On the Rate Loss of Multiple Description Source Codes

Hanying Feng, *Member, IEEE,* and Michelle Effros, *Senior Member, IEEE*

***Abstract*—The rate loss of a multiresolution source code (MRSC) describes the difference between the rate needed to achieve distortion D_i in resolution i and the rate-distortion function $R(D_i)$. This paper generalizes the rate loss definition to multiple description source codes (MDSCs) and bounds the MDSC rate loss for arbitrary memoryless sources. For a two-description MDSC (2DSC), the rate loss of description i with distortion D_i is defined as $L_i = R_i - R(D_i)$, $i = 1, 2$, where R_i is the rate of the ith description; the joint rate loss associated with decoding the two descriptions together to achieve central distortion D_0 is measured either as $L_0 = R_1 + R_2 - R(D_0)$ or as $L_{12} = L_1 + L_2$. We show that for any memoryless source with variance σ^2, there exists a 2DSC for that source with $L_1 \leq \frac{1}{2}$ or $L_2 \leq \frac{1}{2}$ and a) $L_0 \leq 1$ if $D_0 \leq D_1 + D_2 - \sigma^2$, b) $L_{12} \leq 1$ if $1/D_0 \leq 1/D_1 + 1/D_2 - 1/\sigma^2$, c) $L_0 \leq L_{G0} + 1.5$ and $L_{12} \leq L_{G12} + 1$ otherwise, where L_{G0} and L_{G12} are the joint rate losses of a Gaussian source with variance σ^2.**

***Index Terms*—Lossy source coding, rate-distortion function, Shannon lower bound.**

I. Introduction

A MULTIPLE description source code (MDSC) is a compression system built for a lossy packet-based network. For problems where retransmission of lost packets is prohibitively expensive (due, for example, to delay constraints or limits on retransmission requests) the receiver may wish to build a data reconstruction with an incomplete subset of the transmitted packets. The goal in MDSC design is to achieve a code that yields good rate-distortion performance under a variety of packet-loss scenarios. Fig. 1 shows a two-packet MDSC (2DSC). The performance is given by $(R_1, R_2, D_0, D_1, D_2)$, where (R_i, D_i) are the expected rate and distortion for packet $i \in \{1, 2\}$ and $D_0 \leq \min\{D_1, D_2\}$ is the distortion in jointly decoding the two packets.

The *rate loss* of an MDSC is its performance penalty relative to a single-resolution code with the same distortion. More precisely, the rate loss of a 2DSC achieving performance $(R_1, R_2, D_0, D_1, D_2)$ in the limit of large coding dimension is $L = (L_0, L_1, L_2)$, where $L_i = R_i - R(D_i)$ $(i \in \{0, 1, 2\})$ (here $R_0 = R_1 + R_2$) and $R(D)$ is the rate-distortion function for the given source.

This paper describes source-independent upper bounds for the MDSC rate loss. Rate loss bounds are useful for several reasons. 1) They describe the performance degradation associated with using the given code rather than the best traditional code with the same distortion. For example, small, constant upper bounds on the rate loss of multiresolution source codes (MRSCs) (see, e.g., [1], [2]) put to rest fears that there might exist some source on which the cost of multiresolution coding can be made arbitrarily large. 2) They give new achievability results that provide elegant and often tight inner bounds on the region of achievable rates and distortions. These bounds are simpler to analyze than existing alternatives for which solution requires a complex multidimensional optimization for every source (e.g., [3] and [4]). 3) Since the exact rate-distortion regions for MDSCs are not known in general, the rate loss also gives a good bound on the distance between the best existing achievability results and converses. (The rate-distortion function provides a natural converse for general sources. In particular

$$R_1 \geq R(D_1) \tag{1}$$

$$R_2 \geq R(D_2) \tag{2}$$

$$R_1 + R_2 \geq R(D_0) \tag{3}$$

for any $(R_1, R_2, D_0, D_1, D_2)$ achievable by a 2DSC.)

The remainder of this paper is organized as follows. Section II introduces background material, notation, and definitions. Section III lists our main results; the proofs of these results appear in Section IV. Supporting lemmas and their proofs appear in the Appendix.

II. Preliminaries

Let $\{X_i\}_{i=1}^{\infty}$ be a real-valued independent and identically distributed (i.i.d.) source with probability density function (pdf) $f_X(x)$. Let d be a real-valued nonnegative difference distortion measure, $d(x, y) = \rho(x - y)$ for any $x, y \in \boldsymbol{R}$ and some function $\rho : \boldsymbol{R} \to [0, \infty)$. Assume that ρ is continuous and that there exists a reference letter $y^* \in \boldsymbol{R}$ such that $E_x d(x, y^*) < \infty$. For any $x^n, y^n \in \boldsymbol{R}^n$, define

$$d_n(x^n, y^n) = \frac{1}{n} \sum_{i=1}^{n} d(x_i, y_i).$$

The *rate-distortion function* for source $\{X_i\}_{i=1}^{\infty}$ and distortion measure d is

$$R(D) = \inf_{Y: Ed(X,Y) \leq D} I(X; Y)$$

Manuscript received May 22, 2002; revised January 4, 2004. This work was supported in part by the National Science Foundation under Grants CCR-9909026 and CCR-0220039, and by the Lee Center for Advanced Networking at the California Institute of Technology. The material in this paper was presented in part at the IEEE International Symposium on Information Theory, Lausanne, Switzerland, June/July 2002.

H. Feng is with the Department of Electrical Engineering, Mail Stop 4075, Stanford University, Stanford, CA 94305 USA (e-mail: fhy@stanford.edu).

M. Effros is with the Department of Electrical Engineering, MC 136-93, California Institute of Technology, Pasadena, CA 91125 USA (e-mail: effros@caltech.edu).

Communicated by R. Zamir, Associate Editor for Source Coding.
Digital Object Identifier 10.1109/TIT.2004.840900

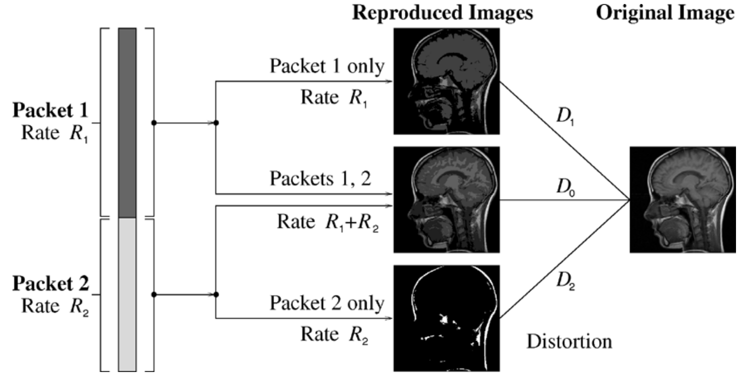


Fig. 1. A 2DSC. Decoding the first binary description with R_1 bits per symbol (bps) yields a reproduction with distortion D_1, decoding another description with R_2 bps yields a reproduction of distortion D_2, and decoding both of them jointly yields a reproduction of distortion D_0.

which characterizes the minimum rate required to describe source X with distortion not exceeding D. In the arguments that follow, we frequently assume that there exists a conditional pdf $f(y|x)$ that achieves $R(D)$. This assumption simplifies the exposition considerably but is not a necessary condition for any of our results.

An (n, M_1, M_2) 2DSC consists of two individual encoder/decoder pairs and a joint decoder. For packet $i \in \{1,2\}$, the encoder/decoder pair $(f_n^{(i)}, g_n^{(i)})$ are mappings

$$f_n^{(i)} : \boldsymbol{R}^n \to \{1,\dots,M_i\} \text{ and } g_n^{(i)} : \{1,\dots,M_i\} \to \boldsymbol{R}^n$$

with distortion $Ed_n(X^n, g_n^{(i)}(f_n^{(i)}(X^n)))$ and rate $\frac{1}{n}\log M_i$. The joint decoder $g_n^{(0)}$ is a mapping

$$g_n^{(0)} : \{1,\dots,M_1\} \times \{1,\dots,M_2\} \to \boldsymbol{R}^n$$

with distortion $Ed_n(X^n, g_n^{(0)}(f_n^{(1)}(X^n), f_n^{(2)}(X^n)))$ and total rate $\frac{1}{n}\log(M_1M_2)$.

The rate-distortion vector (R_1,R_2,D_0,D_1,D_2) is 2DSC-achievable if for any $\epsilon > 0$ and for sufficiently large n, there exists an (n, M_1, M_2) 2DSC such that

$$Ed_n(X^n, g_n^{(1)}(f_n^{(1)}(X^n))) \le D_1+\epsilon,\ \frac{1}{n}\log M_1 \le R_1+\epsilon$$
$$Ed_n(X^n, g_n^{(2)}(f_n^{(2)}(X^n))) \le D_2+\epsilon,\ \frac{1}{n}\log M_2 \le R_2+\epsilon$$
$$Ed_n(X^n, g_n^{(0)}(f_n^{(1)}(X^n), f_n^{(2)}(X^n))) \le D_0+\epsilon.$$

For each (D_0,D_1,D_2), define

$$\mathcal{R}(D_0,D_1,D_2) = \{(R_1,R_2) : (R_1,R_2,D_0,D_1,D_2) \text{ is 2DSC-achievable}\}.$$

Theorem 1, given later, describes a set of 2DSC-achievable rate-distortion vectors. The result for finite alphabets comes from [3]. A generalization to well-behaved continuous sources appears in [5]. The result is not tight in general [6], but is tight when $R_1+R_2 = R(D_0)$ (the "no excess total rate" case) [7]. There exist some additional achievability results and converses for the special case where $R_1 = R(D_1)$ and $R_2 = R(D_2)$ (the "no excess marginal rate" case) [8]. Calculating the full achievable region under the constraint $R_1+R_2 = R(D_0)$ or the constraint $R_1 = R(D_1)$ and $R_2 = R(D_2)$ is nontrivial even for simple sources due to the minimization in the characterization of the achievable region. A more general form for stationary sources appears in [4]. Unfortunately, even for memoryless sources, this result is given as a limit as the coding dimension n grows without bound, making the result difficult to apply except in very special cases (e.g., for a nonergodic mixture of Gaussians). This paper builds on the single-letter result described below.

Theorem 1: [3, Theorem 1], [5, Theorem 1] For any i.i.d. source $\{X_i\}_{i=1}^\infty$ with density $f_X(x)$ and distortion measure d, (R_1,R_2,D_0,D_1,D_2) is 2DSC-achievable if there exists a conditional probability $Q_{Y_0,Y_1,Y_2|X}$ such that

$$\begin{aligned} R_1 &\ge I(X;Y_1), & Ed(X,Y_1) &\le D_1\\ R_2 &\ge I(X;Y_2), & Ed(X,Y_2) &\le D_2\\ R_1+R_2 &\ge I(X;Y_0,Y_1,Y_2)+I(Y_1;Y_2), & Ed(X,Y_0) &\le D_0. \end{aligned}$$

Following Zamir's approach from [9], [1] and [2] give source-independent, constant bounds on the rate loss for MRSCs. This paper relies on similar tools.

III. Main Results

For notational simplicity, assume (without loss of generality) that $E(X) = 0$. Further assume that $d(x,\hat{x}) = (x-\hat{x})^2$ (the mean squared error (mse) distortion measure), that $0 < \sigma^2 < \infty$, that the differential entropy $h(X)$ is finite, and that $0 < D_0 < D_1, D_2 \le \sigma^2$.

A. Rate Loss Bounds

Partition the space

$$\{(D_0,D_1,D_2) : 0 < D_0 < D_1, D_2 \le \sigma^2\}$$

of possible distortion vectors into three regions

$$\mathcal{D}_1 = \{(D_0,D_1,D_2) : 0 < D_0 \le D_1+D_2-\sigma^2\}$$
$$\mathcal{D}_2 = \{(D_0,D_1,D_2) : D_1+D_2-\sigma^2 < D_0 < (1/D_1+1/D_2-1/\sigma^2)^{-1}\}$$
$$\mathcal{D}_3 = \{(D_0,D_1,D_2) : D_0 \ge (1/D_1+1/D_2-1/\sigma^2)^{-1}\}.$$

Here

$$D_1+D_2-\sigma^2 \le (1/D_1+1/D_2-1/\sigma^2)^{-1}$$

by Lemma 1 in the Appendix. In region $\mathcal{D}_1$, $R(D_1)+R(D_2) \le R(D_0)$ by Lemma 2 in the Appendix; thus, the constraint $R_1+R_2 \ge R(D_0)$ (3) dominates the constraint $R_1+R_2 \ge R(D_1)+R(D_2)$ implied by bounds (1) and (2). Thus, region $\mathcal{D}_1$ is called the D_0-critical region. It can be shown that in region $\mathcal{D}_3$, called the (D_1,D_2)-critical region, $R(D_0) \le R(D_1)+R(D_2)+1$;

in this case, the constraint $R_1 + R_2 \geq R(D_1) + R(D_2)$ provides a good converse, and bounds on the new rate-loss term $L_{12} = R_1 + R_2 - (R(D_1) + R(D_2))$ provide a good measure of achievable performance. The relationship between L_0 and L_{12} is

$$L_0 + R(D_0) = L_{12} + R(D_1) + R(D_2). \quad (4)$$

In region $\mathcal{D}_2$, called the mid-region, either constraint can dominate and therefore bounds on both L_0 and L_{12} are potentially useful. (To show that either bound can dominate in $\mathcal{D}_2$, consider a Gaussian source with variance σ^2. If $D_1 = D_2 = \sigma^2/4$ and $D_0 \to 0$, then constraint (3) dominates; while if $D_1 = D_2 \to 0$ and $D_0 = (2/D_1 + 2/D_2 - 2/\sigma^2)^{-1}$, then the constraint $R_1 + R_2 \geq R(D_1) + R(D_2)$ dominates. In either case, the gap between the two constraints can be arbitrarily large.)

Given a fixed vector (D_0, D_1, D_2), let $d_i = D_i/\sigma^2$ for all $i \in \{0, 1, 2\}$ denote the normalized distortions. Next define $L_{\min}$, L_{G0}, and L_{G12} as shown in (5)–(7) at the bottom of the page. Given a Gaussian source N with variance σ^2, $L_{\min}$ is the rate-distortion function for N at distortion $\max\{D_1, D_2\}$ and L_{G0} and L_{G12} are the rate losses L_0 and L_{12} for N assuming that $(D_0, D_1, D_2) \in \mathcal{D}_2$ by [3], [10] and Lemma 3 in the Appendix.

The following results bound the achievable rate loss for the three regions.

Theorem 2: For any $i \in \{1, 2\}$ and any $(D_0, D_1, D_2) \in \mathcal{D}_1 \bigcup \mathcal{D}_2$

$$\begin{aligned} L_i &\leq \frac{1}{2}\log(2 - d_i) \leq \frac{1}{2} \\ L_0 &\leq \frac{1}{2}\log(2 - d_0) + \min\Big\{L_{\min}, \\ &\qquad R(\max\{D_1, D_2\}) + \frac{1}{2}\log(2 - \max\{d_1, d_2\})\Big\} \\ &\leq \min\Big\{L_{\min} + \frac{1}{2}, R(\max\{D_1, D_2\}) + 1\Big\} \end{aligned}$$

are simultaneously achievable.

Theorem 3: For any $i \in \{1, 2\}$ and any $(D_0, D_1, D_2) \in \mathcal{D}_2$, $L_i \leq \frac{1}{2}$ and $L_{12} \leq L_{G12} + 1$ are simultaneously achievable.

Theorem 4: For any $(D_0, D_1, D_2) \in \mathcal{D}_3$

$$L_1 \leq \frac{1}{2}\log(2 - d_1) \leq \frac{1}{2} \quad \text{and} \quad L_2 \leq \frac{1}{2}\log(2 - d_2) \leq \frac{1}{2}$$

are simultaneously achievable.

Theorem 2 leads to Corollaries 1–4 for distinct subsets of the region $\mathcal{D}_1 \bigcup \mathcal{D}_2$ covered there.

Corollary 1: For any $i \in \{1, 2\}$ and any $(D_0, D_1, D_2) \in \mathcal{D}_1$

$$L_i \leq \frac{1}{2}\log(2 - d_i) \leq \frac{1}{2}$$

and

$$L_0 \leq \frac{1}{2}\log[2(2 - d_0)/(1 + d_0)] \leq 1$$

are simultaneously achievable.

Corollary 2: For any $i \in \{1, 2\}$ and any $(D_0, D_1, D_2) \in \mathcal{D}_2$, $L_i \leq \frac{1}{2}$ and $L_0 \leq L_{G0} + 1.5$ are simultaneously achievable.

Corollary 3: For any $i \in \{1, 2\}$ and any $(D_0, D_1, D_2) \in \mathcal{D}_2$ with $d_1 \geq \frac{1}{2}$ or $d_2 \geq \frac{1}{2}$, $L_i \leq \frac{1}{2}$ and $L_0 \leq 1$ are simultaneously achievable.

Corollary 4: For any $i \in \{1, 2\}$ and any $(D_0, D_1, D_2) \in \mathcal{D}_2$

$$L_i \leq \frac{1}{2} \quad \text{and} \quad L_{12} \leq \frac{1}{2}\log\left[\min\left\{\frac{d_1}{d_0}, \frac{d_2}{d_0}\right\}\right] + 1$$

are simultaneously achievable.

Theorems 5 and 6 bound the distance between the upper and lower bounds on the rate loss in region $\mathcal{D}_2$ under the assumption that $\max\{d_1, d_2\} < 1/2$. (These are the only conditions under which the previously described theorems fail to give a constant bound.)

We first define the entropy power of X as

$$P_X = 2^{2h(X)}/(2\pi e)$$

and define function $LG(\delta_0, \delta_1, \delta_2)$ as given in (8) also at the bottom of the page. Notice that $P_X \leq \sigma^2$ and $L_{G0} = LG(D_0/\sigma^2, D_1/\sigma^2, D_2/\sigma^2)$.

In [10], it is shown that if $0 < D_0 < D_1, D_2 \leq P_X$ and $D_1 + D_2 - P_X < D_0 < (1/D_1 + 1/D_2 - 1/P_X)^{-1}$, then the region described by

$$\begin{aligned} \Big\{(R_1, R_2) : R_1 &\geq R(D_1), \\ R_2 &\geq R(D_2), \\ R_1 + R_2 &\geq \frac{1}{2}\log\frac{P_X}{D_0} + LG\left(\frac{D_0}{P_X}, \frac{D_1}{P_X}, \frac{D_2}{P_X}\right)\Big\} \end{aligned}$$

is a converse for the achievable region, which also implies that

$$L_0 \geq \frac{1}{2}\log\frac{P_X}{D_0} + LG\left(\frac{D_0}{P_X}, \frac{D_1}{P_X}, \frac{D_2}{P_X}\right) - R(D_0). \quad (9)$$

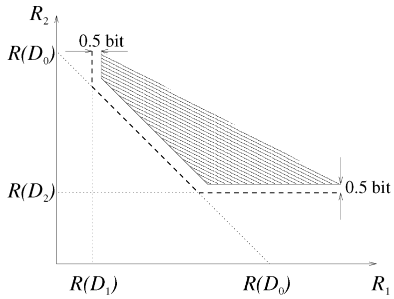
In the next two theorems, we compare the upper bound on the rate loss drawn from the previous theorems to the lower bound shown in (9).

$$L_{\min} = \min\left\{\frac{1}{2}\log\frac{1}{d_1}, \frac{1}{2}\log\frac{1}{d_2}\right\} = \frac{1}{2}\log\frac{\sigma^2}{\max\{D_1, D_2\}} \quad (5)$$

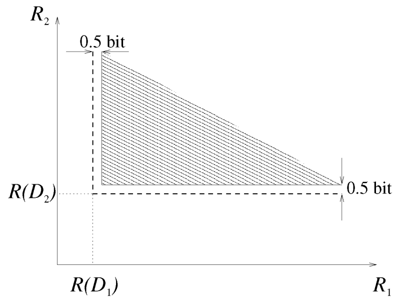
$$L_{G0} = \frac{1}{2}\log\frac{(1 - d_0)^2}{(1 - d_0)^2 - (\sqrt{(1 - d_1)(1 - d_2)} - \sqrt{(d_1 - d_0)(d_2 - d_0)})^2} \quad (6)$$

$$L_{G12} = \frac{1}{2}\log\frac{(1 - d_0)^2 d_1 d_2}{(1 - d_0)^2 d_1 d_2 - (d_0\sqrt{(1 - d_1)(1 - d_2)} - \sqrt{(d_1 - d_0)(d_2 - d_0)})^2} \quad (7)$$

$$LG(\delta_0, \delta_1, \delta_2) = \frac{1}{2}\log\frac{(1 - \delta_0)^2}{(1 - \delta_0)^2 - (\sqrt{(1 - \delta_1)(1 - \delta_2)} - \sqrt{(\delta_1 - \delta_0)(\delta_2 - \delta_0)})^2}. \quad (8)$$



(a)



(b)

Fig. 2. (a) Achievability result and converse from Corollary 1 for $(D_0, D_1, D_2) \in \mathcal{D}_1$. (b) Achievability result and converse from Theorem 4 on $(D_0, D_1, D_2) \in \mathcal{D}_3$.

Theorem 5: For any $i \in \{1, 2\}$ and any $(D_0, D_1, D_2) \in \mathcal{D}_2$ for which $\max\{d_1, d_2\} < 1/2$, there exists an $(R_1, R_2) \in \mathcal{R}(D_0, D_1, D_2)$ with $L_i \le \frac{1}{2}\log(2 - d_i) \le \frac{1}{2}$, and the distancebetween the upper bound and the lower bound for L_0 is no more than

$$\min\left\{\log(2\pi e\sigma^2) - 2h(X)+1, \frac{1}{2}\log(2\pi e\sigma^2)-h(X)+1.5\right\}.$$

Theorem 6: For any $i \in \{1, 2\}$ and any $(D_0, D_1, D_2) \in \mathcal{D}_2$ with $\max\{d_1, d_2\} < 1/2$, if the Shannon lower bound (SLB) is tight at distortions D_1 and D_2, i.e.,

$$R(D_j) = h(X) - \frac{1}{2}\log(2\pi e D_j), \qquad \text{for all } j \in \{1, 2\}$$

then there exists an $(R_1, R_2) \in \mathcal{R}(D_0, D_1, D_2)$ with $L_i \le \frac{1}{2}\log(2 - d_i) \le \frac{1}{2}$, and the distance between the upper bound and the lower bound for L_0 is less than or equal to 2.

B. Discussion of the Rate Loss Bounds

If X is Gaussian, then for any $(D_0, D_1, D_2) \in \mathcal{D}_1$

$$\mathcal{R}(D_0, D_1, D_2) = \{(R_1, R_2) : R_1 \ge R(D_1), \\ R_2 \ge R(D_2), R_1 + R_2 \ge R(D_0)\}$$

by [3]. Corollary 1 gives the corresponding result for an arbitrary i.i.d. source; if X is i.i.d., then for any $(D_0, D_1, D_2) \in \mathcal{D}_1$

$$\mathcal{R}(D_0, D_1, D_2) \supseteq \left\{(R_1, R_2) : R_1 \ge R(D_1) + \frac{1}{2}, \right. \\ \left. R_2 \ge R(D_2) + \frac{1}{2}, R_1 + R_2 \ge R(D_0) + 1\right\}. \quad (10)$$

Fig. 2(a) illustrates this property. The dashed lines trace the converse given by the bounds (1)–(3); any rate pair below this bound is not MDSC-achievable. The shaded region surrounded by solid lines is the achievability result described in (10); any rate pair inside this region is MDSC-achievable. In the special case where $D_2 = \sigma^2$, MDSC reduces to MRSC, and all of the bounds of [2] apply. Corollary 1's bound on L_0 depends only on d_0; it is tight when $D_0 = \sigma^2$. Fig. 3 plots this bound.

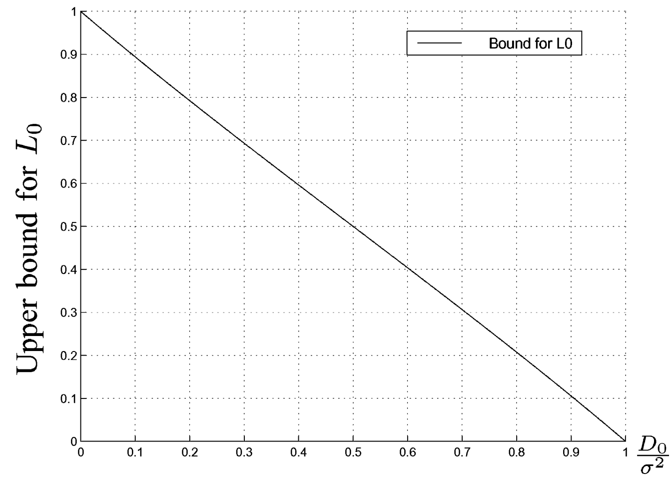


Fig. 3. Bound for L_0 in Corollary 1.

Corollary 2 similarly relates the rate loss in region $\mathcal{D}_2$ to the rate loss of the Gaussian source in the same region. While the latter rate loss may be large in general, it is small when $\max(d_1, d_2) \ge \frac{1}{2}$ (e.g., see Corollary 5 in the Appendix). Corollary 3 generalizes this property to all i.i.d. sources. Corollary 4 parallels Corollary 2 but treats L_{12} instead of L_0.

By Lemmas 3 and 4 in the Appendix, Theorem 3 is tighter than both Corollary 4 and bound

$$L_{12} \le 1 + \frac{1}{2}\log\frac{d_1 d_2}{d_0} + \frac{1}{2}\log\frac{1}{d_1 + d_2 - d_1 d_2} \quad (11)$$

drawn from (33) in the proof of Lemma 4 in the Appendix. Corollary 4 and the bound (11) are easier to analyze. The difference between Theorem 3 and the farther of Corollary 4 and (11) is always less than 1 bit.

When $R_1 = R(D_1)$ and $R_2 = R(D_2)$, achieving $D_0 = \min\{D_1, D_2\}$ is trivial, and it seems intuitively reasonable to expect $D_0 < \min\{D_1, D_2\}$. Theorem 4 justifies this intuition to some extent, demonstrating that for any source and any fixed D_1 and D_2, there always exists a multiple description code with $R_1 \approx R(D_1)$, $R_2 \approx R(D_2)$, and $D_0 \le (1/D_1 + 1/D_2 - 1/\sigma^2)^{-1}$. Fig. 2(b) illustrates $\mathcal{R}(D_0, D_1, D_2)$ for $(D_0, D_1, D_2) \in \mathcal{D}_3$. The dashed lines depict the converse characterized by (1) and (2) (here $R(D_0)$ is not drawn since $R(D_0)$ is typically less than $R(D_1)+R(D_2)$ and always less than $R(D_1)+R(D_2)+1$); the shaded region surrounded by solid lines is the achievable rates given by Theorem 4.

It can be shown that for smooth sources, if the SLB is not tight for $R(D)$ at any distortion $D > 0$, it is asymptotically tight as $D \to 0$. So Theorem 6 generally works for the high-resolution scenario. However, for some specific scenario, a better result has been shown in [10]: if the source is smooth and $D_1, D_2 \to 0$ and D_1/D_0 and D_2/D_0 are held fixed, then the upper bound and the lower bound for L_0 from [10] coincide asymptotically, i.e., the distance between the upper and lower bounds approaches 0 in the high-resolution limit.

The preceding results yield as a by-product some interesting inequalities about the rate-distortion function of an arbitrary i.i.d. source

$$R(D_1)+R(D_2)\le R(D_0)+1, \quad \text{if } (D_0,D_1,D_2)\in\mathcal{D}_1 \tag{12}$$

$$\begin{aligned}R(D_0)-L_{G12}-1 &\le R(D_1)+R(D_2)\\ &\le R(D_0)+L_{G0}+1.5,\\ &\quad \text{if } (D_0,D_1,D_2)\in\mathcal{D}_2\end{aligned} \tag{13}$$

$$R(D_1)+R(D_2)\le R(D_0)+1, \text{ if } (D_0,D_1,D_2)\in\mathcal{D}_2 \text{ and } \max\{D_1,D_2\}\ge \sigma^2/2 \tag{14}$$

$$R(D_1)+R(D_2)\ge R(D_0)-1, \quad \text{if } (D_0,D_1,D_2)\in\mathcal{D}_3. \tag{15}$$

Here (12) comes from Corollary 1; (13) comes from Corollary 2 and Theorem 3; (14) comes from Corollary 3; and (15) comes from Theorem 4. In fact, tighter versions of (12) and (14) can be drawn immediately from the properties of the rate-distortion function (e.g., convexity), as shown in Lemmas 2 and 5. The other inequalities are less obvious.

IV. Proofs of Main Results

This section contains proofs of the theorems. Generally speaking, the proofs involve finding a Gaussian approximation of the optimizing reproduction distributions for Y_0, Y_1, and Y_2 and bounding the optimal rate loss by the rate loss of the Gaussian. The reconstructions in Theorem 2 satisfy $X\to Y_0\to (Y_1,Y_2)$. The reconstructions in Theorems 3 and 4, like those in [10], satisfy $X\to (Y_1,Y_2)\to Y_0$.

Let $\beta_i = 1-d_i$ for $i=0,1,2$. Further, define $\sigma_1^2=\beta_1 D_1-\beta_1^2 D_0/\beta_0$, $\sigma_2^2=\beta_2 D_2-\beta_2^2 D_0/\beta_0$, and $\Gamma=-\beta_1\beta_2 D_0/\beta_0$. Finally, let notation $A\sim\mathcal{N}(0,\sigma^2)$ specify that A is a Gaussian random variable with mean 0 and variance σ^2 and use notation $A\perp\!\!\!\perp B$ to specify that random variables A and B are independent.

Theorem 2: For any $i\in\{1,2\}$ and any $(D_0,D_1,D_2)\in\mathcal{D}_1\bigcup\mathcal{D}_2$

$$\begin{aligned}L_i &\le \frac{1}{2}\log(2-d_i)\le\frac{1}{2}\\ L_0 &\le \frac{1}{2}\log(2-d_0)+\min\Big\{L_{\min},\\ &\qquad R(\max\{D_1,D_2\})+\frac{1}{2}\log(2-\max\{d_1,d_2\})\Big\}\\ &\le \min\Big\{L_{\min}+\frac{1}{2}, R(\max\{D_1,D_2\})+1\Big\}.\end{aligned}$$

are simultaneously achievable.

Proof: Make the following assumptions.

1) Use U_0, U_1, and U_2 to denote the random variables that achieve $R(D_0)$, $R(D_1)$, and $R(D_2)$, respectively, i.e., $Ed(X,U_i)\le D_i$, and $I(X;U_i)=R(D_i)$ for $i=0,1,2$.
2) Let $N_i\sim\mathcal{N}(0,\sigma_i^2)$ for all $i\in\{0,1,2\}$, where $\sigma_0^2=\beta_0 D_0$ and σ_1^2 and σ_2^2 are defined as given earlier.
3) Set $(N_0,N_1,N_2)\perp\!\!\!\perp(X,U_0,U_1,U_2))$, $N_0\perp\!\!\!\perp(N_1,N_2)$, and $E(N_1N_2)=\Gamma=-\beta_1\beta_2 D_0/\beta_0$.

First, verify that these assumptions are reasonable by showing that $\sigma_1^2,\sigma_2^2\ge 0$ and $[E(N_1N_2)]^2\le\sigma_1^2\sigma_2^2$, i.e., Cauchy's inequality holds for N_1 and N_2. The latter is true from Lemma 6 in the Appendix. To prove the former, note that

$$\begin{aligned}\sigma_1^2 &= \beta_1 D_1-\beta_1^2 D_0/\beta_0\\ &=\frac{\beta_1}{\beta_0}(\beta_0 D_1-\beta_1 D_0)\\ &=\frac{\beta_1}{\beta_0}\left(\left(1-\frac{D_0}{\sigma^2}\right)D_1-\left(1-\frac{D_1}{\sigma^2}\right)D_0\right)\\ &=\frac{\beta_1}{\beta_0}(D_1-D_0)\ge 0.\end{aligned}$$

Similarly, $\sigma_2^2\ge 0$.

Second, define

$$\begin{aligned}Y_0&=\beta_0 X+N_0\\ Y_1&=\frac{\beta_1}{\beta_0}Y_0+N_1=\beta_1 X+\frac{\beta_1}{\beta_0}N_0+N_1\\ Y_2&=\frac{\beta_2}{\beta_0}Y_0+N_2=\beta_2 X+\frac{\beta_2}{\beta_0}N_0+N_2.\end{aligned}$$

It can be shown that $Ed(X,Y_i)\le D_i$, $i=0,1,2$. Therefore, according to Theorem 1, the rate pair (R_1,R_2) is achievable if $R_1\ge I(X;Y_1)$, $R_2\ge I(X;Y_2)$, and $R_1+R_2\ge I(X;Y_0,Y_1,Y_2)+I(Y_1;Y_2)$. Following the proof of [1, Theorem 6], the rate loss at decoder 1 is

$$\begin{aligned}L_1 &= I(X;Y_1)-I(X;U_1)\\ &= I(X;Y_1|U_1)-I(X;U_1|Y_1) && (16)\\ &\le I(X;Y_1|U_1)\\ &= I\left(X;\beta_1 X+\frac{\beta_1}{\beta_0}N_0+N_1\middle| U_1\right)\\ &= I\left(X-U_1;\beta_1(X-U_1)+\frac{\beta_1}{\beta_0}N_0+N_1\middle| U_1\right) && (17)\\ &\le I\left(X-U_1;\beta_1(X-U_1)+\frac{\beta_1}{\beta_0}N_0+N_1\right) && (18)\\ &= h\left(\beta_1(X-U_1)+\frac{\beta_1}{\beta_0}N_0+N_1\right)-h\left(\frac{\beta_1}{\beta_0}N_0+N_1\right)\\ &\le \frac{1}{2}\log\left(2\pi e(\beta_1^2 D_1+\beta_1 D_1)\right)-\frac{1}{2}\log(2\pi e\beta_1 D_1) && (19)\\ &= \frac{1}{2}\log(2-d_1)\\ &\le\frac{1}{2}\end{aligned}$$

where (16) follows by applying the chain rule twice to $I(X;Y_1,U_1)$ to obtain

$$\begin{aligned}I(X;Y_1,U_1)&=I(X;Y_1)+I(X;U_1|Y_1)\\ &=I(X;U_1)+I(X;Y_1|U_1).\end{aligned}$$

Equation (17) follows since $h(A|B)=h(A-B|B)$ and $h(A|B,C)=h(A-\alpha B|B,C)$ for any constant α, (18) follows since $(N_0,N_1)\perp\!\!\!\perp(X,U_1)$ implies that

$$U_1\to X-U_1\to\beta_1(X-U_1)+\beta_1 N_0/\beta_0+N_1$$

forms a Markov chain, and (19) follows since the Gaussian distribution maximizes the differential entropy under the constraint that $Ed(X,U_1)\le D_1$.

Similarly, the rate loss at decoder 2 is bounded as

$$L_2 = I(X;Y_2) - R(D_2) \le \frac{1}{2}\log(2-d_2) \le \frac{1}{2}. \tag{20}$$

Finally, the total rate at the joint decoder of the optimal code is

$$\begin{aligned} R_1 + R_2 &\le I(X;Y_0,Y_1,Y_2) + I(Y_1;Y_2) \\ &= I(X;Y_0) + I(Y_1;Y_2) \end{aligned} \tag{21}$$

because $X \to Y_0 \to (Y_1,Y_2)$ form a Markov chain. Thus, the rate loss at the joint decoder is

$$\begin{aligned} L_0 &= R_1 + R_2 - I(X;U_0) \\ &\le [I(X;Y_0) - I(X;U_0)] + I(Y_1;Y_2) \\ &\le \frac{1}{2}\log(2-d_0) + I(Y_1;Y_2) \end{aligned} \tag{22}$$

where the derivation of (22) parallels that of (19).

Since $\beta_1 N_0/\beta_0 + N_1$ and $\beta_2 N_0/\beta_0 + N_2$ are Gaussian random variables and

$$\begin{aligned} &E\left\{\left(\frac{\beta_1}{\beta_0}N_0 + N_1\right)\left(\frac{\beta_2}{\beta_0}N_0 + N_2\right)\right\} \\ &\quad = \frac{\beta_1\beta_2}{\beta_0^2}\beta_0 D_0 + E(N_1N_2) \\ &\quad = 0 \end{aligned}$$

then these two random variables are independent. From the definitions of Y_1 and Y_2, $Y_1 \to X \to Y_2$ forms a Markov chain. Therefore,

$$\begin{aligned} I(Y_1;Y_2) &\le I(X;Y_2) \qquad (23) \\ &= I\left(X; \beta_2 X + \frac{\beta_2}{\beta_0}N_0 + N_2\right) \\ &= h\left(\beta_2 X + \frac{\beta_2}{\beta_0}N_0 + N_2\right) - h\left(\frac{\beta_2}{\beta_0}N_0 + N_2\right) \\ &\le \frac{1}{2}\log\frac{\beta_2^2\sigma^2 + \beta_2 D_2}{\beta_2 D_2} \\ &= \frac{1}{2}\log\frac{\sigma^2}{D_2}. \qquad (24) \end{aligned}$$

On the other hand, by the symmetry of (24)

$$I(Y_1;Y_2) \le \frac{1}{2}\log(\sigma^2/D_1).$$

Thus, in summary

$$I(Y_1;Y_2) \le L_{\min}. \tag{25}$$

As a consequence

$$L_0 \le \frac{1}{2}\log(2-d_0) + L_{\min}.$$

We can also compare $I(Y_1;Y_2)$ to $R(D_1)$ or $R(D_2)$, i.e., by combining (22), (23), and (20), we can show that

$$L_0 \le \frac{1}{2}\log(2-d_0) + R(D_2) + \frac{1}{2}\log(2-d_2)$$

which is a decreasing function of D_2 for fixed D_0 and σ^2.

By symmetry

$$\begin{aligned} L_0 \le \frac{1}{2}\log(2-d_0) &+ R(\max\{D_1,D_2\}) \\ &+ \frac{1}{2}\log(2-\max\{d_1,d_2\}). \qquad \square \end{aligned}$$

Corollary 1: For any $i \in \{1,2\}$ and any $(D_0,D_1,D_2) \in \mathcal{D}_1$

$$L_i \le \frac{1}{2}\log(2-d_i) \le \frac{1}{2}$$

and

$$L_0 \le \frac{1}{2}\log[2(2-d_0)/(1+d_0)] \le 1$$

are simultaneously achievable.

Proof: From Theorem 2, for any $(D_0,D_1,D_2) \in \mathcal{D}_1$

$$\begin{cases} L_1 \le \frac{1}{2}\log(2-d_1) \\ L_2 \le \frac{1}{2}\log(2-d_2) \\ L_0 \le \frac{1}{2}\log(2-d_0) + L_{\min} \end{cases} \tag{26}$$

is achievable.

From the definition of $\mathcal{D}_1$, $D_1 + D_2 \ge D_0 + \sigma^2$; thus,

$$\max\{D_1,D_2\} \ge \frac{D_1+D_2}{2} \ge \frac{D_0+\sigma^2}{2}.$$

Then from (5)

$$L_{\min} = \frac{1}{2}\log\frac{\sigma^2}{\max\{D_1,D_2\}} \le \frac{1}{2}\log\frac{2}{1+d_0}$$

which proves this corollary. $\square$

Corollary 2: For any $i \in \{1,2\}$ and any $(D_0,D_1,D_2) \in \mathcal{D}_2$, $L_i \le \frac{1}{2}$ and $L_0 \le L_{G0} + 1.5$ are simultaneously achievable.

Proof: This corollary is an immediate result of Lemma 4 in the Appendix and Theorem 2. $\square$

Corollary 3: For any $i \in \{1,2\}$ and any $(D_0,D_1,D_2) \in \mathcal{D}_2$ with $d_1 \ge \frac{1}{2}$ or $d_2 \ge \frac{1}{2}$, $L_i \le \frac{1}{2}$ and $L_0 \le 1$ are simultaneously achievable.

Proof: This corollary is an immediate result of Theorem 2. $\square$

Corollary 4: For any $i \in \{1,2\}$ and any $(D_0,D_1,D_2) \in \mathcal{D}_2$

$$L_i \le \frac{1}{2} \quad \text{and} \quad L_{12} \le \frac{1}{2}\log\left[\min\left\{\frac{d_1}{d_0},\frac{d_2}{d_0}\right\}\right] + 1$$

are simultaneously achievable.

Proof: From Theorem 2, there exists an achievable rate pair (R_1,R_2) with $L_1 \le \frac{1}{2}$ and $L_0 \le \min\{R(D_1),R(D_2)\}+1$. From (4)

$$\begin{aligned} L_{12} &\le 1 + \min\{R(D_0)-R(D_1), R(D_0)-R(D_2)\} \\ &\le 1 + \frac{1}{2}\log\frac{\min\{D_1,D_2\}}{D_0} \end{aligned}$$

where the last inequality comes from Lemma 7 in the Appendix. $\square$

Define

$$\Gamma' = \frac{\beta_1\beta_2 D_0 - \sqrt{\beta_1\beta_2(D_1-D_0)(D_2-D_0)}}{\beta_0}$$

$$\alpha_1 = \frac{\beta_0(\beta_1 D_2 - \Gamma')}{\beta_1(\beta_2 D_1 + \beta_1 D_2 - 2\Gamma')}$$
$$\alpha_2 = \frac{\beta_0(\beta_2 D_1 - \Gamma')}{\beta_2(\beta_2 D_1 + \beta_1 D_2 - 2\Gamma')}$$

which are useful in proving Theorem 3.

Theorem 3: For any $i \in \{1,2\}$ and any $(D_0, D_1, D_2) \in \mathcal{D}_2$, $L_i \le \frac{1}{2}$ and $L_{12} \le L_{G12} + 1$ are simultaneously achievable.

Proof: Define

$$Y_1 = \beta_1 X + N_1'$$
$$Y_2 = \beta_2 X + N_2'$$

where $N_i' \sim \mathcal{N}(0, \beta_i D_i)$ and $N_i' \perp\!\!\!\perp (X, U_0, U_1, U_2)$ for $i = 1, 2$, and $E(N_1' N_2') = \Gamma'$. From Lemmas 8 and 9, these definitions are reasonable and $\alpha_1, \alpha_2 > 0$.

Next, define $Y_0 = \alpha_1 Y_1 + \alpha_2 Y_2$. Since $\alpha_1\beta_1 + \alpha_2\beta_2 = \beta_0$, $Y_0 = \beta_0 X + N_0'$, where $N_0' = \alpha_1 N_1' + \alpha_2 N_2'$. Notice that $N_0' \perp\!\!\!\perp (X, U_0, U_1, U_2)$, and from Lemma 9, $N_0' \sim \mathcal{N}(0, \beta_0 D_0)$. Therefore, it can be shown that $Ed(X, Y_i) \le D_i$, $i = 0, 1, 2$.

Again, the rate pair (R_1, R_2) is achievable if $R_1 \ge I(X;Y_1)$, $R_2 \ge I(X;Y_2)$, and $R_1 + R_2 \ge I(X;Y_0,Y_1,Y_2) + I(Y_1;Y_2)$, and

$$L_1 = I(X;Y_1) - I(X;U_1) \le \frac{1}{2}\log(2 - d_1) \le \frac{1}{2}$$
$$L_2 = I(X;Y_2) - I(X;U_2) \le \frac{1}{2}\log(2 - d_2) \le \frac{1}{2}.$$

In addition, since Y_0 is a function of Y_1 and Y_2, we know $I(X;Y_0,Y_1,Y_2) = I(X;Y_1,Y_2)$, and

$$\begin{aligned}\left(\frac{\Gamma'}{\sqrt{\beta_1 D_1}\sqrt{\beta_2 D_2}}\right)^2 &= \left(\frac{\beta_1\beta_2 D_0 - \sqrt{\beta_1\beta_2(D_1 - D_0)(D_2 - D_0)}}{\beta_0\sqrt{\beta_1\beta_2 D_1 D_2}}\right)^2 \\ &= \left(\frac{d_0\sqrt{\beta_1\beta_2} - \sqrt{(d_1 - d_0)(d_2 - d_0)}}{\beta_0\sqrt{d_1 d_2}}\right)^2 \end{aligned} \tag{27}$$

thus,

$$\begin{aligned} L_{12} &\le I(X;Y_1,Y_2) + I(Y_1;Y_2) - R(D_1) - R(D_2) \\ &= I(X;Y_1) + I(X;Y_2|Y_1) + I(Y_1;Y_2) - R(D_1) \\ &\quad - R(D_2) \\ &= I(X;Y_1) + I(X,Y_1;Y_2) - R(D_1) - R(D_2) \\ &= [I(X;Y_1) - R(D_1)] + [I(X;Y_2) - R(D_2)] \\ &\quad + I(Y_1;Y_2|X) \\ &= [I(X;Y_1) - R(D_1)] + [I(X;Y_2) - R(D_2)] \\ &\quad + I(N_1';N_2') \\ &\le \frac{1}{2}\log(2 - d_1) + \frac{1}{2}\log(2 - d_2) + I(N_1';N_2') \\ &< 1 + I(N_1';N_2') \\ &= 1 - \frac{1}{2}\log\left[1 - \left(\frac{\Gamma'}{\sqrt{\beta_1 D_1}\sqrt{\beta_2 D_2}}\right)^2\right] \\ &= 1 + L_{G12} \end{aligned}$$

where the last equation follows from (27) and the definition of L_{G12}. □

We introduce three new quantities

$$\begin{aligned} \Delta &= \beta_0\beta_1\beta_2 D_0 D_1 D_2\left(\frac{1}{D_1} + \frac{1}{D_2} - \frac{1}{D_0} - \frac{1}{\sigma^2}\right) \\ &= \beta_0\beta_1\beta_2(D_0 D_2 + D_0 D_1 - D_1 D_2 - d_0 D_1 D_2) \\ \alpha_1' &= \frac{\beta_0\beta_1 D_2 + \sqrt{\Delta}}{\beta_1(\beta_2 D_1 + \beta_1 D_2)} \\ \alpha_2' &= \frac{\beta_0\beta_2 D_1 - \sqrt{\Delta}}{\beta_2(\beta_2 D_1 + \beta_1 D_2)} \end{aligned}$$

where $\Delta \ge 0$ in $\mathcal{D}_3$.

Theorem 4: For any $(D_0, D_1, D_2) \in \mathcal{D}_3$,

$$L_1 \le \frac{1}{2}\log(2 - d_1) \le \frac{1}{2} \quad \text{and} \quad L_2 \le \frac{1}{2}\log(2 - d_2) \le \frac{1}{2}$$

are simultaneously achievable.

Proof: Let

$$Y_1 = \beta_1 X + N_1'$$
$$Y_2 = \beta_2 X + N_2'$$
$$Y_0 = \alpha_1' Y_1 + \alpha_2' Y_2$$

where $N_i' \sim \mathcal{N}(0, \beta_i D_i)$ and $N_i' \perp\!\!\!\perp (X, U_0, U_1, U_2)$ for $i = 1, 2$, and $N_1' \perp\!\!\!\perp N_2'$ (note that N_1' and N_2' were not independent in the previous proof). It can be verified that

$$Y_0 = \alpha_1' Y_1 + \alpha_2' Y_2 = \beta_0 X + N_0$$

where $N_0 = \alpha_1' N_1' + \alpha_2' N_2' \sim \mathcal{N}(0, \beta_0 D_0)$ (from Lemma 10) and $N_0 \perp\!\!\!\perp (X, U_0, U_1, U_2)$. Thus, it can be shown that Y_1, Y_2, and Y_0 satisfy the distortion requirements. From Theorem 1, the rate pair (R_1, R_2) is achievable if $R_1 \ge I(X;Y_1)$, $R_2 \ge I(X;Y_2)$, and

$$\begin{aligned} R_1 + R_2 &\ge I(X;Y_0,Y_1,Y_2) + I(Y_1;Y_2) \\ &= I(X;Y_1,Y_2) + I(Y_1;Y_2) \\ &= I(X;Y_1) + I(X;Y_2) + I(Y_1;Y_2|X). \end{aligned}$$

Again, $L_1 \le \frac{1}{2}\log(2-d_1) \le \frac{1}{2}$ and $L_2 \le \frac{1}{2}\log(2-d_2) \le \frac{1}{2}$. Bounding L_{12} gives us

$$\begin{aligned} L_{12} &= R_1 + R_2 - R(D_1) - R(D_2) \\ &\le I(X;Y_1) + I(X;Y_2) + I(Y_1;Y_2|X) - R(D_1) - R(D_2) \\ &\le \frac{1}{2}\log(2-d_1) + \frac{1}{2}\log(2-d_2) + I(N_1';N_2') \\ &= \frac{1}{2}\log(2-d_1) + \frac{1}{2}\log(2-d_2) \qquad (28) \\ &\le 1 \end{aligned}$$

where (28) comes from the fact that N_1' and N_2' are independent; therefore, $I(N_1';N_2') = 0$.

Here, the bound for L_{12} is redundant from the bounds on L_1 and L_2. □

Theorem 5: For any $i \in \{1,2\}$ and any $(D_0, D_1, D_2) \in \mathcal{D}_2$ with $\max\{d_1, d_2\} < 1/2$, there exists an $(R_1, R_2) \in$

$\mathcal{R}(D_0, D_1, D_2)$ with $L_i \le \frac{1}{2}\log(2-d_i) \le \frac{1}{2}$, and the distance between the upper bound and the lower bound for L_0 is no more than

$$\min\left\{\log(2\pi e\sigma^2)-2h(X)+1, \frac{1}{2}\log(2\pi e\sigma^2)-h(X)+1.5\right\}.$$

Proof: Here, we want to compare the lower bound and upper bound for L_0 and bound from above the distance between them. Let K_0 denote this distance. Without loss of generality, assume that $D_1 \le D_2$. We are only interested in region $\mathcal{D}_2$ with $D_2 < \sigma^2/2$, i.e., $D_2 < \sigma^2/2$ and

$$D_1 + D_2 - \sigma^2 < 0 < D_0 < (1/D_1 + 1/D_2 - 1/\sigma^2)^{-1}.$$

From Lemma 4, $LG(d_0, d_1, d_2) > 0$ in region $\mathcal{D}_2$. From the proof of Theorem 2 (i.e., (19)–(21), and (24)), there exists an $(R_1, R_2) \in \mathcal{R}(D_0, D_1, D_2)$ with

$$\begin{aligned} L_1 &\le \frac{1}{2}\log(2-d_1) \le \frac{1}{2} \\ L_2 &\le \frac{1}{2}\log(2-d_2) \le \frac{1}{2} \\ L_0 &\le I(X;Y_0) + \frac{1}{2}\log\frac{\sigma^2}{D_2} - R(D_0) \end{aligned} \tag{29}$$

where Y_0 is defined as $Y_0 = \beta_0 X + N_0$, $N_0 \sim \mathcal{N}(0, \beta_0 D_0)$, and $N_0 \perp\!\!\!\perp X$. Thus,

$$I(X;Y_0) \le \frac{1}{2}\log\frac{\sigma^2}{D_0}. \tag{30}$$

First consider the case where $D_2 < P_X/2$. In this configuration, $D_1 + D_2 - P_X < 0 < D_0$. Further

$$\frac{1}{D_1} + \frac{1}{D_2} - \frac{1}{D_0} < \frac{1}{\sigma^2} \le \frac{1}{P_X}.$$

Together (29), (30), and (9) give the following bound on the distance K_0 between the upper and lower bounds for L_0:

$$\begin{aligned} K_0 &\le \frac{1}{2}\log\frac{\sigma^2}{D_0} + \frac{1}{2}\log\frac{\sigma^2}{D_2} - \frac{1}{2}\log\frac{P_X}{D_0} \\ &\quad - LG\left(\frac{D_0}{P_X}, \frac{D_1}{P_X}, \frac{D_2}{P_X}\right) \\ &= \frac{1}{2}\log\frac{\sigma^2}{P_X} + \frac{1}{2}\log\frac{\sigma^2}{D_2} - LG\left(\frac{D_0}{P_X}, \frac{D_1}{P_X}, \frac{D_2}{P_X}\right). \end{aligned}$$

From Lemma 11

$$LG\left(\frac{D_0}{P_X}, \frac{D_1}{P_X}, \frac{D_2}{P_X}\right) > \frac{1}{2}\log\frac{P_X^2}{4D_2(P_X - D_2)}.$$

As a result

$$\begin{aligned} K_0 &\le \frac{1}{2}\log\frac{\sigma^2}{P_X} + \frac{1}{2}\log\frac{4\sigma^2(P_X - D_2)}{P_X^2} \\ &\le \frac{1}{2}\log\frac{\sigma^2}{P_X} + \frac{1}{2}\log\frac{4\sigma^2}{P_X} \\ &= \log\frac{\sigma^2}{P_X} + 1. \end{aligned}$$

If $D_2 \ge P_X/2$, then use 0 as the lower bound for L_0 and use the SLB

$$R(D_0) \ge \frac{1}{2}\log\frac{P_X}{D_0}.$$

From (29) and (30)

$$\begin{aligned} K_0 &\le L_0 \\ &\le \frac{1}{2}\log\frac{\sigma^2}{D_0} + \frac{1}{2}\log\frac{\sigma^2}{D_2} - R(D_0) \\ &\le \frac{1}{2}\log\frac{\sigma^2}{D_0} + \frac{1}{2}\log\frac{\sigma^2}{D_2} - \frac{1}{2}\log\frac{P_X}{D_0} \\ &= \frac{1}{2}\log\frac{\sigma^2}{P_X} + \frac{1}{2}\log\frac{\sigma^2}{D_2} \\ &\le \frac{1}{2}\log\frac{\sigma^2}{P_X} + \frac{1}{2}\log\frac{2\sigma^2}{P_X} \\ &= \log\frac{\sigma^2}{P_X} + \frac{1}{2}. \end{aligned}$$

Thus, in both cases

$$K_0 \le \log\left(\frac{\sigma^2}{P_X}\right) + 1 = 2\left[\frac{1}{2}\log(2\pi e\sigma^2) - h(X)\right] + 1.$$

In fact, there is a better converse for the achievable region. It has been shown (e.g., [11] and [10]) that the inequality

$$R_1 + R_2 \ge SR(e_0, e_1, e_2, R_1 + R_2)$$

is equivalent to the inequality

$$R_1 + R_2 \ge -\frac{1}{2}\log e_0 + LG(e_0, e_1, e_2)$$

if $0 < e_0 < e_1, e_2 \le 1$ and

$$e_0 + e_1 - 1 < e_0 < (1/e_1 + 1/e_2 - 1)^{-1}$$

where $SR(e_0, e_1, e_2, R_1 + R_2)$ is a function defined as:

$$\begin{aligned} SR(e_0, e_1, e_2, R_1 + R_2) &= \frac{1}{2}\log\frac{1}{e_0} \\ &\quad -\frac{1}{2}\log\left(1-\left(\sqrt{(1-e_1)(1-e_2)}-\sqrt{e_1e_2-2^{-2(R_1+R_2)}}\right)^2\right). \end{aligned}$$

By [10]

$$\begin{aligned} \{(R_1, R_2) &: R_1 \ge R(D_1), R_2 \ge R(D_2), \\ &R_1 + R_2 \ge SR(d_0^*, D_1/P_X, D_2/P_X, R_1 + R_2)\} \end{aligned}$$

is also a converse for the achievable region if $0 < D_0 < D_1$, $D_2 \le P_X$, and $D_1 + D_2 - P_X < P_X d_0^* < (1/D_1 + 1/D_2 - 1/P_X)^{-1}$, where d_0^* is the "effective distortion" at the joint decoder defined as

$$d_0^* = 2^{-2R(D_0)}.$$

Note that since $R(D_0) = \frac{1}{2}\log(1/d_0^*)$, this converse is equivalent to

$$\begin{aligned} \{(R_1, R_2) &: R_1 \ge R(D_1), R_2 \ge R(D_2), \\ &R_1 + R_2 \ge R(D_0) + LG(d_0^*, D_1/P_X, D_2/P_X)\}. \end{aligned} \tag{31}$$

We will bound the distance between the achievability result and this new converse. Again, assume $D_1 \le D_2$ and first

TABLE I
SUMMARY OF THE RATE LOSS BOUNDS

Region	Gaussian Source	General Source
$\mathcal{D}_1 = \left\{(D_0, D_1, D_2) : 0 < D_0 \le D_1 + D_2 - \sigma^2\right\}$	$L_1 = L_2 = 0$ $L_0 = 0$	$L_1, L_2 \le \frac{1}{2}$ $L_0 \le 1$
$\mathcal{D}_2 = \left\{(D_0, D_1, D_2) : D_1 + D_2 - \sigma^2 < D_0 < \frac{1}{\frac{1}{D_1}+\frac{1}{D_2}-\frac{1}{\sigma^2}}\right\}$	$L_1 = L_2 = 0$ $L_0 = L_{G0}$ $L_{12} = L_{G12}$	$L_1, L_2 \le \frac{1}{2}$ $L_0 \le L_{G0} + 1.5$ $L_{12} \le L_{G12} + 1$
$\mathcal{D}_3 = \left\{(D_0, D_1, D_2) : D_0 \ge \frac{1}{\frac{1}{D_1}+\frac{1}{D_2}-\frac{1}{\sigma^2}}\right\}$	$L_1 = L_2 = 0$ $L_{12} = 0$	$L_1, L_2 \le \frac{1}{2}$ $L_{12} \le 1$

verify that (31) is a converse in region $\mathcal{D}_2$. In region $\mathcal{D}_2$, since $R(D_0) \ge h(X) - \frac{1}{2}\log(2\pi e D_0)$ and $P_X \le \sigma^2$

$$\begin{aligned} P_X 2^{-2R(D_0)} &\le D_0 < \frac{1}{1/D_1 + 1/D_2 - 1/\sigma^2} \\ &\le \frac{1}{1/D_1 + 1/D_2 - 1/P_X}. \end{aligned}$$

Thus, if $D_2 < P_X/2$, $D_1 + D_2 - P_X < 0 < D_0$, and from Lemma 11

$$L_0 \ge LG(d_0^*, D_1/P_X, D_2/P_X) \ge \frac{1}{2}\log\frac{P_X^2}{4D_2(P_X - D_2)}.$$

From Theorem 2, there exists (R_1, R_2) with $L_1 \le \frac{1}{2}$, $L_2 \le \frac{1}{2}$, and

$$\begin{aligned} K_0 &\le \frac{1}{2} + \frac{1}{2}\log\frac{\sigma^2}{D_2} - LG(d_0^*, D_1/P_X, D_2/P_X) \\ &\le \frac{1}{2} + \frac{1}{2}\log\frac{4\sigma^2(P_X - D_2)}{P_X^2} \\ &\le \frac{1}{2} + \frac{1}{2}\log\frac{4\sigma^2}{P_X}. \end{aligned}$$

Thus, $K_0 \le \frac{1}{2}\log(2\pi e\sigma^2) - h(X) + 1.5$. Similarly, if $D_2 \ge P_X/2$

$$\begin{aligned} K_0 \le L_0 &\le \frac{1}{2} + \frac{1}{2}\log(\sigma^2/D_2) \\ &\le \frac{1}{2}\log(2\pi e\sigma^2) - h(X) + 1. \end{aligned}$$

In summary, in this region, K_0 is always bounded by $\frac{1}{2}\log(2\pi e\sigma^2) - h(X) + 1.5$, which is better than the previous bound if $h(X)$ is much less than $\frac{1}{2}\log(2\pi e\sigma^2)$. □

Theorem 6: For any $i \in \{1, 2\}$ and any $(D_0, D_1, D_2) \in \mathcal{D}_2$ with $\max\{d_1, d_2\} < \frac{1}{2}$, if the SLB is tight at distortions D_1 and D_2, i.e.,

$$R(D_j) = h(X) - \frac{1}{2}\log(2\pi e D_j), \qquad \text{for any } j \in \{1, 2\}$$

then there exists an $(R_1, R_2) \in \mathcal{R}(D_0, D_1, D_2)$ with $L_i \le \frac{1}{2}\log(2 - d_i) \le \frac{1}{2}$, and the distance between the upper bound and the lower bound for L_0 is less than or equal to 2.

Proof: Since the SLB is tight at distortion D_1

$$D_1 = \frac{1}{2\pi e}2^{2[h(X) - R(D_1)]} = P_X d_1^*$$

where $d_1^* = 2^{-2R(D_1)}$ is the "effective distortion" at decoder 1. Similarly, $D_2 = P_X d_2^*$, where $d_2^* = 2^{-2R(D_2)}$. Then, from (31)

$$\begin{aligned} R_1 + R_2 &\ge R(D_0) + LG(d_0^*, D_1/P_X, D_2/P_X) \\ &= R(D_0) + LG(d_0^*, d_1^*, d_2^*). \end{aligned}$$

Again, assume that $D_1 \le D_2$ without loss of generality. Because $R(D)$ is a nonincreasing function of D, so $0 < d_0^* < d_1^* \le d_2^*$. Thus, if $d_2^* < \frac{1}{2}$, then from Lemma 11

$$\begin{aligned} LG(d_0^*, d_1^*, d_2^*) &\ge \frac{1}{2}\log\frac{1}{d_2^*} + \frac{1}{2}\log\frac{1}{4(1 - d_2^*)} \\ &= R(D_2) + \frac{1}{2}\log\frac{1}{4(1 - d_2^*)} \\ &\ge R(D_2) - 1. \end{aligned}$$

From Theorem 2, $R_1 + R_2 = R(D_0) + R(D_2) + 1$ is achievable; thus, the difference between the upper bound and this new lower bound for L_0 is

$$K_0 \le [R(D_0) + R(D_2) + 1] - [R(D_0) + R(D_2) - 1] = 2.$$

If $d_2^* \ge \frac{1}{2}$, then $R(D_2) \le \frac{1}{2}$ according to the definition of d_2^*. Thus, $R_1 + R_2 = R(D_0) + R(D_2) + 1 \le R(D_0) + 1.5$ is achievable, which implies that

$$K_0 \le L_0 \le R_1 + R_2 - R(D_0) < 1.5.$$

In either case, the distance between the upper and lower bound is no larger than 2. □

V. Conclusion

In this paper, we bound the rate loss of the MDSC for general i.i.d. sources and the mse distortion measure. The results include both constant bounds and bounds that depend on the rate loss of a Gaussian source of the same variance. For convenience, we list our main results in Table I (we also include the achievable region on Gaussian sources in each distortion region for comparison). The rate loss bounds are useful both because they characterize the potential penalties associated with using MDSCs and because they provide new achievability results on the MDSC-achievable regions. These new achievability results can be easily analyzed for any source for which we can find the rate-distortion bound, and that achievability result is shown to be quite tight in some cases.

APPENDIX

Lemma 1: If $0 < D_1, D_2 \leq \sigma^2$, then

$$D_1 + D_2 - \sigma^2 \leq \left(\frac{1}{D_1} + \frac{1}{D_2} - \frac{1}{\sigma^2}\right)^{-1}.$$

Proof: We show the lemma by the following inequalities:

$$\begin{aligned}(D_1 + D_2 - \sigma^2)&\left(\frac{1}{D_1} + \frac{1}{D_2} - \frac{1}{\sigma^2}\right)\\ &= (D_1 + D_2 - \sigma^2)\frac{D_1 + D_2 - D_1 D_2/\sigma^2}{D_1 D_2}\\ &= \frac{(D_1 + D_2)^2 - (D_1 + D_2)(\sigma^2 + D_1 D_2/\sigma^2) + D_1 D_2}{D_1 D_2}\\ &= \frac{-(D_1 + D_2)(\sigma^2 - D_1)(\sigma^2 - D_2)/\sigma^2 + D_1 D_2}{D_1 D_2}\\ &\leq 1\end{aligned}$$

where the last inequality comes from the fact that $0 < D_1, D_2 \leq \sigma^2$. □

Lemma 2: Suppose $R(D)$ is the rate-distortion function of an arbitrary i.i.d. source $0 < D_0 < D_1, D_2 \leq \sigma^2$, and $D_1 + D_2 - D_0 \geq \sigma^2$. Then

$$R(D_1) + R(D_2) \leq R(D_0).$$

Proof: Since

$$D_1 = \frac{\sigma^2 - D_1}{\sigma^2 - D_0} D_0 + \left(1 - \frac{\sigma^2 - D_1}{\sigma^2 - D_0}\right)\sigma^2$$

then from the convexity of the rate-distortion function

$$\begin{aligned}R(D_1) &\leq \frac{\sigma^2 - D_1}{\sigma^2 - D_0} R(D_0) + \left(1 - \frac{\sigma^2 - D_1}{\sigma^2 - D_0}\right) R(\sigma^2)\\ &= \frac{\sigma^2 - D_1}{\sigma^2 - D_0} R(D_0).\end{aligned}$$

By symmetry

$$R(D_2) \leq \frac{\sigma^2 - D_2}{\sigma^2 - D_0} R(D_0).$$

Thus,

$$\begin{aligned}R(D_1) &+ R(D_2) - R(D_0)\\ &\leq \frac{\sigma^2 - D_1}{\sigma^2 - D_0} R(D_0) + \frac{\sigma^2 - D_2}{\sigma^2 - D_0} R(D_0) - R(D_0)\\ &= \frac{\sigma^2 + D_0 - D_1 - D_2}{\sigma^2 - D_0} R(D_0)\\ &\leq 0\end{aligned}$$

since $D_1 + D_2 - D_0 \geq \sigma^2$. □

Lemma 3 describes the relation between L_{G12} and L_{G0}, which also implies that L_{G12} equals L_{12} for the Gaussian source.

Lemma 3: In $\mathcal{D}_2$

$$L_{G12} = \frac{1}{2}\log(d_1 d_2) - \frac{1}{2}\log(d_0) + L_{G0}.$$

Proof: First, let

$$\begin{aligned}\Theta_1 &= (1 - d_0)^2 d_1 d_2\\ &\quad - \left(d_0\sqrt{(1 - d_1)(1 - d_2)} - \sqrt{(d_1 - d_0)(d_2 - d_0)}\right)^2\\ \Theta_2 &= d_0(1 - d_0)^2\\ &\quad - d_0\left(\sqrt{(1 - d_1)(1 - d_2)} - \sqrt{(d_1 - d_0)(d_2 - d_0)}\right)^2.\end{aligned}$$

We can show that

$$\begin{aligned}\Theta_1 - &\Theta_2\\ &= (1 - d_0)^2 d_1 d_2 - d_0^2(1 - d_1)(1 - d_2) -\\ &\quad (d_1 - d_0)(d_2 - d_0) - d_0(1 - d_0)^2\\ &\quad + d_0(1 - d_1)(1 - d_2) + d_0(d_1 - d_0)(d_2 - d_0)\\ &= (1 - d_0)^2 d_1 d_2 + d_0(1 - d_0)(1 - d_1)(1 - d_2)\\ &\quad - d_0(1 - d_0)^2 - (1 - d_0)(d_1 - d_0)(d_2 - d_0)\\ &= (1 - d_0)\left[(1 - d_0) d_1 d_2 + d_0(1 - d_1)(1 - d_2)\right.\\ &\quad \left.- d_0(1 - d_0) - (d_1 - d_0)(d_2 - d_0)\right]\\ &= 0\end{aligned}$$

where the last equation comes from the fact that

$$\begin{aligned}(1 - d_0) d_1 d_2 &- d_0(1 - d_0) + d_0(1 - d_1)(1 - d_2)\\ &= d_1 d_2 + d_0^2 - d_0 d_1 - d_0 d_2\\ &= (d_1 - d_0)(d_2 - d_0).\end{aligned}$$

Therefore,

$$\begin{aligned}L_{G12} &= \frac{1}{2}\log\frac{(1 - d_0)^2 d_1 d_2}{\Theta_1}\\ &= \frac{1}{2}\log\frac{(1 - d_0)^2 d_1 d_2}{\Theta_2}\\ &= \frac{1}{2}\log\frac{d_1 d_2}{d_0} + L_{G0}\end{aligned}$$

and the lemma is proved. □

Lemma 4: In $\mathcal{D}_2$, $0 < L_{G0} < L_{\min} < L_{G0} + 1$.

Proof: From the definition of $\mathcal{D}_2$

$$\frac{1}{d_0} > \frac{1}{d_1} + \frac{1}{d_2} - 1$$

and $0 < d_0 < d_1, d_2 < 1$ (otherwise, for example, if $d_1 = 1$, then $d_1 + d_2 - 1 = d_2 > d_0$, which contradicts the definition of $\mathcal{D}_2$), which implies

$$\begin{aligned}d_0^2(1 - d_1)(1 - d_2) &- (d_1 d_2 - d_2 d_0 - d_1 d_0 + d_0^2)\\ &= (1 - d_0) d_0 d_1 d_2 \left(\frac{1}{d_1} + \frac{1}{d_2} - \frac{1}{d_0} - 1\right)\\ &< 0\end{aligned} \tag{32}$$

i.e., $d_0\sqrt{(1 - d_1)(1 - d_2)} < \sqrt{(d_1 - d_0)(d_2 - d_0)}$, or equivalently

$$\begin{aligned}\sqrt{(1 - d_1)(1 - d_2)} &- \sqrt{(d_1 - d_0)(d_2 - d_0)}\\ &< (1 - d_0)\sqrt{(1 - d_1)(1 - d_2)}.\end{aligned}$$

On the other hand

$$\sqrt{(1-d_1)(1-d_2)} - \sqrt{(d_1-d_0)(d_2-d_0)} > 0$$

since $1+d_0 > d_1+d_2$. Therefore,

$$\begin{aligned} 0 &< L_{G0} \\ &< \frac{1}{2}\log\frac{(1-d_0)^2}{(1-d_0)^2-(1-d_0)^2(1-d_1)(1-d_2)} \\ &= \frac{1}{2}\log\frac{1}{1-(1-d_1)(1-d_2)} \\ &= \frac{1}{2}\log\frac{1}{d_1+d_2-d_1d_2} \qquad (33)\\ &< \frac{1}{2}\log\frac{1}{\max\{d_1,d_2\}} \end{aligned}$$

where the last inequality is from the fact that $0 < d_1, d_2 < 1$. Thus, $0 < L_{G0} < L_{\min}$.

To prove the other inequality, we first assume $d_2 \le d_1$ without loss of generality. If $1+d_0-2d_1 < 0$, then $d_1 > (1+d_0)/2 > 1/2$, and

$$L_{\min} = \frac{1}{2}\log\frac{1}{\max\{d_1,d_2\}} = \frac{1}{2}\log\frac{1}{d_1} < \frac{1}{2} < L_{G0}+1.$$

Otherwise, $1+d_0-2d_1 \ge 0$. Note that for fixed d_0 and d_1

$$\sqrt{(1-d_1)(1-d_2)} - \sqrt{(d_1-d_0)(d_2-d_0)}$$

is a positive and monotonically decreasing function of d_2 in $\mathcal{D}_2$, therefore,

$$\begin{aligned} \sqrt{(1-d_1)(1-d_2)} - \sqrt{(d_1-d_0)(d_2-d_0)} &\ge \sqrt{(1-d_1)(1-d_1)} - \sqrt{(d_1-d_0)(d_1-d_0)} \\ &= 1+d_0-2d_1 \\ &\ge 0 \end{aligned}$$

and

$$\begin{aligned} L_{G0} &\ge \frac{1}{2}\log\frac{(1-d_0)^2}{(1-d_0)^2-(1+d_0-2d_1)^2} \\ &= \frac{1}{2}\log\frac{(1-d_0)^2}{4(1-d_1)(d_1-d_0)} \end{aligned}$$

thus,

$$\begin{aligned} L_{\min}-L_{G0} &\le \frac{1}{2}\log\frac{1}{d_1} - \frac{1}{2}\log\frac{(1-d_0)^2}{4(1-d_1)(d_1-d_0)} \\ &= 1-\log(1-d_0)+\frac{1}{2}\log\frac{(1-d_1)(d_1-d_0)}{d_1} \\ &= 1-\log(1-d_0)+\frac{1}{2}\log\left[1+d_0-\left(d_1+\frac{d_0}{d_1}\right)\right] \\ &\le 1-\log(1-d_0)+\frac{1}{2}\log(1+d_0-2\sqrt{d_0}) \qquad (34)\\ &= 1-\log(1-d_0)+\log(1-\sqrt{d_0}) \\ &< 1 \end{aligned}$$

where (34) follows from the arithmetic/geometric mean inequality $a+b \ge 2\sqrt{ab}$ for any nonnegative a and b, and the last inequality comes from the fact that $0 < d_0 < 1$ thus $d_0 < \sqrt{d_0}$. □

Corollary 5 follows immediately from this lemma.

Corollary 5: For any $(D_0, D_1, D_2) \in \mathcal{D}_2$, if $d_1 \ge \frac{1}{2}$ or $d_2 \ge \frac{1}{2}$, then $0 < L_{G0} < \frac{1}{2}$.

Lemma 5: Suppose $R(D)$ is the rate-distortion function of an arbitrary i.i.d. source and $0 < D_0 < D_1, D_2 \le \sigma^2$, then

$$R(D_1)+R(D_2) \le R(D_0)+\frac{1}{2}$$

if $D_1+D_2-\sigma^2 < D_0 < (1/D_1+1/D_2-1/\sigma^2)^{-1}$ and $\max\{D_1,D_2\} \ge \sigma^2/2$.

Proof: Without loss of generality, we assume $D_1 \ge \sigma^2/2$. From Shannon's upper bound for the rate-distortion function

$$R(D_1) \le \frac{1}{2}\log\frac{\sigma^2}{D_1} \le \frac{1}{2}.$$

Since $R(D)$ is a nonincreasing function of D and $D_2 \ge D_0$, $R(D_2) \le R(D_0)$. Thus, $R(D_1)+R(D_2) \le \frac{1}{2}+R(D_0)$. □

Lemma 6: For any $(D_0, D_1, D_2) \in \mathcal{D}_1 \bigcup \mathcal{D}_2$, $\Gamma^2 \le \sigma_1^2\sigma_2^2$.

Proof: In $\mathcal{D}_1$ and $\mathcal{D}_2$

$$\begin{aligned} &\beta_0 D_2 D_1 - \beta_1 D_2 D_0 - \beta_2 D_1 D_0 \\ &\quad = D_2D_1 - \frac{D_2D_1D_0}{\sigma^2} - D_2D_0 + \frac{D_2D_1D_0}{\sigma^2} - D_1D_0 + \frac{D_2D_1D_0}{\sigma^2} \\ &\quad = D_2D_1 - D_2D_0 - D_1D_0 + \frac{D_2D_1D_0}{\sigma^2} \\ &\quad = D_2D_1D_0\left(\frac{1}{D_0}+\frac{1}{\sigma^2}-\frac{1}{D_1}-\frac{1}{D_2}\right) \\ &\quad > 0. \end{aligned}$$

As a consequence

$$\begin{aligned} \Gamma^2 &= \frac{\beta_1^2\beta_2^2D_0^2}{\beta_0^2} \\ &\le \frac{\beta_1^2\beta_2^2D_0^2}{\beta_0^2} + \frac{\beta_1\beta_2}{\beta_0}(\beta_0D_2D_1-\beta_1D_2D_0-\beta_2D_1D_0) \\ &= \left(\beta_1D_1-\frac{\beta_1^2D_0}{\beta_0}\right)\left(\beta_2D_2-\frac{\beta_2^2D_0}{\beta_0}\right) \\ &= \sigma_1^2\sigma_2^2. \end{aligned}$$

This proves the lemma. □

Lemma 7: [2, Lemma 1] Suppose $R(D)$ is the rate-distortion function of an arbitrary i.i.d. source $\{X_i\}_{i=1}^{\infty}$ with an mse distortion measure and $0 < D_2 < D_1$, then

$$R(D_2)-R(D_1) \le \frac{1}{2}\log\frac{D_1}{D_2}.$$

Proof: See [2] for details. □

Lemmas 8 and 9 describe the properties of α_1, α_2, and Γ'.

Lemma 8: In $\mathcal{D}_2$

$$-\sqrt{\beta_1\beta_2D_1D_2} < \Gamma' < 0.$$

Proof: From (32), we can see that

$$d_0^2\beta_1\beta_2 < (d_1-d_0)(d_2-d_0).$$

Therefore, $\Gamma' < 0$. On the other hand, since $0 < d_0 < d_1, d_2 < 1$ in $\mathcal{D}_2$

$$\begin{aligned} &2\sqrt{\beta_1\beta_2d_1d_2} > 0 > d_1(d_2-1)+d_2(d_1-1) \\ &\quad = 2d_1d_2-d_1-d_2 \\ &\Rightarrow 2d_0(1-d_0)\sqrt{\beta_1\beta_2d_1d_2} \\ &\quad > 2d_1d_2d_0-2d_1d_2d_0^2-d_1d_0+d_1d_0^2-d_2d_0+d_2d_0^2 \end{aligned}$$

$$\Rightarrow (1-d_0)^2 d_1 d_2 + (1-d_1)(1-d_2)d_0^2 + 2d_0(1-d_0)\sqrt{\beta_1\beta_2 d_1 d_2} > (d_1-d_0)(d_2-d_0)$$
$$\Rightarrow (1-d_0)^2 d_1 d_2 + 2d_0(1-d_0)\sqrt{\beta_1\beta_2 d_1 d_2} + \beta_1\beta_2 d_0^2 > (d_1-d_0)(d_2-d_0) > 0$$
$$\Rightarrow (1-d_0)\sqrt{d_1 d_2} + d_0\sqrt{\beta_1\beta_2} > \sqrt{(d_1-d_0)(d_2-d_0)}.$$

As a consequence

$$\begin{aligned}\Gamma' &= \frac{\beta_1\beta_2 D_0 - \sqrt{\beta_1\beta_2(D_1-D_0)(D_2-D_0)}}{\beta_0} \\ &> \frac{\sqrt{\beta_1\beta_2}}{\beta_0}[-(1-d_0)\sqrt{D_1 D_2}] \\ &= -\sqrt{\beta_1\beta_2 D_1 D_2}\end{aligned}$$

which concludes the proof. □

Lemma 9: In $\mathcal{D}_2$, $\alpha_1 > 0$, $\alpha_2 > 0$, and

$$\alpha_1^2\beta_1 D_1 + \alpha_2^2\beta_2 D_2 + 2\alpha_1\alpha_2\Gamma' = \beta_0 D_0.$$

Proof: We define $D_{10} = D_1 - D_0 > 0$, $D_{20} = D_2 - D_0 > 0$, and $\Lambda = \sqrt{\beta_1 D_{20}} + \sqrt{\beta_2 D_{10}} > 0$, then the numerator of α_1 is

$$\begin{aligned}\beta_0\beta_1 D_2 - \beta_0\Gamma' &= \beta_0\beta_1 D_2 - \beta_1\beta_2 D_0 + \sqrt{\beta_1\beta_2 D_{10} D_{20}} \\ &= \beta_1 D_{20} + \sqrt{\beta_1\beta_2 D_{10} D_{20}} \\ &= \Lambda\sqrt{\beta_1 D_{20}}.\end{aligned}$$

By symmetry, $\beta_0\beta_2 D_1 - \beta_0\Gamma' = \Lambda\sqrt{\beta_2 D_{10}}$. Further, $\beta_2 D_1 + \beta_1 D_2 - 2\Gamma' = \Lambda^2/\beta_0$. Thus,

$$\begin{aligned}\alpha_1 &= \frac{\beta_0(\beta_1 D_2 - \Gamma')}{\beta_1(\beta_2 D_1 + \beta_1 D_2 - 2\Gamma')} = \frac{\beta_0\sqrt{\beta_1 D_{20}}}{\beta_1\Lambda} > 0 \\ \alpha_2 &= \frac{\beta_0(\beta_2 D_1 - \Gamma')}{\beta_2(\beta_2 D_1 + \beta_1 D_2 - 2\Gamma')} = \frac{\beta_0\sqrt{\beta_2 D_{10}}}{\beta_2\Lambda} > 0.\end{aligned}$$

Since

$$\begin{aligned}\Gamma' &= \frac{\beta_1\beta_2 D_0 - \sqrt{\beta_1\beta_2 D_{10} D_{20}}}{\beta_0} \\ &= \frac{\sqrt{\beta_1\beta_2}}{\beta_0}\left[D_0\sqrt{\beta_1\beta_2} - \sqrt{D_{10} D_{20}}\right]\end{aligned}$$

we have

$$\begin{aligned}&\alpha_1^2\beta_1 D_1 + \alpha_2^2\beta_2 D_2 + 2\alpha_1\alpha_2\Gamma' \\ &= \frac{1}{\Lambda^2}\left[\beta_0^2 D_1 D_{20} + \beta_0^2 D_2 D_{10} + \frac{2\beta_0^2\sqrt{\beta_1\beta_2 D_{10} D_{20}}\Gamma'}{\beta_1\beta_2}\right] \\ &= \frac{\beta_0}{\Lambda^2}\left[\beta_0 D_1 D_{20} + \beta_0 D_2 D_{10} + 2D_0\sqrt{\beta_1\beta_2 D_{10} D_{20}} - 2D_{10} D_{20}\right] \\ &= \frac{\beta_0 D_0}{\Lambda^2}\left[\beta_1 D_{20} + \beta_2 D_{10} + 2\sqrt{\beta_1\beta_2 D_{10} D_{20}}\right] \qquad (35) \\ &= \beta_0 D_0\end{aligned}$$

where (35) comes from the fact that

$$\begin{aligned}\beta_0 D_1 - D_{10} &= (1 - D_0/\sigma^2)D_1 - (D_1 - D_0) \\ &= D_0(1 - D_1/\sigma^2) = \beta_1 D_0\end{aligned}$$

and $\beta_0 D_2 - D_{20} = \beta_2 D_0$ by symmetry, and the last equation follows from the fact that

$$\begin{aligned}&\beta_1 D_{20} + \beta_2 D_{10} + 2\sqrt{\beta_1\beta_2 D_{10} D_{20}} \\ &\qquad = (\sqrt{\beta_1 D_{20}} + \sqrt{\beta_2 D_{10}})^2 = \Lambda^2.\end{aligned}$$ □

Lemma 10: In $\mathcal{D}_3$, $\alpha_1'^2\beta_1 D_1 + \alpha_2'^2\beta_2 D_2 = \beta_0 D_0$.

Proof:

$$\begin{aligned}&\alpha_1'^2\beta_1 D_1 + \alpha_2'^2\beta_2 D_2 \\ &= \left[\frac{\beta_0\beta_1 D_2 + \sqrt{\Delta}}{\beta_1(\beta_2 D_1 + \beta_1 D_2)}\right]^2 \beta_1 D_1 \\ &\quad + \left[\frac{\beta_0\beta_2 D_1 - \sqrt{\Delta}}{\beta_2(\beta_2 D_1 + \beta_1 D_2)}\right]^2 \beta_2 D_2 \\ &= \left[\frac{\beta_0^2\beta_1^2 D_2^2 + 2\beta_0\beta_1 D_2\sqrt{\Delta} + \Delta}{\beta_1(\beta_2 D_1 + \beta_1 D_2)^2}\right] D_1 \\ &\quad + \left[\frac{\beta_0^2\beta_2^2 D_1^2 - 2\beta_0\beta_2 D_1\sqrt{\Delta} + \Delta}{\beta_2(\beta_2 D_1 + \beta_1 D_2)^2}\right] D_2 \\ &= \frac{\beta_0^2 D_1 D_2(\beta_1 D_2 + \beta_2 D_1) + \Delta(\beta_1 D_2 + \beta_2 D_1)/\beta_1\beta_2}{(\beta_2 D_1 + \beta_1 D_2)^2} \\ &= \frac{\beta_0^2 D_1 D_2 + \beta_0(D_0 D_2 + D_0 D_1 - D_1 D_2 - d_0 D_1 D_2)}{\beta_2 D_1 + \beta_1 D_2} \\ &= \beta_0\frac{D_0 D_2 + D_0 D_1 - 2d_0 D_1 D_2}{\beta_2 D_1 + \beta_1 D_2} \\ &= \beta_0 D_0\frac{D_2 + D_1 - 2D_1 D_2/\sigma^2}{\beta_2 D_1 + \beta_1 D_2} \\ &= \beta_0 D_0\end{aligned}$$

where the last equation comes from the definitions of β_1 and β_2. □

Lemma 11: If $0 < \delta_0 < \delta_1 \le \delta_2 < \frac{1}{2}$, then

$$LG(\delta_0, \delta_1, \delta_2) > \frac{1}{2}\log\frac{1}{4\delta_2(1-\delta_2)}.$$

Proof: First, since $\frac{1}{2} < 1 - \delta_2 \le 1 - \delta_1 < 1$ and $0 < \delta_0 < \delta_1 \le \delta_2 < \frac{1}{2}$, we have

$$\sqrt{(1-\delta_1)(1-\delta_2)} - \sqrt{(\delta_1-\delta_0)(\delta_2-\delta_0)} > 0.$$

In addition, $\sqrt{(1-\delta_1)(1-\delta_2)} - \sqrt{(\delta_1-\delta_0)(\delta_2-\delta_0)}$ is a decreasing function of δ_1 for fixed δ_0 and δ_2. Since $\delta_1 \le \delta_2$

$$\begin{aligned}&\sqrt{(1-\delta_1)(1-\delta_2)} - \sqrt{(\delta_1-\delta_0)(\delta_2-\delta_0)} \\ &\qquad \ge \sqrt{(1-\delta_2)(1-\delta_2)} - \sqrt{(\delta_2-\delta_0)(\delta_2-\delta_0)} \\ &\qquad = 1 + \delta_0 - 2\delta_2 \\ &\qquad > 0\end{aligned}$$

where the last inequality follows since $\delta_0 > 0$ and $\delta_2 < \frac{1}{2}$.

As a result

$$\begin{aligned}LG(\delta_0, \delta_1, \delta_2) &\ge LG(\delta_0, \delta_2, \delta_2) \\ &= \frac{1}{2}\log\frac{(1-\delta_0)^2}{4(1-\delta_2)(\delta_2-\delta_0)}.\end{aligned}$$

Taking a partial derivative of the argument inside the logarithm gives

$$\frac{\partial}{\partial \delta_0}\left(\frac{(1-\delta_0)^2}{4(1-\delta_2)(\delta_2-\delta_0)}\right) = \frac{(1-\delta_0)(1+\delta_0-2\delta_2)}{4(1-\delta_2)(\delta_2-\delta_0)^2}.$$

Since $\delta_0 < \delta_2 < 1/2$ and $1+\delta_0 > 2\delta_2$, this derivative is always positive. Therefore,

$$\begin{aligned} LG(\delta_0,\delta_1,\delta_2) > &\ \frac{1}{2}\log \frac{(1-\delta_0)^2}{4(1-\delta_2)(\delta_2-\delta_0)}\bigg|_{\delta_0=0} \\ &= \frac{1}{2}\log\frac{1}{4(1-\delta_2)\delta_2} \end{aligned}$$

which gives us the desired result. □

Acknowledgment

The authors are grateful to the Associate Editor and anonymous reviewers for their detailed suggestions and comments.

References

[1] L. Lastras and T. Berger, "All sources are nearly successively refinable," *IEEE Trans. Inf. Theory*, vol. 47, no. 3, pp. 918–926, Mar. 2001.

[2] H. Feng and M. Effros, "Improved bounds for the rate loss of multiresolution source codes," *IEEE Trans. Info. Theory*, vol. 49, no. 4, pp. 809–821, Apr. 2003.

[3] A. A. El Gamal and T. M. Cover, "Achievable rates for multiple descriptions," *IEEE Trans. Inf. Theory*, vol. IT-28, no. 6, pp. 851–857, Nov. 1982.

[4] M. Fleming and M. Effros, "The rate distortion region for the multiple description problem," in *Proc. IEEE Int. Symp. Information Theory*, Sorrento, Italy, Jun. 2000, p. 208.

[5] R. Venkataramani, G. Kramer, and V. K. Goyal, "Bounds on the achievable region for certain multiple description coding problems," in *Proc. IEEE Int. Symp. Information Theory*, Washington, D.C., Jun. 2001, p. 148.

[6] Z. Zhang and T. Berger, "New results for multiple descriptions," *IEEE Trans. Inf. Theory*, vol. IT-33, no. 4, pp. 502–521, Jul. 1987.

[7] R. Ahlswede, "The rate-distortion region for multiple description without excess rate," *IEEE Trans. Inf. Theory*, vol. IT-31, no. 6, pp. 721–726, Nov. 1985.

[8] Z. Zhang and T. Berger, "Multiple description source coding with no excess marginal rate," *IEEE Trans. Inf. Theory*, vol. 41, no. 2, pp. 349–357, Mar. 1995.

[9] R. Zamir, "The rate loss in the Wyner–Ziv problem," *IEEE Trans. Inf. Theory*, vol. 42, no. 6, pp. 2073–2084, Nov. 1996.

[10] ——, "Gaussian codes and Shannon bounds for multiple descriptions," *IEEE Trans. Inf. Theory*, vol. 45, no. 6, pp. 2629–2636, Nov. 1999.

[11] L. Ozarow, "On a source-coding problem with two channels and three receivers," *Bell Syst. Tech. J.*, vol. 59, no. 10, pp. 1909–1921, Dec. 1980.