

# Discrete Adaptive Transmission for Fading Channels

Lang Lin, Roy D. Yates, and Predrag Spasojević\*  
WINLAB, Rutgers University  
73 Brett Rd., NJ 08854-8060  
{llin, ryates, spasojev}@winlab.rutgers.edu

July 14, 2001

## Abstract

Throughput maximization of an adaptive transmission system with a finite number of transmitted power levels and code rates for communication over slow fading channels is analyzed based on the concept of information outage. Properties of throughput maximizing policies lead to an iterative algorithm which suggests good system designs. Numerical results show that carefully designed discrete adaptive transmission systems with a small number of power levels and code rates can achieve throughput values close to ergodic capacity.

## 1 Introduction

In third generation cellular systems (see [1–3] and [4]), adaptation on the transmitter side is one of the core technologies leading to power efficient designs of wireless data communication systems. The seminal work [5] shows that the ergodic capacity of a fading channel can be achieved by employing an adaptive transmission system with variable transmitted power and variable code rate assignment.

An optimal adaptive transmission system [5] requires knowledge of the current channel state at the transmitter. In addition, both transmitted power and code rate assignments must adapt continuously to changes in the channel state. Both of these requirements have been widely adopted in further work on information theoretic aspects of communication over fading channels [6, 7]. Unfortunately, these requirements are hard to satisfy in practice. In order to simplify an adaptive transmission system, an adaptive  $M$ -ary quadrature amplitude modulation (MQAM) with a finite number of modulation levels and an adaptive trellis-coded modulation (TCM) scheme with a finite number of code rates are proposed in [8] and [9], respectively. More recently, an adaptive transmission design based on outdated channel information and either MQAM or TCM is proposed in [10].

---

\*corresponding author

Since the performance improvement offered by an adaptive approach is promising, this paper examines throughput maximization of an adaptive system with a finite number of power levels and code rates. In this problem, the challenge is twofold. First, the throughput maximization problem for a discrete adaptive system is hard to formulate, especially when channel coding is involved. Second, the problem requires joint optimization of a number of parameters.

We assume a slow multiplicative fading environment where the channel is constant during the transmission of a codeword and the channel state space is partitioned into a finite number of quantization intervals. It is assumed that, instead of the exact channel state, the transmitter learns the quantization interval of the current channel state. From this information, the transmitter selects a corresponding pair from a finite set of code rate and power level pairs to encode and transmit the information message, respectively. For convenience, we only consider the case where the number of quantization levels, code rates, and power levels are equal even though it is possible for these to differ in a specific system design.

For a multiplicative fading channel with AWGN, the coding theorem in [5] states that an adaptive transmission system can achieve ergodic capacity by using the water-filling power assignment and channel codes which achieve the maximum mutual information at each channel state. Hence, we employ the same class of codes at each quantization level as [5] and for these codes, there are a number of related coding theorem references [11–13].

For the proposed discrete system, it is possible that the instantaneous mutual information corresponding to a channel state is less than the assigned code rate. In this case, an information outage event occurs. The information outage is an intrinsic characteristic of slow fading channels and due to this, the strict sense Shannon capacity is zero in slow fading channels [14]. During an outage, a transmission is not considered reliable and, thus, it is frequently convenient to assume that the transmitted data can be ignored [11]. This assumption leads to the *capacity versus outage* problem, which focuses on the tradeoff between the outage probability and the supportable rate [7, 15]. The practice of ignoring data received during an outage is supported by the fact that the outage probability matches well the error probability of actual codes [16] and [17]. Therefore, we characterize the performance of a system design by the *average reliable throughput*, defined as the average of the data rate assuming zero rate when the channel is in outage. Henceforth, our central topic is finding the maximum average reliable throughput, also termed capacity, of a discrete adaptive system communicating over a slow fading channel.

Following the formulation of the finite-level design problem, we focus on exploring the optimal (capacity-achieving) policies where a *policy* is defined as an ensemble of the channel state space partition and corresponding power and rate allocations. In this work, we will show that for an optimal policy, outages occur only for a set of worst channel states within the first quantization interval. In other quantization intervals, we will see that the assigned rates will support the worst channel state of that interval. We will find that an optimal policy can be uniquely characterized by the channel state space partition, first interval rate assignment, and the power level assignments. The optimal power level assignment has a water-filling character. When the number of levels approaches infinity, reasonably designed discrete adaptive transmission schemes with a water-filling power assignment and an equal probability partition of the channel state space can achieve the ergodic capacity.

We emphasize that finding an optimal policy is still a challenging non-convex optimiza-

tion problem. Since brute force searching over a space of all possible policies at a high resolution has complexity that increases exponentially with the number of quantization levels, we present an iterative algorithm that numerically evaluates capacity lower bounds and determines corresponding suboptimal policies. The computation complexity of the algorithm is linear in the number of levels and the achieved lower bounds can be very close to the true capacity.

## 2 System Model and Problem Formulation

We consider a multiplicative flat fading channel model similar to that in [5]. The complex received signal

$$Y = \sqrt{S}X + W, \quad (1)$$

where  $S$  is the channel (fading) state,  $X$  is the complex transmitted signal, and  $W$  is a circularly symmetric additive white Gaussian noise (AWGN) with variance  $N_0$ . The channel state  $S$  is a real random variable of unit mean with a probability density function (PDF)  $f(s)$ , a cumulative density function (CDF)  $F(s)$ , and a domain  $\mathcal{S} = \{s|s \geq 0\}$ . It is also assumed that the fading is sufficiently *slow* that the channel state is constant during the transmission of a codeword.

The proposed adaptive transmission system quantizes any channel state  $s$  to one of  $L$  levels  $v_0 < v_1 < \dots < v_{L-1}$ , where  $v_0 = 0$ . The  $L$  channel state quantization intervals are denoted by  $\mathcal{V}_l = [v_l, v_{l+1})$  for  $l = 0, \dots, L-1$ , where  $v_L = \infty$ . Note that the set  $\{\mathcal{V}_0, \dots, \mathcal{V}_{L-1}\}$  partitions  $\mathcal{S}$ . When the channel state  $s \in \mathcal{V}_l$ , the encoder at the transmitter generates codewords of a code rate  $r_l$  and the codewords are transmitted at a power level  $p_l = E\{|X|^2 | s \in \mathcal{V}_l\}$ , where  $E\{\cdot\}$  denotes expectation. Since (1) is an AWGN channel for any given  $s \in \mathcal{S}$ , the corresponding maximum mutual information is given by  $\log(1 + p_l s / N_0)$ . Adopting the notation  $R(\phi) = \log(1 + \phi / N_0)$ , the maximum mutual information associated with any state  $s \in \mathcal{V}_l$  is  $R(p_l s)$ .

The set of quantization levels  $\{v_l\}$  and the corresponding set of power and rate assignment pairs  $\{(p_l, r_l)\}$  define an adaptive transmission policy. More precisely, the triple of  $L$  by 1 vectors

$$(\mathbf{p}, \mathbf{v}, \mathbf{r}) = ([p_0, \dots, p_{L-1}]^\top, [v_0, \dots, v_{L-1}]^\top, [r_0, \dots, r_{L-1}]^\top) \quad (2)$$

specifies a *transmission policy*. For any  $s \in \mathcal{V}_l$ , given a power level and code rate assignment pair  $(p_l, r_l)$ , it is not guaranteed that the information will be successfully received if  $R(p_l s) < r_l$ . Hence, following the established outage probability definition [15], we define the conditional outage probability as

$$P_{\text{out}}(r_l, p_l | l) = \Pr [R(p_l s) < r_l | s \in \mathcal{V}_l]. \quad (3)$$

In other words,  $P_{\text{out}}(r_l, p_l | l)$  is the probability that the assigned rate  $r_l$  is not achievable with the power level  $p_l$  given that  $s \in \mathcal{V}_l$ .

Given a policy  $(\mathbf{p}, \mathbf{v}, \mathbf{r})$ , the *average reliable throughput* is

$$R_L(\mathbf{p}, \mathbf{v}, \mathbf{r}) = \sum_{l=0}^{L-1} \Pr [s \in \mathcal{V}_l] [1 - P_{\text{out}}(r_l, p_l | l)] r_l. \quad (4)$$

Note that this definition hinges on the assumption that no information is successfully received during an outage. Here, we adopt the following convention:  $p_l = 0$  implies that  $r_l = 0$  and, consequently, there is no decoding error and no outage when no transmission is attempted. Employing the notation  $F(s_1, s_2) = F(s_2) - F(s_1) = \Pr [s_1 \leq S < s_2]$ , the *average power* for the policy  $(\mathbf{p}, \mathbf{v}, \mathbf{r})$  is

$$\rho(\mathbf{p}, \mathbf{v}, \mathbf{r}) = \sum_{l=0}^{L-1} F(v_l, v_{l+1})p_l. \quad (5)$$

If the average power is constrained to be upper bounded by  $\bar{p}$ , the *set of feasible  $L$ -level transmission policies* is

$$\pi_L(\bar{p}) = \{(\mathbf{p}, \mathbf{v}, \mathbf{r}) \mid \rho(\mathbf{p}, \mathbf{v}, \mathbf{r}) \leq \bar{p}\}. \quad (6)$$

We define the maximum average reliable throughput over all  $L$ -level policies to be

$$C_L = \max_{(\mathbf{p}, \mathbf{v}, \mathbf{r}) \in \pi_L(\bar{p})} R_L(\mathbf{p}, \mathbf{v}, \mathbf{r}). \quad (7)$$

A corresponding capacity achieving policy is referred to as an *optimal policy*  $(\mathbf{p}^*, \mathbf{v}^*, \mathbf{r}^*)$ . The theme of this paper is finding the capacity  $C_L$  and the properties of the corresponding optimal policy  $(\mathbf{p}^*, \mathbf{v}^*, \mathbf{r}^*)$  given a fading distribution.

### 3 Properties of Optimal Policies

In this section, we present sketches of power allocations as a function of the channel state and illustrate a number of useful properties of optimal policies. These properties are helpful in simplifying the optimization problem (7). One possible policy assignment is depicted in Fig. 1, where the heights of boxes indicate the power levels assigned to their respective quantization intervals. Within each interval  $\mathcal{V}_l$ , the shaded region indicates the *outage interval* where the channel state does not support reliable communication with the assigned rate  $r_l$  given implicitly through  $q_l$ , as is explained later in this section.

Any outage interval is contiguous within its respective quantization interval  $\mathcal{V}_l$  and would include the left end point  $v_l$  of  $\mathcal{V}_l$ . Specifically, if a channel state  $s_1 \in \mathcal{V}_l$  does not support reliable communication at rate  $r_l$ , no channel state  $s < s_1$  in  $\mathcal{V}_l$  will support reliable communication at rate  $r_l$ . Thus, there can be at most  $L$  outage intervals. In the following, we develop an implicit characterization of the assigned rates for an optimal policy.

**Lemma 1** *For an optimal policy  $(\mathbf{p}^*, \mathbf{v}^*, \mathbf{r}^*)$ , we have that*

$$r_l^* \in [R(p_l^* v_l^*), R(p_l^* v_{l+1}^*)]. \quad (8)$$

This proof, as well as proofs of other lemmas and theorems, can be found in Appendix A.

Without loss of generality, we assume that any policy of interest  $(\mathbf{p}, \mathbf{v}, \mathbf{r})$  also satisfies the condition (8). Consequently, for any policy  $(\mathbf{p}, \mathbf{v}, \mathbf{r})$  for which (8) holds and any quantization interval  $\mathcal{V}_l$ , strict monotonicity of  $R(\cdot)$  implies that we can find a unique channel state  $q_l \in \mathcal{V}_l$  such that  $r_l = R(p_l q_l)$ . This defines a one to one mapping between channel states

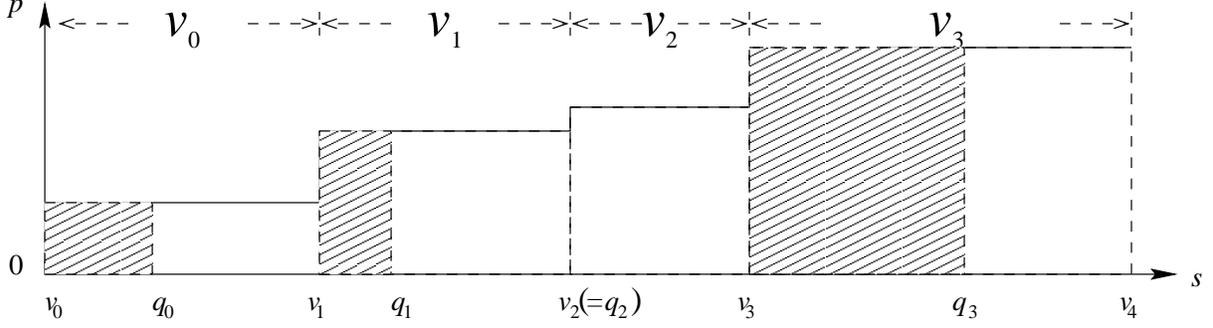


Figure 1: Illustration of an arbitrary power assignment. The rate assignment is implicitly shown through  $q_l$ . The shaded regions are the channel outage intervals.

$\{q_0, \dots, q_{L-1}\}$  and the respective rate assignments  $\{r_0, \dots, r_{L-1}\}$  for a given power policy. Hence,  $q_l$  is the worst channel state in  $\mathcal{V}_l$  that still allows for reliable communication at a rate  $r_l$ . Therefore, it will be more convenient to redefine transmission policies of interest as vector triples  $(\mathbf{p}, \mathbf{v}, \mathbf{q})$ . Accordingly, such a policy assignment can be illustrated by using a plot such as the one in Fig. 1, where all vector parameters are depicted.

With the introduction of  $\mathbf{q}$ , the conditional outage probability in (3) can be interpreted as the probability that  $s \in \mathcal{V}_l$  is worse than the worst reliable channel  $q_l$ , i.e.,

$$\begin{aligned} P_{\text{out}}(r_l, p_l | l) &= \Pr [R(p_l s) < R(p_l q_l) | s \in \mathcal{V}_l] \\ &= F(v_l, q_l) / F(v_l, v_{l+1}). \end{aligned} \quad (9)$$

With (4) and (10), the average rate can be expressed in terms of the vector  $\mathbf{q}$  as

$$R_L(\mathbf{p}, \mathbf{v}, \mathbf{q}) = \sum_{l=0}^{L-1} F(q_l, v_{l+1}) R(p_l q_l). \quad (11)$$

The next lemma, which follows from the monotonicity of  $R(p_l q_l)$  with respect to  $p_l$ , says that an optimal policy meets the average power constraint with equality.

**Lemma 2** *For an optimal policy  $(\mathbf{p}^*, \mathbf{v}^*, \mathbf{q}^*)$ , we have that*

$$\rho(\mathbf{p}^*, \mathbf{v}^*, \mathbf{q}^*) = \sum_{l=0}^L F(v_l^*, v_{l+1}^*) p_l^* = \bar{p}. \quad (12)$$

As apparent from (11) and Fig. 1, an outage interval does not contribute to the overall average rate. Thus, one could intuitively assume that an optimal policy should minimize such intervals. An extreme case is  $q_l = v_l$ , which implies that there is no outage interval.

**Theorem 1** *Given an arbitrary policy  $(\mathbf{p}, \mathbf{v}, \mathbf{q})$ , there exists a policy  $(\mathbf{p}', \mathbf{v}', \mathbf{q})$  such that  $v'_l = q_l$  for all  $l > 0$  and  $R_L(\mathbf{p}', \mathbf{v}', \mathbf{q}) \geq R_L(\mathbf{p}, \mathbf{v}, \mathbf{q})$ .*

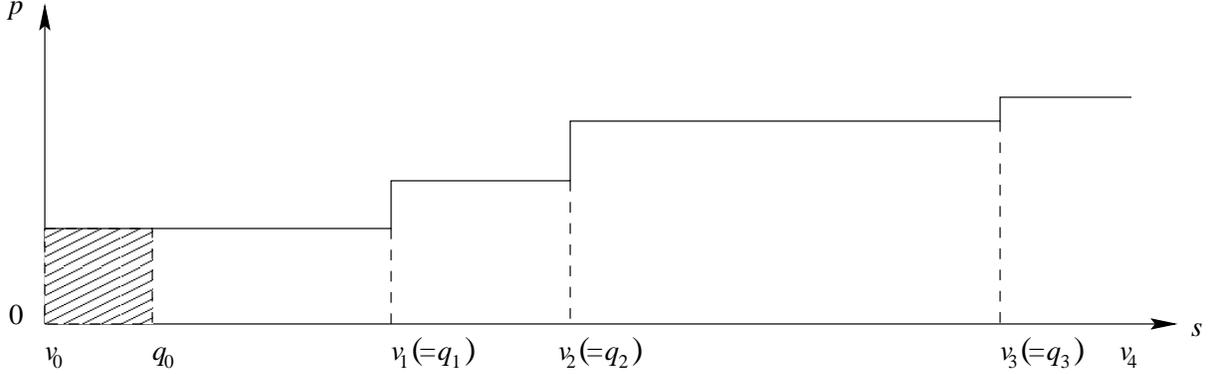


Figure 2: The policy achieves  $C_L$  will have  $q_1 = v_1$  and  $q_3 = v_3$ .

In the special case when a policy has an increasing power allocation depicted in Fig. 2, it is relatively simple to demonstrate Theorem 1. As noted earlier from (11) and Fig. 1, if  $v_{l+1} < q_{l+1}$ , then channel states  $s$  in the outage interval  $[v_{l+1}, q_{l+1})$  do not contribute to the average reliable throughput. Thus, we can increase  $v_{l+1}$  to eliminate the outage interval without changing the value of  $p_l$ . In this case, the average reliable throughput increases, while the average transmitted power remains unchanged. Since this procedure can be applied to any interval  $l > 0$ , for an optimal policy, an outage can occur only in the  $\mathcal{V}_0$  interval. In the general case of arbitrary power levels, Theorem 1 is due to the concavity and monotonicity of  $R(\cdot)$ ; the proof appears in Appendix A.

Theorem 1 demonstrates that policies of interest are fully determined by vectors  $\mathbf{p}$  and  $\mathbf{q}$  with the quantization intervals given by

$$\mathcal{V}_l = \begin{cases} [0, q_1) & l = 0 \\ [q_l, q_{l+1}) & 1 \leq l \leq L - 1 \end{cases} \quad (13)$$

where we assume the convention  $q_L = \infty$ .

In the following, we assume that the CDF  $F(s)$  is a strictly increasing function of  $s$  and, instead of optimizing the policy  $(\mathbf{p}, \mathbf{q})$ , we equivalently optimize a policy  $(\mathbf{p}, \mathbf{a})$ , where  $a_l = F(q_l)$  for all  $l$ . Using the shorthand  $q(a) = F^{-1}(a)$ , the average reliable throughput in terms of the pair  $(\mathbf{p}, \mathbf{a})$  has the form:

$$R_L(\mathbf{p}, \mathbf{a}) = \sum_{l=0}^{L-1} (a_{l+1} - a_l) R(p_l q(a_l)). \quad (14)$$

The average power constraint is

$$\rho(\mathbf{p}, \mathbf{a}) = a_0 p_0 + \sum_{l=0}^{L-1} (a_{l+1} - a_l) p_l \leq \bar{p}. \quad (15)$$

Among the set of  $L$ -level policies,

$$\pi_L(\bar{p}) = \{(\mathbf{p}, \mathbf{a}) | \rho(\mathbf{p}, \mathbf{a}) = \bar{p}\}, \quad (16)$$

a feasible policy from (16) is an optimal policy if it achieves the  $L$ -level capacity

$$C_L = \max_{(\mathbf{p}, \mathbf{a}) \in \pi_L(\bar{p})} R_L(\mathbf{p}, \mathbf{a}). \quad (17)$$

Hence, Theorem 1 now implies the following corollary.

**Corollary 1** *For an optimal policy  $(\mathbf{p}^*, \mathbf{a}^*)$ , either only  $\mathcal{V}_0^*$  includes an outage interval  $[0, q(a_0^*))$  or  $p_0^* = 0$ .*

Although the new optimization problem (17) is somewhat simpler than (7), it is still non-convex and difficult to solve in the general case. The following theorem provides a further characterization of the optimal policies. It shows that, given any adaptation partition now defined by  $\mathbf{q}$ , the optimal power assignment is a water-filling assignment [12].

**Theorem 2** *Given a partition  $\mathbf{a}$ , the optimal power allocation is water-filling,*

$$p_l = \begin{cases} \left( \frac{a_l - a_0}{\lambda a_l} - \frac{N_0}{q(a_0)} \right)^+ & l = 0 \\ \left( \frac{1}{\lambda} - \frac{N_0}{q(a_l)} \right)^+ & l > 0 \end{cases}, \quad (18)$$

where the water-filling level  $\lambda$  is chosen to satisfy the average power constraint  $\rho(\mathbf{p}, \mathbf{a}) = \bar{p}$ .

Note that the powers  $p_l$  allocated according to Theorem 2 are non-decreasing in  $l$ . Also, given an arbitrary partition  $\mathbf{a}$ , water-filling may result a collection of  $l' > 1$  intervals  $\{\mathcal{V}_l | l \leq l'\}$  with power  $p_l = 0$ . However, there is no benefit in terms of the average reliable throughput to design policies with more than one zero power interval. The following lemma shows that the average reliable throughput will increase by dividing an interval with non-zero power.

**Lemma 3** *A quantization interval  $[a_l, a_{l+1})$  with  $p_l > 0$  can be split into two intervals,  $[a_l, x)$  and  $[x, a_{l+1})$ , such that the throughput contributed by channel state  $s \in [q(a_l), q(a_{l+1}))$  will strictly increase while the power assignment for  $s \in [q(a_l), q(a_{l+1}))$  stays the same.*

As illustrated in Fig. 3, greater efficiency can result from repartitioning an arbitrary policy by merging all intervals with zero power into a single interval and subdividing a non-zero power interval. Thus Lemma 3 yields the following corollary.

**Corollary 2** *For an optimal policy  $(\mathbf{p}^*, \mathbf{a}^*)$  only  $p_0^*$  can be zero.*

## 4 An Asymptotically Optimal Policy

The ergodic capacity of a fading channel [5] can be written as

$$C = \int_{s_0}^{\infty} \log \left( \frac{s}{s_0} \right) f(s) ds, \quad (19)$$

where  $s_0$  is a cut-off value which is strictly positive for a finite average power constraint

$$\int_{s_0}^{\infty} p(s) f(s) ds = \bar{p}, \quad (20)$$

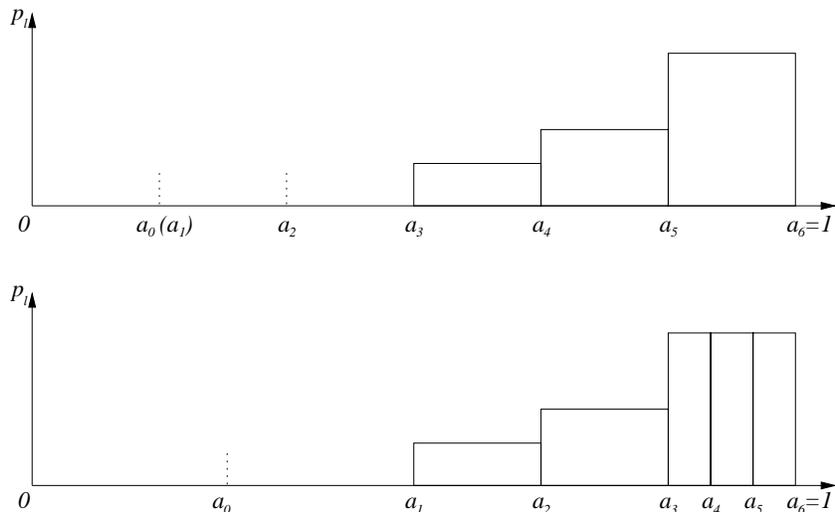


Figure 3: Illustration of the repartitioning step for  $L = 6$  and  $m = 3$ .

where  $p(s)$  is a continuous water-filling power assignment given by

$$p(s) = N_0 \left( \frac{1}{s_0} - \frac{1}{s} \right)^+ . \quad (21)$$

Intuitively, for our discrete adaptive system, when  $L$  increases to infinity, the corresponding capacity  $C_L$  should converge to the ergodic capacity. For a reasonably good policy  $(\mathbf{p}, \mathbf{a})$ , it is expected that the corresponding average reliable throughput  $R_L(\mathbf{p}, \mathbf{a})$  should also converge to the ergodic capacity. This property is referred to as *asymptotic optimality*.

Although we gained some insight on how to construct good policies in the previous section, the design of an optimal partition  $\mathbf{a}^*$  remains unknown. Here, we verify the asymptotic optimality of a policy based on a channel state partition which is uniform in probability. This design is a building block and the starting point of our iterative algorithm in the next section. We compare the ergodic capacity with the average reliable throughput of an  $L$ -level policy  $(\mathbf{p}^\dagger, \mathbf{a}^\dagger)$  with the rate/quantization level assignment

$$a_l^\dagger = \begin{cases} 1/L, & l = 0 \\ l/L, & l = 1, 2, \dots, L-1 \end{cases} , \quad (22)$$

and the water-filling power assignment (18). Here, we deliberately set  $a_0^\dagger = a_1^\dagger$  which leads to  $p_0^\dagger = 0$  after water-filling. Although suboptimal, this choice will simplify subsequent arguments. Based on (18), the rest of  $p_l^\dagger$ 's can be written as

$$p_l^\dagger = N_0 \left( \frac{1}{\lambda^\dagger N_0} - \frac{1}{q(a_l^\dagger)} \right)^+ , \quad (23)$$

Since  $(\mathbf{p}^\dagger, \mathbf{a}^\dagger)$  is sub-optimal,  $R_L^\dagger = R_L(\mathbf{p}^\dagger, \mathbf{a}^\dagger)$  satisfies

$$R_L^\dagger < C_L < C \quad (24)$$

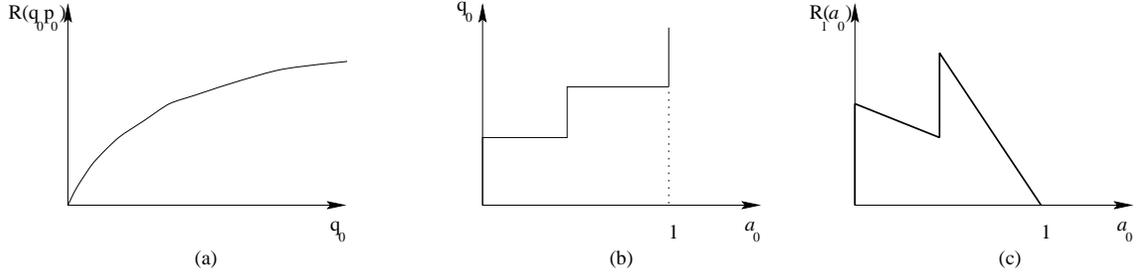


Figure 4: Local optima ( $q_0 = q(a_0)$ ).

The similarity between the power allocation functions  $p(s)$  in (21) and  $p_l^\dagger$  in (23) allows us to show in the following theorem that  $R_L^\dagger$  is asymptotically optimal.

**Theorem 3**  $\lim_{L \rightarrow \infty} R_L^\dagger = C$ .

The implication of Theorem 3 is that for sufficiently large  $L$ , the optimization over the partition  $\mathbf{q}$  is of less importance as long as the power allocation is water-filling and the code rates are chosen accordingly. On the other hand, an arbitrarily chosen partition  $\mathbf{a}$  may not lead to an asymptotically optimal policy. Moreover, this emphasizes that the joint partition and power-rate optimization will offer more significant improvement when  $L$  is small.

## 5 Iterative Policy Improvement

Finding an optimal policy  $(\mathbf{p}^*, \mathbf{a}^*)$  that achieves the capacity  $C_L$  is a challenging optimization problem. The following simple example illustrates that the problem could be non-convex and that there might be multiple local maxima. In this example, we assume  $L = 1$  so that the equality in the average power constraint yields  $p_0 = \bar{p}$ . The objective function (average reliable throughput) is

$$R_1(a_0) = (1 - a_0)R(q_0 p_0), \quad (25)$$

where  $q_0 = q(a_0)$  and  $R(q_0 p_0)$  is a monotonically increasing function shown in Fig. 4(a). Let  $q_0$  be the one shown in Fig. 4(b), then  $R_1(a_0)$  has two local maxima and is not concave as shown in Fig. 4(c).

One way of finding the optimal policy is brute force maximization of (17) over a quantized space of all possible pairs of power and partition assignments. This approach entails quantization of continuous policy variables  $\mathbf{p}$  and  $\mathbf{a}$  and can only be taken for a very small number of quantization levels  $L$  since its complexity increases exponentially in  $L$ . In Fig. 6, we present an iterative algorithm that finds a good policy  $(\mathbf{p}^\ddagger, \mathbf{a}^\ddagger)$  with throughput  $R_L^\ddagger = R_L(\mathbf{p}^\ddagger, \mathbf{a}^\ddagger)$ . The first two steps initialize the algorithm with the asymptotically optimal policy  $(\mathbf{p}^\dagger, \mathbf{a}^\dagger)$  with throughput  $R_L^\dagger$ . The rest of the algorithm consists of the following three local optimization techniques:

- Water-filling (Theorem 2) to optimize power allocation given a partition,
- Water-spilling to optimize the intervals while satisfying by the power constraint (16),

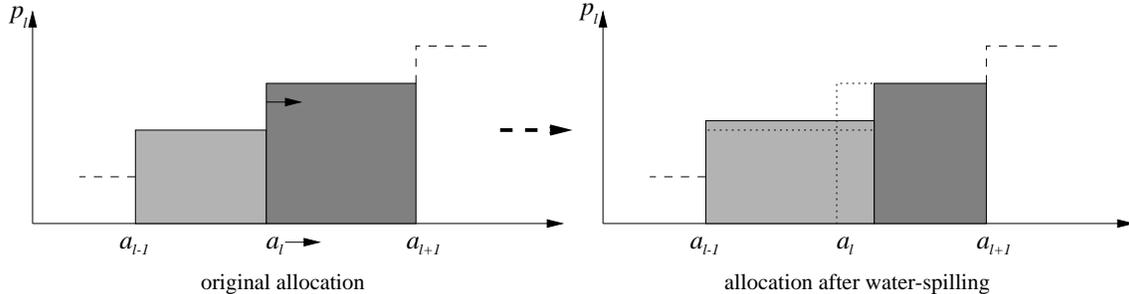


Figure 5: Illustration of the water-spilling step. As the boundary  $a_l$  moves to the right, power from interval  $\mathcal{V}_l = [a_l, a_{l+1}]$  spills over the boundary to fill the interval  $\mathcal{V}_{l-1}$ . The figure on the left shows the increase in  $p_{l-1}$  as the boundary  $a_l$  is increased.

- Repartitioning (Lemma 3) to re-allocate zero power intervals.

The convergence of the algorithm follows from the fact that the average reliable throughput will be nondecreasing at each step of the algorithm and is upper-bounded by  $C$  [5]. Thus,

$$R_L^\dagger \leq R_L^\ddagger \leq C, \quad (26)$$

We note that a water-filling step may result in several zero power assignments ( $p_l = 0$ ), which is suboptimal according to Corollary 2. We will see that water-spilling may not be able to remedy this. Thus we employ repartitioning, as illustrated in Fig. 3, as an intermediate step. Under repartitioning, intervals with zero power are merged to satisfy Corollary 2 and  $\mathcal{V}_{L-1}$  is partitioned following Lemma 3. These two operations ensure that the average reliable throughput will increase after this step.

For iterative adjustment of the partitions, we employ a local optimization technique called *water-spilling* that varies boundaries  $a_l$  one at a time. Fig. 5 depicts a policy before and after such local optimization. The water-spilling technique is designed to satisfy the average power constraint (15). Since the quantization intervals in Fig. 5 are represented in terms of  $\mathbf{a}$ , the average power assigned to a given interval is equal to the area of its respective rectangle. To ensure that the power constraint (15) is satisfied regardless of the change in  $a_l$ , the sum of the areas of the shaded rectangles must remain the same. Consequently, when we shift  $a_l$  to the right (or, equivalently, increase the rate assignment  $r_l$ ), power spills from the interval  $l$  to raise the power  $p_{l-1}$ .

$$p_{l-1}(a_l) = \begin{cases} [\bar{p}_l - p_l(a_{l+1} - a_l)] / (a_l - a_{l-1}) & l > 1 \\ [\bar{p}_l - p_l(a_{l+1} - a_l)] / a_l & l = 1 \end{cases}, \quad (27)$$

where  $\bar{p}_l$  is the average power over  $\mathcal{V}_{l-1}$  and  $\mathcal{V}_l$ . Consequently, water-spilling suggests maximizing the average reliable throughput over a single variable  $a_l$ .

Using  $F'(s)$  and  $F''(s)$  to denote the first and the second order derivatives of a function  $F(s)$  with respect to  $S$ , respectively. We show in Appendix B that if the fading CDF  $F(s)$  satisfies

$$\frac{2}{1 - F(s)} + \frac{F''(s)}{[F'(s)]^2} \geq 0, \quad s \in \mathcal{S}, \quad (28)$$

the water-spilling objective function is concave in  $a_l$ . Therefore, any convex searching algorithm can be used in this step.

- 
1.  $k = 0$ . Choose interval boundaries  $a_l^{(k)} = l/L$  for  $l = 1, \dots, L$  and set  $a_0^{(k)} = 1/L$ .
  2. *water-filling*: Given  $\mathbf{a}^{(k)}$ , find  $\mathbf{p}^{(k)}$  from the water-filling assignment (18).
  3. *repartitioning*: If  $p_1^{(k)} = 0$ , let  $n = \min\{m | p_m^{(k)} > 0\}$ . Define the new policy  $(\mathbf{p}, \mathbf{a})$  as

$$\begin{aligned}
a_l &= \begin{cases} a_n^{(k)}/2, & l = 0 \\ a_{l+n-1}^{(k)}, & l = 1, \dots, L-n \\ a_{L-1}^{(k)} + (1 - a_{L-1}^{(k)})\frac{l-n}{L-n}, & \text{otherwise} \end{cases} \\
p_l &= \begin{cases} 0, & l = 0 \\ p_{l+n-1}^{(k)}, & l = 1, \dots, L-n \\ p_{L-1}^{(k)}, & \text{o.w.} \end{cases}
\end{aligned}$$

Otherwise,  $\mathbf{p} = \mathbf{p}^{(k)}$  and  $\mathbf{a} = \mathbf{a}^{(k)}$ .

4. *water-spilling*: For  $l = L-1, L-2, \dots, 1$ :

- Using (27), let  $(p'_{l-1}, a'_l) = \arg \max_{(p_{l-1}(a_l), a_l)} R_L(\mathbf{p}, \mathbf{a})$ .
- Set

$$\begin{aligned}
\mathbf{p} &= [p_0, \dots, p_{l-2}, p'_{l-1}, p'_l, \dots, p'_{L-2}, p_{L-1}]^\top \\
\mathbf{a} &= [a_0, \dots, a_{l-1}, a'_l, a'_{l+1}, \dots, a'_{L-1}]^\top
\end{aligned}$$

Let  $a'_0 = \arg \max_{a_0} R_L(\mathbf{p}, \mathbf{a})$ . Set  $\mathbf{a}^{(k)} = [a'_0, a'_1, \dots, a'_{L-1}]^\top$  and  $\mathbf{p}^{(k)} = \mathbf{p}$ .

5. If  $R_L(\mathbf{p}^{(k)}, \mathbf{a}^{(k)}) - R_L(\mathbf{p}^{(k-1)}, \mathbf{a}^{(k-1)}) < \epsilon$  let  $\mathbf{p}^\ddagger = \mathbf{p}^{(k)}$  and  $\mathbf{a}^\ddagger = \mathbf{a}^{(k)}$ . Define  $R_L^\ddagger = R_L(\mathbf{p}^\ddagger, \mathbf{a}^\ddagger)$  and stop. Otherwise, set  $\mathbf{a}^{(k+1)} = \mathbf{a}^{(k)}$ ,  $k = k + 1$ , and go back to step 2.
- 

Figure 6: The iterative algorithm combining water-filling, repartitioning, and water-spilling.

## 6 Numerical results

In this section, we present numerical results for the capacity  $C_L$  and lower bounds of  $C_L$  obtained by brute force maximization over a quantized space of all possible policies and by using the proposed iterative algorithm, respectively. We compare  $C_L$  and lower bounds with the ergodic capacity  $C$  for two different fading models: Rayleigh and log-normal fading.

### 6.1 Rayleigh Fading Channel

For a Rayleigh fading channel, the fading CDF is

$$F_R(s) = 1 - e^{-s}, \quad s \in \mathcal{S}. \quad (29)$$

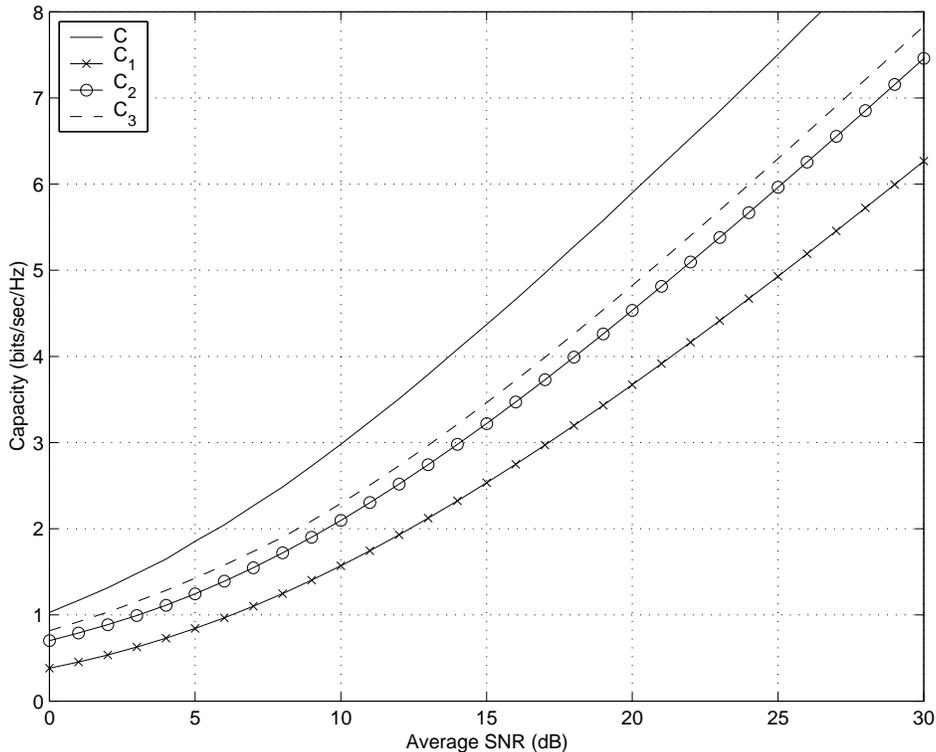


Figure 7:  $C_L$  with  $L = 1, 2, 3$  for Rayleigh fading.

In this case, it is easily verified that the concave condition (28) holds. Fig. 7 shows there is a 6 to 7 dB gap between the curves of  $C$  and  $C_1$  for a capacity around 1 to 2 bits/sec/Hz. As  $C_1$  is the capacity for constant-rate constant-power transmission, the gap indicates potential gains of the adaptive transmission system. By applying an  $L = 2$  adaptive transmission policy, the required SNR can be reduced by approximately 3 dB in comparison with that by using a constant-power and constant-rate policy. In other words, a 2-level adaptive system can eliminate about half of the SNR gap between the curves of the ergodic and the non-adaptive capacity. Furthermore, increasing  $L$  from 2 to 3 yields another 1 dB reduction in the SNR requirement. Note that  $C_2$  and  $C_3$  were obtained by brute force maximization (searching).

Fig. 8 provides a comprehensive comparison among  $C$ ,  $C_L$ ,  $R_L^\dagger$ , and  $R_L^\ddagger$ . Note that, for a small  $L = 2$ , there is only a negligible difference between  $C_L$  and  $R_L^\ddagger$ . For a relatively large  $L = 10$ ,  $R_L^\ddagger$  is within 1 dB from  $C$  when SNR is less than 20 dB. Since  $R_L^\ddagger$  is a lower bound of  $C_L$ , we expect that  $C_L$  is very close to  $C$  with  $L > 10$ .

Meanwhile, it is worthwhile pointing out that the improvement of  $R_L^\dagger$  over  $R_L^\ddagger$  is a result of iterative optimization. This improvement is quite substantial when  $L$  is small and the SNR is high. However, when the SNR is less than 10 dB, there is only a negligible difference between  $R_L^\dagger$  and  $R_L^\ddagger$ , where the average reliable throughput is about 2 bits/sec/Hz or less. On the other hand, even with high SNRs, if  $L$  is sufficiently large,  $R_L^\dagger$  and  $R_L^\ddagger$  will also be

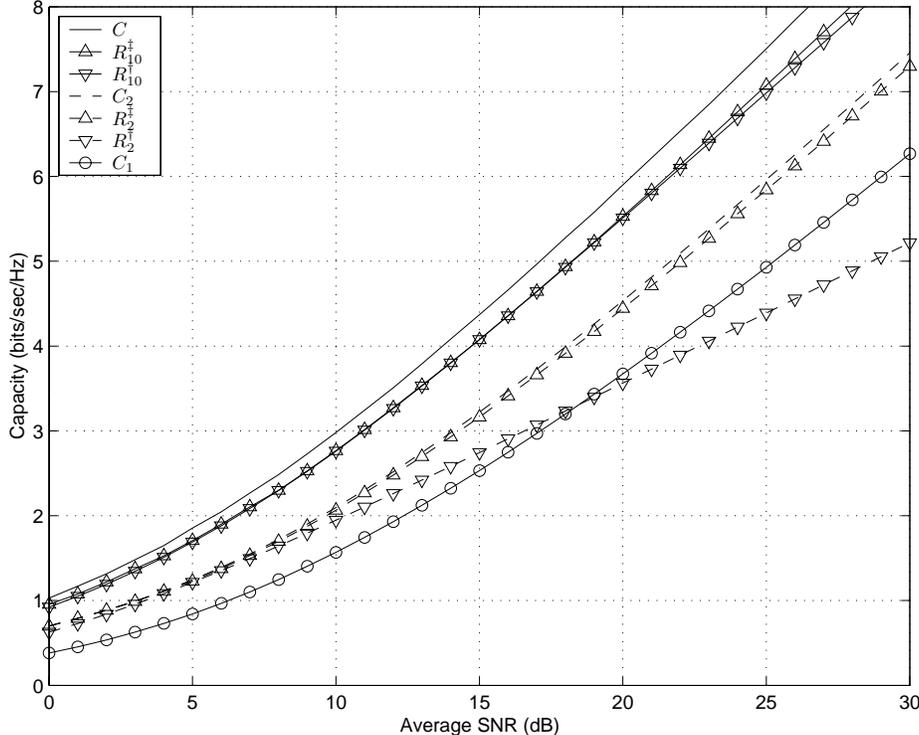


Figure 8: A comparison of  $C$ ,  $C_L$ ,  $R_L^\dagger$ , and  $R_L^\ddagger$  for Rayleigh fading.

very close. This is an indication of the asymptotic optimality of equal probability partitions. Also, since  $(\mathbf{p}^\dagger, \mathbf{a}^\dagger)$  is not a good policy for  $L = 2$  where  $p_0^\dagger = 0$  and  $a_1^\dagger = 1 - a_1^\ddagger = 1/2$ , even  $C_1$  outperforms  $R_L^\dagger$  at high SNRs. Although space considerations do not permit inclusion of the results, similar conclusions hold for the Nakagami fading channel with a parameter  $m = 2$  [18].

## 6.2 Log-Normal Fading Channel

With a mean  $\mu$  and a variance  $\sigma$ , a PDF of a log-normal random variable  $s$  [19] is

$$f_{\text{LN}}(s) = \frac{1}{\sqrt{2\pi\sigma}} \frac{10}{\log 10} \frac{1}{s} e^{-\frac{(10 \log_{10} s - \mu)^2}{2\sigma^2}}. \quad (30)$$

Using a numerical approach, we find that (28) holds for  $\sigma \leq 6$ .

In Fig. 9, a comprehensive comparison with  $C$ ,  $C_L$ ,  $R_L^\dagger$ , and  $R_L^\ddagger$  for the log-normal channel with  $\sigma = 6$  is shown. In comparison with the case of the Rayleigh fading channel, it requires a relative larger  $L$  for an adaptive system design in order to achieve a performance close to  $C$  in the log-normal fading channel.

According to [19], a typical value of  $\sigma$  is 8 dB. When  $\sigma = 8$ , the log-normal distribution does not satisfy the concavity condition (28). However, in this case, a line-search algorithm can be used to substitute a convex searching algorithm at the water-spilling step. The difference between the capacity for two different  $\sigma$  values is small.

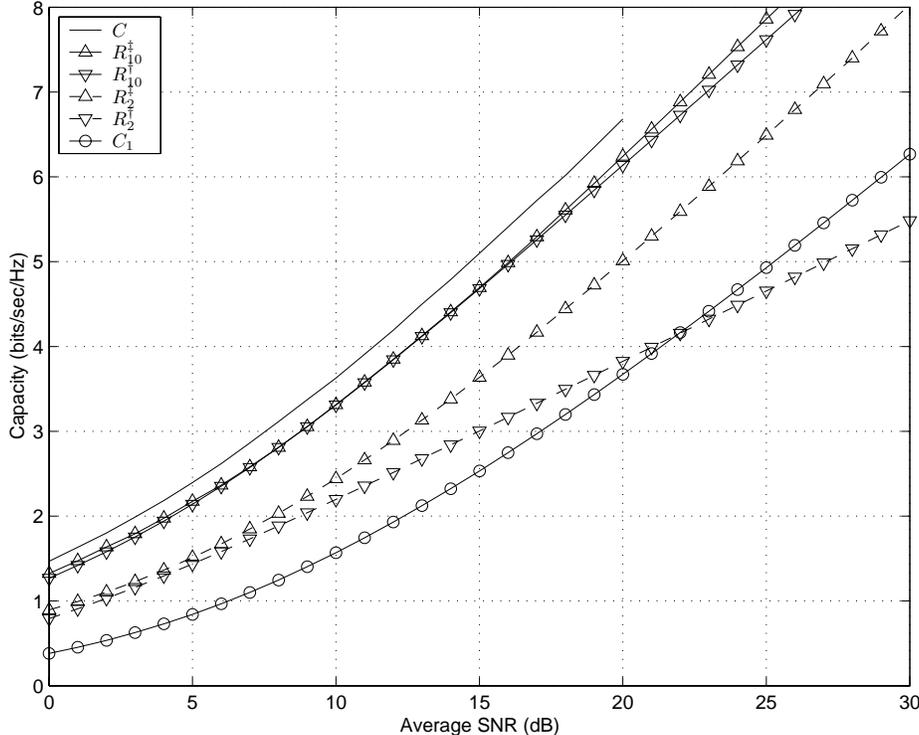


Figure 9: A comparison of  $C$ ,  $C_L$ ,  $R_L^\dagger$ , and  $R_L^\dagger$  for log-normal fading ( $\sigma = 6$ ).

## 7 Conclusion

Following the seminal work [5] and the recently introduced concepts of information outage and outage probability [11, 14, 15], we construct the throughput maximization problem over possible candidate policies, which are the ensembles of channel state space partitions and power-rate allocations. In addition, we observe several properties of optimal policies achieving the maximum throughput. These properties give insight for understanding and simplifying the throughput maximization problem.

Although we do not construct optimal policies, we suggest three options based on the observed properties of the optimal policies. The first approach is to exhaustively search for good policies over possible policies with a good resolution. Unfortunately, this is only effective for the cases where  $L$  is relatively small since its computation complexity increases exponentially in  $L$ . Another approach is to design a reasonably good policy, as shown in Section 4, which provides a very good approximation to an optimal policy when  $L$  is large. The last one is an iterative local search algorithm which is particularly useful where the first approach is impractical and the second one is not good enough.

Finally, the most important message conveyed by this research is that a carefully designed discrete adaptive system with a small number of power levels and code rates can achieve the results close to those obtained by the well-known optimum continuous adaptive transmission system in several slow fading environments.

## A Proofs

### Proof: Lemma 1

When  $p_l^*$  is zero, (8) is obvious. Otherwise, suppose there is an optimal policy  $(\mathbf{p}, \mathbf{v}, \mathbf{r})$  for which  $p_l > 0$  and  $r_l \notin [R(p_l v_{l+1}), R(p_l v_l)]$  for some  $l$ . We construct the policy  $(\mathbf{p}, \mathbf{v}, \mathbf{r}')$  with  $\mathbf{r}' = [r_0, \dots, r_{l-1}, r'_l, r_{l+1}, \dots, r_{L-1}]^\top$  with  $r'_l = R(p_l v_l)$ . It is straightforward to show that  $(\mathbf{p}, \mathbf{v}, \mathbf{r}') \in \pi_L(\bar{p})$ . We consider two cases. First, if  $r_l > R(p_l v_{l+1})$ , then

$$R_L(\mathbf{p}, \mathbf{v}, \mathbf{r}') - R_L(\mathbf{p}, \mathbf{v}, \mathbf{r}) = F(v_l, v_{l+1})r'_l > 0, \quad (31)$$

which is a contradiction. Second, if  $r_l < R(p_l v_l)$ ,

$$R_L(\mathbf{p}, \mathbf{v}, \mathbf{r}') - R_L(\mathbf{p}, \mathbf{v}, \mathbf{r}) = (r'_l - r_l)F(v_l, v_{l+1}) > 0, \quad (32)$$

which is also a contradiction.

### Proof: Lemma 2

Given a policy  $(\mathbf{p}, \mathbf{v}, \mathbf{q})$ , with average power  $\sum_{l=0}^{L-1} F(v_l, v_{l+1})p_l = \bar{p} - \epsilon$ , we can construct a new policy  $(\mathbf{p}', \mathbf{v}, \mathbf{q}) \in \pi_L(\bar{p})$  where  $\mathbf{p}' = [p_0 + \epsilon, \dots, p_{L-1} + \epsilon]^\top$ . This policy achieves

$$R_L(\mathbf{p}', \mathbf{v}, \mathbf{q}) = \sum_{l=0}^{L-1} F(q_l, v_{l+1})R((p_l + \epsilon)q_l) > R_L(\mathbf{p}, \mathbf{v}, \mathbf{q}). \quad (33)$$

□

### Proof: Theorem 1

Given an arbitrary policy  $(\mathbf{p}, \mathbf{v}, \mathbf{q}) \in \pi_L(\bar{p})$ , suppose that there is an  $l$  such that  $v_{l+1} < q_{l+1}$  and  $0 \leq l < L - 1$ . We will construct a new policy  $(\mathbf{p}', \mathbf{v}', \mathbf{q}) \in \pi_L(\bar{p})$  such that  $v'_{l+1} = q_{l+1}$  and  $R_L(\mathbf{p}', \mathbf{v}', \mathbf{q}) \geq R_L(\mathbf{p}, \mathbf{v}, \mathbf{q})$ . If such scenarios appear more than once, we can repeat the same construction for each such  $l$ .

Let

$$\hat{p} = F(q_{l+1}, v_{l+2})p_{l+1} + \sum_{i \neq \{l, l+1\}} F(v_i, v_{i+1})p_i, \quad (34)$$

denote the portion of the average transmitted power  $\rho(\mathbf{p}, \mathbf{v}, \mathbf{q})$  associated with channel states  $s \notin [v_l, q_{l+1}]$ . Therefore, we can write

$$\rho(\mathbf{p}, \mathbf{v}, \mathbf{q}) = \hat{p} + F(v_l, v_{l+1})p_l + F(v_{l+1}, q_{l+1})p_{l+1} \quad (35)$$

$$= \hat{p} + F(v_l, q_{l+1})\tilde{p}, \quad (36)$$

where

$$\tilde{p} = \{F(v_l, v_{l+1})p_l + F(v_{l+1}, q_{l+1})p_{l+1}\} / F(v_l, q_{l+1}). \quad (37)$$

Based on  $\tilde{p}$ , we construct a new policy  $(\mathbf{p}', \mathbf{v}', \mathbf{q})$  with  $\mathbf{v}' = [v_0, \dots, v_l, q_{l+1}, v_{l+2}, \dots, v_{L-1}]^\top$  and  $\mathbf{p}' = [p_0, \dots, p_{l-1}, \tilde{p}, p_{l+1}, \dots, p_{L-1}]^\top$ . In the policy  $(\mathbf{p}', \mathbf{v}', \mathbf{q})$ , transmitted power  $\tilde{p}$  is used for all states  $s \in [v_l, q_{l+1}]$ . Hence,

$$\rho(\mathbf{p}', \mathbf{v}', \mathbf{q}) = \hat{p} + F(v_l, q_{l+1})\tilde{p}. \quad (38)$$

It follows from (36) that  $\rho(\mathbf{p}', \mathbf{v}', \mathbf{q}) = \rho(\mathbf{p}, \mathbf{v}, \mathbf{q})$ . The corresponding average reliable throughput is

$$R_L(\mathbf{p}', \mathbf{v}', \mathbf{q}) = \hat{R} + F(q_l, q_{l+1})R(q_l \tilde{p}), \quad (39)$$

where  $\hat{R} = \sum_{i \neq l} F(q_i, v_{i+1})R(q_i p_i)$  denotes the contributions to  $R_L(\mathbf{p}, \mathbf{v}, \mathbf{q})$  from channel states  $s \notin \mathcal{V}_l$ . Defining  $\alpha = F(q_l, v_{l+1})/F(q_l, q_{l+1})$ , we observe from (37) that

$$\tilde{p} = \frac{p_l F(v_l, q_l) + \alpha p_l F(q_l, q_{l+1}) + p_{l+1} F(v_{l+1}, q_{l+1})}{F(v_l, q_{l+1})} \quad (40)$$

$$\geq \frac{p_l F(v_l, q_l) + \alpha p_l F(q_l, q_{l+1})}{F(v_l, q_{l+1})} \quad (41)$$

Since  $\alpha \leq 1$ ,

$$\tilde{p} \geq \frac{\alpha p_l (F(v_l, q_l) + F(q_l, q_{l+1}))}{F(v_l, q_{l+1})} = \alpha p_l. \quad (42)$$

From (39), (42) and the fact that  $R(\cdot)$  is monotonic increasing, we observe that

$$R_L(\mathbf{p}', \mathbf{v}', \mathbf{q}) \geq \hat{R} + F(q_l, q_{l+1})R[q_l(\alpha p_l + 0)] \quad (43)$$

$$\stackrel{(a)}{\geq} \hat{R} + F(q_l, q_{l+1})[\alpha R(q_l p_l) + (1 - \alpha)R(0)] \quad (44)$$

$$\stackrel{(b)}{=} \hat{R} + F(q_l, q_{l+1})\alpha R(q_l p_l) \quad (45)$$

$$= \hat{R} + F(q_l, v_{l+1})R(q_l p_l) = R_L(\mathbf{p}, \mathbf{v}, \mathbf{q}). \quad (46)$$

Note that inequality (a) is due to the concavity of  $R(\cdot)$  while (b) holds because  $R(0) = 0$ .  $\square$

### Proof: Theorem 2

Following the standard Lagrange procedure, we have

$$J = \lambda \bar{p} + \sum_{l=0}^{L-1} (a_{l+1} - a_l) [R(p_l q(a_l)) - \lambda p_l] - \lambda a_0 p_0. \quad (47)$$

For  $l > 0$ ,

$$\frac{\partial J}{\partial p_l} = (a_{l+1} - a_l) \left( \frac{q(a_l)/N_0}{1 + p_l q(a_l)/N_0} - \lambda \right) = 0. \quad (48)$$

This leads to

$$p_l = \left( \frac{1}{\lambda} - \frac{N_0}{q(a_l)} \right)^+. \quad (49)$$

Similarly, if  $a_0 > 0$ , we have

$$\frac{\partial J}{\partial p_0} = (a_1 - a_0) \frac{q(a_0)/N_0}{1 + p_0 q(a_0)/N_0} - \lambda a_1 = 0, \quad (50)$$

and, thus,

$$p_0 = \left( \frac{a_1 - a_0}{\lambda a_1} - \frac{N_0}{q(a_0)} \right)^+. \quad (51)$$

If  $a_0 = 0$ , then  $p_0 = 0$ . Based on (16),  $1/\lambda$  can be solved and (18) follows immediately.  $\square$

**Proof: Lemma 3**

$$(a_l - a_{l-1})R(p_l q(a_l)) = (a_l - x)R(p_l q(a_l)) + (x - a_{l-1})R(p_l q(a_l)) \quad (52)$$

$$< (a_l - x)R(p_l q(x)) + (x - a_{l-1})R(p_l q(a_l)). \quad (53)$$

$\square$

**Proof: Theorem 3**

For a sufficiently large  $L$ , there exists an  $l_0(L) > 0$  such that  $s_0 \in \mathcal{V}_{l_0(L)}^\dagger$ . Even though  $l_0(L)$  is a function of  $L$ , we will drop  $(L)$  to simplify the following derivations. Using the notation  $q_l^\dagger = q(a_l^\dagger)$ , we can write  $q_{l_0}^\dagger \leq s_0 \leq q_{l_0+1}^\dagger$ . It follows from (21) that

$$p(s) \geq N_0 \left( \frac{1}{q_{l_0+1}^\dagger} - \frac{1}{s} \right)^+ \quad (54)$$

For  $s \in \mathcal{V}_l^\dagger$ ,  $s \geq q_l^\dagger$ , so that

$$p(s) \geq N_0 \left( \frac{1}{q_{l_0+1}^\dagger} - \frac{1}{q_l^\dagger} \right)^+ \quad s \in \mathcal{V}_l^\dagger \quad (55)$$

With the same average power  $\bar{p}$  for both the discrete power assignment (23) and the continuous policy (21), we have

$$\frac{1}{L} \sum_{l=1}^{L-1} N_0 \left( \frac{1}{\lambda^\dagger N_0} - \frac{1}{q_l^\dagger} \right)^+ = \sum_{l=0}^{L-1} \int_{\mathcal{V}_l^\dagger} p(s) f(s) ds \quad (56)$$

$$\geq \frac{1}{L} \sum_{l=1}^{L-1} N_0 \left( \frac{1}{q_{l_0+1}^\dagger} - \frac{1}{q_l^\dagger} \right)^+ \quad (57)$$

We observe that the inequality (57) implies  $q_{l_0+1}^\dagger > \lambda^\dagger N_0$ . It follows from (23) that

$$p_l^\dagger \geq N_0 \left( \frac{1}{q_{l_0+1}^\dagger} - \frac{1}{q_l^\dagger} \right)^+ \quad (58)$$

Now we derive a lower bound to  $R_L^\dagger$ .

$$R_L^\dagger = \sum_{l=1}^{L-1} \frac{1}{L} R(p_l^\dagger q_l^\dagger) \geq \sum_{l=l_0+1}^{L-1} \frac{1}{L} R(p_l^\dagger q_l^\dagger) \quad (59)$$

Applying the lower bound (58), we obtain

$$R_L^\dagger \geq R_{l_0}(L) = \frac{1}{L} \sum_{l=l_0+1}^{L-1} \log \left( \frac{q_l^\dagger}{q_{l_0+1}^\dagger} \right) \quad (60)$$

Next, we upperbound  $C$  by  $R_{l_0}(L)$  plus terms that will go to zero with increasing  $L$ . Since  $s_0 \geq q_{l_0}^\dagger$ , we see from (19) that

$$C \leq \int_{q_{l_0}^\dagger}^{\infty} \log\left(\frac{s}{q_{l_0}^\dagger}\right) f(s) ds \quad (61)$$

$$= \sum_{l=l_0+1}^{L-1} \int_{\mathcal{V}_{l-1}^\dagger} \log\left(\frac{s}{q_{l_0}^\dagger}\right) f(s) ds + I_L^\dagger \quad (62)$$

where  $I_L^\dagger$  denotes the integral

$$I_L^\dagger = \int_{\mathcal{V}_{L-1}^\dagger} \log\left(\frac{s}{q_{l_0}^\dagger}\right) f(s) ds \quad (63)$$

Below, we will take some additional care to upperbound  $I_L^\dagger$ . Since  $s \in \mathcal{V}_{l-1}^\dagger$  implies  $s \leq q_l^\dagger$ , we have from (62) that

$$C \leq \frac{1}{L} \sum_{l=l_0+1}^{L-1} \log\left(\frac{q_l^\dagger}{q_{l_0}^\dagger}\right) + I_L^\dagger \quad (64)$$

$$= \frac{1}{L} \sum_{l=l_0+1}^{L-1} \log\left(\frac{q_l^\dagger q_{l_0+1}^\dagger}{q_{l_0+1}^\dagger q_{l_0}^\dagger}\right) + I_L^\dagger \quad (65)$$

$$\leq R_{l_0}(L) + \log\left(\frac{q_{l_0+1}^\dagger}{q_{l_0}^\dagger}\right) + I_L^\dagger \quad (66)$$

Using the shorthand  $\mathcal{V}_{L-1}^\dagger$  for the event  $S \in \mathcal{V}_{L-1}^\dagger$ , we employ (63) to write

$$I_L^\dagger = \Pr[\mathcal{V}_{L-1}^\dagger] \left( E\{\log S | \mathcal{V}_{L-1}^\dagger\} - \log q_{l_0}^\dagger \right) \quad (67)$$

Since  $\Pr[\mathcal{V}_{L-1}^\dagger] = 1/L$ , and since the log function is concave,

$$I_L^\dagger \leq \frac{1}{L} \log E\{S | \mathcal{V}_{L-1}^\dagger\} - \frac{\log q_{l_0}^\dagger}{L} \quad (68)$$

The conditional expectation can also be upperbounded as

$$E\{S | \mathcal{V}_{L-1}^\dagger\} = \frac{1}{\Pr[\mathcal{V}_{L-1}^\dagger]} \int_{\mathcal{V}_{L-1}^\dagger} s f(s) ds \quad (69)$$

$$\leq L \int_0^\infty s f(s) ds = LE\{S\} \quad (70)$$

From (68) and (70), we have that

$$I_L^\dagger \leq \frac{1}{L} \log\left(\frac{LE\{S\}}{q_{l_0}^\dagger}\right) \quad (71)$$

Finally, we observe that  $q_{l_0}^\dagger \leq s_0 \leq q_{l_0+1}^\dagger$  implies that

$$F(s_0) - F(q_{l_0}^\dagger) \leq F(q_{l_0+1}^\dagger) - F(q_{l_0}^\dagger) = \frac{1}{L} \quad (72)$$

Continuity of the distribution function  $F(\cdot)$  implies that  $\lim_{L \rightarrow \infty} q_{l_0}^\dagger = \lim_{L \rightarrow \infty} q_{l_0+1}^\dagger = s_0$ . This permits us to conclude that

$$\lim_{L \rightarrow \infty} \log \left( \frac{q_{l_0+1}^\dagger}{q_{l_0}^\dagger} \right) = 0 \quad (73)$$

Similarly, (71) implies

$$\lim_{L \rightarrow \infty} I_L^\dagger = 0 \quad (74)$$

Applying (73) and (74) to (66), we see that  $C \leq \lim_{L \rightarrow \infty} R_{l_0}(L)$ . Since  $C \geq R_L^\dagger$  and (60) implies  $\lim_{L \rightarrow \infty} R_L^\dagger \geq \lim_{L \rightarrow \infty} R_{l_0}(L)$ , the theorem follows.

## B Concavity of Water-Spilling

If we let  $a_l$  to be an arbitrary value  $a \in (a_{l-1}, a_{l+1})$ , the water-spilling objective, the rate sum on  $\mathcal{V}_{l-1}$  and  $\mathcal{V}_l$ , equals

$$\tilde{R}_l = (a - a_{l-1})R(p_{l-1}(a)q(a_{l-1})) + (a_{l+1} - a)R(p_l q(a)). \quad (75)$$

Let  $\phi'$  and  $\phi''$  be the first and the second order derivatives with respect to  $a$ , respectively. Now, we have

$$\begin{aligned} \tilde{R}_l'' &= 2R'(p_{l-1}q(a_{l-1})) + (a - a_{l-1})R''(p_{l-1}q(a_{l-1})) \\ &\quad - 2R'(p_l q(a_{l-1})) + (a_{l+1} - a)R''(p_l q(a)). \end{aligned} \quad (76)$$

For the first two terms in (76), after simple manipulations, we obtain

$$2R'(p_{l-1}q(a_{l-1})) + (a - a_{l-1})R''(p_{l-1}q(a_{l-1})) = \frac{-\{[q(a_{l-1})]'\}^2}{\left\{\frac{N_0}{q(a_{l-1})} + p_{l-1}\right\}^2} < 0. \quad (77)$$

For the second two terms in (76),

$$-2R'(p_l q(a_{l-1})) + (a_{l+1} - a)R''(p_l q(a)) = \frac{-2[q(a)]' + (a_{l+1} - a)[q(a)]''}{\frac{N_0}{p_l} + q(a)} + \frac{-\{[q(a_{l-1})]'\}^2}{\left\{\frac{N_0}{p_l} + q(a)\right\}^2}. \quad (78)$$

The value of  $\tilde{R}_l''$  depends on the fading distributions as well as the power  $\bar{p}_l$ . Moreover, it is hard to gain any insight on the sign of  $\tilde{R}_l''$  based on its closed form

$$\tilde{R}_l'' = \frac{-(p_{l-1}')^2}{\left\{\frac{N_0}{q(a_{l-1})} + p_{l-1}\right\}^2} + \frac{-2[q(a)]' + (a_{l+1} - a)[q(a)]''}{\frac{N_0}{p_l} + q(a)} + \frac{-\{[q(a)]'\}^2}{\left\{\frac{N_0}{p_l} + q(a)\right\}^2}. \quad (79)$$

Note, if the numerator of the second term in (79),  $-2[q(a)]' + (a_{l+1} - a)[q(a)]''$ , is negative,  $\tilde{R}_l''$  will be negative and, henceforth, a sufficient condition for our objective function  $\tilde{R}_l$  to be concave is

$$-2[q(a)]' + (a_{l+1} - a)[q(a)]'' \leq 0. \quad (80)$$

It is worthwhile pointing out that  $-2[q(a)]' + (a_{l+1} - a)[q(a)]''$  only depends on the distribution of  $S$ . Since CDF  $F(s)$  is a strictly decreasing function we have that  $[q(a)]' > 0$  and, thus, the sufficient condition can be simplified to

$$\frac{2}{1 - F(s)} + \frac{F''(s)}{[F'(s)]^2} \geq 0, \quad (81)$$

where  $F(s) = a$  for any  $s \in (q(a_{l-1}), q(a_{l+1}))$  and  $a_{l+1} - a$  is substituted by  $1 - F(s)$ . In Section 6, we show that several fading distributions satisfy (28).

The last operation at the fourth step is to maximize the rate over  $\mathcal{V}_0$  by varying  $a_0$  only. The objective function  $(a_1 - a_0)R(a_0 p_0)$  is also concave if (28) is satisfied.

## References

- [1] F. Adachi, M. Sawahashi, and H. Suda, "Wideband DS-CDMA for next-generation mobile communications systems," *IEEE Communications Magazine*, vol. 36, no. 9, pp. 56–69, Sep. 1998.
- [2] D. Knisely, S. Kumar, S. Laha, and S. Nanda, "Evolution of wireless data services: IS-95 to CDMA2000," *IEEE Communications Magazine*, vol. 36, no. 10, pp. 140–149, Oct. 1998.
- [3] P. Bender, P. Black, M. Grob, R. Padovani, N. Sindhushyana, and S. Viterbi, "CDMA/HDR: a bandwidth efficient high speed wireless data service for nomadic users," *IEEE Communications Magazine*, vol. 38, no. 7, pp. 70–77, Jul. 2000.
- [4] K. Balachandran, R. Ejzak, S. Nanda, S. Vitebskiy, and S. Seth, "GPRS-136: high-rate packet data service for North American TDMA digital cellular systems," *IEEE Personal Communications*, vol. 6, no. 3, pp. 34–47, Jun. 1999.
- [5] A. Goldsmith and P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Transactions on Information Theory*, vol. 43, no. 6, pp. 1986–1992, Nov. 1997.
- [6] D. Tse and S. Hanly, "Multi-access fading channels: Part I: Polymatroid structure, optimal resource allocation and throughput capacities," *IEEE Transactions on Information Theory*, vol. 44, no. 7, pp. 2796–2815, Nov. 1998.
- [7] G. Caire, G. Taricco, and E. Biglieri, "Optimum power control over fading channels," *IEEE Transactions on Information Theory*, vol. 45, no. 5, pp. 1468–1489, Jul. 1999.

- [8] A. Goldsmith and S. Chua, “Variable-rate variable-power MQAM for fading channels,” *IEEE Transactions on Communications*, vol. 45, no. 10, pp. 1218–1229, Oct. 1997.
- [9] A. Goldsmith and S. Chua, “Adaptive coded modulation for fading channels,” *IEEE Transactions on Communications*, vol. 46, no. 5, pp. 595–602, May 1998.
- [10] D. Goeckel, “Adaptive coding for time-varying channels using outdated fading estimates,” *IEEE Transactions on Communications*, vol. 47, no. 6, pp. 844–855, Jun. 1999.
- [11] E. Biglieri, J. Proakis, and S. Shamai, “Fading channels: Information-theoretic and communications aspects,” *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2619–2692, Oct. 1998.
- [12] T. Cover and J. Thomas, *Elements of Information Theory*, John Wiley & Sons, Inc., 1991.
- [13] G. Caire and S. Shamai, “On the capacity of some channels with channel state information,” *IEEE Transactions on Information Theory*, vol. 45, no. 6, pp. 2007–2019, Sep. 1999.
- [14] L. Ozarow, S. Shamai, and A. Wyner, “Information theoretic considerations for cellular mobile radio,” *IEEE Transactions on Vehicular Technology*, vol. 43, no. 2, pp. 359–378, May 1994.
- [15] E. Telatar, “Capacity of multi-antenna Gaussian channels,” *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, Nov.-Dec. 1999.
- [16] R. Knopp and P. Humblet, “On coding for block fading channels,” *IEEE Transactions on Information Theory*, vol. 46, no. 1, pp. 189–205, Jan. 2000.
- [17] E. Malkamäki and H. Leib, “Coded diversity on block-fading channels,” *IEEE Transactions on Information Theory*, vol. 45, no. 2, pp. 771–781, Mar. 1999.
- [18] J. G. Proakis, *Digital Communications, 3rd Edition*, McGraw-Hill, 1995.
- [19] T. Rappaport, *Wireless Communications: Principles & Practice*, Prentice Hall, 1996.