

# Polar Codes for Multiple Descriptions

Qi Shi, Lin Song, Chao Tian, *Senior Member, IEEE*, Jun Chen, *Member, IEEE*,  
and Sorina Dumitrescu, *Senior Member, IEEE*

**Abstract**—A polar coding scheme is proposed for the multiple description coding (MDC) problem and is shown to be able to achieve a certain rate pair on the dominant line of the achievable rate region determined by El Gamal and Cover. This scheme is an adaptation of the one developed by Şaşıoğlu *et al.* for the multiple access channel (MAC) to the MDC setting. The analysis of the proposed scheme contains two new ingredients: 1) a certain MDC-MAC duality and 2) an auxiliary random process that involves both the mutual information and the Bhattacharyya parameter. The decorrelation effect of the polar transform is also investigated.

**Index Terms**—Duality, multiple access channel, multiple description coding, polar codes, rate region.

## I. INTRODUCTION

**P**OLAR codes, invented by Arikan [1], are “the first provably capacity-achieving codes for any symmetric input discrete memoryless channel that have low encoding and decoding complexity” [2]. Specifically, the encoding and decoding complexity of polar codes is of order  $O(n \log n)$ , where  $n$  is the code block length. Furthermore, Arikan and Telatar [3] upper-bounded the block error probability to order  $O(2^{-n^\beta})$ , where  $\beta \in (0, \frac{1}{2})$ . The most intriguing aspect of this new coding technique is that through recursive channel splitting and combining,  $n$  uses of the same memoryless channel are converted to successive uses of  $n$  different channels, and, asymptotically, each one of these converted channels is extremal in the sense that it is either a perfect channel or a pure noise channel. Subsequently, Korada and Urbanke proposed a lossy source coding scheme based on polar codes [4]; see also [2], [5]. More recently, Şaşıoğlu *et al.* [6] extended the polar coding technique to the two-user multiple access channel (MAC).

Manuscript received September 23, 2013; revised October 29, 2014; accepted October 29, 2014. Date of publication November 7, 2014; date of current version December 22, 2014. Q. Shi, L. Song, and J. Chen were supported in part by the Early Researcher Award through the Ontario Province and in part by the Natural Science and Engineering Research Council of Canada. This paper was presented at the 2014 IEEE International Symposium on Information Theory.

Q. Shi was with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON L8S 4L8, Canada. He is now with Inspur Group Company, Ltd., Jinan 250101, China (e-mail: shiqi@inspur.com).

L. Song was with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON L8S 4L8, Canada. She is now with the Institute of Network Coding, Chinese University of Hong Kong, Hong Kong (e-mail: linsong@inc.cuhk.edu.hk).

C. Tian was with AT&T Laboratories-Research, Bedminster, NJ 07921 USA. He is now with the University of Tennessee, Knoxville, TN 37996 USA (e-mail: chao.tian@utk.edu).

J. Chen and S. Dumitrescu are with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON L8S 4L8, Canada (e-mail: junchen@ece.mcmaster.ca; sorina@ece.mcmaster.ca).

Communicated by D. Burshtein, Associate Editor for Coding Techniques. Digital Object Identifier 10.1109/TIT.2014.2368561

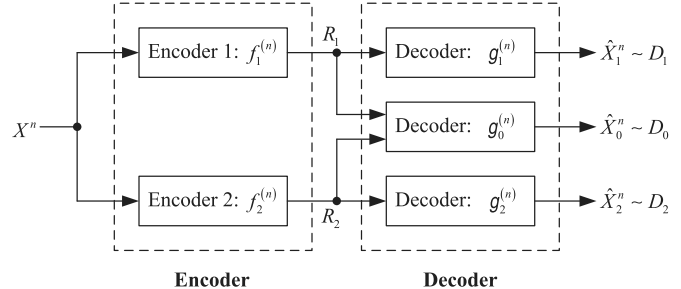


Fig. 1. The system diagram of multiple description coding.

Motivated by the success in [6], we apply the polar coding technique to the multiple description coding (MDC) problem. In MDC (see Fig. 1), a single source  $X$  is encoded into two descriptions of rate  $R_1$  and rate  $R_2$ , respectively, such that the reconstruction distortion based on description  $i$  is  $D_i$ ,  $i = 1, 2$ , and the reconstruction distortion based on both descriptions is  $D_0$ . The goal is to find efficient coding schemes to achieve the optimal tradeoff between  $(R_1, R_2)$  and  $(D_0, D_1, D_2)$ . Unfortunately, the optimal rate-distortion tradeoff in MDC is unknown except for certain special cases. In this work we focus on the achievable rate pairs subject to distortion constraints  $(D_0, D_1, D_2)$  determined by El Gamal and Cover [8] (sometimes referred to as the EGC rate region) and propose a MDC scheme based on polar codes.

We show that the proposed scheme can achieve a certain rate pair on the dominant line of the EGC rate region. The analysis of the proposed scheme is substantially facilitated by a certain MDC-MAC duality, which enables us to translate several results in [6] and [7, Sec. 6] to the MDC setting. However, the existing analysis techniques in the MAC setting do not lead to an effective bound on one particular mutual information term, which is critical in the MDC setting. This problem is resolved via the analysis of an auxiliary random process obtained by adding the Bhattacharyya parameter to the aforementioned mutual information. To the best of our knowledge, this is the first case in polar coding that such an analysis is used, and this technique may be of independent interest to the readers. Our analysis also reveals a decorrelation effect of the polar transform. Specifically, it is shown that, under certain mild conditions, the polar transform can asymptotically remove the symbol-wise dependency between the two descriptions (which should be contrasted with the fact that the block-wise dependency is preserved).

The remainder of this paper is organized as follows. We establish a certain MDC-MAC duality in Section II.

A joint polarization scheme for the MDC problem can be found in Section III. The decorrelation effect of the polar transform is investigated in Section IV. We conclude the paper in Section V.

For any positive integer  $q$ , we use  $\oplus$  and  $\ominus$  to denote modulo- $q$  addition and subtraction, respectively, where the value of  $q$  should be clear from the context. The cardinality of a set  $\mathcal{S}$  is denoted by  $|\mathcal{S}|$ . We use  $\|\cdot\|$  to denote the  $L_1$  norm. Throughout this paper, the logarithm function is to base 2 unless stated otherwise.

## II. MDC-MAC DUALITY

Let  $\{X_i : i \geq 1\}$  be an i.i.d. random process with marginal distribution  $p_X(x)$  defined over finite alphabet  $\mathcal{X}$ . Let  $d_j : \mathcal{X} \times \hat{\mathcal{X}}_j \rightarrow [0, d_{\max}]$ ,  $j = 0, 1, 2$ , be bounded distortion measures, where  $\hat{\mathcal{X}}_j$ ,  $j = 0, 1, 2$ , are the reconstruction alphabets.

*Definition 1:* A rate-distortion tuple  $(R_1, R_2, D_0, D_1, D_2)$  is said to be achievable if, for every  $\epsilon > 0$ , there exist encoding functions  $f_j^{(n)} : \mathcal{X}^n \rightarrow \mathcal{C}_j$ ,  $j = 1, 2$ , and decoding functions  $g_0^{(n)} : \mathcal{C}_1 \times \mathcal{C}_2 \rightarrow \hat{\mathcal{X}}_0^n$  and  $g_j^{(n)} : \mathcal{C}_j \rightarrow \hat{\mathcal{X}}_j^n$ ,  $j = 1, 2$ , such that

$$\begin{aligned} \frac{1}{n} \log |\mathcal{C}_j| &\leq R_j + \epsilon, \quad j = 1, 2, \\ \frac{1}{n} \sum_{i=1}^n \mathbb{E}[d_i(X_i, \hat{X}_{j,i})] &\leq D_j + \epsilon, \quad j = 0, 1, 2, \end{aligned}$$

where  $\hat{X}_0^n = g_0^{(n)}(f_1^{(n)}(X^n), f_2^{(n)}(X^n))$  and  $\hat{X}_i^n = g_i^{(n)}(f_i^{(n)}(X^n))$ ,  $i = 1, 2$ .

The following result provides a sufficient condition on the achievability of  $(R_1, R_2, D_0, D_1, D_2)$ . It was shown in [9] that the convex hull of this achievable region is equivalent to the one derived by El Gamal and Cover [8].

*Theorem 1:* A rate-distortion tuple  $(R_1, R_2, D_0, D_1, D_2)$  is achievable if there exist some probability mass function  $p_X(x)p(y, z|x)$  and deterministic mappings  $\phi_j$ ,  $j = 0, 1, 2$ , such that

$$\begin{aligned} R_1 &\geq I(X; Y), \\ R_2 &\geq I(X; Z), \\ R_1 + R_2 &\geq I(X; Y, Z) + I(Y; Z), \\ D_0 &\geq \mathbb{E}[d_0(X, \phi_0(Y, Z))], \\ D_1 &\geq \mathbb{E}[d_1(X, \phi_1(Y))], \\ D_2 &\geq \mathbb{E}[d_2(X, \phi_2(Z))]. \end{aligned}$$

*Remark:* The random variables  $Y$  and  $Z$  can be understood as the (single-letter) representations of the source  $X$  in the two descriptions, and the decoding functions  $\phi_j$ ,  $j = 0, 1, 2$ , reconstruct  $X$  based on these representations.

Let  $q_j$  be a prime number and define  $\mathbb{F}_{q_j} = \{0, 1, \dots, q_j - 1\}$ ,  $j = 1, 2$ . In this work we focus on a generic probability mass function  $P(x, y, z) = p_X(x)p(y, z|x)$  with the property that the induced random variables  $Y$  and  $Z$  are uniformly distributed over  $\mathbb{F}_{q_1}$  and  $\mathbb{F}_{q_2}$ , respectively. Such a restriction is justified by the fact that every random variable can be approximated arbitrarily well by a random variable uniformly distributed over a sufficiently large alphabet through

a deterministic mapping. We shall write  $P(x, y, z)$  simply as  $P$  when no ambiguity arises.

A probability mass function  $Q(s, u, v)$  is referred to as a MDC distribution if the induced random variables  $U$  and  $V$  are uniformly distributed over  $\mathbb{F}_{q_1}$  and  $\mathbb{F}_{q_2}$ , respectively. For any MDC distribution  $Q$ , define

$$\begin{aligned} I^{(0)}(Q) &= I(S; U, V) + I(U; V), \\ I^{(1)}(Q) &= I(S; U), \quad \overline{I^{(1)}}(Q) = I(S, U; V), \\ I^{(2)}(Q) &= I(S; V), \quad \overline{I^{(2)}}(Q) = I(S, V; U), \\ \mathcal{K}(Q) &= (I^{(0)}(Q), I^{(1)}(Q), I^{(2)}(Q)), \\ \mathcal{J}(Q) &= \{(R_1, R_2) : R_1 + R_2 \geq I^{(0)}(Q), \\ &\quad R_j \geq I^{(j)}(Q), j = 1, 2\}. \end{aligned}$$

Note that

$$I^{(0)}(Q) = I^{(j)}(Q) + \overline{I^{(j)}}(Q), \quad j = 1, 2. \quad (1)$$

Moreover, since  $I^{(0)}(Q) \geq I^{(1)}(Q) + I^{(2)}(Q)$ , the set

$$\mathcal{F}(Q) \triangleq \{(R_1, R_2) \in \mathcal{J}(Q) : R_1 + R_2 = I^{(0)}(Q)\}$$

is non-empty. It is clear that  $P$  is a MDC distribution. We shall refer to  $\mathcal{J}(P)$  as the EGC rate region and  $\mathcal{F}(P)$  as its dominant line.

*Definition 2:* Given a MDC distribution  $Q$  over  $\mathcal{S} \times \mathbb{F}_{q_1} \times \mathbb{F}_{q_2}$ , let  $Q^-$  and  $Q^+$  be two new MDC distributions over  $\mathcal{S}^2 \times \mathbb{F}_{q_1} \times \mathbb{F}_{q_2}$  and  $(\mathcal{S}^2 \times \mathbb{F}_{q_1} \times \mathbb{F}_{q_2}) \times \mathbb{F}_{q_1} \times \mathbb{F}_{q_2}$ , respectively, such that

$$\begin{aligned} Q^-((s_1, s_2), u_1, v_1) &= \sum_{u_2 \in \mathbb{F}_{q_1}} \sum_{v_2 \in \mathbb{F}_{q_2}} Q(s_2, u_2, v_2) Q(s_1, u_1 \oplus u_2, v_1 \oplus v_2), \\ Q^+((s_1, s_2, u_1, v_1), u_2, v_2) &= Q(s_2, u_2, v_2) Q(s_1, u_1 \oplus u_2, v_1 \oplus v_2), \end{aligned}$$

where we have written  $(s_1, s_2)$  in parentheses to indicate that they should be viewed as a single vector, and similarly for  $(s_1, s_2, u_1, v_1)$ .

Applying this process to  $P$  results in two new MDC distributions  $P^-$  and  $P^+$ . We can obtain  $2^k$  MDC distributions

$$P^{-\dots-}, \dots, P^{+\dots+}$$

by repeating this  $k$  times. Similarly to [3] and [6], let  $S_1, S_2, \dots$ , be an i.i.d. sequence of random variables taking values in the set  $\{-, +\}$  with  $\Pr(S_i = -) = \Pr(S_i = +) = \frac{1}{2}$ . Define

$$\begin{aligned} P_0 &= P, \\ P_k &= P_{k-1}^{S_k}, \quad k \geq 1. \end{aligned}$$

We shall characterize the polarization behavior of  $\{I^{(j)}(P_k) : k \geq 0\}$ ,  $j = 0, 1, 2$ , and  $\{\overline{I^{(j)}}(P_k) : k \geq 0\}$ ,  $j = 1, 2$  by leveraging the corresponding results in the MAC setting via a certain MDC-MAC duality. Given random variables  $(X, Y, Z)$  induced by the generic probability mass function  $P(x, y, z)$  in the MDC problem, it is tempting to view  $P(x|y, z)$  as a MAC with inputs  $(Y, Z)$  and output  $X$ . Unfortunately, in the MAC setting the two inputs are required to be independent whereas  $Y$  and  $Z$  in general are not in MDC.

Nevertheless, one can create desired independence via a dithering step. Let  $Z'$  be a random variable uniformly distributed over  $\mathbb{F}_{q_2}$  and independent of  $(X, Y, Z)$ . Define  $\tilde{X} = (X, Z')$ ,  $\tilde{Y} = Y$ , and  $\tilde{Z} = Z \oplus Z'$ . Note that  $\tilde{Y}$  and  $\tilde{Z}$  are independent. Therefore, we can view  $\tilde{P}$  as a MAC with inputs  $(\tilde{Y}, \tilde{Z})$  and output  $\tilde{X}$ , where  $\tilde{P}$  the conditional distribution of  $\tilde{X}$  given  $(\tilde{Y}, \tilde{Z})$ . It is clear that

$$I(\tilde{X}; \tilde{Y}) = I(X; Y), \quad (2)$$

$$I(\tilde{X}, \tilde{Z}; \tilde{Y}) = I(X, Z; Y). \quad (3)$$

Moreover, we have

$$\begin{aligned} I(\tilde{X}; \tilde{Z}) &= I(X, Z'; Z \oplus Z') \\ &= H(Z) - H(Z \oplus Z' | X, Z') \\ &= H(Z) - H(Z | X) \\ &= I(X; Z) \end{aligned} \quad (4)$$

and

$$\begin{aligned} I(\tilde{X}, \tilde{Y}; \tilde{Z}) &= I(X, Z', Y; Z \oplus Z') \\ &= H(Z) - H(Z \oplus Z' | X, Z', Y) \\ &= H(Z) - H(Z | X, Y) \\ &= I(X, Y; Z). \end{aligned} \quad (5)$$

For any two-user MAC  $\tilde{Q}$  with input alphabets  $\mathbb{F}_{q_1}$  and  $\mathbb{F}_{q_2}$ , define

$$\begin{aligned} I^{(0)}(\tilde{Q}) &= I(\tilde{U}, \tilde{V}; \tilde{S}), \\ I^{(1)}(\tilde{Q}) &= I(\tilde{U}; \tilde{S}), \quad \overline{I^{(1)}}(\tilde{Q}) = I(\tilde{V}; \tilde{S}, \tilde{U}), \\ I^{(2)}(\tilde{Q}) &= I(\tilde{S}; \tilde{V}), \quad \overline{I^{(2)}}(\tilde{Q}) = I(\tilde{U}; \tilde{S}, \tilde{V}), \\ \mathcal{J}(\tilde{Q}) &= \{(R_1, R_2) \in \mathbb{R}_+^2 : R_1 + R_2 \leq I^{(0)}(\tilde{Q}), \\ &\quad R_j \leq \overline{I^{(j)}}(\tilde{Q}), j = 1, 2\}, \end{aligned}$$

where  $\tilde{S}$  is the output of MAC  $\tilde{Q}$  generated by independent inputs  $\tilde{U}$  and  $\tilde{V}$  uniformly distributed over  $\mathbb{F}_{q_1}$  and  $\mathbb{F}_{q_2}$ , respectively. Note that

$$I^{(0)}(\tilde{Q}) = I^{(j)}(\tilde{Q}) + \overline{I^{(j)}}(\tilde{Q}), \quad j = 1, 2. \quad (6)$$

Given a MDC distribution  $Q$ , we say  $\tilde{Q}$  is a dual MAC<sup>1</sup> of  $Q$  (see Fig. 2) if<sup>2</sup>

$$\begin{aligned} I^{(j)}(Q) &= I^{(j)}(\tilde{Q}), \quad j = 0, 1, 2, \\ \overline{I^{(j)}}(Q) &= \overline{I^{(j)}}(\tilde{Q}), \quad j = 1, 2. \end{aligned}$$

It follows by (2)–(5) that  $\tilde{P}$  is a dual MAC of  $P$ .

**Definition 3:** Given a MAC  $\tilde{Q} : \mathbb{F}_{q_1} \times \mathbb{F}_{q_2} \rightarrow \tilde{S}$ , let  $\tilde{Q}^- : \mathbb{F}_{q_1} \times \mathbb{F}_{q_2} \rightarrow \tilde{S}^2$  and  $\tilde{Q}^+ : \mathbb{F}_{q_1} \times \mathbb{F}_{q_2} \rightarrow \tilde{S}^2 \times \mathbb{F}_{q_1} \times \mathbb{F}_{q_2}$  be two new MACs such that

$$\begin{aligned} \tilde{Q}^-((\tilde{s}_1, \tilde{s}_2) | \tilde{u}_1, \tilde{v}_1) \\ &= \frac{1}{q_1 q_2} \sum_{\tilde{u}_2 \in \mathbb{F}_{q_1}} \sum_{\tilde{v}_2 \in \mathbb{F}_{q_2}} \tilde{Q}(\tilde{s}_2 | \tilde{u}_2, \tilde{v}_2) \tilde{Q}(\tilde{s}_1 | \tilde{u}_1 \oplus \tilde{u}_2, \tilde{v}_1 \oplus \tilde{v}_2), \end{aligned}$$

$$\begin{aligned} \tilde{Q}^+((\tilde{s}_1, \tilde{s}_2, \tilde{u}_1, \tilde{v}_1) | \tilde{u}_2, \tilde{v}_2) \\ &= \frac{1}{q_1 q_2} \tilde{Q}(\tilde{s}_2 | \tilde{u}_2, \tilde{v}_2) \tilde{Q}(\tilde{s}_1 | \tilde{u}_1 \oplus \tilde{u}_2, \tilde{v}_1 \oplus \tilde{v}_2). \end{aligned}$$

<sup>1</sup>For a given MDC distribution  $Q$ , its dual MAC is not unique.

<sup>2</sup>In view of (1) and (6),  $I^{(0)}(Q) = I^{(0)}(\tilde{Q})$  is implied by  $I^{(j)}(Q) = I^{(j)}(\tilde{Q})$  and  $\overline{I^{(j)}}(Q) = \overline{I^{(j)}}(\tilde{Q})$ ,  $j = 1, 2$ .

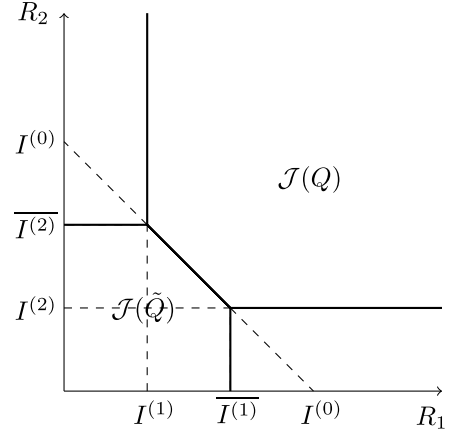


Fig. 2. The duality between MDC distribution  $Q$  and MAC  $\tilde{Q}$ , where  $I^{(j)} = I^{(j)}(Q) = I^{(j)}(\tilde{Q})$ ,  $j = 0, 1, 2$ , and  $\overline{I^{(j)}} = \overline{I^{(j)}}(Q) = \overline{I^{(j)}}(\tilde{Q})$ ,  $j = 1, 2$ .

Applying this process to  $\tilde{P}$  results in two new MACs  $\tilde{P}^-$  and  $\tilde{P}^+$ . We can obtain  $2^k$  MACs

$$\tilde{P}^{\dots}, \dots, \tilde{P}^{\dots+}$$

by repeating this  $k$  times. Following the derivation of (2)–(5), one can verify that

$$I^{(j)}(P^{s^k}) = I^{(j)}(\tilde{P}^{s^k}), \quad j = 0, 1, 2, \quad (7)$$

$$\overline{I^{(j)}}(P^{s^k}) = \overline{I^{(j)}}(\tilde{P}^{s^k}), \quad j = 1, 2, \quad (8)$$

for any  $s^k \in \{-, +\}^k$ . Therefore, with our specific choice of  $\tilde{P}$ , the duality is preserved under the polarization procedure. Define

$$\begin{aligned} \tilde{P}_0 &= \tilde{P}, \\ \tilde{P}_k &= \tilde{P}_{k-1}^{S_k}, \quad k \geq 1. \end{aligned}$$

It is shown in [6] that

$$\mathbb{E}[I^{(0)}(\tilde{P}_k)] = I^{(0)}(\tilde{P}), \quad k \geq 0,$$

which, together with (7), implies

$$\mathbb{E}[I^{(0)}(P_k)] = I^{(0)}(P), \quad k \geq 0. \quad (9)$$

One can obtain the following result by translating [6, Th. 2] and [7, Sec. 6.1.2] to the MDC setting via (7) and (8). It provides a characterization of the polarization behavior  $\{I^{(j)}(P_k) : k \geq 0\}$ ,  $j = 0, 1, 2$ . Note that the polarization behavior of  $\{\overline{I^{(j)}}(P_k) : k \geq 0\}$ ,  $j = 1, 2$ , is determined by that of  $\{I^{(j)}(P_k) : k \geq 0\}$ ,  $j = 0, 1, 2$ , via (1).

**Theorem 2:** The process  $\{\mathcal{K}(P_k) : k \geq 0\}$  converges almost surely.

- 1) When  $q_1$  and  $q_2$  are non-identical primes, the limit of  $\{\mathcal{K}(P_k) : k \geq 0\}$  belongs to the set

$$\{(0, 0, 0), (\Delta_1, \Delta_1, 0), (\Delta_2, 0, \Delta_2), (\Delta_1 + \Delta_2, \Delta_1, \Delta_2)\}$$

with probability 1, where  $\Delta_j = \log q_j$ ,  $j = 1, 2$ .

- 2) When  $q_1 = q_2 = q$  for some prime number  $q$ , the limit of  $\{\mathcal{K}(P_k) : k \geq 0\}$  belongs to the set

$$\{(0, 0, 0), (\Delta, \Delta, 0), (\Delta, 0, \Delta), (2\Delta, \Delta, \Delta), (\Delta, 0, 0)\}$$

with probability 1, where  $\Delta = \log q$ .

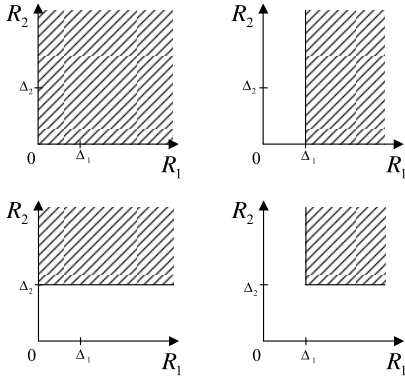


Fig. 3. The four extremal rate regions when  $q_1 \neq q_2$ .

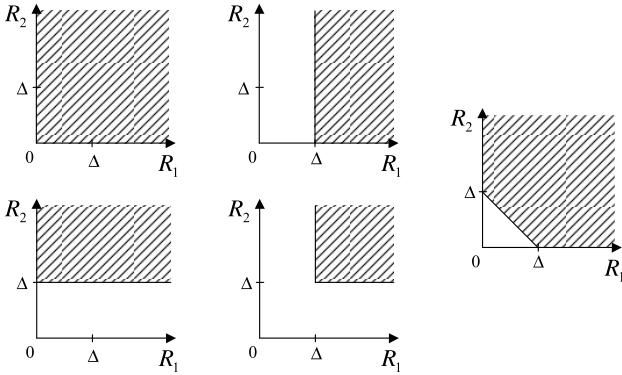


Fig. 4. The five extremal rate regions when  $q_1 = q_2$ .

Theorem 2 implies that, when  $q_1 \neq q_2$ , for most of  $P^{s^k}$ ,  $s^k \in \{-, +\}^k$ , the rate region  $\mathcal{J}(P^{s^k})$  approaches one of the four extremal ones, as depicted in Fig. 3. The rate region associated with  $(0, 0, 0)$  corresponds to the case that both descriptions reveal nothing about the source, thus are useless; the rate regions associated with  $(\Delta_1, \Delta_1, 0)$  and  $(\Delta_2, 0, \Delta_2)$  correspond to the cases that one description is useless, while the other is informative; the last case  $(\Delta_1 + \Delta_2, \Delta_1, \Delta_2)$  is when both descriptions are informative. From a coding perspective,  $(0, 0, 0)$  corresponds to the case that both coded symbols can be set arbitrarily, as long as they are revealed to both the encoder and the decoder;  $(\Delta_1, \Delta_1, 0)$  and  $(\Delta_2, 0, \Delta_2)$  correspond to the cases that the coded symbol in one of the descriptions is fully determined by the source vector, while the other can be set arbitrarily;  $(\Delta_1 + \Delta_2, \Delta_1, \Delta_2)$  corresponds to the case that both symbols are fully determined. Note that we have an additional extremal case<sup>3</sup>  $(\Delta, 0, 0)$  when  $q_1 = q_2$  (see Fig. 4), corresponding to the scenario where the two descriptions jointly are informative whereas each individual description by itself is useless. From a coding perspective, this case means that any one symbol in the two descriptions can be set arbitrarily, and the other is then fully determined. This will become clearer in Section III when we describe the coding scheme in more details.

<sup>3</sup>This extremal case can indeed arise. For example, let  $X = Y \oplus Z$ , where  $Y$  and  $Z$  are mutually independent and uniformly distributed over  $\mathbb{F}_q$ ; it can be verified that  $\mathcal{K}(P^{s^k}) = (\Delta, 0, 0)$  for all for  $s^k \in \{-, +\}^k$ .

### III. POLAR CODES FOR MULTIPLE DESCRIPTIONS WITH JOINT POLARIZATION

#### A. $q_1 \neq q_2$

Now we proceed to describe the polar coding algorithm for the case  $q_1 \neq q_2$ . Let  $(X_i, Y_i, Z_i)$ ,  $i = 1, \dots, n$ , be  $n$  independent copies of  $(X, Y, Z)$  distributed according to  $P(x, y, z)$ , where  $n = 2^k$  is the code length. Let  $B_n$  denote the  $n \times n$  “bit reversal” permutation matrix in [1], and let  $G_n = G_1^{\otimes k}$  be the  $k$ -th power Kronecker product of the matrix

$$G_1 \triangleq \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}.$$

Define  $P_{(i)}$  to be the joint distribution of  $((X^n, U^{i-1}, V^{i-1}), U_i, V_i)$ ,  $i = 1, \dots, n$ , where  $U^n$  and  $V^n$  are defined via the polar transform:

$$U^n = Y^n (B_n G_n)^{-1}, \quad V^n = Z^n (B_n G_n)^{-1}. \quad (10)$$

Note that [1] (see also [6])

$$P_{(1)} = P^{-\dots-}, \dots, P_{(n)} = P^{+\dots+}.$$

Theorem 2 ensures that for almost all  $P_{(i)}$ 's, the triple  $\mathcal{K}(P_{(i)})$  is close to one of the extremals when  $n$  is sufficiently large.

We shall partition the indices of  $U^n$  into two sets (a frozen set and an information set), and similarly for  $V^n$ . More precisely, fix some small  $\epsilon > 0$  and let  $\mathcal{F}_1$  and  $\mathcal{F}_2$  be subsets of  $\{1, 2, \dots, n\}$  defined as follows.

- 1) If  $\overline{I^{(2)}}(P_{(i)}) \leq \epsilon$ , then  $i \in \mathcal{F}_1$ ; otherwise,  $i \notin \mathcal{F}_1$ .
- 2) If  $\overline{I^{(1)}}(P_{(i)}) \leq \epsilon$ , then  $i \in \mathcal{F}_2$ ; otherwise,  $i \notin \mathcal{F}_2$ .

We denote the joint distribution of  $(X^n, U^n, V^n)$  as  $p_{X^n, U^n, V^n}(x^n, u^n, v^n)$ , which is specified by the product distribution  $\prod_{i=1}^n P(x_i, y_i, z_i)$  and the polar transform in (10); the marginals of  $p_{X^n, U^n, V^n}(x^n, u^n, v^n)$  are written in a similar way; when clear from the context, we shall omit the subscripts.

For each  $i \in \mathcal{F}_1$ , randomly set the value of  $u_i$  according to the uniform distribution over  $\mathbb{F}_{q_1}$ ; similarly, for each  $i \in \mathcal{F}_2$ , randomly set the value of  $v_i$  according to the uniform distribution over  $\mathbb{F}_{q_2}$ ;  $u_{\mathcal{F}_1} \triangleq (u_i)_{i \in \mathcal{F}_1}$  and  $v_{\mathcal{F}_2} \triangleq (v_i)_{i \in \mathcal{F}_2}$  are referred to as the frozen symbols of  $u^n$  and  $v^n$ , respectively, and are revealed to both the encoder and the decoder. We define the randomized encoding function  $f_{u_{\mathcal{F}_1}, v_{\mathcal{F}_2}}^{(n)} : \mathcal{X}^n \rightarrow \mathbb{F}_{q_1}^{n-|\mathcal{F}_1|} \times \mathbb{F}_{q_2}^{n-|\mathcal{F}_2|}$  as follows.

For  $i = 1, 2, \dots, n$ , if  $i \notin \mathcal{F}_1$ , then  $u_i$  takes value  $a \in \mathbb{F}_{q_1}$  with probability

$$\frac{p_{X^n, U^i, V^{i-1}}(x^n, (u^{i-1}, a), v^{i-1})}{p_{X^n, U^{i-1}, V^{i-1}}(x^n, u^{i-1}, v^{i-1})};$$

if  $i \notin \mathcal{F}_2$ , then  $v_i$  takes value  $a \in \mathbb{F}_{q_2}$  with probability

$$\frac{p_{X^n, U^i, V^i}(x^n, u^i, (v^{i-1}, a))}{p_{X^n, U^i, V^{i-1}}(x^n, u^i, v^{i-1})}.$$

Descriptions 1 and 2 consist of  $u_{\mathcal{F}_1^c} \triangleq (u_i)_{i \notin \mathcal{F}_1}$  and  $v_{\mathcal{F}_2^c} \triangleq (v_i)_{i \notin \mathcal{F}_2}$ , respectively.

With only the first description, the decoder forms  $y^n = u^n B_n G_n$ ; it further applies  $\phi_1$  in Theorem 1 to each symbol of  $y^n$  individually and then concatenates the outputs

as the reconstruction. Similarly, with only the second description, the decoder forms  $z^n = v^n B_n G_n$ ; it further applies  $\phi_2$  in Theorem 1 to  $z^n$  and then concatenates the outputs as the reconstruction. When both descriptions are available, the decoder can apply  $\phi_0$  in Theorem 1 to  $(y_i, z_i), i = 1, 2, \dots, n$ , and then concatenate the outputs as the reconstruction.

The detailed performance analysis can be found in Appendix A, where it is shown that the proposed scheme can asymptotically achieve a certain rate pair on the dominant line of the EGC rate region.

### B. $q_1 = q_2$

Now consider the case<sup>4</sup> where  $q_1 = q_2 = q$  for some prime number  $q$ . Due to the existence of the additional extremal case  $(\Delta, 0, 0)$ , the polar coding scheme and the associated performance analysis need to be modified. In particular, we use  $\eta(i)$  to indicate whether or not  $\mathcal{K}(P_{(i)})$  is close to  $(\Delta, 0, 0), i = 1, 2, \dots, n$ . Specifically, we fix some small  $\epsilon > 0$  and define

$$\eta(i) = \begin{cases} 1, & \|\mathcal{K}(P_{(i)}) - (\Delta, 0, 0)\| \leq \epsilon \\ 0, & \text{otherwise.} \end{cases}$$

Further define the frozen sets  $\mathcal{F}_1 \subseteq \{1, 2, \dots, n\}$  and  $\mathcal{F}_2 \subseteq \{1, 2, \dots, n\}$  as follows.

- 1)  $\eta(i) = 0$ : if  $\overline{I^{(2)}}(P_{(i)}) \leq \epsilon$ , then  $i \in \mathcal{F}_1$ , otherwise,  $i \notin \mathcal{F}_1$ ; if  $\overline{I^{(1)}}(P_{(i)}) \leq \epsilon$ , then  $i \in \mathcal{F}_2$ , otherwise,  $i \notin \mathcal{F}_2$ .
- 2)  $\eta(i) = 1$ :  $i \in \mathcal{F}_1$  and  $i \notin \mathcal{F}_2$ .

With this new definition of  $\mathcal{F}_1$  and  $\mathcal{F}_2$ , the encoding operation remains the same as that in Section III-A.

In contrast to the other four extremal cases where the problem largely boils down to the point-to-point setting, the case  $(\Delta, 0, 0)$  is a non-degenerate one. For this reason, the performance analysis becomes more difficult. In fact, one encounters a similar situation in the MAC setting. However, there is a crucial difference: to handle the additional extremal case  $(\Delta, 0, 0)$ , one needs to have an effective bound on the rate of polarization for  $\{I^{(j)}(P_k), k \geq 0\}, j = 1, 2$ , in MDC whereas such a bound is only available for  $\{I^{(j)}(\tilde{P}_k), k \geq 0\}, j = 1, 2$ , in the dual MAC setting [6].

The following theorem provides the main analytical tool for dealing with this additional case. Its proof hinges upon the analysis of an auxiliary random process that involves both the mutual information and the Bhattacharyya parameter.

*Theorem 3:* For any  $\beta \in (0, \frac{1}{2})$ ,

$$\begin{aligned} & \lim_{k \rightarrow \infty} \Pr \left( \|\mathcal{K}(P_k) - (\Delta, 0, 0)\| \leq 2^{-2^{k\beta}} \right) \\ &= \Pr \left( \lim_{k \rightarrow \infty} \mathcal{K}(P_k) = (\Delta, 0, 0) \right) \\ &= \lim_{k \rightarrow \infty} \Pr \left( \|\mathcal{K}(P_k) - (\Delta, 0, 0)\| \leq 2^{-2^{k\beta}}, \right. \\ & \quad \left. \lim_{k' \rightarrow \infty} \mathcal{K}(P_{k'}) = (\Delta, 0, 0) \right). \end{aligned}$$

<sup>4</sup>This case can in principle be circumvented in view of the fact that random variables with identical alphabets can be approximated by those with non-identical alphabets. However, as we shall show next, expanding the alphabet size to make such approximations is in fact unnecessary, which is also less desirable from a practical perspective.

*Proof:* It suffices to show that for any  $\beta \in (0, \frac{1}{2})$

$$\begin{aligned} & \lim_{k \rightarrow \infty} \Pr \left( |I^{(0)}(P_k) - \Delta| \leq 2^{-2^{k\beta}}, \lim_{k' \rightarrow \infty} \mathcal{K}(P_{k'}) = (\Delta, 0, 0) \right) \\ &= \lim_{k \rightarrow \infty} \Pr \left( I^{(j)}(P_k) \leq 2^{-2^{k\beta}}, \lim_{k' \rightarrow \infty} \mathcal{K}(P_{k'}) = (\Delta, 0, 0) \right) \\ &= \Pr \left( \lim_{k \rightarrow \infty} \mathcal{K}(P_k) = (\Delta, 0, 0) \right), \quad j = 1, 2. \end{aligned} \quad (11)$$

It follows by (7) and (8) that (11) is equivalent to

$$\begin{aligned} & \lim_{k \rightarrow \infty} \Pr \left( |I^{(0)}(\tilde{P}_k) - \Delta| \leq 2^{-2^{k\beta}}, \lim_{k' \rightarrow \infty} \mathcal{K}(\tilde{P}_{k'}) = (\Delta, 0, 0) \right) \\ &= \lim_{k \rightarrow \infty} \Pr \left( I^{(j)}(\tilde{P}_k) \leq 2^{-2^{k\beta}}, \lim_{k' \rightarrow \infty} \mathcal{K}(\tilde{P}_{k'}) = (\Delta, 0, 0) \right) \\ &= \Pr \left( \lim_{k \rightarrow \infty} \mathcal{K}(\tilde{P}_k) = (\Delta, 0, 0) \right), \quad j = 1, 2. \end{aligned} \quad (12)$$

It is known [6] that

$$\begin{aligned} & \lim_{k \rightarrow \infty} \Pr \left( |\overline{I^{(j)}}(\tilde{P}_k) - \Delta| \leq 2^{-2^{k\beta}}, \lim_{k' \rightarrow \infty} \mathcal{K}(\tilde{P}_{k'}) = (\Delta, 0, 0) \right) \\ &= \Pr \left( \lim_{k \rightarrow \infty} \mathcal{K}(\tilde{P}_k) = (\Delta, 0, 0) \right), \quad j = 1, 2. \end{aligned}$$

Hence, for the purpose of proving (12), it suffices to show that

$$\begin{aligned} & \lim_{k \rightarrow \infty} \Pr \left( I^{(j)}(\tilde{P}_k) \leq 2^{-2^{k\beta}}, \lim_{k' \rightarrow \infty} \mathcal{K}(\tilde{P}_{k'}) = (\Delta, 0, 0) \right) \\ &= \Pr \left( \lim_{k \rightarrow \infty} \mathcal{K}(\tilde{P}_k) = (\Delta, 0, 0) \right), \quad j = 1, 2. \end{aligned}$$

By symmetry, we shall only consider the case  $j = 1$ .

For any  $\alpha \in \mathbb{F}_q \setminus \{0\}$  and  $p_{S,U,V}$  over  $\mathcal{S} \times \mathbb{F}_q^2$ , define

$$Z_\alpha(p_{S,U,V}) = Z(\alpha U \oplus V | S),$$

where

$$\begin{aligned} & Z(\alpha U \oplus V | S) \\ &= \frac{1}{q-1} \sum_{\substack{a, a' \in \mathbb{F}_q \\ a \neq a'}} \sum_s \sqrt{p_{\alpha U \oplus V, S}(a, s) p_{\alpha U \oplus V, S}(a', s)}. \end{aligned}$$

Let  $C_0(\alpha) = I^{(1)}(\tilde{P}_0) + Z(\tilde{P}_0)$ . For  $k \geq 1$ , define

$$C_k(\alpha) = \begin{cases} I^{(1)}(\tilde{P}_{k-1}^-) + Z_\alpha(\tilde{P}_{k-1}^+), & S_k = - \\ I^{(1)}(\tilde{P}_{k-1}^+) + Z_\alpha(\tilde{P}_{k-1}^-), & S_k = +. \end{cases}$$

In view of the fact that  $\lim_{k \rightarrow \infty} I^{(1)}(\tilde{P}_k) \in \{0, \Delta\}$  and  $\lim_{k \rightarrow \infty} Z_\alpha(\tilde{P}_k) \in \{0, 1\}$  almost surely [6], one can readily show that

$$\Pr \left( \lim_{k \rightarrow \infty} C_k(\alpha) \in \{0, 1, \Delta, \Delta + 1\} \right) = 1.$$

Therefore, if there exists some  $K > 0$  such that

$$C_k(\alpha) \leq \begin{cases} K[C_{k-1}(\alpha)]^2, & S_k = - \\ K C_{k-1}(\alpha), & S_k = +, \end{cases} \quad (13)$$

then it follows from Lemma 4 in Appendix A that

$$\begin{aligned} & \lim_{k \rightarrow \infty} \Pr \left( C_k(\alpha) \leq 2^{-2^{k\beta}}, \lim_{k' \rightarrow \infty} C_{k'}(\alpha) = 0 \right) \\ &= \Pr \left( \lim_{k \rightarrow \infty} C_k(\alpha) = 0 \right). \end{aligned} \quad (14)$$

It suffices to verify (13) for  $k = 1$ . Let  $(\tilde{X}_1, \tilde{Y}_1, \tilde{Z}_1)$  and  $(\tilde{X}_2, \tilde{Y}_2, \tilde{Z}_2)$  be two independent copies of  $(\tilde{X}, \tilde{Y}, \tilde{Z})$  (see Section II for the definition of  $(\tilde{X}, \tilde{Y}, \tilde{Z})$ ), and define

$$\begin{aligned}\tilde{U}_1 &= \tilde{Y}_1 \ominus \tilde{Y}_2, & \tilde{U}_2 &= \tilde{Y}_2, \\ \tilde{V}_1 &= \tilde{Z}_1 \ominus \tilde{Z}_2, & \tilde{V}_2 &= \tilde{Z}_2.\end{aligned}$$

Note that

$$\begin{aligned}I^{(1)}(\tilde{P}_0^+) + Z_\alpha(\tilde{P}_0^-) &= I(\tilde{X}_1, \tilde{X}_2, \tilde{U}_1, \tilde{V}_1; \tilde{U}_2) + Z(\alpha\tilde{U}_1 \oplus \tilde{V}_1 | \tilde{X}_1, \tilde{X}_2) \\ &= I(\tilde{X}_1, \tilde{X}_2, \tilde{U}_1; \tilde{U}_2) + I(\tilde{V}_1; \tilde{U}_2 | \tilde{X}_1, \tilde{X}_2, \tilde{U}_1) \\ &\quad + Z(\alpha\tilde{U}_1 \oplus \tilde{V}_1 | \tilde{X}_1, \tilde{X}_2) \\ &\leq I(\tilde{X}_1, \tilde{X}_2; \tilde{U}_1, \tilde{U}_2) + H(\tilde{V}_1 | \tilde{X}_1, \tilde{X}_2, \tilde{U}_1) \\ &\quad + Z(\alpha\tilde{U}_1 \oplus \tilde{V}_1 | \tilde{X}_1, \tilde{X}_2) \\ &= 2I^{(1)}(\tilde{P}_0) + H(\alpha\tilde{U}_1 \oplus \tilde{V}_1 | \tilde{X}_1, \tilde{X}_2, \tilde{U}_1) \\ &\quad + Z(\alpha\tilde{U}_1 \oplus \tilde{V}_1 | \tilde{X}_1, \tilde{X}_2) \\ &\leq 2I^{(1)}(\tilde{P}_0) + H(\alpha\tilde{U}_1 \oplus \tilde{V}_1 | \tilde{X}_1, \tilde{X}_2) \\ &\quad + Z(\alpha\tilde{U}_1 \oplus \tilde{V}_1 | \tilde{X}_1, \tilde{X}_2) \\ &\leq 2I^{(1)}(\tilde{P}_0) + [(q-1)(\log e) + 1]Z(\alpha\tilde{U}_1 \oplus \tilde{V}_1 | \tilde{X}_1, \tilde{X}_2) \\ &\leq 2I^{(1)}(\tilde{P}_0) + [(q-1)(\log e) + 1]qZ_\alpha(\tilde{P}_0),\end{aligned}\tag{15}$$

where (15) is due to the fact that

$$\begin{aligned}H(\alpha\tilde{U}_1 \oplus \tilde{V}_1 | \tilde{X}_1, \tilde{X}_2) &\leq \log(1 + (q-1)Z(\alpha\tilde{U}_1 \oplus \tilde{V}_1 | \tilde{X}_1, \tilde{X}_2)) \\ &\leq (q-1)(\log e)Z(\alpha\tilde{U}_1 \oplus \tilde{V}_1 | \tilde{X}_1, \tilde{X}_2),\end{aligned}$$

and (16) is due to the fact [6, Eq. (14)] that  $Z_\alpha(\tilde{P}_0^-) \leq qZ_\alpha(\tilde{P}_0)$ ; moreover,

$$\begin{aligned}I^{(1)}(\tilde{P}_0^-) + Z_\alpha(\tilde{P}_0^+) &\leq \kappa(q)[I^{(1)}(\tilde{P}_0)]^2 + q[Z_\alpha(\tilde{P}_0)]^2,\end{aligned}\tag{17}$$

where (17) follows from Lemma 5 in Appendix A and the fact [6, Eq. (14)] that  $Z_\alpha(\tilde{P}_0^+) \leq q[Z_\alpha(\tilde{P}_0)]^2$ . Therefore, (13) holds with  $K = \max\{[(q-1)(\log e) + 1]q, \kappa(q)\}$ . Hence, (14) is proved. It is known [6] that

$$\begin{aligned}\Pr\left(\lim_{k \rightarrow \infty} \mathcal{K}(\tilde{P}_k) = (\Delta, 0, 0)\right) &= \Pr\left(\lim_{k \rightarrow \infty} \min_{\alpha \in \mathbb{F}_q \setminus \{0\}} Z_\alpha(\tilde{P}_k) = 0, \lim_{k \rightarrow \infty} \mathcal{K}(\tilde{P}_k) = (\Delta, 0, 0)\right),\end{aligned}$$

which implies

$$\begin{aligned}\Pr\left(\lim_{k \rightarrow \infty} \mathcal{K}(\tilde{P}_k) = (\Delta, 0, 0)\right) &= \Pr\left(\lim_{k \rightarrow \infty} \min_{\alpha \in \mathbb{F}_q \setminus \{0\}} C_k(\alpha) = 0, \lim_{k \rightarrow \infty} \mathcal{K}(\tilde{P}_k) = (\Delta, 0, 0)\right).\end{aligned}\tag{18}$$

Given any  $\epsilon > 0$  and  $\delta \in (0, 1)$ , there exists a  $k_0$  such that

$$\begin{aligned}\Pr\left(C_k(\alpha) \leq \delta, \lim_{k' \rightarrow \infty} C_{k'}(\alpha) = 0\right) &\geq \Pr\left(C_k(\alpha) \leq \delta\right) - \epsilon\end{aligned}\tag{19}$$

for all  $k \geq k_0$  and  $\alpha \in \mathbb{F}_q \setminus \{0\}$ . In view of (18), given any  $\epsilon > 0$  and  $\delta > 0$ , there exists a  $k'_0$  such that

$$\begin{aligned}\Pr\left(\min_{\alpha \in \mathbb{F}_q \setminus \{0\}} C_k(\alpha) \leq \delta, \lim_{k' \rightarrow \infty} \mathcal{K}(\tilde{P}_{k'}) = (\Delta, 0, 0)\right) &\geq \Pr\left(\lim_{k' \rightarrow \infty} \mathcal{K}(\tilde{P}_{k'}) = (\Delta, 0, 0)\right) - \epsilon\end{aligned}\tag{20}$$

for all  $k \geq k'_0$ . Let  $k^* = \max\{k_0, k'_0\}$ . We have

$$\begin{aligned}\Pr\left(\lim_{k \rightarrow \infty} \mathcal{K}(\tilde{P}_k) = (\Delta, 0, 0)\right) &\leq \Pr\left(C_{k^*}(\alpha) \leq \delta, \lim_{k \rightarrow \infty} \mathcal{K}(\tilde{P}_k) = (\Delta, 0, 0)\right) + \epsilon \\ &\leq \Pr\left(C_{k^*}(\alpha) \leq \delta, \lim_{k \rightarrow \infty} C_k(\alpha) = 0, \lim_{k \rightarrow \infty} \mathcal{K}(\tilde{P}_k) = (\Delta, 0, 0)\right) + 2\epsilon \\ &\leq \Pr\left(\lim_{k \rightarrow \infty} C_k(\alpha) = 0, \lim_{k \rightarrow \infty} \mathcal{K}(\tilde{P}_k) = (\Delta, 0, 0)\right) + 2\epsilon\end{aligned}\tag{21}$$

for some  $\alpha \in \mathbb{F}_q \setminus \{0\}$ , where (21) and (22) are due to (20) and (19), respectively. Therefore,

$$\begin{aligned}\Pr\left(\lim_{k \rightarrow \infty} \mathcal{K}(\tilde{P}_k) = (\Delta, 0, 0)\right) &= \Pr\left(\lim_{k \rightarrow \infty} C_k(\alpha) = 0, \lim_{k \rightarrow \infty} \mathcal{K}(\tilde{P}_k) = (\Delta, 0, 0)\right)\end{aligned}$$

for some  $\alpha \in \mathbb{F}_q \setminus \{0\}$ , which, together with (14), implies

$$\begin{aligned}\Pr\left(\lim_{k \rightarrow \infty} \mathcal{K}(\tilde{P}_k) = (\Delta, 0, 0)\right) &= \lim_{k \rightarrow \infty} \Pr\left(\min_{\alpha \in \mathbb{F}_q \setminus \{0\}} C_k(\alpha) \leq 2^{-2k^\beta}, \lim_{k' \rightarrow \infty} \mathcal{K}(\tilde{P}_{k'}) = (\Delta, 0, 0)\right).\end{aligned}$$

The proof is complete in view of the fact that  $C_k(\alpha) \geq I^{(1)}(\tilde{P}_k)$  for all  $\alpha \in \mathbb{F}_q \setminus \{0\}$ . ■

With the aid of Theorem 3, one can readily show that the proposed scheme can asymptotically achieve a certain rate pair on the dominant line of the EGC rate region. The detailed performance analysis is relegated to Appendix B.

*Remark:* In fact, when  $\eta(i) = 1$ , one can assign  $i$  into one of  $\mathcal{F}_1$  and  $\mathcal{F}_2$  in an arbitrary manner, and modify the encoding operation accordingly. Specifically, the encoding operation remains the same if  $\eta(i) = 0$ ; for the case  $\eta(i) = 1$ , there are two possible scenarios.

1) If  $i \in \mathcal{F}_1$ , then  $v_i$  takes value  $a \in \mathbb{F}_q$  with probability

$$\frac{p_{X^n, U^i, V^i}(x^n, u^i, (v^{i-1}, a))}{p_{X^n, U^i, V^{i-1}}(x^n, u^i, v^{i-1})}.$$

2) If  $i \in \mathcal{F}_2$ , then  $u_i$  takes value  $a \in \mathbb{F}_q$  with probability

$$\frac{p_{X^n, U^i, V^i}(x^n, (u^{i-1}, a), v^i)}{p_{X^n, U^{i-1}, V^i}(x^n, u^{i-1}, v^i)}.$$

The performance analysis in Appendix B can be applied to this generalized version of the polar coding scheme with no essential change. Due to the aforementioned freedom in the case  $\eta(i) = 1$ , this general scheme can asymptotically achieve a segment of the dominant line of the EGC rate region if  $\Pr(\lim_{k \rightarrow \infty} (I^{(0)}(P_k), I^{(1)}(P_k), I^{(2)}(P_k)) = (\Delta, 0, 0)) > 0$ .

#### IV. THE DECORRELATION EFFECT OF THE POLAR TRANSFORM

As shown in Section II, the MDC problem considered in this work can be viewed as the source coding dual of the MAC coding problem considered in [6]. Underlying this duality is a dithering-based argument that removes the dependency between the two (single-letter) descriptions without affecting the key mutual information quantities. This dithering-based argument is motivated by our analysis of the polarization behavior of symbol-wise mutual information between the two descriptions. It will be seen that, under mild conditions, the polar transform in (10) has the effect of removing the symbol-wise dependency between the two descriptions, which is somewhat surprising in view of the fact that the block-wise mutual information is invariant under the transformation. For the purpose of proving this result, we establish a necessary and sufficient condition for irreducibility and aperiodicity of a class of Markov chains defined over  $\mathbb{F}_{q_1} \times \mathbb{F}_{q_2}$ .

For any MDC distribution  $Q(s, u, v)$ , we define

$$\begin{aligned}\tilde{I}(Q) &= I(S; U, V), \\ \hat{I}(Q) &= I(U; V).\end{aligned}$$

Note that

$$I^{(0)}(P_k) = \tilde{I}(P_k) + \hat{I}(P_k), \quad k \geq 0,$$

where  $\hat{I}(P_k)$  can be viewed as a measure of the symbol-wise dependency between the two descriptions after the transformation. This section is devoted to investigating the polarization behavior of  $\{\hat{I}(P_k) : k \geq 0\}$ . The polarization behavior of  $\{\tilde{I}(P_k) : k \geq 0\}$  can be deduced from that of  $\{I^{(0)}(P_k) : k \geq 0\}$  (see Theorem 2) and  $\{\hat{I}(P_k) : k \geq 0\}$ .

Let  $Y$  and  $Z$  be two single-letter descriptions of source  $X$ . We assume that  $Y$  and  $Z$  are uniformly distributed over  $\mathbb{F}_{q_1}$  and  $\mathbb{F}_{q_2}$ , respectively, and denote their joint distribution by  $p_{Y,Z}$ . Define  $U^-, U^+, V^-,$  and  $V^+$  such that

$$\begin{aligned}Y &= U^- \oplus U^+, & \tilde{Y} &= U^+, \\ Z &= V^- \oplus V^+, & \tilde{Z} &= V^+, \end{aligned}$$

or equivalently,

$$\begin{aligned}U^- &= Y \ominus \tilde{Y}, & U^+ &= \tilde{Y}, \\ V^- &= Z \ominus \tilde{Z}, & V^+ &= \tilde{Z}, \end{aligned}$$

where  $(\tilde{Y}, \tilde{Z})$  is an independent copy of  $(Y, Z)$ . We further define

$$\begin{aligned}U^{*-} &= U^* \ominus \tilde{U}^*, & U^{*+} &= \tilde{U}^*, \\ V^{*-} &= V^* \ominus \tilde{V}^*, & V^{*+} &= \tilde{V}^*, \end{aligned}$$

where  $* \in \{-, +\}$  and  $(\tilde{U}^*, \tilde{V}^*)$  is an independent copy of  $(U^*, V^*)$ . Recursively, we can define  $U^{s^k}$  and  $V^{s^k}$  for every  $s^k \in \{-, +\}^k$ ,  $k \geq 0$ , where  $U^{s^k} = Y$  and  $V^{s^k} = Z$  when  $k = 0$ . For every  $s^k \in \{-, +\}^k$ , let  $\tau(s^k)$  denote the number of minus signs ( $-$ ) in  $s^k$ . Note that if  $\tau(s^k) = 0$ , then  $(U^{s^k}, V^{s^k})$  is distributed according to  $p_{Y,Z}$ ; if  $\tau(s^k) > 0$ , then  $(U^{s^k}, V^{s^k})$  is the sum of  $2^{\tau(s^k)-1}$  independent copies of  $(Y, Z)$  and  $2^{\tau(s^k)-1}$  independent copies of  $(q_1 \ominus Y, q_2 \ominus Z)$ , where the sum is modulo- $q_j$  for the  $j$ th entry,  $j = 1, 2$ . Define

$\tilde{h}(s^k) = H(U^{s^k}, V^{s^k})$  for  $s^k \in \{-, +\}^k$ . Let  $S_1, S_2, \dots$ , be an i.i.d. sequence of random variables taking values in the set  $\{-, +\}$  with  $\Pr(S_i = -) = \Pr(S_i = +) = \frac{1}{2}$ . Clearly,

$$\hat{I}(P_k) = \Delta_1 + \Delta_2 - \tilde{h}(S^k), \quad k \geq 0.$$

Let  $T$  be a distribution over  $\mathbb{F}_{q_1} \times \mathbb{F}_{q_2}$ . Define  $\text{supp}(T) = \{(a, b) \in \mathbb{F}_{q_1} \times \mathbb{F}_{q_2} : T(a, b) > 0\}$ . Let  $\{W_\ell : \ell \geq 0\}$  be a Markov chain induced by  $T$  such that  $\Pr(W_{\ell+1} = (a, b) | W_\ell = (a', b')) = T(a \ominus a', b \ominus b')$  for all  $(a, b) \in \mathbb{F}_{q_1} \times \mathbb{F}_{q_2}$  and  $(a', b') \in \mathbb{F}_{q_1} \times \mathbb{F}_{q_2}$ ,  $\ell \geq 0$ . It will be seen that the polarization behavior of  $\{\hat{I}(P_k) : k \geq 0\}$  depends critically on the properties of this Markov chain with  $T$  chosen to be  $p_{Y,Z}$ .

##### A. $q_1 \neq q_2$

*Lemma 1:* When  $q_1$  and  $q_2$  are non-identical primes, the Markov chain  $\{W_\ell : \ell \geq 0\}$  is irreducible and aperiodic if and only if there exist  $(a_i, b_i) \in \text{supp}(T)$ ,  $i = 1, 2, 3, 4$ , such that  $a_1 \neq a_2$  and  $b_3 \neq b_4$  (some of these four vectors can be identical).

*Proof:* If the condition specified in Lemma 1 is not satisfied, then one of the following cases must be true.

- 1) There exists an  $a^* \in \mathbb{F}_{q_1}$  such that  $a = a^*$  for all  $(a, b) \in \text{supp}(T)$ .
- 2) There exists a  $b^* \in \mathbb{F}_{q_2}$  such that  $b = b^*$  for all  $(a, b) \in \text{supp}(T)$ .

It suffices to consider case 1). Clearly, the Markov chain  $\{W_\ell : \ell \geq 0\}$  is reducible if  $a^* = 0$ . Now assume  $a^* \neq 0$ . For any  $(\tilde{a}_i, \tilde{b}_i) \in \text{supp}(T)$ ,  $i = 1, 2, \dots, L$ , such that  $(\tilde{a}_1 \oplus \dots \oplus \tilde{a}_L, \tilde{b}_1 \oplus \dots \oplus \tilde{b}_L) = (0, 0)$ , one can easily see that  $L$  must be a multiple of  $q_1$ . Therefore, the Markov chain  $\{W_\ell : \ell \geq 0\}$  cannot be aperiodic.

If the condition specified in Lemma 1 is satisfied (i.e., there exist  $(a_i, b_i) \in \text{supp}(T)$ ,  $i = 1, 2, 3, 4$ , such that  $a_1 \neq a_2$  and  $b_3 \neq b_4$ ), then one of the following cases must be true.

- i) There exists  $(a, b) \in \text{supp}(T)$  such that  $a \neq 0$  and  $b \neq 0$ .
- ii) There exist  $(a, 0) \in \text{supp}(T)$  and  $(0, b) \in \text{supp}(T)$  such that  $a \neq 0$  and  $b \neq 0$ .

The Markov chain  $\{W_\ell : \ell \geq 0\}$  is irreducible in light of the fact that in case i)

$$\begin{aligned}\{(ca \bmod q_1, cb \bmod q_2) : c = 0, 1, \dots, q_1q_2 - 1\} \\ = \mathbb{F}_{q_1} \times \mathbb{F}_{q_2},\end{aligned}$$

and in case ii)

$$\{(ca \bmod q_1, db \bmod q_2) : c \in \mathbb{F}_{q_1}, d \in \mathbb{F}_{q_2}\} = \mathbb{F}_{q_1} \times \mathbb{F}_{q_2}.$$

Since  $a_1 \neq a_2$  and  $b_3 \neq b_4$ , it follows that  $ca_1 \oplus \tilde{c}a_2 = 0$  and  $db_3 \oplus \tilde{d}b_4 = 0$  for some  $c \in \mathbb{F}_{q_1}$ ,  $\tilde{c} \in \mathbb{F}_{q_1}$ ,  $d \in \mathbb{F}_{q_2}$ , and  $\tilde{d} \in \mathbb{F}_{q_2}$  such that  $c + \tilde{c} \leq q_1 - 1$  and  $d + \tilde{d} \leq q_2 - 1$ . Note that

$$\begin{aligned}(q_2(ca_1 + \tilde{c}a_2) \bmod q_1, q_2(cb_1 + \tilde{c}b_2) \bmod q_2) &= (0, 0), \\ (q_1(da_3 + \tilde{d}a_4) \bmod q_1, q_1(db_3 + \tilde{d}b_4) \bmod q_2) &= (0, 0), \\ (q_1q_2a_1 \bmod q_1, q_1q_2b_1 \bmod q_2) &= (0, 0).\end{aligned}$$

The Markov chain  $\{W_\ell : \ell \geq 0\}$  is aperiodic in light of the fact that the greatest common divisor of  $q_2(c + \tilde{c})$ ,  $q_1(d + \tilde{d})$ , and  $q_1q_2$  is 1. This completes the proof of Lemma 1. ■

*Theorem 4:* When  $q_1$  and  $q_2$  are non-identical primes, the process  $\{\hat{I}(P_k), k \geq 0\}$  converges to 0 almost surely; furthermore,

$$\lim_{k \rightarrow \infty} \Pr(\hat{I}(P_k) \leq 2^{-2k\beta}) = 1$$

for any  $\beta \in (0, \frac{1}{2})$ .

*Proof:* Let  $\{W_\ell : \ell \geq 0\}$  be the Markov chain induced by  $p_{Y,Z}$ . One can readily show by invoking Lemma 1 that  $\{W_\ell : \ell \geq 0\}$  is irreducible and aperiodic. Moreover, it is easy to see that the equilibrium distribution of  $\{W_\ell : \ell \geq 0\}$  is  $\text{unif}(\mathbb{F}_{q_1} \times \mathbb{F}_{q_2})$ , where  $\text{unif}(\mathbb{F}_{q_1} \times \mathbb{F}_{q_2})$  denotes the uniform distribution over  $\mathbb{F}_{q_1} \times \mathbb{F}_{q_2}$ . Therefore, in light of the well-known fact (see [10, Th. 1]) that the state distribution of a finite-state irreducible and aperiodic Markov chain converges to the equilibrium distribution with a geometric rate, there exist constants  $C > 0$  and  $\gamma \in (0, 1)$  such that

$$\|p_{W_\ell} - \text{unif}(\mathbb{F}_{q_1} \times \mathbb{F}_{q_2})\| \leq C\gamma^\ell, \quad \ell \geq 0,$$

for any initial state distribution  $p_{W_0}$ . Note that if  $\tau(s^k) = 0$ , then  $p_{U^{s^k}, V^{s^k}} = p_{W_0}$  (with  $W_0 = (0, 0)$ ); if  $\tau(s^k) > 0$ , then  $p_{U^{s^k}, V^{s^k}} = p_{W_{2^{\tau(s^k)-1}}}$  with  $W_0$  being the sum of  $2^{\tau(s^k)-1}$  independent copies of  $(q_1 \ominus Y, q_2 \ominus Z)$  (the sum is modulo- $q_j$  for the  $j$ th entry,  $j = 1, 2$ ). As a consequence,

$$\|p_{U^{s^k}, V^{s^k}} - \text{unif}(\mathbb{F}_{q_1} \times \mathbb{F}_{q_2})\| \leq C\gamma^{2^{\tau(s^k)-1}},$$

which, together with [11, Lemma 2], implies

$$I(U^{s^k}; V^{s^k}) \leq (q_1q_2 \log e)C\gamma^{2^{\tau(s^k)-1}} \quad (23)$$

for every  $s^k \in \{-, +\}^k$ . Therefore, we have

$$\begin{aligned} \Pr\left(\lim_{k \rightarrow \infty} \hat{I}(P_k) = 0\right) \\ \geq 1 - \sum_{k=1}^{\infty} \Pr(S_{k'} = + \text{ for all } k' \geq k) \\ = 1, \end{aligned}$$

which implies that  $\{\hat{I}(P_k), k \geq 0\}$  converges to 0 almost surely. By the weak law of large numbers,

$$\lim_{k \rightarrow \infty} \Pr(\tau(S^k) \geq \beta k) = 1, \quad \beta \in (0, \frac{1}{2}). \quad (24)$$

It can be shown by combining (23) and (24) that

$$\lim_{k \rightarrow \infty} \Pr(\hat{I}(P_k) \leq 2^{-2k\beta}) = 1$$

for any  $\beta \in (0, \frac{1}{2})$ . This completes the proof of Theorem 4. ■

### B. $q_1 = q_2$

*Lemma 2:* When  $q_1 = q_2 = q$  for some prime number  $q$ , the Markov chain  $\{W_\ell : \ell \geq 0\}$  is irreducible and aperiodic if and only if  $\text{supp}(T)$  contains (at least) three different vectors  $(a_i, b_i)$ ,  $i = 1, 2, 3$ , with the following properties:

- 1)  $(a_1, b_1)$  and  $(a_2, b_2)$  are linearly independent;
- 2)  $(a_3, b_3) = c(a_1, b_1) \oplus d(a_2, b_2)$  with  $c \oplus d \neq 1$ .

*Proof:* It is clear that the irreducibility of  $\{W_\ell : \ell \geq 0\}$  is equivalent to the fact that  $\text{supp}(T)$  contains a basis of  $\mathbb{F}_q^2$ , i.e., two linearly independent vectors  $(a_1, b_1)$  and  $(a_2, b_2)$ .

Note that the aperiodicity of  $\{W_\ell : \ell \geq 0\}$  is equivalent to the existence of  $(\tilde{a}_i, \tilde{b}_i) \in \text{supp}(T)$  and  $e_i \in \mathbb{F}_q$ ,  $i = 1, 2, \dots, L$ , for some  $L$  such that  $e_1(\tilde{a}_1, \tilde{b}_1) \oplus \dots \oplus e_L(\tilde{a}_L, \tilde{b}_L) = (0, 0)$  and  $e_1 \oplus \dots \oplus e_L \neq 0$ . Now assume that  $(a_3, b_3) = c(a_1, b_1) \oplus d(a_2, b_2)$  with  $c \oplus d \neq 1$  for some  $(a_3, b_3) \in \text{supp}(T)$ . It is easy to see that  $(a_3, b_3) \oplus (q \ominus c)(a_1, b_1) \oplus (q \ominus d)(a_2, b_2) = (0, 0)$  and  $1 \oplus (q \ominus c) \oplus (q \ominus d) \neq 0$ . Therefore,  $\{W_\ell : \ell \geq 0\}$  is aperiodic. Conversely, if there exist  $(\tilde{a}_i, \tilde{b}_i) \in \text{supp}(T)$  and  $e_i \in \mathbb{F}_q$ ,  $i = 1, 2, \dots, L$ , for some  $L$  such that  $e_1(\tilde{a}_1, \tilde{b}_1) \oplus \dots \oplus e_L(\tilde{a}_L, \tilde{b}_L) = (0, 0)$  and  $e_1 \oplus \dots \oplus e_L \neq 0$ , then, by writing  $(\tilde{a}_i, \tilde{b}_i) = c_i(a_1, b_1) \oplus d_i(a_2, b_2)$ ,  $i = 1, 2, \dots, L$ , we obtain  $(e_1c_1 \oplus \dots \oplus e_Lc_L)(a_1, b_1) \oplus (e_1d_1 \oplus \dots \oplus e_Ld_L)(a_2, b_2) = (0, 0)$ . The linear independence of  $(a_1, b_1)$  and  $(a_2, b_2)$  implies that  $e_1c_1 \oplus \dots \oplus e_Lc_L = 0$  and  $e_1d_1 \oplus \dots \oplus e_Ld_L = 0$ , which further implies  $e_1(c_1 \oplus d_1) \oplus \dots \oplus e_L(c_L \oplus d_L) = 0$ . We must have  $c_i \oplus d_i \neq 1$  for some  $i$  since otherwise  $e_1 \oplus \dots \oplus e_L = 0$ , which is contradictory with the assumption that  $e_1 \oplus \dots \oplus e_L \neq 0$ . This completes the proof of Lemma 2. ■

It is clear that  $q \leq |\text{supp}(p_{Y,Z})| \leq q^2$  when  $q_1 = q_2 = q$  for some prime number  $q$ ; furthermore, there are three possible cases for  $\text{supp}(p_{Y,Z})$ .

- 1) All vectors in  $\text{supp}(p_{Y,Z})$  are linearly dependent.
- 2)  $\text{supp}(p_{Y,Z})$  contains two linearly independent vectors  $(y, z)$  and  $(y', z')$ ; furthermore, every vector in  $\text{supp}(p_{Y,Z})$  can be written as  $c(y, z) \oplus d(y', z')$  for some  $c$  and  $d$  with  $c \oplus d = 1$ .
- 3)  $\text{supp}(p_{Y,Z})$  contains two linearly independent vectors  $(y, z)$  and  $(y', z')$ ; furthermore, there exists a vector in  $\text{supp}(p_{Y,Z})$  which can be written as  $c(y, z) \oplus d(y', z')$  for some  $c$  and  $d$  with  $c \oplus d \neq 1$ .

*Theorem 5:* For cases 1) and 2),

$$\hat{I}(P_k) = \Delta, \quad k \geq 0.$$

For case 3), the process  $\{\hat{I}(P_k) : k \geq 0\}$  converges to 0 almost surely; furthermore,

$$\lim_{k \rightarrow \infty} \Pr(\hat{I}(P_k) \leq 2^{-2k\beta}) = 1$$

for any  $\beta \in (0, \frac{1}{2})$ .

*Proof:* It is easy to see that in case 1)

$$\text{supp}(p_{Y,Z}) = \text{span}\{(y, z)\}$$

for some  $(y, z) \neq (0, 0)$ ; as a consequence,

$$\text{supp}(p_{U^{s^k}, V^{s^k}}) = \text{span}\{(y, z)\}$$

for every  $s^k \in \{-, +\}^k$ . It is also easy to see that in case 2)

$$\text{supp}(p_{Y,Z}) = \{c(y, z) \oplus d(y', z') : c \oplus d = 1\}$$



for some linearly independent  $(y, z)$  and  $(y', z')$ ; therefore,

$$\begin{aligned} & \text{supp}(p_{U^{s^k}, V^{s^k}}) \\ &= \begin{cases} \{c(y, z) \oplus d(y', z') : c \oplus d = 1\}, & \tau(s^k) = 0 \\ \{c(y, z) \oplus d(y', z') : c \oplus d = 0\}, & \tau(s^k) > 0 \end{cases} \end{aligned}$$

for every  $s^k \in \{-, +\}^k$ . Hence, we have  $|\text{supp}(p_{U^{s^k}, V^{s^k}})| = q$  in both case 1) and case 2), which implies that  $U^{s^k}$  and  $V^{s^k}$  are related by a bijection, and consequently

$$I(U^{s^k}; V^{s^k}) = \Delta$$

for every  $s^k \in \{-, +\}^k$ .

The proof for case 3) is essentially the same as that for Theorem 4 and is thus omitted.  $\blacksquare$

*Remark:* If  $Y$  and  $Z$  are not related by a bijection, then  $\text{supp}(p_{Y,Z})$  must contain three different pairs  $(y, z)$ ,  $(y, z')$ , and  $(\tilde{y}, \tilde{z})$  with  $y \neq 0$  and  $y \neq \tilde{y}$ . It is clear that  $(y, z)$  and  $(y, z')$  are linearly independent. Now write  $(\tilde{y}, \tilde{z}) = c(y, z) \oplus d(y, z')$ . Since  $\tilde{y} = cy \oplus dy$  and  $y \neq \tilde{y}$ , it follows that  $c \oplus d \neq 1$ . As a consequence,  $\text{supp}(p_{Y,Z})$  must belong to case 3) if  $Y$  and  $Z$  are not related by a bijection. However, the converse is not true. Indeed, it is easy to find examples in which  $\text{supp}(p_{Y,Z})$  belongs to case 3) even though  $Y$  and  $Z$  are related by a bijection. Note that if  $Y = \psi(Z)$  for some bijective function  $\psi$ , then one can simply generate one description of rate  $I(X; Y)$  and transform it to the second description using  $\psi$ ; the total rate of this simple scheme is  $2I(X; Y)$ , which is less than or equal to the EGC sum rate  $I(X; Y, Z) + I(Y; Z)$  (the inequality is strict unless  $Y$  and  $Z$  are determined by  $X$ ).

## V. CONCLUSION

A joint polarization scheme for the MDC problem is proposed in this paper and is shown to be able to asymptotically achieve a certain rate pair on the dominant line of the EGC rate region. However, this is by no means the only possible polar coding scheme for the MDC problem. It is expected that polar codes can be incorporated to the rate splitting method [12] to achieve the entire EGC rate region; one can also obtain a similar scheme by adapting the method developed in [13] to the MDC setting.

### APPENDIX A

#### PERFORMANCE ANALYSIS: $q_1 \neq q_2$

Recall that the probability distribution  $p(x^n, u^n, v^n)$  is induced by the product distribution  $\prod_{i=1}^n P(x_i, y_i, z_i)$  and the polar transform in (10). Clearly if our encoding procedure replicates this probability distribution, then at the decoder, we can recover  $y^n$  and  $z^n$  which are distributed jointly with  $x^n$  according to  $\prod_{i=1}^n P(x_i, y_i, z_i)$ , and thus meet the distortion constraints  $(D_0, D_1, D_2)$  as given in Theorem 1; moreover, Theorem 2 and (9) ensure that a certain rate pair on the dominant line of the EGC rate region is achieved asymptotically. However, the encoding procedure does not completely replicate  $p(x^n, u^n, v^n)$ , but only closely approximates it (because the almost independent symbols are chosen completely independently instead). As such, our goal is to

show that this approximation does not cause significant performance degradation in terms of the achieved distortions, i.e., the excess distortions can be bounded.

The following lemma (which is a consequence of Pinsker's inequality) is needed for bounding the excess distortions.

*Lemma 3 [5, Lemma 5]:* Let  $W$  denote the transition probability of a discrete channel, and  $I(W)$  denote the mutual information between the input  $X$  and the output  $Y$  when  $X$  is uniformly distributed in the alphabet  $\mathcal{X}$ , then

$$\sum_{x \in \mathcal{X}} \mathbb{E} \left| \frac{1}{|\mathcal{X}|} - p(x|Y) \right| \leq \sqrt{(2 \log^{-1} e) I(W)}.$$

Let  $\hat{p}(x^n, u^n, v^n)$  be a probability distribution defined as follows

$$\hat{p}(x^n, u^n, v^n) = p(x^n) \prod_{i=1}^n \hat{p}(u_i, v_i | x^n, u^{i-1}, v^{i-1}),$$

where

$$\hat{p}(u_i, v_i | x^n, u^{i-1}, v^{i-1}) = \begin{cases} \frac{1}{q_1 q_2}, & i \in \mathcal{F}_1 \text{ and } i \in \mathcal{F}_2 \\ \frac{q_1}{q_2} p(u_i | x^n, u^{i-1}, v^{i-1}), & i \notin \mathcal{F}_1 \text{ and } i \in \mathcal{F}_2 \\ \frac{q_2}{q_1} p(v_i | x^n, u^i, v^{i-1}), & i \in \mathcal{F}_1 \text{ and } i \notin \mathcal{F}_2 \\ p(u_i, v_i | x^n, u^{i-1}, v^{i-1}), & i \notin \mathcal{F}_1 \text{ and } i \notin \mathcal{F}_2. \end{cases}$$

For the coding procedure described in Section III-A, we can write the resultant reconstruction distortion based on the first description as

$$\hat{D}_1 = \mathbb{E}_{\hat{p}}[d_1^{(n)}(X^n, \phi_1^{(n)}(U^n B_n G_n))],$$

where  $d_1^{(n)}$  and  $\phi_1^{(n)}$  are, respectively, the  $n$ -letter extensions of  $d_1$  and  $\phi_1$  in Theorem 1. Similarly, the reconstruction distortion based on the second description and the reconstruction distortion based on both descriptions are given by

$$\begin{aligned} \hat{D}_2 &= \mathbb{E}_{\hat{p}}[d_2^{(n)}(X^n, \phi_2^{(n)}(V^n B_n G_n))], \\ \hat{D}_0 &= \mathbb{E}_{\hat{p}}[d_0^{(n)}(X^n, \phi_0^{(n)}(U^n B_n G_n, V^n B_n G_n))], \end{aligned}$$

where  $d_0^{(n)}$ ,  $d_2^{(n)}$ ,  $\phi_0^{(n)}$ , and  $\phi_2^{(n)}$  are, respectively, the  $n$ -letter extensions of  $d_0$ ,  $d_2$ ,  $\phi_0$ , and  $\phi_2$  in Theorem 1. Note that

$$\begin{aligned} D_1 &\geq \mathbb{E}[d_1(X, \phi_1(Y))] \\ &= \mathbb{E}_p[d_1^{(n)}(X^n, \phi_1^{(n)}(U^n B_n G_n))] \triangleq D_1^*, \\ D_2 &\geq \mathbb{E}[d_2(X, \phi_2(Z))] \\ &= \mathbb{E}_p[d_2^{(n)}(X^n, \phi_2^{(n)}(V^n B_n G_n))] \triangleq D_2^*, \\ D_0 &\geq \mathbb{E}[d_0(X, \phi_0(Y, Z))] \\ &= \mathbb{E}_p[d_0^{(n)}(X^n, \phi_0^{(n)}(U^n B_n G_n, V^n B_n G_n))] \triangleq D_0^*. \end{aligned}$$

Thus we only need to compare  $D_i^*$  and  $\hat{D}_i$  for  $i = 0, 1, 2$ . For this purpose we can write

$$\begin{aligned} |\hat{D}_1 - D_1^*| &= |\mathbb{E}_{\hat{p}}[d_1^{(n)}(X^n, \phi_1^{(n)}(U^n B_n G_n))] \\ &\quad - \mathbb{E}_p[d_1^{(n)}(X^n, \phi_1^{(n)}(U^n B_n G_n))]| \\ &\leq d_{\max} \sum_{x^n, u^n, v^n} |\hat{p}(x^n, u^n, v^n) - p(x^n, u^n, v^n)|. \end{aligned}$$

Similarly, we have

$$|\hat{D}_2 - D_2^*| \leq d_{\max} \sum_{x^n, u^n, v^n} |\hat{p}(x^n, u^n, v^n) - p(x^n, u^n, v^n)|,$$

$$|\hat{D}_0 - D_0^*| \leq d_{\max} \sum_{x^n, u^n, v^n} |\hat{p}(x^n, u^n, v^n) - p(x^n, u^n, v^n)|.$$

We can further write

$$\begin{aligned} & \sum_{x^n, u^n, v^n} |\hat{p}(x^n, u^n, v^n) - p(x^n, u^n, v^n)| \\ &= \sum_{x^n, u^n, v^n} p(x^n) |\hat{p}(u^n, v^n | x^n) - p(u^n, v^n | x^n)| \\ &= \sum_{x^n, u^n, v^n} p(x^n) \left| \prod_{i=1}^n p(u_i, v_i | x^n, u^{i-1}, v^{i-1}) \right. \\ & \quad \left. - \prod_{i=1}^n \hat{p}(u_i, v_i | x^n, u^{i-1}, v^{i-1}) \right| \\ &= \sum_{x^n, u^n, v^n} p(x^n) \left| \sum_{i=1}^n (p(u_i, v_i | x^n, u^{i-1}, v^{i-1}) \right. \\ & \quad \left. - \hat{p}(u_i, v_i | x^n, u^{i-1}, v^{i-1})) \right. \\ & \quad \cdot \left( \prod_{j=1}^{i-1} p(u_j, v_j | x^n, u^{j-1}, v^{j-1}) \right. \\ & \quad \left. \prod_{j=i+1}^n \hat{p}(u_j, v_j | x^n, u^{j-1}, v^{j-1}) \right) \Big| \quad (25) \\ &\leq \sum_{i=1}^n \sum_{x^n, u^n, v^n} p(x^n) \left| (p(u_i, v_i | x^n, u^{i-1}, v^{i-1}) \right. \\ & \quad \left. - \hat{p}(u_i, v_i | x^n, u^{i-1}, v^{i-1})) \right. \\ & \quad \cdot \left( \prod_{j=1}^{i-1} p(u_j, v_j | x^n, u^{j-1}, v^{j-1}) \right. \\ & \quad \left. \prod_{j=i+1}^n \hat{p}(u_j, v_j | x^n, u^{j-1}, v^{j-1}) \right) \Big| \\ &= \sum_{i=1}^n \sum_{x^n, u^i, v^i} p(x^n, u^{i-1}, v^{i-1}) |(p(u_i, v_i | x^n, u^{i-1}, v^{i-1}) \\ & \quad - \hat{p}(u_i, v_i | x^n, u^{i-1}, v^{i-1}))|, \end{aligned}$$

where in (25) the following telescoping expansion [4] is applied

$$\prod_{i=1}^n A_i - \prod_{i=1}^n B_i = \sum_{i=1}^n (A_i - B_i) \prod_{j=1}^{i-1} A_j \prod_{j=i+1}^n B_j.$$

Now one can readily obtain the following upper bounds on the excess distortions

$$|\hat{D}_j - D_j^*| \leq d_{\max} \sum_{i=1}^n E_i, \quad j = 0, 1, 2,$$

where

$$E_i = \sum_{u_i=0}^{q_i-1} \sum_{v_i=0}^{q_2-1} \mathbb{E}_p |p(u_i, v_i | X^n, U^{i-1}, V^{i-1}) \\ - \hat{p}(u_i, v_i | X^n, U^{i-1}, V^{i-1})|.$$

We need to bound  $E_i$  for each of the four cases.

- Case 1:  $i \in \mathcal{F}_1$  and  $i \in \mathcal{F}_2$ . It can be shown that

$$\begin{aligned} E_i &= \sum_{u_i=0}^{q_1-1} \sum_{v_i=0}^{q_2-1} \mathbb{E}_p \left| \frac{1}{q_1 q_2} - p(u_i, v_i | X^n, U^{i-1}, V^{i-1}) \right| \\ &= \sum_{x^n, u^i, v^i} \left| \frac{1}{q_1 q_2} p(x^n, u^{i-1}, v^{i-1}) - p(x^n, u^i, v^i) \right| \\ &= \sum_{x^n, u^i, v^i} \left| \frac{1}{q_1 q_2} p(x^n, u^{i-1}, v^{i-1}) - p(x^n, u^i, v^i) \right. \\ & \quad \left. - \frac{1}{q_2} p(x^n, u^i, v^{i-1}) + \frac{1}{q_2} p(x^n, u^i, v^{i-1}) \right| \\ &\leq \sum_{x^n, u^i, v^i} \left| \frac{1}{q_1 q_2} p(x^n, u^{i-1}, v^{i-1}) \right. \\ & \quad \left. - \frac{1}{q_2} p(x^n, u^i, v^{i-1}) \right| \\ & \quad + \sum_{x^n, u^i, v^i} \left| \frac{1}{q_2} p(x^n, u^i, v^{i-1}) - p(x^n, u^i, v^i) \right| \\ &= \sum_{v_i=0}^{q_2-1} \frac{1}{q_2} \sum_{x^n, u^i, v^{i-1}} \left| \frac{1}{q_1} p(x^n, u^{i-1}, v^{i-1}) \right. \\ & \quad \left. - p(x^n, u^i, v^{i-1}) \right| \\ & \quad + \sum_{x^n, u^i, v^i} \left| \frac{1}{q_2} p(x^n, u^i, v^{i-1}) - p(x^n, u^i, v^i) \right| \\ &\leq \sqrt{(2 \log^{-1} e) I(U_i; X^n, U^{i-1}, V^{i-1})} \\ & \quad + \sqrt{(2 \log^{-1} e) I(V_i; X^n, U^i, V^{i-1})} \quad (26) \\ &= \sqrt{(2 \log^{-1} e) I^{(1)}(P_{(i)})} + \sqrt{(2 \log^{-1} e) \overline{I}^{(1)}(P_{(i)})} \\ &\leq \sqrt{(2 \log^{-1} e) \overline{I}^{(2)}(P_{(i)})} + \sqrt{(2 \log^{-1} e) \overline{I}^{(1)}(P_{(i)})} \\ &\leq 2\sqrt{(2 \log^{-1} e) \epsilon}, \end{aligned}$$

where (26) follows from Lemma 3.

- Case 2:  $i \notin \mathcal{F}_1$  and  $i \in \mathcal{F}_2$ . We have

$$\begin{aligned} E_i &= \sum_{v_i=0}^{q_2-1} \mathbb{E}_p \left| \frac{1}{q_2} - p(v_i | X^n, U^i, V^{i-1}) \right| \\ &\leq \sqrt{(2 \log^{-1} e) \overline{I}^{(1)}(P_{(i)})} \\ &\leq \sqrt{(2 \log^{-1} e) \epsilon}. \end{aligned}$$

- Case 3:  $i \in \mathcal{F}_1$  and  $i \notin \mathcal{F}_2$ . We have

$$\begin{aligned} E_i &= \sum_{u_i=0}^{q_1-1} \mathbb{E}_p \left| \frac{1}{q_1} - p(u_i | X^n, U^{i-1}, V^{i-1}) \right| \\ &\leq \sqrt{(2 \log^{-1} e) I^{(1)}(P_{(i)})} \\ &\leq \sqrt{(2 \log^{-1} e) \overline{I}^{(2)}(P_{(i)})} \\ &\leq \sqrt{(2 \log^{-1} e) \epsilon}. \end{aligned}$$

- Case 4:  $i \notin \mathcal{F}_1$  and  $i \notin \mathcal{F}_2$ . For this case, we have  $E_i = 0$ .

Therefore,

$$|\hat{D}_j - D_j^*| \leq 2d_{\max} n \sqrt{(2 \log^{-1} e) \epsilon}, \quad j = 0, 1, 2.$$

Choosing  $\epsilon = 2^{-n^\beta}$  with  $\beta \in (0, \frac{1}{2})$  yields that

$$|\hat{D}_j - D_j^*| = O(2^{-n^{\beta'}}), \quad j = 0, 1, 2,$$

for any  $\beta' \in (0, \beta)$ .

The following two lemmas provide a way to determine the limiting rates of the two descriptions.

*Lemma 4 [3, Th. 1]:* Let  $S_1, S_2, \dots$ , be an i.i.d. sequence of random variables taking values in the set  $\{-, +\}$  with  $\Pr(S_i = -) = \Pr(S_i = +) = \frac{1}{2}$ . Let  $Z_0, Z_1, \dots$ , be a non-negative valued random process where  $Z_0$  is a constant,  $Z_k$  is a function of  $S_1, \dots, S_k$ , and

$$Z_{k+1} = \begin{cases} \leq K Z_k^2, & S_{k+1} = - \\ \leq K Z_k, & S_{k+1} = + \end{cases}$$

for some finite  $K > 0$ . Suppose also that  $Z_k$  converges almost surely to  $Z_\infty$ . Then for any  $\beta < \frac{1}{2}$ ,

$$\begin{aligned} \lim_{k \rightarrow \infty} \Pr(Z_k \leq 2^{-2^{k\beta}}) \\ &= \Pr(Z_\infty = 0) \\ &= \lim_{k \rightarrow \infty} \Pr(Z_k \leq 2^{-2^{k\beta}}, Z_\infty = 0). \end{aligned}$$

*Lemma 5 [5, Lemma 6]:* For any  $q$ , there is a constant  $\kappa(q)$  such that for any  $q$ -ary input channel  $W$

$$I(W^-) \leq \kappa(q)[I(W)]^2, \quad I(W^+) \leq 2I(W),$$

where

$$W^-((s_1, s_2)|u_1) = \frac{1}{q} \sum_{u_2 \in \mathbb{F}_q} W(s_2|u_2)W(s_1|u_1 \oplus u_2),$$

$$W^+((s_1, s_2, u_1)|u_2) = \frac{1}{q} W(s_2|u_2)W(s_1|u_1 \oplus u_2).$$

We shall show that

$$\overline{I^{(1)}}(P_k^-) \leq \kappa(q_2)[\overline{I^{(1)}}(P_k)]^2, \quad (27)$$

$$\overline{I^{(2)}}(P_k^-) \leq \kappa(q_1)[\overline{I^{(2)}}(P_k)]^2, \quad (28)$$

$$\overline{I^{(j)}}(P_k^+) \leq 2\overline{I^{(j)}}(P_k), \quad j = 1, 2, \quad (29)$$

which, in light of Lemma 4, implies that for any  $\beta \in (0, \frac{1}{2})$

$$\begin{aligned} \lim_{k \rightarrow \infty} \Pr(\overline{I^{(j)}}(P_k) \leq 2^{-2^{k\beta}}) \\ &= \Pr\left(\lim_{k \rightarrow \infty} \overline{I^{(j)}}(P_k) = 0\right) \\ &= \lim_{k \rightarrow \infty} \Pr\left(\overline{I^{(j)}}(P_k) \leq 2^{-2^{k\beta}}, \lim_{k' \rightarrow \infty} \overline{I^{(j)}}(P_{k'}) = 0\right), \\ & \quad j = 1, 2. \quad (30) \end{aligned}$$

It suffices to verify (27)–(29) for  $k = 0$ . By symmetry, we shall only consider the case  $j = 1$ . Indeed,<sup>5</sup>

$$\begin{aligned} \overline{I^{(1)}}(P_0^+) &= I(X_1, X_2, U_1, U_2, V_1; V_2) \\ &= I(X_1, X_2, Y_1, Y_2, V_1; V_2) \\ &\leq 2\overline{I^{(1)}}(P_0) \end{aligned} \quad (31)$$

<sup>5</sup>With a slight abuse of notation, here we define  $U_1 = Y_1 \oplus Y_2$ ,  $U_2 = Y_2$ ,  $V_1 = Z_1 \oplus Z_2$ , and  $V_2 = Z_2$ .

and

$$\begin{aligned} \overline{I^{(1)}}(P_0^-) &= I(X_1, X_2, U_1; V_1) \\ &\leq I(X_1, X_2, U_1, U_2; V_1) \\ &= I(X_1, X_2, Y_1, Y_2; V_1) \\ &\leq \kappa(q_2)[\overline{I^{(1)}}(P_0)]^2, \end{aligned} \quad (32)$$

where (31) and (32) follow from Lemma 5 with  $W$  chosen to be the channel from  $Z$  to  $(X, Y)$ .

Note that the rate of description  $i$  is  $\hat{R}_i \triangleq \frac{n - |\mathcal{F}_i|}{n} \Delta_i$ ,  $i = 1, 2$ . We have

$$\begin{aligned} \lim_{n \rightarrow \infty} \hat{R}_1 &= \lim_{n \rightarrow \infty} \frac{n - |\mathcal{F}_1|}{n} \Delta_1 \\ &= \lim_{k \rightarrow \infty} \left[ 1 - \Pr\left(\overline{I^{(2)}}(P_k) \leq 2^{-2^{k\beta}}\right) \right] \Delta_1 \\ &= \left[ 1 - \Pr\left(\lim_{k \rightarrow \infty} \overline{I^{(2)}}(P_k) = 0\right) \right] \Delta_1 \\ &= \left[ \Pr\left(\lim_{k \rightarrow \infty} I^{(0)}(P_k) = \Delta_1\right) \right. \\ & \quad \left. + \Pr\left(\lim_{k \rightarrow \infty} I^{(0)}(P_k) = \Delta_1 + \Delta_2\right) \right] \Delta_1, \end{aligned} \quad (33)$$

where (33) is due to (30), and (34) follows from (1) as well as Theorem 2. Similarly,

$$\begin{aligned} \lim_{n \rightarrow \infty} \hat{R}_2 &= \left[ \Pr\left(\lim_{k \rightarrow \infty} I^{(0)}(P_k) = \Delta_2\right) \right. \\ & \quad \left. + \Pr\left(\lim_{k \rightarrow \infty} I^{(0)}(P_k) = \Delta_1 + \Delta_2\right) \right] \Delta_2. \end{aligned}$$

Therefore,

$$\begin{aligned} \lim_{n \rightarrow \infty} \hat{R}_1 + \hat{R}_2 &= \Pr\left(\lim_{k \rightarrow \infty} I^{(0)}(P_k) = \Delta_1\right) \Delta_1 \\ & \quad + \Pr\left(\lim_{k \rightarrow \infty} I^{(0)}(P_k) = \Delta_2\right) \Delta_2 \\ & \quad + \Pr\left(\lim_{k \rightarrow \infty} I^{(0)}(P_k) = \Delta_1 + \Delta_2\right) (\Delta_1 + \Delta_2) \\ &= \lim_{k \rightarrow \infty} \mathbb{E}[I^{(0)}(P_k)] \\ &= I^{(0)}(P), \end{aligned} \quad (35)$$

where (35) is due to (9).

Thus, for any  $\delta > 0$  and  $\beta \in (0, \frac{1}{2})$ , when  $n$  is sufficiently large, the proposed scheme achieves rate-distortion tuple  $(R_1^* + \delta, R_2^* + \delta, D_0^* + \epsilon(\beta), D_1^* + \epsilon(\beta), D_2^* + \epsilon(\beta))$ , where  $\epsilon(\beta)$  is of order  $O(2^{-n^\beta})$  and  $(R_1^*, R_2^*)$  is on the dominant line of the EGC rate region.

## APPENDIX B

### PERFORMANCE ANALYSIS: $q_1 = q_2$

Now we proceed to bound the excess distortions  $|\hat{D}_j - D_j^*|$ ,  $j = 0, 1, 2$ . Recall that

$$|\hat{D}_j - D_j^*| \leq d_{\max} \sum_{i=1}^n E_i, \quad j = 0, 1, 2,$$

where

$$E_i = \sum_{u_i=0}^{q-1} \sum_{v_i=0}^{q-1} \mathbb{E}_p |p(u_i, v_i|X^n, U^{i-1}, V^{i-1})$$

$$- \hat{p}(u_i, v_i|X^n, U^{i-1}, V^{i-1})|.$$

For the case  $\eta(i) = 0$ , one can readily show by following the analysis in Appendix A that

$$E_i \leq 2\sqrt{(2\log^{-1} e)\epsilon}.$$

For the case  $\eta(i) = 1$ , we have

$$\begin{aligned} E_i &= \sum_{u_i=0}^{q-1} \mathbb{E}_p \left| \frac{1}{q} - p(u_i|X^n, U^{i-1}, V^{i-1}) \right| \\ &\leq \sqrt{(2\log^{-1} e)I^{(1)}(P_{(i)})} \\ &\leq \sqrt{(2\log^{-1} e)\epsilon}. \end{aligned}$$

Therefore,

$$|\hat{D}_j - D_j^*| \leq 2d_{\max} n \sqrt{(2\log^{-1} e)\epsilon}, \quad j = 0, 1, 2,$$

Choosing  $\epsilon = 2^{-n^\beta}$  with  $\beta \in (0, \frac{1}{2})$  yields that

$$|\hat{D}_j - D_j^*| = O(2^{-n^{\beta'}}), \quad j = 0, 1, 2,$$

for any  $\beta' \in (0, \beta)$ .

Finally, we shall compute the limiting rates. Note that

$$\begin{aligned} \lim_{n \rightarrow \infty} \hat{R}_1 &= \lim_{n \rightarrow \infty} \frac{n - |\mathcal{F}_1|}{n} \Delta \\ &= \lim_{k \rightarrow \infty} \Pr \left( \|\mathcal{K}(P_k) - (\Delta, 0, 0)\| > 2^{-2^{k\beta}}, \right. \\ &\quad \left. \overline{I^{(2)}}(P_k) > 2^{-2^{k\beta}} \right) \Delta, \\ \lim_{n \rightarrow \infty} \hat{R}_2 &= \lim_{n \rightarrow \infty} \frac{n - |\mathcal{F}_2|}{n} \Delta \\ &= \lim_{k \rightarrow \infty} \Pr \left( \|\mathcal{K}(P_k) - (\Delta, 0, 0)\| \leq 2^{-2^{k\beta}} \right) \Delta \\ &\quad + \Pr \left( \|\mathcal{K}(P_k) - (\Delta, 0, 0)\| > 2^{-2^{k\beta}}, \right. \\ &\quad \left. \overline{I^{(1)}}(P_k) > 2^{-2^{k\beta}} \right) \Delta. \end{aligned}$$

It then follows by (30) and Theorem 3 that

$$\begin{aligned} \lim_{n \rightarrow \infty} \hat{R}_1 &= \Pr \left( \lim_{k \rightarrow \infty} \mathcal{K}(P_k) \neq (\Delta, 0, 0), \right. \\ &\quad \left. \lim_{k \rightarrow \infty} \overline{I^{(2)}}(P_k) \neq 0 \right) \Delta, \\ \lim_{n \rightarrow \infty} \hat{R}_2 &= \Pr \left( \lim_{k \rightarrow \infty} \mathcal{K}(P_k) = (\Delta, 0, 0) \right) \Delta \\ &\quad + \Pr \left( \lim_{k \rightarrow \infty} \mathcal{K}(P_k) \neq (\Delta, 0, 0), \right. \\ &\quad \left. \lim_{k \rightarrow \infty} \overline{I^{(1)}}(P_k) \neq 0 \right) \Delta. \end{aligned}$$

In light of Theorem 2,

$$\begin{aligned} &\Pr \left( \lim_{k \rightarrow \infty} \mathcal{K}(P_k) \neq (\Delta, 0, 0), \lim_{k \rightarrow \infty} \overline{I^{(2)}}(P_k) \neq 0 \right) \\ &= \Pr \left( \lim_{k \rightarrow \infty} \mathcal{K}(P_k) \in \{(\Delta, \Delta, 0), (2\Delta, \Delta, \Delta)\} \right), \\ &\Pr \left( \lim_{k \rightarrow \infty} \mathcal{K}(P_k) \neq (\Delta, 0, 0), \lim_{k \rightarrow \infty} \overline{I^{(1)}}(P_k) \neq 0 \right) \\ &= \Pr \left( \lim_{k \rightarrow \infty} \mathcal{K}(P_k) \in \{(\Delta, 0, \Delta), (2\Delta, \Delta, \Delta)\} \right). \end{aligned}$$

Therefore, we have

$$\begin{aligned} \lim_{n \rightarrow \infty} \hat{R}_1 + \hat{R}_2 &= \Pr \left( \lim_{k \rightarrow \infty} I^{(0)}(P_k) = \Delta \right) \Delta \\ &\quad + \Pr \left( \lim_{k \rightarrow \infty} I^{(0)}(P_k) = 2\Delta \right) 2\Delta \\ &= \lim_{k \rightarrow \infty} \mathbb{E}[I^{(0)}(P_k)] \\ &= I^{(0)}(P). \end{aligned}$$

#### ACKNOWLEDGMENT

The authors would like to thank E. Şaşıoğlu for helpful discussions. They would also like to thank the two anonymous reviewers for their constructive comments which significantly improve the presentation of the paper.

#### REFERENCES

- [1] E. Arikan, "Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 7, pp. 3051–3073, Jul. 2009.
- [2] S. B. Korada, "Polar codes for channel and source coding," Ph.D. dissertation, School Comput. Commun. Sci., École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland, 2009.
- [3] E. Arikan and I. E. Telatar, "On the rate of channel polarization," in *Proc. IEEE Symp. Inf. Theory*, Seoul, Korea, Jun./Jul. 2009, pp. 1493–1495.
- [4] S. B. Korada and R. L. Urbanke, "Polar codes are optimal for lossy source coding," *IEEE Trans. Inf. Theory*, vol. 56, no. 4, pp. 1751–1768, Apr. 2010.
- [5] M. Karzand and I. E. Telatar, "Polar codes for  $q$ -ary source coding," in *Proc. IEEE Int. Symp. Inf. Theory*, Austin, TX, USA, Jun. 2010, pp. 909–912.
- [6] E. Şaşıoğlu, E. Telatar, and E. M. Yeh, "Polar codes for the two-user multiple-access channel," *IEEE Trans. Inf. Theory*, vol. 59, no. 10, pp. 6583–6592, Oct. 2013.
- [7] E. Şaşıoğlu, "Polarization and polar codes," *Found. Trends Commun. Inf. Theory*, vol. 8, no. 4, pp. 259–381, 2011.
- [8] A. E. Gamal and T. M. Cover, "Achievable rates for multiple descriptions," *IEEE Trans. Inf. Theory*, vol. IT-28, no. 6, pp. 851–857, Nov. 1982.
- [9] 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. Inf. Theory*, vol. 57, no. 3, pp. 1443–1456, Mar. 2011.
- [10] E. Seneta, "On the limiting set of nonnegative matrix products," *Statist. Probab. Lett.*, vol. 2, no. 3, pp. 159–163, May 1984.
- [11] E. Şaşıoğlu, "An entropy inequality for  $q$ -ary random variables and its application to channel polarization," in *Proc. IEEE Int. Symp. Inf. Theory*, Austin, TX, USA, Jun. 2010, pp. 1360–1363.
- [12] Y. Zhang, S. Dumitrescu, J. Chen, and Z. Sun, "LDGM-based multiple description coding for finite alphabet sources," *IEEE Trans. Commun.*, vol. 60, no. 12, pp. 3671–3682, Dec. 2012.
- [13] E. Arikan, "Polar coding for the Slepian–Wolf problem based on monotone chain rules," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2012, pp. 566–570.

**Qi Shi** received the B.E. and M.S. degrees in Electrical and Computer Engineering from McMaster University, Hamilton, ON, Canada, in 2011 and 2013, respectively. He is currently a software testing engineer at Inspur, Jinan, Shandong, China.

**Lin Song** received the B.Sc. degree in Electronic Information Engineering, the M.S. degree in Communication and Information Systems from the Harbin Institute of Technology, Harbin, Heilongjiang, China, in 2006 and 2008, respectively, and the Ph.D. degree in Electrical and Computer Engineering from McMaster University, Hamilton, ON, Canada, in 2012. She was a postdoctoral research assistant at McMaster University, Hamilton, ON, Canada from September 2012 to August 2013. She is currently with the Institute of Network Coding at the Chinese University of Hong Kong, Hong Kong, as a postdoctoral fellow. Her research interests include network information theory and coding theory.

**Chao Tian** (S'00–M'05–SM'12) received the B.E. degree in Electronic Engineering from Tsinghua University, Beijing, China, in 2000 and the M.S. and Ph. D. degrees in Electrical and Computer Engineering from Cornell University, Ithaca, NY in 2003 and 2005, respectively.

Dr. Tian was a postdoctoral researcher at Ecole Polytechnique Federale de Lausanne (EPFL) from 2005 to 2007, and then was a member of technical staff-research at AT&T Labs-Research in New Jersey. He joined the Department of Electrical Engineering and Computer Science at the University of Tennessee Knoxville in 2014, where he is now an Associate Professor. He was an Associate Editor for *The IEEE SIGNAL PROCESSING LETTERS* from 2012 to 2014. His research interests include data storage systems, multiuser information theory, joint source-channel coding, quantization design and analysis, as well as image/video coding and processing.

Dr. Tian received the Liu Memorial Award at Cornell University in 2004, and the AT&T Key Contributor Award in 2010, 2011 and 2013.

**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 September 2005 to July 2006, and a Postdoctoral Fellow at the IBM Thomas J. Watson Research Center, Yorktown Heights, NY, from July 2006 to August 2007. Since September 2007 he has been with the Department of Electrical and Computer Engineering at McMaster University, Hamilton, ON, Canada, where he is currently an Associate Professor. 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 Researcher Award from the Province of Ontario in 2010, and the IBM Faculty Award in 2010. He is currently serving as an Associate Editor for *Shannon Theory* for the *IEEE TRANSACTIONS ON INFORMATION THEORY*.

**Sorina Dumitrescu** (M'05–SM'13) received the B.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 held an NSERC University Faculty Award during 2007-2012.