Correspondence

The Wyner–Ziv Problem With Multiple Sources

Michael Gastpar, *Member, IEEE*

***Abstract*—This correspondence provides bounds on the rate-distortion region for the distributed compression scenario where two (or more) sources are compressed separately for a decoder that has access to side information. Conclusive rate-distortion results are found for the case where the sources are conditionally independent, given the side information.**

***Index Terms*—Distributed source coding, rate-distortion, side information, source networks, Wyner–Ziv problem.**

I. Introduction

In the problem studied in this correspondence, two (or more) dependent discrete memoryless sources have to be compressed separately from each other in a lossy fashion, i.e., with respect to a fidelity criterion. In addition, the decoder has access to a side information stream that is statistically dependent on ("correlated with") the two sources that need to be compressed. This situation is illustrated in Fig. 1. Clearly, each encoder could compress its respective source, *ignoring* both the side information and the other encoder's presence. For such a scheme, the smallest rates are well known from (standard) single-source rate-distortion theory, see, e.g., [1] or [2, Theorem 2.2.3].

It is well known, however, that both the fact that the two sources are dependent as well as the side information Z permit to lower the necessary rates (or the incurred distortions). The first gain, stemming from the fact that the sources are dependent, has been studied and found by Slepian and Wolf [3] for the case of lossless compression. They considered the system of Fig. 1 without the side information Z. Their surprising result is that the total rate needed for separate (lossless) compression of two sources is the same as the rate needed for joint compression of the two sources (i.e., their joint entropy). When the compression is lossy, the dependence between the sources still permits to lower the rates (see, e.g., [4]), but no conclusive results are available to date. The second gain, stemming from the fact that side information is available at the decoder, has been extensively studied for the case of lossless compression (see e.g., [5, p. 458], with recent extensions to network cases, see, e.g., [6], [7]), and for the case of lossy compression of a single source by Wyner and Ziv [8], [9].

This correspondence investigates the Wyner–Ziv problem with two (and more) discrete[1] memoryless dependent sources and a side information stream, modeled by the sequence $\{(S_{1,k}, S_{2,k}, Z_k)\}_{k=1}^{\infty}$ of independent copies of a triplet of dependent random variables (S_1, S_2, Z) which take values in finite sets $\mathcal{S}_1$, $\mathcal{S}_2$, and $\mathcal{Z}$, respectively, distributed according to a fixed and known pmf

$$p(s_1, s_2, z). \tag{1}$$

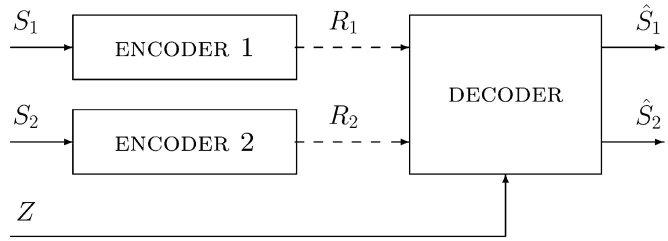

Fig. 1. Separate compression of two sources with side information at the decoder.

The encoder outputs are binary sequences which appear at rates R_1 and R_2 bits per input symbol, respectively. The decoder output is a sequence of pairs $\{(\hat{S}_{1,k}, \hat{S}_{2,k})\}_{k=1}^{\infty}$ whose components take values in finite reproduction alphabets $\hat{\mathcal{S}}_1, \hat{\mathcal{S}}_2$. The encoding is done in blocks of length n, and the fidelity criteria, for $m = 1, 2$, take the shape

$$E\left[\frac{1}{n}\sum_{k=1}^{n} d_m(S_{m,k}, \hat{S}_{m,k})\right] \tag{2}$$

where $d_m(s_m, \hat{s}_m) \geq 0$, $s_m \in \mathcal{S}_m$, $\hat{s}_m \in \hat{\mathcal{S}}_m$, is a given distortion function, and $E[\cdot]$ denotes the expectation operator. We define $\mathcal{R}_{S_1,S_2|Z}^{\mathrm{WZ}}(D_1, D_2)$ as the set of rate pairs (R_2, R_2) for which the system of Fig. 1 can operate when n is large and the average distortions given in (2) are arbitrarily close to D_m, for $m = 1, 2$. The superscript $^{\mathrm{WZ}}$ honors [8] and emphasizes the fact that the side information is *not* available at the encoders.

This correspondence provides bounds to the rate-distortion region $\mathcal{R}_{S_1,S_2|Z}^{\mathrm{WZ}}(D_1, D_2)$. More precisely, after defining notation and conventions in Section II, an inner bound (that is, an achievable rate-distortion region)

$$\mathcal{R}_a(D_1, D_2) \subseteq \mathcal{R}_{S_1,S_2|Z}^{\mathrm{WZ}}(D_1, D_2)$$

is determined in Section III. The coding technique that leads to this region is an extension of the code of [3] (and its extension to the case of lossy compression, given in [4]), combined with the code construction of Wyner and Ziv [8] (based on [3]).

In Section IV, an outer bound

$$\mathcal{R}_c(D_1, D_2) \supseteq \mathcal{R}_{S_1,S_2|Z}^{\mathrm{WZ}}(D_1, D_2)$$

is determined. This region does not appear to coincide with the presented inner bound $\mathcal{R}_a(D_1, D_2)$. This is not surprising since the corresponding rate regions for the problem *without* side information [4] could not be shown to coincide, either.

However, for one special case, namely, when the sources are conditionally independent given the side information, it is shown in Section V, that the two regions *do* coincide. By contrast to earlier work (as reported in [11], [12], [10]), this correspondence provides conclusive rate-distortion results for a *symmetric* case, i.e., *both* sources are

Manuscript received June 6, 2002; revised August 6, 2004. The material in this correspondence was presented in part at the Joint Conference on Communications and Coding, Saas Fee, Switzerland, March 2002, and at the 2003 International Symposium on Information Theory, Yokohama, Japan, June/July 2003.

The author is with the Department of Electrical Engineering and Computer Science, University of California, Berkeley, CA 94720-1770 USA (e-mail: gastpar@eecs.berkeley.edu).

Communicated by P. Narayan, Associate Editor for Shannon Theory.
Digital Object Identifier 10.1109/TIT.2004.836707

[1]The conjecture is that our main results extend without modification to the case of continuous alphabets; however, the proofs in this correspondence are limited to discrete alphabets, in line with the majority of the results in this area. Notable exceptions to this include Wyner's extension [9] of [8] to continuous alphabets, and [10].

encoded, and *both* sources have to be reconstructed with respect to a fidelity criterion.[2]

The results of this correspondence have natural applications to distributed compression and signal processing. One such extension is described in [14]. At the same time, they are also useful to obtain certain rate regions for multiple-relay channels. As a matter of fact, Fig. 1 can be understood as a relay network: The two boxes labeled "encoder" are two relays that observe S_1 and S_2, respectively, and the strategy is for the relays to compress their observations for a final destination that observes Z. Since S_1, S_2, and Z were all produced by the source node in the relay network, they are generally correlated. This is further explained in [15].

II. Notation and Conventions

In this correspondence, random variables are denoted by capital letters, such as X, and their realizations by lower case letters, such as x. The (discrete and finite) alphabet in which the random variable X takes values is denoted by $\mathcal{X}$, and its cardinality by $|\mathcal{X}|$. The pmf of the random variable X is denoted by $p_X(x)$, or simply $p(x)$ when this does not create any confusion. Sequences $x^n = \{x_i\}_{i=1}^n$ will be considered, where each component is sampled independently from a fixed pmf $p(x)$.

Moreover, we will make use of the concept of strong typicality as introduced in [4]. We use the definition as in [5, p. 358], which we quote for convenience. For a sequence x^n and any $a \in \mathcal{X}$, define $N(a|x^n)$ to be the number of occurrences of the symbol a in the sequence x^n. The sequence x^n is said to be ϵ-strongly typical with respect to a pmf $p_X(x)$ if, i) for all $a \in \mathcal{X}$ with $p_X(a) > 0$, $\left|\frac{1}{n}N(a|x^n) - p_X(a)\right| < \epsilon/|\mathcal{X}|$, and ii) for all $a \in \mathcal{X}$ with $p_X(a) = 0$, $N(a|x^n) = 0$. For a given pmf $p(x)$, the set of ϵ-strongly typical sequences will be denoted by $A_\epsilon^{*(n)}(X)$, or simply $A_\epsilon^{*(n)}$, when this does not create any confusion.

Finally, the probability that a considered event $\mathcal{E}$ occurs is denoted as $\mathrm{Prob}\{\mathcal{E}\}$.

III. An Achievable Rate Region

An achievable rate region can be obtained by extending the coding scheme introduced by Slepian and Wolf [3], and elegantly generalized to the concept of "binning" by Cover [16]. In summary, the code leading to the inner bound to the rate-distortion region by Theorem 2 below is the cascade of a suitable vector quantizer with a binning operation for the codeword indices. In particular, the encoder of the source S_1 must apply a binning operation with respect to both the codeword index of source S_2 *and* the side information Z, and the encoder of the source S_2 must do likewise. Given the two bin indices selected by the two encoders, the decoder then uses the side information to undo the binning and retrieve the correct quantization cell indices. But this requires the codewords associated with the two bin indices to be jointly typical with the side information. More precisely, the fact that for each bin index, the corresponding associated codeword is jointly typical with its corresponding source sequence must imply that the two associated codewords and the side information form a jointly typical triplet. The key to such an implication is the Markov lemma.[3]

[2]Earlier work has found that the inner and outer regions of [4] coincide for cases where one of the two sources is either not to be reconstructed, or encoded perfectly (see [11], [12], [10]). Moreover, certain conclusive results are available for the CEO problem [13].

[3]The lemma is quoted without proof in [5, Lemma 14.8.1, p. 436]; with reference to [4], [2]. In [4], the Markov lemma (Lemma 4.1) has a prominent role as well as a proof.

Lemma 1 (Markov Lemma): Consider the three random variables W_1, S_1, and Z whose joint pmf satisfies

$$p(w_1, s_1, z) = p(w_1|s_1)p(s_1, z).$$

For a fixed pair of sequences $(s_1^n, z^n) \in A_\epsilon^{*(n)}(S_1, Z)$, draw a sequence W_1^n according to $\prod_{i=1}^n p_{W_1|S_1}(w_{1,i}|s_{1,i})$. Then

$$\lim_{n\to\infty} \mathrm{Prob}\left\{(W_1^n, s_1^n, z^n) \in A_\epsilon^{*(n)}(W_1, S_1, Z)\right\} = 1.$$

To understand the lemma, note that from the facts that i) a sequence z^n is jointly typical with s_1^n, and ii) s_1^n is jointly typical with w_1^n, it does not yet follow that the three sequences form a jointly typical triplet. One sufficient condition to ensure that they *do* form a jointly typical triplet is precisely the stated Markov relationship. This lemma is at the heart of the presented achievable rate region; the necessary extension is given in Appendix I. We did not find a weaker condition that permits to infer the same conclusions.

Once the decoder has retrieved the correct codeword indices, it can use the side information a second time in order to remove a part of the quantization noise. This is possible because the side information is available in unquantized form. The rates that can be achieved by this coding scheme can be expressed as follows.

Theorem 2: $\mathcal{R}_a(D_1, D_2) \subseteq \mathcal{R}^{\mathrm{WZ}}_{S_1,S_2|Z}(D_1, D_2)$, where $\mathcal{R}_a(D_1, D_2)$ is the set of all rate pairs (R_1, R_2) such that there exists a pair (W_1, W_2) of discrete random variables with

$$p(w_1, w_2, s_1, s_2, z) = p(w_1|s_1)p(w_2|s_2)p(s_1, s_2, z)$$

for which the following conditions are satisfied:

$$R_1 \geq I(S_1, S_2; W_1|Z, W_2) \quad (3)$$

$$R_2 \geq I(S_1, S_2; W_2|Z, W_1) \quad (4)$$

$$R_1 + R_2 \geq I(S_1, S_2; W_1, W_2|Z) \quad (5)$$

and for which there exist functions $g_1(\cdot)$ and $g_2(\cdot)$ such that

$$E[d_1(S_1, g_1(W_1, W_2, Z))] \leq D_1 \text{ and} \quad (6)$$

$$E[d_2(S_2, g_2(W_2, W_1, Z))] \leq D_2. \quad (7)$$

Remark 1: In slight extension of [8, Theorem A2], it can also be shown that it is sufficient to consider auxiliary random variables (W_1, W_2) over alphabets $\mathcal{W}_m$ of cardinalities

$$|\mathcal{W}_m| \leq |\mathcal{S}_m| + 1 \quad (8)$$

for $m = 1, 2, \ldots, M$.

The proof of this theorem is deferred to Appendix I.

The rate region $\mathcal{R}_a$ can easily be extended to more than two sources. For brevity and in order to concentrate on the main result (Section V), we omit an explicit statement and refer to [17].

IV. A General Outer Bound

In this section, we present a region $\mathcal{R}_c(D_1, D_2)$ which contains the desired rate-distortion region $\mathcal{R}^{\mathrm{WZ}}_{S_1,S_2|Z}(D_1, D_2)$. The region $\mathcal{R}_c(D_1, D_2)$ follows from standard outer bounding arguments; it is a slight extension of the arguments given in [4, Theorem 6.2].

Theorem 3 (Outer Bound):

$$\mathcal{R}_c(D_1, D_2) \supseteq \mathcal{R}^{\mathrm{WZ}}_{S_1,S_2|Z}(D_1, D_2)$$

where $\mathcal{R}_c(D_1, D_2)$ is the set of all rate pairs (R_1, R_2) such that there exists a pair (W_1, W_2) of discrete random variables over alphabets $\mathcal{W}_1$ and $\mathcal{W}_2$ with cardinalities $|\mathcal{W}_1||\mathcal{W}_2| \leq 1 + |\mathcal{S}_1||\mathcal{S}_2|$, distributed such that $p(w_1|s_1, s_2, z) = p(w_1|s_1)$ and $p(w_2|s_1, s_2, z) = p(w_2|s_2)$, for which the following conditions are satisfied:

$$R_1 \geq I(S_1, S_2; W_1|Z, W_2) \tag{9}$$

$$R_2 \geq I(S_1, S_2; W_2|Z, W_1) \tag{10}$$

$$R_1 + R_2 \geq I(S_1, S_2; W_1 W_2|Z) \tag{11}$$

and for which there exist functions $g_1(\cdot)$ and $g_2(\cdot)$ such that

$$E[d_1(S_1, g_1(W_1, W_2, Z))] \leq D_1 \text{ and} \tag{12}$$

$$E[d_2(S_2, g_2(W_2, W_1, Z))] \leq D_2. \tag{13}$$

The proof of this theorem is deferred to Appendix II. The extension to the case of more than two sources is explicitly stated in [17].

The region $\mathcal{R}_c(D_1, D_2)$ given in Theorem 3 cannot generally be shown to coincide with $\mathcal{R}_a(D_1, D_2)$, and hence, no conclusive rate-distortion result can be given for the Wyner–Ziv rate-distortion problem with multiple sources. More precisely, the mutual information expressions in Theorems 2 and 3 are exactly the same both for $\mathcal{R}_c(D_1, D_2)$ and $\mathcal{R}_a(D_1, D_2)$; the difference occurs only in the degrees of freedom in choosing the auxiliary random variables W_1 and W_2. More precisely, in Theorem 2, the joint pmf must factor as

$$p(s_1, s_2, z, w_1, w_2) = p(s_1, s_2, z)p(w_1|s_1)p(w_2|s_2).$$

By contrast, in Theorem 3, the joint pmf $p(s_1, s_2, z, w_1, w_2)$ must merely satisfy the marginal constraints

$$p(s_1, s_2, z, w_1) = p(s_1, s_2, z)p(w_1|s_1)$$

and

$$p(s_1, s_2, z, w_2) = p(s_1, s_2, z)p(w_2|s_2).$$

An illustrative example of the difference between these two sets is given in [4] and concerns *mixtures* of the form

$$p(s_1, s_2, z, w_1, w_2) = p(s_1, s_2, z) \sum_{k=1}^{M} \lambda_k p^{(k)}(w_1|s_1)p^{(k)}(w_2|s_2) \tag{14}$$

where $\sum_{k=1}^{M} \lambda_k = 1$. It is immediately clear that such a pmf always satisfies the constraints of Theorem 3, but not necessarily those of Theorem 2.

V. Partial Converse: Conditionally Independent Sources

While the two rate regions derived in this correspondence, $\mathcal{R}_a(D_1, D_2)$ and $\mathcal{R}_c(D_1, D_2)$, do not coincide in general, we now analyze a special case in which they indeed do coincide, hence establishing a conclusive rate-distortion result. This special case is when S_1 and S_2 are independent given Z, i.e., when

$$p(s_1, s_2, z) = p(s_1|z)p(s_2|z)p(z). \tag{15}$$

The first step in our derivation is to rewrite the inner bound to the rate-distortion region, introducing the simplifications due to the assumption that S_1 and S_2 are conditionally independent given Z.

Corollary 4: If S_1 and S_2 are conditionally independent given Z

$$\mathcal{R}_a(D_1, D_2) \subseteq \mathcal{R}^{\mathrm{WZ}}_{S_1,S_2|Z}(D_1, D_2)$$

where $\mathcal{R}_a(D_1, D_2)$ is the set of all rate pairs (R_1, R_2) such that there exists a pair (W_1, W_2) of discrete random variables with

$$p(w_1, w_2, s_1, s_2, z) = p(w_1|s_1)p(w_2|s_2)p(s_1|z)p(s_2|z)p(z)$$

for which the following conditions are satisfied:

$$R_1 \geq I(S_1; W_1) - I(Z; W_1) \tag{16}$$

$$R_2 \geq I(S_2; W_2) - I(Z; W_2) \tag{17}$$

and for which there exist functions $g_1(\cdot)$ and $g_2(\cdot)$ such that

$$E[d_1(S_1, g_1(W_1, Z))] \leq D_1 \text{ and} \tag{18}$$

$$E[d_2(S_2, g_2(W_2, Z))] \leq D_2. \tag{19}$$

Proof: The achievability of the rate-distortion region of this corollary can be inferred from the Wyner–Ziv rate-distortion function of a single source. It can also be inferred from Theorem 2 by noting that the rate conditions of can be expressed equivalently as (43)–(45). The term $I(W_1; W_2|Z)$ in the sum rate bound is zero. Therefore, the sum rate bound, (45), becomes merely the sum of the two side bounds, (43), (44), and hence can be omitted. The side bounds can be simplified by writing out $I(Z, W_2; W_1) = I(Z; W_1) + I(W_1; W_2|Z)$, where the last term is again zero. For the distortion D_1, we know from Theorem 2 that if $p(w_1, w_2, s_1, s_2, z)$ satisfies (16) and (17), then the distortion D_1 is achievable if there exists a function $\tilde{g}_1(w_1, w_2, z)$ such that $E\left[d_1(S_1, \tilde{g}_1(W_1, W_2, Z)\right] \leq D_1$. Consider

$$\begin{aligned} &E\left[d_1(S_1, \tilde{g}_1(W_1, W_2, Z))\right] \\ &\quad = E_{W_2,S_2,Z}\left[E_{W_1,S_1|W_2,S_2,Z}\left[d_1(S_1, \tilde{g}_1(W_1, W_2, Z))\right]\right] \\ &\quad = E_{W_2,S_2,Z}\left[E_{W_1,S_1|Z}\left[d_1(S_1, \tilde{g}_1(W_1, W_2, Z))\right]\right] \end{aligned}$$

where the last equation follows because $p(w_1, s_1|z, s_2, w_2) = p(w_1, s_1|z)$. But then, define $w_2^*(z)$ such that

$$w_2^*(z) \in \arg\min_{w_2} E_{S_1,W_1|Z=z}\left[d_1(S_1, \tilde{g}_1(W_1, w_2, z))\right].$$

Clearly, for any $z \in \mathcal{Z}$

$$\begin{aligned} &E_{W_2,S_2|Z=z}\left[E_{S_1,W_1|Z=z}\left[d_1(S_1, \tilde{g}_1(W_1, W_2, z))\right]\right] \\ &\qquad \geq E_{S_1,W_1|Z=z}\left[d_1(S_1, \tilde{g}_1(W_1, w_2^*(z), z))\right]. \end{aligned} \tag{20}$$

Therefore, introduce $g_1(w_1, z) = \tilde{g}_1(w_1, w_2^*(z), z)$, and note that $E\left[d_1(S_1, g_1(W_1, Z)\right] \leq D_1$. In other words, for any $p(w_1|s_1)$ satisfying (16), there exists $\tilde{g}_1(w_1, w_2, z)$ satisfying (6) if and only if there exists $g_1(w_1, z)$ satisfying (18). The proof for the second encoder goes along the same lines. □

The outer bound to the rate region derived in this correspondence, $\mathcal{R}_c(D_1, D_2)$, can also be replaced by a simpler bound in the special case when S_1 and S_2 are conditionally independent given Z, as follows.

Corollary 5: If S_1 and S_2 are conditionally independent given Z, $\mathcal{R}'_c(D_1, D_2) \supseteq \mathcal{R}_c(D_1, D_2)$, and hence $\mathcal{R}'_c(D_1, D_2) \supseteq \mathcal{R}_{S_1,S_2|Z}(D_1, D_2)$, where $\mathcal{R}'_c(D_1, D_2)$ is the set of all rate pairs (R_1, R_2) such that there exists a pair (W_1, W_2) of discrete random variables with $p(w_1|s_1, s_2, z) = p(w_1|s_1)$ and $p(w_2|s_1, s_2, z) = p(w_2|s_2)$ for which the following conditions are satisfied:

$$R_1 \geq I(S_1; W_1) - I(Z; W_1) \tag{21}$$

$$R_2 \geq I(S_2; W_2) - I(Z; W_2) \tag{22}$$

and for which there exist functions $g_1(\cdot)$ and $g_2(\cdot)$ such that

$$E[d_1(S_1, g_1(W_1, Z))] \leq D_1, \text{ and} \tag{23}$$

$$E[d_2(S_2, g_2(W_2, Z))] \leq D_2. \tag{24}$$

Proof: This lower bound can be established by noting that the rate of Encoder 1 is at least the rate for the single-source Wyner–Ziv problem when the side information at the decoder is Z and S_2, which is known to be [8, Theorem 1]

$$R_1 = \min I(S_1; W_1) - I(Z, S_2; W_1) \tag{25}$$

where the minimum is over all auxiliary random variables W_1 whose distribution satisfies

$$p(w_1, s_1, z, s_2) = p(w_1|s_1)p(s_1, z, s_2)$$

and for which there exists a function $\tilde{g}_1(w_1, s_2, z)$ such that

$$E\left[d_1(S_1, \tilde{g}_1(W_1, S_2, Z)\right] \le D_1. \tag{26}$$

First, the rate expression can be simplified by writing out

$$I(Z, S_2; W_1) = I(Z; W_1) + I(S_2; W_1|Z) \tag{27}$$

where the last mutual information expression is zero because $p(s_2, z, w_1) = p(s_2|z)p(w_1|z)p(z)$ (Markov property). For the distortion, consider a fixed choice of $p(w_1|s_1)$ and corresponding $\tilde{g}_1(\cdot)$ satisfying (25) and (26). The expectation in (26) can be computed as

$$\begin{aligned} &E\left[d_1(S_1, \tilde{g}_1(W_1, S_2, Z)\right] \\ &\quad = E_{S_2,Z}\left[E_{S_1,W_1|S_2,Z}\left[d_1(S_1, \tilde{g}_1(W_1, S_2, Z))\right]\right] \\ &\quad = E_{S_2,Z}\left[E_{S_1,W_1|Z}\left[d_1(S_1, \tilde{g}_1(W_1, S_2, Z))\right]\right] \end{aligned} \tag{28}$$

where the last equation follows because $p(w_1, s_1|z, s_2) = p(w_1, s_1|z)$. But then, define $s_2^*(z)$ such that

$$s_2^*(z) \in \arg\min_{s_2} E_{S_1,W_1|Z=z}\left[d_1(S_1, \tilde{g}_1(W_1, s_2, z))\right].$$

Clearly, for any $z \in \mathcal{Z}$

$$\begin{aligned} &E_{S_2|Z=z}\left[E_{S_1,W_1|Z=z}\left[d_1(S_1, \tilde{g}_1(W_1, S_2, z))\right]\right] \\ &\qquad \ge E_{S_1,W_1|Z=z}\left[d_1(S_1, \tilde{g}_1(W_1, s_2^*(z), z))\right]. \end{aligned} \tag{29}$$

Therefore, introduce $g_1(w_1, z) = \tilde{g}_1(w_1, s_2^*(z), z)$, and note that $E\left[d_1(S_1, g_1(W_1, Z)\right] \le D_1$. In other words, for any W_1 satisfying (25), if there exists $\tilde{g}_1(w_1, s_2, z)$ satisfying (26), then, there exists $g_1(w_1, z)$ satisfying (23). Thus, any coding scheme for Encoder 1 must satisfy (21) and (23). For Encoder 2, the same proof argument establishes (22) and (24). □

The main result of this section follows by combining Corollaries 4 and 5, and by observing that the additional degrees of freedom in Corollary 5 (more particularly, the additional freedom in choosing the auxiliary random variables W_1 and W_2) do not permit to lower the involved mutual information and distortion functionals. It follows that the rate regions described by Corollaries 4 and 5 are actually the same, and hence, that they correspond to the desired rate-distortion region $\mathcal{R}^{\text{WZ}}_{S_1,S_2|Z}(D_1, D_2)$.

Theorem 6: If S_1 and S_2 are conditionally independent given Z, then

$$\mathcal{R}_a(D_1, D_2) = \mathcal{R}_c(D_1, D_2) = \mathcal{R}^{\text{WZ}}_{S_1,S_2|Z}(D_1, D_2). \tag{30}$$

Proof: Both rate regions have the same shape *except* that the auxiliary random variables W_1 and W_2 in Corollary 5 have more degrees of freedom. However, since all of the involved mutual information functionals only depend on the joint marginals of (S_1, W_1, Z) and (S_2, W_2, Z), the additional degrees of freedom cannot lower their values.

More precisely, for a choice of (W_1, W_2) satisfying the conditions of Corollary 5 (and the corresponding rate pair $(R_1, R_2) \in \mathcal{R}_c(D_1, D_2)$), the resulting joint pmf can be written as

$$p(w_1, w_2, s_1, s_2, z) = p(w_1, w_2|s_1, s_2)p(s_1|z)p(s_2|z)p(z).$$

But then, construct the auxiliary random variables (W_1', W_2') such that

$$p_{W_1'|S_1}(w_1|s_1) = \sum_{w_2, s_2} p(w_1, w_2|s_1, s_2)p(s_2|s_1) \tag{31}$$

$$p_{W_2'|S_2}(w_2|s_2) = \sum_{w_1, s_1} p(w_1, w_2|s_1, s_2)p(s_1|s_2). \tag{32}$$

The joint pmf

$$p(w_1', w_2', s_1, s_2, z) = p(w_1'|s_1)p(w_2'|s_2)p(s_1|z)p(s_2|z)p(z)$$

induces the same marginal distributions on (S_1, W_1, Z) and (S_2, W_2, Z). Trivially, therefore, the mutual information and the distortion functionals in Corollaries 4 and 5 must assume the same values, but W_1' and W_2' are achievable since they satisfy the conditions of Corollary 4. □

By the nature of the arguments leading to Theorem 6, it is clear that the result carries over to the case of more than two sources. This is explicitly stated in [17].

VI. Conclusion

Distributed lossy compression as shown in Fig. 1 (without the side information) is a long-standing open problem. The best known achievable rate region is the one given in [4], and it does not coincide with any converse bound; no conclusive rate-distortion results can be given. In this correspondence, we investigate distributed lossy compression with side information. In extension of [4], we present inner and outer bounds to the rate-distortion region, and we establish conclusive rate-distortion results for the scenario where the sources are conditionally independent, given the side information. The extension of our results to the case of more than two sources appears in [17].

Appendix I
Proof of Theorem 2

Before sketching the proof of Theorem 2, we consider the following extension of the Markov lemma (Lemma 1).

Lemma 7 (Extended Markov Lemma): Consider the five random variables W_1, W_2, S_1, S_2, and Z whose joint pmf satisfies

$$p(w_1, w_2, s_1, s_2, z) = p(w_1|s_1)p(w_2|s_2)p(s_1, s_2, z).$$

For a fixed triple of sequences $(s_1^n, s_2^n, z^n) \in A_\epsilon^{*(n)}(S_1, S_2, Z)$, draw a sequence W_1^n according to $\prod_{i=1}^n p_{W_1|S_1}(w_{1,i}|s_{1,i})$, and a sequence W_2^n (independently of W_1^n) according to $\prod_{i=1}^n p_{W_2|S_2}(w_{2,i}|s_{2,i})$. Then

$$\lim_{n\to\infty} \text{Prob}\left\{(W_1^n, W_2^n, s_1^n, s_2^n, z^n) \in A_\epsilon^{*(n)}(W_1, W_2, S_1, S_2, Z)\right\} = 1. \quad (33)$$

Proof: Consider the three random variables $(W_1, W_2), (S_1, S_2)$, and Z. Clearly, (W_1, W_2) and Z are conditionally independent, given (S_1, S_2). In order to apply Lemma 1, consider a fixed pair of sequences $((s_1^n, s_2^n), z^n) \in A_\epsilon^{*(n)}$, and select a pair of sequences (W_1^n, W_2^n) by drawing from the pmf

$$\prod_{i=1}^n p_{W_1,W_2|S_1,S_2}(w_{1,i}, w_{2,i}|s_{1,i}, s_{2,i}). \quad (34)$$

From Lemma 1, we know that then

$$\lim_{n\to\infty} \text{Prob}\left\{(W_1^n, W_2^n, s_1^n, s_2^n, z^n) \in A_\epsilon^{*(n)}\right\} = 1.$$

But by assumption, $p(w_1, w_2|s_1, s_2) = p(w_1|s_1)p(w_2|s_2)$, and therefore, drawing (W_1^n, W_2^n) from the pmf in (34) is the same as drawing W_1^n from $\prod_{i=1}^n p_{W_1|S_1}(w_{1,i}|s_{1,i})$, and W_2^n from $\prod_{i=1}^n p_{W_2|S_2}(w_{2,i}|s_{2,i})$, which completes the proof. □

An extension of a second technical lemma is needed, which, for weakly typical sequences, is given, e.g., in [5, Theorem 8.6.1]. The version for strongly typical sequences can be found in [4, p. 198] and [5, Theorem 13.6.2].

Lemma 8: If $(\tilde{X}^n, \tilde{Y}^n, \tilde{Z}^n) \sim p(x^n)p(y^n)p(z^n)$, i.e., they have the same marginals as $p(x^n, y^n, z^n)$ but they are independent, then

$$\text{Prob}\left\{(\tilde{X}^n, \tilde{Y}^n, \tilde{Z}^n) \in A_\epsilon^{*(n)}\right\} \le 2^{-n(I(Z,X;Y)+I(Z,Y;X)-I(X;Y|Z)-\epsilon_0)} \quad (35)$$

where $\epsilon_0 \to 0$ as $\epsilon \to 0$ and $n \to \infty$.

Outline of the Proof: The lemma can be established along the lines of [5, Theorem 13.6.2]. We here give the explicit argument for the case of weak typicality, i.e., extending the proof of [5, Theorem 8.6.1],

$$\begin{aligned}\text{Prob}&\left\{(\tilde{X}^n, \tilde{Y}^n, \tilde{Z}^n) \in A_\epsilon^{(n)}\right\}\\ &= \sum_{(x^n,y^n,z^n)\in A_\epsilon^{(n)}} p(x^n)p(y^n)p(z^n)\\ &\le 2^{n(H(X,Y,Z)+\epsilon)}2^{-n(H(X)-\epsilon)}2^{-n(H(Y)-\epsilon)}2^{-n(H(Z)-\epsilon)}\\ &= 2^{-n(H(X)+H(Y)+H(Z)-H(X,Y,Z)-4\epsilon)} \end{aligned} \quad (36)$$

where $A_\epsilon^{(n)}$ denotes the typical set as defined in [5, p. 51]. The expression in the exponent can be rewritten as $I(Z,X;Y) + I(Z,Y;X) - I(X;Y|Z)$, which completes the proof for weakly typical sequences.

Proof: (Proof Sketch of Theorem 2.) For $m = 1, 2$, fix $p(w_m|s_m)$ as well as $g_m(w_1, w_2, z)$ in such a way that $E[d_m(S_m, g_m(W_1, W_2, Z))] \le D_m$. Generate $2^{nR'_m}$ codewords of length n, sampled independent and identically distributed (i.i.d.) from the marginal pmf $p(w_m)$. Label these as $w_m^n(u_m)$, with $u_m \in \{1, 2, \ldots, 2^{nR'_m}\}$. Provide 2^{nR_m} bins, indexed by $t_m \in \{1, 2, \ldots, 2^{nR_m}\}$. Randomly assign to every codeword $w_m^n(u_m)$ one of the bin indices t_m. Denote the set of codeword indices u_m with bin index t_m as $B_m(t_m)$.

Encoding: For $m = 1, 2$, given a source sequence S_m^n, encoder m looks for a codeword $W_m^n(u_m)$ such that $(S_m^n, W_m^n(u_m)) \in A_\epsilon^{*(n)}$. The probability of finding such a codeword can be made arbitrarily close to 1 (as $n \to \infty$) as long as

$$R'_m > I(S_m; W_m). \quad (37)$$

Encoder m sends the index t_m for which $u_m \in B_m(t_m)$.

Decoding: The decoder looks for a pair $(W_1^n(u_1), W_2^n(u_2))$ such that $u_1 \in B_1(t_1)$, $u_2 \in B_2(t_2)$, and

$$(W_1^n(u_1), W_2^n(u_2), Z^n) \in A_\epsilon^{*(n)}. \quad (38)$$

If a unique (u_1, u_2) is found, the decoder outputs $(\hat{S}_1^n, \hat{S}_2^n)$, and it is also true that then $(S_1^n, S_2^n, W_1^n(u_1), W_2^n(u_2), Z^n) \in A_\epsilon^{*(n)}$. By the construction of $\hat{S}_1^n$ and $\hat{S}_2^n$, this implies (using standard proof techniques, see, e.g., [4, p. 212]) that

$$\begin{aligned}\frac{1}{n}\sum_{k=1}^n d(S_{1,k}, \hat{S}_{1,k}) &= \frac{1}{n}\sum_{a,b,c,d} d(a, g_1(b,c,d))N(a,b,c,d|S_1^n, W_1^n, W_2^n, Z^n)\\ &\le \frac{1}{n}\sum_{a,b,c,d} d(a, g_1(b,c,d))\\ &\quad \left(np_{S_1,W_1,W_2,Z}(a,b,c,d) + n\epsilon\frac{1}{|\mathcal{S}_1||\mathcal{W}_1||\mathcal{W}_2||\mathcal{Z}|}\right)\end{aligned}$$

where the second and third sum are over all $a \in \mathcal{S}_1$, $b \in \mathcal{W}_1$, $c \in \mathcal{W}_2$, $d \in \mathcal{Z}$. By assumption on the pmf $p_{S_1,W_1,W_2,Z}(\cdot)$, the last expression tends to D_1 as $n \to \infty$. For the distortion D_2, the argument follows along the same lines.

In order to guarantee that the decoder finds a unique pair of codeword indices (u_1, u_2), the following two error events must be considered.

1) The two true codewords and the side information do not form a jointly typical triplet: The joint pmf $p(w_1, w_2, s_1, s_2, z)$ has the structure required in Theorem 2, and hence, by Lemma 7, the probability of this event tends to zero as $n \to \infty$.
2) There exists an alternative choice of two codewords that, together with the side information, form a jointly typical triplet: Denoting the correct codewords indices by (u_1, u_2) and their corresponding bin indices by (t_1, t_2), consider an alternative choice in the first component, (u'_1, u_2). By [5, Lemma 13.6.2]

$$\text{Prob}\left\{\left(W_1^n(u'_1), W_2^n(u_2), Z^n\right) \in A_\epsilon^{*(n)}\right\} \le 2^{-n(I(W_1;W_2,Z)-\epsilon_1)} \quad (39)$$

where ϵ_1 tends to zero as $\epsilon \to 0$ and $n \to \infty$ The number of codewords with bin index t_1 is at most $2^{n(R_1-R'_1)} + \delta_1(n)$, where $\delta_1(n)$ tends to zero as $n \to \infty$, and hence,

$$\begin{aligned}\text{Prob}&\left\{\exists u'_1 \in B(t_1), u'_1 \ne u_1 : (w_1(u'_1), w_2(u_2), z) \in A_\epsilon^{*(n)}\right\}\\ &\le \left(2^{n(R_1-R'_1)} + \delta_1(n)\right)2^{-n(I(W_1;W_2,Z)-\epsilon_1)}.\end{aligned}$$

The same derivation applies to the case of an alternative choice in the second component, (u_1, u'_2). Finally, if both indices are in error, Lemma 8 is needed, yielding

$$\text{Prob}\left\{\left(W_1^n(su'_1), W_2^n(u'_2), Z^n\right) \in A_\epsilon^{*(n)}\right\} \le 2^{-n(I(Z,W_2;W_1)+I(Z,W_1;W_2)-I(W_1;W_2|Z)-\epsilon_{12})}$$

where ϵ_{12} tends to zero as $\epsilon \to 0$ and $n \to \infty$. The number of codewords with bin indices (t_1, t_2) is at most $2^{n(R_1-R'_1+R_2-R'_2)} + \delta_{12}(n)$, where $\delta_{12}(n)$ tends to zero as

$n \to \infty$. Hence, the probability of the error event can be bounded as

$$\begin{aligned}\text{Prob}\Big\{ & \exists(u_1', u_2'), u_1' \in B(t_1), u_1' \neq u_1, u_2' \in B(t_2), \\ & u_2' \neq u_2 : \left(W_1^n(u_1'), W_2^n(u_2'), Z^n\right) \in A_\epsilon^{*(n)} \Big\} \\ & \leq \left(2^{n(R_1 - R_1' + R_2 - R_2')} + \delta_{12}(n)\right) \\ & \leq 2^{-n(I(Z,W_2;W_1)+I(Z,W_1;W_2)-I(W_1;W_2|Z)-\epsilon_{12})}.\end{aligned}$$

In conclusion, this error event will have vanishingly small error probability (as n tends to infinity) when the following rate conditions are satisfied:

$$R_1' - R_1 < I(Z, W_2; W_1) \tag{40}$$
$$R_2' - R_2 < I(Z, W_1; W_2) \tag{41}$$
$$\begin{aligned}R_1' - R_1 + R_2' - R_2 < I(Z, W_2; W_1) + I(Z, W_1; W_2) \\ - I(W_1; W_2|Z).\end{aligned} \tag{42}$$

Combining (40)–(42) with (37) yields

$$R_1 \geq I(S_1; W_1) - I(Z, W_2; W_1) \tag{43}$$
$$R_2 \geq I(S_2; W_2) - I(Z, W_1; W_2) \tag{44}$$
$$\begin{aligned}R_1 + R_2 \geq I(S_1; W_1) - I(Z, W_2; W_1) + I(S_2; W_2) \\ - I(Z, W_1; W_2) + I(W_1; W_2|Z).\end{aligned} \tag{45}$$

To see that these are the conditions claimed in Theorem 2, observe that

$$I(S_1, S_2; W_1|Z, W_2) = I(S_1; W_1|Z, W_2)$$

since $I(S_2; W_1|Z, S_1, W_2) = 0$. But then, write out $I(S_1, Z, W_2; W_1)$ in two different ways as

$$\begin{aligned}I(S_1, Z, W_2; W_1) &= I(Z, W_2; W_1) + I(S_1; W_1|Z, W_2) \\ &= I(S_1; W_1) + I(Z, W_2; W_1|S_1)\end{aligned}$$

where the last term is zero, which implies that

$$I(S_1; W_1) - I(Z, W_2; W_1) = I(S_1, S_2; W_1|Z, W_2).$$

Similar transformations show that (44) and (45) are equal to (4) and (5), respectively. □

Appendix II
Proof of Theorem 3

Along the lines of the proof of [4, Theorem 6.2], suppose that the encoders and the decoder achieve average distortions $(D_{1,i}, D_{2,i})$ in the reproduction of $(X_{1,i}, X_{2,i})$, and hence, overall average distortions

$$D_1 = \frac{1}{n}\sum_{i=1}^n D_{1,i} \quad \text{and} \quad D_2 = \frac{1}{n}\sum_{i=1}^n D_{2,i}.$$

The goal is to show that this implies that the corresponding coding rates $(R_1, R_2) \in \mathcal{R}_c(D_1, D_2)$. For an observed source sequence S_1^n, Encoder 1 must provide the decoder with an index, denoted by T_1. The following is true about T_1:

$$\begin{aligned}nR_1 &\geq H(T_1) \\ &\overset{(a)}{\geq} H(T_1|T_2, Z^n) \\ &\geq H(T_1|T_2, Z^n) - H(T_1|T_2, Z^n, S_1^n, S_2^n) \\ &\overset{(b)}{=} I(S_1^n S_2^n; T_1|T_2, Z^n) \\ &\overset{(c)}{=} \sum_{i=1}^n I\left(S_{1,i}, S_{2,i}; T_1|T_2, Z^n, S_{1,1}^{i-1}, S_{2,1}^{i-1}\right) \\ &\overset{(b)}{=} \sum_{i=1}^n H\left(S_{1,i}, S_{2,i}|T_2, Z^n, S_{1,1}^{i-1}, S_{2,1}^{i-1}\right) \\ &\quad - H\left(S_{1,i}, S_{2,i}|T_1, T_2, Z^n, S_{1,1}^{i-1}, S_{2,1}^{i-1}\right) \\ &\overset{(d)}{=} \sum_{i=1}^n H\left(S_{1,i}, S_{2,i}|W_{2,i}, Z_i\right) \\ &\quad - H\left(S_{1,i}, S_{2,i}|W_{1,i}, W_{2,i}, Z_i\right) \\ &\overset{(b)}{=} \sum_{i=1}^n I\left(S_{1,i}, S_{2,i}; W_{1,i}|W_{2,i}, Z_i\right).\end{aligned} \tag{46}$$

For (a), recall that further conditioning cannot increase entropy; (b) is the definition of mutual information; and (c) is the chain rule for mutual information [5, Theorem 2.5.2]. For (d), we define

$$W_{1,i} = \left(T_1, S_{1,1}^{i-1}, S_{2,1}^{i-1}, Z_1^{i-1}, Z_{i+1}^n\right)$$

and

$$W_{2,i} = \left(T_2, S_{1,1}^{i-1}, S_{2,1}^{i-1}, Z_1^{i-1}, Z_{i+1}^n\right).$$

Note that with this definition, it is indeed true that $W_{1,i}$ is conditionally independent of $(S_{2,i}, Z_i)$ for given $S_{1,i}$, and the corresponding is true for $W_{2,i}$.

For the sum rate bound, we find similarly

$$\begin{aligned}n(R_1 + R_2) &\geq H(T_1, T_2) \\ &\overset{(a)}{\geq} H(T_1, T_2|Z^n) \\ &= I(T_1, T_2; S_1^n, S_2^n|Z^n) \\ &\overset{(b)}{=} \sum_{i=1}^n H\left(S_{1,i}, S_{2,i}|Z^n, S_1^{i-1}, S_2^{i-1}\right) \\ &\quad - H\left(S_{1,i}, S_{2,i}|T_1, T_2, S_{1,1}^{i-1}, S_{2,1}^{i-1}, Z^n\right) \\ &\overset{(c)}{=} \sum_{i=1}^n H\left(S_{1,i}, S_{2,i}|Z_i\right) \\ &\quad - H\left(S_{1,i}, S_{2,i}|T_1, T_2, S_{1,1}^{i-1}, S_{2,1}^{i-1}, Z^n\right) \\ &\overset{(d)}{=} \sum_{i=1}^n H\left(S_{1,i}, S_{2,i}|Z_i\right) \\ &\quad - H\left(S_{1,i}, S_{2,i}|W_{1,i}, W_{2,i}, Z_i\right) \\ &= \sum_{i=1}^n I\left(S_{1,i}S_{2,i}; W_{1,i}W_{2,i}|Z_i\right).\end{aligned} \tag{47}$$

For (a), recall that further conditioning cannot increase entropy; (b) is the chain rule for mutual information; and (c) holds because $(S_{1,i}, S_{2,i})$ is conditionally independent of Z_1^{i-1}, Z_{i+1}^n, S_1^{i-1}, S_2^{i-1}, given Z_i. For (d), we use again the definitions of W_1 and W_2. Note that the sum rate bound can be proved just like in the case of the single-source Wyner–Ziv problem, see, e.g., [5, p. 440].

To complete the argument, define rate pairs $(R_{1,i}, R_{2,i})$ that satisfy the constraints given by the summands with index i in the sums in (46) (and its counterpart for R_2) as well as in (47), and incur distortions $(D_{1,i}, D_{2,i})$. The result now follows from a standard convexity argument that is no different from the one used in the proof of [4, Theorem 6.2]. It can be shown that $\mathcal{R}_c(D_1, D_2)$ is convex in the sense that if

$$(R_{1,i}, R_{2,i}) \in \mathcal{R}_c(D_{1,i}, D_{2,i})$$

and

$$(R_{1,j}, R_{2,j}) \in \mathcal{R}_c(D_{1,j}, D_{2,j})$$

then

$$(\lambda(R_{1,i}, R_{2,i}) + (1-\lambda)(R_{1,j}, R_{2,j})) \in \mathcal{R}_c(\lambda(D_{1,i}, D_{2,i}) + (1-\lambda)(D_{1,j}, D_{2,j}))$$

implying that if the code achieves D_1 and D_2, the corresponding rates R_1 and R_2 must satisfy the conditions stated in Theorem 3.

The bound on the cardinality of the alphabets of the auxiliary random variables is a direct consequence of [8, Theorem A2]. □

Acknowledgment

The author would like to acknowledge the excellent manuscript of Prof. T. Berger [4], as well as stimulating discussions with Dr. P. L. Dragotti, Dr. G. Kramer, and Prof. M. Vetterli. Finally, the comments of the reviewers and the Associate Editor helped to improve this manuscript.

References

[1] T. Berger, *Rate Distortion Theory: A Mathematical Basis for Data Compression*. Englewood Cliffs, NJ: Prentice-Hall, 1971.
[2] I. Csiszár and J. Körner, *Information Theory: Coding Theory for Discrete Memoryless Systems*. New York: Academic, 1981.
[3] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inform. Theory*, vol. IT-19, pp. 471–480, July 1973.
[4] T. Berger, "Multiterminal source coding," presented at the CISM Summer School on the Information Theory Approach to Communications, July 1977.
[5] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.
[6] A. D. Wyner, J. K. Wolf, and F. M. J. Willems, "Communication via a processing broadcast satellite," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1243–1249, June 2002.
[7] I. Csiszár and P. Narayan, "The secret key capacity for multiple terminals," in *Proc. IEEE Int. Symp. Information Theory*, Lausanne, Switzerland, June/July 2002, pp. 27–27.
[8] A. D. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. Inform. Theory*, vol. IT-22, pp. 1–11, Jan. 1976.
[9] A. D. Wyner, "The rate-distortion function for source coding with side information at the decoder—II: General sources," *Inform. Contr.*, vol. 38, pp. 60–80, 1978.
[10] Y. Oohama, "Gaussian multiterminal source coding," *IEEE Trans. Inform. Theory*, vol. 43, pp. 1912–1923, Nov. 1997.
[11] A. H. Kaspi and T. Berger, "Rate-distortion for correlated source with partially separated encoders," *IEEE Trans. Inform. Theory*, vol. IT-28, pp. 828–840, Nov. 1982.
[12] T. Berger and R. W. Yeung, "Multiterminal source encoding with one distortion criterion," *IEEE Trans. Inform. Theory*, vol. IT-35, pp. 228–236, Mar. 1989.
[13] Y. Oohama, "The rate-distortion function for the quadratic Gaussian CEO problem," *IEEE Trans. Inform. Theory*, vol. 44, pp. 1057–1070, May 1998.
[14] M. Gastpar, P. L. Dragotti, and M. Vetterli, "The distributed Karhunen–Loève transform," in *Proc. 2002 Int. Workshop on Multimedia Signal Processing*, St. Thomas, U.S. Virgin Islands, Dec. 2002.
[15] M. Gastpar, G. Kramer, and P. Gupta, "The multiple-relay channel: Coding and antenna-clustering capacity," in *Proc. IEEE Int. Symp. Information Theory*, Lausanne, Switzerland, June/July 2002, pp. 136–136.
[16] T. M. Cover, "A proof of the data compression theorem of Slepian and Wolf for ergodic sources," *IEEE Trans. Inform. Theory*, vol. IT-21, pp. 226–228, Mar. 1975.
[17] M. Gastpar, "On Wyner–Ziv networks," in *Proc. 37th Asilomar Conf. Signals, Systems, and Computers*, Asilomar, CA, Nov. 2003.

On Properties of Rate-Reliability-Distortion Functions

Ashot N. Harutyunyan and
Evgueni A. Haroutunian, *Associate Member, IEEE*

***Abstract*—Some important properties of the rate-reliability-distortion function of discrete memoryless source (DMS) are established. For the binary source and Hamming distortion measure this function is derived and analyzed. Even that elementary case suffices to show the nonconvexity of the rate-reliability-distortion function in the reliability (error exponent) argument.**

***Index Terms*—Convexity, error exponent, Hamming distance, rate-distortion function, rate-reliability-distortion function, reliability, time-sharing argument.**

I. Introductory Notes

We treat the performance bound in the source coding problem under fidelity and *reliability* (error exponent) criteria, namely, the rate-reliability-distortion function. In this concise correspondence, we summarize some basic facts and results on the properties of that function and the concept at all.

Shannon [11] defined the rate-distortion function as the minimal coding rate that can be asymptotically achieved for transmission of an information source data with an average distortion less than a predetermined threshold. It is important also the study of the rate-distortion problem under an additional coding characteristic—an exponential decay in the error probability. The maximum error exponent as a function of coding rate and distortion, characterizing the same source coding system was studied by Marton in [10]. An alternative order dependence of the three parameters was introduced by Haroutunian and Mekoush in [7], defining the rate-reliability-distortion function as the minimal rate at which the messages of a source can be encoded and then reconstructed by the receiver with an exponentially decreasing error probability with increasing codeword length. In this approach, the achievability of the coding rate R is considered as a function of a fixed distortion level $\Delta \geq 0$ and an error exponent $E > 0$.

A number of publications of the coauthors and their collaborators during the past years have been devoted to the development of this idea (and the equivalent approach applied to channel coding introduced in [6]) toward the multiuser source coding problems. Among those are the works concerning the multiple descriptions problem [9], successive refinement of information [8]. Recently, this approach was adopted by Tuncel and Rose [12].

Essentially, as an advantage of the approach considered in [7], we can emphasize the technical ease of treatment of the coding rate as a function of distortion and error exponent which, at the same time, allows to convert readily the results from the rate-reliability-distortion area to the rate-distortion ones looking at the extremal values of the reliability, e.g., $E \to 0$, $E \to \infty$. The importance of that fact is especially pronounced and evident when one deals with a multidimensional

Manuscript received May 26, 2003; revised June 22, 2004. The work of E. A Haroutunian was supported by INTAS under Grant 00-738. The material in this correspondence was presented in part at the Symposium on Information Theory and Some Friendly Neighbors—Ein Wunschkonzert, ZiF, Bielefeld University, Bielefeld, Germany, August 2003.

A. N. Harutyunyan is with the Epygi Labs AM LLC, 375026 Yerevan, Armenia (e-mail: ashot.harutyunyan@epygilab.am).

E. A. Haroutunian is with the Institute for Informatics and Automation Problems of the Armenian National Academy of Sciences and the Yerevan State University, 375044 Yerevan, Armenia (e-mail: evhar@ipia.sci.am).

Communicated by R. W. Yeung, Associate Editor for Shannon Theory.
Digital Object Identifier 10.1109/TIT.2004.836706