2005 Conference on Information Sciences and Systems, The Johns Hopkins University, March 16–18, 2005

Achieving the Multiple Description Rate-Distortion Region with Lattice Quantization

Jun Chen, Chao Tian, Toby Berger[1], and Sheila S. Hemami
School of Electrical and Computer Engineering
Cornell University, Ithaca, NY 14853
jc353@cornell.edu, {ctian, berger, hemami}@ece.cornell.edu

Abstract —

We propose a multiple description coding scheme based on the quantization splitting method. This scheme is able to achieve the whole EGC region. For the Gaussian source, we show that any rate pair on the dominant face of the EGC region is achievable with three or fewer lattice quantizers.

I. Introduction

The problem of multiple descriptions (MD) has been investigated for many years. The MD literature is vast; see, for example, [1–4].

Let $\{X(t)\}_{t=1}^{\infty}$ be an i.i.d. random process with $X(t) \sim P(x)$ for all t. Let $d(\cdot,\cdot) : \mathcal{X} \times \mathcal{X} \to \mathcal{R}^+$ be a distortion measure.

Definition 1. *The quintuple $(R_1, R_2, D_1, D_2, D_3)$ is called achievable, if $\forall \varepsilon > 0$, $\exists n_0$ such that $\forall n > n_0$ there exist encoding functions:*

$$f_i^{(n)} : \mathcal{X}^n \to \mathcal{C}_i^{(n)} \quad \log|\mathcal{C}_i^{(n)}| \le n(R_i + \varepsilon) \quad i = 1, 2,$$

and decoding functions:

$$g_i^{(n)} : \mathcal{C}_i^{(n)} \to \mathcal{X}^n \quad i = 1, 2$$

$$g_3^{(n)} : \mathcal{C}_1^{(n)} \times \mathcal{C}_2^{(n)} \to \mathcal{X}^n$$

such that for $\hat{X}_i^n = g_i^{(n)}(f_i^{(n)}(X^n)), i = 1, 2$, and for $\hat{X}_3^n = g_3^{(n)}(f_1^{(n)}(X^n), f_2^{(n)}(X^n))$, we have

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

The MD rate-distortion region, denoted by $\mathcal{Q}$, is the set of all achievable quintuples.

Definition 2 (EGC region [1]). $(R_1, R_2, D_1, D_2, D_3) \in \mathcal{Q}$ *if there exist random variables $\hat{X}_1, \hat{X}_2$ and $\hat{X}_3$ jointly distributed with the generic source variable X such that*

1. $R_i \ge I(X; \hat{X}_i)$, $i = 1, 2$;
2. $R_1 + R_2 \ge I(X; \hat{X}_1, \hat{X}_2, \hat{X}_3) + I(\hat{X}_1; \hat{X}_2)$;
3. $\mathbb{E}d(X, \hat{X}_i) \le D_i$, $i = 1, 2, 3$.

We denote the EGC region by $\mathcal{Q}_{EGC}$.

[1]This research is supported in part by NSF Grant CCR-033 0059 and a grant from the National Academies Keck Futures Initiative (NAKFI).

Ozarow [2] showed that $\mathcal{Q}_{EGC} = \mathcal{Q}$ for Gaussian sources. Ahlswede [3] showed that the EGC region is also tight for the "no excess sum-rate" case. Zhang and Berger [4] constructed a counterexample for which $\mathcal{Q}_{EGC} \subsetneqq \mathcal{Q}$.

Many practical MD schemes have been proposed [5–15]. Since it has been shown [16] that the Gaussian MD bound is asymptotically tight in the high rate regime for all memoryless continuous sources, the performance of MD schemes is often compared with the optimum solutions for the quadratic memoryless Gaussian case. The main contribution of this paper is that we propose a lattice-based quantization system which is able to achieve the whole Gaussian MD region.

The rest of this paper is divided into 3 sections. In Section II, we propose a "quantization splitting"-based MD scheme, which is able to achieve the whole EGC region. In Section III, we show that for the quadratic Gaussian case, our scheme can be realized with three or fewer lattice quantizers. We conclude the paper in Section IV.

II. Achieving the EGC Region with Quantization Splitting

We rewrite the EGC region in the following form[2]:

1. $R_i \ge I(X; \hat{X}_i)$, $i = 1, 2$;
2. $R_1 + R_2 \ge I(X; \hat{X}_1, \hat{X}_2) + I(\hat{X}_1; \hat{X}_2) + I(X, \hat{X}_3|\hat{X}_1, \hat{X}_2)$.

The $I(X, \hat{X}_3|\hat{X}_1, \hat{X}_2)$ term is the rate used for the superimposed refinement, which can be separated from other parts of the EGC scheme. Henceforth, we shall refer to the following region as the simplified EGC (SEGC) region:

1. $R_i \ge I(X; \hat{X}_i)$, $i = 1, 2$;
2. $R_1 + R_2 \ge I(X; \hat{X}_1, \hat{X}_2) + I(\hat{X}_1; \hat{X}_2)$.

The SEGC region is in general different from the EGC region since the $I(X, \hat{X}_3|\hat{X}_1, \hat{X}_2)$ term is removed. But it captures the main ingredients of the EGC region. Furthermore, as we will show later, the SEGC region is identical with the EGC region for the quadratic Gaussian case. The typical shape of the SEGC region is shown in Fig. 1. Since

$$\begin{aligned} & I(X; \hat{X}_1) + I(X; \hat{X}_2) \\ = \;& I(X; \hat{X}_1, \hat{X}_2) + I(\hat{X}_1; \hat{X}_2) - I(\hat{X}_1; \hat{X}_2|X) \\ \le \;& I(X; \hat{X}_1, \hat{X}_2) + I(\hat{X}_1; \hat{X}_2), \end{aligned}$$

the sum-rate constraint is always effective. We call

$$\Big\{(R_1, R_2) : R_1 + R_2 = I(X; \hat{X}_1, \hat{X}_2) + I(\hat{X}_1; \hat{X}_2), \quad R_i \ge I(X; \hat{X}_i), i = 1, 2\Big\}$$

[2]Rigorously speaking, in order to determine the EGC region, we need to take the union over all feasible joint distributions $P(x, \hat{x}_1, \hat{x}_2, \hat{x}_3)$ and then convexify the resulting region. In order to simplify the analysis, we only consider the EGC region with respect to a fixed joint distribution $P(x, \hat{x}_1, \hat{x}_2, \hat{x}_3)$.

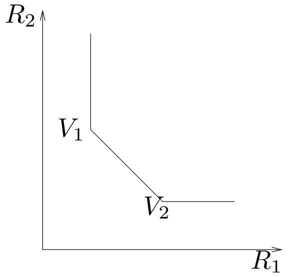

Figure 1: A typical SEGC region.

the dominant face of the SEGC region. Any rate pair inside the SEGC region is inferior to some rate pair on the dominant face in the sense of compression efficiency. Hence in searching for the optimal scheme, we can restrict our attention to rate pairs on the dominant face without loss of generality.

The dominant face of the SEGC region has two vertices, which are given respectively by

V_1: $R_1 = I(X; \hat{X}_1)$, $R_2 = I(X, \hat{X}_1; \hat{X}_2)$;

V_2: $R_1 = I(X, \hat{X}_2; \hat{X}_1)$, $R_2 = I(X; \hat{X}_2)$.

By symmetry, we shall only consider V_1. We first outline two schemes that can achieve V_1.

1. Successive Encoding Scheme: Encoder 1 independently generates $2^{nI(X;\hat{X}_1)}$ codewords $\{\hat{X}_1^n(j)\}_{j=1}^{2^{nI(X;\hat{X}_1)}}$ according to the marginal distribution $P(\hat{x}_1)$. Encoder 2 independently generates $2^{nI(X,\hat{X}_1;\hat{X}_2)}$ codewords $\{\hat{X}_2^n(k)\}_{k=1}^{2^{nI(X,\hat{X}_1;\hat{X}_2)}}$ according to the marginal distribution $P(\hat{x}_2)$.

   Encoding Procedure: Given X^n, encoder 1 finds the codeword $\hat{X}_1^n(j^*)$ such that X^n and $\hat{X}_1^n(j^*)$ are jointly typical, then encoder 2 finds the codeword $\hat{X}_2^n(k^*)$ such that $\hat{X}_2^n(k^*)$ is jointly typical with X^n and $\hat{X}_1^n(j^*)$. Index j^* is transmitted to the decoder through channel 1 and index k^* is transmitted to the decoder through channel 2.

   For this coding scheme, encoder 1 does the encoding first and then encoder 2. The main complexity of this scheme is on encoder 2 since it needs to construct a codebook that covers the $(X, \hat{X}_1)$-space instead of just the X-space. So intuitively, for this scheme, encoder 1 is a simple quantizer while encoder 2 is a complicated quantizer.

2. Two-Stage Encoding Scheme: We write R_2 as $R_2 = I(X; \hat{X}_2) + I(\hat{X}_1; \hat{X}_2|X)$.

   Encoder 1 independently generates $2^{nI(X;\hat{X}_1)}$ codewords $\{\hat{X}_1^n(j)\}_{j=1}^{2^{nI(X;\hat{X}_1)}}$ according to the marginal distribution $P(\hat{x}_1)$. Encoder 2 independently generates $2^{nI(\hat{X}_1;\hat{X}_2|X)}$ codebooks, with each codebook containing $2^{nI(X;\hat{X}_2)}$ codewords. That is, codebook i has codewords $\{\hat{X}_2^n(i,k)\}_{k=1}^{2^{nI(X;\hat{X}_2)}}$, $i = 1, 2, \cdots, 2^{nI(\hat{X}_1;\hat{X}_2|X)}$.

   Encoding Procedure:

   (a) Given X^n, encoder 1 finds the codeword $\hat{X}_1^n(j^*)$ such that X^n and $\hat{X}_1^n(j^*)$ are jointly typical. Encoder 2 finds, in each codebook i, a codeword $\hat{X}_2^n(i, k_i^*)$ such that $\hat{X}_2^n(i, k_i^*)$ is jointly typical with X^n, $i = 1, 2, \cdots, 2^{nI(\hat{X}_1;\hat{X}_2|X)}$. Encoder 2 then forms a new codebook with these $2^{nI(\hat{X}_1;\hat{X}_2|X)}$ selected codewords.

   (b) In this newly-formed codebook, encoder 2 finds a codeword $\hat{X}_2^n(i^*, k_{i^*}^*)$ such that $\hat{X}_2^n(i^*, k_{i^*}^*)$ is jointly typical with X^n and $\hat{X}_1^n(j^*)$.

   Index j^* is transmitted to the decoder through channel 1. Indices i^* and $k_{i^*}^*$ are transmitted to the decoder through channel 2.

   Encoder 1 is essentially the same as the one in the successive encoding scheme. The main difference is on encoder 2. Instead of constructing a giant codebook that covers the $(X, \hat{X}_1)$-space in the successive encoding scheme, encoder 2 , in the two-stage encoding scheme, generates $2^{nI(\hat{X}_1;\hat{X}_2|X)}$ small codebooks, each of which is able to cover the X-space. From the practical viewpoint, it is often easier to construct a bunch of simple quantizers than a single complicated quantizer. But the two-stage coding scheme requires an additional complexity on the selection of quantizers. That is, encoder 2 need select the output of one quantizer to transmit, and such a selection is determined by not only X^n, but also the output of encoder 1.

Now we proceed to study an arbitrary rate pair on the dominant face of the SEGC region. We shall outline four coding schemes.

1. Time Sharing: Any point on the dominant face can be viewed as a linear combination of V_1 and V_2. Hence the coding scheme for the rate pair on the dominant face follows directly from timesharing the encoding schemes for V_1 and V_2.

2. Joint Encoding [1]: Encoder 1 generates 2^{nR_1} codewords $\{\hat{X}_1^n(j)\}_{j=1}^{2^{nR_1}}$ according to the marginal distribution $P(\hat{x}_1)$. Encoder 2 generates 2^{nR_2} codewords $\{\hat{X}_2^n(k)\}_{k=1}^{2^{nR_2}}$ according to the marginal distribution $P(\hat{x}_2)$.

   Encoding Procedure: Given X^n, encoder 1 and encoder 2 find the codewords $\hat{X}_1^n(j^*)$ and $\hat{X}_2^n(k^*)$ such that X^n, $\hat{X}_1^n(j^*)$ and $\hat{X}_2^n(k^*)$ are jointly typical. Index j^* is transmitted to the decoder through channel 1 and index k^* is transmitted to the decoder through channel 2.

3. Two-Stage Encoding: Encoder 1 independently generates $2^{n[R_1 - I(X;\hat{X}_1)]}$ codebooks, with each codebook containing $2^{nI(X;\hat{X}_1)}$ codewords, i.e., codebook i contains codewords $\{\hat{X}_1^n(i,j)\}_{j=1}^{2^{nI(X;\hat{X}_1)}}$, $i = 1, 2, \cdots, 2^{n[R_1 - I(X;\hat{X}_1)]}$. Similarly, encoder 2 independently generates $2^{n[R_2 - I(X;\hat{X}_2)]}$ codebooks, with each codebook containing $2^{nI(X;\hat{X}_2)}$ codewords. That is, codebook k has codewords $\{\hat{X}_2^n(k,l)\}_{k=1}^{2^{nI(X;\hat{X}_2)}}$, $k = 1, 2, \cdots, 2^{n[R_2 - I(X;\hat{X}_2)]}$.

   Encoding Procedure:

   (a) Given X^n, encoder 1 finds, in each codebook i, a codeword $\hat{X}_1^n(i, j_i^*)$ such that $\hat{X}_1^n(i, j_i^*)$ is jointly typical with X^n, $i = 1, 2, \cdots, 2^{n[R_1 - I(X;\hat{X}_1)]}$. Encoder 1 then forms a new codebook with these $2^{n[R_1 - I(X;\hat{X}_1)]}$ selected codewords. Similarly encoder 2 finds, in each codebook k, a codeword $\hat{X}_2^n(k, l_k^*)$ such that $\hat{X}_2^n(k, l_k^*)$ is jointly typical with X^n, $i = 1, 2, \cdots, 2^{n[R_2 - I(X;\hat{X}_2)]}$. Encoder 2 then forms a new codebook with these $2^{n[R_2 - I(X;\hat{X}_2)]}$ selected codewords.

(b) In these two newly-formed codebooks, encoder 1 and encoder 2 find codewords $\hat{X}_1^n(i^*, j^*_{i^*})$ and $\hat{X}_2^n(k^*, l^*_{k^*})$ such that X^n, $\hat{X}_1^n(i^*, j^*_{i^*})$ and $\hat{X}_2^n(k^*, l^*_{k^*})$ are jointly typical.

Indices i^* and $j^*_{i^*}$ are transmitted to the decoder through channel 1. Indices k^* and $l^*_{k^*}$ are transmitted to the decoder through channel 2.

The second stage of this scheme is similar to that of the joint encoding scheme, except that the size of its search space is $2^{n[R_1+R_2-I(X;\hat{X}_1)-I(X;\hat{X}_2)]} = 2^{nI(\hat{X}_1;\hat{X}_2|X)}$ while the size of the search space for the joint encoding scheme is $2^{n(R_1+R_2)} = 2^{n[I(X;\hat{X}_1,\hat{X}_2)+I(\hat{X}_1;\hat{X}_2)]}$. But this reduction in the size of the search space is at the cost of introducing an additional encoding stage—the first stage.

4. Quantization Splitting: We need the following result before presenting the coding scheme.

Lemma 1. *For any rate pair (R_1, R_2) on the dominant face of the SEGC region, we can find $\hat{X}'_2$ with $(X, \hat{X}_1) \to \hat{X}_2 \to \hat{X}'_2$ such that*

$$\begin{aligned} R_1 &= I(X, \hat{X}'_2; \hat{X}_1), \\ R_2 &= I(X, \hat{X}'_2) + I(X, \hat{X}_1; \hat{X}_2|\hat{X}'_2), \end{aligned}$$

Symmetrically, we can find $\hat{X}'_1$ with $(X, \hat{X}_2) \to \hat{X}_1 \to \hat{X}'_1$ such that

$$\begin{aligned} R_1 &= I(X, \hat{X}'_1) + I(X, \hat{X}_2; \hat{X}_1|\hat{X}'_1), \\ R_2 &= I(X, \hat{X}'_1; \hat{X}_2). \end{aligned}$$

Proof. We shall only prove the first form, the other one follows by symmetry.

If we let $\hat{X}'_2 =$ constant, then

$$R_1 = I(X; \hat{X}_1), \quad R_2 = I(X, \hat{X}_1; \hat{X}_2),$$

which corresponds to Vertex 1.

If we let $\hat{X}'_2 = \hat{X}_2$, then

$$R_1 = I(X, \hat{X}_2; \hat{X}_1), \quad R_2 = I(X, \hat{X}_2),$$

which corresponds to Vertex 2.

Clearly, we can construct a class of transition probabilities $P_\epsilon(\hat{x}'_2|\hat{x}_2)$ indexed by ϵ such that $I(X; \hat{X}'_2)$ changes continuously from 0 to $I(X; \hat{X}_2)$ as ϵ changes from 0 to 1. Now we only need to verify

$$R_1 + R_2 = I(X; \hat{X}_1, \hat{X}_2) + I(\hat{X}_1; \hat{X}_2).$$

It follows that

$$\begin{aligned} & R_1 + R_2 \\ = \; & I(X, \hat{X}'_2; \hat{X}_1) + I(X; \hat{X}'_2) + I(X, \hat{X}_1; \hat{X}_2|\hat{X}'_2) \\ = \; & I(X, \hat{X}'_2; \hat{X}_1) + I(X; \hat{X}'_2) + I(X; \hat{X}_2|\hat{X}'_2) \\ & + I(\hat{X}_1; \hat{X}_2|X, \hat{X}'_2) \\ = \; & I(X, \hat{X}_2, \hat{X}'_2; \hat{X}_1) + I(X; \hat{X}_2, \hat{X}'_2). \end{aligned}$$

Since $(X, \hat{X}_1) \to \hat{X}_2 \to \hat{X}'_2$, we have

$$\begin{aligned} & I(X, \hat{X}_2, \hat{X}'_2; \hat{X}_1) + I(X; \hat{X}_2, \hat{X}'_2) \\ = \; & I(X, \hat{X}_2; \hat{X}_1) + I(X; \hat{X}_2) \\ = \; & I(X, \hat{X}_1) + I(X; \hat{X}_2) + I(\hat{X}_1; \hat{X}_2|X) \\ = \; & I(X; \hat{X}_1, \hat{X}_2) + I(\hat{X}_1; \hat{X}_2), \end{aligned}$$

which completes the proof. □

Now we are ready to state the coding scheme based on the quantization splitting.

We will only discuss the first representation, i.e.,

$$\begin{aligned} R_1 &= I(X, \hat{X}'_2; \hat{X}_1), \\ R_2 &= I(X, \hat{X}'_2) + I(X, \hat{X}_1; \hat{X}_2|\hat{X}'_2). \end{aligned}$$

Encoder 1 independently generates $2^{nI(X,\hat{X}'_2;\hat{X}_1)}$ codewords $\{\hat{X}_1^n(i)\}_{i=1}^{2^{nI(X,\hat{X}'_2;\hat{X}_1)}}$ according to the marginal distribution $P(\hat{x}_1)$. Encoder 2 independently generates $2^{nI(X;\hat{X}'_2)}$ codewords $\{\hat{X}'^n_2(j)\}_{j=1}^{2^{nI(X;\hat{X}'_2)}}$ according to the marginal distribution $P(\hat{x}'_2)$. For each codeword $\hat{X}'^n_2(j)$, encoder 2 independently generates $2^{nI(X,\hat{X}_1;\hat{X}_2|\hat{X}'_2)}$ codewords $\{\hat{X}_2^n(j,k)\}_{k=1}^{2^{nI(X,\hat{X}_1;\hat{X}_2|\hat{X}'_2)}}$ according to the marginal conditional distribution $P(\hat{x}_2|\hat{x}'_2)$.

Encoding Procedure: Given X^n, encoder 2 finds the codeword $\hat{X}'^n_2(j^*)$ such that X^n and $\hat{X}'^n_2(j^*)$ are jointly typical. Then encoder 1 finds the codeword $\hat{X}_1^n(i^*)$ such that $\hat{X}_1^n(i^*)$ is jointly typical with X^n and $\hat{X}'^n_2(j^*)$. Finally, encoder 2 finds the codeword $\hat{X}_2^n(j^*, k^*)$ such that $\hat{X}_2^n(j^*, k^*)$ is jointly typical with X^n, $\hat{X}_1^n(i^*)$ and $\hat{X}'^n_2(j^*)$. Index i^* is transmitted to the decoder through channel 1. Indices j^* and k^* are transmitted to the decoder through channel 2.

This approach is a natural generalization of the successive encoding scheme for the vertices of the SEGC region.

We can view $\hat{X}'_2$ as a coarse description of X and view $\hat{X}_2$ as a fine description of X. The idea of introducing an auxiliary coarse description to convert a joint coding scheme to a successive coding scheme has been widely used in the distributed source coding problem [17, 18].

III. Gaussian Case

In the preceding section, we gave an information theoretic analysis of the SEGC region. We shall apply those general results to the quadratic Gaussian case. As we will see, those general results possess particularly simple interpretations in the Gaussian case.

Let $\{X(t)\}_{t=1}^{\infty}$ be an i.i.d. Gaussian process with $X(t) \sim \mathcal{N}(0, \sigma_X^2)$ for all t. Let $d(\cdot,\cdot)$ be the squared distortion measure. It was shown in [1, 2] that $(R_1, R_2, D_1, D_2, D_3) \in \mathcal{Q}$ if and only if

$$\begin{aligned} R_i &\geq \frac{1}{2}\log\frac{\sigma_X^2}{D_i}, \quad i = 1, 2, \\ R_1 + R_2 &\geq \frac{1}{2}\log\frac{\sigma_X^2}{D_3} + \frac{1}{2}\log\psi(D_1, D_2, D_3), \end{aligned}$$

where

$$\psi(D_1,D_2,D_3) = \begin{cases} 1, & D_3 < D_1+D_2-\sigma_X^2 \\ \frac{\sigma_X^2 D_3}{D_1 D_2}, & D_3 > \left(\frac{1}{D_1}+\frac{1}{D_2}-\frac{1}{\sigma_X^2}\right)^{-1} \\ \frac{(\sigma_X^2-D_3)^2}{(\sigma_X^2-D_3)^2-[\sqrt{(\sigma_X^2-D_1)(\sigma_X^2-D_2)}-\sqrt{(D_1-D_3)(D_2-D_3)}]^2}, & \text{o.w.} \end{cases}$$

The case $D_3 < D_1 + D_2 - \sigma_X^2$ and the case $D_3 > \left(1/D_1 + 1/D_2 - 1/\sigma_X^2\right)^{-1}$ are degenerated. It is easy to verify that for any $(R_1, R_2, D_1, D_2, D_3) \in \mathcal{Q}$ with $D_3 < D_1 + D_2 - \sigma_X^2$, we can find $D_1^* \le D_1$, $D_2^* \le D_2$ such that $(R_1, R_2, D_1^*, D_2^*, D_3) \in \mathcal{Q}$ and $D_3 = D_1^* + D_2^* - \sigma_X^2$. Similarly, for any $(R_1, R_2, D_1, D_2, D_3) \in \mathcal{Q}$ with $D_3 > \left(1/D_1 + 1/D_2 - 1/\sigma_X^2\right)^{-1}$, we can find $D_3^* = \left(1/D_1 + 1/D_2 - 1/\sigma_X^2\right)^{-1} < D_3$ such that $(R_1, R_2, D_1, D_2, D_3^*) \in \mathcal{Q}$.

An alternative way is to write the rate-distortion region in the $D-R$ form instead of the $R-D$ form. That is, $(R_1, R_2, D_1, D_2, D_3) \in \mathcal{Q}$ if and only if

$$\begin{aligned} D_i &\ge \sigma_X^2 2^{-2R_i}, \quad i = 1,2, \\ D_3 &\ge \frac{\sigma_X^2 2^{-(R_1+R_2)}}{1 - |\sqrt{\Pi} - \sqrt{\Delta}|^{+2}} \end{aligned}$$

where $\Pi = (1 - D_1/\sigma_X^2)(1 - D_2/\sigma_X^2)$, $\Delta = D_1D_2/\sigma_X^4 - 2^{-2(R_1+R_2)}$, and $|x|^+ = \max(x, 0)$. Again, the case $\sqrt{\Pi} - \sqrt{\Delta} < 0$ (i.e., $\sigma_X^2 + \sigma_X^2 2^{-2(R_1+R_2)} < D_1 + D_2$) is degenerated since we can find $D_1^* \le D_1$ and $D_2^* \le D_2$ such that $\sigma_X^2 + \sigma_X^2 2^{-2(R_1+R_2)} = D_1^* + D_2^*$ and $D_i^* \ge \sigma_X^2 2^{-2R_i}$, $i = 1,2$.

Henceforth we shall only consider the case $\left(1/D_1 + 1/D_2 - 1/\sigma_X^2\right)^{-1} \ge D_3 \ge D_1 + D_2 - \sigma_X^2$. Only in this subregion, D_1, D_2 and D_3 are all effective. Let $U = X + T_0 + T_1$, $W = X + T_0 + T_2$, where (T_1, T_2), T_0, X are zero-mean, jointly Gaussian and independent, and the covariance matrix of (T_0, T_1, T_2) is

$$C_T = \begin{pmatrix} \sigma_{T_0}^2 & 0 & 0 \\ 0 & \sigma_{T_1}^2 & -\sigma_{T_1}\sigma_{T_2} \\ 0 & -\sigma_{T_1}\sigma_{T_2} & \sigma_{T_2}^2 \end{pmatrix}.$$

Let

$$\begin{aligned} \hat{X}_1 &= \mathbb{E}(X|U) = \frac{\sigma_X^2}{\sigma_X^2 + \sigma_{T_0}^2 + \sigma_{T_1}^2} U, \\ \hat{X}_2 &= \mathbb{E}(X|W) = \frac{\sigma_X^2}{\sigma_X^2 + \sigma_{T_0}^2 + \sigma_{T_2}^2} W, \\ \hat{X}_3 &= \mathbb{E}(X|\hat{X}_1, \hat{X}_2) = \alpha \hat{X}_1 + \beta \hat{X}_2, \end{aligned}$$

where

$$\begin{aligned} \alpha &= \frac{\sigma_{T_2}(\sigma_X^2 + \sigma_{T_0}^2 + \sigma_{T_1}^2)}{(\sigma_{T_1} + \sigma_{T_2})(\sigma_X^2 + \sigma_{T_0}^2)}, \\ \beta &= \frac{\sigma_{T_1}(\sigma_X^2 + \sigma_{T_0}^2 + \sigma_{T_2}^2)}{(\sigma_{T_1} + \sigma_{T_2})(\sigma_X^2 + \sigma_{T_0}^2)}. \end{aligned}$$

By setting $\mathbb{E}(X - \hat{X}_i)^2 = D_i$, $i = 1,2,3$, we get

$$\begin{aligned} \sigma_{T_0}^2 &= \frac{D_3 \sigma_X^2}{\sigma_X^2 - D_3}, \\ \sigma_{T_i}^2 &= \frac{D_i \sigma_X^2}{\sigma_X^2 - D_i} - \frac{D_3 \sigma_X^2}{\sigma_X^2 - D_3}, \quad i = 1,2. \end{aligned}$$

Now it is easy to verify that

$$I(X; \hat{X}_i) = \frac{1}{2} \log \frac{\sigma_X^2}{D_i}, \quad i = 1,2,$$

and

$$\begin{aligned} & I(X; \hat{X}_1, \hat{X}_2) + I(\hat{X}_1; \hat{X}_2) \\ = & \frac{1}{2} \log \frac{\sigma_X^2}{D_3} + \frac{1}{2} \log \psi(D_1, D_2, D_3). \end{aligned}$$

Hence for the quadratic Gaussian case, the SEGC region is the same as the EGC region and there is no need to introduce $\hat{X}_3$ (more precisely, $\hat{X}_3$ can be represented as a deterministic function of $\hat{X}_1$ and $\hat{X}_2$).

Now we proceed to study the coding scheme for V_1 of the SEGC region. Firstly, we can compute that

$$\begin{aligned} & \mathbb{E}(\hat{X}_2|X, \hat{X}_1) = a_1 X + a_2 \hat{X}_1, \\ & \mathbb{E}(\mathbb{E}(\hat{X}_2|X, \hat{X}_1)|\hat{X}_2) = a_3 \hat{X}_2, \\ & \mathrm{var}(\mathbb{E}(\hat{X}_2|X, \hat{X}_1)|\hat{X}_2) \\ = & \frac{\sigma_X^4}{(\sigma_X^2 + \sigma_{T_0}^2 + \sigma_{T_2}^2)^2} \left[\sigma_X^2 + \frac{(\sigma_{T_0}^2 - \sigma_{T_1}\sigma_{T_2})^2}{\sigma_{T_0}^2 + \sigma_{T_1}^2} \right] \\ & - \frac{\sigma_X^4 \left[\sigma_X^2 \sigma_{T_1}^2 + \sigma_X^2 \sigma_{T_0}^2 + (\sigma_{T_0}^2 - \sigma_{T_1}\sigma_{T_2})^2 \right]^2}{(\sigma_{T_0}^2 + \sigma_{T_1}^2)^2 (\sigma_X^2 + \sigma_{T_0}^2 + \sigma_{T_2}^2)^3}, \end{aligned}$$

where

$$\begin{aligned} a_1 &= \frac{\sigma_X^2(\sigma_{T_1}^2 + \sigma_{T_1}\sigma_{T_2})}{(\sigma_X^2 + \sigma_{T_0}^2 + \sigma_{T_2}^2)(\sigma_{T_0}^2 + \sigma_{T_1}^2)}, \\ a_2 &= \frac{(\sigma_{T_0}^2 - \sigma_{T_1}\sigma_{T_2})(\sigma_X^2 + \sigma_{T_0}^2 + \sigma_{T_1}^2)}{(\sigma_X^2 + \sigma_{T_0}^2 + \sigma_{T_2}^2)(\sigma_{T_0}^2 + \sigma_{T_1}^2)}, \\ a_3 &= \frac{\sigma_X^2 \sigma_{T_1}^2 + \sigma_X^2 \sigma_{T_0}^2 + \sigma_{T_0}^4 - 2\sigma_{T_0}^2 \sigma_{T_1}\sigma_{T_2} + \sigma_{T_1}^2 \sigma_{T_2}^2}{(\sigma_{T_0}^2 + \sigma_{T_1}^2)(\sigma_X^2 + \sigma_{T_0}^2 + \sigma_{T_2}^2)}. \end{aligned}$$

For Vertex 1, we have

$$\begin{aligned} R_1 &= I(X; \hat{X}_1), \\ R_2 &= I(X, \hat{X}_1; \hat{X}_2) \\ &= I(\mathbb{E}(\hat{X}_2|X, \hat{X}_1); \hat{X}_2) \\ &= I(\mathbb{E}(\hat{X}_2|X, \hat{X}_1); \mathbb{E}(\mathbb{E}(\hat{X}_2|X, \hat{X}_1)|\hat{X}_2)), \end{aligned}$$

where the third equality follows from the fact that $(X, \hat{X}_1) \to \mathbb{E}(\hat{X}_2|X, \hat{X}_1) \to \hat{X}_2$ form a Markov chain. Actually the above equations imply[3] that encoder 2, instead of generating a codebook that covers the $(X, \hat{X}_1)$-space, just needs a codebook that covers the $\mathbb{E}(\hat{X}_2|X, \hat{X}_1)$-space, which is considerably simpler. The system diagram of the above coding scheme is shown in Fig. 2.

Now we shall study an arbitrary rate pair on the dominant face of the SEGC region. Here we adopt the successive encoding scheme based on the quantization splitting method.

Let $V = W + T_3$ with $T_3 \sim \mathcal{N}(0, \sigma_{T_3}^2)$ independent of everything else. Let

$$\begin{aligned} \hat{X}_2' &= \mathbb{E}(X|V) = \frac{\sigma_X^2}{\sigma_X^2 + \sigma_{T_0}^2 + \sigma_{T_2}^2 + \sigma_{T_3}^2} V, \\ \Delta X &= X - \mathbb{E}(X|\hat{X}_2') = X - \hat{X}_2', \\ \Delta \hat{X}_1 &= \hat{X}_1 - \mathbb{E}(\hat{X}_1|\hat{X}_2') = \hat{X}_1 - b_1 \hat{X}_2', \\ \Delta \hat{X}_2 &= \hat{X}_2 - \mathbb{E}(\hat{X}_2|\hat{X}_2') = \hat{X}_2 - \hat{X}_2'. \end{aligned}$$

[3]For a more rigorous justification, one can invoke the Markov lemma [19].

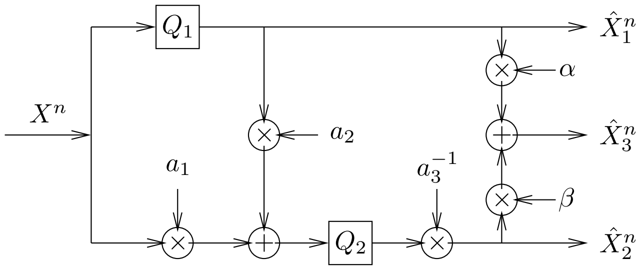

Figure 2: Quantization scheme for V_1.

Note: $\Delta X, \Delta\hat{X}_1$ and $\Delta\hat{X}_2$ are independent of $\hat{X}_2'$.

We can compute that

$$\begin{aligned}
\mathbb{E}(\hat{X}_1|X,\hat{X}_2') &= b_2X + b_3\hat{X}_2',\\
\mathbb{E}(\mathbb{E}(\hat{X}_1|X,\hat{X}_2')|\hat{X}_1) &= b_4\hat{X}_1,\\
\mathbb{E}(\Delta\hat{X}_2|\Delta X,\Delta\hat{X}_1) &= b_5\Delta X + b_6\Delta\hat{X}_1,\\
\mathbb{E}(\mathbb{E}(\Delta\hat{X}_2|\Delta X,\Delta\hat{X}_1)|\Delta\hat{X}_2) &= b_7\Delta\hat{X}_2,
\end{aligned}$$

and

$$\mathrm{var}(X|\hat{X}_2') = \mathbb{E}(\Delta X)^2 = \frac{\sigma_X^2(\sigma_{T_0}^2+\sigma_{T_2}^2+\sigma_{T_3}^2)}{\sigma_X^2+\sigma_{T_0}^2+\sigma_{T_2}^2+\sigma_{T_3}^2},$$

$$\begin{aligned}
&\mathrm{var}(\mathbb{E}(\hat{X}_1|X,\hat{X}_2')|\hat{X}_1)\\
=\;& \frac{\sigma_X^4}{(\sigma_X^2+\sigma_{T_0}^2+\sigma_{T_1}^2)^2}\left[\sigma_X^2 + \frac{(\sigma_{T_0}^2-\sigma_{T_1}\sigma_{T_2})^2}{\sigma_{T_0}^2+\sigma_{T_2}^2+\sigma_{T_3}^2}\right]\\
&-\frac{\sigma_X^4\left[\sigma_X^2(\sigma_{T_0}^2+\sigma_{T_2}^2+\sigma_{T_3}^2)+(\sigma_{T_0}^2-\sigma_{T_1}\sigma_{T_2})^2\right]^2}{(\sigma_{T_0}^2+\sigma_{T_2}^2+\sigma_{T_3}^2)^2(\sigma_X^2+\sigma_{T_0}^2+\sigma_{T_1}^2)^3},
\end{aligned}$$

$$\begin{aligned}
&\mathrm{var}(\mathbb{E}(\Delta\hat{X}_2|\Delta X,\Delta\hat{X}_1)|\Delta\hat{X}_2)\\
=\;& \frac{1}{\left(\mathbb{E}(\Delta X)^2\mathbb{E}(\Delta\hat{X}_1)^2-(\mathbb{E}(\Delta X\Delta\hat{X}_1))^2\right)^2}\\
&\times\Big[\mathbb{E}(\Delta X)^2\left(\mathbb{E}(\Delta X\Delta\hat{X}_2)\mathbb{E}(\Delta\hat{X}_1)^2\right.\\
&\left.-\mathbb{E}(\Delta\hat{X}_1\Delta\hat{X}_2)\mathbb{E}(\Delta X\Delta\hat{X}_1)\right)^2\\
&+\mathbb{E}(\Delta\hat{X}_1)^2\left(\mathbb{E}(\Delta\hat{X}_1\Delta\hat{X}_2)\mathbb{E}(\Delta X)^2\right.\\
&\left.-\mathbb{E}(\Delta X\Delta\hat{X}_1)\mathbb{E}(\Delta X\Delta\hat{X}_2)\right)^2\\
&+2\mathbb{E}(\Delta X\Delta\hat{X}_1)\left(\mathbb{E}(\Delta X\Delta\hat{X}_2)\mathbb{E}(\Delta\hat{X}_1)^2\right.\\
&\left.-\mathbb{E}(\Delta\hat{X}_1\Delta\hat{X}_2)\mathbb{E}(\Delta X\Delta\hat{X}_1)\right)\\
&\times\left(\mathbb{E}(\Delta\hat{X}_1\Delta\hat{X}_2)\mathbb{E}(\Delta X)^2\right.\\
&\left.-\mathbb{E}(\Delta X\Delta\hat{X}_1)\mathbb{E}(\Delta X\Delta\hat{X}_2)\right)\\
&-\frac{1}{\left(\mathbb{E}(\Delta X)^2\mathbb{E}(\Delta\hat{X}_1)^2-(\mathbb{E}(\Delta X\Delta\hat{X}_1))^2\right)^2\mathbb{E}(\Delta\hat{X}_2)^2}\\
&\times\left((\mathbb{E}(\Delta X\Delta\hat{X}_2))^2\mathbb{E}(\Delta\hat{X}_1)^2+(\mathbb{E}(\Delta\hat{X}_1\Delta\hat{X}_2))^2\mathbb{E}(\Delta X)^2\right.\\
&\left.-2\mathbb{E}(\Delta X\Delta\hat{X}_1)\mathbb{E}(\Delta X\Delta\hat{X}_2)\mathbb{E}(\Delta\hat{X}_1\Delta\hat{X}_2)\right)^2,
\end{aligned}$$

where

$$\begin{aligned}
b_1 &= \hat{X}_1 - \frac{\sigma_X^2+\sigma_{T_0}^2-\sigma_{T_1}\sigma_{T_2}}{\sigma_X^2+\sigma_{T_0}^2+\sigma_{T_1}^2},\\
b_2 &= \frac{\sigma_X^2(\sigma_{T_2}^2+\sigma_{T_3}^2+\sigma_{T_1}\sigma_{T_2})}{(\sigma_X^2+\sigma_{T_0}^2+\sigma_{T_1}^2)(\sigma_{T_0}^2+\sigma_{T_2}^2+\sigma_{T_3}^2)},\\
b_3 &= \frac{(\sigma_{T_0}^2-\sigma_{T_1}\sigma_{T_2})(\sigma_X^2+\sigma_{T_0}^2+\sigma_{T_2}^2+\sigma_{T_3}^2)}{(\sigma_X^2+\sigma_{T_0}^2+\sigma_{T_1}^2)(\sigma_{T_0}^2+\sigma_{T_2}^2+\sigma_{T_3}^2)},\\
b_4 &= \frac{\sigma_X^2(\sigma_{T_0}^2+\sigma_{T_2}^2+\sigma_{T_3}^2)+(\sigma_{T_0}^2-\sigma_{T_1}\sigma_{T_2})^2}{(\sigma_X^2+\sigma_{T_0}^2+\sigma_{T_1}^2)(\sigma_{T_0}^2+\sigma_{T_2}^2+\sigma_{T_3}^2)},\\
b_5 &= \frac{\mathbb{E}(\Delta X\Delta\hat{X}_2)\mathbb{E}(\Delta\hat{X}_1)^2-\mathbb{E}(\Delta\hat{X}_1\Delta\hat{X}_2)\mathbb{E}(\Delta X\Delta\hat{X}_1)}{\mathbb{E}(\Delta X)^2\mathbb{E}(\Delta\hat{X}_1)^2-(\mathbb{E}(\Delta X\Delta\hat{X}_1))^2},\\
b_6 &= \frac{\mathbb{E}(\Delta\hat{X}_1\Delta\hat{X}_2)\mathbb{E}(\Delta X)^2-\mathbb{E}(\Delta X\Delta\hat{X}_1)\mathbb{E}(\Delta X\Delta\hat{X}_2)}{\mathbb{E}(\Delta X)^2\mathbb{E}(\Delta\hat{X}_1)^2-(\mathbb{E}(\Delta X\Delta\hat{X}_1))^2},\\
b_7 &= \frac{1}{\mathbb{E}(\Delta X)^2\mathbb{E}(\Delta\hat{X}_1)^2\mathbb{E}(\Delta\hat{X}_2)^2-(\mathbb{E}(\Delta X\Delta\hat{X}_1))^2\mathbb{E}(\Delta\hat{X}_2)^2}\\
&\quad\times\left((\mathbb{E}(\Delta X\Delta\hat{X}_2))^2\mathbb{E}(\Delta\hat{X}_1)^2+(\mathbb{E}(\Delta\hat{X}_1\Delta\hat{X}_2))^2\mathbb{E}(\Delta X)^2\right.\\
&\quad\left.-2\mathbb{E}(\Delta X\Delta\hat{X}_1)\mathbb{E}(\Delta X\Delta\hat{X}_2)\mathbb{E}(\Delta\hat{X}_1\Delta\hat{X}_2)\right),
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{E}(\Delta X)^2 &= \frac{\sigma_X^2(\sigma_{T_0}^2+\sigma_{T_2}^2+\sigma_{T_3}^2)}{\sigma_X^2+\sigma_{T_0}^2+\sigma_{T_2}^2+\sigma_{T_3}^2},\\
\mathbb{E}(\Delta\hat{X}_1)^2 &= \frac{\sigma_X^4}{\sigma_X^2+\sigma_{T_0}^2+\sigma_{T_1}^2}\\
&\quad- \frac{\sigma_X^4(\sigma_X^2+\sigma_{T_0}^2-\sigma_{T_1}\sigma_{T_2})^2}{(\sigma_X^2+\sigma_{T_0}^2+\sigma_{T_1}^2)^2(\sigma_X^2+\sigma_{T_0}^2+\sigma_{T_2}^2+\sigma_{T_3}^2)},\\
\mathbb{E}(\Delta\hat{X}_2)^2 &= \frac{\sigma_X^4}{\sigma_X^2+\sigma_{T_0}^2+\sigma_{T_2}^2}-\frac{\sigma_X^4}{\sigma_X^2+\sigma_{T_0}^2+\sigma_{T_2}^2+\sigma_{T_3}^2},\\
\mathbb{E}(\Delta X\Delta\hat{X}_1) &= \frac{\sigma_X^4}{\sigma_X^2+\sigma_{T_0}^2+\sigma_{T_1}^2}\\
&\quad- \frac{\sigma_X^4(\sigma_X^2+\sigma_{T_0}^2-\sigma_{T_1}\sigma_{T_2})}{(\sigma_X^2+\sigma_{T_0}^2+\sigma_{T_1}^2)(\sigma_X^2+\sigma_{T_0}^2+\sigma_{T_2}^2+\sigma_{T_3}^2)},\\
\mathbb{E}(\Delta X\Delta\hat{X}_2) &= \frac{\sigma_X^4}{\sigma_X^2+\sigma_{T_0}^2+\sigma_{T_2}^2}-\frac{\sigma_X^4}{\sigma_X^2+\sigma_{T_0}^2+\sigma_{T_2}^2+\sigma_{T_3}^2},\\
\mathbb{E}(\Delta\hat{X}_1\Delta\hat{X}_2) &= \frac{\sigma_X^4(\sigma_X^2+\sigma_{T_0}^2-\sigma_{T_1}\sigma_{T_2})}{(\sigma_X^2+\sigma_{T_0}^2+\sigma_{T_1}^2)(\sigma_X^2+\sigma_{T_0}^2+\sigma_{T_2}^2)}\\
&\quad- \frac{\sigma_X^4(\sigma_X^2+\sigma_{T_0}^2-\sigma_{T_1}\sigma_{T_2})}{(\sigma_X^2+\sigma_{T_0}^2+\sigma_{T_1}^2)(\sigma_X^2+\sigma_{T_0}^2+\sigma_{T_2}^2+\sigma_{T_3}^2)}.
\end{aligned}$$

By Lemma 1, we have

$$\begin{aligned}
R_1 &= I(X,\hat{X}_2';\hat{X}_1)\\
&= I(\mathbb{E}(\hat{X}_1|X,\hat{X}_2');\hat{X}_1)\\
&= I(\mathbb{E}(\hat{X}_1|X,\hat{X}_2');\mathbb{E}(\mathbb{E}(\hat{X}_1|X,\hat{X}_2')|\hat{X}_1)),\\
R_2 &= I(X;\hat{X}_2') + I(X,\hat{X}_1;\hat{X}_2|\hat{X}_2')\\
&= I(X;\hat{X}_2') + I(\Delta X,\Delta\hat{X}_1;\Delta\hat{X}_2|\hat{X}_2')\\
&= I(X;\hat{X}_2') + I(\Delta X,\Delta\hat{X}_1;\Delta\hat{X}_2)\\
&= I(X;\hat{X}_2') + I(\mathbb{E}(\Delta\hat{X}_2|\Delta X,\Delta\hat{X}_1);\Delta\hat{X}_2)\\
&= I(\mathbb{E}(\Delta\hat{X}_2|\Delta X,\Delta\hat{X}_1);\mathbb{E}(\mathbb{E}(\Delta\hat{X}_2|\Delta X,\Delta\hat{X}_1)|\Delta\hat{X}_2))\\
&\quad+I(X;\hat{X}_2').
\end{aligned}$$

We have used the fact that $(X,\hat{X}_2')\to\mathbb{E}(\hat{X}_1|X,\hat{X}_2')\to\hat{X}_1$ and $(\Delta X,\Delta\hat{X}_1)\to\mathbb{E}(\Delta\hat{X}_2|\Delta X,\Delta\hat{X}_1)\to\Delta\hat{X}_2$.

The system diagram of the above scheme is shown in Fig. 3.

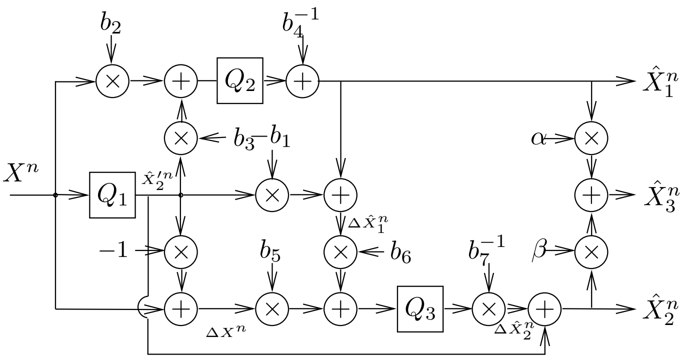

Figure 3: Quantization scheme for an arbitrary rate pair.

IV. Conclusion

Although we have adopted an information theoretic approach, all of our results can be interpreted from the perspective of lattice quantization [20–23]. A more systematic approach based on the connection between the Gram-Schmidt orthogonalization and sequential (dithered) quantization can be found in [24].

References

[1] A. A. El Gamal and T. M. Cover, "Achievable rates for multiple descriptions," *IEEE Trans. on Inform. Theory*, vol.IT-28, pp. 851-857, Nov. 1982.

[2] 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.

[3] R. Ahlswede, "The rate-distortion region for multiple descriptions without excess rate," *IEEE Trans. on Inform. Theory*, vol. IT-31, pp. 721-726, Nov. 1985.

[4] Z. Zhang and T. Berger, "New results in binary multiple descriptions," *IEEE Trans. on Inform. Theory*, vol. IT-33, pp. 502-521, July 1987.

[5] S. S. Pradhan and K. Ramchandran, "On the optimality of block orthogonal transforms for multiple description coding of Gaussian vector sources," *IEEE Signal Processing Letter*, vol. 7, no. 4, pp. 76-78, April 2000.

[6] V. K Goyal and J. Kovacevic, "Generalized multiple description coding with correlating transforms," *IEEE Trans. on Inform. Theory*, vol. 47, no. 6, pp. 2199-2224, Sept. 2001.

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

[8] V. A. Vaishampayan, "Design of multiple description scalar quantizers," *IEEE Trans. Inform. Theory*, vol. 39, pp. 821-834, May 1993.

[9] V. A. Vaishampayan, "Design of entropy-constrained multiple-description scalar quantizers," *IEEE Trans. Inform. Theory*, vol. 40, pp. 245-251, Jan. 1994.

[10] V. A. Vaishampayan and J. C. Batllo, "Asymptotic analysis of multiple- description quantizers," *IEEE Trans. Inform. Theory*, vol. 44, pp. 278-284, Jan. 1998.

[11] V. A. Vaishampayan, N. Sloane, and S. Servetto, "Multiple description vector quantization with lattice codebooks: Design and analysis," *IEEE Trans. Inform. Theory*, vol. 47, pp. 1718-1734, July 2001.

[12] Suhas N. Diggavi, N. J. A. Sloane and Vinay A. Vaishampayan, "Asymmetric multiple description lattice vector quantizers", *IEEE Trans. Inform. Theory*, vol. 48, pp. 174-191, Jan. 2002.

[13] Y. Frank-Dayan and R. Zamir, "Dithered lattice-based quantizers for multiple descriptions," *IEEE Trans. Inform. Theory*, vol. 48, NO. 1, pp. 192-204, Jan. 2002.

[14] C. Tian and S. S. Hemami, "Universal multiple description scalar quantizer: analysis and design," *IEEE Trans. Information Theory*, vol. 50, no. 9, pp. 2089–2102, Sep. 2004.

[15] C. Tian and S. S. Hemami, "Optimality and sub-optimality of multiple description vector quantization with a lattice codebook," *IEEE Trans. Information Theory*, vol. 50, no. 10, pp. 2458–2468, Oct. 2004.

[16] R. Zamir, "Gaussian codes and Shannon bounds for multiple descriptions," *IEEE Trans. Inform. Theory*, vol. 45, pp. 2629-2635, Nov. 1999.

[17] B. Rimoldi and R. Urbanke, "Asynchronous Slepian-Wolf coding via source-splitting," in *IEEE International Symposium on Information Theory*, Ulm, Germany, June 29-July 4 1997, p. 271.

[18] J. Chen and T. Berger, "Successive Wyner-Ziv coding scheme and its application to the quadratic gaussian CEO problem," *IEEE Trans. on Inform Theory*, submitted for publication.

[19] T. Berger, "Multiterminal source coding," in *The Information Theory Approach to Communications* (G. Longo, ed.), vol. 229 of *CISM Courses and Lectures*, pp. 171-231, Springer-Verlag, Vienna/New York, 1978.

[20] R. Zamir and M. Feder, "On universal quantization by randomized uniform/lattice quantizer," *IEEE Trans. Inform. Theory*, vol. 38, pp. 428-436, Mar. 1992.

[21] R. Zamir and M. Feder, "On lattice quantization noise," *IEEE Trans. Inform. Theory*, vol. 42, pp. 1152-1159, July 1996.

[22] R. Zamir and M. Feder, "Information rates of pre/post filtered dithered quantizers," *IEEE Trans. Inform. Theory*, vol. 42, pp. 1340-1353, Sept. 1996.

[23] R. Zamir, S. Shamai and U. Erez, "Nested linear/lattice codes for structured multiterminal binning," *IEEE Trans. Info. Theory*, vol. 48, pp. 1250-1276, June 2002.

[24] J. Chen, C. Tian, T. Berger and S. S. Hemami, "Multiple description quantization via Gram-Schmidt orthogonalization," available on line: http://www.people.cornell.edu/pages/jc353/.