angeles, CA 90089-2565 USA (e-mail: zzhang@milly.usc.edu).

Communicated by K. Zeger, Editor At Large.

Publisher Item Identifier S 0018-9448(99)03157-0.

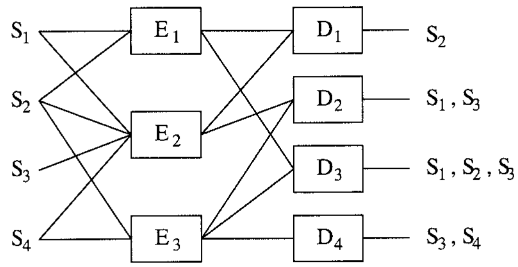

Fig. 1. A distributed source coding system.

is denoted by E_l, $l \in \mathcal{E}$, and the mth decoder is denoted by D_m, $m \in \mathcal{D}$. The sets $\mathcal{A}$, $\mathcal{B}$, and F_m, $m \in \mathcal{D}$ specify the distributed source coding system as follows: E_l has access to S_j if and only if $(j, l) \in \mathcal{A}$, D_m has access to E_l if and only if $(l, m) \in \mathcal{B}$, and D_m reconstructs S_j, $j \in F_m$.

Let us illustrate the above notations by a small example. For the distributed source coding system in Fig. 1,

$$\begin{aligned}\mathcal{S} &= \{1, 2, 3, 4\}, \quad \mathcal{E} = \{1, 2, 3\}, \quad \mathcal{D} = \{1, 2, 3, 4\}\\ \mathcal{A} &= \{(1, 1), (1, 2), (2, 1), (2, 2), (2, 3), (3, 3), (4, 2),\\ &\quad (4, 3)\}\\ \mathcal{B} &= \{(1, 1), (1, 3), (2, 1), (2, 2), (3, 2), (3, 3), (3, 4)\}\\ F_1 &= \{2\}, \quad F_2 = \{1, 3\}, \quad F_3 = \{1, 2, 3\}, \quad F_4 = \{3, 4\}.\end{aligned}$$

To facilitate the description of our model, we define

$$U_l = \{j \in \mathcal{S}\colon \ (j, l) \in \mathcal{A}\}, \qquad l \in \mathcal{E}$$

and

$$V_m = \{l \in \mathcal{E}\colon \ (l, m) \in \mathcal{B}\}, \qquad m \in \mathcal{D}.$$

U_l contains the indices of the sources which are accessed by E_l, and V_m contains the indices of the encoders which are accessed by D_m. Let

$$d_m\colon \ \left(\prod_{j \in F_m} \mathcal{X}_j\right) \times \left(\prod_{j \in F_m} \mathcal{X}_j\right) \to \{0, 1\}$$

be the Hamming distortion measure $m \in \mathcal{D}$; i.e., for any x and x' in $(\prod_{j \in F_m} \mathcal{X}_j) \times (\prod_{j \in F_m} \mathcal{X}_j)$

$$d_m(x, x') = \begin{cases} 0, & \text{if } x = x' \\ 1, & \text{if } x \neq x'. \end{cases}$$

Let $\boldsymbol{X}_j = (X_{j1}, \cdots, X_{jn})$. An $(n, (\eta_l, l \in \mathcal{E}), (\Delta_m, m \in \mathcal{D}))$ code is defined by

$$\begin{aligned} T_l\colon &\quad \prod_{j \in U_l} \mathcal{X}_j^n \to \{0, 1, \cdots, \eta_l - 1\}, && l \in \mathcal{E}\\ W_m\colon &\quad \prod_{l \in V_m} \{0, 1, \cdots, \eta_l - 1\} \to \prod_{j \in F_m} \mathcal{X}_j^n, && m \in \mathcal{D}\end{aligned}$$

and

$$\Delta_m = n^{-1} \mathrm{E} \sum_{k=1}^{n} d_m((X_{jk}, j \in F_m), (\hat{X}_{jk}, j \in F_m)), \qquad m \in \mathcal{D}$$

where

$$((\hat{X}_{j1}, \cdots, \hat{X}_{jn}), j \in F_m) = W_m(T_l(\boldsymbol{X}_j, j \in U_l), l \in V_m).$$

An $|\mathcal{E}|$-tuple $(R_l, l \in \mathcal{E})$ is admissible if for every $\epsilon > 0$, there exists for sufficiently large n an $(n, (\eta_l, l \in \mathcal{E}), (\Delta_m, m \in \mathcal{D}))$ code such that

$$n^{-1} \log \eta_l \leq R_l + \epsilon, \qquad \text{for all } l \in \mathcal{E}$$

and

$$\Delta_m \leq \epsilon, \qquad \text{for all } m \in \mathcal{D}.$$

Let $\boldsymbol{R} = (R_l, l \in \mathcal{E})$, and let

$$\mathcal{R} \stackrel{\text{def}}{=} \{\boldsymbol{R}\colon \boldsymbol{R} \text{ is admissible}\}$$

be the admissible coding rate region (henceforth the coding rate region if there is no ambiguity). The goal of this paper is to characterize $\mathcal{R}$. In Section II, we prove an inner bound $\mathcal{R}_{\text{in}}$ for $\mathcal{R}$. In Section III, we prove an outer bound $\mathcal{R}_{\text{out}}$ for $\mathcal{R}$. In Section IV, we give geometrical interpretations of $\mathcal{R}_{\text{in}}$ and $\mathcal{R}_{\text{out}}$.

The regions $\mathcal{R}_{\text{in}}$ and $\mathcal{R}_{\text{out}}$ are defined in terms of Γ_N^* and $\overline{\Gamma}_N^*$, which are fundamental regions recently defined by Yeung [10]. So far, there has not been a full characterization of either Γ_N^* or $\overline{\Gamma}_N^*$, so explicit evaluation of $\mathcal{R}_{\text{in}}$ and $\mathcal{R}_{\text{out}}$ is not possible. Based on the geometrical interpretation of these regions, we obtain an outer bound for $\mathcal{R}$ called the LP bound (for linear programming bound) which can be evaluated explicitly. This is described in Section V. In this section, we also show that the LP bound is tight for all the special cases for which the admissible coding rate region is known. Concluding remarks are in Section VI.

II. Inner Bound

Before we define $\mathcal{R}_{\text{in}}$, we first introduce some notation from the framework for information inequalities developed in [10]. Let N be a finite set of discrete random variables whose joint distribution is unspecified, and let $Q(N) = 2^N \backslash \phi$. Note that $|Q(N)| = 2^{|N|} - 1$. Let h_G, $G \in Q(N)$ denote the coordinates of $\mathbb{R}^{2^{|N|}-1}$. A vector $\boldsymbol{h} \stackrel{\text{def}}{=} (h_G, G \in Q(N))$ is said to be *constructible* if there exists a joint distribution for $(X, X \in N)$ such that $H(X, X \in G) = h_G$. We then define

$$\Gamma_N^* = \{\boldsymbol{h} \in \mathbb{R}^{2^{|N|}-1}\colon \boldsymbol{h} \text{ is constructible}\}.$$

To simplify notation in the sequel, for any nonempty G, G', $G'' \in Q(N)$, we further define

$$\begin{aligned} h_{G|G'} &= h_{GG'} - h_{G'}\\ i_{G;G'} &= h_G - h_{G|G'}\\ i_{G;G'|G''} &= h_{G|G''} - h_{G|G'G''}\end{aligned}$$

where we have used juxtaposition to denote the union of two sets. In using the above definitions, we will not distinguish elements and singletons of N; i.e., for a random variable $X \in N$, h_X is the same as $h_{\{X\}}$.

Let Y_j, $j \in \mathcal{S}$ and Z_l, $l \in \mathcal{E}$ be discrete random variables whose joint distribution is unspecified, and let

$$N = \{Y_j, j \in \mathcal{S}; Z_l, l \in \mathcal{E}\}$$

i.e., the set containing all the random variables Y_j, $j \in \mathcal{S}$ and Z_l, $l \in \mathcal{E}$. Y_j is an auxiliary random variable associated with the source S_j, and Z_l is an auxiliary random variable associated with $T_l(\boldsymbol{X}_j,\, j \in U_l)$, the output of the encoder E_l. The actual meaning of Y_j and Z_l will become clear later. Let $\mathcal{R}'$ be the set of all $\boldsymbol{R} = (R_l,\, l \in \mathcal{E})$ such that there exists $\boldsymbol{h} \in \Gamma_N^*$ which satisfies the following conditions:

$$h_{(Y_j,\, j\in\mathcal{S})} = \sum_{j\in\mathcal{S}} h_{Y_j} \tag{1}$$

$$h_{Z_l|(Y_j,\, j\in U_l)} = 0, \qquad \text{for } l \in \mathcal{E} \tag{2}$$

$$h_{(Y_j,\, j\in F_m)|(Z_l,\, l\in V_m)} = 0, \qquad \text{for } m \in \mathcal{D} \tag{3}$$

$$h_{Y_j} > H(X_j), \qquad \text{for } j \in \mathcal{S} \tag{4}$$

$$R_l \geq h_{Z_l}, \qquad \text{for } l \in \mathcal{E}. \tag{5}$$

Note that (1)–(3) are hyperplanes, and (4) is an open halfspace in $\mathbb{R}^{2^{|N|}-1}$. We then define $\mathcal{R}_{\text{in}} = \overline{\mathcal{R}'}$.

Theorem 1: $\mathcal{R}_{\text{in}} \subset \mathcal{R}$.

Note that $\mathcal{R}'$ can be defined more conventionally as the set of all $\boldsymbol{R}$ such that there exist auxiliary random variables $Y_j,\, j \in \mathcal{S}$ and $Z_l,\, l \in \mathcal{E}$ satisfying

$$H(Y_j,\, j \in \mathcal{S}) = \sum_{j\in\mathcal{S}} H(Y_j) \tag{6}$$

$$H(Z_l|Y_j,\, j \in U_l) = 0, \qquad \text{for } l \in \mathcal{E} \tag{7}$$

$$H(Y_j,\, j \in F_m|Z_l,\, l \in V_m) = 0, \qquad \text{for } m \in \mathcal{D} \tag{8}$$

$$H(Y_j) > H(X_j), \qquad \text{for } j \in \mathcal{S} \tag{9}$$

$$R_l \geq H(Z_l), \qquad \text{for } l \in \mathcal{E}. \tag{10}$$

Although this alternative definition is more intuitive, the region so defined appears to be totally different from case to case. On the other hand, defining $\mathcal{R}'$ in terms of Γ_N^* enables us to study all cases on the same footing. In particular, if $\tilde{\Gamma}_N$ is an explicit inner bound on Γ_N^*, upon replacing Γ_N^* by $\tilde{\Gamma}_N$ in the definition of $\mathcal{R}'$, we immediately obtain an explicit inner bound on $\mathcal{R}_{\text{in}}$ for all cases. (Unfortunately, no explicit inner bound on Γ_N^* for $|N| \geq 5$ is available at this point; a nontrivial inner bound for $|N| = 4$ has been obtained in [6] and [12].) The introduction of $\mathcal{R}_{\text{LP}}$ in Section V as an explicit outer bound on $\mathcal{R}_{\text{out}}$ is in the same spirit.

Let us give the motivations of the conditions in (1)–(5) before we prove the theorem, although the meaning of these conditions cannot be fully explained until we come to the proof. The condition (1) corresponds to the assumption that the sources $S_j,\, j \in \mathcal{S}$ are independent. The condition (2) corresponds to the fact that the encoder E_l has access to the sources $S_j,\, j \in U_l$. The condition (3) corresponds to the requirement that the sources $S_j,\, j \in F_m$ can be reconstructed by the decoder D_m. The condition (4) means that the entropy of the auxiliary random variable Y_j is strictly greater than the entropy rate of the source S_j, and the condition (5) means that the coding rate of the encoder E_l is greater than or equal to the entropy of the auxiliary random variable Z_l.

We will state the following lemma before we present the proof of the theorem. Since this lemma is a standard result, its proof will be omitted. We first recall the definitions of strong typicality of sequences [2]. A sequence $\boldsymbol{x} \in \mathcal{X}^n$ is *δ-typical* with respect to a distribution $p(x)$ if for all $x \in \mathcal{X}$

$$\left|\frac{1}{n}N(x|\boldsymbol{x}) - p(x)\right| < \frac{\delta}{|\mathcal{X}|}$$

where $N(x|\boldsymbol{x})$ is the number of occurrences of the symbol x in $\boldsymbol{x}$. Similarly, a pair of sequences $(\boldsymbol{x},\, \boldsymbol{y}) \in \mathcal{X}^n \times \mathcal{Y}^n$ is δ-typical with respect to a distribution $p(x,\, y)$ if for all $(x,\, y) \in \mathcal{X} \times \mathcal{Y}$

$$\left|\frac{1}{n}N(x,\, y|\boldsymbol{x},\, \boldsymbol{y}) - p(x,\, y)\right| < \frac{\delta}{|\mathcal{X}||\mathcal{Y}|}.$$

In the sequel, the δ-typical notations in [4] will be adopted.

Lemma 1: Let $(\boldsymbol{x},\, \boldsymbol{y})$ be n-vectors drawn according to

$$p(\boldsymbol{x},\, \boldsymbol{y}) = \prod_{i=1}^{n} p(x_i,\, y_i). \tag{11}$$

If $(\tilde{\boldsymbol{X}},\, \tilde{\boldsymbol{Y}}) \sim p(\boldsymbol{x})p(\boldsymbol{y})$, then

$$P\{(\tilde{\boldsymbol{X}},\, \tilde{\boldsymbol{Y}}) \in T^n_{[XY]\delta}\} \leq 2^{-n(I(X;Y)-O(\delta))}$$

where $O(\delta)$ denotes any function $\gamma(\delta)$ such that $|\gamma(\delta)| < c|\delta|$ for some constant $c > 0$ in a neighborhood of $\delta = 0$.

Proof of Theorem 1: We will prove the theorem by describing a random coding scheme and showing that it has the desired performance. Let ν, δ be small positive quantities and n be a large integer to be specified later. Let $\boldsymbol{h} \in \mathcal{R}'$. Then there exist random variables $Y_j,\, j \in \mathcal{S}$ and $Z_l,\, l \in \mathcal{E}$ such that the left-hand sides of (1)–(5) are the corresponding Shannon information measures; i.e., (1)–(5) can be written as (6)–(10).

The encoding scheme is described in the following steps:

1. Let p_j be the distribution of Y_j, and let $\boldsymbol{y} = (y_1,\, \cdots,\, y_n)$. For each $j \in \mathcal{S}$, independently generate $\lfloor 2^{n(H(Y_j)-\nu)} \rfloor$ vectors in $\mathcal{Y}_j^n$ according to the distribution

$$p(\boldsymbol{y}) = \prod_{k=1}^{n} p_j(y_k) \tag{12}$$

where $\boldsymbol{y} \in \mathcal{Y}_j^n$. Let Ω_j be the set of the $\lfloor 2^{n(H(Y_j)-\nu)} \rfloor$ vectors generated, and denote these vectors by

$$\underline{\omega}_j(i_j), \qquad 0 \leq i_j \leq |\Omega_j| - 1.$$

2. Let $\zeta_j \stackrel{\text{def}}{=} |T^n_{[X_j]\delta}|$ and index the vectors in $T^n_{[X_j]\delta}$ by the set $\{1,\, 2,\, \cdots,\, \zeta_j\}$. Since $H(Y_j) > H(X_j)$ by (4), by taking δ and ν small enough and n sufficiently large, we can make

$$\begin{aligned}\zeta_j &= |T^n_{[X_j]\delta}| < 2^{n(H(X_j)+O(\delta))} < \lfloor 2^{n(H(Y_j)-\nu)} \rfloor - 1 \\ &= |\Omega_j| - 1.\end{aligned} \tag{13}$$

Define a mapping

$$\rho_j : \mathcal{X}_j^n \to \{0,\, 1,\, \cdots,\, \zeta_j\}$$

as follows. For any $\boldsymbol{x} \in \mathcal{X}_j^n$, if $\boldsymbol{x} \notin T^n_{[X_j]\delta}$, then $\rho_j(\boldsymbol{x}) = 0$; if $\boldsymbol{x} \in T^n_{[X_j]\delta}$, then $\rho_j(\boldsymbol{x})$ is equal to the index of $\boldsymbol{x}$ in $T^n_{[X_j]\delta}$ defined above. We note that for $\boldsymbol{x} \notin T^n_{[X_j]\delta}$, $\underline{\omega}_j(\rho_j(\boldsymbol{x}))$ is a constant vector in Ω_j, and for $\boldsymbol{x} \in T^n_{[X_j]\delta}$, $\underline{\omega}_j(\rho_j(\boldsymbol{x}))$ are distinct vectors in Ω_j (for all $\boldsymbol{x} \in T^n_{[X_j]\delta}$, $\rho_j(\boldsymbol{x}) \le |\Omega_j| - 1$ by (13), so $\underline{\omega}_j(\rho_j(\boldsymbol{x}))$ is properly defined).

3. Let q_l be the conditional distribution of Z_l given $(Y_j, j \in U_l)$. Let $\eta_l \stackrel{\text{def}}{=} |T^n_{[Z_l]\delta}|$ and index the vectors in $T^n_{[Z_l]\delta}$ by the set $\{0, 1, \cdots, \eta_l - 1\}$. For each $l \in \mathcal{E}$, define a mapping

$$\psi_l : \prod_{j \in U_l} \Omega_j \to \{0, 1, \cdots, \eta_l - 1\}$$

randomly as follows. For each $(\boldsymbol{y}_j, j \in \mathcal{U}_l) \in \prod_{j\in U_l} \Omega_j$, send it through a discrete memoryless channel with transition probability q_l by using the channel n times. If the received vector is not in $T^n_{[Z_l]\delta}$, then $\psi_l(\boldsymbol{y}_j, j \in \mathcal{U}_l) = 0$, otherwise, $\psi_l(\boldsymbol{y}_j, j \in \mathcal{U}_l)$ is equal to the index of the received vector in $T^n_{[Z_l]\delta}$ defined above.

4. The encoding function

$$T_l\colon \prod_{j \in U_l} \mathcal{X}_j^n \to \{0, 1, \cdots, \eta_l - 1\}$$

is defined as follows. For each

$$(\boldsymbol{x}_j, j \in U_l) \in \prod_{j \in U_l} \mathcal{X}_j^n$$

if $\rho_j(\boldsymbol{x}_j) = 0$ for some $j \in U_l$, then $T_l(\boldsymbol{x}_j, j \in U_l) = 0$, otherwise,

$$T_l(\boldsymbol{x}_j, j \in U_l) = \psi_l(\underline{\omega}_j(\rho_j(\boldsymbol{x}_j)), j \in U_l).$$

The decoding scheme is described in the following steps:

1) For each $m \in \mathcal{D}$, define a mapping

$$\xi_m : \prod_{l \in V_m} \{0, 1, \cdots, \eta_l - 1\} \to \prod_{j \in F_m} \{0, 1, \cdots, \zeta_j\}$$

as follows. Let $\boldsymbol{z}_l(i)$ be the vector in $T^n_{[Z_l]\delta}$ with index i (cf. step 3 in the encoding scheme). For any

$$(t_l, l \in V_m) \in \prod_{l \in V_m} \{0, 1, \cdots, \eta_l - 1\}$$

if there exists a unique set of indices

$$(i_j, j \in F_m) \in \prod_{j \in F_m} \{0, 1, \cdots, \zeta_j\}$$

such that

$$((\underline{\omega}_j(i_j), j \in F_m), (\boldsymbol{z}_l(t_l), l \in V_m)) \in T^n_{[(Y_j, j \in F_m), (Z_l, l \in V_m)]\delta}$$

then

$$\xi_m(t_l, l \in V_m) = (i_j, j \in F_m)$$

otherwise

$$\xi_m(t_l, l \in V_m) = \underbrace{(0, \cdots, 0)}_{|F_m|}.$$

2) For each $j \in \mathcal{E}$, define a mapping

$$\mu_j : \{0, 1, \cdots, \zeta_j\} \to \mathcal{X}_j^n$$

by

$$\mu_j(i) = \begin{cases} \rho_j^{-1}(i), & \text{if } 1 \le i \le \zeta_j \\ \boldsymbol{x}_{j0}, & \text{otherwise} \end{cases} \tag{14}$$

where $\boldsymbol{x}_{j0}$ is an arbitrary constant vector in $\mathcal{X}_j^n$ (cf. step 2 of the encoding scheme).

3) The decoding function

$$W_m : \prod_{l \in V_m} \{0, 1, \cdots, \eta_l - 1\} \to \prod_{j \in F_m} \mathcal{X}_j^n$$

is defined as

$$W_m(t_l, l \in V_m) = (\mu_j(\hat{i}_j), j \in F_m)$$

where

$$(\hat{i}_j, j \in F_m) \stackrel{\text{def}}{=} \xi_m(t_l, l \in V_m).$$

We now analyze the performance of this random code. Let ϵ be any positive quantity. First, for all $l \in \mathcal{E}$, since

$$\eta_l = |T^n_{[Z_l]\delta}| < 2^{n(H(Z_l) + O(\delta))} \tag{15}$$

we have

$$n^{-1} \log \eta_l < H(Z_l) + O(\delta) \le H(Z_l) + \epsilon \tag{16}$$

by taking δ sufficiently small. Therefore,

$$n^{-1} \log \eta_l \le R_l + \epsilon, \qquad \text{for all } l \in \mathcal{E} \tag{17}$$

by (10). We now introduce the following notations:

$$\begin{aligned} \boldsymbol{Y}_j &= \underline{\omega}_j(\rho_j(\boldsymbol{X}_j)) \\ T_l &= T_l(\boldsymbol{X}_j, j \in U_l) \\ \boldsymbol{Z}_l &= \boldsymbol{z}_l(T_l) \\ \hat{\boldsymbol{Y}}_j &= \underline{\omega}_j(\hat{i}_j). \end{aligned}$$

Define the events

$$\begin{aligned} E_1 &= \{\rho_j(\boldsymbol{X}_j) \ne 0, j \in \mathcal{S}\} \\ E_2 &= \{((\boldsymbol{Y}_j, j \in \mathcal{S}), (\boldsymbol{Z}_l, l \in \mathcal{E})) \in T^n_{[(Y_j, j \in \mathcal{S}), (Z_l, l \in \mathcal{E})]\delta}\} \\ E_{2m} &= \{((\boldsymbol{Y}_j, j \in F_m), (\boldsymbol{Z}_l, l \in V_m)) \\ &\qquad \in T^n_{[(Y_j, j \in F_m), (Z_l, l \in V_m)]\delta}\} \\ E_{3m} &= \{\hat{i}_j = \rho_j(\boldsymbol{X}_j), j \in F_m\}. \end{aligned}$$

E_1 is the event that $\boldsymbol{X}_j$ is δ-typical with respect to $p(x_j)$ for all $j \in \mathcal{S}$. E_{3m} is the event that Decoder m decodes correctly the index assigned to $\boldsymbol{X}_j$ for all $j \in F_m$.

By the weak law of large numbers, we see that for sufficiently large n

$$P\{E_1^c\} = P\{\boldsymbol{X}_j \notin T^n_{[X_j]\delta} \text{ for some } j \in \mathcal{S}\} \le \frac{\epsilon}{3}. \tag{18}$$

Let $p((y_j,\, j \in \mathcal{S}),\, (z_l,\, l \in \mathcal{E}))$ be the joint distribution of $((Y_j,\, j \in \mathcal{S}),\, (Z_l,\, l \in \mathcal{E}))$. By (6) and (7), we have

$$\begin{aligned} p((y_j,\, j \in \mathcal{S}),\, (z_l,\, l \in \mathcal{E})) &= P\{Y_j = y_j,\, j \in \mathcal{S}\} P\{Z_l = z_l,\, l \in \mathcal{E} | Y_j = y_j,\, j \in \mathcal{S}\} \\ &= \prod_{j \in \mathcal{S}} p_j(y_j) \prod_{l \in \mathcal{E}} q_l(z_l | y_j,\, j \in U_l). \end{aligned} \tag{19}$$

Since $((\boldsymbol{Y}_j,\, j \in \mathcal{S}),\, (\boldsymbol{Z}_l,\, l \in \mathcal{E}))$ is drawn i.i.d. according to the above distribution, we see from the weak law of large numbers that for a sufficiently large n

$$P\{E_2^c\} \le \frac{\epsilon}{3}. \tag{20}$$

By noting that E_2 implies $E_{2m},\, m \in \mathcal{D}$, we see that for any $m \in \mathcal{D}$

$$P\{E_{2m}^c\} \le P\{E_2^c\} \le \frac{\epsilon}{3}. \tag{21}$$

Further, since E_1 and E_{2m} are independent events, we have

$$P\{E_{2m}^c | E_1\} = P\{E_{2m}^c\} \le \frac{\epsilon}{3}. \tag{22}$$

Now for any $m \in \mathcal{D}$ and sufficiently large n, we have

$$\begin{aligned} \Delta_m &= \mathrm{E} n^{-1} \sum_{k=1}^{n} d_m((X_{jk},\, j \in F_m),\, (\hat{X}_{jk},\, j \in F_m)) \\ &\le \mathrm{E}[\cdot|E_1] P[E_1] + \mathrm{E}[\cdot|E_1^c] P\{E_1^c\} \\ &\le \mathrm{E}[\cdot|E_1] + 1 \cdot P\{E_1^c\} \\ &\le \mathrm{E}[\cdot|E_1 E_{2m}] P\{E_{2m}|E_1\} \\ &\quad + \mathrm{E}[\cdot|E_1 E_{2m}^c] P\{E_{2m}^c|E_1\} + \frac{\epsilon}{3} \\ &\le \mathrm{E}[\cdot|E_1 E_{2m}] + 1 \cdot P\{E_{2m}^c|E_1\} + \frac{\epsilon}{3} \\ &\le \mathrm{E}[\cdot|E_1 E_{2m} E_{3m}] P\{E_{3m}|E_1 E_{2m}\} \\ &\quad + \mathrm{E}[\cdot|E_1 E_{2m} E_{3m}^c] P\{E_{3m}^c|E_1 E_{2m}\} + \tfrac{2}{3}\epsilon \\ &\le \mathrm{E}[\cdot|E_1 E_{2m} E_{3m}] + P\{E_{3m}^c|E_1 E_{2m}\} + \tfrac{2}{3}\epsilon. \end{aligned} \tag{23}$$

Given E_1, E_{2m}, and E_{3m}

$$\begin{aligned} W_m(t_l,\, l \in V_m) &= (\mu_j(\hat{i}_j),\, j \in F_m) \\ &= (\mu_j(\rho_j(\boldsymbol{X}_j)),\, j \in F_m) \\ &= (\boldsymbol{X}_j,\, j \in F_m). \end{aligned} \tag{24}$$

Note that the last equality follows from (14) because

$$1 \le \rho_j(\boldsymbol{X}_j) \le \zeta_j \tag{25}$$

given E_1. Therefore,

$$\mathrm{E}[\cdot|E_1 E_{2m} E_{3m}] = 0. \tag{26}$$

The discussion that follows is a variation of the Channel Coding Theorem. We claim that $(\boldsymbol{Z}_l,\, l \in V_m)$ can be obtained by sending $(\boldsymbol{Y}_j,\, j \in F_m)$ through a discrete memoryless channel with transition probability $\tilde{q}_m$, where $\tilde{q}_m$ is the conditional distribution of $(Z_l,\, l \in V_m)$ given $(Y_j,\, j \in F_m)$. To simply notation, we assume $\boldsymbol{y}_j \in \mathcal{Y}_j^n$ and denote $\mathcal{S} \backslash F_m$ by F_m^c in the following. To see that the claim is true, consider (27) at the top of the following page. Note that in 1) in (27)

$$\begin{aligned} P\{\boldsymbol{Y}_j = \boldsymbol{y}_j,\, j \in F_m\} &= \prod_{j \in F_m} P\{\boldsymbol{Y}_j = \boldsymbol{y}_j\} \\ &= \prod_{j \in F_m} \prod_{k=1}^{n} P\{Y_{jk} = y_{jk}\} \\ &= \prod_{j \in F_m} \prod_{k=1}^{n} P\{Y_j = y_{jk}\}. \end{aligned} \tag{28}$$

Assuming without loss of generality that $P\{Y_j = y\} > 0$ for all $y \in \mathcal{Y}_j$

$$P\{\boldsymbol{Y}_j = \boldsymbol{y}_j,\, j \in F_m\} > 0. \tag{29}$$

Therefore its reciprocal is well-defined.

We now bound the probability $P\{E_{3m}^c | E_1 E_{2m}\}$ in (23). By the symmetry of the problem, we assume without loss of generality that $\rho_j(\boldsymbol{X}_j) = 1$ for $j \in F_m$. To simplify notation, define the sets

$$\Phi = \prod_{j \in F_m} \{0,\, 1,\, \cdots,\, |\Omega_j| - 1\} \backslash (\underbrace{1,\, \cdots,\, 1}_{|F_m|})$$

and

$$\Lambda_\Psi = \{(i_j,\, j \in F_m) \in \Phi \colon i_j = 1 \text{ iff } j \in \Psi\}$$

where $\Psi \in 2^{F_m} \backslash F_m$. Note that $\{\Lambda_\Psi,\, \Psi \in 2^{F_m} \backslash F_m\}$ is a partition of Φ, and

$$\begin{aligned} |\Lambda_\Psi| &= \prod_{j \in F_m \backslash \Psi} (|\Omega_j| - 1) < \prod_{j \in F_m \backslash \Psi} 2^{n(H(Y_j) - \nu)} \\ &= 2^{n(H(Y_j,\, j \in F_m \backslash \Psi) - (|F_m| - |\Psi|)\nu)} \\ &\le 2^{n(H(Y_j,\, j \in F_m \backslash \Psi) - \nu)} \end{aligned} \tag{30}$$

where the last inequality follows because Ψ is always a proper subset of F_m. Then

$$\begin{aligned} &P\{E_{3m}^c | E_1 E_{2m}\} \\ &\quad = P\{((\underline{\omega}_j(i_j),\, j \in F_m),\, (\boldsymbol{z}_l(t_l),\, l \in V_m)) \\ &\qquad \in T^n_{[(Y_j,\, j \in F_m),\, (Z_l,\, l \in V_m)]\delta} \text{ for some } (i_j,\, j \in F_m) \in \Phi\} \\ &\quad \le \sum_{(i_j,\, j \in F_m) \in \Phi} P\{((\underline{\omega}_j(i_j),\, j \in F_m),\, (\boldsymbol{z}_l(t_l),\, l \in V_m)) \\ &\qquad \in T^n_{[(Y_j,\, j \in F_m),\, (Z_l,\, l \in V_m)]\delta}\} \\ &\quad = \sum_{\Psi} \sum_{(i_j,\, j \in F_m) \in \Lambda_\Psi} P\{((\underline{\omega}_j(i_j),\, j \in F_m),\, (\boldsymbol{z}_l(t_l),\, l \in V_m)) \\ &\qquad \in T^n_{[(Y_j,\, j \in F_m),\, (Z_l,\, l \in V_m)]\delta}\}. \end{aligned} \tag{31}$$

We now prove an identity by considering the information diagram [8] for the random variables $(Y_j,\, j \in F_m \backslash \Psi)$, $(Y_j,\, j \in F_m \cap \Psi)$, and $(Z_l,\, l \in V_m)$ in Fig. 2. By (8), we see from the diagram that

$$H(Y_j, j \in F_m \backslash \Psi | (Y_j, j \in F_m \cap \Psi), (Z_l,\, l \in V_m)) = 0 \tag{32}$$

$$I(Y_j, j \in F_m \backslash \Psi; Y_j, j \in F_m \cap \Psi | Z_l, l \in V_m) = 0 \tag{33}$$

$$
\begin{aligned}
&P\{\boldsymbol{Z}_l=\boldsymbol{z}_l,\, l\in V_m|\boldsymbol{Y}_j=\boldsymbol{y}_j,\, j\in F_m\}\\
&\quad=\sum_{(\boldsymbol{y}_j,\, j\in F_m^c)} P\{(\boldsymbol{Z}_l=\boldsymbol{z}_l,\, l\in V_m),\, (\boldsymbol{Y}_j=\boldsymbol{y}_j,\, j\in F_m^c)|\boldsymbol{Y}_j=\boldsymbol{y}_j,\, j\in F_m\}\\
&\quad=\sum_{(-)} P\{\boldsymbol{Z}_l=\boldsymbol{z}_l,\, l\in V_m|\boldsymbol{Y}_j=\boldsymbol{y}_j,\, j\in \mathcal{S}\}P\{\boldsymbol{Y}_j=\boldsymbol{y}_j,\, j\in F_m^c|\boldsymbol{Y}_j=\boldsymbol{y}_j,\, j\in F_m\}\\
&\quad=\sum_{(-)} P\{\boldsymbol{Z}_l=\boldsymbol{z}_l,\, l\in V_m|\boldsymbol{Y}_j=\boldsymbol{y}_j,\, j\in \mathcal{S}\}P\{\boldsymbol{Y}_j=\boldsymbol{y}_j,\, j\in F_m^c\}\\
&\quad=^{1)} P\{\boldsymbol{Y}_j=\boldsymbol{y}_j,\, j\in F_m\}^{-1}\sum_{(-)} P\{\boldsymbol{Z}_l=\boldsymbol{z}_l,\, l\in V_m|\boldsymbol{Y}_j=\boldsymbol{y}_j,\, j\in \mathcal{S}\}P\{\boldsymbol{Y}_j=\boldsymbol{y}_j,\, j\in F_m^c\}\\
&\qquad\cdot P\{\boldsymbol{Y}_j=\boldsymbol{y}_j,\, j\in F_m\}\\
&\quad=P\{\cdot\}^{-1}\sum_{(-)} P\{\boldsymbol{Z}_l=\boldsymbol{z}_l,\, l\in V_m|\boldsymbol{Y}_j=\boldsymbol{y}_j,\, j\in \mathcal{S}\}P\{\boldsymbol{Y}_j=\boldsymbol{y}_j,\, j\in \mathcal{S}\}\\
&\quad=P\{\cdot\}^{-1}\sum_{(-)} P\{(\boldsymbol{Y}_j=\boldsymbol{y}_j,\, j\in \mathcal{S}),\, (\boldsymbol{Z}_l=\boldsymbol{z}_l,\, l\in V_m)\}\\
&\quad=P\{\cdot\}^{-1}\sum_{(\boldsymbol{y}_j,\, j\in F_m^c)}\prod_{k=1}^{n} P\{(Y_{jk}=y_{jk},\, j\in \mathcal{S}),\, (Z_{lk}=z_{lk},\, l\in V_m)\}\\
&\quad=P\{\cdot\}^{-1}\sum_{(y_{j1},\, j\in F_m^c)}\cdots\sum_{(y_{jn},\, j\in F_m^c)}\prod_{k} P\{(Y_{jk}=y_{jk},\, j\in \mathcal{S}),\, (Z_{lk}=z_{lk},\, l\in V_m)\}\\
&\quad=P\{\cdot\}^{-1}\prod_{k}\sum_{(y_{jk},\, j\in F_m^c)} P\{(Y_{jk}=y_{jk},\, j\in \mathcal{S}),\, (Z_{lk}=z_{lk},\, l\in V_m)\}\\
&\quad=P\{\cdot\}^{-1}\prod_{k}\sum_{(y_{jk},\, j\in F_m^c)} P\{Y_{jk}=y_{jk},\, j\in \mathcal{S}\}P\{Z_{lk}=z_{lk},\, l\in V_m|Y_{jk}=y_{jk},\, j\in \mathcal{S}\}\\
&\quad=P\{\cdot\}^{-1}\prod_{k}\sum_{(y_{jk},\, j\in F_m^c)} P\{Y_{jk}=y_{jk},\, j\in F_m\}P\{Y_{jk}=y_{jk},\, j\in F_m^c\}\\
&\qquad\cdot P\{Z_{lk}=z_{lk},\, l\in V_m|Y_{jk}=y_{jk},\, j\in \mathcal{S}\}\\
&\quad=P\{\cdot\}^{-1}\prod_{k}P\{Y_{jk}=y_{jk},\, j\in F_m\}\sum_{(y_{jk},\, j\in F_m^c)} P\{Y_{jk}=y_{jk},\, j\in F_m^c|Y_{jk}=y_{jk},\, j\in F_m\}\\
&\qquad\cdot P\{Z_{lk}=z_{lk},\, l\in V_m|Y_{jk}=y_{jk},\, j\in \mathcal{S}\}\\
&\quad=P\{\cdot\}^{-1}\prod_{k}P\{Y_{jk}=y_{jk},\, j\in F_m\}\\
&\qquad\times\sum_{(y_{jk},\, j\in F_m^c)} P\{(Z_{lk}=z_{lk},\, l\in V_m),\, (Y_{jk}=y_{jk},\, j\in F_m^c)|Y_{jk}=y_{jk},\, j\in F_m\}\\
&\quad=P\{\cdot\}^{-1}\prod_{k}P\{Y_{jk}=y_{jk},\, j\in F_m\}P\{Z_{lk}=z_{lk},\, l\in V_m|Y_{jk}=y_{jk},\, j\in F_m\}\\
&\quad=P\{\cdot\}^{-1}\left(\prod_{k=1}^{n}P\{Y_{jk}=y_{jk},\, j\in F_m\}\right)\prod_{k=1}^{n}\tilde{q}_m(z_{lk},\, l\in V_m|y_{jk},\, j\in F_m)\\
&\quad=P\{\cdot\}^{-1}P\{\cdot\}\prod_{k=1}^{n}\tilde{q}_m(z_{lk},\, l\in V_m|y_{jk},\, j\in F_m)\\
&\quad=\prod_{k=1}^{n}\tilde{q}_m(z_{lk},\, l\in V_m|y_{jk},\, j\in F_m).
\end{aligned}
\tag{27}
$$

---

and

$$H(Y_j,\, j\in F_m\cap\Psi|(Y_j,\, j\in F_m\backslash\Psi),(Z_l,\, l\in V_m))=0. \tag{34}$$

By (6), $(Y_j,\, j\in F_m\backslash\Psi)$ and $(Y_j,\, j\in F_m\cap\Psi)$ are independent, so

$$I(Y_j,\, j\in F_m\backslash\Psi;\, Y_j,\, j\in F_m\cap\Psi)=0. \tag{35}$$

From the diagram, we see that this implies

$$I(Y_j,\, j\in F_m\backslash\Psi;\, Y_j,\, j\in F_m\cap\Psi;\, Z_l,\, l\in V_m)=0 \tag{36}$$

(cf. the definition of the mutual information among more than two random variables in [8]). We mark each of these atoms with measure zero by an asterisk. Then we see immediately

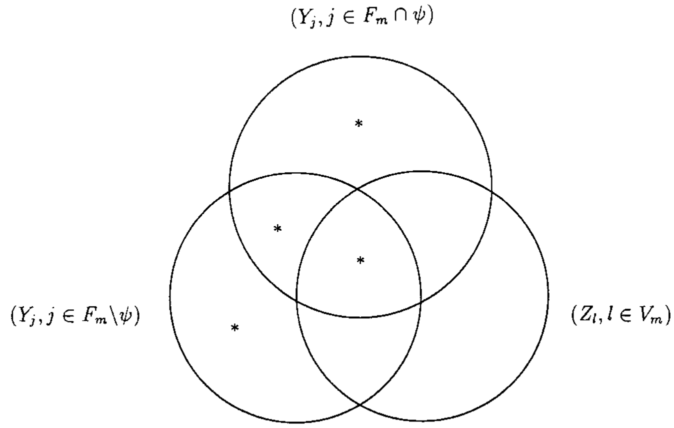

Fig. 2. The information diagram for $(Y_j,\, j \in F_m \backslash \Psi)$, $(Y_j,\, j \in F_m \cap \Psi)$, and $(Z_l,\, l \in V_m)$.

that

$$I(Y_j,\, j \in F_m \backslash \Psi;\, (Y_j,\, j \in F_m \cap \Psi),\, (Z_l,\, l \in V_m)) = H(Y_j,\, j \in F_m \backslash \Psi). \quad (37)$$

Now for $(i_j,\, j \in F_m) \in \Lambda_\Psi$, by Lemma 1 and (37)

$$\begin{aligned} P\{((\underline{\omega}_j(i_j),\, j \in F_m),\, (\boldsymbol{z}_l(t_l),\, l \in V_m)) &\in T^n_{[(Y_j,\, j \in F_m),\, (Z_l,\, l \in V_m)]\delta}\} \\ &\le 2^{-n(I((Y_j,\, j \in F_m \backslash \Psi);\, (Y_j,\, j \in F_m \cap \Psi),\, (Z_l,\, l \in V_m)) - O(\delta))} \\ &= 2^{-n(H(Y_j,\, j \in F_m \backslash \Psi) - O(\delta))}. \end{aligned} \quad (38)$$

From (30), (31), and (38), we have

$$\begin{aligned} P\{E^c_{3m} | E_1 E_{2m}\} &\le \sum_{\Psi} \sum_{(i_j,\, j \in F_m) \in \Lambda_\Psi} 2^{-n(H(Y_j,\, j \in F_m \backslash \Psi) - O(\delta))} \\ &< \sum_{\Psi} 2^{n(H(Y_j,\, j \in F_m \backslash \Psi) - \nu)} \cdot 2^{-n(H(Y_j,\, j \in F_m \backslash \Psi) - O(\delta))} \\ &= (2^{|F_m|} - 1) 2^{-n(\nu - O(\delta))} \le \frac{\epsilon}{3} \end{aligned} \quad (39)$$

when δ is small enough that $O(\delta) < \nu$ and n is sufficiently large. Finally, from (23) and (39), we obtain

$$\Delta_m \le \epsilon, \qquad \text{for all } m \in \mathcal{D}. \quad (40)$$

Hence, from (17) and (40), $\boldsymbol{R} = (R_l,\, l \in \mathcal{E})$ is admissible, which implies $\mathcal{R}_{\text{in}} \subset \mathcal{R}$. □

III. Outer Bound

Let $\mathcal{R}_{\text{out}}$ be the set of all $\boldsymbol{R} = (R_l,\, l \in \mathcal{E})$ such that there exists $\boldsymbol{h} \in \overline{\Gamma}^*_N$ satisfying the following conditions:

$$h_{(Y_j,\, j \in \mathcal{S})} = \sum_{j \in \mathcal{S}} h_{Y_j} \quad (41)$$

$$h_{Z_l | (Y_j,\, j \in U_l)} = 0, \qquad \text{for } l \in \mathcal{E} \quad (42)$$

$$h_{(Y_j,\, j \in F_m) | (Z_l,\, l \in V_m)} = 0, \qquad \text{for } m \in \mathcal{D} \quad (43)$$

$$h_{Y_j} \ge H(X_j), \qquad \text{for } j \in \mathcal{S} \quad (44)$$

$$R_l \ge h_{Z_l}, \qquad \text{for } l \in \mathcal{E}. \quad (45)$$

Theorem 2: $\mathcal{R} \subset \mathcal{R}_{\text{out}}$

We note that the definition of $\mathcal{R}'$ (recall that $\mathcal{R}_{\text{in}} = \overline{\mathcal{R}'}$) is very similar to the definition of $\mathcal{R}_{\text{out}}$ except that

1) Γ^*_N is replaced by $\overline{\Gamma}^*_N$.
2) The inequality in (4) is strict while the inequality in (44) is nonstrict.

It is clear that $\mathcal{R}_{\text{in}}$ is a subset of $\mathcal{R}_{\text{out}}$. As a consequence of the above discrepancies, there is a gap between $\mathcal{R}_{\text{in}}$ and $\mathcal{R}_{\text{out}}$. It is not apparent that the gap between the two regions has zero measure in general.

Proof of Theorem 2: Let $\boldsymbol{R} = (R_l,\, l \in \mathcal{E})$ be admissible. Then for any $\epsilon > 0$, there exists for sufficiently large n an $(n,\, (\eta_l,\, l \in \mathcal{E}),\, (\Delta_m,\, m \in \mathcal{D}))$ code such that

$$n^{-1} \log \eta_l \le R_l + \epsilon, \qquad \text{for all } l \in \mathcal{E} \quad (46)$$

and

$$\Delta_m \le \epsilon, \qquad \text{for all } m \in \mathcal{D}. \quad (47)$$

We first consider such a code for a fixed ϵ.

$$\begin{aligned} H(\boldsymbol{X}_j,\, j \in F_m | T_l,\, l \in V_m) &\le^{1)} H(\boldsymbol{X}_j,\, j \in F_m | \hat{\boldsymbol{X}}_j,\, j \in F_m) \\ &\le^{2)} n(h_b(\Delta_m) + \Delta_m \log |\mathcal{X}_j|) \\ &\le^{3)} n(h_b(\epsilon) + \epsilon \log |\mathcal{X}_j|) \\ &= nO(\epsilon). \end{aligned} \quad (48)$$

In the above, 1) follows because $(\hat{\boldsymbol{X}}_j,\, j \in F_m)$ is a function of $(T_l,\, l \in V_m)$, 2) is the vector version of Fano's inequality where $h_b(\cdot)$ denotes the binary entropy function, and 3) follows from (47). From (46), we also have

$$n(R_l + \epsilon) \ge \log \eta_l \ge H(T_l) \quad (49)$$

for all $l \in \mathcal{E}$. Thus for this code, we have

$$H(\boldsymbol{X}_j,\, j \in \mathcal{S}) = \sum_{j \in \mathcal{S}} H(\boldsymbol{X}_j) \quad (50)$$

$$H(T_l | \boldsymbol{X}_j,\, j \in U_l) = 0, \qquad \text{for } l \in \mathcal{E} \quad (51)$$

$$H(\boldsymbol{X}_j,\, j \in F_m | T_l,\, l \in V_m) \le nO(\epsilon), \qquad \text{for } m \in \mathcal{D} \quad (52)$$

$$H(\boldsymbol{X}_j) \ge nH(X_j), \qquad \text{for } j \in \mathcal{S} \quad (53)$$

$$n(R_l + \epsilon) \ge H(T_l), \qquad \text{for } l \in \mathcal{E}. \quad (54)$$

The inequality in (53) is in fact an equality, which follows from the i.i.d. assumption of the source S_j. We note the one-to-one correspondence between (50)–(54) and (41)–(45). By letting $Y_j = \boldsymbol{X}_j$ and $Z_l = T_l$, we see that there exists $\boldsymbol{h} \in \Gamma^*_N$ such that

$$h_{(Y_j,\, j \in \mathcal{S})} = \sum_{j \in \mathcal{S}} h_{Y_j} \quad (55)$$

$$h_{Z_l | (Y_j,\, j \in U_l)} = 0, \qquad \text{for } l \in \mathcal{E} \quad (56)$$

$$h_{(Y_j,\, j \in F_m) | (Z_l,\, l \in V_m)} \le nO(\epsilon), \qquad \text{for } m \in \mathcal{D} \quad (57)$$

$$h_{Y_j} \ge nH(X_j), \qquad \text{for } j \in \mathcal{S} \quad (58)$$

$$n(R_l + \epsilon) \ge h_{Z_l}, \qquad \text{for } l \in \mathcal{E}. \quad (59)$$

It was shown in the recent work of Zhang and Yeung [11] that $\overline{\Gamma}^*_N$ is a convex cone. Therefore, if $\boldsymbol{h} \in \Gamma^*_N$, then $n^{-1}\boldsymbol{h} \in \overline{\Gamma}^*_N$.

Dividing (55)–(59) by n and replacing $n^{-1}\boldsymbol{h}$ by $\boldsymbol{h}$, we see that there exists $\boldsymbol{h} \in \overline{\Gamma}_N^*$ such that

$$h_{(Y_j,\, j\in\mathcal{S})} = \sum_{j\in\mathcal{S}} h_{Y_j} \tag{60}$$

$$h_{Z_l|(Y_j,\, j\in U_l)} = 0, \qquad \text{for } l \in \mathcal{E} \tag{61}$$

$$h_{(Y_j,\, j\in F_m)|(Z_l,\, l\in V_m)} \le O(\epsilon), \qquad \text{for } m \in \mathcal{D} \tag{62}$$

$$h_{Y_j} \ge H(X_j), \qquad \text{for } j \in \mathcal{S} \tag{63}$$

$$R_l + \epsilon \ge h_{Z_l}, \qquad \text{for } l \in \mathcal{E}. \tag{64}$$

We then let $\epsilon \to 0$ to conclude that there exists $\boldsymbol{h} \in \overline{\Gamma}_N^*$ which satisfies (41)–(45), completing the proof.

IV. Geometrical Interpretations of $\mathcal{R}_{\rm in}$ and $\mathcal{R}_{\rm out}$

In Sections II and III, $\mathcal{R}_{\rm in}$ and $\mathcal{R}_{\rm out}$ are specified in a way which facilitates analysis. In this section, we will present geometrical interpretations of these regions which give further insight.

For a set $A \subset \mathbb{R}^{|\mathcal{E}|}$, define

$$\Upsilon(A) = \{\boldsymbol{r} \in \mathbb{R}^{|\mathcal{E}|}\colon \ \ \boldsymbol{r} \ge \boldsymbol{r}' \text{ for some } \boldsymbol{r}' \in A\}$$

where we write $\boldsymbol{r} \ge \boldsymbol{r}'$ to mean $\boldsymbol{r}$ greater than or equal to $\boldsymbol{r}'$ componentwise. For a set $B \subset \mathbb{R}^{2^{|N|}-1}$, define $\text{proj}_{(h_{Z_l},\, l\in\mathcal{E})}(B)$ to be the projection of the set B on the coordinates h_{Z_l}, $l \in \mathcal{E}$. Define the sets

$$C_1 = \left\{\boldsymbol{h} \in \mathbb{R}^{2^{|N|}-1}\colon h_{(Y_j,\, j\in\mathcal{S})} = \sum_{j\in\mathcal{S}} h_{Y_j}\right\} \tag{65}$$

$$C_2 = \{\boldsymbol{h} \in \mathbb{R}^{2^{|N|}-1}\colon h_{Z_l|(Y_j,\, j\in U_l)} = 0 \text{ for } l \in \mathcal{E}\} \tag{66}$$

$$C_3 = \{\boldsymbol{h} \in \mathbb{R}^{2^{|N|}-1}\colon h_{(Y_j,\, j\in F_m)|(Z_l,\, l\in V_m)} = 0 \text{ for } m \in \mathcal{D}\} \tag{67}$$

$$C_4 = \{\boldsymbol{h} \in \mathbb{R}^{2^{|N|}-1}\colon h_{Y_j} > H(X_j) \text{ for } j \in \mathcal{S}\}. \tag{68}$$

Note that the independence relation in (1) is interpreted as the hyperplane C_1 in $\mathbb{R}^{2^{|N|}-1}$. Similarly, the Markov conditions and the functional dependencies in (2) and (3) are interpreted as the hyperplanes C_2 and C_3, respectively. Then we see that

$$\mathcal{R}_{\rm in} = \overline{\Upsilon(\text{proj}_{(h_{Z_l},\, l\in\mathcal{E})}(\Gamma_N^* \cap C_1 \cap C_2 \cap C_3 \cap C_4))}. \tag{69}$$

Similarly, we see that

$$\mathcal{R}_{\rm out} = \Upsilon(\text{proj}_{(h_{Z_l},\, l\in\mathcal{E})}(\overline{\Gamma}_N^* \cap C_1 \cap C_2 \cap C_3 \cap \overline{C_4})). \tag{70}$$

The bounds $\mathcal{R}_{\rm in}$ and $\mathcal{R}_{\rm out}$ are specified in terms of Γ_N^* and $\overline{\Gamma}_N^*$, respectively. So far, there exists no full characterization of either Γ_N^* or $\overline{\Gamma}_N^*$ [11], [12]. Therefore, although these bounds are in single-letter form, they cannot be evaluated explicitly. Nevertheless, in view of the geometrical interpretation of $\mathcal{R}_{\rm out}$, one can easily obtain an outer bound on $\mathcal{R}_{\rm out}$ which can be evaluated explicitly. This will be described in the next section.

V. The LP Bound

Following [10], we define Γ_N to be the set of $\boldsymbol{h} \in \mathbb{R}^{2^{|N|}-1}$ such that for any nonempty $G, G', G'' \in Q(N)$ (recall that $Q(N) = 2^N \backslash \phi$)

$$h_G \ge 0 \tag{71}$$

$$h_{G|G'} \ge 0 \tag{72}$$

$$i_{G;G'} \ge 0 \tag{73}$$

$$i_{G;G'|G''} \ge 0. \tag{74}$$

These inequalities are referred to as the *basic inequalities*, which are linear inequalities in $\mathbb{R}^{2^{|N|}-1}$. It is easy to see that $\Gamma_N^* \subset \Gamma_N$. Since Γ_N is a closed set, $\overline{\Gamma}_N^* \subset \Gamma_N$. Therefore, upon replacing $\overline{\Gamma}_N^*$ in the definition of $\mathcal{R}_{\rm out}$ by Γ_N, we immediately obtain an outer bound on $\mathcal{R}_{\rm out}$. We call this outer bound the LP bound (for *linear programming* bound) and denote it by $\mathcal{R}_{\rm LP}$. Thus

$$\mathcal{R}_{\rm LP} = \Upsilon(\text{proj}_{(h_{Z_l},\, l\in\mathcal{E})}(\Gamma_N \cap C_1 \cap C_2 \cap C_3 \cap \overline{C_4})). \tag{75}$$

The geometrical interpretation of $\mathcal{R}_{\rm out}$ in the last section can also be applied to $\mathcal{R}_{\rm LP}$, so it is not difficult to see that $\mathcal{R}_{\rm LP}$ can be evaluated explicitly.

The multilevel diversity coding problem (with independent data streams) studied in [9] is a special case of the problem we study in the current paper. For such a problem with K levels ($K \ge 2$), there are K sources, $S_1, \cdots, S_K$, whose importance decreases in the order $S_1 > \cdots > S_K$. Each encoder has access to all the sources. Further, Decoder m belongs to Level $L(m)$, where $1 \le L(m) \le K$, and it reconstructs $S_1, \cdots, S_{L(m)}$, the most important $L(m)$ sources. The problem is specified by

$$\mathcal{S} = \{1, \cdots, K\}$$

$$\mathcal{A} = \mathcal{S} \times \mathcal{E}$$

$$L : \mathcal{D} \to \{1, \cdots, K\}$$

and for all $m \in \mathcal{D}$

$$F_m = \{1, \cdots, L(m)\}.$$

So far, the coding rate region of the problem we study in this paper has been determined for certain special cases [5], [7], [9], [13], and these are all special cases of multilevel diversity coding. The coding rate region for each of these cases has the form

$$\left\{(R_1, \cdots, R_K) : \sum_{l=1}^{K} A_l R_l \ge \sum_{j=1}^{K} f_j(\boldsymbol{A}) H(X_j) \quad \text{for all } \boldsymbol{A} \subset \mathcal{C}\right\} \tag{76}$$

where $\boldsymbol{A} = [A_1, \cdots, A_K]$, $\mathcal{C}$ is some subset of $(\mathbb{R}^+)^{2^{|N|}-1}$, and $f_j(\boldsymbol{A})$, $1 \le j \le K$, are some nonnegative functions depending only on $\boldsymbol{A}$. Further, the necessity of the coding rate region (i.e., the converse coding theorem) for all these cases can be proved by invoking the basic inequalities.

We claim that the LP bound is tight for all the special cases for which the coding rate region has the form in (76) and the converse coding theorem is a consequence of the basic inequalities. Among these is the K-encoder symmetrical multilevel diversity coding (SMDC) problem recently studied by the authors [13]. In the rest of the section, we will prove the tightness of the LP bound for the SMDC problem; the techniques involved in proving the other cases are exactly the same. We further conjecture that the LP bound is tight for any special case as long as the converse coding theorem is a consequence of the basic inequalities, but we cannot prove it.

In the following we will use v_l to denote the lth component of a vector $\boldsymbol{v} \in \{0, 1\}^K$. In the SMDC problem

$$\begin{aligned} \mathcal{S} &= \{1, \cdots, K\} \\ \mathcal{E} &= \{1, \cdots, K\} \\ \mathcal{D} &= \{\boldsymbol{v} \in \{0, 1\}^K : |\boldsymbol{v}| \geq 1\} \\ \mathcal{A} &= \mathcal{S} \times \mathcal{E} \\ \mathcal{B} &= \{(l, \boldsymbol{v}) : v_l = 1\} \end{aligned}$$

and

$$F(\boldsymbol{v}) = \{1, \cdots, |\boldsymbol{v}|\}.$$

With the above specifications, we implicitly have $L(\boldsymbol{v}) = |\boldsymbol{v}|$. From [13], the coding rate region for the SMDC problem is given by

$$\mathcal{R}_{\text{SMDC}} = \left\{(R_1, \cdots, R_K) : \sum_{l=1}^{K} A_l R_l \geq \sum_{j=1}^{K} f_j(\boldsymbol{A}) H(X_j) \quad \text{for all } \boldsymbol{A} \geq 0 \right\}.$$

Comparing with (76), we see that $\mathcal{C} = (\mathbb{R}^+)^{2^{|N|}-1}$ for the SMDC problem. Note that the definition of $\mathcal{R}_{\text{SMDC}}$ implies $R_l \geq 0$ for $1 \leq l \leq K$. This can be seen by letting $\boldsymbol{A}$ be the K-vector whose components are equal to 0 except that the lth component is equal to 1, so that R_l is lower-bounded by 0.

Lemma 2: $\mathcal{R}_{\text{SMDC}} = \Upsilon(\mathcal{R}_{\text{SMDC}})$.

Proof: Trivial.

In the proof for the necessity of $\mathcal{R}_{\text{SMDC}}$ in [13], the authors consider a code with blocklength n. Careful examination of the proof for the zero-error case reveals that the same region can be obtained by considering a code with $n = 1$. (This is due to the assumption that the sources are i.i.d.) In the following, we use $Z_{\boldsymbol{v}}$ to denote the collection of random variables Z_l such that $v_l = 1$ and $1 \leq l \leq K$. Basically, it is proved in [13] that

Proposition 1: For any discrete random variables Y_j, $1 \leq j \leq K$ and $Z_{\boldsymbol{v}}$, $\boldsymbol{v} \in \{0, 1\}^K$, if

1) $H(Y_1, \cdots, Y_K) = \sum_{j=1}^{K} H(Y_j)$;
2) $H(Y_1, \cdots, Y_l | Z_{\boldsymbol{v}}) = 0$ for all $\boldsymbol{v}$ such that $|\boldsymbol{v}| = l$, then

$$\sum_{l=1}^{K} A_l H(Z_l) \geq \sum_{j=1}^{K} f_j(\boldsymbol{A}) H(Y_j) \tag{77}$$

for all $\boldsymbol{A} \geq 0$.

Let us now state another proposition.

Proposition 2: For any discrete random variables Y_j, $1 \leq j \leq K$ and $Z_{\boldsymbol{v}}$, $\boldsymbol{v} \in \{0, 1\}^K$, if

1) $H(Y_1, \cdots, Y_K) = \sum_{j=1}^{K} H(Y_j)$;
2) $H(Y_1, \cdots, Y_l | Z_{\boldsymbol{v}}) = 0$ for all $\boldsymbol{v}$ such that $|\boldsymbol{v}| = l$;
3) $H(Y_j) \geq H(X_j)$ for $1 \leq j \leq K$.

then

$$\sum_{l=1}^{K} A_l H(Z_l) \geq \sum_{j=1}^{K} f_j(\boldsymbol{A}) H(X_j) \tag{78}$$

for all $\boldsymbol{A} \geq 0$

Note that in the above proposition, the quantities $H(X_j)$, $1 \leq j \leq K$ are interpreted as constants.

Lemma 3: Propositions 1 and 2 are equivalent.

Proof: We first show that Proposition 1 implies Proposition 2. Proposition 1 says that 1) and 2) implies (77). Together with 3), we have

$$\sum_{l=1}^{K} A_l H(Z_l) \geq \sum_{j=1}^{K} f_j(\boldsymbol{A}) H(Y_j) \geq \sum_{j=1}^{K} f_j(\boldsymbol{A}) H(X_j)$$

since $f_j(\boldsymbol{A}) \geq 0$, which proves Proposition 2.

We now show that Proposition 2 implies Proposition 1. Suppose 1) and 2) are satisfied. We let $H(X_j) = H(Y_j)$ for $1 \leq j \leq K$ so that 3) is satisfied with equality. Then by Proposition 2, we have

$$\sum_{l=1}^{K} A_l H(Z_l) \geq \sum_{j=1}^{K} f_j(\boldsymbol{A}) H(X_j) = \sum_{j=1}^{K} f_j(\boldsymbol{A}) H(Y_j).$$

Thus we see that 1) and 2) imply (77), which proves Proposition 1. ☐

The constraints 1), 2), and 3) in Proposition 2 correspond to the sets C_1, C_3, and $\overline{C_4}$, respectively, defined in Section IV, which are subsets of $\mathbb{R}^{2^{|N|}-1}$. Since the Proof of Proposition 1 in [13] involves only the basic inequalities (i.e., Proposition 1, and hence Proposition 2, are implied by the basic inequalities), using the results in [10], we see that

$$\Gamma_N \cap C_1 \cap C_3 \cap \overline{C_4} \subset \left\{\boldsymbol{h} \in \mathbb{R}^{2^{|N|}-1} : \sum_{l=1}^{K} A_l h_{Z_l} \geq \sum_{j=1}^{K} f_j(\boldsymbol{A}) H(X_j) \quad \text{for all } \boldsymbol{A} \geq 0 \right\}. \tag{79}$$

So

$$\Gamma_N \cap C_1 \cap C_2 \cap C_3 \cap \overline{C_4} \subset \left\{\boldsymbol{h} \in \mathbb{R}^{2^{|N|}-1} : \sum_{l=1}^{K} A_l h_{Z_l} \geq \sum_{j=1}^{K} f_j(\boldsymbol{A}) H(X_j) \quad \text{for all } \boldsymbol{A} \geq 0 \right\}. \tag{80}$$

By projecting onto the coordinates $h_{Z_l}, l \in \mathcal{E}$, we have

$$\begin{aligned}
&\mathrm{proj}_{(h_{Z_l}, l\in\mathcal{E})}(\Gamma_N \cap C_1 \cap C_2 \cap C_3 \cap \overline{C_4}) \\
&\quad \subset \left\{ (h_{Z_1}, \cdots, h_{Z_K}) \colon \sum_{l=1}^{K} A_l h_{Z_l} \ge \sum_{j=1}^{K} f_j(\boldsymbol{A}) H(X_j) \right. \\
&\qquad\qquad \left. \text{for all } \boldsymbol{A} \ge 0 \right\} \\
&\quad = \left\{ (R_1, \cdots, R_K) \colon \sum_{l=1}^{K} A_l R_l \ge \sum_{j=1}^{K} f_j(\boldsymbol{A}) H(X_j) \right. \\
&\qquad\qquad \left. \text{for all } \boldsymbol{A} \ge 0 \right\} = \mathcal{R}_{\mathrm{SMDC}}.
\end{aligned}$$

Then

$$\begin{aligned}
\mathcal{R}_{\mathrm{LP}} &= \Upsilon(\mathrm{proj}_{(h_{Z_l}, l\in\mathcal{E})}(\Gamma_N \cap C_1 \cap C_2 \cap C_3 \cap \overline{C_4})) \\
&\subset \Upsilon(\mathcal{R}_{\mathrm{SMDC}}) = \mathcal{R}_{\mathrm{SMDC}}
\end{aligned}$$

where the last step follows from Lemma 2. Since $\mathcal{R}_{\mathrm{LP}}$ is an outer bound on $\mathcal{R}_{\mathrm{SMDC}}$, we have

$$\mathcal{R}_{\mathrm{SMDC}} \subset \mathcal{R}_{\mathrm{LP}} \subset \mathcal{R}_{\mathrm{SMDC}}$$

which implies $\mathcal{R}_{\mathrm{LP}} = \mathcal{R}_{\mathrm{SMDC}}$. Thus $\mathcal{R}_{\mathrm{LP}}$ is tight. Compared with $\mathcal{R}_{\mathrm{SMDC}}$, $\mathcal{R}_{\mathrm{LP}}$ has the advantage that it is more explicit and can easily be evaluated.

VI. Conclusion

In this paper, we introduce a new multiterminal source coding problem called the distributed source coding problem. Our model is more general than all previously reported models on multilevel diversity coding [5], [7], [9], [13]. We mention that an even more general model has been formulated in [1].

We have obtained an inner bound $\mathcal{R}_{\mathrm{in}}$ and an outer bound $\mathcal{R}_{\mathrm{out}}$ on the coding rate region. Both $\mathcal{R}_{\mathrm{in}}$ and $\mathcal{R}_{\mathrm{out}}$ are implicit in the sense that they are specified in terms of Γ_N^* and $\overline{\Gamma}_N^*$, respectively, which are fundamental regions in the entropy space yet to be determined. Our work is an application of the results in [10] to information theory problems.

We have also obtained an explicit outer bound $\mathcal{R}_{\mathrm{LP}}$ for the coding rate region. We have shown that this bound is tight for a class of special cases, including all those for which the coding rate region is known.

$\mathcal{R}_{\mathrm{LP}}$ would be tight if $\Gamma_N^* = \Gamma_N$, but this has recently been disproved by the authors [11], [12]. Nevertheless, we believe that it is tight for most cases. A problem for future research is to determine the conditions for which $\mathcal{R}_{\mathrm{LP}}$ is tight.

We point out that the random code we have constructed in Section III has an arbitrarily small probability of error which does not go away even if we are allowed to compress the sources first by variable rate codes. This characteristic is undesirable for many applications. A challenging problem for future research is to construct simple zero-error algebraic codes for distributed source coding.

Acknowledgment

The authors wish to thank Reviewer A for the detailed comments on the manuscript.

References

[1] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, "Network information flow theory," *IEEE Trans. Inform. Theory*, to be published.
[2] T. Berger, "Multiterminal source coding," in *The Information Theory Approach to Communications*, G. Longo, Ed. New York: Springer-Verlag, 1977.
[3] T. M. Cover and J. A. Thomas, *Elements of Information Theory.* New York: Wiley, 1991.
[4] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems.* New York: Academic, 1981.
[5] K. P. Hau, "Multilevel diversity coding with independent data streams," M.Phil. thesis, The Chinese University of Hong Kong, May 1995.
[6] F. Matúš and M. Studený, "Conditional independences among four random variables I," *Combin., Prob. Comput.*, vol. 4, pp. 269–278, 1995.
[7] J. R. Roche, R. W. Yeung, and K. P. Hau, "Symmetrical multilevel diversity coding," *IEEE Trans. Inform. Theory*, vol. 43, pp. 1059–1064, May 1997.
[8] R. W. Yeung, "A new outlook on Shannon's information measures," *IEEE Trans. Inform. Theory*, vol. 37, pp. 466–474, May 1991.
[9] ——, "Multilevel diversity coding with distortion," *IEEE Trans. Inform. Theory*, vol. 41, pp. 412–422, May, 1995.
[10] ——, "A framework for linear information inequalities," *IEEE Trans. Inform. Theory*, vol. 43, pp. 1924–1934, Nov. 1997.
[11] Z. Zhang and R. W. Yeung, "A non-Shannon-type conditional inequality of information quantities," *IEEE Trans. Inform. Theory*, vol. 43, pp. 1982–1986, Nov. 1997.
[12] ——, "On characterization of entropy function via information inequalities," *IEEE Trans. Inform. Theory*, vol. 44, pp. 1440–1452, July 1998.
[13] R. W. Yeung and Z. Zhang, "On symmetrical multilevel diversity coding," *IEEE Trans. Inform. Theory*, vol. 45, pp. 609–621, Mar. 1999.