

A Union Bound on the Error Probability of Binary Codes Over Block-Fading Channels

Salam A. Zummo, *Member, IEEE*, Ping-Cheng Yeh, *Member, IEEE*, and Wayne E. Stark, *Fellow, IEEE*

Abstract—Block-fading is a popular channel model that approximates the behavior of different wireless communication systems. In this paper, a union bound on the error probability of binary-coded systems over block-fading channels is proposed. The bound is based on uniform interleaving of the coded sequence prior to transmission over the channel. The distribution of error bits over the fading blocks is computed. For a specific distribution pattern, the pairwise error probability is derived. Block-fading channels modeled as Rician and Nakagami distributions are studied. We consider coherent receivers with perfect and imperfect channel side information (SI) as well as noncoherent receivers employing square-law combining. Throughout the paper, imperfect SI is obtained using pilot-aided estimation. A lower bound on the performance of iterative receivers that perform joint decoding and channel estimation is obtained assuming the receiver knows the correct data and uses them as pilots. From this, the tradeoff between channel diversity and channel estimation is investigated and the optimal channel memory is approximated analytically. Furthermore, the optimal energy allocation for pilot signals is found for different channel memory lengths.

Index Terms—Block fading, block interference, channel estimation, convolutional codes, interleaving, Nakagami, pilot-aided, Rayleigh, Rician, union bound.

I. INTRODUCTION

A SERIOUS challenge to having good communication quality in wireless communication systems is the time-varying multipath fading environment, which causes the signal to be attenuated randomly and, consequently, the received signal-to-noise ratio (SNR) to vary severely. The fading distribution varies according to the propagation environment. For example, if a line-of-sight exists between the transmitter and receiver in addition to the multipath reception, the fading process is modeled by a Rician distribution [1]. Another popular model for the fading process is the Nakagami distribution [2], which provides a family of distributions that are well matched to measurements under different propagation environments [3], [4].

Error-correcting codes and diversity techniques are standard approaches to mitigate multipath fading. In these techniques, an attempt to providing the receiver with independent fading

realizations of the channel is made. In coded systems, the channel diversity is defined as the number of independent fading realizations available to decode a codeword. Typically, the coded bits are interleaved prior to transmission over the channel in order to distribute errors resulting from consecutive deeply faded bits evenly over the codeword. The performance of coded systems over infinitely interleaved fading channels is commonly analyzed using the union bound as in [5] and [6]. In [3] and [7], the performance of noncoherent and coherent diversity systems over multipath Nakagami fading channels was derived assuming infinite interleaving.

In delay-sensitive applications, infinite interleaving becomes an impractical assumption. Therefore, channel models that exhibit memory such as the block-fading channel [8] are used to model wireless systems including frequency-hopped spread-spectrum (FH-SS) [9], time-division multiplexing (TDM) systems, and orthogonal frequency division multiplexing (OFDM). In this model, a frame consists of blocks of bits that undergo independent fading, where the fading is constant for signals within each block. As an effort to optimize codes for block-fading channels, Leung *et al.* [10] proposed a class of codes suitable for block-fading channels referred to as multifrequency trellis codes. These codes are not practical for block-fading channels with large number of fading blocks. The performance of multifrequency codes was analyzed by Malkamaki *et al.* [11]. Also, several trellis and block codes suitable for block-fading channels with a small number of fading blocks were presented in [12]. In [13], convolutional codes were optimized for block-fading channels. The performance of incremental redundancy convolutional codes over block-fading channels was derived in [14].

In coherent receivers, channel side information (SI) is needed for decoding, i.e., the fading phase and amplitude. If the receiver knows the channel SI perfectly, large channel diversity improves the system performance. Clearly, this is a hypothetical case only. In practice, the receiver needs to estimate the channel. In this case, long channel memory permits better channel estimation. Therefore, longer channel memory improves the performance if the frame size is infinite. However, if the frame size is finite, there exists a fundamental tradeoff between the channel diversity and channel estimation [15]. As the channel memory length increases, the channel diversity is reduced, but the channel estimation becomes more accurate. On the other hand, short channel memory increases the channel diversity, which enables the decoder to average out the channel behavior at the cost of less accurate channel estimation.

In [15], Worthen *et al.* used the error exponent to find the optimal channel memory length of a coded system over some

Manuscript received October 28, 2003; revised December 20, 2004. This work was supported in part by King Fahd University of Petroleum and Minerals, National Taiwan University, and the Office of Naval Research under Grant N00014-03-1-0232. The review of this paper was coordinated by Prof. J. Shea.

S. A. Zummo is with the Electrical Engineering Department, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia (e-mail: zummo@kfupm.edu.sa).

P. Yeh is with the Department of Electrical Engineering, National Taiwan University, Taipei, Taiwan 10617, R.O.C. (e-mail: pcyeh@ntu.edu.tw).

W. E. Stark is with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109 USA (e-mail: stark@eecs.umich.edu).

Digital Object Identifier 10.1109/TVT.2005.858173

simple block memory channels. However, a method to analyze the performance of specific codes over block-fading channels with arbitrary frame size and channel memory length is needed. Such a method is crucial in optimizing the channel memory of a coded system employing practical channel estimation techniques. In this paper, we derive a union bound on the performance of binary convolutional codes over block-fading channels. We assume uniform interleaving of the coded sequence prior to transmission over the channel. Based on this assumption, the distribution of error bits over the fading blocks is derived and the corresponding pairwise error probability is derived under different channel SI assumptions. The proposed bound is used with pilot-aided channel estimation to optimize the channel memory at which the system should operate. One issue that determines the performance of pilot-aided channel estimation is the fraction of energy devoted to the pilot signals. This problem was investigated for different communication systems (see, for example, [16]–[19]). In this paper, we investigate analytically the optimal pilot energy for pilot-aided channel estimation in block-fading channels.

The paper is organized as follows. In Section II, the coded system model is described. Then, the union bound on the error probability of convolutional codes over block-fading channels is derived in Section III. In Sections IV and V, the pairwise error probabilities are derived for Rician and Nakagami fading distributions, respectively, and results are discussed therein. The main conclusions are discussed in Section VI.

II. SYSTEM MODEL

The general block diagram of a binary coded system over block-fading channels is shown in Fig. 1. The transmitter consists of a convolutional encoder, a random interleaver, and a modulator. Time is divided into frames of duration NT seconds, where T is the transmission interval of a bit and N is the number of bits transmitted in a frame. In each time interval of duration KT , a rate- R_c encoder maps K information bits into N coded bits, where $R_c = K/N$ is the code rate. Each coded bit is modulated to generate a signal using either binary phase-shift keying (BPSK) or binary frequency-shift keying (BFSK). The channel we adopt is a block-fading channel in which each frame is subject to F independent fading realizations, resulting in a block of $m = \lceil N/F \rceil$ signals being affected by the same fading realization. If the channel coherence time is longer than the transmission duration of each block, the channel is reasonably assumed to be constant [20]. The coded bits are interleaved prior to transmission over the channel in order to spread out burst errors in the decoder, which result from low instantaneous SNR at the the demodulator output due to fading.

Coherent or noncoherent detection can be employed at the receiver. In coherent receivers, the matched filter sampled output at time lT in the f th fading block is given by

$$y_{f,l} = \sqrt{E_s} h_f s_{f,l} + z_{f,l} \quad (1)$$

where E_s is the average received signal energy, $s_{f,l} = (-1)^{c_{f,l}}$, where $c_{f,l}$ is the corresponding coded bit, and $z_{f,l}$

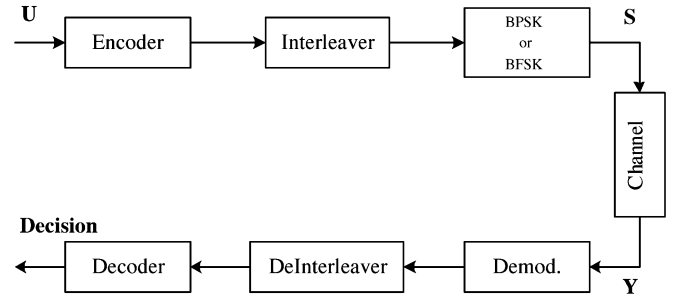


Fig. 1. Structure of a binary coded system.

is an AWGN sample with a complex normal $\mathcal{CN}(0, N_0)$ distribution. The coefficient h_f is the channel gain in fading block f modeled as $\mathcal{CN}(0, 1)$. The channel gain can be written as $h_f = a_f \exp(j\theta_f)$, where θ_f is a uniformly distributed phase and a_f is the amplitude, which is assumed in the paper to have a Rician or a Nakagami distribution.

The receiver employs maximum likelihood (ML) sequence decoding which minimizes the frame error probability. In this rule, the decoder chooses the codeword $\mathbf{S} = \{s_{f,l}, f = 1, \dots, F, l = 1, \dots, m\}$ that maximizes the likelihood function $p(\mathbf{Y} | \mathbf{S})$, where $\mathbf{Y} = \{y_{f,l}, f = 1, \dots, F, l = 1, \dots, m\}$. If perfect SI is available at the receiver, the decoder chooses the codeword \mathbf{S} that maximizes the metric

$$\mathbf{m}(\mathbf{Y}, \mathbf{S}) = \sum_{f=1}^F \sum_{l=1}^m \text{Re}\{y_{f,l}^* h_f s_{f,l}\} \quad (2)$$

where $\text{Re}\{\cdot\}$ represents the real part of a complex number. Note that the perfect SI is a hypothetical assumption that is used to predict the best performance of the code. In practice, the channel is estimated at the receiver as will be discussed in Section IV-B.

In noncoherent systems, BFSK is used where the carrier frequency of the modulated signal is set to one of two frequencies according to whether the coded bit is $c = 0, 1$. For each received signal, square-law combining [21] is employed, whose outputs for $c = 0, 1$ are represented by

$$\begin{aligned} r_{f,l}^{(I,c)} &= \sqrt{E_s} a_f \delta(c_{f,l}, c) \cos(\theta_f) + \eta_{f,l}^{(I,c)} \\ r_{f,l}^{(Q,c)} &= \sqrt{E_s} a_f \delta(c_{f,l}, c) \sin(\theta_f) + \eta_{f,l}^{(Q,c)} \end{aligned} \quad (3)$$

where $r_{f,l}^{(I,c)}$ and $r_{f,l}^{(Q,c)}$ for $c = 0, 1$ are respectively the correlation of the received signal with the in-phase and quadrature dimensions of the signal corresponding to a coded bit c . In (3), θ_f is the unknown phase of the received signals in block f , $\delta(x, y) = 1$ if $x = y$ and $\delta(x, y) = 0$, otherwise, and $\eta_{f,l}^{(I,0)}, \eta_{f,l}^{(Q,0)}, \eta_{f,l}^{(I,1)}$ and $\eta_{f,l}^{(Q,1)}$ are independent random variables with normal distribution, i.e., $\mathcal{N}(0, N_0/2)$ distribution. The decoder chooses the codeword \mathbf{S} that maximizes

$$\mathbf{m}(\mathbf{R}, \mathbf{S}) = \sum_{f=1}^F \sum_{l=1}^m \left(r_{f,l}^{(I,c)} \right)^2 + \left(r_{f,l}^{(Q,c)} \right)^2 \quad (4)$$

where $\mathbf{R} = \{r_{f,l}^{(I,c)}, r_{f,l}^{(Q,c)}, f = 1, \dots, F, l = 1, \dots, m\}$. Note that this decoder makes no use of channel SI in decoding, and is suboptimal with respect to minimizing the frame error probability.

III. PERFORMANCE ANALYSIS

In this section, a union bound on the bit error probability of convolutional codes over block-fading channels is derived. Throughout the paper, the subscripts c, u , and b are used to denote conditional, unconditional, and bit error probabilities, respectively. For linear convolutional codes with k input bits, the bit error probability is upper bounded [5] as

$$P_b \leq \frac{1}{k} \sum_{d=d_{\min}}^N w_d P_u(d) \quad (5)$$

where d_{\min} is the minimum distance of the code, $P_u(d)$ is the unconditional pairwise error probability defined as the probability of decoding a received sequence as a weight- d codeword given that the all-zero codeword was transmitted. In (5), $w_d = \sum_{i=1}^N i A_{i,d}$ is the number of codewords with output weight d , where $A_{i,d}$ is the number of codewords with output weight d and input weight i . The weight distribution $\{w_d\}_{d=d_{\min}}^N$ is obtained directly from the weight enumerator of the code [5].

A. Union Bound for Block-Fading Channels

In block-fading channels, $P_u(d)$ in (5) is a function of the distribution of the d nonzero bits over the F fading blocks. This distribution is quantified assuming uniform channel interleaving of the coded bits over the fading blocks in a frame. Denote the number of fading blocks with weight v by f_v and define $w = \min(m, d)$, then the fading blocks are distributed according to the pattern $\mathbf{f} = \{f_v\}_{v=0}^w$ if

$$F = \sum_{v=0}^w f_v, \quad d = \sum_{v=1}^w v f_v. \quad (6)$$

Denote by $L = F - f_0$ the number of fading blocks with nonzero weights. Then, $P_u(d)$ is determined by averaging over all possible fading block patterns as

$$P_u(d) = \sum_{L=\lceil d/m \rceil}^d \sum_{f_1=0}^{L_1} \sum_{f_2=0}^{L_2} \dots \sum_{f_w=0}^{L_w} P_u(d|\mathbf{f}) p_d(\mathbf{f}) \quad (7)$$

where

$$L_v = \min \left\{ L - \sum_{r=1}^{v-1} f_r, \frac{d - \sum_{r=1}^{v-1} r f_r}{v} \right\}, \quad 1 \leq v \leq w. \quad (8)$$

The probability of a fading block pattern for a specific codeword weight d is computed using combinatorics as

$$p_d(\mathbf{f}) = \frac{\binom{m}{1}^{f_1} \binom{m}{2}^{f_2} \dots \binom{m}{w}^{f_w}}{\binom{mF}{d}} \cdot \frac{F!}{f_0! f_1! \dots f_w!}. \quad (9)$$

The left factor of $p_d(\mathbf{f})$ in (9) is the probability of distributing d nonzero bits over F blocks with f_v blocks having v bits for possible values of v . The right term of $p_d(\mathbf{f})$ is the number of combinations of $\mathbf{f} = \{f_v\}_{v=0}^w$ among the F fading blocks. Using (7)–(9), the union bound on the bit error probability of convolutional codes over a block fading channels is found by substituting (7) in (5).

It should be noted that carefully designed interleavers may outperform the uniform interleaver. However, analyzing coded systems with specific interleavers is much more complicated. Note that the number of summations involved in computing $P_u(d)$ in (7) increases as the channel memory length increases. This makes the computation of (5) when summing over all $d \leq N$ for long channel memory a time consuming task. A good approximation to the union bound is obtained by truncating (5) to a small value of $d_{\max} < N$. This results in an approximation to the error probability rather than an upper bound. However, it is well known that the low-weight terms in the union bound dominate the performance at high SNR values, where the bound is more useful. For low SNR values, simulations can always be performed easily to check the performance. However, at high SNR the bound becomes more accurate because of the domination of the low-weight terms. Thus the bound truncation does not affect the result at high SNR values for which the performance analysis is most important.

B. Pairwise Error Probability

The conditional pairwise error probability $P_c(d|\mathbf{f})$ is defined as the probability of decoding a received sequence \mathbf{Y} as a weight- d codeword $\hat{\mathbf{S}}$ given that the all-zero codeword \mathbf{S} was transmitted and conditioned on the channel fading gains and the fading block pattern \mathbf{f} . It is given by

$$P_c(d|\mathbf{f}) = \Pr(\mathbf{m}(\mathbf{Y}, \mathbf{S}) - \mathbf{m}(\mathbf{Y}, \hat{\mathbf{S}}) < 0 | \mathbf{H}, \mathbf{S}, \mathbf{f}) \quad (10)$$

where $\mathbf{H} = \{h_f\}_{f=1}^F$. For a specific receiver, the unconditional pairwise error probability $P_u(d|\mathbf{f})$ is found by substituting the corresponding decoding metric in (10) and then averaging over the fading statistics. The rest of the paper is devoted to deriving expressions for $P_u(d|\mathbf{f})$ of coded systems over block-fading channels with different receivers and fading distributions.

IV. PAIRWISE ERROR PROBABILITY FOR RICIAN FADING

In this section, we derive the pairwise error probability of coded systems over Rician block-fading channels. Rician fading arises if there is a line-of-sight between the transmitter and the receiver [1]. In this model, the received signal is composed of two signal-dependent components, namely, the specular and diffuse components. The specular component is due to the

line-of-sight reception, whereas the diffuse component results from multipath reception. In this case, the channel gain in each fading block h_f is modeled as a complex Gaussian variable with $\mathcal{CN}(b, 1)$ distribution, where b represents the specular component of the channel. Thus the amplitude a_f has a Rician distribution with a normalized density function [6] given by

$$f_{a_f}(a) = 2a(1+K)\exp[-K - a^2(1+K)] \times I_0\left(2a\sqrt{K(1+K)}\right), \quad a \geq 0 \quad (11)$$

where $K = b^2$ is the energy of the specular component, and $I_0(\cdot)$ is the zero-order modified Bessel function of the first kind. In this context, K denotes the ratio of the specular component energy to the diffuse component energy. When $K = 0$, the specular component is zero, resulting in the well-known Rayleigh fading distribution. In the following, $P_u(d|\mathbf{f})$ is derived for coherent detection with perfect and imperfect SI available at the receiver. Furthermore, a square-law combining receiver is considered with Rayleigh fading.

A. Coherent Detection—Perfect SI

Recall that the received signal over a block-fading channel is given by (1) and the corresponding ML decoding rule is given by (2). Substituting the metric (2) in (10), the conditional pairwise error probability for coherent detection with perfect SI is given by

$$P_c(d|\mathbf{f}) = \Pr\left(\sum_{f=1}^L a_f \sum_{l=1}^m \operatorname{Re}\{y_{f,l}\} < 0 \mid \mathbf{H}, \mathbf{S}, \mathbf{f}\right). \quad (12)$$

The distribution of $\operatorname{Re}\{y_{f,l}\}$ conditioned on a_f is Gaussian with mean $\sqrt{E_s}a_f s_{f,l}$ and variance N_0 . Thus, $P_c(d|\mathbf{f})$ simplifies to

$$P_c(d|\mathbf{f}) = Q\left(\sqrt{2R_c\gamma_b \sum_{v=1}^w v \sum_{i=1}^{f_v} a_i^2}\right) \quad (13)$$

where $\gamma_b = E_b/N_0$ is the SNR per information bit. Note that the average energy per bit is given by $E_b = R_c E_s$, where R_c is the code rate. To find the unconditional pairwise error probability $P_u(d|\mathbf{f})$, (13) is averaged over the statistics of the fading amplitudes in (11). An exact expression of the pairwise error probability is found using the integral form of the Q -function, $Q(x) = \frac{1}{\pi} \int_0^{\pi/2} e^{-x^2/2\sin^2\theta} d\theta$ [22], resulting in

$$\begin{aligned} P_u(d|\mathbf{f}) &= \frac{1}{\pi} \mathbb{E}_{\mathbf{A}} \left[\int_0^{\pi/2} \exp\left(-\frac{R_c\gamma_b}{\sin^2\theta} \sum_{v=1}^w v \sum_{i=1}^{f_v} a_i^2\right) d\theta \right] \\ &= \frac{1}{\pi} \int_0^{\pi/2} \prod_{v=1}^w \left(\frac{1+K}{1+K+vR_c\gamma_b/\sin^2\theta} \right)^{f_v} \\ &\quad \times \exp\left(-\frac{Kv f_v R_c\gamma_b/\sin^2\theta}{1+K+vR_c\gamma_b/\sin^2\theta}\right) d\theta \quad (14) \end{aligned}$$

where the product results from the independence of the fading variables in different fading blocks. Since the integral in (14)

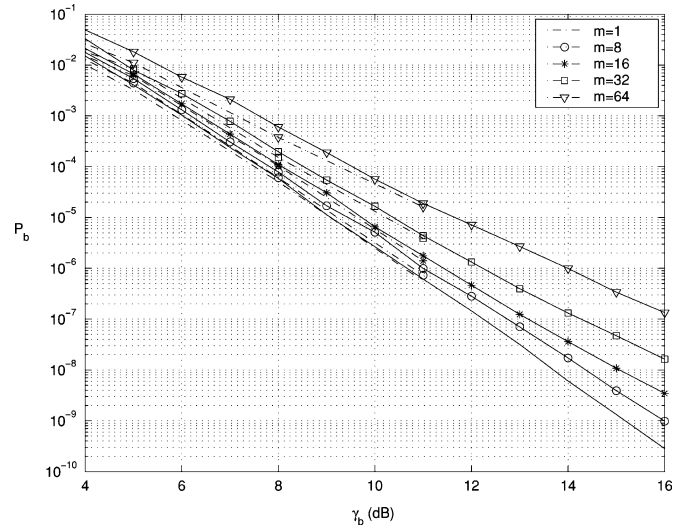


Fig. 2. Bit error probability of a rate-1/2 (23,35) convolutional code with perfect SI and a frame size $N = 1024$ for channel memory lengths $m = 1, 8, 16, 32, 64$ (solid: approximation using the union bound, dash: simulation).

is definite, its computation is straightforward using standard numerical integration packages.

Throughout the paper, the union bound was evaluated for a rate-1/2 (23,35) convolutional code with a frame size of $N = 1024$ coded bits. As discussed in Section III-A, the union bound is truncated to sum over codewords with a distance $d_{\max} \leq 12$ in order to reduce the computational complexity. The bound is compared to simulation results, in which the channel interleaver is chosen randomly and is changed every ten frames to simulate the effect of the uniform interleaver. This results in the performance of the average-performing interleaver over block-fading channels. In coherent systems, BPSK signaling is employed, and (14) is used in (7) and (5) to compute the union bound.

Fig. 2 shows the results for Rayleigh fading channels with perfect SI and different channel memory lengths. Since simulating very low error rates is too difficult, we plot simulation curves down to error rates around $P_b = 10^{-6}$. However, the analytical curves are shown for all SNR values. We observe that the approximation is very close to the simulation curves for a wide range of channel memory lengths. Also, the approximation starts to be loose as the SNR decreases. It is well established that the union bound diverges at SNR values lower than the cutoff rate of the channel [23]. However, the value of the analytical results is more interesting for high SNR values, where simulation is difficult to conduct. In the rest of the paper, analytical results are shown for high SNR values to make the presentation more clear.

Fig. 3 shows the SNR required for the convolutional code to achieve $P_b = 10^{-4}$ versus the specular-to-diffuse ratio K of a Rician fading channel. We observe that increasing the energy of the line-of-site component of the channel reduces the effect of the diversity provided by the independent fading blocks. This is expected, since increasing K causes the channel to become less random, which reduces the need for diversity at the decoder. We conclude that when perfect SI is available at the receiver,

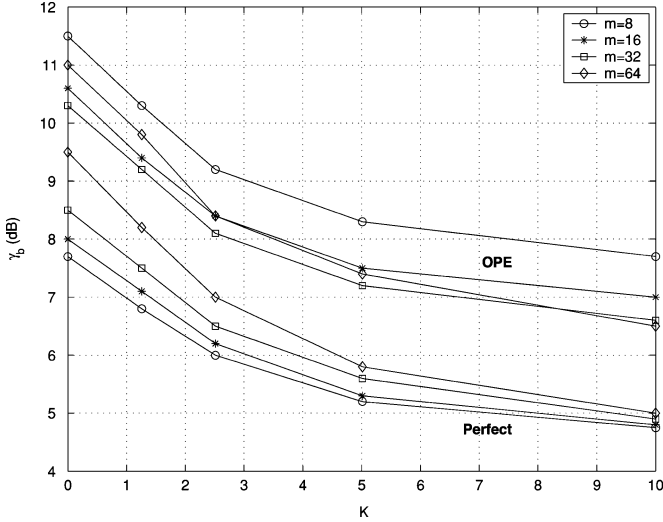


Fig. 3. SNR required for a rate-1/2 (23, 35) convolutional code to achieve $P_b = 10^{-4}$ versus the specular-to-diffuse ratio K (linear scale) for memory lengths $m = 8, 16, 32, 64$.

smaller specular-to-diffuse ratio makes the performance more sensitive to the lack of channel diversity.

B. Coherent Detection—Imperfect SI

For coherent detection with imperfect SI, it is necessary to estimate the channel SI. This is achieved by transmitting a pilot signal with energy E_p in each fading block. The corresponding received signal is given by

$$y_{f,p} = \sqrt{E_p} h_f + z_{f,p}. \quad (15)$$

The ML estimator for h_f is given by $\hat{h}_f = y_{f,p} / \sqrt{E_p} = h_f + e_f$, where $e_f = z_{f,p} / \sqrt{E_p}$ is the estimation error. The distribution of e_f is $\mathcal{CN}(0, \sigma_e^2)$, where $\sigma_e^2 = N_0 / E_p$. The correlation coefficient between the actual channel gain and its estimate is given by

$$\mu = \frac{E[(h_f - b)(\hat{h}_f - b)^*]}{\sqrt{\text{Var}(h_f)\text{Var}(\hat{h}_f)}} = \frac{1}{\sqrt{1 + \sigma_e^2}}. \quad (16)$$

In order to implement an ML sequence decoding rule, the likelihood function of the channel observations (received and pilot signals) conditioned on the transmitted codeword $p(\mathbf{Y}, \hat{\mathbf{H}}|\mathbf{S})$ should be maximized. In [24], the ML decoding rule was shown to be difficult to implement in a Viterbi receiver. Therefore, a suboptimal decoding metric that maximizes the likelihood function $p(\mathbf{Y}|\hat{\mathbf{H}}, \mathbf{S})$ is used. It is given by choosing a codeword \mathbf{S} that maximizes the metric

$$\mathbf{m}(\mathbf{Y}, \mathbf{S}) = \sum_{f=1}^F \sum_{l=1}^m \text{Re}\{y_{f,l}^* \hat{h}_f s_{f,l}\}. \quad (17)$$

Substituting (17) in (10), the conditional pairwise error probability for the suboptimal decoder becomes

$$P_c(d|\mathbf{f}) = \Pr \left(\sum_{f=1}^L \sum_{l=1}^m \text{Re}\{y_{f,l}^* \hat{h}_f\} < 0 \mid \hat{\mathbf{H}}, \mathbf{S}, \mathbf{f} \right). \quad (18)$$

In order to find (18), we need to find the distribution of $y_{f,l}$ conditioned on \hat{h}_f , which is a complex Gaussian random variable with a mean $\sqrt{E_s} s_{f,l} E[h|\hat{h}]$ and a variance $N_0 + (1 - \mu^2)E_s$, where $E[h|\hat{h}] = \mu/\sigma(\hat{h}_f - b) + b$ and $\sigma^2 = \text{Var}(\hat{h}_f) = 1 + \sigma_e^2$. Thus, the conditional pairwise error probability for the suboptimal decoder is given by

$$P_c(d|\mathbf{f}) = Q \left(\sqrt{\frac{2E_s \sum_{f=1}^F d_f \left| \frac{\mu}{\sigma}(\hat{h}_f - b) + b \right|^2}{N_0 + (1 - \mu^2)E_s}} \right) \quad (19)$$

where d_f is the number of nonzero error bits in fading block f . Define the normalized complex Gaussian random variable $\zeta_f = (\hat{h}_f - b)/\sigma + b/\mu$ with a distribution $\mathcal{CN}(b/\mu, 1)$. Then, the conditional pairwise error probability simplifies to

$$P_c(d|\mathbf{f}) = Q \left(\sqrt{\frac{2\mu^2 R_c \gamma_b \sum_{v=1}^w v \sum_{i=1}^{f_v} |\zeta_i|^2}{1 + R_c \gamma_b (1 - \mu^2)}} \right). \quad (20)$$

Comparing (13) and (20), we conclude that the pairwise error probability for the case of imperfect SI is given by (14), with γ_b and K being replaced by $\hat{\gamma}_b = \mu^2 \gamma_b / (1 + R_c \gamma_b (1 - \mu^2))$ and K/μ^2 , respectively.

Two scenarios can be considered for the channel estimation using pilot signals with $E_p = E_s$. The first case results from only pilot estimation (OPE) with an estimation error variance of $\sigma_e^2 = N_0 / E_s$. The second case considers a lower bound on the performance of receivers employing iterative joint decoding and channel estimation. In such receivers, the decoding results are used to improve the channel estimates, which are used to improve the decoding results. This process is repeated iteratively. In general, the more reliable the decoding results, the more accurate the channel estimation. A lower bound on the performance of iterative receivers is obtained if the signals in each fading block are known with probability one. In this case they can be considered as pilot signals resulting in an estimation error of variance $\sigma_e^2 = N_0 / (mE_s)$. This case is referred to as correct data estimation (CDE). Similar channel estimation scenarios were used in [17] and [18] for channel estimation with LDPC codes.

In simulating systems with pilot-aided channel estimation, one coded bit is punctured every m coded bits to account for the rate reduction resulting from inserting a pilot signal every $m - 1$ signals. This affects the whole distance distribution of the resulting code and may reduce the minimum distance of the code. The resultant code rate after puncturing is given by

$$\tilde{R}_c = \frac{mR_c}{m-1}. \quad (21)$$

Table I shows the code rates \tilde{R}_c and the minimum distances of the punctured codes for different channel memory lengths. Also,

TABLE I
RATES, MINIMUM DISTANCES AND PUNCTURING LOCATIONS WITHIN EACH
FADING BLOCK OF THE PUNCTURED RATE-1/2 (23,35)
CONVOLUTIONAL CODE

m	\bar{R}_c	Puncturing Location	d_{\min}
4	0.667	3	4
8	0.571	7	5
16	0.533	15	6
32	0.516	31	6
64	0.508	63	6

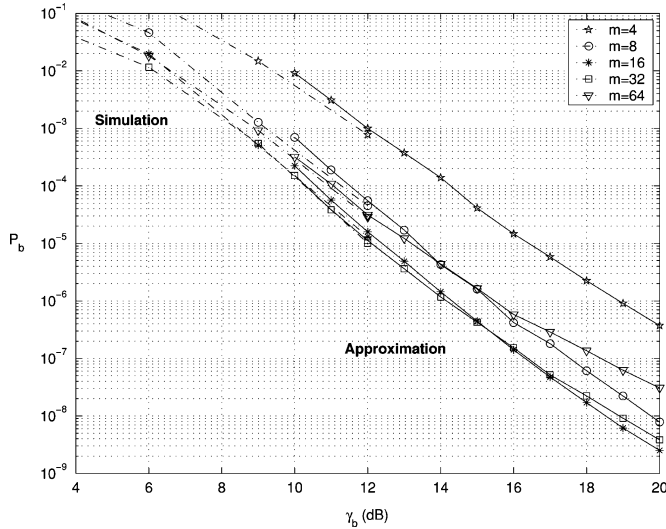


Fig. 4. Bit error probability of a rate-1/2 (23,35) convolutional code with imperfect SI (OPE receiver with $E_p = E_s$) and a frame size $N = 1024$ for channel memory lengths $m = 4, 8, 16, 32,$ and 64 .

the puncturing pattern is presented in the table. According to the table, the code rate increases with reduced channel memory length, which decreases the error correction capability of the code. Thus, systems with short channel memory are expected to have more channel diversity at the cost of lower minimum distance and worse channel estimation quality. On the other hand, longer channel memory results in more powerful codes, as well as better channel estimation, at the cost of less channel diversity.

Note that optimizing the channel memory is under the control of the system designer. As mentioned in the introduction, the block-fading model is used frequently to model communication systems such as FH-SS, TDM, and OFDM systems. In these systems, the channel memory is under the control of the system designer. More specifically, the designer can optimize the memory length by finding the most appropriate number of hops that the transmitter should hop within a codeword in FH-SS systems, and the number of time slots in a TDM frame over which a codeword is transmitted in TDM systems.

Fig. 4 shows the results for the rate-1/2 (23,35) convolutional code with imperfect SI with an OPE receiver over Rayleigh fading channels. Note that the energy of the pilot is taken into account in the SNR axis. Unlike perfect SI, imperfect SI results in a clear tradeoff between channel diversity and estimation. From the figure, the cases of $m = 16$ and $m = 32$ are the best performing systems, where the former becomes better than the

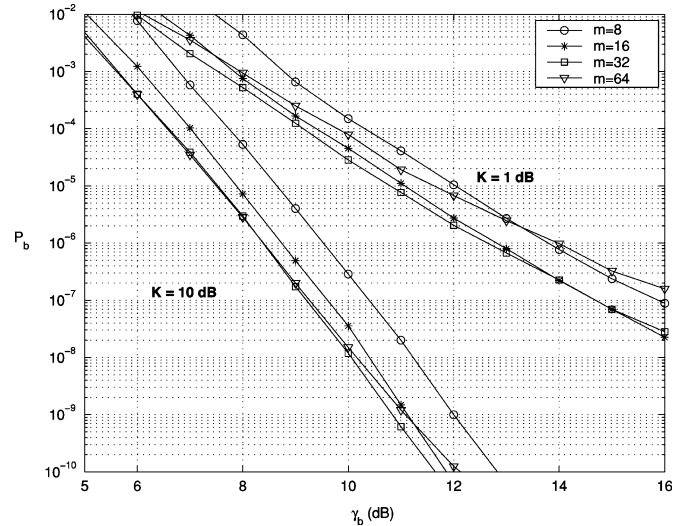


Fig. 5. Approximation of the bit error probability of a rate-1/2 (23,35) convolutional code over a Rician fading channel with $K = 1, 10$ dB, imperfect SI (OPE receiver), and a frame size $N = 1024$ for memory lengths $m = 8, 16, 32,$ and 64 .

later for an SNR values exceeding 14 dB. This suggests that the optimal channel memory length is between $m = 16$ and $m = 32$. Also, the case of $m = 8$ is worse than the case of $m = 64$ at low SNR, and starts to improve as the SNR increases. This is because the resulting code for $m = 8$ is less powerful than the code for $m = 64$ but has larger amount of channel diversity. Although the case of $m = 64$ has the best code and channel estimation quality, it lacks enough channel diversity to perform better than the other cases. This is expected since the number of transmitted pilot signals is reduced as the channel memory gets longer, and hence, the system becomes more energy efficient and the code become stronger.

Fig. 5 shows the results of imperfect SI with an OPE receiver over different Rician fading channels. We observe that systems with long channel memory perform better as the energy of the specular component of the channel increases. This is because as K increases, the channel becomes less faded, which reduces the need for the decoder to average over the statistics of the channel. Therefore, the channel diversity becomes less crucial, causing systems with long channel memory to outperform systems with short memory. Another reason for this is the larger fraction of energy spent on pilot signals in systems with short channel memory lengths than in systems with long memory. This is obvious for the case of $K = 10$ dB, where the performance of $m = 64$ is nearly optimal for most of the SNR values. On the other hand, the case of $m = 8$ is the worst everywhere when $K = 10$ dB, where it outperforms the case of $m = 64$ when $K = 1$ dB. From Fig. 3, the optimal channel memory value for an OPE receiver with $E_p = E_s$ at a bit error rate of 10^{-4} is $m = 32$ for a Rayleigh fading channel; i.e., $K = 0$, where it is $m = 64$ for a Rician channel with $K = 10$. Also, note that the case of $m = 8$ outperforms the case of $m = 64$ when the channel is more faded, where the inverse occurs for channels that are less faded, i.e., larger values of K .

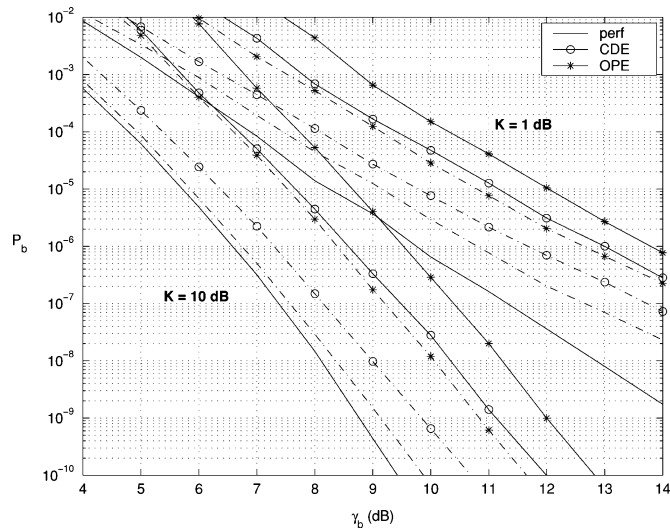


Fig. 6. Approximation of the bit error probability of a rate-1/2 (23,35) convolutional code over a Rician fading channel with $K = 1, 10$ dB, frame size $N = 1024$, and memory lengths $m = 8, 32$ using perfect and imperfect SI with $E_p = E_s$ (solid: $m = 8$, dash: $m = 32$).

Fig. 6 shows a comparison of systems employing perfect SI, OPE, and CDE assumptions with channel memory lengths $m = 8$ and $m = 32$ over Rician fading channels with $K = 1, 10$ dB. It is clear that as the channel memory gets longer, the SNR degradation due to imperfection in the channel SI reduces. This is because long channel memory causes less penalty in the rate and energy than short channel memory does, as well as an improved channel SI under the CDE assumption. In general, the optimal memory tends to increase under the CDE assumption compared to the OPE receiver due to the improved channel estimation [24]. Moreover, the SNR loss in OPE receivers with long channel memory increases with increased energy of the specular component of the channel. When the channel is estimated using a pilot signal, the channel estimation error adds a fading component to the channel gain at the decoder. The effect of this new fading component increases as the energy of the specular component increases of the channel, which degrades the performance of OPE receivers more as K increases.

When the energy allocated for the pilot signal is varied, the performance of an OPE receiver is expected to change as a function of the channel memory. The energy per information bit can be written as

$$E_b = \frac{(m-1)E_s + E_p}{mR_c}. \quad (22)$$

Thus, for a fixed channel memory, there exists an optimal value for pilot energy. This is illustrated in Fig. 7, where the SNR required for the coded system to achieve a bit error probability of $P_b = 10^{-4}$ is shown versus the pilot-to-signal energy ratio E_p/E_s in decibels. We observe that as the channel memory length increases the optimal value for E_p/E_s increases. This is expected since longer channel memory permits the allocation of more energy to the pilot signal. On the other hand, when the channel memory is short, a wiser usage of the available energy seems to transmit the information signals rather than to estimate

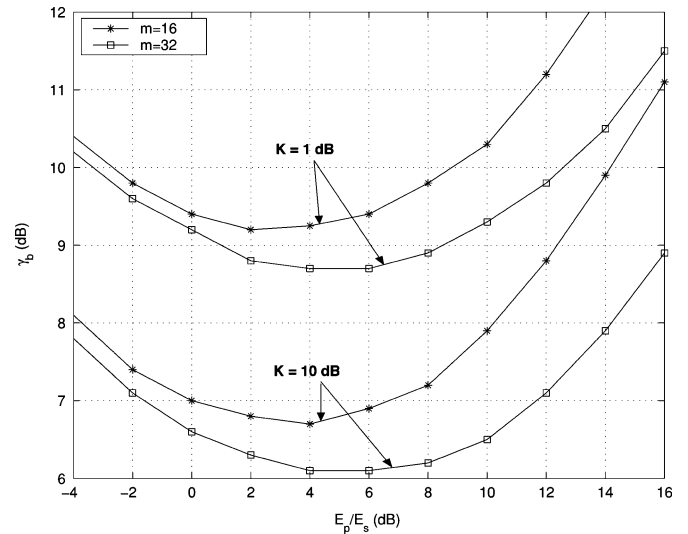


Fig. 7. SNR required for a rate-1/2 (23, 35) convolutional code to achieve $P_b = 10^{-4}$ versus E_p/E_s for the OPE receiver with $E_p = E_s$ and memory lengths $m = 16$ and 32 .

the channel. We conclude from Fig. 7 that optimizing the pilot energy results in an SNR gain slightly less than 1 dB over the case, where $E_p = E_s$. This SNR gain increases as the channel memory increases, since longer memory increases the amount of energy that can be devoted to channel estimation, which improves the overall performance. Moreover, the optimal pilot energy is almost independent of the fading nature of the channel, i.e., independent of the energy of the specular component K of the channel. This is because the amount of energy available in each fading block, which can be used in estimating the channel, is the controlling factor of the optimal pilot energy allocation. Clearly, this energy is a function of the channel memory length only. Also, the SNR gain resulting from optimizing the pilot energy is almost independent of the channel fading behavior. Same observations were reported in [17]–[19].

C. Noncoherent Detection

In this section, we derive a Chernoff bound on the pairwise error probability of coded systems employing square-law combiner over Rayleigh block-fading channels. Recall that the outputs of the square-law combiner and the corresponding decoding metrics are given by (3) and (4), respectively. Substituting the metric (4) in (10) yields

$$P_c(d|\mathbf{f}) = \Pr \left(\sum_{f=1}^F d_f \left(|r_f^{(I,1)}|^2 + |r_f^{(Q,1)}|^2 - |r_f^{(I,0)}|^2 - |r_f^{(Q,0)}|^2 \right) > 0 \mid \mathbf{H}, \mathbf{S}, \mathbf{f} \right) \quad (23)$$

where d_f is the number of error bits in the fading block f . The variables $\{r_f^{(I,0)}, r_f^{(Q,0)}\}$ and $\{r_f^{(I,1)}, r_f^{(Q,1)}\}$ are zero mean Gaussian random variables with variances equal to $1/2(E_s + N_0)$ and $1/2N_0$, respectively. Let $|r_f^{(c)}|^2 = |r_f^{(I,c)}|^2 + |r_f^{(Q,c)}|^2$ for $c = 0, 1$ and define $\kappa = \sum_{f=1}^F d_f (|r_f^{(1)}|^2 - |r_f^{(0)}|^2)$. Then,

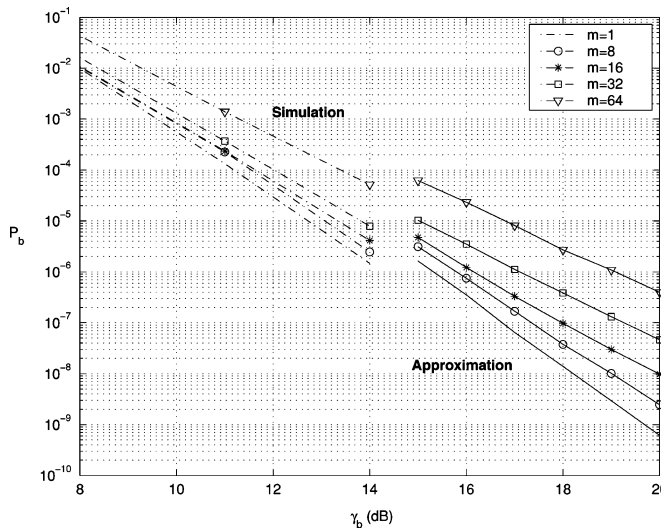


Fig. 8. Bit error probability of a rate- $\frac{1}{2}$ (23, 35) convolutional code with noncoherent detection and a frame size $N = 1024$ for channel memory lengths $m = 1, 8, 16, 32,$ and 64 .

the unconditional pairwise error probability is

$$P_u(d|\mathbf{f}) = \int_0^\infty p(\kappa) d\kappa \leq E_\kappa[e^{\lambda\kappa}]. \quad (24)$$

Due to the complicated form of the pdf $p(\kappa)$ [20], the Chernoff bound was used in (24) to upper bound $P_u(d|\mathbf{f})$, where $\lambda > 0$ is the Chernoff parameter that should be optimized for the tightest bound. Substituting for κ in the Chernoff bound and collecting terms having the same distance results in

$$P_u(d|\mathbf{f}) \leq \prod_{v=1}^w E \left[e^{\lambda v |r_v^{(1)}|^2} \right]^{f_v} E \left[e^{-\lambda v |r_v^{(0)}|^2} \right]^{f_v}. \quad (25)$$

The Chernoff parameter λ is optimized as in [21], and the resulting Chernoff bound for the pairwise error probability simplifies to

$$P_u(d|\mathbf{f}) \leq \prod_{v=1}^w [4D_v(1 - D_v)]^{f_v} \quad (26)$$

where $D_v = 1/(2 + vR_c\gamma_b)$. The Chernoff bound approximation for convolutionally encoded BFSK signals with square-law combining is shown in Fig. 8. Note that we show the approximation for high SNR values to avoid curves to overlap resulting in unclear presentation. We observe that the approximation is not as close to the simulation curves as in the case of coherent receivers. This is mainly because of the use of the Chernoff bounding technique. This makes our approximation not as useful in predicting the error probability of square-law combining. However, the proposed approximation predicts precisely the performance loss due to channel memory. Note that the union bound for the performance of square-law combiner can be evaluated more precisely for the case of $m = 1$, i.e., without the need for the Chernoff bound [20]. Since the Chernoff bound predicts the performance loss due to channel memory, it can be compared with the more precise union bound for $m = 1$ case in order to predict the performance of the square-law combiner in a Rayleigh block-fading environment.

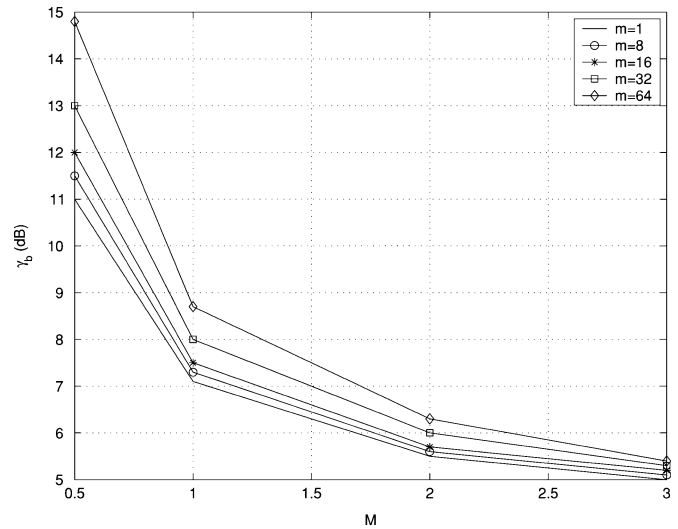


Fig. 9. SNR required for a rate- $\frac{1}{2}$ (23, 35) convolutional code with perfect SI to achieve $P_b = 10^{-3}$ versus the Nakagami parameter M for memory lengths $m = 1, 8, 16, 32,$ and 64 .

V. PAIRWISE ERROR PROBABILITY FOR NAKAGAMI FADING

In this section, we derive the pairwise error probability of coded systems over Nakagami block-fading channels. In Nakagami block-fading channels, the fading amplitude in each fading block is Nakagami distributed with a normalized density function [2] given by

$$f_{a_f}(a) = \frac{2M^M}{\Gamma(M)\Omega^M} a^{2M-1} e^{-Ma^2/\Omega}, \quad a \geq 0, M \geq 0.5 \quad (27)$$

where $\Omega = E[a^2] = 1$, $M = \Omega^2/\text{Var}[a]$ is the fading parameter, and $\Gamma(\cdot)$ is the Gamma function. As M increases, the fading becomes less severe and reaches the nonfading scenario when $M \rightarrow \infty$, i.e., AWGN channels. The Nakagami distribution covers a wide range of fading distributions including Rayleigh fading when $M = 1$, and the single-sided Gaussian distribution when $M = 0.5$. In the following, only coherent detection with perfect SI is considered.

If the channel SI is known perfectly at the receiver, the ML decoder chooses the codeword \mathbf{S} that maximizes the metric in (2). As in Section IV, the conditional pairwise error probability is given by (13). Averaging over the statistics of the fading amplitude in (27), an exact expression of $P_u(d|\mathbf{f})$ [25] is found as

$$P_u(d|\mathbf{f}) = \frac{1}{\pi} \int_0^{\frac{\pi}{2}} \prod_{v=1}^w \left(\frac{1}{1 + \frac{vR_c\gamma_b}{M \sin^2 \theta}} \right)^{M f_v} d\theta. \quad (28)$$

The union bound for Nakagami block-fading channels is computed by substituting (28) in (7). The effect of the fading parameter on the performance of coherent detection with perfect SI is shown in Fig. 9. We observe that as the fading parameter increases, the SNR loss due to channel memory reduces. This is expected, since as the fading parameter increases, the diversity becomes less important since the channel approaches the AWGN channel.

VI. CONCLUSION

In this paper, a union bound on the performance of binary-coded systems over block-fading channels was proposed. The bound is based on uniform interleaving of the coded sequence prior to transmission over the channel and was evaluated for block-fading channels with different fading distributions. Coherent and noncoherent receivers were considered. Results show that the SNR degradation due to channel memory reduces as the channel becomes less fading. Moreover, the interesting situation of imperfect SI obtained from pilot-aided channel estimation was considered. Using the CDE assumption, a lower bound on the performance of receivers employing iterative decoding and channel estimation was obtained. Moreover, the tradeoff between channel estimation and channel diversity was investigated. Results show that the optimal channel memory increases as the channel becomes less fading. Moreover, it was shown that the optimal pilot energy allocation is a function of the channel memory length and does not depend on the channel distribution.

REFERENCES

- [1] S. Rice, "Statistical properties of a sine wave plus random noise," *Bell Syst. Tech. J.*, vol. 27, pp. 109–157, Jan. 1948.
- [2] M. Nakagami, "The m -distribution—a general formula of intensity distribution of fading," in *Statistical Methods in Radio Wave Propagation*. New York: Pergamon, 1960.
- [3] E. Al-Hussaini and A. Al-Bassiouni, "Performance of MRC diversity systems for the detection of signals with Nakagami fading," *IEEE Trans. Commun.*, vol. COM-33, pp. 1315–1319, Dec. 1985.
- [4] H. Hashemi, "The indoor radio propagation channel," *Proc. IEEE*, vol. 81, no. 7, pp. 943–968, Jul. 1993.
- [5] A. Viterbi, "Convolutional codes and their performance in communication systems," *IEEE Trans. Commun.*, vol. COM-19, pp. 751–772, Oct. 1971.
- [6] D. Divsalar and M. Simon, "Trellis coded modulation for 4800–9600 bits/s transmission over a fading satellite channel," *IEEE J. Sel. Areas Commun.*, vol. SAC-5, pp. 162–175, Feb. 1987.
- [7] U. Charash, "Reception through nakagami fading multipath channels with random delays," *IEEE Trans. Commun.*, vol. COM-27, pp. 657–670, Apr. 1979.
- [8] R. J. McEliece and W. E. Stark, "Channels with block interference," *IEEE Trans. Inf. Theory*, vol. IT-30, no. 1, pp. 44–53, Jan. 1984.
- [9] W. E. Stark, "Coding for frequency-hopped spread-spectrum communication with partial-band interference—Part I: Capacity and cutoff rate," *IEEE Trans. Commun.*, vol. COM-33, pp. 1036–1044, Oct. 1985.
- [10] Y. Leung, S. Wilson, and J. Ketchum, "Multifrequency trellis coding with low delay for fading channels," *IEEE Trans. Commun.*, vol. 41, no. 10, pp. 1450–1459, Oct. 1993.
- [11] E. Malkamaki and H. Leib, "Evaluating the performance of convolutional codes over block-fading channels," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1643–1646, Jul. 1999.
- [12] R. Knopp and P. Humblet, "On coding for block-fading channels," *IEEE Trans. Inf. Theory*, vol. 46, no. 1, pp. 189–205, Jan. 2000.
- [13] M. Chiani, A. Conti, and V. Tralli, "Further results on convolutional code search for block-fading channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 6, pp. 1312–1318, Jun. 2004.
- [14] C. Leanderson and G. Caire, "The performance of incremental redundancy schemes based on convolutional codes in the block-fading gaussian collision channel," *IEEE Trans. Wireless Commun.*, vol. 3, no. 3, pp. 843–854, May 2004.
- [15] A. Worthen and W. Stark, "On the channel memory-diversity tradeoff in communication systems," in *Information, Coding and Mathematics*. Boston, MA: Kluwer, 2002.
- [16] M. Chiani, A. Conti, and C. Fontana, "Improved performance in TD-CDMA mobile radio system by optimizing energy partition in channel estimation," *IEEE Trans. Commun.*, vol. 51, no. 3, pp. 352–355, Mar. 2003.
- [17] K. Fu and A. Anastasopoulos, "Performance analysis of LDPC codes for time-selective complex fading channels," in *Proc. Global Telecommunications Conf., GLOBECOM*, 2002, pp. 1279–1283.
- [18] R. Nuriyev and A. Anastasopoulos, "Rotationally invariant and rotationally robust codes for the AWGN and the noncoherent channel," *IEEE Trans. Commun.*, vol. 51, no. 12, pp. 2001–2010, Dec. 2003.
- [19] L. Chong and L. Milstein, "The effects of channel-estimation errors on a space-time spreading CDMA system with dual transmit and dual receive diversity," *IEEE Trans. Commun.*, vol. 52, no. 7, pp. 1145–1151, Jul. 2004.
- [20] J. G. Proakis, *Digital Communications*, 4th ed. New York: McGraw-Hill, 2000.
- [21] J. Wozencraft and I. Jacobs, *Principles of Communication Engineers*. New York: Wiley, 1965.
- [22] M. Simon and D. Divsalar, "Some new twists to problems involving the Gaussian probability integral," *IEEE Trans. Commun.*, vol. 46, no. 2, pp. 200–210, Feb. 1998.
- [23] R. Gallager, *Information Theory and Reliable Communication*. New York: Wiley, 1968.
- [24] S. A. Zummo, "Coding and Channel Estimation for Block-Fading Channels," Communications and Signal Processing Laboratory (CSPL), Univ. of Michigan, Ann Arbor, Tec. Rep., Jun. 2003.
- [25] M. Alouini and M. Simon, "A unified approach to the performance analysis of digital communication over generalized fading channels," *Proc. Inst. Elect. Eng.*, vol. 86, no. 9, pp. 1860–1877, Sep. 1998.



Salam A. Zummo (M'03) received the B.Sc. and M.Sc. degrees (with highest honors) in electrical engineering from King Fahd University of Petroleum and Minerals (KFUPM), Dhahran, Saudi Arabia, in 1998 and 1999, respectively. He received the Ph.D. degree from the University of Michigan, Ann Arbor, in June 2003.

He is currently an Assistant Professor of Electrical Engineering at KFUPM. His research interests include error control coding, coded modulation, iterative receivers, interference modeling and analysis,

and cross-layer design of wireless communication networks

Dr. Zummo was the recipient of the British Council Summer Research Fellowship in June 2004.



Ping-Cheng Yeh (M'05) received the B.S. degree in mathematics and the M.S. degree in electrical engineering from the National Taiwan University, Taipei, Taiwan, R.O.C., in 1996 and 1998, respectively. He received the Ph.D. degree from the University of Michigan, Ann Arbor, in 2005.

He is currently an Assistant Professor in the Department of Electrical Engineering at the National Taiwan University. His research interests include channel coding, coded modulation, directional antennas, and cross-layer design in wireless networks.



Wayne E. Stark (F'98) received the B.S. (with highest honors), M.S., and Ph.D. degrees in electrical engineering from the University of Illinois, Urbana, in 1978, 1979, and 1982, respectively.

Since September 1982, he has been a faculty member in the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, where he is currently Professor. His research interests are in the areas of coding and communication theory, especially for spread-spectrum and wireless communication networks.

Dr. Stark was Editor for Communication Theory of the IEEE TRANSACTIONS ON COMMUNICATION in the area of Spread-Spectrum Communications from 1984 to 1989. He was involved in the planning and organization of the 1986 International Symposium on Information Theory which was held in Ann Arbor, Michigan. He was selected by the National Science Foundation as a 1985 Presidential Young Investigator. He was the principal investigator of a Army Research Office Multidisciplinary University Research Initiative (MURI) project on Low Energy Mobile Communications. He received the IEEE Military Communications Conference Technical Achievement Award in 2002. He is a member of Eta Kappa Nu, Phi Kappa Phi and Tau Beta Pi.