Source Polarization

Erdal Arıkan
Bilkent University, Ankara, Turkey

***Abstract*—The notion of source polarization is introduced and investigated. This complements the earlier work on channel polarization. An application to Slepian-Wolf coding is also considered. The paper is restricted to the case of binary alphabets. Extension of results to non-binary alphabets is discussed briefly.**

***Index Terms*—Polar codes, source polarization, channel polarization, source coding, Slepian-Wolf coding.**

I. Introduction

We introduce the notion of "source polarization" which complements "channel polarization" that was studied in [1]. One immediate application of source polarization is the design of polar codes for lossless source coding. Lossless source coding using polar codes has already been considered extensively in the pioneering works [2] and [3], which reduced this problem to one of channel polarization using the duality between the two problems. The approach in this paper is direct and offers an alternative (primal) viewpoint.

This paper is restricted mostly to binary memoryless sources. We indicate in the end briefly the possible generalizations to non-binary sources.

We use the notation of [1]. In particular, we write u^N to denote a vector $(u_1, \ldots, u_N)$ and u_i^j to denote the sub-vector $(u_i, \ldots, u_j)$ for any $1 \leq i \leq j \leq N$. If $j < i$, u_i^j is the null vector. The logarithm is to the base 2 unless otherwise indicated. We write $X \sim \text{Ber}(p)$ to denote a Bernoulli random variable (RV) with values in $\{0, 1\}$ and $P_X(1) = p$. The entropy $H(X)$ of such a RV is denoted sometimes as $\mathcal{H}(p) = -p \log p - (1-p)\log(1-p)$.

II. Polarization of binary memoryless sources with side information

Let $(X, Y) \sim P_{X,Y}$ be an arbitrary pair of random variables over $\mathcal{X} \times \mathcal{Y}$ with $\mathcal{X} = \{0, 1\}$ and $\mathcal{Y}$ an arbitrary countable set. Throughout this section, we regard (X, Y) as a memoryless source, with X as the part to be compressed and Y in the role of "side-information" about X. We consider a sequence $\{(X_i, Y_i)\}_{i=1}^{\infty}$ of independent drawings from (X, Y) and write (X^N, Y^N) to denote the first N elements of this sequence, for any integer $N \geq 1$.

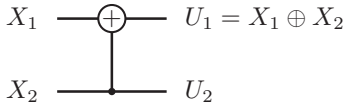

Fig. 1. Basic source transformation.

The basic idea of source polarization is contained in the transformation shown in Fig. 1, where "$\oplus$" denotes addition mod-2. The operation $(X_1, X_2) \to (U_1, U_2)$ performed by the circuit preserves entropy, *i.e.*,

$$\begin{aligned} H(U_1, U_2|Y_1, Y_2) &= H(X_1, X_2|Y_1, Y_2) \\ &= 2H(X|Y), \end{aligned} \tag{1}$$

but is polarizing in the sense that

$$H(U_1|Y_1, Y_2) \geq H(X|Y) \geq H(U_2|Y_1, Y_2, U_1). \tag{2}$$

It is easy to show that equalities hold here if and only if $H(X|Y)$ equals 0 or 1. Thus, unless the entropies at the input of the circuit are already perfectly polarized, the entropies at the output will polarize further.

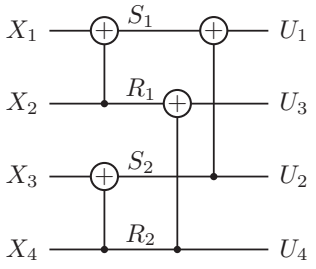

Fig. 2. Four-by-four source transformation.

Figure 2 shows the recursive continuation of the construction to the case where four independent copies of (X, Y) are processed. The entropy conservation law states that

$$H(U^4|Y^4) = H(X^4|Y^4) = 4H(X|Y).$$

Using the chain rule, we may split the output entropy as

$$H(U^4|Y^4) = \sum_{i=1}^{4} H(U_i|Y^4, U^{i-1}).$$

Note that the variables U^4 are assigned to the output terminals of the circuit in Fig. 2 in a shuffled order. This is motivated by the observation that, with this ordering, the pair (U_1, U_2) is obtained from two i.i.d. RVs, namely, (S_1, S_2), by the same two-by-two construction as in Fig. 1. A similar remark applies to the relationship between (U_3, U_4) and (R_1, R_2). These observations lead to the the following inequalities, which are special cases of those in (2).

$$\begin{aligned} H(U_1|Y^4) &\geq H(S_1|Y_1^2) \\ &= H(S_2|Y_3^4) \geq H(U_2|Y^4, U^1), \end{aligned}$$

$$H(U_3|Y^4,U^2) \geq H(R_1|Y_1^2,S_1) = H(R_2|Y_3^4,S_2) \geq H(U_4|Y^4,U^3).$$

There is no general inequality between $H(U_2|Y^4,U^1)$ and $H(U_3|Y^4,U^2)$. The conclusion to be drawn is that polarization is enhanced further by repeating the basic construction.

For any $N = 2^n$, $n \geq 1$, the general form of the source polarization transformation is defined algebraically as

$$G_N = \left[\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right]^{\otimes n} B_N \tag{3}$$

where "$^{\otimes n}$" denotes the nth Kronecker power and B_N is the "bit-reversal" permutation (see [1]). It is easy to check that the transforms in Figures 1 and 2 conform to $U^N = X^N G_N$. The main result on source polarization for binary alphabets is the following.

Theorem 1. *Let (X,Y) be a source as above. For any $N = 2^n$, $n \geq 1$, let $U^N = X^N G_N$. Then, for any $\delta \in (0,1)$, as $N \to \infty$,*

$$\frac{\left|\left\{i \in [1,N] \colon H(U_i|Y^N,U^{i-1}) \in (1-\delta,1]\right\}\right|}{N} \to H(X|Y)$$

and

$$\frac{\left|\left\{i \in [1,N] \colon H(U_i|Y^N,U^{i-1}) \in [0,\delta)\right\}\right|}{N} \to 1 - H(X|Y).$$

We omit the full proof but sketch the idea, which follows the proof of the channel polarization result in [1]. The first step is to define a tree random process for tracking the evolution of the conditional entropy terms $\{H(U_i|Y^N,U^{i-1})\}$. The analysis is aided by an accompanying supermartingale based on the source Bhattacharyya parameters. For the basic source $(X,Y) \sim P_{X,Y}$, this parameter is defined as

$$Z(X|Y) = 2 \sum_y P_Y(y) \sqrt{P_{X|Y}(0|y) P_{X|Y}(1|y)}.$$

The source Bhattacharyya parameters satisfy the following as they undergo the two-by-two polarization transformation.

Proposition 1. *Let (X,Y) be a source as above, and (X_1,Y_1) and (X_2,Y_2) two independent drawings from (X,Y). Then,*

$$Z(X_1 \oplus X_2|Y^2) \;\leq\; 2Z(X|Y) - Z(X|Y)^2$$

and

$$Z(X_2|Y^2, X_1 \oplus X_2) = Z(X|Y)^2.$$

We omit the proof of this result since it is very similar to the proof of a similar inequality on channel Bhattacharyya parameters given in [1]. Thus, we have the inequality

$$Z(U_1|Y^2) + Z(U_2|Y^2,U^1) \leq 2Z(X|Y)$$

which is the basis of the Bhattacharyya supermartingale. Convergence results about the Bhattacharyya supermartingale may be translated into similar results for the entropy martingale through the following pair of inequalities.

Proposition 2. *For (X,Y) a source as above, the following inequalities hold*

$$Z(X|Y)^2 \leq H(X|Y) \tag{4}$$

$$H(X|Y) \leq \log(1 + Z(X|Y)). \tag{5}$$

Either both inequalities are strict or both hold with equality. For equality to hold, it is necessary and sufficient that X conditioned on Y is either deterministic or Ber($\frac{1}{2}$).

The proof is given in the appendix.

These inequalities serve the purpose of showing that $H(X|Y)$ is near 0 or 1 if and only if $Z(X|Y)$ is near 0 or 1, respectively. Hence, the parameters $\{H(U_i|Y^N,U^{i-1})\}_{i=1}^N$ and $\{Z(U_i|Y^N,U^{i-1})\}_{i=1}^N$ polarize simultaneously.

For coding theorems, it is important to have a rate of convergence result.

Definition 1. *Let (X,Y) be a source as above, and let $R > 0$. For $N = 2^n$, $n \geq 1$, let $E_{X|Y}(N,R)$ denote a subset of $\{1,\ldots,N\}$ such that $|E_{X|Y}(N,R)| = \lceil NR \rceil$ and $Z(U_i|Y^N,U^{i-1}) \leq Z(U_j|Y^N,U^{j-1})$ for all $i \in E_{X|Y}(N,R)$ and $j \notin E_{X|Y}(N,R)$. We refer to $E_{X|Y}(N,R)$ as a "high-entropy" (index) set of rate R and block-length N. For the special case where Y is absent or unavailable, we write $E_X(N,R)$ to denote the high-entropy set of X only. When N and R are clear from the context, we simplify the notation by writing $E_{X|Y}$ or E_X.*

Theorem 2. *Let (X,Y) be a source as above and $R > H(X|Y)$ be fixed. Consider a sequence of high-entropy sets $\{E_{X|Y}(N,R) : N = 2^n, n \geq 1\}$. For any such sequence, any fixed $\beta < \frac{1}{2}$, and asymptotically in N, we have*

$$\sum_{i \in E_{X|Y}^c(N,R)} Z(U_i|Y^N,U^{i-1}) = O(2^{-N^\beta}). \tag{6}$$

We omit the proof, which is covered by the results of [4].

III. Lossless source coding

Let (X,Y) be a source as in the previous section and (X^N,Y^N) denote an output block of length $N \geq 1$ produced by this source. Shannon's lossless source coding theorem states that an encoder can compress (X^N,Y^N) into a codeword of length roughly $NH(X|Y)$ bits so that a decoder observing the codeword and Y^N can recover X^N reliably, provided N is sufficiently large. We now describe a method based on polarization that achieves this compression bound. In the absence of any side information Y^N, the method given here is algorithmically identical to the source coding method proposed in [2] and [3]; however, our viewpoint is different. Instead of reducing the source coding problem to a channel coding problem by exploiting a duality relationship between the two problems, we use direct arguments based solely on source polarization.

Fix $N = 2^n$ for some $n \geq 1$. Fix $R > H(X|Y)$ and a high-entropy set $E_{X|Y} = E_{X|Y}(N,R)$.

Encoding: Given a realization $X^N = x^N$, compute $u^N = x^N G_N$ and output $u_{E_{X|Y}}$ as the compressed word. (Note that

the encoder does not require knowledge of the realization of Y^N to implement this scheme.)

Decoding: Having received $u_{E_{X|Y}}$ and observed the realization $Y^N = y^N$, the decoder sequentially builds an estimate $\hat{u}^N$ of u^N by the rule

$$\hat{u}_i = \begin{cases} u_i & \text{if } i \in E_{X|Y} \\ 0 & \text{if } i \in E^c_{X|Y} \text{ and } L^{(i)}_N(y^N, \hat{u}^{i-1}) \geq 1 \\ 1 & \text{else} \end{cases}$$

where

$$L^{(i)}_N(y^N, \hat{u}^{i-1}) = \frac{\Pr(U_i = 0 | Y^N = y^N, U^{i-1} = \hat{u}^{i-1})}{\Pr(U_i = 1 | Y^N = y^N, U^{i-1} = \hat{u}^{i-1})}$$

is a likelihood ratio, which can be computed recursively using the formulas:

$$L_N{}^{(2i-1)}(y^N, u^{2i-2}) = \frac{L^{(i)}_{N/2}(y^{N/2}, u_o^{2i-2} \oplus u_e^{2i-2}) L^{(i)}_{N/2}(y^N_{N/2+1}, u_e^{2i-2}) + 1}{L^{(i)}_{N/2}(y^{N/2}, u_o^{2i-2} \oplus u_e^{2i-2}) + L^{(i)}_{N/2}(y^N_{N/2+1}, u_e^{2i-2})}$$

and

$$L^{(2i)}_N(y^N, u^{2i-1}) = L^{(i)}_{N/2}(y^{N/2}, u_o^{2i-2} \oplus u_e^{2i-2})^{\delta_i} L^{(i)}_{N/2}(y^N_{N/2+1}, u_e^{2i-2})$$

where u_o^{2i-2} and u_e^{2i-2} denote, respectively, the parts of u^{2i-2} with odd and even indices, and δ_i equals 1 or -1 according to u_{2i-1} being 0 or 1, respectively. Having constructed $\hat{u}^N$, the decoder outputs $\hat{x}^N = \hat{u}^N G_N^{-1}$ as the estimate of x^N. (It is easy to verify that $G_N^{-1} = G_N$.)

Performance: The performance of the decoder is measured by the probability of error

$$P_e = \Pr(\hat{U}^N \neq U^N) = \Pr(\hat{U}_{E^c_{X|Y}} \neq U_{E^c_{X|Y}}),$$

which can be upper-bounded by standard (union-bound) techniques as

$$P_e \leq \sum_{i \in E^c_{X|Y}(N,R)} Z(U_i | Y^N, U^{i-1}). \tag{7}$$

The following is a simple corollary to Theorem 2 and (7).

Theorem 3. *For any fixed $R > H(X|Y)$ and $\beta < \frac{1}{2}$, the probability of error for the above polar source coding method is bounded as $P_e = O(2^{-N^\beta})$.*

Complexity: The complexity of encoding and that of decoding are both $O(N \log N)$.

IV. Application to channel coding: Duality

The above source coding scheme can be used to design a capacity-achieving code for any binary-input memoryless channel. Let such a channel be defined by the transition probabilities $W(y|x)$, $x \in \mathcal{X} = \{0, 1\}$ and $y \in \mathcal{Y}$. Consider the block coding scheme shown in Fig. 3, where signals flow from right to left. Here, $N = 2^n$, $n \geq 1$, is the code block length; U^N denotes the message vector, $X^N = U^N G_N$ the channel input vector, and Y^N the channel output vector. Due

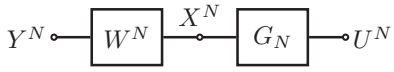

Fig. 3. Channel coding.

to memorylessness, $W^N(y^N|x^N) = \prod_{i=1}^N W(y_i|x_i)$ for any $x^N \in \mathcal{X}^N$, $y^N \in \mathcal{Y}^N$.

We turn the triple (U^N, X^N, Y^N) into a joint ensemble of random vectors by assigning the probabilities $\Pr(X^N = x^N) = 2^{-N}$ for all $x^N \in \{0, 1\}^N$. Under this assignment, (X^N, Y^N) may be regarded as independent samples from a source $(X, Y) \sim Q(x)W(y|x)$ where Q is the uniform distribution on $\{0, 1\}$. We let $I(W) = I(X; Y)$ denote the symmetric channel capacity and fix $R < I(W)$. This implies that $1 - R > H(X|Y)$. Let $E_{X|Y} = E_{X|Y}(N, 1 - R)$ denote a high-entropy set of rate $(1 - R)$ for the source (X, Y). The following coding scheme achieves reliable communication at rate R over the channel W.

Encoding: Prepare a binary source vector U^N as follows. Pick the pattern $U_{E_{X|Y}}$ at random from the uniform distribution and make it available to the decoder ahead of the session. In each round, fill $U_{E^c_{X|Y}}$ with uniformly chosen data bits. (Thus, $\lfloor NR \rfloor$ bits are sent in each round, for a data transmission rate of roughly R.) Encode U^N into a channel codeword by computing $X^N = U^N G_N$ and transmit X^N over the channel W.

Decoding: Having received Y^N, use the source decoder of the previous section to produce an estimate $\hat{U}_{E^c_{X|Y}}$ of the data bits $U_{E^c_{X|Y}}$.

Analysis: The error probability $\Pr(\hat{U}_{E^c_{X|Y}} \neq U_{E^c_{X|Y}})$ is bounded as $O(2^{-N^\beta})$ for any fixed $\beta < \frac{1}{2}$ since the source coding rate is $1 - R > H(X|Y)$. The complexity of the scheme is bounded as $O(N \log N)$.

Remark. The above argument reduces the channel coding problem for achieving the symmetric capacity $I(W)$ of a binary-input channel W to a source coding problem for a source $(X, Y) \sim QW$ where Q is uniform on $\{0, 1\}$. This reduction exploits the duality of the two problems. This dual approach provides an alternative proof of the channel coding results of [1]. It also complements the duality arguments in [2] and [3], where the source coding problem for a Ber(p) source was reduced to a channel coding problem for a binary symmetric channel with cross-over probability p.

V. Slepian-Wolf Coding

The above source coding method can be easily extended to the Slepian-Wolf setting [5]. Suppose $\{(X_i, Y_i)\}_{i=1}^\infty$ are independent samples from a source (X, Y) where both X and Y are binary RVs. In the Slepian-Wolf scenario, there are two encoders and one decoder. Fix a block-length $N = 2^n$, $n \geq 1$, and rates R_x and R_y for the two encoders. Encoder 1 observes X^N only and maps it to an integer $i_x \in [1, 2^{NR_x}]$, encoder 2 observes Y^N only and maps it to an integer $i_y \in [1, 2^{NR_y}]$. The decoder in the system observes (i_x, i_y)

and tries to recover (X^N, Y^N) with vanishing probability of error. The well-known Slepian-Wolf theorem states that this is possible provided $R_x \geq H(X|Y)$, $R_y \geq H(Y|X)$, and $R_x + R_y \geq H(X,Y)$.

It is straightforward to design a polar coding scheme that achieves the corner point $(H(X|Y), H(Y))$ of the Slepian-Wolf rate region. Fix $R_y > H(Y)$ and $R_x > H(X|Y)$. For $N = 2^n$, $n \geq 1$, consider a pair of high-entropy sets $E_Y = E_Y(N, R_y)$ and $E_{X|Y} = E_{X|Y}(N, R_x)$.

Encoding: Given a realization $X^N = x^N$, encoder 1 calculates $u^N = x^N G_N$ and sends $u_{E_{X|Y}}$ to the common decoder. Given a realization $Y^N = y^N$, encoder 2 calculates $v^N = y^N G_N$ and sends v_{E_Y}.

Decoding: The decoder first applies the decoding algorithm of Section III to obtain an estimate $\hat{y}^N$ of y^N from v_{E_Y}. Next, the decoder applies the same algorithm to obtain an estimate of x^N using $\hat{y}^N$ (as a substitute for the actual realization y^N) and $u_{E_{X|Y}}$.

We omit the analysis of this scheme since it essentially consists of two single-user source coding schemes of the type treated in Section III.

It is clear that polar coding can achieve all points of the Slepian-Wolf region by time-sharing between the corner points $(H(X), H(X|Y))$ and $(H(X|Y), H(Y))$.

We should remark that polar coding for Slepian-Wolf problem was first studied in [6], [2], and [3] under the assumptions that $X, Y \sim \text{Ber}(\frac{1}{2})$, and $X \oplus Y \sim \text{Ber}(p)$.

The above approach to Slepian-Wolf coding reduces the problem to single-user source coding problems. A direct appoach would be to have each encoder apply polar transforms locally, with encoder 1 computing $U^N = X^N G_N$ and encoder 2 computing $V^N = Y^N G_N$. Preliminary analyses show that such local operations polarize X_1^N and Y_1^N not only individually but also in a joint sense. A detailed study of such schemes is left for future work.

VI. Polarization of non-binary memoryless sources

Theorem 4. *Let $X \sim P_X$ be a memoryless source over $\mathcal{X} = \{0, 1, \ldots, q-1\}$ for some prime $q \geq 2$. For $n \geq 1$ and $N = 2^n$, let $X^N = (X_1, \ldots, X_N)$ be N independent drawings from the source X. Let $U^N = X^N G_N$ where G_N is as defined in* (3) *but the matrix operation is now carried out in GF(q). Then, the polarization limits in Theorem 1 remain valid provided the entropy terms are calculated with respect to base-q logarithms.*

If q is not prime, the theorem may fail. Consider X over $\{0,1,2,3\}$ with $P_X(0) = P_X(2) = \frac{1}{2}$. Then, it is straightforward to check that U^N has the same distribution as X^N for all N. On closer inspection, we realize that X is actually a binary source under disguise. More precisely, X is already polarized over $\{0,2\}$, which is a subfield of $GF(4)$, and vectors over this subfield are closed under multiplication by G_N.

The preceding example illustrates the difficulties in making a general statement regarding source polarization over arbitrary alphabets. If we introduce some randomness into the construction as in [7], it is possible to polarize sources over arbitrary alphabets, still maintaining the $O(N \log N)$ complexity of the construction.

Acknowledgment

Helpful discussions with E. Şaşoğlu and S. B. Korada are gratefully acknowledged. This work was supported in part by The Scientific and Technological Research Council of Turkey (TÜBİTAK) under contract no. 107E216, and in part by the European Commission FP7 Network of Excellence NEWCOM++ (contract no. 216715).

VII. Appendix

A. Proof of Inequality (4)

First we prove that $Z(X)^2 \leq H(X)$ for any $X \sim \text{Ber}(p)$ with equality if and only if $p \in \{0, \frac{1}{2}, 1\}$. Let $F(p) = H(Z) - Z(X)^2 = -p\log_2(p) - (1-p)\log_2(1-p) - 4p(1-p)$, and compute

$$\frac{dF}{dp} = \frac{1}{\ln 2}\left[-\ln p + \ln(1-p)\right] - 4 + 8p,$$

$$\frac{d^2F}{dp^2} = \frac{1}{\ln 2}\left[-\frac{1}{p} - \frac{1}{1-p}\right] + 8,$$

$$\frac{d^3F}{dp^3} = \frac{1}{\ln 2}\left[\frac{1}{p^2} - \frac{1}{(1-p)^2}\right].$$

Inspection of the third order derivative shows that dF/dp is strictly convex for $p \in [0, \frac{1}{2})$ and strictly concave for $p \in (\frac{1}{2}, 1]$. Thus, $dF/dp = 0$ can have at most one solution in each interval $[0, \frac{1}{2})$ and $(\frac{1}{2}, 1]$. Since $dF/dp = 0$ at $p = \frac{1}{2}$, the number of zeros of dF/dp over $[0,1]$ is at most three. Thus, $F(p)$ can have at most three zeros over $[0,1]$. Since $F(p) = 0$ for $p \in \{0, \frac{1}{2}, 1\}$, there can be no other zeros.

Thus, for any pair of random variables (X, Y) with X binary, if we condition on $Y = y$, we have

$$Z(X|Y=y)^2 \leq H(X|Y=y).$$

Averaging over Y, and by Jensen's inequality, we obtain (4).

B. Proof of Inequality (5)

Recall that the Rényi entropy of order α ($\alpha > 0$, $\alpha \neq 1$) for a RV X is defined as

$$H_\alpha(X) = \frac{1}{1-\alpha}\log\sum_x P_X(x)^\alpha$$

and has the following properties [8].

- $H_\alpha(X)$ is strictly decreasing in α unless P_X is uniform on its support $\text{Supp}(X) = \{x : P_X(x) > 0\}$.
- $H(X) = \lim_{\alpha \to 1} H_\alpha(X)$.

Now suppose $X \sim \text{Ber}(p)$ and note that

$$H_{\frac{1}{2}}(X) = \log\left[\sum_x \sqrt{P_X(x)}\right]^2 = \log(1 + Z(X)).$$

Thus, we have

$$H(X) \leq H_{\frac{1}{2}}(X) = \log(1 + Z(X)).$$

It follows that, for any jointly distributed pair (X, Y) with X binary and any sample value $Y = y$

$$H(X|Y = y) \leq \log(1 + Z(X|Y = y)).$$

Averaging over Y and by Jensen's inequality, we obtain (5).

REFERENCES

[1] E. Arıkan, "Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," *IEEE Trans. Inform. 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. 2009 IEEE Int. Symp. Inform. Theory*, (Seoul, South Korea), pp. 1488–1492, 28 June - 3 July 2009.

[3] S. B. Korada, *Polar codes for channel and source coding*. PhD thesis, EPFL, Lausanne, 2009.

[4] E. Arıkan and E. Telatar, "On the rate of channel polarization," in *Proc. 2009 IEEE Int. Symp. Inform. Theory*, (Seoul, South Korea), pp. 1493–1495, 28 June - 3 July 2009.

[5] D. Slepian and J. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inform. Theory*, vol. 19, pp. 471–480, Jul. 1973.

[6] N. Hussami, S. B. Korada, and R. L. Urbanke, "Polar codes for channel and source coding." http://arxiv.org/abs/0901.2370, 2009.

[7] E. Şaşoğlu, E. Telatar, and E. Arıkan, "Polarization for arbitrary discrete memoryless channels," *CoRR*, vol. abs/0908.0302, 2009.

[8] I. Csiszár, "Generalized cutoff rates and Rényi's information measures," *IEEE Trans. Inform. Theory*, vol. 41, pp. 26–34, Jan. 1995.