Achieving the Multiple Description Rate-Distortion Region with Lattice Quantization

Jun Chen, Chao Tian, Toby Berger¹, 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)\}\textsuperscript{∞}_{t=1} be an i.i.d. random process with \(X(t) \sim P(x)\) for all \(t\). Let \(d(\cdot, \cdot) : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}^+\) 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:

\[
f_i^{(n)} : \mathcal{X}^n \rightarrow \mathcal{C}_i^{(n)} \log |\mathcal{C}_i^{(n)}| \leq n(R_i + \varepsilon) \quad i = 1, 2,
\]
and decoding functions:

\[
g_i^{(n)} : \mathcal{C}_i^{(n)} \rightarrow \mathcal{X}^n \quad i = 1, 2
\]

\[
g_3^{(n)} : \mathcal{C}_3^{(n)} \times \mathcal{C}_2^{(n)} \rightarrow \mathcal{X}^n
\]

such that for \(\hat{X}_1^n = g_1^{(n)}(f_1^{(n)}(X^n)), i = 1, 2,\) and for \(\hat{X}_2^n = g_3^{(n)}(f_1^{(n)}(X^n), f_2^{(n)}(X^n))\), we have

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

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 \(X_1, \hat{X}_2, \hat{X}_3\) jointly distributed with the generic source variable \(X\) such that

1. \(R_i \geq I(X; \hat{X}_i), i = 1, 2\);
2. \(R_1 + R_2 \geq I(X; \hat{X}_1, \hat{X}_2, \hat{X}_3) + I(\hat{X}_1; \hat{X}_2)\);
3. \(Ed(X, \hat{X}_i) \leq D_i, i = 1, 2, 3\).

We denote the EGC region by \(\mathcal{Q}_{\text{EGC}}\).

Ozarow [2] showed that \(\mathcal{Q}_{\text{EGC}} = \mathcal{Q}\) 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 \(\mathcal{Q}_{\text{EGC}} \subsetneq \mathcal{Q}\).

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\(^\dagger\):

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

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

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

\[
\begin{align*}
I(X; \hat{X}_1) + I(X; \hat{X}_2) &= I(X; \hat{X}_1, \hat{X}_2) + I(\hat{X}_1; \hat{X}_2) - I(\hat{X}_1; \hat{X}_2|X) \\
&\leq I(X, \hat{X}_1, \hat{X}_2) + I(\hat{X}_1; \hat{X}_2),
\end{align*}
\]

the sum-rate constraint is always effective. We call

\[
\{(R_1, R_2) : R_1 + R_2 = I(X; \hat{X}_1, \hat{X}_2) + I(\hat{X}_1; \hat{X}_2), R_i \geq I(X; \hat{X}_i), i = 1, 2\}
\]

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

\[^2\]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 \( 2^{nR_1(X; X_1)} \) codewords \( \{\hat{X}^1_i(j)\}_{j=1}^{2^n(X; X_1)} \) according to the marginal distribution \( P(\hat{x}_1). \) Encoder 2 independently generates \( 2^{nR_2(X_1; X_2)} \) codewords \( \{\hat{X}^2_i(k)\}_{k=1}^{2^n(X_1; X_2)} \) according to the marginal distribution \( P(\hat{x}_2). \)

   **Encoding Procedure:** Given \( X^n \), encoder 1 finds the codeword \( \hat{X}^1_i(j^*) \) such that \( X^n \) and \( \hat{X}^1_i(j^*) \) are jointly typical. Then encoder 2 finds the codeword \( \hat{X}^2_i(k^*) \) such that \( \hat{X}^2_i(k^*) \) is jointly typical with \( X^n \) and \( \hat{X}^1_i(j^*) \).

2. **Two-Stage Encoding Scheme:** We write \( R_2 = I(X; X_2) + I(\hat{X}_1; \hat{X}_2|X). \)

   Encoder 1 independently generates \( 2^{nR_1(X; X_1)} \) codewords \( \{\hat{X}^1_i(j)\}_{j=1}^{2^n(X; X_1)} \) according to the marginal distribution \( P(\hat{x}_1). \) Encoder 2 independently generates \( 2^{nR_2(X_1; X_2|X)} \) codewords, with each codeword containing \( 2^{nR_2(X_1; X_2)} \) codewords. That is, codebook \( i \) has codewords \( \{\hat{X}^2_i(k)\}_{k=1}^{2^n(X_1; X_2)} \), \( i = 1, 2, \ldots, 2^{nR_1(X; X_1)}. \)

   **Encoding Procedure:**

   (a) Given \( X^n \), encoder 1 finds the codeword \( \hat{X}^1_i(j^*) \) such that \( X^n \) and \( \hat{X}^1_i(j^*) \) are jointly typical. Encoder 2 finds, in each codebook \( i \), a codeword \( \hat{X}^2_i(k^*_i) \) such that \( \hat{X}^2_i(k^*_i) \) is jointly typical with \( X^n, i = 1, 2, \ldots, 2^{nR_1(X; X_1)}. \) Encoder 2 then forms a new codebook with these \( 2^{nR_1(X; X_2)} \) selected codewords.

   (b) In this newly-formed codebook, encoder 2 finds a codeword \( \hat{X}^2_i(j^*_i, k^*_i) \) such that \( \hat{X}^2_i(j^*_i, k^*_i) \) is jointly typical with \( X^n \) and \( \hat{X}^1_i(j^*_i) \).

   Index \( j^*_i \) 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 \( 2^{nR_2(X; X_2)} \) 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 \( \{\hat{X}^1_i(j)\}_{j=1}^{2^n(X; X_1)} \) according to the marginal distribution \( P(\hat{x}_1). \) Encoder 2 generates \( 2^{nR_2} \) codewords \( \{\hat{X}^2_i(k)\}_{k=1}^{2^n(X_1; X_2)} \) according to the marginal distribution \( P(\hat{x}_2). \)

   **Encoding Procedure:** Given \( X^n \), encoder 1 and encoder 2 find the codewords \( \hat{X}^1_i(j^*) \) and \( \hat{X}^2_i(k^*) \) such that \( X^n, \hat{X}^1_i(j^*) \) and \( \hat{X}^2_i(k^*) \) 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 \( 2^{nR_1-I(X; X_1)} \) codebooks, with each codebook containing \( 2^{nR_1(X; X_1)} \) codewords, i.e., codebook \( i \) contains codewords \( \{\hat{X}^1_i(i, j)\}_{j=1}^{2^n(X; X_1)}, i = 1, 2, \ldots, 2^{nR_1-I(X; X_1)}. \) Similarly, encoder 2 independently generates \( 2^{nR_2-I(X; X_2)} \) codebooks, with each codebook containing \( 2^{nR_2(X; X_2)} \) codewords. That is, codebook \( k \) has codewords \( \{\hat{X}^2_i(k, l)\}_{l=1}^{2^n(X; X_2)}, k = 1, 2, \ldots, 2^{nR_2-I(X; X_2)}. \)

   **Encoding Procedure:**

   (a) Given \( X^n \), encoder 1 finds, in each codebook \( i \), a codeword \( \hat{X}^1_i(i, j^*_i) \) such that \( \hat{X}^1_i(i, j^*_i) \) is jointly typical with \( X^n, i = 1, 2, \ldots, 2^{nR_1-I(X; X_1)}. \) Encoder 1 then forms a new codebook with these \( 2^{nR_1-I(X; X_1)} \) selected codewords. Similarly encoder 2 finds, in each codebook \( k \), a codeword \( \hat{X}^2_i(k, l^*_i) \) such that \( \hat{X}^2_i(k, l^*_i) \) is jointly typ-
ical with $X^n, i = 1, 2, \ldots, 2^{n[R_2 - I(X; X_2)]}$ En-
coder 2 then forms a new codebook with these
$2^{n[R_2 - I(X; X_2)]}$ selected codewords.

(b) In these two newly-formed codebooks, encoder
1 and encoder 2 find codewords $X^n_1(\hat{t}^*, \hat{j}_i^*)$ and
$X^n_2(\hat{k}^*, \hat{l}_k^*)$ such that $X^n_1(\hat{t}^*, \hat{j}_i^*)$ and
$X^n_2(\hat{k}^*, \hat{l}_k^*)$ 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 $2^n[1 + R_2 - I(X; X_1) - I(X; X_2)] = 2^nI(X_1; X_2 | X)$
while the size of the search space for the joint encoding
scheme is $2^n[R_1 + R_2]$. 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 $X^n_2$ with $(X, X_1) 
\rightarrow X^n_2 \rightarrow X^n_1$ such that

$$R_1 = I(X, X_2^1; \hat{X}_1),$$
$$R_2 = I(X, X_2^2) + I(X, \hat{X}_1; X_2 | X_2),$$

Symmetrically, we can find $X^n_1$ with $(X, \hat{X}_2) 
\rightarrow \hat{X}_1 \rightarrow X^n_1$ such that

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

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

If we let $X^n_2 = \text{constant}$, then

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

which corresponds to Vertex 1.

If we let $X^n_2 = X_2$, 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 prob-
abilities $P_{\epsilon}(X^n_2 | X^n_2)$ indexed by $\epsilon$ such that
$I(X; X^n_2)$ changes continuously from 0 to $I(X; \hat{X}_2)$ as $\epsilon$ 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

$$R_1 + R_2 = I(X, X_2^1; \hat{X}_1) + I(X, \hat{X}_2^1) + I(X, \hat{X}_1; X_2 | X_2^1)^2$$
$$= I(X, X_2^1; \hat{X}_1) + I(X, \hat{X}_2^1) + I(X, \hat{X}_2; X_2 | X_2^1)$$
$$+ I(X; \hat{X}_1; \hat{X}_2 | X_2)$$
$$= I(X, \hat{X}_1; \hat{X}_2^1; \hat{X}_2; \hat{X}_2; X_2).$$

Since $(X, \hat{X}_1) \rightarrow \hat{X}_2 \rightarrow X^n_2$, we have

$$I(X, X_2^1; \hat{X}_1) + I(X; X_2^1, \hat{X}_2)$$
$$= I(X, X_2^1; \hat{X}_1) + I(X; \hat{X}_2)$$
$$= I(X, \hat{X}_1) + I(X, \hat{X}_2) + I(\hat{X}_1; \hat{X}_2 | X)$$
$$= I(X; \hat{X}_1, \hat{X}_2) + I(\hat{X}_1; \hat{X}_2),$$

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.,

$$R_1 = I(X, \hat{X}_1),$$
$$R_2 = I(X, X_2^2) + I(X, \hat{X}_1; X_2 | X_2).$$

Encoder 1 independently generates $2^{n[R_1 + R_2]}$
codewords $\{\hat{X}_1(i)\}_{i=0}^{2^n}$ according to the marginal
distribution $P(\hat{X}_1(i))$. Encoder 2 independently generates
$2^{n[R_1 + R_2]}$ codewords $\{X_2^2(j)\}_{j=0}^{2^n}$
according to the marginal distribution $P(X_2^2(j))$.

Encoding Procedure: Given $X^n$, encoder 2 finds the
codeword $X_2^2(j^*)$ such that $X^n$ and $X_2^2(j^*)$ are jointly

typical. Then encoder 1 finds the codeword $\hat{X}_1(i^*)$ such
that $\hat{X}_1(i^*)$ is jointly typical with $X^n$ and $X_2^2(j^*)$.

Finally, encoder 2 finds the codeword $X_2^2(j^*, k^*)$ such
that $X_2^2(j^*, k^*)$ is jointly typical with $X^n$, $\hat{X}_1(i^*)$
and $X_2^2(j^*)$. 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 suc-
cessive encoding scheme for the vertices of the SEGC
region.

We can view $\hat{X}_2^1$ as a coarse description of $X$ and view
$\hat{X}_1$ 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 $\{X(t)\}_{t=1}^{n}$ 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

$$R_i \geq \frac{1}{2} \log \frac{\sigma_X^2}{D_i}, \quad i = 1, 2,$$

$$R_1 + R_2 \geq \frac{1}{2} \log \frac{\sigma_X^2}{D_1} + \frac{1}{2} \log \psi(D_1, D_2, D_3),$$

where $\psi(D_1, D_2, D_3)$ is defined by

$$\psi(D_1, D_2, D_3) = \frac{1}{2} \log \frac{1}{D_1} + \frac{1}{2} \log \frac{1}{D_2} + \frac{1}{2} \log \frac{1}{D_3}.$$
where
\[
\psi(D_1, D_2, D_3) = \begin{cases} 
1, & D_3 < D_1 + D_2 - \sigma_X^2 \\
\frac{\sigma_X^2 D_2}{\sigma_X^2 - D_3}, & D_3 > \left( \frac{1}{\sigma_X^2} + \frac{1}{\sigma_D^2} - \frac{1}{\sigma_1^2} \right)^{-1} \\
(\sigma_X^2 - D_3)^2 - \sqrt{(\sigma_X^2 - D_3)(\sigma_X^2 - D_2)(\sigma_X^2 - D_2)} > 0, \text{ o.w.} 
\end{cases}
\]

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

An alternative way is to write the rate-distortion region in the \( D - R \) form instead of the \( R - D \) form. That is, \((R_1, R_2, D_1, D_2, D_3) \in \mathbb{Q}\) if and only if
\[
\begin{aligned}
D_1 &\geq \frac{\sigma_X^2 2^{-2R_1}}{1 - \frac{1}{\sqrt{\Pi} - \sqrt{\Delta + \beta^2}}}, \quad i = 1, 2, \\
D_3 &\geq \frac{\sigma_X^2 2^{-2(\alpha R_1 + \beta R_2)}}{1 - \frac{1}{\sqrt{\Pi} - \sqrt{\Delta + \beta^2}}}
\end{aligned}
\]

where \( \Pi = (1 - D_1/\sigma_X^2)(1 - D_2/\sigma_X^2), \Delta = D_1 D_2 / \sigma_X^2 - 2^{-2(\alpha R_1 + \beta R_2)} \) and \( |x| = \max(|x|, 0) \). Again, the case \( \sqrt{\Pi} - \sqrt{\Delta} < 0 \) is degenerated since we can find \( D_1^* \leq D_1 \) and \( D_2^* \leq D_2 \) such that \( \sigma_X^2 + \frac{\sigma_X^2 2^{-2(R_1 + R_2)}}{\sigma_1^2} = D_1^* + D_2^* \) and \( D_3^* \geq \sigma_X^2 2^{-2R_1 - \beta R_2}, \quad i = 1, 2 \).

Henceforth we shall only consider the case \((1/D_1 + 1/D_2 - 1/\sigma_X^2)^{-1} \geq D_3 \geq D_1 + D_2 - \sigma_X^2 \). 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
\[
C_T = \begin{pmatrix}
\sigma_T^2 & 0 & 0 \\
0 & \sigma_T^2 & -\sigma_T \sigma_{T_2} \\
0 & -\sigma_T \sigma_{T_2} & \sigma_{T_2}^2
\end{pmatrix}.
\]

Let
\[
\begin{aligned}
X_1 &= \mathbb{E}(X | U) = \frac{\sigma_X^2}{\sigma_X^2 + \sigma_T^2 + \sigma_{T_1}^2} U, \\
X_2 &= \mathbb{E}(X | W) = \frac{\sigma_X^2}{\sigma_X^2 + \sigma_T^2 + \sigma_{T_2}^2} W, \\
X_3 &= \mathbb{E}(X | X_1, X_2) = \alpha X_1 + \beta X_2,
\end{aligned}
\]
where
\[
\begin{aligned}
\alpha &= \frac{\sigma_T (\sigma_X^2 + \sigma_T^2 + \sigma_{T_1}^2)}{(\sigma_T + \sigma_{T_1}) (\sigma_X^2 + \sigma_T^2)}, \\
\beta &= \frac{\sigma_T (\sigma_X^2 + \sigma_T^2 + \sigma_{T_2}^2)}{(\sigma_T + \sigma_{T_2}) (\sigma_X^2 + \sigma_T^2)}.
\end{aligned}
\]

By setting \( \mathbb{E}(X - \hat{X}_i)^2 = D_i, \ i = 1, 2, 3 \), we get
\[
\begin{aligned}
\sigma_{X_1}^2 &= \frac{D_3 \sigma_T^2}{\sigma_X^2 - D_3}, \\
\sigma_{X_2}^2 &= \frac{D_3 \sigma_T^2}{\sigma_X^2 - D_3},
\end{aligned}
\]

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
\[
I(X; \hat{X}_1, \hat{X}_2) + I(\hat{X}_1; \hat{X}_2)
\]
\[
= \frac{1}{2} \log \frac{\sigma_X^2}{D_3} + \frac{1}{2} \log \psi(D_1, D_2, D_3).
\]

Hence for the quadratic Gaussian case, the SEGC region is the same as the EGC region and there is no need to introduce \( X_3 \) (more precisely, \( X_3 \) can be represented as a deterministic function of \( X_1 \) and \( X_2 \)).

Now we proceed to study the coding scheme for \( V_1 \) of the SEGC region. Firstly, we can compute that
\[
\begin{aligned}
\mathbb{E}(X_2 | X, X_1) &= a_1 X + a_2 X_1, \\
\mathbb{E}(X_2 | X, X_1) &= a_3 X, \\
\text{var}(\mathbb{E}(X_2 | X, X_1)) &= \frac{\sigma_X^2}{(\sigma_X^2 + \sigma_T^2 + \sigma_{T_2}^2)} \left[ \frac{\sigma_X^2 + (\sigma_{T_2}^2 - \sigma_T^2 \sigma_{T_2}^2)^2}{\sigma_T^2 + \sigma_{T_2}^2} \right] \\
&= \frac{\sigma_X^2}{(\sigma_T + \sigma_{T_2})^2} \left[ \frac{\sigma_X^2 + (\sigma_T^2 - \sigma_T^2 \sigma_{T_2}^2)^2}{\sigma_T^2 + \sigma_{T_2}^2} \right].
\end{aligned}
\]

For Vertex 1, we have
\[
R_1 = I(X; \hat{X}_1), \\
R_2 = I(X; \hat{X}_1; \hat{X}_2), \\
I(\mathbb{E}(X_2 | X, \hat{X}_1); \hat{X}_2) &= I(\mathbb{E}(X_2 | X, \hat{X}_1); \mathbb{E}(X_2 | X, \hat{X}_1)),
\]
where the third equality follows from the fact that \( (X, \hat{X}_1) \rightarrow \mathbb{E}(X_2 | X, \hat{X}_1) \rightarrow \hat{X}_2 \) 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 \( \mathbb{E}(X_2 | X, \hat{X}_1) \)-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 \( T_3 \sim \mathcal{N}(0, \sigma_T^2) \) independent of everything else. Let
\[
\begin{aligned}
\hat{X}_2' &= \mathbb{E}(X | V) = \frac{\sigma_X^2}{\sigma_X^2 + \sigma_T^2 + \sigma_{T_2}^2} V, \\
\Delta X &= X - \mathbb{E}(X | \hat{X}_2'), \\
\Delta \hat{X}_1 &= \hat{X}_1 - \mathbb{E}(\hat{X}_1 | \hat{X}_2') = \hat{X}_1 - b_1 \Delta \hat{X}_2, \\
\Delta \hat{X}_2 &= \hat{X}_2 - \mathbb{E}(\hat{X}_2 | \hat{X}_2') = \hat{X}_2 - \Delta \hat{X}_1.
\end{aligned}
\]

\(^3\) For a more rigorous justification, one can invoke the Markov lemma [19].
Note: $\Delta X, \Delta \tilde{X}_1$ and $\Delta \tilde{X}_2$ are independent of $\tilde{X}_j^2$.

We can compute

\[ E(X_1|X_0, X_1^2) = b_2 X + b_3 \tilde{X}_2, \]

\[ E(X_1|X_0, X_1^2, |X_1) = b_4 \tilde{X}_1, \]

\[ E(\Delta \tilde{X}_2|\Delta X, \Delta \tilde{X}_1|\Delta X_2) = b_7 \Delta \tilde{X}_2, \]

and

\[
\text{var}(X_1|X_0, X_1^2) = E(\Delta X)^2 = \frac{\sigma_X^2 (\sigma_{X_0}^2 + \sigma_{X_1}^2 + \sigma_{X_2}^2)}{\sigma_X^2 + \sigma_{X_0}^2 + \sigma_{X_1}^2 + \sigma_{X_2}^2},
\]

\[
\text{var}(E(\Delta \tilde{X}_2|\Delta X, \Delta \tilde{X}_1)|\Delta X_2) = \frac{\sigma_X^4}{(\sigma_X^2 + \sigma_{X_0}^2 + \sigma_{X_1}^2 + \sigma_{X_2}^2)^2} \left[ \frac{\sigma_X^2 + (\sigma_{X_0}^2 - \sigma_{X_1}^2 \sigma_{X_2})^2}{\sigma_{X_0}^2 + \sigma_{X_1}^2 + \sigma_{X_2}^2} \right] - \frac{\sigma_X^4}{(\sigma_X^2 + \sigma_{X_0}^2 + \sigma_{X_1}^2 + \sigma_{X_2}^2)^2} \left[ (\sigma_{X_0}^2 + \sigma_{X_1}^2 + \sigma_{X_2}^2)^2 (\sigma_X^2 + \sigma_{X_0}^2 + \sigma_{X_1}^2 + \sigma_{X_2}^2)^2 \right],
\]

\[
E(\Delta \tilde{X}_2)^2 = \frac{\sigma_X^4}{(\sigma_X^2 + \sigma_{X_0}^2 + \sigma_{X_1}^2 + \sigma_{X_2}^2)^2} \left[ \frac{\sigma_X^2 + (\sigma_{X_0}^2 - \sigma_{X_1}^2 \sigma_{X_2})^2}{\sigma_{X_0}^2 + \sigma_{X_1}^2 + \sigma_{X_2}^2} \right] - \frac{\sigma_X^4}{(\sigma_X^2 + \sigma_{X_0}^2 + \sigma_{X_1}^2 + \sigma_{X_2}^2)^2} \left[ (\sigma_{X_0}^2 + \sigma_{X_1}^2 + \sigma_{X_2}^2)^2 (\sigma_X^2 + \sigma_{X_0}^2 + \sigma_{X_1}^2 + \sigma_{X_2}^2)^2 \right],
\]

\[
\frac{E\Delta X \Delta \tilde{X}_1}{E(\Delta \tilde{X}_2)^2} = \frac{\sigma_X^4}{(\sigma_X^2 + \sigma_{X_0}^2 + \sigma_{X_1}^2 + \sigma_{X_2}^2)^2} \left[ \frac{\sigma_X^2 + (\sigma_{X_0}^2 - \sigma_{X_1}^2 \sigma_{X_2})^2}{\sigma_{X_0}^2 + \sigma_{X_1}^2 + \sigma_{X_2}^2} \right] - \frac{\sigma_X^4}{(\sigma_X^2 + \sigma_{X_0}^2 + \sigma_{X_1}^2 + \sigma_{X_2}^2)^2} \left[ (\sigma_{X_0}^2 + \sigma_{X_1}^2 + \sigma_{X_2}^2)^2 (\sigma_X^2 + \sigma_{X_0}^2 + \sigma_{X_1}^2 + \sigma_{X_2}^2)^2 \right],
\]

\[
E\Delta X \Delta \tilde{X}_2 = \frac{\sigma_X^4}{(\sigma_X^2 + \sigma_{X_0}^2 + \sigma_{X_1}^2 + \sigma_{X_2}^2)^2} \left[ \frac{\sigma_X^2 + (\sigma_{X_0}^2 - \sigma_{X_1}^2 \sigma_{X_2})^2}{\sigma_{X_0}^2 + \sigma_{X_1}^2 + \sigma_{X_2}^2} \right] - \frac{\sigma_X^4}{(\sigma_X^2 + \sigma_{X_0}^2 + \sigma_{X_1}^2 + \sigma_{X_2}^2)^2} \left[ (\sigma_{X_0}^2 + \sigma_{X_1}^2 + \sigma_{X_2}^2)^2 (\sigma_X^2 + \sigma_{X_0}^2 + \sigma_{X_1}^2 + \sigma_{X_2}^2)^2 \right],
\]

By Lemma 1, we have

\[
R_1 = \frac{E(X_0, X_2; X_1)}{I(E(X_0, X_2); X_1)} = I(E(X_0, X_2); X_1),
\]

\[
R_2 = \frac{E(X_0, X_2)}{I(E(X_0, X_2); X_1)} = \frac{E(X_0, X_2)}{I(E(X_0, X_2); X_1)},
\]

We have used the fact that $(X_0, X_2) \rightarrow (X_0|X_1, X_2) \rightarrow X_1$ and $(X_0, X_1) \rightarrow (X_0|X_1, X_2) \rightarrow X_2$. 

where

\[
b_1 = X_1 - \sigma_X^2 + \sigma_{X_0}^2 - \sigma_{X_2}^2,
\]

\[
b_2 = \frac{\sigma_X^2}{\sigma_{X_0}^2 + \sigma_{X_1}^2 + \sigma_{X_2}^2}
\]

\[
b_3 = \frac{\sigma_X^2}{\sigma_{X_0}^2 + \sigma_{X_1}^2 + \sigma_{X_2}^2}
\]

\[
b_4 = \frac{\sigma_X^2}{\sigma_{X_0}^2 + \sigma_{X_1}^2 + \sigma_{X_2}^2}
\]

\[
b_5 = \frac{E(\Delta X \Delta \tilde{X}_2)^2 - (E(\Delta X \Delta \tilde{X}_1))^2}{E(\Delta X)^2 E(\Delta \tilde{X}_1)^2}
\]

\[
b_6 = \frac{E(\Delta X \Delta \tilde{X}_1)^2 - (E(\Delta X \Delta \tilde{X}_2))^2}{E(\Delta X)^2 E(\Delta \tilde{X}_2)^2}
\]

\[
b_7 = \frac{E(\Delta X)^2 E(\Delta \tilde{X}_1)^2 - (E(\Delta X \Delta \tilde{X}_2))^2}{E(\Delta X)^2 E(\Delta \tilde{X}_2)^2 - (E(\Delta X \Delta \tilde{X}_1))^2}
\]
The system diagram of the above scheme is shown in Fig. 3.

![System Diagram](image)

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


