A Lower Bound on Minimum Distance of Convolutional Polar Codes

Ruslan Morozov, Peter Trifonov
Saint Petersburg Polytechnic University, Russia
{rmorozov,petert}@dcn.icc.spbstu.ru

***Abstract*—A lower bound on minimum distance of convolutional polar codes is provided. The bound is obtained from the minimum weight of generalized cosets of the codes generated by bottom rows of the polarizing matrix. Minimum weight of cosets is also used for construction of convolutional polar subcodes, which provide lower FER under list SC decoding compared to polar subcodes and convolutional polar codes.**

***Index Terms*—convolutional polar codes, convolutional polar subcodes, polar codes, minimum distance, coset.**

I. Introduction

In this paper we consider convolutional polar codes (CvPCs) [1], which are shown to provide substantially better performance under successive cancellation (SC) decoding compared to classical polar codes [2]. In [1] both open-boundary and periodic-boundary CvPCs are presented, in this paper by CvPCs we always mean open-boundary CvPCs.

The problem of computing the minimum distance of an arbitrary linear code is NP-complete. For moderate-length codes the minimum distance can be obtained by the method presented in [3]. For Arikan polar codes, the minimum distance is given by the smallest unfrozen row weight. This is due to the fact that the minimum weight of the coset, given by the i-th row of matrix of the polarizing transformation, of the code, given by all rows below the i-th row, is equal to the weight of the i-th row. For the convolutional polarizing transformation (CvPT) it is not true in general. We propose a method for computing the minimum weight of coset generated by the i-th row of the CvPT matrix. Using these values, we derive a lower bound on the minimum distance of a CvPC. The minimum coset weights can be also used for generalizing the construction of randomized polar subcodes [4] to the case of CvPCs and obtaining convolutional polar subcodes (CvPSs) with lower FER under list SC decoding [5], [6], [7].

II. Background

A. Notations

The following notations are used throughout the paper. $\mathbb{F}$ denotes $GF(2)$. For integer n we denote $[n] = \{0, 1, \ldots, n-1\}$. For vector a symbol $a_b^c = (a_b, a_{b+1}, \ldots, a_c)$. For two vectors a and b we denote their concatenation by (a, b). For $m \times\ n$ matrix A and sets $\mathcal{X} \subseteq [m]$, $\mathcal{Y} \subseteq [n]$, by $A_{\mathcal{X},\mathcal{Y}}$ we denote the submatrix of A with rows with indices from set $\mathcal{X}$ and columns with indices from set $\mathcal{Y}$, indexing of rows and columns starts with zero. Similar notations are applied to vectors as well. If $\mathcal{X} = *$ or $\mathcal{Y} = *$, this means that all rows or all columns of the original matrix are in the submatrix. Furthermore, $A_{\overline{\mathcal{X}},\overline{\mathcal{Y}}}$ denotes submatrix of A consisting of rows and columns with indices that are not in $\mathcal{X}$ and $\mathcal{Y}$, respectively. The vector of i zeroes is denoted by $\mathbf{0}^i$, or just by $\mathbf{0}$ if i is clear from the context.

B. A Linear Block Code Representation and SC Decoding

Consider a binary linear block code in the form

$$\left\{u_0^{n-1}G^{(n)} \middle| u_{\mathcal{I}} \in \mathbb{F}^k, u_{\mathcal{F}} = \mathbf{0}\right\}, \mathcal{I} \subseteq [n], |\mathcal{I}| = k, \quad (1)$$

where $G^{(n)}$ is an $n \times n$ non-singular binary matrix, $\mathcal{I}$ is called the information set and $\mathcal{F} = [n] \backslash \mathcal{I}$ is called the frozen set. The generator matrix of such code is $G^{(n)}_{\mathcal{I},*}$. Note that any (n, k) linear code with generator matrix G can be expressed as (1) using $G^{(n)}$, s.t. $G = G^{(n)}_{\mathcal{I},*}$ for some $\mathcal{I} \subseteq [n]$. For example, classical polar codes [2] have $G^{(n)} = F^{\otimes m}$ for $n = 2^m$.

For such code representation, the successive cancellation (SC) decoding method can be defined. Consider transmission of codeword $c_0^{n-1} = u_0^{n-1}G^{(n)}$ through a binary-input memoryless channel $\mathcal{W} : \mathbb{F} \to \mathcal{Y}$. Let y_0^{n-1} be the output of this channel. After demodulation, the probabilities $W(c_i|y_i) = \mathcal{W}(y_i|c_i)/\left(\mathcal{W}(y_i|0) + \mathcal{W}(y_i|1)\right)$ for $c_i \in \mathbb{F}$ are provided to the decoding algorithm. Given the prior hard decisions $\hat{u}_0 \ldots \hat{u}_{\varphi-1}$, at phase φ the SC decoding algorithm calculates probabilities $W_n^{(\varphi)}(\hat{u}_0^{\varphi-1}, u_\varphi|y_0^{n-1})$, defined as

$$W_n^{(\varphi)}(u_0^{\varphi}|y_0^{n-1}) = \sum_{u_{\varphi+1}^{n-1} \in \mathbb{F}^{n-\varphi-1}} W^n(u_0^{n-1}G^{(n)}|y_0^{n-1}), \quad (2)$$

where $W^n(c_0^{n-1}|y_0^{n-1}) = \prod_{i=0}^{n-1} W(c_i|y_i)$. The channels $W_n^{(\varphi)} : \mathcal{Y} \to \mathbb{F}^{\varphi+1}$ are called *bit subchannels*. Then, the hard decision on u_φ is made by

$$\hat{u}_\varphi = \begin{cases} 0, & \varphi \in \mathcal{F} \\ \arg\max\limits_{u_\varphi \in \mathbb{F}} W_n^{(\varphi)}(\hat{u}_0^{\varphi-1}, u_\varphi|y_0^{n-1}), & \varphi \notin \mathcal{F}. \end{cases}$$

SC decoding can be defined for any linear code, if an efficient method for computing $W_n^{(\varphi)}(u_0^{\varphi}|y_0^{n-1})$ is available. However, SC decoding can provide reasonable performance only if $G^{(n)}$ is polarizing, i.e. the capacities of bit subchannels $W_n^{(\varphi)}$ converge to 0 or 1 with $n \to \infty$.

C. Convolutional Polar Codes

Convolutional polar codes [1] (CvPCs) are a family of linear block codes, for which $G^{(n)}$, $n = 2^m$, is equal to the matrix of the convolutional polarizing transformation (CvPT)

$$Q^{(n)} = \left(X^{(n)}Q^{(n/2)}, Z^{(n)}Q^{(n/2)}\right), \tag{3}$$

where $Q^{(1)} = (1)$, $X^{(l)}$ and $Z^{(l)}$ are $l \times l/2$ matrices, defined for even l as

$$X^{(l)}_{i,j} = \begin{cases} 1, & \text{if } 2j \le i \le 2j+2 \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

$$Z^{(l)}_{i,j} = \begin{cases} 1, & \text{if } 2j < i \le 2j+2 \\ 0, & \text{otherwise} \end{cases} \tag{5}$$

For example, $X^{(4)} = \begin{pmatrix}1110\\0011\end{pmatrix}^T$, $Z^{(4)} = \begin{pmatrix}0110\\0001\end{pmatrix}^T$. Expansion (3) corresponds to one *layer* of the CvPT. The m-th layer of the CvPT is a mapping of vector u_0^{n-1} to vectors $x_0^{n/2-1} = u_0^{n-1}X^{(n)}$ and $z_0^{n/2-1} = u_0^{n-1}Z^{(n)}$.

III. A Lower Bound on The Minimum Distance of Linear Codes

A. Basic Definitions

Let $\mathbb{S}_n$ be the set of all linear subspaces of $\mathbb{F}^n$.

Denote $a_0^{l-1} \cdot b_0^{l-1} = \sum_{i=0}^{l-1} a_i b_i$, where $a_i, b_i \in \mathbb{F}$. For vectors $b^{(0)}, \dots, b^{(l-1)} \in \mathbb{F}^t$, denote by $\langle b^{(0)}, \dots, b^{(l-1)}\rangle$ the linear subspace of $\mathbb{F}^t$ with basis vectors $b^{(i)}$. For the empty basis we assume $\langle\rangle = \{\mathbf{0}^t\}$, where t is clear from the context. By abuse of notation, we write $x_0x_1\dots x_{t-1}$ for $x_i \in \mathbb{F}$ to denote a vector $(x_0, x_1, \dots, x_{t-1}) \in \mathbb{F}^t$.

Example 1. It can be seen that $\mathbb{S}_2 = \{\langle\rangle, \langle 10\rangle, \langle 01\rangle, \langle 11\rangle, \langle 10, 01\rangle\}$, and $|\mathbb{S}_3| = 16$.

B. Minimum Weight of Cosets and the Minimum Distance

Definition 1. Given an $n \times n$ non-singular matrix $G^{(n)}$, for a vector $p \in \mathbb{F}^j$ define a generalized coset $\mathcal{C}_n^{(\varphi)}(p)$ as

$$\mathcal{C}_n^{(\varphi)}(p) = \left\{u_0^{n-1}G^{(n)} | u_0^{\varphi-1} = \mathbf{0} \ \wedge \ p \cdot u_\varphi^{\varphi+j-1} = 1\right\}. \tag{6}$$

Remark 1. In the case of $j > n - \varphi$, we assume in (6) that $u_l = 0$ for $l \ge n$.

Denote $d_n^{(\varphi)} = \min_{c \in \mathcal{C}_n^{(\varphi)}(1)} \mathbf{wt}(c)$. Observe that for all $j > 0$ one has $\mathcal{C}_n^{(\varphi)}(p) = \mathcal{C}_n^{(\varphi)}(p, \mathbf{0}^j)$, which implies $d_n^{(\varphi)} = \min_{c \in \mathcal{C}_n^{(\varphi)}(1, \mathbf{0}^j)} \mathbf{wt}(c)$. Note that throughout the paper we write $\mathcal{C}_n^{(\varphi)}(p, \mathbf{0}^j)$ instead of formally correct $\mathcal{C}_n^{(\varphi)}((p, \mathbf{0}^j))$, meaning that p and $\mathbf{0}^j$ are concatenated into one argument vector of size $j+1$.

Lemma 1. *If a linear code with minimum distance d is generated by rows of $G^{(n)}$ with indices from $\mathcal{I} \subseteq [n]$, then*

$$d \ge \min_{\varphi \in \mathcal{I}} d_n^{(\varphi)}. \tag{7}$$

Proof. Consider a minimum-weight codeword $c_0^{n-1} = u_0^{n-1}G^{(n)}$, $\mathbf{wt}(c_0^{n-1}) = d$. Let ψ be the first position of non-zero element in u_0^{n-1}. Thus, $\psi \in \mathcal{I}$, $u_\psi = 1$, $u_0^{\psi-1} = \mathbf{0}$, which implies $c_0^{n-1} \in \mathcal{C}_n^{(\psi)}(1)$ and $d = \mathbf{wt}(c_0^{n-1}) \ge d_n^{(\psi)} \ge \min_{\varphi \in \mathcal{I}} d_n^{(\varphi)}$. □

C. Recoverable and Erased Vectors

Consider transmission of a codeword $c_0^{n-1} = u_0^{n-1}G^{(n)}$ of a code with frozen set $\mathcal{F} = [\varphi]$, $u_0^{\varphi-1} = \mathbf{0}$ and dimension $k = n - \varphi$ over a BEC.

The set of erased positions $\mathcal{E} \subseteq [n]$ is called an *erasure configuration*. When erasure configuration $\mathcal{E}$ occurs, the values $c_{\overline{\mathcal{E}}} = u_\varphi^{n-1}\hat{G}$ are available for the receiver, where $\hat{G} = G^{(n)}_{[\varphi],\overline{\mathcal{E}}}$ is the $k \times r$ submatrix of $G^{(n)}$ without rows from $[\varphi]$ and without columns from $\mathcal{E}$, $r = n - |\mathcal{E}|$. Denote by $\mathcal{U}$ the set of all $\hat{u}_\varphi^{n-1}$ such that $\hat{u}_\varphi^{n-1}\hat{G} = c_{\overline{\mathcal{E}}}$. One can see that

$$\mathcal{U} = \left\{u_\varphi^{n-1} + a_0^{k-1} | a_0^{k-1} \in \mathrm{cs}^\perp(\hat{G})\right\}, \tag{8}$$

where for a set of vectors $\mathcal{A} \subseteq \mathbb{F}^t$, by $\mathcal{A}^\perp \subseteq \mathbb{F}^t$ we denote the set of vectors $x_0^{t-1} : \forall y_0^{t-1} \in \mathcal{A} : x_0^{t-1} \cdot y_0^{t-1} = 0$, and $\mathrm{cs}(A)$ is the column space of matrix A. The value u_φ^{n-1} can be unambiguously recovered by the receiver after erasure configuration $\mathcal{E}$ iff $|\mathcal{U}| = 1$, i.e. $\mathcal{U} = \left\{u_\varphi^{n-1}\right\}$.

More generally, consider the recoverability of the value of a linear combination $p_0^{k-1} \cdot u_\varphi^{n-1}$ after erasure configuration $\mathcal{E}$. The set of values of $p_0^{k-1} \cdot \hat{u}_\varphi^{n-1}$ for all $\hat{u}_\varphi^{n-1} \in \mathcal{U}$ is given by

$$\left\{p_0^{k-1} \cdot (u_\varphi^{n-1} + a_0^{k-1}) \mid a_0^{k-1} \in \mathrm{cs}^\perp(\hat{G})\right\}. \tag{9}$$

We say that vector p_0^{k-1} is $(\mathcal{E}, \varphi)$-*recoverable*, if the corresponding linear combination $p_0^{k-1} \cdot u_\varphi^{n-1}$ can be recovered unambiguously for given $c_{\overline{\mathcal{E}}}$, i.e., the set (9) contains only the correct value $p_0^{k-1} \cdot u_\varphi^{n-1}$. Expanding the brackets in (9), one can see that p_0^{k-1} is $(\mathcal{E}, \varphi)$-recoverable iff $\forall a_0^{k-1} \in \mathrm{cs}^\perp(\hat{G}) : p_0^{k-1} \cdot a_0^{k-1} = 0$, which leads to $p_0^{k-1} \in \mathrm{cs}^{\perp\perp}(\hat{G}) = \mathrm{cs}(\hat{G})$. Thus, the set of $(\mathcal{E}, \varphi)$-recoverable vectors is a linear space, which is equal to $\mathrm{cs}(\hat{G}) \in \mathbb{S}_k$.

Definition 2. Let $s \in \mathbb{S}_j$ be the space of all p_0^{j-1}, such that $(p_0^{j-1}, \mathbf{0}^{k-j})$ is $(\mathcal{E}, \varphi)$-recoverable. In this case, s is called the $(\mathcal{E}, \varphi, j)$-*space* and is denoted by $\chi_n^{(\varphi,j)}(\mathcal{E})$, and $\mathcal{E}$ is called an (s, φ, j)-*configuration*. The set of all (s, φ, j)-configurations is denoted by $\xi_n^{(\varphi,j)}(s)$. Thus,

$$\chi_n^{(\varphi,j)}(\mathcal{E}) = \left\{p_0^{j-1} \mid (p_0^{j-1}, \mathbf{0}^{k-j}) \in \mathrm{cs}\left(G^{(n)}_{[\varphi],\overline{\mathcal{E}}}\right)\right\}, \tag{10}$$

$$\xi_n^{(\varphi,j)}(s) = \left\{\mathcal{E} \mid \chi_n^{(\varphi,j)}(\mathcal{E}) = s\right\}. \tag{11}$$

If $\mathcal{A}$ is a set, denote by $2^{\mathcal{A}}$ the set of all subsets of $\mathcal{A}$. Thus, function $\chi_n^{(\varphi,j)} : 2^{[n]} \to \mathbb{S}_j$ maps an erasure configuration, which is a subset of $[n]$, to a linear subspace of $\mathbb{F}^j$, and $\xi_n^{(\varphi,j)}$ returns the inverse image of $\chi_n^{(\varphi,j)}$. Note that $\chi_n^{(\varphi,j)}$ is not injective, so $\xi_n^{(\varphi,j)} : \mathbb{S}_j \to 2^{2^{[n]}}$.

In words, $\chi_n^{(\varphi,j)}(\mathcal{E})$ defines the set of vectors p_0^{j-1}, for which the value of linear combination $p_0^{j-1}\bullet u_\varphi^{\varphi+j-1}$ can be recovered after erasure configuration $\mathcal{E}$, provided that $u_0^{\varphi-1}=\mathbf{0}$. Conversely, $\xi_n^{(\varphi,j)}(s)$ defines the set of erasure configurations, after which the linear combination $p_0^{j-1}\cdot u_\varphi^{\varphi+j-1}$ can be deduced by the receiver if *and only if* $p\in s$.

Remark 2. Let $j>k$, i.e. $j=k+h$ for some $h>0$. In this case, the conditional part of definition (10) is inconsistent. Recall Remark 1, where we assume that symbols u_{n+h} for $h\geq 0$ are equal to zero. Hence, these symbols are always perfectly known for the receiver, so any $\mathcal{E}$ does not erase any symbol u_{n+h}. Observe that any vector from $\mathbb{F}^j\setminus\chi_n^{(\varphi,j)}(\mathcal{E})$ must be *not* $(\mathcal{E},\varphi)$-recoverable, so for any $\mathcal{E}$ and $q_0^{h-1}\in\mathbb{F}^h$, we must include vector $(\mathbf{0}^k,q_0^{h-1})$ in the set $\chi_n^{(\varphi,k+h)}(\mathcal{E})$. This leads to

$$\chi_n^{(\varphi,k+h)}(\mathcal{E})=\left\{(p,q)\mid p\in\chi_n^{(\varphi,k)}(\mathcal{E}),q\in\mathbb{F}^h\right\}.$$

Similarly, we assume that $\xi_n^{(\varphi,k+h)}(s)=\emptyset$ for all s which do not contain $(\mathbf{0}^k,q)$ for some $q\in\mathbb{F}^h$.

Example 2. Consider $(s,0,2)$-configurations for the case of $n=2$, $c_0^1=u_0^1Q^{(2)}=(u_0+u_1,u_1)$. For erasure configuration $\mathcal{E}=\{0\}$, the only non-zero vector which is $(\mathcal{E},0)$-recoverable is $p=(0,1)$. That is, if symbol c_0 is erased, one can recover unambiguously only $u_1=c_1$. This means that $\{0\}\in\xi_2^{(0,2)}(\langle 01\rangle)$. All $(s,0,2)$-configurations are

$$\xi_2^{(0,2)}(\langle 01\rangle)=\{\{0\}\},\xi_2^{(0,2)}(\langle 10\rangle)=\emptyset,\xi_2^{(0,2)}(\langle 11\rangle)=\{\{1\}\},$$
$$\xi_2^{(0,2)}(\langle\rangle)=\{\{0,1\}\},\xi_2^{(0,2)}(\mathbb{F}^2)=\{\emptyset\}. \tag{12}$$

That is, there are no erasure configurations, such that only $\langle 10\rangle$ (i.e. symbol u_0) is unambiguously recoverable, and the whole vector u_0^1 can be unambiguously recovered only if there are no erasures. For the same case, the $(\mathcal{E},0,2)$-spaces are

$$\chi_2^{(0,2)}(\emptyset)=\mathbb{F}^2,\chi_2^{(0,2)}(\{0\})=\langle 01\rangle,$$
$$\chi_2^{(0,2)}(\{1\})=\langle 11\rangle,\chi_2^{(0,2)}(\{0,1\})=\langle\rangle.$$

Example 3. Consider the case of $\varphi=2$, $j=2$, $n=4$ and $c_0^3=u_0^3Q^{(4)}=(u_0+u_1+u_3,u_2+u_3,u_1+u_2+u_3,u_3)$. Since $\varphi=2$ implies $u_0^1=\mathbf{0}$, one has $c_0=c_3=u_3$, $c_1=c_2=u_2+u_3$. One can restore u_3 if c_0 or c_3 is not erased, and one cannot restore any other non-zero linear combination of u_0^3 only if both c_1 and c_2 are erased. Thus, $\xi_4^{(2,2)}(\langle 01\rangle)=\{\{1,2\},\{0,1,2\},\{1,2,3\}\}$.

D. Coset Minimum Weight and Erasure Configurations

For a subspace $s\in\mathbb{S}_j$, we denote the minimal cardinality of (s,φ,j)-configuration by

$$\delta_n^{(\varphi,j)}(s)=\min_{\mathcal{E}\in\xi_n^{(\varphi,j)}(s)}|\mathcal{E}|, \tag{13}$$

assuming that the minimum over the empty set is $+\infty$.

Theorem 1. *Let $\varphi\in[n]$ and $j>0$. For any $p\in\mathbb{F}^j$,*

$$\min_{c\in\mathcal{C}_n^{(\varphi)}(p)}\mathbf{wt}(c)=\min_{s\in\mathbb{S}_j:p\notin s}\delta_n^{(\varphi,j)}(s).$$

Proof. Denote $\mathcal{A}=\left\{\mathrm{supp}(c)\middle|c\in\mathcal{C}_n^{(\varphi)}(p)\right\}$,

$$\mathcal{B}=\bigcup_{s\in\mathbb{S}_j:p\notin s}\xi_n^{(\varphi,j)}(s)=\bigcup_{s\in\mathbb{S}_j:p\notin s}\left\{\mathcal{E}\mid\chi_n^{(\varphi,j)}(\mathcal{E})=s\right\}$$
$$=\left\{\mathcal{E}\mid p\notin\chi_n^{(\varphi,j)}(\mathcal{E})\right\}.$$

Then the theorem can be reformulated as $\min_{\Omega\in\mathcal{A}}|\Omega|=\min_{\mathcal{E}\in\mathcal{B}}|\mathcal{E}|$.

If $\Omega\in\mathcal{A}$, then there exists u_φ^{n-1}, such that $p\bullet u_\varphi^{\varphi+j-1}=1$ and $\Omega=\mathrm{supp}(c_0^{n-1})$ for $c_0^{n-1}=(\mathbf{0}^\varphi,u_\varphi^{n-1})G^{(n)}$. In this case $c_{\overline{\Omega}}=\mathbf{0}$ and the all-zero value $\hat{u}_\varphi^{n-1}=\mathbf{0}$ also belongs to set (8) of possible values of u_φ^{n-1} for the given $c_{\overline{\Omega}}$, but $p\bullet\hat{u}_\varphi^{\varphi+j-1}=0$. Thus, the value of $p\bullet u_\varphi^{\varphi+j-1}$ is not recoverable after erasure configuration Ω, which implies $p\notin\chi_n^{(\varphi,j)}(\Omega)$ and, hence, $\Omega\in\mathcal{B}$. So, $\Omega\in\mathcal{A}\implies\Omega\in\mathcal{B}$ and $\min\limits_{\Omega\in\mathcal{A}}|\Omega|\geq\min\limits_{\mathcal{E}\in\mathcal{B}}|\mathcal{E}|$.

If $\mathcal{E}\in\mathcal{B}$, then $p\notin\chi_n^{(\varphi,j)}(\mathcal{E})$, which by Definition 2 implies $(p,\mathbf{0}^{k-j})\notin\mathrm{cs}(\hat{G})$ and $\exists a_0^{k-1}\in\mathrm{cs}^\perp(\hat{G}):(p,\mathbf{0}^{k-j})\cdot a_0^{k-1}=1$, which implies $p\bullet a_0^{j-1}=1$. Denote $\hat{c}_0^{n-1}=(\mathbf{0}^\varphi,a_0^{k-1})G^{(n)}$. Since $p\bullet a_0^{j-1}=1$, by Definition 1 one has $\hat{c}_0^{n-1}\in\mathcal{C}_n^{(\varphi)}(p)$, and therefore $\mathrm{supp}(\hat{c})\in\mathcal{A}$. On the other hand, $\hat{c}_{\overline{\mathcal{E}}}=a_0^{k-1}\hat{G}=\mathbf{0}$, which means $\mathrm{supp}(\hat{c})\subseteq\mathcal{E}$. So, $\forall\mathcal{E}\in\mathcal{B}\ \exists\Omega\in\mathcal{A}:\Omega\subseteq\mathcal{E}$, hence, $\min\limits_{\Omega\in\mathcal{A}}|\Omega|\leq\min\limits_{\mathcal{E}\in\mathcal{B}}|\mathcal{E}|$. □

Corollary 1. *For any $j>0$:*

$$d_n^{(\varphi)}=\min\left\{\delta_n^{(\varphi,j)}(s)\middle|s\in\mathbb{S}_j:(1,\mathbf{0}^{j-1})\notin s\right\}.$$

IV. Bound on the Minimum Distance of Convolutional Polar Codes

A. Computing the Minimum Coset Weight

The structure of the CvPT $Q^{(n)}$, $n=2^m$, enables one to compute easily $\delta_n^{(\varphi,j)}(s)$, defined in (13), for $j=3$. By computing values of $\delta_n^{(\varphi,3)}(s)$, one can obtain values of $d_n^{(\varphi)}$ by Corollary 1 and a lower bound on the minimum distance by Lemma 1.

Consider transmission of $c_0^{n-1}=u_0^{n-1}Q^{(n)}$, such that $u_0^{\varphi-1}=\mathbf{0}$, through a BEC and let the erasure configuration be $\mathcal{E}$. The intuition behind recursive computing of $\delta_n^{(\varphi,3)}(s)$ is as follows.

Consider the case of $\varphi=2\psi+1<n-1$. Denote $x_0^{n/2-1}=u_0^{n-1}X^{(n)}$, $z_0^{n/2-1}=u_0^{n-1}Z^{(n)}$, $\mathcal{E}'=\mathcal{E}\cap[\frac{n}{2}]$, $\mathcal{E}''=\left\{i\geq 0|i+\frac{n}{2}\in\mathcal{E}\right\}$. Recall that $\chi_n^{(2\psi+1,3)}(\mathcal{E})$ is the set of all p_0^2, such that the value of $p_0^2\bullet u_{2\psi+1}^{2\psi+3}$ can be deduced from c_0^{n-1} after erasure configuration $\mathcal{E}$. Similarly, $\chi_{n/2}^{(\psi,3)}(\mathcal{E}')$ and $\chi_{n/2}^{(\psi,3)}(\mathcal{E}'')$ are the sets of q_0^2 and r_0^2, s.t. $q_0^2\bullet x_\psi^{\psi+2}$ and $r_0^2\bullet z_\psi^{\psi+2}$ are recoverable from $c_0^{n/2-1}$ and $c_{n/2}^{n-1}$ after erasure configurations $\mathcal{E}'$ and $\mathcal{E}''$, under assumption $x_0^{\psi-1}=\mathbf{0}$ and $z_0^{\psi-1}=\mathbf{0}$, respectively. By (4)–(5) one obtains $x_i=u_{2i}+u_{2i+1}+u_{2i+2}$ and $z_i=u_{2i+1}+u_{2i+2}$ for $i<\frac{n}{2}-1$, which, together with $u_0^{2\psi}=\mathbf{0}$, implies $x_0^{\psi-1}=z_0^{\psi-1}=\mathbf{0}$, so the above assumption holds. Furthermore, since u_0^{n-1} was processed by the m-th layer of the CvPT before the transmission, the value

of elements of $u_{2\psi+1}^{2\psi+3}$, as well as the value of any linear combination $p_0^2 \cdot u_{2\psi+1}^{2\psi+3}$, can be deduced only from known linear combinations of elements of x_ψ^{n-1} and z_ψ^{n-1}. However, for any $x_{\psi+3}^{n/2-1}$, $z_{\psi+3}^{n/2-1}$, and $u_{2\psi+1}^{2\psi+3}$, one can find $u_{2\psi+4}^{n-1}$, such that $(\mathbf{0}^{2\psi+1}, u_{2\psi+1}^{n-1}) = \left(\mathbf{0}^\psi, x_\psi^{n/2-1}, \mathbf{0}^\psi, z_\psi^{n/2-1}\right) Q^{(n)}$ as follows: set u_{2i+2} to $x_{i+1} + z_{i+1}$ for $i = \frac{n}{2}-2, \dots, \psi+1$, set u_{n-1} to $z_{n/2-1}$, and set u_{2i+1} to $z_i + u_{2i+2}$ for $i = \frac{n}{2}-2, \dots, \psi+2$. So, for any $p \in \mathbb{F}^3$, even complete knowledge of $x_{\psi+3}^{n/2-1}$ and $z_{\psi+3}^{n/2-1}$ does not provide the value $p \cdot u_{2\psi+1}^{2\psi+3}$. Thus, recoverable linear combinations $q_0^2 \cdot x_\psi^{\psi+2}$ and $r_0^2 \cdot z_\psi^{\psi+2}$ contain all information about recoverable linear combinations $p_0^2 \cdot u_{2\psi+1}^{2\psi+3}$, and therefore $\chi_n^{(2\psi+1,3)}(\mathcal{E})$ can be uniquely deduced from given $\chi_{n/2}^{(\psi,3)}(\mathcal{E}')$ and $\chi_{n/2}^{(\psi,3)}(\mathcal{E}'')$. The similar consideration for $\varphi = 2\psi+2$ leads to the fact that $\chi_n^{(2\psi+2,3)}(\mathcal{E})$ can also be deduced from $\chi_{n/2}^{(\psi,3)}(\mathcal{E}')$ and $\chi_{n/2}^{(\psi,3)}(\mathcal{E}'')$.

Let $\psi = \left\lfloor \frac{\varphi-1}{2} \right\rfloor$, $\mathbb{S}_3 = \{\mathcal{T}_i\}_{i=0}^{15}$. For any $l \in [16]$, consider $(\mathcal{T}_l, \varphi, 3)$-erasure configuration $\mathcal{E}$ for which the minimum in (13) is achieved, i.e. $\chi_n^{(\varphi,3)}(\mathcal{E}) = \mathcal{T}_l$ and $|\mathcal{E}| = \delta_n^{(\varphi,3)}(\mathcal{T}_l)$. Obviously, $|\mathcal{E}| = |\mathcal{E}'| + |\mathcal{E}''|$. Let $\chi_{n/2}^{(\psi,3)}(\mathcal{E}') = \mathcal{T}_i$, $\chi_{n/2}^{(\psi,3)}(\mathcal{E}'') = \mathcal{T}_j$. Then, $\mathcal{E}'$ and $\mathcal{E}''$ are also minimum-weight $(\mathcal{T}_i, \psi, 3)$- and $(\mathcal{T}_j, \psi, 3)$- erasure configurations, respectively, i.e. $|\mathcal{E}'| = \delta_{n/2}^{(\psi,3)}(\mathcal{T}_i)$, and $|\mathcal{E}''| = \delta_{n/2}^{(\psi,3)}(\mathcal{T}_j)$. We know that $\mathcal{T}_l$ can be deduced from $\mathcal{T}_i$ and $\mathcal{T}_j$, i.e., for each φ and n there is a function $\mathbf{T}_n^{(\varphi)}(i,j)$, which returns $\mathcal{T}_l$ for given i and j, and for considered minimum-weight $\mathcal{E}$, $\mathcal{E}'$, $\mathcal{E}''$ one can obtain $\delta_n^{(\varphi,3)}(\mathbf{T}_n^{(\varphi)}(i,j)) = \delta_{n/2}^{(\psi,3)}(\mathcal{T}_i) + \delta_{n/2}^{(\psi,3)}(\mathcal{T}_j)$.

It appears that $\mathbf{T}_n^{(\varphi)} = \mathbf{T}_{n'}^{(\varphi')}$ if $\varphi \equiv \varphi' \mod 2$, i.e., there are only two different functions $\mathbf{T}_n^{(\varphi)}$: one for odd φ and another one for even φ. They are defined as $\mathbf{T}_o, \mathbf{T}_e : [16] \times [16] \to \mathbb{S}_3$, such that

$$\begin{aligned}\mathbf{T}_o(i,j) =& \{p_0^2 | \exists p' \in \mathcal{T}_i, p'' \in \mathcal{T}_j : \\ & (p_0^2, 0, 0)^T = X_{\overline{[1]},*}^{(6)} p'^T + Z_{\overline{[1]},*}^{(6)} p''^T\} \qquad (14)\\ \mathbf{T}_e(i,j) =& \{p_0^2 | \exists p' \in \mathcal{T}_i, p'' \in \mathcal{T}_j : \\ & (p_0^2, 0)^T = X_{\overline{[2]},*}^{(6)} p'^T + Z_{\overline{[2]},*}^{(6)} p''^T\}. \qquad (15)\end{aligned}$$

The above consideration form the following theorem.

Theorem 2. *Denote $\Delta_{n,l}^{(\varphi)} = \delta_n^{(\varphi,3)}(\mathcal{T}_l)$ for $l \in [16]$, $n = 2^m$. Then, for a CvPT, for $0 \le \psi < \frac{n}{2}$:*

$$\Delta_{n,l}^{(2\psi+1)} = \min_{i,j} \left\{ \Delta_{n/2,i}^{(\psi)} + \Delta_{n/2,j}^{(\psi)} \middle| \mathbf{T}_o(i,j) = \mathcal{T}_l \right\}, \qquad (16)$$

$$\Delta_{n,l}^{(2\psi)} = \min_{i,j} \left\{ \Delta_{n/2,i}^{(\psi-1)} + \Delta_{n/2,j}^{(\psi-1)} \middle| \mathbf{T}_e(i,j) = \mathcal{T}_l \right\}. \qquad (17)$$

The base of the recursion is

$$\delta_1^{(0,1)}(\langle\rangle) = 1,\ \delta_1^{(0,1)}(\langle 1\rangle) = 0. \qquad (18)$$

Remark 3. Note that (16)–(17) include the cases of $\Delta_{n,l}^{(n-2)} = \delta_n^{(n-2,3)}(\mathcal{T}_l)$ and $\Delta_{n,l}^{(n-1)} = \delta_n^{(n-1,3)}(\mathcal{T}_l)$. They can be obtained according to Remark 2 as follows. For $s \in \mathbb{S}_{i+h}$, denote $s|_i = \left\{ p_0^{i-1} \mid p_0^{i+h-1} \in s \right\}$. Since any erasure configuration does not erase u_{n-1+h} for any $h > 0$, one obtains

$$\delta_n^{(n-i,i+h)}(s) = \begin{cases} \delta_n^{(n-i,i)}(s|_i), & \text{if } \forall p \in \mathbb{F}^h : (\mathbf{0}^i, p) \in s \\ +\infty, & \text{otherwise} \end{cases}$$

The same assumption is applied for deriving $\Delta_{1,l}^{(0)} = \delta_1^{(0,3)}(\mathcal{T}_l)$ from the values $\delta_1^{(0,1)}(s)$ for $s \in \mathbb{S}_1$, which are given by the base (18) of the recursion.

Remark 4. Formula (17) in the case of $\Delta_{n,l}^{(0)}$ leads to computing $\Delta_{n/2,i}^{(-1)} = \delta_{n/2}^{(-1,3)}(\mathcal{T}_i)$, which is equal to the minimum number of erased symbols which erases value $p \cdot u_{-1}^2$ iff $p \in \mathcal{T}_i$. For symbol u_{-i}, $i > 0$, we do not employ the same assumption as in Remark 3. If one assumes that symbols with negative indices are always known and employs functions $\mathbf{T}_o$ and $\mathbf{T}_e$, one would obtain that input symbols on the current layer of convolutional polarizing transformation u_{-2}, u_{-1}, and input symbol $x_{-1} = u_{-2} + u_{-1} + u_0$ on the next layer are always known, which implies that u_0 is always known. This would result in incorrect value of $\Delta_{n,l}^{(0)}$. Thus, we assume that u_{-i} for $i > 0$ are always erased, which leads to

$$\chi_n^{(-i,j)}(\mathcal{E}) = \left\{ (\mathbf{0}^i, p) \mid p \in \chi_n^{(0,j-i)}(\mathcal{E}) \right\}, 0 < i \le j.$$

Proof. For erasure configuration $\mathcal{E} \subseteq [n]$, denote $\mathcal{E}' = \mathcal{E} \cap [n/2]$ and $\mathcal{E}'' = \{j - n/2 \mid j \in \mathcal{E} \setminus [n/2]\}$. We now consider the case of $\varphi = 2\psi+1$ and prove (16).

Note that $u_0^{2\psi} = \mathbf{0}^{2\psi+1}$ implies $x_0^{\psi-1} = z_0^{\psi-1} = \mathbf{0}^\psi$. By (3) one obtains $\hat{Q} = \left(\hat{X}\hat{Q}',\ \hat{Z}\hat{Q}''\right)$, where $\hat{Q} = Q_{\overline{[2\psi+1]},\overline{\mathcal{E}}}^{(n)}$, $\hat{Q}' = Q_{\overline{[\psi]},\overline{\mathcal{E}'}}^{(n/2)}$, $\hat{Q}'' = Q_{\overline{[\psi]},\overline{\mathcal{E}''}}^{(n/2)}$, $\hat{X} = X_{\overline{[2\psi+1]},\overline{[\psi]}}^{(n)}$, $\hat{Z} = Z_{\overline{[2\psi+1]},\overline{[\psi]}}^{(n)}$. By (10), $p_0^2 \in \chi_n^{(\varphi,3)}(\mathcal{E})$ iff there exists q:

$$(p_0^2, \mathbf{0}^{k-3})^T = \hat{Q}q^T = (\hat{X}\hat{Q}', \hat{Z}\hat{Q}'')q^T = \hat{X}\hat{Q}'q'^T + \hat{Z}\hat{Q}''q''^T,$$

where $q = (q', q'')$, $k = n - \varphi$, which implies, in particular,

$$(\hat{X}\hat{Q}'q'^T)_{\overline{[3]}} = (\hat{Z}\hat{Q}''q''^T)_{\overline{[3]}}. \qquad (19)$$

Denote $a = q'\hat{Q}'^T$, $b = q''\hat{Q}''^T$. Thus, $a \in \mathrm{cs}(\hat{Q}')$, $b \in \mathrm{cs}(\hat{Q}'')$. Then (19) implies $a\hat{X}^T_{\overline{[3]},*} = b\hat{Z}^T_{\overline{[3]},*}$, so from (4)–(5) one obtains the system of equations

$$\begin{cases} a_i + a_{i+1} = b_i, & i = 1 \dots n/2 - \psi - 2 \\ a_i = b_i, & i = 2 \dots n/2 - \psi - 1 \end{cases} \qquad (20)$$

It is easy to see that (20) implies $a_i = b_i = 0$ for $i \ge 3$. Let $k' = n/2 - \psi$. By the above consideration, for any $p \in \mathbb{F}^3$ one has $(p, \mathbf{0}^{k-3}) \in \mathrm{cs}(\hat{Q})$ iff there exists $p', p'' \in \mathbb{F}^3$, s.t. $(p', \mathbf{0}^{k'-3}) \in \mathrm{cs}(\hat{Q}')$, $(p'', \mathbf{0}^{k'-3}) \in \mathrm{cs}(\hat{Q}'')$, and

$$(p, \mathbf{0}^3) = p' \begin{pmatrix} 110000 \\ 011100 \\ 000111 \end{pmatrix} + p'' \begin{pmatrix} 110000 \\ 001100 \\ 000011 \end{pmatrix}. \qquad (21)$$

Note that two last elements of vector in the left-hand side equals 0, and two last columns in the right hand size of (21) are identical, so these columns can be removed. The resulting

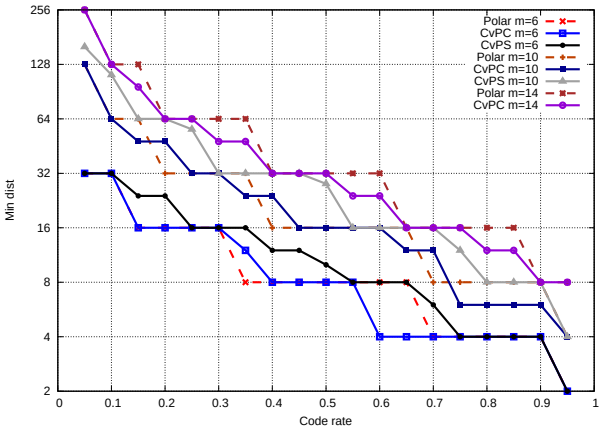

Fig. 1: Minimum distance of polar codes, CvPCs and CvPSs

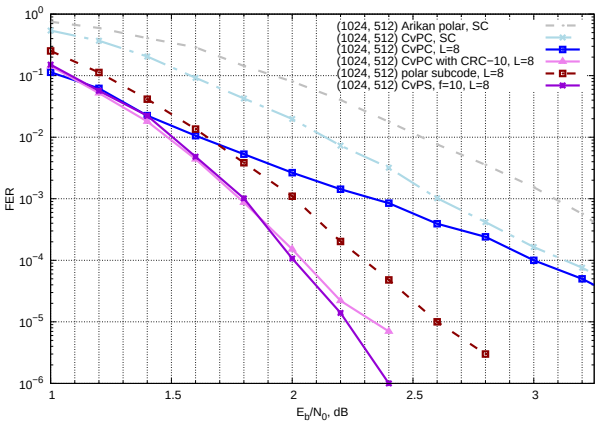

Fig. 2: Performance of $(1024, 512)$ CvPS with $|\mathcal{D}| = 10$

matrices are equal to transposed $X^{(6)}_{\overline{[1]},*}$ and $Z^{(6)}_{\overline{[1]},*}$, respectively. Recalling (10), one obtains that $\chi_n^{(2\psi+1,3)}(\mathcal{E})$ consists of all p_0^2, for which there exist $p' \in \chi_{n/2}^{(\psi,3)}(\mathcal{E}')$, $p'' \in \chi_{n/2}^{(\psi,3)}(\mathcal{E}'')$:

$$(p_0^2, \mathbf{0}^2)^T = X^{(6)}_{\overline{[1]},*} p'^T + Z^{(6)}_{\overline{[1]},*} p''^T. \tag{22}$$

Observe that (22) is equivalent to the equation in the right-hand side of (14). Obviously, $|\mathcal{E}| = |\mathcal{E}'|+|\mathcal{E}''|$ and the minimal cardinality of $(\mathcal{T}_l, 2\psi+1, 3)$-configuration $|\mathcal{E}|$ for each $\mathcal{T}_l \in \mathbb{S}_3$ can be found exactly as it is stated in (16).

Equality (17) can be proved similarly. □

In Fig. 1 the lower bound on minimum distance, computed by (7), for CvPCs of lengths 2^6, 2^{10}, 2^{14} is presented. The codes are obtained via the Monte-Carlo method by minimization of the E_b/N_0 needed to achieve the SC decoding error probability 10^{-3}. For comparison, we also report the results for Arikan polar codes, which are optimized in the same way.

By employing the low-weight codeword search algorithm presented in [3], we verified that the bound is exact for CvPCs with $m = 5, \ldots, 13$, rates $\frac{1}{20}, \ldots \frac{19}{20}$ and target FER of SC decoding 10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}, and 10^{-6}.

B. Convolutional Polar Subcodes

By Lemma 1, any codeword $c_0^{n-1} = u_0^{n-1} G^{(n)}$ of weight d corresponds to vector u_0^{n-1} with at least one symbol $u_i = 1, i \in \mathcal{I}$, such that $d_n^{(i)} \leq d$. In the case of polar codes, $d_n^{(i)}$ is equal to the weight of the i-th row of $A^{(n)}$. In the case of CvPCs one can obtain $d_n^{(i)}$ by Theorem 2.

Polar subcodes, which has low list SC decoding [8] error probability, was proposed in [4]. Polar subcodes are obtained as generalization of polar codes with some symbols $u_\varphi, \varphi \in \mathcal{D}$, called dynamic frozen symbols, not set to zero, but set to linear combinations of previous symbols $u_i, i < \varphi$. Set $\mathcal{D}$ consists of indices i of unfrozen subchannels of minimum weight $d_n^{(i)}$. With the knowledge of $d_n^{(i)}$ for CvPT, this approach can be immediately extended to the case of CvPT. The detailed description of the construction is presented in the extended version of the paper [9].

Fig. 2 presents the performance of a $(1024, 512)$ CvPS with $|\mathcal{D}| = 10$, a polar code and a randomized polar subcode [4] in AWGN channel. The polar code and the polar subcode are constructed for AWGN channel with $E_b/N_0 = 2$ dB using Gaussian approximation of density evolution [10], and the CvPC and the CvPS are constructed for the same channel using Monte-Carlo simulations for subchannels qualities. One can see that the CvPS outperforms the polar subcode, the CvPC and the CvPC concatenated with CRC-10. More detailed complexity and performance comparison can be found in [9].

V. Conclusions

In this paper a method for evaluation coset minimum weight of convolutional polar codes is provided, and tight lower bound on minimum distance of convolutional polar codes is derived. The results are used for construction of convolutional polar subcodes, which outperform CvPCs under list SC decoding.

References

[1] A. J. Ferris, C. Hirche, and D. Poulin, "Convolutional polar codes," *CoRR*, vol. abs/1704.00715, 2017. Available: http://arxiv.org/abs/1704.00715

[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, Jul. 2009.

[3] A. Canteaut and F. Chabaud, "A new algorithm for finding minimum-weight words in a linear code: Application to McEliece's cryptosystem and to narrow-sense BCH codes of length 511," *IEEE Transactions on Information Theory*, vol. 44, no. 1, pp. 367–378, Jan. 1998.

[4] P. Trifonov and G. Trofimiuk, "A randomized construction of polar subcodes," in *IEEE International Symposium on Information Theory*, Jun. 2017, pp. 1863–1867.

[5] R. Morozov and P. Trifonov, "Efficient SC decoding of convolutional polar codes," in *International Symposium on Information Theory and its Applications*, Oct. 2018, pp. 442–446.

[6] T. Prinz and P. Yuan, "Successive cancellation list decoding of BMERA codes with application to higher-order modulation," in *International Symp. Turbo Codes and Iterative Inf. Processing*, Dec. 2018.

[7] H. Saber, Y. Ge, R. Zhang, W. Shi, and W. Tong, "Convolutional polar codes: LLR-based successive cancellation decoder and list decoding performance," in *International Symposium on Information Theory*, Jun. 2018, pp. 1480–1484.

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

[9] R. Morozov and P. Trifonov, "On distance properties of convolutional polar codes," *IEEE Transactions On Communications, accepted*.

[10] P. Trifonov, "Efficient design and decoding of polar codes," *IEEE Trans. on Communications*, vol. 60, no. 11, pp. 3221 – 3227, Nov. 2012.