

Conditions for Robustness of Polar Codes in the Presence of Channel Mismatch

Mine Alsan

Information Theory Laboratory
Ecole Polytechnique Fédérale de Lausanne
CH-1015 Lausanne, Switzerland
Email: mine.alsan@epfl.ch

Abstract

A challenging problem related to the design of polar codes is “robustness against channel parameter variations”, as stated in Arıkan’s original work. In this paper, we describe how the problem of robust polar code design can be viewed as a mismatch decoding problem. We study channel conditions under which a polar encoder/decoder designed for a mismatched B-DMC can be used to communicate reliably.

I. INTRODUCTION

In 2007, Arıkan [1] proposed polar codes as an appealing error correction method based on a phenomenon called channel polarization. This class of codes are proved to achieve the symmetric capacity of any binary discrete memoryless channel (B-DMC) using low complexity encoders and decoders, and their block error probability is shown to decrease exponentially in the square root of the blocklength [2].

Two basic channel transformations lie at the heart of channel polarization. The two successive channels characterized by these transformations $W^- : \mathcal{X} \rightarrow \mathcal{Y}^2$ and $W^+ : \mathcal{X} \rightarrow \mathcal{Y}^2 \times \mathcal{X}$ can be defined by the following transition probabilities

$$W^-(y_1 y_2 | u_1) = \sum_{u_2 \in \mathcal{X}} \frac{1}{2} W(y_1 | u_1 \oplus u_2) W(y_2 | u_2)$$
$$W^+(y_1 y_2 u_1 | u_2) = \frac{1}{2} W(y_1 | u_1 \oplus u_2) W(y_2 | u_2).$$

For a blocklength $N = 2$, these channels would be indexed as $W_2^{(1)}$ and $W_2^{(2)}$. In general $N = 2^n$, and the channels $W_N^{(i)}$ are synthesized by the recursive applications of these plus/minus transformations until sufficiently polarized, i.e. they are perfect or completely noisy channels.

The polarization idea is used to propose polar codes, and the recursive process leads to efficient encoding and decoding structures. On the encoder side, uncoded data bits are sent only through those perfect channels. For the rest, bits are fixed beforehand and revealed to the decoder as well. On the decoder side, the synthesized channels lend themselves to a particular decoding procedure referred to as successive cancellation decoder (SCD). At the i -th stage, on those good channels, the SCD estimates the channel input u_i with law $W_N^{(i)}(y_1^N u_1^{i-1} | u_i)$ according to maximum likelihood decision rule for the i -th channel using the previous estimates \hat{u}_1^{i-1} , and supplies the new estimate \hat{u}_i to the next stages. The analysis carried in [1] shows that this SCD performs with vanishing error probability.

A particular aspect of polar codes is that they are channel specific designs. The polarization process is adjusted to the particular channel at hand, whence the index set of the synthesized good channels. This set, referred as the information set \mathcal{A} , is required both by the encoder and decoder. The situation in which this knowledge is partially missing have been already addressed. Let W and V be two given B-DMCs. The following two cases are known to lead to an ordering $\mathcal{A}_V \subset \mathcal{A}_W$: If V is a binary erasure channel (BEC) with larger Bhattacharyya parameter than the channel W , or V is a stochastically degraded version of W [1]. These results help the designer to use safely the information set designed for the channel V for communication over W .

On the other hand, a critical point is the assumption of the availability of the channel knowledge at the decoder. Indeed, the described SCD not only requires the information set but also the exact channel knowledge to function. Therefore, if the true channel is unknown, the code design, including the decoding rule, should be based on a mismatched channel [3]. In this work, we assume the same SCD rule is kept, but instead of the true channel law a different one is employed in the decision procedure. As usual, we want to communicate reliably over the unknown communication channel using the polar code designed for the mismatched channel.

The article follows with the problem statement in the next section, where we derive expressions to analyze the error performance of polar codes designed for mismatched channels. In the subsequent section, the main results are explored. The final section is the conclusions.

II. PROBLEM STATEMENT

We want to assess the performance of polar codes over an unknown channel for mismatched designs. For that purpose, we revisit suitable expressions derived for the average probability of error under SCD with respect to a mismatched channel in [4]. These derivations follow the matched counterparts in [1] with a slight modification in the decision rule.

The SCD described in the introduction is closely tied to a channel splitting operation. After channel combining, the splitting synthesizes the channels whose transition probabilities are given by:

$$W_N^{(i)}(y_1^N u_1^{i-1} | u_i) = \sum_{u_{i+1}^N} \frac{1}{2^{N-i}} W_N(y_1^N | u_1^N).$$

We define the likelihood ratio (LR) of a given B-DMC W as $L_W(y) = \frac{W(y|1)}{W(y|0)}$. Decision functions similar to ML decoding rule can then be defined by

$$d_W^{(i)}(y_1^N, \hat{u}_1^{i-1}) = \begin{cases} 0, & \text{if } L_{W_N^{(i)}}(y_1^N, \hat{u}_1^{i-1}) < 1 \\ 1, & \text{if } L_{W_N^{(i)}}(y_1^N, \hat{u}_1^{i-1}) > 1 \\ *, & \text{if } L_{W_N^{(i)}}(y_1^N, \hat{u}_1^{i-1}) = 1 \end{cases}$$

where $*$ is chosen from the set $\{0, 1\}$ by a fair coin flip.

The polar SCD will decode the received output in N stages using a chain of estimators from $i = 1, \dots, N$, each depending on the previous ones. The estimators are defined as

$$\hat{u}_i = \begin{cases} u_i, & \text{if } i \in \mathcal{A}_N^c \\ d_W^{(i)}(y_1^N, \hat{u}_1^{i-1}), & \text{if } i \in \mathcal{A}_N. \end{cases}$$

Let $Pe(W, V, \mathcal{A}_N)$ denote the best achievable block error probability over the ensemble of all possible choices of the set \mathcal{A}_N^c when $|\mathcal{A}_N| = \lfloor NR \rfloor$ under mismatched successive cancellation

decoding with respect to the channel V when the true channel is W . Then, one can show that

$$\begin{aligned}
Pe(W, V, \mathcal{A}_N) &= \mathbb{P}_W \left[\bigcup_{i \in \mathcal{A}_N} \left\{ \hat{U}_1^{i-1} = u_1^{i-1} \cap \hat{U}_i \neq u_i \right\} \right] \\
&= \mathbb{P}_W \left[\bigcup_{i \in \mathcal{A}_N} \left\{ \hat{U}_1^{i-1} = u_1^{i-1} \cap d_V^{(i)}(y_1^N, \hat{U}_1^{i-1}) \neq u_i \right\} \right] \\
&\leq \mathbb{P}_W \left[\bigcup_{i \in \mathcal{A}_N} \left\{ d_V^{(i)}(y_1^N, u_1^{i-1}) \neq u_i \right\} \right] \\
&\leq \sum_{i \in \mathcal{A}_N} Pe_N^{(i)}(W, V)
\end{aligned}$$

where we have defined $Pe_N^{(i)}(W, V)$ as

$$\begin{aligned}
Pe_N^{(i)}(W, V) &\triangleq \sum_{y_1^N, u_1^N} \frac{1}{2^N} W_N(y_1^N | u_1^N) \mathbf{1} \left\{ \frac{V_N^{(i)}(y_1^N, u_1^{i-1} | u_i \oplus 1)}{V_N^{(i)}(y_1^N, u_1^{i-1} | u_i)} > 1 \right\} \\
&\quad + \frac{1}{2} \sum_{y_1^N, u_1^N} \frac{1}{2^N} W_N(y_1^N | u_1^N) \mathbf{1} \left\{ \frac{V_N^{(i)}(y_1^N, u_1^{i-1} | u_i \oplus 1)}{V_N^{(i)}(y_1^N, u_1^{i-1} | u_i)} = 1 \right\}.
\end{aligned}$$

with $\mathbf{1}\{\cdot\}$ denoting the indicator function as usual.

For symmetric channels, the next proposition can be proved similarly to Corollary 1 in [1].

Proposition 1: Let W and V be symmetric B-DMCs. Then,

$$\begin{aligned}
Pe_N^{(i)}(W, V) &= \sum_{y_1^N} W(y_1^N | 0_1^N) \mathbf{1} \{ L_{V_N^{(i)}}(y_1^N, 0_1^N) > 1 \} \\
&\quad + \frac{1}{2} \sum_{y_1^N} W(y_1^N | 0_1^N) \mathbf{1} \{ L_{V_N^{(i)}}(y_1^N, 0_1^N) = 1 \}. \quad (1)
\end{aligned}$$

For shorthand notation we will use $L_{V_N^{(i)}}(y_1^N) \triangleq L_{V_N^{(i)}}(y_1^N, 0_1^{i-1})$. However, one should keep in mind that this is the LR given that the all zero sequence have been sent through the channel.

We define the function

$$\mathbf{H} \left(L_{V_N^{(i)}}(y_1^N) \right) \triangleq \mathbf{1} \{ L_{V_N^{(i)}}(y_1^N) > 1 \} + \frac{1}{2} \mathbf{1} \{ L_{V_N^{(i)}}(y_1^N) = 1 \},$$

so that

$$Pe_N^{(i)}(W, V) = \sum_{y_1^N} W(y_1^N | 0_1^N) \mathbf{H} \left(L_{V_N^{(i)}}(y_1^N) \right) = \mathbb{E}_W \left[\mathbf{H} \left(L_{V_N^{(i)}}(y_1^N) \right) \right].$$

We also define the ‘‘complement’’ function of \mathbf{H} as

$$\overline{\mathbf{H}} \left(L_{V_N^{(i)}}(y_1^N) \right) \triangleq \mathbf{1}\{L_{V_N^{(i)}}(y_1^N) < 1\} + \frac{1}{2}\mathbf{1}\{L_{V_N^{(i)}}(y_1^N) = 1\}.$$

We will often use $\mathbb{P}_W[\cdot] \triangleq \mathbb{E}_{W(y_1^N|0_1^N)}[\mathbf{1}\{\cdot\}]$ notation. Whenever we state a result concerning the parameters $Pe_N^{(i)}$, it should be understood that $N = 2^n$, $n = 1, 2, \dots$, and for a fixed N the possible values for the indices are $i = 1, \dots, N$. To be concise, we will avoid repeating these.

The next two propositions explore the recursive structure of the LR computations after applying the polar transformations.

Proposition 2: [1] The LRs satisfy the recursion

$$L_{V_{2N}^{(2i-1)}}(y_1^{2N}) = \frac{L_{V_N^{(i)}}(y_1^N) + L_{V_N^{(i)}}(y_{N+1}^{2N})}{1 + L_{V_N^{(i)}}(y_1^N)L_{V_N^{(i)}}(y_{N+1}^{2N})},$$

$$L_{V_{2N}^{(2i)}}(y_1^{2N}) = L_{V_N^{(i)}}(y_1^N)L_{V_N^{(i)}}(y_{N+1}^{2N}).$$

Hence, the computed LRs can be seen as a symmetric functions $f(L_{V_N^{(i)}}(y_1^N), L_{V_N^{(i)}}(y_{N+1}^{2N}))$ of the arguments.

Proposition 3: The quantities $Pe_{2N}^{(i)}(W, V)$ can be recursively computed as

$$Pe_{2N}^{(2i-1)}(W, V) = \sum_{y_1^{2N}} W(y_1^{2N}|0_1^{2N}) \mathbf{H} \left(\frac{L_{V_N^{(i)}}(y_1^N) + L_{V_N^{(i)}}(y_{N+1}^{2N})}{1 + L_{V_N^{(i)}}(y_1^N)L_{V_N^{(i)}}(y_{N+1}^{2N})} \right),$$

$$Pe_{2N}^{(2i)}(W, V) = \sum_{y_1^{2N}} W(y_1^{2N}|0_1^{2N}) \mathbf{H} \left(L_{V_N^{(i)}}(y_1^N)L_{V_N^{(i)}}(y_{N+1}^{2N}) \right).$$

Proof of Proposition 3: The expression directly follows from Proposition 1, and the recursive structure stated in Proposition 2. \blacksquare

Upper Bounds to $Pe_N^{(i)}(W, V)$: Now, we define two channel parameters that can be used to upper bound $Pe_N^{(i)}(W, V)$ for symmetric channels. The first one is simply

$$Pe_N^{(i)}(W, V) \leq \mathbb{P}_{W_0} \left[L_{V_N^{(i)}}(y_1^N) \geq 1 \right].$$

The second parameter, analogous to the Bhattacharyya parameter defined for the matched scenario and referred to as the mismatched version of this quantity, can be defined for symmetric channels as

$$Z(W, V) = \sum_y W(y|0) \sqrt{L_V(y)}.$$

One can easily show $Pe_N^{(i)}(W, V) \leq Z_N^{(i)}(W, V) \triangleq Z(W_N^{(i)}, V_N^{(i)})$ by extending the definition to the i -th synthesized channels.

Mismatch Performance Analysis over BSCs using an Approximation on the Decoder Side:

Consider the case where W and V are BSCs with crossover probabilities $\epsilon_W, \epsilon_V \leq 0.5$. Theorem 1 in [4] shows that replacing the minus polar transformation with the min approximation as initially proposed in [5] for efficient hardware implementations of polar codes, i.e,

$$\log(L_{V_{2^{i-1}}}^{(2^i)}(y_1^{2^N})) = -\text{sign}(\ell_1 * \ell_2) \min\{|\ell_1|, |\ell_2|\} \quad (2)$$

where $\ell_1 \triangleq \log(L_{V_N^{(i)}}(y_1^N))$, $\ell_2 \triangleq \log(L_{V_{N+1}^{(i)}}(y_{N+1}^{2^N}))$, results in the LRs of the synthesized channels $W_N^{(i)}$ and $V_N^{(i)}$ to be ordered for each $i = 1$ as

$$\begin{aligned} 1 &\leq \tilde{L}_{V_N^{(i)}}(y_1^N) \leq \tilde{L}_{W_N^{(i)}}(y_1^N), \\ \text{or } \tilde{L}_{W_N^{(i)}}(y_1^N) &\leq \tilde{L}_{V_N^{(i)}}(y_1^N) \leq 1, \end{aligned}$$

where the symbol $\tilde{\cdot}$ indicates computations use the approximation. Indeed, the LRs of the worst BSC will be closer to 1. So, the decoder estimate for a given output realization will be identical whether the computations are performed with respect to the LRs of the channel W , or the channel V , as long as $\epsilon_W, \epsilon_V \leq 0.5$. In this case: $\tilde{P}e_N^{(i)}(W, V) = \tilde{P}e_N^{(i)}(W)$, for any $i = 1, \dots, N$, whence the information sets of the matched and mismatched designs are identical.

Even though we do not expect in the more general cases the ordering of the LRs to hold, averaging might once again be the ‘savior’ of the code performance so that we have at least the hope to show “ $Pe_N^{(i)}(W, V) \leq Pe_N^{(i)}(V)$ ” or “ $Pe_N^{(i)}(W, V) \leq Z(V_N^{(i)})$ ” type of relations. Provided any such orderings hold for all $i \in \mathcal{A}_N$, one can guarantee reliability in the mismatch scenario: the mismatch polar code designed for the channel V could be used to communicate reliably over the channel W , achieving rates up to the symmetric capacity of the mismatched channel V .

III. RESULTS

One way to guarantee reliability in the mismatch scenario would be to ensure the sign of the difference $Pe_N^{(i)}(W, V) - Pe_N^{(i)}(V)$ is preserved throughout the recursive application of the polarization transformations. In this section, we first push the analysis as far as possible with no particular assumptions on the channels. Then, we impose conditions that would ensure the design goal is met.

To investigate the sign of the difference $Pe_N^{(i)}(W, V) - Pe_N^{(i)}(V)$, we introduce a set of propositions.

In the next proposition and subsequent corollary we derive equivalent expressions to compute the difference $Pe_N^{(i)}(W, V) - Pe_N^{(i)}(V)$ using the symmetry in the recursive computations of the LRs.

Proposition 4: For symmetric B-DMCs W and V , we have

$$Pe_{2N}^{(i)}(W, V) - Pe_{2N}^{(i)}(V) = \sum_{y_1^N} [W(y_1^N | 0_1^N) - V(y_1^N | 0_1^N)] \times \left(\sum_{y_{N+1}^{2N}} [W(y_{N+1}^{2N} | 0_1^N) + V(y_{N+1}^{2N} | 0_1^N)] \mathbf{H} \left(L_{V_{2N}^{(i)}}(y_1^{2N}) \right) \right). \quad (3)$$

Proof of Proposition 4: We develop the RHS of Equation (3)

$$\begin{aligned} & \sum_{y_1^{2N}} [W(y_1^{2N} | 0_1^{2N}) - V(y_1^{2N} | 0_1^{2N})] \mathbf{H}\{L_{V_{2N}^{(i)}}(y_1^{2N}) \geq 1\} \\ & + \sum_{y_1^{2N}} W(y_1^N | 0_1^N) V(y_{N+1}^{2N} | 0_1^N) \mathbf{H}\{f(L_{V_N^{(i)}}(y_1^N), L_{V_N^{(i)}}(y_{N+1}^{2N})) \geq 1\} \\ & - \sum_{y_1^{2N}} W(y_{N+1}^{2N} | 0_1^N) V(y_1^N | 0_1^N) \mathbf{H}\{f(L_{V_N^{(i)}}(y_{N+1}^{2N}), L_{V_N^{(i)}}(y_1^N)) \geq 1\} \\ & = Pe_{2N}^{(i)}(W, V) - Pe_{2N}^{(i)}(V), \end{aligned}$$

where we used the symmetry of the function f described in Proposition 2. ■

Corollary 1: For symmetric B-DMCs W and V we have

$$Pe_N^{(i)}(W, V) - Pe_N^{(i)}(V) = \sum_{y_1^N} [W(y_1 | 0) - V(y_1 | 0)] [W(y_2 | 0) + V(y_2 | 0)] \times \prod_{i=2}^n (W(y_{2^{i-1}+1} | 0) \dots W(y_{2^i} | 0) + V(y_{2^{i-1}+1} | 0) \dots V(y_{2^i} | 0)) \mathbf{H}\{L_{V_{2N}^{(i)}}(y_1^{2N}) \geq 1\}.$$

Now, we look for properties which are preserved under the one step transformations. These are surveyed in the following two propositions.

Proposition 5: For a symmetric B-DMC channel V such that the condition

$$\mathbb{P}_{V_0} \left[L_{V_N^{(i)}}(y_1^N) < 1 \right] \geq \mathbb{P}_{V_0} \left[L_{V_N^{(i)}}(y_1^N) > 1 \right]$$

holds for a given i , the basic polarization transformations preserve the inequality, i.e. for $j = 2i - 1, 2i$, we have

$$\mathbb{P}_{V_0} \left[L_{V_{2N}^{(j)}}(y_1^{2N}) < 1 \right] \geq \mathbb{P}_{V_0} \left[L_{V_{2N}^{(j)}}(y_1^{2N}) > 1 \right].$$

Proof of Proposition 5: For simplicity we omit the subscripts in $L_{V_N^{(i)}}(y_1^N) \triangleq L(y_1^N)$ and \mathbb{P}_{V_0} . Note that by symmetry in the construction of polar codes $\mathbb{P}[L(y_1^N) < 1] = \mathbb{P}[L(y_{N+1}^{2N}) < 1]$.

$$\begin{aligned} \mathbb{P}\left[\frac{L(y_1^N) + L(y_{N+1}^{2N})}{1 + L(y_1^N)L(y_{N+1}^{2N})} < 1\right] &= \mathbb{P}[L(y_1^N) < 1] \mathbb{P}[L(y_{N+1}^{2N}) < 1] + \mathbb{P}[L(y_1^N) > 1] \mathbb{P}[L(y_{N+1}^{2N}) > 1] \\ &= \mathbb{P}[L(y_1^N) < 1]^2 + \mathbb{P}[L(y_1^N) > 1]^2, \end{aligned}$$

and

$$\begin{aligned} \mathbb{P}\left[\frac{L(y_1^N) + L(y_{N+1}^{2N})}{1 + L(y_1^N)L(y_{N+1}^{2N})} > 1\right] &= \mathbb{P}[L(y_1^N) < 1] \mathbb{P}[L(y_{N+1}^{2N}) > 1] + \mathbb{P}[L(y_1^N) > 1] \mathbb{P}[L(y_{N+1}^{2N}) < 1] \\ &= 2\mathbb{P}[L(y_1^N) < 1] \mathbb{P}[L(y_1^N) > 1]. \end{aligned}$$

By noting that the difference of these equals

$$\left(\mathbb{P}[L(y_1^N) < 1] - \mathbb{P}[L(y_1^N) > 1]\right)^2 \geq 0,$$

the claim for the minus transformation is proved.

For the plus transformation, we use a property [7] following by the symmetry of the channels

$$\mathbb{P}[L(y_1^N) = \ell] = \frac{1}{\ell} \mathbb{P}\left[L(y_1^N) = \frac{1}{\ell}\right].$$

We define the following notations

$$\begin{aligned} \mathbb{P}[L(y_1^N) \geq 1] &\triangleq \mathbb{P}[L(y_1^N) > 1] + \frac{1}{2} \mathbb{P}[L(y_1^N) = 1] \\ \mathbb{P}[L(y_1^N) \leq 1] &\triangleq \mathbb{P}[L(y_1^N) < 1] + \frac{1}{2} \mathbb{P}[L(y_1^N) = 1] \end{aligned}$$

$$\begin{aligned}
& \mathbb{P} [L(y_1^N)L(y_{N+1}^{2N}) \leq 1] \\
&= \sum_{\ell_1 < 1} \sum_{\ell_2 < 1} \mathbb{P} [L(y_1^N) = \ell_1] \mathbb{P} [L(y_{N+1}^{2N}) = \ell_2] + \sum_{\ell_1 < 1} \sum_{1 \geq \ell_2 < 1/\ell_1} \mathbb{P} [L(y_1^N) = \ell_1] \mathbb{P} [L(y_{N+1}^{2N}) = \ell_2] \\
&\quad + \sum_{\ell_1 \geq 1} \sum_{\ell_2 \leq 1/\ell_1} \mathbb{P} [L(y_1^N) = \ell_1] \mathbb{P} [L(y_{N+1}^{2N}) = \ell_2] - \frac{1}{2} \mathbb{P} [L(y_1^N) = 1]^2 \\
&= \mathbb{P} [L(y_1^N) < 1]^2 - \frac{1}{2} \mathbb{P} [L(y_1^N) = 1]^2 \\
&\quad + \sum_{\ell_1 > 1} \sum_{1 \leq \ell_2 < \ell_1} \ell_1 \mathbb{P} [L(y_1^N) = \ell_1] \mathbb{P} [L(y_{N+1}^{2N}) = \ell_2] \\
&\quad + \sum_{\ell_1 \geq 1} \sum_{\ell_2 \geq \ell_1} \ell_2 \mathbb{P} [L(y_1^N) = \ell_1] \mathbb{P} [L(y_{N+1}^{2N}) = \ell_2] \\
&= \mathbb{P} [L(y_1^N) < 1]^2 - \frac{1}{2} \mathbb{P} [L(y_1^N) = 1]^2 \\
&\quad + \sum_{\ell_1 > 1} \sum_{1 < \ell_2 < \ell_1} \ell_1 \mathbb{P} [L(y_1^N) = \ell_1] \mathbb{P} [L(y_{N+1}^{2N}) = \ell_2] + \mathbb{P} [L(y_1^N) = 1] \sum_{\ell_1 > 1} \ell_1 \mathbb{P} [L(y_1^N) = \ell_1] \\
&\quad + \sum_{\ell_1 > 1} \sum_{\ell_2 \geq \ell_1} \ell_2 \mathbb{P} [L(y_1^N) = \ell_1] \mathbb{P} [L(y_{N+1}^{2N}) = \ell_2] + \mathbb{P} [L(y_1^N) = 1] \sum_{\ell_2 > 1} \ell_2 \mathbb{P} [L(y_1^N) = \ell_2] + \mathbb{P} [L(y_1^N) = 1]^2 \\
&= \mathbb{P} [L(y_1^N) < 1]^2 + \frac{1}{4} \mathbb{P} [L(y_1^N) = 1]^2 + \mathbb{P} [L(y_1^N) = 1] \mathbb{P} [L(y_1^N) < 1] \\
&\quad + \sum_{\ell_1 > 1} \sum_{\ell_2 > 1} \mathbb{P} [L(y_1^N) = \ell_1] \mathbb{P} [L(y_{N+1}^{2N}) = \ell_2] \max\{\ell_1, \ell_2\} \tag{4} \\
&\quad + \mathbb{P} [L(y_1^N) = 1] \sum_{\ell_2 > 1} \ell_2 \mathbb{P} [L(y_1^N) = \ell_2] + \frac{1}{4} \mathbb{P} [L(y_1^N) = 1]^2 \\
&= \mathbb{P} [L(y_1^N) \leq 1]^2 + \sum_{\ell_1 \geq 1} \sum_{\ell_2 \geq 1} \mathbb{P} [L(y_1^N) = \ell_1] \mathbb{P} [L(y_{N+1}^{2N}) = \ell_2] \max\{\ell_1, \ell_2\} \tag{5}
\end{aligned}$$

where we abuse the notation to define

$$\begin{aligned}
& \sum_{\ell_1 \geq 1} \sum_{\ell_2 \geq 1} \mathbb{P} [L(y_1^N) = \ell_1] \mathbb{P} [L(y_{N+1}^{2N}) = \ell_2] \max\{\ell_1, \ell_2\} \\
&= \sum_{\ell_1 > 1} \sum_{\ell_2 > 1} \mathbb{P} [L(y_1^N) = \ell_1] \mathbb{P} [L(y_{N+1}^{2N}) = \ell_2] \max\{\ell_1, \ell_2\} \\
&\quad + \mathbb{P} [L(y_1^N) = 1] \sum_{\ell_2 > 1} \ell_2 \mathbb{P} [L(y_1^N) = \ell_2] + \frac{1}{4} \mathbb{P} [L(y_1^N) = 1]^2.
\end{aligned}$$

In the same spirit, we define

$$\begin{aligned}
& \sum_{\ell_1 \geq 1} \sum_{\ell_2 \geq 1} \mathbb{P} [L(y_1^N) = \ell_1] \mathbb{P} [L(y_{N+1}^{2N}) = \ell_2] \min\{\ell_1, \ell_2\} \\
&= \sum_{\ell_1 > 1} \sum_{\ell_2 > 1} \mathbb{P} [L(y_1^N) = \ell_1] \mathbb{P} [L(y_{N+1}^{2N}) = \ell_2] \min\{\ell_1, \ell_2\} \\
&\quad + \mathbb{P} [L(y_1^N) = 1] \mathbb{P} [L(y_1^N) > 1] + \frac{1}{4} \mathbb{P} [L(y_1^N) = 1]^2,
\end{aligned}$$

and we note that

$$\begin{aligned}
& \sum_{\ell_1 \geq 1} \sum_{\ell_2 \geq 1} \mathbb{P} [L(y_1^N) = \ell_1] \mathbb{P} [L(y_{N+1}^{2N}) = \ell_2] (\max\{\ell_1, \ell_2\} + \min\{\ell_1, \ell_2\}) \\
&= \sum_{\ell_1 \geq 1} \sum_{\ell_2 \geq 1} \mathbb{P} [L(y_1^N) = \ell_1] \mathbb{P} [L(y_{N+1}^{2N}) = \ell_2] (\ell_1 + \ell_2) \\
&= 2 \sum_{\ell_1 \geq 1} \mathbb{P} [L(y_1^N) = \ell_1] \ell_1 \sum_{\ell_2 \geq 1} \mathbb{P} [L(y_{N+1}^{2N}) = \ell_2] \\
&= 2 \mathbb{P} [L(y_1^N) \leq 1] \mathbb{P} [L(y_1^N) \geq 1]. \tag{6}
\end{aligned}$$

As

$$\begin{aligned}
1 &= \mathbb{P} [L(y_1^N)L(y_{N+1}^{2N}) \leq 1] + \mathbb{P} [L(y_1^N)L(y_{N+1}^{2N}) \geq 1] \\
&= (\mathbb{P} [L(y_1^N) \leq 1] + \mathbb{P} [L(y_1^N) \geq 1])^2 \\
&= \mathbb{P} [L(y_1^N) \leq 1]^2 + \mathbb{P} [L(y_1^N) \geq 1]^2 + 2\mathbb{P} [L(y_1^N) \leq 1] \mathbb{P} [L(y_1^N) \geq 1]
\end{aligned}$$

must hold, we get

$$\begin{aligned}
& \mathbb{P} [L(y_1^N)L(y_{N+1}^{2N}) \geq 1] \\
&= \mathbb{P} [L(y_1^N) \geq 1]^2 + \sum_{\ell_1 \geq 1} \sum_{\ell_2 \geq 1} \mathbb{P} [L(y_1^N) = \ell_1] \mathbb{P} [L(y_{N+1}^{2N}) = \ell_2] \min\{\ell_1, \ell_2\} \tag{7}
\end{aligned}$$

Therefore, we (5) and (7) proves that

$$\mathbb{P} [L(y_1^N)L(y_{N+1}^{2N}) < 1] \geq \mathbb{P} [L(y_1^N)L(y_{N+1}^{2N}) > 1]$$

holds as claimed. ■

Proposition 6: Let W and V be symmetric B-DMCs such that

$$\mathbb{E}_W \left[\mathbf{H} \left(L_{V_N^{(i)}}(y_1^N) \right) \right] \leq \mathbb{E}_V \left[\mathbf{H} \left(L_{V_N^{(i)}}(y_1^N) \right) \right], \tag{8}$$

and

$$\mathbb{P}_V \left[L_{V_N^{(i)}}(y_1^N) < 1 \right] \geq \mathbb{P}_V \left[L_{V_N^{(i)}}(y_1^N) > 1 \right] \quad (9)$$

hold. Then,

$$\mathbb{P}_W \left[L_{V_N^{(i)}}(y_1^N) < 1 \right] \geq \mathbb{P}_W \left[L_{V_N^{(i)}}(y_1^N) > 1 \right].$$

Proof of Proposition 6: We have

$$\begin{aligned} & \mathbb{P}_W [L(y_1^N) > 1] + \frac{1}{2} \mathbb{P}_W [L(y_1^N) = 1] - \mathbb{P}_V [L(y_1^N) > 1] - \frac{1}{2} \mathbb{P}_V [L(y_1^N) = 1] \\ &= \mathbb{P}_V [L(y_1^N) < 1] + \frac{1}{2} \mathbb{P}_V [L(y_1^N) = 1] - \mathbb{P}_W [L(y_1^N) < 1] - \frac{1}{2} \mathbb{P}_V [L(y_1^N) = 1] \leq 0, \end{aligned}$$

where the negativity follows by the assumption in (8). Therefore, adding both sides gives

$$\mathbb{P}_W [L(y_1^N) > 1] - \mathbb{P}_V [L(y_1^N) > 1] + \mathbb{P}_V [L(y_1^N) < 1] - \mathbb{P}_W [L(y_1^N) < 1] \leq 0.$$

Hence,

$$\mathbb{P}_W [L(y_1^N) < 1] - \mathbb{P}_W [L(y_1^N) > 1] \geq \mathbb{P}_V [L(y_1^N) < 1] - \mathbb{P}_V [L(y_1^N) > 1] \geq 0,$$

where the non-negativity follows by the assumption in (9). ■

The final proposition of this section plays a key role in the subsequent theorem.

Proposition 7: The quantities $Pe_{2N}^{(i)}(W, V) - Pe_{2N}^{(i)}(V)$ can be recursively computed for $\forall i$ as

$$Pe_{2N}^{(2i-1)}(W, V) - Pe_{2N}^{(2i-1)}(V) = \sum_{y_1^N} [W(y_1^N | 0_1^N) - V(y_1^N | 0_1^N)] \mathbf{H} \left(L_{V_N^{(i)}}(y_1^N) \right) K_N, \quad (10)$$

where

$$K_N = \left(\begin{array}{c} \sum_{\substack{y_{N+1}^{2N} \\ L_{V_N^{(i)}}(y_{N+1}^{2N}) < 1}} [W(y_{N+1}^{2N} | 0_1^N) + V(y_{N+1}^{2N} | 0_1^N)] - \sum_{\substack{y_{N+1}^{2N} \\ L_{V_N^{(i)}}(y_{N+1}^{2N}) > 1}} [W(y_{N+1}^{2N} | 0_1^N) + V(y_{N+1}^{2N} | 0_1^N)] \end{array} \right), \quad (11)$$

and

$$\begin{aligned} Pe_{2N}^{(2i)}(W, V) - Pe_{2N}^{(2i)}(V) &= \sum_{y_1^{2N}} [W(y_1^N | 0_1^N) - V(y_1^N | 0_1^N)] [W(y_{N+1}^{2N} | 0_1^N) + V(y_{N+1}^{2N} | 0_1^N)] \times \\ & \quad \mathbf{H} \left(L_{V_N^{(i)}}(y_1^N, 0_1^{i-1}) L_{V_N^{(i)}}(y_{N+1}^{2N}, 0_1^{i-1}) \right). \end{aligned} \quad (12)$$

Proposition 7: For simplicity we omit the subscripts in $L_{V^{(i)}}(y_1^N) \triangleq L(y_1^N)$. First observe that

$$\mathbf{H} \left(\frac{L(y_1^N) + L(y_{N+1}^{2N})}{1 + L(y_1^N)L(y_{N+1}^{2N})} \right) = \begin{cases} \frac{1}{2}, & \text{if } L(y_1^N) = 1, \\ & \text{or } L(y_{N+1}^{2N}) = 1 \\ 1, & \text{if } L(y_1^N) < 1 \text{ and } L(y_{N+1}^{2N}) > 1, \\ & \text{or } L(y_1^N) > 1 \text{ and } L(y_{N+1}^{2N}) < 1 \\ 0, & \text{if } L(y_1^N) < 1 \text{ and } L(y_{N+1}^{2N}) < 1, \\ & \text{or } L(y_1^N) > 1 \text{ and } L(y_{N+1}^{2N}) > 1 \end{cases}$$

We use the expression derived in Proposition (4) to get

$$\begin{aligned} & Pe_{2N}^{(2i-1)}(W, V) - Pe_{2N}^{(2i-1)}(V) \\ &= \sum_{\substack{y_1^N: \\ L(y_1^N)=1}} [W(y_1^N|0_1^N) - V(y_1^N|0_1^N)] \times 1 \\ &+ \sum_{\substack{y_1^N: \\ L(y_1^N)>1}} [W(y_1^N|0_1^N) - V(y_1^N|0_1^N)] \times \sum_{\substack{y_{N+1}^{2N}: \\ L(y_{N+1}^{2N}) \leq 1}} [W(y_{N+1}^{2N}|0_1^N) + V(y_{N+1}^{2N}|0_1^N)] \\ &+ \sum_{\substack{y_1^N: \\ L(y_1^N)<1}} [W(y_1^N|0_1^N) - V(y_1^N|0_1^N)] \times \sum_{\substack{y_{N+1}^{2N}: \\ L(y_{N+1}^{2N}) \geq 1}} [W(y_{N+1}^{2N}|0_1^N) + V(y_{N+1}^{2N}|0_1^N)] \\ &= \sum_{\substack{y_1^N: \\ L(y_1^N)=1}} [W(y_1^N|0_1^N) - V(y_1^N|0_1^N)] \times 1 \\ &+ \sum_{\substack{y_1^N: \\ L(y_1^N)>1}} [W(y_1^N|0_1^N) - V(y_1^N|0_1^N)] \times \sum_{y_{N+1}^{2N}} [W(y_{N+1}^{2N}|0_1^N) + V(y_{N+1}^{2N}|0_1^N)] \bar{\mathbf{H}}(L(y_{N+1}^{2N})) \\ &+ \sum_{\substack{y_1^N: \\ L(y_1^N)<1}} [W(y_1^N|0_1^N) - V(y_1^N|0_1^N)] \times \sum_{y_{N+1}^{2N}} [W(y_{N+1}^{2N}|0_1^N) + V(y_{N+1}^{2N}|0_1^N)] \mathbf{H}(L(y_{N+1}^{2N})). \end{aligned}$$

By substituting $\bar{\mathbf{H}}(L(y_1^N)) = 1 - \mathbf{H}(L(y_1^N))$ and regrouping the terms, we obtain

$$\begin{aligned} & Pe_{2N}^{(2i-1)}(W, V) - Pe_{2N}^{(2i-1)}(V) \\ &= \sum_{y_1^N} [W(y_1^N|0_1^N) - V(y_1^N|0_1^N)] 2\mathbf{H}(L(y_1^N)) + \sum_{y_1^N} [W(y_1^N|0_1^N) - V(y_1^N|0_1^N)] [1 - 2\mathbf{H}(L(y_1^N))] \times \\ & \quad \sum_{y_{N+1}^{2N}} [W(y_{N+1}^{2N}|0_1^N) + V(y_{N+1}^{2N}|0_1^N)] \mathbf{H}(L(y_{N+1}^{2N})), \end{aligned}$$

where we used the fact that $1 - 2\mathbf{H}(L(y_1^N)) = \mathbf{1}\{L(y_1^N) < 1\} - \mathbf{1}\{L(y_1^N) > 1\}$. Now, note that the term in the second summation with the 1 sums to 0. Hence, we get

$$\begin{aligned} & Pe_{2N}^{(2i-1)}(W, V) - Pe_{2N}^{(2i-1)}(V) \\ &= \sum_{y_1^N} [W(y_1^N | 0_1^N) - V(y_1^N | 0_1^N)] \mathbf{H}(L(y_1^N)) \left[2 - \sum_{y_{N+1}^{2N}} [W(y_{N+1}^{2N} | 0_1^N) + V(y_{N+1}^{2N} | 0_1^N)] 2\mathbf{H}(L(y_{N+1}^{2N})) \right] \\ &= \sum_{y_1^N} [W(y_1^N | 0_1^N) - V(y_1^N | 0_1^N)] \mathbf{H}(L(y_1^N)) \sum_{y_{N+1}^{2N}} [W(y_{N+1}^{2N} | 0_1^N) + V(y_{N+1}^{2N} | 0_1^N)] [1 - 2\mathbf{H}(L(y_{N+1}^{2N}))]. \end{aligned}$$

We recover Equation (10) upon noticing K_N defined in (11) equals

$$K_N = \sum_{y_{N+1}^{2N}} [W(y_{N+1}^{2N} | 0_1^N) + V(y_{N+1}^{2N} | 0_1^N)] [1 - 2\mathbf{H}(L(y_{N+1}^{2N}))],$$

as $1 - 2\mathbf{H}(L(y_{N+1}^{2N})) = \mathbf{1}\{L(y_{N+1}^{2N}) < 1\} - \mathbf{1}\{L(y_{N+1}^{2N}) > 1\}$. ■

We state our first main result.

Theorem 1: Let W and V be B-DMCs such that for a given i , the following conditions hold:

A)

$$\mathbb{P}_V [L_{V_N^{(i)}}(y_1^N) < 1] \geq \mathbb{P}_V [L_{V_N^{(i)}}(y_1^N) > 1],$$

B)

$$Pe_N^{(i)}(W, V) - Pe_N^{(i)}(V) \leq 0.$$

Then, the minus polar transformation preserves the above conditions in the sense that at the next level, they hold for the $2i - 1$ -th indice. On the other hand, while the plus transformation preserves condition A, condition B may not be preserved in general.

Proof of Theorem 1:

A[±]) We know condition A is preserved by Proposition 5 for both transformations.

B⁻) For the minus transformation, we have by Proposition 7

$$Pe_{2N}^{(2i-1)}(W, V) - Pe_{2N}^{(2i-1)}(V) = [Pe_N^{(i)}(W, V) - Pe_N^{(i)}(V)] K_N. \quad (13)$$

Now, we claim that $K_N \geq 0$, from which the sign of $Pe_{2N}^{(2i-1)}(W, V) - Pe_{2N}^{(2i-1)}(V) \leq 0$ follows. To prove the claim, note that by equation (11), the constant K_N equals to

$$\begin{aligned} & \mathbb{P}_{W_0} [L_{V_N^{(i)}}(y_{N+1}^{2N}) < 1] + \mathbb{P}_{V_0} [L_{V_N^{(i)}}(y_{N+1}^{2N}) < 1] \\ & - \mathbb{P}_{W_0} [L_{V_N^{(i)}}(y_{N+1}^{2N}) > 1] - \mathbb{P}_{V_0} [L_{V_N^{(i)}}(y_{N+1}^{2N}) > 1]. \end{aligned}$$

Then, the non-negativity of K_N follows by condition A, and Proposition 6 which shows conditions A and B imply

$$\mathbb{P}_W \left[L_{V_N^{(i)}}(y_1^N) < 1 \right] \geq \mathbb{P}_W \left[L_{V_N^{(i)}}(y_1^N) > 1 \right].$$

B⁺) We give a counterexample: Let W be a binary symmetric channel (BSC) of crossover probability 0.3, and V a B-DMC such that the LRs take the values $\{1/\ell, 1, \ell\}$ such that $\ell > 1$ with probabilities $\{0.4, 0.5, 0.1\}$. One can check that although conditions A and B are satisfied, condition B fails to hold for the channel V^+ . ■

In Theorem 1, we studied the one step preservation properties related to the channel parameter $Pe_N^{(i)}$, and saw that we need to have some constraints on the mismatch channel to be used if we want to ensure condition B is preserved under both transformations. Before we proceed with the second theorem, in which one such constraint is imposed, we make a small digression.

Consider the mismatched Bhattacharyya parameter $Z(W, V)$. The parameter we obtain after applying the plus polar transformation is given by

$$\begin{aligned} Z(W^+, V^+) &= \sum_{y_1 y_2 u_1} W^+(y_1 y_2 u_1 | u_2 = 0) \sqrt{\frac{V^+(y_1 y_2 u_1 | u_2 = 1)}{V^+(y_1 y_2 u_1 | u_2 = 0)}} \\ &= \sum_{y_1 y_2} \frac{1}{2} \left(W(y_1 | 0) W(y_2 | 0) \sqrt{\frac{V(y_1 | 1) V(y_2 | 1)}{V(y_1 | 0) V(y_2 | 0)}} \right) \\ &\quad + \sum_{y_1 y_2} \frac{1}{2} \left(W(y_1 | 1) W(y_2 | 0) \sqrt{\frac{V(y_1 | 0) V(y_2 | 1)}{V(y_1 | 1) V(y_2 | 0)}} \right) \\ &= \sum_{y_1} W(y_1 | 0) \sqrt{\frac{V(y_1 | 1)}{V(y_1 | 0)}} \sum_{y_2} W(y_2 | 0) \sqrt{\frac{V(y_2 | 1)}{V(y_2 | 0)}} \\ &= \mathbb{E}_{W_0(y_1) W_0(y_2)} \left[\sqrt{L_V(y_1) L_V(y_2)} \right]. \\ &= Z(W, V)^2, \end{aligned}$$

where we used the symmetry property of W and V . Similarly, $Z_{2N}^{(2i)}(W, V) = Z_N^{(i)}(W, V)^2$ holds as well. As we know from [1] that $Z(V_{2N}^{(2i)}) = Z(V_N^{(i)})^2$, this time we can easily show that the difference of the Bhattacharyya parameters will preserve its sign after applying the plus transformation, i.e.,

$$Z_N^{(i)}(W, V) - Z_N^{(i)}(V) \leq 0 \Rightarrow Z_{2N}^{(2i)}(W, V) - Z_{2N}^{(2i)}(V) \leq 0.$$

In the next theorem, we explore the possible connection of such a result with Theorem 1.

Theorem 2: Let W and V be B-DMCs such that for any $N = 2^n$, $n = 1, 2, \dots$, and for any $i = 1, \dots, N$, the channels satisfy

$$Pe_N^{(i)}(W, V) - Pe_N^{(i)}(V) < 0 \quad \text{iff} \quad Z_N^{(i)}(W, V) - Z_N^{(i)}(V) < 0.$$

Then, the condition B of Theorem 1 is preserved under both polar transformations.

The theorem statement simply tells that if the Bhattacharyya upper bounds follow the same behavior as their $Pe_N^{(i)}$ counterparts; which can occur if for instance they are sufficiently tight for both the matched and mismatched error probabilities, then as long as we design the polar code for a mismatched channel V such that $Pe(W, V) \leq Pe(V)$ is satisfied, we are safe to use the code over the channel W . Although Theorem 2 provides a partial solution to the design problem, unfortunately it is non-constructive at this stage. We would need to study which channels could satisfy these type of constraints.

A. Performance Analysis over Channels that Satisfy a Certain Stochastic Dominance Order

In this part, we impose slightly stronger conditions on the channels such that the design goal is met.

Theorem 3: Let W and V be two symmetric B-DMCs which satisfy the following conditions:

- (i) $\mathbb{P}_V [L_V(y_1) \leq 1] \geq \mathbb{P}_V [L_V(y_1) \geq 1]$,
- (ii) $\mathbb{P}_W [\mathbf{1}\{L_V(y_1) \geq 1\}] \leq \mathbb{P}_V [\mathbf{1}\{L_V(y_1) \geq 1\}]$,
- (iii) $\mathbb{P}_W [\mathbf{1}\{L_V(y_1) \leq 1\}] \geq \mathbb{P}_V [\mathbf{1}\{L_V(y_1) \geq 1\}]$.

Then, for any given $N = 2^n$ with $n = 1, 2, \dots$, and any given $i = 1, \dots, N$, $Pe_N^{(i)}(W, V) \leq Z(V_N^{(i)})$ holds. Moreover, $Pe_N^{(i)}(W, V) \leq Pe_N^{(i)}(V)$ holds for $\forall i \in \mathcal{A}_N$.

Note that for B-DMCs W and V such that no initial output has a LR which equals to one, the assumptions (ii) and (iii) of Theorem 3 can be merged into a single initial condition as $Pe(W, V) \leq Pe(V)$.

Theorem 3 will be proved as a corollary to Proposition 8 and the subsequent Theorem 4.

Proposition 8: The process $\mathbb{P}_W [L_{V_N^{(i)}}(y_1^{2^N}) = 1]$ is a bounded submartingale in $[0, 1]$ which converges almost surely to the values $\{0, 1\}$.

Proof of Proposition 8: The boundedness claim is trivial. Let $L_1 = L_{V_N^{(i)}}(y_1^N)$ and $L_2 = L_{V_N^{(i)}}(y_{N+1}^{2N})$ for simplicity. We first note that

$$\begin{aligned} \mathbb{P}_W \left[\frac{L_1 + L_2}{1 + L_1 L_2} = 1 \right] &= 2\mathbb{P}_W [L = 1] - \mathbb{P}_W [L = 1]^2, \\ \mathbb{P}_W [L_1 L_2 = 1] &\geq \mathbb{P}_W [L = 1]^2 \end{aligned} \quad (14)$$

where we used the fact that $\mathbb{P}_W [L = 1] \triangleq \mathbb{P}_W [L_1 = 1] = \mathbb{P}_W [L_2 = 1]$. Therefore,

$$\mathbb{P}_W [L_1 L_2 = 1] + \mathbb{P}_W \left[\frac{L_1 + L_2}{1 + L_1 L_2} = 1 \right] \geq 2\mathbb{P}_W [L_1 = 1].$$

This inequality proves the process is a submartingale. By general results on bounded martingales, we know the process converges almost surely [1]. One can complete the proof that the convergence is to the extremes, using the relation in (14) in a similar fashion to the proof of the convergence to the extremes of the Bhattacharyya parameters' process attached to the polarization transformations carried in [1, Proposition 9] since

$$\mathbb{E}_\pm [|\mathbb{P}_W [L^\pm = 1] - \mathbb{P}_W [L = 1]|] \geq \frac{1}{2}\mathbb{P}_W [L = 1](1 - \mathbb{P}_W [L = 1])$$

holds, and when the left side of this inequality goes to zero, $\{0, 1\}$ are the only possible values $\mathbb{P}_W [L = 1]$ can take. \blacksquare

Theorem 4: Let W and V be B-DMCs such that for a given i , the following conditions hold:

A)

$$\mathbb{P}_V [L_{V_N^{(i)}}(y_1^N) \leq 1] \geq \mathbb{P}_V [L_{V_N^{(i)}}(y_1^N) \geq 1],$$

B)

$$\mathbb{P}_W [L_{V_N^{(i)}}(y_1^N) \geq 1] \leq \mathbb{P}_V [L_{V_N^{(i)}}(y_1^N) \geq 1],$$

C)

$$\mathbb{P}_W [L_{V_N^{(i)}}(y_1^N) \leq 1] \geq \mathbb{P}_V [L_{V_N^{(i)}}(y_1^N) \leq 1].$$

Then, the polar transforms preserve the above three conditions in the sense at the next level, they hold for the $2i$ -th and $2i - 1$ -th indices.

To prove Theorem 4 we need the following two propositions. Before these, we introduce a notation.

Given two B-DMCs W and V , we denote by $\mathbf{H}(L_V(y))^W \prec_{SD} \mathbf{H}(L_V(y))^V$ if the distribution of the random variable $\mathbf{H}(L_V(y))$ under the distribution $W(y|0)$ is stochastically dominated by

the distribution under $V(y|0)$. For a definition of stochastic dominance, see for instance [6, Chapter 1.2, Theorem B]. Note that by definition the condition implies

$$\mathbb{E}_W[F(\mathbf{H}(L_V(y)))] \leq \mathbb{E}_V[F(\mathbf{H}(L_V(y)))]$$

holds for any non-decreasing function $F(\cdot)$ for which the expectations exist. As an example, the cases where W and V are BSCs with crossover probabilities $\epsilon_W \leq \epsilon_V \leq 0.5$ satisfy $\mathbf{H}(L_V(y))^W \prec_{SD} \mathbf{H}(L_V(y))^V$ order.

Proposition 9: For any B-DMCs W and V , we have

$$\begin{aligned} \mathbf{1}\{L_V(y_1) \geq 1\}^W \prec_{SD} \mathbf{1}\{L_V(y_1) \geq 1\}^V \\ \text{iff } \mathbb{P}_W[\mathbf{1}\{L_V(y_1) \geq 1\}] \leq \mathbb{P}_V[\mathbf{1}\{L_V(y_1) \geq 1\}]. \end{aligned}$$

Similarly, we have

$$\begin{aligned} \mathbf{1}\{L_V(y_1) \leq 1\}^W \succ_{SD} \mathbf{1}\{L_V(y_1) \leq 1\}^V \\ \text{iff } \mathbb{P}_W[\mathbf{1}\{L_V(y_1) \leq 1\}] \geq \mathbb{P}_V[\mathbf{1}\{L_V(y_1) \leq 1\}]. \end{aligned}$$

Proof of Proposition 9: The proposition follows simply noting the random variable with the indicator function is binary valued, so the two conditions are equivalent in each case. ■

Proposition 10: The random variable $\mathbb{E}_W \left[\mathbf{1}\{L_{V_{2N}^{(2i)}}(y_1^{2N}) \geq 1\} | \mathbf{1}\{L_{V_N^{(i)}}(y_1^N) \geq 1\} \right]$ $\left(\mathbb{E}_W \left[\mathbf{1}\{L_{V_{2N}^{(2i)}}(y_1^{2N}) \leq 1\} | \mathbf{1}\{L_{V_N^{(i)}}(y_1^N) \leq 1\} \right] \right)$ is non-decreasing in $\mathbf{1}\{L_{V_N^{(i)}}(y_1^N) \geq 1\}$ $\left(\mathbf{1}\{L_{V_N^{(i)}}(y_1^N) \leq 1\} \right)$. The random variable $\mathbb{E}_W \left[\mathbf{1}\{L_{V_{2N}^{(2i-1)}}(y_1^{2N}) \geq 1\} | \mathbf{1}\{L_{V_N^{(i)}}(y_1^N) \geq 1\} \right]$ $\left(\mathbb{E}_W \left[\mathbf{1}\{L_{V_{2N}^{(2i-1)}}(y_1^{2N}) \leq 1\} | \mathbf{1}\{L_{V_N^{(i)}}(y_1^N) \leq 1\} \right] \right)$ is non-decreasing in $\mathbf{1}\{L_{V_N^{(i)}}(y_1^N) \geq 1\}$ $\left(\mathbf{1}\{L_{V_N^{(i)}}(y_1^N) \leq 1\} \right)$ if the following condition holds:

$$\mathbb{P}_W \left[\mathbf{1}\{L_{V_N^{(i)}}(y_1^N) < 1\} \right] \geq \mathbb{P}_W \left[\mathbf{1}\{L_{V_N^{(i)}}(y_1^N) > 1\} \right]. \quad (15)$$

Proof of Proposition 10: The claims for the plus operations are trivial. For the minus operation, the claims follows by noting that

$$\begin{aligned} \mathbb{E} \left[\mathbf{1}\{L_{V_{2N}^{(2i-1)}}(y_1^{2N}) \geq 1\} | \mathbf{1}\{L_{V_N^{(i)}}(y_1^N) \geq 1\} = 0 \right] \\ = \mathbb{E} \left[\mathbf{1}\{L_{V_{2N}^{(2i-1)}}(y_1^{2N}) \leq 1\} | \mathbf{1}\{L_{V_N^{(i)}}(y_1^N) \leq 1\} = 0 \right] \\ = \mathbb{P}_W \left[L_{V_N^{(i)}}(y_{N+1}^{2N}) \geq 1 \right] \end{aligned}$$

and both

$$\begin{aligned} \mathbb{E} \left[\mathbf{1}\{L_{V_{2N}^{(2i-1)}}(y_1^{2N}) \geq 1\} | \mathbf{1}\{L_{V_N^{(i)}}(y_1^N) \geq 1\} = 1 \right], \\ \mathbb{E} \left[\mathbf{1}\{L_{V_{2N}^{(2i-1)}}(y_1^{2N}) \leq 1\} | \mathbf{1}\{L_{V_N^{(i)}}(y_1^N) \leq 1\} = 1 \right] \\ \geq \mathbb{P}_W \left[L_{V_N^{(i)}}(y_{N+1}^{2N}) \leq 1 \right] \end{aligned}$$

So, the condition in (15) is sufficient to prove the monotonicity claims. \blacksquare

Proof of Theorem 4:

A $^\pm$) We know condition A is preserved by Proposition 5 for both transformations.

B $^\pm$) Using Proposition 4 with the $\mathbf{1}\{L_{V_{2N}^{(i)}}(y_1^{2N}) \geq 1\}$ random variable instead of $\mathbf{H} \left(L_{V_{2N}^{(i)}}(y_1^{2N}) \right)$, we can write

$$\begin{aligned} \mathbb{P}_W \left[\mathbf{1}\{L_{V_{2N}^{(i)}}(y_1^{2N}) \geq 1\} \right] - \mathbb{P}_V \left[\mathbf{1}\{L_{V_{2N}^{(i)}}(y_1^{2N}) \geq 1\} \right] \\ = \sum_{y_1^N} \left[W(y_1^N | 0_1^N) - V(y_1^N | 0_1^N) \right] \mathbb{E}_{W+V} \left[\mathbf{1}\{L_{V_{2N}^{(i)}}(y_1^{2N}) \geq 1\} | \mathbf{1}\{L_{V_N^{(i)}}(y_1^N) \geq 1\} \right] \end{aligned}$$

where we have defined

$$\begin{aligned} \mathbb{E}_{W+V} \left[\mathbf{1}\{L_{V_{2N}^{(i)}}(y_1^{2N}) \geq 1\} | \mathbf{1}\{L_{V_N^{(i)}}(y_1^N) \geq 1\} \right] \\ = \sum_{y_{N+1}^{2N}} \left[W(y_{N+1}^{2N} | 0_1^N) + V(y_{N+1}^{2N} | 0_1^N) \right] \mathbf{1}\{L_{V_{2N}^{(i)}}(y_1^{2N}) \geq 1\}. \quad (16) \end{aligned}$$

Moreover, we know that the condition B implies via Proposition 9 that the random variable satisfy $\mathbf{1}\{L_{V_N^{(i)}}(y_1^N) \geq 1\}^W \prec_{SD} \mathbf{1}\{L_{V_N^{(i)}}(y_1^N) \geq 1\}^V$. So, we will be done if we show that the random variable defined in (16) obtained after applying both the plus and minus transformations are non-decreasing transformations in $\mathbf{1}\{L_{V_N^{(i)}}(y_1^N) \geq 1\}$. We will consider the cases the expectations are taken under W and V separately. For $\mathbb{E}_V \left[\mathbf{1}\{L_{V_{2N}^{(i)}}(y_1^{2N}) \geq 1\} | \mathbf{1}\{L_{V_N^{(i)}}(y_1^N) \geq 1\} \right]$, we know by taking $W = V$ in Proposition 10 and by condition A this claim holds. For $\mathbb{E}_W \left[\mathbf{1}\{L_{V_{2N}^{(i)}}(y_1^{2N}) \geq 1\} | \mathbf{1}\{L_{V_N^{(i)}}(y_1^N) \geq 1\} \right]$, we know once again by Proposition 10 that this is true for the plus transformation and is also true for the minus transformation if we have

$$\mathbb{P}_W [L_V(y_1) \leq 1] \geq \mathbb{P}_W [L_V(y_1) \geq 1]. \quad (17)$$

So, now we show that (17) holds. Taking the difference of the inequalities in B and C, we get

$$\begin{aligned} \mathbb{P}_W [\mathbf{1}\{L_V(y_1) \leq 1\}] - \mathbb{P}_W [\mathbf{1}\{L_V(y_1) \geq 1\}] \\ \geq \mathbb{P}_V [\mathbf{1}\{L_V(y_1) \leq 1\}] - \mathbb{P}_V [\mathbf{1}\{L_V(y_1) \geq 1\}] \geq 0 \end{aligned}$$

where the non-negativity follows by condition A.

C[±]) The proof can be carried following similar steps as in part B[±]) showing that the transformations defined by $\mathbb{E}_{W+V} \left[\mathbf{1}\{L_{V_{2N}^{(i)}}(y_1^{2N}) \leq 1\} | \mathbf{1}\{L_{V_N^{(i)}}(y_1^N) \leq 1\} \right]$ are also non-decreasing in $\mathbf{1}\{L_{V_N^{(i)}}(y_1^N) \leq 1\}$ using Proposition 10, condition A, and Equation (17). ■

Now, we are ready to prove Theorem 3.

Proof of Theorem 3: Assume the conditions (i), (ii), and (iii) hold. Then, by Theorem 4, the conditions are preserved for the synthetic channels created by the polar transformations. Hence, for $\forall i$

$$\mathbb{P}_W \left[L_{V_N^{(i)}}(y_1) \geq 1 \right] \leq \mathbb{P}_V \left[L_{V_N^{(i)}}(y_1) \geq 1 \right].$$

Knowing the bounds $Pe_N^{(i)}(W, V) \leq \mathbb{P}_W \left[L_{V_N^{(i)}}(y_1) \geq 1 \right]$ and $\mathbb{P}_V \left[L_{V_N^{(i)}}(y_1) \geq 1 \right] \leq Z(V_N^{(i)})$ hold, the relation $Pe_N^{(i)}(W, V) \leq Z(V_N^{(i)})$ is proved.

On the other hand, Proposition 8 shows that once channels are sufficiently polarized, either $\mathbb{P}_W \left[L_{V_N^{(i)}}(y_1^{2N}) = 1 \right] \approx 1$ or $\mathbb{P}_W \left[L_{V_N^{(i)}}(y_1^{2N}) = 1 \right] \approx 0$. Moreover, one can easily find that the first case lead to a completely noisy channel, and only the second case can possibly lead to a perfect channel under mismatched decoding. In this last case, as we have

$$Pe_N^{(i)}(W, V) = \mathbb{P}_W \left[L_{V_N^{(i)}}(y_1^{2N}) \geq 1 \right],$$

it turns out that, for those indices $i \in \mathcal{A}_N$ which correspond to the good channels' picked by the polar code designed for the channel V , we expect to have

$$Pe_N^{(i)}(W, V) \leq Pe_N^{(i)}(V), \quad \forall i \in \mathcal{A}_N$$

as claimed. This completes the proof of the theorem. ■

Conclusions

We took a designer's perspective to analyze the performance of mismatched polar codes in which we tried to identify circumstances under which the polar code design using Arıkan's original construction method [1] for a given B-DMC can be used reliably for a mismatched channel. More implications of the results presented in this report will be discussed elsewhere.

REFERENCES

- [1] E. Arıkan, "Channel polarization: a method for constructing capacity-achieving codes for symmetric binary-input memoryless channels", *IEEE Trans. Inf. Theor.*, vol. 55, no. 7, pp. 3051-3073, 2009.
- [2] E. Arıkan, and E. Telatar, "On the Rate of Channel Polarization", 2008, <http://arxiv.org/abs/0807.3806v2> [cs.IT].
- [3] I. Csiszár, and P. Narayan, "Channel capacity for a given decoding metric", *IEEE Trans. Inf. Theor.*, vol. 41, no. 1, pp. 35 -43, 1995.
- [4] M. Alsan, "Performance of Polar Codes over BSCs", *International Symposium on Information Theory and its Applications (ISITA2012)*, 2012.
- [5] C. Leroux, and I. Tal, and A. Vardy, and W.J. Gross, "Hardware architectures for successive cancellation decoding of polar codes", *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2011.
- [6] R. Szekli, "Stochastic Ordering and Dependence in Applied Probability", *Lecture Notes in Statistics*, Springer-Verlag, 1995.
- [7] E. Telatar, Private Communications.