



HAL
open science

A Pre-Transformation Method to Increase the Minimum Distance of Polar-Like Codes

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

► **To cite this version:**

Samet Gelincik, Philippe Mary, Anne Savard, Jean-Yves Baudais. A Pre-Transformation Method to Increase the Minimum Distance of Polar-Like Codes. 2022. hal-03557597v2

HAL Id: hal-03557597

<https://hal.science/hal-03557597v2>

Preprint submitted on 13 Apr 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A Pre-Transformation Method to Increase the Minimum Distance of Polar-Like Codes

Samet Gelincik, *Member, IEEE*, Philippe Mary, *Member, IEEE*, Anne Savard,
Member, IEEE and Jean-Yves Baudais

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. The proofs give a constructive method to choose the row triples to be merged together to increase the information length of the code and they follow from partitioning the row indices of the polar encoding matrix with respect to the recursive structure imposed by the binary representation of row indices. Moreover, our findings are combined with the method presented in [2] to further reduce the number of minimum weight codewords. Numerical results

S. Gelincik, P. Mary and J.-Y. Baudais are with Univ Rennes, INSA Rennes, CNRS, IETR-UMR 6164, F-35000 Rennes, France

A. Savard is with IMT Nord Europe, Institut Mines Télécom, Centre for Digital Systems, F-59653 Villeneuve d'Ascq, France.

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. Part of the content of this paper has been submitted to ISIT 2022 [1].

show that the designed codes perform close to the meta-converse bound at short blocklengths and better than the polarization-adjusted-convolutional polar codes with the same parameters.

Index Terms

Polar codes, Reed Muller codes, minimum distance, finite block length.

I. INTRODUCTION

New usages and services of 5G and beyond wireless systems, such as machine type communication or ultra-reliable low latency communications are pushing the limits of channel coding by requiring efficient error correcting codes at short to moderate block lengths. Indeed, These use-cases involve communicating objects that either occasionally transmit short packets at low power transmission to increase the device lifetime or because to meet stringent latency constraint [3].

Polar codes, the first provably asymptotically capacity achieving error correcting codes over binary input memoryless channels [4] with explicit construction, are currently used over the control channels of 5G networks [5]. They also are envisioned for ultra-reliable low-latency communications and massive machine-type communications [6] thanks to their low complexity successive cancellation based decoder. Unfortunately, standard polar codes do not show outstanding performance at short-to-moderate block lengths due to their poor minimum distance and a non-complete polarization [7]. Several methods, such as enhanced-Bose–Chaudhuri–Hocquenghem subcodes [8] and low-weight-bit polar codes [9], have hence been proposed to improve their distance spectrum. 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 [10] and the obtained performances were further enhanced by optimizing the CRC polynomial to improve the minimum distance of the obtained codes [11], [12]. The later was considered as the best code design in terms of Frame Error Rate (FER) up to the introduction of polarized adjusted convolutional (PAC) polar codes in [13].

PAC polar codes [13], by choosing the information set of the polar codes according to the Reed-Muller (RM) rule, i.e. the rows of the polarization matrix with the highest Hamming weights, perform very close to the second-order rate approximation of the binary-input additive white Gaussian noise in the short block length regime. It is a special case of the convolutional pre-transformation with an upper-triangular matrix, which has been proven not to reduce the minimum distance of underlying RM code while reducing the number of minimum weight codewords if properly designed [14]. Since polar codes are tailored for a given channel, authors in [15] proposed a genetic algorithm to obtain the frozen set that minimizes the bit or block error rate of the code over additive white Gaussian noise channel and Rayleigh channel. The obtained polar code does not require CRC and the proposed method either outperform CRC-aided polar codes occasionally or performs equally well in terms of bit and block error rate. Since the construction proposed in [15] is signal to noise ratio dependent, authors in [16] proposed a genetic algorithm, extended from [15], that relies on the distance spectrum of the code rather than an error rate criterion. This technique allows to achieve the normal approximation of the second-order rate in Gaussian noise channel. However, genetic algorithms are rather complex and strongly depend on the population size and its initial state. Contrarily, in our work, we give an explicit method to increase the minimum distance while adding additional information bits on the underlying RM code.

In [2], we proposed a method to decrease the number of low weight codewords compared to RM and PAC codes in the short block length regime. The approach relies on encoding some information bits by the sum of two or three rows of the polar encoding matrix. The pairs and triplets of the merged rows are determined via the connection between the binary representation of the selected row indices and their common 1 bit-positions. Notably, our designed codes achieve the same performance as PAC polar codes of the same parameters without extra computational complexity.

In this paper, we extend our previously proposed method [2] and state sufficient conditions

to increase the information length of some polar-like codes, i.e. increasing the rate for a fixed given codeword length, where the information set is built according to the RM rule. Moreover, we explicitly give the corresponding pre-transformation matrix to sustain the same minimum distance as the RM code. The analysis is conducted by extending the method introduced in [17] that partitions the row indices of the encoding matrix according to the indices of their binary representation. Numerical simulations show that our proposed code design outperforms the PAC polar codes with the same parameters in terms of FER, and performs close to the meta-converse (MC) bound thanks to the decrease in the number of minimum weight codewords achieved by our previously proposed algorithm in [2, Algorithm 1].

The rest of the paper is organized as follows. Section II introduces the notations and the important definitions that are used in the proofs of our findings. Section III presents our main contribution with the statement of three theorems. Theorems 3 and 4 state the conditions for which the row merging increases the Hamming weight of the resulting row for the case where there is no common 1-bit position among the binary representations of the row indices and when there are such common positions, respectively. Theorem 5 is the main result of the paper and claims that it is possible to increase the information length without decreasing the minimum distance of the code. In Section IV, we present our proposed polar-like code design which increases the number of information bits while exhibiting a high minimum distance. Section V compares the performance of our proposed design with PAC polar codes and with the meta-converse bound. Finally, conclusions are drawn in Section VI.

II. PRELIMINARIES

A. Notations

The positions of elements in a vector of length N is indexed from 0 to $N - 1$. Any vector of length N is considered as a row vector and is denoted by \mathbf{x} or \mathbf{x}^{N-1} . The j^{th} entry of the vector \mathbf{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 $\mathbf{x} \in \mathbb{F}_2^{1 \times N}$ and index set $\mathcal{A} \subset \mathcal{N}$, $\mathbf{x}_{\mathcal{A}}$ denotes the vector consisting of the elements of \mathbf{x} at the positions indexed by \mathcal{A} . The matrices are denoted by uppercase sans serif font, e.g., \mathbf{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 $0 \leq j < 2^n$, its n -bit binary representation is denoted by the vector \mathbf{b}_j^n , or \mathbf{b}_j if it is clear enough from the context. The ℓ -th bit position of \mathbf{b}_j is denoted by $b_{j,\ell}$, $0 \leq \ell < n$ 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 $(j_1, j_2) \in [0, 2^n)^2$:

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

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

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

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 $\mathbf{G} = \mathbf{G}_2^{\otimes n}$ where

$$\mathbf{G}_2 := \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \quad (3)$$

is the corresponding kernel matrix and \otimes is Kronecker product. The j th row \mathbf{g}_j of \mathbf{G} can be represented by

$$\mathbf{g}_j = \hat{\mathbf{g}}_{b_{j,n-1}} \otimes \hat{\mathbf{g}}_{b_{j,n-2}} \otimes \cdots \otimes \hat{\mathbf{g}}_{b_{j,0}} \quad (4)$$

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

$$r_{j,\ell} = \begin{cases} 0^{2^{\ell-1}} & \text{if } b_{j,\ell} = 0 \\ [r_{j,0}r_{j,1} \cdots r_{j,\ell-1}] & \text{if } b_{j,\ell} = 1 \end{cases} \quad (5)$$

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

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

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

Definition 1. *The projection of a row \mathbf{g}_j of the polar encoding matrix onto indices of \mathcal{M}_ℓ^c is denoted by \mathbf{g}_j^ℓ and given as*

$$\mathbf{g}_j^\ell := \hat{\mathbf{g}}_{b_{j,n-1}} \otimes \cdots \otimes \hat{\mathbf{g}}_{b_{j,\ell+1}} \otimes \hat{\mathbf{g}}_{b_{j,\ell-1}} \cdots \otimes \hat{\mathbf{g}}_{b_{j,0}} \quad (7)$$

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

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

The following definition is the generalization of Definition 1.

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

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

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 \mathbf{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

$$\mathbf{g}_j|_{\cap_{\ell \in \mathcal{B}_0} \mathcal{M}_\ell \cap_{\hat{\ell} \in \mathcal{B}/\mathcal{B}_0} \mathcal{M}_{\hat{\ell}}^c} = \begin{cases} 0^{|\mathcal{B}|^{-1}} & \text{if } \bar{\cap}_{\ell \in \mathcal{B}_0} b_{j,\ell} = 0 \\ \mathbf{g}_j^{\mathcal{B}} & \text{if } \bar{\cap}_{\ell \in \mathcal{B}_0} b_{j,\ell} = 1 \end{cases} \quad (10)$$

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

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

where \mathcal{F} is the index set of the frozen bit positions, and $\mathcal{A} = \mathcal{N} \setminus \mathcal{F}$ is the information set. For classical polar codes under SC decoding, the set \mathcal{A} is the set of the most reliable bit sub-channels [4]. However in this paper, we allow to choose the information set differently. From this perspective, a RM(n, r) code of degree r can be seen as a polar-like code of information set

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

In [18], the minimum distance of a polar-like code is given by

$$d(\mathbf{C}) = \min_{i \in \mathcal{A}} i_1(\mathbf{g}_i) \stackrel{(a)}{=} 2^{\min_{i \in \mathcal{A}} i_1(\mathbf{b}_i)} \quad (13)$$

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

The pre-transformed polar-like codes [14] is obtained through a pre-transformation matrix $\mathsf{T} \in \mathbb{F}_2^{N \times N}$

$$\mathsf{C}_P = \{\mathsf{c} = \mathsf{u}\mathsf{T}\mathsf{G} : \mathsf{u} \in \mathbb{F}_2^n, \mathsf{u}_{\mathcal{F}} = \mathbf{0}\} \quad (14)$$

where T is an upper triangular matrix with $T_{i,i} = 1, i \in \mathcal{N}$ 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 $|\{i : T_{i,j}, i \in \mathcal{N}\}| \in \{1, 2\} \forall j \in \mathcal{F}_d$, 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

$$\mathsf{c} = \mathsf{u}\mathsf{T}\mathsf{G} = \mathsf{u}\tilde{\mathsf{G}} \quad (15)$$

with

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

where $\tilde{\mathbf{t}}_i$ 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 and $i \in \mathcal{N} \setminus \mathcal{T}$. Then, by $\mathbf{g}_{\mathcal{T}}$ and $\mathbf{g}_{\{i, \mathcal{T}\}}$, we denote

$$\mathbf{g}_{\mathcal{T}} = \bigoplus_{t \in \mathcal{T}} \mathbf{g}_t \quad \text{and} \quad \mathbf{g}_{\{i, \mathcal{T}\}} = \mathbf{g}_i \oplus \mathbf{g}_{\mathcal{T}} \quad (17)$$

A. Preliminary Theorems

For the sake of completeness, we first state Theorem 2 of [2] and give a corollary that will be exploited later on in this paper.

Theorem 1. [2, Theorem 2] Let $\mathcal{T} \subseteq \mathcal{N}$ be any subset of row indices of polar-like code generator matrix \mathbf{G}_N . Then, the Hamming weight of the sum of the rows \mathbf{g}_j , $j \in \mathcal{T}$ is given by

$$i_1(\mathbf{g}_{\mathcal{T}}) = \sum_{w=1}^{|\mathcal{T}|} (-2)^{w-1} \sum_{\mathcal{T}^w \subset \mathcal{T}} 2^{i_1(\bar{\cap}_{j \in \mathcal{T}^w} \mathbf{b}_j)} \quad (18)$$

where \mathcal{T}^w is any subset of \mathcal{T} with w elements.

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} : $\mathbf{b}_{\tilde{j}} = \Pi(\mathbf{b}_j)$, $j \in \mathcal{T}$ and $\tilde{j} \in \tilde{\mathcal{T}}$. Then,

$$i_1(\mathbf{g}_{\mathcal{T}}) = i_1(\mathbf{g}_{\tilde{\mathcal{T}}}) \quad (19)$$

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

$$i_1(\bar{\cap}_{j \in \mathcal{T}^w} \mathbf{b}_j) = i_1(\bar{\cap}_{j \in \mathcal{T}^w} \Pi(\mathbf{b}_j)) = i_1(\bar{\cap}_{\tilde{j} \in \Pi \mathcal{T}^w} \mathbf{b}_{\tilde{j}}) = i_1(\bar{\cap}_{\tilde{j} \in \tilde{\mathcal{T}}^w} \mathbf{b}_{\tilde{j}}) \quad (20)$$

then, by (18), 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 2. For any given $\mathcal{T} \subseteq \mathcal{N}$ the Hamming weight of $\mathbf{g}_{\mathcal{T}}$ is lower bounded by

$$i_1(\mathbf{g}_{\mathcal{T}}) \geq \max_{\ell \in [0, n]} i_1(\mathbf{g}_{\mathcal{T}^{\ell}}) \quad (21)$$

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

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

$$i_1(\mathbf{u} \oplus \mathbf{v}) + i_1(\mathbf{v}) = i_1(\mathbf{u}) + i_1(\mathbf{v}) - 2 \cdot i_1(\mathbf{u} \bar{\cap} \mathbf{v}) + i_1(\mathbf{v})$$

$$= i_1(\mathbf{u}) + 2 \cdot \underbrace{(i_1(\mathbf{v}) - i_1(\mathbf{u} \bar{\cap} \mathbf{v}))}_{\geq 0} \geq i_1(\mathbf{u}). \quad (22)$$

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

$$i_1(\mathbf{g}_j) = \begin{cases} i_1(\mathbf{g}_j^\ell) & \text{if } b_{j,\ell} = 0 \\ 2 \cdot i_1(\mathbf{g}_j^\ell) & \text{if } b_{j,\ell} = 1 \end{cases} \quad (23)$$

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

$$\begin{aligned} i_1(\mathbf{g}_{\mathcal{T}}) &= (\mathbf{g}_{\mathcal{T}} | \mathcal{M}_\ell^c) + (\mathbf{g}_{\mathcal{T}} | \mathcal{M}_\ell) \stackrel{(a)}{=} i_1\left(\bigoplus_{j \in \mathcal{T}} \mathbf{g}_j^\ell\right) + i_1\left(\bigoplus_{j \in \mathcal{T}} \mathbf{g}_j^\ell \mathbb{I}\{b_{j,\ell} = 1\}\right) \\ &= i_1\left(\bigoplus_{j \in \mathcal{T}} \mathbf{g}_j^\ell \mathbb{I}\{b_{j,\ell} = 0\}\right) + i_1\left(\bigoplus_{j \in \mathcal{T}} \mathbf{g}_j^\ell \mathbb{I}\{b_{j,\ell} = 1\}\right) + i_1\left(\bigoplus_{j \in \mathcal{T}} \mathbf{g}_j^\ell \mathbb{I}\{b_{j,\ell} = 1\}\right) \\ &\stackrel{(b)}{\geq} i_1\left(\bigoplus_{j \in \mathcal{T}} \mathbf{g}_j^\ell \mathbb{I}\{b_{j,\ell} = 0\}\right) = i_1\left(\bigoplus_{j \in \mathcal{T}_\ell^0} \mathbf{g}_j^\ell\right) \stackrel{(c)}{=} i_1\left(\bigoplus_{j \in \mathcal{T}_\ell^0} \mathbf{g}_j\right) \end{aligned} \quad (24)$$

where (a) is due to (7) and (8), (b) is due to (22) and (c) is due to (23). ■

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 $(i, j) \in \mathcal{N}_\ell, \ell \geq 2, k \in \mathcal{N}_2$ and $i_1(\mathbf{b}_i \bar{\cap} \mathbf{b}_j) = i_1(\mathbf{b}_i \bar{\cap} \mathbf{b}_k) = i_1(\mathbf{b}_j \bar{\cap} \mathbf{b}_k) = 0$. Moreover, let $\bar{\mathbf{C}}$ be another polar-like code that encodes an additional information bit by $\mathbf{g}_i \oplus \mathbf{g}_j \oplus \mathbf{g}_k$, i.e.

$$\bar{\mathbf{C}} := \{\mathbf{C}\} \cup \{\mathbf{c} : \mathbf{c} = \mathbf{g}_{\{i,j,k\}} \oplus \mathbf{g}_{\mathcal{T}}, \mathcal{T} \subseteq \mathcal{A}\}. \quad (25)$$

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

$$\begin{aligned} d(\bar{\mathbf{C}}) &= \min\{d(\mathbf{C}), \min_{\mathcal{T} \subseteq \mathcal{A}} i_1(\mathbf{g}_{\{i,j,k\}} \oplus \mathbf{g}_{\mathcal{T}})\} \\ &= d(\mathbf{C}) = 2^{\ell+1} \end{aligned} \quad (26)$$

Proof: The proof is given in Appendix VII-B. ■

B. Merging Three Rows with Common 1-bit Positions

The following theorem is a generalization of Theorem 3 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 4. *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 $\mathcal{P}_1(\mathbf{b}_i \bar{\cap} \mathbf{b}_j) = \mathcal{P}_1(\mathbf{b}_i \bar{\cap} \mathbf{b}_k) = \mathcal{P}_1(\mathbf{b}_j \bar{\cap} \mathbf{b}_k) \neq \emptyset$, $(i, j) \in \mathcal{N}_\ell$, $k \in \mathcal{N}_{i_1(\mathbf{b}_i \bar{\cap} \mathbf{b}_j)+2}$, $\ell \geq i_1(\mathbf{b}_k)$. Let the code $\bar{\mathbf{C}}$ be:*

$$\bar{\mathbf{C}} := \{\mathbf{C}\} \cup \{\mathbf{c} : \mathbf{c} = \mathbf{g}_{\{i,j,k\}} \oplus \mathbf{g}_{\mathcal{T}}, \mathcal{T} \subseteq \mathcal{A}\} \quad (27)$$

Then,

$$d(\bar{\mathbf{C}}) = d(\mathbf{C}) = 2^{\ell+1} \quad (28)$$

Proof: Since

$$d(\bar{\mathbf{C}}) = \min\{d(\mathbf{C}), \min_{\mathcal{T} \subseteq \mathcal{A}} i_1(\mathbf{g}_{\{i,j,k\}} \oplus \mathbf{g}_{\mathcal{T}})\} \quad (29)$$

it is sufficient to prove the following statement

$$i_1(\mathbf{g}_{\{i,j,k\}} \oplus \mathbf{g}_{\mathcal{T}}) \geq 2^{\ell+1}, \quad \forall \mathcal{T} \subseteq \mathcal{A}. \quad (30)$$

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

$$\tilde{\mathcal{T}} := \{t : \mathcal{P}_1(\mathbf{b}_t) \cap \mathcal{P}_0(\mathbf{b}_i \bar{\cup} \mathbf{b}_j \bar{\cup} \mathbf{b}_k) \neq \emptyset, t \in \mathcal{T}\} \quad (31)$$

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

$$\begin{aligned} i_1(\mathbf{g}_{\{i,j,k\}} \oplus \mathbf{g}_{\hat{\mathcal{T}}} \oplus \mathbf{g}_{\tilde{\mathcal{T}}}) &\stackrel{(a)}{\geq} \max_{p_0 \in \mathcal{P}_0(\mathbf{b}_i \bar{\cup} \mathbf{b}_j \bar{\cup} \mathbf{b}_k)} i_1(\mathbf{g}_{\{i,j,k\}} \oplus \mathbf{g}_{\hat{\mathcal{T}}} \bigoplus_{t \in \tilde{\mathcal{T}}} \mathbf{g}_t \mathbb{I}\{b_{t,p_0} = 0\}) \\ &\stackrel{(b)}{\geq} \max_{p_1 \in \mathcal{P}_0(\mathbf{b}_i \bar{\cup} \mathbf{b}_j \bar{\cup} \mathbf{b}_k) \setminus p_0} i_1(\mathbf{g}_{\{i,j,k\}} \oplus \mathbf{g}_{\hat{\mathcal{T}}} \bigoplus_{t \in \tilde{\mathcal{T}}} \mathbf{g}_t \mathbb{I}\{b_{t,p_0} = b_{t,p_1} = 0\}) \\ &\vdots \end{aligned}$$

$$\begin{aligned}
&\stackrel{(c)}{\geq} i_1(\mathbf{g}_{\{i,j,k\}} \oplus \mathbf{g}_{\hat{\mathcal{T}}} \bigoplus_{t \in \tilde{\mathcal{T}}} \mathbf{g}_t \mathbb{I}\{b_{t,p_0} = b_{t,p_1} = \dots = b_{t,p_{n-2\ell-1}} = 0\}) \\
&\stackrel{(d)}{=} i_1(\mathbf{g}_{\{i,j,k\}} \oplus \mathbf{g}_{\hat{\mathcal{T}}}) \tag{32}
\end{aligned}$$

where $\{p_0, p_1, \dots, p_{n-2\ell-1}\} = \mathcal{P}_0(\mathbf{b}_i \bar{\cup} \mathbf{b}_j \bar{\cup} \mathbf{b}_k)$, where (a), (b) and (c) follow from the repeated application of Theorem 2, and (d) comes from (31), which implies that there is no $t \in \tilde{\mathcal{T}}$ such that $\mathcal{P}_1(\mathbf{b}_t) \cap \{p_0, p_1, \dots, p_{n-2\ell-1}\} = \emptyset$. This means that the Hamming weight of $\mathbf{g}_{\{i,j,k,\mathcal{T}\}}$ is lower bounded by the Hamming weight of $\mathbf{g}_{\{i,j,k,\hat{\mathcal{T}}\}}$. Therefore, in the following, we will proceed the proof for $\hat{\mathcal{T}}$.

Now, assume that $\mathcal{W} = \mathcal{P}_1(\mathbf{b}_i \bar{\cap} \mathbf{b}_j)$, i.e., $\mathcal{W} = \mathcal{P}_1(\mathbf{b}_i \bar{\cap} \mathbf{b}_j \bar{\cap} \mathbf{b}_k)$ as well, by assumption. Then, by partitioning the row indices of the polar encoding matrix respect to binary bit positions \mathcal{W} , we obtain the following expression

$$\begin{aligned}
i_1(\mathbf{g}_{\{i,j,k\}} \oplus \mathbf{g}_{\hat{\mathcal{T}}}) &= i_1(\mathbf{g}_{\{i,j,k\}} \oplus \mathbf{g}_{\hat{\mathcal{T}}} | \mathcal{M}_{\mathcal{W}(|\mathcal{W}|-1)}^c \cap \mathcal{M}_{\mathcal{W}(|\mathcal{W}|-2)}^c \cap \dots \cap \mathcal{M}_{\mathcal{W}(0)}^c) \\
&+ i_1(\mathbf{g}_{\{i,j,k\}} \oplus \mathbf{g}_{\hat{\mathcal{T}}} | \mathcal{M}_{\mathcal{W}(|\mathcal{W}|-1)}^c \cap \mathcal{M}_{\mathcal{W}(|\mathcal{W}|-2)}^c \cap \dots \cap \mathcal{M}_{\mathcal{W}(0)}) \\
&\quad \vdots \\
&+ i_1(\mathbf{g}_{\{i,j,k\}} \oplus \mathbf{g}_{\hat{\mathcal{T}}} | \mathcal{M}_{\mathcal{W}(|\mathcal{W}|-1)} \cap \mathcal{M}_{\mathcal{W}(|\mathcal{W}|-2)} \cap \dots \cap \mathcal{M}_{\mathcal{W}(0)}) \\
&\stackrel{(a)}{=} i_1(\mathbf{g}_{\{i,j,k\}}^{\mathcal{W}} \bigoplus_{t \in \hat{\mathcal{T}}} \mathbf{g}_t^{\mathcal{W}}) + i_1(\mathbf{g}_{\{i,j,k\}}^{\mathcal{W}} \bigoplus_{t \in \hat{\mathcal{T}}} \mathbf{g}_t^{\mathcal{W}} \mathbb{I}\{b_{t,\mathcal{W}(0)} = 1\}) \\
&\quad \vdots \\
&+ i_1(\mathbf{g}_{\{i,j,k\}}^{\mathcal{W}} \bigoplus_{t \in \hat{\mathcal{T}}} \mathbf{g}_t^{\mathcal{W}} \mathbb{I}\{b_{t,\mathcal{W}(|\mathcal{W}|-1)} = b_{t,\mathcal{W}(|\mathcal{W}|-2)} = \dots = b_{t,\mathcal{W}(0)} = 1\}) \tag{33}
\end{aligned}$$

where (a) is due to (9) and (10). Since $\mathcal{P}_1(\mathbf{b}_i) \setminus \mathcal{P}_1(\mathbf{b}_i \bar{\cap} \mathbf{b}_j)$, $\mathcal{P}_1(\mathbf{b}_j) \setminus \mathcal{P}_1(\mathbf{b}_i \bar{\cap} \mathbf{b}_j)$ and $\mathcal{P}_1(\mathbf{b}_k) \setminus \mathcal{P}_1(\mathbf{b}_i \bar{\cap} \mathbf{b}_k)$ comply with the conditions of Theorem 3, each term of the partition is lower bounded by $2^{\ell-|\mathcal{W}|+1}$. Then,

$$i_1(\mathbf{g}_{\{i,j,k\}} \oplus \mathbf{g}_{\hat{\mathcal{T}}}) \geq 2^{i_1(\mathbf{b}_i \bar{\cap} \mathbf{b}_j)} \cdot (2^{\ell-i_1(\mathbf{b}_i \bar{\cap} \mathbf{b}_j)+1}) = 2^{\ell+1} \tag{34}$$

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

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 4 to have the following form

$$\mathcal{P}_1(\mathbf{b}_i \bar{\cup} \mathbf{b}_j \bar{\cup} \mathbf{b}_k) \succ \mathcal{P}_0(\mathbf{b}_i \bar{\cup} \mathbf{b}_j \bar{\cup} \mathbf{b}_k) \quad (35)$$

and

$$\mathcal{P}_1(\mathbf{b}_i \bar{\cap} \mathbf{b}_j \bar{\cap} \mathbf{b}_k) \succ \mathcal{P}_1(\mathbf{b}_i \bar{\cup} \mathbf{b}_j \bar{\cup} \mathbf{b}_k) \setminus \mathcal{P}_1(\mathbf{b}_i \bar{\cap} \mathbf{b}_j \bar{\cap} \mathbf{b}_k) \quad (36)$$

and

$$\begin{aligned} \mathcal{P}_1(\mathbf{b}_k) \setminus \mathcal{P}_1(\mathbf{b}_i \bar{\cap} \mathbf{b}_j) &\succ \mathcal{P}_1(\mathbf{b}_j) \setminus \mathcal{P}_1(\mathbf{b}_i \bar{\cap} \mathbf{b}_j), \\ \mathcal{P}_1(\mathbf{b}_k) \setminus \mathcal{P}_1(\mathbf{b}_i \bar{\cap} \mathbf{b}_j) &\succ \mathcal{P}_1(\mathbf{b}_i) \setminus \mathcal{P}_1(\mathbf{b}_i \bar{\cap} \mathbf{b}_j) \end{aligned} \quad (37)$$

and

$$\begin{aligned} \mathcal{P}_1(\mathbf{b}_i) \setminus \mathcal{P}_1(\mathbf{b}_i \bar{\cap} \mathbf{b}_j) &\not\succeq \mathcal{P}_1(\mathbf{b}_j) \setminus \mathcal{P}_1(\mathbf{b}_i \bar{\cap} \mathbf{b}_j), \\ \mathcal{P}_1(\mathbf{b}_j) \setminus \mathcal{P}_1(\mathbf{b}_i \bar{\cap} \mathbf{b}_j) &\not\succeq \mathcal{P}_1(\mathbf{b}_i) \setminus \mathcal{P}_1(\mathbf{b}_i \bar{\cap} \mathbf{b}_j) \end{aligned} \quad (38)$$

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

$$b_{\Pi_p^\theta, v} = b_{p, v - \theta + n \pmod{n}} \quad (39)$$

The following theorem is the main result of this paper.

Theorem 5. *Let \mathbf{C} be a polar-like code with information set $\mathcal{A} = \bigcup_{p=\ell+1}^m \mathcal{N}_p$. Let (i, j, k) be a triple satisfying the conditions of Theorem 4 and (35), (36), (37), (38). Let $\bar{\mathbf{C}}$ be a code*

$n \backslash r$	2	3	4	5
6	(1, 23, 16)	–	–	–
7	–	(2, 66, 16)	(1, 100, 8)	–

Table I

NUMBER OF ADDITIONAL INFORMATION BITS m THAT CAN BE ADDED ON TOP OF THE INFORMATION LENGTH k FOR THE MINIMUM DISTANCE d ACCORDING TO THE RECURSION NUMBER n AND ORDER r OF THE UNDERLYING RM-POLAR CODE.

obtained by encoding each of the extra $m \leq t_0 + t_1 + 1$ information bits with a merged row triple $\mathbf{g}_{\{\Pi_i^\theta, \Pi_j^\theta, \Pi_k^\theta\}}$. Then,

$$d(\bar{\mathbf{C}}) = d(\mathbf{C}) \quad (40)$$

Proof: The proof is given in Appendix VII-C. ■

The following section explains how Theorem 5 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 4, (35), (36), (37) and (38). For any $m \in [1, t_0 + t_1 + 1]$,

- Each of $m - 1$ triples, i.e., $\{(i_0, j_0, k_0), \dots, (i_{m-2}, j_{m-2}, k_{m-2})\}$, corresponds to one of consecutive left-circular shifts 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 in the same distance spectrum since the underlying information set is chosen by RM rule.

- Algorithm 1 of [2] is applied to obtain the pairs (t, v) , where $t \in \mathcal{N}_{\ell+1}$, $v \in \mathcal{N}_{\ell}$, $v > t$, $\ell = i_1(\mathbf{b}_i)$, 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 of this evidence is complex and 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 the dynamic frozen bit. The pre-transformation matrix T , is such that

$$T_{a,a} = T_{a,b} = T_{a,c} = T_{t,v} = 1 \quad (41)$$

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 of [2] to obtain pairs instead of triples.

Table I summarizes the characteristics of the codes that we can construct with our method. Each entry of the table is a triple (m, k, d) where m is the number of information bits that can be added on the initial k information bits and d is the minimum distance. The code parameters are given according to two other parameters, (n, r) representing the recursive number and the order of the RM polar codes, respectively. Moreover, ‘-’ means that Theorem 5 cannot be applied for the corresponding $\text{RM}(n, r)$. For $n = 6$ and $n = 7$, we have codewords of length 64 and 128 respectively. It can be seen that, for instance, for a block length 128 and $r = 3$, the code rate can be extended from $23/64$ to $24/64$ while keeping the same minimum distance, which is an interesting improvement at this short block length.

V. SIMULATION RESULTS

We numerically compare in Figure 1 our proposed design (PD) with PAC codes and the saddle-point approximation of the MC (SMC) bound [19] for the binary input additive white

Gaussian noise 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 [2, Algorithm 1] to obtain (t, v) pairs such that $i_1(\mathbf{b}_t \bar{\cap} \mathbf{b}_v) = 1$. 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 [2, Algorithm 1] to obtain (t, v) pairs such that $i_1(\mathbf{b}_t \bar{\cap} \mathbf{b}_v) = 0$.

For PAC codes, the additional information indices are chosen as the most reliable bit sub-channel indices from the set \mathcal{N}_ℓ , which are the highest indices due to partial ordering [20]. We optimize the polynomial of the convolutional code with memory length 7 to minimize the number of minimum weight codewords. We implemented the algorithm [21] with a large list size, i.e. $5 \cdot 10^4$, and we choose the one 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.

Figure 1 plots the frame error rate (FER) versus E_b/N_0 for a code length of $N = 128$ in an additive white Gaussian noise channel and two information length, i.e. $k = 66$ and $k = 100$. Our proposed design (PD) outperforms PAC codes for the entire range of E_b/N_0 , since, at short block lengths, the minimum distance plays an important role in the SCL decoding with large list sizes. In particular, for $(128, 66)$ code, while our design performs within 0.25 dB of SMC bound at FER of $3 \cdot 10^{-4}$, PAC code achieves the same performance with 0.4 dB of additional power. For $(128, 100)$ codes, our design outperforms PAC code of about 0.2 dB at FER of 10^{-3} .

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. Our findings allow to reduce the number of minimum weight codewords of polar-like codes that perform closer to the MC bound than PAC codes with the same system parameters. We believe that this work may lead to a new method for code design, particularly at short block lengths, with interesting

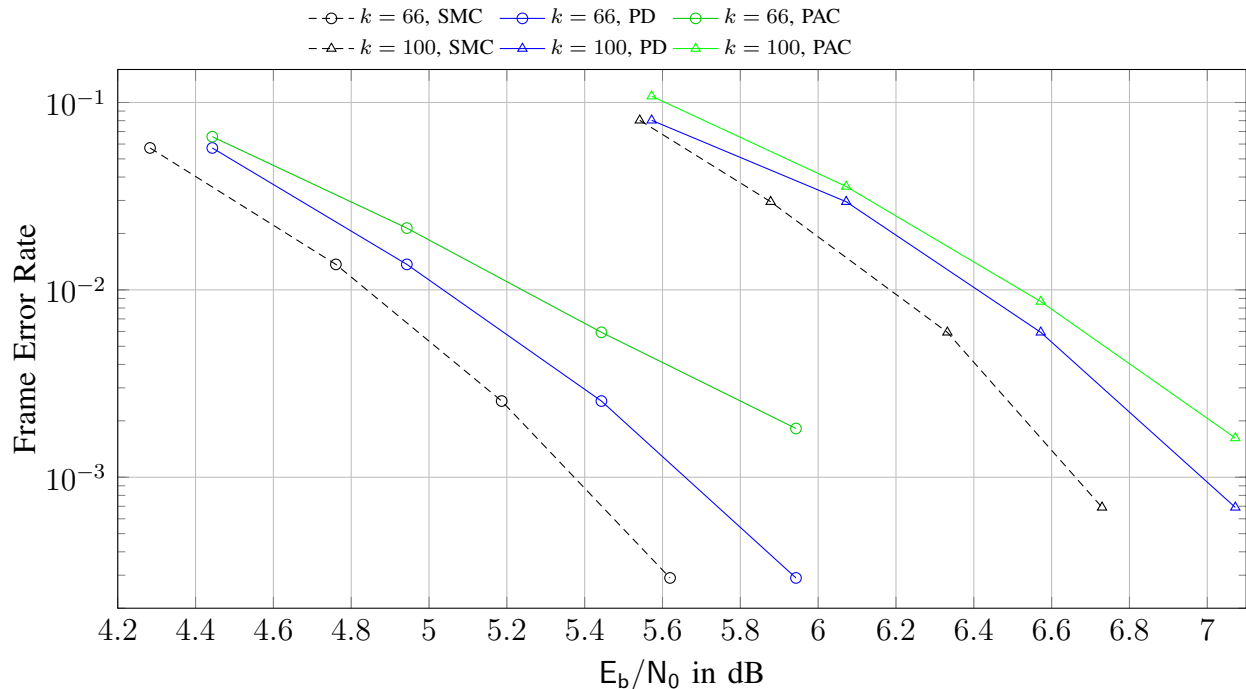


Figure 1. FER of our proposed scheme (PD), compared to SMC and PAC codes, $N = 128$, $k \in \{66; 100\}$.

performance. The extension of this work to moderate block lengths is under investigation.

REFERENCES

- [1] S. Gelincik, P. Mary, A. Savard, and J. Y. Baudais, “Increasing the minimum distance of polar-like codes with pre-transformation,” in *Submitted to ISIT*, 2022.
- [2] S. Gelincik, P. Mary, J.-Y. Baudais, and A. Savard, “Achieving pac code performance without extra computational complexity,” to appear in *IEEE ICC, Seoul, Korea*, 2022.
- [3] M. Shirvanimoghaddam, M. S. Mohammadi, R. Abbas, A. Minja, C. Yue, B. Matuz, G. Han, Z. Lin, W. Liu, Y. Li, S. Johnson, and B. Vucetic, “Short block-length codes for ultra-reliable low latency communications,” *IEEE Communications Magazine*, vol. 57, no. 2, pp. 130–137, 2019.
- [4] 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.
- [5] 3rd Generation Partnership Project (3GPP), “Multiplexing and Channel Coding,” *3GPP 38.212 V.15.3.0*, 2018.
- [6] 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.

- [7] M. Rowshan, A. Burg, and E. Viterbo, "Polarization-adjusted convolutional (pac) codes: Sequential decoding vs list decoding," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 2, pp. 1434–1447, 2021.
- [8] P. Trifonov and V. Miloslavskaya, "Polar subcodes," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 2, pp. 254–266, 2016.
- [9] P. Yuan, T. Prinz, G. Boecherer, O. Iscan, R. Boehnke, and W. Xu, "Polar code construction for list decoding," in *SCC 2019; 12th International ITG Conference on Systems, Communications and Coding*, 2019, pp. 1–6.
- [10] I. Tal and A. Vardy, "List decoding of polar codes," *IEEE Trans. Inf. Theory*, vol. 61, no. 5, pp. 2213–2226, 2015.
- [11] Q. Zhang, A. Liu, X. Pan, and K. Pan, "Crc code design for list decoding of polar codes," *IEEE Commun. Lett.*, vol. 21, no. 6, pp. 1229–1232, 2017.
- [12] 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.
- [13] E. Arkan, "From sequential decoding to channel polarization and back again," *arXiv 1908.09594*, September 2019.
- [14] B. Li, H. Zhang, and J. Gu, "On pre-transformed polar codes," *arXiv 1912.06359*, September 2019.
- [15] A. Elkelesh, M. Ebada, S. Cammerer, and S. t. Brink, "Decoder-tailored polar code design using the genetic algorithm," *IEEE Trans. Commun.*, vol. 67, no. 7, pp. 4521–4534, 2019.
- [16] T. Tonnellier and W. J. Gross, "On systematic polarization-adjusted convolutional (PAC) codes," *IEEE Commun. Lett.*, vol. 25, no. 7, pp. 2128–2132, 2021.
- [17] R. Polyanskaya, M. Davletshin, and N. Polyanskii, "Weight distributions for successive cancellation decoding of polar codes," *IEEE Transactions on Communications*, vol. 68, no. 12, pp. 7328–7336, 2020.
- [18] 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.
- [19] 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, 2020.
- [20] C. Schürch, "A partial order for the synthesized channels of a polar code," in *2016 IEEE International Symposium on Information Theory (ISIT)*, 2016, pp. 220–224.
- [21] 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.

VII. APPENDICES

A. Preliminary Theorems

In this section, we provide preliminary results that are useful in the proof of Theorem 3 and Theorem 5.

Theorem 6. For any pairs $(i, j) \in \mathcal{N}_\ell$, $\mathcal{N}_\ell := \{t : i_1(\mathbf{b}_t) = \ell, t \in \mathcal{N}\}$, such that $i_1(\mathbf{b}_i \bar{\cap} \mathbf{b}_j) = 0$, the combination of $\mathbf{g}_i \oplus \mathbf{g}_j$ with higher hamming weight rows of polar encoding matrix is lower bounded by Hamming weight of $\mathbf{g}_i \oplus \mathbf{g}_j$:

$$i_1(\mathbf{g}_i \oplus \mathbf{g}_j \oplus \mathbf{g}_\mathcal{T}) \geq i_1(\mathbf{g}_i \oplus \mathbf{g}_j) \stackrel{(a)}{=} 2^{\ell+1} - 2 \quad (42)$$

where $\mathcal{T} \subseteq \cup_{p=\ell+1}^n \mathcal{N}_p$ and (a) is by Theorem 1.

Proof: For $\ell = 1$, the statement of the theorem turns to be trivial due to the fact that $i_1(\mathbf{g}_{\{i,j,\mathcal{T}\}}) \geq 2^{\min_{k \in \{i,j,\mathcal{T}\}} i_1(\mathbf{b}_k)} = i_1(\mathbf{g}_i \oplus \mathbf{g}_j)$ for $\ell = 1$. Hence, in the following, the proof is conducted for $\ell \geq 2$.

Due to the symmetry imposed by Corollary 1, without losing generality we assume that $\mathcal{P}_1(\mathbf{b}_i) = [0, \ell - 1]$ and $\mathcal{P}_1(\mathbf{b}_j) = [\ell, 2\ell - 1]$. In the following we use this assumption for the ease of presentation.

We can divide the index set into two subsets such that

$$\hat{\mathcal{T}} := \{k : \mathcal{P}_1(\mathbf{b}_k) \cap [2 \cdot \ell, n) = \emptyset, k \in \mathcal{T}\} \quad (43)$$

and $\tilde{\mathcal{T}} = \mathcal{T} \setminus \hat{\mathcal{T}}$. Then, by Theorem 2

$$\begin{aligned} i_1(\mathbf{g}_i \oplus \mathbf{g}_j \bigoplus_{k \in \hat{\mathcal{T}}} \mathbf{g}_k \bigoplus_{t \in \tilde{\mathcal{T}}} \mathbf{g}_t) &\geq \max_{p_0 \in [2 \cdot \ell, n)} i_1(\mathbf{g}_i \oplus \mathbf{g}_j \bigoplus_{k \in \hat{\mathcal{T}}} \mathbf{g}_k \bigoplus_{t \in \tilde{\mathcal{T}}} \mathbf{g}_t \mathbb{I}\{b_{t,p_0} = 0\}) \\ &\geq \max_{p_1 \in [2 \cdot \ell, n) \setminus p_0} i_1(\mathbf{g}_i \oplus \mathbf{g}_j \bigoplus_{k \in \hat{\mathcal{T}}} \mathbf{g}_k \bigoplus_{t \in \tilde{\mathcal{T}}} \mathbf{g}_t \mathbb{I}\{b_{t,p_0} = b_{t,p_1} = 0\}) \\ &\vdots \\ &\stackrel{(a)}{\geq} i_1(\mathbf{g}_i \oplus \mathbf{g}_j \bigoplus_{k \in \hat{\mathcal{T}}} \mathbf{g}_k \bigoplus_{t \in \tilde{\mathcal{T}}} \mathbf{g}_t \mathbb{I}\{b_{t,p_0} = b_{t,p_1} = \dots = b_{t,p_{n-2\ell-1}} = 0\}) \\ &\stackrel{(b)}{=} i_1(\mathbf{g}_i \oplus \mathbf{g}_j \bigoplus_{k \in \hat{\mathcal{T}}} \mathbf{g}_k) \end{aligned} \quad (44)$$

where (a) comes from the repeated application of Theorem 2 in $[2\ell, n)$, and (b) comes from (43), which implies for any $t \in \tilde{\mathcal{T}}$, $\mathcal{P}_1(\mathbf{b}_t) \cap [2 \cdot \ell, n) \neq \emptyset$. Let us continue for $\hat{\mathcal{T}}$ which is divided

into two subsets such that

$$\hat{\mathcal{T}}^0 := \{k : \mathcal{P}_1(\mathbf{b}_k) \not\supseteq \mathcal{P}_1(\mathbf{b}_i), \mathcal{P}_1(\mathbf{b}_k) \not\supseteq \mathcal{P}_1(\mathbf{b}_j), k \in \hat{\mathcal{T}}\} \quad (45)$$

and

$$\hat{\mathcal{T}}^1 := \{k : \mathcal{P}_1(\mathbf{b}_k) \supseteq \mathcal{P}_1(\mathbf{b}_i) \text{ or } \mathcal{P}_1(\mathbf{b}_k) \supseteq \mathcal{P}_1(\mathbf{b}_j), k \in \hat{\mathcal{T}}\} \quad (46)$$

Two cases have to be investigated, i.e. when $\hat{\mathcal{T}}^0$ is not empty and $\hat{\mathcal{T}}^0$ is empty.

Case 1, $\hat{\mathcal{T}}^0 \neq \emptyset$: The elements of this set are partitioned with respect to the Hamming weights of their binary representations:

$$\hat{\mathcal{T}}^0 = \bigcup_{d=1}^{n-\ell} \hat{\mathcal{T}}_{\ell+d}^0, \quad \hat{\mathcal{T}}_{\ell+d}^0 := \{k : i_1(\mathbf{b}_k) = \ell + d, k \in \hat{\mathcal{T}}^0\} \quad (47)$$

Now, let us assume that d^* be the minimum number such that $\hat{\mathcal{T}}_{\ell+d}^0$ is not empty, i.e., $d^* = \min\{d : \hat{\mathcal{T}}_{\ell+d}^0 \neq \emptyset, n - \ell \geq d \geq 1\}$. Note that for any $t \in \hat{\mathcal{T}}_{\ell+d^*}^0$, $\exists p_0 \in \mathcal{P}_0(\mathbf{b}_i) \cap \mathcal{P}_0(\mathbf{b}_t) \cap \mathcal{P}_1(\mathbf{b}_j)$ and $p_1 \in \mathcal{P}_1(\mathbf{b}_i) \cap \mathcal{P}_0(\mathbf{b}_j) \cap \mathcal{P}_0(\mathbf{b}_t)$ such that

$$\begin{aligned} & i_1(\mathbf{g}_i \oplus \mathbf{g}_j \oplus \mathbf{g}_{\hat{\mathcal{T}}^0} \oplus \mathbf{g}_{\hat{\mathcal{T}}^1}) \\ & \stackrel{(a)}{\geq} i_1(\mathbf{g}_i \mathbb{I}\{b_{i,p_0} = 0\} \oplus \mathbf{g}_j \mathbb{I}\{b_{j,p_0} = 0\} \oplus \mathbf{g}_t \mathbb{I}\{b_{t,p_0} = 0\} \bigoplus_{k \in \hat{\mathcal{T}}^0 \setminus t} \mathbf{g}_k \mathbb{I}\{b_{k,p_0} = 0\} \bigoplus_{k \in \hat{\mathcal{T}}^1} \mathbf{g}_k \mathbb{I}\{b_{k,p_0} = 0\}) \\ & \stackrel{(b)}{\geq} i_1(\mathbf{g}_i \oplus \mathbf{g}_t \bigoplus_{k \in \hat{\mathcal{T}}^0 \setminus t} \mathbf{g}_k \mathbb{I}\{b_{k,p_0} = 0\} \bigoplus_{k \in \hat{\mathcal{T}}^1} \mathbf{g}_k \mathbb{I}\{b_{k,p_0} = 0\}) \\ & \stackrel{(c)}{\geq} i_1(\mathbf{g}_i \mathbb{I}\{b_{i,p_0} = b_{i,p_1} = 0\} \mathbf{g}_t \mathbb{I}\{b_{t,p_0} = b_{t,p_1} = 0\} \bigoplus_{k \in \hat{\mathcal{T}}^0 \setminus t} \mathbf{g}_k \mathbb{I}\{b_{k,p_0} = b_{k,p_1} = 0\} \\ & \quad \bigoplus_{k \in \hat{\mathcal{T}}^0} \mathbf{g}_k \mathbb{I}\{b_{k,p_0} = b_{k,p_1} = 0\}) \\ & = i_1(\bigoplus_{k \in \hat{\mathcal{T}}^0} \mathbf{g}_k \mathbb{I}\{b_{k,p_0} = b_{k,p_1} = 0\}) \stackrel{(d)}{\geq} 2^{\ell+d^*} \end{aligned} \quad (48)$$

where (a) and (c) come from Theorem 2 and from the conditions imposed by (45) and (46), and (b) comes from the fact $b_{j,p_0} = 1$ by hypothesis. Moreover (d) is due to the fact that $i_1(\mathbf{g}_S) \geq \min_{j \in S} 2^{i_1(\mathbf{b}_j)}$, i.e., by (13). That means if $\hat{\mathcal{T}}^0$ is not empty, the lower bound is satisfied whatever $\hat{\mathcal{T}}^1$ is empty or not.

Case 2, $\hat{\mathcal{T}}^0 = \emptyset$: The elements of $\hat{\mathcal{T}}^1$ are partitioned such as

$$\hat{\mathcal{T}}_i^1 = \begin{cases} \{k : \mathcal{P}_1(\mathbf{b}_k) \supset \mathcal{P}_1(\mathbf{b}_i), k \in \hat{\mathcal{T}}^1\} \setminus \{2^{2^\ell} - 1\} & \text{if } 2^{2^\ell} - 1 \in \hat{\mathcal{T}}^1 \\ \{k : \mathcal{P}_1(\mathbf{b}_k) \supset \mathcal{P}_1(\mathbf{b}_i), k \in \hat{\mathcal{T}}^1\} & \text{otherwise} \end{cases} \quad (49)$$

and

$$\hat{\mathcal{T}}_j^1 = \{k : \mathcal{P}_1(\mathbf{b}_k) \supset \mathcal{P}_1(\mathbf{b}_j), k \in \hat{\mathcal{T}}^1\} \quad (50)$$

Then, we obtain the total Hamming weight by partitioning the indices of row vectors into 2^ℓ subsets with respect to binary representation indices $[0, \ell - 1]$ and using Definition 2:

$$\begin{aligned} i_1(\mathbf{g}_i \oplus \mathbf{g}_j \oplus \mathbf{g}_{\hat{\mathcal{T}}_i^0} \oplus \mathbf{g}_{\hat{\mathcal{T}}_i^1}) &= i_1(\mathbf{g}_i \oplus \mathbf{g}_j \oplus \mathbf{g}_{\hat{\mathcal{T}}_i^0} \oplus \mathbf{g}_{\hat{\mathcal{T}}_i^1} | \mathcal{M}_{\ell-1}^c \cap \mathcal{M}_{\ell-2}^c \cap \dots \cap \mathcal{M}_1^c \cap \mathcal{M}_0^c) \\ &\quad + i_1(\mathbf{g}_i \oplus \mathbf{g}_j \oplus \mathbf{g}_{\hat{\mathcal{T}}_i^0} \oplus \mathbf{g}_{\hat{\mathcal{T}}_i^1} | \mathcal{M}_{\ell-1}^c \cap \mathcal{M}_{\ell-2}^c \cap \dots \cap \mathcal{M}_1^c \cap \mathcal{M}_0) \dots \\ &\quad \dots + i_1(\mathbf{g}_i \oplus \mathbf{g}_j \oplus \mathbf{g}_{\hat{\mathcal{T}}_i^0} \oplus \mathbf{g}_{\hat{\mathcal{T}}_i^1} | \mathcal{M}_{\ell-1} \cap \mathcal{M}_{\ell-2} \cap \dots \cap \mathcal{M}_1 \cap \mathcal{M}_0) \end{aligned} \quad (51)$$

By Definition 2, we can write (51) in a more compact form by denoting \mathcal{M}_p^c with 0 and \mathcal{M}_p with 1 for $p \in [0, \ell - 1]$, then

$$\begin{aligned} i_1(\mathbf{g}_i \oplus \mathbf{g}_j \oplus \mathbf{g}_{\hat{\mathcal{T}}_i^1} \oplus \mathbf{g}_{\hat{\mathcal{T}}_j^1}) &= \sum_{k=0}^{2^\ell-1} i_1(\mathbf{g}_i^{[0, \ell-1]} \mathbb{I}\{\mathcal{P}_1(\mathbf{b}_i) \supset \mathcal{P}_1(\mathbf{b}_k)\} \oplus \mathbf{g}_j^{[0, \ell-1]} \mathbb{I}\{\mathcal{P}_1(\mathbf{b}_j) \supset \mathcal{P}_1(\mathbf{b}_k)\} \\ &\quad \bigoplus_{t \in \hat{\mathcal{T}}_i^1} \mathbf{g}_t^{[0, \ell-1]} \mathbb{I}\{\mathcal{P}_1(\mathbf{b}_t) \supset \mathcal{P}_1(\mathbf{b}_k)\} \bigoplus_{t \in \hat{\mathcal{T}}_j^1} \mathbf{g}_t^{[0, \ell-1]} \mathbb{I}\{\mathcal{P}_1(\mathbf{b}_t) \supset \mathcal{P}_1(\mathbf{b}_k)\}) \end{aligned}$$

$$\stackrel{(a)}{=} \sum_{k=0}^{2^\ell-1} i_1(\mathbf{g}_{\{i, \hat{\mathcal{T}}_i^1\}}^{[0, \ell-1]}) \bigoplus_{t \in \{j, \hat{\mathcal{T}}_j^1\}} \mathbf{g}_t^{[0, \ell-1]} \mathbb{I}\{\mathcal{P}_1(\mathbf{b}_t) \supset \mathcal{P}_1(\mathbf{b}_k)\} \quad (52)$$

where (a) is by Definition 2 and conditions imposed by (49) and (50), i.e., $\mathcal{P}_1(\mathbf{b}_t) \supset \mathcal{P}_1(\mathbf{b}_k)$ for any $t \in \{i, \hat{\mathcal{T}}_i^1\}$.

Note that since $i_1(\mathbf{g}_t^{[0, \ell-1]}) = 2^\ell$, $t \in \{j, \hat{\mathcal{T}}_j^1\}$, we have

$$i_1\left(\bigoplus_{t \in \{j, \hat{\mathcal{T}}_j^1\}} \mathbf{g}_t^{[0, \ell-1]} \mathbb{I}\{\mathcal{P}_1(\mathbf{b}_t) \supset \mathcal{P}_1(\mathbf{b}_k)\}\right) = \begin{cases} 2^\ell & \text{if } |\{t : \mathcal{P}_1(\mathbf{b}_t) \supset \mathcal{P}_1(\mathbf{b}_k), t \in \{j, \hat{\mathcal{T}}_j^1\}\}| \text{ is odd,} \\ 0 & \text{otherwise} \end{cases} \quad (53)$$

and

$$1 \stackrel{(a)}{\leq} i_1(\mathbf{g}_{\{i, \hat{\mathcal{T}}_i^1\}}^{[0, \ell-1]}) \stackrel{(b)}{\leq} 2^\ell - 1 \quad (54)$$

where (a) is by (13) and (b) is by the property of polar encoding matrix that is $\mathbf{g}_{t, 2^n-1} = 0$, $0 \leq t \leq 2^n - 2$ and $\mathbf{g}_{t, 2^n-1} = 1$, if $t = 2^n - 1$. This can be deduced once it is recognized that, by (9), each $\mathbf{g}_t^{[0, \ell-1]}$, $t \in \{i, \hat{\mathcal{T}}_i^1\}$ is a polar matrix row and $\mathbf{g}_{t, 2^\ell-1}^{[0, \ell-1]} = 0$, $t \in \{i, \hat{\mathcal{T}}_i^1\}$ since $\mathcal{P}_1(\mathbf{b}_t) \not\supseteq [\ell, 2 \cdot \ell - 1]$ by (49), which imposes (b).

By (53), we can write (52) as

$$i_1(\mathbf{g}_i \oplus \mathbf{g}_j \oplus \mathbf{g}_{\hat{\mathcal{T}}_i^1} \oplus \mathbf{g}_{\hat{\mathcal{T}}_j^1}) = \alpha \cdot i_1(\mathbf{g}_{\{i, \hat{\mathcal{T}}_i^1\}}^{[0, \ell-1]}) + (2^\ell - \alpha) \cdot (2^\ell - i_1(\mathbf{g}_{\{i, \hat{\mathcal{T}}_i^1\}}^{[0, \ell-1]})). \quad (55)$$

Note that the α is such that

$$1 \stackrel{(a)}{\leq} \alpha \stackrel{(b)}{\leq} 2^\ell - 1 \quad (56)$$

where (a) is due to the fact that, when $k = 2^\ell - 1$ and by (10), $\mathbf{g}_{\{i, \hat{\mathcal{T}}_i^1\}}^{[0, \ell-1]}$ is the only term in (52), i.e.

$$i_1(\mathbf{g}_{\{i, \hat{\mathcal{T}}_i^1\}}^{[0, \ell-1]}) \bigoplus_{t \in \{j, \hat{\mathcal{T}}_j^1\}} \mathbf{g}_t \mathbb{I}\{\mathcal{P}_1(\mathbf{b}_t) \supset \mathcal{P}_1(\mathbf{b}_{2^\ell-1})\} = i_1(\mathbf{g}_{\{i, \hat{\mathcal{T}}_i^1\}}^{[0, \ell-1]}) \quad (57)$$

and (b) comes from the fact that \mathbf{g}_{t^*} , $t^* = \max\{j, \mathcal{T}_j^1\}$ is the only element from $\{j, \mathcal{T}_j^1\}$ such that (52) simplifies to

$$i_1(\mathbf{g}_{\{i, \hat{\mathcal{T}}_i^1\}}^{[0, \ell-1]} \bigoplus_{t \in \{j, \mathcal{T}_j^1\}} \mathbf{g}_t^{[0, \ell-1]} \mathbb{I}\{\mathcal{P}_1(\mathbf{b}_t) \supset \mathcal{P}_1(\mathbf{b}_{t^*})\}) = i_1(\mathbf{g}_{\{i, \hat{\mathcal{T}}_i^1\}}^{[0, \ell-1]} \oplus \mathbf{g}_{t^*}^{[0, \ell-1]}) = 2^\ell - i_1(\mathbf{g}_{\{i, \hat{\mathcal{T}}_i^1\}}^{[0, \ell-1]}) \quad (58)$$

by using (10). When (54) and (56) are satisfied, (55) cannot be less than $2^{\ell+1} - 2$. This ends the proof for Theorem 6. \blacksquare

Theorem 7. For any pairs $(i, j) \in \mathcal{N}_\ell$, $\ell \geq 2$, the combination of $\mathbf{g}_i \oplus \mathbf{g}_j$ with higher Hamming weight rows of polar encoding matrix is lower bounded by Hamming weight of $\mathbf{g}_i \oplus \mathbf{g}_j$

$$i_1(\mathbf{g}_i \oplus \mathbf{g}_j \oplus \mathbf{g}_\mathcal{T}) \geq i_1(\mathbf{g}_i \oplus \mathbf{g}_j) \quad (59)$$

where $\mathcal{T} \subseteq \cup_{p=\ell+1}^n \mathcal{N}_p$.

Proof: We can divide the index set into two subsets such that

$$\tilde{\mathcal{T}} := \{k : \mathcal{P}_1(\mathbf{b}_k) \cap \mathcal{P}_0(\mathbf{b}_i \bar{\cap} \mathbf{b}_j) \neq \emptyset, k \in \mathcal{T}\} \quad (60)$$

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

Applying the same reasoning as in Theorem 6 in (44), the proof is conducted for $\hat{\mathcal{T}}$.

When $i_1(\mathbf{b}_i \bar{\cap} \mathbf{b}_j) = \ell - 1$, the proof is trivial since

$$i_1(\mathbf{g}_{\hat{\mathcal{T}} \cup \{i, j\}}) \geq \min_{k \in \hat{\mathcal{T}} \cup \{i, j\}} 2^{i_1(\mathbf{b}_k)} = 2^{i_1(\mathbf{b}_i)} = 2^\ell \quad (61)$$

and $i_1(\mathbf{g}_i \oplus \mathbf{g}_j) = 2^\ell$ by [2, Theorem 2] if $i_1(\mathbf{b}_i \bar{\cap} \mathbf{b}_j) = \ell - 1$.

Now assume that $i_1(\mathbf{b}_i \bar{\cap} \mathbf{b}_j) = |\mathcal{P}_1(\mathbf{b}_i \bar{\cap} \mathbf{b}_j)| \leq \ell - 2$ and let $\mathcal{W} = \mathcal{P}_1(\mathbf{b}_i \bar{\cap} \mathbf{b}_j)$, which is the index set of common one bit positions of \mathbf{b}_i and \mathbf{b}_j . Then, by partitioning the row indices of the polar encoding matrix, we obtain the following expression

$$\begin{aligned} i_1(\mathbf{g}_{\{i, j\}} \oplus \mathbf{g}_{\hat{\mathcal{T}}}) &= i_1(\mathbf{g}_{\{i, j\}} \oplus \mathbf{g}_{\hat{\mathcal{T}}} | \mathcal{M}_{\mathcal{W}(|\mathcal{W}|-1)}^c \cap \mathcal{M}_{\mathcal{W}(|\mathcal{W}|-2)}^c \cap \cdots \cap \mathcal{M}_{\mathcal{W}(0)}^c) \\ &\quad + i_1(\mathbf{g}_{\{i, j\}} \oplus \mathbf{g}_{\hat{\mathcal{T}}} | \mathcal{M}_{\mathcal{W}(|\mathcal{W}|-1)}^c \cap \mathcal{M}_{\mathcal{W}(|\mathcal{W}|-2)}^c \cap \cdots \cap \mathcal{M}_{\mathcal{W}(0)}) \end{aligned} \quad (62)$$

$$\begin{aligned}
& \vdots \\
& + i_1(\mathbf{g}_{\{i,j\}} \oplus \mathbf{g}_{\hat{\mathcal{T}}} | \mathcal{M}_{\mathcal{W}(|\mathcal{W}|-1)} \cap \mathcal{M}_{\mathcal{W}(|\mathcal{W}|-2)} \cap \cdots \cap \mathcal{M}_{\mathcal{W}(0)}) \\
& \stackrel{(a)}{=} i_1(\mathbf{g}_{\{i,j\}}^{\mathcal{W}} \bigoplus_{k \in \hat{\mathcal{T}}} \mathbf{g}_k^{\mathcal{W}}) + i_1(\mathbf{g}_{\{i,j\}}^{\mathcal{W}} \bigoplus_{k \in \hat{\mathcal{T}}} \mathbf{g}_k^{\mathcal{W}} \mathbb{I}\{b_{k,\mathcal{W}(0)} = 1\}) \\
& \vdots \\
& + i_1(\mathbf{g}_{\{i,j\}}^{\mathcal{W}} \bigoplus_{k \in \hat{\mathcal{T}}} \mathbf{g}_k^{\mathcal{W}} \mathbb{I}\{b_{k,\mathcal{W}(|\mathcal{W}|-1)} = b_{k,\mathcal{W}(|\mathcal{W}|-2)} = \cdots = b_{k,\mathcal{W}(0)} = 1\})
\end{aligned}$$

where (a) is due to (9) and (10). By Theorem 6, each term of the partition is greater than $2^{\ell-|\mathcal{W}|+1} - 2$ since $|\mathcal{P}_1(\mathbf{b}_k) \cap \mathcal{P}_1(\mathbf{b}_i \oplus \mathbf{b}_j)| > \ell - |\mathcal{W}|$ for any $k \in \hat{\mathcal{T}}$ by assumption. Then,

$$\begin{aligned}
i_1(\mathbf{g}_i \oplus \mathbf{g}_j \oplus \mathbf{g}_{\hat{\mathcal{T}}}) &\geq 2^{i_1(\mathbf{b}_i \bar{\cap} \mathbf{b}_j)} \cdot (2^{\ell-i_1(\mathbf{b}_i \bar{\cap} \mathbf{b}_j)+1} - 2) = 2^{\ell+1} - 2^{i_1(\mathbf{b}_i \bar{\cap} \mathbf{b}_j)+1} \\
&\stackrel{(a)}{=} i_1(\mathbf{g}_i \oplus \mathbf{g}_j)
\end{aligned} \tag{63}$$

where (a) is due to Theorem 1. ■

B. Proof of Theorem 3

The proof relies on Theorem 6, given at the beginning of the appendices. Theorem 6 states that, for any given pair of rows with the same Hamming weight and no intersection in their binary representations, the Hamming weights of combination of the given pair with any subset of rows with higher Hamming weights is lower bounded by the Hamming weight of the given pair.

Since

$$d(\bar{\mathbf{C}}) = \min\{d(\mathbf{C}), \min_{\mathcal{T} \subseteq \mathcal{A}} i_1(\mathbf{g}_{\{i,j,k\}} \oplus \mathbf{g}_{\mathcal{T}})\} \tag{64}$$

it is sufficient to prove the following statement

$$i_1(\mathbf{g}_{\{i,j,k\}} \oplus \mathbf{g}_{\mathcal{T}}) \geq 2^{\ell+1}, \quad \forall \mathcal{T} \subseteq \mathcal{A} \tag{65}$$

to prove the theorem.

We can divide the index set into two subsets such that

$$\tilde{\mathcal{T}} := \{t : \mathcal{P}_1(\mathbf{b}_t) \cap \mathcal{P}_0(\mathbf{b}_i \bar{\cup} \mathbf{b}_j \bar{\cup} \mathbf{b}_k) \neq \emptyset, t \in \mathcal{T}\} \quad (66)$$

and $\hat{\mathcal{T}} = \mathcal{T} \setminus \tilde{\mathcal{T}}$. Then by Theorem 2

$$\begin{aligned} i_1(\mathbf{g}_{\{i,j,k\}} \oplus \mathbf{g}_{\hat{\mathcal{T}}} \oplus \mathbf{g}_{\tilde{\mathcal{T}}}) &\geq \max_{p_0 \in \mathcal{P}_0(\mathbf{b}_i \bar{\cup} \mathbf{b}_j \bar{\cup} \mathbf{b}_k)} i_1(\mathbf{g}_{\{i,j\}} \oplus \mathbf{g}_{\hat{\mathcal{T}}} \bigoplus_{t \in \tilde{\mathcal{T}}} \mathbf{g}_t \mathbb{I}\{b_{t,p_0} = 0\}) \\ &\geq \max_{p_1 \in \mathcal{P}_0(\mathbf{b}_i \bar{\cup} \mathbf{b}_j \bar{\cup} \mathbf{b}_k) \setminus \{p_0\}} i_1(\mathbf{g}_{\{i,j,k\}} \oplus \mathbf{g}_{\hat{\mathcal{T}}} \bigoplus_{t \in \tilde{\mathcal{T}}} \mathbf{g}_t \mathbb{I}\{b_{t,p_0} = b_{t,p_1} = 0\}) \\ &\vdots \\ &\geq i_1(\mathbf{g}_{\{i,j,k\}} \oplus \mathbf{g}_{\hat{\mathcal{T}}} \bigoplus_{t \in \tilde{\mathcal{T}}} \mathbf{g}_t \mathbb{I}\{b_{t,p_0} = b_{t,p_1} = \dots = b_{t,p_{n-2\ell-1}} = 0\}) \\ &\stackrel{(a)}{=} i_1(\mathbf{g}_i \oplus \mathbf{g}_j \oplus \mathbf{g}_k \oplus \mathbf{g}_{\hat{\mathcal{T}}}) \end{aligned} \quad (67)$$

where $\{p_0, p_1, \dots, p_{n-2\ell-1}\} = \mathcal{P}_0(\mathbf{b}_i \bar{\cup} \mathbf{b}_j \bar{\cup} \mathbf{b}_k)$ and (a) comes from (66), which implies for any $t \in \tilde{\mathcal{T}}$, $\mathcal{P}_1(\mathbf{b}_t) \cap \mathcal{P}_0(\mathbf{b}_i \bar{\cup} \mathbf{b}_j \bar{\cup} \mathbf{b}_k) \neq \emptyset$. This result means that the Hamming weight of $\mathbf{g}_{\{i,j,k,\mathcal{T}\}}$ is lower bounded by the Hamming weight of $\mathbf{g}_{\{i,j,k,\hat{\mathcal{T}}\}}$. Therefore, in the following, the proof is conducted for $\hat{\mathcal{T}}$.

We divide $\hat{\mathcal{T}}$ into three subsets

$$\hat{\mathcal{T}}_{k,1} = \{t : \mathcal{P}_1(\mathbf{b}_t) \supset \mathcal{P}_1(\mathbf{b}_k), t \in \hat{\mathcal{T}}\} \quad (68)$$

$$\hat{\mathcal{T}}_{k,2} = \{t : |\mathcal{P}_1(\mathbf{b}_t) \cap \mathcal{P}_1(\mathbf{b}_k)| = 1, t \in \hat{\mathcal{T}}\} \quad (69)$$

$$\tilde{\mathcal{T}}_0 = \hat{\mathcal{T}} \setminus \{\hat{\mathcal{T}}_{k,1} \cup \hat{\mathcal{T}}_{k,2}\} \quad (70)$$

In the following, we show that the theorem holds whatever the sets $\hat{\mathcal{T}}_{k,1}$ and $\hat{\mathcal{T}}_{k,2}$ are empty or not.

Case 1, $\hat{\mathcal{T}}_{k,1} \neq \emptyset$ and $\hat{\mathcal{T}}_{k,2} \neq \emptyset$: Since, $k \in \mathcal{N}_2$, we denote $\{p_1, p_2\} = \mathcal{P}_1(\mathbf{b}_k)$. Then, by partitioning the row indices of the polar encoding matrix, we obtain the following expression:

$$i_1(\mathbf{g}_i \oplus \mathbf{g}_j \oplus \mathbf{g}_k \oplus \mathbf{g}_{\hat{\mathcal{T}}}) = i_1(\mathbf{g}_i \oplus \mathbf{g}_j \oplus \mathbf{g}_k \oplus \mathbf{g}_{\hat{\mathcal{T}}} | \mathcal{M}_{p_1}^c \cap \mathcal{M}_{p_2}^c) + i_1(\mathbf{g}_i \oplus \mathbf{g}_j \oplus \mathbf{g}_k \oplus \mathbf{g}_{\hat{\mathcal{T}}} | \mathcal{M}_{p_1}^c \cap \mathcal{M}_{p_2})$$

$$\begin{aligned}
& + i_1(\mathbf{g}_i \oplus \mathbf{g}_j \oplus \mathbf{g}_k \oplus \mathbf{g}_{\mathcal{T}} | \mathcal{M}_{p_1} \cap \mathcal{M}_{p_2}^c) + i_1(\mathbf{g}_i \oplus \mathbf{g}_j \oplus \mathbf{g}_k \oplus \mathbf{g}_{\mathcal{T}} | \mathcal{M}_{p_1} \cap \mathcal{M}_{p_2}) \\
\stackrel{(a)}{=} & i_1(\mathbf{g}_i^{\{p_1, p_2\}} \oplus \mathbf{g}_j^{\{p_1, p_2\}} \oplus \mathbf{g}_k^{\{p_1, p_2\}} \bigoplus_{\theta \in \hat{\mathcal{T}}_{k,1}} \mathbf{g}_\theta^{\{p_1, p_2\}} \bigoplus_{\theta \in \hat{\mathcal{T}}_{k,2}} \mathbf{g}_\theta^{\{p_1, p_2\}} \bigoplus_{\theta \in \tilde{\mathcal{T}}_0} \mathbf{g}_\theta^{\{p_1, p_2\}}) \\
& + i_1(\mathbf{g}_k^{\{p_1, p_2\}} \bigoplus_{\theta \in \hat{\mathcal{T}}_{k,1}} \mathbf{g}_\theta^{\{p_1, p_2\}} \bigoplus_{\theta \in \hat{\mathcal{T}}_{k,2}} \mathbf{g}_\theta^{\{p_1, p_2\}} \mathbb{I}\{b_{\theta, p_2} = 1\}) \\
& + i_1(\mathbf{g}_k^{\{p_1, p_2\}} \bigoplus_{\theta \in \hat{\mathcal{T}}_{k,1}} \mathbf{g}_\theta^{\{p_1, p_2\}} \bigoplus_{\theta \in \hat{\mathcal{T}}_{k,2}} \mathbf{g}_\theta^{\{p_1, p_2\}} \mathbb{I}\{b_{\theta, p_1} = 1\}) + i_1(\mathbf{g}_k^{\{p_1, p_2\}} \bigoplus_{\theta \in \hat{\mathcal{T}}_{k,1}} \mathbf{g}_\theta^{\{p_1, p_2\}}) \\
\stackrel{(b)}{\geq} & i_1(\mathbf{g}_i^{\{p_1, p_2\}} \oplus \mathbf{g}_j^{\{p_1, p_2\}} \bigoplus_{\theta \in \hat{\mathcal{T}}_{k,2}} \mathbf{g}_\theta^{\{p_1, p_2\}} \mathbb{I}\{b_{\theta, p_2} = 0\} \bigoplus_{\theta \in \tilde{\mathcal{T}}_0} \mathbf{g}_\theta^{\{p_1, p_2\}}) \\
& + i_1(\mathbf{g}_k^{\{p_1, p_2\}} \bigoplus_{\theta \in \hat{\mathcal{T}}_{k,1}} \mathbf{g}_\theta^{\{p_1, p_2\}} \bigoplus_{\theta \in \hat{\mathcal{T}}_{k,2}} \mathbf{g}_\theta^{\{p_1, p_2\}} \mathbb{I}\{b_{\theta, p_1} = 1\}) + i_1(\mathbf{g}_k^{\{p_1, p_2\}} \bigoplus_{\theta \in \hat{\mathcal{T}}_{k,1}} \mathbf{g}_\theta^{\{p_1, p_2\}}) \\
\stackrel{(c)}{\geq} & i_1(\mathbf{g}_i^{\{p_1, p_2\}} \oplus \mathbf{g}_j^{\{p_1, p_2\}} \bigoplus_{\theta \in \hat{\mathcal{T}}_{k,2}} \mathbf{g}_\theta^{\{p_1, p_2\}} \mathbb{I}\{b_{\theta, p_2} = 0\} \bigoplus_{\theta \in \tilde{\mathcal{T}}_0} \mathbf{g}_\theta^{\{p_1, p_2\}}) \\
& + i_1(\bigoplus_{\theta \in \hat{\mathcal{T}}_{k,2}} \mathbf{g}_\theta^{\{p_1, p_2\}} \mathbb{I}\{b_{\theta, p_1} = 1\}) \stackrel{(d)}{\geq} 2^{\ell+1} \tag{71}
\end{aligned}$$

where (a) is due to (9) and (10), (b) and (c) are due to the fact that $i_1(\mathbf{u} \oplus \mathbf{v}) + i_1(\mathbf{v}) \geq i_1(\mathbf{u})$, i.e., by (22), and (d) is by (13), i.e., the Hamming weight of any row at each expression is at least 2^ℓ . This means that if neither $\hat{\mathcal{T}}_{k,1}$ nor $\hat{\mathcal{T}}_{k,2}$ is empty, the lower bound is satisfied whatever $\tilde{\mathcal{T}}_0$.

Case 2, $\hat{\mathcal{T}}_{k,1} = \hat{\mathcal{T}}_{k,2} = \emptyset$: By partitioning the row indices with respect to $\{p_1, p_2\} = \mathcal{P}_1(\mathbf{b}_k)$, we obtain the following expression:

$$\begin{aligned}
i_1(\mathbf{g}_i \oplus \mathbf{g}_j \oplus \mathbf{g}_k \oplus \mathbf{g}_{\mathcal{T}}) & = i_1(\mathbf{g}_i \oplus \mathbf{g}_j \oplus \mathbf{g}_k \oplus \mathbf{g}_{\mathcal{T}} | \mathcal{M}_{p_1}^c \cap \mathcal{M}_{p_2}^c) + i_1(\mathbf{g}_i \oplus \mathbf{g}_j \oplus \mathbf{g}_k \oplus \mathbf{g}_{\mathcal{T}} | \mathcal{M}_{p_1}^c \cap \mathcal{M}_{p_2}) \\
& + i_1(\mathbf{g}_i \oplus \mathbf{g}_j \oplus \mathbf{g}_k \oplus \mathbf{g}_{\mathcal{T}} | \mathcal{M}_{p_1} \cap \mathcal{M}_{p_2}^c) + i_1(\mathbf{g}_i \oplus \mathbf{g}_j \oplus \mathbf{g}_k \oplus \mathbf{g}_{\mathcal{T}} | \mathcal{M}_{p_1} \cap \mathcal{M}_{p_2}) \\
& = i_1(\mathbf{g}_i^{\{p_1, p_2\}} \oplus \mathbf{g}_j^{\{p_1, p_2\}} \oplus \mathbf{g}_k^{\{p_1, p_2\}} \bigoplus_{\theta \in \tilde{\mathcal{T}}_0} \mathbf{g}_\theta^{\{p_1, p_2\}}) + i_1(\mathbf{g}_k^{\{p_1, p_2\}}) + 2 \cdot i_1(\mathbf{g}_k^{\{p_1, p_2\}}) \\
& \quad \underbrace{\geq i_1(\mathbf{g}_i^{\{p_1, p_2\}} \oplus \mathbf{g}_j^{\{p_1, p_2\}} \bigoplus_{\theta \in \tilde{\mathcal{T}}_0} \mathbf{g}_\theta^{\{p_1, p_2\}})}_{\geq i_1(\mathbf{g}_i^{\{p_1, p_2\}} \oplus \mathbf{g}_j^{\{p_1, p_2\}} \bigoplus_{\theta \in \tilde{\mathcal{T}}_0} \mathbf{g}_\theta^{\{p_1, p_2\}}) \text{ by (22)}}
\end{aligned}$$

$$\geq i_1(\mathbf{g}_i^{\{p_1, p_2\}} \oplus \mathbf{g}_j^{\{p_1, p_2\}} \bigoplus_{\theta \in \tilde{\mathcal{T}}_0} \mathbf{g}_\theta^{\{p_1, p_2\}}) + 2 \cdot i_1(\mathbf{g}_k^{\{p_1, p_2\}}) \stackrel{(a)}{=} 2^{\ell+1} \quad (72)$$

where (a) comes from the fact that the rows $\mathbf{g}_i^{\{p_1, p_2\}}$, $\mathbf{g}_j^{\{p_1, p_2\}}$ and $\mathbf{g}_{\tilde{\mathcal{T}}_0}^{\{p_1, p_2\}}$ comply with the conditions of Theorem 6. Hence, $i_1(\mathbf{g}_{\{i, j, \tilde{\mathcal{T}}_0\}}^{\{p_1, p_2\}}) \geq 2^{\ell+1} - 2$, and $2 \cdot i_1(\mathbf{g}_k^{\{p_1, p_2\}}) = 2$.

The same result can be shown for the other cases of $\hat{\mathcal{T}}_{k,1}$, $\hat{\mathcal{T}}_{k,2}$ by following similar steps.

C. Proof of Theorem 5

The proof of this theorem is based on Theorem 2 and Theorem 7. The latter is a generalization of Theorem 6 for the case where there are some common 1-bit indices in the intersection of the binary representations of the pair.

For $m = t_0 + t_1 + 1$, the code $\bar{\mathbf{C}}$ is given as

$$\bar{\mathbf{C}} = \{\mathbf{C}\} \bigcup_{\mathcal{D} \subseteq [0, t_0 + t_1]} \{\mathbf{c} : \mathbf{c} = \bigoplus_{\theta \in \mathcal{D}} \mathbf{g}_{\{\Pi_i^\theta, \Pi_j^\theta, \Pi_k^\theta\}} \oplus \mathbf{g}_{\mathcal{T}}, \mathcal{T} \subseteq \mathcal{A}\} \quad (73)$$

and hence, it is sufficient to prove the following statement

$$i_1\left(\bigoplus_{\theta \in \mathcal{D}} \mathbf{g}_{\{\Pi_i^\theta, \Pi_j^\theta, \Pi_k^\theta\}} \oplus \mathbf{g}_{\mathcal{T}}\right) \geq 2^{\ell+1}, \quad (74)$$

for any $\mathcal{D} \subseteq [0, t_0 + t_1]$ and $\mathcal{T} \subseteq \mathcal{A}$.

The proof is done for the whole set $\{\mathbf{g}_{\Pi_i^d} \oplus \mathbf{g}_{\Pi_j^d} \oplus \mathbf{g}_{\Pi_k^d}\}$, $d \in [0, t_0 + t_1]$ and the case $i_0(\mathbf{b}_i \bar{\cup} \mathbf{b}_j \bar{\cup} \mathbf{b}_k) > i_1(\mathbf{b}_i \bar{\cap} \mathbf{b}_j \bar{\cap} \mathbf{b}_k)$. The same result can be found by following similar steps for any combination of $\{\mathbf{g}_{\Pi_i^d} \oplus \mathbf{g}_{\Pi_j^d} \oplus \mathbf{g}_{\Pi_k^d}\}$, $d \in [0, t_0 + t_1]$ and other cases such as $i_0(\mathbf{b}_i \bar{\cup} \mathbf{b}_j \bar{\cup} \mathbf{b}_k) = i_1(\mathbf{b}_i \bar{\cap} \mathbf{b}_j \bar{\cap} \mathbf{b}_k)$ or $i_0(\mathbf{b}_i \bar{\cup} \mathbf{b}_j \bar{\cup} \mathbf{b}_k) = 0$.

For the considered case, the main steps are summarized in (75) on top of the next page, where (a) is due to the fact that $\mathcal{M}_{n-2}^c \cap \mathcal{M}_{n-2} = \emptyset$, (b) is due to (8), (c) is due to the fact that the minimum Hamming weight of any row involved in the right-hand side of the equation is $2^{\ell-1}$, (d) is due to Theorem 2, (e) is due to (35), (36) and (37), (f) is since $\Pi_i^{t_1}, \Pi_j^{t_1}$ and $\{t : b_{t, n-1} = b_{t, t_0} = \dots = b_{t, t_1} = 0, t \in \mathcal{T}\}$ comply with the conditions of Theorem 7, (g) is due to Theorem 1 and (h) is due to the assumption $i_1(\mathbf{b}_i \bar{\cap} \mathbf{b}_j) \leq \ell - 2$.

The proof can be conducted the same way for all $m < t_0 + t_1 + 1$ and the proof is complete.

$$\begin{aligned}
& i_1 \left(\bigoplus_{d=0}^{t_0+t_1} \mathfrak{g}_{\{\Pi_i^d, \Pi_j^d, \Pi_k^d\}} \oplus \mathfrak{g}_{\mathcal{T}} \right) \stackrel{(a)}{=} i_1 \left(\bigoplus_{d=0}^{t_0+t_1} \mathfrak{g}_{\{\Pi_i^d, \Pi_j^d, \Pi_k^d\}} \oplus \mathfrak{g}_{\mathcal{T}} | \mathcal{M}_{n-2}^c \right) + i_1 \left(\bigoplus_{d=0}^{t_0+t_1} \mathfrak{g}_{\{\Pi_i^d, \Pi_j^d, \Pi_k^d\}} \oplus \mathfrak{g}_{\mathcal{T}} | \mathcal{M}_{n-2} \right) \\
& \stackrel{(b)}{=} i_1 \left(\bigoplus_{d=0}^{t_0+t_1} \mathfrak{g}_{\{\Pi_i^d, \Pi_j^d, \Pi_k^d\}}^{n-2} \oplus \mathfrak{g}_{\mathcal{T}}^{n-2} \right) + i_1 \left(\bigoplus_{d=0}^{t_0+t_1} \bigoplus_{t \in \{\Pi_i^d, \Pi_j^d, \Pi_k^d\}} \mathfrak{g}_t^{n-2} \mathbb{I}\{b_{t,n-2} = 1\} \bigoplus_{t \in \mathcal{T}} \mathfrak{g}_t^{n-2} \mathbb{I}\{b_{t,n-2} = 1\} \right) \\
& \stackrel{(c)}{\geq} i_1 \left(\bigoplus_{d=0}^{t_0+t_1} \mathfrak{g}_{\{\Pi_i^d, \Pi_j^d, \Pi_k^d\}}^{n-2} \oplus \mathfrak{g}_{\mathcal{T}}^{n-2} \right) + 2^{\ell-1} \\
& \stackrel{(d)}{\geq} i_1 \left(\bigoplus_{d=0}^{t_0+t_1} \bigoplus_{t \in \{\Pi_i^d, \Pi_j^d, \Pi_k^d\}} \mathfrak{g}_t^{n-2} \mathbb{I}\{b_{t,n-1} = b_{t,t_0} = \dots = b_{t,t_1} = 0\} \bigoplus_{t \in \mathcal{T}} \mathfrak{g}_t^{n-2} \mathbb{I}\{b_{t,n-1} = b_{t,t_0} = \dots = b_{t,t_1} = 0\} \right) \\
& \hspace{25em} + 2^{\ell-1} \\
& \stackrel{(e)}{=} i_1 \left(\mathfrak{g}_{\Pi_i^{t_1}}^{n-2} \oplus \mathfrak{g}_{\Pi_j^{t_1}}^{n-2} \bigoplus_{t \in \mathcal{T}} \mathfrak{g}_t^{n-2} \mathbb{I}\{b_{t,n-1} = b_{t,t_0} = \dots = b_{t,t_1} = 0\} \right) + 2^{\ell-1} \\
& \stackrel{(f)}{\geq} i_1 \left(\mathfrak{g}_{\Pi_i^{t_1}}^{n-2} \oplus \mathfrak{g}_{\Pi_j^{t_1}}^{n-2} \right) + 2^{\ell-1} \stackrel{(g)}{=} 2^{\ell+1} - 2^{i_1(\mathbf{b}_i \bar{\cap} \mathbf{b}_j) + 1} + 2^{\ell-1} \stackrel{(h)}{\geq} 2^{\ell+1} \tag{75}
\end{aligned}$$