Cutoff Rate Optimal Binary Inputs with Imperfect CSI

Saswat Misra, Ananthram Swami, *Senior Member, IEEE,* and Lang Tong, *Senior Member, IEEE*

***Abstract*— We use the cutoff rate to study the optimal binary input distributions for the Rayleigh flat-fading channel with imperfect receiver channel state information (CSI). First, we evaluate the cutoff rate and analyze the optimal binary input as a function of the CSI quality and receiver SNR. Next, we study the limiting distributions – BPSK and On-Off Keying (OOK) – and derive an analytic design rule that allows adaptive switching between these two as the receiver CSI changes. We establish the virtues of a modulation scheme that employs only these limiting distributions, rather than the full spectrum of binary inputs. Finally, we use our results to design an adaptive modulation scheme for Pilot Symbol Assisted Modulation systems. We show that switching between just BPSK and equiprobable-OOK is nearly optimal for moderate to large SNR, and that switching between BPSK and generalized-OOK is nearly optimal for all SNR.**

***Index Terms*— Adaptive modulation, cutoff rate, correlated fading, Doppler, imperfect CSI.**

I. Introduction

BINARY input distributions are often assumed when studying the reliable rates of communications systems, either through channel capacity or other related metrics. The widespread analysis of binary inputs follows from their tractability and optimality, or near optimality, at low SNR under varying amounts of receiver channel state information [2], [34], [19]. We consider reliable rates (i.e., those for which the probability of decoding error can be made arbitrarily small) for communications over a discrete-time Rayleigh flat-fading channel. We assume that the transmitter can select among the class of binary input distributions, and that imperfect (or partial) channel state information (CSI) is available at the receiver.

When perfect receiver CSI is available, it is well known that antipodal signaling (BPSK) maximizes the capacity of this channel (among binary inputs). Conversely, without CSI at the receiver, On-Off keying (OOK) has been shown to be capacity achieving [2]. However, when only imperfect receiver CSI is available, it is not clear as to which strategy, even among these two, is optimal. In this paper, we investigate this intermediate case.

We use the cutoff rate R_o to characterize reliable rates. Cutoff rate analysis has been common since its reintroduction in [26], and studies have been conducted for perfect receiver CSI in [21] (independent fading), [23] and [27] (temporally correlated fading), and for no CSI multiple-input multiple-output (MIMO) systems in [18]. The cutoff rate is a lower bound on capacity that also provides a bound on the random coding exponent (thereby characterizing the entire rate vs. performance curve) via $P_e \leq 2^{-N(R_o-R)}$, where R is the rate, and P_e the probability of decoding error for length N codewords [13],[30, Sections I and II]. Although certain encoding-decoding structures can achieve rates greater than R_o (e.g., turbo coding with iterative decoding), the cutoff rate remains a metric of interest for these systems, as well as others [6]. For example, in sequential decoding, the cutoff rate specifies the largest rate for which decoding complexity remains finite [4]. The cutoff rate often leads to a tractable analysis that often would not be possible through direct evaluation of the random coding exponent or the capacity.

In Section II we introduce a Rayleigh fading channel with imperfect receiver CSI and find the corresponding cutoff rate under binary signaling. Our model has been abstracted away from the particular estimator used. Rather, these mechanics are absorbed into a single parameter, the normalized variance of the channel estimate, termed the *CSI quality*. In Section III, we analyze the optimal binary input as a function of the SNR and CSI quality available at the receiver. We establish the cutoff rate optimality of the limiting distributions – BPSK and On-Off Keying – and find an analytic design rule that allows adaptive switching between these two based on the receiver CSI quality. We then establish the virtues of a modulation scheme that employs only these limiting distributions, rather than the full spectrum of binary inputs. In the remaining sections, we introduce an explicit Pilot Symbol Assisted Modulation (PSAM) front-end, and use it to illustrate how results from Sections II and III can be applied to design an adaptive modulation scheme. In Section IV we include temporal correlation in the channel model, and find the cutoff rate under a PSAM scheme with minimum mean square error (MMSE) estimation. In Section IV-C, we discuss adaptive modulation strategies and show that switching between just BPSK and equiprobable-OOK nearly achieves optimal binary signaling for moderate (≈ 0 dB) to large SNR. Further, we show that switching between just BPSK and generalized-OOK is nearly optimal for all SNR. Finally, we provide a summary

Manuscript received October 7, 2004; revised September 11, 2005; accepted December 15, 2005. The associate editor coordinating the review of this paper and approving it for publication was F. Daneshgaran. Parts of this paper were presented at the *37th Annual Conference on Information Sciences and Systems*, Baltimore, March 2003, and the *IEEE International Symposium on Information Theory*, Chicago, June 2004.

S. Misra is with the Army Research Laboratory, Adelphi, MD 20783 USA and the School of Electrical and Computer Engineering, Cornell University, Ithaca, NY 14853 USA (e-mail: smisra@arl.army.mil).

A. Swami is with the Army Research Laboratory, 2800 Powder Mill Rd., Adelphi, MD 20783-1197 USA (e-mail: aswami@arl.army.mil).

Lang Tong is with the School of Electrical and Computer Engineering, Cornell University, 384 Rhodes Hall, 35 Rosina Drive, Ithaca, NY 14850 USA (e-mail: ltong@ece.cornell.edu).

Digital Object Identifier 10.1109/TWC.2006.04665

and discussion in Section V.

In this paper, we do not consider the design of higher order inputs (which has been done in [2], [19], [10], and [34] for the capacity) or optimal inputs when the channel is peak-constrained (e.g., see [30], [31]). Here, we use the cutoff rate to (i) study the behavior of the optimal binary inputs when only imperfect CSI is available, and (ii) apply this analysis to the design of a tractable adaptive modulation scheme for PSAM based communications systems.

Notation and Definitions

We use the following notation and definitions: (a) $x \sim \mathcal{CN}\left(\mu, \sigma^2\right)$ denotes a complex Gaussian random variable x with mean μ and with independent real and imaginary parts, each having variance $\sigma^2/2$, (b) $|\mathcal{A}|$ is the magnitude of the complex number $\mathcal{A}$, (c) $\mathcal{E}[.]$ is the expectation operator, and (d) superscript "H" denotes complex conjugation.

II. System Model

A. Channel Model

We consider single-user communications over a time-varying Rayleigh flat-fading channel. The received signal y_k is given by

$$y_k = \sqrt{E}\, h_k s_k + n_k, \tag{1}$$

where k denotes discrete time, $h_k \sim \mathcal{CN}(0, \sigma_h^2)$ models independent and identically distributed (i.i.d.) fading, E is the average symbol energy used at the transmitter, and $n_k \sim \mathcal{CN}(0, \sigma_N^2)$ models additive white Gaussian noise (AWGN). The binary channel input $s_k \in \{A, -B\}$ is assumed to be real-valued, without loss of generality, and subject to a unit-energy constraint $pA^2 + (1-p)B^2 = 1$, where $0 \le p \le 1$ is the probability of transmitting A. Without loss of generality, we assume that $1 \le A \le \infty$ and $0 \le B \le 1$. We assume that $\sigma_N^2 \neq 0$ and $\sigma_h^2 \neq 0$.

During each symbol interval, the receiver obtains imperfect CSI in the form of a channel estimate, $\widehat{h}_k$, and so (1) can be rewritten,

$$y_k = \sqrt{E}\, \widehat{h}_k s_k + \sqrt{E}\, \widetilde{h}_k s_k + n_k,$$

where $\widetilde{h}_k = h_k - \widehat{h}_k$ is the residual error in the channel estimate. We assume that both the estimate and the residual error are zero-mean Gaussian and independent, i.e., $\widehat{h}_k \sim \mathcal{CN}(0, \widehat{\sigma}^2)$, $\widetilde{h}_k \sim \mathcal{CN}(0, \widetilde{\sigma}^2)$, and $\widehat{\sigma}^2 + \widetilde{\sigma}^2 = \sigma_h^2$. MMSE estimation schemes exist that satisfy these assumptions, and we describe one such PSAM-based scheme in Section IV. However, this information is not needed for the discussion contained in this section or in Section III.

The receiver employs the soft decision ML decoder which treats s_k as the channel input and the pair $(y_k, \widehat{h}_k)$ as the channel output. That is, letting $\mathbf{s} = (s_1, \ldots, s_N)$ denote a transmitted codeword, and $\mathbf{y} = (y_1, \ldots, y_N)$ and $\widehat{\mathbf{h}} = (\widehat{h}_1, \ldots, \widehat{h}_N)$ denote the observation and channel estimate during the span of a codeword, the decision rule maximizes the posteriori probability of the observation,

$$\max_{\mathbf{s} \in \mathcal{Q}} P(\, \mathbf{y}, \widehat{\mathbf{h}} \mid \mathbf{s}),$$

where $\mathcal{Q}$ is the set of all possible length N input sequences. Lastly, we will find it useful to define the *CSI quality* as the normalized variance of the channel estimate at the receiver,

$$\omega \triangleq \widehat{\sigma}^2 / \sigma_h^2.$$

Note that $\omega = 0$ denotes no CSI, while $\omega = 1$ denotes perfect CSI.

B. Cutoff Rate

The cutoff rate, measured in bits per channel use, is [26] (also see [23] for time-selective fading channels with perfect receiver CSI),

$$R_o = -\min_{Q} \log_2 \int_{y_k} \int_{\widehat{h}_k} \left[\sum_{s_k \in \{A, -B\}} Q(s_k) \times \sqrt{P(y_k, \widehat{h}_k \mid s_k)} \right]^2 d\widehat{h}_k \, dy_k, \tag{2}$$

where $Q(A) = p$, $Q(-B) = 1 - p$, and where $P(y_k, \widehat{h}_k | s_k)$ is the probability distribution function (pdf) of the received signal and channel estimate, conditioned upon the transmitted signal.[1] We evaluate (2) in Appendix A and obtain

$$R_o = -\min_{\mathcal{C}(p,A,B)} \log_2 \left[1 + 2p\,(1-p) \times \left(\frac{\sqrt{1 + \kappa\,(1-\omega)\,A^2}\sqrt{1 + \kappa\,(1-\omega)\,B^2}}{1 + \frac{\kappa}{2}\left(1 - \frac{\omega}{2}\right)(A^2 + B^2) + \frac{\kappa\,\omega}{2} AB} - 1 \right) \right], \tag{3}$$

where $\mathcal{C}(p, A, B) \triangleq \{(p, A, B) : 0 \le p \le 1,\ 1 \le A < \infty,\ 0 \le B \le 1,\ pA^2 + (1-p)B^2 = 1\}$ is the constraint set on the input. We have defined the *received SNR* as

$$\kappa \triangleq E \frac{\sigma_h^2}{\sigma_N^2}.$$

III. Optimal Binary Inputs

The optimal binary input (p^*, A^*, B^*) as a function of the CSI quality ω and SNR κ is found from (3) through the minimization

$$\min_{\mathcal{C}(p,A,B)} p\,(1-p) \times \left(\frac{\sqrt{1 + \kappa\,(1-\omega)\,A^2}\sqrt{1 + \kappa\,(1-\omega)\,B^2}}{1 + \frac{\kappa}{2}\left(1 - \frac{\omega}{2}\right)(A^2 + B^2) + \frac{\kappa\,\omega}{2} AB} - 1 \right). \tag{4}$$

The behavior of the this input is shown in Fig. 1 parameterized by ω.

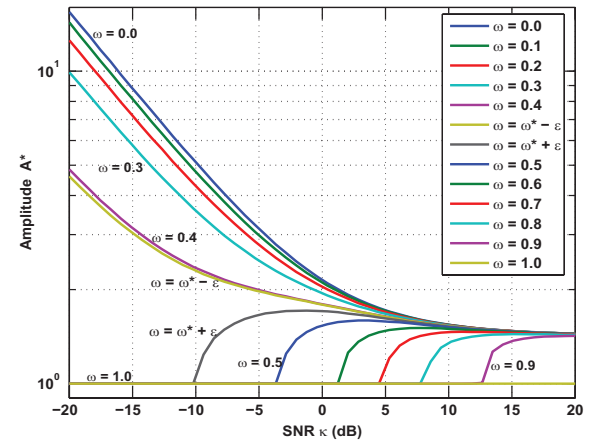
Define the *transitional CSI quality* by

$$\overset{\star}{\omega} \triangleq 1 - \frac{1}{\sqrt{3}}. \tag{5}$$

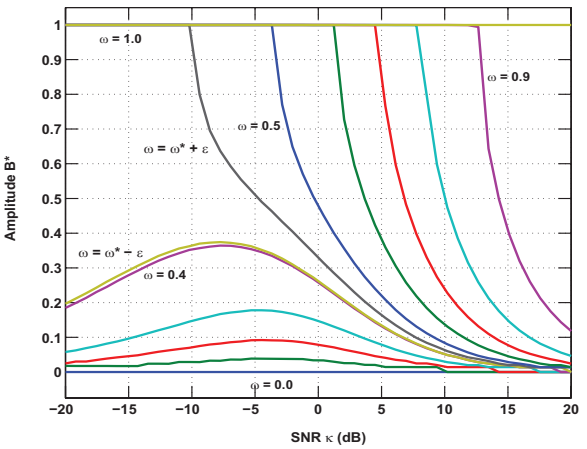
Then the behavior of the optimal binary input is characterized by the following statements:

R1. *For small SNR ($\kappa << 1$)*: If the CSI quality is below the $\overset{\star}{\omega}$ threshold ($\omega < \overset{\star}{\omega}$), then a solution resembling

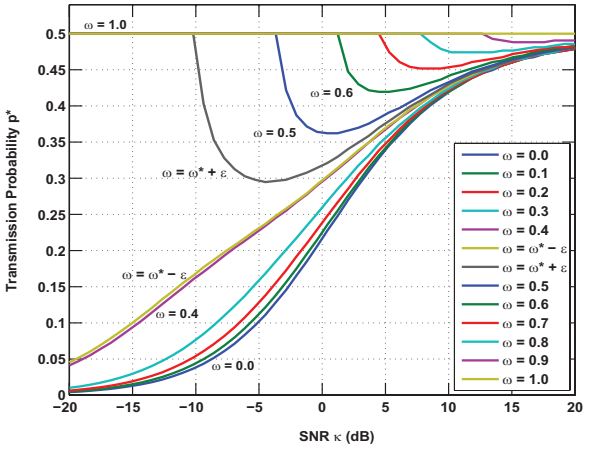
[1]Here, the term cutoff rate implies optimization over the inputs. Later, it will imply certain fixed inputs.



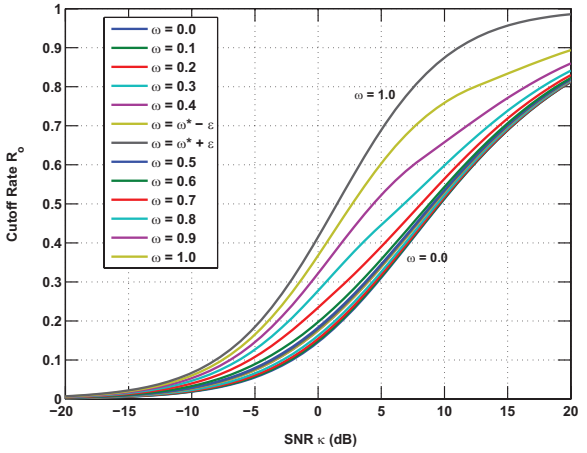
(a) A^* vs. $\mathcal{K}$ for several values of ω



(b) B^* vs. $\mathcal{K}$ for several values of ω



(c) p^* vs. $\mathcal{K}$ for several values of ω



(d) R_o vs. $\mathcal{K}$ for several values of ω

Fig. 1. The optimal binary input (p^*, A^*, B^*) and cutoff rate R_o versus SNR $\mathcal{K}$ for several values of CSI quality ω. The quantity $\epsilon \triangleq 0.01$ is used to illustrate behavior around the $\omega = \overset{\star}{\omega}$ transition point discussed in the text.

OOK with large amplitude is optimal. As $\kappa \to 0$, $\lim_{p\to 0} \text{OOK}(p)$ is optimal.[2] *However*, if the CSI quality exceeds the $\overset{\star}{\omega}$ threshold ($\omega > \overset{\star}{\omega}$), the optimal distribution is BPSK, $A = B = 1, p = 1/2$. Proofs are given in Appendix B.

R2. *For large SNR* ($\kappa >> 1$): If $\omega < \overset{\star}{\omega}$, then from Fig. 1, A^* decreases with SNR to the value $\sqrt{2}$. If $\omega > \overset{\star}{\omega}$, A^* remains fixed at 1 before diverging from the BPSK solution with increasing SNR. After divergence, A^* increases in SNR, reaches a peak, and then also decreases to the value $\sqrt{2}$. As SNR $\to \infty$, OOK$(1/2)$ is shown to be optimal for any CSI quality ($\omega \neq 1$) in Appendix B.

R3. *The optimal transmission probability satisfies* $p^* \leq 1/2$. A sketch of the proof follows: let $\mathcal{I}_1 \triangleq (p,\ A_1,\ B)$ be an arbitrary triple with $p > 1/2$. We show that the alternative solution $\mathcal{I}_2 \triangleq (1-p,\ A_2,\ B)$, where $A_2 = \sqrt{\frac{1-pB^2}{1-p}}$ due to the energy constraint results in a smaller value of (4).

That $\overset{\star}{\omega} = 1 - 1/\sqrt{3}$ is the low SNR ($\kappa \to 0$) transition value is shown in Appendix B. However, the cutoff rate can be seen to be well behaved around $\omega = \overset{\star}{\omega}$ in Fig. 1(d). In Fig. 2, we plot the optimal input parameterized by the SNR κ. Again, the presence of the $\overset{\star}{\omega}$ threshold at low SNR is evident.

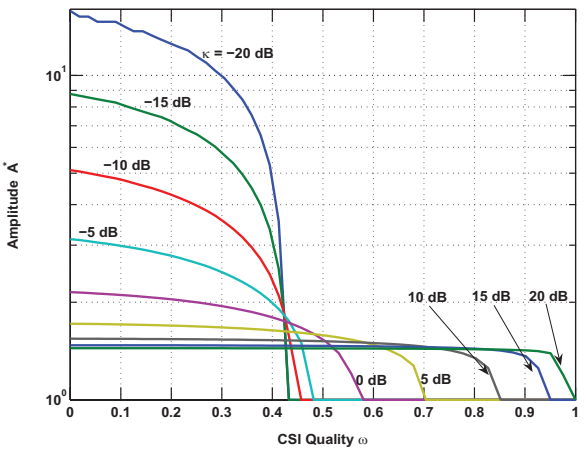
Having examined the general behavior of the optimal binary input, we now focus on the limiting cases of OOK and BPSK.

A. OOK Cutoff Rate

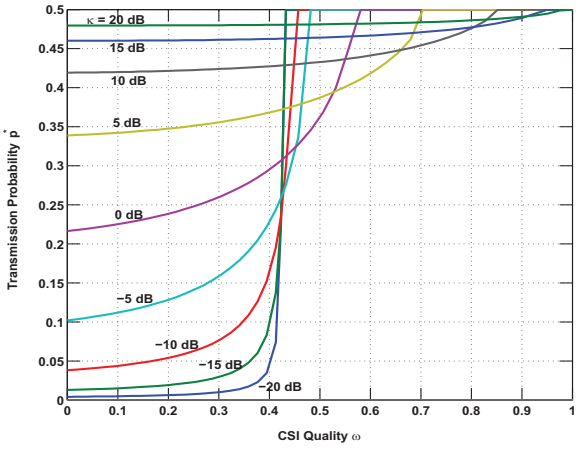
The cutoff rate of OOK was derived for the AWGN channel and hard decision ML decoder in [15]. Here, we derive the OOK cutoff rate for the fading channel with imperfect CSI and soft decision ML decoder. Consider first the no CSI case ($\omega = 0$). We find that OOK(p) modulation maximizes the cutoff rate at all SNR κ (as shown in Appendix C), and so it remains to determine p^*. Setting $A^2 = \frac{1}{p}$, $B = 0$, and $\omega = 0$ in (4), p^* is given by

$$p^* = \min_{0<p<\frac{1}{2}} p(1-p)\left[\frac{\sqrt{1+\kappa\frac{1}{p}}}{1+\kappa\frac{1}{2p}} - 1\right]$$

[2]Henceforth, we will use OOK(p_θ) to denote the binary alphabet with $p = p_\theta$, $A = 1/\sqrt{p_\theta}$, and $B = 0$.

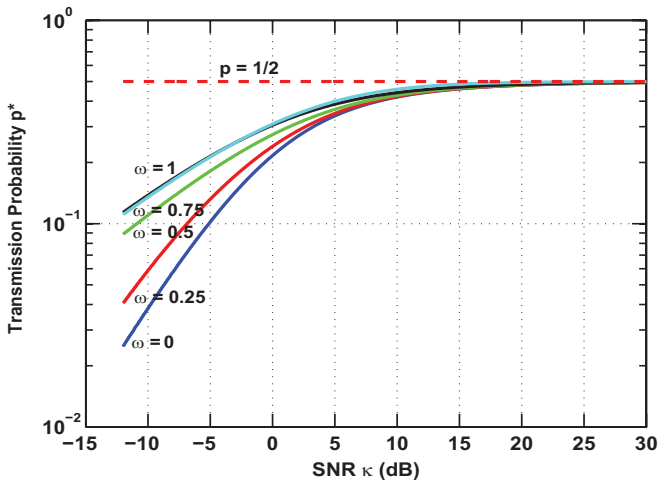


(a) A^* vs. ω for several values of SNR $\mathcal{K}$

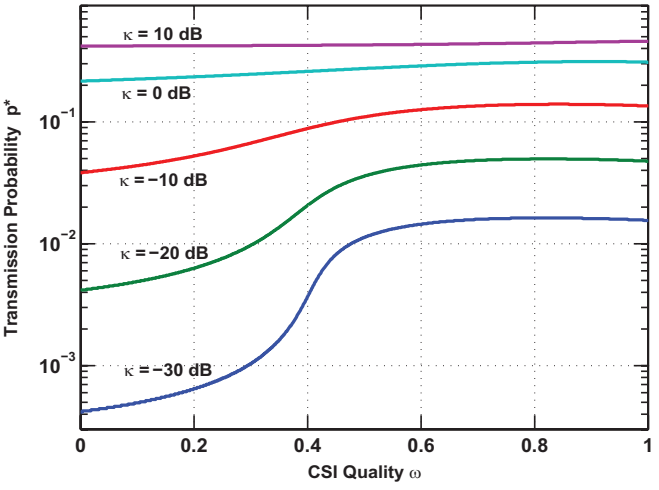


(b) p^* vs. ω for several values of SNR $\mathcal{K}$

Fig. 2. The optimal binary input (A^*, p^*) versus $0 \leq \omega \leq 1$ for several values of SNR $\mathcal{K}$.



(a)



(b)

Fig. 3. The optimal OOK probability p^* versus (a) SNR $\mathcal{K}$ (dB) and (b) CSI quality.

$$= \left\{ p:\ 0 \leq p \leq \frac{1}{2}\ ,\ 2(\kappa+1)p^4 + \left(\frac{11\kappa^2}{4} + \kappa - 1\right) p^3 + \left(\kappa^3 - \frac{3\kappa^2}{2} - \kappa\right) p^2 - \kappa^3 p + \frac{\kappa^3}{4} = 0 \right\}. \quad (6)$$

Solving (6) yields p^* explicitly (as the valid root of the fourth-order polynomial), and provides an easy characterization of the optimal transmission probability (equivalently, the optimal signal energy) as a function of the SNR.

At low SNR ($\kappa \ll 1$), the transmission probability is linear in SNR, i.e., $p^* = \beta\,\kappa$, where

$$\beta \triangleq \frac{(19+3\sqrt{33})^{\frac{2}{3}} - 2(19+3\sqrt{33})^{\frac{1}{3}} + 4}{6(19+3\sqrt{33})^{\frac{1}{3}}} = 0.419\ldots. \quad (7)$$

Returning now to the case of partial CSI, (3) yields the OOK cutoff rate for arbitrary ω as

$$R_{o,K} = -\min_{0<p<1} \log_2 \left[1 + 2p(1-p) \times \left(\frac{\sqrt{1+\kappa(1-\omega)\frac{1}{p}}}{1+\kappa\,(2-w)\,\frac{1}{4p}} - 1 \right) \right]. \quad (8)$$

Analytic maximization of (8) over p leads to a high-order polynomial that does not have an explicit solution as a function of κ and ω. We plot p^* as a function of κ in Fig. 3(a) and as a function of ω in Figure 3(b). It can be verified that as $\kappa \to \infty$, $p^* \to 1/2$, and that as $\kappa \to 0, p^* \to 0$. The transmission probability p^* is seen to be non-monotonic in ω at low SNR. A second-order Taylor series expansion of the expression within the curly brackets in (8) at $\kappa = 0$, yields ($\omega \neq 0$)

$$p^* = \arg\min_{0<p<1} -\frac{\kappa(1-p)\left[4\,p\,\omega + \kappa(3\omega^2 - 6\omega + 2)\right]}{16p},$$

from which we get

$$p^* = \frac{\sqrt{\kappa}}{2}\sqrt{\frac{-2+6\omega-3\omega^2}{w}} \quad \text{for } \omega > \overset{\star}{\omega}, \quad (9)$$

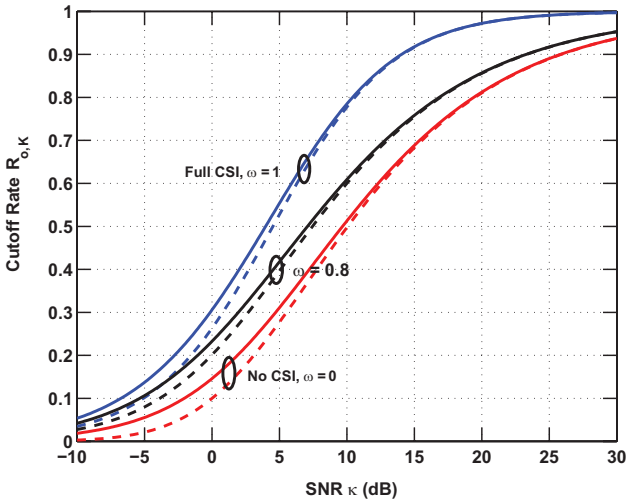


Fig. 4. The OOK cutoff rate $R_{o,K}$ vs. SNR $\mathcal{K}$ (dB) for no CSI (ω = 0), perfect CSI ($\omega = 1$) and imperfect CSI ($\omega = 0.8$), when $p = p^*$ (solid lines, from (8)) and $p = 1/2$ (dashed lines, from (10)).

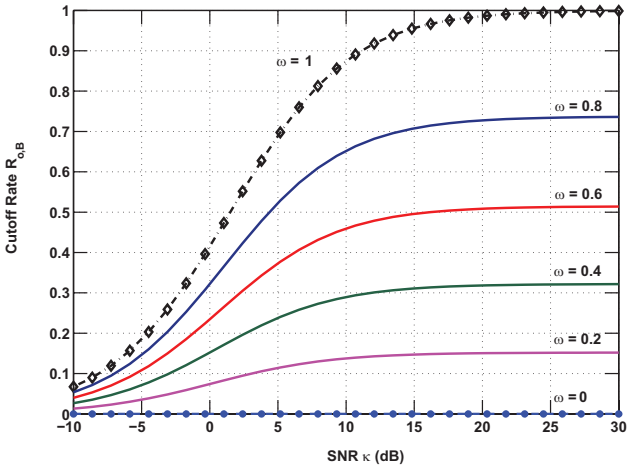


Fig. 5. The BPSK cutoff rate $R_{o,B}$ vs. SNR $\mathcal{K}$ (dB) for different values of the CSI quality ω.

which is decreasing for $\omega \in (\sqrt{2/3}, 1)$.[3] As κ increases, the amplitude A^* decreases (since $A^2 = \frac{1}{p}$). From Fig. 3, we see that for fixed κ, the amplitude A^* is a decreasing function of ω. For moderate to large values of κ, letting $p = 1/2$ is a reasonable approximation to p^*. Using $p = 1/2$, the equiprobable-OOK cutoff rate is found to be

$$R_{o,K} = -\log_2 \left\{ \frac{1}{2} + \frac{1}{2} \left[\frac{\sqrt{1 + 2\kappa(1-\omega)}}{1 + \kappa\left(1 - \frac{\omega}{2}\right)} \right] \right\}. \qquad (10)$$

In Fig. 4, we plot $R_{o,K}$ for $p = p^*$ (solid lines) and $p = 1/2$ (dashed lines) for no CSI ($\omega = 0$), perfect CSI ($\omega = 1$), and imperfect CSI ($\omega = 0.8$). Note that the cutoff rate of both OOK($\frac{1}{2}$) and OOK(p^*) approaches 1 at high SNR for any ω.

B. BPSK Cutoff Rate

Consider the case of perfect CSI. We let $\omega = 1$ in (4), which leads readily to the BPSK solution, $A = B$, and $p = 1/2$. For arbitrary ω, the cutoff rate of BPSK is

$$R_{o,B} = -\log_2 \left\{ \frac{1}{2} + \frac{1}{2} \left[\frac{1 + \kappa(1-\omega)}{1 + \kappa} \right] \right\}. \qquad (11)$$

Fig. 5 shows the cutoff rate as a function of the received SNR κ and the CSI quality ω. Note that the CSI quality imposes an asymptotic ceiling on $R_{o,B}$, and that at high SNR, the cutoff rate saturates to

$$R_{o,B} = -\log_2 \left\{ 1 - \frac{\omega}{2} \right\}.$$

As expected, the cutoff rate is zero when there is no CSI.

To study the relative impact of imperfect CSI on BPSK and OOK, it is instructive to consider the statistics of y_k under the two hypotheses:

$$y_k | \widehat{h}_k, s_k \sim \mathcal{CN}\left(\sqrt{E}\, \widehat{h}_k s_k,\ \sigma_N^2 (1 + s_k^2\, \kappa(1-\omega)) \right),$$

[3]A third-order Taylor expansion yields an expression for p^* that is valid for all ω; it is omitted for brevity.

where $s_k \in \{-1, 1\}$ for BPSK and $s_k \in \{0, \sqrt{2}\}$ for OOK(1/2). When the SNR is large enough, i.e., $\kappa \gg \frac{1}{1-\omega}$, the channel estimation error dominates, and it is easy to verify that the BPSK performance saturates.

Thus, we expect that OOK is optimal at large κ, and that BPSK is optimal for small κ. Next, we quantify the SNR at which one should switch from BPSK to OOK as a function of estimator quality ω.

C. Transitional SNR

We confine our interest to BPSK (optimal for perfect CSI) and OOK (optimal for no CSI) and provide the transitional SNR $\overset{\star}{\kappa}$, above which OOK is optimal, and below which BPSK is optimal. This result provides an initial characterization of the intermediate region where imperfect CSI is available, and provides an *analytic basis* for an adaptive modulation scheme in which the transmitter can select between OOK and BPSK based on the SNR κ and CSI quality ω available at the receiver as described in subsequent sections. For OOK(1/2), the *transitional SNR* $\overset{\star}{\kappa}$ is found by equating (11) and (10) and solving for κ. Doing so yields the solution

$$\overset{\star}{\kappa}(\omega) = \left\{ \kappa :\ \kappa \geq 0,\ \frac{(2-\omega)^2(1-\omega)^2}{4}\,\kappa^3 - \frac{(1-\omega)}{2} \times \left[(10 - 3\omega)\omega - 4 \right]\kappa^2 + \left(\frac{13\omega^2}{4} - 5\omega + 1 \right)\kappa - \omega = 0 \right\},$$

for which the explicit solution is

$$\overset{\star}{\kappa}(\omega) = \frac{(a+b)^{1/3} + (a-b)^{1/3} - 2\left(4 - 10\omega + 3\omega^2\right)}{3(2-\omega)^2(1-\omega)}, \qquad (12)$$

with the definitions

$$a \triangleq 81\omega^6 - 468\omega^5 + 828\omega^4 - 640\omega^3 + 624\omega^2 - 192\omega + 64,$$
$$b \triangleq 6\sqrt{3}(\omega - 2)^2 w^2 \sqrt{61\omega^4 - 208\omega^3 + 168\omega^2 - 64\omega + 16}.$$

The transitional SNR $\overset{\star}{\kappa}$ depends on the CSI quality, and is shown in Fig. 6 (dashed line). At the end points, $\omega = \{0, 1\}$,

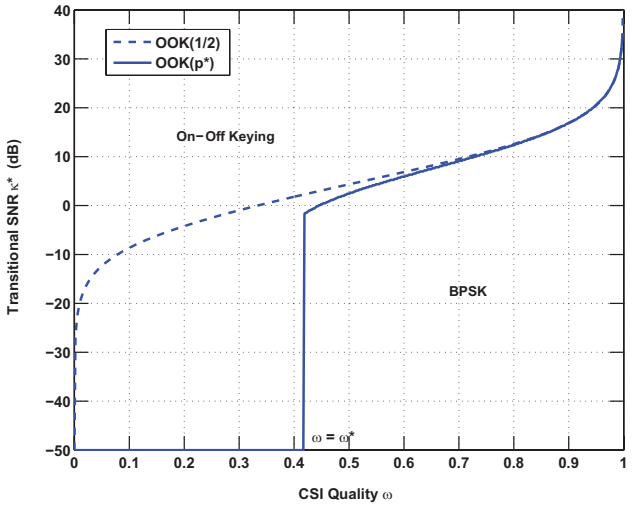


Fig. 6. The transitional SNR $\overset{\star}{\kappa}$ above which OOK is optimal and below which BPSK is optimal, for: equiprobable-OOK and BPSK (dashed line) and OOK(p^*) and BPSK (solid line).

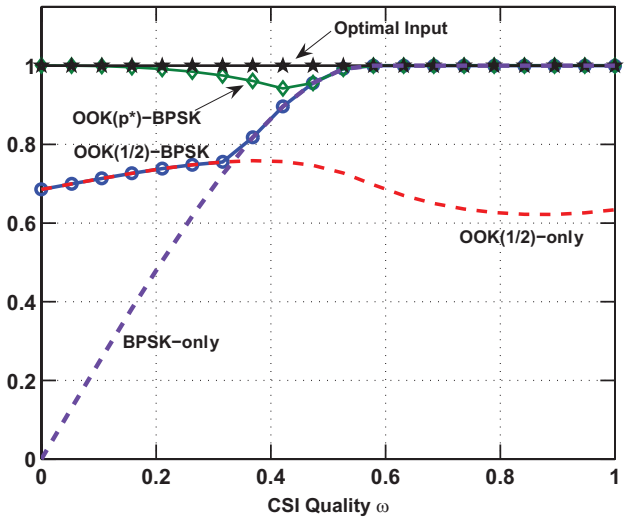


Fig. 7. The normalized cutoff rate for the Rayleigh flat-fading channel with imperfect CSI as a function of the CSI quality ω, and for five different binary transmission strategies with $\mathcal{K} = 0$ dB.

we find the anticipated results: $\overset{\star}{\kappa}(0) = 0$, implying that equiprobable-OOK is preferred to BPSK at any SNR when no CSI is available, and $\lim_{\omega\to 1} \overset{\star}{\kappa}(\omega) = \infty$, implying that BPSK is preferred to equiprobable-OOK when perfect CSI is available.

In Fig. 6, we repeat the threshold curve for OOK(p^*) (solid line). To find this region we equate (11) and (8) and solve for $\overset{\star}{\kappa}$ numerically. As expected, optimizing over p results in OOK(p) being preferred to BPSK over a wider range of SNR for fixed ω. Interestingly, we find that there is a threshold value of CSI below which BPSK is not useful. A low SNR analysis once again reveals this value to be $\overset{\star}{\omega} = 1 - 1/\sqrt{3}$.[4]

D. Adaptive Modulation Based on CSI Quality

The performance of the adaptive modulation schemes described in Section III-C is shown in Fig. 7. As upper and lower bounds, we plot the cutoff rate of optimal binary signaling (determined from (4)), and for the BPSK–only and OOK($\frac{1}{2}$)–only schemes when $\kappa = 0$ dB. Each curve has been normalized by the cutoff rate of optimal binary signaling. The OOK($\frac{1}{2}$)–BPSK scheme simply traces out the best of the BPSK and OOK cutoff rates. The BPSK–only scheme performs arbitrarily poorly for small ω (as expected due to its saturating behavior at large SNR, see Fig. 5), while the OOK($\frac{1}{2}$)–only scheme is seen to be suboptimal by up to ~ 40 percent for large values of ω. In contrast, *the OOK(p^*)–BPSK scheme performs nearly as well as optimal binary signaling over the entire range of* ω. To understand this behavior, we partition the (κ, ω) plane into three regions in Fig. 8: (a) the region where BPSK is within one-percent of optimal, (b) the region where OOK(p^*) is within one-percent of optimal, and (c) the remaining region.[5] Over most of the (κ, ω) plane, we

[4]This reoccurrence of $\overset{\star}{\omega}$ is to be expected. Earlier in this section we found that, among all binary inputs, only OOK(p^*) or BPSK is optimal at low SNR.

[5]While BPSK is optimal over a (κ, ω) region, OOK is optimal only on a line (from Fig. 1). This necessitates the use of optimality regions in the figure.

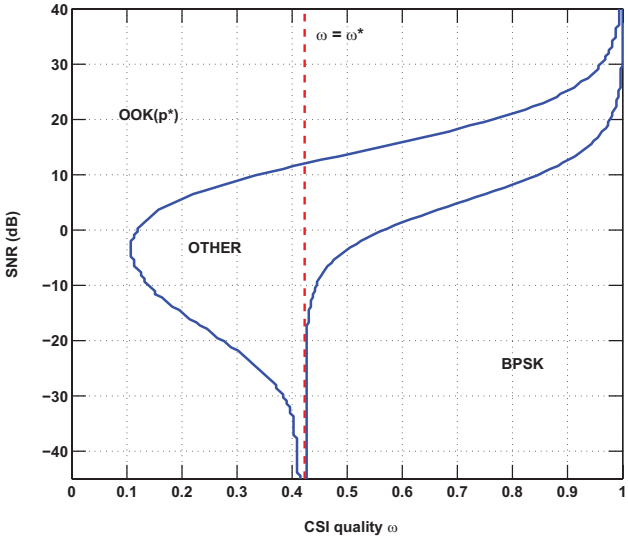


Fig. 8. Partitioning of the $(\mathrm{SNR}(\mathcal{K}), \mathrm{CSI}(\omega))$ plane into three regions: (a) where BPSK is within one-percent of optimal, (b) where OOK(p^*) is within one-percent of optimal, and (c) the remaining region.

see that either BPSK or OOK(p^*) is indeed nearly optimal. Comparing Fig. 8 to Fig. 6, we find that BPSK retains nearly its entire optimality region even when arbitrary binary inputs are admitted. In contrast, OOK(p^*) loses a portion of its optimality region in this scenario. In particular, the region loosely described by $\{(\kappa, \omega) : -20 \text{ dB} \le \kappa \le 10 \text{ dB}, \ 0.2 \le \omega \le \overset{\star}{\omega}\}$ is now allocated to region (c).

E. Sensitivity Analysis

We study the sensitivity of the cutoff rate to the binary input used for four limiting cases: (a) large κ ($\kappa = 30$ dB), large ω ($\omega = 0.95$) (b) small κ ($\kappa = -10$ dB), large ω, (c) large κ, small ω ($\omega = 0.1$), and (d) small κ, small ω.

Large ω. For large κ, the cutoff rate is sensitive to the choice of p, but not to the choice of B. For example, with $B = 1$,

the cutoff rate increases more than 300 percent as p increases from 0.1 to 0.5. On the other hand, with $p = 0.4$, the cutoff rate increases by only 4 percent as B varies from 0 to 1. For small κ, the cutoff rate is sensitive to choice of both p and B. For example, with $B = 1$, the cutoff rate increases by 280 percent as p increases from 0.1 to 0.5. With $p = 0.5$, the cutoff rate increases by a factor of approximately 200 percent as B varies from 0 to 1.

Small ω. At both small and large κ, the cutoff rate is sensitive to the choice of B when p is chosen optimally. When p is chosen suboptimally, sensitivity decreases. For example, at small κ the cutoff rate increases by 400 percent as B increases from 0 to 1. At large κ, the increase is 800 percent. In each case, sensitivity to B diminishes if p is chosen suboptimally. Overall, it is clear that optimization can provide large gains in the cutoff rate.

IV. Adaptive Modulation Scheme for PSAM

In this section we consider a temporally correlated flat-fading channel and apply the results derived in Section III, in particular those of Section III-C, to the design of an adaptive modulation scheme for PSAM-based communications. We then analyze the performance of the proposed scheme relative to some clear alternatives.

In PSAM [8], [32], known pilot symbols are multiplexed with data symbols for transmission through the communications channel. At the receiver, knowledge of these pilots is used to form channel estimates, which aid the detection of the data both directly (by modifying the detection rule based on the channel estimate) and indirectly (e.g., by allowing for estimate-directed modulation, power control, and media access). In general, there is no guarantee that PSAM-based approaches are optimal, and PSAM has been shown to be suboptimal when the channel coherence time is small and/or the SNR small from various perspectives [17], [25], [2], [11]. Nevertheless, the technique is of great practical significance. In addition to providing implementable receiver structures, PSAM facilitates accurate timing and synchronization. PSAM has been incorporated into many commercial and Military standards, and optimized approaches to PSAM have been studied from the perspectives of frequency and timing offset estimation [22], [14], bit-error rate (BER) [8], [11], [7], [12], and the channel capacity or its bounds [17], [3], [24], [29].

A. System Model

We generalize the Rayleigh fading channel of (1) to include temporal correlation. The observation equation is

$$y_k = \sqrt{E} h_k s_k + n_k,$$

where $h_k \sim \mathcal{CN}\left(0, \sigma_h^2\right)$ now exhibits temporal correlation described by the normalized correlation function $R_h(\tau) \triangleq \frac{1}{\sigma_h^2}\mathcal{E}\left[h_k h_{k+\tau}^H\right]$. We assume that training is sent with period T at times $k = mT, m \in \mathcal{Z}$ and that $s_{mT} = +1$. In each data slot $mT + \ell$ $(1 \le \ell \le T-1)$, an MMSE estimate of the channel $\widehat{h}_{mT+\ell}$ is made at the receiver using some subset $\mathcal{N}$ of past and future training symbol observations, so that

$$\widehat{h}_{mT+\ell} = \mathcal{E}[h_{mT+\ell} | \{y_{nT}\},\ n \in \mathcal{N} \subseteq \mathcal{Z}], \tag{13}$$

$1 \le \ell \le T-1,\ m \in \mathcal{Z}$. The system equation in the m^{th} frame, i.e., $mT \le k \le (m+1)T - 1$, is then

$$y_k = \begin{cases} \sqrt{E}\, h_{mT} + n_{mT}, & \text{(pilot)}, \\ \sqrt{E}\, \left(\widehat{h}_{mT+\ell} + \widetilde{h}_{mT+\ell}\right) s_{mT+\ell} + n_{mT+\ell} & \text{(data)}. \end{cases} \tag{14}$$

The use of an MMSE estimator implies that the estimate $\widehat{h}_{mT+\ell}$ and the estimation error $\widetilde{h}_{mT+\ell}$ are zero-mean, jointly Gaussian, and independent with variances $\widehat{\sigma}_\ell^2$ and $\sigma_h^2 - \widehat{\sigma}_\ell^2$, respectively; $\widehat{h}_{mT+\ell} \sim \mathcal{CN}\left(0, \widehat{\sigma}_\ell^2\right)$ and $\widetilde{h}_{mT+\ell} \sim \mathcal{CN}\left(0, \sigma_h^2 - \widehat{\sigma}_\ell^2\right)$. To characterize the partial CSI provided by the estimator ℓ slots from the last-pilot, we define the CSI quality *in the ℓ^{th} slot*

$$\omega_\ell \triangleq \widehat{\sigma}_\ell^2 / \sigma_h^2, \tag{15}$$

for $\ell = 1, \ldots, T-1$. The CSI quality $\omega_\ell,\ 0 \le \omega_\ell \le 1$, captures the impact of the channel correlation $R_h(\tau)$, estimator $\mathcal{N}$, and SNR κ on the statistical quality of channel estimates at the receiver. The variance of any estimator can be found by noting that, $\widehat{h}_{mT+\ell}$ is the expected value of one Gaussian vector conditioned upon another. The CSI quality is then readily obtained via (15).

Given the periodic nature of the training, it is natural to let the binary signaling scheme vary from data slot to data slot, with period T. Therefore, defining $[k] \triangleq k \mod T$, we allow the input s_k to be selected from a real-valued binary signal set $\mathcal{S}_{[k]} = \{A_{[k]}, -B_{[k]}\}$ subject to a unit average-energy constraint: $p_{[k]} A_{[k]}^2 + (1 - p_{[k]}) B_{[k]}^2 = 1$, where $p_{[k]}$ is the probability of transmitting $A_{[k]}$ (note that $\mathcal{S}_0 = \{+1\}$). We assume that $A_{[k]}$ and $B_{[k]}$ are real-valued, and that $1 \le A_{[k]} < \infty$ and $0 \le B_{[k]} \le 1$. Finally, we assume that codewords can occur in integers multiples of a frame length, i.e., $N = n(T-1), n = 1, 2, \ldots$, and are decoded using the ML decoder which treats $s_1, \ldots, s_{T-1}$ as the channel input and the pair $(\widehat{h}_1, \ldots, \widehat{h}_{T-1}\ ;\ y_1, \ldots, y_{T-1})$ as the channel output.

We consider a system in which perfect interleaving [5], [23], [27] is performed at the transmitter and channel estimation is performed before deinterleaving at the receiver. The system equation under interleaving is still given by (14), but now $h_k \sim \mathcal{CN}\left(0, \sigma_h^2\right)$ and $n_k \sim \mathcal{CN}\left(0, \sigma_N^2\right)$ are i.i.d. sequences representing the interleaved channel and noise. Interleaving implies that $\widehat{h}_k$ and $\widetilde{h}_k$ are independent sequences in k and with respect to each other. However, the marginal statistics of the channel estimate and estimation error are preserved, i.e., $\widehat{h}_{mT+\ell} \sim \mathcal{CN}\left(0, \widehat{\sigma}_\ell^2\right)$ and $\widetilde{h}_{mT+\ell} \sim \mathcal{CN}\left(0, \sigma_h^2 - \widehat{\sigma}_\ell^2\right)$.

B. Cutoff Rate

In [28], we have previously studied the cutoff rate of a PSAM communications system. Under the assumption of perfect interleaving, and by modification of the proof in [28] to the case of generalized binary inputs, the cutoff rate of the system of Section IV-A can be seen to be given by

$$R_o = -\frac{1}{T} \sum_{\ell=1}^{T-1} \min_{c_\ell(p,A,B)} \log_2 \left[1 + 2p_\ell \left(1 - p_\ell\right) \right.$$

$$\times \left(\frac{\sqrt{1+\kappa\,(1-\omega_\ell)\,A_\ell^2}\sqrt{1+\kappa\,(1-\omega_\ell)\,B_\ell^2}}{1+\frac{\kappa}{2}\,\left(1-\frac{\omega_\ell}{2}\right)(A_\ell^2+B_\ell^2)+\frac{\kappa\,\omega_\ell}{2}A_\ell B_\ell} - 1 \right) \Bigg], \tag{16}$$

where $\mathcal{C}_\ell(p,A,B) \triangleq \{(p_\ell, A_\ell, B_\ell) : 0 \le p_\ell \le 1, 1 \le A_\ell < \infty, 0 \le B_\ell \le 1,\ p_\ell A_\ell^2 + (1-p_\ell)B_\ell^2 = 1\}$ is the constraint set on the ℓth input.

Comparing the cutoff rate of the i.i.d. channel (3) to that of a PSAM system operating over the temporally correlated fading channel under interleaving (16), it is clear that the latter can be interpreted as consisting of $T-1$ parallel data-channels, where the ℓ^{th} channel consists of all data slots occurring ℓ positions after the most recent pilot. The ℓ^{th} $(1 \le \ell \le T-1)$ term in the sum of (16) represents the cutoff rate in one of $T-1$ data channels, with CSI quality ω_ℓ and SNR κ. Therefore, letting $\omega = \omega_\ell$, we can apply the analysis given in Section III on a per-channel basis. This motivates design of a PSAM system in which the optimal binary distribution $(A_\ell^*, B_\ell^*, p_\ell^*)$ is used in each data channel. Next, we combine the cutoff rate (16) with the optimal input analysis of Section III to design adaptive modulation schemes in which the transmitter selects the modulation in each data slot based on the partial CSI ω_ℓ and SNR κ at the receiver.

C. Adaptive Modulation Scheme

Adaptive transmission techniques for fading channels have been well studied in the literature (e.g., see [7], [16], [9], and the references therein). Typically, a subset of the key transmission parameters – power, rate, modulation shape and size, and bandwidth – is adapted based on some instantaneous measure of the channel quality, which may determined by the fading, noise, or interference level at the receiver. This knowledge is typically provided to the transmitter via a feedback link, which introduces its own noise and/or delay to the process. However, when PSAM is employed over continuously time-varying fading channels, the transmitter need not adapt to instantaneous channel quality measurements, since it can adapt instead to the statistical quality of the channel estimates – which varies with the estimate's position relative to pilot symbols. Further, if the transmitter has knowledge of the channel Doppler spectra, $R_h(\tau)$, it can compute this statistical quality without requiring explicit feedback.

In [1], this methodology is proposed and used to characterize PSAM-based achievable rates for the Gauss-Markov correlation model. An adaptive binary modulation scheme was proposed in which the transmitter selects the achievable rate optimal binary input in each data slot. It was shown that adaptive modulation provides significant gains in the achievable rates relative to static modulation when the channel fading is slower and/or the estimator more accurate. Here, we consider an adaptive binary modulation scheme based on the cutoff rate. Our aim is to compare the performance of the simple two-distribution modulation techniques derived in Section III-C to optimal binary signaling (from Section III). Specifically, we propose:

C1. The OOK($\frac{1}{2}$)–BPSK adaptive system in which equiprobable-OOK is used in each sub-channel where it is preferred to BPSK, and where BPSK is used otherwise. Denote the cutoff rate of this system by R_{HYB1}. This scheme is implemented through the analytic switching rule derived in (12).

C2. The OOK(p^*)–BPSK adaptive system in which generalized-OOK is used in each sub-channel where it is preferred to BPSK, and where BPSK is used otherwise. Denote this cutoff rate by R_{HYB2}. This scheme is implemented using the solid curve shown in Fig. 6. In practice, this curve can be implemented in hardware at low cost.

As benchmarks, we compare the performance of our schemes to:

C3. *Optimal binary signaling*, in which each data slot is assigned the cutoff rate optimal binary input as determined from (4). This will be computed numerically. Denote the cutoff rate of this system by R_{BIN}. This scheme provides an upper bound on the performance of C1 and C2.

C4. The BPSK–only system. Denote the cutoff rate of this system by R_{BPSK}. This scheme provides a lower bound on the performance of C1 and C2.

C5. The OOK($\frac{1}{2}$)–only system which uses OOK($\frac{1}{2}$) in each sub-channel. Denote the cutoff rate of this system by R_{OOK}. This scheme also provides a lower bound on the performance of C1 and C2.

We show that our adaptive modulation scheme, based on switching between just two inputs[6], captures the optimality of scheme C3 over a wide range of SNR, while requiring a fraction of the complexity.

Simulation. We consider two estimators (see [28],[1]): The causal $(\mathbf{1},\mathbf{0})$ estimator $\mathcal{N} = \{m\}$, for which

$$\omega_\ell^{(1,0)} = \left|R_h^2(\ell)\right| \frac{\kappa}{1+\kappa},$$

and the non-causal $(\mathbf{1},\mathbf{1})$ estimator $\mathcal{N} = \{m, m+1\}$, for which

$$\omega_\ell^{(1,1)} = \frac{(\kappa^2+\kappa)\Big(\,|R_h(\ell)|^2 + |R_h(T-\ell)|^2\,\Big)}{(\kappa+1)^2 - \kappa^2|R_h(T)|^2} - \frac{2\kappa^2\ \text{Re}\Big\{R_h(\ell)\ R_h(T-\ell)\ R_h^*(T)\Big\}}{(\kappa+1)^2 - \kappa^2|R_h(T)|^2},$$

where Re $\{.\}$ denotes the real part. We assume that the channel correlation is described by the well-known Jakes model [20], for which $R_h(\tau) = J_o(2\pi f_D T_D \tau)$, where $J_o(.)$ is the zeroth-order Bessel function of the first kind, and where $f_D T_D$ is the normalized Doppler spread. We let $f_D T_D = 1/50$, and $T = 7$. Fig. 9(a) plots the cutoff rate versus SNR (dB) for each of the schemes C1-C5 for the $(\mathbf{1},\mathbf{0})$ estimator (each curve has been normalized by the cutoff rate of optimal binary signaling, scheme C3). Note that:

R4. For small SNR, BPSK outperforms OOK($\frac{1}{2}$). At high SNR, the reverse is true.

R5. *The performance of the OOK(p^*)–BPSK adaptive strategy is nearly identical to that of optimal binary signaling.*

[6]Note that this switching may be oscillatory, e.g., producing an BPSK–OOK–BPSK or OOK–BPSK–OOK behavior, if ω_ℓ is non-monotonic in ℓ. This may be the case if the channel correlation $R_h(\tau)$ is non-monotonic and/or a noncausal estimator is used.

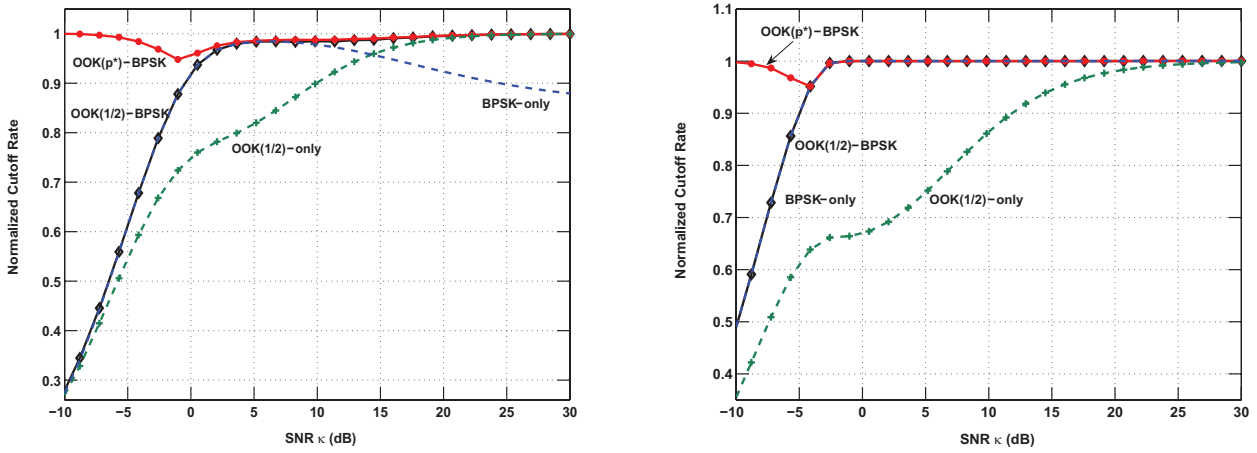


Fig. 9. The (normalized) cutoff rate of PSAM when using the (a) $(\mathbf{1},\mathbf{0})$ or (b) $(\mathbf{1},\mathbf{1})$ estimator for five different binary transmission strategies. Parameters: $R_h(\tau) = J_o(2\pi f_D T_D \tau)$, $f_D T_D = 1/50$, and $T = 7$. All curves are normalized to the range $[0,1]$ by the optimal binary input, which is not plotted explicitly.

Therefore, using only two types of inputs, BPSK and the OOK family, is nearly optimal.

R6. Limited to the OOK($\frac{1}{2}$)–BPSK scheme, performance is nearly identical to optimal binary signaling for moderate to high SNR. This implies that nearly optimal transmission can be achieved even under transmitter peak-to-average power ratio (PAPR) constraints, simply by switching between two constellations when the SNR is moderate to large (in this example, $k \geq 2$ dB).

In Fig. 9(b) we repeat the analysis for the $(\mathbf{1},\mathbf{1})$ estimator. This estimator provides *at least* the CSI quality of the $(\mathbf{1},\mathbf{0})$ estimator, implying that BPSK will be preferred to OOK in at least as many data slots.[7] Here, we find that BPSK is preferred to OOK in every data slot. Note that for SNR greater than -2 dB, the optimality of scheme C3 is captured by the simple OOK($\frac{1}{2}$)–BPSK adaptive scheme.

V. Discussion

We have studied cutoff rate optimal binary inputs for the Rayleigh flat-fading channel with imperfect receiver CSI. First, we evaluated the cutoff rate for i.i.d. fading (3), and analyzed the optimal binary input as a function of the CSI quality and SNR at the receiver (summarized in remarks R1-R3). It was seen that there exists a CSI quality threshold, $\overset{\star}{\omega}$, which characterizes the phase transition in the optimal input versus the CSI quality at low SNR. Next, we considered the limiting distributions – BPSK and OOK. Under OOK, Equations (6), (7), (9) show that the cutoff rate provides a simple characterization of the probability versus location of the non-zero mass point as a function of the CSI quality and SNR. We derived a transitional SNR $\overset{\star}{\kappa}$ (see (12) and Fig. 6) that enables adaptive switching between these distributions based on the CSI quality, and have shown that this scheme is nearly optimal. Next, we applied our results to adaptive modulation design in PSAM communications over a temporally correlated channel. We have shown that switching between just BPSK and equiprobable-OOK nearly achieves optimal binary signaling at moderate ($\approx$ 0 dB) to large SNR, and that switching between BPSK and generalized-OOK is nearly optimal at all SNR.

While explicit expressions for the optimal transmission probability, such as (6), (7), and (9), are not readily available in a capacity analysis, similar trends have been observed [2]. Explicit expressions for the transitional SNR or CSI quality as defined here do not readily follow for the mutual information, and may not exist (given the sometimes drastically different behavior of these two metrics, e.g., see [27]). As has been previously noted in various contexts for the capacity, we have found an extreme form of "peaky" OOK to be cutoff rate optimal at low SNR when no CSI is available (and as long as the CSI remains below the $\overset{\star}{\omega}$ threshold).

We have restricted our attention to binary inputs. However, even at low SNR, binary inputs are not second-order optimal [33], and a study of optimal higher order inputs for imperfect CSI may be of interest.

Appendix

A. Derivation of the Cutoff Rate, Equation (3)

Define $\mathcal{S} \triangleq \{A, -B\}$ and omit the subscript k for brevity. Starting from (2), we get

$$\int_y \int_{\widehat{h}} \left[\sum_{s\in\mathcal{S}} Q(s) \sqrt{P(y,\widehat{h} \mid s)} \right]^2 d\widehat{h}\, dy = \sum_{v,w\in\,\mathcal{S}} \sum Q(v)Q(w)\mathcal{E}_{\widehat{h}} \left[\int_y \sqrt{P(y|v,\widehat{h})P(y|w,\widehat{h})} dy \right].$$

Note that $y|v,\widehat{h} \sim \mathcal{CN}\left(\sqrt{E}\widehat{h}v, E\widetilde{\sigma}^2 v^2 + \sigma_N^2\right)$ and similarly for $y|w,\widehat{h}$. Thus, we get

$$\int_y \sqrt{P(y|v,\widehat{h})P(y|w,\widehat{h})}\, dy = \exp\left[-\frac{1}{4} \frac{E|\widehat{h}|^2 (v-w)^2}{E\widetilde{\sigma}^2 \left(\frac{v^2+w^2}{2}\right) + \sigma_N^2} \right]$$

[7] A similar conclusion holds for any estimator that provides uniformly as good or better CSI quality than a simpler one.

$$\times \frac{\sqrt{E\widetilde{\sigma}^2 v^2+\sigma_N^2}\sqrt{E\widetilde{\sigma}^2 w^2+\sigma_N^2}}{E\widetilde{\sigma}^2\left(\frac{v^2+w^2}{2}\right)+\sigma_N^2}.$$

Following (2), we take the expectation of the above with respect to $\widehat{h} \sim \mathcal{CN}(0,\widehat{\sigma}^2)$, yielding

$$\frac{\sqrt{\sigma_N^2+\widetilde{\sigma}^2 E v^2}\sqrt{\sigma_N^2+\widetilde{\sigma}^2 E w^2}}{\sigma_N^2+\widetilde{\sigma}^2 E\left(\frac{v^2+w^2}{2}\right)+\frac{1}{4}\widehat{\sigma}^2 E\left(v-w\right)^2}.$$

Dividing the numerator and denominator by σ_N^2 and substituting the result into (2) yields

$$R_o = -\min_{\mathcal{C}(p,A,B)} \log_2\left[1+2p\left(1-p\right) \times\left(\frac{\sqrt{1+\kappa\left(1-\omega\right)A^2}\sqrt{1+\kappa\left(1-\omega\right)B^2}}{1+\frac{1}{2}\,\kappa\,\left(1-\omega\right)\left(A^2+B^2\right)+\frac{1}{4}\,\kappa\,\omega\left(A+B\right)^2}-1\right)\right], \tag{17}$$

where $\mathcal{C}(p,A,B)$ is the constraint set on the input. Simple algebraic manipulation yields (3).

B. Proof of Optimal Inputs - Section III

Proof of R1. Low SNR analysis ($\kappa << 1$):

Assume $\omega \neq 1$ (if $\omega = 1$, the solution to (4) is easily seen to be BPSK). Retaining the first two terms in a Taylor series expansion of (4) about $\kappa = 0$, we get

$$\max_{\mathcal{C}(p,A,B)} \mathcal{J}_{\mathrm{L}} \triangleq p\left(1-p\right)\left(A+B\right)^2 \times\left[\omega+\kappa\left\{\frac{1}{4}(A-B)^2(3\omega^2-6\omega+2)-\omega AB\right\}\right]. \tag{18}$$

† ***If*** $\omega < \overset{\star}{\omega}$. Consider the case where $B = 0$. The low SNR cost function (18) becomes

$$\mathcal{J}_{\mathrm{L}} = (1-p)\left[1+\frac{\kappa}{4}\frac{\phi(w)}{p}\right],$$

where $\phi(w) \triangleq 3\omega^2 - 6\omega + 2$, and where we have used the energy constraint. The cost function becomes arbitrarily large if $p \to 0$, with $pA^2 = 1$, provided $\phi(\omega) > 0$. Therefore, at low SNR and for $\omega < \overset{\star}{\omega}$, $\lim_{p\to 0} \mathrm{OOK}(p)$ is the optimal input.

† ***If*** $\omega > \overset{\star}{\omega}$. It follows that $3\omega^2-6\omega+2 < 0$. A consequence is that the optimal A in (18) must be finite. If not, $\mathcal{J}_{\mathrm{L}}$ will take on an arbitrarily large negative value. Since A is finite and κ small, we omit the κ term in (18). Removing other irrelevant terms, we seek

$$\max_{\mathcal{C}(p,A,B)} p(1-p)(A+B)^2. \tag{19}$$

The solution to (19) is any input of the form $\mathcal{C}_{p_0} = \left(p_0, \sqrt{\frac{1-p_0}{p_0}}, \sqrt{\frac{p_0}{1-p_0}}\right)$ where $p_0 \in (0, 1/2]$. Next, using $\mathcal{C}_{p_0}$ as a candidate set of possible solutions, we increase κ slightly, to consider the κ term and determine which $p_0 \in (0,1/2]$ maximizes (18) when $A = \sqrt{\frac{1-p_0}{p_0}}$ and $B = \sqrt{\frac{p_0}{1-p_0}}$. Substituting $\mathcal{C}_{p_0}$ into the above, and removing irrelevant terms (note that $p(1-p)(A+B)^2 = AB = 1$ for solutions in $\mathcal{C}_{p_0}$), we must solve

$$\max_{\mathcal{C}(p,A,B)} \left(\sqrt{\frac{p}{1-p}}-\sqrt{\frac{1-p}{p}}\right)^2(3\omega^2-6\omega+2),$$

which is maximized for $p_0 = 1/2$ (since $3\omega^2 - 6\omega + 2 < 0$). Therefore, the optimal input distribution is $\mathcal{C}_{1/2} = (\frac{1}{2}, 1, 1)$, or BPSK.

Proof of R2. High SNR Analysis ($\kappa \to \infty$):

Again, assume that $\omega \neq 1$. At high SNR, the minimization problem (4) becomes

$$\min_{\mathcal{C}(p,A,B)} \mathcal{J}_{\mathrm{H}} \triangleq p(1-p) \times\left\{\frac{\sqrt{1-w}\;A\;\gamma_B(\kappa)}{\frac{\sqrt{\kappa}}{2}\left[\left(1-\frac{\omega}{2}\right)\left(A^2+B^2\right)+\omega AB\right]}-1\right\},$$

where $\gamma_B(\kappa) = \{1, \text{ if } B = 0, \quad B\sqrt{\kappa(1-\omega)}, \text{ if } B \neq 0\}$. Note that $\mathcal{J}_{\mathrm{H}} \geq -p(1-p) \geq -1/4$, with equality for $A = \sqrt{2}$, $B = 0$, and $p = 1/2$. Therefore, OOK($\frac{1}{2}$) is optimal as $\kappa \to \infty$.

C. On-Off Keying Optimal for no CSI - Section III

Setting $\omega = 0$, we seek to minimize (4) with the constraint set $\mathcal{C}(p,A,B)$. Let $p = p_0 \leq \frac{1}{2}$ be fixed, and let $x = B^2$. Using the energy constraint, the minimization problem becomes

$$\min_{0\leq x\leq 1} f(x) = \min_{0\leq x\leq 1} \frac{\sqrt{\left(\frac{\kappa+p_0}{p_0}\right)+x\left[\frac{(2p_0-1)\kappa+\kappa^2}{p_0}\right]-x^2\kappa^2\frac{1-p_0}{p_0}}}{\left(\frac{p_0+\kappa/2}{p_0}\right)+x\frac{\kappa}{2}\left(\frac{2p_0-1}{p_0}\right)}. \tag{20}$$

It can be verified that $\partial f(x)/\partial x \geq 0$ for $x \in [0,1]$, implying that (20) is minimized for $x = B^2 = 0$. Therefore, when $\omega = 0$, an On-Off keying solution is optimal.

References

[1] I. Abou-Fayçal, M. Médard, and U. Madhow, "Binary adaptive coded pilot symbol assisted modulation over Rayleigh fading channels without feedback," *IEEE Trans. Commun.*, vol. 53, no. 6, pp. 1036–1046, June 2005.

[2] I. Abou-Fayçal, M. Trott, and S. Shamai. "The capacity of discrete-time memoryless Rayleigh-fading channels," *IEEE Trans. Inform. Theory*, vol. 47, no. 4, pp. 1290–1301, May 2001.

[3] S. Adireddy, L. Tong, and H. Viswanathan, "Optimal placement of training for frequency selective block-fading channels," *IEEE Trans. Inform. Theory*, vol. 48, no. 8, pp. 2338–2353, Aug. 2002.

[4] E. Arikan, "An upper bound on the cutoff rate of sequential decoding," *IEEE Trans. Inform. Theory*, vol. 34, no. 1, pp. 55–63, Jan. 1988.

[5] J. Baltersee, G. Fock, and H. Meyr, "An information theoretic foundation of synchronized detection," *IEEE Trans. Commun.*, vol. 49, no. 12, pp. 2115–2123, Dec. 2001.

[6] E. Biglieri, J. Proakis, and S. Shamai, "Fading channels: Information-theoretic and communications aspects," *IEEE Trans. Inform. Theory*, vol. 44, no. 6, pp. 2619–2692, Oct. 1998.

[7] X. Cai and G. Giannakis, "Adapative PSAM accounting for channel estimation and prediction errors," *IEEE Trans. Wireless Commun.*, vol. 4, no. 1, pp. 24–256, Jan. 2005.

[8] J. K. Cavers, "An analysis of pilot symbol assisted modulation for Rayleigh fading channels [mobile radio]," *IEEE Trans. Veh. Technol.*, vol. 40, no. 4, pp. 686–693, Nov. 1991.

[9] J. K. Cavers, "Variable-rate transmission for Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 20, no. 2, pp. 15–22, Feb. 1972.

[10] R.-R. Chen, B. Hajek, R. Koetter, and U. Madhow "On fixed input distributions for noncoherent communication over high SNR Rayleigh fading channels," *IEEE Trans. Inform. Theory*, vol. 50, no. 12, pp. 3390–3396, Dec. 2004.

[11] M. Dong, L. Tong, and B. Sadler, "Optimal insertion of pilot symbols for transmissions over time-varying flat fading channels," *IEEE Trans. Signal Processing*, vol. 52, no. 5, pp. 1403–1418, May 2004.

[12] X. Dong and L. Xiao, "Symbol error probability of two-dimensional signaling in Ricean fading with imperfect channel estimation," *IEEE Trans. Veh. Technol.*, vol. 54, no. 2, Mar. 2005.

[13] R. Gallager, *Information Theory and Reliable Communication*. New York: John Wiley and Sons, 1968.
[14] M. Garcia and J. Paez-Borrallo, "Tracking of time misalignments for OFDM systems in multipath fading channels," *IEEE Trans. Consumer Electron.*, vol. 48, no. 4, pp. 982–989, Nov. 2002.
[15] J. M. Geist, "The cutoff rate for on-off keying," *IEEE Trans. Commun.*, vol. 39, no. 8, pp. 1179–1181, Aug. 1991.
[16] D. Goeckel, "Adaptive coding for time-varying channels using outdated fading estimates," *IEEE Trans. Commun.*, vol. 47, no. 6, pp. 844–855, June 1999.
[17] B. Hassibi and B. Hochwald, "How much training is needed in multiple-antenna wireless links?," *IEEE Trans. Inform. Theory*, vol. 49, no. 4, pp. 951–963, Apr. 2003.
[18] A. O. Hero and T. L. Marzetta, "Cutoff rate and signal design for the quasi-static Rayleigh fading space-time channel," *IEEE Trans. Inform. Theory*, vol. 47, no. 6, pp. 2400–2416, Sep. 2001.
[19] J. Huang and S. Meyn, "Characterization and computation of optimal distributions for channel coding," in *Proc. 37th Annual Conf. Information Sciences Syst.*, Mar. 2003.
[20] W. C. Jakes, Jr., *Microwave Mobile Communication*. New York: Wiley, 1974.
[21] S. Jamali, and T. Le-Ngoc, *Coded-Modulation Techniques for Fading Channels*. Norwell, MA: Kluwer Publishers, 1994.
[22] W. Kuo and M. P. Fitz, "Frequency offset compensation of pilot symbol assisted modulation in frequency flat fading," *IEEE Trans. Commun.*, vol. 45, no. 11, pp. 1412–1416, Nov. 1997.
[23] K. Leeuwin-Boulle and J. C. Belfiore, "The cutoff rate of time-correlated fading channels," *IEEE Trans. Inform. Theory*, vol. 39, no. 2, pp. 612–617, Mar. 1993.
[24] X. Ma, G. Giannakis, and S. Ohno, "Optimal training for block transmissions over doubly selective wireless fading channels," *IEEE Trans. Signal Processing*, vol. 51, no. 5, pp. 1351–1366, May 2003.
[25] T. Marzetta and B. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading," *IEEE Trans. Inform. Theory*, vol. 45, no. 1, pp. 139–157, Jan. 1999.
[26] J. Massey, "Coding and moulation in digital communications," in *Proc. Int. Zurich Seminar Digital Commun.*, pp. E2(1)-E2(4), Mar. 1974.
[27] R. McEliece and W. Stark, "Channels with block interference," *IEEE Trans. Inform. Theory*, vol. 30, no. 1, pp. 44–53, Jan. 1984.
[28] S. Misra, A. Swami, and L. Tong, "Optimal training for time-selective wireless fading channels using cutoff rate," *EURASIP J. Applied Signal Processing*, vol. 2006, Article ID 47245, 2006.
[29] S. Ohno and G. Giannakis, "Capacity maximizing MMSE-optimal pilots for wireless OFDM over frequency-selective block Rayeligh-fading channels," *IEEE Trans. Inform. Theory*, vol 50, no. 9, pp. 2138–2145, Sep. 2004.
[30] A. Saleh and J. Salz, "On the computational cutoff rate, R_0, for the peak-power-limited Gaussian channel," *IEEE Trans. Commun.*, vol. 35, no. 1, pp. 13–20, Jan. 1987.
[31] S. Shamai and I. Bar-David, "The capacity of average and peak-power-limited quadrature Gaussian channels," *IEEE Trans. Inform. Theory*, vol. 41, no. 4, pp. 1060–1071, July 1995.
[32] L. Tong, B. Sadler, and M. Dong, "Pilot-assisted wireless transmissions," *IEEE Signal Processing Mag.*, pp. 12–25, Nov. 2004.
[33] S. Verdú, "Spectral efficiency in the wideband regime," *IEEE Trans. Inform. Theory*, vol. 48, no. 6, pp. 1319–1343, June 2002.
[34] S. Verdú, "On channel capacity per unit cost," *IEEE Trans. Inform. Theory*, vol. 36, no. 5, pp. 1019–1030, Sep. 1990.

Saswat Misra was born in College Park, MD, USA in 1978. He received the B.S. in electrical engineering from the University of Maryland at College Park in 2000 and the M.S. in electrical engineering from the University of Illinois at Urbana-Champaign in 2002. Since 2002, he has been a Research Scientist at the Army Research Laboratory (ARL) in Adelphi, MD in the Communications and Network Systems division. Mr. Misra has previously worked on optimal training design for wireless communication systems; an area in which he has published several papers and holds two patents (pending). Since Fall 2005, he has been a Ph.D. candidate at Cornell University. He is currently studying routing and security issues in wireless military networks.

Ananthram Swami (SM'96) is a Senior Research Scientist and Fellow of the Army Research Laboratory, Adelphi, MD, USA, where he works in the broad area of signal processing for communications. He received the B.S. degree from the Indian Institute of Technology, Bombay, India; the M.S. degree from Rice University, Houston, TX; and the Ph.D. degree from the University of Southern California, Los Angeles, all in electrical engineering.

Dr. Swami has held positions with Unocal Corporation, the University of Southern California, CS-3, and Malgudi Systems. He was a Statistical Consultant to the California Lottery, and has held visiting faculty positions at INP, Toulouse, France. Dr. Swami is chair of the IEEE Signal Processing Society's TC on Signal Processing for Communications, an associate editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, and of the IEEE TRANSACTIONS ON SIGNAL PROCESSING. He was co-organizer and co-chair of a 1999 ASA-IMA Workshop on Heavy-Tailed Phenomena, and of a 2002 ONR/NSF/ARO/CTA Workshop on "Future Challenges to Wireless Communications and Networking." He was co-guest editor of a 2004 special issue of the IEEE Signal Processing Magazine (SPM) on "Signal Processing for Networking." He is also co-guest editor of an upcoming IEEE SPM special issue on "Distributed Signal Processing in Sensor Networks," an EURASIP JASP special issue on "Reliable Communications over Rapidly Time-varying Channels," and an EURASIP JWCN special issue on "Wireless Mobile Ad Hoc Networks."

Lang Tong (F'05) is a Professor in the School of Electrical and Computer Engineering, Cornell University, Ithaca, NY, USA. He received the B.E. degree from Tsinghua University, Beijing, China, in 1985, and M.S. and Ph.D. degrees in electrical engineering in 1987 and 1991, respectively, from the University of Notre Dame, Notre Dame, Indiana. He was a Postdoctoral Research Affiliate at the Information Systems Laboratory, Stanford University in 1991. He was also the 2001 Cor Wit Visiting Professor at the Delft University of Technology.

Dr. Tong received the Young Investigator Award from the Office of Naval Research in 1996, the Outstanding Young Author Award from the IEEE Circuits and Systems Society, the 2004 Best Paper Award (with M. Dong) from the IEEE Signal Processing Society, and the 2005 Leonard G. Abraham Prize Paper Award (with P. Venkitasubramaniam and S. Adireddy) from the IEEE Communications Society. His areas of interest include statistical signal processing, wireless communications, communication networks and sensor networks, and information theory.