# Precoding for Outage Probability Minimization on Block Fading Channels

Dieter Duyck, Joseph J. Boutros, and Marc Moeneclaey

arXiv:1103.5348v2 [cs.IT] 12 May 2011

**Abstract**

The outage probability limit is a fundamental and achievable lower bound on the word error rate of coded communication systems affected by fading. This limit is mainly determined by two parameters: the diversity order and the coding gain. With linear precoding, the maximum achievable coding rate yielding full diversity on a block fading channel can exceed the upper limit given by the standard Singleton bound. However, the effect of precoding on the coding gain is not well known, mainly due to the complicated expression of the outage probability. This paper establishes simple upper bounds on the outage probability, from which the optimal precoding matrices minimizing these upper bounds can be determined. For discrete alphabets, it is shown that the combination of constellation expansion and precoding is sufficient to closely approach the minimum possible outage achieved by an i.i.d. Gaussian input alphabet, thus essentially maximizing the coding gain.

## I. INTRODUCTION

In many applications, such as frequency-hopping (GSM, EDGE), time-interleaving (DVB-T), OFDMA (WiMax and LTE), H-ARQ with cross-packet coding [16], [8], and cooperative communications [27], [28], [19], [7], the channel can be modelled as a flat block fading (BF) channel [2], where the fading gain is piecewise constant over the duration of a transmitted packet.

Dieter Duyck and Marc Moeneclaey are with the Department of Telecommunications and Information processing, Ghent University, St-Pietersnieuwstraat 41, B-9000 Gent, Belgium, {dduyck,mm}@telin.ugent.be. Joseph J. Boutros is with Texas A&M University at Qatar, PO Box 23874 Doha, Qatar, boutros@tamu.edu

Due to motion of the transmitter, receiver or objects between the transmitter and receiver, the fading gains vary from one packet to the next and are considered unknown at the transmitter side. The fraction of codewords where decoding fails to wipe out all errors is referred to as the average word error rate (WER). When displayed on a log-scale versus the average signal-to-noise ratio (SNR) in decibel, the high-SNR slope of the WER is called the diversity order. Since the diversity order determines how fast the error rate decreases with the SNR, it is then a key parameter of the communication system.

The Singleton bound limits the coding rate $R_c$ of standard coded modulations achieving full diversity [17], [21], [11]. Because of their reduced spectral efficiency, full-diversity error-correcting codes are bandwidth consuming. Fortunately, the spectral efficiency can be increased beyond the Singleton bound by using component interleaving combined with linear precoding [13], [14], [12], which modifies the marginal distribution of the components of a multidimensional constellation at the channel input. When using linear precoding with component interleaving, full diversity can be achieved without error-correcting code. Therefore, it has been almost exclusively studied for uncoded schemes (see [1], [2] and references therein, e.g. [3]), except for few papers [13], [18] investigating coded transmission schemes and recent work on pairing subchannels when channel state information at the transmitter (CSIT) is available [22], [23]. However, besides yielding full diversity with a higher rate, linear precoding can also improve the coding gain of coded systems. So the effect of linear precoding on the diversity order is well understood, but there is no sufficient insight in the effect of linear precoding on coding gain of coded systems.

Before studying the optimization of the WER for practical schemes with linear precoding, it is important to understand the influence of linear precoding on the performance limits of the communication channel. The outage probability limit is an achievable lower bound on the WER of coded systems in the limit of large block length [2], [24] and is given by [29, section 5.4],

$$P_{out}(\gamma, P, R) = \Pr\big(I(\boldsymbol{\alpha}, \gamma, P) < R\big),$$

where $I(\boldsymbol{\alpha}, \gamma, P)$ is the instantaneous mutual information between transmitted symbols and received symbols as a function of the set of fading gains $\boldsymbol{\alpha}$ observed during the transmission of one codeword, the average SNR $\gamma$, and the precoding matrix $P$. By choosing a well designed precoding matrix $P$, the outage probability can be minimized. But even in the simple case

of an independent and identically distributed (i.i.d.) Gaussian input alphabet of the channel, a closed form expression for the outage probability is not known yet. Therefore, only a brute force optimization can be applied to optimize the outage probability. Such an optimization is often intractable when the number of fading gains per codeword is larger than two and/or large constellations (e.g. 16 points) are used. A simple approach could be designing $P$ such that the mean of $I(\boldsymbol{\alpha}, \gamma, P)$ over the fading distribution is maximized, in the hope that the area under the left tail of its probability density function (pdf) would be minimized. However, the ergodic mutual information $\mathbb{E}_{\boldsymbol{\alpha}}[I(\boldsymbol{\alpha}, \gamma, P)]$ contains no information on the diversity order, and, due to the limited spectral efficiency of finite discrete input alphabets, for increasing $\gamma$ rapidly converges to a maximum that does not depend on the precoding matrix. Therefore, the maximization of the ergodic mutual information fails to provide the optimum precoding matrix [9].

In this paper, we first study the effect of linear precoding of discrete input alphabets on the outage probability. A unitary $B \times B$ precoding matrix has $B^2$ degrees of freedom, so that the optimization of the outage probability is multivariate. In a brute force optimization, Monte Carlo simulations are required to take into account the distribution of all fading gains when calculating the outage probability. Here, the analysis uses a geometric approach that leads to upper bounds on the outage probability, which are easier to optimize, because it is no longer necessary to perform a Monte Carlo simulation based on the fading gain distribution. Therefore, with a minimal computational effort, the optimal precoding matrix minimizing the upper bound on the outage probability limit can be determined. These results serve as a basis for the design of practical coded systems with linear precoding [6], which is discussed in final section before conclusions.

## II. SYSTEM MODEL

At the transmitter output, a packet is represented as a real or complex column vector $\boldsymbol{\chi} = [\boldsymbol{\chi}(1)^T, \ldots, \boldsymbol{\chi}(B)^T]^T$ of dimension $N$, consisting of $B$ blocks that each contain $N/B$ symbols, where $(.)^T$ indicates transposition; the b-th block of the packet is $\boldsymbol{\chi}(b) = [\chi(b)_1, \ldots, \chi(b)_{\frac{N}{B}}]^T$ with $\mathbb{E}[|\chi(b)_n|^2] = 1$. The channel is memoryless with additive white Gaussian noise and multiplicative real fading. The fading coefficients are only known at the decoder side. The

received signal vector $\boldsymbol{\mu}(b)$ corresponding to the transmitted block $\boldsymbol{\chi}(b)$ is

$$\boldsymbol{\mu}(b) = \alpha_b \boldsymbol{\chi}(b) + \boldsymbol{\omega}(b), \quad b = 1, \ldots, B. \tag{1}$$

The fading coefficients $\{\alpha_b, b = 1, \ldots, B\}$ are i.i.d. In the numerical results, we consider the fading coefficients to be Rayleigh distributed, with $\mathbb{E}[\alpha_b^2] = 1$, but the analysis in this paper is valid for arbitrary distributions. The noise vector $\boldsymbol{\omega}(b)$ consists of $N/B$ independent noise samples which are complex Gaussian distributed, $\omega(b)_n \sim \mathcal{CN}(0, 2\sigma^2)$. The average signal-to-noise ratio is $\gamma = \frac{1}{2\sigma^2}$.

The transmitted vector $\boldsymbol{\chi}$ is obtained from the information bits through a sequence of operations. Assuming a binary encoder with coding rate $R_c$, a packet of $K$ information bits is encoded into $K/R_c$ coded bits. The binary codeword is split into $K/(mR_c)$ strings each containing $m$ bits. In a standard coded communications system, the components of the transmitted vector $\boldsymbol{\chi}(b)$ are obtained by directly mapping each string of $m$ coded bits to one of $M = 2^m$ points belonging to a 1-dimensional real or complex space; the corresponding spectral efficiency $R$ in bits per channel use (bpcu) is given by $R = mR_c$. When using precoding combined with component interleaving, each string of $m$ bits is mapped to one of $M = 2^m$ points belonging to a B-dimensional real or complex space; the corresponding B-dimensional M-point constellation $\Omega_z$ is denoted $M$-$\mathcal{R}^B$ or $M$-$\mathcal{C}^B$, respectively. Denoting as $\mathbf{z}(n) = [z(n)_1, \ldots, z(n)_B]^T$ the B-dimensional vector that results from mapping the n-th string of $m$ coded bits, the linear precoding involves the computation

$$\mathbf{x}(n) = P\mathbf{z}(n), \quad n = 1, \ldots, \frac{N}{B} \tag{2}$$

where $P$ is a non-singular precoding matrix of dimension $B \times B$. The precoder output vectors $\mathbf{x}(n) = [x(n)_1, \ldots, x(n)_B]$ belong to a B-dimensional M-point constellation $\Omega_x$ which results from a linear transformation (through $P$) of $\Omega_z$. Finally, component interleaving is applied: the n-th component of the b-th block of the transmitted vector $\boldsymbol{\chi}$ equals the b-th component of the n-th precoder output vector, i.e., $\chi(b)_n = x(n)_b$ (Fig. 1). Hence, the $B$ components of $\mathbf{x}(n)$ experience independent fading when transmitted over the BF channel. Taking into account that $K$ information bits are transformed into a transmitted vector $\boldsymbol{\chi}$ with $N = KB/(mR_c)$ components, the overall spectral efficiency is $R = R_c\frac{m}{B}$ bpcu. Note that there are several ways to achieve a given spectral efficiency $R$. For example, $R = 0.9$ bpcu for $B = 2$ can be achieved
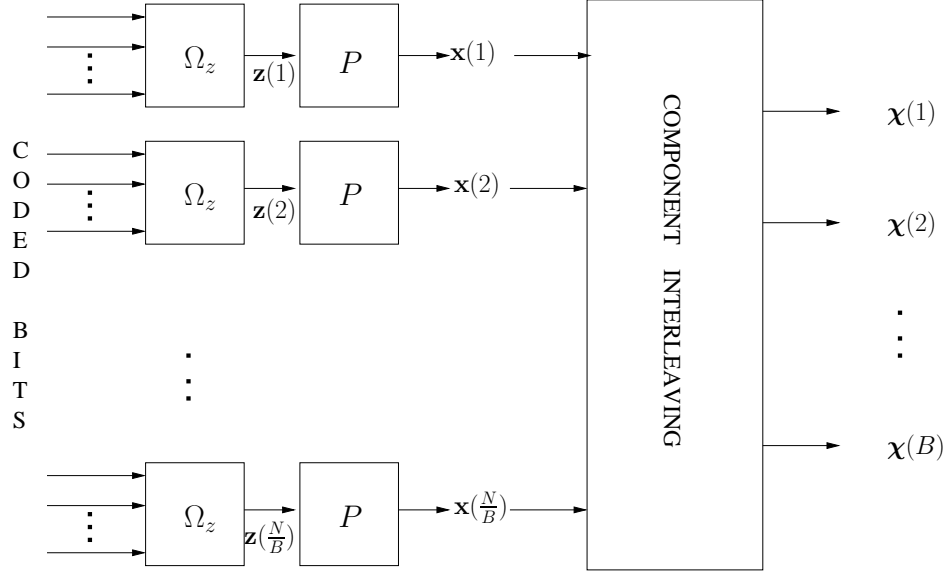
Fig. 1. The coded bits are mapped to multidimensional symbols $\mathbf{z}(n)$ in a constellation $\Omega_z$. Each symbol $\mathbf{z}(n)$ is transformed to $\mathbf{x}(n)$, where the different components are placed in the corresponding blocks in $\boldsymbol{\chi}$.

by a coded communication system with precoding and component interleaving ($R = R_c \frac{m}{B}$) by choosing $m = 3$ and $R_c = 0.6$, whereas a standard communication system ($R = R_c m$) achieves $R = 0.9$ bpcu for $m = 1$ and $R_c = 0.9$. Also other combinations of $m$ and $R_c$ are possible. Note that a standard communication system with $R = m' R_c$ can be viewed as a special case of a multidimensional modulation, where $\Omega_z$ is a Cartesian product of $B$ identical constellations belonging to $2^{m'}$-$\mathcal{R}^1$ or $2^{m'}$-$\mathcal{C}^1$ and $P = I$, where $I$ is the identity matrix, so that $\Omega_x = \Omega_z$; these B-dimensional constellations contain $M = 2^{m' B}$ points.

We can reformulate Eq. (1) in terms of $\mathbf{x}(n)$ as

$$\mathbf{y}(n) = \boldsymbol{\alpha} \cdot \mathbf{x}(n) + \mathbf{w}(n), \quad n = 1, \ldots, \frac{N}{B}, \tag{3}$$

where $y(n)_b = \mu(b)_n$, $(\boldsymbol{\alpha})_b = \alpha_b$, $w(n)_b = \omega(b)_n$ and $\boldsymbol{\alpha} \cdot \mathbf{x}(n)$ denotes component-wise multiplication: $(\boldsymbol{\alpha} \cdot \mathbf{x}(n))_b = \alpha_b x(n)_b$.

As precoding allows to increase the spectral efficiency associated with full diversity, this technique has been extensively studied in previous works, but almost exclusively for *uncoded* schemes [1], [2], [3]. Here, we study this system for *coded* schemes and we choose the precoding matrix $P$ and the constellation $\Omega_z$ minimizing the outage probability.

A precoding matrix $P$ that is unitary is an obvious choice because it does not decrease the

capacity of a Gaussian channel. In this paper, we restrict our study to real precoding matrices, hence $P$ is orthogonal. When $B = 2$, $P$ is a rotation matrix where the rotation angle $\theta$ is the degree of freedom:

$$P = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}. \tag{4}$$

However, rotation matrices are difficult to construct for higher dimensions. In Sec. III, it will be shown that for $B > 2$ it is sufficient to consider orthogonal circulant precoding matrices. We denote its first row as $(p_0, ..., p_{B-1})$. The second row is a cyclic shift to the right of the first row, and so on. Because the columns of the $B \times B$ Fourier matrix $F$ are the eigenvectors of any circulant matrix, we can construct $P$ as follows:

$$P = F\Lambda F^H, \tag{5}$$

where $(F)_{m,n} = \frac{1}{\sqrt{B}} \exp\left(\frac{-2j\pi mn}{B}\right)$, $m, n \in \{0, \ldots, B-1\}$, and $\Lambda$ is a diagonal matrix containing the eigenvalues of $P$. The condition for $P$ being orthogonal is $\Lambda^H \Lambda = I_B$, or the $B$ eigenvalues of $P$ must have a squared magnitude of 1. It is easy to find that

$$\lambda_n = \sum_{l=0}^{B-1} p_l \exp\left(\frac{-j2\pi nl}{B}\right). \tag{6}$$

Now, it follows that

$$p_l = \frac{1}{B} \sum_{m=0}^{B-1} \lambda_m \exp\left(\frac{j2\pi ml}{B}\right). \tag{7}$$

As the eigenvalues must have a magnitude of 1, we have $\lambda_n = \exp(j\theta_n)$. In order to obtain a real-valued $P$, we take $\lambda_0$ real (i.e., $\lambda_0 = 1$ or $\lambda_0 = -1$) and $\lambda_{B-n} = (\lambda_n)^*$ (i.e., $\theta_{B-n} = -\theta_n$) for $n = 1, \ldots, B - 1$. When $B$ is even, this implies $\lambda_{B/2} = 1$ or $\lambda_{B/2} = -1$. For $B > 2$, $P$ is determined by $\lfloor (B-1)/2) \rfloor$ continuous parameters that can be optimized.

Note that for $B = 3$, $P$ constructed as above, with $\lambda_0 = 1$ and $\lambda_1 = \exp(j\theta_1)$, corresponds to a 3-dimensional rotation with angle $\theta_1$ around the fixed axis $\frac{1}{\sqrt{3}}(1, 1, 1)$,

$$P = \frac{1}{3} \begin{bmatrix} 1 + 2k & 1 - k - \sqrt{3}l & 1 - k + \sqrt{3}l \\ 1 - k + \sqrt{3}l & 1 + 2k & 1 - k - \sqrt{3}l \\ 1 - k - \sqrt{3}l & 1 - k + \sqrt{3}l & 1 + 2k \end{bmatrix}, \tag{8}$$
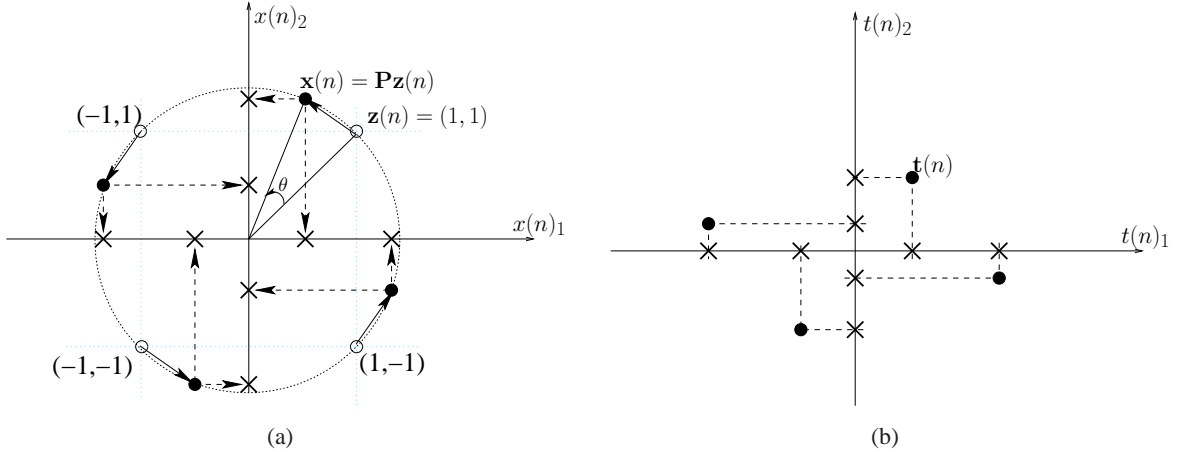
Fig. 2. Displaying the rotation at the transmitter (a) and at the receiver (without noise) (b) for $B = 2$. The empty (filled) circles represent $\Omega_z$ ($\Omega_x$). The components of $\mathbf{t}(n)$ are obtained by scaling the components $\mathbf{x}(n)$ by their respective fading gain (here $\alpha_2 < \alpha_1$), as expressed by the component-wise multiplication $\boldsymbol{\alpha} \cdot \mathbf{x}(n)$. The crosses on the coordinate axes are the transmitted and received vector components, respectively.

where $k = \cos(\theta_1)$ and $l = \sin(\theta_1)$.

Fig. 2 illustrates the effect of a rotation for $B = 2$ when a $4$-$\mathcal{R}^2$ constellation is used as $\Omega_z$. The transmitted components are affected by their corresponding fading gain, which is expressed by the component-wise multiplication $\mathbf{t}(n) = \boldsymbol{\alpha} \cdot \mathbf{x}(n)$, which is shown at the right side in Fig. 2. We say that $\mathbf{t}(n)$ belongs to the *faded* constellation $\Omega_t$. The point $\mathbf{t}(n) = [t(n)_1, ..., t(n)_B]$ is shown for a particular fading point $\boldsymbol{\alpha}$. When $\alpha_1 \neq \alpha_2$, the constellation $\Omega_t$ can be interpreted as a distorted QPSK constellation (i.e., a constellation in which both components do not have the same magnitude).

Considering $\mathbf{t}(n)$, an equivalent channel model can be formulated:

$$\mathbf{y}(n) = \mathbf{t}(n) + \mathbf{w}(n), \quad n = 1, \ldots, \frac{N}{B}, \tag{9}$$

which yields more insight and will be useful in the proofs of propositions of this paper. This system model represents a Gaussian vector channel with input $\mathbf{t}(n)$. This means that for a particular fading point, the block fading channel can be interpreted as a virtual[1] Gaussian channel, with a discrete input alphabet $\Omega_t$.

---

[1]We use the term *virtual* because the fading gains of the actual channel are incorporated in the constellation $\Omega_t$.

## III. ANALYSIS OF OUTAGE PROBABILITY IN THE FADING SPACE

For the remainder of the paper, we will drop the index $n$ in the vectors $\mathbf{z}(n)$, $\mathbf{x}(n)$, $\mathbf{t}(n)$, $\mathbf{y}(n)$ and $\mathbf{w}(n)$, as the time index is not important when considering mutual information. We write random variables using upper case letters corresponding to the lower case letters used for their realizations. The mutual information $I(\boldsymbol{\alpha}, \gamma, P)$ at a certain fading point $\boldsymbol{\alpha}$ between the transmitted B-dimensional symbol $\mathbf{x}$ (uniformly distributed over $\Omega_x$) and the corresponding received vector $\mathbf{y}$ is given by [12], [10]

$$I(\boldsymbol{\alpha}, \gamma, P) = \frac{1}{B} I(\mathbf{X}; \mathbf{Y} | \boldsymbol{\alpha}, \gamma), \tag{10}$$

where the scaling factor $\frac{1}{B}$ is added because the $B$ blocks in the channel timeshare a time-interval [5, Section 9.4], [29, Section 5.4.4]. The mutual information $I(\mathbf{X}; \mathbf{Y} | \boldsymbol{\alpha}, \gamma)$ is

$$I(\mathbf{X}; \mathbf{Y} | \boldsymbol{\alpha}, \gamma) = m - 2^{-m} \sum_{\mathbf{x} \in \Omega_x} \mathbb{E}_{\mathbf{y} | \mathbf{x}} \left[ \log_2 \left( \sum_{\mathbf{x}' \in \Omega_x} \exp \left[ \frac{d^2(\mathbf{y}, \boldsymbol{\alpha} \cdot \mathbf{x}) - d^2(\mathbf{y}, \boldsymbol{\alpha} \cdot \mathbf{x}')}{2\sigma^2} \right] \right) \right], \tag{11}$$

where $d^2(\mathbf{v}, \mathbf{u}) = \sum_{b=1}^{B} |v_b - u_b|^2$. However, more insight can be gained when considering $\mathbf{t}$. From Eq. (9), it is clear that for a certain fading point, the mutual information of this virtual channel, with input $\Omega_t$, is the same as the mutual information of the actual channel, with input $\Omega_x$,

$$I(\boldsymbol{\alpha}, \gamma, P) = \frac{1}{B} I(\mathbf{T}; \mathbf{Y} | \boldsymbol{\alpha}, \gamma). \tag{12}$$

Hence, the fading point $\boldsymbol{\alpha}$ maximizing (minimizing) the mutual information $I(\mathbf{X}; \mathbf{Y} | \boldsymbol{\alpha}, \gamma)$ corresponds to the fading point that distorts constellation $\Omega_t$ in the best (worst) way at the input of a Gaussian vector channel. This interpretation allows to exploit the many results from literature on the mutual information of Gaussian channels. Therefore, Eq. (12) will be useful in the following.

The outage probability is the probability that the instantaneous mutual information is less than the transmitted information bitrate $R$ [29, section 5.4],

$$P_{out}(\gamma, P, R) = \Pr\big(I(\boldsymbol{\alpha}, \gamma, P) < R\big). \tag{13}$$

Our main goal is to find the precoding matrix $P$ that minimizes the outage probability $P_{out}(\gamma, P, R)$,

$$P_{opt} = \text{Arg} \min_P P_{out}(\gamma, P, R).$$

(a) Examples of an outage boundary when $\Omega_z = 4 \cdot \mathcal{R}^2$ and the rotation angle is $\theta = 0, 10, 27$ degrees. The region $V_o$ is coloured red for $\theta = 27$ degrees.

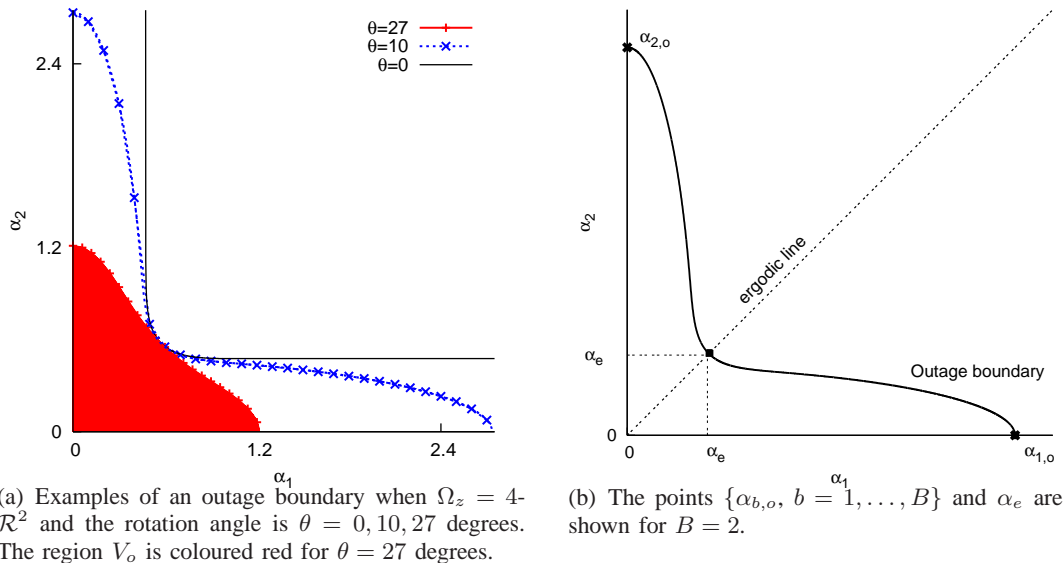(b) The points $\{\alpha_{b,o}, \ b = 1, \ldots, B\}$ and $\alpha_e$ are shown for $B = 2$.

Fig. 3. The outage boundary limits the region $V_o$ in the fading space which corresponds to an information theoretic outage event. The information rate is $R = 0.9$ bpcu. The average SNR is fixed to $\gamma = 8$dB.

A closed form expression for $I(\boldsymbol{\alpha}, \gamma, P)$ does not exist and Eq. (11) is difficult to analyze because of the presence of mathematical expectations. Therefore, in the following two sections, we will develop simple bounds on the outage probability and will approximate $P_{opt}$ by the optimal precoding matrix that minimizes the bounds on the outage probability.

But first we will introduce a framework in this section that allows to gain insight on the meaning of the outage probability. The considered framework is the fading space [4], which is a subset of the B-dimensional Euclidean space, $\mathbb{R}^{+B}$, of the real positive fading gains. The outage probability equals the probability that the fading point $\boldsymbol{\alpha}$ belongs to a region $V_o$, which is such that $I(\boldsymbol{\alpha}, \gamma, \theta) < R$ for all $\boldsymbol{\alpha}$ in $V_o$ (Fig. 3(a)):

$$P_{out}(\gamma, P, R) = \int_{\boldsymbol{\alpha} \in V_o} p(\boldsymbol{\alpha}) d\boldsymbol{\alpha}, \tag{14}$$

where $p(\boldsymbol{\alpha})$ is the joint pdf of the fading gains $\alpha_1, \ldots, \alpha_B$. We say that the region $V_o$ is limited by an *outage boundary* $B_o(\gamma, P, R)$ (see Fig. 3(a) for $B = 2$), defined by

$$I(\boldsymbol{\alpha}, \gamma, P) = R, \quad \forall \ \boldsymbol{\alpha} \in B_o(\gamma, P, R).$$

In general, we can consider a certain boundary in the fading space limiting a region $V$. This

boundary is denoted by $B(V)$. The outage boundary $B_o(\gamma, P, R)$ corresponds to a region $V_o$, as defined in (14), so that the outage boundary is $B_o(\gamma, P, R) = B(V_o)$.

*Definition 1:* We define $\alpha_{b,o}$ by the magnitude of the intersection between the outage boundary and the axis $\alpha_b$. More precisely, $I\left(\boldsymbol{\alpha}|_{\alpha_i=0, i\neq b, \alpha_b=\alpha_{b,o}}, \gamma, P\right) = R$. By convention, $\alpha_{b,o} = +\infty$ if the axis $\alpha_b$ is an asymptote for the outage boundary.

*Definition 2:* We define $\alpha_e$ as the value of the components of the intersection between the outage boundary and the line $\alpha_1 = \ldots = \alpha_B$ (also known as the *ergodic line*). More precisely, $I\left(\boldsymbol{\alpha}|_{\alpha_i=\alpha_e, i\in[1,\ldots,B]}, \gamma, P\right) = R$.

The defined points are illustrated in Fig. 3(b) for $B = 2$. In the remainder of the paper, we denote the points $\boldsymbol{\alpha}|_{\alpha_i=0, i\neq b, \alpha_b=\alpha_{b,o}}$ by $\boldsymbol{\alpha}_{b,o}$ and $\boldsymbol{\alpha}|_{\alpha_i=\alpha_e, i\in[1,\ldots,B]}$ by $\boldsymbol{\alpha}_e$.

*Proposition 1:* On a BF channel with $B = 2$, with a discrete input alphabet and with linear precoding, the magnitudes $\{\alpha_{b,o}, \ b = 1, 2\}$ are equal if the constellation is invariant under a rotation of 90 degrees and the precoding matrix is orthogonal.

On a BF channel with $B > 2$, with a discrete input alphabet and with linear precoding, the magnitudes $\{\alpha_{b,o}, \ b = 1, \ldots, B\}$ are equal if the constellation is invariant under a cyclic shift of the components of the points of the constellation and the precoding matrix is an orthogonal circulant matrix.

*Proof:* See Appendix A. ∎

Notice that the condition of Prop. 1 is a sufficient condition and not a necessary condition. In the remainder of this paper, it is assumed that the constellation used at the transmitter fulfils Prop. 1. The magnitudes $\{\alpha_{b,o}, b = 1, \ldots, B\}$ will then simply be denoted by $\alpha_o$. This also means that the projection of $\Omega_x$ on either coordinate axes yields the same set of points, which we denote by $\mathcal{S}_p$, where $p$ stands for projection.

Note that a multidimensional constellation that fulfils Prop. 1 (i.e., its projection on each coordinate axis yields the same set of points, see Appendix A) has an interesting property. For these constellations, the function $I(X_b; Y_b|\alpha_b = \alpha, \gamma)$, which is the mutual information of a point-to-point channel with fading coefficient $\alpha_b = \alpha$, average SNR $\gamma$ and with discrete input $X_b$, does not depend on $b$. As a consequence, we will represent this mutual information by $I_{\mathcal{S}_p}(\alpha^2\gamma, P)$.

*Definition 3:* Boundary $B_1 = B(V_1)$ is said to outer bound outage boundary $B_2 = B(V_2)$ if $V_2 \subseteq V_1$. The opposite holds for inner bounding.

The outage boundary in the fading space has a simple but interesting property: if an outage boundary *outer bounds (inner bounds)* another outage boundary, than its corresponding outage probability is larger (smaller) (see Eq. (14)). For example, consider the input alphabet $\mathbf{Z} \sim \mathcal{N}(\mathbf{0}, I)$, which we denote as an i.i.d. Gaussian input alphabet[2], where $\mathbf{0}$ is a column vector of zeros and $I$ is the identity matrix. The outage boundary corresponding to i.i.d. Gaussian inputs, denoted by $B_o(gauss)$, inner bounds the outage boundary corresponding to a discrete input alphabet, denoted by $B_o(discrete)$ [5]. Therefore, the outage probability corresponding to i.i.d. Gaussian inputs is a lower bound on the outage probability corresponding to a discrete input alphabet. Consequently, by minimizing the outage probability corresponding to a discrete input alphabet, $B_o(discrete)$ can at most approach $B_o(gauss)$ (see Fig. 4).
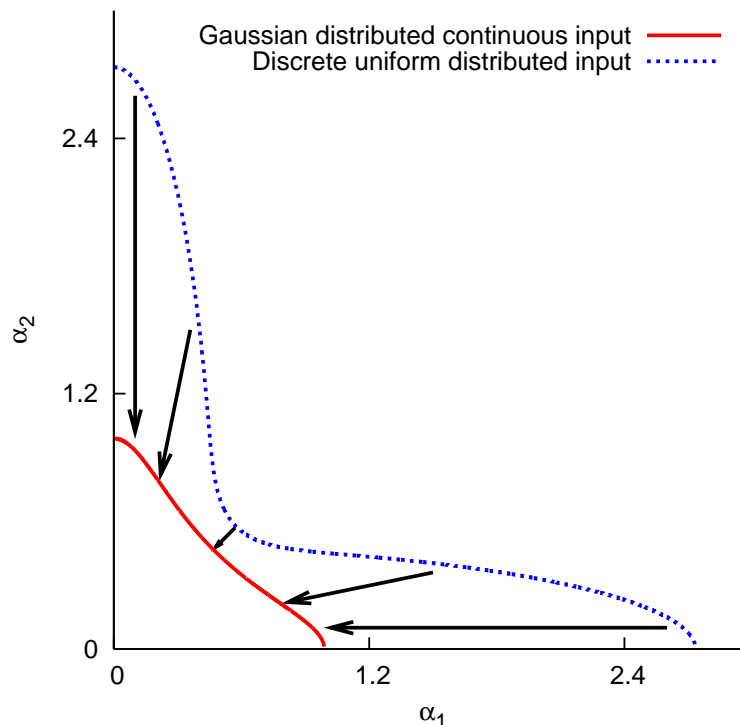


Fig. 4. The outage boundary $B_o(discrete)$ is inner bounded by $B_o(gauss)$. Optimizing the precoding matrix can at most make $B_o(discrete)$ approach $B_o(gauss)$ as illustrated by the arrows.

In the following two sections, we will determine boundaries with simple shapes outer (inner) bounding $B_o(discrete)$, which are then much easier to optimize. For the outer boundary, this

[2]We mainly refer to real i.i.d. Gaussian input alphabets in this paper. In Sec. VII-A2, we extend this to complex i.i.d. Gaussian input alphabet.

can be done by determining a surface in the fading space, $U(\boldsymbol{\alpha}) = 0$, satisfying

$$I(\boldsymbol{\alpha}, \gamma, P) \geq R, \text{ for all } \boldsymbol{\alpha} \text{ satisfying } U(\boldsymbol{\alpha}) = 0. \tag{15}$$

For the inner boundary, the greater than or equal sign is replaced by a smaller than or equal sign. For example, for $B = 2$, we will prove that a circular arc touching $B_o(discrete)$ at $\alpha_1 = \alpha_o$ (on the horizontal axis) and $\alpha_2 = \alpha_o$ (on the vertical axis) satisfies Eq. (15).

## IV. BOUNDS ON THE OUTAGE PROBABILITY WITHOUT LINEAR PRECODING

As a first step, we will establish upper and lower bounds on the outage probability of a communication system without linear precoding. This will set the stage for the bounds with linear precoding in Sec. V.

We will prove that the outage boundary is outer bounded by a $B$-hypersphere touching the outage boundary on the axes of the fading space, hence with radius $\alpha_o$. A $B$-hypersphere $U(\boldsymbol{\alpha}) = 0$ is a generalization of a sphere to $B$ dimensions,

$$U(\boldsymbol{\alpha}) = \sum_{b=1}^{B} \alpha_b^2 - \alpha_o^2.$$

Note that this is only possible for constellations fulfilling the conditions of Prop. 1. For other constellations, the outer bounding $B$-hypersphere will have a radius of $\max_{b \in [1,...,B]} \alpha_o$ and will therefore be less tight.

Next, we will prove that the outage boundary is inner bounded by a $B$-hypersphere touching the outage boundary at $\boldsymbol{\alpha}_e$.

*Lemma 1:* On a BF channel with a discrete multidimensional input alphabet $\mathbf{X}$, the mutual information $I(\boldsymbol{\alpha}, \gamma, P)$ is upper bounded as follows

$$I(\boldsymbol{\alpha}, \gamma, P) \leq \frac{1}{B} \sum_{b=1}^{B} I_{\mathcal{S}_p}\left(\alpha_b^2 \gamma, P\right). \tag{16}$$

*Proof:* See Appendix B. ∎

A special case is the case without precoding ($P = I$) so that $\Omega_x = \Omega_z$. If in that case $\Omega_z$ is the Cartesian product of $\Omega_{z_1}, \ldots, \Omega_{z_B}$ [3], $\Omega_z = (\mathcal{S}_p)^B$, then all variables $X_b = Z_b$ are independent

---

[3] We denote the projection of $\Omega_z$ on the $b$-th coordinate axis as $\Omega_{z_b}$. For constellations fulfilling Prop. 1, $\Omega_{z_b} = \mathcal{S}_p$.

for all $b$, so that equality holds in Eq. (16). E.g., $\Omega_z = 4\text{-}\mathcal{R}^2$, as shown in Fig. 2, is the Cartesian product of two BPSK constellations, denoted as BPSK$\times$BPSK.

For an i.i.d. Gaussian input alphabet, the instantaneous mutual information with linear precoding is the same as the instantaneous mutual information without linear precoding, because $P\mathbf{z} = \mathbf{x} \sim \mathcal{N}(\mathbf{0}, I)$ when $P$ is orthogonal, so that precoding does not change the distribution of the input vector. Therefore, the instantaneous mutual information for an i.i.d. Gaussian input alphabet is

$$I\left(\boldsymbol{\alpha}, \gamma, P\right) = I\left(\boldsymbol{\alpha}, \gamma, P = I\right) = \frac{1}{B} \sum_{b=1}^{B} 0.5 \, \log_2(1 + 2\gamma\alpha_b^2). \tag{17}$$

*Proposition 2:* The outage boundary $B_o(gauss)$ of a BF channel with an i.i.d. Gaussian input alphabet is outer bounded by the hypersurface of the $B$-hypersphere $\alpha_1^2 + \alpha_2^2 + \ldots + \alpha_B^2 = \frac{4^{BR}-1}{2\gamma}$ and inner bounded by the hypersurface of the $B$-hypersphere $\alpha_1^2 + \alpha_2^2 + \ldots + \alpha_B^2 = \frac{4^R-1}{2\gamma}$.

*Proof:* In Appendix C, it is proved that for $\boldsymbol{\alpha}$ on a hypersurface:

- $I\left(\boldsymbol{\alpha}, \gamma\right)|_{\alpha_1 = \ldots = \alpha_B} \geq I\left(\boldsymbol{\alpha}, \gamma\right)$
- $I\left(\boldsymbol{\alpha}, \gamma\right)|_{\alpha_b = \alpha_o} \leq I\left(\boldsymbol{\alpha}, \gamma\right)$,

because the mutual information is a sum of concave functions of the instantaneous SNRs $\gamma\alpha_b^2$ (Eq. (17)). Calculating $\alpha_o$ and $\alpha_e$, according to Definitions 1 & 2, yields the radii of both $B$-hyperspheres. ∎

The inner boundary touches the outage boundary in the point $\boldsymbol{\alpha}_e$, so that it does not depend on $B$ (the entire codeword is affected by the same fading gain $\alpha_e$). The outer boundary touches the outage boundary in the points $\boldsymbol{\alpha}_{b,o}$, $\forall \, b$; its dependence on $B$ follows from the fact that $B - 1$ is equal to the number of erased channels.

*Proposition 3:* On a BF channel with a discrete input alphabet $\Omega_z$ that is a Cartesian product of one-dimensional constellations, the outage boundary $B_o(discrete)$ is outer bounded by the hypersurface of the $B$-hypersphere $\alpha_1^2 + \alpha_2^2 + \ldots + \alpha_B^2 = \alpha_o^2$ touching it at the axes of the fading space at $\boldsymbol{\alpha}_{b,o}$, $\forall \, b$. Also, $B_o(discrete)$ is inner bounded by the hypersurface of the $B$-hypersphere $\alpha_1^2 + \alpha_2^2 + \ldots + \alpha_B^2 = \alpha_e^2$ touching it at $\boldsymbol{\alpha}_e$.

*Proof:* Lemma 1 proved that the mutual information is upper bounded by $\frac{1}{B} \sum_{b=1}^{B} I_{\mathcal{S}_p}\left(\alpha_b^2 \gamma, P\right)$, where this upper bound coincides with the exact expression in the case that $\Omega_z$ is a Cartesian product. Using the relation between the mutual information $I(\text{SNR}) = I_{\mathcal{S}_p}\left(\alpha_b^2 \gamma, P\right)$ and the minimum mean-square error (MMSE) in estimating the input symbol, $X_b \in \mathcal{S}_p$, given the output

symbol $Y_b$ [15],

$$\frac{\mathrm{d}}{\mathrm{d\,SNR}} I(\mathrm{SNR}) = \frac{1}{2} \mathrm{MMSE(SNR)}, \tag{18}$$

it is easily proved that $I_{\mathcal{S}_p}(\alpha_b^2 \gamma, P)$ is a concave function (the second derivative is negative) of the instantaneous SNR $\gamma \alpha_b^2$, because the MMSE is a decreasing function of the SNR. Therefore, the proof is the same as for Prop. 2. ∎

Notice that the techniques of the proofs of Props. 2 and 3 cannot be used for discrete input alphabets $\Omega_z$ that are not a Cartesian product of one-dimensional constellations, because in that case the upper bound $\frac{1}{B} \sum_{b=1}^{B} I_{\mathcal{S}_p}(\alpha_b^2 \gamma, P)$ does not coincide with the exact expression of the mutual information. However, this case is merely a particular case ($P = I$) of a *precoded* discrete input alphabet, which is covered in the next section.

## V. BOUNDS ON THE OUTAGE PROBABILITY WITH LINEAR PRECODING

Propositions 2 and 3 mainly state that the outage boundary for a channel with an i.i.d. Gaussian input alphabet or a discrete input alphabet without precoding is outer (inner) bounded by the hypersurface of a $B$-hypersphere touching it at $\boldsymbol{\alpha}_{b,o}$, $\forall\, b$ ($\boldsymbol{\alpha}_e$). We conjecture that this property still holds for a communication system with a discrete alphabet with linear precoding at the input of the channel. First, we will give new detailed proofs for low and high instantaneous SNR of this property. Then, a more intuitive explanation will be given to provide more insight.

*Proposition 4:* On a BF channel at low and high instantaneous SNR, with a discrete input alphabet and with linear precoding, the outage boundary $B_o(discrete)$ is outer bounded by the hypersurface of the $B$-hypersphere $\alpha_1^2 + \alpha_2^2 + \ldots + \alpha_B^2 = \alpha_o^2$ touching it at the axes of the fading space.

*Proof:* See Appendix D. ∎

This outer boundary corresponds with an upper bound on the outage probability. Minimizing this upper bound is simply achieved by minimizing $\alpha_o$. In the following proposition, we will establish an inner bound on the outage probability.

*Proposition 5:* On a BF channel at low and high instantaneous SNR, with a discrete input alphabet and with linear precoding, the outage boundary $B_o(discrete)$ is inner bounded by the hypersurface of the $B$-hypersphere $\alpha_1^2 + \alpha_2^2 + \ldots + \alpha_B^2 = B\alpha_e^2$.

*Proof:* See Appendix E. ∎

This inner boundary corresponds with a lower bound on the outage probability. Minimizing this lower bound is simply achieved by minimizing $\alpha_e$.

To get more insight, we give an illustration of Props. 4 and 5. We consider the $4\text{-}\mathcal{R}^2$ constellation $\Omega_x$ from Fig. 2(a), with $\theta = \pi/6$. Further, we take $\alpha_1 = \alpha_o cos(\lambda)$ and $\alpha_2 = \alpha_o sin(\lambda)$, i.e., $|\boldsymbol{\alpha}| = \alpha_o$, so that $\boldsymbol{\alpha}$ is on the outer boundary of the outage boundary. Fig. 5 shows the corresponding mutual information $I(\boldsymbol{\alpha}, \gamma, \theta)$ as a function of $\gamma$ for various values of $\lambda$. We observe that the mutual information increases when $\lambda$ increases from 0 to $\pi/4$. The projections of the constellation $\Omega_t$ on the horizontal and vertical coordinate axes have variances $\alpha_1^2 = \alpha_o^2 cos^2(\lambda)$ and $\alpha_2^2 = \alpha_o^2 sin^2(\lambda)$ respectively, whereas the total variance of $\Omega_t$ equals $\alpha_o^2$, irrespective of $\lambda$. When $\lambda = \pi/4$, $\Omega_t$ is equivalent to a common QPSK constellation: the variances of both projections equal $\frac{\alpha_o^2}{2}$, and the corresponding mutual information is maximum. When $\lambda$ decreases from $\pi/4$ to 0, the difference between the two variances increases, $\Omega_t$ becomes an increasingly more distorted QPSK constellation, and the mutual information decreases. At $\lambda = 0$ we get $\alpha_1 = \alpha_o$ and $\alpha_2 = 0$; $\Omega_t$ (which is now a scaling of $\mathcal{S}_p$) reduces to a 4-PAM constellation and the corresponding mutual information is minimum. This illustrates the general principle that performance is optimized when the transmit power is equally split over the available identical channels and the performance is worst when the transmit power is completely used for only one channel, which is exactly what is claimed in Props. 4 and 5 for high and low instantaneous SNR. More specifically, for $\boldsymbol{\alpha}$ on the hypersurface of a hypersphere, it is proved in Prop. 4 that the mutual information is the smallest at $\boldsymbol{\alpha}_{b,o}$, and it is proved in Prop. 5 that the mutual information is the largest at $\boldsymbol{\alpha}_e$.

Finally, we show that the precoded constellations achieve full diversity for a rate $R$, provided that $2^{BR}$ does not exceed the number of points contained in the projection of $\Omega_x$ on a coordinate axis.

*Proposition 6:* For any coding rate $0 < R_c < 1$, the outage probability of a Rayleigh distributed block fading channel with a discrete input alphabet and linear precoding exhibits full diversity, i.e., $P_{out}(\gamma, P, R) \propto 1/\gamma^B$.

*Proof:* See Appendix F. ∎

It should be noted that the proof of Prop. 6 assumes the existence of $I_{\mathcal{S}_p}^{-1}(BR, P)$; hence, the number of points in $\mathcal{S}_p$ must not be less than $2^{BR}$. When considering the $4\text{-}\mathcal{R}^2$ constellation from Fig. 2(a), $\mathcal{S}_p$ contains only 2 points when the rotation angle $\theta$ is a multiple of $\pi/2$, and
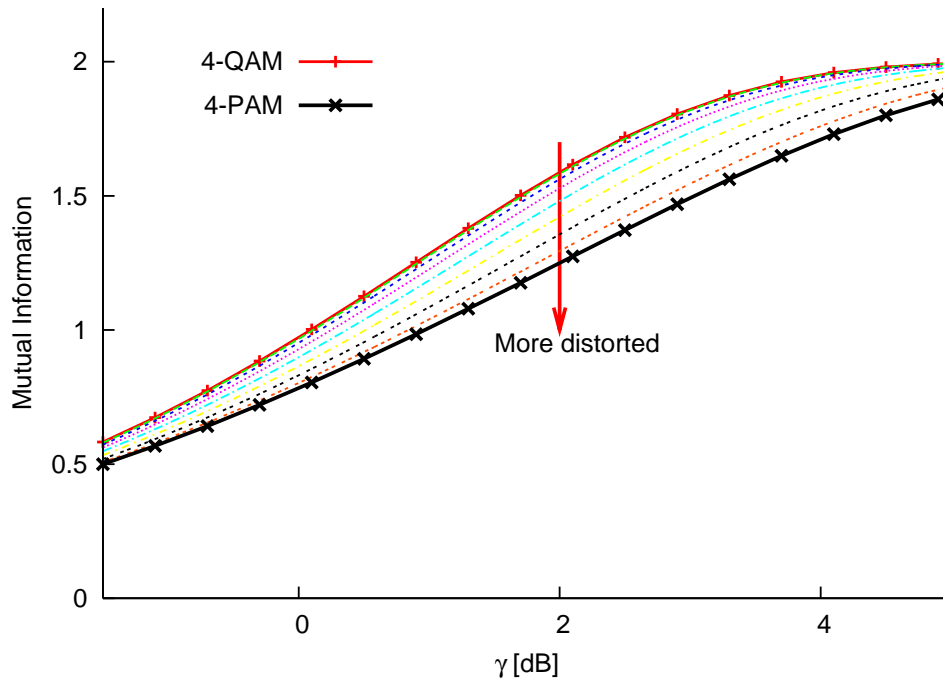
Fig. 5. The mutual information is monotonically non-increasing as the constellation becomes more distorted. The upper and lower curve correspond to $\lambda = \pi/4$ and $\lambda = 0$ respectively.

4 points otherwise. Therefore, the maximum rate that corresponds to full diversity is $R = 0.5$ when the rotation angle $\theta$ is a multiple of $\pi/2$, and $R = 1$ otherwise.

## VI. OPTIMIZATION OF THE OUTAGE PROBABILITY OF PRECODED CONSTELLATIONS

In the previous section, we proved for high and low instantaneous SNR that the outage boundary of block fading channels with precoded constellations is outer bounded by a hypersurface of a $B$-hypersphere with center in the origin and touching the outage boundary on the axes of the fading space. This hypersurface corresponds to an upper bound on the outage probability of the channel (see the paragraph after Def. 3). Instead of minimizing the actual outage probability, it is easier to minimize the upper bound on the outage probability. This optimization will allow the actual outage probability to closely approach a lower bound on the outage probability, i.e., the outage probability corresponding to an i.i.d. Gaussian input alphabet, as we will see in the numerical results.

The $B$-hypersphere is completely determined by one variable, its radius $\alpha_o$. We denote the outer boundary on the outage boundary by $B_{up}(\alpha_o)$ and the corresponding upper bound on the
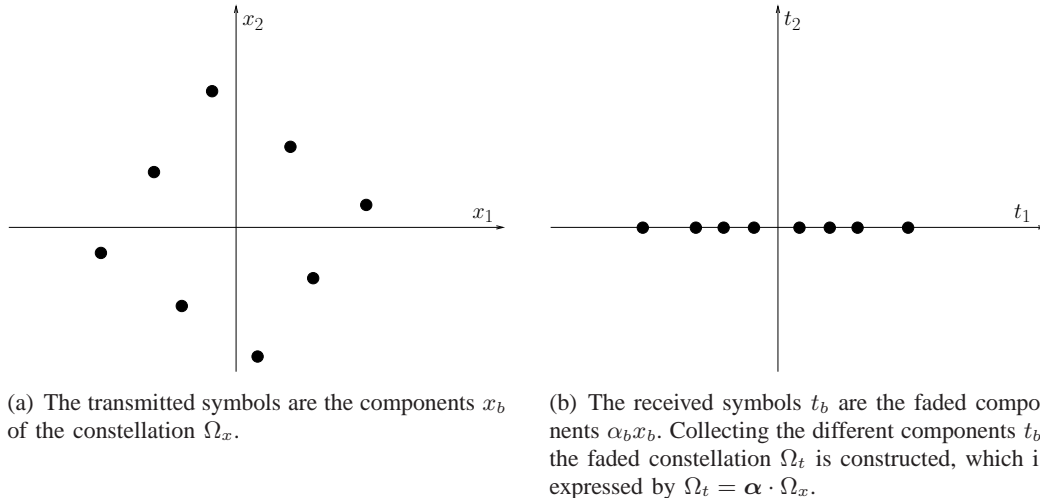
(a) The transmitted symbols are the components $x_b$ of the constellation $\Omega_x$.

(b) The received symbols $t_b$ are the faded components $\alpha_b x_b$. Collecting the different components $t_b$, the faded constellation $\Omega_t$ is constructed, which is expressed by $\Omega_t = \boldsymbol{\alpha} \cdot \Omega_x$.

Fig. 6. The transmitted ($\Omega_x$) and faded constellation ($\Omega_t$) are shown for the case of the transmission of real symbols, $B = 2$ and $\Omega_z = 8\text{-}\mathcal{R}^2$. For $\Omega_t = \boldsymbol{\alpha} \cdot \Omega_x$, the fading point $(\alpha_{1,o}, 0)$, the intersection between the outage boundary and the axis $\alpha_2 = 0$, is used.

outage probability by $P_{up}(\alpha_o)$:

$$P_{up}(\alpha_o) = \int_{\boldsymbol{\alpha} \in V_{up}(\alpha_o)} p(\boldsymbol{\alpha}) d\boldsymbol{\alpha}, \tag{19}$$

where $V_{up}(\alpha_o)$ is the considered $B$-hypersphere limited by $B_{up}(\alpha_o)$. From Eq. (19), it is clear that the region $V_{up}(\alpha_o)$ has to be made as small as possible to minimize $P_{up}(\alpha_o)$. Therefore, the optimization target is to minimize the radius $\alpha_o$.

Because $I_{\mathcal{S}_p}(\alpha_o^2 \gamma, P) = BR$, the minimization of $\alpha_o$ (and, therefore, the minimization of the upper bound on the outage probability) is achieved by selecting the constellation $\mathcal{S}_p$ requiring the least energy to achieve a rate $BR$, $I_{\mathcal{S}_p}(., P) = BR$ on a Gaussian channel. This involves a proper selection of both the constellation $\Omega_z$ and the precoding matrix $P$. Note that this optimization is much simpler than the direct minimization of the outage probability as given by Eqs. (13) and (10), because the outage probability is hard to evaluate, especially when the number of fading gains and constellation points is large. Furthermore, no insight is gained by the latter approach, so that it would not be clear which constellation $\Omega_z$ should be taken.

Also, the outage boundary of block fading channels with precoded constellations is inner bounded by a hypersurface of a $B$-hypersphere with center in the origin and touching the outage boundary in the point $\boldsymbol{\alpha}_e$. Hence, $\alpha_e$ determines the radius of this $B$-hypersphere, $\sqrt{B}\alpha_e$. In this point, the faded constellation $\Omega_t$ is balanced, as in Fig. 6(a). We will denote this balanced

constellation $\Omega_t = \boldsymbol{\alpha}_e \cdot \Omega_x$ by $\mathcal{S}_e$. Hence, the hypersurface corresponds to a lower bound on the outage probability which is minimized by optimizing the mutual information of $\mathcal{S}_e$.

In this section, we will show that the radius of the outer boundary, $\alpha_o$, and the radius of the inner boundary, $\sqrt{B}\alpha_e$, can be minimized by combining a simple optimization of the precoding matrix $P$ with a constellation expansion.

### A. Optimization of the precoding matrix

Because it follows from Eq. (10) that the value $\alpha_e$ is insensitive to orthogonal transformations, the selection of the precoding matrix $P$ affects only the radius of the outer boundary $\alpha_o$. Let us denote by $\mathcal{O}$ the set of parameters from $P$ over which we will minimize $\alpha_o$. For $B = 2$ and $B = 3$, the only degree of freedom is the rotation angle (see (4) and (8)). For $B > 3$, more degrees of freedom can be exploited to minimize $\alpha_o$. For the numerical results, we restrict ourselves to $B \leq 3$.

The mutual information of $\mathcal{S}_p$ can be rewritten as $I_{\mathcal{S}_p}(\alpha_o^2 \gamma, \mathcal{O})$, which yields $\alpha_o^2 = \frac{I_{\mathcal{S}_p}^{-1}(BR,\mathcal{O})}{\gamma}$. Changing the value of $\mathcal{O}$ (e.g. the rotation angle $\theta$ for $B = 2$) will change the distances between the points in $\mathcal{S}_p$ and so change its mutual information. For a fixed spectral efficiency $R$ and fixed average SNR $\gamma$, minimizing the radius yields the optimization criterion

$$\mathcal{O}_{opt} = \arg \min_{\mathcal{O}} I_{\mathcal{S}_p}^{-1}(\mathcal{O}, BR).$$

The optimization is performed by means of a simulation, due to the lack of closed form expressions of the mutual information. Because the constellation is one-dimensional, the computational effort is minimal. We apply this optimization for different scenarios in Sec. VII.

### B. Constellation expansion

As the number of information bits per channel use is $R = mR_c/B$, there are different combinations of $m$ and $R_c$ yielding the same $R$. Taking into account that $R_c \leq 1$, the minimum value of $m$ equals $\lceil BR \rceil$, with a corresponding coding rate $R_c = \frac{BR}{\lceil BR \rceil}$.

The number of points in the constellation is $|\Omega_t|$. Increasing the constellation size of $\Omega_z$ will render a constellation $\Omega_t$ with more points, both for $\mathcal{S}_e$ and $\mathcal{S}_p$. This higher order constellation may need less energy to achieve the same rate, both for the balanced case (optimization of $\alpha_e$) as the distorted case (optimization of $\alpha_o$). However, the decoding complexity increases as

well as the complexity of optimization, so that there is a trade-off between performance and complexity. The higher the constellation size, the smaller the horizontal SNR-gap between the outage probabilities corresponding to a precoded discrete input alphabet and i.i.d. Gaussian input alphabet. However, the improvement in performance becomes smaller and smaller, as illustrated in Sec. VII.

## VII. NUMERICAL RESULTS

### A. Numerical results for $B = 2$

When $B = 2$, $\mathcal{O} = \theta$, and the optimization criterion for the upper bound on the outage probability is to find $\theta$ so that $I_{\mathcal{S}_p}^{-1}(\theta, BR)$ is minimized. Next, a constellation expansion is performed to further minimize the upper bound as well as the lower bound on the outage probability.

*1) Real constellations:* Assume that a transmission rate $R = 0.9$ bpcu is aimed. First, we consider the optimization of the rotation angle $\theta$, see Fig. 7. On the left $y$-axis, we show the instantaneous SNR per symbol, $\gamma_s = \alpha_o^2 \gamma$, so that $I_{\mathcal{S}_p}(\gamma_s, \theta) = BR$. The minimum SNR per symbol $\gamma_s$ that is needed to transmit $R = 0.9$ bpcu for $\gamma = 8$dB is achieved by an i.i.d. Gaussian input alphabet:

$$\gamma_s = \frac{2^{4R} - 1}{2}.$$

This fundamental minimum can be approached when using a precoded discrete input $\Omega_z = 4$-$\mathcal{R}^2$ ($R_c = 0.9$) with rotation angle $\theta = 27$ degrees. Now, we apply a constellation expansion to further reduce $\alpha_o$ (see Fig. 7) and $\alpha_e$ (see Fig. 8(a)). For example, $\gamma_s$ for the rotated constellation $\Omega_z = 8$-$\mathcal{R}^2$ ($R_c = 0.6$) approaches the theoretical minimum very closely for rotation angles within $[0 \ldots 9]$ degrees. An expansion to $\Omega_z = 16$-$\mathcal{R}^2$ ($R_c = 0.45$ and $\theta_{opt} \in [35, 45]$ degrees) only slightly improves the performance. The optimization of $\gamma_s$ decreases the volume of the region $V_o$, which is illustrated in Fig. 3(a) for $\Omega_z = 4$-$\mathcal{R}^2$. Fig. 8(a) illustrates that constellation expansion is sufficient to reduce the value $\alpha_e$, which is very close to the theoretical minimum. Therefore, the constellation is not further shaped to minimize the lower bound of the outage probability.

The information theoretic approach used in this paper does not lead to the same optimized rotations as in the case of algebraic constructions of uncoded constellations. In [3] and [1],
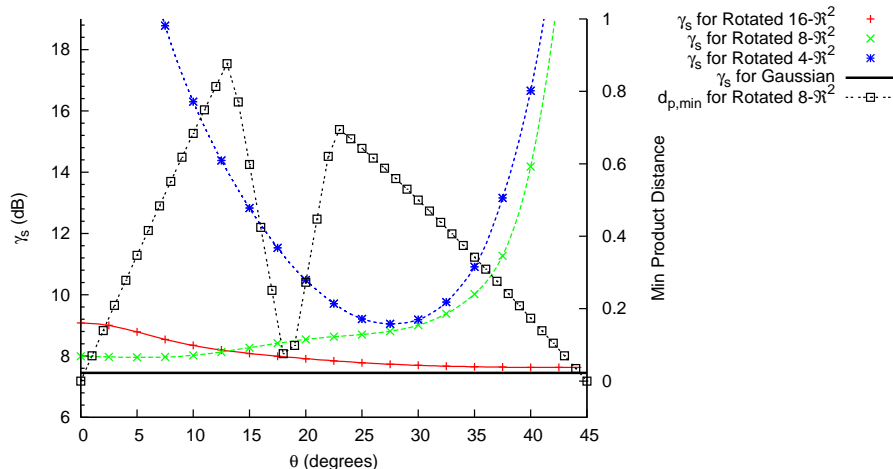
Fig. 7. The optimization of the radius of the outer boundary is shown for $\Omega_z = 4\text{-}\mathcal{R}^2$, $R = 0.9$ bpcu and $\gamma = 8$dB. The y-axis at the left denotes the instantaneous SNR per symbol, $\gamma_s = \alpha_o^2\gamma$, and the right y-axis denotes the minimum product distance. The thick black line without markers represents the fundamental minimum SNR per symbol, $\gamma_s$, that is needed to transmit $R = 0.9$ bpcu for $\gamma = 8$dB, i.e., when using an i.i.d. Gaussian input alphabet. The effect of constellation expansion on the radius $\alpha_o$ is shown by optimization of $\Omega_z = 8\text{-}\mathcal{R}^2$ and $\Omega_z = 16\text{-}\mathcal{R}^2$. The profile of the minimum product distance, $d_{p,min}$, is shown for $\Omega_z = 8\text{-}\mathcal{R}^2$.

multidimensional rotations have been optimized for uncoded infinite constellations transmitted on ergodic fading channels. As a simple illustration, we show in Fig. 7 the minimum product distance $d_{p,min}$ [29] of the uncoded $8\text{-}\mathcal{R}^2$ versus the rotation angle. The optimum and the profile of $d_{p,min}$ and those of $\gamma_s$ do not match. The minimum product distance approach is not suitable for coded schemes. For example, the minimum product distance is zero for $\theta = 0$ degrees, while this rotation angle is optimal in terms of outage probability.

The outage probabilities of the considered multidimensional constellations are shown in Fig. 8(b). This confirms that constellation expansion together with the optimization of the precoding parameter is sufficient to approach the outage probability with an i.i.d. Gaussian input alphabet very closely. It also shows that the constellation $\Omega_x = 8\text{-}\mathcal{R}^2$ with $\theta \in [0 \ldots 9]$ degrees represents the best trade-off between performance and complexity.

*2) Extension to complex constellations:* All the proofs in this paper are valid for complex constellations. This means that also for complex constellations, the outage boundary is outer bounded by a $B$-hypersphere, determined by one variable, its radius. We restrict our attention to real precoding matrices. Consider Eq. (2), where $\mathbf{z}$ is now a complex vector and $P$ is real.
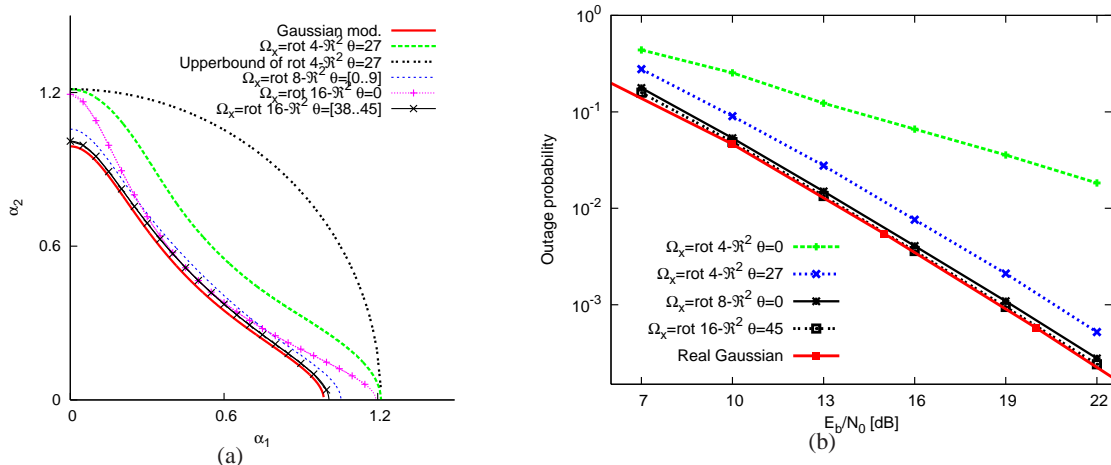
Fig. 8. The outage boundaries (left) and outage probabilities (right) of $\Omega_z = 4\text{-}\mathcal{R}^2$, $\Omega_z = 8\text{-}\mathcal{R}^2$ and $\Omega_z = 16\text{-}\mathcal{R}^2$ with and without optimized rotation angle are shown. The spectral efficiency is $R = 0.9$ bpcu and $\gamma = 8dB$.

For complex symbols, this can be rewritten as

$$\mathbf{x} = P\mathcal{R}\{\mathbf{z}\} + jP\mathbb{I}\{\mathbf{z}\}, \tag{20}$$

where $j^2 = -1$, $\mathcal{R}\{.\}$ and $\mathbb{I}\{.\}$ take the real and complex part respectively. This means that the real and imaginary part of the complex vector are each precoded by the same matrix $P$.

Assume that a transmission rate $R = 1.8$ bpcu is aimed. Initially, we take the constellation $\Omega_z = 16\text{-}\mathcal{C}^2$ ($R_c = 0.9$), which can be build as the Cartesian product of two 4-QAM constellations (16-$\mathcal{C}^2$=4-QAM×4-QAM). As for real constellations, the rotation angle $\theta$ can be optimized, see Fig. 9(a). The gap to the outage probability corresponding with an i.i.d. Gaussian input alphabet can be closed by a constellation expansion and a new optimization of the rotation angle. The same strategy as for real constellations could be applied by only adding one bit in the multidimensional constellation, which would extend $\Omega_z = 16\text{-}\mathcal{C}^2$ to $\Omega_z = 32\text{-}\mathcal{C}^2$. However, $\Omega_z = 32\text{-}\mathcal{C}^2$ cannot be written as the Cartesian product of two constellations and is therefore less convenient to generate (32 points would have to be placed properly in a 4-dimensional space). For simplicity, the constellation expansion is done by adding one bit per component which extends $\Omega_z = 16\text{-}\mathcal{C}^2$ (=4-QAM×4-QAM, $R_c = 0.9$) to $\Omega_z = 64\text{-}\mathcal{C}^2$ (8-QAM×8-QAM [4], $R_c = 0.6$). The optimization of $\theta$ and the optimized outage probabilities are shown in Fig. 9.

---

[4]The 8-QAM constellations have the same form as in Fig. 6(a).

(a) The optimization of the rotation angle for complex symbols is shown.



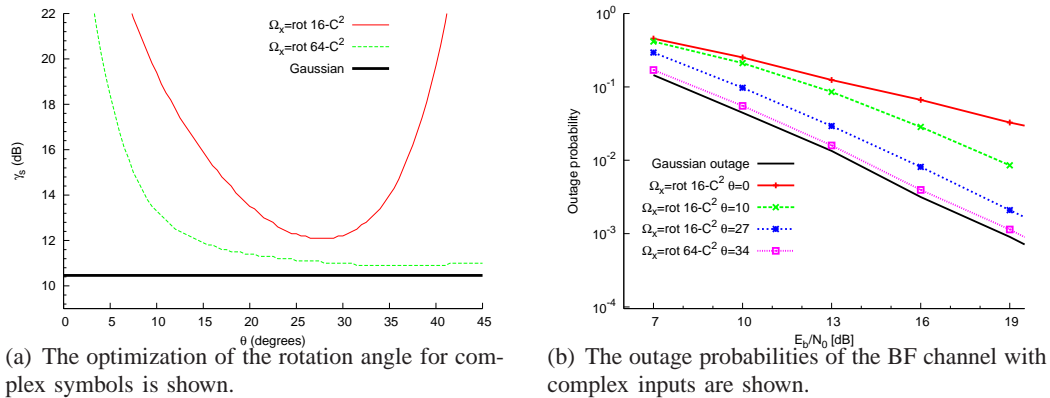(b) The outage probabilities of the BF channel with complex inputs are shown.

Fig. 9. The optimization of $\theta$ and the optimized outage probabilities are shown when complex symbols are transmitted. The transmitted rate is $R = 1.8$ bpcu.

Note that, for $\Omega_z = 16\text{-}\mathcal{C}^2$, the profile of the rotation angle $\theta$, and so the optimum rotation angle, is the same as for $\Omega_z = 4\text{-}\mathcal{R}^2$ and $R = 0.9$ bpcu. This can be explained as follows. When $\mathcal{R}\{\mathbf{z}\}$ and $\mathbb{I}\{\mathbf{z}\}$ are drawn from the same real constellation $\Psi_z$, then $\mathbf{z}$ belongs to a constellation $\Omega_z = \Psi_z + j\Psi_z$. Consequently, $\mathbf{x}$ belongs to a constellation $\Omega_x = \Psi_x + j\Psi_x$, where $\Psi_x$ is obtained by applying the precoding matrix $P$ to the constellation $\Psi_z$. From the chain rule of mutual information [5], we obtain

$$I_{\Omega_x}\left(\gamma\right) = 2I_{\Psi_x}\left(\frac{\gamma}{2}\right). \tag{21}$$

Hence, the precoding matrix $P$ that is optimum for a real constellation $\Psi_z$ and rate $R$ is also optimum for a complex constellation $\Omega_z = \Psi_z + j\Psi_z$ and rate $2R$. The corresponding SNR for the complex constellation is 3 dB higher than for the real constellation. For $\Omega_z = 16\text{-}\mathcal{C}^2$ and $R = 1.8$ bpcu, the corresponding real constellation is $\Psi_z = 4\text{-}\mathcal{R}^2$. Therefore, the profiles in Figs. 9(a) and 7 are the same for both constellations, except for an upward translation of 3 dB.

We also tested the performance for higher spectral efficiencies. Consider for example a system requiring a spectral efficiency of $R = 3.6$ bpcu. Here, the same techniques can be used. First, $\Omega_z = 256\text{-}\mathcal{C}^2$ (16-QAM$\times$16-QAM, $R_c = 0.9$) is optimized, followed by $\Omega_z = 1024\text{-}\mathcal{C}^2$ (32-QAM$\times$32-QAM [5], $R_c = 0.72$). The results are given in Fig. 10. The same observations hold as
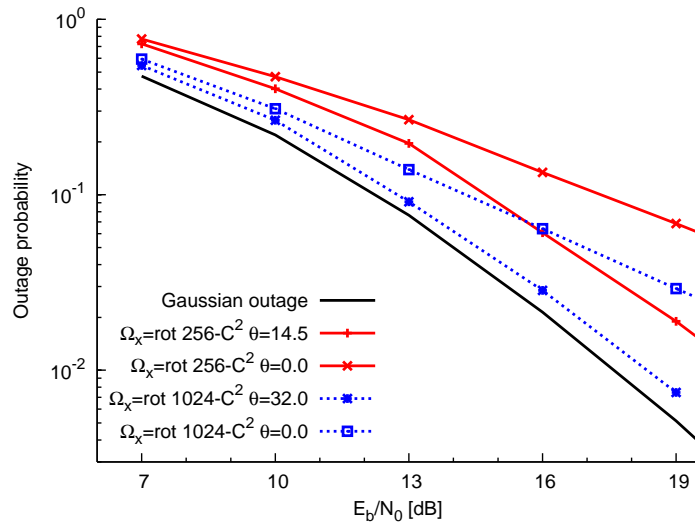
[5]The well known cross 32-QAM constellations are used.

Fig. 10. Outage probabilities of the BF channel with input $\Omega_z = 256\text{-}\mathcal{C}^2$ and $\Omega_z = 1024\text{-}\mathcal{C}^2$. The spectral efficiency is $R = 3.6$ bpcu.

for the previous numerical results.

## B. Numerical results for $B = 3$

When $B = 3$, $\mathcal{O} = \theta_1$ (see Eq. (8)), and the optimization criterion for the upper bound on the outage probability is to find $\theta_1$ so that $I_{\mathcal{S}_p}^{-1}(\theta_1, BR)$ is minimized. Next, a constellation expansion is performed to further minimize the upper bound as well as the lower bound on the outage probability.

We aim to transmit $R = 0.9$ bpcu. We construct constellations $\Omega_z$ that are invariant to a cyclic shift of the components of the constellation points. Hence, the constellations are invariant to a rotation of $2\pi/3$ with respect to the bisector $(1, 1, 1)$ (this can be verified by evaluating $P$ in Eq. (8) for this rotation angle). An example of such constellations are the Cartesian products of three identical one-dimensional constellations, such as $8\text{-}\mathcal{R}^3 = (\text{BPSK})^3$ (constellation points are corners of a cube) and $64\text{-}\mathcal{R}^3 = (4\text{-PAM})^3$ (constellation points are the corners of four nested cubes). More sophisticated constellations can also be considered. For example, $16\text{-}\mathcal{R}^3$ can be constructed by considering 5 sets of three constellation points $\{(a_j, b_j, c_j), (b_j, c_j, a_j), (c_j, a_j, b_j)\}$ [6] for $j = 1, ..., ..., 5$, and adding a constellation point located on the bisector. To build the actual constellations, a few design parameters have to be specified, such as the distances between

[6]The three constellation points are in a plane perpendicular to the bisector.
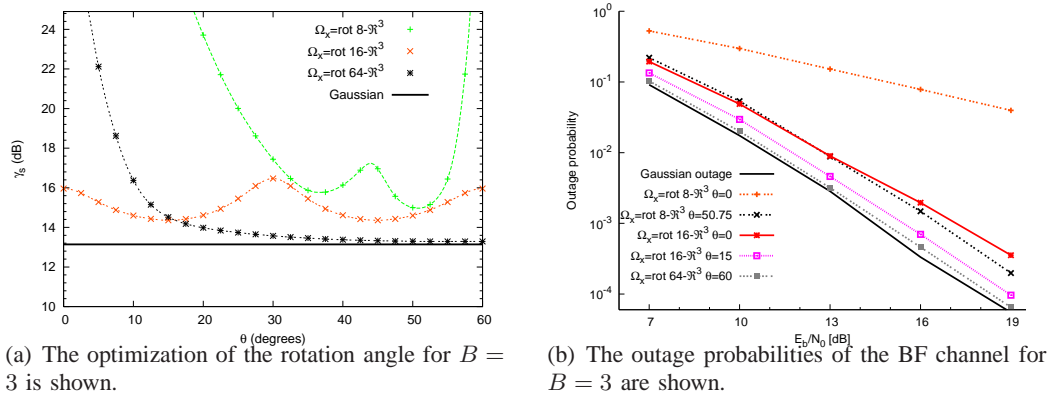
(a) The optimization of the rotation angle for $B = 3$ is shown.

(b) The outage probabilities of the BF channel for $B = 3$ are shown.

Fig. 11. The optimization of $\theta_1$ and the optimized outage probabilities are shown for $B = 3$. The transmitted rate is $R = 0.9$ bpcu.

the planes perpendicular to the bisector, the radii of the circles containing three points of the constellation in the planes, how many points on the bisector are taken and how many points in groups of three are taken. These design parameters will impact on the final performance.

Because it is not the main topic of the paper, we will not elaborate on many different designs for the multidimensional constellations. We compare the performance of $8\text{-}\mathcal{R}^3$ ($R_c = 0.9$), $16\text{-}\mathcal{R}^3$ ($R_c = 0.675$) and $64\text{-}\mathcal{R}^3$ ($R_c = 0.45$) in Fig. 11. Note that the symmetry point in Fig. 11(a) is not $\theta_1 = 45$ degrees, as for $B = 2$, but it is $60$ degrees. The $8\text{-}\mathcal{R}^3$ constellation is equal to the Cartesian product of three BPSK constellations and the $64\text{-}\mathcal{R}^3$ constellation is the Cartesian product of three 4-PAM constellations. For the $16\text{-}\mathcal{R}^3$ constellation, we chose to take a circle of three points in a plane containing the origin perpendicular to the bisector, and two circles with 6 points, each in a plane next to the first plane, perpendicular to the bisector. Finally the origin is also chosen as a constellation point. The rounded coordinates of the points of the constellation $\Omega_z$ are $\{(0.0, 1.0, -1.0), (1.9, 0.12, 0.12), (1.3, -0.5, 1.3), (0.5 - 1.3 - 1.3), (-1.9, -0.1, -0.1), (0, 0, 0)\}$ as well as the cyclic shifts of these coordinates.

## C. Practical implications

Although this work may seem very theoretical, it has an important practical consequence. The WER of a practical system is lower bounded by the outage probability (the lower bound is achievable). To minimize the WER, first its lower bound must be minimized. Therefore, the multidimensional constellation $\Omega_z$ and the rotation angle interval for the practical code should
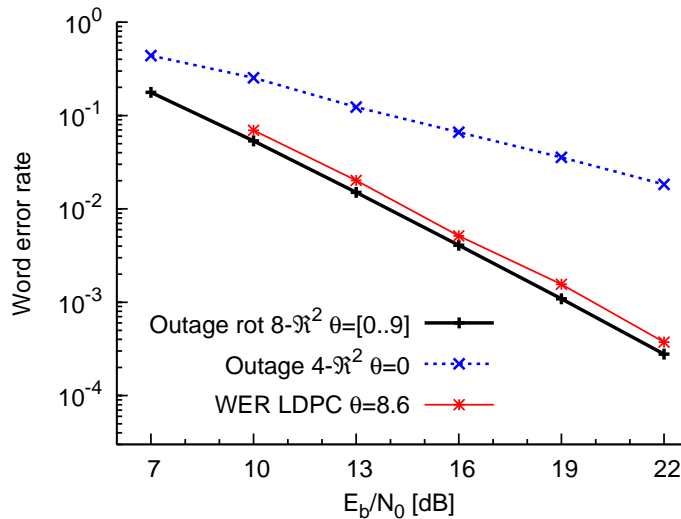
Fig. 12. The optimized outage probability can be approached very closely by the WER of a practical error-correcting code. The block length is $N = 5000$, but the results are valid for all block lengths (tested for $N = 2000$ and $N = 10000$).

be taken as obtained in this work. Next, the labelling, the rotation angle within the rotation angle interval obtained in this paper, and the error-correcting code must be determined. Of course, the last three optimizations are the topics of another work [6], but it is important to understand that these optimizations are based on what is presented in this paper. In Fig. 12, we show for $B = 2$ and $R = 0.9$ bpcu that the optimized outage probability can be approached very closely by the WER of a practical system.

## VIII. CONCLUSIONS

We have studied the effect of linear precoding on the outage probability of block fading channels. We have analyzed the outage boundaries in the fading space and established outer and inner boundaries with simple shapes which yield an easy optimization of the outage probability for a discrete constellation, for an arbitrary number of blocks in the fading channel, real or complex constellations, low or high spectral efficiency. The combination of a constellation expansion and an optimized precoding matrix, has shown to be sufficient to closely approach the outage probability corresponding to an i.i.d. Gaussian input alphabet. With this work, the practical code performance can be optimized by admitting the parameters obtained here.

## APPENDIX

### A. Symmetry conditions for multidimensional constellations

The points $\{\boldsymbol{\alpha}_{b,o}, b = 1, \ldots, B\}$ correspond to the case that all fading gains are zero, except one, whose value is the scaling factor of the projection of the multidimensional constellation on the b-th coordinate axis $x_b$, so that the mutual information between $\mathbf{X}$ and $\mathbf{Y}$ is equal to the spectral efficiency $BR$. In other words, if the projection of the multidimensional constellation on each coordinate axis yields the same set of points, then the magnitudes of the points $\{\boldsymbol{\alpha}_{b,o}, b = 1, \ldots, B\}$ are equal.

First, we restrict our attention to the case that $B = 2$. Consider the constellation point $\mathbf{z}^{(i)} = (u_1^{(i)}, u_2^{(i)}) \in \Omega_z$. The projection of the multidimensional constellation on each coordinate axis yields the same set of points if for each point $\mathbf{z}^{(i)}$, the points $\mathbf{z}^{(j)} = (u_1^{(j)}, u_2^{(j)})$ and $\mathbf{z}^{(q)} = (u_1^{(q)}, u_2^{(q)})$ exist, $i, j, q \in [1, \ldots, 2^m]$; $j, q \neq i$, so that

$$\begin{cases} \cos(\theta)u_1^{(i)} - \sin(\theta)u_2^{(i)} = \sin(\theta)u_1^{(j)} + \cos(\theta)u_2^{(j)} \\ \sin(\theta)u_1^{(i)} + \cos(\theta)u_2^{(i)} = \cos(\theta)u_1^{(q)} - \sin(\theta)u_2^{(q)}. \end{cases}$$

In other words, $x_1^{(i)} = x_2^{(j)}$ and $x_2^{(i)} = x_1^{(q)}$, where $\mathbf{x}^{(i)}, \mathbf{x}^{(j)}$ and $\mathbf{x}^{(q)}$ are the corresponding points of $\mathbf{z}^{(i)}, \mathbf{z}^{(j)}$ and $\mathbf{z}^{(q)}$ in $\Omega_x$. It can be easily verified that this is always fulfilled if

$$\begin{cases} (u_1^{(i)}, u_2^{(i)}) = (u_2^{(j)}, -u_1^{(j)}) \\ (u_1^{(i)}, u_2^{(i)}) = (-u_2^{(q)}, u_1^{(q)}), \end{cases}$$

or in other words, the constellation is invariant under a rotation of $\pi/2$, which is obtained after a cyclic shift and a reflection, which proves what was claimed. An example of such a constellation is the constellation shown in Fig. 2(a).

Now consider the case that $B > 2$. Consider the $B$-dimensional constellation $\Omega_z$ that contains $M$ points. When $\mathbf{z}$ belongs to $\Omega_z$, then also $\mathbf{z}^{(1)}, \ldots, \mathbf{z}^{(B-1)}$ belong to $\Omega_z$, where $\mathbf{z}^{(b)}$ is obtained from $\mathbf{z}$ by a $b$-fold upward cyclic shift of the components of $\mathbf{z}$: $\mathbf{z}^{(b)} = C^b \mathbf{z}$, where $C$ is obtained as a cyclic shift to the right of the columns of the $B \times B$ identity matrix. Note that the number of constellation points does not need to be a multiple of $B$: a subset of the constellation may consist of an arbitrary number of constellation points of the type $[z, z, \ldots, z]^T$ which remain invariant under a cyclic shift.

Consider an orthogonal circulant $B \times B$ precoding matrix $P$. Therefore, $P = CPC^T$ (a circulant matrix remains the same when applying a left cyclic shift to the columns and an upward cyclic shift to the rows). The transformation of $\mathbf{z}^{(b)}$ is

$$P\mathbf{z}^{(b)} = PC^b\mathbf{z} = (CPC^T)C^b\mathbf{z} = CPC^{b-1}\mathbf{z} = \ldots = C^b P\mathbf{z} = C^b\mathbf{x} = \mathbf{x}^{(b)}, \qquad (22)$$

where we exploit that $C$ is an orthogonal matrix.

Consider the matrix $\left(\mathbf{x}, \mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(B-1)}\right)$. As the $(i+1)$-th row is obtained as a cyclic shift to the left of the $i$-th row, the set of components in a row is the same for each row. A constellation point in $\Omega_z$ of the type $(z, z, ..., z)^T$ is transformed into a constellation point in $\Omega_x$ of the type $(x, x, ..., x)^T$. We conclude that the projection of the constellation $\Omega_x$ on any of the coordinate axes yields the same set of points.

*B. Proof of Lemma 1*

The mutual information $I\left(\boldsymbol{\alpha}, \gamma, P = I\right)$ is equal to $\frac{1}{B}I(\mathbf{X}; \mathbf{Y}|\boldsymbol{\alpha}, \gamma)$, which is equal to $\frac{1}{B}I(\mathbf{Y}; \mathbf{X}|\boldsymbol{\alpha}, \gamma)$. This can be split in the sum of $B$ terms through the chain rule [5]:

$$I\left(\boldsymbol{\alpha}, \gamma, P = I\right) = \frac{1}{B}\left(I\left(Y_1; \mathbf{X}|\alpha_1, \gamma\right) + I\left(Y_2; \mathbf{X}|Y_1, \alpha_1, \alpha_2, \gamma\right) + I\left(Y_B; \mathbf{X}|Y_1, \ldots, Y_{B-1}, \boldsymbol{\alpha}, \gamma\right)\right).$$

The mutual information

$$
\begin{aligned}
I\left(Y_b; \mathbf{X} | Y_1, \ldots, Y_{b-1}, \boldsymbol{\alpha}, \gamma\right) &= \mathbb{E}\left[\log \frac{\Pr\left(\mathbf{x} | y_1, \ldots, y_b\right)}{\Pr\left(\mathbf{x} | y_1, \ldots, y_{b-1}\right)}\right] \\
&= \mathbb{E}\left[\log \frac{p\left(y_b | y_1, \ldots, y_{b-1}, \mathbf{x}\right)}{p\left(y_b | y_1, \ldots, y_{b-1}\right)}\right] \\
&= \mathbb{E}\left[\log \frac{p\left(y_b | x_b\right)}{p\left(y_b | y_1, \ldots, y_{b-1}\right)}\right] \\
&= I\left(X_b; Y_b | Y_1, \ldots, Y_{b-1}, \boldsymbol{\alpha}, \gamma\right),
\end{aligned}
$$

which can be further elaborated to

$$
\begin{aligned}
&I\left(X_b; Y_b | Y_1, \ldots, Y_{b-1}, \boldsymbol{\alpha}, \gamma\right) \\
&= H\left(Y_b | Y_1, \ldots, Y_{b-1}, \boldsymbol{\alpha}, \gamma\right) - H\left(Y_b | X_b, Y_1, \ldots, Y_{b-1}, \boldsymbol{\alpha}, \gamma\right).
\end{aligned}
$$

Because

$$
H\left(Y_b | Y_1, \ldots, Y_{b-1}, \boldsymbol{\alpha}, \gamma\right) \leq H\left(Y_b | \boldsymbol{\alpha}, \gamma\right) \tag{23}
$$

and

$$
H\left(Y_b | X_b, Y_1, \ldots, Y_{b-1}, \boldsymbol{\alpha}, \gamma\right) = H\left(Y_b | X_b, \boldsymbol{\alpha}, \gamma\right), \tag{24}
$$

it is clear that

$$
I\left(X_b; Y_b | Y_1, \ldots, Y_{b-1}, \boldsymbol{\alpha}, \gamma\right) \leq I\left(X_b; Y_b | \alpha_b, \gamma\right), \tag{25}
$$

where the upper bound equals $I_{\mathcal{S}_p}\left(\alpha_b^2 \gamma, P\right)$.

### C. Proof of Props. 2 and 3

Consider a function $F(x)$ that is concave for $x \geq 0$. Hence, for arbitrary $L$,

$$
F\left(\sum_{l=1}^{L} p_l \gamma_l\right) \geq \sum_{l=1}^{L} p_l F(\gamma_l), \tag{26}
$$

for $\sum_{l=1}^{L} p_l = 1$, $p_l \geq 0$ and $\gamma_l \geq 0$ for $l = 1, \ldots, L$, by Jensen's inequality. In addition, assume that $F(0) = 0$.

We construct the function $\sum_{b=1}^{B} F(\gamma_b)$ where $\sum_{b=1}^{B} \gamma_b = C$, $\gamma_b \geq 0$ and $\gamma_b = \gamma \alpha_b^2$ for $b = 1, \ldots, B$. Hence, $\boldsymbol{\alpha}$ is on the surface of a hypersphere with squared radius $C/\gamma$.

From Eq. (26) with $L = B$, we obtain

$$\sum_{b=1}^{B} F(\gamma_b) = B\left(\frac{1}{B}\sum_{b=1}^{B}F(\gamma_b)\right) \leq BF\left(\frac{1}{B}\sum_{b=1}^{B}\gamma_b\right) = B.F\left(\frac{C}{B}\right)$$

The maximum value $B.F\left(\frac{C}{B}\right)$ is achieved for $\gamma_b = \frac{C}{B}$, $b = 1, \ldots, B$.

Further, using Eq. (26) with $L = 2$,

$$F(\gamma_b) = F\left(\frac{\gamma_b}{C}C + \frac{C - \gamma_b}{C}\,0\right) \geq \frac{\gamma_b}{C}F(C) + \frac{C - \gamma_b}{C}F(0) = \frac{\gamma_b}{C}F(C)$$

Summing over $b$ yields

$$\sum_{b=1}^{B} F(\gamma_b) \geq \frac{F(C)}{C}\sum_{b=1}^{B}\gamma_b = F(C).$$

The minimum value $F(C)$ is achieved when, for given $l \in \{1, \ldots, B\}$, $\gamma_l = C$ and $\gamma_b = 0$, for $b \neq l$.

### D. Proof of Prop. 4

The instantaneous SNR is $\frac{\mathbb{E}[|\mathbf{t}^2|]}{\mathbb{E}[|\mathbf{w}^2|]} = \gamma|\boldsymbol{\alpha}|^2$. We provide a proof for two SNR regimes: low and high instantaneous SNR.

High instantaneous SNR:

Recall that $I(\boldsymbol{\alpha}, \gamma, P) = \frac{1}{B}I(\mathbf{T}; \mathbf{Y}|\boldsymbol{\alpha}, \gamma)$. In [20] and [25], it is proven that for high instantaneous SNR, maximizing (minimizing) the minimal Euclidean distance $d_{min}$ of a constellation maximizes (minimizes) the mutual information on a Gaussian channel. For high instantaneous SNR, the mutual information is [25]

$$I(\mathbf{T}; \mathbf{Y}|\boldsymbol{\alpha}, \gamma) \approx m - 2\frac{K\pi}{2^m d_{min}^2(\boldsymbol{\alpha})}Q(d_{min}(\boldsymbol{\alpha})\sqrt{\gamma/2}), \tag{27}$$

where $d_{min}(\boldsymbol{\alpha})$ is the minimal distance of the constellation $\Omega_t$ and $K$ is the number of pairs of points at minimum distance in the constellation $\Omega_t$.

Consider two points $\mathbf{x}^{(i)} = (u_1^{(i)}, \ldots, u_B^{(i)})$ and $\mathbf{x}^{(j)} = (u_1^{(j)}, \ldots, u_B^{(j)})$ from the constellation $\Omega_x$. The corresponding points $\mathbf{t}^{(i)}$ and $\mathbf{t}^{(j)}$ from the faded constellation $\Omega_t$ have a squared Euclidean distance given by

$$\left|\mathbf{t}^{(i)} - \mathbf{t}^{(j)}\right|^2 = \sum_{b=1}^{B}\alpha_b^2\left(u_b^{(i)} - u_b^{(j)}\right)^2. \tag{28}$$

Let us denote by $b^*$ the value of $b \in \{1, \ldots, B\}$ for which $\left(u_b^{(i)} - u_b^{(j)}\right)^2$ is minimum. For the B-hypersphere $\sum_{b=1}^{B} \alpha_b^2 = \alpha_o^2$, we obtain

$$\left| \mathbf{t}^{(i)} - \mathbf{t}^{(j)} \right|^2 \geq \alpha_o^2 \left( u_{b^*}^{(i)} - u_{b^*}^{(j)} \right)^2, \tag{29}$$

where equality in Eq. (29) is achieved when $\alpha_{b^*} = \alpha_o$. When $b^*$ is not unique, equality in Eq. (29) holds when $\alpha_{b^*} = \alpha_o$ holds for any $b^*$ that minimizes $\left( u_b^{(i)} - u_b^{(j)} \right)^2$. Hence, the minimum distance $d_{min}(\boldsymbol{\alpha})$ for the constellation $\Omega_t$ is given by

$$d_{min}(\boldsymbol{\alpha}) = \alpha_o \min_{\substack{i,j \in \{1,\ldots,M\} \\ i \neq j}} \min_{b \in \{1,\ldots,B\}} \left| u_b^{(i)} - u_b^{(j)} \right| \tag{30}$$

As the constellation $\Omega_x$ is such that its projection on either coordinate axis yields the same set of points, the minimum distance $d_{min}(\boldsymbol{\alpha})$ is achieved for either fading gain equal to $\alpha_o$ (and the remaining $B - 1$ fading gains equal to 0).

Low SNR:

Consider the mutual information $I(\boldsymbol{\alpha}, \gamma, P)$ of the constellation $\Omega_t$, given the fading gains $\boldsymbol{\alpha}$ (Eq. (10)). This expression can be reformulated in terms of $\gamma$:

$$I(\boldsymbol{\alpha}, \gamma, P) = \frac{m}{B} - \frac{2^{-m}}{B} \sum_{\mathbf{x} \in \Omega_x} \mathbb{E}_{\mathbf{y}|\mathbf{x}} \left[ \log_2 \left( \sum_{\mathbf{x}' \in \Omega_x} \exp\left[ \gamma \left( d^2(\mathbf{y}, \boldsymbol{\alpha} \cdot \mathbf{x}) - d^2(\mathbf{y}, \boldsymbol{\alpha} \cdot \mathbf{x}') \right) \right] \right) \right],$$

where $\mathbb{E}_{\mathbf{y}|\mathbf{x}}$ can be replaced by an expectation over the noise, $\mathbb{E}_{\mathbf{w}}$, $w_b \sim \mathcal{N}(0, 1/(2\gamma))$. The argument of the exponential functions can be simplified, so that

$$I(\boldsymbol{\alpha}, \gamma, P) = \frac{m}{B} - \frac{2^{-m}}{B} \sum_{\mathbf{x} \in \Omega_x} \mathbb{E}_{\mathbf{w}} \left[ \log_2 \left( \sum_{\mathbf{x}' \in \Omega_x} \exp\left[ -\gamma d^2(\boldsymbol{\alpha} \cdot \mathbf{x}, \boldsymbol{\alpha} \cdot \mathbf{x}') + \sum_{b=1}^{B} (2\gamma w_b f(b)) \right] \right) \right],$$

where $f(b) = \alpha_b(x_b - x_b')$. This expression can be further simplified by approximating the exponential functions and logarithms by their respective Taylor series for small $\gamma \alpha_b^2$, $\forall b = 1, \ldots, B$.. Next, the expectation of the expression over the random vector $\mathbf{w}$ can be replaced by an expectation over $\gamma \mathbf{w}$, where $\gamma \mathbf{w} \sim \mathcal{N}(\mathbf{0}, \frac{\gamma}{2}\mathbf{I})$ ($\mathbf{I}$ is the identity matrix). Therefore, we can drop all terms that are proportional to $\mathbb{E}_{\gamma w_b}[\gamma w_b]$ and replace $\mathbb{E}_{\gamma w_b}[(\gamma w_b)^2]$ in all terms proportional

to $\mathbb{E}_{\gamma w_b}\left[(\gamma w_b)^2\right]$ by $\frac{\gamma}{2}$. Now, after some calculus,we obtain

$$I\left(\boldsymbol{\alpha}, \gamma, P\right) = \frac{\gamma \sum_b \alpha_b^2 Var(X_b)}{B\log(2)} + \sum_{b=1}^{B} o(\gamma \alpha_b^2),$$

where $Var(X(b))$ is the variance of the $b$-th component of the points of constellation $\Omega_x$. The validity of this approximation for small instantaneous SNR has been verified numerically. As the projection of $\Omega_x$ on either coordinate axis yields the same set of points, this variance is independent of $b$. Hence, for small instantaneous SNR, the mutual information remains constant for the set of fading points where $\sum_{b=1}^{B} \alpha_b^2$ is constant. By the definition of $\alpha_o$, it is clear that for low instantaneous SNR, the outage boundary coincides with the hypersurface of the B-hypersphere $\sum_{b=1}^{B} \alpha_b^2 = \alpha_o^2$.

*E. Proof of Prop. 5*

For low SNR, the outage boundary coincides with the hypersurface of the B-hypersphere $\sum_{b=1}^{B} \alpha_b^2 = \alpha_o^2$ (see proof op Prop. 4), which in turn coincides with the hypersurface of the B-hypersphere $\sum_{b=1}^{B} \alpha_b^2 = B\alpha_e^2$ in this case.

For high SNR, we must prove that the minimal distance of $\Omega_t$ is maximized in $\boldsymbol{\alpha}_e$, or

$$\boldsymbol{\alpha}_e = \arg \max_{\substack{\boldsymbol{\alpha} \\ |\boldsymbol{\alpha}|^2 = B\alpha_e^2}} \left( \min_{i,j} \sum_{b=1}^{B} \alpha_b^2 \left( u_b^{(i)} - u_b^{(j)} \right)^2 \right) \tag{31}$$

For $\boldsymbol{\alpha} = \boldsymbol{\alpha}_e$, we obtain $\sum_{b=1}^{B} \alpha_b^2 \left( u_b^{(i)} - u_b^{(j)} \right)^2 = \alpha_e^2 \sum_{b=1}^{B} \left( u_b^{(i)} - u_b^{(j)} \right)^2$, which is minimized to $\alpha_e^2 d^2$ when $\mathbf{x}^{(i)}$ and $\mathbf{x}^{(j)}$ are points at minimum distance $d$ in the constellation $\Omega_x$. For any $\boldsymbol{\alpha}$ with $|\boldsymbol{\alpha}|^2 = B\alpha_e^2$, $\boldsymbol{\alpha} \neq \boldsymbol{\alpha}_e$, we will prove that $\min_{i,j} \sum_{b=1}^{B} \alpha_b^2 \left( u_b^{(i)} - u_b^{(j)} \right)^2 \leq \alpha_e^2 d^2$, which is what is claimed. We first restrict our attention to the case $B > 2$. We consider $B$ pairs $\left\{ \left( \mathbf{x}^{(n)}, \mathbf{x}'^{(n)} \right), n = 0, \ldots, B-1 \right\}$ of points in $\Omega_x$ and their corresponding points $\left\{ \left( \mathbf{t}^{(n)}, \mathbf{t}'^{(n)} \right), n = 0, \ldots, B-1 \right\}$ in $\Omega_t$, such that $\mathbf{x}^{(n)} = T^n \mathbf{x}$ and $\mathbf{x}'^{(n)} = T^n \mathbf{x}'$, with $T$ denoting the matrix operator that causes a downward cyclic shift. We select $\mathbf{x}$ and $\mathbf{x}'$ such that $|\mathbf{x} - \mathbf{x}'| = d$; obviously, $|\mathbf{x}^{(n)} - \mathbf{x}'^{(n)}| = d$ for $n = 0, \ldots, B-1$. Because $\sum_{b=1}^{B} \left( u_b^{(n)} - u_b'^{(n)} \right)^2 = \sum_{n=0}^{B-1} \left( u_b^{(n)} - u_b'^{(n)} \right)^2$, we obtain that, for any $\boldsymbol{\alpha}$ with $|\boldsymbol{\alpha}|^2 = B\alpha_e^2$, the average distance $\mathbb{E}_n|\mathbf{t}^{(n)} -$

$\mathbf{t}'^{(n)}|^2$ is

$$\frac{1}{N} \sum_{n=0}^{B-1} \left( \sum_{b=1}^{B} \alpha_b^2 \left( u_b^{(n)} - u_b'^{(n)} \right)^2 \right) = \frac{1}{N} \sum_{b=1}^{B} \alpha_b^2 \left( \underbrace{\sum_{n=0}^{B-1} \left( u_b^{(n)} - u_b'^{(n)} \right)^2}_{d^2} \right) = d^2 \alpha_e^2. \tag{32}$$

Hence, $\min_n \sum_{b=1}^{B} \alpha_b^2 \left( u_b^{(n)} - u_b'^{(n)} \right)^2 \leq d^2 \alpha_e^2$. For $B = 2$, the proof has to be slightly modified: we consider 4 pairs $\left\{ \left( \mathbf{x}^{(n)}, \mathbf{x}'^{(n)} \right), n = 0, \ldots, B-1 \right\}$ of points in $\Omega_x$, such that $\mathbf{x}^{(n)} = T^n \mathbf{x}$ and $\mathbf{x}'^{(n)} = T^n \mathbf{x}'$, with $T$ denoting the matrix operator representing a counter clockwise rotation of $\pi/2$. We select $\mathbf{x}$ and $\mathbf{x}'$ such that $|\mathbf{x} - \mathbf{x}'| = d$; obviously, $|\mathbf{x}^{(n)} - \mathbf{x}'^{(n)}| = d$ for $n = 0, \ldots, 3$. For any $\boldsymbol{\alpha}$ with $|\boldsymbol{\alpha}|^2 = 2\alpha_e^2$, we have, considering that $T^2 = -I$,

$$\frac{1}{4} \sum_{n=0}^{3} \left( \sum_{b=0}^{1} \alpha_b^2 \left( u_b^{(n)} - u_b'^{(n)} \right)^2 \right) = \frac{1}{4} \sum_{b=0}^{1} \alpha_b^2 \left( \underbrace{\sum_{n=0}^{3} \left( u_b^{(n)} - u_b'^{(n)} \right)^2}_{2d^2} \right) = d^2 \alpha_e^2. \tag{33}$$

The remainder of the proof follows the same lines as for $B > 2$.

### F. Proof of Prop. 6

Let us find an expression for $\alpha_o$. Therefore, we consider that all fading gains are zero, but one, so that the rate achieved by $\Omega_t$ (which is now a scaling of $\mathcal{S}_p$) over that single slot must be at least $RB$, or $I_{\mathcal{S}_p}(\alpha_o^2 \gamma, P) = BR$, so that

$$\alpha_o^2 = \frac{I_{\mathcal{S}_p}^{-1}(BR, P)}{\gamma}, \tag{34}$$

Now, as the outage boundary $B(V_o)$ is outer bounded by the hypersurface of the hypersphere with radius $\alpha_o$, we have

$$P_{out}(\gamma, P, R) \leq \Pr(\alpha_1^2 + \alpha_2^2 + \ldots + \alpha_B^2 < \alpha_o^2). \tag{35}$$

Because the fading gains $\{\alpha_b, \ b = 1, \ldots, B\}$ are i.i.d. Rayleigh, the cumulative distribution function of the chi-square distribution with $2B$ degrees of freedom is [26]

$$\Pr(\alpha_1^2 + \alpha_2^2 + \ldots + \alpha_B^2 < x) = 1 - e^{-x} \sum_{k=0}^{B-1} \frac{x^k}{k!}.$$

The right hand side can be manipulated as follows:

$$1 - e^{-x} \sum_{k=0}^{B-1} \frac{x^k}{k!} = \frac{e^x - \sum_{k=0}^{B-1} \frac{x^k}{k!}}{e^x} = \frac{\sum_{k=B}^{\infty} \frac{x^k}{k!}}{\sum_{m=0}^{\infty} \frac{x^m}{m!}} = \frac{x^B \sum_{k=0}^{\infty} \frac{x^k}{(B+k)!}}{\sum_{m=0}^{\infty} \frac{x^m}{m!}} \leq x^B,$$

so that from Eqs. (34) and (35)

$$P_{out}(\gamma, P, R) \leq \alpha_o^{2B} \propto \frac{1}{\gamma^B}. \tag{36}$$

Hence, with increasing $\gamma$ the outage probability goes to zero not slower than $\gamma^{-B}$, so the diversity order is at least $B$. As there are only $B$ i.i.d. fading gains involved in the transmission of a codeword, the diversity order is exactly $B$.

## REFERENCES

[1] D. Bayer-Fluckiger, F. Oggier, E. Viterbo, "New algebraic constructions of rotated $Z^n$-lattice constellations for the Rayleigh fading channel ," *IEEE Trans. on Inf. Theory*, vol. 50, no. 4, pp. 702-714, 2004.

[2] E. Biglieri, J. Proakis, and S. Shamai, "Fading channels: information-theoretic and communications aspects," *IEEE Trans. on Inf. Theory*, vol. 44, no. 6, pp. 2619-2692, Oct. 1998.

[3] J.J. Boutros and E. Viterbo, "Signal space diversity: a power- and bandwidth-efficient diversity technique for the Rayleigh fading channel," *IEEE Trans. on Inf. Theory*, vol. 44, no. 4, pp. 1453-1467, July 1998.

[4] J.J. Boutros, A. Guillén i Fàbregas, and E. Calvanese, "Analysis of coding on non-ergodic block fading channels," *Allerton Conf. on Communication and Control*, Illinois, 2005.

[5] T.M. Cover and J.A. Thomas, *Elements of Information Theory*, New York, Wiley, 2006.

[6] D. Duyck, J.J. Boutros, and M. Moeneclaey, "Coded modulations for block fading channels," *IEEE Symposium on Communications and Vehicular Technology in the Benelux*, Twente, The Netherlands, Nov. 2010.

[7] D. Duyck, J.J. Boutros, and M. Moeneclaey, "Low-Density Graph Codes for Slow Fading Relay Channels," *Accepted in IEEE Trans. on Inf. Theory*, Download from telin.ugent.be/~dduyck/publications/paper_ldpc_cooperative.pdf

[8] D. Duyck, D. Capirone, C. Hausl, M. Moeneclaey, "Design of Diversity-Achieving LDPC Codes for H-ARQ with Cross-Packet Channel Coding," *Personal Indoor Mobile Radio Communications (PIMRC)*, Istanbul, Turkey, Sept. 2010.

[9] D. Duyck, J.J. Boutros, M. Moeneclaey, "Rotated Modulations for Outage Probability Minimization: a fading space approach," ISIT (Intern. Symp. on Inf. Theory), Texas, June 2010.

[10] A. Guillén i Fàbregas, *Concatenated codes for block-fading channels*, Ph.D. thesis, EPFL, June 2004.

[11] A. Guillén i Fàbregas and G. Caire, "Coded modulation in the block-fading channel: coding theorems and code construction," *IEEE Trans. on Inf. Theory*, vol. 52, no. 1, pp. 91-114, Jan. 2006.

[12] A. Guillén i Fàbregas and G. Caire, "Multidimensional coded modulation in block-fading channels," *IEEE Trans. on Inf. Theory*, vol. 54, no. 5, pp. 2367-2372, 2008.

[13] N. Gresset, J.J. Boutros, L. Brunel, "Optimal linear precoding for BICM over MIMO channels," *International Symposium on Information Theory (ISIT)*, Chicago, Illinois, July 2004.

[14] N. Gresset, L. Brunel, J.J. Boutros, "Space-time coding techniques with bit interleaved coded modulation over block-fading MIMO channels," *IEEE Trans. on Inf. Theory,* vol. 54, no. 5, pp. 2156-2178, May 2008.

[15] D. Guo, S. Shamai, and S. Verdú, "Mutual information and minimum mean-square error in Gaussian channels," *IEEE Trans. on Inf. Theory*, vol. 51, no. 4, pp. 1261-1282, Apr. 2005.

[16] C. Hausl and A. Chindapol, "Hybrid ARQ with Cross-Packet Channel Coding," *IEEE Comm. Letters*, vol. 11, no. 5, pp. 434-436, May 2007.

[17] R. Knopp and P.A. Humblet, "On coding for block fading channels," *IEEE Trans. on Inf. Theory*, vol. 46, no. 1, pp. 189-205, Jan. 2000.

[18] G.M. Kraidy, N. Gresset, and J.J. Boutros, "Information theoretical versus algebraic constructions of linear unitary precoders for non-ergodic multiple antenna channels," *Canadian Workshop on Information Theory*, Montreal, Canada, June 2005.

[19] J.N. Laneman, D. Tse, and G.W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062-3080, Dec. 2004.

[20] A. Lozano, A. M. Tulino, and S. Verdú, "Optimum power allocation for parallel Gaussian channels with arbitrary input distributions," *IEEE Transactions on Information Theory*, vol. 52, no. 7, pp. 3033 3051, July 2006.

[21] E. Malkamaki and H. Leib, "Evaluating the performance of convolutional codes over block fading channels," *IEEE Trans. on Inf. Theory*, vol. 45, no. 5, pp. 1643-1646, Jul. 1999.

[22] S.K. Mohammed, E. Viterbo, Y. Hong, and A. Chockalingam, "Precoding by Pairing Subchannels to Increase MIMO Capacity with Discrete Input Alphabets," *accepted for publication in IEEE Trans. on Inf. Theory*, Jan. 2011.

[23] S.K. Mohammed, E. Viterbo, Y. Hong, and A. Chockalingam, "MIMO Precoding with X- and Y-Codes," *accepted for publication in IEEE Trans. on Inf. Theory*, Oct. 2010.

[24] L.H. Ozarow, S. Shamai and A.D. Wyner, "Information theoretic considerations for cellular mobile radio," *IEEE Trans. on Vehicular Technology*, vol. 43, no. 2, pp. 359-379, May 1994.

[25] F. Pérez-Cruz, M.R.D. Rodrigues and S. Verdú, "Optimal precoding for Digital Subscriber Lines," *Proc. Intern. Conf. on Comm. (ICC)*, Beijing, China, May 2008.

[26] J.G. Proakis, "Digital Communications," McGraw-Hill, 4th ed., 2000.

[27] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity-Part I: System description," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1927-1938, Nov. 2003, DOI: 10.1109/TCOMM.2003.818096.

[28] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity-Part II: Implementation aspects and performance analysis," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1939-1948, Nov. 2003.

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