Applied Mathematics & Information Sciences
An International Journal

http://dx.doi.org/10.12785/amis/090443

On Fast Channel Polarization of Double-layer Binary Discrete Memoryless Channels

Wei Duan[1], *Dazu Huang*[2], *Ying Guo*[3] *and Moon Ho Lee*[1,*]

[1] The Institute of Information and Communication, Chonbuk National University, Chonju 561-756, Korea.
[2] Department of Information Management, Hunan University of Finance and Economics, Changsha 410205, China.
[3] School of Information Science and Engineering Central South University, Changsha 410205, China.

Received: 13 Nov. 2014, Revised: 13 Feb. 2015, Accepted: 14 Feb. 2015
Published online: 1 Jul. 2015

Abstract: Polar codes are linear codes which split input channels to increase its transition performance and provably achieve the capacity of symmetric binary discrete memoryless channels (B-DMC). The idea of Polar codes is related to the recursive construction of Reed-Muller codes on the basis of 2-order square matrix G_2, can achieve the symmetric capacity of arbitrary binary-input discrete memoryless channels and to create from N independent copies of a B-DMC W, N different channels through a linear transformation. It has already been mentioned that in principle larger matrices can be employed to construct polar codes with better performances. In this paper we consider a problem of systematic constructions of polar codes based on fast channel polarization of binary discrete memoryless channel, which is an idea approach to construct code sequences as splitting input channels to increase the cutoff rate. We analyzes a novel polar channel coding and decoding approach by using the 4×4 matrix $G_4 = G_2^{\otimes 2}$ as a core on dual binary discrete memoryless channels (D-BDMC). In this paper, we characterize its parameters for a given core square standard matrix G_4 and derive upper and lower bounds on achievable exponents of derived polar codes based on $G_{4^n} = G_4^{\otimes n}$ with block-length 4^n, through which the performance can be improved with lower encoding and decoding complexity and achieve explicit construction. We investigate polarization schemes whose salient features may be decoded with a maximize likelihood (ML) decoder, which render the schemes analytically tractable and provide powerful low-complexity coding algorithms. Moreover, we give a general family of polar codes based on Reed-Mull codes with fast channel polarization.

Keywords: Polar code, dual binary discrete memoryless channels, Reed-Mull code.

1 Introduction

Polar codes, introduced by Arikan [1], achieve the capacity of arbitrary binary-input symmetric DMCs. Moreover, they have low encoding and decoding complexity and an explicit construction. The channel polarization may be consisted of code sequences using a belief propagation (BP) decoder with symmetric high rate capacity in given binary-input discrete memoryless channels (B-DMC). It is a commonplace phenomenon that is almost impossible to avoid as long as several similar channels are combined in a sufficient density with certain elegant connections [2]. The investigation of channel polarization not only has become an interesting theoretical problem, but also have lots of practical applications in signal sequence transforms, data processing, signal processing, and code coding theory [3], [4]. Following Ariank's paper [1], authors of [5] had introduced a list successive-cancellation (SCL) decoding algorithm with consideration of L SC decoding paths, where the results showed that performance of SCL was very close to that of maximum-likelihood (ML) decoding. Then, to decrease the time complexity of the SCL, another decoding algorithm derived from SC called stack successive-cancellation (SCS) was proposed in [6]. Furthermore, it was proven in [8] that, with cyclic redundancy check (CRC) aided, SCL even outperformed more than some turbo codes.

The proposed symmetric capacity is the highest coding rate achievable subject to using the input alphabets of the channel with equal probability. It is known that polar code is the first provably capacity achieving codes for an arbitrary B-DMC with low encoding and decoding complexity [7,9,10,11]. In [12], the polarization

* Corresponding author e-mail: moonho@jbnu.ac.kr

phenomenon has been studied for arbitrary kernel matrices, rather than Ariank's original 2×2 polarization kernel, and error exponents were derived for each such kernel. The construction of polar codes is based on the transformation of $G_2^{\otimes n}$ to a block of $N=2^n$ bits that are transmitted to the output of independent copies of a B-DMC W, where the notion '$\otimes$' denotes Kronecker product. As n grows large, the channels seen by individual bits start being polarized. Namely, some channels approach to either a noiseless channel or a pure-noise channel, where the fraction of channels becoming noiseless is close to the symmetric capacity $I(W)$ that is the high rate of reliable communication channel.

Motivated by a fascinating aspect of Shannons channel coding theorem that shows the existence of capacity achieving code sequences, we show a novel construction of provably capacity-achieving sequences with low coding complexities with BP decoders. This paper is an attempt to meet this elusive goal for B-DMC, which is an extension of work where channel combining and splitting were used to to improve the sum cutoff rate [1], [3], [4].

In this paper, we consider the construction of polar codes based on the transformation of $G_4^{\otimes n}$ to a block of $N=4^n$ bits transmitted to independent copies of a dual binary discrete memoryless channels (D-BDMC) W_4. A channel is called as D-BDMC of it is composed of two pair of channels W_2 that are used as a core for channel polarization by Arikan [1]. As n grows large, the channels seen by individual bits start being fast polarized with a lower complex computation.

The organization of the paper is as follows: Section II reviews polar code and describes our proposed polarization construction. In Section III, we propose a novel polar coding algorithm. The simulation results are presented to show the good performance of the proposed method in Section IV, and section V concludes the paper.

2 Polarization Construction

In this section, we briefly review polar codes and consider a novel construction of polar codes which is based on D-BDMC and generate them from four copies of a generic B-DMC W with special characters.

First, we define two important parameters of symmetric B-DMCs: the mutual information and the Bhattacharyya parameter. Consider a generic B-DMC $W:\mathscr{X}\mapsto\mathscr{Y}$ with input alphabet $\mathscr{X}=\{0,1\}$, output alphabet $\mathscr{Y}$ and transition probabilities $W(y|x)$ for $x\in\mathscr{X}$ and $y\in\mathscr{Y}$, there are two channel parameters [1], i.e., the symmetric capacity

$$I(W)=\sum_{x\in\mathscr{X}}\sum_{y\in\mathscr{Y}}\frac{1}{2}W(y|x)\log\frac{W(y|x)}{\frac{1}{2}W(y|0)+\frac{1}{2}W(y|1)}, \qquad (1)$$

and the Bhattacharyya parameter

$$Z(W)=\sum_{y\in\mathscr{Y}}\sqrt{W(y|0)W(y|1)}. \qquad (2)$$

The two parameters are much useful while consider the measurements of rate and reliability of D-BDMC. $I(W)$ is a measure of rate in a channel. It is well-known that reliable communication is possible over a symmetric B-DMC at any rates up to $I(W)$. The Bhattacharyya parameter is a measure of the reliability of a channel since $Z(W)$ is an upper bound on the probability of maximum-likelihood (ML) decision error for uncoded transmission over W.

Furthermore, note that $Z(W)$ and $1-I(W)$ are expectations of the functions $f(x)=2\sqrt{x(1-x)}$ and $g(x)=-xlog(x)-(1-x)log(1-x)$ over the distribution P, respectively.

Denote the length of codewords we will transmit over W by $n=2^m$. Given $\mathbf{y}=(y_1,y_1,...y_n)\in\mathscr{Y}^n$ and $\mathbf{u}=(u_1,u_1,...u_n)\in\mathscr{X}^n$, thus the total underlying memoryless channel can be expressed as

$$W^n(\mathbf{y}|\mathbf{u})=\prod_{i=1}^{n}W(y_i|x_i).$$

Thus, the corresponds to n independent uses of the channel W. Let G_2 be the standard polarization matrix, shown as

$$G_2=\begin{pmatrix}1&0\\1&1\end{pmatrix},$$

for which $G_2^{-1}=G_2$.

Let $G_2^{\otimes m}$ be the m-fold Kronecker product of G and B_n be the $n\times n$ bit-reversal permutation matrix. Thus, the transition probabilities is

$$W_i(\mathbf{y},\mathbf{u}_{i-1}|u_i)=\sum_{v\in(0,1)}\frac{1}{2^n}W^n(y|(\mathbf{u}_{i-1},u_i,\mathbf{v})B_nG_2^{\otimes m}) \qquad (3)$$

The maximum-likelihood decision rule for estimating u_i is

$$\widehat{u_i}=max\{W_i(\mathbf{y},\mathbf{u}_{i-1}|0),W_i(\mathbf{y},\mathbf{u}_{i-1}|1)\}. \qquad (4)$$

This is the decision rule used in successive cancellation decoding.

In this paper, we use the notation $\mathbf{a}_1^N$ to denote a row vector $(a_1,\cdots,a_N)$. For a given vector $\mathbf{a}_1^N$, we write $\mathbf{a}_i^j$ to denote the subvector $(a_i,\cdots,a_j)$ for $j>i$. Moreover, we write $\mathbf{a}_{\mathscr{A}}$ to denote the subvector $(a_i:i\in\mathscr{A})$, and write $\mathbf{a}_{1,o}^j$ to denote the subvector $(a_i:1\le i\le j,\ i$ odd) with odd indices, and $\mathbf{a}_{1,e}^j$ to denote the subvector $(a_i:1\le i\le j,\ i$ even) with even indices. Also, we write $\mathbf{a}_{1,l}^j$ to denote the subvector $(a_i:1\le i\le j,\ i=4l+k)$ with indices $i=4l+k$ for $k\in\{1,2,3,4\}$. The notation W^N is used for denoting the combined channel corresponding to N uses of W, and hence one has $W^N:\mathscr{X}^N\mapsto\mathscr{Y}^N$ with

$$W^N(\mathbf{y}_1^N|\mathbf{x}_1^N)=\prod_{i=1}^{N}W(y_i|x_i).$$

Channel polarization over D-BDMC is an operation by which one manufactures out of independent N copies

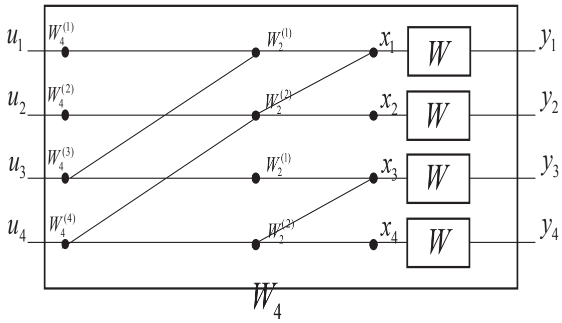


Fig. 1: The combined channel W_4 with its lower-level channel W.

of a given B-DMC W for $N = 4^n$ generate a second set of N channels $\{W_N^i : 1 \leq i \leq N\}$ that show an affection of channel polarization in a sense that, as n becomes large, for some indices $1 \leq i \leq N$ the symmetric capacity terms $\{I(W_N^i) : 1 \leq i \leq N\}$ tend towards 0 or 1. This operation consists of a channel combining phase and a channel splitting phase.

For channel combining phase, it combines N copies of a given B-DMC W in a recursive manner to produce a combined vector channel W_N for $N = 4^n$. This recursion begins at the zero level with one copy of W and hence we let $W_1 = W$. Consequently the first level of the recursion combines four independent copies of W as shown in Fig. 1 and obtains the channel $W_4 : \mathscr{X}^4 \mapsto \mathscr{Y}^4$ with the transition probabilities being calculated as

$$W_4(\mathbf{y}_1^4|\mathbf{u}_1^4) = W(y_1|\oplus_{i=1}^4 u_i)W(y_2|u_2 \oplus u_4) \quad (5)$$
$$\cdot W(y_3|u_3 \oplus u_4)W(y_4|u_4), \quad (6)$$

where $\oplus_{i=1}^n u_i \triangleq u_1 \oplus \cdots \oplus u_n$. In Fig.2, R_4 is the permutation operation that maps an input (s_1, s_2, s_3, s_4) to $v_1^4 = (s_1, s_2, s_3, s_4)$. The mapping $W_4 : \mathbf{u}_1^4 \to \mathbf{y}_1^4$ from input of W_4 to output of W^4 can be denoted by

$$W_4(\mathbf{y}_1^4|\mathbf{u}_1^4) = W^4(\mathbf{y}_1^4|\mathbf{x}_1^4) = W^4(\mathbf{y}_1^4|\mathbf{u}_1^4 G_4),$$

where $\mathbf{x}_1^4 = \mathbf{u}_1^4 G_4$ and the element of G_4 determine whether an element of u^4 apperas as a summand in the encoded word or nor, at here we consider the

$$G_4 = G_2^{\otimes 2} = \begin{pmatrix} 1\,0\,0\,0 \\ 1\,1\,0\,0 \\ 1\,0\,1\,0 \\ 1\,1\,1\,1 \end{pmatrix},$$

for which $G_4^{-1} = G_4$. Then $\mathbf{u}_1^4 G_4$ is defined as

$$[u_1 u_2 u_3 u_4] \times \begin{pmatrix} 1\,0\,0\,0 \\ 1\,1\,0\,0 \\ 1\,0\,1\,0 \\ 1\,1\,1\,1 \end{pmatrix} = \begin{pmatrix} u_1 + u_2 + u_3 + u_4 \\ u_2 + u_4 \\ u_3 + u_4 \\ u_4 \end{pmatrix},$$

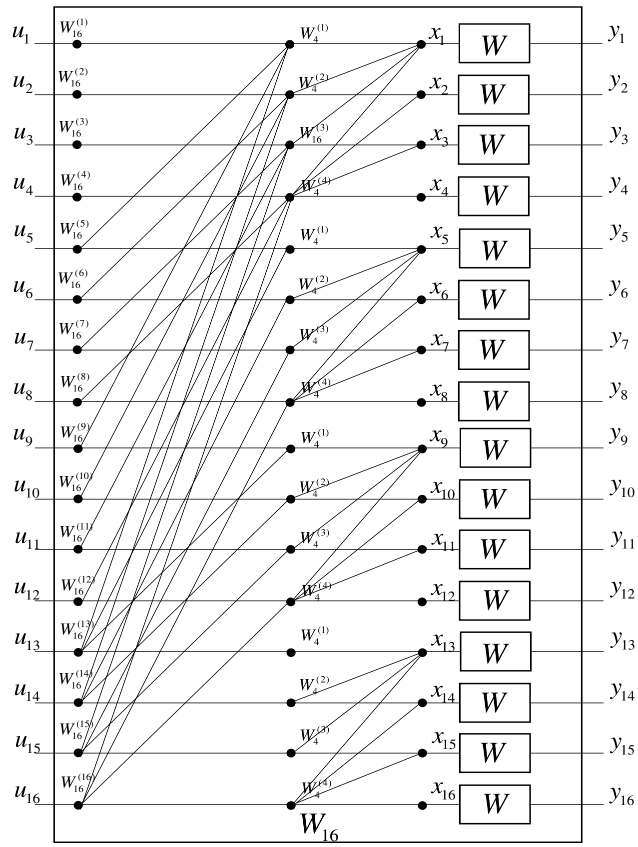


Fig. 2: The transformation of the combined channel W_{16} from its lower-level channel W_4.

Using this convention, we can define a group code based on the given binary matrix without actually defining a multiplication operation for the group.

The second level of the recursion is shown in Fig.2 where four independent copies of W_4 are combined to create the channel W_{4^2} with transition probabilities

$$W_{4^2}(\mathbf{y}_1^{16}|\mathbf{u}_1^{16}) = W_4(y_1^4|\oplus_{i=1}^4 u_i, \oplus_{i=5}^8 u_i, \oplus_{i=9}^{12} u_i, \oplus_{i=13}^{16} u_i)$$
$$\cdot W_4(y_5^8|\oplus_{i=2}^3 u_i, \oplus_{i=6}^7 u_i, \oplus_{i=10}^{11} u_i, \oplus_{i=14}^{15} u_i)$$
$$\cdot W_4(y_9^{12}|\oplus_{i=3}^4 u_i, \oplus_{i=7}^8 u_i, \oplus_{i=11}^{12} u_i, \oplus_{i=15}^{16} u_i)$$
$$\cdot W_4(y_{13}^{16}|u_4, u_8, u_{12}, u_{16}),$$

The general form of the recursion is shown in Fig. 3 where four independent copies of $W_{N/4}$ are combined to produce the channel W_N for $N = 4^n$. It is obvious that the mapping $W_N : \mathbf{u}_1^N \to \mathbf{y}_1^N$ from the input of the synthesized channel to the input of the underlying raw channels is linear over $GF(2)$. Thus it is represented by a matrix G_N so that

$$W_N(\mathbf{y}_1^N|\mathbf{u}_1^N) = W^N(\mathbf{y}_1^N|\mathbf{v}_1^N) = W^N(\mathbf{y}_1^N|\mathbf{u}_1^N G_N),$$

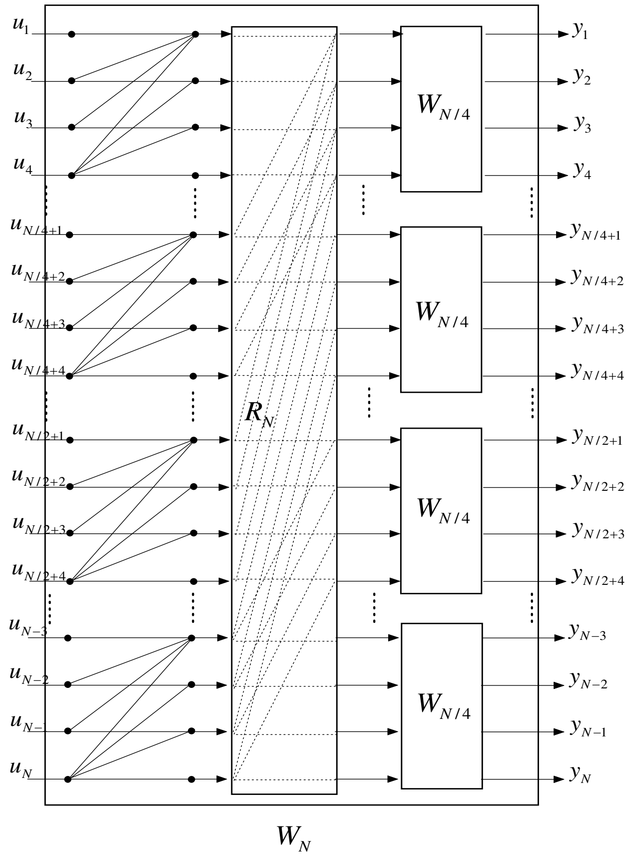


Fig. 3: A recursive construction of W_N based on four lower-order channels $W_{N/4}$.

where $\mathbf{y}_1^N \in \mathscr{Y}^N$, $\mathbf{u}_1^N \in \mathscr{X}^N$, and $G_N = G_4^{\otimes n}$.

For channel splitting phase, having synthesized the vector channel W_N out of W^N, we split W_N back into a set of binary-input coordinate channels $W_N^{(i)} : \mathscr{X} \to \mathscr{Y}^N \times \mathscr{X}^{i-1}$ for any $1 \leq i \leq N$, which is defined by the transition probabilities

$$W_N^{(i)}(\mathbf{y}_1^N, \mathbf{u}_1^{i-1}|u_i) = \sum_{\mathbf{u}_{i+1}^N} \frac{1}{4^{N-1}} W_N(\mathbf{y}_1^N|\mathbf{u}_1^N). \tag{7}$$

where the notation $(\mathbf{y}_1^N, \mathbf{u}_1^{i-1})$ represents the output of $W_N^{(i)}$ for the given input u_i. It's clear that when u_1^N is a priori uniform on $\mathscr{X}^N$, then $W_N^{(i)}$ is the effective channel seen by the ith decision element. In the basic case, $N = 2$, using chain rule of mutual information, the linear transform between (u_1, u_2) and (X_1, X_2) is a one-to-one mapping, which is shown as follow

$$I(u_1, u_2; Y_1, Y_2) = I(u_1; Y_1, Y_2) + I(u_1; Y_1, Y_2, I(u_1). \tag{8}$$

It's clear that

$$I(y_1, y_2 \mid u_1) + I(y_1, y_2, u_1 \mid u_2) = 2I(W), \tag{9}$$

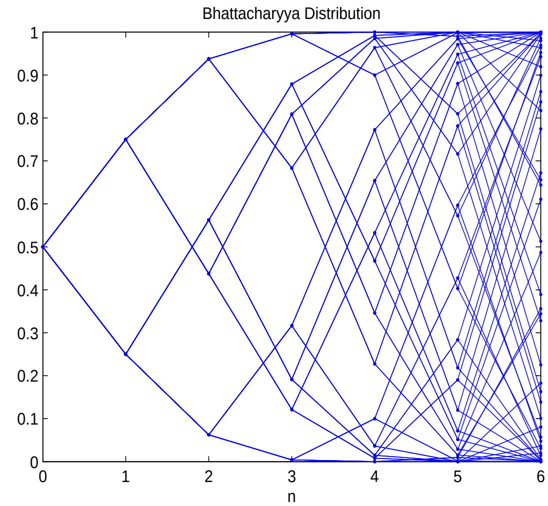


Fig. 4: Bhattacharyya bound distribution with increase of n and beginning 0.5.

$$Z(y_1, y_2 \mid u_1) \leq 2Z(W) - Z(W)^2, \tag{10}$$

$$Z(y_1, y_2, u_1 \mid u_2) = Z(W)^2, \tag{11}$$

To analyze the behavior of these channels, we may define a random process processing as the Bhattacharyya bound distribution with increase of layers n i.e., the layer of 2^3 is 3, in Fig. 4.

Proposition 1: For any D-BDMC with respect to $N = 4^n$, the channels $W_N^{(i)}$ can be polarized with large n, i.e., for any fixed $\delta \in (0,1)$, as n goes to infinite, the fraction of indices $i \in \{1, \cdots, N\}$ for which $W_N^{(i)} \in (1-\delta, 1]$ goes to $I(W)$ and the fraction for which $W_N^{(i)} \in [0, \delta]$ goes to $I(W)$.

The channel capacity $I(W_N^{(i)})$ for any $i \in \{1, \cdots, N\}$ can be calculated using the following recursive relations

$$\begin{aligned} I(W_N^{(4i-3)}) &= I(W_{N/4}^{(i)})^4 \\ I(W_N^{(4i-2)}) &= I(W_N^{(4i-1)}) = I(W_{N/4}^{(i)})^2, \\ I(W_N^{(4i)}) &= 4I(W_{N/4}^{(i)}) - 2I(W_{N/4}^{(i)})^2 - I(W_{N/4}^{(i)})^4, \end{aligned} \tag{12}$$

where $I(W_1^{(1)}) = I(W)$. In addition, to describe the performance of the proposed polar codes, the speed with which the polarization takes hold as a function of N is important. Thus, we consider to define the following parameters over splitting channel

$$Z(W_N^{(i)}) = \sum_{\mathbf{y}_1^N} \sum_{\mathbf{u}_1^{i-1}} \sqrt{W_N^{(i)}(\mathbf{y}_1^N, \mathbf{u}_1^{i-1}|0) W_N^{(i)}(\mathbf{y}_1^N, \mathbf{u}_1^{i-1}|1)}, \tag{13}$$

For any given D-BDMC based on any B-DMC W, the goal is to show that the blockwise channel transformation can be broken recursively into single-step channel transformations. Thus, we define a single-step transformation of four independent copies of a binary-input channel W as follows

$$(W,W,W,W) \rightarrow (W_4^{(1)},W_4^{(2)},W_4^{(3)},W_4^{(4)}), \qquad (14)$$

where

$$\begin{aligned} W_4^{(1)}(\mathbf{y}_1^4|u_1) &= \sum_{\mathbf{u}_2^4} \frac{1}{4} W(y_1|\oplus_{i=1}^4 u_i) W(y_2|u_2 \oplus u_4) \\ &\quad \cdot W(y_3|u_3 \oplus u_4) W(y_4|u_4) \\ W_4^{(2)}(\mathbf{y}_1^4, u_1|u_2) &= \sum_{\mathbf{u}_3^4} \frac{1}{4} W(y_1|\oplus_{i=1}^4 u_i) W(y_2|u_2 \oplus u_4) \\ &\quad \cdot W(y_3|u_3 \oplus u_4) W(y_4|u_4) \\ W_4^{(3)}(\mathbf{y}_1^4, u_2|u_3) &= \sum_{u_4} \frac{1}{4} W(y_1|\oplus_{i=1}^4 u_i) W(y_2|u_2 \oplus u_4) \\ &\quad \cdot W(y_3|u_3 \oplus u_4) W(y_4|u_4) \\ W_4^{(4)}(\mathbf{y}_1^4, u_3|u_4) &= \frac{1}{4} W(y_1|\oplus_{i=1}^4 u_i) W(y_2|u_2 \oplus u_4) \\ &\quad \cdot W(y_3|u_3 \oplus u_4) W(y_4|u_4) \end{aligned} \qquad (15)$$

Generally, for any $N = 4^n$ and $1 \le i \le N$ the above definition can be similarly extended with the following results

$$(W_N^{(i)},W_N^{(i)},W_N^{(i)},W_N^{(i)}) \rightarrow (W_{4N}^{(4i-3)},W_{4N}^{(4i-2)},W_{4N}^{(4i-1)},W_{4N}^{(4i)}), \qquad (16)$$

where

$$\begin{aligned} &W_{4N}^{(4i-3)}(\mathbf{y}_1^{4N}, \mathbf{u}_1^{4i-4}|u_{4i-3}) \\ &= \sum_{\mathbf{u}_{4i-2}^{4i}} \frac{1}{4} W_N^{(i)}(\mathbf{y}_1^N, \oplus_{i=j}^4 \mathbf{u}_{1,j}^{4i-4}| \oplus_{j=4i-3}^{4i} u_j) \\ &\quad W_N^{(i)}(\mathbf{y}_{N+1}^{2N}, \mathbf{u}_{1,2}^{4i-4} \oplus \mathbf{u}_{1,4}^{4i-4}|u_{4i-2} \oplus u_{4i}) \\ &\quad W_N^{(i)}(\mathbf{y}_{2N+1}^{3N}, \mathbf{u}_{1,3}^{4i-4} \oplus \mathbf{u}_{1,4}^{4i-4}|u_{4i-1} \oplus u_{4i}) \\ &\quad W_N^{(i)}(\mathbf{y}_{3N+1}^{4N}, \mathbf{u}_{1,4}^{4i-4}|u_{4i}) \end{aligned} \qquad (17)$$

$$\begin{aligned} &W_{4N}^{(4i-2)}(\mathbf{y}_1^{4N}, \mathbf{u}_1^{4i-3}|u_{4i-2}) \\ &= \sum_{\mathbf{u}_{4i-3}^{4i}} \frac{1}{4} W_N^{(i)}(\mathbf{y}_1^N, \oplus_{i=j}^4 \mathbf{u}_{1,j}^{4i-4}| \oplus_{j=4i-3}^{4i} u_j) \\ &\quad W_N^{(i)}(\mathbf{y}_{N+1}^{2N}, \mathbf{u}_{1,2}^{4i-4} \oplus \mathbf{u}_{1,4}^{4i-4}|u_{4i-2} \oplus u_{4i}) \\ &\quad W_N^{(i)}(\mathbf{y}_{2N+1}^{3N}, \mathbf{u}_{1,3}^{4i-4} \oplus \mathbf{u}_{1,4}^{4i-4}|u_{4i-1} \oplus u_{4i}) \\ &\quad W_N^{(i)}(\mathbf{y}_{3N+1}^{4N}, \mathbf{u}_{1,4}^{4i-4}|u_{4i}) \end{aligned} \qquad (18)$$

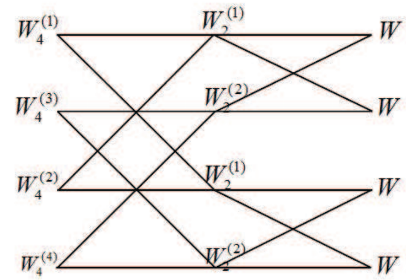


Fig. 5: The channel transformation with $N = 4$ channels.

$$\begin{aligned} &W_{4N}^{(4i-1)}(\mathbf{y}_1^{4N}, \mathbf{u}_1^{4i-3}|u_{4i-1}) \\ &= \sum_{u_{4i}} \frac{1}{4} W_N^{(i)}(\mathbf{y}_1^N, \oplus_{i=j}^4 \mathbf{u}_{1,j}^{4i-4}| \oplus_{j=4i-3}^{4i} u_j) \\ &\quad W_N^{(i)}(\mathbf{y}_{N+1}^{2N}, \mathbf{u}_{1,2}^{4i-4} \oplus \mathbf{u}_{1,4}^{4i-4}|u_{4i-2} \oplus u_{4i}) \\ &\quad W_N^{(i)}(\mathbf{y}_{2N+1}^{3N}, \mathbf{u}_{1,3}^{4i-4} \oplus \mathbf{u}_{1,4}^{4i-4}|u_{4i-1} \oplus u_{4i}) \\ &\quad W_N^{(i)}(\mathbf{y}_{3N+1}^{4N}, \mathbf{u}_{1,4}^{4i-4}|u_{4i}) \end{aligned} \qquad (19)$$

$$\begin{aligned} &W_{4N}^{(4i)}(\mathbf{y}_1^{4N}, \mathbf{u}_1^{4i-3}|u_{4i}) \\ &= \frac{1}{4} W_N^{(i)}(\mathbf{y}_1^N, \oplus_{i=j}^4 \mathbf{u}_{1,j}^{4i-4}| \oplus_{j=4i-3}^{4i} u_j) \\ &\quad W_N^{(i)}(\mathbf{y}_{N+1}^{2N}, \mathbf{u}_{1,2}^{4i-4} \oplus \mathbf{u}_{1,4}^{4i-4}|u_{4i-2} \oplus u_{4i}) \\ &\quad W_N^{(i)}(\mathbf{y}_{2N+1}^{3N}, \mathbf{u}_{1,3}^{4i-4} \oplus \mathbf{u}_{1,4}^{4i-4}|u_{4i-1} \oplus u_{4i}) \\ &\quad W_N^{(i)}(\mathbf{y}_{3N+1}^{4N}, \mathbf{u}_{1,4}^{4i-4}|u_{4i}) \end{aligned} \qquad (20)$$

where $\mathbf{u}_{1,j}^w = (u_{4k+j} : 1 \le 4k+j \le w; j = 1,2,3,4)$ for any $w > 1$. The full set of such transformation form a fabric as shown in Fig[Processing] for $N = 4$. The figure starts with two copies of transformation $(W,W) \mapsto (W_2^{(1)},W_2^{(2)})$ and in the next layer, each representing a channel transformation of the form $(W_2^{(1)},W_2^{(2)}),(W_2^{(1)},W_2^{(2)}) \mapsto (W_4^{(1)},W_4^{(3)},W_4^{(2)},W_4^{(4)})$. Independent copies are halved at each step.

From the above discussions, it is shown that the blockwise channel transformation from W^N to $W_N^{(i)}$ breaks out at a local level into single-step channel transformations. As an example, we consider a set of such transformations illustrated in Fig.2 for $N = 16$, in which the figure starts with four copies of the transformation

$$(W,W,W,W) \rightarrow (W_4^{(1)},W_4^{(2)},W_4^{(3)},W_4^{(4)})$$

and continues in butterfly-like patterns. This transformation can be generalized to represent another family of channel transformation

$$(W_{4^i}^{(k)},W_{4^i}^{(k)},W_{4^i}^{(k)},W_{4^i}^{(k)}) \rightarrow (W_{4^{i+1}}^{(4k-3)},W_{4^{i+1}}^{(4k-2)},W_{4^{i+1}}^{(4k-1)},W_{4^{i+1}}^{(4k)}).$$

The four channels at the right endpoints are always identical and independent. There are 16 independent copies of W at the rightmost level. Consider the second level to the left, there are four independent copies of $\{W_4^{(k)} : k = 1,2,3,4\}$, and so on. Each step to the left quadruples the number of channel types, but quarters the number of independent copies.

Proposition 2: For any D-BDMC and $N = 4^n$ the transformation

$$(W_N^{(i)},W_N^{(i)},W_N^{(i)},W_N^{(i)}) \to (W_{4N}^{(4i-3)},W_{4N}^{(4i-2)},W_{4N}^{(4i-1)},W_{4N}^{(4i)})$$

is rate-preserving and reliability-improving in the sense that

$$\sum_{j=0}^{3} I(W_{4N}^{(4i-j)}) = 4I(W_N^{(i)})$$
$$\sum_{j=0}^{3} Z(W_{4N}^{(4i-j)}) \leq 4Z(W_N^{(i)}) \tag{21}$$

Channel splitting moves the rate and reliability of channels away from the center in the sense that

$$I(W_{4N}^{(4i-3)}) \leq I(W_{4N}^{(4i-2)}) = I(W_{4N}^{(4i-1)}) \leq I(W_N^{(i)}) \leq I(W_{4N}^{(4i)})$$
$$Z(W_{4N}^{(4i-3)}) \geq Z(W_{4N}^{(4i-2)}) = Z(W_{4N}^{(4i-1)}) \geq Z(W_N^{(i)}) \geq Z(W_{4N}^{(4i)})$$

with equality iff equals $I(W) = 0$ or $I(W) = 1$. The reliability of channels satisfies the following condition

$$\begin{aligned} Z(W_{4N}^{(4i-3)}) &\leq 4Z(W_N^{(i)}) - Z(W_N^{(i)})^4 \\ Z(W_{4N}^{(4i-2)}) &\leq 2Z(W_N^{(i)}) - Z(W_N^{(i)})^2 \\ Z(W_{4N}^{(4i-1)}) &\leq 2Z(W_N^{(i)}) - Z(W_N^{(i)})^2 \\ Z(W_{4N}^{(4i)}) &= 2Z(W_N^{(i)})^2 + Z(W_N^{(i)})^4. \end{aligned} \tag{22}$$

The cumulative rate and reliability of the channels satisfy the condition

$$\sum_{i=1}^{N} Z(W_N^{(i)}) = N(I(W)); \sum_{i=1}^{N} Z(W_N^{(i)}) \leq NZ(W). \tag{23}$$

In a special case of a BDMC $W : \mathscr{X} \to \mathscr{Y}$ such that $W(y|0) = W(y|1)$, we have $I(W) = 1-\varepsilon$ and $Z(W) = \varepsilon$, and the term $Z(W)$ is called as erasure probability ε for erasure channel W. Thus we have the following results.

Proposition 3: Consider a binary erasure channel W with erasure probability ε, each polarized channel $W_N^{(i)}$ has the erasure probability $\varepsilon_N^{(i)}$ that can be calculated from the recursive relation

$$\begin{aligned} \varepsilon_{4k}^{(4j-3)} &= 4\varepsilon_k^{(j)} - (\varepsilon_k^{(j)})^4, \\ \varepsilon_{4k}^{(4j-2)} &= \varepsilon_{4k}^{(4j-1)} = 2\varepsilon_k^{(j)} - (\varepsilon_k^{(j)})^2, \\ \varepsilon_{4k}^{(4j)} &= (\varepsilon_k^{(j)})^4 + 2(\varepsilon_k^{(j)})^2. \end{aligned} \tag{24}$$

where $\varepsilon_1^{(1)} = Z(W) = \varepsilon$.

3 Polar Coding Algorithm

In this section, we will consider the encoding and decoding of the present polar codes. Therefore for $N = 4^n$ we present an explicit algebraic expression for encoder, the generator matrix G_N for polar coding. When consider the encoding operation $\mathbf{u}_1^N G_N$ of the polar codes, we exploit the fast transform methods popular in signal processing.

We will carry out the class of G_N-coset codes for a class of polar codes. Recall that individual G_N-coset codes are identified by $(N,K,\mathscr{A},\mu_{\mathscr{A}^c})$. For the fixed parameters $(N,K,\mathscr{A})$, we keep a free $\mu_{\mathscr{A}^c} \in \mathscr{X}^{N-K}$.

As an example, for $N = 16$ we take $G_{16} = G_4^{\otimes 2}$, i.e.,

$$G_{16} = \begin{bmatrix} 1000000000000000 \\ 1100000000000000 \\ 1010000000000000 \\ 1111000000000000 \\ 1000100000000000 \\ 1100110000000000 \\ 1010101000000000 \\ 1111111100000000 \\ 1000000010000000 \\ 1100000011000000 \\ 1010000010100000 \\ 1111000011110000 \\ 1000100010001000 \\ 1100110011001100 \\ 1010101010101010 \\ 1111111111111111 \end{bmatrix} \tag{25}$$

First we compute the reliability terms of channel polarization on the base of the vector $z(N) = (z_{N,1}, z_{N,2}, \cdots, z_{N,N})$ through the recursion

$$z_{4k,j} = \begin{cases} 4z_{k,j} - z_{k,j}^4, & \text{for } 1 \leq j \leq k; \\ 2z_{k,j-k} - z_{k,j-k}^2, & \text{for } k+1 \leq j \leq 2k; \\ 2z_{k,j-2k} - z_{k,j-2k}^2, & \text{for } 2k+1 \leq j \leq 3k; \\ 2z_{k,j-3k}^2 + z_{k,j-3k}^4, & \text{for } 3k+1 \leq j \leq 4k, \end{cases} \tag{26}$$

for any $k = 1,4,4^2,\cdots,4^{n-1}$ starting with $z_{1,1} = 1/4$. Next, we form a permutation $\pi_N = (i_1,\cdots,i_N)$ of the set $(1,\cdots,N)$ so that for any $1 \leq j < k \leq N$, the inequality $z_{N,i_j} < z_{N,i_k}$ is true.

The generator matrix $\mathscr{G}_P(N,K)$ of an (N,K) polar code is defined as the sub-matrix of G_N consisting of rows with indices $\{i_1,\cdots,i_K\} \subseteq \{1,\cdots,N\}$. It is easy to see that the computational complexity of this code construction method is $O(N\log_4 N)$.

Consider the matrix G_{16}, we have

$$z_{16} = (0.004, 0.016, 0.035, 0.049, 0.016, 0.063, 0.063, 0.141, 0.016, 0.063, 0.063, 0.141, 0.035, 0.141, 0.141, 0.316),$$

which gives

$$\pi_{16} = (16,15,14,12,8,13,4,11,10,7,6,4,9,5,3,2,1).$$

Thus a polar code can be constructed with the parameters $(16,5,\{16,15,14,12,8\})$, which has the generator matrix

$$G_{16} = \begin{bmatrix} 1111111100000000 \\ 1111000011110000 \\ 1100110011001100 \\ 1010101010101010 \\ 1111111111111111 \end{bmatrix}, \quad (27)$$

which is also the generator of a Reel-Muller code with the parameters $(16,5,8)$.

Proposition 4: For any polar code (N,K) of a block length $N = 4^n$ on D-BDMC, there is an encoder and decoder with the computational complexity $O(N\log_4 N)$.

Consider the decoder for an arbitrary G_N-coset code with parameter $(N,K,\mathscr{A},\mu_{\mathscr{A}^c})$. The source vector $\mathbf{u}_1^N$ consists of a random sub-vector $\mu_{\mathscr{A}}$ and a frozen sub-vector $\mu_{\mathscr{A}^c}$. The vector $\mathbf{u}_1^N$ is transmitted across W_N and a channel output $\mathbf{y}_1^N$ is obtained with probability $W_N(\mathbf{y}_1^N|\mathbf{u}_1^N)$. The decoder generates an estimate $\hat{\mathbf{u}}_1^N$ of $\mathbf{u}_1^N$ for the given $(\mathbf{y}_1^N,\mu_{\mathscr{A}^c})$.

The SC decoder generates an estimate $\hat{u}^N$ of u^N by observing the channel output y^N. The decoder takes N decisions for each u_i. If u_i is a frozen bit, the decoder will fix $\hat{u}_i$ to its known value. If u_i is an information bit, the decoder waits to estimate all the previous bits. For one hand, if $i \in \mathscr{A}^c$, i.e., the element u_i is known, then the i-th decision element is $\hat{u}_i = u_i$. For another, if $i \in \mathscr{A}$, then the i-th decision element has to be waited until it has received the previous decisions $\hat{\mathbf{u}}_1^{i-1}$. After that, we computes the likelihood ratio (LR) as follows

$$L_N^{(i)}(\mathbf{y}_1^N,\hat{\mathbf{u}}_1^{i-1}) = \frac{W(\mathbf{y}_1^N,\hat{\mathbf{u}}_1^{i-1}|0)}{W(\mathbf{y}_1^N,\hat{\mathbf{u}}_1^{i-1}|1)}, \quad (28)$$

and generates its decision as

$$\hat{u}_i = \begin{cases} 0, \text{ if } L_N^{(i)}(\mathbf{y}_1^N,\hat{\mathbf{u}}_1^{i-1}) \geq 1; \\ 1, \text{ otherwise.} \end{cases} \quad (29)$$

which is then sent to all succeeding decision elements. This processing is a single-pass algorithm, with no revision of estimates. The complexity of this algorithm is determined essentially by the complexity of computing the LRs.

A straightforward calculation using the recursive formulas (17)-(20) gives

$$\begin{aligned}
&L_N^{(4i-3)}(\mathbf{y}_1^N,\hat{\mathbf{u}}_1^{4i-4}) \\
&\quad= \frac{\prod_{j=1}^4 \mathscr{L}_N^{(i)}(j) + \sum_{j_1<j_2}^4 \mathscr{L}_N^{(i)}(j_1)\mathscr{L}_N^{(i)}(j_2) + 1}{\sum_{j=1}^4 \mathscr{L}_N^{(i)}(j) + \sum_{j_1<j_2<j_3}^4 \mathscr{L}_N^{(i)}(j_1)\mathscr{L}_N^{(i)}(j_2)\mathscr{L}_N^{(i)}(j_3)} \\
&L_N^{(4i-2)}(\mathbf{y}_1^N,\hat{\mathbf{u}}_1^{4i-3}) \\
&\quad= \frac{(\mathscr{L}_N^{(i)}(1))^{\phi(1)} \prod_{j=2}^4 \mathscr{L}_N^{(i)}(j) + P_1 + 1}{\mathscr{L}_N^{(i)}(1)^{\phi(1)} + \sum_{j=2}^4 \mathscr{L}_N^{(i)}(j) + P_2} \\
&L_N^{(4i-1)}(\mathbf{y}_1^N,\hat{\mathbf{u}}_1^{4i-2}) \\
&\quad= \frac{\prod_{k=1}^2 (\mathscr{L}_N^{(i)}(k))^{\phi(k)} \prod_{j=3}^4 \mathscr{L}_N^{(i)}(j) + P_3 + 1}{\sum_{j=1}^2 \mathscr{L}_N^{(i)}(k)^{\phi(k)} + \sum_{j=3}^4 \mathscr{L}_N^{(i)}(j) + P_4} \\
&L_N^{(4i)}(\mathbf{y}_1^N,\hat{\mathbf{u}}_1^{4i-1}) = \prod_{j=1}^3 (\mathscr{L}_N^{(i)}(j))^{\phi(j)} \cdot \mathscr{L}_N^{(i)}(4), \quad (30)
\end{aligned}$$

where $\phi(k) = 1 - 2\hat{u}_{4(i-1)+k}$. In [1], it is shown that the block error probability of the SC decoder decays to zero for rates below $I(W)$; consequently, polar codes achieve the capacity of symmetric B-DMCs using the SC decoder. The notations $P(i)$ for $i \in \{1,2,3,4\}$ may be expressed as

$$\begin{aligned}
P_1 &= \sum_{j_2}^4 \mathscr{L}_N^{(i)}(1)^{\phi(1)} \mathscr{L}_N^{(i)}(j_2) + \sum_{j_1<j_2}^4 \mathscr{L}_N^{(i)}(j_1)\mathscr{L}_N^{(i)}(j_2), \\
P_2 &= \sum_{j_2<j_3}^4 \mathscr{L}_N^{(i)}(1)^{\phi(1)} \mathscr{L}_N^{(i)}(j_2)\mathscr{L}_N^{(i)}(j_3) + \sum_{j_1<j_2<j_3}^4 \prod_{k=1}^3 \mathscr{L}_N^{(i)}(j_k) \\
P_3 &= \sum_{k=1}^2 \sum_{k<j_2}^4 \mathscr{L}_N^{(i)}(k)^{\phi(k)} \mathscr{L}_N^{(i)}(j_2) + \sum_{2<j_1<j_2}^4 \mathscr{L}_N^{(i)}(j_1)\mathscr{L}_N^{(i)}(j_2) \\
P_4 &= \prod_{k=1}^2 \mathscr{L}_N^{(i)}(k)^{\phi(k)} (\sum_{j=3}^4 \mathscr{L}_N^{(i)}(j)) + (\sum_{j=1}^2 \mathscr{L}_N^{(i)}(k)^{\phi(k)}) \prod_{j=3}^4 \mathscr{L}_N^{(i)}(j),
\end{aligned}$$

where $j_1 \neq 1$, $j_2 \neq 1$, and $\mathscr{L}_N^{(i)}(j)$ are function defined by

$$\begin{aligned}
&\mathscr{L}_N^{(i)}(1) = L_N^{(i)}(\mathbf{y}_1^{N/4}, \oplus_{i=j}^4 \hat{\mathbf{u}}_{1,j}^{4i-4}) \\
&\mathscr{L}_N^{(i)}(2) = L_N^{(i)}(L_N^{(i)}(\mathbf{y}_{N/4+1}^{N/2}, \hat{\mathbf{u}}_{1,2}^{4i-4} \oplus \hat{\mathbf{u}}_{1,4}^{k}), \\
&\mathscr{L}_N^{(i)}(3) = L_N^{(i)}(\mathbf{y}_{N/2+1}^{3N/4}, \hat{\mathbf{u}}_{1,3}^{k} \oplus \hat{\mathbf{u}}_{1,4}^{4i-4}), \\
&\mathscr{L}_N^{(i)}(4) = L_N^{(i)}(\mathbf{y}_{3N+1}^{4N}, \hat{\mathbf{u}}_{1,4}^{4i-4}). \quad (31)
\end{aligned}$$

Thus, the calculation of an LR at length N is reduced to the calculation of two LRs at length $N/4$. This recursion can be continued down to block length 1, at which the LRs are calculated with the formula $L_1^{(1)} = W(y_i|0)/W(y_i|1)$.

By transmitting the information bits over B-DMC W, polar-code sequences of block-length $N = 4^n$ can be constructed starting with any polarizing core matrix G_4.It is clear that the encoding and successive cancellation decoding complexities of such codes can be much lower than that of the previous schemes with complexities $O(NlogN)$. The steps of how to build the encoding and decoding construction are summarized as shown as follow:

1: Derive the relation between the channels $W_{l^{n+1}}^{(i)}$ and $W_{l^n}^{(i)}$.

2: Derive the recursive formulas between $Z(W_{l^{n+1}}^{(i)})$ and $Z(W_{l^n}^{(i)})$.

3: Derive $L(W_{l^{n+1}}^{(i)})$ and $L(W_{l^n}^{(i)})$ for ML decoding.

4: Go to 1.

4 Simulation and Discussion

In this section, we show the performance of the proposed scheme in terms of the BER performance and bounds on probability of block error with different block lengths with some simulation results. In Fig.7, we investigate the error performances under ML estimation of the polar codes for (64,32) by the use of G_2 and proposed method G_4, where BPSK modulation is applied. In this case, the proposed method achieved equivalent performance as existing polar code. Moreover, as shown in Fig.8, our proposed method does better error performance than existing polar code at high SNR regime with increasing block length, where we fixed the codes length as (128,64) in this case. Fig. 9 shows the rate versus reliability tradeoff for W using polar codes with block lengths $N \in \{4^3, 4^4, 4^5\}$. This figure is obtained by using codes whose information sets are of the form $A(\eta) = \{i \in \{1,...,N\} : Z(W_N^{(i)} < \eta)\}$, where $0 \leq \eta \leq 1$ is a variable threshold parameter. The bounds are plot of $R(\eta) = |A(\eta)|/N$ versus $B(\eta) = \sum_{i \in A(\eta)} Z(W_N^{(i)})$. The parameter is varied over a subset of $[0,1]$ to obtain the curves. It shows that the achievable capacity as the block length is increase [1].

5 Conclusion

In this paper, we suggest a polar code at block-length $N = 4^n$, based on a specific matrix G_4 and suggest the overall encoding/decoding structures and systems of the Polar code on D-BDMC, to improve the logarithm expression of encoding/decoding of the Polar code with the previous code which was proposed by Arikan. Also we derive the steps of how to build the encoding and decoding construction in detail. By transmitting the information bits over the BDMC channel W, polar codes of blocklength 4^n can be efficiently constructed starting with any polarizing matrix $G_4^{\otimes n}$. The complexity of the proposed encoding scheme are much lower that the previous which is proposed by Arikan. It can be shown that polar codes of blocklength 4^n can be constructed from generator matrices of the form G_{4^n} in which G_4 is a polarizing matrix of size 4. The encoding and successive cancellation decoding complexities of such codes are $O(N \log_4 N)$, which is better than Arikan's codes with complexities $O(N \log_2 N)$ for any $N \geq 4$.

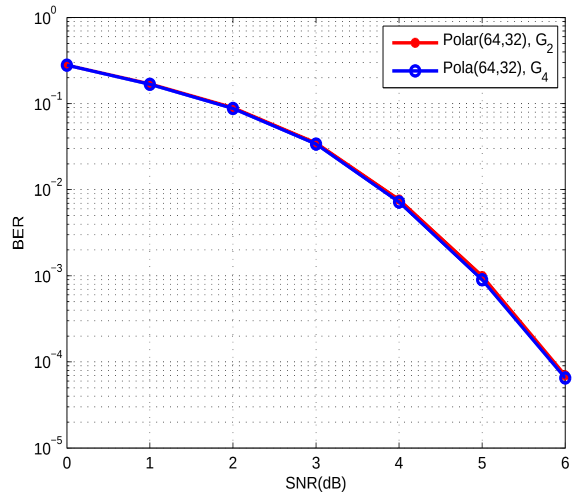


Fig. 6: Bit error rates for (64,32) polar G_2 and G_4 codes on BPSK channel.

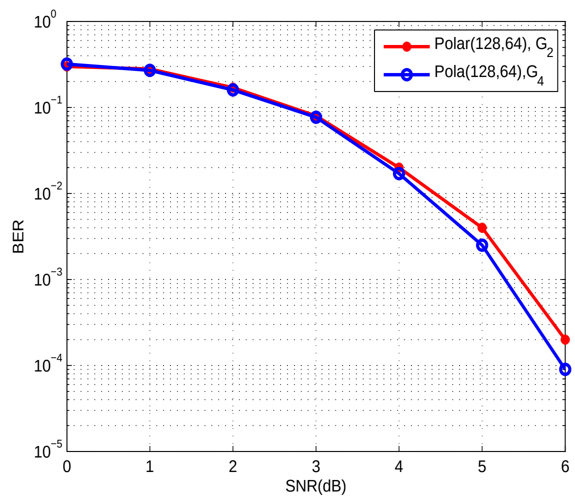


Fig. 7: Bit error rates for (128,64) polar G_2 and G_4 codes on BPSK channel.

Acknowledgements

This work was supported by MEST 2012-002521 NRF and and partly by Scientific Research Fund of Hunan

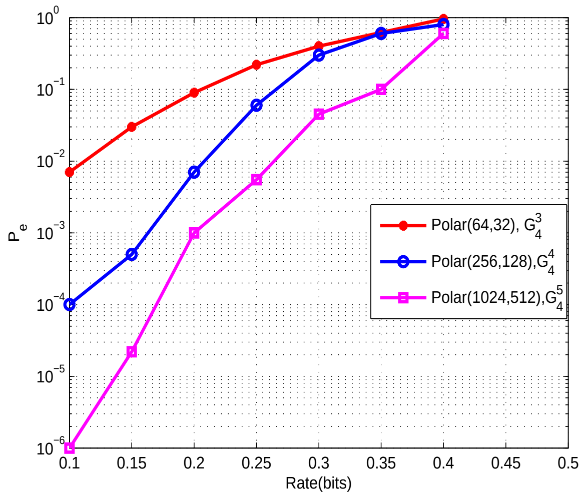


Fig. 8: Rate versus reliability for Polar code at block lengths G_{4^3}, G_{4^4}, G_{4^5}.

Provincial Education Department (13A010), the construct program of the key discipline in Hunan province and the National Natural Science Foundation of China (61272495, 61379153).

References

[1] E. Arikan, Channel Polarization : A Method for Constructing Capacity-Achieving codes for Symmetric Binary-Input Memoryless Channel, IEEE Transactions on Information Theory, Vol .55, No.7, July 2009.

[2] E. Arikan, Channel Polarization: A Method for Constructing Capacity-Achieving Codes, ISIT 2008, Toronto, Canada, July 6-11, 2008.

[3] E. Arikan, Channel Combining and Splitting for Cutoff Rate Improvement, IEEE Transactions on Information Theory, VOL.52, NO.2, FEB 2006.

[4] E. Arikan, performance comparison of polar-codes and Reed-Muller codes. IEEE. Comm. Letters, vol. 12, pp. 447-449, June 2008.

[5] I.Tal andA.Vardy, List decoding of polar codes, in Proceedings of the IEEE International Symposium on Information Theory (ISIT 11), pp. 1C5, St Petersburg, Russia,August 2011.

[6] K. Chen, K. Niu, and J. R. Lin, Improved successive cancellation decoding of polar codes, IEEE Transactions on Communications, vol. 61, no. 8, pp. 3100C3107, 2013.

[7] Ying Guo, Moon Ho Lee and Jun Li, ”A novel channel polarization on binary discrete memoryless channels,” IEEE International Conference on Communication Systems (ICCS), Singapore, 17-19 Nov. 2010.

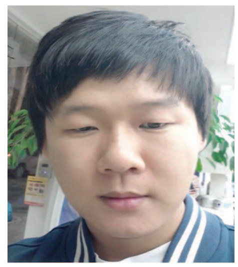
[8] K.Niu andK.Chen, CRC-aided decoding of polar codes, IEEE Communications Letters, vol. 16, no. 10, pp. 1668C1671, 2012.

[9] R. Mori, T. Tanaka, Performance and Construction of Polar codes on Symmetric Binary-Input Memoryless Channels, IEEE ISIT 2009, Seoul, Korea, June 28, 2009.

[10] R. Urbanke, S. Korada, Polar Codes: Characterization of Exponent, Bounds and Constructions, IEEE ISIT 2009, Seoul, Korea, June 28, 2009.

[11] E. Arikan, On the rate of channel polarization, IEEE ISIT 2009, Seoul, Korea, June 28, 2009.

[12] S.B. Korada, E. Sasoglu and R. Urbanke, ”Polar codes: Characterization of exponent, bounds, and constructions,” IEEE Trans. Inform. Theory, vol. 56, pp. 6253-6264, 2010.



Wei Duan received B.S degree in 2008 in Cheng Du University of Technology, China, and M.S degree in 2012 in Chonbuk National University, Korea. He is currently working towards the Ph.D. degree at the Chonbuk National University. His research interests include interference alignment (IA), space-time coding and error correcting codes.

Dazu Huang received the M.S. in 2003 in Wuhan University, China, and the Ph.D. in 2009 in the Central South University, China. His research interests include quantum information processing, digital signal processing, and wireless communication. He is now a professor in Hunan University of Finance and Economics, China.

Ying Guo received the M.S. in 2003 in Kunming University of Science and Technology, China, and the Ph.D. in 2006 in the Shanghai Jiaotong University, China. His research interests include quantum information processing and communications, wireless communication. He is now a professor in Central South University, China.

Moon Ho Lee is a professor and former chair of the Department of Electronics Engineering in Chonbuk National University, Korea. He received the Ph.D. degree from Chonnam National University, Korea in 1984, and from the University of Tokyo, Japan in 1990, both Electrical Engineering, He was in University of Minnesota, U.S.A, from 1985 to 1986 as a post-doctor. He has been working in Namyang MBC broadcasting with chief engineer from 1970 to 1980, after then he joined to Chonbuk National University as a Professor. Dr. Lee has made significant original contributions in the areas of mobile communication code design, channel coding, and multi-dimensional source and channel coding. He has authored 41 books, 160 SCI papers in international journals, and 240 papers in domestic journals, and delivered 350 papers at international conferences. Dr. Lee is a member of the National Academy of Engineering in Korea and a Foreign Fellow of the Bulgaria Academy of Sciences. He is the inventor of Jacket Matrix and it in Wikipedia was cited over 89,151 times, Aug. 15, 2014.