Vector Gaussian Multiple Description with Individual and Central Receivers[1]

Hua Wang and Pramod Viswanath [2]

Abstract

L multiple descriptions of a vector Gaussian source for individual and central receivers are investigated. The sum rate of the descriptions with covariance distortion measure constraints, in a positive semidefinite ordering, is exactly characterized. For two descriptions, the entire rate region is characterized. The key component of the solution is a novel information-theoretic inequality that is used to lower bound the achievable multiple description rates. Jointly Gaussian descriptions are optimal in achieving the limiting rates. We also show the robustness of this description scheme: the distortions achieved are no larger when used to describe any non-Gaussian source with the same covariance matrix.

1 Introduction

In the multiple description problem, an information source is encoded into L packets and these packets are sent through parallel communication channels. There are several receivers, each of which can receive a subset of the packets and needs to reconstruct the information source from this subset. In the most general case, there are $2^L - 1$ receivers and the packets received at each receiver correspond to one of $2^L - 1$ subsets of $\{1, \ldots, L\}$. A long standing open problem in the literature [1–10] is to characterize the information-theoretic region of packet encoding rates subject to the specified distortion constraints on the reconstructions. Practical multiple description codes have been discussed in [11–18] and recent work [19,20] has considered the multiple description problem in the context of the distributed source coding scenario. From a fundamental view point, however, optimal descriptions of even the Gaussian source with quadratic distortion measures have not been fully characterized. Only in the special case of two descriptions of a scalar Gaussian source with quadratic distortion measures, the entire rate region has been characterized in [1].

We view the information source as a stationary and ergodic process. To simplify the study, we consider block memoryless information sources in this paper, i.e., we model

[1]This research was sponsored in part by NSF CCR-0325924 and a Vodafone US Foundation Fellowship.

[2]The authors are with the Department of Electrical and Computer Engineering and the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Urbana IL 61801; e-mail: {huawang,pramodv}@uiuc.edu

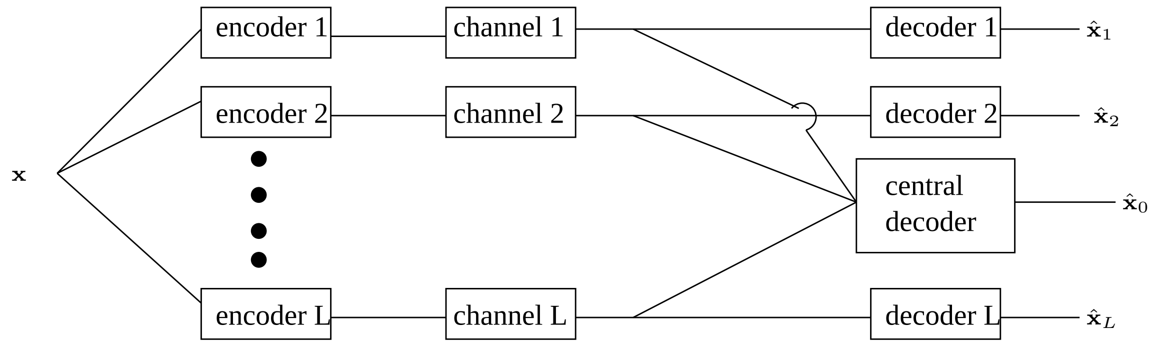

Figure 1: MD problem with only individual reconstructions and central reconstruction

the memory in information source by dividing the source sequence into blocks, and while the different source samples inside the same block are jointly distributed, the blocks themselves are independent and identically distributed. We can think of each block as a random vector, and the whole source sequence as a *memoryless vector* stochastic process. Specifically, our focus in this paper is on L descriptions of a memoryless vector *Gaussian* source.

In this paper, we consider only $L+1$ receivers – L individual and a single common receiver (cf. Figure 1). Each receiver needs to reconstruct the original source such that the empirical covariance matrix of the difference is less than, in the sense of a positive semidefinite ordering, a "distortion" matrix. This form of distortion constraint is quite general; all quadratic distortion constraints can be handled via the covariance matrix distortion constraint. In this setting, the symmetric rate multiple description problem of a scalar Gaussian source with symmetric distortion constraints has been characterized in [7, 8, 10], but a complete understanding of all other rate-distortion settings is open.

Our main result is an *exact* characterization of the sum rate for any specified $L+1$ distortion matrix constraints. With $L = 2$, we characterize the entire rate region. A natural Gaussian multiple description scheme is shown to be optimal in these scenarios. Our contribution is two fold:

- First, we derive a novel information-theoretic inequality that provides a lower bound to the sum of the description rates. The key step is to avoid using the entropy power inequality, which was a central part of the proof of two descriptions of the scalar Gaussian source in [1]: the vector entropy power inequality is tight only with a certain covariance alignment condition, which arbitrary distortion matrix requirements do not necessarily allow.
- Second, we show that jointly Gaussian descriptions actually achieve the lower bound not by resorting to a direct calculation and comparison, which appears to be difficult

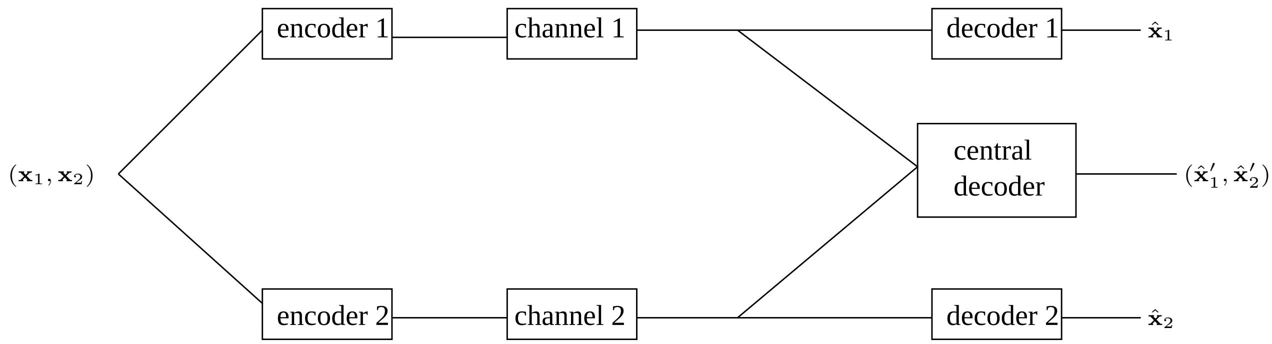

Figure 2: Multiple Descriptions with separate distortion constraints.

for $L > 2$, but instead by arguing the equivalence of certain optimization problems.

The upshot is that Gaussian multiple descriptions are optimal in terms of the sum rate of descriptions of a vector *Gaussian* source. It turns out that this description scheme is *robust*: for any other memoryless vector source with a fixed covariance matrix, the distortion achieved for the Gaussian source is the *largest* (in a strong positive semidefinite ordering sense).

Consider another problem of two descriptions of a pair of jointly Gaussian memoryless sources as depicted in Figure 2. There are two encoders that describe this source to three receivers: receiver i gets the description of encoder i, with $i = 1, 2$ and the third receiver receives both the descriptions. Suppose receiver i is interested in reconstructing the ith marginal of the jointly Gaussian source, with $i = 1, 2$. The third receiver is interested in reconstructing the entire vector source. This description problem is closely related to the vector Gaussian description problem that is the main focus of this paper. We exploit this connection and characterize the rate region where the reconstructions have a constraint on the covariance of error at each of the receivers (in the sense of a positive semidefinite order).

We have organized the results in this paper as follows. In Section 2 we give a formal description of the problem and summarize our main result. The derivation of an achievable sum rate is in Section 3. In Section 4 we derive a lower bound to the sum rate. We show that the achievable sum rate is equal to the lower bound in Section 5, thus completing the characterization of the sum rate. For certain special cases, we can derive more structure to the optimal Gaussian multiple description scheme: in Section 6 we focus on the *scalar* Gaussian source and in Section 7 we focus on *two* descriptions of a vector Gaussian source. In both these cases, we provide detailed and explicit structure to the optimal Gaussian multiple description scheme.

Moving to the description problem described in Figure 2, we use the previous results to resolve this problem: again, Gaussian multiple descriptions are optimal in achieving the rate region; this is done in Section 8.1. Finally, while the characterization of the rate region of general multiple descriptions of the Gaussian source (with each receiver having access to some subset of the descriptions) is still open, we can use the insights derived via our sum rate characterization to solve this problem for a nontrivial set of covariance distortion constraints; this is done in Section 8.2. In Section 8.3 and Section 8.4 we study the case where the source is not Gaussian. We provide upper and lower bound to sum rate, and show that Gaussian source is the hardest to compress. The robustness of the Gaussian distributed architecture is shown in Section 8.3.

A note about the notation in this paper: we use lower case letters for scalars, lower case and bold face for vectors, upper case and bold face for matrices. The superscript t denotes matrix transpose. We use $\mathbf{I}_N$ and $\mathbf{0}$ to denote the identity matrix and the all zero matrix respectively, and $\text{diag}\{p_1, \ldots, p_n\}$ to denote a diagonal matrix with the diagonal entries equal to $p_1, \ldots, p_n$. The partial order $\succ$ ($\succcurlyeq$) denotes positive definite (semidefinite) ordering: $\mathbf{A} \succ \mathbf{B}$ ($\mathbf{A} \succcurlyeq \mathbf{B}$) means that $\mathbf{A} - \mathbf{B}$ is a positive definite (semidefinite) matrix. We write $\mathcal{N}(\mu, \mathbf{Q})$ to denote a Gaussian random vector with mean μ and covariance $\mathbf{Q}$. All logarithms in this paper are to the natural base.

2 Problem Setting and Main Results

2.1 Problem Setting

The information source $\{\mathbf{x}[m]\}$ is an i.i.d. random process with the marginal distribution $\mathcal{N}(0, \mathbf{K}_x)$, i.e., a collection of i.i.d. Gaussian random vectors. Denoting the dimension of $\{\mathbf{x}[m]\}$ by N, we suppose that $\mathbf{K}_x$ is an $N \times N$ positive definite matrix. There are L encoding functions at the source, encoder l encodes a source sequence, of length n, $\mathbf{x}^n = (\mathbf{x}[1], \ \ldots, \ \mathbf{x}[n])^t$ to a source codeword $f_l^{(n)}(\mathbf{x}^n)$, for $l = 1 \ldots L$. This codeword $f_l^{(n)}(\mathbf{x}^n)$ is sent through lth communication channel at the rate $R_l = \frac{1}{n} \log |C_l^{(n)}|$, where $C_l^{(n)}$ is the code book of encoder l.

There are L individual receivers and one central receiver. For $l = 1, \ \ldots \ L$, the lth individual receiver uses its information (the output of the lth channel) to generate an estimation $\hat{\mathbf{x}}_l^n = g_l^{(n)} \left(f_l^{(n)}(\mathbf{x}^n) \right)$ of the source sequence $\mathbf{x}^n$. The central receiver uses the output of all the L channels to generate an estimate $\hat{\mathbf{x}}_0^n$ of the source sequence $\mathbf{x}^n$. Since we are interested in covariance constraints, the decoders' maps can be restricted, without loss of generality, to be the minimal mean square error (MMSE) estimation of the source sequence based on the received codewords. So,

$$\begin{aligned}\hat{\mathbf{x}}_l^n &= \mathbb{E}\left[\mathbf{x}^n | f_l^{(n)}(\mathbf{x}^n)\right], \quad l = 1,\ \ldots,\ L \\ \hat{\mathbf{x}}_0^n &= \mathbb{E}\left[\mathbf{x}^n | f_1^{(n)}(\mathbf{x}^n),\ \ldots,\ f_L^{(n)}(\mathbf{x}^n)\right].\end{aligned} \tag{1}$$

If the reconstructed sequences satisfy the covariance constraints

$$\begin{aligned}\frac{1}{n}\sum_{m=1}^{n}\mathbb{E}\Big[(\mathbf{x}[m]-\hat{\mathbf{x}}_l[m])^t(\mathbf{x}[m]-\hat{\mathbf{x}}_l[m])\Big] &\preccurlyeq \mathbf{D}_l, \quad l = 1,\ \ldots,\ L, \\ \frac{1}{n}\sum_{m=1}^{n}\mathbb{E}\Big[(\mathbf{x}[m]-\hat{\mathbf{x}}_0[m])^t(\mathbf{x}[m]-\hat{\mathbf{x}}_0[m])\Big] &\preccurlyeq \mathbf{D}_0,\end{aligned} \tag{2}$$

then we say that multiple descriptions with distortion constraints $(\mathbf{D}_1,\ \ldots,\ \mathbf{D}_L,\ \mathbf{D}_0)$ are achievable at the rate tuple $(R_1,\ \ldots,\ R_L)$.

The closure of the convex hull of the set of all achievable rate tuples is called the rate-distortion region and is denoted by $\mathcal{R}_*(\mathbf{K}_x,\ \mathbf{D}_1,\ \ldots,\ \mathbf{D}_L,\ \mathbf{D}_0)$. Throughout this paper, we suppose that $\mathbf{0} \prec \mathbf{D}_0 \prec \mathbf{D}_l \prec \mathbf{K}_x,\ \forall l = 1, \ldots, L$.[3]

2.2 A Natural Achievable Scheme

There is a natural *Gaussian random multiple description scheme* for the multiple description problem described in the previous section. Let $\mathbf{w}_1,\ \cdots,\ \mathbf{w}_L$ be N dimensional zero mean jointly Gaussian random vectors independent of $\mathbf{x}$, with the positive definite covariance matrix $(\mathbf{w}_1,\ \cdots,\ \mathbf{w}_L)$ denoted by $\mathbf{K}_w$. Defining

$$\mathbf{u}_l = \mathbf{x} + \mathbf{w}_l, \quad l = 1,\ \ldots,\ L,$$

we consider $\mathbf{K}_w$ such that

$$\begin{aligned}\mathrm{Cov}[\mathbf{x}|\mathbf{u}_l] &\stackrel{\text{def}}{=} \mathbb{E}\Big[(\mathbf{x}-\mathbb{E}[\mathbf{x}|\mathbf{u}_l])^t(\mathbf{x}-\mathbb{E}[\mathbf{x}|\mathbf{u}_l])\Big] \preccurlyeq \mathbf{D}_l, \quad l = 1,\ \ldots,\ L, \\ \mathrm{Cov}[\mathbf{x}|\mathbf{u}_1,\ \ldots,\ \mathbf{u}_L] &\stackrel{\text{def}}{=} \mathbb{E}\Big[(\mathbf{x}-\mathbb{E}[\mathbf{x}|\mathbf{u}_1,\ \ldots,\ \mathbf{u}_L])^t(\mathbf{x}-\mathbb{E}[\mathbf{x}|\mathbf{u}_1,\ \ldots,\ \mathbf{u}_L])\Big] \preccurlyeq \mathbf{D}_0.\end{aligned} \tag{3}$$

To construct the code book for the lth description, draw e^{nR_l} $\mathbf{u}_l^n$ vectors randomly according to the marginal of $\mathbf{u}_l$. The encoders observe the source sequence $\mathbf{x}^n$, look for

[3] That $\mathbf{D}_0 \preccurlyeq \mathbf{D}_l$, is without loss of generality is seen by applying the data processing inequality for MMSE estimation errors; having more access to information can only reduce the covariance of the error in a positive semidefinite sense. Similarly, $\mathbf{K}_x \preccurlyeq \mathbf{D}_0$ is also not interesting; here we simplify this condition and take $\mathbf{D}_0 \prec \mathbf{K}_x$.

codewords $(\mathbf{u}_1^n, \ldots, \mathbf{u}_L^n)$ that are jointly typical with $\mathbf{x}^n$ and send the index of the resulting $\mathbf{u}_l^n$ through the lth channel, respectively. The lth individual receiver uses this index and generates a reproduction sequence $\mathbb{E}[\mathbf{x}^n|\mathbf{u}_l^n]$ for $l = 1 \ldots L$, the central receiver uses all the L indices to generate a reproduction sequence $\mathbb{E}[\mathbf{x}^n|\mathbf{u}_1^n, \ldots, \mathbf{u}_L^n]$. The following lemma gives the achievable rate region by using this scheme.

***Lemma* 1.** *For every $\mathbf{K}_w$ satisfying* (3)*, the rate tuple $(R_1, \ldots, R_L)$ satisfying*

$$\sum_{l \in S} R_l \geq \left[\sum_{l \in S} h(\mathbf{u}_l)\right] - h(\mathbf{u}_l, l \in S|\mathbf{x}) = \frac{1}{2} \log \frac{\prod_{l \in S} |\mathbf{K}_x + \mathbf{K}_{w_l}|}{|\mathbf{K}_{w_S}|}, \quad \forall S \subseteq \{1, \ldots, L\} \tag{4}$$

is achievable, where $\mathbf{K}_{w_S}$ is the covariance matrix for all $\mathbf{w}_l, l \in S$, and $\mathbf{K}_{w_l} = \mathbb{E}[\mathbf{w}_l^t \mathbf{w}_l]$. In particular, the achievable sum rate is

$$\min_{\mathbf{K}_w} \frac{1}{2} \log \frac{\prod_{l=1}^{L} |\mathbf{K}_x + \mathbf{K}_{w_l}|}{|\mathbf{K}_w|}. \tag{5}$$

Proof. This follows from [7, Theorem 1]. For completeness, we provide a sketch of the proof in Section 3. □

We denote this ensemble of descriptions, throughout this paper, as the Gaussian description scheme and when embellished with time sharing, as the Gaussian description *strategy*. Later we will show that Gaussian description schemes can achieve the optimal sum rate.

2.3 Lower Bound to Sum Rate

We have the following lower bound on the sum rate for multiple description with individual and central receivers for an i.i.d $\mathcal{N}(0, \mathbf{K}_x)$ Gaussian source.

***Theorem* 1.** *Given distortion constraints $(\mathbf{D}_1, \ldots, \mathbf{D}_L, \mathbf{D}_0)$, the sum rate satisfies*

$$\sum_{l=1}^{L} R_l \geq \sup_{\mathbf{K}_z \succ \mathbf{0}} \quad \frac{1}{2} \log \left(\frac{|\mathbf{K}_x||\mathbf{K}_x + \mathbf{K}_z|^{(L-1)}|\mathbf{D}_0 + \mathbf{K}_z|}{|\mathbf{D}_0| \prod_{l=1}^{L} |\mathbf{D}_l + \mathbf{K}_z|} \right). \tag{6}$$

Proof. See Section 4. □

2.4 Optimal Sum Rate

We show that the achievable sum rate (127) matches the lower bound (6), thus characterizing the optimal sum rate.

Theorem 2. *The achievable sum rate* (127) *is equal to the lower bound* (6)*. Thus the natural Gaussian achievable scheme is optimal in achieving the sum rate.*

Proof. See Section 5. □

2.5 Rate Region for Two Description Problem

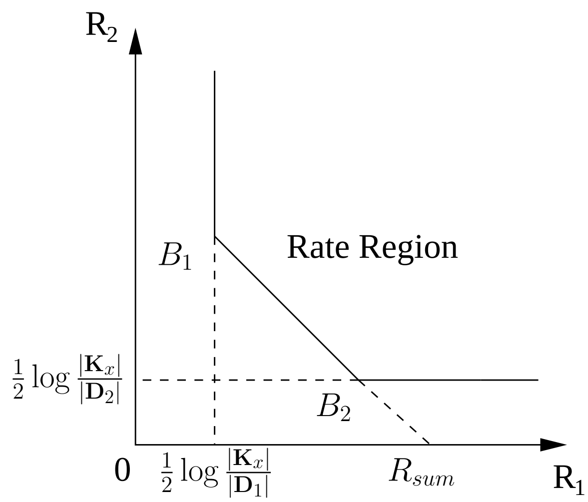

Figure 3: Rate region for two description problem

For two descriptions, we can characterize the entire rate region.

Theorem 3. *Given distortion constraints* $(\mathbf{D}_1,\ \mathbf{D}_2,\ \mathbf{D}_0)$*, the rate region for the two description problem for an i.i.d.* $\mathcal{N}(0, \mathbf{K}_x)$ *vector Gaussian source is*

$$\mathcal{R}_*(\mathbf{K}_x,\ \mathbf{D}_1,\ \mathbf{D}_2,\ \mathbf{D}_0) = \left\{ \begin{array}{l} (R_1,\ R_2): \\ R_l \geq \dfrac{1}{2}\log\dfrac{|\mathbf{K}_x|}{|\mathbf{D}_l|}, \quad l = 1,\ 2 \\ R_1 + R_2 \geq \sup\limits_{\mathbf{K}_z \succ \mathbf{0}} \dfrac{1}{2}\log\dfrac{|\mathbf{K}_x||\mathbf{K}_x + \mathbf{K}_z||\mathbf{D}_0 + \mathbf{K}_z|}{|\mathbf{D}_0||\mathbf{D}_1 + \mathbf{K}_z||\mathbf{D}_2 + \mathbf{K}_z|} \end{array} \right\}. \tag{7}$$

Further, this region is achieved by the Gaussian description strategy.

Proof. This is a generalization of the classical result [1] on the problem of two descriptions of a scalar Gaussian source. The achievability follows from Lemma 1. The proof of the optimality of the Gaussian description strategy is in Section 7.2. □

In some situations, we can explicitly solve for the optimizing $\mathbf{K}_z$ in Equation (7).

***Proposition* 1.** *Let* $\mathbf{K}_{w_l} = [\mathbf{D}_l^{-1} - \mathbf{K}_x^{-1}]^{-1}$ *for* $l = 0, 1, 2$. *If the distortion constraints* $(\mathbf{D}_1,\ \mathbf{D}_2,\ \mathbf{D}_0)$ *satisfy* $\mathbf{D}_0 + \mathbf{K}_x - \mathbf{D}_1 - \mathbf{D}_2 \succ \mathbf{0}$ *and* $\mathbf{D}_0^{-1} + \mathbf{K}_x^{-1} - \mathbf{D}_1^{-1} - \mathbf{D}_2^{-1} \succ \mathbf{0}$, *then the optimizing* $\mathbf{K}_z$ *in the sum rate of equation* (7) *is*

$$\mathbf{K}_z = \mathbf{K}_x(\mathbf{K}_x - \mathbf{A}^*)^{-1}\mathbf{K}_x - \mathbf{K}_x, \tag{8}$$

where $\mathbf{A}^*$ *is given by the solution of the* matrix Riccati *equation*

$$(\mathbf{K}_{w_1} + \mathbf{A}^*)(\mathbf{K}_{w_1} - \mathbf{K}_{w_0})^{-1}(\mathbf{K}_{w_1} + \mathbf{A}^*) = 2\mathbf{A}^* + \mathbf{K}_{w_1} + \mathbf{K}_{w_2} \tag{9}$$

for $\mathbf{A}^*$ *whose solution is*

$$\mathbf{A}^* = (\mathbf{K}_{w_1} - \mathbf{K}_{w_0})^{\frac{1}{2}} \left[(\mathbf{K}_{w_1} - \mathbf{K}_{w_0})^{-\frac{1}{2}}(\mathbf{K}_{w_2} - \mathbf{K}_{w_0})(\mathbf{K}_{w_1} - \mathbf{K}_{w_0})^{-\frac{1}{2}}\right]^{\frac{1}{2}} (\mathbf{K}_{w_1} - \mathbf{K}_{w_0})^{\frac{1}{2}} - \mathbf{K}_{w_0}, \tag{10}$$

Proof. The proof is in Section 7.1. □

An illustration of the rate region is shown in Figure 3. Letting R_{sum} denote the optimal sum rate, the two corner points in Figure 3 are

$$\begin{aligned} B_1 &= \left(\frac{1}{2}\log\frac{|\mathbf{K}_x|}{|\mathbf{D}_1|}, R_{sum} - \frac{1}{2}\log\frac{|\mathbf{K}_x|}{|\mathbf{D}_1|}\right), \quad \text{and} \\ B_2 &= \left(R_{sum} - \frac{1}{2}\log\frac{|\mathbf{K}_x|}{|\mathbf{D}_2|}, \frac{1}{2}\log\frac{|\mathbf{K}_x|}{|\mathbf{D}_2|}\right). \end{aligned} \tag{11}$$

3 Proof of Achievable Sum Rate

In this section we first give a sketch of the achievable sum rate (127). A rigorous proof can be found in [7]. We then discuss an important combinatorial property of the achievable region.

3.1 Achievable Rate Region

By using the Gaussian description scheme described in Section 2.2, we can see that given the source sequence $\mathbf{x}^n$, as long as we can find a combination of codewords $(\mathbf{u}_1^n, \ldots, \mathbf{u}_L^n)$ that is jointly typical with $\mathbf{x}^n$, all the receivers can generate reproduction sequences that satisfy their given distortion constraints. An intuitive way to understand (4) is the following: since $(\mathbf{u}_1^n, \ldots, \mathbf{u}_L^n)$ is jointly typical with $\mathbf{x}^n$, then for any $S \subseteq \{1, \ldots, L\}$, we have that $\mathbf{u}_l^n, l \in S$ is jointly typical with $\mathbf{x}^n$. Now the probability that a randomly generated combination of codewords $\mathbf{u}_l^n, l \in S$ are jointly typical with $\mathbf{x}^n$ is roughly

$$\frac{e^{nh(\mathbf{u}_l, l \in S|\mathbf{x})}}{\prod_{l \in S} e^{nh(\mathbf{u}_l)}}, \tag{12}$$

and the number of possible combination of codewords $\mathbf{u}_l^n, l \in S$ are $\prod_{l \in S} e^{nR_l}$. Thus, as long as

$$\sum_{l \in S} R_l \geq \left[\sum_{l \in S} h(\mathbf{u}_l)\right] - h(\mathbf{u}_l, l \in S|\mathbf{x}), \tag{13}$$

we can find a combination of codewords $\mathbf{u}_l^n, l \in S$ that are jointly typical with $\mathbf{x}^n$. Rigorously speaking, we need to show that as long as (13) is satisfied for all S, then for any given source sequence $\mathbf{x}^n$ we can find a combination of codewords $(\mathbf{u}_1^n, \ldots, \mathbf{u}_L^n)$ such that $\mathbf{u}_l^n, l \in S$ are jointly typical with $\mathbf{x}^n$ for all $S \subseteq \{1, \ldots, L\}$. The *second moment method* [21] is commonly used to address this aspect, and a proof of (13) can be found in [7].

Evaluating (13) based on the Gaussian distribution of $\mathbf{x}$ and $\mathbf{u}_1, \ldots, \mathbf{u}_L$, we get that all the rate tuples $(R_1, \ldots, R_L)$ satisfying

$$\sum_{l \in S} R_l \geq \left[\sum_{l \in S} h(\mathbf{u}_l)\right] - h(\mathbf{u}_l, l \in S|\mathbf{x}) = \frac{1}{2} \log \frac{\prod_{l \in S} |\mathbf{K}_x + \mathbf{K}_{w_l}|}{|\mathbf{K}_{w_S}|}, \quad \forall S \subseteq \{1, \ldots, L\} \tag{14}$$

are achievable by the Gaussian description scheme. In particular, we have that the achievable sum rate is

$$\left[\sum_{l=1}^{L} h(\mathbf{u}_l)\right] - h(\mathbf{u}_1, \ldots, \mathbf{u}_L|\mathbf{x}) = \frac{1}{2} \log \frac{\prod_{l=1}^{L} |\mathbf{K}_x + \mathbf{K}_{w_l}|}{|\mathbf{K}_w|}. \tag{15}$$

We can get the optimal sum rate achieved by Gaussian description schemes through minimizing (15) over all covariance matrix $\mathbf{K}_w$ satisfying the distortion constraint and get (127).

The resulting distortions $(\mathbf{D}_1^*,\ \ldots,\ \mathbf{D}_L^*,\ \mathbf{D}_0^*)$ by using the Gaussian description scheme can be calculated as

$$
\begin{aligned}
\mathbf{D}_l^* =& \mathrm{Cov}[\mathbf{x}|\mathbf{u}_l] = [\mathbf{K}_x^{-1} + \mathbf{K}_{w_l}^{-1}]^{-1}, \quad l = 1,\ \ldots,\ L, \\
\mathbf{D}_0^* =& \mathrm{Cov}[\mathbf{x}|\mathbf{u}_1,\ \ldots,\ \mathbf{u}_L] = [\mathbf{K}_x^{-1} + (\mathbf{I}_N,\ \ldots,\ \mathbf{I}_N)\mathbf{K}_w^{-1}(\mathbf{I}_N,\ \ldots,\ \mathbf{I}_N)^t]^{-1}.
\end{aligned} \tag{16}
$$

3.2 Combinatorial Property of the Achievable Region

The achievable region given in (13) has useful combinatorial properties; in particular it belongs to the class of *contra-polymatroids* [22, 23]. Certain rate regions of the multiple access channel [24] and distributed source coding problems [25] are also known to have this specific combinatorial property.

To see this, let

$$
\phi(S) \overset{\text{def}}{=} \left[\sum_{l \in S} h(\mathbf{u}_l) \right] - h(\mathbf{u}_l, l \in S|\mathbf{x}), \quad S \subseteq \{1,\ \ldots,\ L\}. \tag{17}
$$

We can readily verify that

$$
\begin{aligned}
\phi(S \cup \{t\}) \geq \phi(S), \quad \forall t \in \{1,\ \ldots,\ L\}, \\
\phi(S \cup T) + \phi(S \cap T) \geq \phi(S) + \phi(T).
\end{aligned} \tag{18}
$$

By definition of contra-polymatroids [22, 23], we conclude that the achievable rate region of a Gaussian multiple description scheme is a contra-polymatroid. The key advantage of this combinatorial property is that we can exactly characterize the vertices of the achievable rate region (13). Letting π to be a permutation on $\{1,\ \ldots,\ L\}$, define

$$
\begin{aligned}
& b_1^{(\pi)} \overset{\text{def}}{=} \phi(\{\pi(1)\}), \\
& b_i^{(\pi)} \overset{\text{def}}{=} \phi(\{\pi(1), \pi(2), \ldots, \pi(i)\}) - \phi(\{\pi(1), \pi(2), \ldots, \pi(i-1)\}), \quad i = 2,\ \ldots,\ L
\end{aligned} \tag{19}
$$

and $\mathbf{b}^{(\pi)} = \left(b_1^{(\pi)}, \ldots, b_L^{(\pi)} \right)$, then the $L!$ points $\{\mathbf{b}^{(\pi)}, \pi \text{ a permutation}\}$ are the vertices of the contra-polymatroid (13).

4 Lower Bound to the Sum Rate

By fairly procedural steps, we have the following lower bound to the sum rate of the multiple descriptions:

$$\begin{aligned} n\sum_{l=1}^{L} R_l \geq \sum_{l=1}^{L} H(C_l) &= \left[\sum_{l=1}^{L} H(C_l)\right] - H(C_1,\ \ldots,\ C_L|\mathbf{x}^n) \\ &= \left[\sum_{l=1}^{L} H(C_l)\right] - H(C_1,\cdots,C_L) + H(C_1,\ \ldots,\ C_L) - H(C_1,\ \ldots,\ C_L|\mathbf{x}^n) \\ &= I(C_1;C_2;\ldots;C_L) + I(C_1,\ \ldots,\ C_L;\mathbf{x}^n), \end{aligned} \tag{20}$$

where we have defined

$$I(C_1;C_2;\ldots;C_L) \stackrel{\text{def}}{=} \left[\sum_{l=1}^{L} H(C_l)\right] - H(C_1,\ \ldots,\ C_L) = \sum_{l=2}^{L} I(C_l;C_1\ldots C_{l-1}), \tag{21}$$

and called it the symmetric mutual information between $C_1,\ \ldots,\ C_L$. Note that $I(C_1;C_2;\ldots;C_L) \geq 0$ and is also well defined even when $C_1,\ \ldots,\ C_L$ are continuous random variables. We have the following information theoretic inequality which gives a lower bound to the sum of symmetric mutual information between $(C_1, C_2, \ldots, C_L)$ and mutual information between $C_1, C_2, \ldots, C_L$ and $\mathbf{x}^n$ for given covariance constraints.

Lemma 2. *Let* $\mathbf{x}^n = (\mathbf{x}[1],\ \ldots,\ \mathbf{x}[n])$*, where* $\mathbf{x}[m]$*'s are i.i.d.* $\mathcal{N}(\mathbf{0}, \mathbf{K}_x)$ *Gaussian random vectors for* $m = 1,\ \ldots,\ n$*. Let* $C_1,\ \ldots,\ C_L$ *be random variables jointly distributed with* $\mathbf{x}^n$*. Let* $\hat{\mathbf{x}}_0^n = \mathbb{E}[\mathbf{x}^n|C_1,\ \ldots,\ C_L]$ *and* $\hat{\mathbf{x}}_l^n = \mathbb{E}[\mathbf{x}^n|C_l]$ *for* $l = 1,\ \ldots,\ L$*. Given positive definite matrices* $\mathbf{D}_1,\ \ldots,\ \mathbf{D}_L,\ \mathbf{D}_0$*, if*

$$\begin{aligned} &\frac{1}{n}\sum_{m=1}^{n} \mathbb{E}[(\mathbf{x}[m] - \hat{\mathbf{x}}_l[m])^t(\mathbf{x}[m] - \hat{\mathbf{x}}_l[m])] \preccurlyeq \mathbf{D}_l, \quad l = 1,\ \ldots,\ L, \\ &\frac{1}{n}\sum_{m=1}^{n} \mathbb{E}[(\mathbf{x}[m] - \hat{\mathbf{x}}_0[m])^t(\mathbf{x}[m] - \hat{\mathbf{x}}_0[m])] \preccurlyeq \mathbf{D}_0, \end{aligned} \tag{22}$$

then

$$I(C_1;C_2;\ldots;C_L) + I(C_1,\ \ldots,\ C_L;\mathbf{x}^n) \geq \sup_{\mathbf{K}_z \succ \mathbf{0}} \frac{n}{2} \log \frac{|\mathbf{K}_x||\mathbf{K}_x + \mathbf{K}_z|^{(L-1)}|\mathbf{D}_0 + \mathbf{K}_z|}{|\mathbf{D}_0| \prod_{l=1}^{L} |\mathbf{D}_l + \mathbf{K}_z|}. \tag{23}$$

Furthermore, there exists a Gaussian distribution of $(C_1, \ldots, C_L, \mathbf{x}^n)$ *such that the inequality in* (23) *is tight.*

Proof. See Appendix B. □

This is a fundamental information-theoretic inequality which involves only the joint distribution[4] between $C_1, C_2, \ldots, C_L$ and $\mathbf{x}^n$ and bounds on mean square error estimation of $\mathbf{x}^n$ from $C_1, C_2, \ldots, C_L$. We can now use Lemma 2 to derive a lower bound to the sum rate

$$\sum_{l=1}^{L} R_l \geq \sup_{\mathbf{K}_z \succ \mathbf{0}} \frac{1}{2} \log \frac{|\mathbf{K}_x||\mathbf{K}_x + \mathbf{K}_z|^{(L-1)}|\mathbf{D}_0 + \mathbf{K}_z|}{|\mathbf{D}_0| \prod_{l=1}^{L} |\mathbf{D}_l + \mathbf{K}_z|}. \tag{24}$$

By letting $L = 1$ in the lemma above, we can derive a simple lower bound to the rate of the individual descriptions as well:

$$\begin{aligned} R_l &\geq \frac{1}{n} H(C_l) = \frac{1}{n}\big(H(C_l) - H(C_l|\mathbf{x}^n)\big) \\ &= \frac{1}{n} I(\mathbf{x}^n; C_l) \\ &\geq \frac{1}{2} \log \frac{|\mathbf{K}_x|}{|\mathbf{D}_l|}, \quad l = 1, \ \ldots, \ L. \end{aligned} \tag{25}$$

This bound is actually the point-to-point rate-distortion function for individual receivers, since each individual receiver only faces a point-to-point compression problem.

From the proof in Appendix B we can see that for any positive definite $\mathbf{K}_z$,

$$\frac{1}{2} \log \frac{|\mathbf{K}_x||\mathbf{K}_x + \mathbf{K}_z|^{(L-1)}|\mathbf{D}_0 + \mathbf{K}_z|}{|\mathbf{D}_0| \prod_{l=1}^{L} |\mathbf{D}_l + \mathbf{K}_z|}$$

is a lower bound to the sum rate of the multiple descriptions. Two special choices of $\mathbf{K}_z$ are of particular interest:

- Letting $\mathbf{K}_z = \epsilon \mathbf{I}_N$ and $\epsilon \to 0^+$, we have the following lower bound:

$$\sum_{l=1}^{L} R_l \geq \frac{1}{2} \log \frac{|\mathbf{K}_x|^L}{|\mathbf{D}_1| \ldots |\mathbf{D}_L|}. \tag{26}$$

  This bound is actually the summation of the bounds on the individual rates.

[4] This inequality holds even when $C_1, C_2, \ldots, C_L$ are not simply functions of $\mathbf{x}^n$ and can also be continuous random variables.

- Letting some eigenvalues of $\mathbf{K}_z$ to go to infinity, we have the following lower bound:

$$\sum_{l=1}^{L} R_l \geq \frac{1}{2} \log \frac{|\mathbf{K}_x|}{|\mathbf{D}_0|}. \tag{27}$$

  This bound is the point-to-point rate-distortion function when we only have the central distortion constraint.

We will see later that for some distortion constraints $(\mathbf{D}_1, \ldots, \mathbf{D}_L, \mathbf{D}_0)$, (26) and (27) can be tight.

5 Comparison of Upper Bound and the Lower Bound

Our goal is to show that the Gaussian description scheme achieves the lower bound to the sum rate, i.e., we need to show that two optimization problems (127) and (6) have the same optimal value. In general it does not seem facile to do a direct calculation and comparison. We forgo this strategy and, instead, provide an alternative characterization of the achievable sum rate using which a comparison with the lower bound is much easier.

Similar to the derivation of the lower bound (in Appendix B), we consider an $\mathcal{N}(0, \mathbf{K}_z)$ Gaussian random vector $\mathbf{z}$, independent of $\mathbf{x}$ and all $\mathbf{w}_l$'s. Defining $\mathbf{y} = \mathbf{x} + \mathbf{z}$, we have the following achievable sum rate:

$$\begin{aligned}
\sum_{l=1}^{L} R_l &= \left[\sum_{l=1}^{L} h(\mathbf{u}_l)\right] - h(\mathbf{u}_1, \ldots, \mathbf{u}_L|\mathbf{x}) \\
&= \left[\sum_{l=1}^{L} h(\mathbf{u}_l)\right] - h(\mathbf{u}_1, \ldots, \mathbf{u}_L) + h(\mathbf{u}_1, \ldots, \mathbf{u}_L) - h(\mathbf{u}_1, \ldots, \mathbf{u}_L|\mathbf{x}) \\
&= \left[\sum_{l=1}^{L} h(\mathbf{u}_l)\right] - h(\mathbf{u}_1, \cdots, \mathbf{u}_L) + I(\mathbf{u}_1, \ldots, \mathbf{u}_L; \mathbf{x}) \\
&\overset{(a)}{\geq} \left[\sum_{l=1}^{L} h(\mathbf{u}_l)\right] - h(\mathbf{u}_1, \cdots, \mathbf{u}_L) + I(\mathbf{u}_1, \ldots, \mathbf{u}_L; \mathbf{x}) \\
&\quad - \left(\left[\sum_{l=1}^{L} h(\mathbf{u}_l|\mathbf{y})\right] - h(\mathbf{u}_1, \ldots, \mathbf{u}_L|\mathbf{y})\right) \\
&= \left[\sum_{l=1}^{L} \big(h(\mathbf{y}) - h(\mathbf{y}|\mathbf{u}_l)\big)\right] - h(\mathbf{y}) + h(\mathbf{y}|\mathbf{u}_1, \ldots, \mathbf{u}_L) + h(\mathbf{x}) - h(\mathbf{x}|\mathbf{u}_1, \ldots, \mathbf{u}_L)
\end{aligned}$$

$$= h(\mathbf{x}) + (L-1)h(\mathbf{y}) - \left[\sum_{l=1}^{L} h(\mathbf{y}|\mathbf{u}_l)\right] + h(\mathbf{y}|\mathbf{u}_1, \ldots, \mathbf{u}_L) - h(\mathbf{x}|\mathbf{u}_1, \ldots, \mathbf{u}_L)$$

$$= \frac{1}{2}\log \frac{\left|\mathbf{K}_x\right|\left|\mathbf{K}_x + \mathbf{K}_z\right|^{(L-1)}\left|\mathrm{Cov}[\mathbf{x}|\mathbf{u}_1, \ldots, \mathbf{u}_L] + \mathbf{K}_z\right|}{\left|\mathrm{Cov}[\mathbf{x}|\mathbf{u}_1, \ldots, \mathbf{u}_L]\right| \prod_{l=1}^{L}\left|\mathrm{Cov}[\mathbf{x}|\mathbf{u}_l] + \mathbf{K}_z\right|}, \tag{28}$$

where the last step is from a procedural Gaussian MMSE calculation.

Note that if we have

$$\left[\sum_{l=1}^{L} h(\mathbf{u}_l|\mathbf{y})\right] - h(\mathbf{u}_1, \ldots, \mathbf{u}_L|\mathbf{y}) = 0, \tag{29}$$

then (a) in (28) is actually an equality. Thus, if our choice of $\mathbf{K}_w$ and $\mathbf{K}_z$ satisfy the following two conditions:

- (29) is true.
- distortion constraints are met with equality, i.e.,

$$\begin{aligned} &\mathrm{Cov}[\mathbf{x}|\mathbf{u}_l] = \mathbf{D}_l, \quad l = 1, \ldots, L, \\ &\mathrm{Cov}[\mathbf{x}|\mathbf{u}_1, \ldots, \mathbf{u}_L] = \mathbf{D}_0, \end{aligned} \tag{30}$$

then the upper bound matches the lower bound and we have characterized the sum rate.

In the following we examine under what circumstances the above two conditions are true. First, we give a necessary and sufficient condition for (29) to be true.

Proposition 2. *There exists a positive definite $\mathbf{K}_z$ such that (29) is true if and only if $\mathbf{K}_w$, the covariance matrix of $(\mathbf{w}_1, \cdots, \mathbf{w}_L)$, takes the following form*

$$\mathbf{K}_w = \begin{pmatrix} \mathbf{K}_{w_1} & -\mathbf{A} & -\mathbf{A} & \ldots & -\mathbf{A} \\ -\mathbf{A} & \mathbf{K}_{w_2} & -\mathbf{A} & \ldots & -\mathbf{A} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ -\mathbf{A} & \ldots & -\mathbf{A} & \mathbf{K}_{w_{L-1}} & -\mathbf{A} \\ -\mathbf{A} & \ldots & -\mathbf{A} & -\mathbf{A} & \mathbf{K}_{w_L} \end{pmatrix}, \tag{31}$$

where

$$\mathbf{A} = \mathbf{K}_x - \mathbf{K}_x(\mathbf{K}_x + \mathbf{K}_z)^{-1}\mathbf{K}_x \tag{32}$$

for this covariance matrix $\mathbf{K}_z$.

Proof. See Appendix C. □

Next, we look at the conditions for (30) to be true. From (16), we have

$$\begin{aligned} \mathbf{D}_l^{-1} &= \mathrm{Cov}[\mathbf{x}|\mathbf{u}_l]^{-1} = \mathbf{K}_x^{-1} + \mathbf{K}_{w_l}^{-1}, \quad l = 1,\ \ldots,\ L \\ \mathbf{D}_0^{-1} &= \mathrm{Cov}[\mathbf{x}|\mathbf{u}_1,\ \ldots,\ \mathbf{u}_L]^{-1} = \mathbf{K}_x^{-1} + (\mathbf{I}_N,\ \ldots,\ \mathbf{I}_N)\mathbf{K}_w^{-1}(\mathbf{I}_N,\ \ldots,\ \mathbf{I}_N)^t. \end{aligned} \tag{33}$$

$(\mathbf{I}_N,\ \mathbf{I}_N,\ \ldots,\ \mathbf{I}_N)\mathbf{K}_w^{-1}(\mathbf{I}_N,\ \mathbf{I}_N,\ \ldots,\ \mathbf{I}_N)^t$ is calculated in the following lemma.

***Lemma* 3.** *Let* $\mathbf{K}_w$ *be given by* (31) *but with an arbitrary* $\mathbf{A} \succeq \mathbf{0}$. *If* $\mathbf{K}_w \succ \mathbf{0}$, *then*

$$(\mathbf{I}_N,\ \mathbf{I}_N,\ \ldots,\ \mathbf{I}_N)\mathbf{K}_w^{-1}(\mathbf{I}_N,\ \mathbf{I}_N,\ \ldots,\ \mathbf{I}_N)^t = \left[\left(\sum_{l=1}^{L}(\mathbf{K}_{w_l} + \mathbf{A})^{-1}\right)^{-1} - \mathbf{A}\right]^{-1}. \tag{34}$$

Proof. See Appendix D. □

Using this lemma, from (33) we arrive at

$$\left[(\mathbf{D}_0^{-1} - \mathbf{K}_x^{-1})^{-1} + \mathbf{A}\right]^{-1} = \sum_{l=1}^{L}\left[(\mathbf{D}_l^{-1} - \mathbf{K}_x^{-1})^{-1} + \mathbf{A}\right]^{-1}. \tag{35}$$

Defining

$$\mathbf{K}_{w_0} = (\mathbf{D}_0^{-1} - \mathbf{K}_x^{-1})^{-1}, \tag{36}$$

(33) is equivalent to

$$\left[\mathbf{K}_{w_0} + \mathbf{A}\right]^{-1} = \sum_{l=1}^{L}\left[\mathbf{K}_{w_l} + \mathbf{A}\right]^{-1}. \tag{37}$$

Thus, if there exists a positive definite solution $\mathbf{A}$ to (37), and the corresponding $\mathbf{K}_w$ is positive definite, then the distortion constraints are met with equality, i.e., (30) holds. It turns out that as long as $\mathbf{A}$ is a solution to (37), the resulting $\mathbf{K}_w$ is always positive definite; we state this formally below.

***Lemma* 4.** *If for some* $\mathbf{K}_{w_0} \succ \mathbf{0}$ *and* $\mathbf{A} \succ \mathbf{0}$ (37) *is true, then the covariance matrix* $\mathbf{K}_w$ *defined in* (31) *is positive definite.*

Proof. See Appendix E. □

We summarize the state of affairs in the following theorem.

***Theorem* 4.** *Given distortion constraints* $(\mathbf{D}_1,\ \ldots,\ \mathbf{D}_L, \mathbf{D}_0)$, *let*

$$\mathbf{K}_{w_l} = (\mathbf{D}_l^{-1} - \mathbf{K}_x^{-1})^{-1}, \quad l = 0,\ 1,\ \ldots,\ L. \tag{38}$$

If there exists a solution $\mathbf{A}^*$ *to* (37) *and* $\mathbf{0} \prec \mathbf{A}^* \prec \mathbf{K}_x$, *then the Gaussian description scheme with* $\mathbf{K}_w$ *defined in* (31) *with* $\mathbf{A} = \mathbf{A}^*$ *achieves the optimal sum rate, and the optimal* $\mathbf{K}_z$ *for lower bound* (6) *is* $\mathbf{K}_z = \mathbf{K}_x(\mathbf{K}_x - \mathbf{A}^*)^{-1}\mathbf{K}_x - \mathbf{K}_x$.

Now we proceed to the proof of Theorem 2. From Theorem 4 we know that the Gaussian description scheme achieves the optimal sum rate if the given distortion constraints $(\mathbf{D}_1, \ldots, \mathbf{D}_L, \mathbf{D}_0)$ satisfy the condition for Theorem 4, and we can calculate the optimal $\mathbf{K}_w$ by solving a matrix equation. To complete the proof of Theorem 2, we need also consider the case that for arbitrarily given distortion constraints, (37) may not have a solution $\mathbf{A}^*$ such that $\mathbf{0} \prec \mathbf{A}^* \prec \mathbf{K}_x$. In the following, we show that in this case there exists a Gaussian description scheme that achieves the sum rate lower bound, and results in distortions $(\mathbf{D}_1^*, \ldots, \mathbf{D}_L^*, \mathbf{D}_0^*)$ such that $\mathbf{D}_l^* \preccurlyeq \mathbf{D}_l$ for $l = 0, 1, \ldots, L$.

Proof of Theorem 2. We draw the connection between solution to (37) and the solution to an optimization problem. First, note that by a linear transformation at the encoders and the decoders, we have the following result on rate region for multiple description with individual and central receivers.

***Proposition* 3.**

$$R_*(\mathbf{K}_x,\ \mathbf{D}_1,\ \ldots,\ ,\mathbf{D}_L,\ \mathbf{D}_0) = R_*(\mathbf{I}_N,\ \mathbf{K}_x^{-\frac{1}{2}}\mathbf{D}_1\mathbf{K}_x^{-\frac{1}{2}},\ \ldots,\ ,\mathbf{K}_x^{-\frac{1}{2}}\mathbf{D}_L\mathbf{K}_x^{-\frac{1}{2}},\ \mathbf{K}_x^{-\frac{1}{2}}\mathbf{D}_0\mathbf{K}_x^{-\frac{1}{2}}). \tag{39}$$

Thus, throughout this proof we will suppose, for notation simplicity, that $\mathbf{K}_x = \mathbf{I}_N$.

Given distortion constraints $(\mathbf{D}_1, \ldots\ \mathbf{D}_L,\ \mathbf{D}_0)$, let

$$\mathbf{K}_{w_l} = (\mathbf{D}_l^{-1} - \mathbf{I}_N)^{-1}, \quad l = 0,\ 1,\ \ldots,\ L, \tag{40}$$

and define

$$f(\mathbf{A}) \stackrel{\text{def}}{=} \quad [\mathbf{K}_{w_0} + \mathbf{A}]^{-1} - \sum_{l=1}^{L} [\mathbf{K}_{w_l} + \mathbf{A}]^{-1}, \tag{41}$$

$$F(\mathbf{A}) \stackrel{\text{def}}{=} \ \log|\mathbf{K}_{w_0} + \mathbf{A}| - \sum_{l=1}^{L} \log|\mathbf{K}_{w_l} + \mathbf{A}|. \tag{42}$$

Note that

$$\frac{dF(\mathbf{A})}{d\mathbf{A}} = f(\mathbf{A}). \tag{43}$$

Consider the following optimization problem:

$$\max_{\mathbf{0}\preccurlyeq\mathbf{A}\preccurlyeq\mathbf{I}_N} \qquad F(\mathbf{A}). \tag{44}$$

Now, since $F(\mathbf{A})$ is a continuous map and $\mathbf{0} \preccurlyeq \mathbf{A} \preccurlyeq \mathbf{I}_N$ is a compact set, there exists an optimal solution $\mathbf{A}^*$ to (44) where $\mathbf{A}^*$ satisfies the Karush-Kuhn-Tucker (KKT) conditions [28, Section 5.5.3]: there exist $\mathbf{\Lambda}_1 \succcurlyeq \mathbf{0}$ and $\mathbf{\Lambda}_2 \succcurlyeq \mathbf{0}$ such that

$$f(\mathbf{A}^*) + \mathbf{\Lambda}_1 - \mathbf{\Lambda}_2 = \mathbf{0} \tag{45}$$

$$\mathbf{\Lambda}_1\mathbf{A}^* = \mathbf{0} \tag{46}$$

$$\mathbf{\Lambda}_2(\mathbf{A}^* - \mathbf{I}_N) = \mathbf{0}. \tag{47}$$

Now $\mathbf{A}^*$ falls into the following four cases.

Case 1: $\mathbf{0} \prec \mathbf{A}^* \prec \mathbf{I}$. Alternatively, 0 and 1 are not eigenvalues of $\mathbf{A}^*$. In this case, $\mathbf{\Lambda}_1 = \mathbf{0}$ and $\mathbf{\Lambda}_2 = \mathbf{0}$; thus the KKT conditions in (45) reduce to

$$f(\mathbf{A}^*) = \mathbf{0}.$$

Equivalently,

$$[\mathbf{K}_{w_0} + \mathbf{A}^*]^{-1} = \sum_{l=1}^{L} [\mathbf{K}_{w_l} + \mathbf{A}^*]^{-1}. \tag{48}$$

From Theorem 4, the Gaussian description scheme with covariance matrix for $\mathbf{w}_1, \ldots, \mathbf{w}_L$ being

$$\mathbf{K}_w = \begin{pmatrix} \mathbf{K}_{w_1} & -\mathbf{A}^* & -\mathbf{A}^* & \ldots & -\mathbf{A}^* \\ -\mathbf{A}^* & \mathbf{K}_{w_2} & -\mathbf{A}^* & \ldots & -\mathbf{A}^* \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ -\mathbf{A}^* & \ldots & -\mathbf{A}^* & \mathbf{K}_{w_{L-1}} & -\mathbf{A}^* \\ -\mathbf{A}^* & \ldots & -\mathbf{A}^* & -\mathbf{A}^* & \mathbf{K}_{w_L} \end{pmatrix} \tag{49}$$

achieves the lower bound to the sum rate. Thus in this case, we have characterized the optimality of the Gaussian description scheme parameterized by (49) in terms of achieving the sum rate.

Case 2: $\mathbf{0} \preccurlyeq \mathbf{A}^* \prec \mathbf{I}_N$. Alternatively, some eigenvalues of $\mathbf{A}^*$ are 0, but no eigenvalues of $\mathbf{A}^*$ are 1. Thus $\mathbf{\Lambda}_2 = \mathbf{0}$ and the KKT conditions in (45) reduce to

$$(\mathbf{K}_{w_0} + \mathbf{A}^*)^{-1} - \sum_{l=1}^{L} (\mathbf{K}_{w_l} + \mathbf{A}^*)^{-1} + \mathbf{\Lambda}_1 = \mathbf{0}, \tag{50}$$

for some $\mathbf{\Lambda}_1 \succcurlyeq \mathbf{0}$ satisfying $\mathbf{\Lambda}_1 \mathbf{A}^* = \mathbf{0}$. The *key idea* now is to see that the distortion constraint on the central receiver is too loose and we can in fact achieve a *lower* distortion (in the sense of positive semidefinite ordering) for the same sum rate. We first identify this lower distortion: defining

$$\mathbf{K}_{w_0}^* = \left(\mathbf{K}_{w_0}^{-1} + \mathbf{\Lambda}_1\right)^{-1}, \tag{51}$$

consider the smaller distortion matrix on the central receiver

$$\mathbf{D}_0^* = \left({\mathbf{K}_{w_0}^*}^{-1} + \mathbf{I}_N\right)^{-1} = \left(\mathbf{I}_N + \mathbf{K}_{w_0}^{-1} + \mathbf{\Lambda}_1\right)^{-1} = (\mathbf{D}_0^{-1} + \mathbf{\Lambda}_1)^{-1} \prec \mathbf{D}_0. \tag{52}$$

This new distortion matrix on the central receiver satisfies two key properties, that we state as a lemma.

***Lemma* 5.**

$$(\mathbf{K}_{w_0} + \mathbf{A}^*)^{-1} + \mathbf{\Lambda}_1 = (\mathbf{K}_{w_0}^* + \mathbf{A}^*)^{-1}, \tag{53}$$

$$\frac{|\mathbf{D}_0 + \mathbf{K}_z|}{|\mathbf{D}_0|} = \frac{|\mathbf{D}_0^* + \mathbf{K}_z|}{|\mathbf{D}_0^*|}. \tag{54}$$

Proof. See Appendix F. □

Comparing (50) with (53), we have

$$\left[\mathbf{K}_{w_0}^* + \mathbf{A}^*\right]^{-1} = \sum_{l=1}^{L} \left[\mathbf{K}_{w_l} + \mathbf{A}^*\right]^{-1}. \tag{55}$$

Now, the corresponding $\mathbf{K}_z = (\mathbf{I}_N - \mathbf{A}^*)^{-1} - \mathbf{I}_N$ is singular. If it had not been, then by Theorem 4 we could have concluded that the Gaussian description scheme achieves the lower bound to the sum rate. We now address this technical difficulty.

Our first observation is that there exists $\delta > 0$ such that for all $\epsilon \in (0, \delta)$ we have $\mathbf{0} \prec \mathbf{A} + \epsilon\mathbf{I}_N \prec \mathbf{I}_N$, and $0 \prec \mathbf{K}_{w_0}^* - \epsilon\mathbf{I}_N$, $\mathbf{0} \prec \mathbf{K}_{w_l} - \epsilon\mathbf{I}_N$, and we can rewrite (55) as

$$\left[(\mathbf{K}_{w_0}^* - \epsilon\mathbf{I}_N) + (\mathbf{A}^* + \epsilon\mathbf{I}_N)\right]^{-1} = \sum_{l=1}^{L} \left[(\mathbf{K}_{w_l} - \epsilon\mathbf{I}_N) + (\mathbf{A}^* + \epsilon\mathbf{I}_N)\right]^{-1}. \tag{56}$$

Thus if the distortion constraints were $(\mathbf{D}_1(\epsilon), \ldots, \mathbf{D}_L(\epsilon), \mathbf{D}_0(\epsilon))$ with

$$\begin{aligned} \mathbf{D}_l(\epsilon) &= \left[(\mathbf{K}_{w_l} - \epsilon\mathbf{I}_N)^{-1} + \mathbf{I}_N\right]^{-1}, \quad l = 1, \ldots, L, \\ \mathbf{D}_0(\epsilon) &= \left[(\mathbf{K}_{w_0}^* - \epsilon\mathbf{I}_N)^{-1} + \mathbf{I}_N\right]^{-1}, \end{aligned}$$

then $\mathbf{A}^* + \epsilon\mathbf{I}_N$ is a solution to (56). This situation corresponds to that discussed in Case I; we can conclude that sum rate for this modified distortion multiple description problem is

$$\frac{1}{2} \log \frac{|\mathbf{I}_N + \mathbf{K}_z(\epsilon)|^{(L-1)} |\mathbf{D}_0(\epsilon) + \mathbf{K}_z(\epsilon)|}{|\mathbf{D}_0(\epsilon)| \prod_{l=1}^{L} |\mathbf{D}_l(\epsilon) + \mathbf{K}_z(\epsilon)|}, \tag{57}$$

where $\mathbf{K}_z(\epsilon) = \left[\mathbf{I}_N - (\mathbf{A}^* + \epsilon\mathbf{I}_N)\right]^{-1} - \mathbf{I}_N$. We would like to let ϵ approach zero and consider the limiting multiple description problem. In particular, we show that

$$\mathbf{D}_l(\epsilon) \rightarrow \mathbf{D}_l, \quad l = 1, \ldots, L, \tag{58}$$

$$\mathbf{D}_0(\epsilon) \rightarrow \mathbf{D}_0^*, \tag{59}$$

as $\epsilon \rightarrow 0$ in Appendix G. Further, we show that

$$\mathbf{K}_z(\epsilon) \rightarrow (\mathbf{I}_N - \mathbf{A}^*)^{-1} - \mathbf{I}_N, \tag{60}$$

as $\epsilon \rightarrow 0$ in Appendix H. Thus we can conclude that the sum rate approaches, using (54),

$$\frac{1}{2} \log \frac{|\mathbf{I}_N + \mathbf{K}_z|^{(L-1)} |\mathbf{D}_0 + \mathbf{K}_z|}{|\mathbf{D}_0| \prod_{l=1}^{L} |\mathbf{D}_l + \mathbf{K}_z|}, \tag{61}$$

as $\epsilon \to 0$; here $\mathbf{K}_z = (\mathbf{I}_N - \mathbf{A}^*)^{-1} - \mathbf{I}_N$. We observe that this sum rate is achievable using the Gaussian multiple description scheme. Further, this sum rate is identical to the lower bound to sum rate for the original distortions $(\mathbf{D}_1, \ldots, \mathbf{D}_L, \mathbf{D}_0)$. Thus we conclude the optimality of the Gaussian description scheme in this case as well.

Case 3: $\mathbf{0} \prec \mathbf{A}^* \preccurlyeq \mathbf{I}_N$. Alternatively, some eigenvalues of $\mathbf{A}^*$ are 1, but no eigenvalues of $\mathbf{A}^*$ are 0. In this case, the $\mathbf{\Lambda}_1 = \mathbf{0}$ and the KKT conditions in (45) reduce to

$$(\mathbf{K}_{w_0} + \mathbf{A}^*)^{-1} - \sum_{l=1}^{L} (\mathbf{K}_{\mathbf{w}_l} + \mathbf{A}^*)^{-1} - \mathbf{\Lambda}_2 = 0, \tag{62}$$

for some $\mathbf{\Lambda}_2 \succcurlyeq 0$ satisfying $\mathbf{\Lambda}_2(\mathbf{A}^* - \mathbf{I}_N) = \mathbf{0}$. Defining

$$\mathbf{K}_{w_l}^* = \left[(\mathbf{K}_{\mathbf{w}_l} + \mathbf{I}_N)^{-1} + \mathbf{\Lambda}_2\right]^{-1} - \mathbf{I}_N, \tag{63}$$

we have, as in (53), that

$$(\mathbf{K}_{\mathbf{w}_l} + \mathbf{A}^*)^{-1} + \mathbf{\Lambda}_2 = (\mathbf{K}_{\mathbf{w}_l}^* + \mathbf{A}^*)^{-1}. \tag{64}$$

The observation

$$(\mathbf{K}_{\mathbf{w}_l} + \mathbf{A}^*)^{-1} + \mathbf{\Lambda}_2 = [(\mathbf{K}_{\mathbf{w}_l} + \mathbf{I}_N) + (\mathbf{A}^* - \mathbf{I}_N)]^{-1} + \mathbf{\Lambda}_2, \tag{65}$$

combined with the proof of (53) suffices to justify (64). Now, from (64),

$$(\mathbf{K}_{w_0} + \mathbf{A}^*)^{-1} - \sum_{l=1}^{L-1} (\mathbf{K}_{w_l} + \mathbf{A}^*)^{-1} - (\mathbf{K}_{w_L}^* + \mathbf{A}^*)^{-1} = \mathbf{0}. \tag{66}$$

As in the previous case, the key step is to identify smaller distortion matrices at each of the individual receivers (ordered in the positive semidefinite sense) that are achievable at the same sum rate:

$$\mathbf{D}_l^* = \left[{\mathbf{K}_{w_l}^*}^{-1} + \mathbf{I}_N\right]^{-1}, \quad l = 1, \ldots, L. \tag{67}$$

To see that this is indeed a smaller distortion matrix, observe that since $\mathbf{K}_w$ is positive definite, it follows that $\mathbf{K}_{w_l}^* \succ 0$ and

$$\begin{aligned}
\mathbf{D}_l^* &= \left[{\mathbf{K}_{w_l}^*}^{-1} + \mathbf{I}_N\right]^{-1} \\
&= \left[\left(\left((\mathbf{K}_{w_l} + \mathbf{I}_N)^{-1} + \mathbf{\Lambda}_2\right)^{-1} - \mathbf{I}_N\right)^{-1} + \mathbf{I}_N\right]^{-1} \\
&= \left[\mathbf{I}_N - (\mathbf{K}_{w_l} + \mathbf{I}_N)^{-1} - \mathbf{\Lambda}_2\right] \\
&= [\mathbf{I}_N + \mathbf{K}_{w_l}]^{-1} - \mathbf{\Lambda}_2 \\
&= \mathbf{D}_l - \mathbf{\Lambda}_2, \quad l = 1, \ldots, L.
\end{aligned} \tag{68}$$

Since $\mathbf{\Lambda}_2 \succcurlyeq \mathbf{0}$, it follows that $\mathbf{0} \prec \mathbf{D}_l^* \preccurlyeq \mathbf{D}_l,\ l = 1,\ \ldots,\ L$. Define

$$\begin{aligned} \mathbf{D}_l(\epsilon) &= \left[(\mathbf{K}_{w_l} + \epsilon\mathbf{I}_N)^{-1} + \mathbf{I}_N\right]^{-1}, \quad l = 0,\ 1,\ \ldots,\ L-1, \\ \mathbf{D}_L(\epsilon) &= \left[(\mathbf{K}_{w_L}^* + \epsilon\mathbf{I}_N)^{-1} + \mathbf{I}_N\right]^{-1}, \end{aligned} \tag{69}$$

then there exists $\delta > 0$ such that for all $\epsilon \in (0, \delta)$ we have $\mathbf{0} \prec \mathbf{A}^* - \epsilon\mathbf{I}_N \prec \mathbf{I}_N$, and $\mathbf{0} \prec \mathbf{D}_l(\epsilon) \prec \mathbf{I}_N$. We can rewrite (66) as

$$[(\mathbf{K}_{w_0} + \epsilon\mathbf{I}_N) + (\mathbf{A}^* - \epsilon\mathbf{I}_N)]^{-1} = \sum_{l=1}^{L-1} [(\mathbf{K}_{w_l} + \epsilon\mathbf{I}_N) + (\mathbf{A}^* - \epsilon\mathbf{I}_N)]^{-1} + \left[(\mathbf{K}_{w_L}^* + \epsilon\mathbf{I}_N) + (\mathbf{A}^* - \epsilon\mathbf{I}_N)\right]^{-1}. \tag{70}$$

Thus if the distortion constraints were $(\mathbf{D}_1(\epsilon),\ \ldots,\ \mathbf{D}_L(\epsilon),\ \mathbf{D}_0(\epsilon))$, then $\mathbf{A}^* - \epsilon\mathbf{I}_N$ is a solution to (70). This situation corresponds to that discussed in Case I; we conclude that the sum rate for this modified distortion multiple description problem is

$$\frac{1}{2} \log \frac{|\mathbf{I}_N + \mathbf{K}_z(\epsilon)|^{(L-1)} |\mathbf{D}_0(\epsilon) + \mathbf{K}_z(\epsilon)|}{|\mathbf{D}_0(\epsilon)| \prod_{l=1}^{L} |\mathbf{D}_l(\epsilon) + \mathbf{K}_z(\epsilon)|}, \tag{71}$$

where $\mathbf{K}_z(\epsilon) = [\mathbf{I}_N - (\mathbf{A}^* - \epsilon\mathbf{I}_N)]^{-1} - \mathbf{I}_N$. We would like to let ϵ approach zero and consider the limiting multiple description problem. Similar to equations (58) and (59), we have

$$\begin{aligned} \mathbf{D}_l(\epsilon) &\to \mathbf{D}_l, \quad l = 1, \ldots L, \\ \mathbf{D}_0(\epsilon) &\to \mathbf{D}_0^*. \end{aligned} \tag{72}$$

Further, we show that

$$\lim_{\epsilon \to 0} \frac{|\mathbf{I}_N + \mathbf{K}_z(\epsilon)|^{(L-1)} |\mathbf{D}_0(\epsilon) + \mathbf{K}_z(\epsilon)|}{| \prod_{l=1}^{L} |\mathbf{D}_l(\epsilon) + \mathbf{K}_z(\epsilon)|} = 1 \tag{73}$$

in Appendix I. We can now conclude that the sum rate approaches

$$\frac{1}{2} \log \frac{1}{|\mathbf{D}_0|} \tag{74}$$

as ϵ approaches 0. In other words, the point-to-point rate-distortion function for central receiver with distortion $\mathbf{D}_0$ can be achieved by using the Gaussian description scheme, and the resulting distortion is $(\mathbf{D}_1,\ \ldots,\ \mathbf{D}_L^*,\ \mathbf{D}_0)$ where $\mathbf{0} \prec \mathbf{D}_L^* \preccurlyeq \mathbf{D}_L$. In conclusion, the Gaussian description scheme is also optimal in this case.

Case 4: $\mathbf{0} \preccurlyeq \mathbf{A}^* \preccurlyeq \mathbf{I}_N$. i.e., both 0 and 1 are eigenvalues of $\mathbf{A}^*$. In this case, the KKT conditions are: there exist $\mathbf{\Lambda}_1 \succcurlyeq 0$ and $\mathbf{\Lambda}_2 \succcurlyeq 0$ such that equations (45), (46) and (47) hold. We can combine equations (53) and (64) to get

$$(\mathbf{K}_{w_0}^* + \mathbf{A}^*)^{-1} = \sum_{l=1}^{L-1} (\mathbf{K}_{w_l} + \mathbf{A}^*)^{-1} + (\mathbf{K}_{w_L}^* + \mathbf{A}^*)^{-1}, \tag{75}$$

where

$$\begin{aligned} \mathbf{K}_{w_0}^* &= \left(\mathbf{K}_{w_0}^{-1} + \mathbf{\Lambda}_1\right)^{-1}, \\ \mathbf{K}_{w_L}^* &= \left[(\mathbf{K}_{w_L} + \mathbf{I}_N)^{-1} + \mathbf{\Lambda}_2\right]^{-1} - \mathbf{I}_N. \end{aligned} \tag{76}$$

As in cases 2 and 3, we want to show the optimality of the Gaussian multiple description scheme through a limiting procedure. We do this by first perturbing $\mathbf{A}^*$ so that it has no eigenvalue equal to 0 or 1 as follows.

Without loss of generality, suppose that $\mathbf{A}^*$ has p eigenvalues equal to 0 and q eigenvalues equal 1, where $p > 0$ and $q > 0$, and there exists $N \times N$ orthogonal matrix $\mathbf{Q}$ such that

$$\mathbf{Q}\mathbf{A}^*\mathbf{Q}^t = \text{diag}\{\underbrace{0, \ldots, 0}_{p}, \underbrace{1, \ldots, 1}_{q}, a_{p+q+1}, \ldots, a_N\}, \tag{77}$$

with $0 < a_{p+q+1} < 1, \ldots, 0 < a_N < 1$. We need to perturb the eigenvalues of $\mathbf{A}^*$ away from both 0 and 1. Towards this, we define two $N \times N$ diagonal matrices:

$$\begin{aligned} \mathbf{E}_1 &= \text{diag}(\underbrace{1, \ldots, 1}_{p}, \underbrace{0, \ldots, 0, 0, \ldots, 0}_{N-p}), \\ \mathbf{E}_2 &= \text{diag}(\underbrace{0, \ldots, 0}_{p}, \underbrace{1, \ldots, 1}_{q}, 0, \ldots, 0), \end{aligned} \tag{78}$$

Also define

$$\begin{aligned} \mathbf{A}^*(\epsilon_1, \epsilon_2) &= \mathbf{A}^* + \mathbf{Q}^t(\epsilon_1 \mathbf{E}_1 - \epsilon_2 \mathbf{E}_2)\mathbf{Q}, \\ \mathbf{K}_z(\epsilon_1, \epsilon_2) &= (\mathbf{I}_N - \mathbf{A}^*(\epsilon_1, \epsilon_2))^{-1} - \mathbf{I}_N, \\ \mathbf{K}_{w_l}(\epsilon_1, \epsilon_2) &= \mathbf{K}_{w_l} - \mathbf{Q}^t(\epsilon_1 \mathbf{E}_1 - \epsilon_2 \mathbf{E}_2)\mathbf{Q}, \quad l = 1, \ldots, L-1, \\ \mathbf{K}_{w_L}(\epsilon_1, \epsilon_2) &= \mathbf{K}_{w_L}^* - \mathbf{Q}^t(\epsilon_1 \mathbf{E}_1 - \epsilon_2 \mathbf{E}_2)\mathbf{Q}, \\ \mathbf{K}_{w_0}(\epsilon_1, \epsilon_2) &= \mathbf{K}_{w_0}^* - \mathbf{Q}^t(\epsilon_1 \mathbf{E}_1 - \epsilon_2 \mathbf{E}_2)\mathbf{Q}. \end{aligned} \tag{79}$$

Further, defining

$$\mathbf{D}_l(\epsilon_1, \epsilon_2) = (\mathbf{I}_N + \mathbf{K}_{w_l}(\epsilon_1, \epsilon_2))^{-1}, \quad l = 1, \ldots, L, \tag{80}$$

there exists $\delta > 0$ such that for all $\epsilon_1 \in (0, \delta)$ and $\epsilon_2 \in (0, \delta)$ we have $\mathbf{0} \prec \mathbf{A}^*(\epsilon_1, \epsilon_2) \prec \mathbf{I}_N$, and $\mathbf{0} \prec \mathbf{D}_l(\epsilon_1, \epsilon_2) \prec \mathbf{I}_N$. Now, we can rewrite (75) as

$$\Big[\mathbf{K}_{w_0}(\epsilon_1, \epsilon_2) + \mathbf{A}^*(\epsilon_1, \epsilon_2)\Big]^{-1} = \sum_{l=1}^{L} \Big[\mathbf{K}_{w_l}(\epsilon_1, \epsilon_2) + \mathbf{A}^*(\epsilon_1, \epsilon_2)\Big]^{-1}. \tag{81}$$

Thus if the distortion constraints were $(\mathbf{D}_1(\epsilon_1, \epsilon_2),\ \ldots,\ \mathbf{D}_L(\epsilon_1, \epsilon_2),\ \mathbf{D}_0(\epsilon_1, \epsilon_2))$, then $\mathbf{A}^*(\epsilon_1, \epsilon_2)$ is a solution to (81). This situation corresponds to that discussed in Case I; we conclude that the sum rate for this modified distortion multiple description problem is

$$\frac{1}{2} \log \frac{|\mathbf{I}_N + \mathbf{K}_z(\epsilon_1, \epsilon_2)|^{(L-1)} |\mathbf{D}_0(\epsilon_1, \epsilon_2) + \mathbf{K}_z(\epsilon_1, \epsilon_2)|}{|\mathbf{D}_0(\epsilon_1, \epsilon_2)| \prod_{l=1}^{L} |\mathbf{D}_l(\epsilon_1, \epsilon_2) + \mathbf{K}_z(\epsilon_1, \epsilon_2)|}, \tag{82}$$

where $\mathbf{K}_z(\epsilon_1, \epsilon_2) = [\mathbf{I}_N - \mathbf{A}^*(\epsilon_1, \epsilon_2)]^{-1} - \mathbf{I}_N$. We would like to let ϵ_1 and ϵ_2 approach zero and consider the limiting multiple description problem. Similar to equations (58) and (59), when ϵ_1 and ϵ_2 approach 0, we get

$$\begin{aligned} \mathbf{D}_l(\epsilon_1, \epsilon_2) &\to \mathbf{D}_l, \quad l = 1,\ \ldots,\ L-1, \\ \mathbf{D}_L(\epsilon_1, \epsilon_2) &\to \mathbf{D}_L^*, \\ \mathbf{D}_0(\epsilon_1, \epsilon_2) &\to \mathbf{D}_0^*, \end{aligned} \tag{83}$$

where $\mathbf{D}_L^* = \mathbf{D}_L - \mathbf{\Lambda}_2$ as in case 3 and $\mathbf{D}_0^* = [\mathbf{D}_0^{-1} + \mathbf{\Lambda}_1^{-1}]^{-1}$ as in case 2. Further, we show that

$$\lim_{\epsilon_2 \to 0} \lim_{\epsilon_1 \to 0} \frac{1}{2} \log \frac{|\mathbf{I}_N + \mathbf{K}_z(\epsilon_1, \epsilon_2)|^{(L-1)} |\mathbf{D}_0(\epsilon_1, \epsilon_2) + \mathbf{K}_z(\epsilon_1, \epsilon_2)|}{|\mathbf{D}_0(\epsilon_1, \epsilon_2)| \prod_{l=1}^{L} |\mathbf{D}_l(\epsilon_1, \epsilon_2) + \mathbf{K}_z(\epsilon_1, \epsilon_2)|} = \frac{1}{2} \log \frac{1}{|\mathbf{D}_0|} \tag{84}$$

in Appendix J. We conclude that the sum rate approaches

$$\frac{1}{2} \log \frac{1}{|\mathbf{D}_0|} \tag{85}$$

as ϵ_1 and ϵ_2 approach 0. Thus the point-to-point rate-distortion function for central receiver with distortion $\mathbf{D}_0$ can be achieved by using the Gaussian description scheme, and the resulting distortions are $(\mathbf{D}_1,\ \ldots,\ \mathbf{D}_L^*,\ \mathbf{D}_0^*)$ where $\mathbf{0} \prec \mathbf{D}_L^* \preccurlyeq \mathbf{D}_L$ and $\mathbf{0} \prec \mathbf{D}_0^* \preccurlyeq \mathbf{D}_0$. In other words, the Gaussian multiple description scheme is also optimal in this case. □

To summarize, we see that the Gaussian description scheme achieves the limiting sum rate. The limiting sum rate is the solution to an optimization problem. For some specific distortion constraints, the sum rate can be characterized as the solution to a matrix polynomial equation (Case 1). In the following we study two examples: the scalar Gaussian source and two descriptions of the vector Gaussian source.

6 Scalar Gaussian Source

Here we suppose that the information source is an i.i.d. sequence of $\mathcal{N}(0, \sigma_x^2)$ scalar Gaussian random variables. Let individual distortion constraints be $(d_1, \ldots, d_L)$ and the central distortion constraints be d_0, where $0 < d_0 < d_l < \sigma_x^2$ for $l = 1, \ldots, L$. Consider the Gaussian description scheme with the following covariance matrix for $w_1, \ldots, w_l$:

$$\mathbf{K}_w = \begin{pmatrix} \sigma_1^2 & -a & -a & \ldots & -a \\ -a & \sigma_2^2 & -a & \ldots & -a \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ -a & \ldots & -a & \sigma_{L-1}^2 & -a \\ -a & \ldots & -a & -a & \sigma_L^2 \end{pmatrix}. \tag{86}$$

Consider the condition for Theorem 4 to hold: to meet the individual distortion constraint with equality, we need

$$\sigma_l^2 = (d_l^{-1} - \sigma_x^{-2})^{-1} = \frac{d_l \sigma_x^2}{\sigma_x^2 - d_l}, \quad l = 1, \ldots, L. \tag{87}$$

Letting

$$\sigma_0^2 \stackrel{\text{def}}{=} (d_0^{-1} - \sigma_x^{-2})^{-1} = \frac{d_0 \sigma_x^2}{\sigma_x^2 - d_0}, \tag{88}$$

we need

$$\left[\sigma_0^2 + a\right]^{-1} = \sum_{l=1}^{L} \left[\sigma_l^2 + a\right]^{-1} \tag{89}$$

to have a solution $a^* \in (0, \sigma_x^2)$, to meet the central distortion constraint with equality. Towards this, define

$$f(a) \stackrel{\text{def}}{=} \frac{1}{\sigma_0^2 + a} - \sum_{l=1}^{L} \frac{1}{\sigma_l^2 + a}, \tag{90}$$

and we have

$$\begin{aligned} f(0) &= \frac{1}{\sigma_0^2} - \sum_{l=1}^{L} \frac{1}{\sigma_l^2} = \frac{1}{d_0} + \frac{L-1}{\sigma_x^2} - \sum_{l=1}^{L} \frac{1}{d_l}, \\ f(\sigma_x^2) &= \frac{1}{\sigma_0^2 + \sigma_x^2} - \sum_{l=1}^{L} \frac{1}{\sigma_l^2 + \sigma_x^2} = \frac{1}{\sigma_x^4} \left(\sum_{l=1}^{L} d_l - d_0 - (L-1)\sigma_x^2 \right). \end{aligned} \tag{91}$$

Using induction, we can show that

$$\left(\sum_{l=1}^{L} \frac{1}{d_l} - \frac{L-1}{\sigma_x^2} \right)^{-1} \geq \sum_{l=1}^{L} d_l - (L-1)\sigma_x^2. \tag{92}$$

Thus we have

$$f(0) \leq 0 \Rightarrow f(\sigma_x^2) \leq 0,$$
$$f(\sigma_x^2) \geq 0 \Rightarrow f(0) \geq 0.$$

Then given distortions $(d_1, \ \ldots, \ d_L, \ d_0)$, $f(0)$ and $f(\sigma_x^2)$ falls into the following three cases.

Case 1: $f(0) > 0$ and $f(\sigma_x^2) < 0$.

In this case, since $f(a)$ is a continuous function, from intermediate value theorem [29, Page 48] we know that there exists an $a^* \in (0, \sigma_x^2)$ such that $f(a^*) = 0$. In this case the condition for Theorem 4 holds and from Theorem 4 we know that Gaussian description scheme with covariance matrix for $w_1, \ \ldots, \ w_l$ being (86) with $a = a^*$ achieves the optimal sum rate.

Case 2: $f(0) \leq 0$. Alternatively, $\frac{1}{d_0} + \frac{L-1}{\sigma_x^2} - \sum_{l=1}^{L} \frac{1}{d_l} \leq 0$.

In this case, the condition for Theorem 4 does not hold. But the Gaussian description scheme can still achieve the sum rate. To see this, choosing $a = 0$ in $\mathbf{K}_w$ we can meet individual distortions with equality and get a central distortion d_0'. From (33) we have

$$\begin{aligned}\frac{1}{d_0'} &= \frac{1}{\sigma_x^2} + (1\,1\ \ldots\ 1)K_w^{-1}(1\,1\ \ldots\ 1)^t \\ &= \frac{1}{\sigma_x^2} + \sum_{l=1}^{L} \frac{1}{\sigma_l^2} = \sum_{l=1}^{L} \frac{1}{d_l} - \frac{L-1}{\sigma_x^2} \\ &\geq \frac{1}{d_0}.\end{aligned} \tag{93}$$

Hence we have achieved distortion $(d_1, \ldots, d_L, d_0')$ where $d_0' \leq d_0$, and from (15) the achievable sum rate is

$$\sum_{l=1}^{L} R_l \geq \frac{1}{2} \log \frac{\sigma_x^{2L}}{d_1 d_2 \cdots d_L}, \tag{94}$$

which equals the sum of our bounds on individual rates.

Case 3: $f(\sigma_x^2) \geq 0$, Alternatively, $\sum_{l=1}^{L} d_l - d_0 - (L-1)\sigma_x^2 \geq 0$.

In this case, the conditions for Theorem 4 do not hold as well. But the Gaussian description strategy still achieves the sum rate. To see this, note that we can find a d_L' such that $0 < d_L' \leq d_L$ and

$$\sum_{l=1}^{L-1} d_l + d_L' - d_0 - (L-1)\sigma_x^2 = 0, \tag{95}$$

and we choose $a = \sigma_x^2$, $\sigma_l^2 = (d_l^{-1} - \sigma_x^{-2})^{-1}$ for $l = 1, \cdots, L-1$, and $\sigma_L^2 = (d'^{-1}_L - \sigma_x^{-2})^{-1}$ in K_w. Defining $\sigma_0^2 = (d_0^{-1} - \sigma_x^{-2})^{-1}$, (95) is equivalent to the following equation:

$$\left[\sigma_0^2 + \sigma_x^2\right]^{-1} = \sum_{l=1}^{L} \left[\sigma_l^2 + \sigma_x^2\right]^{-1}. \tag{96}$$

From Lemma 4, our choice of K_w is positive definite. Thus the resulting distortions are $(d_1, \ \ldots, \ d_{L-1}, \ d'_L, d_0)$, where $0 < d'_L \le d_L$.

Using the determinant equation

$$\begin{vmatrix} \sigma_1^2 & -\sigma_x^2 & -\sigma_x^2 & -\sigma_x^2 & \ldots & -\sigma_x^2 \\ -\sigma_x^2 & \sigma_2^2 & -\sigma_x^2 & -\sigma_x^2 & \ldots & -\sigma_x^2 \\ -\sigma_x^2 & -\sigma_x^2 & \sigma_3^2 & -\sigma_x^2 & \ldots & -\sigma_x^2 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ -\sigma_x^2 & \ldots & -\sigma_x^2 & -\sigma_x^2 & \sigma_{L-1}^2 & -\sigma_x^2 \\ -\sigma_x^2 & \ldots & -\sigma_x^2 & -\sigma_x^2 & -\sigma_x^2 & \sigma_L^2 \end{vmatrix} = \left(1 - \sum_{l=1}^{L} \frac{\sigma_x^2}{\sigma_l^2 + \sigma_x^2}\right) \prod_{l=1}^{L} (\sigma_l^2 + \sigma_x^2) \tag{97}$$

and (96), we have an achievable sum rate

$$\sum_{l=1}^{L} R_l = \frac{1}{2} \log \frac{\sigma_x^2}{d_0}. \tag{98}$$

We conclude that in this case the point-to-point rate-distortion bound for the central receiver is achievable.

In summary, we have shown that the Gaussian description scheme achieves the lower bound on the sum rate[5]. Further, the sum rate can be calculated either trivially (by choosing $a^* = 0$ in case 2 or $a^* = 1$ in case 3) or by solving a polynomial equation in a single variable (case 1).

7 Two-Description Problem

In this section, we first show that when there are only two descriptions, we can explicitly solve (37) for some cases of distortion constraints. Then we show that we can characterize the entire optimal rate region for two description problem.

[5]If all the individual distortion constraints are equal and all descriptions have the same rate, our result reduce to the the symmetrical rate point result given in [7, Section V].

7.1 Explicit Solutions for Sum Rate

With only two descriptions, we can explicitly solve (37), thus generalizing the corresponding solution for the scalar Gaussian source, derived in [1].

Suppose the distortion constraints are denoted by $(\mathbf{D}_1,\ \mathbf{D}_2,\ \mathbf{D}_0)$ and let

$$\mathbf{K}_w = \begin{pmatrix} \mathbf{K}_{w_1} & -\mathbf{A}^* \\ -\mathbf{A}^* & \mathbf{K}_{w_2} \end{pmatrix}.$$

We now solve (33), which is equivalent to (37), for $\mathbf{K}_{w_1}$, $\mathbf{K}_{w_2}$ and $\mathbf{A}^*$. From (33) we get

$$\mathbf{K}_{w_l} = (\mathbf{D}_l^{-1} - \mathbf{K}_x^{-1})^{-1}, \quad l = 1,\ 2, \tag{99}$$

and

$$\mathbf{D}_0^{-1} = \mathbf{K}_x^{-1} + (\mathbf{I}_N\ \mathbf{I}_N)\mathbf{K}_w^{-1}(\mathbf{I}_N\ \mathbf{I}_N)^t. \tag{100}$$

Expanding out $\mathbf{K}_w^{-1}$ using Lemma 7 in Appendix A, we get

$$\mathbf{D}_0^{-1} - \mathbf{K}_x^{-1} = \mathbf{K}_{w_1}^{-1} + (\mathbf{I}_N + \mathbf{K}_{w_1}^{-1}\mathbf{A}^*)(\mathbf{K}_{w_2} - \mathbf{A}^*\mathbf{K}_{w_1}^{-1}\mathbf{A}^*)^{-1}(\mathbf{I}_N + \mathbf{A}^*\mathbf{K}_{w_1}^{-1}). \tag{101}$$

Taking inverse on both sides, we have

$$(\mathbf{D}_0^{-1} - \mathbf{K}_x^{-1})^{-1} = \mathbf{K}_{w_1} - (\mathbf{K}_{w_1} + \mathbf{A}^*)(\mathbf{K}_{w_1} + \mathbf{K}_{w_2} + 2\mathbf{A}^*)^{-1}(\mathbf{K}_{w_1} + \mathbf{A}^*). \tag{102}$$

Defining $\mathbf{K}_{w_0}$ as

$$\mathbf{K}_{w_0} \overset{\text{def}}{=} [\mathbf{D}_0^{-1} - \mathbf{K}_x^{-1}]^{-1}, \tag{103}$$

we find (102) is equivalent to

$$\mathbf{K}_{w_1} - \mathbf{K}_{w_0} = (\mathbf{K}_{w_1} + \mathbf{A}^*)(\mathbf{K}_{w_1} + \mathbf{K}_{w_2} + 2\mathbf{A}^*)^{-1}(\mathbf{K}_{w_1} + \mathbf{A}^*). \tag{104}$$

Defining

$$\mathbf{X} \overset{\text{def}}{=} \mathbf{K}_{w_1} + \mathbf{A}^*,$$

we have that (104) is equivalent to

$$\mathbf{K}_{w_1} - \mathbf{K}_{w_0} = \mathbf{X}(2\mathbf{X} + \mathbf{K}_{w_2} - \mathbf{K}_{w_1})^{-1}\mathbf{X}, \tag{105}$$

which is further equivalent to

$$\mathbf{X}(\mathbf{K}_{w_1} - \mathbf{K}_{w_0})^{-1}\mathbf{X} = 2\mathbf{X} + \mathbf{K}_{w_2} - \mathbf{K}_{w_1}. \tag{106}$$

This is a version of the so-called *algebraic Riccati equation*; the corresponding Hamiltonian is readily seen to be positive semidefinite and we can even write down the following explicit solution:

$$\begin{aligned} \mathbf{X} =& \mathbf{K}_{w_1} - \mathbf{K}_{w_0} \\ &+ (\mathbf{K}_{w_1} - \mathbf{K}_{w_0})^{\frac{1}{2}} \left[(\mathbf{K}_{w_1} - \mathbf{K}_{w_0})^{-\frac{1}{2}}(\mathbf{K}_{w_2} - \mathbf{K}_{w_0})(\mathbf{K}_{w_1} - \mathbf{K}_{w_0})^{-\frac{1}{2}}\right]^{\frac{1}{2}} (\mathbf{K}_{w_1} - \mathbf{K}_{w_0})^{\frac{1}{2}}. \end{aligned} \tag{107}$$

Thus

$$\mathbf{A}^* = (\mathbf{K}_{w_1}-\mathbf{K}_{w_0})^{\frac{1}{2}} \left[(\mathbf{K}_{w_1}-\mathbf{K}_{w_0})^{-\frac{1}{2}}(\mathbf{K}_{w_2}-\mathbf{K}_{w_0})(\mathbf{K}_{w_1}-\mathbf{K}_{w_0})^{-\frac{1}{2}}\right]^{\frac{1}{2}} (\mathbf{K}_{w_1}-\mathbf{K}_{w_0})^{\frac{1}{2}}-\mathbf{K}_{w_0}. \tag{108}$$

Now, if $\mathbf{0} \prec \mathbf{A}^* \prec \mathbf{K}_x$ then we can appeal to Theorem 4 and arrive at the explicit Gaussian description scheme parameterized by $\mathbf{K}_w$ that achieves the sum rate. Analogous to the scalar case (cf. [1]), we have the following sufficient condition for when this is true.

***Proposition* 4.** *If the distortion constraints* $(\mathbf{D}_1,\ \mathbf{D}_2,\ \mathbf{D}_0)$ *satisfy*

$$\begin{aligned} &\mathbf{D}_0 + \mathbf{K}_x - \mathbf{D}_1 - \mathbf{D}_2 \succ \mathbf{0} \\ \text{and}\quad &\mathbf{D}_0^{-1} + \mathbf{K}_x^{-1} - \mathbf{D}_1^{-1} - \mathbf{D}_2^{-1} \succ \mathbf{0}, \end{aligned} \tag{109}$$

then $\mathbf{0} \prec \mathbf{A}^* \prec \mathbf{K}_x$.

Proof. See Appendix K. □

We now consider the cases that are not covered by the conditions in Proposition 4.

- When
$$\mathbf{D}_0^{-1} + \mathbf{K}_x^{-1} - \mathbf{D}_1^{-1} - \mathbf{D}_2^{-1} \preccurlyeq \mathbf{0}, \tag{110}$$
we can choose $\mathbf{A}^* = \mathbf{0}$ to achieve the sum of point-to-point individual rate-distortion functions. Thus in this case, the sum rate is equal to this natural lower bound.

- When
$$\mathbf{D}_0 + \mathbf{K}_x - \mathbf{D}_1 - \mathbf{D}_2 \preccurlyeq \mathbf{0}, \tag{111}$$
we can choose $\mathbf{A}^* = \mathbf{K}_x$ to achieve the point-to-point rate distortion-function for central receiver, also a natural lower bound.

- When neither $\mathbf{D}_0 + \mathbf{K}_x - \mathbf{D}_1 - \mathbf{D}_2$ nor $\mathbf{D}_0^{-1} + \mathbf{K}_x^{-1} - \mathbf{D}_1^{-1} - \mathbf{D}_2^{-1}$ is positive or negative semidefinite (this case cannot happen in the scalar case), we cannot use Theorem 4, and the trivial choice of $\mathbf{A}^* = \mathbf{0}$ or $\mathbf{A}^* = \mathbf{K}_x$ does not meet the lower bound. By Theorem 2, the Gaussian description scheme also achieves the lower bound on the sum rate.

If we let the source to be scalar Gaussian, our result reduces to Ozarow's solution to sum rate of the two-description problem for a scalar Gaussian source [1]: this is because the last case described above does not happen in the scalar case.

7.2 Rate Region for Two Descriptions

Applying Theorem 2 to the case of $L = 2$, i.e., the two description problem, we can see that Gaussian description scheme achieves the optimal sum rate. It also turns out that in the two-description problem, we can show that the Gaussian description strategy achieves the entire rate region.

Proof of Theorem 3. From Section 4 we have a outer bound to the rate region for the two description problem

$$\mathcal{R}_{out}(\mathbf{K}_x,\ \mathbf{D}_1,\ \mathbf{D}_2,\ \mathbf{D}_0) = \left\{ \begin{array}{l} (R_1,\ R_2): \\ R_l \geq \frac{1}{2}\log\frac{|\mathbf{K}_x|}{|\mathbf{D}_l|}, \quad l = 1,\ 2 \\ R_1 + R_2 \geq \sup_{\mathbf{K}_z \succ \mathbf{0}} \frac{1}{2}\log\frac{|\mathbf{K}_x||\mathbf{K}_x + \mathbf{K}_z||\mathbf{D}_0 + \mathbf{K}_z|}{|\mathbf{D}_0||\mathbf{D}_1 + \mathbf{K}_z||\mathbf{D}_2 + \mathbf{K}_z|} \end{array} \right\}. \quad (112)$$

Following the discussion in Section 5, we show in the following that the Gaussian description strategy (Gaussian multiple description schemes and the time sharing between them) achieves the outer bound to the rate region.

Let

$$\mathbf{K}_{w_l} = (\mathbf{D}_l^{-1} - \mathbf{K}_x^{-1})^{-1}, \quad l = 0,\ 1,\ 2 \quad (113)$$

and

$$F(\mathbf{A}) = \log|\mathbf{K}_{w_0} + \mathbf{A}| - \log|\mathbf{K}_{w_1} + \mathbf{A}| - \log|\mathbf{K}_{w_2} + \mathbf{A}|. \quad (114)$$

Now consider the optimization problem:

$$\max_{\mathbf{0} \preccurlyeq \mathbf{A} \preccurlyeq \mathbf{K}_x} \quad F(\mathbf{A}). \quad (115)$$

As in Section 5, the optimal solution $\mathbf{A}^*$ falls into four cases.

Case 1: $\mathbf{0} \prec \mathbf{A}^* \prec \mathbf{K}_x$. In this case, we know from Section 3 that the rate pair (R_1, R_2) in the following set

$$\left\{ \begin{array}{l} (R_1,\ R_2): \\ R_l \geq \frac{1}{2}\log\frac{|\mathbf{K}_x + \mathbf{K}_{w_l}|}{|\mathbf{K}_{w_l}|}, \quad l = 1,\ 2 \\ R_1 + R_2 \geq \frac{1}{2}\log\frac{|\mathbf{K}_x + \mathbf{K}_{w_1}||\mathbf{K}_x + \mathbf{K}_{w_2}|}{|\mathbf{K}_w|} \end{array} \right\} \quad (116)$$

is achievable using the Gaussian multiple description scheme with the covariance matrix of $\mathbf{w}_1,\ \mathbf{w}_2$ being

$$\mathbf{K}_w = \begin{pmatrix} \mathbf{K}_{w_1} & -\mathbf{A}^* \\ -\mathbf{A}^* & \mathbf{K}_{w_2} \end{pmatrix}.$$

Denoting the resulting distortions as $(\mathbf{D}_1,\ \mathbf{D}_2,\ \mathbf{D}_0)$, we readily calculate

$$\frac{1}{2}\log\frac{|\mathbf{K}_x+\mathbf{K}_{w_l}|}{|\mathbf{K}_{w_l}|}=\frac{1}{2}\log|\mathbf{K}_x||\mathbf{K}_{w_l}^{-1}+\mathbf{K}_x^{-1}|=\frac{1}{2}\log\frac{|\mathbf{K}_x|}{|\mathbf{D}_l|}$$

for $l=1,\ 2$. From Theorem 4 we know that the lower bound to sum rate is achieved using this Gaussian description scheme. Thus, in this case, the Gaussian description scheme achieves the rate region. As an aside, we note in this case that, $\mathbf{A}^*$ satisfies

$$\left[\mathbf{K}_{w_0}^*+\mathbf{A}^*\right]^{-1}=\left[\mathbf{K}_{w_1}+\mathbf{A}^*\right]^{-1}+\left[\mathbf{K}_{w_2}+\mathbf{A}^*\right],$$

and, from the discussion in Section 7.1, that a sufficient condition for this case to happen is (109).

Case 2: $\mathbf{0}\preccurlyeq\mathbf{A}^*\prec\mathbf{K}_x$. This case is similar to case 1: the Gaussian description scheme with covariance matrix for $\mathbf{w}_1,\ \mathbf{w}_2$ being

$$\mathbf{K}_w=\begin{pmatrix}\mathbf{K}_{w_1} & -\mathbf{A}^*\\ -\mathbf{A}^* & \mathbf{K}_{w_2}\end{pmatrix}.$$

achieves the lower bound on the rate region. We note that in this case the resulting distortions are $(\mathbf{D}_1,\ \mathbf{D}_2,\ \mathbf{D}_0^*)$, with $\mathbf{D}_0^*\preccurlyeq\mathbf{D}_0$. Further, we know from the discussion in 7.1, that a sufficient condition for this case to happen is

$$\mathbf{D}_0^{-1}+\mathbf{K}_x^{-1}-\mathbf{D}_1^{-1}-\mathbf{D}_2^{-1}\preccurlyeq\mathbf{0}.$$

Case 3: $\mathbf{0}\prec\mathbf{A}^*\preccurlyeq\mathbf{K}_x$. In this case, we know from the discussion of the corresponding case 3 in Section 5 that for another two-description problem with distortions $(\mathbf{D}_1,\ \mathbf{D}_2^*,\ \mathbf{D}_0)$ such that $\mathbf{D}_2^*\preccurlyeq\mathbf{D}_2$, the Gaussian description scheme with covariance matrix for $\mathbf{w}_1,\ \mathbf{w}_2$ being

$$\mathbf{K}_w=\begin{pmatrix}\mathbf{K}_{w_1} & -\mathbf{A}^*\\ -\mathbf{A}^* & \mathbf{K}_{w_2}^*\end{pmatrix}$$

achieves the lower bound to sum rate $\left(\frac{1}{2}\log\frac{|\mathbf{K}_x|}{|\mathbf{D}_0|}\right)$ to the original distortions $(\mathbf{D}_1,\ \mathbf{D}_2,\ \mathbf{D}_0)$. We can see, from the contra-polymatroid structure of the achievable region of the Gaussian description scheme, that the corner point

$$B_1=\left(\frac{1}{2}\log\frac{|\mathbf{K}_x|}{|\mathbf{D}_1|},\frac{1}{2}\log\frac{|\mathbf{K}_x|}{|\mathbf{D}_0|}-\frac{1}{2}\log\frac{|\mathbf{K}_x|}{|\mathbf{D}_1|}\right)$$

in Figure 3 is achievable by this Gaussian description scheme.

Now observe that the discussion in case 3 of Section 5 is symmetric with respect to the individual receivers. Thus, by exchanging the role of receiver 1 and receiver 2, we can achieve the other corner point

$$B_2=\left(\frac{1}{2}\log\frac{|\mathbf{K}_x|}{|\mathbf{D}_0|}-\frac{1}{2}\log\frac{|\mathbf{K}_x|}{|\mathbf{D}_2|},\frac{1}{2}\log\frac{|\mathbf{K}_x|}{|\mathbf{D}_2|}\right)$$

in Figure 3 by another appropriate Gaussian description scheme. Finally, time sharing between these two Gaussian multiple description schemes allows us to achieve the lower bound on the rate region. As an aside, we note, as a consequence of the discussion in Section 7.1, that a sufficient condition for this case to happen is

$$\mathbf{D}_0 + \mathbf{K}_x - \mathbf{D}_1 - \mathbf{D}_2 \preccurlyeq \mathbf{0}.$$

Case 4: $\mathbf{0} \preccurlyeq \mathbf{A}^* \preccurlyeq \mathbf{K}_x$. In this case, we know, from the discussion of the corresponding case 4 in Section 5, that for another two description problem with distortions $(\mathbf{D}_1,\ \mathbf{D}_2^*,\ \mathbf{D}_0^*)$ such that $\mathbf{D}_2^* \preccurlyeq \mathbf{D}_2$ and $\mathbf{D}_0^* \preccurlyeq \mathbf{D}_0$, the Gaussian description scheme with covariance matrix for $\mathbf{w}_1,\ \mathbf{w}_2$ being

$$\mathbf{K}_w = \begin{pmatrix} \mathbf{K}_{w_1} & -\mathbf{A}^* \\ -\mathbf{A}^* & \mathbf{K}_{w_2}^* \end{pmatrix}$$

achieves the lower bound to sum rate $\left(\frac{1}{2}\log\frac{|\mathbf{K}_x|}{|\mathbf{D}_0|}\right)$ to the original distortions $(\mathbf{D}_1,\ \mathbf{D}_2,\ \mathbf{D}_0)$. Using an argument entirely analogous to that applied that the Gaussian description strategy achieves the rate region. □

To summarize: the Gaussian description strategy achieves the rate region for the two description problem. For a class of distortion constraints, the corner points of the rate region can be characterized by solving a matrix polynomial equation, as already seen in Section 7.1.

8 Discussions

Although multiple description for individual and central receivers is a special case of the most general multiple description problem, the solution to this problem sheds substantial insight to the issue-at-large. In this section, we discuss two instances of other multiple description problems that can be resolved using the insights developed so far. In particular, we discuss the problem of two descriptions with separate distortion constraints and the general multiple description problem for some special sets of distortion constraints.

8.1 Two Description with Separate Distortion Constraints

The problem of two descriptions with separate distortion constraints is illustrated in Figure 2. Suppose the vector Gaussian source $\mathbf{x}[m] = (\mathbf{x}_1[m], \mathbf{x}_2[m])$, the dimension of $\mathbf{x}_1[m]$ is N_1 and the dimension of $\mathbf{x}_2[m]$ is N_2. This implies that the dimension of $\mathbf{x}[m]$ is $N = N_1 + N_2$. Let $\mathbf{K}_x = \mathbb{E}[\mathbf{x}[m]^t\mathbf{x}[m]]$, $\mathbf{K}_{x_1} = \mathbb{E}[\mathbf{x}_1[m]^t\mathbf{x}_1[m]]$, and $\mathbf{K}_{x_2} = \mathbb{E}[\mathbf{x}_2[m]^t\mathbf{x}_2[m]]$.

There are two encoders at the source providing two descriptions of $\mathbf{x}[m]$. There are three receivers: the individual receivers 1 and 2 are only interested in generating reproduction of $\mathbf{x}_1[m]$ with mean square distortion constraint $\mathbf{D}_1$ (an $N_1 \times N_1$ positive definite matrix) from description 1 and $\mathbf{x}_2[m]$ with mean square distortion constraint $\mathbf{D}_2$ (an $N_2 \times N_2$ positive definite matrix) from description 2, respectively. The central receiver uses both descriptions to generate a reproduction of $\mathbf{x}[m]$ with the error covariance meeting a distortion constraint $\mathbf{D}_0$ (an $N \times N$ positive definite matrix) from both descriptions.

This situation is closely related to the two description problem and we can harness our results thus far to completely characterize the rate region of the problem at hand.

***Theorem* 5.** *The rate region of two description with separate distortion constraints is*

$$\mathcal{R}(\mathbf{D}_1,\ \mathbf{D}_2,\ \mathbf{D}_0) = \bigcup_{\Upsilon(\mathbf{D}_1',\ \mathbf{D}_2')} \mathcal{R}_*(\mathbf{D}_1',\ \mathbf{D}_2',\ \mathbf{D}_0), \tag{117}$$

where $\Upsilon(\mathbf{D}_1',\ \mathbf{D}_2')$ *is defined as*

$$\Upsilon(\mathbf{D}_1',\ \mathbf{D}_2') \overset{def}{=} \left\{\ (\mathbf{D}_1',\ \mathbf{D}_2') :\ (\mathbf{D}_1')_{\{1,\dots,N_1\}} \preccurlyeq \mathbf{D}_1,\ (\mathbf{D}_2')_{\{N_1+1,\dots,N\}} \preccurlyeq \mathbf{D}_2\ \right\}. \tag{118}$$

Proof. It is clear that any rate pair $(R_1, R_2) \in \mathcal{R}_*(\mathbf{D}_1',\ \mathbf{D}_2',\ \mathbf{D}_0)$ for some $(\mathbf{D}_1',\ \mathbf{D}_2') \in \Upsilon(\mathbf{D}_1',\ \mathbf{D}_2')$ is in the rate region for the two description with separate distortion constraints, and so

$$\mathcal{R}_*(\mathbf{D}_1',\ \mathbf{D}_2',\ \mathbf{D}_0) \subseteq \mathcal{R}(\mathbf{D}_1,\ \mathbf{D}_2,\ \mathbf{D}_0).$$

On the other hand, although receiver 1 (2) is only interested in reconstructing $\mathbf{x}_1$ ($\mathbf{x}_2$), they can actually reconstruct the entire source $\mathbf{x}$ based on their received descriptions. Hence, any coding scheme for the two description with separate distortion constraints will result in some achievable distortions $(\mathbf{D}_1',\ \mathbf{D}_2',\ \mathbf{D}_0')$ with $(\mathbf{D}_1',\ \mathbf{D}_2') \in \Upsilon(\mathbf{D}_1',\ \mathbf{D}_2')$ and $\mathbf{D}_0' \preccurlyeq \mathbf{D}_0$. Thus any rate pair $(R_1, R_2) \in \mathcal{R}(\mathbf{D}_1,\ \mathbf{D}_2,\ \mathbf{D}_0)$ achieved by this coding scheme is in the rate region $\mathcal{R}_*(\mathbf{D}_1',\ \mathbf{D}_2',\ \mathbf{D}_0)$ for the two description problem. Thus

$$\mathcal{R}(\mathbf{D}_1,\ \mathbf{D}_2,\ \mathbf{D}_0) \subseteq \bigcup_{\Upsilon(\mathbf{D}_1',\ \mathbf{D}_2')} \mathcal{R}_*(\mathbf{D}_1',\ \mathbf{D}_2',\ \mathbf{D}_0).$$

From equivalence of the two regions in (117), the proof is now complete. □

8.2 General Gaussian Multiple Description Problem for Special Choices of Distortion Constraints

Consider the general Gaussian multiple description problem with source covariance $\mathbf{K}_x$ and $2^L - 1$ distortion constraints D_S for each $S \subseteq \{1,\ \dots,\ L\}$.

Following arguments similar to that used in arriving at the lower bound (24) for sum rate, we have an outer bound on the rate region:

$$\mathcal{R}_{out}(\mathbf{K}_x,\ \mathbf{D}_1,\ \ldots,\ \mathbf{D}_L,\ \mathbf{D}_0) = \left\{ \begin{array}{l} (R_1,\ \ldots,\ R_L): \\ \sum\limits_{l\in S} R_l \geq \frac{1}{2}\log \frac{|\mathbf{K}_x||\mathbf{K}_x+\mathbf{K}_z|^{(|S|-1)}|\mathbf{D}_S+\mathbf{K}_z|}{|\mathbf{D}_S|\prod\limits_{l\in S}|\mathbf{D}_l+\mathbf{K}_z|}, \quad \forall S\subseteq\{1,\ \ldots,\ L\} \end{array} \right\}. \tag{119}$$

Following arguments similar to those used in arriving at the upper bound (28) for the sum rate, we can use a Gaussian description scheme with covariance matrix of $\mathbf{w}_l$'s ($\mathbf{K}_w$) taking the form (31), any tuple $(R_1,\ \ldots,\ R_L)$ satisfying

$$\left\{ \begin{array}{l} (R_1,\ \ldots,\ R_L): \\ \sum\limits_{l\in S} R_l \geq \frac{1}{2}\log \frac{\left|\mathbf{K}_x\right|\left|\mathbf{K}_x+\mathbf{K}_z\right|^{(|S|-1)}\left|\mathrm{Cov}[\mathbf{x}|\mathbf{u}_l,\ l\in S]+\mathbf{K}_z\right|}{\left|\mathrm{Cov}[\mathbf{x}|\mathbf{u}_l,\ l\in S]\right|\prod\limits_{l\in S}\left|\mathrm{Cov}[\mathbf{x}|\mathbf{u}_l]+\mathbf{K}_z\right|}, \quad \forall S\subseteq\{1,\ \ldots,\ L\} \end{array} \right\} \tag{120}$$

is achievable. Thus if we can find a $\mathbf{K}_w$ of the form in (31) such that all of the 2^L-1 distortion constraints are met with equality, i.e.,

$$\mathbf{D}_S = \mathrm{Cov}[\mathbf{x}|\mathbf{u}_l,\ l\in S] = [\mathbf{K}_x^{-1}+(\mathbf{I}_N,\ \ldots,\ \mathbf{I}_N)\mathbf{K}_{w_S}^{-1}(\mathbf{I}_N,\ \ldots,\ \mathbf{I}_N)^t]^{-1}, \quad \forall S\subseteq\{1,\ \ldots,\ L\}, \tag{121}$$

where $\mathbf{K}_{w_S}$ is the covariance matrix for all $\mathbf{K}_{w_l}, l\in S$, then the achievable region matches the outer bound and we would have characterized the rate region of the multiple description problem.

From the above discussion, we see that for some choice of distortion constraints of the multiple description problem, we can indeed do this: First choose $L+1$ distortions $(\mathbf{D}_1,\ \mathbf{D}_2,\ \ldots,\ \mathbf{D}_L,\ \mathbf{D}_0)$ such that they satisfy the condition for Theorem 4 for the multiple description problem with individual and central receivers. Next we can solve for the $\mathbf{K}_w$ which is the covariance matrix of $(\mathbf{w}_1,\ \ldots,\ \mathbf{w}_L)$ for the sum-rate-achieving Gaussian description scheme. For any other $S\subseteq\{1,\ \ldots,\ L\}$, this scheme results in distortion $\mathbf{D}_S = [\mathbf{K}_x^{-1}+(\mathbf{I}_N,\ \ldots,\ \mathbf{I}_N)\mathbf{K}_{w_S}^{-1}(\mathbf{I}_N,\ \ldots,\ \mathbf{I}_N)^t]^{-1}$. Finally we choose these $\mathbf{D}_S$'s as the other distortion constraints. Now we have a general multiple description problem with 2^L-1 distortion constraints D_S for each $S\subseteq\{1,\ \ldots,\ L\}$, and hence we can find a $\mathbf{K}_w$ of form (31) such that all of the 2^L-1 distortion constraints are met with equality. Thus (119) is actually the rate region and it can be achieved by a Gaussian description scheme.

8.3 Robustness of Gaussian Multiple Descriptions

In this section, we demonstrate the robustness of the ensemble of the Gaussian multiple description schemes by showing that the distortion achieved through describing any memoryless vector source using this scheme is no worse (in the sense of positive semidefinite ordering) than when the source itself were Gaussian. In particular, this implies that the Gaussian memoryless vector source is the *hardest* to multiply describe among all memoryless vector sources with the same covariance matrix. This latter result could be viewed as a generalization of the result in [5] which focused on two descriptions of a memoryless scalar source.

***Proposition* 5.** *Consider a memoryless vector source $\mathbf{x}^n$ with marginal covariance matrix $\mathbf{K}_x$. Let $\mathbf{w}_1, \cdots, \mathbf{w}_L$ be N dimensional zero mean jointly Gaussian random vectors independent of $\mathbf{x}$, with the positive definite covariance matrix of $(\mathbf{w}_1, \cdots, \mathbf{w}_L)$ denoted by $\mathbf{K}_w$. Then, treating the source statistics as Gaussian and using the Gaussian multiple description encoder described in Section 2.2 parameterized by $\mathbf{K}_w$, and reconstructing the sources at the decoders via MMSE estimation the achieved distortion $(\mathbf{D}_1, \cdots, \mathbf{D}_L, \mathbf{D}_0)$ satisfies*

$$\mathbf{D}_l \preceq \mathbf{D}_l^G, \quad l = 0, 1, \cdots, L \tag{122}$$

where $(\mathbf{D}_1^G, \cdots, \mathbf{D}_L^G, \mathbf{D}_0^G)$ is the distortion achieved by the same architecture when $\mathbf{x}^n$ is a memoryless vector Gaussian source.

Proof. Using the natural achievable scheme described in Section 2.2, let

$$\mathbf{u}_l = \mathbf{x} + \mathbf{w}_l, \quad l = 1, \ \ldots, \ L.$$

For the description rate tuple $(R_1, \ \ldots, \ R_L)$ to be sufficient to convey $\mathbf{u}_l$ to each receiver l, we need satisfying

$$\sum_{l \in S} R_l \geq \left[\sum_{l \in S} h(\mathbf{u}_l) \right] - h(\mathbf{u}_l, l \in S | \mathbf{x}). \tag{123}$$

Now, since $h(\mathbf{u}_l)$ is maximized when $\mathbf{x}$ is Gaussian simultaneously for each l, we have

$$\left[\sum_{l \in S} h(\mathbf{u}_l) \right] - h(\mathbf{u}_l, l \in S | \mathbf{x}) \leq \frac{1}{2} \log \frac{\prod_{l \in S} |\mathbf{K}_x + \mathbf{K}_{w_l}|}{|\mathbf{K}_{w_S}|}, \quad \forall S \subseteq \{1, \ \ldots, \ L\}. \tag{124}$$

Thus the description rates for a non-Gaussian source are only smaller in general than for a Gaussian source.

Now we turn to the reconstructions:

$$\begin{aligned} \hat{\mathbf{x}}_l &= \alpha_l \mathbf{u}_l, \quad l = 1, \ \ldots, \ L, \\ \hat{\mathbf{x}}_0 &= \sum_{i=1}^{L} \beta_l \mathbf{u}_l, \end{aligned} \tag{125}$$

where $\alpha_1, \cdots, \alpha_L$ and $\beta_1, \cdots, \beta_L$ are chosen so that $\alpha_l \mathbf{u}_l$ is the linear MMSE estimation of $\mathbf{x}$ given $\mathbf{u}_l$ for $l = 1, \ldots, L$, and $\sum_{i=1}^{L} \beta_l \mathbf{u}_l$ is the linear MMSE estimation of $\mathbf{x}$ given $\mathbf{u}_1, \cdots, \mathbf{u}_L$. We can see that $\alpha_1, \cdots, \alpha_L$ and $\beta_1, \cdots, \beta_L$ do not change with the source distribution for fixed Gaussian description scheme. The distortions achieved are

$$\begin{aligned} \mathbf{D}_l =& \mathbb{E}\Big[(\mathbf{x} - \hat{\mathbf{x}}_l)^t (\mathbf{x} - \hat{\mathbf{x}}_l])\Big], \quad l = 1, \ldots, L, \\ \mathbf{D}_0 =& \mathbb{E}\Big[(\mathbf{x} - \hat{\mathbf{x}}_0)^t (\mathbf{x} - \hat{\mathbf{x}}_0)\Big]. \end{aligned} \tag{126}$$

Note that if $\mathbf{x}^n$ is memoryless vector Gaussian, then linear MMSE is the MMSE estimation. On the other hand, if $\mathbf{x}^n$ is not Gaussian then doing an MMSE estimation (instead of just linear MMSE) potentially leads to smaller distortions. We conclude the proof of the claim in (122). □

Now, using the optimality properties of the Gaussian multiple description scheme derived earlier in this paper we can conclude the following worst-case property of multiply describing Gaussian sources.

***Corollary* 1.** *Consider a memoryless vector source $\mathbf{x}^n$ with marginal covariance matrix $\mathbf{K}_x$. The minimal sum rate in describing for individual and central receiver so as to satisfy distortion constraint $(\mathbf{D}_1, \cdots, \mathbf{D}_L, \mathbf{D}_0)$ is upper bounded by*

$$\min_{\mathbf{K}_w} \frac{1}{2} \log \frac{\prod_{l=1}^{L} |\mathbf{K}_x + \mathbf{K}_{w_l}|}{|\mathbf{K}_w|}, \tag{127}$$

where the minimization is over all $\mathbf{K}_w$ satisfying (126). In other words, the Gaussian source is the hardest to multiply describe in this setting.

8.4 A Lower Bound on the Sum Rate

In this section, we point out that a lower bound on sum rate of multiply describing a general memoryless vector source for individual and central receivers can be derived readily by following the proof of Lemma 2. Towards this, we observe that every step in Appendix B still holds for arbitrarily distributed memoryless vector source $\mathbf{x}^n$ except for equation (135), which becomes

$$\begin{aligned} h(\mathbf{x}^n) =& \frac{1}{2} \log (2\pi e)^{Nn} P_{\mathbf{x}}^n, \\ h(\mathbf{y}^n) =& \frac{1}{2} \log (2\pi e)^{Nn} P_{\mathbf{y}}^n \geq \frac{1}{2} \log (2\pi e)^{Nn} \left(P_{\mathbf{x}}^{1/N} + |\mathbf{K}_z|^{1/N} \right)^{Nn}, \end{aligned} \tag{128}$$

where we have denoted the normalized entropy power of an N-dimensional random vector $\mathbf{x}$ by

$$P_{\mathbf{x}} = \frac{e^{2h(\mathbf{x})}}{(2\pi e)^N}. \tag{129}$$

Following the rest of the steps in Appendix B, we have the following lower bound on the sum rate of multiply describing an arbitrarily distributed memoryless vector source.

***Proposition* 6.** *Consider a memoryless vector source $\mathbf{x}^n$ with marginal covariance matrix $\mathbf{K}_x$. The minimal sum rate in multiply describing this source for individual and central receiver, so as to satisfy a distortion constraint $(\mathbf{D}_1, \cdots, \mathbf{D}_L, \mathbf{D}_0)$, is lower bounded by*

$$\sup_{\mathbf{K}_z \succeq \mathbf{0}} \frac{1}{2} \log \frac{P_{\mathbf{x}} \left(P_{\mathbf{x}}^{1/N} + |\mathbf{K}_z|^{1/N}\right)^{N(L-1)} |\mathbf{D}_0 + \mathbf{K}_z|}{|\mathbf{D}_0| \prod_{l=1}^{L} |\mathbf{D}_l + \mathbf{K}_z|}. \tag{130}$$

Appendix

A Useful Matrix Lemmas

In this appendix we provide some useful results in matrix analysis that are extensively used in this paper.

***Lemma* 6 (Matrix Inversion Lemma).** *[30, Theorem 2.5] Let $\mathbf{A}$ be an $m \times m$ nonsingular matrix and $\mathbf{B}$ be an $n \times n$ nonsingular matrix and let $\mathbf{C}$ and $\mathbf{D}$ be $m \times n$ and $n \times m$ matrices, respectively. If the matrix $\mathbf{A} + \mathbf{CBD}$ is nonsingular, then*

$$(\mathbf{A} + \mathbf{CBD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{C}(\mathbf{B}^{-1} + \mathbf{DA}^{-1}\mathbf{C})^{-1}\mathbf{DA}^{-1}$$

***Lemma* 7.** *[30, Theorem 2.3] Suppose that the partitioned matrix*

$$\mathbf{M} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}$$

is invertible and that the inverse is partitioned as

$$\mathbf{M}^{-1} = \begin{pmatrix} \mathbf{X} & \mathbf{Y} \\ \mathbf{U} & \mathbf{V} \end{pmatrix}.$$

If $\mathbf{A}$ is a nonsingular principal sub-matrix of $\mathbf{M}$, then

$$\begin{aligned} \mathbf{X} &= \mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{B}(\mathbf{D} - \mathbf{CA}^{-1}\mathbf{B})^{-1}\mathbf{CA}^{-1}, \\ \mathbf{Y} &= -\mathbf{A}^{-1}\mathbf{B}(\mathbf{D} - \mathbf{CA}^{-1}\mathbf{B})^{-1}, \\ \mathbf{U} &= -(\mathbf{D} - \mathbf{CA}^{-1}\mathbf{B})^{-1}\mathbf{CA}^{-1}, \\ \mathbf{V} &= (\mathbf{D} - \mathbf{CA}^{-1}\mathbf{B})^{-1}. \end{aligned} \tag{131}$$

Lemma 8. *[30, Theorem 6.13] Let* $\mathbf{E} \in \mathbb{M}_n$ *be a positive definite matrix and let* $\mathbf{F}$ *be an* $n \times m$ *matrix. Then for any* $m \times m$ *positive definite matrix* $\mathbf{G}$*,*

$$\begin{pmatrix} \mathbf{E} & \mathbf{F} \\ \mathbf{F}^t & \mathbf{G} \end{pmatrix} \succ 0 \Longleftrightarrow \mathbf{G} \succ \mathbf{F}^t \mathbf{E}^{-1} \mathbf{F}. \tag{132}$$

Lemma 9. *[30, Theorem 6.8 and 6.9] Let* $\mathbf{A}$ *and* $\mathbf{B}$ *be positive definite matrices such that* $\mathbf{A} \succ \mathbf{B}$ $(\mathbf{A} \succcurlyeq \mathbf{B})$*. Then,*

$$\begin{aligned} &|\mathbf{A}| \succ |\mathbf{B}| \quad (|\mathbf{A}| \succcurlyeq |\mathbf{B}|), \\ &\mathbf{A}^{-1} \prec \mathbf{B}^{-1} \quad (\mathbf{A}^{-1} \preccurlyeq \mathbf{B}^{-1}), \\ &\mathbf{A}^{1/2} \succ \mathbf{B}^{1/2} \quad (\mathbf{A}^{1/2} \succcurlyeq \mathbf{B}^{1/2}). \end{aligned} \tag{133}$$

B Proof of Lemma 2

We start the proof by first closely following the steps of [1,7]. However, the key difference between our approach and that in [1,7] is that at certain point (equation (137)), instead of using entropy power inequality as in [1,7], we use worst case noise result [31, Lemma II.2]. It is well known that the vector entropy power inequality is tight only when a certain covariance alignment condition is satisfied. Our approach avoids this route thus enabling a tighter lower bound.

We first define an i.i.d. random process $\{\mathbf{z}[m]\}$, $m = 1, \ldots, n$ of $\mathcal{N}(0, \mathbf{K}_z)$ Gaussian random vectors, where $\mathbf{z}[m]$, $m = 1, \ldots, n$ are independent of $\mathbf{x}^n$ and C_l, $l = 1, \ldots, L$. Form a random process $\mathbf{y}^n = (\mathbf{y}[1], \ldots, \mathbf{y}[n])^t$ by

$$\mathbf{y}[m] = \mathbf{x}[m] + \mathbf{z}[m], \quad m = 1, \ldots, n.$$

It follows that $\{\mathbf{y}[m]\}$ is an i.i.d. random process of $\mathcal{N}(0, \mathbf{K}_y)$ Gaussian random vectors, where $\mathbf{K}_y = \mathbf{K}_x + \mathbf{K}_z$. Then

$$
\begin{aligned}
&I(C_1; C_2; \ldots; C_L) + I(C_1, \ldots, C_L; \mathbf{x}^n) \\
&\quad = \sum_{l=1}^{L} H(C_l) - H(C_1, \cdots, C_L) + I(C_1, \ldots, C_L; \mathbf{x}^n) \\
&\quad \geq \sum_{l=1}^{L} H(C_l) - H(C_1, \cdots, C_L) + I(C_1, \ldots, C_L; \mathbf{x}^n) \\
&\qquad - \Big(\sum_{l=1}^{L} H(C_l|\mathbf{y}^n) - H(C_1, \ldots, C_L|\mathbf{y}^n) \Big) \\
&\quad = \sum_{l=1}^{L} (h(\mathbf{y}^n) - h(\mathbf{y}^n|C_l)) - h(\mathbf{y}^n) + h(\mathbf{y}^n|C_1, \ldots, C_L) + h(\mathbf{x}^n) - h(\mathbf{x}^n|C_1, \ldots, C_L) \\
&\quad = h(\mathbf{x}^n) + (L-1)h(\mathbf{y}^n) - \sum_{l=1}^{L} h(\mathbf{y}^n|C_l) + h(\mathbf{y}^n|C_1, \ldots, C_L) - h(\mathbf{x}^n|C_1, \ldots, C_L).
\end{aligned}
\tag{134}
$$

Since $\mathbf{x}^n$ and $\mathbf{y}^n$ are Gaussian vectors, for the first two terms in (134), we have

$$
\begin{aligned}
h(\mathbf{x}^n) &= \frac{1}{2} \log(2\pi e)^{Nn} |\mathbf{K}_x|^n, \\
h(\mathbf{y}^n) &= \frac{1}{2} \log(2\pi e)^{Nn} |\mathbf{K}_y|^n = \frac{1}{2} \log(2\pi e)^{Nn} |\mathbf{K}_x + \mathbf{K}_z|^n.
\end{aligned}
\tag{135}
$$

We also have the following bound on $h(\mathbf{y}^n|C_l)$ for $l = 1, \ldots, L$:

$$\begin{aligned}
h(\mathbf{y}^n|C_l) &\leq \sum_{m=1}^{n} h(\mathbf{y}[m]|C_l) \\
&\leq \sum_{m=1}^{n} \frac{1}{2}\log(2\pi e)^N \left|\mathrm{Cov}[\mathbf{y}[m]|C_l]\right| \\
&\leq \frac{1}{2}\log(2\pi e)^{Nn} + \frac{n}{2}\log\left|\frac{1}{n}\sum_{m=1}^{n}\mathrm{Cov}[\mathbf{y}[m]|C_l]\right| \\
&= \frac{1}{2}\log(2\pi e)^{Nn} + \frac{n}{2}\log\left|\frac{1}{n}\sum_{m=1}^{n}\mathrm{Cov}[(\mathbf{x}[m]+\mathbf{z}[m])|C_l]\right| \\
&= \frac{1}{2}\log(2\pi e)^{Nn} + \frac{n}{2}\log\left|\frac{1}{n}\sum_{m=1}^{n}\mathrm{Cov}[\mathbf{x}[m]|C_l]+\mathbf{K}_z\right| \\
&\leq \frac{1}{2}\log(2\pi e)^{Nn} + \frac{n}{2}\log|\mathbf{D}_l+\mathbf{K}_z| \\
&= \frac{1}{2}\log(2\pi e)^{Nn}|\mathbf{D}_l+\mathbf{K}_z|^n.
\end{aligned} \tag{136}$$

Next we bound the difference of the last two terms of (134). Different from [1, 7], we do not use entropy power inequality to bound the difference.

$$\begin{aligned}
&h(\mathbf{y}^n|C_1, \ldots, C_L) - h(\mathbf{x}^n|C_1, \ldots, C_L) \\
&\quad = h(\mathbf{y}^n|C_1, \ldots, C_L) - h(\mathbf{x}^n|\mathbf{z}^n, C_1, \ldots, C_L) \\
&\quad = h(\mathbf{y}^n|C_1, \ldots, C_L) - h(\mathbf{y}^n|\mathbf{z}^n, C_1, \ldots, C_L) \\
&\quad = I(\mathbf{y}^n; \mathbf{z}^n|C_1, \ldots, C_L).
\end{aligned} \tag{137}$$

Letting

$$\mathbf{K}_c[m] \stackrel{\text{def}}{=} \mathrm{Cov}[\mathbf{x}[m] - \hat{\mathbf{x}}_0[m]], \tag{138}$$

we have

$$
\begin{aligned}
I(\mathbf{y}^n;\mathbf{z}^n|C_1,\ \dots,\ C_L) &= h(\mathbf{z}^n|C_1,\ \dots,\ C_L) - h(\mathbf{z}^n|\mathbf{y}^n, C_1,\ \dots,\ C_L)\\
&= h(\mathbf{z}^n) - h(\mathbf{z}^n|\mathbf{y}^n - \hat{\mathbf{x}}_0^n, C_1,\ \dots,\ C_L)\\
&\geq h(\mathbf{z}^n) - h(\mathbf{z}^n|\mathbf{y}^n - \hat{\mathbf{x}}_0^n)\\
&= \sum_{m=1}^{n} \big(h(\mathbf{z}[m]) - h(\mathbf{z}[m]|\mathbf{z}[1],\ \dots,\ \mathbf{z}[m-1], \mathbf{y}^n - \hat{\mathbf{x}}_0^n)\big)\\
&\geq \sum_{m=1}^{n} \big(h(\mathbf{z}[m]) - h(\mathbf{z}[m]|\mathbf{y}[m] - \hat{\mathbf{x}}_0[m])\big)\\
&= \sum_{m=1}^{n} I(\mathbf{z}[m];\mathbf{x}[m] - \hat{\mathbf{x}}_0[m] + \mathbf{z}[m])\\
&\overset{(a)}{\geq} \sum_{m=1}^{n} \frac{1}{2}\log\frac{|\mathbf{K}_c[m] + \mathbf{K}_z[m]|}{|\mathbf{K}_c[m]|}\\
&\overset{(b)}{\geq} \frac{n}{2}\log\frac{|\mathbf{D}_0 + \mathbf{K}_z|}{|\mathbf{D}_0|},
\end{aligned}
\tag{139}
$$

where (a) is from (138) and [31, Lemma II.2]. The justification for (b) is from the convexity of $\log\frac{|\mathbf{A}+\mathbf{B}|}{|\mathbf{A}|}$ in $\mathbf{A}$ and (22). From (137) and (139) we have

$$
h(\mathbf{y}^n|C_1,\ \dots,\ C_L) - h(\mathbf{x}^n|C_1,\ \dots,\ C_L) \geq \frac{n}{2}\log\frac{|\mathbf{D}_0 + \mathbf{K}_z|}{|\mathbf{D}_0|}. \tag{140}
$$

Combining (134), (135) and (140), we have

$$
I(C_1;\ C_2;\ \dots;\ C_L) + I(C_1,\ \dots,\ C_L;\mathbf{x}^n) \geq \frac{n}{2}\log\frac{|\mathbf{K}_x||\mathbf{K}_x + \mathbf{K}_z|^{(L-1)}|\mathbf{D}_0 + \mathbf{K}_z|}{|\mathbf{D}_0|\prod_{l=1}^{L}|\mathbf{D}_l + \mathbf{K}_z|}. \tag{141}
$$

By taking the supremum over all positive definite $\mathbf{K}_z$, we can sharpen the lower bound in (141) and get (23).

The introduction of $\mathbf{z}^n$ and $\mathbf{y}^n$ is due to Ozarow [1]. Note that the only step with inequality in equation (134) is actually equality if we can find a $\mathbf{y}^n$ such that conditioned on $\mathbf{y}^n$, the codewords for different descriptions corresponding to the same source sequence $\mathbf{x}^x$ are independent. Existence of such a $\mathbf{y}^n$ is the key observation in comparing lower bound and achievable sum rate. There are different ways to introduce this conditional independence and which may lead to different lower bounds (one example is the "bootstrapping" technique in [7]). However, in our situation the lower bound derived here is tightest possible. Thus the specific way in which we have induced conditional independence is optimal for the sum rate of our problem. Conditional independence seems to be the crucial property in the solution to some other multiterminal source coding problems, such as the distributed source coding problem [26, 27].

C Proof of Proposition 2

Conditioned on $\mathbf{y}$, the collection of random variables $(\mathbf{u}_1, \ldots, \mathbf{u}_L)$ are Gaussian and thus we have

$$\sum_{l=1}^{L} h(\mathbf{u}_l|\mathbf{y}) - h(\mathbf{u}_1, \ldots, \mathbf{u}_L|\mathbf{y}) = \frac{1}{2} \log \frac{\prod_{l=1}^{L} |\mathrm{Cov}[\mathbf{u}_l|\mathbf{y}]|}{\left|\mathrm{Cov}[\mathbf{u}_1, \ldots, \mathbf{u}_L|\mathbf{y}]\right|}. \tag{142}$$

From MMSE of $\mathbf{u}_l$ from $\mathbf{y}$ we have

$$\mathrm{Cov}[\mathbf{u}_l|\mathbf{y}] = \mathbf{K}_x + \mathbf{K}_{w_l} - \mathbf{K}_x(\mathbf{K}_x + \mathbf{K}_z)^{-1}\mathbf{K}_x, \quad l = 1, \ldots, L \tag{143}$$

and

$$\mathrm{Cov}(\mathbf{u}_1, \ldots, \mathbf{u}_L|\mathbf{y}) = \mathbf{J} \otimes \mathbf{K}_x + \mathbf{K}_w - \mathbf{J} \otimes \left(\mathbf{K}_x(\mathbf{K}_x + \mathbf{K}_z)^{-1}\mathbf{K}_x\right), \tag{144}$$

where $\mathbf{J}$ is an $L \times L$ matrix of all ones and $\otimes$ is the Kronecker Product [30, Section 6.5].

By Fischer inequality (the block matrix version of Hadamard inequality, see [30, Theorem 6.10]) we know that $\prod_{l=1}^{L} |\mathrm{Cov}[\mathbf{u}_l|\mathbf{y}]| = \left|\mathrm{Cov}[\mathbf{u}_1, \ldots, \mathbf{u}_L|\mathbf{y}]\right|$ if and only if the off-diagonal block matrices of $\mathrm{Cov}[\mathbf{u}_1, \ldots, \mathbf{u}_L|\mathbf{y}]$ are all zero matrices. Thus we have

$$\sum_{l=1}^{L} h(\mathbf{u}_l|\mathbf{y}) - h(\mathbf{u}_1, \ldots, \mathbf{u}_L|\mathbf{y}) = 0$$

if and only if

$$\mathbf{K}_x - \mathbf{A} = \mathbf{K}_x(\mathbf{K}_x + \mathbf{K}_z)^{-1}\mathbf{K}_x, \tag{145}$$

or equivalently, if and only if

$$\mathbf{K}_z = \mathbf{K}_x(\mathbf{K}_x - \mathbf{A})^{-1}\mathbf{K}_x - \mathbf{K}_x. \tag{146}$$

To get a valid $\mathbf{K}_z \succ \mathbf{0}$, we need the additional condition $\mathbf{0} \prec \mathbf{A} \prec \mathbf{K}_x$.

D Proof of Lemma 3

First we assume $\mathbf{A} \succ 0$, and hence by Lemma 6 in Appendix A we have

$$\begin{aligned}
&\left[\mathbf{A}^{-1} + (\mathbf{I}_N\ \mathbf{I}_N\ \dots\ \mathbf{I}_N)\,\mathbf{K}_w^{-1}\,(\mathbf{I}_N\ \mathbf{I}_N\ \dots\ \mathbf{I}_N)^t\right]^{-1} \\
=&\mathbf{A} - \mathbf{A}\,(\mathbf{I}_N\ \mathbf{I}_N\ \dots \mathbf{I}_N)\left[\mathbf{K}_w + (\mathbf{I}_N\ \mathbf{I}_N\ \dots\ \mathbf{I}_N)^t\,\mathbf{A}\,(\mathbf{I}_N\ \mathbf{I}_N\ \dots\ \mathbf{I}_N)\right]^{-1}(\mathbf{I}_N\ \mathbf{I}_N\ \dots\ \mathbf{I}_N)^t\,\mathbf{A} \\
=&\mathbf{A} - \mathbf{A}\,(\mathbf{I}_N\ \mathbf{I}_N\ \dots \mathbf{I}_N)\left[\mathrm{diag}\{\mathbf{K}_{w_1} + \mathbf{A},\ \mathbf{K}_{w_2} + \mathbf{A},\ \dots\ \mathbf{K}_{w_L} + \mathbf{A}\}\right]^{-1}(\mathbf{I}_N\ \mathbf{I}_N\ \dots \mathbf{I}_N)^t\,\mathbf{A} \\
=&\mathbf{A} - \mathbf{A}\sum_{l=1}^{L}[\mathbf{K}_{w_l} + \mathbf{A}]^{-1}\mathbf{A}.
\end{aligned} \tag{147}$$

Thus,

$$\begin{aligned}
&(\mathbf{I}_N\ \mathbf{I}_N\ \dots\ \mathbf{I}_N)\mathbf{K}_w^{-1}\,(\mathbf{I}_N\ \mathbf{I}_N\ \dots\ \mathbf{I}_N)^t \\
&\quad= \left[\mathbf{A} - \mathbf{A}\sum_{l=1}^{L}(\mathbf{K}_{w_l} + \mathbf{A})^{-1}\mathbf{A}\right]^{-1} - \mathbf{A}^{-1} \\
&\quad= \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{A}\left[-\left(\sum_{l=1}^{L}(\mathbf{K}_{w_l} + \mathbf{A})^{-1}\right)^{-1} + \mathbf{A}\mathbf{A}^{-1}\mathbf{A}\right]^{-1}\mathbf{A}\mathbf{A}^{-1} - \mathbf{A}^{-1} \\
&\quad= \left[\left(\sum_{l=1}^{L}(\mathbf{K}_{w_l} + \mathbf{A})^{-1}\right)^{-1} - \mathbf{A}\right]^{-1}.
\end{aligned} \tag{148}$$

When $\mathbf{A}$ is singular, we can choose $\delta > 0$ such that $\mathbf{A} + \epsilon\mathbf{I}_N \succ 0$ for $\epsilon \in (0, \delta)$, and thus we have the previous argument. By letting $\epsilon \to 0^+$, we have the result.

E Proof of Lemma 4

We use induction. First consider the matrix

$$\mathbf{\Delta}_2 = \begin{pmatrix} \mathbf{K}_{w_1} & -\mathbf{A} \\ -\mathbf{A} & \mathbf{K}_{w_2} \end{pmatrix}. \tag{149}$$

We have

$$
\begin{aligned}
\mathbf{\Delta}_2 \succ \mathbf{0} &\iff \mathbf{K}_{w_2} \succ \mathbf{A}\mathbf{K}_{w_1}^{-1}\mathbf{A} \\
&\iff \mathbf{K}_{w_2} + \mathbf{A} \succ \mathbf{A}\mathbf{K}_{w_1}^{-1}\mathbf{A} + \mathbf{A} \\
&\iff (\mathbf{K}_{w_2} + \mathbf{A})^{-1} \prec (\mathbf{A}\mathbf{K}_{w_1}^{-1}\mathbf{A} + \mathbf{A})^{-1} \\
&\iff (\mathbf{K}_{w_2} + \mathbf{A})^{-1} \prec \mathbf{A}^{-1} - (\mathbf{K}_{w_1} + \mathbf{A})^{-1} \\
&\iff (\mathbf{K}_{w_1} + \mathbf{A})^{-1} + (\mathbf{K}_{w_2} + \mathbf{A})^{-1} \prec \mathbf{A}^{-1} \\
&\overset{(a)}{\Longleftarrow} \sum_{l=1}^{L} (\mathbf{K}_{w_l} + \mathbf{A})^{-1} \prec \mathbf{A}^{-1} \\
&\overset{(b)}{\iff} (\mathbf{K}_{w_0} + \mathbf{A})^{-1} \prec \mathbf{A}^{-1} \\
&\iff \mathbf{K}_{w_0} + \mathbf{A} \succ \mathbf{A} \\
&\iff \mathbf{K}_{w_0} \succ 0,
\end{aligned}
\tag{150}
$$

where (a) is only one direction because

$$
(\mathbf{K}_{w_1} + \mathbf{A})^{-1} + (\mathbf{K}_{w_2} + \mathbf{A})^{-1} \prec \sum_{l=1}^{L} (\mathbf{K}_{w_l} + \mathbf{A})^{-1} \tag{151}
$$

and (b) is from (37).

Next we define

$$
\mathbf{\Delta}_k = \begin{pmatrix}
\mathbf{K}_{w_1} & -\mathbf{A} & -\mathbf{A} & \dots & -\mathbf{A} \\
-\mathbf{A} & \mathbf{K}_{w_2} & -\mathbf{A} & \dots & -\mathbf{A} \\
\dots & \dots & \dots & \dots & \dots \\
-\mathbf{A} & \dots & -\mathbf{A} & \mathbf{K}_{w_{k-1}} & -\mathbf{A} \\
-\mathbf{A} & \dots & -\mathbf{A} & -\mathbf{A} & \mathbf{K}_{w_k}
\end{pmatrix} \tag{152}
$$

and suppose $\mathbf{\Delta}_k \succ 0$ for $k = 3, \ldots, l-1$. Then

$$\begin{aligned}
\mathbf{\Delta}_l \succ 0 &\iff \mathbf{K}_{w_l} \succ \mathbf{A}(\mathbf{I}_N, \mathbf{I}_N, \ldots, \mathbf{I}_N)\Delta_{l-1}^{-1}(\mathbf{I}_N, \mathbf{I}_N, \ldots, \mathbf{I}_N)^t\mathbf{A} \\
&\iff \mathbf{K}_{w_l} \succ \mathbf{A}\left[\left(\sum_{k=1}^{l-1}(\mathbf{K}_{W_k}+\mathbf{A})^{-1}\right)^{-1} - \mathbf{A}\right]^{-1}\mathbf{A} \\
&\iff \mathbf{K}_{w_l} + \mathbf{A} \succ \mathbf{A}\left[\left(\sum_{k=1}^{l-1}(\mathbf{K}_{w_k}+\mathbf{A})^{-1}\right)^{-1} - \mathbf{A}\right]^{-1}\mathbf{A} + \mathbf{A} \\
&\iff (\mathbf{K}_{w_l} + \mathbf{A})^{-1} \prec \left[\mathbf{A}\left[\left(\sum_{k=1}^{l-1}(\mathbf{K}_{w_k}+\mathbf{A})^{-1}\right)^{-1} - \mathbf{A}\right]^{-1}\mathbf{A} + \mathbf{A}\right]^{-1} \\
&\iff (\mathbf{K}_{w_l} + \mathbf{A})^{-1} \prec \mathbf{A}^{-1} - \left[\left(\sum_{k=1}^{l-1}(\mathbf{K}_{w_k}+\mathbf{A})^{-1}\right)^{-1} - \mathbf{A} + \mathbf{A}\right]^{-1} \\
&\iff (\mathbf{K}_{w_l} + \mathbf{A})^{-1} \prec \mathbf{A}^{-1} - \sum_{k=1}^{l-1}(\mathbf{K}_{w_k}+\mathbf{A})^{-1} \\
&\overset{(c)}{\Longleftarrow} \sum_{k=1}^{L}(\mathbf{K}_{w_k}+\mathbf{A})^{-1} \prec \mathbf{A}^{-1} \\
&\overset{(d)}{\iff} (\mathbf{K}_{w_0}+\mathbf{A})^{-1} \prec \mathbf{A}^{-1} \\
&\iff \mathbf{K}_{w_0}+\mathbf{A} \succ \mathbf{A} \\
&\iff \mathbf{K}_{w_0} \succ 0,
\end{aligned} \tag{153}$$

where (c) is only one direction because

$$\sum_{k=1}^{l}(\mathbf{K}_{w_k}+\mathbf{A})^{-1} \prec \sum_{k=1}^{L}(\mathbf{K}_{w_k}+\mathbf{A})^{-1}, \quad \text{for} \quad l < L, \tag{154}$$

and (d) is from (37).

F Proof of Lemma 5

$$\begin{aligned}
\left[(\mathbf{K}_{w_0}+\mathbf{A}^*)^{-1}+\boldsymbol{\Lambda}_1\right]^{-1} &= \left[(\mathbf{K}_{w_0}+\mathbf{A}^*)^{-1}(\mathbf{I}_N+(\mathbf{K}_{w_0}+\mathbf{A}^*)\boldsymbol{\Lambda}_1)\right]^{-1} \\
&\overset{(a)}{=} (\mathbf{I}_N+\mathbf{K}_{w_0}\boldsymbol{\Lambda}_1)^{-1}(\mathbf{K}_{w_0}+\mathbf{A}^*) \\
&= (\mathbf{I}_N+\mathbf{K}_{w_0}\boldsymbol{\Lambda}_1)^{-1}(\mathbf{K}_{w_0}+\mathbf{A}^*-(\mathbf{I}_N+\mathbf{K}_{w_0}\boldsymbol{\Lambda}_1)\mathbf{A}^*)+\mathbf{A}^* \\
&\overset{(b)}{=} (\mathbf{I}_N+\mathbf{K}_{w_0}\boldsymbol{\Lambda}_1)^{-1}\mathbf{K}_{w_0}+\mathbf{A}^* \\
&= \left(\mathbf{K}_{w_0}^{-1}(\mathbf{I}_N+\mathbf{K}_{w_0}\boldsymbol{\Lambda}_1)\right)^{-1}+\mathbf{A}^* \\
&= \left(\mathbf{K}_{w_0}^{-1}+\boldsymbol{\Lambda}_1\right)^{-1}+\mathbf{A}^*,
\end{aligned} \tag{155}$$

where (a) and (b) are from $\boldsymbol{\Lambda}_1\mathbf{A}^*=\mathbf{0}$.

$$\begin{aligned}
\frac{|\mathbf{D}_0^*+\mathbf{K}_z|}{|\mathbf{D}_0^*|} &= |\mathbf{I}_N+\mathbf{D}_0^{*-1}\mathbf{K}_z| \\
&= |\mathbf{I}_N+(\mathbf{D}_0^{-1}+\boldsymbol{\Lambda}_1)\mathbf{K}_z| \\
&= |\mathbf{I}_N+\mathbf{D}_0^{-1}\mathbf{K}_z+\boldsymbol{\Lambda}_1\mathbf{K}_z| \\
&= |\mathbf{I}_N+\mathbf{D}_0^{-1}\mathbf{K}_z+\boldsymbol{\Lambda}_1\left((\mathbf{I}_N-\mathbf{A}^*)^{-1}-\mathbf{I}_N\right)| \\
&\overset{(c)}{=} |\mathbf{I}_N+\mathbf{D}_0^{-1}\mathbf{K}_z+\boldsymbol{\Lambda}_1(\mathbf{I}_N-\mathbf{A}^*)\left((\mathbf{I}_N-\mathbf{A}^*)^{-1}-\mathbf{I}_N\right)| \\
&= |\mathbf{I}_N+\mathbf{D}_0^{-1}\mathbf{K}_z| \\
&= \frac{|\mathbf{D}_0+\mathbf{K}_z|}{|\mathbf{D}_0|},
\end{aligned} \tag{156}$$

where (c) is from $\boldsymbol{\Lambda}_1\mathbf{A}^*=\mathbf{0}$.

G Proof of Equations (58) and (59)

We first prove the following lemma.

Lemma 10. *Let* $\mathbf{D}$ *be an* $N\times N$ *matrix such that* $\mathbf{0}\prec\mathbf{D}\prec\mathbf{I}_N$. *Let* $\mathbf{K}=(\mathbf{D}^{-1}-\mathbf{I}_N)^{-1}$. *Choose* $\epsilon>0$ *such that* $\mathbf{K}-\epsilon\mathbf{I}_N\succ\mathbf{0}$. *Define*

$$\mathbf{D}(\epsilon)\overset{def}{=}\left[(\mathbf{K}-\epsilon\mathbf{I}_N)^{-1}+\mathbf{I}_N\right]^{-1}.$$

Then, there exist constants $b_1\geq b_2>0$, *such that*

$$\mathbf{D}-b_1\epsilon\mathbf{I}_N+o(\epsilon)\prec\mathbf{D}(\epsilon)\prec\mathbf{D}-b_2\epsilon\mathbf{I}_N+o(\epsilon)$$

Proof. There exists an $N \times N$ orthogonal matrix $\mathbf{Q}$ such that

$$\mathbf{QKQ}^t = \text{diag}\{k_1, \ \dots, \ k_N\},$$

where $k_i > 0$ are eigenvalues of $\mathbf{K}$. We have

$$\begin{aligned}\mathbf{QDQ}^t &= \mathbf{Q}(\mathbf{K}^{-1} + \mathbf{I}_N)^{-1}\mathbf{Q}^t \\ &= \text{diag}\left\{\frac{k_1}{1+k_1}, \ \dots, \ \frac{k_N}{1+k_N}\right\},\end{aligned}$$

and

$$\begin{aligned}\mathbf{QD}(\epsilon)\mathbf{Q}^t &= \mathbf{Q}\Big[(\mathbf{K} - \epsilon\mathbf{I}_N)^{-1} + \mathbf{I}_N\Big]^{-1}\mathbf{Q}^t \\ &= \Big[(\text{diag}\{k_1, \ \dots, \ k_N\} - \epsilon\mathbf{I}_N)^{-1} + \mathbf{I}_N\Big]^{-1} \\ &= \text{diag}\left\{\frac{k_1 - \epsilon}{1 + k_1 - \epsilon}, \ \dots, \ \frac{k_N - \epsilon}{1 + k_N - \epsilon}\right\} \\ &= \text{diag}\left\{\frac{k_1}{1+k_1} - \frac{\epsilon}{(1+k_1)^2} + o(\epsilon), \ \dots, \ \frac{k_N}{1+k_N} - \frac{\epsilon}{(1+k_N)^2} + o(\epsilon)\right\}.\end{aligned}$$

We now have

$$\mathbf{QDQ}^t - b_1\epsilon\mathbf{I}_N + o(\epsilon) \prec \mathbf{QD}(\epsilon)\mathbf{Q}^t \prec \mathbf{QDQ}^t - b_2\epsilon\mathbf{I}_N + o(\epsilon),$$

where $b_1 \geq b_2 > 0$ are some constants. Hence

$$\mathbf{D} - b_1\epsilon\mathbf{I}_N + o(\epsilon) \prec \mathbf{D}(\epsilon) \prec \mathbf{D} - b_2\epsilon\mathbf{I}_N + o(\epsilon).$$

□

Equations (58) and (59) are a direct consequence of this lemma.

H Proof of Equation (60)

We first prove the following lemma.

Lemma 11. *Let $\mathbf{A}$ be an $N \times N$ matrix such that $\mathbf{0} \preccurlyeq \mathbf{A} \prec \mathbf{I}_N$. Let $\mathbf{K}_z = (\mathbf{I}_N - \mathbf{A})^{-1} - \mathbf{I}_N$. Choose $\epsilon > 0$ such that $\mathbf{A} + \epsilon\mathbf{I}_N \prec \mathbf{I}_N$. Define*

$$\mathbf{K}_z(\epsilon) \stackrel{def}{=} [\mathbf{I}_N - (\mathbf{A} + \epsilon\mathbf{I}_N)]^{-1} - \mathbf{I}_N.$$

Then, there exist constants $c_1 \geq c_2 > 0$ such that

$$\mathbf{K}_z - c_1\epsilon\mathbf{I}_N + o(\epsilon) \prec \mathbf{K}_z(\epsilon) \prec \mathbf{K}_z - c_2\epsilon\mathbf{I}_N + o(\epsilon).$$

Proof. There exists an $N \times N$ orthogonal matrix $\mathbf{Q}$ such that

$$\mathbf{Q}\mathbf{A}\mathbf{Q}^t = \mathrm{diag}\{a_1,\ \ldots,\ a_N\}$$

where $a_i > 0$ are the eigenvalues of $\mathbf{A}$. We have

$$\begin{aligned}\mathbf{Q}\mathbf{K}_z\mathbf{Q}^t &= \mathbf{Q}((\mathbf{I}_N - \mathbf{A})^{-1} - \mathbf{I}_N)\mathbf{Q}^t \\ &= \mathrm{diag}\left\{\frac{a_1}{1-a_1},\ \ldots,\ \frac{a_N}{1-a_N}\right\},\end{aligned}$$

and

$$\begin{aligned}\mathbf{Q}\mathbf{K}_z(\epsilon)\mathbf{Q}^t &= \mathbf{Q}((\mathbf{I}_N - (\mathbf{A} + \epsilon\mathbf{I}_N))^{-1} - \mathbf{I}_N)\mathbf{Q}^t \\ &= \mathrm{diag}\left\{\frac{a_1+\epsilon}{1-a_1-\epsilon},\ \ldots,\ \frac{a_N+\epsilon}{1-a_N-\epsilon}\right\} \\ &= \mathrm{diag}\left\{\frac{a_1}{1-a_1} - \frac{(2a_1-1)\epsilon}{(1-a_1)^2} + o(\epsilon),\ \ldots,\ \frac{a_N}{1-a_N} - \frac{(2a_N-1)\epsilon}{(1-a_N)^2} + o(\epsilon)\right\}.\end{aligned}$$

We now have

$$\mathbf{Q}\mathbf{K}_z\mathbf{Q}^t - c_1\epsilon\mathbf{I}_N + o(\epsilon) \prec \mathbf{Q}\mathbf{K}_z(\epsilon)\mathbf{Q}^t \prec \mathbf{Q}\mathbf{K}_z\mathbf{Q}^t - c_2\epsilon\mathbf{I}_N + o(\epsilon),$$

where $c_1 \geq c_2 > 0$ are some constants. Hence

$$\mathbf{K}_z - c_1\epsilon\mathbf{I}_N + o(\epsilon) \prec \mathbf{K}_z(\epsilon) \prec \mathbf{K}_z - c_2\epsilon\mathbf{I}_N + o(\epsilon).$$

□

Equation (60) is a direct result of this lemma.

I Proof of equation (73)

We first prove the following lemma.

Lemma 12. *Let $\mathbf{A}$ be an $N \times N$ matrix such that $\mathbf{0} \prec \mathbf{A} \preccurlyeq \mathbf{I}_N$. Choose $\epsilon > 0$ such that $\mathbf{A} - \epsilon\mathbf{I}_N \succ \mathbf{0}$. Define*

$$\mathbf{K}_z(\epsilon) \stackrel{def}{=} [\mathbf{I}_N - (\mathbf{A} - \epsilon\mathbf{I}_N)]^{-1} - \mathbf{I}_N.$$

Then, for any $\mathbf{E}$ and $\mathbf{F}$ such that $\mathbf{0} \prec \mathbf{E} \preccurlyeq \mathbf{I}_N$ and $\mathbf{0} \prec \mathbf{F} \preccurlyeq \mathbf{I}_N$, we have

$$\lim_{\epsilon \to 0} \frac{|\mathbf{E} + \mathbf{K}_z(\epsilon)|}{|\mathbf{F} + \mathbf{K}_z(\epsilon)|} = 1.$$

Proof. There exists an $N \times N$ orthogonal matrix $\mathbf{Q}$ such that

$$\mathbf{Q}\mathbf{A}\mathbf{Q}^t = \mathrm{diag}\{a_1, \ldots, a_N\},$$

where $0 < a_i \leq 1$ are eigenvalues of $\mathbf{A}$. Without loss of generality, we suppose $a_1 = 1, \ldots, a_p = 1, a_{p+1} < 1, \ldots, a_N < 1$.

We have

$$\begin{aligned}\mathbf{Q}\mathbf{K}_z(\epsilon)\mathbf{Q}^t &= \mathbf{Q}((\mathbf{I}_N - (\mathbf{A} - \epsilon\mathbf{I}_N))^{-1} - \mathbf{I}_N)\mathbf{Q}^t \\ &= \mathrm{diag}\left\{\frac{1-\epsilon}{\epsilon}, \ldots, \frac{1-\epsilon}{\epsilon}, \frac{a_{p+1}-\epsilon}{1-a_{p+1}+\epsilon}, \frac{a_N-\epsilon}{1-a_N+\epsilon}\right\},\end{aligned}$$

and since

$$\frac{|\mathbf{I}+\mathbf{K}_z(\epsilon)|}{|\mathbf{K}_z(\epsilon)|} \geq \frac{|\mathbf{E}+\mathbf{K}_z(\epsilon)|}{|\mathbf{F}+\mathbf{K}_z(\epsilon)|} \geq \frac{|\mathbf{K}_z(\epsilon)|}{|\mathbf{I}+\mathbf{K}_z(\epsilon)|},$$

we have

$$\lim_{\epsilon\to 0}\frac{|\mathbf{E}+\mathbf{K}_z(\epsilon)|}{|\mathbf{F}+\mathbf{K}_z(\epsilon)|} = 1.$$

□

Equation (73) is a direct consequence of this lemma.

J Proof of Equation (84)

We would like to have a property similar to (54), as ϵ_1 approaches zero, and a property similar to (73), as ϵ_2 approaches zero. To see this is the case, we need the following lemma.

***Lemma* 13.**

$$\mathbf{\Lambda}_1\mathbf{K}_z(\epsilon_1 = 0, \epsilon_2) = \mathbf{0}$$

Proof. Since

$$\mathbf{Q}\mathbf{\Lambda}_1\mathbf{Q}^t\mathbf{Q}\mathbf{A}^*\mathbf{Q}^t = \mathbf{0}$$

and

$$\begin{aligned}\mathbf{Q}\mathbf{A}^*\mathbf{Q}^t &= \mathrm{diag}(\underbrace{0, \ldots, 0}_{p}, \underbrace{1, \ldots, 1}_{q}, a_{p+q+1}, \ldots, a_s) \\ \mathbf{Q}\mathbf{A}^*\mathbf{Q}^t - \epsilon_2\mathbf{E}_2 &= \mathrm{diag}(\underbrace{0, \ldots, 0}_{p}, \underbrace{1-\epsilon_2, \ldots, 1-\epsilon_2}_{q}, a_{p+q+1}, \ldots, a_s),\end{aligned}$$

we have that

$$\mathbf{Q}\mathbf{A}^*\mathbf{Q}^t(\mathbf{Q}\mathbf{A}^*\mathbf{Q}^t - \epsilon_2\mathbf{E}_2) = \mathbf{0}.$$

Thus

$$\begin{aligned}
\mathbf{Q}\mathbf{\Lambda}_1\mathbf{K}_z(\epsilon_1 = 0, \epsilon_2)\mathbf{Q}^t &= \mathbf{Q}\mathbf{\Lambda}_1\mathbf{Q}^t\mathbf{Q}\left((\mathbf{I}_N - \mathbf{A}^* + \mathbf{Q}^t\epsilon_2\mathbf{E}_2\mathbf{Q})^{-1} - \mathbf{I}_N\right)\mathbf{Q}^t \\
&= \mathbf{Q}\mathbf{\Lambda}_1\mathbf{Q}^t\left((\mathbf{I}_N - \mathbf{Q}\mathbf{A}^*\mathbf{Q}^t + \epsilon_2\mathbf{E}_2)^{-1} - \mathbf{I}_N\right) \\
&= \mathbf{Q}\mathbf{\Lambda}_1\mathbf{Q}^t(\mathbf{I}_N - \mathbf{Q}\mathbf{A}^*\mathbf{Q}^t + \epsilon_2\mathbf{E}_2)\left((\mathbf{I}_N - \mathbf{Q}\mathbf{A}^*\mathbf{Q}^t + \epsilon_2\mathbf{E}_2)^{-1} - \mathbf{I}_N\right) \\
&= \mathbf{0}.
\end{aligned}$$

□

Using this lemma, we can show a property similar to (54) as ϵ_1 approaches zero. First note that similar to case 2, we have

$$\mathbf{D}_0^{-1} + \mathbf{\Lambda}_1 - e_2\epsilon_2\mathbf{I}_N + o(\epsilon_2) \prec \mathbf{D}_0^{-1}(\epsilon_1, \epsilon_2) \prec \mathbf{D}_0^{-1} + \mathbf{\Lambda}_1 + e_1\epsilon_1\mathbf{I}_N + o(\epsilon_1)$$

where $e_1 > 0$ and $e_2 > 0$ are constants. Hence we have

$$\begin{aligned}
\frac{|\mathbf{D}_0(\epsilon_1 = 0, \epsilon_2) + \mathbf{K}_z(\epsilon_1 = 0, \epsilon_2)|}{|\mathbf{D}_0(\epsilon_1 = 0, \epsilon_2)|} &= |\mathbf{I}_N + \mathbf{D}_0^{-1}(\epsilon_1 = 0, \epsilon_2)\mathbf{K}_z(\epsilon_1 = 0, \epsilon_2)| \\
&\geq |\mathbf{I}_N + (\mathbf{D}_0^{-1} + \mathbf{\Lambda}_1 - e_2\epsilon_2\mathbf{I}_N)\mathbf{K}_z(\epsilon_1 = 0, \epsilon_2)| \\
&= |\mathbf{I}_N + \mathbf{D}_0^{-1}\mathbf{K}_z(\epsilon_1 = 0, \epsilon_2) - e_2\epsilon_2\mathbf{K}_z(\epsilon_1 = 0, \epsilon_2)| \\
&= \frac{|\mathbf{D}_0 + \mathbf{K}_z(\epsilon_1 = 0, \epsilon_2) - e_2\epsilon_2\mathbf{D}_0\mathbf{K}_z(\epsilon_1 = 0, \epsilon_2)|}{|\mathbf{D}_0|}.
\end{aligned}$$

Similarly, we have

$$\frac{|\mathbf{D}_0(\epsilon_1 = 0, \epsilon_2) + \mathbf{K}_z(\epsilon_1 = 0, \epsilon_2)|}{|\mathbf{D}_0(\epsilon_1 = 0, \epsilon_2)|} \leq \frac{|\mathbf{D}_0 + \mathbf{K}_z(\epsilon_1 = 0, \epsilon_2)|}{|\mathbf{D}_0|}.$$

Thus

$$\begin{aligned}
&\lim_{\epsilon_2 \to 0}\lim_{\epsilon_1 \to 0}\frac{1}{2}\log\frac{|\mathbf{I}_N + \mathbf{K}_z(\epsilon_1, \epsilon_2)|^{(L-1)}|\mathbf{D}_0(\epsilon_1, \epsilon_2) + \mathbf{K}_z(\epsilon_1, \epsilon_2)|}{|\mathbf{D}_0(\epsilon_1, \epsilon_2)|\prod_{l=1}^{L}|\mathbf{D}_l(\epsilon_1, \epsilon_2) + \mathbf{K}_z(\epsilon_1, \epsilon_2)|} \\
&\qquad= \lim_{\epsilon_2 \to 0}\frac{1}{2}\log\frac{|\mathbf{I}_N + \mathbf{K}_z(\epsilon_1 = 0, \epsilon_2)|^{(L-1)}|\mathbf{D}_0 + \mathbf{K}_z(\epsilon_1 = 0, \epsilon_2)|}{|\mathbf{D}_0|\prod_{l=1}^{L}|\mathbf{D}_l(\epsilon_1 = 0, \epsilon_2) + \mathbf{K}_z(\epsilon_1 = 0, \epsilon_2)|} \qquad (157) \\
&\qquad= \frac{1}{2}\log\frac{1}{|\mathbf{D}_0|},
\end{aligned}$$

where the last step is similar to (73).

K Proof of Proposition 4

Consider equation (108), which is rewritten in the following

$$\mathbf{A}^* = (\mathbf{K}_{w_1} - \mathbf{K}_{w_0})^{\frac{1}{2}} \left[(\mathbf{K}_{w_1} - \mathbf{K}_{w_0})^{-\frac{1}{2}}(\mathbf{K}_{w_2} - \mathbf{K}_{w_0})(\mathbf{K}_{w_1} - \mathbf{K}_{w_0})^{-\frac{1}{2}}\right]^{\frac{1}{2}} (\mathbf{K}_{w_1} - \mathbf{K}_{w_0})^{\frac{1}{2}} - \mathbf{K}_{w_0}. \tag{158}$$

We find that

$$\begin{aligned}
& \mathbf{A}^* \succ \mathbf{0} \\
\Longleftrightarrow\ & (\mathbf{K}_{w_1} - \mathbf{K}_{w_0})^{\frac{1}{2}} \left[(\mathbf{K}_{w_1} - \mathbf{K}_{w_0})^{-\frac{1}{2}}(\mathbf{K}_{w_2} - \mathbf{K}_{w_0})(\mathbf{K}_{w_1} - \mathbf{K}_{w_0})^{-\frac{1}{2}}\right]^{\frac{1}{2}} (\mathbf{K}_{w_1} - \mathbf{K}_{w_0})^{\frac{1}{2}} \succ \mathbf{K}_{w_0} \\
\Longleftrightarrow\ & \left[(\mathbf{K}_{w_1} - \mathbf{K}_{w_0})^{-\frac{1}{2}}(\mathbf{K}_{w_2} - \mathbf{K}_{w_0})(\mathbf{K}_{w_1} - \mathbf{K}_{w_0})^{-\frac{1}{2}}\right]^{\frac{1}{2}} \succ (\mathbf{K}_{w_1} - \mathbf{K}_{w_0})^{-\frac{1}{2}}\mathbf{K}_{w_0}(\mathbf{K}_{w_1} - \mathbf{K}_{w_0})^{-\frac{1}{2}} \\
\Longleftarrow\ & (\mathbf{K}_{w_1} - \mathbf{K}_{w_0})^{-\frac{1}{2}}(\mathbf{K}_{w_2} - \mathbf{K}_{w_0})(\mathbf{K}_{w_1} - \mathbf{K}_{w_0})^{-\frac{1}{2}} \\
& \quad \succ (\mathbf{K}_{w_1} - \mathbf{K}_{w_0})^{-\frac{1}{2}}\mathbf{K}_{w_0}(\mathbf{K}_{w_1} - \mathbf{K}_{w_0})^{-1}\mathbf{K}_{w_0}(\mathbf{K}_{w_1} - \mathbf{K}_{w_0})^{-\frac{1}{2}} \\
\Longleftrightarrow\ & \mathbf{K}_{w_2} - \mathbf{K}_{w_0} \succ \mathbf{K}_{w_0}\left(\mathbf{K}_{w_1} - \mathbf{K}_{w_0}\right)^{-1}\mathbf{K}_{w_0} \\
\Longleftrightarrow\ & \mathbf{K}_{w_2} - \mathbf{K}_{w_0} \succ \mathbf{K}_{w_0}\left(-\mathbf{I}_N + (\mathbf{K}_{w_1} - \mathbf{K}_{w_0})^{-1}\right)\mathbf{K}_{w_1} \\
\Longleftrightarrow\ & \mathbf{K}_{w_2} - \mathbf{K}_{w_0} \succ -\mathbf{K}_{w_0} + \mathbf{K}_{w_0}(\mathbf{K}_{w_1} - \mathbf{K}_{w_0})^{-1}\mathbf{K}_{w_1} \\
\Longleftrightarrow\ & \mathbf{K}_{w_2} \succ \mathbf{K}_{w_0}(\mathbf{K}_{w_1} - \mathbf{K}_{w_0})^{-1}\mathbf{K}_{w_1} \\
\Longleftrightarrow\ & \mathbf{K}_{w_2} \succ \mathbf{K}_{w_0}\mathbf{K}_{w_0}^{-1}(\mathbf{K}_{w_0}^{-1} - \mathbf{K}_{w_1}^{-1})^{-1}\mathbf{K}_{w_1}^{-1}\mathbf{K}_{w_1} \\
\Longleftrightarrow\ & \mathbf{K}_{w_2} \succ (\mathbf{K}_{w_0}^{-1} - \mathbf{K}_{w_1}^{-1})^{-1} \\
\Longleftrightarrow\ & \mathbf{K}_{w_1}^{-1} + \mathbf{K}_{w_2}^{-1} \prec \mathbf{K}_{w_0}^{-1} \\
\Longleftrightarrow\ & \mathbf{D}_0^{-1} + \mathbf{K}_x^{-1} - \mathbf{D}_1^{-1} - \mathbf{D}_2^{-1} \succ \mathbf{0}.
\end{aligned} \tag{159}$$

We thus have

$$\mathbf{D}_0^{-1} + \mathbf{K}_x^{-1} - \mathbf{D}_1^{-1} - \mathbf{D}_2^{-1} \succ \mathbf{0} \Rightarrow \mathbf{A}^* \succ \mathbf{0}. \tag{160}$$

The proof of

$$\mathbf{D}_0 + \mathbf{K}_x - \mathbf{D}_1 - \mathbf{D}_2 \succ \mathbf{0} \Rightarrow \mathbf{A}^* \prec \mathbf{K}_x \tag{161}$$

is similar and hence is omitted.

References

[1] 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.

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

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

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

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

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

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

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

[9] H. Feng and M. Effros, "On the rate loss of multiple description source codes," *IEEE Trans. Inform. Theory*, vol. 51, no. 2, pp. 671–683, Feb. 2005.

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

[11] V. A. Vaishampayan, "Design of multiple description scalar quantizers," *IEEE Trans. Inform. Theory*, vol. 39, no. 3, pp. 821–834, May 1993.

[12] V. A. Vaishampayan and J. Domaszewicz, "Design of entropy-constrained multiple-description scalar quantizers," *IEEE Trans. Inform. Theory*, vol. 40, no. 1, pp. 245–250, Jan. 1994.

[13] V. A. Vaishampayan and J. C. Batllo, "Asymptotic analysis of multiple-description quantizers," *IEEE Trans. Inform. Theory*, vol. 44, no. 1, pp. 278–284, Jan. 1998.

[14] V. A. Vaishampayan, N. Sloane, and S. Servetto, "Multiple description vector quantizers with lattice codebooks: design and analysis," *IEEE Trans. Inform. Theory*, vol. 47, no. 4, pp. 1718–1734, July 2001.

[15] V. K. Goyal, "Multiple description coding: compression meets the network," *IEEE Signal Processing Mag.*, vol. 18, pp. 74–93, Sept. 2001.

[16] S. N. Diggavi, N. J. A. Sloane, and V. A. Vaishampayan, "Asymmetric multiple description lattice vector quantizers," *IEEE Trans. Inform. Theory*, vol. 48, no. 1, pp. 174–191, Jan. 2002.

[17] V. K. Goyal, J. A. Kelner, and J. Kovacevic, "Multiple description vector quantization with a coarse lattice," *IEEE Trans. Inform. Theory*, vol. 48, no. 3, pp. 781–788, Mar. 2002.

[18] C. Tian and S. S. Hemami, "Universal multiple description scalar quantizer: analysis and design," *IEEE Trans. Inform. Theory*, vol. 50, no. 9, pp. 2737–2751, Sept. 2004.

[19] P. Ishwar, R. Puri, S. S. Pradhan, and K. Ramchandran, "On compression for robust estimation in sensor networks," in *Proc. IEEE Int. Symp. Inform. Theory (ISIT)*, Yokohama, Japan, June-July 2003.

[20] J. Chen and T. Berger, "Robust distributed source coding," submitted to *IEEE Trans. Inform. Theory*, 2005.

[21] N. Alon and J. H. Spencer, *The probabilistic Method, 2nd edition.* New York: Wiley, 2000.

[22] J. Edmonds, "Submodular functions, matroids and certain polyhedra," *Proc. Calgary Int. Conf. Combinatorial Structures and Applications*, pp. 69–87, Calgary, Alta, June 1969.

[23] D. J. A. Welsh, *Matroid Theory*, Academic Press, London, 1976.

[24] D. Tse and S. Hanly, "Multi-access fading channels: part I: polymatroid structure,optimal resource allocation and throughput capacities," *IEEE Trans. Inform. Theory*, vol. 44, no. 7, pp. 2796–2815, Nov. 1998.

[25] P. Viswanath, "Sum rate of a class of Gaussian multiterminal source coding problems," in *Advances in Network Information Theory*, P. Gupta, G. Kramer and A. Wijngaarden editors, Rutgers, NJ, 2004, pp. 43–60.

[26] Y. Oohama, "Rate-distortion theory for Gaussian multiterminal source coding systems with several side informations at the decoder," *IEEE Trans. Inform. Theory*, vol. 51, no. 7, pp. 2577–2593, Jul. 2005.

[27] A. B. Wagner, S. Tavildar and P. Viswanath, "Rate Region of the Quadratic Gaussian Two-Terminal Source-Coding Problem", submitted to *IEEE Trans. Inform. Theory*, Feb. 2006.

[28] S Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge, UK, 2004.

[29] H. L. Royden, *Real Analysis*, 3rd edition, Prentice-Hall Inc., NJ, 1988.

[30] F. Zhang, *Matrix Theory: Basic Results and Techniques.* Springer, 1999.

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