Optimum power control over fading channels

Giuseppe Caire Giorgio Taricco Ezio Biglieri

Dipartimento di Elettronica • Politecnico • Corso Duca degli Abruzzi 24 • I-10129 Torino (Italy)

e-mail: <name>@polito.it

This work was supported by the Italian Space Agency (ASI).

December 9, 1998

Abstract

We study optimal constant-rate coding schemes for a block-fading channel with strict transmission delay constraint, under the assumption that both the transmitter and the receiver have perfect channel-state information. We show that the information outage probability is minimized by concatenating a standard "Gaussian" code with an optimal power-controller, which allocates the transmitted power dynamically to the transmitted symbols. We solve the minimum outage probability problem under different constraints on the transmitted power and we derive the corresponding power allocation strategies. In addition, we propose an algorithm that approaches the optimal power allocation when the fading statistics are not known. Numerical examples for different fading channels are provided, and some applications discussed. In particular, we show that minimum outage probability and delay-limited capacity are closely related quantities, and we find a closed-form expression for the delay-limited capacity of the Rayleigh block-fading channel with transmission over two independent blocks. We also discuss repetition diversity and its relation with direct-sequence or multicarrier spread-spectrum transmission. The optimal power allocation strategy in this case corresponds to selection diversity at the transmitter. From the single-user point of view considered in this paper, there exists an optimal repetition diversity order (or spreading factor) that minimizes the information outage probability for given rate, power, and fading statistics.

Keywords: Block-fading channel, power control, channel capacity, information outage probability.

1 Introduction

Fostered by the increasing importance of mobile wireless systems, of late a great deal of valuable scholarly work has gone into assessing the information-theoretic limits of the channels which model the mobile environment. This paper inscribes itself in this framework, its main goal being the derivation of optimal constant-rate transmission schemes for the block-fading Gaussian (BF-AWGN) channel under a strict transmission delay constraint.

The BF-AWGN channel, introduced in [1, 2], belongs to the general class of block-interference channels described in [3]. It is motivated by the fact that, in many wireless communication situations, changes in the propagation environment occur on a very slow time scale with respect to the signaling rate. In the BF-AWGN channel, blocks of N symbols undergo the same "channel state" (defined by the fading gain), which is random but constant over the whole block. A code word of length $n = MN$ spans M blocks (a group of M blocks will be referred to as a *frame*). Blocks can be thought of as separated in time (e.g., in a TDMA system as in [1]), as separated in frequency (e.g., in a multicarrier system), or as separated both in time and in frequency (e.g., in a slow time-frequency hopping system as in [4, 5]). We consider the case where all the channel states in a frame are known to the transmitter *before* transmitting a code word. Then, our model is better suited to the case of a multicarrier system with M parallel subchannels, possibly located at non-adjacent carrier frequencies—otherwise, if blocks were transmitted in different time intervals, the transmitter Channel-State Information (CSI) would be non-causal.[1]

The physical validity of the BF-AWGN model is discussed in [1], where it is observed that the code word transmission delay is essentially determined by interleaving. For a fixed block length N, the number of blocks M spanned by a code word is related to the system interleaving depth. Therefore, M can be considered as a measure of the overall transmission delay. For typical practical systems, N is fairly large [1, 5]. Then, it makes sense to study the BF-AWGN channel performance limits as $N \to \infty$, distinguishing between the "delay-unconstrained" case, where also $M \to \infty$ and the (interleaving) "delay-limited" case, where M is fixed and finite (we refer to this case as the M-block BF-AWGN channel).

With no delay constraint, the capacity derived in [9] is the relevant performance limit indicator.

[1]From a practical viewpoint, CSI at the transmitter can be provided either by a dedicated feedback channel (some existing systems already implement a fast power control feedback channel [6, 7]) or by time-division duplex [8], where the uplink and the downlink time-share the same M subchannels and the fading gains can be estimated from the incoming signal.

This applies for example to variable-rate systems, like wireless data networks [10]. On the other hand, most of today's mobile radio systems carry real-time speech (cellular telephony), for which constant-rate, delay-limited transmission should be considered. In this case, *information outage probability*, defined as the probability that the *instantaneous mutual information* of the channel is below the transmitted code rate, is the appropriate performance limit indicator [1, 2, 11]. As a matter of fact, outage probability predicts surprisingly well the error probability of actual codes for practical values of N (say $N \simeq 100$) [4, 12, 13].

The M-block BF-AWGN channel is typically not *information stable*, since it is characterized by a finite number M of random channel states. Then, its capacity should be studied in the general framework of [14], extended to continuous input, output and state spaces by standard discretization and partitioning arguments [15][Ch. 7]. We follow the terminology of [16] and refer to the capacity of the M-block BF-AWGN channel as to the "delay-limited" capacity. Outage probability and delay-limited capacity are closely related. Namely, the delay-limited capacity can be obtained as the maximum rate at which the minimum outage probability is zero.

Here we solve the problem of minimizing the outage probability of the M-block BF-AWGN channel under the assumption that both transmitter and receiver have perfect CSI. We show that the minimum outage probability can be achieved by transmitting a fixed code book, randomly generated with i.i.d. Gaussian symbols, and by suitably allocating the transmitted power to the blocks. The optimal power allocation strategy is derived subject to different constraints on the transmitted power. Solving the minimum outage probability problem, also yields a way of computing the delay-limited capacity. In particular, an explicit formula is derived for the delay-limited capacity of the 2-block BF-AWGN channel with Rayleigh i.i.d. blocks. This shows that a small delay ($M = 2$) can buy a considerable increase in transmission reliability (the delay-limited capacity is zero for $M = 1$). Moreover, the delay-limited capacity of this channel is only 5 dB away from the capacity of the AWGN channel, and 2.5 dB away from the delay-unconstrained capacity of the Rayleigh fading channel ($M \to \infty$).

All the results listed above assume optimal coding over the M-block frame. A suboptimal coding scheme is *repetition diversity* [1], which consists of repeating the same code word of length N over all the M blocks. This can be viewed as the concatenation of a code of length N with a repetition code of length M.[2] In spite of its obvious suboptimality, repetition diversity might be an option because of its simplicity. Moreover, it is useful as a model to study the performance of spread-spectrum in a

[2]Note the difference with respect to space diversity: this may be thought of as a technique to modify the fading gain statistics after combining, and hence does not decrease the code rate.

single-user, frequency-selective fading channel [5, 17]. In fact, the M-block BF-AWGN channel can be seen as a simplified model for a frequency-selective fading channel with M subbands, each of which can be considered as frequency-flat, and repetition diversity can be interpreted as spreading by a factor M (the same symbol is repeated, or spread, over M subbands). With an appropriate signal space representation, this model can be applied either to direct-sequence and to multicarrier CDMA formats [12].

We show that the outage probability minimizing scheme for repetition diversity is equivalent to selection diversity at the transmitter, where a single subband is used in every frame and power is allocated optimally. This result agrees with the findings of [12, 18], where a similar strategy for maximizing the rate-sum of a multiple-access channel with fading and transmitter CSI is derived. Also, we show that for both optimal and constant power allocation there exists an optimal diversity order (or spreading factor) which minimizes the outage probability, and that the outage probability approaches 1 as $M \to \infty$. These simple facts show that the common belief that "spread-spectrum signals are more robust against frequency-selective fading" should be reconsidered, at least when transmitter CSI is available (an even more striking result along this line was recently obtained in [19]: it shows that, when CSI is not perfect and a "peakiness" constraint on the signal power spectral density is imposed, an arbitrarily large signal bandwidth expansion drives the capacity to zero even for optimal coding.)

The paper is organized as follows. In Section 2 we formally define the BF-AWGN channel, recall known results about its capacity, introduce outage probability and delay-limited capacity, and provide a coding theorem which gives operational meaning to these quantities. In Section 3 we solve the minimum outage probability problem, while in Section 4 the general solution is applied to simple on-off and Rayleigh independent fading channels. Finally, in Section 5 we summarize our main conclusions.

Notation. Here we define the notation used throughout this paper:

- $\mathbb{R}_+$ indicates the non-negative real line and $\mathbb{R}_+^M$ is the non-negative orthant of the M-dimensional real Euclidean space.
- Given two random M-vectors $\mathbf{x}$ and $\mathbf{y}$, we denote their joint and conditional cdf's by the short-hand notations $F(\mathbf{x}, \mathbf{y})$ and $F(\mathbf{x} \mid \mathbf{y})$, respectively, instead of the more complete but cumbersome

$$
\begin{aligned}
F_{\mathbf{x},\mathbf{y}}(\mathbf{u}, \mathbf{v}) &\triangleq P(x_0 \le u_0, \dots, x_{M-1} \le u_{M-1}, y_0 \le v_0, \dots, y_{M-1} \le v_{M-1}) \\
F_{\mathbf{x}|\mathbf{y}}(\mathbf{u} \mid \mathbf{v}) &\triangleq P(x_0 \le u_0, \dots, x_{M-1} \le u_{M-1} \mid y_0 = v_0, \dots, y_{M-1} = v_{M-1})
\end{aligned}
$$

- Given a set $\mathcal{A}$ in the probability space of a random variable x, the indicator function of the event $\{x \in \mathcal{A}\}$ is denoted by $\chi_{\mathcal{A}}$.
- Given a vector $\mathbf{a}$ of length M, we denote its arithmetic mean by $\langle \mathbf{a} \rangle \stackrel{\Delta}{=} \frac{1}{M}\sum_{m=0}^{M-1} a_m$.
- $[x]_+ \stackrel{\Delta}{=} \max(x, 0)$.
- $\mathrm{Ei}\,(n, x) \stackrel{\Delta}{=} \int_1^\infty \frac{e^{-xt}}{t^n} dt$ $(\mathrm{Re}\{x\} > 0)$.
- $\Pi_k(x) \stackrel{\Delta}{=} e^{-x}\sum_{j=0}^{k-1} x^j/j! = \Gamma(k,x)/\Gamma(k)$ where $\Gamma(k,x) = \int_x^\infty u^{k-1}e^{-u}du$ and $\Gamma(k) = \Gamma(k,0)$.
- The normal distribution with mean μ and variance σ^2 is denoted by $\mathcal{N}(\mu, \sigma^2)$, and $\sim$ means "distributed as".

2 Channel model, capacity and outage probability

Let $\mathbf{x}, \mathbf{y}$, and $\mathbf{z}$ be vectors in $\mathbb{R}^{MN}$ representing the channel input, output, and noise sequences, respectively, where $\mathbf{z}$ is an i.i.d. $\sim \mathcal{N}(0,1)$ sequence. We can arrange the components of $\mathbf{x}, \mathbf{y}$, and $\mathbf{z}$ as $M \times N$ arrays denoted $\mathbf{X}, \mathbf{Y}$, and $\mathbf{Z}$, respectively. The transmission of a code word over the (real) BF-AWGN channel spans exactly one frame of M fading blocks. This can be written concisely as

$$\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{Z} \tag{1}$$

where $\mathbf{A} = \mathrm{diag}(\sqrt{\alpha_0}, \ldots, \sqrt{\alpha_{M-1}})$ is an $M \times M$ matrix whose diagonal elements are the fading amplitudes over the current frame. Input symbols on the same row of $\mathbf{X}$ experience the same fading coefficient, i.e., they are transmitted over the same fading block. We denote by $\boldsymbol{\alpha} \in \mathbb{R}_+^M$ the sequence $\{\alpha_m\}_{m=0}^{M-1}$ of fading powers in the frame. As usual, $\boldsymbol{\alpha}$ is assumed to be statistically independent of $\mathbf{x}$ and $\mathbf{z}$. The transmission of a long (infinite) sequence of code words is characterized by a sequence of frames $\{\boldsymbol{\alpha}^{(k)}\}_{k=-\infty}^{\infty}$. This can be seen as a vector random process, which we assume to be ergodic and have first-order cdf $F(\boldsymbol{\alpha})$ (the marginal cdf's of $F(\boldsymbol{\alpha})$ may be different).

2.1 Capacity with no delay constraints

In this section we consider BF-AWGN channels with block length $N \leq \infty$ and M arbitrarily large. We assume that $\{\alpha_m\}_{m=0}^{M-1}$ is asymptotically ergodic [20] as $M \to \infty$.[3] These channels form a family, indexed by the block length $N = 1, 2, \ldots$. It is well-known that, with perfect CSI, all channels in

[3]This assumption is sufficient to ensure the information stability of the channel [14].

the family have the same capacity, independent of N [3]. In this case, the following result holds (see [1, 2, 3, 9]):

Proposition 1. *Under the average transmitted power constraint* $\mathrm{E}[x^2] \leq \mathcal{P}$ *(x is a component of vector* $\mathbf{x}$*), the capacity of the BF-AWGN channel is given by:*

1. *With perfect CSI at the receiver and no CSI at the transmitter,*

$$C_{\text{const}}(\mathcal{P}) = \mathrm{E}\left[\frac{1}{2}\log(1+\alpha\mathcal{P})\right] \tag{2}$$

2. *With perfect CSI at both the receiver and the transmitter,*

$$C_{\text{opt}}(\mathcal{P}) = \max_{\gamma} \mathrm{E}\left[\frac{1}{2}\log(1+\alpha\gamma)\right] \tag{3}$$

where the maximization is over the power allocation functions $\gamma = \gamma(\alpha)$ *such that* $\mathrm{E}[\gamma] \leq \mathcal{P}$.

□

Remark. For all $N = 1, 2, \ldots$, capacities (2) and (3) are achieved by sequences of codes with block length MN, as $M \to \infty$. Capacity (2) is achieved by random codes whose symbols are generated independently according to the Gaussian distribution $\mathcal{N}(0, \mathcal{P})$. In the sequel, this coding scheme will be referred to as "single-codebook, constant-power" transmission. In [9], the proof that (3) is achievable was obtained by using a coding scheme based on multiplexing different code books with different rates and average powers, where the multiplexer and the corresponding demultiplexer are driven by the fading process $\{\alpha_m\}$. We refer to such scheme as "multiple-codebook, variable-power" (or also "variable-rate, variable-power") transmission [9]. Now, the latter is not necessary in order to achieve (3). In fact, the same capacity can also be achieved by a single codebook with i.i.d. symbols $\sim \mathcal{N}(0,1)$, where the m-th block of N symbols is scaled by $\sqrt{\gamma(\alpha_m)}$ before transmission [21]. From a practical point of view, this scheme (referred to as "single-codebook, variable-power" transmission) is especially appealing, since it can be simply implemented by concatenating a conventional "Gaussian" encoder with a power-controller driven by the transmitter CSI.

A final remark on the results stated in Proposition 1 is about terminology. Although $C_{\text{const}}(\mathcal{P})$ and $C_{\text{opt}}(\mathcal{P})$ are expressed in the form of ensemble averages, there is no reason to refer to these quantities as "average" capacities. In fact, because of ergodicity, the *information density* [14] of sample pairs of input and output sequences converges to $C_{\text{const}}(\mathcal{P})$ (or to $C_{\text{opt}}(\mathcal{P})$, depending on the input distribution) with probability 1 as $M \to \infty$ and given N. Thus, $C_{\text{const}}(\mathcal{P})$ and $C_{\text{opt}}(\mathcal{P})$ are not averages of achievable rates over an ensemble of channel realizations, but rather achievable rates *for all* channel realizations, with probability 1. □

The power allocation function achieving $C_{\rm opt}(\mathcal{P})$ is [9]

$$\gamma(\alpha) = \left[\frac{1}{\lambda} - \frac{1}{\alpha}\right]_+ \tag{4}$$

where λ is the solution of the constraint equation

$$\int_\lambda^\infty \left(\frac{1}{\lambda} - \frac{1}{\alpha}\right) dF(\alpha) = \mathcal{P} \tag{5}$$

The resulting expression for $C_{\rm opt}(\mathcal{P})$ is

$$C_{\rm opt}(\mathcal{P}) = \int_\lambda^\infty \frac{1}{2}\log(\alpha/\lambda) dF(\alpha) \tag{6}$$

Example: Capacity with Rayleigh fading. For further reference (this result will be used later on, to assess the effect of finite interleaving), consider a Rayleigh fading channel with D-th order independent space diversity and maximal-ratio combining. The channel gain after combining has pdf [22]

$$f(\alpha) = \frac{D^D \alpha^{(D-1)} \exp(-D\alpha)}{(D-1)!} \tag{7}$$

where we assume a normalized average channel gain $\mathrm{E}[\alpha] = 1$. The capacity with constant power is computed via integration by parts of (2) (see [23])

$$C_{\rm const}(\mathcal{P}) = \frac{1}{2}\left(\Pi_D(-D/\mathcal{P})\mathrm{Ei}\,(1, D/\mathcal{P}) + \sum_{k=1}^{D-1} \frac{1}{k}\Pi_k(D/\mathcal{P})\Pi_{D-k}(-D/\mathcal{P})\right)$$

Eq. (5) in this case becomes, for $D = 1$,

$$\frac{1}{\lambda}e^{-\lambda} - \mathrm{Ei}\,(1, \lambda) = \mathcal{P}$$

and, for $D > 1$,

$$\frac{1}{\lambda}\Pi_D(D\lambda) - \frac{D}{D-1}\Pi_{D-1}(D\lambda) = \mathcal{P}$$

The capacity with optimal power allocation is given by [23]

$$C_{\rm opt}(\mathcal{P}) = \frac{1}{2}\left(\mathrm{Ei}\,(1, D\lambda) + \sum_{k=1}^{D-1} \frac{1}{k}\Pi_k(D\lambda)\right)$$

where λ is the solution of (5), and if $D = 1$ the summation in the RHS above is void.

Fig. 1 shows $C_{\text{const}}(\mathcal{P})$ and $C_{\text{opt}}(\mathcal{P})$ for $D = 1,2,3$. The AWGN channel capacity $C_{\text{awgn}}(\mathcal{P}) = \frac{1}{2}\log(1+\mathcal{P})$ is shown for comparison.[4] Optimal power allocation offers only a small improvement over constant power for rates above 0.5 bit/symbol, especially for $D > 1$. A moderate space diversity order (e.g., $D = 3$) is sufficient to approach the AWGN capacity by less than 1 dB. Finally, while $C_{\text{const}}(\mathcal{P}) \leq C_{\text{awgn}}(\mathcal{P})$ because of Jensen's inequality, no similar inequality exists for $C_{\text{opt}}(\mathcal{P})$, which may be larger than $C_{\text{awgn}}(\mathcal{P})$ for low $\mathcal{P}$ (for $D = 1$, the cross-over is about $\mathcal{P} = 0$ dB).

2.2 Performance limits under a delay constraint

We now focus on delay-limited transmission at constant rate R. For a given finite M, we consider a family of M-block BF-AWGN channels indexed by the block length $N = 1, 2, \ldots$. As discussed in Section 1, we are interested in the limiting performance as $N \to \infty$. Next, we introduce the information outage probability and the capacity of the M-block BF-AWGN channel. Then, we provide a coding theorem which gives operational meaning to these quantities.

Definition 1: Instantaneous mutual information. *The maximum instantaneous mutual information $I_M(\boldsymbol{\alpha},\boldsymbol{\gamma})$ of the M-block BF-AWGN channel for a given sequence of input powers $\boldsymbol{\gamma} = (\gamma_0, \ldots, \gamma_{M-1})$, where $\gamma_m \triangleq E[|x_{mN+n}|^2]$ for all $n = 0, 1, \ldots,$ is given by*

$$I_M(\boldsymbol{\alpha},\boldsymbol{\gamma}) \triangleq \frac{1}{M}\sum_{m=0}^{M-1}\frac{1}{2}\log(1+\alpha_m\gamma_m) \tag{8}$$

This is achieved by independent Gaussian inputs $x_{mN+n} \sim \mathcal{N}(0,\gamma_m)$. $\Diamond$

The sequence of instantaneous powers $\boldsymbol{\gamma}$ defines the allocation of the transmitted power to the M blocks. Since the transmitter has perfect CSI, $\boldsymbol{\gamma}$ depends on $\boldsymbol{\alpha}$ (we shall use equivalently the notations $\boldsymbol{\gamma}$ and $\boldsymbol{\gamma}(\boldsymbol{\alpha})$). In general, this dependence is expressed by the conditional cdf $F(\boldsymbol{\gamma} \mid \boldsymbol{\alpha})$. Then, $\boldsymbol{\gamma}(\boldsymbol{\alpha})$ is a random function of $\boldsymbol{\alpha}$ (deterministic functions are special cases). In this paper, we consider the following class of power allocation functions:

Definition 2: Probabilistic stationary memoryless power allocation. *The class of probabilistic stationary memoryless power allocation functions is the set of all time-invariant random functions*

[4] All the numerical results in this paper can be translated immediately into results for the more standard circularly-symmetric complex channel [1] with average energy per symbol E_s, noise power spectral density N_0 and signaling rate W symbols/s by letting $\mathcal{P} = E_s/N_0$ and by multiplying the information rates by $2W$. In this way, the information rates are expressed in information units per second.

$\boldsymbol{\gamma} = \boldsymbol{\gamma}(\boldsymbol{\alpha})$, $\mathbb{R}_+^M \to \mathbb{R}_+^M$, whose conditional cdf $F(\boldsymbol{\gamma}|\boldsymbol{\alpha})$ has the following property. Consider a sequence of M-block frames, indexed by $k = 0, 1, \ldots, \nu-1$. Let $\boldsymbol{\alpha}^{(k)}$ and $\boldsymbol{\gamma}^{(k)}$ be the channel gain and the power allocation vectors in the k-th frame. Then,

$$F(\boldsymbol{\gamma}^{(0)}, \cdots, \boldsymbol{\gamma}^{(\nu-1)} \mid \boldsymbol{\alpha}^{(0)}, \cdots, \boldsymbol{\alpha}^{(\nu-1)}) = \prod_{k=0}^{\nu-1} F(\boldsymbol{\gamma}^{(k)} \mid \boldsymbol{\alpha}^{(k)}) \tag{9}$$

◊

We define the *short-term* average transmitted power as the normalized power sum over a frame of M blocks. The resulting short-term power constraint is given by

$$\frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} |x_{mN+n}|^2 \leq \mathcal{P} \tag{10}$$

Next, consider a sequence of M-block frames. We define the *long-term* average transmitted power as the normalized power sum over the whole sequence. The resulting long-term power constraint is given by

$$\frac{1}{K} \sum_{k=0}^{K-1} \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \left| x_{mN+n}^{(k)} \right|^2 \leq \mathcal{P} \tag{11}$$

Under the assumptions of stationary and memoryless power allocation and ergodic fading, the RHS of (10) converges to $\langle \boldsymbol{\gamma} \rangle$ with probability 1, and (11) converges to $E[\langle \boldsymbol{\gamma} \rangle]$ where expectation is with respect the joint distribution of $(\boldsymbol{\alpha}, \boldsymbol{\gamma})$. Then, we have

Definition 3: Power constraints. *For an arbitrary non-negative constant $\mathcal{P}$, the short-term power constraint is given by*

$$\langle \boldsymbol{\gamma} \rangle \leq \mathcal{P} \quad \textit{with probability 1} \tag{12}$$

and the long-term power constraint is given by

$$\mathrm{E}[\langle \boldsymbol{\gamma} \rangle] \leq \mathcal{P} \tag{13}$$

◊

For a given function $\boldsymbol{\gamma}(\boldsymbol{\alpha})$ satisfying (12) or (13), $I_M(\boldsymbol{\alpha}, \boldsymbol{\gamma})$ is a random variable. Following [1] we define the information outage probability as

Definition 4: Information outage probability. *Let $\boldsymbol{\gamma}$ satisfy (12) or (13). The information outage probability evaluated at rate R is given by:*

$$P_{\rm out}(R, \mathcal{P}) \stackrel{\Delta}{=} P(I_M(\boldsymbol{\alpha}, \boldsymbol{\gamma}) < R) \tag{14}$$

$\diamondsuit$

In the following, we are mainly concerned with the minimization of $P_{\rm out}(R, \mathcal{P})$ with respect to the choice of the power allocation function in the class defined by Definition 2. More precisely, minimization is over the conditional cdf's $F(\boldsymbol{\gamma}|\boldsymbol{\alpha})$ satisfying (12), (13) or both. The minimum outage probability is intimately related to the capacity of the M-block BF-AWGN channel. In order to distinguish the channel capacity without delay constraints from the capacity of the M-block channel, the latter is referred to as the "delay-limited" capacity [16]. We have

Definition 5: Delay-limited capacity. *The delay-limited capacity of the M-block BF-AWGN channel, subject to a short-term (resp., long-term) power constraint, is given by*

$$C_{\rm delay}(\mathcal{P}) \stackrel{\Delta}{=} \sup_{\boldsymbol{\gamma}} \inf_{\boldsymbol{\alpha} \in \mathbb{R}_+^M} I_M(\boldsymbol{\alpha}, \boldsymbol{\gamma}) \tag{15}$$

where the supremum is over all $F(\boldsymbol{\gamma}|\boldsymbol{\alpha})$ satisfying (12) (resp., (13)). $\diamondsuit$

The next proposition is a coding theorem which follows as an application of the results of [14].

Proposition 2. *The maximum ϵ-achievable [5] rate of the M-block BF-AWGN channel subject to a short-term (resp., long-term) power constraint is given by*

$$C_\epsilon(\mathcal{P}) = \sup_{\boldsymbol{\gamma}} \sup\{R : P_{\rm out}(R, \mathcal{P}) \leq \epsilon\} \tag{16}$$

where the supremum is over all $F(\boldsymbol{\gamma}|\boldsymbol{\alpha})$ satisfying (12) (resp., (13)). Also, the delay-limited capacity (15) is given by $C_{\rm delay}(\mathcal{P}) = \lim_{\epsilon \to 0_+} C_\epsilon(\mathcal{P})$.

Proof. See Appendix A. $\square$

Remark. As a consequence of Proposition 2 and Definition 1, we notice that the minimum outage probability is achieved by random codes whose symbols in the m-th block are i.i.d. $\sim \mathcal{N}(0, \gamma_m)$, for the optimal choice of $\boldsymbol{\gamma}$. Any such code can be obtained as the concatenation of a code with i.i.d.

[5] See Appendix A for the definition of ϵ-achievable rate used in this paper.

symbols $\sim \mathcal{N}(0,1)$ with a *power-controller* scaling the m-th block of N symbols by $\sqrt{\gamma_m}$. Therefore, all minimum outage probabilities (and delay-limited capacities) obtained in this paper can be achieved by single-codebook, variable-power transmission schemes. $\square$

Example: Delay-limited capacity with Rayleigh fading. By specializing the results of [16] to our case, we obtain immediately the delay-limited capacity in the case $M = 1$ as

$$C_{\text{delay}}(\mathcal{P}) = \frac{1}{2}\log\left(1 + \frac{\mathcal{P}}{\mathrm{E}[1/\alpha]}\right)$$

In the case of Rayleigh fading with D-th order independent space diversity, from the above we obtain

$$C_{\text{delay}}(\mathcal{P}) = \frac{1}{2}\log\left(1 + \frac{D-1}{D}\mathcal{P}\right)$$

Fig. 2 compares $C_{\text{delay}}(\mathcal{P})$ for $D = 2, 3$ with the AWGN capacity. Note that without diversity, i.e., for $D = 1$, the delay-limited capacity of the Rayleigh fading channel is zero: in fact, $\mathrm{E}[1/\alpha] = \infty$, which expresses the fact that "channel inversion" requires transmission of an infinite average power. The results of [16] do not apply to the case of M-block transmission, for finite $M > 1$. In the rest of the paper we solve the minimum outage probability problem for general M-block transmission and we provide an explicit expression of the delay-limited Rayleigh fading channel capacity for $M = 2$ and $D = 1$.

3 Minimum outage probability

In this section we solve the minimum outage probability problem for both short-term and long-term constraints. In the first case, we obtain the standard "water-filling" solution, which maximizes $I_M(\boldsymbol{\alpha}, \boldsymbol{\gamma})$ for any given $\boldsymbol{\alpha}$ (some extra technicalities, related to the uniqueness of the solution, are needed). Next, we solve the more interesting (and challenging) problem of outage probability minimization under long-term constraint. Further, we briefly discuss the case where both short- and long-term constraints are present, and exhibit a general method for computing the delay-limited capacity. Finally, we analyze the suboptimal scheme based on repetition diversity. [6]

[6]Since $\boldsymbol{\gamma} = \boldsymbol{\gamma}(\boldsymbol{\alpha})$ is a random function, defined by its conditional cdf $F(\boldsymbol{\gamma}|\boldsymbol{\alpha})$, all optimization problems in the following are to be intended with respect to $F(\boldsymbol{\gamma}|\boldsymbol{\alpha})$. However, it is convenient to give the solution directly in terms of $\boldsymbol{\gamma}$, rather than in terms of its conditional cdf.

3.1 Short-term power constraint

The constrained minimization problem can be stated as

$$\begin{cases} \text{Minimize} & P(I_M(\boldsymbol{\alpha},\boldsymbol{\gamma}) < R) \\ \text{Subject to} & \langle\boldsymbol{\gamma}\rangle \leq \mathcal{P} \quad \text{with probability 1} \end{cases} \tag{17}$$

It should be rather intuitive that the solution to the above problem must maximize $I_M(\boldsymbol{\alpha},\boldsymbol{\gamma})$ for all $\boldsymbol{\alpha}$ (or at least, for all $\boldsymbol{\alpha}$ in a certain subset of $\mathbb{R}_+^M$), so that (17) can be reduced to a mutual-information maximization problem, whose solution is well-known [15]. Formally, we have the following proposition:

Proposition 3. *Problem (17) is solved by*

$$\widehat{\boldsymbol{\gamma}}(\boldsymbol{\alpha}) = \begin{cases} \boldsymbol{\gamma}^{\rm st}(\boldsymbol{\alpha}) & \textit{if } \boldsymbol{\alpha} \notin \mathcal{U}(R,\mathcal{P}) \\ \mathbf{g}(\boldsymbol{\alpha}) & \textit{if } \boldsymbol{\alpha} \in \mathcal{U}(R,\mathcal{P}) \end{cases} \tag{18}$$

where $\mathbf{g}(\boldsymbol{\alpha})$ *is an arbitrary function* $\mathbb{R}_+^M \to \mathbb{R}_+^M$ *such that* $\langle\mathbf{g}(\boldsymbol{\alpha})\rangle \leq \mathcal{P}$ *with probability 1,* $\boldsymbol{\gamma}^{\rm st}(\boldsymbol{\alpha})$ *is the solution of the maximization problem*

$$\begin{cases} \text{Maximize} & I_M(\boldsymbol{\alpha},\boldsymbol{\gamma}) \\ \text{Subject to} & \langle\boldsymbol{\gamma}\rangle \leq \mathcal{P} \end{cases} \tag{19}$$

and where $\mathcal{U}(R,\mathcal{P})$*, the* outage region*, is given by*

$$\mathcal{U}(R,\mathcal{P}) = \left\{\boldsymbol{\alpha} \in \mathbb{R}_+^M : I_M(\boldsymbol{\alpha},\boldsymbol{\gamma}^{\rm st}(\boldsymbol{\alpha})) < R\right\} \tag{20}$$

Since $\mathbf{g}(\boldsymbol{\alpha})$ *is arbitrary, the solution of (17) is in general not unique.*

Proof. See Appendix B. □

For the sake of completeness and for future use we obtain the explicit form of $\boldsymbol{\gamma}^{\rm st}(\boldsymbol{\alpha})$ and of the outage region $\mathcal{U}(R,\mathcal{P})$. The solution of (19) is readily obtained by using Lagrange multipliers and Kuhn-Tucker conditions [15]. Let the region $\mathcal{Q} \subset \mathbb{R}_+^M$ be defined by

$$\mathcal{Q} = \left\{\boldsymbol{\alpha} \in \mathbb{R}_+^M : \alpha_0 \geq \cdots \geq \alpha_{M-1}\right\} \tag{21}$$

and assume $\boldsymbol{\alpha} \in \mathcal{Q}$. Then, the m-th component of $\boldsymbol{\gamma}^{\rm st}(\boldsymbol{\alpha})$ is given by

$$\gamma_m^{\rm st}(\boldsymbol{\alpha}) = \left[\lambda^{\rm st}(\mu,\boldsymbol{\alpha}) - \frac{1}{\alpha_m}\right]_+ \qquad \text{for } m = 0,\dots,M-1 \tag{22}$$

where

$$\lambda^{\text{st}}(\mu, \boldsymbol{\alpha}) = \frac{1}{\mu} \sum_{\ell=0}^{\mu-1} \frac{1}{\alpha_\ell} + \frac{M}{\mu} \mathcal{P} \tag{23}$$

and μ is the unique integer in $\{1, \ldots, M\}$ such that $1/\alpha_m \le \lambda^{\text{st}}(\mu, \boldsymbol{\alpha})$ for $m < \mu$ and $1/\alpha_m > \lambda^{\text{st}}(\mu, \boldsymbol{\alpha})$ for $m \ge \mu$.

The function $\boldsymbol{\gamma}^{\text{st}}(\boldsymbol{\alpha})$, defined by (22) and (23) for $\boldsymbol{\alpha} \in \mathcal{Q}$, is extended to the whole $\mathbb{R}_+^M$ by sorting the components of $\boldsymbol{\alpha}$ in non-increasing order, computing (22) on the sorted fading vector, and applying the inverse permutation to the resulting power allocation vector. Similarly, in order to study the outage region $\mathcal{U}(R, \mathcal{P})$, it is convenient to give an explicit expression for the intersection $\mathcal{U}(R, \mathcal{P}) \cap \mathcal{Q}$ and obtain the inequalities defining $\mathcal{U}(R, \mathcal{P})$ outside $\mathcal{Q}$ by permuting the coordinate axes. Hence, without loss of generality, we shall consider only the case $\boldsymbol{\alpha} \in \mathcal{Q}$.

The region $\mathcal{Q}$ is divided into M subregions $\mathcal{V}_i$ defined by the inequalities

$$\begin{cases} \lambda^{\text{st}}(i, \boldsymbol{\alpha}) \ge 1/\alpha_m & m = 0, \ldots, i-1 \\ \lambda^{\text{st}}(i, \boldsymbol{\alpha}) < 1/\alpha_m & m = i, \ldots, M-1 \end{cases} \tag{24}$$

Since μ in (23) exists and is unique, $\{\mathcal{V}_i : i = 1, \ldots, M\}$ is a partition of $\mathcal{Q}$. Notice that, for $\boldsymbol{\alpha} \in \mathcal{V}_i$, exactly i blocks are used for transmission (i.e., are given positive instantaneous power). The minimum outage probability is obtained as the probability that $\boldsymbol{\alpha}$ belongs to the outage region $\mathcal{U}(R, \mathcal{P})$, whose intersection with $\mathcal{V}_i$ is given explicitly by

$$\mathcal{U}(R, \mathcal{P}) \cap \mathcal{V}_i = \left\{ \boldsymbol{\alpha} \in \mathcal{Q} \ : \ \frac{1}{M} \sum_{m=0}^{i-1} \frac{1}{2} \log \left(\alpha_m \lambda^{\text{st}}(i, \boldsymbol{\alpha}) \right) < R \right\} \quad \text{for } i = 1, \ldots, M \tag{25}$$

When $\boldsymbol{\alpha} \in \mathcal{U}(R, \mathcal{P})$, an outage event occurs irrespective of the choice of $\mathbf{g}(\boldsymbol{\alpha})$. Then, the most sensible choice is to set $\mathbf{g}(\boldsymbol{\alpha}) = \mathbf{0}$, i.e., to turn off transmission. However, this is not the only choice, unless some other constraints are taken into account.

3.2 Long-term power constraint

The constrained minimization problem can be stated as

$$\begin{cases} \text{Minimize} & P(I_M(\boldsymbol{\alpha}, \boldsymbol{\gamma}) < R) \\ \text{Subject to} & \mathrm{E}[\langle \boldsymbol{\gamma} \rangle] \le \mathcal{P} \end{cases} \tag{26}$$

Before stating the main result of this section, which provides a general solution for (26), we need some lemmas and definitions.

Consider the minimization problem, dual of (19)

$$\begin{cases} \text{Minimize} & \langle \boldsymbol{\gamma} \rangle \\ \text{Subject to} & I_M(\boldsymbol{\alpha}, \boldsymbol{\gamma}) \geq R \end{cases} \tag{27}$$

We have

Lemma 1. *Assume $\boldsymbol{\alpha} \in \mathcal{Q}$. Then the m-th component of $\boldsymbol{\gamma}^{\rm lt}(\boldsymbol{\alpha})$, solution of (27), is given by*

$$\gamma_m^{\rm lt}(\boldsymbol{\alpha}) = \left[\lambda^{\rm lt}(\mu, \boldsymbol{\alpha}) - \frac{1}{\alpha_m} \right]_+ \quad \textit{for } m = 0, \ldots, M-1 \tag{28}$$

where

$$\lambda^{\rm lt}(\mu, \boldsymbol{\alpha}) = \left(\frac{e^{2MR}}{\prod_{\ell=0}^{\mu-1} \alpha_\ell} \right)^{1/\mu} \tag{29}$$

and where μ is the unique integer in $\{1, \ldots, M\}$ such that $1/\alpha_m \leq \lambda^{\rm lt}(\mu, \boldsymbol{\alpha})$ for $m < \mu$ and $1/\alpha_m > \lambda^{\rm lt}(\mu, \boldsymbol{\alpha})$ for $m \geq \mu$.

Proof. See Appendix C. □

The function $\boldsymbol{\gamma}^{\rm lt}(\boldsymbol{\alpha})$, defined by (28) and (29) for $\boldsymbol{\alpha} \in \mathcal{Q}$, is extended to the whole $\mathbb{R}_+^M$ by component permutation. Hence, without loss of generality, we shall consider only the case $\boldsymbol{\alpha} \in \mathcal{Q}$.

Lemma 2. *$\boldsymbol{\gamma}^{\rm lt}(\boldsymbol{\alpha})$ is continuous for $\boldsymbol{\alpha} \in \mathbb{R}_+^M, \boldsymbol{\alpha} \neq \mathbf{0}$ and $\langle \boldsymbol{\gamma}^{\rm lt}(\boldsymbol{\alpha}) \rangle$ is a non-increasing function of α_m, for all $0 \leq m \leq M-1$.*

Proof. See Appendix C. □

Let u be a non-negative real random variable with cdf $F(u)$ and consider the maximization of the linear functional subject to linear constraints

$$\begin{cases} \text{Maximize} & \mathrm{E}[w(u)] \\ \text{Subject to} & 0 \leq w(u) \leq 1 \text{ and } \mathrm{E}[u\, w(u)] = \mathcal{P} \end{cases} \tag{30}$$

Then,

Lemma 3. *The solution of (30) is*

$$\widehat{w}(u) = \begin{cases} 1 & \textit{for } u < s^* \\ w^* & \textit{for } u = s^* \\ 0 & \textit{for } u > s^* \end{cases} \tag{31}$$

where for $s \in \mathbb{R}_+$, we let

$$\mathcal{P}(s) = \int_{[0,s)} u \, dF(u) \ , \qquad \overline{\mathcal{P}}(s) = \int_{[0,s]} u \, dF(u) \tag{32}$$

and where w^* is given by

$$w^* = \frac{\mathcal{P} - \mathcal{P}(s^*)}{\overline{\mathcal{P}}(s^*) - \mathcal{P}(s^*)} \tag{33}$$

with $s^* = \sup\{s : \mathcal{P}(s) < \mathcal{P}\}$.

Proof. See Appendix C. $\square$

Note that in Lemma 3 $w^* = 0$ if $\mathcal{P} = \mathcal{P}(s^*)$ and $w^* \to 1$ if $\mathcal{P} \to \overline{\mathcal{P}}(s^*)$. Moreover, if $F(u)$ is continuous in $u = s^*$, then $\overline{\mathcal{P}}(s^*) - \mathcal{P}(s^*) = s^* P(u = s^*) = 0$, so that $\{u = s^*\}$ is a set of probability measure zero and the value of w^* can be any real in $[0,1]$ without affecting neither the objective nor the constraint of (30).

For $s \in \mathbb{R}_+$ we define the regions

$$\begin{aligned} \mathcal{R}(s) &= \left\{ \boldsymbol{\alpha} \in \mathbb{R}_+^M : \left\langle \boldsymbol{\gamma}^{\rm lt}(\boldsymbol{\alpha}) \right\rangle < s \right\} \\ \overline{\mathcal{R}}(s) &= \left\{ \boldsymbol{\alpha} \in \mathbb{R}_+^M : \left\langle \boldsymbol{\gamma}^{\rm lt}(\boldsymbol{\alpha}) \right\rangle \leq s \right\} \end{aligned} \tag{34}$$

The *boundary surface* $\mathcal{B}(s)$ of $\overline{\mathcal{R}}(s)$ is the set of points $\boldsymbol{\alpha}$ such that $\left\langle \boldsymbol{\gamma}^{\rm lt}(\boldsymbol{\alpha}) \right\rangle = s$. Then, we define the two average power sums $\mathcal{P}(s)$ and $\overline{\mathcal{P}}(s)$ as

$$\mathcal{P}(s) = \int_{\mathcal{R}(s)} \left\langle \boldsymbol{\gamma}^{\rm lt}(\boldsymbol{\alpha}) \right\rangle dF(\boldsymbol{\alpha}) \ , \qquad \overline{\mathcal{P}}(s) = \int_{\overline{\mathcal{R}}(s)} \left\langle \boldsymbol{\gamma}^{\rm lt}(\boldsymbol{\alpha}) \right\rangle dF(\boldsymbol{\alpha}) \tag{35}$$

and the power sum threshold s^* by

$$s^* = \sup\{s : \mathcal{P}(s) < \mathcal{P}\} \tag{36}$$

Finally, we define the weight w^* by

$$w^* = \frac{\mathcal{P} - \mathcal{P}(s^*)}{\overline{\mathcal{P}}(s^*) - \mathcal{P}(s^*)} \tag{37}$$

We are now ready for

Proposition 4. *Problem (26) is solved by*

$$\widehat{\boldsymbol{\gamma}}(\boldsymbol{\alpha}) = \begin{cases} \boldsymbol{\gamma}^{\rm lt}(\boldsymbol{\alpha}) & \text{if } \boldsymbol{\alpha} \in \mathcal{R}(s^*) \\ \mathbf{0} & \text{if } \boldsymbol{\alpha} \notin \overline{\mathcal{R}}(s^*) \end{cases} \tag{38}$$

while if $\boldsymbol{\alpha} \in \mathcal{B}(s^*)$, $\widehat{\boldsymbol{\gamma}}(\boldsymbol{\alpha}) = \boldsymbol{\gamma}^{\rm lt}(\boldsymbol{\alpha})$ with probability w^* and $\widehat{\boldsymbol{\gamma}}(\boldsymbol{\alpha}) = \mathbf{0}$ with probability $1 - w^*$, where $\mathcal{R}(s), \overline{\mathcal{R}}(s), \mathcal{B}(s), s^*$ and w^* are defined by (34–37).

Proof. See Appendix D. □

Remark. The power allocation $\widehat{\gamma}(\boldsymbol{\alpha})$ given by Proposition 4 corresponds to setting a threshold s^* such that, if the power sum $\langle \boldsymbol{\gamma}^{\rm lt}(\boldsymbol{\alpha}) \rangle > s^*$, transmission is turned off, while if $\langle \boldsymbol{\gamma}^{\rm lt}(\boldsymbol{\alpha}) \rangle < s^*$ transmission is turned on and the power is allocated according to $\boldsymbol{\gamma}^{\rm lt}(\boldsymbol{\alpha})$, i.e., the allocation that requires the minimum power sum in order to avoid an outage event (see (27)). The threshold s^* is chosen so that the long-term power is actually equal to $\mathcal{P}$. If $F(\boldsymbol{\alpha})$ is not continuous, a randomization may be needed if $\langle \boldsymbol{\gamma}^{\rm lt}(\boldsymbol{\alpha}) \rangle = s^*$, i.e., if $\boldsymbol{\alpha} \in \mathcal{B}(s^*)$. Again, the probability w^* of transmitting in this case is chosen so that the long-term power is equal to $\mathcal{P}$.

The optimal power allocation strategy of Proposition 4 has an *economic interpretation*: If fading is very bad, the power required to compensate for it would cost too much in terms of average power. Then, it is more convenient to accept an outage and turn off transmission. In this way, we save power for compensating for more favorable fading conditions. □

The resulting minimum outage probability is given by

$$\widehat{P}_{\rm out}(R, \mathcal{P}) = 1 - w^* P(\boldsymbol{\alpha} \in \mathcal{B}(s^*)) - P(\boldsymbol{\alpha} \in \mathcal{R}(s^*)) \tag{39}$$

In order to compute (39) we need an explicit form for $\mathcal{R}(s)$. Again, it is convenient to give an explicit expression for the intersection $\mathcal{R}(s) \cap \mathcal{Q}$ and obtain the inequalities defining $\mathcal{R}(s)$ outside $\mathcal{Q}$ by permuting the coordinate axes. The region $\mathcal{Q}$ is divided into M subregions $\mathcal{W}_i$ defined by the inequalities

$$\begin{cases} \lambda^{\rm lt}(i, \boldsymbol{\alpha}) \geq 1/\alpha_m & m = 0, \ldots, i-1 \\ \lambda^{\rm lt}(i, \boldsymbol{\alpha}) < 1/\alpha_m & m = i, \ldots, M-1 \end{cases} \tag{40}$$

Since μ in (29) exists and is unique, $\{\mathcal{W}_i : i = 1, \ldots, M\}$ is a partition of $\mathcal{Q}$. Notice that, for $\boldsymbol{\alpha} \in \mathcal{W}_i$, exactly i blocks are used for transmission (i.e., are given positive instantaneous power). Thus, the intersection of $\mathcal{R}(s)$ with $\mathcal{W}_i$ is given by

$$\mathcal{R}(s) \cap \mathcal{W}_i = \left\{ \boldsymbol{\alpha} \in \mathbb{R}_+^M \ : \ \frac{i}{M} \left[\lambda^{\rm lt}(i, \boldsymbol{\alpha}) - \frac{1}{i} \sum_{m=0}^{i-1} \frac{1}{\alpha_m} \right] < s \right\} \quad \text{for } i = 1, \ldots, M \tag{41}$$

The intersections $\overline{\mathcal{R}}(s) \cap \mathcal{W}_i$ are obtained by replacing $<$ by $\leq$.

By comparing (25) and (41) it is immediate to see that (although the regions $\mathcal{V}_i$ and $\mathcal{W}_i$ do not coincide) the region $\mathbb{R}_+^M - \overline{\mathcal{R}}(s^*)$ is actually equal to $\mathcal{U}(R, \mathcal{P})$ defined by (25), calculated for $\mathcal{P} = s^*$. Then, we may define a more general outage region $\mathcal{U}(R, s) = \mathbb{R}_+^M - \overline{\mathcal{R}}(s)$ (or equivalently, by using

(25) with $\mathcal{P} = s$). If $F(\boldsymbol{\alpha})$ is continuous, the minimum outage probability under either the short-term or long-term constraint is given by the same integral

$$\int_{\mathcal{U}(R,s)} dF(\boldsymbol{\alpha}) \tag{42}$$

computed for $s = \mathcal{P}$ (short-term) and for $s = s^*$ (long-term). If $F(\boldsymbol{\alpha})$ has discontinuities, $\mathcal{B}(s^*)$ may have positive probability measure. Then, randomization with probability w^* on the boundary must be taken into account when computing the outage probability (with long-term constraint only).[7]

3.3 Minimum outage probability under both long-term and short-term constraints

Consider the problem

$$\left\{ \begin{array}{ll} \text{Minimize} & P(I_M(\boldsymbol{\alpha},\boldsymbol{\gamma}) < R) \\ \text{Subject to} & \mathrm{E}[\langle\boldsymbol{\gamma}\rangle] \leq \mathcal{P} \\ \text{and to} & \langle\boldsymbol{\gamma}\rangle \leq \mathcal{P}^* \quad \text{with probability 1} \end{array} \right. \tag{43}$$

where $\mathcal{P}^* > \mathcal{P}$, otherwise the first constraint is irrelevant. Proposition 4 can be easily modified in order to provide the solution of (43). We have

Proposition 5. *Problem (43) is solved by*

$$\widehat{\boldsymbol{\gamma}}(\boldsymbol{\alpha}) = \left\{ \begin{array}{ll} \boldsymbol{\gamma}^{\mathrm{lt}}(\boldsymbol{\alpha}) & \textit{if } \boldsymbol{\alpha} \in \mathcal{R}(\widehat{s}) \\ \mathbf{0} & \textit{if } \boldsymbol{\alpha} \notin \overline{\mathcal{R}}(\widehat{s}) \end{array} \right. \tag{44}$$

where $\widehat{s} \stackrel{\Delta}{=} \min\{s^*, \mathcal{P}^*\}$. *For* $\boldsymbol{\alpha} \in \mathcal{B}(\widehat{s})$ *we distinguish two cases: i) If* $\widehat{s} = s^*$, *then* $\widehat{\boldsymbol{\gamma}}(\boldsymbol{\alpha}) = \boldsymbol{\gamma}^{\mathrm{lt}}(\boldsymbol{\alpha})$ *with probability* w^* *and* $\widehat{\boldsymbol{\gamma}}(\boldsymbol{\alpha}) = \mathbf{0}$ *with probability* $1 - w^*$, *where* w^* *is given by (37). ii) If* $\widehat{s} = \mathcal{P}^*$, *then* $\widehat{\boldsymbol{\gamma}}(\boldsymbol{\alpha}) = \boldsymbol{\gamma}^{\mathrm{lt}}(\boldsymbol{\alpha})$ *with probability 1.*

Proof. In essence, the above proposition states that one of the two constraints in (43) is always redundant, so that it is always possible to minimize the outage probability with respect to one constraint while satisfying the other. This is obviously a solution of the problem.

If $s^* \leq \mathcal{P}^*$, then $\langle\widehat{\boldsymbol{\gamma}}(\boldsymbol{\alpha})\rangle \leq s^* \leq \mathcal{P}^*$ for all $\boldsymbol{\alpha}$, so that the short-term constraint is automatically satisfied, and since in this case $\widehat{s} = s^*$, $\widehat{\boldsymbol{\gamma}}(\boldsymbol{\alpha})$ coincides with (38), which minimizes the outage probability under the long-term constraint.

[7] Even though typical fading models have continuous cdf, there are practical cases of interest where the fading distribution appears as discrete. This occurs for example when the transmitter can use only a coarsely quantized information about the fading levels, or in mobile satellite systems when the fading is modeled as an on-off process depending on the presence or absence of a line-of-sight propagation path [24].

If $\mathcal{P}^* < s^*$, then from the definition (36) of s^* it follows that $\mathrm{E}[\langle\widehat{\boldsymbol{\gamma}}(\alpha)\rangle] = \overline{\mathcal{P}}(\mathcal{P}^*) \leq \mathcal{P}$, so that the long-term constraint is automatically satisfied. Moreover, the outage probability with power allocation $\widehat{\boldsymbol{\gamma}}(\boldsymbol{\alpha})$ is given by $\int_{\mathcal{U}(R,\mathcal{P}^*)} dF(\boldsymbol{\alpha})$, which is the minimum outage probability with short-term constraint $\mathcal{P}^*$ (this follows from the fact that $\mathcal{U}(R,s) = \mathbb{R}_+^M - \overline{\mathcal{R}}(s)$, as shown above). Thus, $\widehat{\boldsymbol{\gamma}}(\boldsymbol{\alpha})$ minimizes the outage probability under the short-term constraint. □

By letting $\mathcal{P} \to \infty$ in (43), so that the long-term constraint becomes irrelevant, Proposition 5 provides an alternative solution to the minimum outage probability problem with short-term constraint. This differs substantially from the solution of Proposition 3. In fact, the solution of Proposition 3 maximizes the instantaneous mutual information by keeping a constant power sum equal to the short-term power, while the solution of Proposition 5 minimizes the power sum by keeping a constant mutual information equal to the code rate. Also, note that when $\mathcal{P} \to \infty$, the zero power allocation in (44) can be replaced by any arbitrary function $\mathbf{g}(\boldsymbol{\alpha})$ such that $\langle \mathbf{g}(\boldsymbol{\alpha})\rangle = \mathcal{P}^*$.

3.4 General solution for the delay-limited capacity

From (42), the zero-outage condition is achieved when $\int_{\mathcal{U}(R,s)} dF(\boldsymbol{\alpha}) = 0$, where $s = \mathcal{P}$ for the short-term constraint and $s = s^*$ for the long-term constraint. In the latter case, if $F(\boldsymbol{\alpha})$ has discontinuities and $\mathcal{B}(s^*)$ has a positive probability measure, the condition $w^* = 1$ must be added. This means that transmission is never turned off. Define $\delta = \inf\{|\boldsymbol{\alpha}| : F(\boldsymbol{\alpha}) > 0\}$. If $\delta > 0$, there exists $s_\delta < \infty$ such that, for all $s \geq s_\delta$, $\mathcal{U}(R,s)$ is contained in a ball of radius δ centered at the origin. Therefore, $\int_{\mathcal{U}(R,s)} dF(\boldsymbol{\alpha}) = 0$ for all $s \geq s_\delta$. In this case, the rate R can be reliably transmitted with short-term power s_δ and long-term power $\overline{\mathcal{P}}(s_\delta)$. Then, (R, s_δ) and $(R, \overline{\mathcal{P}}(s_\delta))$ are points on the delay-limited capacity curves subject to the short-term and to the long-term constraints, respectively.

If $\delta = 0$, the delay-limited capacity subject to the short-term constraint is zero, while the delay-limited capacity subject to the long-term constraint can be computed from the limit $\overline{\mathcal{P}}_\infty \triangleq \lim_{s\to\infty} \overline{\mathcal{P}}(s)$. If $\overline{\mathcal{P}}_\infty < \infty$, reliable transmission at rate R is possible with minimum required long-term power $\overline{\mathcal{P}}_\infty$. Then, $(R, \overline{\mathcal{P}}_\infty)$ is a point on the delay-limited capacity curve.

3.5 Adaptive power allocation

For given M and $\mathcal{P}$, the optimal power allocation of Proposition 4 depends on the fading statistics only through the threshold s^*. For known fading statistics, this value can be computed in advance. However, in actual implementations the fading statistics may not be known *a priori*. In this case, the threshold s^* must be estimated from the fading samples. A simple heuristic algorithm is the following:

1. Initialize $s_0^* = \mathcal{P}$.

2. For $k = 1, 2, \ldots$, let

$$s_k^* = s_{k-1}^* \left[1 + \epsilon \left(\mathcal{P} - \frac{1}{k} \sum_{i=0}^{k-1} \left\langle \widehat{\boldsymbol{\gamma}}(\boldsymbol{\alpha}^{(i)}) \right\rangle \right) \right] \tag{45}$$

   and use the threshold s_k^* to determine $\widehat{\boldsymbol{\gamma}}(\boldsymbol{\alpha}^{(k)})$, given by (38) with $\boldsymbol{\alpha} = \boldsymbol{\alpha}^{(k)}$ and $s^* = s_k^*$.

The above algorithm is motivated by the following argument. Let $u = \langle \boldsymbol{\gamma}^{\text{lt}}(\boldsymbol{\alpha}) \rangle$, and assume that its cdf $F(u)$ is continuous and increasing, so that the pdf $f(u)$ exists and is positive for all $u > 0$ (from Lemma 2, a necessary and sufficient condition for the existence and positivity of $f(u)$ is that the fading has continuous cdf and takes values of the whole $\mathbb{R}_+^M$. The most important fading models, like Rayleigh, Rice, Nakagami and Log-normal [10] belong to this class). We consider the cost function $G(s) = (\mathcal{P} - \mathcal{P}(s))^2$, where $\mathcal{P}(s)$ defined in (35) can be rewritten as

$$\mathcal{P}(s) = \int_0^s u f(u)\, du$$

Since $\mathcal{P}(s)$ is an increasing function of s, it is straightforward to see that $G(s)$ has a unique minimum at s^* solution of $\mathcal{P}(s) = \mathcal{P}$. (This solution does not exist only if $\mathcal{P} > \lim_{s\to\infty} \mathcal{P}(s)$, in which case s^* is infinite and the optimal power allocation is trivially obtained by transmitting always with power allocation $\boldsymbol{\gamma}^{\text{lt}}(\boldsymbol{\alpha})$ without any threshold. Also in this case (45) yields the correct estimate: in fact, $s_k^* \to \infty$ as $k \to \infty$.) Moreover, $G(s)$ is convex for s close to s^*. The unique minimum s^* can be approximated by using the recursion

$$s_k^* = s_{k-1}^* - \epsilon \left. \frac{d}{ds} G(s) \right|_{s = s_{k-1}^*} \tag{46}$$

for a sufficiently small step size ϵ. The derivative of $G(s)$ is given by

$$\frac{d}{ds} G(s) = -2 s f(s) \left(\mathcal{P} - \int_0^s u f(u) du \right) \tag{47}$$

The integral in (47) is the (ensemble) average transmitted power with the long-term power allocation with threshold s. By invoking ergodicity, it can be approximated by the time average

$$\int_0^s u f(u) du \approx \frac{1}{k} \sum_{i=0}^{k-1} \left\langle \widehat{\boldsymbol{\gamma}}(\boldsymbol{\alpha}^{(i)}) \right\rangle \tag{48}$$

The term $f(s)$ in (47) is unknown. However, since it is positive it can be replaced by a positive constant without changing the sign of $dG(s)/ds$. Finally, by using (48) in (47), by replacing $f(s)$ by 1/2 and by substituting the approximate derivative into (46) we obtain (45).

The convergence of the algorithm (for a decreasing sequence of step sizes) is not guaranteed because $G(s)$ is not globally convex. However, with the initialization $s_0^* = \mathcal{P}$, divergent behaviors or convergence to local minima were never observed in our experiments. As usual, the step size must be chosen by trading acquisition time for tracking precision. A gear-shifting strategy which uses a step size in the acquisition mode and another one in the tracking mode might be a good option. In order to track a slowly-varying fading statistics, the algorithm may be modified by using an exponentially decaying window with an appropriate forgetting factor. An interesting feature of the above algorithm is that initially, i.e., for small k, $s_k^* \simeq \mathcal{P}$, so that a short-term power constraint is applied. As long as k increases, and hence the estimate of the fading statistics improves, s_k^* approaches s^*, and a long-term power constraint is applied.

3.6 Repetition diversity

We conclude this section by exhibiting expressions for the outage probability and for the optimal power allocation with repetition diversity. The instantaneous mutual information of the repetition diversity channel with gains $\boldsymbol{\alpha}$ and power allocation $\boldsymbol{\gamma}$ and with optimal (maximal-ratio) combining at the receiver is given by

$$I_M^{\text{rep}}(\boldsymbol{\alpha},\boldsymbol{\gamma}) = \frac{1}{2M}\log\left(1+\sum_{m=0}^{M-1}\alpha_m\gamma_m\right) \tag{49}$$

For constant-power transmission this reduces to $\frac{1}{2M}\log(1+\mathcal{P}\sum_{m=0}^{M-1}\alpha_m)$. The outage probability is again defined as the probability $P(I_M^{\text{rep}}(\boldsymbol{\alpha},\boldsymbol{\gamma}) < R)$. In the following we let $\alpha_{\max} \stackrel{\Delta}{=} \max_\ell \alpha_\ell$. By the same arguments developed in the previous two sections, it is immediate to show that the optimal power allocation functions subject to the short-term and the long-term power constraints are given by

- Short-term: for $m = 0, \ldots, M-1$ let

$$\widehat{\gamma}_m = \begin{cases} M\mathcal{P} & \text{if} \quad \alpha_m = \alpha_{\max} \\ 0 & \text{otherwise} \end{cases}$$

- Long-term: for $m = 0, \ldots, M-1$ let

$$\widehat{\gamma}_m = \begin{cases} (e^{2MR}-1)/\alpha_m & \text{if } \alpha_m = \alpha_{\max} > \rho^* \\ 0 & \text{otherwise} \end{cases}$$

where we define, in analogy with (35),

$$\mathcal{P}^{\rm rep}(\rho) = \int_{(\rho,\infty)} \frac{e^{2MR}-1}{\alpha_{\max}} dF(\alpha_{\max}) \ , \qquad \overline{\mathcal{P}}^{\rm rep}(\rho) = \int_{[\rho,\infty)} \frac{e^{2MR}-1}{\alpha_{\max}} dF(\alpha_{\max})$$

and where the threshold ρ^* is given by $\rho^* = \inf\{\rho : \mathcal{P}^{\rm rep}(\rho) < \mathcal{P}\}$. If $F(\alpha_{\max})$ has discontinuities, randomization might be needed if $\alpha_{\max} = \rho^*$. In this case, $\widehat{\gamma}_m = (e^{2MR}-1)/\alpha_m$ with probability w^* if $\alpha_m = \alpha_{\max} = \rho^*$, and zero elsewhere. The value of w^* is given by

$$w^* = \frac{\mathcal{P} - \mathcal{P}^{\rm rep}(\rho^*)}{\overline{\mathcal{P}}^{\rm rep}(\rho^*) - \mathcal{P}^{\rm rep}(\rho^*)}$$

In analogy with the case of optimal coding, we can define the delay-limited capacity with repetition diversity (see (15)). The same discussion carried out in Section 3.4 applies in this case. In particular, if $\delta = \inf\{\alpha_{\max} : F(\alpha_{\max}) > 0\} = 0$, the delay-limited capacity subject to the long-term constraint can be computed from the limit $\overline{\mathcal{P}}_0^{\rm rep} = \lim_{\rho\to 0} \overline{\mathcal{P}}^{\rm rep}(\rho)$. If $\overline{\mathcal{P}}_0^{\rm rep} < \infty$, reliable transmission at rate R is possible with minimum required long-term power $\overline{\mathcal{P}}_0^{\rm rep}$. Then, $(R, \overline{\mathcal{P}}_0^{\rm rep})$ is a point on the delay-limited capacity curve.

The optimal power allocation with either short-term or long-term constraint corresponds to selection diversity at the transmitter. Thus, selection combining and maximal-ratio combining at the receiver (with the assumption of perfect CSI) are equivalent, since the signal is transmitted over only one of the blocks. By looking at the M-block channel as a simplified model for spread-spectrum signals over a frequency selective channel [5, 17, 12], this shows that it is expedient to spend all the available power in a single subband, rather than spreading the signal power over the entire channel bandwidth.[8]

4 Applications

In this section we apply the results previously derived to two examples of channel statistics: the independent on-off and the independent Rayleigh BF-AWGN channels. In addition, a closed-form expression for the delay-limited capacity in the case of Rayleigh fading with $M = 2$ blocks is obtained.

4.1 On-off fading

A simple fading model for the M-block BF-AWGN channel assumes that each block is either totally faded or unfaded, independently of the others. The gains α_m are Bernoulli i.i.d. with $P(\alpha_m = 0) = p$.

[8] With this observation we do not intend to question the merits of classical CDMA. For example, the selection diversity technique proposed here would be easy to intercept by eavesdroppers. Moreover, CDMA is used in multiuser systems, where multiple-access interference rather than fading is the main impairment.

We refer to this channel as *on-off BF-AWGN*.[9]

Under the short-term average power constraint, the optimal power allocation strategy is

$$\widehat{\gamma}_m(\boldsymbol{\alpha}) = \begin{cases} \frac{M}{\mu}\mathcal{P} & \text{if } \alpha_m = 1 \\ 0 & \text{if } \alpha_m = 0 \end{cases} \tag{50}$$

where $\mu = \sum_{m=0}^{M-1} \alpha_m$ is the Hamming weight of $\boldsymbol{\alpha}$, binomially distributed with parameters (M, p). The resulting minimum outage probability is

$$\widehat{P}_{\text{out}}(R, \mathcal{P}) = P(\mu < \mu_0)$$

where μ_0 is the smallest integer for which $\frac{\mu}{2M} \log\left(1 + \frac{M\mathcal{P}}{\mu}\right) \geq R$. Under the long-term average power constraint, the regions $\mathcal{R}(s)$ can be defined in terms of the weight μ of $\boldsymbol{\alpha}$ as

$$\mathcal{R}(s) = \left\{\boldsymbol{\alpha} \in \{0,1\}^M : \frac{\mu}{M}(e^{2MR/\mu} - 1) < s\right\}$$

Let $\mu(s)$ be the largest μ such that $\frac{\mu}{M}(e^{2MR/\mu} - 1) \geq s$. We have

$$\begin{aligned} \mathcal{P}(s) &= \sum_{\mu=\mu(s)+1}^{M} \binom{M}{\mu} p^{M-\mu}(1-p)^{\mu} \frac{\mu}{M}(e^{2MR/\mu} - 1) \\ \overline{\mathcal{P}}(s) &= \sum_{\mu=\mu(s)}^{M} \binom{M}{\mu} p^{M-\mu}(1-p)^{\mu} \frac{\mu}{M}(e^{2MR/\mu} - 1) \end{aligned}$$

From the above expressions, using (36) and (37) we can determine s^* and w^* for every pair $(R, \mathcal{P})$. The solution $\boldsymbol{\gamma}^{\text{lt}}(\boldsymbol{\alpha})$ of (27) is given by

$$\gamma_m^{\text{lt}}(\boldsymbol{\alpha}) = \begin{cases} e^{2MR/\mu} - 1 & \text{if } \alpha_m = 1 \\ 0 & \text{if } \alpha_m = 0 \end{cases}$$

Finally, by letting $\mu^* = \mu(s^*)$, the optimal power allocation strategy given by Propositions 4 can be written as

$$\widehat{\boldsymbol{\gamma}}(\boldsymbol{\alpha}) = \begin{cases} \boldsymbol{\gamma}^{\text{lt}}(\boldsymbol{\alpha}) & \text{if } \mu > \mu^* \\ \mathbf{0} & \text{if } \mu < \mu^* \end{cases} \tag{51}$$

[9] This model does not represent only an extreme simplification of fading. Consider for example a multicarrier multimedia network where fixed-rate services (e.g., voice and video) share several subcarriers with variable-rate bursty services (e.g., file transfer). A common centralized multiple-access protocol allocates the subcarriers to the different services. A fixed-rate service transmits information at rate R and makes use of M subcarriers, but some of these subcarriers may be assigned dynamically to some other variable-rate service with higher priority, with a certain probability. We are interested in the minimum outage probability of fixed-rate users in such system.

while if $\mu = \mu^*$ we transmit with power allocation $\boldsymbol{\gamma}^{\mathrm{lt}}(\boldsymbol{\alpha})$ with probability w^* and we turn off transmission with probability $1 - w^*$. The resulting minimum outage probability is

$$\widehat{P}_{\mathrm{out}}(R, \mathcal{P}) = P(\mu < \mu^*) + (1 - w^*)P(\mu = \mu^*)$$

The outage probability with constant power is given by

$$P_{\mathrm{out}}(R, \mathcal{P}) = P\left(\mu < \frac{2MR}{\log(1 + \mathcal{P})}\right) \tag{52}$$

With repetition diversity we obtain the following outage probabilities:

- Repetition diversity with constant power:

$$P_{\mathrm{out}}(R, \mathcal{P}) = P\left(\mu < \frac{e^{2MR} - 1}{\mathcal{P}}\right) \tag{53}$$

- Repetition diversity with optimal short-term power allocation:

$$P_{\mathrm{out}}(R, \mathcal{P}) = \begin{cases} 1 & \text{if } \frac{1}{2M}\log(1 + M\mathcal{P}) < R \\ p^M & \text{else} \end{cases} \tag{54}$$

- Repetition diversity with optimal long-term power allocation:

$$P_{\mathrm{out}}(R, \mathcal{P}) = \begin{cases} 1 - \frac{M\mathcal{P}}{e^{2MR}-1} & \text{if } \frac{1}{2M}\log(1 + \frac{M\mathcal{P}}{1-p^M}) < R \\ p^M & \text{else} \end{cases} \tag{55}$$

Fig. 3 shows the outage probability vs. $\mathcal{P}$ in the cases discussed above for $p = 0.1$, $M = 8$ and $R = 0.5$ bit/symbol. In all cases, the asymptotic outage floor as $\mathcal{P} \to \infty$ is equal to p^M. The delay-limited capacity of the on-off BF-AWGN channel is zero: in fact $P_{\mathrm{out}}(R, \mathcal{P}) \geq p^M$.

4.2 Rayleigh fading with $M = 1$ and $M = 2$

We consider now a BF-AWGN channel modeled as a set of M independent Rayleigh fading channels. The vector $\boldsymbol{\alpha}$ of the fading gains has independent exponentially-distributed components with mean 1.

Case $M = 1$. For both constant and short-term power transmission, the outage probability is given by [1]

$$\widehat{P}_{\mathrm{out}}(R, \mathcal{P}) = 1 - \exp(-(e^{2R} - 1)/\mathcal{P})$$

(notice that for $M = 1$, constant and optimal power allocation under the short-term constraint always coincide). With optimal power allocation and long-term constraint, s^* is the solution of

$$(e^{2R} - 1)\mathrm{Ei}\left(1, \frac{e^{2R} - 1}{s}\right) = \mathcal{P}$$

and the resulting outage probability is

$$\widehat{\mathcal{P}}_{\text{out}}(R, \mathcal{P}) = 1 - \exp(-(e^{2R} - 1)/s^*)$$

The corresponding power allocation function is given by

$$\widehat{\gamma}(\alpha_0) = \begin{cases} (e^{2R} - 1)/\alpha_0 & \alpha_0 > s^* \\ 0 & \text{otherwise} \end{cases}$$

Fig. 4 shows the above outage probabilities vs. $\mathcal{P}$ for $R = 0.5$ bit/symbol. We observe that optimal power allocation with long-term constraint decreases dramatically the outage probability with respect to constant-power transmission (the average-power saving is about 22 dB at $P_{\text{out}} = 10^{-3}$). This result differs conspicuously from the one obtained from capacity calculations, which show a very small gain offered by optimal power allocation (see Fig. 1). However, as anticipated in Section 2, the delay-limited capacity of this channel for $M = 1$ is zero since $\overline{\mathcal{P}}_\infty = \infty$ for all $R > 0$.

Case $M = 2$. From the inequalities (41) defining $\mathcal{R}(s)$, we obtain the curve $\mathcal{B}(s)$ (the boundary of the region $\overline{\mathcal{R}}(s)$), which for $\alpha_0 \geq \alpha_1$ is given by

$$\begin{aligned} \alpha_1 &= \beta(\alpha_0, s) \quad \text{for} \quad \alpha_0 \leq e^{4R}\alpha_1 \\ \alpha_0 &= x_1 \quad\quad\quad \text{for} \quad \alpha_0 > e^{4R}\alpha_1 \end{aligned}$$

We define $x_0 = (e^{2R} - 1)/s$, $x_1 = (e^{4R} - 1)/(2s)$ and

$$\beta(\alpha_0, s) = \alpha_0 \left(\frac{e^{2R} + \sqrt{e^{4R} - 1 - 2s\alpha_0}}{1 + 2s\alpha_0}\right)^2$$

For $\alpha_0 < \alpha_1$, the equations defining $\mathcal{B}(s)$ are obtained by exchanging α_0 and α_1. Fig. 5 shows $\mathcal{B}(s)$ for $R = 0.5$ and $s = 0, 5$ and 10 dB. As s increases, the outage region $\mathcal{U}(R, s)$ (i.e., the region below the curve $\mathcal{B}(s)$) shrinks but keeps the same shape.

In order to compute the outage probability with short-term power constraint, we have to integrate $f(\boldsymbol{\alpha}) = e^{-\alpha_0 - \alpha_1}$ over $\mathcal{U}(R, \mathcal{P})$. We obtain

$$\widehat{P}_{\text{out}}(R, \mathcal{P}) = \left[1 + e^{-2x_0} - 2e^{-x_1} - 2\int_{x_0}^{x_1} e^{-z - \beta(z,s)}\, dz\right]_{s=\mathcal{P}} \quad (56)$$

As for the outage probability with long-term power constraint, we compute first $\mathcal{P}(s)$ as a function of s, then, for each value of $\mathcal{P}$, we compute s^* as the solution of $\mathcal{P}(s) = \mathcal{P}$. Finally, we obtain $\widehat{P}_{\rm out}(R, \mathcal{P})$ from (56) evaluated at $s = s^*$. From (35) we get

$$\begin{aligned}
\mathcal{P}(s) \quad = \quad & (e^{4R}-1)\left(\mathrm{Ei}\,(1,x_1) - \mathrm{Ei}\left(1, \frac{\sinh(4R)}{s}\right)\right) + \pi e^{2R}\left(1-\mathrm{erf}(\sqrt{x_0})^2\right) - \\
& -\mathrm{Ei}\left(1, x_1 e^{-4R}\right) e^{-x_1} + \mathrm{Ei}\,(1,x_0) e^{-x_0} - \\
& -\left\{ 2\sqrt{\pi} e^{2R}\left(\int_{x_0}^{x_1} \frac{e^{-z}\mathrm{erf}(\sqrt{\beta(z,s)})}{\sqrt{z}} dz + \int_{x_1}^{\infty} \frac{e^{-z}\mathrm{erf}(\sqrt{z}e^{-2R})}{\sqrt{z}} dz\right) + \right. \\
& \left. + \int_{x_0}^{x_1}\left(\frac{1}{z}e^{-\beta(z,s)} + \mathrm{Ei}\,(1,\beta(z,s))\right) e^{-z} dz \right\}
\end{aligned} \tag{57}$$

These integrals, as well as the equation $\mathcal{P}(s) = \mathcal{P}$, must be solved numerically. Fig. 6 shows $\mathcal{P}(s)$ vs. s for $R = 0.5$ bit/symbol. Both the result of (57) and of Monte Carlo simulation are included, and show perfect agreement. The horizontal asymptote is the delay-limited capacity power limit.

The optimum power allocation function is given explicitly as follows:

$$\left\{\begin{array}{llllll}
\text{If} & 0 < \alpha_0 < e^{-4R}\alpha_1 & \text{and} & \alpha_1 > (e^{4R}-1)/2s^* & \text{then:} & \widehat{\gamma}_0 = 0 \qquad \widehat{\gamma}_1 = (e^{4R}-1)/\alpha_1 \\
 & & \text{and} & \alpha_1 < (e^{4R}-1)/2s^* & & \widehat{\gamma}_0 = 0 \qquad \widehat{\gamma}_1 = 0 \\
\text{If} & e^{-4R}\alpha_1 < \alpha_0 < \alpha_1 & \text{and} & \alpha_0 > \beta(\alpha_1, s^*) & \text{then:} & \widehat{\gamma}_0 = \frac{e^{2R}}{\sqrt{\alpha_0\alpha_1}} - \frac{1}{\alpha_0} \qquad \widehat{\gamma}_1 = \frac{e^{2R}}{\sqrt{\alpha_0\alpha_1}} - \frac{1}{\alpha_1} \\
 & & \text{and} & \alpha_0 < \beta(\alpha_1, s^*) & & \widehat{\gamma}_0 = 0 \qquad \widehat{\gamma}_1 = 0 \\
\text{If} & \alpha_1 < \alpha_0 < e^{4R}\alpha_1 & \text{and} & \alpha_1 > \beta(\alpha_0, s^*) & \text{then:} & \widehat{\gamma}_0 = \frac{e^{2R}}{\sqrt{\alpha_0\alpha_1}} - \frac{1}{\alpha_0} \qquad \widehat{\gamma}_1 = \frac{e^{2R}}{\sqrt{\alpha_0\alpha_1}} - \frac{1}{\alpha_1} \\
 & & \text{and} & \alpha_1 < \beta(\alpha_0, s^*) & & \widehat{\gamma}_0 = 0 \qquad \widehat{\gamma}_1 = 0 \\
\text{If} & \alpha_0 > e^{4R}\alpha_1 & \text{and} & \alpha_0 > (e^{4R}-1)/2s^* & \text{then:} & \widehat{\gamma}_0 = (e^{4R}-1)/\alpha_0 \qquad \widehat{\gamma}_1 = 0 \\
 & & \text{and} & \alpha_0 < (e^{4R}-1)/2s^* & & \widehat{\gamma}_0 = 0 \qquad \widehat{\gamma}_1 = 0
\end{array}\right.$$

The outage probability for constant power allocation in the general case of M independent blocks is immediately calculated from the M-fold convolution of the distribution of $\log(1+\alpha\mathcal{P})$ [1]. With repetition diversity we obtain explicit expressions for the outage probability for any M:

- Repetition diversity with constant power:

$$P_{\rm out}(R, \mathcal{P}) = 1 - \Pi_{M-1}\left(\frac{e^{2MR}-1}{\mathcal{P}}\right) \tag{58}$$

- Repetition diversity with optimal short-term power allocation:

$$P_{\rm out}(R, \mathcal{P}) = \left(1 - \exp\left(-\frac{e^{2MR}-1}{M\mathcal{P}}\right)\right)^M \tag{59}$$

- Repetition diversity with optimal long-term power allocation:

$$P_{\rm out}(R,\mathcal{P}) = (1-\exp(-\rho^*))^M \tag{60}$$

where ρ^* is the solution of

$$\mathcal{P}(\rho) = (e^{2MR}-1)\sum_{k=0}^{M-1}\binom{M-1}{k}(-1)^k \mathrm{Ei}\,(1,(k+1)\rho) = \mathcal{P}$$

Figs. 7, 8, and 9 show the above outage probabilities vs. $\mathcal{P}$ for $R = 0.5$, 1.0 and 2.0 bit/symbol. The power gain obtained by optimal long-term power allocation with respect to constant power is still very large, especially for large R, but it is reduced with respect to the case $M = 1$ (e.g., for $R = 0.5$ it is about 15.5 dB at $P_{\rm out} = 10^{-3}$). This fact can be explained by observing that, as $M \to \infty$, both outage probabilities for optimal and for constant power allocations converge to their capacity limits, i.e., they are close to 1 if $R > C_{\rm opt}(\mathcal{P})$ (resp., if $R > C_{\rm const}(\mathcal{P})$) and zero if $R < C_{\rm opt}(\mathcal{P})$ (resp., if $R < C_{\rm const}(\mathcal{P})$). Since $C_{\rm opt}(\mathcal{P})$ and $C_{\rm const}(\mathcal{P})$ are very close, we expect that the gain of optimal with respect to constant power allocation reduces as M increases. These limits hold under the assumption that the sequence $\boldsymbol{\alpha}$ is asymptotically ergodic as $M \to \infty$. Moreover, it is immediate to see that both short-term and long-term power allocation yield the $C_{\rm opt}(\mathcal{P})$ capacity limit (in fact, as $M \to \infty$ the two constraints coincide), while constant power yields $C_{\rm const}(\mathcal{P})$. A simple direct proof of convergence, based on the analysis of the limiting outage probability as $M \to \infty$ can be given irrespective of the fading statistics as long as ergodicity holds (we skip this for the sake of brevity).

Optimal power allocation under the short-term constraint does not provide any significant gain with respect to constant power. As M increases, the outage probability with short-term constraint is always very close to that for constant power. This helps understanding why optimal power allocation does not yield significant gains in terms of capacity. On the other hand, optimal power allocation under the long-term constraint provides dramatic power gains for small M (i.e., in strictly delay-limited conditions.). Long-term power allocation yields zero outage probability for $\mathcal{P} \geq \overline{\mathcal{P}}_\infty$ (optimal coding) and $\mathcal{P} \geq \overline{\mathcal{P}}_0^{\rm rep}$ (repetition diversity). These power thresholds (denoted by the vertical solid lines in the figures) are the delay-limited capacity limits, i.e., the average powers required to transmit reliably at rate R over the 2-block BF-AWGN channel for optimal coding and repetition diversity, respectively.

We can compute explicitly the delay-limited capacity limits by letting $s^* \to \infty$ (optimal coding) and $\rho^* \to 0$ (repetition diversity). From the asymptotic expansion [25]

$$\mathrm{Ei}\,(1,z) = -\Gamma - \log z + O(z), \qquad z \to 0$$

(Γ denotes the Euler's constant) and from the fact that the integrals in the expression (57) of $\mathcal{P}(s)$ vanish as $s \to \infty$, we find:

- Optimal coding:

$$\overline{\mathcal{P}}_{\infty} = (e^{4R} - 1) \log\left(1 + e^{-4R}\right) + e^{2R}(\pi - 4 \arctan(e^{-2R})) - 4R$$

- Repetition diversity:

$$\overline{\mathcal{P}}_0^{\text{rep}} = (e^{4R} - 1) \log 2$$

Fig. 10 shows the delay-limited capacity and, for comparison, the AWGN capacity and the capacity of a Rayleigh fading channel with optimal and constant power allocation. We could not obtain a closed-form evaluation of the delay-limited capacity for $M > 2$. However, this can be computed by Monte Carlo simulation of $\mathcal{P}(s)$ for very large s and different values of R. Results for $M = 4$ are shown in Fig. 10. With optimal coding, the delay-limited capacity is only 5 dB away from the AWGN capacity and 2.5 dB from the Rayleigh capacity for high rates. This gap is further reduced at low rates. For $M = 4$, these gaps are reduced by 1.2 dB (high rates). Repetition diversity suffers from a large delay-limited capacity penalty, especially for high rates, while it yields good results for low rates. This result shows that transmitter CSI and optimal power allocation may have a substantial impact on the performance of delay-limited systems, and that reliable transmission can be achieved by allowing for a very small delay ($M = 2, 4$) and a small power penalty with respect to unconstrained-delay transmission ($M \to \infty$).

Finally, Fig. 11 shows the long-term average transmitted power when algorithm (45) is used to estimate the optimal threshold s^*. This curve was obtained by Monte Carlo simulation with $\mathcal{P} = 2$ dB, $M = 2$, and $R = 0.5$ bit/symbol. We used a constant step size $\epsilon = 0.01$. The algorithm is able to approach the desired long-term power even if the fading statistics is totally unknown a priori.

4.3 Optimal M for repetition diversity

From the outage probability expressions (53–55) and (58–60) we observe that, for both the on-off and the Rayleigh fading statistics, there is an optimal M minimizing the outage probability with repetition diversity. This optimal M depends on R, $\mathcal{P}$, on the type of power allocation and on the fading statistics, and can be determined numerically from the outage probability expressions.

By viewing the M-block BF-AWGN as a very simple model for a frequency selective channel with M subbands and repetition diversity as spreading (see Section 1), the above fact shows that there exists an optimal spreading factor above which it is not convenient to spread the signal (at least from the single-user outage probability point of view considered here). On the contrary, when optimal coding is used, the outage probability of either constant and optimal short-term and long-term power allocation generally decreases as M increases (to be precise, the outage probability is strictly decreasing for Rayleigh fading, while it is not in the on-off channel for constant and short-term power allocation). Tables 1 and 2 show $P_{\rm out}(R, \mathcal{P})$ for some values of M, $R = 0.5$ bit/symbol and $P = 10$ dB for the on-off and the independent Rayleigh BF-AWGN channels, respectively. In Table 2, the zero outage values correspond to M, $\mathcal{P}$ and R for which $R \leq C_{\rm delay}(\mathcal{P})$.

5 Conclusions

In this paper we considered the problem of transmitting at constant rate over a slowly-varying fading Gaussian channel under a strict transmission delay constraint. As in [1, 2, 4, 13], we adopted the BF-AWGN channel model, where the transmission of a code word spans a finite number M of fading blocks. However, differently from the above referenced works, we assumed that the transmitter has perfect knowledge of the fading gains over all the M blocks *before* transmitting a code word, so that it can compensate for the fading and reduce the information outage probability.

We showed that minimum outage probability can be achieved by a single-codebook, variable-power scheme, obtained by concatenating a conventional "Gaussian" code with a power-controller, which scales the power of the transmitted symbols according to an optimal power allocation strategy. This scheme is particularly appealing for practical applications, since it requires no variable-rate coding or multiplexing of several codebooks. Thus, the minimum outage probability problem reduces to finding optimal power allocation strategies.

We have derived the optimum power allocation strategy for both optimal coding and repetition diversity, under different constraints on the transmitted power. In the case of long-term power constraint, the optimal power allocation depends on the fading statistics through the value of a power threshold above which transmission must be turned off. We presented an adaptive algorithm that estimates this threshold without prior knowledge of the fading statistics. We also showed that the minimum outage probability problem is closely related to the delay-limited capacity problem, and that our solution yields the delay-limited capacity of the M-block BF-AWGN channel. This method is general, although closed-form results appear to be difficult to obtain for correlated fading and/or $M > 2$. Nevertheless, Monte Carlo simulation can be used to compute both the minimum outage probability and the delay-limited capacity in the general case.

For repetition diversity, we showed that the optimal power allocation strategy is selection diversity at the transmitter, and that there exists an optimal diversity order which minimizes the outage probability. This may lead to some interesting considerations for spread-spectrum in frequency-selective fading channels, if we view repetition diversity as a simplified model for spreading.

As an application, we presented closed-form results for a simple on-off fading model and for the independent Rayleigh fading channel with $M = 1$ and $M = 2$ blocks. Optimal power allocation under a long-term power constraint yields impressive performance improvements with respect to constant-power transmission. This shows that transmitter CSI and dynamic power allocation has a huge impact on the performance of delay-limited wireless systems, if instantaneously large power peaks are possible.

We conclude by pointing out some suggestions for further research.

- Minimum outage probability and delay-limited capacity are achieved only in the limit for $N \to \infty$. However, in practice, N must be finite and instantaneously large power peaks can last only for a few frames. Then, more practical optimization problems should consider, for example, the maximization of the error exponent of the M-block BF-AWGN channel [2, 13] for fixed and finite N [26] and the construction of actual low-complexity coding schemes suited to dynamic power allocation [27].

- A key assumption of this work is the availability of perfect CSI at the transmitter. A more realistic assumption is that the transmitter is provided with (noisy) measurements of the past fading gains and uses some form of prediction to allocate power to the current frame. Outage probability minimization in this case is another very interesting problem.

- Our outage probability minimization results can be combined with transmit antenna diversity [11, 28, 29, 30]. This combination is the subject of a forthcoming companion paper [31] (see also [32]).

Acknowledgements

The authors wish to thank the anonymous reviewers of this manuscript for their comments and suggestions.

APPENDIX

A Proof of Proposition 2

This proof is largely based on the results and notations of [14]. In particular, we indicate by liminf i.p. and by limsup i.p. the liminf and the limsup in probability, defined in [14]. Also, when an equality (resp., inequality) holds "for sufficiently large N", we use $\underset{N}{=}$ (resp., $\underset{N}{\leq}$). Moreover, when not otherwise specified, equalities (resp., inequalities) between random variables are to be intended in probability, i.e., for two sequences of random variables $\{A_N\}$ and $\{B_N\}$, $A_N \underset{N}{=} B_N$ and $A_N \underset{N}{\leq} B_N$ mean that, for arbitrary non-negative ϵ, δ and sufficiently large N, $P(|A_N - B_N| \leq \epsilon) \geq 1 - \delta$ and $P(A_N - B_N \leq -\epsilon) \geq 1 - \delta$, respectively.

Consider the M-block BF-AWGN channel and let $(\mathbf{a}, \mathbf{b})$ be a pair of input and output sequences of length MN generated according to the joint pdf $p_{\mathbf{y}|\mathbf{x},\boldsymbol{\alpha}}(\mathbf{b}|\mathbf{a}) q_{\mathbf{x}|\boldsymbol{\alpha}}(\mathbf{a})$, where

$$p_{\mathbf{y}|\mathbf{x},\boldsymbol{\alpha}}(\mathbf{b}|\mathbf{a}) = \prod_{m=0}^{M-1} \prod_{n=0}^{N-1} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left|b_{mN+n} - \sqrt{\alpha_m} a_{mN+n}\right|^2\right)$$

is the MN-th order channel transition pdf conditioned on $\boldsymbol{\alpha}$ and where $q_{\mathbf{x}|\boldsymbol{\alpha}}(\mathbf{a})$ is the MN-th order input distribution. This is also conditioned on $\boldsymbol{\alpha}$ since the transmitter has perfect CSI and can choose the input distribution accordingly.

The normalized information density of $(\mathbf{a}, \mathbf{b})$ is defined by [14]

$$i_{MN}(\mathbf{a}, \mathbf{b}) \triangleq \frac{1}{MN} \log \frac{p_{\mathbf{y}|\mathbf{x},\boldsymbol{\alpha}}(\mathbf{b}|\mathbf{a})}{p_{\mathbf{y}|\boldsymbol{\alpha}}(\mathbf{b})}$$

where $p_{\mathbf{y}|\boldsymbol{\alpha}}(\mathbf{b})$ is the marginal conditional pdf of the channel output sequence.

In order to get rid of some annoying technicalities related to the continuity of ϵ-achievable rates with respect to ϵ, in this paper we follow the alternative more "regular" definition of achievability given in [14][Sect. IV]. Namely, a rate R is ϵ-achievable if there exists a sequence of (n, M_n, ϵ_n)-codes such that $\liminf_{n\to\infty} \frac{1}{n} \log M_n \geq R$ and $\limsup_{n\to\infty} \epsilon_n \leq \epsilon$. While this has almost no impact on the operational significance of outage probability and capacity, it allows us to characterize the maximum ϵ-achievable rate (ϵ-capacity) and the channel capacity by the compact expressions [14][Th. 6]

$$\begin{aligned} C_\epsilon &= \sup_{\mathbf{X}} \sup\{R : F_I(R) \leq \epsilon\} \quad \text{for } 0 \leq \epsilon < 1 \\ C &= \lim_{\epsilon\to 0_+} C_\epsilon \end{aligned} \tag{61}$$

where $\sup_{\mathbf{X}}$ denotes the supremum over the input distributions and $F_I(R)$ is the limiting cdf, for large block length, of the normalized information density.

Armed with these results, all what we have to show is that, for any choice of the input pdf $q_{\mathbf{x}|\boldsymbol{\alpha}}(\mathbf{a})$ satisfying a given input power constraint, the limit of normalized information density cdf is lower bounded by the cdf of the instantaneous mutual information $I_M(\boldsymbol{\alpha},\boldsymbol{\gamma})$ defined in (8), and that $I_M(\boldsymbol{\alpha},\boldsymbol{\gamma})$ is the limit in probability of $i_{MN}(\mathbf{x},\mathbf{y})$ when $\mathbf{x}$ is conditionally Gaussian with $(mN+n)$-th component $\sim \mathcal{N}(0,\gamma_m)$.

Fix an arbitrary vector $\boldsymbol{\gamma} = (\gamma_0,\dots,\gamma_{M-1})$. For all $R \geq 0$, we want to show that

$$P(I_M(\boldsymbol{\alpha},\boldsymbol{\gamma}) < R) \underset{N}{\leq} P(i_{MN}(\mathbf{x},\mathbf{y}) < R) \tag{62}$$

where the input distribution is constrained by:

$$\limsup_{N\to\infty} \text{i.p.} \frac{1}{N}\sum_{n=0}^{N-1} |x_{mN+n}|^2 \leq \gamma_m \quad \text{for } m = 0,\dots,M-1 \tag{63}$$

Since the channel is conditionally memoryless given $\boldsymbol{\alpha}$, from [14][Th. 10] we obtain that

$$i_{MN}(\mathbf{x},\mathbf{y}) \underset{N}{\leq} \frac{1}{MN}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} \log \frac{p_{y|x,\boldsymbol{\alpha}}(y_{mN+n}|x_{mN+n})}{p_{y|\boldsymbol{\alpha}}(y_{mN+n})} \tag{64}$$

where $p_{y|x,\boldsymbol{\alpha}}(b_{mN+n}|a_{mN+n})$ is the 1-st order $(mN+n)$-th channel transition pdf and $p_{y|\boldsymbol{\alpha}}(b_{mN+n})$ is the 1-st order $(mN+n)$-th output marginal pdf. Since the channel has additive noise and the noise is stationary and ergodic, we can write the RHS of (64) as

$$-\frac{1}{MN}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} \log p_{y|\boldsymbol{\alpha}}(b_{mN+n}) - h(z) \tag{65}$$

for sufficiently large N, where $h(z)$ is the differential entropy of the noise. Let $\widetilde{\mathbf{y}}$ be the output sequence resulting from conditionally independent Gaussian inputs $\widetilde{x}_{mN+n} \sim \mathcal{N}(0,\gamma_m)$, and let $g_m(b_{mN+n})$ denote the resulting conditional $(mN+n)$-th marginal output pdf, given by

$$g_m(b_{mN+n}) = \frac{1}{\sqrt{2\pi(\alpha_m\gamma_m+1)}} \exp\left(-\frac{|b_{mN+n}|^2}{2(\alpha_m\gamma_m+1)}\right)$$

We can write

$$\begin{aligned} \frac{1}{MN}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} \log \frac{g_m(\widetilde{y}_{mN+n})}{p_{y|\boldsymbol{\alpha}}(y_{mN+n})} &\underset{N}{\leq} \frac{1}{MN}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} \log \frac{g_m(y_{mN+n})}{p_{y|\boldsymbol{\alpha}}(y_{mN+n})} \\ &\underset{N}{\leq} 0 \end{aligned} \tag{66}$$

where the first inequality of (66) holds since

$$\frac{1}{MN}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} |\widetilde{y}_{mN+n}|^2 \underset{N}{=} \frac{1}{M}\sum_{m=0}^{M-1} \alpha_m\gamma_m + 1$$

and, from (63),

$$\frac{1}{MN}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}|y_{mN+n}|^2 \underset{N}{\leq} \frac{1}{M}\sum_{m=0}^{M-1}\alpha_m\gamma_m + 1$$

The second inequality of (66) follows from [14][Th. 8, (a)]. By collecting (64), (65) and (66), we obtain

$$i_{MN}(\mathbf{x},\mathbf{y}) \underset{N}{\leq} i_{MN}(\widetilde{\mathbf{x}},\widetilde{\mathbf{y}})$$

where $\widetilde{\mathbf{x}}$ is an independent Gaussian sequence such that $\widetilde{x}_{mN+n} \sim \mathcal{N}(0,\gamma_m)$ and $\widetilde{\mathbf{y}}$ is the resulting output sequence. From the weak law of large numbers, it is immediate to show the following equalities

$$\begin{aligned} i_{MN}(\widetilde{\mathbf{x}},\widetilde{\mathbf{y}}) &\underset{N}{=} I_M(\boldsymbol{\alpha},\boldsymbol{\gamma}) \\ \frac{1}{N}\sum_{n=0}^{M-1}|\widetilde{x}_{mN+n}|^2 &\underset{N}{=} \gamma_m \quad \text{for } m=0,\ldots,M-1 \end{aligned}$$

Then, (62) follows. Since the γ_m's are arbitrary, condition (63) expresses any input power constraint. Hence, the supremum over $\mathbf{X}$ in (61) can be replaced without loss of optimality by the supremum over all power allocation functions $\boldsymbol{\gamma}$, having established that the optimal input distribution is conditionally independent Gaussian, entirely specified by $\boldsymbol{\gamma}$.

Finally, Proposition 2 follows from (61) by noting that $P_{\text{out}}(R,\mathcal{P})$ is the left-continuous version of $F_I(R)$, and that left or right continuity have no effect on the supremum of $\{R : P_{\text{out}}(R,\mathcal{P}) \leq \epsilon\}$.

B Proof of Proposition 3

We write the outage probability as

$$P(I_M(\boldsymbol{\alpha},\boldsymbol{\gamma}) < R) = E[\chi_{\{I_M(\boldsymbol{\alpha},\boldsymbol{\gamma})<R\}}] \tag{67}$$

Consider the solution $\boldsymbol{\gamma}^{\text{st}} = \boldsymbol{\gamma}^{\text{st}}(\boldsymbol{\alpha})$ to the problem (19). This is a deterministic function of $\boldsymbol{\alpha}$. Define

$$\chi^{\text{st}}(\boldsymbol{\alpha}) \stackrel{\Delta}{=} \chi_{\{I_M(\boldsymbol{\alpha},\boldsymbol{\gamma}^{\text{st}}(\boldsymbol{\alpha}))<R\}}$$

and notice that $\mathcal{U}(R,\mathcal{P})$ given in (20) can be written as $\mathcal{U}(R,\mathcal{P}) = \{\boldsymbol{\alpha} \in \mathbb{R}_+^M : \chi^{\text{st}}(\boldsymbol{\alpha}) = 1\}$. For any random function $\boldsymbol{\gamma}(\boldsymbol{\alpha})$ satisfying the short-term constraint, we have

$$\chi^{\text{st}}(\boldsymbol{\alpha}) \leq \chi_{\{I_M(\boldsymbol{\alpha},\boldsymbol{\gamma}(\boldsymbol{\alpha}))<R\}} \quad \text{with probability 1} \tag{68}$$

In fact, for all $\boldsymbol{\alpha}$ such that $\chi^{\text{st}}(\boldsymbol{\alpha}) = 0$, the above inequality is trivially satisfied, while for all $\boldsymbol{\alpha}$ such that $\chi^{\text{st}}(\boldsymbol{\alpha}) = 1$ also $\chi_{\{I_M(\boldsymbol{\alpha},\boldsymbol{\gamma})<R\}} = 1$ with probability 1 (otherwise, for some $\boldsymbol{\alpha}$, there would be

$\boldsymbol{\gamma}(\boldsymbol{\alpha})$ for which $I_M(\alpha, \boldsymbol{\gamma}(\boldsymbol{\alpha})) > I_M(\boldsymbol{\alpha}, \boldsymbol{\gamma}^{\rm st}(\boldsymbol{\alpha}))$, which contradicts the definition of $\boldsymbol{\gamma}^{\rm st}$). Now, let $\mathbf{g}(\boldsymbol{\alpha})$ be any random function satisfying the short-term constraint, and rewrite (18) as

$$\widehat{\boldsymbol{\gamma}}(\boldsymbol{\alpha}) = (1 - \chi^{\rm st}(\boldsymbol{\alpha}))\boldsymbol{\gamma}^{\rm st}(\boldsymbol{\alpha}) + \chi^{\rm st}(\boldsymbol{\alpha})\mathbf{g}(\boldsymbol{\alpha})$$

Clearly, $\langle \widehat{\boldsymbol{\gamma}} \rangle \leq \mathcal{P}$ with probability 1 and

$$\chi^{\rm st}(\boldsymbol{\alpha}) = \chi_{\{I_M(\boldsymbol{\alpha}, \widehat{\boldsymbol{\gamma}}(\boldsymbol{\alpha})) < R\}}$$

From the above equality, by using (68) into (67), we get

$$P(I_M(\boldsymbol{\alpha}, \boldsymbol{\gamma}) < R) \geq E[\chi^{\rm st}(\boldsymbol{\alpha})] = P(I_M(\boldsymbol{\alpha}, \widehat{\boldsymbol{\gamma}}) < R)$$

for all $\boldsymbol{\gamma}$ satisfying the short-term constraint.

C Proof of Lemmas

C.1 Proof of Lemma 1

Let us set $\gamma_m = y_m^2$, for $m = 0, \ldots, M-1$, let $\eta \in \mathbb{R}$ be the Lagrange multiplier, and consider the Lagrangian functional

$$\Phi(t) = \sum_{m=0}^{M-1} \left\{ (y_m + t\psi_m)^2 - \eta \frac{1}{2} \log(1 + \alpha_m (y_m + t\psi_m)^2) \right\}$$

We have a minimum if and only if $\Phi'(0) = 0$ and $\Phi''(0) > 0$ for arbitrary functions $\psi_m = \psi_m(\boldsymbol{\alpha})$. We obtain

$$\Phi'(0) = \sum_{m=0}^{M-1} \psi_m \left\{ 2y_m - \eta \frac{\alpha_m y_m}{1 + \alpha_m y_m^2} \right\}$$

and

$$\Phi''(0) = \sum_{m=0}^{M-1} \psi_m^2 \left\{ 2 - \eta \frac{\alpha_m (1 - \alpha_m y_m^2)}{(1 + \alpha_m y_m^2)^2} \right\}$$

The choice

$$y_m^2 = \left[\frac{\eta}{2} - \frac{1}{\alpha_m} \right]_+ \tag{69}$$

yields $\Phi'(0) = 0$ for all ψ_m's. Since $\langle \boldsymbol{\gamma} \rangle$ is decreasing with $\boldsymbol{\gamma}$, the minimum is obtained when the constraint is satisfied with equality. By substituting (69) into $\Phi''(0)$ we get

$$\Phi''(0) = \sum_{m=0}^{M-1} B_m \psi_m^2$$

where

$$B_m = \begin{cases} \frac{8}{\eta}\left[\frac{\eta}{2} - \frac{1}{\alpha_m}\right] & \text{for } y_m^2 > 0 \\ -2\alpha_m\left[\frac{\eta}{2} - \frac{1}{\alpha_m}\right] & \text{for } y_m^2 = 0 \end{cases}$$

By comparing the expressions of y_m^2 and of B_m and by recalling that η must be positive in order to satisfy the constraint equation, we see that $\Phi''(0) > 0$. Then, the solution (69) is the (unique) minimum. Finally, by letting $2/\eta = \lambda(\boldsymbol{\alpha})$, independent of m, we obtain the solution in the form

$$\gamma_m^{\text{lt}}(\boldsymbol{\alpha}) = \left[\lambda(\boldsymbol{\alpha}) - \frac{1}{\alpha_m}\right]_+ \tag{70}$$

where $\lambda(\boldsymbol{\alpha})$ can be obtained by solving the constraint equation $I_M(\boldsymbol{\alpha}, \widehat{\boldsymbol{\gamma}}) = R$. This proves eq. (28) of Lemma 1.

Solution of the short-term constraint equation. Here we consider the equation

$$\frac{1}{M}\sum_{m=0}^{M-1}\frac{1}{2}\log\left(1 + \alpha_m\left[\lambda(\boldsymbol{\alpha}) - \frac{1}{\alpha_m}\right]_+\right) = R \tag{71}$$

Let us assume, without loss of generality, that $\boldsymbol{\alpha} \in \mathcal{Q}$ i.e., $\alpha_m \geq \alpha_{m+1}$ for $m = 0, \ldots, M-2$ and that at least $\alpha_0 > 0$ (otherwise we have outage in any case). Let us define

$$\mu = |\{m : 1/\alpha_m \leq \lambda(\boldsymbol{\alpha})\}|$$

($|\mathcal{S}|$ denotes the cardinality of the set $\mathcal{S}$). Hence, we see that

$$\lambda(\boldsymbol{\alpha}) = \lambda^{\text{lt}}(\mu, \boldsymbol{\alpha}) \triangleq \left(\frac{e^{2MR}}{\prod_{m=0}^{\mu-1}\alpha_m}\right)^{1/\mu}$$

Let $\delta[\mu] = \frac{1}{\mu}(2MR - \sum_{m=0}^{\mu-1}\log(1/\alpha_m))$ and $\nu[\mu] = |\{m : 1/\alpha_m \leq e^{\delta[\mu]}\}|$. Then, (71) is solved by finding μ over $\{1, \ldots, M\}$ such that $\nu[\mu] = \mu$. In the following we prove the existence and the uniqueness of such solution.

Existence. Since $1/\alpha_0 \leq e^{2MR}/\alpha_0 = e^{\delta[1]}$, we get $\nu[\mu] \geq 1$. Also, by definition, $\nu[M] \leq M$. We want to show that the equation

$$\nu[\mu] = \mu \tag{72}$$

has a solution, for any assigned $\boldsymbol{\alpha} \neq \mathbf{0}$. We have

$$\delta[\mu] - \delta[\mu+1] = \frac{1}{\mu+1}(\delta[\mu] - \log(1/\alpha_\mu)) \tag{73}$$

Then, $\nu[\mu] > \mu \Rightarrow \delta[\mu] - \log(1/\alpha_\mu) \geq 0 \Rightarrow \delta[\mu] \geq \delta[\mu+1]$. Hence, for $\nu[\mu] > \mu$, the sequence $\nu[\mu]$ is non-increasing. Moreover, from (73) we get

$$\delta[\mu+1] - \log(1/\alpha_\mu) = \frac{\mu}{\mu+1}(\delta[\mu] - \log(1/\alpha_\mu))$$

so that, $\nu[\mu] > \mu \Rightarrow \delta[\mu+1] - \log(1/\alpha_\mu) \geq 0 \Rightarrow \nu[\mu+1] \geq \mu+1$. Then, either $\nu[\mu] > \mu$ for all $\mu = 1, \ldots, M-1$, which implies that $\nu[M] = M$, or $\nu[\mu] = \mu$ for some $\mu < M$. In both cases the solution exists.

Uniqueness. Assume there exist two values μ_1 and μ_2, with $\mu_1 < \mu_2$, such that $\nu[\mu_1] = \mu_1$ and $\nu[\mu_2] = \mu_2$. Then,

$$\log(1/\alpha_{\mu_2-1}) > \delta[\mu_1] \quad \text{and} \quad \log(1/\alpha_{\mu_2-1}) \leq \delta[\mu_2]$$

By inverting the sign of the former inequality and adding the result to the latter, we obtain

$$(\mu_2 - \mu_1)\log(1/\alpha_{\mu_2-1}) < \sum_{m=\mu_1}^{\mu_2-1} \log(1/\alpha_m)$$

which contradicts the non-increasing assumption on $\boldsymbol{\alpha}$.

C.2 Proof of Lemma 2

Because of symmetry, without loss of generality we can consider only $\boldsymbol{\alpha} \in \mathcal{Q}$. It is immediate to see that $\boldsymbol{\gamma}^{\mathrm{lt}}(\boldsymbol{\alpha})$ is continuous for all $\boldsymbol{\alpha} \neq \mathbf{0}$ in the interior of each $\mathcal{W}_i$ (the regions $\mathcal{W}_i$ are defined by (40)). Then, we need only to check continuity on the boundaries. Consider the surface $\mathcal{S}_i$, boundary of $\mathcal{W}_i$ and $\mathcal{W}_{i-1}$, defined by

$$\mathcal{S}_i = \left\{ \boldsymbol{\alpha} \in \mathcal{W}_i : \lambda^{\mathrm{lt}}(i, \boldsymbol{\alpha}) = \frac{1}{\alpha_{i-1}} \right\}$$

From the equality defining $\mathcal{S}_i$ we obtain

$$\lambda^{\mathrm{lt}}(i, \boldsymbol{\alpha}) = \lambda^{\mathrm{lt}}(i-1, \boldsymbol{\alpha}) \tag{74}$$

Then, let $\mathbf{a}_0 \in \mathcal{S}_i$ and $\mathbf{a}$ be an arbitrary point of $\mathcal{W}_{i-1}$. For $i-1 \leq m \leq M-1$, $\gamma_m^{\mathrm{lt}}(\mathbf{a}) = \gamma_m^{\mathrm{lt}}(\mathbf{a}_0) = 0$ and for $0 \leq m \leq i-2$, $\lim_{\mathbf{a} \to \mathbf{a}_0} \gamma_m^{\mathrm{lt}}(\mathbf{a}) = \gamma_m^{\mathrm{lt}}(\mathbf{a}_0)$ because of (74). Then, $\boldsymbol{\gamma}^{\mathrm{lt}}(\boldsymbol{\alpha})$ is continuous also in $\boldsymbol{\alpha} \in \mathcal{S}_i, \boldsymbol{\alpha} \neq \mathbf{0}$.

Since $\boldsymbol{\gamma}^{\mathrm{lt}}(\boldsymbol{\alpha})$ is continuous, we can show that $\langle \boldsymbol{\gamma}^{\mathrm{lt}}(\boldsymbol{\alpha}) \rangle$ is non-increasing in α_m by showing that $\frac{\partial}{\partial \alpha_m} \langle \boldsymbol{\gamma}^{\mathrm{lt}}(\boldsymbol{\alpha}) \rangle \leq 0$ in the interior of each $\mathcal{W}_i$, for $i = 1, \ldots, M$. By using (28) and (29) and by differentiating with respect to α_m we get

$$\frac{\partial}{\partial \alpha_m}\left\langle \boldsymbol{\gamma}^{\mathrm{lt}}(\boldsymbol{\alpha})\right\rangle = \begin{cases} \dfrac{1}{M\alpha_m}\left[\dfrac{1}{\alpha_m} - \lambda^{\mathrm{lt}}(i,\boldsymbol{\alpha})\right] & m = 0,\dots,i-1 \\ 0 & m = i,\dots,M-1 \end{cases}$$

Since $\boldsymbol{\alpha} \in \mathcal{W}_i$, the RHS of the above equation is non-positive for all $m = 0,\dots,M-1$.

C.3 Proof of Lemma 3

For $s \in \mathbb{R}_+$, let $\mathcal{P}(s)$ and $\overline{\mathcal{P}}(s)$ be defined by (32), $\widehat{w}(u)$ given by (31). The function $\widehat{w}(u)$ defined above satisfies the constraint with equality,

$$\begin{aligned} E[u\,\widehat{w}(u)] &= \mathcal{P}(s^*) + w^* s^* P(u = s^*) \\ &= \mathcal{P}(s^*) + \frac{\mathcal{P} - \mathcal{P}(s^*)}{\overline{\mathcal{P}}(s^*) - \mathcal{P}(s^*)} s^* P(u = s^*) \\ &= \mathcal{P} \end{aligned}$$

and achieves the objective function value

$$E[\widehat{w}(u)] = \int_{[0,s^*)} dF(u) + w^* P(u = s^*) \tag{75}$$

We are to show that any other $0 \le w(u) \le 1$ such that $E[w(u)]$ is larger than (75) must violate the constraint. For any such $w(u)$, we have

$$\begin{aligned} E[u\,w(u)] - \mathcal{P} &= E[u\,w(u)] - E[u\,\widehat{w}(u)] \\ &= \int_{(s^*,\infty)} u\,w(u)dF(u) + s^*(w(s^*) - w^*)P(u = s^*) - \int_{[0,s^*)} u\,(1 - w(u))dF(u) \\ &\ge s^*\left\{\int_{(s^*,\infty)} w(u)dF(u) + (w(s^*) - w^*)P(u = s^*) - \int_{[0,s^*)} (1 - w(u))dF(u)\right\} \\ &\ge s^*\left\{E[w(u)] - E[\widehat{w}(u)]\right\} \end{aligned} \tag{76}$$

Then, if $E[w(u)] > E[\widehat{w}(u)]$ we get $E[u\,w(u)] > \mathcal{P}$, i.e., $w(u)$ violates the constraint.

D Proof of Proposition 4

The proof is organized in two steps. First, we show that the solution of (26) must be in the form

$$\widehat{\boldsymbol{\gamma}}(\boldsymbol{\alpha}) = \begin{cases} \boldsymbol{\gamma}^{\mathrm{lt}}(\boldsymbol{\alpha}) & \text{with probability } \widehat{w}(\boldsymbol{\alpha}) \\ \mathbf{0} & \text{with probability } 1 - \widehat{w}(\boldsymbol{\alpha}) \end{cases} \tag{77}$$

where $\boldsymbol{\gamma}^{\text{lt}}(\boldsymbol{\alpha})$ is the solution of (27) and where $\widehat{w}(\boldsymbol{\alpha})$ is a weight function $\mathbb{R}_+^M \to [0,1]$, solution of the problem

$$\begin{cases} \text{Maximize} & \mathrm{E}[w(\boldsymbol{\alpha})] \\ \text{Subject to} & 0 \leq w(\boldsymbol{\alpha}) \leq 1 \;\text{ and }\; \mathrm{E}[\langle \widehat{\boldsymbol{\gamma}}(\boldsymbol{\alpha})\rangle\, w(\boldsymbol{\alpha})] = \mathcal{P} \end{cases} \tag{78}$$

Then, we use Lemma 1, 2, and 3 along with the definitions (34 - 37) to complete the proof.

Step 1. The outage probability resulting from the power allocation function given by (77) is

$$\widehat{P}_{\text{out}}(R,\mathcal{P}) = 1 - \mathrm{E}[\widehat{w}(\boldsymbol{\alpha})] \tag{79}$$

Let $\boldsymbol{\gamma}$ be an arbitrary power allocation in the class of probabilistic stationary memoryless power allocation functions satisfying $E[\langle\boldsymbol{\gamma}\rangle] \leq \mathcal{P}$, and consider the region $\mathcal{A}(\boldsymbol{\alpha}, R) = \{\boldsymbol{\gamma} \in \mathbb{R}_+^M : I_M(\boldsymbol{\alpha},\boldsymbol{\gamma}) \geq R\}$. The outage probability resulting from $\boldsymbol{\gamma}$ is

$$P_{\text{out}}(R,\mathcal{P}) = 1 - P(\boldsymbol{\gamma} \in \mathcal{A}(\boldsymbol{\alpha}, R)) \tag{80}$$

We are to show that $\widehat{P}_{\text{out}}(R,\mathcal{P}) \leq P_{\text{out}}(R,\mathcal{P})$, where $\widehat{P}_{\text{out}}(R,\mathcal{P})$ is given by (79). To this purpose, let $\chi_{\mathcal{A}}$ denote the indicator function of $\{\boldsymbol{\gamma} \in \mathcal{A}(\boldsymbol{\alpha}, R)\}$ and define the weight function

$$w(\boldsymbol{\alpha}) = \mathrm{E}[\chi_{\mathcal{A}}|\boldsymbol{\alpha}] \tag{81}$$

where expectation is with respect to $F(\boldsymbol{\gamma}|\boldsymbol{\alpha})$. Now, $0 \leq w(\boldsymbol{\alpha}) \leq 1$, so that $w(\boldsymbol{\alpha})$ is a valid weight function. Then, define the new power allocation

$$\boldsymbol{\gamma}'(\boldsymbol{\alpha}) = \begin{cases} \boldsymbol{\gamma}^{\text{lt}}(\boldsymbol{\alpha}) & \text{with probability } \; w(\boldsymbol{\alpha}) \\ \mathbf{0} & \text{with probability } \; 1 - w(\boldsymbol{\alpha}) \end{cases}$$

where $\boldsymbol{\gamma}^{\text{lt}}$ is the solution of (27). By definition of $\boldsymbol{\gamma}^{\text{lt}}$, the outage probability $P'_{\text{out}}(R,\mathcal{P})$ resulting from $\boldsymbol{\gamma}'$ is equal to $P_{\text{out}}(R,\mathcal{P})$. In fact,

$$\begin{aligned} P'_{\text{out}}(R,\mathcal{P}) &= 1 - \mathrm{E}[w(\boldsymbol{\alpha})] \\ &= 1 - \mathrm{E}[\mathrm{E}[\chi_{\mathcal{A}}|\boldsymbol{\alpha}]] \\ &= 1 - \mathrm{E}[\chi_{\mathcal{A}}] \\ &= 1 - P(\boldsymbol{\gamma} \in \mathcal{A}(\boldsymbol{\alpha}, R)) \end{aligned} \tag{82}$$

Moreover, $\boldsymbol{\gamma}'$ satisfies the long-term power constraint. In fact,

$$\begin{aligned}
\mathcal{P} &\geq \mathrm{E}[\langle\boldsymbol{\gamma}\rangle] \\
&\geq \mathrm{E}[\chi_{\mathcal{A}}\,\langle\boldsymbol{\gamma}\rangle] \\
&\overset{\mathrm{a}}{\geq} \mathrm{E}[\chi_{\mathcal{A}}\left\langle\boldsymbol{\gamma}^{\mathrm{lt}}\right\rangle] \\
&= \mathrm{E}[\mathrm{E}[\chi_{\mathcal{A}}\left\langle\boldsymbol{\gamma}^{\mathrm{lt}}\right\rangle|\boldsymbol{\alpha}]] \\
&\overset{\mathrm{b}}{=} \mathrm{E}[\left\langle\boldsymbol{\gamma}^{\mathrm{lt}}\right\rangle \mathrm{E}[\chi_{\mathcal{A}}|\boldsymbol{\alpha}]] \\
&= \mathrm{E}[\left\langle\boldsymbol{\gamma}^{\mathrm{lt}}\right\rangle w(\boldsymbol{\alpha})] \\
&= \mathrm{E}[\langle\boldsymbol{\gamma}'\rangle]
\end{aligned}$$

where (a) follows by noting that for all $\boldsymbol{\alpha}$ such that $\chi_{\mathcal{A}} = 1$, then $\left\langle\boldsymbol{\gamma}^{\mathrm{lt}}\right\rangle \leq \langle\boldsymbol{\gamma}\rangle$ by definition of $\boldsymbol{\gamma}^{\mathrm{lt}}$, and (b) follows from the fact that $\boldsymbol{\gamma}^{\mathrm{lt}}$ is a deterministic function of $\boldsymbol{\alpha}$ (see Lemma 1). Since $w(\boldsymbol{\alpha})$ is just one of the possible weight functions satisfying $\mathrm{E}[\left\langle\boldsymbol{\gamma}^{\mathrm{lt}}\right\rangle w(\boldsymbol{\alpha})] \leq \mathcal{P}$, we have $\mathrm{E}[\widehat{w}(\boldsymbol{\alpha})] \geq \mathrm{E}[w(\boldsymbol{\alpha})]$ where $\widehat{w}(\boldsymbol{\alpha})$ is the solution of (78) (note that the solution of (78) must satisfy the constraint with equality). From the last inequality and from (82) we get $\widehat{P}_{\mathrm{out}}(R,\mathcal{P}) \leq P_{\mathrm{out}}(R,\mathcal{P})$.

Step 2. The explicit expression of $\boldsymbol{\gamma}^{\mathrm{lt}}(\boldsymbol{\alpha})$ is given by Lemma 1. We use Lemma 2 and Lemma 3 in order to find the explicit expression of the optimal weight function $\widehat{w}(\boldsymbol{\alpha})$. The power sum $u = \left\langle\boldsymbol{\gamma}^{\mathrm{lt}}(\boldsymbol{\alpha})\right\rangle$ is a function $\mathbb{R}_+^M \to \mathbb{R}_+$. For a given $F(\boldsymbol{\alpha})$, the power sum is a non-negative random variable with cdf

$$F(u) = P(\left\langle\boldsymbol{\gamma}^{\mathrm{lt}}(\boldsymbol{\alpha})\right\rangle \leq u) = \int_{\overline{\mathcal{R}}(u)} dF(\boldsymbol{\alpha})$$

We can apply Lemma 3 to u and find s^* and w^* which yield the maximum of $E[w(u)]$ subject to $E[u\ w(u)] \leq \mathcal{P}$. Finally, from Lemma 2 we have that $\boldsymbol{\gamma}^{\mathrm{lt}}(\boldsymbol{\alpha})$ is continuous and that $\left\langle\boldsymbol{\gamma}^{\mathrm{lt}}(\boldsymbol{\alpha})\right\rangle$ has a gradient all of whose components are non-positive. This implies that for any $0 \leq s < s' < \infty$, the regions $\mathcal{R}(s)$ and $\mathcal{R}(s')$ defined by (34) satisfy $\mathcal{R}(s) \subseteq \mathcal{R}(s')$. Thus, the intervals $u < s^*$, $u = s^*$, and $u > s^*$ on the non-negative u-axis correspond to the regions $\mathcal{R}(s^*)$, to the boundary surface $\mathcal{B}(s^*)$, and to the complement region $\mathbb{R}_+^M - \overline{\mathcal{R}}(s^*)$, respectively.

References

[1] L. Ozarow, S. Shamai, and A. D. Wyner, "Information theoretic considerations for cellular mobile radio," *IEEE Trans. Vehic. Tech.,* Vol. 43, No. 2, pp. 359–378, May 1994.

[2] G. Kaplan and S. Shamai (Shitz), "Error probabilities for the block-fading Gaussian channel," *A.E.Ü.,* Vol. 49, No. 4, pp. 192–205, 1995.

[3] R. McEliece and W. Stark, "Channels with block interference," *IEEE Trans. on Inform. Theory,* Vol. 30, No. 1, pp. 44–53, Jan. 1984.

[4] R. Knopp, P. A. Humblet, "Maximizing diversity on block fading channels," *Proceedings of ICC '97,* Montréal, Canada, 8–12 June, 1997.

[5] G. Caire, R. Knopp and P. Humblet, "System capacity of F-TDMA cellular systems," to appear on *IEEE Trans. on Comm.,* 1998.

[6] T. Rappaport, *Wireless Communications.* Englewood Cliffs, NJ: Prentice-Hall, 1996.

[7] Proposed EIA/TIA Interim Standard, *Wideband Spread Spectrum Digital Cellular System Dual-Mode Mobile Station-Base Station Compatibility Standard*, 1992.

[8] ARIB, *Personal Handy Phone System ARIB Standard, Version 2*, 1995.

[9] A. Goldsmith and P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Trans. on Inform. Theory,* Vol. 43, No. 6, pp. 1968 – 1992, Nov. 1997.

[10] A. Goldsmith, *Design and Performance of High-Speed Communication Systems over Time-Varying Radio Channels*, Ph.D. Thesis, University of California at Berkeley, Berkeley, CA, USA, 1994.

[11] G. Foschini Jr. and M. Gans, "On limits of wireless communication in a fading environment when using multiple antennas," *Wireless Personal Communications,* submitted for publication, 1997.

[12] R. Knopp, *Coding and Multiple Access over fading channels*, Ph.D. Thesis, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland, 1997.

[13] E. Malkamäki and H. Leib, "Coded diversity schemes on block fading Rayleigh channels," *Proc. of IEEE Int. Conf. on Universal Personal Communications (ICUPC'97),* San Diego, CA, Oct. 1997.

[14] S. Verdú and T. S. Han, "A general formula for channel capacity," *IEEE Trans. Inform. Theory,* Vol. 40, No. 4, pp. 1147–1157, July 1994.

[15] R. Gallager, *Information Theory and Reliable Communication.* New York: J. Wiley & Sons, 1968.

[16] D. Tse and V. Hanly, "Multi-access Fading Channels. Part II: Delay-limited capacities," *IEEE Trans. Inform. Theory,* submitted for publication, 1996.

[17] G. Caire, E. Biglieri and G. Taricco, "System capacity of a single-cell CDMA mobile communication system," *Proc. of the IEEE Communication Theory Mini-Conference, Globecom'97,* Phoenix, AZ, Nov. 1997.

[18] R. Knopp, P. A. Humblet, "Information capacity and power-control in single-cell multiuser communications," *Proc. of the IEEE International Conference on Communications, ICC '95*, pp. 331–335, Seattle, WA, June 18–22, 1995.

[19] R. Gallager and M. Médard, "Bandwidth scaling for fading channels," *Proc. of ISIT '97*, p. 471, Ulm, Germany, June 29–July 4, 1997.

[20] R. M. Gray, *Entropy and Information Theory.* New York: Springer-Verlag, 1990.

[21] G. Caire and S. Shamai, "On the capacity of some channels with channel state information," Technion Technical Report EE PUB #1122, Nov. 1997. Also submitted to *IEEE Trans. Inform. Theory,* Nov. 1997.

[22] M. Schwartz, W. Bennett, and S. Stein, *Communication Systems and Techniques.* New York: McGraw-Hill, 1966.

[23] M.S. Alouini and A. Goldsmith, "Capacity of Rayleigh fading channels under different adaptive transmission and diversity-combining techniques," *IEEE Trans. Vehic. Tech.* submitted for publication, Nov. 1996.

[24] E. Lutz, D. Cygan, M. Dippold, F. Dolainsky, and W. Papke "The land mobile satellite communication channel–Recording, statistics, and channel model," *IEEE Trans. Vehic. Tech.,* Vol. 40, No. 2, pp. 1209–1218, May 1991.

[25] I. Gradshteyn and I. Ryzhik, *Tables of Integrals, Series, and Products.* New York: Academic Press, 1980.

[26] E. Biglieri, G. Caire, G. Taricco, "Random coding bounds for the block-fading channel with simple power allocation schemes" *Conf. on Infomation Sciences and Systems (CISS '98),* Princeton University, Princeton, NJ, March 18 – 20, 1998.

[27] R. Knopp and G. Caire, "Simple power-controlled modulation schemes for fixed rate transmission over fading channels," *IEEE Comm. Letters*, submitted for publication, Sept. 1997.

[28] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. on Inform. Theory* Vol. 44, No. 2, pp. 744 – 765, March 1998.

[29] A. Narula, M. D. Trott, and G. Wornell, "Information theoretic analysis of multiple antenna transmitter diversity," *IEEE Trans. on Inform. Theory,* submitted for publication, Nov. 1996.

[30] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Unpublished manuscript*, 1997.

[31] G. Caire, G. Taricco and E. Biglieri, "Limiting performance of block-fading channels with multiple antennas," in preparation, June 1998.

[32] G. Caire, G. Taricco and E. Biglieri, "Minimum information outage rate for slowly-varying fading channels," *International Symposium on Information Theory (ISIT '98),* Boston, MA, Aug. 16 – 21, 1998.

M	Const.	Opt. short-term	Opt. long-term	Rep. const.	Rep. short-term	Rep. long-term
1	1.0e-1	1.0e-1	1.0e-1	1.0e-1	1.0e-1	1.0e-1
2	1.0e-2	1.0e-2	1.0e-2	1.0e-2	1.0e-2	1.0e-2
3	1.0e-3	1.0e-3	1.0e-3	1.0e-3	1.0e-3	1.0e-3
4	3.7e-3	1.0e-4	1.0e-4	3.7e-3	1.0e-4	1.0e-4
5	4.6e-4	1.0e-5	1.0e-5	8.1e-2	1.0e-5	1.0e-5
6	5.5e-5	5.5e-5	1.0e-6	1.0	1.0	4.7e-2
7	1.7e-4	6.4e-6	1.0e-7	1.0	1.0	4.4e-1
8	2.3e-5	7.3e-7	1.0e-8	1.0	1.0	6.8e-1
9	2.9e-6	8.2e-8	1.0e-09	1.0	1.0	8.2e-1
10	3.7e-7	9.1e-9	1.0e-10	1.0	1.0	9.0e-1
11	1.2e-6	1.0e-9	1.0e-11	1.0	1.0	9.4e-1
12	1.6e-7	5.4e-9	1.0e-12	1.0	1.0	9.7e-1

Table 1: Outage probability of the on-off BF-AWGN channel with M blocks, $p = 0.1$, $R = 0.5$ bit/symbol and $\mathcal{P} = 10$ dB.

M	Const.	Opt. short-term	Opt. long-term	Rep. const.	Rep. short-term	Rep. long-term
1	9.5e-2	9.5e-2	2.5e-4	9.5e-2	9.5e-2	2.5e-4
2	2.1e-2	1.5e-2	0.0	3.6e-2	1.9e-2	0.0
3	6.0e-3	2.8e-3 *	0.0	3.4e-2	9.0e-3	0.0
4	1.6e-3	6.6e-4 *	0.0	6.5e-2	9.5e-3	0.0
5	4.9e-4	1.3e-4 *	0.0	2.0e-1	2.1e-2	0.0
6	1.4e-4	3.0e-5 *	0.0	6.0e-1	7.5e-2	0.0
7	4.4e-5	6.4e-6 *	0.0	9.6e-1	2.8e-1	0.0
8	1.3e-5	1.8e-6 *	0.0	1.0	7.1e-1	1.6e-1
9	4.1e-6	...	0.0	1.0	9.6e-1	4.1e-1
10	1.2e-6	...	0.0	1.0	1.0	6.1e-1
11	3.9e-7	...	0.0	1.0	1.0	7.5e-1
12	1.2e-7	...	0.0	1.0	1.0	8.4e-1

Table 2: Outage probability of the Rayleigh BF-AWGN channel with M blocks, $R = 0.5$ bit/symbol and $\mathcal{P} = 10$ dB. The values marked by * are obtained by Monte Carlo simulation (a minimum of 1000 outage events were counted for each value).

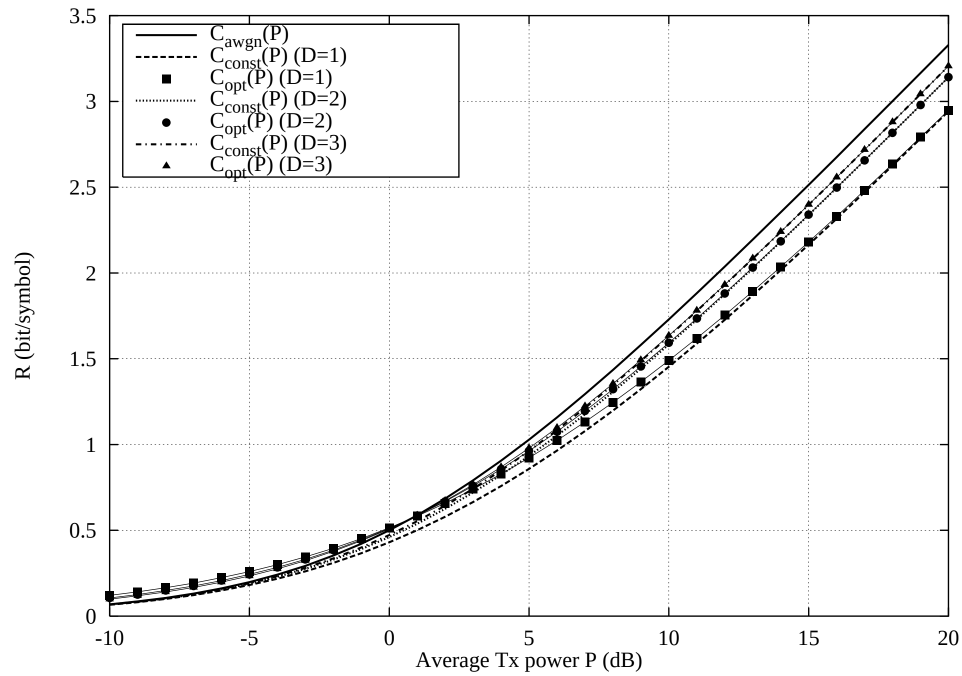

Figure 1: AWGN and Rayleigh fading channel capacity with optimal and constant power allocation and space diversity order $D = 1, 2, 3$.

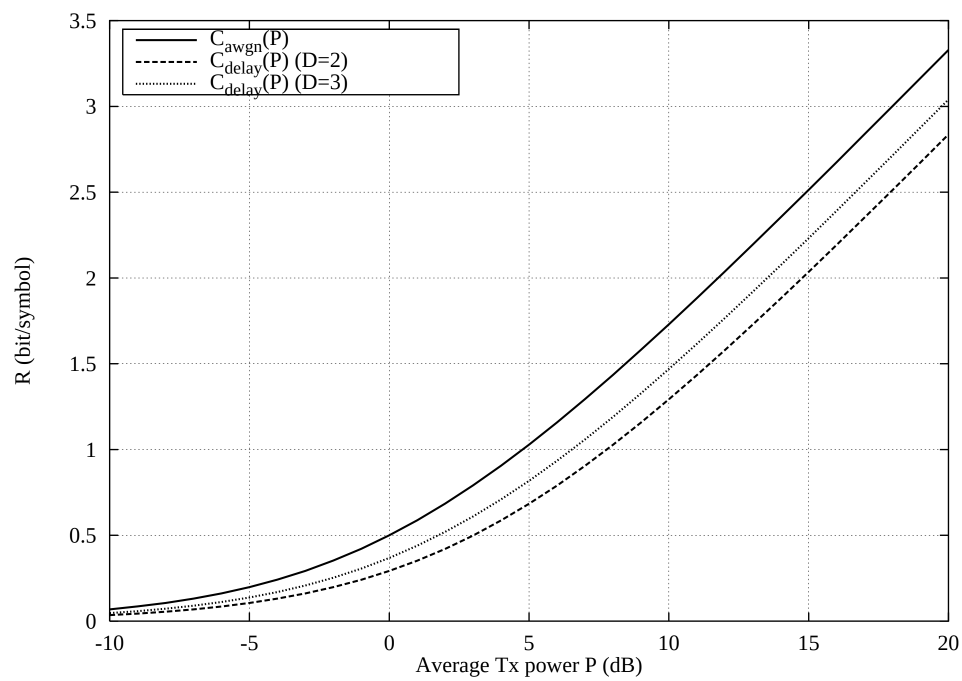

Figure 2: Rayleigh fading channel delay-limited capacity with space diversity order $D = 2, 3$. The AWGN capacity is shown for comparison.

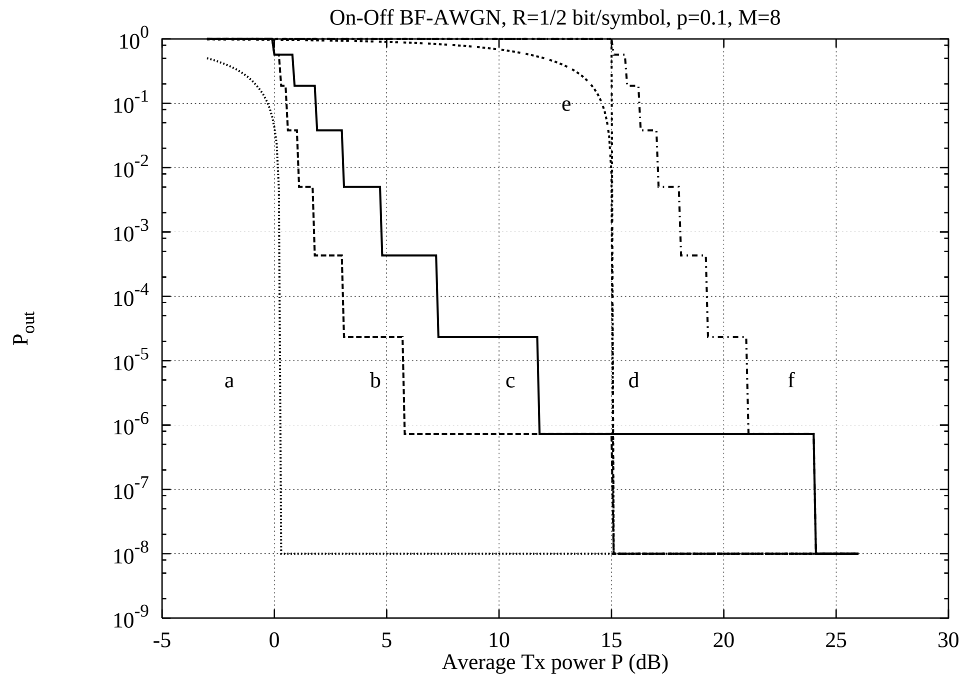

Figure 3: Outage probability vs. $\mathcal{P}$ for the on-off BF-AWGN channel with $p = 0.1$, $M = 8$ and $R = 0.5$ bit/symbol. a) Optimal long-term power. b) Optimal short-term power. c) Constant power. d) Repetition diversity, short-term power. e) Repetition diversity, long-term power. f) Repetition diversity, constant power.

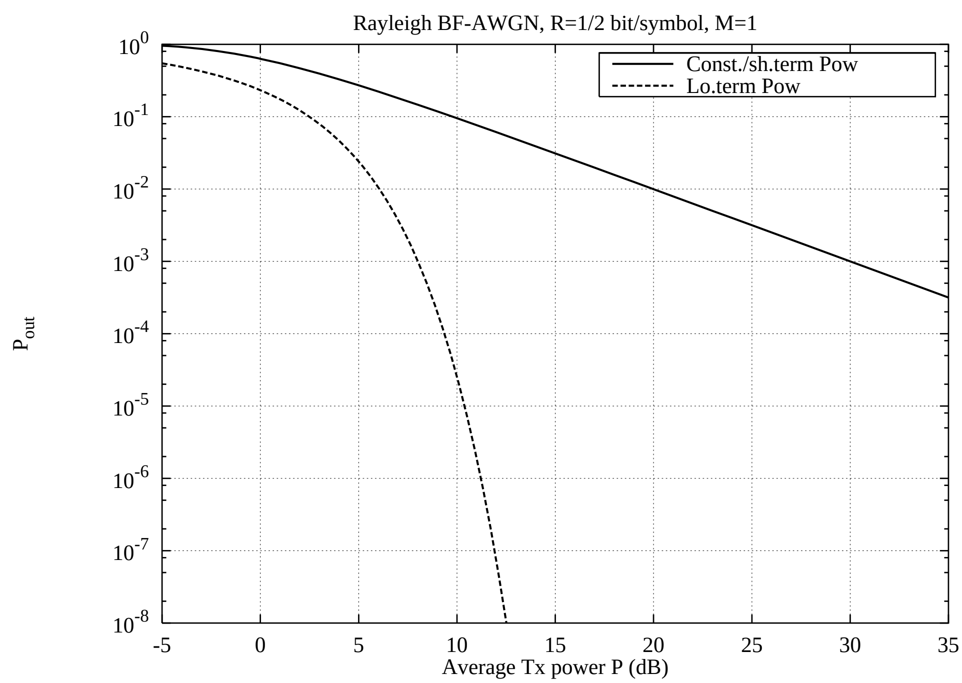

Figure 4: Outage probability vs. $\mathcal{P}$ for the Rayleigh fading channel with $M = 1$ and $R = 0.5$ bit/symbol.

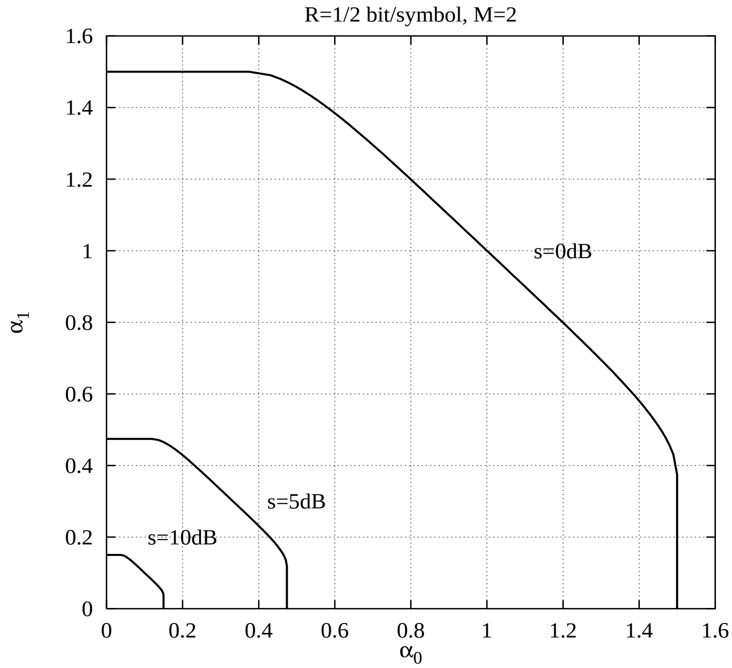

Figure 5: Curves $\mathcal{B}(s)$ for $M = 2$, $s = 0, 5$ and 10 dB and $R = 0.5$ bit/symbol.

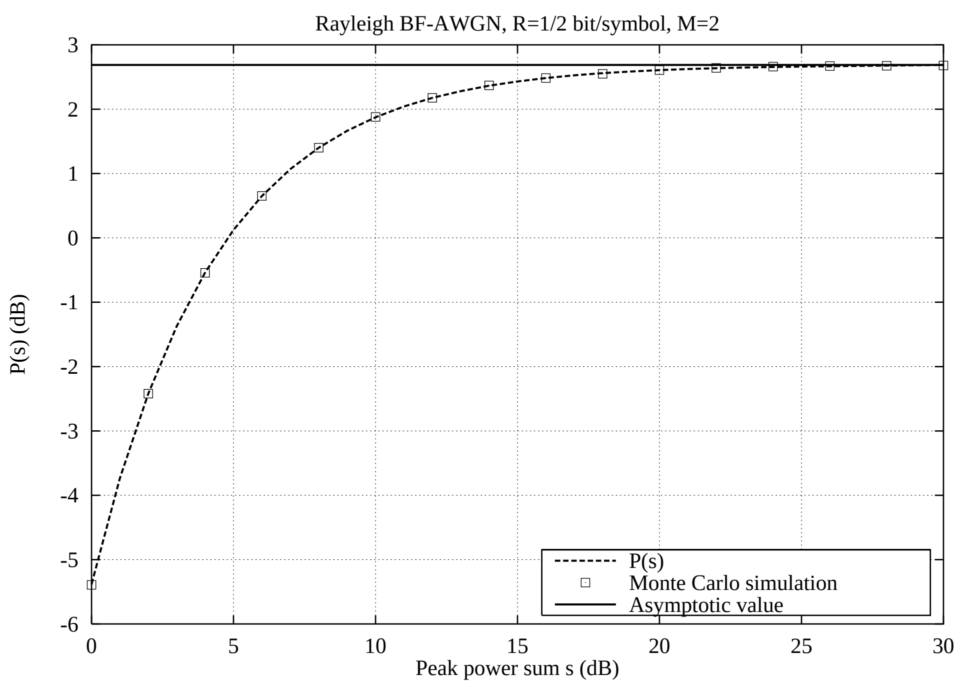

Figure 6: $\mathcal{P}(s)$ vs. s for the Rayleigh fading channel with $M = 2$ and $R = 0.5$ bit/symbol. Monte Carlo simulation points are shown with the analytical curve.

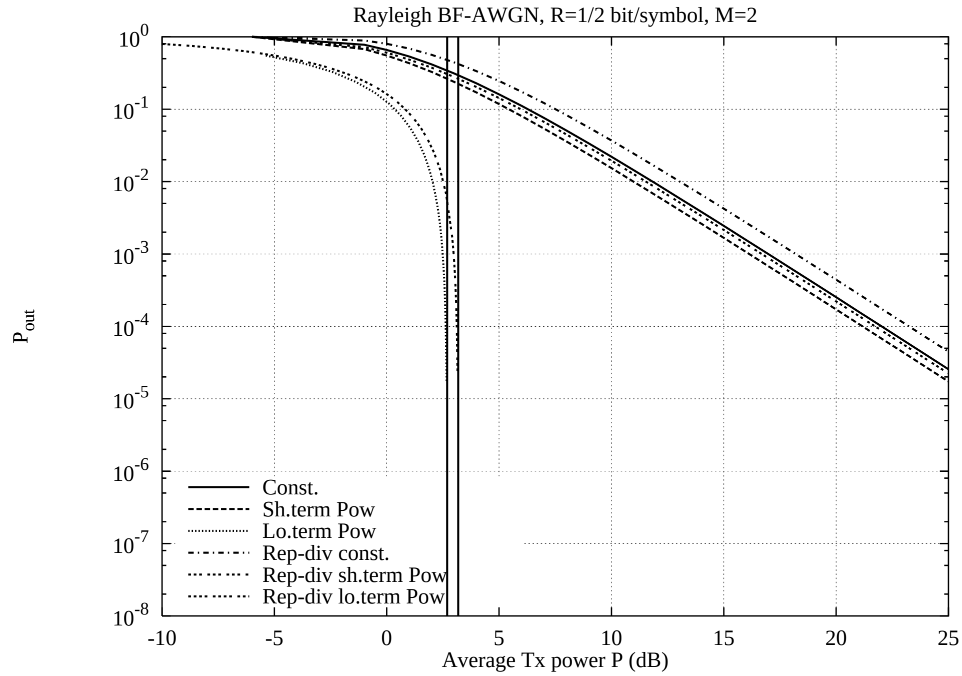

Figure 7: Outage probability vs. $\mathcal{P}$ for the Rayleigh BF-AWGN channel with $M = 2$, $R = 0.5$ bit/symbol.

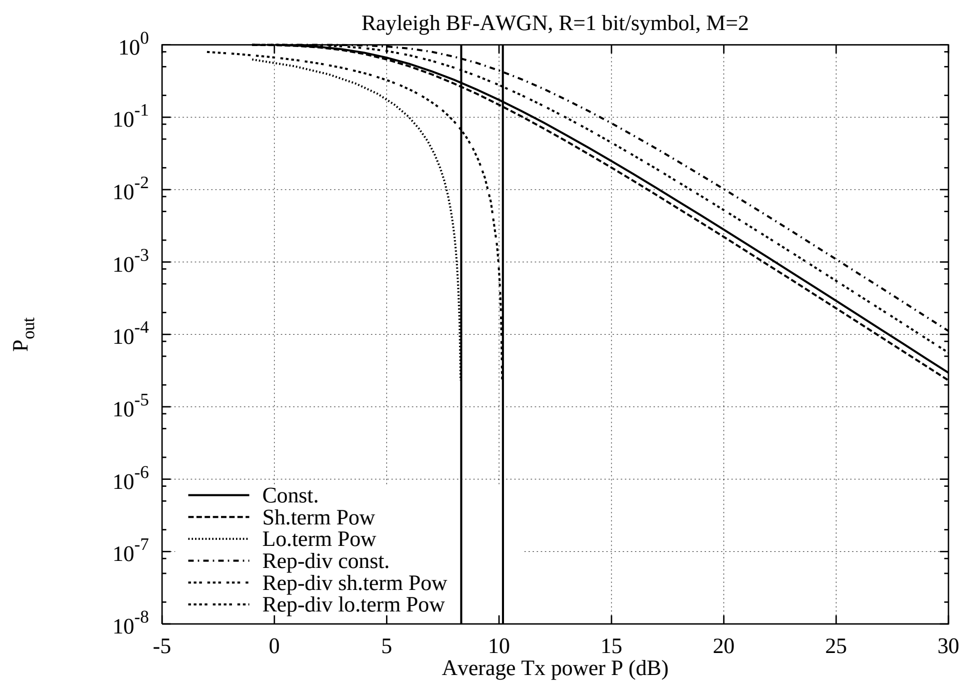

Figure 8: Outage probability vs. $\mathcal{P}$ for the Rayleigh BF-AWGN channel with $M = 2$, $R = 1.0$ bit/symbol.

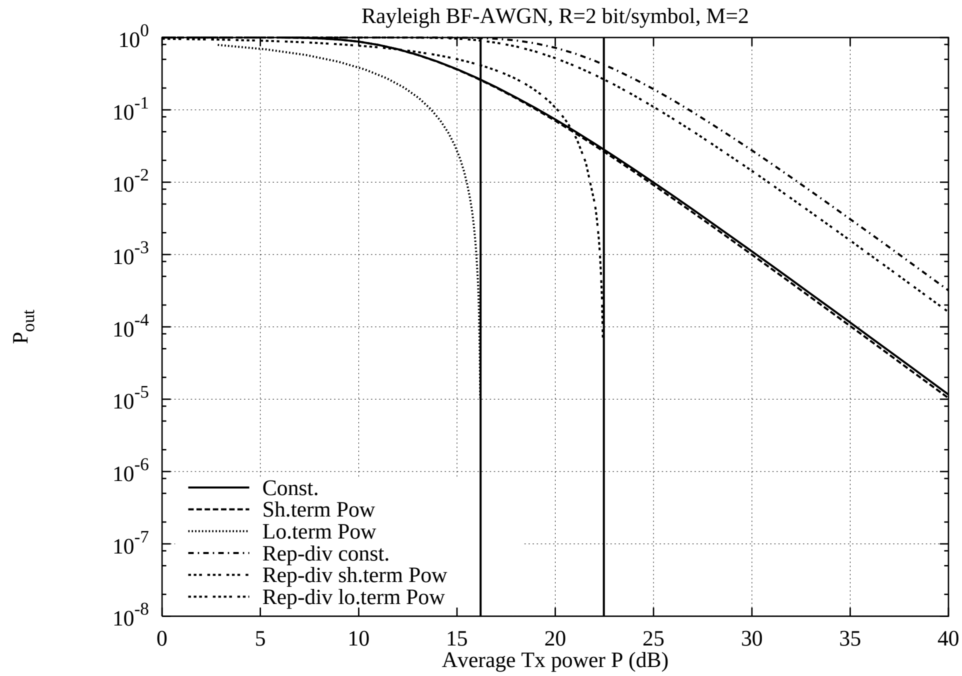

Figure 9: Outage probability vs. $\mathcal{P}$ for the Rayleigh BF-AWGN channel with $M = 2$, $R = 2.0$ bit/symbol.

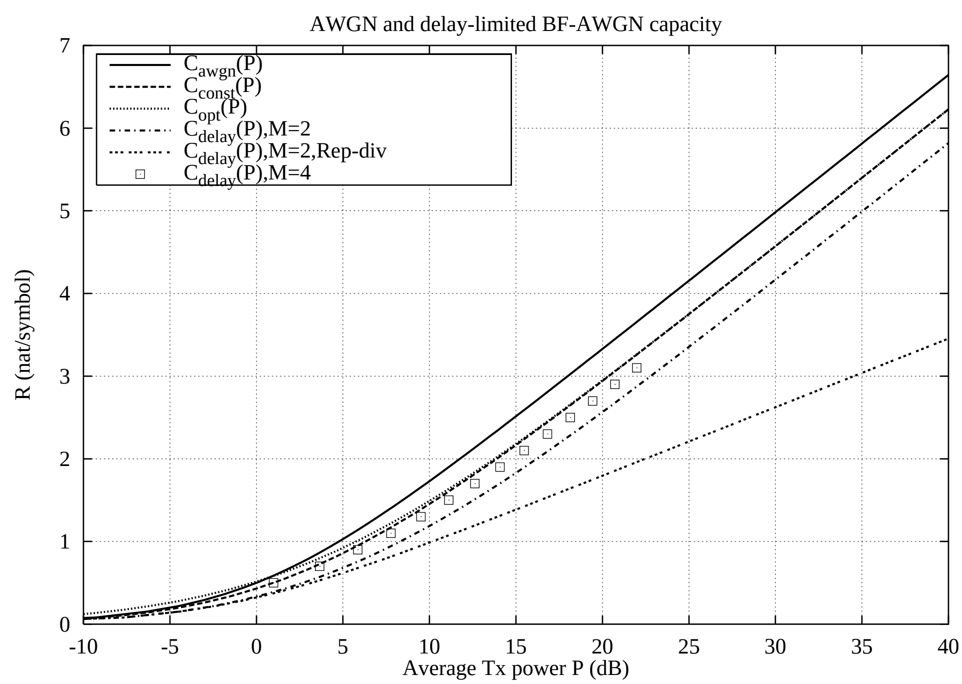

Figure 10: Delay-limited capacity of the Rayleigh BF-AWGN channel with $M = 2$ with optimal coding and repetition diversity. The AWGN capacity and the capacity of a Rayleigh fading channel without delay constraint are shown for comparison. Monte Carlo simulation for optimal coding and $M = 4$ is also included.

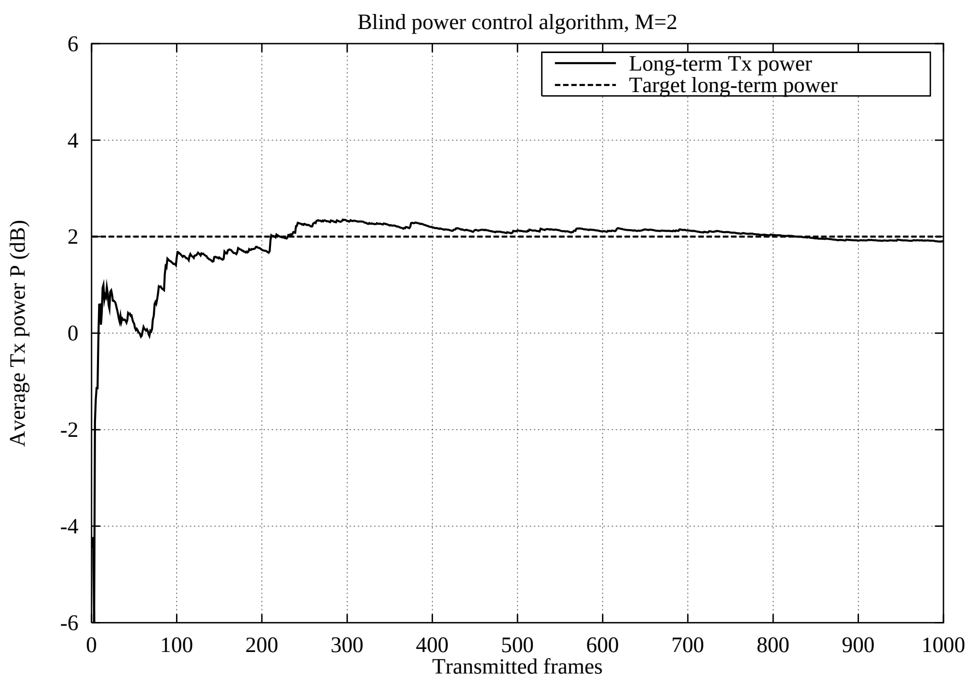

Figure 11: Long-term transmitted power resulting from the adaptive power control algorithm for $\mathcal{P} = 2$ dB, $M = 2$, $R = 0.5$ bit/symbol and $\epsilon = 0.01$. The horizontal line indicates the target long-term power $\mathcal{P} = 2$ dB.

Footnotes

1. From a practical viewpoint, CSI at the transmitter can be provided either by a dedicated feedback channel (some existing systems already implement a fast power control feedback channel [6, 7]) or by time-division duplex [8], where the uplink and the downlink time-share the same M subchannels and the fading gains can be estimated from the incoming signal.

2. Note the difference with respect to space diversity: this may be thought of as a technique to modify the fading gain statistics after combining, and hence does not decrease the code rate.

3. This assumption is sufficient to ensure the information stability of the channel [14].

4. All the numerical results in this paper can be translated immediately into results for the more standard circularly-symmetric complex channel [1] with average energy per symbol E_s, noise power spectral density N_0 and signaling rate W symbols/s by letting $\mathcal{P} = E_s/N_0$ and by multiplying the information rates by $2W$. In this way, the information rates are expressed in information units per second.

5. See Appendix A for the definition of ϵ-achievable rate used in this paper.

6. Since $\boldsymbol{\gamma} = \boldsymbol{\gamma}(\boldsymbol{\alpha})$ is a random function, defined by its conditional cdf $F(\boldsymbol{\gamma}|\boldsymbol{\alpha})$, all optimization problems in the following are to be intended with respect to $F(\boldsymbol{\gamma}|\boldsymbol{\alpha})$. However, it is convenient to give the solution directly in terms of $\boldsymbol{\gamma}$, rather than in terms of its conditional cdf.

7. Even though typical fading models have continuous cdf, there are practical cases of interest where the fading distribution appears as discrete. This occurs for example when the transmitter can use only a coarsely quantized information about the fading levels, or in mobile satellite systems when the fading is modeled as an on-off process depending on the presence or absence of a line-of-sight propagation path [24].

8. With this observation we do not intend to question the merits of classical CDMA. For example, the selection diversity technique proposed here would be easy to intercept by eavesdroppers. Moreover, CDMA is used in multiuser systems, where multiple-access interference rather than fading is the main impairment.

9. This model does not represent only an extreme simplification of fading. Consider for example a multicarrier multimedia network where fixed-rate services (e.g., voice and video) share several

subcarriers with variable-rate bursty services (e.g., file transfer). A common centralized multiple-access protocol allocates the subcarriers to the different services. A fixed-rate service transmits information at rate R and makes use of M subcarriers, but some of these subcarriers may be assigned dynamically to some other variable-rate service with higher priority, with a certain probability. We are interested in the minimum outage probability of fixed-rate users in such system.