The Exponent of a Polarizing Matrix Constructed from the Kronecker Product
Myung-Kyu Lee and Kyeongcheol Yang

Abstract—The asymptotic performance of a polar code under successive cancellation decoding is determined by the exponent of its polarizing matrix. We first prove that the partial distances of a polarizing matrix constructed from the Kronecker product are simply expressed as a product of those of its component matrices. We then show that the exponent of the polarizing matrix is shown to be a weighted sum of the exponents of its component matrices. These results may be employed in the design of a large polarizing matrix with high exponent.

Index Terms—Polar codes, channel polarization, rate of polarization, partial distances, exponent, Kronecker product.

I. INTRODUCTION
Channel polarization introduced by Arıkan [1] is a method to construct a class of capacity-achieving codes, called polar codes, for symmetric binary-input discrete memoryless channels (BI-DMCs). Since polar codes are constructed by a well-defined rule and are provably capacity-achieving, they have attracted much attention. The probability of block error for polar coding based on Arıkan’s construction under successive cancellation (SC) decoding was analyzed by Arıkan and Telatar [2]. Mori and Tanaka employed density evolution in this framework, and the results were extended to general polarizing matrices [3]. Recently, Korada et al. constructed new polar codes using larger matrices than the $2 \times 2$ matrix proposed by Arıkan and analyzed their polarization rate via the partial distances and exponent [4].

A method to construct polar codes of length $l = l_1 l_2 \ldots l_N$ is to employ a generator matrix of the form $A_1 \otimes \cdots \otimes A_N$, where $\otimes$ denotes the Kronecker product and each $A_i$ is an $l_i \times l_i$ polarizing matrix [5]. One interesting problem is to analyze the characteristics of such a polarizing matrix. In this paper, we study the partial distances and the exponent of a polarizing matrix $A \otimes B$ where $A$ and $B$ are $l_1 \times l_1$ and $l_2 \times l_2$ polarizing matrices, respectively. We first prove that the partial distances of $A \otimes B$ are directly determined by those of $A$ and $B$. These results can be generalized to a polarizing matrix of the form $A_1 \otimes \cdots \otimes A_N$. Finally, we give design examples to illustrate that our results may be employed in the design of a large polarizing matrix with high exponent.

The outline of the paper is as follows. In Section II, we give some basic notation and definitions, and review briefly the partial distances and the exponent of a polarizing matrix. In Section III, we introduce Hamming weight functions associated with the Kronecker and Hadamard products. Our main results on the partial distances and the exponent of a polarizing matrix constructed from the Kronecker product are given in Section IV. In Section V, some design examples are presented. Finally, we give some concluding remarks in Section VI.

II. PRELIMINARIES

A. Basic Notation and Definitions
Let $\mathbb{F}$ be a field and $\mathbb{F}^l$ the $l$-dimensional vector space of all $l$-tuple vectors over $\mathbb{F}$. Given two vectors $a = [a_1, a_2, \ldots, a_l]$ and $b = [b_1, b_2, \ldots, b_l]$, the Hadamard product $a \circ b$ and the vector addition $a + b$ are defined as

\[
\begin{align*}
\mathbf{a} \circ \mathbf{b} & \triangleq [a_1 b_1, a_2 b_2, \ldots, a_l b_l], \\
\mathbf{a} + \mathbf{b} & \triangleq [a_1 + b_1, a_2 + b_2, \ldots, a_l + b_l],
\end{align*}
\]

respectively. Clearly, the vector addition and the Hadamard product are associative and commutative, that is,

\[
\begin{align*}
\mathbf{a} + (\mathbf{b} + \mathbf{c}) & = (\mathbf{a} + \mathbf{b}) + \mathbf{c}, \\
\mathbf{a} \circ (\mathbf{b} \circ \mathbf{c}) & = (\mathbf{a} \circ \mathbf{b}) \circ \mathbf{c},
\end{align*}
\]

for any $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{F}^l$. It is also easily checked that the Hadamard product is distributive over the addition, that is,

\[
\mathbf{a} \circ (\mathbf{b} + \mathbf{c}) = \mathbf{a} \circ \mathbf{b} + \mathbf{a} \circ \mathbf{c}
\]

for any $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{F}^l$.

For two vectors $\mathbf{a} = [a_1, a_2, \ldots, a_l]$ and $\mathbf{b} = [b_1, b_2, \ldots, b_m]$ over $\mathbb{F}$, the Kronecker product $\mathbf{a} \otimes \mathbf{b}$ is the vector of length $lm$, given by

\[
\mathbf{a} \otimes \mathbf{b} \triangleq [a_1 b_1, a_2 b_2, \ldots, a_l b_l] = [a_1 b_1, a_1 b_2, \ldots, a_1 b_m, a_2 b_1, a_2 b_2, \ldots, a_2 b_m, \ldots, a_l b_1, a_l b_2, \ldots, a_l b_m].
\]

The Kronecker product is associative, i.e., $(\mathbf{a} \otimes \mathbf{b}) \otimes \mathbf{c} = \mathbf{a} \otimes (\mathbf{b} \otimes \mathbf{c})$ for any $\mathbf{a} \in \mathbb{F}^l, \mathbf{b} \in \mathbb{F}^m, \mathbf{c} \in \mathbb{F}^n$. It is also distributive over the addition, that is,

\[
\mathbf{a} \otimes (\mathbf{b} + \mathbf{c}) = \mathbf{a} \otimes \mathbf{b} + \mathbf{a} \otimes \mathbf{c}
\]

for any $\mathbf{a} \in \mathbb{F}^l$ and any $\mathbf{b}, \mathbf{c} \in \mathbb{F}^m$.

Given an $m \times n$ matrix $A = (a_{ij})$ and an $r \times s$ matrix $B = (b_{ij})$ over $\mathbb{F}$, the Kronecker product of $A$ and $B$, denoted
by $A \otimes B$, is defined as the $nr \times ns$ matrix given by
\[
A \otimes B \equiv \begin{bmatrix}
a_{11}B & a_{12}B & \cdots & a_{1n}B \\
a_{21}B & a_{22}B & \cdots & a_{2n}B \\
\vdots & \vdots & & \vdots \\
a_{m1}B & a_{m2}B & \cdots & a_{mn}B
\end{bmatrix}.
\]

If we partition $A$ and $B$ on a row basis, that is,
\[
A = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix}, \quad B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_r \end{bmatrix},
\]
where $a_i$ and $b_j$ are the $i$th and $j$th rows of $A$ and $B$, respectively, then $A \otimes B$ may be expressed as
\[
A \otimes B = \begin{bmatrix} a_1 \otimes b_1 \\ a_2 \otimes b_2 \\ \vdots \\ a_m \otimes b_r \end{bmatrix}.
\]

Clearly, the Kronecker product of matrices is associative, that is,
\[
A \otimes (B \otimes C) = (A \otimes B) \otimes C
\]
for any matrices $A, B, C$. For simple notation, let $A^{\otimes n}$ denote the $n$th Kronecker power of $A$, given by
\[
A^{\otimes n} = A \otimes A \otimes \cdots \otimes A, \quad n \text{ times}.
\]

### B. Partial Distances and Exponent of a Polarizing Matrix

From now on, we are restricted only to the binary field $\mathbb{F}_2 = \{0, 1\}$. For a binary vector $a$, we denote $w(a)$ by its (Hamming) weight, that is, the number of nonzero components in $a$. Let $\text{supp}(a)$ be the support of $a = [a_1, a_2, \ldots, a_l]$, given by
\[
\text{supp}(a) = \{1 \leq i \leq l \mid a_i \neq 0\}.
\]
Clearly, $w(a) = |\text{supp}(a)|$. The (Hamming) distance $d(a, b)$ between two binary vectors $a$ and $b$ of length $l$ is defined as the number of positions at which the corresponding symbols are different in the two vectors. In particular,
\[
d(a, b) = w(a + b).
\]
\tag{1}

Consider the binary linear code $C$ generated by $g_1, \ldots, g_k \in \mathbb{F}_2^l$, denoted by $C = (g_1, \ldots, g_k)$. The minimum distance between $C$ and a vector $b \in \mathbb{F}_2^l$, denoted by $d(b, C)$, is defined as
\[
d(b, C) = \min_{c \in C} d(b, c).
\]
The coset of $C$ containing $b$ is defined as the set given by
\[
b + C = \{b + c \mid c \in C\}.
\]

### Definition 1

Given an $l \times l$ binary matrix $G = [g_1^T, g_2^T, \ldots, g_l^T]^T$, the partial distances $D_{G,i}$, $i = 1, \ldots, l$ are defined as
\[
D_{G,i} \triangleq d(g_i, g_{i+1}, \ldots, g_l), \quad i = 1, \ldots, l - 1
\]
\[
D_{G,l} \triangleq d(g_l, 0)
\]
where $(\cdot)^T$ is the transpose operation and $\mathbf{0}$ denotes the all-zero vector.

### Theorem 2

For any BI-DMC and any $l \times l$ polarizing matrix $G$ with partial distances $\{D_{G,i}\}_{i=1}^l$, the rate of polarization $E(G)$ is given by
\[
E(G) = \frac{1}{l} \sum_{i=1}^l \log_l D_{G,i}.
\]

For convenience, it is referred to as the exponent of the matrix $G$. It is known in \cite{4} that when $n$ is sufficiently large, the block error probability of a polar code constructed by $G^{\otimes n}$ under SC decoding, $P_e(l^n)$ can be bounded as
\[
P_e(l^n) \leq 2^{-l^n\beta}
\]
for any positive number $\beta \leq E(G)$. Due to this property, the exponent of a polarizing matrix can be employed as a meaningful performance measure of the corresponding polar code under SC decoding.

### III. Weight Functions Associated with the Kronecker and Hadamard Products

The weights of the addition, the Hadamard product and the Kronecker product of two binary vectors are well-known or easily computed. The following lemma will be useful in computing the weight of a more complicated combination of many binary vectors.

#### Lemma 3

i) For any $a, b \in \mathbb{F}_2^l$,
\[
w(a + b) = w(a) + w(b) - 2w(a \circ b).
\]

ii) For any $a \in \mathbb{F}_2^l$,
\[
w(\mathbf{0} \circ a) = w(a) - w(\mathbf{0}).
\]

iii) For any $a, b \in \mathbb{F}_2^l$,
\[
w(a \circ b) \leq \min(w(a), w(b))
\]
with equality iff $\text{supp}(a) \subseteq \text{supp}(b)$ or vice versa.

iv) For any $a_1, a_2 \in \mathbb{F}_2^l$ and any $b_1, b_2 \in \mathbb{F}_2^m$,
\[
w((a_1 \otimes b_1) \circ (a_2 \otimes b_2)) = w(a_1 \circ a_2)w(b_1 \circ b_2).
\]

v) For any $a_1, a_2 \in \mathbb{F}_2^l$ and any $b_1, b_2 \in \mathbb{F}_2^m$,
\[
w((a_1 \otimes b_1) \circ (a_2 \otimes b_2)) = w(a_1)w(b_1) + w(a_2)w(b_2) - 2w(a_1 \circ a_2)w(b_1 \circ b_2).
\]

**Proof:** i), ii) and iii) are obvious. iv) comes from ii) and the fact
\[
(a_1 \otimes b_1) \circ (a_2 \otimes b_2) = (a_1 \circ a_2) \otimes (b_1 \circ b_2).
\]
v) is directly obtained by applying i) and iv).

The following three lemmas can be easily derived by applying the mathematical induction and Lemma 5.

**Lemma 4.** For any \(a_1, \ldots, a_K \in \mathbb{F}_2^n\),

\[
w(a_1 + \cdots + a_K) = \sum_{i=1}^{K} w(a_i) - 2 \cdot \sum_{1 \leq i < j \leq K} w(a_i \circ a_j) + 4 \cdot \sum_{1 \leq i < j < i_3 \leq K} w(a_i \circ a_j \circ a_{i_3}) + \cdots + (-2)^{K-1} w(a_1 \circ a_2 \cdots \circ a_K).
\]

**Lemma 5.** For any \(a_1, \ldots, a_K \in \mathbb{F}_2^n\) and any \(b_1, \ldots, b_K \in \mathbb{F}_2^n\),

\[
w \left( (a_1 \otimes b_1) \circ \cdots \circ (a_K \otimes b_K) \right) = w(a_1 \circ \cdots \circ a_K)w(b_1 \circ \cdots \circ b_K).
\]

**Lemma 6.** For any \(a_1, \ldots, a_K \in \mathbb{F}_2^n\) and any \(b_1, \ldots, b_K \in \mathbb{F}_2^n\),

\[
w \left( \sum_{i=1}^{K} a_i \otimes b_i \right) = \sum_{i=1}^{K} w(a_i)w(b_i) - 2 \cdot \sum_{1 \leq i < j \leq K} w(a_i \circ a_j)w(b_i \circ b_j) + 4 \cdot \sum_{1 \leq i < j < i_3 \leq K} w(a_i \circ a_j \circ a_{i_3})w(b_i \circ b_j \circ b_{i_3}) + \cdots + (-2)^{K-1} w(a_1 \circ \cdots \circ a_K)w(b_1 \circ \cdots \circ b_K).
\]

In order to analyze the partial distances of a polarizing matrix \(A \otimes B\) in the next section, we need to introduce two kinds of weight functions, that is, the weight exclusion function and the weight difference function. More specifically, these two functions will be employed in proving that the partial distances of \(A \otimes B\) are expressed as a product of those of \(A\) and \(B\).

**Definition 7.** Let \(a_i \in \mathbb{F}_2^n\) for \(1 \leq i \leq K\). For \(K = 1\), let \(f_1(a_1) = w(a_1)\). For \(K \geq 2\), the weight exclusion function \(f_K(a_1; a_2, \ldots, a_K)\) is defined as

\[
f_K(a_1; a_2, \ldots, a_K) \triangleq w(a_1) - \sum_{i=2}^{K} w(a_1 \circ a_i) + \sum_{2 \leq i < j \leq K} w(a_1 \circ a_i \circ a_j) - \sum_{2 \leq i < j < i_3 \leq K} w(a_1 \circ a_i \circ a_j \circ a_{i_3}) + \cdots + (-1)^{K-1} w(a_1 \circ a_2 \cdots \circ a_K).
\]

**Lemma 8.** For any \(a_1, \ldots, a_K \in \mathbb{F}_2^n\),

\[
f_K(a_1; a_2, \ldots, a_K) = w(a_1 \circ a_2 \cdots \circ a_K)
\]

where \(\bar{a}\) denotes the complement of \(a = [a_1, a_2, \ldots, a_i]\), that is,

\[\bar{a} = [1 + a_1, 1 + a_2, \ldots, 1 + a_i].\]

In particular, \(f_K(a_1; a_2, \ldots, a_K) \geq 0\).

**Proof:** Let \(S_i\) be the support of \(a_i\). Clearly, \(w(\bar{a}) = l - |S_i| = |S_i^C|\), where \(S_i^C\) denotes the complement set of \(S_i\). Using the inclusion-exclusion principle, we have

\[
f_K(a_1; a_2, \ldots, a_K) = |S_1| - |(S_1 \cap S_2) \cup (S_1 \cap S_3) \cup \cdots \cup (S_1 \cap S_K)| = |S_1| - |S_1 \cap (S_2 \cup S_3 \cup \cdots \cup S_K)| = |S_1 \cap (S_2 \cup \cdots \cup S_K)^C| = |S_1 \cap S_2^C \cap \cdots \cap S_K^C| = w(a_1 \circ \bar{a}_2 \cdots \circ \bar{a}_K).
\]

**Lemma 9.** For any \(a_1, \ldots, a_K \in \mathbb{F}_2^n\),

\[
f_1(a_1) = f_K(a_1; a_2, a_3, \ldots, a_K) + f_{K-1}(a_1 \circ a_2; a_3, \ldots, a_K) + \cdots + f_{K-1}(a_1 \circ a_2 \circ \cdots \circ a_{K-1}; a_K) + f_{K-2}(a_1 \circ a_2 \circ \cdots \circ a_{K-2}; a_K) + \cdots + f_1(a_1 \circ a_2 \circ \cdots \circ a_{K-2} \circ \cdots \circ a_K).
\]

**Proof:** Note that \(a = a \circ (b + \bar{b})\) for any \(a, b \in \mathbb{F}_2^n\), since \(b + \bar{b} = [1, 1, \ldots, 1]\). Therefore,

\[
w(a) = w(a \circ b) + w(a \circ \bar{b}) = \sum_{x \in \{b, \bar{b}\}} w(a \circ x).
\]

Applying the above relation to \(a_1\) repeatedly, we have

\[
w(a_1) = \sum_{x_2 \in \{a_2, a_2^C\}} \sum_{x_3 \in \{a_3, a_3^C\}} \cdots \sum_{x_K \in \{a_K, a_K^C\}} w(a_1 \circ x_2 \circ \cdots \circ x_K).
\]

Using the commutativity of the Hadamard product and the definition of \(f_K\), we complete the proof.

**Definition 10.** Let \(a_1, \ldots, a_K \in \mathbb{F}_2^n\) and \(b_1, \ldots, b_K \in \mathbb{F}_2^n\).

For \(K = 1\), let \(g_1(a_1; b_1) = 0\). For \(K \geq 2\), the weight difference function is defined as

\[
g_K(a_1; a_2, \ldots, a_K; b_1; b_2, \ldots, b_K) \triangleq w(\sum_{i=1}^{K} a_i \otimes b_i) - w(a_1 \otimes b_1).
\]

Note that \(g_K\) can be expressed as a linear combination of \(f_i\)'s. For example, if we take \(K = 2\), we get

\[
g_2(a_1; a_2; b_1; b_2) = w(a_2)w(b_2) - 2w(a_1 \circ a_2)w(b_1 \circ b_2) = w(a_2)w(b_2) + w(a_1 \circ a_2) - w(a_1)w(b_1 \circ b_2) - w(a_1 \circ a_2)w(b_1 \circ b_2).
\]
Such an expression as in (4) plays a key role in proving that $g_K \geq 0$ under some conditions.

**Lemma 11.** For a positive integer $K$, let $a_1, \ldots, a_K \in \mathbb{F}_2^n$ such that $w(a_i) \leq w(a_1 + \sum_{i=2}^K e_i a_i)$ for any $e_i \in \mathbb{F}_2$. Then $g_K(a_1; a_2, \ldots, a_K; b_1; b_2, \ldots, b_K) \geq 0$ for any $b_1, b_2, \ldots, b_K \in \mathbb{F}_2^m$. In particular, $g_K = 0$ if $b_i = 0$ for all $i \geq 2$.

**Proof:** We first show that $g_K$ can be expressed as a linear combination of $f_i$'s. It is true for $g_1$ by definition. The expression for $g_2$ is given in (4). In order to illustrate such an expression by a more example, if we take $K = 3$, we have

$$g_3(a_1; a_2, a_3; b_1; b_2, b_3) = w(a_2)w(b_2) + w(a_3)w(b_3) - 2w(a_1 \circ a_2)w(b_1 \circ b_2) - 2w(a_1 \circ a_3)w(b_1 \circ b_3) - 2w(a_2 \circ a_3)w(b_2 \circ b_3) + 4w(a_1 \circ a_2 \circ a_3)w(b_1 \circ b_2 \circ b_3). \quad (5)$$

By Lemma 4, we get

$$4w(a_1 \circ a_2 \circ a_3) = w(a_1 + a_2 + a_3) - w(a_1) - w(a_2) - w(a_3) + 2w(a_1 \circ a_2) + 2w(a_2 \circ a_3) + 2w(a_1 \circ a_3). \quad (6)$$

Plugging (6) into (5), we obtain

$$g_3(a_1; a_2, a_3; b_1; b_2, b_3) = w(a_2)[w(b_2) - w(b_1 \circ b_2 \circ b_3)] + w(a_3)[w(b_3) - w(b_1 \circ b_2 \circ b_3)] - 2w(a_1 \circ a_2)[w(b_1 \circ b_2) - w(b_1 \circ b_2 \circ b_3)] - 2w(a_1 \circ a_3)[w(b_1 \circ b_3) - w(b_1 \circ b_2 \circ b_3)] - 2w(a_2 \circ a_3)[w(b_2 \circ b_3) - w(b_1 \circ b_2 \circ b_3)] + [w(a_1 + a_2 + a_3) - w(a_1)]w(b_1 \circ b_2 \circ b_3).$$

From the definition of $f_K$, we have

$$f_1(b_1 \circ b_2 \circ b_3) = w(b_1 \circ b_2 \circ b_3),$$

$$f_2(b_1 \circ b_2; b_3) = w(b_1 \circ b_2) - w(b_1 \circ b_2 \circ b_3),$$

$$f_3(b_1 \circ b_3; b_2) = w(b_1 \circ b_3) - w(b_1 \circ b_2 \circ b_3),$$

$$f_3(b_2 \circ b_1; b_3) = w(b_2 \circ b_1) - w(b_1 \circ b_2 \circ b_3),$$

$$f_3(b_3; b_1, b_2) = w(b_3) - w(b_1 \circ b_3) - w(b_1 \circ b_2 \circ b_3) + w(b_1 \circ b_2 \circ b_3).$$

Using these relations and the relation $-2w(a_i + a_j) = w(a_i + a_j) - w(a_i) - w(a_j)$, we get

$$g_3(a_1; a_2, a_3; b_1; b_2, b_3) = w(a_1 + a_2 + a_3) - w(a_1 + a_j) - w(a_i) - w(a_j) + [w(a_1 + a_2) - w(a_1)]f_2(b_1 \circ b_2; b_3) + [w(a_1 + a_3) - w(a_1)]f_2(b_1 \circ b_3; b_2) + [w(a_2 + a_3) - w(a_2)]f_2(b_2 \circ b_3; b_1) + w(a_2)f_3(b_2; b_1, b_3) + w(a_3)f_3(b_3; b_1, b_2).$$

In the same procedure as above, it is possible to express $g_K$ as

$$g_K(a_1; a_2, \ldots, a_K; b_1; b_2, \ldots, b_K) = [w(a_1 + a_2 + \cdots + a_K) - w(a_1)]f_1(b_1 \circ b_2 \circ \cdots \circ b_K) + [w(a_1 + a_2 + \cdots + a_{K-1}) - w(a_1)]f_2(b_1 \circ b_2 \circ \cdots \circ b_{K-1}; b_K) + [w(a_1 + \cdots + a_{K-2} + a_K) - w(a_1)]f_2(b_1 \circ \cdots \circ b_{K-2} \circ b_K; b_1) + \cdots + w(a_2 + a_3 + \cdots + a_K)f_2(b_2 \circ b_3 \circ \cdots \circ b_K; b_1, b_2) + \cdots + w(a_2)f_K(b_2; b_1, b_3, \ldots, b_K) + \cdots + w(a_K)f_K(b_K; b_1, b_2, \ldots, b_{K-1}). \quad (7)$$

As a second step, we note that the first factor in each term of $g_K$ is larger than or equal to 0 by the assumption on $a_1, \ldots, a_K$ and $f_i \geq 0$ for any $1 \leq i \leq K$ by Lemma 8. Therefore, we complete the proof. \qed

**IV. MAIN RESULTS**

Let $A$ be an $l_1 \times l_1$ polarizing matrix with partial distances $\{D_{A,i}\}_{i=1}^{l_1}$ and $B$ an $l_2 \times l_2$ polarizing matrix with partial distances $\{D_{B,i}\}_{i=1}^{l_2}$, given by

$$A = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_{l_1} \end{bmatrix}, \quad B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_{l_2} \end{bmatrix},$$

where $a_i$ is the $i$th row of $A$ and $b_j$ is the $j$th row of $B$. Note that $A \otimes B$ is an $l_1 l_2 \times l_1 l_2$ polarizing matrix and every integer $k$ with $1 \leq k \leq l_1 l_2$ can be uniquely expressed as $k = (i-1)l_2 + j$ with $1 \leq i \leq l_1$ and $1 \leq j \leq l_2$. Our first problem is to determine the partial distances of the polarizing matrix $A \otimes B$ in terms of those of $A$ and $B$.

**Theorem 12.** The partial distances of the polarizing matrix $A \otimes B$ are given by

$$D_{A \otimes B,i \otimes j} = D_{A,i} \cdot D_{B,j}$$

for $1 \leq i \leq l_1$ and $1 \leq j \leq l_2$.

**Proof:** We divide our problem into two cases depending on the index $i$.

Case 1) $i = 1$: Using the relation in (4), the $j$th partial distance of $A \otimes B$ is given by

$$D_{A \otimes B,j} = \min_{x_1, x_2, \ldots, x_{l_1}} \left\{ w(a_1 \otimes x_1 + \sum_{k=2}^{l_1} a_k \otimes x_k), \right\} \quad 1 \leq j \leq l_2 \quad (8)$$
where $x_1 \in b_j + (b_{j+1}, b_{j+2}, \ldots, b_{l_1})$, and $\bar{x}_2, \bar{x}_3, \ldots, \bar{x}_{l_1} \in (b_1, b_2, \ldots, b_{l_2})$. Let $a_D \in a_1 + (a_2, a_3, \ldots, a_{l_1})$ be a binary vector with minimum weight $D_{A,1}$, i.e., $D_{A,1} = w(a_D)$. Then

$$
a_D = a_1 + \sum_{k=2}^{l_1} \epsilon_k a_k
$$

with $\epsilon_k \in F_2$ for $2 \leq k \leq l_1$ and the partial distance $D_{A \otimes B} \in \mathcal{B}$ may be rewritten as

$$
D_{A \otimes B} = \min_{x_1, x_2, \ldots, x_{l_1}} w\left(a_D \otimes x_1 + \sum_{k=2}^{l_1} a_k \otimes x_k\right) \quad (9)
$$

$$
= \min_{x_1} \left(\min_{x_2, \ldots, x_{l_1}} w\left(a_D \otimes x_1 + \sum_{k=2}^{l_1} a_k \otimes x_k\right)\right) \quad (10)
$$

where $x_k = \bar{x}_k + \epsilon_k x_1$ for $k = 2, \ldots, l_1$. Using the weight difference function $g_K$ in Definition 10 we may express $w(\cdot)$ in (10) as follows:

$$
w\left(a_D \otimes x_1 + \sum_{k=2}^{l_1} a_k \otimes x_k\right) = w(a_D \otimes x_1) + g_1(a_D; a_2, \ldots, a_{l_1}; x_1; x_2, \ldots, x_{l_1}).
$$

By the choice of $a_D$ with $w(a_D) = D_{A,1}$ and Lemma 11 it is easily checked that for any $x_1 \in b_j + (b_{j+1}, \ldots, b_{l_1})$ and any $x_2, \ldots, x_{l_1} \in (b_1, \ldots, b_{l_2})$,

$$
g_1(a_D; a_2, \ldots, a_{l_1}; x_1; x_2, \ldots, x_{l_1}) \geq 0
$$

where the equality holds if $x_k = 0$ for all $k \geq 2$. Therefore, for a given binary vector $x_1$

$$
\min_{x_2, \ldots, x_{l_1}} w\left(a_D \otimes x_1 + \sum_{k=2}^{l_1} a_k \otimes x_k\right) = w(a_D \otimes x_1)
$$

This relation reduces (10) to

$$
D_{A \otimes B} = \min_{x_1} D_{A,1} \cdot w(x_1)
$$

$$
= D_{A,1} \cdot \min_{x \in (b_{j+1}, \ldots, b_{l_1})} w(b_j + x)
$$

$$
= D_{A,1} \cdot D_{B,j}
$$

for any $1 \leq j \leq l_2$.

Case 2) $i \geq 2$: Let $A^{(i)}$ be the $(l_1 - i + 1) \times l_1$ submatrix of $A$, given by

$$
A^{(i)} = [a_1^T, a_2^T, \ldots, a_{l_1}^T]^T.
$$

Then $D_{A \otimes B, (i-1)l_2+j} = D_{A^{(i)} \otimes B,j}$. In a similar approach as in Case 1), we have

$$
D_{A^{(i)} \otimes B,j} = D_{A^{(i)},1} \cdot D_{B,j}
$$

for any $1 \leq j \leq l_2$. Note that the first factor $D_{A^{(i)},1}$ is exactly equal to $D_{A,i}$. Therefore, we complete the proof.

**Theorem 13.** The exponent of the polarizing matrix $A \otimes B$ is given by

$$
E(A \otimes B) = \frac{E(A)}{\log_2 l_1 l_2} + \frac{E(B)}{\log_2 l_1 l_2}.
$$

**Proof:** By Theorems 2 and 12 we have

$$
E(A \otimes B) = \frac{1}{l_1 l_2} \sum_{i=1}^{l_1} \sum_{j=1}^{l_2} \log \frac{l_1 l_2}{l_1 l_2} D_{A \otimes B, (i-1)l_2+j}
$$

$$
= \frac{1}{l_1 l_2} \sum_{i=1}^{l_1} \sum_{j=1}^{l_2} \log \frac{l_1 l_2}{l_1 l_2} D_{A,i} \cdot D_{B,j}
$$

$$
= \frac{1}{l_1 l_2} \sum_{i=1}^{l_1} \log \frac{l_1 l_2}{l_1 l_2} D_{A,i} + \frac{1}{l_2} \sum_{j=1}^{l_2} \log \frac{l_1 l_2}{l_1 l_2} D_{B,j}
$$

$$
= \frac{1}{l_1} \sum_{i=1}^{l_1} \log \frac{l_1 l_2}{l_1 l_2} D_{A,i} + \frac{1}{l_2} \sum_{j=1}^{l_2} \log \frac{l_1 l_2}{l_1 l_2} D_{B,j}.
$$

**Remark:** $E(A \otimes B) = E(B \otimes A)$ even though $A \otimes B \neq B \otimes A$ in general.

**Corollary 14.** The exponent of the polarizing matrix $A \otimes B$ is an internally dividing point of $E(A)$ and $E(B)$. That is,

$$
E(A \otimes B) = \frac{\alpha}{1 + \alpha} E(A) + \frac{1}{1 + \alpha} E(B)
$$

where $\alpha = \log_2 l_1 \geq 0$.

**Proof:** By Theorem 13 we have

$$
E(A \otimes B) = \frac{E(A)}{1 + \log_2 l_2} + \frac{E(B)}{1 + \log_2 l_1}.
$$

Without loss of generality, we may assume that $l_1 \leq l_2$. Let $\alpha = \log_2 l_1$. Then $0 \leq \alpha \leq 1$ and

$$
E(A \otimes B) = \frac{1}{1 + 1/\alpha} E(A) + \frac{1}{1 + 1/\alpha} E(B)
$$

$$
= \frac{\alpha}{1 + \alpha} E(A) + \frac{1}{1 + \alpha} E(B).
$$

**Corollary 15.** Let $A_1, A_2$ be polarizing matrices of size $l_1 \times l_1$ and let $B_1, B_2$ be polarizing matrices of size $l_2 \times l_2$. Assume that $E(A_1) \geq E(A_2)$ and $E(B_1) > E(B_2)$, or $E(A_1) > E(A_2)$ and $E(B_1) \geq E(B_2)$. Then

$$
E(A_1 \otimes B_1) > E(A_2 \otimes B_2).
$$

Corollary 15 tells us that a polarizing matrix with higher exponent should be selected as a component matrix when we construct a polarizing matrix with higher rate of polarization from the Kronecker product.

Theorems 12 and 13 can be generalized to a polarizing matrix $A = A_1 \otimes A_2 \otimes \cdots \otimes A_N$ of length $l = l_1 l_2 \cdots l_N$ where $A_i$ is an $l_i \times l_i$ polarizing matrix for $i = 1, 2, \ldots, N$. 


Let $k$ be an integer with $1 \leq k \leq l_1l_2 \cdots l_N$. Then the $k$th partial distance of the polarizing matrix $A_1 \otimes A_2 \otimes \cdots \otimes A_N$ is given by

$$D_{A_1 \otimes A_2 \otimes \cdots \otimes A_N,k} = D_{A_1,i_1}D_{A_2,i_2} \cdots D_{A_N,i_N}$$

where $k = (i_1 - 1)l_2l_3 \cdots l_N + (i_2 - 1)l_3l_4 \cdots l_N + \cdots + (i_{N-1} - 1)l_N + i_N$ with $1 \leq i_j \leq l_j$ for $j = 1, 2, \ldots, N$.

**Proof:** Since the Kronecker product is associative, i.e., $A \otimes (B \otimes C) = (A \otimes B) \otimes C$, the statement can be easily derived in a recursive way.

Theorem 17. The exponent of the polarizing matrix $A_1 \otimes A_2 \otimes \cdots \otimes A_N$ is given by

$$E(A_1 \otimes A_2 \otimes \cdots \otimes A_N) = \sum_{i=1}^{N} \frac{E(A_i)}{\log_{l_i} l_1l_2 \cdots l_N}.$$

**Proof:** It is similar to the Proof of Theorem 13.

Corollary 18. Let $A$ be an $l \times l$ polarizing matrix. For any integer $N \geq 1$, the exponent of the $N$th Kronecker power of $A$, $A^\otimes N = A \otimes \cdots \otimes A$, is given by

$$E(A^\otimes N) = E(A).$$

### V. Design Examples

In order to illustrate the relationship between the exponent of a polarizing matrix constructed from the Kronecker product and the error rate of the corresponding polar code, some design examples are presented in this section. The following matrices are employed as a component matrix for larger polarizing matrices:

$$G_2 = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \quad G_{3,L} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix},$$

$$G_{3,H} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

where $G_2$ is proposed by Arikan [1]. $G_{3,L}$ is introduced in [4] and $G_{3,H}$ is newly designed. Using these matrices, we construct two polarizing matrices of size $6 \times 6$ given by

$$G_{6,L} = G_2 \otimes G_{3,L}, \quad G_{6,H} = G_2 \otimes G_{3,H}.$$

The partial distances and the exponents of the above matrices are given in Table I. Since $E(G_{3,H}) > E(G_{3,L})$, we have $E(G_{6,H}) > E(G_{6,L})$ as shown in Corollary 18.

We designed four half-rate polar codes whose generator matrices are $G_{6,L}^\otimes 4, G_{6,H}^\otimes 4, G_{6,L}^\otimes 5, G_{6,H}^\otimes 5$, respectively, and whose frozen bits are optimized to the binary erasure channel with erasure rate $1/2$. It is assumed that the coded bits are modulated to binary phase-shift keying (BPSK) symbols and then transmitted over an additive white Gaussian noise (AWGN) channel. Fig. 1 shows the block error rates of these polar codes under SC decoding, where $E_b$ is the received signal energy per information bit and $N_0$ is the one-sided power spectral density of the AWGN. The polar codes with $G_{6,H}$ as a component polarizing matrix have much lower error rates than those with $G_{6,L}$ in the high signal-to-noise power ratio (SNR) region. This result shows that when a polarizing matrix is constructed from the Kronecker product, it is required to select a polarizing matrix with high exponent as a component matrix.

Korada et al. [4] constructed new polarizing matrices of size $m \times m$ for $m \leq 31$ obtained by shortening a BCH code of length 31. For our reference, we denote such a $m \times m$ matrix by $G_{S,m}$. The exponent of $G_{S,m}$ provides a lower bound on the maximum exponent for polarizing matrices of size $m \times m$, defined as

$$E_m = \max_{G \in \{0,1\}^{m \times m}} E(G),$$

in a constructive way. Note that polarizing matrices with $m > 31$ may be constructed from the method proposed in [4]. However, it is a very difficult problem to calculate their exponents, since a search space for computing their

![Fig. 1. Block error rates of half-rate polar codes whose lengths are 1296 and 7776 bits over an AWGN channel.](image-url)
partial distances becomes significantly large. The difficulty may be overcome by employing the Kronecker product. As an example, for $32 \leq l \leq 128$, the exponents of $l \times l$ polarizing matrices of the form $G_2 \otimes G_3 \otimes G_m$ are easily calculated by Theorem and are presented in Table Note that these exponents may become a good lower bound on $E_l$ for $32 \leq l \leq 128$.

### VI. Conclusions

We derived the partial distances and the exponent of a polarizing matrix constructed from the Kronecker product. Our results can be employed in the design of a polarizing matrix with high exponent when it is constructed from the Kronecker product. It is expected that our approach can be generalized to the calculation of the partial distances and the exponent of a nonbinary polar code.

---

For a simple example, the size $l$ is restricted to $32 \leq l \leq 128$. Polarizing matrices of size $l \times l$ for $l \geq 129$ can be constructed in a similar approach.

---

### 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, July 2009.

[2] E. Arıkan and E. Telatar, “On the rate of channel polarization,” in Proc. IEEE Int. Symp. Inf. Theory, Seoul, South Korea, July 2009, pp. 1478-1482.

[3] 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, July 2009.

[4] S. B. Korada, E. Şaşoğlu, and R. Urbanke, “Polar codes: Characterization of exponent, bounds, and constructions,” IEEE Trans. Inf. Theory, vol. 56, no. 12, pp. 6253-6264, Dec. 2010.

[5] S. B. Korada and E. Şaşoğlu, “A class of transformations that polarize binary-input memoryless channels,” in Proc. IEEE Int. Symp. Inf. Theory, Seoul, South Korea, July 2009, pp. 1478-1482.

[6] Andrews, G. E. Number Theory. Philadelphia, PA: Saunders, 1971.

[7] E. Arıkan, “A performance comparison of polar codes and Reed-Muller codes,” IEEE Commun. Lett., vol. 12, no. 6, pp. 447-449, June 2008.