Robust Distributed Source Coding

Jun Chen, *Member, IEEE*, and Toby Berger, *Fellow, IEEE*

***Abstract*—We consider a distributed source coding system in which several observations must be encoded separately and communicated to the decoder by using limited transmission rate. We introduce a robust distributed coding scheme which flexibly trades off between system robustness and compression efficiency. The optimality of this coding scheme is proved for various special cases.**

***Index Terms*—CEO problem, common information, distributed source coding, multiple description coding.**

I. Introduction

THERE are many situations in which data collected at different sites must be encoded separately and sent to a fusion center for subsequent processing. In their landmark paper [1], Slepian and Wolf showed that it is possible to exploit the correlation between the data received at different sites even if the encoders are operated in a completely distributed manner. They characterized the achievable rate region for the case where the fusion center (or the decoder) must reproduce separately encoded memoryless sources with arbitrarily small error probability. Their result was further generalized by Cover [2] via his renowned random binning argument. There have been many attempts to extend the Slepian–Wolf coding problem by considering general distortion criteria on the source reconstruction. Wyner and Ziv [3] derived the rate–distortion function for the case in which the decoder has access to side information. A general distributed source coding problem was formulated by Berger [4] and Tung [5]. However, except for certain special cases [6]–[9], the problem is still widely open. A variant of the distributed source coding problem, often referred to as the CEO problem, has received considerable attention in recent years [10]–[16]. In particular, for the quadratic Gaussian case, the rate–distortion region has been completely characterized by Oohama [17] and Prabhakaran *et al.* [18].

Multiple description coding is another important class of source coding problems. In the multiple description problem, the encoder constructs several descriptions of the source and sends them separately through unreliable links; the decoder tries to reconstruct the source given any subset of the descriptions, and the reconstruction quality may depend on the specific subset that has been received (or equivalently, we can associate each description with an encoder and each subset of descriptions with a decoder). Early contributions to this problem can be found in [19]–[22]. The first general result was by El Gamal and Cover (EGC) on the two-description achievable rate–distortion region [23]. The EGC region was shown to be tight for the quadratic Gaussian case by Ozarow [21] and for the no excess sum–rate case by Ahlswede [24]. Further results can be found in [25]–[31].

Distributed source coding problems of the Slepian–Wolf type and their extensions emphasize compression efficiency but ignore system robustness. A distributed source coding scheme which is optimal in the sense of compression efficiency can be very sensitive to the encoder failure, i.e., the performance of the whole system may degenerate dramatically when one of the encoders is subject to a failure. In contrast, the multiple description problem does consider system robustness. However, it is a centralized source coding problem and its coding schemes in general cannot be applied in the distributed source coding scenario. Therefore, it is of considerable interest to study robust distributed source coding schemes that are able to trade off between two important parameters: system robustness and compression efficiency.

The remainder of this paper is organized as follows. In Section II, we describe the system model and the problem formulation. Some motivations are given in Section III. In Section IV, we first consider two different scenarios, namely, the distributed source coding scenario and the centralized source coding scenario, for which the corresponding coding schemes are given. Then we propose a unified approach by developing a coding scheme based on the idea of common information. The quadratic Gaussian case is studied in Section V. The inner bound and the outer bound of the rate–distortion region are derived, which coincide in several special cases. We conclude the paper in Section VI.

II. System Model and Problem Formulation

Consider the distributed source coding system shown in Fig. 1. Let $\{X(t),Y_1(t),Y_2(t)\}_{t=1}^{\infty}$ be a memoryless source with zero-order joint probability distribution $P_{XY_1Y_2}$ on $\mathcal{X}\times\mathcal{Y}_1\times\mathcal{Y}_2$, where $\mathcal{X}$ is the common alphabet of the random variables $X(t)$ for $t=1,2,\cdots$, and $\mathcal{Y}_i$ $(i=1,2)$ is the common alphabet of the random variables $Y_i(t)$ for $t=1,2,\cdots$. The target data sequence $\{X(t)\}_{t=1}^{\infty}$ cannot be observed directly. Instead, two corrupted versions of $\{X(t)\}_{t=1}^{\infty}$, i.e., $\{Y_1(t)\}_{t=1}^{\infty}$ and $\{Y_2(t)\}_{t=1}^{\infty}$, are observed by encoder 1 and encoder 2, respectively. Encoder i encodes a block $y_i^n=[y_i(1),\ldots,y_i(n)]$ of length n from its observed data by using a source code $c_i^{(n)}=f_{E,i}^{(n)}(y_i^n)$ of rate $\frac{1}{n}\log|\mathcal{C}_i^{(n)}|$, $i=1,2$. Decoder i reconstructs the target sequence $x^n=[x(1),\ldots,x(n)]$ by implementing a mapping $f_{D,i}^{(n)}:\mathcal{C}_i^{(n)}\rightarrow\mathcal{X}^n$, $i=1,2,3$, where $\mathcal{C}_3^{(n)}=\mathcal{C}_1^{(n)}\times\mathcal{C}_2^{(n)}$.

Manuscript received April 23, 2006; revised March 12, 2008.

J. Chen is with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON L8S 4K1, Canada (e-mail: junchen@ece.mcmaster.ca).

T. Berger is with the Department of Electrical and Computer Engineering, University of Virginia, Charlottesville, VA 22904 USA (e-mail: tb6n@virginia.edu).

Communicated by M. Effros, Associate Editor for Source Coding.

Digital Object Identifier 10.1109/TIT.2008.926389

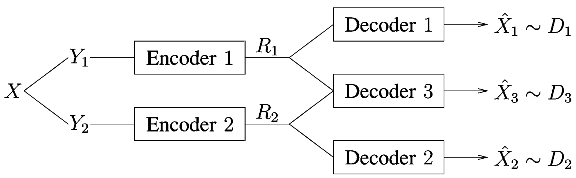

Fig. 1. Robust distributed source coding system.

Definition 1: The quintuple $(R_1, R_2, D_1, D_2, D_3)$ is said to be achievable if for any $\epsilon > 0$ and all sufficiently large n, there exist encoders

$$f_{E,i}^{(n)} : \mathcal{Y}_i^n \to \mathcal{C}_i^{(n)} \quad \log|\mathcal{C}_i^{(n)}| \le n(R_i + \epsilon), \qquad i = 1, 2$$

and decoders

$$f_{D,i}^{(n)} : \mathcal{C}_i^{(n)} \to \mathcal{X}^n, \qquad i = 1, 2$$
$$f_{D,3}^{(n)} : \mathcal{C}_1^{(n)} \times \mathcal{C}_2^{(n)} \to \mathcal{X}^n$$

such that for $\hat{X}_i^n = f_{D,i}^{(n)}(f_{E,i}^{(n)}(Y_i^n))$, $i = 1, 2$, and for $\hat{X}_3^n = f_{D,3}^{(n)}(f_{E,1}^{(n)}(Y_1^n), f_{E,2}^{(n)}(Y_2^n))$

$$\frac{1}{n}\mathbb{E}\sum_{t=1}^{n} d(X(t), \hat{X}_i(t)) \le D_i + \epsilon, \qquad i = 1, 2, 3.$$

Here $d(\cdot,\cdot) : \mathcal{X} \times \mathcal{X} \to [0, d_{\max}]$ is a given distortion measure. Let $\mathcal{Q}$ denote the set of all achievable quintuples.

Remark: More generally, one can allow different distortion measures at different decoders. We choose the current definition mainly to simplify notation.

Definition 2: Let $D_{\max} = \min_{x_0 \in \mathcal{X}} \mathbb{E}d(X, x_0)$ and

$$D_1^*(R_1, R_2) = \min\{D_1 : (R_1, R_2, D_1, D_{\max}, D_{\max}) \in \mathcal{Q}\}$$
$$D_2^*(R_1, R_2) = \min\{D_2 : (R_1, R_2, D_{\max}, D_2, D_{\max}) \in \mathcal{Q}\}$$
$$D_3^*(R_1, R_2) = \min\{D_3 : (R_1, R_2, D_{\max}, D_{\max}, D_3) \in \mathcal{Q}\}.$$

Since $D_1^*(R_1, R_2)$ does not depend on R_2 and $D_2^*(R_1, R_2)$ does not depend on R_1, we shall denote them by $D_1^*(R_1)$ and $D_2^*(R_2)$, respectively. Note that $D_i^*(R_i)$ $(i = 1, 2)$ is the distortion–rate function of the remote source coding problem [32], [33] with $\{X(t)\}_{t=1}^{\infty}$ as the hidden source and $\{Y_i(t)\}_{t=1}^{\infty}$ as the noisy obervation. Specifically, we have

$$D_i^*(R_i) = \min_{P_{\hat{X}_i|Y_i}: X - Y_i - \hat{X}_i, I(Y_i;\hat{X}_i) \le R_i} \mathbb{E}d(X, \hat{X}_i), \quad i = 1, 2.$$

Correspondingly, the rate–distortion function of the remote source coding problem is defined as

$$R_i^*(D_i) = \min_{P_{\hat{X}_i|Y_i}: X - Y_i - \hat{X}_i, \mathbb{E}d(X,\hat{X}_i) \le D_i} I(Y_i; \hat{X}_i), \quad i = 1, 2.$$

Here $X - Y_i - \hat{X}_i$ means X, Y_i, and $\hat{X}_i$ form a Markov chain, i.e., X and $\hat{X}_i$ are independent conditioned on Y_i.

Our model was first introduced by Ishwar *et al.* in [34]. An analogous problem called multilevel diversity coding has been studied in [35]–[38], which is a centralized source coding problem since all the encoders have the same observation. A distributed version of multilevel diversity coding was introduced in [39], where only the lossless case was treated.

III. Motivations

Our problem reduces to the CEO problem if $\min(D_1, D_2) \ge D_{\max}$ and reduces to the multiple description problem if there exist deterministic functions $f_1(\cdot)$ and $f_2(\cdot)$ such that

$$X(t) = f_1(Y_1(t)) = f_2(Y_2(t))$$

with probability one for $t = 1, 2, \cdots$. So it is instructive to review the coding schemes for the CEO problem and the multiple description problem.

For the CEO problem, the distortion constraint is only imposed on the reconstruction of the target sequence at decoder 3. The largest known achievable rate–distortion region[1] for the CEO problem is the set of (R_1, R_2, D_3) for which there exist random variables (W_1, W_2) jointly distributed with the generic source variables (X, Y_1, Y_2) such that

1) $W_1 - Y_1 - (X, Y_2, W_2)$ and $W_2 - Y_2 - (X, Y_1, W_1)$;
2) $R_1 \ge I(Y_1; W_1|W_2), R_2 \ge I(Y_2; W_2|W_1), R_1 + R_2 \ge I(Y_1, Y_2; W_1, W_2)$;
3) there exists a function $f : \mathcal{W}_1 \times \mathcal{W}_2 \to \mathcal{X}$ such that $\mathbb{E}d(X, \hat{X}) \le D$, where $\hat{X} = f(W_1, W_2)$.

The proof of the achievability of this rate–distortion region is based on the idea of random binning. The main feature of the random binning scheme is outlined as follows: there are many bins at each encoder and many codewords in each bin; instead of directly sending the codeword, each encoder sends the index of the bin which contains the codeword that this encoder wants to convey to the decoder; upon receiving the indices of bins from all the encoders, the decoder picks one codeword from each bin such that these codewords are jointly typical.

There are two important parameters for each encoder: the number of bins and the number of codewords. Roughly speaking, the number of bins determines the rate of the encoder while the number of codewords is associated with the description ability of the encoder. When the system is optimized in the sense of compression efficiency, the number of bins is minimized at each encoder if the number of its codewords is fixed (or equivalently, the number of codewords is maximized at each encoder if the number of bins at that encoder is fixed). Note that there exists a tradeoff between the maximum number of codewords at different encoders if the number of bins is fixed at each encoder (or, equivalently, a tradeoff between the minimum number of bins at different encoders if the number of codewords is fixed at each encoder). But intuitively this optimization is achieved at the price of sacrificing the robustness of the whole system: if the decoder only receives the data from one of the encoders, then it may not be able to recover the correct codeword since the decoder only gets a bin index from one encoder and there are many codewords in that bin. In general, we can improve the robustness of the distributed source coding system by reducing the number of codewords in each bin, which is a way to trade compression efficiency for system robustness. This is essentially the main idea of the robust distributed source coding scheme proposed in [34].

[1] By a timesharing argument, the convex hull of this region is also achievable.

Now we proceed to discuss multiple description coding. The multiple description problem has been studied for years and many coding schemes have been proposed. The common feature of the existing multiple description coding schemes is as follows: encoder i ($i=1,2$), sends a vector, say $(C_{i,1}^{(n)}, C_{i,2}^{(n)})$; decoder i ($i=1,2$) can only decode the $C_{i,1}^{(n)}-$ part; decoder 3 can decode both $(C_{1,1}^{(n)}, C_{1,2}^{(n)})$ and $(C_{2,1}^{(n)}, C_{2,2}^{(n)})$. Clearly, this idea is also applicable in distributed source coding. In the next section, we propose a robust distributed coding scheme by combining the random binning technique and the ideas from multiple description coding.

IV. Achievable Rate–Distortion Regions

A. An Achievable Rate–Distortion Region

Theorem 1: $(R_1, R_2, D_1, D_2, D_3)$ is achievable if there exist random variables (U_1, U_2, W_1, W_2) jointly distributed with the generic source variables (X, Y_1, Y_2) such that the following conditions are satisfied:

1) $(U_1, W_1) - Y_1 - (X, Y_2, U_2, W_2)$ and $(U_2, W_2) - Y_2 - (X, Y_1, U_1, W_1)$;
2) $(R_1, R_2) \in \mathcal{R}(U_1, U_2, W_1, W_2)$, where

$$\begin{aligned}\mathcal{R}(U_1, U_2, W_1, W_2) &= \Big\{(R_1,R_2): R_1 \geq I(Y_1;U_1) + I(Y_1;W_1|U_1,U_2,W_2),\\ &\qquad R_2 \geq I(Y_2;U_2) + I(Y_2;W_2|U_1,U_2,W_1),\\ &\qquad R_1+R_2 \geq I(Y_1;U_1) + I(Y_2;U_2)\\ &\qquad\qquad + I(Y_1,Y_2;W_1,W_2|U_1,U_2)\Big\};\end{aligned}$$

3) there exist functions $f_i : \mathcal{U}_i \to \mathcal{X}$ ($i=1,2$) and $f_3 : \mathcal{U}_1 \times \mathcal{W}_1 \times \mathcal{U}_2 \times \mathcal{W}_2 \to \mathcal{X}$ such that $\mathbb{E}d(X, \hat{X}_i) \leq D_i$, $i=1,2,3$, where $\hat{X}_1 = f_1(U_1)$, $\hat{X}_2 = f_2(U_2)$ and $\hat{X}_3 = f_3(U_1, W_1, U_2, W_2)$.

If $\mathcal{C}$ denotes the set of these achievable quintuples, then time sharing yields that $\mathcal{Q}_{\rm in} \triangleq \mathrm{conv}(\mathcal{C})$ is also an achievable region, i.e., $\mathcal{Q}_{\rm in} \subseteq \mathcal{Q}$.

Proof: See Appendix I. □

Remark:

1) One can readily show by invoking the support lemma [40, p. 310] that $\mathcal{U}_1$ must have $|\mathcal{Y}_1|-1$ letters to preserve the probability distribution P_{Y_1} and five more to preserve

$$\begin{aligned}&I(Y_1;U_1) + I(Y_1;W_1|U_1,U_2,W_2),\\ &I(Y_2;W_2|U_1,U_2,W_1),\\ &I(Y_1;U_1) + I(Y_1,Y_2;W_1,W_2|U_1,U_2),\end{aligned}$$

D_1, and D_3; $\mathcal{W}_1$ must have $|\mathcal{Y}_1||\mathcal{U}_1|-1$ letters to preserve the probability distribution $P_{Y_1U_1}$ and four more to preserve $I(Y_1;W_1|U_1,U_2,W_2)$, $I(Y_2;W_2|U_1,U_2,W_1)$, $I(Y_1,Y_2;W_1,W_2|U_1,U_2)$, and D_3. Therefore, it suffices to have $|\mathcal{U}_1| = |\mathcal{Y}_1|+4$ and

$$|\mathcal{W}_1| = |\mathcal{Y}_1||\mathcal{U}_1| - 1 + 4 = |\mathcal{Y}_1|^2 + 4|\mathcal{Y}_1| + 3.$$

Similarly, it suffices to have $|\mathcal{U}_2| = |\mathcal{Y}_2|+4$ and $|\mathcal{W}_2| = |\mathcal{Y}_2|^2 + 4|\mathcal{Y}_2| + 3$.

2) It is easy to verify that $\mathcal{R}(U_1, U_2, W_1, W_2)$ is a contra-polymatroid.[2]
3) Let $W_i' = (U_i, W_i)$, $i=1,2$. It is easy to see that $\mathcal{R}(U_1, U_2, W_1', W_2') = \mathcal{R}(U_1, U_2, W_1, W_2)$. Therefore, there is no loss of generality in assuming $U_i - W_i - Y_i$ and defining $f_3(\cdot)$ on $\mathcal{W}_1 \times \mathcal{W}_2$.

A counterexample constructed by Körner and Marton [42] shows that $\mathcal{Q}_{\rm in} \subsetneqq \mathcal{Q}$ in general. Actually even some special cases of our problem such as the multiple description problem and the CEO problem are longstanding open problems. However, a conclusive result can be obtained for the following case.

Corollary 1: We have

$$\begin{aligned}\min\{R_1&: \text{there exists } R_2 \text{ such that } (R_1,R_2,D_1,D_{\max},D_3)\in\mathcal{Q}\}\\ &= \min[I(Y_1;U_1) + I(Y_1;W_1|Y_2,U_1)]\end{aligned}$$

for any D_1 and D_3, where the second minimization is over the set of random variables (U_1, W_1) jointly distributed with the generic source variables (X, Y_1, Y_2) such that the following conditions are satisfied:

1) $(U_1, W_1) - Y_1 - (X, Y_2)$;
2) there exist functions $f: \mathcal{U}_1 \to \mathcal{X}$, $g: \mathcal{U}_1 \times \mathcal{W}_1 \times \mathcal{Y}_2 \to \mathcal{X}$ such that $\mathbb{E}d(X, f(U_1)) \leq D_1$, $\mathbb{E}d(X, g(U_1, W_1, Y_2)) \leq D_3$;
3) $|\mathcal{U}_1| = |\mathcal{Y}_1|+2$, $|\mathcal{W}_1| = (|\mathcal{Y}_1|+1)^2$.

Proof: Since here we are concerned about minimizing R_1 under the distortion constraints D_1 and D_3, there is no loss of generality in assuming that R_2 is large enough so that $\{Y_2(t)\}_{t=1}^{\infty}$ can be recovered losslessly at decoder 3. In this case, our problem becomes the remote Heegard–Berger problem. Its direct coding theorem can be easily reduced from Theorem 1 while the converse coding theorem can be proved along the same line as that in [43]. □

B. Remote Multiple Description Coding

If there exist $f_1(\cdot)$ and $f_2(\cdot)$ such that $Y = f_1(Y_1) = f_2(Y_2)$ with probability one and $X - Y - (Y_1, Y_2)$, our problem becomes the remote multiple description problem. In this case, the existing multiple description coding scheme can be adopted after a slight modification.

Theorem 2:

1) $(R_1, R_2, D_1, D_2, D_3)$ is achievable if there exist random variables $(\hat{X}_0, \hat{X}_1, \hat{X}_2, \hat{X}_3)$ jointly distributed with the generic source variables (X, Y) such that the following properties are satisfied:
   a) $X - Y - (\hat{X}_0, \hat{X}_1, \hat{X}_2, \hat{X}_3)$;
   b) $R_1 + R_2 \geq 2I(Y;\hat{X}_0) + I(\hat{X}_1;\hat{X}_2|\hat{X}_0) + I(Y;\hat{X}_1,\hat{X}_2,\hat{X}_3|\hat{X}_0)$, $R_i \geq I(Y;\hat{X}_0,\hat{X}_i)$, $i=1,2$;
   c) $\mathbb{E}d(X, \hat{X}_i) \leq D_i$, $i=1,2,3$.

   If $\mathcal{C}'$ denotes the set of these achievable quintuples, then time sharing yields that $\mathrm{conv}(\mathcal{C}')$ is also an achievable region.

[2] See [41] for the definition of contra-polymatroid.

2) Let $\mathcal{C}^*$ denote the subset of $\mathcal{C}'$ containing all those quintuples satisfying a)-c) and the additional constraints that $\hat{X}_1$ is independent of $\hat{X}_2$, and $\hat{X}_0$ is a constant. Define

$$R^*(D) = \min_{P_{\hat{X}|Y}: X-Y-\hat{X}, \mathbb{E}d(X,\hat{X}) \le D} I(Y;\hat{X})$$

$$\mathcal{Q}(D_3) = \{(R_1,R_2,D_1,D_2,D_3) \in \mathcal{Q} : R_1 + R_2 = R^*(D_3)\}$$

$$\mathrm{conv}(\mathcal{C}^*)(D_3) = \{(R_1,R_2,D_1,D_2,D_3) \in \mathrm{conv}(\mathcal{C}^*) : R_1 + R_2 = R^*(D_3)\}.$$

We have

$$\mathcal{Q}(D_3) = \mathrm{conv}(\mathcal{C}^*)(D_3).$$

Proof: Part 1) follows from the Markov lemma and Theorem 1 (specialized to the two-encoder case) in [29]. Part 2) can be proved via a continuity argument similar to that in [24] by replacing Shannon's rate–distortion function $R(D_3)$ with $R^*(D_3)$ and noticing the following Markov relation:

$$X(t) - Y(t) - (Y_1^n, Y_2^n) - (f_{E,1}^{(n)}(Y_1^n), f_{E,2}^{(n)}(Y_2^n)) - (\hat{X}_1(t), \hat{X}_2(t), \hat{X}_3(t)). \quad \square$$

C. A Unified Approach

Theorem 1 is associated with a distributed source coding scheme while Theorem 2 is associated with a centralized source coding scheme. For a centralized coding system, since the two encoders have the same inputs, one knows exactly the operation of the other and thus they can have arbitrary cooperation. In this sense, the encoders in a centralized coding system should be viewed as the different functionalities of a single encoder. For a distributed coding system, since the two encoders have different inputs, one does not know for sure about the operation of the other. As a consequence, the types of cooperation between the two encoders in a distributed system are very limited. On the other hand, since centralized coding can be viewed as a special case of distributed coding, one would expect a unified approach to both of them. However, it is easy to see that Theorem 1, when applied to the centralized case (i.e., $Y_1 = Y_2 = Y$ with probability one), does not coincide with Theorem 2. That is to say, Theorem 2 is not a "centralized" version of Theorem 1. Now it is natural to ask whether there exists a distributed source coding scheme that subsumes the centralized source coding scheme in Theorem 2 as a special case.

We shall suggest a unified approach which incorporates these two schemes in a single framework. The main ingredient is a concept called the common part (or the common information) of two dependent random variables in the sense of Gács and Körner [44] and Witsenhausen [45]. The following definition is quoted from [46].

Definition 3: The common part Z of two random variables Y_1 and Y_2 is defined by finding the maximum integer k such that there exist functions $f : \mathcal{Y}_1 \to \{1,2,\cdots,k\}$ and $g : \mathcal{Y}_2 \to \{1,2,\cdots,k\}$ with $\Pr(f(Y_1) = i) > 0$, $\Pr(g(Y_2) = i) > 0$, $i = 1,2,\cdots,k$, such that $f(Y_1) = g(Y_2)$ with probability one and then defining $Z = f(Y_1)$.

With this definition, it is obvious that encoder 1 and encoder 2 can agree on the value of Z with probability one. Therefore, they can use an efficient centralized coding scheme (of Theorem 2 type) for the common part Z and then superimpose a distributed coding scheme (of Theorem 1 type). This observation immediately leads to the following theorem.

Theorem 3: Let Z be the common part of Y_1 and Y_2. The quintuple (R_1,R_2,D_1,D_2,D_3) is achievable if there exist random variables $(U_1,U_2,W_1,W_2,Z_0,Z_1,Z_2,Z_3)$ jointly distributed with the generic source variables (X,Y_1,Y_2,Z) such that the following conditions are satisfied:

1) $(X,Y_1,Y_2) - Z - (Z_0,Z_1,Z_2)$;
2) $U_1 - (Y_1,Z_0,Z_1) - (X,Y_2,Z_2,U_2)$ and $U_2 - (Y_1,Z_0,Z_2) - (X,Y_1,Z_1,U_1)$;
3) $Z_3 - (Z,Z_0,Z_1,Z_2) - (X,Y_1,Y_2,U_1,U_2)$;
4) $W_1 - (Y_1,Z_0,Z_1,Z_2,Z_3,U_1) - (X,Y_2,U_2,W_2)$ and $W_2 - (Y_2,Z_0,Z_1,Z_2,Z_3,U_2) - (X,Y_1,U_1,W_1)$;
5) $(R_1,R_2) \in \mathcal{R}(U_1,U_2,W_1,W_2,Z_0,Z_1,Z_2,Z_3)$, where $\mathcal{R}(U_1,U_2,W_1,W_2,Z_0,Z_1,Z_2,Z_3)$ is the following set:

$$\begin{aligned}\{(R_1,R_2) : R_1 &\ge I(Y_1;Z_0,Z_1,U_1) \\ &\quad + I(Y_1;W_1|Z_0,Z_1,Z_2,Z_3,U_1,U_2,W_2), \\ R_2 &\ge I(Y_2;Z_0,Z_2,U_2) \\ &\quad + I(Y_2;W_2|Z_0,Z_1,Z_2,Z_3,U_1,U_2,W_1), \\ R_1 + R_2 &\ge 2I(Z;Z_0) + I(Z_1;Z_2|Z_0) \\ &\quad + I(Z;Z_1,Z_2,Z_3|Z_0) \\ &\quad + I(Y_1;U_1|Z_0,Z_1) + I(Y_2;U_2|Z_0,Z_2) \\ &\quad + I(Y_1,Y_2;W_1,W_2|Z_0,Z_1,Z_2,Z_3,U_1,U_2)\};\end{aligned}$$

6) there exist functions: $f_i : \mathcal{U}_i \to \mathcal{X}$, $i = 1,2$, and $f_3 : \mathcal{U}_1 \times \mathcal{W}_1 \times \mathcal{U}_2 \times \mathcal{W}_2 \to \mathcal{X}$ such that $\mathbb{E}d(X,\hat{X}_i) \le D_i$, $i = 1,2,3$, where $\hat{X}_1 = f_1(U_1)$, $\hat{X}_2 = f_2(U_2)$, and $\hat{X}_3 = f_3(U_1,W_1,U_2,W_2)$.

If $\mathcal{C}''$ denotes the set of these achievable quintuples, then time sharing yields that $\mathrm{conv}(\mathcal{C}'')$ is also an achievable region.

Proof: This result can be proved by combining the ideas from Theorem 1 and Theorem 2. The details are omitted. $\square$

Remark:

1) Theorem 3 can be reduced to Theorem 1 by setting $(Z_0,Z_1,Z_2,Z_3) =$ constant. If $X - Z - (Y_1,Y_2)$, then Theorem 3 can be specialized to Theorem 2 by setting $(U_1,U_2,W_1,W_2) =$ constant and noticing that there is no loss of generality in requiring Z_1,Z_2,Z_3 to assume values in $\mathcal{X}$.
2) The conventional distributed source coding scheme [4], [5] does not consider the common part (even if it does exist) of the observations at different encoders and thus requires very restricted long Markov chain conditions on the auxiliary random variables. As we have seen in Theorem 3, the long Markov chain conditions are not always necessary, at least in the case when the observations at different encoders share a common part.
3) Theorem 3 essentially suggests an approach to bridging the distributed source coding scheme and the centralized source coding scheme. However, it is possible that Y_1 and Y_2 do not share any common part even when they are highly correlated. Hence, it is of special interest to see

whether there exists a general coding scheme that can transition smoothly from a distributed scheme to a centralized scheme when Y_1 and Y_2 become more and more correlated but no common part exists.

V. Gaussian Case

In this section, we apply the general results obtained in the previous section to analyze the Gaussian case with squared distortion measure. Although most of the results in Section IV are proved for the finite alphabet case with bounded distortion measure, they can be extended to the Gaussian case with squared distortion measure by standard techniques [8], [47].

Let $\{X(t), Y_1(t)=X(t)+N_1(t), Y_2(t)=X(t)+N_2(t)\}_{t=1}^{\infty}$ be a sequence of independent and identically distributed (i.i.d.) zero-mean Gaussian vectors, where $X(t)$, $N_1(t)$ and $N_2(t)$ are mutually independent with variances $\sigma_X^2, \sigma_{N_1}^2$, and $\sigma_{N_2}^2$, respectively.[3] Let

$$D_{i,\min} = \mathrm{var}(X(t)|Y_i(t)) = \left(1/\sigma_X^2 + 1/\sigma_{N_i}^2\right)^{-1}, \quad i=1,2$$
$$D_{3,\min} = \mathrm{var}(X(t)|Y_1(t),Y_2(t)) = \left(1/\sigma_X^2 + 1/\sigma_{N_1}^2 + 1/\sigma_{N_2}^2\right)^{-1}.$$

Without loss of generality, we shall focus on the region $\{(R_1,R_2,D_1,D_2,D_3) \in \mathcal{Q} : D_3 \le \min(D_1,D_2), D_{i,\min} \le D_i \le \sigma_X^2, i=1,2,3\}$. For convenience, we denote this region by $\mathcal{Q}$.

A. An Inner Bound of the Rate–Distortion Region

We shall derive an inner bound of the rate–distortion region for the quadratic Gaussian case by evaluating the achievable region in Theorem 1. Let W_1, U_1, W_2, U_2 be the auxiliary random variables jointly distributed with the generic source variables X, Y_1, Y_2 such that $U_1 = Y_1 + T_{11}$, $U_2 = Y_2 + T_{21}$, $W_1 = Y_1 + T_{12}$, and $W_2 = Y_2 + T_{22}$. Here $T_{11}, T_{12}, T_{21}, T_{22}$ are zero-mean Gaussian random variables with variances $\sigma_{T_{11}}^2, \sigma_{T_{12}}^2, \sigma_{T_{21}}^2, \sigma_{T_{22}}^2$, respectively, and they are independent of X, Y_1, Y_2; moreover, T_{11}, T_{12} are independent of T_{21}, T_{22}.

Let $W_i^* = \mathbb{E}(Y_i|U_i,W_i)$, $i=1,2$. It is easy to verify that

$$\mathcal{R}(U_1,U_2,W_1,W_2) = \mathcal{R}(U_1,U_2,W_1^*,W_2^*)$$

and

$$\begin{aligned}\mathbb{E}(X-\mathbb{E}(X|W_1,W_2,U_1,U_2))^2 &= \mathbb{E}(X-\mathbb{E}(X|W_1^*,W_2^*,U_1,U_2))^2\\ &= \mathbb{E}(X-\mathbb{E}(X|W_1^*,W_2^*))^2.\end{aligned}$$

Therefore, there is no loss of generality in assuming $U_i - W_i - Y_i$ ($i=1,2$), i.e., we can assume $T_{i1} = T_{i2} + \Delta T_i$ ($i=1,2$), where $\Delta T_1 \sim \mathcal{N}(0, \sigma_{T_{11}}^2 - \sigma_{T_{12}}^2)$ and $\Delta T_2 \sim \mathcal{N}(0, \sigma_{T_{21}}^2 - \sigma_{T_{22}}^2)$ are mutually independent, and they are independent of $X, Y_1, Y_2, T_{12}, T_{22}$.

Now it follows from Theorem 1 that the following region is achievable:

$$\mathcal{Q}_{\mathrm{in}} = \mathrm{conv}\left(\bigcup_{\sigma_{T_{11}}^2 \ge \sigma_{T_{12}}^2, \sigma_{T_{21}}^2 \ge \sigma_{T_{22}}^2} \mathcal{C}\left(\sigma_{T_{11}}^2, \sigma_{T_{12}}^2, \sigma_{T_{21}}^2, \sigma_{T_{22}}^2\right)\right)$$

[3]Unless specified otherwise, $\sigma_X^2, \sigma_{N_1}^2$, and $\sigma_{N_2}^2$ are assumed to be positive.

where

$$\begin{aligned}&\mathcal{C}\left(\sigma_{T_{11}}^2, \sigma_{T_{12}}^2, \sigma_{T_{21}}^2, \sigma_{T_{22}}^2\right)\\ &= \Bigg\{(R_1,R_2,D_1,D_2,D_3):\\ &\quad \frac{1}{D_i} \le \frac{1}{\sigma_X^2} + \frac{1}{\sigma_{N_i}^2 + \sigma_{T_{i1}}^2}, i=1,2,\\ &\quad \frac{1}{D_3} \le \frac{1}{\sigma_X^2} + \frac{1}{\sigma_{N_1}^2 + \sigma_{T_{12}}^2} + \frac{1}{\sigma_{N_2}^2 + \sigma_{T_{22}}^2},\\ &\quad R_1 \ge \frac{1}{2}\log\frac{\sigma_{U_1}^2\left(\sigma_{W_1}^2\sigma_{W_2}^2 - \sigma_X^4\right)}{\sigma_{T_{12}}^2\left(\sigma_{U_1}^2\sigma_{W_2}^2 - \sigma_X^4\right)},\\ &\quad R_2 \ge \frac{1}{2}\log\frac{\sigma_{U_2}^2\left(\sigma_{W_1}^2\sigma_{W_2}^2 - \sigma_X^4\right)}{\sigma_{T_{22}}^2\left(\sigma_{W_1}^2\sigma_{U_2}^2 - \sigma_X^4\right)},\\ &\quad R_1+R_2 \ge \frac{1}{2}\log\frac{\sigma_{U_1}^2\sigma_{U_2}^2\left(\sigma_{W_1}^2\sigma_{W_2}^2 - \sigma_X^4\right)}{\sigma_{T_{12}}^2\sigma_{T_{22}}^2\left(\sigma_{U_1}^2\sigma_{U_2}^2 - \sigma_X^4\right)}\Bigg\}\end{aligned}$$

and

$$\sigma_{U_i}^2 = \sigma_X^2 + \sigma_{N_i}^2 + \sigma_{T_{i1}}^2, \sigma_{W_i}^2 = \sigma_X^2 + \sigma_{N_i}^2 + \sigma_{T_{i2}}^2, \quad i=1,2.$$

B. An Outer Bound of the Rate-Distortion Region

Let $\theta(t) = X(t) - S(t)$, $t=1,2,\cdots$, where

$$\begin{aligned}S(t) &= \mathbb{E}(X(t)|Y_1(t),Y_2(t))\\ &= \frac{D_{3,\min}}{\sigma_{N_1}^2}Y_1(t) + \frac{D_{3,\min}}{\sigma_{N_2}^2}Y_2(t)\end{aligned}$$

and $\theta(t) \sim \mathcal{N}(0, D_{3,\min})$ is independent of $Y_1(t)$ and $Y_2(t)$. Let $d_X = \sigma_X^2 - D_{3,\min}$ and $d_i = D_i - D_{3,\min}$, $i=1,2,3$. Define

$$\mathcal{Q}_{\mathrm{out}} = \bigcup_{(r_{11},r_{12},r_{21},r_{22})\in\Sigma_{\mathrm{out}}} \mathcal{C}_{\mathrm{out}}(r_{11},r_{12},r_{21},r_{22})$$

where $\mathcal{C}_{\mathrm{out}}(r_{11},r_{12},r_{21},r_{22})$, $\lambda\left(\sigma_X^2, D_1, D_2, D_3, r_{12}, r_{22}\right)$, ζ, and Σ_{out} are given in the expression at the top of the following page.

Theorem 4: $\mathcal{Q} \subseteq \mathcal{Q}_{\mathrm{out}}$.

Proof: This result can be proved by combining the techniques developed in [21], [17], [18]. The details are left to Appendix II. □

C. On the Tightness of $\mathcal{Q}_{\mathrm{in}}$ and $\mathcal{Q}_{\mathrm{out}}$

Note that $\{(R_1,R_2,D_3) : (R_1,R_2,\sigma_X^2,\sigma_X^2,D_3) \in \mathcal{Q}\}$ is the rate–distortion region of the quadratic Gaussian CEO problem. It is known [17], [18] that

$$\{(R_1,R_2,D_3) : (R_1,R_2,\sigma_X^2,\sigma_X^2,D_3) \in \mathcal{Q}\} = \bigcup_{r_{12}\ge 0, r_{22}\ge 0} \mathcal{C}_{\mathrm{CEO}}(r_{12},r_{22}) \quad (1)$$

where

$$\mathcal{C}_{\mathrm{CEO}}(r_{12},r_{22}) = \Bigg\{(R_1,R_2,D_3):$$
$$R_1 \ge r_{12} + \frac{1}{2}\log\frac{1}{D_3} - \frac{1}{2}\log\left(\frac{1}{\sigma_X^2} + \frac{1-\exp(-r_{22})}{\sigma_{N_2}^2}\right),$$

$$\mathcal{C}_{\text{out}}(r_{11},r_{12},r_{21},r_{22})=\Bigg\{(R_1,R_2,D_1,D_2,D_3):\frac{1}{D_i}\le\frac{\exp(2r_{i1})}{\sigma_X^2},R_i\ge r_{i1}+r_{i2},i=1,2,$$
$$\frac{1}{D_3}\le\frac{1}{\sigma_X^2}+\frac{1-\exp(-2r_{12})}{\sigma_{N_1}^2}+\frac{1-\exp(-2r_{22})}{\sigma_{N_2}^2},$$
$$r_{11}+r_{21}\ge\frac{1}{2}\log\frac{\sigma_X^2}{D_3}+\lambda(\sigma_X^2,D_1,D_2,D_3,r_{12},r_{22})\Bigg\}$$

$$\lambda\left(\sigma_X^2,D_1,D_2,D_3,r_{12},r_{22}\right)=\begin{cases}0, & \zeta\le d_1+d_2-d_X\\ \frac{1}{2}\log\frac{d_X\zeta}{d_1d_2}, & \zeta\ge\left(\frac{1}{d_1}+\frac{1}{d_2}-\frac{1}{d_X}\right)^{-1}\\ \frac{1}{2}\log\frac{(d_X-\zeta)^2}{(d_X-\zeta)^2-[\sqrt{(d_X-d_1)(d_X-d_2)}-\sqrt{(d_1-\zeta)(d_2-\zeta)}]^2}, & \text{otherwise}\end{cases}$$

$$\zeta=D_3D_{3,\min}\left(\frac{\exp(-2r_{12})}{\sigma_{N_1}^2}+\frac{\exp(-2r_{22})}{\sigma_{N_2}^2}\right)$$

and

$$\Sigma_{\text{out}}=\left\{(r_{11},r_{12},r_{21},r_{22})\in\mathbb{R}_+^4:\frac{1}{\sigma_X^2}\exp(2r_{i1})\le\frac{1}{\sigma_X^2}+\frac{1-\exp(-2r_{i2})}{\sigma_{N_i}^2},i=1,2\right\}.$$

$$R_2\ge r_{22}+\frac{1}{2}\log\frac{1}{D_3}-\frac{1}{2}\log\left(\frac{1}{\sigma_X^2}+\frac{1-\exp(-r_{12})}{\sigma_{N_1}^2}\right),$$
$$R_1+R_2\ge r_{12}+r_{22}+\frac{1}{2}\log\frac{\sigma_X^2}{D_3},$$
$$\frac{1}{D_3}\le\frac{1}{\sigma_X^2}+\frac{1-\exp(-2r_{12})}{\sigma_{N_1}^2}+\frac{1-\exp(-2r_{22})}{\sigma_{N_2}^2}\Bigg\}.$$

Corollary 2: $\mathcal{Q}_{\text{in}}$ and $\mathcal{Q}_{\text{out}}$ coincide for the quadratic Gaussian CEO problem, i.e.,

$$\begin{aligned}&\{(R_1,R_2,D_3):\left(R_1,R_2,\sigma_X^2,\sigma_X^2,D_3\right)\in\mathcal{Q}_{\text{in}}\}\\ &\quad=\{(R_1,R_2,D_3):\left(R_1,R_2,\sigma_X^2,\sigma_X^2,D_3\right)\in\mathcal{Q}\} &(2)\\ &\quad=\{(R_1,R_2,D_3):\left(R_1,R_2,\sigma_X^2,\sigma_X^2,D_3\right)\in\mathcal{Q}_{\text{out}}\}. &(3)\end{aligned}$$

Proof: Equation (2) follows from the fact

$$\bigcup_{\sigma_{T_{11}}^2=\sigma_{T_{21}}^2=\infty,\sigma_{T_{12}}^2,\sigma_{T_{22}}^2}\mathcal{C}\left(\sigma_{T_{11}}^2,\sigma_{T_{12}}^2,\sigma_{T_{21}}^2,\sigma_{T_{22}}^2\right)$$

is the rate–distortion region of the quadratic Gaussian CEO problem [17], [18].

Now we proceed to prove (3). Since $\mathcal{Q}\subseteq\mathcal{Q}_{\text{out}}$, it suffices to show that

$$\{(R_1,R_2,D_3):(R_1,R_2,\sigma_X^2,\sigma_X^2,D_3)\in\mathcal{Q}_{\text{out}}\}\subseteq\{(R_1,R_2,D_3):(R_1,R_2,\sigma_X^2,\sigma_X^2,D_3)\in\mathcal{Q}\}.$$

For any $(R_1,R_2,D_1,D_2,D_3)\in\mathcal{Q}_{\text{out}}$ with $D_1=D_2=\sigma_X^2$, we have

$$\begin{aligned}\zeta&=D_3D_{3,\min}\left(\frac{\exp(-2r_{12})}{\sigma_{N_1}^2}+\frac{\exp(-2r_{22})}{\sigma_{N_2}^2}\right)\\ &\le D_3D_{3,\min}\left(\frac{1}{\sigma_{N_1}^2}+\frac{1}{\sigma_{N_2}^2}\right)\\ &\le\sigma_X^2D_{3,\min}\left(\frac{1}{\sigma_{N_1}^2}+\frac{1}{\sigma_{N_2}^2}\right)\\ &=\sigma_X^2D_{3,\min}\left(\frac{1}{D_{3,\min}}-\frac{1}{\sigma_X^2}\right)\\ &=\sigma_X^2-D_{3,\min}\\ &=d_1+d_2-d_X\end{aligned}$$

which implies that $\lambda(\sigma_X^2,D_1,D_2,D_3,r_{12},r_{22})=0$. Therefore, we have

$$\begin{aligned}&\{(R_1,R_2,D_3):(R_1,R_2,\sigma_X^2,\sigma_X^2,D_3)\in\mathcal{Q}_{\text{out}}\}\\ &\quad=\bigcup_{(r_{11},r_{12},r_{21},r_{22})\in\Sigma_{\text{out}}}\mathcal{C}'_{\text{out}}(r_{11},r_{12},r_{21},r_{22}) \qquad (4)\end{aligned}$$

where

$$\mathcal{C}'_{\text{out}}(r_{11},r_{12},r_{21},r_{22})=\Bigg\{(R_1,R_2,D_3):$$
$$R_i\ge r_{i1}+r_{i2},\quad i=1,2,$$
$$r_{11}+r_{21}\ge\frac{1}{2}\log\frac{\sigma_X^2}{D_3},$$
$$\frac{1}{D_3}\le\frac{1}{\sigma_X^2}+\frac{1-\exp(-2r_{12})}{\sigma_{N_1}^2}+\frac{1-\exp(-2r_{22})}{\sigma_{N_2}^2}\Bigg\}.$$

Comparing (1) and (4), it is easy to see that $\{(R_1,R_2,D_3):(R_1,R_2,\sigma_X^2,\sigma_X^2,D_3)\in\mathcal{Q}_{\text{out}}\}\subseteq\{(R_1,R_2,D_3):(R_1,R_2,\sigma_X^2,\sigma_X^2,D_3)\in\mathcal{Q}\}$. Thus, the proof is complete. $\square$

Note that $\mathcal{Q}|_{\sigma_{N_1}^2=\sigma_{N_2}^2=0}$ is the rate–distortion region of the Gaussian multiple description problem. It was proved by Ozarow [21] that $(R_1,R_2,D_1,D_2,D_3)\in\mathcal{Q}|_{\sigma_{N_1}^2=\sigma_{N_2}^2=0}$ if and only if

$$R_i\ge\frac{1}{2}\log\frac{\sigma_X^2}{D_i},\qquad i=1,2$$
$$R_1+R_2\ge\frac{1}{2}\log\frac{\sigma_X^2}{D_3}+\gamma\left(\sigma_X^2,D_1,D_2,D_3\right)$$

where we define $\gamma(\sigma_X^2,D_1,D_2,D_3)$ in the expression at the bottom of the following page. Here we use the expression of $\mathcal{Q}|_{\sigma_{N_1}^2=\sigma_{N_2}^2=0}$ given in [23] with some minor corrections by

[48], [49]; it is well known that this expression is equivalent to the one derived by Ozarow.

The following corollary states that $\mathcal{Q}_{\text{out}}$ converges to $\mathcal{Q}|_{\sigma_{N_1}^2=\sigma_{N_2}^2=0}$ as $(\sigma_{N_1}^2,\sigma_{N_2}^2)\to(0,0)$.

Corollary 3: We have

1) $\mathcal{Q}_{\text{out}}\subseteq\mathcal{Q}|_{\sigma_{N_1}^2=\sigma_{N_2}^2=0}$;
2) for any interior point (R_1,R_2,D_1,D_2,D_3) of $\mathcal{Q}|_{\sigma_{N_1}^2=\sigma_{N_2}^2=0}$, there exists an $\epsilon>0$ such that

$$(R_1,R_2,D_1,D_2,D_3)\in\mathcal{Q}_{\text{out}}$$

whenever $\max(\sigma_{N_1}^2,\sigma_{N_2}^2)\leq\epsilon$.

Proof: Part 1) can be deduced from Corollary 5, which will be proved later. So we shall only prove part 2) here.

Given $\delta>0$, define

$$\mathcal{Q}_{\text{out}}^{(\delta)}=\bigcup_{(r_{11},r_{12},r_{21},r_{22})\in\Sigma_{\text{out}},r_{12}=r_{22}=\delta}\mathcal{C}_{\text{out}}(r_{11},r_{12},r_{21},r_{22}).$$

It is clear that $\mathcal{Q}_{\text{out}}^{(\delta)}\subseteq\mathcal{Q}_{\text{out}}$ for any $\delta>0$. Moreover, we have $(R_1,R_2,D_1,D_2,D_3)\in\mathcal{Q}_{\text{out}}^{(\delta)}$ if and only if there exists $(r_{11},r_{21})\in\mathbb{R}_+^2$ such that

$$\begin{aligned}
&\frac{1}{D_i}\leq\frac{\exp(2r_{i1})}{\sigma_X^2}, \qquad i=1,2,\\
&R_i\geq r_{i1}+\delta, \qquad i=1,2\\
&\frac{1}{D_3}\leq\frac{1}{\sigma_X^2}+\frac{1-\exp(-2\delta)}{\sigma_{N_1}^2}+\frac{1-\exp(-2\delta)}{\sigma_{N_2}^2} &&(5)\\
&r_{11}+r_{21}\geq\frac{1}{2}\log\frac{D_3}{\sigma_X^2}+\lambda(\sigma_X^2,D_1,D_2,D_3,\delta,\delta)\\
&\frac{1}{\sigma_X^2}\exp(2r_{i1})\leq\frac{1}{\sigma_X^2}+\frac{1-\exp(-2\delta)}{\sigma_{N_i}^2}, \qquad i=1,2. &&(6)
\end{aligned}$$

First let $(\sigma_{N_1}^2,\sigma_{N_2}^2)$ converge to $(0,0)$ with δ fixed. We have $D_{3,\min}\to0$ and $\zeta\to D_3\exp(-2\delta)$; moreover, constraints (5) and (6) become inactive. Then send δ to 0. It follows by continuity that $\lambda(\sigma_X^2,D_1,D_2,D_3,\delta,\delta)$ converges to $\gamma(\sigma_X^2,D_1,D_2,D_3)$. The proof is complete. □

In contrast, it can be verified that $\mathcal{Q}_{\text{in}}$ does not converge to $\mathcal{Q}|_{\sigma_{N_1}^2=\sigma_{N_2}^2=0}$ as $(\sigma_{N_1}^2,\sigma_{N_2}^2)\to(0,0)$. The reason is that in order to achieve $\mathcal{Q}|_{\sigma_{N_1}^2=\sigma_{N_2}^2=0}$, the quantization errors introduced by different encoders must be negatively correlated [49]; however, for $\mathcal{Q}_{\text{in}}$, we have $\mathbb{E}(T_{1i}T_{2j})=0$ $(i=1,2;j=1,2)$, i.e., the quantization errors of the two encoders are uncorrelated.[4] Since $\mathcal{Q}$ is inner semicontinuous with respect to $(\sigma_{N_1}^2,\sigma_{N_2}^2)$ [50], it implies that $\mathcal{Q}_{\text{in}}$ is not tight.

[4] More fundamentally, this difference is due to the fact that the single-letter long Markov chain constraint is required for the existing distributed source coding schemes while it is not needed for centralized schemes.

The following corollary shows that $\mathcal{Q}_{\text{in}}$ and $\mathcal{Q}_{\text{out}}$ coincide in some subregions.

Corollary 4: We have

$$\begin{aligned}
&\{(D_1,D_2,D_3):(R_1^*(D_1),R_2^*(D_2),D_1,D_2,D_3)\in\mathcal{Q}\}\\
&\quad=\left\{(D_1,D_2,D_3):\frac{1}{D_3}\leq\frac{1}{D_1}+\frac{1}{D_2}-\frac{1}{\sigma_X^2}\right\}.
\end{aligned}$$

Remark: It is known [16] that

$$R_i^*(D_i)=\frac{1}{2}\log\frac{\sigma_X^4}{D_i\sigma_X^2-\sigma_X^2\sigma_{N_i}^2+D_i\sigma_{N_i}^2}, \qquad i=1,2.$$

Proof: For any $(R_1,R_2,D_1,D_2,D_3)\in\mathcal{Q}_{\text{out}}$, we have $R_i\geq r_{i1}+r_{i2}$, $D_i\geq\sigma_X^2\exp(-2r_{i1})$, and

$$1/\sigma_X^2+(1-\exp(-2r_{i2}))/\sigma_{N_i}^2\geq\exp(2r_{i1})/\sigma_X^2, \quad i=1,2.$$

Thus, if $R_i=R_i^*(D_i)$, $i=1,2$, then

$$\begin{aligned}
r_{i1}&=\frac{1}{2}\log\frac{\sigma_X^2}{D_i}\\
r_{i2}&=\frac{1}{2}\log\frac{\sigma_X^2D_i}{D_i\sigma_X^2-\sigma_X^2\sigma_{N_i}^2+D_i\sigma_{N_i}^2}.
\end{aligned}$$

Therefore, we have

$$\begin{aligned}
\frac{1}{D_3}&\leq\frac{1}{\sigma_X^2}+\frac{1-\exp(-2r_{12})}{\sigma_{N_1}^2}+\frac{1-\exp(-2r_{22})}{\sigma_{N_2}^2}\\
&=\frac{1}{D_1}+\frac{1}{D_2}-\frac{1}{\sigma_X^2}.
\end{aligned}$$

Now we proceed to prove the achievability part. Given D_1 and D_2, choose $\sigma_{T_{i1}}^2=\sigma_{T_{i2}}^2$ $(i=1,2)$ such that

$$\frac{1}{D_i}=\frac{1}{\sigma_X^2}+\frac{1}{\sigma_{N_i}^2+\sigma_{T_{i1}}^2}, \qquad i=1,2.$$

It can be seen that if

$$\begin{aligned}
R_i&=\frac{1}{2}\log\frac{\sigma_X^2+\sigma_{N_i}^2+\sigma_{T_{i1}}^2}{\sigma_{T_{i1}}^2}\\
&=\frac{1}{2}\log\frac{\sigma_X^4}{D_i\sigma_X^2-\sigma_X^2\sigma_{N_i}^2+D_i\sigma_{N_i}^2}\\
&=R_i^*(D_i), \qquad i=1,2\\
\frac{1}{D_3}&\leq\frac{1}{\sigma_X^2}+\frac{1}{\sigma_{N_1}^2+\sigma_{T_{11}}^2}+\frac{1}{\sigma_{N_2}^2+\sigma_{T_{21}}^2}\\
&=\frac{1}{D_1}+\frac{1}{D_2}-\frac{1}{\sigma_X^2}
\end{aligned}$$

then $(R_1,R_2,D_1,D_2,D_3)\in\mathcal{Q}_{\text{in}}$. The proof is complete. □

$$\gamma(\sigma_X^2,D_1,D_2,D_3)=\begin{cases}0, & D_3\leq D_1+D_2-\sigma_X^2\\ \frac{1}{2}\log\frac{\sigma_X^2D_3}{D_1D_2}, & D_3\geq\left(\frac{1}{D_1}+\frac{1}{D_2}-\frac{1}{\sigma_X^2}\right)^{-1}\\ \frac{1}{2}\log\frac{(\sigma_X^2-D_3)^2}{(\sigma_X^2-D_3)^2-[\sqrt{(\sigma_X^2-D_1)(\sigma_X^2-D_2)}-\sqrt{(D_1-D_3)(D_2-D_3)}]^2}, & \text{otherwise.}\end{cases}$$

D. Gaussian Remote Heegard–Berger Problem

The following theorem provides a complete characterization of the rate–distortion region for the Gaussian remote Heegard–Berger problem (i.e., the case where $\{Y_2(t)\}_{t=1}^{\infty}$ is directly available at decoder 2 and decoder 3).

Theorem 5: Let

$$\mathcal{Q}_e = \{(R_1, D_1, D_2, D_3) : (R_1, \infty, D_1, D_2, D_3) \in \mathcal{Q}\}.$$

We have $(R_1, D_1, D_2, D_3) \in \mathcal{Q}_e$ if and only if $D_2 \geq D_{2,\min}$ and R_1 is defined as in the expression at the bottom of the page.

Proof: Since $R_2 = \infty$, we can assume that $\{Y_2(t)\}_{t=1}^{\infty}$ is directly available at decoder 2 and decoder 3. Hence, any $D_2 \geq D_{2,\min}$ is achievable. Now only (R_1, D_1, D_3) remains to be characterized. The achievability part follows directly by evaluating $\mathcal{Q}_{\text{in}}$ with $\sigma^2_{T_{21}} = \sigma^2_{T_{22}} = 0$. For the converse, it is clear that $R_1 \geq R_1^*(D_1)$, which resolves the case $D_3 > D_1\sigma^2_{N_2}/(D_1 + \sigma^2_{N_2})$. For the case $D_3 \leq D_1\sigma^2_{N_2}/(D_1 + \sigma^2_{N_2})$, the details are left to Appendix III. □

Remark: The converse cannot be reduced from $\mathcal{Q}_{\text{out}}$, which shows that the outer bound in not tight.

Theorem 5 implies that

$$D_3^*(R_1, \infty) = \frac{\sigma_X^4\sigma_{N_2}^4 \exp(-2R_1) + \sigma_X^2\sigma_{N_1}^2\sigma_{N_2}^2\left(\sigma_X^2 + \sigma_{N_2}^2\right)}{\left(\sigma_X^2 + \sigma_{N_2}^2\right)\left(\sigma_X^2\sigma_{N_1}^2 + \sigma_X^2\sigma_{N_2}^2 + \sigma_{N_1}^2\sigma_{N_2}^2\right)}$$

and $\min\{D_1 : (R_1, \infty, D_1, D_2, D_3^*(R_1, \infty)) \in \mathcal{Q}\} = \sigma_X^2$ for $D_2 \geq D_{2,\min}$. That is to say, for the Gaussian remote Heegard–Berger problem, if decoder 3 achieves the minimum D_3 for a given R_1, then it is impossible for decoder 1 to make a nontrivial estimation of $\{X(t)\}_{t=1}^{\infty}$.

E. Gaussian Remote Multiple Description Coding

Now consider the case when both encoder 1 and encoder 2 can observe $\{Y_1(t)\}_{t=1}^{\infty}$ and $\{Y_2(t)\}_{t=1}^{\infty}$ simultaneously. Clearly, the rate–distortion region of this problem (denoted by $\mathcal{Q}'$) is an outer bound of $\mathcal{Q}$.

If we assume encoder 1 and encoder 2 can only observe $\{S(t)\}_{t=1}^{\infty}$, i.e., $\{\mathbb{E}(X(t)|Y_1(t), Y_2(t))\}_{t=1}^{\infty}$, and let $\mathcal{Q}''$ be the rate–distortion region for this case, then clearly we have $\mathcal{Q}'' \subseteq \mathcal{Q}'$ since $\{S(t)\}_{t=1}^{\infty}$ can be computed from $\{Y_1(t)\}_{t=1}^{\infty}$ and $\{Y_2(t)\}_{t=1}^{\infty}$.

Theorem 6: $\mathcal{Q}' = \mathcal{Q}'' = \mathcal{T}$, where

$$\mathcal{T} = \left\{(R_1, R_2, D_1, D_2, D_3) : R_1 + R_2 \geq \frac{1}{2}\log\frac{d_X}{d_3} + \gamma(d_X, d_1, d_2, d_3), R_i \geq \frac{1}{2}\log\frac{d_X}{d_i}, i = 1, 2\right\}.$$

Proof: Let $\mathcal{Q}_S$ denote the multiple description rate–distortion region with $\{S(t)\}_{t=1}^{\infty}$ as the source. The desired result follows from the simple observation that $(R_1, R_2, D_1, D_2, D_3) \in \mathcal{Q}'$ (or $\mathcal{Q}''$) if and only if

$$(R_1, R_2, D_1 - D_{3,\min}, D_2 - D_{3,\min}, D_3 - D_{3,\min}) \in \mathcal{Q}_S$$

(i.e., $(R_1, R_2, d_1, d_2, d_3) \in \mathcal{Q}_S$). □

Remark: Theorem 6 continues to hold when $N_1(t)$ and $N_2(t)$ are correlated (with correlation coefficient ρ_N). In this case, we have

$$D_{3,\min} = \left(\frac{1}{\sigma_X^2} + \frac{\sigma_{N_1}^2 + \sigma_{N_2}^2 - 2\rho_N\sigma_{N_1}\sigma_{N_2}}{(1 - \rho_N^2)\sigma_{N_1}^2\sigma_{N_2}^2}\right)^{-1}.$$

Specifically, $D_{3,\min} = (1/\sigma_X^2 + 1/\sigma_N^2)^{-1}$ if $\sigma^2_{N_1} = \sigma^2_{N_2} = \sigma^2_N$ and $\rho_N = 1$; otherwise, $D_{3,\min} = 0$ if $\rho_N^2 = 1$.

Corollary 5: $\mathcal{Q}_{\text{out}} \subsetneqq \mathcal{T}$.

Proof: For any $(R_1, R_2, D_1, D_2, D_3) \in \mathcal{Q}_{\text{out}}$, we have $R_i \geq r_{i1} + r_{i2}$, $D_i \geq \sigma_X^2 \exp(-2r_{i1})$, and

$$1/\sigma_X^2 + (1 - \exp(-2r_{i2}))/\sigma^2_{N_i} \geq \exp(2r_{i1})/\sigma_X^2, \quad i = 1, 2.$$

Therefore

$$\begin{aligned} R_i &\geq r_{i1} - \frac{1}{2}\log\left(\frac{\sigma^2_{N_i}}{\sigma^2_X} - \frac{\sigma^2_{N_i}\exp(2r_{i1})}{\sigma^2_X} + 1\right) \\ &\geq \frac{1}{2}\log\frac{\sigma^2_X}{D_i} - \frac{1}{2}\log\left(\frac{\sigma^2_{N_i}}{\sigma^2_X} - \frac{\sigma^2_{N_i}}{D_i} + 1\right) \\ &= R_i^*(D_i), \qquad i = 1, 2. \end{aligned} \tag{7}$$

Note that $\frac{1}{2}\log\frac{d_X}{d_i}$ is the rate–distortion function of the remote source coding problem with $\{S(t)\}_{t=1}^{\infty}$ (or equivalently, $\{Y_1(t)\}_{t=1}^{\infty}$ and $\{Y_2(t)\}_{t=1}^{\infty}$) as the noisy observation and $\{X(t)\}_{t=1}^{\infty}$ as the hidden source. Clearly, we have

$$R_i^*(D_i) \geq \frac{1}{2}\log\frac{d_X}{d_i}, \qquad i = 1, 2. \tag{8}$$

Moreover, given any $(R_1, R_2, D_1, D_2, D_3) \in \mathcal{Q}_{\text{out}}$, we have

$$\begin{aligned} R_1 + R_2 &\geq r_{11} + r_{21} + r_{12} + r_{22} \\ &\geq r_{12} + r_{22} + \frac{1}{2}\log\frac{\sigma^2_X}{D_3} \\ &\quad + \lambda\left(\sigma^2_X, D_1, D_2, D_3, r_{12}, r_{22}\right) \\ &= r_{12} + r_{22} + \frac{1}{2}\log\frac{\sigma^2_X}{D_3} - \frac{1}{2}\log\frac{d_X}{\zeta} \\ &\quad + \frac{1}{2}\log\frac{d_X}{\zeta} + \lambda\left(\sigma^2_X, D_1, D_2, D_3, r_{12}, r_{22}\right) \end{aligned}$$

for some $(r_{12}, r_{22}) \in \mathbb{R}^2_+$ satisfying

$$\frac{1}{D_3} \leq \frac{1}{\sigma^2_X} + \frac{1 - \exp(-2r_{12})}{\sigma^2_{N_1}} + \frac{1 - \exp(-2r_{22})}{\sigma^2_{N_2}}. \tag{9}$$

$$R_1 \geq \begin{cases} \frac{1}{2}\log\frac{\sigma_X^4\sigma_{N_2}^4}{\left(\sigma^2_{N_2}+D_1\right)\left(D_3\sigma^2_X\sigma^2_{N_1}+D_3\sigma^2_X\sigma^2_{N_2}+D_3\sigma^2_{N_1}\sigma^2_{N_2}-\sigma^2_X\sigma^2_{N_1}\sigma^2_{N_2}\right)}, & D_3 \leq \frac{D_1\sigma^2_{N_2}}{D_1+\sigma^2_{N_2}} \\ R_1^*(D_1), & D_3 > \frac{D_1\sigma^2_{N_2}}{D_1+\sigma^2_{N_2}}. \end{cases}$$

Note that both $r_{12} + r_{22} + \frac{1}{2}\log\zeta$ and $\frac{1}{2}\log\frac{d_X}{\zeta} + \lambda(\sigma_X^2, D_1, D_2, D_3, r_{12}, r_{22})$ are monotone increasing functions of (r_{12}, r_{22}). Thus, for the optimization problem

$$\min_{(r_{12},r_{22})\in\mathbb{R}_+^2} r_{12}+r_{22}+\frac{1}{2}\log\frac{\sigma_X^2}{D_3}-\frac{1}{2}\log\frac{d_X}{\zeta} +\frac{1}{2}\log\frac{d_X}{\zeta}+\lambda(\sigma_X^2,D_1,D_2,D_3,r_{12},r_{22})$$

subject to constraint (9), the inequality in (9) can be replaced by equality, which yields

$$\begin{aligned}\zeta &= D_3 D_{3,\min}\left(\frac{\exp(-2r_{12})}{\sigma_{N_1}^2}+\frac{\exp(-2r_{22})}{\sigma_{N_2}^2}\right)\\ &= D_3 D_{3,\min}\left(\frac{1}{D_{3,\min}}-\frac{1}{D_3}\right)\\ &= d_3\end{aligned}$$

and $\lambda(\sigma_X^2, D_1, D_2, D_3, r_{12}, r_{22}) = \gamma(d_X, d_1, d_2, d_3)$. Therefore, we have

$$R_1+R_2 \geq \min_{(r_{12},r_{22})\in\mathcal{F}(D_3)} r_{12}+r_{22}+\frac{1}{2}\log\frac{\sigma_X^2}{D_3}+\gamma(d_X,d_1,d_2,d_3)$$

where

$$\mathcal{F}(D_3)=\left\{(r_{12},r_{22})\in\mathbb{R}_+^2 : \frac{1}{D_3} = \frac{1}{\sigma_X^2}+\frac{1-\exp(-2r_{12})}{\sigma_{N_1}^2}+\frac{1-\exp(-2r_{22})}{\sigma_{N_2}^2}\right\}.$$

Define

$$R_{\mathrm{CEO}}(D_3)=\min_{(r_{12},r_{22})\in\mathcal{F}(D_3)} r_{12}+r_{22}+\frac{1}{2}\log\frac{\sigma_X^2}{D_3}.$$

We have

$$R_1+R_2 \geq R_{\mathrm{CEO}}(D_3)+\gamma(d_X,d_1,d_2,d_3). \quad (10)$$

Note that $R_{\mathrm{CEO}}(D_3)$ is actually the sum–rate distortion function of the quadratic Gaussian CEO problem [16]–[18] with $\{X(t)\}_{t=1}^\infty$ as the hidden source and $(\{Y_1(t)\}_{t=1}^\infty, \{Y_2(t)\}_{t=1}^\infty)$ as the noisy observations at two separate encoders, which implies

$$R_{\mathrm{CEO}}(D_3) \geq \frac{1}{2}\log\frac{d_X}{d_3}. \quad (11)$$

It can be readily seen that $\mathcal{Q}_{\mathrm{out}} \subseteq \mathcal{T}$ by combining (7), (8), (10), and (11). To prove that $\mathcal{Q}_{\mathrm{out}}$ is strictly contained in $\mathcal{T}$, one just needs to notice that the inequalities in (8) and (11) are strict if $\sigma_{N_1}^2 > 0, \sigma_{N_2}^2 > 0$. □

F. On the Correlated Noise Case

It is worth noting that the assumption on the mutual independence of $N_1(t)$ and $N_2(t)$ can be relaxed. Indeed, the correlated noise case can sometimes be reduced to the uncorrelated case.

Example 1: If $0 < \mathbb{E}[N_1(t)N_2(t)] < \min(\sigma_{N_1}^2, \sigma_{N_2}^2)$, then we can write $N_i(t) = N'(t) + N_i'(t)$ $(i = 1, 2)$ such that $\sigma_{N'}^2 \triangleq \mathrm{var}(N'(t)) = \mathbb{E}[N_1(t)N_2(t)]$, $\mathbb{E}[N'(t)N_i'(t)] = 0$ $(i = 1, 2)$, and $\mathbb{E}[N_1'(t)N_2'(t)] = 0$. Let $\mathcal{Q}_{X'}$ be the rate–distortion region with $\{X'(t)\}_{t=1}^\infty$ as the hidden source and $(\{Y_1(t)\}_{t=1}^\infty, \{Y_2(t)\}_{t=1}^\infty)$ as the noisy observations at two separate encoders, where $X'(t) = X(t) + N'(t)$. Note that $Y_i(t) = X'(t) + N_i'(t)$ $(i = 1, 2)$, and now the effective noise $N_1'(t)$ and $N_2'(t)$ are mutually independent. Furthermore, we have $X(t) - X'(t) - (Y_1(t), Y_2(t))$. It can be easily shown that $(R_1, R_2, D_1, D_2, D_3) \in \mathcal{Q}_{X'}$ if and only if $(R_1, R_2, D_1', D_2', D_3') \in \mathcal{Q}$, where

$$D_i' = \frac{\sigma_X^4}{(\sigma_X^2+\sigma_{N'}^2)^2}D_i + \frac{\sigma_X^2\sigma_{N'}^2}{\sigma_X^2+\sigma_{N'}^2}, \qquad i=1,2,3.$$

Example 2: If $-\sigma_X^2 < \mathbb{E}[N_1(t)N_2(t)] < 0$, then we can write $X(t) = X''(t) + N''(t)$ such that $\sigma_{N''}^2 \triangleq \mathrm{var}(N''(t)) = -\mathbb{E}[N_1(t)N_2(t)]$ and $\mathbb{E}[X''(t)N''(t)] = 0$. Let $\mathcal{Q}_{X''}$ be the rate–distortion region with $\{X''(t)\}_{t=1}^\infty$ as the hidden source and $(\{Y_1(t)\}_{t=1}^\infty, \{Y_2(t)\}_{t=1}^\infty)$ as the noisy observations at two separate encoders. Define $N_i''(t) = N_i(t) + N''(t)$ $(i = 1, 2)$. Note that $Y_i(t) = X''(t) + N_i''(t)$ $(i = 1, 2)$, and now the effective noise values $N_1''(t)$ and $N_2''(t)$ are mutually independent. Furthermore, we have $X''(t) - X(t) - (Y_1(t), Y_2(t))$. It can be readily shown that $(R_1, R_2, D_1, D_2, D_3) \in \mathcal{Q}$ if and only if $(R_1, R_2, D_1'', D_2'', D_3'') \in \mathcal{Q}_{X''}$, where

$$D_i'' = \frac{(\sigma_X^2-\sigma_{N''}^2)^2}{\sigma_X^4}D_i + \frac{(\sigma_X^2-\sigma_{N''}^2)\sigma_{N''}^2}{\sigma_X^2}, \quad i=1,2,3.$$

Therefore, in the preceding examples, it suffices to characterize $\mathcal{Q}_{X'}$ and $\mathcal{Q}_{X''}$; the results derived for the uncorrelated noise case can be directly applied.

VI. Conclusion

We proposed a robust distributed source coding scheme which flexibly trades off between system robustness and compression efficiency. The achievable rate–distortion region of this scheme was analyzed in detail for the quadratic Gaussian case. We also derived an outer bound on the rate–distortion region, which was leveraged to establish the optimality of the proposed scheme in several special cases.

Appendix I
Proof of Theorem 1

The proof of Theorem 1 employs techniques which have already been established in the literature, especially in [43], [51]–[53]. Hence, we only give a sketch here.

For each (U_1, U_2, W_1, W_2) satisfying conditions 1) and 3), we shall prove the admissibility of the rate tuple (R_1, R_2), where

$$\begin{aligned}R_1 &= I(Y_1;U_1)+I(Y_1;W_1|U_1,U_2,W_2),\\ R_2 &= I(Y_2;U_2)+I(Y_2;W_2|U_1,U_2).\end{aligned}$$

Then by symmetry, the rate tuple (R_1', R_2') with

$$\begin{aligned}R_1' &= I(Y_1;U_1)+I(Y_1;W_1|U_1,U_2),\\ R_2' &= I(Y_2;U_2)+I(Y_2;W_2|U_1,U_2,W_1)\end{aligned}$$

is also admissible. Theorem 1 follows by time-sharing (R_1, R_2) and (R_1', R_2').

It was shown in [43] that for any positive ϵ and sufficiently large n with

$$|\mathcal{C}_1^{(n)}| \le \exp(n(I(Y_1;U_1) + I(Y_1;W_1|U_1,U_2,W_2) + \epsilon))$$

decoder 1 and decoder 3 can recover u_1^n and decoder 1 can construct $\hat{x}_1^n$ with $\hat{x}_1(t) = f_1(u_1(t)), t = 1, 2, \ldots, n$, such that

$$\frac{1}{n}\mathbb{E}\sum_{t=1}^{n} d(X(t), \hat{X}_1(t)) \le D_1 + \epsilon;$$

moreover, provided u_2^n and w_2^n are available to decoder 3, it can further recover w_1^n and use $w_1^n, w_2^n, u_1^n, u_2^n$ to construct $\hat{x}_3^n$ with $\hat{x}_3(t) = f_3(u_1(t), w_1(t), u_1(t), w_2(t)), t = 1, 2, \ldots, n$, such that the average distortion is less than or equal to $D_3 + \epsilon$.

Similarly, with

$$|\mathcal{C}_2^{(n)}| \le \exp(n(I(Y_2;U_2) + I(Y_2;W_2|U_1,U_2) + \epsilon))$$

decoder 2 and decoder 3 can recover u_2^n and decoder 2 can construct $\hat{x}_2^n$ with $\hat{x}_2(t) = f_2(u_2(t)), t = 1, 2, \cdots, n$, such that

$$\frac{1}{n}\mathbb{E}\sum_{t=1}^{n} d(X(t), \hat{X}_2(t)) \le D_2 + \epsilon;$$

moreover, provided u_1^n is available to decoder 3, it can further recover w_2^n.

In summary, decoder i recovers u_i^n $(i = 1, 2)$, and decoder 3 recovers $u_1^n, u_2^n, w_1^n, w_2^n$ with the decoding order $(u_1^n, u_2^n) \to w_2^n \to w_1^n$. The proof is complete.

APPENDIX II
PROOF OF THEOREM 4

The following lemma was proved in [17], [18].

Lemma 1: Let

$$r_{i1} = I(X^n; f_{E,i}^{(n)}(Y_i^n))/n$$
$$r_{i2} = I(Y_i^n; f_{E,i}^{(n)}(Y_i^n)|X^n)/n, \qquad i = 1, 2.$$

We have

$$\frac{1}{\sigma_X^2}\exp(2r_{i1}) \le \frac{1}{\sigma_X^2} + \frac{1-\exp(2r_{i2})}{\sigma_{N_i}^2}, \qquad i = 1, 2$$
$$\frac{1}{\sigma_X^2}\exp\left(\frac{2}{n}I\left(X^n; f_{E,1}^{(n)}(Y_1^n), f_{E,2}^{(n)}(Y_2^n)\right)\right) \le \frac{1}{\sigma_X^2} + \frac{1-\exp(-2r_{12})}{\sigma_{N_i}^2} + \frac{1-\exp(-2r_{22})}{\sigma_{N_2}^2}.$$

Now we are ready to prove Theorem 4.

Proof: By the data processing inequality and the rate–distortion theorem, we have

$$I\left(X^n; f_{E,i}^{(n)}(Y_i^n)\right) \ge I\left(X^n; \hat{X}_i^n\right) \ge \frac{n}{2}\log\frac{\sigma_X^2}{D_i}, \quad i = 1, 2 \tag{12}$$

$$I\left(X^n; f_{E,1}^{(n)}(Y_1^n), f_{E,2}^{(n)}(Y_2^n)\right) \ge I\left(X^n; \hat{X}_3^n\right) \ge \frac{n}{2}\log\frac{\sigma_X^2}{D_3}. \tag{13}$$

It follows from (12), (13), and Lemma 1 that

$$\frac{1}{D_i} \le \frac{\exp(2r_{i1})}{\sigma_X^2}, \qquad i = 1, 2$$
$$\frac{1}{D_3} \le \frac{1}{\sigma_X^2} + \frac{1-\exp(-2r_{12})}{\sigma_{N_1}^2} + \frac{1-\exp(-2r_{22})}{\sigma_{N_2}^2}.$$

In view of the fact that $X^n - Y_i^n - f_{E,i}^{(n)}(Y_i^n), i = 1, 2$, we have

$$\begin{aligned} R_i &\ge \frac{1}{n}I\left(Y_i^n; f_{E,i}^{(n)}(Y_i^n)\right) \\ &= \frac{1}{n}I\left(X^n, Y_i^n; f_{E,i}^{(n)}(Y_i^n)\right) \\ &= \frac{1}{n}I\left(X^n; f_{E,i}^{(n)}(Y_i^n)\right) \\ &\quad + \frac{1}{n}I\left(Y_i^n; f_{E,i}^{(n)}(Y_i^n)\middle| X^n\right) \\ &= r_{i1} + r_{i2}, \qquad i = 1, 2. \end{aligned}$$

Now we proceed to derive a lower bound on $r_{11} + r_{21}$. Note that

$$\begin{aligned} n(r_{11} + r_{21}) &= I\left(X^n; f_{E,1}^{(n)}(Y_1^n)\right) + I\left(X^n; f_{E,2}^{(n)}(Y_2^n)\right) \\ &\overset{(a)}{=} I\left(X^n; f_{E,1}^{(n)}(Y_1^n), f_{E,2}^{(n)}(Y_2^n)\right) \\ &\quad + I\left(f_{E,1}^{(n)}(Y_1^n); f_{E,2}^{(n)}(Y_2^n)\right) \\ &\quad - I\left(f_{E,1}^{(n)}(Y_1^n); f_{E,2}^{(n)}(Y_2^n)\middle| X^n\right) \\ &\overset{(b)}{=} I\left(X^n; f_{E,1}^{(n)}(Y_1^n), f_{E,2}^{(n)}(Y_2^n)\right) \\ &\quad + I\left(f_{E,1}^{(n)}(Y_1^n); f_{E,2}^{(n)}(Y_2^n)\right) \end{aligned} \tag{14}$$

where (a) follows from the identity

$$I(A;BC) = I(A;B) + I(A;C) + I(B;C|A) - I(B;C) \tag{15}$$

and (b) is because $f_{E,1}^{(n)}(Y_1^n) - X^n - f_{E,2}^{(n)}(Y_2^n)$. Substituting (13) into (14), we get

$$n(r_{11} + r_{21}) \ge \frac{n}{2}\log\frac{\sigma_X^2}{D_3} + I\left(f_{E,1}^{(n)}(Y_1^n); f_{E,2}^{(n)}(Y_2^n)\right). \tag{16}$$

To lower-bound $I\left(f_{E,1}^{(n)}(Y_1^n); f_{E,2}^{(n)}(Y_2^n)\right)$, we introduce an auxiliary random vector Z^n such that $Z(t) = S(t) + M(t)$, $t = 1, 2, \cdots, n$, where $\{M(t)\}_{t=1}^n$ are i.i.d zero-mean Gaussian random variables with variance σ_M^2 (which will be optimized later). We assume that M^n is independent of (X^n, Y_1^n, Y_2^n). Since θ^n is indepedent of (Y_1^n, Y_2^n) and thus independent of $(\hat{X}_1^n, \hat{X}_2^n)$, we have

$$\begin{aligned} D_i &\ge \frac{1}{n}\sum_{t=1}^{n}\mathbb{E}\left(X(t) - \hat{X}_i(t)\right)^2 \\ &= \frac{1}{n}\sum_{t=1}^{n}\mathbb{E}(S(t) + \theta(t) - \hat{X}_i(t))^2 \\ &= \frac{1}{n}\sum_{t=1}^{n}\mathbb{E}\left(S(t) - \hat{X}_i(t)\right)^2 + D_{3,\min} \end{aligned}$$

i.e.,

$$\frac{1}{n}\sum_{t=1}^{n}\mathbb{E}\left(S(t) - \hat{X}_i(t)\right)^2 \le D_i - D_{3,\min} = d_i, \qquad i = 1, 2.$$

Since

$$\frac{1}{n}\sum_{t=1}^{n}\mathbb{E}\left(Z(t)-\hat{X}_i(t)\right)^2 = \frac{1}{n}\sum_{t=1}^{n}\mathbb{E}\left(S(t)-\hat{X}_i(t)\right)^2+\frac{1}{n}\sum_{t=1}^{n}\mathbb{E}M^2(t)\le d_i+\sigma_M^2, \quad i=1,2$$

it follows from the rate–distortion theorem that

$$I\left(Z^n;\hat{X}_i^n\right)\ge\frac{n}{2}\log\left(\frac{d_X+\sigma_M^2}{d_i+\sigma_M^2}\right),\qquad i=1,2.$$

Now applying the identity (15) to $I\left(f_{E,1}^{(n)}(Y_1^n);f_{E,2}^{(n)}(Y_2^n)\right)$, we get

$$\begin{aligned}I\left(f_{E,1}^{(n)}(Y_1^n);f_{E,2}^{(n)}(Y_2^n)\right)&=I\left(Z^n;f_{E,1}^{(n)}(Y_1^n)\right)+I\left(Z^n;f_{E,2}^{(n)}(Y_2^n)\right)\\&\quad+I\left(f_{E,1}^{(n)}(Y_1^n);f_{E,2}^{(n)}(Y_2^n)\middle| Z^n\right)\\&\quad-I\left(Z^n;f_{E,1}^{(n)}(Y_1^n),f_{E,2}^{(n)}(Y_2^n)\right)\\&\ge I\left(Z^n;\hat{X}_1^n\right)+I\left(Z^n;\hat{X}_2^n\right)\\&\quad-I\left(Z^n;f_{E,1}^{(n)}(Y_1^n),f_{E,2}^{(n)}(Y_2^n)\right)\\&\ge\frac{n}{2}\log\left(\left(\frac{d_X+\sigma_M^2}{d_1+\sigma_M^2}\right)\left(\frac{d_X+\sigma_M^2}{d_2+\sigma_M^2}\right)\right)\\&\quad-I\left(Z^n;f_{E,1}^{(n)}(Y_1^n),f_{E,2}^{(n)}(Y_2^n)\right).\end{aligned}\tag{17}$$

The term $I(Z^n;f_{E,1}^{(n)}(Y_1^n),f_{E,2}^{(n)}(Y_2^n))$ can be upper-bounded as follows:

$$\begin{aligned}I\left(Z^n;f_{E,1}^{(n)}(Y_1^n),f_{E,2}^{(n)}(Y_2^n)\right)&=h(Z^n)-h\left(Z^n\middle|f_{E,1}^{(n)}(Y_1^n),f_{E,2}^{(n)}(Y_2^n)\right)\\&=\frac{n}{2}\log\left[2\pi e\left(d_X+\sigma_M^2\right)\right]\\&\quad-h\left(S^n+M^n\middle|f_{E,1}^{(n)}(Y_1^n),f_{E,2}^{(n)}(Y_2^n)\right)\\&\overset{(c)}{\le}\frac{n}{2}\log\left(2\pi e\left(d_X+\sigma_M^2\right)\right)\\&\quad-\frac{n}{2}\log\left(\exp\left(\frac{2}{n}h\left(S^n\middle|f_{E,1}^{(n)}(Y_1^n),f_{E,2}^{(n)}(Y_2^n)\right)\right)\right.\\&\quad\left.+2\pi e\sigma_M^2\right)\end{aligned}\tag{18}$$

where (c) follows from the conditional version of the entropy power inequality [54]. Note that

$$\begin{aligned}h\left(S^n\middle|f_{E,1}^{(n)}(Y_1^n),f_{E,2}^{(n)}(Y_2^n)\right)&=h\left(S^n\middle|X^n,f_{E,1}^{(n)}(Y_1^n),f_{E,2}^{(n)}(Y_2^n)\right)\\&\quad+I\left(X^n;S^n\middle|f_{E,1}^{(n)}(Y_1^n),f_{E,2}^{(n)}(Y_2^n)\right)\\&=h\left(S^n\middle|X^n,f_{E,1}^{(n)}(Y_1^n),f_{E,2}^{(n)}(Y_2^n)\right)\\&\quad+I\left(X^n;S^n,f_{E,1}^{(n)}(Y_1^n),f_{E,2}^{(n)}(Y_2^n)\right)\\&\quad-I\left(X^n;f_{E,1}^{(n)}(Y_1^n),f_{E,2}^{(n)}(Y_2^n)\right)\\&=h\left(S^n\middle|X^n,f_{E,1}^{(n)}(Y_1^n),f_{E,2}^{(n)}(Y_2^n)\right)\\&\quad+I(X^n;S^n)-I\left(X^n;f_{E,1}^{(n)}(Y_1^n),f_{E,2}^{(n)}(Y_2^n)\right)\end{aligned}$$

where the last equality follows from the fact $X^n - S^n - (f_{E,1}^{(n)}(Y_1^n),f_{E,2}^{(n)}(Y_2^n))$. Therefore, we have

$$\begin{aligned}\exp\left(\frac{2}{n}h\left(S^n\middle|f_{E,1}^{(n)}(Y_1^n),f_{E,2}^{(n)}(Y_2^n)\right)\right)&=\exp\left(\frac{2}{n}h\left(S^n\middle|X^n,f_{E,1}^{(n)}(Y_1^n),f_{E,2}^{(n)}(Y_2^n)\right)\right)\\&\quad\times\exp\left(\frac{2}{n}I(X^n;S^n)\right)\\&\quad\times\exp\left(-\frac{2}{n}I\left(X^n;f_{E,1}^{(n)}(Y_1^n),f_{E,2}^{(n)}(Y_2^n)\right)\right)\\&=\frac{\sigma_X^2}{D_{3,\min}}\exp\left(\frac{2}{n}h\left(S^n\middle|X^n,f_{E,1}^{(n)}(Y_1^n),f_{E,2}^{(n)}(Y_2^n)\right)\right)\\&\quad\times\exp\left(-\frac{2}{n}I\left(X^n;f_{E,1}^{(n)}(Y_1^n),f_{E,2}^{(n)}(Y_2^n)\right)\right).\end{aligned}\tag{19}$$

Now we proceed to derive a lower bound on $\exp(\frac{2}{n}h(S^n|X^n,f_{E,1}^{(n)}(Y_1^n),f_{E,2}^{(n)}(Y_2^n).))$. Since Y_1^n and Y_2^n are independent conditioned on $(X^n,f_{E,1}^{(n)}(Y_1^n),f_{E,2}^{(n)}(Y_2^n))$, it follows by the conditional version of the entropy power inequality [54] that

$$\begin{aligned}\exp\left(\frac{2}{n}h\left(S^n\middle|X^n,f_{E,1}^{(n)}(Y_1^n),f_{E,2}^{(n)}(Y_2^n)\right)\right)&\ge\sum_{i=1}^{2}\exp\left(\frac{2}{n}h\left(\frac{D_{3,\min}}{\sigma_{N_i}^2}Y_i^n\middle|X^n,f_{E,1}^{(n)}(Y_1^n),f_{E,2}^{(n)}(Y_2^n)\right)\right)\\&=D_{3,\min}^2\sum_{i=1}^{2}\frac{1}{\sigma_{N_i}^4}\exp\left(\frac{2}{n}h\left(Y_i^n\middle|X^n,f_{E,1}^{(n)}(Y_1^n),f_{E,2}^{(n)}(Y_2^n)\right)\right)\\&=D_{3,\min}^2\sum_{i=1}^{2}\frac{1}{\sigma_{N_i}^4}\exp\left(\frac{2}{n}h\left(Y_i^n\middle|X^n,f_{E,i}^{(n)}(Y_1^n)\right)\right)\\&=D_{3,\min}^2\sum_{i=1}^{2}\frac{1}{\sigma_{N_i}^4}\exp\left(\frac{2}{n}h(Y_i^n|X^n)-\frac{2}{n}I\left(Y_i^n;f_{E,i}^{(n)}(Y_i^n)\middle|X^n\right)\right)\\&=2\pi eD_{3,\min}^2\sum_{i=1}^{2}\frac{\exp\left(-2r_{i2}\right)}{\sigma_{N_i}^2}.\end{aligned}\tag{20}$$

Thus, by (19) and (20)

$$\begin{aligned}\exp\left(\frac{2}{n}h\left(S^n\middle|f_{E,1}^{(n)}(Y_1^n),f_{E,2}^{(n)}(Y_2^n)\right)\right)&\ge2\pi e\sigma_X^2D_{3,\min}\left(\sum_{i=1}^{2}\frac{\exp\left(-2r_{i2}\right)}{\sigma_{N_i}^2}\right)\\&\quad\times\exp\left(-\frac{2}{n}I\left(X^n;f_{E,1}^{(n)}(Y_1^n),f_{E,2}^{(n)}(Y_2^n)\right)\right).\end{aligned}\tag{21}$$

Combining (18) and (21) yields that

$$\begin{aligned}I\left(Z^n; f_{E,1}^{(n)}(Y_1^n), f_{E,2}^{(n)}(Y_2^n)\right) &\le \frac{n}{2}\log(d_X+\sigma_M^2))\\ &-\frac{n}{2}\log\left(\sigma_X^2 D_{3,\min}\left(\sum_{i=1}^{2}\frac{\exp(-2r_{i2})}{\sigma_{N_i}^2}\right)\right.\\ &\left.\times\exp\left(-\frac{2}{n}I\left(X^n; f_{E,1}^{(n)}(Y_1^n), f_{E,2}^{(n)}(Y_2^n)\right)\right)+\sigma_M^2\right).\end{aligned} \tag{22}$$

Substituting (22) into (17) and then using (14), we have

$$\begin{aligned}&I\left(f_{E,1}^{(n)}(Y_1^n); f_{E,2}^{(n)}(Y_2^n)\right)\\ &\ge \frac{n}{2}\log\left(\frac{d_X+\sigma_M^2}{(d_1+\sigma_M^2)(d_2+\sigma_M^2)}\right)\\ &\quad+\frac{n}{2}\log\left(\sigma_X^2 D_{3,\min}\left(\sum_{i=1}^{2}\frac{\exp(-2r_{i2})}{\sigma_{N_i}^2}\right)\right.\\ &\quad\left.\times\exp\left(-\frac{2}{n}I\left(X^n; f_{E,1}^{(n)}(Y_1^n), f_{E,2}^{(n)}(Y_2^n)\right)\right)+\sigma_M^2\right)\\ &\ge \frac{n}{2}\log\left(\frac{d_X+\sigma_M^2}{(d_1+\sigma_M^2)(d_2+\sigma_M^2)}\right)\\ &\quad+\frac{n}{2}\log\left(\sigma_X^2 D_{3,\min}\left(\sum_{i=1}^{2}\frac{\exp(-2r_{i2})}{\sigma_{N_i}^2}\right)\right.\\ &\quad\times\exp\left(\frac{2}{n}I\left(f_{E,1}^{(n)}(Y_1^n); f_{E,2}^{(n)}(Y_2^n)\right)\right)\\ &\quad\left.\times\exp(-2(r_{11}+r_{21}))+\sigma_M^2\right)\end{aligned}$$

which can be rewritten as (23) at the bottom of the page. Combining (16) and (23) yields the second expression at the bottom of the page, which implies

$$r_{11}+r_{21}\ge\frac{1}{2}\log\frac{\sigma_X^2}{D_3}+\eta(\sigma_M^2,d_1,d_2,\zeta)$$

where

$$\zeta = D_{3,\min}D_3\left(\sum_{i=1}^{2}\frac{\exp(-2r_{i2})}{\sigma_{N_i}^2}\right)$$

$$\eta(\sigma_M^2,d_X,d_1,d_2,\zeta)=\frac{1}{2}\log\frac{(d_X+\sigma_M^2)(\zeta+\sigma_M^2)}{(d_1+\sigma_M^2)(d_2+\sigma_M^2)}.$$

It can be verified that $\sup_{\sigma_M^2>0}\eta(\sigma_M^2,d_X,d_1,d_2,\zeta)$ can be defined as shown in the third expression at the bottom of the page, and thus the proof is complete. □

Appendix III
Proof of Theorem 5

Note that

$$\begin{aligned}nR_1 &\ge H\left(f_{E,1}^{(n)}(Y_1^n)\right)=I\left(Y_1^n; f_{E,1}^{(n)}(Y_1^n)\right)\\ &=I\left(X^n,Y_1^n; f_{E,1}^{(n)}(Y_1^n)\right)\\ &=I\left(X^n,Y_1^n; f_{E,1}^{(n)}(Y_1^n),Y_2^n\right)\\ &\quad-I\left(X^n,Y_1^n;Y_2^n\,\middle|\,f_{E,1}^{(n)}(Y_1^n)\right)\\ &=I\left(X^n; f_{E,1}^{(n)}(Y_1^n),Y_2^n\right)+I\left(Y_1^n; f_{E,1}^{(n)}(Y_1^n),Y_2^n\,\middle|\,X^n\right)\\ &\quad-I\left(X^n,Y_1^n,f_{E,1}^{(n)}(Y_1^n);Y_2^n\right)+I\left(f_{E,1}^{(n)}(Y_1^n);Y_2^n\right)\\ &=I\left(X^n; f_{E,1}^{(n)}(Y_1^n),Y_2^n\right)+I\left(Y_1^n; f_{E,1}^{(n)}(Y_1^n)\,\middle|\,X^n\right)\\ &\quad+I\left(Y_1^n;Y_2^n\,\middle|\,X^n,f_{E,1}^{(n)}(Y_1^n)\right)\\ &\quad-I\left(X^n,Y_1^n;Y_2^n\right)+I\left(f_{E,1}^{(n)}(Y_1^n);Y_2^n\right).\end{aligned} \tag{24}$$

Now we proceed to bound each term separately. By the data processing inequality and the rate–distortion theorem, we have

$$I\left(X^n; f_{E,1}^{(n)}(Y_1^n),Y_2^n\right)\ge I\left(X^n;\hat{X}_3^n\right)\ge\frac{n}{2}\log\frac{\sigma_X^2}{D_3}. \tag{25}$$

$$\begin{aligned}&\exp\left(\frac{2}{n}I\left(f_{E,1}^{(n)}(Y_1^n); f_{E,2}^{(n)}(Y_2^n)\right)\right)\\ &\qquad\ge\frac{(d_X+\sigma_M^2)\,\sigma_M^2}{(d_1+\sigma_M^2)(d_2+\sigma_M^2)-\sigma_X^2 D_{3,\min}(d_X+\sigma_M^2)\exp(-2(r_{11}+r_{21}))\left(\sum_{i=1}^{2}\frac{\exp(-2r_{i2})}{\sigma_{N_i}^2}\right)}.\end{aligned} \tag{23}$$

$$\exp(2(r_{11}+r_{21}))\ge\frac{\sigma_X^2\sigma_M^2(d_X+\sigma_M^2)}{D_3\left((d_1+\sigma_M^2)(d_2+\sigma_M^2)-\sigma_X^2 D_{3,\min}(d_X+\sigma_M^2)\exp(-2(r_{11}+r_{21}))\left(\sum_{i=1}^{2}\frac{\exp(-2r_{i2})}{\sigma_{N_i}^2}\right)\right)}$$

$$\sup_{\sigma_M^2>0}\eta(\sigma_M^2,d_X,d_1,d_2,\zeta)=\begin{cases}0, & \zeta\le d_1+d_2-d_X\\ \frac{1}{2}\log\frac{d_X\zeta}{d_1d_2}, & \zeta\ge\left(\frac{1}{d_1}+\frac{1}{d_2}-\frac{1}{d_X}\right)^{-1}\\ \frac{1}{2}\log\frac{(d_X-\zeta)^2}{(d_X-\zeta)^2-[\sqrt{(d_X-d_1)(d_X-d_2)}-\sqrt{(d_1-\zeta)(d_2-\zeta)}]^2}, & \text{otherwise}\end{cases}$$

$$I\left(Y_1^n; f_{E,1}^{(n)}\left(Y_1^n\right)\middle| X^n\right) \geq \frac{n}{2}\log\frac{\sigma_X^2\sigma_{N_2}^2 D_3}{\left(D_3\sigma_X^2\sigma_{N_1}^2 + D_3\sigma_X^2\sigma_{N_2}^2 + D_3\sigma_{N_1}^2\sigma_{N_2}^2 - \sigma_X^2\sigma_{N_1}^2\sigma_{N_2}^2\right)}. \quad (27)$$

$$R_1 \geq \frac{1}{2}\log\frac{\sigma_X^4\sigma_{N_2}^4}{\left(\sigma_{N_2}^2 + D_1\right)\left(D_3\sigma_X^2\sigma_{N_1}^2 + D_3\sigma_X^2\sigma_{N_2}^2 + D_3\sigma_{N_1}^2\sigma_{N_2}^2 - \sigma_X^2\sigma_{N_1}^2\sigma_{N_2}^2\right)}. \quad (31)$$

Invoking Lemma 1 with $f_{E,2}^{(n)}(Y_2^n) = Y_2^n$, we get

$$\frac{1}{\sigma_X^2} + \frac{1}{\sigma_{N_2}^2} + \frac{1 - \exp\left(-\frac{2}{n}I\left(Y_1^n; f_{E,1}^{(n)}(Y_1^n)\middle| X^n\right)\right)}{\sigma_{N_1}^2} \geq \frac{1}{\sigma_X^2}\exp\left(\frac{2}{n}I\left(X^n; f_{E,1}^{(n)}(Y_1^n), Y_2^n\right)\right). \quad (26)$$

Combining (25) and (26) and after some simple calculation, we obtain (27) at the top of the page. Since $Y_1^n - X^n - Y_2^n$, it follows that

$$I\left(Y_1^n; Y_2^n \middle| X^n, f_{E,1}^{(n)}(Y_1^n)\right) = 0 \quad (28)$$

and

$$I\left(X^n, Y_1^n; Y_2^n\right) = I\left(X^n; Y_2^n\right) = \frac{n}{2}\log\frac{\sigma_X^2 + \sigma_{N_2}^2}{\sigma_{N_2}^2}. \quad (29)$$

To bound the term $I(f_{E,1}^{(n)}(Y_1^n); Y_2^n)$, we first note that

$$\begin{aligned}\frac{1}{n}\sum_{t=1}^{n}\mathbb{E}(Y_2(t) - \hat{X}_1(t))^2 &= \frac{1}{n}\sum_{t=1}^{n}\mathbb{E}(X(t) - \hat{X}_1(t))^2 + \frac{1}{n}\sum_{t=1}^{n}\mathbb{E}N_2^2(t)\\ &\leq D_1 + \sigma_{N_2}^2;\end{aligned}$$

now it follows by the data processing inequality and the rate–distortion theorem that

$$I\left(f_{E,1}^{(n)}(Y_1^n); Y_2^n\right) \geq I\left(\hat{X}_1^n; Y_2^n\right) \geq \frac{n}{2}\log\frac{\sigma_X^2 + \sigma_{N_2}^2}{D_1 + \sigma_{N_2}^2}. \quad (30)$$

Substituting (25)–(30) back to (24), we get (31), also at the top of the page, and thus, the proof is complete.

Remark: The main technical difference between the derivation here and the one used to prove Theorem 4 is the way to lower-bound $I(f_{E,1}^{(n)}(Y_1^n); f_{E,2}^{(n)}(Y_2^n))$. Since for the Gaussian remote Heegard–Berger problem, it boils down to lower-bounding $I(f_{E,1}^{(n)}(Y_1^n); Y_2^n)$, one can adopt the straightforward approach as shown above (cf. (30)).

ACKNOWLEDGMENT

The authors wish to thank the Associate Editor and the two anonymous reviewers for their valuable comments and suggestions.

REFERENCES

[1] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inf. Theory*, vol. IT-19, no. 4, pp. 471–480, Jul. 1973.
[2] T. M. Cover, "A proof of the data compression theorem of Slepian and Wolf for ergodic sources," *IEEE Trans. Inf. Theory*, vol. IT-21, no. 2, pp. 226–228, Mar. 1975.
[3] A. D. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. Inf. Theory*, vol. IT-22, no. 1, pp. 1–10, Jan. 1976.
[4] T. Berger, "Multiterminal source coding," in *The Information Theory Approach to Communications (CISM Courses and Lectures, no. 229)*, G. Longo, Ed. Vienna/New York: Springer-Verlag, 1978, pp. 171–231.
[5] S. Y. Tung, "Multiterminal Source Coding," Ph.D. dissertation, School of Electrical Engineering, Cornell Univ, Ithaca, NY, May 1978.
[6] T. Berger and R. W. Yeung, "Multiterminal source encoding with one distortion criterion," *IEEE Trans. Inf. Theory*, vol. 35, no. 2, pp. 228–236, Mar. 1989.
[7] M. Gastpar, "The Wyner–Ziv problem with multiple sources," *IEEE Trans. Inf. Theory*, vol. 50, no. 11, pp. 2762–2768, Nov. 2004.
[8] Y. Oohama, "Gaussian multiterminal source coding," *IEEE Trans. Inf. Theory*, vol. 43, no. 6, pp. 1912–1923, Nov. 1997.
[9] P. Viswanath, "Sum rate of multiterminal Gaussian source coding," in *Advances in Network Information Theory*, P. Gupta, G. Kramer, and A. Wijngaarden, Eds. Providence, RI: Amer. Math. Soc., 2004, pp. 43–60, DIMACS Series in Discrete Mathematics and Theoretical Computer Science.
[10] S. I. Gel'fand and M. S. Pinsker, "Coding of sources on the basis of observations with incomplete information," *Probl. Inf. Transm.*, vol. 15, no. 2, pp. 115–125, 1979.
[11] T. J. Flynn and R. M. Gray, "Encoding of correlated observations," *IEEE Trans. Inf. Theory*, vol. IT-33, no. 6, pp. 773–787, Nov. 1987.
[12] T. Berger, Z. Zhang, and H. Viswanathan, "The CEO problem," *IEEE Trans. Inf. Theory*, vol. 42, no. 3, pp. 887–902, May 1996.
[13] H. Viswanathan and T. Berger, "The quadratic Gaussian CEO problem," *IEEE Trans. Inf. Theory*, vol. 43, no. 5, pp. 1549–1559, Sep. 1997.
[14] S. C. Draper and G. W. Wornell, "Side information aware coding strategies for sensor networks," *IEEE J. Select. Areas Commun.*, vol. 22, no. 6, pp. 966–976, Aug. 2004.
[15] Y. Oohama, "The rate-distortion function for the quadratic Gaussian CEO problem," *IEEE Trans. Inf. Theory*, vol. 44, no. 3, pp. 1057–1070, May 1998.
[16] J. Chen, X. Zhang, T. Berger, and S. B. Wicker, "An upper bound on the sum-rate distortion function and its corresponding rate allocation schemes for the CEO problem," *IEEE J. Select. Areas Commun.*, vol. 22, no. 6, pp. 977–987, Aug. 2004.
[17] Y. Oohama, "Rate–distortion theory for Gaussian multiterminal source coding systems with several side informations at the decoder," *IEEE Trans. Inf. Theory*, vol. 51, no. 7, pp. 2577–2593, Jul. 2005.
[18] V. Prabhakaran, D. N. C. Tse, and K. Ramchandran, "Rate region of the quadratic Gaussian CEO problem," in *Proc. IEEE Int. Symp. Information Theory*, Chicago, IL, Jun./Jul. 2004, p. 119.
[19] H. Witsenhausen, "On source networks with minimal breakdown degradation," *Bell Syst. Tech. J.*, vol. 59, no. 6, pp. 1083–1087, July–Aug. 1980.
[20] J. Wolf, A. Wyner, and J. Ziv, "Source coding for multiple descriptions," *Bell Syst. Tech. J.*, vol. 59, no. 8, pp. 1417–1426, Oct. 1980.
[21] 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.

[22] H. S. Witsenhausen and A. D. Wyner, "Source coding for multiple descriptions II: A binary source," *Bell Syst. Tech. J.*, vol. 60, pp. 2281–2292, Nov. 1981.

[23] 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.

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

[25] R. Ahlswede, "On multiple descriptions and team guessing," *IEEE Trans. Inf. Theory*, vol. IT-32, no. 4, pp. 543–549, Jul. 1986.

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

[27] R. Zamir, "Gaussian codes and Shannon bounds for multiple descriptions," *IEEE Trans. Inf. Theory*, vol. 45, no. 7, pp. 2629–2635, Nov. 1999.

[28] F. W. Fu and R. W. Yeung, "On the rate-distortion region for multiple descriptions," *IEEE Trans. Inf. Theory*, vol. 48, no. 7, pp. 2012–2021, Jul. 2002.

[29] R. Venkataramani, G. Kramer, and V. K. Goyal, "Multiple description coding with many channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 9, pp. 2106–2114, Sep. 2003.

[30] S. S. Pradhan, R. Puri, and K. Ramchandran, "n-channel symmetric multiple descriptions–part I: (n, k) source–channel erasure codes," *IEEE Trans. Inf. Theory*, vol. 50, no. 1, pp. 47–61, Jan. 2004.

[31] R. Puri, S. S. Pradhan, and K. Ramchandran, "n-channel symmetric multiple descriptions—part II: An achievable rate–distortion region," *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1377–1392, Apr. 2005.

[32] J. Wolf and J. Ziv, "Transmission of noisy information to a noisy receiver with minimum distortion," *IEEE Trans. Inf. Theory*, vol. IT-16, no. 4, pp. 406–411, Jul. 1970.

[33] T. Berger, *Rate Distortion Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1971.

[34] P. Ishwar, R. Puri, K. Ramchandran, and S. S. Pradhan, "On rate-constrained distributed estimation in unreliable sensor networks," *IEEE J. Select. Areas Commun*, vol. 23, no. 4, pp. 765–775, Apr. 2005.

[35] J. R. Roche, "Distributed Information Storage," Ph.D. dissertation, Stanford Univ., Stanford, CA, Mar. 1992.

[36] R. W. Yeung, "Multilevel diversity coding with distortion," *IEEE Trans. Inf. Theory*, vol. 41, no. 2, pp. 412–422, Mar. 1995.

[37] J. R. Roche, R. W. Yeung, and K. P. Hau, "Symmetrical multilevel diversity coding," *IEEE Trans. Inf. Theory*, vol. 43, no. 3, pp. 1059–1064, May 1997.

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

[39] R. W. Yeung and Z. Zhang, "Distributed source coding for satellite communications," *IEEE Trans. Inf. Theory*, vol. 45, no. 4, pp. 1111–1120, May 1999.

[40] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. New York: Academic, 1981.

[41] J. Edmonds, "Submodular functions, matroids and certain polyhedra," in *Combinatorial Structures and Their Applications (Proc. Calgary Int. Conf. 1969)*, R. Guy, H. Hanani, N. Sauer, and J. Schonheim, Eds. New York: Gordon and Breach, 1970, pp. 69–87.

[42] J. Körner and K. Marton, "How to encode the modulo-two sum of binary sources," *IEEE Trans. Inf. Theory*, vol. IT-25, no. 2, pp. 219–221, Mar. 1979.

[43] C. Heegard and T. Berger, "Rate distortion when side information may be absent," *IEEE Trans. Inf. Theory*, vol. IT-31, no. 6, pp. 727–734, Nov. 1985.

[44] P. Gács and J. Körner, "Common information is far less than mutual information," *Probl. Contr. Inf. Theory*, vol. 2, pp. 149–162, 1973.

[45] H. S. Witsenhausen, "On sequences of pairs of dependent random variables," *SIAM J. Appl. Math.*, vol. 28, pp. 100–113, Jan. 1975.

[46] T. M. Cover, A. A. El Gamal, and M. Salehi, "Multiple access channels with arbitrarily correlated sources," *IEEE Trans. Inf. Theory*, vol. IT-26, no. 6, pp. 648–657, Nov. 1980.

[47] A. D. Wyner, "The rate-distortion function for source coding with side information at the decoder-II: General sources," *Inf. Contr.*, vol. 38, pp. 60–80, Jul. 1978.

[48] H. Feng and M. Effros, "On the achievable region for multiple description source codes on gaussian sources," in *Proc. IEEE Int. Symp. Information Theory*, Yokohama, Japan, Jun./Jul. 2003, p. 195.

[49] J. Chen, C. Tian, T. Berger, and S. S. Hemami, "Multiple description quantization via Gram–Schmidt orthogonalization," *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5197–5217, Dec. 2006.

[50] J. Chen and A. B. Wagner, "A semicontinuity theorem and its application to network source coding," in *Proc. IEEE Int. Symp. Information Theory*, Toronto, ON, Canada, Jul. 2008, submitted for publication.

[51] T. S. Han and K. Kobayashi, "A unified achievable rate region for a general class of multiterminal source coding systems," *IEEE Trans. Inf. Theory*, vol. IT-26, no. 3, pp. 277–288, May 1980.

[52] T. Berger and R. W. Yeung, "Multiterminal source encoding with encoder breakdown," *IEEE Trans. Inf. Theory*, vol. 35, no. 2, pp. 237–244, Mar. 1989.

[53] R. Ahlswede and J. Körner, "Source coding with side information and converse for degraded broadcast channel," *IEEE Trans. Inf. Theory*, vol. IT-21, no. 6, pp. 629–637, Dec. 1975.

[54] N. M. Blachman, "The convolution inequality for entropy powers," *IEEE Trans. Inf. Theory*, vol. IT-11, no. 2, pp. 267–271, Apr. 1965.