Fixed-Threshold Polar Codes

Jing Guo
University of Cambridge
jg582@cam.ac.uk

Albert Guillén i Fàbregas
ICREA & Universitat Pompeu Fabra
University of Cambridge
guillen@ieee.org

Jossy Sayir
University of Cambridge
jossy.sayir@eng.cam.ac.uk

***Abstract*— We study a family of polar codes whose frozen set is such that it discards the bit channels for which the mutual information falls below a certain (fixed) threshold. We show that if the threshold, which might depend on the code length, is bounded appropriately, a coding theorem can be proved for the underlying polar code. We also give accurate closed-form upper and lower bounds to the minimum distance of the resulting code when the design channel is the binary erasure channel.**

I. Introduction

Channel polarization, introduced by Arıkan [1], is a phenomenon by which, given a binary-input discrete memoryless channel, virtual channels between the bits at the input of a linear encoder and the channel output sequence are created, such that the mutual information in each of these channels converges to either zero or one as the code length tends to infinity; the proportion of channels with mutual information close to one converges to the original channel's mutual information. These virtual channels are created by recursively applying channel combining and splitting steps.

Polar codes of rate $R = \frac{K}{N}$ are linear codes whose generator matrix is such that its rows induce the K virtual channels with highest mutual information among all N possible channels. The scheme behaves as if uncoded bits were sent through these channels. This construction, together with polarization, explicitly gives a code of rate close to the mutual information of the channel with vanishing error probability.

We propose a different construction of polar codes. Instead of choosing the best K virtual channels, we choose all channels whose mutual information is above a certain threshold which might depend on the code length. This new construction is shown to preserve the capacity-achieving property of Arıkan's original construction as long as the threshold function is bounded appropriately. This construction induces accurate closed-form upper and lower bounds to the minimum distance of the resulting codes when the design channel is the binary erasure channel (BEC). Our results sharpen existing bounds in the literature on the minimum distance of polar codes [2].

The paper is organized as follows. Notation and preliminaries are given in Section II. Fixed-threshold polar codes are discussed in Section III. Proofs of our main results can be found in Section IV.

This work has been funded in part by the European Research Council under ERC grant agreement 259663, by the FP7 Network of Excellence NEWCOM# under grant agreement 318306 and by the Spanish Ministry of Economy and Competitiveness under grant TEC2012-38800-C03-03.

II. Notation and Preliminaries

In this section, we introduce our notation. We also state some relevant results in the literature that are needed. Consider a binary discrete memoryless channel (B-DMC) $W : \{0,1\} \to \mathcal{Y}$ with transition probability $W(y|x)$, where $\mathcal{Y}$ is the output alphabet. The channel input-output mutual information with equiprobable inputs is denoted by $I(W)$ and the corresponding Bhattacharyya parameter is denoted by $Z(W)$. Let N be the channel block length and define $\boldsymbol{x} = (x_1, \ldots, x_N)$ and $\boldsymbol{y} = (y_1, \ldots, y_N)$ to be the length-N input and output sequences, respectively. The vector channel from $\boldsymbol{x}$ to $\boldsymbol{y}$ is defined as $W^N(\boldsymbol{y}|\boldsymbol{x})$. Consider the matrix

$$\boldsymbol{G}_2 = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \tag{1}$$

and denote by $\boldsymbol{G}_N = \boldsymbol{G}_2^{\otimes n}$ the $N \times N$ matrix, corresponding to the Kronecker product by itself $n = \log_2 N$ times.

The information bits are denoted by $\boldsymbol{u} = (u_1, \ldots, u_N)$, with $u_i \in \{0,1\}$ for $i = 1, \ldots, N$. Then, we define $W_N(\boldsymbol{y}|\boldsymbol{u}) = W^N(\boldsymbol{y}|\boldsymbol{u}\,\boldsymbol{G}_N)$ as the vector channel induced from the information bits when applying the linear transformation $\boldsymbol{G}_N$. This step is commonly termed channel combining. Channel splitting is a procedure that generates N binary-input channels out of $W_N(\boldsymbol{y}|\boldsymbol{u})$ assuming a successive decoder. More precisely, the i-th channel, for $i = 1, \ldots, N$, is generated as follows

$$W_N^{(i)}(\boldsymbol{y}, u_1, \ldots, u_{i-1}|u_i) = \sum_{u_{i+1},\ldots,u_N} \frac{1}{2^{N-i}} W_N(\boldsymbol{y}|\boldsymbol{u}) \tag{2}$$

and assumes that the bits $u_1, \ldots, u_{i-1}$ are known. Similarly, we denote the Bhattacharyya parameter of the i-th channel, for $i = 1, \ldots, N$, as $Z_N^{(i)} = Z\big(W_N^{(i)}(\boldsymbol{y}, u_1, \ldots, u_{i-1}|u_i)\big)$. We let Z_N denote the random variable corresponding to $Z_N^{(i)}$. For future use, we define the channel inverse labeling functions $b_j : \{1, \ldots, N\} \to \{0,1\}$ for $j = 1, \ldots, n$, such that $b_j(i)$ returns the j-th bit in the binary label of channel i (natural labeling is assumed), for $j = 1, \ldots, n$ and $i = 1, \ldots, N$.

Lemma 1 (Channel Polarization [1]): For any B-DMC W, the channels $\{W_N^{(i)}\}$ for $i = 1, \ldots, N$ have the property that

$$\lim_{N\to\infty} \frac{\Big|i \in \{1,\ldots,N\} : I\big(W_N^{(i)}\big) \in (1-\delta, 1]\Big|}{N} = I(W),$$
$$\lim_{N\to\infty} \frac{\Big|i \in \{1,\ldots,N\} : I\big(W_N^{(i)}\big) \in [0, \delta)\Big|}{N} = 1 - I(W). \tag{3}$$

In words, Lemma 1 implies that as $N \to \infty$ the mutual information of the channels $W_N^{(i)}$ for $i = 1, \dots, N$ tends to either one or zero, and that the fraction of channels with mutual information close to one is in the limit $I(W)$. In terms of convergence rate, we have the following.

Lemma 2 ([3]): For any B-DMC W and constant $0 < \beta < \frac{1}{2}$, we have

$$\lim_{N\to\infty} \Pr\left(Z_N \le 2^{-N^\beta}\right) = I(W). \tag{4}$$

The following is a corollary of [4, Th. 3].

Lemma 3: For any B-DMC W and constant $0 < \beta < \frac{1}{2}$, we have

$$\lim_{N\to\infty} \Pr\left(Z_N < 1 - 2^{-N^\beta}\right) = I(W). \tag{5}$$

A. Polar Codes

In order to construct a polar code of rate $R = \frac{K}{N}$, exactly K rows of $\boldsymbol{G}$ must be selected. Different methods for choosing these rows yield different codes. In particular, Arıkan's original construction, first fixes the rate R, and then selects the rows of $\boldsymbol{G}$ whose indices give highest $I\big(W_N^{(i)}\big)$ over all $i = 1, \dots, N$.

Alternatively, polar codes can be defined through the set of discarded rows. This is called the *frozen* set $\mathcal{F}$; its complement is denoted by $\mathcal{F}^c$. If the rate is fixed, the cardinality of the frozen set is obviously $|\mathcal{F}| = N - K$ and the frozen set is

$$\mathcal{F} = \big\{i = 1, \dots, N \,|\, I\big(W_N^{(i)}\big) < \theta_N\big\} \tag{6}$$

where θ_N is the mutual information threshold that makes $|\mathcal{F}| = N - K$. Arıkan's threshold function is difficult to characterize analytically and even to compute numerically. Fig. 1 illustrates the evolution of the threshold of Arıkan's construction for $R = 0.3, 0.4$ in a BEC with $I(W) = \frac{1}{2}$.

As shown in [2, Lemma 6.2], the minimum distance of polar codes can be expressed as a function of the frozen set. Let $w_{\mathrm{H}}(b_1, \dots, b_n)$ denote the function that outputs the Hamming weight of bits $b_1, \dots, b_n$.

Lemma 4 ([2]): For any choice of frozen set $\mathcal{F}$, the minimum distance $d_{\min}$ of polar codes is given by

$$d_{\min} = \min_{i\in\mathcal{F}^c} 2^{w_{\mathrm{H}}(b_1(i),\dots,b_n(i))}. \tag{7}$$

III. Fixed-threshold construction

We propose a different method for constructing polar codes. Arıkan's original codes, have a fixed rate, and the frozen set is determined by all sub-channels with mutual information smaller than a certain threshold, such that $|\mathcal{F}| = N - K$ (see (6)). In the proposed construction, a mutual information threshold function θ_N is fixed, and the frozen set is given as

$$\mathcal{F} = \big\{i = 1, \dots, N \,|\, I\big(W_N^{(i)}\big) < \theta_N\big\}. \tag{8}$$

The threshold function θ_N considered here is a closed-form function of N, thus easily computable. As a result, the number

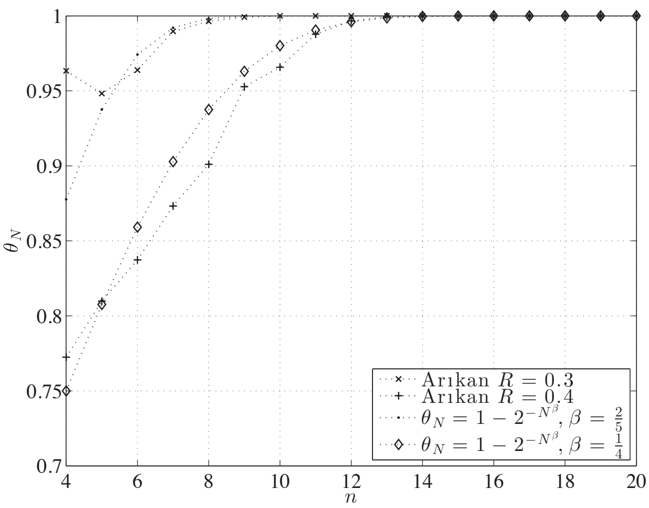

Fig. 1. Thresholds of Arıkan's fixed-rate construction for $R = 0.3, 0.4$ and fixed-threshold polar codes with threshold θ_N over the BEC with $I(W) = \frac{1}{2}$.

of channels in the frozen set is in general not fixed by the rate, and hence, the rate itself depends on N. Fig. 1 also shows the threshold functions $\theta_N = 1 - 2^{-N^\beta}$ for $\beta = \frac{2}{5}, \frac{1}{4}$.

A. Coding Theorem

For fixed-threshold polar codes, the rate R of the code is in general a function of N. In order for these codes to have a valid coding theorem, the threshold function θ_N must be such that, as N grows, the rate R converges to $I(W)$ and the error probability $P_e \to 0$. We first show the conditions under which R converges to $I(W)$.

Lemma 5 (Rate convergence): Consider a B-DMC W, and a fixed-threshold polar code of length N and threshold function θ_N. Let R_N denote the rate of the code. If there exists an N_0 such that for $N > N_0$ such that the threshold function satisfies the bounds

$$\alpha_1 2^{-N^{\beta_1}} \le \theta_N \le 1 - \alpha_2 2^{-N^{\beta_2}}, \tag{9}$$

where $\alpha_m > 0$, $0 < \beta_m < \frac{1}{2}$, $m = 1, 2$, then

$$\lim_{N\to\infty} R_N = I(W). \tag{10}$$

Lemma 5 shows that asymptotically, any threshold meeting the bounds in (9) is such that $R_N \to I(W)$. In particular, any constant threshold $\theta_N \in (0, 1)$ will result in $R_N \to I(W)$. Fig. 2 illustrates the results of Lemma 5. We can see that the speed of convergence depends on the threshold θ_N.

It is implicit in the figure that some threshold functions that fulfill the bounds in Lemma 5 approach the mutual information from above, i.e., $R_N = I(W) + \delta_N$ for some $\delta_N > 0$ such that $\delta_N \to 0$. By the converse to the channel coding theorem [5], such codes will have an error probability bounded away from zero. The following result characterizes the set of threshold functions that yield vanishing error probability.

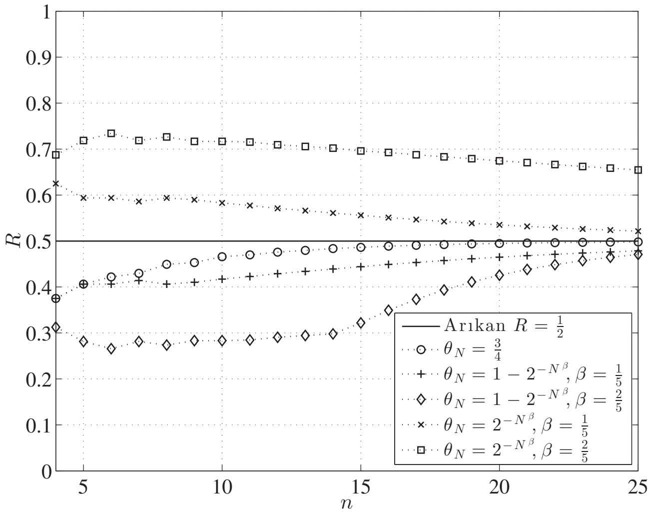

Fig. 2. Rate convergence for different threshold functions θ_N over a BEC with $I(W)=\frac{1}{2}$.

Theorem 1 (Coding Theorem): Consider a B-DMC W, and a fixed-threshold polar code of length N and threshold function θ_N. Let $P_e(N,\theta_N)$ denote the error probability of this code. Let $\gamma_1,\gamma_2>0$, $\beta_1>2$, $0<\beta_2<\frac{1}{2}$ be fixed. If there exists an N_0 such that for $N>N_0$, the threshold function is such that

$$1-\gamma_1 N^{-\beta_1} \le \theta_N \le 1-\gamma_2 2^{-N^{\beta_2}}. \tag{11}$$

then,

$$\lim_{N\to\infty} P_e(N,\theta_N)=0. \tag{12}$$

Theorem 1 shows that as $N\to\infty$, thresholds close to 1 will yield polar codes for which both the rate converges to $I(W)$ and the probability of error vanishes; this rules out constant-threshold codes.

B. Minimum Distance

In this section, we give closed-form upper and lower bounds to the minimum distance of fixed-threshold polar codes when the design channel is the BEC. As the results will show, the bounds are very easy to compute and accurately characterize the minimum distance. BEC designs perform well in multiple channels, including the binary-symmetric channel and the additive white Gaussian noise channel (AWGN) [6]; the performance of BEC codes over the latter channel is shown to be very close to that of AWGN-tailored polar codes. Our bounds, stated in Theorem 2, follow after the following results, the first of which being an extension of [3, Lemma 1].

Lemma 6: Let $A:\mathbb{R}\to\mathbb{R}: x\mapsto 2x-x^2$ and $B:\mathbb{R}\to\mathbb{R}: x\mapsto x^2$. Suppose a sequence of numbers $x_0,x_1,\ldots,x_n$ is defined by specifying $x_0: 0\le x_0\le 1$ and the recursion $x_{i+1}=f_i(x_i)$ with $f_i\in\{A,B\}$. Suppose $|\{0\le i\le n-1: f_i=A\}|=k$ and $|\{0\le i\le n-1: f_i=B\}|=n-k$. Then,

$$x_n \le B^{(n-k)}\big(A^{(k)}(x_0)\big), \tag{13}$$

where $A^{(k)}$ denotes the application of function A k times.

Proof: Observe that both $A(x)$ and $B(x)$ are monotonically increasing functions over $x\in[0,1]$. We first have that

$$A\big(B(x)\big)\le B\big(A(x)\big). \tag{14}$$

The upper bound in (13) corresponds to choosing

$$\begin{aligned} f_0 &= \ldots = f_{k-1}=A,\\ f_k &= \ldots = f_{n-1}=B. \end{aligned}$$

Suppose $\{f_i\}$ is not chosen as above. Then, there exists $j\in\{1,\ldots,n-1\}$ for which $f_{j-1}=B$ and $f_j=A$. According to Eq. (14), we can always achieve a larger value by swapping f_j and f_{j-1}. ■

Lemma 7: Consider the N channels $W_N^{(i)}$ for $i=1,\ldots,N$ resulting from $n=\log N$ steps of channel combining and splitting with matrix $\boldsymbol{G}_2$ over a BEC with mutual information $I(W)$. Consider the binary label of each channel $\big(b_1(i),\ldots,b_n(i)\big)$, its corresponding Hamming weight $w_{\mathrm{H}}\big(b_1(i),\ldots,b_n(i)\big)$, and define

$$\mathcal{S}_k=\big\{i=1,\ldots,N \,|\, w_{\mathrm{H}}\big(b_1(i),\ldots,b_n(i)\big)=k\big\} \tag{15}$$

to be the set of channel indices such that the Hamming weight of its binary label is equal to k for $i=1,\ldots,N$ and $k=1,\ldots,n$. Define

$$i_k^\star \triangleq \arg\max_{i\in\mathcal{S}_k} I\big(W_N^{(i)}\big). \tag{16}$$

Then,

$$\big(b_1(i_k^\star),\ldots,b_n(i_k^\star)\big)=\big(\underbrace{1,\ldots,1}_{k},\underbrace{0,\ldots,0}_{n-k}\big). \tag{17}$$

In words, the above result states that for a fixed Hamming weight k of the binary label of the channel index$i=1,\ldots,N$, placing the k ones in the most significant bit positions yields the maximum mutual information over all such channels.

Proof: We notice that $A(x)=2x-x^2$ and $B(x)=x^2$ are the recursive formulas to calculate mutual information of the bit channels for a BEC of mutual information $I(W)$. Then, for $i\in\mathcal{S}_k$ Lemma 6 states that

$$I\big(W_N^{(i)}\big)\le B^{(n-k)}\Big(A^{(k)}\big(I(W)\big)\Big). \tag{18}$$

and that channel $i_k^\star$ achieves the bound (18) with equality. ■

Now, assume the minimum distance $d_{\min}$ is such that $d_{\min}=2^{k_0}$. Then, from the definition of the frozen set in (8) and Lemma 4 k_0 must be such that

$$I\Big(W_N^{(i^\star_{k_0-1})}\Big)<\theta_N\le I\Big(W_N^{(i^\star_{k_0})}\Big). \tag{19}$$

Note that, for a fixed threshold θ_N, k_0 is the minimum (resp. maximum) integer solution of the upper bound (resp. lower bound) in (19).

Theorem 2: Consider a BEC with mutual information $I(W)=1-\epsilon$. The minimum distance $d_{\min}=2^{k_0}$ of fixed-threshold polar codes of length N and threshold θ_N is such

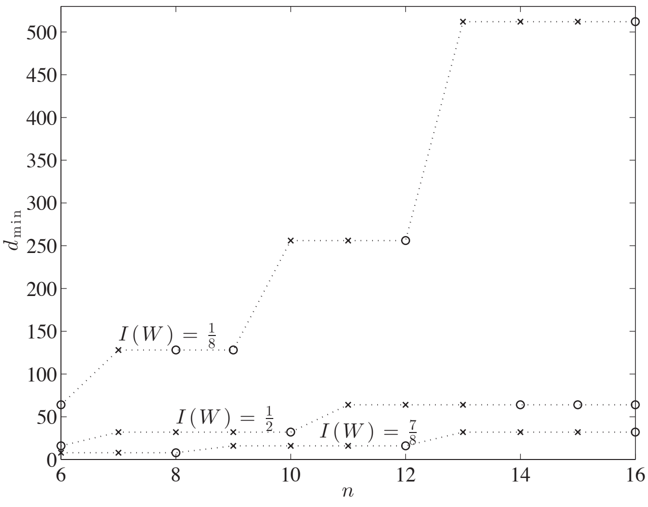

Fig. 3. Minimum distance of fixed-threshold polar codes with $\theta_N = 1 - \frac{1}{N}2^{-N^\beta}$ with $\beta = \frac{2}{5}$ over the BEC.

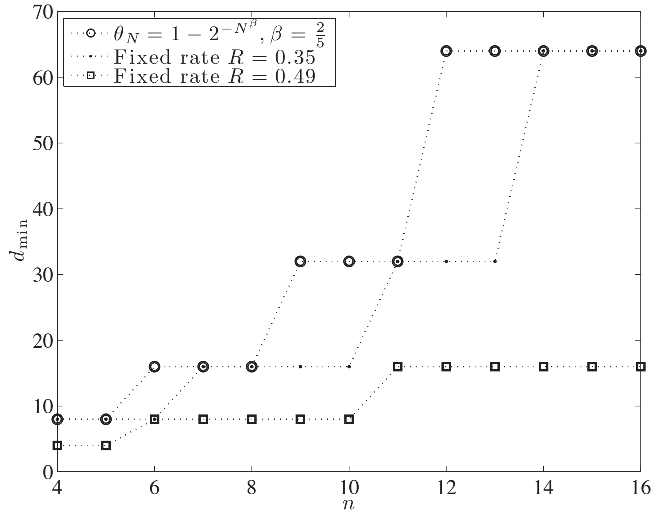

Fig. 4. Minimum distance $d_{\min}$ for Arıkan's fixed-rate polar codes with $R = 0.35, 0.49$ and for fixed-threshold polar code with threshold function $\theta_N = 1 - 2^{-N^\beta}$ with $\beta = \frac{2}{5}$ for a BEC with $I(W) = \frac{1}{2}$.

that k_0 can be bounded as

$$k_0 \leq \left\lceil \log_2\big(n - \log_2(c_1 c_2)\big) - c_0 \right\rceil, \tag{20}$$

$$k_0 \geq \left\lfloor \log_2\big(n - \log_2(c_1)\big) - c_0 \right\rfloor, \tag{21}$$

where $c_0 = \log_2\{-\log_2(\epsilon)\}, c_1 = -\ln\theta_N, c_2 = -\frac{\epsilon}{\ln(1-\epsilon)}$.

As opposed to the original polar codes [1], where the minimum distance can only be computed once the frozen set is determined, the bounds given in Theorem 2 are given as closed-form expressions and are very simple to calculate. In Fig. 3 we show the minimum distance bounds of fixed-threshold polar codes with threshold function $\theta_N = 1 - \frac{1}{N}2^{-N^\beta}$ with $\beta = \frac{1}{5}$ in a BEC channel with $I(W) = \frac{7}{8}, \frac{1}{2}$ and $\frac{1}{8}$. The dotted lines correspond to the actual minimum distance. Crosses (resp. circles) correspond to points where the upper (resp. lower) bound gives the actual $d_{\min}$ of the code. As we can see from the figure, the bounds are tight in the sense that either the upper or lower bound gives the actual $d_{\min}$ of the code.

Fig. 4 compares the actual minimum distance of fixed-threshold polar codes with that of Arıkan's fixed-rate polar codes. By checking the rate-convergence in Fig. 2, we observe that for fixed-threshold polar-codes, the slower the convergence to $I(W)$, the higher the minimum distance. We also note that the regions where the minimum distances of fixed-threshold and fixed-rate polar codes coincide, is the region where their corresponding rates are close. In this case, both constructions would give approximately the same code.

IV. Proofs

A. Proof of Lemma 5

Consider first the upper bound in (9) and let $\theta_N = 1 - \alpha_2 2^{-N^{\beta_2}}$ with $0 < \beta_2 < \frac{1}{2}$. For $I\big(W_N^{(i)}\big) \geq \theta_N$, using [1, Eq. (2)] and $1 - x^2 \leq 2(1-x)$ gives

$$Z\big(W_N^{(i)}\big) \leq \sqrt{2\alpha_2}\, 2^{-2^{(n\beta_2 - 1)}}. \tag{22}$$

There exists an N_1 such that, for $N > N_1$, we can find $0 < \beta' < \frac{1}{2}$, that

$$\sqrt{2\alpha_2}\, 2^{-2^{(n\beta_2-1)}} \leq 2^{-N^{\beta'}}. \tag{23}$$

The above implies that

$$\Pr\Big(I\big(W_N^{(i)}\big) \geq \theta_N\Big) \leq \Pr\Big(Z\big(W_N^{(i)}\big) \leq 2^{-N^{\beta'}}\Big). \tag{24}$$

Similarly, for $I\big(W_N^{(i)}\big) < \theta_N$, we use $\ln(1+x) \leq x$ and [1, Eq. (1)] to show that

$$Z\big(W_N^{(i)}\big) > \frac{\alpha_2 2^{-N^{\beta_2}}}{\log_2 e}. \tag{25}$$

There exists an N_2 such that for $N > N_2$, we can find $0 < \tilde{\beta} < \frac{1}{2}$, that

$$\frac{\alpha_2 2^{-N^{\beta_2}}}{\log_2 e} \geq 2^{-N^{\tilde{\beta}}}, \tag{26}$$

and hence

$$\Pr\Big(I\big(W_N^{(i)}\big) < \theta_N\Big) \leq \Pr\Big(Z\big(W_N^{(i)}\big) > 2^{-N^{\tilde{\beta}}}\Big). \tag{27}$$

According to the fixed-threshold construction,

$$R_N = \frac{\Big|i \in \{1,\ldots,N\} : I\big(W_N^{(i)}\big) \in [\theta_N, 1]\Big|}{N}. \tag{28}$$

We also have that

$$R_N + \frac{\Big|i \in \{1,\ldots,N\} : I\big(W_N^{(i)}\big) \in [0, \theta_N)\Big|}{N} = 1. \tag{29}$$

Taking limits

$$\lim_{N\to\infty} R_N = \lim_{N\to\infty} \frac{\left|i \in \{1,\dots,N\} : I\big(W_N^{(i)}\big) \in [\theta_N, 1]\right|}{N} \tag{30}$$

$$\leq \lim_{N\to\infty} \frac{\left|i \in \{1,\dots,N\} : Z\big(W_N^{(i)}\big) \leq 2^{-N^{\beta'}}\right|}{N} \tag{31}$$

$$= I(W) \tag{32}$$

where (30) follows from (24) and the last step follows from Lemma 2. Similarly,

$$\lim_{N\to\infty} \frac{\left|i \in \{1,\dots,N\} : I\big(W_N^{(i)}\big) \in [0, \theta_N)\right|}{N} \tag{33}$$

$$\leq \lim_{N\to\infty} \frac{\left|i \in \{1,\dots,N\} : Z\big(W_N^{(i)}\big) > 2^{-N^{\bar{\beta}}}\right|}{N} \tag{34}$$

$$= 1 - I(W). \tag{35}$$

Therefore, we have that $\lim_{N\to\infty} R_N = I(W)$ as desired.

Convergence of the lower bound is proved similarly invoking Lemma 3. Given that both upper and lower bounds on θ_N are such $\lim_{N\to\infty} R_N = I(W)$, any threshold function that for large enough N can be bounded as (9) will have that the corresponding R_N will converge to $I(W)$.

B. Proof of Theorem 1

First, we notice that for a given channel W and block length N, the smaller θ_N, the higher the rate, and the larger the error probability $P_e(N, \theta_N)$. Therefore, we concentrate on the lower bound and prove that for $\theta_N = 1 - \gamma_1 N^{-\beta_1}, \beta_1 > 2$ that

$$\lim_{N\to\infty} P_e(N, \theta_N) = 0. \tag{36}$$

Using [1, Eq. (2)] and $1 - x^2 \leq 2(1-x)$ get that

$$Z\left(W_N^{(i)}\right) \leq \sqrt{2\gamma_1} N^{-\frac{\beta_1}{2}}. \tag{37}$$

Thus,

$$P_e(N, \theta_N) \leq \sum_{i\in\mathcal{F}^c} Z\left(W_N^{(i)}\right), \tag{38}$$

$$\leq N \max_{i\in\mathcal{F}^c} Z\left(W_N^{(i)}\right), \tag{39}$$

$$\leq \sqrt{2\gamma_1} N^{1-\frac{\beta_1}{2}}, \tag{40}$$

$$\tag{41}$$

Since $\beta_1 > 2$, we have that

$$\lim_{N\to\infty} \sqrt{2\gamma_1} N^{1-\frac{\beta_1}{2}} = 0, \tag{42}$$

and hence

$$\lim_{N\to\infty} P_e(N, \theta_N) = 0. \tag{43}$$

C. Proof for Theorem 2

We prove the upper bound. Consider

$$\big(b_1(i_k^\star), \dots, b_n(i_k^\star)\big) = \big(\underbrace{1,\dots,1}_{k}, \underbrace{0,\dots,0}_{n-k}\big). \tag{44}$$

To find $d_{\min} = 2^{k_0}$ we need to find the smallest k such that $I\big(W_N^{(i_k^\star)}\big) \geq \theta_N$. For the BEC, we have that

$$I\big(W_N^{(i_k^\star)}\big) = \Big(1 - \epsilon^{2^k}\Big)^{2^{n-k}} \tag{45}$$

where we have defined $\epsilon = 1 - I(W)$. If we take logarithms we have that k must be such that

$$2^{n-k} \ln\Big(1 - \epsilon^{2^k}\Big) \geq \ln \theta_N. \tag{46}$$

It can be shown that

$$\ln\Big(1 - \epsilon^{2^k}\Big) \geq c\,\epsilon^{2^k} \tag{47}$$

where $c = \frac{\ln(1-\epsilon)}{\epsilon}$, and that $I\big(W_N^{(i_k^\star)}\big)$ given in (45) non-negative and increasing in k. This implies that the minimum k that solves (46) is upper bounded by the minimum k that solves

$$2^{n-k} c\,\epsilon^{2^k} \geq \ln \theta_N, \tag{48}$$

which is in turn the same that solves

$$2^k \geq \frac{n-k}{-\log_2 \epsilon} + \frac{\log_2 \xi_N}{\log_2 \epsilon} \tag{49}$$

where $\xi_N \triangleq \frac{\ln \theta_N}{c}$. Further taking logarithms we have that

$$k \geq \log_2(n - k - \log_2 \xi_N) - \log_2(-\log_2 \epsilon). \tag{50}$$

Observe that upper bounding the R.H.S. of (50) by $\log_2(n - \log_2 \xi_N) - \log_2(-\log_2 \epsilon)$ implies an upper bound to the solution for k. Hence,

$$k^\star = \lceil \log_2(n - \log_2 \xi_N) - \log_2(-\log_2 \epsilon) \rceil \tag{51}$$

is an integer solution of (50), proving the upper bound. The lower bound is proved similarly, using that $\ln(1-x) \leq -x$ for $x \leq 0$ to bound the L.H.S. of (47).

Acknowledgment

The authors thank Erdal Arıkan for helpful discussions and meetings facilitated by the FP7 Network of Excellence NEWCOM# and its predecessor NEWCOM++.

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, no. 7, pp. 3051–3073, July 2009.

[2] S. B. Korada, *Polar codes for channel and source coding*, Ph.D. thesis, Ecole Polytechnique Fédérale de Lausanne, Switzerland, 2009.

[3] E. Arıkan and E. Telatar, "On the rate of channel polarization," in *IEEE Int. Symp. Inf. Theory*, June 2009, pp. 1493–1495.

[4] S. H. Hassani, R. Mori, T. Tanaka, and R. Urbanke, "Rate-dependent analysis of the asymptotic behavior of channel polarization," *IEEE Trans. Inf. Theory*, Nov. 2012.

[5] R. G. Gallager, *Information theory and reliable communication*, Wiley, 1968.

[6] R. Mori and T. Tanaka, "Performance of polar codes with the construction using density evolution," *IEEE Commun. Letters*, vol. 13, no. 7, pp. 519–521, Jul. 2009.