Gaussian Multiple Description Coding with Low-Density Generator Matrix Codes

Jun Chen, Ying Zhang, Sorina Dumitrescu

Abstract

It is shown that the coding problem for an arbitrary point on the dominant face of an L-description El Gamal-Cover (EGC) region can be converted to that for a vertex of a K-description EGC region for some $K \leq 2L-1$, where the latter problem can be solved via successive coding. A practical scheme is proposed for the quadratic Gaussian case by reducing each step in successive coding to a Gaussian quantization operation and implementing such an operation using low-density generator matrix codes. The effectiveness of this scheme is verified through extensive simulation experiments.

Index Terms

Lossy source coding, low-density generator matrix, message-passing algorithm, multiple description coding, quantization splitting.

I. Introduction

Multiple description coding is a classic problem in network information theory. El Gamal and Cover [1] established a general inner bound of the 2-description rate region, commonly referred to as the EGC region. Ozarow [2] proved that the EGC region is tight in the quadratic Gaussian case. In fact, it has been shown, by refining and generalizing Ozarow's proof technique, that a natural extension of the EGC region to the L-description case is tight for Gaussian multiple description coding with individual and central distortion constraints [3]–[5].

©2010 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

The authors are with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON L8S 4K1, Canada (email: junchen@mail.ece.mcmaster.ca, yingzhang@grads.ece.mcmaster.ca, sorina@mail.ece.mcmaster.ca).

Driven by a wide range of potential applications (say, multimedia data transmission over lossy networks), the code design aspect of multiple description coding has also received considerable attention. Indeed, a number of practical multiple description coding schemes have been proposed over the last two decades by leveraging various quantization and signal processing techniques, e.g., scalar quantization [6]–[12], trellis-coded quantization [13], lattice vector quantization [14]–[19], correlating transforms [20], [21], and Delta-Sigma quantization [22]. It is worth noting that research work on code design has been largely focused on the quadratic Gaussian case, which is the only case where the 2-description rate region has been completely characterized. However, this is not a severe limitation since coding schemes developed for Gaussian sources often have certain performance guarantee when adopted for other continuous-valued sources due to the extremal properties of the Gaussian distribution [23].

In view of the recent success of low-density generator matrix (LDGM) codes for lossy source coding [24]–[27], it is natural to investigate the application of these codes to multiple description coding. As a first step in this direction, we shall explore an alternative implementation of the Gaussian 2-description coding scheme proposed in [19] with lattice codes replaced by LDGM codes as well as the extension to the general L-description case. Although we shall follow the general strategy of [19], [28] by reducing the Gaussian multiple description problem to a sequence of Gaussian quantization problems via Gram-Schmidt orthogonalization, there are several noteworthy conceptual differences.

1) A direct calculation of the coefficients in Gram-Schmidt orthogonalization, though possible for the 2-description case [19], appears to be cumbersome in the general setting. Instead, we shall exploit the special structure of the covariance matrix associated with the sum-rate optimal EGC region, which not only simplifies the calculation, but also leads to an efficient implementation of Gram-Schmidt orthogonalization.
2) We give a new interpretation of the quantization splitting method developed in [19], [28] by eliminating the use of conditional codebooks.

The remainder of this paper is organized as follows. In Section II we show that the coding problem for an arbitrary point on the dominant face of an L-description EGC region can be converted to that for a vertex of a K-description EGC region for some $K \leq 2L-1$, where the latter problem can be solved via successive coding; moreover, in the quadratic Gaussian case

each step in successive coding can be reduced to a Gaussian quantization operation. In Section III we propose a practical Gaussian multiple description coding scheme by implementing each Gaussian quantization operation using LDGM codes. The effectiveness of this scheme is verified through extensive simulation experiments. We conclude the paper in Section IV.

II. EGC Region, Successive Coding, and Quantization Splitting

Let $\{X(t)\}_{t=1}^{\infty}$ be a stationary and memoryless process with marginal distribution p_X on $\mathcal{X}$ and $d : \mathcal{X} \times \hat{\mathcal{X}} \to [0, \infty)$ be a distortion measure, where $\mathcal{X}$ and $\hat{\mathcal{X}}$ are respectively the source alphabet and the reconstruction alphabet. We say a rate tuple $(R_1, \cdots, R_L)$ is achievable subject to individual distortion constraints $d_{\{\ell\}}$, $\ell = 1, \cdots, L$, and central distortion constraint $d_{\{1,\cdots,L\}}$ if given any $\epsilon > 0$ there exist, for all sufficiently large n, encoding functions $f_\ell^{(n)} : \mathcal{X}^n \to \{1, \cdots, \lfloor 2^{n(R_\ell+\epsilon)} \rfloor\}$, $\ell = 1, \cdots, L$, and decoding functions $g_{\{\ell\}}^{(n)} : \{1, \cdots, \lfloor 2^{n(R_\ell+\epsilon)} \rfloor\} \to \hat{\mathcal{X}}^n$, $\ell = 1, \cdots, L$, and $g_{\{1,\cdots,L\}}^{(n)} : \prod_{\ell=1}^{L} \{1, \cdots, \lfloor 2^{n(R_\ell+\epsilon)} \rfloor\} \to \hat{\mathcal{X}}^n$ such that

$$\frac{1}{n}\sum_{t=1}^{n} \mathbb{E}[d(X(t), \hat{X}_{\{\ell\}}(t))] \le d_{\{\ell\}} + \epsilon, \quad \ell = 1, \cdots, L,$$
$$\frac{1}{n}\sum_{t=1}^{n} \mathbb{E}[d(X(t), \hat{X}_{\{1,\cdots,L\}}(t))] \le d_{\{1,\cdots,L\}} + \epsilon,$$

where $\hat{X}_{\{\ell\}}^n = g_{\{\ell\}}^{(n)}(f_\ell^{(n)}(X^n))$, $\ell = 1, \cdots, L$, and $\hat{X}_{\{1,\cdots,L\}}^n = g_{\{1,\cdots,L\}}^{(n)}(f_1^{(n)}(X^n), \cdots, f_L^{(n)}(X^n))$. The union of such rate tuples, denoted as $\mathcal{R}(d_{\{1\}}, \cdots, d_{\{L\}}, d_{\{1,\cdots,L\}})$, is called the L-description rate region subject to individual distortion constraints $d_{\{\ell\}}$, $\ell = 1, \cdots, L$, and central distortion constraint $d_{\{1,\cdots,L\}}$. Note that more generally one can impose a distortion constraint for each non-empty subset of $\{1, \cdots, L\}$. We choose the current definition of the L-description rate region for the following reasons: 1) the L-description rate region subject to individual and central distortion constraints has been completely characterized in the quadratic Gaussian case; 2) the current definition is suitable for the case where the distortion constraints are only imposed for the worst case scenario (i.e., reconstruction based on a single description) and the best case scenario (i.e., reconstruction based on the complete set of descriptions).

A. EGC Region and Successive Coding

For any L auxiliary random variables $U_{\{1\}}, \cdots, U_{\{L\}}$ jointly distributed with the generic source variable X, we define $\mathcal{R}(p_{U_{\{1\}},\cdots,U_{\{L\}}|X})$ as the set of rate tuples $(R_1, \cdots, R_L)$ satisfying

$$\sum_{\ell\in\mathcal{A}} R_\ell \geq \sum_{\ell\in\mathcal{A}} H(U_{\{\ell\}}) - H(U_{\{\ell\}}, \ell \in \mathcal{A}|X),$$
$$\emptyset \subset \mathcal{A} \subseteq \{1, \cdots, L\}.$$

Let $\mathcal{P}(d_{\{1\}}, \cdots, d_{\{L\}}, d_{\{1,\cdots,L\}})$ be the set of conditional distributions $p_{U_{\{1\}},\cdots,U_{\{L\}}|X}$ such that

$$\mathbb{E}[d(X, g_{\{\ell\}}(U_{\{\ell\}}))] \leq d_{\{\ell\}}, \quad \ell = 1, \cdots, L, \tag{1}$$

$$\mathbb{E}[d(X, g_{\{1,\cdots,L\}}(U_{\{1\}}, \cdots, U_{\{L\}}))] \leq d_{\{1,\cdots,L\}} \tag{2}$$

for some functions $g_{\{\ell\}}$, $\ell = 1, \cdots, L$, and $g_{\{1,\cdots,L\}}$. The EGC region $\mathcal{R}_{EGC}(d_{\{1\}}, \cdots, d_{\{L\}}, d_{\{1,\cdots,L\}})$ is defined[1] as the convex closure of

$$\bigcup_{p_{U_{\{1\}},\cdots,U_{\{L\}}|X}\in\mathcal{P}(d_{\{1\}},\cdots,d_{\{L\}},d_{\{1,\cdots,L\}})} \mathcal{R}(p_{U_{\{1\}},\cdots,U_{\{L\}}|X}).$$

It is known [3] that

$$\mathcal{R}_{EGC}(d_{\{1\}}, \cdots, d_{\{L\}}, d_{\{1,\cdots,L\}})$$
$$\subseteq \mathcal{R}(d_{\{1\}}, \cdots, d_{\{L\}}, d_{\{1,\cdots,L\}}).$$

Roughly speaking, one may view $U_{\{\ell\}}$, $\ell = 1, \cdots, L$, as L descriptions of source X; moreover, $g_{\{1\}}(U_{\{\ell\}})$, $\ell = 1, \cdots, L$, can be interpreted as the reconstructions based on individual descriptions while $g_{\{1,\cdots,L\}}(U_{\{1\}}, \cdots, U_{\{L\}})$ can be interpreted as the reconstruction based on the complete set of descriptions.

Henceforth we shall primarily focus on $\mathcal{R}(p_{U_{\{1\}},\cdots,U_{\{L\}}|X})$ and simply refer to it as the EGC region when no confusion can arise. As observed in [4], [28], $\mathcal{R}(p_{U_{\{1\}},\cdots,U_{\{L\}}|X})$ is a contra-polymatroid and its vertices can be easily characterized. Specifically, $(R_1(\pi), \cdots, R_L(\pi))$ is a

[1]Note that $\mathcal{R}_{EGC}(d_{\{1\}}, \cdots, d_{\{L\}}, d_{\{1,\cdots,L\}})$ is consistent with the standard definition of the EGC region for the 2-description case [1]. More precisely, $\mathcal{R}_{EGC}(d_{\{1\}}, \cdots, d_{\{L\}}, d_{\{1,\cdots,L\}})$ becomes the EGC* region (an antecedent version of the EGC region) when $L = 2$, which is known to be equivalent to the EGC region [29].

vertex of $\mathcal{R}(p_{U_{\{1\}},\cdots,U_{\{L\}}|X})$ for every permutation π on $\{1,\cdots,L\}$, where

$$R_{\pi(1)}(\pi) = I(X;U_{\{\pi(1)\}}), \tag{3}$$

$$R_{\pi(\ell)}(\pi) = I(X,U_{\{\pi(1)\}},\cdots,U_{\{\pi(\ell-1)\}};U_{\{\pi(\ell)\}}),$$
$$\ell = 2,\cdots,L. \tag{4}$$

The dominant face of $\mathcal{R}(p_{U_{\{1\}},\cdots,U_{\{L\}}|X})$, denoted as $\mathcal{D}(p_{U_{\{1\}},\cdots,U_{\{L\}}|X})$, is the set of rate tuples in $\mathcal{R}(p_{U_{\{1\}},\cdots,U_{\{L\}}|X})$ satisfying

$$\sum_{\ell=1}^{L} R_\ell = \sum_{\ell=1}^{L} H(U_{\{\ell\}}) - H(U_{\{1\}},\cdots,U_{\{L\}}|X).$$

It can be readily verified by leveraging (3) and (4) that all the vertices are on the dominant face.

Note that the vertices of $\mathcal{R}(p_{U_{\{1\}},\cdots,U_{\{L\}}|X})$ are of particular importance since every rate tuple in $\mathcal{R}(p_{U_{\{1\}},\cdots,U_{\{L\}}|X})$ is dominated in a component-wise manner by some rate tuple in $\mathcal{D}(p_{U_{\{1\}},\cdots,U_{\{L\}}|X})$ and the latter can be expressed as a convex combination of no more than L vertices. As pointed out in [19], [28], the expression of the vertices (see (3) and (4)) suggests a successive coding scheme which can be roughly described as follows[2]: for vertex $(R_1(\pi),\cdots,R_L(\pi))$, one first uses X to produce $U_{\{\pi(1)\}}$, then successively from $\ell=2$ to L, uses $(X,U_{\{\pi(1)\}},\cdots,U_{\{\pi(\ell-1)\}})$ to produce $U_{\{\pi(\ell)\}}$. Furthermore, every rate tuple in $\mathcal{D}(p_{U_{\{1\}},\cdots,U_{\{L\}}|X})$ is achievable via suitable timesharing of such successive coding schemes.

Now we proceed to propose an efficient implementation of the aforementioned successive coding scheme in the quadratic Gaussian case, where p_X is a Gaussian distribution with mean zero and variance σ_X^2, and $d(\cdot,\cdot)$ is the standard squared error distortion measure. In this setting it is known [5] that $\mathcal{R}_{EGC}(d_{\{1\}},\cdots,d_{\{L\}},d_{\{1,\cdots,L\}}) = \mathcal{R}(d_{\{1\}},\cdots,d_{\{L\}},d_{\{1,\cdots,L\}})$; moreover, it suffices to assume that $U_{\{1\}},\cdots,U_{\{L\}}$ are zero-mean and jointly Gaussian with the generic source variable X. By exploiting the properties of the Gaussian distribution, the following simplified version of the successive coding scheme, referred to as the successive quantization scheme, was developed in [19], [28]. Without loss of generality, we shall assume $\pi(\ell)=\ell$, $\ell=1,\cdots,L$.

[2]For simplicity, here we describe the scheme in the form of single-letter operation. However, it should be noted that to approach the information-theoretic limits, one has to implement such a scheme over long blocks.

Using the Gram-Schmidt orthogonalization procedure, we can write

$$U_{\{\ell\}} = \hat{U}_{\{\ell\}} + \Delta_\ell, \quad \ell = 1, \cdots, L,$$

where

$$\hat{U}_{\{1\}} = \mathbb{E}[U_{\{1\}}|X],$$
$$\hat{U}_{\{\ell\}} = \mathbb{E}[U_{\{\ell\}}|X, U_{\{1\}}, \cdots, U_{\{\ell-1\}}], \quad \ell = 2, \cdots, L,$$

and $X, \Delta_1, \cdots, \Delta_L$ are jointly independent and Gaussian. Now (3) and (4) can be rewritten as

$$R_\ell(\pi) = I(\hat{U}_{\{\ell\}}; U_{\{\ell\}}) = I(\hat{U}_{\{\ell\}}; \hat{U}_{\{\ell\}} + \Delta_\ell),$$
$$\ell = 1, \cdots, L. \tag{5}$$

As observed in [19], [28], one can readily obtain an L-step successive quantization scheme by interpreting $\hat{U}_{\{\ell\}}$, $U_{\{\ell\}}$, and Δ_ℓ in (5) respectively as the quantization input, the quantization output, and the quantization error at step ℓ. Note that the explicit expressions of $\hat{U}_{\{1\}}, \cdots, \hat{U}_{\{L\}}$ in terms of $(X, U_{\{1\}}, \cdots, U_{\{L\}})$ depend on the covariance matrix of $(X, U_{\{1\}}, \cdots, U_{\{L\}})$, which in turn depends on distortion constraints $d_{\{1\}}, \cdots, d_{\{L\}}$, and $d_{\{1,\cdots,L\}}$. A direct derivation of such expression, though possible for the case $L = 2$ [19], appears to be cumbersome for general L. Fortunately, it turns out that the special structure of the optimal covariance matrix of $(X, U_{\{1\}}, \cdots, U_{\{L\}})$ allows for an efficient implementation of Gram-Schmidt orthogonalization as well as a simple calculation of the relevant coefficients.

Without loss of generality, we shall assume $0 < d_{\{\ell\}} \leq \sigma_X^2$, $\ell = 1, \cdots, L$, and $0 < d_{\{1,\cdots,L\}} \leq \sigma_X^2$. Define

$$R_\Sigma(d_{\{1\}}, \cdots, d_{\{L\}}, d_{\{1,\cdots,L\}}) = \min\Big\{ \sum_{\ell=1}^{L} R_\ell :$$
$$(R_1, \cdots, R_L) \in \mathcal{R}(d_{\{1\}}, \cdots, d_{\{L\}}, d_{\{1,\cdots,L\}}) \Big\}.$$

It is known [4], [5] that

$$R_\Sigma(d_{\{1\}}, \cdots, d_{\{L\}}, d_{\{1,\cdots,L\}}) =$$

$$\max_{a\in[0,\sigma_X^2]} \frac{1}{2}\log\left(\frac{\sigma_X^{4L-2}(\sigma_X^2 d_{\{1,\cdots,L\}} - a d_{\{1,\cdots,L\}} + a\sigma_X^2)}{d_{\{1,\cdots,L\}}\prod_{\ell=1}^{L}(\sigma_X^2 d_{\{\ell\}} - a d_{\{\ell\}} + a\sigma_X^2)}\right).$$

In particular,

$$\begin{aligned} &R_\Sigma(d_{\{1\}},\cdots,d_{\{L\}},d_{\{1,\cdots,L\}}) \\ &= \begin{cases} \frac{1}{2}\sum\limits_{\ell=1}^{L}\log(\frac{\sigma_X^2}{d_{\{\ell\}}}), & d_{\{1,\cdots,L\}} \geq \overline{d}_{\{1,\cdots,L\}} \\ \frac{1}{2}\log(\frac{\sigma_X^2}{d_{\{1,\cdots,L\}}}), & d_{\{1,\cdots,L\}} \leq \underline{d}_{\{1,\cdots,L\}} \end{cases}, \end{aligned}$$

where

$$\begin{aligned} \overline{d}_{\{1,\cdots,L\}} &= \left(\sum_{\ell=1}^{L}\frac{1}{d_{\{\ell\}}} - \frac{L-1}{\sigma_X^2}\right)^{-1}, \\ \underline{d}_{\{1,\cdots,L\}} &= \sum_{\ell=1}^{L} d_\ell - (L-1)\sigma_X^2. \end{aligned}$$

Therefore, if $d_{\{1,\cdots,L\}} \geq \overline{d}_{\{1,\cdots,L\}}$, then one can decrease $d_{\{1,\cdots,L\}}$ until $d_{\{1,\cdots,L\}} = \overline{d}_{\{1,\cdots,L\}}$ without affecting $R_\Sigma(d_{\{1\}},\cdots,d_{\{L\}},d_{\{1,\cdots,L\}})$; similarly, if $d_{\{1,\cdots,L\}} \leq \underline{d}_{\{1,\cdots,L\}}$, then one can decrease one of $d_{\{\ell\}}$, $\ell = 1,\cdots,L$, until $d_{\{1,\cdots,L\}} = \underline{d}_{\{1,\cdots,L\}}$ without affecting $R_\Sigma(d_{\{1\}},\cdots,d_{\{L\}},d_{\{1,\cdots,L\}})$. As a consequence, there is no loss of generality in assuming $\underline{d}_{\{1,\cdots,L\}} \leq d_{\{1,\cdots,L\}} \leq \overline{d}_{\{1,\cdots,L\}}$. In this case we have

$$\begin{aligned} &R_\Sigma(d_{\{1\}},\cdots,d_{\{L\}},d_{\{1,\cdots,L\}}) \\ &= \frac{1}{2}\log\left(\frac{\sigma_X^{4L-2}(\sigma_X^2 d_{\{1,\cdots,L\}} - \hat{a} d_{\{1,\cdots,L\}} + \hat{a}\sigma_X^2)}{d_{\{1,\cdots,L\}}\prod_{\ell=1}^{L}(\sigma_X^2 d_{\{\ell\}} - \hat{a} d_{\{\ell\}} + \hat{a}\sigma_X^2)}\right), \end{aligned}$$

where $\hat{a} \in [0,\sigma_X^2]$ is the solution to the following equation

$$\left(\frac{\sigma_X^2 d_{\{1,\cdots,L\}}}{\sigma_X^2 - d_{\{1,\cdots,L\}}} + a\right)^{-1} = \sum_{\ell=1}^{L}\left(\frac{\sigma_X^2 d_{\{\ell\}}}{\sigma_X^2 - d_{\{\ell\}}} + a\right)^{-1}.$$

In particular, when $d_{\{1\}} = \cdots = d_{\{L\}} = d$, we have

$$\hat{a} = \frac{\sigma_X^4 d - L\sigma_X^4 d_{\{1,\cdots,L\}} - \sigma_X^2 d d_{\{1,\cdots,L\}} + L\sigma_X^2 d d_{\{1,\cdots,L\}}}{(L-1)(\sigma_X^2 - d)(\sigma_X^2 - d_{\{1,\cdots,L\}})}$$

and

$$R_\Sigma(d,\cdots,d,d_{\{1,\cdots,L\}}) = \frac{1}{2}\log\left(\frac{(L-1)^{L-1}\sigma_X^2(\sigma_X^2-d_{\{1,\cdots,L\}})^L}{L^L d_{\{1,\cdots,L\}}(\sigma_X^2-d)(d-d_{\{1,\cdots,L\}})^{L-1}}\right).$$

Note that in the current setting (1) and (2) can be rewritten as

$$\mathbb{E}[(X-\mathbb{E}[X|U_{\{\ell\}}])^2] \le d_{\{\ell\}}, \quad \ell=1,\cdots,L, \tag{6}$$

$$\mathbb{E}[(X-\mathbb{E}[X|U_{\{1\}},\cdots,U_{\{L\}}])^2] \le d_{\{1,\cdots,L\}}. \tag{7}$$

It is known [4], [5] that if $U_{\{1\}},\cdots,U_{\{L\}}$ are zero-mean and jointly Gaussian with X such that

$$\mathbb{E}[XU_{\{\ell\}}] = \sigma_X^2, \quad \ell=1,\cdots,L, \tag{8}$$

$$\mathbb{E}[U_{\{\ell\}}U_{\{\ell'\}}] = \begin{cases} \sigma_X^2+\sigma_{\{\ell\}}^2, & \ell=\ell' \\ \sigma_X^2-\hat{a}, & \ell\neq\ell' \end{cases}, \tag{9}$$

where $\sigma_{\{\ell\}}^2 = \frac{\sigma_X^2 d_{\{\ell\}}}{\sigma_X^2-d_{\{\ell\}}}$, then rate tuples in $\mathcal{D}(p_{U_{\{1\}},\cdots,U_{\{L\}}|X})$ achieve the minimum sum rate $R_\Sigma(d_{\{1\}},\cdots,d_{\{L\}},d_{\{1,\cdots,L\}})$ and distortion constraints (6) and (7) are satisfied; the corresponding $\mathcal{R}(p_{U_{\{1\}},\cdots,U_{\{L\}}|X})$ will be referred to as the sum-rate optimal Gaussian EGC region. Now we proceed to give an explicit construction of such $(U_{\{1\}},\cdots,U_{\{L\}})$. Let

$$\sigma_{\{1,\cdots,\ell\}}^2 = \left(\sum_{i=1}^{\ell}\left(\sigma_{\{i\}}^2+\hat{a}\right)^{-1}\right)^{-1} - \hat{a}, \quad \ell=2,\cdots,L.$$

Note that

$$\hat{a} = \sqrt{(\sigma_{\{1,\cdots,\ell-1\}}^2-\sigma_{\{1,\cdots,\ell\}}^2)(\sigma_{\{\ell\}}^2-\sigma_{\{1,\cdots,\ell\}}^2)} - \sigma_{\{1,\cdots,\ell\}}^2, \quad \ell=2,\cdots,L.$$

Let $N_{\{1,\cdots,L\}}$ and $N'_{\{1,\cdots,\ell\}}$, $\ell=2,\cdots,L$, be L zero-mean Gaussian random variables, where the variance of $N_{\{1,\cdots,L\}}$ is $\sigma_{\{1,\cdots,L\}}^2$ and the rest have unit variance. We assume that X, $N_{\{1,\cdots,L\}}$, and $N'_{\{1,\cdots,\ell\}}$, $\ell=2,\cdots,L$, are jointly independent. One can successively construct

$$N_{\{1,\cdots,\ell\}}$$

$$= N_{\{1,\cdots,\ell+1\}} + \sqrt{\sigma^2_{\{1,\cdots,\ell\}} - \sigma^2_{\{1,\cdots,\ell+1\}}} N'_{\{1,\cdots,\ell+1\}}$$

from $\ell = L-1$ to 1. Now let

$$\begin{aligned} &U_{\{1,\cdots,\ell\}} = X + N_{\{1,\cdots,\ell\}}, \quad \ell = 1,\cdots,L, \\ &U_{\{\ell\}} = U_{\{1,\cdots,\ell\}} - \sqrt{\sigma^2_{\{\ell\}} - \sigma^2_{\{1,\cdots,\ell\}}} N'_{\{1,\cdots,\ell\}}, \\ &\quad \ell = 2,\cdots,L. \end{aligned}$$

It can be verified that the constructed $(U_{\{1\}},\cdots,U_{\{L\}})$ satisfies (8) and (9). Note that $(U_1,\cdots,U_{\{\ell\}}) - U_{\{1,\cdots,\ell\}} - (X, U_{\{\ell+1\}},\cdots,U_{\{L\}})$ form a Markov chain, $\ell = 2,\cdots,L$. Therefore, we have

$$\begin{aligned} &\hat{U}_{\{1\}} = X, \\ &\hat{U}_{\{\ell\}} = \mathbb{E}[U_{\{\ell\}}|X, U_{\{1,\cdots,\ell-1\}}] = \gamma_{\ell-1}X + \beta_{\ell-1}U_{\{1,\cdots,\ell-1\}}, \\ &\quad \ell = 2,\cdots,L, \end{aligned} \tag{10}$$

where

$$\begin{aligned} \gamma_{\ell-1} &= 1 + \frac{\hat{a}}{\sigma^2_{\{1,\cdots,\ell-1\}}}, \\ \beta_{\ell-1} &= -\frac{\hat{a}}{\sigma^2_{\{1,\cdots,\ell-1\}}}. \end{aligned}$$

It is easy to see that

$$U_{\{1,\cdots,\ell\}} = \eta_{\ell-1}U_{\{1,\cdots,\ell-1\}} + \bar{\eta}_{\ell-1}U_{\{\ell\}}, \quad \ell = 2,\cdots,L, \tag{11}$$

where

$$\begin{aligned} \eta_{\ell-1} &= 1 - \bar{\eta}_{\ell-1} \\ &= \frac{\sqrt{\sigma^2_{\{\ell\}} - \sigma^2_{\{1,\cdots,\ell\}}}}{\sqrt{\sigma^2_{\{1,\cdots,\ell-1\}} - \sigma^2_{\{1,\cdots,\ell\}}} + \sqrt{\sigma^2_{\{\ell\}} - \sigma^2_{\{1,\cdots,\ell\}}}}. \end{aligned}$$

This recurrence relation leads to an efficient implementation of Gram-Schmidt orthogonalization

(see Fig. 1). It can also be verified that

$$\mathbb{E}[X|U_{\{\ell\}}] = \alpha_{\{\ell\}} U_{\{\ell\}}, \quad \ell = 1, \cdots, L, \tag{12}$$

$$\begin{aligned}
&\mathbb{E}[X|U_{\{1\}}, \cdots, U_{\{\ell\}}] = \mathbb{E}[X|U_{\{1,\cdots,\ell\}}] = \alpha_{\{1,\cdots,\ell\}} U_{\{1,\cdots,\ell\}}, \\
&\quad \ell = 2, \cdots, L,
\end{aligned} \tag{13}$$

$$\begin{aligned}
&\mathbb{E}[\Delta_1^2] = \sigma_{\{1\}}^2, \\
&\mathbb{E}[\Delta_\ell^2] = \sigma_{\{\ell\}}^2 - \frac{\hat{a}^2}{\sigma_{\{1,\cdots,\ell-1\}}^2}, \quad \ell = 2, \cdots, L,
\end{aligned}$$

where

$$\begin{aligned}
\alpha_{\{\ell\}} &= \frac{\sigma_X^2}{\sigma_X^2 + \sigma_{\{\ell\}}^2}, \\
\alpha_{\{1,\cdots,\ell\}} &= \frac{\sigma_X^2}{\sigma_X^2 + \sigma_{\{1,\cdots,\ell\}}^2}.
\end{aligned}$$

In particular, when $d_{\{1\}} = \cdots = d_{\{\ell\}} = d$, we have

$$\begin{aligned}
&\sigma_{\{1\}}^2 = \cdots = \sigma_{\{L\}}^2 = \frac{\sigma_X^2 d}{\sigma_X^2 - d} \triangleq \sigma^2, \\
&\sigma_{\{1,\cdots,\ell\}}^2 = \frac{1}{\ell}\sigma^2 - \frac{\ell-1}{\ell}\hat{a} = \frac{\sigma_X^2 d}{\ell(\sigma_X^2 - d)} - \frac{\ell-1}{\ell}\hat{a}, \\
&\quad \ell = 2, \cdots, L, \\
&U_{\{1,\cdots,\ell\}} = \frac{\ell-1}{\ell} U_{\{1,\cdots,\ell-1\}} + \frac{1}{\ell} U_{\{\ell\}} = \frac{1}{\ell}\sum_{i=1}^{\ell} U_{\{i\}}, \\
&\quad \ell = 2, \cdots, L, \\
&\mathbb{E}[\Delta_1^2] = \sigma^2, \\
&\mathbb{E}[\Delta_\ell^2] = \sigma^2 - \frac{(\ell-1)\hat{a}^2}{\sigma^2 - (\ell-2)\hat{a}}, \quad \ell = 2, \cdots, L.
\end{aligned}$$

It should be pointed out that although we have mainly focused on the minimum sum rate, one can obtain similar results in a more general setting. Indeed, it can be shown by leveraging the construction in [5] that the Gram-Schmidt orthogonalization procedure can be simplified in essentially the same manner for a vertex that achieves a general minimum weighted sum rate. The details are left to the interested reader.

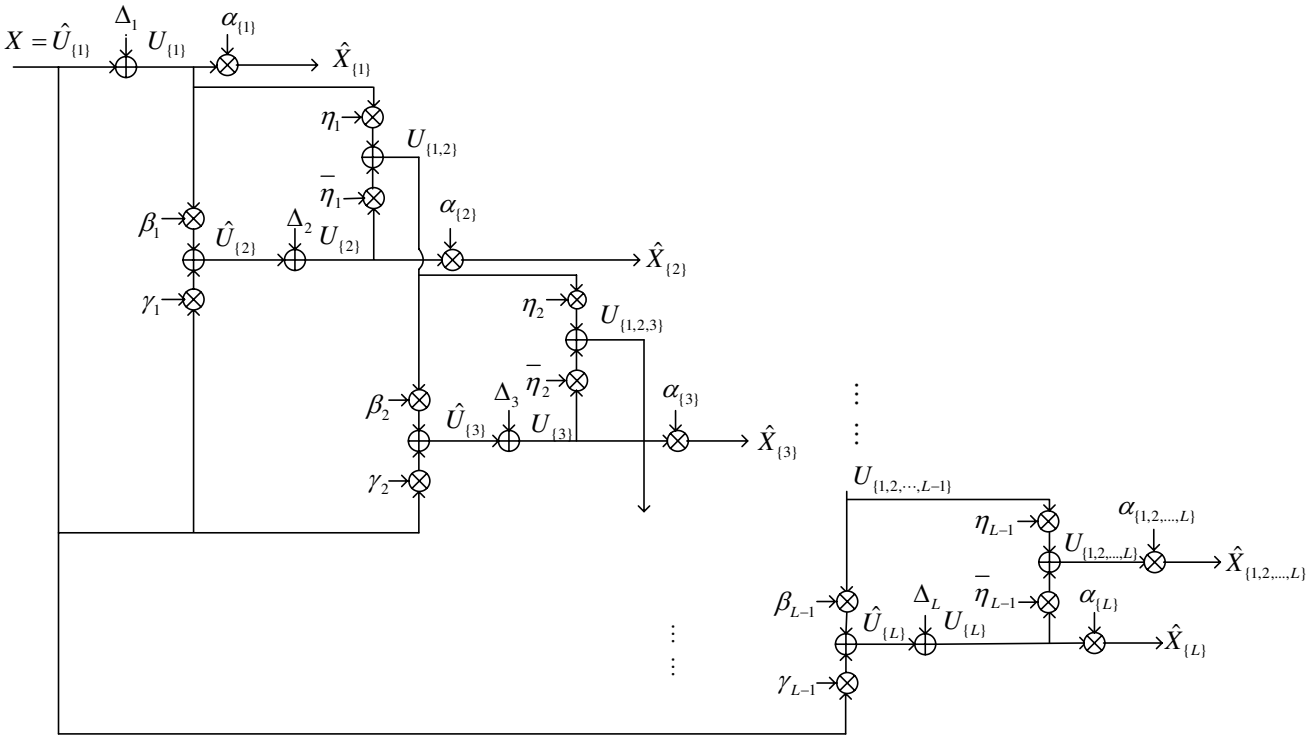

Fig. 1. Successive quantization scheme for a vertex of the sum-rate optimal Gaussian EGC region.

B. Quantization Splitting

As mentioned earlier, every rate tuple on the dominant face of an EGC region is achievable via timesharing of the successive coding schemes for vertices. Alternatively, one can use the splitting method developed in [19], [28].

Given $(X, U_{\{1\}}, \cdots, U_{\{L\}})$, we say $(U_{\{1\},1}, \cdots, U_{\{L\},1})$ is split from $(U_{\{1\}}, \cdots, U_{\{L\}})$ if $U_{\{\ell\},1} - U_{\{\ell\}} - (X, U_{\{\ell'\},1}, U_{\{\ell'\}}, \ell' \neq \ell)$ form a Markov chain for all ℓ. Let $\mathcal{U} = \{U_{\{1\},1}, U_{\{1\}}, \cdots, U_{\{L\},1}, U_{\{L\}}\}$. We say μ is a well-ordered permutation on $\mathcal{U}$ if $\mathcal{U}_{\{\ell\},1}$ is placed before $\mathcal{U}_{\{\ell\}}$ for all ℓ. For any $U \in \mathcal{U}$, let $\{U\}_\mu^-$ denote the set of random variables placed before U in μ.

It is known [19], [28] that for any $(R_1, \cdots, R_L) \in \mathcal{D}(p_{U_{\{1\}},\cdots,U_{\{L\}}|X})$, one can find $(U_{\{1\},1}, \cdots, U_{\{L\},1})$ split from $(U_{\{1\}}, \cdots, U_{\{L\}})$ and a well-ordered permutation μ such that

$$R_\ell = R_{\ell,1} + R_{\ell,2}, \quad \ell = 1, \cdots, L,$$

where

$$R_{\ell,1} = I(X, \{U_{\{\ell\},1}\}_{\mu}^{-}; U_{\{\ell\},1}),$$
$$R_{\ell,2} = I(X, \{U_{\{\ell\}}\}_{\mu}^{-}; U_{\{\ell\}}|U_{\{\ell\},1});$$

moreover, at least one $U_{\{\ell\},1}$ can be set to zero[3] and removed from μ.

Note that $R_{\ell,2}$ is expressed as a conditional mutual information. This is why in [19], [28] conditional codebooks are used in the random coding argument for the splitting method. In fact, an inspection of the random coding argument in [19], [28] reveals that the resulting scheme requires one to construct and store $2^{nR_{\ell,1}}$ conditional codebooks, each of size $2^{nR_{\ell,2}}$, for the ℓ-th description. Here we shall give a new interpretation of the splitting method by converting the expression of $R_{\ell,2}$ from a conditional form to an unconditional form and consequently eliminating the use of conditional codebooks[4]. In view of Lemma 1 in [29], there exist random variables $U_{\{1\},2}, \cdots, U_{\{L\},2}$ jointly distributed with $(X, U_{\{1\},1}, U_{\{1\}}, \cdots, U_{\{L\},1}, U_{\{L\}})$ such that the following properties are satisfied for all ℓ:

P1) $U_{\{\ell\},2}$ is independent of $U_{\{\ell\},1}$;

P2) $U_{\{\ell\}}$ is a deterministic function of $U_{\{\ell\},1}$ and $U_{\{\ell\},2}$;

P3) $U_{\{\ell\},2} - (U_{\{\ell\},1}, U_{\{\ell\}}) - (X, U_{\{\ell'\},1}, U_{\{\ell'\},2}, U_{\{\ell'\}}, \ell' \neq \ell)$ form a Markov chain.

Let $\mathcal{U}' = \{U_{\{1\},1}, U_{\{1\},2}, \cdots, U_{\{L\},1}, U_{\{L\},2}\}$ and μ' be a permutation on $\mathcal{U}'$ induced by μ with $U_{\{\ell\}}$ replaced by $U_{\{\ell\},2}$ at the corresponding positions. For any $U \in \mathcal{U}'$, let $\{U\}_{\mu'}^{-}$ denote the set of random variables placed before U in μ'. By P2) and P3), we can rewrite $R_{\ell,1}$ as

$$R_{\ell,1} = I(X, \{U_{\{\ell\},1}\}_{\mu'}^{-}; U_{\{\ell\},1}).$$

Moreover, it follows by P1), P2), and P3) that

$$\begin{aligned}
&I(X, \{U_{\{\ell\},2}\}_{\mu'}^{-}; U_{\{\ell\},2}) \\
&= I(X, \{U_{\{\ell\},2}\}_{\mu'}^{-}; U_{\{\ell\},2}|U_{\{\ell\},1}) + I(U_{\{\ell\},1}; U_{\{\ell\},2}) \\
&= I(X, \{U_{\{\ell\},2}\}_{\mu'}^{-}; U_{\{\ell\},2}|U_{\{\ell\},1})
\end{aligned}$$

[3]In this case we have $R_{\ell,1} = 0$ and $R_{\ell} = R_{\ell,2} = I(X, \{U_{\{\ell\}}\}_{\mu}^{-}; U_{\{\ell\}})$.

[4]More precisely, the new interpretation allows one to replace those $2^{nR_{\ell,1}}$ conditional codebooks with a single codebook of size $2^{nR_{\ell,2}}$.

$$= I(X, \{U_{\{\ell\}}\}_{\mu}^{-}; U_{\{\ell\}}|U_{\{\ell\},1}).$$

Therefore, we have

$$R_{\ell,2} = I(X, \{U_{\{\ell\},2}\}_{\mu'}^{-}; U_{\{\ell\},2}).$$

Now by ordering $R_{1,1}, R_{1,2}, \cdots, R_{L,1}, R_{L,2}$ according to μ', one can readily see that the coding problem for an arbitrary point on the dominant face of an L-description EGC region can be converted to that for a vertex of a K-description EGC region for some $K \leq 2L-1$ (due to the fact that at least one $U_{\{\ell\},1}$ can be set to zero and removed from μ and μ'), where the latter problem can be solved via successive coding. Note that we essentially split each description into two coarse descriptions; moreover, according to P2), the original description can be recovered from the two coarse descriptions.

More concrete results can be obtained in the quadratic Gaussian case. In this setting there is no loss of generality in assuming that $U_{\{1\},1}, U_{\{1\},2}, \cdots, U_{\{L\},1}, U_{\{L\},2}$ are zero-mean and jointly Gaussian with $(X, U_{\{1\}}, \cdots, U_{\{L\}})$. Specifically, we can let $U_{\{\ell\},1} = U_{\{\ell\}} + Z_\ell$ and $U_{\{\ell\},2} = U_{\{\ell\}} - b_\ell Z_\ell$, $\ell = 1, \cdots, L$, where Z_ℓ is a Gaussian random variable with mean zero and variance $\sigma_{Z_\ell}^2$, and $b_\ell = \frac{\mathbb{E}[U_{\{\ell\}}^2]}{\sigma_{Z_\ell}^2}$; moreover, $Z_1, \cdots, Z_L$, and $(X, U_{\{1\}}, \cdots, U_{\{L\}})$ are jointly independent. The values of $\sigma_{Z_1}^2, \cdots, \sigma_{Z_L}^2$ are determined by $(R_1, \cdots, R_L)$. Note that in the extreme case when $\sigma_{Z_\ell}^2 = \infty$, we let $U_{\{\ell\},1} = 0$ and $U_{\{\ell\},2} = U_{\{\ell\}}$; similarly, when $\sigma_{Z_\ell}^2 = 0$, we let $U_{\{\ell\},1} = U_{\{\ell\}}$ and $U_{\{\ell\},2} = 0$. It is easy to verify that P1), P2), and P3) are satisfied; in particular, we have $U_{\{\ell\}} = \tau_\ell U_{\{\ell\},1} + \bar{\tau}_\ell U_{\{\ell\},2}$, $\ell = 1, \cdots, L$, where $\tau_\ell = 1 - \bar{\tau}_\ell = \frac{b_\ell}{b_\ell + 1}$. To obtain a successive quantization scheme, one can apply the Gram-Schmidt orthogonalization procedure to $(U_{\{1\},1}, U_{\{1\},2}, \cdots, U_{\{L\},1}, U_{\{L\},2})$ with the projection order specified by μ'.

Now we proceed to give a detailed treatment of the case $L = 2$. It is known [2], [4], [5] that there is no loss of generality in assuming $\underline{d}_{\{1,2\}} \leq d_{\{1,2\}} \leq \overline{d}_{\{1,2\}}$; moreover, in this setting $\mathcal{R}(d_{\{1\}}, d_{\{2\}}, d_{\{1,2\}}) = \mathcal{R}(p_{U_{\{1\}}, U_{\{2\}}|X})$, where $p_{U_{\{1\}}, U_{\{2\}}|X}$ is the conditional Gaussian distribution specified by (8) and (9) with

$$\sigma_{\{\ell\}}^2 = \frac{\sigma_X^2 d_{\{\ell\}}}{\sigma_X^2 - d_{\{\ell\}}}, \quad \ell = 1, 2,$$

$$\hat{a} = \sqrt{\prod_{\ell=1}^{2}\left(\frac{\sigma_X^2 d_{\{\ell\}}}{\sigma_X^2 - d_{\{\ell\}}} - \frac{\sigma_X^2 d_{\{1,2\}}}{\sigma_X^2 - d_{\{1,2\}}}\right)} - \frac{\sigma_X^2 d_{\{1,2\}}}{\sigma_X^2 - d_{\{1,2\}}}.$$

Note that for any $(R_1, R_2) \in \mathcal{D}(p_{U_{\{1\}},U_{\{2\}}|X})$, we can set $U_{\{2\},1} = 0$, $U_{\{2\},2} = U_{\{2\}}$, and write

$$\begin{aligned} R_1 &= R_{1,1} + R_{1,2}, \\ R_2 &= I(X, U_{\{1\},1}; U_{\{2\}}), \end{aligned} \tag{14}$$

where

$$R_{1,1} = I(X; U_{\{1\},1}), \tag{15}$$

$$R_{1,2} = I(X, U_{\{1\},1}, U_{\{2\}}; U_{\{1\},2}). \tag{16}$$

The Gram-Schmidt orthogonalization procedure yields

$$\begin{aligned} U_{\{1\},1} &= \mathbb{E}[U_{\{1\},1}|X] + \tilde{\Delta}_1, \\ U_{\{2\}} &= \mathbb{E}[U_{\{2\}}|X, U_{\{1\},1}] + \tilde{\Delta}_2, \\ U_{\{1\},2} &= \mathbb{E}[U_{\{1\},2}|X, U_{\{1\},1}, U_{\{2\}}] + \tilde{\Delta}_3, \end{aligned}$$

where $X, \tilde{\Delta}_1, \tilde{\Delta}_2, \tilde{\Delta}_3$ are jointly independent. Therefore, we can rewrite $R_{1,1}$, R_2, and $R_{1,2}$ as

$$\begin{aligned} R_{1,1} &= I(\mathbb{E}[U_{\{1\},1}|X]; U_{\{1\},1}) \\ &= I(\mathbb{E}[U_{\{1\},1}|X]; \mathbb{E}[U_{\{1\},1}|X] + \tilde{\Delta}_1), \\ R_2 &= I(\mathbb{E}[U_{\{2\}}|X, U_{\{1\},1}]; U_{\{2\}}) \\ &= I(\mathbb{E}[U_{\{2\}}|X, U_{\{1\},1}]; \mathbb{E}[U_{\{2\}}|X, U_{\{1\},1}] + \tilde{\Delta}_2), \\ R_{1,2} &= I(\mathbb{E}[U_{\{1\},2}|X, U_{\{1\},1}, U_{\{2\}}]; U_{\{1\},2}) \\ &= I(\mathbb{E}[U_{\{1\},2}|X, U_{\{1\},1}, U_{\{2\}}]; \mathbb{E}[U_{\{1\},2}|X, U_{\{1\},1}, U_{\{2\}}] \\ &\quad + \tilde{\Delta}_3). \end{aligned}$$

It can be verified that

$$\mathbb{E}[U_{\{1\},1}|X] = X,$$

$$
\begin{aligned}
&\mathbb{E}[U_{\{2\}}|X, U_{\{1\},1}] \\
&= X + \mathbb{E}[U_{\{2\}} - X|U_{\{1\},1} - X] \\
&= X - \frac{\hat{a}}{\sigma_{\{1\}}^2 + \sigma_{Z_1}^2}(U_{\{1\},1} - X) \\
&= \nu_1 X + \nu_2 U_{\{1\},1}, \qquad (17)
\end{aligned}
$$

$$
\begin{aligned}
&\mathbb{E}[U_{\{1\},2}|X, U_{\{1\},1}, U_{\{2\}}] \\
&= X + \mathbb{E}[U_{\{1\},2} - X|U_{\{1\},1} - X, U_{\{2\}} - X] \\
&= X - \frac{\sigma_X^2 \sigma_{\{2\}}^2 + \hat{a}^2}{(\sigma_{\{1\}}^2 + \sigma_{Z_1}^2)\sigma_{\{2\}}^2 - \hat{a}^2}(U_{\{1\},1} - X) \\
&\quad - \frac{\sigma_X^2 \hat{a} + (\sigma_{\{1\}}^2 + \sigma_{Z_1}^2)\hat{a}}{(\sigma_{\{1\}}^2 + \sigma_{Z_1}^2)\sigma_{\{2\}}^2 - \hat{a}^2}(U_{\{2\}} - X) \\
&= \nu_3 X + \nu_4 U_{\{1\},1} + \nu_5 U_{\{2\}}, \qquad (18)
\end{aligned}
$$

and

$$
\begin{aligned}
\mathbb{E}[\tilde{\Delta}_1^2] &= \sigma_{\{1\}}^2 + \sigma_{Z_1}^2, \\
\mathbb{E}[\tilde{\Delta}_2^2] &= \sigma_{\{2\}}^2 - \frac{\hat{a}^2}{\sigma_{\{1\}}^2 + \sigma_{Z_1}^2}, \\
\mathbb{E}[\tilde{\Delta}_3^2] &= \sigma_{\{1\}}^2 + \frac{(\sigma_X^2 + \sigma_{\{1\}}^2)^2}{\sigma_{Z_1}^2} \\
&\quad - \frac{\sigma_X^4 \sigma_{\{2\}}^2 + 2\sigma_X^2 \hat{a}^2 + (\sigma_{\{1\}}^2 + \sigma_{Z_1}^2)\hat{a}^2}{(\sigma_{\{1\}}^2 + \sigma_{Z_1}^2)\sigma_{\{2\}}^2 - \hat{a}^2},
\end{aligned}
$$

where

$$
\begin{aligned}
\nu_1 &= \frac{\sigma_{\{1\}}^2 + \sigma_{Z_1}^2 + \hat{a}}{\sigma_{\{1\}}^2 + \sigma_{Z_1}^2}, \\
\nu_2 &= -\frac{\hat{a}}{\sigma_{\{1\}}^2 + \sigma_{Z_1}^2}, \\
\nu_3 &= \frac{(\sigma_X^2 + \sigma_{\{1\}}^2 + \sigma_{Z_1}^2)(\sigma_{\{2\}}^2 + \hat{a})}{(\sigma_{\{1\}}^2 + \sigma_{Z_1}^2)\sigma_{\{2\}}^2 - \hat{a}^2}, \\
\nu_4 &= -\frac{\sigma_X^2 \sigma_{\{2\}}^2 + \hat{a}^2}{(\sigma_{\{1\}}^2 + \sigma_{Z_1}^2)\sigma_{\{2\}}^2 - \hat{a}^2},
\end{aligned}
$$

$$\nu_5 = -\frac{\sigma_X^2 \hat{a} + (\sigma_{\{1\}}^2 + \sigma_{Z_1}^2)\hat{a}}{(\sigma_{\{1\}}^2 + \sigma_{Z_1}^2)\sigma_{\{2\}}^2 - \hat{a}^2}.$$

Note that

$$R_2 = h(U_{\{2\}}) - h(\tilde{\Delta}_2) = \frac{1}{2}\log\left(\frac{\sigma_X^2 + \sigma_{\{2\}}^2}{\sigma_{\{2\}}^2 - \frac{\hat{a}^2}{\sigma_{\{1\}}^2 + \sigma_{Z_1}^2}}\right),$$

which implies

$$\sigma_{Z_1}^2 = \frac{\hat{a}^2}{\sigma_{\{2\}}^2 - 2^{-2R_2}(\sigma_X^2 + \sigma_{\{2\}}^2)} - \sigma_{\{1\}}^2.$$

In particular, when $d_{\{1\}} = d_{\{2\}} = d$ and $R_1 = R_2 = R$, we have

$$R = \frac{1}{4}\log\left(\frac{(\sigma_X^2 + \sigma^2)^2}{\sigma^4 - \hat{a}^2}\right),$$

and consequently,

$$\sigma_{Z_1}^2 = \frac{\hat{a}^2}{\sigma^2 - 2^{-2R}(\sigma_X^2 + \sigma^2)} - \sigma^2 = \sqrt{\sigma^4 - \hat{a}^2},$$

where

$$\sigma^2 = \frac{\sigma_X^2 d}{\sigma_X^2 - d},$$
$$\hat{a} = \frac{\sigma_X^2 d}{\sigma_X^2 - d} - \frac{2\sigma_X^2 d_{\{1,2\}}}{\sigma_X^2 - d_{\{1,2\}}}.$$

The quantization splitting system for the 2-description case is depicted in Fig. 2.

III. A Practical Scheme Based on LDGM Codes

A. Gaussian Quantization with LDGM Codes

As discussed previously, each stage ℓ in the successive quantization scheme reduces to a Gaussian quantization operation interpreted as the forward channel $U_{\{\ell\}} = \hat{U}_{\{\ell\}} + \Delta_\ell$, which we implement based on LDGM codes. Therefore, let us first describe the proposed Gaussian quantization scheme in a general setting.

Consider an i.i.d. Gaussian source $\hat{U} \sim \mathcal{N}(0, \sigma_{\hat{U}}^2)$ and an additive Gaussian noise channel $U = \hat{U} + \Delta$, where $\Delta \sim \mathcal{N}(0, \sigma^2)$ denotes the noise. Our goal is to construct an n-block quantizer of rate $R = I(\hat{U}; U) + \epsilon$ and codebook $\mathcal{C} = \{u_i^n : 1 \le i \le 2^{nR}\}$, to approximate this

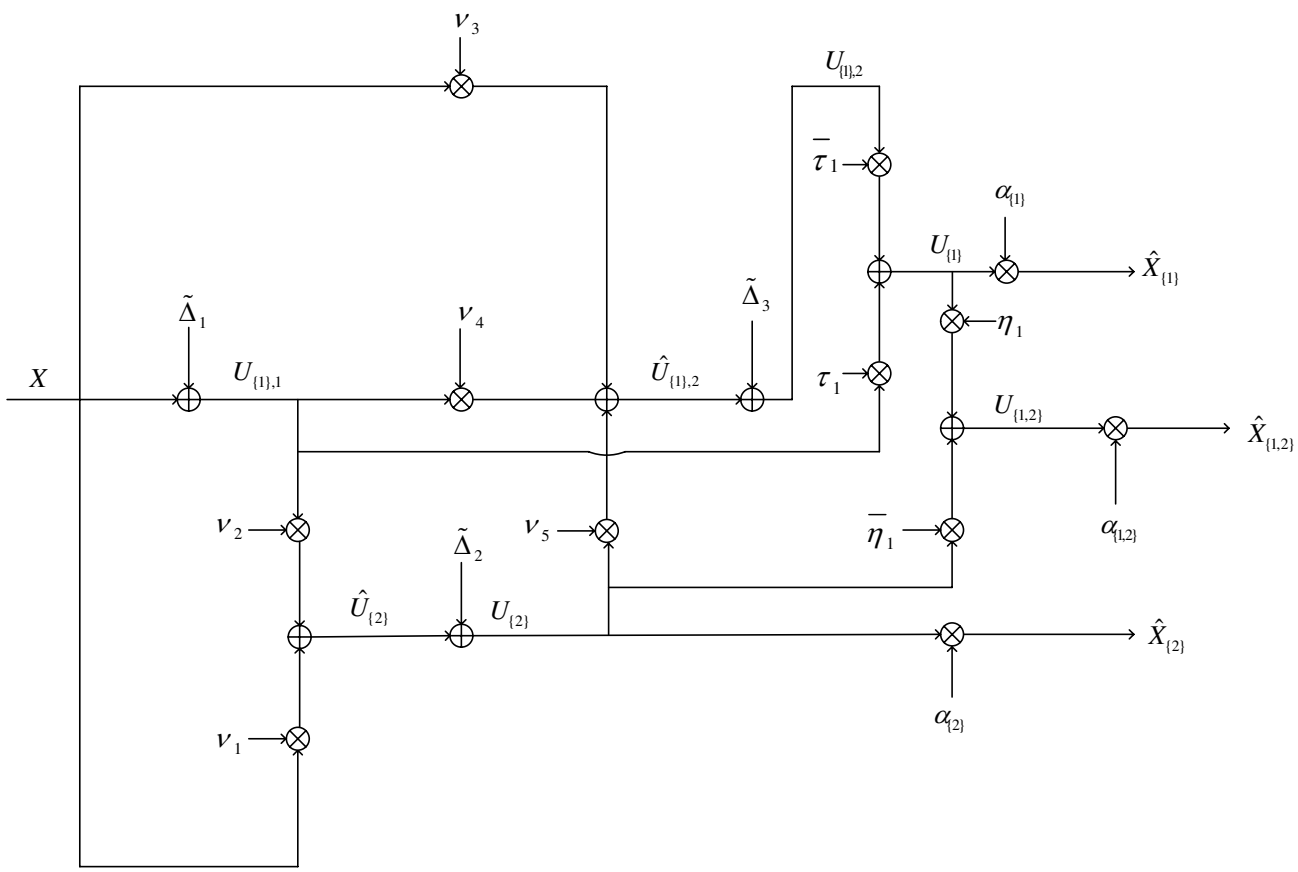

Fig. 2. Quantization splitting scheme for the Gaussian 2-description case.

Gaussian channel. To this end we need first to construct a finite random variable $\tilde{U}$ over some finite alphabet $\tilde{\mathcal{U}}$, to approximate U. An additional requirement on $\tilde{U}$, to be explained shortly, is that a positive integer ω exists such that $2^{\omega} p_{\tilde{U}}(\tilde{u})$ is a positive integer for every $\tilde{u} \in \tilde{\mathcal{U}}$. Next we design a multilevel LDGM code [27] to generate the codebook $\mathcal{C} \subseteq \tilde{\mathcal{U}}^n$, such that the marginal distribution of the codewords approximates $\tilde{U}$, and thus approximates U as well; finally, given an input sequence $\hat{u}^n = \hat{u}(1) \cdots \hat{u}(n)$, the quantizer output $u^n = u(1) \cdots u(n)$ is selected from $\mathcal{C}$ using a message passing algorithm.

The multilevel LDGM code is best described by means of its associated factor graph (Fig. 3). The factor graph is composed of n source nodes $\{S_1, \cdots, S_n\}$, m variable nodes $\{V_1, \cdots, V_m\}$, where $m = nR$, $n\omega$ check nodes $\{C_1, \cdots, C_{n\omega}\}$, and n network nodes $\{N_1, \cdots, N_n\}$. Every

variable node V_k is associated to information bit $v(k)$ and is connected by an edge to every check node C_h such that $G(h,k) = 1$, where G is some low-density generator matrix of dimension $n\omega \times m$ over $GF(2)$. Each network node N_l is connected by an edge to check nodes $C_l, C_{l+n}, \cdots, C_{l+(\omega-1)n}$, and to source node S_l. Source node S_l is associated to the l-th symbol $\hat{u}(l)$ in the input sequence, and N_l is associated to the l-th symbol $u(l)$ of the output sequence (i.e., codeword). Given an assignment of bit values to the variable nodes, the bit value $c(h)$ assigned to the check node C_h is computed as the modulo 2 addition of the values in neighboring variable nodes. The value $u(l)$ in the network node is computed via a mapping applied to the neighboring check nodes. In other words, $u(l) = \phi(c(l), c(n+l), \cdots, c(n(\omega-1)+l)$, where $\phi : \{0,1\}^\omega \to \tilde{\mathcal{U}}$ is chosen such that $\frac{|\phi^{-1}(\tilde{u})|}{2^\omega} = p_{\tilde{U}}(\tilde{u})$ for every $\tilde{u} \in \tilde{\mathcal{U}}$. Such a mapping is guaranteed to exist since $2^\omega p_{\tilde{U}}(\tilde{u})$, $\tilde{u} \in \tilde{\mathcal{U}}$, are positive integers. Finally, the codebook consists of all possible sequences u^n constructed at the network nodes for all possible assignments of bit values to the variable nodes.

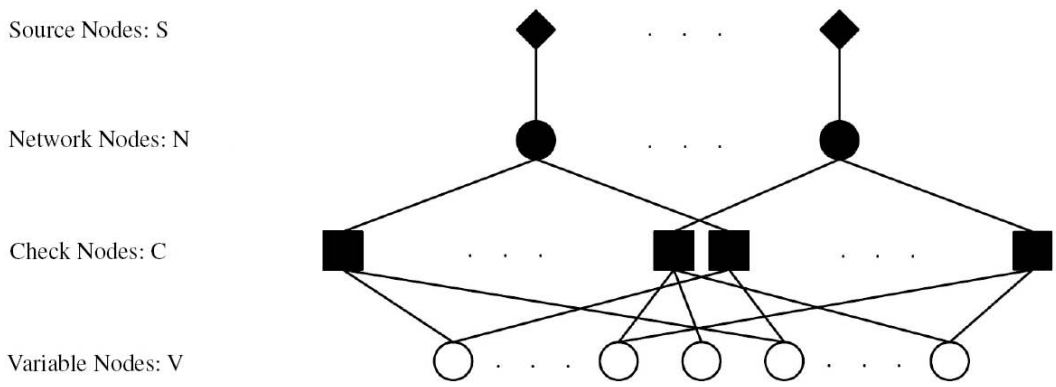

Fig. 3. Factor graph associated to the multilevel LDGM code.

Notice that the MMSE estimator of $\hat{U}$ given the variable U is $\mathbb{E}[\hat{U}|U] = \alpha U$, where $\alpha = \frac{\sigma^2_{\hat{U}}}{\sigma^2_{\hat{U}} + \sigma^2}$. Thus, if the quantizer output is u^n then the optimal source reconstruction is αu^n. Therefore, given the quantizer input sequence $\hat{u}^n$, we formulate the quantizer encoder problem as the problem of selecting the output sequence $u^n \in \mathcal{C}$ such that the mean squared error between $\hat{u}^n$ and αu^n to be minimized.

As a heuristic algorithm to solve the encoder problem we use the belief propagation with decimation. The algorithm is similar to the message passing algorithm described in [30], [31]. The algorithm description is presented in Fig. 5, and the equations to compute the messages in

Fig. 4. The notation $M_{A\to B}$ is for the message passed by node A to node B. All messages have two components, except for messages sent by the source nodes, which have $|\tilde{\mathcal{U}}|$ components, one for each symbol $\tilde{u} \in \tilde{\mathcal{U}}$. Finally, any check node, whose adjacent variable nodes are all fixed, computes the message to send, using:

$$M(b) = \frac{1}{\exp(\delta) + \exp(-\delta)}(c\exp((-1)^{1-b}\delta) + (1-c)\exp((-1)^b\delta)), \tag{19}$$

for $b \in \{0,1\}$, where c equals the fixed value of the check node.

$$\begin{aligned}
M_{S_l\to N_l}(\tilde{u}) =& \exp(-\lambda(\hat{u}(l) - \alpha\tilde{u})^2), \\
& \text{for all } \tilde{u} \in \tilde{\mathcal{U}}, 1 \le l \le n, \\
M_{N_l\to C_{l+sn}}(b) =& \sum_{\tilde{u}\in\tilde{\mathcal{U}}} M_{S_l\to N_l}(\tilde{u}) \sum_{\substack{b^\omega\in\{0,1\}^\omega \\ b(s+1)=b \\ \phi(b^\omega)=\tilde{u}}} \prod_{\substack{j=0 \\ j\neq s}}^{\omega-1} M_{C_{l+jn}\to N_l}(b(j)) \\
M_{C_{l+sn}\to N_l}(b) =& \frac{1}{2} + \frac{(-1)^b}{2} \prod_{k\in\mathcal{B}_v(l+sn)} (M_{V_k\to C_{l+sn}}(0) - M_{V_k\to C_{l+sn}}(1)) \\
& \text{for all } b \in \{0,1\}, 1 \le l \le n, 0 \le s \le \omega-1, \\
M_{C_{l+sn}\to V_k}(b) =& \frac{1}{2} + \frac{(-1)^b}{2}(M_{N_l\to C_{l+sn}}(0) - M_{N_l\to C_{l+sn}}(1)) \prod_{i\in\mathcal{B}_v(l+sn)\setminus\{k\}} (M_{V_i\to C_{l+sn}}(0) - M_{V_i\to C_{l+sn}}(1)) \\
M_{V_k\to C_{l+sn}}(b) =& \prod_{q\in\mathcal{A}_c(k)\setminus\{l+sn\}} M_{C_q\to V_k}(b) \\
& \text{for all } b \in \{0,1\}, k \in \mathcal{B}_v(l+sn), 1 \le l \le n, 0 \le s \le \omega-1.
\end{aligned}$$

Fig. 4. Message passing equations. After applying these equations, the components of each message are normalized to sum up to 1. $\mathcal{A}_c(k)$ denotes the set of indices q such that C_q is adjacent to node V_k, and $\mathcal{B}_v(q)$ denotes the set of of indices k such that V_k is adjacent to C_q.

B. A Note on Optimal $\tilde{U}$

The performance of aforedescribed scheme depends on how well the message passing algorithm solves the encoder problem, but also on the choice of random variable $\tilde{U}$. The choice

of $\tilde{U}$ is constrained by the requirement that a positive integer ω exists such that $2^{\omega}p_{\tilde{U}}(\tilde{u})$ is a positive integer for every $\tilde{u} \in \tilde{\mathcal{U}}$. Notice that this condition also constrains the size of the alphabet $\tilde{\mathcal{U}}$ to be at most 2^{ω}. Therefore it is interesting to consider the problem of optimizing $\tilde{U}$ subject to fixed ω. In order to formulate this problem we will disregard the dependence on the particular behavior of the message passing algorithm. Moreover, we will assume that the LDGM encoder approximates a theoretical random coding scheme[5]. Then the problem of optimizing $\tilde{U}$ is equivalent to optimizing the alphabet $\tilde{\mathcal{U}}$ and the conditional probability $p_{\tilde{U}|\hat{U}}$, formulated as follows

$$\min_{\tilde{\mathcal{U}}, p_{\tilde{U}|\hat{U}}} \qquad \mathbb{E}[(\hat{U} - \alpha\tilde{U})^2] \tag{20}$$

subject to $I(\hat{U};\tilde{U}) = I(\hat{U};U), \tilde{\mathcal{U}} \subseteq \mathbb{R}$, and $2^{\omega}p_{\tilde{U}}(\tilde{u}) \in \mathbb{N}$ for all $\tilde{u} \in \tilde{\mathcal{U}}$. This optimization problem has similar flavor to the problems considered in the context of alphabet constrained rate-distortion theory for continuous-valued sources in [35], [36], but appears to be more difficult due to the additional integer constraint. On the other hand, as a practical solution to LDGM code design, one can modify the requirements to fit the problems solved in [35], [36]. Specifically, one can replace first $\hat{U}$ by the output $\breve{U}$ of a fine scalar quantizer; then drop the last constraint in (20), impose instead the condition that $p_{\tilde{U}}(\tilde{u})$ are equal to some fixed values, and use the algorithm of [36] to determine the optimal alphabet $\tilde{\mathcal{U}}$. Alternatively, upon replacing $\hat{U}$ by $\breve{U}$, one can fix only the size of the alphabet $\tilde{\mathcal{U}}$ and determine the probabilities $p_{\tilde{U}}(\tilde{u})$ via the algorithm of [35]; then choose an integer ω such that $2^{\omega}p_{\tilde{U}}(\tilde{u})$ are close to some integer values.

We leave the quest for a solution algorithm to problem (20) and/or the investigation of the performance of the aforementioned strategies for future work. In our experiments we confine ourselves to a simple heuristic for the selection of $\tilde{U}$, inspired by the central limit theorem and by the intuition that $\tilde{U}$ has to be a good approximation of variable U. Let $W_1, \cdots, W_{\omega}$ be ω independent random variables, uniformly distributed over the alphabet $\{-1, 1\}$. Define $\bar{W}_{\omega} = \frac{\sum_{j=1}^{\omega} W_i \sqrt{\sigma_{\hat{U}}^2 + \sigma^2}}{\sqrt{\omega}}$. According to the central limit theorem, the sequence of random variables $\bar{W}_{\omega}$ converges to $\mathcal{N}(0, \sigma_{\hat{U}}^2 + \sigma^2)$ as $\omega \to \infty$. Therefore, we choose $\tilde{U} = \bar{W}_{\omega}$.

[5]Such an assumption is supported by the results of [27]. Although the argument in [27] is for discrete-valued sources and bounded distortion measures, it can be extended to cover the quadratic Gaussian case using standard techniques.

C. Successive Quantization

As discussed in Section II-A, the successive quantization scheme for vertices of the sum-rate optimal Gaussian EGC region follows the block diagram in Fig. 1. The operation $U_{\{\ell\}} = \hat{U}_{\{\ell\}} + \Delta_\ell$, at the ℓ-th step, $\ell = 1, \cdots, L$, is implemented using an n-block multilevel LDGM code as described in Section III-A, for $U = U_{\{\ell\}}$, $\hat{U} = \hat{U}_{\{\ell\}}$, and $\Delta = \Delta_\ell$. The input sequence $\hat{u}^n_{\{\ell\}} = \hat{u}_{\{\ell\}}(1), \cdots, \hat{u}_{\{\ell\}}(n)$ coincides with the source sequence $x^n = x(1) \cdots x(n)$ for $\ell = 1$, and for $\ell > 1$, it is computed based on the sequences $u^n_{\{\kappa\}}$ output at all previous stages $1 \leq \kappa \leq \ell - 1$, according to the recursive equations (10) and (11), applied symbol by symbol. The sequence $u^n_{\{\ell\}}$ output by the quantizer at stage ℓ, is found using the belief propagation algorithm described in Section III-A. The index i_ℓ formed out of the nR_ℓ information bits corresponding to the selected output is transmitted as the ℓ-th description.

The decoder corresponding to the ℓ-th description receives index i_ℓ and constructs the corresponding codeword $u^n_{\{\ell\}}$ using the factor graph for the ℓ-th stage LDGM code. The source reconstruction $\hat{x}^n_{\{\ell\}}$ is formed by $\hat{x}_{\{\ell\}}(l) = \alpha_{\{\ell\}} u_{\{\ell\}}(l)$, $1 \leq l \leq n$, according to (12).

Finally, the central decoder receives all indices $i_1, \cdots, i_L$, recovers $u^n_{\{1\}}, \cdots, u^n_{\{L\}}$, and based on them constructs the sequence $u^n_{\{1,\cdots,L\}}$ using (11) recursively. Then the source reconstruction is generated according to (13).

D. Quantization Splitting

The procedure of quantization splitting for the 2-description case is illustrated in the block diagram of Fig. 2. It consists of three successive n-block quantizers implemented using multilevel LDGM codes, as described in Section III-A. Specifically, the first quantizer models the forward channel $U_{\{1\},1} = X + \tilde{\Delta}_1$. Its input is the source sequence x^n and its output is denoted by $u^n_{\{1\},1}$. The second quantizer approximates the channel $U_{\{2\}} = \hat{U}_{\{2\}} + \tilde{\Delta}_2$. Its n-block input sequence $\hat{u}^n_{\{2\}}$ is constructed based on x^n and $u^n_{\{1\},1}$ according to (17) as shown in the block diagram. Its output is denoted by $u^n_{\{2\}}$. Finally, the third quantizer models the channel $U_{\{1\},2} = \hat{U}_{\{1\},2} + \tilde{\Delta}_3$, with input $\hat{u}^n_{\{1\},2}$ generated from x^n, $u^n_{\{1\},1}$, and $u^n_{\{2\}}$ using (18). Its output is denoted by $u^n_{\{1\},2}$.

Let i_1, i_2, i_3 denote the information bit sequences corresponding to the outputs of the three quantizers, respectively. Then indices i_1 and i_3 form the first description, while i_2 constitutes the second description. The decoder of the first description receives i_1 and i_3, recovers $u^n_{\{1\},1}$ and $u^n_{\{1\},2}$, based on which it generates $u^n_{\{1\}}$, and further $\hat{x}^n_{\{1\}}$ as the source reconstruction, using the

operations described in Fig. 2. The decoder of the second description receives index i_2, recovers $u^n_{\{2\}}$, and generates the source reconstruction $\hat{x}^n_{\{2\}}$ according to Fig. 2. When both descriptions are received at the decoder, the sequence $u^n_{\{1,2\}}$ is generated from $u^n_{\{1\}}$ and $u^n_{\{2\}}$, which is used to further generate the source reconstruction $\hat{x}^n_{\{1,2\}}$ as in Fig. 2.

E. Experimental Results

We have tested the proposed successive quantization and quantization splitting scheme for an i.i.d. zero-mean unit-variance Gaussian source. We have considered input sequences of various lengths $n = 100, 1000$ and 10000. In all our tests we set $\eta = 0.9$, and $\omega = 4$. The degree distributions of the low-density generator matrices used in our simulations are obtained from the website (http://lthcwww.epfl.ch.research/ldpcopt) or by implementing the algorithm in [32]. We use damping as in [25], [31] in our message passing algorithm, if the messages do not converge after 30 iterations.

Tables I and II present the simulation results of the LDGM-based successive quantization scheme for a 2-description symmetric and asymmetric distortion tuples, respectively. Tables III and IV exhibit the results of the proposed scheme for an L-description symmetric distortion tuple with $L = 3$ and $L = 4$, respectively. In each table $(d_{\{1\}}, \cdots, d_{\{L\}}, d_{\{1,\cdots,L\}})$ denotes the target distortion tuple, while $\hat{d}_{\{1\}}$, $\cdots$, $\hat{d}_{\{L\}}$, and $\hat{d}_{\{1,\cdots,L\}}$ denote the empirical distortions; $(R_1, \cdots, R_L)$ denotes the rate pair used in the experiments, which corresponds to a vertex of the sum-rate optimal Gaussian EGC region, defined by (5); λ_i and δ_i are for the parameters λ and δ, respectively, used in the message passing algorithm at the encoding stage i, $i = 1, \cdots, L$. Next we list the values of the parameters for $L = 4$, which did not fit in Table IV. For $n = 10000$, we have $\lambda_1 = \cdots = \lambda_4 = 1.7$ and $\delta_1 = \cdots = \delta_4 = 1.9$, while for $n = 1000$ and $n = 100$, we have $\lambda_1 = \lambda_2 = 1.6$, $\lambda_3 = \lambda_4 = 1.7$, $\delta_1 = \delta_2 = 1.8$ and $\delta_3 = \delta_4 = 1.9$.

Table V presents the result obtained using the proposed quantization splitting scheme for a 2-description problem with symmetric rates and symmetric distortions: $(d_{\{1\}}, d_{\{2\}}, d_{\{1,2\}})$ denotes the target distortion triple and $\hat{d}_{\{1\}}$, $\hat{d}_{\{2\}}$, $\hat{d}_{\{1,2\}}$ are the empirical distortions; $R_{1,1}, R_{1,2}, R_2$ represent the rates defined in (15), (16), and (14); $\lambda_{1,1}$, $\delta_{1,1}$ and $\lambda_{1,2}$, $\delta_{1,2}$ are the parameters of the LDGM codes used for the first description, in other words for encoder of stage 1 and encoder of stage 3, respectively; λ_2 and δ_2 are the parameters of the LDGM code used for the second description, i.e., for encoder of stage 2.

From the results showed in the tables, we can observe that the empirical distortions are very close to the theoretical distortion bounds.

Notice that the results in Tables III and IV show a tendency for the gap between the empirical and the target distortion to increase with the index of the description. This increase could be due to the error propagation from an encoding stage to the subsequent one in the successive coding scheme. Another possible reason could be the suboptimal choice of the LDGM code or of the parameters involved in the message passing algorithm. However, in the event that error propagation is the real cause, it is worth discussing methods to reduce its effect.

Some error propagation from an encoding stage to the subsequent one is expected to appear in any successive coding scheme, be it practical or theoretical (based on random codebooks), not only when using LDGM codes. In theoretical schemes, as $n \to \infty$, the error in each encoding stage, and thus, in each description, should go to 0. This suggests increasing n as a practical solution to keep the errors in all descriptions as well as the error in the central reconstruction below some desired value. The fact that indeed in practice the errors in all descriptions can go down as n increases, is validated in our experiments - see the results in Tables I-V. On the other hand, if increasing the value of n is not an option, then another practical solution to limit the errors at higher descriptions, is to slightly increase the rates at later encoding stages.

We would also like to mention that the "pure" successive coding scheme is applied only at the corner points of the rate region, while for rate tuples inside the dominant face of the rate region time-sharing of successive codes for corner points or quantization splitting have to be used. Both latter techniques lessen the effect of error propagation by distributing errors more evenly among descriptions. In particular, for the case of symmetric rates and symmetric distortion constraints, which is the most common case in practice and is also the most addressed case in the literature, the errors in all descriptions will be approximately equal as a consequence of equal time-sharing of the codes for corner points.

IV. Conclusion

We have proposed a practical Gaussian multiple description coding scheme based on LDGM codes for the scenario where the distortion constraints are only imposed on individual descriptions and the complete set of descriptions. The effectiveness of the proposed scheme is verified through simulation. It is worth noting for the more general scenario studied in [33], [34], the Gaussian

n	(R_1, R_2)	$(d_{\{1\}}, d_{\{2\}}, d_{\{1,2\}})$	$\hat{d}_{\{1\}}$	$\hat{d}_{\{2\}}$	$\hat{d}_{\{1,2\}}$	λ_1	λ_2	δ_1	δ_2
10000	$(1, 1.015)$	$(0.25, 0.25, 0.125)$	0.267	0.262	0.135	1.7	1.7	1.9	1.9
1000	$(1, 1.015)$	$(0.25, 0.25, 0.125)$	0.267	0.264	0.136	1.6	1.6	1.8	1.8
100	$(1, 1.015)$	$(0.25, 0.25, 0.125)$	0.270	0.272	0.148	1.6	1.6	1.8	1.8

TABLE I
PARAMETERS AND RESULTS FOR A 2-DESCRIPTION CASE WITH SYMMETRIC DISTORTIONS.

n	(R_1, R_2)	$(d_{\{1\}}, d_{\{2\}}, d_{\{1,2\}})$	$\hat{d}_{\{1\}}$	$\hat{d}_{\{2\}}$	$\hat{d}_{\{1,2\}}$	λ_1	λ_2	δ_1	δ_2
10000	$(1.161, 0.914)$	$(0.2, 0.25, 0.1)$	0.211	0.266	0.111	2.6	1.7	2.8	1.9
10000	$(1, 1.075)$	$(0.25, 0.2, 0.1)$	0.262	0.219	0.111	1.7	2.5	1.9	2.7
1000	$(1.161, 0.914)$	$(0.2, 0.25, 0.1)$	0.212	0.269	0.113	2.8	1.7	2.8	1.9
1000	$(1, 1.075)$	$(0.25, 0.2, 0.1)$	0.265	0.220	0.114	1.8	2.6	1.9	2.8
100	$(1.161, 0.914)$	$(0.2, 0.25, 0.1)$	0.219	0.277	0.121	2.8	1.7	2.8	1.8
100	$(1, 1.075)$	$(0.25, 0.2, 0.1)$	0.271	0.225	0.122	1.8	2.6	1.9	2.8

TABLE II
PARAMETERS AND RESULTS FOR A 2-DESCRIPTION CASE WITH ASYMMETRIC DISTORTIONS.

n	(R_1, R_2, R_3)	$(d_{\{1\}}, d_{\{2\}}, d_{\{3\}}, d_{\{1,2,3\}})$	$\hat{d}_{\{1\}}$	$\hat{d}_{\{2\}}$	$\hat{d}_{\{3\}}$	$\hat{d}_{\{1,2,3\}}$	λ_1	λ_2	λ_3	δ_1	δ_2	δ_3
10000	$(1.161, 1.165, 1.169)$	$(0.2, 0.2, 0.2, 0.067)$	0.210	0.210	0.212	0.076	1.7	1.7	1.7	1.9	1.9	1.9
1000	$(1.161, 1.165, 1.169)$	$(0.2, 0.2, 0.2, 0.067)$	0.212	0.214	0.217	0.079	1.6	1.6	1.7	1.8	1.9	1.9
100	$(1.161, 1.165, 1.169)$	$(0.2, 0.2, 0.2, 0.067)$	0.219	0.222	0.226	0.088	1.6	1.6	1.7	1.8	1.9	1.9

TABLE III
PARAMETERS AND RESULTS FOR A 3-DESCRIPTION CASE WITH SYMMETRIC DISTORTIONS.

n	(R_1, R_2, R_3, R_4)	$(d_{\{1\}}, d_{\{2\}}, d_{\{3\}}, d_{\{4\}}, d_{\{1,2,3,4\}})$	$\hat{d}_{\{1\}}$	$\hat{d}_{\{2\}}$	$\hat{d}_{\{3\}}$	$\hat{d}_{\{4\}}$	$\hat{d}_{\{1,2,3,4\}}$
10000	$(1.161, 1.163, 1.165, 1.168)$	$(0.2, 0.2, 0.2, 0.2, 0.05)$	0.209	0.209	0.211	0.213	0.059
1000	$(1.161, 1.163, 1.165, 1.168)$	$(0.2, 0.2, 0.2, 0.2, 0.05)$	0.210	0.211	0.213	0.214	0.061
100	$(1.161, 1.163, 1.165, 1.168)$	$(0.2, 0.2, 0.2, 0.2, 0.05)$	0.218	0.219	0.222	0.227	0.075

TABLE IV
RESULTS FOR A 4-DESCRIPTION CASE WITH SYMMETRIC DISTORTIONS.

n	$(R_{1,1}, R_{1,2}, R_2)$	$(d_{\{1\}}, d_{\{2\}}, d_{\{1,2\}})$	$\hat{d}_{\{1\}}$	$\hat{d}_{\{2\}}$	$\hat{d}_{\{1,2\}}$	$\lambda_{1,1}$	$\lambda_{1,2}$	λ_2	$\delta_{1,1}$	$\delta_{1,2}$	δ_2
10000	$(0.661, 0.346, 1.007)$	$(0.25, 0.25, 0.125)$	0.268	0.269	0.138	1.0	0.3	1.5	1.2	0.6	1.7
1000	$(0.661, 0.346, 1.007)$	$(0.25, 0.25, 0.125)$	0.272	0.271	0.141	1.0	0.3	1.5	1.2	0.6	1.7
100	$(0.661, 0.346, 1.007)$	$(0.25, 0.25, 0.125)$	0.278	0.276	0.147	1.0	0.3	1.5	1.2	0.6	1.7

TABLE V
PARAMETERS AND RESULTS FOR A 2-DESCRIPTION CASE WITH SYMMETRIC RATES AND SYMMETRIC DISTORTIONS, USING QUANTIZATION SPLITTING.

multiple description problem has both a source coding aspect (in the form of quantization) and a channel coding aspect (in the form of binning). In a certain sense, only the source coding aspect is addressed in the present work. Therefore, it is of considerable interest to develop more advanced multiple description techniques by incorporating practical binning methods into the proposed scheme.

REFERENCES

[1] A. A. El Gamal and T. M. Cover, "Achievable rates for multiple descriptions," *IEEE Trans. Inf. 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. Venkataramani, G. Kramer, and V. K. Goyal, "Multiple description coding with many channels," *IEEE Trans. Inf. Theory*, vol. 49, pp. 2106-2114, Sep. 2003.

[4] H. Wang and P. Viswanath, "Vector gaussian multiple description with individual and central receivers," *IEEE Trans. Inf. Theory*, vol. 53, pp. 2133-2153, Jun. 2007.

[5] J. Chen, "Rate region of Gaussian multiple description coding with individual and central distortion constraints," *IEEE Trans. Inf. Theory*, vol. 55, pp. 3991-4005, Sep. 2009.

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

[7] V. A. Vaishampayan and J. Domaszewicz, "Design of entropy-constrained multiple-description scalar quantizers," *IEEE Trans. Inf. Theory*, vol. 40, pp. 245-250, Jan. 1994.

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

[9] C. Tian and S. S. Hemami, "Universal multiple description scalar quantization: Analysis and design," *IEEE Trans. Inf. Theory*, vol. 50, pp. 2089-2102, Sep. 2004.

[10] S. Dumitrescu and X. Wu, "Optimal two-description scalar quantizer design," *Algorithmica*, vol. 41, no. 4, pp. 269-287, Feb. 2005.

[11] S. Dumitrescu and X. Wu, "Lagrangian optimization of two-description scalar quantizers," *IEEE Trans. Inf. Theory*, vol. 53, pp. 3990-4012, Nov. 2007.

[12] D. Muresan and M. Effros, "Quantization as histogram segmentation: optimal scalar quantizer design in network systems," *IEEE Trans. Inf. Theory*, vol. 54, pp. 344-366, Jan. 2008.

[13] H. Jafarkhani and V. Tarokh, "Multiple description trellis-coded quantization," *IEEE Trans. Commun.*, vol. 47, pp. 799-803, Jun. 1999.

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

[15] S. N. Diggavi, N. J. A. Sloane, and V. A. Vaishampayan, "Asymmetric multiple description lattice vector quantizers", *IEEE Trans. Inf. Theory*, vol. 48, pp. 174-191, Jan. 2002.

[16] Y. Frank-Dayan and R. Zamir, "Dithered lattice-based quantizers for multiple descriptions," *IEEE Trans. Inform. Theory*, vol. 48, pp. 192-204, Jan. 2002.

[17] X. Huang and X. Wu, "Optimal index assignment for multiple description lattice vector quantization," *Proc. of DCC 2006*, pp. 272-281, 2006.

[18] J. Østergaard, J. Jensen, and R. Heusdens, "n-channel entropy-constrained multiple-description lattice vector quantization," in *IEEE Trans. Inf. Theory*, pp. 1956-1973, May 2006.

[19] J. Chen, C. Tian, T. Berger, and S. S. Hemami, "Multiple description quantization via Gram-Schmidt orthogonalization," *IEEE Trans. Inf. Theory*, vol. 52, pp. 5197-5217, Dec. 2006.

[20] Y. Wang, M. T. Orchard, V. A. Vaishampayan, and A. R. Reibman, "Multiple description coding using pairwise correlating transforms," *IEEE Trans. on Image Processing*, vol. 10, pp. 351-366, Mar. 2001.

[21] V. K. Goyal and J. Kovačević, "Generalized multiple description coding with correlating transforms," *IEEE Trans. on Inf. Theory*, vol. 47, pp. 2199-2224, Sep. 2001.

[22] J. Østergaard and R. Zamir, "Multiple descriptions by dithered Delta-Sigma quantization," *IEEE Trans. Inf. Theory*, vol. 55, pp. 4661-4675, Oct. 2009.

[23] R. Zamir, "Gaussian codes and Shannon bounds for multiple descriptions," *IEEE Trans. on Inf. Theory*, vol. 45, pp. 2629-2636, Nov. 1999.

[24] E. Martinian and J. S. Yedidia, "Iterative quantization using codes on graphs," in *Proc. Allerton Conf. Commun. Contr. Computing*, Oct. 2003.

[25] T. Filler and J. Friedrich, "Binary quantization using belief propagation with decimation over factor graphs of LDGM codes," in *Proc. Allerton Conf. Commun. Contr. Computing*, Sep. 2007.

[26] M. J. Wainwright, E. Maneva, and E. Martinian, "Lossy source compression using low-density generator matrix codes: Analysis and algorithms," *IEEE Trans. Inf. Theory*, vol. 56, pp. 1351-1368, Mar. 2010.

[27] Z. Sun, M. Shao, J. Chen, K. M. Wong, and X. Wu, "Achieving the rate-distortion bound with low-density generator matrix codes," *IEEE Trans. on Commun.*, vol. 58, pp. 1643-1653, Jun. 2010.

[28] X. Zhang, J. Chen, S. B. Wicker, and T. Berger, "Successive coding in multiuser information theory," *IEEE Trans. on Inf. Theory*, vol. 53, pp. 2246-2254, Jun. 2007.

[29] J. Wang, J. Chen, L. Zhao, P. Cuff, and H. Permuter, "On the role of the refinement layer in multiple description coding and scalable cding," *IEEE Trans. on Inf. Theory*, vol. 57, pp. 1443-1456, Mar. 2011.

[30] J. Chen, S. Dumitrescu, Y. Zhang, J. Wang, "Robust multiresolution coding," *IEEE Trans. on Commun.*, vol. 58, pp. 3186-3195, Nov. 2010.

[31] Y. Zhang, S. Dumitrescu, J. Chen, and Z. Sun, "LDGM-based multiple description coding with Hamming distortion measure," submitted for publication.

[32] S.-Y. Chung, G. D. Forney, T. J. Richardson and R. Urbanke, "On the design of low-desity parity check codes within 0.0045 dB of the Shannon limit," *IEEE Communications Letters*, vol. 5, pp. 58-60, Feb. 2001.

[33] S. S. Pradhan, R. Puri, and K. Ramchandran, "n-channel symmetric multiple descriptions—Part I: (n, k) source-channel erasure codes," *IEEE Trans. Inf. Theory*, vol. 50, pp. 47-61, Jan. 2004.

[34] R. Puri, S. S. Pradhan, and K. Ramchandran, "n-channel symmetric multiple descriptions—Part II: An achievable rate-distortion region," *IEEE Trans. Inf. Theory*, vol. 51, pp. 1377-1392, Apr. 2005.

[35] W. A, Finamore, and W. A. Pearlman, "Optimal encoding of discrete-time continuous-amplitude memoryless sources with finite output alphabets", *IEEE Trans. Inf. Theory*, vol. 26, pp. 144-155, Mar. 1980.

[36] W. A. Pearlman, and A. Chekima, "Source coding bounds using quantizer reproduction levels", *IEEE Trans. Inf. Theory*, vol. 30, pp. 559-567, May 1984.

$max_{iter} \leftarrow 100$, $\eta \leftarrow 0.9$
$\mathcal{A}_c(k) \leftarrow \{q|G(q,k)=1\}$, for all $k = 1, \cdots, m$
$\mathcal{B}_v(q) \leftarrow \{k|G(q,k)=1\}$, for all $q = 1, \cdots, n\omega$
$\mathcal{NV} \leftarrow \{1, \cdots, m\}$
for $l = 1$ to n **do**
  Compute $M_{S_l \to N_l}$ as in Fig. 4
  for $s = 0$ to $\omega - 1$ **do**
    $M_{C_{l+sn} \to N_l} \leftarrow (\frac{1}{2}, \frac{1}{2})$
    for $k \in \mathcal{B}_v(l+sn)$ **do**
      $M_{C_{l+sn} \to V_k} \leftarrow (\frac{1}{2}, \frac{1}{2})$
    end for
  end for
end for
while $\mathcal{NV} \neq \emptyset$ **do**
  for $i = 1$ to max_{iter} **do**
    Compute $M_{N_l \to C_{l+sn}}$, $M_{V_k \to C_{l+sn}}$ as in Fig. 4
    for all $l = 1, \cdots, n$, $s = 0, \cdots, \omega - 1$, $k \in \mathcal{B}_v(l+sn) \cap \mathcal{NV}$.
    for $l = 1$ to n **do**
      for $s = 0$ to $\omega - 1$ **do**
        if $\mathcal{B}_v(l+sn) \cap \mathcal{NV} \neq \emptyset$ **then**
          Compute $M_{C_{l+sn} \to N_l}$ as in Fig. 4
          for $k \in \mathcal{B}_v(l+sn) \cap \mathcal{NV}$ **do**
            Compute $M_{C_{l+sn} \to V_k}$ as in Fig. 4
          end for
        else
          Compute $M_{C_{l+sn} \to N_l}$ as in Eq. (19)
        end if
      end for
    end for
    if $M_{C_q \to V_k}$ converge for all $q = 1, \cdots, n\omega$, $k \in \mathcal{B}_v(q) \cap \mathcal{NV}$ **then**
      Break
    end if
  end for
  for $k \in \mathcal{NV}$ **do**
    $p(V_k = b) \leftarrow \prod_{q \in \mathcal{A}_c(k)} M_{C_q \to V_k}(b)$, for $b \in \{0,1\}$
    $p(V_k = b) \leftarrow \frac{p(V_k=b)}{p(V_k=0)+p(V_k=1)}$, for $b \in \{0,1\}$
    if $|p(V_k = 0) - p(V_k = 1)| > \eta$ **then**
      $\mathcal{NV} \leftarrow \mathcal{NV} - \{k\}$
      $\hat{b} \leftarrow \text{argmax}_{b \in \{0,1\}} p(V_k = b)$, $V_k \leftarrow \hat{b}$
      $M_{V_k \to C_q}(\hat{b}) \leftarrow 1$, $M_{V_k \to C_q}(1 - \hat{b}) \leftarrow 0$, for all $q \in \mathcal{A}_c(k)$
    end if
  end for
end while

Fig. 5. Message passing algorithm. $\mathcal{NV}$ denotes the set of indices k of currently non-decimated variable nodes V_k.