A Performance Comparison of Polar Codes and Reed-Muller Codes

Erdal Arıkan, *Senior Member, IEEE*

***Abstract*— Polar coding is a code construction method that can be used to construct capacity-achieving codes for binary-input channels with certain symmetries. Polar coding may be considered as a generalization of Reed-Muller (RM) coding. Here, we demonstrate the performance advantages of polar codes over RM codes under belief-propagation decoding.**

***Index Terms*— Polar codes, Reed-Muller (RM) codes, channel coding, forward error correction.**

I. Introduction

POLAR coding is a code construction method that can achieve the capacity of symmetric binary-input discrete memoryless channels such as the binary symmetric channel (BSC) and binary erasure channel (BEC). This technique was introduced and theoretically analyzed in [1]. The aim of this letter is to give some experimental results to show the practical utility of polar coding. Specifically, we demonstrate how to use the polar coding idea to improve the performance of Reed Muller (RM) codes [2], [3] without increasing their encoding and decoding complexity.

II. RM codes

Let $G_{RM}(n,n)$ denote the generator matrix of an nth order RM code of block-length $N = 2^n$. Using the well-known Plotkin construction for RM codes, we may take

$$G_{RM}(n,n) = F^{\otimes n} \tag{1}$$

where $F = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ and $F^{\otimes n}$ denotes the nth tensor power of F. The rth order RM code RM(r,n) can then be defined as the linear code with generator matrix $G_{RM}(r,n)$ which is obtained by taking the rows of $G_{RM}(n,n)$ with Hamming weights $\geq 2^{n-r}$. For example, $G_{RM}(3,3)$ is given by

$$G_{RM}(3,3) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \tag{2}$$

and RM$(1,3)$ is the code with generator matrix

$$G_{RM}(1,3) = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \tag{3}$$

III. Polar codes

For any $N = 2^n$, $n \geq 1$, and $1 \leq K \leq N$, an (N,K) polar code is a block code whose generator matrix $G_P(N,K)$ is a $K \times N$ submatrix of $F^{\otimes n}$ constructed in accordance with the following procedure. First, we compute the vector $z_N = (z_{N,1}, \ldots, z_{N,N})$ through the recursion

$$z_{2k,j} = \begin{cases} 2\, z_{k,j} - z_{k,j}^2 & \text{for } 1 \leq j \leq k \\ z_{k,j-k}^2 & \text{for } k+1 \leq j \leq 2k \end{cases} \tag{4}$$

for $k = 1, 2, 2^2, \ldots, 2^{n-1}$, starting with $z_{1,1} = 1/2$. Next, we form a permutation $\pi_N = (i_1, \ldots, i_N)$ of the set $(1, \ldots, N)$ so that, for any $1 \leq j < k \leq N$, the inequality $z_{N,i_j} \leq z_{N,i_k}$ is true. The generator matrix $G_P(N,K)$ for an (N,K) polar code is defined as the submatrix of $F^{\otimes n}$ consisting of rows with indices $i_1, \ldots, i_K$. It is easy to see that the computational complexity of this code construction method is $\mathcal{O}(N \log N)$.

As an example, consider the case $n = 3$. Then, we have z_8 =(0.996, 0.684, 0.809, 0.121, 0.879, 0.191, 0.316, 0.004), which gives $\pi_8 = (8,4,6,7,2,3,5,1)$. Thus, an $(N,K) = (8,5)$ polar code has the generator matrix

$$G_P(8,5) = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \tag{5}$$

Since the ranking of the rows of $F^{\otimes 3}$ by π_8 happens to coincide with the ranking of them by their Hamming weights, there is no difference between polar coding and RM coding for $n = 3$. The same holds also for $n = 4$. However, for $n = 5$, we have π_{32} = (32, 16, 24, 28, 30, 31, 8, 12, 20, 14, 22, 26, 15, 23, 27, 4, 29, 6, 10, 7, 18, 11, 19, 13, 21, 2, 25, 3, 5, 9, 17, 1), while the weights of the corresponding rows are (32, 16, 16, 16, 16, 16, 8, 8, 8, 8, 8, 8, 8, 8, 8, 4, 8, 4, 4, 4, 4, 4, 4, 4, 4, 2, 4, 2, 2, 2, 2, 1). Thus, the polar code with parameter $(N,K) = (32,16)$ differs from the RM code with the same parameter in that it employs a weight-4 row of $F^{\otimes 5}$ and leaves out a weight-8 row. The non-equivalence of RM codes and polar codes becomes commonplace as the code order n increases and significant performance advantages begin to emerge in favor of polar codes, as we will show in the next section by experimental results.

As Forney [4] showed, RM codes can be regarded as *codes on graphs*, and, hence, decoded by BP decoders. Since polar

Manuscript received January 4, 2008. The associate editor coordinating the review of this letter and approving it for publication was A. Haimovich. This work was supported in part by The Scientific and Technological Research Council of Turkey (TÜBİTAK) under project no 107E216, and in part by the EC under FP7 Network of Excellence project NEWCOM++.

E. Arıkan is with the Department of Electrical-Electronics Engineering, Bilkent University, Ankara, 06800, Turkey (e-mail: arikan@ee.bilkent.edu.tr).

Digital Object Identifier 10.1109/LCOMM.2008.080017.

codes are subcodes of the full RM(n, n) codes, they, too, can be decoded by BP decoders. In this paper, we consider solely BP decoding of RM and polar codes. Using Forney's factor graph representation for RM codes, it is easy to see that the encoding and BP-decoding complexities of RM and polar codes are both $\mathcal{O}(N \log N)$ where N is the code block-length.

IV. Experimental results

First, we give a result, as presented in Table I, for the case $(N, K) = (32, 16)$, which is the smallest instance for which RM and polar codes differ. This table gives the bit error rate (BER) performance of (32,16) RM and polar codes over BECs with various erasure probabilities. (The final row of the table corresponds to a different type of polar coding scheme that will be discussed in the next section.) Each entry in this table was obtained by simulating the transmission of 10,000 codewords. To minimize statistical fluctuations, we used the same set of 10,000 channel realizations for all three entries in each column. Maximum number of iterations in BP decoding was set to 60, allowing each node in the code's factor graph to be visited up to 60 times.

TABLE I

BER for $(N, K) = (32, 16)$ codes on a BEC.

	Erasure probability				
Type of Code	0.05	0.15	0.25	0.35	0.45
RM(2,5)	0.00000	0.00039	0.01169	0.06525	0.24507
Polar	0.00001	0.00056	0.00702	0.05005	0.16722
Adaptive polar	0.00000	0.00039	0.00702	0.05005	0.16722

Table I does not provide sufficient evidence to conclude that polar codes are better than RM codes. However, as we consider codes with larger block-lengths, it does not take too long to obtain such evidence. To that end, we will consider codes of length $N = 256$ with dimensions K as given in Table II. Here, the code dimensions are chosen to correspond to those of RM$(r, 8)$ codes by setting $K = \sum_{i=0}^{r} \binom{8}{i}$.

TABLE II

Dimensions of codes used in simulations.

r	K	Rate
1	9	0.04
2	37	0.14
3	93	0.36
4	163	0.64
5	219	0.86
6	247	0.96

In Table III we show simulation results for RM codes of length $N = 256$ over a BEC. Results for competing polar codes are presented in Table IV. Comparison of the two sets of results shows that polar codes have a clear performance advantage. In these tables, and the ones that follow, we obtained each entry by simulating the transmission of 1000 codewords, with the maximum number of iterations in the BP decoder equal to 60. Blank entries in tables correspond to cases where code rate is above channel capacity.

Performance results for RM and polar codes on a BSC are shown in Tables VI and VII, respectively. Performances under antipodal signaling (in which code bits are mapped to channel inputs ±1) over an additive Gaussian noise channel (AGNC) are shown in Tables VIII and IX. Inspection of these tables again shows a distinct performance advantage for polar codes over RM codes. We have observed experimentally that such advantages become even more pronounced as the code block-length is increased.

V. Channel-specific code construction

One important property of the code construction rule (4) is that it is channel-independent; we use the same rule for constructing codes for any binary-input channel. Results in [1] show that if one uses polar code construction rules that are tailored to the specific channel on which the code will be used, one may expect better performance at the expense of more complexity in code construction (but not in encoding and decoding of the resulting code). In fact, it is proved in [1] that such channel-specific rules can achieve channel capacity.

The channel-specific polar code construction rule given in [1] has a simple form for the class of BECs. For a BEC with erasure probability ϵ, the rule is the same as above except the recursion in (4) begins with $z_{1,1} = \epsilon$. Thus, the code construction method that we considered above is actually tailored for a BEC with erasure rate 1/2, but it happens to work well for other channels as well. This is significant in that it shows the robustness of the rule against uncertainties and variations in channel parameters.

If we construct $(32, 16)$ polar codes tailored to specific BECs, we obtain the performance given in the last row of Table I, which matches the best performance available by either the RM code or the *fixed* polar code for the same scenario. Advantages of channel-tailored polar coding is illustrated further by Table V, which gives performance results for channel-adapted polar codes of length $N = 256$. Comparison of Tables V and IV shows that significant BER improvements are possible by channel-specific constructions.

Unfortunately, the exact code construction rule for arbitrary binary-input channels is too complicated to be given here. We will conclude by suggesting a heuristic method instead: given an arbitrary binary-input channel with capacity C bits, use the polar code that is matched to the BEC with erasure rate $\epsilon = 1 - C$, i.e., the BEC that has same capacity as the given channel. This rule has yielded good results in experiments.

References

[1] E. Arıkan, "Channel polarization: a method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," submitted for publication, Oct. 2007.

[2] D. E. Muller, "Application of boolean algebra to switching circuit design and to error correction," *IRE Trans. Electron. Computers*, vol. EC-3, pp. 6–12, Sept. 1954.

[3] I. Reed, "A class of multiple-error-correcting codes and the decoding scheme," *IRE Trans. Inform. Theory*, vol. 4, pp. 39–44, Sept. 1954.

[4] G. D. Forney Jr., "Codes on graphs: normal realizations," *IEEE Trans. Inform. Theory*, vol. IT-47, pp. 520–548, Feb. 2001.

TABLE III

RM code over BEC. $N = 256$, Trials = 1000, Max. iter. = 60.

	Erasure probability								
Rate	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90
0.04	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.07200
0.14	0.00000	0.00000	0.00000	0.00000	0.00000	0.01195	0.24395	0.47465	-
0.36	0.00000	0.00000	0.00933	0.21003	0.45208	0.48368	-	-	-
0.64	0.00083	0.23588	0.45156	-	-	-	-	-	-
0.86	0.30484	-	-	-	-	-	-	-	-

TABLE IV

Polar code over BEC. $N = 256$, Trials = 1000, Max. iter. = 60.

	Erasure probability								
Rate	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90
0.04	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.01378
0.14	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00308	0.20427	-
0.36	0.00000	0.00000	0.00000	0.00011	0.02087	0.29740	-	-	-
0.64	0.00000	0.00363	0.16155	-	-	-	-	-	-
0.86	0.10419	-	-	-	-	-	-	-	-

TABLE V

Channel-specific polar code over BEC. $N = 256$, Trials = 1000, Max. iter. = 60.

	Erasure probability								
Rate	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90
0.04	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00689
0.14	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00130	0.11181	-
0.36	0.00000	0.00000	0.00000	0.00005	0.02087	0.29740	-	-	-
0.64	0.00000	0.00339	0.16155	-	-	-	-	-	-
0.86	0.08167	-	-	-	-	-	-	-	-

TABLE VI

RM code over BSC. $N = 256$, Trials = 1000, Max. iter. = 60.

	Error probability										
Rate	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.36	0.37	0.38	0.39
0.04	0.00000	0.00000	0.00000	0.00000	0.00000	0.01111	0.10756	0.14389	0.21100	0.25778	0.30400
0.14	0.00000	0.00000	0.00584	0.12159	0.34368	-	-	-	-	-	-
0.36	0.00086	0.12829	0.39794	-	-	-	-	-	-	-	-

TABLE VII

Polar code over BSC. $N = 256$, Trials = 1000, Max. iter. = 60.

	Error probability										
Rate	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.36	0.37	0.38	0.39
0.04	0.00000	0.00000	0.00000	0.00000	0.00000	0.00344	0.04667	0.07844	0.11889	0.17722	0.21378
0.14	0.00000	0.00000	0.00000	0.01230	0.14297	-	-	-	-	-	-
0.36	0.00000	0.02418	0.21481	-	-	-	-	-	-	-	-

TABLE VIII

RM code using antipodal signaling over AGNC. $N = 256$, Trials = 1000, Max. iter. = 60.

	Signal-to-noise ratio (dB)										
Rate	-10.00	-8.00	-6.00	-4.00	-2.00	0.00	2.00	4.00	6.00	8.00	10.00
0.04	0.09300	0.02400	0.00089	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
0.14	-	-	0.42068	0.23214	0.02949	0.00014	0.00000	0.00000	0.00000	0.00000	0.00000
0.36	-	-	-	-	-	0.30597	0.02986	0.00000	0.00000	0.00000	0.00000
0.64	-	-	-	-	-	-	0.40810	0.10399	0.00109	0.00000	0.00000
0.86	-	-	-	-	-	-	-	-	0.12259	0.00051	0.00000
0.96	-	-	-	-	-	-	-	-	-	0.05732	0.00044

TABLE IX

Polar code using antipodal signaling over AGNC. $N = 256$, Trials = 1000, Max. iter. = 60.

	Signal-to-noise ratio (dB)										
Rate	-10.00	-8.00	-6.00	-4.00	-2.00	0.00	2.00	4.00	6.00	8.00	10.00
0.04	0.04844	0.00400	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
0.14	-	-	0.23784	0.03135	0.00038	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
0.36	-	-	-	-	-	0.04568	0.00088	0.00000	0.00000	0.00000	0.00000
0.64	-	-	-	-	-	-	0.25902	0.00787	0.00000	0.00000	0.00000
0.86	-	-	-	-	-	-	-	-	0.04048	0.00000	0.00000
0.96	-	-	-	-	-	-	-	-	-	0.05671	0.00068