

Recursive Trellis Processing of Large Polarization Kernels

Peter Trifonov
ITMO University, Russia
Email: pvtrifonov@itmo.ru

Abstract—A reduced complexity algorithm is presented for computing the log-likelihood ratios arising in the successive cancellation decoder for polar codes with large kernels of arbitrary dimension. The proposed algorithm exploits recursive trellis representation of the codes generated by submatrices of the polarization kernel, and enables polar codes based on large kernels to provide better performance compared to the codes based on Arikan kernel with the same decoding complexity.

I. INTRODUCTION

Polar codes [1], have already made their way into the 5G specification. However, their role is limited there for encoding of small data blocks, since long polar codes do not compete well with e.g. LDPC codes. Polar codes with large kernels were shown to have asymptotically optimal scaling exponent [2]. The SC decoding algorithm for such codes has complexity $O(n \log n)$ operations of computing LLRs. Each such operation, called kernel processing or kernel marginalization [3], has complexity $O(2^l)$ if implemented straightforwardly, where l is kernel dimension. An approximate kernel processing method based on box-and-match decoder was suggested in [4]. Some kernels of dimension 16 and 32 were published together with efficient processing algorithms [5]. The corresponding polar codes were shown to have lower decoding complexity compared to codes based on Arikan kernel with the same performance. However, no efficient processing techniques for generic kernels have been published so far.

In this paper we present such algorithm, and report its complexity for some kernels available in the literature. We show that the proposed approach has much lower complexity compared to the Viterbi algorithm.

II. BACKGROUND

A. Polar codes

Polar code [6] is a set of vectors $c_0^{n-1} = u_0^{n-1} K^{\otimes m}$, where K is a non-singular $l \times l$ matrix called polarization kernel, $n = l^m$, $u_i = 0$ for $i \in \mathcal{F}$, $\mathcal{F} \subset [n]$ is a frozen set, and $[n] = \{0, \dots, n-1\}$. This definition can be generalized to obtain mixed kernel polar codes with the codewords given by $c_0^{n-1} = u_0^{n-1} (K_{l_1} \otimes K_{l_2} \otimes \dots \otimes K_{l_m})$ [7], [8], where K_{l_i} is a kernel of dimension l_i . Unless stated otherwise, we consider here the case of all kernels being the same.

Decoding of polar codes can be implemented by the successive cancellation algorithm, which makes decisions

$$\hat{u}_i = \begin{cases} 0, & i \in \mathcal{F} \\ \arg \max_{u_i \in \mathbb{F}_2} \mathcal{W}_m^{(i)}(\hat{u}_0^{i-1} \bullet u_i | r_0^{n-1}), & i \notin \mathcal{F}, \end{cases}$$

where $\mathcal{W}_m^{(i)}(u_0^i | r_0^{n-1}) = \sum_{u_{i+1}^{n-1}} \prod_{j=0}^{n-1} W_0^{(0)}((u_0^{n-1} K^{\otimes l})_i | r_i)$, $\mathcal{W}_0^{(0)}(c|r) = W_0^{(0)}(c|r) = \frac{W(r|c)}{2W(r)}$, $W(r|c)$ is the channel transition probability, and \bullet denotes the concatenation operator. These probabilities can be recursively computed as

$$\mathcal{W}_m^{(li+s)}(u_0^{li+s} | r_0^{n-1}) = \sum_{u_{li+s+1}^{l-1}} \prod_{j=0}^{l-1} \mathcal{W}_{m-1}^{(i)}\left(\left(u_{lt}^{l(t+1)-1} K\right)_j, t \in [i+1] | r_{j,l}^{n-1}\right), \quad (1)$$

where $r_{j,l}^{n-1} = (r_j, r_{j+l}, \dots, r_{j+n-l})$. This operation is known as kernel processing or kernel marginalization [9], [3].

It was suggested in [4] to approximate the above probabilities as $\mathcal{W}_m^{(i)}(u_0^i | r_0^{n-1}) \approx W_m^{(i)}(u_0^i | r_0^{n-1}) = \max_{u_{i+1}^{n-1}} \prod_{j=0}^{n-1} W_0^{(0)}((u_0^{n-1} K^{\otimes l})_j | r_j)$, so that

$$W_m^{(li+s)}(u_0^{li+s} | r_0^{n-1}) = \max_{u_{s+1}^{l-1}} \prod_{j=0}^{l-1} W_{m-1}^{(i)}\left(\left(u_{lt}^{l(t+1)-1} K\right)_j, t \in [i+1] | r_{j,l}^{n-1}\right). \quad (2)$$

Decoding can be implemented in the LLR domain using LLRs

$$S_m^{(i)}(u_0^{i-1}, r_0^{n-1}) = \ln \frac{W_m^{(i)}(u_0^{i-1} \bullet 0 | r_0^{n-1})}{W_m^{(i)}(u_0^{i-1} \bullet 1 | r_0^{n-1})}.$$

Assume for the sake of simplicity $m = 1$. It can be seen that

$$\begin{aligned} S_1^{(i)}(u_0^{i-1}, r_0^{n-1}) &= \max_{u_{i+1}^{l-1}} \sum_{j=0}^{l-1} \ln W_0^{(0)}\left(\left(\left(u_0^{i-1} \bullet 0 \bullet u_{i+1}^{l-1}\right) K\right)_j | r_j\right) \\ &\quad - \max_{u_{i+1}^{l-1}} \sum_{j=0}^{l-1} \ln W_0^{(0)}\left(\left(\left(u_0^{i-1} \bullet 1 \bullet u_{i+1}^{l-1}\right) K\right)_j | r_j\right) \\ &= \frac{1}{2} \max_{u_{i+1}^{l-1}} Q\left(\left(u_0^{i-1} \bullet 0 \bullet u_{i+1}^{l-1}\right) K, r_0^{l-1}\right) \\ &\quad - \frac{1}{2} \max_{u_{i+1}^{l-1}} Q\left(\left(u_0^{i-1} \bullet 1 \bullet u_{i+1}^{l-1}\right) K, r_0^{l-1}\right), \quad (3) \end{aligned}$$

where $Q(c_0^{l-1}, r_0^{l-1}) = \sum_{j=0}^{l-1} (-1)^{c_j} S_0^{(0)}(r_j)$ is the correlation function. The objective of the present paper is to provide an efficient method for computing LLRs $S_1^{(i)}(u_0^{i-1}, r_0^{n-1})$.

B. Recursive maximum likelihood decoding algorithm

It can be seen that computing (3) reduces to ML decoding in the cosets of the code generated by $l - i - 1$ last rows of

matrix K , provided that all symbols $u_s, s \in [l]$, are equiprobable binary random values. An efficient recursive maximum likelihood decoding algorithm for linear block codes was suggested in [10]. The idea is to recursively partition the received noisy vector into a number of sections $[x, y]$, identify for each section a number of most likely vectors $c_x^{y-1} \in \mathbb{F}_2^{y-x}$ corresponding to the received values r_x^{y-1} , and combine them to obtain most likely vectors for longer sections.

Given a linear code C , let $C_{h,h'}$ be its subcode, such that all its codewords have non-zero symbols only in positions $h \leq i < h'$. Let $p_{h,h'}(C)$ be a linear code obtained by puncturing all symbols, except those in positions $h \leq i < h'$, from codewords of C . Let us further define $s_{h,h'}(C) = p_{h,h'}(C_{h,h'})$, i.e. a code obtained from C by shortening it on all symbols except those with indices $h \leq i < h'$. The codes $s_{h,h'}(C)$ and $p_{h,h'}(C)$ are referred to as *section codes*. Consider a minimal trellis of code C , and its section corresponding to symbols from x to y . It is possible to show that the paths between two adjacent states in this section correspond to a coset in $p_{x,y}(C)/s_{x,y}(C)$ [10]. This coset may appear in the trellis several times. Hence, one can simplify maximum likelihood (ML) decoding by pre-computing the metrics of these paths. That is, for each coset $D \in p_{x,y}(C)/s_{x,y}(C)$ one needs to identify the most probable element $l(D)$, and store its correlation $Q(D) = Q(l(D), r_x^{y-1})$. Let the *composite branch table (CBT)* $T_{x,y}$ be an array containing values $T_{x,y}[v].l = l(D)$ and $T_{x,y}[v].q = Q(D)$, where v is an index of D . In the case of conventional decoding of an (n, k) code, $p_{0,n}(C)/s_{0,n}(C)$ contains a single element, so the corresponding CBT has one entry $T_{0,n}[0]$, which gives a solution of the ML decoding problem.

The straightforward approach to construction of a composite branch table for some code C is to enumerate all codewords of $p_{x,y}(C)$, and find the most probable one for each coset in $p_{x,y}(C)/s_{x,y}(C)$. We assume that this method is used for $y - x \leq 2$. However, more efficient approach was suggested in [10] for the case of $y - x \geq 2$. Consider some $z : x < z < y$. Let the generator matrix of $p_{x,y}(C)$ be represented as

$$G_{x,y}^{(p)} = \begin{pmatrix} G_{x,z}^{(s)} & 0 \\ 0 & G_{z,y}^{(s)} \\ G_{x,y}^{(00)} & G_{x,y}^{(01)} \\ G_{x,y}^{(10)} & G_{x,y}^{(11)} \end{pmatrix}, \quad (4)$$

where $G_{x,y}^{(s)} = \begin{pmatrix} G_{x,z}^{(s)} & 0 \\ 0 & G_{z,y}^{(s)} \\ G_{x,y}^{(00)} & G_{x,y}^{(01)} \\ G_{x,y}^{(10)} & G_{x,y}^{(11)} \end{pmatrix}$ is a generator matrix of $s_{x,y}(C)$, and $G_{x,y}^{(00)}, G_{x,y}^{(01)}$ are some $k''_{x,y} \times (z - x)$ and $k''_{x,y} \times (y - z)$ matrices, respectively, where $k'_{x,y} = k'_{x,y}(C)$ and $k''_{x,y} = k''_{x,y}(C)$ are some code-dependent integers. There is a one-to-one correspondence between vectors $vG'_{x,y}$, where $G'_{x,y} = \begin{pmatrix} G_{x,y}^{(10)} & G_{x,y}^{(11)} \end{pmatrix}$ is a $k'_{x,y} \times (y - x)$ matrix, and cosets $D \in p_{x,y}(C)/s_{x,y}(C)$. Here $k'_{x,y}, k''_{x,y}$ are some integers, which depend on code structure, and can be obtained from the minimum span form of its generator matrix.

Hence, we write $T_{x,y}[v].l := l(D)$ and $T_{x,y}[v].q = Q(D)$, with D being a coset corresponding to v . It can be seen that

$$\begin{aligned} T_{x,y}[v].q &= \max_{c_x^{y-1} \in D} Q(c_x^{y-1}, r_x^{y-1}) \\ &= \max_{w \in \mathbb{F}_2^{k''_{x,y}}} (T_{x,z}[a].q + T_{z,y}[b].q), v \in \mathbb{F}_2^{k'_{x,y}} \end{aligned} \quad (5)$$

where a and b are indices of the cosets $D' \in p_{x,z}(C)/s_{x,z}(C)$ and $D'' \in p_{z,y}(C)/s_{z,y}(C)$, respectively, such that $(w \ v) \begin{pmatrix} G_{x,y}^{(00)} \\ G_{x,y}^{(10)} \end{pmatrix} \in D'$ and $(w \ v) \begin{pmatrix} G_{x,y}^{(01)} \\ G_{x,y}^{(11)} \end{pmatrix} \in D''$. Such vectors a, b can be identified from the system of equations

$$\begin{aligned} (a' \ a) \begin{pmatrix} G_{x,z}^{(s)} \\ G'_{x,z} \end{pmatrix} &= (w \ v) \begin{pmatrix} G_{x,y}^{(00)} \\ G_{x,y}^{(10)} \end{pmatrix} \\ (b' \ b) \begin{pmatrix} G_{z,y}^{(s)} \\ G'_{z,y} \end{pmatrix} &= (w \ v) \begin{pmatrix} G_{x,y}^{(01)} \\ G_{x,y}^{(11)} \end{pmatrix}, \end{aligned}$$

where a', b' are not used elsewhere. The solution is given by $a = (w \ v) \hat{G}_{x,y}$ and $b = (w \ v) \tilde{G}_{x,y}$ for some matrices $\hat{G}_{x,y}$ and $\tilde{G}_{x,y}$. The corresponding most likely coset representative is given by $T_{x,y}[v].l = T_{x,z}[\hat{a}].l \bullet T_{z,y}[\hat{b}].l$, where \hat{a}, \hat{b} are the values of a and b , which deliver maximum in (5).

The complexity of this calculation is $O(2^{k'_{x,y} + k''_{x,y}})$. It can be further reduced by exploiting the tricks suggested in [10]. The overall decoding complexity strongly depends on the sectionalization method being used, i.e. a rule for selection of the partitioning point z for some x, y . This approach, known as recursive maximum likelihood decoding (RMLD), was shown to be more efficient compared to the Viterbi algorithm [10].

III. RECURSIVE TRELLIS PROCESSING

A. Extended kernel codes

Polar codes can be considered as a result of recursive application of the construction of generalized concatenated codes [11], [12]. These codes rely on non-systematic inner codes. An optimal SISO decoding algorithm for non-systematically encoded linear block codes was presented in [13]. This algorithm can be easily tailored to implement computation of (3). To do this, consider an extended $(l + 1, l - i)$ code $\bar{C}^{(i)}$ generated by matrix $G^{(i)}$. First l columns of this matrix are obtained by taking $l - i$ last rows of kernel K . The last column has 1 in the 0-th row, and zeroes in the remaining positions. Assuming that $u_0^{i-1} = 0$, one obtains that computing (3) is equivalent to finding the most probable codewords of code $\bar{C}^{(i)}$ having 0 and 1 in the last symbol. This can be implemented by running the Viterbi algorithm over the trellis of $\bar{C}^{(i)}$, assuming that the last codeword symbol is erased. The same trellises, although with different labeling, can be used to implement decoding in the cosets of $\bar{C}^{(i)}$, which arise in the case of $u_0^{i-1} \neq 0$.

Example 1. Consider Arikan kernel $F_2 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$. One obtains $G^{(0)} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$ and $G^{(1)} = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix}$.

B. Recursive processing of polarization kernels

We propose to compute (3) by applying the RMLD algorithm to the cosets of the extended kernel codes. Let $\bar{\mathcal{C}}(u_0^{i-1}) = w + \bar{\mathcal{C}}^{(i)}$ be the coset of $\mathcal{C}^{(i)}$ given by prior decisions u_0^{i-1} , where $w = (u_0^{i-1}K_{0,\dots,i-1}, 0)$, and $K_{0,\dots,i-1}$ is the matrix consisting of i top rows of K . For any section $[x, y]$, the cosets associated with states in the recursive trellis for $\bar{\mathcal{C}}(u_0^{i-1})$ are obtained from those for $\bar{\mathcal{C}}^{(i)}$ as $D(u_0^{i-1}) = \{f + w_x^{y-1} | f \in D\}$, $D \in p_{x,y}(\bar{\mathcal{C}}^{(i)})/s_{x,y}(\bar{\mathcal{C}}^{(i)})$.

It can be seen that $p_{0,l}(\bar{\mathcal{C}}^{(i)})/s_{0,l}(\bar{\mathcal{C}}^{(i)})$ contains two cosets, which correspond to $u_i = 0$ and $u_i = 1$. Hence,

$$S_1^{(i)}(u_0^{i-1}, r_0^{l-1}) = \frac{T_{0,l}[0].q - T_{0,l}[1].q}{2}, \quad (6)$$

where $T_{0,l}$ is the composite branch table constructed for $\bar{\mathcal{C}}(u_0^{i-1})$ given noisy vector r_0^{l-1} . It is assumed in what follows that *CBT* entries contain only q values, and the corresponding coset representatives $l(D)$ are omitted.

Furthermore, we propose to reuse the composite branch tables, or their parts, obtained at successive phases i . To do this, we need to identify how CBTs evolve with i , find a way to handle prior decisions u_0^{i-1} , design efficient algorithms for construction of CBTs for short sections, and obtain an optimal sectionalization strategy.

1) *Reusing the CBTs*: Assume that the same sectionalization is used for all phases i . Obviously, $p_{x,y}(\bar{\mathcal{C}}^{(i+1)}) \subset p_{x,y}(\bar{\mathcal{C}}^{(i)})$ and $s_{x,y}(\bar{\mathcal{C}}^{(i+1)}) \subset s_{x,y}(\bar{\mathcal{C}}^{(i)})$, $i \in [l-1]$ for any x, y , such that $0 \leq x < y \leq l$.

Lemma 1. *If $p_{x,y}(\bar{\mathcal{C}}^{(i+1)}) = p_{x,y}(\bar{\mathcal{C}}^{(i)})$ and $s_{x,y}(\bar{\mathcal{C}}^{(i+1)}) = s_{x,y}(\bar{\mathcal{C}}^{(i)})$, then for any $u_i \in \mathbb{F}_2$ the composite branch table $T_{x,y}$ constructed for $\bar{\mathcal{C}}(u_0^{i-1})$ is identical to the one constructed for $\bar{\mathcal{C}}(u_0^i)$, denoted $T'_{x,y}$, for the same received vector r_0^{l-1} .*

Proof. Since both section codes are identical for phases i and $i+1$, the CBTs have the same size. $p_{x,y}(\bar{\mathcal{C}}^{(i+1)}) = p_{x,y}(\bar{\mathcal{C}}^{(i)})$ implies that $\kappa = (K_{i,x}, \dots, K_{i,y-1}) \in p_{x,y}(\bar{\mathcal{C}}^{(i+1)})$. Hence, for any $u_i \in \mathbb{F}_2$ both $p_{x,y}(\bar{\mathcal{C}}^{(i)})/s_{x,y}(\bar{\mathcal{C}}^{(i)})$ and $p_{x,y}(\bar{\mathcal{C}}^{(i+1)})/s_{x,y}(\bar{\mathcal{C}}^{(i+1)})$ have cosets containing $u_i\kappa$, $u_i \in \mathbb{F}_2$. Hence, both $T_{x,y}$ and $T'_{x,y}$ contain the same values. \square

The above lemma suggests that one does not need to recompute the CBTs for those sections, where section codes do not change from phase i to phase $i+1$.

In what follows, we assume that

$$s_{x,z}(\mathcal{C}^{(i+1)}) = s_{x,z}(\mathcal{C}^{(i)}) \text{ and } s_{z,y}(\mathcal{C}^{(i+1)}) = s_{z,y}(\mathcal{C}^{(i)}). \quad (7)$$

Let us assume temporarily that $u_0^{i-1} = 0$.

Even if section codes do change, it is still possible to reuse some results obtained at prior phases. Let $k'_{i,x,y} = k'_{x,y}(\mathcal{C}^{(i)})$ and $k''_{i,x,y} = k''_{x,y}(\mathcal{C}^{(i)})$. First, observe that if $k''_{i+1,x,y} = k''_{i,x,y}$, but $k'_{i+1,x,y} < k'_{i,x,y}$, then $p_{x,y}(\mathcal{C}^{(i+1)})/s_{x,y}(\mathcal{C}^{(i+1)}) \subset p_{x,y}(\mathcal{C}^{(i)})/s_{x,y}(\mathcal{C}^{(i)})$, so that the corresponding CBT at phase $i+1$ can be obtained as a subvector of the CBT at phase i .

Second, we propose to implement maximization recursively, and keep all intermediate results. More specifically, we propose to rewrite (5) as

$$\begin{aligned} T_{x,y}[v].q &= \max_{w \in \mathbb{F}_2^{k''_{i,x,y}}} (T_{x,z}[a].q + T_{z,y}[b].q) \\ &= \max_{w_{k''_{i,x,y}-1}} \dots \max_{w_1} \max_{w_0} (T_{x,z}[a].q + T_{z,y}[b].q). \end{aligned} \quad (8)$$

Instead of storing in the CBT the final results of maximization in (8), we propose to keep the intermediate results of maximization for all w . These values can be arranged in a binary tree for each $v \in \mathbb{F}_2^{k''_{i,x,y}}$, so that a path from a root in this tree can be specified by values $w_{k''_{i,x,y}-1}, w_{k''_{i,x,y}-2}, \dots, w_0$. By maximization forest we denote the set of such trees obtained at some phase for a given section. The subtrees within a forest can be indexed by variables w, v .

We propose to use the above described maximization tree constructed at some phase i_0 to obtain CBTs for all $i \geq i_0$, where (7) holds. Let $i_1 > i_0$ be the smallest integer, where this does not hold.

Lemma 2. *Let $G''_{j,x,y} = \begin{pmatrix} G_{x,y}^{(00)} & G_{x,y}^{(01)} \end{pmatrix}$ and $G'_{j,x,y} = \begin{pmatrix} G_{x,y}^{(10)} & G_{x,y}^{(11)} \end{pmatrix}$ be the matrices obtained from (4) for code $\mathcal{C}^{(j)}$ for any j . If all matrices $G''_{i,x,y}$ are nested, so that $G''_{i+1,x,y}$ occupies top rows of $G''_{i,x,y}$ for any $i : i_0 \leq i < i_1$, then the maximization forest for phase $i : i_0 \leq i \leq i_1$ can be obtained by taking the subtrees of the trees in the forest constructed at phase i_0 , given by values w and v satisfying the equation*

$$(w \ v) \begin{pmatrix} G''_{i_0,x,y} \\ G'_{i_0,x,y} \end{pmatrix} = (\bar{w} \ \bar{v}) \begin{pmatrix} G''_{i,x,y} \\ G'_{i,x,y} \end{pmatrix}, \quad (9)$$

where $\bar{w} \in \mathbb{F}_2^{k''_{i_0,x,y}}$, $\bar{v} \in \mathbb{F}_2^{k'_{i_0,x,y}}$ denote the subtree indices in the forest at phase i_0 .

Proof. Assumption (7) ensures that the maximization trees for all i are obtained from the same values.

Observe that the matrices in (9) together with the generator matrices of $s_{x,z}(\mathcal{C}^{(i_0)})$ and $s_{z,y}(\mathcal{C}^{(i_0)})$ constitute generator matrices of $p_{x,y}(\mathcal{C}^{(i_0)})$ and $p_{x,y}(\mathcal{C}^{(i)})$, given by (4). Since $p_{x,y}(\mathcal{C}^{(i)}) \subset p_{x,y}(\mathcal{C}^{(i_0)})$, for any \bar{w}, \bar{v} there is a unique solution $(w \ v) = (\bar{w} \ \bar{v}) M_{i,x,y}$ of (9), where $M_{i,x,y}$ is a matrix, which can be constructed at the design time. Since $G_{i,x,y}$ occupies $k''_{i,x,y}$ top rows of $G''_{i_0,x,y}$, one obtains $w_j = \bar{w}_j$, $0 \leq j < k''_{i,x,y}$. Hence, for each \bar{v} the corresponding maximization tree indeed appears as a subtree indexed by w_j of the v -th maximization tree. \square

It is always possible to obtain $G''_{j,x,y}$ in the form required by Lemma 2 by applying elementary row operations.

2) *Handling prior decisions*: The SC decoder needs to take into account at phase i the values u_0^{i-1} . This reduces to decoding in cosets of section codes. The corresponding coset representative for section $[x, y]$ can be computed as a linear combination of subvectors $(K_{j,x} \ K_{j,x+1} \ \dots \ K_{j,y-1})$, $0 \leq j < i$, of the kernel.

Consider first the case of section $[x, y]$, where the maximization forest (8) is constructed from scratch. If there exists a solution of

$$(f_{j,x,y} \quad h_{j,x,y}) \begin{pmatrix} G_{i,x,y}^{(s)} \\ G'_{i,x,y} \end{pmatrix} = (K_{j,x} \quad K_{j,x+1} \quad \dots \quad K_{j,y-1}), \quad (10)$$

then the CBT for section $[x, y]$ contains at position $h_{j,x,y}$ the required coset representative. In this case we assume $h_{j,x',y'} = 0$ for all $x < x' < y' < y$. Otherwise, we assume $h_{j,x,y} = 0$. Given a vector of prior decisions u_0^{i-1} , we obtain the corresponding position offset at section $[x, y]$ as $h_{x,y} = \sum_{j=0}^{i-1} u_j h_{j,x,y}$, so that (8) becomes

$$T_{x,y}[v].q = \max_w (T_{x,z}[a + h_{x,z}].q + T_{z,y}[b + h_{z,y}].q). \quad (11)$$

The values $h_{x,y}$ are similar to partial sums, which arise in the SC decoder of Arikan polar codes.

For those sections, where the CBT is obtained by taking subtrees of (8), we need to check if there is a solution of

$$\omega_{j,x,y} \begin{pmatrix} G_{i_0,x,y}^{(s)} \\ G'_{i_0,x,y} \end{pmatrix} = (K_{j,x} \quad K_{j,x+1} \quad \dots \quad K_{j,y-1}). \quad (12)$$

If this equation does not have a solution for some $j < i$, the corresponding coset representatives have already been accounted for at smaller sections while constructing the CBT for section $[x, y]$ at phase i_0 . and we assume $\omega_{j,x,y} = 0$. Otherwise, the corresponding CBT can be obtained from the maximization forest constructed at phase $i_0 < i$ by taking entries

$$(w \quad v) = (\bar{w} \quad \bar{v}) M_{i,x,y} + \sum_{j=0}^{i-1} u_j \omega_{j,x,y}.$$

3) *Special trellises*: It can be seen that the result of (6) does not change if the same value is subtracted from all CBT entries at any section. Doing this enables one to compute some CBTs with smaller number of operations compared to (11). The following special cases were identified:

$$1) k'_{i,x,y} = k''_{i,x,y} = 1, \hat{G}_{i,x,y} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \tilde{G}_{i,x,y} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Let $a = \frac{T_{x,z}[0].q - T_{x,z}[1].q}{2}$, $b = \frac{T_{z,y}[0].q - T_{z,y}[1].q}{2}$. We propose to set $T_{x,y}[0].q = \text{sgn}(a) \text{sgn}(b) \min(|a|, |b|)$ and $T_{x,y}[1].q = -T_{x,y}[0].q$. If one can guarantee that $T_{x,z}[1].q = -T_{x,z}[0].q$, then it is possible to further simplify computation by setting $a = T_{x,z}[0].q$. Similar simplification applies to b . This trick allows one to construct section CBT using just 1 comparison operation, instead of 2 comparison and 4 summations for a straightforward implementation.

$$2) k'_{i,x,y} = 1, k''_{i,x,y} = 0, \hat{G}_{i,x,y} = \tilde{G}_{i,x,y} = (1). \text{ Using the same definitions as above, one obtains } T_{x,y}[0].q = (-1)^{h_{x,z}a} + (-1)^{h_{z,y}b}, T_{x,y}[1].q = -T_{x,y}[0].q.$$

$$3) k'_{i,x,y} = 0, k''_{i,x,y} = 2, \hat{G}_{i,x,y} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \tilde{G}_{i,x,y} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Using the same definitions as above, we propose to

$$\text{set } T_{x,y}[0].q = a + b, T_{x,y}[1].q = b - a, T_{x,y}[2].q = -T_{x,y}[1].q, T_{x,y}[3].q = -a.$$

- 4) If there is all-1 row in $G_{i,x,y}^{(s)}$, then one step of maximization can be avoided in (8) by taking the absolute values of the corresponding terms. This reduces the maximization tree construction complexity by a factor of 2.

Simplification tricks 1 and 2 together with the results in [14] fully establish the equivalence of the proposed approach and the min-sum SC algorithm for the case of $K = B_\mu \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}^{\otimes \mu}$, where B_μ is the bit reversal permutation matrix.

4) *Optimal sectionalization*: The total complexity of kernel processing with the proposed algorithm is given by

$$C = \sum_{i=0}^{l-1} (\delta_i + c_{i,0,l}),$$

where δ_i is the complexity of computing the final LLR from the obtained CBT, and $c_{i,x,y}$ is the complexity of construction of the CBT for section $[x, y]$ at phase i . In general, the former operation reduces to computing (6), i.e. costs 1 subtraction. However, if above described special trellises 1 or 2 arise at section $[0, l]$, then $\delta_i = 0$. Furthermore,

$$c_{i,x,y} = \begin{cases} m_{ixy}, & \text{if CBTs for subsections can be reused} \\ m_{ixy} + c_{i,x,z} + c_{i,z,y}, & \text{otherwise,} \end{cases}$$

where

$$m_{ixy} = \begin{cases} 0, & \text{if forest reuse is possible} \\ M_j, & \text{if type-}j \text{ special trellis is encountered,} \\ 2^{k'_{ixy} + k''_{ixy} - f_{ixy}} + 2^{k'_{ixy}} (2^{k''_{ixy} - f_{ixy}} - 1), & \text{otherwise,} \end{cases}$$

f_{ixy} is equal 1 if the above described 4-th simplification trick is applicable, and 0 otherwise, and M_j is the complexity of type- j special trellis, $1 \leq j \leq 3$. The first and second terms in the latter expression represent the number of summations and comparisons, respectively. For each section $[x, y]$ one should find splitting position z , which minimizes the overall complexity.

This can be done offline by the optimization algorithm given in [10]. Observe that sectionalization should be optimized jointly for all phases i .

IV. NUMERIC RESULTS

Table I presents processing complexity for some polarization kernels. Here K_l denotes a kernel of size l . For comparison, we present the complexity of the window-based and Viterbi processing algorithms. To the best of our knowledge, no other processing algorithm was published for these kernels. It can be seen that for kernels given in [15] the proposed approach has slightly higher complexity compared to the window processing algorithm. However, extension of the latter algorithm to kernels of size other than 2^μ is non-obvious, while the recursive trellis algorithm can be applied to any

TABLE I: Kernel processing complexity

Kernel K_l	$E(K_l)$	$\mu(K_l)$	State of the art			Proposed	
			Method	Add	Comp.	Add	Comp.
$K_{16}B_4$ [15]	0.51828	3.45	window	95	86	131	105
$K_{32}B_5$ [15]	0.521936	3.417	window	297	274	406	262
K_{32}^r [16]	0.52194	3.42111	Viterbi	4536	9072	355	191
K_{32}^{bch} [17]	0.53656	3.1221	Viterbi	99745	199490	28455	26161
K_{32}^{enbch} [18]	0.53656	3.1221	window	2864420		32183	29873
K_{20}^* [16]	0.506169	3.43827	Viterbi	7524	15054	2893	2001
K_{20} [16]	0.49943	3.64931	Viterbi	1866	3756	289	189
K_{24}^* [16]	0.51577	3.3113	Viterbi	9922	19860	1621	1207
K_{24} [16]	0.502911	3.61903	Viterbi	2102	4218	241	124

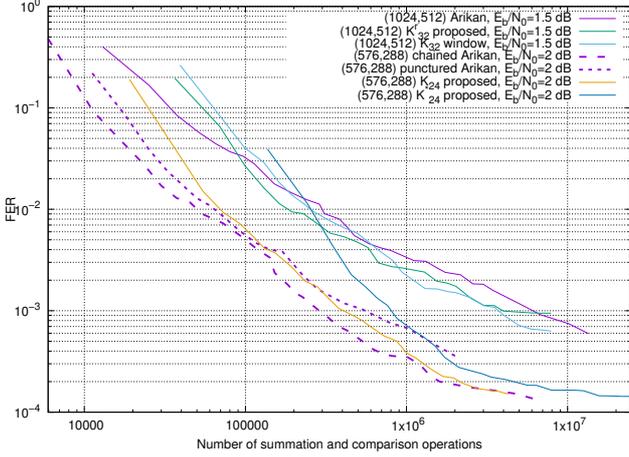


Fig. 1: Performance and SCL decoding complexity

kernel. Observe that the the proposed approach provides huge complexity reduction compared to the Viterbi algorithm.

Figure 1 presents the performance and complexity of the successive cancellation list decoding algorithm with various list size for some polar subcodes [19]. It can be seen that the proposed approach allows one to obtain better performance compared to the codes based on Arikan kernel with the same decoding complexity. More specifically, (1024, 512) code based on kernel K_{32}^r appears to be better than the code based on Arikan kernel starting from $FER = 3 \cdot 10^{-2}$, which corresponds to list size 3 and 7 for K_{32}^r and Arikan kernels, respectively. Polar subcode (576, 288) based on kernel K_{24} outperform punctured Arikan polar subcode starting from $FER = 3 \cdot 10^{-3}$, which corresponds to list size 8 and 20, respectively, and approaches the performance of the chained polar subcode [20]. High processing complexity of kernel K_{24}^* allows the corresponding polar subcode to outperform punctured polar subcode with Arikan kernel only at $FER = 4 \cdot 10^{-4}$, which corresponds to list size 14 and 192, respectively.

Figure 2 presents performance (solid lines) and average decoding complexity (dashed lines) of rate 1/2 polar subcodes and 5G LDPC codes. Sequential [21] and shuffled belief propagation [22] algorithms were used for decoding of polar subcodes and LDPC codes, respectively. Complexity is reported in terms of the number of summation and comparison operations for polar subcodes, and number of summations and calls to $\log \tanh(x/2)$ for LDPC codes. Polar subcodes were

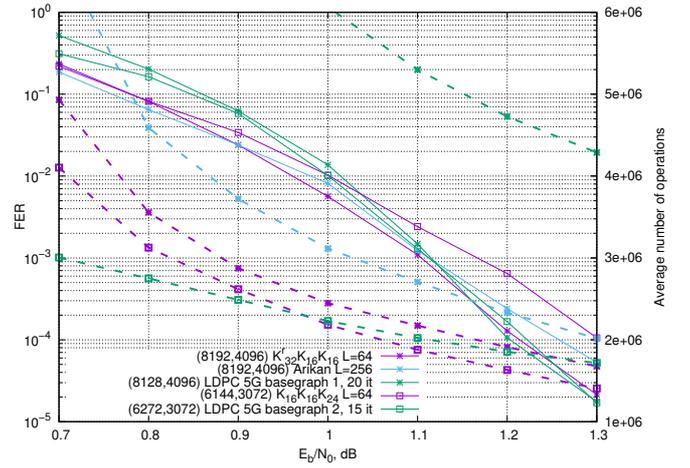


Fig. 2: Performance and decoding complexity of long codes

constructed using a mixture of kernels, as shown in the plot. It can be seen that polar subcodes provide almost the same performance as LDPC codes. At high SNR, the decoding complexity of polar subcodes appears to be lower compared to the corresponding LDPC codes. Observe also, that the code based on the Arikan kernel requires very large list size to obtain the performance comparable to the LDPC code. It also has higher decoding complexity compared to the code based on large kernels, and inferior performance at high SNR.

V. CONCLUSIONS

In this paper a novel processing algorithm for large polarization kernels was proposed. It relies on extensive reuse of the intermediate results arising in the recursive maximum likelihood decoding algorithm for the codes generated by submatrices of the polarization kernel. The proposed algorithm can be applied to kernels of arbitrary dimension, and has much lower complexity compared to the Viterbi algorithm. Derivation of a processor for a given kernel according to the proposed method involves matrix manipulations, which should be performed once at the design time. Actual kernel processing reduces to summation and comparison of the elements of some arrays, where the indices of the operands are obtained as XOR of some pre-computed values and partial sums given by the SC algorithm decisions. The obtained algorithm is fully deterministic.

The proposed algorithm enables polar subcodes with well-designed large kernels to provide better performance/complexity tradeoff compared to the codes based on Arikan kernel as well as LDPC codes. However, finding large kernels with good polarization properties and low processing complexity remains, in general, an open problem.

ACKNOWLEDGEMENT

This work is partially supported by the Ministry of Science and Higher Education of Russian Federation, passport of goszadanie no. 2019-0898.

REFERENCES

- [1] E. Arkan, "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] 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.
- [3] V. Bioglio and I. Land, "On the marginalization of polarizing kernels," in *Proceedings of International Symposium on Turbo Codes and Iterative Information Processing*, 2018.
- [4] V. Miloslavskaya and P. Trifonov, "Sequential decoding of polar codes with arbitrary binary kernel," in *Proceedings of IEEE Information Theory Workshop*. Hobart, Australia: IEEE, 2014, pp. 377–381.
- [5] 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.
- [6] 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.
- [7] 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.
- [8] V. Bioglio, F. Gabry, I. Land, and J. C. Belfiore, "Multi-kernel polar codes: Concept and design principles," *IEEE Transactions on Communications*, vol. 68, no. 9, pp. 5350–5362, 2020.
- [9] P. Trifonov, "Binary successive cancellation decoding of polar codes with Reed-Solomon kernel," in *Proceedings of IEEE International Symposium on Information Theory*. Honolulu, USA: IEEE, 2014, pp. 2972 – 2976.
- [10] T. Fujiwara, H. Yamamoto, T. Kasami, and S. Lin, "A trellis-based recursive maximum-likelihood decoding algorithm for binary linear block codes," *IEEE Transactions On Information Theory*, vol. 44, no. 2, March 1998.
- [11] E. Blokh and V. Zyablov, "Coding of generalized concatenated codes," *Problems of Information Transmission*, vol. 10, no. 3, pp. 45–50, 1974.
- [12] P. Trifonov, "Efficient design and decoding of polar codes," *IEEE Transactions on Communications*, vol. 60, no. 11, pp. 3221 – 3227, November 2012.
- [13] H. Griesser and V. R. Sidorenko, "A posteriori probability decoding of nonsystematically encoded block codes," *Problems of Information Transmission*, vol. 38, no. 3, 2002.
- [14] P. Trifonov, "Trellis-based decoding techniques for polar codes with large kernels," in *Proceedings of IEEE Information Theory Workshop*, 2019, invited paper.
- [15] G. Trofimiuk and P. Trifonov, "Window processing of binary polarization kernels," *IEEE Transactions on Communications*, 2021, to appear.
- [16] G. Trofimiuk, "A search method for large polarization kernels," in *Proceedings of IEEE International Symposium on Information Theory*, 2021.
- [17] E. Moskovskaya and P. Trifonov, "Design of BCH polarization kernels with reduced processing complexity," *IEEE Communications Letters*, vol. 24, no. 7, pp. 1383–1386, July 2020.
- [18] F. Abbasi and E. Viterbo, "Large kernel polar codes with efficient window decoding," *IEEE Transactions On Vehicular Technology*, vol. 69, no. 11, November 2020.
- [19] P. Trifonov, "On construction of polar subcodes with large kernels," in *Proceedings of IEEE International Symposium on Information Theory*, Paris, France, 2019.
- [20] —, "Randomized chained polar subcodes," in *Proceedings of IEEE Wireless Communications and Networking Conference Workshops*. Barcelona, Spain: IEEE, 2018, pp. 292–297.
- [21] —, "A score function for sequential decoding of polar codes," in *Proceedings of IEEE International Symposium on Information Theory*, Vail, USA, 2018.
- [22] J. Zhang and M. P. Fossorier, "Shuffled iterative decoding," *IEEE Transactions on Communications*, vol. 53, no. 2, February 2005.