# On the Correlation Between Polarized BECs

Mani Bastani Parizi and Emre Telatar
EPFL, Lausanne, Switzerland
Email: {mani.bastaniparizi,emre.telatar}@epfl.ch

*Abstract*—We consider the $2^n$ channels synthesized by the $n$-fold application of Arıkan's polar transform to a binary erasure channel (BEC). The synthetic channels are BECs themselves, and we show that, asymptotically for almost all these channels, the pairwise correlations between their erasure events are extremely small: the correlation coefficients vanish faster than any exponential in $n$. Such a fast decay of correlations allows us to conclude that the union bound on the block error probability of polar codes is very tight.

## I. INTRODUCTION

Channel Polarization is a technique recently introduced by Arıkan [1] as a means of constructing capacity achieving codes for binary discrete memoryless channels (B-DMCs). The underlying principle of channel polarization is the following: Let $W : \mathcal{X} \longrightarrow \mathcal{Y}$ be a B-DMC with input alphabet $\mathcal{X} = \mathbb{F}_2$. From two independent copies of $W$ synthesize $W^- : \mathcal{X} \longrightarrow \mathcal{Y}^2$ and $W^+ : \mathcal{X} \longrightarrow \mathcal{Y}^2 \times \mathcal{X}$ as:

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

$$W^+(y_1, y_2, u_1 | u_2) = \frac{1}{2} W(y_1 | u_1 \oplus u_2) W(y_2 | u_2).$$

As the superscripts suggest $W^-$ turns out to be a B-DMC worse than $W$ while $W^+$ is a better B-DMC compared to $W$. This transform can be repeated $n$ times to get $N = 2^n$ B-DMCs $W_n^{(\mathbf{s})}, \mathbf{s} \in \{-, +\}^n$. Arıkan shows that (i) the transformation preserves the mutual information, (ii) $W_n^{(\mathbf{s})}$s approach to "extremal" channels, i.e., either noiseless or useless channels. In particular, the fraction of almost noiseless channels is equal to the symmetric capacity of the original B-DMC $W$. Based on these properties Arıkan constructs *polar codes* by sending uncoded data bits only on (almost) noiseless channels and arbitrary (but known to receiver) bits on the remaining channels. The channels used to transmit information are referred to as "information" channels and the rest are called "frozen" channels. A successive cancellation decoder has been proposed by Arıkan to decode the information bits with complexity $O(N \log N)$ and shown to have a block error probability that behaves roughly as $O\left(2^{-\sqrt{N}}\right)$ (cf. [2]).

The set of Binary Erasure Channels (BECs) is stable under Polarization in the sense that if $W$ is a BEC, then $W^+$ and $W^-$ are also BECs. We denote a BEC with erasure probability $\epsilon$ as $\mathrm{BEC}(\epsilon)$. One can establish a one-to-one relationship between a $\mathrm{BEC}(\epsilon)$ and an "erasure indicator" random variable $E$ such that $E \in \{0, 1\}$ and $\mathbb{P}[E = 1] = \epsilon$. The polar transform of a BEC is hence equivalent to taking two independent copies of $E$ and creating the erasure indicators of $W^-$ and $W^+$.

**Lemma 1** (Polar Transform of BEC [1, Proposition 6]). *If $W$ is a BEC with erasure probability $\epsilon$, applying the polar transform $(W, W) \mapsto (W^-, W^+)$ produces two BECs $W^+$ with erasure probability $\epsilon^2$ and $W^-$ with erasure probability $2\epsilon - \epsilon^2$. Moreover, $W^-$ erases iff either copy of $W$ erases, and $W^+$ erases iff both copies of $W$ erase.*

**Corollary 1.** *The erasure indicators of $W^-$ and $W^+$, denoted by $E^-$ and $E^+$, are constructed from two independent copies of $E$, denoted by $E$ and $E'$, as:*

$$E^- = \max\{E, E'\} = E + E' - EE' \tag{1a}$$

$$E^+ = \min\{E, E'\} = EE'. \tag{1b}$$

While two copies of $E$ are independent (and hence uncorrelated), $E^+$ and $E^-$ are correlated: $E^+ = 1$ implies $E^- = 1$. On the other side, by polarization $W_n^{(\mathbf{s})}$s (and equivalently $E_n^{(\mathbf{s})}$s) become deterministic as $n \to \infty$. Hence it looks like $E_n^{(\mathbf{s})}$ and $E_n^{(\mathbf{t})}$ would become uncorrelated for $\mathbf{s} \neq \mathbf{t}$, where $\mathbf{s}$ and $\mathbf{t}$ are sign sequences of length $n$ used for indexing the channels. In particular it is easy to see that $\mathbb{E}\left[E_n^{(\mathbf{s})} E_n^{(\mathbf{t})}\right] - \mathbb{E}\left[E_n^{(\mathbf{s})}\right]\mathbb{E}\left[E_n^{(\mathbf{t})}\right]$ is small for almost every $\mathbf{s}, \mathbf{t}$.

In this paper we provide upper bounds on correlation *coefficients* defined as:

$$\rho_n^{(\mathbf{s}, \mathbf{t})} \triangleq \frac{\mathbb{E}\left[E_n^{(\mathbf{s})} E_n^{(\mathbf{t})}\right] - \mathbb{E}\left[E_n^{(\mathbf{s})}\right]\mathbb{E}\left[E_n^{(\mathbf{t})}\right]}{\sqrt{\mathrm{var}\left[E_n^{(\mathbf{s})}\right] \mathrm{var}\left[E_n^{(\mathbf{t})}\right]}} \tag{2}$$

and exploit these bounds and the inclusion–exclusion principle to find lower bounds on the block error probability of polar codes. In particular, our bounds are strong enough to show that the sum of the Bhattacharyya parameters of the information channels is a tight estimate of the block error probability.

## II. NOTATION

Throughout this manuscript, we use uppercase letters (like $X$) to indicate a random variable, and its lowercase version ($x$) for a realization of that random variable. The boldface letters denote matrices, vectors or sequences which will be clear from the context.

We denote the sets by script-style uppercase letters like $\mathcal{S}$ and by $|\mathcal{S}|$ we mean the cardinality of $\mathcal{S}$.

We use the *bar* notation defined as $\overline{x} \triangleq 1 - x$ for the sake of brevity and refer to $\overline{x}$ as the "complement" of $x$.

For sign sequences $\mathbf{s} \in \{-, +\}^*$ and $\mathbf{t} \in \{-, +\}^*$, $\mathrm{CP}[\mathbf{s}, \mathbf{t}]$ denotes their common prefix. Furthermore, let $|\mathbf{s}|$ denote the length of a sequence $\mathbf{s}$.

## III. PROPERTIES OF CORRELATION COEFFICIENTS

As we mentioned in Section I, we are interested in analyzing the matrix of correlation coefficients of the erasure indicator vector $\mathbf{E}_n = \left[E_n^{(\mathbf{s})}r : \mathbf{s} \in \{-,+\}^n\right]$. It is more convenient to index the $N = 2^n$ elements of that vector using sign sequences $\mathbf{s} \in \{-,+\}^n$ instead of mapping the sign sequences to integers and using the natural indexing. We will use the same indexing for the $N^2$ elements of the correlation coefficients matrix.

Arıkan has already shown that the vector $\mathbf{Z}_n = \mathbb{E}[\mathbf{E}_n]$ can be computed via a single-step recursion. More precisely, having $\mathbf{Z}_{n-1}$ we can compute the elements of $\mathbf{Z}_n$ as:

$$Z_n^{(\mathbf{s}-)} = 2Z_{n-1}^{(\mathbf{s})} - \left(Z_{n-1}^{(\mathbf{s})}\right)^2 \tag{3a}$$

$$Z_n^{(\mathbf{s}+)} = \left(Z_{n-1}^{(\mathbf{s})}\right)^2 \tag{3b}$$

for $\forall \mathbf{s} \in \{-,+\}^{n-1}$ with $Z_0 = \epsilon$. Note that (3a) and (3b) can also be derived by taking the expectation from both sides of (1a) and (1b) and using the independence between $E$ and $E'$.

Interestingly, the correlation coefficients matrix $\rho_n = \left[\rho_n^{(\mathbf{s},\mathbf{t})} : \mathbf{s},\mathbf{t} \in \{-,+\}^n\right]$ can also be computed via a single-step recursion as we see in this section.

It is useful to rewrite (1a) and (1b) as

$$E^- = \overline{\overline{E} \times \overline{E'}} \tag{4a}$$

$$E^+ = E \times E' \tag{4b}$$

and subsequently (3a) and (3b) as:

$$\overline{Z_n^{(\mathbf{s}-)}} = \overline{Z_{n-1}^{(\mathbf{s})}}^2 \tag{5a}$$

$$Z_n^{(\mathbf{s}+)} = Z_{n-1}^{(\mathbf{s})}{}^2 \tag{5b}$$

to see the symmetry between 'minus' and 'plus' transforms.

Recall that the "covariance" of random variables $X$ and $Y$ is defined as:

$$\mathrm{cov}[X,Y] \triangleq \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]. \tag{6}$$

**Lemma 2.** *Let $X$ and $Y$ be arbitrary random variables and set $U \triangleq \overline{X}$ and $V \triangleq \overline{Y}$. Then:*

$$\mathrm{var}[U] = \mathrm{var}[X]. \tag{7}$$

*Moreover,*

$$\mathrm{cov}[U,V] = \mathrm{cov}[X,Y] \tag{8a}$$

$$\mathrm{cov}[X,V] = \mathrm{cov}[U,Y] = -\mathrm{cov}[X,Y] \tag{8b}$$

*Proof:* It is clear that $\mathbb{E}[U] = 1 - \mathbb{E}[X]$ and $\mathbb{E}[V] = 1 - \mathbb{E}[Y]$. (7) is also trivial since $\mathrm{var}[aX + b] = |a|^2 \mathrm{var}[X]$ for any constants $a$ and $b$. Furthermore:

$$\mathbb{E}[UV] = \mathbb{E}[(1-X)(1-Y)] = 1 - \mathbb{E}[X] - \mathbb{E}[Y] + \mathbb{E}[XY]$$

hence

$$\mathrm{cov}[U,V] = \mathbb{E}[UV] - \mathbb{E}[U]\mathbb{E}[V]$$
$$= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = \mathrm{cov}[X,Y]$$

which proves (8a). Likewise,

$$\mathbb{E}[UY] = \mathbb{E}[(1-X)Y] = \mathbb{E}[Y] - \mathbb{E}[XY]$$

which shows $\mathrm{cov}[U,Y] = \mathbb{E}[UY] - \mathbb{E}[U]\mathbb{E}[Y] = -\mathbb{E}[XY] + \mathbb{E}[X]\mathbb{E}[Y] = -\mathrm{cov}[X,Y]$. The same argument applies to $\mathrm{cov}[X,V]$ which proves (8b). ∎

**Corollary 2.** *Let $X,Y,U$ and $V$ be defined as in Lemma 2 and $\rho[X,Y] \triangleq \frac{\mathrm{cov}[X,Y]}{\sqrt{\mathrm{var}[X]\mathrm{var}[Y]}}$ denote the correlation coefficient between random variables $X$ and $Y$, then:*

$$\rho[U,V] = \rho[X,Y] \tag{9a}$$

$$\rho[X,V] = \rho[U,Y] = -\rho[X,Y] \tag{9b}$$

**Lemma 3.** *The covariance matrix of the random vector $\mathbf{E}_n$, $\mathbf{C}_n \triangleq \left[C_n^{(\mathbf{s},\mathbf{t})} : \mathbf{s},\mathbf{t} \in \{-,+\}^n\right]$ where*

$$C_n^{(\mathbf{s},\mathbf{t})} \triangleq \mathrm{cov}\left[E_n^{(\mathbf{s})}, E_n^{(\mathbf{t})}\right],$$

*can be computed in terms of $\mathbf{C}_{n-1}$ and $\mathbf{Z}_{n-1}$ as follows:*

$$C_n^{(\mathbf{s}-,\mathbf{t}-)} = 2\overline{Z_{n-1}^{(\mathbf{s})}}\,\overline{Z_{n-1}^{(\mathbf{t})}}C_{n-1}^{(\mathbf{s},\mathbf{t})} + C_{n-1}^{(\mathbf{s},\mathbf{t})^2}, \tag{10a}$$

$$C_n^{(\mathbf{s}-,\mathbf{t}+)} = 2\overline{Z_{n-1}^{(\mathbf{s})}}Z_{n-1}^{(\mathbf{t})}C_{n-1}^{(\mathbf{s},\mathbf{t})} - C_{n-1}^{(\mathbf{s},\mathbf{t})^2}, \tag{10b}$$

$$C_n^{(\mathbf{s}+,\mathbf{t}-)} = 2Z_{n-1}^{(\mathbf{s})}\overline{Z_{n-1}^{(\mathbf{t})}}C_{n-1}^{(\mathbf{s},\mathbf{t})} - C_{n-1}^{(\mathbf{s},\mathbf{t})^2}, \tag{10c}$$

$$C_n^{(\mathbf{s}+,\mathbf{t}+)} = 2Z_{n-1}^{(\mathbf{s})}Z_{n-1}^{(\mathbf{t})}C_{n-1}^{(\mathbf{s},\mathbf{t})} + C_{n-1}^{(\mathbf{s},\mathbf{t})^2}. \tag{10d}$$

*It is clear that $C_0 = \epsilon\bar{\epsilon}$ where $\epsilon$ is the erasure probability of the underlying BEC.*

*Proof:* We first prove (10d) and then show how the rest of results easily follow using Lemma 2.

Recall that $E_n^{(\mathbf{s}+)} = E_{n-1}^{(\mathbf{s})} \times E_{n-1}^{(\mathbf{s})\,\prime}$ and $E_n^{(\mathbf{t}+)} = E_{n-1}^{(\mathbf{t})} \times E_{n-1}^{(\mathbf{t})\,\prime}$. Furthermore, $\mathbb{E}\left[E_{n-1}^{(\mathbf{s})}\right] = Z_{n-1}^{(\mathbf{s})}$ and $\mathbb{E}\left[E_{n-1}^{(\mathbf{t})}\right] = Z_{n-1}^{(\mathbf{t})}$:

$$\mathrm{cov}\left[E_n^{(\mathbf{s}+)}, E_n^{(\mathbf{t}+)}\right] = \mathbb{E}\left[E_{n-1}^{(\mathbf{s})}E_{n-1}^{(\mathbf{s})\,\prime}E_{n-1}^{(\mathbf{t})}E_{n-1}^{(\mathbf{t})\,\prime}\right]$$
$$- \mathbb{E}\left[E_{n-1}^{(\mathbf{s})}E_{n-1}^{(\mathbf{s})\,\prime}\right]\mathbb{E}\left[E_{n-1}^{(\mathbf{t})}E_{n-1}^{(\mathbf{t})\,\prime}\right]$$
$$\overset{(*)}{=} \mathbb{E}\left[E_{n-1}^{(\mathbf{s})}E_{n-1}^{(\mathbf{t})}\right]^2 - \mathbb{E}\left[E_{n-1}^{(\mathbf{s})}\right]^2\mathbb{E}\left[E_{n-1}^{(\mathbf{t})}\right]^2$$
$$= \left(\mathbb{E}\left[E_{n-1}^{(\mathbf{s})}E_{n-1}^{(\mathbf{t})}\right] - Z_{n-1}^{(\mathbf{s})}Z_{n-1}^{(\mathbf{t})}\right)^2$$
$$+ 2Z_{n-1}^{(\mathbf{s})}Z_{n-1}^{(\mathbf{t})}\left(\mathbb{E}\left[E_{n-1}^{(\mathbf{s})}E_{n-1}^{(\mathbf{t})}\right] - Z_{n-1}^{(\mathbf{s})}Z_{n-1}^{(\mathbf{t})}\right)$$
$$= C_{n-1}^{(\mathbf{s},\mathbf{t})^2} + 2Z_{n-1}^{(\mathbf{s})}Z_{n-1}^{(\mathbf{t})}C_{n-1}^{(\mathbf{s},\mathbf{t})}.$$

Note that in (*) we have used the independence between the indicator variables with *prime* and the ones without that and the fact that they are both identical copies of the same random variable.

Now observe that $E_n^{(\mathbf{s}-)} = \overline{\overline{E_{n-1}^{(\mathbf{s})}} \times \overline{E_{n-1}^{(\mathbf{s})}}^{\,\prime}}$ and $E_n^{(\mathbf{t}-)} = \overline{\overline{E_{n-1}^{(\mathbf{t})}} \times \overline{E_{n-1}^{(\mathbf{t})}}^{\,\prime}}$.

To compute $C_n^{(\mathbf{s}-,\mathbf{t}-)}$, using (8a) we have:

$$\text{cov}\big[E_n^{(\mathbf{s}-)}, E_n^{(\mathbf{t}-)}\big] = \text{cov}\left[\overline{\overline{E_{n-1}^{(\mathbf{s})}} \times \overline{E_{n-1}^{(\mathbf{s})}}'}, \overline{\overline{E_{n-1}^{(\mathbf{t})}} \times \overline{E_{n-1}^{(\mathbf{t})}}'}\right]$$

$$= \text{cov}\left[\overline{E_{n-1}^{(\mathbf{s})}} \times \overline{E_{n-1}^{(\mathbf{s})}}', \overline{E_{n-1}^{(\mathbf{t})}} \times \overline{E_{n-1}^{(\mathbf{t})}}'\right]$$

$$\stackrel{(*)}{=} C_{n-1}^{(\mathbf{s},\mathbf{t})2} + 2\overline{Z_{n-1}^{(\mathbf{s})}} \overline{Z_{n-1}^{(\mathbf{t})}} C_{n-1}^{(\mathbf{s},\mathbf{t})}$$

where (*) follows by observing that we are essentially computing the same covariance as the one we just computed to show (10d) considering the facts that (i) $\text{cov}\left[\overline{E_{n-1}^{(\mathbf{s})}}, \overline{E_{n-1}^{(\mathbf{t})}}\right] = \text{cov}\big[E_{n-1}^{(\mathbf{s})}, E_{n-1}^{(\mathbf{t})}\big]$ (using (8a) once again) and (ii) $\mathbb{E}\left[\overline{E_{n-1}^{(\mathbf{s})}}\right] = \overline{Z_{n-1}^{(\mathbf{s})}}$ and $\mathbb{E}\left[\overline{E_{n-1}^{(\mathbf{t})}}\right] = \overline{Z_{n-1}^{(\mathbf{t})}}$.

Likewise (10b) (similarly (10c)) follows using (8b):

$$\text{cov}\big[E_n^{(\mathbf{s}-)}, E_n^{(\mathbf{t}+)}\big] = \text{cov}\left[\overline{\overline{E_{n-1}^{(\mathbf{s})}} \times \overline{E_{n-1}^{(\mathbf{s})}}'}, E_{n-1}^{(\mathbf{t})} \times E_{n-1}^{(\mathbf{t})}{}'\right]$$

$$= -\text{cov}\left[\overline{E_{n-1}^{(\mathbf{s})}} \times \overline{E_{n-1}^{(\mathbf{s})}}', E_{n-1}^{(\mathbf{t})} \times E_{n-1}^{(\mathbf{t})}{}'\right]$$

$$\stackrel{(*)}{=} -\big(C_{n-1}^{(\mathbf{s},\mathbf{t})2} - 2\overline{Z_{n-1}^{(\mathbf{s})}} Z_{n-1}^{(\mathbf{t})} C_{n-1}^{(\mathbf{s},\mathbf{t})}\big).$$

Once again in (*) we are computing the same form of covariance as the one we did to show (10d) considering the fact that $\text{cov}\left[\overline{E_{n-1}^{(\mathbf{s})}}, E_{n-1}^{(\mathbf{t})}\right] = -\text{cov}\big[E_{n-1}^{(\mathbf{s})}, E_{n-1}^{(\mathbf{t})}\big] = -C_{n-1}^{(\mathbf{s},\mathbf{t})}$ (by (8b)). ∎

**Corollary 3.** *The correlation coefficients matrix of the random vector* $\mathbf{E}_n$, *defined as* $\rho_n \triangleq \left[\rho_n^{(\mathbf{s},\mathbf{t})}\right]$ *(where* $\rho_n^{(\mathbf{s},\mathbf{t})}$ *is defined in (2)) can be computed in terms of* $\rho_{n-1}$ *and* $\mathbf{Z}_{n-1}$ *as:*

$$\rho_n^{(\mathbf{s}-,\mathbf{t}-)} = 2\sqrt{\frac{\overline{Z_{n-1}^{(\mathbf{s})}}}{1+\overline{Z_{n-1}^{(\mathbf{s})}}}}\sqrt{\frac{\overline{Z_{n-1}^{(\mathbf{t})}}}{1+\overline{Z_{n-1}^{(\mathbf{t})}}}}\rho_{n-1}^{(\mathbf{s},\mathbf{t})}$$
$$+ \sqrt{\frac{Z_{n-1}^{(\mathbf{s})}}{1+\overline{Z_{n-1}^{(\mathbf{s})}}}}\sqrt{\frac{Z_{n-1}^{(\mathbf{t})}}{1+\overline{Z_{n-1}^{(\mathbf{t})}}}}\rho_{n-1}^{(\mathbf{s},\mathbf{t})2} \quad (11a)$$

$$\rho_n^{(\mathbf{s}-,\mathbf{t}+)} = 2\sqrt{\frac{\overline{Z_{n-1}^{(\mathbf{s})}}}{1+\overline{Z_{n-1}^{(\mathbf{s})}}}}\sqrt{\frac{Z_{n-1}^{(\mathbf{t})}}{1+Z_{n-1}^{(\mathbf{t})}}}\rho_{n-1}^{(\mathbf{s},\mathbf{t})}$$
$$- \sqrt{\frac{Z_{n-1}^{(\mathbf{s})}}{1+\overline{Z_{n-1}^{(\mathbf{s})}}}}\sqrt{\frac{\overline{Z_{n-1}^{(\mathbf{t})}}}{1+Z_{n-1}^{(\mathbf{t})}}}\rho_{n-1}^{(\mathbf{s},\mathbf{t})2} \quad (11b)$$

$$\rho_n^{(\mathbf{s}+,\mathbf{t}-)} = 2\sqrt{\frac{Z_{n-1}^{(\mathbf{s})}}{1+Z_{n-1}^{(\mathbf{s})}}}\sqrt{\frac{\overline{Z_{n-1}^{(\mathbf{t})}}}{1+\overline{Z_{n-1}^{(\mathbf{t})}}}}\rho_{n-1}^{(\mathbf{s},\mathbf{t})}$$
$$- \sqrt{\frac{\overline{Z_{n-1}^{(\mathbf{s})}}}{1+Z_{n-1}^{(\mathbf{s})}}}\sqrt{\frac{Z_{n-1}^{(\mathbf{t})}}{1+\overline{Z_{n-1}^{(\mathbf{t})}}}}\rho_{n-1}^{(\mathbf{s},\mathbf{t})2} \quad (11c)$$

$$\rho_n^{(\mathbf{s}+,\mathbf{t}+)} = 2\sqrt{\frac{Z_{n-1}^{(\mathbf{s})}}{1+Z_{n-1}^{(\mathbf{s})}}}\sqrt{\frac{Z_{n-1}^{(\mathbf{t})}}{1+Z_{n-1}^{(\mathbf{t})}}}\rho_{n-1}^{(\mathbf{s},\mathbf{t})}$$

$$+ \sqrt{\frac{\overline{Z_{n-1}^{(\mathbf{s})}}}{1+Z_{n-1}^{(\mathbf{s})}}}\sqrt{\frac{\overline{Z_{n-1}^{(\mathbf{t})}}}{1+Z_{n-1}^{(\mathbf{t})}}}\rho_{n-1}^{(\mathbf{s},\mathbf{t})2} \quad (11d)$$

*Clearly* $\rho_0 = 1$.

*Proof:* Once again we only prove (11d) and the rest follow by the symmetry using Corollary 2. Since $E_n^{(\mathbf{s})}$'s are $\{0,1\}$ valued RVs with mean $Z_n^{(\mathbf{s})}$:

$$\text{var}[E_n^{(\mathbf{s})}] = Z_n^{(\mathbf{s})}\overline{Z_n^{(\mathbf{s})}}. \quad (12)$$

Setting $C_n^{(\mathbf{s},\mathbf{t})} = \rho_n^{(\mathbf{s},\mathbf{t})}\sqrt{Z_n^{(\mathbf{s})}\overline{Z_n^{(\mathbf{s})}}Z_n^{(\mathbf{t})}\overline{Z_n^{(\mathbf{t})}}}$ in both sides of (10d) and using the fact that $Z_n^{(\mathbf{s}+)} = Z_{n-1}^{(\mathbf{s})2}$ (similarly $Z_n^{(\mathbf{t}+)} = Z_{n-1}^{(\mathbf{t})2}$) we get:

$$\rho_n^{(\mathbf{s}+,\mathbf{t}+)}\sqrt{Z_{n-1}^{(\mathbf{s})2}\overline{\left(Z_{n-1}^{(\mathbf{s})2}\right)}Z_{n-1}^{(\mathbf{t})2}\overline{\left(Z_{n-1}^{(\mathbf{t})2}\right)}} =$$
$$2Z_{n-1}^{(\mathbf{s})}Z_{n-1}^{(\mathbf{t})}\sqrt{Z_{n-1}^{(\mathbf{s})}\overline{Z_{n-1}^{(\mathbf{s})}}Z_{n-1}^{(\mathbf{t})}\overline{Z_{n-1}^{(\mathbf{t})}}}\rho_{n-1}^{(\mathbf{s},\mathbf{t})}$$
$$+ \big(Z_{n-1}^{(\mathbf{s})}\overline{Z_{n-1}^{(\mathbf{s})}}Z_{n-1}^{(\mathbf{t})}\overline{Z_{n-1}^{(\mathbf{t})}}\big)\rho_{n-1}^{(\mathbf{s},\mathbf{t})2}$$

Eliminating $Z_{n-1}^{(\mathbf{s})}Z_{n-1}^{(\mathbf{t})}$ from both sides and observing that $\sqrt{\frac{x\overline{x}}{1-x^2}} = \sqrt{\frac{x}{1+x}}$ and $\frac{\overline{x}}{\sqrt{1-x^2}} = \sqrt{\frac{\overline{x}}{1+x}}$ proves the claim. ∎

The property of being computable by a single-step recursion generalizes to higher order statistics:

**Lemma 4.** *In general the $m$-th order moments of the random variables* $E_n^{(\mathbf{s}^n)}, \mathbf{s}^n \in \{-,+\}^n$ *can be computed from the $m$-th order moments of random variables* $E_{n-1}^{(\mathbf{s}^{n-1})}, \mathbf{s}^{n-1} \in \{-,+\}^{n-1}$.

*Proof:* By the $m$-th order moment we mean:

$$\mathbb{E}\left[E_n^{(\mathbf{s}_1^n)} E_n^{(\mathbf{s}_2^n)} \cdots E_n^{(\mathbf{s}_m^n)}\right]$$

for some set of indices $\mathbf{s}_1^n, \mathbf{s}_2^n, \cdots, \mathbf{s}_m^n$ which are *not* necessarily distinct.

Let $\mathbf{s}^{n-1}$ denote the subsequence of $\mathbf{s}^n$ including its first $n-1$ elements and observe that for any $k \in \{1, 2, \ldots, m\}$, $E_n^{(\mathbf{s}_k^n)}$ is linear in each of $E_{n-1}^{(\mathbf{s}_k^{n-1})}$ and $E_{n-1}^{(\mathbf{s}_k^{n-1})'}$ (cf. (1a) and (1b)). This means in the expansion of $E_n^{(\mathbf{s}_1^n)} E_n^{(\mathbf{s}_2^n)} \cdots E_n^{(\mathbf{s}_m^n)}$ we will have the terms in the form of $E_{n-1}^{(\mathbf{s}_1^{n-1})} E_{n-1}^{(\mathbf{s}_2^{n-1})} \cdots E_{n-1}^{(\mathbf{s}_l^{n-1})} \times E_{n-1}^{(\mathbf{s}_1^{n-1'})'} E_{n-1}^{(\mathbf{s}_2^{n-1'})'} \cdots E_{n-1}^{(\mathbf{s}_{l'}^{n-1'})'}$ for some $l \le m$ and $l' \le m$.

The independence of the variables with prime and the one without prime implies that the expectation of such product will be product of two expectations each of which is at most an $m$-th order moment of the random variables $E_{n-1}^{(\mathbf{s}^{n-1})}$. ∎

One can derive the properties stated in the sequel on $\rho_n^{(\mathbf{s},\mathbf{t})}$ according to the aforementioned recursions:

**Property 1.**

$$0 \le \rho_n^{(\mathbf{s},\mathbf{t})} \le \min\left\{\sqrt{\frac{\overline{Z_n^{(\mathbf{s})}Z_n^{(\mathbf{t})}}}{Z_n^{(\mathbf{s})}\overline{Z_n^{(\mathbf{t})}}}}, \sqrt{\frac{Z_n^{(\mathbf{s})}\overline{Z_n^{(\mathbf{t})}}}{\overline{Z_n^{(\mathbf{s})}}Z_n^{(\mathbf{t})}}}\right\} \quad (13)$$

Property 1 follows as a corollary of the following property on $C_n^{(\mathbf{s},\mathbf{t})}$:

**Property.**

$$0 \leq C_n^{(\mathbf{s},\mathbf{t})} \leq \min\left\{\overline{Z_n^{(\mathbf{s})}}Z_n^{(\mathbf{t})}, Z_n^{(\mathbf{s})}\overline{Z_n^{(\mathbf{t})}}\right\} \qquad (14)$$

*Proof:* We prove the claim by induction on $n$. The claim is trivially true for $n = 0$ since:

$$0 \leq C_0 = \mathrm{var}[E_0] = \epsilon\bar{\epsilon} \leq \min\{\bar{\epsilon}\epsilon, \epsilon\bar{\epsilon}\}$$

where $\epsilon$ is the erasure probability of the underlying BEC.

Now, assuming (14) holds for $n - 1$, we shall show:

$$0 \leq C_n^{(\mathbf{s}-,\mathbf{t}-)} \leq \min\left\{\overline{Z_n^{(\mathbf{s}-)}}Z_n^{(\mathbf{t}-)}, Z_n^{(\mathbf{s}-)}\overline{Z_n^{(\mathbf{t}-)}}\right\}. \qquad (15a)$$

$$0 \leq C_n^{(\mathbf{s}-,\mathbf{t}+)} \leq \min\left\{\overline{Z_n^{(\mathbf{s}-)}}Z_n^{(\mathbf{t}+)}, Z_n^{(\mathbf{s}-)}\overline{Z_n^{(\mathbf{t}+)}}\right\}. \qquad (15b)$$

$$0 \leq C_n^{(\mathbf{s}+,\mathbf{t}-)} \leq \min\left\{\overline{Z_n^{(\mathbf{s}+)}}Z_n^{(\mathbf{t}-)}, Z_n^{(\mathbf{s}+)}\overline{Z_n^{(\mathbf{t}-)}}\right\}. \qquad (15c)$$

$$0 \leq C_n^{(\mathbf{s}+,\mathbf{t}+)} \leq \min\left\{\overline{Z_n^{(\mathbf{s}+)}}Z_n^{(\mathbf{t}+)}, Z_n^{(\mathbf{s}+)}\overline{Z_n^{(\mathbf{t}+)}}\right\}. \qquad (15d)$$

As (10d) is obtained by replacing both $Z_n^{(\mathbf{t})}$ and $Z_n^{(\mathbf{s})}$ by their complements and (10c) is obtained by swapping $\mathbf{s}$ and $\mathbf{t}$ in (10b) we only need to prove (15a) and (15b) and the rest follow by symmetry. Furthermore, positivity of $C_n^{(\mathbf{s}-,\mathbf{t}-)}$ and $C_n^{(\mathbf{s}-,\mathbf{t}+)}$ is clear by the assumption (14) (for $n-1$) and the combination formulae (10a) and (10b). So, we only verify the upper-bounds.

Let $a \triangleq Z_{n-1}^{(\mathbf{s})}$, $b \triangleq Z_{n-1}^{(\mathbf{t})}$ and $c \triangleq C_{n-1}^{(\mathbf{s},\mathbf{t})}$, for the sake of brevity. Note that by definition $0 \leq a \leq 1$ and $0 \leq b \leq 1$. However, if either $a$ or $b$ is extremal, by assumption (14), $c = 0$ and the claim is trivial. So, for the rest of the proof, we safely assume $0 < a < 1$ and $0 < b < 1$.

- To prove (15a) we have to show:

$$2a\bar{b}c + c^2 \leq \min\{\bar{a}^2(2b - b^2), (2a - a^2)\bar{b}^2\}.$$

The above inequality is symmetric in $a$ and $b$ hence without loss of generality we can assume $a \geq b$ which implies $\bar{b}a \geq \bar{a}b$ and also $(2a - a^2)\bar{b}^2 \geq \bar{a}^2(2b - b^2)$. The LHS of the above inequality is increasing in $c$, hence once we verify the inequality for maximum possible value of $c$ we are done. Replacing $c$ with $\bar{a}b$ we get:

$$2a\bar{b}\bar{a}b + (\bar{a}b)^2 \leq (\bar{a})^2(2b - b^2).$$

Simplifying $\bar{a}^2$ from both sides yields $2b - b^2 \leq 2b - b^2$.

- To prove (15b) we need to show:

$$2\bar{a}bc - c^2 \leq \min\{\bar{a}^2 b^2, (2a - a^2)(1 - b^2)\}$$
$$= \min\{(\bar{a}b)^2, a\bar{b}(1 + \bar{a})(1 + b)\}$$

As $c \leq \bar{a}b$ the LHS is an increasing function of $c$ and we only need to verify the inequality for maximum possible value of $c$.

  - If $\bar{a}b \leq a\bar{b}$, the LHS of the inequality will be $(\bar{a}b)^2$ at $c = \bar{a}b$ and:

$$(\bar{a}b)^2 \leq (\bar{a}b) \times (a\bar{b}) < [(1 + \bar{a})(1 + b)] \times [a\bar{b}]$$

- If $a\bar{b} \leq \bar{a}b$, then the LHS of our inequality at $c = a\bar{b}$ will be equal to:

$$2\bar{a}b \times a\bar{b} - (a\bar{b})^2 = a\bar{b}[2\bar{a}b - a\bar{b}]$$
$$= a\bar{b}[\bar{a}b + \bar{a} + b - 1]$$
$$= a\bar{b}[(1 + \bar{a})(1 + b) - 2]$$
$$\leq a\bar{b}(1 + \bar{a})(1 + b)$$

Furthermore, as the LHS is increasing in $c$, at $c = a\bar{b}$ it will be less than $(\bar{a}b)^2$ (its value at $c = \bar{a}b$). ∎

*Remark.* This upper-bound shows for almost all choices of $\mathbf{s}$ and $\mathbf{t}$, $C_n^{(\mathbf{s},\mathbf{t})} = \mathbb{E}\big[E_n^{(\mathbf{s})}E_n^{(\mathbf{t})}\big] - \mathbb{E}\big[E_n^{(\mathbf{s})}\big]\mathbb{E}\big[E_n^{(\mathbf{t})}\big]$ goes to zero as $n$ gets large.

**Property 2.** *For* $\mathbf{s}, \mathbf{t} \in \{-,+\}^{n-1}$ *and* $s_n, t_n \in \{-,+\}$

$$\rho_n^{(\mathbf{s}s_n, \mathbf{t}t_n)} \leq \rho_{n-1}^{(\mathbf{s},\mathbf{t})}$$

*with equality iff*
 (i) $\rho_{n-1}^{(\mathbf{s},\mathbf{t})} = 0$, *or*
 (ii) $s_n = t_n$ *and* $\rho_{n-1}^{(\mathbf{s},\mathbf{t})} = 1$ *and* $Z_{n-1}^{(\mathbf{s})} = Z_{n-1}^{(\mathbf{t})}$, *or*
 (iii) $Z_{n-1}^{(\mathbf{s})} = b_{s_n}$ *and* $Z_{n-1}^{(\mathbf{t})} = b_{t_n}$, *where* $b_+ = 1$ *and* $b_- = 0$.

*Proof:* The case of $\rho_{n-1}^{(\mathbf{s},\mathbf{t})} = 0$ is trivial. Otherwise, we consider the ratio $\rho_n^{(\mathbf{s}s_n, \mathbf{t}t_n)}/\rho_{n-1}^{(\mathbf{s},\mathbf{t})}$. Using (11a) to (11d) this ratio is as shown in (16). Let $a \triangleq Z_{n-1}^{(\mathbf{s})}$, $b \triangleq Z_{n-1}^{(\mathbf{t})}$ and $r \triangleq \rho_{n-1}^{(\mathbf{s},\mathbf{t})}$ and observe that:
1) If $(s_n, t_n) = (+, +)$, applying the Cauchy-Schwarz inequality to the RHS of (16) we get:

$$\left(\sqrt{\frac{2a}{1+a}}\sqrt{\frac{2b}{1+b}} + \sqrt{\frac{\bar{a}r}{1+a}}\sqrt{\frac{\bar{b}r}{1+b}}\right)$$
$$\leq \sqrt{\frac{2a + r\bar{a}}{1+a}}\sqrt{\frac{2b + r\bar{b}}{1+b}}$$

For $a \in [0, 1]$, $b \in [0, 1]$ and $r \in [0, 1]$, each of the square-roots are strictly smaller than 1 unless $r = 1$ [1] or $a = b = 1$. Furthermore, the equality conditions for Cauchy-Schwarz inequality imply $\sqrt{a/\bar{a}} = \sqrt{b/\bar{b}}$ which in turn implies $a = b$. Therefore, we can conclude that if $(s_n, t_n) = (+, +)$, $\rho_n^{(\mathbf{s}s_n, \mathbf{t}t_n)}/\rho_{n-1}^{(\mathbf{s},\mathbf{t})} \leq 1$ with equality iff $(Z_{n-1}^{(\mathbf{s})} = Z_{n-1}^{(\mathbf{t})}$ and $\rho_{n-1}^{(\mathbf{s},\mathbf{t})} = 1)$ or $(Z_{n-1}^{(\mathbf{s})} = Z_{n-1}^{(\mathbf{t})} = 1)$. The same argument can also be applied to the case of $(s_n, t_n) = (-, -)$.
2) If $(s_n, t_n) = (+, -)$, the RHS of (16) can be bounded as:

$$2\sqrt{\frac{a}{1+a}}\sqrt{\frac{\bar{b}}{1+\bar{b}}} - \sqrt{\frac{\bar{a}}{1+a}}\sqrt{\frac{b}{1+\bar{b}}}r$$
$$\leq 2\sqrt{\frac{a}{1+a}}\sqrt{\frac{\bar{b}}{1+\bar{b}}} \leq 1.$$

---

[1] As each of them is in the form of $\sqrt{\frac{1+x+(r-1)\bar{x}}{1+x}}$ which is smaller than one since the numerator is less than the denominator.

$$\frac{\rho_n^{(\mathbf{s}s_n,\mathbf{t}t_n)}}{\rho_{n-1}^{(\mathbf{s},\mathbf{t})}} = \begin{cases} 2\sqrt{\frac{Z_{n-1}^{(\mathbf{s})}}{1+Z_{n-1}^{(\mathbf{s})}}}\sqrt{\frac{Z_{n-1}^{(\mathbf{t})}}{1+Z_{n-1}^{(\mathbf{t})}}} + \sqrt{\frac{\overline{Z_{n-1}^{(\mathbf{s})}}}{1+Z_{n-1}^{(\mathbf{s})}}}\sqrt{\frac{\overline{Z_{n-1}^{(\mathbf{t})}}}{1+Z_{n-1}^{(\mathbf{t})}}}\rho_{n-1}^{(\mathbf{s},\mathbf{t})} & \text{if } (s_n,t_n)=(+,+), \\[12pt] 2\sqrt{\frac{Z_{n-1}^{(\mathbf{s})}}{1+Z_{n-1}^{(\mathbf{s})}}}\sqrt{\frac{\overline{Z_{n-1}^{(\mathbf{t})}}}{1+Z_{n-1}^{(\mathbf{t})}}} - \sqrt{\frac{\overline{Z_{n-1}^{(\mathbf{s})}}}{1+Z_{n-1}^{(\mathbf{s})}}}\sqrt{\frac{Z_{n-1}^{(\mathbf{t})}}{1+Z_{n-1}^{(\mathbf{t})}}}\rho_{n-1}^{(\mathbf{s},\mathbf{t})} & \text{if } (s_n,t_n)=(+,-), \\[12pt] 2\sqrt{\frac{\overline{Z_{n-1}^{(\mathbf{s})}}}{1+Z_{n-1}^{(\mathbf{s})}}}\sqrt{\frac{Z_{n-1}^{(\mathbf{t})}}{1+Z_{n-1}^{(\mathbf{t})}}} - \sqrt{\frac{Z_{n-1}^{(\mathbf{s})}}{1+Z_{n-1}^{(\mathbf{s})}}}\sqrt{\frac{\overline{Z_{n-1}^{(\mathbf{t})}}}{1+Z_{n-1}^{(\mathbf{t})}}}\rho_{n-1}^{(\mathbf{s},\mathbf{t})} & \text{if } (s_n,t_n)=(-,+), \\[12pt] 2\sqrt{\frac{\overline{Z_{n-1}^{(\mathbf{s})}}}{1+Z_{n-1}^{(\mathbf{s})}}}\sqrt{\frac{\overline{Z_{n-1}^{(\mathbf{t})}}}{1+Z_{n-1}^{(\mathbf{t})}}} + \sqrt{\frac{Z_{n-1}^{(\mathbf{s})}}{1+Z_{n-1}^{(\mathbf{s})}}}\sqrt{\frac{Z_{n-1}^{(\mathbf{t})}}{1+Z_{n-1}^{(\mathbf{t})}}}\rho_{n-1}^{(\mathbf{s},\mathbf{t})} & \text{if } (s_n,t_n)=(-,-). \end{cases} \tag{16}$$

The last inequality follows by observing that $\sqrt{\frac{x}{1+x}} \leq \frac{1}{\sqrt{2}}$ for $x \in [0,1]$ with equality iff $x = 1$. Furthermore, it is easy to see that the equality in all above chain of weak inequalities happens iff $(a,b) = (1,0)^2$. By symmetry, this argument also applies to the case of $(s_n,t_n) = (-,+)$. ∎

**Property 3.** *If* $\mathbf{s} \neq \mathbf{t}$ *then* $\rho_n^{(\mathbf{s},\mathbf{t})} \leq \frac{1}{3}$.

*Proof:* Let $\mathbf{p} \triangleq \mathrm{CP}\,[\mathbf{s},\mathbf{t}]$ be the common prefix of $\mathbf{s}$ and $\mathbf{t}$ and $m \triangleq |\mathbf{p}|$ its length. Then $s_{m+1} \neq t_{m+1}$ and Property 2 together with either (11b) or (11c) result in:

$$\rho_n^{(\mathbf{s},\mathbf{t})} \leq \rho_{m+1}^{(\mathbf{p}s_{m+1},\mathbf{p}t_{m+1})}$$
$$= 2\sqrt{\frac{Z_m^{(\mathbf{p})}}{1+Z_m^{(\mathbf{p})}}}\sqrt{\frac{\overline{Z_m^{(\mathbf{p})}}}{1+\overline{Z_m^{(\mathbf{p})}}}} - \sqrt{\frac{\overline{Z_m^{(\mathbf{p})}}}{1+\overline{Z_m^{(\mathbf{p})}}}}\sqrt{\frac{Z_m^{(\mathbf{p})}}{1+Z_m^{(\mathbf{p})}}}$$
$$= \sqrt{\frac{Z_m^{(\mathbf{p})}\overline{Z_m^{(\mathbf{p})}}}{\left(1+Z_m^{(\mathbf{p})}\right)\left(1+\overline{Z_m^{(\mathbf{p})}}\right)}} = \sqrt{\frac{Z_m^{(\mathbf{p})}\overline{Z_m^{(\mathbf{p})}}}{2+Z_m^{(\mathbf{p})}\overline{Z_m^{(\mathbf{p})}}}} \leq \frac{1}{3}$$

with equality iff $Z_m^{(\mathbf{p})} = \frac{1}{2}$. ∎

## IV. CONVERGENCE OF CORRELATION COEFFICIENTS

In the previous section we showed how correlation coefficients can be computed efficiently by single-step recursions and derived some algebraic properties of them. In this section we show that correlation coefficients converge to zero.

**Lemma 5.** *Let* $\mathbf{s}$ *and* $\mathbf{t}$ *be infinite sign sequences such that* $\mathbf{s} \neq \mathbf{t}$ *and* $\mathbf{s}^n$ *and* $\mathbf{t}^n$ *be the subsequences corresponding to their first* $n$ *elements respectively. Then* $\lim_{n\to\infty}\rho_n^{(\mathbf{s}^n,\mathbf{t}^n)} = 0$.

*Proof:* Let $m = |\mathrm{CP}\,[\mathbf{s},\mathbf{t}]|$ and $a_n \triangleq \rho_n^{(\mathbf{s}^n,\mathbf{t}^n)}$. For $n > m$, by Properties 1 and 3 we know $a_n \in [0,1/3]$ and by Property 2 it is decreasing. Hence, $a_n$ is a convergent sequence. Suppose its limit is $a^* > 0$. This implies for every $\varepsilon > 0$ there exist a $n_0$ such that for $n > n_0$, $a_n/a_{n-1} \geq 1 - \varepsilon$. By the continuity of (16), we must have $|Z_{n-1}^{(\mathbf{s}^{n-1})} - b_{s_n}| < \delta$ and $|Z_{n-1}^{(\mathbf{t}^{n-1})} - b_{t_n}| < \delta$ for all $n > n_0$ according to equality condition (iii) of Property 2, where $\delta$ is a quantity approaching zero as $\varepsilon$ gets small. This implies $s_n = s^*$ and $t_n = t^*$ for all $n > n_0$

[2]By Property 1 this condition implies $\rho_{n-1}^{(\mathbf{s},\mathbf{t})} = 0$.

because the evolutions of $Z$ do not allow $Z$ to jump from one extreme to the other. Without loss of generality, assume $s^* = +$ which in turn requires $Z_{n-1}^{(\mathbf{s}^{n-1})} > 1 - \delta$. Now we have an incompatible situation: $s_n = +$ for all $n > n_0$ will drive $Z_n^{(\mathbf{s}^n)}$ to 0. This shows $a_n$ cannot converge to a non-zero value. ∎

Additionally we can show that the average of the elements of the correlation coefficients matrix is exponentially small in $n$.

**Lemma 6.** *For any* $\mathbf{s},\mathbf{t} \in \{-,+\}^{n-1}$,

$$\frac{1}{4}\sum_{(s,t)\in\{-,+\}^2}\rho_n^{(\mathbf{s}s,\mathbf{t}t)} \leq \frac{2}{3}\rho_{n-1}^{(\mathbf{s},\mathbf{t})}.$$

*Proof:* Let $a = Z_{n-1}^{(\mathbf{s})}$, $b = Z_{n-1}^{(\mathbf{t})}$, $f(x) \triangleq \frac{1}{\sqrt{2}}\left[\sqrt{\frac{x}{1+x}} + \sqrt{\frac{\overline{x}}{1+\overline{x}}}\right]$, and $g(x) \triangleq \frac{1}{2}\left[\sqrt{\frac{\overline{x}}{1+x}} - \sqrt{\frac{x}{1+\overline{x}}}\right]$. Using (11a) to (11d) one can easily verify that:

$$\frac{1}{4}\sum_{(s,t)\in\{-,+\}^2}\rho_n^{(\mathbf{s}s,\mathbf{t}t)} = f(a)f(b)\rho_{n-1}^{(\mathbf{s},\mathbf{t})} + g(a)g(b)\rho_{n-1}^{(\mathbf{s},\mathbf{t})^2}$$
$$= \left[f(a)f(b) + g(a)g(b)\rho_{n-1}^{(\mathbf{s},\mathbf{t})}\right]\rho_{n-1}^{(\mathbf{s},\mathbf{t})}.$$

Now, observe that both sides of the above are positive and:

$$\left[f(a)f(b) + g(a)g(b)\rho_{n-1}^{(\mathbf{s},\mathbf{t})}\right]^2$$
$$\overset{(*)}{\leq} \left[f(a)^2 + \rho_{n-1}^{(\mathbf{s},\mathbf{t})}g(a)^2\right]\left[f(b)^2 + \rho_{n-1}^{(\mathbf{s},\mathbf{t})}g(b)^2\right]$$
$$\leq \left[f(a)^2 + g(a)^2\right]\left[f(b)^2 + g(b)^2\right]$$

where (*) follows from the Cauchy-Schwarz inequality. It is easy to see $f(x)^2 + g(x)^2 = \frac{1}{2}\left(1 + \sqrt{\frac{x\overline{x}}{(1+x)(1+\overline{x})}}\right)$ which is maximized at $x = \frac{1}{2}$ (for $x \in [0,1]$) with value $\frac{2}{3}$. ∎

**Corollary 4.** *The average of the normalized correlation matrix elements satisfies:*

$$\frac{1}{4^n}\sum_{\mathbf{s},\mathbf{t}\in\{-,+\}^n}\rho_n^{(\mathbf{s},\mathbf{t})} \leq \left(\frac{2}{3}\right)^n$$

*Proof:* The result follows by applying Lemma 6 $n$ times and observing that $\rho_0 = 1$. ∎

## V. RATE OF CONVERGENCE

Corollary 4 implies that for large enough $n$, almost all of non-diagonal entries of $\rho_n$ are small. However, the bound it gives is not strong enough to show the asymptotic tightness of the union bound on the block error probability of polar codes. For that, one has to show (i) that the correlations decay like $O\big(2^{-(1+\alpha)n}\big)$ for some $\alpha > 0$, and (ii) that this bound applies not just to the average value of $\rho_n^{(\mathbf{s},\mathbf{t})}$ but to $\max_{\mathbf{t} \neq \mathbf{s}} \rho_n^{(\mathbf{s},\mathbf{t})}$ for the $\mathbf{s}$'s and $\mathbf{t}$'s which index the information channels.

To this end, we establish a probabilistic framework similar to that used in [1] for proving the channel polarization theorem.

Let $S_1, S_2, \ldots,$ be i.i.d Bernoulli $\left(\frac{1}{2}\right)$ random variables such that $S_i \in \{-, +\}$, define $\mathbf{S}^n \triangleq (S_1, S_2, \ldots, S_n)$ and $\mathcal{F}_n \triangleq \sigma(\mathbf{S}^n)$ as the $\sigma$-algebra generated by random vector $\mathbf{S}^n$. We consider the random variables $Z_n^{(\mathbf{S})} = \mathbb{E}\big[E_n^{(\mathbf{S}^n)}|\mathbf{S}^n\big]$ and $\rho_n^{(\mathbf{S}^n,\mathbf{t}^n)}$ for $\mathbf{t}^n \in \{-, +\}^n$ which are all $\mathcal{F}_n$ measurable.

We show that for any $\alpha > 0$, $\max_{\mathbf{t}^n \neq \mathbf{S}^n} \rho_n^{(\mathbf{S}^n,\mathbf{t}^n)} \leq 2^{-(1+\alpha)n}$ with very high probability for sufficiently large $n$.

### A. Closely related $\mathbf{s}$ and $\mathbf{t}$

Let us first focus on $\rho_n^{(\mathbf{s},\mathbf{t})}$ for $\mathbf{s}$ and $\mathbf{t}$ sharing a long common prefix. Recall that $|\,\mathrm{CP}\,[\mathbf{s},\mathbf{t}]\,|$ denotes the length of this prefix.

**Lemma 7.** *Fix $\alpha > 0$. Set $m_n \triangleq 4\log\big(2(1+\alpha)n - 1\big)$. Then:*

$$\lim_{n \to \infty} \mathbb{P}\left[\max_{\mathbf{t}^n \neq \mathbf{S}^n : |\mathrm{CP}[\mathbf{S}^n,\mathbf{t}^n]| \geq m_n} \rho_n^{(\mathbf{S}^n,\mathbf{t}^n)} \leq 2^{-(1+\alpha)n}\right] = 1$$

*Proof:* Let $\mathbf{P} = \mathrm{CP}\,[\mathbf{S}^n,\mathbf{t}^n]$ and $n_0 = |\mathbf{P}|$. Observe that $\mathbf{P}$ is a uniformly chosen sign sequence in $\{-, +\}^{n_0}$. According to Property 2, $\rho_{n_0}^{(\mathbf{P},\mathbf{P})} = 1$ and:

$$\rho_n^{(\mathbf{S}^n,\mathbf{t}^n)} < \rho_{n_0+1}^{\big(\mathbf{P}S_{n_0+1},\mathbf{P}t_{n_0+1}\big)} = \sqrt{\frac{Z_{n_0}^{(\mathbf{P})}\overline{Z_{n_0}^{(\mathbf{P})}}}{2 + Z_{n_0}^{(\mathbf{P})}\overline{Z_{n_0}^{(\mathbf{P})}}}}$$

$$\leq \min\left\{\sqrt{\frac{1}{2}Z_{n_0}^{(\mathbf{P})}}, \sqrt{\frac{1}{2}\overline{Z_{n_0}^{(\mathbf{P})}}}\right\}.$$

Results of [2] show that for any fixed $0 < \beta < 1/2$ and $\delta > 0$ there exist a $m_0$ such that for $n_0 \geq m_0$

$$\mathbb{P}\left[Z_{n_0}^{(\mathbf{P})} \in [2^{-N_0^\beta}, 1 - 2^{-N_0^\beta}]\right] < \delta$$

where $N_0 = 2^{n_0}$.

In particular we take $\beta = \frac{1}{4}$ in the above bound and take $n$ large enough so that $m_n \geq m_0$. Hence $n_0 \geq m_n \geq m_0$, and with probability at least $1 - \delta$, $Z_{n_0}^{(\mathbf{P})}$ is extremal. Together with $2^{-N_0^{1/4}} \leq 2^{-2(1+\alpha)n+1}$ we get

$$\mathbb{P}\left[\rho_n^{(\mathbf{S}^n,\mathbf{t}^n)} \leq 2^{-(1+\alpha)n}\right] \geq 1 - \delta. \qquad \blacksquare$$

### B. Distantly related $\mathbf{s}$ and $\mathbf{t}$

A more involved task is find and upper-bound on $\rho_n^{(\mathbf{s},\mathbf{t})}$ when $\mathbf{s}$ and $\mathbf{t}$ do not have a long common prefix. For this purpose we first seek an upper-bound on $\rho_n^{(\mathbf{S}^n,\mathbf{t}^n)}/\rho_{n-1}^{(\mathbf{S}^{n-1},\mathbf{t}^{n-1})}$ only

in terms of $\mathbf{S}^{n-1}$, $S_n$ and $p_n = |\mathrm{CP}\,[\mathbf{S}^n,\mathbf{t}^n]|$, denoted as $\chi\big(\mathbf{S}^{n-1}, S_n, p_n\big)$.

To this end, let:

$$M\left(S_n, t_n, \rho_{n-1}^{(\mathbf{S}^{n-1},\mathbf{t}^{n-1})}, Z_{n-1}^{(\mathbf{S})}, Z_{n-1}^{(\mathbf{t})}\right) \triangleq \frac{\rho_n^{(\mathbf{S}^n,\mathbf{t}^n)}}{\rho_{n-1}^{(\mathbf{S}^{n-1},\mathbf{t}^{n-1})}}.$$

$M(s, t, r, a, b)$ takes four possible forms according to (16), each of which can be bounded as:

$$M(+, t, r, a, b) \leq \min\left\{1, \sqrt{2a} + r\right\}$$
$$M(-, t, r, a, b) \leq \min\left\{1, \sqrt{2\overline{a}} + r\right\}$$

using Lemma 8 (and triangle inequality if $s \neq t$):

**Lemma 8.** *Let $f(x) \triangleq \sqrt{\frac{x}{1+x}}$ and $g(x) \triangleq \sqrt{\frac{\overline{x}}{1+x}}$. Define*

$$F(r, a, b) \triangleq 2f(a)f(b) + g(a)g(b)r.$$

*Then*

$$F(r, a, b) \leq \min\left\{1, \sqrt{2a} + r\right\}, \qquad (17)$$

*for all $0 \leq r \leq 1, 0 \leq a \leq 1, 0 \leq b \leq 1$.*

*Proof:* Observe that $F(r, a, b) \geq 0$ by construction and:

$$F(r, a, b)^2 = (2f(a)f(b) + g(a)g(b)r)^2$$
$$\overset{r<1}{\leq} (2f(a)f(b) + g(a)g(b))^2$$
$$\overset{(*)}{\leq} (2f(a)^2 + g(a)^2)(2f(b)^2 + g(b)^2)$$

where $(*)$ follows by Cauchy-Schwarz inequality. Furthermore, $2f(x)^2 + g(x)^2 = \frac{2x}{1+x} + \frac{\overline{x}}{1+x} = 1$ which proves $F(r, a, b) \leq 1$.

It is also easy to verify $f(x) \leq \frac{1}{\sqrt{2}}$ and $g(x) \leq 1$ for $\forall x \in [0, 1]$. Hence:

$$F(r, a, b) \leq \sqrt{2}f(a) + r \leq \sqrt{2a} + r$$

where the last inequality follows by observing that $\sqrt{\frac{x}{1+x}} \leq \sqrt{x}$ since $x \geq 0$. $\blacksquare$

Observe that the upper-bounds on $M$ depend only $Z_{n-1}^{(\mathbf{S})}$ and $\rho_{n-1}^{(\mathbf{S}^{n-1},\mathbf{t}^{n-1})}$. Let us also define

$$\rho_{n,p}^{(\mathbf{s}^n,*)} \triangleq \max_{\mathbf{t}^n \neq \mathbf{s}^n : |\mathrm{CP}[\mathbf{s}^n,\mathbf{t}^n]| \leq p} \rho_n^{(\mathbf{s}^n,\mathbf{t}^n)}.$$

Consequently we may choose:

$$\chi\big(\mathbf{S}^{n-1}, +, p_n\big) = \min\left\{1, \sqrt{2Z_{n-1}^{(\mathbf{S})}} + \rho_{n-1,p_n}^{(\mathbf{S}^{n-1},*)}\right\} \quad (18\mathrm{a})$$

$$\chi\big(\mathbf{S}^{n-1}, -, p_n\big) = \min\left\{1, \sqrt{2\overline{Z_{n-1}^{(\mathbf{S})}}} + \rho_{n-1,p_n}^{(\mathbf{S}^{n-1},*)}\right\} \quad (18\mathrm{b})$$

Now we would like to show that $\min_{s_n} \chi\big(\mathbf{S}^{n-1}, s_n, p_n\big)$ gets arbitrarily small with very high probability. For this, we first need the following lemma:

**Lemma 9.** *For any sequence $p_n$ such that $\lim_{n\to\infty} \frac{n}{2} - p_n = \infty$ and any fixed $\gamma > 0$,*

$$\lim_{n \to \infty} \mathbb{P}\left[\forall i \geq \frac{n}{2} : \rho_{i,p_n}^{(\mathbf{S}^i,*)} \leq \gamma\right] = 1. \qquad (19)$$

*Proof:* Observe that for fixed $p$, $\rho_{i,p}^{(\mathbf{s}^i,*)}$ is decreasing in $i$ (if $i > p$). Hence $\rho_{n/2,p_n}^{(\mathbf{s}^{n/2},*)} \leq \gamma$ implies $\rho_{i,p_n}^{(\mathbf{s}^n,*)} \leq \gamma$ for all $i \geq n/2$.

Suppose $\mathbf{s}$ is a sequence such that for some $\mathbf{t} \neq \mathbf{s}$ with $|\mathrm{CP}\,[\mathbf{s},\mathbf{t}]| \leq p_n$, $\rho_{n/2}^{(\mathbf{s}^{n/2},\mathbf{t}^{n/2})} > \gamma$. Recall that $\mathbf{s}^i$ (resp. $\mathbf{t}^i$) denotes the subsequence of $\mathbf{s}$ (resp. $\mathbf{t}$) including its first $i$ elements.

Define $a_i \triangleq \rho_i^{(\mathbf{s}^i,\mathbf{t}^i)}$ and $m_i \triangleq a_i/a_{i-1}$. It is clear that $a_{p_n+1} \leq \frac{1}{3}$ and $a_i$ is decreasing for $i > p_n$ by Properties 3 and 2.

For any $0 < \varepsilon < 1$, $a_{n/2} > \gamma$ implies that the number of indices $i \in \{p_n+2, p_n+3, \ldots, \frac{n}{2}\}$ for which $m_i \leq 1 - \varepsilon$ is at most $\frac{\log(3\gamma)}{\log(1-\varepsilon)}$.

Let $l = \frac{n}{2} - p_n - 1$, take $\varepsilon = 1/\sqrt{l}$, and observe that the number of indices for which $m_i \leq 1 - 1/\sqrt{l}$ is at most

$$\frac{\log(3\gamma)}{\log(1-1/\sqrt{l})} \leq \frac{-\log(3\gamma)}{1/\sqrt{l}} = c_\gamma\sqrt{l},$$

where $c_\gamma$ is a constant that depends on $\gamma$ only. These indices partition the interval $[p_n+2 : \frac{n}{2}]$ into at most $c_\gamma\sqrt{l}$ segments, one of those must have a length at least $c_\gamma^{-1}\sqrt{l}$. Let us only consider this "long" segment:

The fact that $m_i \geq 1-1/\sqrt{l}$ on this segment implies the sign sequence $s_{p_n+2}, \ldots, s_{n/2}$ must be constant on this segment (cf. Proof of Lemma 5). The set of sequences of length $l$ which have a run of the same sign for an interval of length $c_\gamma^{-1}\sqrt{l}$ has probability at most $2l \cdot 2^{-c_\gamma^{-1}\sqrt{l}}$. However, by assumption $l = \frac{n}{2} - p_n - 1$ goes to infinity as $n$ gets large. Hence the probability of having such a $\mathbf{s}$ sequence gets arbitrarily small when $n$ gets large. ∎

**Lemma 10.** *For any sequence $p_n$ such that $\lim_{n\to\infty}\frac{n}{2} - p_n = \infty$ and any fixed $\alpha > 0$*

$$\lim_{n\to\infty}\mathbb{P}\left[\forall i > \frac{n}{2} : \min_{s_i}\chi\left(\mathbf{S}^{i-1}, s_i, p_n\right) \leq 2^{-4(1+\alpha)}\right] = 1.$$

*Proof:* Let

$$\mathcal{G}_R(n) \triangleq \left\{\forall i \geq \frac{n}{2} : \rho_{i,p_n}^{(\mathbf{S}^i,*)} \leq 2^{-(5+4\alpha)}\right\}.$$

Observe that Lemma 9 implies for any $\delta > 0$ there exist a $n_0$ such that $\mathbb{P}\left[\mathcal{G}_R(n)\right] \geq 1 - \delta/2$ for $n \geq n_0$.

Let

$$\mathcal{G}_Z(n) \triangleq \left\{\forall i \geq \frac{n}{2} : Z_i^{(\mathbf{S})} \notin \left[2^{-(11+8\alpha)}, 1-2^{-(11+8\alpha)}\right]\right\}.$$

Likewise, the convergence of $Z$ process implies that there exist a $n_1$ such that for any $n \geq n_1$ $\mathbb{P}\left[\mathcal{G}_Z(n)\right] \geq 1 - \delta/2$.

Now (18a) and (18b) imply that for $\mathbf{S} \in \mathcal{G}_R(n) \cap \mathcal{G}_Z(n)$, $\forall i > \frac{n}{2}$, either $\chi\left(\mathbf{S}^{i-1}, +, p_n\right) \leq 2^{-4(1+\alpha)}$ or $\chi\left(\mathbf{S}^{i-1}, -, p_n\right) \leq 2^{-4(1+\alpha)}$. For $n \geq \max\{n_0, n_1\}$, $\mathbb{P}\left[\mathcal{G}_R(n) \cap \mathcal{G}_Z(n)\right] \geq 1 - \delta$ which proves the claim. ∎

**Lemma 11.** *Fix $\alpha > 0$ and let $m_n \triangleq 4\log\left(2(1+\alpha)n - 1\right)$ (as in Lemma 7). Then:*

$$\lim_{n\to\infty}\mathbb{P}\left[\max_{\mathbf{t}\neq\mathbf{S}:|\mathrm{CP}[\mathbf{S},\mathbf{t}]|<m_n}\rho_n^{(\mathbf{S},\mathbf{t})} \leq 2^{-(1+\alpha)n}\right] = 1$$

*Proof:* For any $p$, let us define the random variable $B_{n,p} \triangleq \mathbb{1}\left[S_n = \arg\min_s\chi\left(\mathbf{S}^{n-1}, s, p\right)\right]$. It is easy to see that $\mathbb{P}\left[B_{n,p} = 1|\mathcal{F}_{n-1}\right] = \mathbb{P}\left[B_{n,p} = 0|\mathcal{F}_{n-1}\right] = \frac{1}{2}$.

Fix $\varepsilon > 0$ and let

$$\mathcal{G}_B(n, p, \varepsilon) \triangleq \left\{\frac{1}{n/2}\sum_{i=n/2+1}^{n} B_{i,p} \geq \frac{1-\varepsilon}{2}\right\}.$$

Observe that $\mathbb{P}\left[\mathcal{G}_B(n, p, \varepsilon)\right]$ is independent of $p$ and by the Weak Law of Large Numbers for any $\delta > 0$ there exist a $n_0$ such that $\mathbb{P}\left[\mathcal{G}_B(n, p, \varepsilon)\right] \geq 1 - \delta/2$ for $n \geq n_0$.

Fix $\alpha' > 0$ and define

$$\mathcal{G}_\chi(n) \triangleq \left\{i > \frac{n}{2} : \min_{s_i}\chi(\mathbf{S}^{i-1}, s_i, m_n) \leq 2^{-4(1+\alpha')}\right\}$$

Since $\lim_{n\to\infty}\frac{n}{2} - m_n = \infty$, in view of Lemma 10, there exist $n_1$ such that $\mathbb{P}\left[\mathcal{G}_\chi(n)\right] \geq 1 - \delta/2$ for $n \geq n_1$.

For $n \geq \max\{n_0, n_1\}$, $\mathbb{P}\left[\mathcal{G}_B(n, m_n, \varepsilon) \cap \mathcal{G}_\chi(n)\right] \geq 1 - \delta$ and for $\mathbf{S}^n \in \mathcal{G}_B(n, m_n, \varepsilon) \cap \mathcal{G}_\chi(n)$ and any $\mathbf{t}^n \neq \mathbf{S}^n$ such that $|\mathrm{CP}\,[\mathbf{S}^n, \mathbf{t}^n]| < m_n$ we have:

$$\begin{aligned}
\log\left(\rho_n^{(\mathbf{S}^n,\mathbf{t}^n)}\right) &\leq \log\left(\rho_{n/2}^{(\mathbf{S}^{n/2},\mathbf{t}^{n/2})}\right) \\
&\quad + \sum_{i=n/2+1}^{n}\log\left(\chi\left(\mathbf{S}^{i-1}, S_i, m_n\right)\right) \\
&\overset{(*)}{\leq} \sum_{i=n/2+1}^{n} -4(1+\alpha')B_{i,m_n} \\
&\leq -n(1-\varepsilon)(1+\alpha').
\end{aligned}$$

In the above, $(*)$ follows from the fact that $0 \geq \rho_n^{(\mathbf{s},\mathbf{t})} \leq 1$ and observing that if $B_{i,m_n} = 1$ then $\chi(\mathbf{S}^{i-1}, S_i, m_n) \leq 2^{-4(1+\alpha')}$ (as $\mathbf{S} \in \mathcal{G}_\chi(n)$), otherwise $\chi(\mathbf{S}^{i-1}, S_i, m_n) \leq 1$ hence:

$$\log\left(\chi\left(\mathbf{S}^{i-1}, S_i, m_n\right)\right) \leq -4(1+\alpha')B_{i,m_n}.$$

For $\mathbf{S} \in \mathcal{G}_B(n, m_n, \varepsilon)$, $\sum_{i=n/2+1}^{n} B_{i,m_n} \geq \frac{n(1-\varepsilon)}{4}$.

Choosing $\alpha'$ and $\varepsilon$ such that $(1-\varepsilon)(1+\alpha') \geq (1+\alpha)$ proves the claim. ∎

**Theorem 1.** *For any $\alpha > 0$.*

$$\lim_{n\to\infty}\mathbb{P}\left[\max_{\mathbf{t}\neq\mathbf{S}}\rho_n^{(\mathbf{S},\mathbf{t})} \leq 2^{-n(1+\alpha)}\right] = 1. \qquad (20)$$

*Proof:* The proof follows by combining the results of Lemma 7 and Lemma 11. ∎

## VI. Lower Bound on Probability of Error of Polar Codes

In this section, we use our results on correlations among polarized BECs to give lower-bounds on block error probability of Polar Codes over BEC. Recall the analysis of error of the code: The error event $\mathcal{E}$ is the union of error events in each of information channels: $\mathcal{E} = \bigcup_{\mathbf{s}\in\mathcal{A}}\mathcal{E}_\mathbf{s}$ where $\mathcal{A} \subset \{-, +\}^n$ is the set of information bits and $\mathcal{E}_\mathbf{s}$ denotes the error in $W_n^{(\mathbf{s})}$.

For a BEC — with a pessimistic assumption on decoder — a decision error happens exactly when an erasure happens. [3] Therefore, $\mathcal{E}_{\mathbf{s}} = \left\{ E_n^{(\mathbf{s})} = 1 \right\}$ and the union bound gives us:

$$\mathbb{P}\left[\mathcal{E}\right] \leq \sum_{\mathbf{s} \in \mathcal{A}} Z_n^{(\mathbf{s})} \tag{21}$$

A trivial lower-bound on the probability of decoding error is obtained by observing that $\mathcal{E} \supseteq \mathcal{E}_{\mathbf{s}}$, hence, $\mathbb{P}\left[\mathcal{E}\right] \geq \mathbb{P}\left[\mathcal{E}_{\mathbf{s}}\right]$ for any $\mathbf{s} \in \mathcal{A}$. In particular,

$$\mathbb{P}\left[\mathcal{E}\right] \geq \max_{\mathbf{s} \in \mathcal{A}} \mathbb{P}\left[\mathcal{E}_{\mathbf{s}}\right] = \max_{\mathbf{s} \in \mathcal{A}} Z_n^{(\mathbf{s})}. \tag{22}$$

However, having the second order statistics, one can use the inclusion–exclusion principle to obtain a much tighter lower-bound on probability of error.

**Lemma 12.** *Let $W$ be a $\mathsf{BEC}\,(\epsilon)$ and $\mathcal{C}_n$ be a polar code of block-length $N = 2^n$ with information bits $\mathcal{A}_n$. The block error probability of such a code, $P_e(\mathcal{C}_n)$ is lower-bounded as:*

$$P_e(\mathcal{C}_n) \geq \sum_{\mathbf{s} \in \mathcal{A}_n} Z_n^{(\mathbf{s})} - \frac{1}{2} \sum_{\substack{\mathbf{s},\mathbf{t} \in \mathcal{A}_n: \\ \mathbf{s} \neq \mathbf{t}}} \left[ Z_n^{(\mathbf{s})} Z_n^{(\mathbf{t})} \right.$$
$$\left. + \rho_n^{(\mathbf{s},\mathbf{t})} \sqrt{Z_n^{(\mathbf{s})} \overline{Z_n^{(\mathbf{s})}}} \sqrt{Z_n^{(\mathbf{t})} \overline{Z_n^{(\mathbf{t})}}} \right] \tag{23}$$

*where $\mathbf{Z}_n$ vector and $\rho_n$ matrix can be computed via single-step recursions explained in Section III.*

*Proof:* The result follows by applying the inclusion–exclusion principle to lower-bound the probability of $\bigcup_{\mathbf{s} \in \mathcal{A}_n} \mathcal{E}_{\mathbf{s}}$. ∎

While the lower-bound given by Lemma 12 is already useful in practice (see Section VII), we seek for a lower-bound that is theoretically more significant.

**Theorem 2.** *Let $W$ be a $\mathsf{BEC}\,(\epsilon)$ and $R < 1 - \epsilon$. Let $\mathcal{C}_n$ be a polar code of block length $N = 2^n$ with information bits $\mathcal{A}_n$ such that $|\mathcal{A}_n| = \lceil NR \rceil$. Let $P(N, R, \epsilon)$ be the sum of $\lceil NR \rceil$ smallest elements of the vector $\mathbf{Z}_n$. Then, for any fixed $\delta > 0$ and sufficiently large $n$:*

$$(1 - \delta)\, P\left(N, (1-\delta)R, \epsilon\right) \leq P_e(\mathcal{C}_n) \leq P(N, R, \epsilon).$$

*Proof:* The upper-bound is already known and we only need to prove the lower-bound. Let

$$\mathcal{D}_n = \left\{ \mathbf{s} \in \{-, +\}^n : \max_{\mathbf{t} \neq \mathbf{s}} \rho_n^{(\mathbf{s},\mathbf{t})} \leq \delta 2^{-n} \right\}$$

By Theorem 1 we know that $\lim_{n \to \infty} \frac{|\mathcal{D}_n|}{N} = 1$. Let, $\mathcal{C}'_n$ be the polar code defined by the information bits $\mathcal{A}'_n = \mathcal{A}_n \cap \mathcal{D}_n$ and $S'_n \triangleq \sum_{\mathbf{s} \in \mathcal{A}'_n} Z_n^{(\mathbf{s})}$. It is clear that $\lim_{n \to \infty} \frac{|\mathcal{A}'_n|}{|\mathcal{A}_n|} = 1$, $S'_n \leq P(N, R, \epsilon)$ (as $\mathcal{A}_n$ contains $\lceil NR \rceil$ smallest elements of $\mathbf{Z}_n$), and $P_e(\mathcal{C}'_n) \leq P_e(\mathcal{C}_n)$ as $\mathcal{C}'_n$ is a sub-code of $\mathcal{C}_n$.

Choose $n$ large enough such that $\frac{|\mathcal{A}'_n|}{|\mathcal{A}_n|} \geq 1 - \delta$ and $P(N, R, \epsilon) \leq \delta$ (note that this is possible since $R < 1 - \epsilon$ and the results of [2] suggest that $P(N, R, \epsilon) = O\left(2^{-\sqrt{N}}\right)$). By (23):

$$S'_n - P_e(\mathcal{C}_n) \leq S'_n - P_e(\mathcal{C}'_n)$$
$$\leq \frac{1}{2} \sum_{\substack{\mathbf{s},\mathbf{t} \in \mathcal{A}'_n: \\ \mathbf{s} \neq \mathbf{t}}} \left[ Z_n^{(\mathbf{s})} Z_n^{(\mathbf{t})} + \rho_n^{(\mathbf{s},\mathbf{t})} \sqrt{Z_n^{(\mathbf{s})} \overline{Z_n^{(\mathbf{s})}}} \sqrt{Z_n^{(\mathbf{t})} \overline{Z_n^{(\mathbf{t})}}} \right].$$

Observe that $\rho_n^{(\mathbf{s},\mathbf{t})} \leq \delta/N$ for all $\mathbf{s}, \mathbf{t}$ in the above summation, $\sum_{\mathbf{s},\mathbf{t} \in \mathcal{A}'_n : \mathbf{s} \neq \mathbf{t}} Z_n^{(\mathbf{s})} Z_n^{(\mathbf{t})} \leq \sum_{\mathbf{s},\mathbf{t} \in \mathcal{A}'_n} Z_n^{(\mathbf{s})} Z_n^{(\mathbf{t})} = S_n'^{\,2}$, and

$$\sum_{\mathbf{s},\mathbf{t} \in \mathcal{A}'_n : \mathbf{s} \neq \mathbf{t}} \sqrt{Z_n^{(\mathbf{s})} \overline{Z_n^{(\mathbf{t})}}} \sqrt{Z_n^{(\mathbf{t})} \overline{Z_n^{(\mathbf{t})}}}$$
$$\leq \sum_{\mathbf{s},\mathbf{t} \in \mathcal{A}'_n : \mathbf{s} \neq \mathbf{t}} \sqrt{Z_n^{(\mathbf{s})}} \sqrt{Z_n^{(\mathbf{t})}} \leq \sum_{\mathbf{s},\mathbf{t} \in \mathcal{A}'_n} \sqrt{Z_n^{(\mathbf{s})}} \sqrt{Z_n^{(\mathbf{t})}}$$
$$= \left[ \sum_{\mathbf{s} \in \mathcal{A}'_n} \sqrt{Z_n^{(\mathbf{s})}} \right]^2 \overset{(*)}{\leq} |\mathcal{A}'_n| \sum_{\mathbf{s} \in \mathcal{A}'_n} Z_n^{(\mathbf{s})} \leq N S'_n,$$

where (*) follows by the Cauchy-Schwarz inequality [4].

Therefore,

$$S'_n - P_e(\mathcal{C}_n) \leq \frac{1}{2} \left[ S_n'^{\,2} + \delta S'_n \right] \leq \delta S'_n,$$

where the last inequality follows by observing that $S'_n \leq P(N, R, \epsilon) \leq \delta$. As a result,

$$(1 - \delta) S'_n \leq P_e(\mathcal{C}_n)$$

$\mathcal{C}'_n$ is a code of rate $R' \geq (1 - \delta)R$ and by definition $S'_n \geq P(N, R', \epsilon) \geq P\left(N, (1 - \delta)R, \epsilon\right)$. Hence we can lower-bound the LHS of the above by substituting $S'_n$ with $P\left(N, (1 - \delta)R, \epsilon\right)$ which completes the proof. ∎

## VII. Numerical Results

In this section we provide a numerical example which confirms our theoretical results. We have considered Polar Codes of different rates on a $\mathsf{BEC}\,(0.5)$ and computed the upper-bound of (21), the trivial lower-bound of (22) and the tighter lower-bound of (23). We emphasize that we have exactly computed the lower-bound on the error probability by computing the correlation coefficients. We did the computations for block lengths of $N = 4096$ ($n = 12$) and $N = 16384$ ($n = 14$).

As shown in Table I, the proposed lower bound is much tighter than the trivial one. Moreover, the results show that the lower bound is very close to the upper bound of (21). This confirms that $P(N, R, \epsilon)$ (as defined in Theorem 2) is indeed a very good estimation for block error probability of Polar Codes over BEC.

---

[3] A practical decoder can break the ties randomly which increases the chance of correctly decoding the bit to $\frac{1}{2}$. An analysis analogous to the one we do in this section applies to such a decoder.

[4] For any set of $m$ numbers $x_i, i = 1, 2, \cdots, m$:

$$\left( \sum_{i=1}^{m} x_i \right)^2 \leq m \sum_{i=1}^{m} x_i^2$$

.

| $R$ | $\sum_{\mathbf{s}\in\mathcal{A}_n} Z_n^{(\mathbf{s})}$ | $\max_{\mathbf{s}\in\mathcal{A}_n} Z_n^{(\mathbf{s})}$ | Lower-bound (23) |
|---|---|---|---|
| 0.2 | $4.04 \cdot 10^{-18}$ | $3.43 \cdot 10^{-19}$ | $4.04 \cdot 10^{-18}$ |
| 0.25 | $1.87 \cdot 10^{-11}$ | $9.25 \cdot 10^{-13}$ | $1.87 \cdot 10^{-11}$ |
| 0.3 | $5.4 \cdot 10^{-7}$ | $2.29 \cdot 10^{-8}$ | $5.4 \cdot 10^{-7}$ |
| 0.35 | $8.14 \cdot 10^{-4}$ | $2.11 \cdot 10^{-5}$ | $8.12 \cdot 10^{-4}$ |
| 0.4 | $0.17$ | $3.49 \cdot 10^{-3}$ | $0.14$ |

(a) $N = 4096$

| $R$ | $\sum_{\mathbf{s}\in\mathcal{A}_n} Z_n^{(\mathbf{s})}$ | $\max_{\mathbf{s}\in\mathcal{A}_n} Z_n^{(\mathbf{s})}$ | Lower-bound (23) |
|---|---|---|---|
| 0.2 | $9.32 \cdot 10^{-36}$ | $4.72 \cdot 10^{-37}$ | $9.32 \cdot 10^{-36}$ |
| 0.25 | $1.32 \cdot 10^{-22}$ | $3.54 \cdot 10^{-24}$ | $1.32 \cdot 10^{-22}$ |
| 0.3 | $2.32 \cdot 10^{-13}$ | $5.4 \cdot 10^{-15}$ | $2.32 \cdot 10^{-13}$ |
| 0.35 | $2.63 \cdot 10^{-7}$ | $3.61 \cdot 10^{-9}$ | $2.63 \cdot 10^{-7}$ |
| 0.4 | $5.47 \cdot 10^{-3}$ | $4.91 \cdot 10^{-5}$ | $5.43 \cdot 10^{-3}$ |

(b) $N = 16384$

TABLE I: Bounds on Block Error Probability of Polar Code on BEC $(0.5)$

REFERENCES

[1] E. Arıkan, "Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," *IEEE Transactions on Information Theory*, vol. 55, no. 7, pp. 3051 –3073, July 2009.

[2] E. Arıkan and E. Telatar, "On the rate of channel polarization," in *IEEE International Symposium on Information Theory (ISIT), 2009*, July 2009, pp. 1493 –1495.