On the Performance of Adaptive Modulation in Cellular Systems

Xiaoxin Qiu, *Member, IEEE*, and Kapil Chawla, *Senior Member, IEEE*

***Abstract*— Adaptive modulation techniques have the potential to substantially increase the spectrum efficiency and to provide different levels of service to users, both of which are considered important for third-generation cellular systems. In this work, we propose a general framework to quantify the potential gains of such techniques. Specifically, we study the throughput performance gain that may be achieved by combining adaptive modulation and power control. Our results show that: 1) using adaptive modulation even without any power control provides a significant throughput advantage over using signal-to-interference-plus-noise ratio (SINR) balancing power control and 2) combining adaptive modulation and a suitable power control scheme leads to a significantly higher throughput as compared to no power control or using SINR-balancing power control. The first observation is especially important from an implementation point of view. Adjusting the modulation level without changing the transmission power requires far fewer measurements and feedback as compared to the SINR-balancing power control or the optimal power control. Hence, it is significantly easier to implement. Although presented in the context of adaptive modulation, the results also apply to other variable rate transmission techniques, e.g., rate adaptive coding schemes, coded modulation schemes, etc. This work provides valuable insight into the performance of variable rate transmission techniques in multi-user environments.**

***Index Terms*— Adaptive modulation, cellular system, power control, variable rate transmission.**

I. Introduction

With the projected demand for multimedia services, the ability to provide spectrally efficient and flexible data rate access is one of the important design considerations of future wireless systems. One approach to satisfy both of these requirements is to adapt the modulation and transmission power according to the instantaneous propagation conditions, interference scenarios, and traffic or data rate requirements. This technique is called *adaptive modulation* [1]–[4], which is part of the V.34 modem standard [5], and is currently used in modems to maintain an acceptable bit error rate (BER) over poor quality telephone lines. Adaptive modulation techniques have recently been proposed for two-way data transmission over cable [6]. Similar ideas have also been applied to improve the BER on mobile channels [3]. For instance, variable-rate QAM has been proposed for several third-generation wireless communications systems [7].

It has been shown in [2]–[4] that adaptive modulation effectively improves the BER performance on radio channels which suffer shadowing and fading. In [2], it is demonstrated that in a single user case, using adaptive modulation can provide a 5–10-dB gain over a fixed rate system having only power control. In [8], the authors study a so-called *average area spectral efficiency* (ASE), defined as the cumulative data rate per hertz per unit area, which can be achieved by using adaptive modulation. Using simulations and analysis, the impact of various system configurations and settings on the ASE is quantified. In particular, the authors obtain the channel reuse patterns that maximize the ASE for different interference scenarios. In [9], the authors propose an adaptive modulation-based cellular system for personal multimedia communications. Using system simulations, the authors show that this system can achieve spectrum efficiencies that are three times higher than a baseline system without adaptive modulation.

We adopt a different approach to study adaptive modulation in a multi-user environment. The goal of our study is to investigate the theoretical performance limit of adaptive modulation, to gain some insight into this technique itself, and possibly to provide some guidance for real implementations. In this paper, three scenarios are studied: 1) signal-to-interference-plus-noise ratio (SINR) balancing power control [10], [11]; 2) adaptive modulation without power control; and 3) adaptive modulation with power control. In particular, for the third scenario, we develop a framework to maximize the overall throughput for a given set of users in a cellular system, by jointly optimizing the modulation and the transmission power. We consider two objectives for the optimization. The first objective is of the same form as the throughput, but has a limitation that it has several local maxima. We propose an iterative algorithm to optimize this objective. We find that while the iterations converge to a desirable solution most of the time, it is also possible for them to converge to an undesirable solution. Keeping this in mind, we define a slightly different objective function. We propose an iterative algorithm to optimize this second objective and show that it always converges to its global maximum. In addition, we also show that compared to optimizing the total throughput directly, this second objective leads to a limited degradation in throughput. Through numerical studies, we demonstrate that significant throughput gains may be achieved using either of these two schemes. In particular, we find that the throughput may almost be doubled compared to using SINR-balancing power control.

Paper approved by A. Goldsmith, the Editor for Wireless Communication of the IEEE Communication Society. Manuscript received February 18, 1998; revised August 4, 1998 and November 3, 1998. This paper was presented in part at the IEEE ICUPC'98, Florence, Italy, October 1998.

The authors are with AT&T Laboratories–Research, Red Bank, NJ 07701-7033 USA (e-mail: xqiu@research.att.com; kapil@research.att.com).

Publisher Item Identifier S 0090-6778(99)05005-9.

Although we present our work in the context of adaptive modulation, the results are also applicable to other variable-rate transmission techniques, e.g., rate adaptive coding schemes, coded modulation schemes, etc. It is expected that most systems will use a combination of these techniques to provide variable rate capabilities to users. For example, the EDGE [Enhanced Data rates for GSM (Global System for Mobile Communications) Evolution] system [12], [13], proposed as a wireless data enhancement for GSM and IS-136 (North American Time Division Multiple Access System), uses a combination of modulation and coding to achieve variable data rates. In [14], the authors show that this technique significantly increases the spectrum efficiency, as compared to "standard" GSM.

The rest of the paper is organized as follows. In Section II, we discuss the system under consideration and the performance measure of interest. In Section III, we develop a framework to optimize the system throughput, characterize the solution, and propose iterative algorithms to achieve the optimum. In Section IV, we present some numerical results that quantify the throughput performance of the proposed schemes. In Section V, we offer some concluding remarks.

II. System Model

Consider a cellular system consisting of a finite number of cells and a fixed number of independent channels that may be assigned to users. Assume that a given channel is concurrently being used by N different links or transmit-receive pairs. The links in question may be a mixture of up- and down-links. However, for concreteness, and without loss of generality, we focus on downlink transmissions.

Let $\{U_i\}$ and $\{B_i\}, i = 1 \cdots N$, denote a set of cochannel users and their serving bases, respectively. Let $\boldsymbol{G} = \{G_{ij}\}$ denote the $N \times N$ path gain matrix, where G_{ij} is defined as the path gain from base j to user i. Let P_j denote the transmission power of base j. The power received at user i from base j, R_{ij}, is given by

$$R_{ij} = G_{ij} P_j.$$

If $i = j$, this is the signal; otherwise, this is interference. For a given path gain matrix $\boldsymbol{G}$ and a transmission power vector $\boldsymbol{P} = \{P_i\}$, the SINR of user i, γ_i, is equal to

$$\gamma_i(\boldsymbol{P}) = \frac{G_{ii} P_i}{\sum_{j \neq i} G_{ij} P_j + n_i} = \frac{G_{ii} P_i}{I_i} \tag{1}$$

where n_i denotes the receiver thermal noise power and I_i denotes the total noise and interference power received by user i.

Adaptive modulation provides the system with the ability to match the effective bit rate or throughput to the interference or channel conditions of specific users. The actual choice of the constellation size, and therefore the bit rate, depends upon system and user specifics like the user mobility, the propagation condition, the interference scenario, the receiver structure, the target physical layer performance, etc. However, for a given system and a target performance, it is reasonable to assume that the throughput is a function of the received SINR at the user.

Let T_i denote the *user throughput* of user i, which is defined here as the number of bits that can be successfully sent to this user within each transmitted symbol. For our model, we assume that T_i increases with the SINR

$$T_i(\gamma_i) = \log_2(1 + k \cdot \gamma_i(\boldsymbol{P})) \tag{2}$$

where k is a constant. We show in Appendix A that this is a reasonable choice for both additive white Gaussian noise (AWGN) and Rayleigh fading environments.

This formulation has a form similar to the Shannon capacity of the given set of users. If W denotes the available bandwidth in hertz, γ denotes the signal-to-noise ratio (SNR) and C denotes the Shannon capacity, it is well known that

$$C = W \log_2(1 + \gamma). \tag{3}$$

Equations (3) and (2) are similar. Therefore, a solution that maximizes $T(\boldsymbol{P})$ for $k = 1$ also maximizes the Shannon capacity. Further, our formulation may also be used to estimate the throughput gain that may be achieved by using variable rate techniques other than adaptive modulation, as long as their throughput can be approximated by a similar expression.

Equation (2) assumes that the throughput is "continuous," that is, it can be any real number. Note that in any implementation, the number of bits transmitted within each symbol is restricted to a finite number of values. For example, when using M-QAM modulation, the constellation size is restricted to $M = 2^j, j = 1, 2, \cdots, J$, making the bits per symbol equal to one of a small set of values. However, to make the analytical model tractable, we *begin* by assuming that T_i can be any real number. We relax this assumption in the numerical results.

For a specific transmission power vector $\boldsymbol{P}$, the *total throughput* T is given by

$$\begin{aligned} T(\boldsymbol{P}) &= \sum_i \log_2(1 + k \cdot \gamma_i(\boldsymbol{P})) \\ &= \log_2\left(\prod_i (1 + k \cdot \gamma_i(\boldsymbol{P}))\right). \end{aligned} \tag{4}$$

The problem that we address is the following: *given a set of users and the corresponding path gain matrix $\boldsymbol{G}$, what is the maximum throughput that can be achieved?*

III. Maximum Throughput

In this section, we develop a framework to jointly optimize the modulation and the transmission power, in order to maximize the total throughput for a given set of users. We formulate two optimization problems and propose two iterative algorithms that achieve the corresponding optima. The objective of the first optimization has the same form as the throughput, but may have more than one local maxima. Therefore, depending on the initial vector, the corresponding iterative algorithm may converge to a suboptimal solution. The objective of the second optimization is slightly altered from the first, but has a unique maximum. We show that the corresponding iterative algorithm always converges to

this global maximum, while leading to a limited degradation compared to the optimal throughput.

A. Objective I

Let $\boldsymbol{G}$ be a positive path gain matrix, i.e., $G_{ij} > 0, \forall i, j \in \{1, \cdots, N\}$. Let $\Omega_l = \{\boldsymbol{P}: \boldsymbol{P}_{\min} \leq \boldsymbol{P} \leq \boldsymbol{P}_{\max}\}$ denote the set of feasible transmission power vectors, where $\boldsymbol{P}_{\min} = [P_1^{\min} P_2^{\min} \cdots P_N^{\min}]^T$ and $\boldsymbol{P}_{\max} = [P_1^{\max} P_2^{\max} \cdots P_N^{\max}]^T$ represent the minimum and maximum transmission power limits, respectively. Our goal is to maximize the total throughput of a given set of users under the transmission power constraint. We first consider the following optimization problem, where the objective to be maximized, called *Objective I*, is the total throughput $T(\boldsymbol{P})$

$$\max_{\boldsymbol{P}} T(\boldsymbol{P}) \quad \text{s.t. } \boldsymbol{P} \in \Omega_l. \tag{5}$$

This problem is a constrained, nonlinear optimization problem [15]. Except in special cases, it is very difficult to find closed-form solutions. Iterative algorithms, which find the solution numerically, are often used instead.

Before we propose the iterative algorithm, we characterize the objective. The following lemma may be proved from the definition of $T(\boldsymbol{P})$.

Lemma 1: The objective function $T(\boldsymbol{P})$ has the following properties over the space $\Omega = \{\mathbf{P} : \mathbf{P} \geq \mathbf{0}\}$ and therefore also over the set Ω_l:

1) $T(\boldsymbol{P}) \geq 0$.
2) $T(\boldsymbol{P})$ is continuous and differentiable.
3) $\forall \alpha > 1$, $T(\alpha \boldsymbol{P}) > T(\boldsymbol{P})$.

Further, Ω_l is a compact set, as it is a closed and bounded subset of Ω. From the *Maximum-Minimum Theorem*[1] for continuous functions over compact sets, the following lemma is true.

Lemma 2: The function $T(\boldsymbol{P})$ has a maximum on the set Ω_l.

1) Iterative Algorithm I: Let $\boldsymbol{P}^I$ denote a *local* maximum of (5). As $T(\boldsymbol{P})$ is differentiable, it follows that for any *feasible* direction $\boldsymbol{d}$ at $\boldsymbol{P}^{\mathrm{I}}$ [15]

$$\nabla T(\boldsymbol{P}^{\mathrm{I}}) \cdot \boldsymbol{d} \leq 0 \tag{6}$$

where the gradient of a function is defined as

$$\nabla T(\boldsymbol{P}) = \left[\frac{\partial T(\boldsymbol{P})}{\partial P_1} \quad \frac{\partial T(\boldsymbol{P})}{\partial P_2} \quad \cdots \quad \frac{\partial T(\boldsymbol{P})}{\partial P_N} \right]$$

and a direction $\boldsymbol{d}$ is *feasible* at $\boldsymbol{P}^{\mathrm{I}}$ if there is a $\beta > 0$: $(\boldsymbol{P}^{\mathrm{I}} + \beta \boldsymbol{d}) \in \Omega_l$.

The following may be derived from the definition of $T(\boldsymbol{P})$ and (1):

$$\frac{\partial T(\boldsymbol{P})}{\partial P_i} = T(\boldsymbol{P}) \left[\frac{k\gamma_i}{P_i(1+k\gamma_i)} - \sum_{j \neq i} \frac{k\gamma_j}{(1+k\gamma_j)} \cdot \frac{G_{ji}}{I_j} \right] \tag{7}$$

[1] Theorem 4.15 [16, p. 89]: Suppose f is a continuous real function on a compact metric space X, and

$$M = \sup_{p \in X} f(p), \quad m = \inf_{p \in X} f(p).$$

Then, there exist points $p, q \in X$ such that $f(p) = M$ and $f(q) = m$.

where, as before, $I_j = \Sigma_{m \neq j} G_{jm} P_m + n_j$ denotes the total interference received by user j. It follows from (6) that for any P_i^{I} that is in the interior of Ω_l, i.e., if $P_i^{\min} < P_i^{\mathrm{I}} < P_i^{\max}$,

$$\frac{k\gamma_i}{P_i^{\mathrm{I}}(1+k\gamma_i)} - \sum_{j \neq i} \frac{k\gamma_j}{(1+k\gamma_j)} \cdot \frac{G_{ji}}{I_j} = 0. \tag{8}$$

However, if $P_i^{\mathrm{I}} = P_i^{\min}$ or $P_i^{\mathrm{I}} = P_i^{\max}$, then (8) need not be true. It follows from Lemma 1 3) that $P_i^{\mathrm{I}} = P_i^{\max}$ for at least one i, i.e., any local maximum of (5) touches the "outer" boundary of Ω_l.

Using (8), we propose the following sets of iterations. Let

$$\begin{aligned} \tilde{P}_i(n+1) &= \frac{1}{\displaystyle\sum_{j \neq i} \frac{G_{ji}}{I_j(\boldsymbol{P}(n))} \cdot \frac{\gamma_j(\boldsymbol{P}(n))}{(1+k\gamma_j(\boldsymbol{P}(n)))} \cdot \frac{(1+k\gamma_i(\boldsymbol{P}(n)))}{\gamma_i(\boldsymbol{P}(n))}} \\ &\equiv X_i(\boldsymbol{P}(n)) \end{aligned} \tag{9}$$

where $\boldsymbol{P}(n)$ denotes the transmission power vector at the nth iteration, and the functions $I_j(\boldsymbol{P})$ and $\gamma_j(\boldsymbol{P})$ are defined in (1). The transmission power of user i at the $(n+1)$th iteration is given by

$$P_i(n+1) = \begin{cases} P_i^{\min}, & \text{if } X_i(\boldsymbol{P}(n)) < P_i^{\min} \\ P_i^{\max}, & \text{if } X_i(\boldsymbol{P}(n)) > P_i^{\max} \\ X_i(\boldsymbol{P}(n)), & \text{otherwise.} \end{cases} \tag{10}$$

To simplify the notation, define a vector function $\boldsymbol{Z}^{\mathbf{I}}(\boldsymbol{P}) = [Z_1^{\mathrm{I}}(\boldsymbol{P}) \quad Z_2^{\mathrm{I}}(\boldsymbol{P}) \quad \cdots \quad Z_N^{\mathrm{I}}(\boldsymbol{P})]^T$, where $Z_i^{\mathbf{I}}(\boldsymbol{P})$ denotes the right-hand side of (10). Thus, the overall iteration can be represented as

$$\boldsymbol{P}(n+1) = \boldsymbol{Z}^{\mathbf{I}}(\boldsymbol{P}(n)). \tag{11}$$

2) Convergence of Iterative Algorithm I: A vector $\boldsymbol{Q} \in \Omega_l$ is a *fixed point* of the iteration if $\boldsymbol{Q} = \boldsymbol{Z}^{\mathbf{I}}(\boldsymbol{Q})$. We show next that the set of fixed points of these iterations is also the set of local maxima of the function $T(\boldsymbol{P})$.

Lemma 3: A vector $\boldsymbol{Q} \in \Omega_l$ is a fixed point of the iteration defined by (11) *if and only if* $T(\boldsymbol{Q})$ is a local maximum of the objective function $T(\boldsymbol{P})$.

Proof: We prove only the *if* part. The proof for the converse is similar.

Assume that $T(\boldsymbol{P})$ achieves a local maximum at $\boldsymbol{Q}$. It follows from (6) that $\nabla T(\boldsymbol{Q}) \cdot \boldsymbol{d} \leq 0$ for any feasible direction $\boldsymbol{d}$ at $\boldsymbol{Q}$. Consider a direction along axis i. Clearly, $P_i^{\min} \leq Q_i \leq P_i^{\max}$. This can be split into three cases.

1) $P_i^{\min} < Q_i < P_i^{\max}$: As Q_i is an "interior point," it satisfies (8), i.e.,

$$\frac{k\gamma_i}{Q_i(1+k\gamma_i)} - \sum_{j \neq i} \frac{k\gamma_j}{(1+k\gamma_j)} \cdot \frac{G_{ji}}{I_j} = 0$$

where I_j is evaluated at $\boldsymbol{Q}$. Rearranging this equation, and comparing to the definition of $\boldsymbol{Z}^{\mathbf{I}}(\boldsymbol{P})$, we find that $Z_i^{\mathrm{I}}(\boldsymbol{Q}) = Q_i$.

2) $Q_i = P_i^{\min}$: In this case, the only feasible direction is an increase along the i-axis, i.e., $d_i > 0$. It follows that

$\partial T(\boldsymbol{Q})/\partial Q_i \leq 0$. Therefore, from (7), we get

$$T(\boldsymbol{Q})\left[\frac{k\gamma_i}{Q_i(1+k\gamma_i)} - \sum_{j\neq i}\frac{k\gamma_j}{(1+k\gamma_j)}\cdot\frac{G_{ji}}{I_j}\right] \leq 0.$$

Noting that $T(\boldsymbol{Q})$ should be greater than zero, and comparing the part in square-brackets to the definition of $\boldsymbol{Z}^{\mathrm{I}}(\boldsymbol{P})$, we find that $Z_i^{\mathrm{I}}(\boldsymbol{Q}) = P_i^{\min} = Q_i$.

3) $Q_i = P_i^{\max}$: This is very similar to the previous case, and may be verified similarly.

The above proof is true for any i. Therefore, if $T(\boldsymbol{P})$ achieves a local maximum at $\boldsymbol{Q}, \boldsymbol{Z}^{\mathbf{I}}(\boldsymbol{Q}) = \boldsymbol{Q}$, i.e., $\boldsymbol{Q}$ is a fixed point of the iteration. □

It may be shown numerically that function $T(\boldsymbol{P})$ may have more than one local maxima and that the values of the function at these maxima may be very different. Depending upon the initial vector, it is therefore possible for these iterations to converge to a highly suboptimal solution.

B. Objective II

In order to avoid the possibility of converging to a highly suboptimal solution, we consider a slightly different objective function from that considered in (5). For this so-called *Objective II*, we then propose iterations that are guaranteed to converge to its global optimum. The resulting solution also has other desirable features as will be clear shortly.

Let

$$O(\boldsymbol{P}) = \prod_i \gamma_i(\boldsymbol{P}).$$

We define a new optimization problem as follows:

$$\max_{\boldsymbol{P}} O(\boldsymbol{P}) \quad \text{s.t. } \boldsymbol{P} \in \Omega_l. \tag{12}$$

Let $\boldsymbol{P}^{\mathrm{II}}$ denote the transmission power vector that maximizes (12). Clearly, $\boldsymbol{P}^{\mathrm{II}}$ is not optimal for Objective I, or equivalently, it may not achieve the optimal throughput. However, due to the functional similarity of $O(\boldsymbol{P})$ and $T(\boldsymbol{P})$, the solution $\boldsymbol{P}^{II}$ provides close to optimal performance.

Before proposing the iterative algorithm, we state some results that characterize the new objective $O(\boldsymbol{P})$ and are useful in obtaining the solution. The following lemmas follow from the definition of $O(\boldsymbol{P})$ and from the observation that Ω_l is compact.

Lemma 4: The objective function $O(\boldsymbol{P})$ has the following properties over the space $\Omega = \{\boldsymbol{P}: \boldsymbol{P} \geq \boldsymbol{0}\}$ and therefore also over the set Ω_l.

1) $O(\boldsymbol{P}) \geq 0$.
2) $O(\boldsymbol{P})$ is continuous and differentiable.
3) $O(\boldsymbol{P}) = 0$, if $P_i = 0$ for any i.
4) $\forall \alpha > 1$, $O(\alpha\boldsymbol{P}) > O(\boldsymbol{P})$.

Lemma 5: The function $O(\boldsymbol{P})$ has a maximum on the set Ω_l.

1) Iterative Algorithm II: As for (5), it is difficult to obtain a closed-form solution to (12). We use an approach similar to that in Section III-A.1 to derive an iterative algorithm for the solution.

Because $\boldsymbol{P}^{II}$ optimizes (12), it follows that for any feasible direction $\boldsymbol{d}$ at $\boldsymbol{P}^{\mathrm{II}}$,

$$\nabla O(\boldsymbol{P}^{II})\cdot\boldsymbol{d} \leq 0 \tag{13}$$

with

$$\nabla O(\boldsymbol{P}) = \left[\frac{\partial O(\boldsymbol{P})}{\partial P_1} \quad \frac{\partial O(\boldsymbol{P})}{\partial P_2} \quad \cdots \quad \frac{\partial O(\boldsymbol{P})}{\partial P_N}\right].$$

It may be shown that

$$\frac{\partial O(\boldsymbol{P})}{\partial P_i} = O(\boldsymbol{P})\left[\frac{1}{P_i} - \sum_{j\neq i}\frac{G_{ji}}{I_j}\right]. \tag{14}$$

It follows from (13) that for any P_i^{II} that is in the interior of Ω_l, i.e., if $P_i^{\min} < P_i^{\mathrm{II}} < P_i^{\max}$,

$$\frac{1}{P_i^{\mathrm{II}}} - \sum_{j\neq i}\frac{G_{ji}}{I_j} = 0 \tag{15}$$

or equivalently

$$\sum_{j\neq i}\frac{G_{ji}P_i^{\mathrm{II}}}{I_j} = 1. \tag{16}$$

However, if $P_i^{\mathrm{II}} = P_i^{\min}$ or $P_i^{\mathrm{II}} = P_i^{\max}$, (16) need not be true.

Note that the term $(G_{ji}P_i/I_j)$ is the contribution of user i to the total interference at user j. If we ignore the noise term, i.e., $n_j = 0$, *independent* of the choice of $\boldsymbol{P}$, the cumulative sum of all these contributions by all the users is always equal to N, that is

$$\sum_i\sum_{j\neq i}\frac{G_{ji}P_i}{I_j} = N.$$

Therefore, the set of conditions implied by (16) imposes a type of balance on the system: the interference contributions of each user to all the other users have to sum to one. Hence, we call this set of equations the *interference balancing equations*.

However, because of the nonzero thermal noise, $n_i > 0$, and Lemma 4 4), the optimal solution $\boldsymbol{P}^{\mathrm{II}}$ touches the "outer" boundary of Ω_l, that is, $P_i^{\mathrm{II}} = P_i^{\max}$ for at least one i. As will be clear soon, it also follows that there is at least one value of i for which the left-hand side of (16) is strictly less than one. Nevertheless, we use these interference balancing equations in the proposed iterations.

Using (15), we propose the following sets of iterations. Let

$$\tilde{P}_i(n+1) = \frac{1}{\displaystyle\sum_{j\neq i}\frac{G_{ji}}{I_j(\boldsymbol{P}(n))}} \equiv Y_i(\boldsymbol{P}(n)) \tag{17}$$

where $\boldsymbol{P}(n)$ is the transmission power vector at the nth iteration, and the so-called *interference* function $\boldsymbol{Y}(\boldsymbol{P}) = [Y_1(\boldsymbol{P})\, Y_2(\boldsymbol{P}) \cdots Y_N(\boldsymbol{P})]^T$ is a vector function defined as above. Then, because of the transmission power limits

$$P_i(n+1) = \begin{cases} P_i^{\min}, & \text{if } Y_i(\boldsymbol{P}(n)) < P_i^{\min} \\ P_i^{\max}, & \text{if } Y_i(\boldsymbol{P}(n)) > P_i^{\max} \\ Y_i(\boldsymbol{P}(n)), & \text{otherwise.} \end{cases} \tag{18}$$

To simplify the notation, define a vector function $\boldsymbol{Z}^{\mathbf{II}}(\boldsymbol{P}) = [Z_1^{\mathrm{II}}(\boldsymbol{P}) \quad Z_2^{\mathrm{II}}(\boldsymbol{P}) \quad \cdots \quad Z_N^{\mathrm{II}}(\boldsymbol{P})]^T$, where $Z_i^{\mathbf{II}}(\boldsymbol{P})$ denotes the right-hand side of (18). Thus, the overall iteration can be represented as

$$\boldsymbol{P}(n+1) = \boldsymbol{Z}^{\mathbf{II}}(\boldsymbol{P}(n)). \tag{19}$$

2) Convergence of Iterative Algorithm II: Before we prove the convergence of the Iterative Algorithm II, we summarize some important properties of functions $\mathbf{Y}(\mathbf{P})$ and $\boldsymbol{Z}^{\mathbf{II}}(\boldsymbol{P})$. These properties may be verified from the definition of $\boldsymbol{Y}(\boldsymbol{P})$. For all vectors $\boldsymbol{P} \in \boldsymbol{\Omega}$,

1) *Positivity:* $\boldsymbol{Y}(\boldsymbol{P}) > \mathbf{0}$;
2) *Monotonicity:* If $\boldsymbol{P} \geq \boldsymbol{P}'$, then $\boldsymbol{Y}(\boldsymbol{P}) \geq \boldsymbol{Y}(\boldsymbol{P}')$;
3) *Scalability:* $\forall \alpha > 1, \alpha \boldsymbol{Y}(\boldsymbol{P}) > \boldsymbol{Y}(\alpha P)$.

An interference function that has the above properties is called a *standard* interference function [17]. Hence, $\boldsymbol{Y}(\boldsymbol{P})$ is standard. It follows from Theorems 7 and 8 of [17] that $\boldsymbol{Z}^{\mathbf{II}}(\boldsymbol{P})$, which is essentially $\boldsymbol{Y}(\boldsymbol{P})$ with maximum and minimum power constraints, is also a *standard* interference function. Iterations of the type defined in (19) that use standard interference functions have very desirable convergence properties. Using these properties, we show next that the iterations of (19) converge to the desired solution.

Lemma 6: A vector $\boldsymbol{Q} \in \Omega_l$ is a fixed point of the iteration defined by (19) if and only if $O(\boldsymbol{Q})$ is a local maximum of the objective function $O(\boldsymbol{P})$.

Proof: The proof is similar to that of Lemma 3.

Lemma 7: Starting from any initial vector $\boldsymbol{P}^0 \in \Omega$, the iterations defined by (19) converge to a *unique* fixed point in Ω_l.

Proof: This proof uses a set of results derived by Yates in [17]. For ease of reference, we summarize these results in Appendix B. Consider the iterations $\boldsymbol{P}(n+1) = \boldsymbol{Z}^{\mathbf{II}}(\boldsymbol{P}(n))$: we have discussed above that $\boldsymbol{Z}^{\mathbf{II}}(\boldsymbol{P})$ is a so-called standard interference function.

In addition, from Lemmas 5 and 6, we are assured of the existence of at least one fixed point of the function $\boldsymbol{Z}^{\mathbf{II}}(\boldsymbol{P})$, or in the notation of [17], $\boldsymbol{Z}^{\mathbf{II}}(\boldsymbol{P})$ is *feasible*.[2] It therefore follows from Theorem 2 of [17] that the iterations of (19) converge to a unique fixed point. □

We now state our basic theorem.

Theorem 1: Starting from any initial vector $\boldsymbol{P}^0 \in \Omega$, the iterations defined by $\boldsymbol{P}(n+1) = \boldsymbol{Z}^{\mathbf{II}}(\boldsymbol{P}(n))$ converge to a unique fixed point, $\boldsymbol{P}^{\mathbf{II}} \in \Omega_l$, which achieves the global maximum of $O(\boldsymbol{P})$.

Proof: From Lemma 5, we are assured of the existence of a global maximum of $O(\boldsymbol{P})$. From Lemma 6, we deduce that this global maximum is achieved at a fixed point. Lemma 7 assures us that the iterations defined in (19) converge to a *unique* fixed point. Since the global maximum is a fixed point, it must be the only fixed point. It follows therefore that all iterations must necessarily converge to the global maximum. □

IV. Numerical Results

We present performance results of four schemes: *adaptive modulation without power control*; *SINR-balancing power control*; *adaptive modulation with power control—Objective I*; and *adaptive modulation with power control—Objective II*. The last two schemes have been discussed in the previous section, whereas the first two schemes are simulated to provide a performance baseline. All four schemes include some form of adaptive modulation but differ in the resulting system complexity, as will be clear shortly. In every case, the performance measure of interest is the throughput, $T(\boldsymbol{P})$, which the scheme can provide. The general procedure for evaluating this throughput is the same for all schemes and may be summarized by the following steps.

1) Find the "optimal" power vector $\boldsymbol{P}^*$ for the given scheme.
2) Evaluate the SINR for each user at $\boldsymbol{P}^*$.
3) Choose the appropriate constellation size for each user, based upon the SINR.
4) Evaluate the resulting total throughput $T(\boldsymbol{P}^*)$.

The schemes differ in the optimal power vector and the technique used to find this power vector. Except for first scheme, i.e., the adaptive modulation without power control, finding this power vector is the most involved step, and hence largely determines the complexity of the scheme. We summarize next the technique used in each scheme to obtain this power vector.

A. Description of the Schemes

Adaptive Modulation Without Power Control: This scheme is the simplest to implement since it does not require any real-time power control as is necessary in the next three schemes. The transmission powers of all the users are fixed to the maximum permitted, i.e., $P_i = P_i^{\max}$. The modulation level of each user is then adjusted according to the received SINR. The receiver estimates the interference and therefore the appropriate modulation level, and feeds this information to the transmitter.

SINR-Balancing Power Control: As an additional performance baseline, we also evaluate the performance of SINR-balancing power control. This type of power control, which is usually called "Optimal Power Control," has been extensively studied in the literature [10], [11]. In our case, power control is used to ensure that all the cochannel users have the same SINR. The best possible constellation size is then chosen according to this achievable SINR. That is, the appropriate modulation is used on a user-set basis, rather than on a per-user basis. Note that this usually provides higher throughput than a real system with SINR-balancing power control and a *fixed* modulation.

In this study, we use a technique proposed by Zander [10] to evaluate the best achievable SINR for a given set of users. This technique uses the approximation that there is no thermal noise. Therefore, we simulate this scheme only for interference-limited scenarios, where ignoring the thermal noise may be justified. Appendix C provides a brief description of this method.

In practice, this power control may be implemented in a distributed manner [11]. However, repeated real-time measurements, feedback, and power adjustments are necessary. Therefore, implementing this would require significant mea-

[2] $\boldsymbol{Z}^{\mathrm{II}}(\boldsymbol{P})$ is said to be feasible if for some $\boldsymbol{P} \in \Omega$, $\boldsymbol{P} \geq \boldsymbol{Z}^{\mathrm{II}}(P)$.

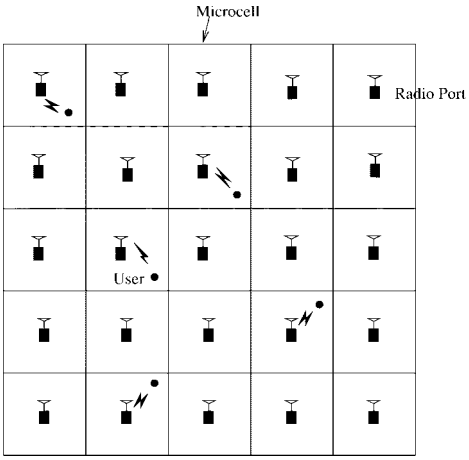


Fig. 1. The radio port arrangement of the system simulated.

surement and signaling overhead and increase the system complexity.

Adaptive Modulation with Power Control—Objective I: This refers to Iterative Algorithm I defined by (11). As discussed earlier, this iterative algorithm may converge to different fixed points, depending upon the initial power vector. For the results that follow, this initial vector is generated randomly, so that there is no guarantee that the resulting fixed point is the desired global optimum. Nevertheless, we find that with most starting vectors, the iterations converge to local maxima with throughputs close to the optimal. Therefore, on a statistical basis, these results are a good indicator of the optimal throughput.

Adaptive Modulation with Power Control—Objective II: This refers to Iterative Algorithm II defined by (19), which is guaranteed to converge to the global optimum of the objective $O(\boldsymbol{P})$, independent of the initial power vector. The resulting power vector $\boldsymbol{P}^{\mathbf{II}}$, while being the optimal solution of (12), may be suboptimal for maximizing the throughput. Nevertheless, its near optimal performance and assured global convergence make it desirable. As for the previous case, the initial power vector is generated randomly.

The latter two schemes are more complex than the SINR-balancing power control and therefore even more difficult to implement. As for SINR-balancing power control, repeated real-time measurements, feedback, and power adjustments are necessary. However, as opposed to SINR-balancing power control, some of these measurements may be very difficult to obtain. For example, the value of G_{ji}, which is necessary for updating the power of user i, cannot be inferred based on measurements at user i. In contrast, SINR-balancing power control *can* be implemented based upon local SINR measurements at each user [11]. Nevertheless, these schemes provide the performance limits of adaptive modulation in a multi-user environment. A simple, distributed implementation of adaptive modulation with a suitable power control is a good area for study.

B. Simulation Settings

Fig. 1 shows the system that is simulated. The system models a microcellular environment with a square grid of 25 radio ports. The distance between a radio port and each of its closest four neighbors is 400 m. A cell wraparound technique is used to avoid edge effects. A frequency reuse of one is assumed, i.e., each channel can be reused in every radio port. One channel is simulated, which may represent a certain frequency in a frequency division multiplex (FDM) system or a specific time-slot and frequency in an FDM/TDMA (time division multiple access) system. In an FDM/TDMA system, slot synchronization is assumed and propagation delays are ignored. Each cell can use this channel to communicate with one user on the downlink. N denotes the number of cochannel users. These N users are uniformly distributed in the system service area. Different choices of N represent different *loading* scenarios, with $N/25$ defined as the specific measure of the system load. The nominal setting for N is eight, so that the nominal system load is 0.32. *Site selection* is assumed, i.e., each user communicates with the strongest radio port. Omnidirectional antennas are assumed, both at the radio ports and at the user terminals. Unless stated otherwise, the minimum and maximum transmit powers $P_i^{\min}$ and $P_i^{\max}$ are set to -10 and 20 dBm, respectively.

The propagation model assumes operation at microwave frequencies in a suburban residential environment. It captures the effects of median path loss and shadowing [18]. The median path loss $L(r)$ is given by [19]

$$L(r)\,(\mathrm{dB}) = L(r_0) + 10 * \alpha \log \frac{r}{r_0} \tag{20}$$

where r_0 is a close-in reference distance and r is the distance in meters between the transmitter and the receiver. The path loss exponent α is set to 4.5. We set $r_0 = 10$ m, and $L(r_0)$ is *artificially* set to 0 dB, i.e., *all path loss and signal measurements are relative to* $L(r_0)$. It follows that the median path loss between a port and the corner of its nominal square coverage area is 65 dB.

Shadow fading is modeled as the sum of two independent log-normal components [20]: a user-location-specific component and a path-specific component. The user-location-specific component is the same for all radio paths to or from a given user. The path-specific components are different and independent for all radio paths. The user-location-specific component has a standard deviation of $\sigma_L = 6$ dB, and the path-specific component has a standard deviation of $\sigma_P = 8$ dB. This yields a combined shadow fading with $\sigma = \sqrt{\sigma_L^2 + \sigma_P^2} = 10$ dB. Some form of receiver diversity is assumed to mitigate the effect of multipath fading. Therefore, multipath fading is not modeled explicitly.

The user throughput in bits per symbol is estimated assuming M-QAM in an AWGN environment and for a target BER of 10^{-3}. As we are considering the performance with a finite number of cochannel users, the channel is not Gaussian. However, treating it as an AWGN channel usually corresponds to a worst case assumption. Therefore, we expect that the results obtained below are lower bounds. Two cases are considered: a continuous constellation size case and a discrete constellation size case.

TABLE I
CONSTELLATION SIZE, REQUIRED SINR, AND THROUGHPUT (BITS/SYMBOL)

M	2	4	8	16	32	64
Min. SINR (dB)	6	10	14	18	21	24
$T = \log_2 M$ (bits/sym.)	1	2	3	4	5	6

- *Continuous*: Equations (23) and (24) in Appendix A are used to estimate the throughput for each user. Hence, we permit the constellation size M to be any positive number.
- *Discrete*: The constellation size is restricted to a finite set. This set, and the corresponding required SINR's are shown in Table I. These SINR's are estimated from (23) and (24).

We first present results for an interference-limited system, where the thermal noise is set to a value much lower than the interference, and then study the impact of increasing the noise floor.

C. Interference-Limited Scenario

To simulate an interference-limited scenario, the noise floor is set to -80 dBm.[3] With the power set to $P_i^{\max} = 20$ dBm, the median signal level at the corner of the nominal coverage area of a port is equal to -45 dBm. Therefore, even without considering the beneficial effect of site selection, the median SNR at a cell corner is 35 dB.

1) An Illustrative Example: In order to gain some insight, we begin with a simple example. Consider a set of five users and the corresponding path gain matrix $\boldsymbol{G}$, generated using the simulation settings described earlier. For convenience, the path gain values are expressed in decibels

$$\boldsymbol{G}(\text{dB}) = \begin{bmatrix} -60.5 & -61.4 & -86.2 & -79.7 & -82.3 \\ -65.7 & -53.8 & -90.0 & -68.3 & -105.3 \\ -114.1 & -107.4 & -60.3 & -67.7 & -75.9 \\ -84.4 & -58.8 & -83.7 & -36.0 & -93.1 \\ -97.0 & -86.3 & -99.5 & -100.5 & -53.2 \end{bmatrix} \tag{21}$$

We compute the resulting throughput performance of the four schemes under consideration and summarize the results in Table II. The table shows the optimal power vector for each scheme, the resulting SINR for each user, and the total throughput. All power settings are in dBm's, and all SINR's are in decibels. The throughput is computed for both continuous and discrete constellation size cases. For the SINR-balancing scheme, the thermal noise and the maximum and minimum power limits are ignored. In all the other cases, both thermal noise and power limits are taken into consideration. A random initial power vector is used for both Iterative Algorithms I and II.

There are several points to note from this example.

1) The SINR-balancing scheme results in the lowest throughput, while both schemes that "optimally" combine the adaptive modulation with power control provide the highest throughput. The throughput of adaptive modulation without power control is somewhere in the middle, but still significantly better than that of the SINR-balancing scheme.
2) If discrete constellation sizes are considered, Iterative Algorithm II achieves the best total throughput performance (16 bits) for the given path gain matrix. Note that this need not be true with other path gain matrices.
3) There are two different local maxima to which Iterative Algorithm I converges, depending upon the initial vector. These are shown in Table II as Solutions 1 and 2. With random initial vectors, Iterative Algorithm I converges to Solution 1 around 95% of the time and to Solution 2 the remaining 5% of the time. It is unlikely, though possible, that there are other local maxima of (5) for the given gain matrix.
4) Iterative Algorithm I tends to achieve extreme power settings, i.e., it tends to assign a user either the maximum or the minimum power. On the other hand, Iterative Algorithm II tends to achieve more of a balance among users.

It is not surprising that the SINR-balancing scheme provides a lower throughput than using adaptive modulation even without any power control. By requiring the SINR performance to be equalized across users, we are forcing the scheme to raise the powers of the users that are more vulnerable to interference at the expense of the less vulnerable users. While this makes sense when all users use the same modulation and therefore require the same target SINR, it does not with adaptive modulation, which can adapt to the available SINR. For example, consider user 5. Without any power control, this user has an SINR equal to 32.2 dB. SINR balancing significantly reduces its transmission power in order to equalize the performance among users. On the other hand, the optimal schemes for both Objectives I and II do almost the opposite: they further amplify the SINR differences by allowing this user to transmit at the maximum power. This results in an overall increase in the throughput.

One difference between Objectives I and II is that the maxima of Objective I tend to push the transmission power of some users to the minimum permitted. This is because the contribution of each user to Objective I, $\log_2(1 + k\gamma_i)$, is lower bounded by 0, while the contribution of each user to Objective II, $\log_2(\gamma_i)$, decreases rapidly as $\gamma_i \to 0$. Therefore, with Objective II, there is a large penalty in setting a user's power to the minimum permitted. In this sense, Objective I is more "greedy" than Objective II.

One limitation of both the proposed objectives is that neither of them restricts the SINR to be within a "useful" range. In other words, neither objective accounts for the maximum available constellation size. For instance, user 5 achieves an SINR of 41.5 dB in Objective II even though the highest level modulation only requires an SINR of 24 dB. This explains the marked difference between the throughput for continuous constellations and that for discrete constellations.

2) General Results: The above observations are based upon a specific set of users and the corresponding path gain matrix. In order to corroborate these observations and to quantify the performance in general, we repeat this procedure for 1000 different sets of users, with eight users per set, i.e., a system loading of 0.32. Each set of users is generated as before, and the path gain matrix is evaluated accordingly. We

[3]Note that all power levels are relative to $L(r_0)$.

TABLE II
TRANSMISSION POWERS AND RESULTING SINR'S FOR THE FOUR SCHEMES

	SINR-Balancing		No Pow. Cont.		Objective I				Objective II	
					Solution 1		Solution 2			
User	$\mathbf{P}^\star$	SINR	$\mathbf{P}$	SINR	$\mathbf{P}^I$	SINR	$\mathbf{P}^I$	SINR	$\mathbf{P}^{II}$	SINR
1	20.0	6.4	20.0	0.8	17.5	14.6	-10.0	-29.2	16.2	13.3
2	14.5	6.4	20.0	9.9	-10.0	-18.6	20.0	34.3	-0.6	-6.2
3	-2.8	6.4	20.0	6.8	-10.0	-23.2	20.0	15.6	20.0	11.5
4	-1.9	6.4	20.0	22.8	20.0	48.1	-10.0	-7.2	13.8	35.3
5	-11.0	6.4	20.0	32.2	20.0	41.9	20.0	32.7	20.0	41.5
$T(\mathbf{P}^\star)$ (Cont.)	**5.8**	-	**18.3**		-	**29.4**	-	**22.2**	-	**27.1**
$T(\mathbf{P}^\star)$ (Disc.)	**5**	-	**13**		-	**15**	-	**15**	-	**16**

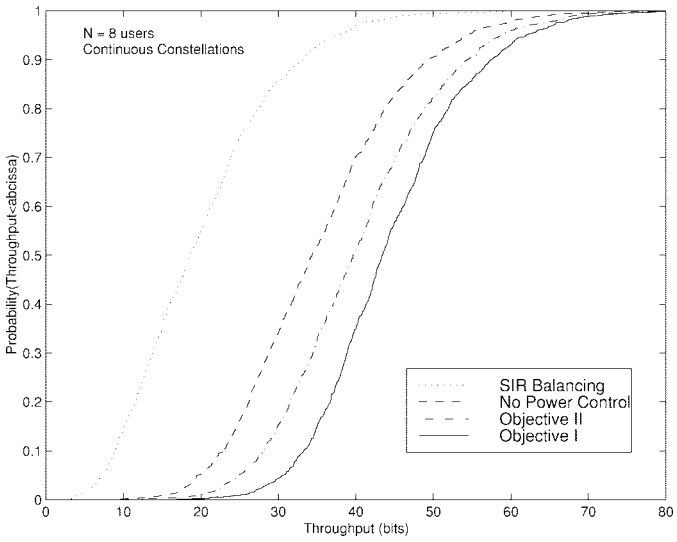


Fig. 2. Throughput comparison of the schemes assuming continuous constellations.

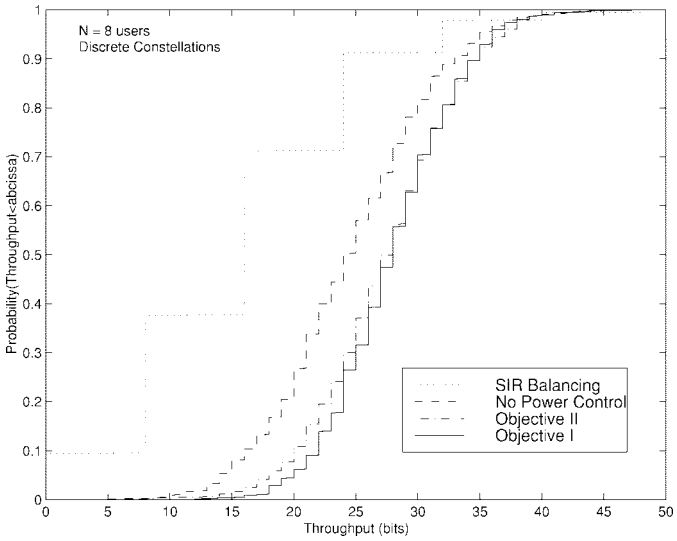


Fig. 3. Throughput comparison of the schemes for discrete M-QAM constellations.

again simulate an interference-limited scenario: the noise floor is set to -80 dBm; and the power limits are $P_i^{\max} = 20$ dBm and $P_i^{\min} = -10$ dBm. Figs. 2–4 summarize the results.

Figs. 2 and 3 show the cumulative distribution functions (CDF's) of the total throughput for the continuous and discrete

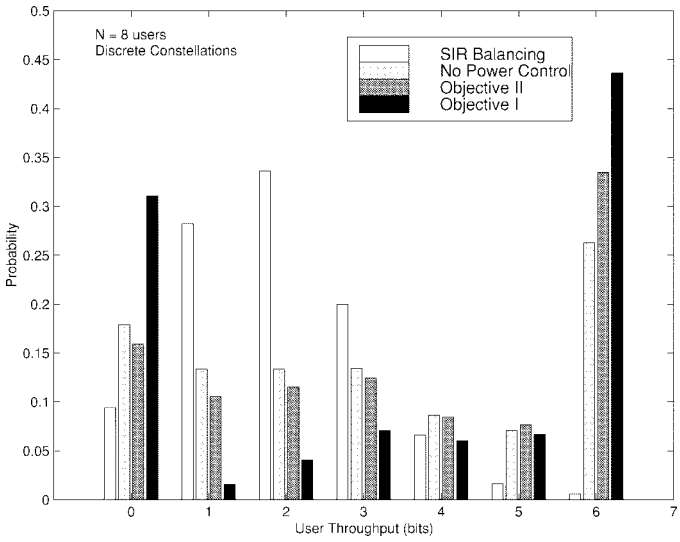


Fig. 4. Histograms of the user-throughputs of the schemes.

constellation size cases, respectively. Both figures show the same trend.

1) The SINR-balancing power control achieves the lowest throughput.
2) Using adaptive modulation without any power control provides significantly higher throughput than the SINR-balancing scheme.
3) Further throughput improvement can be obtained by combining the adaptive modulation with power control.

In particular, for the discrete constellation size case, the median value of the throughput is around 16 bits with SINR-balancing power control, around 24 bits using adaptive modulation without power control, and around 28 bits when adaptive modulation and power control are suitably combined, i.e., using the schemes corresponding to Objectives I or II.

As discussed earlier, the SINR-balancing scheme does include adaptive modulation, except that the adaptation is done on a user-set basis, rather than on a per-user basis. A system with fixed modulation would provide an even lower throughput than any of these schemes. This throughput may also be estimated from the figure. For example, if the modulation is fixed at 4-QAM, the figure shows that 38% of users would not be able to achieve the required SINR.

Fig. 4 shows the probability distribution of the throughput experienced by individual users. The four schemes show markedly different patterns. With the SINR-balancing scheme,

TABLE III
SENSITIVITY OF THE AVERAGE THROUGHPUT TO LOAD VARIATIONS

Load	SINR-Balancing	No Pow. Cont.	Objective I	Objective II
0.32	15	25	28	28
0.48	12	29	36	34
0.64	8	32	43	38

most of the users achieve a throughput of 1–3 bits/symbol, and relatively few of them achieve 0 bits/symbol or more than 3 bits/symbol. With the other three schemes, more users can use a higher level modulation and achieve a better throughput. In the case of adaptive modulation without power control, the distribution decays slowly with increasing throughput, with a discrepancy in this trend at a throughput of 6 bits/symbol. The distribution corresponding to Objective II begins to show a bimodal behavior, and this is further amplified in the distribution corresponding to Objective I. For example, in the latter case, around 32% of users have zero throughput and 44% have a throughput of 6 bits/symbol. This "greedy" behavior of Objectives I and II is not necessarily unfair from a system point of view. Appropriate admission control and resource assignment policies can and should be used to ensure fairness in the system.

Table III lists the average throughputs of the four schemes as the system load varies from 0.32–0.64, i.e., the number of cochannel users N varies from 8 to 16. Discrete constellations are assumed to obtain the results. It can be seen that for each case, the average throughput increases from the SINR-balancing scheme to the adaptive modulation without power control scheme to the scheme corresponding to Objective II to the scheme corresponding to Objective I. The table also shows that the relative gain of adaptive modulation over SINR-balancing power control increases with the load, i.e., it is relatively more beneficial to implement adaptive modulation in a highly loaded system than in a lightly loaded system. In addition, the discrepancy between the throughput of Objectives I and II increases with increasing load: Objective I becomes more beneficial for higher loads.

Further, with increasing load, each scheme shows a different trend. The average throughput for the SINR-balancing scheme reduces as the load increases. This is the result of the increased interference. On the other hand, the average throughput of the other three schemes increases with increasing load. This increase is most pronounced for the scheme corresponding to Objective I.

D. Impact of Thermal Noise and Power Limits

In our numerical investigations so far, we have modeled an interference-limited system. In this section, we study the impact on performance of: 1) increasing the thermal noise and 2) reducing the dynamic range of the transmit power. As discussed earlier, it is difficult to evaluate the performance of the SINR-balancing scheme in the presence of thermal noise. Therefore, we present performance results for the remaining three schemes.

Fig. 5 shows the performance of these three schemes as the system goes from an interference-limited system to a noise-limited system. This is achieved by increasing the thermal noise floor from -80 to -40 dBm. The maximum and minimum power limits are fixed at 20 and -10 dBm, respectively. Therefore, without considering the effect of site-selection, the median SNR at the corner of a cell is varied from 35 to -5 dB. The figure plots the average throughput versus the thermal noise floor for a load of 0.32, i.e., for $N = 8$.

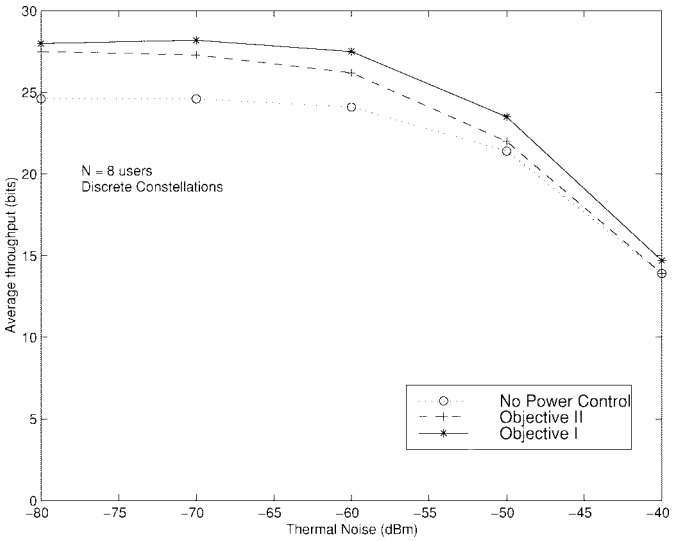


Fig. 5. Average throughput of the schemes in interference- and noise-limited environments.

From the figure, we see that the average throughput of the schemes stays essentially unchanged until the thermal noise increases beyond -60 dBm, i.e., the median SNR at the boundary decreases below 15 dB. As the thermal noise is increased further, the throughput of all the schemes suffers, and the gap between throughputs of different schemes narrows as well. This is not surprising because, as the system becomes more noise-limited, all these schemes would tend to set the transmission power for each user to the maximum permitted, $P_i^{\max}$.

Fig. 6 shows the performance of these schemes as the dynamic range of transmission powers varies. This is achieved by varying $P_i^{\min}$ from -10 to 20 dBm, while $P_i^{\max}$ stays fixed at 20 dBm. The figure shows the average throughput as a function of $P_i^{\min}$. Clearly, the throughput of adaptive modulation without power control should stay unchanged, as shown in the figure. The throughput of the other two schemes is essentially unaffected for $-10 \leq P_i^{\min} \leq 0$ dBm. However, as $P_i^{\min}$ is increased further, the throughput begins to decrease to that of the no power control case. We conclude that a 20-dB transmission power range, which is not unrealistic, can lead to a significant throughput improvement.

V. CONCLUSIONS

In this paper, we propose a general framework to study the performance of adaptive modulation in cellular systems. Specifically, we study the throughput performance gain that may be achieved by combining adaptive modulation and power control. We propose two sets of iterative algorithms: the first converges to the optimum solution for a set of initial vectors, but may converge to suboptimal solutions with

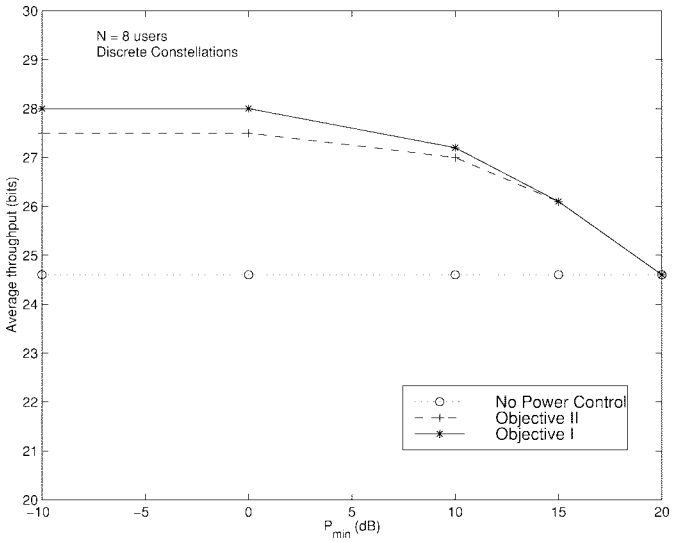


Fig. 6. Average throughput of the schemes versus the dynamic range of transmission power.

some other starting vectors; the second always converges to a close-to-optimal solution. We characterize these optimal and close-to-optimal solutions. This work provides valuable insight into the performance of adaptive modulation in multi-user environments.

Our results show that: 1) using adaptive modulation even without any power control provides a significant throughput advantage over using SINR-balancing power control; 2) combining adaptive modulation and power control leads to a significantly higher throughput as compared to using no power control or using SINR-balancing power control; and 3) most of the throughput gains may be realized with a modest transmission power range.

We show that this framework is valid in both AWGN and Rayleigh fading environments. Hence, it may be applied to both fixed and mobile wireless systems. In addition, although we pose this work in the context of adaptive modulation, it applies equally well to other physical layer, variable rate techniques. Some examples are coded-modulation or rate-adaptive coding schemes.

As regards implementation, using adaptive modulation without power control appears to be the most promising. It provides a throughput that is close to optimal and certainly higher than that provided by SINR-balancing power control. Furthermore, it is also the simplest one to implement among all the schemes considered. As discussed earlier, the repeated real-time measurements and iterations necessary make any of the other schemes difficult to implement. In this sense, the results show that adaptive modulation obviates the need for real-time power control while providing a significant throughput gain.

In practice, just selecting the appropriate modulation even without any power control may require multiple SINR measurements and modulation adjustments. This is particularly true in packet-switched systems, where the interference scenario changes rapidly. Therefore, how to adapt the modulation in a fast changing environment is an interesting topic for research.

Appendix A
Throughput of M-QAM in AWGN and Rayleigh Fading Channels

AWGN Channels: Consider the performance of variable-rate M-QAM modulation in an AWGN channel. In [21], the BER of M-QAM modulation in an AWGN channel is shown to be well approximated by

$$\text{BER} \approx 0.2 \exp^{[-1.5\gamma/(M-1)]} \tag{22}$$

for $0 \le \gamma \le 30$ dB, where γ is the SINR. In particular, for $M = 2$, (22) is a lower bound to the exact BER; and for $M \ge 4$, (22) is an upper bound. Rearranging the above equation, and assuming equality in place of the inequality, we get

$$M(\gamma) = 1 + \frac{1.5}{-\ln(5\text{BER})}\gamma = 1 + k\gamma \tag{23}$$

where $k = 1.5/(-\ln(5\text{BER}))$ is a constant for a specific BER requirement. With this approximation, the throughput of user i is

$$T_i(\gamma_i) = \log_2 M(\gamma_i) = \log_2(1 + k\gamma_i). \tag{24}$$

In practice, M is restricted to some specific integers (e.g., two, four, eight, etc.).

Rayleigh Fading Channels: We derive below the performance of variable-rate M-QAM modulation in a Rayleigh fading channel. Let $\overline{\gamma}$ and $\overline{\text{BER}}$ denote the average SINR and BER, respectively. The probability density function of the instantaneous SINR γ is given by

$$p(\gamma) = \frac{1}{\overline{\gamma}} \exp^{-\gamma/\overline{\gamma}}, \qquad \gamma \ge 0.$$

Using (22), we have

$$\begin{aligned} \overline{\text{BER}} &\approx \int_0^\infty c_1 \exp^{[-c_2\gamma/(M-1)]} \frac{1}{\overline{\gamma}} \exp^{-\gamma/\overline{\gamma}} \, d\gamma \\ &= \frac{c_1}{\left(1 + \dfrac{c_2\overline{\gamma}}{M-1}\right)} \end{aligned} \tag{25}$$

where $c_1 = 0.2$ and $c_2 = 1.5$. Rearranging the above equation, and assuming equality in place of the inequality, we get

$$M(\overline{\gamma}) = 1 + \frac{c_2\overline{\gamma}}{\left(\dfrac{c_1}{\overline{\text{BER}}} - 1\right)} = 1 + c_3\overline{\gamma} \tag{26}$$

where $c_3 = c_2/(c_1/\overline{\text{BER}} - 1)$.

Note the similarity of (26) and (23). Of course, for a given target BER, there is a large difference between the values of k and c_3. For example, setting $\overline{\text{BER}} = 0.01$, (26) reduces to $M(\overline{\gamma}) = 1 + 0.08\overline{\gamma}$. Similarly, setting BER $= 0.01$, (23) reduces to $M(\gamma) = 1 + 0.5\gamma$. Nevertheless, it follows that (2) is also a good approximation for a Rayleigh fading channel.

APPENDIX B
USEFUL RESULTS FOR CONVERGENCE PROOF

Yates proposed a framework for uplink power control in cellular systems in [17]. In order to prove convergence of the iterations defined in (19), we use some of these results. For ease of reference, we state these results below.

Let $\boldsymbol{P}$ denote the transmission power vector and $\boldsymbol{I}(\boldsymbol{P})$ denote a so-called *interference function*. The interference function $\boldsymbol{I}(\boldsymbol{P})$ is standard if it has the properties listed in Section III-B.2. For a class of useful power control schemes and for an appropriately defined interference function, the interference constraints may be represented by

$$\boldsymbol{P} \geq \boldsymbol{I}(\boldsymbol{P}).$$

A so-called *standard power control algorithm* is then defined by the following iteration:

$$\boldsymbol{P}(\boldsymbol{n}+\boldsymbol{1}) = \boldsymbol{I}(\boldsymbol{P}(\boldsymbol{n})).$$

The following results can be shown to be true for these algorithms.

Theorem 1 in [17]: If the standard power control algorithm has a fixed point, then that fixed point is unique.

Lemma 1 in [17]: If $\boldsymbol{P}$ is a feasible power vector, then $\boldsymbol{I}^n(\boldsymbol{P})$ is a monotonically decreasing sequence of feasible power vectors that converges to a unique fixed point $\boldsymbol{P}^*$.

Lemma 2 in [17]: If $\boldsymbol{I}(\boldsymbol{P})$ is feasible, then starting from the all-zero vector $\boldsymbol{z}$, the standard power control algorithm produces a monotonically increasing sequence of power vectors $\boldsymbol{I}^n(\boldsymbol{z})$ that converges to the fixed point $\boldsymbol{P}^*$.

Theorem 2 in [17]: If $\boldsymbol{I}(\boldsymbol{P})$ is feasible, then for any initial power vector $\boldsymbol{P}$, the standard power control algorithm converges to a unique fixed point $\boldsymbol{P}^*$.

APPENDIX C
SINR-BALANCING POWER CONTROL

SINR-balancing power control, or so-called *optimal power control*, has been extensively studied in the literature [10], [11]. It has been shown to be very effective in reducing cochannel interference, and therefore increasing capacity. In particular, under appropriate conditions, it has the following desirable properties: i) it maximizes the minimum SINR among a given set of users; and ii) for a given target SINR, it minimizes the total transmission power of users.

Zander outlines a powerful technique to obtain the maximum SINR that may be achieved for a given set of users under the approximation that there is *no thermal noise* [10]. Let $\gamma^* = \max\{\gamma\colon \gamma_i \geq \gamma, \forall i\}$ denote the maximum balanced SINR that can be achieved by the given set of users, where γ_i is the SINR of user i. Given the path gain matrix $\boldsymbol{G}$, γ^* is shown to be given by

$$\gamma^* = \frac{1}{\lambda^* - 1} \tag{27}$$

where λ^* is the largest eigenvalue of a matrix that may be obtained from $\boldsymbol{G}$. The transmission power vector that achieves this SINR is the corresponding eigenvector.

ACKNOWLEDGMENT

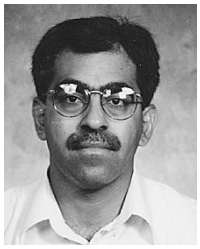
The authors would like to thank J. Chuang, M. Clark, L. Greenstein, and P. Henry for many helpful discussions and comments. They would also like to thank the Editor and the anonymous reviewers for their insightful comments that improved the paper significantly.

REFERENCES

[1] G. J. Pottie, "System design choices in personal communications," *IEEE Personal Commun.*, vol. 2, pp. 50–67, Oct. 1995.
[2] A. J. Goldsmith and S.-G. Chua, "Variable-rate variable-power MQAM for fading channels," *IEEE Trans. Commun.*, vol. 45, pp. 1218–1230, Oct. 1997.
[3] W. T. Webb and R. Steele, "Variable rate QAM for mobile radio," *IEEE Trans. Commun.*, vol. 43, pp. 2223–2230, July 1995.
[4] H. Matsuoka, S. Sampei, N. Morinaga, and Y. Kamio, "Adaptive modulation system with variable coding rate concatenated code for high-quality multi-media communication systems," in *Proc. IEEE VTC'96*, pp. 487–491.
[5] CCITT Draft Recommendation V.34: A Modem Operating at Data Signaling Rates of up to 28800 bits/s for Use on the General Switched Telephone Network and on Leased Point to Point 2-Wire Telephone-Type Circuits, 1993.
[6] Cable Television Laboratories, Inc. (1998). DOCSIS: Radio frequency interface specification: Interim Specification SP-RFI-104-980724. [Online]. Available HTTP: http://www.cablemodem.com/
[7] "IMT-2000: Standards efforts of the ITU," *IEEE Personal Commun.*, vol. 4, Aug. 1997.
[8] M.-S. Alouini and A. Goldsmith, "Area spectral efficiency of cellular mobile radio systems," in *Proc. IEEE VTC'97*, pp. 652–656.
[9] S. Sampei, S. Komaki, and N. Morinaga, "Adaptive modulation/TDMA scheme for personal multi-media communication systems," in *Proc. IEEE Globecom '94*, pp. 989–993.
[10] J. Zander, "Performance of optimum transmitter power control in cellular radio systems," *IEEE Trans. Veh. Technol.*, vol. 41, pp. 57–62, Feb. 1992.
[11] G. J. Foschini and Z. Miljanic, "A simple distributed autonomous power control algorithm and its convergence," *IEEE Trans. Veh. Technol.*, vol. 42, pp. 641–646, Nov. 1993.
[12] ETSI TDoc SMG2 95/97, *EDGE Feasibility Study. Work Item 184: Improved Data Rates through Optimized Modulation*, version 0.3, Dec. 1997.
[13] P. Schramm, H. Andreasson, C. Edholm, N. Edvardssona, M. Hook, S. Javerbring, F. Muller, and J. Skold, "Radio interface performance of EDGE, a proposal for enhanced data rates in existing digital cellular systems," in *Proc. IEEE VTC'98*, pp. 1064–1068.
[14] A. Furuskar, M. Frodigh, H. Olofsson, and J. Skold, "System performance of EDGE, a proposal for enhanced data rates in existing digital cellular systems," in *Proc. IEEE VTC'98*, pp. 1284–1289.
[15] D. G. Luenberger, *Optimization by Vector Space Methods.* New York: Wiley, 1969.
[16] W. Rudin, *Principles of Mathematical Analysis*. New York: McGraw-Hill, 1976.
[17] R. D. Yates, "A framework for uplink power control in cellular radio systems," *IEEE J. Select. Areas Commun.*, vol. 13, pp. 1341–1348, Sept. 1995.
[18] S. Ariyavisitakul, T. E. Darcie, L. J. Greenstein, M. R. Phillips, and N. K. Shankaranarayanan, "Performance of simulcast wireless techniques for personal communication systems," *IEEE J. Select. Areas Commun.*, vol. 14, pp. 632–643, May 1996.
[19] D. C. Cox, R. R. Murray, and A. W. Norris, "800-MHz attenuation measured in and around suburban houses," *AT&T Bell Lab. Tech. J.*, vol. 63, pp. 921–954, July–Aug. 1984.
[20] H. W. Arnold, D. C. Cox, and R. R. Murray, "Macroscopic diversity performance measured in the 800-MHz portable radio communications environment," *IEEE Trans. Antennas Propagat.*, vol. 36, pp. 277–280, Feb. 1988.
[21] M.-S. Alouini and A. Goldsmith, "Adaptive modulation over Nakagami fading channels," *Wireless Personal Commun.*, to be published.

Xiaoxin Qiu (M'93) received the B.E. and M.E. degrees from Tsinghua University, China, in 1990 and 1991, respectively, and the Ph.D. degree from University of Southern California, Los Angeles, in 1996.

She joined the Broadband Wireless Systems Research Department, AT&T Laboratories, Red Bank, NJ, in 1996. Her research interests are in the areas of wireless communications networks, personal communication systems, and multimedia communications.



Kapil Chawla (M'86–SM'98) received the B.Tech. degree from the Indian Institute of Technology, Kanpur, in 1984 and the M.S. and Ph.D. degrees from the University of Illinois at Urbana-Champaign in 1987 and 1991, all in electrical engineering.

He joined AT&T Bell Laboratories in 1990, where he worked on several aspects of wireless and wireline communications, including systems engineering, performance analysis, and network planning. He is currently with the Wireless Communications Research Department of AT&T Laboratories, Red Bank, NJ. His research interests are in the areas of resource assignment in wireless networks and of cellular system engineering.