

Cutoff Rate Analysis of the Gauss-Markov Fading Channel with Binary Inputs and Partial CSI at the Receiver

Saswat Misra
Army Research Lab
Adelphi, MD
smisra@arl.army.mil

Ananthram Swami
Army Research Lab
Adelphi, MD
aswami@arl.army.mil

Lang Tong
Cornell University
Ithaca, NY
ltong@ece.cornell.edu

Abstract — We use cutoff rate to study the rate-optimal binary input distributions for use on the Rayleigh fading channel when only partial CSI is available at the receiver. First, we consider iid fading, and analyze the cutoff rate under partial CSI for BPSK and On-Off Keying (OOK) modulation. Then, we analytically determine the optimal distribution as a function of the faded SNR and the CSI quality available at the receiver. Lastly, we consider a temporally-correlated Gauss-Markov fading model, and discuss attainable rates for this channel when a PSAM estimation scheme is combined with adaptive binary modulation based on the CSI quality available at the receiver.

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 popularity of binary input distributions for analysis lies in their analytic simplicity, and optimality (or near optimality) at low SNR ([11], [1]).

Consider a system operating over the memoryless discrete-time Rayleigh fading channel with AWGN in which the transmitter can select among the general class of binary input distributions, subject to an average energy constraint per symbol. When full CSI is available at the receiver, it is well known that antipodal signaling (BPSK) maximizes the capacity of this system. Conversely, without any CSI available at the receiver, on-off keying (OOK) has been shown to be optimal [1]. However, when only partial CSI is available at the receiver, it is not clear as to which strategy, even among these two, is optimal.

In this paper, we investigate this intermediate case, when only partial CSI is available the receiver. We assume that partial CSI is provided to the receiver through an MMSE channel estimate. First, we consider the memoryless Rayleigh fading channel. We confine our interest to the two limiting constellations of BPSK and OOK, and analytically find the design rule which specifies, for a given SNR and CSI quality, when OOK is better than BPSK and vice-versa. Finally, we consider a system operating over the temporally-correlated Gauss-Markov fading channel, in which a PSAM estimation scheme is used for channel estimation [8]. We analyze the performance of an adaptive modulation scheme in which the modulate-type is adapted according to the SNR and CSI quality available at the receiver.

We will use the cutoff rate R_o as an information-theoretic metric for characterizing the reliable rates R for the systems we study. The use of cutoff rate R_o in place of capacity has been common since its reintroduction in [7], and discussions on

cutoff rate for no CSI and full CSI may be found in [5] and [6]. The cutoff rate is a lower bound on capacity and the function $R_o - R$ can be used in place of the random coding exponent to characterize the entire rate vs. performance curve. The use of cutoff rate in place of capacity often leads to tractable results that would not be available otherwise. For low-complexity, low-delay systems, the cutoff rate is often viewed as a “practical channel capacity.”

II. THE RAYLEIGH FADING CHANNEL WITH PARTIAL CSI

System Model

We start by considering a single-user, time-varying fading channel. We will consider the iid Rayleigh model described via

$$y_k = \sqrt{E} h_k s_k + n_k, \quad (1)$$

where k denotes discrete time, $h_k \sim \mathcal{CN}(0, \sigma_h^2)$ models iid fading, y_k is the signal at the receiver, E is the average symbol energy used at the transmitter, and $n_k \sim \mathcal{CN}(0, \sigma_N^2)$ models AWGN.¹ The channel input $s_k \in \{A_1, A_2\}$ is binary modulation subject to the normalized energy constraint $pA_1^2 + (1-p)A_2^2 = 1$, where $0 \leq p \leq 1$ is the probability of transmitting A_1 . We assume that A_1, A_2 are real, and that $\sigma_N^2 \neq 0$, $\sigma_h^2 \neq 0$.

During each symbol interval, we assume that the receiver has an estimate of the channel, \hat{h}_k , so that (1) can be written as

$$y_k = \sqrt{E} \hat{h}_k s_k + \sqrt{E} \tilde{h}_k s_k + n_k, \quad (2)$$

where $\tilde{h}_k = h_k - \hat{h}_k$ is the residual error in the channel estimate. Specifically, we will assume that an MMSE estimator is used. Because the MMSE estimator of the Gaussian channel is linear, both the estimate and the residual error are Gaussian, i.e., $\hat{h}_k \sim \mathcal{CN}(0, \hat{\sigma}^2)$ and $\tilde{h}_k \sim \mathcal{CN}(0, \tilde{\sigma}^2)$. Also, this implies that $\hat{\sigma}^2 + \tilde{\sigma}^2 = \sigma_h^2$. We let

$$\omega = \hat{\sigma}^2 / \sigma_h^2$$

denote the quality of the channel estimate.² Note that $\omega = 0$ denotes no CSI, and $\omega = 1$ denotes perfect CSI at the receiver. In Section III, we describe one manner in which the channel estimate can be obtained, however, this information is not relevant to the following discussion.

¹In this paper $\mathcal{CN} \sim (\mu, \sigma^2)$ denotes a complex Gaussian random variable with mean μ and with independent real and imaginary parts, each having variance $\sigma^2/2$.

²Henceforth, we will refer to the CSI quality as the *estimator quality*, and denote it by ω .

General Binary Signaling

For general binary signaling, we have shown that the unquantized ergodic cutoff rate for a memoryless fading channel with partial CSI at the receiver is [10]

$$R_o = \max_Q -\log_2 \sum_{s_1=\{A_1, A_2\}} \sum_{s_2=\{A_1, A_2\}} Q(s_1)Q(s_2) \times \mathcal{E}_{\hat{h}} \left[\int_{y \in Y} \sqrt{P(y|s_1, \hat{h})P(y|s_2, \hat{h})} dy \right], \quad (3)$$

where $Q(A_1) = p$, $Q(A_2) = 1 - p$, $\mathcal{E}_{\hat{h}}$ denotes expectation w.r.t. the pdf of the channel estimate, \hat{h} , and $P(y|s, \hat{h})$ is the pdf of the received signal, conditioned upon the transmitted signal and the channel estimate. For the Rayleigh fading channel in AWGN described by (2), this yields

$$R_o = \max_{pA_1^2 + (1-p)A_2^2 = 1} -\log_2 \{1 + 2p(1-p) [\Gamma(A_1, A_2) - 1]\},$$

where

$$\Gamma(A_1, A_2) = \frac{\{[1 + \kappa(1-\omega)A_1^2][1 + \kappa(1-\omega)A_2^2]\}^{1/2}}{\{1 + \frac{\kappa}{2}[(1-\omega)(A_1^2 + A_2^2) + \frac{1}{2}\omega(A_1 - A_2)^2]\}}, \quad (4)$$

and where $\kappa \triangleq \sigma_h^2 E / \sigma_N^2$ is the faded SNR.

Cutoff Rate for BPSK

First, we consider the case where full CSI is available at the receiver. Letting $\omega = 1$ in (4), we find that the maximum occurs at $A_1 = -A_2$, and $p = 1/2$. Therefore, for full CSI at the receiver, BPSK modulation maximizes the cutoff rate. The unquantized ergodic cutoff-rate for BPSK is given by

$$R_{o,B} = -\log_2 \left\{ \frac{1}{2} + \frac{1}{2} \left[\frac{1 + \kappa(1-\omega)}{1 + \kappa} \right] \right\}. \quad (5)$$

Figure 1 shows the cutoff rate as a function of the faded SNR κ and the estimator quality ω . We make the following observations:

1. The estimator quality ω places an asymptotic ceiling on $R_{o,B}$. For large κ the cutoff rate saturates to

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

2. When $\omega = 0$ (i.e., no CSI is available), information transmission is not possible. This is because the statistics of y_k at the receiver are independent of s_k . I.e.,

$$y_k | s_k = 1 \sim \mathcal{CN}(0, E\sigma_h^2 + \sigma_N^2), \\ y_k | s_k = -1 \sim \mathcal{CN}(0, E\sigma_h^2 + \sigma_N^2).$$

3. More generally, the statistics of y_k under the two hypotheses, and conditioned upon the known part of the channel \hat{h}_k , are

$$y_k | \hat{h}_k, s_k = 1 \sim \mathcal{CN}(\sqrt{E} \hat{h}_k, E\tilde{\sigma}^2 + \sigma_N^2), \\ y_k | \hat{h}_k, s_k = -1 \sim \mathcal{CN}(-\sqrt{E} \hat{h}_k, E\tilde{\sigma}^2 + \sigma_N^2).$$

The ability to distinguish between the two hypotheses is only through the difference in the means, and therefore

it is critical that $\omega > 0$. When $\kappa \gg \frac{1}{1-\omega}$ (i.e., when $E\tilde{\sigma}^2 \gg \sigma_N^2$), the statistics become

$$y_k | \hat{h}_k, s_k = 1 \sim \mathcal{CN}(\sqrt{E} \hat{h}_k, E\tilde{\sigma}^2), \\ y_k | \hat{h}_k, s_k = -1 \sim \mathcal{CN}(-\sqrt{E} \hat{h}_k, E\tilde{\sigma}^2).$$

Increasing κ scales the variance and power in the mean equally, and so for large κ , i.e., $\kappa \gg \frac{1}{1-\omega}$, performance saturates as depicted in Fig. 1.

Next, we consider the case of no CSI at the receiver.

Cutoff Rate for OOK

When no CSI is available at the receiver (i.e., $\omega = 0$), it has been shown in [1] that the capacity-optimal input distribution at low SNR is OOK. Also, it was shown that, restricted to binary distributions, OOK is capacity optimal at all SNR. Here, if we assume equiprobable inputs (i.e., $p = 1/2$), we find that OOK modulation also maximizes the cutoff rate of (4) for all κ .

The general expression for OOK is obtained by setting $A_1 = A$ and $A_2 = 0$ in (4). This yields the unquantized ergodic cutoff rate

$$R_{o,K} = \max_{0 < p < 1} -\log_2 \left\{ 1 + 2p(1-p) \left[\frac{\sqrt{1 + \kappa(1-\omega)\frac{1}{p}}}{1 + \kappa(2-\omega)\frac{1}{4p}} - 1 \right] \right\}, \quad (6)$$

where we have used that fact that $A^2 = \frac{1}{p}$ due to the energy constraint. In general, it is not possible to analytically maximize (6) over p , as it leads to solving a high-order polynomial that has no explicit solution as a function of κ and ω . However, it can be verified analytically that as $\kappa \rightarrow \infty$, $p^* = 1/2$, and that as $\kappa \rightarrow 0$, $p^* = 0$ (this corresponds to no information transmission). In general $0 \leq p^* \leq 1/2$. In Figure 2, we plot p^* as a function of κ for several values of ω . We see that for moderate to large values of κ , letting $p = 1/2$ is a reasonable approximation to p^* . Using $p = 1/2$, the cutoff rate becomes

$$R_{o,K} = -\log_2 \left\{ \frac{1}{2} + \frac{1}{2} \left[\frac{\sqrt{1 + 2\kappa(1-\omega)}}{1 + \kappa(1-\frac{\omega}{2})} \right] \right\}. \quad (7)$$

Since $A^2 = \frac{1}{p}$, we can equivalently repeat Figure 2 in terms of the optimal OOK amplitude A^* versus faded SNR κ as shown in Figure 3. Note that as κ increases, the optimal OOK amplitude A^* decreases and (correspondingly) p^* increases. This trend was shown in [1] for the capacity metric and for no CSI ($\omega = 0$). From the figure, we see that when $\omega > 0$ this general trend is still true, and that for fixed κ , the optimal amplitude A^* is a decreasing function of ω . As κ becomes large, $A^* \rightarrow \sqrt{2}$, and correspondingly, $p^* \rightarrow \frac{1}{2}$.

OOK offers significant advantages over BPSK for achieving high transmission rates. In Figure 4, we repeat Figure 1, this time for OOK, with $p = 1/2$, as given by (7). As $\kappa \rightarrow \infty$, $R_{o,K}$ approaches 1 regardless of ω . In particular, the cutoff rate is non-zero when $\omega = 0$, which was not the case for BPSK.

To gain some insight into the figure, we look at the statistics of y_k under the two hypotheses

$$y_k | \hat{h}_k, s_k = 0 \sim \mathcal{CN}(0, \sigma_N^2), \\ y_k | \hat{h}_k, s_k = \sqrt{2} \sim \mathcal{CN}(\sqrt{2E} \hat{h}_k, 2E\tilde{\sigma}^2 + \sigma_N^2).$$

The distance between the means is obviously reduced compared to that for BPSK, however the variance terms are now distinct. We expect that if the difference in the variance terms is large enough (i.e., if κ is large enough), then OOK will be able to outperform BPSK despite the decreased separation between means. Conversely, for small κ (when the variance terms are nearly identical), we expect BPSK to outperform OOK. The general tradeoff is shown in Figure 5 for $\omega = 0.8$. In the figure, we also plot OOK where the optimum p is used for each value of κ . Note that, even for small κ , the difference between the two OOK curves is small; note also that there is a significant gain in using OOK instead of BPSK at moderate-to-large κ .

In particular, for the no CSI scenario ($\omega = 0$) the hypotheses become

$$y_k | \hat{h}_k, s_k = 0 \sim \mathcal{CN}(0, \sigma_N^2),$$

$$y_k | \hat{h}_k, s_k = \sqrt{2} \sim \mathcal{CN}(0, 2E\sigma_h^2 + \sigma_N^2),$$

and so, unlike BPSK, information transmission is possible, particularly for large κ . This is confirmed in Figure 4.

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

Comparing BPSK and OOK

Here we provide results for the transitional value of the faded SNR $\bar{\kappa}$, above which OOK is optimal, and below which BPSK is optimal. In doing so, we provide an initial characterization of the intermediate region where partial CSI is available.

In [10], we will provide results which detail the optimal binary modulation scheme as a function of κ and ω . Here, we confine our interest to the two limiting cases of BPSK (optimal for full CSI) and OOK (optimal for no CSI). The results presented here give an *analytic basis* for an adaptive modulation scheme in which the transmitter can select between OOK and BPSK based on the faded SNR κ and estimator quality ω available at the receiver.

The transitional faded SNR $\bar{\kappa}$ is found by equating (5) and (7). This yields a third-order polynomial. Retaining the relevant root yields

$$\bar{\kappa} = f(\omega),$$

where

$$f(\omega) = \left[\frac{(a+b)^{1/3} + (a-b)^{1/3} - 2(4-10\omega+3\omega^2)}{3(2-\omega)^2(1-\omega)} \right],$$

and where

$$a = 81\omega^6 - 468\omega^5 + 828\omega^4 - 640\omega^3 + 624\omega^2 - 192\omega + 64,$$

$$b = 6\sqrt{3}(\omega-2)^2\omega^2\sqrt{61\omega^4 - 208\omega^3 + 168\omega^2 - 64\omega + 16}. \quad (8)$$

The function $f(\omega)$ depends on the estimator quality, and is shown in Figure 6. At the end points, $\omega = \{0, 1\}$, our results agree with existing theory:

1. Observe that $f(0) = 0$. Therefore, when no CSI is available, it is always optimal to use OOK instead of BPSK. This is in agreement with the results of [1], which are for the $\omega = 0$ case.

2. It can be verified that

$$\lim_{\omega \rightarrow 1} \bar{\kappa} = \infty,$$

which confirms that, when full CSI is available, BPSK is always optimal independent of the faded SNR. This is in agreement with the well known fact that, for AWGN channels, the use of OOK is suboptimal to BPSK for a fixed average symbol energy constraint.

III. THE GAUSS-MARKOV FADING CHANNEL WITH PARTIAL CSI

Here, we generalize the Rayleigh fading channel of (1) to incorporate temporal correlation. We will model this temporally correlated channel as a first-order Gauss-Markov process described via

$$y_k = \sqrt{E} h_k s_k + n_k \quad (\text{observation}),$$

$$h_k = \alpha h_{k-1} + z_k \quad (\text{state}), \quad (9)$$

with the same definitions as in (1), but now with $0 < \alpha < 1$ describing the temporal correlation. The parameter α is related to the normalized Doppler spread. The driving noise is $z_k \sim \mathcal{CN}(0, \sigma_h^2(1 - \alpha^2))$. We let $s_k \in \{-1, 1\}$ for BPSK modulation, or $s_k \in \{0, \sqrt{2}\}$ for OOK modulation.

Channel Estimation

To provide partial CSI at the receiver, we use pilot-symbol assisted modulation (PSAM) and MMSE channel estimation. We will use a periodic PSAM methodology in which a single training (pilot) symbol is periodically inserted into the symbol stream every T symbol slots. Motivation for sending one pilot symbol, rather than many, may be found in [2], [4], and [3]. This PSAM channel estimation scheme was first utilized in [8] to characterize reliable rates of the Gauss-Markov channel using a capacity metric.

At the receiver, a MMSE channel estimator is used as follows: Upon the reception of a training symbol, the current channel state is estimated, and the next $T - 1$ channel states are predicted using only the *most recent* training symbol. The MMSE estimator equations are³:

$$\hat{h}_{mT} = \mathcal{E}[h_{mT} | y_{mT}] = \frac{\sqrt{E}\sigma_h^2}{\sigma_N^2 + E\sigma_h^2} y_{mT},$$

$$\hat{h}_{mT+\ell} = \alpha^\ell \hat{h}_{mT} \quad \text{for } 1 \leq \ell \leq T-1. \quad (10)$$

The pdf of the estimator (10) is given by:

$$\hat{h}_{mT+\ell} \sim \mathcal{CN}(0, \hat{\sigma}_\ell^2), \quad \hat{\sigma}_\ell^2 = \alpha^{2\ell} \sigma_h^2 \frac{\kappa}{1 + \kappa}, \quad (11)$$

for $0 \leq \ell \leq T-1$. Note that the quality of the estimator decays exponentially with ℓ . The quality of the channel estimator is now a function of ℓ and also has period T . In particular (for $\ell \neq 0$)

$$w_\ell = \frac{\hat{\sigma}_\ell^2}{\sigma_h^2} = \alpha^{2\ell} \frac{\kappa}{1 + \kappa}, \quad \text{where } \ell = k \bmod T.$$

Note that by treating the Gauss-Markov channel in this manner we have essentially created T separate sub-channels,

³Note that the training symbols are given the same energy as the data symbols.

where two successive realizations of the same sub-channel are separated in time by T symbol periods as shown in Figure 8. Therefore, if we consider separate coding on each of the $T - 1$ data-bearing sub-channels via codebook multiplexing (see Figure 9), we can define an *average cutoff rate* R_A as the average cutoff rate over the set of T channels,

$$R_A = \frac{1}{T} \sum_{l=1}^{T-1} R_{o,l}$$

where $R_{o,l}$ is the cutoff rate of the l^{th} channel. For more details on R_A , see [9].

BPSK/OOK Adaptive Modulation

We will now consider an adaptive modulation scheme where the transmitter chooses between BPSK and OOK for each sub-channel, based on the faded SNR κ and the quality of the channel estimator ω at the receiver. The CSI feedback required at the transmitter consists only of κ , and α , the Doppler parameter. Recall from the previous subsection that the estimator quality ω is determined by the parameters α and κ . Equation (8) dictates the choice between BPSK and OOK.

In [8], the authors used capacity as a metric and considered a similar analysis, where the transmitter was free to choose the optimal binary distribution for each sub-channel (among all possible binary distributions). Because of the intractability of both the capacity metric and the input design rule, the results presented in [8] were purely numerical. Here, we consider a scheme that alternates between OOK and BPSK using only analytic design rules. We compare our BPSK/OOK adaptive modulation system to:

- C1. The BPSK-only system proposed in [9], where the optimal BPSK modulation is used in each sub-channel. Denote the cutoff rate of this system by $R_{A,\text{BPSK}}$.
- C2. The OOK-only system which uses OOK with $p = 1/2$ in each sub-channel. Denote the cutoff rate of this system by $R_{A,\text{OOK}}$; see (7).
- C3. The p -optimal OOK-only system in which OOK is used in each sub-channel, but is optimized over the probability p at each value of the faded SNR κ . This will be computed numerically for comparison. Denote the cutoff rate of this system by R_{A,OOK_p} .

The system has T statistically distinct sub-channels, of which $T - 1$ are information-bearing. From (8), we know that OOK is optimal for the l^{th} sub-channel if (for $1 \leq l \leq T - 1$)

$$\kappa \geq f(\omega_\ell) = f\left(\alpha^{2\ell} \frac{\kappa}{1 + \kappa}\right). \quad (12)$$

The right-hand side of the above equation is a decreasing function of ℓ . Therefore, sub-channels are initially assigned BPSK modulation. Eventually the transitional SNR is reached, and the remaining sub-channels are assigned OOK modulation. The cutoff rate $R_{o,H}$ for our BPSK/OOK adaptive modulation scheme is

$$R_{A,H} = -\frac{1}{T} \sum_{\ell=0}^{\mathcal{L}} \log_2 \left\{ \frac{1}{2} + \frac{1}{2} \left[\frac{1 + \kappa(1 - \omega_\ell)}{1 + \kappa} \right] \right\} - \frac{1}{T} \sum_{\ell=\mathcal{L}+1}^{T-1} \log_2 \left\{ \frac{1}{2} + \frac{1}{2} \left[\frac{\sqrt{1 + 2\kappa(1 - \omega_\ell)}}{1 + \kappa \left(1 - \frac{\omega_\ell}{2}\right)} \right] \right\}$$

where $1 \leq \mathcal{L} \leq T - 1$ is determined from (12). After substitution of ω_ℓ , this yields

$$R_{A,H} = \frac{T-1}{T} - \frac{1}{T} \sum_{\ell=0}^{\mathcal{L}} \log_2 \left\{ 2 - \alpha^{2\ell} \frac{\kappa^2}{(1 + \kappa)^2} \right\} - \frac{1}{T} \sum_{\ell=\mathcal{L}+1}^{T-1} \log_2 \left\{ 1 + \frac{\sqrt{1 + 2\kappa \left(1 - \alpha^{2\ell} \frac{\kappa}{1 + \kappa}\right)}}{1 + \kappa \left(1 - \frac{\alpha^{2\ell} \kappa}{2(1 + \kappa)}\right)} \right\}.$$

For a fair comparison, each system in C1-C3 should be optimized for the training period T that optimizes its cutoff rate. Indeed, for the OOK based systems, it is possible that sending no training data may be optimal since the cutoff rate is non-zero when $\omega_\ell = 0$. This is a significant issue that we will address in [10]. Here, we will consider a system with $\alpha = 0.98$, and use the results of [9], which show that for a BPSK-only system, the optimal training period at large κ is $T = 7$ and that the cutoff-rate saturates to 0.71.⁴ We will let $T = 7$.

In Figure 7, we plot the cutoff rate of the BPSK scheme $R_{A,\text{BPSK}}$, the unoptimized OOK scheme $R_{A,\text{OOK}}$, the optimized OOK scheme R_{A,OOK_p} , and the BPSK/OOK adaptive modulation scheme $R_{A,H}$. For small κ , BPSK is optimal for all sub-channels (from (8)), and so the average cutoff-rate $R_{A,H}$ is equal to $R_{A,\text{BPSK}}$. For the intermediate values of κ , BPSK is optimal for the initial sub-channels, while OOK is optimal for the latter ones. In this region, $R_{A,H}$ is larger than $R_{A,\text{BPSK}}$. To find the κ above which $R_{A,H}$ becomes larger than $R_{A,\text{BPSK}}$ we solve for κ in the following equation

$$\kappa = f\left(\alpha^{2(T-1)} \frac{\kappa}{1 + \kappa}\right), \quad (13)$$

which, with $\alpha = .98$ and $T = 7$, indicates that the OOK/BPSK hybrid scheme outperforms the BPSK-only scheme starting at $\kappa \approx 9.8$ dB. This is confirmed in the figure.

For large κ , $R_{A,H}$ is equal to $R_{A,\text{OOK}}$ since OOK outperforms BPSK in all sub-channels. To find the κ above which the OOK only scheme performs as well as the BPSK/OOK adaptive scheme we solve for κ in

$$\kappa = f\left(\alpha^2 \frac{\kappa}{1 + \kappa}\right), \quad (14)$$

which yields the intersection point to be $\kappa \approx 20.73$ dB.

At any value of the faded SNR κ , the BPSK/OOK scheme performs *at least* as well as the best of the BPSK or OOK only approaches, and for some intermediate range of κ , the adaptive scheme performs better than the best of either the OOK only or BPSK only schemes. We see that the cutoff rates for the two OOK approaches $R_{A,\text{OOK}}$ and R_{A,OOK_p} are roughly equal; note again that substantial gains are obtained by using OOK in the latter slots. A more extensive analysis will be given in [10].

⁴In [9], we consider a BPSK-only system, and study a more general problem in which we find the optimal training period and energy allocation between the data and training symbols subject to an average symbol energy constraint. Here, we assume that the data and training energies are equal (as in [8]), however, the value of the optimal training period and saturation cutoff-rate do not change from the values given in [9].

IV. DISCUSSION

We have considered the use of BPSK (optimal for no CSI) and OOK (shown to be the cutoff rate optimal binary strategy at all SNR) for use on Rayleigh fading channel, when only partial CSI is available at the receiver. We found the optimal modulation type as a function of faded SNR and CSI quality at the receiver. We also discussed the use of adaptive BPSK/OOK modulation on Gauss-Markov channel in which each data sub-channel uses the modulation type that is optimal within that sub-channel based on the faded SNR and CSI quality available. We have shown that use of this adaptive modulation scheme results in transmission rates that, at each SNR, are least as large as the larger of the BPSK-only or OOK-only modulation rates.

We are currently studying the general binary case (i.e., not limited to OOK or BPSK), channel estimators considered here, and more general channel models such as the Ricean and Nakagami fading channels. Results on optimal allocation of energy, where the data sub-channels may have variable energy and variable modulation, as well as the optimal training period, subject to an average energy constraint, will be reported elsewhere. Extensions to the frequency selective case (e.g., using an OFDM scheme) and to a general MIMO model are also being studied.

REFERENCES

- [1] I. Abou-Faycal, M. Trott, S. Shamai. "The Capacity of Discrete-Time Memoryless Rayleigh-Fading Channels," *IEEE Trans. Info. Theory*, May 2001.
- [2] S. Adireddy, L. Tong, H. Viswanathan, "Optimal Placement of Training for Frequency Selective Block-Fading Channels", *IEEE Trans. Info. Theory*, 48(8), 2338–2353, Aug 2002.
- [3] J.K. Cavers, "An analysis of pilot symbol assisted modulation for Rayleigh fading channels [mobile radio]", *IEEE Trans. Veh. Tech.*, 40(4), 686-693, Nov 1991.
- [4] M. Dong, L. Tong, and B. Sadler, "Training placement for tracking fading channels," *Proc. IEEE ICASSP*, May 2002.
- [5] A.O. Hero and T.L. Marzetta, "Cutoff rate and signal design for the quasi-static Rayleigh fading space-time channel", *IEEE Trans. Info. Theory*, 47(6), 2400-16, Sept 2001.
- [6] S. Jamali, and T. Le-Ngoc, *Coded-Modulation Techniques for Fading Channels*. Kluwer Academic Publishers, April 1994.
- [7] J. Massey, "Coding and Modulation in Digital Communications," In *Proc. 1974 Int. Zurich Seminar., Digital Communication*, March 1974, pp.E2.1-E2.4.
- [8] M. Medard, I. Abou-Faycal, and U. Madhow, "Adaptive coded modulation without channel feedback for pilot symbol assisted modulation," *Proc. 38th Annual Allerton Conference on Communication, Control, and Computing*, Oct 2002.
- [9] S. Misra, A. Swami, and L. Tong, "Cutoff Rate of the Gauss-Markov channel with PSAM estimation and BPSK modulation.," to appear in *Proc. IEEE Workshop on Signal Processing Advances in Wireless Communications*, Rome, June 2003.
- [10] S. Misra, A. Swami, and L. Tong, "Binary Input Distributions for Fading Channels with Partial CSI at the Receiver.," Journal Version, In preparation.
- [11] J. Proakis, *Digital Communications*, McGraw Hill, 4ed.

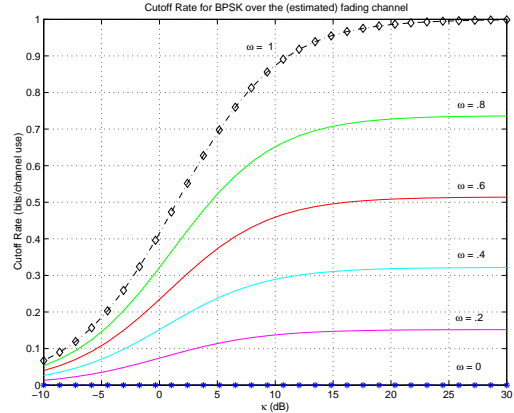


Figure 1: The BPSK cutoff rate $R_{o,B}$ vs. faded SNR κ (dB), for different values of the estimator quality ω .

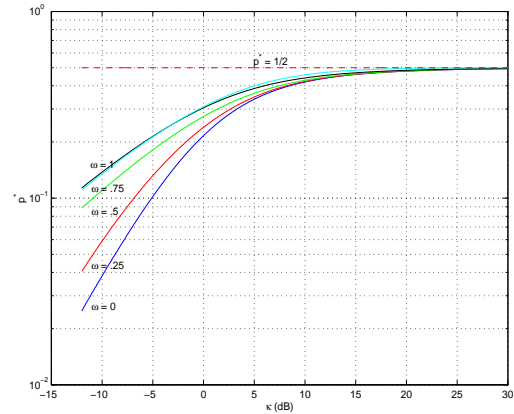


Figure 2: The optimal probability of transmitting a '1' (A), p^* vs. faded SNR κ (dB), for different values of the estimator quality ω .

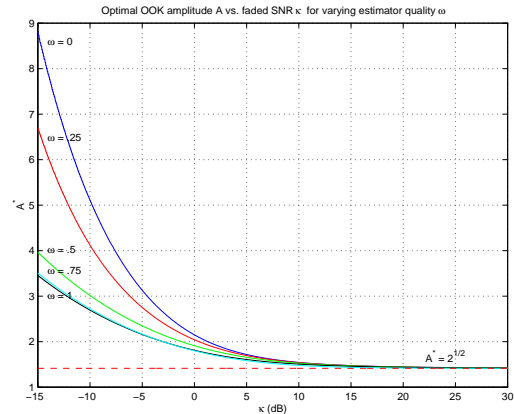


Figure 3: The optimal OOK amplitude A^* vs. faded SNR κ (dB), for different values of the estimator quality ω .

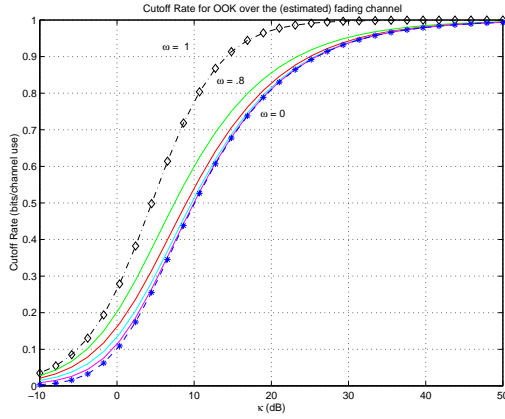


Figure 4: The OOK cutoff rate $R_{o,K}$ vs. faded SNR κ (dB), for different values of the estimator quality ω .

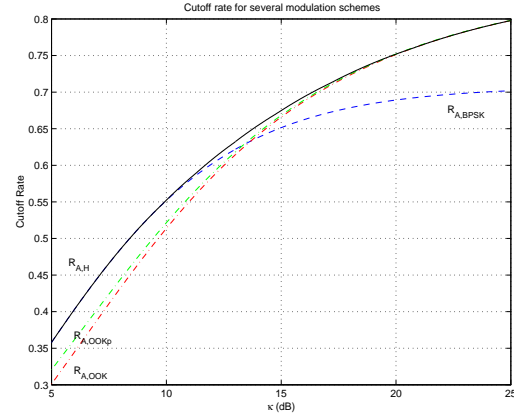


Figure 7: A comparison of the cutoff rate for several modulation techniques. $R_{A,BPSK}$ denotes BPSK only. $R_{A,OOK}$ denotes the use of OOK with $p = 1/2$. $R_{A,OOKp}$ denotes the use of OOK where p has been maximized at each value of the faded SNR κ . $R_{A,H}$ denotes the BPSK/OOK adaptive modulation scheme.

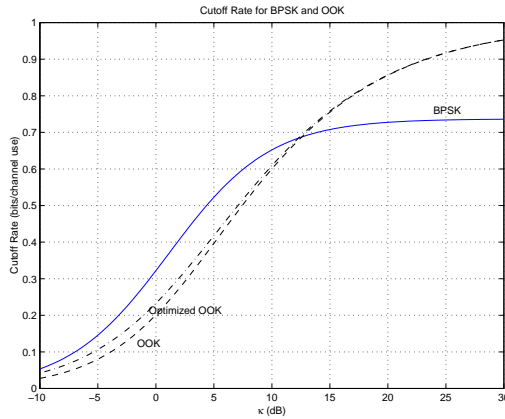


Figure 5: A comparison of the BPSK cutoff rate $R_{o,B}$, the unoptimized OOK cutoff rate $R_{o,K}$, and the optimized OOK cutoff rate vs. κ (dB) for $\omega = 0.8$.

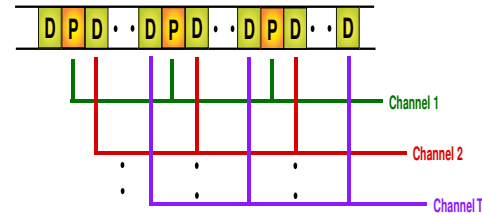


Figure 8: Depicts the separation of the Gauss-Markov channel into T distinct sub-channels

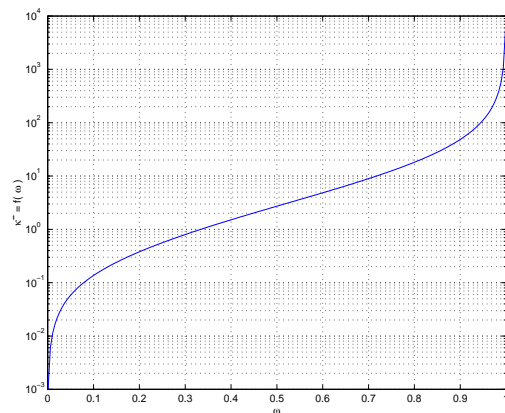


Figure 6: The transitional faded SNR $\bar{\kappa} = f(\omega)$. For larger faded SNR κ OOK is cutoff-rate optimal, and for smaller faded SNR, BPSK is optimal.

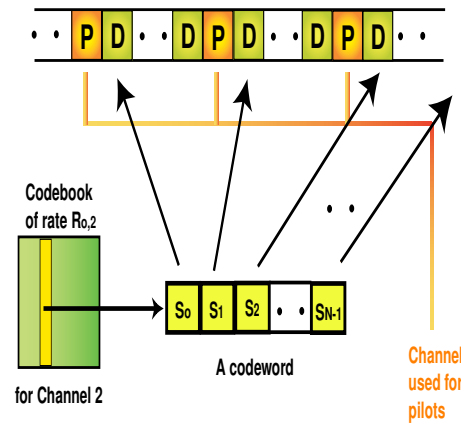


Figure 9: Separate coding for each of the $T - 1$ data-bearing channels. The ℓ^{th} channel uses a codebook of rate $R_{o,\ell}$ for $1 \leq \ell \leq T - 1$.