# LDGM-based Multiple Description Coding for Finite Alphabet Sources

Ying Zhang, Sorina Dumitrescu, *Member, IEEE*, Jun Chen, *Member, IEEE*, Zhibin Sun

*Abstract*—This work presents an LDGM-based practical successive coding scheme for the multiple description (MD) problem for finite alphabet sources. The scheme, which targets the Zhang-Berger (ZB) rate-distortion region, is shown to be asymptotically optimal with joint typicality encoding, while as a practical encoding solution a message passing algorithm is adopted.

We further discuss in more detail the application of the coding scheme in three cases of the MD problem with the Hamming distortion measure: 1) no excess sum-rate for binary sources, 2) successive refinement, and 3) no excess marginal rate for the uniform binary source. In the no excess sum-rate case some progress is made in the characterization of fundamental limits by deriving the analytical expression of the distortion region for general binary sources, and of the auxiliary variables needed to achieve its boundary. The exact expression of the Zhang-Berger upper bound to the central distortion is also provided for the case of no excess marginal rate for the uniform binary source.

The proposed LDGM-based coding scheme is tested in practice for all three aforementioned cases. The experimental results show very good performance, demonstrating its ability to approach the theoretical rate-distortion limits or the available upper bounds.

*Index Terms*—Lossy source coding, low-density generator matrix code, message-passing algorithm, multiple description coding.

## I. INTRODUCTION

In the multiple description (MD) problem, a source sequence is compressed into two descriptions, which are constructed in such a way that an adequate reconstruction of the source is possible based on each description while the two descriptions together can lead to better reconstruction quality. Information theoretical results on this problem can be found in [1]–[7].

Research on practical MD coding has intensified in recent years due to its applications in modern communications systems [8]–[12]. However, the design of practical schemes able to approach the known fundamental limits for finite alphabet sources has received little attention. For such sources, practical codes based on low-density generator matrices (LDGM),

The authors are with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, L8S 4K1, Canada (email: yingzhang@grads.ece.mcmaster.ca, sorina@mail.ece.mcmaster.ca, junchen@mail.ece.mcmaster.ca, sunz4@mcmaster.ca).

combined with various message passing algorithms, have been successful in approaching the rate-distortion bound for single description source coding. In [13], [14], LDGM codes are used to encode a uniform binary source, while in [17] a multilevel LDGM coding scheme for finite alphabet sources with general distortion measure is proposed.

Motivated by the success of LDGM codes for the single description problem we attempt to apply them to the MD case. However, for finite alphabet sources with general distortion measure the characterization of the MD rate-distortion region is not known. Therefore, we turn our attention to the available inner bounds of the MD region. The first general inner bound, referred to as the EGC region, was provided by El Gamal and Cover in [3]. This was shown to be tight for the no excess sum-rate case by Ahlswede [4]. Later Zhang and Berger [5] proposed a different inner bound, termed the ZB region, and showed that it contains points not included in the EGC region. Recently, Wang *et al.* [7] proved that the ZB region includes the EGC region and established the ZB region as the best inner bound known to date. Motivated by the latter result we design an LDGM-based coding scheme for this region. The scheme consists of three sequential encoders and is applicable to corner points of a certain rate region derived from the ZB region, while any other rate pair can be obtained through timesharing of two such corner points. We prove that the coding scheme is asymptotically optimal when the joint typicality rule is employed at each encoder, whereas as a practical solution for the encoding problem we propose a message passing algorithm.

The proposed coding scheme needs knowledge of the auxiliary variables involved in the ZB region. In order to determine the optimal auxiliary variables achieving rate-distortion points on the boundary of the ZB region, numerical techniques can be used to solve the underlying non-convex optimization problem. On the other hand, availability of analytical solutions would spare this step and thus simplify the design of the coding system. Therefore, next we turn our attention to several cases of the MD problem with Hamming distortion measure where such optimal auxiliary variables can be computed analytically.

The first case is the no excess sum-rate case for binary sources. In this case the characterization of the rate-distortion MD region is available, as the EGC region, however, the computation of the rate-distortion limits is not simple and analytical expressions are only partly available for the uniform binary source. Interestingly, it turns out that the the distortion region and the auxiliary variables achieving their boundary can be computed analytically for general binary sources, and we derive its exact expression.

Further, we consider the case of no excess marginal rate for the uniform binary source. For this case Zhang and Berger proposed in [6] a method to compute an upper bound for the central distortion given the values of the side distortions. We improve Zhang and Berger's result by providing the analytical expression of this upper bound and of the achieving auxiliary variables.

Another special case of the MD problem, where the fundamental limits are known, is the problem of successive refinement (SR) for finite alphabet sources [18]–[20]. In this case the second description is of no interest alone, but only in conjunction with the first description. We show how the proposed LDGM-based coding scheme can be adapted to the SR problem.

Finally, we test the proposed coding scheme in each of the aforementioned cases of the MD problem. Our simulations show very good performance demonstrating thus its ability to approach in practice the theoretical rate-distortion limits or the available upper bounds.

We mention that LDGM-based successive coding schemes for the MD problem were also presented in [21], [22]. The scheme in [22] is specifically designed for a Gaussian source with squared-error distortion measure, leveraging on particularities of the MD problem conferred by the specific source and distortion measure, and cannot be extended to the case of finite alphabet sources with general distortion. On the other hand, the MD scheme of [21] is tailored for the EGC region, which can be strictly smaller than the ZB region, targeted in our work. For instance, Zhang and Berger's upper bound on the central distortion, in the no excess marginal rate case for the uniform binary source, cannot be achieved using the MD code of [21]. Additionally, beside the differences in the coding scheme, another contribution of the current work versus [21], [22] consists in the analytical results emphasized in previous paragraphs.

The paper is organized as follows. Section II reviews the MD problem and the ZB region. The following section presents the proposed sequential coding scheme for the ZB region. Section IV addresses the no excess sum-rate case for binary sources and derives the analytical expression of the distortion region and boundary achieving auxiliary variables. Section V discusses the upper bound on the central distortion in the no excess marginal rate case for the uniform binary source. Section VI reviews the SR rate-distortion region and clarifies the application of the proposed coding scheme in this case. Finally, Section VII presents our simulation results and Section VIII concludes the paper.

## II. THE MD PROBLEM

The formal definition of the MD problem is given as follows. Consider a sequence $X^n = (X(1), X(2), \cdots, X(n))$ drawn from an i.i.d. source $X$ with generic distribution $p_X$. Let $d : \mathcal{X} \times \hat{\mathcal{X}} \to [0, \infty)$ be a distortion measure, where $\mathcal{X}$ and $\hat{\mathcal{X}}$ are the source alphabet and the reconstruction alphabet, respectively. Throughout this work, the alphabets $\mathcal{X}$ and $\hat{\mathcal{X}}$ are assumed to be finite. The quintuple $(R_1, R_2, d_1, d_2, d_0)$ is said achievable, if for all sufficiently large $n$, there exist encoding functions

$$f_t^{(n)} : \mathcal{X}^n \to \{1, 2, \cdots, \lfloor 2^{nR_t} \rfloor\}, \quad t = 1, 2,$$

and decoding functions

$$g_t^{(n)} : \{1, 2, \cdots, \lfloor 2^{nR_t} \rfloor\} \to \hat{\mathcal{X}}^n, \quad t = 1, 2,$$
$$g_0^{(n)} : \{1, 2, \cdots, \lfloor 2^{nR_1} \rfloor\} \times \{1, 2, \cdots, \lfloor 2^{nR_2} \rfloor\} \to \hat{\mathcal{X}}^n,$$

such that

$$\mathbb{E}\left[\frac{1}{n}\sum_{l=1}^{n} d(X(l), \hat{X}_t(l))\right] \leq d_t, \quad t = 0, 1, 2,$$

where

$$\hat{X}_t^n = g_t^{(n)}(f_t^{(n)}(X^n)), \quad t = 1, 2,$$
$$\hat{X}_0^n = g_0^{(n)}(f_1^{(n)}(X^n), f_2^{(n)}(X^n)).$$

The MD rate-distortion region, denoted by $\mathcal{RD}$, is the closure of the set of all achievable quintuples $(R_1, R_2, d_1, d_2, d_0)$.

Zhang and Berger [5] proposed an inner bound of the MD rate-distortion region termed the ZB region. The ZB region, denoted by $\mathcal{RD}_{\mathrm{ZB}}$, is the set of quintuples $(R_1, R_2, d_1, d_2, d_0)$ such that there exist random variables $X_c$, $X_1$, $X_2$ jointly distributed with $X$ and functions $\psi_t$, $t = 0, 1, 2$, satisfying

$$R_t \geq I(X; X_c, X_t), \quad t = 1, 2, \tag{1}$$
$$R_1 + R_2 \geq 2I(X; X_c) + I(X; X_1, X_2 | X_c)$$
$$\qquad + I(X_1; X_2 | X_c), \tag{2}$$
$$\mathbb{E}d(X, \psi_t(X_c, X_t)) \leq d_t, \quad t = 1, 2, \tag{3}$$
$$\mathbb{E}d(X, \psi_0(X_c, X_1, X_2)) \leq d_0. \tag{4}$$

The question whether $\mathcal{RD}_{\mathrm{ZB}}$ is tight is an open problem. However, $\mathcal{RD}_{\mathrm{ZB}}$ is the best inner bound known so far for $\mathcal{RD}$. Therefore, in the next section we propose a practical coding scheme based on LDGM codes for the ZB region.

## III. AN LDGM-BASED CODING SCHEME FOR THE CORNER RATE POINTS ASSOCIATED TO THE ZB REGION

### A. Sequential Random Coding Scheme

Given the auxiliary random variables $X_c, X_1, X_2$ jointly distributed with $X$, the distortion triple $(d_1, d_2, d_0)$ obeying (3) and (4), can be achieved by the set of rate pairs $(R_1, R_2)$ satisfying relations (1) and (2). Let us denote by $\mathcal{R}(p_{X_c X_1 X_2 | X})$ this rate region, and by $\mathcal{F}(p_{X_c X_1 X_2 | X})$ its dominant face, which is defined as the set of rate pairs $(R_1, R_2) \in \mathcal{R}(p_{X_c X_1 X_2 | X})$ for which (2) holds with equality. Without loss of generality, we shall focus on $\mathcal{F}(p_{X_c X_1 X_2 | X})$ since every point in $\mathcal{R}(p_{X_c X_1 X_2 | X})$ is dominated, in a componentwise sense, by some point $(R_1, R_2) \in \mathcal{F}(p_{X_c X_1 X_2 | X})$. Moreover, the two corner points $E_1$ and $E_2$ of $\mathcal{R}(p_{X_c X_1 X_2 | X})$ (i.e., the end points of $\mathcal{F}(p_{X_c X_1 X_2 | X})$) are of particular importance. Specifically, one can achieve an arbitrary point on the dominant face by timesharing $E_1$ and $E_2$. Precisely, the coordinates of $E_1$ are

$$R_1(E_1) = I(X; X_c) + I(X; X_1 | X_c), \tag{5}$$
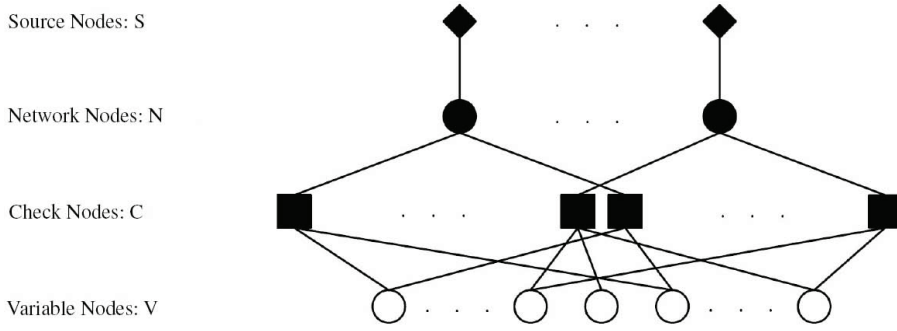$$R_2(E_1) = I(X; X_c) + I(X, X_1; X_2 | X_c).$$

Fig. 1. Factor graph associated to the multilevel LDGM code.

The coordinates of $E_2$ are obtained by swapping indices 1 and 2 in the above relations.

The successive encoding scheme directly suggested by (5) needs conditional codebooks, which complicate the implementation. To avoid this we use a variable substitution as follows. According to [7, Lemma 1], there exist random variables $U_t{}^1$, $t = 1, 2$, taking values in the finite sets $\mathcal{U}_t$, respectively, with $|\mathcal{U}_t| \leq |\mathcal{X}_c|(|\mathcal{X}_t| - 1) + 1$, $t = 1, 2$, and functions $\pi_t : \mathcal{X}_c \times \mathcal{U}_t \to \mathcal{X}_t$, $t = 1, 2$, such that

C1) $U_t$ is independent of $X_c$;
C2) $X_t = \pi_t(X_c, U_t)$;
C3) $(X, X_2)-(X_c, X_1)-U_1$ and $(X, X_1, U_1)-(X_c, X_2)-U_2$ form Markov chains.

The above relations imply that

$$I(X, X_c; U_1) = I(X_c; U_1) + I(X; U_1|X_c) = I(X; X_1|X_c),$$
$$I(X, X_c, X_1; U_2) = I(X_c; U_2) + I(X, X_1; U_2|X_c)$$
$$= I(X, X_1; X_2|X_c).$$

Hence, (5) can be written equivalently as

$$R_1(E_1) = I(X; X_c) + I(X, X_c; U_1),$$
$$R_2(E_1) = I(X; X_c) + I(X, X_c, X_1; U_2).$$

Next we present a sequential coding scheme approaching the corner point $E_1$. This system consists of three encoders, labeled $0, 1$ and $2$.

**Codebook generation.** Encoder 0 randomly generates a codebook $\mathcal{C}_0 = \{x_{c,i}^n\}_{i=1}^{2^{n(I(X;X_c)+\epsilon)}}$ according to the distribution $\prod_{l=1}^n p_{X_c}(\cdot)$. Encoder 1 randomly generates a codebook $\mathcal{C}_1 = \{u_{1,j}^n\}_{j=1}^{n(I(X,X_c;U_1)+\epsilon)}$, according to the distribution $\prod_{l=1}^n p_{U_1}(\cdot)$. Encoder 2 randomly generates a codebook $\mathcal{C}_2 = \{u_{2,k}^n\}_{k=1}^{n(I(X,X_c,X_1;U_2)+\epsilon)}$, according to the distribution $\prod_{l=1}^n p_{U_2}(\cdot)$.

**Encoding.** Given the source sequence $x^n$, encoder 0 selects a codeword $x_{c,i}^n$ which is jointly strongly typical with $x^n$. Encoder 1 finds an index $j$ such that $u_{1,j}^n$ is jointly strongly typical with $x^n$ and $x_{c,i}^n$. Finally, encoder 2 chooses an index $k$ such that $u_{2,k}^n$ is jointly strongly typical with $x^n$, $x_{c,i}^n$, and $x_{1,i,j}^n$, where $x_{1,i,j}^n = (\pi_1(x_{c,i}(l), u_{1,j}(l)))_{l=1}^n$. The pairs $(i, j)$ and $(i, k)$ form descriptions 1 and 2, respectively.

**Decoding.** Decoder 1 receives index pair $(i, j)$ and

takes $g_1^{(n)}(i, j) = (\psi_1(x_{1,i,j}^n(l)))_{l=1}^n$ as the reconstruction. Decoder 2 receives index pair $(i, k)$ and reconstructs the source as $g_2^{(n)}(i, k) = (\psi_2(x_{2,i,k}^n(l)))_{l=1}^n$, where $x_{2,i,k}^n(l) = \pi_2(x_{c,i}(l), u_{2,k}(l))$. The central decoder obtains the reconstruction as $g_0^{(n)}(i, j, i, k) = (\psi_0(x_{c,i}(l), x_{1,i,j}^n(l), x_{2,i,k}^n(l)))_{l=1}^n$.

Following the above sequential coding system we propose an LDGM-based coding scheme for the corner rate points $E_1$ and $E_2$, which uses a multilevel LDGM code to generate the codebook at each encoding stage. Further, we show that when the joint typicality rule is used at each encoder the proposed scheme achieves optimality asymptotically. Finally, we present the message passing algorithm used as a practical heuristic solution for the encoding problem at each stage.

### B. Asymptotically Optimal LDGM-based Code

A multilevel LDGM code of rate $R = \frac{m}{n}$ is specified by a low density generator matrix $G$ of dimension $m \times n\omega$ over $\mathbb{F}_2$ and a mapping $\phi : \mathbb{F}_2^\omega \to \mathcal{Z}$, for some positive integer $\omega$ and finite set $\mathcal{Z}$. Define now the mapping $\Phi : \mathbb{F}_2^{n\omega} \to \mathcal{Z}^n$ such that for any $c^{n\omega} \in \mathbb{F}_2^{n\omega}$, the $l$-th component of $\Phi(c^{n\omega})$ is $\phi(c(l), c(n+l), \cdots, c(n(\omega-1)+l))$, for $1 \leq l \leq n$. Then the multilevel LDGM code specified by $G$ and $\phi$ is defined as the set

$$\mathcal{C}(G, \phi) \triangleq \{\Phi(v^m G)|v^m \in \mathbb{F}_2^m\},$$

where $v^m G$ denotes the product between the row vector $v^m$ and the matrix $G$.

Consider now the joint probability distribution $p_{Y, \phi(\tilde{Y})}$, where $Y$ takes values in some finite alphabet $\mathcal{Y}$, while $\tilde{Y}$ takes values in $\mathbb{F}_2^\omega$. Then we define the joint typicality encoder associated to $\mathcal{C}(G, \phi)$ and $p_{Y, \phi(\tilde{Y})}$ as the mapping $\mathcal{E}_{G, \phi, p_{Y, \phi(\tilde{Y})}} : \mathcal{Y}^n \to \mathbb{F}_2^m$, where

$$\mathcal{E}_{G, \phi, p_{Y, \phi(\tilde{Y})}}(y^n) \triangleq$$
$$\operatorname{argmin}_{v^m \in \mathbb{F}_2^m} \{\alpha : (y^n, \Phi(v^m G)) \in \mathcal{T}_{p_{Y, \phi(\tilde{Y})}}^{\alpha, n}\}. \quad (6)$$

Note that for every $\alpha > 0$ and every distribution $p_{V_1, V_2}$, $\mathcal{T}_{p_{V_1, V_2}}^{\alpha, n}$ denotes the set of pairs of sequences $(v_1^n, v_2^n)$ which are $\alpha$-strongly jointly typical according to $p_{V_1, V_2}{}^2$. Likewise, $\mathcal{T}_{p_{V_1}}^{\alpha, n}$ denotes the set of sequences $v_1^n$ which are $\alpha$-strongly typical according to $p_{V_1}$.

---

[1]The construction of such variables is shown in the proof of Lemma 1 in [7].

[2]The definition of strongly typical sequences is as in [24].

$$M_{S_l \to N_l}(z) = \exp(-\gamma(y(l), z)), \text{ for all } z \in \mathcal{Z}, 1 \le l \le n,$$

$$M_{N_l \to C_{l+sn}}(b) = \sum_{z \in \mathcal{Z}} M_{S_l \to N_l}(z) \sum_{\substack{b^\omega \in \{0,1\}^\omega \\ b(s+1)=b \\ \phi(b^\omega)=z}} \prod_{\substack{j=0 \\ j \ne s}}^{\omega-1} M_{C_{l+jn} \to N_l}(b(j+1)),$$

$$M_{C_{l+sn} \to N_l}(b) = \frac{1}{2} + \frac{(-1)^b}{2} \prod_{k \in \mathcal{B}_v(l+sn)} (M_{V_k \to C_{l+sn}}(0) - M_{V_k \to C_{l+sn}}(1)),$$

$$\text{for all } b \in \{0,1\}, 1 \le l \le n, 0 \le s \le \omega - 1,$$

$$M_{C_{l+sn} \to V_k}(b) = \frac{1}{2} + \frac{(-1)^b}{2} (M_{N_l \to C_{l+sn}}(0) - M_{N_l \to C_{l+sn}}(1)) \prod_{i \in \mathcal{B}_v(l+sn) \setminus \{k\}} (M_{V_i \to C_{l+sn}}(0) - M_{V_i \to C_{l+sn}}(1)),$$

$$M_{V_k \to C_{l+sn}}(b) = \prod_{q \in \mathcal{A}_c(k) \setminus \{l+sn\}} M_{C_q \to V_k}(b),$$

$$\text{for all } b \in \{0,1\}, k \in \mathcal{B}_v(l+sn), 1 \le l \le n, 0 \le s \le \omega - 1.$$

Fig. 2. Message passing equations. The message passed by node $A$ to node $B$ is denoted by $M_{A \to B}$. It is a vector with two components for $A \ne S_l$, and with $|\mathcal{Z}|$ components for $A = S_l$. 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 $C_q$. The values of the parameters $\gamma(y, z)$ used in our simulations are specified in Section VII.

Further, as in [17] we consider an LDGM code ensemble specified by a random generator matrix $\mathbf{G}$ with elements selected independently from $\mathbb{F}_2$ using a Bernoulli distribution with parameter $p_n$, where $p_n \to 0$ and $np_n \to \infty$ as $n \to \infty$. Finally, following [17], [15] and [16], we obtain the following result, which is crucial for our development and whose proof is given in Appendix A.

**Lemma 1.** Consider the LDGM code ensemble as described above for some $R$, $\omega$, $\phi$, $Y$, and $\tilde{Y}$ such that $\tilde{Y}$ is uniformly distributed over $\mathbb{F}_2^\omega$ and $R \ge I(Y; \tilde{Y}) + \epsilon_0$, for some arbitrary $\epsilon_0 > 0$. Furthermore, assume that the random vector $Y^n$ representing the input at the encoder satisfies the equality

$$\lim_{n \to \infty} \mathbb{P}(Y^n \in \mathcal{T}_{p_Y}^{\delta,n}) = 1, \tag{7}$$

for any $\delta > 0$. Then, for any $\delta > 0$ one has

$$\lim_{n \to \infty} \mathbb{P}((Y^n, \Phi(\mathcal{E}_{\mathbf{G}, \phi, p_{Y,\phi(\tilde{Y})}}(Y^n)\mathbf{G})) \in \mathcal{T}_{p_{Y,\phi(\tilde{Y})}}^{\delta,n}) = 1. \tag{8}$$

In order to proceed with the specifics of the proposed LDGM-based code for each encoding stage, we need first the following result which was proved in [17] as a part of the proof of Theorem 1.

**Lemma 2 ( [17]).** Consider arbitrary jointly distributed random variables $Y$ and $Z$ over the finite alphabets $\mathcal{Y}$ and $\mathcal{Z}$, respectively. Then, for any positive integer $\omega$ there exist a random variable $\tilde{Y}_\omega$ uniformly distributed over $\mathbb{F}_2^\omega$ and a mapping $\phi_\omega : \mathbb{F}_2^\omega \to \mathcal{Z}$ such that

$$\lim_{\omega \to \infty} p_{Y, \phi_\omega(\tilde{Y}_\omega)}(y, z) = p_{Y,Z}(y, z)$$

for every $(y, z) \in \mathcal{Y} \times \mathcal{Z}$.

In virtue of the above result for each triple of positive integers $\omega_0, \omega_1, \omega_2$ there exist random variables $\tilde{X}_{0,\omega_0}$, $\tilde{X}_{1,\omega_0,\omega_1}$ and $\tilde{X}_{2,\omega_0,\omega_1,\omega_2}$ uniformly distributed over $\mathbb{F}_2^{\omega_0}$, $\mathbb{F}_2^{\omega_1}$, and $\mathbb{F}_2^{\omega_2}$ respectively, and mappings $\phi_0 : \mathbb{F}_2^{\omega_0} \to \mathcal{X}_c$, $\phi_1 : \mathbb{F}_2^{\omega_1} \to \mathcal{U}_1$,

and $\phi_2 : \mathbb{F}_2^{\omega_2} \to \mathcal{U}_2$ such that

$$\lim_{\omega_0 \to \infty} p_{X, \phi_0(\tilde{X}_{0,\omega_0}), U_1, U_2}(x, x_c, u_1, u_2)$$
$$= p_{X, X_c, U_1, U_2}(x, x_c, u_1, u_2), \tag{9}$$

$$\lim_{\omega_1 \to \infty} p_{X, \phi_0(\tilde{X}_{0,\omega_0}), \phi_1(\tilde{X}_{1,\omega_0,\omega_1}), U_2}(x, x_c, u_1, u_2)$$
$$= p_{X, \phi_0(\tilde{X}_{0,\omega_0}), U_1, U_2}(x, x_c, u_1, u_2), \tag{10}$$

$$\lim_{\omega_2 \to \infty} p_{X, \phi_0(\tilde{X}_{0,\omega_0}), \phi_1(\tilde{X}_{1,\omega_0,\omega_1}), \phi_2(\tilde{X}_{2,\omega_0,\omega_1,\omega_2})}(x, x_c, u_1, u_2)$$
$$= p_{X, \phi_0(\tilde{X}_{0,\omega_0}), \phi_1(\tilde{X}_{1,\omega_0,\omega_1}), U_2}(x, x_c, u_1, u_2), \tag{11}$$

for every $(x, x_c, u_1, u_2) \in \mathcal{X} \times \mathcal{X}_c \times \mathcal{U}_1 \times \mathcal{U}_2$.

Let $\omega_j$, respectively $\phi_j$, denote the value of $\omega$, respectively the mapping $\phi$ used at encoder $j$, for $j = 0, 1, 2$. The encoding mapping at each stage is defined as in (6) where

a) $Y = X$ and $\tilde{Y} = \tilde{X}_{0,\omega_0}$ at encoder 0;

b) $Y = (X, \phi_0(\tilde{X}_{0,\omega_0}))$ and $\tilde{Y} = \tilde{X}_{1,\omega_0,\omega_1}$ at encoder 1;

c) $Y = (X, \phi_0(\tilde{X}_{0,\omega_0}), \pi_1(\phi_0(\tilde{X}_{0,\omega_0}), \phi_1(\tilde{X}_{1,\omega_0,\omega_1})))$ and $\tilde{Y} = \tilde{X}_{2,\omega_0,\omega_1,\omega_2}$ at encoder 2.

The rates at the three encoders are as in the previous subsection, i.e.

$R_{(0)} = I(X; X_c) + \epsilon$ at encoder 0;

$R_{(1)} = I(X, X_c : U_1) + \epsilon$ at encoder 1;

$R_{(2)} = I(X, X_c, X_1; U_2) + \epsilon$ at encoder 2.

Lemmas 1 and 2 immediately imply the asymptotical optimality of the proposed LDGM-based coding scheme, which is stated next.

**Theorem 1.** The sequential LDGM-based coding scheme described in this subsection ensures that the target distortion values $d_1, d_2$ and $d_0$ for the side and central descriptions, respectively, can be approached arbitrarily closely as $n, \omega_0, \omega_1, \omega_2 \to \infty$.

*Proof of Theorem 1.* Notice that in view of relations (9-11) and of the continuity of the information mapping $I(\cdot, \cdot)$, there

variables $X_t$, $t = 0, 1, 2$, jointly distributed with $X$ such that

$$I(X_1; X_2) = 0,$$
$$R_t \geq I(X; X_t), \quad t = 1, 2,$$
$$R_1 + R_2 \geq I(X; X_0, X_1, X_2),$$
$$d_t \geq \mathbb{E}[d(X, X_t)], \quad t = 0, 1, 2.$$

Our next result represents some progress towards an explicit characterization of $\mathcal{RD}$ in the no excess sum-rate case for general binary sources with Hamming distortion measure (i.e., $d(x, \hat{x}) = d_H(x, \hat{x}) = 0$ if $x = \hat{x}$, and 1 otherwise). Let $\mathcal{X} = \hat{\mathcal{X}} = \{0, 1\}$ and $p_X(0) = \delta$, where $0 < \delta \leq 1/2$. Let $\mathcal{D}$ denote the distortion region, i.e.,

$$\mathcal{D} \triangleq \{(d_1, d_2, d_0) : \exists (R_1, R_2) \text{ such that } R_1 + R_2 = R(d_0)$$
$$\text{and } (R_1, R_2, d_1, d_2, d_0) \in \mathcal{RD}\}.$$

For $d_0 \geq 0$ let us additionally define

$$\mathcal{D}(d_0) \triangleq \{(d_1, d_2) : (d_1, d_2, d_0) \in \mathcal{D}\}.$$

Clearly,

$$\mathcal{D} = \cup_{d_0 : d_0 \geq 0} \{(d_1, d_2, d_0) : (d_1, d_2) \in \mathcal{D}(d_0)\}.$$

Now we proceed to characterize $\mathcal{D}(d_0)$. With no loss of generality, we shall assume $0 \leq d_0 < \delta$. Since $\mathcal{D}(d_0)$ is convex, it suffices to characterize its supporting lines. Two of the supporting lines are already known, namely $d_1 = d_0$ and $d_2 = d_0$. Therefore, to complete the task, it is enough to solve the following optimization problem, for all $\alpha > 0$,

$$\min_{(d_1, d_2) \in \mathcal{D}(d_0)} \alpha d_1 + d_2. \tag{13}$$

Note that $R(d_0)$ is a strictly convex function of $d_0$ for $d_0 \in [0, \delta)$. Furthermore, if

$$R(d_0) = R_1 + R_2 \geq I(X; X_0, X_1, X_2), \quad \text{and}$$
$$d_0 \geq \mathbb{E}[d(X, X_0)],$$

then one must have $X = X_0 \oplus Z$, where $\oplus$ denotes the modulo-2 addition operation, and $Z \sim \text{Ber}(d_0)$ is independent of $(X_0, X_1, X_2)$. Therefore, it suffices to specify $p_{X_1 X_2 | X_0}$ in order to determine $p_{X X_0 X_1 X_2}$ due to the fact that $p_{X X_0}$ is completely determined once $d_0$ is given and that $X - X_0 - (X_1, X_2)$ form a Markov chain. Now in view of (12), it can be readily shown that

$$\min_{(d_1, d_2) \in \mathcal{D}(d_0)} \alpha d_1 + d_2 \geq$$
$$\min_{p_{X_1 X_2 | X_0} : I(X_1; X_2) = 0} \alpha \mathbb{E}[d(X, X_1)] + \mathbb{E}[d(X, X_2)].$$

It turns out that the inequality can be replaced by an equality. To see this, for any $p_{X_1 X_2 | X_0}$ such that the induced $X_1$ and $X_2$ are independent, let

$$R_1 = I(X; X_1),$$
$$R_2 = I(X, X_1; X_2) + I(X; X_0 | X_1, X_2),$$
$$d_t = \mathbb{E}[d(X, X_t)], \quad t = 0, 1, 2.$$

It can be verified that $(R_1, R_2, d_1, d_2, d_0) \in \mathcal{Q}$ and $R_1 + R_2 = R(d_0)$. Therefore, (13) is equivalent to

$$\min_{p_{X_1 X_2 | X_0} : I(X_1; X_2) = 0} \alpha \mathbb{E}[d(X, X_1)] + \mathbb{E}[d(X, X_2)]. \tag{14}$$

Solving (14) for all $\alpha > 0$ leads to the following result, whose proof is deferred to Appendix B.

**Theorem 2.** The following relation holds

$$\mathcal{D}(d_0) = \{(d_1, d_2) : (d_1 + 1 - 2d_0 - \delta)(d_2 + 1 - 2d_0 - \delta)$$
$$\geq (1 - d_0 - \delta)(1 - 2d_0), d_1 \geq d_0, d_2 \geq d_0\}.$$

Moreover, for each $\alpha > 0$, the solution $(d_1, d_2)$ to (13) is unique, and a corresponding solution to (14) is as specified in Fig. 4. Additionally, when $\delta < 1/2$, the solution to (14) is unique, while for $\delta = 1/2$ and $d_1 \neq d_0, d_2 \neq d_0$, there is only one alternative solution, which is obtained by swapping 0 and 1 in Fig. 4.

$$p_{XX_1X_2}(0, 1, 1) = \frac{d_0}{1 - 2d_0}(1 - \delta - d_0),$$
$$p_{XX_1X_2}(1, 0, 0) = \frac{d_0}{1 - 2d_0}(\delta + d_0 - d_1 - d_2),$$
$$p_{XX_1X_2}(1, 1, 0) = \frac{d_0}{1 - 2d_0}(d_1 - d_0),$$
$$p_{XX_1X_2}(1, 0, 1) = \frac{d_0}{1 - 2d_0}(d_2 - d_0),$$
$$p_{XX_1X_2}(0, 1, 0) = \frac{1 - d_0}{1 - 2d_0}(d_1 - d_0),$$
$$p_{XX_1X_2}(0, 0, 1) = \frac{1 - d_0}{1 - 2d_0}(d_2 - d_0),$$
$$p_{XX_1X_2}(0, 0, 0) = \frac{1 - d_0}{1 - 2d_0}(\delta + d_0 - d_1 - d_2),$$
$$p_{XX_1X_2}(1, 1, 1) = \frac{1 - d_0}{1 - 2d_0}(1 - \delta - d_0),$$
$$X_0 = \psi_0(X_1, X_2) = \left\{ \begin{array}{ll} 1, & X_1 = X_2 = 1 \\ 0, & \text{otherwise} \end{array} \right. ;$$

Fig. 4. Auxiliary random variables achieving the lower boundary of $\mathcal{D}(d_0)$.

Note that for the case of the uniform binary source, part of the conclusion of Theorem 2 can be inferred from [5, Section IV]. Fig. 5 depicts the distortion region $\mathcal{D}(d_0)$ for a binary source with $p_X(0) = 1/4$, $d_0 = 0$, and $d_0 = 0.013$.

Consider now a distortion pair $(d_1, d_2)$ on the lower boundary[3] of $\mathcal{D}(d_0)$, let $p_{X_0 X_1 X_2 | X}$ and $\psi_0(X_1, X_2)$ be specified by Fig. 4, $X_c = 0$, and $\psi_t(X_t) = X_t$, for $t = 1, 2$. Then relations (3) and (4) are satisfied. Further, define the rate region

$$\mathcal{R}(d_1, d_2, d_0) \triangleq \{(R_1, R_2) : R_1 + R_2 = R(d_0), \text{ and}$$
$$(R_1, R_2, d_1, d_2, d_0) \in \mathcal{RD}\}.$$

It can be shown by leveraging Theorem 2 that $\mathcal{R}(d_1, d_2, d_0)$ coincides with the set of rate pairs satisfying $R_1 \geq I(X; X_1)$, $R_2 \geq I(X; X_2)$, $R_1 + R_2 = R(d_0)$. Thus, $\mathcal{R}(d_1, d_2, d_0)$ coincides with the dominant face $\mathcal{F}(p_{X_c X_1 X_2 | X})$ defined in

---

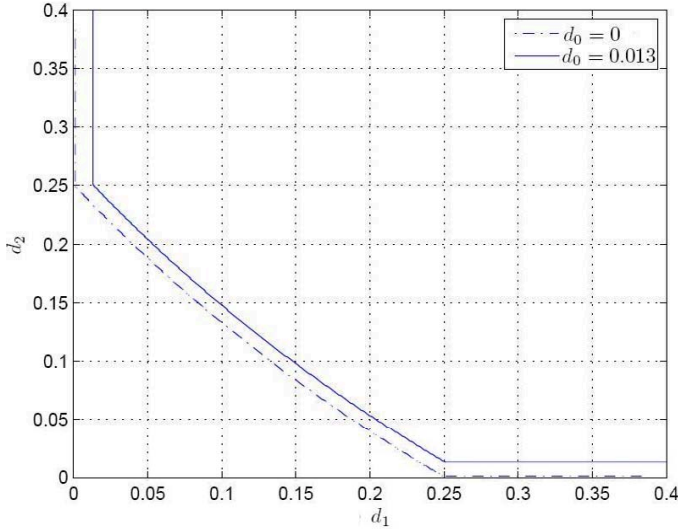[3]The lower boundary of $\mathcal{D}(d_0)$ is defined as the set of solutions to (13) for $\alpha > 0$.

Fig. 5. Distortion region $\mathcal{D}(d_0)$ for a binary source with $p_X(0) = 1/4$, $d_0 = 0$, and $d_0 = 0.013$.

Subsection III-A, and the coding scheme proposed in Section III, with $U_1 = X_1$ and $U_2 = X_2$, can be used to approach its extremities.

## V. No Excess Marginal Rate Case for the Uniform Binary Source with Hamming Distortion

The term "no excess marginal rate" refers to the case when $R_1 = R(d_1)$ and $R_2 = R(d_2)$. An interesting question in such a case is, given the distortion pair $(d_1, d_2)$, what is the minimum distortion $d_0$? Precisely, let us define

$$d_0(d_1, d_2) \triangleq \min\{d_0 : (R(d_1), R(d_2), d_1, d_2, d_0) \in \mathcal{RD}\}.$$

Zhang and Berger proposed in [6] an upper bound for $d_0(d_1, d_2)$ by choosing the auxiliary random variables $X_c, X_1, X_2$ in the definition of the ZB region such that $X - X_t - X_c$, $t = 1, 2$, and $X_1 - (X, X_c) - X_2$ form Markov chains, $I(X; X_t) = R(d_t)$, $\mathbb{E}[d(X, X_t)] \leq d_t$, $\psi_t(X_t, X_c) = X_t$, for $t = 1, 2$, and $\psi_0(X_c, X_1, X_2) = X_1$, if $X_1 = X_2$, and $X_c$, otherwise.

Further they showed that for the uniform binary source with Hamming distortion the proposed upper bound is strictly better than the upper bound derived from the EGC region, namely than $\min(d_1, d_2)$. Notice that the case of interest is when $\min(d_1, d_2) > 0$. Now let

$$\mathbb{P}(X_c = X|X) = 1 - s, \quad \mathbb{P}(X_c \neq X|X) = s,$$
$$\mathbb{P}(X_t = X|X) = 1 - d_t, \quad \mathbb{P}(X_t \neq X|X) = d_t,$$
$$\mathbb{P}(X_c = X_t|X) = 1 - s_t, \quad \mathbb{P}(X_c \neq X_t|X) = s_t, \quad t = 1, 2,$$

for some $s$ such that $\max(d_1, d_2) \leq s \leq \frac{1}{2}$, and $s_t = \frac{s - d_t}{1 - 2d_t}$, $t = 1, 2$. Then the upper bound is computed in [6] as

$$UB(d_1, d_2) = \inf_{\max(d_1, d_2) \leq s \leq \frac{1}{2}} \mathbb{P}(X \neq \psi_0(X_c, X_1, X_2)). \quad (15)$$

We argue that the mapping $\psi_0$ given in [6] is not optimal and provide the optimal mapping $\psi_0$ which minimizes the central distortion, along with the optimal value $s_0$ achieving the infimmum in (15). Let us assume without restricting the generality that $d_1 \leq d_2$. Define $\alpha(d_1, d_2) \triangleq \frac{d_1(1 - d_2)}{d_1 + d_2 - 2d_1 d_2}$. Then relations $d_2 \leq \alpha(d_1, d_2) \leq 1/2$ hold. In Appendix C we show that the mapping $\psi_0$ which minimizes the central distortion is the following

$$\psi_0(x_c, x_1, x_2) = \begin{cases} x_1, & x_1 = x_2 \\ 1 - x_c, & \text{if } x_1 \neq x_2, \quad d_2 \leq s \leq \alpha(d_1, d_2) \\ x_1, & \text{if } x_1 \neq x_2, \quad \alpha(d_1, d_2) < s \leq 1/2 \end{cases}. \quad (16)$$

Then the upper bound (15) becomes

$$UB(d_1, d_2) =$$
$$\min_{d_2 \leq s \leq \alpha(d_1, d_2)} \Big( 1 - s - \frac{(1 - d_1)(1 - d_2)(1 - s - d_1)(1 - s - d_2)}{(1 - s)(1 - 2d_1)(1 - 2d_2)}$$
$$+ \frac{d_1 d_2 (1 - s - d_1)(1 - s - d_2)}{s(1 - 2d_1)(1 - 2d_2)} \Big). \quad (17)$$

Define $\beta(d_1, d_2) \triangleq \frac{d_1 + d_2 - 2d_1 d_2}{d_1 d_2 (1 - d_1)(1 - d_2)}$. As argued in Appendix C, the solution $s_0$ to (17) is

$$s_0 = \frac{1}{2} - \frac{1}{2} \sqrt{1 - 4 \frac{\sqrt{1 + \beta(d_1, d_2)} - 1}{\beta(d_1, d_2)}}. \quad (18)$$

Thus, $UB(d_1, d_2)$ can be recovered by substituting (18) into the expression in (17).

Further, to apply the the coding scheme proposed in Section III for the distortion triple $(d_1, d_2, d_0 = UB(d_1, d_2))$, we define $U_t \triangleq X_c \oplus X_t$, $t = 1, 2$. It follows easily that C1 and C3 are satisfied, while C2 holds since $X_t = \pi_t(X_c, U_t) \triangleq X_c \oplus U_t$, $t = 1, 2$.

## VI. Successive Refinement with Hamming Distortion

Successive refinement (SR) can be regarded as a special form of multiple description coding in which the distortion constraint on the second description (i.e., $d_2$) is not imposed. In this scenario it is common to refer to the first description as the base layer and the second description as the refinement layer. The successive refinement coding rate-distortion region $\mathcal{RD}_{\text{SR}}$ is given by

$$\mathcal{RD}_{\text{SR}} = \{(R_1, R_2, d_1, d_0) : (R_1, R_2, d_1, \infty, d_0) \in \mathcal{RD}\}.$$

An important case is when $(R(d_1), R(d_0) - R(d_1), d_1, d_0) \in \mathcal{RD}_{\text{SR}}$. The sources for which this relation holds for all admissible $d_1 > d_0$ are called in [19] successively refinable sources. Equitz and Cover showed in [19] that $(R(d_1), R(d_0) - R(d_1), d_1, d_0) \in \mathcal{RD}_{\text{SR}}$ if and only if there are random variables $X_c$ and $X_1$, jointly distributed with the generic source variable $X$, such that

$$R(d_1) = I(X; X_c), \quad R(d_0) = I(X; X_c, X_1), \quad (19)$$
$$\mathbb{E}[d(X, X_c)] \leq d_1, \quad \mathbb{E}[d(X, X_1)] \leq d_0,$$

and $X - X_1 - X_c$ form a Markov chain. They further proved, based on the results of [25], that finite alphabet sources with Hamming distortion are successively refinable and that the joint distribution $p_{XX_cX_1}$ is specified by the relations in Fig. 6.

Given $p_{XX_cX_1}$ as in Fig. 6, the rate-distortion tuple satisfying (19) can be achieved by using a sequential coding scheme adapted from Section III. The scheme consists only

$$p_{X_t}(x) = \frac{(p_X(x) - \lambda_t)^+}{\sum_{x' \in \mathcal{X}}(p_X(x') - \lambda_t)^+}, \quad x \in \mathcal{X}, t \in \{c, 1\},$$

$$p_{X|X_1}(x|x_1) = \begin{cases} 1 - d_0, & x = x_1 \in \mathcal{X}_1^+ \\ \lambda_1, & x \neq x_1, x \in \mathcal{X}_1^+, x_1 \in \mathcal{X}_1^+ \\ p_X(x) & x \notin \mathcal{X}_1^+, x_1 \in \mathcal{X}_1^+ \end{cases},$$

$$p_{X_1|X_c}(x_1|x_c) = \begin{cases} \frac{1 - d_1 - \lambda_1}{1 - d_0 - \lambda_1}, & x_1 = x_c \in \mathcal{X}_c^+ \\ \frac{\lambda_c - \lambda_1}{1 - d_0 - \lambda_1}, & x_1 \neq x_c, x_1 \in \mathcal{X}_c^+, x_c \in \mathcal{X}_c^+ \\ \frac{p_X(x_1) - \lambda_1}{1 - d_0 - \lambda_1} & x_1 \in \mathcal{X}_1^+, x_1 \notin \mathcal{X}_c^+, x_c \in \mathcal{X}_c^+ \end{cases},$$

where $\lambda_1 \in [0, \lambda_c]$, $\lambda_c$ are determined by

$$\sum_{x_t \in \mathcal{X}_t^+} p_{X_t}(x_t) p_{X|X_t}(x|x_t) = p_X(x), x \in \mathcal{X}, t \in \{c, 1\},$$

and $\mathcal{X}_t^+ = \{x \in \mathcal{X} : p_X(x) - \lambda_t > 0\}, t \in \{c, 1\}$.

Fig. 6. Relations defining the joint distribution $p_{XX_cX_1}$ for SR.

| $(R_1, R_2)$ | | $(d_1, d_2, d_0)$ | $\hat{d}_1$ | $\hat{d}_2$ | $\hat{d}_0$ |
|---|---|---|---|---|---|
| (0.500, 0.500) | $E_1$ | (0.147, 0.272, 0) | 0.154 | 0.279 | 0.008 |
| (0.500, 0.500) | $E_1$ | (0.156, 0.266, 0) | 0.156 | 0.269 | 0.008 |
| (0.383, 0.617) | $E_1$ | (0.207, 0.207, 0) | 0.209 | 0.212 | 0.007 |
| (0.617, 0.383) | $E_2$ | (0.207, 0.207, 0) | 0.212 | 0.209 | 0.007 |
| (0.500, 0.500) | $T$ | (0.207, 0.207, 0) | 0.211 | 0.211 | 0.007 |
| (0.445, 0.555) | $E_1$ | (0.174, 0.242, 0) | 0.176 | 0.247 | 0.009 |
| (0.676, 0.324) | $E_2$ | (0.174, 0.242, 0) | 0.175 | 0.246 | 0.009 |
| (0.500, 0.500) | $T$ | (0.174, 0.242, 0) | 0.176 | 0.247 | 0.009 |
| (0.349, 0.551) | $E_1$ | (0.215, 0.215, 0.013) | 0.219 | 0.216 | 0.021 |
| (0.551, 0.349) | $E_2$ | (0.215, 0.215, 0.013) | 0.216 | 0.219 | 0.021 |
| (0.450, 0.450) | $T$ | (0.215, 0.215, 0.013) | 0.218 | 0.218 | 0.021 |
| (0.422, 0.478) | $E_1$ | (0.174, 0.258, 0.013) | 0.181 | 0.258 | 0.022 |
| (0.619, 0.281) | $E_2$ | (0.174, 0.258, 0.013) | 0.177 | 0.259 | 0.021 |
| (0.450, 0.450) | $T$ | (0.174, 0.258, 0.013) | 0.180 | 0.258 | 0.022 |

TABLE I

TEST RESULTS IN THE NO EXCESS SUM-RATE CASE FOR THE UNIFORM BINARY SOURCE: $(d_1, d_2, d_0)$ IS A TARGET DISTORTION TRIPLE; $(R_1, R_2)$ IS THE PAIR OF RATES USED BY THE CODE; $\hat{d}_1, \hat{d}_2, \hat{d}_0$ ARE THE EMPIRICAL DISTORTIONS.

| $(R_1, R_2)$ | | $(d_1, d_2, d_0)$ | $\hat{d}_1$ | $\hat{d}_2$ | $\hat{d}_0$ |
|---|---|---|---|---|---|
| (0.492, 0.319) | $E_1$ | (0.116, 0.116, 0) | 0.120 | 0.120 | 0.014 |
| (0.319, 0.492) | $E_2$ | (0.116, 0.116, 0) | 0.120 | 0.120 | 0.014 |
| (0.424, 0.387) | $T$ | (0.116, 0.116, 0) | 0.120 | 0.120 | 0.014 |
| (0.502, 0.309) | $E_1$ | (0.060, 0.176, 0) | 0.065 | 0.181 | 0.014 |
| (0.162, 0.649) | $E_2$ | (0.176, 0.060, 0) | 0.184 | 0.065 | 0.016 |
| (0.373, 0.439) | $T$ | (0.104, 0.132, 0) | 0.110 | 0.137 | 0.015 |
| (0.291, 0.423) | $E_1$ | (0.123, 0.123, 0.013) | 0.130 | 0.128 | 0.025 |
| (0.423, 0.291) | $E_2$ | (0.123, 0.123, 0.013) | 0.128 | 0.130 | 0.025 |
| (0.369, 0.344) | $T$ | (0.123, 0.123, 0.013) | 0.129 | 0.129 | 0.025 |
| (0.492, 0.222) | $E_1$ | (0.060, 0.191, 0.013) | 0.064 | 0.197 | 0.027 |
| (0.122, 0.591) | $E_2$ | (0.191, 0.060, 0.013) | 0.198 | 0.066 | 0.027 |
| (0.377, 0.337) | $T$ | (0.101, 0.151, 0.013) | 0.105 | 0.156 | 0.027 |

TABLE II

TEST RESULTS IN THE NO EXCESS SUM-RATE CASE FOR A BINARY SOURCE WITH $p_X(0) = 1/4$: $(d_1, d_2, d_0)$ IS A TARGET DISTORTION TRIPLE; $(R_1, R_2)$ IS THE PAIR OF RATES USED BY THE CODE; $\hat{d}_1, \hat{d}_2, \hat{d}_0$ ARE THE EMPIRICAL DISTORTIONS.

of encoders 0 and 1, whose outputs represent the base, respectively the refinement layer. To clarify the details of the scheme we only need to specify $U_1$. Assume without restricting the generality that $\mathcal{X} = \{0, 1, \cdots, \zeta - 1\}$ and $\mathcal{X}_c^+ = \{0, 1, \cdots, \tau - 1\}$, for some integers $\zeta \geq \tau > 0$. Then let $U_1 \triangleq \sigma(X_c, X_1)$, where $\sigma(X_c, X_1) = X_1 \oplus_\tau (\tau - X_c)$, if $X_1 \in \mathcal{X}_c^+$, and $\sigma(X_c, X_1) = X_1$ otherwise, and $\oplus_\tau$ denotes the modulo-$\tau$ addition. It can be easily verified that conditions C1-C3 hold, where $\pi_1(X_c, U_1) = U_1 \oplus_\tau X_c$ if $U_1 \in \mathcal{X}_c^+$, and $\pi_1(X_c, U_1) = U_1$ otherwise.

## VII. EXPERIMENTAL RESULTS

We have tested the proposed LDGM-based coding scheme in each of the three cases of the MD problem discussed in the previous section. The degree distributions of the LDGM codes are from the website (http://lthcwww.epfl.ch.research/ldpcopt) or obtained by implementing the algorithm in [26]. We have used damping as in [14], [17] in our message passing algorithm, if the messages do not converge after 30 iterations. The length $n$ of the input sequence is 10,000. Next we present the experimental results for each case. The values of the empirical distortions are averaged over 100 runs.

### A. No Excess Sum-Rate Case for Binary Sources

We have targeted distortion pairs $(d_1, d_2)$ on the lower boundary of $\mathcal{D}(d_0)$ for a uniform binary source (UBS) and for a non-uniform binary source (NBS) with $p_X(0) = 1/4$. In both cases we have considered $d_0 = 0$ and $d_0 = 0.013$. In our tests the sum-rate $R_1 + R_2$ equals $R(d_0)$ in each case. The value of the threshold $\eta$ is 0.9, while $\delta$ is 1.8 for UBS and 1.6 for NBS. At encoder 1 we set $\gamma(0, 0) = \gamma(1, 1) = 0$, and a) $\gamma(0, 1) = 1.0$, $\gamma(1, 0) \in [3.2, 3.6]$ for UBS; b) $\gamma(0, 1) \in [1.0, 2.0]$, $\gamma(1, 0) = 2.8$ for NBS. At encoder 2 we use $\gamma((0, 1), 0) = \gamma((1, 1), 1) = \gamma((0, 0), 0) = \gamma((1, 0), 1) = 0$ and $\gamma((0, 1), 1) = \gamma((1, 1), 0) = 3.2$, while: a) $\gamma((0, 0), 1) = \gamma((1, 0), 0) \in [0.0, 0.6]$ for UBS; and b) $\gamma((0, 0), 1) = \gamma((1, 0), 0) = 0.0$ for NBS. Tables I and II present the results for UBS, respectively NBS. The second

column indicates whether the pair of rates $(R_1, R_2)$ is an $E_1$ or $E_2$ corner point, or it is obtained by timesharing ($T$) the two corner points. As observed from Tables I and II, the distortions are very close to the theoretical lower bounds.

### B. No Excess Marginal Rate Case for the Uniform Binary Source

We have considered three target distortion triples $(d_1, d_2, d_0 = UB(d_1, d_2))$: $d_1 = d_2 = 0.1, 0.2, 0.3$. At encoder 0, we have set $\gamma(x, x_c) = 0$ if $x = x_c$, and 0.5 if $x \neq x_c$. At encoders 1 and 2, $\gamma((x, x_c), u_1) = \rho(x \oplus x_c, u_1)$ and $\gamma((x, x_c, u_1), u_2) = \rho(x \oplus x_c, u_2)$, respectively, where $\rho(x \oplus x_c, u_t) = 0$ if $x \oplus x_c = u_t$, and 1.2 if $x \oplus x_c \neq u_t$, $t = 1, 2$. The value of $\delta$ is 0.5 at encoder 0 and 0.8 at encoders 1 and 2. The threshold $\eta$ is set to 0.99. In all the cases $R_1 = R(d_1) = R_2 = R(d_2)$. The experimental results, which are summarized in Table III, show very good performance.

| $d_1 = d_2$ | $UB(d_1, d_2)$ | $\hat{d}_1$ | $\hat{d}_2$ | $\hat{d}_0$ |
|---|---|---|---|---|
| 0.1 | 0.0626 | 0.1159 | 0.1136 | 0.0783 |
| 0.2 | 0.1574 | 0.2161 | 0.2170 | 0.1720 |
| 0.3 | 0.2658 | 0.3173 | 0.3134 | 0.2823 |

TABLE III

TEST RESULTS IN THE NO EXCESS MARGINAL RATE CASE FOR THE UNIFORM BINARY SOURCE: $d_1 = d_2$ ARE TARGET DISTORTION VALUES FOR THE SIDE DESCRIPTIONS; $UB(d_1, d_2)$ IS THE TARGET VALUE FOR THE CENTRAL DISTORTION; $\hat{d}_1, \hat{d}_2, \hat{d}_0$ ARE THE EMPIRICAL DISTORTIONS.

### C. Successive Refinement

We have tested the proposed SR coding scheme for two binary and one ternary sources. The binary sources are 1) uniform binary source; 2) non-uniform binary source with $p_X(0) = 0.25$. In both cases $\delta = 1.6$, and a) at encoder 0, $\gamma(x, x_c) = 0$ if $x = x_c$, and 2.8 if $x \neq x_c$; b) at encoder 1, $\gamma((x, x_c), u_1) = 0$ if $x \oplus x_c = u_1$, and 2.8 otherwise. For the uniform ternary source $\delta = 1.1$ and a) at encoder 0, $\gamma(x, x_c) = 0$ if $x = x_c$, and $\rho$ if $x \neq x_c$; b) at encoder 1, $\gamma((x, x_c), u_1) = 0$ if $x \oplus_3 (3 - x_c) = u_1$, and $\rho$ otherwise, where $\rho \in [1.91, 2.31]$. The value of $\eta$ is 0.9 or 0.99. The results of our experiments are presented in Figs. 7, 8 and 9. As seen in these figures, the empirical distortions are very close to the theoretical limits.
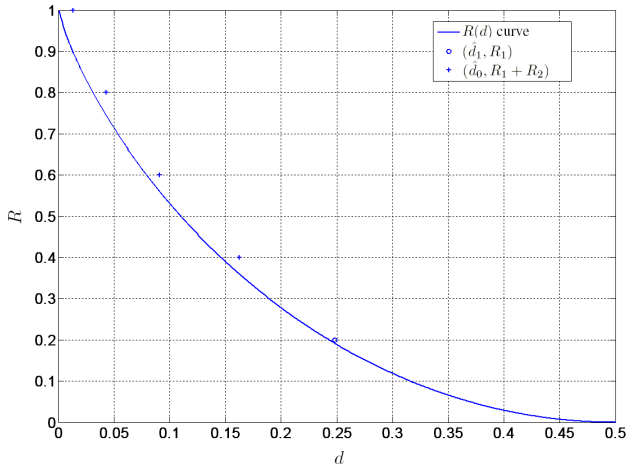


Fig. 7. Simulation results in the SR case for the uniform binary source at $R_1 = 0.2$ and various $R_2$ values.

## VIII. CONCLUSION

This work presents a practical MD coding scheme based on LDGM codes for the Zhang-Berger region. The asymptotical optimality of the scheme is proved for a joint typicality encoder, while a message passing algorithm is used as a practical encoding solution. The application of this scheme to three cases of the MD problem is discussed in more detail: 1) the no excess sum-rate case for binary sources; 2) the no excess marginal rate case for the uniform binary source; 3) the problem of successive refinement. In all cases, the Hamming distortion is the fidelity criterion. To aid the code
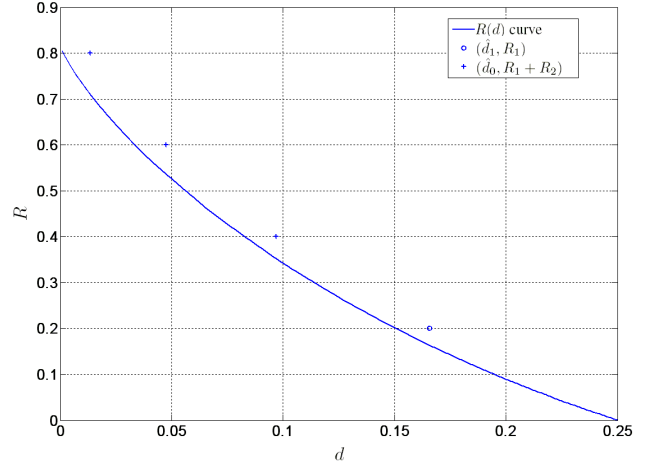


Fig. 8. Simulation results in the SR case for a non-uniform binary source with $p_X(0) = 0.25$ at $R_1 = 0.2$ and various $R_2$ values.
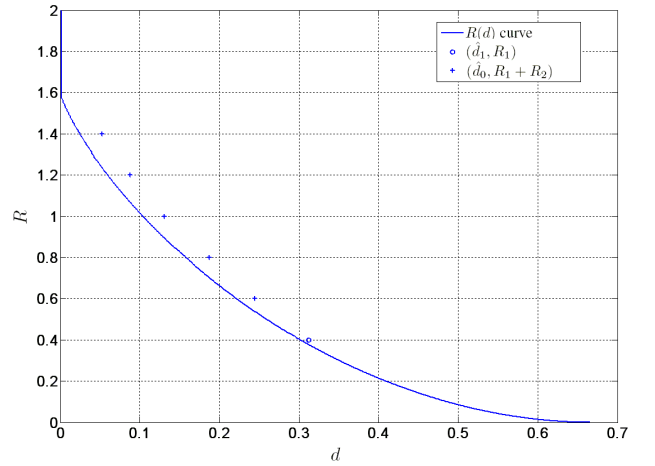


Fig. 9. Simulation results in the SR case for the uniform ternary source at $R_1 = 0.4$ and various $R_2$ values.

design for the first of the aforementioned cases, we provide the analytical expression of the distortion region, which was known previously only for the uniform binary source. Our simulations show very good results, validating the ability of the proposed coding scheme to approach in practice the theoretical rate-distortion limits or the available upper bounds.

## APPENDIX A
## PROOF OF LEMMA 1

For each $1 \leq l \leq n$, let $\varrho_l$ denote the random $m \times \omega$ matrix consisting of columns $l, l+n, l+2n, \cdots, l+(\omega-1)n$ of the random matrix $\mathbf{G}$, and define $\tilde{Y}^n(v^m) = (v^m \varrho_l)_{l=1}^n$ for every $v^m \in \mathbb{F}_2^m$. Further, for each $y^n \in \mathcal{Y}^n$ and $\delta > 0$ define

$$\sigma_n(y^n, \mathbf{G}, \delta) \triangleq \sum_{v^m \in \mathbb{F}_2^m} \mathbb{I}((y^n, \tilde{Y}^n(v^m)) \in \mathcal{T}_{p_{Y, \tilde{Y}}}^{\delta, n}).$$

It is easy to verify that if $(y^n, \tilde{y}^n) \in \mathcal{T}_{p_{Y, \tilde{Y}}}^{\frac{\delta}{|\mathcal{Z}|}, n}$, then $(y^n, \phi(\tilde{y}^n)) \in \mathcal{T}_{p_{Y, \tilde{Y}}}^{\delta, n}$. Therefore, in order to prove Lemma 1 it

is sufficient to show that $\lim_{n\to\infty} \mathbb{P}(\sigma_n(Y^n, \mathbf{G}, \delta) > 0) = 1$, for any $\delta > 0$.

According to the proof of Theorem 2 in [17] the following holds

$$\lim_{n\to\infty} \frac{1}{n} \log_2 \mathbb{P}(\sigma_n(Y^n, \mathbf{G}, \delta) > 0) \geq 0, \qquad (20)$$

for any $\delta > 0$. We point out that the proof in [17] was carried out assuming that the sequence $Y^n$ was i.i.d according to the probability $p_Y$. However, a careful examination reveals that only property (7) was needed in the proof[4].

To complete the proof we will construct a bounded difference martingale sequence. For this, for each $y^n \in \mathcal{Y}^n$ and matrix $G$ define the mapping

$$\rho(y^n, G) \triangleq$$
$$\min\{\frac{n\alpha}{|\mathcal{Y}||\tilde{\mathcal{Y}}|} : (y^n, \tilde{y}^n(v^m)) \in \mathcal{T}_{p_{Y,\tilde{Y}}}^{\alpha,n} \text{ for some } v^m \in \mathbb{F}_2^m\},$$

where $\tilde{y}^n(v^m)$ denotes the realization of $\tilde{Y}^n(v^m)$ for the given matrix $G$. Consider now a sequence of random variables $\{B_{0,n}, B_{1,n}, \cdots, B_{n,n}\}$ defined as follows:

$$B_{0,n} = \mathbb{E}\{\rho(Y^n, \mathbf{G})\},$$
$$B_{i,n} = \mathbb{E}\{\rho(Y^n, \mathbf{G})|Y_1, \cdots, Y_i, \varrho_1, \cdots, \varrho_i\}, \quad 1 \leq i \leq n.$$

Notice that $B_{n,n} = \rho(Y^n, \mathbf{G})$. Further, it can be easily shown that the inequality $|B_{i+1,n} - B_{i,n}| \leq 1$ holds for all $i, 0 \leq i \leq n - 1$. Then according to Azuma's inequality [27] it follows that

$$\mathbb{P}\{|B_{n,n} - B_{0,n}| > n\tau\} < 2e^{-\frac{n\tau^2}{2}}, \qquad (21)$$

for any $n, \tau > 0$. Further, using a proof by contradiction one obtains the equality $\lim_{n\to\infty} \frac{1}{n} B_{0,n} = 0$[5] based on (20) and (21). Finally, the previous equality corroborated with (21) immediately implies that $\lim_{n\to\infty} \mathbb{P}\{B_{n,n} > n\delta\} = 0$ for all $\delta > 0$, which completes the proof.

## APPENDIX B
## PROOF OF THEOREM 2

Since $\mathcal{D}(d_0)$ is symmetric, it suffices to solve (14) for $\alpha \geq 1$. Note that

$$\mathbb{E}[d(X, X_t)] = \mathbb{E}[d(X_0 \oplus Z, X_t)]$$
$$= \mathbb{E}[\mathbb{E}[d(X_0 \oplus Z, X_t)|Z]]$$
$$= (1 - d_0)\mathbb{E}[d(X_0, X_t)] + d_0(1 - \mathbb{E}[d(X_0, X_t)])$$
$$= (1 - 2d_0)\mathbb{E}[d(X_0, X_t)] + d_0,$$

for $t = 1, 2$. Therefore problem (14) is equivalent to

$$\min_{p_{X_1 X_2|X_0}:I(X_1;X_2)=0} \alpha\mathbb{E}[d(X_0, X_1)] + \mathbb{E}[d(X_0, X_2)]. \quad (22)$$

To tackle this minimizaiton problem, we first fix the marginal distribution of $X_1$, in other words, let $p_{X_1}(0) = \epsilon$, for some $\epsilon, 0 < \epsilon < 1$, and solve the parameterized problem $P(\epsilon)$, then optimize over all possible values of $\epsilon$. In order to proceed let

---

[4]In the proof of Theorem 2 of [17] the quantity $I(X; f(\tilde{X}))$ has to be replaced by $I(X; \tilde{X})$.

[5]The proof of this relation proceeds as the proof of relation (74) in [16].

---

us adopt the following notation

$$y_1 = p_{X_0 X_1 X_2}(0, 1, 1), \quad y_2 = p_{X_0 X_1 X_2}(1, 0, 0),$$
$$y_3 = p_{X_0 X_1 X_2}(1, 1, 0), \quad y_4 = p_{X_0 X_1 X_2}(1, 0, 1),$$
$$y_5 = p_{X_0 X_1 X_2}(0, 1, 0), \quad y_6 = p_{X_0 X_1 X_2}(0, 0, 1),$$
$$y_7 = p_{X_0 X_1 X_2}(0, 0, 0), \quad y_8 = p_{X_0 X_1 X_2}(1, 1, 1).$$

Then

$$d_1^* \triangleq \mathbb{E}[d(X_0, X_1)] = y_1 + y_2 + y_4 + y_5,$$
$$d_2^* \triangleq \mathbb{E}[d(X_0, X_2)] = y_1 + y_2 + y_3 + y_6. \qquad (23)$$

Let $\mathbf{y} = (y_1, \cdots, y_8)$ and $\delta^* \triangleq p_{X_0}(0) = \frac{\delta - d_0}{1 - 2d_0} \leq 1/2$. Now problem (22) becomes

$$\min_{0 < \epsilon < 1} \min_{\mathbf{y}} C(\epsilon, \mathbf{y}) = (\alpha + 1)y_1 + (\alpha + 1)y_2$$
$$+ y_3 + \alpha y_4 + \alpha y_5 + y_6 \qquad (24)$$

$$\text{subject to } \sum_{i=1}^{8} y_i = 1 \qquad (25)$$
$$y_1 + y_5 + y_6 + y_7 = \delta^*$$
$$y_2 + y_4 + y_6 + y_7 = \epsilon$$
$$y_2 + y_7 = \frac{\epsilon}{1 - \epsilon}(y_3 + y_5)$$
$$y_i \geq 0, \quad 1 \leq i \leq 8.$$

The second constraint in (25) is obtained from $p_{X_0}(0) = \delta^*$, while the third follows from $p_{X_1}(0) = \epsilon$. The fourth is derived from $p_{X_1}(0)p_{X_2}(0) = p_{X_1 X_2}(0, 0)$, which holds due to the independence of $X_1$ and $X_2$. First we solve problem $P(\epsilon)$ for the case when $\epsilon \leq 1/2$. From the four equations in (25) one obtains

$$y_5 = -y_1 + y_2 + y_4 + \delta^* - \epsilon, \qquad (26)$$
$$y_8 = -y_2 - y_3 - y_4 + 1 - \delta^*,$$
$$y_6 = \frac{\epsilon}{1 - \epsilon}y_1 - \frac{\epsilon}{1 - \epsilon}y_2 - \frac{\epsilon}{1 - \epsilon}y_3 - \frac{1}{1 - \epsilon}y_4 + \frac{\epsilon(1 - \delta^*)}{1 - \epsilon},$$
$$y_7 = -\frac{\epsilon}{1 - \epsilon}y_1 + \left(\frac{\epsilon}{1 - \epsilon} - 1\right)y_2 + \frac{\epsilon}{1 - \epsilon}y_3 + \frac{\epsilon}{1 - \epsilon}y_4 +$$
$$\frac{\epsilon(\delta^* - \epsilon)}{1 - \epsilon}.$$

Plugging them into the cost function yields

$$C(\epsilon, \mathbf{y}) = \alpha(\delta^* - \epsilon) + \frac{\epsilon(1 - \delta^*)}{1 - \epsilon} + \left(1 + \frac{\epsilon}{1 - \epsilon}\right)y_1$$
$$+ \left(2\alpha + 1 - \frac{\epsilon}{1 - \epsilon}\right)y_2 + \left(1 - \frac{\epsilon}{1 - \epsilon}\right)y_3$$
$$+ \left(2\alpha - \frac{1}{1 - \epsilon}\right)y_4. \qquad (27)$$

It can be easily verified that the coefficients of all variables in (27) are non-negative when $\epsilon \leq 1/2$, and are strictly positive when $\epsilon < 1/2$. Thus, $C(\epsilon, \mathbf{y}) \geq g_1(\epsilon) \triangleq \alpha(\delta^* - \epsilon) + \frac{\epsilon(1 - \delta^*)}{1 - \epsilon}$, for any $\mathbf{y}$ satisfying (25). If, additionally, $\epsilon \leq \delta^*$, then the vector $\mathbf{y}$ obtained using (26) and $y_1 = y_2 = y_3 = y_4 = 0$, has non-negative components, hence it is a feasible solution achieving the lower bound $g_1(\epsilon)$, and therefore, it is a solution

to $P(\epsilon)$. Now denote by $C^*(\epsilon)$ the value of the cost function of $P(\epsilon)$ at optimality. The conclusion of the above discussion is stated next.

**Assertion 1.** Let $\alpha \geq 1$. For $0 < \epsilon \leq 1/2$, one has $C^*(\epsilon) \geq g_1(\epsilon)$. Additionally, when $0 < \epsilon \leq \delta^*$, one has $C^*(\epsilon) = g_1(\epsilon)$ and a solution to $P(\epsilon)$ is $\mathbf{y}(\epsilon)$ defined by

$$y_1(\epsilon) = y_2(\epsilon) = y_3(\epsilon) = y_4(\epsilon) = 0,$$
$$y_5(\epsilon) = \delta^* - \epsilon, \quad y_6(\epsilon) = \frac{\epsilon(1 - \delta^*)}{1 - \epsilon},$$
$$y_7(\epsilon) = \frac{\epsilon(\delta^* - \epsilon)}{1 - \epsilon}, \quad y_8(\epsilon) = 1 - \delta^*.$$

Furthermore, when $0 < \epsilon \leq \delta^*$ and $\epsilon \neq 1/2$, $\mathbf{y}(\epsilon)$ is the unique solution to $P(\epsilon)$.

The next result can be easily verified.

**Assertion 2.** For $1 \leq \alpha \leq 1/(1 - \delta^*)$,

$$\min_{0 < \epsilon \leq 1/2} g_1(\epsilon) = g_1(\epsilon_1),$$

where $\epsilon_1 \triangleq 1 - \sqrt{(1 - \delta^*)/\alpha} \in (0, 1/2)$, and $\epsilon_1$ is the unique point of minimum.

Next we will analyze the case when $\epsilon > 1/2$. By solving the variables $y_3, y_4, y_7, y_8$ from the four equalities in (25) and replacing them in the cost function, one obtains

$$C(\epsilon, \mathbf{y}) = \alpha(\epsilon - \delta^*) + \frac{\delta^*(1 - \epsilon)}{\epsilon} + \left(2\alpha + 1 - \frac{1 - \epsilon}{\epsilon}\right) y_1 + \left(1 + \frac{1 - \epsilon}{\epsilon}\right) y_2 + \left(2\alpha - \frac{1}{\epsilon}\right) y_5 + \left(1 - \frac{1 - \epsilon}{\epsilon}\right) y_6. \quad (28)$$

Because $1/2 < \epsilon < 1$ and $\alpha \geq 1$, the coefficients of variables in (28) are non-negative; therefore, we obtain that $C^*(\epsilon) \geq g_2(\epsilon) \triangleq \alpha(\epsilon - \delta^*) + \frac{\delta^*(1 - \epsilon)}{\epsilon}$. Further, the following assertion can be easily verified.

**Assertion 3.** Let $\alpha \geq 1$ and $1/2 < \epsilon < 1$. Then

$$C^*(\epsilon) \geq \min_{1/2 < \epsilon < 1} g_2(\epsilon) \geq \min_{0 < \epsilon < 1} g_2(\epsilon) = g_2(\epsilon_2),$$

where $\epsilon_2 \triangleq \sqrt{\delta^*/\alpha} \in (0, 1)$.

Finally, we need one more result, which is stated next.

**Assertion 4.** For $\delta < 1/2$, $g_1(\epsilon_1) < g_2(\epsilon_2)$ holds, while for $\delta = 1/2$, one has $g_1(\epsilon_1) = g_2(\epsilon_2)$.

*Proof of Assertion 4.* The nontrivial case is when $\delta < 1/2$. Relation $g_1(\epsilon_1) < g_2(\epsilon_2))$ is equivalent to $2\sqrt{\alpha}(\sqrt{1 - \delta^*} - \sqrt{\delta^*}) < (\alpha + 1)((1 - \delta^*) - \delta^*)$. By multiplying both sides of the above inequality with $1/(\sqrt{1 - \delta^*} - \sqrt{\delta^*})$, which is positive for $\delta < 1/2$, it reduces to $2\sqrt{\alpha} < (\alpha + 1)(\sqrt{1 - \delta^*} + \sqrt{\delta^*})$, which holds due to $0 < 2\sqrt{\alpha} \leq \alpha + 1$ and $1 < \sqrt{1 - \delta^*} + \sqrt{\delta^*}$. $\square$

Further, by combining Assertions 1-4, we conclude that when $1 \leq \alpha \leq 1/(1 - \delta^*)$ a solution to problem (24) is $(\epsilon_1, \mathbf{y}(\epsilon_1))$. Additionally, when $\delta < 1/2$ this solution is unique. This solution yields the following values for $d_1, d_2$ (via (23))

$$d_1 = d_0 + (1 - 2d_0)(\sqrt{1 - \delta^*}(1/\sqrt{\alpha} - \sqrt{1 - \delta^*})), \quad (29)$$
$$d_2 = d_0 + (1 - 2d_0)(\sqrt{1 - \delta^*}(\sqrt{\alpha} - \sqrt{1 - \delta^*})).$$

Further, eliminating $\alpha$ from the above relations leads to

$$(d_1 + 1 - 2d_0 - \delta)(d_2 + 1 - 2d_0 - \delta)$$
$$= (1 - d_0 - \delta)(1 - 2d_0). \quad (30)$$

Notice that for any pair $(d_1, d_2)$ with $d_0 \leq d_1 \leq d_2$, satisfying (30), the value of $\alpha$ for which equations (29) hold, is $\alpha = \frac{(1 - d_0 - \delta)(1 - 2d_0)}{(d_1 + 1 - 2d_0 - \delta)^2}$, which clearly satisfies $1 \leq \alpha \leq 1/(1 - \delta^*)$. Thus, we have proved that the set of solutions to problem (13) for all $1 \leq \alpha \leq 1/(1 - \delta^*)$, equals the portion of hyperbola (30) corresponding to $d_0 \leq d_1 \leq d_2$. The solution pair $(d_1, d_2)$ corresponding to $\alpha = 1/(1 - \delta^*)$ is $(d_1, d_2) = (d_0, \delta)$. Since the value of $d_1$ in the solution to (13) is non-increasing as $\alpha$ increases, and cannot go below $d_0$, it follows that for any $\alpha > 1/(1 - \delta^*)$ the unique solution to problem (13) is $(d_1, d_2) = (d_0, \delta)$. Using further the fact that $\mathcal{D}(d_0)$ is symmetric the first conclusion of Theorem 2 follows.

To complete the proof note that, for any pair $(d_1, d_2)$, with $d_1 \leq d_2$, satisfying (30), the joint probability distribution $p_{X_0 X_1 X_2}$ can be recovered from the solution vector $\mathbf{y}(\epsilon_1)$. Further, $p_{X X_1 X_2}$ can be evaluated using the Markov chain $X - X_0 - (X_1, X_2)$, leading to relations in Fig. 4. Additionally, since $\mathcal{D}(d_0)$ is symmetric and relations in Fig. 4 are symmetric in $(X_1, d_1)$ and $(X_2, d_2)$, one concludes that they also hold for $d_1 > d_2$.

We have already established the uniqueness of the conditional probability $p_{X_0 X_1 X_2 | X}$ when $\delta < 1/2$. For the case $\delta = 1/2$ the range of interest for $\alpha$ is $1 \leq \alpha \leq 1/(1 - \delta^*) = 2$. Then $\epsilon_2 = \sqrt{\frac{1}{2\alpha}} \geq 1/2$ and by Assertions 1-4, one has $g_2(\epsilon_2) = \min_{0 < \epsilon < 1} C^*(\epsilon)$. Note that when $\epsilon = \epsilon_2$ and $\alpha < 2$ the coefficients of $y_1, y_2, y_5, y_6$ in (28) are positive. Therefore, in order to achieve the minimum value $g_2(\epsilon_2)$ the variables $y_1, y_2, y_5, y_6$ must be set to 0. By letting $\epsilon = \epsilon_2$ and $y_1 = y_2 = y_5 = y_6 = 0$ in (25), one obtains the only other solution $(\epsilon_2, \mathbf{y}(\epsilon_2))$ to problem (24). It can be verified that the random variables $X_1'$ and $X_2'$ derived from $(\epsilon_2, \mathbf{y}(\epsilon_2))$ satisfy $p_{X X_1' X_2'}(x, x_1, x_2) = p_{X X_1 X_2}(x \oplus 1, x_1 \oplus 1, x_2 \oplus 1)$. $\square$

## APPENDIX C
### PROOF OF CLAIMS IN SECTION V

Using the fact that $X - X_t - X_c$, $t = 1, 2$, and $X_1 - (X, X_c) - X_2$ form Markov chains, one obtains

$$p_{X X_c X_1 X_2}(x, x_c, x_1, x_2)$$
$$= \frac{p_{X_1}(x_1) p_{X | X_1}(x | x_1) p_{X_c | X_1}(x_c | x_1) p_{X_2}(x_2) p_{X | X_2}(x | x_2) p_{X_c | X_2}(x_c | x_2)}{p_{X_c}(x_c) p_{X | X_c}(x | x_c)}.$$

Then the mapping $\psi_0(\cdot)$ which minimizes the central distortion must satisfy

$$\psi_0(x_c, x_1, x_2) = \arg \max_{x = 0, 1} p_{X X_c X_1 X_2}(x, x_c, x_1, x_2)$$
$$= \arg \max_{b = 0, 1} h(b),$$

where $h(b) \triangleq \frac{p_{X | X_1}(b | x_1) p_{X | X_2}(b | x_2)}{p_{X | X_c}(b | x_c)}$. Notice that $p_{X | X_t}(b | x_t) = p_{X_t | X}(x_t | b)$ and $p_{X | X_c}(b | x_c) = p_{X_c | X}(x_c | b)$. Using the fact that the function $\frac{1 - x}{x}$ is decreasing for $x > 0$, and that $\frac{1 - x}{x} \geq 1$ for $0 < x \leq 1/2$, combined with $0 < d_1 \leq d_2 \leq s$, it follows easily that $\frac{h(b)}{h(1 - b)} \geq 1$,

for the following cases: 1) $x_1 = x_2 = x_c = b$; 2) $x_1 = x_2 = b, x_c = 1 - b$; 3) $x_1 = b, x_2 = x_c = 1 - b$.

Consider now the case $x_1 = x_c = b$ and $x_2 = 1 - b$. Then $\frac{h(1-b)}{h(b)} = \frac{d_1}{1-d_1} \frac{1-d_2}{d_2} \frac{1-s}{s}$. Clearly, $\frac{h(1-b)}{h(b)} \geq 1$ holds if and only if $s \leq \alpha(d_1, d_2)$. Notice that $d_2 \leq \alpha(d_1, d_2) \leq 1/2$. Based on the above arguments one concludes that (16) is valid.

Now let us analyze the central distortion when $s \leq \alpha(d_1, d_2)$. The following relations hold

$$\mathbb{P}(X \neq \psi_0(X_c, X_1, X_2))$$
$$= \mathbb{P}(X_c = X) - \mathbb{P}(X = X_c = X_1 = X_2)$$
$$+ \mathbb{P}(X \neq X_c, X_c = X_1 = X_2)$$
$$= 1 - s - \frac{(1-d_1)(1-d_2)(1-s-d_1)(1-s-d_2)}{(1-s)(1-2d_1)(1-2d_2)}$$
$$+ \frac{d_1 d_2 (1-s-d_1)(1-s-d_2)}{s(1-2d_1)(1-2d_2)}.$$

Let $f(s)$ denote the last expression. The function $f(\cdot)$ is strictly convex over the interval $(d_2, \alpha(d_1, d_2))$ and $f(d_2) = f(\alpha(d_1, d_2))$. Then its unique point of minimum $s_0 \in (d_2, \alpha(d_1, d_2))$ is the solution to $f'(s) = 0$. The above equation is equivalent to $\beta(d_1, d_2)s^2(1-s)^2 + 2s(1-s) - 1 = 0$, whose only real solution smaller than $1/2$ is given by (18). $\square$

## REFERENCES

[1] L. Ozarow, "On a source coding problem with two channels and three receivers", *Bell Syst. Tech. J.*, vol. 59, pp. 1909-1921, 1980.

[2] J. K. Wolf, A. D. Wyner, and J. Ziv, "Source coding for multiple descriptions", *Bell Syst. Tech. J.*, vol. 59, pp. 1417-1426, 1980.

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

[4] 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.

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

[6] Z. Zhang and T. Berger, "Multiple description source coding with no excess marginal rate," *IEEE Trans. on Inform. Theory*, vol. 41, no. 2, pp. 502-521, March. 1995.

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

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

[9] V. K Goyal and J. Kovacevic, "Generalized multiple description coding with correlating transforms," *IEEE Trans. on Inform. Theory*, vol. 47, pp. 2199-2224, Sept. 2001.

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

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

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

[13] M. Wainwright and E. Maneva, "Lossy source compression using low-density generator matrix codes: analysis and algorithms," *IEEE Trans. on Inform. Theory*, vol. 56, no. 3, pp. 1351–1368, Mar. 2010.

[14] T. Filler and J. Friedrich, "Binary quantization using belief propagation with decimation over factor graphs of ldgm codes," *Proc. Allerton Conf. Comm., Control, and Computing*, Monticello, IL, Sept. 2007.

[15] M. J. Wainwright and E. Martinian, "Low-density graph codes that are optimal for binning and coding with side information", *IEEE Trans. on Inform. Theory*, vol. 55, no. 3, pp. 1061–1079, Mar. 2009.

[16] A. Gupta and S. Verdu, "Nonlinear sparse-graph codes for lossy compression", *IEEE Trans. on Inform. Theory*, vol. 55, no. 5, pp. 1961–1975, May 2009.

[17] 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.

[18] V. N. Koshelev, "Hierarchical coding of discrete sources," *Probl. Pered. Inform.*, vol. 16, pp. 31-49, 1980.

[19] W. H. R. Equitz and T. M. Cover, "Successive refinement of information," in *IEEE Trans. on Inform. Theory*, vol. 37, pp. 269-275, Mar. 1991.

[20] B. Rimoldi, "Successive refinement of information: Characterization of the achievable rates," *IEEE Trans. on Inform. Theory*, vol. 40, pp. 253-259, Jan. 1994.

[21] J. Chen, S. Dumitrescu, Y. Zhang, and J. Wang, "Robust multiresolution coding", *IEEE Trans. on Commun.*, vol. 58, no. 11, pp. 3186 - 3195, Nov. 2010.

[22] J. Chen, Y. Zhang, and S. Dumitrescu, "Gaussian multiple description coding with low-density generator matrix codes", *IEEE Trans. on Commun.*, vol. 60, no. 3, pp. 676 - 687, Mar. 2012.

[23] Y. Zhang, "Practical LDGM-based Multiple Description Coding," Ph.D. thesis, McMaster University, 2011.

[24] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd Ed., John Wiley & Sons, New Jersey, 2006.

[25] J. T. Pinkston, "An application of rate-distortion theory to a converse to the coding theorem," *IEEE Trans. on Inform. Theory*, vol. IT-15, pp. 269-275, Jan. 1969.

[26] 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.

[27] K. Azuma, "Weighted sums of certain dependent random variables", *Tohoku Math. J.*, vol. 19, pp. 357–367, 1967.

**Ying Zhang** received the B.Eng degree in Automation from Northeastern University, Shenyang, China in 2004, the M.Eng degree in Pattern Recognition and Intelligence System from Tianjin University, Tianjin, China in 2007 and the Ph.D. degree in Electrical and Computer Engineering from McMaster University, Hamilton, ON, Canada in 2011. Her research interests include information theory and multimedia communications.

**Sorina Dumitrescu** (M'05) received the B.A.Sc. and Ph. D. degrees in mathematics from the University of Bucharest, Romania, in 1990 and 1997, respectively.

From 2000 to 2002 she was a Postdoctoral Fellow in the Department of Computer Science at the University of Western Ontario, London, Canada. Since 2002 she has been with the Department of Electrical and Computer Engineering at McMaster University, Hamilton, Canada, where she held Postdoctoral, Research Associate, and Assistant Professor positions, and where she is currently an Associate Professor. Her current research interests include multimedia coding and communications, network-aware data compression, multiple description codes, joint source-channel coding, signal quantization. Her earlier research interests were in formal languages and automata theory. Dr. Dumitrescu has held an NSERC University Faculty Award during 2007-2012.

PLACE
PHOTO
HERE

**Jun Chen** (S'03-M'06) received the B.E. degree with honors in communication engineering from Shanghai Jiao Tong University, Shanghai, China, in 2001 and the M.S. and Ph.D. degrees in electrical and computer engineering from Cornell University, Ithaca, NY, in 2004 and 2006, respectively.

He was a Postdoctoral Research Associate in the Coordinated Science Laboratory at the University of Illinois at Urbana-Champaign, Urbana, IL, from 2005 to 2006, and a Postdoctoral Fellow at the IBM Thomas J. Watson Research Center, Yorktown Heights, NY, from 2006 to 2007. He is currently an Assistant Professor of Electrical and Computer Engineering at McMaster University, Hamilton, ON, Canada. He holds the Barber-Gennum Chair in Information Technology. His research interests include information theory, wireless communications, and signal processing.

He received several awards for his research, including the Josef Raviv Memorial Postdoctoral Fellowship in 2006, the Early Research Award from the Province of Ontario in 2010, and the IBM Faculty Award in 2010.

PLACE
PHOTO
HERE

**Zhibin Sun** received the B.Eng. and M.A.Sc. degrees in electrical engineering from McMaster University, Hamilton, ON, Canada, in 2007 and 2009, respectively.