Transactions Letters

Outage Capacity and Optimal Power Allocation for Multiple Time-Scale Parallel Fading Channels

Subhrakanti Dey, *Senior Member, IEEE,* and Jamie Evans, *Member, IEEE*

***Abstract*— In this paper, we address the optimal power allocation problem for minimizing capacity outage probability in multiple time-scale parallel fading channels. Extending ideas from [1], we derive the optimal power allocation scheme for parallel fading channels with fast Rayleigh fading, as a function of the slow fading gains. Numerical results are presented to demonstrate the outage performance of this scheme for lognormal slow fading on two parallel channels.**

***Index Terms*— Fading channels, multiple time-scale, optimization, outage capacity, power control.**

I. Introduction

DETERMINING the information theoretic capacity of fading wireless channels has been an important area of research over the past decade. Various notions of capacity for single-user fading channels include *ergodic capacity* [2], *delay-limited capacity* [3] and *capacity versus outage probability* [4], [5]. If the transmitter has the channel state information (CSI), then transmit power could be controlled as a function of the channel to maximize the capacity. The paper [2] (see also [6]) looked at the problem of maximizing ergodic capacity subject to an average power constraint, and showed that the optimal power control law was *waterfilling* on the inverse of the channel gain (more power is allocated when the channel gain is high, than when the channel gain is low). Another problem that suggests itself is to design power allocation policies that minimize outage probability on a given fading channel. This problem (amongst others) was addressed in [5] where it was shown that the best power allocation scheme was to use no transmit power if the channel gain falls below a threshold and to use *channel inversion* above the threshold (more power is allocated when the channel gain is low, than when the channel gain is high).

The power allocation policies resulting from maximizing ergodic capacity and from minimizing the probability of outage are very different and represent two ends of the spectrum. To bridge this gap, an optimal power and rate allocation problem was considered in [7] where long term average capacity is optimized with respect to a deterministic power allocation policy subject to a constraint on outage along with the standard average power constraint. This additional outage constraint was motivated by the idea that in an integrated network, non-real time applications will benefit from maximizing the ergodic capacity and at the same time, real time applications (such as voice and video) will benefit from a Quality of Service (QoS) guarantee on the maximum outage probability. The optimal power allocation for this problem was shown to be a mixture of channel inversion and water-filling allocation. Extensions of this problem to parallel fading channels with random power allocation policies (to include discrete fading distributions) have been considered in [8]. The results of [7] were generalized to a class of fading channels in [1] where the channels have a two-time scale nature. The slow variation in these fading channel is due to distance based attenuation and shadow fading and the resulting slow fading channel gain is assumed to be known at both transmitter and receiver. The fast fading gain (resulting from local mobility and multipath fading) is assumed to be known at the receiver but unknown at the transmitter, however the transmitter does have access to the statistics of the fast fading distribution, which is restricted to the Rayleigh distribution in [1].

While the results of [1] are restricted to single fading channels only, it is important to extend these results to obtain optimal power allocation in two time-scale parallel fading channels. The generic notion of *parallel channels* represents various different wireless transmission technologies such as multi-antenna, multi-carrier such as OFDM, MC-CDMA etc and various other diversity based transmission schemes many of which will form the basis of next generation wireless communication technologies. In order to achieve these results, a necessary step is to obtain optimal power allocation policies for outage probability minimization in two-time scale parallel fading channels. Indeed, this is the focus of this paper.

In this paper, under the usual long codeword assumption (as in [7]), we adopt the concept of a block-ergodic capacity (BEC) defined in [1] for fast Rayleigh fading and define the corresponding BEC for two-time scale parallel fading channels. We then define outage probability as the probability that this BEC falls below a minimum basic rate. We derive the optimal power allocation scheme that minimizes this outage probability subject to an average long term power constraint

Manuscript received November 11, 2005; revised March 23, 2006; accepted April 28, 2006. The associate editor coordinating the review of this letter and approving it for publication was R. Mallik. This work was supported by the Australian Research Council.

The authors are with the ARC Special Research Centre for Ultra-Broadband Information Networks (CUBIN), Department of Electrical and Electronic Engineering, University of Melbourne, Victoria 3010, Australia (e-mail:{s.dey, j.evans}@ee.unimelb.edu.au). CUBIN is an affiliated program of National ICT Australia (NICTA).

Digital Object Identifier 10.1109/TWC.2007.05910.

as a function of the slow fading gains of the parallel channels. The resulting power allocation policies for the various channels are given as a set of simultaneous nonlinear equations that can be solved numerically. We then present some simulation results for the case of two parallel (with lognormal slow fading and fast Rayleigh fading). The results demonstrate the superior outage probability performance of the optimal power allocation scheme compared to a constant power transmission scheme.

II. Channel Model and Outage Capacity

The concept of a block-fading Gaussian channel (BF-AWGN) was introduced in [4]. Essentially it refers to a a wireless propagation environment where the channel state over a block of possibly N symbols remains constant but varies from one block to another according to some random process. This idea was generalized to an M-block BF-AWGN channel in [5], where a codeword spans M blocks of N symbols each. This M-block channel model can be used to describe a multicarrier system with M parallel subchannels where the M channel gains are known at the transmitter before transmission. We consider a two-time scale version of this M-block BF-AWGN channel in the current paper.

By two-time scale fading we mean that the channel gain can be expressed as the product of a slowly-varying component and a quickly-varying component. The slowly-varying component is random but is constant over a block of length N symbols (the block length is much shorter than the coherence time of the slowly-varying channel gain). The quickly-varying component is random and varies considerably over a block of length of N symbols (the block length is much larger than the coherence time of the quickly-varying channel gain).

We consider M parallel blocks of N symbols to constitute a frame. We will work in the limit of large block length N, but assume the number of parallel channels M is finite. A codeword consists of all MN symbols within one frame: codewords do not span multiple frames.

Within a frame we can write the n-th received symbol in block m

$$y_m(n) = \sqrt{g_m f_m(n)}\, b_m(n) + z_m(n)$$

where g_m is a slow-fading gain, $f_m(n)$ is a fast-fading gain, $b_m(n)$ is a channel input symbol and $z_m(n)$ is the channel noise. The additive noise terms are assumed to be independent and identically distributed Gaussian random variables with mean zero and variance one. The fast-fading gains are assumed to be independent across blocks with identical marginal distributions following an exponential distribution (Rayleigh fading) with mean one. Within a block we assume that the process $\{f_m(n)\}_{n=0}^{\infty}$ is ergodic. The slow-fading gains are independent of the fast-fading gains and are constant within each block. The vector random process formed by considering the M slow-fading gains as they vary from one frame to the next is assumed to be ergodic with stationary first-order cumulative distribution function $Q(\mathbf{g})$ where $\mathbf{g} = [g_0, g_1, \cdots, g_{M-1}]^T$.

We assume that the slow-fading gains are known to both the transmitter and the receiver, while the fast-fading gains are known to the receiver but not to the transmitter. This allows for the possibility of varying the transmit power as a function of the slow-fading gain vector $\mathbf{g}$. The transmit power is represented by the vector $\mathbf{p} = [p_0(\mathbf{g}), p_1(\mathbf{g}), \cdots, p_{M-1}(\mathbf{g})]^T$.

With these preliminaries we can now define (under the assumption that the block lengths are large) a vector channel version of the block-ergodic capacity (BEC) introduced in [1]:

$$R_b(\mathbf{g}, \mathbf{p}) = E\left[\frac{1}{2M}\sum_{m=0}^{M-1} \log\left(1 + g_m f p_m(\mathbf{g})\right) \mid \mathbf{g}\right] \quad (1)$$

where, without loss of generality, the average power of the background white noise has been taken to be unity and f denotes the independent across the blocks and identically exponentially distributed (with mean one) fast fading process. The conditional expectation E denotes the expectation over the distribution of the fast fading process f, given $\mathbf{g}$.

This now leads to the following definitions of *Outage Capacity* and *Outage probability*:

Outage capacity: The outage capacity is defined as the maximum achievable BEC over M fading blocks, denoted by $C_\epsilon^b(P_{av})$ where

$$C_\epsilon^b(P_{av}) = \max_{\mathbf{p}} r, \ \textit{subject to } P(R_b(\mathbf{g}, \mathbf{p}) < r) \le \epsilon,$$
$$E[\frac{1}{M}\sum_{m=0}^{M-1} p_m(\mathbf{g})] \le P_{av}, \ p_m(\mathbf{g}) \ge 0, \ \forall m \quad (2)$$

The *outage probability* for a given BEC r_0 is defined as

$$O(\mathbf{p}) = P(R_b(\mathbf{g}, \mathbf{p}) < r_0) \quad (3)$$

Recalling that the fast fading process f is exponentially distributed with mean unity over all blocks, one can easily show that

$$R_b(\mathbf{g}, \mathbf{p}) = \frac{1}{2M}\sum_{m=0}^{M-1} e^{x_m} E_1(x_m) \quad (4)$$

where $x_m = \frac{1}{p_m(\mathbf{g}) g_m}$ and $E_1(x) = \int_x^\infty \frac{e^{-s}}{s} ds$.
Given a vector $\mathbf{x} = (x_1\ x_2\ \ldots\ x_M)$, we denote the arithmetic mean $\frac{1}{M}\sum_{i=1}^{M} x_m$ by $< \mathbf{x} >$. We also assume (without loss of generality) that $g_0 \ge g_1 \ge \ldots \ge g_{M-1}$. In the following the notation log will denote the natural logarithm.

III. Power Control for Minimizing Outage Probability

In this section, we are primarily concerned with the following optimization problem:
Problem P1: *Given* r_0*, minimize* $P(R_b(\mathbf{g}, \mathbf{p}) < r_0)$ *over the choice of power allocation functions* $\mathbf{p}$ *subject to* $E[< \mathbf{p} >] \le P_{av}$*,* $\mathbf{p} \ge 0$.

Remark 1: Note that the above problem is the multiple time-scale fading analog of the outage probability minimization problem with a long-term power constraint considered in [5] for a single time-scale fading situation.
As in [5], in order to solve **P1**, we first concern ourselves with the following problem:
Problem P2: *Minimize* $< \mathbf{p} >$ *subject to* $R_b(\mathbf{g}, \mathbf{p}) \ge r_0$, $\mathbf{p} \ge 0$.
The solution to Problem **P2** is given by the following Lemma:

Lemma 3.1: Suppose $g_0 \geq g_1 \geq \ldots \geq g_{M-1}$. Then the m-th component of the optimal power allocation for Problem **P2** is given by

$$p_m^*(\mathbf{g}) = \begin{cases} \lambda_1(\mathbf{g})(1 - x_m^* e^{x_m^*} E_1(x_m^*)) & if\ g_m > \frac{1}{\lambda_1(\mathbf{g})} \\ 0 & otherwise \end{cases} \tag{5}$$

where $x_m^* = \frac{1}{g_m p_m^*(\mathbf{g})}$ and

$$\lambda_1(\mathbf{g}) = \frac{\sum_{m=0}^{\mu-1} \frac{1}{g_m {x_m^*}^2}}{\sum_{m=0}^{\mu-1} \frac{1}{x_m^*} - 2Mr_0} \tag{6}$$

and μ is the unique integer such that $g_m > \frac{1}{\lambda_1(\mathbf{g})}, \forall m < \mu$ and $g_m \leq \frac{1}{\lambda_1(\mathbf{g})}, \forall m \geq \mu$.

Proof: First note that the optimization problem **P2** is a convex optimization problem since the objective function is linear in $\mathbf{p}$ and the constraint is convex (follows from (4) and the fact that $R_b(\mathbf{g}, \mathbf{p})$ is a concave function of $\mathbf{p}$ (can be shown easily from the Dominated Convergence Theorem (see [1])).

This allows us to use the standard Lagrangian based function (for some $\lambda > 0$)

$$L(\mathbf{p}, \lambda) = \frac{1}{M}[\sum_{m=0}^{M-1}(p_m - \frac{\lambda}{2} e^{x_m} E_1(x_m))],\ x_m = \frac{1}{g_m p_m} \tag{7}$$

Taking the partial derivative with respect to p_m, one can then show that

$$\frac{\partial L}{\partial p_m} = [1 - \lambda_1(\mathbf{g}) g_m x_m (1 - x_m e^{x_m} E_1(x_m))] \tag{8}$$

where $\lambda_1(\mathbf{g}) = \frac{\lambda}{2}$.

It is then straightforward to show that the power allocation given by Lemma 3.1 satisfies the Kuhn-Tucker optimality conditions (see [9], page 74 for details). To solve for $\lambda_1(\mathbf{g})$, note that the constraint in Problem **P2** has to be satisfied with equality. To prove this, assume $R_b(\mathbf{g}, \mathbf{p}) > r_0$. Then, one can decrease any of the positive powers such that the constraint is met with equality, and in the process reducing the value of the objective function. Thus, at the optimal solution, we must have $R_b(\mathbf{g}, \mathbf{p}) = r_0$.

The only thing remaining to show is that there exists a unique μ such that $g_m > \frac{1}{\lambda_1(\mathbf{g})}, \forall m < \mu$ and $g_m \leq \frac{1}{\lambda_1(\mathbf{g})}, \forall m \geq \mu$. The proof of this is slightly tedious. Therefore it is relegated to the Appendix. ■

Computation of optimal power: The elements of the optimal power vector: $p_m^*(\mathbf{g}), m = 0, 1, \ldots, M-1$ satisfy a set of simultaneous nonlinear equations given by (5), (6). One has to resort to numerical methods to compute the optimal power vector. Note that one can rewrite (5) as

$$\frac{1}{g_m \lambda_1(\mathbf{g})} = x_m^*(1 - x_m^* e^{x_m^*} E_1(x_m^*)),$$

$$p_m^*(\mathbf{g}) = \begin{cases} \frac{1}{g_m x_m^*}, & if\ g_m > \frac{1}{\lambda_1(\mathbf{g})} \\ 0 & otherwise \end{cases} \tag{9}$$

We used an iterative method by repeatedly applying (6) and (9) (note that for a given $\lambda_1(\mathbf{g})$ there is a unique solution to x_m^* from (9)) with suitable initial values for x_m^* and $\lambda_1(\mathbf{g})$ until a prespecified relative accuracy was reached. We ran this algorithm via a MATLAB7 code on an Intel Pentium M 755 (2.0 GHz) processor. On an average, it took only 1.5364 seconds to calculate the optimal power allocation for $M = 64$ channels with a final relative accuracy (in $R_b(\mathbf{g}, \mathbf{p})$) of the order of 10^{-5}. It was also seen to be remarkably fast in computing the optimal power allocation for even larger values of M.

In order to proceed, we present another result that is similar to its counterpart in [5].

Denote the vector $(p_0^*(\mathbf{g})\ p_1^*(\mathbf{g}) \ldots p_{M-1}^*(\mathbf{g}))$ by $\mathbf{p}^*$.

Lemma 3.2: The optimal power allocation function $p_m^*(\mathbf{g}),\ m = 0, 1, \ldots, M-1$ (as defined in Lemma 3.1 above) is a continuous function of $\mathbf{g}$. In addition, $\langle \mathbf{p}^* \rangle$ is a nonincreasing function of $g_m,\ m = 0, 1, \ldots, M-1$, in fact the following is true; for all $p_j^*(\mathbf{g}),\ j = 0, 1, \ldots, M-1$:

$$\frac{\partial \langle \mathbf{p}^* \rangle}{\partial g_j} = -\frac{1}{M} \frac{p_j^*(\mathbf{g})}{g_j} \tag{10}$$

.

Proof: The proof of the first statement is similar to that of [5] and is excluded. The proof of the second statement is a direct consequence of the fact $\sum_{i=0}^{\mu-1} e^{x_i^*} E_1(x_i^*) = 2Mr_0,\ x_i^* = \frac{1}{g_i p_i^*(\mathbf{g})}$ and is given in the Appendix. ■

Remark 2: Note that if there is no transmission on the j-th channel, that is, $p_j^*(\mathbf{g}) = 0$, then the average sum-power over all channels is a constant as a function of g_j, that is, the derivative in (10) is equal to zero.

Remark 3: Notice that the optimal power policy here is a function of the slow fading gain $\mathbf{g}$, which is assumed to be known exactly at the transmitter. While this is a more preferable option than tracking the fast fading gain, in practice, one has to estimate $\mathbf{g}$ by averaging the fast fading process. Error in this estimation process can give rise to inaccuracies in the values of $\mathbf{g}$ used to compute the optimal power allocations, thus contributing to sub-optimal performance. Given that $\mathbf{g}$ is the slowly varying gain, we can assume that it can be estimated fairly accurately, in which case the performance of the power allocation algorithm will be close to optimal. A rigorous analysis taking into account the statistics of the estimation error process is, however, beyond the scope of this paper.

Example: We provide an example here to demonstrate the power profiles for $M = 2$, where $g_0 = 1$ is kept fixed and g_1 is increased and g_1 is varied between 0.0001 to $g_0 = 1$. r_0 is taken to be 1 bit per symbol.

For this case, Fig. 1 illustrates the optimal power allocation functions p_0^*, p_1^* and the average power sum $\frac{p_0^* + p_1^*}{2}$ as a function of g_1. Note that the powers are expressed in units of the noise variance (assuming the noise variance is one unit). It is clearly visible that p_1^* remains zero (no transmission) until g_1 attains a threshold and then transmissions are on at both channels $M = 1, 2$. As g_1 approaches g_0, p_1^* approaches p_0^*. A similar graph can be obtained by keeping g_1 fixed and varying g_0. It is also seen that the average power sum $\frac{p_0^* + p_1^*}{2}$ is a nonincreasing function of g_1.

As in [5], we define the following regions:

$$\begin{aligned} \mathcal{R}(s) &= \{\mathbf{g} \in \mathbb{R}_+^M : \langle \mathbf{p}^* \rangle < s\}, \\ \bar{\mathcal{R}}(s) &= \{\mathbf{g} \in \mathbb{R}_+^M : \langle \mathbf{p}^* \rangle \leq s\} \end{aligned} \tag{11}$$

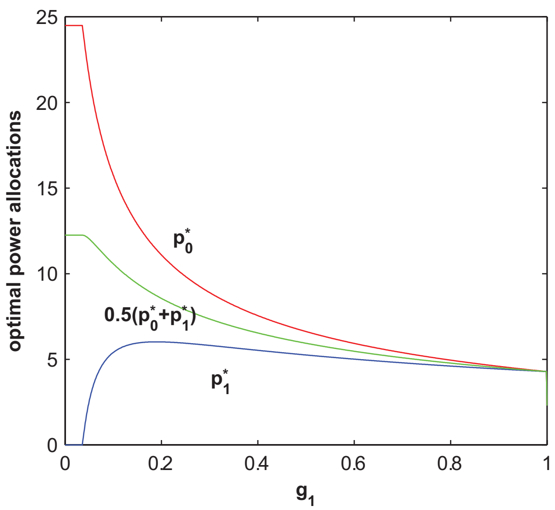

Fig. 1. Plot of optimal power allocation versus the slow fading gain g_1 ($M = 2, g_0 = 1$).

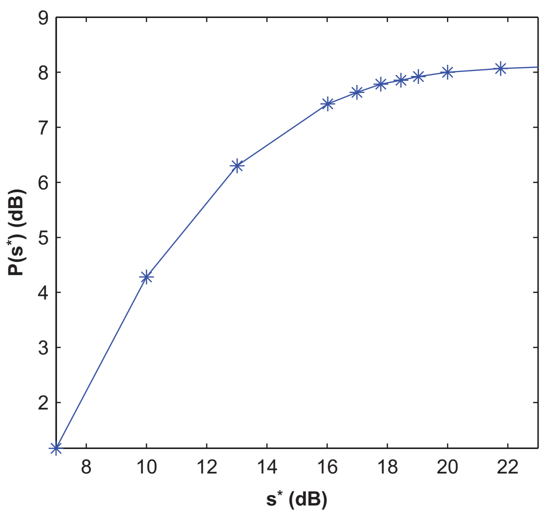

Fig. 2. Plot of $\mathcal{P}(s^*)$ versus s^* ($M = 2$).

and the boundary surface of $\mathcal{B}(s)$ of $\bar{\mathcal{R}}(s)$ as the set of points $\mathbf{g}$ such that $\langle \mathbf{p}^* \rangle = s$. We also define the two average power sums $\mathcal{P}(s)$ and $\bar{\mathcal{P}}(s)$ as (recall that the cdf of $\mathbf{g}$ is given by $Q(\mathbf{g})$)

$$\mathcal{P}(s) = \int_{\mathcal{R}(s)} \langle \mathbf{p}^* \rangle dQ(\mathbf{g}),\ \bar{\mathcal{P}}(s) = \int_{\bar{\mathcal{R}}(s)} \langle \mathbf{p}^* \rangle dQ(\mathbf{g}) \quad (12)$$

Finally the threshold s^* and weight w^* are defined as

$$s^* = \sup\{s : \mathcal{P}(s) < P_{av}\},\ w^* = \frac{P_{av} - \mathcal{P}(s^*)}{\bar{\mathcal{P}}(s^*) - \mathcal{P}(s^*)} \quad (13)$$

With Lemmas 3.1 and 3.2 and the previous definitions ((11)-(13)), we can now state the main result of this paper:

Theorem 1: **Problem P1** is solved by the following power allocation scheme:

$$\mathbf{p}_{opt} = \begin{cases} \mathbf{p}^*, & if\ \mathbf{g} \in \mathcal{R}(s^*) \\ 0, & if\ \mathbf{g} \notin \bar{\mathcal{R}}(s^*) \end{cases} \quad (14)$$

while if $\mathbf{g} \in \mathcal{B}(s^*)$, $\mathbf{p}_{opt} = \mathbf{p}^*$ with probability w^* and $\mathbf{p}_{opt} = 0$ with probability $1 - w^*$.

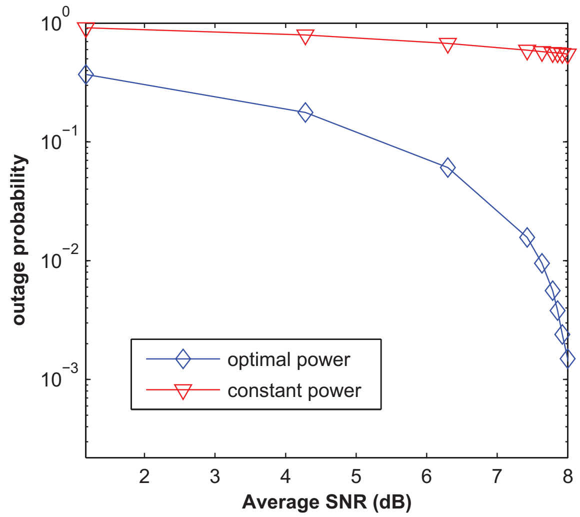

Fig. 3. Plot of outage probability versus average SNR ($M = 2$).

Proof: The proof requires an intermediate functional optimization result (see Lemma 3 (pp. 1474) in [5])) that explains the reason for the randomization in the power allocation scheme if $\mathbf{g} \in \mathcal{B}(s^*)$. The rest of the proof is similar to the proof of the main result (Proposition 4) in [5] and is excluded for lack of space. ■

Remark 4: Note that if the cdf $Q(\mathbf{g})$ is continuous, then there is no need for randomization as the probability that $\mathbf{g} \in \mathcal{B}(s^*)$ is zero. Randomization is needed when the cdf is discontinuous and has discrete point masses. The minimum outage probability for a general $Q(\mathbf{g})$ is given by $1 - w^* P(\mathbf{g} \in \mathcal{B}(s^*)) - P(\mathbf{g} \in \mathcal{R}(s^*))$.

Remark 5: Note that the single channel case ($M = 1$) was dealt with in [1] separately, and the same result can also be obtained as a special case of Theorem 1.

IV. Simulation Studies

In this section, we present some simulation studies conducted with a $M = 2$ block-fading channels with fast Rayleigh fading. The slow fading gains g_0, g_1 are distributed with independent lognormal distributions (standard for shadow fading), such that $\log g_0$ is distributed with mean 0 and variance $\sigma_{g_0} = 3dB$, and $\log g_1$ is distributed with mean 0 and standard deviation $\sigma_{g_1} = 1dB$. The fast fading gains for both channels are independently exponentially distributed with mean 1 (Rayleigh fading). The minimum basic rate is taken to be $r_0 = 1$ bit/sec/Hz. The average signal-to-noise ratio (SNR) P_{av} is varied between 1 dB and 8 dB. The following results were obtained through Monte Carlo simulations over 100000 realizations of the slow fading gains and averaged over 5 random sets.

Fig 2 shows how $\mathcal{P}(s^*)$ varies with s^*. This graph can be used to obtain the optimal threshold s^* for a given P_{av}. Note that for continuous probability distribution functions of the slow fading gains, $\mathcal{P}(s^*) = P_{av}$.

Fig. 3 shows the outage probability achieved by the optimal power allocation and compares it with that achieved by constant power allocation ($P_{av}/2$ in both channels all the time). Clearly, the constant power allocation performs very poorly.

With the optimal power allocation, an outage probability of 0.01 is achieved at an average power of approximately 7.6 dB.

APPENDIX I
PROOF OF EXISTENCE AND UNIQUENESS OF μ IN LEMMA 3.1

Recall that $\mu = \left|\{m : g_m > \frac{1}{\lambda_1(\mathbf{g})}\}\right|$ where $\lambda_1(\mathbf{g})$ is given by (6). Following similar techniques in [5], we define

$$\nu(\mu) = \left|\{m : g_m > \frac{\sum_{m=0}^{\mu-1} \frac{1}{x_m^*} - 2Mr_0}{\sum_{m=0}^{\mu-1} \frac{1}{g_m {x_m^*}^2}}\}\right|$$

where $|.|$ of a set indicates the number of elements in that set.

We now need to show that there exists a unique integer solution to the equation $\nu(\mu) = \mu$. Before we begin, recall that by definition, $1 \leq \nu(\mu) \leq M$. Now define

$$\delta(\mu) = \frac{\sum_{m=0}^{\mu-1} \frac{1}{x_m^*} - 2Mr_0}{\sum_{m=0}^{\mu-1} \frac{1}{g_m {x_m^*}^2}}$$

It can be shown (after some algebra) that

$$(\delta(\mu+1) - \delta(\mu)) \left(\sum_{m=0}^{\mu-1} \frac{1}{g_m {x_m^*}^2}\right) = \frac{1}{x_\mu^*}\left[1 - \delta(\mu+1)p_\mu^*(\mathbf{g})\right] \quad (15)$$

It is trivial to see that $\delta(1) < \frac{1}{p_0^*(\mathbf{g})}$. Now if one assumes that $\delta(\mu) < \frac{1}{p_{\mu-1}^*(\mathbf{g})}$, after some more algebraic manipulation, one can show that $\delta(\mu+1) < \frac{1}{p_\mu^*(\mathbf{g})}$. It follows by induction then that for any $\mu \in \{0, 2, \ldots, M-1\}$ that $\delta(\mu+1) < \frac{1}{p_\mu^*(\mathbf{g})}$.

Using this with (15), one obtains $\delta(\mu+1) > \delta(\mu)$. By definition, $\nu(\mu) = |\{m : g_m > \delta(\mu)\}|$. Since $g_0 \geq g_1 \geq \ldots \geq g_{M-1}$, we obtain (using the above facts) $\nu(\mu+1) \leq \nu(\mu)$. Combining this with the fact that $1 \leq \nu(\mu) \leq M$, it follows that $\nu(\mu) = \mu$ has a unique solution.

Proof of (10):
From Lemma 3.1, we know that

$$\sum_{i=0}^{\mu-1} e^{x_i^*} E_1(x_i^*) = 2Mr_0$$

Differentiating both sides w.r.t g_j for $j = 1, 2, \ldots, M-1$, we have (after some rearranging)

$$\sum_i \left(\frac{1}{x_i^*}\frac{\partial x_i^*}{\partial g_j}\right)(1 - x_i^* e^{x_i^*} E_1(x_i^*)) = 0$$

Using (5), one can then write $\sum_i p_i^*(\mathbf{g}) \frac{\partial \log x_i^*}{\partial g_j} = 0$. Using the fact $\log x_i^* = -\log g_i - \log p_i^*(\mathbf{g})$, one then has the following:

$$\frac{\partial \log x_i^*}{\partial g_j} = \begin{cases} -\frac{1}{g_i} - \frac{1}{p_i^*(\mathbf{g})}\frac{\partial p_i^*(\mathbf{g})}{\partial g_i} & if \;\; j = i \\ \frac{1}{p_i^*(\mathbf{g})}\frac{\partial p_i^*(\mathbf{g})}{\partial g_j} & if \;\; j \neq i \end{cases} \quad (16)$$

Substituting the above result in $\sum_i p_i^*(\mathbf{g}) \frac{\partial \log x_i^*}{\partial g_j} = 0$, one obtains

$$\frac{\partial p_j^*(\mathbf{g})}{\partial g_j} + \sum_{i \neq j} \frac{\partial p_i^*(\mathbf{g})}{\partial g_j} = -\frac{p_j^*(\mathbf{g})}{g_j}$$

(10) now follows immediately.

REFERENCES

[1] S. Dey and J. Evans, "Optimal power control over multiple time-scale fading channels with service outage constraints," *IEEE Trans. Commun.*, vol. 53, no. 4, pp. 708–717, Apr. 2005.

[2] A. J. Goldsmith and P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Trans. Inform. Theory*, vol. 43, pp. 1986–1992, Nov. 1997.

[3] S. V. Hanly and D. N. C. Tse, "Multi-access fading channels - Part II: Delay-limited capacities," *IEEE Trans. Inform. Theory*, vol. 44, pp. 2816–2831, Nov. 1998.

[4] L. H. Ozarow, S. Shamai, and A. D. Wyner, "Information theoretic considerations for cellular mobile radio," *IEEE Trans. Veh. Technol.*, vol. 43, pp. 359–378, May 1994.

[5] G. Caire, G. Taricco, and E. Biglieri, "Optimum power control over fading channels," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1468–1489, July 1999.

[6] G. Caire and S. Shamai, "On the capacity of some channels with channel state information," *IEEE Trans. Inform. Theory*, vol. 45, pp. 2007–2019, Sep. 1999.

[7] J. Luo, L. Lin, R. Yates, and P. Spasojevic, "Service outage based power and rate allocation," *IEEE Trans. Inform. Theory*, vol. 49, pp. 323–330, Jan. 2003.

[8] J. Luo, R. Yates, and P. Spasojevic, "Service outage based power and rate allocation for parallel fading channels," *IEEE Trans. Inform. Theory*, vol. 51, pp. 2594–2611, July 2005.

[9] M. J. Fryer and J. V. Greenman, *Optimization Theory: Applications in OR and Economics*. London, UK: Edward Arnold Ltd., 1987.