Systematic Polar Coding

Erdal Arıkan, *Senior Member, IEEE*

Abstract—Polar codes were originally introduced as a class of non-systematic linear block codes. This paper gives encoding and decoding methods for systematic polar coding that preserve the low-complexity nature of non-systematic polar coding while guaranteeing the same frame error rate. Simulation results are given to show that systematic polar coding offers significant advantages in terms of bit error rate performance.

Index Terms—Polar codes, systematic polar codes, successive cancellation decoding, error propagation.

I. Introduction

POLAR coding is a linear block coding method that has been introduced recently [1] and is of interest currently for its provably capacity-achieving performance with low-complexity encoding and decoding algorithms in a diverse set of scenarios. For a representative list of papers discussing potential applications of polar coding, we cite [2], [3], [4], [5], [6], [7]. Hardware implementations are discussed in [8].

Polar codes in their standard form are non-systematic codes, in other words, the information bits do not appear as part of the codeword transparently. Since any linear code can be turned into a systematic code, polar codes can also be encoded systematically. However, it is not clear immediately if this can be done while retaining the low-complexity nature of polar coding. It is also unclear at first if there are any significant advantages that arise from systematic encoding of polar codes. In this note we describe a systematic variant of polar coding that preserves the low-complexity properties of standard non-systematic polar coding while significantly improving its bit-error rate (BER) performance. Systematic polar coding may also prove useful in turbo-style receiver designs where a polar decoder exchanges information with other modules iteratively. For such reasons, it is likely that systematic polar coding will become the preferred method of polar coding. The methods presented here are also applicable to Reed-Muller codes and may be of independent interest in that context.

II. Systematic Encoding in General

We consider coding schemes defined by a linear transformation over a field $\mathbb{F}$:

$$\mathbf{x} = \mathbf{u}\mathbf{G}, \quad \mathbf{x}, \mathbf{u} \in \mathbb{F}^N, \mathbf{G} \in \mathbb{F}^{N \times N} \tag{1}$$

For polar coding, we are mainly interested in the case where $\mathbb{F}$ is the binary field, $\mathbb{F} = \mathbb{F}_2$, but the results in this section hold for arbitrary fields, finite or infinite.

In non-systematic coding, $\mathbf{u}$ is regarded as the source word and $\mathbf{x}$ as the codeword. Actually, we use (1) to define a family of codes whose rates are adjusted by splitting the source word into two parts $\mathbf{u} = (\mathbf{u}_{\mathcal{A}}, \mathbf{u}_{\mathcal{A}^c})$ for some $A \subset \{1, \ldots, N\}$ so that the first part $\mathbf{u}_{\mathcal{A}} = (u_i : i \in \mathcal{A})$ consists of user data that is free to change in each round of transmission, while the second part $\mathbf{u}_{\mathcal{A}^c} = (u_i : i \in \mathcal{A}^c)$ consists of digits that are frozen at the beginning of the session and made known to the decoder. The mapping (1) can then be written as

$$\mathbf{x} = \mathbf{u}_{\mathcal{A}}\mathbf{G}_{\mathcal{A}} + \mathbf{u}_{\mathcal{A}^c}\mathbf{G}_{\mathcal{A}^c} \tag{2}$$

where $\mathbf{G}_{\mathcal{A}}$ and $\mathbf{G}_{\mathcal{A}^c}$ are the submatrices of $\mathbf{G}$ consisting of rows with indices in $\mathcal{A}$ and $\mathcal{A}^c$, respectively. This mapping defines a non-systematic encoder $\mathbf{u}_{\mathcal{A}} \mapsto \mathbf{x} = \mathbf{u}_{\mathcal{A}}\mathbf{G}_{\mathcal{A}} + \mathbf{c}$ where $\mathbf{c} \triangleq \mathbf{u}_{\mathcal{A}^c}\mathbf{G}_{\mathcal{A}^c}$ is a fixed vector. The code rate can be adjusted by choosing the size of the set $\mathcal{A}$.

Let us now fix a code, as specified by a non-systematic encoder as in (2), and consider various possible systematic encoders for this code. To that end, let us split the codeword into two parts by writing $\mathbf{x} = (\mathbf{x}_{\mathcal{B}}, \mathbf{x}_{\mathcal{B}^c})$, where $\mathcal{B}$ is an arbitrary subset of $\{1, \ldots, N\}$, and rewrite (2) as

$$\mathbf{x}_{\mathcal{B}} = \mathbf{u}_{\mathcal{A}}\mathbf{G}_{\mathcal{A}\mathcal{B}} + \mathbf{u}_{\mathcal{A}^c}\mathbf{G}_{\mathcal{A}^c\mathcal{B}} \tag{3}$$

$$\mathbf{x}_{\mathcal{B}^c} = \mathbf{u}_{\mathcal{A}}\mathbf{G}_{\mathcal{A}\mathcal{B}^c} + \mathbf{u}_{\mathcal{A}^c}\mathbf{G}_{\mathcal{A}^c\mathcal{B}^c} \tag{4}$$

where $\mathbf{G}_{\mathcal{A}\mathcal{B}}$ denotes the submatrix of $\mathbf{G}$ consisting of the array of elements $(G_{i,j})$ with $i \in \mathcal{A}$ and $j \in \mathcal{B}$, and similarly for the other submatrices. We now seek systematic encoders where $\mathbf{x}_{\mathcal{B}}$ plays the role $\mathbf{u}_{\mathcal{A}}$ played in non-systematic encoding as the data carrier, while $\mathbf{u}_{\mathcal{A}^c}$ is fixed as before. More precisely, for any given non-systematic encoder with parameter $(\mathcal{A}, \mathbf{u}_{\mathcal{A}^c})$, we say that a systematic encoder with parameter $(\mathcal{B}, \mathbf{u}_{\mathcal{A}^c})$ exists if equations (3) and (4) establish a one-to-one correspondence between the sets of possible values of $\mathbf{u}_{\mathcal{A}}$ and $\mathbf{x}_{\mathcal{B}}$.

Proposition 1: For any polar code defined by some non-systematic encoder with parameter $(\mathcal{A}, \mathbf{u}_{\mathcal{A}^c})$, there exists a systematic encoder with parameter $(\mathcal{B}, \mathbf{u}_{\mathcal{A}^c})$ if (and only if) $\mathcal{A}$ and $\mathcal{B}$ have the same number of elements and $\mathbf{G}_{\mathcal{A}\mathcal{B}}$ is an invertible matrix.

Proof is immediate and omitted. (The only-if part is not used in the sequel.) If the sufficiency conditions of Proposition 1 are satisfied, a systematic encoder with parameter $(\mathcal{B}, \mathbf{u}_{\mathcal{A}^c})$ can implement the mapping $\mathbf{x}_{\mathcal{B}} \mapsto \mathbf{x} = (\mathbf{x}_{\mathcal{B}}, \mathbf{x}_{\mathcal{B}^c})$ by computing

$$\mathbf{u}_{\mathcal{A}} = (\mathbf{x}_{\mathcal{B}} - \mathbf{u}_{\mathcal{A}^c}\mathbf{G}_{\mathcal{A}^c\mathcal{B}})(\mathbf{G}_{\mathcal{A}\mathcal{B}})^{-1}$$

and then inserting $\mathbf{u}_{\mathcal{A}}$ into (4) to obtain $\mathbf{x}_{\mathcal{B}^c}$.

III. SYSTEMATIC ENCODING FOR POLAR CODES

We now specialize the results of the previous section to polar codes. We first fix the field as the binary field, $\mathbb{F} = \mathbb{F}_2$. We recall that for any $n \geq 1$, polar codes of block size $N = 2^n$ can be defined so that they have an encoder as in (1) with

$$\mathbf{G}_N = \mathbf{F}^{\otimes n}, \qquad \mathbf{F} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \qquad (5)$$

where $\mathbf{F}^{\otimes n}$ denotes the nth Kronecker power of $\mathbf{F}$. Some important properties of generator matrices for polar codes are as follows.

1) Generator matrices are related by the recursion

$$\mathbf{G}_N = \begin{bmatrix} \mathbf{G}_{N/2} & \mathbf{0}_{N/2} \\ \mathbf{G}_{N/2} & \mathbf{G}_{N/2} \end{bmatrix} \qquad (6)$$

    where $\mathbf{0}_{N/2}$ is the all-zero matrix of size $(N/2) \times (N/2)$.
2) $\mathbf{G}_N$ is lower-triangular with ones on the diagonal, hence it is invertible (in fact, the inverse of $\mathbf{G}_N$ is itself).
3) Any submatrix $(\mathbf{G}_N)_{\mathcal{A}\mathcal{A}}$ of $\mathbf{G}_N$, with $\mathcal{A} \subset \{1, \ldots, N\}$, is also lower-triangular and has ones on the diagonal, so it is also invertible.

The third property implies that the sufficiency conditions of Proposition 1 hold for polar codes if we choose $\mathcal{A} = \mathcal{B}$. In accordance with this observation, we will leave $\mathcal{A}$ subject to choice but set $\mathcal{B} = \mathcal{A}$ in the rest of the paper.

A. Complexity

Proposition 2: Systematic encoding for polar codes can be carried out in time $\mathcal{O}(N \log(N))$.

Proof: We will make use of the recursive structure of the generator matrix as given by (6). Let $\mathcal{A} \subset \{1, \ldots, N\}$ be given. For systematic coding, we wish to solve

$$\mathbf{x} = \mathbf{u}\mathbf{G}_N \qquad (7)$$

given $(\mathbf{u}_{\mathcal{A}^c}, \mathbf{x}_{\mathcal{A}})$. Let $\mathbf{x} = (\mathbf{x}^{(1)}, \mathbf{x}^{(2)})$ where $\mathbf{x}^{(1)} = (x_1, \ldots, x_{N/2})$ and $\mathbf{x}^{(2)} = (x_{N/2+1}, \ldots, x_N)$, and similarly let $\mathbf{u} = (\mathbf{u}^{(1)}, \mathbf{u}^{(2)})$. Then, (7) can be written as

$$(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) = (\mathbf{u}^{(1)}, \mathbf{u}^{(2)}) \begin{bmatrix} \mathbf{G}_{N/2} & \mathbf{0}_{N/2} \\ \mathbf{G}_{N/2} & \mathbf{G}_{N/2} \end{bmatrix} \qquad (8)$$

which contains, as a subproblem, solving a similar problem of half the size, namely, finding a solution to

$$\mathbf{x}^{(2)} = \mathbf{u}^{(2)}\mathbf{G}_{N/2} \qquad (9)$$

for a given $(\mathbf{u}_{\mathcal{A}_2^c}^{(2)}, \mathbf{x}_{\mathcal{A}_2}^{(2)})$, where

$$\mathcal{A}_2 = \{i \in \{1, \ldots, N/2\} : i + N/2 \in \mathcal{A}\}$$

Let $\chi_{N/2}$ denote the worst case complexity of solving this problem over all possible choices for the set $\mathcal{A}_2$. Supposing that (9) has been solved, we can transform problem (8) into the form

$$\mathbf{x}^{(1)} = \mathbf{u}^{(1)}\mathbf{G}_{N/2} + \mathbf{u}^{(2)}\mathbf{G}_{N/2}$$
$$= \mathbf{u}^{(1)}\mathbf{G}_{N/2} + \mathbf{x}^{(2)}$$

where $\mathbf{x}^{(2)}$ is now known. This is simply another instance of the encoding problem, also at code length $N/2$, as can be seen more clearly by writing

$$(\mathbf{x}^{(1)} - \mathbf{x}^{(2)}) = \mathbf{u}^{(1)}\mathbf{G}_{N/2} \qquad (10)$$

We wish to solve (10) given $\mathbf{u}_{\mathcal{A}_1^c}^{(1)}$ and $(\mathbf{x}^{(1)} - \mathbf{x}^{(2)})_{\mathcal{A}_1}$ where

$$\mathcal{A}_1 = \{i \in \{1, \ldots, N/2\} : i \in \mathcal{A}\}$$

Since this problem, too, can be solved in time $\chi_{N/2}$, the overall complexity χ_N of the original problem is bounded as

$$\chi_N \leq 2\chi_{N/2} + \alpha N \qquad (11)$$

where the term αN, for some constant α, accounts for the work required in transforming problem instances back and forth; namely, first subtracting $\mathbf{x}_{\mathcal{A}_1}^{(2)}$ from $\mathbf{x}_{\mathcal{A}_1}^{(1)}$ to prepare (10), and then adding $\mathbf{x}^{(2)}$ to $(\mathbf{x}^{(1)} - \mathbf{x}^{(2)})$ once (10) has been solved. The recursive inequalities (11) give the complexity bound $\chi_N \leq \alpha N \log_2(N)$. ∎

Another method for systematic encoding is to use a successive cancellation decoder as an encoder. For this, one pretends that $\mathbf{x}$ has been sent across a binary erasure channel (BEC) and that the user data part $\mathbf{x}_{\mathcal{A}}$ has been received intact while the remaining part $\mathbf{x}_{\mathcal{A}^c}$ has been fully erased. One initializes the decoder suitably to reflect full knowledge of $(\mathbf{u}_{\mathcal{A}^c}, \mathbf{x}_{\mathcal{A}})$ and complete uncertainty about $(\mathbf{u}_{\mathcal{A}}, \mathbf{x}_{\mathcal{A}^c})$. It can be shown by recursive arguments that the decoder will always find $\mathbf{u}_{\mathcal{A}}$ correctly.

B. Permuted generator matrices

Polar codes may also be defined by generator matrices which differ from (5) by a post permutation operation, namely,

$$\mathbf{G}_N = \mathbf{F}^{\otimes n}\mathbf{\Pi}_N$$

where $\mathbf{\Pi}_N$ is an $N \times N$ permutation matrix. In fact, the standard form of polar coding as described in [1] applies such a permutation that corresponds to the *bit-reversal* operation. In such cases, a systematic encoder can be constructed by choosing the sets $\mathcal{A}$ and $\mathcal{B}$ so that $\mathcal{B}$ is the image of $\mathcal{A}$ under the permutation represented by $\mathbf{\Pi}_N$.

IV. PERFORMANCE

In this part, we give simulation results to demonstrate the performance advantages of systematic coding of polar codes. Simulations reported here were conducted using a polar code with length $N = 256$ and rate 1/2 over an additive white Gaussian noise (AWGN) channel. The modulation was binary phase shift keying (BPSK). Specifically, the codeword $\mathbf{x} \in \mathbb{F}_2^N$ produced by the polar encoder was mapped to a BPSK signal $\mathbf{s} \in \{-1, +1\}^N$ by the rule

$$s_i = \begin{cases} +1 & \text{for } x_i = 0 \\ -1 & \text{for } x_i = 1 \end{cases}$$

and the equivalent AWGN channel was

$$\mathbf{y} = \mathbf{s} + \mathbf{z}$$

where $\mathbf{z} = (z_1, \ldots, z_N)$ is an i.i.d. set of Gaussian random variables with mean zero and variance $N_0/2$.

The decoder used in the simulations was a successive cancellation (SC) decoder, as described in [1]. In fact, we employed the same SC decoder for both systematic and non-systematic codes. In both cases, the decoder took as input $(\mathbf{y}, \mathbf{u}_{\mathcal{A}^c})$ and produced an estimate $\hat{\mathbf{u}}$ of $\mathbf{u}$. For non-systematic
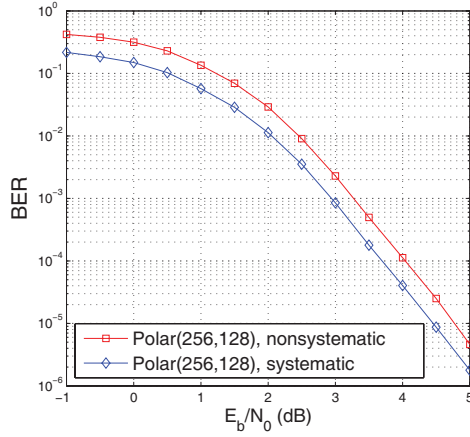
Fig. 1.   BER vs. E_b/N_0 for systematic and non-systematic polar codes under successive cancellation decoding: BPSK modulation over AWGN channel.
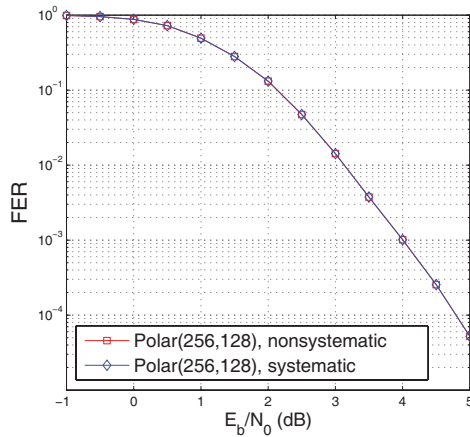

Fig. 2.   FER vs. E_b/N_0 for systematic and non-systematic polar codes under successive cancellation decoding: BPSK modulation over AWGN channel.

coding, the decoder stopped after putting out $\hat{\mathbf{u}}_{\mathcal{A}}$. For systematic coding, the decoder had an extra step of computing an estimate $\hat{\mathbf{x}} = \hat{\mathbf{u}}\mathbf{G}_N$ of $\mathbf{x}$, and produced $\hat{\mathbf{x}}_{\mathcal{A}}$ as output.

For non-systematic codes, bit error rate (BER) and frame error rate (FER) statistics were compiled by comparing $\hat{\mathbf{u}}_{\mathcal{A}}$ with $\mathbf{u}_{\mathcal{A}}$; for systematic codes, they were compiled by comparing $\hat{\mathbf{x}}_{\mathcal{A}}$ with $\mathbf{x}_{\mathcal{A}}$. Clearly, the two codes have the same FER under the particular decoder used here. In simulations, we used identical sets of test vectors for both systematic and non-systematic codes to ensure that their FER statistics coincided; this helped focus on the variation in BER statistics.

Figures 1 and 2 show, respectively, BER and FER vs. E_b/N_0 results. The parameter E_b stands for the transmitted signal energy per source bit and equals 2 here. The set $\mathcal{A}$ for polar coding was optimized for operation at $E_b/N_0 = 1$ (0 dB) using the Tal-Vardy method [9]. (Note that $E_b/N_0 = 0.188$ dB is the Shannon limit here.) Simulations with polar codes optimized for operation at various other E_b/N_0 values did not show any significant variation from the results reported here.

In summary, this simulation study confirms that systematic polar codes achieve better BER performance compared to non-systematic polar codes. The performance improvement is especially pronounced at low E_b/N_0 values where the BER curves are relatively flat.

V. Discussion

The development of systematic polar coding was motivated by the desire to reduce the susceptibility of non-systematic polar codes to error propagation under SC decoding. Systematic codes are expected to be more robust against error propagation than their non-systematic counterparts since the information bits in systematic coding are directly observed through the channel. Our results are consistent with this expectation.

It is surprising, however, that systematic polar coding showed an improvement in BER under a decoding method that computed $\hat{\mathbf{x}}$ indirectly, by first computing $\hat{\mathbf{u}}$ as if the code were non-systematic. One would expect that any decoding errors in $\hat{\mathbf{u}}$ (which suffers from error propagation) would be amplified in the re-encoding step $\hat{\mathbf{x}} = \hat{\mathbf{u}}\mathbf{G}_N$. Paradoxically, the simulation results show that this is not the case. Apparently, the SC decoder inherently pays more attention to minimizing the errors in $\hat{\mathbf{x}}$ than those in $\hat{\mathbf{u}}$. A better understanding of this issue is a subject for further study.

An alternative to SC decoding of polar codes is belief propagation (BP) decoding, as discussed in [10]. We repeated the above experiments with a BP decoder but did not obtain any results to recommend BP decoding over SC decoding; in fact, BP decoding performed noticeably worse than SC decoding at high E_b/N_0. In essence, we may view SC decoding as a special form of BP decoding with a carefully crafted schedule of belief updates. It is conceivable that there exist some other well-crafted schedules for BP decoding that give better performance than SC decoding at comparable complexity. Currently, we view SC decoding as the method of choice for low-complexity decoding of polar codes—systematic or non-systematic.

References

[1] E. Arıkan, "Channel polarization: a method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," *IEEE Trans. Inf. Theory*, vol. 55, pp. 3051–3073, July 2009.

[2] N. Hussami, S. B. Korada, and R. Urbanke, "Performance of polar codes for channel and source coding," in *Proc. IEEE Int. Symp. Inform. Theory*, pp. 1488–1492, June 2009.

[3] S. B. Korada, "Polar codes for channel and source coding," Ph.D. thesis, EPFL, Lausanne, 2009.

[4] E. Şaşoğlu, E. Telatar, and E. Yeh, "Polar codes for the two-user multiple-access channel," in *Proc. IEEE Inform. Theory Workshop*, Jan. 2010.

[5] E. Abbe and E. Telatar, "MAC polar codes and matroids," in *Proc. Inform. Theory and App. Workshop*, pp. 1-8, Jan. 2010.

[6] E. Hof, I. Sason, and S. Shamai, "Polar coding for reliable communications over parallel channels," in *Proc. IEEE Inform. Theory Workshop*, Aug. 2010.

[7] H. Mahdavifar and A. Vardy, "Achieving the secrecy capacity of wiretap channels using polar codes," to appear in *IEEE Trans. Inf. Theory*. Also available in revised form on arXiv:1007.3568.

[8] C. Leroux, I. Tal, A. Vardy, and W. J. Gross, "Hardware architectures for successive cancellation decoding of polar codes," *ICASSP 2011*, May 2011.

[9] I. Tal and A. Vardy, "How to construct polar codes." Online: arXiv:1105.6164v1, 31 May 2011.

[10] E. Arıkan, "A performance comparison of polar codes and Reed-Muller codes," *IEEE Commun. Lett.*, vol. 12, no. 6, pp. 447-449, June 2008.