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 $\Gamma_{\mathcal{S}}^e$ and $\Gamma_{\mathcal{N}}^e$, respectively, which are fundamental regions in the entropy space recently defined by Yeung. So far, there has not been a full characterization of $\Gamma_{\mathcal{N}}^e$, 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 $\pm 70^\circ$ 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 $i$th component of a vector $\mathbf{x}$ 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 $\mathcal{X}$ to denote the 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) $\mathcal{F}_{mn} \subset 2^{\mathcal{S} \setminus \mathcal{B}}$, $m \in \mathcal{D}$, which specify the reconstruction requirements of the decoders.

The $j$th 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 $|X_j| < \infty$. The $l$th encoder
is denoted by $E_l$, $l \in \mathcal{E}$, and the $m$th decoder is denoted by $D_m$, $m \in \mathcal{D}$. The sets $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{F}_m$, $m \in \mathcal{D}$ 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 $S_j$, $j \in \mathcal{F}_m$.

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

$$\mathcal{S} = \{1, 2, 3, 4\}, \quad \mathcal{E} = \{1, 2, 3\}, \quad \mathcal{D} = \{1, 2, 3, 4\}$$

$$\mathcal{A} = \{(1, 1), (1, 2), (2, 1), (2, 2), (2, 3), (3, 3), (4, 2), (4, 3)\}$$

$$\mathcal{B} = \{(1, 1), (1, 3), (2, 1), (2, 2), (3, 2), (3, 3), (3, 4)\}$$

$$\mathcal{F}_1 = \{2\}, \quad \mathcal{F}_2 = \{1, 3\}, \quad \mathcal{F}_3 = \{1, 2, 3\}, \quad \mathcal{F}_4 = \{3, 4\}.$$

To facilitate the description of our model, we define $U_l = \{j \in \mathcal{S} : (j, l) \in \mathcal{A}\}$, $l \in \mathcal{E}$, and $V_m = \{l \in \mathcal{E} : (l, m) \in \mathcal{B}\}$, $m \in \mathcal{D}$. $U_l$ 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

$$d_m : \left( \prod_{j \in \mathcal{F}_m} X_j \right) \times \left( \prod_{j \in \mathcal{F}_m} X_j \right) \to \{0, 1\}$$

be the Hamming distortion measure $m \in \mathcal{D}$; i.e., for any $x$ and $x'$ in $\left( \prod_{j \in \mathcal{F}_m} X_j \right) \times \left( \prod_{j \in \mathcal{F}_m} X_j \right)$,

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

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

$$T_l : \left( \prod_{j \in U_l} X_j^\eta_l \right) \to \{0, 1, \ldots, \eta_l - 1\}, \quad l \in \mathcal{E}$$

$$W_m : \left( \prod_{l \in V_m} \{0, 1, \ldots, \eta_l - 1\} \right) \to \left( \prod_{j \in \mathcal{F}_m} X_j^\eta_l \right), \quad m \in \mathcal{D}$$

and

$$\Delta_m = n^{-1} \sum_{k=1}^n d_m((X_{jk}, j \in \mathcal{F}_m), (X_{jk}, j \in \mathcal{F}_m)), \quad m \in \mathcal{D}$$

where

$$((X_{j1}, \ldots, X_{jn}), j \in \mathcal{F}_m) = W_m(T_l(X_j, j \in U_l), l \in V_m).$$

An $(\mathcal{E}, 1)$-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

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

and

$$\Delta_m \leq \epsilon, \quad \text{for all } m \in \mathcal{D}.$$
with the source $S_j$, and $Z_l$ is an auxiliary random variable associated with $T_l(X_j, j \in U_l)$, 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 $R = (R_l, l \in \mathcal{E})$ such that there exists $h \in \Gamma_N^*$ which satisfies the following conditions:

\begin{align}
&h(Y_j, j \in \mathcal{S}) = \sum_{j \in \mathcal{S}} h(Y_j) \quad (1) \\
&h(Z_l|Y_j, j \in U_l) = 0, \quad \text{for } l \in \mathcal{E} \quad (2) \\
&h(Y_j, j \in F_m|Z_l, l \in V_m) = 0, \quad \text{for } m \in \mathcal{D} \quad (3) \\
&h(Y_j) > H(X_j), \quad \text{for } j \in \mathcal{S} \quad (4) \\
&R_l \geq h(Z_l), \quad \text{for } l \in \mathcal{E}. \quad (5)
\end{align}

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

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

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

\begin{align}
H(Y_j, j \in \mathcal{S}) &= \sum_{j \in \mathcal{S}} H(Y_j) \quad (6) \\
H(Z_l|Y_j, j \in U_l) &= 0, \quad \text{for } l \in \mathcal{E} \quad (7) \\
H(Y_j, j \in F_m|Z_l, l \in V_m) &= 0, \quad \text{for } m \in \mathcal{D} \quad (8) \\
H(Y_j) &> H(X_j), \quad \text{for } j \in \mathcal{S} \quad (9) \\
R_l &\geq h(Z_l), \quad \text{for } l \in \mathcal{E}. \quad (10)
\end{align}

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 $\Gamma_N^*$ enables us to study all cases on the same footing. In particular, if $\Gamma_N^*$ is an explicit inner bound on $\mathcal{R}_{\text{in}}$, upon replacing $\Gamma_N^*$ 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 $\Gamma_N^*$ 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 $\mathcal{R}_{\text{in}}$ 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 $x \in \mathcal{X}^n$ is $\delta$-typical with respect to a distribution $p(x)$ if for all $x \in \mathcal{X}$

\[
\left| \frac{1}{n} N(x|z) - p(x) \right| < \frac{\delta}{|\mathcal{X}|}
\]

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

\[
\left| \frac{1}{n} N(x, y|z, \mathcal{X}) - p(x, y) \right| < \frac{\delta}{|\mathcal{X}| |\mathcal{Y}|}
\]

In the sequel, the $\delta$-typical notations in [4] will be adopted.

**Lemma 1:** Let $(x, y)$ be $n$-vectors drawn according to

\[
p(x, y) = \prod_{i=1}^n p(x_i, y_i).
\]

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

\[
P\{ (\tilde{X}, \tilde{Y}) \in T_{\mathcal{X}Y}^n \} \leq 2^{-n(T(X;Y)-O(\delta))}
\]

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 $\delta, \epsilon$ be small positive quantities and $n$ be a large integer to be specified later. Let $h \in \mathcal{R}'$. 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 $y = (y_1, \ldots, y_n)$. For each $j \in \mathcal{S}$, independently generate $\lfloor 2^{H(Y_j) - \nu} \rfloor$ vectors in $\mathcal{Y}_j^\nu$ according to the distribution

\[
p(y) = \prod_{k=1}^n p_j(y_k) \quad (12)
\]

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

\[
\omega_j(\iota_j), \quad 0 \leq \iota_j \leq |\Omega_j| - 1.
\]

2. Let $\zeta_j = \lfloor 2^{H(Y_j) + O(\delta)} \rfloor$ and index the vectors in $T_{\mathcal{X}Y}^n$ by the set $\{1, 2, \ldots, \zeta_j\}$. Since $H(Y_j) > H(X_j)$ by (4), by taking $\delta$ and $\nu$ small enough and $n$ sufficiently large, we can make

\[
\zeta_j = \lfloor 2^{H(Y_j) + O(\delta)} \rfloor < \lfloor 2^{n(H(Y_j) - \nu)} \rfloor - 1 = |\Omega_j| - 1.
\]

Define a mapping

\[
\rho_j : \mathcal{X}_j^\nu \to \{0, 1, \ldots, \zeta_j\}
\]

as follows. For any $x \in \mathcal{X}_j^\nu$, if $x \notin T_{\mathcal{X}Y}^n$, then $\rho_j(x) = 0$; if $x \in T_{\mathcal{X}Y}^n$, then $\rho_j(x)$ is equal to the index of $x$ in $T_{\mathcal{X}Y}^n$ defined above. We note that for $x \notin T_{\mathcal{X}Y}^n, \omega_j(\rho_j(x))$ is a constant vector in $\Omega_j$, and...
for \( x \in \mathcal{X}^n_{j} \), \( \omega_j(\rho_j(x)) \) are distinct vectors in \( \Omega_j \) (for all \( x \in \mathcal{X}^n_{j} \), \( \rho_j(x) \leq |\Omega_j| - 1 \) by (13), so \( \omega_j(\rho_j(x)) \) is properly defined).

3. Let \( \rho \) be the conditional distribution of \( Z_t \) given \( (Y_j, j \in U_t) \). Let \( \eta \) \( \equiv \sum_{i=1}^{|\Omega_j|} \eta \) and index the vectors in \( \mathcal{X}^n_{j} \) by the set \( \{0, 1, \ldots, \eta - 1\} \). For each \( l \in \mathcal{E} \), define a mapping

\[
\psi_l : \prod_{j \in \mathcal{E}_l} \Omega_j \to \{0, 1, \ldots, \eta - 1\}
\]

randomly as follows. For each \( (y_j, j \in U_t) \in \prod_{j \in \mathcal{E}_l} \Omega_j \), send it through a discrete memoryless channel with transition probability \( \rho_j \) by using the channel \( n \) times. If the received vector is not in \( \mathcal{X}^n_{j} \), then \( \psi_l(y_j, j \in U_t) = 0 \), otherwise, \( \psi_l(y_j, j \in U_t) \) is equal to the index of the received vector in \( \mathcal{X}^n_{j} \) defined above.

4. The encoding function

\[
T_l : \prod_{j \in \mathcal{E}_l} \mathcal{X}^n_{j} \to \{0, 1, \ldots, \eta - 1\}
\]

is defined as follows. For each

\[
(y_j, j \in U_t) \in \prod_{j \in \mathcal{E}_l} \mathcal{X}^n_{j}
\]

if \( \rho_j(x_j) = 0 \) for some \( j \in U_t \), then \( T_l(x_j, j \in U_t) = 0 \), otherwise,

\[
T_l(x_j, j \in U_t) = \psi_l(\omega_j(\rho_j(x_j)), j \in U_t).
\]

The decoding scheme is described in the following steps:

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

\[
\xi_m : \prod_{l \in \mathcal{V}_m} \{0, 1, \ldots, \eta_l - 1\} \to \prod_{j \in \mathcal{F}_m} \{0, 1, \ldots, \eta_j\}
\]

as follows. Let \( z_l(i) \) be the vector in \( \mathcal{X}^n_{j} \) with index \( i \) (cf. step 3 in the encoding scheme). For any

\[
(y_j, j \in U_t) \in \prod_{l \in \mathcal{V}_m} \{0, 1, \ldots, \eta_l - 1\}
\]

if there exists a unique set of indices

\[
(i_j, j \in \mathcal{F}_m) \in \prod_{j \in \mathcal{F}_m} \{0, 1, \ldots, \eta_j\}
\]

such that

\[
(\omega_j(i_j), j \in \mathcal{F}_m), (z_l(i), l \in \mathcal{V}_m)) \in \mathcal{T}^n_{\{Y_j, j \in \mathcal{F}_m\}, \{z_l, l \in \mathcal{V}_m\}}
\]

then

\[
\xi_m(t_l, l \in \mathcal{V}_m) = (i_j, j \in \mathcal{F}_m)
\]

otherwise

\[
\xi_m(t_l, l \in \mathcal{V}_m) = (0, \ldots, 0).
\]

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

\[
\mu_j : \{0, 1, \ldots, \eta_j\} \to \mathcal{X}^n_{j}
\]

by

\[
\mu_j(i) = \begin{cases} \rho_j^{-1}(i), & \text{if } 1 \leq i \leq \eta_j \\ x_{j,0}, & \text{otherwise} \end{cases}
\]

where \( x_{j,0} \) is an arbitrary constant vector in \( \mathcal{X}^n_{j} \) (cf. step 2 of the encoding scheme).

3) The decoding function

\[
W_m : \prod_{l \in \mathcal{V}_m} \{0, 1, \ldots, \eta_l - 1\} \to \prod_{j \in \mathcal{F}_m} \mathcal{X}^n_{j}
\]

is defined as

\[
W_m(t_l, l \in \mathcal{V}_m) = (\mu_j(\xi_m(t_l)), j \in \mathcal{F}_m)
\]

where

\[
(i_j, j \in \mathcal{F}_m) \equiv \xi_m(t_l, l \in \mathcal{V}_m).
\]

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

\[
|\mathcal{T}_{\{Z_t\}}| = 2^{-nH(Z_t) + O(\delta)}
\]

we have

\[
n^{-1} \log n \eta \leq H(Z_t) + O(\delta)
\]

by taking \( \delta \) sufficiently small. Therefore,

\[
n^{-1} \log n \eta \leq R_{t} + \epsilon,
\]

for all \( l \in \mathcal{E} \) by (10). We now introduce the following notations:

\[
Y_j = \omega_j(\rho_j(X_j))
\]

\[
T_l = T_l(X_j, j \in U_t)
\]

\[
Z_l = z_l(T_l)
\]

\[
\hat{Y}_j = \omega_j(i_j).
\]

Define the events

\[
E_1 = \{\rho_j(X_j) \neq 0, j \in S\}
\]

\[
E_2 = \{(Y_j, j \in S), (Z_l, l \in \mathcal{E}) \in \mathcal{T}^n_{\{Y_j, j \in S\}, \{Z_l, l \in \mathcal{E}\}}\}
\]

\[
E_{2m} = \{(Y_j, j \in \mathcal{F}_m), (Z_l, l \in \mathcal{V}_m) \in \mathcal{T}^n_{\{Y_j, j \in \mathcal{F}_m\}, \{Z_l, l \in \mathcal{V}_m\}}\}
\]

\[
E_{3m} = \{i_j = \rho_j(X_j), j \in F_m\}.
\]

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

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

\[
P(E_1) = P(X_j \notin \mathcal{T}^n_{\{Y_j\}}) \leq \frac{\epsilon}{3},
\]

(18)
Let \( p((y_j, j \in S), (z_l, l \in E)) \) be the joint distribution of \(((Y_j, j \in S), (Z_l, l \in E))\). By (6) and (7), we have
\[
p((y_j, j \in S), (z_l, l \in E)) = P\{Y_j = y_j, j \in S\} P\{Z_l = z_l, l \in E\}\}
= \prod_{j \in S} p_j(y_j) \prod_{l \in E} q_l(z_l | y_j, j \in U_l).
\]
(19)

Since \(((Y_j, j \in S), (Z_l, l \in E))\) is drawn i.i.d. according to the above distribution, we see from the weak law of large numbers that for a sufficiently large \( n \)
\[
P\{E_2\} \leq \epsilon/3.
\]
(20)

By noting that \( E_2 \) implies \( E_{2m}, m \in \mathcal{D}, \) we see that for any \( m \in \mathcal{D} \)
\[
P\{E_{2m}\} \leq P\{E_2\} \leq \epsilon/3.
\]
(21)

Further, since \( E_1 \) and \( E_{2m} \) are independent events, we have
\[
P\{E_{2m} | E_1\} = P\{E_{2m}\} \leq \epsilon/3.
\]
(22)

Now for any \( m \in \mathcal{D} \) and sufficiently large \( n \), we have
\[
\Delta_m = E \sum_{k=1}^{n} d_m((X_{jk}, j \in F_m), (X_{jk}, j \in F_m))
\leq E[E_1 P\{E_{2m}\}] + E[E_2 P\{E_{2m}\}]
\leq E[E_1] + E[P\{E_{2m}\}]
\leq E[E_1 E_{2m}] P\{E_{2m} | E_1\}
+ E[E_2 E_{2m}] P\{E_{2m} | E_2\}
+ \epsilon/3
\leq E[E_2 E_{2m}] + 1 \cdot P\{E_{2m} | E_1\} + \epsilon/3
\leq E[E_2 E_{2m}] + P\{E_{2m} | E_2\} + \epsilon/3
\leq E[E_2 E_{2m}] + 1 \cdot P\{E_{2m} | E_2\} + \epsilon/3
\leq E[E_2 E_{2m}] + 1 \cdot P\{E_{2m} | E_2\} + \epsilon/3
\leq E[E_2 E_{2m}] + 1 \cdot P\{E_{2m} | E_2\} + \epsilon/3
\leq E[E_2 E_{2m}] + 1 \cdot P\{E_{2m} | E_2\} + \epsilon/3
\leq E[E_2 E_{2m}] + 1 \cdot P\{E_{2m} | E_2\} + \epsilon/3
\leq E[E_2 E_{2m}] + 1 \cdot P\{E_{2m} | E_2\} + \epsilon/3
\leq E[E_2 E_{2m}] + 1 \cdot P\{E_{2m} | E_2\} + \epsilon/3
\leq E[E_2 E_{2m}] + 1 \cdot P\{E_{2m} | E_2\} + \epsilon/3
\leq E[E_2 E_{2m}] + 1 \cdot P\{E_{2m} | E_2\} + \epsilon/3
\leq E[E_2 E_{2m}] + 1 \cdot P\{E_{2m} | E_2\} + \epsilon/3
\leq E[E_2 E_{2m}] + 1 \cdot P\{E_{2m} | E_2\} + \epsilon/3
\leq E[E_2 E_{2m}] + 1 \cdot P\{E_{2m} | E_2\} + \epsilon/3
\leq E[E_2 E_{2m}] + 1 \cdot P\{E_{2m} | E_2\} + \epsilon/3
\leq E[E_2 E_{2m}] + 1 \cdot P\{E_{2m} | E_2\} + \epsilon/3
\leq E[E_2 E_{2m}] + 1 \cdot P\{E_{2m} | E_2\} + \epsilon/3
\leq E[E_2 E_{2m}] + 1 \cdot P\{E_{2m} | E_2\} + \epsilon/3
\leq E[E_2 E_{2m}] + 1 \cdot P\{E_{2m} | E_2\} + \epsilon/3
\leq E[E_2 E_{2m}] + 1 \cdot P\{E_{2m} | E_2\} + \epsilon/3
\leq E[E_2 E_{2m}] + 1 \cdot P\{E_{2m} | E_2\} + \epsilon/3
\leq E[E_2 E_{2m}] + 1 \cdot P\{E_{2m} | E_2\} + \epsilon/3
\leq E[E_2 E_{2m}] + 1 \cdot P\{E_{2m} | E_2\} + \epsilon/3
\leq E[E_2 E_{2m}] + 1 \cdot P\{E_{2m} | E_2\} + \epsilon/3
\leq E[E_2 E_{2m}] + 1 \cdot P\{E_{2m} | E_2\} + \epsilon/3
\leq E[E_2 E_{2m}] + 1 \cdot P\{E_{2m} | E_2\} + \epsilon/3
\leq E[E_2 E_{2m}] + 1 \cdot P\{E_{2m} | E_2\} + \epsilon/3
\leq E[E_2 E_{2m}] + 1 \cdot P\{E_{2m} | E_2\} + \epsilon/3
\leq E[E_2 E_{2m}] + 1 \cdot P\{E_{2m} | E_2\} + \epsilon/3
\leq E[E_2 E_{2m}] + 1 \cdot P\{E_{2m} | E_2\} + \epsilon/3
\leq E[E_2 E_{2m}] + 1 \cdot P\{E_{2m} | E_2\} + \epsilon/3
\leq E[E_2 E_{2m}] + 1 \cdot P\{E_{2m} | E_2\} + \epsilon/3
\leq E[E_2 E_{2m}] + 1 \cdot P\{E_{2m} | E_2\} + \epsilon/3
\]
\begin{align}
P\{Z_l = z_l, l \in V_m | Y_j = y_j, j \in F_m\} \\
= \sum_{(y_j, j \in F_m)} P\{(Z_l = z_l, l \in V_m), (Y_j = y_j, j \in F_m) | Y_j = y_j, j \in F_m\} \\
= \sum_{(-)} P\{Z_l = z_l, l \in V_m | Y_j = y_j, j \in S\} P\{Y_j = y_j, j \in F_m\} \\
= \sum_{(-)} P\{Z_l = z_l, l \in V_m | Y_j = y_j, j \in S\} P\{Y_j = y_j, j \in F_m\}
\end{align}

\begin{align}
= 1 \cdot \sum_{(-)} P\{Y_j = y_j, j \in F_m\}^{-1} \sum_{(-)} P\{Z_l = z_l, l \in V_m | Y_j = y_j, j \in S\} P\{Y_j = y_j, j \in F_m\} \\
\cdot P\{Y_j = y_j, j \in F_m\} \\
= P\{\cdot\}^{-1} \sum_{(-)} P\{Z_l = z_l, l \in V_m | Y_j = y_j, j \in S\} P\{Y_j = y_j, j \in S\} \\
= P\{\cdot\}^{-1} \sum_{(-)} P\{Y_j = y_j, j \in S\} P\{Z_l = z_l, l \in V_m\}
\end{align}

\begin{align}
= P\{\cdot\}^{-1} \prod_{(y_j, j \in F_m)} \prod_{k=1}^{n} P\{Y_j = y_{jk}, j \in S\}, (Z_{lk} = z_{lk}, l \in V_m) \\
= P\{\cdot\}^{-1} \prod_{(y_j, j \in F_m)} \prod_{k=1}^{n} P\{Y_j = y_{jk}, j \in S\}, (Z_{lk} = z_{lk}, l \in V_m) \\
= P\{\cdot\}^{-1} \prod_{(y_j, j \in F_m)} \prod_{k=1}^{n} P\{Y_j = y_{jk}, j \in S\} P\{Z_l = z_l, l \in V_m | Y_j = y_{jk}, j \in S\} \\
= P\{\cdot\}^{-1} \prod_{(y_j, j \in F_m)} \prod_{k=1}^{n} P\{Y_j = y_{jk}, j \in F_m\} P\{Y_j = y_{jk}, j \in F_m\} \\
\cdot P\{Z_l = z_l, l \in V_m | Y_j = y_{jk}, j \in S\} \\
= P\{\cdot\}^{-1} \prod_{k} P\{Y_j = y_{jk}, j \in F_m\} \sum_{(y_j, j \in F_m)} P\{Y_j = y_{jk}, j \in F_m | Y_j = y_{jk}, j \in F_m\} \\
\cdot P\{Z_l = z_l, l \in V_m | Y_j = y_{jk}, j \in S\} \\
= P\{\cdot\}^{-1} \prod_{k} P\{Y_j = y_{jk}, j \in F_m\} \\
\times \sum_{(y_j, j \in F_m)} P\{Z_l = z_l, l \in V_m, (Y_j = y_{jk}, j \in F_m) | Y_j = y_{jk}, j \in F_m\} \\
= P\{\cdot\}^{-1} \prod_{k} P\{Y_j = y_{jk}, j \in F_m\} P\{Z_l = z_l, l \in V_m | Y_j = y_{jk}, j \in F_m\} \\
= P\{\cdot\}^{-1} \prod_{k=1}^{n} P\{Y_j = y_{jk}, j \in F_m\} \prod_{k=1}^{n} \tilde{q}_m(z_{lk}, l \in V_m | y_{jk}, j \in F_m) \\
= P\{\cdot\}^{-1} P\{\cdot\} \prod_{k=1}^{n} \tilde{q}_m(z_{lk}, l \in V_m | y_{jk}, j \in F_m) \\
= \prod_{k=1}^{n} \tilde{q}_m(z_{lk}, l \in V_m | y_{jk}, j \in F_m).
\end{align}

From the diagram, we see that this implies
\begin{align}
H(Y_j, j \in F_m \cap \Psi | (Y_j, j \in F_m \cap \Psi), (Z_l, l \in V_m)) = 0.
\end{align}

By (6), (Y_j, j \in F_m \cap \Psi) and (Y_j, j \in F_m \cap \Psi) are independent, so
\begin{align}
I(Y_j, j \in F_m \cap \Psi; Y_j, j \in F_m \cap \Psi) = 0.
\end{align}

By (34), (Y_j, j \in F_m \cap \Psi) and (Z_l, l \in V_m) are independent, so
\begin{align}
I(Y_j, j \in F_m \cap \Psi; Z_l, l \in V_m) = 0.
\end{align}

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

$$P((i_j, j \in F_m); (z_l, l \in V_m)) \in \mathbb{T}_{\lambda}(Y_{j}, Y_{j} \in F_m)\Psi, (Z_l, l \in V_m))$$

$$\leq 2^{-n(H(Y_{j}, j \in F_m); (Z_l, l \in V_m) - O(\delta))}$$

$$= 2^{-n(H(Y_{j}, j \in F_m); (Z_l, l \in V_m) - O(\delta))}$$

(38)

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

$$P(E_{1}E_{2}E_{3}E_{4}E_{5}) \leq \sum_{\Psi} \sum_{Z_l, l \in V_m} 2^{-n(H(Y_{j}, j \in F_m); (Z_l, l \in V_m) - O(\delta))}$$

$$\leq \sum_{\Psi} 2^{-n(H(Y_{j}, j \in F_m); (Z_l, l \in V_m) - O(\delta))}$$

$$= (2^{|F_m|} - 1) 2^{-n(H(Y_{j}, j \in F_m); (Z_l, l \in V_m) - O(\delta))} \leq \frac{\epsilon}{\lambda}$$

(39)

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

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

Hence, from (17) and (40), $R = (R_l, l \in \mathcal{E})$ is admissible, which implies $R_{m} \subset R$. \hfill \Box

III. OUTER BOUND

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

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

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

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

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

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

Theorem 2: $R \subset R_{\text{out}}$

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

1) $R_{\lambda}$ is replaced by $R_{\lambda}^*$.

2) The inequality in (4) is strict while the inequality in (44) is nonstrict.

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

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

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

and

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

We first consider such a code for a fixed $\epsilon$.

$$H(X_j, j \in F_m; l \in V_m) \leq H(Y_j, j \in F_m) \leq 2^{-n(H(X_j, j \in F_m; l \in V_m) - \log |X_j|)} \leq 3^{-n(H(Y_j, j \in F_m; l \in V_m) - \log |X_j|)}$$

(48)

In the above, 1) follows because $(X_j, j \in F_m)$ 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

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

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

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

$$H(T_l | X_j, j \in \mathcal{S}) = 0, \quad \text{for } l \in \mathcal{E} \quad (51)$$

$$H(X_j, j \in F_m; l \in V_m) \leq nO(\epsilon), \quad \text{for } m \in \mathcal{D} \quad (52)$$

$$H(X_j) \geq nH(X_j), \quad \text{for } j \in \mathcal{S} \quad (53)$$

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

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 $Y_j = X_j$ and $Z_l = T_l$, we see that there exists $h \in \Gamma_{N}^*$ such that

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

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

$$h_{Y_j, j \in \mathcal{F}_m} (Z_l, l \in V_m) \leq nO(\epsilon), \quad \text{for } m \in \mathcal{D} \quad (57)$$

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

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

It was shown in the recent work of Zhang and Yeung [11] that $\Gamma_{N}^*$ is a convex cone. Therefore, if $h \in \Gamma_{N}^*$, then $n^{-1}h \in \Gamma_{N}^*$.
Dividing (55)–(59) by $n$ and replacing $n^{-1}h$ by $h$, we see that there exists $h \in \Gamma_N^*$ such that

$$h(\gamma_j, j \in \mathcal{S}) = \sum_{j \in \mathcal{S}} h\gamma_j$$

$$h\mathcal{Z}_l(\gamma_j, j \in \mathcal{U}_l) = 0, \quad \text{for } l \in \mathcal{E}$$

$$h(\gamma_j, j \in \mathcal{F}_m)\mathcal{Z}_l(t, \mathcal{C}_m) \leq O(\epsilon), \quad \text{for } m \in \mathcal{D}$$

$$h\gamma_j > H(X_j), \quad \text{for } j \in \mathcal{S}$$

$$R_l + \epsilon \geq h\mathcal{Z}_l, \quad \text{for } l \in \mathcal{E}.$$ (60) (61) (62) (63) (64)

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

IV. GEOMETRICAL INTERPRETATIONS OF $\mathcal{R}_{\text{in}}$ AND $\mathcal{R}_{\text{cat}}$

In Sections II and III, $\mathcal{R}_{\text{in}}$ and $\mathcal{R}_{\text{cat}}$ 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

$$\mathcal{T}(A) = \{r \in \mathbb{R}^{[\mathcal{E}]}; \ r \geq r' \text{ for some } r' \in A\}$$

where we write $r \geq r'$ to mean $r$ greater than or equal to $r'$ componentwise. For a set $B \subset \mathbb{R}^{2^{[\mathcal{N}]}-1}$, define $\mathcal{T}(B) = \{r \in \mathbb{R}^{[\mathcal{E}]}; \ h(\gamma_j, j \in \mathcal{S}) = \sum_{j \in \mathcal{S}} h\gamma_j\}$.

$$C_1 = \{h \in \mathbb{R}^{2^{[\mathcal{N}]}-1}; \ h(\gamma_j, j \in \mathcal{S}) = \sum_{j \in \mathcal{S}} h\gamma_j\}$$

$$C_2 = \{h \in \mathbb{R}^{2^{[\mathcal{N}]}-1}; \ h\mathcal{Z}_l(\gamma_j, j \in \mathcal{U}_l) = 0 \text{ for } l \in \mathcal{E}\}$$

$$C_3 = \{h \in \mathbb{R}^{2^{[\mathcal{N}]}-1}; \ h(\gamma_j, j \in \mathcal{F}_m)\mathcal{Z}_l(t, \mathcal{C}_m) = 0 \text{ for } m \in \mathcal{D}\}$$

$$C_4 = \{h \in \mathbb{R}^{2^{[\mathcal{N}]}-1}; \ h\gamma_j > H(X_j) \text{ for } j \in \mathcal{S}\}.$$ (65) (66) (67) (68)

Note that the independence relation in (1) is interpreted as the hyperplane $C_1$ in $\mathbb{R}^{2^{[\mathcal{N}]}-1}$. 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

$$\mathcal{R}_{\text{in}} = \mathcal{T}(\mathcal{C}_1 \cap \mathcal{C}_2 \cap \mathcal{C}_3 \cap \mathcal{C}_4).$$ (69)

Similarly, we see that

$$\mathcal{R}_{\text{cat}} = \mathcal{T}(\mathcal{C}_1 \cap \mathcal{C}_2 \cap \mathcal{C}_3 \cap \mathcal{C}_4).$$ (70)

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

V. THE LP BOUND

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

$$h_G \geq 0$$

$$h_G \geq 0$$

$$i_{G} \geq 0$$

$$i_{G} \geq 0.$$ (71) (72) (73) (74)

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

$$\mathcal{R}_{\text{LP}} = \mathcal{T}(\mathcal{C}_1 \cap \mathcal{C}_2 \cap \mathcal{C}_3 \cap \mathcal{C}_4).$$ (75)

The geometrical interpretation of $\mathcal{R}_{\text{cat}}$ in the last section can also be applied to $\mathcal{R}_{\text{LP}}$, so it is not difficult to see that $\mathcal{R}_{\text{LP}}$ 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 ($K \geq 2$), there are $K$ sources, $S_1, \ldots, S_K$, 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 $1 \leq L(m) \leq K$, and it reconstructs $S_1, \ldots, S_{L(m)}$, the most important $L(m)$ sources. The problem is specified by

$$S = \{1, \ldots, K\}$$

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

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

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

$$F_m = \{1, \ldots, L(m)\}.$$ (76)

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

$$\left\{R_1, \ldots, R_K : \sum_{l=1}^{K} A_l R_l \geq \sum_{j=1}^{K} f_j(\mathcal{A}) H(X_j) \right\}$$

for all $\mathcal{A} \subset \mathcal{C}$} (76)

where $\mathcal{A} = \{A_1, \ldots, A_K\}$ is some subset of $\mathcal{C}$, and $f_j(\mathcal{A}), 1 \leq j \leq K$, are some nonnegative functions depending only on $\mathcal{A}$. 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 $\mathbf{v}$ to denote the $i$th component of a vector $\mathbf{v} \in \{0, 1\}^K$. In the SMDC problem

$$S = \{1, \cdots, K\}$$
$$E = \{1, \cdots, K\}$$
$$D = \{\mathbf{v} \in \{0, 1\}^K : |\mathbf{v}| \geq 1\}$$
$$A = S \times E$$
$$B = \{(l, \mathbf{v}) : \mathbf{v}_l = 1\}$$

and

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

With the above specifications, we implicitly have $L(\mathbf{v}) = |\mathbf{v}|$.

From [13], the coding rate region for the SMDC problem is given by

$$R_{\text{SMDC}} = \left\{ (R_1, \cdots, R_K) : \sum_{l=1}^K A_l R_l \geq \sum_{j=1}^K f_j(\mathbf{A}) H(X_j) \right\}$$

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

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

Lemma 2: $R_{\text{SMDC}} = \mathcal{Y}(R_{\text{SMDC}})$.

Proof: Trivial.

In the proof for the necessity of $R_{\text{SMDC}}$ 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 $Z_\mathbf{v}$ to denote the collection of random variables $Z_\mathbf{v}$ such that $\mathbf{v}_l = 1$ 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 $Z_\mathbf{v}, \mathbf{v} \in \{0, 1\}^K$, if

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

then

$$\sum_{l=1}^K A_l H(Z_l) \geq \sum_{j=1}^K f_j(\mathbf{A}) H(Y_j) \quad \text{for all } \mathbf{A} \geq 0. \quad (77)$$

Let us now state another proposition.

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

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

then

$$\sum_{l=1}^K A_l H(Z_l) \geq \sum_{j=1}^K f_j(\mathbf{A}) H(X_j) \quad \text{for all } \mathbf{A} \geq 0. \quad (78)$$

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

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) imply (77). Together with 3), we have

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

since $f_j(\mathbf{A}) \geq 0$, 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

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

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 $C_1$, $C_2$, and $C_3$, 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

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

(79)

So

$$\Gamma_N \cap C_1 \cap \mathcal{C}_2 \cap \mathcal{C}_3 \cap \mathcal{C}_4 \subset \left\{ h \in \mathbb{R}^{2^{|N|}-1} : \sum_{l=1}^K A_l h Z_l \geq \sum_{j=1}^K f_j(\mathbf{A}) H(X_j) \right\} \quad \text{for all } \mathbf{A} \geq 0 \right\}.$$  

(80)
By projecting onto the coordinates \( h_{Z_1}, I \in E \), we have
\[
\text{proj}_{\{h_{Z_1}, I \in E\}}(\Gamma_N \cap C_1 \cap C_2 \cap C_3 \cap \overline{C_4})
\subset \left\{ \begin{array}{l}
(h_{Z_1}, \cdots, h_{Z_K}) : \\
\sum_{i=1}^{K} A_i h_{Z_i} \geq \sum_{j=1}^{K} f_j(A) H(X_j)
\end{array} \right\}
\]
for all \( A \geq 0 \)
\[
= \left\{ (R_1, \cdots, R_K) : \\
\sum_{i=1}^{K} A_i R_i \geq \sum_{j=1}^{K} f_j(A) H(X_j)
\right\}
\]
where the last step follows from Lemma 2. Since \( R_{LP} \) is an outer bound on \( R_{SMDC} \), we have
\[
R_{SMDC} \subset R_{LP} \subset R_{SMDC}
\]
which implies \( R_{LP} = R_{SMDC} \). Thus \( R_{LP} \) is tight. Compared with \( R_{SMDC} \), \( R_{LP} \) 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 \( R_{in} \) and an outer bound \( R_{out} \) on the coding rate region. Both \( R_{in} \) and \( R_{out} \) are implicit in the sense that they are specified in terms of \( \Gamma_N^* \) 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 \( R_{LP} \) 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.

\( R_{LP} \) 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 \( R_{LP} \) 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


