Channel Capacity and Error Exponents of Variable Rate Adaptive Channel Coding for Rayleigh Fading Channels

Vincent K. N. Lau, *Member, IEEE*

***Abstract*—We have evaluated the information theoretical performance of *variable rate* adaptive channel coding for Rayleigh fading channels. The channel states are detected at the receiver and fed back to the transmitter by means of a noiseless feedback link. Based on the channel state informations, the transmitter can adjust the channel coding scheme accordingly. Coherent channel and arbitrary channel symbols with a fixed average transmitted power contraint are assumed. The *channel capacity* and the *error exponent* are evaluated and the optimal rate control rules are found for Rayleigh fading channels with feedback of channel states. It is shown that the variable rate scheme can only increase the channel error exponent. The effects of additional practical constraints and finite feedback delays are also considered. Finally, we compare the performance of the variable rate adaptive channel coding in high bandwidth-expansion systems (CDMA) and high bandwidth-efficiency systems (TDMA).**

I. Introduction

ERROR correction codes have been widely used to combat the effect of Rayleigh fading in mobile radio channels. In traditional FEC schemes [1], [2], fixed rate codes were used which failed to explore the time varying nature of the channel. To keep the performance at a desirable level, they were designed for the average or worst case situation. To better exploit the time varying nature of the channel, adaptive channel coding based on feedback channel state has been proposed. The performance of uncoded variable rate and power transmission schemes for Rayleigh fading channel based on the feedback of channel state information has been considered in [3]–[7]. Many practical adaptive error correction codes have been proposed in recent years to reduce the bit error rate and to increase throughput of the mobile radio channels [8]–[14], [15]. In this paper, we model a general scheme of variable rate adaptive channel coding which varies the code rate according to the channel condition and explore the fundamental reasons why there is a performance improvement over fixed-rate coding.

We investigate the information theoretical performance, namely the channel capacity and the error exponent, of Rayleigh fading channels using variable rate adaptive channel coding (VRAECC) with constant transmitted power.[1] Channel capacity describes the maximum allowable bit rate for reliable transmission across a channel. Error exponent describes how fast error probability drops w.r.t. block length. The channel states are detected at the receiver and fed back to the transmitter by means of a noiseless feedback link. Based on the channel state informations, the transmitter can adjust the *rate* of the channel coding scheme accordingly. We try to answer the following questions in this paper.

- Is channel capacity or error exponent increased by using VRAECC?
- What are the optimal rate control functions that maximize the error exponent?
- Does VRAECC perform better in high or low bandwidth expansion?

An equivalent discrete time channel model is developed in Section II. For simplicity, coherent detection and ideal interleaving are assumed. The error exponent and channel capacity of a Rayleigh fading channel with feedback of channel state using *constant input* VRACE are evaluated in Sections III and IV, respectively. Numerical results are presented and discussed in Section V. Finally, we conclude with a brief summary in Section VI.

II. Channel Modeling and Induced State Distribution

A. Physical Channel Model

The physical Rayleigh fading channel is a bandlimited continuous-time channel in which the channel input can be modeled by a bandlimited complex random process. The random process is segmented into a number of channel symbols with the ith channel symbol $\tilde{X}_i(t)$ having a variable duration T_i, as shown in Fig. 1. To maintain generality, no modulation format is specified. Variable rate channel encoder is integrated

Paper approved by J. Huber, the Editor for Coding and Coded Modulation of the IEEE Communications Society. Manuscript received July 7, 1997; revised May 8, 1998 and February 19, 1999. This work was supported in part by the University of Hong Kong Area of Fundamentals in Information Technology. This paper was presented in part at the Sixth IMA Conference on Cryptography and Coding, Cirencester, U.K., December 1997.

The author is with the Department of Electrical and Electronic Engineering, University of Hong Kong, Hong Kong (e-mail: knlau@eee.hku.hk).

Publisher Item Identifier S 0090-6778(99)07467-X.

[1] In particular, channel capacity of downlink fading channels with variable power schemes has been investigated in [16]. It is shown that the optimal power distribution that maximize the channel capacity is achieved by water-filing in time domain. However, due to the rapid power control required to compensate for the channel fading in the scheme, it is not feasible with nowadays power-amplifier technology. Variable rate transmission, on the other hand, is more feasible and hence, we focus on the optimal variable rate schemes in this paper. As shown in Section IV, variable rate schemes cannot increase the channel capacity. Instead, we aim at increasing the error exponent with variable rate schemes. Optimal rate control rule is in the sense of maximizing the error exponent.

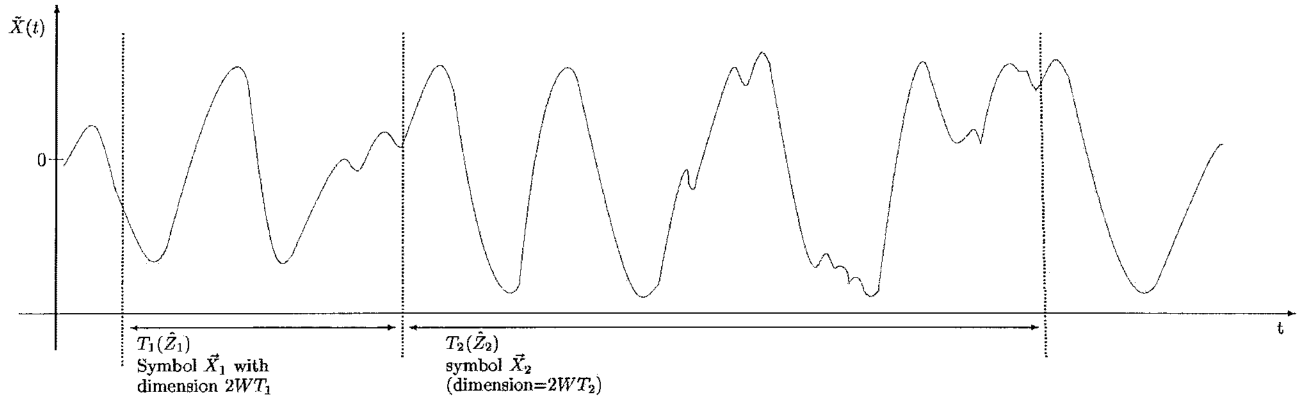


Fig. 1. Segmentation of a bandlimited random process $\tilde{X}(t)$ into a channel symbol $\vec{X}$.

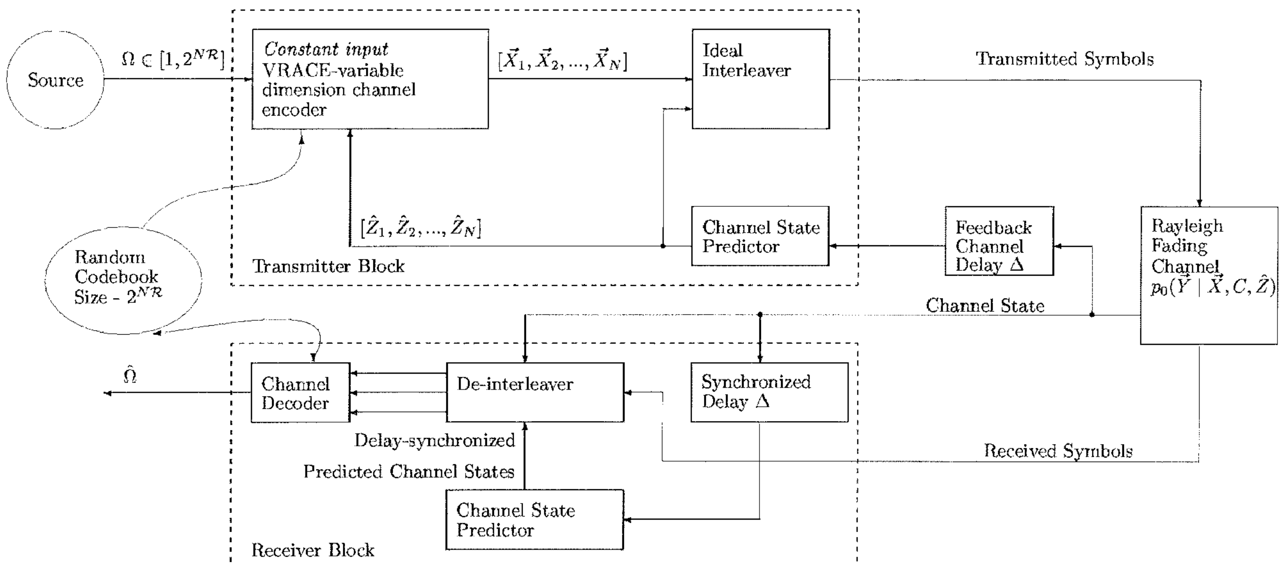


Fig. 2. Equivalent discrete-time channel model for *variable duration* scheme.

into variable throughput modulator with an average power constraint P_0 only. The ith channel output (in complex low-pass equivalent) $\tilde{Y}_i(t)$ is given by

$$\tilde{Y}_i(t) = C_i(t)\tilde{X}_i(t) + \tilde{N}_i(t) \tag{1}$$

where $C_i(t)$ and $\tilde{N}_i(t)$ are channel fading attenuation[2] and complex white Gaussian noise for the ith channel symbol, respectively.

It has been shown that a continuous-time complex signal which is *approximately* time-limited to T_i and bandlimited to W can be represented by a $2WT_i$-dimensional vector in the signal space spanned by the Prolate spheroidal wave functions [17]. Hence, the ith channel symbol is represented by a $2WT_i$-dimensional vector $\vec{X}_i$. Assume that T_i and W are both much smaller than the coherence time, and the coherence bandwidth[3] $C(t)$ can be considered as a constant in every dimension of the signal space. Hence, the continuous-time model is reduced to a discrete-time model.

[2] Since coherent detection is assumed, channel phase variation is corrected by the receiver and hence, WLOG, the channel phase reference is set to zero.

[3] We have a flat fading channel.

B. Equivalent Discrete-Time Channel Model

The equivalent channel is a discrete-time, continuous-input and continuous-output channel with feedback. There is a channel state C_i associated with the ith channel symbol. The channel state is available to the receiver and known to the transmitter via a feedback channel with a certain unavoidable delay, Δ seconds. For each C_i, there is a corresponding prediction, denoted by $\hat{Z}_i$, at the transmitter. The channel states C_i and C_j (and hence, $\hat{Z}_i$ and $\hat{Z}_j$) are correlated but through ideal interleaving, they become i.i.d. and the channel becomes a memoryless channel.

The ith channel output is given by

$$\vec{Y}_i = C_i\vec{X}_i + \vec{N}_i \tag{2}$$

where $\vec{N}_i$ is an uncorrelated Gaussian noise variable with variance $\mathcal{E}[||\vec{N}_i||^2] = \eta_0 WT_i$ and η_0 is the white noise spectral density. We assume $\mathcal{E}[C_i^2] = 1$.

For the *variable duration* scheme, symbol duration of the ith channel symbol T_i is varying according to the predicted channel state $\hat{Z}_i$. The channel model is illustrated in Fig. 2. Each channel symbol carries a constant number of information bits $\mathcal{R}$ with a varying dimension $2WT_i$, which is a function

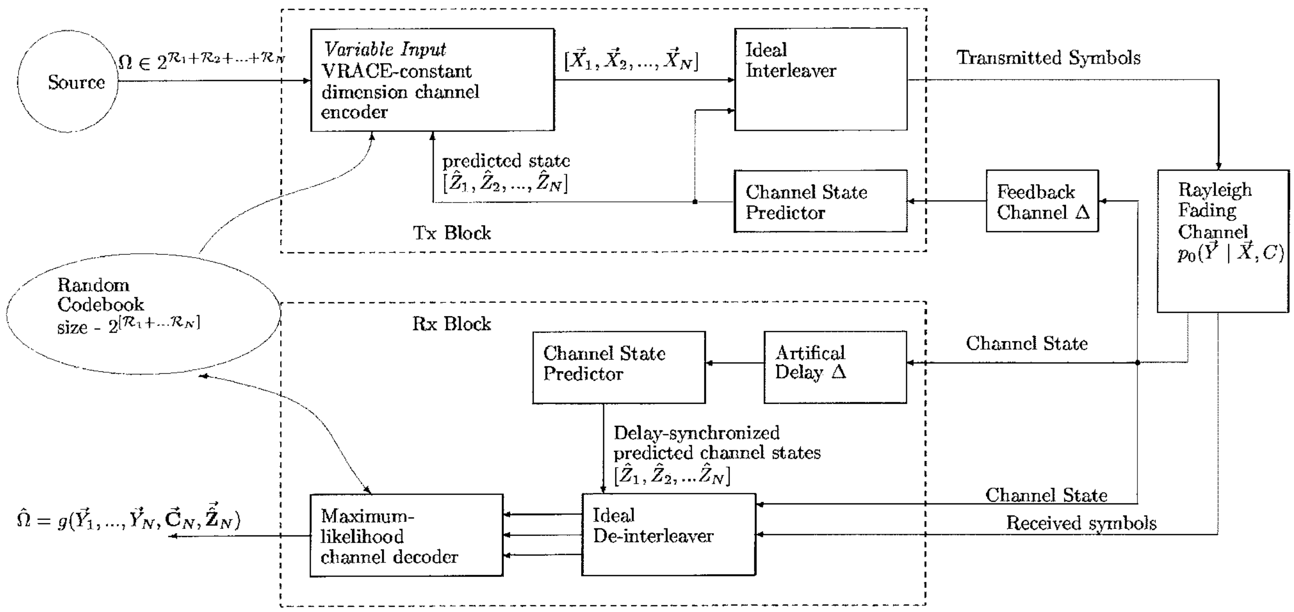


Fig. 3. Equivalent discrete-time channel model for *variable input* scheme.

of predicted channel state $\hat{Z}_i$. Random block coding [18] with block length N is used. An index $\Omega \in [1, 2^{N\mathcal{R}}]$ is fed into the *variable dimension* channel encoder, giving out a codeword of N channel symbols. Each channel symbol in a codeword $\vec{X}_i$ is generated randomly according to a continuous distribution $Q(\vec{x}_i \mid \hat{z}_i)$. This forms a random codebook of size $2^{N\mathcal{R}}$ which is known both to the transmitter and the receiver. The channel can be described by a channel transition density $p_0(\vec{y}_i \mid \vec{x}_i c_i \hat{z}_i)$. The dependence of Q and p_0 on $\hat{Z}_i$ comes from the dependence of the symbol dimension $2WT_i$ on $\hat{Z}_i$.

For the *variable input* scheme, the ith symbol duration T_i is a constant given by T_s. Fig. 3 illustrates the equivalent discrete-time channel model. The ith channel symbol carries $\mathcal{R}_i$ information bits and variable throughput is achieved by varying $\mathcal{R}_i$ w.r.t. the predicted state $\hat{Z}_i$. Random block coding is used and the ith channel symbol is generated randomly according to a continuous time distribution, $Q(\vec{X}_i \mid \hat{Z}_i)$. The overall size of the random codebook is $2^{[\mathcal{R}_1+\cdots+\mathcal{R}_N]}$ which is both known to the transmitter and the receiver. The channel can be described by a transition probability $p_0(\vec{Y}_i \mid \vec{X}_i, C_i)$. Note that due to the constant symbol duration, the channel transition probability is independent of the predicted state $\hat{Z}_i$.

For both schemes, channel outputs, together with channel states and predicted states, are fed into a deinterleaver and a maximum-likelihood decoder at the receiver.[4] The decoder produces an estimate of the transmitted index $\hat{\Omega}$ and an error occurs when $\Omega \neq \hat{\Omega}$.

C. Induced State Distribution

For the *variable duration* scheme, the sequence of symbol duration $\{T_1(\hat{z}_1), \cdots, T_N(\hat{z}_N)\}$ is varying according to the sequence of predicted channel states $\vec{\hat{\mathbf{Z}}}_N = \{\hat{Z}_1, \cdots, \hat{Z}_N\}$, and hence, it induces a probability density on $\vec{\hat{\mathbf{Z}}}_N$ which is different from the original fading density in general. For simplicity, we assume a simple prediction rule $\hat{z}(t) = c(t-\Delta)$. Hence, $\hat{z}(t)$ is an ergodic random process and it is shown in Appendix A that the induced probability density on $\vec{\hat{\mathbf{Z}}}_N$, denoted by $\mathcal{P}_N(\vec{\hat{\mathbf{z}}}_N)$, is given by

$$\mathcal{P}_N(\hat{z}_1, \hat{z}_2, \cdots, \hat{z}_N) = \mathcal{P}(\hat{z}_1)\mathcal{P}(\hat{z}_2)\cdots\mathcal{P}(\hat{z}_N) \tag{3}$$

where

$$\mathcal{P}(\hat{z}) \stackrel{\text{def}}{=} \lim_{\mathcal{N}\to\infty} \frac{\mathcal{N}_{\hat{z}}}{\mathcal{N}\delta} = \frac{f(\hat{z})}{E_f T(\hat{z})}. \tag{4}$$

$\mathcal{N}_{\hat{z}}$ is the number of symbols with $\hat{z} \in [\hat{z}, \hat{z} + \delta]$ in a sequence of $\mathcal{N}$ symbols, $f(\hat{z})$ is the fading density and $E_f = \int_{\hat{z}} (1/T(\hat{z})) f(\hat{z})\, d\hat{z}$. Furthermore, E_f is shown to be the average symbol rate (number of channel symbols per second). Since given $\hat{Z}$, the symbol duration T_i is constant, the conditional density $\mathcal{P}(c \mid \hat{z})$ is not affected by the varying symbol duration and is given by [21]

$$\begin{aligned}\mathcal{P}(c \mid \hat{z}) &= f(c \mid \hat{z}) \\ &= \frac{2c}{(1-\lambda^2)} \exp\left[-\frac{c^2 + \lambda^2 \hat{z}^2}{(1-\lambda^2)}\right] I_0\left(2c\hat{z}\frac{\lambda}{1-\lambda^2}\right)\end{aligned} \tag{5}$$

where $\lambda^2 = J_0(2\pi f_d \Delta)$, f_d is the Doppler spread, J_0 is the zeroth-order Bessel function, and I_0 is the zeroth-order modified Bessel function.

For the *variable input* scheme, the symbol duration is constant and hence, the *induced density* reduces to original standing fading density.

III. Error Exponent for Variable Rate Schemes

We shall bound the average codeword error probability by an average error exponent using Gallager's approach [22]. Given a sequence of channel states $\vec{\mathbf{c}}_N$ and a sequence of

[4] Note that because of the variable throughput, the interleaving task is nontrival. For the *variable duration* scheme, interleaving is done by *parallel interleaving* as illustrated in [19]. For the *variable input* scheme, interleaving is illustrated in [20]. For the sake of simplicity, we assumed ideal interleaving without further discussing the actual schemes.

predicted channel states $\vec{\mathbf{z}}_N$, the channel is equivalent to an additive white Gaussian noise (AWGN) channel. Hence, the conditional codeword error probability $P_e(\vec{\mathbf{c}}_N, \vec{\mathbf{z}}_N)$ is bounded by [22]

$$P_e(\vec{\mathbf{c}}_N, \vec{\mathbf{z}}_N) \leq 2^{-N[E_N(\rho,Q,\vec{\mathbf{c}}_N,\vec{\mathbf{z}}_N)-\rho\mathcal{R}]} \tag{6}$$

for all $N \geq 1$ and $\rho \in [0,1]$. $E_N(\rho, Q, \vec{\mathbf{c}}_N, \vec{\mathbf{z}}_N)$ is called the Gallager's error exponent and is given by

$$E_N(\rho, Q, \vec{\mathbf{c}}_N, \vec{\mathbf{z}}_N) = -\frac{1}{N}\sum_{j=1}^{N}\log_2\left\{\int_{\vec{y}_j}\left[\int_{\vec{x}_j} Q(\vec{x}_j \mid \hat{z}_j) \times p_0(\vec{y}_j \mid \vec{x}_j c_j \hat{z}_j)^{\frac{1}{\rho+1}}\, d\vec{x}_j\right]^{\rho+1} d\vec{y}_j\right\}. \tag{7}$$

To determine the average error probability, we have to uncondition (6) w.r.t. $\vec{\mathbf{c}}_N$ and $\vec{\mathbf{z}}_N$. The average codeword error probability is given by

$$\bar{P}_e = \int_{\vec{\mathbf{z}}_N}\int_{\vec{\mathbf{c}}_N} P_e(\vec{\mathbf{c}}_N, \vec{\mathbf{z}}_N)\mathcal{P}(c_1, \hat{z}_1)\cdots\mathcal{P}(c_N, \hat{z}_N)\, d\vec{\mathbf{c}}_N\, d\vec{\mathbf{z}}_N \tag{8}$$

where $\mathcal{P}(c_i, \hat{z}_i)$ is the joint density of c_i and $\hat{z}_i$.

By symmetry, the ith channel symbol $\vec{X}_i$ has $2WT_i$ i.i.d. components $[X_{i,1}, \cdots, X_{i,2WT_i}]$. Hence, $Q(\vec{x}_i \mid \hat{z}_i)$ and $p_0(\vec{y}_i \mid \vec{x}_i c_i \hat{z}_i)$ can be expressed using product forms as

$$Q(\vec{x}_i \mid \hat{z}_i) = \prod_{j=1}^{2WT_i} Q(x_{i,j} \mid \hat{z}_i) \tag{9}$$

$$p_0(\vec{y}_i \mid \vec{x}_i c_i \hat{z}_i) = \prod_{j=1}^{2WT_i} p_0(y_{i,j} \mid x_{i,j} c_i). \tag{10}$$

Separating (8) into product of N integrals and using (6), (9), and (10), the average error probability $\bar{P}_e$ is given by

$$\bar{P}_e \leq 2^{-N[\bar{E}'(\rho,Q)-\rho\mathcal{R}]} \tag{11}$$

for all $\rho \in [0,1]$ and $Q(x \mid \hat{z})$ where $\bar{E}'(\rho, Q)$ is given by

$$\bar{E}'(\rho, Q) = -\log_2\left\{\int_{\hat{z}}\int_c \mathcal{P}(c, \hat{z})\left[\int_y\left(\int_x Q(x \mid \hat{z}) \times p_0(y \mid xc)^{\frac{1}{\rho+1}}\, dx\right)^{\rho+1} dy\right]^{2WT(\hat{z})} dc\, d\hat{z}\right\}. \tag{12}$$

A. Variable Duration Scheme

Since $\{T_1(\hat{z}_1), \cdots, T_N(\hat{z}_N)\}$ is an i.i.d. sequence and by the weak law of large numbers [23]

$$\frac{1}{N}(T_1(\hat{z}_1) + \cdots + T_N(\hat{z}_N))$$

will converge in probability to $\mathcal{E}[T_i] = 1/E_f$, where the expectation is taken w.r.t. the induced density $\mathcal{P}(\hat{Z})$. The average information bit rate (bits per second) $\bar{\mathcal{R}}_b$ is given by

$$\bar{\mathcal{R}}_b \stackrel{\text{def}}{=} \frac{N\mathcal{R}}{T_1 + T_2 + \cdots + T_N} \stackrel{N\to\infty}{\to} \mathcal{R}E_f. \tag{13}$$

Hence, we have

$$\mathcal{R} = \bar{\mathcal{R}}_b/E_f \tag{14}$$

Expressing (11) and (12) in terms of $\bar{\mathcal{R}}_b$, the average error probability $\bar{P}_e$ is given by

$$\bar{P}_e \leq 2^{-\frac{N}{E_f}[\bar{E}(\rho,Q)-\rho\bar{\mathcal{R}}_b]} \tag{15}$$

where $\bar{E}(\rho, Q)$ is given by

$$\begin{aligned}\bar{E}(\rho, Q) &= E_f\bar{E}'(\rho, Q)\\ &= -E_f\log_2\left\{\frac{1}{E_f}\int_c\int_{\hat{z}}\frac{f(c \mid \hat{z})f(\hat{z})}{T(\hat{z})}\right.\\ &\quad\times\left[\int_y\left(\int_x Q(x \mid \hat{z})p_0(y \mid xc)^{\frac{1}{\rho+1}}dx\right)^{\rho+1} dy\right]^{2WT(\hat{z})}\\ &\quad\left.\times\, dc\, d\hat{z}\right\}.\end{aligned} \tag{16}$$

To obtain a tight error bound, we have to minimize $\bar{P}_e$ w.r.t. Q, ρ and $T(\hat{c})$. Since ρ is a parameter which is not measurable in practice, the functions $Q(x \mid \hat{z})$ and $T(\hat{c})$ are independent of the parameter ρ and hence, the optimization is decoupled.

1) Optimization w.r.t. ρ: Define the *average error exponent* $\bar{E}_r(\rho, Q, \bar{\mathcal{R}}_b)$ as

$$\bar{E}_r(\rho, Q, \bar{\mathcal{R}}_b) = [\bar{E}(\rho, Q) - \rho\bar{\mathcal{R}}_b]. \tag{17}$$

We first prove the following lemmas.

Lemma 1: For any $T(\hat{c})$ and $Q(X \mid \hat{Z})$, $\bar{E}$ in (16) satisfies the following properties:

(i) $0 < \dfrac{\partial\bar{E}}{\partial\rho} \leq E_f I(\vec{X}; \vec{Y} \mid C\hat{Z})$. Equality holds iff $\rho = 0$.

(ii) $\dfrac{\partial^2\bar{E}}{\partial\rho^2} < 0$.

Proof: Refer to Appendix C-1. □

Let

$$R_0 = \left.\frac{\partial\bar{E}}{\partial\rho}\right|_{\rho=0} \tag{18}$$

$$R_1 = \left.\frac{\partial\bar{E}}{\partial\rho}\right|_{\rho=1}. \tag{19}$$

The following summarizes the general result of the optimization w.r.t. ρ.

Lemma 2: For any $T(\hat{c})$ and $Q(X \mid \hat{Z})$, the optimal error exponent is given by

$$\bar{E}_r^*(Q, \bar{\mathcal{R}}_b) \stackrel{\text{def}}{=} \max_\rho \bar{E}_r(\rho, Q, \bar{\mathcal{R}}_b) = \bar{E}(\rho^*, Q) - \rho^*\bar{\mathcal{R}}_b \tag{20}$$

where $\rho^* \in [0,1]$, a function of $\bar{\mathcal{R}}_b$, is given by

$$\begin{cases}\left.\frac{\partial\bar{E}}{\partial\rho}\right|_{\rho=\rho^*} = \bar{\mathcal{R}}_b & R_1 \leq \bar{\mathcal{R}}_b \leq R_0\\ \rho^* = 1 & R_1 \leq \bar{\mathcal{R}}_b \leq 0.\end{cases} \tag{21}$$

Proof: Refer to Appendix C-2. □

Collecting the above results, we have the following theorem.

Theorem 1: For any $Q(X \mid \hat{Z})$ and $T(\hat{c})$, if $\bar{\mathcal{R}}_b < R_0$, then $\lim_{N\to\infty}\bar{P}_e = 0$.

Proof: Consider

$$\frac{\partial\bar{E}_r^*(Q, \bar{\mathcal{R}}_b)}{\partial\bar{\mathcal{R}}_b} = \frac{\frac{\partial\bar{E}_r^*}{\partial\rho^*}}{\frac{\partial\bar{\mathcal{R}}_b}{\partial\rho^*}} = \frac{\left[\frac{\partial\bar{E}}{\partial\rho^*} - \bar{\mathcal{R}}_b\right] - \rho^*\frac{\partial\bar{\mathcal{R}}_b}{\partial\rho^*}}{\frac{\partial\bar{\mathcal{R}}_b}{\partial\rho^*}} = -\rho^* < 0$$

for $\bar{\mathcal{R}}_b \in [R_1, R_0]$. $\frac{\partial\bar{E}_r^*(Q,\bar{\mathcal{R}}_b)}{\partial\bar{\mathcal{R}}_b} = -1$ for $\bar{\mathcal{R}}_b \in [0, R_1]$.

Hence, $\bar{E}_r^*$ is strictly decreasing w.r.t. $\bar{\mathcal{R}}_b$. At $\bar{\mathcal{R}}_b = R_0$, $\rho^* = 0$ and $\bar{E}_r^*(Q^*, \bar{\mathcal{R}}_b) = 0$. If $\bar{\mathcal{R}}_b < R_0$, we have $\bar{E}_r^*(\bar{\mathcal{R}}_b) > \bar{E}_r^*(R_0) = 0$ and hence from (15), $\bar{P}_e \to 0$ as $N \to \infty$. □

2) Optimization w.r.t.$Q(x \mid \hat{z})$: For simplicity, take Q to be a capacity achieving distribution,[5] Q^*, which maximizes the mutual information. Given a channel state C_i and a predicted state $\hat{Z}_i$, the channel is memoryless and is equivalent to an AWGN channel. By symmetry, the capacity achieving distribution [22], [1] $Q^*(x_{i,j} \mid \hat{z}_i)$ would be a zero-mean Gaussian density with variance $\sigma_X^2 = P_0/(2W)$. Since the variance σ_X^2 is independent of $\hat{Z}_i$, $Q^*(x_{i,j} \mid \hat{z}_i)$ is independent of $\hat{Z}_i$; we shall drop the conditional notation of $\hat{Z}$ in Q^* hereafter.

The remaining problem is to minimize $\bar{P}_e$ w.r.t. $T(\hat{z})$.

3) Optimization w.r.t. $T(\hat{z})$: In this section, we minimize $\bar{P}_e$ w.r.t. $T(\hat{z})$. For any given $\bar{\mathcal{R}}_b$, take $\rho = \rho^*$ as in (21) and $Q(x) = Q^*(x)$. Two situations, namely negligible feedback delay and significant feedback delay, are considered as follows.

a) Small feedback delay: We assume feedback delay is small relative to the channel coherence time. Therefore, $\hat{z}(t) = c(t - \Delta) \approx c(t)$ and

$$f(c \mid \hat{z}) \to \delta(c - \hat{z}). \tag{22}$$

By (22), rewrite $\bar{P}_e$ in (15) as

$$\bar{P}_e \le 2^{\frac{N\rho^*\bar{\mathcal{R}}_b}{E_f}} \frac{1}{E_f^N} \prod_{j=1}^{N} \left\{ \int_{\hat{z}_j} \frac{f(\hat{z}_j)}{T_j(\hat{z}_j)} \left[\int_y \left(\int_x Q^*(x) \times p_c(y \mid x\hat{z})^{\frac{1}{\rho^*+1}} \, dx \right)^{\rho^*+1} dy \right]^{2WT_j(\hat{z}_j)} d\hat{z}_j \right\}. \tag{23}$$

The optimization problem is equivalent to choosing $\{T_1, T_2, \cdots, T_N\}$ that minimize $\bar{P}_e$ in (23) under the constraints

$$\frac{1}{E_f} \int_{\hat{z}_i} \frac{f(\hat{z}_i)}{T(\hat{z}_i)} \, d\hat{z}_i = 1 \quad \forall i \in [1, N] \tag{24}$$

$$T(\hat{z}_i) \ge T_l \quad \forall i \in [1, N] \tag{25}$$

$$T(\hat{z}_i) \le T_p \quad \forall i \in [1, N]. \tag{26}$$

Constraint (24) is due to the fact that total area under the induced density $\mathcal{P}(\hat{z}_i)$ should be equal to one. Constraint (25) is to set a lower limit on $T(\hat{z})$ so that $2WT_i \ge 2$. Constraint (26) is to set a peak limit on $T(\hat{z})$ so that symbol duration is smaller than the channel coherence time.

By the Calculus of Variations, it is shown in Appendix B that the optimal[6] $T(\hat{z})$ is given by

$$T(\hat{z}, \bar{\mathcal{R}}_b) = \begin{cases} \dfrac{A_0(\rho^*)}{\log_2 \left[\int_y \left[\int_x Q^*(x) p_0(y \mid x\hat{z})^{1/(\rho^*+1)} \, dx \right]^{\rho^*+1} dy \right]}, & \hat{z}_l \ge \hat{z} \ge \hat{z}_p \\ T_p, & \hat{z}_p \ge \hat{z} \ge 0 \\ T_l, & \hat{z} \ge \hat{z}_l \end{cases} \tag{27}$$

[5] Although the capacity achieving distribution will not, in general, optimize $E\bar{(}9)(\rho^*, Q)$ for all $\bar{\mathcal{R}}_b$ (and hence ρ^*), the obtained error bound can serve as an upper bound.

[6] Since ρ^* is a function of $\bar{\mathcal{R}}_b$, $T(\hat{z})$ is a function of $\bar{\mathcal{R}}_b$ as well.

where $\hat{z}_p$ is given by the solution of

$$T_P = \frac{A_0(\rho^*)}{\log_2 \left[\int_y \left[\int_x Q^*(x) p_0(y \mid x\hat{z}_p)^{1/(\rho^*+1)} \, dx \right]^{\rho^*+1} dy \right]} \tag{28}$$

$\hat{z}_l$ is given by the solution of

$$T_l = \frac{A_0(\rho^*)}{\log_2 \left[\int_y \left[\int_x Q^*(x) p_0(y \mid x\hat{z}_l)^{1/(\rho^*+1)} \, dx \right]^{\rho^*+1} dy \right]}. \tag{29}$$

Intuitively, a longer symbol duration should be used to encode $\mathcal{R}$ information bits when the predicted state $\hat{z}$ is small.

Substituting (27) into the constraints (24), (25), and (26), the constant $A_0(\rho^*)$ is given by

$$\begin{aligned} A_0(\rho^*) &= \frac{1}{E_f} \left\{ \int_{\hat{z}_p}^{\hat{z}_l} \log_2 \left[\int_y \left[\int_x Q^*(x) \times p_0(y \mid x\hat{z})^{\frac{1}{\rho^*+1}} \, dx \right]^{\rho^*+1} dy \right] \right. \\ &\quad \cdot f(\hat{z}) \, d\hat{z} + \log_2 \left[\int_y \left[\int_x Q^*(x) p_0(y \mid x\hat{z}_p)^{\frac{1}{\rho^*+1}} \, dx \right]^{\rho^*+1} dy \right] \\ &\quad \times \int_0^{\hat{z}_p} f(\hat{z}) \, d\hat{z} + \log_2 \left[\int_y \left[\int_x Q^*(x) \right. \right. \\ &\quad \left. \left. \cdot\, p_0(y \mid x\hat{z}_l)^{\frac{1}{\rho^*+1}} \, dx \right]^{\rho^*+1} dy \right] \times \left. \int_{\hat{z}_l}^{\infty} f(\hat{z}) \, d\hat{z} \right\}. \end{aligned} \tag{30}$$

The optimal error exponent $\bar{E}_r^*(Q^*, \bar{\mathcal{R}}_b)$ is found by solving the simultaneous equations of (20), (21) in Lemma 2, as well as (27) numerically for any given $\bar{\mathcal{R}}_b \in [0, R_0]$.

b) Large feedback delay: When feedback delay is large, (22) no longer holds. Since the integrand of $\bar{P}_e$ is not separable w.r.t. c and $\hat{z}$, it is not possible to obtain a closed-form expression for the optimal $T(\hat{z})$. We shall investigate the effect of feedback delay on the performance using the control rule in (27) instead.

4) Overall Result of Optimizations for Variable Duration Scheme: Given $\bar{\mathcal{R}}_b$, take ρ^*, $Q^*(x \mid \hat{z})$ and $T(\hat{z}, \bar{\mathcal{R}}_b)$ to be the optimal parameter, the capacity achieving distribution and the optimal symbol duration control rule, respectively. The solution of the average error exponent $\bar{E}_r^*(Q^*, \bar{\mathcal{R}}_b, \Delta)$ with delay Δ is given by the following.

i) $0 \le \bar{\mathcal{R}}_b \le R_1$: Using $T(\hat{z}, \bar{\mathcal{R}}_b)$ in (27) with $\rho^* = 1$, the solution is given by

$$\bar{E}_r^*(Q^*, \bar{\mathcal{R}}_b, \Delta) = \bar{E}(1, Q^*, \Delta) - \bar{\mathcal{R}}_b. \tag{31}$$

ii) $R_1 < \bar{\mathcal{R}}_b \le R_0$: Using $T(\hat{z}, \bar{\mathcal{R}}_b)$ in (27), the solution is given by the following nonlinear parametric equations:

$$\begin{cases} \bar{\mathcal{R}}_b = \left. \frac{\partial \bar{E}(\rho, Q^*, \Delta)}{\partial \rho} \right|_{\rho=\rho^*} \\ \bar{E}_r^*(Q^*, \bar{\mathcal{R}}_b, \Delta) = \bar{E}(\rho^*, Q^*, \Delta) - \rho^* \bar{\mathcal{R}}_b \end{cases}. \tag{32}$$

B. Variable Input Schemes

Because of the constant symbol duration, the *induced density* in Section II-C is reduced to the standard fading density $f(c, \hat{z})$. Using similar techniques to Section III-A, the average codeword error probability $\bar{P}_e$ is bounded by

$$\bar{P}_e \le 2^{-N\bar{E}_r(\rho, Q, \mathcal{R})} \tag{33}$$

where

$$\bar{E}_r(\rho, Q, \mathcal{R}) = -\log_2\left\{\int_c\int_{\hat{z}} 2^{-[E(\rho,Q,c,\hat{z})-\rho\mathcal{R}]} f(c,\hat{z})\,dc\,d\hat{z}\right\} \tag{34}$$

$$E(\rho, Q, c, \hat{z}) = 2WT_s\log_2\left\{\int_y\left[\int_x Q(x \mid \hat{z})p_0(y \mid xc)^{\frac{1}{\rho+1}}\,dx\right]^{\rho+1} dy\right\}. \tag{35}$$

To obtain a tight error bound, we have to minimize $\bar{P}_e$ w.r.t. ρ, $Q(x \mid \hat{z})$, and $\mathcal{R}(\hat{z})$. Since $Q(x \mid \hat{z})$, $\mathcal{R}(\hat{z})$ and ρ are independent of each other, the optimization problem is uncoupled.

1) Optimization w.r.t. ρ: Minimization of $\bar{P}_e$ is equivalent to maximization of $\bar{E}_r(\rho, Q^*, \mathcal{R})$. Define the *optimal error exponent* $\bar{E}_r^*(Q^*, \mathcal{R})$ as

$$\bar{E}_r^*(Q^*, \mathcal{R}) = \max_\rho \bar{E}_r(\rho, Q^*, \mathcal{R}) = \bar{E}_r(\rho^*, Q^*, \mathcal{R}). \tag{36}$$

We first introduce the following lemma.

Lemma 3: For any $\mathcal{R}(\hat{z})$, $\partial^2\bar{E}_r/\partial\rho^2 < 0$ for all ρ.

Proof: Refer to Appendix D-1. □

The result of the optimization is summarized in Lemma 4.

Lemma 4: For any $\mathcal{R}(\hat{z})$, let ρ_1 be the roots of the equation

$$\int_0^\infty\int_0^\infty f(c,\hat{z})2^{-[E(\rho,Q,c,\hat{z})-\rho\mathcal{R}(\hat{z})]}\left(\frac{\partial E}{\partial\rho} - \mathcal{R}(\hat{z})\right)dc\,d\hat{z} = 0. \tag{37}$$

The optimal parameter ρ^* is given by

$$\rho^* = \begin{cases}\rho_1, & \rho_1 \in [0,1]\\ 1, & \rho_1 > 1.\end{cases} \tag{38}$$

Proof: The necessary condition for ρ_1 to be the optimal parameter is $(\partial\bar{E}_r(\rho, Q^*, \mathcal{R}))/\partial\rho|_{\rho_1} = 0$. By Lemma 3, $\partial^2\bar{E}_r/\partial\rho^2 < 0$, and hence, the stationary point obtained, ρ_1, corresponds to the absolute maximum point. Furthermore, $\partial\bar{E}_r/\partial\rho$ is a decreasing function of ρ. Suppose $\rho_1 > \rho_2$, we have $\partial\bar{E}_r/\partial\rho|_{\rho_2} > \partial\bar{E}_r/\partial\rho|_{\rho_1} = 0$. Therefore, $\bar{E}_r$ is an increasing function w.r.t. ρ for $\rho < \rho_1$. Since the error bound in (33) is valid only when $\rho \in [0,1]$, the optimal parameter ρ^* is given by $\rho^* = 1$ if $\rho_1 > 1$. □

Collecting the results from the two lemmas, we have the following theorem.

Theorem 2: For any $\mathcal{R}(\hat{z})$, if the average code rate $\bar{\mathcal{R}} \stackrel{\text{def}}{=} \int_0^\infty \mathcal{R}(\hat{z})f(\hat{z})\,d\hat{z}$, is less than or equal to $I(\vec{X}; \vec{Y} \mid C, \hat{Z})$, then $\lim_{N\to\infty}\bar{P}_e = 0$.

Proof: Refer to Appendix D-2. □

2) Optimization w.r.t. $Q(x \mid \hat{z})$: Similar to Section III-A-1, take $Q(x \mid \hat{z})$ as the capacity achieving distribution $Q^*(x)$, which is a Gaussian distribution with variance $\sigma_X^2 = P_0/2W$. The remaining problem is to minimize $\bar{P}_e$ w.r.t. $\mathcal{R}(\hat{z})$.

3) Optimization w.r.t. $\mathcal{R}(\hat{z})$: In this section, we minimize $\bar{P}_e$ w.r.t. $\mathcal{R}(\hat{z})$. Using a normalized rate control, $\mathcal{R}(\hat{z}) = \bar{\mathcal{R}}r(\hat{z})$, we find the optimal normalized rate control function, $r(\hat{z})$. For any given $\bar{\mathcal{R}}$, take $\rho = \rho^*$ as in (37) and $Q(x) = Q^*(x)$. Unlike Section III-A3, a general expression for the optimal rate control function for both negligible and nonnegligible feedback delays is derived.

Using the normalized rate control, $\mathcal{R}(\hat{z}) = \bar{\mathcal{R}}r(\hat{z})$, $\bar{P}_e$ in (33) is given by

$$\bar{P}_e \le \left\{\int_{\hat{z}_1}\int_{c_1} f(c_1,\hat{z}_1)2^{-\left[E(\rho^*,Q^*,c_1)-\rho\bar{\mathcal{R}}r_1(\hat{z}_1)\right]}\,dc_1\,d\hat{z}_1\right\}\cdots \left\{\int_{\hat{z}_N}\int_{c_N} f(c_N,\hat{z}_N)2^{-\left[E(\rho^*,Q^*,c_N)-\rho\bar{\mathcal{R}}r_N(\hat{z}_N)\right]}dc_N\,d\hat{z}_N\right\}. \tag{39}$$

The optimization problem is equivalent to choosing $\{r_1, r_2, \cdots, r_N\}$ that minimizes $\bar{P}_e$ in (39) under the constraints

$$\int_{\hat{z}_i} f(\hat{z}_i)r_i(\hat{z}_i)d\hat{z}_i = 1 \qquad \forall i \in [1,N] \tag{40}$$

$$r_i(\hat{z}_i) \ge 0 \qquad \forall i \in [1,N]. \tag{41}$$

By the Calculus of Variations, it is shown in Appendix E that the optimal control rule is given by

$$r(\hat{z}) = \begin{cases}\frac{1}{\rho^*\bar{\mathcal{R}}}[G(\rho^*, Q^*, \hat{z}_l) - G(\rho^*, Q^*, \hat{z})], & \hat{z} \ge \hat{z}_l\\ 0, & 0 \le \hat{z} < \hat{z}_l\end{cases} \tag{42}$$

where $G(\rho^*, Q^*, \hat{z})$ is given by

$$G(\rho^*, Q^*, \hat{z}) = \log_2\left\{\int_0^\infty f(c \mid \hat{z})2^{-E(\rho^*,Q^*,c)}\,dc\right\} \tag{43}$$

and $\hat{z}_l$ is given by the solution of

$$\frac{1}{\rho^*\bar{\mathcal{R}}}\left\{G(\rho^*, Q^*, \hat{z}_l)\int_{\hat{z}_l}^\infty f(\hat{z})\,d\hat{z} - \int_{\hat{z}_l}^\infty G(\rho^*, Q^*, \hat{z})f(\hat{z})\,d\hat{z}\right\} = 1. \tag{44}$$

Intuitively, at small feedback delay, $r(\hat{z})$ is an increasing function of $\hat{z}$ and the control law implies that more information bits per symbol should be carried if the predicted state $\hat{z}$ is good. On the other hand, at large feedback delay, $\hat{z}$ and c are independent and $r(\hat{z})$ tends to be independent of $\hat{z}$, suggesting that fixed-rate control will be optimal if the predicted state is not accurate.

4) Overall Result of Optimizations for Variable Input Scheme: Given an average code rate $\bar{\mathcal{R}}$, take ρ^*, Q^*, and $r(\hat{z})$ to be the optimal parameter, the capacity achieving distribution, and the optimal rate control law, respectively. Let $R_0 = I(\vec{X}; \vec{Y} \mid C, \hat{Z})$ and $R_1 > 0$ be the solution of the equation

$$\int_0^\infty\int_0^\infty f(c,\hat{z})2^{-[E(\rho,Q,c,\hat{z})-R_1r(\hat{z})]} \times\left(\left.\frac{\partial E}{\partial\rho}\right|_{\rho=1} - R_1r(\hat{z})\right)dc\,d\hat{z} = 0. \tag{45}$$

The optimal error exponent $\bar{E}_r^*(Q^*, \bar{\mathcal{R}}, \Delta)$ at a feedback delay Δ is given by the following.

i) $R_1 \ge \bar{\mathcal{R}} \ge 0$: Using the rate control rule in (42) with $\rho^* = 1$, the solution is given by

$$\bar{E}_r^*(Q^*, \bar{\mathcal{R}}, \Delta) = \bar{E}_r(1, Q^*, \bar{\mathcal{R}}, \Delta). \tag{46}$$

ii) $R_0 \geq \bar{\mathcal{R}} > R_1$: Using the rate control rule in (42), the solution is obtained by solving the following system of three nonlinear simultaneous equations[7] numerically

$$\frac{1}{\rho^*\bar{\mathcal{R}}}\left\{G(\rho^*, Q^*, \hat{z}_l)\int_{\hat{z}_l}^{\infty} f(\hat{z})d\hat{z} - \int_{\hat{z}_l}^{\infty} G(\rho^*, Q^*, \hat{z})f(\hat{z})\,d\hat{z}\right\} = 1$$
$$\int_0^{\infty}\int_0^{\infty} f(c,\hat{z})2^{-[E(\rho,Q,c,\hat{z})-\rho\bar{\mathcal{R}}r(\hat{z})]}\left(\frac{\partial E}{\partial \rho} - \bar{\mathcal{R}}r(\hat{z})\right)dc\,d\hat{z} = 0$$
$$\bar{E}_r^*(Q^*, \bar{\mathcal{R}}, \Delta) = \bar{E}_r(\rho^*, Q^*, \bar{\mathcal{R}}, \Delta). \quad (47)$$

IV. Channel Capacity for Variable Rate Schemes

In this section, we derive a general expression for channel capacity of Rayleigh fading channel with variable rate transmission. Channel capacity is defined as follows.

Definition 1: A channel is said to have a channel capacity $\mathcal{C} > 0$, if

(i) for every $\epsilon > 0$ and $\delta > 0$, $\exists$ a channel code of rate[8] $\mathcal{R} = \mathcal{C} - \epsilon$ with block length N such that the error probability is bounded above by δ $\forall N > n_0$ for some $n_0 > 0$.

(ii) for every $\epsilon > 0$, all codes with rate $\mathcal{R} = \mathcal{C} + \epsilon$ cannot have asymptotically zero error probability as $N \to \infty$.

Lemma 5 (Converse): The channel capacity $\mathcal{C}$ (in bits/symbol) of a Rayleigh fading channel, with feedback of channel states to the transmitter using variable rate transmission, is upper bounded by $\mathcal{C}_0$

$$\mathcal{C} \leq \mathcal{C}_0 \stackrel{\text{def}}{=} \max_{Q(\vec{X}|\hat{Z})} I(\vec{X};\vec{Y} \mid C\hat{Z}). \quad (48)$$

Proof: Refer to Appendix C-3. □

Lemma 6 (Achievability): The channel capacity $\mathcal{C}$ (in bits/symbol) of a Rayleigh fading channel, with feedback of channel states to the transmitter using variable rate transmission, is lower bounded by $\mathcal{C}_0$.

Proof: Assume that $\mathcal{C} < \mathcal{C}_0$ and let $\mathcal{C} < \mathcal{R} < \mathcal{C}_0$. For the *variable duration* scheme, by Theorem 1, there is at least a code of rate $\mathcal{R}$ that has asymptotically zero error probability as $N \to \infty$. However, this contradicts the definition of channel capacity which states that no such code exists with $\mathcal{R} > \mathcal{C}$. For the *variable input* scheme, by Theorem 2, there is at least a code of average rate $\bar{\mathcal{R}}$ that has asymtopically zero error probability as $N \to \infty$. However, this contradicts the definition of channel capacity which states that no such code exists with $\bar{\mathcal{R}} > \mathcal{C}$. Hence, the result follows. □

Combining the above two lemmas, we have the following theorem.

Theorem 3: For any symbol duration control law $T(\hat{z})$ or rate control law $\mathcal{R}(\hat{z})$, the channel capacity of a Rayleigh fading channel with feedback of channel states to the transmitter using variable rate transmission is equal to $\mathcal{C}_0$ in (48).

[7]The unknowns are ρ^*, $\hat{z}_l$, and $\bar{E}_r^*$.

[8]Code rate is expressed as number of information bits per channel symbol and is given by $\mathcal{R}$ in our system.

Using the capacity achieving distribution Q^*, the mutual information is given by

$$I(X_{i,j}; Y_{i,j} \mid C_i\hat{Z}_i) = \frac{1}{2}\int_c\int_{\hat{z}} \log_2\left(1 + \frac{c^2 P_0}{W\eta_0}\right) f(c \mid \hat{z})\mathcal{P}(\hat{z})\,d\hat{z}\,dc. \quad (49)$$

For the *variable duration* scheme, the feedback channel capacity $\mathcal{C}$ (in bits/sec) becomes

$$\begin{aligned}\mathcal{C} &= WE_f \int_c\int_{\hat{z}} T(\hat{z})\log_2\left(1 + \frac{c^2P_0}{W\eta_0}\right) \\ &\quad\times f(c \mid \hat{z})\left[\frac{1}{E_f}\frac{f(\hat{z})}{T(\hat{z})}\right]d\hat{z}\,dc \\ &= W\int_c f(c)\log_2\left(1 + \frac{c^2P_0}{W\eta_0}\right)\int_{\hat{z}} f(\hat{z} \mid c)\,d\hat{z}\,dc \\ &= W\int_c f(c)\log_2\left(1 + \frac{c^2P_0}{W\eta_0}\right)dc \\ &= \mathcal{C}_{(nfb)}\end{aligned} \quad (50)$$

where $\mathcal{C}_{(nfb)}$ is the fixed-duration channel capacity without feedback. For the *variable input* scheme, $\mathcal{C}$ (in bits/sec) becomes $(I(\vec{X};\vec{Y} \mid C\hat{Z}))/T_s$, which is equal to the fixed-rate channel capacity, $\mathcal{C}_{(nfb)}$ as well. Hence, the variable rate schemes cannot increase the channel capacity of Rayleigh fading channels.

V. Results and Discussions

In a microcellular environment at 2 GHz with mobiles moving at a maximum speed of 75 km/h, the coherence time is around 1 ms and the coherence bandwidth is around 2 MHz. We choose the symbol rate E_f to be 40 ksym/s and the system bandwidth to be 800 kHz. Hence, the system bandwidth is smaller than the coherence bandwidth and the average symbol duration ($1/E_f = 0.025$ ms) is much smaller than the coherence time. These justify the flat-fading assumption made in the channel model. Since $W = 800$ kHz, $T_l = 5\ \mu$s is sufficient to ensure $2WT_i > 2$. For a fixed-duration system, the symbol duration $T(\hat{z})$ is constant and is equal to $1/E_f$. Hence, $T_s = 1/E_f$ is taken to be the reference symbol duration. The channel normalized fading rate is $f_dT_s = 2.5 \times 10^{-3}$.

A. Variable Rate Channel Capacity

It is shown in Section IV that channel capacity is not increased by variable rate control. This is intuitively correct since in a large block, we have either the total block duration approaches a constant value of N/E_f [refer to (13)] for the *variable duration* scheme or the total number of information bits transmitted in a large block approaches the average code rate, $\bar{\mathcal{R}}$ for the *variable input* scheme. Hence, there is no difference with fixed-rate coding schemes asymptotically. The channel capacity in the example is equal to 616 kb/s at reference SNR, $E_b/\eta_0 = 3$ dB.

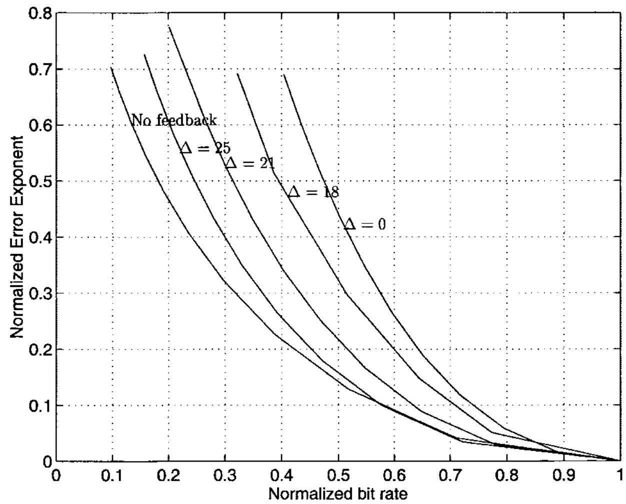


Fig. 4. Error exponent of *variable duration* scheme for various feedback delays. Delays are expressed as number of reference symbols.

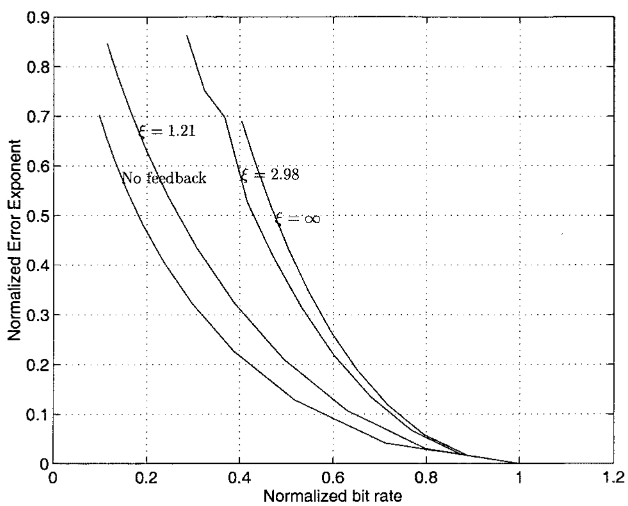


Fig. 5. Error exponent of *variable duration* scheme for various peak time constraints.

B. Variable Rate Error Exponent

Although channel capacity is not increased by variable rate channel coding, the error exponent is increased significantly compared with the fixed-duration error exponent. For the *variable duration* scheme from Fig. 4, the improvement in error exponent is three times the fixed-duration case at $\bar{\mathcal{R}}_b = 0.5\mathcal{C} = 308$ kb/s under ideal situations ($\Delta = 0$, $T_p = \infty$, $T_l = 5\ \mu$s). The performance improvement is degraded to 2.1 times and 1.13 times if the feedback delay is 18 symbols and 25 symbols, respectively. The effect of peak time constraint T_p is shown in Fig. 5. Define the ratio of peak to average symbol duration ξ as $\xi \stackrel{\text{def}}{=} T_p/(\mathcal{E}[T]) = E_f T_p$. With $\xi = 2.98$ and $\xi = 1.21$, the improvement in the error exponent is 2.46 and 1.40 times, respectively.

For the *variable input* scheme from Fig. 6, the improvement in error exponent is 2.67 times at $\bar{\mathcal{R}}_b = 0.5\mathcal{C} = 308$ kb/s with negligible feedback delay compared with fixed-rate code. The performance is degraded to 2.27 times and 1.73 times if the feedback delay is 18 symbols and 23 symbols respectively. At a feedback delay of 100 symbols, error exponent of *variable input* scheme approaches the fixed-rate error exponent because the optimal rate control rule would be a fixed rate control at such large feedback delay. The bandwidth expansion used in the above calculation is $W/\bar{\mathcal{R}}_b = 1.3$. Note that a 2-time increase in error exponent means a 2-time reduction in coding complexity (e.g. block length N) to achieve the same error probability.

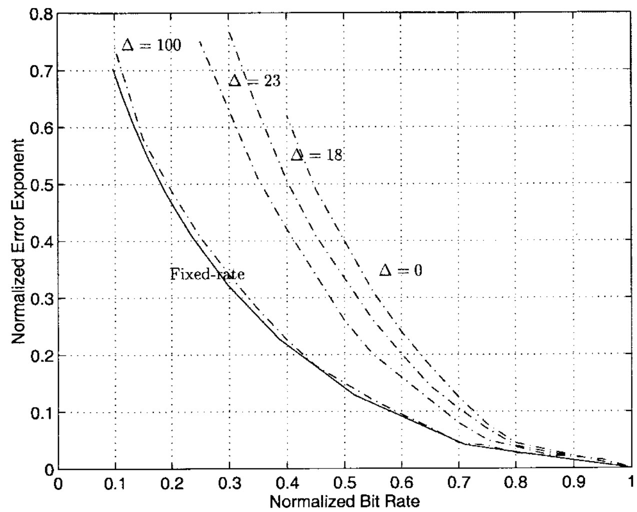


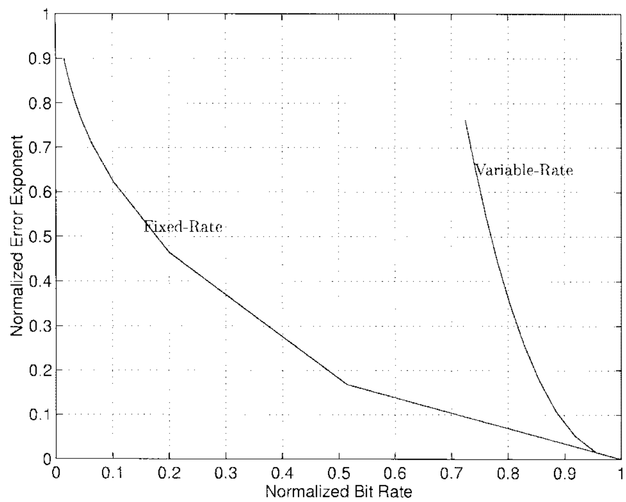
Fig. 6. Error exponent of *variable input* scheme atvarious feedback delays Δ (in number of symbols).

C. Bandwidth Expansion Consideration

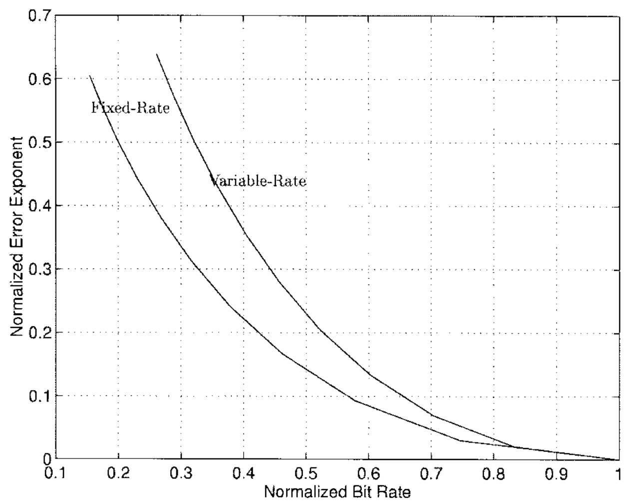
We consider two extreme cases, a bandwidth expansion of 0.25 which models TDMA systems and a bandwidth expansion of 20 which models CDMA systems. Error exponents of the *variable duration* scheme for small and large bandwidth expansion systems are shown in Fig. 7(a) and (b). For the system with small bandwidth expansion (TDMA), there is a significant 5.62-time increase in error exponent at $\bar{\mathcal{R}}_b = 0.8\mathcal{C}$. For the system with large bandwidth expansion (CDMA), there is only a 1.5-time improvement in error exponent relative to fixed-duration scheme at the same $\bar{\mathcal{R}}_b$.

The error exponents for the *variable input* scheme with small (0.25) and large bandwidth expansion (20) systems are shown in Fig. 8(a) and (b). For systems with small bandwidth expansion (TDMA), there is a significant 5-time increase in error exponent at $\bar{\mathcal{R}}_b = 0.8\mathcal{C}$. For systems with large bandwidth expansion (CDMA), there is only a 1.1-time improvement in error exponent relative to fixed-rate schemes at the same $\bar{\mathcal{R}}_b$.

Therefore, variable rate channel coding is more effective in high bandwidth-efficiency systems. This means that only a limited gain can be achieved in high bandwidth-expansion systems when very powerful capacity achieving codes are used as the component codes in the construction of variable rate adaptive codes. However, a significant gain should be expected in high bandwidth-efficiency systems even when very powerful component codes are used.

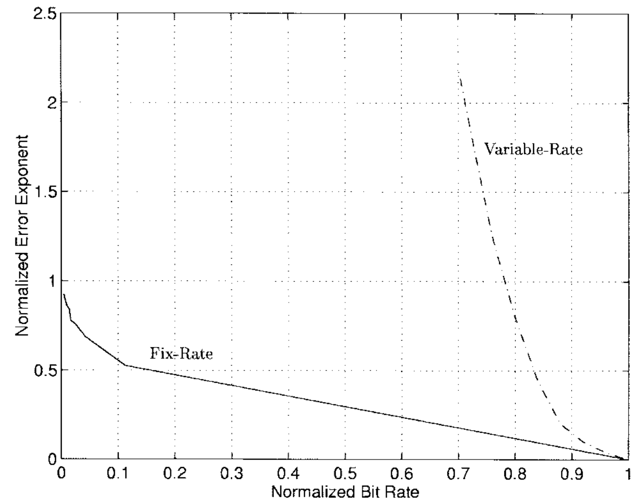


(a)

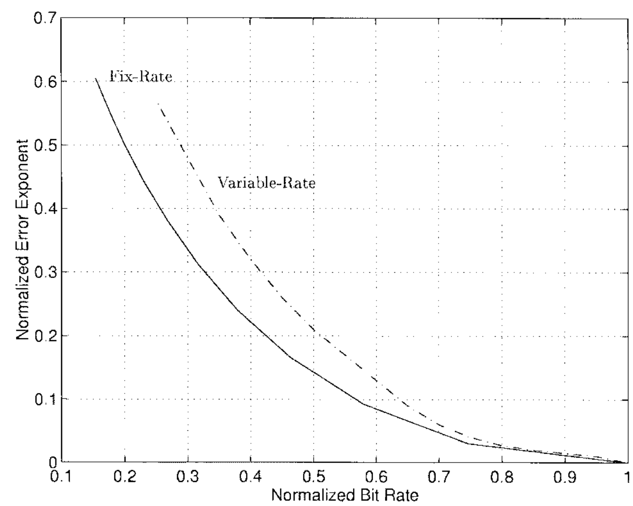


(b)

Fig. 7. Error exponents of *variable duration* scheme in large and small bandwidth expansion.



(a)



(b)

Fig. 8. Error exponents of *variable input* scheme in large and small bandwidth expansion.

VI. Summary

In this paper, we have evaluated the channel capacity and the error exponent of variable rate Rayleigh fading channel using *variable duration* and *variable input* schemes. Optimal symbol duration control law and optimal input rate control law are derived taking into account of feedback delay. Performance degradation w.r.t. feedback delay is also investigated. We found that channel capacity was not increased by variable rate coding schemes for any control law. On the other hand, there was a significant increase in error exponent for both schemes. This means that less complex codes can be found to achieve the same $\bar{P}_e$ using variable rate adaptive coding. Hence, instead of aiming at maximizing the channel capacity by previous approaches, we should aim at maximizing the error exponent with variable rate adaptive channel coding.

For the dependence of the improvement on bandwidth expansion, we found that improvement was limited at large bandwidth expansion. On the contrary, significant gain resulted when bandwidth expansion was small. This suggests that variable rate channel coding schemes have limited intrinsic gains in CDMA systems compared with TDMA systems.

Appendix A
Induced Probability Density on $\vec{\hat{\mathbf{Z}}}_N$

A. One Dimension

We first prove (4) in Section II.

Proof: Define the induced probability density $\mathcal{P}(\hat{z})$ as

$$\mathcal{P}(\hat{z})\delta \stackrel{\text{def}}{=} \lim_{\mathcal{N}\to\infty} \frac{\mathcal{N}_{\hat{z}}}{\mathcal{N}}$$

where $\mathcal{N}_{\hat{z}}$ is the number of symbols with $\hat{Z} \in [\hat{z}, \hat{z}+\delta]$ in a sequence of $\mathcal{N}$ symbols, and δ is a small increment in $\hat{z}$.

Observe $\hat{Z}(t)$ over a long time duration $\mathcal{T}_0$. Suppose $\hat{Z}(t)$ is an ergodic random process, we have

$$\lim_{\mathcal{T}_0\to\infty} \frac{\mathcal{T}_{\hat{z}}}{\mathcal{T}_0} = f(\hat{z})\delta$$

where $\mathcal{T}_{\hat{z}}$ is the total time that $\hat{Z}(t) \in [\hat{z}, \hat{z}+\delta]$ and $f(\hat{z})$ is the fading density. Hence, $\mathcal{N}_{\hat{z}} = \mathcal{T}_{\hat{z}}/T(\hat{z}) = \mathcal{T}_0 f(\hat{z})\delta/T(\hat{z})$

where $T(\hat{z})$ is the symbol duration (a function of $\hat{z}$) and $\mathcal{N} = \int_{\hat{z}} \mathcal{N}_{\hat{z}}\, d\hat{z} = \mathcal{T}_0 \int_{\hat{z}} f(\hat{z})/T(\hat{z})\, d\hat{z}$. Therefore,

$$\mathcal{P}(\hat{z}) = \frac{f(\hat{z})}{E_f T(\hat{z})} \tag{51}$$

where

$$E_f = \int_{\hat{z}} \frac{1}{T(\hat{z})} f(\hat{z})\, d\hat{z}. \tag{52}$$

□

We shall show that E_f is the average symbol rate. Suppose we observe $\hat{Z}(t)$ over a very long time interval $\mathcal{T}_0$, then $\mathcal{T}_{\hat{z}} = f(\hat{z})\delta\mathcal{T}_0$ and there are $\mathcal{T}_{\hat{z}}/T(\hat{z})$ symbols during this time duration. By definition, the average symbol rate is the number of symbols transmitted per unit time and is given by $\int_{\hat{z}} \mathcal{T}_{\hat{z}}/T(\hat{z})\, d\hat{z}$ which is equal to E_f.

We shall illustrate the use of this induced probability density in the following example.

Example 1: Suppose the symbol error probability $P_e(\hat{z})$ is conditioned on $\hat{z}$. To obtain the unconditional error probability, we have to use the *induced density* in the integration given by

$$\bar{P}_e = \int_{\hat{z}} P_e(\hat{z})\mathcal{P}(\hat{z})\, d\hat{z}.$$

Proof: The average number of symbol errors given that $\hat{Z} \in [\hat{z}, \hat{z}+\delta]$ is given by $\mathcal{N}_{\hat{z}} P_e(\hat{z})$. By definition, the average error probability is the total number of symbol errors divided by the total number of symbols transmitted and is given by

$$\bar{P}_e = \lim_{\mathcal{N}\to\infty} \frac{\int_{\hat{z}} \mathcal{N}_{\hat{z}} P_e(\hat{z})\, d\hat{z}}{\mathcal{N}} = \int_{\hat{z}} \mathcal{P}(\hat{z}) P_e(\hat{z})\, d\hat{z}.$$

□

B. N Dimensions

For the ideal interleaved channel, $\hat{z}_i$ and $\hat{z}_j$ are independent. Hence, the N-dimensional induced probability density on $\vec{\hat{\mathbf{Z}}}_N$ is the product of N one-dimensional induced densities

$$\mathcal{P}_N(\hat{z}_1, \hat{z}_2, \cdots, \hat{z}_N) = \frac{1}{E_f^N} \frac{f(\hat{z}_1) f(\hat{z}_2)\cdots f(\hat{z}_N)}{T(\hat{z}_1)T(\hat{z}_2)\cdots T(\hat{z}_N)}. \tag{53}$$

Extending Example 1 to the N-dimensional and ideal interleaving case, the unconditioned error probability is given by

$$\begin{aligned}\bar{P}_e &= \lim_{\mathcal{N}\to\infty} \frac{\int_{\vec{\hat{\mathbf{z}}}_N} \mathcal{N}_{\hat{z}_1}\cdots\mathcal{N}_{\hat{z}_N} P_e(\hat{z}_1,\cdots,\hat{z}_N)\, d\vec{\hat{\mathbf{z}}}_N}{\mathcal{N}^N} \\ &= \int_{\vec{\hat{\mathbf{z}}}_N} \mathcal{P}_N(\hat{z}_1,\cdots,\hat{z}_N) P_e(\hat{z}_1,\cdots,\hat{z}_N)\, d\vec{\hat{\mathbf{z}}}_N.\end{aligned}$$

Appendix B
Optimal Control Rule for *Variable Duration* Scheme

The problem is to choose $\{T_1(\hat{z}_1), \cdots, T_N(\hat{z}_N)\}$ that minimize $\bar{P}_e$ under the constraints of (24), (25), and (26). We form the ith Lagrange multiplier as

$$\begin{aligned}L_i(T_i) = \frac{2^{\frac{N\rho^*\mathcal{R}_b}{E_f}}}{E_f^N} \prod_{\substack{j=1\\ j\neq i}}^{N} \left\{ \int_0^\infty \frac{f(\hat{z}_j)}{T_j} G(\rho^*, \hat{z}_j)^{2WT_j}\, d\hat{z}_j \right\} \\ \times \frac{f(\hat{z}_i)}{T_i} G(\rho^*, \hat{z}_i)^{2WT_i} - \beta \frac{1}{E_f} \frac{f(\hat{z}_i)}{T_j}\end{aligned} \tag{54}$$

where

$$G(\rho^*, \hat{z}_i) = \left[\int_y \left(\int_x Q^*(x) p_0(y \mid x\hat{z}_i)^{\frac{1}{\rho^*+1}}\, dx \right)^{\rho^*+1} dy \right].$$

By the Calculus of Variations [24], the necessary condition for $T_i(\hat{z}_i)$ to be the optimal control is $\frac{\partial L_i}{\partial T_i} = 0$ for all $i \in [1, N]$. This implies

$$G(\rho^*, \hat{z}_i)^{2WT_i} \log_2(G(\rho^*, \hat{z}_i)^{2WT_i}) = K_0$$

where K_0 is a constant independent of $\hat{z}_i$. This means that $G(\rho^*, \hat{z}_i)^{2WT_i}$ must be a constant for every i. Hence, the optimal $T(\hat{z})$ is given by

$$T(\hat{z}) = \frac{K}{E_f \log_2 \left[\int_y \left[\int_x Q^*(x) p_0(y \mid x\hat{z})^{\frac{1}{\rho^*+1}}\, dx \right]^{\rho^*+1} dy \right]}$$

where K is determined by substituting back into the constraints (24)–(26).

Appendix C
Proof of Lemmas for *Variable Duration* Scheme

A. Proof of Lemma 1

Proof: From (16) and (7), $\bar{E}(\rho, Q)$ can be expressed as

$$\bar{E}(\rho, Q) = -E_f \log_2 \left[\int_c \int_{\hat{z}} 2^{-E_N(\rho, Q, c, \hat{z})} \frac{f(c, \hat{z})}{T(\hat{z})}\, dc\, d\hat{z} \right]. \tag{55}$$

(i)

$$\frac{\partial \bar{E}}{\partial \rho} = E_f \frac{\int_c \int_{\hat{z}} \frac{f(c,\hat{z})}{T(\hat{z})} 2^{-[E_N(\rho,Q,c,\hat{z})]} \frac{\partial E_N}{\partial \rho} dc\, d\hat{z}}{\int_c \int_{\hat{z}} \frac{f(c,\hat{z})}{T(\hat{z})} 2^{-[E_N(\rho,Q,c,\hat{z})]}\, dc\, d\hat{z}}.$$

Since $\partial E_N/\partial\rho > 0$ [22], [1], we have $\partial\bar{E}/\partial\rho > 0$. Since

$$\left.\frac{\partial E_N}{\partial \rho}\right|_{\rho=0} = I(\vec{X}; \vec{Y} \mid C = c, \hat{Z} = \hat{z})$$

and $E_N(0, Q, c, \hat{z}) = 0$, we have

$$\begin{aligned}\left.\frac{\partial \bar{E}}{\partial \rho}\right|_{\rho=0} &= E_f \int_c \int_{\hat{z}} \frac{f(c,\hat{z})}{T(\hat{z})E_f} I(\vec{X}; \vec{Y} \mid C = c, \hat{Z} = \hat{z})\, dc\, d\hat{z} \\ &= E_f I(\vec{X}; \vec{Y} \mid C\hat{Z}).\end{aligned}$$

(ii) Since [22], [1]

$$\begin{aligned}&E_N(\alpha_1\rho_1 + \alpha_2\rho_2, Q, c, \hat{z}) \\ &\qquad > \alpha_1 E_N(\rho_1, Q, c, \hat{z}) + \alpha_2 E_N(\rho_2, Q, c, \hat{z})\end{aligned}$$

for any $\alpha_1+\alpha_2=1$, we have

$$\begin{aligned}
&\bar{E}(\alpha_1\rho_1+\alpha_2\rho_2,Q)\\
&\quad > -E_f\log_2\left[\int_c\int_{\hat z}\mathcal{P}(\hat z)f(c\mid\hat z)\right.\\
&\qquad\left.\times\, 2^{-[\alpha_1E_N(\rho_1)+\alpha_2E_N(\rho_2)]}\,dc\,d\hat z\right]\\
&\quad = -E_f\log_2\left\{\int_c\int_{\hat z}\left[\mathcal{P}(\hat z)f(c\mid\hat z)2^{-[E_N(\rho_1)]}\right]^{\alpha_1}\right.\\
&\qquad\left.\times\left[\mathcal{P}(\hat z)f(c\mid\hat z)2^{-[E_N(\rho_2)]}\right]^{\alpha_2}dc\,d\hat z\right\}\\
&\quad \ge -E_f\log_2\left\{\left[\int_c\int_{\hat z}\mathcal{P}(\hat z)f(c\mid\hat z)2^{-[E_N(\rho_1)]}\,dc\,d\hat z\right]^{\alpha_1}\right.\\
&\qquad\left.\times\left[\int_c\int_{\hat z}\mathcal{P}(\hat z)f(c\mid\hat z)2^{-[E_N(\rho_2)]}\,dc\,d\hat z\right]^{\alpha_2}\right\}\\
&\quad = \alpha_1\bar{E}(\rho_1,Q)+\alpha_2\bar{E}(\rho_2,Q)
\end{aligned}$$

where the last inequality follows from the Holder's inequality. Hence, $\bar{E}(\rho,Q)$ is a convex function in ρ and the result follows. □

B. Proof of Lemma 2

Proof: The necessary condition for $\rho=\rho^*$ to maximize $\left[\bar{E}(\rho,Q)-\rho\bar{\mathcal{R}}_b\right]$ is $\frac{\partial[\bar{E}-\rho\bar{\mathcal{R}}_b]}{\partial\rho}\big|_{\rho=\rho^*}=0$. This gives

$$\left.\frac{\partial\bar{E}}{\partial\rho}\right|_{\rho=\rho^*}=\bar{\mathcal{R}}_b. \tag{56}$$

By Lemma 1(ii), we have $\frac{\partial^2\bar{E}}{\partial\rho^2}<0$. This verifies that the stationary point obtained, ρ^*, corresponds to the absolute maximum and $\bar{\mathcal{R}}_b$ in (56) is a strictly decreasing function of ρ^*. Hence, as ρ^* increases from 0 to 1, $\bar{\mathcal{R}}_b$ decreases from R_0 to R_1 and $\bar{E}_r$ maximizes at ρ^*. For $\bar{\mathcal{R}}_b<R_0$, $\rho^*>1$. Since $\frac{\partial\bar{E}}{\partial\rho}$ is a decreasing function of ρ, we have

$$\left.\frac{\partial\bar{E}}{\partial\rho}\right|_{\rho=\rho_1}>\left.\frac{\partial\bar{E}}{\partial\rho}\right|_{\rho=\rho^*}=\bar{\mathcal{R}}_b>0$$

for $0\le\rho_1<\rho^*$. Since the error bound in (6) is valid only when $\rho\in[0,1]$, $\bar{E}_r$ maximizes at $\rho_1=1$. □

C. Proof of Lemma 5

Proof: The estimated index $\hat\Omega$ is given by

$$\hat\Omega=g(\vec{Y}_1,\cdots,\vec{Y}_N,\vec{\mathbf{C}}_N,\vec{\hat{\mathbf{Z}}}_N)$$

where $g(\vec{Y}_1,\cdots,\vec{Y}_N,\vec{\mathbf{C}}_N,\vec{\hat{\mathbf{Z}}}_N)$ is a general decoding function. The average codeword error probability $\bar{P}_e$ is given by

$$\bar{P}_e=\Pr[\Omega\ne\hat\Omega]. \tag{57}$$

By Fano's inequality

$$H(\Omega\mid\vec{Y}_1,\cdots,\vec{Y}_N,\vec{\mathbf{C}}_N,\vec{\hat{\mathbf{Z}}}_N)\le 1+\bar{P}_eN\mathcal{R}. \tag{58}$$

Since $I(\Omega;\vec{Y}_1,\cdots,\vec{Y}_N,\vec{\mathbf{C}}_N,\vec{\hat{\mathbf{Z}}}_N)=H(\Omega)-H(\Omega\mid\vec{Y}_1,\cdots,\vec{Y}_N,\vec{\mathbf{C}}_N,\vec{\hat{\mathbf{Z}}}_N)$, $H(\Omega)=N\mathcal{R}$ because of equiprobable input and

$$\begin{aligned}
&I(\Omega;\vec{Y}_1,\cdots,\vec{Y}_N,\vec{\mathbf{C}}_N,\vec{\hat{\mathbf{Z}}}_N)\\
&\quad=I(\Omega;\vec{Y}_1,\cdots,\vec{Y}_N\mid\vec{\mathbf{C}}_N,\vec{\hat{\mathbf{Z}}}_N)+I(\Omega;\vec{\mathbf{C}}_N,\vec{\hat{\mathbf{Z}}}_N)\\
&\quad=I(\Omega;\vec{Y}_1,\cdots,\vec{Y}_N\mid\vec{\mathbf{C}}_N\vec{\hat{\mathbf{Z}}}_N)\quad(I(\Omega;\vec{\mathbf{C}}_N,\vec{\hat{\mathbf{Z}}}_N)=0)
\end{aligned}$$

we have

$$\mathcal{R}\le\frac{1}{N}I(\Omega\vec{Y}_1,\cdots,\vec{Y}_N\mid\vec{\mathbf{C}}_N,\vec{\hat{\mathbf{Z}}}_N)+\epsilon \tag{59}$$

where $\epsilon=\frac{1}{N}[1+\bar{P}_eN\mathcal{R}]$. Given any particular $\vec{\mathbf{C}}_N=\vec{\mathbf{c}}_N$ and $\vec{\hat{\mathbf{Z}}}_N=\vec{\hat{\mathbf{z}}}_N$, we have $\Omega\to\{\vec{X}_1,\cdots,\vec{X}_N\}\to\{\vec{Y}_1,\cdots,\vec{Y}_N\}$ and by the data processing inequality [18], we have

$$\begin{aligned}
&I(\Omega;\vec{Y}_1,\cdots,\vec{Y}_N\mid\vec{\mathbf{C}}_N=\vec{\mathbf{c}}_N,\vec{\hat{\mathbf{C}}}_N=\vec{\hat{\mathbf{z}}}_N)\\
&\quad\le I(\vec{X}_1,\cdots,\vec{X}_N;\vec{Y}_1,\cdots,\vec{Y}_N\mid\vec{\mathbf{C}}_N=\vec{\mathbf{c}}_N,\vec{\hat{\mathbf{Z}}}_N=\vec{\hat{\mathbf{z}}}_N).
\end{aligned}$$

Hence

$$\begin{aligned}
&I(\Omega;\vec{Y}_1,\cdots,\vec{Y}_N\mid\vec{\mathbf{C}}_N,\vec{\hat{\mathbf{Z}}}_N)\\
&\quad=\int_{\vec{\mathbf{c}}_N,\vec{\hat{\mathbf{z}}}_N}\mathcal{P}_N(\vec{\hat{\mathbf{z}}}_N,\vec{\mathbf{c}}_N)I(\Omega;\vec{Y}_1,\cdots,\vec{Y}_N\mid\\
&\qquad\vec{\mathbf{C}}_N=\vec{\mathbf{c}}_N,\vec{\hat{\mathbf{Z}}}_N=\vec{\hat{\mathbf{z}}}_N)\,d\vec{\mathbf{c}}_N\,d\vec{\hat{\mathbf{z}}}_N\\
&\quad\le\int_{\vec{\mathbf{c}}_N,\vec{\hat{\mathbf{z}}}_N}\mathcal{P}_N(\vec{\hat{\mathbf{z}}}_N,\vec{\mathbf{c}}_N)I(\vec{X}_1,\cdots,\vec{X}_N;\vec{Y}_1,\cdots,\vec{Y}_N\mid\\
&\qquad\vec{\mathbf{C}}_N=\vec{c}_N,\vec{\hat{\mathbf{Z}}}_N=\vec{\hat{\mathbf{z}}}_N)\,d\vec{\mathbf{c}}_N\,d\vec{\hat{\mathbf{z}}}_N\\
&\quad=I(\vec{X}_1,\cdots,\vec{X}_N;\vec{Y}_1,\cdots,\vec{Y}_N\mid\vec{\mathbf{C}}_N,\vec{\hat{\mathbf{Z}}}_N)\\
&\quad=NI(\vec{X};\vec{Y}\mid C,\hat{Z}).
\end{aligned}\tag{60}$$

Substituting (60) into (59) and maximizing w.r.t. $Q(\vec{X}\mid\hat{A})$, we found that if $\lim_{N\to\infty}\bar{P}_e=0$, we have $\epsilon\to 0$ and

$$\mathcal{R}\le\mathcal{C}_0=\max_{Q(\vec{X}\mid\hat{Z})}I(\vec{X};\vec{Y}\mid C,\hat{Z}). \tag{61}$$

Assume that $\mathcal{C}>\mathcal{C}_0$. By definition of channel capacity, $\exists$ a code with rate $\mathcal{R}$ and $\mathcal{C}>\mathcal{R}>\mathcal{C}_0$ such that $\lim_{N\to\infty}\bar{P}_e=0$. However, this contradicts (61) which states that

$$\lim_{N\to\infty}\bar{P}_e=0\Rightarrow\mathcal{R}\le\mathcal{C}_0.$$

Hence, the lemma follows. □

APPENDIX D
PROOF OF LEMMAS AND THEOREM FOR *VARIABLE INPUT* SCHEME

A. Proof of Lemma 3

Proof: Since [22], [1]

$$E(\alpha_1\rho_1+\alpha_2\rho_2,Q,c,\hat z)>\alpha_1E(\rho_1,Q,c,\hat z)+\alpha_2E(\rho_2,Q,c,\hat z)$$

for any $\alpha_1+\alpha_2=1$, we have

$$\begin{aligned}
&\bar{E}_r(\alpha_1\rho_1+\alpha_2\rho_2,Q)\\
&\quad>-\log_2\left[\int_{\hat z}\int_c f(c,\hat z)\right.\\
&\qquad\left.\times\,2^{-\{\alpha_1[E(\rho_1)-\rho_1\mathcal{R}]+\alpha_2[E(\rho_2)-\rho_2\mathcal{R}]\}}dc\,d\hat z\right]\\
&\quad=-\log_2\left\{\int_c\int_{\hat z}\left[f(c,\hat z)2^{-[E(\rho_1)-\rho_1\mathcal{R}]}\right]^{\alpha_1}\right.\\
&\qquad\left.\times\left[f(c,\hat z)2^{-[E(\rho_2)-\rho_2\mathcal{R}]}\right]^{\alpha_2}dc\,d\hat z\right\}\\
&\quad\ge-\log_2\left\{\left[\int_c\int_{\hat c}f(c,\hat z)2^{-[E(\rho_1)-\rho_1\mathcal{R}]}\,dc\,d\hat z\right]^{\alpha_1}\right.\\
&\qquad\left.\times\left[\int_c\int_{\hat z}f(c,\hat z)2^{-[E(\rho_2)-\rho_2\mathcal{R}]}\,dc\,d\hat z\right]^{\alpha_2}\right\}\\
&\quad=\alpha_1\bar{E}_r(\rho_1,Q)+\alpha_2\bar{E}_r(\rho_2,Q)
\end{aligned}$$

where the last inequality follows from the Holder's inequality. Hence, $\bar{E}_r(\rho, Q)$ is a convex function in ρ and the result follows. □

B. Proof of Theorem 2

Proof: Express the code rate $\mathcal{R}(\hat{z})$ as

$$\mathcal{R}(\hat{z}) = \bar{\mathcal{R}} r(\hat{z}) \tag{62}$$

where $\int_0^\infty r(\hat{z}) f(\hat{z})\, d\hat{z} = 1$. Let ρ_1 be the roots of (37). From Lemma 4, if $\rho_1 > 1$, the optimal parameter $\rho^* = 1$ and we have

$$\frac{\partial \bar{E}_r^*}{\partial \bar{\mathcal{R}}} = -\frac{\int_{\hat{z}} \int_c f(c,\hat{z}) r(\hat{z}) 2^{-[E(1,Q,c,\hat{z}) - \bar{\mathcal{R}} r(\hat{z})]}\, dc\, d\hat{z}}{\int_{\hat{z}} \int_c f(c,\hat{z}) 2^{-[E(1,Q,c,\hat{z}) - \bar{\mathcal{R}} r(\hat{z})]}\, dc\, d\hat{z}} < 0$$

if $\rho_1 \in [0,1]$, the optimal parameter, $\rho^* = \rho_1$, and we have

$$\frac{\partial \bar{E}_r^*}{\partial \bar{\mathcal{R}}} = \frac{\partial \rho^*}{\partial \bar{\mathcal{R}}} \frac{\int_c \int_{\hat{z}} f(c,\hat{z}) 2^{-[E(\rho^*,Q,c,\hat{z}) - \rho^* \bar{\mathcal{R}} r(\hat{z})]} \left[\frac{\partial E}{\partial \rho^*} - \bar{\mathcal{R}} r(\hat{z})\right] dc\, d\hat{z}}{\int_{\hat{z}} \int_c f(c,\hat{z}) 2^{-[E(\rho^*,Q,c,\hat{z}) - \rho^* \bar{\mathcal{R}} r(\hat{z})]}\, dc\, d\hat{z}} - \rho^* \frac{\int_{\hat{z}} \int_c f(c,\hat{z}) r(\hat{z}) 2^{-[E(\rho^*,Q,c,\hat{z}) - \rho^* \bar{\mathcal{R}} r(\hat{z})]}\, dc\, d\hat{z}}{\int_{\hat{z}} \int_c f(c,\hat{z}) 2^{-[E(\rho^*,Q,c,\hat{z}) - \rho^* \bar{\mathcal{R}} r(\hat{z})]}\, dc\, d\hat{z}}.$$

From Lemma 4, $\int_c \int_{\hat{z}} f(c,\hat{z}) 2^{-[E(\rho^*,Q,c,\hat{z}) - \rho^* \bar{\mathcal{R}} r(\hat{z})]} [\frac{\partial E}{\partial \rho^*} - \bar{\mathcal{R}} r(\hat{z})]\, dc\, d\hat{z} = 0$. Hence, $\frac{\partial \bar{E}_r^*}{\partial \bar{\mathcal{R}}} < 0$ and $\bar{E}_r^*(Q, \bar{\mathcal{R}})$ is strictly decreasing w.r.t. $\bar{\mathcal{R}}$. At $\rho^* = 0$, from (37), $\bar{\mathcal{R}} = I(\vec{X}; \vec{Y} \mid C\hat{Z})$ and $\bar{E}_r^* = 0$. Therefore, if $0 \le \bar{\mathcal{R}} < I(\vec{X}; \vec{Y} \mid C\hat{Z})$, $\bar{E}_r^* > 0$. Take $N \to \infty$, $\bar{P}_e = 0$. □

APPENDIX E
OPTIMAL CONTROL RULE FOR *VARIABLE INPUT* SCHEME

The optimization problem is to choose $\{\mathcal{R}_1, \cdots, \mathcal{R}_N\}$ that minimizes $\bar{P}_e$ in (39) under the constraints of (40) and (41). The ith Lagrange multiplier is given by

$$L_i(r_i) = \left\{ \prod_{\substack{j=1 \\ j \ne i}}^N \int_0^\infty f(c_j, \hat{z}_j) 2^{-[E(\rho^*, Q^*, c_j) - \rho \bar{\mathcal{R}} r_j(\hat{z}_j)]}\, dc_j\, d\hat{z}_j \right\} \times 2^{G(\rho^*, Q^*, \hat{z}_i) + \rho^* \bar{\mathcal{R}} r_i(\hat{z}_i)} - \beta r_i(\hat{z}_i). \tag{63}$$

By the Calculus of Variations [24], the necessary condition for $r_i(\hat{z}_i)$ to minimize $\bar{P}_e$ under the constraints of (40) and (41) is $\frac{\partial L}{\partial r_i} = 0$ which is given by

$$2^{[G(\rho^*, Q^*, \hat{z}_i) + \rho^* \bar{\mathcal{R}} r_i(\hat{z}_i)]} = \text{ constant.} \tag{64}$$

This gives

$$r_i(\hat{z}_i) = \frac{1}{\rho^* \bar{\mathcal{R}}} [K_0 - G(\rho^*, Q^*, \hat{z}_i)]$$

WLOG, we drop the index i. Since $r(\hat{z})$ is an increasing function of $\hat{z}$ and the constraint (41) requires r to be nonnegative, $r(\hat{z})$ is given by

$$r(\hat{z}) = \begin{cases} \frac{1}{\rho^* \bar{\mathcal{R}}} [G(\rho^*, Q^*, \hat{z}_l) - G(\rho^*, Q^*, \hat{z})], & \hat{z} \ge \hat{z}_l \\ 0, & 0 \le \hat{z} < \hat{z}_l \end{cases}$$

where $\hat{z}_l > 0$ is a constant determined by the constraint (40) and is given by the roots of the equation

$$\frac{1}{\rho^* \bar{\mathcal{R}}} \left\{ G(\rho^*, Q^*, \hat{z}_l) \int_{\hat{z}_l}^\infty f(\hat{z})\, d\hat{z} - \int_{\hat{z}_l}^\infty G(\rho^*, Q^*, \hat{z}) f(\hat{z})\, d\hat{z} \right\} = 1.$$

ACKNOWLEDGMENT

The author would like to thank Dr. S. V. Maric and Dr. M. D. Macleod for their advice in this work.

REFERENCES

[1] S. G. Wilson, *Digital Modulation and Coding*, 1st ed. Englewood Cliffs, NJ: Prentice-Hall, 1996.
[2] J. G. Proakis, *Digital Communications*, 3rd ed. New York: McGraw-Hill, 1995.
[3] S. G. Chua and A. Goldsmith, "Variable-rate variable-power MQAM for fading channels," in *Proc. IEEE VTC*, Apr. 1996, pp. 815–819.
[4] W. T. Webb and R. Steele, "Variable rate QAM for mobile radio," *IEEE Trans. Commun.*, vol. 43, pp. 2223–2230, July 1995.
[5] J. K. Cavers, "Variable rate transmission for Rayleigh fading channel," *IEEE Trans. Commun.*, vol. COM-20, pp. 15–21, Feb. 1972.
[6] S. Sampei, S. Komarki, and N. Morinaga, "Adaptive modulation/TDMA scheme for large capacity personal multi-media communication systems," *IEICE Trans. Commun.*, vol. E-77B, pp. 1096–1103, Sept. 1994.
[7] Y. Kamio, S. Sampei, and H. Sasaoka, "Performance of modulation level controlled adaptive modulation under limited transmission delay time for land mobile communications," presented at the IEEE Vehicular Technology Conf. (VTC'95), July 1995.
[8] S. Alamouti and S. Kallel, "Adaptive trellis-coded multiple-phase-shift keying for Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 42, pp. 2305–2314, June 1994.
[9] K. N. Lau, "Variable rate adaptive modulation for DS-CDMA," *IEEE Trans. Commun.*, vol. 47, pp. 577–589, Apr. 1999.
[10] B. Vucetic, "An adaptive coding scheme for time-varying channels," *IEEE Trans. Commun.*, vol. 39, pp. 653–663, May 1991.
[11] T. Ue, S. Sampei, and N. Morinaga, "Symbol rate and modulation level controlled adaptive modulation/TDMA/TDD for personal communication systems," in *Proc. IEEE VTC'95*, July 1995, pp. 306–310.
[12] H. Matsuoka, S. Sampei, and N. Morinaga, "Adaptive modulation system with punctured convolutional code for high quality personal communication systems," in *Proc. IEEE ICUPC'95*, Nov. 1995, pp. 22–26.
[13] J. Hagenauer, N. Seshadri, and C. W. Sundberg, "Variable rate subband speech coding and matched channel coding for mobile radio channel," in *Proc. IEEE VTC'88*, June 1988, pp. 139–146.
[14] J. Hagenauer, N. Seshadri, and C. W. Sundberg, "The performance of rate-compatible punctured convolutional codes for future digital mobile radio," in *Proc. IEEE VTC'88*, June 1988, pp. 22–29.
[15] M. B. Pursley and S. D. Sandberg, "Variable rate coding for meteor-burst communications," *IEEE Trans. Commun.*, vol. 37, pp. 1105–1112, Nov. 1989.
[16] A. J. Goldsmith, "The capacity of downlink fading channels with variable rate and power," *IEEE Trans. Veh. Technol.*, vol. 46, pp. 569–580, Aug. 1997.
[17] C. Benedetto, Biglieri, Castellani, *Digital Transmission Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1987.
[18] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 1st ed. New York: Wiley, 1991.
[19] K. N. Lau, "Variable rate adaptive coded M-ary orthogonal modulation for DS-CDMA," *Bell Syst. Tech. J.*, Apr. 1999.
[20] K. N. Lau and M. D. Malcolm, "Variable rate adaptive trellis-coded modulation for high bandwidth efficiency applications," in *Proc. IEEE VTC'98*, May 1998, pp. 22–29.
[21] W. C. Jakes, *Mobile Radio Systems*, 1st ed. New York: McGraw-Hill, 1974.
[22] R. G. Gallager, *Information Theory and Reliable Communication*, 2nd ed. New York: Wiley, 1968.
[23] P. Billingsley, *Ergodic Theory and Information*, 1st ed. New York: Wiley, 1965.
[24] L. A. Pars, *An Introduction to the Calculus of Variations*, 1st ed. New York: Heinemann, 1962.

Vincent K. N. Lau (M'92), for photograph and biography, see p. 589 of the April 1999 issue of this TRANSACTIONS.