Polarization of Quasi-Static Fading Channels

Joseph J. Boutros
Dept. of Electrical & Computer Eng.
Texas A&M University at Qatar
Education City, 23874, Doha, Qatar
Email: boutros@tamu.edu

Ezio Biglieri
Dept. of Information & Communication Tech.
Universitat Pompeu Fabra
Barcelona, 08018, Spain
Email: e.biglieri@ieee.org

***Abstract*—This work investigates polar coding for block-fading channels. We show that polarization does occur at infinity for three types of channel multiplexers. Nevertheless, the polarization process is not unique, as it is shaped by the choice of the multiplexer. The fading-plane approach is used to study the outage behavior of polar coding at a fixed transmission rate. Two types of multiplexers are shown to provide full diversity at finite and infinite code length.**

I. Introduction

The construction of codes on graphs for quasi-static fading channels [3][4] proceeds from a random ensemble of low-density parity-check codes [5] and makes it suitable for fading by introducing a deterministic structure that guarantees a high or a maximal diversity order for information bits. In wireless communications [6][7], the standard binary-input memoryless Gaussian channel (BIAWGN) corresponds to the special and rare case (a zero-probability event) where all fading coefficients are equal. For all binary-input memoryless symmetric (BMS) channels, including the BIAWGN, many asymptotically optimal channel coding methods do exist, such as threshold saturation via spatial coupling [8] for convolutional low-density parity-check codes [9], and channel polarization which achieves capacity by converting the channel into multiple extremal channels [2].

This paper presents an investigation of polar code design for block-fading channels. Polar coding [2][10], which is deterministic by construction, does not involve any random ensemble as an extra degree of freedom for performance optimization. Furthermore, in the rich recent literature on polar codes, authors are mainly considering a single BMS channel. On a block-fading (BF) channel, as described later in this paper, polarization occurs on multiple parallel channels of the same type, each having a different mutual information [1]. Our work aims at answering the following basic questions:

- How does polarization operate on BF channels?
- Which permutation of coded bits (this is referred to as a *channel multiplexer* in [11]) will yield the "best" polarization on the BF channel?
- How does outage probability depend on signal-to-noise ratio and polar code length? (Here we assume that information transmission has a fixed rate, thus making outage events possible.)

The work of J.J. Boutros was supported by the QNRF project NPRP 5-597-2-241. The work of E. Bliglieri was supported by project TEC2012-34642.

In this paper, mutual information [1] is the main tool for analysis. Our paper starts with a description of the main BF channel features. Next, we introduce diagonal, horizontal, and uniform permutations (Sections III and IV), and we explain how polarization works on BF channels. Outage analysis and mutual information functions are studied in Section V.

II. The BIAWGN Channel with Fading

We first introduce the channel model and the notations. Binary modulation and coherent detection are assumed [7]. The channel is defined by its input–output relationship $y = \alpha x + \eta$, where x denotes the binary channel input, y its real output, η a realization of additive white Gaussian noise $\sim \mathcal{N}(0, \sigma^2 = N_0/2)$, and α the real nonnegative fading gain subject to the constraint $\mathbb{E}[\alpha^2] = 1$. The input symbol x takes on values $\pm\sqrt{RE_b} = \pm A$, where E_b is the average energy per bit, and $R \in [0, 1]$ is the coding rate. The signal-to-noise ratio (SNR) is denoted by $\gamma = E_b/N_0$.

Due to channel symmetry, the study of mutual information may suppose that $x = +\sqrt{RE_b}$, i.e. the all-zero codeword is sent. For a fixed fading α, the conditional log-ratio quantity $L \in \mathbb{R}$ at the channel output is

$$L = \log(\frac{p(y|\alpha, x = +A)}{p(y|\alpha, x = -A)} = \frac{2A\alpha}{\sigma^2}\, y \;\sim\; \mathcal{N}\left(4\alpha^2 R\gamma, 8\alpha^2 R\gamma\right).$$

Let $p = p(L|\alpha)$ denote the conditional Gaussian probability density function associated with L. The corresponding BMS fading channel, denoted from now on as $W(p)$, has mutual information

$$\mathcal{I}(p) = 1 - E_p[\log_2(1 + e^{-L}))] = I(\alpha^2\gamma), \qquad (1)$$

where $E_p[\,\cdot\,]$ denotes mathematical expectation over p. The first notation $\mathcal{I}(p)$ means that $\mathcal{I}$ is a linear functional from the L-density space to the real interval $[0, 1]$. The second notation $I(\alpha^2\gamma)$ indicates that I is a real function of the real instantaneous SNR $\alpha^2\gamma$.

On an ergodic fading channel, the fading is i.i.d. from one channel use to another. The L-density observed by the decoder is $p(L) = \int_{\alpha=0}^{+\infty} p(\alpha)p(L|\alpha)d\alpha$. For example, on an ergodic Rayleigh fading channel where $p(\alpha) = 2\alpha e^{-\alpha^2}$, the output L-density $p(L)$ is

$$\frac{1}{\sqrt{4R\gamma(1+R\gamma)}} \exp\left(\frac{L}{2}\left(1 - sgn(L)\sqrt{\frac{R\gamma+1}{R\gamma}}\right)\right) \qquad (2)$$

The capacity $\mathcal{I}(p(L))$ of the BMS channel defined by the density $p(L)$ given in (2) can be achieved using standard polar coding techniques. The comparison of the ergodic Rayleigh channel with the BIWGN would be interesting. Instead, we are going to consider channels of type $W(p = p(L|\alpha))$ where fading is constant over a period longer than a symbol duration.

For simplicity, we restrict our attention, as we did in [3], to a BF channel with two distinct fading values. This choice implies that, within a codeword of length N bits, $N/2$ symbols are affected by a fading gain α_1, and the remaining $N/2$ by a fading gain α_2. The two random variables α_1 and α_2 are i.i.d. There are $\binom{N}{2}$ possible associations of the two gains α_1, α_2 with the N symbol positions, each corresponding to a subset of the permutations of N symbols. Each such permutation may be embedded into the code itself, while keeping the channel unchanged, or embedded into the channel in the form of a multiplexer, while the code structure is left invariant. Throughout this paper, we follow the second viewpoint.

The two conditional L-densities at the output of our BF channel are $p_1 = p(L_1|\alpha_1)$ for the channel $W(p_1)$ with fading α_1, and $p_2 = p(L_2|\alpha_2)$ for the channel $W(p_2)$ with fading α_2. p_1 and p_2 are Gaussian, viz., $L_1 \sim \mathcal{N}\left(4\alpha_1^2 R\gamma, 8\alpha_1^2 R\gamma\right)$ and $L_2 \sim \mathcal{N}\left(4\alpha_2^2 R\gamma, 8\alpha_2^2 R\gamma\right)$. We assume no specific probability distribution for the fading coefficients $\alpha_i \in \mathbb{R}^+$, which may be for example Rayleigh-, or Rice-, or Nakagami-distributed. The only constraint we impose is the behavior of the tail of $p(\alpha)$ near the origin, which is assumed to yield $\mathbb{P}(\alpha^2\gamma \leq 1) = O(\gamma^{-d_0})$ as $\gamma \to \infty$. The exponent d_0 may be called the channel intrinsic diversity order. The main goal of full-diversity coding, as discussed in [3], is to achieve a state diversity [12] of order two, i.e., an error probability after decoding which decays as γ^{-2d_0}.

By considering the standard 2×2 polarization kernel $G_2 = \left[\begin{smallmatrix}1 & 0\\ 1 & 1\end{smallmatrix}\right]$, a codeword of length $N = 2^n$ bits is generated as $X = (X_1, X_2, \ldots, X_N) = (U_1, U_2, \ldots, U_N)G_2^{\otimes n}$, $X_i \in \mathbb{F}_2$ and $U_i \in \mathbb{F}_2$. The real transmitted symbol for a bit X_i is $x = (1-2X_i)\sqrt{R\gamma}$. We hasten to mention that X does not include the bit-reversal matrix used in [2]. The order of bits in codeword X is the natural order after polar coding at the multiplexer (MUX) input. Also, the encoding circuit shown in Figures 2 and 3 is the mirrored version of those found in the literature. The mirrored version is convenient for the BF channel and it is made possible thanks to the property $\left(G_2^{\otimes n}\right)^{-1} = G_2^{\otimes n}$.

Fig. 1 illustrates the two parallel fading channels with L-densities p_1 and p_2. The polarization process will directly depend on the order of symbols as defined by the multiplexer. Changing the multiplexer may completely reshape the polarization of the BF channel, as shown in the following sections.

Finally, let us recall some notations used to combine L-densities. Density evolution usually employed for the asymptotic study of low-density parity-check codes ensembles [5] is also one of the methods for polar codes construction [13]. The standard convolution of two densities, associated to the sum $L_1 + L_2$ of two log-ratio messages, is written as $p_1 \circledast p_2$.

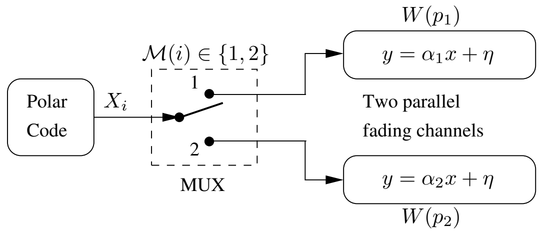

Figure 1. Model of polar coding on a BF channel.

The latter is the usual operation at bitnode level in codes on graphs. At checknode level, the sum of two bits (in $\mathbb{F}_2$) has the log-ratio message $L = 2\tanh^{-1}(\tanh(\frac{L_1}{2})\tanh\frac{L_2}{2}))$ and its density is written as $p_1 \boxast p_2$, also known as R-convolution.

III. Polarization with Diagonal Multiplexing

Assume that the polar code has length $N = 2^n$ bits. As recalled in the previous section, the bits in the codeword $X = UG_2^{\otimes n}$ have a natural order. Let us define the set $\aleph_N = \{1, 2, \ldots, N\} \subset \mathbb{N}$. Now, the diagonal multiplexer is equivalent to the mapping $\mathcal{M}_d : \aleph_N \to \mathbb{N}$, given by $\mathcal{M}_d(i) = 1$ for i odd and $\mathcal{M}_d(i) = 2$ for i even. Through the multiplexer, the binary element X_i is transmitted over the channel $W(p_{\mathcal{M}_d(i)})$ with fading coefficient $\alpha_{\mathcal{M}_d(i)}$, where L-densities p_1 and p_2 have been introduced in the previous section. Diagonal multiplexing on the two parallel channels $W(p_1)$ and $W(p_2)$ is shown in Fig. 2 for $N = 8$.

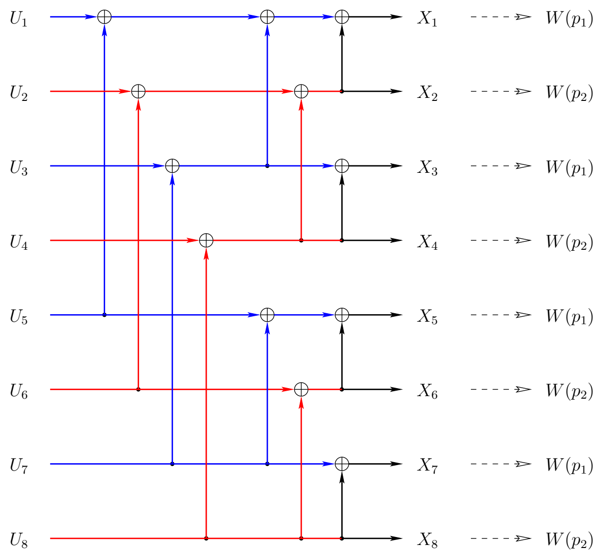

Figure 2. Illustration of diagonal multiplexing for $N = 8$ bits. The switch in the multiplexer does the sequence of positions (12121212).

For a better understanding of the polarization process, the notion of splitting is recalled and the notion of combining is introduced. In the definitions below, L-densities are not necessarily Gaussian. We have general L-densities consistent with the condition $p(-x) = p(x)e^{-x}$ for positive x [5].

Definition 1: (Splitting [2]) Let $W(p)$ be a BMS channel. The polarization kernel G_2 splits $W(p)$ into two channels $W^+ = W(p \circledast p)$ and $W^- = W(p \boxast p)$.

Definition 2: (Combining) Let $W(p_1)$ and $W(p_2)$ be two BMS channels. The polarization kernel G_2 converts $W(p_1)$ and $W(p_2)$ into $W^+ = W(p_1 \circledast p_2)$ and $W^- = W(p_1 \boxast p_2)$. This notion of combining has been brought out in [14] within the general framework of polarization acceleration.

From $X = UG_2^{\otimes n}$, as described in Fig. 2, the polarization process starts by combining the two parallel fading channels $W(p_1)$ and $W(p_2)$ with mutual information $\mathcal{I}(p_1) = I(\alpha_1^2\gamma)$ and $\mathcal{I}(p_2) = I(\alpha_2^2\gamma)$ respectively. Similar to splitting [2], there is no mutual information loss in combining:

$$\mathcal{I}(p_1) + \mathcal{I}(p_2) = \mathcal{I}(p_1 \circledast p_2) + \mathcal{I}(p_1 \boxast p_2). \tag{3}$$

Combining leads to MRC and SDC channels described below.

• *The Maximum Ratio Combining (MRC) Channel.*
The MRC channel is $W^+ = W(p_1 \circledast p_2)$. The density after convolution $p_1 \circledast p_2 \sim \mathcal{N}(4R(\alpha_1^2 + \alpha_2^2)\gamma, 8R(\alpha_1^2 + \alpha_2^2)\gamma)$ is equivalent to a message $L_1 + L_2 \propto (\alpha_1 y_1 + \alpha_2 y_2)$. This operation is known as Maximum Ratio Combining [6][7]. The mutual information is $\mathcal{I}(p_1 \circledast p_2) = I((\alpha_1^2 + \alpha_2^2)\gamma)$.

• *The Selection Diversity Combining (SDC) Channel.*
The SDC channel is $W^- = W(p_1 \boxast p_2)$, where $p_1 = p(L_1|\alpha_1)$ and $p_2 = p(L_2|\alpha_2)$ as in the MRC case above. Its L-density $p_1 \boxast p_2$ corresponds to the message $L = 2\tanh^{-1}(\tanh(\frac{L_1}{2})\tanh\frac{L_2}{2}))$. At low error rates, a good approximation would be $|L| = \min(|L_1|, |L_2|)$, i.e. the confidence value is imposed by the worst message. Selecting the channel with the highest fading value is known as Selection Diversity Combining [6][7]. Here, the opposite operation is made, yet the name SDC is maintained.

Proposition 1: For any point (α_1, α_2) in the fading plane $\mathbb{R}^+ \times \mathbb{R}^+$, polarization occurs on a BF channel with diagonal multiplexing as $n \to \infty$. The number of perfect channels is a fraction $\frac{1}{2}I(\alpha_1^2\gamma) + \frac{1}{2}I(\alpha_2^2\gamma)$ of N.
Proof: Let $\ell = 1 \ldots n$ denote the splitting and combining steps. At $\ell = 1$, the kernel G_2 starts combining the two fading channels $W(p_1)$ and $W(p_2)$. This operation is illustrated in black color on the right in Fig. 2. After combining, two independent and parallel polarization processes are engaged for $\ell = 2 \ldots n$. The SDC channel $W(p_1 \boxast p_2)$ involves $N/2$ bits and polarizes as any BMS [2]— this first process is shown in blue color on Fig. 2. The number of perfect channels is a fraction $\mathcal{I}(p_1 \boxast p_2)$ of $N/2$. Similarly, the second polarization process, illustrated in red color, involves $N/2$ bits under the MRC channel $W(p_1 \circledast p_2)$ and yields a fraction $\mathcal{I}(p_1 \circledast p_2)$ of perfect channels. Finally, we obtain $\frac{1}{2}\mathcal{I}(p_1 \boxast p_2) + \frac{1}{2}\mathcal{I}(p_1 \circledast p_2) = \frac{1}{2}I(\alpha_1^2\gamma) + \frac{1}{2}I(\alpha_2^2\gamma)$ as a fraction of N. *QED.*

IV. Horizontal and Uniform Multiplexing

The horizontal multiplexer is built from the mapping $\mathcal{M}_h : \aleph_N \to \mathbb{N}$, given by $\mathcal{M}_h(i) = 1$ for $1 \leq i \leq N/2$ and $\mathcal{M}_h(i) = 2$ for $N/2 < i \leq N$. Horizontal multiplexing on the two parallel channels $W(p_1)$ and $W(p_2)$ is shown in Fig. 3 for $N = 8$.

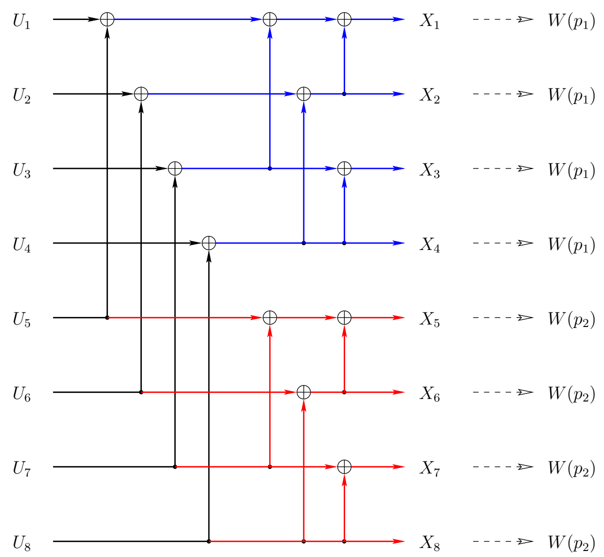

Figure 3. Illustration of horizontal multiplexing for $N = 8$ bits. The switch in the multiplexer does the sequence of positions (11112222).

The uniform multiplexer makes a random equiprobable selection between positions 1 and 2 of the switch. Its mapping $\mathcal{M}_u$ satisfies $\mathcal{P}(\mathcal{M}_u(i) = 1) = \mathcal{P}(\mathcal{M}_u(i) = 2) = \frac{1}{2}$, $\forall i \in \aleph_N$.

Proposition 2: For any point (α_1, α_2) in the fading plane $\mathbb{R}^+ \times \mathbb{R}^+$, polarization occurs on a BF channel with both horizontal and uniform multiplexing as $n \to \infty$. The number of perfect channels is a fraction $\frac{1}{2}I(\alpha_1^2\gamma) + \frac{1}{2}I(\alpha_2^2\gamma)$ of N.
Proof: The proof is similar to that for the diagonal multiplexer but with a big difference in the polarization process. (i) For horizontal multiplexing, two splitting processes of length $n-1$ start in parallel on both $W(p_1)$ and $W(p_2)$. As for any BMS channel, this leads to the polarization of $W(p_1)$ and $W(p_2)$ separately, when $n \gg 1$. At the final step, a combining process recombines densities from the two splitting processes such that mutual information pairs are converted as follows: $(0, 0)$ and $(1, 1)$ are unchanged, $(1, 0)$ and $(0, 1)$ are both converted into $(0, 1)$. This leads to the announced result. In Fig. 3, the polarization of $W(p_1)$ via an $n-1$-stage splitting is shown in blue color, the one for $W(p_2)$ is shown in red. The final combining stage is drawn in black on the left side. (ii) For the uniform multiplexer, because of the equiprobable selection of $W(p_1)$ and $W(p_2)$, the polar code observes a channel with L-density $\frac{1}{2}p_1 + \frac{1}{2}p_2$. Since $\mathcal{I}(p)$ is a linear functional, we find a standard polarization of a BMS channel with mutual information $\frac{1}{2}\mathcal{I}(p_1) + \frac{1}{2}\mathcal{I}(p_2)$. *QED.*

Polarization on BF channels will be analyzed via the fading plane tool, as described in the next section.

V. Fading Plane Analysis of Polarization

In the sequel, we fix the coding rate to $R = \frac{1}{2}$, which is the largest achievable rate of a full-diversity code for the channel considered here [3]. For a given point (α_1, α_2) in the fading plane $\mathbb{R}^+ \times \mathbb{R}^+$, a mutual information outage (MIO) occurs if $\frac{1}{2}I(\alpha_1^2\gamma) + \frac{1}{2}I(\alpha_2^2\gamma) < R = \frac{1}{2}$.

Definition 3: The MIO region $R_o(\mathcal{C},\gamma)$ for the BF channel is $R_o(\mathcal{C},\gamma) = \left\{(\alpha_1,\alpha_2) : I(\alpha_1^2\gamma) + I(\alpha_2^2\gamma) \leq 1\right\}$. The MIO boundary is $B_o(\mathcal{C},\gamma) = \left\{(\alpha_1,\alpha_2) : I(\alpha_1^2\gamma) + I(\alpha_2^2\gamma) = 1\right\}$. The symbol $\mathcal{C}$ stands for capacity. An illustration of B_o is given in Fig. 4 below.

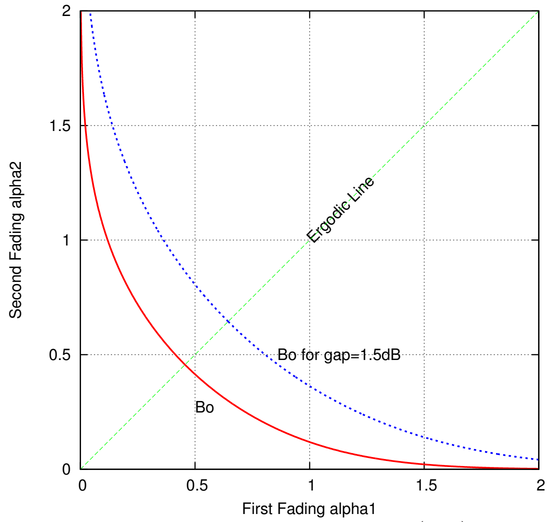

Figure 4. Mutual information outage boundary $B_o(\mathcal{C},\gamma)$ (red curve), for $\gamma = 7$dB. We also show $B_o(\mathcal{C},g\gamma)$ (blue curve) for SNR gap g=1.5dB.

At high SNR γ, the MIO probability has double state diversity order $P_o(\mathcal{C},\gamma) = \int_{R_o(\mathcal{C},\gamma)} p(\alpha_1)p(\alpha_2)d\alpha_1 d\alpha_2 \propto \frac{1}{(\gamma^{d_0})^2}$. For example, in Rayleigh channels ($d_0 = 1$), $P_o(\mathcal{C},\gamma) \approx \frac{4}{\gamma^2}$ and 73% of this outage is due to "unbalanced fading" [4]. Unbalanced fading plays a critical role also on general fading channels. Indeed, the probability of a fading area located near the tails of $B_o(\mathcal{C},\gamma)$ varies as $\frac{1}{(\gamma^{d_0})^1}$, whereas the fading area near the origin has probability $O(\frac{1}{(\gamma^{d_0})^2})$. Thanks to the symmetry in polar coding ($\circledast$ and $\boxtimes$ are Abelian operators), it is sufficient to assume that $0 \leq \alpha_2 \leq \alpha_1 < \infty$.
Let $t = \frac{\alpha_1}{\alpha_2} \in [1,\infty)$ be the unbalance factor. Now, we make a parametric representation of the lower part of $B_o(\mathcal{C},\gamma)$ under the ergodic line. For $t = 1$ the point is on the intersection of B_o and the ergodic line. By increasing t, the fading point moves to the right and downwards.

Proposition 3: Let $t = \frac{\alpha_1}{\alpha_2} \geq 1$. On the outage boundary $B_o(\mathcal{C},\gamma)$, we have $I(\alpha_1^2\gamma) = I(\chi(t))$ and $I(\alpha_2^2\gamma) = I(\psi(t))$, both being independent of SNR γ.
Proof: We have $I(\alpha_1^2\gamma) + I(\alpha_2^2\gamma) = 1$ when $(\alpha_1,\alpha_2) \in B_o(\mathcal{C},\gamma)$. Introducing the unbalance factor t, we get $I(\alpha_1^2\gamma) + I(\frac{\alpha_1^2\gamma}{t^2}) = 1$. The unbalance factor can be solved as

$$t = \left(\frac{\alpha_1^2\gamma}{I^{-1}\left(1 - I(\alpha_1^2\gamma)\right)}\right)^{\frac{1}{2}} = \chi^{-1}(\alpha_1^2\gamma), \qquad (4)$$

which is equivalent to $\alpha_1^2\gamma = \chi(t)$. Similarly, $I(t^2\alpha_2^2\gamma) + I(\alpha_2^2\gamma) = 1$, then t can be solved as

$$t = \left(\frac{I^{-1}\left(1 - I(\alpha_2^2\gamma)\right)}{\alpha_2^2\gamma}\right)^{\frac{1}{2}} = \psi^{-1}(\alpha_2^2\gamma), \qquad (5)$$

which is equivalent to $\alpha_2^2\gamma = \psi(t)$. *QED.*

From the proof of proposition 3, at a fixed SNR γ, the parametric representation of $B_o(\mathcal{C},\gamma)$ is given by fading points with squared components $(\alpha_1^2 = \frac{\chi(t)}{\gamma}, \alpha_2^2 = \frac{\psi(t)}{\gamma})$.

All functions involved in BF channel polarization can be determined and studied versus the unbalance factor t, where fading is located on $B_o(\mathcal{C},\gamma)$. The mutual information of the MRC channel is $\mathcal{I}(p_1 \circledast p_2) = I((\alpha_1^2+\alpha_2^2)\gamma) = I(\chi(t)+\psi(t))$ and is plotted versus t in Fig. 5. For the SDC channel, $\mathcal{I}(p_1 \boxtimes p_2) = 1 - \mathcal{I}(p_1 \circledast p_2)$ is also shown on Fig. 5. Perusal of these plots shows that MRC and SDC channels are already close to an extremal channel, mainly in the critical unbalanced regime for large t. At the first splitting step for MRC and SDC, we get $\Delta_1(I) = I(W^+) - I(W^-) = 2(I(2\chi(t) + 2\psi(t)) - I(\chi(t)+\psi(t)))$ and $\Delta_2(I) = 2(\mathcal{I}((p_1 \boxtimes p_2)^{\circledast 2}) - \mathcal{I}(p_1 \boxtimes p_2))$ respectively. These functions are different but they have very close numerical values as shown in Fig. 5. In the unbalanced regime, we also find a good behavior of these Δ functions, i.e. smaller Δ implies that mutual information converges towards 0 or 1. The uniform multiplexer has a unique Δ function because it has a single polarization process. For the uniform multiplexer, we can write

$$\Delta(I) = I((\alpha_1^2+\alpha_2^2)\gamma) + \frac{1}{2}I(2\alpha_1^2\gamma) + \frac{1}{2}I(2\alpha_2^2\gamma) - 1.$$

The above expression corresponds to the curve in the middle of Fig. 5, it is seen that it is very close to $\frac{1}{2}$ for all t. It should be clear that the uniformly multiplexed BF channel is not naturally pre-polarized like the diagonally or the horizontally multiplexed one. For horizontal multiplexing, a behavior similar to the diagonal is illustrated in Fig. 6.

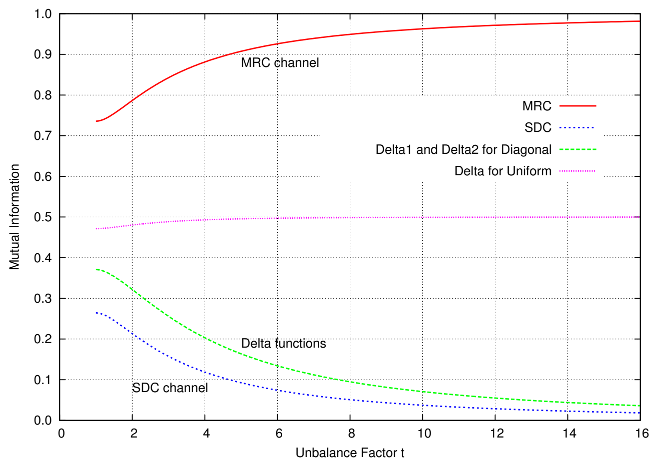

Figure 5. Functions $\mathcal{I}(p_1 \circledast p_2)$ (MRC), $\mathcal{I}(p_1 \boxtimes p_2)$ (SDC), $\Delta_1(I)$, and $\Delta_2(I)$ encountered in polarization with diagonal multiplexing.

After n steps, i.e., with a polar code of length $N = 2^n$, let I_i be the mutual information of channel i. For a given fading point at SNR γ, the sequence of mutual information values sorted in increasing order will be denoted by $\tilde{I}_i$, $i = 1 \ldots N$.

Definition 4: The MIO region for a length-N polar code is $R_o(\mathcal{A},\gamma) = \left\{(\alpha_1,\alpha_2) : \sum_{i=N/2+1}^{N} \tilde{I}_i \leq \frac{1}{2} - \frac{1}{N}\right\}$. The MIO boundary is $B_o(\mathcal{A},\gamma) = \left\{(\alpha_1,\alpha_2) : \sum_{i=N/2+1}^{N} \tilde{I}_i = \frac{1}{2} - \frac{1}{N}\right\}$

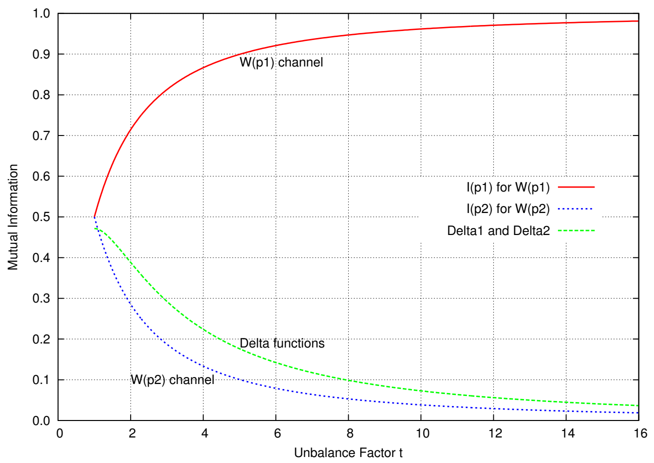

Figure 6. Functions $\mathcal{I}(p_1)$ (channel $W(p_1)$), $\mathcal{I}(p_2)$ (channel $W(p_2)$)), $\Delta_1(I)$, and $\Delta_2(I)$ encountered in polarization with horizontal multiplexing.

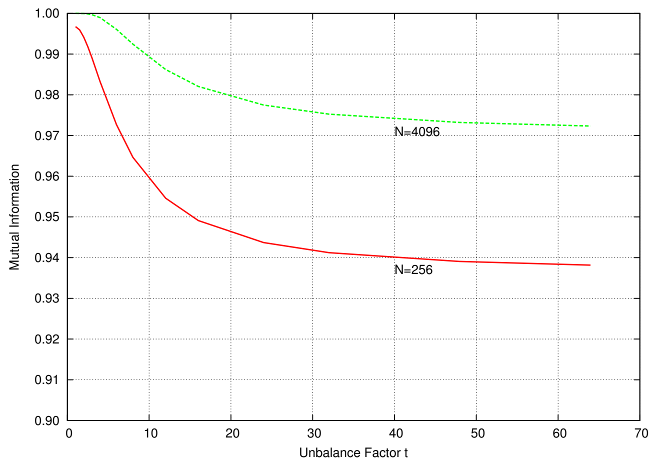

Figure 7. Uniform multiplexer. Mutual information sum for the best $N/2$ channels, $2/N \times \sum_{i=N/2+1}^{N} \tilde{I}_i$, versus the unbalance factor t.

(The symbol $\mathcal{A}$ stands for Arıkan). The rate $\frac{1}{2}-\frac{1}{N}$ corresponds to the loss of 1 bit out of $N/2$. The MIO probability for a polar code becomes $P_o(\mathcal{A},\gamma) = \int_{R_o(\mathcal{A},\gamma)} p(\alpha_1)p(\alpha_2)d\alpha_1 d\alpha_2$.

Proposition 4: (Sufficient Condition) For large N, if there exists a finite $g \in \mathbb{R}^+$ such that $\sum_{i=N/2+1}^{N} \tilde{I}_i \geq \frac{1}{2}-\frac{1}{N}$ for all $(\alpha_1,\alpha_2) \in B_o(\mathcal{C}, g\gamma)$, then $P_o(\mathcal{A},\gamma) \leq P_o(\mathcal{C}, g\gamma)$. In other words, the polar code has full diversity and its coding gain is at a gap of at most $10\log_{10}(g)$dB from the BF channel MIO probability.

Proof: Since the polar code is achieving a rate higher than $\frac{1}{2}-\frac{1}{N}$ on the outage boundary $B_o(\mathcal{C}, g\gamma)$, then we can write that $B_o(\mathcal{C}, g\gamma) \not\subset R_o(\mathcal{A},\gamma)$, which leads to $R_o(\mathcal{A},\gamma) \subset R_o(\mathcal{C}, g\gamma)$, then after integration over the fading distribution we get $P_o(\mathcal{A},\gamma) \leq P_o(\mathcal{C}, g\gamma)$. *QED.*

In practice, the sufficient condition in Prop. 4 can be easily verified with the help of the parametric description of the outage boundary from Prop. 3. Some relevant numerical results are given in Table I. The gap is infinite for uniform multiplexing at finite N. Diagonal multiplexing slightly outperforms horizontal multiplexing.

Length N	Diagonal	Horizontal	Uniform
256	1.65 dB	1.90 dB	> 6 dB
4096	1.40 dB	1.60 dB	> 6 dB
Full diversity ($N < \infty$)	Yes	Yes	No
Full diversity ($N = \infty$)	Yes	Yes	Yes

Table I
GAP BETWEEN MUTUAL INFORMATION OUTAGE PROBABILITIES $P_o(\mathcal{A})$ AND $P_o(\mathcal{C})$, AS DEFINED IN PROP. 4, FOR $\mathcal{M}_d$, $\mathcal{M}_h$, AND $\mathcal{M}_u$.

VI. CONCLUSIONS AND FUTURE WORK

Answering the three questions raised in the introduction, we were able to describe how block-fading channel polarization operates in conjunction with diagonal, horizontal, and uniform multiplexers (the question whether a less trivial multiplexer may exhibit a better performance at finite code lengths is left to future work). Diagonal and horizontal multiplexers are found to have full diversity at finite code lengths. The failure of the uniform multiplexer to achieve full diversity cannot be deduced from Prop. 4, which involves only a sufficient condition, but can be observed from Fig. 7, where the achieved rate is plotted versus the unbalance factor t. Finally, all multiplexers do polarize for any fading-gain pair at infinite length, i.e., the channel outage probability for rate-$1/2$ can be achieved with a zero signal-to-noise ratio gap.

REFERENCES

[1] C. E. Shannon, "A mathematical theory of communication," *Bell System Techn. J.*, vol. 27, pp. 379–423 and 623–656, Jul. and Oct. 1948.

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

[3] J.J. Boutros, A. Guillén i Fàbregas, E. Biglieri, and G. Zémor, "Low-Density Parity-Check Codes for Nonergodic Block-Fading Channels," *IEEE Trans. Inform. Theory*, vol. 56, no. 9, pp. 4286-4300, Sept. 2010.

[4] J.J. Boutros, "Diversity and coding gain evolution in graph codes," *Information Theory and Appl.*, pp. 34-43, UCSD, San Diego, Feb. 2009.

[5] T.J. Richardson and R.L. Urbanke, *Modern Coding Theory*, Cambridge University Press, 2008.

[6] E. Biglieri, *Coding for Wireless Channels*, New York, Springer, 2005.

[7] D.N.C. Tse and P. Viswanath, *Fundamentals of Wireless Communication*, Cambridge University Press, 2005.

[8] S. Kudekar, T. Richardson, and R. Urbanke, "Spatially Coupled Ensembles Universally Achieve Capacity under Belief Propagation," arXiv:1201.2999v1, cs.IT, 14 January 2012.

[9] A.J. Felström and K.S. Zigangirov, "Time-varying periodic convolutional codes with low-density parity-check matrix," *IEEE Trans. Inform. Theory*, vol. 45, no. 5, pp. 2181-2190, Sept. 1999.

[10] I. Tal and A. Vardy, "How to construct polar codes," arXiv:1105.6164v2, cs.IT, May 2011.

[11] J.J. Boutros, E. Calvanese Strinati, and A. Guillén i Fàbregas, "Turbo code design for block fading channels," *Allerton's Conference*, Monticello, Illinois, Sept 2004.

[12] J.J. Boutros, G.M. Kraidy, and N. Gresset, "Near outage limit space-time coding for MIMO channels," *Inaugural ITA workshop*, UCSD, San Diego, California, Feb. 2006.

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

[14] J.J. Boutros and E. Biglieri, "Convergence acceleration of polar code construction," in preparation.