# On the Design of Universal Receivers for Nonselective Rician-Fading Channels

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

***Abstract*—The purpose of this paper is to illustrate the issues involved in designing the demodulator portion of a universal receiver for unknown or time-varying channels by means of a specific example. We consider the class of nonselective Rician fading channels with additive white Gaussian noise. The optimal receiver for a Rician channel depends on the parameters of the channel, and the collection of optimal receivers for channels in the class of interest forms an infinite receiver class. We find a *finite* number of receivers in this receiver class with the property that, regardless of the parameters of the channel in effect, at least one of these receivers provides a symbol error probability that is within a specified deviation from the *optimal* symbol error probability for the channel. These receivers are then used in parallel to perform a symbol-by-symbol demodulation of the received signal. The receiver output that gives the most reliable reproduction of the transmitted sequence is identified by means of a data verification mechanism. The resulting system is a universal receiver. Methods for data verification are developed in other papers. In this paper, we develop an algorithm for finding the required finite set of receivers. Typical issues, such as the tradeoff between the number of parallel receivers and the allowed deviation from optimality, are discussed.**

## I. Introduction

THIS paper is concerned with the design of a communication system that will operate well in the presence of nonselective Rician fading and additive white Gaussian noise (AWGN). For a Rician-fading channel, the received signal can have both a specular and a scatter (Rayleigh-faded) signal component. The strengths of these components, together with the noise power spectral density, determine the optimal, or minimum probability of error, receiver. Since these parameters are, in general, unknown or time-varying, it is not possible to design a single receiver that achieves optimal performance over the entire class of Rician fading channels. Our goal is to design a *universal receiver* that achieves nearly optimal performance over this class of channels (as in [2], the term *channel* refers to the cascade of the transmitter and the physical channel). The concept of a universal receiver for unknown or time-varying channels was introduced in a general setting in [2] and [4]. In this paper, we illustrate the detailed design of the demodulator portion of such a universal receiver for the specific example of nonselective Rician-fading channels.

Paper approved by J. Weitzen, the Editor for Fading/Equalization of the IEEE Communications Society. Manuscript received June 15, 1991; revised December 15, 1991. This work was supported by the Army Research Office under Contract DAAL-03-87-K-0097 and under Grant DAAL-03-91-G-0154.

U. Madhow is with the Coordinated Science Laboratory, University of Illinois, 1308 West Main Street, Urbana, IL 61801 USA.

M. B. Pursley is with Clemson University, 102 Riggs Hall, Clemson, SC 29634–0915 USA.

IEEE Log Number 9400571.

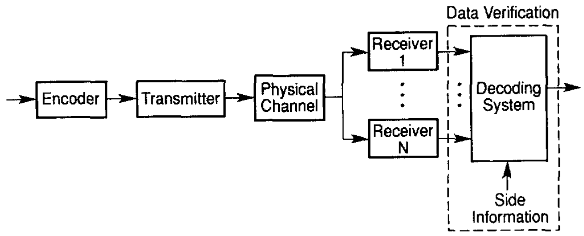


Fig. 1. A parallel implementation of a universal receiver.

Our implementation approach for a universal receiver is as follows. First find a finite set of receivers with the property that, for any channel in the class of interest, the performance of at least one of these receivers is within a specified deviation of the optimal performance for that channel. This part of the design is called the *detection aspect* of the universal receiver design. The receivers in this set are then employed in a parallel configuration to demodulate the received signal. The identity of the receiver that is best matched to the channel parameters is not known *a priori*, and some of the other receivers may perform very poorly for the channel in effect. The problem of choosing the output of the best receiver is termed the *data verification problem*, since it consists of determining which receiver is producing an output sequence which is the most reliable reproduction of the transmitted sequence. Data verification techniques that rely primarily on error-control coding are discussed in [2]. In a specific application, it may also be possible to develop additional side information to aid in data verification, and this is discussed for Rician channels in [3].

The block diagram of the universal receiver is shown in Fig. 1. For the purpose of this paper, we assume that the data verification problem can be solved. Our sole concern here is with the detection aspect of the universal receiver design.

The class of Rician-fading channels is modeled as a parametric channel class $C$ in which each channel is identified with an $m$-tuple that specifies completely the statistical characterization of the channel. If the values of the channel parameters are known, the optimal receiver can be determined. The optimal receiver is characterized by a single parameter, and the receiver class $R$ can be defined as the range of possible values of this parameter. This turns out to be an infinite set.

Suppose channel $c$ is in effect, and receiver $b$ is used; that is, $c$ is an $m$-tuple specifying the true values of the channel parameters for the channel in effect and $b$ is the value of the parameter for the receiver being used. Denote the error probability in such a situation by $P_E(c;\, b)$. For a given channel $c$, there is an optimal receiver in $R$. The corresponding minimum error probability is

$$P_E^*(c) = \min_{b \in R} P_E(c;\, b).$$

It is desired to attain a performance within a prescribed deviation of this optimal performance, and the allowable deviation from optimality is specified by a *degradation function* $h$ [2] as follows. For any channel $c$ in the channel class, the error probability attained by the universal receiver should not exceed $h[P_E^*(c)]$. The detection aspect of the design consists of finding a finite number of receivers, denoted by $b_1, b_2, \ldots, b_N$, with the property that at least one of the receivers performs within the specified degradation for any channel in the class of interest. Thus, we require that

$$\min_{1 \le i \le N} P_E(c;\, b_i) \le h[P_E^*(c)].$$

The goal is to choose the $b_i$ in a systematic manner, and general procedures are given in [2] for accomplishing this. These procedures require certain conditions on the channel class, receiver class, and the degradation function. In some applications, however, the conditions given in [2] do not hold, and specific features of the problem must be exploited in order to design a universal receiver. This is true for the Rician channel class considered in this paper. In fact, one of the points we wish to illustrate by means of this example is that the details of a given application must be considered in the design of a universal receiver; the formulation in [2] provides only general guidelines. Furthermore, the specific nature of the Rician example enables us to explore the tradeoff between the required deviation from optimality and the number of receivers required in the parallel configuration, and it permits consideration of the complexity of implementing a parallel configuration of receivers compared to a single receiver. Such issues cannot be explored in the general setting of [2].

Even though the performance gain achieved by the universal receiver may not be sufficient to justify is additional complexity for this specific example, the Rician-fading channel is a good choice for illustrative purposes because it provides a channel class for which the universal receiver design can be carried out primarily by analytical techniques. Our primary purpose here, and in [3], is to employ a concrete example to explore some of the issues involved in using the concepts of universality and parallelism in the design of communication systems.

The detailed system model is presented in Section II. The universal receiver design problem is formulated for Rician fading channels in Section III. In Section IV, an algorithm is given for the universal receiver design. Section V contains some numerical results, together with a discussion of implementation issues. Some concluding remarks are given in Section VI.

## II. System Model

The signal received over a Rician-fading channel may have one or both of the following signal components: a specular component and a scatter component. We assume the receiver can acquire the phase of the specular component, if it is present, but the phase of the scatter component may change from bit to bit. The phase of the scatter component is modeled as a random variable whose value is not known to the receiver. If there is only a scatter component, the receiver has no phase reference for the received signal, so that it is not possible to use antipodal signaling for channels on which Rayleigh fading of this kind may occur. Thus, it is assumed in this paper that orthogonal signaling (e.g., FSK or DPSK) is used, because such signals are appropriate for noncoherent demodulation. The signals used are narrowband and of duration $T$. The signal $s_i$, transmitted when the binary symbol $i$ $(i = 0,\, 1)$ is sent, is given by

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

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

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

$f_c$ is the carrier frequency, and $\phi$ 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

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

where $n(t)$ is white Gaussian noise with two-sided power spectral density $N_0/2$. The first term above is the specular component of the signal, and $\alpha$ is the channel gain for the specular component. The second term is the scatter component of the signal, and $S$ and $\Theta$ are random variables with joint density

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

Without loss of generality, the energy $E$ of the transmitted signal is normalized to 1, replacing $\alpha$ by $\alpha E^{1/2}$ and $\sigma^2$ by $\sigma^2 E$.

The receiver has perfect knowledge of the baseband waveforms $x_0(t)$ and $x_1(t)$. When there is a specular component, the receiver is assumed to be locked to its phase $\phi$. The receiver's decision is based on the statistics $L_{ci}$ and $L_{si}$ given by

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

and

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

for $i = 0$ and $i = 1$. The corresponding block diagram is given in Fig. 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 $N(m,\, \nu)$ a Gaussian distribution with mean $m$ and variance $\nu$. The

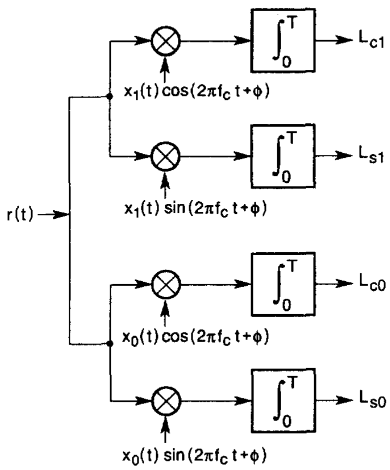


Fig. 2. Block diagram for generation of correlator outputs.

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

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

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

If a 0 is sent, interchange $L_{c1}$ and $L_{c0}$, and $L_{s1}$ and $L_{s0}$, in (3). The minimum probability of error decision rule can now be found. If there is a scatter component (i.e., $\beta > 0$), the decision rule is given by [6], [7]

$$L_{c1}^2 + L_{s1}^2 + 2b_{\text{opt}} L_{c1} \gtrless_{H_0}^{H_1} L_{c0}^2 + L_{s0}^2 + 2b_{\text{opt}} L_{c0}, \tag{4}$$

where

$$b_{\text{opt}} = \alpha/\beta. \tag{5}$$

Setting $b_{\text{opt}} = 0$ in (4) yields the *noncoherent receiver*. If there is no scatter component ($\beta = 0$), the minimum probability of error decision rule is given by

$$L_{c1} \gtrless_{H_0}^{H_1} L_{c0}. \tag{6}$$

The receiver corresponding to this decision rule is called the *coherent receiver*. Note that (6) can be obtained from (4) by dividing both sides by $b_{\text{opt}}$ and letting $b_{\text{opt}} \to \infty$.

It is useful, at this point, to note some key parameters of interest. The channel is completely specified by the parameters $\alpha$, $\beta$, and $N_0$, since the dependence on $E$ has already been removed by the normalization $E = 1$. Implementation of the optimal decision rule requires knowledge of the parameter $b_{\text{opt}}$, which, in turn, depends on the channel parameters. Another parameter that is often considered in the literature is $\gamma = \alpha/\sigma$, which is a measure of the relative strengths of the specular and scatter components of the signal. When there is no specular component, $\gamma = 0$, and when there is no scatter component, $\gamma = \infty$. Intermediate cases lead to positive real values of $\gamma$. Finally, we note that the overall signal-to-noise ratio (SNR) is given by $(\alpha^2 + 2\sigma^2)/N_0$.

It has been argued [6] that the noncoherent receiver works fairly well over the entire range of $\gamma$, so that it can be used as a *robust* receiver when the system parameters are either variable or unknown. However, our goal is to achieve nearly *optimal* performance over the range of $\gamma$. In particular, for large $\gamma$, there is a potential gain in performance that can be obtained by a universal receiver. Of course, realizing this potential performance gain requires a solution to the data verification problem, and this is given in [3]. In view of the added complexity, the resulting performance gain (see [3]) may not be enough to justify the use of a universal receiver for this specific channel class. However, as noted in the introduction, Rician fading channels are ideal for our purpose of illustrating the idea of a universal receiver using a practical, yet analytically tractable, channel model.

## III. Problem Formulation

In this section, we define the channel class to be considered in the universal receiver design, and we then characterize the class of available receivers. For the design, it is necessary to have an expression for the error probability as a function of the receiver and the channel in effect. After supplying this expression, the degradation function is specified. A mathematical formulation of the design problem concludes this section. An algorithm for the design is given in the next section.

The channel class $C$ is defined as

$$C = \{(\alpha,\, \beta,\, N_0) \mid 0 \le \alpha < \infty,\ 0 \le \beta < \infty,\ N_{0,\,\min} \le N_0 \le N_{0,\,\max}\}. \tag{7}$$

The only knowledge that the system designer has of the channel parameters is the range for noise power spectral density $N_0$. Note that the requirement in the general formulation [2] that $C$ be compact is not satisfied.

The receivers, or more specifically, decision rules, considered in the design are all of the form

$$L_{c1}^2 + L_{s1}^2 + 2bL_{c1} \gtrless_{H_0}^{H_1} L_{c0}^2 + L_{s0}^2 + 2bL_{c0}. \tag{8}$$

Each receiver is completely characterized by the parameter $b$, and the number $b$ will henceforth be identified with the corresponding receiver. The value $b = 0$ corresponds to the noncoherent receiver. The coherent receiver is identified with $b = \infty$. The receiver class $R$ is specified by the range of $b$, which is the entire nonnegative extended real line $[0, \infty]$. This is the set of minimum probability of error receivers for the channels in $C$.

Let $P_E(c;\, b)$ denote the probability of error that results when the receiver $b \in [0, \infty)$ is used for the channel $c = (\alpha,\, \beta,\, N_0)$. For $\beta > 0$, this error probability is shown in the Appendix to be

$$\begin{aligned} P_E(c;\, b) = {} & [(k_4^2 - 1)/\pi] \int_0^\infty \frac{1}{(z^2+1)(z^2+k_4^2)} \\ & \cdot \exp\left(-\frac{k_5(z^2+1+2\delta)}{(1+\delta)(z^2+k_4^2)}\right) \\ & \cdot \left[\cos\left(\frac{2\delta k_5 z}{(1+\delta)(z^2+k_4^2)}\right)\right. \\ & \left. - z\sin\left(\frac{2\delta k_5 z}{(1+\delta)(z^2+k_4^2)}\right)\right] dz \end{aligned} \tag{9}$$

where $k_4 = 1 + 2\beta^{-1}$, $k_5 = \alpha(\alpha + 2b)/N_0\beta$, and

$$\delta = (1 + \beta^{-1})\frac{\alpha(\alpha + 2b)}{(\alpha + b)^2 + b^2(\beta + 1)} - 1.$$

Turin [6] previously derived an expression for the optimal error probability $P_E^*(c)$ for a given set of channel parameters $c = (\alpha, \beta, N_0)$. This error probability is defined by

$$P_E^*(c) = \min_{b \in [0, \infty]} P_E(c; b) = P(c; b_{\text{opt}}(c))$$

where $b_{\text{opt}}(c) = \alpha/\beta$ is the optimal receiver for the channel $c$ (see (4)–(5)). When $\beta > 0$, $P_E^*(c)$ is obtained by setting $\delta = 0$ and $k_5 = k_5^* = (1 + 2\beta^{-1})\gamma^2/2$ in (9). The constant $k_4$ is independent of $b$, and depends only on $c$. While the optimal error probability can also be expressed in terms of known transcendental functions [6], the general expression in (9) is not amenable to such simplification and has to be evaluated by numerical integration. The exception is the special case of the noncoherent receiver ($b = 0$), for which the error probability is given by [6]

$$P_E(c; 0) = (\beta + 2)^{-1}\exp[-\alpha^2/N_0(\beta + 2)].$$

Unlike (9), this formula is valid for $\beta = 0$ also, and the result for $\beta = 0$ is

$$P_E(c; 0) = \tfrac{1}{2}\exp(-\alpha^2/2N_0),$$

the well-known expression for the error probability of the noncoherent receiver on an AWGN channel.

We will also need the expression for the error probability $P_E(c; \infty)$ when the coherent receiver, which uses the decision rule given in (6), is used for a given channel $c = (\alpha, \beta, N_0)$. It follows easily from (3) that

$$P_E(c; \infty) = 1 - \Phi\{\alpha/[N_0(\beta + 2)/2]^{1/2}\} \tag{10}$$

where $\Phi(\cdot)$ is the standard Gaussian distribution function. When $\beta = 0$, the coherent receiver is optimal, and $P_E^*(c)$ is obtained from the expression (10).

The performance measure considered here is the symbol error probability at the output of the demodulator. In practice, there is a range of interest for the symbol error probability $p$; namely, it is required to achieve $p \le p_{\max}$ and there is no interest in achieving symbol error probabilities lower than $p_{\min}$. The only constraint on $p_{\min}$ is $p_{\min} > 0$, but in practice typical values of $p_{\min}$ might be in the range $10^{-4}$ to $10^{-12}$, depending on the application. Channels in the given class $C$ for which the optimal error probability exceeds $p_{\max}$ are termed *bad* channels, and will not be considered in the design. Channels other than these are called *good* channels, and these are the channels of interest in the design.

With these considerations in mind, the universal receiver design problem can now be stated as follows: find a finite number of nonnegative extended real numbers, denoted by $b_1, \ldots, b_N$, so that, given any $c \in C$ for which $P_E^*(c) \le p_{\max}$,

$$\min_{1 \le i \le N} P_E(c; b_i) \le \max\{p_{\min}, h(P_E^*(c))\} \tag{11}$$

where $h$ is a degradation function, which for the present application is an increasing, continuous function that satisfies $h(x) > x$ for all $x \in (0, 1/2]$. Note that even though $h(0) = 0$ is allowed, the effective degradation function is $h^*(x) = \max\{p_{\min}, h(x)\}$, which satisfies $h^*(x) > x$ for all $x \in [0, 1/2]$. Thus, if we find receivers $b_1, \ldots, b_N$ satisfying (11), it is guaranteed that for any good channel, at least one of these receivers attains an error probability that either does not exceed $p_{\min}$ or stays within a specified degradation of the optimal error probability. An example of a degradation function is

$$h(x) = kx, \quad 0 \le x \le 1/2, \tag{12}$$

where $k > 1$ is a design parameter. This function allows a multiplicative degradation of the optimal error probability. However, as seen from (11), when the optimal error probability is less than $k^{-1}p_{\min}$, we are content with attaining an error probability of $p_{\min}$. Another possibility for a degradation function is

$$h(x) = \{\max[k_{\min}, -r\log_{10}(x)]\}x, \quad 0 \le x \le 1/2 \tag{13}$$

where $k_{\min} > 1$ and $r \ge 0$ are design constants. Here, too, a multiplicative degradation of the optimal error probability is allowed, but the degradation factor is permitted to depend on the optimal error probability. Specifically, the lower the optimal error probability, the larger the allowed multiplicative degradation. For instance, if $k_{\min} = 1.5$ and $r = 1$, then $h$ takes the value $1.5 \times 10^{-1}$ when the optimal error probability is $10^{-1}$, and the value $6 \times 10^{-6}$ when the optimal error probability is $10^{-6}$. Note that for $r = 0$, (13) reduces to (12).

The choice of the degradation function $h$ for a given application depends on how effectively the data verification problem can be solved. If the allowed deviation from optimality is too small, a large number of receivers may be required in the parallel configuration, and it may be difficult to identify reliably the best receiver. In such a situation, it may be better to allow a larger deviation from optimality and use fewer receivers in parallel. The choice of $h$ may also depend on whether feedback is available for adapting the rate of the error-control code to the symbol error probability at the outputs of the parallel receivers. If so, we would try to make the error probability as small as possible for any given channel, and may use a fixed multiplicative degradation as in (12). If, on the other hand, the code rate is fixed, a larger relative degradation may be permitted for better channels, and $h$ may be as in (13) with $r > 0$. Both choices of $h$ are considered in the numerical evaluation.

## IV. Design Algorithm

Although the conditions in the general formulation of [2] are not satisfied in the present application, it is useful to start with a result [2, Proposition 2] which can be stated roughly as follows. If the channel and receiver classes are compact, and the performance measure is continuous in the channel and receiver parameters, then the desired universal receiver design can be obtained by partitioning the receiver class. The resulting design procedure is given in [2], and will be referred to in this paper as the *general design procedure*.

In this section, we develop a *specific design algorithm* for Rician-fading channels. The need for a specific design algorithm is based on two considerations. First, as noted above, the sufficient conditions required to apply the general design procedure are not satisfied by the class of Rician-fading channels. Second, for Rician-fading channels, it is necessary to consider implementation details that cannot be considered in the general framework in which the general design procedure is stated. In what follows, we first deal with the difficulties that arise because the sufficient conditions of [2] are not satisfied. Next, we show that the implementation of the algorithm requires the solution of a series of subproblems, we describe and solve a typical subproblem, and we use the solution to obtain the specific design algorithm for Rician-fading channels.

For any receiver $b$ and any subset $S$ of the channel class $C$, it is said that $b$ is *nearly optimal* for $S$ if it performs within the specified degradation for all channels in $S$, that is,

$$P_E(c;\, b) \le \max\{p_{\min},\, h[P_E^*(c)]\} \text{ for all } c \in S. \tag{14}$$

Recall that a channel $c$ is a triple $(\alpha,\, \beta,\, N_0)$, and that a receiver $b$ corresponds to the ratio $\alpha/\beta$ for some such channel in $C$. Define, for any receiver $b^* \in R$, the set $C(b^*)$ of good channels in $C$ for which $b^*$ is optimal, that is,

$$C(b^*) = \{c \in C: b_{\text{opt}}(c) = b^*,\ P_E^*(c) \le p_{\max}\}. \tag{15}$$

Our approach is to find $b_1, \ldots, b_N$ so that, given any receiver $b^* \in R$, at least one of the $b_i$ is nearly optimal for all the channels in $C(b^*)$. Since any given channel in $C$ has an optimal receiver, this means that one of these $N$ receivers is nearly optimal for any good channel in $C$, as required by (11). Thus, $R$ is partitioned into intervals such that, for each interval in the partition, there is a $b_i$ that is nearly optimal for every good channel for which some receiver in the interval is optimal. This is the same idea as in the general design procedure, but for our present application, neither the channel class $C$ nor the receiver class $R$ are compact. The following proposition overcomes the difficulty with dealing with the receiver class if we insist that the coherent receiver is always one of the receivers in the parallel configuration. The proof of this proposition is given in the Appendix.

*Proposition 1:* There exists $b_0$ such that the coherent receiver is nearly optimal for all channels in $C(b)$, $b \in [b_0, \infty]$, that is, for any such channel $c$, $P(c, \infty) \le \max\{p_{\min},\, h[P_E^*(c)]\}$.

Denote by $U$ the smallest such $b_0$. For the purpose of developing the algorithm, any such $b_0$ can be chosen as $U$; choosing the smallest possible $U$ only enhances the efficiency of the algorithm. The fact that there is indeed a smallest such $b_0$ can be shown using the continuity of the bit error probability as a function of the channel and receiver parameters. The coherent receiver is nearly optimal for all good channels for which any receiver in $[U, \infty]$ is optimal. Thus, in choosing the other receivers, the receiver class can be restricted to the compact interval $[0, U]$, and the channel class can be considered to be the set of good channels for which any receiver in $[0, U]$ is optimal. This channel class is not compact, because of the range for $\alpha$ and $\beta$. However, as $\alpha$ and $\beta$ become large, the error probability goes to zero for any receiver in $[0, U]$. Since the error probability is not required to be less than $p_{\min}$, $\alpha$ and $\beta$ can be restricted to lie in a suitably chosen compact set. This enables us to show, using Proposition 2 of [2], that the receiver class $[0, U]$ can be partitioned, and this result is stated as Proposition 2 in the following.

Proposition 2 is a result concerning the function $n(\cdot)$ defined on $[0, \infty)$ by

$$\begin{aligned} n(b) = \sup\{r > 0\colon &\text{ For all } b' \in [b,\, b + r],\\ &P_E(c;\, b) \le \max[p_{\min},\, h(P_E^*(c))]\\ &\text{for all } c \in C(b')\}. \end{aligned}$$

Thus, if $b' < b + n(b)$, then the receiver $b$ is nearly optimal for all good channels for which $b'$ is optimal. Proposition 2, which is proved in the Appendix, states that $n(b)$ is bounded away from zero so that all intervals of the form $[b,\, b + n(b)]$ have length greater than some fixed positive number $\eta$.

*Proposition 2:* Let $\eta = \inf\{n(b)\colon\ 0 \le b \le U\}$, then $\eta > 0$.

Thus, we can start with $b_1 = 0$ and let $b_{i+1} = b_i + n(b_i)$ until the entire interval $[0, U]$ has been covered, and it will require at most $\lceil U/\eta \rceil$ receivers. These receivers, together with the coherent receiver, satisfy the design criterion (11). However, the numbers $U$ and $n(b_i)$ must be computed for use in the design. In order to do so, it is necessary to solve a series of subproblems. We first consider a typical subproblem.

Given any two receivers $b$ and $b^*$, it is required to check whether $b$ is nearly optimal for all channels in $C(b^*)$, the class of good channels for which $b^*$ is optimal. To this end, define

$$\Delta(b,\, b^*) = \min_{c \in C}[h(P_E^*(c)) - P_E(c;\, b)],$$

subject to

$$b_{\text{opt}}(c) = \alpha/\beta = b^*, \quad P_E^*(c) \le p_{\max} \tag{16}$$

and

$$P_E(c;\, b) \ge p_{\min} \tag{17}$$

where $c = (\alpha,\, \beta,\, N_0)$. It is easy to see from (14) and (15) that $b$ is nearly optimal for $C(b^*)$ if and only if $\Delta(b,\, b^*) \ge 0$. The set $C(b^*)$ is defined by the constraint (16), and the constraint (17) reflects the fact that it is not required to attain an error probability below $p_{\min}$. Denote by $C(b,\, b^*)$ the subset of $C(b^*)$ that satisfies the constraints (16)–(17). Then

$$\Delta(b,\, b^*) = \min_{c \in C(b,\, b^*)}[h(P_E^*(c)) - P_E(c;\, b)]. \tag{18}$$

This minimization problem is solved by giving a simple characterization of $C(b,\, b^*)$. For $c \in C(b,\, b^*)$, it follows from (16) that $\beta = \alpha/b^*$ where it is assumed that $b^* > 0$. The error probability when receiver $b$ is used for such channels is denoted by $P_b(\alpha,\, N_0,\, b^*)$, and is given by

$$P_b(\alpha,\, N_0,\, b^*) = P_E[(\alpha,\, \alpha/b^*,\, N_0);\, b]. \tag{19}$$

The corresponding optimal error probability is given by $P_{b^*}(\alpha,\, N_0,\, b^*)$. Obviously, $P_{b^*}(\alpha,\, N_0,\, b^*) \le P_b(\alpha,\, N_0,\, b^*)$. We now need the following proposition, the proof of which is given in the Appendix.

*Proposition 3:* Suppose that $b^* \in (0, \infty)$, and that either $b = \infty$ or $0 \le b \le b^*$. Then $P_b$ (and, as a special case, $P_{b^*}$) is decreasing in $\alpha$.

Fix $b$ and $b^*$ as in Proposition 3. We now exploit Proposition 3 to specify the range of $\alpha$ corresponding to the constraint set $C(b, b^*)$ for the optimization in (18). Fix a given value of $N_0 \in [N_{0,\min}, N_{0,\max}]$. Let $\alpha^*(N_0)$ be the value of $\alpha$ for which the error probability using receiver $b$ equals $p_{\min}$. Since we are not interested in attaining error probabilities smaller than $p_{\min}$, larger values of $\alpha$ need not be considered, by virtue of Proposition 3. Similarly, let $\alpha_*(N_0)$ be the value of $\alpha$ for which the error probability using the optimal receiver $b^*$ equals $p_{\max}$. Since we do not consider bad channels (channels for which the optimal error probability exceeds $p_{\max}$) in our design, it is not necessary to consider values of $\alpha$ lower than $\alpha_*(N_0)$, again invoking Proposition 3. Formally, define $\alpha^*(N_0)$ and $\alpha_*(N_0)$ by

$$P_b(\alpha^*, N_0, b^*) = p_{\min},$$

and

$$P_{b^*}(\alpha_*, N_0, b^*) = p_{\max}$$

where the dependence of $\alpha^*$ and $\alpha_*$ on $b$ and $b^*$ is suppressed. For the purpose of computing $\Delta(b, b^*)$, for each $N_0 \in [N_{0,\min}, N_{0,\max}]$, we can restrict attention to $\alpha \in [\alpha_*(N_0), \alpha^*(N_0)]$. Since $p_{\min} < p_{\max}$, it is easy to see that $\alpha_*(N_0) < \alpha^*(N_0)$, so that the interval of interest is nontrivial. Now, define (suppressing dependence on $b$ and $b^*$ as before)

$$J(N_0) = \min\{h[P_{b^*}(\alpha, N_0, b^*)] - P_b(\alpha, N_0, b^*) | \alpha_*(N_0) \le \alpha \le \alpha^*(N_0)\}.$$

Then

$$\Delta(b, b^*) = \min\{J(N_0) | N_{0,\min} \le N_0 \le N_{0,\max}\}.$$

Thus, $\Delta(b, b^*)$ can be computed by two successive unconstrained one-dimensional minimizations over compact intervals. First, we minimize over $\alpha \in [\alpha_*(N_0), \alpha^*(N_0)]$ for each $N_0 \in [N_{0,\min}, N_{0,\max}]$ to obtain $J(N_0)$. Then, $J(N_0)$ is minimized over $N_0 \in [N_{0,\min}, N_{0,\max}]$ to obtain $\Delta(b, b^*)$. The sign of $\Delta(b, b^*)$ determines if the receiver $b$ is nearly optimal for $C(b^*)$.

In addition to the preceding development, two conjectures are needed to prove that the algorithm always produces a set of receivers that satisfy (11).

*Conjecture C1:* Suppose $b_2^* < b_1^* < \infty$. If the coherent receiver ($b = \infty$) is nearly optimal for $C(b_2^*)$, then it is nearly optimal for $C(b_1^*)$.

*Conjecture C2:* Suppose $b \in [0, \infty)$ and $b < b_1^* < b_2^* < \infty$. If the receiver $b$ is nearly optimal for $C(b_2^*)$, then $b$ also is nearly optimal for $C(b_1^*)$.

Conjecture C1 is reasonable because the relative strength of the specular component is larger for channels in $C(b_1^*)$ than for channels in $C(b_2^*)$, so that the coherent receiver is better matched to channels in the former class. According to C1, if $\Delta(\infty, b_2^*) \ge 0$, it is not necessary to check the value of $\Delta(\infty, b_1^*)$ at any $b_1^* \in (b_2^*, \infty)$. Conjecture C2 is reasonable because the relative strength of the scatter component is larger for channels in $C(b_1^*)$ than for channels in $C(b_2^*)$, and it is larger for channels in $C(b)$ than for channels in $C(b_1^*)$. Thus, the receiver $b$ is better matched to channels in $C(b_1^*)$ than to channels in $C(b_2^*)$. According to C2, if $\Delta(b, b_2^*) \ge 0$, it is not necessary to check the value of $\Delta(b, b_1^*)$ at any $b_1^* \in (b, b_2^*)$.

We are now ready to give the specific design algorithm for finding a set of receivers $b_1, \ldots, b_N$ that satisfy (11).

*Step 1:* Set $V = \min\{b^* \ge 0\colon \Delta(\infty, b^*) \ge 0$ for all $b^* \ge b_0\}$. From C1 and the discussion immediately following Proposition 1, it follows that $V = U$, and that the coherent receiver is nearly optimal for $C(b^*)$ for all $b^* \in [V, \infty)$.

*Step 2:* Set $i = 1$ and $b_1 = 0$ to ensure that the noncoherent receiver is always included in the parallel configuration.

*Step 3:* If $\Delta(b_i, V) \ge 0$, go to Step 5. According to C2, this means that the receiver $b_i$ is nearly optimal for $C(b^*)$ for all $b^* \in [b_i, V]$.

*Step 4:* Set $b_{i+1} = \max\{b^* > b_i\colon \Delta(b_i, b^*) \ge 0\}$. According to C2, this is equivalent to setting $b_{i+1} = b_i + n(b_i)$. Replace $i$ by $(i+1)$ and go to Step 3.

*Step 5:* Set $N = i + 1$, $b_N = \infty$, and stop.

The algorithm gives an explicit statement of the design procedure developed earlier in the section. The two conjectures are needed for the correctness of the algorithm; its convergence is guaranteed by Propositions 1 and 2. To compute $V$ in Step 1, the value of $b^*$ is decreased until $\Delta(\infty, b^*) < 0$, after which $b^*$ is increased until the first time $\Delta(\infty, b^*) \ge 0$. Since the changes in the value of $b^*$ are discrete, it is necessary that C1 be true to obtain a correct implementation of Step 1. Similarly, given $b_i$, to compute $b_{i+1}$, we find $b^*$ such that $\Delta(b_i, b^*) < 0$, and then decrease the value of $b^*$ until $\Delta(b_i, b^*) \ge 0$. Again, since the changes in $b^*$ are discrete, C2 is needed to ensure that Step 4 is implemented correctly. Both our intuition and numerical results indicate that these conjectures are true, but analytical proofs have not been found. In any case, the algorithm can be used and the final result can be verified independently.

## V. Numerical Results and Implementation Issues

For the numerical results the design algorithm is carried out for several choices of the degradation function. The channel class that we consider is of the form (7), and is given by

$$C = \{(\alpha, \beta, N_0) \mid 0 \le \alpha < \infty,\ 0 \le \beta < \infty,\ 100 \le N_0 \le 1000\}.$$

Take $p_{\max} = 10^{-1}$ and $p_{\min} = 10^{-8}$. The numerical results obtained for various degradation functions of the form in (12) and (13) are presented in Table I.

TABLE I
Some Typical Designs

| Degradation function, $h(x)$ | $V$ | Placement of receivers, $b_i$ |
|---|---|---|
| $2x$ | 41 | 0, 11, $\infty$ |
| $3x$ | 26 | 0, 20, $\infty$ |
| $4x$ | 21 | 0, $\infty$ |
| $\{\max[1.2, -\log_{10}(x)]\}x$ | 34 | 0, 5, 21, $\infty$ |
| $\{\max[2, -\log_{10}(x)]\}x$ | 19 | 0, 17, $\infty$ |
| $\{\max[3, -\log_{10}(x)]\}x$ | 16 | 0, $\infty$ |

It should be noted that, if the other design parameters are fixed, the design obtained by the algorithm can be scaled to reflect scaling in the values of $N_{0,\min}$ and $N_{0,\max}$. The key property is

$$P_E[(r\alpha,\, r^2\beta,\, r^2N_0);\, rb] = P_E[(\alpha,\, \beta,\, N_0);\, b], \quad r > 0.$$

Thus, if $N_{0,\min}$ and $N_{0,\max}$ are scaled by a factor $r^2$, then the receivers obtained originally must be scaled by a factor $r$ for the same degradation function, provided all the other design parameters are fixed. This is especially relevant if there are changes in the gain in the receiver front end. As long as the same gain can be applied to the receiver parameters $b_1, \ldots, b_N$, the universal receiver will operate as designed. For instance, suppose that the noise spectral densities are scaled down by a factor of 100, so that

$$C = \{(\alpha,\, \beta,\, N_0) \mid 0 \le \alpha < \infty,\; 0 \le \beta < \infty,\; 1 \le N_0 \le 10\}.$$

If, in addition, $h(x) = 2x$, then we can deduce from Table I (scaling the receiver parameters down by a factor of 10) that $V = 4.1$, and that the receivers are placed at 0, 1.1 and $\infty$. Scaling is important if the values of $N_0$ are so small that round-off errors become significant in the implementation of the algorithm. The spectral densities can be scaled up for running the algorithm, and the result can subsequently be scaled back down to obtain the required design.

It is worth noting that the receivers that are placed in parallel have many components in common. Specifically, all the receivers require generation of the correlator outputs $L_{c1}$, $L_{s1}$, $L_{c0}$, and $L_{s0}$ [see Fig. 2 and eqs. (1)–(2)]. The difference between different receivers is in how these correlator outputs are processed [see (8)]. As a result, one need actually implement only the four correlators of Fig. 2 regardless of the number of parallel receivers. The resulting universal receiver is illustrated in Fig. 3; which shows the block diagram for a parallel configuration of $N$ receivers with $b_1 = 0$ (noncoherent receiver) and $b_N = \infty$ (coherent receiver). More generally, we expect that in many applications, receivers that are optimal for different channel conditions differ only in the way they process certain basic statistics, and these statistics can be generated by a single subsystem. Thus, by identifying the common operations for the different receivers, we can avoid duplicating them in the hardware implementation, so that the cost of building a parallel configuration of $N$ receivers may be only marginally more than that of building a single receiver.

## VI. Conclusion

We have illustrated the detection aspect of a universal receiver design for the class of Rician-fading channels. This class was selected for illustrative purposes only; as is seen in [3], the gain in performance over the noncoherent receiver is only of the order of 1 dB. In addition to providing a concrete example that supplements the theoretical development in [2], the universal receiver design also illustrates tradeoffs and implementation issues that are expected to be important in other applications as well. However, it is expected that the performance gains achievable with a universal receiver are greater for more complex channel classes.

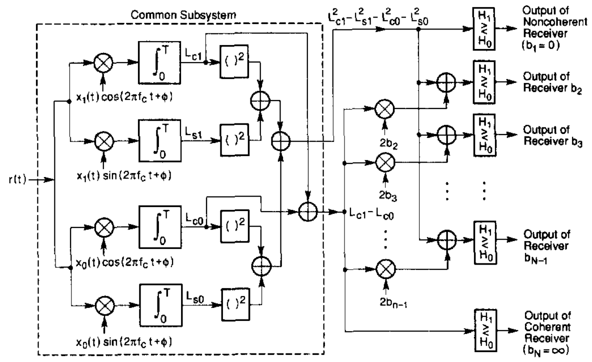

Fig. 3. Implementation of universal receiver.

The results of this paper are encouraging in that the universal approach is shown to apply even when the sufficient conditions of [2] do not hold. However, even after expending a significant amount of effort in developing the algorithm for the universal receiver design, it is necessary to rely on heuristics, in the form of conjectures C1 and C2. In fact, simply using an intuitive argument without the benefit of any design algorithm leads us to a good design for Rician fading channels: since the noncoherent receiver gives good performance over the entire range of fading characteristics, and the coherent receiver is nearly optimal when the scatter component is small, placing them in parallel is a good idea. The fact that this is a valid universal receiver design for a degradation function $h(s) = 4s$ (see Table I) simply adds a formal justification for our choice. We believe that in many situations, even when it is not possible to apply analytical methods such as those in [2] or those in this paper, the basic philosophy of employing receivers in parallel is a good approach for channels with unknown or time-varying characteristics. In fact, the channel classes for which we can gain the most from using parallelism are probably those that are less well-defined or more complicated than for the example considered here. In such situations, we expect that heuristic rather than analytical methods will be used to arrive at good choices for the receivers in the parallel configuration.

## Appendix

We first derive the expression (9) for the error probability when a given receiver $b$ is used for a given channel $c = (\alpha,\, \beta,\, N_0)$, with $\beta > 0$. This is followed by the proof of Proposition 3. Following that, two lemmas are proved; one concerns the continuity of the error probability as a function of the channel parameters, and the other is a simple result in real analysis. Finally, the proofs of Propositions 1 and 2, in which these two lemmas are used, are given. Due to the symmetry of the system, the error probabilities of interest equal the conditional error probabilities given that a 1 is sent. Throughout this appendix, therefore, all the relevant distributions and probabilities will be subjected to this conditioning, but for convenience, this is not reflected in the notation.

### *A.1 Derivation of the Error Probability Expression*

As explained earlier, we take $E = 1$. The analysis closely follows that in [6], the difference being that here the receiver $b$ need not be the optimal receiver for the channel $c$.

The decision rule (8), corresponding to receiver $b$, can be written as

$$(L_{c1} + b)^2 + L_{s1}^2 \gtrless_{H_0}^{H_1} (L_{c0} + b)^2 + L_{s0}^2.$$

Define, as in [6],

$$W_1 = L_{c1}+b,\ W_2 = L_{s1},\ W_3 = L_{c0}+b,\ W_4 = L_{s0},\quad \text{(A.1)}$$

and

$$D = W_1^2 + W_2^2 - W_3^2 - W_4^2. \quad \text{(A.2)}$$

Since we are implicitly conditioning on a 1 being sent, the error probability is the probability that $D < 0$. It follows from (3) that the $W_i$ are conditionally independent Gaussian random variables, with

$$W_1 \sim N(\alpha + b,\ N_0(\beta + 1)/2),$$

$$W_2 \sim N(0,\ N_0(\beta + 1)/2), \quad \text{(A.3)}$$

$$W_3 \sim N(b,\ N_0/2), \qquad W_4 \sim N(0,\ N_0/2). \quad \text{(A.4)}$$

The random variable $D$ is a quadratic form of the Gaussian random vector $\boldsymbol{W} = [W_1,\ W_2,\ W_3,\ W_4]^T$, so that its moment generating function is given by [6]

$$F_D(s) = E[e^{sD}] = |I - 2sMQ|^{-1/2} \cdot \exp\{-\tfrac{1}{2}\langle \boldsymbol{W}\rangle^T M^{-1}[I - (I - 2sMQ)^{-1}]\langle \boldsymbol{W}\rangle\} \quad \text{(A.5)}$$

where $I$ is the $4 \times 4$ identity matrix, $Q = \mathrm{diag}(1,\ 1,\ -1,\ -1)$ is the matrix of the quadratic form,

$$M = (E\{W_i W_j\}) = (N_0/2)\mathrm{diag}(\beta + 1,\ \beta + 1,\ 1,\ 1)$$

is the moment matrix of $\boldsymbol{W}$, $\langle \boldsymbol{W}\rangle = E\{\boldsymbol{W}\} = [\alpha + b,\ 0,\ b,\ 0]^T$, and $|\cdot|$ is the determinant function. Substituting in (A.5) yields, after some algebra, that

$$F_D(s) = [1 - k_3 s(1 + k_2^* s)]^{-1} \cdot \exp\{k_1 s(1 + k_2 s)/[1 - k_3 s(1 + k_2^* s)]\} \quad \text{(A.6)}$$

where $k_1 = (\alpha + 2b)\alpha$, $k_2^* = N_0(\beta + 1)/\beta$, $k_3 = N_0\beta$, and

$$k_2 = \frac{N_0[(\alpha + b)^2 + b^2(\beta + 1)]}{(\alpha + 2b)\alpha}.$$

The error probability can be written as

$$P(D < 0) = -(2\pi j)^{-1} \int_{-j\infty}^{j\infty} s^{-1} F_D(s)\, ds \quad \text{(A.7)}$$

where the path of integration is indented to the left at the $j$-axis singularities. Let $\delta_2 = (2k_2^*)^{-1}$ and

$$k_4 = (1 + 4k_2^*/k_3)^{1/2} = 1 + 2\beta^{-1}.$$

It is seen from (A.6) that $s^{-1}F_D(s)$ has singularities at the origin and at $\delta_2(-1 \pm k_4)$. Moving the integration path to the left by $\delta_2$ does not, therefore, change the value of the integral in (A.7), so that

$$P(D < 0) = -(2\pi j)^{-1} \int_{-j\infty-\delta_2}^{+j\infty-\delta_2} s^{-1} F_D(s)\, ds.$$

Making the substitution $s = \delta_2(jz - 1)$, it is obtained, after substantial simplification, that

$$P(D < 0) = [(k_4^2 - 1)/2\pi] \int_{-\infty}^{+\infty} \frac{1 + jz}{(z^2 + 1)(z^2 + k_4^2)} \cdot \exp\left\{\frac{-k_5(z^2 + 1 + 2\delta - 2jz\delta)}{(1 + \delta)(z^2 + k_4^2)}\right\} dz$$

where $k_5 = k_1/k_3$ and $\delta = (k_2^*/k_2) - 1$. The imaginary part of the integrand is odd and therefore integrates to zero. Using the fact that the real part of the integrand is even, the error probability is easily seen to reduce to the required formula (9). Note that this expression reduces to the formula (93) obtained in [6] for the optimal error probability on substituting $b = b_{\mathrm{opt}} = \alpha/\beta$. In this situation, $\delta = 0$ (since $k_2 = k_2^*$) and $k_5 = k_5^* = (\alpha + 2b_{\mathrm{opt}})b_{\mathrm{opt}}/N_0$. The value of $k_4$ is independent of $b$ and hence remains unchanged. Substituting, the following expression is obtained for the optimal error probability:

$$[(k_4^2 - 1)/\pi] \int_0^{\infty} \frac{1}{(z^2 + 1)(z^2 + k_4^2)} \exp\left(-\frac{k_5^*(z^2 + 1)}{(z^2 + k_4^2)}\right) dz.$$

□

### *A.2 Proof of Proposition 3*

Consider the error probability $P_b(\alpha,\ N_0,\ b^*)$ (defined in (19)) when receiver $b$ is used for a channel $c = (\alpha,\ \beta,\ N_0)$ for which the optimal receiver is $b^* > 0$. By (5), $\beta = \alpha/b^*$, and substituting in the error probability expression (10) for the coherent receiver, it is obtained that

$$P_\infty(\alpha,\ N_0,\ b^*) = 1 - \Phi\{\alpha/[N_0(\alpha/b^* + 2)/2]^{1/2}\}.$$

Taking partial derivatives with respect to $\alpha$ and simplifying,

$$\frac{\partial}{\partial\alpha} P_\infty(\alpha,\ N_0,\ b^*) = -\tfrac{1}{2}(b^*/\pi N_0)^{1/2} \tfrac{\alpha + 4b^*}{(\alpha + 2b^*)^{3/2}} \cdot \exp\left(-\frac{\alpha^2 b^*}{N_0(\alpha + 2b^*)}\right) < 0.$$

This proves the proposition for $b = \infty$. For $b < \infty$, the expression (9) for the error probability is in the form of an integral, and it is not clear upon differentiating with respect to $\alpha$ that the derivative is negative. It is convenient, therefore, to adopt an indirect approach, and consider the expression for the error probability in terms of the random variables $W_i$ and $D$ defined by (A.1)–(A.2). We have

$$P_b(\alpha,\ N_0,\ b^*) = P[D < 0] = P[W_1^2 + W_2^2 - W_3^2 - W_4^2 < 0].$$

Substituting $\beta = \alpha/b^*$, it is evident from (A.3)–(A.4) that only the distributions of $W_1$ and $W_2$ depend on $\alpha$. It is shown that $|W_1|$ and $|W_2|$ are stochastically increasing in $\alpha$. This implies that $W_1^2$ and $W_2^2$ are stochastically increasing in $\alpha$. Since the $W_i$ are independent, this means that $D$ is stochastically increasing in $\alpha$, from which it follows that $P_b(\alpha, N_0, b^*)$ is decreasing in $\alpha$.

From (A.3), it is simple to compute, for $t \geq 0$,

$$\begin{aligned} P[|W_1| \leq t] &= P[-t \leq W_1 \leq t] \\ &= \Phi\left(\frac{t-\alpha-b}{(N_0(\alpha/b^*+2)/2)^{1/2}}\right) \\ &\quad - \Phi\left(\frac{-t-\alpha-b}{(N_0(\alpha/b^*+2)/2)^{1/2}}\right). \end{aligned}$$

Taking partial derivatives with respect to $\alpha$ yields on simplification

$$\begin{aligned} &\frac{\partial}{\partial\alpha} P[|W_1| \leq t] \\ &\quad = -1/2(b^*/\pi N_0)^{1/2}(\alpha+b^*)^{-3/2} \\ &\quad \cdot \exp\left(-\frac{(t+\alpha-b)^2 b^*}{N_0(\alpha+b^*)}\right) \\ &\quad \cdot \left\{(\alpha+2b^*-b)\left[1-\exp\left(-\frac{4t(\alpha+b)b^*}{N_0(\alpha+b^*)}\right)\right]\right. \\ &\quad \left. + t\left[1+\exp\left(-\frac{4t(\alpha+b)b^*}{N_0(\alpha+b^*)}\right)\right]\right\} \leq 0, \end{aligned}$$

for $b^* \geq b$. This shows that $|W_1|$ is stochastically increasing in $\alpha$. Similarly, $|W_2|$ is seen to be stochastically increasing in $\alpha$ by noting that

$$\begin{aligned} P[|W_2| \leq t] &= P[-t \leq W_2 \leq t] \\ &= \Phi[t/(N_0(\alpha/b^*+2)/2)^{1/2}] \\ &\quad - \Phi[-t(N_0(\alpha/b^*+2)/2)^{1/2}]. \end{aligned}$$

which yields on differentiating with respect to $\alpha$

$$\begin{aligned} \frac{\partial}{\partial\alpha} P[|W_2| \leq t] &= t(b^*/\pi N_0)^{1/2}(\alpha+b^*)^{-3/2} \\ &\quad \cdot \exp\left(-\frac{t^2 b^*}{N_0(\alpha+b^*)}\right) \leq 0. \end{aligned}$$

Both the above inequalities are strict for almost all $t$, which can be used to show that $P_b(\alpha, N_0, b^*)$ is strictly decreasing in $\alpha$. This completes the proof of Proposition 3. □

### *A.3 Proofs of Propositions 1 and 2*

We first state and prove two lemmas that are used in the proofs of Propositions 1 and 2.

*Lemma A.1:* For $b^* > 0$, $P_b(\alpha, N_0, b^*)$ (and, as a special case, $P_{b^*}(\alpha, N_0, b^*)$) is a continuous function of $\alpha$ and $N_0$.

*Proof:* For $b = \infty$, the lemma follows directly on examining the expression (10) for the error probability. For $b < \infty$, the probability of error equals $P[D < 0]$ where $D$ is a continuous function of the $W_i$ (see (A.1)–(A.2)). Substituting $\beta = \alpha/b^*$, it is seen that (A.3)–(A.4) that the densities of the $W_i$ are pointwise continuous in $\alpha \geq 0$ and $N_0 \in [N_{0,\min}, N_{0,\max}]$ (assuming $N_{0,\min} > 0$). By Scheffe's theorem [1, p. 218], this implies $L_1$-continuity of the densities in these parameters, which in turn implies that the $W_i$ are continuous in distribution with respect to variations in $\alpha$ and $N_0$. Since $D$ is a continuous function of the $W_i$, it follows that $D$ is continuous in distribution as $\alpha$ and $N_0$ vary, which proves the required continuity of the error probability. □

The next lemma is a simple result in real analysis that will be used in the proof of Proposition 2. The proof is very simple, and is included only for lack of a convenient reference.

*Lemma A.2:* Suppose $r: X \times Y \to (-\infty, \infty)$ is a continuous function where $X$ and $Y$ are metric spaces, and let $t$ be a real number. Suppose that for each $x \in X$, there is a unique $y \in Y$ such that $r(x, y) = t$, and define the function $s: X \to Y$ by $r(x, s(x)) = t$. Then $s$ is continuous on $X$.

*Proof:* Suppose $x_n \to x$. By hypothesis, $y_n = s(x_n)$ is defined uniquely by $r(x_n, y_n) = t$, and $y = s(x)$ is defined uniquely by $r(x, y) = t$. Let $y'$ be any limit point of $(y_n)$. There is a subsequence $y_{n_j} \to y'$, so that, by the continuity of $r$, $t = \lim_j r(x_{n_j}, y_{n_j}) = r(x, y')$. Since $r(x, y) = t$, this implies that $y' = y$. Thus, the only limit point of $(y_n)$, is $y$, so that $s(x_n) = y_n \to y = s(x)$, which proves the continuity of $s$. □

It remains to prove Propositions 1 and 2.

*A.3.1) Proof of Proposition 1:* For a channel $c$ with optimal receiver $b^*$, the error probability when the coherent receiver ($b = \infty$) is used is given by

$$\begin{aligned} P_\infty(\alpha, N_0, b^*) &= 1 - \Phi\{\alpha/[N_0(\alpha/b^*+2)/2]^{1/2}\} \\ &\to 1 - \Phi(\alpha/N_0^{1/2}) \text{ as } b^* \to \infty. \end{aligned} \quad \text{(A.8)}$$

The optimal error probability can be written as

$$P_{b^*}(\alpha, N_0, b^*) = P[U_{b^*} + V_{b^*} < 0]$$

where $U_{b^*} = (L_{c1}^2 + L_{s1}^2 - L_{c0}^2 - L_{s0}^2)/2b^*$, and $V_{b^*} = L_{c1} - L_{c0}$. Using the distributions of $L_{ci}, L_{si}$ $(i = 0, 1)$ given in (3), and substituting $\beta = \alpha/b^*$, it is easy to show that, as $b^* \to \infty$, $E\{U_{b^*}^2\} \to 0$, and $V_{b^*} \to N(\alpha, N_0)$ in distribution. By Slutsky's theorem [5, p. 19], $(U_{b^*} + V_{b^*})$ is seen to converge in distribution to an $N(\alpha, N_0)$ random variable as $b^* \to \infty$. This implies that

$$P_{b^*}(\alpha, N_0, b^*) \to 1 - \Phi(\alpha/N_0^{1/2}) \text{ as } b^* \to \infty. \quad \text{(A.9)}$$

Given a number $b_1^* > 0$, define $\alpha_{\max}$ by $P_\infty(\alpha_{\max}, N_{0,\max}, b_1^*) = p_{\min}$. Since $\Phi$ is increasing, it is easy to see from (A.8) that $P_\infty(\alpha, N_0, b^*)$ is decreasing in $b^*$ and increasing in $N_0$. Further, it is known from Proposition 3 that $P_\infty(\alpha, N_0, b^*)$ is decreasing in $\alpha$. Hence, for $b^* > b_1^*$, if $\alpha > \alpha_{\max}$ and $N_0 \in [N_{0,\min}, N_{0,\max}]$, then $P_\infty(\alpha, N_0, b^*) < p_{\min}$. Hence, given $b_1^*$, in order to determine whether the coherent receiver is nearly optimal for all channels in $C(b^*)$, $b^* \geq b_1^*$, it suffices to restrict attention to $\alpha \in [0, \alpha_{\max}]$, which implies that $\beta \in [0, \alpha_{\max}/b_1^*]$. Therefore, to prove Proposition 1, it suffices to show that there is a $b_0^* \geq b_1^*$ such that, for $b^* \geq b_0^*$, for all $\alpha \in [0, \alpha_{\max}]$, and for all $N_0 \in [N_{0,\min}, N_{0,\max}]$,

$$h[P_{b^*}(\alpha, N_0, b^*)] - P_\infty(\alpha, N_0, b^*) \geq 0. \quad \text{(A.10)}$$

This is shown by contradiction. If there is no such $b_0^*$, there are sequences $b_n^* \to \infty$ $(b_n^* \geq b_1^*)$, $\alpha_n \in [0, \alpha_{\max}]$, and $N_0^{(n)} \in [N_{0,\min}, N_{0,\max}]$ which satisfy

$$h[P_{b^*}(\alpha_n, N_0^{(n)}, b_n^*)] - P_\infty(\alpha_n, N_0^{(n)}, b_n^*) < 0. \quad \text{(A.11)}$$

By the compactness of $[0, \alpha_{\max}]$ and $[N_{0,\min}, N_{0,\max}]$, it is assumed, without loss of generality, that $\alpha_n \to \alpha \in [0, \alpha_{\max}]$

and $N_0^{(n)} \to N_0 \in [N_{0,\min}, N_{0,\max}]$ by passing to convergent subsequences. Since $b_n^* \to \infty$, we have, letting $n \to \infty$ in (A.11), and using (A.8), (A.9) and Lemma A.1, that $h(x_0) - x_0 \leq 0$ where $x_0 = 1 - \Phi(\alpha/N_0^{1/2}) \in (0, 1/2]$. This gives the required contradiction and proves (A.10), since $h(x) > x$ for all $x \in (0, 1/2]$. □

*A.3.2) Proof of Proposition 2:* As before, consider a channel $c = (\alpha, \beta, N_0)$ with optimal receiver $b^*$, but now $b^* \in [0, U]$. Since $b^* = \alpha/\beta$, it is convenient to express $\alpha$ in terms of $\beta$ rather than the other way around in order to enable us to consider $b^* = 0$. Thus, we substitute $\alpha = \beta b^*$ and denote the error probability when receiver $b < \infty$ is used for such a channel by $P_b(\beta, N_0, b^*)$. Following the proof of Proposition 3, it is easy to show that this is a strictly decreasing function of $\beta$ for $b^* \geq b$. Also, it can be shown, arguing as in the proof of Lemma A.1, that $P_b(\beta, N_0, b^*)$ is a continuous function of $b$, $b^*$, $\beta$, and $N_0$, by showing the continuity in distribution of the random variable $D$ with respect to variations in these parameters.

Given $N_0$, $b$, and $b^*$, define $\beta^* = \beta^*(N_0, b, b^*)$ by $P_b(\beta^*, N_0, b^*) = p_{\min}$. Now, $P_b(\beta, N_0, b^*)$ is strictly decreasing in $\beta$, $P_b(0, N_0, b^*) = 0.5 > p_{\min}$, and $\lim_{\beta \to \infty} P_b(\beta, N_0, b^*) = 0 < p_{\min}$. Thus, the function $\beta^*$ is well-defined. Since $P_b(\beta, N_0, b^*)$ is continuous in $b$, $b^*$, $\beta$, and $N_0$. Lemma A.2 can be applied to see that $\beta^*$ is a continuous function of $N_0$, $b$, and $b^*$.

Define

$$\beta_{\max} = \max\{\beta^*(N_0, b, b^*) \mid N_{0,\min} \leq N_0 \leq N_{0,\max}, \\ 0 \leq b, \ b^* \leq U\}.$$

Since the error probability is decreasing in $\beta$, $P_b(\beta, N_0, b^*) < p_{\min}$ for $\beta > \beta_{\max}$. Thus, for partitioning the compact receiver space $R' = [0, U]$, attention is restricted to $\beta \in [0, \beta_{\max}]$. Since $\alpha = \beta b^*$, the channel parameters can be specified by $\beta$, $b^*$, and $N_0$, and the following channel class needs to be considered:

$$C' = \{(\beta, b^*, N_0) \mid 0 \leq \beta \leq \beta_{\max}, \ 0 \leq b^* \leq U, \\ N_{0,\min} \leq N_0 \leq N_{0,\max}\}.$$

This channel class is compact, and so is the associated receiver class $[0, U]$. Thus, since the error probability $P_b(\beta, N_0, b^*)$ is continuous in its parameters, the conditions for the general design procedure in [2] hold, and Proposition 2 in [2] implies that $\inf_{b \in [0, U]} n(b) > 0$. This completes the proof of Proposition 2 and concludes the Appendix. □

## REFERENCES


[1] P. Billingsley, *Probability and Measure.* New York: Wiley, 1986.
[2] U. Madhow and M. B. Pursley, "A parallel systems approach to universal receivers," *IEEE Trans. Inform. Theory*, vol. 37, pp. 291–306, Mar. 1991.
[3] ——, "Universal receivers with side information from the demodulators: an example for nonselective Rician fading channels," *IEEE Trans. Commun.*, vol. 42, pp. 2395–2405, July 1994.
[4] M. B. Pursley and U. Madhow, "Parallel structures for universal communication receivers," in *Proc. 1990 Bilkent Int. Conf. New Trends in Commun., Contr. Signal Processing*, vol. I, July 2–5, 1990, Bilkent Univ., Ankara, Turkey, pp. 215–227.
[5] R. J. Serfling, *Approximation Theorems of Mathematical Statistics.* New York: Wiley, 1980.
[6] 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.
[7] H. L. Van Trees, *Detection, Estimation and Modulation Theory.* New York: Wiley, 1968, vol. 1.



**Upamanyu Madhow** for a photograph and biography, please see the July 1994 issue of this TRANSACTIONS, p. 2404.

**Michael B. Pursley** (S'68–M'68–SM'77–F'82) for a photograph and biography, please see the July 1994 issue of this TRANSACTIONS, p. 2405.