Adaptive Coding for Time-Varying Channels Using Outdated Fading Estimates

Dennis L. Goeckel

***Abstract*—The idea of using knowledge of the current channel fading values to optimize the transmitted signal in wireless communication systems has attracted substantial research attention in recent years. However, the practicality of this adaptive signaling has been questioned due to the variation of the wireless channel over time, which results in a different channel at the time of data transmission than at the time of channel estimation. By characterizing the effects of fading channel variation on the adaptive signaling paradigm, it is demonstrated here that these misgivings are well founded, as the channel variation greatly alters the nature of the problem. The main goal of this paper is to employ this characterization of the effects of the channel variation to design adaptive signaling schemes that are effective for the time-varying channel. The design of uncoded adaptive quadrature amplitude modulation (QAM) systems is considered first, and it demonstrates the need to consider the channel variation in system design. This is followed by the main contribution of this paper; using only a single outdated fading estimate when neither the Doppler frequency nor the exact shape of the autocorrelation function of the channel fading process is known, adaptive trellis-coded modulation schemes are designed that can provide a significant increase in bandwidth efficiency over their nonadaptive counterparts on time-varying channels.**

***Index Terms*— Adaptive coding, fading channels, quadrature amplitude modulation, time-varying channels, trellis-coded modulation.**

I. Introduction

A. Adaptive Signaling for Fading Channels

THE goal of wireless communication systems is to match the throughput of their wired counterparts. However, the limited bandwidth available and time-varying multipath fading inherent to the wireless link make this a difficult problem. One possible method of significantly increasing the bandwidth efficiency of communications systems operating over multipath fading channels is to employ adaptive signaling schemes, which use knowledge of the current channel fading values to optimize the transmitted signal.

Adaptive signaling for wireless channels is partially motivated by the throughput gains afforded by the V.34 standard for data transmission over telephone lines, where the characteristics of the channel are well-established and the channel can be probed to obtain a reliable channel quality estimate [1]; the transmitter then uses this estimate to choose the appropriate signaling method. An alternative method, which can be applied to a wireline channel with any frequency characteristic, is the use of multicarrier modulation with the modulation on each subcarrier matched to the corresponding subchannel's received signal-to-noise ratio [2], [3]. It is then a straightforward extension to propose the same idea for Rayleigh fading channels, except that the modulation on each subcarrier is varied adaptively based on subchannel fading estimates fed back from the receiver; thus, this idea has received moderate research interest historically [4]–[6] and has been of significant research interest in recent years [7]–[14]. However, the utility of these results has been questioned due to the time-varying nature of the wireless channel, which results in a different channel at the time of data transmission than at the time of channel estimation. Note that even a wireless channel with a fixed transmitter and receiver will exhibit variation due to the movement of objects in the environment. In fact, it is the relatively fast variation of the channel over time that delineates the wireless channel from many of its wireline counterparts.

In this work, the effect of fading channel variation on the adaptive coding model is formally addressed, and this characterization is used to design both uncoded and coded systems that are effective for a time-varying channel. The channel fading will be modeled as a multiplicative time-varying complex Gaussian process, and the variability of the wireless channel over time will be reflected in the autocorrelation function of the Gaussian process. With knowledge of this autocorrelation function, the characterization of the effect of the channel variation on the adaptive signaling model is straightforward. However, it illustrates the key point that the probability density function of the current channel fading amplitude is Rician when conditioned on the outdated fading estimates, with the ratio of the specular component to the diffuse component depending not only on the statistical characteristics of the channel and the delay between channel estimation and data transmission, but also on the value of the fading estimates. This suggests an adaptive signaling scheme that instead of only adapting to a modified received signal-to-noise ratio (SNR), as in schemes previously considered where the current channel fading values are assumed to be perfectly known [10], adapts instead for a channel which varies from a low-SNR nearly Rayleigh channel to a high-SNR strongly Rician channel. This characterization will also reveal

Paper approved by B. Vucetic, the Editor for Modulation of the IEEE Communications Society. Manuscript received October 10, 1997; revised November 11, 1998. This work was supported in part by the National Science Foundation under Grant NCR-9714597. This paper was presented in part at the 35th Annual Allerton Conference on Communication, Control, and Computing, Monticello, IL USA, September 1997.

The author is with the Electrical and Computer Engineering Department, University of Massachusetts, Amherst, MA 01003-5110 USA (e-mail: goeckel@ecs.umass.edu).

Publisher Item Identifier S 0090-6778(99)05011-4.

the key role that transmitter knowledge of the autocorrelation function of the channel fading process plays in adaptive coding. Since nonadaptive schemes on slowly fading channels are not dependent on the autocorrelation function of the fading process if sufficient interleaving is employed, adaptive codes must provide the same guarantees against uncertainty in the autocorrelation function. This motivates the formal definition of strongly robust adaptive signaling schemes, which guarantee performance across a class of possible autocorrelation functions for the channel fading process.

B. Related Work of Other Researchers

Until very recently, only [6] had considered the effects of the channel variation on adaptive system *design*, and the investigation in [6] was limited to adaptive binary uncoded modulation without any consideration of robustness to the uncertainty at the transmitter about the autocorrelation function of the channel fading process. In very recent work independent of the work in [15] and this paper, [16] has considered the design of trellis-coded modulation for time-varying channels. The system of [16] employs high-order predictors that use past fading values to estimate the current channel status and then signals based on this prediction. In contrast to the work contained in this paper, the thresholds that determine which constellation to employ for a given prediction are not specified analytically, and, more importantly, the high-order predictors require precise transmitter knowledge of the autocorrelation function of the channel fading process, which is generally not available in a wireless system; thus, as in previous work on adaptive signaling for wireless channels, there has been no consideration of robustness.

In recent work on adaptive coding, [9], [13], and [14] briefly address the effects of channel variation on adaptive signaling systems that are designed assuming the channel does not vary over time. In [9], the claim is made that slowly fading channels do not change rapidly enough to greatly affect adaptive system performance and that if they do change rapidly, prediction can be employed to accurately estimate the current channel fading values. The characterization of the effects of the channel variation presented here will question these claims, demonstrating that the channel variation can greatly impact the performance of adaptive codes, even in systems with low mobility and small delay between channel estimation and data transmission. As noted above, prediction relies on an accurate knowledge of the autocorrelation function of the channel fading process, which is generally not available at the transmitter. In [13], it is briefly noted through simulation that the channel variation in a certain environment increases the bit error probability by an order of magnitude. This is substantial because the bit error probabilities being considered are quite high (10^{-3}), and the characterization of the effects of the channel variation presented here will indicate that the effect will be much greater at lower error probabilities, thus greatly affecting trellis-coded modulation schemes with parallel branches. Finally, in recent work independent of the work in this paper, [14] considered an adaptive system that is designed using a single channel estimate assuming a time-

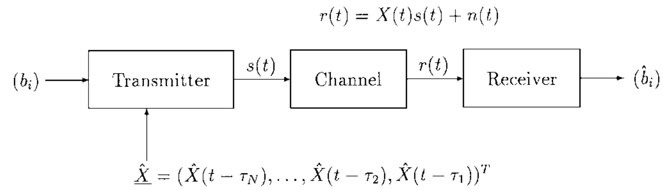

Fig. 1. System diagram.

invariant channel and characterized the performance when the system is being employed over a time-varying Nakagami fading channel. The numerical results in [14] indicate that the system works reasonably well only if the product of the delay between channel estimation and data transmission and the Doppler frequency is kept below 0.01. Thus, whereas work previous to [15] (with the exception of [6]) has employed design methods assuming perfect knowledge of the current channel fading values and then (perhaps) characterized the effect of channel variability on the performance of systems employing these methods, this work represents a significant departure in that the variability is accounted for in the design process.

Based on previous work, it may appear to the reader that accommodating an adaptive rate system is not worth the effort. It can be shown that transmitter knowledge of the current channel fading values does not increase channel capacity unless the transmit power is adapted [17], and even if the transmit power is adapted, the gain in channel capacity is small [12]. Furthermore, the recent establishment of trellis-coded modulation that operates close to the cutoff rate [18] might imply that there is little gain available even at reasonable complexities. However, it will be demonstrated in this paper that adaptive coding can provide significant gains relative to the codes presented in [18] at equal complexity for systems with low mobility, in particular indoor wireless systems in the 900-MHz ISM band with walking speed mobility.

C. Organization of Paper

In Section II, a general framework for adaptive signaling over time-varying channels is introduced, and the effects of the channel variation are considered. The impact of the uncertainty in the statistics of the channel fading at the transmitter is also discussed, which leads to the definition of strongly robust signaling. The design of strongly robust uncoded adaptive quadrature amplitude modulation (QAM) schemes for the time-varying channel is presented in Section III. In Section IV, the results of the previous sections are employed to develop a design method for strongly robust adaptive trellis-coded modulation operating over time-varying channels. Finally, Section V presents the conclusions.

II. System Model and Characterization

A. System Model

A baseband system model for the problem is shown in Fig. 1. The independent and identically distributed (IID) sequence (b_i) is the stream of information bits to be transmitted over the channel, and $(\hat{b}_i)$ is the corresponding stream of information bit estimates at the output of the receiver. Each information bit is assumed to be equally likely to be 0 or 1. The transmitted signal is given by $s(t) = \Sigma_{k=-\infty}^{\infty} z_k p(t - kT_s)$, where z_k is the kth (complex) data symbol, $p(t)$ is a pulse shape that results in no intersymbol interference in the samples (spaced at T_s) of the output of the matched filter at the receiver, and $1/T_s$ is the symbol rate. The additive noise $n(t)$ is white Gaussian noise with two-sided power spectral density $N_0/2$. The fading is modeled as a complex multiplier $X(t)$, thus implying a frequency-nonselective channel appropriate for a narrow-band wireless channel or a single subchannel of a multicarrier system. The Gaussian wide-sense stationary uncorrelated scattering (GWSSUS) fading model [19] is assumed, where the independent component Gaussian processes are zero mean with autocorrelation function $R_X(\tau)$; that is, $X(t) = X_R(t) + jX_I(t)$ is complex Gaussian, where $X_R(t)$ and $X_I(t)$ are the respective real and imaginary parts of $X(t), j = \sqrt{-1}$, $E[X_R(t)] = E[X_I(t)] = 0$, and $E[X_R(t)X_R(t+\tau)] = E[X_I(t)X_I(t+\tau)] = R_X(\tau)$. Thus, the channel is modeled as a Rayleigh fading channel, which is appropriate for narrow-band mobile systems or indoor systems without a line-of-sight component [20]. If there is a line-of-sight component, the fading in indoor systems is Rician [21], [22], but it will be clear from the characterization of the effects of the channel variation that extensions to the Rician case are trivial. Furthermore, it can be shown that when there is uncertainty about the value of the Rician factor, the Rayleigh assumption will lead to signaling for the worst case. It is also assumed that $X(t)$ varies slowly enough that analysis can be performed by assuming it is constant over the support of the pulse $p(t)$ of a single symbol. It is believed that this assumption will be true in any system where the use of adaptive coding is considered; however, as in the consideration of Rician fading, the modifications when this assumption is altered are conceptually minimal. Coherent reception with perfect carrier phase estimation and perfect fading value estimation *at the receiver* is assumed throughout this work.

The key difference between the system model of Fig. 1 and that of a standard communication system is the availability at the transmitter of the vector $\hat{\boldsymbol{X}} = \hat{\boldsymbol{X}}_R + j\hat{\boldsymbol{X}}_I$ of channel fading estimates, where $\tau_{i+1} > \tau_i, \forall i$. The availability of true channel estimates as opposed to estimates of a filtered version of the channel is based on the assumption that the channel varies slowly enough to be assumed constant over the duration of a symbol (and thus estimation) period. These fading estimates can be obtained via literal feedback of measured fading values from the receiver or can be estimated using any signal sent from the current receiver to the current transmitter: a pilot signal at the end of initial handshaking, packet acknowledgment signals, or data sent from the current receiver during the previous slot in a time-division duplex (TDD) system.

For a specific example of such an adaptive system, consider the transmission of data from a mobile terminal to the access point in a wireless local area network (WLAN) that employs a multicarrier strategy. The system protocol generally involves handshaking that ends with the access point sending the mobile terminal a message that informs the mobile terminal that it can begin to send data. The mobile terminal can use this message sent from the access point to measure the current fading on each subcarrier and to prepare its transmitted signal per the recipe in this paper. The mobile terminal then codes the (now outdated) channel measurements for each subcarrier with a low-rate nonadaptive code, interleaves the resulting coded bits across subcarriers to obtain frequency diversity, and sends the resulting bits as a prelude to the actual coded data. The access point decodes the channel measurements for each subcarrier, determines the sizes of the signal sets employed on each subcarrier at each delay by running an algorithm on the channel measurements identical to that being employed at the mobile terminal, and then decodes the data. The overhead for the transmission of the channel measurements will be considered briefly in Section IV-C.

B. Characterization of the Channel Variation

Adaptive signaling employs (perhaps imperfect) knowledge at the transmitter of the current fading values to select a signal set from which to draw the current symbol. Consider the choice of the signal set for the kth symbol. Recalling the assumption that analysis can assume the channel fading is constant over the duration of a single symbol, let $Y = |X(kT_s)|$ be the amplitude of the fading that multiplies the kth transmitted symbol. Assume temporarily that $R_X(\tau)$ is known at the transmitter, and let N be the number of outdated estimates employed. Define the N by N autocorrelation matrix $\Sigma_{\hat{\boldsymbol{X}}}$ of the real or imaginary component of the vector of outdated channel estimates as

$$\Sigma_{\hat{\boldsymbol{X}}} = E\left[\hat{\boldsymbol{X}}_R \hat{\boldsymbol{X}}_R^T\right] = E\left[\hat{\boldsymbol{X}}_I \hat{\boldsymbol{X}}_I^T\right]$$

and the length N column correlation vector $\boldsymbol{\rho}$ as

$$\boldsymbol{\rho} = E\left[\hat{\boldsymbol{X}}_R X_R(kT_s)\right] = E\left[\hat{\boldsymbol{X}}_I X_I(kT_s)\right].$$

Since the focus here is on characterizing the effects of the channel variation independent of the channel estimation algorithm employed, assume that the outdated fading estimates are made perfectly; that is, $\hat{X}(kT_s - \tau_i) = X(kT_s - \tau_i), i = 1, \cdots, N$. Since $X_R(kT_s)$ and the variables in $\hat{\boldsymbol{X}}_R$ are jointly Gaussian, $X_R(kT_s)$ is Gaussian when conditioned on $\hat{\boldsymbol{X}}_R = \boldsymbol{x}_R$ with mean $\boldsymbol{\rho}^T \Sigma_{\hat{\boldsymbol{X}}}^{-1} \boldsymbol{x}_R$ and variance $\sigma^2 = R_X(0) - \boldsymbol{\rho}^T \Sigma_{\hat{\boldsymbol{X}}}^{-1} \boldsymbol{\rho}$ (see, for example, [23, p. 216]). Likewise, $X_I(kT_s)$ is Gaussian when conditioned on $\hat{\boldsymbol{X}}_I = \boldsymbol{x}_I$ with mean $\boldsymbol{\rho}^T \Sigma_{\hat{\boldsymbol{X}}}^{-1} \boldsymbol{x}_I$ and variance $\sigma^2 = R_X(0) - \boldsymbol{\rho}^T \Sigma_{\hat{\boldsymbol{X}}}^{-1} \boldsymbol{\rho}$. Thus, $Y = |X(kT_s)|$ is Rician when conditioned on $\hat{\boldsymbol{X}}$ with conditional probability density function given by

$$p_{Y|\hat{\boldsymbol{X}}}(y|\boldsymbol{x}) = \frac{y}{\sigma^2} e^{-(y^2 + s^2/2\sigma^2)} I_0\left(\frac{ys}{\sigma^2}\right) \qquad y \geq 0 \quad (1)$$

where $I_0(\cdot)$ is the zeroth-order modified Bessel function and the noncentrality parameter is given by $s^2 = (\boldsymbol{\rho}^T \Sigma_{\hat{\boldsymbol{X}}}^{-1} \boldsymbol{x}_R)^2 + (\boldsymbol{\rho}^T \Sigma_{\hat{\boldsymbol{X}}}^{-1} \boldsymbol{x}_I)^2$. Note that σ^2 is the mean squared error of a minimum mean squared error (MMSE) predictor of the in-phase or quadrature fading value of interest, and s^2 is the sum of the squares of the MMSE prediction of $X_R(kT_s)$ and the

MMSE prediction of $X_I(kT_s)$. The result of this section can be reconciled with the approach of [15] using the partitioned matrix inversion identity [24, p. 18].

For interpretation of (1), consider the $N = 1$ case. Defining $\rho = R_X(\tau_1)$, and normalizing such that $E[(X_R(kT_s))^2] = E[(X_I(kT_s))^2] = 1$,[1] one obtains $s^2 = |\boldsymbol{x}|^2\rho^2$ and $\sigma^2 = 1 - \rho^2$. Define the specular-to-diffuse component (Rician) factor of (1) by $K = (s^2/2\sigma^2) = (|\boldsymbol{x}|^2\rho^2/2(1-\rho^2))$. For a fixed correlation ρ, the effective channel conditioned on a single outdated fading estimate varies from a low-SNR nearly Rayleigh fading channel to a high-SNR strongly Rician fading channel, with the type of fading dependent on the value of the outdated estimate. A similar interpretation holds for $N > 1$.

C. Robustness to Autocorrelation Function Uncertainties

If $R_X(\tau)$ is known exactly at the transmitter, using a large number of outdated fading channel estimates yields a linear MMSE predictor for the current in-phase and quadrature fading values that works well (i.e., $\sigma^2 \approx 0$) for systems with modest mobility; however, in wireless systems, $R_X(\tau)$ is generally not known at the transmitter. Not only does the shape of the autocorrelation function vary over different channels [22, pp. 88–89], but the Doppler frequency is unknown. Spectral estimation relies on a long history of channel estimates and cannot provide a guarantee of performance. Nonadaptive codes for slowly fading channels are not affected by the autocorrelation function of the channel fading process if sufficient interleaving is employed; thus, if adaptive codes are going to be compared to nonadaptive codes, there must be a guarantee of performance across possible autocorrelation functions. This leads to the following definition.

Definition: An adaptive signaling scheme is *strongly robust* with respect to $\mathcal{R}$ if it operates below the desired bit error probability for all autocorrelation functions $R_X(\tau) \in \mathcal{R}$.

The class $\mathcal{R}$ can be determined from propagation measurements. For example, for the 900-MHz indoor channel with walking speed mobility, where the shape of the autocorrelation function represents a more low-pass process than that corresponding to the mobile autocorrelation function $R_X(\tau) = J_0(2\pi f_d \tau)$, and the Doppler frequency f_d does not exceed 6 Hz [21], [22], a suitable definition for $\mathcal{R}$ might be all functions in $\{R_X(\tau)\colon R_X(\tau) \geq J_0(2\pi 6\tau)\}$ that satisfy the properties of an autocorrelation function, where $J_0(\cdot)$ is the zeroth-order Bessel function of the first kind.

In this paper, the design methods considered will guarantee strongly robust performance for each observation value $\hat{\boldsymbol{X}} = \boldsymbol{x}$, which guarantees strong robustness of the method. However, in general, the design of strongly robust schemes can be complicated if more than a single outdated fading estimate is employed (i.e., $N > 1$). Thus, in this paper, the methods introduced will be applicable to two cases:

1) strongly robust adaptive signaling with $N = 1$;
2) adaptive signaling with arbitrary N when $R_X(\tau)$ is known.

[1] Note that this "simplification" will make the average received energy twice the average transmitted energy, which will be accounted for below.

Design for the second case is a simplification of design for the first case; thus, strongly robust adaptive signaling with a single outdated fading estimate will be considered. The $N = 1$ case is of considerable practical interest, in particular for bursty high-speed data applications, where only a single fading estimate is generally available from initial handshaking. Furthermore, if the two-sided exponential function cannot be excluded from $\mathcal{R}$, the maximum error probability of a scheme can be no less than that obtained when the autocorrelation function of the channel fading is the worst case two-sided exponential autocorrelation function, where no performance is gained by increasing the number of estimates beyond $N = 1$.

III. Adaptive Uncoded QAM for Time-Varying Channels

In this section, strongly robust adaptive uncoded modulation will be designed. Per Section II-C, only the strongly robust $N = 1$ case will be considered, which requires performance to be guaranteed for all $\rho \in [\rho_{\min}, 1]$, where $\rho_{\min}$ is the minimum value of $R_X(\tau_1)$. For the example of Section II-C, $\rho_{\min} = J_0(2\pi 6\tau_1)$.

A. Design Rules

Assume for illustrative purposes that the set of candidate signal sets are 0-QAM (no data transmitted), 2-QAM, 4-QAM, 16-QAM, and 64-QAM with two-dimensional Gray mapping as considered in [13]. Temporarily, no energy adaptation is employed; in other words, the average energy of the signal constellation employed will be constant across symbols. Although the standard water-pouring energy allocation method can be employed [2], [13], [25], it is desirable to have a method that can be readily adapted to adaptive trellis-coded modulation systems operating over the time-varying channel; the proposed energy adaptation scheme will be presented in Section III-B.

Let P_b be the target bit error probability for the system, which operates at the average received signal-to-noise ratio E_s/N_0, where E_s is the average received energy per QAM symbol. Specification of the adaptive transmitter requires finding $\tilde{M}(h), \forall h$, where $\tilde{M}(h)$ is the number of signals in the QAM signal set employed when $|\hat{X}(kT_s - \tau_1)| = h$. If $\tilde{M}(h)$ is chosen such that P_b is maintained for each h

$$\tilde{M}(h) = \max\left\{M\colon \sup_{\rho_{\min}\leq\rho\leq 1} \tilde{P}_M\left(\frac{E_s}{N_0}, h, \rho\right) \leq P_b\right\} \quad (2)$$

where $\tilde{P}_M((E_s/N_0), h, \rho)$ is defined as the bit error probability of the M-QAM signal set at average received SNR E_s/N_0 when $R_X(\tau_1) = \rho$ and $|\hat{X}(kT_s - \tau_1)| = h$. Assume that maximum likelihood symbol detection given the current channel fading amplitude is employed on the samples of the matched filter output at the receiver. Conditioned on the value y of the current fading amplitude, the probability of bit error for M-QAM (which is upper bounded by the symbol error probability), $M > 2$, is upper bounded as [26, p. 225]

$$P_M\left(y^2\frac{E_s}{N_0}\right) \le 4Q\left(\sqrt{\frac{3}{2(M-1)}\frac{E_s}{N_0}y^2}\right) \le 2\exp\left(-\frac{3}{4(M-1)}\frac{E_s}{N_0}y^2\right) \quad (3)$$

where $Q(x) \triangleq (1/\sqrt{2\pi})\int_x^\infty e^{-(u^2/2)}\,du$. The fact that the average received energy is twice the average transmitted energy has been recalled, and the second inequality is obtained by employing the bound $Q(\alpha) \le \frac{1}{2}e^{-(\alpha^2/2)}$. Results indicate that employing (3) for the calculation of $\tilde{P}_M((E_s/N_0), h, \rho)$ results in very conservative designs, as do the results from employing the generally tighter bound $Q(\alpha) \le (1/\sqrt{2\pi}\alpha)e^{-(\alpha^2/2)}$ to obtain the analog of (3). Thus, the tight approximation [13]

$$P_M\left(y^2\frac{E_s}{N_0}\right) \approx 0.2\exp\left(-\frac{3}{4(M-1)}\frac{E_s}{N_0}y^2\right) \quad (4)$$

will be employed for all M for much of the design work for uncoded systems. Note that (4) is not an upper bound in all cases considered as it will be implicitly employed for all y and applied to $M = 2$. Using (4) yields

$$\begin{aligned}\tilde{P}_M&\left(\frac{E_s}{N_0}, h, \rho\right)\\ &= E\left[P_M\left(Y^2\frac{E_s}{N_0}\right)\Bigg|\left|\hat{X}(kT_s-\tau_1)\right| = h\right]\\ &\approx \begin{cases} \dfrac{0.2\exp\left[-\dfrac{h^2\rho^2}{2(1-\rho^2)}\left(1-\dfrac{1}{1+\frac{3}{2}\frac{E_s}{N_0}\frac{(1-\rho^2)}{(M-1)}}\right)\right]}{1+\frac{3}{2}\frac{E_s}{N_0}\frac{(1-\rho^2)}{(M-1)}} & \rho < 1\\ 0.2\exp\left(-\dfrac{3}{4}\dfrac{E_s}{N_0}\dfrac{h^2}{(M-1)}\right) & \rho = 1\end{cases}\end{aligned} \quad (5)$$

where the approximation is obtained by substituting (1) and (4) into the equality and evaluating the expectation over Y using [27, 6.614.3].

From (2), (5) must be evaluated at its supremum on $\rho \in [\rho_{\min}, 1]$. Since the right side of (5) is a continuous function on this closed interval, it achieves its maximum on this interval at a point which will be denoted ρ^*. The following solution is found by standard calculus techniques. Let

$$\tilde{\rho} = \begin{cases} 0 & h \ge \sqrt{2}\\ \sqrt{\left(1+\dfrac{2(M-1)}{3}\dfrac{N_0}{E_s}\right)\dfrac{(2-h^2)}{2}} & 0 \le h \le \sqrt{2}.\end{cases}$$

The worst case autocorrelation is then given by

$$\rho^* = \begin{cases}\rho_{\min} & \tilde{\rho} \le \rho_{\min}\\ \tilde{\rho} & \rho_{\min} < \tilde{\rho} < 1\\ 1 & 1 \le \tilde{\rho}.\end{cases} \quad (6)$$

The signal set is specified using (5) and (6) in $\tilde{M}(h) = \max\{M\colon\ \tilde{P}_M((E_s/N_0), h, \rho^*) \le P_b\}$. Note that $\tilde{M}(h)$ is nondecreasing in h. Thus, the adaptive scheme can be specified by the values h_m, $m = 2, 4, 16, 64$, where h_m is defined as the threshold such that for $h \ge h_m$, m-QAM can be employed. For a fixed ρ, (5) can be solved explicitly to find h_m, but since ρ^* depends on h, the resulting equation is not useful in general if strong robustness is desired. However, in many situations, particularly for small P_b, it can be shown analytically that one need only consider $\rho = \rho_{\min}$, and thus the thresholds can be found explicitly.

B. Energy Adaptation

The rate of the method of the previous section is limited by the following reason: for all h such that $h_m < h < h_{2m}$, the estimate is more favorable than that required to use m-QAM but not favorable enough to use $(2m)$-QAM. A solution to this problem is to employ energy adaptation. First, a signal set is chosen according to the previous section with no energy adaptation. Then (5) and (6) are used to decide the minimum energy required to maintain P_b given the channel estimate h, thus employing a method similar to the power-pruning of [3]. This minimum energy is employed for the current symbol and the remainder of the energy is saved for the next symbol. This algorithm, which requires at most a two-dimensional table look-up, will be employed for all results below.

C. Numerical Results

In this section, although design is done for the worst case ρ for each h as prescribed by (2), the generation of the current channel fading value in the simulations is done assuming that $\rho = \rho_{\min}$, which leads to the worst case performance when averaged across all estimates for each case considered here.

Fig. 2 displays the probability of bit error of the scheme developed in Section III-A and the probability of bit error for a scheme that is designed based on a static channel (i.e., one that assumes $\rho = 1$). When designing for the static channel at $P_b = 10^{-5}$, the more conservative estimate $P_2(y^2(E_s/N_0)) \approx \frac{1}{2}\exp(-(E_s/2N_0)y^2)$ is employed, as the approximation of (4) is very poor at this point. Note that the probability of error targets are missed considerably by the scheme that assumes a static channel when there is even modest channel variability, particularly for low target bit error probabilities; this is due to (1) being a Rician density. Although uncoded systems will generally not operate at bit error probabilities as low as $P_b = 10^{-5}$, trellis-coded modulation schemes will operate in such ranges, and the results of Fig. 2 apply to the parallel branches of such schemes, thus indicating that adaptive trellis-coded modulation schemes should take into account the variability of the wireless channel.

One method of achieving the target bit error probability by the scheme that assumes a static channel is to add energy mar-

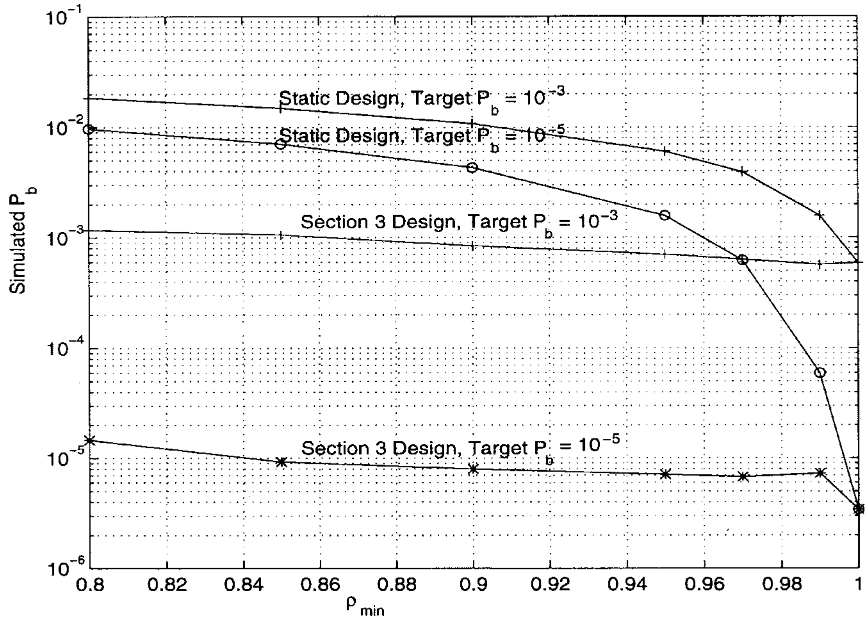

Fig. 2. Simulated probability of bit error of adaptive uncoded QAM design schemes versus $\rho_{\min}$ at average received signal-to-noise ratio $(E_s/N_0) = 15$ dB per QAM symbol. For target $P_b = 10^{-3}$ curves, each data point is the result of 10^6 trials. For target $P_b = 10^{-5}$ curves, each data point is the result of 10^7 trials. The static design scheme assumes a time-invariant channel, and it is observed that it misses the target bit error probability by a significant margin.

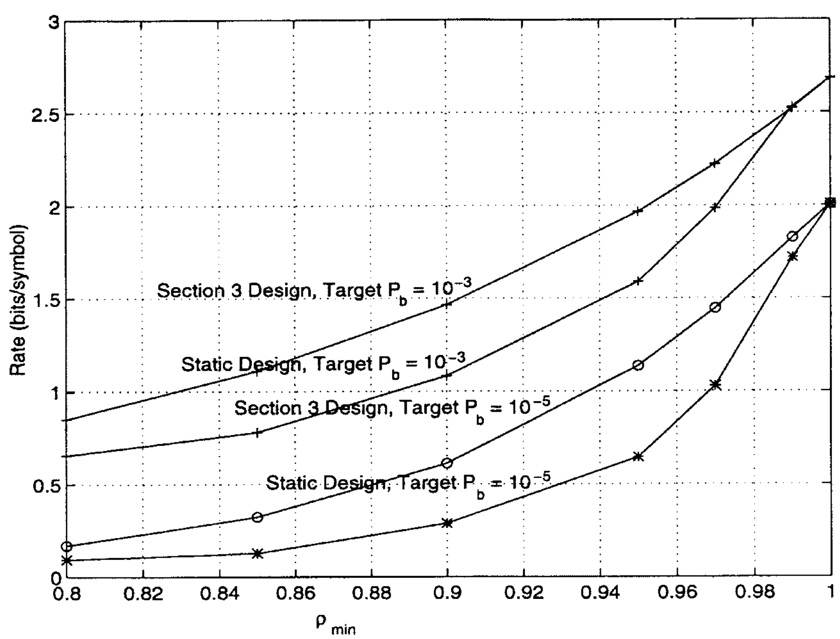

Fig. 3. Simulated rate of adaptive uncoded QAM design schemes versus $\rho_{\min}$ at average received signal-to-noise ratio $(E_s/N_0) = 15$ dB per QAM symbol. For target $P_b = 10^{-3}$ curves, each data point is the result of 10^6 trials. For target $P_b = 10^{-5}$ curves, each data point is the result of 10^7 trials. The static design scheme employs an energy margin to meet the target bit error probability or the bit error probability of the scheme of Section III, whichever is larger. The rate increases by employing the scheme of Section III are not only substantial but also are conservative estimates, as the simulated bit error probabilities of these schemes are no greater than those of the static design scheme at every point and generally are substantially smaller.

gin during the energy adaptation process; that is, instead of using the energy that is predicted to give the proper performance, add a fixed margin (in decibels) to this energy at each symbol until simulation results indicate that the probability of bit error is at the desired point. If this is done to the scheme that assumes a static channel until it achieves the target P_b or the simulated P_b of the method of Section III-A, whichever is *larger*, the resulting rates are as shown in Fig. 3. The results demonstrate conservatively the rate increase possible by considering the variation of the channel explicitly in system design. Finally, Figs. 4 and 5 plot the rate of the scheme of Section III-A versus signal-to-noise ratio for a number of values of $\rho_{\min}$ at target bit error probabilities $P_b = 10^{-3}$ and $P_b = 10^{-5}$, respectively. The close correspondence of the $\rho_{\min} = 1.0$ curve on Fig. 4 to the results in [13] suggests that the flexible energy allocation scheme of Section III-B is effective.

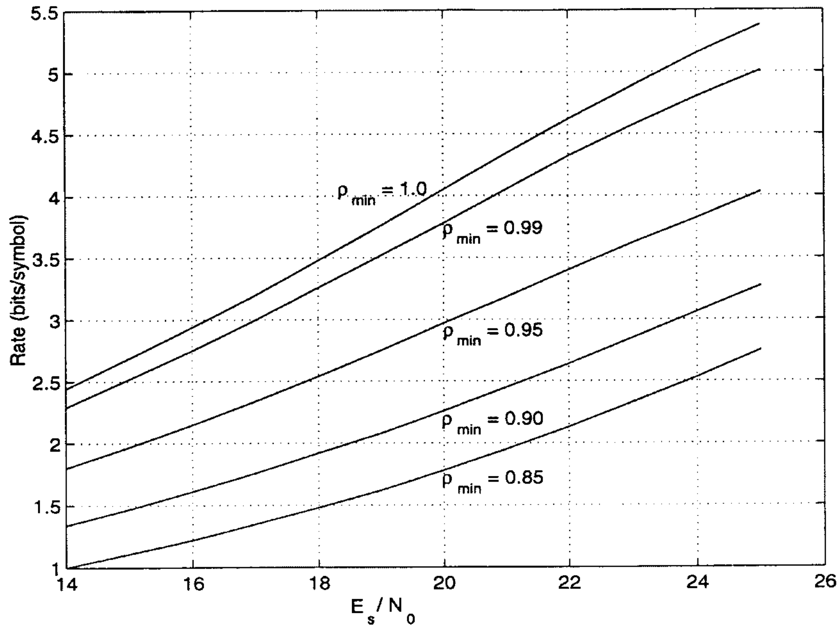

Fig. 4. Rate in bits per QAM symbol versus average received SNR E_s/N_0 per QAM symbol for the scheme of Section III at target error probability $P_b = 10^{-3}$. Each data point is the result of 10^6 trials. All data points meet the target bit error probability, except for approximately one-half of the points on the $\rho_{\min} = 0.85$ curve; in this latter case, all simulated bit error probabilities are less than 1.1 P_b. Note the more conservative results in [15] are due to the fact that 0-QAM signals were sent at average energy E_s in [15].

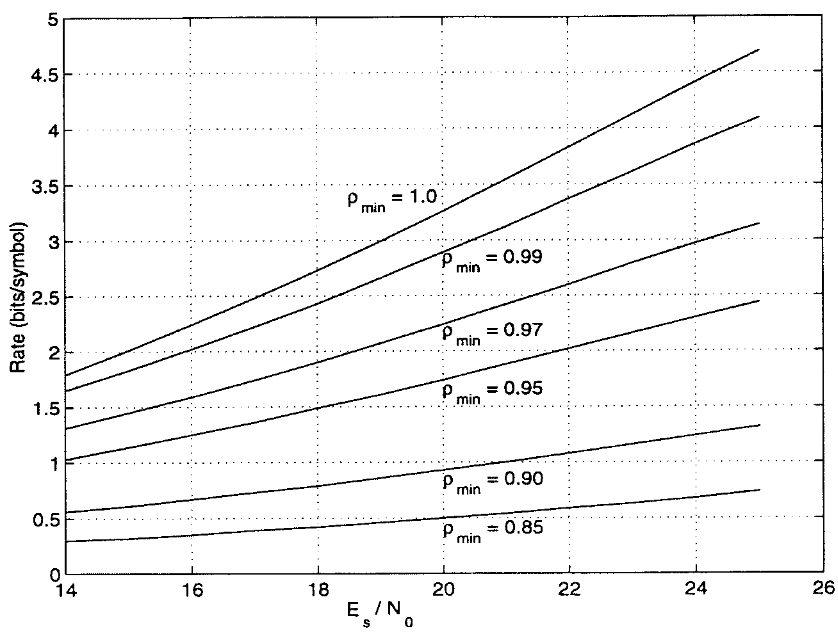

Fig. 5. Rate in bits per QAM symbol versus average received SNR E_s/N_0 per QAM symbol for the scheme of Section III at target bit error probability $P_b = 10^{-5}$. Each data point is the result of 10^7 trials. All data points meet the target bit error probability, except for approximately one-half of the data points on the $\rho_{\min} = 0.85$ curve; in this latter case, all simulated bit error probabilities are less than 1.3 P_b. Note the more conservative results in [15] are due to the fact that 0-QAM signals were sent at average energy E_s in [15].

IV. Adaptive Trellis-Coded Modulation for Time-Varying Channels

A. Paradigm for Adaptive Trellis-Coded Modulation

Adaptive signaling is viewed as a powerful method for highly bandwidth efficient modulation, and thus the focus here is on adaptive trellis-coded modulation as in [7], [9], and [10]. The adaptive trellis-coded modulation paradigm that will be followed is shown in Fig. 6. For this paradigm, the number of subsets 2^n is kept constant across all signal sets; only the number of signals per subset is adapted. The operation of the trellis coder is as follows: sets of k_1 information bits are input to the convolutional coder, which produces n bit subset selectors that are interleaved. The kth transmitted symbol z_k is then chosen as follows. The subset selector $\tilde{S}_k$ is taken from the interleaver output. The signal constellation $\tilde{\zeta}_k$ (and thus the subsets) is chosen from the allowable signal sets

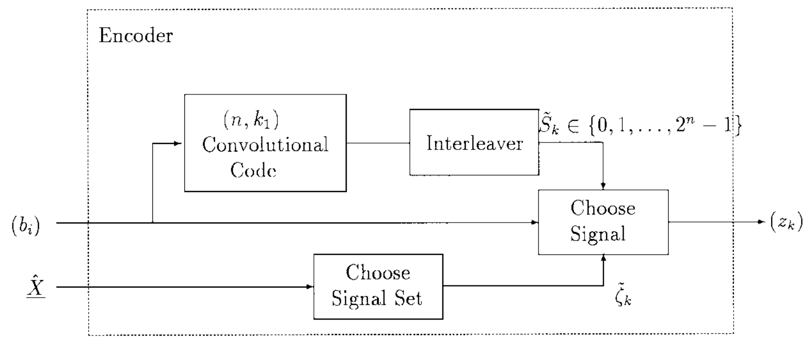

Fig. 6. The adaptive trellis coding paradigm.

$\zeta_1, \zeta_2, \cdots, \zeta_L$ (with possibly scaling for energy adaptation) based on the information about the current fading provided by $\hat{\boldsymbol{X}}$. Finally, the appropriate number of data bits (base two logarithm of the size of a subset of $\tilde{\zeta}_k$) are used to choose z_k from subset $\tilde{S}_k$ of $\tilde{\zeta}_k$. There are two design considerations that must be considered: how to design the trellis structure and subset mapping, and how to choose the signal set $\tilde{\zeta}_k$ based on $\hat{\boldsymbol{X}}$. The first item is solved well in the literature for nonadaptive coding; thus, it will be assumed here that there exists a nominal (nonadaptive) trellis-coded modulation scheme with signal set ζ_0 that achieves the desired bit error probability P_b. The same trellis structure and subset mapping of this nominal scheme will be employed.

Consider the choice of the signal set $\tilde{\zeta}_k$. Let $S_l^r, r = 0, \cdots, 2^n - 1$ be the rth subset of signal set ζ_l, where the subsets are found by standard Ungerboeck partitioning [28]. Let g_l be the *intrasubset* minimum normalized Euclidean distance squared of ζ_l, and let $u_l^{m,r}, m = 0, \cdots, 2^n - 1, r = 0, \cdots, 2^n - 1$ be the *intersubset* minimum normalized Euclidean distance squared between signals in subset m and r of ζ_l. In other words, letting $d_E(x, v)$ be the Euclidean distance between x and v

$$g_l = \min_{S_l^r} \min_{x, v \in S_l^r, x \neq v} \frac{d_E^2(x, v)}{E_l}$$

and

$$u_l^{m,r} = \min_{x \in S_l^m, v \in S_l^r} \frac{d_E^2(x, v)}{E_l}$$

where E_l is the average energy of signal set ζ_l. When the current channel fading value $Y = y$ is assumed to be known exactly from $\hat{\boldsymbol{X}}$, the following recipe can be employed: let $\tilde{\zeta}_k$, the signal set for the kth transmitted signal, be the largest signal set ζ_l such that $y^2 g_l \geq g_0$ and $y^2 u_l^{m,r} \geq u_0^{m,r}, \forall m, r$. A similar method for the case when the current channel fading value is known exactly is formalized in [10], where the signal set point density of the nominal scheme is maintained, thus preserving the minimum intrasubset and intersubset distances. This technique allows the system to not signal during deep fades, thus mitigating the key inhibiting factor to fading channel communications. The form of (1) as a Rician density questions whether the same gains will still be possible for adaptive trellis-coded modulation on a time-varying wireless channel.

B. Design Rules

Throughout this section, perfect interleaving of the subset selectors $\{\tilde{S}_k\}$ will be assumed. This may sound like it does not fit the model, which inherently lacks large time diversity at reasonable delay. However, if the coding of the subchannels of a multicarrier system is being considered, the desired diversity can be obtained by interleaving in frequency. It will also be assumed that the decoder performs maximum likelihood symbol sequence estimation with the samples of the output of the receiver matched filter as input and knowledge of the channel fading values.

Per the paradigm, a nominal (nonadaptive) trellis-coded modulation scheme is assumed that employs signal set ζ_0 and provides the desired bit error probability P_b at some signal-to-noise ratio $(E_s/N_0)_0$ for the environment (Rayleigh fading or AWGN, for example) for which the nominal scheme was designed. When the channel impairment faced by the adaptive scheme is the same as the nominal code, maintaining intersubset and intrasubset Euclidean distances is sufficient to obtain the probability of error of the nominal code. However, the conditional distribution of (1) varies from Rayleigh to strongly Rician, thus revealing that there will not be a nominal code to which the type of fading is matched for all outdated channel estimates. Hence, the design method employed is the maintenance of the intersubset and intrasubset *differences*, which will be defined as the ability to decide between symbols from different subsets and the same subset, respectively.

To consider the motivation behind employing the maintenance of intersubset and intrasubset differences, assume that the trellis of the nominal trellis-coded modulation scheme has no parallel branches. This will be a desirable (but not necessary) property of the nominal scheme, as when h is small, (1) approaches the density of a Rayleigh random variable, and this trellis structure allows data to still be transmitted in many cases; when h is large, (1) is the density of a strongly Rician random variable, and parallel branches can be added. Suppose that a signal set ζ_l is being considered for use when $|\hat{X}(kT_s - \tau_1)| = h$ is observed. Consider first maintaining the bit error probability of the nominal code for the convolutionally coded bits in the adaptive trellis coded modulation scheme. For the nominal scheme, the probability of the Viterbi algorithm choosing the symbol path $\hat{\boldsymbol{w}} = (\hat{w}_1, \hat{w}_2, \cdots)$ when the symbol path $\boldsymbol{w} = (w_1, w_2, \cdots)$ was the correct path can be tightly bounded by considering the

Chernoff bound [29]

$$P(\boldsymbol{w} \rightarrow \hat{\boldsymbol{w}}) \leq \prod_{i \in \eta(\boldsymbol{w}, \hat{\boldsymbol{w}})} D_0(w_i, \hat{w}_i) \tag{7}$$

where $\eta(\boldsymbol{w}, \hat{\boldsymbol{w}})$ is the index set of positions where $\boldsymbol{w}$ and $\hat{\boldsymbol{w}}$ differ, and $D_0(w, v)$ is the Bhattacharyya bound [30] on the error probability of choosing w when v is correct for the nominal code. For example, $D_0(x, v) = e^{-(d_E^2(w,v)/4E_s)(E_s/N_0)_0}$ if the nominal code was designed for the AWGN channel and $D_0(x, v) = 1/(1 + (d_E^2(w, v)/4E_s)(E_s/N_0)_0$ if the nominal code was designed for a Rayleigh fading channel. To approximately maintain the bit error probability of the nominal scheme for the convolutionally coded bits in the adaptive scheme, the right side of (7) will be maintained for all paths $\hat{\boldsymbol{z}}$ and $\boldsymbol{z}$ in the adaptive scheme that traverse the same sequence of states in the trellis as $\hat{\boldsymbol{w}}$ and $\boldsymbol{w}$, respectively. To accomplish this, it is sufficient that

$$\sup_{\rho \geq \rho_{\min}} E_Y \left[e^{-(d_E^2(Y z_i, Y \hat{z}_i)/8E_s)(E_s/N_0)} \middle| \left| \hat{X}(kT_s - \tau_1) \right| = h \right] \leq D_0(w_i, \hat{w}_i) \tag{8}$$

for all possible values of w_i and $\hat{w}_i$, where z_i and $\hat{z}_i$ are the two symbols from ζ_l such that:

1) z_i is from the same subset as w_i;
2) $\hat{z}_i$ is from the same subset as $\hat{w}_i$;
3) z_i and $\hat{z}_i$ are separated by the minimum Euclidean distance of all signals that satisfy conditions 1 and 2.

This is equivalent to maintaining the minimum intersubset differences between all subsets or, more precisely, for all $m = 0, \cdots, 2^n - 1$ and $r = 0, \cdots, 2^n - 1$, $m \neq r$, guaranteeing that

$$\sup_{\rho \geq \rho_{\min}} \int_0^\infty e^{-y^2 u_l^{m,r}(E_s/8N_0)} p_{Y|h}(y|h)\, dy \leq D_0(0, u_0^{m,r} E_s). \tag{9}$$

The form of the left side of (9) is identical to equations of Section III-A, and thus its manipulation is omitted.

Conditioned on the assumption that there are no errors in the sequence of trellis states chosen by the Viterbi algorithm, achieving the desired bit error probability for uncoded bits in the adaptive trellis-coded modulation scheme is easy to consider. If the signal set ζ_l implies parallel branches in the trellis, an intrasubset distance will be required that implies an upper bound to the probability of choosing the wrong signal in a subset is kept below the desired bit error probability

$$\sup_{\rho \geq \rho_{\min}} \int_0^\infty Q\left(\sqrt{y^2 g_l \frac{E_s}{4N_0}}\right) p_{Y|h}(y|h)\, dy \leq P_b \tag{10}$$

which is maintained if an upper bound to the left side is maintained. The upper bound is obtained by employing the bound $Q(\alpha) \leq \frac{1}{2} e^{-(\alpha^2/2)}$ to obtain the condition

$$\sup_{\rho \geq \rho_{\min}} \frac{1}{2} \int_0^\infty e^{-y^2 g_l (E_s/8N_0)} p_{Y|h}(y|h)\, dy \leq P_b. \tag{11}$$

The form of the left side of (11) is identical to equations of Section III-A, and thus its manipulation is omitted.

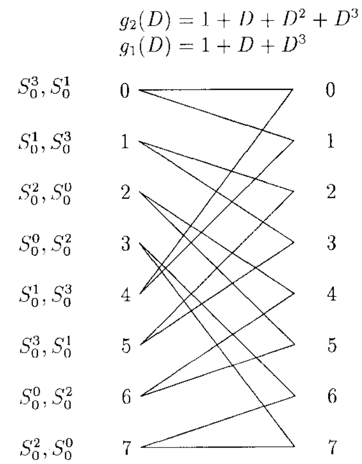

Fig. 7. Trellis for 2 bits per QAM symbol eight-state I-Q Code from [15].

C. Numerical Results and Analysis

For a nominal trellis-coded modulation scheme, the 2 bits per QAM symbol eight-state I-Q code from [18], which operates at the desired bit error probability $P_b = 10^{-5}$ at average received SNR $(E_s/N_0)_0 = 17.5$ dB per QAM symbol on a Rayleigh fading channel, is employed. In this nominal scheme, a 16-QAM symbol is formed from the output of two coded modulators that employ rate one-half convolutional codes and 4-ary ASK. There are no parallel branches in the trellis diagram (shown in Fig. 7) of each of these component coded 4-ary ASK schemes, which employ the subsets $S_0^0 = \{-3\}$, $S_0^1 = \{-1\}$, $S_0^2 = \{+1\}$, and $S_0^3 = \{+3\}$. Let the candidate signal sets $\zeta_1, \zeta_2, \zeta_3, \zeta_4$ be the set of M-ary ASK constellations, $M = 4, 8, 16, 32$, respectively. Define the four subsets for each of $\zeta_1, \zeta_2, \zeta_3, \zeta_4$ using standard Ungerboeck partitioning

$$\begin{aligned}
M = 4:\quad & S_1^0 = \{-3\}, S_1^1 = \{-1\}, S_1^2 = \{+1\}, S_1^3 = \{+3\} \\
M = 8:\quad & S_2^0 = \{-7, +1\}, S_2^1 = \{-5, +3\} \\
& S_2^2 = \{-3, +5\}, S_2^3 = \{-1, +7\} \\
M = 16:\quad & S_3^0 = \{-15, -7, +1, +9\} \\
& S_3^1 = \{-13, -5, +3, +11\} \\
& S_3^2 = \{-11, -3, +5, +13\} \\
& S_3^3 = \{-9, -1, +7, +15\} \\
M = 32:\quad & S_4^0 = \{-31, -23, -15, -7, +1, +9, +17, +25\} \\
& S_4^1 = \{-29, -21, -13, -5, +3, +11, +19, +27\} \\
& S_4^2 = \{-27, -19, -11, -3, +5, +13, +21, +29\} \\
& S_4^3 = \{-25, -17, -9, -1, +7, +15, +23, +31\}.
\end{aligned}$$

For a given h, (9) can be checked for each of the three possible intersubset differences; however, for this nominal code, the separation between trellis paths leaving and remerging at a node is always separated by the intersubset difference between S_0^0 and S_0^2. Thus, only this difference will be guaranteed and a small energy margin will be employed; it has been found that this results in a better scheme than guaranteeing all of the intersubset differences when the trellis of the nominal scheme

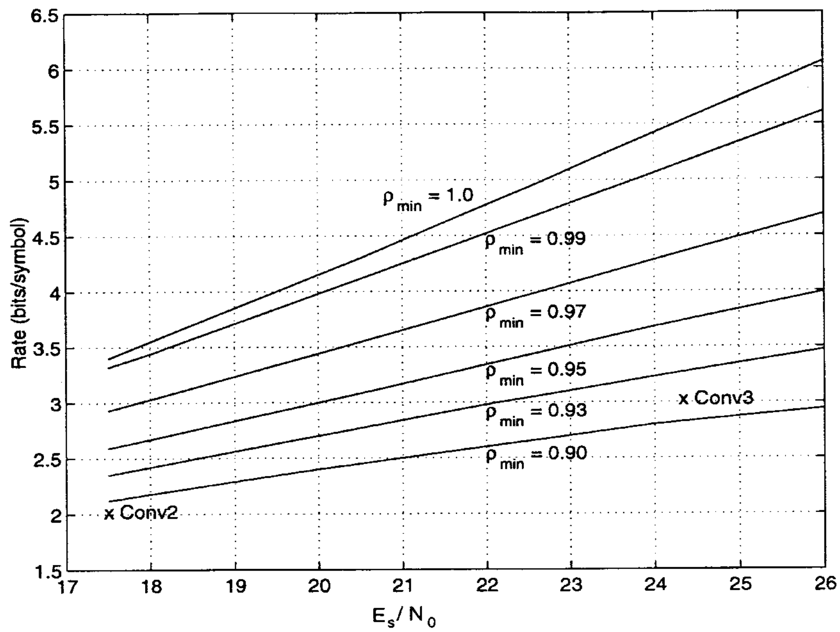

Fig. 8. Simulated rate of adaptive eight-state trellis-coded modulation in bits per QAM symbol as a function of the average received SNR E_s/N_0 per QAM symbol. "Conv2" and "Conv3" correspond to eight-state I-Q codes from [18] that operate at bit error probability 10^{-5}. The simulated P_b for each data point is less than the target 10^{-5} as shown in Table I. Each data point is the result of the simulation of 10^7 QAM symbols.

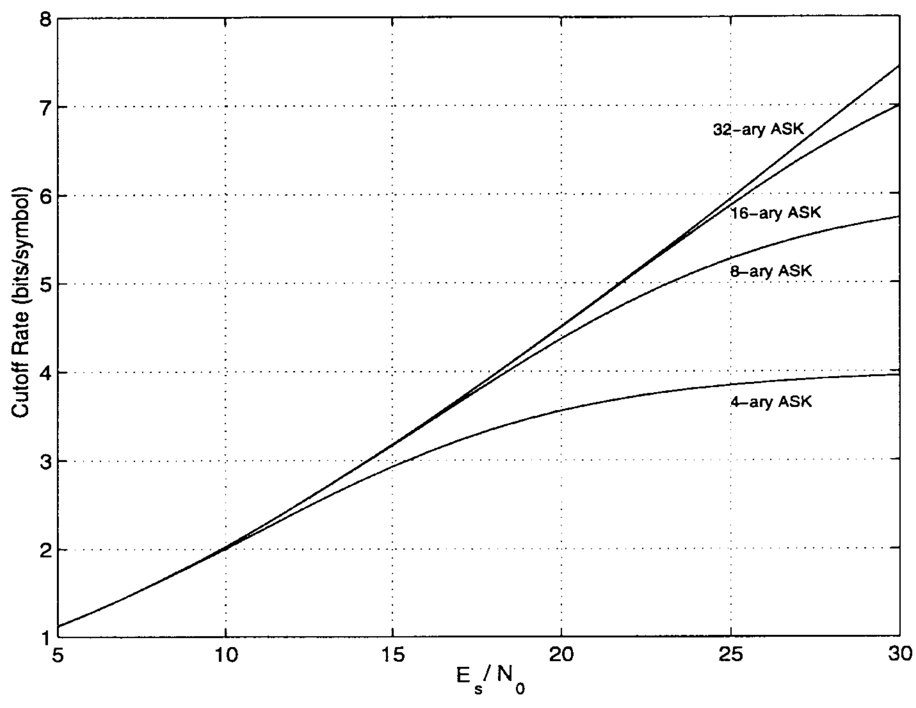

Fig. 9. The cutoff rate in bits per QAM symbol for perfectly interleaved QAM over a Rayleigh fading channel versus the average received SNR E_s/N_0 per QAM symbol, where the QAM signals are formed by two M-ASK signal streams that employ separate encoders and decoders. Note that the cutoff rate shown in [18] is the cutoff rate for interleaved QAM symbols. The cutoff rate when the I and Q dimensions are effectively interleaved is approximately 1 dB lower at the rates of interest and can be found in [31].

has no parallel branches. Making this intersubset guarantee for each l yields h_l^{tr}, the threshold such that for $h \geq h_l^{\mathrm{tr}}$, ζ_l can be employed according to the intersubset differences. Likewise, (11) can be used to find h_l^{par}, the threshold such that for $h \geq h_l^{\mathrm{par}}$, ζ_l can be employed according to the intrasubset differences. The threshold for h above which ζ_l can be used is then given by $\max(h_l^{\mathrm{tr}}, h_l^{\mathrm{par}})$. Energy adaptation will be employed similarly to the method described in Section III-B; the minimum energy that is required to guarantee (9) and (11) is used for this symbol; the remainder of the energy is saved for the next symbol.

Numerical results for the rate of this adaptive trellis-coded modulation scheme are shown in Fig. 8. Despite the time-varying nature of the channel and uncertainties in the

TABLE I
SAMPLE SIMULATED BIT ERROR PROBABILITIES OF ADAPTIVE 8-STATE TRELLIS-CODED MODULATION AS A FUNCTION OF THE AVERAGE RECEIVED SNR $\frac{E_s}{N_0}$ PER QAM SYMBOL EACH ENTRY IS A RESULT OF THE SIMULATION OF 10^7 QAM SYMBOLS

ρ_{min}	Margin	$\frac{E_s}{N_0}$	Simulated P_b
1.0	1.5 dB	18.0 dB	5.93×10^{-6}
		20.0 dB	8.10×10^{-6}
		22.0 dB	9.91×10^{-6}
		24.0 dB	9.52×10^{-6}
		26.0 dB	9.05×10^{-6}
0.99	1.5 dB	18.0 dB	6.79×10^{-6}
		20.0 dB	6.84×10^{-6}
		22.0 dB	7.92×10^{-6}
		24.0 dB	3.91×10^{-6}
		26.0 dB	4.53×10^{-6}
0.97	0.75 dB	18.0 dB	6.39×10^{-6}
		20.0 dB	5.22×10^{-6}
		22.0 dB	3.93×10^{-6}
		24.0 dB	3.78×10^{-6}
		26.0 dB	2.98×10^{-6}
0.95	0.50 dB	18.0 dB	6.82×10^{-6}
		20.0 dB	3.19×10^{-6}
		22.0 dB	4.60×10^{-6}
		24.0 dB	1.90×10^{-6}
		26.0 dB	1.76×10^{-6}
0.93	0.25 dB	18.0 dB	7.21×10^{-6}
		20.0 dB	4.81×10^{-6}
		22.0 dB	5.17×10^{-6}
		24.0 dB	4.45×10^{-6}
		26.0 dB	1.38×10^{-6}
0.90	0.0 dB	18.0 dB	2.75×10^{-6}
		20.0 dB	3.49×10^{-6}
		22.0 dB	1.54×10^{-6}
		24.0 dB	2.09×10^{-6}
		26.0 dB	8.15×10^{-7}

autocorrelation function of the fading process, there is still a substantial bandwidth efficiency gain versus the nominal scheme on which the codes are based, even when only a single outdated fading estimate is employed. The high bandwidth efficiency, use of only a single outdated fading estimate, and robustness of these codes makes them particularly suitable for high-speed data systems. For example, at $\rho_{\min} = 0.99$ and $\rho_{\min} = 0.97$, Fig. 8 reveals approximately a 67 and 50% bandwidth efficiency increase, respectively, over state-of-the-art nonadaptive codes. If the autocorrelation class example of Section II-C is considered, this implies a large throughput gain if data transmission follows channel estimation by 5–10 ms, a time during which a number of packets can be transmitted and acknowledgment received in a high-speed wireless data system. The reader may suppose that the coding gains shown in Fig. 8 are not due to the adaptive coding but instead are due to the larger constellations employed by the adaptive scheme, which have a higher cutoff rate. However, whereas the I-Q codes of [18] operate within a few decibels of the perfectly interleaved 4-ASK (no energy adaptation) cutoff rate with a 64-state trellis, the codes constructed here operate near the (no energy adaptation) cutoff rate shown in Fig. 9 for interleaved 32-ASK with only an eight-state trellis. Thus, although there is very little gain in capacity for adaptive coding [12], adaptive coding does provide significant gain when system complexity is limited, even for a time-varying channel.

Finally, the overhead required to send the outdated channel measurements in the sample WLAN application given in Section II-A is addressed. Assume the multicarrier system described in Section II-A sends 20 000 QAM symbols per second per subcarrier, there is a delay between channel estimation and the start of data transmission of 1 ms, $(E_s/N_0) = 17.5$ dB, and $\rho_{\min} = J_0(2\pi 6\tau_1)$. Per Fig. 8, a nonadaptive system that employs the codes of [18] can send two information bits per QAM symbol at $P_b = 10^{-5}$, which yields 40 000 bits per second per subcarrier. In particular, between 1 ms and 9.2 ms of receiving the signal to send data, the nonadaptive system can transmit 328 information bits per subcarrier. Next, consider the adaptive system. Suppose 8-bit representations of the outdated channel measurement for each subcarrier are coded with a constraint length four, rate 1/2 convolutional code whose output bits are interleaved across subcarriers and then mapped to 4-QAM signal sets; this conservatively results in an error probability at the signal-to-noise ratios of interest that has a negligible effect on the overall system bit error probability. Under this scheme, the overhead for transmission of the channel measurements is 8 QAM symbols per subcarrier, which requires 0.4 ms. At $\tau_1 = 9.2$ ms, $\rho_{\min} = 0.97$. Since $\rho_{\min} = 0.97$ lower bounds $\rho_{\min} = J_0(2\pi 6\tau_1)$ for all $\tau_1 < 9.2$ ms, this implies that a lower bound to the number of information bits per subcarrier that the adaptive system can send between 1 and 9.2 ms is given by (9.2 ms − 1.4 ms) × 2.95 bits/symbol × 20 symbols/ms = 460 bits. The data rate of the adaptive system will actually be significantly higher because of the higher correlation corresponding to smaller delays for symbols sent early in the transmission. Recall that the receiver is able to determine the signal set used at each symbol delay τ_1 because it has knowledge of the outdated channel measurement and employs the same algorithm as the transmitter employed to choose the signal sets. Finally, it is generally not necessary to send the channel measurement for each subcarrier, as the channel response can be strongly correlated in frequency. Thus, if a reduction in overhead is desired, outdated channel estimates for a subset of the subcarriers can be transmitted, and any additional uncertainty due to channel variation over frequency can be addressed in the same manner as the uncertainty due to channel variation over time has been addressed in this paper.

V. CONCLUSIONS

Although the theory of adaptive coding demonstrates large throughput gains over nonadaptive coding, it has been avoided in practice because of perceived problems with the variation of the fading channel between channel estimation and data transmission. In this paper, it has been shown that these perceptions are well-founded: the delay between channel estimation and data transmission greatly alters the nature of the problem. However, it has also been demonstrated that by designing with this variability in mind, the system designer can specify adaptive coding schemes that display large throughput gains over conventional schemes. This was done requiring knowledge of neither the Doppler frequency nor the exact shape of the autocorrelation function of the channel fading process.

ACKNOWLEDGMENT

The author is indebted to G. Ananthaswamy for suggesting the placement of the interleaver in Fig. 6, a placement which was noted independently in [16]. The author would also like to thank K. Chugg, whose early review of this work improved the exposition in Section II-B.

REFERENCES

[1] M. Eyuboglu, C. Forney, P. Dong, and G. Long, "Advanced modulation techniques for V.Fast," *Eur. Trans. Telecommun.*, vol. 4, pp. 243–256, May–June 1993.

[2] I. Kalet, "The multitone channel," *IEEE Trans. Commun.*, vol. 37, pp. 119–124, Feb. 1989.

[3] E. Feig, "Practical aspects of DFT-based frequency division multiplexing for data transmission," *IEEE Trans. Commun.*, vol. 38, pp. 929–932, July 1990.

[4] J. Hayes, "Adaptive feedback communications," *IEEE Trans. Commun. Technol.*, vol. COM-16, pp. 29–34, Feb. 1968.

[5] V. Hentinen, "Error performance for adaptive transmission on fading channels," *IEEE Trans. Commun.*, vol. COM-22, pp. 1331–1337, Sept. 1974.

[6] J. Cavers, "Variable-rate transmission for Rayleigh fading channels," *IEEE Trans. Commun.*, vol. COM-20, pp. 15–22, Feb. 1972.

[7] J. Bingham, "Multicarrier modulation for data transmission: An idea whose time has come," *IEEE Commun. Mag.*, pp. 5–14, May 1990.

[8] B. Vucetic, "An adaptive coding scheme for time-varying channels," *IEEE Trans. Commun.*, vol. 39, pp. 653–663, May 1991.

[9] S. Alamouti and S. Kallel, "Adaptive trellis-coded multiple-phase-shift keying for Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 42, pp. 2305–2314, June 1994.

[10] A. Goldsmith, "Variable-rate coded MQAM for fading channels," in *Proc. IEEE Global Communications Conf.—Communication Theory Miniconf.*, 1994, pp. 186–190.

[11] L. Cimini, Jr., "Performance studies for high-speed indoor wireless communications," *Wireless Personal Commun.*, vol. 2, pp. 67–85, 1995.

[12] A. Goldsmith, "Capacity and dynamic resource allocation in broadcast fading channels," in *Proc. Allerton Conf. Communications, Control, and Computing*, 1995, pp. 915–924.

[13] A. Goldsmith and S. Chua, "Variable-rate variable-power MQAM for fading channels," *IEEE Trans. Commun.*, vol. 45, pp. 1218–1230, Oct. 1997.

[14] M. S. Alouini and A. Goldsmith, "Adaptive M-QAM modulation over Nakagami fading channels," in *Proc. IEEE Global Communications Conf.—Communication Theory Miniconf.*, Nov. 1997, pp. 218–223.

[15] D. Goeckel, "Robust adaptive coded modulation for time-varying channels with delayed feedback," in *Proc. 35th Annu. Allerton Conf. Communication, Control, and Computing*, Sept. 1997, pp. 370–379.

[16] V. Lau and M. Macleod, "Variable rate adaptive trellis coded QAM for high bandwidth efficiency applications under Rayleigh fading channel," in *Proc. 48th Annu. Vehicular Technology Conf.*, May 1998, pp. 348–352.

[17] T. Ericson, "A Gaussian channel with slow fading," *IEEE Trans. Inform. Theory*, vol. IT-16, pp. 353–355, May 1970.

[18] S. Al-Semari and T. Fuja, "I-Q TCM: Reliable communication over the Rayleigh fading channel close to the cutoff rate," *IEEE Trans. Inform. Theory*, vol. 43, pp. 250–262, Jan. 1997.

[19] P. Bello, "Characterization of randomly time-variant linear channels," *IEEE Trans. Commun. Syst.*, vol. COM-11, pp. 360–393, Dec. 1963.

[20] H. Bertoni, W. Honcharenko, L. Maciel, and H. Xia, "UHF propagation prediction for wireless personal communications," *Proc. IEEE*, vol. 82, pp. 1333–1359, Sept. 1994.

[21] T. Rappaport and C. McGillen, "UHF fading in factories," *IEEE J. Select. Areas Commun.*, vol. 7, pp. 40–48, Jan. 1989.

[22] K. Pahlavan and A. Levesque, *Wireless Information Networks*. New York: Wiley, 1995.

[23] H. V. Poor, *An Introduction to Signal Detection and Estimation*. New York: Springer-Verlag, 1988.

[24] R. Horn and C. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1985.

[25] R. Gallager, *Information Theory and Reliable Communication*. New York: Wiley, 1968.

[26] S. Benedetto, E. Biglieri, and V. Castellani, *Digital Transmission Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1987.

[27] I. Gradshteyn and I. Rhyzhik, *Table of Integrals, Series, and Products*. New York: Academic, 1980.

[28] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Trans. Inform. Theory*, vol. IT-28, pp. 55–67, Jan. 1982.

[29] D. Divsalar and M. Simon, "Trellis coded modulation for 4800-9600 bits/s transmission over a fading mobile satellite channel," *IEEE J. Select. Areas Commun.*, vol. SAC-5, pp. 162–175, Feb. 1987.

[30] A. Viterbi and J. Omura, *Principles of Digital Communication and Coding*. New York: McGraw-Hill, 1979.

[31] B. Jeličić and S. Roy, "Cutoff rates for coordinate interleaved QAM over Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 44, pp. 1231–1233, Oct. 1996.

Dennis L. Goeckel received the B.S.C.E.E. degree from Purdue University, West Lafayette, IN, in 1992. From 1992 to 1995, he was a National Science Foundation Graduate Fellow at the University of Michigan, Ann Arbor, where he received the M.S.E.E. degree in 1993 and the Ph.D. degree in 1996, both in electrical engineering with a specialty in communications systems.

In September 1996, he assumed his current position as an Assistant Professor in the Electrical and Computer Engineering Department at the University of Massachusetts, Amherst. His current research interest is in the design of digital communication systems, particularly for wireless communication applications.