Finite-State Markov Channel—A Useful Model for Radio Communication Channels

Hong Shen Wang, *Member, IEEE*, and Nader Moayeri, *Senior Member, IEEE*

***Abstract*—In this paper, we first study the behavior of a finite-state channel where a binary symmetric channel is associated with each state and Markov transitions between states are assumed. Such a channel is referred to as a finite-state Markov channel (FSMC). By partitioning the range of the received singal-to-noise ratio into a finite number of intervals, FSMC models can be constructed for Rayleigh fading channels. Theoretical approach is conducted to show the usefulness of FSMC's compared to that of two-state Gilbert–Elliott channels. The crossover probabilities of the binary symmetric channels associated with its states are calculated. We use the second-order statistics of the received SNR to approximate the Markov transition probabilities. The validity and accuracy of the model are confirmed by the state equilibrium equations and computer simulation.**

I. Introduction

THE STUDY OF the finite-state Markov channel (FSMC) emerges from early work of Gilbert [1] and Elliott [2]. They study a two-state Markov channel known as the Gilbert–Elliott channel. In their channel model, each state corresponds to a specific channel quality which is either noiseless or totally noisy. In general, a binary symmetric channel (BSC) with a given crossover probability can be associated with each state so that the channel quality for each state can be identified. The Gilbert–Elliott channel is the special case where the crossover probabilities of the BSC's are 0 and 0.5, respectively. In some cases, modeling a radio communication channel as a two-state Gilbert–Elliott channel is not adequate when the channel quality varies dramatically. A straightforward solution is to form a channel model with more than two states. This idea motivates our study of the extension from a two-state channel to a finite-state one.

Let $\mathcal{S} = \{s_0, s_1, \cdots, s_{K-1}\}$ denote a finite set of states and $\{S_n\}, n = 0, 1, 2, \cdots$, be a constant Markov process. Since the constant Markov process has the property of stationary transitions, the transition probability is independent of the time index n and can be written as

$$t_{j,k} = \Pr(S_{n+1} = s_k | S_n = s_j) \quad (1)$$

Manuscript received August 2, 1993; revised November 17, 1993; accepted June 27, 1994. This work was supported in part by the Wireless Information Network Laboratory, Rutgers University. This work was presented in part at the 1993 International Symposium on Communications, Hsinchu, Taiwan, ROC, December, 1993.

H. S. Wang is with the Department of Information Engineering, the Chinese University of Hong Kong, Hong Kong.

N. Moayeri is with the Hewlett-Packard Laboratories, Palo Alto, CA USA.

IEEE Log Number 9407343.

for all $n = 0, 1, 2, \ldots$ and $k, j \in \{0, 1, 2, \cdots, K-1\}$. With this definition, we can define a $K \times K$ state transition probability matrix $\mathbf{T}$ with its elements $t_{j,k}$ as in (1). Note that a state transition probability matrix has the property that the sum of the elements on each row is equal to 1, or

$$\sum_{l=0}^{K-1} t_{k,l} = 1, \quad \forall k \in \{0, 1, 2, \cdots, K-1\}. \quad (2)$$

Moreover, with the stationary transition property, the probability of state k at any permissible time index n without any state information at other time indices can also be defined as

$$p_k = \Pr(S_n = s_k), \quad k \in \{0, 1, 2, \cdots, K-1\}. \quad (3)$$

A $K \times 1$ steady state probability vector $\mathbf{p}$ can be defined with its elements p_k as in (3). In many cases, this vector can be served as the set of initial state probabilities. Note that (1) and (3) must satisfy the equilibrium condition which states that for any given state k, the incoming flow and outgoing flow must be equal. That is, for all $k \in \{0, 1, 2, \cdots, K-1\}$,

$$\sum_{j=0}^{K-1} p_j t_{j,k} = \sum_{l=0}^{K-1} p_k t_{k,l}.$$

With (2), we have

$$\sum_{j=0}^{K-1} p_j t_{j,k} = p_k, \quad \forall k \in \{0, 1, 2, \cdots, K-1\}$$

or simply

$$\mathbf{p}^t \mathbf{T} = \mathbf{p}^t \quad (4)$$

where $\mathbf{p}^t$ is the transpose of $\mathbf{p}$.

A complete description of a finite-state Markov channel requires additional information on the channel quality for each state. Define a $K \times 1$ crossover probability vector $\mathbf{e}$ with its elements $e_k, k \in \{0, 1, 2, \cdots, K-1\}$, being the crossover probability of the binary symmetric channel associated with state k. A FSMC is then uniquely defined by $\mathbf{T}, \mathbf{p}$ and $\mathbf{e}$. The overall average error probability e of the FSMC is then

$$e = \mathbf{p}^t \mathbf{e} = \sum_{k=0}^{K-1} p_k e_k. \quad (5)$$

The choices of $\mathbf{T}, \mathbf{p}$ and $\mathbf{e}$ may not be arbitrary. Actually, in addition to (2), (4), and (5), there are four obvious constraints imposed on $\mathbf{p}$ and $\mathbf{e}$ as follows:

1. $0 < p_k \leq 1, \ \forall k \in \{0, 1, 2, \cdots, K-1\}$.
2. $\sum_{k=0}^{K-1} p_k = 1$.
3. $0 \leq e_k \leq 0.5, \ \forall k \in \{0, 1, 2, \cdots, K-1\}$.
4. $e_i \neq e_j, \text{if } i \neq j \ \forall i, j \in \{0, 1, 2, \cdots, K-1\}$.

Most studies on the performance of the Gilbert–Elliott channel were based on variations of $\mathbf{T}, \mathbf{p}$, and $\mathbf{e}$ under the constraints mentioned above. For example, in [3], the authors calculate the capacity of a Gilbert–Elliott channel with $\mathbf{p}$ and $\mathbf{e}$ fixed while varying $\mathbf{T}$. It is easy to see that with $\mathbf{p}$ fixed, the ratio between $t_{0,1}$ and $t_{1,0}$ is also fixed. Under these conditions, they define a measure of the channel memory as

$$\mu = 1 - t_{0,1} - t_{1,0} \tag{6}$$

and calculate the capacity as a monotonically increasing function of μ in a reasonable range, i.e., $\mu \in [0, 1)$.

At this point, we shall emphasize that it may seem to be a fair comparison if one fixes the average error probability e and varies other channel parameters under the constraints mentioned above. The conclusion that can be drawn from this kind of analysis is that with the average error probability and the number of states fixed, the *modeled* finite-state channel with more memory content may have greater capacity. As in [3], greater capacity is observed when the value of μ in (6) increases. It is then natural to ask how can the performance be optimized with respect to the channel parameters subject to the four constraints given above. An attempt to solve this question motivates our study on establishing a connection between Rayleigh fading channels and their FSMC models such that the transmission technologies thereby designed can be applied to the channel efficiently. As what will be presented later, a partitioning of the received signal to noise ratio (SNR) into a finite number of intervals leads to an FSMC model. With the partitioning, the elements in $\mathbf{T}, \mathbf{p}$, and $\mathbf{e}$ can be obtained and the corresponding FSMC model is established. Similar ideas have been proposed in [4] where soft dicision statistical distributions obtained from the output of a slowly varying Rayleigh fading channel are compared between the simulated channels and the models. In this paper, analytical results approximate the transition matrix $\mathbf{T}$ are derived and compared with computer simulations to show the accuracy of the model. Furthermore, state equilibrium condition has been shown that verifies the validity of the FSMC model. The contribution of analytical approximation using level crossing rate is important for real-time applications where time consuming simulation is not possible. The capacity evaluation and the performance of an efficient coding technique based on this channel model can be found in [5].

In Section II, we define the capacity and coding distortion as criterions of optimization. In Section III, we confirm the assertions mentioned above by attempting to find the FSMC that results in the maximal channel capacity or the minimal coding distortion. The issues of modeling, verification, and comparison between analytical and simulation model are presented in Section IV followed by conclusions in Section V.

II. Optimization Criterions

The capacity of the finite-state Markov channel can be treated from two viewpoints. When the channel state information (CSI) is available, the capacity C^{CSI} is simply the average capacity over all the states [6], or

$$C^{CSI} = \sum_{k=0}^{K-1} p_k[1 - h(e_k)] \tag{7}$$

where $h(\cdot)$ is the binary entropy function defined as

$$h(e) = e \log \frac{1}{e} + (1-e) \log \frac{1}{1-e}. \tag{8}$$

On the other hand, it is quite difficult to calculate the capacity of the finite-state Markov channel when CSI is not available. In [1], Gilbert calculates the capacity of the Gilbert–Elliott channel, which is a two-state Markov channel with one good and one bad state. Results are shown for the crossover probability of these two states equal to 0 and 0.5, respectively. In [3], Mushkin and Bar-David present a method for calculating the capacity of the Gilbert–Elliott channel with all possible crossover probabilities. The generalization of this method to the case of a K-state Markov channel is presented in [5] while no closed-form solution is possible. Therefore, we use the capacity with channel state information C^{CSI} as our first criterion.

Farvardin and Vaishampayan [7] studied a joint source and channel coding system where the output of a scalar quantizer is transmitted over a binary symmetric channel. In such a system, the encoder is specified by its quantization regions and the corresponding channel input letters. The decoder simply maps a received channel output letter to a reproduction value. An iterative algorithm is developed using two necessary optimality conditions derived by Kurtenbach and Wintz [8]. The resulting noisy channel quantizer is at least locally optimal in the sense of minimizing the overall distortion for a given rate. For the finite-state Markov channel, the necessary optimality conditions are derived in [5] with the assumption that the receiver has the knowledge of the current channel state information. Let the quantizer rate be R_j bits per sample. Suppose we have a K-state Markov channel and the encoder maps the source input x to an index $m(i), i \in \{1, 2, \cdots, 2^{R_j}\}$, which is transmitted through the channel. Upon receiving the index $m(j)$ (which may not be the same as $m(i)$ due to channel error) and the current channel state information $k \in \{0, 1, 2, \cdots, K-1\}$, the decoder releases $y_{j,k}$ as the reproduction of x. The encoder and decoder mappings, Q_e and Q_d, can then be written as

$$Q_e : \mathbf{R} \to \mathcal{I},$$

and

$$Q_d : \mathcal{I} \times \mathcal{K} \to \mathcal{C}$$

where $\mathcal{K} = \{0, 1, \cdots, K-1\}$ and $\mathcal{I} = \{0,1\}^{R_j}$ is the set of all binary words of length R_j.

Let $p_k, k = 0, 1, \cdots, K-1$ be the steady state probability that the K-state channel is in state k and $p(m(j) \mid m(i), k)$ be the conditional probability that index $m(j)$ is received given that index $m(i)$ is transmitted and the channel is in state k. It was shown in [5] that when the channel state information (CSI) is available to the receiver, the mean squared error distortion is given by

$$D = \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{k=0}^{K-1} p(m(j) \mid m(i), k) p_k \times \int_{V(i)} p_X(x)(x - y_{j,k})^2 dx \tag{9}$$

where $N = 2^{R_j}$ and $V(i)$ is the quantization region for the index $m(i)$ defined as

$$V(i) = \{x \, : \, x \in \mathbf{R}, \, Q_e(x) = m(i)\}, \quad m(i) \in \mathcal{I}.$$

It can further be shown that the necessary optimality conditions, also known as the *nearest neighbor condition* and *centroid condition*, are given by

$$V(i) = \left\{ x : \sum_{j=1}^{N} \sum_{k=0}^{K-1} p(m(j) \mid m(i), k) p_k (x - y_{j,k})^2 \leq \sum_{j=1}^{N} \sum_{k=0}^{K-1} p(m(j) \mid m(l), k) p_k (x - y_{j,k})^2; \; \forall l \neq i \right\} \tag{10}$$

and

$$y_{j,k} = \frac{\sum_{i=1}^{N} p(m(j) \mid m(i), k) p_k \int_{V(i)} x p_X(x) dx}{\sum_{i=1}^{N} p(m(j) \mid m(i), k) p_k \int_{V(i)} p_X(x) dx}. \tag{11}$$

Note that these two equations formulate the necessary conditions for an optimal quantizer when the channel state information is available to the decoder only. They are similar to the *Nearest Neighbor Condition* and *Centroid Condition* [9]–[11]. In the absence of the channel state information, the channel is simply equivalent to a binary symmetric channel. In this case, previous results and methods [8], [7] are applicable. On the other hand, if the channel state information is available to both the encoder and decoder, then the optimal quantizer for a K-state channel is equivalent to that for K binary symmetric channels [12]. The detailed proofs for (9)–(11) can be found in [5]. We present the results here and conclude that a quantizer is said to be optimal if both the nearest neighbor condition and centroid condition are satisfied. Consequently, an optimal quantizer for the finite-state Markov channel satisfies (10) and (11) with CSI available to the decoder.

III. Optimizing The Finite-State Markov Channel

As mentioned in Section I, a finite-state Markov channel (FSMC) can be characterized by its state transition matrix $\mathbf{T}$, steady state probability vector $\mathbf{p}$, and crossover probability vector $\mathbf{e}$. Given a fixed average error probability e as in (5), for an FSMC, an effort is made to find the model that optimizes the system performance in the sense of maximizing the channel capacity and minimizing the coding distortion. Since the channel capacity without channel state information (CSI) does not have a closed-form solution, here we consider the capacity with CSI as in (7). On the other hand, the distortion measure, as shown in (9), is obtained by an optimal quantizer over the FSMC assuming that CSI is available to decoder. The results of these optimization procedures motivate the modeling issue in Section IV.

When the average error probability e for a finite-state Markov channel is given, one may want to optimize the quantization system by varying the steady state probability vector $\mathbf{p}$, the crossover probability vector $\mathbf{e}$, and, as a consequence, the state transition matrix $\mathbf{T}$. Even the number of states, K, may be increased to improve the performance. According to the four constraints given in Section I, an arbitrary choice of $\mathbf{p}, \mathbf{e}$ and $\mathbf{T}$ may not correspond to a valid FSMC model. In this section, we try to show that even with valid $\mathbf{p}$ and $\mathbf{e}$, the effort of optimization, in the sense of maximizing the channel capacity or minimizing the coding distortion, is conclusively ineffectual. The capacity issue will be treated first.

A. Capacity Viewpoint

In (7), Section II, we presented the channel capacity for an FSMC when the channel state information is available. The following lemma says that with the average error probability unchanged, each state can be split into two substates in such a way that the capacity contributed by the two substates can be made greater than or equal to the capacity contributed by the original state. More importantly, the capacity contributed by the substates reaches a maximum value when the two substates correspond to a noiseless binary symmetric channel (crossover probability equal to 0) and an entirely noisy channel (crossover probability equal to 0.5).

Lemma 1: Suppose the average error probability e for a K-state Markov channel is fixed. Referring to (7), we can write $C_k = p_k[1 - h(e_k)]$ as the capacity contributed by state k when the channel state information is available. Then by splitting state k into substates $k0$ and $k1$ with

$$e_{k0} = 0.5, \quad p_{k0} = 2e_k p_k, \quad e_{k1} = 0, \quad \textit{and} \quad p_{k1} = p_k(1 - 2e_k)$$

the capacity contributed by these two substates is maximized. The gain in capacity due to splitting is

$$p_k[h(e_k) - 2e_k].$$

Proof: In order to partition the state k into two substates, say, state $k0$ and state $k1$, the following equations must be true to guarantee that the average error probability in (5) remains the same.

$$p_k = p_{k0} + p_{k1} \quad \text{and} \quad e_k p_k = e_{k0} p_{k0} + e_{k1} p_{k1}. \tag{12}$$

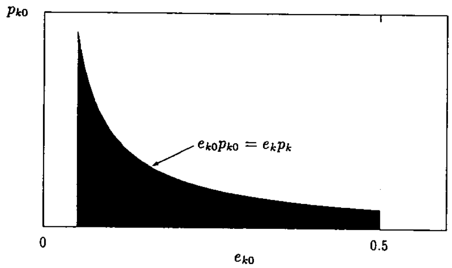

Fig. 1. Optimization region for p_{k0} and e_{k0} in the proofs of Lemma 1 and Lemma 2.

Also, the four constraints in (5) are applicable to e_{k0}, e_{k1}, p_{k0}, and p_{k1} as well. The contribution to the capacity of the substates $k0$ and $k1$ can then be written as

$$\begin{aligned} C_{ks} &= p_{k0}[1 - h(e_{k0})] + p_{k1}[1 - h(e_{k1})] \\ &= p_k - p_{k0}h(e_{k0}) - p_{k1}h(e_{k1}) \end{aligned} \quad (13)$$

where $h(\cdot)$ is the binary entropy function defined in (8). Therefore, maximizing C_{ks} is equivalent to maximizing C'_{ks} defined as

$$\begin{aligned} C'_{ks} &= p_{k0}(e_{k0}\log e_{k0} + (1-e_{k0})\log(1-e_{k0})) \\ &\quad + p_{k1}(e_{k1}\log e_{k1} + (1-e_{k1})\log(1-e_{k1})). \end{aligned} \quad (14)$$

From (12), we have

$$p_{k1} = p_k - p_{k0} \text{ and } e_{k1} = \frac{e_kp_k - e_{k0}p_{k0}}{p_k - p_{k0}}. \quad (15)$$

Substituting (15) into (14), we have

$$\begin{aligned} C'_{ks} &= p_{k0}(e_{k0}\log e_{k0} + (1-e_{k0})\log(1-e_{k0})) \\ &\quad + (p_k - p_{k0})\left\{(\frac{e_kp_k - e_{k0}p_{k0}}{p_k - p_{k0}})\log(\frac{e_kp_k - e_{k0}p_{k0}}{p_k - p_{k0}}) \right. \\ &\quad \left. + (1 - \frac{e_kp_k - e_{k0}p_{k0}}{p_k - p_{k0}})\log(1 - \frac{e_kp_k - e_{k0}p_{k0}}{p_k - p_{k0}})\right\}. \end{aligned}$$

Without any loss of generality, let $e_{k0} > e_k$. We are interested in finding the maximum value of C'_{ks} subject to the condition that e_{k0} and p_{k0} should take values in the shaded region shown in Fig. 1.

By taking the partial derivative of C'_{ks} with respect to e_{k0} and p_{k0}, we have

$$\frac{\partial C'_{ks}}{\partial e_{k0}} = \frac{p_{k0}}{\ln 2}\ln\left\{\frac{e_{k0}(p_k - p_{k0} - e_kp_k + e_{k0}p_{k0})}{(1-e_{k0})(e_kp_k - e_{k0}p_{k0})}\right\}$$

and

$$\begin{aligned} \frac{\partial C'_{ks}}{\partial p_{k0}} &= \frac{1}{\ln 2}\left\{e_{k0}\ln\frac{e_{k0}(p_k - p_{k0})}{e_kp_k - e_{k0}p_{k0}}\right\} \\ &\quad + (1-e_{k0})\ln\frac{(1-e_{k0})(p_k - p_{k0})}{p_k - p_{k0} - e_kp_k + e_{k0}p_{k0}}. \end{aligned} \quad (16)$$

By setting $\frac{\partial C'_{ks}}{\partial p_{k0}}$ equal to zero, we have either $p_{k0} = 0$ or $e_{k0} = e_k$. Therefore, the maximum value of C'_{ks} can occur only at the boundaries of the shaded region in Fig. 1. Note that since the conditions shown in (12) must hold, the boundaries $p_{k0} = 0$ and $e_{k0} = e_k$ both imply $C_{ks} = C_k$. We then need to look at the boundaries of $e_{k0} = 0.5$ and $e_{k0}p_{k0} = e_kp_k$ only. Evaluating (16) at $e_{k0} = 0.5$, we have

$$\left.\frac{\partial C'_{ks}}{\partial p_{k0}}\right|_{e_{k0}=0.5} = \frac{1}{2\ln 2} \times \ln\left\{\frac{1}{4}(p_k - p_{k0})^2 \Big/ \left[(e_kp_k - \frac{1}{2}p_{k0})(p_k - \frac{1}{2}p_{k0} - e_kp_k)\right]\right\}.$$

It is easy to see that the difference between the numerator and the denominator is

$$(\frac{1}{2} - e_k)^2p_k^2 > 0.$$

Therefore, on the boundary of $e_{k0} = 0.5$, C'_{ks} is monotonically increasing with p_{k0} and reaches its maximum at

$$p_{k0} = \frac{e_kp_k}{e_{k0}} = 2e_kp_k. \quad (17)$$

On the boundary of $e_{k0}p_{k0} = e_kp_k$, we can evaluate C'_{ks} with $p_{k0} = e_kp_k/e_{k0}$ yielding

$$C^*_{ks} = \frac{e_kp_k}{e_{k0}}(e_{k0}\log e_{k0} + (1-e_{k0})\log(1-e_{k0})).$$

The behavior of C^*_{ks} as a function of e_{k0} can be obtained by

$$\frac{\partial C^*_{ks}}{\partial e_{k0}} = -\frac{e_kp_k}{e_{k0}^2\ln 2}\ln(1-e_{k0}) > 0.$$

Therefore, C^*_{ks} is monotonically increasing with e_{k0} and reaches its maximum at the same point as (17). We conclude that by letting

$$\begin{aligned} e_{k0} &= 0.5 \\ p_{k0} &= e_kp_k/e_{k0} = 2e_kp_k \\ e_{k1} &= 0 \\ p_{k1} &= p_k - p_{k0} = p_k(1 - 2e_k) \end{aligned} \quad (18)$$

the capacity C_{ks} achieves its maximum by substituting (18) into (13) yielding

$$\tilde{C}_{ks} = p_k - p_{k0}h(e_{k0}) = p_k(1 - 2e_k). \quad (19)$$

Comparing (19) and $C_k = p_k(1 - h(e_k))$, it remains to show that $2e_k \le h(e_k)$ for $0 \le e_k \le 0.5$. Let us write

$$\Delta(x) = h(x) - 2x, \quad 0 \le x \le 0.5.$$

Then, since $\Delta(0) = \Delta(0.5) = 0$, and

$$\frac{\partial^2\Delta(x)}{\partial x^2} = -\frac{1}{\ln 2}\left(\frac{1}{x} + \frac{1}{1-x}\right) < 0$$

we have

$$\Delta(e_k) = h(e_k) - 2e_k \ge 0, \text{ for } 0 \le e_k \le 0.5.$$

Thus, the increase in capacity due to splitting state k into substates $k0$ and $k1$ is

$$\tilde{C}_{ks} - C_k = p_k(h(e_k) - 2e_k).$$

Since there is no closed form solution for the capacity without channel state information, a statement similar to the one in Lemma 1 is not possible.

B. Distortion Viewpoint:

In Section II, the necessary conditions for optimality of a quantizer over an FSMC is given. Due to the complicated structure of (10) and (11), we are unable to show a similar result as Lemma 1 for minimum distortion coding. Actually, the average distortion in (9) can be calculated only if the encoder and the decoder are well defined. Unfortunately, the optimal encoder and decoder can only be obtained through an iterative algorithm. An exception exists if we know exactly what the optimal encoder is. This happens when there are only two encoding regions separated by the mean μ_X of the source random variable X and when X is symmetrically distributed about μ_X. This, in general, implies a rate one quantizer with encoding regions

$$V(1) = (-\infty, \mu_X] \quad and \quad V(2) = (\mu_X, \infty). \tag{20}$$

Under these conditions, the following lemma is true.

Lemma 2: Suppose the average error probability e is fixed for a K-state Markov channel. Let the quantization rate be one bit per sample and the source be a symmetrically distributed random variable X. Referring to (9), we can write

$$D_k = p_k \sum_{i=1}^{2}\sum_{j=1}^{2} p(m(j)|m(i), k) \int_{V(i)} (x - y_{j,k})^2 p_X(x)dx \tag{21}$$

as the distortion contributed by state k when the channel state information is available to the decoder. Then by splitting state k into substates k0 and k1 with

$$e_{k0} = 0.5, \quad p_{k0} = 2e_k p_k, \quad e_{k1} = 0, \quad and \quad p_{k1} = p_k(1 - 2e_k)$$

the distortion contributed by these two substates is minimized. The decrease in distortion due to splitting is

$$8m^2 e_k p_k (1 - 2e_k)$$

where

$$m = \int_{\mu_X}^{\infty} x p_X(x)dx. \tag{22}$$

Proof: Given the quantization regions for the FSMC, the output reproductions $y_{1,k}$ and $y_{2,k}$ of a rate one bit per sample optimal quantizer can be found by substituting (20) into (11) with $N = 2$ and

$$p(m(j)|m(i), k) = \begin{cases} e_k, & \text{if } m(i) \neq m(j) \\ 1 - e_k, & \text{if } m(i) = m(j) \end{cases}. \tag{23}$$

For simplicity, we assume that X is zero-mean, i.e., $\mu_X = 0$. Then the optimal output reproductions are given by

$$\begin{aligned} y_{1,k} &= -2m(1 - 2e_k), \\ y_{2,k} &= -y_{1,k} = 2m(1 - 2e_k) \end{aligned} \tag{24}$$

where m is defined in (22). By substituting (23) and (24) into (21), we have

$$D_k = p_k[\sigma_x^2 - 4m^2(1 - 2e_k)^2] \tag{25}$$

where σ_x^2 is the variance of X. Let D_{ks} be the contribution to distortion due to splitting state k into substates $k0$ and $k1$. Then, similar to the proof of Lemma 1, we can apply (5), (12), (15) and the four constraints in (5) to write D_{ks} as a function of e_{k0} and p_{k0} as

$$\begin{aligned} D_{ks} = {} & p_k\sigma_x^2 - 4p_{k0}m^2(1 - 2e_{k0})^2 \\ & - 4(p_k - p_{k0})m^2 \left\{1 - 2\frac{e_k p_k - e_{k0}p_{k0}}{p_k - p_{k0}}\right\}^2. \end{aligned} \tag{26}$$

Without any loss of generality, let $e_{k0} > e_k$. We are interested in finding the minimum value of D_{ks} subject to the condition that e_{k0} and p_{k0} should take values in the shaded region shown in Fig. 1. By taking the partial derivative of D_{ks} with respect to e_{k0} and p_{k0}, we have

$$\frac{\partial D_{ks}}{\partial e_{k0}} = -32m^2 \frac{p_{k0}p_k(e_{k0} - e_k)}{p_k - p_{k0}} \tag{27}$$

and

$$\frac{\partial D_{ks}}{\partial p_{k0}} = -16m^2 \frac{p_k^2(e_{k0} - e_k)^2}{(p_k - p_{k0})^2}. \tag{28}$$

Setting (27) and (28) equal to zero, we have either $p_{k0} = 0$ or $e_{k0} = e_k$. Similar to the discussion in the proof of Lemma 1, the minimum value of D_{ks} can occur only at the boundaries of the shaded region in Fig. 1; actually only the boundaries of $e_{k0} = 0.5$ and $e_{k0}p_{k0} = e_k p_k$ need to be investigated. Evaluating (28) at $e_{k0} = 0.5$, we have

$$\left.\frac{\partial D_{ks}}{\partial p_{k0}}\right|_{e_{k0}=0.5} = -16m^2 \frac{p_k^2(\frac{1}{2} - e_k)^2}{(p_k - p_{k0})^2} < 0.$$

Therefore, D_{ks} is monotonically decreasing with p_{k0} on the boundary of $e_{k0} = 0.5$ and reaches its minimum at

$$p_{k0} = \frac{e_k p_k}{e_{k0}} = 2e_k p_k.$$

We now look at the behavior of D_{ks} on the boundary of $e_{k0}p_{k0} = e_k p_k$. Substituting $p_{k0} = e_k p_k / e_{k0}$ into (26), we have

$$\begin{aligned} D_{ks}^* = D_{ks}|_{p_{k0}=e_k p_k/e_{k0}} = {} & p_k[\sigma_x^2 - 4m^2(1 - 4e_k)] \\ & - 16m^2 e_k p_k e_{k0}. \end{aligned} \tag{29}$$

Since the first term in (29) is a constant, it is obvious that by letting $e_{k0} = 0.5$, D_{ks}^* is minimized. We then conclude that with e_{k0}, p_{k0}, e_{k1}, and p_{k1} assigned as in (18) the distortion D_{ks} is minimized. By substituting (18) into (26), we have the minimum value of D_{ks} given by

$$\tilde{D}_{ks} = p_k[\sigma_x^2 - 4m^2(1 - 2e_k)]. \tag{30}$$

It is then easy to see that the difference between (25) and (30) is

$$D_k - \tilde{D}_{ks} = 8m^2 e_k p_k (1 - 2e_k).$$

The quantitative results from Lemma 1 and Lemma 2 are depicted in Figs. 2 and 3 where the increase of capacity and the decrease of distortion due to state splitting are drawn.

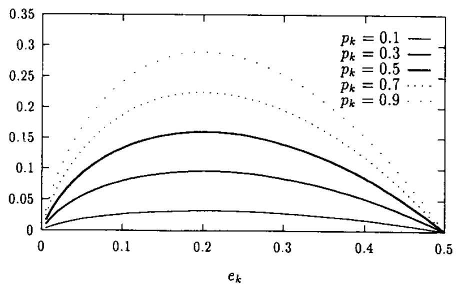

Fig. 2. The increase in capacity by splitting the state k into two substates as shown in Lemma 1.

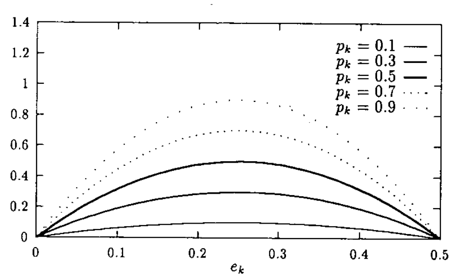

Fig. 3. The decrease in distortion by splitting the state k into two substates as shown in Lemma 2. (assuming $m = 1$)

With these lemmas, the following two theorems can be easily proved.

Theorem 1: Suppose the average error probability e *for a* K*-state Markov channel is fixed. Let* C *be the channel capacity defined as in (7). The result of maximizing* C *with respect to* $\mathbf{p}, \mathbf{e}$, *and the number of states* K *subject to (5) and the constraints in (5) is a two-state Markov channel with*

$$e_0 = 0.5, \quad p_0 = 2e$$

$$e_1 = 0, \quad p_1 = 1 - 2e. \tag{31}$$

Theorem 2: Suppose the average error probability e *for a* K*-state Markov channel is fixed. Let* D *be the average distortion defined as in (9). The result of minimizing* D *with respect to* $\mathbf{p}, \mathbf{e}$ *and the number of states* K *subject to (5) and the constraints in (5) is a two-state Markov channel with*

$$e_0 = 0.5, \quad p_0 = 2e$$

and

$$e_1 = 0, \quad p_1 = 1 - 2e.$$

Given Lemmas 1 and 2, the proof for Theorems 1 and 2 are stated as the following: Given a K-state Markov channel, one can optimize the channel model, in the sense of increasing the channel capacity or decreasing the coding distortion, by splitting each of those states satisfying $0 < e_k < 0.5$ into two substates with a noiseless one and an entirely noisy one. Then, the resulting finite-state Markov channel shall contain only two kinds of binary symmetric channel, where the crossover probability equals to 0 and 0.5, respectively. The resulting FSMC is eventually a two-state Markov channel with $\mathbf{p}$ and $\mathbf{e}$ satisfying (31). The theorems are then proved.

Note that in this section, the FSMC is optimized with respect to $\mathbf{p}$, $\mathbf{e}$, and K excluding the state transition matrix $\mathbf{T}$. This is due to the assumption that CSI is available. Consequently, with the instantaneous CSI available, the conditional probability of the next state given the current state, and therefore the transition matrix $\mathbf{T}$, is not important at all.

IV. The Rayleigh Fading Channel and Its Finite-State Model

From the theorems in Section III, we found that unreasonable results can be concluded if the finite-state model does not represent a real physical channel. Therefore, the methodology in establishing the relationship between physical channels and their finite-state models is important. In this section, we describe a typical radio communication channel, namely the Rayleigh fading channel, that produces time-varying received SNR characterizing the channel quality in terms of the average error probability. By partitioning the range of the received SNR into a finite number of intervals, a finite-state model for the Rayleigh fading channel is built. The physical characteristics and the time-varying behavior of a Rayleigh fading channel are introduced. Experimental results supporting the finite-state model can be found in [13], [14] but no prediction has been made for future states based on the present state, and possibly some of the previous ones. Obviously, the prediction is useful and acquirable due to the memory that exists in the physical channel. Here, we use a Markov chain to describe the transition activities between states from one observation time to the next. A finite-state Markov channel can then be used to model the Rayleigh fading channel. The transition probabilities are approximated and the state equilibrium equations are verified.

A. Rayleigh Fading Channels

The digital cellular radio transmission environment usually consists of a large number of scatterers that result in multiple propagation paths. Associated with each path is a propagation delay and an attenuation factor depending on the obstacles in the path that reflects the electromagnetic waves. When a continuous waveform (CW) is transmitted, the multipath effect results in the fluctuation of the received signal envelope that is Rayleigh distributed. This channel is known as a Rayleigh fading channel. The time variations of the signal level is characterized by the Doppler frequency effect, which is due to the motion of the mobile terminal.

Let A denote the received signal to noise ratio which is proportional to the square of the signal envelope. The p.d.f. of A is exponential [15], [16] and can be written as

$$p_A(a) = \frac{1}{\rho} \exp\left\{-\frac{a}{\rho}\right\} \tag{32}$$

for $a \geq 0$, where

$$\rho = E[A].$$

Following the derivation in [17], let f_m be the maximum Doppler frequency defined as

$$f_m = \frac{v}{\lambda}$$

where v is the speed of the vehicle and λ is the wavelength. Then, let N_a be the expected number of times per second the received SNR A passes downward across a given level a, we have

$$N_a = \sqrt{\frac{2\pi a}{\rho}} f_m \exp\left\{-\frac{a}{\rho}\right\}. \tag{33}$$

As mentioned in Section I, a finite-state Markov channel can be uniquely defined by the state transition matrix, the initial state probability vector, and the error probability vector. It was also noted that any partition of the received signal to noise ratio into a finite number of intervals forms a finite-state channel model. Let $0 = A_0 < A_1 < A_2 < \cdots < A_K = \infty$ be the thresholds of the received signal to noise ratio. Then the Rayleigh fading channel is said to be in state $s_k, k = 0, 1, 2, \cdots, K-1$, if the received SNR is in the interval $[A_k, A_{k+1})$. Associated with each state, there is a binary symmetric channel with crossover probability e_k. With the assumption of discrete channel structure, modulation and demodulation are considered as an inherent part of the channel. Given a specific digital modulation scheme, the average error probability is a function of the received signal to noise ratio. The crossover probability e_k for each state can then be related to the received SNR thresholds. In this study, the binary phase shift keying (BPSK) is assumed with coherent demodulation [18,19,20,21]. The error probability as a function of the received SNR can be written as

$$e_m(a) = 1 - F\left(\sqrt{2a}\right)$$

where

$$F(\alpha) = \int_{-\infty}^{\alpha} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{x^2}{2}\right\} dx.$$

With the p.d.f. of the received SNR as in (32), the steady state probability and the crossover probability for each state are

$$\begin{aligned} p_k &= \int_{A_k}^{A_{k+1}} \frac{1}{\rho} \exp\left\{-\frac{a}{\rho}\right\} \\ &= \exp\left\{-\frac{A_k}{\rho}\right\} - \exp\left\{-\frac{A_{k+1}}{\rho}\right\} \end{aligned} \tag{34}$$

and

$$e_k = \frac{\left[\int_{A_k}^{A_{k+1}} \frac{1}{\rho} \exp\left\{-\frac{a}{\rho}\right\}\left(1 - F(\sqrt{2a})\right) da\right]}{\left[\int_{A_k}^{A_{k+1}} \frac{1}{\rho} \exp\left\{-\frac{a}{\rho}\right\} da\right]}. \tag{35}$$

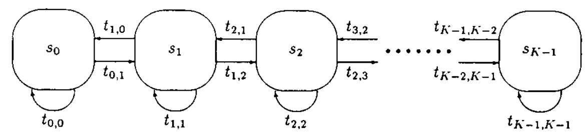

Fig. 4 K-state noisy channel with Markov transitions modeling a Rayleigh fading channel.

Equation (35) has been shown in the Appendix to be

$$e_k = \frac{\gamma_k - \gamma_{k+1}}{p_k} \tag{36}$$

where

$$\begin{aligned} \gamma_k &= \exp\left\{-\frac{A_k}{\rho}\right\}\left(1 - F\left(\sqrt{2A_k}\right)\right) \\ &\quad + \sqrt{\frac{\rho}{\rho+1}} F\left(\sqrt{\frac{2A_k(\rho+1)}{\rho}}\right). \end{aligned} \tag{37}$$

A similar calculation shows that the average error probability e is

$$e = \frac{1}{2}\left(1 - \sqrt{\frac{\rho}{\rho+1}}\right)$$

as in [16]. This expression can also be obtained by substituting (36) and (37) into (5).

B. Channel Modeling and Verification

To calculate the transition probabilities $t_{j,k}$ defined in (1), we make the following assumptions. We first assume that the Rayleigh fading channel is slow enough that the received SNR remains at a certain level for the time duration of a channel symbol. Furthermore, the channel states associated with consecutive symbols are assumed to be neighboring states. In other words, each state can have no more than three outgoing and incoming transitions as shown in Fig. 4. That is,

$$t_{j,k} = 0, \quad \forall\, |j-k| > 1.$$

Now, consider a communication system with a transmission rate of $R_t$symbols per second. There are, on the average,

$$R_t^{(k)} = R_t \times p_k \tag{38}$$

symbols per second transmitted during which the channel is in state s_k. Due to the slow fading assumption, we can conclude that the level crossing rate in (33) at A_k and/or A_{k+1} is much smaller than the value $R_t^{(k)}$. The transition probability $t_{k,k+1}$ can then be approximated by the ratio of the expected level crossing at A_{k+1} divided by the average symbols per second the SNR falls in the interval associated with state s_k. Similarly, the transition probability $t_{k,k-1}$ can be approximated by the ratio of the expected level crossing at A_k divided by the average symbols per second the SNR falls in the interval associated with state s_k. Specifically, let $N_k, k = 1, 2, 3, \cdots, K-1$, be the expected number of times

TABLE I
ANALYTICAL VALUES OF THE TRANSITION PROBABILITIES FOR EIGHT-STATE MARKOV CHANNEL WITH THE MAXIMUM DOPPLER FREQUENCY f_m EQUAL TO 10 AND 100 Hz

	$f_m = 10$ Hz			$f_m = 100$ Hz		
	$t_{k,k-1}$	$t_{k,k}$	$t_{k,k+1}$	$t_{k,k-1}$	$t_{k,k}$	$t_{k,k+1}$
k=0	–	0.999359	0.000641	–	0.993588	0.006412
k=1	0.000641	0.998552	0.000807	0.006412	0.985521	0.008067
k=2	0.000807	0.998334	0.000859	0.008067	0.983341	0.008592
k=3	0.000859	0.998306	0.000835	0.008592	0.983060	0.008348
k=4	0.000835	0.998420	0.000745	0.008348	0.984205	0.007447
k=5	0.000745	0.998665	0.000590	0.007447	0.986650	0.005903
k=6	0.000590	0.999048	0.000361	0.005903	0.990483	0.003615
k=7	0.000361	0.999639	–	0.003615	0.996385	–

TABLE II
SIMULATION VALUES OF THE TRANSITION PROBABILITIES FOR EIGHT-STATE MARKOV CHANNEL WITH THE MAXIMUM DOPPLER FREQUENCY f_m EQUAL TO 10 AND 100 Hz

	$f_m = 10$ Hz			$f_m = 100$ Hz		
	$t_{k,k-1}$	$t_{k,k}$	$t_{k,k+1}$	$t_{k,k-1}$	$t_{k,k}$	$t_{k,k+1}$
k=0	–	0.999304	0.000696	–	0.993198	0.006802
k=1	0.000690	0.998419	0.000891	0.006774	0.984704	0.008523
k=2	0.000879	0.998170	0.000951	0.008570	0.982362	0.009068
k=3	0.000894	0.998217	0.000890	0.008973	0.982441	0.008586
k=4	0.000876	0.998349	0.000775	0.008705	0.983559	0.007736
k=5	0.000777	0.998590	0.000633	0.007672	0.986202	0.006126
k=6	0.000637	0.998945	0.000418	0.006295	0.989790	0.003914
k=7	0.000384	0.999616	–	0.003878	0.996122	–

per second the received SNR passes downward across the threshold A_k. Then, from (33), we have

$$N_k = \sqrt{\frac{2\pi A_k}{\rho}} f_m \exp\left\{-\frac{A_k}{\rho}\right\}.$$

The Markov transition probabilities are then approximated by

$$t_{k,k+1} \approx \frac{N_{k+1}}{R_t^{(k)}}, \quad k = 0, 1, 2, \cdots, K-2 \tag{39}$$

and

$$t_{k,k-1} \approx \frac{N_k}{R_t^{(k)}}, \quad k = 1, 2, 3, \cdots, K-1. \tag{40}$$

The values of $t_{0,0}, t_{K-1,K-1}$, and $t_{k,k}, k = 1, 2, 3, \cdots, K-2$ are given by

$$t_{0,0} = 1 - t_{0,1}, \quad t_{K-1,K-1} = 1 - t_{K-1,K-2}$$

and

$$t_{k,k} = 1 - t_{k,k-1} - t_{k,k+1}, \quad k = 1, 2, 3, \cdots, K-2$$

respectively.

To verify this finite-state Markov channel model, we introduce the state equilibrium equations for each state. At state s_0, it is only necessary to show that

$$p_0 \times t_{0,1} = p_1 \times t_{1,0}. \tag{41}$$

By substituting (38)–(40) into (41), we see that both sides of (41) equal N_1/R_t. Similarly, both $p_{K-2} \times t_{K-2,K-1}$ and $p_{K-1} \times t_{K-1,K-2}$ equal N_{K-1}/R_t. This verifies the equilibrium at state s_{K-1}. For states $s_1, s_2, \cdots, s_{K-2}$, we need to show that

$$p_k \times (t_{k,k-1} + t_{k,k+1}) = p_{k-1} \times t_{k-1,k} + p_{k+1} \times t_{k+1,k} \tag{42}$$

for $k = 1, 2, 3, \cdots, K-2$. Again, by substituting (38)–(40) into (42), we see that the left hand side of (42) is given by

$$\frac{R_t^{(k)}}{R_t} \times \left(\frac{N_{k+1}}{R_t^{(k)}} + \frac{N_k}{R_t^{(k)}}\right) \tag{43}$$

and the right hand side is

$$\frac{R_t^{(k-1)}}{R_t} \times \frac{N_k}{R_t^{(k-1)}} + \frac{R_t^{(k+1)}}{R_t} \times \frac{N_{k+1}}{R_t^{(k+1)}}. \tag{44}$$

The equality between (43) and (44) shows the state equilibrium condition at states $s_1, s_2, \cdots, s_{K-2}$. We then conclude that the state equilibrium equations hold.

To show the accuracy of our analytical model, computer simulation is performed following the model in [15]. In our simulation, the number of waves is equal to 50 and the maximum Doppler frequency f_m is equal to 10 Hz and 100 Hz. We use an eight-state Markov channel model for 60 s simulation. The transmission rate is equal to 10^5 symbols/s (symbol duration of 10 μs) and therefore 6×10^6 observations are obtained. For the purpose of verification, the choice of thresholds $\{A_1, A_2, \cdots, A_{K-1}\}$ is arbitrary. We choose the one such that $p_0 = p_1 = \cdots = p_{K-1}$ for our simulation. Note that due to the non-linearity between SNR and e_k, the SNR intervals may have to be non-uniform to be useful. In fact, optimization of thresholds for a particular environment under mean square error criterion has been addressed in [5] and the SNR intervals are indeed non-uniform.

The transition probability $t_{i,j}$'s from analytical approximation and computer simulation are presented in Tables I and II. From these tables, the accuracy of our analytical model is verified. The assumption that limits transitions to their neighboring states is again confirmed by the simulation.

Note that the signaling rate used in [4] is 1 kbit/s which eventually makes the fading fast. As a result, the transitions between states other than neighboring ones are observed there. For communication environments such as cellular mobile communication and cordless telephone, the transmission rate applied here are more pratical.

V. CONCLUSIONS

In this paper, the Rayleigh fading channel is modeled as finite-state Markov channel (FSMC). The relationship between the channel and its model is investigated. The validity and accuracy of the FSMC as a model for the Rayleigh fading channel is also shown through the state equilibrium equations and computer simulation. With this useful model, the effect of fading to the coding efficiency and the channel capacity can be further investigated. Moreover, it is clear that the partitioning of the received SNR has certain effect over the resulting FSMC model. When this is known to transmitter or receiver, the search of a partitioning that improves the system performance is an interesting problem.

Appendix
Proof of (36) and (37)

It is easy to see that the denominator in (35) is simply the steady state probability p_k given in (34). The numerator of (35) is

$$\begin{aligned}\int_{A_k}^{A_{k+1}}&\left(1-\int_{-\infty}^{\sqrt{2a}}\frac{1}{\sqrt{2\pi}}e^{-x^2/2}dx\right)\frac{1}{\rho}e^{-a/\rho}\,da\\ &=\int_{A_k}^{A_{k+1}}\frac{1}{\rho}e^{-a/\rho}\,da\\ &\quad-\int_{A_k}^{A_{k+1}}\int_{-\infty}^{\sqrt{2a}}\frac{1}{\sqrt{2\pi}}e^{-x^2/2}\frac{1}{\rho}e^{-a/\rho}\,dx\,da. \quad (45)\end{aligned}$$

The second term of the right hand side in (45) can be simplified by changing the order of integration yielding

$$\begin{aligned}\int_{\sqrt{2A_k}}^{\sqrt{2A_{k+1}}}&\int_{x^2/2}^{A_{k+1}}\frac{1}{\rho}e^{-a/\rho}da\,\frac{1}{\sqrt{2\pi}}e^{-x^2/2}dx\\ &+\int_{\infty}^{\sqrt{2A_k}}\int_{A_k}^{A_{k+1}}\frac{1}{\rho}e^{-a/\rho}da\,\frac{1}{\sqrt{2\pi}}e^{-x^2/2}dx\\ &=\int_{\sqrt{2A_k}}^{\sqrt{2A_{k+1}}}\frac{1}{\sqrt{2\pi}}e^{(-x^2/2)((\rho+1)/\rho)}\,dx\\ &\quad-e^{-A_{k+1}/\rho}\int_{\sqrt{2A_k}}^{\sqrt{2A_{k+1}}}\frac{1}{\sqrt{2\pi}}e^{-x^2/2}\,dx\\ &\quad+\left(e^{-A_k/\rho}-e^{-A_{k+1}/\rho}\right)\int_{\infty}^{\sqrt{2A_k}}\frac{1}{\sqrt{2\pi}}e^{-x^2/2}\,dx\\ &=\sqrt{\frac{\rho}{\rho+1}}\left[F\left(\sqrt{\frac{2A_{k+1}(\rho+1)}{\rho}}\right)\right.\\ &\quad\left.-F\left(\sqrt{\frac{2A_k(\rho+1)}{\rho}}\right)\right]-e^{-A_{k+1}/\rho}F\left(\sqrt{2A_{k+1}}\right)\\ &\quad+e^{-A_k/\rho}F\left(\sqrt{2A_k}\right). \quad (46)\end{aligned}$$

With (34), (46), and the definition of γ_k given in (37), we can write the right hand side of (45) as $\gamma_k-\gamma_{k+1}$. Equations (36) and (37) are then proved.

References

[1] E. N. Gilbert, "Capacity of a burst-noise channel," *Bell Syst. Tech. J.*, vol. 39, pp. 1253–1265, Sept. 1960.

[2] E. O. Elliott, "Estimates of error rates for codes on burst-noise channels," *Bell Syst. Tech. J.*, vol. 42, pp. 1977–1997, Sept. 1963.

[3] M. Mushkin and I. Bar-David, "Capacity and coding for the Gilbert–Elliott channels," *IEEE Trans. Info. Theory*, vol. 35, pp. 1277–1290, Nov. 1989.

[4] F. Swarts and H. C. Ferreira, "Markov characterization of channels with soft decision outputs," *IEEE Trans. Commun.*, vol. 41, pp. 678–682, May 1993.

[5] H. S. Wang and N. Moayeri, "Modeling, capacity, and joint source/channel coding for Rayleigh fading channels," in *Proc. 43rd IEEE Vehicular Technol. Conf.*, May 1993.

[6] R. J. McEliece and W. E. Stark, "Channels with block interference," *IEEE Trans. Info. Theory*, vol. IT-30, pp. 44–53, Jan. 1984.

[7] N. Farvardin and V. Vaishampayan, "Optimal quantizer design for noisy channels: an approach to combined source-channel coding," *IEEE Trans. Info. Theory*, vol. IT-33, pp. 827–838, Nov. 1987.

[8] A. J. Kurtenbach and P. A. Wintz, "Quantizing for noisy channels," *IEEE Trans. Commun. Tech.*, vol. COM-17, pp. 291–302, Apr. 1969.

[9] S. P. Lloyd, "Least squares quantization in PCM," *IEEE Trans. Info. Theory*, vol. IT-28, pp. 129–136, Mar. 1982.

[10] J. Max, "Quantization for minimum distortion," *IEEE Trans. Info. Theory*, vol. IT-6, pp. 7–12, Mar. 1960.

[11] J. Lukaszewicz and H. Steinhaus, "On measuring by comparison," *Zastos Mat.*, vol. 2, pp. 225–231, 1955.

[12] H. S. Wang and N. Moayeri, "Variable rate transmission over a meteor burst channel," in *Proc. 25th Ann. Conf. Info. Sci. Syst.*, 1991.

[13] B. D. Fritchman, "A binary channel characterization using partitioned Markov chains," *IEEE Trans. Info. Theory*, vol. IT-13, Apr. 1967.

[14] E. Lutz, D. Cygan, M. Dippold, F. Dolainsky, and W. Papke, "The land mobile satellite communication channel—Recording, statistics, and channel model," *IEEE Trans. Veh. Technol.*, vol. 40, May 1991.

[15] W. C. Y. Lee, *Mobile Communications Engineering.* New York: McGraw-Hill, 1982.

[16] J. G. Proakis, *Digital Communications*, 2nd ed. New York: McGraw-Hill, 1989.

[17] W. C. Jakes, *Microwave Mobile Communications.* New York: Wiley, 1974.

[18] M. Yokoyama, "BPSK systems with sounder to combat Rayleigh fading in mobile radio communication," *IEEE Trans. Veh. Technol.*, vol. VT-34, Feb. 1985.

[19] F. Davarian, "Comments on BPSK systems with sounder to combat Rayleigh fading in mobile radio communication" *IEEE Trans. Veh. Technol.*, vol. VT-34, Nov. 1985.

[20] A. Bateman, "A general analysis of bit error probability for reference-based BPSK mobile data transmission," *IEEE Trans. Commun.*, vol. 37, pp. 398–402, Apr. 1989.

[21] D. Subaslinghe-Dias and K. Feher, "Miller coded pilot aided modulation schemes for digital mobile radio," in *Proc. 40th IEEE Vehicular Technol. Conf.*, May 1990, pp. 465–470.

Hong Shen Wang (S'89–M'92) received the B.C. degree in statistics from the College of Law and Commerce, National Chung-Hsing University, Taipei, Taiwan, in 1982. After fulfilling the military duty in 1984, he joined Rutgers, The State University of New Jersey, where he received the M.S. degrees in statistics and electrical engineering, and the Ph.D. degree in electrical engineering in 1986, 1988, and 1992, respectively.

Since 1992, he has been a lecturer of the Department of Information Engineering, the Chinese University of Hong Kong. His interests are in coding, information theory, and wireless communications.

Nader Moayeri (S'79–M'80–S'80–M'85–S'85–M'86–SM'90) studied electrical engineering at the Sharif University of Technology, Tehran, Iran, from 1974 to 1978. He received the M.S.E.E., M.S. CICE, and Ph.D. degrees in electrical engineering-systems from the University of Michigan, Ann Arbor, MI, in 1980, 1981, and 1986, respectively.

He was with the Department of Electrical and Computer Engineering at Rutgers, The State University of New Jersey, from 1986 to 1994. Since June 1994, he has been with the Hewlett-Packard Laboratories, Palo Alto, CA. His research interests are in data compression, mobile radio communications, information theory, and routing and flow control in data networks.

Dr. Moayeri served the Princeton/Central Jersey Section of the IEEE in various capacities, among which was Section Chair (1992–1993) and founder and first Chairman (1989–1993) of the Information Theory Chapter. He was a Co-Chairman of the Joint DIMACS/IEEE Workshop on Coding and Quantization, held at Rutgers University in October 1992.