Randomized Polar Subcodes with Optimized Error Coefficient

Peter Trifonov, *Member, IEEE*

***Abstract*—A method for construction of polar subcodes with reduced error coefficient is presented. The proposed approach relies on explicit enumeration of low-weight non-zero codewords in a polar code, and construction of dynamic freezing constraints which define a subcode not containing most of these codewords. The obtained codes provide a large performance gain in the high-SNR region compared to non-optimized polar subcodes and polar codes with CRC.**

***Index Terms*—Polar codes, polar subcodes, weight distribution.**

I. Introduction

Polar codes is a novel class of capacity-achieving codes, having low-complexity construction, encoding and decoding algorithms [1]. Classical polar codes have rather low minimum distance, and the successive cancellation (SC) decoding algorithm is highly suboptimal for finite length codes. Therefore, improved code constructions were suggested, such as polar codes with CRC [2], [3] and polar subcodes [4], [5], together with improved decoding algorithms [2], [6]. These constructions were shown to provide excellent performance for block length up to a few thousands bits, although for longer block length they are outperformed by LDPC codes. As a result, polar codes were adopted in 5G for use in the control channel only [7], [8].

Very high scaling exponent of Arıkan's polar codes [9] does not allow them to compete well at large block length with other modern code constructions, such as LDPC codes. Polar codes with large kernels [10] were shown to asymptotically achieve optimal scaling exponent [11]. In general, the decoding complexity of such codes is very high. However, low-complexity decoding algorithms were derived for some specific kernels [12], [13], [14]. Together with the successive cancellation list (SCL) decoding method, these techniques allow decoding with much lower complexity compared to codes based on the Arıkan's kernel with the same performance.

Next generation machine-to-machine communication requires very low block error rate (e.g. less than 10^{-5}). This requirement must be taken into account in the design of the corresponding error correcting codes. Unfortunately, there are still no rigorous theoretical tools for performance evaluation of the SCL decoding algorithm. Simulations suggest that the performance of polar codes and their improved versions under SCL decoding depends both on their SC and ML decoding error probability. Increasing list size allows the performance of the SCL decoder to approach the performance of the ML decoder. At high SNR, the ML decoding error probability depends on the number of low weight non-zero (LWNZ) codewords in the code and, in particular, its error coefficient. It was shown in [15] that the scaling exponent, which is known to be quite high for Arikan's polar codes, does not improve when adding a finite list to the MAP decoder. This means that the construction of polar codes needs to be modified somehow to obtain better performance.

In this paper we present a method for construction of polar subcodes with reduced error coefficient, and show that the obtained codes provide better performance under SCL and sequential decoding compared to polar codes with CRC, randomized and algebraic polar subcodes. The main results of the paper include:

- Analysis of minimum distance of polar codes with generic kernels (Theorem 1).
- Analysis of the impact of type-B dynamic frozen symbols on the minimum distance of polar codes (Theorem 2).
- Extension of the construction of randomized polar subcodes to arbitrary kernels.

The paper is organized as follows. Section II presents an overview of polar codes and polar subcodes. The proposed code construction method is derived in Section III. Simulation results are presented in Section IV.

II. Background

A. Channel polarization

Let F_l be an invertible $l \times l$ matrix not permutation-equivalent to an upper triangular matrix. $(n = l^m, k)$ polar code with kernel F_l is a set of vectors $\mathbf{c} = c_0^{n-1} = u_0^{n-1}A$, where $A = F_l^{\otimes m}$, $u_i = 0, i \in \mathcal{F}$, $\mathcal{F} \subset [n]$ is the set of frozen symbol indices (frozen set), $m \in \mathbb{N}$, and $[n] = \{0, \dots, n-1\}$. Let $\mathbf{U}_i, \mathbf{C}_i, \mathbf{Y}_i$ be random variables corresponding to input values of the polarizing transformation, channel input and output symbols, respectively. Matrix A together with a binary-input memoryless channel $W(\mathbf{Y}|\mathbf{C})$ gives rise to synthetic bit subchannels

$$W_m^{(i)}(\mathbf{Y}_0^{n-1}, \mathbf{U}_0^{i-1}|\mathbf{U}_i) = \sum_{U_{i+1}^{n-1} \in \mathbb{F}_2^{n-i-1}} \frac{W(\mathbf{Y}_0^{n-1}|\mathbf{U}_0^{n-1}A)}{2^{n-1}},$$

where $W(\mathbf{Y}_0^{n-1}|\mathbf{C}_0^{n-1}) = \prod_{j=0}^{n-1} W(\mathbf{Y}_j|\mathbf{C}_j)$. The capacities C_i of these subchannels converge with $m \to \infty$ to 0 or 1, and the fraction of almost noiseless subchannels converges to the capacity C of the original channel W [10].

Efficient techniques for evaluation of the reliabilities of these subchannels were suggested in [16], [17] for the case

This work was supported by Government of Russian Federation (grant 08-08).

This work was partially presented at IEEE International Symposium on Information Theory 2019.

P. Trifonov is with ITMO University, Russia. E-mail: pvtrifonov@itmo.ru.

of the Arıkan's kernel $F_2 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$, and in [18], [19], [20] for larger kernels. For a *classical* polar code, the frozen set $\mathcal{F}$ is selected as the set of indices i of $n-k$ least reliable bit subchannels $W_m^{(i)}$, where k is code dimension. That is, assuming that r_j is the sequence of distinct integers, such that $C_{r_0} \leq C_{r_1} \leq \dots \leq C_{r_{n-1}}$, the frozen set is constructed as $\mathcal{F} = \{r_0, \dots, r_{n-k-1}\}$.

The SC decoding error probability of a polar code of rate $R < C$ satisfies $2^{-n^{\beta'}} \leq P_{SC} \leq 2^{-n^{\beta}}$ for sufficiently large n and any $\beta, \beta' : \beta < E(F_l) < \beta'$, where

$$E(F_l) = \frac{1}{l} \sum_{i=0}^{l-1} \log_l D_i$$

is the rate of polarization of kernel F_l, and D_i is the i-th partial distance of F_l, i.e. the minimum distance between the i-th row of F_l and the space generated by rows $i+1, \dots, l-1$ [10].

Several kernels of different dimensions can be combined in the construction of a polarizing transformation [21], i.e. it can be defined as

$$A = F_{l_0} \otimes F_{l_1} \otimes \cdots F_{l_{m-1}}.$$

This results in a polar code of length $n = \prod_{i=0}^{m-1} l_i$. Rate of polarization for such transformation can be computed from $E(F_{l_i})$ as described in [22].

B. Low-weight codewords of polar codes with Arıkan's kernel

Polar codes with Arıkan's kernel were shown to be instances of decreasing monomial codes [23]. That is, their codewords are given by evaluation vectors of Zhegalkin polynomials of certain type. Recall, that Zhegalkin polynomial is a polynomial in variables $x_0, \dots, x_{m-1}$ over the integers modulo 2 [24].

Indeed, the rows of the Arıkan matrix $F_2^{\otimes m}$ can be considered as evaluation vectors of some Zhegalkin monomials in variables $x_0, x_1, \dots, x_{m-1}$ in all points of $\mathbb{F}_2^m$.

Example 1. *For $m = 3$, the Arıkan polarizing matrix is given by*

$$A = \begin{array}{cc} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix} & \begin{matrix} x_0x_1x_2 \\ x_1x_2 \\ x_0x_2 \\ x_2 \\ x_0x_1 \\ x_1 \\ x_0 \\ 1 \end{matrix} \end{array}$$

Assuming that $\mathcal{F} = \{0, 1, 2\}$, one obtains that any codeword $\mathbf{c} = u_0^{n-1}A$ of $(8, 5)$ polar code is given by a vector of values of some polynomial

$$u(X) = u(x_0, x_1, x_2) = u_7 \cdot 1 + u_6 \cdot x_0 + u_5 \cdot x_1 + u_4 \cdot x_0 x_1 + u_3 \cdot x_2$$

in points $(1, 1, 1), \dots, (0, 0, 0)$.

The frozen set $\mathcal{F}$ for Arıkan polar code is known to satisfy the following partial order constraints [25], [23]. For any $i = \sum_{j=0}^{m-1} i_j 2^j \notin \mathcal{F}$, where $i_j \in \{0, 1\}$, one has

$$i_s = 0 \Rightarrow (i + 2^s) \notin \mathcal{F}, 0 \leq s < m \quad (1)$$

$$i_s = 0, i_t = 1, s > t \Rightarrow (i + 2^s - 2^t) \notin \mathcal{F}. \quad (2)$$

This was used in [26] to derive an algorithm for construction of polar codes with sublinear complexity.

Given a polynomial $u(x_0^{m-1})$ corresponding to a codeword of a polar code, define $v(x_0^{m-1}) = u(Bx_0^{m-1} + b)$, where B is some non-singular $m \times m$ matrix, and $b \in \mathbb{F}_2^m$. The evaluation vector of $v(x_0^{m-1})$ is obtained by permuting the evaluation vector of $u(x_0^{m-1})$. However, $v(x_0^{m-1})$ does not necessarily correspond to a valid codeword of the considered polar code.

It was shown in [23] that the automorphism group of a polar code includes lower triangular affine group $LTA(m, 2)$, i.e. $v(x_0^{m-1})$ is guaranteed to correspond to a valid codeword if B is an arbitrary $m \times m$ lower-triangular matrix with 1's on the diagonal, and b is an arbitrary binary vector.

Furthermore, all distinct weight-d codewords of an Arıkan polar code with minimum distance $d = 2^r$, where $r = \min_{i \notin \mathcal{F}} \mathrm{wt}(i)$, can be obtained as evaluation vectors of polynomials $g_i(Bx_0^{m-1} + b), i \in \{0, \dots, 2^m - 1\} \setminus \mathcal{F}$, where matrices B satisfy

$$B_{tj} = 0 \text{ if } t \notin \mathrm{ind}(g_i) \vee j \in \mathrm{ind}(g_i),$$

and vectors b are such that $t \notin \mathrm{ind}(g_i) \Rightarrow b_t = 0$. Here $\mathrm{wt}(i)$ denotes the number of non-zero bits in the base-2 expansion of integer i, $g_i(x_0^{m-1}) = \prod_{j=0}^{m-1} x_j^{1-i_j}$ are the monomials corresponding to the rows of Arıkan matrix $F_2^{\otimes m}$ of weight d. and $\mathrm{ind}(g) = \{t : x_t | g\}$ for some monomial $g = g(x_0^{m-1})$. This implies that any unfrozen symbol i, such that $\mathrm{wt}(i) = r$, induces a number of weight-2^r codewords, such that $u_i = 1, u_j = 0, j < i$.

The total number of such codewords, i.e. the error coefficient of the polar code, is given by [23]

$$w_d = 2^{m-r} \sum_{\substack{i \in [n] \setminus \mathcal{F} \\ \mathrm{wt}(i) = r}} 2^{|\lambda_i|}, \quad (3)$$

where $\lambda_i = \mathrm{ind}(g_i)$ is the list of indices of zero bits in integer i, and

$$|(i_0, \dots, i_{m-r-1})| = \sum_{j=0}^{m-r-1} (i_j - j).$$

C. Polar subcodes

Classical polar codes exhibit quite poor finite-length performance due to their low minimum distance. Substantially better performance can be obtained by employing polar subcodes [4], [27]. The construction of polar subcodes relies on dynamic frozen symbols, i.e. input symbols of a polarizing transformation, which are given by

$$u_{j_i} = \sum_{s < j_i} V_{i,s} u_s, i \in \mathcal{F}, \quad (4)$$

where V is a $(n-k) \times n$ constraint matrix, such that the last non-zero elements of its rows are located in distinct columns $j_i, 0 \leq i < n-k$. The constraint matrix can be obtained either from a check matrix of an algebraic code (e.g. extended BCH) with sufficiently large minimum distance, or constructed by a pragmatic randomized algorithm [5], so that most of the low-weight codewords of the original polar code do not satisfy (4), and are therefore eliminated from the obtained polar subcode.

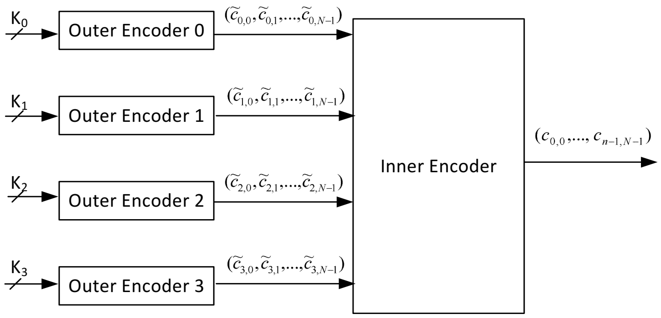


Fig. 1: Generalized concatenated encoder

The constraint matrix of a classical polar code consists of $n-k$ rows of weight 1, where 1's are located in columns corresponding to frozen symbols.

Decoding of polar subcodes can be implemented by the SCL or sequential algorithms [2], [6], [28].

D. Generalized concatenated codes

A generalized concatenated code (GCC) [29], [30] over $\mathbb{F}_q$ is defined using a family of nested inner (n, k_i, d_i) codes $\mathcal{C}^{(i)} : \mathcal{C}^{(0)} \supset \mathcal{C}^{(1)} \supset \dots \supset \mathcal{C}^{(\nu-1)}$, and a family of outer (N, K_i, D_i) codes $\mathbb{C}^{(i)}$, where the i-th outer code is defined over $\mathbb{F}_{q^{k_i-k_{i+1}}}, 0 \leq i < \nu, k_\nu = 0$. In this paper we assume that $k_i = k_{i+1}+1, \nu = n$. Let $\mathcal{G}$ be a $n \times n$ matrix, such that its rows $i, \dots, n-1$ generate code $\mathcal{C}^{(i)}$. GCC encoding is performed as follows. First, partition a data vector into n blocks of size $K_i, 0 \leq i < n$. Second, encode these blocks with codes $\mathbb{C}^{(i)}$ to obtain codewords $(\tilde{c}_{i,0}, \dots, \tilde{c}_{i,N-1})$. Finally, multiply vectors $(\tilde{c}_{0,j}, \dots, \tilde{c}_{n-1,j}), 0 \leq j < N$, by $\mathcal{G}$ to obtain a GCC codeword $(c_{0,0}, \dots, c_{n-1,0}, c_{0,1}, \dots, c_{n-1,N-1})$. This process is illustrated in Figure 1.

A GCC generator matrix can be obtained as

$$G = \begin{pmatrix} G^{(0)} \otimes \mathcal{G}_{0,-} \\ G^{(1)} \otimes \mathcal{G}_{1,-} \\ \vdots \\ G^{(n-1)} \otimes \mathcal{G}_{n-1,-} \end{pmatrix},$$

where $G^{(i)}$ is a generator matrix of $\mathbb{C}^{(i)}$, and $\mathcal{G}_{i,-}$ denotes the i-th row of $\mathcal{G}$. It is possible to show that this encoding method results in a $(Nn, \sum_{i=0}^{n-1} K_i, \geq \min_i d_i D_i)$ linear block code.

III. Polar subcodes with reduced error coefficient

The error probability of the Tal-Vardy list decoding algorithm for polar (sub)codes is given by

$$P(L) = P_{ML} + P(\mathcal{E}(L)|\mathbb{C}), \tag{5}$$

where P_{ML} is the maximum likelihood decoding error probability of the considered code, $\mathbb{C}$ is the event corresponding to the maximum likelihood decoder producing the correct codeword $u_0^{n-1}A_m$, and $\mathcal{E}(L)$ is the event corresponding to the score of the correct vector becoming lower than the scores of L incorrect vectors $\hat{u}_0^i$ at some intermediate phase, so that the correct vector is killed by the decoder. The value of P_{ML} can be estimated via the union bound as

$$P_{ML} \leq \sum_{i=d}^{n} w_i Q\left(\sqrt{2i\frac{E_s}{N_0}}\right) = \sum_{i=d}^{n} w_i Q\left(\sqrt{2iR\frac{E_b}{N_0}}\right),$$

where w_i is the number of codewords of weight i. At high SNR it depends primarily on d, minimum distance of the code, and w_d, the error coefficient. To the best of author knowledge, there are still no techniques for computing $P(\mathcal{E}(L)|\mathbb{C})$. However, experiments show that this quantity increases with P_{SC}, the error probability of the successive cancellation decoder. Hence, obtaining a code with good performance under SCL decoding requires both P_{SC} and P_{ML} to be sufficiently small.

In what follows, we propose a method for construction of a subcode of a polar code, such that most of the low-weight codewords, which appear in the original polar code, are not included into the polar subcode. At the same time, we aim to keep the successive cancellation decoding error probability of the obtained code as low as possible.

A. Low-weight codewords of polar codes

In this section we characterize the information vectors of LWNZ codewords of polar codes obtained from a large class of kernels. Essentially, we show that, for a sufficiently general class of kernels, any such codeword is generated by at least one low-weight row of the polarizing matrix A.

Theorem 1. *Consider an (n,k) polar code $\mathcal{C}$ given by a polarizing transformation $A = F_{l_0} \otimes \dots \otimes F_{l_{m-1}}$ and the set of frozen symbol indices $\mathcal{F}$, where F_{l_i} is an l_i-dimensional kernel, such that its partial distances $D_{i,j}$ satisfy*

$$D_{i,j} = \mathrm{wt}((F_{l_i})_j), 0 \leq j < l_i, 0 \leq i < m, \tag{6}$$

$n = \prod_{i=0}^{m-1} l_i$, and $(F_{l_i})_j$ is the j-th row of F_{l_i}. Then:

1) *The minimum distance of the polar code is $d = \min_{s \notin \mathcal{F}} \mathrm{wt}(A_s)$, where A_s is the s-th row of matrix A, $0 \leq s < n$.*
2) *Any codeword $\mathbf{c} = u_0^{n-1}A$ of weight d, where d is the minimum distance, has $u_s = 1$ for some $s : \mathrm{wt}(A_s) = d$.*

Proof:

It is sufficient to consider the case of $D_{i,j} \leq D_{i,j+1}, 0 \leq j < l_i - 1$, since otherwise the corresponding rows of F_{l_i} can be swapped [10, Proposition 15] together with appropriate modification of the set $\mathcal{F}$, so that the code remains the same. In this case $D_{i,j}$ is the minimum distance of the code generated by rows $j, \dots, l_i - 1$ of F_{l_i}.

Both statements for $m = 1$ follow from (6). Assume that the theorem holds for some $A = \bigotimes_{i=0}^{m-1} F_{l_i}$ with $m \geq 1$, and consider the polarizing transformation $A' = \bigotimes_{i=0}^{m} F_{l_i}$. Then $\mathrm{wt}(A'_{tl_m+j}) = \mathrm{wt}(A_t)\,\mathrm{wt}((F_{l_m})_j)$, where $j \in [l_m], t \in [n]$. Consider a polar code with polarizing transformation A' and the frozen set $\mathcal{F}$. It can be considered as a GCC with outer (l_m, K_t, D_t) codes $\mathbb{C}^{(t)}$ and inner (n, k_t, d_t) codes $\mathcal{C}^{(t)}$, where $\mathbb{C}^{(t)}$ is generated by $F_{l_m,j} : tl_m + j \notin \mathcal{F}$, and $\mathcal{C}^{(t)}$ is a polar code with polarizing transformation A and the set of frozen symbol indices $\mathcal{F}^{(t)} = \{r | \forall j : 0 \leq j < l_m : l_m r + j \in \mathcal{F}, t \leq r < n\}, 0 \leq t < n$.

Then the minimum distance of the GCC is $d \geq d_t D_t$, where $d_t = \min_{j \notin \mathcal{F}^{(t)}} \mathrm{wt}(A_j)$, and $D_t \geq \min_{j: l_m t + j \notin \mathcal{F}} D_{m,j}$. This bound is achieved with equality, since there is a weight-d codeword in $\mathcal{C}$ given by a row $tl_m + j \notin \mathcal{F}$ of A'.

The second statement also holds for any m, since if a codeword of $\mathcal{C}$ has $u_t = 0$ for all $t : \mathrm{wt}(A_t) = d$, then it belongs to a code with the set of frozen symbol indices $\mathcal{F}' = \mathcal{F} \setminus \{t | \mathrm{wt}(A_t) = d\}$, which has minimum distance $d' > d$, according to the first statement. ■

The above theorem allows to design constraints on the input symbols of the polarizing transformation, such that most of the LWNZ codewords of a polar code do not satisfy them, obtaining thus a subcode of the polar code with low error coefficient.

B. Type-A dynamic frozen symbols

To obtain an (n, k) code with good performance under SCL decoding, we need to eliminate low-weight non-zero (LWNZ) codewords from a polar code. This can be done by introducing dynamic freezing constraints (4). To obtain an (n, k, d) code $\mathcal{C}$, we propose to construct first an $(n, k + f_A, \overline{d})$ polar code $\overline{\mathcal{C}}$ (parent code) with constraint matrix $\overline{V}$, and obtain the constraint matrix for code $\mathcal{C}$ as $V = \begin{pmatrix} \overline{V} \\ V^{(A)} \end{pmatrix}$. Here $V^{(A)}$ is a $f_A \times n$ matrix, which defines dynamic freezing constraints, such that most of the LWNZ codewords of the parent code do not satisfy them, so that $\mathcal{C}$ does not contain these codewords. Let us further define $\overline{\mathcal{F}}$ as the set of frozen symbol indices given by $\overline{V}$.

To reduce the probability of the correct path being killed by the SCL decoder at an early phase, these constraints need to be imposed on symbols u_{s_i} with the smallest possible indices s_i in such way, so that low-weight codewords are still eliminated. Theorem 1 suggests that this can be done by extending the construction of [5], i.e. by selecting s_i as f_A maximal indices, such that $\mathrm{wt}(A_{s_i}) = \overline{d}$, where $\overline{d}$ is the minimum distance of the parent polar code $\overline{\mathcal{C}}$, and selecting $V_{i,j}^{(A)}, 0 \leq j < s_i, 0 \leq i < f_A$, as independent random binary values. Furthermore, we set $V_{i,s_i}^{(A)} = 1$ and $V_{i,j}^{(A)} = 0, j > s_i$. Indeed, such choice of s_i ensures that for any LWNZ codeword $\mathbf{c} = u_0^{n-1} A \in \overline{\mathcal{C}}$ it is possible to find such values $V_{i,j}^{(A)}$, so that u_0^{n-1} does not satisfy (4), i.e. $\mathbf{c} \notin \mathcal{C}$. For small values of f_A it may not be possible to eliminate all LWNZ codewords in this way, but the results given in [5] suggest that even for randomized selection of $V^{(A)}$ the obtained codes provide quite good performance compared to polar codes with CRC and LDPC codes.

However, more careful design is possible. We propose to explicitly enumerate (almost) all non-zero codewords $c^{(p)}, 0 \leq p < P$, of the parent code with weight up to $\Delta \geq d$. For polar codes with Arıkan's kernel and $\Delta = d$, this can be done as described in Section II-B. For codes based on other kernels, LWNZ codewords can be obtained with the algorithm presented in [31]. Let $u^{(p)} = c^{(p)} A^{-1}$ be the information vectors corresponding to the obtained LWNZ codewords, and let $\mathcal{P} = \left\{ c^{(p)} \right\}$. We say that codeword $c^{(p)}$ is pruned by constraint i if

$$u_{s_i}^{(p)} \neq \sum_{t < s_i} V_{i,t}^{(A)} u_t^{(p)}. \tag{7}$$

Let $\mathcal{P}(V_i^{(A)})$ be the set of codewords pruned from $\mathcal{P}$ by the constraint given by the i-th row of $V^{(A)}$. We propose to select for each i the coefficients $V_{i,j}^{(A)}$ so that $|\mathcal{P}(V_i^{(A)})|$ is maximized. This can be implemented by generating randomly T vectors of coefficients $(V_{i,0}^{(A)}, \ldots, V_{i,n-1}^{(A)})$, such that $V_{i,s_i}^{(A)} = 1$ and $V_{i,t}^{(A)} = 0, t > s_i$, checking (7) for each codeword in $\mathcal{P}$, and selecting the vector which prunes the maximal number of low-weight codewords. Having obtained $\mathcal{P}(V_i^{(A)})$, we set $\mathcal{P} := \mathcal{P} \setminus \mathcal{P}(V_i^{(A)})$, and proceed with construction of the dynamic freezing constraints for the next suitable index $s_{i'} < s_i$.

The particular indices s_i to be considered in the above described elimination process should be as small as possible, so that the SCL decoder can process them at earliest possible phases, reducing thus the probability of the correct path being killed. On the other hand, constraints (4) should involve all symbols u_t which may give rise to a LWNZ codeword, so that as many as possible such codewords are pruned by each constraint.

Let $v_0 < v_1 < \ldots$ be distinct weights of rows of A, which correspond to non-frozen symbols in the parent code. For any t we define $f^{(t)}$ as the maximal index of a symbol, such that $\mathrm{wt}(A_{f^{(t)}}) = v_t$, and $f^{(t)}$ is not yet frozen, i.e. $f^{(t)} \notin \overline{\mathcal{F}}$, and $f^{(s)}$ is not equal to any of the previously selected s_i. If no such symbol exists, we assume $f^{(t)} = -\infty$. Let us set initially $\rho = 0$. We propose to set s_i as $f^{(t)}$, provided that $f^{(t)} \in [n]$, and the number of LWNZ codewords, which can be pruned for this s_i as described above, is at least $(\frac{1}{2} - \delta)|\mathcal{P}|$, where δ is a small value. The latter constraint ensures that no dynamic freezing constraints are imposed on information symbols which give rise to too few LWNZ codewords, and in many cases results in much lower error coefficient of the obtained code. Otherwise, we set $\rho := \rho + 1$ and repeat this selection process until f_A dynamic freezing constraints are constructed.

The parameter f_A should be selected as the smallest integer, which results in the number of LWNZ codewords in the obtained code to achieve some target value (e.g. 0). The dynamic freezing constraints and the corresponding dynamic frozen symbols obtained in this way are referred to as type-A ones.

Example 2. *Consider construction of a $(16, 7)$ polar subcode with the Arıkan's kernel for the BEC with erasure probability 0.5. Let $f_A = 2$. The corresponding parent $(16, 9)$ polar code has frozen set $\overline{\mathcal{F}} = \{0, 1, 2, 3, 4, 5, 8\}$. It has 44 codewords of weight 4. The corresponding information vectors are shown in Table I. Let $s_0 = 12$ be the index of the 0-th dynamic frozen symbol. It can be verified that $V_0^{(A)} = (0111111111111000)$ prunes 24 LWNZ codewords. The indices of surviving codewords are 2,3,5,7,10,11,13,15,18,19,21,24,26,27,29,32,34,35,38,39. Let $s_1 = 10$ be the index of the 1-st dynamic frozen symbol. It can be verified that $V_1^{(A)} = (0010100101100000)$ prunes 12 codewords. The indices of surviving codewords are 2, 7, 10, 15, 18, 24, 26, 32. Hence, the constraint matrix of the obtained*

TABLE I: Information vectors u_0^{15} for weight-4 codewords of $(16, 9, 4)$ polar code

ID	u_0^{15}	ID	u_0^{15}	ID	u_0^{15}	ID	u_0^{15}	ID	u_0^{15}	ID	u_0^{15}	ID	u_0^{15}	ID	u_0^{15}	ID	u_0^{15}
1	0000001000000000	6	0000001100001100	11	0000001000100010	16	0000001100111111	21	0000000001010000	26	0000000001100110	31	0000000001011111	36	0000000000111100	41	0000000000001000
2	0000001000001000	7	0000001100110000	12	0000001000101010	17	0000000001000000	22	0000000001110000	27	0000000001001100	32	0000000001111111	37	0000000000100010	42	0000000000001100
3	0000001000100000	8	0000001100111100	13	0000001100000011	18	0000000001100000	23	0000000001011010	28	0000000001101100	33	0000000000100000	38	0000000000101010	43	0000000000001010
4	0000001000101000	9	0000001000000010	14	0000001100001111	19	0000000001001000	24	0000000001111000	29	0000000001010101	34	0000000000101000	39	0000000000110011	44	0000000000001111
5	0000001100000000	10	0000001000001010	15	0000001100110011	20	0000000001101010	25	0000000001000100	30	0000000001110111	35	0000000000110000	40	0000000000111111		

polar subcode is

$$V = \left(\begin{array}{cccccccccccccccc}
1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&1&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&1&0&0&0&0&0&0&0\\
\hline
0&1&1&1&1&1&1&1&1&1&1&1&1&0&0&0\\
0&0&1&0&1&0&0&1&0&1&1&0&0&0&0&0
\end{array}\right)$$

Observe that it is possible to zero out the columns with indices in $\overline{\mathcal{F}}$ in the bottom part of this matrix by performing appropriate row operations.

C. Type-B dynamic frozen symbols

In this section we further refine the construction of polar subcodes by reducing the probability of the SCL decoder killing the correct path at an early phase.

The probability of $\mathcal{E}(L)$, the event corresponding to the score of the correct path becoming lower than the scores of L incorrect paths at some intermediate phase of SCL decoding, depends on how fast the scores of incorrect paths decrease after these paths diverge from the correct one. Therefore, we propose to introduce an additional set of dynamic frozen symbols, mapped onto relatively reliable bit subchannels, so that for an incorrect path the values of these symbols would deviate with high probability from those used by the encoder, causing thus the path score to drop.

Let C_i be the capacity (or some other reliability measure) of bit subchannel $W_m^{(i)}$, and consider sequence $r_i : C_{r_0} \leq C_{r_1} \leq \cdots \leq C_{r_{n-1}}$.

It was suggested in [5] to impose non-trivial dynamic freezing constraints on f_B symbols u_{r_i}, which are going to be transmitted over subchannels with indices $r_{n-k-f_A-1}, \ldots, r_{n-k-f_A-f_B}$, i.e. most reliable bit subchannels which would carry static frozen symbols in a classical polar code. Here f_B is a parameter of the code construction.

Such dynamic freezing constraints are referred to as type-B ones. Here we show that for a certain class of polarizing transformations this does not decrease the minimum distance.

Theorem 2. *Consider an (n, k, d) polar subcode $\mathcal{C}$ based on the polarizing matrix $A = F_{l_0} \otimes \cdots \otimes F_{l_{m-1}}$ with constraint matrix $V = \begin{pmatrix} V' \\ V'' \\ V^{(A)} \end{pmatrix}$, such that the kernels F_{l_i} satisfy* (6)*, both V' and $\begin{pmatrix} V' \\ V'' \end{pmatrix}$ are constraint matrices of some (n, k', d) and (n, k'', d) polar codes, all rows of V' and V'' have weight 1, s_i is the position of last non-zero entry in the i-th row of V, all s_i are distinct, and $\mathrm{wt}(A_{s_i}) = d, n - k' \leq i < n - k''$.*

Let $\widetilde{V} = \begin{pmatrix} V' \\ V^{(B)} \\ V^{(A)} \end{pmatrix}$ be a matrix, such that $\widetilde{V}_{i,t} = V''_{i,t}, s_i \leq t < n$, and $\widetilde{V}_{i,t}, t < s_i, n - k' \leq i < n - k''$, are arbitrary binary values[1].

Then $\widetilde{V}$ defines another $(n, k, \widetilde{d})$ polar subcode $\widetilde{\mathcal{C}}$, where $\widetilde{d} \geq d$.

Proof: By construction of $\widetilde{V}$, the sets of frozen symbol indices for both C and $\widetilde{\mathcal{C}}$ are identical. Let $\mathcal{F}$ denote this set.

According to [22, Theorem 7], partial distances of matrix A satisfy $D_i = \mathrm{wt}(A_i) = \prod_{j=0}^{m-1} D_{j,i_j}$, where A_i is the i-th row of A, and $i = \sum_{j=0}^{m-1} i_j \prod_{t=m-j}^{m-1} l_t$. Given a non-zero codeword $\mathbf{c} = u_0^{n-1} A$ from either $\mathcal{C}$ or $\widetilde{\mathcal{C}}$, let $t \notin \mathcal{F}$ be the position of the first non-zero element in vector u_0^{n-1}, so that $\mathbf{c} \in A_t + \mathcal{S}_t$, where $\mathcal{S}_t$ is the linear space generated by $A_{t+1}, \ldots, A_{n-1}$. The minimum distance between the coset $A_t + \mathcal{S}_t$ and $\mathcal{S}_t$ is equal D_t. Hence, $\mathrm{wt}(\mathbf{c}) \geq D_t$. Since the frozen set $\mathcal{F}$ is identical for $\mathcal{C}$ and $\widetilde{\mathcal{C}}$, one obtains that $\widetilde{d} \geq d = \min_{t \notin \mathcal{F}} D_t$. ∎

D. Summary of the proposed code construction method

Here we present a summary of the proposed method for construction of (n, k) polar subcode with an $n \times n$ polarizing matrix A satisfying the assumptions of Theorem 1:

1) Compute the capacities (or other reliability measures) of bit subchannels $W_m^{(i)}$. Let r_0^{n-1} be the sequence of subchannel indices in the ascending order of their reliability.
2) Construct parent polar code $\mathcal{C}'$ with the frozen set $\mathcal{F}' = \{r_0, \ldots, r_{n-k-f_A-f_B-1}\}$, where f_A and f_B are parameters of the code construction. Let V' be the corresponding constraint matrix, i.e. $V'_{i,s} = \begin{cases} 1, & \text{if } s = r_i, \\ 0 & \text{otherwise,} \end{cases}$ and let $j_i = r_i, 0 \leq i < n - k - f_A - f_B$, be the indices of frozen symbols.
3) Compute minimum distance d and enumerate (almost) all codewords $\mathbf{c}$ of weight d in $\mathcal{C}'$. Recall, that for codes with Arıkan's kernel all such codewords can be enumerated explicitly as described in Section II-B, while in a more general case computer search can be used to obtain them. Let $\mathcal{P}$ be the set of such codewords.
4) Construct matrix $V^{(B)}$, such that $V^{(B)}_{i,r_{n-k-f_A-f_B+i}} = 1, V^{(B)}_{i,j} = 0, j > r_{n-k-f_A-f_B+i}, 0 \leq i < f_B$, while the remaining elements of $V^{(B)}$ are selected as independent random binary values. Exclude from $\mathcal{P}$ codewords $\mathbf{c} = \mathbf{u}A$, such that $V^{(B)}\mathbf{u}^T \neq 0$.
   If a generic low-weight codeword search algorithm like [31] is used, then it may be easier to construct first a code $\overline{C}$ with constraint matrix $\overline{V} = \begin{pmatrix} V' \\ V^{(B)} \end{pmatrix}$, and search for LWNZ codewords in $\overline{C}$ instead of $\mathcal{C}'$.

[1] That is, $V^{(B)}$ is obtained by replacing some zeroes in V'' with random values.

TABLE II: Number of weight-8 codewords in $(1024, 768, \geq 8)$ polar subcodes with Arıkan's kernel

f_A	parent code	unoptimized subcode	optimized subcode
0	41344	41344	41344
1	57728	28808	28696
2	57728	14380	14108
3	57728	7152	6840
4	57728	3490	3258
5	57728	1736	1484
6	57728	813	632
7	61824	379	240
8	70016	149	48
9	70016	68	7
10	70016	24	0

5) Let $v_0 < v_1 < \ldots$ be distinct weights of rows of A, which correspond to non-frozen symbols in the parent code. Select the indices $s_i \notin \overline{\mathcal{F}}$ of type-A DFS as f_A maximal integers, such that $\mathrm{wt}(A_{s_i}) = v_j$ for smallest possible j. Construct $f_A \times n$ matrix $V^{(A)}$, such that $V^{(A)}_{i,s_i} = 1, V^{(A)}_{i,j} = 0, j > s_i$, and for each i the values $V^{(A)}_{i,t}, t < s_i$ are successively selected (e.g. by T iterations of randomized search) to eliminate as many as possible codewords from $\mathcal{P}$. Integer s_i is accepted as a type-A DFS index only if the corresponding constraint results in elimination of at least $(0.5 - \delta) \cdot 100\%$ of the remaining LWNZ codewords, where δ is a small value. Let $\widetilde{V} = \begin{pmatrix} V' \\ V^{(B)} \\ V^{(A)} \end{pmatrix}$ be the constraint matrix of the obtained polar subcode.

The number of type-A dynamic frozen symbols f_A can be adjusted depending on the residual number of low-weight codewords after step 5. Unfortunately, there is no simple way (besides simulations) to find an optimal value of the number f_B of type-B dynamic frozen symbols.

E. Preconditioned polar subcodes

A simple way to improve the minimum distance of polar codes satisfying the constraints of Theorem 1 is to enforce at steps 1 and 2 $i \in \mathcal{F}'$ for all $i : \mathrm{wt}(A_i) < d$ for some integer d. This results in a code with minimum distance at least d. In the case of Arıkan kernel, such codes still satisfy the partial order (1)–(2), and their low-weight codewords can be obtained as described above. However, such polar codes typically have huge error coefficient.

We propose to use such polar codes as parent ones, and employ the above described algorithm to obtain *preconditioned* polar subcodes with much lower error coefficient and improved performance under near-ML decoding.

IV. Numerical results

Table II presents the number of weight-8 codewords of parent $(1024, 768 + t)$ polar codes and their $(1024, 768)$ subcodes with f_A type-A and $f_B = 0$ type-B dynamic frozen symbols. It can be seen that the error coefficient of parent codes increases very slowly with f_A. This justifies application of type-A dynamic frozen symbols for reduction of the error coefficient. It can be also seen that increasing f_A results in lower error coefficient for the codes obtained by the proposed optimization procedure.

TABLE III: Distance properties of $(1024, 768)$ polar subcodes with Arıkan's kernel

Code construction	d	w_d
unoptimized, $f_A = 11, f_B = 0$	8	9
unoptimized, $f_A = 11, f_B = 53$	8	≥ 6
optimized, $f_A = 11, f_B = 0$	12	≥ 5658
optimized, $f_A = 11, f_B = 53$	12	≥ 2682
optimized, $f_A = 16, f_B = 50$	12	≥ 548
BCH(12)	12	≥ 5632
CRC-11	12	30303
CRC-16	12	996
preconditioned $d = 16$, unoptimized, $f_A = 0, f_B = 0$	16	38821056
preconditioned $d = 16$, unoptimized, $f_A = 10, f_B = 0$	16	52820
preconditioned $d = 16$, unoptimized, $f_A = 10, f_B = 54$	16	≥ 2962
preconditioned $d = 16$, optimized, $f_A = 10, f_B = 0$	16	51324
preconditioned $d = 16$, optimized, $f_A = 10, f_B = 54$	16	≥ 2630

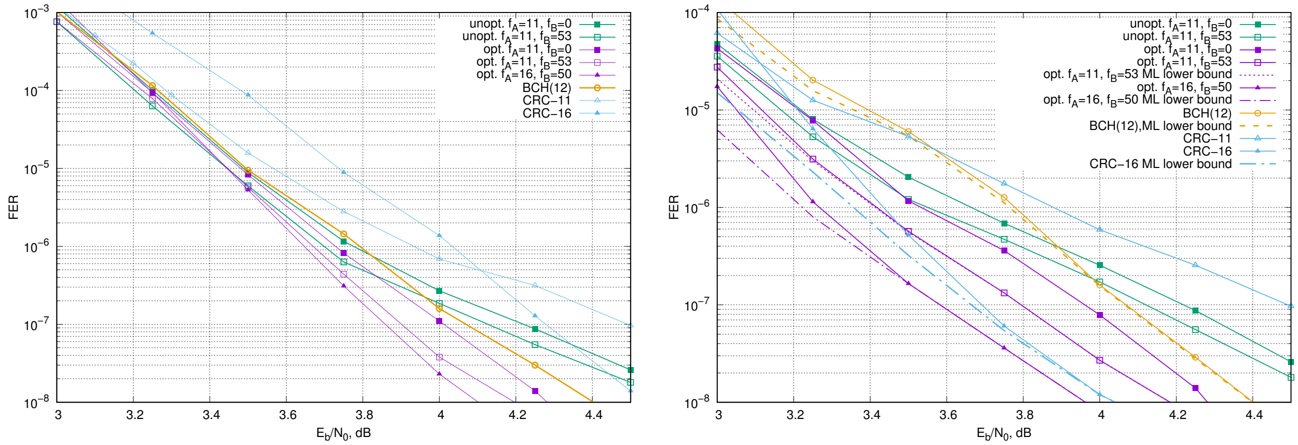
We present the performance of polar subcodes obtained with the proposed method. Simulations were performed for the case of BPSK modulation and AWGN channel. Decoding was done by the sequential algorithm [6], [28], with at most L visits per phase (equivalent to list size in [2]).

Figure 2 presents the performance of $(1024, 768)$ polar subcodes with Arıkan's kernel. The codes were constructed for $E_b/N_0 = 3$ dB. We report the performance of unoptimized randomized polar subcodes [5], and the codes obtained with the proposed code construction algorithm with $T = 1000$. For comparison, we report also the performance of polar codes with CRC, as well as a polar subcode of the extended BCH code with minimum distance 12 [4]. CRC-f_A can be considered as an alternative way to select a quasi-random linear subcode of the parent $(n, k + f_A, \overline{d})$ polar code, so that most of the LWNZ codewords are eliminated. For some codes we report a lower bound on the maximum likelihood decoding error probability, which is obtained as the fraction of events corresponding to the codeword obtained by the decoder being closer to the received vector than the actual transmitted codeword.

It can be seen that the proposed optimized polar subcode with $f_A = 11$ provides 0.5 dB performance gain with respect to the unoptimized one. It also far outperforms polar code with CRC and polar subcode of the extended BCH code. Observe that type-B dynamic frozen symbols provide significant performance improvement for large L. Even larger performance gain can be obtained by setting $f_A = 16$. However, for small L at low SNR polar subcode with $f_A = 16$ has inferior performance compared to the one obtained for $f_A = 11$. Increasing the value of f_A results in much more significant performance penalty at low SNR for polar codes with CRC.

It can be also seen that at high SNR the performance of the sequential decoder with $L = 32$ is almost the same as in the case of $L = 4096$.

Table III presents the minimum distance and error coefficient of the considered codes. Entries $\geq w$ were obtained



(a) $L = 32$ (b) $L = 4096$

Fig. 2: Performance of $(1024, 768)$ polar subcodes with Arıkan's kernel

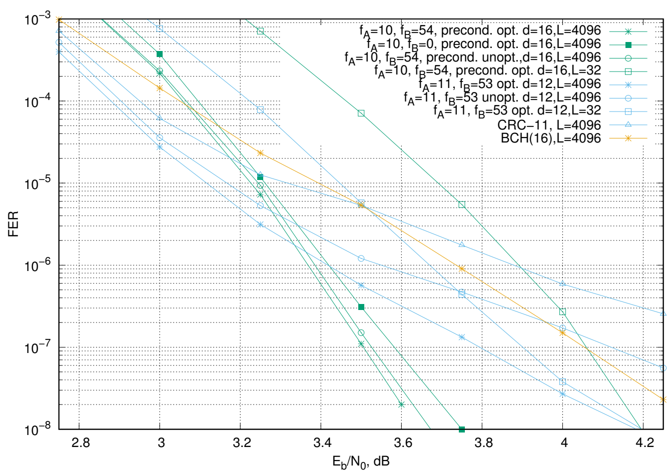


Fig. 3: Performance of $(1024, 768)$ preconditioned polar subcodes with Arıkan's kernel

by the algorithm in [31], which was allowed to run for 10^7 iterations. It can be seen that the proposed approach allows construction of codes with minimum distance $d = 12$ with much smaller error coefficient compared to the BCH-based construction, as well as compared to polar codes with CRC. Furthermore, type-B dynamic frozen symbols indeed provide some reduction of the error coefficient. This explains the performance gain reported in Figure 2.

Figure 3 presents the performance of preconditioned polar subcodes. These codes have inferior performance at low SNR compared to the case of $\overline{\mathcal{F}}$ constructed solely based on subchannel reliability[2]. This is compensated by 0.4 dB gain with respect to the non-preconditioned code at $FER = 2{\cdot}10^{-8}$ for $L = 4096$. The total gain with respect to the code obtained without preconditioning and optimization is approximately 0.7 dB. Observe also that the proposed preconditioned codes with design distance $d = 16$ outperform a polar subcode of the extended BCH code with the same design distance, as well as the codes obtained without preconditioning (these codes have minimum distance $d = 12$). However, for $L = 32$ the preconditioned polar subcode exhibits huge performance loss compared to the code without preconditioning. It can be also seen from Table III that with $f_A = 10$ type-A dynamic frozen symbols the proposed optimization method allows one to reduce the error coefficient only by 3% for $f_B = 0$ and approximately by 12% for $f_B = 54$. This results in a marginal performance gain for the proposed codes, although the sequential decoder with $L = 4096$ still does not provide maximum likelihood decoding of the preconditioned codes.

Polar codes were shown in [32] not to have an error floor. This result was obtained for the case of the SC decoding algorithm. However, the results shown in Figure 3 clearly show that improved versions of polar codes, such as polar codes with CRC and polar subcodes, may exhibit an error floor under list successive cancellation or sequential decoding, although their performance is much better compared to polar codes with the same parameters. Indeed, the slope of the FER curves for the polar code with CRC and polar subcode with $d = 12$ is substantially different at E_b/N_0 values 2.8 dB and 3.6 dB. The reason for this apparent contradiction is that the SC algorithm is very far from maximum likelihood decoding, especially at low SNR. List/sequential SC decoding may provide substantially better performance in this region. On the other hand, in the high SNR region the slope of the FER curve depends on the minimum distance of the code, which is still given by $O(\sqrt{n})$, even for improved code constructions. However, the proposed code construction, especially the preconditioning method, may substantially reduce the error coefficient, and in some cases increase (non-asymptotically) the minimum distance of the obtained codes, and therefore change the location of the error floor.

Figure 4 presents the performance of rate-$1/2$ codes. The results are pretty much similar to those obtained for rate $3/4$ codes, i.e. optimized polar subcodes far outperform non-optimized ones. Optimized polar subcodes without precondi-

[2]The latter polar subcodes have minimum distance 12.

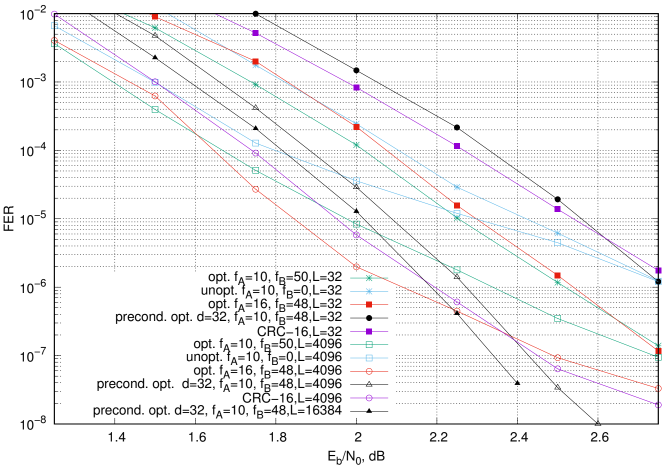


Fig. 4: Performance of $(1024, 512)$ codes

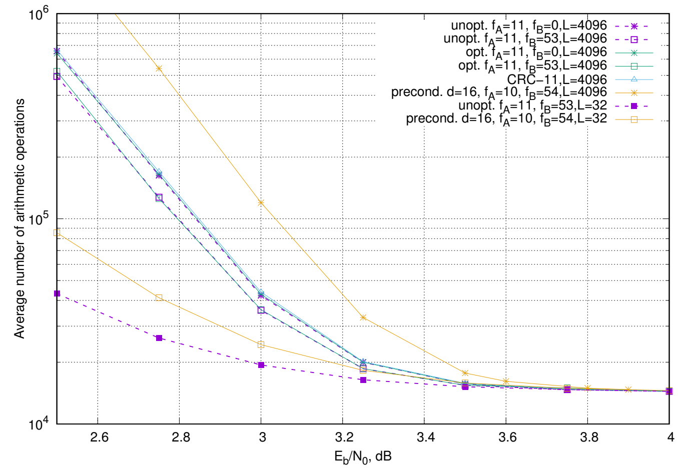


Fig. 5: Complexity of sequential decoding of $(1024, 768)$ polar subcodes with Arıkan kernel

tioning were found to have minimum distance $d = 24$. Due to lack of an analytical tool for enumeration of codewords of an Arıkan polar code of weight more than its minimum distance, the number of weight-24 codewords in the obtained polar subcode with $f_A = 16, f_B = 48$ could not be optimized. It was found to be equal to $w_{24} \approx 347$. This is still less than the error coefficient of the polar code with CRC-16 ($w'_{24} \approx 717$). However, the latter code performs slightly better at high SNR. This should be attributed to the differences in higher-order components of weight spectrum.

At high SNR the best performance is provided by the preconditioned polar subcode, which has higher minimum distance $d = 32$ than all other considered codes. Even $L = 16384$ is not sufficient to implement near-ML decoding of the preconditioned code. However, increasing L from 4096 to 16384 provides 0.1 dB performance gain, while the average sequential decoding complexity in the latter case is still $3.98 \cdot 10^4$ operations at $E_b/N_0 = 2.4$ dB. This is quite close to the decoding complexity of the non-preconditioned optimized code, which is equal to $1.91 \cdot 10^4$, $1.89 \cdot 10^4$ and $1.62 \cdot 10^4$ for $L = 16384, 4096, 32$, respectively. Observe for the non-preconditioned code for $L = 4096$ all observed error events at this SNR are ML decoding errors, so this code cannot benefit from $L > 4096$.

Figure 5 illustrates the average complexity (i.e. number of summation and comparison operations) of the sequential decoding algorithm for the considered codes. It can be seen that for $E_b/N_0 \geq 3.5$ dB, where the proposed codes start to outperform non-optimized ones, the average decoding complexity is very close to $n \log_2 n$ operations even for $L = 4096$. Furthermore, introducing type-B dynamic frozen symbols (i.e. setting $f_B > 0$) results in reduced decoding complexity at low E_b/N_0. The decoding complexity of a polar code with CRC is very close to that of polar subcodes with $f_B = 0$. Preconditioned code has somewhat higher decoding complexity compared to the case of $\mathcal{F}'$ obtained solely based on subchannel reliability, but at high SNR its decoding complexity converges to the same value as for the case of non-preconditioned codes..

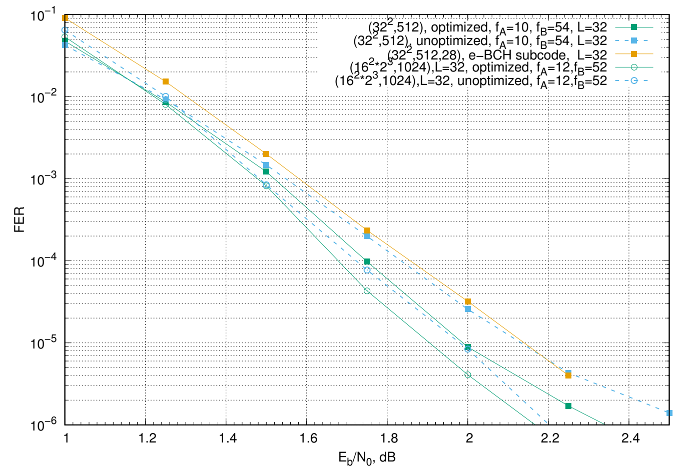


Fig. 6: Performance of polar subcodes with large kernels

Hence, one can exploit excellent performance of the proposed codes at high SNR with very small decoding complexity. This comes at the expense of large memory requirements and high decoder latency. These problems were addressed for the case of SCL-like decoding in [33].

Figure 6 illustrates the performance of polar (sub)codes with various kernels under SCL decoding. Here $(l_0^{m_0} l_1^{m_1}, k)$ denotes a k-dimensional code based on the polarizing transformation $A = F_{l_0}^{\otimes m_0} \otimes F_{l_1}^{\otimes m_1}$, and F_{l_i} is a $l_i \times l_i$ kernel. F_{16} and F_{32} are the kernels with scaling exponents 3.45 and 3.202, having rate of polarization 0.518 and 0.529, respectively, reported in [12]. The proposed optimization method results in up to 0.2 dB gain compared to a randomized construction without optimized selection of $V^{(A)}$. Observe that the obtained $(1024, 512)$ code outperforms the polar subcode of an extended BCH code [4], where low-weight codewords are eliminated algebraically. It can be also seen that the pruned mixed-kernel $(2048, 1024)$ polar subcode outperforms the one with Arıkan's kernel only, while the SCL decoding complexity of these codes is 10^6 and $2 \cdot 10^6$ operations, respectively.

V. Conclusions

In this paper a method for construction of polar subcodes with reduced error coefficient was presented. The obtained codes were shown to provide at high SNR substantially better performance compared to unoptimized randomized polar subcodes, polar subcodes of extended BCH codes, and polar codes with CRC.

The proposed code construction method relies on enumeration of low-weight codewords in a parent polar code. This can be done analytically for the case of codes with Arıkan's kernel, but requires expensive computer search for other kernels. Explicit characterization of the set of their low-weight codewords would enable faster and more accurate construction of polar subcodes with such kernels.

The proposed codes provide the most significant performance gain in the high SNR region. By employing the sequential decoding algorithm, this gain can be obtained with decoding complexity close to that of plain successive cancellation algorithm.

Acknowledgement

The author thanks the editor and anonymous reviewers for their detailed comments, which have helped to greatly improve the presentation of the results.

References

[1] E. Arıkan, "Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," *IEEE Transactions on Information Theory*, vol. 55, no. 7, pp. 3051–3073, July 2009.
[2] I. Tal and A. Vardy, "List decoding of polar codes," *IEEE Transactions On Information Theory*, vol. 61, no. 5, pp. 2213–2226, May 2015.
[3] K. Niu and K. Chen, "CRC-aided decoding of polar codes," *IEEE Communications Letters*, vol. 16, no. 10, pp. 1668–1671, October 2012.
[4] P. Trifonov and V. Miloslavskaya, "Polar subcodes," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 2, pp. 254–266, February 2016.
[5] P. Trifonov and G. Trofimiuk, "A randomized construction of polar subcodes," in *Proceedings of IEEE International Symposium on Information Theory*. Aachen, Germany: IEEE, 2017, pp. 1863–1867.
[6] V. Miloslavskaya and P. Trifonov, "Sequential decoding of polar codes," *IEEE Communications Letters*, vol. 18, no. 7, pp. 1127–1130, 2014.
[7] "TS 36.212 V13.2.0 (2016-08). Technical Specification. LTE; Evolved Universal Terrestrial Radio Access (E-UTRA); Multiplexing and channel coding," August 2016.
[8] V. Bioglio, C. Condo, and I. Land, "Design of polar codes in 5g new radio," *IEEE Communications Surveys & Tutorials*, 2020.
[9] M. Mondelli, R. Urbanke, and S. H. Hassani, "Unified scaling of polar codes: Error exponent,scaling exponent, moderate deviations, and error floors," in *Proceedings of IEEE International Symposium on Information Theory*, 2015.
[10] S. B. Korada, E. Sasoglu, and R. Urbanke, "Polar codes: Characterization of exponent, bounds, and constructions," *IEEE Transactions on Information Theory*, vol. 56, no. 12, pp. 6253–6264, December 2010.
[11] A. Fazeli, S. H. Hassani, M. Mondelli, and A. Vardy, "Binary linear codes with optimal scaling: Polar codes with large kernels," in *Proceedings of IEEE Information Theory Workshop*, 2018.
[12] G. Trofimiuk and P. Trifonov, "Reduced complexity window processing of binary polarization kernels," in *Proceedings of IEEE International Symposium on Information Theory*, Paris, France, July 2019.
[13] G. Trofimiuk and P. Trifonov, "Efficient decoding of polar codes with some 16×16 kernels," in *Proceedings of IEEE Information Theory Workshop*, 2018.
[14] V. Bioglio and I. Land, "On the marginalization of polarizing kernels," in *Proceedings of International Symposium on Turbo Codes and Iterative Information Processing*, 2018.
[15] M. Mondelli, S. H. Hassani, and R. Urbanke, "Scaling exponent of list decoders with applications to polar codes," *IEEE Transactions On Information Theory*, vol. 61, no. 9, September 2015.
[16] P. Trifonov, "Efficient design and decoding of polar codes," *IEEE Transactions on Communications*, vol. 60, no. 11, pp. 3221 – 3227, November 2012.
[17] I. Tal and A. Vardy, "How to construct polar codes," *IEEE Transactions on Information Theory*, vol. 59, no. 10, pp. 6562–6582, October 2013.
[18] V. Miloslavskaya and P. Trifonov, "Design of polar codes with arbitrary kernels," in *Proceedings of IEEE Information Theory Workshop*, 2012, pp. 119–123.
[19] H. Yao, A. Fazeli, and A. Vardy, "Explicit polar codes with small scaling exponent," in *Proceedings of IEEE International Symposium on Information Theory*, 2019.
[20] P. Trifonov, "On construction of polar subcodes with large kernels," in *Proceedings of IEEE International Symposium on Information Theory*, Paris, France, 2019.
[21] N. Presman, O. Shapira, and S. Litsyn, "Mixed-kernels constructions of polar codes," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 2, pp. 239–253, Feb 2016.
[22] M.-K. Lee and K. Yang, "The exponent of a polarizing matrix constructed from the kronecker product," *Des. Codes Cryptography*, vol. 70, no. 3, pp. 313–322, Mar. 2014. [Online]. Available: http://dx.doi.org/10.1007/s10623-012-9689-z
[23] M. Bardet, V. Dragoi, A. Otmani, and J.-P. Tillich, "Algebraic properties of polar codes from a new polynomial formalism," in *Proceedings of IEEE International Symposium on Information Theory*, 2016.
[24] I. Zhegalkin, "On the technique of calculating propositions in symbolic logic," *Matematicheskii Sbornik*, vol. 34, no. 1, pp. 9–28, 1927, in Russian.
[25] C. Schürch, "A partial order for the synthesized channels of a polar code," in *Proceedings of IEEE International Symposium on Information Theory*, 2016.
[26] M. Mondelli, S. H. Hassani, and R. Urbanke, "Construction of polar codes with sublinear complexity," *IEEE Transactions On Information Theory*, vol. 65, no. 5, May 2019.
[27] P. Trifonov, "Randomized chained polar subcodes," in *Proceedings of IEEE Wireless Communications and Networking Conference Workshops*. Barcelona, Spain: IEEE, 2018, pp. 292–297.
[28] ——, "A score function for sequential decoding of polar codes," in *Proceedings of IEEE International Symposium on Information Theory*, Vail, USA, 2018.
[29] E. Blokh and V. Zyablov, "Coding of generalized concatenated codes," *Problems of Information Transmission*, vol. 10, no. 3, pp. 45–50, 1974.
[30] V. Zyablov, S. Shavgulidze, and M. Bossert, "An introduction to generalized concatenated codes," *European transactions on telecommunications*, vol. 10, no. 6, pp. 609–622, 1999.
[31] 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, January 1998.
[32] M. Mondelli, S. H. Hassani, and R. Urbanke, "Unified scaling of polar codes: Error exponent, scaling exponent, moderate deviations, and error floors," *IEEE Transactions On Information Theory*, vol. 62, no. 12, December 2016.
[33] S. A. Hashemi, M. Mondelli, S. H. Hassani, C. Condo, R. Urbanke, and W. Gross, "Decoder partitioning: Towards practical list decoding of polar codes," *IEEE Transactions On Communications*, vol. 66, no. 9, September 2018.

Peter Trifonov (S'02,M'05) was born in St.Petersburg, USSR in 1980. He received the MSc and PhD (Candidate of Science) degrees from Saint Petersburg Polytechnic University in 2003 and 2005, and Dr.Sc degree from the Institute for Information Transmission Problems in 2018. His research interests include coding theory and its applications in telecommunications and storage systems. Currently he is a professor at the ITMO University in Saint Petersburg, Russia. He is an editor at IEEE Transaction on Communications.