Asymmetric Multilevel Diversity Coding and Asymmetric Gaussian Multiple Descriptions

Soheil Mohajer, *Student Member, IEEE*, Chao Tian, *Member, IEEE*, and Suhas N. Diggavi, *Member, IEEE*

Abstract

We consider the asymmetric multilevel diversity (A-MLD) coding problem, where a set of $2^K - 1$ information sources, ordered in a decreasing level of importance, is encoded into K messages (or descriptions). There are $2^K - 1$ decoders, each of which has access to a non-empty subset of the encoded messages. Each decoder is required to reproduce the information sources up to a certain importance level depending on the combination of descriptions available to it. We obtain a single letter characterization of the achievable rate region for the 3-description problem. In contrast to symmetric multilevel diversity coding, source-separation coding is not sufficient in the asymmetric case, and ideas akin to network coding need to be used strategically. Based on the intuitions gained in treating the A-MLD problem, we derive inner and outer bounds for the rate region of the asymmetric Gaussian multiple description (MD) problem with three descriptions. Both the inner and outer bounds have a similar geometric structure to the rate region template of the A-MLD coding problem, and moreover, we show that the gap between them is small, which results in an approximate characterization of the asymmetric Gaussian three description rate region.

I. Introduction

In the symmetric multilevel diversity coding (MLD) problem [9], K source sequences are encoded into K descriptions, which are sent to the decoders through noiseless channels. These source sequences have a decreasing levels of importance, and each decoder has access to a non-empty subset of the descriptions. The goal of the encoder is to produce the descriptions such that each decoder with k available descriptions is able to reconstruct the k most important source sequences. The symmetric MLD problem was motivated

S. Mohajer and S. N. Diggavi are with the School of Computer and Communication Sciences, Ecole Polytechnique Federale de Lausanne, Switzerland. C. Tian is with AT&T Labs-Research, Florham Park, New Jersey, USA.

by fault-tolerant storage for disk arrays and for incremental priority encoding on packet erasure channels; see [9] for more details. The MLD problem with three levels was solved by Roche *et al*. in [9], and the result was later extended by Yeung and Zhang [10] to an arbitrary number of levels. It was shown that source-separation coding[1] is optimal for the symmetric problem. This means that each source sequence can be compressed separately, and then the descriptions are obtained by concatenating the compressed source sequences appropriately.

In this work we formulate the asymmetric multilevel diversity (A-MLD) coding problem. The problem can be understood as a refined version of symmetric MLD coding problem, and it is naturally applicable in distributed disk storage applications with asymmetric (unequal) reliabilities, in contrast to symmetric (equal) reliabilities which motivate the symmetric MLD problem. Similarly, for packet erasure applications, the erasure probabilities for the sub-packets may not be equal because the paths over which they are sent may have different reliabilities. As such, in both applications, we may wish to utilize not just the number of the encoders which are accessible, but also their identities, since the descriptions are no longer symmetric. Therefore, the difference between the MLD and A-MLD problem is that in the asymmetric version the levels of reconstruction is determined by the specific combination of descriptions available to them, not just the number of descriptions.

More precisely, 2^K-1 source sequences are encoded into K descriptions at the encoder. The 2^K-1 decoders are ordered in a specific way, and the goal of the encoder is to produce the descriptions such that the k-th decoder is able to reconstruct the k most important source sequences, for $k=1,\ldots,2^K-1$. In this work, we only consider the 3-description case and provide a complete characterization of the achievable rate region. In particular we show that source-separation coding coding is *not* optimal for this problem, and the source sequences in different levels have to be jointly encoded (like in *network coding*) in an optimal coding strategy. We also show that the scheme using *linear* combinations of these compressed sequences is optimal. We note that various special cases of 3-description problem were studied in[2] [11], where, however, only no more than *three* information sources were considered. The characterization we provide in this work strictly subsumes those considered in [11].

Let us now turn to a closely related problem, namely the multiple description (MD) problem. In this problem a source is mapped into K descriptions and sent to 2^K-1 decoders, just as in the A-MLD

[1]This was called superposition coding in these papers. In order not to confuse this with the common terminology of broadcast channels, the new terminology has been adopted here, as suggested by R. Yeung.

[2]We would like to thank R. Yeung for bringing this work to our attention.

coding problem. The decoders are required to reconstruct the source sequence within certain distortions using the available descriptions. The MD rate region characterization is long-standing open problem in information theory with a long history [1]–[3]. Despite many important results, the problem is still open, even for the quadratic Gaussian case with only three descriptions. Using the intuitions gained in treating the A-MLD problem as well as the sum-rate lower bound for symmetric Gaussian MD problem recently discovered in [5], we develop inner and outer bounds for the MD rate region, both of which bear similar geometric structure to the A-MLD coding rate region. Moreover, the gap between the bounds is small (less than 1.3 bits in terms of the Euclidean distance between the bounding planes), yielding an approximate characterization. One surprising consequence of this result is that the proposed simple architecture based on successive refinement (SR) [14] and A-MLD coding is in fact close to optimality. From an engineering viewpoint, this suggests that one can design simple and flexible MD codes that are (approximately) optimal.

One important observation leading to this work is the intimate connection between the multilevel diversity (MLD) coding problem and the MD problem observed in [7], where we showed that for the symmetric MD problem, achievable rate region based on SR coding coupled with symmetric multilevel diversity (S-MLD) coding provides good approximation to the MD rate region under symmetric distortion constraints; perhaps more interestingly, the achievable rate region has the same geometric structure as that of the symmetric MLD coding rate region. In fact, the symmetric MLD coding result is essential for establishing the symmetric MD result in [7]. The result in [7] suggests a general approach in treating lossy source coding problems: first solve a corresponding a lossless version of the problem, then extend the results and intuitions to its lossy counterpart to yield an approximate characterization. This is exactly our motivation to formulate the A-MLD coding problem, and indeed the result given in this work further illustrates the effectiveness of this approach.

The paper is organized as follows. In Section II, we introduce the notations and provide a formal definition of the problems. In Section III, we present the main results of the paper. We prove the main theorem for rate region characterization of the A-MLD problem in Section IV. In Section V, we focus on deriving the outer and inner bounds for the rate region of the A-MD problem. Finally, Section VI concludes the paper. Some of the detailed and technical proofs are given in the appendix.

II. Notations and Problem Formulation

In this section we provide formal definitions for both the asymmetric multilevel diversity (A-MLD) and the asymmetric multiple description (A-MD) coding problems. Since we need to use the result of the

A-MLD problem when treating the A-MD problem, we may use different notations for these problems in order to avoid confusion.

A. Asymmetric Multilevel Diversity Coding

Let $\{(V_{1,t}, V_{2,t}, \ldots, V_{2^K-1,t})\}_{t=1,2,\ldots}$ be an independent and identically distributed process sampled from a finite size alphabet $\mathcal{V}_1 \times \mathcal{V}_2 \times \cdots \times \mathcal{V}_{2^K-1}$ with time index t. This can be considered as $2^K - 1$ pieces of independent data streams, namely, $\{V_{1,t}\}, \ldots, \{V_{2^K-1,t}\}$, where each data stream is an independently and identically distributed sequence. The data streams are ordered with decreasing importance, e.g., consecutive refinements of a single source. We use V_i^n to denote a length n sequence of V_i, namely, $V_i^n = (V_{i,1}, \ldots, V_{i,n})$.

Define the vector random variables U_j as $U_j \triangleq (V_1, \ldots, V_j)$ for $j = 1, \ldots, 2^K - 1$, and $U_0 \triangleq 0$. We use U_j^n to denote length n sequences of U_j. We may simply use U^n to denote $U_{2^K-1}^n = (V_1^n, \ldots, V_{2^K-1}^n)$ for brevity. Note that U_j^n is a two-dimensional array, whose elements are independent of each other along both directions, $i = 1, \ldots, j$, and $t = 1, \ldots, n$.

The Shannon entropy rate of the source V_k is denoted by h_k. We also denote the entropy of U_j by H_j, where the independence of sources V_k's implies

$$H_j = H(U_j) = H(V_1, \ldots, V_j) = \sum_{i=1}^{j} H(V_i) = \sum_{i=1}^{j} h_i. \tag{1}$$

The A-MLD problem can be described as follows. Consider $2^K - 1$ source sequences which are fed to a single encoder. The encoder produces K descriptions, denoted as $\Gamma_1, \Gamma_2 \ldots, \Gamma_K$ to encode the source sequences. The descriptions are sent over K perfect channel. There are $2^K - 1$ decoders, each has access to a non-empty subset of the descriptions, $\mathcal{S} \subseteq \{\Gamma_1, \Gamma_2 \ldots, \Gamma_K\}$, and wishes to decode losslessly the source data streams below a certain *level*, which is a function of the description set $\mathcal{S}$. Fig. 1 illustrates the problem setting for $K = 3$, and a specific decoding requirement for the decoders.

Formally, we define the notion of *ordering level* to connect the decoding requirement of the decoders to their available description subsets as follows.

***Definition** 1:* A valid *ordering level* (or simply *ordering*) on the non-empty subsets[3] of $\{\Gamma_1, \Gamma_2 \ldots, \Gamma_K\}$ is a one-to-one mapping $\mathscr{L} : \mathcal{P}(\{\Gamma_1, \Gamma_2 \ldots, \Gamma_K\}) \setminus \emptyset \longrightarrow \{1, \ldots, 2^K - 1\}$ satisfying

(i) $\mathscr{L}(\{\Gamma_1\}) < \mathscr{L}(\{\Gamma_2\}) < \cdots < \mathscr{L}(\{\Gamma_K\})$,

(ii) $\mathcal{S} \subset \mathcal{T}$ implies $\mathscr{L}(\mathcal{S}) < \mathscr{L}(\mathcal{T})$,

[3]For the rest of this paper, by *subset* we always mean a non-empty subset although it is not precisely mentioned.

where $\mathcal{P}(M)$ is the power set of M.

The ordering level will be used to determine the decoding requirements of the decoders, *e.g.*, a decoder with a set of descriptions $\mathcal{S}$ needs to decode the first $\mathscr{L}(\mathcal{S})$ source streams. Condition (i) is given to avoid permuted repetition of the levels, where without loss of generality, we assume an initial ordering on the single description decoders. Condition (ii) is a natural fact that if $\mathcal{S}$ is a subset of $\mathcal{T}$, then the corresponding decoder can not do better than what decoder $\mathcal{T}$ can. We may simplify the notation occasionally, by omitting the braces, *e.g.*, $\mathscr{L}(\Gamma_1,\Gamma_2) = \mathscr{L}(\{\Gamma_1,\Gamma_2\})$. The inverse mapping $\mathscr{L}^{-1}(k)$ is well defined, which is the subset of descriptions whose ordering level is k.

An $(n;\mathscr{L};M_i, i\in\{1,2,\ldots,K\})$ MLD-code is defined by a set of encoding functions

$$F_i : \mathcal{V}_1^n \times \mathcal{V}_2^n \times \cdots \times \mathcal{V}_{2^K-1}^n \longrightarrow \{1,2,\ldots,M_i\}, \quad i\in\{1,2,\ldots,K\}, \tag{2}$$

and decoding functions

$$G_{\mathcal{S}} : \prod_{j:\Gamma_j\in\mathcal{S}} \{1,\ldots,M_j\} \longrightarrow \mathcal{V}_1^n \times \mathcal{V}_2^n \times \cdots \times \mathcal{V}_{\mathscr{L}(\mathcal{S})}^n, \quad \mathcal{S}\subseteq\{\Gamma_1,\ldots,\Gamma_K\}, \tag{3}$$

where $\prod$ denotes a set product. We define

$$\hat{U}^n_{\mathscr{L}(\mathcal{S})}(\mathcal{S}) \triangleq G_{\mathcal{S}}(F_j(U^n); j:\Gamma_j\in\mathcal{S}) \tag{4}$$

and $\hat{V}_i^n(\mathcal{S})$ is the corresponding part of $\hat{U}^n_{\mathscr{L}(\mathcal{S})}(\mathcal{S})$, for $i\le\mathscr{L}(\mathcal{S})$.

A rate tuple $\mathbf{R}^{\mathscr{L}} = (R_1,R_2,\ldots,R_K)$ is called admissible for a prescribed ordering $\mathscr{L}$, if for any $\varepsilon>0$ and sufficiently large n, there exist an $(n;\mathscr{L};M_i, i\in\{1,2,\ldots,K\})$ MLD-code such that

$$\frac{1}{n}\log M_i \le R_i+\varepsilon, \quad i\in\{1,2,\ldots,K\}, \tag{5}$$

and

$$\Pr(\hat{V}_i^n(\mathcal{S}) \ne V_i^n) < \varepsilon \quad \forall\mathcal{S}\subseteq\{\Gamma_1,\ldots,\Gamma_K\}, \text{ and } \forall i\le\mathscr{L}(\mathcal{S}). \tag{6}$$

The main goal in the (lossless) multilevel diversity coding problem is to characterize $\mathcal{R}_{\text{MLD}}$, the set of all achievable rate tuples $(R_i; i\in\{1,2,\ldots,K\})$ in terms of the entropy of the source sequences and the given ordering level. We denote such rate region by $\mathcal{R}^{\mathscr{L}}_{\text{MLD}}$ for a specific ordering.

In this paper we consider this problem for three descriptions ($K=3$) and give a complete characterization of the rate region. It is straightforward to show that there are eight possible orderings for $\{\Gamma_1,\Gamma_2,\Gamma_3\}$, which are shown in Table I. We may further divide each ordering into sub-regimes to simplify the problem for each case. The results of this work are general and hold for all possible orderings. However, in order to

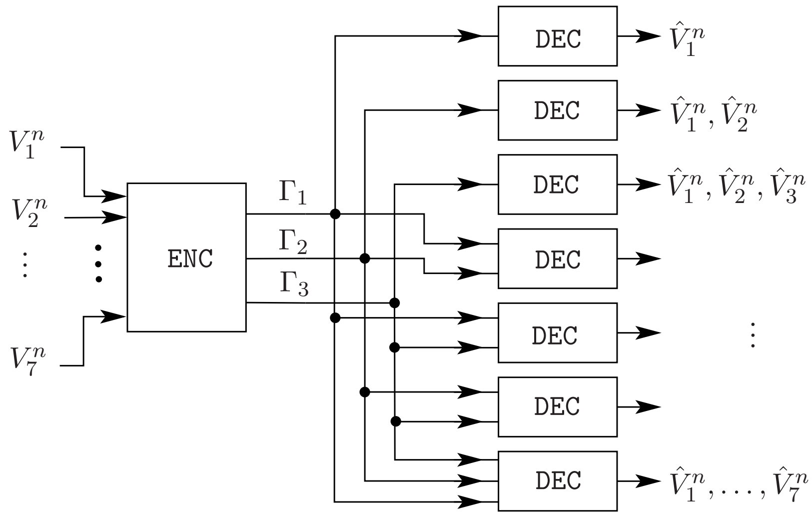

Fig. 1. The 3-description asymmetric multilevel diversity coding problem for ordering level $\mathscr{L}_1$.

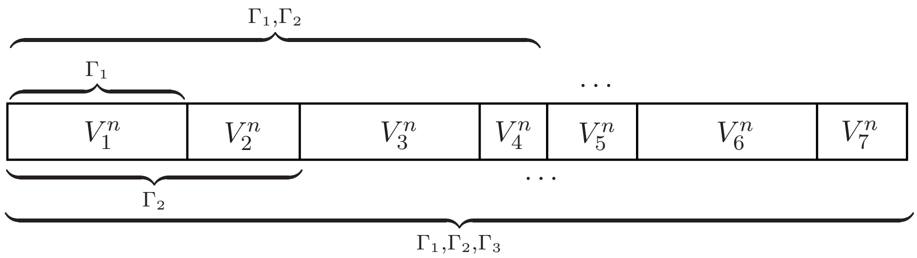

Fig. 2. Levels assigned to the description subsets determines the recoverable source subsequences. The requirements corresponding to the ordering level $\mathscr{L}_1$ are shown in this figure.

illustrate the result, we may specialize some of the arguments/theorems to the ordering level $\mathscr{L}_1$ defined as

$$\mathscr{L}_1(\Gamma_1) = 1, \quad \mathscr{L}_1(\Gamma_2) = 2, \quad \mathscr{L}_1(\Gamma_3) = 3,$$
$$\mathscr{L}_1(\Gamma_1, \Gamma_2) = 4, \quad \mathscr{L}_1(\Gamma_1, \Gamma_3) = 5, \quad \mathscr{L}_1(\Gamma_2, \Gamma_3) = 6, \quad \mathscr{L}_1(\Gamma_1, \Gamma_2, \Gamma_3) = 7.$$

The setting of the problem for the ordering level $\mathscr{L}_1$ is shown in Fig. 1.

Fig. 2 shows the subset of source streams which should be recovered by each subset of descriptions in $\mathscr{L}_1$ setting.

B. Asymmetric Gaussian Multiple-Description Coding

Let $\{X(t)\}_{t=1,2,...}$ be a sequence of independent and identically distributed zero mean and unit variance real-valued Gaussian source, *i.e.*, $\mathcal{X} = \mathbb{R}$, with time index t. Moreover, the reconstruction alphabet is also assumed to be $\mathbb{R}$. The vector $X(1), X(2), \ldots, X(n)$ is denoted by X^n. We use capital letters for random variables, and the corresponding lower-case letters for their realization. The quality of the reconstruction is measured by the quadratic distance between the original sequence x^n and the reconstructed one $\hat{x}^n$. Formally, we define the distortion as

$$d(x^n, \hat{x}^n) = \frac{1}{n} \sum_{k=1}^{n} |x(k) - \hat{x}(k)|^2, \tag{7}$$

In a general multiple description setting, the encoders produces K descriptions, namely $\Gamma_1, \Gamma_2, \ldots, \Gamma_K$ based on the source sequence and sends them to the decoders through noiseless channels. Each decoder receives a non-empty subset of the descriptions, and has to reconstruct the source sequence $\hat{x}^n$ which satisfies a certain level of fidelity.

In a manner similar to the last subsection, we denote each decoder by the corresponding set of available descriptions. Each decoder $\mathcal{S}$ has a distortion constraint $D_{\mathcal{S}}$, and needs to reconstruct the source such that the corresponding expected distortion does not exceed this constraint. The main goal in this problem is to characterize the set of admissible rates of the descriptions in a way that such reconstructions are possible. We present a formal definition of the problem next.

An $(n; M_i, i \in \{1, \ldots, K\}; \Delta_{\mathcal{S}}, \mathcal{S} \subseteq \{\Gamma_1, \ldots, \Gamma_K\})$ MD-code is defined as a set of encoding functions

$$F_i : \mathcal{X}^n \longrightarrow \{1, 2, \ldots, M_i\}, \quad i \in \{1, 2, \ldots, K\}, \tag{8}$$

and $2^K - 1$ decoding functions

$$G_{\mathcal{S}} : \prod_{j:\Gamma_j \in \mathcal{S}} \{1, \ldots, M_j\} \longrightarrow \mathcal{X}^n, \quad \mathcal{S} \subseteq \{\Gamma_1, \ldots, \Gamma_K\}, \tag{9}$$

with

$$\Delta_{\mathcal{S}} = \mathbb{E} d(X^n, \hat{X}^n_{\mathcal{S}}), \quad \mathcal{S} \subseteq \{\Gamma_1, \ldots, \Gamma_K\}, \tag{10}$$

where

$$\hat{X}^n_{\mathcal{S}} = G_{\mathcal{S}}(F_j(X^n), j : \Gamma_j \in \mathcal{S}). \tag{11}$$

Again, $\prod$ denotes set product, and $\mathbb{E}$ is the expectation operator.

A rate tuple $\mathbf{R} = (R_1, R_2, \ldots, R_K)$ is called $\mathbf{D} = (D_{\mathcal{S}}; \mathcal{S} \subseteq \{\Gamma_1, \ldots, \Gamma_K\})$-admissible if for every $\varepsilon > 0$ and sufficiently large n, there exists an $(n; M_i, i \in \{1, \ldots, K\}; \Delta_{\mathcal{S}}, \mathcal{S} \subseteq \{\Gamma_1, \ldots, \Gamma_K\})$ MD-code such that

$$\frac{1}{n} \log M_i \leq R_i + \varepsilon, \quad i \in \{1, \ldots, K\}, \tag{12}$$

and

$$\Delta_{\mathcal{S}} \leq D_{\mathcal{S}} + \varepsilon, \quad \mathcal{S} \subseteq \{\Gamma_1, \ldots, \Gamma_K\}. \tag{13}$$

We denote by $\mathcal{R}_{\text{MD}}(\mathbf{D})$ the set of all $\mathbf{D}$-admissible rate tuples, which we seek to characterize.

Let $\mathcal{T}$ and $\mathcal{S}$ be two description sets, satisfying $\mathcal{T} \subseteq \mathcal{S} \subseteq \{\Gamma_1, \ldots, \Gamma_K\}$. It is clear that the decoder with access to $\mathcal{S}$ can reconstruct the source sequence as *well* as the one with access to $\mathcal{T}$ does, even if $D_{\mathcal{S}} \geq D_{\mathcal{T}}$. The following lemma shows that slightly modification of the distortion vector in order to satisfy such property does not change the admissible rate region.

Lemma 1: For a given distortion vector $\mathbf{D}$, define $\tilde{\mathbf{D}}$ as $\tilde{\mathbf{D}} = (\tilde{D}_{\mathcal{S}}; \mathcal{S} \subseteq \{\Gamma_1, \ldots, \Gamma_K\})$, where

$$\tilde{D}_{\mathcal{S}} = \min_{\mathcal{T}: \mathcal{T} \subseteq \mathcal{S}} D_{\mathcal{T}}.$$

Then $\mathcal{R}_{\text{MD}}(\tilde{\mathbf{D}}) = \mathcal{R}_{\text{MD}}(\mathbf{D})$.

Proof of Lemma 1: It is clear that $\tilde{D}_{\mathcal{S}} \leq D_{\mathcal{S}}$ for all $\mathcal{S} \subseteq \{\Gamma_1, \ldots, \Gamma_K\}$, and therefore $\mathcal{R}_{\text{MD}}(\tilde{\mathbf{D}}) \subseteq \mathcal{R}_{\text{MD}}(\mathbf{D})$. So, it remains to prove $\mathcal{R}_{\text{MD}}(\mathbf{D}) \subseteq \mathcal{R}_{\text{MD}}(\tilde{\mathbf{D}})$. Let $\mathbf{R} \in \mathcal{R}_{\text{MD}}(\mathbf{D})$ be an admissible rate tuple for $\mathbf{D}$, and $(n; M_i; \Delta_{\mathcal{S}})$ be a code for a given ε which achieves the distortion constraints $\mathbf{D}$, with encoding functions $\{F_i\}$ and decoding functions $\{G_{\mathcal{S}}\}$. We can easily modify the decoding functions and obtain a code which satisfies $\tilde{\mathbf{D}}$. By the definition of $\tilde{\mathbf{D}}$, for all $\mathcal{S}$ we have $\tilde{D}_{\mathcal{S}} = D_{\tilde{\mathcal{S}}}$, where

$$\tilde{\mathcal{S}} \triangleq \arg \min_{\mathcal{T}: \mathcal{T} \subseteq \mathcal{S}} D_{\mathcal{T}}.$$

Define

$$\tilde{X}^n_{\mathcal{S}} = \tilde{G}_{\mathcal{S}}(F_j(X^n); j : \Gamma_j \in \mathcal{S}) \triangleq \hat{X}^n_{\tilde{\mathcal{S}}}.$$

Obviously,

$$\mathbb{E} d(X^n, \tilde{X}^n_{\mathcal{S}}) = \mathbb{E} d(X^n, \hat{X}^n_{\tilde{\mathcal{S}}}) \leq D_{\tilde{\mathcal{S}}} + \varepsilon = \tilde{D}_{\mathcal{S}} + \varepsilon.$$

Thus the similar code with the modified decoding functions satisfies the constraint tuple $\tilde{\mathbf{D}}$, and therefore $\mathbf{R} \in \mathcal{R}_{\text{MD}}(\tilde{\mathbf{D}})$. ■

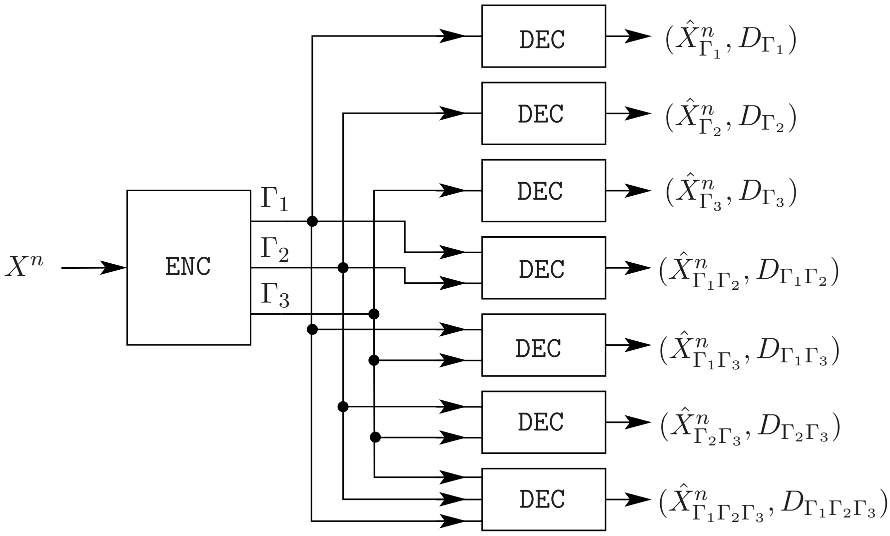

Fig. 3. The three description source coding problem with ordering level $\mathscr{L}_1$.

Given this lemma, we can assume, without loss of generality, that $D_{\mathcal{T}} \leq D_{\mathcal{S}}$ for all $\mathcal{S} \subseteq \mathcal{T}$. These distortion constraints then induce an ordering on the decoders, or equivalently on their associated subset of descriptions.

In this work, again we focus on the three description ($K = 3$) problem, and present the results in general form, *i.e.,* regardless the exact ordering. Occasionally we shall provide the proof details only for the specific sorted distortion constraints

$$D_{\Gamma_1} \geq D_{\Gamma_2} \geq D_{\Gamma_3} \geq D_{\Gamma_1\Gamma_2} \geq D_{\Gamma_1\Gamma_3} \geq D_{\Gamma_2\Gamma_3} \geq D_{\Gamma_1\Gamma_2\Gamma_3},$$

which induces the ordering

$$\mathscr{L}(\Gamma_1) < \mathscr{L}(\Gamma_2) < \mathscr{L}(\Gamma_3) < \mathscr{L}(\Gamma_1\Gamma_2) < \mathscr{L}(\Gamma_1\Gamma_3) < \mathscr{L}(\Gamma_2\Gamma_3) < \mathscr{L}(\Gamma_1\Gamma_2\Gamma_3)$$

on the subsets of descriptions, which is exactly the aforementioned ordering $\mathscr{L}_1$. Fig. 3 shows the setting of this problem for the ordering $\mathscr{L}_1$. It is worth mentioning that the distortion constraints may also induce different ordering of subsets of the descriptions. All possible ordering functions are listed in Table I.

III. The Main Results

In this section we present the main results of the paper. We state the theorems in a unified way which hold for all orderings, and also specialize it to the ordering $\mathscr{L}_1$ to facilitate understanding and further discussion. We start with the admissible rate region of the A-MLD problem, $\mathcal{R}_{\text{MLD}}$, and then give an approximate characterization of the rate region of the A-MD problem based on the coding scheme inspired by the A-MLD problem.

A. The Admissible Rate Region of 3*-Description Asymmetric Multilevel Diversity Coding*

The following theorem characterizes the admissible rate region of the asymmetric multilevel diversity coding problem for an arbitrary ordering level.

***Theorem** 1:* Let $\tilde{V} = (V_1, \ldots, V_7)$ be a given sequence of sources with entropy sequence $\mathbf{H} = (H_1, \ldots, H_7) = (H(V_1), H(V_1, V_2), \ldots, H(V_1, \ldots, V_7))$. For a given ordering level $\mathscr{L}$, the rate region $\mathcal{R}^{\mathscr{L}}_{\text{MLD}}(\mathbf{H})$ is the set of all non-negative triples (R_1, R_2, R_3) which satisfy

$$R_i \geq H_{\mathscr{L}(\Gamma_i)}, \qquad i = 1, 2, 3 \tag{P1}$$

$$R_i + R_j \geq H_{\min\{\mathscr{L}(\Gamma_i), \mathscr{L}(\Gamma_j)\}} + H_{\mathscr{L}(\Gamma_i, \Gamma_j)}, \qquad i \neq j \tag{P2}$$

$$\begin{aligned} 2R_i + R_j + R_k \geq\ & H_{\min\{\mathscr{L}(\Gamma_i), \mathscr{L}(\Gamma_j)\}} + H_{\min\{\mathscr{L}(\Gamma_i), \mathscr{L}(\Gamma_k)\}} \\ & + H_{\min\{\mathscr{L}(\Gamma_i, \Gamma_j), \mathscr{L}(\Gamma_i, \Gamma_k)\}} + H_{\mathscr{L}(\Gamma_i, \Gamma_j, \Gamma_k)}, \qquad i \neq j \neq k \end{aligned} \tag{P3}$$

$$R_1 + R_2 + R_3 \geq H_{\mathscr{L}(\Gamma_1)} + H_{\min\{\mathscr{L}(\Gamma_1, \Gamma_2), \mathscr{L}(\Gamma_3)\}} + H_{\mathscr{L}(\Gamma_1, \Gamma_2, \Gamma_3)}, \tag{P4}$$

$$\begin{aligned} R_1 + R_2 + R_3 \geq\ & H_{\mathscr{L}(\Gamma_1)} + \frac{1}{2} H_{\mathscr{L}(\Gamma_2)} + \frac{1}{2} H_{\min\{\mathscr{L}(\Gamma_1, \Gamma_2), \mathscr{L}(\Gamma_1, \Gamma_3), \mathscr{L}(\Gamma_2, \Gamma_3)\}} \\ & + H_{\mathscr{L}(\Gamma_1, \Gamma_2, \Gamma_3)}. \end{aligned} \tag{P5}$$

In the following corollary, we specialize the bounds for the specific ordering $\mathscr{L}_1$.

***Corollary** 1:* For the ordering level $\mathscr{L}_1$, the admissible rate region of the three-description A-MLD problem is given by the set of all rate triples (R_1, R_2, R_3) which satisfy

$$R_1 \geq H(V_1), \tag{Q1}$$

$$R_2 \geq H(V_1) + H(V_2), \tag{Q2}$$

$$R_3 \geq H(V_1) + H(V_2) + H(V_3), \tag{Q3}$$

$$R_1 + R_2 \geq 2H(V_1) + H(V_2) + H(V_3) + H(V_4), \tag{Q4}$$

$$R_1 + R_3 \geq 2H(V_1) + H(V_2) + H(V_3) + H(V_4) + H(V_5), \tag{Q5}$$

$$R_2 + R_3 \geq 2H(V_1) + 2H(V_2) + H(V_3) + H(V_4) + H(V_5) + H(V_6), \tag{Q6}$$

$$2R_1 + R_2 + R_3 \geq 4H(V_1) + 2H(V_2) + 2H(V_3) + 2H(V_4) + H(V_5) + H(V_6) + H(V_7), \tag{Q7}$$

$$R_1 + 2R_2 + R_3 \geq 4H(V_1) + 3H(V_2) + 2H(V_3) + 2H(V_4) + H(V_5) + H(V_6) + H(V_7), \tag{Q8}$$

$$R_1 + R_2 + 2R_3 \geq 4H(V_1) + 3H(V_2) + 2H(V_3) + 2H(V_4) + 2H(V_5) + H(V_6) + H(V_7), \tag{Q9}$$

$$R_1 + R_2 + R_3 \geq 3H(V_1) + 2H(V_2) + 2H(V_3) + H(V_4) + H(V_5) + H(V_6) + H(V_7), \tag{Q10}$$

$$R_1 + R_2 + R_3 \geq 3H(V_1) + 2H(V_2) + \frac{3}{2}H(V_3) + \frac{3}{2}H(V_4) + H(V_5) + H(V_6) + H(V_7). \tag{Q11}$$

B. Approximate Rate Region Characterization of Gaussian Asymmetric 3*-Description Coding*

In the following theorems, we establish outer and inner bounds for the rate region of the Gaussian asymmetric multiple descriptions coding.

Theorem 2: For a given distortion vector $\mathbf{D} = (D_{\Gamma_1}, \ldots, D_{\Gamma_1\Gamma_2\Gamma_3})$, denote by $\underline{\mathcal{R}}_{\text{MD}}(\mathbf{D})$ the set of all rate triples (R_1, R_2, R_3) satisfying

$$R_i \geq \frac{1}{2}\log\frac{1}{D_{\Gamma_i}}, \qquad i = 1,2,3 \tag{$\mathcal{O}-1$}$$

$$R_i + R_j \geq \min\left(\frac{1}{2}\log\frac{1}{D_{\Gamma_i}}, \frac{1}{2}\log\frac{1}{D_{\Gamma_j}}\right) + \frac{1}{2}\log\frac{1}{D_{\Gamma_i\Gamma_j}} - 1, \qquad i \neq j \tag{$\mathcal{O}-2$}$$

$$\begin{aligned}2R_i + R_j + R_k \geq &\min\left(\frac{1}{2}\log\frac{1}{D_{\Gamma_i}}, \frac{1}{2}\log\frac{1}{D_{\Gamma_j}}\right) + \min\left(\frac{1}{2}\log\frac{1}{D_{\Gamma_i}}, \frac{1}{2}\log\frac{1}{D_{\Gamma_k}}\right)\\ &+ \min\left(\frac{1}{2}\log\frac{1}{D_{\Gamma_i\Gamma_j}}, \frac{1}{2}\log\frac{1}{D_{\Gamma_i\Gamma_k}}\right) + \frac{1}{2}\log\frac{1}{D_{\Gamma_i\Gamma_j\Gamma_k}} - 3, \qquad i \neq j \neq k\end{aligned} \tag{$\mathcal{O}-3$}$$

$$\begin{aligned}R_1 + R_2 + R_3 \geq &\frac{1}{2}\log\frac{1}{D_{\Gamma_1}} + \min\left(\frac{1}{2}\log\frac{1}{D_{\Gamma_1\Gamma_2}}, \frac{1}{2}\log\frac{1}{D_{\Gamma_3}}\right)\\ &+ \frac{1}{2}\log\frac{1}{D_{\Gamma_1\Gamma_2\Gamma_3}} - 2,\end{aligned} \tag{$\mathcal{O}-4$}$$

$$\begin{aligned}R_1 + R_2 + R_3 \geq &\frac{1}{4}\log\frac{1}{D_{\Gamma_1}^2 D_{\Gamma_2}} + \min\left(\frac{1}{2}\log\frac{1}{D_{\Gamma_1\Gamma_2}}, \frac{1}{2}\log\frac{1}{D_{\Gamma_1\Gamma_3}}, \frac{1}{2}\log\frac{1}{D_{\Gamma_2\Gamma_3}}\right)\\ &+ \frac{1}{2}\log\frac{1}{D_{\Gamma_1\Gamma_2\Gamma_3}} - \frac{9}{2}.\end{aligned} \tag{$\mathcal{O}-5$}$$

Then any admissible rate triple belongs to $\underline{\mathcal{R}}_{\text{MD}}(\mathbf{D})$, *i.e.*, $\mathcal{R}_{\text{MD}}(\mathbf{D}) \subseteq \underline{\mathcal{R}}_{\text{MD}}(\mathbf{D})$.

The bound stated in this theorem is a consequence of a more general parametric outer bound $\underline{\mathcal{R}}^p_{\text{MD}}(\mathbf{D}, \mathbf{d})$, defined in Theorem 4. However, the current form is more convenient for comparison between the inner and outer bounds. This region is given in the following corollary for the specific ordering $\mathscr{L}_1$.

***Corollary** 2:* Any admissible rate triple for a three-description A-MD with $\mathscr{L}_1$ ordering satisfies

$$R_1 \geq \frac{1}{2}\log\frac{1}{D_{\Gamma_1}}, \qquad (\mathcal{O}'-1)$$

$$R_2 \geq \frac{1}{2}\log\frac{1}{D_{\Gamma_2}}, \qquad (\mathcal{O}'-2)$$

$$R_3 \geq \frac{1}{2}\log\frac{1}{D_{\Gamma_3}}, \qquad (\mathcal{O}'-3)$$

$$R_1 + R_2 \geq \frac{1}{2}\log\frac{1}{D_{\Gamma_1}D_{\Gamma_1\Gamma_2}} - 1, \qquad (\mathcal{O}'-4)$$

$$R_1 + R_3 \geq \frac{1}{2}\log\frac{1}{D_{\Gamma_1}D_{\Gamma_1\Gamma_3}} - 1, \qquad (\mathcal{O}'-5)$$

$$R_2 + R_3 \geq \frac{1}{2}\log\frac{1}{D_{\Gamma_2}D_{\Gamma_2\Gamma_3}} - 1, \qquad (\mathcal{O}'-6)$$

$$2R_1 + R_2 + R_3 \geq \frac{1}{2}\log\frac{1}{D_{\Gamma_1}^2 D_{\Gamma_1\Gamma_2}D_{\Gamma_1\Gamma_2\Gamma_3}} - 3, \qquad (\mathcal{O}'-7)$$

$$R_1 + 2R_2 + R_3 \geq \frac{1}{2}\log\frac{1}{D_{\Gamma_1}D_{\Gamma_2}D_{\Gamma_1\Gamma_2}D_{\Gamma_1\Gamma_2\Gamma_3}} - 3, \qquad (\mathcal{O}'-8)$$

$$R_1 + R_2 + 2R_3 \geq \frac{1}{2}\log\frac{1}{D_{\Gamma_1}D_{\Gamma_2}D_{\Gamma_1\Gamma_3}D_{\Gamma_1\Gamma_2\Gamma_3}} - 3, \qquad (\mathcal{O}'-9)$$

$$R_1 + R_2 + R_3 \geq \frac{1}{2}\log\frac{1}{D_{\Gamma_1}D_{\Gamma_3}D_{\Gamma_1\Gamma_2\Gamma_3}} - \frac{9}{2}, \qquad (\mathcal{O}'-10)$$

$$R_1 + R_2 + R_3 \geq \frac{1}{4}\log\frac{1}{D_{\Gamma_1}^2 D_{\Gamma_2}D_{\Gamma_1\Gamma_2}D_{\Gamma_1\Gamma_2\Gamma_3}^2} - 2. \qquad (\mathcal{O}'-11)$$

Theorem 3 gives an inner bound for the admissible rate region of the three-description A-MD problem.

***Theorem** 3:* For a given distortion vector $\mathbf{D} = (D_{\Gamma_1}, \ldots, D_{\Gamma_1\Gamma_2\Gamma_3})$, let $\overline{\mathcal{R}}_{\text{MD}}(\mathbf{D})$ be the set of all rate triples (R_1, R_2, R_3) satisfying

$$R_i \geq \frac{1}{2}\log\frac{1}{D_{\Gamma_i}}, \qquad i = 1,2,3 \tag{$\mathcal{I}-1$}$$

$$R_i + R_j \geq \min\left(\frac{1}{2}\log\frac{1}{D_{\Gamma_i}}, \frac{1}{2}\log\frac{1}{D_{\Gamma_j}}\right) + \frac{1}{2}\log\frac{1}{D_{\Gamma_i\Gamma_j}}, \qquad i \neq j \tag{$\mathcal{I}-2$}$$

$$\begin{aligned}2R_i + R_j + R_k \geq& \min\left(\frac{1}{2}\log\frac{1}{D_{\Gamma_i}}, \frac{1}{2}\log\frac{1}{D_{\Gamma_j}}\right) + \min\left(\frac{1}{2}\log\frac{1}{D_{\Gamma_i}}, \frac{1}{2}\log\frac{1}{D_{\Gamma_k}}\right)\\ &+ \min\left(\frac{1}{2}\log\frac{1}{D_{\Gamma_i\Gamma_j}}, \frac{1}{2}\log\frac{1}{D_{\Gamma_i\Gamma_k}}\right) + \frac{1}{2}\log\frac{1}{D_{\Gamma_i\Gamma_j\Gamma_k}}, \qquad i \neq j \neq k\end{aligned} \tag{$\mathcal{I}-3$}$$

$$R_1 + R_2 + R_3 \geq \frac{1}{2}\log\frac{1}{D_{\Gamma_1}} + \min\left(\frac{1}{2}\log\frac{1}{D_{\Gamma_1\Gamma_2}}, \frac{1}{2}\log\frac{1}{D_{\Gamma_3}}\right) + \frac{1}{2}\log\frac{1}{D_{\Gamma_1\Gamma_2\Gamma_3}}, \tag{$\mathcal{I}-4$}$$

$$\begin{aligned}R_1 + R_2 + R_3 \geq& \frac{1}{4}\log\frac{1}{D_{\Gamma_1}^2 D_{\Gamma_2}} + \min\left(\frac{1}{2}\log\frac{1}{D_{\Gamma_1\Gamma_2}}, \frac{1}{2}\log\frac{1}{D_{\Gamma_1\Gamma_3}}, \frac{1}{2}\log\frac{1}{D_{\Gamma_2\Gamma_3}}\right)\\ &+ \frac{1}{2}\log\frac{1}{D_{\Gamma_1\Gamma_2\Gamma_3}}.\end{aligned} \tag{$\mathcal{I}-5$}$$

Then any rate triple $\mathbf{R} \in \overline{\mathcal{R}}_{\text{MD}}$ is achievable, *i.e,* $\overline{\mathcal{R}}_{\text{MD}} \subseteq \mathcal{R}_{\text{MD}}$.

The following corollary specifies the above theorem for the ordering level $\mathscr{L}_1$.

Corollary *3:* If the distortion constraints satisfy the ordering level $\mathscr{L}_1$, *i.e.*,

$$D_{\Gamma_1} \geq D_{\Gamma_2} \geq D_{\Gamma_3} \geq D_{\Gamma_1\Gamma_2} \geq D_{\Gamma_1\Gamma_3} \geq D_{\Gamma_2\Gamma_3} \geq D_{\Gamma_1\Gamma_2\Gamma_3},$$

then any rate triple (R_1, R_2, R_3) satisfying

$$R_1 \geq \frac{1}{2} \log \frac{1}{D_{\Gamma_1}}, \tag{$\mathcal{I}'-1$}$$

$$R_2 \geq \frac{1}{2} \log \frac{1}{D_{\Gamma_2}}, \tag{$\mathcal{I}'-2$}$$

$$R_3 \geq \frac{1}{2} \log \frac{1}{D_{\Gamma_3}}, \tag{$\mathcal{I}'-3$}$$

$$R_1 + R_2 \geq \frac{1}{2} \log \frac{1}{D_{\Gamma_1} D_{\Gamma_1\Gamma_2}}, \tag{$\mathcal{I}'-4$}$$

$$R_1 + R_3 \geq \frac{1}{2} \log \frac{1}{D_{\Gamma_1} D_{\Gamma_1\Gamma_3}}, \tag{$\mathcal{I}'-5$}$$

$$R_2 + R_3 \geq \frac{1}{2} \log \frac{1}{D_{\Gamma_2} D_{\Gamma_2\Gamma_3}}, \tag{$\mathcal{I}'-6$}$$

$$2R_1 + R_2 + R_3 \geq \frac{1}{2} \log \frac{1}{D_{\Gamma_1}^2 D_{\Gamma_1\Gamma_2} D_{\Gamma_1\Gamma_2\Gamma_3}}, \tag{$\mathcal{I}'-7$}$$

$$R_1 + 2R_2 + R_3 \geq \frac{1}{2} \log \frac{1}{D_{\Gamma_1} D_{\Gamma_2} D_{\Gamma_1\Gamma_2} D_{\Gamma_1\Gamma_2\Gamma_3}}, \tag{$\mathcal{I}'-8$}$$

$$R_1 + R_2 + 2R_3 \geq \frac{1}{2} \log \frac{1}{D_{\Gamma_1} D_{\Gamma_3}^2 D_{\Gamma_1\Gamma_2\Gamma_3}}, \tag{$\mathcal{I}'-9$}$$

$$R_1 + R_2 + R_3 \geq \frac{1}{2} \log \frac{1}{D_{\Gamma_1} D_{\Gamma_3} D_{\Gamma_1\Gamma_2\Gamma_3}}, \tag{$\mathcal{I}'-10$}$$

$$R_1 + R_2 + R_3 \geq \frac{1}{4} \log \frac{1}{D_{\Gamma_1}^2 D_{\Gamma_2} D_{\Gamma_1\Gamma_2} D_{\Gamma_1\Gamma_2\Gamma_3}^2}, \tag{$\mathcal{I}'-11$}$$

is achievable.

Summarizing the results of Theorem 2 and Theorem 3 gives the following corollary.

***Corollary** 4:*

$$\overline{\mathcal{R}}_{\text{MD}}(\mathbf{D}) \subseteq \mathcal{R}_{\text{MD}}(\mathbf{D}) \subseteq \underline{\mathcal{R}}_{\text{MD}}(\mathbf{D}). \tag{14}$$

The result of this corollary is that the multiple description admissible rate region is bounded between two sets of hyperplanes, which are pair-wise parallel. For each pair of parallel planes, we can compute the distance between them. Denote by $\delta_{(x,y,z)}$ the Euclidean distance between two parallel planes which are orthogonal to the vector (x, y, z). Then for the distortion constraints corresponding to ordering $\mathscr{L}_1$, we have

$$\delta_{(1,0,0)} = 0, \tag{15}$$

$$\delta_{(1,1,0)} \leq \frac{1}{\sqrt{2}} = 0.7071, \tag{16}$$

$$\delta_{(2,1,1)} \leq \frac{3}{\sqrt{6}} = 1.2247, \tag{17}$$

$$\delta_{(1,1,1)} \leq \frac{9}{4\sqrt{3}} = 1.2990, \tag{18}$$

where the denominators are the normalizing factors, corresponding to the length of the vector (x, y, z). This shows that the inner and outer bounds provide an approximate characterization for the admissible rate region, for which the Euclidean distance between the bounds in less than 1.3 in the worst case. Fig. 4 shows a typical pair of inner and outer bounds for $\mathscr{L}_1$ ordering and the case $D_{\Gamma_2} D_{\Gamma_1\Gamma_3} \leq D^2_{\Gamma_3} \leq D_{\Gamma_2} D_{\Gamma_1\Gamma_2}$, which is the lossy counterpart of the lossless A-MLD problem with $h_4 \leq h_3 \leq h_4 + h_5$, discussed in Subsection IV-B, under regime II (see also Fig. 7).

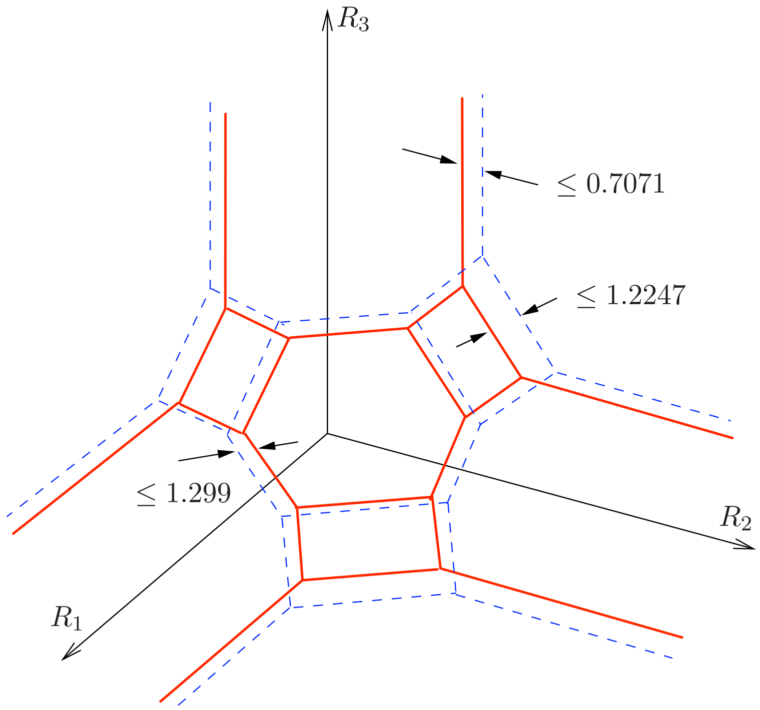

Fig. 4. The inner and outer bound for the admissible rate region of the Gaussian multiple descriptions problem for distortion constraints corresponding to the ordering $\mathscr{L}_1$, and the case $D_{\Gamma_2} D_{\Gamma_1\Gamma_3} \leq D^2_{\Gamma_3} \leq D_{\Gamma_2} D_{\Gamma_1\Gamma_2}$.

IV. Asymmetric Multilevel Diversity Coding

In this section we first prove the converse part of Theorem 1 for all orderings, and then the achievability part for ordering $\mathscr{L}_1$. Similar techniques can be used straightforwardly to prove the achievability for all the other orderings, and therefore complete the proof of Theorem 1.

A. The Converse Proof

In this subsection we show that any admissible rate triple satisfies $(P1)$-$(P5)$. The following important lemma, which simplifies the proof of the theorem, relates the entropy of the original source to the reconstructed one.

***Lemma** 2:* Let $\mathcal{S} \subseteq \{\Gamma_1, \Gamma_2, \Gamma_3\}$ be a subset of descriptions available at a decoder, and $i \leq j \leq \mathscr{L}(\mathcal{S})$. Then

$$H(\mathcal{S}|U_i^n) \geq H(\mathcal{S}|U_j^n) + n(H_j - H_i - \delta_n) \tag{19}$$

where $\delta_n \to 0$ as n increases.

Proof: Note that $i \leq j \leq \mathscr{L}(\mathcal{S})$. Therefore, the decoding requirement for the decoder with access to $\mathcal{S}$ implies that the reconstructed sequence $\hat{U}_j^n(\mathcal{S})$ equals to U_j^n with high probability. Then

$$\begin{aligned} H(\mathcal{S}|U_i^n) &\overset{(a)}{=} H(\mathcal{S}, \hat{U}_j^n(\mathcal{S})|U_i^n) \\ &= H(\mathcal{S}, U_j^n, \hat{U}_j^n(\mathcal{S})|U_i^n) - H(U_j^n|\mathcal{S}, \hat{U}_j^n(\mathcal{S}), U_i^n) \\ &\geq H(\mathcal{S}, U_j^n|U_i^n) - H(U_j^n|\hat{U}_j^n(\mathcal{S})) \\ &\overset{(b)}{=} H(\mathcal{S}|U_j^n) + H(U_j^n|U_i^n) - H(U_j^n|\hat{U}_j^n(\mathcal{S})), \end{aligned} \tag{20}$$

where (a) holds since $\hat{U}_j^n(\mathcal{S})$ is function of $\mathcal{S}$, for $j \leq \mathscr{L}(\mathcal{S})$, and (b) is due to the fact that U_i^n is a subsequence of U_j^n for $j \geq i$. The underlying distribution of U_i^n and U_j^n implies $H(U_j^n|U_i^n) = n(H_j - H_i)$. The last term in (20) can be upper bounded using the Fano's inequality [12, page ...] as

$$H(U_j^n|\hat{U}_j^n(\mathcal{S})) \leq h_B(P_e) + P_e \log(|\mathcal{U}_j^n| - 1) \leq 1 + ncP_e \tag{21}$$

where $P_e = \Pr(\hat{U}_j^n(\mathcal{S}) \neq U_j^n) < \varepsilon$, $h_B(p)$ defined as

$$h_B(p) = -p \log_2(p) - (1-p) \log_2(1-p)$$

is the binary entropy function, and $c = \log |\mathcal{U}_j|$ is a constant. The proof is complete by setting $\delta_n = \frac{1}{n} + cP_e$. ■

The converse proof of Theorem 1: Let (R_1, R_2, R_3) be any admissible rate triple, and Γ_i be a single description with ordering level $\mathscr{L}(\Gamma_i)$. Recall $U_0^n = 0$, and note that $\hat{U}^n_{\mathscr{L}(\Gamma_i)}(\Gamma_i)$ is a function of Γ_i. Thus

$$n(R_i + \varepsilon) \geq H(\Gamma_i) = H(\Gamma_i | U_0^n) \overset{(\star)}{\geq} H(\Gamma_i | U^n_{\mathscr{L}(\Gamma_i)}) + n(H_{\mathscr{L}(\Gamma_i)} - \delta_n) \geq n(H_{\mathscr{L}(\Gamma_i)} - \delta_n). \tag{22}$$

This proves $(P1)$. Note that here and in the rest of this proof all the inequalities labeled by $(\star)$ are due to Lemma 2.

Toward proving $(P2)$, we can write

$$\begin{aligned}
n(R_i + R_j + 2\varepsilon) &\geq H(\Gamma_i) + H(\Gamma_j) \\
&\geq H(\Gamma_i | U_0^n) + H(\Gamma_j | U_0^n) \\
&\overset{(\star)}{\geq} H(\Gamma_i | U^n_{\mathscr{L}(\Gamma_i)}) + n(H_{\mathscr{L}(\Gamma_i)} - \delta_n) + H(\Gamma_j | U^n_{\mathscr{L}(\Gamma_j)}) + n(H_{\mathscr{L}(\Gamma_j)} - \delta_n) \\
&\geq n(H_{\mathscr{L}(\Gamma_i)} + H_{\mathscr{L}(\Gamma_j)} - 2\delta_n) + H(\Gamma_i | U^n_{\max\{\mathscr{L}(\Gamma_i), \mathscr{L}(\Gamma_j)\}}) + H(\Gamma_i | U^n_{\max\{\mathscr{L}(\Gamma_i), \mathscr{L}(\Gamma_j)\}}) \\
&\geq n(H_{\mathscr{L}(\Gamma_i)} + H_{\mathscr{L}(\Gamma_j)} - 2\delta_n) + H(\Gamma_i, \Gamma_j | U^n_{\max\{\mathscr{L}(\Gamma_i), \mathscr{L}(\Gamma_j)\}}) \\
&\overset{(\star)}{\geq} n(H_{\mathscr{L}(\Gamma_i)} + H_{\mathscr{L}(\Gamma_j)} - 2\delta_n) + n(H_{\mathscr{L}(\Gamma_i, \Gamma_j)} - H_{\max\{\mathscr{L}(\Gamma_i), \mathscr{L}(\Gamma_j)\}} - \delta_n) \\
&\quad + H(\Gamma_i, \Gamma_j | U^n_{\mathscr{L}(\Gamma_i, \Gamma_j)}) \\
&\geq n\left[H_{\min\{\mathscr{L}(\Gamma_i), \mathscr{L}(\Gamma_j)\}} + H_{\mathscr{L}(\Gamma_i, \Gamma_j)} - 3\delta_n\right].
\end{aligned} \tag{23}$$

For proving $(P3)$ we can start with

$$
\begin{aligned}
n(2R_i + R_j{+}R_k + 4\varepsilon) &\geq 2H(\Gamma_i) + H(\Gamma_j) + H(\Gamma_k) \\
&\overset{(\star)}{\geq} [H(\Gamma_i|U^n_{\mathscr{L}(\Gamma_i)}) + nH_{\mathscr{L}(\Gamma_i)} - n\delta_n + H(\Gamma_j|U^n_{\mathscr{L}(\Gamma_j)}) + nH_{\mathscr{L}(\Gamma_j)} - n\delta_n] \\
&\quad + [H(\Gamma_i|U^n_{\mathscr{L}(\Gamma_i)}) + nH_{\mathscr{L}(\Gamma_i)} - n\delta_n + H(\Gamma_k|U^n_{\mathscr{L}(\Gamma_k)}) + nH_{\mathscr{L}(\Gamma_k)} - n\delta_n] \\
&\geq n(2H_{\mathscr{L}(\Gamma_i)} + H_{\mathscr{L}(\Gamma_j)} + H_{\mathscr{L}(\Gamma_k)} - 4\delta_n) + H(\Gamma_i, \Gamma_j|U^n_{\max\{\mathscr{L}(\Gamma_i),\mathscr{L}(\Gamma_j)\}}) \\
&\quad + H(\Gamma_i, \Gamma_k|U^n_{\max\{\mathscr{L}(\Gamma_i),\mathscr{L}(\Gamma_k)\}}) \\
&\overset{(\star)}{\geq} n(2H_{\mathscr{L}(\Gamma_i)} + H_{\mathscr{L}(\Gamma_j)} + H_{\mathscr{L}(\Gamma_k)} - 4\delta_n) \\
&\quad + n(H_{\mathscr{L}(\Gamma_i,\Gamma_j)} - H_{\max\{\mathscr{L}(\Gamma_i),\mathscr{L}(\Gamma_j)\}} - \delta_n) + H(\Gamma_i, \Gamma_j|U^n_{\mathscr{L}(\Gamma_i,\Gamma_j)}) \\
&\quad + n(H_{\mathscr{L}(\Gamma_i,\Gamma_k)} - H_{\max\{\mathscr{L}(\Gamma_i),\mathscr{L}(\Gamma_k)\}} - \delta_n) + H(\Gamma_i, \Gamma_k|U^n_{\mathscr{L}(\Gamma_i,\Gamma_k)}) \\
&\geq n(H_{\min\{\mathscr{L}(\Gamma_i),\mathscr{L}(\Gamma_j)\}} + H_{\min\{\mathscr{L}(\Gamma_i),\mathscr{L}(\Gamma_k)\}} + H_{\mathscr{L}(\Gamma_i,\Gamma_j)} + H_{\mathscr{L}(\Gamma_i,\Gamma_k)} - 6\delta_n) \\
&\quad + H(\Gamma_i, \Gamma_j, \Gamma_k|U^n_{\max\{\mathscr{L}(\Gamma_i,\Gamma_j),\mathscr{L}(\Gamma_i,\Gamma_k)\}}) \\
&\overset{(\star)}{\geq} n(H_{\min\{\mathscr{L}(\Gamma_i),\mathscr{L}(\Gamma_j)\}} + H_{\min\{\mathscr{L}(\Gamma_i),\mathscr{L}(\Gamma_k)\}} + H_{\mathscr{L}(\Gamma_i,\Gamma_j)} + H_{\mathscr{L}(\Gamma_i,\Gamma_k)} - 6\delta_n) \\
&\quad + n(H_{\mathscr{L}(\Gamma_i,\Gamma_j,\Gamma_k)} - H_{\max\{\mathscr{L}(\Gamma_i,\Gamma_j),\mathscr{L}(\Gamma_i,\Gamma_k)\}} - \delta_n) \\
&= n\big[H_{\min\{\mathscr{L}(\Gamma_i),\mathscr{L}(\Gamma_j)\}} + H_{\min\{\mathscr{L}(\Gamma_i),\mathscr{L}(\Gamma_k)\}} \\
&\quad + H_{\min\{\mathscr{L}(\Gamma_i,\Gamma_j),\mathscr{L}(\Gamma_i,\Gamma_k)\}} + H_{\mathscr{L}(\Gamma_i,\Gamma_j,\Gamma_k)} - 7\delta_n\big]. \qquad (24)
\end{aligned}
$$

Toward proving $(P4)$ we can write

$$
\begin{aligned}
n(R_1+R_2+R_3+3\varepsilon) &\geq H(\Gamma_1)+H(\Gamma_2)+H(\Gamma_3) \\
&\overset{(\star)}{\geq} n(H_{\mathscr{L}(\Gamma_1)}+H_{\mathscr{L}(\Gamma_2)}+H_{\min\{\mathscr{L}(\Gamma_1,\Gamma_2),\mathscr{L}(\Gamma_3)\}}-3\delta_n) \\
&\quad + H(\Gamma_1|U^n_{\mathscr{L}(\Gamma_1)})+H(\Gamma_2|U^n_{\mathscr{L}(\Gamma_2)})+H(\Gamma_3|U^n_{\min\{\mathscr{L}(\Gamma_1,\Gamma_2),\mathscr{L}(\Gamma_3)\}}) \\
&\geq n(H_{\mathscr{L}(\Gamma_1)}+H_{\mathscr{L}(\Gamma_2)}+H_{\min\{\mathscr{L}(\Gamma_1,\Gamma_2),\mathscr{L}(\Gamma_3)\}}-3\delta_n) \\
&\quad + H(\Gamma_1,\Gamma_2|U^n_{\mathscr{L}(\Gamma_2)})+H(\Gamma_3|U^n_{\min\{\mathscr{L}(\Gamma_1,\Gamma_2),\mathscr{L}(\Gamma_3)\}}) \\
&\overset{(\star)}{\geq} n(H_{\mathscr{L}(\Gamma_1)}+H_{\mathscr{L}(\Gamma_2)}+H_{\min\{\mathscr{L}(\Gamma_1,\Gamma_2),\mathscr{L}(\Gamma_3)\}}-3\delta_n) \\
&\quad + H(\Gamma_1,\Gamma_2|U^n_{\min\{\mathscr{L}(\Gamma_1,\Gamma_2),\mathscr{L}(\Gamma_3)\}})+n(H_{\min\{\mathscr{L}(\Gamma_1,\Gamma_2),\mathscr{L}(\Gamma_3)\}}-H_{\mathscr{L}(\Gamma_2)}-\delta_n) \\
&\quad + H(\Gamma_3|U^n_{\min\{\mathscr{L}(\Gamma_1,\Gamma_2),\mathscr{L}(\Gamma_3)\}}) \\
&\geq n(H_{\mathscr{L}(\Gamma_1)}+2H_{\min\{\mathscr{L}(\Gamma_1,\Gamma_2),\mathscr{L}(\Gamma_3)\}}-4\delta_n) \\
&\quad + H(\Gamma_1,\Gamma_2,\Gamma_3|U^n_{\min\{\mathscr{L}(\Gamma_1,\Gamma_2),\mathscr{L}(\Gamma_3)\}}) \\
&\overset{(\star)}{\geq} n(H_{\mathscr{L}(\Gamma_1)}+2H_{\min\{\mathscr{L}(\Gamma_1,\Gamma_2),\mathscr{L}(\Gamma_3)\}}-4\delta_n) \\
&\quad + n(H_{\mathscr{L}(\Gamma_1,\Gamma_2,\Gamma_3)}-H_{\min\{\mathscr{L}(\Gamma_1,\Gamma_2),\mathscr{L}(\Gamma_3)\}}-\delta_n) \\
&\geq n\left[H_{\mathscr{L}(\Gamma_1)}+H_{\min\{\mathscr{L}(\Gamma_1,\Gamma_2),\mathscr{L}(\Gamma_3)\}}+H_{\mathscr{L}(\Gamma_1,\Gamma_2,\Gamma_3)}-5\delta_n\right]. \qquad (25)
\end{aligned}
$$

We need to consider two different cases in order to obtain the other sum-rate bound in $(P5)$. First consider the case $\mathscr{L}(\Gamma_3)>\mathscr{L}(\Gamma_1,\Gamma_2)$. Note that this implies $\min\{\mathscr{L}(\Gamma_1,\Gamma_2),\mathscr{L}(\Gamma_1,\Gamma_3),\mathscr{L}(\Gamma_2,\Gamma_3)\}=$

$\mathscr{L}(\Gamma_1,\Gamma_2)$. We have

$$
\begin{aligned}
n(R_1+&R_2+R_3+3\varepsilon) \geq H(\Gamma_1)+H(\Gamma_2)+H(\Gamma_3) \\
&\overset{(\star)}{\geq} n(H_{\mathscr{L}(\Gamma_1)}+H_{\mathscr{L}(\Gamma_2)}+H_{\mathscr{L}(\Gamma_3)}-3\delta_n)+H(\Gamma_1|U^n_{\mathscr{L}(\Gamma_1)})+H(\Gamma_2|U^n_{\mathscr{L}(\Gamma_2)})+H(\Gamma_3|U^n_{\mathscr{L}(\Gamma_3)}) \\
&= n(H_{\mathscr{L}(\Gamma_1)}+H_{\mathscr{L}(\Gamma_2)}+H_{\mathscr{L}(\Gamma_3)}-3\delta_n)+\frac{1}{2}[H(\Gamma_1|U^n_{\mathscr{L}(\Gamma_1)})+H(\Gamma_2|U^n_{\mathscr{L}(\Gamma_2)})] \\
&\quad\quad +\frac{1}{2}[H(\Gamma_1|U^n_{\mathscr{L}(\Gamma_1)})+H(\Gamma_3|U^n_{\mathscr{L}(\Gamma_3)})]+\frac{1}{2}[H(\Gamma_2|U^n_{\mathscr{L}(\Gamma_2)})+H(\Gamma_3|U^n_{\mathscr{L}(\Gamma_3)})] \\
&\geq n(H_{\mathscr{L}(\Gamma_1)}+H_{\mathscr{L}(\Gamma_2)}+H_{\mathscr{L}(\Gamma_3)}-3\delta_n) \\
&\quad +\frac{1}{2}[H(\Gamma_1,\Gamma_2|U^n_{\mathscr{L}(\Gamma_2)})+H(\Gamma_1,\Gamma_3|U^n_{\mathscr{L}(\Gamma_3)})+H(\Gamma_2,\Gamma_3|U^n_{\mathscr{L}(\Gamma_3)})] \qquad (26) \\
&\overset{(\star)}{\geq} n(H_{\mathscr{L}(\Gamma_1)}+H_{\mathscr{L}(\Gamma_2)}+H_{\mathscr{L}(\Gamma_3)}-3\delta_n) \\
&\quad +\frac{1}{2}\left[H(\Gamma_1,\Gamma_2|U^n_{\mathscr{L}(\Gamma_1,\Gamma_2)})+n(H_{\mathscr{L}(\Gamma_1,\Gamma_2)}-H_{\mathscr{L}(\Gamma_2)}-\delta_n)\right] \\
&\quad +\frac{1}{2}\left[H(\Gamma_1,\Gamma_3|U^n_{\mathscr{L}(\Gamma_3)})+H(\Gamma_2,\Gamma_3|U^n_{\mathscr{L}(\Gamma_3)})\right] \\
&\overset{(a)}{\geq} n(H_{\mathscr{L}(\Gamma_1)}+\frac{1}{2}H_{\mathscr{L}(\Gamma_2)}+\frac{1}{2}H_{\mathscr{L}(\Gamma_1,\Gamma_2)}+H_{\mathscr{L}(\Gamma_3)}-\frac{7}{2}\delta_n) \\
&\quad +\frac{1}{2}\left[H(\Gamma_1,\Gamma_2|U^n_{\mathscr{L}(\Gamma_3)})+H(\Gamma_1,\Gamma_3|U^n_{\mathscr{L}(\Gamma_3)})+H(\Gamma_2,\Gamma_3|U^n_{\mathscr{L}(\Gamma_3)})\right] \\
&\overset{(b)}{\geq} n(H_{\mathscr{L}(\Gamma_1)}+\frac{1}{2}H_{\mathscr{L}(\Gamma_2)}+\frac{1}{2}H_{\mathscr{L}(\Gamma_1,\Gamma_2)}+H_{\mathscr{L}(\Gamma_3)}-\frac{7}{2}\delta_n)+H(\Gamma_1,\Gamma_2,\Gamma_3|U^n_{\mathscr{L}(\Gamma_3)}) \\
&\overset{(\star)}{\geq} n(H_{\mathscr{L}(\Gamma_1)}+\frac{1}{2}H_{\mathscr{L}(\Gamma_2)}+\frac{1}{2}H_{\mathscr{L}(\Gamma_1,\Gamma_2)}+H_{\mathscr{L}(\Gamma_3)}-\frac{7}{2}\delta_n)+n(H_{\mathscr{L}(\Gamma_1,\Gamma_2,\Gamma_3)}-H_{\mathscr{L}(\Gamma_3)}-\delta_n) \\
&= n\left[H_{\mathscr{L}(\Gamma_1)}+\frac{1}{2}H_{\mathscr{L}(\Gamma_2)}+\frac{1}{2}H_{\mathscr{L}(\Gamma_1,\Gamma_2)}+H_{\mathscr{L}(\Gamma_1,\Gamma_2,\Gamma_3)}-\frac{9}{2}\delta_n\right], \qquad (27)
\end{aligned}
$$

where in (a) we have used $H(\Gamma_1,\Gamma_2|U^n_{\mathscr{L}(\Gamma_1,\Gamma_2)}) \geq H(\Gamma_1,\Gamma_2|U^n_{\mathscr{L}(\Gamma_3)})$, implied by the assumption $\mathscr{L}(\Gamma_3) > \mathscr{L}(\Gamma_1,\Gamma_2)$, and (b) is due to the conditional version of Han's inequality [12, page 491].

For the second case, *i.e.,* $\mathscr{L}(\Gamma_3) < \mathscr{L}(\Gamma_1, \Gamma_2)$, we have from (26),

$$\begin{aligned}
n(R_1{+}R_2 + R_3 + 3\varepsilon) &\geq n(H_{\mathscr{L}(\Gamma_1)} + H_{\mathscr{L}(\Gamma_2)} + H_{\mathscr{L}(\Gamma_3)} - 3\delta_n) \\
&\quad + \frac{1}{2}\left[H(\Gamma_1,\Gamma_2|U^n_{\mathscr{L}(\Gamma_2)}) + H(\Gamma_1,\Gamma_3|U^n_{\mathscr{L}(\Gamma_3)}) + H(\Gamma_2,\Gamma_3|U^n_{\mathscr{L}(\Gamma_3)})\right] \\
&\overset{(\star)}{\geq} n(H_{\mathscr{L}(\Gamma_1)} + H_{\mathscr{L}(\Gamma_2)} + H_{\mathscr{L}(\Gamma_3)} - 3\delta_n) + \frac{1}{2}\Big[H(\Gamma_1,\Gamma_2|U^n_{\min\{\mathscr{L}(\Gamma_1,\Gamma_2),\mathscr{L}(\Gamma_1,\Gamma_3),\mathscr{L}(\Gamma_2,\Gamma_3)\}}) \\
&\quad + H(\Gamma_1,\Gamma_3|U^n_{\min\{\mathscr{L}(\Gamma_1,\Gamma_2),\mathscr{L}(\Gamma_1,\Gamma_3),\mathscr{L}(\Gamma_2,\Gamma_3)\}}) + H(\Gamma_2,\Gamma_3|U^n_{\min\{\mathscr{L}(\Gamma_1,\Gamma_2),\mathscr{L}(\Gamma_1,\Gamma_3),\mathscr{L}(\Gamma_2,\Gamma_3)\}}) \\
&\quad + n(3H_{\min\{\mathscr{L}(\Gamma_1,\Gamma_2),\mathscr{L}(\Gamma_1,\Gamma_3),\mathscr{L}(\Gamma_2,\Gamma_3)\}} - H_{\mathscr{L}(\Gamma_2)} - 2H_{\mathscr{L}(\Gamma_3)} - 3\delta_n)\Big] \\
&\overset{(c)}{\geq} n(H_{\mathscr{L}(\Gamma_1)} + \frac{1}{2}H_{\mathscr{L}(\Gamma_2)} + \frac{3}{2}H_{\min\{\mathscr{L}(\Gamma_1,\Gamma_2),\mathscr{L}(\Gamma_1,\Gamma_3),\mathscr{L}(\Gamma_2,\Gamma_3)\}}) - \frac{9}{2}\delta_n) \\
&\quad + H(\Gamma_1,\Gamma_2,\Gamma_3|U^n_{\min\{\mathscr{L}(\Gamma_1,\Gamma_2),\mathscr{L}(\Gamma_1,\Gamma_3),\mathscr{L}(\Gamma_2,\Gamma_3)\}}) \\
&\geq n(H_{\mathscr{L}(\Gamma_1)} + \frac{1}{2}H_{\mathscr{L}(\Gamma_2)} + \frac{3}{2}H_{\min\{\mathscr{L}(\Gamma_1,\Gamma_2),\mathscr{L}(\Gamma_1,\Gamma_3),\mathscr{L}(\Gamma_2,\Gamma_3)\}} - \frac{9}{2}\delta_n) \\
&\quad + n(H_{\mathscr{L}(\Gamma_1,\Gamma_2,\Gamma_3)} - H_{\min\{\mathscr{L}(\Gamma_1,\Gamma_2),\mathscr{L}(\Gamma_1,\Gamma_3),\mathscr{L}(\Gamma_2,\Gamma_3)\}} - \delta_n) \\
&= n\left[H_{\mathscr{L}(\Gamma_1)} + \frac{1}{2}H_{\mathscr{L}(\Gamma_2)} + \frac{1}{2}H_{\min\{\mathscr{L}(\Gamma_1,\Gamma_2),\mathscr{L}(\Gamma_1,\Gamma_3),\mathscr{L}(\Gamma_2,\Gamma_3)\}} + H_{\mathscr{L}(\Gamma_1,\Gamma_2,\Gamma_3)} - \frac{11}{2}\delta_n\right]. \qquad (28)
\end{aligned}$$

Again we have used the conditional Han's inequality in (c). Putting (27) and (28) together, we obtain the bound $(P5)$. ■

B. Achievability

In the following we will show that the inequalities $(P1)$–$(P5)$ provide a complete characterization of the achievable rate region of the A-MLD problem. However, each individual case given in Table I needs to be considered separately, due to the specific strategy used in the coding scheme. For conciseness, we only present the analysis for the ordering level $\mathscr{L}_1$, and provide the details of the achievability scheme for this specific ordering. More precisely, we show that any rate triple (R_1, R_2, R_3) satisfying $(Q1)$–$(Q11)$ is achievable, *i.e.,* there exist encoding and decoding functions with the desired rates which are able to reconstruct the required subset of the sources from the corresponding descriptions. This implies $\mathcal{R}^{\mathscr{L}}_{\text{MLD}}(\mathbf{H})$ is achievable, and completes the proof of the theorem for the ordering $\mathscr{L}_1$. Similar proof for other orderings can be straightforwardly completed by applying almost identical techniques. Different cases that needed to be considered are listed in Table I.

Note that the $\mathcal{R}^{\mathscr{L}}_{\text{MLD}}$ is a polytopes specified by several hyperplanes in a three-dimensional space. Therefore, the region $\mathcal{R}^{\mathscr{L}}_{\text{MLD}}$ is a *convex* polytopes, and it suffices to show the achievability only for the

TABLE I

THE EIGHT POSSIBLE LEVEL ORDERINGS AND THE CORRESPONDING SUB-REGIMES.

Ordering	Regime
$\mathscr{L}(\Gamma_1) < \mathscr{L}(\Gamma_2) < \mathscr{L}(\Gamma_3) < \mathscr{L}(\Gamma_{12}) < \mathscr{L}(\Gamma_{13}) < \mathscr{L}(\Gamma_{23}) < \mathscr{L}(\Gamma_{123})$	$h_3 \leq h_4$
	$h_4 \leq h_3 \leq h_4 + h_5$
	$h_3 \geq h_4 + h_5$
$\mathscr{L}(\Gamma_1) < \mathscr{L}(\Gamma_2) < \mathscr{L}(\Gamma_3) < \mathscr{L}(\Gamma_{12}) < \mathscr{L}(\Gamma_{23}) < \mathscr{L}(\Gamma_{13}) < \mathscr{L}(\Gamma_{123})$	$h_3 \leq h_4$
	$h_4 \leq h_3 \leq h_4 + h_5$
	$h_3 \geq h_4 + h_5$
$\mathscr{L}(\Gamma_1) < \mathscr{L}(\Gamma_2) < \mathscr{L}(\Gamma_3) < \mathscr{L}(\Gamma_{13}) < \mathscr{L}(\Gamma_{12}) < \mathscr{L}(\Gamma_{23}) < \mathscr{L}(\Gamma_{123})$	$h_3 \leq h_4$
	$h_3 \geq h_4$
$\mathscr{L}(\Gamma_1) < \mathscr{L}(\Gamma_2) < \mathscr{L}(\Gamma_3) < \mathscr{L}(\Gamma_{13}) < \mathscr{L}(\Gamma_{23}) < \mathscr{L}(\Gamma_{12}) < \mathscr{L}(\Gamma_{123})$	$h_3 \leq h_4$
	$h_3 \geq h_4$
$\mathscr{L}(\Gamma_1) < \mathscr{L}(\Gamma_2) < \mathscr{L}(\Gamma_3) < \mathscr{L}(\Gamma_{23}) < \mathscr{L}(\Gamma_{12}) < \mathscr{L}(\Gamma_{13}) < \mathscr{L}(\Gamma_{123})$	$h_3 \leq h_4$
	$h_3 \geq h_4$
$\mathscr{L}(\Gamma_1) < \mathscr{L}(\Gamma_2) < \mathscr{L}(\Gamma_3) < \mathscr{L}(\Gamma_{23}) < \mathscr{L}(\Gamma_{13}) < \mathscr{L}(\Gamma_{12}) < \mathscr{L}(\Gamma_{123})$	$h_3 \leq h_4$
	$h_3 \geq h_4$
$\mathscr{L}(\Gamma_1) < \mathscr{L}(\Gamma_2) < \mathscr{L}(\Gamma_{12}) < \mathscr{L}(\Gamma_3) < \mathscr{L}(\Gamma_{13}) < \mathscr{L}(\Gamma_{23}) < \mathscr{L}(\Gamma_{123})$	$h_3 \leq h_5$
	$h_3 \geq h_5$
$\mathscr{L}(\Gamma_1) < \mathscr{L}(\Gamma_2) < \mathscr{L}(\Gamma_{12}) < \mathscr{L}(\Gamma_3) < \mathscr{L}(\Gamma_{23}) < \mathscr{L}(\Gamma_{13}) < \mathscr{L}(\Gamma_{123})$	$h_3 \leq h_5$
	$h_3 \geq h_5$

corner points [18]; that is because a simple *time-sharing* argument can be used to extend the achievability to any arbitrary point in the region $\mathcal{R}_{\text{MLD}}^{\mathscr{L}}$.

Depending on the relationship of h_3, h_4, and h_5, some of the inequalities in $(Q1)$–$(Q11)$ may be dominated by the others. Note that $(Q10)$ and $(Q11)$ are of the form

$$R_1 + R_2 + R_3 \geq 3H(V_1) + 2H(V_2) + 2H(V_3) + H(V_4) + H(V_5) + H(V_6) + H(V_7),$$
$$R_1 + R_2 + R_3 \geq 3H(V_1) + 2H(V_2) + \frac{3}{2}H(V_3) + \frac{3}{2}H(V_4) + H(V_5) + H(V_6) + H(V_7).$$

It is clear either one of them would be redundant and implied by the other, depending on whether $h_3 \lessgtr h_4$.

Also if $h_3 \geq h_4 + h_5$, inequalities ($Q3$) and ($Q10$) imply

$$\begin{aligned} R_1 + R_2 + 2R_3 &\geq H_1 + H_3 + H_7 + H_3 \\ &\geq H_1 + H_3 + H_7 + H_2 + h_4 + h_5 \\ &= H_1 + H_2 + H_5 + H_7, \end{aligned}$$

which is exactly the inequality given in ($Q9$), *i.e.,* this inequality is redundant in this regime. Thus, we split the achievability proof into three regimes corresponding to the aforementioned conditions, since the proposed encoding schemes are slightly different for these regimes. We show the achievability of the corner points in each case.

To simplify matters, we perform a lossless *pre-coding*, acting on all the seven source sequences V_i^n's as

$$E_i : \mathcal{V}_i^n \longrightarrow \{0,1\}^{\ell_i}$$

for $i = 1, \dots, 7$. This function maps the source sequence V_i^n to $\tilde{V}_i \triangleq E_i(V_i^n)$, which can be used as a new binary source sequence of length ℓ_i. This can be done by using any lossless scheme, and achieves ℓ_i arbitrary close to nh_i for large enough n. With the new source sequences $\tilde{V}_i$, we next perform further coding.

Regime I: $h_3 \geq h_4 + h_5$

As mentioned above, the inequalities ($Q9$) and ($Q11$) are dominated by the others in this regime. Therefore we only need to consider the remaining nine hyperplanes. In the following we list the corner points of $\mathcal{R}_{\text{MLD}}^{\mathscr{L}}$ in this regime. Each corner point with coordinates (R_1, R_2, R_3) is the intersection of (at least) three hyperplane, say (Qi), (Qj), and (Qk). Such point is denoted by $\langle Qi, Qj, Qk \rangle : (R_1, R_2, R_3)$. In order to list all the corner points, we first find the intersection of any three hyperplanes, and then check whether the intersection point satisfies all the other inequalities. We next provide an encoding strategy to achieve the rates prescribed by the corner points of the polytope.

- $X_1 = \langle Q1{,}Q4{,}Q7 \rangle : (H_1, H_4, H_7)$

  This corner point is the intersection of the planes Q_1, Q_4, and Q_7, and determines the individual rates of the descriptions as

  $$(R_1, R_2, R_3) = (H_1, H_4, H_7).$$

  The scheme for achieving this rate tuple is as follows. Γ_1 is exactly the pre-coded sequence of V_1^n, *i.e.,* $\tilde{V}_1$. In order to construct Γ_2 it suffices to concatenate the codewords $\tilde{V}_1$, $\tilde{V}_2$, $\tilde{V}_3$, and $\tilde{V}_4$.

Similarly, Γ_3 is the concatenation of all the seven codewords. That is,

$$\Gamma_1 : \tilde{V}_1, \qquad \Gamma_2 : \tilde{V}_1, \tilde{V}_2, \tilde{V}_3, \tilde{V}_4, \qquad \Gamma_3 : \tilde{V}_1, \tilde{V}_2, \tilde{V}_3, \tilde{V}_4, \tilde{V}_5, \tilde{V}_6, \tilde{V}_7.$$

It is easy to check that the description rates are the same as the rate triple of the corner point, and all the decoding requirements at the seven decoders are satisfied.

We will only determine the rate triples and illustrate the descriptions construction for the remaining corner points.

- $X_2 = \langle Q1,Q5,Q7 \rangle : (H_1, H_7 - h_5, H_5)$

$$\Gamma_1 : \tilde{V}_1, \qquad \Gamma_2 : \tilde{V}_1, \tilde{V}_2, \tilde{V}_3, \tilde{V}_4, \tilde{V}_6, \tilde{V}_7, \qquad \Gamma_3 : \tilde{V}_1, \tilde{V}_2, \tilde{V}_3, \tilde{V}_4, \tilde{V}_5.$$

- $X_3 = \langle Q2,Q4,Q8 \rangle : (H_1 + h_3 + h_4, H_2, H_7)$

$$\Gamma_1 : \tilde{V}_1, \tilde{V}_3, \tilde{V}_4, \qquad \Gamma_2 : \tilde{V}_1, \tilde{V}_2, \qquad \Gamma_3 : \tilde{V}_1, \tilde{V}_2, \tilde{V}_3, \tilde{V}_4, \tilde{V}_5, \tilde{V}_6, \tilde{V}_7.$$

- $X_4 = \langle Q2,Q6,Q8 \rangle : (H_1 + h_3 + h_4 + h_7, H_2, H_6)$

$$\Gamma_1 : \tilde{V}_1, \tilde{V}_3, \tilde{V}_4, \tilde{V}_7, \qquad \Gamma_2 : \tilde{V}_1, \tilde{V}_2, \qquad \Gamma_3 : \tilde{V}_1, \tilde{V}_2, \tilde{V}_3, \tilde{V}_4, \tilde{V}_5, \tilde{V}_6.$$

- $X_5 = \langle Q3,Q5,Q10 \rangle : (H_1 + h_4 + h_5, H_3 + h_6 + h_7, H_3)$

  The encoding schemes for the previous corner points only involve concatenation of different codewords. However, concatenation is not optimal to achieve the rate triple induced by the point X_5, and we need to jointly encode the sources to construct the descriptions. This can be done using a modulo-2 summation of (parts of) the codewords of the same size.

  The description Γ_1 is simply constructed by concatenating $\tilde{V}_1$, $\tilde{V}_4$, and $\tilde{V}_5$. Similarly, Γ_3 is obtained by putting $\tilde{V}_1$, $\tilde{V}_2$, and $\tilde{V}_3$ together. The second description, Γ_2, should be able to help Γ_1 to reconstruct $\tilde{V}_3$ at the decoder with access to $\{\Gamma_1, \Gamma_2\}$, and help Γ_3 to reconstruct $(\tilde{V}_4, \tilde{V}_5)$ at decoder $\{\Gamma_2, \Gamma_3\}$, where $\tilde{V}_3$ is already provided as a part of Γ_3. We can use this fact to construct Γ_2 as follows. Partition[4] the bit stream $\tilde{V}_3$ into $\tilde{V}_{3,1}$ and $\tilde{V}_{3,2}$ of lengths $\ell_3 - (\ell_4 + \ell_5)$ and $\ell_4 + \ell_5$, respectively. Compute the modulo-2 summation (binary `xor`) of the bitstreams $\tilde{V}_{3,2}$ and $(\tilde{V}_4, \tilde{V}_5)$. The description Γ_2 is constructed by concatenating this new bit stream with $\tilde{V}_1$, $\tilde{V}_2$, $\tilde{V}_{3,1}$, $\tilde{V}_6$, and $\tilde{V}_7$.

  The partitioning and encoding[5] are illustrated in Fig. 5.

$$\Gamma_1 : \tilde{V}_1, \tilde{V}_4, \tilde{V}_5, \qquad \Gamma_2 : \tilde{V}_1, \tilde{V}_2, \tilde{V}_{3,1}, \tilde{V}_{3,2} \oplus (\tilde{V}_4, \tilde{V}_5), \tilde{V}_6, \tilde{V}_7, \qquad \Gamma_3 : \tilde{V}_1, \tilde{V}_2, \tilde{V}_3.$$

[4]Since we are in regime I, we have $h_3 \geq h_4 + h_5$ and hence, $\ell_3 \geq \ell_4 + \ell_5$.

[5]Note that for this corner point, a part of the description Γ_2 is given by $\tilde{V}_{3,2} \oplus (\tilde{V}_4, \tilde{V}_5)$, which linearly combines independent (compressed) source sequences, just as the network coding idea in the familiar Butterfly network [13].

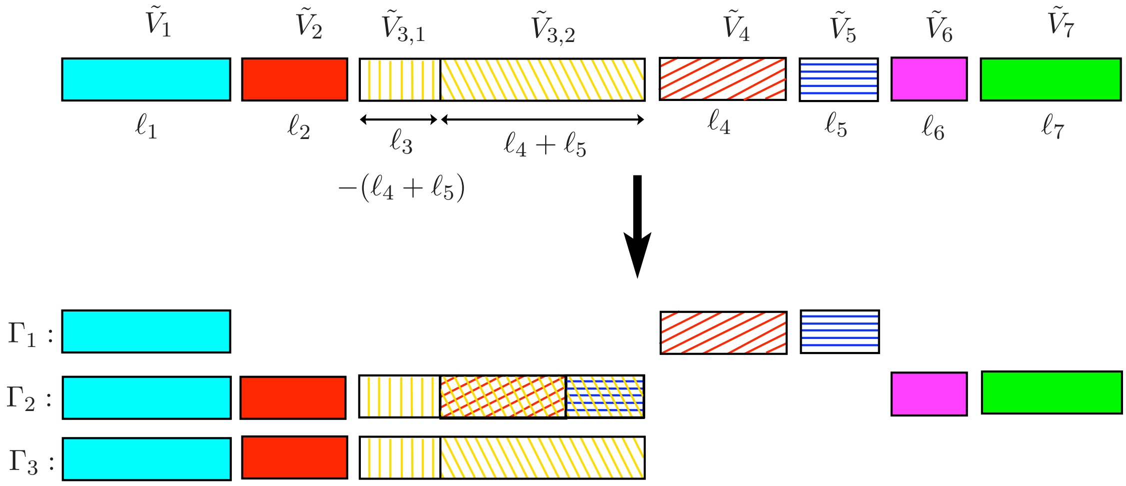

Fig. 5. Linear encoding for the corner-point X_5

- $X_6 = \langle Q3{,}Q6{,}Q10\rangle : (H_1+h_3+h_7, H_2+h_4+h_5+h_6, H_3)$

  Partition $\tilde{V}_3$ into $\tilde{V}_{3,1}$ and $\tilde{V}_{3,2}$ of lengths $\ell_3-(\ell_4+\ell_5)$ and $\ell_4+\ell_5$, respectively.

$$\Gamma_1 : \tilde{V}_1, \tilde{V}_{3,1}, \tilde{V}_{3,2}\oplus(\tilde{V}_4,\tilde{V}_5), \tilde{V}_7 \qquad \Gamma_2 : \tilde{V}_1, \tilde{V}_2, \tilde{V}_4, \tilde{V}_5, \tilde{V}_6, \qquad \Gamma_3 : \tilde{V}_1, \tilde{V}_2, \tilde{V}_3.$$

- $X_7 = \langle Q4{,}Q7{,}Q10\rangle : (H_1+h_4, H_3, H_3+h_5+h_6+h_7)$

  Partition $\tilde{V}_3$ into $\tilde{V}_{3,1}$ and $\tilde{V}_{3,2}$ of lengths $\ell_3-\ell_4$ and ℓ_4, respectively.

$$\Gamma_1 : \tilde{V}_1, \tilde{V}_4, \qquad \Gamma_2 : \tilde{V}_1, \tilde{V}_2, \tilde{V}_{3,1}, \tilde{V}_{3,2}\oplus\tilde{V}_4, \qquad \Gamma_3 : \tilde{V}_1, \tilde{V}_2, \tilde{V}_3, \tilde{V}_5, \tilde{V}_6, \tilde{V}_7.$$

- $X_8 = \langle Q4,\, Q8,\, Q10\rangle : (H_1+h_3, H_2+h_4, H_3+h_5+h_6+h_7)$

  Partition $\tilde{V}_3$ into $\tilde{V}_{3,1}$ and $\tilde{V}_{3,2}$ of lengths $\ell_3-\ell_4$ and ℓ_4, respectively.

$$\Gamma_1 : \tilde{V}_1, \tilde{V}_{3,1}, \tilde{V}_{3,2}\oplus\tilde{V}_4, \qquad \Gamma_2 : \tilde{V}_1, \tilde{V}_2, \tilde{V}_4, \qquad \Gamma_3 : \tilde{V}_1, \tilde{V}_2, \tilde{V}_3, \tilde{V}_5, \tilde{V}_6, \tilde{V}_7.$$

- $X_9 = \langle Q5,\, Q7,\, Q10\rangle : (H_1+h_4, H_3+h_6+h_7, H_3+h_5)$

  Partition $\tilde{V}_3$ into $\tilde{V}_{3,1}$ and $\tilde{V}_{3,2}$ of lengths $\ell_3-\ell_4$ and ℓ_4, respectively.

$$\Gamma_1 : \tilde{V}_1, \tilde{V}_4, \qquad \Gamma_2 : \tilde{V}_1, \tilde{V}_2, \tilde{V}_{3,1}, \tilde{V}_{3,2}\oplus\tilde{V}_4, \tilde{V}_6, \tilde{V}_7, \qquad \Gamma_3 : \tilde{V}_1, \tilde{V}_2, \tilde{V}_3, \tilde{V}_5.$$

- $X_{10} = \langle Q6,\, Q8,\, Q10\rangle : (H_1+h_3+h_7, H_2+h_4, H_3+h_5+h_6)$

  Partition $\tilde{V}_3$ into $\tilde{V}_{3,1}$ and $\tilde{V}_{3,2}$ of lengths $\ell_3-\ell_4$ and ℓ_4, respectively.

$$\Gamma_1 : \tilde{V}_1, \tilde{V}_{3,1}, \tilde{V}_{3,2}\oplus\tilde{V}_4, \tilde{V}_7, \qquad \Gamma_2 : \tilde{V}_1, \tilde{V}_2, \tilde{V}_4, \qquad \Gamma_3 : \tilde{V}_1, \tilde{V}_2, \tilde{V}_3, \tilde{V}_5, \tilde{V}_6.$$

The associated rate region is shown in Fig. 6.

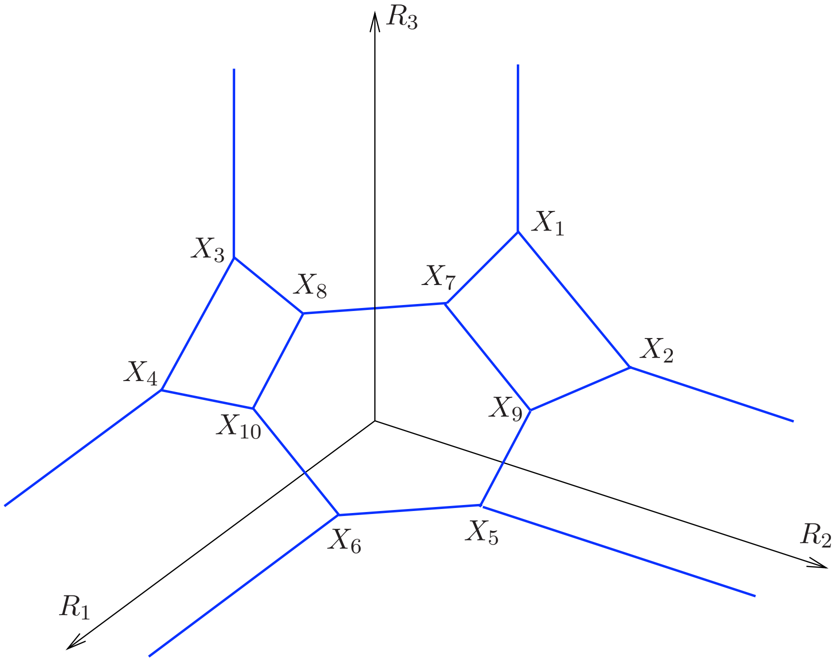

Fig. 6. Rate region for Regime I: $h_3 \geq h_4 + h_5$

Regime II: $h_4 \leq h_3 \leq h_4 + h_5$

In this regime, $(Q11)$ is dominated by $(Q10)$. Therefore, we only have to consider ten hyperplanes. The rates and encoding scheme for the corner points $Y_1 = \langle Q1,Q4,Q7 \rangle$, $Y_2 = \langle Q1,Q5,Q7 \rangle$, $Y_3 = \langle Q2,Q4,Q8 \rangle$, $Y_4 = \langle Q2,Q6,Q8 \rangle$, $Y_7 = \langle Q4,Q7,Q10 \rangle$, $Y_8 = \langle Q4,Q8,Q10 \rangle$, $Y_9 = \langle Q5,Q7,Q10 \rangle$ and $Y_{10} = \langle Q6,Q8,Q10 \rangle$ are exactly the same as that of X_1, X_2, X_3, X_4, X_7, X_8, X_9, and X_{10}, respectively. For the remaining corner points, we next provide the encoding schemes.

- $Y_5 = \langle Q3, Q5, Q9 \rangle : (H_1 + h_4 + h_5, H_2 + h_4 + h_5 + h_6 + h_7, H_3)$

  Partition $\tilde{V}_5$ into $\tilde{V}_{5,1}$ and $\tilde{V}_{5,2}$ of lengths $\ell_3 - \ell_4$ and $\ell_4 + \ell_5 - \ell_3$, respectively. Also partition $\tilde{V}_3$ into $\tilde{V}_{3,1}$ and $\tilde{V}_{3,2}$, of sizes $\ell_3 - \ell_4$ and ℓ_4, respectively.

$$\Gamma_1 : \tilde{V}_1, \tilde{V}_4, \tilde{V}_5, \qquad \Gamma_2 : \tilde{V}_1, \tilde{V}_2, \tilde{V}_{3,2} \oplus \tilde{V}_4, \tilde{V}_{3,1} \oplus \tilde{V}_{5,1}, \tilde{V}_{5,2}, \tilde{V}_6, \tilde{V}_7, \qquad \Gamma_3 : \tilde{V}_1, \tilde{V}_2, \tilde{V}_3.$$

- $Y_6 = \langle Q3, Q6, Q9 \rangle : (H_1 + h_4 + h_5 + h_7, H_2 + h_4 + h_5 + h_6, H_3)$

  Partition $\tilde{V}_5$ into $\tilde{V}_{5,1}$ and $\tilde{V}_{5,2}$ of lengths $\ell_3 - \ell_4$ and $\ell_4 + \ell_5 - \ell_3$, respectively. Also partition $\tilde{V}_3$ into $\tilde{V}_{3,1}$ and $\tilde{V}_{3,2}$, of sizes $\ell_3 - \ell_4$ and ℓ_4, respectively.

$$\Gamma_1 : \tilde{V}_1, \tilde{V}_4, \tilde{V}_5, \tilde{V}_7, \qquad \Gamma_2 : \tilde{V}_1, \tilde{V}_2, \tilde{V}_{3,2} \oplus \tilde{V}_4, \tilde{V}_{3,1} \oplus \tilde{V}_{5,1}, \tilde{V}_{5,2}, \tilde{V}_6, \qquad \Gamma_3 : \tilde{V}_1, \tilde{V}_2, \tilde{V}_3.$$

- $Y_{11} = \langle Q5, Q9, Q10 \rangle : (H_1 + h_3, H_3 + h_6 + h_7, H_2 + h_4 + h_5)$
  Partition $\tilde{V}_5$ into $\tilde{V}_{5,1}$ and $\tilde{V}_{5,2}$ of lengths $\ell_3 - \ell_4$ and $\ell_4 + \ell_5 - \ell_3$, respectively. Also partition $\tilde{V}_3$ into $\tilde{V}_{3,1}$ and $\tilde{V}_{3,2}$, of sizes $\ell_3 - \ell_4$ and ℓ_4, respectively.

$$\Gamma_1 : \tilde{V}_1, \tilde{V}_4, \tilde{V}_{5,1}, \qquad \Gamma_2 : \tilde{V}_1, \tilde{V}_2, \tilde{V}_{3,2} \oplus \tilde{V}_4, \tilde{V}_{3,1} \oplus \tilde{V}_{5,1}, \tilde{V}_6, \tilde{V}_7, \qquad \Gamma_3 : \tilde{V}_1, \tilde{V}_2, \tilde{V}_3, \tilde{V}_{5,2}.$$

- $Y_{12} = \langle Q6, Q9, Q10 \rangle : (H_1 + h_3 + h_7, H_3 + h_6, H_2 + h_4 + h_5)$
  Partition $\tilde{V}_5$ into $\tilde{V}_{5,1}$ and $\tilde{V}_{5,2}$ of lengths $\ell_3 - \ell_4$ and $\ell_4 + \ell_5 - \ell_3$, respectively. Also partition $\tilde{V}_3$ into $\tilde{V}_{3,1}$ and $\tilde{V}_{3,2}$, of sizes $\ell_3 - \ell_4$ and ℓ_4, respectively.

$$\Gamma_1 : \tilde{V}_1, \tilde{V}_4, \tilde{V}_{5,1}, \tilde{V}_7, \qquad \Gamma_2 : \tilde{V}_1, \tilde{V}_2, \tilde{V}_{3,2} \oplus \tilde{V}_4, \tilde{V}_{3,1} \oplus \tilde{V}_{5,1}, \tilde{V}_6, \qquad \Gamma_3 : \tilde{V}_1, \tilde{V}_2, \tilde{V}_3, \tilde{V}_{5,2}.$$

Fig. 7 shows the rate region for this regime.

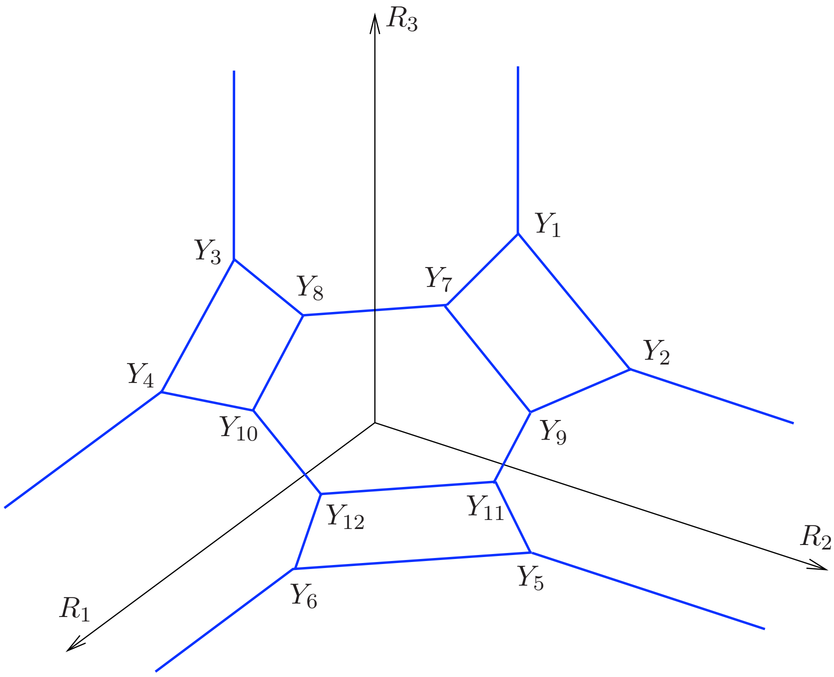

Fig. 7. Rate region for Regime II of ordering level $\mathscr{L}_1$: $h_4 \leq h_3 \leq h_4 + h_5$

Regime III: $h_3 \leq h_4$

It is clear that in this regime ($Q10$) is dominated by ($Q11$), and thus ($Q10$) does not affect the rate region. The remaining ten inequalities characterize the region. The rates and coding schemes for the points $Z_1 = \langle Q1,Q4,Q7 \rangle$, $Z_2 = \langle Q1,Q5,Q7 \rangle$, $Z_3 = \langle Q2,Q4,Q8 \rangle$, and $Z_4 = \langle Q2,Q6,Q8 \rangle$ are exactly the same as that of X_1, X_2, X_3, and X_4, respectively. The rate tuples and the corresponding descriptions for the other corner points are as follows.

- $Z_5 = \langle Q3, Q5, Q9 \rangle : (H_1 + h_4 + h_5, H_2 + h_4 + h_5 + h_6 + h_7, H_3)$
  Partition $\tilde{V}_4$ into $\tilde{V}_{4,1}$ and $\tilde{V}_{4,2}$ of lengths ℓ_3 and $\ell_4 - \ell_3$, respectively.

$$\Gamma_1 : \tilde{V}_1, \tilde{V}_4, \tilde{V}_5, \qquad \Gamma_2 : \tilde{V}_1, \tilde{V}_2, \tilde{V}_3 \oplus \tilde{V}_{4,1}, \tilde{V}_{4,2}, \tilde{V}_5, \tilde{V}_6, \tilde{V}_7, \qquad \Gamma_3 : \tilde{V}_1, \tilde{V}_2, \tilde{V}_3.$$

- $Z_6 = \langle Q3, Q6, Q9 \rangle : (H_1 + h_4 + h_5 + h_7, H_2 + h_4 + h_5 + h_6, H_3)$
  Partition $\tilde{V}_4$ into $\tilde{V}_{4,1}$ and $\tilde{V}_{4,2}$ of lengths ℓ_3 and $\ell_4 - \ell_3$, respectively.

$$\Gamma_1 : \tilde{V}_1, \tilde{V}_4, \tilde{V}_5, \tilde{V}_7, \qquad \Gamma_2 : \tilde{V}_1, \tilde{V}_2, \tilde{V}_3 \oplus \tilde{V}_{4,1}, \tilde{V}_{4,2}, \tilde{V}_5, \tilde{V}_6, \qquad \Gamma_3 : \tilde{V}_1, \tilde{V}_2, \tilde{V}_3.$$

- $Z_7 = \langle Q4, Q7, Q8, Q11 \rangle : \left(H_1 + \frac{h_3+h_4}{2}, H_2 + \frac{h_3+h_4}{2}, H_2 + \frac{h_3+h_4}{2} + h_5 + h_6 + h_7\right)$
  Partition $\tilde{V}_4$ into $\tilde{V}_{4,1}$, $\tilde{V}_{4,2}$, and $\tilde{V}_{4,3}$ of lengths ℓ_3, $\frac{1}{2}(\ell_4 - \ell_3)$ and $\frac{1}{2}(\ell_4 - \ell_3)$, respectively.

$$\Gamma_1 : \tilde{V}_1, \tilde{V}_{4,1}, \tilde{V}_{4,2} \quad \Gamma_2 : \tilde{V}_1, \tilde{V}_2, \tilde{V}_3 \oplus \tilde{V}_{4,1}, \tilde{V}_{4,2} \oplus \tilde{V}_{4,3}, \quad \Gamma_3 : \tilde{V}_1, \tilde{V}_2, \tilde{V}_3, \tilde{V}_{4,3}, \tilde{V}_5, \tilde{V}_6, \tilde{V}_7.$$

- $Z_8 = \langle Q5, Q7, Q9, Q11 \rangle : \left(H_1 + \frac{h_3+h_4}{2}, H_2 + \frac{h_3+h_4}{2} + h_6 + h_7, H_2 + \frac{h_3+h_4}{2} + h_5\right)$
  Partition $\tilde{V}_4$ into $\tilde{V}_{4,1}$, $\tilde{V}_{4,2}$, and $\tilde{V}_{4,3}$ of lengths ℓ_3, $\frac{1}{2}(\ell_4 - \ell_3)$ and $\frac{1}{2}(\ell_4 - \ell_3)$, respectively.

$$\Gamma_1 : \tilde{V}_1, \tilde{V}_{4,1}, \tilde{V}_{4,2} \quad \Gamma_2 : \tilde{V}_1, \tilde{V}_2, \tilde{V}_3 \oplus \tilde{V}_{4,1}, \tilde{V}_{4,2} \oplus \tilde{V}_{4,3}, \tilde{V}_6, \tilde{V}_7, \quad \Gamma_3 : \tilde{V}_1, \tilde{V}_2, \tilde{V}_3, \tilde{V}_{4,3}, \tilde{V}_5.$$

- $Z_9 = \langle Q6, Q8, Q11 \rangle : \left(H_1 + \frac{h_3+h_4}{2} + h_7, H_2 + \frac{h_3+h_4}{2}, H_2 + \frac{h_3+h_4}{2} + h_5 + h_6\right)$
  Partition $\tilde{V}_4$ into $\tilde{V}_{4,1}$, $\tilde{V}_{4,2}$, and $\tilde{V}_{4,3}$ of lengths ℓ_3, $\frac{1}{2}(\ell_4 - \ell_3)$ and $\frac{1}{2}(\ell_4 - \ell_3)$, respectively.

$$\Gamma_1 : \tilde{V}_1, \tilde{V}_{4,1}, \tilde{V}_{4,2}, \tilde{V}_7 \quad \Gamma_2 : \tilde{V}_1, \tilde{V}_2, \tilde{V}_3 \oplus \tilde{V}_{4,1}, \tilde{V}_{4,2} \oplus \tilde{V}_{4,3}, \quad \Gamma_3 : \tilde{V}_1, \tilde{V}_2, \tilde{V}_3, \tilde{V}_{4,3}, \tilde{V}_5, \tilde{V}_6.$$

- $Z_{10} = \langle Q6, Q8, Q11 \rangle : \left(H_1 + \frac{h_3+h_4}{2} + h_7, H_2 + \frac{h_3+h_4}{2} + h_6, H_2 + \frac{h_3+h_4}{2} + h_5\right)$
  Partition $\tilde{V}_4$ into $\tilde{V}_{4,1}$, $\tilde{V}_{4,2}$, and $\tilde{V}_{4,3}$ of lengths ℓ_3, $\frac{1}{2}(\ell_4 - \ell_3)$ and $\frac{1}{2}(\ell_4 - \ell_3)$, respectively.

$$\Gamma_1 : \tilde{V}_1, \tilde{V}_{4,1}, \tilde{V}_{4,2}, \tilde{V}_7 \quad \Gamma_2 : \tilde{V}_1, \tilde{V}_2, \tilde{V}_3 \oplus \tilde{V}_{4,1}, \tilde{V}_{4,2} \oplus \tilde{V}_{4,3}, \tilde{V}_6, \quad \Gamma_3 : \tilde{V}_1, \tilde{V}_2, \tilde{V}_3, \tilde{V}_{4,3}, \tilde{V}_5.$$

  This region and its corner points are shown in Fig. 8.

The coding schemes proposed for these three cases give us the achievability proof of the theorem for the specific ordering $\mathscr{L}_1$. As stated before, the coding scheme for other possible orderings listed in Table I are similar to that of the ordering $\mathscr{L}_1$. There are three main ingredients used in all of them; (1)

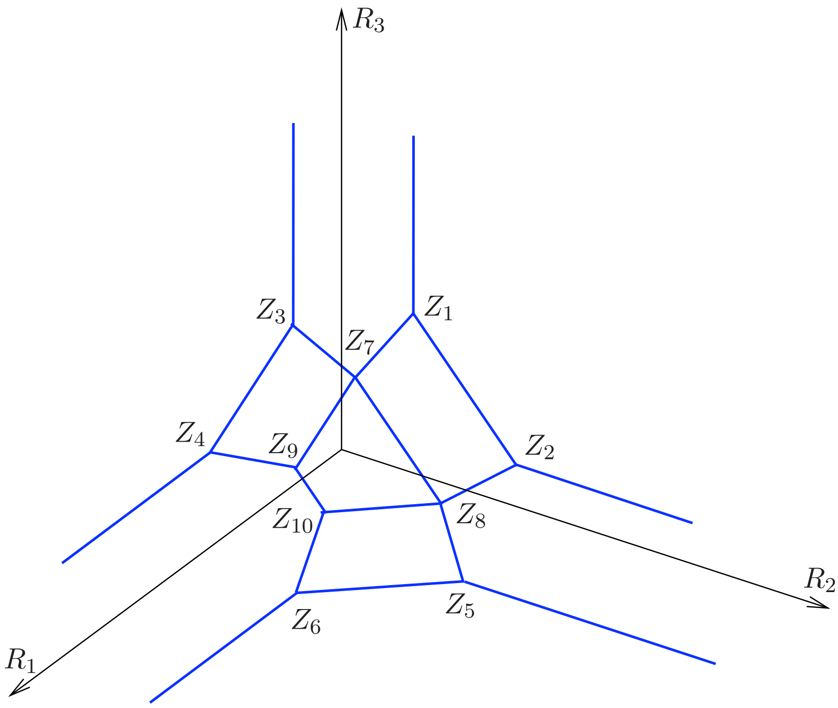

Fig. 8. Rate region for Regime III of the ordering level $\mathscr{L}_1$: $h_3 \leq h_4$

converting the source sequences into bitstreams, (2) partitioning the bit streams into sequences of proper length, and (3) (if required) applying linear coding (binary `xor`) on them. This completes the proof of Theorem 1.

V. Asymmetric Multiple Description Coding

In this section we prove Theorems 2 and 3, which together give an approximate characterization for the admissible rate region of the A-MD problem.

A. An Outer Bound for the Rate Region of A-MD: Proof of Theorem 2

In order to prove this theorem, we first show a parametric outer-bound for the A-MD rate region. Then we specialize the parameters to obtain the bound claimed in the theorem.

We first need to define a set of auxiliary random variables in order to state and prove the parametric bound, which are some noisy versions of the source. The strategy of expanding the probability space by a single auxiliary variable was used to characterize the two descriptions Gaussian MD region [2], and later in [8] extended to include multiple auxiliary random variables with certain built-in Markov structure. We shall continue to use this extended strategy as used in [8].

Let $N_i \sim \mathcal{N}(0, \sigma_i^2),\ i = 1, \dots, 6$, be mutually independent zero-mean Gaussian random variables with variance σ_i^2. They are also assumed to be independent of X. A noisy version of the source, Y_i, is defined as

$$Y_i = X + Z_i, \quad i = 1, \dots, 6 \tag{29}$$

where $Z_i = \sum_{j=i}^{6} N_j$ for $j = 1, \dots, 6$. Thus $d_i \triangleq \sum_{j=i}^{6} \sigma_j^2$ would be the variance of the noises Z_i, for $i = 1, \dots, 6$. We also define $Y_7 = X$ and $d_7 = 0$ for convenience. Note that incremental noises are added to X to build Y_i's, and therefore they form a Markov chain as

$$(\Gamma_1, \Gamma_2, \Gamma_3) \leftrightarrow X^n \leftrightarrow Y_6^n \leftrightarrow Y_5^n \leftrightarrow \cdots \leftrightarrow Y_1^n. \tag{30}$$

The following theorem provides a parametric outer-bound for the rate region of the A-MLD problem, depending on d_i variables, which are the noise variances defined above. Such bound holds for any choice of $d_1 \geq d_2 \geq \cdots \geq d_6 > 0$, and can be further optimized to obtain a good non-parametric outer-bound for the rate region. However, we simply derive the bound in Theorem 2 by setting the values of d_i's.

Theorem 4: For a given distortion vector $\mathbf{D} = (D_{\Gamma_1}, \dots, D_{\Gamma_1\Gamma_2\Gamma_3})$ and a set of variables $d_1 \geq d_2 \geq$

$\cdots \geq d_6 > d_7 = 0$, denote by $\underline{\mathcal{R}}^p_{\text{MD}}(\mathbf{D}, \mathbf{d})$ the set of all rate triples (R_1, R_2, R_3) satisfying

$$R_j \geq \frac{1}{2}\log\frac{1}{D_{\Gamma_j}} \qquad j=1,2,3 \tag{$\mathcal{PO}$−1}$$

$$R_i + R_j \geq \frac{1}{2}\log\frac{1+d_{\mathscr{L}(\Gamma_i)}}{D_{\Gamma_i}+d_{\mathscr{L}(\Gamma_i)}}\frac{1+d_{\mathscr{L}(\Gamma_j)}}{D_{\Gamma_j}+d_{\mathscr{L}(\Gamma_j)}}\frac{(D_{\Gamma_i\Gamma_j}+d_{\max\{\mathscr{L}(\Gamma_i),\mathscr{L}(\Gamma_j)\}})}{(1+d_{\max\{\mathscr{L}(\Gamma_i),\mathscr{L}(\Gamma_j)\}})D_{\Gamma_i\Gamma_j}} \qquad i \neq j \tag{$\mathcal{PO}$−2}$$

$$\begin{aligned}
2R_i + R_j + R_k \geq\ & \frac{1}{2}\log\left(\frac{1+d_{\mathscr{L}(\Gamma_i)}}{D_{\Gamma_i}+d_{\mathscr{L}(\Gamma_i)}}\right)^2\frac{1+d_{\mathscr{L}(\Gamma_j)}}{D_{\Gamma_j}+d_{\mathscr{L}(\Gamma_j)}}\frac{1+d_{\mathscr{L}(\Gamma_k)}}{D_{\Gamma_k}+d_{\mathscr{L}(\Gamma_k)}}\\
&+\frac{1}{2}\log\frac{(1+d_{\mathscr{L}(\Gamma_i,\Gamma_j)})(D_{\Gamma_i\Gamma_j}+d_{\max\{\mathscr{L}(\Gamma_i),\mathscr{L}(\Gamma_j)\}})}{(1+d_{\max\{\mathscr{L}(\Gamma_i),\mathscr{L}(\Gamma_j)\}})(D_{\Gamma_i\Gamma_j}+d_{\mathscr{L}(\Gamma_i,\Gamma_j)})}\\
&+\frac{1}{2}\log\frac{(1+d_{\mathscr{L}(\Gamma_i,\Gamma_k)})(D_{\Gamma_i\Gamma_k}+d_{\max\{\mathscr{L}(\Gamma_i),\mathscr{L}(\Gamma_k)\}})}{(1+d_{\max\{\mathscr{L}(\Gamma_i),\mathscr{L}(\Gamma_k)\}})(D_{\Gamma_i\Gamma_k}+d_{\mathscr{L}(\Gamma_i,\Gamma_k)})}\\
&+\frac{1}{2}\log\frac{(D_{\Gamma_i\Gamma_j\Gamma_k}+d_{\max\{\mathscr{L}(\Gamma_i,\Gamma_j),\mathscr{L}(\Gamma_i,\Gamma_k)\}})}{(1+d_{\max\{\mathscr{L}(\Gamma_i,\Gamma_j),\mathscr{L}(\Gamma_i,\Gamma_k)\}})D_{\Gamma_i\Gamma_j\Gamma_k}} \qquad i\neq j\neq k
\end{aligned} \tag{$\mathcal{PO}$−3}$$

$$\begin{aligned}
R_1+R_2+R_3 \geq\ & \frac{1}{2}\log\frac{1+d_{\mathscr{L}(\Gamma_1)}}{D_{\Gamma_1}+d_{\mathscr{L}(\Gamma_1)}}\frac{1+d_{\mathscr{L}(\Gamma_2)}}{D_{\Gamma_2}+d_{\mathscr{L}(\Gamma_2)}}\frac{1+d_{\min\{\mathscr{L}(\Gamma_1,\Gamma_2),\mathscr{L}(\Gamma_3)\}}}{D_{\Gamma_3}+d_{\min\{\mathscr{L}(\Gamma_1,\Gamma_2),\mathscr{L}(\Gamma_3)\}}}\\
&+\frac{1}{2}\log\frac{(1+d_{\min\{\mathscr{L}(\Gamma_1,\Gamma_2),\mathscr{L}(\Gamma_3)\}})(D_{\Gamma_1\Gamma_2}+d_{\mathscr{L}(\Gamma_2)})}{(1+d_{\mathscr{L}(\Gamma_2)})(D_{\Gamma_1\Gamma_2}+d_{\min\{\mathscr{L}(\Gamma_1,\Gamma_2),\mathscr{L}(\Gamma_3)\}})}\\
&+\frac{1}{2}\log\frac{D_{\Gamma_1\Gamma_2\Gamma_3}+d_{\min\{\mathscr{L}(\Gamma_1,\Gamma_2),\mathscr{L}(\Gamma_3)\}}}{(1+d_{\min\{\mathscr{L}(\Gamma_1,\Gamma_2),\mathscr{L}(\Gamma_3)\}})D_{\Gamma_1\Gamma_2\Gamma_3}}
\end{aligned} \tag{$\mathcal{PO}$−4}$$

$$\begin{aligned}
R_1+R_2+R_3 \geq\ & \frac{1}{2}\log\frac{1+d_{\mathscr{L}(\Gamma_1)}}{D_{\Gamma_1}+d_{\mathscr{L}(\Gamma_1)}}\frac{1+d_{\mathscr{L}(\Gamma_2)}}{D_{\Gamma_2}+d_{\mathscr{L}(\Gamma_2)}}\frac{1+d_{\mathscr{L}(\Gamma_3)}}{D_{\Gamma_3}+d_{\mathscr{L}(\Gamma_3)}}\\
&+\frac{1}{4}\log\frac{(1+d_\alpha)(D_{\Gamma_1\Gamma_2}+d_{\mathscr{L}(\Gamma_2)})}{(1+d_{\mathscr{L}(\Gamma_2)})(D_{\Gamma_1\Gamma_2}+d_\alpha)}+\frac{1}{4}\log\frac{(1+d_\alpha)(D_{\Gamma_1\Gamma_3}+d_{\mathscr{L}(\Gamma_3)})}{(1+d_{\mathscr{L}(\Gamma_3)})(D_{\Gamma_1\Gamma_3}+d_\alpha)}\\
&+\frac{1}{4}\log\frac{(1+d_\alpha)(D_{\Gamma_2\Gamma_3}+d_{\mathscr{L}(\Gamma_3)})}{(1+d_{\mathscr{L}(\Gamma_3)})(D_{\Gamma_2\Gamma_3}+d_\alpha)}+\frac{1}{2}\log\frac{D_{\Gamma_1\Gamma_2\Gamma_3}+d_{\mathscr{L}(\Gamma_3)}}{(1+d_{\mathscr{L}(\Gamma_3)})D_{\Gamma_1\Gamma_2\Gamma_3}},
\end{aligned} \tag{$\mathcal{PO}$−5}$$

where

$$\alpha = \begin{cases} \mathscr{L}(\Gamma_3) & \text{if } \mathscr{L}(\Gamma_3) > \mathscr{L}(\Gamma_1,\Gamma_2),\\ \min\{\mathscr{L}(\Gamma_1,\Gamma_2),\mathscr{L}(\Gamma_1,\Gamma_3),\mathscr{L}(\Gamma_2,\Gamma_3)\} & \text{if } \mathscr{L}(\Gamma_3) < \mathscr{L}(\Gamma_1,\Gamma_2).\end{cases}$$

Then any admissible rate triple belongs to $\underline{\mathcal{R}}^p_{\text{MD}}(\mathbf{D}, \mathbf{d})$, *i.e.*, $\mathcal{R}_{\text{MD}}(\mathbf{D}) \subseteq \underline{\mathcal{R}}^p_{\text{MD}}(\mathbf{D}, \mathbf{d})$, for all choices of $d_1 \geq d_2 \geq \cdots \geq d_6 \geq d_7 = 0$.

The following two lemmas are extracted from [5], whose proofs can be found in Appendix A for completeness. They are useful to bound the mutual information between the noisy versions of the source and the descriptions.

***Lemma** 3:* For any set of descriptions $\mathcal{S} \subseteq \{\Gamma_1, \Gamma_2, \Gamma_3\}$, and noisy version of the source Y_i, $i =$

$1, 2, \ldots, 7$, we have

$$I(\mathcal{S}; Y_i^n) \geq \frac{n}{2} \log \frac{1 + d_i}{D_{\mathcal{S}} + d_i}. \tag{31}$$

***Lemma** 4:* For any subset of the descriptions $\mathcal{S}$, and two noisy versions of the source Y_i and Y_j with $i < j$, we have

$$I(\mathcal{S}; Y_j^n) - I(\mathcal{S}; Y_i^n) \geq \frac{n}{2} \log \frac{(1 + d_j)(D_{\mathcal{S}} + d_i)}{(1 + d_i)(D_{\mathcal{S}} + d_j)}. \tag{32}$$

We will use these results in several points in the proof of Theorem 4, which are indicated by (†). Now, we are ready to prove the parametric outer-bound.

Proof of Theorem 4:

The single description levels inequalities are just straight forward result of Lemma 3. We have

$$nR_i \geq H(\Gamma_i) = H(\Gamma_i) - H(\Gamma_i | X^n) = I(\Gamma_i; X^n) \overset{(\dagger)}{\geq} \frac{n}{2} \log \frac{1}{D_{\Gamma_i}} \tag{33}$$

where we used Lemma 3 for $Y_7 = X$ and the fact $d_7 = 0$ in the last inequality. This proves ($\mathcal{PO}-1$).

The bound for the two description rates in ($\mathcal{PO}-2$) follows from

$$\begin{aligned}
n(R_i + R_j + 2\varepsilon) &\geq H(\Gamma_i) + H(\Gamma_j) \\
&\overset{(a)}{\geq} H(\Gamma_i) + H(\Gamma_j) - H(\Gamma_i, \Gamma_j | X^n) \\
&\quad - \Big[H(\Gamma_i | Y^n_{\max\{\mathscr{L}(\Gamma_i), \mathscr{L}(\Gamma_j)\}}) + H(\Gamma_j | Y^n_{\max\{\mathscr{L}(\Gamma_i), \mathscr{L}(\Gamma_j)\}}) - H(\Gamma_i, \Gamma_j | Y^n_{\max\{\mathscr{L}(\Gamma_i), \mathscr{L}(\Gamma_j)\}}) \Big] \\
&= I(\Gamma_i; Y^n_{\max\{\mathscr{L}(\Gamma_i), \mathscr{L}(\Gamma_j)\}}) + I(\Gamma_j; Y^n_{\max\{\mathscr{L}(\Gamma_i), \mathscr{L}(\Gamma_j)\}}) \\
&\quad + [I(\Gamma_i \Gamma_j; X^n) - I(\Gamma_i \Gamma_j; Y^n_{\max\{\mathscr{L}(\Gamma_i), \mathscr{L}(\Gamma_j)\}})] \\
&\overset{(b)}{\geq} I(\Gamma_i; Y^n_{\mathscr{L}(\Gamma_i)}) + I(\Gamma_j; Y^n_{\mathscr{L}(\Gamma_j)}) + [I(\Gamma_i \Gamma_j; X^n) - I(\Gamma_i \Gamma_j; Y^n_{\max\{\mathscr{L}(\Gamma_i), \mathscr{L}(\Gamma_j)\}})] \\
&\overset{(\dagger)}{\geq} \frac{n}{2} \log \frac{1 + d_{\mathscr{L}(\Gamma_i)}}{D_{\Gamma_i} + d_{\mathscr{L}(\Gamma_i)}} \frac{1 + d_{\mathscr{L}(\Gamma_j)}}{D_{\Gamma_j} + d_{\mathscr{L}(\Gamma_j)}} \frac{(1 + 0)(D_{\Gamma_i \Gamma_j} + d_{\max\{\mathscr{L}(\Gamma_i), \mathscr{L}(\Gamma_j)\}})}{(1 + d_{\max\{\mathscr{L}(\Gamma_i), \mathscr{L}(\Gamma_j)\}})(D_{\Gamma_i \Gamma_j} + 0)}
\end{aligned} \tag{34}$$

where the subtracted terms in (a) are positive due to the fact that Γ_1 and Γ_2 are functions of X^n and non-negativity of mutual information, (b) is by the data processing inequality and the Markov chain in (30). Finally, we have used Lemma 3 and Lemma 4 in (†).

The inequality $(\mathcal{PO}-3)$ can be proved through the following chain of inequalities.

$$
\begin{aligned}
n(2R_i + R_j + R_k + 4\varepsilon) &\geq 2H(\Gamma_i) + H(\Gamma_j) + H(\Gamma_k) \\
&\overset{(a)}{\geq} 2H(\Gamma_i) + H(\Gamma_j) + H(\Gamma_k) - H(\Gamma_i\Gamma_j\Gamma_k|X^n) \\
&\quad - \Big[H(\Gamma_i|Y^n_{\max\{\mathscr{L}(\Gamma_i),\mathscr{L}(\Gamma_j)\}}) + H(\Gamma_j|Y^n_{\max\{\mathscr{L}(\Gamma_i),\mathscr{L}(\Gamma_j)\}}) - H(\Gamma_i\Gamma_j|Y^n_{\max\{\mathscr{L}(\Gamma_i),\mathscr{L}(\Gamma_j)\}})\Big] \\
&\quad - \Big[H(\Gamma_i|Y^n_{\max\{\mathscr{L}(\Gamma_i),\mathscr{L}(\Gamma_k)\}}) + H(\Gamma_k|Y^n_{\max\{\mathscr{L}(\Gamma_i),\mathscr{L}(\Gamma_k)\}}) - H(\Gamma_i\Gamma_k|Y^n_{\max\{\mathscr{L}(\Gamma_i),\mathscr{L}(\Gamma_k)\}})\Big] \\
&\quad - \Big[H(\Gamma_i\Gamma_j|Y^n_{\max\{\mathscr{L}(\Gamma_i,\Gamma_j),\mathscr{L}(\Gamma_i,\Gamma_k)\}}) + H(\Gamma_i\Gamma_k|Y^n_{\max\{\mathscr{L}(\Gamma_i,\Gamma_j),\mathscr{L}(\Gamma_i,\Gamma_k)\}}) \\
&\qquad - H(\Gamma_i\Gamma_j\Gamma_k|Y^n_{\max\{\mathscr{L}(\Gamma_i,\Gamma_j),\mathscr{L}(\Gamma_i,\Gamma_k)\}})\Big] \\
&= I(\Gamma_i;Y^n_{\max\{\mathscr{L}(\Gamma_i),\mathscr{L}(\Gamma_j)\}}) + I(\Gamma_j;Y^n_{\max\{\mathscr{L}(\Gamma_i),\mathscr{L}(\Gamma_j)\}}) \\
&\quad + I(\Gamma_i;Y^n_{\max\{\mathscr{L}(\Gamma_i),\mathscr{L}(\Gamma_k)\}}) + I(\Gamma_k;Y^n_{\max\{\mathscr{L}(\Gamma_i),\mathscr{L}(\Gamma_k)\}}) \\
&\quad + [I(\Gamma_i\Gamma_j;Y^n_{\max\{\mathscr{L}(\Gamma_i,\Gamma_j),\mathscr{L}(\Gamma_i,\Gamma_k)\}}) - I(\Gamma_i\Gamma_j;Y^n_{\max\{\mathscr{L}(\Gamma_i),\mathscr{L}(\Gamma_j)\}})] \\
&\quad + [I(\Gamma_i\Gamma_k;Y^n_{\max\{\mathscr{L}(\Gamma_i,\Gamma_j),\mathscr{L}(\Gamma_i,\Gamma_k)\}}) - I(\Gamma_i\Gamma_k;Y^n_{\max\{\mathscr{L}(\Gamma_i),\mathscr{L}(\Gamma_k)\}})] \\
&\quad + [I(\Gamma_i\Gamma_j\Gamma_k;X^n) - I(\Gamma_i\Gamma_j\Gamma_k;Y^n_{\max\{\mathscr{L}(\Gamma_i,\Gamma_j),\mathscr{L}(\Gamma_i,\Gamma_k)\}})] \qquad (35)
\end{aligned}
$$

where in (a) we have used the fact that all the brackets are non-negative. Now, we will bound each term in (35) individually. The single description terms can be bounded as

$$
I(\Gamma_i;Y^n_{\max\{\mathscr{L}(\Gamma_i),\mathscr{L}(\Gamma_k)\}}) \geq I(\Gamma_i;Y^n_{\mathscr{L}(\Gamma_i)}) \overset{(\dagger)}{\geq} \frac{n}{2}\log\frac{1+d_{\mathscr{L}(\Gamma_i)}}{D_{\Gamma_i}+d_{\mathscr{L}(\Gamma_i)}}, \qquad (36)
$$

and similarly for j and k. Also we can bound the differential terms as

$$
\begin{aligned}
I(\Gamma_i\Gamma_j;Y^n_{\max\{\mathscr{L}(\Gamma_i,\Gamma_j),\mathscr{L}(\Gamma_i,\Gamma_k)\}}) &- I(\Gamma_i\Gamma_j;Y^n_{\max\{\mathscr{L}(\Gamma_i),\mathscr{L}(\Gamma_j)\}}) \\
&\overset{(b)}{\geq} I(\Gamma_i\Gamma_j;Y^n_{\mathscr{L}(\Gamma_i\Gamma_j)}) - I(\Gamma_i\Gamma_j;Y^n_{\max\{\mathscr{L}(\Gamma_i),\mathscr{L}(\Gamma_j)\}}) \\
&\overset{(\dagger)}{\geq} \frac{n}{2}\log\frac{(1+d_{\mathscr{L}(\Gamma_i\Gamma_j)})(D_{\Gamma_i\Gamma_j}+d_{\max\{\mathscr{L}(\Gamma_i),\mathscr{L}(\Gamma_j)\}})}{(1+d_{\max\{\mathscr{L}(\Gamma_i),\mathscr{L}(\Gamma_j)\}})(D_{\Gamma_i\Gamma_j}+d_{\mathscr{L}(\Gamma_i\Gamma_j)})} \qquad (37)
\end{aligned}
$$

where (b) is due to the data processing inequality implied by the Markov chain

$$
(\Gamma_i\Gamma_j) \leftrightarrow Y^n_{\max\{\mathscr{L}(\Gamma_i,\Gamma_j),\mathscr{L}(\Gamma_i,\Gamma_k)\}} \leftrightarrow Y^n_{\mathscr{L}(\Gamma_i\Gamma_j)}
$$

implied by (30). We also have

$$
\begin{aligned}
I(\Gamma_i\Gamma_j\Gamma_k;X^n) &- I(\Gamma_i\Gamma_j\Gamma_k;Y^n_{\max\{\mathscr{L}(\Gamma_i,\Gamma_j),\mathscr{L}(\Gamma_i,\Gamma_k)\}}) \\
&\overset{(\dagger)}{\geq} \frac{n}{2}\log\frac{(1+0)(D_{\Gamma_i\Gamma_j\Gamma_k}+d_{\max\{\mathscr{L}(\Gamma_i,\Gamma_j),\mathscr{L}(\Gamma_i,\Gamma_k)\}})}{(1+d_{\max\{\mathscr{L}(\Gamma_i,\Gamma_j),\mathscr{L}(\Gamma_i,\Gamma_k)\}})(D_{\Gamma_i\Gamma_j\Gamma_k}+0)}. \qquad (38)
\end{aligned}
$$

By replacing (36)–(38) in (35) we get the desired inequality.

In order to derive the sum-rate bound in ($\mathcal{PO}-4$), we can write

$$\begin{aligned} n(R_1+R_2+R_3+3\varepsilon) &\geq H(\Gamma_1)+H(\Gamma_2)+H(\Gamma_3) \\ &\geq H(\Gamma_1)+H(\Gamma_2)+H(\Gamma_3)-H(\Gamma_1\Gamma_2\Gamma_3|X^n) \\ &\quad -\Big[H(\Gamma_1|Y^n_{\mathscr{L}(\Gamma_2)})+H(\Gamma_2|Y^n_{\mathscr{L}(\Gamma_2)})-H(\Gamma_1\Gamma_2|Y^n_{\mathscr{L}(\Gamma_2)})\Big] \\ &\quad -\Big[H(\Gamma_1\Gamma_2|Y^n_{\min\{\mathscr{L}(\Gamma_1,\Gamma_2),\mathscr{L}(\Gamma_3)\}})+H(\Gamma_3|Y^n_{\min\{\mathscr{L}(\Gamma_1,\Gamma_2),\mathscr{L}(\Gamma_3)\}}) \\ &\qquad -H(\Gamma_1\Gamma_2\Gamma_3|Y^n_{\min\{\mathscr{L}(\Gamma_1,\Gamma_2),\mathscr{L}(\Gamma_3)\}})\Big] \\ &\geq I(\Gamma_1;Y^n_{\mathscr{L}(\Gamma_1)})+I(\Gamma_2;Y^n_{\mathscr{L}(\Gamma_2)})+I(\Gamma_3;Y^n_{\min\{\mathscr{L}(\Gamma_1,\Gamma_2),\mathscr{L}(\Gamma_3)\}}) \\ &\quad +\Big[I(\Gamma_1\Gamma_2;Y^n_{\min\{\mathscr{L}(\Gamma_1,\Gamma_2),\mathscr{L}(\Gamma_3)\}})-I(\Gamma_1\Gamma_2;Y^n_{\mathscr{L}(\Gamma_2)})\Big] \\ &\quad +\Big[I(\Gamma_1\Gamma_2\Gamma_3;X^n)-I(\Gamma_1\Gamma_2\Gamma_3;Y^n_{\min\{\mathscr{L}(\Gamma_1,\Gamma_2),\mathscr{L}(\Gamma_3)\}})\Big]. \end{aligned} \tag{39}$$

Again, applying Lemma 3 and Lemma 4 we can bound each term in (39), and obtain ($\mathcal{PO}-4$).

It remains to show the bound in ($\mathcal{PO}-5$). Recall the proof of ($P4$), and consider two cases. If $\mathscr{L}(\Gamma_3) > \mathscr{L}(\Gamma_1,\Gamma_2)$, then Using the similar argument as in the proof of ($\mathcal{PO}-3$), we obtain

$$\begin{aligned} n(R_1+R_2+R_3+3\varepsilon) &\geq H(\Gamma_1)+H(\Gamma_2)+H(\Gamma_3)-H(\Gamma_1\Gamma_2\Gamma_3|X^n) \\ &\quad -\frac{1}{2}\Big[H(\Gamma_1|Y^n_{\mathscr{L}(\Gamma_2)})+H(\Gamma_2|Y^n_{\mathscr{L}(\Gamma_2)})-H(\Gamma_1\Gamma_2|Y^n_{\mathscr{L}(\Gamma_2)})\Big] \\ &\quad -\frac{1}{2}\Big[H(\Gamma_1|Y^n_{\mathscr{L}(\Gamma_3)})+H(\Gamma_3|Y^n_{\mathscr{L}(\Gamma_3)})-H(\Gamma_1\Gamma_3|Y^n_{\mathscr{L}(\Gamma_3)})\Big] \\ &\quad -\frac{1}{2}\Big[H(\Gamma_2|Y^n_{\mathscr{L}(\Gamma_3)})+H(\Gamma_3|Y^n_{\mathscr{L}(\Gamma_3)})-H(\Gamma_2\Gamma_3|Y^n_{\mathscr{L}(\Gamma_3)})\Big] \\ &\quad -\frac{1}{2}\Big[H(\Gamma_1\Gamma_2|Y^n_{\mathscr{L}(\Gamma_3)})+H(\Gamma_1\Gamma_3|Y^n_{\mathscr{L}(\Gamma_3)})+H(\Gamma_2\Gamma_3|Y^n_{\mathscr{L}(\Gamma_3)}) \\ &\qquad -2H(\Gamma_1\Gamma_2\Gamma_3|Y^n_{\mathscr{L}(\Gamma_3)})\Big] \\ &\geq I(\Gamma_1;Y^n_{\mathscr{L}(\Gamma_1)})+I(\Gamma_2;Y^n_{\mathscr{L}(\Gamma_2)})+I(\Gamma_3;Y^n_{\mathscr{L}(\Gamma_3)}) \\ &\quad +\frac{1}{2}[I(\Gamma_1\Gamma_2;Y^n_{\mathscr{L}(\Gamma_3)})-I(\Gamma_1\Gamma_2;Y^n_{\mathscr{L}(\Gamma_2)})] \\ &\quad +[I(\Gamma_1\Gamma_2\Gamma_3;X^n)-I(\Gamma_1\Gamma_2\Gamma_3;Y^n_{\mathscr{L}(\Gamma_3)})], \end{aligned} \tag{40}$$

which gives us the desired inequality by using Lemma 3 and Lemma 4 to bound each individual term. Similarly, for the case where $\mathscr{L}(\Gamma_3) < \mathscr{L}(\Gamma_1,\Gamma_2)$ we can write

$$
\begin{aligned}
n(R_1+R_2+R_3+3\varepsilon) \geq\ & H(\Gamma_1)+H(\Gamma_2)+H(\Gamma_3)-H(\Gamma_1\Gamma_2\Gamma_3|X^n) \\
& -\frac{1}{2}\left[H(\Gamma_1|Y^n_{\mathscr{L}(\Gamma_2)})+H(\Gamma_2|Y^n_{\mathscr{L}(\Gamma_2)})-H(\Gamma_1\Gamma_2|Y^n_{\mathscr{L}(\Gamma_2)})\right] \\
& -\frac{1}{2}\left[H(\Gamma_1|Y^n_{\mathscr{L}(\Gamma_3)})+H(\Gamma_3|Y^n_{\mathscr{L}(\Gamma_3)})-H(\Gamma_1\Gamma_3|Y^n_{\mathscr{L}(\Gamma_3)})\right] \\
& -\frac{1}{2}\left[H(\Gamma_2|Y^n_{\mathscr{L}(\Gamma_3)})+H(\Gamma_3|Y^n_{\mathscr{L}(\Gamma_3)})-H(\Gamma_2\Gamma_3|Y^n_{\mathscr{L}(\Gamma_3)})\right] \\
& -\frac{1}{2}\left[H(\Gamma_1\Gamma_2|Y^n_\beta)+H(\Gamma_1\Gamma_3|Y^n_\beta)+H(\Gamma_2\Gamma_3|Y^n_\beta)-2H(\Gamma_1\Gamma_2\Gamma_3|Y^n_\beta)\right] \\
\geq\ & I(\Gamma_1;Y^n_{\mathscr{L}(\Gamma_1)})+I(\Gamma_2;Y^n_{\mathscr{L}(\Gamma_2)})+I(\Gamma_3;Y^n_{\mathscr{L}(\Gamma_3)}) \\
& +\frac{1}{2}[I(\Gamma_1\Gamma_2;Y^n_\beta)-I(\Gamma_1\Gamma_2;Y^n_{\mathscr{L}(\Gamma_2)})] \\
& +\frac{1}{2}[I(\Gamma_1\Gamma_3;Y^n_\beta)-I(\Gamma_1\Gamma_3;Y^n_{\mathscr{L}(\Gamma_3)})] \\
& +\frac{1}{2}[I(\Gamma_2\Gamma_3;Y^n_\beta)-I(\Gamma_2\Gamma_3;Y^n_{\mathscr{L}(\Gamma_3)})] \\
& +[I(\Gamma_1\Gamma_2\Gamma_3;X^n)-I(\Gamma_1\Gamma_2\Gamma_3;Y^n_\beta)]. \qquad (41)
\end{aligned}
$$

where $\beta = \min\{\mathscr{L}(\Gamma_1,\Gamma_2),\mathscr{L}(\Gamma_1,\Gamma_3),\mathscr{L}(\Gamma_2,\Gamma_3)\}$. Now, we can use the above-mentioned lemmas again to bound each individual term. It is clear that (40) and (41) give ($\mathcal{PO}-5$). ■

Remark 1: Note that there is an one-to-one correspondence between the converse proof of Theorem 1 and that of Theorem 4. In fact, here we use the description subsets and their capability of lossy recovering the noisy source layers, where they have been used to losslessly reconstruct the source levels in the A-MLD.

Now we are ready to prove Theorem 2, which is a direct consequence of Theorem 4.

Proof of Theorem 2: We can choose arbitrary values of d_i's, the variance of the additive noise in Theorem 4, such that $d_1 \geq d_2 \geq \cdots \geq d_6 > 0$. One can optimize the bound in Theorem 4 with respect to the values of d_i's, and obtain a bound isolated from d_i's, by replacing them with the optimal choices. Such bound would be the best that can be found using this method. However instead of solving such a difficult optimization problem, we choose $d_i = D_{\mathscr{L}^{-1}(i)}$, for $i = 1,\ldots,6$. It is clear the d_i's satisfy the desired non-increasing order due to the definition of the ordering level. We will later show that this choice gives a bound which is within constant bit gap from the inner bound in Theorem 3.

The single description rate inequalities are exactly the same. The proof of the other inequalities is by straightforward evaluation of their counterparts in Theorem 4, for $d_i = D_{\mathscr{L}^{-1}(i)}$, and applying simple bounds. We do not repeat the same arguments here, and only illustrate such derivation for one simple case. For the sum of two description rates, we can start with $(\mathcal{PO}{-}2)$ and use $d_i = D_{\mathscr{L}^{-1}(i)}$ to get

$$\begin{aligned} R_i + R_j + 2\varepsilon &\overset{(a)}{\geq} \frac{1}{2}\log\frac{1+D_{\Gamma_i}}{D_{\Gamma_i}+D_{\Gamma_i}}\frac{1+D_{\Gamma_j}}{D_{\Gamma_j}+D_{\Gamma_j}}\frac{D_{\Gamma_i\Gamma_j}+\min(D_{\Gamma_i},D_{\Gamma_j})}{(1+\min(D_{\Gamma_i},D_{\Gamma_j}))D_{\Gamma_i\Gamma_j}} \\ &\overset{(b)}{=} \frac{1}{2}\log\frac{1+\max(D_{\Gamma_i},D_{\Gamma_j})}{4D_{\Gamma_1}D_{\Gamma_j}}\frac{D_{\Gamma_i\Gamma_j}+\min(D_{\Gamma_i},D_{\Gamma_j})}{D_{\Gamma_i\Gamma_j}} \\ &\overset{(c)}{\geq} \frac{1}{2}\log\frac{\min(D_{\Gamma_i},D_{\Gamma_j})}{4D_{\Gamma_i}D_{\Gamma_j}D_{\Gamma_i\Gamma_j}} \\ &= -1 + \frac{1}{2}\log\frac{1}{\max(D_{\Gamma_i},D_{\Gamma_j})} + \frac{1}{2}\log\frac{1}{D_{\Gamma_i\Gamma_j}}, \end{aligned} \tag{42}$$

where we have also used the fact $d_{\max(a,b)} = \min(d_a, d_b)$ in (a) which is implied by decreasing ordering of d_i's, (b) is due to the fact that $(1+x)(1+y) = (1+\min(x,y))(1+\max(x,y))$, and (c) holds since $D_{\mathcal{S}}$'s are non-negative. Similar simple manipulations give the other bounds in Theorem 2.

■

B. A Simple Coding Scheme for 3*-Description A-MD: Proof of Theorem 3*

Our approach to prove Theorem 3 is to present a simple scheme with description rates satisfying $(\mathcal{I}{-}1)$–$(\mathcal{I}{-}5)$ which guarantees the distortion constraints. This scheme is based on the successive refinability of Gaussian sources [14], [15], as well as the asymmetric multilevel diversity coding result presented in the previous section. In the encoding scheme, we first produce seven successive refinement layers of the source, and then encode them losslessly.

Successive refinement coding

Consider the non-increasing sequence of distortion constraints $\mathbf{D}' = (D_{\mathscr{L}^{-1}(1)}, D_{\mathscr{L}^{-1}(2)}, \ldots, D_{\mathscr{L}^{-1}(7)})$. Produce seven layers of successive refinement (SR), Ψ_k for $k = 1, 2, \ldots, 7$, such that one can reconstruct the source sequence within distortion constraint $D_{\mathscr{L}^{-1}(k)}$ using $\Psi_1, \ldots, \Psi_k$. Since the Gaussian source is successively refinable [14], it is clear that Ψ_k can be encoded to a binary block of length arbitrary close to

$$nh'_k \triangleq nR(D_{\mathscr{L}^{-1}(k)}) - nR(D_{\mathscr{L}^{-1}(k-1)}) \tag{43}$$

where $R(D) = -\frac{1}{2}\log D$ is the unit variance Gaussian R-D function, and $D_{\mathscr{L}^{-1}(0)} \triangleq 1$. Note that by using fixed length code in SR coding, these blocks are block-wise independently and identically distributed.

Multilevel diversity coding

Now, it only remains to produce the descriptions such that the decoder at level $\mathscr{L}(\mathcal{S})$ can losslessly recover the precoded bitstream SR layers $\Psi_1, \ldots, \Psi_{\mathscr{L}(\mathcal{S})}$, and then reconstruct the Gaussian source sequence within distortion $D_{\mathcal{S}}$. Encoding and decoding of the precoded SR layers are exactly the A-MLD problem. We can simply use the rate region characterization of the A-MLD problem in Theorem 1 to find the achievable rate region of the proposed scheme for A-MD, where only substitution of $V_k = \Psi_k$ and $U_k = (\Psi_1, \ldots, \Psi_k)$ is needed. Therefore we have

$$h_k = \frac{1}{2} \log \frac{D_{\mathscr{L}^{-1}(k-1)}}{D_{\mathscr{L}^{-1}(k)}} \tag{44}$$

and

$$H_k = \sum_{j=1}^{k} h_j = \frac{1}{2} \log \frac{1}{D_{\mathscr{L}^{-1}(k)}}. \tag{45}$$

Replacing the values of H_k's in Theorem 1, we obtain Theorem 3.

It is worth mentioning that although the successive refinement part of the scheme is well-known, producing the descriptions and their rate characterization is not an easy task without the A-MLD result. As an example, consider a system with ordering level $\mathscr{L}_1$ and assume $D_{\Gamma_2} D_{\Gamma_1\Gamma_3} \leq D^2_{\Gamma_3} \leq D_{\Gamma_2} D_{\Gamma_1\Gamma_2}$. An achievable rate triple is

$$(R_1, R_2, R_3) = (\frac{1}{2} \log \frac{D_{\Gamma_2} D_{\Gamma_2\Gamma_3}}{D_{\Gamma_1} D_{\Gamma_3} D_{\Gamma_1\Gamma_2\Gamma_3}}, \frac{1}{2} \log \frac{D_{\Gamma_1\Gamma_3}}{D_{\Gamma_3} D_{\Gamma_2\Gamma_3}}, \frac{1}{2} \log \frac{D_{\Gamma_3}}{D_{\Gamma_2} D_{\Gamma_1\Gamma_3}}), \tag{46}$$

which corresponds to the corner point Y_{12} in regime II of the A-MLD coding problem. The description encoding for this corner point is illustrated in Fig 9. Clearly, the coding scheme for this point matches that for Y_{12} closely, and the SR encoded information in the 3-rd, 4-th and 5-th layers needs to be strategically re-processed using linear codes. Without the underlining A-MLD coding scheme, it appears difficult to devise this coding operation directly.

VI. Conclusion

We formulated the asymmetric multilevel diversity coding problem, an asymmetric counterpart for the symmetric version of the problem. A complete characterization of the admissible rate region is given for the three-description case. We partition the data and apply linear network coding (binary `xor`) on the partitioned subsequences, as a part of the proposed encoding scheme to achieve the upper bound. It turns out that using such a strategy of jointly encoding the independent data streams is crucial, and the outer bound is not achievable without using it, in contrast to the symmetric problem, in which the source-separation coding is known to be optimal.

Using the intuition gained through A-MLD coding problem, we consider the Gaussian asymmetric three description problem. Inner and outer bounds for the admissible rate region are given, and the difference

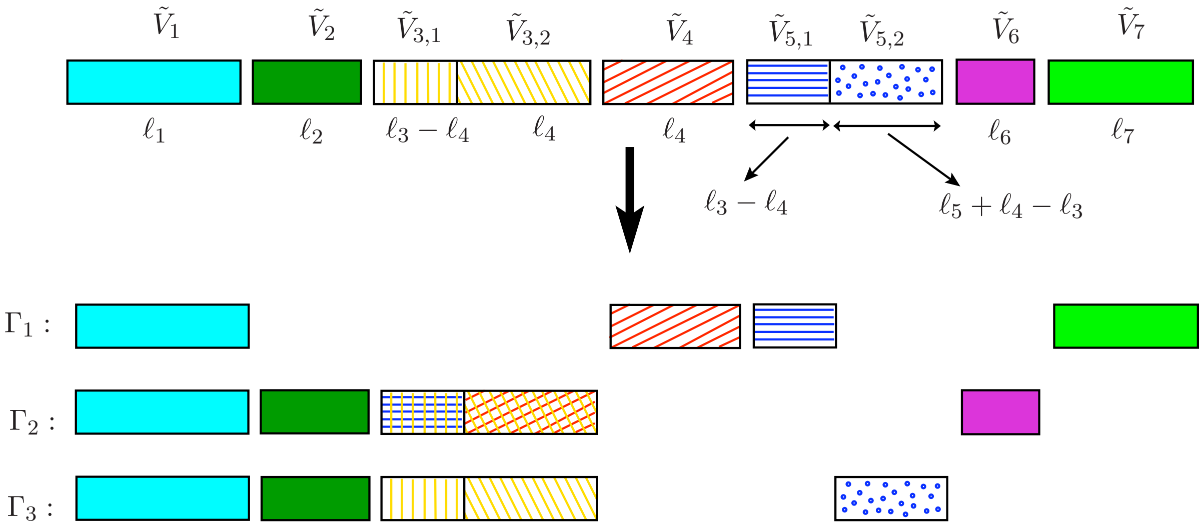

Fig. 9. Description encoding lossless reconstruction for a system with ordering level $\mathscr{L}_1$ and $D_{\Gamma_2} D_{\Gamma_1\Gamma_3} \leq D^2_{\Gamma_3} \leq D_{\Gamma_2} D_{\Gamma_1\Gamma_2}$

between them are shown to be bounded by small universal constants. Though the general asymmetric Gaussian MD rate distortion region is hard to characterize, it is satisfying to see that a simple coding architecture is almost optimal. The A-MLD coding problem plays a key role in establishing these results, which further strengthens the connection between the MLD coding and the MD problem. Philosophically, this work is related to the approximation results obtained in the context of the interference and relay networks [19], [20], and further illustrates the effectiveness of the general approach of first treating the lossless (deterministic) coding problem, and then deriving approximate characterization for its lossy (noisy) counterpart.

Appendix A

Proof of Lemma 3:

$$
\begin{aligned}
I(\mathcal{S};Y_i^n) &= h(Y_i^n) - h(Y_i^n|\mathcal{S}) \\
&= h(Y_i^n) - h(Y_i^n - \hat{X}_{\mathcal{S}}^n|\mathcal{S}) \\
&\geq h(Y_i^n) - h(X^n + Z_i^n - \hat{X}_{\mathcal{S}}^n) \\
&\geq h(Y_i^n) - \sum_{t=1}^{n} h(X(t) - \hat{X}_{\mathcal{S}}(t) + Z_i(t)) \\
&\overset{(a)}{\geq} \frac{n}{2}\log(2\pi e(1+d_i)) - \sum_{t=1}^{n}\frac{1}{2}\log\left(2\pi e(\mathbb{E}(X(t)-\hat{X}_{\mathcal{S}}(t))^2 + d_i)\right) \\
&\overset{(b)}{\geq} \frac{n}{2}\log(1+d_i) - \frac{n}{2}\log\left(\mathbb{E}d(X^n,\hat{X}_{\mathcal{S}}^n) + d_i\right) \\
&\overset{(c)}{\geq} \frac{n}{2}\log\frac{1+d_i}{D_{\mathcal{S}}+d_i}
\end{aligned}
\tag{47}
$$

where (a) is due to the fact that the entropy of any random variable is upper bounded by that of a Gaussian variables with the same variance; (b) is implied by concavity of the function $\log(x)$; and in (c) we have used the fact that $\log(x+a)$ is an increasing function in x. ■

Proof of Lemma 4: Note that

$$
\begin{aligned}
h(Y_i^n|\mathcal{S}) - h(Y_j^n|\mathcal{S}) &\overset{(a)}{=} h(Y_i^n|\mathcal{S}) - h(Y_j^n|\mathcal{S}, Z_i^n - Z_j^n) \\
&\overset{(b)}{=} h(Y_i^n|\mathcal{S}) - h(Y_i^n|\mathcal{S}, Z_i^n - Z_j^n) \\
&= I(Y_i^n; Z_i^n - Z_j^n|\mathcal{S}) \\
&= h(Z_i^n - Z_j^n|\mathcal{S}) - h(Z_i^n - Z_j^n|\mathcal{S}, Y_i^n) \\
&\overset{(c)}{\geq} h(Z_i^n - Z_j^n) - h(Z_i^n - Z_j^n|Y_i^n - \hat{X}_{\mathcal{S}}^n) \\
&\geq \sum_{t=1}^{n} h(Z_i(t) - Z_j(t)) - h(Z_i(t) - Z_j(t)|Y_i(t) - \hat{X}_{\mathcal{S}}(t)) \\
&= \sum_{t=1}^{n} I(Z_i(t) - Z_j(t); X(t) - \hat{X}_{\mathcal{S}}(t) + Z_i(t)) \\
&\overset{(d)}{\geq} \sum_{t=1}^{n}\frac{1}{2}\log\frac{\mathbb{E}(X(t)-\hat{X}_{\mathcal{S}}(t))^2 + d_i}{\mathbb{E}(X(t)-\hat{X}_{\mathcal{S}}(t))^2 + d_j} \\
&\overset{(e)}{\geq} \frac{n}{2}\log\frac{D_{\mathcal{S}}+d_i}{D_{\mathcal{S}}+d_j}
\end{aligned}
\tag{48}
$$

where (a) holds because Y_j^n is independent of $Z_i^n - Z_j^n = N_i^n + \cdots + N_{j-1}^n$ for $i < j$; the equality in (b) is because of $Y_i = Y_j + (Z_i - Z_j)$; (c) is due to the data processing inequality and the fact that $Z_i^n - Z_j^n$ is purely noise and independent of X^n and therefore $\mathcal{S}$; in (d) we use the worst noise lemma in [12], [17]; and (e) is due to convexity and monotonicity of $\log(x+a)/(x+b)$ in x when $a \geq b$. Therefore, we simply have

$$\begin{aligned} I(\mathcal{S}; Y_j^n) - I(\mathcal{S}; Y_i^n) &= h(Y_j^n) - h(Y_j^n|\mathcal{S}) - h(Y_i^n) + h(Y_i^n|\mathcal{S}) \\ &\geq \frac{n}{2} \log \frac{1+d_j}{1+d_i} + \frac{n}{2} \log \frac{D_{\mathcal{S}} + d_i}{D_{\mathcal{S}} + d_j} \\ &= \frac{n}{2} \log \frac{(1+d_j)(D_{\mathcal{S}} + d_i)}{(1+d_i)(D_{\mathcal{S}} + d_j)}. \end{aligned} \tag{49}$$

■

REFERENCES

[1] A. El Gamal and T. Cover, "Achievable Rates for Multiple Descriptions," *IEEE Trans. Inf. Theory*, vol. 28, no. 6, pp. 851–857, Nov. 1982.

[2] L. Ozarow, "On a Source-Coding Problem with Two Channels and Three Receivers," *Bell System Technical Journal*, vol. 59, no. 10, pp. 1909–1921, Dec. 1980.

[3] Z. Zhang and T. Berger, "New Results in Binary Multiple Descriptions," *IEEE Trans. Inf. Theory*, vol. 33, no. 4, pp. 502–821, July 1987.

[4] S. Mohajer, C. Tian, and S. N. Diggavi, "Asymmetric multilevel diversity coding," in *Data Compression Conference*, Snowbird, UT, USA, Mar. 2007, pp. 412–421.

[5] C. Tian, S. Mohajer, and S. N. Diggavi, "On the symmetric Gaussian multiple descriptions rate-distortion function," *Data Compression Conference*, Snowbird, UT, USA, Mar. 2007, pp. 402–411.

[6] S. Mohajer, C. Tian, and S. Diggavi, "Asymmetric Gaussian Multiple Descriptions and Asymmetric multilevel Diversity Coding," in *Proc. 2008 IEEE Int. Symp. Information Theory*, Toronto, ON, Canada, Jul. 2008, pp. 1992–1996.

[7] C. Tian, S. Mohajer, and S. N. Diggavi, "Approximating the Gaussian multiple description rate region under symmetric distortion constraints," in *Proc. 2008 IEEE Int. Symp. Information Theory*, Toronto, ON, Canada, Jul. 2008. pp. 1413–1417.

[8] C. Tian, S. Mohajer and S. N. Diggavi, "Approximating the Gaussian Multiple Description Rate Region under Symmetric Distortion Constraints," *I*EEE Trans. Inf. Theory, vol. 55, no. 8, pp. 3869–3891, Aug. 2009.

[9] J. R. Roche, R. W. Yeung, and K. P. Hau, "Symmetrical multilevel diversity coding," *IEEE Trans. Inf. Theory*, vol. 43, no. 5, pp. 1059–1064, May 1997.

[10] R. W. Yeung and Z. Zhang, "On symmetrical Multilevel diversity coding," *IEEE Trans. Inf. Theory*, vol. 45, no. 2, pp. 609–621, Mar. 1999.

[11] K. P. Hau, Multilevel Diversity Coding with Independent Data Streams, Master's Thesis, The Chinese University of Hong Kong, 1995.

[12] T. M. Cover and J. A. Thomas, *Elements of information theory*, New York: Wiley, 1991.

[13] R. Ahlswede, N. Cai, S. Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. Inf. Theory*, vol. 46, no. 4, pp. 1204–1216, Apr. 2000.

[14] W. H. R. Equitz and T. M. Cover, “Successive refinement of information,” *IEEE Trans. Inf. Theory*, vol. 37, no. 2, pp. 269–275, Mar. 1991.

[15] B. Rimoldi, “ Successive refinement of information: characterization of the achievable rates,” *IEEE Trans. Inf. Theory*, vol. 40, no. 1, pp 253–259, Jan. 1994.

[16] R. Puri, S.S. Pradhan, and K. Ramchandran, “n-channel symmetric multiple descriptions - Part II: an achievable rate-distortion region,” *IEEE Trans. Inf. Theory*, vol. 51, pp. 1377–1392, Apr. 2005.

[17] S. Diggavi and T. M. Cover, “Worst additive noise under covariance constraints,” *IEEE Trans. Inf. Theory*, vol. 47, no. 7, pp. 3072–3081, Nov. 2001.

[18] A. J. Goldman, “Resolution and separation theorems for polyhedral convex sets,” in *Linear Inequalities and Related Systems*, H. W. Kuhn and A. W. Tucker, Eds. Princeton, NJ: Princeton Univ. Press, 1956.

[19] R. Etkin, D. N. C. Tse, and H. Wang, “Gaussian interference channel capacity to within one bit,” *IEEE. Trans. Inf. Theory*, vol. 54, no. 12, pp. 5534-5562, Dec. 2008.

[20] A. S. Avestimehr, S. Diggavi, and D. N. C. Tse, “Wireless Network Information Flow: A Deterministic Approach,” available at `http://arxiv.org/abs/0906.5394`.