Rate-Compatible Polar Codes Based on Polarization-Driven Shortening

Robert M. Oliveira and Rodrigo C. de Lamare, *Senior Member, IEEE*

***Abstract*—This letter presents a polarization-driven (PD) shortening technique for the design of rate-compatible polar codes. The proposed shortening strategy consists of reducing the generator matrix by relating its row index with the channel polarization index. We assume that the shortened bits are known by both the encoder and the decoder and employ successive cancellation for decoding the shortened codes constructed by the proposed PD technique. A performance analysis is then carried out based on the spectrum distance. Simulations show that the proposed PD-based shortened polar codes outperform existing shortened polar codes.**

***Index Terms*—Channel polarization, shortening, polar codes, 5G systems, IoT networks.**

I. Introduction

POLAR codes, proposed by Arikan, are low-complexity capacity-achieving codes based on the phenomenon called channel polarization [1]. A typical construction of conventional polar codes is based on the Kronecker product, which is restricted to the lengths 2^l $(l = 1, 2, \ldots)$. Polar codes with arbitrary lengths can be obtained by shortening or puncturing [3], which will be required for 5G scenarios, where code lengths ranging from 420 to 1920 bits with various rates will be adopted [4] and [5]. Shortened and punctured polar codes can be decoded in a similar way to conventional polar codes.

Various shortening and puncturing methods for polar codes have been proposed in the literature [6]–[17] and evaluated with successive cancellation (SC) or belief propagation (BP) decoding. In [7]–[10] puncturing methods have been reported using BP decoding based on optimization techniques employing retransmission schemes such as Hybrid Automatic Repeat reQuest (HARQ). Different properties of punctured codes have been explored: minimum distance, exponent binding, stop tree drilling, and the reduced generating matrix method [9]–[12]. Schemes that depend on the analysis of density evolution were proposed in [13] and [14]. On the other hand, shortening methods have been studied with SC decoding. With shortening techniques, we freeze a bit channel that receives a fixed zero value. The decoder, however, uses a plus infinity log-likelihood ratio (LLR) for that code bit as it is often assumed that this value is known. The study in [16] proposed a search algorithm to jointly optimize the shortening patterns and the values of the shortened bits. The work in [17] devised a simple shortening method, reducing the generator matrix based on the weight of the columns (CW). In [6] the reversal quasi-uniform puncturing scheme (RQUP) for reducing the generator matrix has been proposed.

Manuscript received July 9, 2018; accepted July 29, 2018. This work was financially supported by CAPES, CNPq, Faperj and CGI, Brazil. The associate editor coordinating the review of this paper and approving it for publication was C. Feng. *(Corresponding author: Rodrigo C. de Lamare.)*

The authors are with the Centre for Telecommunications Studies, Pontifical Catholic University of Rio de Janeiro, Rio de Janeiro 22451-900, Brazil (e-mail: rbtmota@gmail.com; delamare@cetuc.puc-rio.br).

Digital Object Identifier 10.1109/LCOMM.2018.2863375

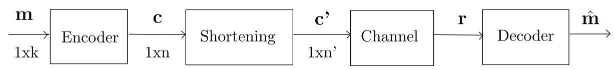

Fig. 1. System model.

In this letter, we propose a polarization-driven (PD) shortening technique based on the Gaussian Approximation (GA) [2], where the channel polarization index determines the channel shortening patterns. In particular, we describe the design of rate-compatible polar codes using the PD method and its application to fifth generation (5G) wireless system scenarios. We also carry out an analysis of the proposed PD method using the Spectrum Distance (SD) [6] and assess the performance of design examples via simulations.

II. System Model and Problem Statement

Fig.1 shows a block diagram of the polar coding system considered in this letter.

In this system, $\mathbf{m}$ is the binary message with k bits that is transmitted. The $n \times n$ generator matrix $\mathbf{G}$ encodes the message $\mathbf{m}$ and produces the codeword $\mathbf{c}$ with n bits. With an appropriate shortening technique, the codeword $\mathbf{c}$ has its length reduced to n', resulting in the shortened codeword $\mathbf{c'}$, where $2^{l-1} < n' < 2^l$, where l is an integer that defines the levels in the polarization tree, $l = \log_2 n$. The shortened codeword $\mathbf{c'}$ is then transmitted over a channel with additive white Gaussian noise (AWGN), resulting in the received vector $\mathbf{r} = \mathbf{c'} + \mathbf{w}$, where $\mathbf{w}$ is the vector corresponding to the noise. In the decoding step, the decoding algorithm observes $\mathbf{r}$ in order to estimate $\mathbf{m}$. We call it an estimated message $\hat{\mathbf{m}}$, and if $\mathbf{m} = \hat{\mathbf{m}}$ we say that the message has been fully recovered. The problem we are interested in solving is how to design shortened codes with the best performance.

III. Polar Coding System

Let $W : X \rightarrow Y$ denote a binary discrete memoryless channel (B-DMC), with input alphabet $X = \{0, 1\}$, output alphabet Y, and the channel transition probability $W(y|x)$, $x \in X$, $y \in Y$. The mutual information of the channel with equiprobable inputs, or symmetric capacity, is defined by [1]

$$I(W) = \sum_{y \in Y} \sum_{x \in X} \frac{1}{2} W(y|x) log \frac{W(y|x)}{\frac{1}{2}W(y|0) + \frac{1}{2}W(y|1)} \quad (1)$$

and the corresponding reliability metric, the Bhattacharyya parameter, is described by [1]

$$Z(W) = Z_0 = \sum_{y \in Y} \sqrt{W(y|0)W(y|1)} \quad (2)$$

Applying the channel polarization transform for n independent uses of W we obtain after channel combining and splitting operations the group of polarized channels $W_n^{(i)} : X \rightarrow Y \times X^{i-1}$, $i = 1, 2, \ldots, n$. The channel polarization index $Z(W_n)$

over AWGN channels is calculated using the GA method [2] with the following recursions:

$$\begin{cases} Z(W_n^{(2i-1)}) = \phi^{-1}(1-(1-\phi(Z(W_{n/2}^{(i)}))^2) \\ Z(W_n^{(2i)}) = 2Z(W_{n/2}^{(i)}), \end{cases} \tag{3}$$

where

$$\phi(x) \triangleq \begin{cases} \exp(-0.4527x^{(0.86)} + 0.0218) & \text{if } 0 < x \leq 10 \\ \sqrt{\frac{\pi}{x}}(1-\frac{10}{7x})\exp(\frac{-x}{4}) & \text{if } x > 10 \end{cases} \tag{4}$$

In general, we use in the notation $\mathbf{a}_1^n$ to designate a vector $(a_1, a_2, \ldots, a_n)$ and $|\mathbf{a}_1^n|$ to refer to its cardinality. The channel polarization theorem [1] states that $I(W_n^{(i)})$ converges to either 0 (completely noisy channels) or 1 (noiseless channels) as $n \to \infty$ and the fraction of noiseless channel tends to $I(W)$, while polarized channels converge to either $Z(W_n^{(i)}) = 1$ or $Z(W_n^{(i)}) = 0$. The vector $\mathbf{c} = (\mathbf{u}_A, \mathbf{u}_{A^c})$, where $\mathbf{m} = \mathbf{u}_A$, for some $A \subset \{1, \ldots, n\}$ denotes the set of information bits and $A^c \subset \{1, \ldots, n\}$ denotes the set of frozen bits. We select the $|\mathbf{m}|$ channels to transmit information bits such that $Z(W_n^{(i)}) \leq Z(W_n^{(j)})$. For encoding, a codeword is generated by $\mathbf{c} = c_1^n = \mathbf{m}\mathbf{B}_n\mathbf{G}_2^{\otimes l}$, where $\mathbf{m}$ is the information sequence, $\mathbf{B}_n$ is the bit-reversal permutation matrix, $\otimes l$ is the l-th Kronecker power and $\mathbf{G}_2 = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ is the kernel matrix. We adopt the SC decoder to estimate the information bits as [1]

$$\hat{u}_i = \arg\max_{u_i \in \{0,1\}} W_n^{(i)}(y_1^n, u_1^{i-1}|u_i), \quad i \in A \tag{5}$$

IV. Proposed Polarization-Driven Shortening

In this section, we detail the proposed PD shortening technique and show how to calculate the polarization channels. Polar codes are nonuniversal [1], i.e., different polar codes are generated depending on the specified value of the signal-to-noise ratio (SNR), known as design-SNR. The design-SNR choice is critical for ensuring good performance in all SNRs of interest and in this work we adopt the design-SNR equal to zero.

The purpose of shortening is to reduce the size of the generator matrix $\mathbf{G}_n$ from $n \times n$ to $n' \times n'$, such that $n' < n$. In particular, the size reduction is obtained by eliminating rows and columns of the matrix $\mathbf{G}_n$. Consider the shortened vector $\mathbf{p}$ which contains the indexes of the rows of the matrix $\mathbf{G}_n$ to be shortened, where $|\mathbf{p}| = n - n'$ indicates its number of elements. Consider the set $\mathbf{G}_n^r$ of all shortened matrices, $r = (1, \ldots, \binom{n}{n'})$. We define

$$\mathbf{G}_n^* = \arg\min_{\mathbf{G}^r} \text{ BER}, \tag{6}$$

where the bit error rate (BER) is adopted. Many works resort to exhaustive searches with an optimization method to determine $\mathbf{G}_n^*$, its the optimal shortened generator matrix. In contrast to prior work, we propose the PD shortening technique for computing $\mathbf{p}$, where the channels with the lowest polarization indexes are eliminated.

TABLE I

Polarization Vector $\mathbf{b}$ for $n = 8$

b	0.992	0.882	0.915	0.578	0.938	0.639	0.715	0.000
index	1	2	3	4	5	6	7	8
After sorting								
a	0.000	0.578	0.639	0.715	0.882	0.915	0.938	0.992
k	8	4	6	7	2	3	5	1

Using the notation in [1], for $n = 8$ we have l stages of polarization (3) are

- stage 1:
  $Z(W^{+})$ and $Z(W^{-})$
- stage 2:
  $Z(W^{++})$, $Z(W^{-+})$, $Z(W^{+-})$ and $Z(W^{--})$
- stage 3:
  $Z(W^{+++})$, $Z(W^{-++})$, $Z(W^{+-+})$, $Z(W^{--+})$, $Z(W^{++-})$, $Z(W^{-+-})$, $Z(W^{+--})$ and Z(W^{---}).

The channels $(W^{+++}, W^{-++}, W^{+-+}, W^{--+}, W^{++-}, W^{-+-}, W^{+--}, W^{---})$ can be written with $(W_0, W_1, W_2, W_3, W_4, W_5, W_6, W_7)$. We define the polarization vector as

$$\mathbf{b} \triangleq \left[\, Z(W_0); Z(W_1); \ldots; Z(W_{n-1}) \,\right]^T. \tag{7}$$

As an example for stage 3 with normalized values, we have $\mathbf{b} = [0.992, 0.882, 0.915, 0.578, 0.938, 0.639, 0.715, 0.000]^T$. The key idea of the proposed PD method is to remove in the generator matrix $\mathbf{G}_n$ the rows that correspond to the channels with smallest values of polarization.

These channels can be obtained by sorting the polarization vector $\mathbf{b}$. The goal of sorting is to determine a permutation $k(1)k(2)\ldots k(n)$ of the indexes $\{1, 2, \ldots, n\}$ that will organize the entries of the polarization vector $\mathbf{b}$ in increasing order [18]:

$$Z(W_{k(1)}) \leq Z(W_{k(2)}) \leq \ldots \leq Z(W_{k(n)}) \tag{8}$$

Consider the sort function $[\mathbf{a}, \mathbf{k}] = \text{sort}(\mathbf{b})$ which implements (8), where $\mathbf{a}$ lists the sorted $\mathbf{b}$ and $\mathbf{k}$ contains the corresponding indexes of $\mathbf{a}$. Table 1 shows an example of the polarization vector $\mathbf{b}$ for $n = 8$, sorting vector $\mathbf{a}$ and the new index $\mathbf{k}$.

The vector $\mathbf{k} = [8, 4, 6, 7, 2, 3, 5, 1]$ contains the indexes of the polarization values of the channels in increasing order, which are used to obtain the shortening vector $\mathbf{p}$ of the proposed PD method:

$$\mathbf{p} = [k(1), \ldots, k(n-n')], \tag{9}$$

with $n - n'$ being the shortening length. In Algorithm 1 we have included a pseudo-code of the proposed PD method with details of the size reduction of the generator matrix $\mathbf{G}_n$ and the shortening of the channels with the lowest polarization values.

We consider now an example with shortened polar codes with length $n' = 5$. For the shortening of $\mathbf{G}_8$ to $\mathbf{G}_5$, the channels with the lowest polarization rank values are W_8, W_4 and W_6, the shortening vector is $\mathbf{p} = (8, 4, 6)$ and $|\mathbf{p}| = 3$. The 1st element of $\mathbf{p}$ is 8, which results in the deletion of the 8th column and the 8th row of $\mathbf{G}_8$. The 2nd element of $\mathbf{p}$ is 4, which requires the elimination of the 4th column and the 4th row. At last, the 3rd element of the $\mathbf{p}$ is 6, which requires the deletion of the 6th column and the 6th row. The matrix $\mathbf{G}_8$

Algorithm 1 Proposed PD algorithm

```
Given a shortened codeword with length n'
Use G_n as the base matrix
Index each column by {1, 2, ..., n}
Index each row by {1, 2, ..., n}
Calculate the polarization channel vector b for n
Calculate [a, k] = sort(b)
Calculate the shortening vector p = [k(1), ..., k(n − n')]
for y = 1 to |p| do
    r_min ← p(y)
    Delete row from G_n with index r_min
    Delete column from G_n with index r_min
end for
```

with the indication of the deletions and the resulting generator matrix $\mathbf{G}_5$ are given by

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 1\,0\,0\,0\,0 \\ 1\,1\,0\,0\,0 \\ 1\,0\,1\,0\,0 \\ 1\,0\,0\,1\,0 \\ 1\,0\,1\,1\,1 \end{pmatrix}$$

As the code has been shortened, the reliability of the bit channels changes and the information set should change accordingly. The $Z(W)$ parameters of the polarized channels shortened are smaller than those of the original polarized channels, we consider in this letter that the order of channel polarization given by (8) does not change after shortening. The shortened codeword **c'** generated with $\mathbf{G}_{n'}$, which contains the bits of the binary message $\mathbf{m} = \mathbf{u}_A$ such that $Z(W_{n'}^{(i)}) \leq Z(W_{n'}^{(j)})$ for all $i \in A$, $j \in A^c$ and $\mathbf{u}_{A^c} = (u_i : i \in A^c | u_i = 0)$ is then transmitted over a channel. The PD shortening method assumes that the channels remaining after shortening keep their polarization ordering. Therefore, any rate for the shortening of the polar code can be arbitrarily chosen as in conventional polar codes, where the channels with the smallest polarization indexes are chosen for the information bits.

V. Analysis

The work in [19] examined maximum likelihood (ML) decoding for polar codes and demonstrated that systematic coding yields better BER performance than non-systematic coding with the same FER performance for both encoding schemes. However, ML decoding is quite costly since it compares all possible codewords for a given polar code using the Hamming distance.

We employ the SD that has been studied in [6] for analysis due to its lower computational cost than ML decoding [19], and its suitability to compare the performance of the proposed PD and existing techniques. This metric is based on the channel polarization tree and the number of paths on the tree with the same number of zeros or ones, respectively. The channel polarization tree is obtained by the recursive process of polarization channel construction [1]. The branch of the tree obtained by $Z(W_n^{(2i-1)})$ in (3) is labeled 1 and the branch obtained by $Z(W_n^{(2i)})$ in (3) is labeled 0. Each tree path refers to a polarized channel W_i and to a row of the generator matrix $\mathbf{G}_n$.

The SD for path weight for the 1s is given by [6]

$$d = \sum_{(k=0)}^{l} P_1(l,k,Q)k = \sum_{(k=0)}^{l} \frac{H_n^{(k)}}{n} k, \qquad (10)$$

where $P_1(l,k,Q) = \frac{H_n^{(k)}}{n}$ is the probability of path weight k with $Q = |\mathbf{p}|$ bits shortening and l refers to the levels in the polarization tree.

The SD for path weight for the 0s is given by [6]

$$\lambda = \sum_{(r=0)}^{l} P_0(l,r,Q)r = \sum_{(r=0)}^{l} \frac{C_n^{(r)}}{n} r, \qquad (11)$$

where $P_0(l,k,Q) = \frac{C_n^{(r)}}{n}$, where $C_l^{(r)} = \binom{l}{r}$ is the probability of path weight r with $Q = |\mathbf{p}|$ bits shortening and l refers to the levels in the polarization tree. The SD for path weight for the 0s used as the main metric to evaluate the performance of the proposed and existing shortening techniques.

The term $C(X) = \sum_{(r=0)}^{l} C_l^{(r)} X^r$ describes the number of paths with a given number of zeros, or alternatively $C(X) = \sum_{(i=1:n)} X^{\mathrm{Pb_i}}$, Pb is the number of zeros of each path. As an example, for a $\mathbf{G}_{16}$, we have $C(X) = X^0 + 4X^1 + 6X^2 + 4X^3 + X^4$, one path with no zero, four paths with 1 zero, six paths with 2 zeros, 4 paths with 3 zeros and one path with 4 zeros, the $\lambda = \frac{1\cdot0+4\cdot1+6\cdot2+4\cdot3+1\cdot4}{16} = 2$.

Given a shortening procedure, the $C(X)$ is updated by removing the paths cut by shortening, each path corresponds to a channel, which in turn corresponds to a row (and column) in the generator matrix $\mathbf{G}$. For $\mathbf{G}_{12}$ with $\mathbf{p} = (14, 15, 16)$, updating $C(X) = 2X^1 + 5X^2 + 4X^3 + 1X^4$ and new $\lambda = \frac{2\cdot1+5\cdot2+4\cdot3+1\cdot4}{16} = 1.75$, always less than the previous value λ.

The set of shortened paths have different weights for each shortening method, existing shortening ($\mathbf{p}_{\mathrm{ep}}$) such that $\mathbf{p}_{\mathrm{PD}} \neq \mathbf{p}_{\mathrm{ep}}$, $|\mathbf{p}_{\mathrm{PD}}| = |\mathbf{p}_{\mathrm{ep}}| = \mathrm{y} > 0$, $\exists$ y we have

$$\sum_{i=1}^{\mathrm{y}} C_{\mathrm{PD(i)}}(X) \leq \sum_{i=1}^{\mathrm{y}} C_{\mathrm{CW(i)}}(X), \qquad (12)$$

with $C_{\mathrm{PD(i)}}(X)$ for PD technique and $C_{\mathrm{CW(i)}}(X)$ for CW technique. We remark that there is a value of y from which the shortened channels will be different. Expanding the above equation and assuming that one chosen path of the PD set is different from that of the CW set, we have

$$\sum_{i=1}^{\mathrm{y}-1} C_{\mathrm{PD(i)}}(X) + \alpha X^{\mathrm{y}} < \sum_{i=1}^{\mathrm{y}-1} C_{\mathrm{CW(i)}}(X) + \beta X^{\mathrm{y}}, \qquad (13)$$

where α and β are integer numbers. For small values of y, where the channels shortened by either method will be the same, we have the equality in (12). We then exploit the fact that $\alpha < \beta$, which yields

$$\alpha X^{\mathrm{y}} < \beta X^{\mathrm{y}}, \qquad (14)$$

proving the inequality in (12).

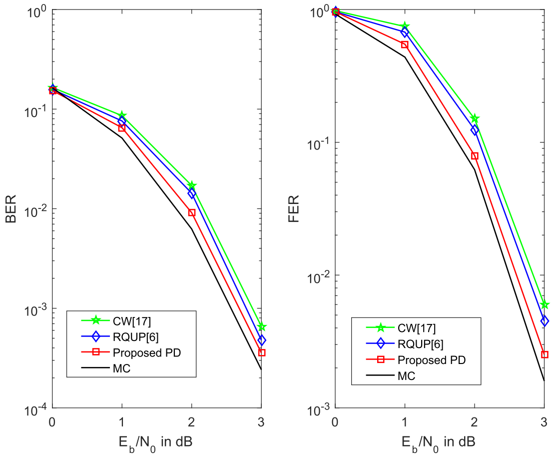

Fig. 2. BER and FER performaces of rate-compatible polar Code n'=480 k=256.

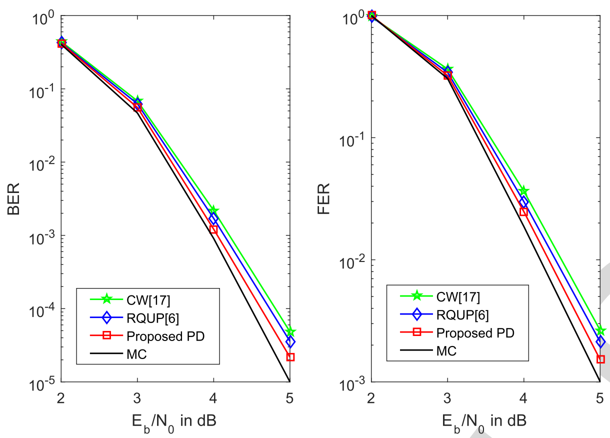

Fig. 3. BER and FER performaces of rate-compatible polar Code n'=1920 k=1600.

VI. Simulations

In this section, we present simulations of rate-compatible polar codes with shortening and a system equipped with the SC decoder, as described for the Internet of Things (IoT) and the Enhanced Mobile Broadband (eMBB) 5G scenarios [3] and [5], which require the use of short to moderate block lengths. We measure the BER and the frame error rate (FER) against the SNR, defined as the ratio of the bit energy, E_b, and the power spectral density, N_0, in dB. In the first example, we consider an IoT scenario with $n = 512$, $n' = 480$ and $k = 256$. In particular, we have reduced the $\mathbf{G}_{512}$ matrix to the $\mathbf{G}_{480}$ matrix using the proposed PD shortening technique. For comparison purposes, we have also included the curves associated with the best performing existing methods CW and RQUP, as can be observed in Fig. 2. The results in Fig. 2 show that the proposed PD technique outperforms the RQUP and the CW techniques by up to 0.25 dB for the same BER and FER performances, and approaches the performance of the mother code (MC) with $n = 512$ and $k = 256$.

In the second example, for eMBB scenario with $n = 2048$, $n' = 1920$ and $k = 1600$. The results in Fig. 3 show that the proposed PD technique outperforms the RQUP and the CW techniques by up to 0.20 dB for the same BER and FER performances, and approaches the performance of the mother code (MC) with $n = 2048$ and $k = 1600$.

TABLE II
Spectrum Distance for Figs. 2 and 3

SDC	Proposed PD	RQUP[6]	CW[17]
Fig. 2	4.53	4.46	4.43
Fig. 3	5.46	5.44	5.43

In the Table II we compare for each simulation the SD values obtained for each curve. Note that the SD of the proposed model PD has a higher value than the CW and RQUP techniques.

VII. Conclusion

We have proposed the PD shortening method, which is based on the channel polarization index, and can bring a performance improvement in shortened polar codes as compared to existing shortening methods in the literature. The use of the spectrum distance as a benchmark for performance comparison has been shown as a valuable tool to indicate the best shortening strategy, while requiring a low computational complexity.

References

[1] 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.

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

[3] V. Bioglio, F. Gabry, and I. Land. (2017). "Low-complexity puncturing and shortening of polar codes." [Online]. Available: https://arxiv.org/abs/1701.06458

[4] A. Osseiran *et al.*, "Scenarios for 5G mobile and wireless communications: The vision of the METIS project," *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 26–35, May 2014.

[5] J. S. Ferreira *et al.*, "GFDM frame design for 5G application scenarios," *J. Commun. Inf. Syst.*, vol. 32, no. 1, 2017, doi: 10.14209/jcis.2017.6.

[6] K. Niu *et al.* (2016). "Rate-compatible punctured polar codes: Optimal construction based on polar spectra." [Online]. Available: https://arxiv.org/abs/1612.01352

[7] S. N. Hong, D. Hui, and I. Marić. (2015). "Capacity-achieving rate-compatible polar codes." [Online]. Available: https://arxiv.org/abs/1510.01776

[8] H. Saber and I. Marsland, "An incremental redundancy Hybrid ARQ scheme via puncturing and extending of polar codes," *IEEE Trans. Commun.*, vol. 63, no. 11, pp. 3964–3973, Nov. 2015.

[9] K. Chen, K. Niu, and J. Lin, "A hybrid ARQ scheme based on polar codes," *IEEE Commun. Lett.*, vol. 17, no. 10, pp. 1996–1999, Oct. 2013.

[10] K. Niu, K. Chen, and J.-R. Lin, "Beyond turbo codes: Rate-compatible punctured polar codes," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2013, pp. 3423–3427.

[11] D. Shin, S.-C. Lim, and K. Yang, "Design of length-compatible polar codes based on the reduction of polarizing matrices," *IEEE Trans. Commun.*, vol. 61, no. 7, pp. 2593–2599, Jul. 2013.

[12] A. Eslami and H. Pishro-Nik, "A practical approach to polar codes," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul./Aug. 2011, pp. 16–20.

[13] A. Eslami and H. Pishro-Nik, "On finite-length performance of polar codes: Stopping sets, error floor, and concatenated design," *IEEE Trans. Commun.*, vol. 61, no. 3, pp. 919–929, Mar. 2013.

[14] J. Kim, J.-H. Kim, and S.-H. Kim, "An Efficient search on puncturing patterns for short polar codes," in *Proc. Int. Conf. Inf. Commun. Technol. Converg. (ICTC)*, Oct. 2015, pp. 182–184.

[15] L. Zhang, Z. Zhang, X. Wang, Q. Yu, and Y. Chen, "On the puncturing patterns for punctured polar codes," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun./Jul. 2014, pp. 121–125.

[16] V. Miloslavskaya, "Shortened polar codes," *IEEE Trans. Inf. Theory*, vol. 61, no. 9, pp. 4852–4865, Sep. 2015.

[17] R. Wang and R. Liu, "A novel puncturing scheme for polar codes," *IEEE Commun. Lett.*, vol. 18, no. 12, pp. 2081–2084, Dec. 2014.

[18] D. E. Knuth, *The Art of Computer Programming*, vol. 3. Reading, MA, USA: Addison-Wesley, 1998.

[19] Z. Liu, K. Chen, K. Niu, and Z. He, "Distance spectrum analysis of polar codes," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2014, pp. 490–495.