Fifty-Fifth Annual Allerton Conference
Allerton House, UIUC, Illinois, USA
October 3-6, 2017

Generalized Partial Orders for Polar Code Bit-Channels

Wei Wu, Bing Fan, and Paul H. Siegel
Electrical and Computer Engineering Department
University of California, San Diego
La Jolla, CA 92093, USA
Email: {wew128, bifan, psiegel}@ucsd.edu

***Abstract*—We study partial orders (POs) for the synthesized bit-channels of polar codes. First, we consider complementary bit-channel pairs whose Bhattacharyya parameters over the binary erasure channel (BEC) exhibit a symmetry property and provide insight into the alignment of polarized sets of bit-channels for the BEC and general binary-input memoryless symmetric (BMS) channels. Next, we give an alternative proof of a recently presented PO for bit-channels and use the underlying idea to extend the bit-channel ordering to some additional cases. In particular, the bit-channel ordering for a given code block-length is used to generate additional bit-channel ordering relationships for larger block-lengths, generalizing previously known POs. Examples are provided to illustrate the new POs.**

I. Introduction

Polar codes, introduced by Arıkan [1], are the first family of codes proved to be capacity-achieving on binary-input memoryless symmetric (BMS) channels with low-complexity encoders and decoders. The code construction starts from a *channel transformation*, where N synthesized bit-channels $W_N^{(i)}$, $i = 0, 1, \ldots, N-1$, are obtained by applying a linear transformation to N independent copies of a BMS channel W. As the block-length N goes to infinity, the synthesized bit-channels become either noiseless or completely noisy. Encoding and decoding (by means of a successive cancellation (SC) decoder, which decodes the bit-channels $W_N^{(i)}$ successively) were shown to have time complexity $O(N \log_2 N)$.

A polar code carries the information on the least noisy bit-channels and freezes the remaining channels to a predetermined value, usually chosen to be zero. However, with the exception of the binary erasure channel (BEC), it is generally difficult to precisely measure the quality of the bit-channel $W_N^{(i)}$ because of the exponentially growing output alphabet size as a function of the bit-channel index. Several methods have been proposed to help select the information-bearing bit-channels: Monte Carlo simulation was discussed in [1], density evolution was used in [2], and a Gaussian approximation for density evolution was proposed in [3]. In [4], Tal and Vardy accurately approximated the error probabilities of bit-channels by using efficient degrading and upgrading quantization schemes.

Another important characteristic of polar codes is that the bit-channel ordering is channel-dependent. Although no general rule is known for completely ordering the bit-channels of a general BMS channel W, some partial orders (POs) that are independent of the underlying channel W have been found for selected bit-channels in [5], [6], and [7]. In [5], an ordering applicable to the bit-channels with different Hamming weights is presented. (The Hamming weight of $W_N^{(i)}$ is defined as the number of ones in the binary expansion of i.) It states that a bit-channel $W_N^{(j)}$ is stochastically degraded with respect to $W_N^{(i)}$ if the positions of 1 in the binary expansion of j are a subset of the positions of 1 in the binary expansion of i. The ordering in [6, Theorem 1] and [7] compares bit-channels with the same Hamming weight. It is based on the observation that a bit-channel $W_N^{(j)}$ is stochastically degraded with respect to $W_N^{(i)}$ if j is obtained by swapping a more significant 1 with a less significant 0 in the binary expansion of i. Both of these orders are partial, in the sense that not all bit-channels $W_N^{(i)}$, $W_N^{(j)}$ are comparable. However, they can still be used to simplify the construction of polar codes [8].

In this paper, we present further results related to bit-channel ordering. We first consider complementary bit-channel pairs of the form $(W_N^{(i)}, W_N^{(N-1-i)})$. By analyzing properties of their Bhattacharyya parameters, we identify a symmetry property on the BEC and provide a condition for the alignment of polarized sets of bit-channels for the BEC and a general BMS channel W. Next, we provide an elementary proof of the main PO in [6] and use the proof idea to extend the PO to a larger subset of bit-channels. In particular, the bit-channel ordering for a given code block-length is used to generate additional bit-channel ordering relationships for larger block-lengths, generalizing previously known POs. We also present several examples for the BEC to illustrate the new POs.

The rest of this paper is organized as follows. In Section II, we present notations and definitions, as well as some basic results relating to key bit-channel parameters. In Section III, we consider bit-channel pairs $(W_N^{(i)}, W_N^{(N-1-i)})$ whose binary expansions represent complementary polarization sequences. In Section IV, we give an elementary proof of the main PO in [6] and use the proof technique to derive some generalized POs. Some specific examples for the BEC are given in

Section V, and further properties of BEC bit-channel ordering are examined. Finally, Section VI discusses some results for general BMS channels relating to bit-channel POs based upon Bhattacharyya parameter and error probability, rather than channel degradation.

II. Preliminaries

Consider a BMS channel given by $W : \mathcal{X} \to \mathcal{Y}$, with input alphabet $\mathcal{X} = \{0, 1\}$, output alphabet $\mathcal{Y}$ and transition probabilities $\{W(y|x) : x \in \mathcal{X}, y \in \mathcal{Y}\}$. Define the *Bhattacharyya parameter* of the channel W as

$$Z(W) = \sum_{y\in\mathcal{Y}} \sqrt{W(y|0)W(y|1)} \tag{1}$$

and the error probability of W with uniform input under maximum-likelihood decoding as

$$P_e(W) = \frac{1}{2} \sum_{y\in\mathcal{Y}} \min\{W(y|0), W(y|1)\}. \tag{2}$$

Note that when $W = \text{BEC}(\epsilon)$, we have $Z(W) = \epsilon$.

The following relations between $Z(W)$ and $P_e(W)$ are given in [9]:

$$1 - \sqrt{1 - Z(W)^2} \leqslant 2P_e(W) \leqslant Z(W). \tag{3}$$

Consider the channel transformation $(W, W) \to (W^0, W^1)$ defined in the following manner. Starting from the BMS channel $W : \{0, 1\} \to \mathcal{Y}$, the channels $W^0 : \{0, 1\} \to \mathcal{Y}^2$ and $W^1 : \{0, 1\} \to \{0, 1\} \times \mathcal{Y}^2$ are defined as

$$W^0(y_1, y_2|x_1) = \sum_{x_2\in\{0,1\}} \frac{1}{2} W(y_1|x_1 \oplus x_2)W(y_2|x_2), \tag{4}$$

$$W^1(y_1, y_2, x_1|x_2) = \frac{1}{2} W(y_1|x_1 \oplus x_2)W(y_2|x_2). \tag{5}$$

When W is a BEC, the channels W^0 and W^1 are also BECs.

This channel transformation yields the following results relating to the Bhattacharyya parameters of the channels W^0, W^1, and W [9]:

$$Z(W)\sqrt{2 - Z(W)^2} \leqslant Z(W^0) \leqslant 2Z(W) - Z(W)^2, \tag{6}$$

$$Z(W^1) = Z(W)^2. \tag{7}$$

The lower bound and upper bound in (6) are achieved when W is a binary symmetric channel (BSC) and a BEC, respectively. In particular, when $W = \text{BEC}(\epsilon)$, we have $W^0 = \text{BEC}(2\epsilon - \epsilon^2)$ and $W^1 = \text{BEC}(\epsilon^2)$. In [10], similar relations involving the error probability are proved:

$$P_e(W^0) = 2P_e(W) - 2P_e(W)^2, \tag{8}$$

$$P_e(W^1) \geqslant 2P_e(W)^2. \tag{9}$$

The channel transformation can be recursively repeated n times to produce $N = 2^n$ bit-channels. For any $0 \leqslant i \leqslant N - 1$, let $b^n = b_1b_2 \cdots b_n$ be the binary expansion of i, where b_1 is the most significant digit. For example, if $n = 3$ and $i = 6$, the corresponding binary expansion of i is $b^3 = b_1b_2b_3 = 110$. We denote the binary complement of b^n by $\bar{b^n} = \bar{b_1}\bar{b_2} \cdots \bar{b_n}$. Here, $b_j \oplus \bar{b_j} = 1$ for any $1 \leqslant j \leqslant n$. In fact, $\bar{b^n}$ is the binary expansion of $N - 1 - i$. The bit-channel $W_N^{(i)}$ can be written as

$$W_N^{(i)} = W^{b^n} \overset{def}{=} (((W^{b_1})^{b_2})^{\cdots})^{b_n}.$$

Usually, the Bhattacharyya parameter is used to measure the quality of a bit-channel. For any $\varepsilon \in (0, \frac{1}{2})$ (ε is usually taken to be very small), the following sets of ε-*good* and ε-*bad* bit-channels are often considered

$$\mathcal{G}_n^\varepsilon(W) \triangleq \{i \in \{0, 1, \cdots, N - 1\} : Z(W_N^{(i)}) \leqslant \varepsilon\}, \tag{10}$$

$$\mathcal{B}_n^\varepsilon(W) \triangleq \{i \in \{0, 1, \cdots, N-1\} : Z(W_N^{(i)}) \geqslant 1-\varepsilon\}. \tag{11}$$

A stronger measure of bit-channel ordering is provided by the channel degradation relation.

Definition 1. *The channel $Q : \mathcal{X} \to \mathcal{Z}$ is stochastically degraded with respect to the channel $W : \mathcal{X} \to \mathcal{Y}$ if there exists a channel $P : \mathcal{Y} \to \mathcal{Z}$ such that*

$$Q(z|x) = \sum_{y\in\mathcal{Y}} W(y|x)P(z|y) \tag{12}$$

for all $z \in \mathcal{Z}$ and $x \in \mathcal{X}$.

We write $Q \preccurlyeq W$ ($W \succcurlyeq Q$) to denote that Q (W) is stochastically degraded (upgraded) with respect to W (Q). In this paper, when we use $\preccurlyeq$ or $\succcurlyeq$ to describe the relation between two channels, we assume they are not equivalent.

Lemma 1. *The channel transformation in (4) and (5) preserves the degradation relation [4, Lemma 5]. Namely, if $Q \preccurlyeq W$, then*

$$Q^0 \preccurlyeq W^0 \quad \textit{and} \quad Q^1 \preccurlyeq W^1. \tag{13}$$

III. Complementary Bit-Channels

In this section, we examine properties of bit-channels obtained by complementary polarization sequences. We begin by considering the basic one-step channel transformation.

Define $f_0(x) = 2x - x^2$, $x \in [0, 1]$, and $f_0^{(n)} = \underbrace{f_0 \circ f_0 \circ \cdots \circ f_0}_{n}$.

Proposition 1. *For any two BMS channels W and V, if $Z(W)+Z(V) \leqslant a$, $a \in [0, 1]$, then $Z(W^1)+Z(V^0) \leqslant f_0(a)$ and $Z(W^0) + Z(V^1) \leqslant f_0(a)$.*

Proof: According to (6) and (7), we have $Z(W^1) = Z(W)^2$ and $Z(V^0) \leqslant 2Z(V) - Z(V)^2$. Since the Bhattacharyya parameter is non-negative, the condition $Z(W) + Z(V) \leqslant a$ implies $0 \leqslant Z(W) \leqslant a - Z(V)$ and $0 \leqslant Z(V) \leqslant a$. Therefore,

$$\begin{aligned} Z(W^1) + Z(V^0) &\leqslant Z(W)^2 + 2Z(V) - Z(V)^2 \\ &\leqslant [a - Z(V)]^2 + 2Z(V) - Z(V)^2 \\ &= a^2 + 2(1 - a)Z(V) \\ &\leqslant a^2 + 2a(1 - a) \\ &= f_0(a). \end{aligned} \tag{14}$$

Interchanging the roles of W and V, we obtain $Z(W^0) + Z(V^1) \leqslant f_0(a)$. ■

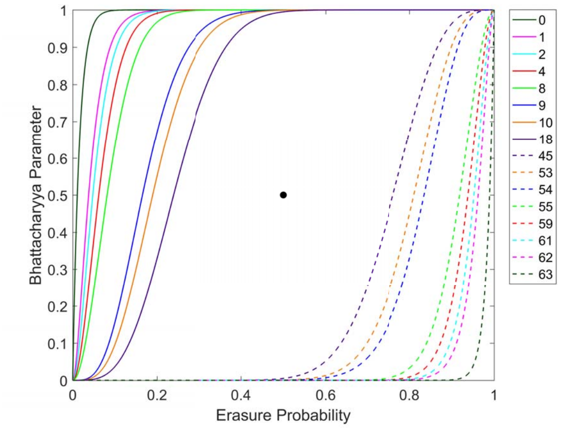

Fig. 1. BEC bit-channels with universal positions when $N = 64$.

Proposition 2. *For any two BMS channels W and V, if $Z(W) + Z(V) \leqslant a$, $a \in [0,1]$, then $Z(W^{b^n}) + Z(V^{\bar{b^n}}) \leqslant f_0^{(n)}(a)$, for all $b^n \in \{0,1\}^n$ and $n \geqslant 1$.*

Proof: Notice that $f_0(a) \in [0,1]$ for any $a \in [0,1]$. The result can be proved by applying Proposition 1 recursively n times. ∎

We have the following three corollaries.

Corollary 1. *If $Z(W)+Z(V) \leqslant 1$, then $Z(W^{b^n})+Z(V^{\bar{b^n}}) \leqslant 1$, for all $b^n \in \{0,1\}^n$ and $n \geqslant 1$.*

Corollary 2. *If $Z(W) \leqslant 1/2$, then $Z(W^{b^n}) + Z(W^{\bar{b^n}}) \leqslant 1$, for all $b^n \in \{0,1\}^n$ and $n \geqslant 1$.*

Remark 1. By Corollary 2, if W satisfies $Z(W) \leqslant 1/2$, then the two bit-channels W^{b^n} and $W^{\bar{b^n}}$ can not both belong to $\mathcal{B}_n^{\varepsilon}(W)$ for any $b^n \in \{0,1\}^n$.

Corollary 3. *Let $W = BEC(\epsilon)$ and $V = W^c \overset{def}{=} BEC(1-\epsilon)$, $\epsilon \in (0,1)$. So $Z(W) + Z(W^c) = 1$. Then $Z(W^{b^n}) + Z((W^c)^{\bar{b^n}}) = 1$, for all $b^n \in \{0,1\}^n$ and $n \geqslant 1$.*

In Fig. 1, we plot the Bhattacharyya parameters for bit-channels obtained by polarizing BECs, with $N = 64$. Each curve is labeled by the corresponding bit-channel index i. The solid and dashed curves with the same color represent complementary bit-channels $(W_N^{(i)}, W_N^{(N-1-i)})$. Their rotational symmetry of order 2 (rotation by an angle of 180°) with respect to the point $(0.5, 0.5)$ follows from Corollary 3. (This symmetry was also noted in a recent paper [11].) The bit-channels shown in the figure have universal positions with respect to bit-channel ordering in the sense that their positions in the complete ordering of the 64 bit-channels are independent of channel erasure probability.

We conclude with some results on the alignment of polarized sets.

Theorem 1. *Let $V = BEC(\epsilon)$, $\epsilon \in (0,1)$. For any BMS channel W, if $Z(W) + Z(V^c) \leqslant 1$, i.e., $Z(W) \leqslant \epsilon$, then the ε-good and ε-bad sets defined in (10) and (11) satisfy*

$$\mathcal{G}_n^{\varepsilon}(W) \supseteq \mathcal{G}_n^{\varepsilon}(V) \quad \text{and} \quad \mathcal{B}_n^{\varepsilon}(W) \subseteq \mathcal{B}_n^{\varepsilon}(V) \tag{15}$$

for all $n \geqslant 1$ and $\varepsilon \in (0, \frac{1}{2})$.

Proof: By Corollary 1, we have $Z(W^{b^n})+Z((V^c)^{\bar{b^n}}) \leqslant 1$, for all $b^n \in \{0,1\}^n$ and $n \geqslant 1$. Since V is a BEC, Corollary 3 can be applied to get $Z(V^{b^n})+Z((V^c)^{\bar{b^n}}) = 1$. Combining the above two relations, we get

$$Z(W^{b^n}) \leqslant Z(V^{b^n}) \tag{16}$$

for all $b^n \in \{0,1\}^n$ and $n \geqslant 1$. Therefore, $Z(V^{b^n}) \leqslant \varepsilon$ implies that $Z(W^{b^n}) \leqslant \varepsilon$, or $\mathcal{G}_n^{\varepsilon}(W) \supseteq \mathcal{G}_n^{\varepsilon}(V)$. Similarly, we find $\mathcal{B}_n^{\varepsilon}(W) \subseteq \mathcal{B}_n^{\varepsilon}(V)$. ∎

Remark 2. Theorem 1 can also be derived by using the results in (6) and (7). Further results about the alignment of polarized sets are given in [12].

IV. Partial Orders

In this section, we give an elementary proof of the main PO in [6] based on mathematical induction. We then use the underlying idea to generalize the PO.

A. New Proof

The following basic lemma follows from the discussion of the case $N = 4$ in [6].

Lemma 2. *For any BMS channel W, when $N = 4$, we have the bit-channel relation $W^{01} \preccurlyeq W^{10}$.*

Remark 3. This implies that $Z(W^{10}) \leqslant Z(W^{01})$ for any BMS channel W, a fact that also follows immediately from (6) and (7), which gives:

$$Z(W^{10}) \leqslant 2Z(W^1) - Z(W^1)^2 = 2Z(W)^2 - Z(W)^4, \tag{17}$$

$$Z(W^{01}) = Z(W^0)^2 \geqslant [Z(W)\sqrt{2 - Z(W)^2}]^2. \tag{18}$$

If W is replaced by W^{p^n}, for all $p^n \in \{0,1\}^n$ and $n \geqslant 1$, Lemma 2 implies

$$W^{p^n 01} \preccurlyeq W^{p^n 10}. \tag{19}$$

Applying Lemma 1 recursively yields

$$W^{p^n 01 q^m} \preccurlyeq W^{p^n 10 q^m} \tag{20}$$

for all $q^m \in \{0,1\}^m$ and $m \geqslant 1$. We restate the main PO in [6] as follows.

Theorem 2. *For any BMS channel W, the degradation relation*

$$W^{p^n 0 r^l 1 q^m} \preccurlyeq W^{p^n 1 r^l 0 q^m} \tag{21}$$

holds for all $p^n \in \{0,1\}^n$, $r^l \in \{0,1\}^l$, $q^m \in \{0,1\}^m$, and $n, l, m \geqslant 0$.

Proof: The proof proceeds by mathematical induction. It is easy to see that when $l = 0$, (21) reduces to (20). Now assume (21) is true for $l = k$, $k \geqslant 0$, i.e.,

$$W^{p^n 0 r^k 1 q^m} \preccurlyeq W^{p^n 1 r^k 0 q^m}$$

holds for all $r^k \in \{0,1\}^k$, $p^n \in \{0,1\}^n$, $q^m \in \{0,1\}^m$, and $n, m \geqslant 0$.

For the induction step, let $l = k+1$, denoting the additional bit by r_{k+1}. There are two cases to consider.

(i) If $r_{k+1} = 0$, the following relations hold:

$$W^{p^n 0 r^{k+1} 1 q^m} = W^{p^n 0 r^k 01 q^m} \overset{(a)}{\preccurlyeq} W^{p^n 0 r^k 10 q^m} \overset{(b)}{\preccurlyeq} W^{p^n 1 r^k 00 q^m} = W^{p^n 1 r^{k+1} 0 q^m}. \quad (22)$$

Here, (a) follows from (20) and (b) is based on the induction hypothesis for $l = k$.

(ii) If $r_{k+1} = 1$, the following relations hold:

$$W^{p^n 0 r^{k+1} 1 q^m} = W^{p^n 0 r^k 11 q^m} \overset{(c)}{\preccurlyeq} W^{p^n 1 r^k 01 q^m} \overset{(d)}{\preccurlyeq} W^{p^n 1 r^k 10 q^m} = W^{p^n 1 r^{k+1} 0 q^m}. \quad (23)$$

Here, (c) is based on the induction hypothesis for $l = k$ and (d) follows from (20).

Combining cases (i) and (ii), we have

$$W^{p^n 0 r^{k+1} 1 q^m} \preccurlyeq W^{p^n 1 r^{k+1} 0 q^m}$$

for all $r^{k+1} \in \{0,1\}^{k+1}$, $p^n \in \{0,1\}^n$, $q^m \in \{0,1\}^m$, and $n, m \geqslant 0$. This completes the induction step. ■

Remark 4. Theorem 2 presents the same PO as in [6, Theorem 1], and (20) leads to the Covering Relation (CR) of the PO proposed in [6].

B. Generalized PO

The PO in [6] applies to bit-channels with the same Hamming weight. With the exception of the PO in [5], little is known about relations among bit-channels with different Hamming weights. We now exploit the idea in the proof of Theorem 2 to derive some additional bit-channel orderings, generalizing the PO in [6].

Theorem 3. *Let a^k be a binary sequence of length k and b^m, c^m be binary sequences of length m, for $k \geqslant 0$ and $m \geqslant 1$. Let W be a BMS channel, and assume $W^{b^m} \preccurlyeq W^{c^m}$. If the condition*

$$W^{b^m a^k 1^n} \preccurlyeq W^{c^m a^k 0^n} \quad (\star)$$

holds for some $n \geqslant 1$, then

$$W^{b^m a^k d^h 1^n} \preccurlyeq W^{c^m a^k d^h 0^n} \quad (24)$$

holds for all $d^h \in \{0,1\}^h$ and $h \geqslant 1$.

Proof: We first let $h = 1$. There are two cases to consider.

(i) If $d_1 = 0$, we have

$$W^{b^m a^k 01^n} \preccurlyeq W^{b^m a^k 1^n 0} \preccurlyeq W^{c^m a^k 0^n 0} = W^{c^m a^k 00^n}. \quad (25)$$

The first degradation relation follows from the PO of Theorem 2, and the second degradation relation is based on applying Lemma 1 to the condition $(\star)$.

(ii) If $d_1 = 1$, we have

$$W^{b^m a^k 11^n} = W^{b^m a^k 1^n 1} \preccurlyeq W^{c^m a^k 0^n 1} \preccurlyeq W^{c^m a^k 10^n}. \quad (26)$$

The first degradation relation is based on applying Lemma 1 to the condition $(\star)$, and the second degradation relation follows from the PO of Theorem 2.

Combining cases (i) and (ii) gives

$$W^{b^m a^k d_1 1^n} \preccurlyeq W^{c^m a^k d_1 0^n}. \quad (27)$$

Repeating this argument for successive inserted bits $d_2, \dots, d_h$ completes the proof. ■

Remark 5. If $W^{b^m} \succcurlyeq W^{c^m}$, then we have

$$W^{b^m a^k 1^n} \succcurlyeq W^{c^m a^k 0^n}$$

holds for all $a^k \in \{0,1\}^k$ and $k \geqslant 0$. Hence, there is no a^k that satisfies the condition $(\star)$.

If $W^{b^m} \preccurlyeq W^{c^m}$, we consider two scenarios, based upon the difference between the Hamming weights of the two binary sequences, $\mathrm{wt}(c^m) - \mathrm{wt}(b^m)$.

(i) $\mathrm{wt}(c^m) - \mathrm{wt}(b^m) \geqslant n$: Applying the PO in [5] and the PO of Theorem 2, we get $W^{b^m 1^n} \preccurlyeq W^{c^m 0^n}$. According to Theorem 3, this implies that the condition $(\star)$ is satisfied for all $a^k \in \{0,1\}^k$ and any $k \geqslant 0$.

(ii) $\mathrm{wt}(c^m) - \mathrm{wt}(b^m) < n$: it is non-trivial since the a^k that satisfies the condition $(\star)$ is dependent on the underlying channel W, the values of (m, n), and the sequences b^m, c^m. A stronger result about the existence of a^k is discussed in Section IV-C when W is a BEC.

The following corollary is immediate.

Corollary 4. *Let $b^m = 0^m$ and $c^m = 1^m$, assume $m < n$. If for some $k^* \geqslant 0$, $W^{0^m a^{k^*} 1^n} \preccurlyeq W^{1^m a^{k^*} 0^n}$ holds for all $a^{k^*} \in \{0,1\}^{k^*}$, then*

$$W^{0^m a^k 1^n} \preccurlyeq W^{1^m a^k 0^n} \quad (28)$$

holds for all $a^k \in \{0,1\}^k$ and $k \geqslant k^*$.

C. The Condition $(\star)$

Theorem 3 provides a generalized PO that does not require the bit-channels to have the same Hamming weight, provided the ordering in the condition $(\star)$ holds. The next proposition shows that, if the channel W is a BEC, then for any $m, n \geqslant 1$, the condition $(\star)$ holds for some sequence a^k, for sufficiently large k.

Proposition 3. *Let $W = BEC(\epsilon)$ for some $\epsilon \in (0,1)$, and $m, n \geqslant 1$. Let b^m, c^m be binary sequences of length m such that $W^{b^m} \preccurlyeq W^{c^m}$. Then, for sufficiently large k, there exists a finite-length sequence $a^k \in \{0,1\}^k$, such that*

$$Z(W^{b^m a^k 1^n}) \geqslant Z(W^{c^m a^k 0^n}). \quad (29)$$

Therefore, $W^{b^m a^k 1^n} \preccurlyeq W^{c^m a^k 0^n}$.

Proof: Consider the two functions

$$f_0(x) = 2x - x^2,$$

$$f_1(x) = x^2,$$

both of which are increasing on $[0, 1]$. Define

$$F_{a^k} = f_{a_k} \circ \cdots \circ f_{a_2} \circ f_{a_1}.$$

Then we have

$$Z(W^{b^m a^k 1^n}) = f_1^{(n)}[F_{a^k}(F_{b^m}(\epsilon))] = [F_{a^k}(F_{b^m}(\epsilon))]^{2^n}, \quad (30)$$

$$Z(W^{c^m a^k 0^n}) = f_0^{(n)}[F_{a^k}(F_{c^m}(\epsilon))] = 1-[1-F_{a^k}(F_{c^m}(\epsilon))]^{2^n}. \quad (31)$$

We want to find a^k such that

$$Z(W^{b^m a^k 1^n}) \geqslant Z(W^{c^m a^k 0^n}), \quad (32)$$

that is

$$[F_{a^k}(F_{b^m}(\epsilon))]^{2^n} + [1 - F_{a^k}(F_{c^m}(\epsilon))]^{2^n} \geqslant 1. \quad (33)$$

Let $(a_\ell)_{\ell \geqslant 1}$ be a binary sequence. Then almost surely, the function F_{a^k} exhibits a threshold behavior as k grows large. To be more precise, according to [9, Lemma 11], there exists $\epsilon^* \in [0, 1]$ such that

$$\lim_{k\to\infty} F_{a^k}(\epsilon) = \begin{cases} 0, & \epsilon \in [0, \epsilon^*) \\ 1, & \epsilon \in (\epsilon^*, 1] \end{cases}. \quad (34)$$

Here, ϵ^* depends on the realization of a^k, and has a uniform distribution on $[0, 1]$. We use this result to complete the proof.

Since $W^{b^m} \preccurlyeq W^{c^m}$, then we have $0 < F_{c^m}(\epsilon) < F_{b^m}(\epsilon) < 1$. We can therefore find a threshold point ϵ^* such that $F_{c^m}(\epsilon) < \epsilon^* < F_{b^m}(\epsilon)$. Therefore, there exists a sequence $(a_\ell)_{\ell \geqslant 1}$ and a sufficiently large k such that

$$F_{a^k}(F_{b^m}(\epsilon)) \geqslant 1 - \delta(n),$$

$$F_{a^k}(F_{c^m}(\epsilon)) \leqslant \delta(n),$$

where $\delta(n) \triangleq 1 - 2^{-\frac{1}{2^n}}$. These inequalities imply that

$$[F_{a^k}(F_{b^m}(\epsilon))]^{2^n} + [1 - F_{a^k}(F_{c^m}(\epsilon))]^{2^n} \geqslant 2[1 - \delta(n)]^{2^n} = 1$$

and thus (32) holds. The degradation relation follows from the fact that the bit-channels are BECs. ■

V. Numerical Examples

In this section, we consider some applications of Theorem 3 and Corollary 4 to BECs in the case where $b^m = 0^m$ and $c^m = 1^m$.

A. Application of Theorem 3

The Bhattacharyya parameters of several synthesized bit-channels for BECs when $N = 32$ are shown in Fig. 2. In contrast to Fig. 1, the intersecting curves are shown. Since the curves can be described explicitly as polynomials in erasure probability ϵ, we can numerically determine the intersection points for any pair of intersecting curves. Table I lists all values of $\epsilon \in (0, 1)$ for which a pair of bit-channels $(W_N^{(i)}, W_N^{(j)})$, denoted by (i, j), have an intersection point. For example, the curves corresponding to $W_{32}^{(7)}$ and $W_{32}^{(20)}$ intersect at two locations, $\epsilon = 0.3077$ and $\epsilon = 0.5772$.

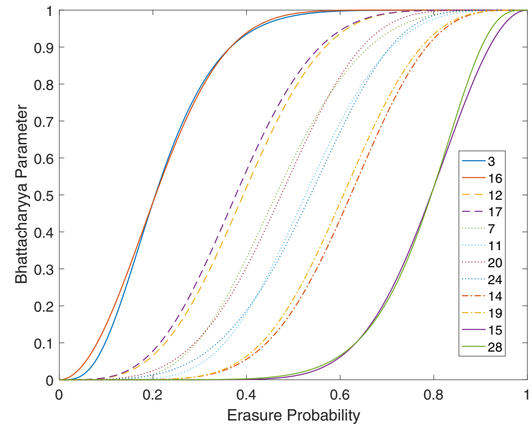

Fig. 2. BEC bit-channels with intersections when $N = 32$.

TABLE I
The intersections of all intersecting pairs when $N = 32$.

(3,16)	(12,17)	(7,20)	(7,24)	(11,24)	(14,19)	(15,28)
0.1997	0.7493	0.3077	0.1831	0.4228	0.2507	0.6358
0.3642		0.5772	0.8169	0.6923		0.8003

Since the curve for $W_{32}^{(7)}$ is above the curve for $W_{32}^{(20)}$ when $0.3077 < \epsilon < 0.5772$, this implies that $W_{32}^{(7)} \preccurlyeq W_{32}^{(20)}$ when $W = \text{BEC}(\epsilon)$ for ϵ in this range. Applying Theorem 3, we obtain the following results.

Example 1. Suppose $W = \text{BEC}(\epsilon)$, where $0.3077 < \epsilon < 0.5772$. We have seen that $W_{32}^{(7)} \preccurlyeq W_{32}^{(20)}$, i.e., $W^{00111} \preccurlyeq W^{10100}$. Then, for $N = 64$, Theorem 3 implies that

$$W^{001011} \preccurlyeq W^{101000}, \text{ i.e., } W_{64}^{(11)} \preccurlyeq W_{64}^{(40)},$$

$$W^{001111} \preccurlyeq W^{101100}, \text{ i.e., } W_{64}^{(15)} \preccurlyeq W_{64}^{(44)}.$$

More generally, for $N = 2^{5+h}$, we have

$$W^{001d^h11} \preccurlyeq W^{101d^h00}$$

holds for all $d^h \in \{0, 1\}^h$ and $h \geqslant 1$.

B. Example of Corollary 4

The minimum possible k^* in Corollary 4 over different (m, n) for BEC(ϵ) is denoted by $k^*_{min}(m, n, \epsilon)$. In Table II, we consider the case where $\epsilon = 0.5$ and list values of k^*_{min} for various (m, n).

TABLE II
Values of k^*_{min} for various (m, n) when $W = \text{BEC}(0.5)$.

m \ n	1	2	3	4	5	6	7	8
1	0	3	8	10	12	13	14	15
2	0	0	0	1	3	4	5	5
3	0	0	0	0	0	0	0	1

The following example illustrates how to interpret the results in Table II.

$$F'(\epsilon) = 2^{m+n}[(1-(1-\epsilon)^{2^m})^{2^n-1}(1-\epsilon)^{2^m-1} - (1-\epsilon^{2^m})^{2^n-1}\epsilon^{2^m-1}] \tag{49}$$

$$= 2^{m+n}[(\frac{1-(1-\epsilon)^{2^m}}{1-\epsilon^{2^m}})^{2^n-1} - (\frac{\epsilon}{1-\epsilon})^{2^m-1}](1-\epsilon)^{2^m-1}(1-\epsilon^{2^m})^{2^n-1}. \tag{50}$$

Example 2. Consider two cases $(m, n) = (1, 2)$ and $(m, n) = (2, 4)$. From Table II we find that the corresponding k^*_{min} values are 3 and 1, respectively. It follows that, for $W = \text{BEC}(0.5)$, we have $W^{0a^k11} \preccurlyeq W^{1a^k00}$ for all $a^k \in \{0,1\}^k$, $k \geqslant 3$. Similarly, $W^{00a^k1111} \preccurlyeq W^{11a^k0000}$ for all $a^k \in \{0,1\}^k$, $k \geqslant 1$.

C. Properties of $k^*_{min}(m, n, \epsilon)$

We now prove some additional properties of the parameter $k^*_{min}(m, n, \epsilon)$ defined in Section V-B.

1) Structural Property of Table II: The following proposition considers "L"-shaped regions in tables for $k^*_{min}(m, n, \epsilon)$, similar to Table II.

k_3	
k_1	k_2

Proposition 4. *The values in any "L"-shaped region satisfy*

$$k_1 \leqslant k_2 \leqslant k_3.$$

Proof: Since W is a BEC, we have

$$W^0 \preccurlyeq W \preccurlyeq W^1. \tag{35}$$

The proof includes two parts.

(i) To prove $k_1 \leqslant k_2$, we need to show

$$W^{0^m a^k 1^{n+1}} \preccurlyeq W^{1^m a^k 0^{n+1}} \Rightarrow W^{0^m a^k 1^n} \preccurlyeq W^{1^m a^k 0^n}.$$

According to (35), we have

$$W^{0^m a^k 1^n} \preccurlyeq W^{0^m a^k 1^{n+1}} \preccurlyeq W^{1^m a^k 0^{n+1}} \preccurlyeq W^{1^m a^k 0^n}. \tag{36}$$

(ii) To prove $k_2 \leqslant k_3$, we need to show

$$W^{0^m a^k 1^n} \preccurlyeq W^{1^m a^k 0^n} \Rightarrow W^{0^{m+1} a^k 1^{n+1}} \preccurlyeq W^{1^{m+1} a^k 0^{n+1}}.$$

Applying Theorem 3 to the assumption gives

$$W^{0^m a^k 0 1^n} \preccurlyeq W^{1^m a^k 0 0^n}. \tag{37}$$

Furthermore, by applying the PO of Theorem 2, we have

$$W^{0^{m+1} a^k 1^n} = W^{0^m 0 a^k 1^n} \preccurlyeq W^{0^m a^k 0 1^n}. \tag{38}$$

Combining (37) and (38) gives

$$W^{0^{m+1} a^k 1^n} \preccurlyeq W^{1^m a^k 0^{n+1}}. \tag{39}$$

By applying Lemma 1 to (39), we get

$$W^{0^{m+1} a^k 1^{n+1}} = W^{0^{m+1} a^k 1^n 1} \preccurlyeq W^{1^m a^k 0^{n+1} 1}. \tag{40}$$

Again, by using the PO, $W^{a^k 0^{n+1} 1} \preccurlyeq W^{1 a^k 0^{n+1}}$, so

$$W^{1^m a^k 0^{n+1} 1} \preccurlyeq W^{1^m 1 a^k 0^{n+1}} = W^{1^{m+1} a^k 0^{n+1}}. \tag{41}$$

Combining (40) and (41), we get

$$W^{0^{m+1} a^k 1^{n+1}} \preccurlyeq W^{1^{m+1} a^k 0^{n+1}}. \tag{42}$$

■

Remark 6. For a fixed ϵ, $k^*_{min}(m, n, \epsilon)$ is increasing as m decreases or as n increases.

2) Pairs (m, n) for which $k^*_{min}(m, n, 0.5) = 0$: Notice that $k^*_{min}(m, n, 0.5) = 0$ means $W^{0^m 1^n} \preccurlyeq W^{1^m 0^n}$, i.e.,

$$Z(W^{0^m 1^n}) \geqslant Z(W^{1^m 0^n}) \tag{43}$$

when $W = \text{BEC}(0.5)$. In general, when $W = \text{BEC}(\epsilon)$, the Bhattacharyya parameters of these two bit-channels can be expressed as

$$Z(W^{0^m 1^n}) = f_1^{(n)}(f_0^{(m)}(\epsilon)) = [1-(1-\epsilon)^{2^m}]^{2^n}, \tag{44}$$

$$Z(W^{1^m 0^n}) = f_0^{(n)}(f_1^{(m)}(\epsilon)) = 1-(1-\epsilon^{2^m})^{2^n}. \tag{45}$$

Combining (43) translates to

$$F(\epsilon) \triangleq (1-\epsilon^{2^m})^{2^n} + [1-(1-\epsilon)^{2^m}]^{2^n} \geqslant 1. \tag{46}$$

Specializing to $\epsilon = 0.5$, the inequality yields an upper bound on n, namely

$$n \leqslant -\log_2[2^m - \log_2(2^{2^m} - 1)] \triangleq U(m). \tag{47}$$

The following proposition, proved in Appendix A, shows that $n = 2^m - 1$ is the largest value for which $k^*_{min}(m, n, 0.5) = 0$.

Proposition 5. *For any $m \geqslant 1$,*

$$2^m - 1 < U(m) < 2^m. \tag{48}$$

3) Range of ϵ where $k^*_{min}(m, n, \epsilon) = 0$: To explore the range of ϵ where $k^*_{min}(m, n, \epsilon) = 0$ for a given pair (m, n), we need to figure out when $F(\epsilon) \geqslant 1$. Since $F(\epsilon)$ is a symmetric function for any pair (m, n), we only need to look at the behavior of $F(\epsilon)$ on the interval $(0, 0.5)$. The derivative of $F(\epsilon)$ is given in (49) and (50). Define $g(\epsilon, m) = \frac{1-(1-\epsilon)^{2^m}}{1-\epsilon^{2^m}}$. Here, we consider three different cases based on the value of the pair (m, n).

(i) $n \leqslant m$: We make use of the following result.

Lemma 3. $1 > g(\epsilon, m+1) > g(\epsilon, m) > 0$, *for any $\epsilon \in (0, 0.5)$ and any $m \geqslant 1$.*

Proof: When $\epsilon \in (0, 0.5)$, we have

$$\begin{aligned} g(\epsilon, m+1) &= \frac{1-(1-\epsilon)^{2^{m+1}}}{1-\epsilon^{2^{m+1}}} \\ &= g(\epsilon, m) \cdot \frac{1+(1-\epsilon)^{2^m}}{1+\epsilon^{2^m}} \\ &> g(\epsilon, m). \end{aligned} \tag{51}$$

It is also clear that $0 < g(\epsilon, m) < 1$ for any $\epsilon \in (0, 0.5)$ and any $m \geqslant 1$. ∎

From Lemma 3, we have

$$[g(\epsilon, m)]^{2^n-1} \geqslant [g(\epsilon, m)]^{2^m-1} > [g(\epsilon, 0)]^{2^m-1} \tag{52}$$

for any $\epsilon \in (0, 0.5)$ and any $m \geqslant 1$. Referring to (50), we see that for $n \leqslant m$, $F'(\epsilon) > 0$ for any $\epsilon \in (0, 0.5)$. That is, $F(\epsilon)$ is increasing on $(0, 0.5)$. Hence,

$$F(\epsilon) \geqslant 1, \text{ for any } \epsilon \in [0, 1].$$

(ii) $m < n \leqslant 2^m - 1$: We will use the following lemma.

Lemma 4. *For any $x \in (0, 1)$ and any $n \geqslant 0$, we have*

$$1 - nx \leqslant (1-x)^n \leqslant 1 - nx + \frac{n(n-1)}{2}x^2. \tag{53}$$

The lower bound is Bernoulli's inequality, and the upper bound follows from the Taylor series of $(1-x)^n$.

Using the expression for $F(\epsilon)$ in (46) and Lemma 4, we get

$$\begin{aligned} F(\epsilon) &\leqslant 1 - 2^n\epsilon^{2^m} + \frac{2^n(2^n-1)}{2}\epsilon^{2^{m+1}} + (2^m\epsilon)^{2^n} && (54) \\ &\leqslant 1 - 2^n\epsilon^{2^m} + \frac{2^n(2^n-1)}{2}\epsilon^{2^{m+1}} + 2^{m2^n}\epsilon^{2^{m+1}} && (55) \\ &= 1 + \epsilon^{2^m}[(\frac{2^n(2^n-1)}{2} + 2^{m2^n})\epsilon^{2^m} - 2^n], && (56) \end{aligned}$$

where (55) follows from $n \geqslant m+1$. By setting $(\frac{2^n(2^n-1)}{2} + 2^{m2^n})\epsilon^{2^m} - 2^n = 0$, we can find a value of ϵ, denoted by $\eta(m, n)$, such that $F(\epsilon) \leqslant 1$. Then, for any $\epsilon \in (0, \eta(m, n))$, we have $F(\epsilon) < 1$. The upper bound $2^m - 1$ on n guarantees that $F(0.5) > 1$, so $F(\epsilon) = 1$ has at least one solution in $[\eta(m, n), 0.5)$. Denoting the largest one by $\epsilon^*_{max}(m, n)$, we conclude that

$$F(\epsilon) \geqslant 1, \text{ for any } \epsilon \in [\epsilon^*_{max}(m, n), 1 - \epsilon^*_{max}(m, n)].$$

(iii) $n \geqslant 2^m$: For any $\epsilon \in (0, 0.5)$, if $F(\epsilon) < 1$ for $n = 2^m$, then $F(\epsilon) < 1$ for $n > 2^m$ since $F(\epsilon)$ decreases as n increases for a fixed m. For $m = 1, 2, 3$, and $n = 2^m$, a computer search confirms that the only real roots of $F'(\epsilon) = 0$ occur at $\epsilon = 0$, 0.5, and 1. This implies that $F(\epsilon) < 1$, for any $\epsilon \in (0, 1)$, as can be seen in the corresponding curves in Fig. 3. Consequently, for $m = 1, 2, 3$, and any $\epsilon \in (0, 1)$, we have

$$W^{0^m 1^{2^m}} \succcurlyeq W^{1^m 0^{2^m}}.$$

We conjecture that this is also true for $m \geqslant 4$.

We can analyze the asymptotic behavior of bit-channels $W^{0^m 1^{2^m}}$ and $W^{1^m 0^{2^m}}$ based on the following lemma.

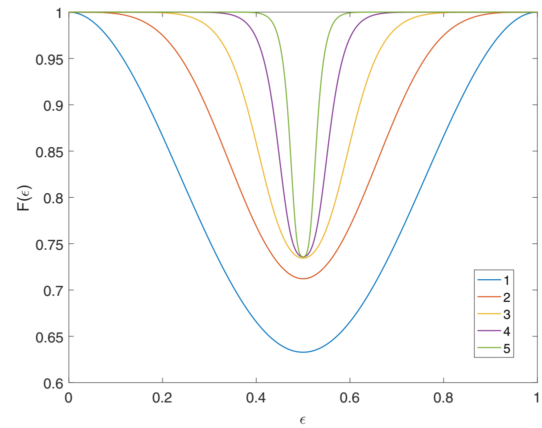

Fig. 3. The function $F(\epsilon)$ when $m = 1, 2, 3, 4, 5$ and $n = 2^m$.

Lemma 5.

$$\lim_{n\to\infty}(1-\epsilon^n)^{2^n} = \begin{cases} 1, & \epsilon \in [0, 0.5) \\ e^{-1}, & \epsilon = 0.5 \\ 0, & \epsilon \in (0.5, 1] \end{cases}. \tag{57}$$

Proof: We use the following well-known inequality:

$$(1+x)^r \leqslant e^{rx}, \text{ for any } x > -1 \text{ and } r > 0. \tag{58}$$

According to (58), we have

$$(1-\epsilon^n)^{2^n} \leqslant e^{-(2\epsilon)^n}. \tag{59}$$

When $\epsilon \in [0, 0.5)$, lower and upper bounds on the limit in (57) follow from Lemma 4 and (59), respectively; namely, we have

$$1 - (2\epsilon)^n \leqslant (1-\epsilon^n)^{2^n} \leqslant e^{-(2\epsilon)^n}. \tag{60}$$

From (60), it follows that $\lim_{n\to\infty}(1-\epsilon^n)^{2^n} = 1$.

When $\epsilon = 0.5$, we use the expression $\lim_{x\to 0}(1+x)^{\frac{1}{x}} = e$ to obtain $\lim_{n\to\infty}(1-\frac{1}{2^n})^{2^n} = e^{-1}$.

When $\epsilon \in (0.5, 1]$, the upper bound in (59) implies $\lim_{n\to\infty}(1-\epsilon^n)^{2^n} = 0$. ∎

Combining (44), (45) and Lemma 5, we find the asymptotic values of the Bhattacharyya parameters for the two bit-channels:

$$\left(Z(W^{0^m 1^{2^m}}), Z(W^{1^m 0^{2^m}})\right) \xrightarrow{m\to\infty} \begin{cases} (0, 0), & \epsilon \in [0, 0.5) \\ (e^{-1}, 1-e^{-1}), & \epsilon = 0.5 \\ (1, 1), & \epsilon \in (0.5, 1] \end{cases}. \tag{61}$$

VI. Discussion

The degradation relation in the condition $(\star)$ of Theorem 3, which states that $W^{b^m a^k 1^n} \preccurlyeq W^{c^m a^k 0^n}$, may be difficult to verify for BMS channels W other than the BECs. It would therefore be desirable to have a similar bit-channel ordering result based upon a more easily verified condition. The following two propositions represent a step in that direction.

Proposition 6. *For any BMS channel* W *and an* $a^k \in \{0,1\}^k$*, if* $Z(W^{b^m a^k 1^n}) \geqslant Z(W^{c^m a^k 0^n})$*, then*

$$Z(W^{b^m a^k 11^n}) \geqslant Z(W^{c^m a^k 10^n}). \quad (62)$$

Proposition 7. *For any BMS channel* W *and an* $a^k \in \{0,1\}^k$*, if* $P_e(W^{b^m a^k 1^n}) \geqslant P_e(W^{c^m a^k 0^n})$*, then*

$$P_e(W^{b^m a^k 01^n}) \geqslant P_e(W^{c^m a^k 00^n}). \quad (63)$$

Remark 7. The proofs of Proposition 6 and Proposition 7 are derived from equations (7) and (8), respectively, using the PO of Theorem 2.

The next two lemmas give some useful relationships between the Bhattacharyya parameter and the error probability for channels ordered according to these parameters.

Lemma 6. *For any two BMS channels* W *and* V*, if* $Z(W) \geqslant \delta_1 \geqslant \delta_2 \geqslant Z(V)$*, where* δ_1 *and* δ_2 *satisfy*

$$1 - \sqrt{1-\delta_1^2} \geqslant \delta_2, \quad (64)$$

then $P_e(W) \geqslant P_e(V)$*.*

Lemma 7. *For any two BMS channels* W *and* V*, if* $P_e(W) \geqslant \lambda_1 \geqslant \lambda_2 \geqslant P_e(V)$*, where* λ_1 *and* λ_2 *satisfy*

$$2\lambda_1 \geqslant \sqrt{1-(1-2\lambda_2)^2}, \quad (65)$$

then $Z(W) \geqslant Z(V)$*.*

Remark 8. Both of the lemmas above are based on (3).

APPENDIX A
PROOF OF PROPOSITION 5

Proof: Referring to (47), we can rewrite $U(m)$ as

$$U(m) = -\log_2[-\log_2(1-\frac{1}{2^{2^m}})]. \quad (66)$$

Set $r(m) = \frac{1}{2^{2^m}}$. Then, from the Taylor series of $\ln(1-x)$, we have

$$-\log_2(1-r(m)) = \frac{1}{\ln 2}\sum_{i=1}^{\infty}\frac{r(m)^i}{i}. \quad (67)$$

The summation in (67) can be bounded from below and above, respectively, by

$$\sum_{i=1}^{\infty}\frac{r(m)^i}{i} > r(m), \quad (68)$$

and

$$\sum_{i=1}^{\infty}\frac{r(m)^i}{i} < \sum_{i=1}^{\infty} r(m)^i = \frac{r(m)}{1-r(m)}. \quad (69)$$

Applying (67) and (68) in (66), we obtain the upper bound

$$\begin{aligned} U(m) &< -\log_2\left[\frac{1}{\ln 2}\cdot r(m)\right] \\ &= 2^m + \log_2 \ln 2 \\ &< 2^m. \end{aligned} \quad (70)$$

On the other hand, applying (67) and (69) in (66), we obtain the lower bound

$$\begin{aligned} U(m) &> -\log_2\left[\frac{1}{\ln 2}\cdot\frac{r(m)}{1-r(m)}\right] && (71)\\ &= \log_2\left[(2^{2^m}-1)\cdot\ln 2\right] && (72)\\ &> \log_2(2^{2^m-1}) && (73)\\ &= 2^m - 1. && (74) \end{aligned}$$

The inequality in (73) is based on the following lemma.

Lemma 8. *For any* $m \geqslant 1$*,*

$$(2^{2^m}-1)\cdot\ln 2 > 2^{2^m-1}. \quad (75)$$

Inequality (75) is equivalent to

$$1 - \frac{1}{2\ln 2} > \frac{1}{2^{2^m}}. \quad (76)$$

As the right-hand side of (76) decreases as m increases, it suffices to check the case $m = 1$, which can be verified numerically.

The proof is complete. ∎

ACKNOWLEDGMENT

The work was supported by the Center for Memory and Recording Research (CMRR) at UC San Diego and NSF Grants CCF-1405119 and CCF-1619053.

REFERENCES

[1] E. Arıkan, “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] R. Mori and T. Tanaka, “Performance and construction of polar codes on symmetric binary-input memoryless channels,” in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Seoul, South Korea, Jul. 2009, pp. 1496-1500.

[3] P. Trifonov, “Efficient design and decoding of polar codes,” *IEEE Trans. Commun.*, vol. 60, no. 11, pp. 3221-3227, Nov. 2012.

[4] I. Tal and A. Vardy, “How to construct polar codes,” *IEEE Trans. Inf. Theory*, vol. 59, no. 10, pp. 6562-6582, Oct. 2013.

[5] R. Mori and T. Tanaka, “Performance of polar codes with the construction using density evolution,” *IEEE Commun. Lett.*, vol. 13, no. 7, pp. 519-521, Jul. 2009.

[6] C. Schürch, “A partial order for the synthesized channels of a polar code,” in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Barcelona, Spain, Jul. 2016, pp. 220-224.

[7] M. Bardet, V. Dragoi, A. Otmani, and J.-P. Tillich, “Algebraic properties of polar codes from a new polynomial formalism,” in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Barcelona, Spain, Jul. 2016, pp. 230-234.

[8] M. Mondelli, S. H. Hassani, R. Urbanke, “Construction of polar codes with sublinear complexity,” in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Aachen, Germany, Jun. 2017, pp. 1853-1857.

[9] S. H. Hassani, K. Alishahi, R. L. Urbanke, “Finite-length scaling for polar codes,” *IEEE Trans. Inf. Theory*, vol. 60, no. 10, pp. 5875-5898, Oct. 2014.

[10] M. El-Khamy, H. Mahdavifar, G. Feygin, J. Lee and I. Kang, “Relaxed polar codes,” *IEEE Trans. Inf. Theory*, vol. 63, no. 4, pp. 1986-2000, Apr. 2017.

[11] E. Ordentlich and R. M. Roth, “On the pointwise threshold behavior of the binary erasure polarization subchannels,” in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Aachen, Germany, Jun. 2017, pp. 859-963.

[12] J. M. Renes, D. Sutter and S. H. Hassani, “Alignment of polarized sets,” *IEEE J. Sel. Area. Comm.*, vol. 34, no. 2, pp. 224-238, Feb. 2016.