

CUTOFF RATE ANALYSIS OF THE GAUSS-MARKOV FADING CHANNEL WITH ADAPTIVE ENERGY ALLOCATION

Saswat Misra¹, Ananthram Swami¹, and Lang Tong²

¹ Army Research Lab, Adelphi, MD 20783, USA

² Cornell University, Ithaca, NY, USA

{smisra,aswami}@arl.army.mil, ltong@ece.cornell.edu

ABSTRACT

We use the cutoff rate to analyze reliable rates over the first-order Gauss-Markov fading channel when an optimized pilot-symbol assisted modulation (PSAM) scheme is used to provide partial CSI to the receiver. The transmitter uses BPSK modulation, and has knowledge of the SNR at the receiver and the normalized doppler spread only. Under an average energy constraint, closed form expressions are found for the optimal training and data energies, for both the adaptive and non-adaptive cases. For the non-adaptive case, a lower bound on the optimal training period, which is exact at moderate to high SNR, is given.

1. INTRODUCTION

We investigate reliable rates for single user transmission over a temporally-correlated Rayleigh fading channel, which we model as a first-order Gauss-Markov process. We use the average cutoff rate as an information-theoretic metric.

At the outset, the channel state is unknown to both the transmitter and receiver. To learn the channel, we use *pilot-symbol assisted modulation* (PSAM) [3], and employ a MMSE channel estimator, thereby providing partial CSI to the receiver. By optimizing the cutoff rate over the training parameters (energy allocation and the training period), we characterize the reliable rates on the Gauss-Markov channel for an optimized PSAM approach.

Previous work on reliable rates for fading channels using an optimized PSAM channel estimation scheme have been primarily for the (block) iid Rayleigh fading scenario, and have looked at the maximization of (bounds on) mutual information [12], [6]. A minimax optimization leads to a Gaussian input distribution, and makes the analysis tractable. It has been shown in [4] that the Gaussian input is ill-suited for consideration on the Gauss-Markov channel, as it produces bounded mutual information at high SNR. The results of [1] indicate that the capacity-achieving distribution for the iid Rayleigh scenario without CSI is a discrete one, and that at low SNR, binary ON-OFF keying (OOK) is optimal.

The paper most related to our work is [10], which also considers reliable rates for the Gauss-Markov fading model when partial CSI is provided to the receiver via a PSAM/MMSE estimation scheme. The mutual information metric is used, and its intractability leads to a numerical analysis. The authors use the optimal binary input distribution at the transmitter, however, the transmission energy is kept constant.

Here, we consider a similar framework, and obtain analytic results for the adaptive energy scenario. Our results detail the optimal energy allocation between the pilot and data, and we find a

bound on the optimal training period (the period between training symbols) that is exact at moderate to high SNR. We assume antipodal BPSK modulation. We consider the case of adaptive modulation in [11].

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 [9], and discussions on cutoff rate for no CSI and full CSI may be found in [7] and [8]. The cutoff rate is a lower bound on capacity and the function $R_o - R$ may 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.”

2. SYSTEM MODEL

We describe the channel model, the training scheme, and the channel estimation algorithm.

2.1. The Gauss-Markov Channel

The channel of interest is a Rayleigh flat fading model. Typically the Jakes model is used to describe the temporal correlation of the fading process. It is known that second- and third-order AR (Gauss-Markov) models provide excellent fits to the Jakes model. The higher-order models are not analytically tractable and so do not provide ready insights. We consider the first order Gauss-Markov fading channel described via

$$\begin{aligned} h_k &= \alpha h_{k-1} + z_k && \text{(state),} \\ y_k &= \sqrt{E_k} h_k s_k + n_k && \text{(observation),} \end{aligned} \quad (1)$$

where k denotes discrete time, $h_k \sim \mathcal{CN}(0, \sigma_h^2)$ models fading, $0 < \alpha < 1$ describes the channel correlation and is related to the normalized Doppler spread; $z_k \sim \mathcal{CN}(0, \sigma_h^2(1 - \alpha^2))$ is driving noise, y_k is the signal at the receiver, E_k is the energy of the k -th symbol, $n_k \sim \mathcal{CN}(0, \sigma_N^2)$ models AWGN, and $s_k \in \{-1, 1\}$ (antipodal BPSK modulation).¹ We assume that $\sigma_N^2, \sigma_h^2 \neq 0$.

2.2. Channel Estimation

Partial CSI is available at the receiver via pilot-symbol assisted modulation and MMSE channel estimation. We will use a periodic

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 $\sigma^2/2$.

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], [5], and [3]. Given periodic training, we assume that the energy allocation is also periodic, i.e., the received data are

$$y_k = \sqrt{E_\ell} h_k + n_k, \quad \ell = k \bmod T.$$

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 are:

$$\begin{aligned} \hat{h}_{mT} &= \mathcal{E} [h_{mT} | y_{mT}] = \frac{\sqrt{E_0}}{\sigma_N^2 / \sigma_h^2 + E_0} y_{mT}, \\ \hat{h}_{mT+\ell} &= \alpha^\ell \hat{h}_{mT} \quad \text{for } 1 \leq \ell \leq T - 1. \end{aligned} \quad (2)$$

The marginal pdf of the estimator (2) is given by:

$$\hat{h}_{mT+\ell} \sim \mathcal{CN}(0, \hat{\sigma}_\ell^2), \quad \hat{\sigma}_\ell^2 = \alpha^{2\ell} \frac{E_0 \sigma_h^2}{\sigma_N^2 / \sigma_h^2 + E_0}, \quad (3)$$

for $0 \leq \ell \leq T - 1$. Note that the performance of the estimator becomes worse with increasing ℓ .

3. AVERAGE CUTOFF RATE

To characterize the rates that can be used reliably on the Gauss-Markov channel, we will use the *average cutoff rate* R_A , which is developed as follows: The estimation scheme separates the channel into T distinct slots per period. Consider the ℓ^{th} slot. Over time, and after estimation, successive realizations of the ℓ^{th} slot have the same marginal statistics (3), and form a channel which can be treated as iid under the assumption of “perfect” interleaving. The first channel is used only for training symbols, and the remaining channels are data channels. Then R_A is simply the average cutoff rate over the set of T channels created by the channel estimation algorithm. Given partial CSI at the receiver, via the channel estimate \hat{h} in (2), we assert that² the ergodic cutoff rate of the ℓ^{th} channel ($\ell \neq 0$) is given by:

$$R_{o,\ell} = -\log_2 \left\{ \frac{1}{2} + \frac{1}{2} \left[\frac{1 + \kappa_\ell (1 - \omega_\ell)}{1 + \kappa_\ell} \right] \right\},$$

where $\omega_\ell \triangleq \hat{\sigma}_\ell^2 / \sigma_h^2$ ($0 \leq \omega_\ell \leq 1$) is the estimator quality and $\kappa_\ell \triangleq \sigma_h^2 E_\ell / \sigma_N^2$ is the received SNR in the ℓ^{th} slot. Figure 1 shows the cutoff rate as a function of ω_ℓ and κ_ℓ . Note that, above a certain value of κ_ℓ , the cutoff rate saturates to a value that depends only on the estimator quality ω_ℓ . The average cutoff rate R_A is

$$R_A \triangleq \frac{1}{T} \sum_{\ell=1}^{T-1} R_{o,\ell} = -\frac{1}{T} \sum_{\ell=1}^{T-1} \log_2 \left[1 - \frac{1}{2} \frac{\alpha^{2\ell} \kappa_\ell \kappa_0}{(1 + \kappa_\ell)(1 + \kappa_0)} \right]. \quad (4)$$

Codebook multiplexing is one scheme that achieves R_A . There are $T - 1$ codebooks. The ℓ^{th} codebook operates independently at rate $R_{o,\ell}$ and transmits an N -length codeword over a period of NT slots. Under the assumption of perfect interleaving, R_A is a lower bound on reliable rates for the Gauss-Markov channel for the following reasons:

²We defer the proof of this assertion, and of several other results, to [11].

- LB1. The PSAM front end is a design constraint and may be rate-suboptimal. The MMSE-1 estimation algorithm is clearly suboptimal, as it uses only the current training symbol.
- LB2. Treating the Gauss-Markov channel as T independently coded sub-channels does not allow for joint coding across sub-channels.

Increasing the energy in the pilot improves channel estimation accuracy and increases R_A ; decreasing the energy in the data symbols decreases R_A . Thus, subject to a total energy constraint, there is an energy allocation problem which we study next.

4. OPTIMAL ENERGY ALLOCATION: EQUAL ENERGY DATA SLOTS

In this section we assume that practical constraints require that all data slots use the same energy κ_D , so that $\kappa_\ell = \kappa_D$ for $\ell \in \{1, \dots, T - 1\}$.³

4.1. Optimal Energy

Here, we will find the data energy κ_D^* and training energy κ_0^* which maximize R_A subject to a constraint on the total energy,

$$\kappa_0 + (T - 1)\kappa_D \leq \kappa_{\text{tot}}. \quad (5)$$

The average cutoff rate is given by (4) with $\kappa_\ell = \kappa_D, \forall \ell \geq 1$.

We can show that the optimal solution satisfies the equality constraint in (5), and is given by

$$\kappa_D^* = \Gamma - \sqrt{\Gamma^2 - \frac{\kappa_{\text{tot}}}{T-1}\Gamma}, \quad \Gamma = \frac{\kappa_{\text{tot}} + 1}{T-2}, \quad (6)$$

for $T > 2$. For $T = 2$, $\kappa_D^* = \kappa_0^* = \kappa_{\text{tot}}/2$.

Note that κ_D^* does not depend on α , however $R_A(\kappa_D^*)$, does. Next, we look at two asymptotic results:

- Low energy regime ($\kappa_{\text{tot}} \ll 1$), we obtain

$$\kappa_D^* = \frac{1}{2} \frac{\kappa_{\text{tot}}}{T-1}, \quad \kappa_0^* = \frac{\kappa_{\text{tot}}}{2},$$

implying that half of the available energy per period should be allocated to the training symbol. The 50-percent training paradigm has also been reported in [12] for a different channel model, metric, and assumptions.

- High energy regime ($\kappa_{\text{tot}} \gg 1$), we have

$$\begin{aligned} \kappa_D^* &= \frac{\kappa_{\text{tot}}}{T-2} \left(1 - \frac{1}{\sqrt{T-1}} \right), \quad \text{and} \\ \kappa_0^* &= \kappa_{\text{tot}} \left[1 - \frac{T-1}{T-2} \left(1 - \frac{1}{\sqrt{T-1}} \right) \right]. \end{aligned} \quad (7)$$

A similar result was found in [12]. For large T , the energy allocated to the training symbol decays as $T^{-1/2}$,

$$\kappa_0^* = \frac{\kappa_{\text{tot}}}{\sqrt{T}}.$$

³Henceforth, we will describe energy quantities by their faded values, denoted by $\kappa = E\sigma_h^2/\sigma_N^2$. The system designer has control over the corresponding unfaded energy E . The conversion between these two quantities is trivial, as the transmitter has knowledge of the noise ratio, σ_N^2/σ_h^2 .

4.2. Optimal Training Period

The training period T is the remaining design parameter. Here, we find the optimal training period T^* for maximization of R_A . For a fair comparison, we generalize the constraint of (5) so that the average energy per slot κ_{av} remains fixed as we vary T :

$$\kappa_0 + (T-1)\kappa_D = \kappa_{\text{tot}} = T\kappa_{\text{av}} \quad (8)$$

Let $R_A^*(T)$ be the optimized average cutoff rate for fixed T . Using (8) in (4), with $\kappa_\ell = \kappa_D^*$ given by (6), and $T\kappa_{\text{av}}$ replacing κ_{tot} we obtain

$$R_A^*(T) = -\frac{1}{T} \sum_{\ell=1}^{T-1} \log_2 \left\{ 1 - \frac{\alpha^{2\ell}}{2} \frac{(T\kappa_{\text{av}} - (T-1)\kappa_D^*)\kappa_D^*}{(1 + T\kappa_{\text{av}} - (T-1)\kappa_D^*)(1 + \kappa_D^*)} \right\} \quad (9)$$

which depends on T , α , and κ_{av} . Increasing T decreases the overhead (rate loss) due to training, but also decreases the channel estimation accuracy, and thus the achievable rate in the latter slots. We seek the T that optimizes $R_A^*(T)$.

In the high energy regime ($\kappa_{\text{av}} \rightarrow \infty$), each term in the summation of (9) enters the saturation region, and we have

$$R_A = -\frac{1}{T} \sum_{\ell=1}^{T-1} \log_2 \left[1 - \frac{\alpha^{2\ell}}{2} \right].$$

Denoting the optimum training period in the high energy regime as T_H^* , we have that

$$T_H^* = \arg \max_T \prod_{\ell=1}^{T-1} \left[1 - \frac{\alpha^{2\ell}}{2} \right]^{-1/T}. \quad (10)$$

which is given explicitly in terms of α in Table 1. Notice that T_H^* is an increasing function of α . This is easily verified from (10) for the high energy case, and can be verified, in general, from (9). This is because larger α implies that estimates of successive channel slots degrade slowly, and therefore more channels can be utilized before we reach the point at which the loss in $R_A^*(T)$ incurred by adding another channel is no longer offset by the potential gain due to less training overhead.

T_H^* is a lower bound on T^* for any κ_{av} (i.e., $T_H^* \leq T^*$). This is because, in the saturation region, the quality loss between successive data slots is magnified (see Figure 1). Therefore, making T larger diminishes $R_A^*(T)$ quickly. In Table 2, we compare the optimum training period in the high energy regime T_H^* , with T^* for several values of κ_{av} . Note also that for $\alpha \leq 0.98$ and for $\kappa_{\text{av}} \geq 10$, the bound is nearly exact. Note that T^* indicates the optimal training period under an average energy constraint; it clearly depends upon the normalized Doppler bandwidth, α .

5. OPTIMAL ENERGY ALLOCATION: VARIABLE ENERGY DATA SLOTS

In this section we seek to optimize R_A in (4) over $\boldsymbol{\kappa} \triangleq \{\kappa_0, \dots, \kappa_{T-1}\}$, subject to the total energy constraint

$$\sum_{\ell=0}^{T-1} \kappa_\ell = \kappa_{\text{tot}}.$$

We will treat the training period T as a fixed parameter.

α	T_H^*
$0 < \alpha < 0.73$	2
$0.74 \leq \alpha < 0.89$	3
$0.90 \leq \alpha < 0.94$	4
$0.95 \leq \alpha < 0.96$	5
$\alpha = 0.99$	9
$\alpha = 0.999$	27
$\alpha = 0.9999$	84

Table 1. The optimal training period at high energy T_H^*

α	T_H^*	T^*			
		$\kappa_{\text{av}} = 1000$	100	10	1
0.50	2	2	2	2	2
0.80	3	3	3	3	3
0.90	4	4	4	4	5
0.95	5	5	5	5	8
0.98	7	7	7	8	13
0.99	9	9	9	11	20
0.999	27	27	28	37	88
0.9999	84	85	90	136	387

Table 2. Comparing T_H^* to T^* for different values of κ_{av}

Our intuition from water-filling over parallel AWGN channels applies here as we are ignoring the correlation between the T channels with respect to coding, and are instead exploiting this correlation in designing the channel estimator. Waterfilling predicts that more energy will be allocated to less noisy channels, and that channels with noise-levels above a threshold will not be used at all. We will see that both of these ideas are preserved, where the noise metric is σ_N^2/σ_h^2 .

5.1. Substitution Function

Optimization of R_A over $\boldsymbol{\kappa}$ does not lead to a closed form solution for the optimal energy allocation. Instead, we propose an approximate solution based on optimizing the substitution function

$$\tilde{R}_A \triangleq \sum_{\ell=1}^{T-1} \alpha^{2\ell} \frac{\kappa_\ell \kappa_0}{(1 + \kappa_0)(1 + \kappa_\ell)}.$$

over $\boldsymbol{\kappa}$. We will denote the optimizer of the substitution function by $\tilde{\boldsymbol{\kappa}}^*$. Let $\boldsymbol{\kappa}^*$ be the optimal energy-vector for R_A in (4). We claim that $\tilde{\boldsymbol{\kappa}}^* \approx \boldsymbol{\kappa}^*$ for the following reasons (proofs and further details are given in [11]):

- A1. The approximate solution is exact (i.e., $\tilde{\boldsymbol{\kappa}}^* = \boldsymbol{\kappa}^*$) as $\alpha \rightarrow 1$ or as $\alpha \rightarrow 0$ or as $\kappa_{\text{tot}} \rightarrow 0$.
- A2. The appropriate Taylor expansion shows that $\tilde{\boldsymbol{\kappa}}^* \approx \boldsymbol{\kappa}^*$ if $\alpha^{2\ell} \ll 1$ or if $\kappa_0 \ll 1$ or if $\kappa_\ell \ll 1, \forall \ell \geq 1$.
- A3. Numerical simulations show that $\tilde{\boldsymbol{\kappa}}^* \approx \boldsymbol{\kappa}^*$ for moderate values of α , at moderate to high values of κ_{tot} (this is the region where no theoretical justification has been given).

Illustrative examples of the above remarks are given in section 5.3.

5.2. Optimal Energy Allocation

The optimal energy vector $\tilde{\boldsymbol{\kappa}}^*$ is specified by the following:

Theorem. (a) Use the first M data slots ($T_A = M$) iff

$$\phi_\alpha(M-1) \leq \kappa_{\text{tot}} < \frac{\phi_\alpha(M)}{1 - \delta(M-T+1)}, \quad (11)$$

where $\delta(x)$ is the Kronecker delta, $1 \leq M \leq T-1$, and

$$\phi_\alpha(N) = \left[\frac{\alpha^{-N} - 1}{1 - \alpha} - \left(N + \frac{1}{2} \right) + \sqrt{\frac{1}{4} + \frac{(\alpha^{-N} - \alpha)(\alpha^{-N} - 1)}{1 - \alpha^2}} \right]. \quad (12)$$

(b) The optimal training energy ($T_A \neq 1$) is given by:

$$\begin{aligned} \kappa_0(T_A) &= -\Delta(T_A + \kappa_{\text{tot}}) \\ &+ \sqrt{(\Delta^2 + \Delta)(T_A + \kappa_{\text{tot}})^2 - (\Delta + 1)(T_A + \kappa_{\text{tot}})}, \end{aligned}$$

where $\Delta = \frac{1}{2} \frac{(1-\alpha)(1+\alpha^{T_A})}{\alpha - \alpha^{T_A}}$.

(c) The data energies ($T_A \neq 1$) are given by, $1 \leq \ell \leq T_A$,

$$\kappa_\ell = \alpha^{\ell-1} \frac{1-\alpha}{1-\alpha^{T_A}} [\kappa_{\text{tot}} - \kappa_0(T_A) + T_A] - 1.$$

(d) If $T_A = 1$, $\kappa_0 = \kappa_1 = \kappa_{\text{tot}}/2$.

Proof. See [11]. \square

The channel assignment strategy of (11) is illustrated in Figure 2 for a system with $\kappa_{\text{tot}} = 50$, $T = 21$, and for several values of α . Consider the curve $\phi_\alpha(M)$. The candidate energy line intersects $\phi_\alpha(M)$ between $M = 10$ and $M = 11$. Therefore, $T_A = 11$ is the optimum number of data paths to activate.

We look at some consequences of the Theorem:

1. T_A is an increasing function of α (see Fig 2). This can be verified by noting that

$$\frac{\partial \phi_\alpha(M)}{\partial \alpha} < 0, \quad \text{for } 0 < \alpha < 1.$$

2. As $\kappa_{\text{tot}} \rightarrow \infty$ all $T-1$ slots become active and

$$\kappa_0(T-1) = \kappa_{\text{tot}} \left[\frac{\sqrt{h(\alpha)} - 1}{h(\alpha) - 1} \right], \quad (13)$$

where $h(\alpha) = \frac{(1+\alpha)(1-\alpha^{T-1})}{(1-\alpha)(1+\alpha^{T-1})}$.

3. As $\kappa_{\text{tot}} \rightarrow 0$ only the first data slot is active and

$$\kappa_0(1) = \frac{E_{\text{tot}}}{2}. \quad (14)$$

5.3. Numerical Simulations

In this section, we show that $\tilde{\boldsymbol{\kappa}}^* \approx \boldsymbol{\kappa}^*$ using numerical techniques. Define the normalized error metric⁴

$$\mathbf{e} \triangleq \frac{\|\boldsymbol{\kappa} - \tilde{\boldsymbol{\kappa}}\|_1}{\kappa_{\text{tot}}}.$$

Figure 3 compares $\tilde{\boldsymbol{\kappa}}^*$ with $\boldsymbol{\kappa}^*$ for $\kappa_{\text{tot}} = 10$ and $T = 6$, and for several values of α . The remarks A1 and A2 predict that the approximate solution performs well for $\alpha = 0.2$ and $\alpha = 0.98$. This is verified in the figure, both graphically and from the \mathbf{e} metric. We observe that the solution is also close for $\alpha = 0.5$ and $\alpha = 0.7$. Note that in all cases the approximate solution correctly predicts the number of active paths T_A .

In Figure 4, we compare $\tilde{\boldsymbol{\kappa}}^*$ and $\boldsymbol{\kappa}^*$ for $\alpha = 0.85$, $T = 6$, and for different values of κ_{tot} . The remarks A1 and A2 predict accuracy for $\kappa_{\text{tot}} = 0.1$. We see that the normalized error \mathbf{e} remains small for the higher values of κ_{tot} as well. Again, the approximate solution correctly predicts T_A in each case.

6. DISCUSSION

We considered the Gauss-Markov fading channel with partial CSI available at the receiver via periodic training. Using average cut-off rate as a metric, we derived a closed-form expression for the optimal allocation of energy between the training and data slots for both the fixed and variable data energy cases. For the fixed data-energy case, we also found a bound on the optimal training period, that is exact at moderate to high SNR. For the variable data-energy case, a well-justified substitution function was used, leading to closed-form expressions for the optimal number of active data slots, and the energy allocation to each slot. We are currently studying optimal binary signaling in place of BPSK, more sophisticated estimation strategies, and MIMO/OFDM extensions.

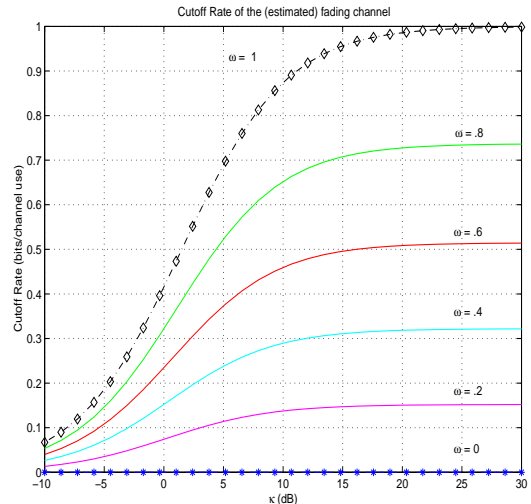


Fig. 1. Cutoff rate R_o vs. received SNR κ (dB) for several values of estimation quality factor ω .

⁴Here $\|\mathbf{a}\|_1$ denotes the 1-norm of the vector \mathbf{a} .

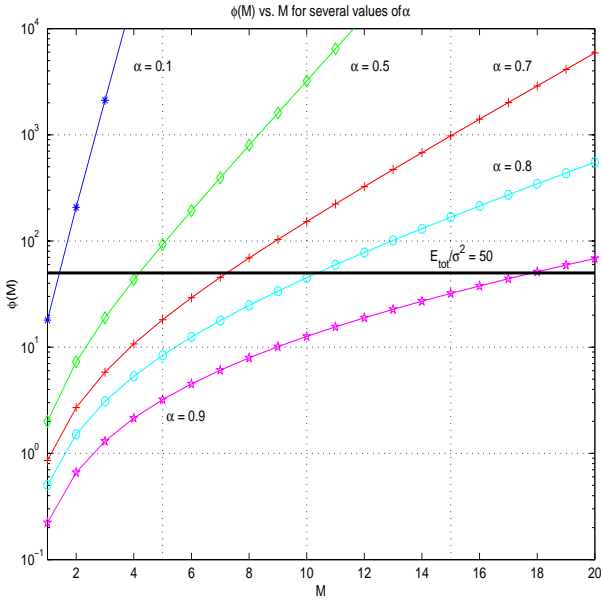


Fig. 2. An illustration of the test for determining T_A .

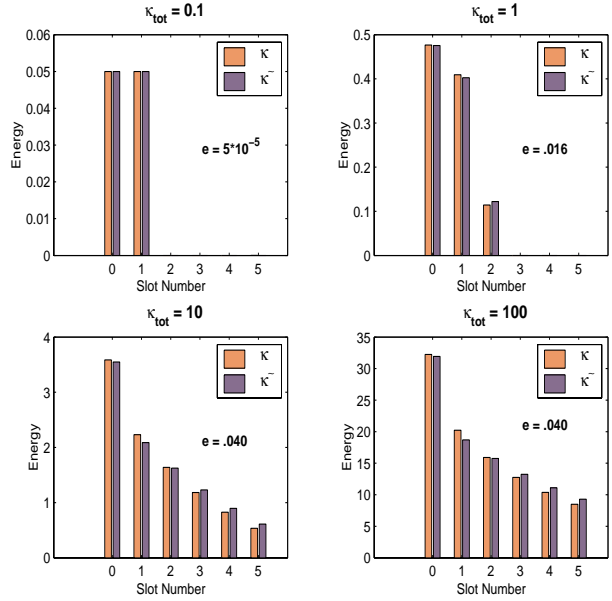


Fig. 4. Comparison of κ^* with $\tilde{\kappa}^*$ for $\alpha = 0.85$ and different values of κ_{tot} .

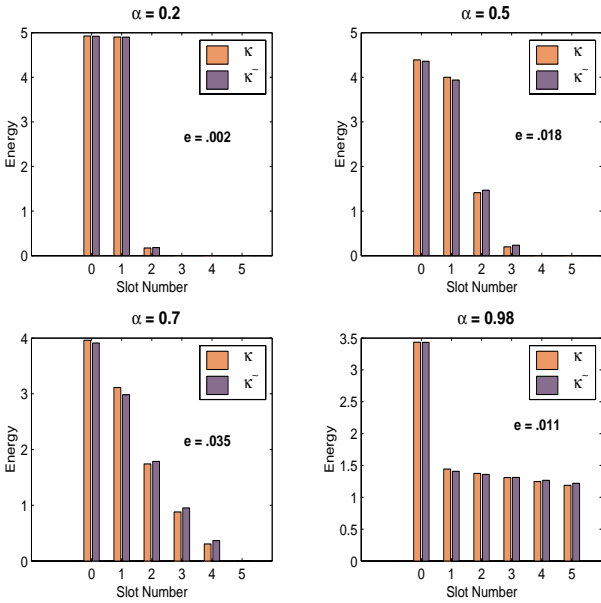


Fig. 3. Comparison of κ^* with $\tilde{\kappa}^*$ for $\kappa_{\text{tot}} = 10$ and different values of α . The x-axis indicates each of the T slots. The y-axis shows the energy placed in each of the slots.

7. 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] R.-R. Chen, B. Hajek, R. Koetter, and U. Madhow "On fixed input distributions for noncoherent communication over high SNR Rayleigh fading channels," Available at <http://tesla.csl.uiuc.edu/~koetter/publications.html>, June 2002.
- [5] M. Dong, L. Tong, and B. Sadler, "Training placement for tracking fading channels," *Proc. IEEE ICASSP*, May 2002.
- [6] B. Hassibi and B. Hochwald, "How Much Training is Needed in Multiple-Antenna Wireless Links?," Available at <http://mars.bell-labs.com/cm/ms/what/mars/papers/training/>, April 2000.
- [7] 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.
- [8] S. Jamali, and T. Le-Ngoc. *Coded-Modulation Techniques for Fading Channels*. Kluwer Academic Publishers, April 1994.
- [9] J. Massey, "Coding and Moulation in Digital Communications," In *Proc. 1974 Int. Zurich Seminar, Digital Communication*, March 1974, pp.E2.1-E2.4.
- [10] 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.
- [11] S. Misra, A. Swami, and L. Tong, "Cutoff Rate Analysis of the Gauss-Markov Fading Channel with Binary Inputs and Partial CSI at the Receiver," *Proc. 37th Annual Conference on Information Sciences and Systems*, March 2003.
- [12] H. Vikalo, B. Hassibi, B. Hochwald, and T. Kailath, "Optimal Training For Frequency-Selective Fading Channels," *IEEE ICASSP*, pp.2105-2108 vol. 4, 2001.