

Universal Receivers with Side Information from the Demodulators: An Example for Nonselective Rician Fading Channels

Upamanyu Madhow and Michael B. Pursley, *Fellow, IEEE*

Abstract— We consider binary orthogonal signaling over a nonselective Rician-fading channel with additive white Gaussian noise. The received signal over such a channel may have both a specular component and a scatter (Rayleigh-faded) component. If there is only a scatter component, the noncoherent receiver is optimal. If there is only a specular component, the optimal receiver is the coherent receiver. In general, the optimal receiver for a Rician channel depends on the strengths of the two signal components and the noise density, and the set of possible optimal receivers is infinite. We consider a system in which the noncoherent receiver and the coherent receiver are employed in a parallel configuration for a symbol-by-symbol demodulation of the received signal. Each sequence of transmitted symbols produces a sequence at the output of each of the parallel receivers. The task of identifying which of these received sequences is a more reliable reproduction of the transmitted sequence is the data verification problem. In this paper, we show that data verification can be accomplished by combining side information from the demodulators with a suitable error-control coding scheme. The resulting system is a universal receiver that provides good performance over the entire range of channel parameters. In particular, the universal receiver performs better than the traditional noncoherent receiver.

I. INTRODUCTION

THE CONCEPT of universal receivers for unknown and time-varying channels has been introduced in [5] and [8]. In this and a companion paper [7], we consider different aspects of the application of this concept to binary signaling over nonselective Rician-fading channels with additive white Gaussian noise (AWGN). The received signal over such channels can have both a specular component and a scatter (Rayleigh-faded) component. Among the class of channels of interest are Rayleigh fading channels, for which there is only a scatter component. Since it is difficult to obtain a phase reference for the received signal in such situations, we must employ noncoherent demodulation, so that antipodal signaling cannot be used. In view of the possibility of Rayleigh fading, we consider orthogonal signaling in this paper, since such signals are appropriate for noncoherent demodulation. However, noncoherent demodulation is not optimal if there is a specular signal component. In general, the optimal receiver for a Rician

fading channel depends on the strengths of both the signal components and the power spectral density of the AWGN, so that there is no single receiver that gives optimal performance for all Rician fading channels. If there is only a specular component, the optimal receiver is the *coherent receiver*, and if there is only a scatter component, the optimal receiver is the *noncoherent receiver*. Thus, the coherent receiver gives nearly optimal performance if the scatter component is small, while the noncoherent receiver gives nearly optimal performance when the specular component is small.

The foregoing discussion motivates us to consider a communication system in which the received signal is demodulated in parallel by the coherent receiver and the noncoherent receiver. Regardless of the parameters of the channel in effect, one of these receivers gives good performance. In this paper, we develop side information that identifies, with high probability, which of these receivers is performing well under the current channel conditions, and show that the resulting system outperforms a single noncoherent receiver.

Our purpose is to use this system as an example to illustrate the concept of a *universal receiver* [5]–[7]. Given a class of channels for which no single receiver gives optimal performance, it is shown in [5] that, under certain conditions, it is possible to find a finite set of receivers with the property that, for any channel in the class of interest, at least one of these receivers gives a performance which is within a prescribed deviation of the *optimal performance* for that channel. The performance measure of interest is usually the symbol error probability. These receivers are then used in parallel to perform a symbol-by-symbol demodulation of the received signal. For each sequence of transmitted symbols, therefore, there is one received sequence for each receiver in the parallel configuration. The task of identifying which receiver output is the most reliable reproduction of the transmitted sequence is referred to as the *data verification problem*. The main purpose of the formulation in [7] is to show that, even though the sufficient conditions in [5] are not satisfied for the class of Rician fading channels, the ideas of [5] can be used to obtain a finite set of receivers as described above for this channel class. The tradeoff between the allowed deviation from optimality and the number of receivers that are required in the parallel configuration is explored in [7]. However, the data verification problem is not considered in [7]. In this paper, we address the data verification problem for a parallel configuration consisting of the coherent and the noncoherent receivers.

Paper approved by J. Weitzel, the Editor for Fading/Equalization of the IEEE Communications Society. Manuscript received August 15, 1991; revised December 15, 1991. This work was supported by the Army Research Office under Grant DAAL-03-91-G-0154.

U. Madhow is with Bellcore, Morristown, NJ 07960 USA.

M. B. Pursley is with Clemson University, 102 Riggs Hall, Clemson, SC 29634 USA.

IEEE Log Number 9401828.

The data verification problem has been considered in a general setting in [5] where, assuming that the channel in effect does not change over the duration of a codeword, the error-detection capability of the error-control code is used to achieve a high probability of rejecting the outputs of mismatched receivers with a high probability. It is seen in [5] that Reed–Solomon codes with bounded distance decoding provide both error-correction and data verification if the error-detection capability of the code is enhanced by not using its full error-correction capability. However, in order to maintain a given error-correction capability, it is necessary to decrease the rate of the code to provide sufficient error-detection capability for data verification.

This approach is not quite satisfactory for the present application. We consider binary signaling in this paper, and binary codes do not have the good distance properties of Reed–Solomon codes, which employ much larger alphabets. A substantial loss in rate would thus be incurred in providing data verification for a universal receiver, and it would be better to use instead the noncoherent receiver, which is known to perform reasonably well for all Rician channels [11]. For a system employing the noncoherent receiver, the coding scheme is required mainly for error-correction, and it may have only moderate error-detection capability. A typical code may be a short binary BCH code. For comparing the universal receiver with the noncoherent receiver, we consider such a code. Data verification is achieved not by decreasing the rate of the code to increase its error-detection capability, but by supplementing its available error-detection capability by developing *side information from the demodulators*. Such side information is based on information within the receivers which is normally discarded once hard decisions are made, and it has the advantage that no transmission overhead is incurred.

The universal receiver with side information from the demodulators performs better than the noncoherent receiver. The performance gain is not substantial, being of the order of 1 dB, and may not justify the added complexity in practice. However, the results of [7] and this paper illustrate the important issues in designing universal receivers, and are intended to motivate exploration of more practical applications.

The system model is given in Section II. The performance of the system with coding but without additional side information is analyzed in Section III, and the results of the analysis are used to motivate the need for side information. A side-information statistic is developed in Section IV, and some numerical results are given to show its efficacy. The performance of the system with coding and side information is analyzed in Section V, and numerical results comparing the performance with that of the noncoherent receiver are given. Section VI contains our conclusions.

II. SYSTEM MODEL

The system model is as in [7] and [11], and further details can be found in those references. Binary orthogonal signaling is used, and the signals are narrowband and of duration T . The signal s_i , transmitted when the binary symbol i ($i = 0, 1$)

is sent, is given by

$$s_i(t) = (2E)^{1/2} x_i(t) \cos(2\pi f_c t + \phi), \quad 0 \leq t \leq T \quad (i = 0, 1)$$

where $x_0(t)$ and $x_1(t)$ are baseband waveforms satisfying

$$\int_0^T x_i(t) x_j(t) dt = \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases}$$

f_c is the carrier frequency, and ϕ is a fixed phase angle. The signals s_0 and s_1 are thus orthogonal and have equal energy E . It is assumed that 0 and 1 are sent with equal probability.

When symbol i is sent, the received signal is given by [7], [11]

$$r(t) = \alpha(2E)^{1/2} x_i(t) \cos(2\pi f_c t + \phi) + S(2E)^{1/2} x_i(t) \cos(2\pi f_c t + \phi + \Theta) + n(t).$$

The first term above is the specular component of the signal, and α is the channel gain for the specular component. The second term is the scatter component of the signal, and S and Θ are random variables with joint density

$$p(s, \theta) = \begin{cases} (s/2\pi\sigma^2) \exp(-s^2/2\sigma^2), & s \geq 0, \theta \in [0, 2\pi] \\ 0, & \text{otherwise.} \end{cases}$$

The third term is AWGN with a two-sided power spectral density of $N_0/2$. Without loss of generality, the energy E of the transmitted signal is normalized to 1 by replacing α by $\alpha E^{1/2}$ and σ^2 by $\sigma^2 E$.

When there is a specular component, the receiver is assumed to be locked to its phase ϕ , so that it can compute the following decision variables by suitable correlations with the received signal:

$$L_{ci} = \int_0^T r(t) 2^{1/2} x_i(t) \cos(2\pi f_c t + \phi) dt \quad (i = 0, 1), \quad (1)$$

and

$$L_{si} = \int_0^T r(t) 2^{1/2} x_i(t) \sin(2\pi f_c t + \phi) dt \quad (i = 0, 1). \quad (2)$$

The correlator outputs L_{c1} , L_{s1} , L_{c0} , and L_{s0} are sufficient statistics for deciding whether a 0 or a 1 has been transmitted.

The signal-to-noise ratio for the scatter component of the received signal is defined as $\beta = 2\sigma^2/N_0$. Denote by $N(m, v)$ a Gaussian distribution with mean m and variance v . The random variables L_{c1} , L_{s1} , L_{c0} , L_{s0} are conditionally independent and Gaussian, given that symbol i is sent. If a 1 is sent, the conditional distributions are given by

$$L_{c1} \sim N(\alpha, N_0(\beta + 1)/2), \quad L_{s1} \sim N(0, N_0(\beta + 1)/2), \quad (3a)$$

$$L_{c0} \sim N(0, N_0/2), \quad L_{s0} \sim N(0, N_0/2). \quad (3b)$$

If a 0 is sent, interchange L_{c1} and L_{c0} , and L_{s1} and L_{s0} , in (3).

The noncoherent receiver, which is optimal if there is no specular component (i.e., $\alpha = 0$), corresponds to the following decision rule.

$$L_{c1}^2 + L_{s1}^2 \underset{H_0}{\overset{H_1}{\geq}} L_{c0}^2 + L_{s0}^2. \quad (4)$$

The coherent receiver, which is optimal if there is no scatter component (i.e., $\beta = 0$), corresponds to the decision rule

$$L_{c1} \underset{H_0}{\overset{H_1}{\geq}} L_{c0}. \quad (5)$$

In general, the optimal decision rule depends on the values of α and β , and, for $\beta > 0$, is given by [7], [11], [12]

$$L_{c1}^2 + L_{s1}^2 + 2b_{\text{opt}}L_{c1} \underset{H_0}{\overset{H_1}{\geq}} L_{c0}^2 + L_{s0}^2 + 2b_{\text{opt}}L_{c0} \quad (6)$$

where

$$b_{\text{opt}} = \alpha/\beta. \quad (7)$$

Note that the noncoherent receiver (4) corresponds to $b_{\text{opt}} = 0$, and the coherent receiver corresponds to $b_{\text{opt}} = \infty$, in the sense that (5) is obtained from (6) by dividing both sides by b_{opt} , and letting $b_{\text{opt}} \rightarrow \infty$.

The triple (α, β, N_0) gives a complete statistical characterization of the channel, in the sense that knowledge of these parameters enables us to implement an optimal receiver and to analyze its performance. However, we assume that α and β are not known, and hence the optimal decision rule cannot be implemented. Instead, we consider a system in which the received signal is demodulated in parallel by the noncoherent receiver and the coherent receiver. Our purpose is to develop a method of selecting the output of the receiver that is performing better at a given time. It is assumed that a rough estimate of the noise power spectral density is available at the destination. Specifically, as in [7], N_0 is assumed to lie in a known interval $[N_{0,\min}, N_{0,\max}]$.

It is convenient to define the parameter $\gamma = \alpha/\sigma$, which is a measure of the relative strengths of the specular and scatter components of the signal, and the parameter $\text{SNR} = (\alpha^2 + 2\sigma^2)/N_0$, which is the overall signal-to-noise ratio. In the numerical results of Section V, we compare the performance of the universal receiver and that of the noncoherent receiver as a function of SNR for different values of γ .

The bit error probability at the output of the noncoherent receiver is [11]

$$p_{nc} = (\beta + 2)^{-1} \exp[-\alpha^2/N_0(\beta + 2)], \quad (8)$$

and the bit error probability at the output of the coherent receiver is [7]

$$p_{sp} = Q\{\alpha/[N_0(\beta + 2)/2]^{1/2}\} \quad (9)$$

where $Q(\cdot)$ is the complementary standard Gaussian distribution function. The task of data verification can be stated as follows: identify, at any given time, which of these two receivers is giving a lower symbol error probability, and choose the output of that receiver. However, data verification is not done on a symbol-by-symbol basis, since it is assumed that the channel parameters, and therefore the identity of the good receiver, do not change significantly over several symbol durations. Specifically, an error-control code is used, and it is assumed that the identity of the good receiver does not change over the duration of a codeword, so that the error-detection capabilities of the code can be exploited for data verification

[5]. This does not suffice for our application, however, and it is necessary to develop additional side information that indicates the relative magnitude of p_{nc} and p_{sp} over the duration of a given codeword.

The code symbols are transmitted using the orthogonal signals described earlier. It is assumed that, given the values of the channel parameters, the channel is memoryless. The received signal is demodulated symbol by symbol by both the coherent and the noncoherent receivers, so that there are two received words corresponding to each transmitted codeword. These are the two inputs to a decoding system whose task is to produce a good estimate for the transmitted codeword. Consider, first, a decoding system (introduced in [5]) that does not utilize any additional side information. Each received word is decoded separately in the usual manner by a bounded distance decoder. Each decoder produces one of three possible outcomes: correct decoding, decoding error, and decoding failure.

The results of the individual decodings are processed using the following *output selection algorithm*. Any output that fails to decode is ignored. If both outputs fail to decode, a failure is declared for the decoding system. If both outputs decode, but the resulting codewords do not match, again a failure is declared for the decoding system. In this case, at least one of the decoded words must be wrong, and since the identity of the good receiver is not known, we cannot choose between the two codewords. If only one of the received words decodes, or if both received words decode into the same codeword, the resulting decoded word is assumed to be correct, and a successful decoding is declared for the decoding system. Thus, if both the decoded words match, but are wrong, the result is a decoding error for the decoding system.

If additional side information is available, it is used to select one or both of the receivers (a fresh selection is made for every codeword). If both receivers are selected by the side information, then the output selection algorithm described above is employed. If only one receiver is selected, the output of the decoding system is simply the output of the decoder for the selected receiver.

Note that, if a receiver is poorly matched to the prevailing channel conditions, the corresponding received word may have a large number of errors. If this received word decodes incorrectly, then, even if the output of the other receiver is decoded correctly, we have a decoding failure for the decoding system if there is no additional side information. Thus, the code should be chosen such that even if there are a large number of errors in a received word, the decoder is much more likely to fail to decode rather than decode into a wrong codeword, so that, with high probability, its output is ignored by the output selection mechanism. Thus, the error-detection capability of the code is used to reject the output of a mismatched receiver. Since the short binary codes considered in this application are not powerful enough for this purpose, it is necessary to develop additional side information that selects, with high probability, the outputs of the receivers that are performing well, in order to direct only the corresponding decoder outputs to the output selection mechanism.

In the next section, we sketch the analysis of the performance of the decoding system with no side information. The results of the analysis are used to motivate the necessity for side information. The results are also used in Section V, where we analyze the performance with side information.

III. PERFORMANCE OF THE UNIVERSAL RECEIVER WITH NO SIDE INFORMATION

The performance measures of interest for the decoding system are the probability of decoding failure Q_F and the probability of decoding error Q_E . Given the values of the channel parameters, we can calculate the symbol error probabilities for each receiver using (8) and (9). Since the channel is memoryless, this gives the distribution of the number of errors in each received word. However, the performance of the decoding system clearly depends on the *joint* distribution of errors at the outputs of the two receivers. This distribution is difficult to calculate, and we develop bounds on the performance that depend only on the symbol error probabilities at the outputs of the receivers. This is consistent with our approach to the data verification problem, in which the goal is to choose the output of the receiver with the lower symbol error probability rather than to exploit the dependence of the outputs of different receivers. The analysis closely follows [5], and, rather than include all the details, we refer to [5] whenever possible. The bound on the decoding failure probability is the same as in [5], but, for our purpose, it is necessary to develop a tighter bound on the decoding error probability than in [5].

It is useful to number the two receivers in accordance with the symbol error probabilities at their output; the first receiver (also referred to as the good receiver) has the smaller symbol error probability $p_1 = \min\{p_{nc}, p_{sp}\}$, and the second receiver has the larger symbol error probability $p_2 = \max\{p_{nc}, p_{sp}\}$. An (n, k) binary code with minimum distance d is used, and the decoders in the decoding system attempt to correct up to t errors where $2t + 1 \leq d$. Let $P_E(u)$ denote the conditional probability of decoding error for a particular decoder given that there are u errors in the decoder input.

Because the channel is memoryless, the performance of the decoder at the output of a particular receiver depends only on the symbol error probability for that receiver. Since n is the number of symbols in a codeword, if the symbol error probability for the receiver is p , the number of errors in the received word is a binomial random variable with parameters n and p . The probability that the decoder for that receiver does not decode correctly, which is a tight upper bound for the probability of decoding failure for that decoder, is given by

$$P_{DF}(p) = \sum_{u=t+1}^n \binom{n}{u} p^u (1-p)^{(n-u)}.$$

The probability of decoding error for the same situation is given by

$$P_{DE}(p) = \sum_{u=d-t}^n P_E(u) \binom{n}{u} p^u (1-p)^{(n-u)}.$$

These functions appear in the bounds on Q_F and Q_E given in the following.

The bound on the probability of decoding failure for the decoding system is given by [5]

$$Q_F \leq P_{DF}(p_1) + P_{DE}(p_2). \quad (10)$$

For the probability of decoding error for the decoding system, we have [5]

$$Q_E \leq P[\text{decoder for good receiver decodes incorrectly}] + P \left[\begin{array}{l} \text{decoder for good receiver} \\ \text{does not decode correctly} \\ \text{and the other decoder} \\ \text{decodes incorrectly} \end{array} \right]. \quad (11)$$

The first term equals $P_{DE}(p_1)$. Some additional notation is required to give an explicit bound for the second term. Let U_i be the number of code symbols in error at the output of the i th receiver. As discussed earlier, U_i is a binomial random variable with parameters n and p_i . The second term depends on the distribution of the random vector $\mathbf{U} = (U_1, U_2)$, which in turn depends on the joint distribution of errors at the outputs of the two receivers, but in the following, we give a bound on the second term that depends only on the marginal distributions of U_1 and U_2 .

Define the indicator function $I_t(u)$ by

$$I_t(u) = \begin{cases} 0 & u \leq t, \\ 1 & u \geq t+1. \end{cases}$$

The conditional probability corresponding to the second term in (11) is

$$P \left[\begin{array}{l} \text{decoder for good receiver} \\ \text{does not decode correctly} \\ \text{and the other decoder} \\ \text{decodes incorrectly} \end{array} \middle| \mathbf{U} = \mathbf{u} \right] = I_t(u_1) P_E(u_2).$$

Removing the conditioning, the second term in (11) can be bounded using Holder's inequality [9, p. 113] as follows:

$$E\{I_t(U_1)P_E(U_2)\} \leq \|I_t(U_1)\|_q \|P_E(U_2)\|_r \quad (12)$$

where $\|X\|_q$ denotes the L_q norm [9, pp. 111–112] of the random variable X , $1 \leq q \leq \infty$, $1 \leq r \leq \infty$, and $q^{-1} + r^{-1} = 1$.

The expressions for the norms of interest are as follows:

$$\begin{aligned} \|I_t(U_1)\|_q &= \left\{ \sum_{u=t+1}^n \binom{n}{u} p_1^u (1-p_1)^{(n-u)} \right\}^{1/q} = [P_{DF}(p_1)]^{1/q}, \\ &= \begin{cases} \left\{ \sum_{u=t+1}^n \binom{n}{u} p_1^u (1-p_1)^{(n-u)} \right\}^{1/q}, & 1 \leq q < \infty, \\ 1, & p_1 > 0, q = \infty, \\ 0, & p_1 = 0, \end{cases} \end{aligned}$$

and

$$\begin{aligned} \|P_E(U_2)\|_r &= \left\{ \sum_{u=d-t}^n [P_E(u)]^r \binom{n}{u} p_2^u (1-p_2)^{(n-u)} \right\}^{1/r}, \\ &= \begin{cases} \left\{ \sum_{u=d-t}^n [P_E(u)]^r \binom{n}{u} p_2^u (1-p_2)^{(n-u)} \right\}^{1/r}, & 1 \leq r < \infty, \\ \max_u P_E(u), & p_2 > 0, r = \infty, \\ 0, & p_2 = 0. \end{cases} \end{aligned}$$

Define $H(p_1, p_2)$ as

$$H(p_1, p_2) = \min \left\{ \|I_t(U_1)\|_q \|P_E(U_2)\|_r \mid 1 \leq q, r \leq \infty, q^{-1} + r^{-1} = 1 \right\}. \quad (13)$$

Using (12) and (13), the second term in the summation in (11) is bounded by $H(p_1, p_2)$, so that

$$Q_E \leq P_{DE}(p_1) + H(p_1, p_2). \quad (14)$$

It is worth noting that, instead of performing the optimization in (13), we could set $q = 1$, $r = \infty$ in (12) to get

$$E\{I_t(U_1)P_E(U_2)\} \leq P_{DF}(p_1) \left[\max_u P_E(u) \right].$$

Since the right-hand side is independent of p_2 , this results in a much simpler worst-case bound for Q_E as follows:

$$Q_E \leq P_{DE}(p_1) + P_{DF}(p_1) \left[\max_u P_E(u) \right].$$

This is the bound that is given in [5]. However, while this is a good bound for the Reed–Solomon codes considered there, for the binary BCH codes considered in this paper, $\max_u P_E(u)$ can be very large; for instance, it equals one if the all-ones word is in the code. Thus, it is necessary to use the more complicated bound of (14).

For our system, the important measure of performance is the decoding failure probability Q_F . The decoding error probability Q_E is important only in that it is required to be several orders of magnitude lower than Q_F for a good system design. If there is no scatter component ($\gamma = \infty$), the coherent receiver is optimal, in which case $p_1 = p_{sp}$ and $p_2 = p_{nc}$. However, the noncoherent receiver performs well enough that the probability of decoding failure for the decoding system is governed by the decoding failure probability for the coherent receiver, and not by the decoding error probability for the noncoherent receiver. That is, $P_{DE}(p_2) \ll P_{DF}(p_1)$ in (10), which implies that $Q_F \approx P_{DF}(p_1)$. This means that for $\gamma = \infty$ the performance of the universal receiver is given by that of the good receiver, even without additional side information. This comment also applies for large finite values of γ .

Consider, however, the situation in which there is only a scatter component (i.e., $\gamma = 0$). The noncoherent receiver is optimal, and $p_1 = p_{nc}$, but the coherent receiver performs very poorly, with $p_2 = p_{sp} = 0.5$. The output of the coherent receiver has a relatively large probability of decoding into an incorrect codeword, so that, even if the output of the noncoherent receiver decodes correctly, we have a decoding failure for the decoding system. This corresponds to the second term in (10) being dominant. Note that this term, which takes the value $P_{DE}(0.5)$ for $\gamma = 0$, does not decrease as SNR increases. On the other hand, the decoding failure probability for the decoder for the noncoherent receiver is given by $P_{DF}(p_{nc})$, which does decrease with SNR. Thus, in this situation, the noncoherent receiver gives better performance than the universal receiver with no side information, and it is necessary to develop additional side information about the

identity of the good receiver in order for the universal receiver to be competitive with the noncoherent receiver.

The requirements on the side information are given as follows. Since the performance of the coherent receiver deteriorates rapidly as γ decreases, it is required that, with high probability, the side information select only the noncoherent receiver for small values of γ . On the other hand, in order to exploit the presence of the coherent receiver for large values of γ , it is required that, with high probability, the selection include the coherent receiver (the noncoherent receiver can also be included in the selection without affecting the decoding failure probability). In the following section, we develop side information that satisfies these requirements.

IV. DEVELOPMENT OF THE SIDE-INFORMATION STATISTIC

Ideally, we would like to have perfect side information, which corresponds to selecting the receiver that has the lowest bit error probability at its output. The side information developed here is a consequence of attempting to achieve this goal using the limited information at our disposal.

Using the bound [12, p. 39]

$$Q(x) \leq \exp(-x^2/2)/x(2\pi)^{1/2},$$

it follows from (9) that

$$p_{sp} \leq \frac{N_0(\beta + 2)^{1/2}}{(2\pi)^{1/2}\alpha} \exp[-\alpha^2/N_0(\beta + 2)]. \quad (15)$$

It is convenient to compare the expression (8) for p_{nc} with the upper bound (15) for p_{sp} rather than with the original expression (9). Since the comparison is biased in favor of the noncoherent receiver, this is also a more conservative approach, because the noncoherent receiver performs quite well for all values of the channel parameters. Note that (15) is used only for the purpose of developing the side information; in all computations involving p_{sp} , the exact value (9) is used.

Our objective, therefore, is for the side information to select the coherent receiver if the value of the bound (15) on p_{sp} is less than the value of p_{nc} given by (8); otherwise, it should select the noncoherent receiver. Defining $T = \alpha/[N_0(\beta + 2)/2]^{3/2}$, if α , β , and N_0 are known, the decision rule can be expressed as follows: select the coherent receiver if $T > (2/\pi)^{1/2}N_0^{-1}$; and select the noncoherent receiver if $T \leq (2/\pi)^{1/2}N_0^{-1}$. Since these parameters are not known, however, this decision rule is approximated by estimating T and using the fact that $N_{0,\min} \leq N_0 \leq N_{0,\max}$.

To this end, let $\hat{T}$ be an estimate for T . The statistic $\hat{T}$ is referred to as the *side-information statistic*. Define $t_* = (2/\pi)^{1/2}N_{0,\max}^{-1}$ and $t^* = (2/\pi)^{1/2}N_{0,\min}^{-1}$. The decision rule is to select only the coherent receiver if $\hat{T} \geq t^*$, select only the noncoherent receiver if $\hat{T} \leq t_*$, and select both receivers if $t_* < \hat{T} < t^*$. Lack of knowledge of the channel parameters affects the decision rule for the side information in two ways: the estimate $\hat{T}$ is used instead of the actual value of T , and the choice of both receivers is allowed for certain values of $\hat{T}$. The outputs of the selected receivers are the input to the decoding system described in Section II.

In the following, we develop a side-information statistic $\hat{T}$ and give its distribution. The selection probabilities are expressed in terms of this distribution, and some numerical results are given to show that the side-information statistic does satisfy the requirements described in the previous section. In the next section, we analyze the performance of the decoding system when this side information is used with a binary code, and we give numerical results for a (32, 16) extended BCH code.

Consider a particular transmitted codeword. Assume that the channel parameters do not change during the transmission of n_s consecutive bits, which include the n bits of the codeword. These bits are indexed by i , which ranges from one to n_s . Let $\mathbf{Z}_i = (L_{c1}^{(i)}, L_{s1}^{(i)}, L_{c0}^{(i)}, L_{s0}^{(i)})$ be random vectors whose components are the correlator outputs that correspond to the i th bit. Since the channel is memoryless, these random vectors are independent and identically distributed, and their distribution depends on the channel parameters only. One can use the $\mathbf{Z}_i$ to attempt to estimate the channel parameters α , β , and N_0 , but for our purposes, it suffices to estimate T . The distribution of the $\mathbf{Z}_i$ depends on whether the i th transmitted bit was a 0 or a 1, and this makes a maximum likelihood estimate of T very complicated. It is preferred, therefore, to use a very simple estimate of T based on a function of the $\mathbf{Z}_i$ that does not depend on the sequence of transmitted bits.

Define the positive real variable λ by $\lambda^2 = N_0(\beta + 2)/2$, and define X_i by $X_i = L_{c1}^{(i)} + L_{c0}^{(i)}$. From (3), we see that the distribution of X_i is independent of whether the i th bit is a 0 or a 1, and is given by

$$X_i \sim N(\alpha, \lambda^2). \quad (16)$$

Then

$$T = \alpha/\lambda^3. \quad (17)$$

The random variables X_i , $1 \leq i \leq n_s$, are independent and identically distributed. Thus, if $\bar{X}$ and S^2 are estimates of the mean and variance of the X_i , respectively, then (16) and (17) suggest the estimate

$$\hat{T} = \bar{X}/S^3. \quad (18)$$

The statistic $\bar{X}$ used here is the sample mean $\bar{X} = n_s^{-1} \sum_{i=1}^{n_s} X_i$. It is easy to see that $\bar{X} \sim N(\alpha, \lambda^2 n_s^{-1})$. The estimate S^2 for the variance is chosen from a class of estimates S_r^2 , $r \geq 0$, defined by $S_r^2 = (n_s - r)^{-1} \sum_{i=1}^{n_s} (X_i - \bar{X})^2$. The standard unbiased estimate for the variance corresponds to $r = 1$ [10, p. 73]. Since the purpose here is to estimate T using (18), however, it is more appropriate to choose r such that $\bar{X}/S_r^3$ is an unbiased estimate of T up to first order in n_s^{-1} . It is not possible to make $\hat{T} = \bar{X}/S_r^3$ a truly unbiased estimate of T .

Thus, it is required to choose r to satisfy

$$E\{\bar{X}/S_r^3\} = \alpha/\lambda^3 + O(n_s^{-2}), \quad n_s \rightarrow \infty. \quad (19)$$

It is shown in the Appendix that $r = 3.5$ satisfies (19), and does not depend on either α or λ . Thus, we consider the side-information statistic $\hat{T} = \bar{X}/S^3$, with $S^2 = (S_{3.5})^2$, so that $E\{\hat{T}\} = T + O(n_s^{-2})$, $n_s \rightarrow \infty$.

It remains to specify the distribution of $\hat{T}$. The density of S^3 is shown in the Appendix to be

$$p_{S^3}(z) = \frac{(n_s - 3.5)^{(n_s-1)/2}}{3\lambda^{n_s-1} 2^{(n_s-3)/2} \Gamma[(n_s-1)/2]} z^{(n_s-4)/3} \cdot \exp[-\frac{1}{2}(n_s - 3.5)z^{2/3}] \quad (20)$$

for $z \geq 0$. Since the sample mean $\bar{X} \sim N(\alpha, \lambda^2 n_s^{-1})$, and is independent of the sample variance S^3 [3, p. 356], the distribution $\hat{F}$ of $\hat{T}$ is given by

$$\hat{F}(t) = P[\bar{X}/S^3 \leq t] = \int_0^\infty \Phi[(zt - \alpha)/n_s^{-1/2}\lambda] p_{S^3}(z) dz \quad (21)$$

where $\Phi(\cdot)$ is the standard Gaussian distribution function.

It can be shown [6] that $\hat{T}$ is an asymptotically normal estimate of T , in the sense that $n_s^{1/2}(\hat{T} - T)$ tends in distribution to a zero mean Gaussian random variable with variance $\xi^2 = (\lambda^2 + 4.5\alpha^2)/\lambda^6$ as $n_s \rightarrow \infty$. This can be used to approximate the distribution of $\hat{T}$ by a Gaussian distribution with mean T and variance $\xi^2 n_s^{-1}$ for finite n_s . However, the values of n_s that are of interest here are relatively small, so that the tails of the actual distribution of $\hat{T}$ decay much more slowly than those of the Gaussian distribution. Thus, it is necessary to use the exact distribution of $\hat{T}$ in computing the selection probabilities in the following.

The selection probabilities for the various receivers are expressed in terms of the distribution of $\hat{T}$ as follows. The probability of selecting only the coherent receiver is

$$P_S(sp) = P[\hat{T} \geq t^*] = 1 - \hat{F}(t^*); \quad (22)$$

the probability of selecting only the noncoherent receiver is

$$P_S(nc) = P[\hat{T} \leq t_*] = \hat{F}(t_*); \quad (23)$$

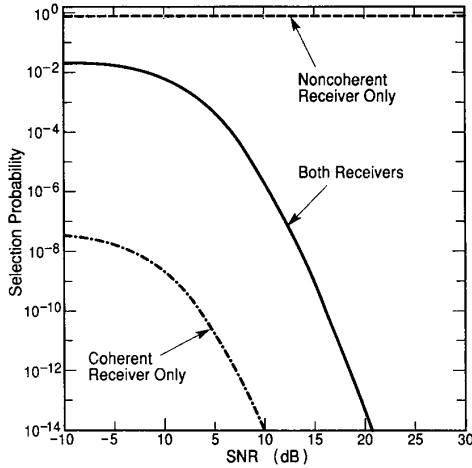
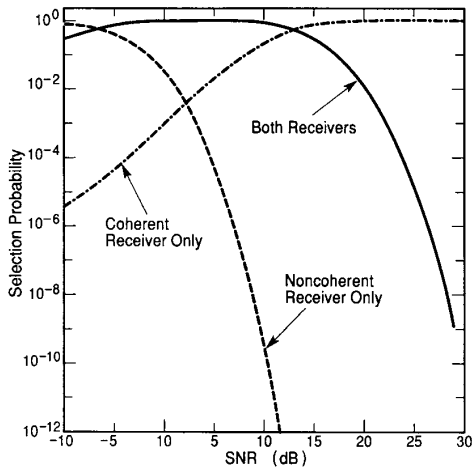
and the probability of selecting both receivers is

$$P_S(\text{both}) = P[t_* < \hat{T} < t^*] = \hat{F}(t^*) - \hat{F}(t_*). \quad (24)$$

These probabilities can be computed using (21).

We now give some numerical results that demonstrate the efficacy of the side information developed above. The numerical results of the next section are for a code with a blocklength of 32. Assuming that only correlator outputs corresponding to bits of the codeword are used for computing the side-information statistic, the value $n_s = 32$ is appropriate. However, the value $n_s = 31$ is used for computational convenience, because the argument of the gamma function in (20) is an integer for odd values of n_s (so that $\Gamma[(n_s-1)/2] = [(n_s-3)/2]!$). We consider $N_{0,\min} = 1$ and $N_{0,\max} = 10$. The selection probabilities given by (22)–(24) are shown as a function of SNR in Figs. 1 and 2 for two representative values of γ , $\gamma = 0$, and $\gamma = \infty$, respectively. Note that, while the error probabilities for the two receivers depend on γ and the SNR alone, the selection probabilities depend on the value of N_0 as well. A single value of $N_0 = 5$ is considered here, but the results obtained are typical for all values of N_0 in the range of interest [6].

The numerical results displayed in Figs. 1 and 2, together with other numerical results in [6], show that the side information from the demodulators satisfies the requirements discussed

Fig. 1. Selection probabilities for $\gamma = 0$.Fig. 2. Selection probabilities for $\gamma = \infty$.

at the end of Section III. The undesirable selection events for $\gamma = 0$ are selecting only the coherent receiver and selecting both receivers. We see from Fig. 1 that the probabilities of these undesirable events, $P_S(sp)$ and $P_S(both)$, respectively, are small, as desired. For $\gamma = \infty$, the side information should, with high probability, select either both receivers or the coherent receiver alone to realize the advantage in SNR from using the two receivers in parallel instead of using the noncoherent receiver alone. We see from Fig. 2 that the side information from the demodulators does satisfy this property.

V. PERFORMANCE OF THE UNIVERSAL RECEIVER WITH SIDE INFORMATION

The performance of the decoding system is analyzed by first conditioning on the selection event, and then removing the conditioning. However, the correlator outputs for the bits of the codeword are being used to provide side information to the decoding system. Therefore, when conditioned on the selection event, the bit errors in the codeword are not conditionally independent, so that the channel is not memoryless. Moreover,

the conditional probabilities of bit error for the two receivers are not equal to the unconditional error probabilities given by the expressions (8) and (9). We will, however, use an approximation to the performance of the decoding system. To obtain this approximation, one assumes that, conditioned on the selection event, the channel is memoryless, with the conditional bit error probabilities for the two receivers given by the expressions (8) and (9). The dependence of bit errors due to conditioning on the side-information statistic $\hat{T} = \bar{X}/S^3$ is weak, because the contribution of the correlator outputs for any given bit to either $\bar{X}$ or S^2 is attenuated by a factor n_s^{-1} . Thus, the approximation should be quite good. Further, the true performance is expected to be better than that predicted by the approximation, and an intuitive justification for this belief is given in the following.

The correlator outputs that are used to make the symbol-by-symbol hard decisions at each receiver are also used to select one or both of the receivers. If the good receiver is selected (i.e., the right decision is made), the correlator outputs used for the selection are likely to be less noisy, so that the performance of the good receiver conditioned on its selection should be better than the unconditional performance used in the approximation. Because the probability of selecting the bad receiver is much smaller than that of selecting the good receiver, the main contribution to the overall performance comes from the term that corresponds to the selection of the good receiver. Consequently, the approximate analysis is conjectured to give upper bounds for the decoding failure probability and the decoding error probability for the universal receiver with side information from the demodulators. Because the channel is memoryless, the approximate analysis is exact if the n_s bits used for the side information do not include any of the bits in the codeword (for instance, if the n_s bits immediately preceding the codeword are used). Intuitively, we expect that by using the bits of the codeword we get better side information about how well each receiver performs over the duration of the codeword than if we used bits not in the codeword.

We denote by $Q(F | J)$ and $Q(E | J)$ the conditional probabilities of decoding failure and decoding error, respectively, given the selection event J . In our notation, $J = sp$ denotes the event that the coherent receiver is selected, $J = nc$ the event that the noncoherent receiver is selected, and $J = both$ the event that both receivers are selected. The performance of the decoding system with side information is therefore given by

$$Q_F = P_S(sp)Q(F | sp) + P_S(nc)Q(F | nc) + P_S(both)Q(F | both), \quad (25)$$

and

$$Q_E = P_S(sp)Q(E | sp) + P_S(nc)Q(E | nc) + P_S(both)Q(E | both). \quad (26)$$

The approximation discussed above is used at this point (for simplicity, this is not reflected in the notation). We have $Q(F | sp) \leq P_{DF}(p_{sp})$ and $Q(F | nc) \leq P_{DF}(p_{nc})$, and $Q(E | sp) = P_{DE}(p_{sp})$ and $Q(E | nc) = P_{DE}(p_{nc})$. If both

receivers are selected, the performance is identical to that of a system with no side information; that is, upper bounds on $Q(F | \text{both})$ and $Q(E | \text{both})$ are given by (10) and (14), respectively. The performance measures Q_F and Q_E can now be computed using (25) and (26).

The numerical results are for a (32, 16) triple-error-correcting extended BCH code with minimum distance 8, obtained by adding an extra parity check bit to the codewords of a (31, 16) triple-error-correcting primitive BCH code. The (32, 16) code is preferred over the (31, 16) code because of its enhanced error-detection capability (at the expense of a single extra parity check bit), but we have found that there is not a very significant difference in the performance of these two codes. The performance of the universal receiver with side information is to be compared to that of the noncoherent receiver alone. Clearly, in the latter instance, the full error correction capability of the code would be used, so $t = 3$ is the appropriate choice for ensuring a fair comparison. We need the conditional probabilities of decoding error $P_E(u)$ in our analysis (see Section III). These can be obtained from the weight distribution of the code [4]. The weight distribution is easily obtained from tables in the literature (see, for instance, [2]). As in the previous section, $n_s = 31$, $N_{0, \min} = 1$, $N_{0, \max} = 10$, and $N_0 = 5$. The performance of three receivers is shown as a function of SNR: the noncoherent receiver, the universal receiver with side information from the demodulators, and the universal receiver with no side information. As noted earlier, the decoding failure probability is a more important performance measure than the decoding error probability, hence we include a graph of the latter in only one of the situations considered in the following (see [6] for additional numerical results).

The decoding failure probability for $\gamma = 0$ is plotted in Fig. 3(a), and the decoding error probability in Fig. 3(b). The performance with no side information is, of course, quite bad, because the coherent receiver performs very poorly in this situation. The performance of the universal receiver with side information from the demodulators, however, is indistinguishable from that of the (optimal) noncoherent receiver. These observations hold for small values of γ as well; as an example, we show the decoding failure probability for $\gamma = 2$ in Fig. 4. Fig. 5 shows the decoding failure probability for $\gamma = \infty$. The performance of the universal receiver with no side information and that of the universal receiver with side information from the demodulators is almost the same, and both these systems are better than the noncoherent receiver. This property is also typical of large finite values of γ , as is illustrated by Fig. 6, which shows the decoding failure probability for $\gamma = 20$.

Another way of comparing the performance of the noncoherent receiver and the universal receiver with side information from the demodulators is in terms of the value of SNR required to attain a given performance level for each receiver. A comparison of the SNR required for these two receivers to attain a decoding failure probability of 10^{-10} is given in Table I (each entry is rounded to a tenth of a dB). The universal receiver with side information from the demodulators is better by almost 1 dB for high to moderate values of γ , and it is no worse for low values of γ .

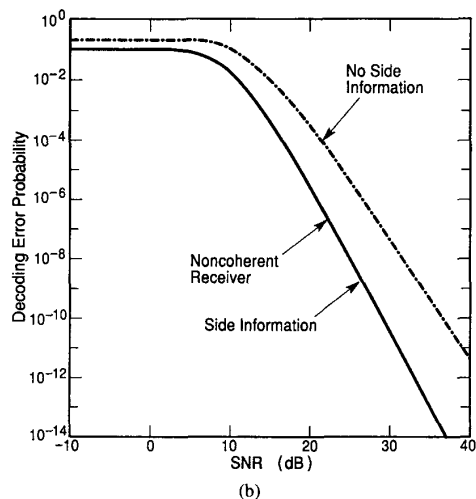
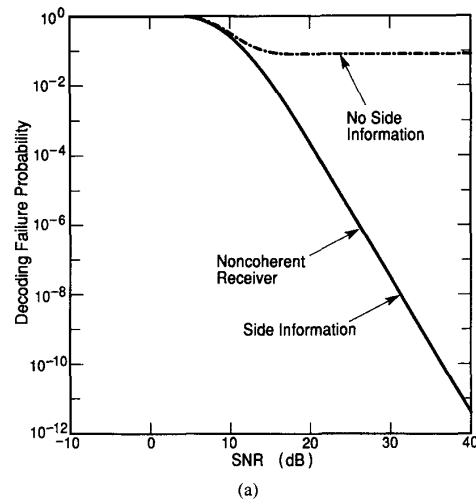


Fig. 3. (a) Decoding failure probability for $\gamma = 0$. (b) Decoding error probability for $\gamma = 0$.

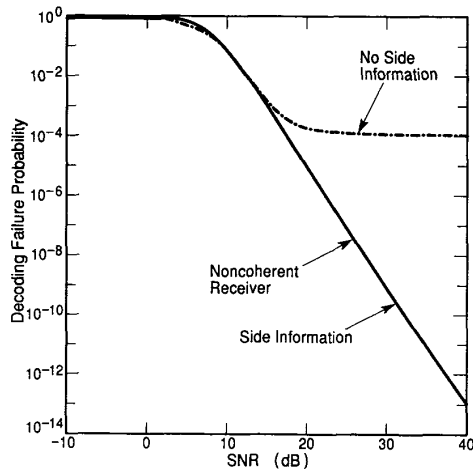
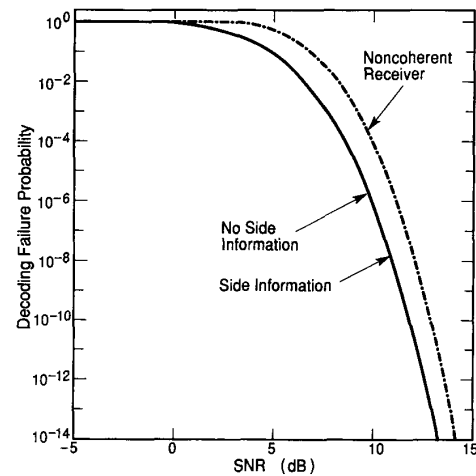
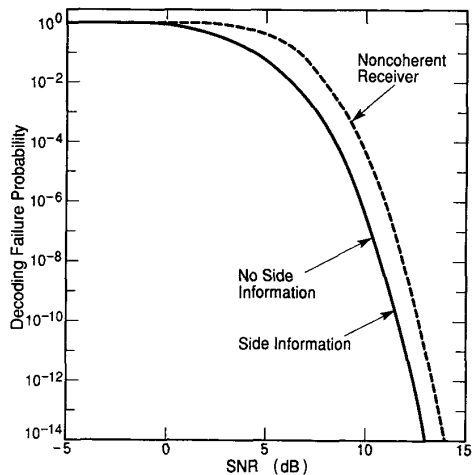
Our results demonstrate that the universal receiver with side information from the demodulators provides a gain in SNR over the noncoherent receiver. They also show that the universal receiver with no side information performs better than the noncoherent receiver for large γ , but it suffers a significant degradation in performance for small γ .

VI. CONCLUSIONS

We have developed side information from the demodulators for a universal receiver for Rician-fading channels, and we have shown that the performance of the resulting system is superior to that of the noncoherent receiver. As noted in the introduction, the example of Rician fading channels is chosen mainly as an illustration of the universal approach, even though the potential gains are limited to about 1 dB. However, the results of the paper are encouraging for precisely the same reason: even though the noncoherent receiver gives such good performance, it is possible to do better using a universal

TABLE I
SNR (dB) REQUIRED TO ATTAIN A DECODING FAILURE PROBABILITY OF 10^{-10} FOR THE NONCOHERENT RECEIVER AND THE UNIVERSAL RECEIVER WITH SIDE INFORMATION FROM THE DEMODULATORS

γ	Noncoherent receiver	Universal receiver	Gain in SNR (dB)
0	36.4	36.4	0
2	32.5	32.5	0
3	24.8	24.8	0
4	17.9	17.7	0.2
5	15.2	14.2	1.0
10	12.6	11.6	1.0
20	12.0	11.1	1.0
∞	11.9	10.9	0.9

Fig. 4. Decoding failure probability for $\gamma = 2$.Fig. 6. Decoding failure probability for $\gamma = 20$.Fig. 5. Decoding failure probability for $\gamma = \infty$.

approach and some ingenuity in developing side information. It is hoped that our results will motivate consideration of the application of universal receivers to practical problems in which potential performance gains are much greater.

Since the demodulation side information depends on estimating a function T of the channel parameters, it is appropriate

to conclude this paper by comparing the universal receiver that employs demodulation side information with the following adaptive system. Since the optimal receiver b_{opt} for a channel (α, β, N_0) is given by $b_{\text{opt}} = \alpha/\beta$ [see (7)], estimates of the channel parameters α and β could be used to estimate the receiver parameter b_{opt} , and the resulting receiver used to demodulate the received signal. One major drawback of this procedure is that it cannot be carried out in real time; before making any decisions about the transmitted bits, it is necessary to make an estimate of b_{opt} based on the corresponding received waveforms. For the universal receiver, on the other hand, the receivers in the parallel configuration make the hard decisions on the bits as the corresponding signals are demodulated, so that the decoders can proceed to decode a given received word as soon as it is received. The side-information statistic can be obtained by the time the decoding is completed, and the output selection algorithm can use this side information along with the outputs of the decoders in making its decision. Another problem with the adaptive approach is that, since the channel parameters may vary with time, only a moderate number of samples can be used to make the estimate. The choice of the optimal receiver is probably more sensitive to errors in estimation than is the demodulation side information used for the universal receiver. Finally, the implementation of an adaptive receiver of this kind is much more complicated than that of a universal receiver with demodulation side information.

APPENDIX

We first show that $r = 3.5$ satisfies (19), and then derive the density of $S^3 = S_{3.5}^3$. Standard results [3, p. 356] for the sample mean and sample variance for Gaussian random variables imply that S_r and $\bar{X}$ are independent, and that the distribution of S_r does not depend on α . This yields

$$E\{\bar{X}/S_r^3\} = E\{\bar{X}\}E\{1/S_r^3\} = \alpha E\{1/S_r^3\},$$

so that it is enough to choose r to satisfy

$$E\{1/S_r^3\} = 1/\lambda^3 + O(n_s^{-2}), \quad n_s \rightarrow \infty. \quad (\text{A.1})$$

It is clear from the foregoing that the value of r satisfying (A.1) does not depend on α . Moreover, it is seen by a simple scaling argument that the required value of r does not depend on λ , either. In order to find the desired value of r , therefore, we can, without loss of generality, let $\alpha = 0$ and $\lambda = 1$. For these values, (16) implies that the independent and identically distributed samples X_i are standard Gaussian. The distribution of S_0 is then given by [3, p. 356]

$$f(s) = A s^{n_s-2} \exp(-n_s s^2/2), \quad (\text{A.2})$$

for $s \geq 0$. The normalization constant A is given by

$$A = 2^{(n_s-3)/2} \Gamma[(n_s-1)/2] / n_s^{(n_s-1)/2}$$

where Γ is the gamma function [1, p. 253].

The density of S_0 is now used to obtain

$$E\{1/S_0^3\} = (n_s/2)^{3/2} \frac{\Gamma((n_s-4)/2)}{\Gamma((n_s-1)/2)}. \quad (\text{A.3})$$

An asymptotic relation for the gamma function [1, p. 257] implies

$$\begin{aligned} z^{b-a} \frac{\Gamma(z+a)}{\Gamma(z+b)} \\ = 1 + \frac{(a-b)(a+b-1)}{2z} + O(z^{-2}), \quad z \rightarrow \infty. \end{aligned}$$

Setting $z = n_s/2$, $a = -2$, and $b = -\frac{1}{2}$ in (A.3) gives

$$E\{1/S_0^3\} = 1 + \frac{21}{4} n_s^{-1} + O(n_s^{-2}), \quad n_s \rightarrow \infty.$$

Since $(n_s - r)S_r^2 = n_s S_0^2$, it follows that, for $n_s \rightarrow \infty$,

$$E\{1/S_r^3\} = (1 - r n_s^{-1})^{3/2} \left(1 + \frac{21}{4} n_s^{-1} + O(n_s^{-2}) \right),$$

which yields

$$E\{1/S_r^3\} = 1 + \left(\frac{21}{4} - \frac{3}{2} r \right) n_s^{-1} + O(n_s^{-2}).$$

Thus, setting $r = 3.5$ cancels the first-order term in n_s^{-1} .

We now return to the original problem, in which $\alpha \geq 0$ and $\lambda^2 > 0$ are arbitrary. The density of S^3 is obtained from the preceding results for standard Gaussian samples as follows. With a slight change of notation, let S'_0 denote a random variable with density f given by (A.2). Then

$$S^3 = \lambda^3 \left(\frac{n_s}{n_s - 3.5} \right)^{3/2} (S'_0)^3.$$

A change of variables in (A.2) now yields the desired expression (20) for the density of S^3 .

REFERENCES

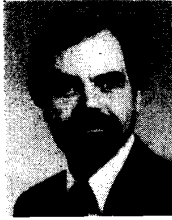
- [1] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*. New York: Dover, 1970.
- [2] E. R. Berlekamp, *Algebraic Coding Theory*. New York: McGraw-Hill, 1968.
- [3] M. Kendall, A. Stuart, and J. K. Ord, *Kendall's Advanced Theory of Statistics*. Oxford, England: Oxford University Press, 1987, vol. 1.
- [4] F. J. MacWilliams, "A theorem on the distribution of weights in a systematic code," *Bell Syst. Tech. J.*, vol. 42, pp. 79-94, Jan. 1963.
- [5] U. Madhow and M. B. Pursley, "A parallel systems approach to universal receivers," *IEEE Trans. Inform. Theory*, vol. 37, pp. 291-306, Mar. 1991.
- [6] U. Madhow, "A parallel systems approach to universal receivers," Ph.D. dissertation, Univ. Illinois at Urbana-Champaign, Aug. 1990.
- [7] U. Madhow and M. B. Pursley, "On the design of universal receivers for nonselective Rician fading channels," *IEEE Trans. Commun.*, to be published, vol. 42, Sept. 1994.
- [8] M. B. Pursley and U. Madhow, "Parallel structures for universal communication receivers," in *Proc. 1990 Bilkent Int. Conf. New Trends Commun. Contr. Signal Processing*, Bilkent Univ., Ankara, Turkey, July 2-5, 1990, vol. 1, pp. 215-227.
- [9] H. L. Royden, *Real Analysis*. New York: Macmillan, 1968.
- [10] R. J. Serfling, *Approximation Theorems of Mathematical Statistics*. New York: Wiley, 1980.
- [11] G. L. Turin, "Error probabilities for binary symmetric ideal reception through nonselective slow fading and noise," *Proc. IRE*, vol. 46, pp. 1603-1619, Sept. 1958.
- [12] H. L. Van Trees, *Detection, Estimation and Modulation Theory*. New York: Wiley, 1968, vol. 1.



Upamanyu Madhow received the B.S. degree in electrical engineering from the Indian Institute of Technology, Kanpur, in 1985, and the M.S. and Ph.D. degrees also in electrical engineering from the University of Illinois, Urbana-Champaign, in 1987 and 1990, respectively.

From August 1990 to July 1991, he was a Visiting Assistant Professor at the University of Illinois. Since August 1991, he has been a Member of the Technical Staff with Bell Communications Research. His current research interests are in communication systems and networks for wireless mobile communications, and high speed computer communication networks.

Dr. Madhow was awarded the President of India Gold Medal for graduating at the top of his undergraduate class. He was the recipient of a University of Illinois fellowship from 1985 to 1986, and a Schlumberger fellowship from 1987 to 1988.



Michael B. Pursley (S'68-M'68-SM'77-F'82) was born in Winchester, IN, on August 10, 1945. He received the B.S. degree (with highest distinction) in electrical engineering and the M.S. degree from Purdue University, Lafayette, IN, in 1967 and 1968, respectively, and the Ph.D. degree in electrical engineering in 1974 from the University of Southern California.

He has several years of industrial experience with the Nortronics Division of Northrop Corporation and the Space and Communications Group of the Hughes Aircraft Company. He was Hughes Doctoral Fellow and a Research Assistant in the Department of Electrical Engineering at the University of Southern California. From January through June of 1974 he was an Acting Assistant Professor in the System Science Department of the University of California, Los Angeles. From June 1974 to July 1993, he was with the Department of Electrical and Computer Engineering and the Coordinated

Science Laboratory at the University of Illinois, Urbana, where he held the rank of Professor since 1980. His research is in the general area of communications and information theory with emphasis on spread-spectrum communications, correlation properties of sequences, mobile radio networks, and multiple-access communication theory.

Dr. Pursley has served as Holcombe Endowed Professor at Clemson University, Clemson, SC. He is a member of Phi Eta Sigma, Tau Beta Pi, and the Institute of Mathematical Statistics. He was a member of the Board of Governors of the IEEE Information Theory Group during 1977-1984 and 1989-1991, and in 1982 he served as the President of the Group. He has served as Program Chairman for the 1979 IEEE International Symposium on Information Theory which was held in Grignano, Italy, and he was a member of the Editorial Board of the PROCEEDINGS OF THE IEEE for the period 1984-1991. He was awarded an IEEE Centennial Medal in 1984. He is Co-Chairman for the 1995 IEEE International Symposium on Information Theory.