Abstract

This paper proposes a mapping strategy and a code construction method for $2^m$-ary polar-coded modulation. In order to find a good mapping for a polar code in an efficient way, we reduce a search space for mapping patterns by exploiting the properties of its polarizing matrix. Once a mapping is selected, the set of unfrozen information bits to define a polar code is automatically determined. Numerical results show that our approach can provide a performance gain of about 0.3 - 0.5 dB over a conventional approach using random mapping and the polar code optimized for binary phase-shift keying (BPSK) modulation, when pulse-amplitude modulation (PAM) with Gray labelling is employed.

Index Terms

Polar codes, pulse-amplitude modulation (PAM), density evolution, frozen bit.

I. INTRODUCTION

Polar codes invented by Arıkan [1] are known as a class of capacity-achieving codes for symmetric binary-input discrete memoryless channels. They were generalized to the case of non-binary input alphabets in [2]. Korada et al. [3] constructed polarizing matrices using larger
matrices than the $2 \times 2$ matrix proposed by Arikan and analyzed their polarization rate via the partial distances.

Given a polarizing matrix, constructing a polar code is equivalent to selecting unfrozen bits among the information bits. In general, it seems very complicated to find an optimal selection of unfrozen bits for arbitrary binary-input channels. For this reason, Arikan suggested a selection method based on channel approximation to the binary erasure channel having the same capacity as that of the given channel [4]. Later, Mori and Tanaka employed density evolution in order to simply find unfrozen bits [5]. Recently, Tal and Vardy proposed an efficient selection method using channel upgrading and degrading [6].

This paper studies mapping selection and code construction for $2^m$-ary polar-coded modulation. Throughout the paper, we distinguish between ‘labelling’ and ‘mapping’. Labelling in $2^m$-ary modulation denotes the rule assigning an $m$-bit pattern to each transmit symbol, e.g. Gray labelling or anti-Gray labelling, while mapping refers to the rule assigning coded bits to $m$ bit positions of transmit symbols. Figure 1 shows two examples of mappings for 4-ary pulse-amplitude modulation (PAM) when the code length is eight. In this case, four coded bits are assigned to the most significant bit (MSB) position and the other bits are assigned to the least significant bit (LSB) position since there are two bit positions corresponding to the MSB and the LSB. Since the reliability of each coded bit varies with its bit position within a transmit symbol, the performance of a polar code depends on the employed mapping as well as the selection of unfrozen bits. This paper addresses the problem of how coded bits are mapped to transmit symbols and which information bits are frozen for $2^m$-ary polar-coded modulation.

In order to find a good mapping for $2^m$-ary polar-coded modulation in an efficient way, we first reduce its search space by exploiting the properties of the polarizing matrix for a polar code. And then, we compute the reliability of each mapping in a reduced search space by applying density evolution and choose a mapping maximizing the reliability, which will be referred to as a maximum partial-polarization mapping. Once a mapping is selected, the set of unfrozen information bits to define a polar code is automatically determined. Numerical results show that our approach provides a performance gain of about 0.3 - 0.5 dB over a conventional approach using random mapping and the polar code optimized for binary phase-shift keying (BPSK) modulation, when PAM with Gray labelling is employed.
II. Polar Codes and System Model

Let $N = l_1 l_2 \cdots l_n$ where $l_1, l_2, \ldots, l_n$ are positive integers and let $F_{l_i}$ be an $l_i \times l_i$ binary polarizing matrix. An $N \times N$ polarizing matrix $G_N$ is constructed as

$$G_N = B_N (F_{l_1} \otimes F_{l_2} \otimes \cdots \otimes F_{l_n})$$

(1)

where $B_N$ is the $N \times N$ permutation matrix\(^1\) defined similarly to the bit-reversal operation for successive cancellation (SC) decoding and $\otimes$ is the Kronecker product [1], [3]. We denote by $u_N = [u_0, u_1, \ldots, u_{N-1}]$ the information bit vector, and by $c_N = [c_0, c_1, \ldots, c_{N-1}]$ the corresponding coded bit vector. The relation between these vectors is given by $c_N = u_N G_N$.

If the number of unfrozen information bits is $K$, the generator matrix $G(N, K)$ for an $(N, K)$ polar code is defined as the $K \times N$ submatrix of $G_N$ consisting of the rows with the indices of unfrozen bits. Given the polarizing matrix $G_N$ in (1), constructing a generator matrix is equivalent to selecting unfrozen bits among information bits.

Consider a $2^m$-ary polar-coded modulation system in Figure 2. The coded bits $c_0, \ldots, c_{N-1}$ are grouped and mapped into $2^m$-ary constellation symbols $x_0, \ldots, x_{N/m-1}$ where $N$ is assumed to be a multiple of $m$. The received signal is given by

$$y_k = x_k + n_k, \quad k = 0, 1, \ldots, N/m - 1$$

where $n_k$'s are independent and identically distributed (i.i.d.) Gaussian noise random variables with zero mean and variance $N_0/2$. The symbol-to-bit metric generator computes the log-likelihood ratio (LLR) of each coded bit from the received signal vector $y = [y_0, y_1, \ldots, y_{N/m-1}]$. These LLRs are the inputs to the polar decoder with SC decoding. The decoder calculates the LLR of the $i$th information bit $u_i$ as

$$L_N^{(i)}(y, \hat{u}_0^{i-1}) = \log \frac{W_N^{(i)}(y, \hat{u}_0^{i-1}|u_i = 0)}{W_N^{(i)}(y, \hat{u}_0^{i-1}|u_i = 1)}$$

(2)

where $\hat{u}_0^{i-1} = [\hat{u}_0, \ldots, \hat{u}_{i-1}]$ denotes the estimate of $u_0^{i-1} = [u_0, u_1, \ldots, u_{i-1}]$ and $W_N^{(i)}(y, \hat{u}_0^{i-1}|u_i)$ denotes the transition probability of the $i$th channel polarized by $G_N$. The LLRs in (2) are computed in a recursive manner [1], [5]. Finally, the decoder determines $\hat{u}_i$ by taking the sign of $L_N^{(i)}(y, \hat{u}_0^{i-1})$.

\(^1\)An analysis of $B_N$ is omitted here since it is a simple extension of results in [1].
The set of coded bits mapped to $2^m$-ary constellation symbol is partitioned into $m$ subsets according to their positions within a transmit symbol. For this reason, the overall channel of this transmission scheme may be decomposed into $m$ equivalent binary-input subchannels [7].

For $0 \leq l \leq m - 1$, let $C_l$ be the subset of the indices of the coded bits transmitted through the $l$th equivalent binary-input subchannel. In particular, $C_0$ and $C_{m-1}$ denote the subsets of coded bits corresponding to the most significant bits (MSBs) and the least significant bits (LSBs), respectively. Note that any mapping pattern can be classified in terms of \{ $C_0, C_1, \ldots, C_{m-1}$ \}.

For example, ‘Mapping B’ in Figure 1 corresponds to $C_0 = \{0, 2, 3, 7\}$ and $C_1 = \{1, 4, 5, 6\}$.

### III. Reduction of the Search Space for Mappings

For $2^m$-ary modulation, the LLR of a coded bit has a probability density function (PDF) depending on the subchannel over which it is transmitted. This implies that the performance of a polar code may be heavily dependent on the employed mapping strategy. Since there exist $m$ subchannels in $2^m$-ary modulation, the number of possible mapping patterns for a polar code of length $N$ is $N! / ((N/m)!)^m$. Therefore, it is impossible to find an optimal mapping maximizing the performance by an exhaustive search when $N$ goes to infinity. In order to avoid this difficulty, we need to reduce a search space for mapping patterns by exploiting the properties of $G_N$ in (1).

As a first step, we choose an integer $J$ as a design parameter which determines the size of a reduced search space. It is required to be a multiple of $m$ and be of the form

$$ J = l_k l_{k+1} \cdots l_n $$

for some $k$ with $1 \leq k \leq n$. Then the polarizing matrix in (1) can be rewritten as

$$ G_N = B_N \left( (B_{N/J}^{-1} G_{N/J}^{-1}) \otimes (B_J^{-1} G_J) \right) = B_N \left( B_{N/J}^{-1} \otimes B_J^{-1} \right) (G_{N/J} \otimes G_J) = B_N \left( B_{N/J}^{-1} \otimes B_J^{-1} \right) (G_{N/J} \otimes I_J) (I_{N/J} \otimes G_J) $$

where the second and third equalities come from the relation $(AB) \otimes (CD) = (A \otimes C)(B \otimes D)$, and $I_a$ denotes the $a \times a$ identity matrix. Note that $G_{N/J} \triangleq B_{N/J} (F_1 \otimes F_2 \otimes \cdots \otimes F_{l_k})$ and $G_J \triangleq B_J (F_{l_k} \otimes F_{l_{k+1}} \otimes \cdots \otimes F_{l_n})$ where $B_{N/J}$ and $B_J$ are the $(N/J) \times (N/J)$ and $J \times J$ permutation matrices, respectively, defined as in (1). After some manipulation, it is checked that the matrix...
$D_N \triangleq B_N \left( B_{N/J}^{-1} \otimes B_J^{-1} \right)$ is the permutation matrix which maps an input $(s_0, s_1, s_2, \ldots, s_{N-1})$ to

\[
\left( s_0, s_{N/J}, s_{2N/J}, \ldots, s_{(J-1)N/J},
\begin{array}{c}
  s_1, s_{N/J+1}, s_{2N/J+1}, \ldots, s_{(J-1)N/J+1}, \\
  \vdots \\
  s_{N/J-1}, s_{2N/J-1}, s_{3N/J-1}, \ldots, s_{N-1}
\end{array}
\right).
\]

Since $D_N$ satisfies $D_N \left( G_{N/J} \otimes I_J \right) = \left( I_J \otimes G_{N/J} \right) D_N$, we finally get

\[
G_N = \left( I_J \otimes G_{N/J} \right) D_N \left( I_{N/J} \otimes G_J \right).
\] (3)

That is, the polarization process by $G_N$ is decomposed into two polarization processes by $G_{N/J}$ and $G_J$, as shown in Figure 3.

The decomposition in (3) tells us that a mapping pattern maximizing the polarization effect of $G_J$ can improve the performance of a polar code based on $G_N$. Therefore, we focus only on good mappings for the set of $J$ coded bits connected to one subblock corresponding to $G_J$ in Figure 3, in order to reduce a search space for mappings on $N$ coded bits. Since there are $N/J$ subblocks corresponding to $G_J$ in the block corresponding to $G_N$, we reduce a search space to $\{C_0, C_1, \ldots, C_{m-1}\}$ such that each $C_l$ has the following specific pattern:

\[
C_l = \{a_{i,l} + Jp \mid 0 \leq i \leq J/m - 1, 0 \leq p \leq N/J - 1\}
\] (4)

where $a_{i,l}$ is the index of the $i$th coded bit belonging to the set $C_l \cap \{0, 1, \ldots, J - 1\}$. This approach can reduce the size of our search space from $N!/\left( \binom{N}{m} \right)^m$ to $J!/\left( \binom{J}{m} \right)^m$, which is independent of $N$. Numerical results, as will be shown in Section V, demonstrate that the mapping obtained by the proposed method in a reduced search space provides a significant performance gain over random mappings.

IV. MAXIMUM PARTIAL-POLARIZATION MAPPING AND CODE CONSTRUCTION

In this section, we explain how to find a good mapping in a reduced search space and construct the corresponding polar codes by properly selecting unfrozen bits. Due to the periodicity of
elements of $C_l$ in (4), the mapping $\{C_0, C_1, \ldots, C_{m-1}\}$ can be expressed in terms of the $l$th subset in $\{0, 1, \ldots, J - 1\}$, given by

$$C^{(J)}_l \triangleq \{a_{i,l} | 0 \leq i \leq J/m - 1\}$$

(5)

for $l = 0, 1, \ldots, m - 1$. Clearly, $C^{(J)} \triangleq \{C^{(J)}_0, \ldots, C^{(J)}_{m-1}\}$ is a partition of $\{0, 1, \ldots, J - 1\}$, that is, $C^{(J)}_i \cap C^{(J)}_j = \phi$ for any $i \neq j$ and $\bigcup_{i=0}^{m-1} C^{(J)}_i = \{0, 1, 2, \ldots, J - 1\}$. For the mapping $C^{(J)}$, the PDF of the LLRs of the coded bits in each subset is easily obtained by using the i.i.d. channel adapter under the assumption that the all-zero code word was transmitted [8]. Based on these PDFs, it is possible to track the PDF of $L^{(i)}_N(y, \hat{u}_1^{-1})$ in (2) by applying density evolution [9] to the corresponding factor graph under SC decoding. Note that density evolution can be applied to an arbitrary polarizing matrix $G_N$ in (1), although only a polarizing matrix with the size of a power of two was dealt with in [5].

Let $p^{(i)}_N(a | C^{(J)})$ be the conditional PDF of $L^{(i)}_N(y, \hat{u}_1^{-1})$ in (2), obtained by density evolution when $C^{(J)}$ is given. The error probability of the $i$th information bit can be computed as

$$P_e(u_i | C^{(J)}) = \lim_{\epsilon \to 0} \left( \int_{-\epsilon}^{-\epsilon} p^{(i)}_N(a | C^{(J)}) da + \frac{1}{2} \int_{-\epsilon}^{\epsilon} p^{(i)}_N(a | C^{(J)}) da \right).$$

(6)

The reliability measure of $C^{(J)}$ is then defined as

$$R(C^{(J)}) \triangleq \min_{A} \sum_{i \in A, |A|=K} P_e(u_i | C^{(J)})$$

(7)

where $A$ runs through all the subsets of $\{0, 1, 2, \ldots, N - 1\}$ as the set of the indices of unfrozen information bits for a polar code. Note that the inner sum in (7) is a union bound on the conditional frame error probability. Given $C^{(J)}$, the optimal selection of unfrozen information bits with respect to $R(C^{(J)})$ is given by

$$A(C^{(J)}) \triangleq \arg\min_{A} \sum_{i \in A, |A|=K} P_e(u_i | C^{(J)}).$$

That is, optimal unfrozen information bits of a polar code for $2^m$-ary modulation vary with the given mapping.

The proposed method to find a good mapping and construct a polar code (or equivalently, find the set of unfrozen bits) is summarized as follows:
• **Step 1:** Set a polarizing matrix $G_N$, the dimension $K$, and the target signal-to-noise ratio $(E_b/N_0)^*$ at which we want to maximize the code performance, where $E_b$ is the transmitted signal energy per information bit.

• **Step 2:** For each $C^{(J)}$, calculate $P_e(u_i | C^{(J)})$ by density evolution at $(E_b/N_0)^*$. Determine $R(C^{(J)})$ in (7).

• **Step 3:** Find an optimal mapping given by

$$C^{(J)}_{opt} = \arg \min_{C^{(J)}} R(C^{(J)})$$

and the set of the indices of unfrozen information bits, $A(C^{(J)}_{opt})$.

The mapping $\{C_0, C_1, \ldots, C_{m-1}\}$ corresponding to $C^{(J)}_{opt}$ maximizes the performance of a polar code among all the considered mappings. It will be referred to as a **maximum partial-polarization (MPP)** mapping with parameter $J$.

### V. Numerical Results

In order to examine the performances of polar codes with proposed mappings over $2^m$-ary modulation, we employ $F_2$, $F_3$, and $F_5$ as component matrices of a polarizing matrix $G_N$, where $F_2$ is proposed by Arikan, and $F_3$ and $F_5$ are given by

$$F_3 = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}, \quad F_5 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix},$$

respectively. These matrices have the maximum exponents [3] among the matrices with the same sizes.

Proposed MPP mappings for $2^m$-ary PAM with Gray labelling are shown in Table I with the design parameters $N, K, G_N, J$ and $(E_b/N_0)^*$ at code rate 1/2. Here, $F^\otimes b$ denotes the $b$th Kronecker power of $F$. The proposed MPP mappings have better reliabilities than any other mappings from the viewpoints of $R(C^{(J)})$. Note that an MPP mapping for given design
parameters is not unique due to the property of $G_N$. For example, in the case of 16-PAM, the mapping that exchanges $C_{0}^{(J)}$ and $C_{3}^{(J)}$ in Table I also becomes an MPP mapping.

The frame error rate (FER) and bit error rate (BER) performances of polar codes in Table I over 4-PAM, 8-PAM or 16-PAM are shown in Figures 4, 5 and 6, respectively, when SC decoding is employed. Here, ‘Code for MPP’ denotes the code constructed from $A(C_{opt}^{(J)})$, while ‘Code for BPSK’ denotes the code constructed under the assumption of BPSK modulation at $(E_b/N_0)^* = 3$ dB. In order to remove correlation among $m$ coded bits of each transmit symbol in the case of MPP mappings, we employ a random interleaver for each $C_i$ as shown in Figure 1. Furthermore, we employ as a reference ‘Random Mapping’ which selects $m$ bits randomly from $N$ coded bits and assigns them to a transmit symbol. Simulation results demonstrate that a combination of the proposed mapping and the corresponding polar code provides a performance gain of about 0.3 - 0.5 dB over that of a random mapping and the polar code optimized for BPSK modulation at BER $= 10^{-3}$. The performances for other design parameters have a similar behavior, although they are not included due to the limit of space.

VI. Concluding Remarks and Further Research

We proposed a mapping strategy and a code construction method for $2^m$-ary polar-coded modulation. Although our approach is not guaranteed to be globally optimal in the case that $J < N$, it provides a performance gain of about 0.3 - 0.5 dB over a conventional approach using random mapping and the polar code optimized for BPSK modulation without the increase of complexity. The gain may decrease a little bit due to the relative reduction of a search space when $J \ll N$.

REFERENCES


The uncoded BER of BPSK modulation at $E_b/N_0 = 3$ dB is almost the same as those of Gray-labelled 4-PAM at $E_b/N_0 = 6$ dB, 8-PAM at $E_b/N_0 = 9$ dB and 16-PAM at $E_b/N_0 = 13$ dB, respectively.


Fig. 1. Two mapping examples for a polar code of length 8 over 4-PAM

Fig. 2. Block diagram of a polar-coded modulation system
Fig. 3. Decomposition of the polarizing matrix $G_N$

Fig. 4. FER and BER performances of rate-$1/2$ polar codes of length 900 over 4-PAM
Fig. 5. FER and BER performances of polar codes of $N = 3456$ over 8-PAM

**TABLE I**

**EXAMPLES OF MPP MAPPING ACCORDING TO DESIGN PARAMETERS SUCH AS $N$, $K$, $G_N$, $J$ AND ($E_b/N_0$)**

<table>
<thead>
<tr>
<th>Mod.</th>
<th>$(N, K)$</th>
<th>$G_N$</th>
<th>$J$</th>
<th>$(E_b/N_0)$</th>
<th>MPP Mapping</th>
</tr>
</thead>
<tbody>
<tr>
<td>4-PAM</td>
<td>(900, 450)</td>
<td>$B_N(F_2^2 \otimes F_2^2 \otimes F_3^{\otimes 2})$</td>
<td>18</td>
<td>6 dB</td>
<td>$C_0^{(J)} = {3, 4, 5, 7, 8, 13, 14, 16, 17}$, $C_1^{(J)} = {0, 1, 2, 6, 9, 10, 11, 12, 15}$</td>
</tr>
<tr>
<td>8-PAM</td>
<td>(3456, 1728)</td>
<td>$B_N(F_2^{\otimes 7} \otimes F_3^{\otimes 3})$</td>
<td>9</td>
<td>9 dB</td>
<td>$C_0^{(J)} = {4, 5, 8}$, $C_1^{(J)} = {1, 2, 7}$, $C_2^{(J)} = {0, 3, 6}$</td>
</tr>
<tr>
<td>16-PAM</td>
<td>(1024, 512)</td>
<td>$B_N(F_2^{\otimes 10})$</td>
<td>8</td>
<td>13 dB</td>
<td>$C_0^{(J)} = {0, 2}$, $C_1^{(J)} = {4, 5}$, $C_2^{(J)} = {6, 7}$, $C_3^{(J)} = {1, 3}$</td>
</tr>
</tbody>
</table>
Fig. 6. FER and BER performances of polar codes of $N = 1024$ over 16-PAM