Efficient Algorithms for Outage Minimization in Parallel Fading Channels with Limited Feedback

YuanYuan He and Subhrakanti Dey, *Senior Member, IEEE*
ARC Special Research Centre for Ultra-Broadband Information Networks (CUBIN)
Department of Electrical and Electronic Engineering
University of Melbourne, Vic. 3010, Australia
e-mail: {yyhe, s.dey}@ee.unimelb.edu.au

Abstract— In this paper, we address the optimal power allocation problem for minimizing the notion of information theoretic outage for an M parallel block-Rayleigh-fading channels using a finite rate power codebook (limited feedback). In contrast to the existing literature on outage minimization for MIMO systems with limited feedback, we derive effective approximations to the channel quantization regions in order to design a number of low-complexity power allocation algorithms for various ranges of average power constraints. Unlike previous work, we show that it is not generally optimal to allocate same power to all channels, and that this is only asymptotically optimal at high average power (average SNR). We also derive a suitable Gaussian approximation based power allocation scheme for large number of parallel channels which has important practical applications in multi-carrier systems such as OFDM. Extensive numerical results illustrate that only a few bits of feedback (for $M = 4$ or $M = 6$) closes the gap substantially in outage performance with the full instantaneous channel information at the transmitter. For large number of channels, less than 1 bit of (broadcast) feedback per channel can achieve the same outage probability (10^{-4}) with approximately only a 2.5 dB average power (or average SNR) gap.

I. INTRODUCTION

Determining the information theoretic capacity of block-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* [1]. For delay-sensitive traffic such as voice and video, the last two notions are rather important. In particular, the notion of outage probability signifies the probability that the capacity of a wireless channel falls below a required rate threshold. In [1], optimal power allocation for outage minimization in the case of parallel fading channels (single user) was obtained with the assumption of full channel state information (CSI) at the transmitter. However, full CSI at the transmitter is hard to obtain due to limited bandwidth in the feedback channel from the receiver to the transmitter, as it is more common to have full CSI at the receiver. Motivated by this, there are a number of works that have looked at outage minimization for fading channels with limited feedback. Such works include [4], [5], [6], [7]. In particular, [7] looks at outage minimization with a finite-rate power codebook for MIMO systems. It assumes however

the same transmit power (as a function of all channel gains) is used in all transmit antennas. This allows the authors of [7] to reduce the finite-rate power codebook design problem to an equivalent scalar quantization problem. Even then, finding the cumulative density function for the equivalent scalar random variable requires computing multi-dimensional probability integrals. In [4], a Gaussian approximation is used to capture the probability distribution of the mutual information for a MIMO system.

In this paper, although we look at an M-parallel fading channel (thus consider a simpler setting than MIMO), based on a simple linearized approximations to the quantized channel space boundaries, we derive a number of novel low-complexity locally optimal or suboptimal finite-rate power codebook design algorithms for outage minimization without having to assume the same transmit power per channel or use the Gaussian approximation in general. Based on these linearized approximations, we show that only in high average power (or average SNR) it is asymptotically optimal to allocate equal transmit power to all channels. We also derive the diversity order for this sub-optimal algorithm. The Gaussian approximation is seen to perform poorly for small number of parallel channels compared to our low-complexity algorithms based on a simple approximation to the quantized regions. We also derive a Gaussian approximation based optimal power allocation scheme for large number of channels (e.g. $M \geq 16$) which has important practical applications to multi-carrier systems such as OFDM. Extensive numerical results are presented which illustrate that only 4 bits of feedback close the gap with the outage performance of the full CSI algorithm substantially for $M = 4$ or $M = 6$. For a large number of channels, our Gaussian approximation based algorithm performs approximately within 2.5 dB (SNR gap) of the full CSI based algorithm at an outage probability of 10^{-4}.

II. CHANNEL MODEL AND OUTAGE MINIMIZATION

We consider an M-parallel flat-fading channel model similar to that in [8], where a transmitted codeword spans M subchannels in one fading block. For each subchannel $i, i = 1, 2, ..., M$, the received signal can be written as:

$$y_i = \sqrt{h_i}x_i + w_i \qquad (1)$$

where h_i denotes the channel power gain and x_i is the channel input symbol. The noise sequences $w_1, \ldots, w_M$ are independent and identically distributed (i.i.d) Gaussian random variables with zero mean and unit variance. It is assumed that the components of channel state vector $\mathbf{h} = (h_1, \ldots, h_M)$ are mutually independent, individually i.i.d and ergodic and fading is sufficiently slow so that the input symbols transmitted over the same fading block experience the same channel state.

Given a channel realization $\mathbf{h}$, and assuming the availability of full channel state information (CSI) at the transmitter and receiver, denote the corresponding power allocation to the M subchannels by the vector $\mathbf{p}(\mathbf{h}) = (p_1(\mathbf{h}), \ldots, p_M(\mathbf{h}))$. Then the maximum mutual information of an M-parallel Gaussian channel is given by [8]

$$r(\mathbf{h}, \mathbf{p}(\mathbf{h})) = \frac{1}{M} \sum_{i=1}^{M} \frac{1}{2} \log(1 + h_i p_i(\mathbf{h})) \qquad (2)$$

where, the rate unit is nats per channel use.

Thus, the outage probability, defined as the probability that the instantaneous mutual information of the channel is less than a pre-specified transmission rate r_0 (nats/channel use), can be expressed as

$$P_{out} = Prob\left[r(\mathbf{h}, \mathbf{p}(\mathbf{h})) < r_0\right] \qquad (3)$$

Under a long term average power constraint defined by $E[\langle \mathbf{p}(\mathbf{h}) \rangle] \leq P_{av}$ (where $\langle x \rangle$ denotes the arithmetic mean of the vector x with length M, namely, $\langle x \rangle = \frac{1}{M} \sum_{i=1}^{M} x$), the outage minimization problem can be described as

$$\min_{\mathbf{p}(\mathbf{h})} \; Prob\left[\frac{1}{M} \sum_{i=1}^{M} \frac{1}{2} \log(1 + h_i p_i(\mathbf{h})) < r_0\right]$$
$$s.t. \quad E[\langle \mathbf{p}(\mathbf{h}) \rangle] \leq P_{av}, \; \mathbf{p}(\mathbf{h}) \geq 0 \qquad (4)$$

The optimal power allocation for this problem can be found explicitly by using convex optimization techniques and was presented in Proposition 4 of [1]. The readers are referred to [1] for further details. Note that here P_{av} can be thought of effectively as the average signal-to-noise ratio (since noise variance has been normalized to unity). In the following we will address the optimal power allocation problem for outage minimization where only partial or limited CSI is available at the transmitter. For the purpose of analysis, we will assume that each channel h_i is exponentially distributed (Rayleigh fading) with mean $\frac{1}{\lambda_i}$. However, the analysis in this paper can be easily extended to other fading models such as Nakagami, log-normal etc. Proofs of various results are excluded due to space limitations. However they can be found in the following document online [10].

III. OPTIMUM QUANTIZED POWER CONTROL WITH FINITE-RATE FEEDBACK

It is well known that having perfect CSI at both transmitter and receiver is hard to satisfy in practical system. In this section, we consider the optimal power allocation procedure for M-parallel flat-fading channels using limited feedback. The finite feedback strategy (see also [4],[7]) is as follows: a fixed power codebook $\mathcal{P} = \{\mathbf{P}_1, \ldots, \mathbf{P}_L\}$ (size L) (designed offline) with rate $B = \log_2 L$ is known at both transmitter and receiver; the feedback channel is error-free and delay-free, L distinct feedback signals can be conveyed to the transmitter to indicate each power levels in codebook $\mathcal{P}$; We assume full CSI at the receiver. Given a channel realization $\mathbf{h}$, the receiver employs a mapping $\mathbf{h} \rightarrow J(\mathbf{h})$ [4], where $J(\mathbf{h}) \in \{1, \ldots, L\}$ is the index signal, and sends J (B bits codeword) to the transmitter via the feedback link. Then, supposing $J = j$, the associated power level $\mathbf{P}_j$ in codebook $\mathcal{P}$ will be employed by the transmitter in this fading block. The key techniques here involve the $\mathbf{h} \rightarrow J(\mathbf{h})$ mapping and the optimal design of the power codebook. [4],[7] provides the optimal $\mathbf{h} \rightarrow J(\mathbf{h})$ mapping result. Our objective is to design the optimal power allocation scheme (optimal power codebook) so as to minimize the outage probability while satisfying the long term average power constraint.

A. Problem Formulation

The basic idea of $\mathbf{h} \rightarrow J(\mathbf{h})$ mapping is that the space defined by all possible sets of channel states $\mathbf{h} = (h_1, \ldots, h_M)$ is partitioned into L regions $\mathcal{R}_1, \mathcal{R}_2, \ldots, \mathcal{R}_L$, resulting in a feedback codeword of $B = \log_2(L)$ bits, by a set of quantization thresholds $\{r_{1j}, \ldots, r_{Mj}\}_{j=1}^{L}$. Let $\mathbf{P}_j = \{p_{1j}, \ldots, p_{Mj}\}$ present the corresponding power levels for $\mathcal{R}_j$, $j = 1, 2, \ldots, L$ (note that power levels for different channels here are distinct as opposed to [4], [7] where the same transmit power was allocated to all transmit antennas in the MIMO setting). [4], [7] also showed that the optimal quantization region has a 'circular' structure. Assuming $\mathbf{P}_1 > \mathbf{P}_2 > \ldots > \mathbf{P}_L$, power level $\mathbf{P}_j$ is used to assure zero outage for all possible channels realizations $\mathbf{h}$ in region $\mathcal{R}_j$, $j = 1, 2, \ldots, L - 1$, implying

$$\mathcal{R}_j = \{\mathbf{h} : r(\mathbf{h}, \mathbf{P}_j) \geq r_0, and \; r(\mathbf{h}, \mathbf{P}_{j+1}) < r_0\}; \quad (5)$$

The last region $\mathcal{R}_L$ has two parts $\mathcal{R}_{L,1}$ and $\mathcal{R}_{L,2}$, where

$$\mathcal{R}_{L,1} = \{\mathbf{h} : r(\mathbf{h}, \mathbf{P}_1) < r_0\}$$
$$\mathcal{R}_{L,2} = \{\mathbf{h} : r(\mathbf{h}, \mathbf{P}_L) \geq r_0\} \qquad (6)$$

and the outage occurs only in $\mathcal{R}_{L,1}$, as illustrated in Fig.1. Thus the boundary between $\mathcal{R}_{j-1}$ and $\mathcal{R}_j$, $j = 2, \ldots, L-1$ is a hypersurface $g(h_1, \ldots, h_{M-1}, \mathbf{P}_j)$, which is obtained by solving for h_M in the following equation:

$$r(\mathbf{h}, \mathbf{P}_j) = \frac{1}{M} \sum_{i=1}^{M} \frac{1}{2} \log(1 + h_i p_{ij}) = r_0 \qquad (7)$$

Namely,

$$g(h_1, \ldots, h_{M-1}, \mathbf{P}_j) = \frac{k - \prod_{i=1}^{M-1}(1 + h_i p_{ij}) + 1}{p_{Mj} \prod_{i=1}^{M-1}(1 + h_i p_{ij})} \qquad (8)$$

where $k = e^{2Mr_0} - 1$. The boundaries between $\mathcal{R}_{L,1}$ and $\mathcal{R}_1$, $\mathcal{R}_{L-1}$ and $\mathcal{R}_{L,2}$ are given by $g(h_1, \ldots, h_{M-1}, \mathbf{P}_1)$ and $g(h_1, \ldots, h_{M-1}, \mathbf{P}_L)$ respectively. Interestingly, from (8), it can be easily verified that the quantization thresholds $r_{ij} =$

763

$\frac{k}{p_{ij}}$, $i \in \{1, 2, \ldots, M\}$, $j \in \{1, 2, \ldots, L\}$. The hypersurface $g(h_1, \ldots, h_{M-1}, \mathbf{P}_1)$ defines the outage region.
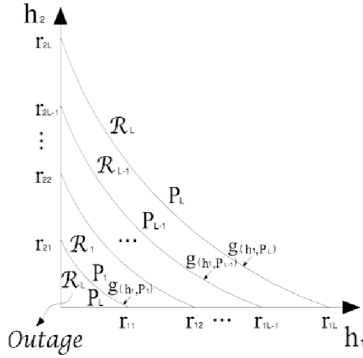

Fig. 1. Structure of the vector quantization region for 2 channels with $B = \log_2 L$ bits of feedback. The entire $\{h_1, h_2\}$ space is quantized into L regions; the region $\mathcal{R}_j$, $j = 1, \ldots, L$ employs power level $\mathbf{P}_j$; the area below the curve $g(h_1, \mathbf{P}_1)$ defines the outage region.

Denoting $F(\mathbf{P}_j)$ as the probability that the channel state $(h_1, \ldots, h_M)$ lies below $g(h_1, \ldots, h_{M-1}, \mathbf{P}_j)$, $j = 1, \ldots, L$, we have

$$F(\mathbf{P}_j) = \int \ldots \int_{\mathcal{R}_{L,1} \cup \ldots \cup \mathcal{R}_{j-1}} f(h_1, \ldots, h_M) dh_1 \ldots dh_M \tag{9}$$

where $f(h_1, \ldots, h_M)$ is the joint probability density function for the channel vector given by $f(h_1, \ldots, h_M) = \prod_{i=1}^{M} \lambda_i e^{-\lambda_i h_i}$ for Rayleigh fading. Thus, the probability that the channel realization $\mathbf{h} \in \mathcal{R}_j$ is $F(\mathbf{P}_{j+1}) - F(\mathbf{P}_j)$ for $j = 1, \ldots, L-1$, and $1 - F(\mathbf{P}_L) + F(\mathbf{P}_1)$ for $j = L$.

The outage minimization problem (4) with limited feedback can thus be formulated as

$$\min_{\{\mathbf{P}_1, \ldots, \mathbf{P}_L\}} P_{out} = F(\mathbf{P}_1)$$

$$s.t. \sum_{j=1}^{L-1} \left(\sum_{i=1}^{M} p_{ij} \right) (F(\mathbf{P}_{j+1}) - F(\mathbf{P}_j))$$

$$+ \left(\sum_{i=1}^{M} p_{iL} \right) (1 - F(\mathbf{P}_L) + F(\mathbf{P}_1)) = M P_{av} \tag{10}$$

It can be easily verified that the average power constraint is satisfied with equality. Problem (10) is in general a nonlinear non-convex optimization problem. Since the function $g(h_1, \ldots, h_{M-1}, \mathbf{P}_j), j = 1, \ldots, L$ is highly nonlinear, solving the above optimization problem with a closed-form expression for $F(\mathbf{P}_j)$ is pretty hard. Although one can use numerical integrals to calculate $F(\mathbf{P}_j)$, and use randomized search techniques to find the optimum solution, the associated computational complexity increases exponentially with the number of feedback bits and channels. Next, we will derive an approximation for $g(h_1, \ldots, h_{M-1}, \mathbf{P}_j)$, such that an analytical (approximate) closed-form expression for $F(\mathbf{P}_j)$ can be easily obtained, (unlike [4] which used a Gaussian distribution to approximate the distribution of the mutual information to get $F(\mathbf{P}_j)$) thus significantly reducing

the computational complexity of solving problem(10). Based on the obtained optimal power allocation using this approximation, one can use Monte Carlo simulations to evaluate the corresponding outage probability performance given by $F(\mathbf{P}_1)$. More details on this can be found in the *Numerical Results* Section.

B. Initial Approach: Optimal straight line approximation (OSLA)

Note that the projection of hypersurface $h_M = g(h_1, \ldots, h_{M-1}, \mathbf{P}_j)$, $j = 1, \ldots, L$ on any arbitrary two channel coordinate plane, i.e. h_n versus h_m, $n, m \in \{1, 2, \ldots, M\}, n \neq m$, is convex. From (8), the projection of hypersurface g on h_n versus h_m coordinate plane is a curve expressed as

$$h_n = \frac{k - h_m p_{mj}}{p_{nj}(1 + h_m p_{mj})} \tag{11}$$

We can approximate the curve (11) by a straight line $r_{nj} r_{mj}$, similar to [6], as shown in Fig.2 (a), in this case

$$h_n = \frac{k - h_m p_{mj}}{p_{nj}}, \quad h_m \in [0, \frac{k}{p_{mj}}] \tag{12}$$

we call this 'OSLA1'. We can also approximate the curve (11) by another straight line $r'_{nj} r'_{mj}$, as displayed in Fig.2(b), which is parallel to $r_{nj} r_{mj}$ and a tangent to the curve (11) at the intersection point 'a'. In this case,

$$h_n = \frac{k' - h_m p_{mj}}{p_{nj}}, \quad h_m \in [0, \frac{k'}{p_{mj}}] \tag{13}$$

where $k' = 2(\sqrt{k+1} - 1)$, $r'_{nj} = \frac{k'}{p_{nj}}, r'_{mj} = \frac{k'}{p_{mj}}$, and point 'a' is $(\frac{k'}{2p_{mj}}, \frac{k'}{2p_{nj}})$. We name this approximation as 'OSLA2'.

Thus, with OSLA1 or OSLA2, the boundaries between $\mathcal{R}_{L,1}$ and $\mathcal{R}_1$, $\mathcal{R}_1$ and $\mathcal{R}_2$, ..., $\mathcal{R}_{L-1}$ and $\mathcal{R}_{L,2}$ can be approximated as

$$g'(h_1, \ldots, h_{M-1}, \mathbf{P}_j) = \frac{K - \sum_{i=1}^{M-1} h_i p_{ij}}{p_{Mj}}, j = 1, \ldots, L \tag{14}$$

where $K = k$ or k' corresponding to two kinds of OSLA. Any channel vector below $g'(h_1, \ldots, h_{M-1}, \mathbf{P}_1)$ is said to be in outage. Since now $g'(h_1, \ldots, h_{M-1}, \mathbf{P}_j)$ is linear, an analytical closed-form approximation for $F(\mathbf{P}_j)$ can be obtained, which is denoted as $F'(\mathbf{P}_j)$. For example, for $M = 2$, an approximation for $F(\mathbf{P}_j)$ using OSLA is given by

$$F(\mathbf{P}_j) \approx -e^{-\frac{\lambda_1 K}{p_{1j}}} - \frac{\lambda_1 p_{2j}}{-\lambda_1 p_{2j} + \lambda_2 p_{1j}} \left(e^{-\frac{\lambda_1 K}{p_{1j}}} - e^{-\frac{\lambda_2 K}{p_{2j}}} \right) \tag{15}$$

It can be easily verified that the area between $r'_{nj} r'_{mj}$ and the curve (11) is smaller then the area between the curve (11) and $r_{nj} r_{mj}$, resulting in OSLA2 being a better approximation than OSLA1 for the curve (11). Simulations have illustrated that the outage performance of the optimal power allocation based on OSLA2 is superior to that based on OSLA1, especially when $M \geq 3$. Thus, all the results on

764
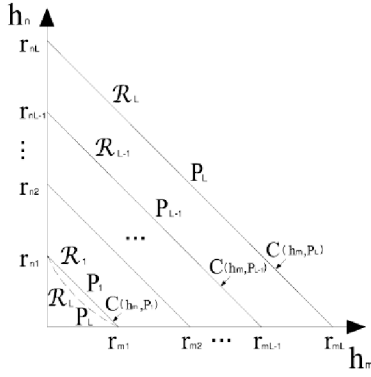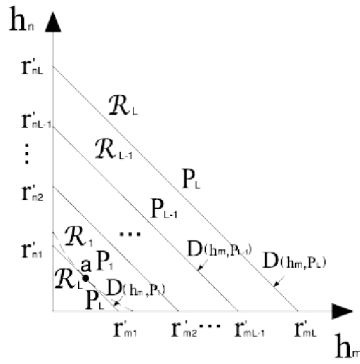
(a) OSLA1

(b) OSLA2

Fig. 2. Two kinds of OSLA: OSLA1 using straight line $r_{nj}r_{mj}$ approximation, OSLA2 approximating the curve with straight line $r'_{nj}r'_{mj}$

OSLA in the *Numerical Results* section are based on OSLA2. With OSLA, problem (10) becomes

$$\min_{\{\mathbf{P}_1,\ldots,\mathbf{P}_L\}} F'(\mathbf{P}_1)$$

$$s.t. \quad \sum_{j=1}^{L-1}(\sum_{i=1}^{M} p_{ij})(F'(\mathbf{P}_{j+1}) - F'(\mathbf{P}_j))$$

$$+ (\sum_{i=1}^{M} p_{iL})(1 - F'(\mathbf{P}_L) + F'(\mathbf{P}_1)) = MP_{av} \quad (16)$$

Although this problem is much more tractable, it is still a challenging nonconvex optimization problem and finding a global minimum is difficult. However, we can employ the Karush-Kuhn-Tucker (KKT) necessary conditions to achieve locally optimal solutions. Using these KKT conditions, we have the following result

Lemma 1: Let $\{p_{1j}{}^*, \ldots, p_{Mj}{}^*\}_{j=1}^{L}$ be an optimal scheme for Problem (16). Then we have

$$\frac{\partial F'(\mathbf{P}_j)}{\partial p_{1j}{}^*} = \ldots = \frac{\partial F'(\mathbf{P}_j)}{\partial p_{Mj}{}^*}, j = 1, \ldots, L \quad (17)$$

Combining the above result with the average power constraint in (16), we have following system of $M \times L$ nonlinear equations.

$$\begin{cases} \sum_{j=1}^{L-1}(p_{1j} + \ldots + p_{Mj})(F'(\mathbf{P}_{j+1}) - F'(\mathbf{P}_j)) \\ + (p_{1L} + \ldots + p_{ML})(1 - F'(\mathbf{P}_L) + F'(\mathbf{P}_1)) = MP_{av} \\ \frac{\partial F'(\mathbf{P}_j)}{\partial p_{1j}{}^*} = \ldots = \frac{\partial F'(\mathbf{P}_j)}{\partial p_{Mj}{}^*}, \quad j = 1, \ldots, L \\ \frac{\partial F'(\mathbf{P}_j)}{\partial p_{Mj}{}^*} = -\frac{F'(\mathbf{P}_{j+1}) - F'(\mathbf{P}_j)}{\sum_{i=1}^{M}(p_{i,j-1} - p_{ij})}, \quad j = 2, \ldots, L-1 \\ \frac{\partial F'(\mathbf{P}_L)}{\partial p_{ML}{}^*} = -\frac{1 - F'(\mathbf{P}_L) + F'(\mathbf{P}_1)}{\sum_{i=1}^{M}(p_{i,L-1} - p_{iL})} \end{cases} \quad (18)$$

A solution to (18) provides a locally optimum power allocation policy $\{\mathbf{P}_j{}^*\}_{j=1}^{L}$. For small number of regions and channels, the above system of nonlinear equations can be solved by various optimization softwares,and here we employ the 1stOpt (First Optimization) software, which uses Levenberg-Marquardt + Universal Global Optimization method and not requires to guess the initial values. However, the complexity of solving the above set of nonlinear equations is still too high for large numbers of feedback bits and channels, especially when $M \times L > 20$. Therefore, we consider two different low-complexity suboptimal schemes described below: one for high average power and one for low average power.

C. High Average Power Approximation ($HP_{av}A$)

Theorem 1: For an arbitrary M, in high average power (as $P_{av} \to \infty$), by using the M-th order Taylor's series approximation for the exponential function, namely, $e^{-x} \approx 1 - x + \frac{x^2}{2!} + \ldots + (-1)^M \frac{x^M}{M!}$, $F'(\mathbf{P}_j), j = 1, \ldots, L$ for M channels can be further approximated as

$$F'(\mathbf{P}_j) \approx \frac{c}{p_{1j}p_{2j} \ldots p_{Mj}} \quad (19)$$

where $c = \frac{\lambda_1 \lambda_2 \ldots \lambda_M (K_M)^M}{M!}$, $K_M = e^{2Mr_0} - 1$ or $2(\sqrt{e^{2Mr_0} - 1 + 1} - 1)$, corresponding to OSLA1 and OSLA2, respectively.

From (19), we have,

$$\frac{\partial F'(\mathbf{P}_j)}{\partial p_{ij}} \approx -\frac{c}{p_{ij}(p_{1j} \ldots p_{Mj})}, i = 1, \ldots, M \quad (20)$$

Apply (20) into (17), we get

$$p_{1j}{}^* \approx \ldots \approx p_{Mj}{}^*, j = 1, \ldots, L \quad (21)$$

which means, in high P_{av}, for each quantization region, the power allocated to each channel is asymptotically equal.

With this result, the above $M \times L$ equations system (18) can be simplified into an L equations system only:

$$\begin{cases} \sum_{j=1}^{L-1} p_j(F'(p_{j+1}) - F'(p_j)) + p_L(1 - F'(p_L) + F'(p_1)) \\ = P_{av} \\ \frac{\partial F'(p_j)}{\partial p_j} = -\frac{F'(p_{j+1}) - F'(p_j)}{(p_{j-1} - p_j)}, \quad j = 2, \ldots, L-1 \\ \frac{\partial F'(p_L)}{\partial p_L} = -\frac{1 - F'(p_L) + F'(p_1)}{(p_{L-1} - p_L)} \end{cases} \quad (22)$$

The $HP_{av}A$ reduces the M-dimensional vector quantization problem into a one-dimensional scalar quantization problem, with corresponding quantization thresholds $r_1, \ldots, r_L$, where

765

$r_j = r_{1j} = \ldots = r_{Mj} = K/p_j$, and remarkably reduces the complexity. We call this suboptimal solution as 'Equal Power Per Channel (EPPC)'. These equations can be easily solved for small values of L by 1stOpt.

For large values of L, one can use the equal average power per region (EPPR) approximation for such a scalar quantization problem, as also used in [7], but first derived by [5] using the *Mean Value Theorem*. This essentially implies that when L goes to infinity, the total average power assigned to each quantization region is asymptotically equal and the performance using this approximation is close to optimum for increasing bits of feedback. In this case, we have

$$p_j(F'(p_{j+1}) - F'(p_j)) = \frac{P_{av}}{L}, j = 1, \ldots, L-1$$

$$p_L(1 - F'(p_L) + F'(p_1)) = \frac{P_{av}}{L} \qquad (23)$$

and above equations can be solved by using the algorithm below (called 'EPPC+EPPR').

The Algorithm EPPC+EPPR:

Initialize $k = 0$, $p_{L_{low}}^{(0)} = \frac{P_{av}}{L}$, $p_{L_{high}}^{(0)} = const$, where the *const* is arbitrary large number.
repeat

   1) $p_{L_{new}}^{(k)} = \frac{p_{L_{low}}^{(k)} + p_{L_{high}}^{(k)}}{2}$;
   2) Fix $p_{L_{new}}^{(k)}$, find $p_{L-1}^{(k)}, \ldots, p_1^{(k)}$ by successively solving $p_j(F'(p_{j+1}) - F'(p_j)) = \frac{P_{av}}{L}, j = L-1, \ldots, 1$ (using similar bisection method), and calculate $P^k = p_{L_{new}}^{(k)}(1 - F'(p_{L_{new}}^{(k)}) + F'(p_1^{(k)}))$;
   3) Check if $P^k > \frac{P_{av}}{L}$, $p_{L_{high}}^{(k+1)} = p_{L_{new}}^{(k)}$, otherwise $p_{L_{low}}^{(k+1)} = p_{L_{new}}^{(k)}$;
   4) $k \leftarrow k + 1$;

until Convergence

Another effective choice for large L is to use the 'Zero Power in Outage Region'(ZPiOR) approximation, which uses power level $\mathbf{P}_L = 0$ with $r_L \to \infty$. Since, when L is large, $r_1 \to 0$, the average power allocated (via the optimal solution) to the outage region $\mathcal{R}_{L,1}$ is negligible. Thus the only difference is that the ZPiOR algorithm uses one less region, the performance loss due to which becomes also negligible as L increases. Thus the ZPiOR approximation (EPPC+ ZPiOR) is asymptotically (as $L \to \infty$) close to the optimal EPPC.

In this case, (22) is simplified as

$$\begin{cases} \sum_{j=1}^{L-1} p_j(F'(p_{j+1}) - F'(p_j)) = P_{av} \\ p_{j-1} = p_j + \frac{F'(p_{j+1}) - F'(p_j)}{-\frac{\partial F'(p_j)}{\partial p_j}}, \ j = 2, \ldots, L-1, F'(p_L) = 1 \end{cases}$$

$$(24)$$

which can be easily solved by using a standard bisection method. In fact, the ZPiOR approximation has a near-optimum performance for large number of regions. Thus, the ZPiOR approximation achieves a better complexity-performance tradeoff than EPPC+EPPR.

However, clearly for low average power, EPPC is not

optimal. Next, we derive a suboptimal scheme for low average power.

D. Low Average Power Approximations ($LP_{av}A$)

Let P_{av}^j present the average power allocated to region $\mathcal{R}_j$. Then the original problem (16) can be reformulated as

$$\min_{\{\mathbf{P}_1, \ldots, \mathbf{P}_L\}} F'(\mathbf{P}_1)$$

$$s.t. \ (\sum_{i=1}^{M} p_{ij})(F'(\mathbf{P}_{j+1}) - F'(\mathbf{P}_j)) = P_{av}^j$$

$$j = 1, \ldots, L-1$$

$$(\sum_{i=1}^{M} p_{iL})(1 - F'(\mathbf{P}_L) + F'(\mathbf{P}_1)) = P_{av}^L$$

$$P_{av}^1 + \ldots + P_{av}^L = M P_{av} \qquad (25)$$

From the average power constraint of (25), we have,

$$\sum_{j=1}^{L} \frac{P_{av}^j}{(p_{1j} + \ldots + p_{Mj})} = 1 \qquad (26)$$

Since $\mathbf{P}_1 > \mathbf{P}_2 > \ldots > \mathbf{P}_L$, we have,

$$\frac{1}{M}(p_{1L} + \ldots + p_{ML}) \leq P_{av} \qquad (27)$$

Thus, when the average power is small ($P_{av} \to 0$), $p_{iL} \to 0, i = 1, \ldots, M$ as well, and the corresponding quantization threshold $r_{iL} \to \infty$. In this case, the region $\mathcal{R}_L$ only includes $\mathcal{R}_{L,1}$ (the outage region) and the corresponding power level $\mathbf{P}_L = 0$, which is exactly the ZPiOR model studied in the previous subsection.

When L is large, we also can employ the EPPR approach for the vector quantization case, and similar to [6], we can show that by using EPPR and above ZPiOR approximation, the optimal power levels can be obtained in a successive manner instead of being jointly solved from (18), which dramatically reduces the complexity and makes the algorithm applicable to larger number of feedback bits.

Therefore, in low P_{av}, problem (25) can be expressed as

$$\min_{\{\mathbf{P}_1, \ldots, \mathbf{P}_{L-1}\}} F'(\mathbf{P}_1)$$

$$s.t. \ (\sum_{i=1}^{M} p_{ij})(F'(\mathbf{P}_{j+1}) - F'(\mathbf{P}_j)) = \frac{M P_{av}}{L-1},$$

$$j = 1, \ldots, L-1, \ F'(\mathbf{P}_L) = 1 \qquad (28)$$

And the $M \times L$ equations system (18) can be simplified to separate M-equation systems for each region $j = 1, \ldots, L-1$,

$$\begin{cases} (p_{1j}^* + \ldots + P_{Mj}^*)(F'(\mathbf{P}_{j+1}^*) - F'(\mathbf{P}_j^*)) = \frac{M P_{av}}{L-1} \\ \frac{\partial F'(\mathbf{P}_j)}{\partial p_{1j}^*} = \ldots = \frac{\partial F'(\mathbf{P}_j)}{\partial p_{Mj}^*} \end{cases}$$

$$(29)$$

The above system of equations implies that solving the power level $\mathbf{P}_j^* = \{p_{ij}^*\}_{i=1}^{M}$ needs knowledge of $F'(\mathbf{P}_{j+1}^*)$. Since, for region $\mathcal{R}_{L-1}$, $F'(\mathbf{P}_L^*) = 1$, the solution of (29) with $(j = L-1)$ determines the optimal power level $\mathbf{P}_{L-1}^*$ and $F'(\mathbf{P}_{L-1}^*)$. Once region $\mathcal{R}_{L-1}$ has been accounted for,

the optimal power level for $\mathcal{R}_{L-2}$ can be identified by substituting the value of $F'(\mathbf{P}_{L-1}^*)$ in (29) with $(j = L-2)$. By recursively repeating the same progress, all the optimal power levels $\{\mathbf{P}_j^*\}_{j=1}^{L-1}$ can be achieved.

In fact, problem (28) can be turned into a separable optimization problem:

Lemma 2: The following optimization problem

$$\min_{\{p_{1j},\dots,p_{Mj}\}} F'(\mathbf{P}_j)$$

$$s.t. \quad (\sum_{i=1}^M p_{ij})(F'(\mathbf{P}_{j+1}) - F'(\mathbf{P}_j)) = \frac{MP_{av}}{L-1},$$
$$j = 1,\dots,L-1 \qquad (30)$$

shares the same KKT necessary conditions and hence the same local optima with problem (28).

Problem (30) can be successively solved by using optimization tools 1stOpt instead of recursively solving systems of M nonlinear equations.

E. Asymptotic Behavior of Outage Probability

Here we briefly comment on the diversity behaviour of the outage minimization algorithm using the optimal straight line approximations (OSLA). The proof is again excluded due to space limitations. Define the diversity gain d as

$$d = -\lim_{P_{av}\to\infty} \frac{\log P_{out}}{\log P_{av}} \qquad (31)$$

Theorem 2: For an arbitrary number of M channels with $\log_2 L$ bits of quantized feedback, using the optimal power allocation employing the OSLA approximations, we have

$$d \approx \sum_{j=1}^L M^j \qquad (32)$$

This result is consistent with similar results in [4] and [5].

IV. LARGE NUMBER OF CHANNELS ANALYSIS

The previous algorithms can be effectively applied to find locally optimal solutions or their approximations for moderate number of parallel channels, such as $M < 10$. Once $M \geq 10$, these algorithms become computationally demanding. Given that practical multi-carrier systems such as OFDM can have $M = 64$ or $M = 128$ sub-carriers, one needs to find outage minimizing power allocation algorithms for large M. Below we provide such an algorithm using a Gaussian approximation for large M in high P_{av}.

Note that in high average power and for large M,

$$\sum_{i=1}^M \log(1 + p_{ij}h_i) \approx \sum_{i=1}^M \log(p_{ij}h_i)$$
$$= \sum_{i=1}^M \log(\frac{p_{ij}}{\lambda_i}) + \sum_{i=1}^M \log(f_i) \quad (33)$$

where under the assumption of Rayleigh fading, the pdf of $f_i = h_i\lambda_i$ is $e^{-f_i}, \forall i$. Recall that $\frac{1}{\lambda_i}$ is the mean of channel gain h_i.

Thus $F(\mathbf{P}_j), j = 1,\dots,L$ for M channels can be expressed as

$$F(\mathbf{P}_j) = Prob(\frac{1}{M}\sum_{i=1}^M \frac{1}{2}\log(1 + p_{ij}h_i) < r_0)$$

$$\approx Prob(\frac{1}{M}\sum_{i=1}^M \log(f_i) < s_j) = V(s_j) \quad (34)$$

where $s_j = 2r_0 - \frac{1}{M}\sum_{i=1}^M \log(\frac{p_{ij}}{\lambda_i}) = c' - \frac{1}{M}\sum_{i=1}^M \log(p_{ij})$, $c' = 2r_0 + \frac{1}{M}\sum_{i=1}^M \log(\lambda_i)$ and the function $V(.)$ denotes the cumulative distribution function (cdf) of $\frac{1}{M}\sum_{i=1}^M \log(f_i)$. It is easy to calculate that the pdf of $z_i = \log(f_i)$ is $e^{-e^{z_i}}e^{z_i}$, which is the well known 'Gumbel Distribution'. It's mean $m = -r$, where r is Euler-Mascheroni constant $(r = 0.5772156649...)$ and variance is $\frac{\pi^2}{6}$.

Since z_i is i.i.d with finite mean and variance, when the number of channels $M \to \infty$, using the *Central Limit Theorem*, the probability distribution of $\frac{1}{M}\sum_{i=1}^M z_i$ can be approximated by a Gaussian distribution with mean and variance given by

$$\mu = E[\frac{1}{M}\sum_{i=1}^M z_i] = \frac{1}{M}\sum_{i=1}^M E[z_i] = E[z_i] = m = -r$$

$$\sigma^2 = Var[\frac{1}{M}\sum_{i=1}^M z_i] = \frac{1}{M^2}Var[\sum_{i=1}^M z_i]$$
$$= \frac{1}{M}Var[z_i] = \frac{1}{M}\frac{\pi^2}{6} \qquad (35)$$

Thus, we have

$$V(s_j) \approx \int_{-\infty}^{s_j} \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx = \frac{1}{2}[1 + erf(\frac{s_j - \mu}{\sigma\sqrt{2}})]$$

The original problem (10) for a large number of channels case can now be restated as

$$\min_{\{\mathbf{P}_1,\dots,\mathbf{P}_L\}} V(s_1)$$

$$s.t. \sum_{j=1}^{L-1}(\sum_{i=1}^M p_{ij})(V(s_{j+1}) - V(s_j))$$
$$+ (\sum_{i=1}^M p_{iL})(1 - V(s_L) + V(s_1)) = MP_{av} \quad (36)$$

Using the KKT necessary conditions, we again get

$$\frac{\partial V(s_j)}{\partial p_{1j}} = \dots = \frac{\partial V(s_j)}{\partial p_{Mj}}, j = 1,\dots,L \qquad (37)$$

Since

$$\frac{\partial V(s_j)}{\partial p_{ij}} = f(s_j)\frac{\partial s_j}{\partial p_{ij}} = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(c'-\frac{1}{M}\sum_{i=1}^M \log(p_{ij})-\mu)^2}{2\sigma^2}}(-\frac{1}{M}\frac{1}{p_{ij}})$$

where $f(s) = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(s-\mu)^2}{2\sigma^2}}$. Similarly, we also have

$$p_{1j} = \dots = p_{Mj}, j = 1,\dots,L \qquad (38)$$

Denote $p_j = p_{ij}, j = 1,\dots,L$. The we have $s_j = c' - \log(p_j)$ and the vector quantization problem (36) can be

converted into the scalar quantization problem below with quantization thresholds $s_1, \ldots, s_L$:

$$\min_{\{p_1,\ldots,p_L\}} V(s_1)$$
$$s.t. \sum_{j=1}^{L-1} p_j(V(s_{j+1}) - V(s_j))$$
$$+ p_L(1 - V(s_L) + V(s_1)) = P_{av} \qquad (39)$$

After employing the corresponding KKT necessary optimality conditions and simplifying, we have the system of L nonlinear equations below:

$$\begin{cases} \sum_{j=1}^{L-1} p_j(V(s_{j+1}) - V(s_j)) + p_L(1 - V(s_L) + V(s_1)) \\ = P_{av} \\ p_{j-1} = p_j(\frac{V(s_{j+1}) - V(s_j)}{f(s_j)} + 1), \quad j = 2, \ldots, L-1 \\ p_{L-1} = p_L(\frac{1 - V(s_L) + V(s_1)}{f(s_L)} + 1) \end{cases}$$
$$(40)$$

When L is not large, one can solve the above equations using 1stOpt software. When L is large (roughly $L > 20$), we can also use the EPPR approximation or the ZPiOR approximation to solve them, as discussed in the section on high P_{av} approximations. Table I below shows the applicability of various algorithms discussed so far according to different ranges of M, L and high or low P_{av}.

TABLE I
PROPOSED POWER ALLOCATION STRATEGIES

Channels Number	$M < 10$	$M \geq 10$
Initial approach	OSLA	-
High P_{av}	EPPC ($L > 20$, EPPC+ZPiOR or EPPC+EPPR)	Gauss. Approx. (GA) ($L > 20$, GA+ZPiOR or GA+EPPR)
Low P_{av}	ZPiOR+EPPR	-

V. NUMERICAL RESULTS

To numerically illustrate the performance of the designed power allocation strategies, we consider an M-parallel (independent) Rayleigh block-fading channels where the mean value of the exponentially distributed fading gain for each channel is assumed to be inversely proportional to the square of the wireless propagation distance d, and the required transmission rate is taken to be $r_0 = 0.25$ nats per channel use. All the results with OSLA are obtained with the approximation OSLA2 as this outperforms OSLA1 (details excluded due to space limitations). Outage performance with full CSI at the transmitter is obtained with the optimal power allocation results presented in [1]. It should be noted that the results illustrate the "real outage" performance of the proposed algorithms (the power codebook designed via the algorithms is used to obtain the average outage probability over a large number of Monte-Carlo simulated channel realizations). As a result, the average power required for a given *real* outage may not be the same as the original average power based on which the power codebook is designed. However, for a given algorithm, the graphs should be used to determine the minimum outage probability obtainable for a given average power and vice versa.
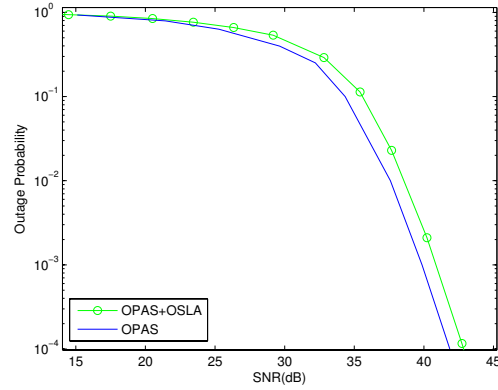

Fig. 3. Outage performance comparison between OPAS and OPAS+OSLA with 2 channels 1 bits feedback.
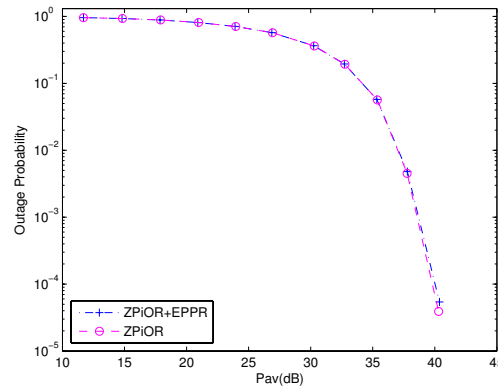

Fig. 4. Performance comparison between ZPiOR + EPPR and ZPiOR schemes (M=2,B=2)

Experiment 1 : The first experiment examines the efficiency of OSLA. Fig. 3 compares the outage performance of the optimal power allocation scheme (OPAS) without any approximation, obtained by an exhaustive search over the space of all possible power allocation policies, with OPAS+OSLA. OPAS+OSLA refers to the OSLA algorithm without any further additional approximations. For all the simulation results to follow, this OPAS+OSLA will be called the "optimal" scheme for comparison purposes. Fig. 3 shows outage results for 2 channels ($d_1 = 40m, d_2 = 60m$) and 1 bit feedback case. It can be observed that for low average power, the performance of OPAS (exhaustive search) and OPAS+OSLA are very close and with increasing average power, OPAS slightly outperforms OPAS+OSLA (less than 0.95 dB power consumption gap at an outage probability of 10^{-2}), but OPAS+OSLA is much less complexity than OPAS. The negligible difference in performance at low P_{av} attributes to the fact that when $P_{av} \to 0$, $\sum_{i=1}^{M} \log(1 + p_{ij}h_i) \approx \sum_{i=1}^{M} p_{ij}h_i$ (OSLA). This clearly demonstrates that OSLA is an efficient approximation technique.

Experiment 2 : This experiment tests the performance of two suboptimal schemes HP_{av}A (EPPC) and LP_{av}A

768
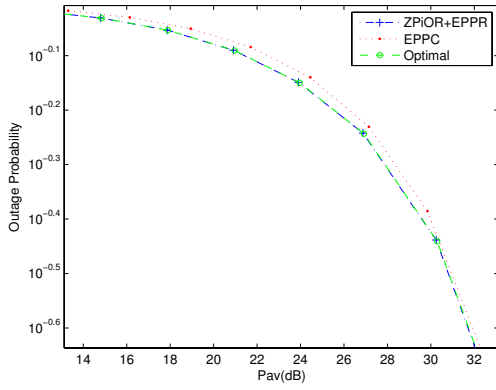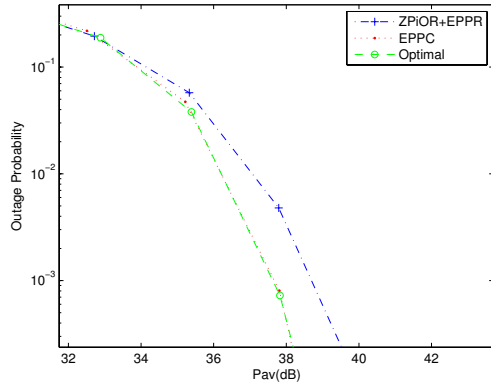
(a) Low $P_{av}(P_{av} < 32.5dB)$

(b) High $P_{av}(P_{av} > 32.5dB)$

Fig. 5. Outage performance of LPavA (ZPiOR+EPPR) and HPavA (EPPC) schemes (M=2,B=2).

(ZPiOR+EPPR). As seen in Fig. 4, ZPiOR+EPPR and ZPiOR achieve almost equivalent outage performances (a similar observation was also made in [5] for the scalar quantization case), which validates the suitability of extending the EPPR approach to our vector quantization scenario. The striking observation in Fig. 5 is that when $P_{av} < 32.5dB$, the performance of ZPiOR+EPPR and the optimal scheme (OPAS+OSLA) are not distinguishable and they outperform EPPC, while when $P_{av} > 32.5dB$, EPPC performs better than ZPiOR+EPPR and is very close to optimal, indicating that ZPiOR+EPPR is a near-optimal solution in low P_{av} (a similar observation is also made by [4] that ZPiOR works relatively well when P_{av} is low) whereas EPPC is an efficient suboptimal scheme for high P_{av}. Thus $\min(P_{out}(ZPiOR+EPPR),P_{out}(EPPC))$ gives near-optimal performance for nearly all values of P_{av}.

Experiment 3 : The third simulation, as illustrated in Fig. 6 for $M = 6$ channels case (**d**=[20m,30m,40m,60m,70m,80m]), studies the effect of increasing the number of feedback bits on the outage performance using the proposed schemes. For comparison, the performance of the optimal power control policy with full CSI [1] is also shown. The results for 1 bit

feedback are obtained using OPAS+OSLA, while for 2 bits and 4 bits of feedback we use both ZPiOR+EPPR and EPPC. The important observation from these two figures is that the introducing one extra bit of feedback substantially reduces the gap with the full CSI performance and only a few bits of feedback can eliminate most of the gap with the full CSI performance. For example, at an outage probability of 10^{-2}, with 4bits feedback $(\min(P_{out}(ZPiOR+EPPR),P_{out}(EPPC)))$, there is only approximately a 2 dB power loss compared to the full CSI case. This confirms that power allocation with limited feedback (only with a few feedback bits) can provide a dramatic performance advantage over no CSI (channel non-adaptive power allocation across all channels).

Additionally, as we can see, EPPC outperforms ZPiOR+EPPR in high P_{av}. Interestingly, we see that this improvement is reduced as B increases : at an outage probability of $10^{-4.4}$ in Fig. 6 , EPPC with 2 and 4 bits feedback provide roughly 0.62 dB and 0.05 dB gains over ZPiOR+EPPR respectively. This is due to the fact that in high P_{av}, as B increases, the performance of ZPiOR+EPPR gradually becomes close to EPPC. In this sense, for a large number of feedback bits, ZPiOR+EPPR can be treated as a suboptimal solution for all ranges of P_{av}. Fig. 7 shows that ZPiOR can also be combined with EPPC as a computationally simpler alternative to EPPC+EPPR for large number of feedback bits.
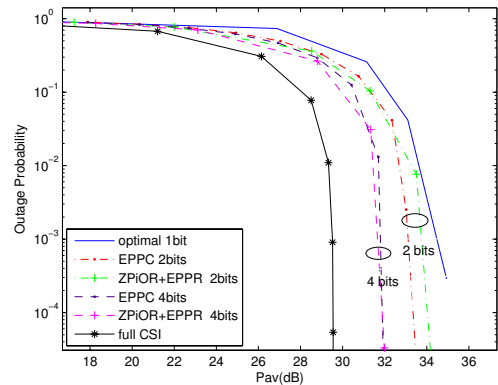

Fig. 6. Effect of more feedback bits on outage performance for 6 channels.

Experiment 4: Fig. 8 compares the outage performance between the EPPC scheme, (implementing the OSLA2 approximation), and the Gaussian approximation (GA). It can be seen very clearly that the EPPC scheme outperforms GA, the benefit of the EPPC scheme becoming more pronounced as P_{av} increases. For instance, with the same B, at an outage probability of 10^{-2}, EPPC with 1 bit requires a power of roughly 2.6 dB less than GA does; and EPPC with 2 and 4 bits feedback provide around 2.2 dB and 1.1 dB power savings over GA respectively. Even only 1 bit(2bits) EPPC can achieve close performance to 2 bits (4 bits) GA in high P_{av}. These results indicate that the OSLA approximation can achieve remarkable performance advantage over GA, espe-
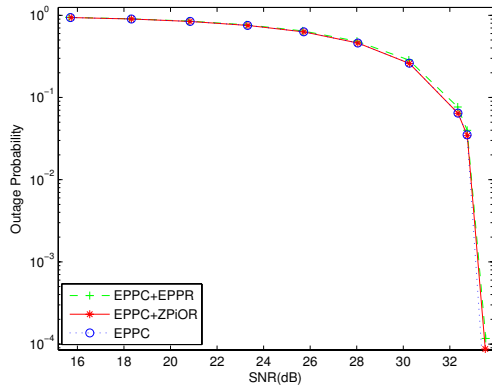
769

Fig. 7. Performance comparison between EPPC with EPPR approximation, EPPC with ZPiOR approximation and optimal EPPC (M=4,B=4)

cially in high P_{av}. However, when the number of channels is large, OSLA becomes computationally prohibitive. And in this case, GA is an efficient alternative, which is consistent with similar observations (for MIMO settings) in [9].

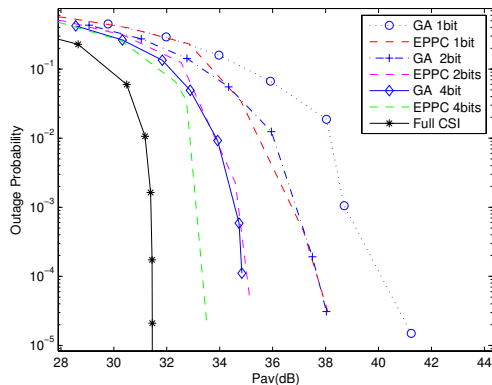Fig. 9 illustrates the outage probability for the case of a

Fig. 8. Outage performance comparison between OSLA and GA (M=4, B=1,2,4 bits).

large M ($M = 16$) using GA, with the values of distances $d_1, \ldots, d_{16}$ randomly obtained from the range [20m,100m]. From Fig. 9, we have a similar observation as Experiment 3 that only a few bits of feedback are required to attain significant savings in power. For instance, to achieve a target expected outage probability 10^{-2}, the average SNR gap between 10 bits feedback and full CSI perfect performance is only about 2.5dB.

REFERENCES

[1] G. Caire, G. Taricco, and E. Biglieri, "Optimum power control over fading channels," *IEEE Trans. Inform. Theory*, vol. 45, no. 5, pp. 1468–1489, July 1999.
[2] A. J. Goldsmith and P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Transactions on Information Theory*, vol. 43, pp. 1986–1992, November 1997.
[3] S. V. Hanly and D. N. C. Tse, "Multi-access fading channels - part II: Delay-limited capacities," *IEEE Transactions on Information Theory*, vol. 44, pp. 2816–2831, November 1998.
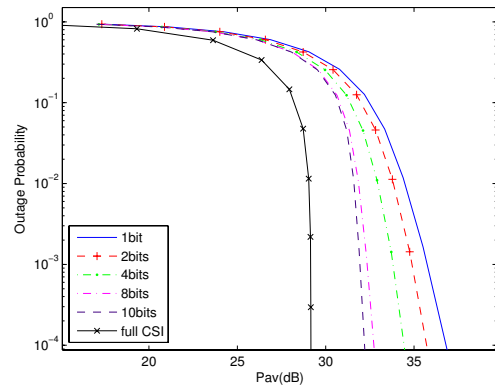
Fig. 9. The outage performance of $M = 16$ channels with GA.

[4] T. T. Kim and M. Skoglund, "DiversityMultiplexing Tradeoff in MIMO Channels With Partial CSIT," *IEEE Transactions on Information Theory*, vol. 53, no. 8, pp. 2743–2759, August 2007.
[5] A. Khoshnevis and A. Sabharwal, "On the asymptotic performance of multiple antenna channels with quantized feedback," *IEEE Transactions on Wireless Communications*, vol. 7, no. 10, pp. 3869–3877, October 2008.
[6] N. Ahmed, M. A. Khojastepour, A. Sabharwal, and B. Aazhang, "Outage minimization with limited feedback for the fading relay channel," *IEEE Trans. Commun.*, vol. 54, no. 4, pp. 659–669, April 2006.
[7] M. A. Khojastepour, G. Yue, X. Wang and M. Madihian, "Optimal power control in MIMO systems with quantized feedback," *IEEE Transactions on Wireless Communications*, vol. 7, no. 12, pp. 4859–4866, Dec. 2008.
[8] J. H. Luo, R. Yates, and P. Spasojević, "Service outage based power and rate allocation for parallel fading channels," *IEEE Trans. Inform. Theory*, vol. 51, no. 7, pp. 2594–2611, July 2005.
[9] B.M. Hochwald, T.L. Marzetta and V. Tarokh, "Multiple-antenna channel hardening and its implications for rate feedback and scheduling," *IEEE Transactions on Information Theory*, vol. 50, no. 9, pp. 1893–1909, Sep. 2004.
[10] http://sites.google.com/site/zhyfutopia/Home/Proofs.pdf?attredirects=0