Preserving the Minimum Distance of Polar-Like Codes while Increasing the Information Length

Samet Gelincik, Philippe Mary, Anne Savard, Jean-Yves Baudais

Preserving the Minimum Distance of Polar-Like Codes while Increasing the Information Length

Samet Gelincik*, Philippe Mary*, Anne Savard†‡ and Jean-Yves Baudais*
* Univ Rennes, INSA Rennes, CNRS, IETR-UMR 6164, F-35000 Rennes, France
† IMT Nord Europe, Institut Mines Télécom, Centre for Digital Systems, F-59653 Villeneuve d'Ascq, France
‡ Univ. Lille, CNRS, Centrale Lille, UPHF, UMR 8520 - IEMN, F-59000 Lille, France

***Abstract*—Reed Muller (RM) codes are known for their good minimum distance. One can use their structure to construct polar-like codes with good distance properties by choosing the information set as the rows of the polarization matrix with the highest Hamming weight, instead of the most reliable synthetic channels. However, the information length options of RM codes are quite limited due to their specific structure. In this work, we present sufficient conditions to increase the information length by at least one bit for some underlying RM codes and in order to obtain pre-transformed polar-like codes with the same minimum distance than lower rate codes. Moreover, our findings are combined with the method presented in [1] to further reduce the number of minimum weight codewords. Numerical results show that the designed codes perform close to the meta-converse bound at short blocklengths and better than the polarized adjusted convolutional polar codes with the same parameters.**

***Index Terms*—Polar codes, Reed Muller codes, minimum distance, code extension**

I. Introduction

Polar codes are the first provably capacity achieving error correction codes with explicit construction and a low complexity decoder, based on successive cancellation for binary input memoryless channels [2]. Thanks to their low complexity decoding, polar codes are now used for control channels in the 5G networks [3] and may be used for other usages in the future wireless network generations, such as ultra-reliable low-latency communications and massive machine-type communications [4]. Eventhough polar codes are asymptotically capacity achieving, they do not show outstanding performance at short-to-moderate blocklenghts due to their poor minimum distance and a non-complete polarization.

To overcome this issue, several methods such as enhanced-Bose–Chaudhuri–Hocquenghem subcodes [5] and low-weight-bit polar codes [6] have been proposed to improve the distance spectrum. In [7], a pre-transformation method based on brute search minimizing successive cancellation error probability for a given minimum distance has been proposed. However, the method is getting complex for blocklengths higher than 64. Cyclic-redundancy-check (CRC) aided successive cancellation list (SCL) decoding, which boosts the performance by choosing the best decoding paths in a hierarchical tree, has been proposed in [8]. This work has even been enhanced by optimizing the used CRC polynomial in order to improve the minimum distance in [9], [10], and has been considered as the best code construction up to the introduction of polarized adjusted convolutional (PAC) polar codes in [11].

The PAC polar codes, proposed by Arikan [11], perform very close to the second-order rate approximation of the binary-input additive white Gaussian noise (BIAWGN) in short blocklength regime. The performance gain comes by choosing the information set of the polar code with the Reed-Muller (RM) rule, i.e., the rows of the polarization matrix that have the highest Hamming weights. Moreover, the convolutional pre-transformation allows to decrease the number of minimum weight codewords. The effect of pre-transformation has been justified in [12], where it has been proved that any pre-transformation with upper-triangular matrix, which is the case for PAC codes, does not reduce the minimum distance and can reduce the number of minimum weight codewords if properly designed. Based on [13], authors in [14] proposed a genetic algorithm tailored to PAC code construction that achieves higher rates than the ones authorized by RM rule and it performs better than the original Arikan's PAC codes.

We recently proposed another way to improve the performance of polar codes in the short blocklength regime by encoding some information bits with the sum of two or three rows of the polar encoding matrix [1]. The pairs and triplets of merged rows are determined via the connection between the binary representations of the selected row indices and their common 1 bits. The designed codes achieve the same performance as PAC codes with same parameters, and without extra computational complexity.

In this work, we extend the method proposed in [1] and state sufficient conditions to increase the information length of some polar-like codes with RM information set, i.e., to increase the rate for a given codeword length, and we explicitly give the corresponding deterministic pre-transformation matrix to sustain the same minimum distance as the RM code. We do the analysis by partitioning the indices of the encoding matrix rows with respect to the recursive structure projected on binary representation of row indices. Our proposed design is shown to perform close to the meta-converse (MC) bound and outperforms the PAC codes with the same parameters.

This work has been partially supported by IRCICA, CNRS USR 3380, Lille, France and the French National Agency for Research (ANR) under grant ANR-16-CE25-0001 ARBurst.

II. Preliminaries

A. Notations

The entries in a vector of length N are indexed from 0 to $N-1$. Any vector of length N is considered as a row vector and is denoted by x or x^{N-1}. The j^{th} entry of the vector x is denoted as x_j. The set of positive integers is $\mathbb{N}$ and the binary field is $\mathbb{F}_2$. The set of integers from j to $k-1$ is represented by $[j,k)$ or $[j,k-1]$. Uppercase calligraphic letters, such as $\mathcal{A}$, are reserved to index sets. Any index set is sorted in the ascending order and $\mathcal{A}(i)$, $i \in [0,|\mathcal{A}|)$ denotes the i-th element of $\mathcal{A}$. Specifically, we set $\mathcal{N} := [0,N)$. For any given two index sets $\mathcal{A}$ and $\mathcal{B}$, $\mathcal{A} \succ \mathcal{B}$ denotes that any element of $\mathcal{A}$ is larger than any element of $\mathcal{B}$, i.e., $\mathcal{A}(i) > \mathcal{B}(j)\ \forall i \in [0,|\mathcal{A}|)$ and $\forall j \in [0,|\mathcal{B}|)$. For a given binary vector $\mathsf{x} \in \mathbb{F}_2^{1\times N}$ and index set $\mathcal{A} \subset \mathcal{N}$, $\mathsf{x}_{\mathcal{A}}$ denotes the vector consisting of the elements of x at the positions indexed by $\mathcal{A}$. The matrices are denoted by uppercase sans serif font, e.g., G. Uppercase boldface letters denote set of vectors, e.g., $\mathbf{C}$. The indicator function is $\mathbb{I}\{\cdot\}$. The sets $\mathcal{P}_1(\cdot)$ and $\mathcal{P}_0(\cdot)$ denote the indices of 1's and 0's of a given vector, respectively.

For any $0 \le j < 2^n$, its n-bit binary representation is denoted by the vector b_j^n, or b_j if it is clear enough from the context. The ℓ-th bit position of b_j is denoted by $b_{j,\ell}$, $0 \le \ell < n$ and the indexing is started from the least significant bit, which is placed at the rightmost position. The number of 1's and 0's in a vector is represented by $i_1(\cdot)$ and $i_0(\cdot)$, respectively.

The operator $\bar{\cup}$ represents the element-wise 'OR' operation of binary vectors such that, for all $(j_1,j_2) \in [0,2^n)^2$:

$$b_{j_1,\ell} \bar{\cup} b_{j_2,\ell} = 1, \text{ if } b_{j_1,\ell} = 1 \text{ or } b_{j_2,\ell} = 1 \tag{1}$$

The operator $\bar{\cap}$ represents the element-wise 'AND' operation of binary vectors such that

$$b_{j_1,\ell} \bar{\cap} b_{j_2,\ell} = 1, \text{ if } b_{j_1,\ell} = b_{j_2,\ell} = 1 \tag{2}$$

The operator $\oplus$ denotes binary addition in $\mathbb{F}_2$.

B. Properties of the Polar Encoding Matrix

For any given $N = 2^n$, $n \in \mathbb{N}$, the polarization matrix is $\mathsf{G} = \mathsf{G}_2^{\otimes n}$ where

$$\mathsf{G}_2 := \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \tag{3}$$

is the corresponding kernel matrix and $\otimes$ is Kronecker product. The jth row g_j of G can be represented by

$$\mathsf{g}_j = \hat{\mathsf{g}}_{b_{j,n-1}} \otimes \hat{\mathsf{g}}_{b_{j,n-2}} \otimes \cdots \otimes \hat{\mathsf{g}}_{b_{j,0}} \tag{4}$$

where $\hat{\mathsf{g}}_0 = [1 \ \ 0]$ and $\hat{\mathsf{g}}_1 = [1 \ \ 1]$. By (4), for a given $n \in \mathbb{N}$, the jth row of G can be divided into n disjoint regions, i.e.,

$$\mathsf{r}_{j,\ell} = \begin{cases} \mathsf{0}^{2^\ell - 1} & \text{if } b_{j,\ell} = 0 \\ [\mathsf{r}_{j,0}\ \mathsf{r}_{j,1}\ \cdots\ \mathsf{r}_{j,\ell-1}] & \text{if } b_{j,\ell} = 1 \end{cases} \tag{5}$$

for $\ell \in [1,n)$ and $\mathsf{r}_{j,0} = \hat{\mathsf{g}}_{b_{j,0}}$. Close inspection of the recursive nature of $\mathsf{r}_{j,\ell}$ reveals that each bit position $\ell \in [0,n)$ of b_j is associated with a set of positions at g_j denoted by the index set $\mathcal{M}_\ell \subset \mathcal{N}$

$$\mathcal{M}_\ell := \{k : b_{k,\ell} = 1, k \in \mathcal{N}\} \tag{6}$$

and $\mathcal{M}_\ell^c := \mathcal{N} \setminus \mathcal{M}_\ell$. The fact that $\mathsf{g}_{j,\mathcal{M}_\ell} = \mathsf{0}^{N/2-1}$ if $b_{j,\ell} = 0$ imposes that $\mathsf{g}_{j,\mathcal{M}_\ell^c}$ is independent from the value of $b_{j,\ell}$ [1]. The following definition highlights this fact.

Definition 1. *The projection of a row g_j of the polar encoding matrix onto indices of $\mathcal{M}_\ell^c$ is denoted by g_j^ℓ and given as*

$$\mathsf{g}_j^\ell := \hat{\mathsf{g}}_{b_{j,n-1}} \otimes \cdots \otimes \hat{\mathsf{g}}_{b_{j,\ell+1}} \otimes \hat{\mathsf{g}}_{b_{j,\ell-1}} \cdots \otimes \hat{\mathsf{g}}_{b_{j,0}} \tag{7}$$

Note that, by (5), $\mathsf{b}_{j,\ell} = 1$ imposes that $[\mathsf{r}_{j,0}\mathsf{r}_{j,1}\cdots\mathsf{r}_{j,\ell-1}]$ is copied to $\mathsf{r}_{j,\ell}$ and $\mathsf{r}_{j,t>\ell}$ is obtained with respect to corresponding bit values. Hence, the projection of g_j onto $\mathcal{M}_\ell$ is the same as $\mathsf{g}_{j,\mathcal{M}_\ell^c}$ if $b_{j,\ell} = 1$

$$\mathsf{g}_{j,\mathcal{M}_\ell} = \begin{cases} \mathsf{0}^{\frac{N}{2}-1} & \text{if } b_{j,\ell} = 0 \\ \mathsf{g}_j^\ell & \text{if } b_{j,\ell} = 1 \end{cases} \tag{8}$$

The following definition is the generalization of Definition 1.

Definition 2. *The projection of row g_j of the polar encoding matrix onto $\cap_{\ell\in\mathcal{B}}\mathcal{M}_\ell^c$ is denoted by $\mathsf{g}_j^{\mathcal{B}}$ and $\mathsf{g}_j|\cap_{\ell\in\mathcal{B}}\mathcal{M}_\ell^c$, and is given as*

$$\begin{aligned} \mathsf{g}_j^{\mathcal{B}} &:= \mathsf{g}_j|\cap_{\ell\in\mathcal{B}}\mathcal{M}_\ell^c \\ &= \hat{\mathsf{g}}_{b_{j,\mathcal{W}(|\mathcal{W}|-1)}} \otimes \hat{\mathsf{g}}_{b_{j,\mathcal{W}(|\mathcal{W}|-2)}} \otimes \cdots \otimes \hat{\mathsf{g}}_{b_{j,\mathcal{W}(0)}} \end{aligned} \tag{9}$$

where $\mathcal{W} := [0,n) \setminus \mathcal{B}$.

Note that, similar to (8), for any subset $\mathcal{B}_0 \subset \mathcal{B}$, the projection of g_j onto $\cap_{\ell\in\mathcal{B}_0}\mathcal{M}_\ell \cap_{\hat{\ell}\in\mathcal{B}/\mathcal{B}_0} \mathcal{M}_{\hat{\ell}}^c$ is given by

$$\mathsf{g}_j|\cap_{\ell\in\mathcal{B}_0}\mathcal{M}_\ell \cap_{\hat{\ell}\in\mathcal{B}\setminus\mathcal{B}_0}\mathcal{M}_{\hat{\ell}}^c = \begin{cases} \mathsf{0}^{\frac{N}{|\mathcal{B}|}-1} & \text{if } \bar{\cap}_{\ell\in\mathcal{B}_0} b_{j,\ell} = 0 \\ \mathsf{g}_j^{\mathcal{B}} & \text{if } \bar{\cap}_{\ell\in\mathcal{B}_0} b_{j,\ell} = 1 \end{cases} \tag{10}$$

C. Row Merging Pre-transformed Polar-like Codes and RM Codes

A polar-like code $(N = 2^n, k) \in \mathbb{N}^2$, is constructed as

$$\mathbf{C} = \{\mathsf{c} = \mathsf{u}\mathsf{G} : \mathsf{u} \in \mathbb{F}_2^n, \mathsf{u}_{\mathcal{F}} = \mathsf{0}\} \tag{11}$$

where $\mathcal{F}$ is the index set of the frozen bit positions, and $\mathcal{A} = \mathcal{N} \setminus \mathcal{F}$ is the information set, with $|\mathcal{A}| = k$ and $|\mathcal{F}| = N - k$. For classical polar codes under SC decoding, the set $\mathcal{A}$ is the set of the most reliable bit sub-channels [2]. However in this paper, we allow to choose the information set differently. From this perspective, a $\mathrm{RM}(n,r)$ code of degree r can be seen as a polar-like code of information set

$$\mathcal{A} = \bigcup_{p=n-r}^{n} \mathcal{N}_p, \quad \mathcal{N}_p := \{t : i_1(\mathsf{b}_t) = p, t \in \mathcal{N}\}. \tag{12}$$

In [15], the minimum distance of a polar-like code is given by

$$d(\mathbf{C}) = \min_{i\in\mathcal{A}} i_1(\mathsf{g}_i) \overset{\text{(a)}}{=} 2^{\min_{i\in\mathcal{A}} i_1(\mathsf{b}_i)} \tag{13}$$

where (a) is due to [1, Theorem 2].

The pre-transformed polar-like codes [12] is obtained through a pre-transformation matrix $\mathsf{T} \in \mathbb{F}_2^{N\times N}$

$$\mathbf{C_P} = \{\mathsf{c} = \mathsf{uTG} : \mathsf{u} \in \mathbb{F}_2^n, \mathsf{u}_{\mathcal{F}} = \mathsf{0}\} \tag{14}$$

where T is an upper triangular matrix with $\mathsf{T}_{i,i} = 1$, $i \in \mathcal{N}$ and $\mathcal{F}_d := \{j : T_{i,j} = 1, i \in \mathcal{N}, j > i\}$ is the set of dynamic frozen bits. If T is restricted such that $|\{i : \mathsf{T}_{i,j}, i \in \mathcal{N}\}| \in \{1, 2\}$ $\forall j \in \mathcal{F}_d$, then T turns out to be a row merging pre-transformation matrix since some information bits are encoded with more than one row of the polarization matrix but any frozen row can be associated with at most one information row

$$\mathsf{c} = \mathsf{uTG} = \mathsf{u}\tilde{\mathsf{G}} \tag{15}$$

with

$$\tilde{\mathsf{g}}_i = \mathsf{g}_i \bigoplus_{j\in\mathcal{P}_1(\tilde{\mathsf{t}}_i)\setminus i} \mathsf{g}_j \tag{16}$$

where $\tilde{\mathsf{t}}_i$ is the i-th row of T.

III. Adding Information Bits to RM Information Set by Sustaining the Same Minimum Distance

In this section, we present how to obtain triples of polarization matrix rows to keep the same minimum distance as the underlying RM code and state the size of information length increment for some given parameters. Let $\mathcal{T} \subseteq \mathcal{N}$ be any subset of row indices of the polarization matrix G. Then, by $\mathsf{g}_{\mathcal{T}}$, we denote

$$\mathsf{g}_{\mathcal{T}} = \bigoplus_{t\in\mathcal{T}} \mathsf{g}_t \tag{17}$$

A. Preliminary Theorems

Let us first state a corollary of [1, Theorem 2] that will be exploited later on in the paper.

Corollary 1. *Let $\Pi : \mathbb{F}_2^n \mapsto \mathbb{F}_2^n$ be a permutation on binary representations of $j \in \mathcal{N}$ and $\tilde{\mathcal{T}}$ be the index set obtained by applying permutation Π to the binary representations of elements of $\mathcal{T}$: $\mathsf{b}_{\tilde{j}} = \Pi(\mathsf{b}_j)$, $j \in \mathcal{T}$ and $\tilde{j} \in \tilde{\mathcal{T}}$. Then,*

$$i_1(\mathsf{g}_{\mathcal{T}}) = i_1(\mathsf{g}_{\tilde{\mathcal{T}}}) \tag{18}$$

Proof: The number of common $1-$bits will not change with permutation for any subset $\mathcal{T}^w \subset \mathcal{T}$, i.e.,

$$\begin{aligned} i_1(\bar{\cap}_{j\in\mathcal{T}^w}\mathsf{b}_j) &= i_1(\bar{\cap}_{j\in\mathcal{T}^w}\Pi(\mathsf{b}_j)) = i_1(\bar{\cap}_{\tilde{j}\in\Pi_{\mathcal{T}^w}}\mathsf{b}_{\tilde{j}}) \\ &= i_1(\bar{\cap}_{\tilde{j}\in\tilde{\mathcal{T}}^w}\mathsf{b}_{\tilde{j}}) \end{aligned} \tag{19}$$

then, by [1, Theorem 2], the Hamming weight does not change. ∎

The following theorem is also used to obtain subsequent results of this paper. It basically states that for any given set of rows of the polarization matrix, the Hamming weight of the sum of all rows is lower bounded by the maximum Hamming weight of the sum of a subset of rows whose binary representations are zero at the corresponding binary indices.

Theorem 1. *For any given $\mathcal{T} \subseteq \mathcal{N}$ the Hamming weight of $\mathsf{g}_{\mathcal{T}}$ is lower bounded by*

$$i_1(\mathsf{g}_{\mathcal{T}}) \geq \max_{\ell\in[0,n)} i_1(\mathsf{g}_{\mathcal{T}_\ell^0}) \tag{20}$$

where $\mathcal{T}_\ell^0 := \{k : b_{k,\ell} = 0, k \in \mathcal{T}\}$.

Proof: For any $\mathsf{u}, \mathsf{v} \in \mathbb{F}_2^{1\times N}$, we have

$$\begin{aligned} i_1(\mathsf{u}\oplus\mathsf{v}) + i_1(\mathsf{v}) &= i_1(\mathsf{u}) + i_1(\mathsf{v}) - 2\cdot i_1(\mathsf{u}\bar{\cap}\mathsf{v}) + i_1(\mathsf{v}) \\ &= i_1(\mathsf{u}) + 2\cdot \underbrace{(i_1(\mathsf{v}) - i_1(\mathsf{u}\bar{\cap}\mathsf{v}))}_{\geq 0} \\ &\geq i_1(\mathsf{u}). \end{aligned} \tag{21}$$

Then, note that for any $j \in \mathcal{N}$

$$i_1(\mathsf{g}_j) = \begin{cases} i_1(\mathsf{g}_j^\ell) & \text{if } b_{j,\ell} = 0 \\ 2\cdot i_1(\mathsf{g}_j^\ell) & \text{if } b_{j,\ell} = 1 \end{cases} \tag{22}$$

for any $\ell \in [0, n)$ due to (7) and (8). Therefore, for any $\ell \in [0, n)$ we can write

$$\begin{aligned} i_1(\mathsf{g}_{\mathcal{T}}) &= (\mathsf{g}_{\mathcal{T}}|\mathcal{M}_\ell^c) + (\mathsf{g}_{\mathcal{T}}|\mathcal{M}_\ell) \\ &\overset{(a)}{=} i_1(\bigoplus_{j\in\mathcal{T}} \mathsf{g}_j^\ell) + i_1(\bigoplus_{j\in\mathcal{T}} \mathsf{g}_j^\ell \mathbb{I}\{b_{j,\ell} = 1\}) \\ &= i_1(\bigoplus_{j\in\mathcal{T}} \mathsf{g}_j^\ell \mathbb{I}\{b_{j,\ell} = 0\} \bigoplus_{j\in\mathcal{T}} \mathsf{g}_j^\ell \mathbb{I}\{b_{j,\ell} = 1\}) \\ &\quad + i_1(\bigoplus_{j\in\mathcal{T}} \mathsf{g}_j^\ell \mathbb{I}\{b_{j,\ell} = 1\}) \\ &\overset{(b)}{\geq} i_1(\bigoplus_{j\in\mathcal{T}} \mathsf{g}_j^\ell \mathbb{I}\{b_{j,\ell} = 0\}) = i_1(\bigoplus_{j\in\mathcal{T}_\ell^0} \mathsf{g}_j^\ell) \overset{(c)}{=} i_1(\bigoplus_{j\in\mathcal{T}_\ell^0} \mathsf{g}_j) \end{aligned} \tag{23}$$

where (a) is due to (7) and (8), (b) is due to (21) and (c) is due to (22). ∎

Theorem 2. *Let $\mathbf{C}$ be a polar-like code with information set $\mathcal{A} = \bigcup_{p=\ell+1}^n \mathcal{N}_p$. and (i, j, k) be a triple such that $(i, j) \in \mathcal{N}_\ell, \ell \geq 2$, $k \in \mathcal{N}_2$ and $i_1(\mathsf{b}_i\bar{\cap}\mathsf{b}_j) = i_1(\mathsf{b}_i\bar{\cap}\mathsf{b}_k) = i_1(\mathsf{b}_j\bar{\cap}\mathsf{b}_k) = 0$. Moreover, let $\bar{\mathbf{C}}$ be another polar-like code that encodes an additional information bit by $\mathsf{g}_i \oplus \mathsf{g}_j \oplus \mathsf{g}_k$, i.e.,*

$$\bar{\mathbf{C}} := \{\mathbf{C}\} \cup \{\mathsf{c} : \mathsf{c} = \mathsf{g}_{\{i,j,k\}} \oplus \mathsf{g}_{\mathcal{T}}, \mathcal{T} \subseteq \mathcal{A}\}. \tag{24}$$

Then, the minimum distance of $\bar{\mathbf{C}}$ is the same as $\mathbf{C}$, i.e.,

$$\begin{aligned} d(\bar{\mathbf{C}}) &= \min\{d(\mathbf{C}), \min_{\mathcal{T}\subseteq\mathcal{A}} i_1(\mathsf{g}_{\{i,j,k\}} \oplus \mathsf{g}_{\mathcal{T}})\} \\ &= d(\mathbf{C}) = 2^{\ell+1} \end{aligned} \tag{25}$$

Proof: The proof is given in Appendices of [16]. ∎

B. Merging Three Rows with Common 1-bit Positions

The following theorem is a generalization of Theorem 2 and states the sufficient conditions on the rows of a triple with some common $1-$bit positions in their binary representations, to be merged together such that the minimum distance of the underlying RM code is preserved.

Theorem 3. *Let $\mathbf{C}$ be a polar-like code with information set $\mathcal{A} = \bigcup_{p=\ell+1}^n \mathcal{N}_p$. and (i, j, k) be a triple such that*

$\mathcal{P}_1(\mathsf{b}_i \bar{\cap} \mathsf{b}_j) = \mathcal{P}_1(\mathsf{b}_i \bar{\cap} \mathsf{b}_k) = \mathcal{P}_1(\mathsf{b}_j \bar{\cap} \mathsf{b}_k) \neq \emptyset$, $(i,j) \in \mathcal{N}_\ell$, $k \in \mathcal{N}_{i_1(\mathsf{b}_i \bar{\cap} \mathsf{b}_j)+2}$, $\ell \geq i_1(\mathsf{b}_k)$. *Let the code* $\bar{\mathbf{C}}$ *be:*

$$\bar{\mathbf{C}} := \{\mathbf{C}\} \cup \{\mathsf{c} : \mathsf{c} = \mathsf{g}_{\{i,j,k\}} \oplus \mathsf{g}_{\mathcal{T}}, \mathcal{T} \subseteq \mathcal{A}\} \tag{26}$$

Then,

$$d(\bar{\mathbf{C}}) = d(\mathbf{C}) = 2^{\ell+1} \tag{27}$$

Proof: Since

$$d(\bar{\mathbf{C}}) = \min\{d(\mathbf{C}), \min_{\mathcal{T} \subseteq \mathcal{A}} i_1(\mathsf{g}_{\{i,j,k\}} \oplus \mathsf{g}_{\mathcal{T}})\} \tag{28}$$

it is sufficient to prove the following statement

$$i_1(\mathsf{g}_{\{i,j,k\}} \oplus \mathsf{g}_{\mathcal{T}}) \geq 2^{\ell+1}, \ \ \forall \mathcal{T} \subseteq \mathcal{A}. \tag{29}$$

For any $\mathcal{T} \subseteq \mathcal{A}$, the index set can be divided into two subsets such that

$$\tilde{\mathcal{T}} := \{t : \mathcal{P}_1(\mathsf{b}_t) \cap \mathcal{P}_0(\mathsf{b}_i \bar{\cup} \mathsf{b}_j \bar{\cup} \mathsf{b}_k) \neq \emptyset, t \in \mathcal{T}\} \tag{30}$$

and $\hat{\mathcal{T}} = \mathcal{T} \setminus \tilde{\mathcal{T}}$. Then,

$$\begin{aligned}
&i_1(\mathsf{g}_{\{i,j,k\}} \oplus \mathsf{g}_{\hat{\mathcal{T}}} \oplus \mathsf{g}_{\tilde{\mathcal{T}}}) \\
&\overset{(a)}{\geq} \max_{p_0 \in \mathcal{P}_0(\mathsf{b}_i \bar{\cup} \mathsf{b}_j \bar{\cup} \mathsf{b}_k)} i_1(\mathsf{g}_{\{i,j,k\}} \oplus \mathsf{g}_{\hat{\mathcal{T}}} \bigoplus_{t \in \tilde{\mathcal{T}}} \mathsf{g}_t \mathbb{I}\{b_{t,p_0} = 0\}) \\
&\overset{(b)}{\geq} \max_{p_1 \in \mathcal{P}_0(\mathsf{b}_i \bar{\cup} \mathsf{b}_j \bar{\cup} \mathsf{b}_k) \setminus p_0} i_1(\mathsf{g}_{\{i,j,k\}} \oplus \mathsf{g}_{\hat{\mathcal{T}}} \bigoplus_{t \in \tilde{\mathcal{T}}} \mathsf{g}_t \mathbb{I}\{b_{t,p_0} = b_{t,p_1} = 0\}) \\
&\qquad \vdots \\
&\overset{(c)}{\geq} i_1(\mathsf{g}_{\{i,j,k\}} \oplus \mathsf{g}_{\hat{\mathcal{T}}} \bigoplus_{t \in \tilde{\mathcal{T}}} \mathsf{g}_t \mathbb{I}\{b_{t,p_0} = b_{t,p_1} = \cdots = b_{t,p_{n-2\ell-1}} = 0\}) \\
&\overset{(d)}{=} i_1(\mathsf{g}_{\{i,j,k\}} \oplus \mathsf{g}_{\hat{\mathcal{T}}})
\end{aligned} \tag{31}$$

where $\{p_0, p_1, \cdots, p_{n-2\ell-1}\} = \mathcal{P}_0(\mathsf{b}_i \bar{\cup} \mathsf{b}_j \bar{\cup} \mathsf{b}_k)$, where (a), (b) and (c) follow from the repeated application of Theorem 1, and (d) comes from (30), which implies that there is no $t \in \tilde{\mathcal{T}}$ such that $\mathcal{P}_1(\mathsf{b}_t) \cap \{p_0, p_1, \cdots, p_{n-2\ell-1}\} = \emptyset$. This means that the Hamming weight of $\mathsf{g}_{\{i,j,k,\mathcal{T}\}}$ is lower bounded by the Hamming weight of $\mathsf{g}_{\{i,j,k,\hat{\mathcal{T}}\}}$. Therefore, in the following, we will proceed the proof for $\hat{\mathcal{T}}$.

Note that by assumption of the theorem $\mathcal{P}_1(\mathsf{b}_i \bar{\cap} \mathsf{b}_j) = \mathcal{P}_1(\mathsf{b}_i \bar{\cap} \mathsf{b}_j \bar{\cap} \mathsf{b}_k)$. Now, assume that $\mathcal{W} = \mathcal{P}_1(\mathsf{b}_i \bar{\cap} \mathsf{b}_j \bar{\cap} \mathsf{b}_k)$, which is the index set of common one bit positions of b_i, b_j and b_k. Then, by partitioning the row indices of the polar encoding matrix, we obtain the following expression

$$\begin{aligned}
&i_1(\mathsf{g}_{\{i,j,k\}} \oplus \mathsf{g}_{\hat{\mathcal{T}}}) \\
&= i_1(\mathsf{g}_{\{i,j,k\}} \oplus \mathsf{g}_{\hat{\mathcal{T}}} | \mathcal{M}^c_{\mathcal{W}(|\mathcal{W}|-1)} \cap \mathcal{M}^c_{\mathcal{W}(|\mathcal{W}|-2)} \cap \cdots \cap \mathcal{M}^c_{\mathcal{W}(0)}) \\
&+ i_1(\mathsf{g}_{\{i,j,k\}} \oplus \mathsf{g}_{\hat{\mathcal{T}}} | \mathcal{M}^c_{\mathcal{W}(|\mathcal{W}|-1)} \cap \mathcal{M}^c_{\mathcal{W}(|\mathcal{W}|-2)} \cap \cdots \cap \mathcal{M}_{\mathcal{W}(0)}) \\
&\qquad \vdots \\
&+ i_1(\mathsf{g}_{\{i,j,k\}} \oplus \mathsf{g}_{\hat{\mathcal{T}}} | \mathcal{M}_{\mathcal{W}(|\mathcal{W}|-1)} \cap \mathcal{M}_{\mathcal{W}(|\mathcal{W}|-2)} \cap \cdots \cap \mathcal{M}_{\mathcal{W}(0)}) \\
&\overset{(a)}{=} i_1(\mathsf{g}^{\mathcal{W}}_{\{i,j,k\}} \bigoplus_{t \in \hat{\mathcal{T}}} \mathsf{g}^{\mathcal{W}}_t) + i_1(\mathsf{g}^{\mathcal{W}}_{\{i,j,k\}} \bigoplus_{t \in \hat{\mathcal{T}}} \mathsf{g}^{\mathcal{W}}_t \mathbb{I}\{b_{t,\mathcal{W}(0)} = 1\}) \\
&\qquad \vdots \\
&+ i_1(\mathsf{g}^{\mathcal{W}}_{\{i,j,k\}} \bigoplus_{t \in \hat{\mathcal{T}}} \mathsf{g}^{\mathcal{W}}_t \mathbb{I}\{b_{t,\mathcal{W}(|\mathcal{W}|-1)} = b_{t,\mathcal{W}(|\mathcal{W}|-2)} = \cdots \\
&\qquad\qquad \cdots = b_{t,\mathcal{W}(0)} = 1\})
\end{aligned} \tag{32}$$

where (a) is due to (9) and (10). Since $\mathcal{P}_1(\mathsf{b}_i) \setminus \mathcal{P}_1(\mathsf{b}_i \bar{\cap} \mathsf{b}_j)$, $\mathcal{P}_1(\mathsf{b}_j) \setminus \mathcal{P}_1(\mathsf{b}_i \bar{\cap} \mathsf{b}_j)$ and $\mathcal{P}_1(\mathsf{b}_k) \setminus \mathcal{P}_1(\mathsf{b}_i \bar{\cap} \mathsf{b}_k)$ comply with the conditions of Theorem 2, each term of the partition is lower bounded by $2^{\ell - |\mathcal{W}|+1}$. Then,

$$\begin{aligned}
i_1(\mathsf{g}_{\{i,j,k\}} \oplus \mathsf{g}_{\hat{\mathcal{T}}}) &\geq 2^{i_1(\mathsf{b}_i \bar{\cap} \mathsf{b}_j)} \cdot 2^{\ell - i_1(\mathsf{b}_i \bar{\cap} \mathsf{b}_j)+1} \\
&= 2^{\ell+1}
\end{aligned} \tag{33}$$

where $|\mathcal{W}| = i_1(\mathsf{b}_i \bar{\cap} \mathsf{b}_j)$. ∎

In the following, we state the sufficient conditions to increase the information length by multiple bits for a fix codeword length. Thanks to the symmetry imposed by Corollary 1, we apply a permutation Π to any given row triple satisfying the conditions of Theorem 3 to have the following form

$$\mathcal{P}_1(\mathsf{b}_i \bar{\cup} \mathsf{b}_j \bar{\cup} \mathsf{b}_k) \succ \mathcal{P}_0(\mathsf{b}_i \bar{\cup} \mathsf{b}_j \bar{\cup} \mathsf{b}_k) \tag{34}$$

and

$$\mathcal{P}_1(\mathsf{b}_i \bar{\cap} \mathsf{b}_j \bar{\cap} \mathsf{b}_k) \succ \mathcal{P}_1(\mathsf{b}_i \bar{\cup} \mathsf{b}_j \bar{\cup} \mathsf{b}_k) \setminus \mathcal{P}_1(\mathsf{b}_i \bar{\cap} \mathsf{b}_j \bar{\cap} \mathsf{b}_k) \tag{35}$$

and

$$\begin{aligned}
&\mathcal{P}_1(\mathsf{b}_k) \setminus \mathcal{P}_1(\mathsf{b}_i \bar{\cap} \mathsf{b}_j) \succ \mathcal{P}_1(\mathsf{b}_j) \setminus \mathcal{P}_1(\mathsf{b}_i \bar{\cap} \mathsf{b}_j), \\
&\mathcal{P}_1(\mathsf{b}_k) \setminus \mathcal{P}_1(\mathsf{b}_i \bar{\cap} \mathsf{b}_j) \succ \mathcal{P}_1(\mathsf{b}_i) \setminus \mathcal{P}_1(\mathsf{b}_i \bar{\cap} \mathsf{b}_j)
\end{aligned} \tag{36}$$

and

$$\begin{aligned}
&\mathcal{P}_1(\mathsf{b}_i) \setminus \mathcal{P}_1(\mathsf{b}_i \bar{\cap} \mathsf{b}_j) \not\succ \mathcal{P}_1(\mathsf{b}_j) \setminus \mathcal{P}_1(\mathsf{b}_i \bar{\cap} \mathsf{b}_j), \\
&\mathcal{P}_1(\mathsf{b}_j) \setminus \mathcal{P}_1(\mathsf{b}_i \bar{\cap} \mathsf{b}_j) \not\succ \mathcal{P}_1(\mathsf{b}_i) \setminus \mathcal{P}_1(\mathsf{b}_i \bar{\cap} \mathsf{b}_j)
\end{aligned} \tag{37}$$

Moreover, let Π^θ_p be a left-circular shift permutation on the index set of binary representation of $p \in \mathcal{N}$, with $\theta \in [0, t_0 + t_1]$, $t_0 = i_0(\mathsf{b}_i \bar{\cup} \mathsf{b}_j \bar{\cup} \mathsf{b}_k)$ and $t_1 = i_1(\mathsf{b}_i \bar{\cap} \mathsf{b}_j \bar{\cap} \mathsf{b}_k)$. We have

$$b_{\Pi^\theta_p, v} = b_{p, v-\theta+n \pmod n} \tag{38}$$

The following theorem is the main result of this paper.

Theorem 4. *Let* $\mathbf{C}$ *be a polar-like code with information set* $\mathcal{A} = \bigcup_{p=\ell+1}^{n} \mathcal{N}_p$. *Let* (i,j,k) *be a triple satisfying the conditions of Theorem 3 and* (34), (35), (36), (37). *Let* $\bar{\mathbf{C}}$ *be a code obtained by encoding each of the extra* $m \leq t_0 + t_1 + 1$ *information bits with a merged row triple* $\mathsf{g}_{\{\Pi^\theta_i, \Pi^\theta_j, \Pi^\theta_k\}}$. *Then,*

$$d(\bar{\mathbf{C}}) = d(\mathbf{C}) \tag{39}$$

Proof: The proof is given in Appendices of [16]. ∎

The following section explains how Theorem 4 is used in order to increase the information length of a polar-like code with RM information set by preserving the minimum distance.

IV. Code Construction

Let us consider a triple (i, j, k) that satisfies the conditions of Theorem 3, (34), (35), (36) and (37). For any $m \in [1, t_0 + t_1 + 1]$,

- $m-1$ triples, $\{(i_0, j_0, k_0), \cdots, (i_{m-2}, j_{m-2}, k_{m-2})\}$, are obtained from the left-circular shift of (i, j, k).
- For all triples, the permutation of their binary representations such that the smallest element among all the triples is maximized, is searched. This prevents from adding more badly polarized bit sub-channels to the information set. Indeed, with Corollary 1, the code constructed by any permutation of $m-$triples has the same distance spectrum since the underlying information set is chosen by RM rule.
- Algorithm 1, proposed in [1] and given below for the sake of completeness, is applied to obtain the pairs (t, v), where $t \in \mathcal{N}_{\ell+1}$, $v \in \mathcal{N}_\ell$, $v > t$, $\ell = i_1(\mathsf{b}_i)$, to decrease the number of minimum weight codewords.

Remark 1. *Even though we have verified experimentally that the application of the third step does not decrease the minimum distance, an explicit proof remains to be done.*

The pre-transformation matrix is constructed by adding the smallest index of each of $m-$triple to the information set and the other two indices are considered as dynamic frozen bits. For any obtained pair (t, v), v is considered as a dynamic frozen bit. The pre-transformation matrix T, is such that

$$T_{a,a} = T_{a,b} = T_{a,c} = T_{t,v} = 1 \tag{40}$$

where $a \in \mathcal{N}$ is the minimum of the triples, and $v \in \mathcal{N}_\ell$ is the associated index to any $t \in \mathcal{N}_{\ell+1}$ by the application of Algorithm 1 for a given $\ell^* = i_1(\mathsf{b}_t \bar{\cap} \mathsf{b}_v) \leq \ell - 1$.

Algorithm 1: Row Merging Pairs for ℓ^*.

Set $\mathcal{K} \leftarrow \mathcal{N}_\ell$;
for $w = 1$ *to* $|\mathcal{N}_{\ell+1}|$ **do**
  Set $t \leftarrow \mathcal{N}_{\ell+1}(w)$;
  Set $\mathcal{K}_t \leftarrow \{z : i_1(b_t \bar{\cap} b_z) = \ell^*, z > t, z \in \mathcal{K}\}$;
  if $\mathcal{K}_t \neq \emptyset$ **then**
    Set $(t, v) \leftarrow (t, \mathcal{K}_t(1))$;
    Set $\mathcal{K} \leftarrow \mathcal{K} \setminus v$;
  end
end

Decoding is performed by SCL decoding with dynamic frozen functions, i.e. instead of setting a frozen bit to zero as in standard SCL decoding, the value of any dynamic frozen bit is set through its corresponding dynamic frozen function.

V. Simulation Results

We numerically compare in Figure 1 our proposed design (PD) with PAC codes under SCL decoding with list size 256 and the saddle-point approximation of the MC (SMC) bound [17] for BIAWGN channel. Our construction for the code $(128, 66)$ is obtained by first adding two extra bits to the $(128, 64)$ polar-like code with RM information set and then by applying Algorithm 1 to obtain the (t, v) pairs such that $i_1(\mathsf{b}_t \bar{\cap} \mathsf{b}_v) = 1$. Similarly, the code $(128, 100)$ is obtained by first adding one extra bit to the polar-like code $(128, 99)$ with RM information set and then by applying Algorithm 1 to obtain the (t, v) pairs such that $i_1(\mathsf{b}_t \bar{\cap} \mathsf{b}_v) = 0$.

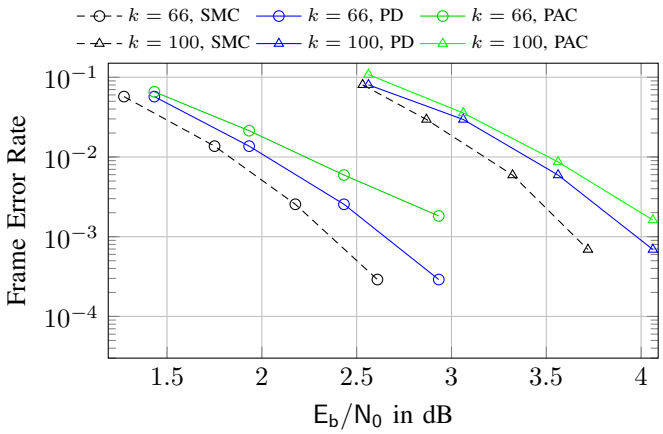

Figure 1. FER of our proposed scheme (PD), compared to SMC and PAC codes, $N = 128$, $k \in \{66; 100\}$.

For PAC codes, the additional information indices are chosen as the most reliable bit subchannel indices from the set $\mathcal{N}_\ell$, which are the highest indices due to partial ordering [18]. We optimize the polynomial of the convolutional code with memory length 7 to minimize the number of minimum weight codewords. The algorithm [19] with a large list size, i.e., $5 \cdot 10^4$, has been implemented and we choose the polynomial that leads to the minimal number of second minimum weight codewords since the number of minimum weight codewords does not change for a few increment of the information length. As a result, we have implemented PAC codes with the polynomials $[1\ 1\ 0\ 0\ 1\ 1\ 0\ 0\ 1]$ and $[1\ 1\ 1\ 1\ 0\ 1\ 1\ 1]$ for information lengths $k = 66$ and $k = 100$, respectively.

It is seen that our design performs better than PAC codes for the entire range of SNR, since, at short blocklengths, the minimum distance plays an important role in the SCL decoding with large list sizes. Moreover, our design performs close to the MC bound especially at high E_b/N_0.

VI. Concluding Remarks

In this work, we proposed a method to increase the information length of a polar-like code while keeping the same minimum distance with the underlying RM code. Moreover, we proposed an heuristic to reduce the number of minimum weight codewords that allows to design codes that perform closer to the MC bound than PAC codes with the same system parameters. We believe that this work allows a better understanding of the polarization matrix properties, which may lead to more efficient code design, particularly for short blocklengths. The extension of this work to moderate blocklengths is under investigation.

REFERENCES

[1] S. Gelincik, P. Mary, J.-Y. Baudais, and A. Savard, "Achieving PAC code performance without extra computational complexity," in *IEEE Internat. Conf. on Communications*, Seoul, Korea, May 2022, pp. 1–6.

[2] E. Arikan, "Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 7, pp. 3051–3073, 2009.

[3] ETSI, *Multiplexing and Channel Coding (3GPP TS 38.212 V15.3.0 Release 15)*, Oct. 2018.

[4] V. Bioglio, C. Condo, and I. Land, "Design of polar codes in 5G new radio," *IEEE Commun. Surveys Tutorials*, vol. 23, no. 1, pp. 29–40, 2021.

[5] P. Trifonov and V. Miloslavskaya, "Polar subcodes," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 2, pp. 254–266, 2016.

[6] P. Yuan, T. Prinz, G. Boecherer, O. Iscan, R. Boehnke, and W. Xu, "Polar code construction for list decoding," in *International ITG Conference on Systems, Communications and Coding*, Germany, Feb. 2019, pp. 1–6.

[7] V. Miloslavskaya and B. Vucetic, "Design of short polar codes for scl decoding," *IEEE Transactions on Communications*, vol. 68, no. 11, pp. 6657–6668, 2020.

[8] I. Tal and A. Vardy, "List decoding of polar codes," *IEEE Trans. Inf. Theory*, vol. 61, no. 5, pp. 2213–2226, 2015.

[9] Q. Zhang, A. Liu, X. Pan, and K. Pan, "CRC code design for list decoding of polar codes," *IEEE Com. Lett.*, vol. 21, no. 6, pp. 1229–1232, 2017.

[10] J. Piao, K. Niu, J. Dai, and C. Dong, "Approaching the normal approximation of the finite blocklength capacity within 0.025 dB by short polar codes," *IEEE Wireless Commun. Lett.*, vol. 9, no. 7, pp. 1089–1092, 2020.

[11] E. Arıkan, "From sequential decoding to channel polarization and back again," arXiv, Sep. 2019. [Online]. Available: https://arxiv.org/abs/1908.09594

[12] B. Li, H. Zhang, and J. Gu, "On pre-transformed polar codes," arXiv, Dec. 2019. [Online]. Available: https://arxiv.org/abs/1912.06359

[13] A. Elkelesh, M. Ebada, S. Cammerer, and S. ten Brink, "Decoder-tailored polar code design using the genetic algorithm," *IEEE Trans. Commun.*, vol. 67, no. 7, pp. 4521–4534, 2019.

[14] T. Tonnellier and W. J. Gross, "On systematic polarization-adjusted convolutional (PAC) codes," *IEEE Commun. Lett.*, vol. 25, no. 7, pp. 2128–2132, 2021.

[15] S. H. Hassani, R. Mori, T. Tanaka, and R. L. Urbanke, "Rate-dependent analysis of the asymptotic behavior of channel polarization," *IEEE Transactions on Information Theory*, vol. 59, no. 4, pp. 2267–2276, 2013.

[16] S. Gelincik, P. Mary, A. Savard, and J. Y. Baudais, "A pre-transformation method to increase the minimum distance of polar-like codes," arXiv, Apr. 2022. [Online]. Available: https://arxiv.org/abs/2202.04366

[17] D. Anada, J. M. Gorce, P. Mary, and S. M. Perlaza, "An upper bound on the error induced by saddlepoint approximations-applications to information theory," *Entropy*, vol. 22, no. 6: 690, Jun. 2020.

[18] C. Schürch, "A partial order for the synthesized channels of a polar code," in *IEEE International Symposium on Information Theory*, Barcelona, Spain, Jul. 2016, pp. 220–224.

[19] B. Li, H. Shen, and D. Tse, "An adaptive successive cancellation list decoder for polar codes with cyclic redundancy check," *IEEE Commun. Lett.*, vol. 16, no. 12, pp. 2044–2047, 2012.