

Variable-Rate Variable-Power MQAM for Fading Channels

Andrea J. Goldsmith, *Member, IEEE*, and Soon-Ghee Chua, *Associate Member, IEEE*

Abstract— We propose a variable-rate and variable-power MQAM modulation scheme for high-speed data transmission over fading channels. We first review results for the Shannon capacity of fading channels with channel side information, where capacity is achieved using adaptive transmission techniques. We then derive the spectral efficiency of our proposed modulation. We show that there is a constant power gap between the spectral efficiency of our proposed technique and the channel capacity, and this gap is a simple function of the required bit-error rate (BER). In addition, using just five or six different signal constellations, we achieve within 1–2 dB of the maximum efficiency using unrestricted constellation sets. We compute the rate at which the transmitter needs to update its power and rate as a function of the channel Doppler frequency for these constellation sets. We also obtain the exact efficiency loss for smaller constellation sets, which may be required if the transmitter adaptation rate is constrained by hardware limitations. Our modulation scheme exhibits a 5–10-dB power gain relative to variable-power fixed-rate transmission, and up to 20 dB of gain relative to nonadaptive transmission. We also determine the effect of channel estimation error and delay on the BER performance of our adaptive scheme. We conclude with a discussion of coding techniques and the relationship between our proposed modulation and Shannon capacity.

Index Terms— Adaptive modulation, fading channels, spectral efficiency.

I. INTRODUCTION

HIGH-SPEED wireless data transmission requires robust and spectrally efficient communication techniques for flat-fading channels. When the channel can be estimated and this estimate sent back to the transmitter, the transmission scheme can be adapted relative to the channel characteristics. Most modulation and coding techniques do not adapt to fading conditions. These nonadaptive methods require a fixed link margin to maintain acceptable performance when the channel quality is poor. Thus, these systems are effectively designed for the worst case channel conditions, resulting in insufficient utilization of the full channel capacity. Adapting to the signal fading allows the channel to be used more efficiently since

power and rate can be allocated to take advantage of favorable channel conditions. In [1], the optimal adaptive transmission scheme which achieves the Shannon capacity of a fading channel was derived. In this work, we develop practical variable-rate variable-power MQAM modulation techniques for fading channels inspired by the capacity results in [1].

Adaptive transmission, which requires accurate channel estimates at the receiver and a reliable feedback path between the receiver and transmitter, was first proposed in the late 1960's [3]. Interest in these techniques was short lived, perhaps due to hardware constraints, lack of good channel estimation techniques, and/or systems focusing on point-to-point radio links without transmitter feedback. The fact that these issues are less constraining in current systems, coupled with the growing demand for spectrally efficient communication, has revived interest in adaptive modulation methods. The basic idea behind adaptive transmission is to maintain a constant E_b/N_0 by varying the transmitted power level [3], symbol transmission rate [4], constellation size [5]–[7], coding rate/scheme [8], or any combination of these parameters [9]–[11]. Thus, without sacrificing bit-error rate (BER), these schemes provide high average spectral efficiency by transmitting at high speeds under favorable channel conditions, and reducing throughput as the channel degrades. The performance of these schemes is further improved by combining them with space diversity [12], [13]. Adaptive techniques are also used for high-speed modems [14], [15], satellite links [16]–[18], and to minimize distortion or satisfy quality-of-service requirements in end-to-end wireless applications [19], [20]. Our approach is novel relative to all of these adaptive techniques in that we optimize *both* the transmission rate and power to maximize spectral efficiency, while satisfying average power and BER constraints. Although we restrict ourselves to MQAM signal constellations, the same adaptive techniques can be applied to lattice-based constellations [21], which exhibit 1–1.5 dB of shaping gain relative to MQAM.

We first determine the spectral efficiency of our variable-rate variable-power MQAM technique and compare it with the Shannon capacity limit. We find that there is a constant power gap between these two spectral efficiencies which is a simple function of the target BER. We also compare the spectral efficiency of our proposed technique with that of two fixed-rate variable-power MQAM schemes using channel inversion and truncated channel inversion. Channel inversion adapts the transmit power to maintain a constant received SNR. This technique suffers a large power penalty since most of the average signal power is used to compensate for deep fades.

Paper approved by S. Roy, the Editor for Communication Theory/Systems of the IEEE Communications Society. Manuscript received March 17, 1996; revised November 20, 1996 and April 23, 1997. This work was supported by NSF CAREER Award NCR-9501452, by the Office of Naval Research under NAV-5X-N149510861, and by a Caltech Summer Undergraduate Research Fellowship (SURF). This paper was presented in part at the IEEE Vehicular Technology Conference, Atlanta GA, April 1996.

A. J. Goldsmith is with the the Department of Electrical Engineering, California Institute of Technology, Pasadena, CA 91125 USA.

S.-G. Chua is with Singapore Technologies Telecommunications Pte Ltd., Singapore 088934.

Publisher Item Identifier S 0090-6778(97)07278-4.

Truncated channel inversion maintains a constant received SNR unless the channel fading falls below a given cutoff level, at which point a signal outage is declared and no signal is sent. This technique has almost the same spectral efficiency as our adaptive modulation. However, the corresponding outage probability can be quite high. Thus, fixed-rate transmission with truncated channel inversion approximates a packet radio protocol, with bursts of high-speed data when the channel is favorable and idle times in between. Nonadaptive modulation exhibits up to 20 dB of power loss relative to these adaptive techniques.

There are several practical constraints which determine when adaptive modulation should be used. If the channel is changing faster than it can be estimated and fed back to the transmitter, adaptive techniques will perform poorly, and other means of mitigating the effects of fading should be used. In Section VII, we find that, for a target BER of 10^{-6} , the BER remains at its target level as long as the total delay of the channel estimator and feedback path is less than $0.001\lambda/v$, where v is the vehicle speed and λ is the signal wavelength. Thus, for a 900-MHz signal, at pedestrian speeds of 3.6 km/h, the total delay should not exceed 0.3 ms, and at vehicle speeds of 90 km/h, the total delay should not exceed 13.3 μ s. The former constraint is within the capabilities of existing estimation techniques and feedback channels, while the latter constraint is more challenging. However, a higher BER target loosens the delay constraint: at 10^{-3} BER, a total delay constraint of less than $0.01\lambda/v$ suffices for good performance. The effects of estimation error are also characterized in Section VII where we find that, assuming perfect AGC, the estimation error must be less than 1 dB to maintain the target BER. In Rayleigh fading, this bound on estimation error can be achieved using the pilot-symbol-assisted estimation technique described in [25] with appropriate choice of parameters.¹ Finally, hardware constraints and pulse-shaping considerations may dictate how often the transmitter can change its rate and/or power. In Section V, we derive a closed-form expression for how often the transmitter must adapt its signal constellation as a function of the Doppler frequency $f_D = v/\lambda$. We find that even at high Doppler frequencies ($f_D = 100$ Hz), for symbol rates of 100 ksymbols/s, the constellation remains constant over tens–hundreds of symbols.

Cellular systems exploit the power falloff with distance of signal propagation to reuse the same frequency channel at spatially separated locations. While frequency reuse provides more efficient use of the limited available spectrum within a given area, it also introduces cochannel interference, which ultimately determines the data rates and corresponding BER's available to each user. Thus, although adaptive modulation techniques increase the spectral efficiency (bits/s/Hz) of a single channel, these techniques may also increase cochannel interference levels in a cellular system, thereby requiring a higher reuse distance to mitigate this increased interference power. Adaptive modulation may therefore reduce the *area*

*spectral efficiency*² of a cellular system, defined as its average bits/s/Hz/km². Indeed, while we show in this paper that channel inversion can significantly reduce the spectral efficiency of a single user relative to optimal adaptation, this inversion is necessary in CDMA cellular systems without multiuser detection to reduce the near–far effect [22], [23]. The area spectral efficiency of FDMA/TDMA cellular systems with the adaptive policies described in this paper are analyzed in [24], where it is shown that power adaptation typically reduces area spectral efficiency, while rate adaptation improves it. We do not consider the effect of cochannel interference in our analysis below. Thus, our results apply to systems without frequency reuse, or to cellular systems where the cochannel interference is significantly mitigated through cell isolation, sectorization, or adaptive antennas.

The remainder of this paper is organized as follows. In the next section, we present the system model. In Section III, we review the Shannon capacity of our fading channel model under different power adaptation strategies. We describe our variable-rate variable-power MQAM modulation technique in Section IV, and optimize its rate and power adaptation to maximize spectral efficiency. The spectral efficiency for restricted constellation sets is obtained in Section V, where we also determine the rate of transmitter adaptation as a function of the channel Doppler frequency. Our simulation results for BER and spectral efficiency are presented in Section VI, along with the spectral efficiency of suboptimal adaptive techniques. We derive expressions for the increase in BER as a function of channel estimation error and the estimator and feedback path delay in Section VIII. We conclude with a discussion of coding techniques for our adaptive policy, and the relationship between our proposed technique and Shannon capacity.

II. SYSTEM MODEL

Consider a discrete-time channel with stationary and ergodic time-varying gain $\sqrt{g[i]}$ and additive white Gaussian noise $n[i]$. Let \bar{S} denote the average transmit signal power, $N_0/2$ denote the noise density of $n[i]$, B denote the received signal bandwidth, and \bar{g} denote the average channel gain. With appropriate scaling of \bar{S} , we can assume that $\bar{g} = 1$. For a constant transmit power \bar{S} , the instantaneous received SNR is $\gamma[i] = \bar{S}g[i]/(N_0B)$ and the average received SNR is $\bar{\gamma} = \bar{S}/(N_0B)$. Suppose, however, that we adapt the transmit power at time i based on the channel estimate $\hat{g}[i]$ or, equivalently, on $\hat{\gamma}[i] = \bar{S}\hat{g}[i]/(N_0B)$. We denote the transmit power at time i with this adaptive scheme by $S(\hat{\gamma}[i])$, and the received power at time i is then $\gamma[i](S(\hat{\gamma}[i])/\bar{S})$. Since $g[i]$ is stationary, the distribution of $\gamma[i]$ is independent of i , and we denote this distribution by $p(\gamma)$. When the context is clear, we will omit the time reference i relative to γ and $S(\gamma)$.

The system model is illustrated in Fig. 1. We assume that an estimate $\hat{g}[i]$ of the channel power gain $g[i]$ at time i is available to the receiver after an estimation time delay of τ_e and that this same estimate is available to the transmitter after

¹There will be some loss of spectral efficiency for this estimation technique since the pilot symbol rate must be subtracted from the transmitted symbol rate.

²Unfortunately, the area spectral efficiency is often referred to as just spectral efficiency, which causes some confusion between the two definitions. In this paper, spectral efficiency refers to the bits/s/Hz of a single-user channel.

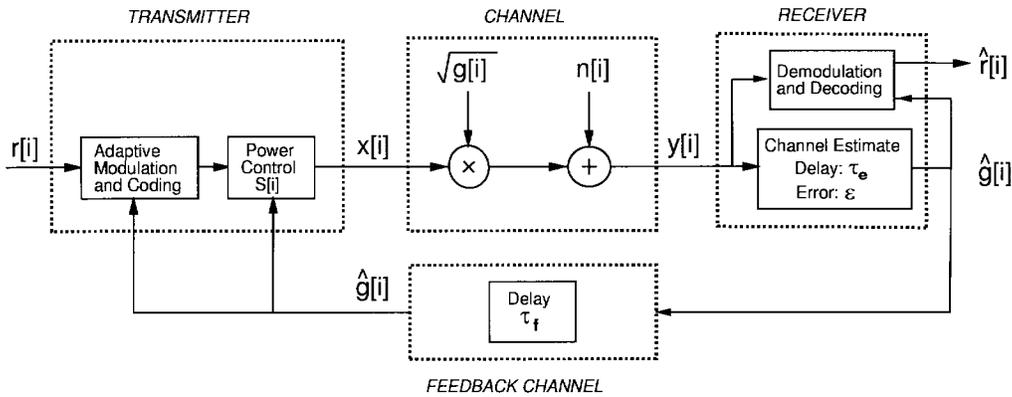


Fig. 1. System model.

a combined estimation and feedback path delay of $\tau = \tau_e + \tau_f$. We also assume ideal coherent phase detection. The channel gain estimation error $\epsilon[i]$ is defined as $\epsilon[i] = \hat{g}[i]/g[i] = \hat{\gamma}[i]/\gamma[i]$. We assume that the feedback path does not introduce any errors, which can be assured by increasing its delay time and using an ARQ transmission protocol. The availability of channel information at the transmitter allows it to adapt its transmission scheme relative to the channel variation. In Sections III–VI, we ignore the effects of estimation error and delay, assuming $\epsilon = 1$ and $\tau = 0$. We relax these assumptions in Section VII, and determine closed-form expressions for the increase in BER resulting from these effects.

We will assume $p(\gamma)$ to be either lognormal or exponential (Rayleigh fading) in the numerical calculations below, although our formulas apply for any distribution of γ . The lognormal distribution arises from attenuation of the transmitted signal by surrounding buildings, and the exponential distribution arises from multipath [26]. Although both types of fading will typically be superimposed on the received signal, we consider the two distributions separately for the following reasons. At low speeds, the lognormal shadowing is essentially constant, and the Rayleigh fading is sufficiently slow so that it can be estimated and fed back to the transmitter within the estimation error and delay constraints outlined in Section VII. At high speeds, these constraints may not be met. In this case, most of the Rayleigh fading can be removed with a sufficient number of diversity branches at the transmitter or receiver, in which case the adaptive modulation need only respond to lognormal channel variations.

The rate of channel variation will dictate how often the transmitter must adapt its rate and/or power, and will also impact the BER increase due to estimation error and delay. For Rayleigh fading, we assume the standard Jakes model for the autocorrelation of the channel power gain over time [26]:

$$A_g(\tau) = J_0^2(2\pi v\tau/\lambda) \quad (1)$$

where v is the mobile user's velocity and λ is the RF wavelength.

The autocorrelation function for lognormal shadowing is not well characterized. However, measurements reported in [27] support an autoregressive model:

$$A_g(\tau) = e^{-v|\tau|/X_c} \quad (2)$$

where X_c is the effective autocorrelation distance of the lognormal shadowing. This distance is on the order of 10–100 m, depending on propagation distance [28].

III. CHANNEL CAPACITY

The capacity of both fading and AWGN channels is limited by the available transmit power and bandwidth. Let $S(\gamma)$ denote the transmit power adaptation policy relative to an instantaneous value of γ , subject to the average power constraint

$$\int_0^\infty S(\gamma)p(\gamma) d\gamma \leq \bar{S}. \quad (3)$$

The capacity of a fading channel with bandwidth B and average power \bar{S} under the assumptions outlined in Section II is derived in [1] to be

$$C = \max_{S(\gamma): \int S(\gamma)p(\gamma) d\gamma = \bar{S}} \int_0^\infty B \log_2 \left(1 + \frac{S(\gamma)\gamma}{\bar{S}} \right) p(\gamma) d\gamma. \quad (4)$$

The power adaptation which maximizes (4) is

$$\frac{S(\gamma)}{\bar{S}} = \begin{cases} \frac{1}{\gamma_0} - \frac{1}{\gamma}, & \gamma \geq \gamma_0 \\ 0, & \gamma < \gamma_0 \end{cases} \quad (5)$$

for some ‘‘cutoff’’ value γ_0 . If $\gamma[i] < \gamma_0$ at time i , then no power is allocated to the i th data transmission. Since γ is time varying, the maximizing transmit power distribution (5) is a ‘‘water-filling’’ formula in time that depends on the fading statistics $p(\gamma)$ only through the cutoff value γ_0 . In particular, when the channel is favorable (γ large), more power is allocated for transmission. Conversely, when the channel is not as good, less power will be transmitted. If the channel quality drops below γ_0 , the channel is not used on that transmission.

Substituting (5) into (3), we see that γ_0 is determined by numerically solving

$$\int_{\gamma_0}^\infty \left(\frac{1}{\gamma_0} - \frac{1}{\gamma} \right) p(\gamma) d\gamma = 1. \quad (6)$$

For some $p(\gamma)$ distributions, a closed-form expression for γ_0 may be found. Once γ_0 is known, we substitute (5) into (4)

to get

$$C = \int_{\gamma_0}^{\infty} B \log_2 \left(\frac{\gamma}{\gamma_0} \right) p(\gamma) d\gamma. \quad (7)$$

The maximum spectral efficiency is defined as C/B , the channel capacity per unit hertz, and is obtained by dividing both sides of (7) by the channel bandwidth B .

The modulation and coding strategy which achieves this capacity is a multiplexing technique whereby the coding and modulation transmitted over the channel with a fade level of $\gamma[i] = \gamma$ are optimized for that fade level. The resulting transmission scheme is both variable power and variable rate. In [1], we show that this multiplexing strategy achieves the channel capacity (7), and that no other transmission method achieves a higher rate with arbitrarily small BER.

Note that for a constant transmit power $S(\gamma) = \bar{S}$, the capacity of (4) reduces to

$$C = \int_0^{\infty} B \log_2(1 + \gamma) p(\gamma) d\gamma \quad (8)$$

which was previously reported by Lee as the average channel capacity [29]. In fact, it is shown in [1] that (8) is the Shannon capacity of the fading channel when the transmitter adapts to the channel variation using a constant-power variable-rate strategy. Surprisingly, the difference between (4) and (8) is a small fraction of a decibel for most types of fading [1]. This negligible impact of power adaptation results from the fact that the capacity formula (4) places no restriction on the input signal constellation. Since we can vary the rate continuously, varying the power as well has little impact. We will see in Section V that when the constellation size is restricted to square constellations (an even number of bits/symbol), restricting the power adaptation as well has a greater impact on spectral efficiency.

A simpler form of power adaptation is total channel inversion, where the transmitter adjusts its power to maintain a constant received power. The power control policy for channel inversion is

$$\frac{S(\gamma)}{\bar{S}} = \frac{\sigma}{\gamma} \quad (9)$$

where σ equals the received SNR that can be maintained subject to the average power constraint \bar{S} . Thus, σ must satisfy

$$\int \frac{\sigma}{\gamma} p(\gamma) = 1 \implies \sigma = \frac{1}{\overline{[1/\gamma]}} \quad (10)$$

where \bar{x} denotes the expected, or average, value of x .

The channel capacity with this power adaptation strategy is derived from the capacity of an AWGN channel with a received SNR of σ :

$$C = B \log_2[1 + \sigma] = B \log_2 \left[1 + \frac{1}{\overline{[1/\gamma]}} \right]. \quad (11)$$

This form of power adaptation greatly simplifies the coding and modulation for the fading channel since the channel with inversion appears to the encoder and decoder as an AWGN channel, independent of the fading statistics. However, with channel inversion, much of the transmit power is used to

compensate for deep fading. In fact, the capacity with channel inversion (11) in Rayleigh fading is zero [1].

A better form of power control is truncated channel inversion, which only compensates for fading above a certain cutoff fade depth γ_0 :

$$\frac{S(\gamma)}{S} = \begin{cases} \frac{\sigma_0}{\gamma}, & \gamma \geq \gamma_0 \\ 0, & \gamma < \gamma_0. \end{cases} \quad (12)$$

Since the channel is only used when $\gamma \geq \gamma_0$, the average power constraint \bar{S} yields $\sigma_0 = 1/\overline{[1/\gamma]}_{\gamma_0}$, where

$$\overline{[1/\gamma]}_{\gamma_0} \triangleq \int_{\gamma_0}^{\infty} \frac{1}{\gamma} p(\gamma) d\gamma. \quad (13)$$

The capacity in this case, obtained by maximizing over all possible γ_0 , is

$$C(S) = \max_{\gamma_0} B \log_2[1 + \sigma_0] p(\gamma \geq \gamma_0). \quad (14)$$

The capacity of this truncated policy with optimized γ_0 is shown in [1, Figs. 3–4] to exhibit a power loss relative to the optimal policy (7) of 3–5 dB in lognormal shadowing and 1–2 dB in Rayleigh fading.³

IV. VARIABLE-RATE VARIABLE-POWER MQAM

Shannon capacity places no restriction on the complexity or delay of the multiplexed transmission scheme which achieves capacity. In fact, Shannon theory does not tell us anything about how to design this scheme. Therefore, the main emphasis of this paper is on practical adaptive modulation methods, and their spectral efficiency relative to the theoretical maximum of the previous section. Specifically, we consider a variable-rate and variable-power modulation method using MQAM signal constellations. We will see that the same optimization of power and rate used in Section III can be applied to our MQAM design. We also obtain a formula for the efficiency difference between our adaptive MQAM technique and the fading channel capacity (7).

Consider a family of MQAM signal constellations with a fixed symbol rate T_s , where M denotes the number of points in each signal constellation and we assume ideal Nyquist data pulses ($\text{sinc}[t/T_s]$) for each constellation.⁴ Let $\bar{S}, B, N_0, \gamma = (\bar{S}g/N_0B)$, and $\bar{\gamma} = (\bar{S}/N_0B)$ be as defined in Section III. Since each of our MQAM constellations has Nyquist data pulses ($B = 1/T_s$), the average E_s/N_0 equals the average SNR:

$$\frac{\bar{E}_s}{N_0} = \frac{\bar{S}T_s}{N_0} = \bar{\gamma}. \quad (15)$$

The spectral efficiency of our modulation scheme equals its data rate per unit bandwidth (R/B). For fixed $M, R = (\log_2 M)/T_s$. The spectral efficiency for fixed M is therefore $\log_2 M$, the number of bits/symbol. This efficiency is typically

³These results assume an SNR range of 5–30 dB, and a standard deviation of 8 dB for the lognormal shadowing. As this standard deviation decreases to zero, the capacity of all of the adaptive techniques approaches that of an AWGN channel.

⁴Practical Nyquist filters with nonzero excess bandwidth will reduce the spectral efficiency.

parameterized by the average transmit power \bar{S} and the BER of the modulation technique.

In [30], the BER for an AWGN channel with MQAM modulation and ideal coherent phase detection is bounded by

$$\text{BER} \leq 2e^{-1.5\bar{\gamma}/(M-1)}. \quad (16)$$

A tighter bound good to within 1 dB for $M \geq 4$ and $0 \leq \bar{\gamma} \leq 30$ dB is

$$\text{BER} \leq 0.2e^{-1.5\bar{\gamma}/(M-1)}. \quad (17)$$

In a fading channel with nonadaptive transmission (constant transmit power and rate), the received SNR varies with time. The BER in this case is obtained by integrating the BER in AWGN over the fading distribution $p(\gamma)$. For BPSK ($M = 2$) in Rayleigh fading, this integration yields $\text{BER} \approx (1/4\bar{\gamma})$ at large SNR's [31]. Without transmitter adaptation, we therefore require $\bar{\gamma} = 24$ dB to obtain a spectral efficiency of 1 at 10^{-3} BER. For $M \geq 4$, we can bound the average BER by integrating over (17):

$$\text{BER} \leq \int 0.2e^{-1.5\gamma/(M-1)}p(\gamma) d\gamma, \quad M \geq 4. \quad (18)$$

Setting $M = 4$ in (18) yields a required average SNR of $\bar{\gamma} = 26$ dB to obtain a spectral efficiency of 2 at 10^{-3} BER. We will see below that adaptive techniques yield much higher spectral efficiencies at these BER and power specifications.

We now consider adapting the transmit power $S(\gamma)$ relative to γ , subject to the average power constraint \bar{S} . The received SNR is then $\gamma S(\gamma)/\bar{S}$, and the BER bound for each value of γ becomes

$$\text{BER}(\gamma) \leq 0.2 \exp \left[\frac{-1.5\gamma}{M-1} \frac{S(\gamma)}{\bar{S}} \right]. \quad (19)$$

We can also adjust M and $S(\gamma)$ to maintain a fixed BER. Rearranging (19) yields the following maximum constellation size for a given BER:

$$M(\gamma) = 1 + \frac{1.5\gamma}{-\ln(5\text{BER})} \frac{S(\gamma)}{\bar{S}}. \quad (20)$$

We maximize spectral efficiency by maximizing

$$E[\log_2 M(\gamma)] = \int \log_2 \left(1 + \frac{1.5\gamma}{-\ln(5\text{BER})} \frac{S(\gamma)}{\bar{S}} \right) p(\gamma) d\gamma \quad (21)$$

subject to the power constraint

$$\int S(\gamma)p(\gamma) d\gamma = \bar{S}. \quad (22)$$

The power control policy that maximizes (21) has the same form as the optimal power control policy (5) which achieves capacity:

$$\frac{S(\gamma)}{\bar{S}} = \begin{cases} \frac{1}{\gamma_0} - \frac{1}{\gamma K}, & \gamma \geq \gamma_0/K \\ 0, & \gamma < \gamma_0/K, \end{cases} \quad (23)$$

where γ_0/K is the optimized cutoff fade depth and

$$K = \frac{-1.5}{\ln(5\text{BER})}. \quad (24)$$

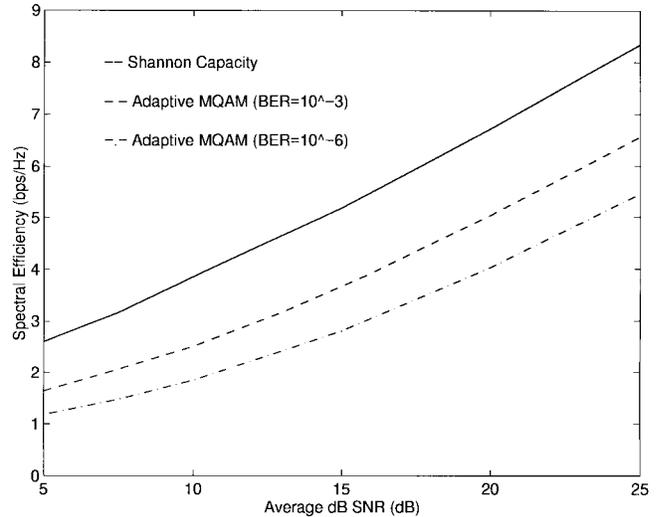


Fig. 2. Efficiency in lognormal shadowing ($\sigma = 8$ dB).

If we define $\gamma_K = \gamma_0/K$ and substitute (23) into (21), we get the maximum spectral efficiency

$$\frac{R}{B} = \int_{\gamma_K}^{\infty} \log_2 \left(\frac{\gamma}{\gamma_K} \right) p(\gamma). \quad (25)$$

Comparing the power adaptations (23) and (5) and the spectral efficiencies (7) and (25), we see that the power adaptation and spectral efficiency for both the optimal transmission scheme and our MQAM technique are the same, with an effective power loss of K in the latter case. In other words, there is a simple relationship between the maximum spectral efficiency of a fading channel and the spectral efficiency of our uncoded adaptive MQAM technique: uncoded MQAM has an effective power loss of K relative to the optimal transmission scheme, independent of the fading distribution. Equivalently, K is the maximum possible coding gain for our adaptive MQAM method. We discuss coding techniques for our adaptive modulation in Section VIII. It is interesting to note that this constant gap between Shannon capacity and the spectral efficiency of MQAM has also been reported for time-invariant channels with ISI and decision-feedback equalization [32], [33].

We compute the efficiency (25) at BER's of 10^{-3} and 10^{-6} for both lognormal shadowing (relative to the average decibel received power and for a standard deviation $\sigma = 8$ dB) and Rayleigh fading in Figs. 2 and 3, respectively. We also plot the capacity (7) in these figures for comparison.

We can also apply the suboptimal policies of total and truncated channel inversion to adaptive MQAM. The spectral efficiency with total channel inversion is obtained by substituting (9) in (20):

$$\frac{R}{B} = \log_2 \left(1 + \frac{-1.5}{\ln(5\text{BER})[1/\gamma]} \right). \quad (26)$$

This spectral efficiency is based on the tight bound (17); if $R/B < 4$, the loose bound (16) must be used and the spectral efficiency recalculated.

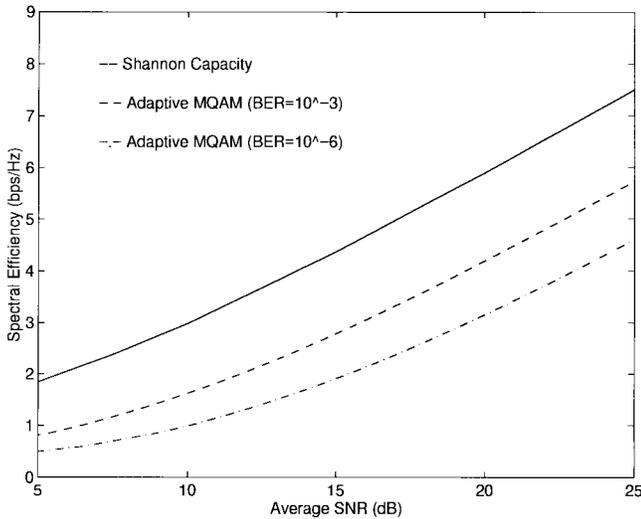


Fig. 3. Efficiency in Rayleigh fading.

With truncated channel inversion, the channel is only used when $\gamma > \gamma_0$. Thus, the spectral efficiency with truncated channel inversion is obtained by substituting (12) into (20) and multiplying by the probability that $\gamma > \gamma_0$. The maximum value is obtained by optimizing relative to the cutoff level γ_0 :

$$\frac{R}{B} = \max_{\gamma_0} \log_2 \left(1 + \frac{-1.5}{\ln(5\text{BER})[1/\gamma]_{\gamma_0}} \right) p(\gamma > \gamma_0). \quad (27)$$

The spectral efficiency of MQAM with these suboptimal policies, in both lognormal and Rayleigh fading, is evaluated in [34, Figs. 3 and 4].

The spectral efficiencies (25)–(27) place no restrictions on the constellation size; indeed, the size is not even restricted to integer values. While transmission at noninteger rates is possible, the complexity is quite high [40]. Moreover, it is difficult in practice to continually adapt the transmit power and constellation size to the channel fading, particularly in fast fading environments. Thus, we now consider restricting the constellation size to just a handful of values. This will clearly impact the spectral efficiency, although, surprisingly, not by very much.

V. CONSTELLATION RESTRICTION

We now restrict ourselves to MQAM constellations of size $M_0 = 0$, $M_1 = 2$, and $M_j = 2^{2(j-1)}$, $j = 2, \dots, N$. We use square constellations for large M due to their inherent spectral efficiency and ease of implementation [31]. We first consider the impact of this restriction on the spectral efficiency of the optimal adaptation policy. We then determine the effect on the channel inversion policies.

A. Optimal Adaptation

We now optimize the variable-rate variable-power MQAM transmission scheme subject to the constellation restrictions described above. Thus, at each symbol time, we transmit a constellation from the set $\{M_j: j = 0, 1, \dots, N\}$: the choice of constellation depends on the fade level γ over that symbol

time. Choosing the M_0 constellation corresponds to no data transmission. For each value of γ , we must decide which constellation to transmit and what the associated transmit power should be. The rate at which the transmitter must change its constellation and power is analyzed below. Since the power adaptation is continuous while the constellation size is discrete, we call this the continuous-power discrete-rate adaptation scheme.

We determine the constellation size associated with each γ by discretizing the range of channel fade levels. Specifically, we divide the range of γ into $N + 1$ fading regions, and associate the constellation M_j with the j th region. The data rate for γ values falling in the j th region is thus $\log_2 M_j$.

We set the region boundaries as follows. Define

$$M(\gamma) = \frac{\gamma}{\gamma_K^*} \quad (28)$$

where $\gamma_K^* > 0$ is a parameter which will later be optimized to maximize spectral efficiency. Note that substituting (23) into (20) yields (28) with $\gamma_K^* = \gamma_K$. Therefore, the appropriate choice of γ_K^* in (28) defines the optimal constellation size for each γ when there is no constellation restriction.

Assume now that γ_K^* is fixed, and define $M_{N+1} = \infty$. To obtain the constellation size M_j for a fixed γ , we first compute $M(\gamma)$ from (28). We then find j such that $M_j \leq M(\gamma) < M_{j+1}$ and assign constellation M_j to this γ value. Thus, for a fixed γ , we transmit the largest constellation in our set $\{M_j: j = 0, \dots, N\}$ that is smaller than $M(\gamma)$. For example, if the fade level γ satisfies $2 \leq \gamma/\gamma_K^* < 4$, we transmit the BPSK signal constellation. The region boundaries are located at $\gamma = \gamma_K^* M_j$, $j = 0, \dots, N + 1$. Clearly, increasing the number of discrete signal constellations N yields a better approximation to the continuous adaptation (20), resulting in a higher spectral efficiency.

Once the regions and associated constellations are fixed, we must find a power control policy which satisfies the BER requirement and the power constraint. By (20), we can maintain a fixed BER for the constellation $M_j > 0$ using the power control policy

$$\frac{S_j(\gamma)}{\bar{S}} = \begin{cases} (M_j - 1) \frac{1}{\gamma_K^*}, & M_j \leq \frac{\gamma}{\gamma_K^*} \leq M_{j+1} \\ 0, & M_j = 0. \end{cases} \quad (29)$$

A fixed BER implies that the received E_s/N_0 for each constellation M_j is constant:

$$\frac{E_s(j)}{N_0} = \frac{\gamma S_j(\gamma)}{\bar{S}} = \frac{M_j - 1}{K} \quad (30)$$

where $S_j(\gamma)/\bar{S}$ is defined in (29). In Table I, we tabulate the constellation size and power adaptation as a function of γ and γ_K^* for five fading regions.

The spectral efficiency for this discrete-rate policy is just the sum of the data rates associated with each of the regions multiplied by the probability that γ falls in that region:

$$\frac{R}{B} = \sum_{j=1}^N \log_2(M_j) p(M_j \leq \gamma/\gamma_K^* < M_{j+1}). \quad (31)$$

TABLE I
RATE AND POWER ADAPTATION FOR FIVE REGIONS

Region(<i>j</i>)	γ Range	M_j	$S_j(\gamma)/\bar{S}$
0	$0 \leq \gamma/\gamma_K^* < 2$	0	0
1	$2 \leq \gamma/\gamma_K^* < 4$	2	$\frac{1}{K\gamma}$
2	$4 \leq \gamma/\gamma_K^* < 16$	4	$\frac{3}{K\gamma}$
3	$16 \leq \gamma/\gamma_K^* < 64$	16	$\frac{15}{K\gamma}$
4	$64 \leq \gamma/\gamma_K^* < \infty$	64	$\frac{63}{K\gamma}$

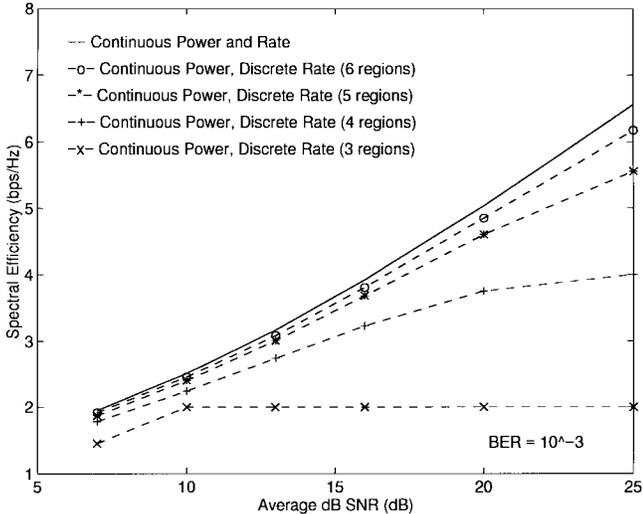


Fig. 4. Discrete-rate efficiency in lognormal shadowing ($\sigma = 8$ dB).

Since M_j is a function of γ_K^* , we can maximize (31) relative to γ_K^* , subject to the power constraint

$$\sum_{j=1}^N \int_{\gamma_K^* M_j}^{\gamma_K^* M_{j+1}} \frac{S_j(\gamma)}{\bar{S}} p(\gamma) d\gamma = 1 \quad (32)$$

where $S_j(\gamma)/\bar{S}$ is defined in (29). There is no closed-form solution for the optimal γ_K^* : in the calculations below, it was found using numerical search techniques.

In Figs. 4 and 5, we show the maximum of (31) versus the number of regions ($N + 1$) for lognormal shadowing and Rayleigh fading, respectively. These figures assume a BER of 10^{-3} . The plots for BER's of 10^{-6} were similar in shape with a smaller spectral efficiency. From Fig. 4, we see that restricting our adaptive policy to just six different signal constellations ($M_j = 0, 2, 4, 16, 64, 256$) results in a spectral efficiency that is within 1 dB of the efficiency obtained with continuous-rate adaptation (25). Fig. 5 also shows approximately 1 dB of power loss for Rayleigh fading using five constellations ($M_j = 0, 2, 4, 16, 64$). Similar results were obtained at a BER of 10^{-6} .

We can simplify our discrete-rate policy even further by using a constant transmit power for each constellation M_j . Thus, each fading region is associated with one signal constellation and one transmit power. We call this the discrete-power discrete-rate policy. Ideally, the fixed transmit power associated with each region should be optimized to maximize spectral efficiency. However, since we do not have a closed-form expression for the spectral efficiency of this policy, we

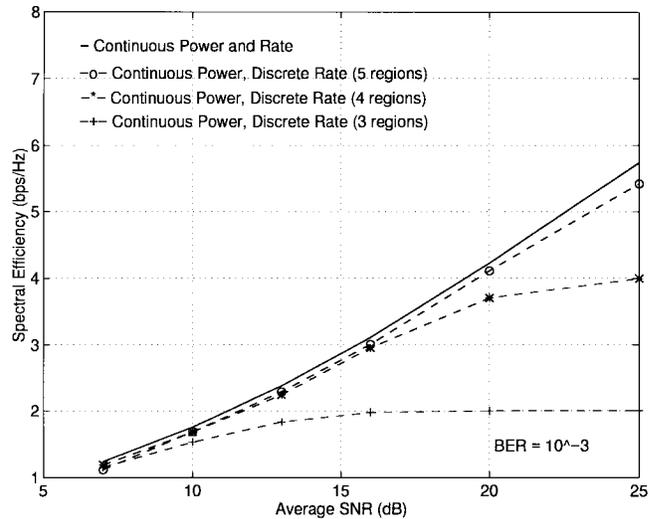


Fig. 5. Discrete-rate efficiency in Rayleigh fading.

cannot perform the optimization. We will present simulation results for this policy in Section VI using suboptimal transmit power values. Even with this suboptimal choice, these simulations demonstrate that keeping the transmit power constant in each region results in less than 2 dB of power loss relative to the continuous-power discrete-rate policy.

The choice of the number of regions to use in the adaptive policy will depend on how fast the channel is changing as well as on the hardware constraints, which dictate how many constellations are available to the transmitter and at what rate the transmitter can change its constellation and power. For constellation adaptation on a per-symbol basis, the number of regions must be chosen such that the channel gain stays within one region over a symbol time. However, hardware constraints and pulse-shaping considerations may dictate that the constellation remain constant over tens or even hundreds of symbols. In addition, power-amplifier linearity requirements and out-of-band emission constraints may restrict the rate at which power can be adapted. An in-depth discussion of hardware implementation issues and a description of a VLSI prototype can be found in [35]. Hardware advances will eventually make today's constraints obsolete. However, determining how long the channel gain remains within a particular region is of interest since it determines the tradeoff between the number of regions and the rate of power and constellation adaptation. We now derive this value.

Let $\bar{\tau}_j$ denote the average time duration that γ stays within the j th fading region. Let $A_j = \gamma_K^* M_j$ for γ_K^* and M_j as defined above. The j th fading region is then defined as $\{\gamma: A_j \leq \gamma < A_{j+1}\}$. We call $\bar{\tau}_j$ the j th average fade region duration (AFRD). This definition is similar to the average fade duration (AFD) [26, eq. (1.3-44)], except that the AFD measures the average time that γ stays below a single level, whereas we are interested in the average time that γ stays between two levels. For the worst case region ($j = 0$), these two definitions coincide.

Determining the exact value of $\bar{\tau}_j$ requires a complex derivation based on the joint density $p(\gamma, \dot{\gamma})$, and remains an open problem. However, we can obtain a good approximation

using the finite-state Markov model derived in [36]. In that paper, fading is approximated as a discrete-time Markov process with time discretized to the symbol period T_s . The underlying assumption of the model is that γ remains within one region over a symbol period, and from a given region, the process can only transition to the same region or to adjacent regions. These assumptions are consistent with our model, where γ stays within one region over a symbol time. The transition probabilities between regions under this assumption are given as

$$\begin{aligned} p_{j,j+1} &= \frac{N_{j+1}T_s}{\pi_j}, & p_{j,j-1} &= \frac{N_jT_s}{\pi_j}, \\ p_{j,j} &= 1 - p_{j,j+1} - p_{j,j-1} \end{aligned} \quad (33)$$

where N_j is the level-crossing rate at A_j and π_j is the steady-state distribution corresponding to the j th region: $\pi_j = p(A_j \leq \gamma < A_{j+1})$. Since the time in which the Markov process stays in a given state is geometrically distributed [37, eq. (2.66)], $\bar{\tau}_j$ is given by

$$\bar{\tau}_j = \frac{T_s}{p_{j,j+1} + p_{j,j-1}} = \frac{\pi_j}{N_{j+1} + N_j}. \quad (34)$$

The value of $\bar{\tau}_j$ is thus a simple function of the level-crossing rate and the fading distribution. While the level-crossing rate is known for Rayleigh fading [26, Sec. 1.3.4], it cannot be obtained for lognormal shadowing since the joint distribution $p(\gamma, \dot{\gamma})$ for this fading type is unknown.

In Rayleigh fading, the level-crossing rate is given by

$$N_j = \sqrt{\frac{2\pi A_j}{\bar{\gamma}}} f_D e^{-A_j/\bar{\gamma}} \quad (35)$$

where $f_D = v/\lambda$ is the Doppler frequency. Substituting (35) into (34), it is easily seen that $\bar{\tau}_j$ is inversely proportional to the Doppler frequency. Moreover, since π_j and A_j do not depend on f_D , if we compute $\bar{\tau}_j$ for a given Doppler frequency f_D , we can compute $\hat{\tau}_j$ corresponding to another Doppler frequency \hat{f}_D as

$$\hat{\tau}_j = \frac{f_D}{\hat{f}_D} \bar{\tau}_j. \quad (36)$$

We show in Table II the $\bar{\tau}_j$ values corresponding to five regions ($M_j = 0, 2, 4, 16, 64$) in Rayleigh fading⁵ for $f_D = 100$ Hz and two average power levels: $\bar{\gamma} = 10$ dB ($\gamma_K^* = 1.22$) and $\bar{\gamma} = 20$ dB ($\gamma_K^* = 1.685$). The AFRD for other Doppler frequencies is easily obtained using the table values and (36). This table indicates that, even at high velocities, for symbol rates of 100 ksymbols/s, the discrete-rate discrete-power policy will maintain the same constellation and transmit power over tens-hundreds of symbols.

Signal variation due to shadowing is generally much slower than variation due to multipath, and so in shadowing, γ will generally remain within the same region over hundreds or thousands of symbols. We can make this statement more precise through a coarse approximation for $\bar{\tau}_j$ based on the autocorrelation function (2). Specifically, if $\bar{\tau}_j \approx 0.1X_c/v$,

⁵The validity of the finite-state Markov model for Rayleigh fading channels has been confirmed in [38].

TABLE II
AFRD $\bar{\tau}_j$ FOR $f_m = 100$ Hz (RAYLEIGH FADING)

Region(j)	$\bar{\gamma} = 10$ dB	$\bar{\gamma} = 20$ dB
0	2.23ms	.737ms
1	.830ms	.301ms
2	3.00ms	1.06ms
3	2.83ms	2.28ms
4	1.43ms	3.84ms

then the correlation between fade levels separated in time by $\bar{\tau}_j$ is 0.9. The decorrelation distance X_c is on the order of tens of meters in microcells and hundreds of meters in macrocells [28], and thus even at high vehicle speeds, the shadowing level will remain roughly constant for tens-hundreds of milliseconds.

B. Suboptimal Policies

A restriction on allowable signal constellations will also affect the total channel inversion and truncated channel inversion policies. Specifically, although the power adaptation policies remain the same, the constellation must be chosen from the signal set $\mathcal{M} = \{0, 2, 4, 16, 64, 256\}$. For total channel inversion, the spectral efficiency with this restriction is thus

$$\frac{R}{B} = \log_2 \left(\left\lfloor 1 + \frac{-1.5}{\ln(5\text{BER})[1/\gamma]} \right\rfloor_{\mathcal{M}} \right) \quad (37)$$

where $\lfloor x \rfloor_{\mathcal{M}}$ denotes the largest number in the set \mathcal{M} less than or equal to x . The spectral efficiency with this policy will be restricted to values of $\log_2 M$, $M \in \mathcal{M}$, with discrete jumps at the $\bar{\gamma}$ values where the spectral efficiency without constellation restriction (26) equals $\log_2 M$ for some $M \in \mathcal{M}$. For truncated channel inversion, the spectral efficiency is given by

$$\frac{R}{B} = \max_{\gamma_0} \log_2 \left(\left\lfloor 1 + \frac{-1.5}{\ln(5\text{BER})[1/\gamma]} \right\rfloor_{\mathcal{M}} \right) p(\gamma > \gamma_0). \quad (38)$$

VI. SIMULATION RESULTS

We use the Communication System Simulation and Analysis Package (COSSAP) by Synopsys for our simulations. Fixed-rate transmitter and receiver modules were used as building blocks for the variable-rate transmitter and receiver simulation. The Rayleigh and lognormal shadowing simulation modules in the COSSAP library were used [39], with velocity entered as a parameter. The velocity was chosen so that, over a symbol time T_s , γ stays within one fading region with high probability. The constellation size transmitted at each symbol time was determined using the discrete-rate adaptive policy outlined in the previous section, assuming perfect instantaneous knowledge of the simulated fade level γ at the transmitter and receiver. We also assumed coherent phase detection at the receiver and Nyquist data pulses with zero ISI. Gray coding was used for bit mapping to the MQAM constellations.

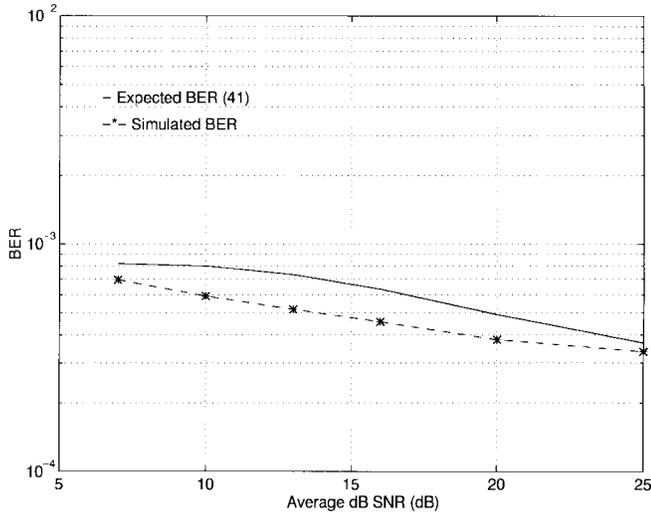


Fig. 6. BER for lognormal shadowing (six regions).

We expect our simulated BER to be slightly smaller than the target BER since (19) is an upper bound. An exact BER expression for MQAM with two-dimensional Gray coding is [41]

$$\text{BER}(M) = \alpha_M \operatorname{erfc} \left(\sqrt{\beta_M \frac{E_b}{N_0}} \right) + \text{HOT's} \quad (39)$$

where α_M and β_M are constants which depend on M and the higher order terms (HOT's) are negligible. Moreover, for our continuous-power discrete-rate policy, the E_b/N_0 for the j th signal constellation is approximately

$$\frac{E_b(j)}{N_0} = \frac{E_s(j)}{N_0} \frac{1}{\log_2 M_j} = \frac{M_j - 1}{K \log_2 M_j}. \quad (40)$$

We obtain the exact BER for our adaptive policy by averaging over the BER (39) for each signal constellation:

$$\text{BER} = \sum_{j=1}^N \alpha_{M_j} \operatorname{erfc} \left(\sqrt{\frac{\beta_{M_j} (M_j - 1)}{K \log_2 M_j}} \right) \cdot \int_{\gamma_K^* M_j}^{\gamma_K^* M_{j+1}} p(\gamma) d\gamma. \quad (41)$$

We plot (41) and the simulated BER in Figs. 6 and 7 for lognormal shadowing and Rayleigh fading, respectively, at a target BER of 10^{-3} . These simulation results are slightly better than the analytical calculation of (41), and both are smaller than the target BER of 10^{-3} , for $\bar{\gamma} > 10$ dB. The BER bound of 10^{-3} breaks down at low SNR's, since (17) is not applicable to BPSK, and we must use the looser bound (16). Since our adaptive policy will use the BPSK constellation often at low SNR's, the BER will be larger than that predicted from the tight bound (17).

The fact that the simulated BER is less than our target at high SNR's implies that the analytical calculations in Figs. 4 and 5 are pessimistic. A slightly higher efficiency could be achieved while still maintaining the target BER of 10^{-3} .

In Figs. 8 and 9, we show the simulated spectral efficiency corresponding to this simulated BER for the continuous-power

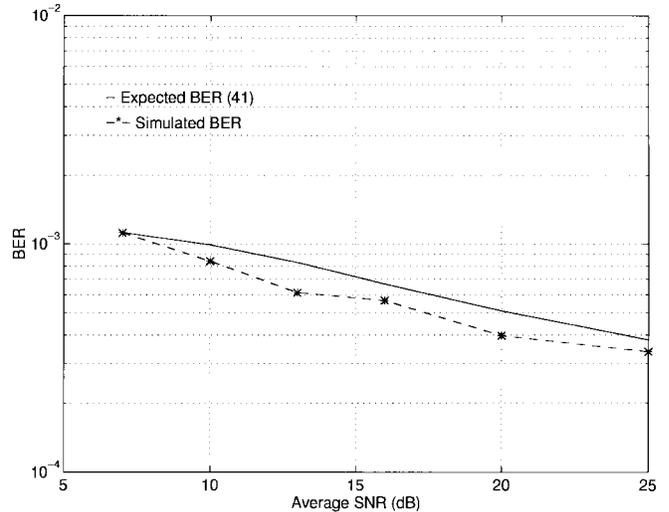


Fig. 7. BER for Rayleigh fading (five regions).

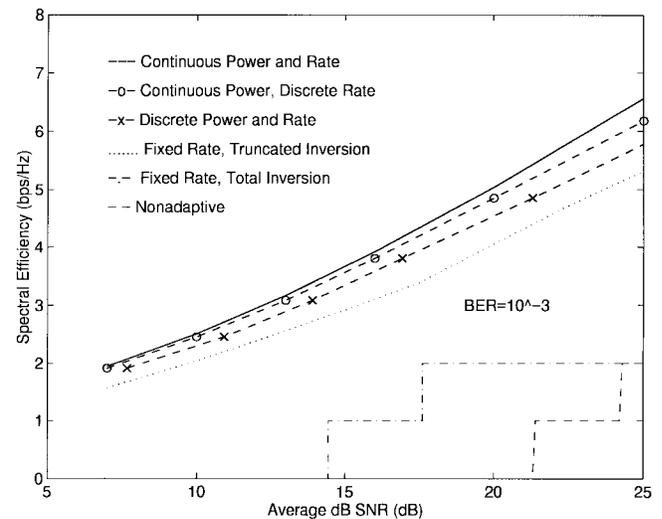


Fig. 8. Efficiency in lognormal shadowing ($\sigma = 8$ dB).

discrete-rate policy. These figures also show the simulated efficiency of the discrete-power discrete-rate policy, where the transmit power for each region was chosen to achieve the same simulated BER as the continuous-power discrete-rate policy. We see that even with this suboptimal choice of power assignment, keeping the power constant for each transmit constellation results in a power loss of just 1–2 dB relative to continuous-power adaptation. For comparison, we also plot the maximum efficiency (25) for continuous power and rate adaptation. Both discrete-rate policies have simulated performance that is within 3 dB of this theoretical maximum.

These figures also show the spectral efficiency of fixed-rate transmission with truncated channel inversion (38). The efficiency of this scheme is quite close to that of the discrete-power discrete-rate policy. However, to achieve this high efficiency, the optimal γ_0 is quite large, with a corresponding outage probability $P_{\text{out}} = p(\gamma \leq \gamma_0)$ ranging from 0.1 to 0.6. Thus, this policy is similar to packet radio, with bursts of high-speed data when the channel conditions are favorable. The efficiency of total channel inversion (37) is

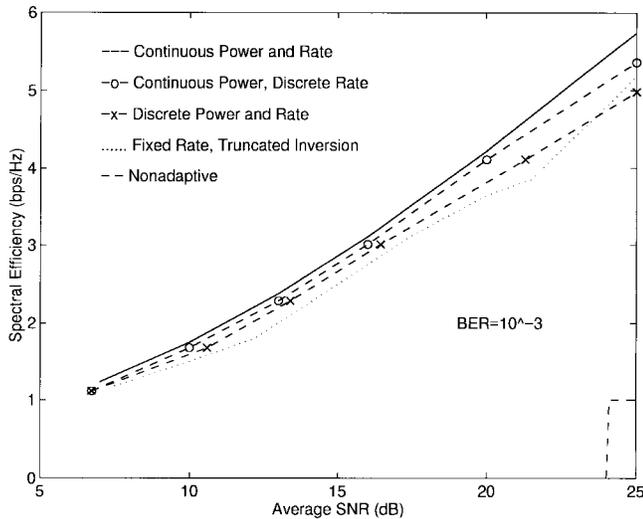


Fig. 9. Efficiency in Rayleigh fading.

also shown for lognormal shadowing: this efficiency equals zero in Rayleigh fading. We also plot the spectral efficiency of nonadaptive transmission, where both the transmission rate and power are constant. The average BER in this case is obtained by integrating the probability of error (39) against the fade distribution $p(\gamma)$. The spectral efficiency is obtained by determining the value of M which yields a 10^{-3} BER for the given value of $\bar{\gamma}$. Nonadaptive transmission clearly suffers a large efficiency loss in exchange for its simplicity. However, if the channel varies rapidly and cannot be accurately estimated, nonadaptive transmission may be the best alternative. Similar curves were obtained for a target BER of 10^{-6} , with roughly the same spectral efficiency loss relative to a 10^{-3} BER as was exhibited in Figs. 2 and 3.

VII. CHANNEL ESTIMATION ERROR AND DELAY

We now relax our earlier assumptions about estimation error and delay to consider the case when the estimation error $\epsilon = \hat{\gamma}/\gamma \neq 1$ and the delay $\tau = \tau_f + \tau_e \neq 0$. The impact of estimation error and delay will be most dramatic on systems adapting to Rayleigh fading since the signal variation due to multipath is much faster than the variation due to shadowing. We first consider the estimation error. Our adaptive modulation scheme uses the channel estimate at both the transmitter and the receiver. At the transmitter, the channel estimate is used to determine the signal constellation and power to be sent. At the receiver, the channel estimate is used to determine which signal constellation and power was sent, and also to scale the decision regions of the MQAM demodulator corresponding to that constellation and power. Note that the channel estimate used to determine the signal constellation and power must be strictly causal since the transmitter must adapt to the changing channel. However, the channel estimate used to scale the decision regions can be noncausal (i.e., it can interpolate between past and future channel estimates or pilot symbols), although this will introduce delay in the decoding. Thus, the channel estimate used for rate and power adaptation $\hat{\gamma}$ will

typically be less reliable than the estimate $\hat{\gamma}$ used for decision-region scaling in the demodulator.

The estimation error therefore affects the BER of the adaptive system in two ways. First of all, the transmit power and rate will be adapted as $S(\hat{\gamma})$ and $M(\hat{\gamma})$ instead of $S(\gamma)$ and $M(\gamma)$, which may cause a change in the target data rate and/or BER. In addition, the decision regions of the MQAM demodulator will be scaled according to the channel estimate $\hat{\gamma}$ instead of its correct value, which will cause an increase in BER. This imperfect scaling of the decision regions is also referred to as imperfect AGC. The effect of imperfect AGC for nonadaptive MQAM using pilot-symbol-assisted modulation (PSAM) in Rayleigh fading was studied in [25], [42]. In these studies, imperfect AGC was found to degrade performance by 1–3 dB over a wide range of Doppler frequencies and MQAM signal constellations ranging from BPSK to 256-QAM. The extension of these results to adaptive modulation with imperfect AGC is beyond the scope of this paper, and we defer this study to future work, although we expect a similar 1–3 dB degradation. We will, however, determine the effect of estimation error on the BER resulting from imperfect transmit power and rate adaptation, assuming perfect AGC. The full impact of estimation error combines the effects of imperfect power and rate adaptation with imperfect AGC.

We now analyze the effect of $\hat{\gamma} \neq \gamma$ on BER, assuming perfect AGC ($\hat{\gamma} = \gamma$). Suppose the transmitter adapts its power and rate relative to a target BER_0 based on the channel estimate $\hat{\gamma}$ instead of the true value γ . From (19), the BER is then bounded by

$$\begin{aligned} \text{BER}(\gamma, \hat{\gamma}) &\leq 0.2 \exp \left[\frac{-1.5\gamma}{M(\hat{\gamma}) - 1} \frac{S(\hat{\gamma})}{\bar{S}} \right] \\ &= 0.2[5\text{BER}_0]^{1/\epsilon} \end{aligned} \quad (42)$$

where the second equality is obtained by substituting the optimal rate (20) and power (23) policies. For $\epsilon = \hat{\gamma}/\gamma = 1$, (42) reduces to the target BER_0 . For $\epsilon \neq 1$, $\epsilon > 1$ yields an increase in BER, and $\epsilon < 1$ yields a decrease in BER.

Given the distribution $p_\epsilon(x)$ for the estimation error, the effect of this error on BER is given by

$$\overline{\text{BER}} \leq \int_0^\infty 0.2[5\text{BER}_0]^{1/x} p_\epsilon(x) dx. \quad (43)$$

The distribution $p_\epsilon(x)$ will depend on the channel estimation technique. Although the estimation error variance for some estimation methods in Rayleigh fading has been obtained [25], analytic and/or empirical distributions for estimation error are generally unknown. In addition, while the estimation error for shadow fading in the presence of Rayleigh fading was characterized in [43], our shadow fading model assumes that the Rayleigh fading is mostly removed through diversity, and the estimation error under this assumption has not been fully characterized. The lack of appropriate models for estimation error therefore precludes a full characterization of its effects. However, if the estimation error stays within some finite range, then we can bound its effects using (42).

We plot the BER increase as a function of a constant ϵ in Fig. 10. This figure shows that for a target BER of 10^{-3} ,

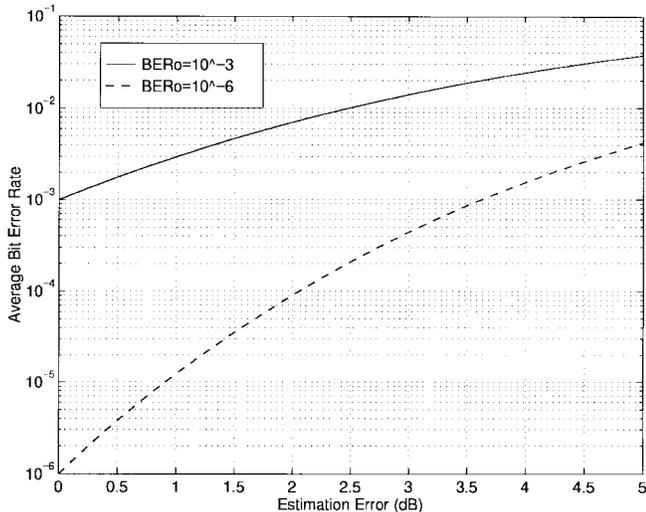


Fig. 10. Effect of estimation error on BER (perfect AGC).

the estimation error should be less than 1 dB, and for a target BER of 10^{-6} , it should be less than 0.5 dB. These values are pessimistic since they assume a constant value of estimation error. Even so, the estimation error can be kept within this range using the PSAM estimation technique described in [25] with appropriate choice of parameters (pilot symbol power, spacing, and interpolation length as a function of the channel Doppler). When the channel is underestimated ($\epsilon < 1$), the BER decreases, but there will also be some loss in spectral efficiency since the mean of the channel estimate $\hat{\gamma}$ will differ from the true mean $\bar{\gamma}$. The effect of this average power estimation error on spectral efficiency is characterized in [44], where it is shown to be negligible. Pilot symbols used for channel estimation will also reduce spectral efficiency, particularly for rapidly varying channels, where the pilot symbols must be sent frequently [25].

Suppose now that the channel is estimated perfectly ($\epsilon = 1$), but the delay τ of the estimation and feedback path is nonzero. Thus, at time t , the transmitter will use the delayed version of the channel estimate $\hat{\gamma}(t) = \gamma(t - \tau)$ to adjust its power and rate. The resulting increase in BER is obtained in the same manner as (42):

$$\begin{aligned} \text{BER}(\gamma(t), \hat{\gamma}(t)) &\leq 0.2 \exp \left[\frac{-1.5\gamma(t)}{M(\hat{\gamma}(t)) - 1} \frac{S(\hat{\gamma}(t))}{\bar{S}} \right] \\ &= 0.2 [5\text{BER}_0]^{\gamma(t-\tau)/\gamma(t)}. \end{aligned} \quad (44)$$

Define $\rho(t, \tau) = \gamma(t - \tau)/\gamma(t)$. Since $\gamma(t)$ is stationary and ergodic, the distribution of $\rho(t, \tau)$ conditioned on $\gamma(t)$ depends only on τ and the value of $\gamma = \gamma(t)$. We denote this distribution by $p_\tau(\rho|\gamma)$. The average BER is obtained by integrating over ρ and γ . Specifically, it is shown in [45] that

$$\text{BER}(\tau) = \int_{\gamma_0}^{\infty} \left[\int_0^{\infty} 0.2 [5\text{BER}_0]^\rho p_\tau(\rho|\gamma) d\rho \right] p(\gamma) d\gamma \quad (45)$$

where γ_0 is the cutoff level of the optimal policy and $p(\gamma)$ is the fading distribution. The distribution $p_\tau(\rho|\gamma)$ will depend on the autocorrelation of the fading process. A closed-form expression for $p_\tau(\rho|\gamma)$ in Nakagami fading (of which Rayleigh

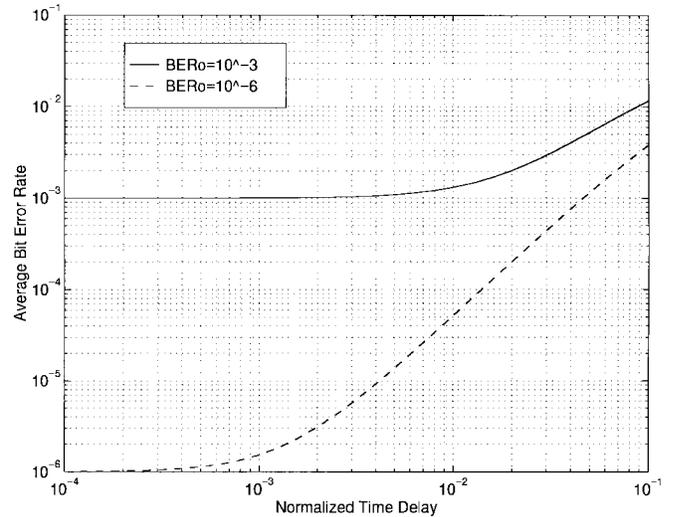


Fig. 11. Effect of normalized delay (τf_D) on BER.

fading is a special case), based on the autocorrelation function (1), is derived in [45]. Using this distribution in (45), we obtain the average BER in Rayleigh fading as a function of the delay parameter τ . A plot of (45) versus the normalized time delay τf_D is shown in Fig. 11. From this figure, we see that the total estimation and feedback path delay must be kept to within $0.001/f_D$ to keep the BER near its desired target.

VIII. CODING ISSUES AND CAPACITY REVISITED

A convolutional or block code can be applied to the uncoded bit stream before modulation to reduce the BER. If adaptive modulation is applied to these coded bits, they will not suffer burst errors typically exhibited on fading channels. Since the adaptive modulation keeps the BER constant under all fading conditions by adjusting the transmit power and rate, the probability of error in a deep fade is the same as with little or no fading, thereby eliminating error bursts and the need for an interleaver. Standard decoding algorithms can be applied to the demodulated bits, although some buffering may be required. Unfortunately, block and convolutional codes are not spectrally efficient, and would therefore reduce some of the efficiency gains of the variable-rate scheme.

A more effective coding scheme is to superimpose a trellis code on top of the adaptive modulation. This superimposed coding technique is investigated in [46], where a simple four-state trellis code yields an asymptotic coding gain of 3 dB and an eight-state code yields an asymptotic gain of 4 dB. It is difficult to obtain more than 4 dB of coding gain using a trellis code of reasonable complexity. Thus, the constant gap (24) between the spectral efficiency of adaptive modulation and Shannon capacity, exhibited in Figs. 2 and 3, cannot be fully closed. This discrepancy between Shannon capacity and achievable rates arises from the lack of complexity and implementation constraints inherent to Shannon theory. However, the derivation and general form of the optimal power and rate adaptation for our MQAM scheme were identical to that of the Shannon capacity analysis. Thus, although we cannot reach the Shannon limit, the intuition and general strategy of

optimal adaptation in a Shannon sense were useful guides in our adaptive modulation design.

IX. CONCLUSIONS

We have proposed a variable-rate and variable-power MQAM modulation technique which adapts to the channel variation. We compare the spectral efficiency of our adaptive method to the theoretical bound on spectral efficiency, and to the efficiency of suboptimal and nonadaptive modulation techniques. We first show that there is a constant gap between the channel capacity and the maximum efficiency of adaptive MQAM which is a simple function of the target BER. We also show that, using just five or six different signal constellations, we get a spectral efficiency within 1 dB of the efficiency without constellation restriction. Moreover, our adaptive technique has a 5–10 dB power gain over variable-power fixed-rate modulation, and up to 20 dB of power gain over nonadaptive modulation. Our technique is sensitive to channel estimation errors and to estimation and feedback path delay, and this must be taken into account in any practical implementation. We also determine how fast the transmitter must change its constellation and power as a function of the channel Doppler frequency. Hardware constraints will dictate whether these adaptation rates are feasible. In the final section, we discuss coding techniques, and reflect on the connection between the Shannon capacity of our channel model and the adaptive MQAM technique we propose.

ACKNOWLEDGMENT

The authors are indebted to the anonymous reviewers for their careful reading and critique of the manuscript. Their insights and suggestions greatly improved the paper. They would also like to thank M.-S. Alouini for valuable discussions on estimation error and delay effects, and L. J. Greenstein for early discussions on adaptive MQAM.

REFERENCES

- [1] A. J. Goldsmith and P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Trans. Inform. Theory*, to be published Nov. 1997. See also [2].
- [2] A. J. Goldsmith, "Design and performance of high-speed communication systems in time-varying radio channels," Ph.D. dissertation, Dept. Elec. Eng. Comput. Sci., Univ. California, Berkeley, 1994.
- [3] J. F. Hayes, "Adaptive feedback communications," *IEEE Trans. Commun. Technol.*, vol. COM-16, pp. 29–34, Feb. 1968.
- [4] J. K. Cavers, "Variable-rate transmission for Rayleigh fading channels," *IEEE Trans. Commun.*, vol. COM-20, pp. 15–22, Feb. 1972.
- [5] S. Otsuki, S. Sampei, and N. Morinaga, "Square-QAM adaptive modulation/TDMA/TDD systems using modulation level estimation with Walsh function," *Electron. Lett.*, vol. 31, pp. 169–171, Feb. 1995.
- [6] W. T. Webb and R. Steele, "Variable rate QAM for mobile radio," *IEEE Trans. Commun.*, vol. 43, pp. 2223–2230, July 1995.
- [7] Y. Kamio, S. Sampei, H. Sasaoka, and N. Morinaga, "Performance of modulation-level-controlled adaptive-modulation under limited transmission delay time for land mobile communications," in *Proc. IEEE VTC'95*, July 1995, pp. 221–225.
- [8] B. Vucetic, "An adaptive coding scheme for time-varying channels," *IEEE Trans. Commun.*, vol. 39, pp. 653–663, May 1991.
- [9] S. M. Alamouti and S. Kallel, "Adaptive trellis-coded multiple-phased-shift keying for Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 42, pp. 2305–2314, June 1994.
- [10] T. Ue, S. Sampei, and N. Morinaga, "Symbol rate and modulation level controlled adaptive modulation/TDMA/TDD for personal communication systems," in *Proc. IEEE VTC'95*, July 1995, pp. 306–310.
- [11] H. Matsuoka, S. Sampei, N. Morinaga, and Y. Kamio, "Symbol rate and modulation level controlled adaptive modulation/TDMA/TDD for personal communication systems," in *Proc. IEEE VTC'96*, Apr. 1996, pp. 487–491.
- [12] S. Sampei, N. Morinaga, and Y. Kamio, "Adaptive modulation/TDMA with a BDDFE for 2 Mbit/s multi-media wireless communication systems," in *Proc. IEEE VTC'95*, July 1995, pp. 311–315.
- [13] M.-S. Alouini and A. J. Goldsmith, "Capacity of Rayleigh fading channels under different adaptive transmission and diversity-combining techniques," *IEEE Trans. Veh. Technol.*, submitted for publication.
- [14] J. A. C. Bingham, "Multicarrier modulation for data transmission: An idea whose time has come," *IEEE Commun. Mag.*, vol. 28, pp. 5–14, May 1990.
- [15] P. S. Chow, J. M. Cioffi, and J. A. C. Bingham, "A practical discrete multitone transceiver loading algorithm for data transmission over spectrally shaped channels," *IEEE Trans. Commun.*, vol. 43, Feb./Mar./Apr. 1995.
- [16] M. Filip and E. Vilar, "Optimum utilization of the channel capacity of a satellite link in the presence of amplitude scintillations and rain attenuation," *IEEE Trans. Commun.*, vol. 38, pp. 1958–1965, Nov. 1990.
- [17] A. M. Monk and L. B. Milstein, "Open-loop power control error in a land mobile satellite system," *IEEE J. Select. Areas. Commun.*, vol. 13, pp. 205–212, Feb. 1995.
- [18] J. L. Rose, "Satellite communications in the 30/20 GHz band," *Satellite Commun.*, On-Line Publications, pp. 155–162, 1985.
- [19] R. V. Cox, J. Hagenauer, N. Seshadri, and C.-E. W. Sundberg, "Subband speech coding and matched convolutional channel coding for mobile radio channels," *IEEE Trans. Signal Processing*, vol. 39, pp. 1717–1731, Aug. 1991.
- [20] L. C. Yun and D. G. Messerschmitt, "Variable quality of service in CDMA systems by statistical power control," in *Proc. IEEE Int. Commun. Conf.*, June 1995, pp. 713–719.
- [21] G. D. Forney, "Trellis shaping," *IEEE Trans. Inform. Theory*, vol. 38, pp. 281–300, Mar. 1992.
- [22] K. S. Gilhousen, I. M. Jacobs, R. Padovani, A. J. Viterbi, L. A. Weaver, Jr., and C. E. Wheatley, III, "On the capacity of a cellular CDMA system," *IEEE Trans. Veh. Technol.*, vol. 40, pp. 303–312, May 1991.
- [23] J. Zander, "Performance of optimum transmitter power control in cellular radio systems," *IEEE Trans. Veh. Technol.*, vol. 41, pp. 57–62, Feb. 1992.
- [24] M.-S. Alouini and A. J. Goldsmith, "Area spectral efficiency of cellular mobile radio systems," in *Proc. IEEE Veh. Technol. Conf.*, May 1997, pp. 652–656. Also, *IEEE Trans. Veh. Technol.*, submitted for publication.
- [25] J. K. Cavers, "An analysis of pilot symbol assisted modulation for Rayleigh fading channels," *IEEE Trans. Veh. Technol.*, pp. 686–693, Nov. 1991.
- [26] W. C. Jakes, Jr., *Microwave Mobile Communications*. New York: Wiley, 1974.
- [27] R. Vijayan and J. M. Holtzman, "Foundations for level crossing analysis of handoff algorithms," in *Proc. IEEE ICC*, June 1993, pp. 935–939.
- [28] M. Gudmundson, "Correlation model for shadow fading in mobile radio systems," *Electron. Lett.*, vol. 27, pp. 2145–2146, Nov. 7, 1991.
- [29] W. C. Y. Lee, "Estimate of channel capacity in Rayleigh fading environment," *IEEE Trans. Veh. Technol.*, vol. 39, pp. 187–189, Aug. 1990.
- [30] G. J. Foschini and J. Salz, "Digital communications over fading radio channels," *Bell Syst. Tech. J.*, pp. 429–456, Feb. 1983.
- [31] J. G. Proakis, *Digital Communications*, 2nd ed. New York: McGraw-Hill, 1989.
- [32] R. Price, "Non-linearly feedback equalized PAM versus capacity for noisy filter channels," in *Proc. IEEE ICC*, June 1972.
- [33] M. V. Eyuboglu, "Detection of coded modulation signals on linear, severely distorted channels using decision feedback noise prediction with interleaving," *IEEE Trans. Commun.*, vol. 36, pp. 401–409, Apr. 1988.
- [34] A. Goldsmith and P. Varaiya, "Increasing spectral efficiency through power control," in *Proc. IEEE ICC*, June 1993, pp. 600–604.
- [35] M. Filip and E. Vilar, "Implementation of adaptive modulation as a fade countermeasure," *Int. J. Satellite Commun.*, vol. 12, pp. 181–191, 1994.
- [36] H. S. Wang and N. Moayeri, "Finite-state Markov channel—A useful model for radio communication channels," *IEEE Trans. Veh. Technol.*, vol. 44, pp. 163–171, Feb. 1995.
- [37] L. Kleinrock, *Queueing Systems, Volume I: Theory*. New York: Wiley, 1975.
- [38] H. S. Wang and P.-C. Chang, "On verifying the first-order Markov assumption for a Rayleigh fading channel model," *IEEE Trans. Veh. Technol.*, vol. 45, pp. 353–357, May 1996.
- [39] *COSSAP Model Libraries*, vol. 1, Ver. 6.7, Synopsys, 1994.

- [40] G. D. Forney, Jr., R. G. Gallager, G. R. Lang, F. M. Longstaff, and S. U. Quereschi, "Efficient modulation for band-limited channels," *IEEE J. Select. Areas Commun.*, vol. SAC-2, pp. 632-647, Sept. 1984.
- [41] M. K. Simon, S. M. Hinedi, and W. C. Lindsey, *Digital Communication Techniques*. Englewood Cliffs, NJ: Prentice-Hall, 1995.
- [42] S. Sampei and T. Sunaga, "Rayleigh fading compensation for QAM in land mobile radio communications," *IEEE Trans. Veh. Technol.*, vol. 42, pp. 137-147, May 1993.
- [43] A. J. Goldsmith, L. J. Greenstein, and G. J. Foschini, "Error statistics of real-time power measurements in cellular channels with multipath and shadowing," *IEEE Trans. Veh. Technol.*, vol. 43, pp. 439-446, Aug. 1994.
- [44] A. J. Goldsmith and L. J. Greenstein, "Effect of average power estimation error on adaptive MQAM modulation," in *Proc. IEEE ICC*, June 1997, pp. 1105-1109.
- [45] M.-S. Alouini and A. J. Goldsmith, "Adaptive M-QAM modulation over Nakagami fading channels," IEEE GLOBECOM Conf., Nov. 1997, submitted.
- [46] S.-G. Chua and A. J. Goldsmith, "Adaptive coded modulation for fading channels," in *Proc. IEEE ICC*, June 1997, pp. 1488-1492. Also, *IEEE Trans. Commun.*, submitted for publication.



Soon Ghee Chua (S'95-A'96) received the B.S. degree with honors in electrical engineering and economics from Caltech in 1996, and the M.S. degree in electrical engineering from Stanford University in 1997. He is a recipient of the Caltech Merit Award and a Stanford Engineering Fellowship.

He is currently with Singapore Technologies, Ltd, where his research interests include modulation, coding, and the design of satellite, cellular, and digital communication networks.



Andrea Goldsmith (S'90-M'95) received the B.S., M.S., and Ph.D. degrees in electrical engineering from the University of California, Berkeley, in 1986, 1991, and 1994, respectively.

She joined the faculty of the California Institute of Technology in 1994, where her current research includes work in multimedia wireless communication techniques, joint source and channel coding, and resource allocation in cellular systems. From 1986 to 1990 she was affiliated with Maxim Technologies, where she worked on packet radio and satellite communication systems, and from 1991 to 1992 she was affiliated with AT&T Bell Laboratories, where she worked on microcell modeling and channel estimation.

Dr. Goldsmith is a recipient of the National Science Foundation CAREER Development Award, an IBM Graduate Fellowship, and the David Griep Memorial Prize from the University of California, Berkeley. She is an editor for the IEEE TRANSACTIONS ON COMMUNICATIONS and the IEEE PERSONAL COMMUNICATIONS MAGAZINE.