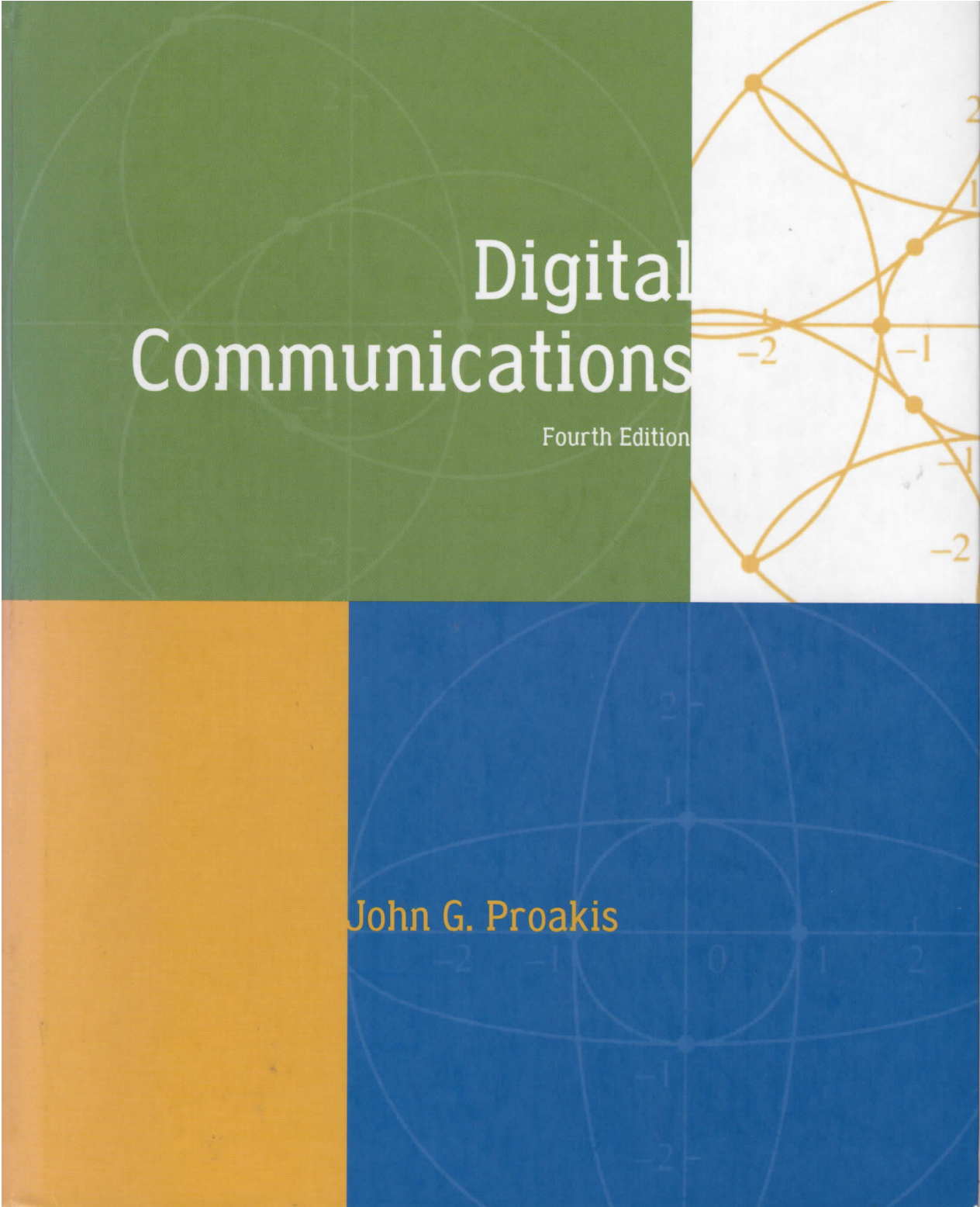




The book cover is divided into four quadrants. The top-left quadrant is green with faint white circles and a coordinate grid. The top-right quadrant is white with orange geometric lines and dots, and a coordinate grid with labels -2, -1, 1, 2. The bottom-left quadrant is solid orange. The bottom-right quadrant is blue with faint white circles and a coordinate grid with labels -2, -1, 0, 1, 2.

Digital Communications

Fourth Edition

John G. Proakis

Digital Communications

Fourth Edition

JOHN G. PROAKIS

*Department of Electrical and Computer Engineering
Northeastern University*



Boston Burr Ridge, IL Dubuque, IA Madison, WI New York San Francisco St. Louis
Bangkok Bogotá Caracas Lisbon London Madrid Mexico City Milan
New Delhi Seoul Singapore
Sydney Taipei Toronto

McGraw-Hill Higher Education 

A Division of The McGraw-Hill Companies

DIGITAL COMMUNICATIONS

Published by McGraw-Hill, an imprint of The McGraw-Hill Companies, Inc., 1221 Avenue of the Americas, New York, NY, 10020. Copyright © 2001, 1995, 1989, 1983, by The McGraw-Hill Companies, Inc. All rights reserved. No part of this publication may be reproduced or distributed in any form or by any means, or stored in a database or retrieval system, without the prior written consent of The McGraw-Hill Companies, Inc., including, but not limited to, in any network or other electronic storage or transmission, or broadcast for distance learning. Some ancillaries, including electronic and print components, may not be available to customers outside the United States.

This book is printed on acid-free paper.

2 3 4 5 6 7 8 9 0 DOC/DOC 0 9 8 7 6 5 4 3 2 1

ISBN 0-07-232111-3

Publisher: *Thomas Casson*

Sponsoring editor: *Catherine Fields Shultz*

Developmental editor: *Emily J. Gray*

Marketing manager: *John Wannemacher*

Project manager: *Craig S. Leonard*

Production supervisor: *Rose Hepburn*

Senior designer: *Kiera Cunningham*

New media: *Christopher Styles*

Compositor: *Interactive Composition Corporation*

Typeface: *10.5/12 Times Roman*

Printer: *Quebecor World Fairfield, Inc.*

Library of Congress Cataloging-in-Publication Data

Proakis, John G.

Digital communications / John G. Proakis.—4th ed.

p. cm.

ISBN 0-07-232111-3

1. Digital communications. I. Title.

TK5103.7.P76 2000

621.382-dc21

00-025305

www.mhhe.com

$$\hat{I}(t) = \sum_{j=-K}^K c_j(t-1)v_{t-j}$$

By changing the index j on $c_j(t-1)$ to run from $j = 0$ to $j = N-1$ and simultaneously defining

$$y(t) = v_{t+K}$$

the estimate $\hat{I}(t)$ becomes

$$\begin{aligned}\hat{I}(t) &= \sum_{j=0}^{N-1} c_j(t-1)y(t-j) \\ &= \mathbf{C}'_N(t-1)\mathbf{Y}_N(t)\end{aligned}\quad (11.4-1)$$

where $\mathbf{C}_N(t-1)$ and $\mathbf{Y}_N(t)$ are, respectively, the column vectors of the equalizer coefficients $c_j(t-1)$, $j = 0, 1, \dots, N-1$, and the input signals $y(t-j)$, $j = 0, 1, 2, \dots, N-1$.

Similarly, in the decision-feedback equalizer, we have tap coefficients $c_j(t)$, $j = 0, 1, \dots, N-1$, where the first $K_1 + 1$ are the coefficients of the feedforward filter and the remaining $K_2 = N - K_1 - 1$ are the coefficients of the feedback filter. The data in the estimate $\hat{I}(t)$ is $v_{t+K_1}, \dots, v_{t+1}, \tilde{I}_{t-1}, \dots, \tilde{I}_{t-K_2}$, where $\tilde{I}_{t-j}$, $1 \leq j \leq K_2$, denote the decisions on previously detected symbols. In this development, we neglect the effect of decision errors in the algorithms. Hence, we assume that $\tilde{I}_{t-j} = I_{t-j}$, $1 \leq j \leq K_2$. For notational convenience, we also define

$$y(t-j) = \begin{cases} v_{t+K_1-j} & (0 \leq j \leq K_1) \\ I_{t+K_1-j} & (K_1 < j \leq N-1) \end{cases} \quad (11.4-2)$$

Thus,

$$\begin{aligned}\mathbf{Y}_N(t) &= [y(t) \quad y(t-1) \cdots y(t-N+1)]^t \\ &= [v_{t+K_1} \cdots v_{t+1} \quad v_t \quad I_{t-1} \cdots I_{t-K_2}]^t\end{aligned}\quad (11.4-3)$$

11.4.1 Recursive Least-Squares (Kalman) Algorithm

The recursive least-squares (RLS) estimation of $\hat{I}(t)$ may be formulated as follows. Suppose we have observed the vectors $\mathbf{Y}_N(n)$, $n = 0, 1, \dots, t$, and we wish to determine the coefficient vector $\mathbf{C}_N(t)$ of the equalizer (linear or decision-feedback) that minimizes the time-average weighted squared error

$$\mathcal{E}_N^{LS} = \sum_{n=0}^t w^{t-n} |e_N(n, t)|^2 \quad (11.4-4)$$

where the error is defined as

$$e_N(n, t) = I(n) - \mathbf{C}'_N(t)\mathbf{Y}_N(n) \quad (11.4-5)$$

and w represents a weighting factor $0 < w < 1$. Thus we introduce exponential weighting into past data, which is appropriate when the channel characteristics are time-variant. Minimization of $\mathcal{E}_N^{LS}$ with respect to the coefficient vector $\mathbf{C}_N(t)$ yields the set of linear equations

$$\mathbf{R}_N(t)\mathbf{C}_N(t) = \mathbf{D}_N(t) \quad (11.4-6)$$

where $\mathbf{R}_N(t)$ is the signal correlation matrix defined as

$$\mathbf{R}_N(t) = \sum_{n=0}^t w^{t-n} \mathbf{Y}_N^*(n) \mathbf{Y}_N'(n) \quad (11.4-7)$$

and $\mathbf{D}_N(t)$ is the cross-correlation vector

$$\mathbf{D}_N(t) = \sum_{n=0}^t w^{t-n} I(n) \mathbf{Y}_N^*(n) \quad (11.4-8)$$

The solution of Equation 11.4-6 is

$$\mathbf{C}_N(t) = \mathbf{R}_N^{-1}(t) \mathbf{D}_N(t) \quad (11.4-9)$$

The matrix $\mathbf{R}_N(t)$ is akin to the statistical autocorrelation matrix $\mathbf{\Gamma}_N$, while the vector $\mathbf{D}_N(t)$ is akin to the cross-correlation vector $\mathbf{\xi}_N$, defined previously. We emphasize, however, that $\mathbf{R}_N(t)$ is not a Toeplitz matrix. We also should mention that, for small values of t , $\mathbf{R}_N(t)$ may be ill conditioned; hence, it is customary to initially add the matrix $\delta \mathbf{I}_N$ to $\mathbf{R}_N(t)$, where δ is a small positive constant and $\mathbf{I}_N$ is the identity matrix. With exponential weighting into the past, the effect of adding $\delta \mathbf{I}_N$ dissipates with time.

Now suppose we have the solution 11.4-9 for time $t-1$, i.e., $\mathbf{C}_N(t-1)$, and we wish to compute $\mathbf{C}_N(t)$. It is inefficient, and, hence, impractical to solve the set of N linear equations for each new signal component that is received. To avoid this, we proceed as follows. First, $\mathbf{R}_N(t)$ may be computed recursively as

$$\mathbf{R}_N(t) = w\mathbf{R}_N(t-1) + \mathbf{Y}_N^*(t) \mathbf{Y}_N'(t) \quad (11.4-10)$$

We call Equation 11.4-10 the *time-update equation* for $\mathbf{R}_N(t)$.

Since the inverse of $\mathbf{R}_N(t)$ is needed in Equation 11.4-9, we use the matrix-inverse identity

$$\mathbf{R}_N^{-1}(t) = \frac{1}{w} \left[\mathbf{R}_N^{-1}(t-1) - \frac{\mathbf{R}_N^{-1}(t-1) \mathbf{Y}_N^*(t) \mathbf{Y}_N'(t) \mathbf{R}_N^{-1}(t-1)}{w + \mathbf{Y}_N'(t) \mathbf{R}_N^{-1}(t-1) \mathbf{Y}_N^*(t)} \right] \quad (11.4-11)$$

Thus $\mathbf{R}_N^{-1}(t)$ may be computed recursively according to Equation 11.4-11.

For convenience, we define $\mathbf{P}_N(t) = \mathbf{R}_N^{-1}(t)$. It is also convenient to define an N -dimensional vector, called the *Kalman gain vector*, as

$$\mathbf{K}_N(t) = \frac{1}{w + \mu_N(t)} \mathbf{P}_N(t-1) \mathbf{Y}_N^*(t) \quad (11.4-12)$$

where $\mu_N(t)$ is a scalar defined as

$$\mu_N(t) = \mathbf{Y}'_N(t)\mathbf{P}_N(t-1)\mathbf{Y}_N^*(t) \quad (11.4-13)$$

With these definitions, Equation 11.4-11 becomes

$$\mathbf{P}_N(t) = \frac{1}{w}[\mathbf{P}_N(t-1) - \mathbf{K}_N(t)\mathbf{Y}'_N(t)\mathbf{P}_N(t-1)] \quad (11.4-14)$$

Suppose we postmultiply both sides of Equation 11.4-14 by $\mathbf{Y}_N^*(t)$. Then

$$\begin{aligned} \mathbf{P}_N(t)\mathbf{Y}_N^*(t) &= \frac{1}{w}[\mathbf{P}_N(t-1)\mathbf{Y}_N^*(t) - \mathbf{K}_N(t)\mathbf{Y}'_N(t)\mathbf{P}_N(t-1)\mathbf{Y}_N^*(t)] \\ &= \frac{1}{w}\{[w + \mu_N(t)]\mathbf{K}_N(t) - \mathbf{K}_N(t)\mu_N(t)\} \\ &= \mathbf{K}_N(t) \end{aligned} \quad (11.4-15)$$

Therefore, the Kalman gain vector may also be defined as $\mathbf{P}_N(t)\mathbf{Y}_N(t)$.

Now we use the matrix inversion identity to derive an equation for obtaining $\mathbf{C}_N(t)$ from $\mathbf{C}_N(t-1)$. Since

$$\mathbf{C}_N(t) = \mathbf{P}_N(t)\mathbf{D}_N(t)$$

and

$$\mathbf{D}_N(t) = w\mathbf{D}_N(t-1) + I(t)\mathbf{Y}_N^*(t) \quad (11.4-16)$$

we have

$$\begin{aligned} \mathbf{C}_N(t) &= \frac{1}{w}[\mathbf{P}_N(t-1) - \mathbf{K}_N(t)\mathbf{Y}'_N(t)\mathbf{P}_N(t-1)][w\mathbf{D}_N(t-1) + I(t)\mathbf{Y}_N^*(t)] \\ &= \mathbf{P}_N(t-1)\mathbf{D}_N(t-1) + \frac{1}{w}I(t)\mathbf{P}_N(t-1)\mathbf{Y}_N^*(t) \\ &\quad - \mathbf{K}_N(t)\mathbf{Y}'_N(t)\mathbf{P}_N(t-1)\mathbf{D}_N(t-1) \\ &\quad - \frac{1}{w}I(t)\mathbf{K}_N(t)\mathbf{Y}'_N(t)\mathbf{P}_N(t-1)\mathbf{Y}_N^*(t) \\ &= \mathbf{C}_N(t-1) + \mathbf{K}_N(t)[I(t) - \mathbf{Y}'_N(t)\mathbf{C}_N(t-1)] \end{aligned} \quad (11.4-17)$$

Note that $\mathbf{Y}'_N(t)\mathbf{C}_N(t-1)$ is the output of the equalizer at time t , i.e.,

$$\hat{I}(t) = \mathbf{Y}'_N(t)\mathbf{C}_N(t-1) \quad (11.4-18)$$

and

$$e_N(t, t-1) = I(t) - \hat{I}(t) \equiv e_N(t) \quad (11.4-19)$$

is the error between the desired symbol and the estimate. Hence, $\mathbf{C}_N(t)$ is updated recursively according to the relation

$$\mathbf{C}_N(t) = \mathbf{C}_N(t-1) + \mathbf{K}_N(t)e_N(t) \quad (11.4-20)$$

The residual MSE resulting from this optimization is

$$\mathcal{E}_{N \min}^{LS} = \sum_{n=0}^t w^{t-n} |I(n)|^2 - \mathbf{C}'_N(t) \mathbf{D}_N^*(t) \quad (11.4-21)$$

To summarize, suppose we have $\mathbf{C}_N(t-1)$ and $\mathbf{P}_N(t-1)$. When a new signal component is received, we have $\mathbf{Y}_N(t)$. Then the recursive computation for the time update of $\mathbf{C}_N(t)$ and $\mathbf{P}_N(t)$ proceeds as follows:

- Compute output:

$$\hat{I}(t) = \mathbf{Y}'_N(t) \mathbf{C}_N(t-1)$$

- Compute error:

$$e_N(t) = I(t) - \hat{I}(t)$$

- Compute Kalman gain vector:

$$\mathbf{K}_N(t) = \frac{\mathbf{P}_N(t-1) \mathbf{Y}'_N(t)}{w + \mathbf{Y}'_N(t) \mathbf{P}_N(t-1) \mathbf{Y}_N^*(t)}$$

- Update inverse of the correlation matrix:

$$\mathbf{P}_N(t) = \frac{1}{w} [\mathbf{P}_N(t-1) - \mathbf{K}_N(t) \mathbf{Y}'_N(t) \mathbf{P}_N(t-1)]$$

- Update coefficients:

$$\begin{aligned} \mathbf{C}_N(t) &= \mathbf{C}_N(t-1) + \mathbf{K}_N(t) e_N(t) \\ &= \mathbf{C}_N(t-1) + \mathbf{P}_N(t) \mathbf{Y}_N^*(t) e_N(t) \end{aligned} \quad (11.4-22)$$

The algorithm described by Equation 11.4-22 is called the *RLS direct form* or *Kalman algorithm*. It is appropriate when the equalizer has a transversal (direct-form) structure.

Note that the equalizer coefficients change with time by an amount equal to the error $e_N(t)$ multiplied by the Kalman gain vector $\mathbf{K}_N(t)$. Since $\mathbf{K}_N(t)$ is N -dimensional, each tap coefficient in effect is controlled by one of the elements of $\mathbf{K}_N(t)$. Consequently rapid convergence is obtained. In contrast, the steepest-descent algorithm, expressed in our present notation, is

$$\mathbf{C}_N(t) = \mathbf{C}_N(t-1) + \Delta \mathbf{Y}_N^*(t) e_N(t) \quad (11.4-23)$$

and the only variable parameter is the step size Δ .

Figure 11.4-1 illustrates the initial convergence rate of these two algorithms for a channel with fixed parameters $f_0 = 0.26$, $f_1 = 0.93$, $f_2 = 0.26$, and a linear equalizer with 11 taps. The eigenvalue ratio for this channel is $\lambda_{\max}/\lambda_{\min} = 11$. All the equalizer coefficients were initialized to zero. The steepest-descent algorithm was implemented with $\Delta = 0.020$. The superiority of the Kalman algorithm is clearly evident. This is especially important in a time-variant channel. For example, the time variations in the characteristics of an (ionospheric) high-frequency (HF) radio channel are too rapid to be equalized by the gradient

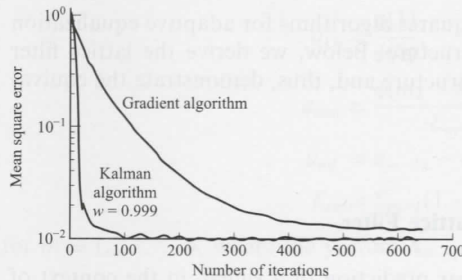


FIGURE 11.4-1
Comparison of convergence rate for the Kalman and gradient algorithms.

algorithm, but the Kalman algorithm adapts sufficiently rapidly to track such variations.

In spite of its superior convergence performance, the Kalman algorithm described above has two disadvantages. One is its complexity. The second is its sensitivity to roundoff noise that accumulates due to the recursive computations. The latter may cause instabilities in the algorithm.

The number of computations or operations (multiplications, divisions, and subtractions) in computing the variables in Equation 11.4-22 is proportional to N^2 . Most of these operations are involved in the updating of $\mathbf{P}_N(t)$. This part of the computation is also susceptible to roundoff noise. To remedy that problem, algorithms have been developed that avoid the computation of $\mathbf{P}_N(t)$ according to Equation 11.4-14. The basis of these algorithms lies in the decomposition of $\mathbf{P}_N(t)$ in the form

$$\mathbf{P}_N(t) = \mathbf{S}_N(t)\mathbf{\Lambda}_N(t)\mathbf{S}_N'(t) \quad (11.4-24)$$

where $\mathbf{S}_N(t)$ is a lower-triangular matrix whose diagonal elements are unity, and $\mathbf{\Lambda}_N(t)$ is a diagonal matrix. Such a decomposition is called a *square-root factorization* (see Bierman, 1977). This factorization is described in Appendix D. In a square-root algorithm, $\mathbf{P}_N(t)$ is not updated as in Equation 11.4-14 nor is it computed. Instead, the time updating is performed on $\mathbf{S}_N(t)$ and $\mathbf{\Lambda}_N(t)$.

Square-root algorithms are frequently used in control systems applications in which Kalman filtering is involved. In digital communications, the square-root Kalman algorithm has been implemented in a decision-feedback-equalized PSK modem designed to transmit at high speed over high-frequency radio channels with a nominal 3-kHz bandwidth. This algorithm is described in the paper by Hsu (1982). It has a computational complexity of $1.5N^2 + 6.5N$ (complex-valued multiplications and divisions per output symbol). It is also numerically stable and exhibits good numerical properties. For a detailed discussion of square-root algorithms in sequential estimation, the reader is referred to the book by Bierman (1977).

It is also possible to derive RLS algorithms with computational complexities that grow linearly with the number N of equalizer coefficients. Such algorithms are generally called *fast RLS algorithms* and have been described in the papers by Carayannis et al. (1983), Cioffi and Kailath (1984), and Slock and Kailath (1991).

Another class of recursive least squares algorithms for adaptive equalization are based on the lattice equalizer structure. Below, we derive the lattice filter structure from the transversal filter structure and, thus, demonstrate the equivalence of the two structures.

11.4.2 Linear Prediction and the Lattice Filter

In Chapter 3, we considered the linear prediction of a signal, in the context of speech encoding. In this section, we shall establish the connection between linear prediction and a lattice filter.

The linear prediction problem may be stated as follows: given a set of data $y(t-1), y(t-2), \dots, y(t-p)$, predict the value of the next data point $y(t)$. The predictor of order p is

$$\hat{y}(t) = \sum_{k=1}^p a_{pk} y(t-k) \quad (11.4-25)$$

Minimization of the MSE, defined as

$$\begin{aligned} \mathcal{E}_p &= E[y(t) - \hat{y}(t)]^2 \\ &= E \left[y(t) - \sum_{k=1}^p a_{pk} y(t-k) \right]^2 \end{aligned} \quad (11.4-26)$$

with respect to the predictor coefficients $\{a_{pk}\}$ yields the set of linear equations

$$\sum_{k=1}^p a_{pk} \phi(k-l) = \phi(l), \quad l = 1, 2, \dots, p \quad (11.4-27)$$

where

$$\phi(l) = E[y(t)y(t+l)]$$

These are called the *normal equations* or the *Yule–Walker equations*.

The matrix Φ with elements $\phi(k-l)$ is a Toeplitz matrix, and, hence, the Levinson–Durbin algorithm described in Appendix A provides an efficient means for solving the linear equations recursively, starting with a first-order predictor and proceeding recursively to the solution of the coefficients for the predictor of order p . The recursive relations for the Levinson–Durbin algorithm are

can be dramatically reduced by means of efficient modulation/coding and demodulation/decoding techniques.

14.1

CHARACTERIZATION OF FADING MULTIPATH CHANNELS

If we transmit an extremely short pulse, ideally an impulse, over a time-varying multipath channel, the received signal might appear as a train of pulses, as shown in Figure 14.1–1. Hence, one characteristic of a multipath medium is the time spread introduced in the signal that is transmitted through the channel.

A second characteristic is due to the time variations in the structure of the medium. As a result of such time variations, the nature of the multipath varies with time. That is, if we repeat the pulse-sounding experiment over and over, we shall observe changes in the received pulse train, which will include changes in the sizes of the individual pulses, changes in the relative delays among the pulses, and, quite often, changes in the number of pulses observed in the received pulse train as shown in Figure 14.1–1. Moreover, the time variations appear to be unpredictable to the user of the channel. Therefore, it is reasonable to characterize the time-variant multipath channel statistically. Toward this end, let us examine the effects of the channel on a transmitted signal that is represented in general as

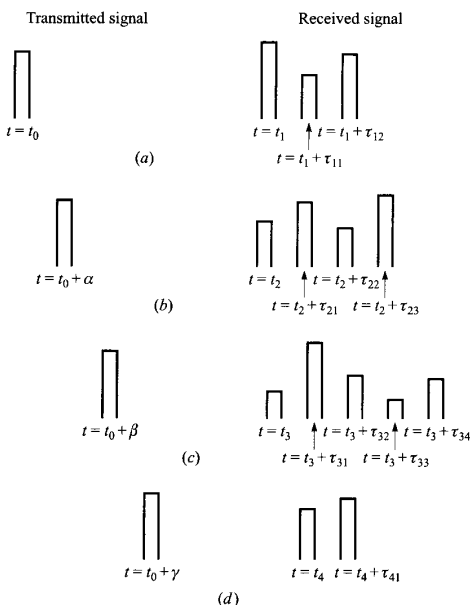
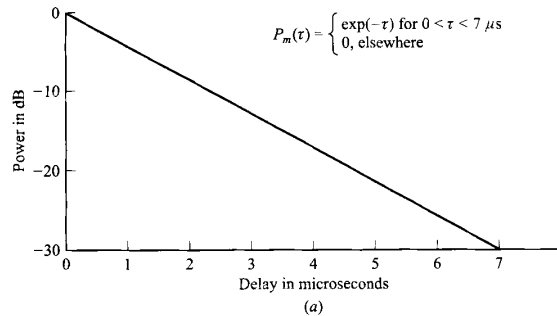
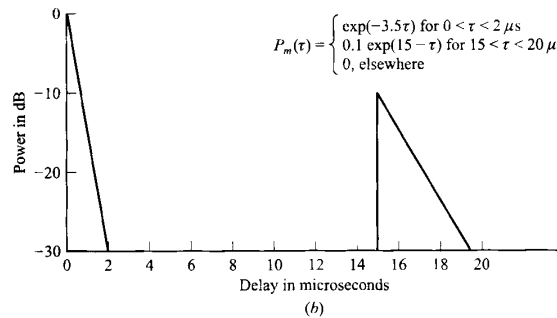


FIGURE 14.1–1

Example of the response of a time-variant multipath channel to a very narrow pulse.

**FIGURE 14.1-7**

Cost 207 average power delay profiles: (a) typical delay profile for suburban and urban areas; (b) typical “bad”-case delay profile for hilly terrain. [From Cost 207 Document 207 TD (86)51 rev 3.]



14.1.2 Statistical Models for Fading Channels

There are several probability distributions that can be considered in attempting to model the statistical characteristics of the fading channel. When there are a large number of scatterers in the channel that contribute to the signal at the receiver, as is the case in ionospheric or tropospheric signal propagation, application of the central limit theorem leads to a Gaussian process model for the channel impulse response. If the process is zero-mean, then the envelope of the channel response at any time instant has a Rayleigh probability distribution and the phase is uniformly distributed in the interval $(0, 2\pi)$. That is

$$p_R(r) = \frac{2r}{\Omega} e^{-r^2/\Omega}, \quad r \geq 0 \quad (14.1-23)$$

where

$$\Omega = E(R^2) \quad (14.1-24)$$

We observe that the Rayleigh distribution is characterized by the single parameter $E(R^2)$.

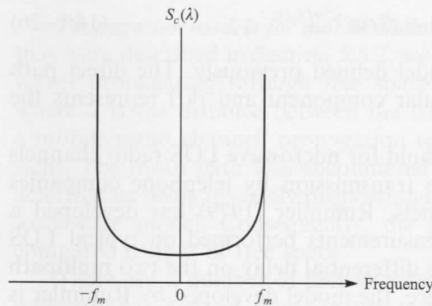


FIGURE 14.1-8
Model of Doppler spectrum for a mobile radio channel.

An alternative statistical model for the envelope of the channel response is the Nakagami- m distribution given by the PDF in Equation 2.1-147. In contrast to the Rayleigh distribution, which has a single parameter that can be used to match the fading channel statistics, the Nakagami- m is a two-parameter distribution, involving the parameter m and the second moment $\Omega = E(R^2)$. As a consequence, this distribution provides more flexibility and accuracy in matching the observed signal statistics. The Nakagami- m distribution can be used to model fading channel conditions that are either more or less severe than the Rayleigh distribution, and it includes the Rayleigh distribution as a special case ($m = 1$). For example, Turin et al. (1972) and Suzuki (1977) have shown that the Nakagami- m distribution is the best fit for data signals received in urban radio multipath channels.

The Rice distribution is also a two-parameter distribution. It may be expressed by the PDF given in Equation 2.1-141, where the parameters are s and σ^2 . Recall that s^2 is called the *noncentrality parameter* in the equivalent chi-square distribution. It represents the power in the nonfading signal components, sometimes called *specular components*, of the received signal.

There are many radio channels in which fading is encountered that are basically line-of-sight (LOS) communication links with multipath components arising from secondary reflections, or signal paths, from surrounding terrain. In such channels, the number of multipath components is small, and, hence, the channel may be modeled in a somewhat simpler form. We cite two channel models as examples.

As the first example, let us consider an airplane to ground communication link in which there is the direct path and a single multipath component at a delay t_0 relative to the direct path. The impulse response of such a channel may be modeled as

$$c(\tau; t) = \alpha\delta(\tau) + \beta(t)\delta[\tau - \tau_0(t)] \quad (14.1-25)$$

where α is the attenuation factor of the direct path and $\beta(t)$ represents the time-variant multipath signal component resulting from terrain reflections. Often, $\beta(t)$ can be characterized as a zero-mean Gaussian random process. The transfer function for this channel model may be expressed as

$$C(f; t) = \alpha + \beta(t)e^{-j2\pi f \tau_0(t)} \quad (14.1-26)$$

This channel fits the Ricean fading model defined previously. The direct path with attenuation α represents the specular component and $\beta(t)$ represents the Rayleigh fading component.

A similar model has been found to hold for microwave LOS radio channels used for long-distance voice and video transmission by telephone companies throughout the world. For such channels, Rummler (1979) has developed a three-path model based on channel measurements performed on typical LOS links in the 6-GHz frequency band. The differential delay on the two multipath components is relatively small, and, hence, the model developed by Rummler is one that has a channel transfer function

$$C(f) = \alpha[1 - \beta e^{-j2\pi(f-f_0)\tau_0}] \quad (14.1-27)$$

where α is the overall attenuation parameter, β is called a shape parameter which is due to the multipath components, f_0 is the frequency of the fade minimum, and τ_0 is the relative time delay between the direct and the multipath components. This simplified model was used to fit data derived from channel measurements.

Rummler found that the parameters α and β may be characterized as random variables that, for practical purposes, are nearly statistically independent. From the channel measurements, he found that the distribution of β has the form $(1 - \beta)^{2.3}$. The distribution of α is well modeled by the lognormal distribution, i.e., $-\log \alpha$ is Gaussian. For $\beta > 0.5$, the mean of $-20 \log \alpha$ was found to be 25 dB and the standard deviation was 5 dB. For smaller values of β , the mean decreases to 15 dB. The delay parameter determined from the measurements was $\tau_0 = 6.3$ ns. The magnitude-square response of $C(f)$ is

$$|C(f)|^2 = \alpha^2[1 + \beta^2 - 2\beta \cos 2\pi(f - f_0)\tau_0] \quad (14.1-28)$$

$|C(f)|$ is plotted in Figure 14.1-9 as a function of the frequency $f - f_0$ for $\tau_0 = 6.3$ ns. Note that the effect of the multipath component is to create a deep attenuation at $f = f_0$ and at multiples of $1/\tau_0 \approx 159$ MHz. By comparison, the typical channel bandwidth is 30 MHz. This model was used by Lundgren and Rummler (1979) to determine the error rate performance of digital radio systems.

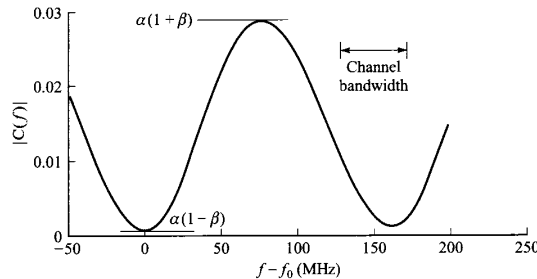


FIGURE 14.1-9
Magnitude frequency
response of LOS channel
model.

Propagation models for mobile radio channels. In the link budget calculations that were described in Section 5.5.2, we had characterized the path loss of radio waves propagating through free space as being inversely proportional to d^2 , where d is the distance between the transmitter and the receiver. However, in a mobile radio channel, propagation is generally neither free space nor line of sight. The mean path loss encountered in mobile radio channels may be characterized as being inversely proportional to d^p , where $2 \leq p \leq 4$, with d^4 being a worst-case model. Consequently, the path loss is usually much more severe compared to that of free space.

There are a number of factors affecting the path loss in mobile radio communications. Among these factors are base station antenna height, mobile antenna height, operating frequency, atmospheric conditions, and presence or absence of buildings and trees. Various mean path loss models have been developed that incorporate such factors. For example, a model for a large city in an urban area is the Hata model, in which the mean path loss is expressed as

$$\begin{aligned} \text{Loss in dB} = & 69.55 + 26.16 \log_{10} f - 13.82 \log_{10} h_t - a(h_r) \\ & + (44.9 - 6.55 \log_{10} h_t) \log_{10} d \end{aligned} \quad (14.1-29)$$

where f is the operating frequency in MHz ($150 < f < 1500$), h_t is the transmitter antenna height in meters ($30 < h_t < 200$), h_r is the receiver antenna height in meters ($1 < h_r < 10$), d is the distance between transmitter and receiver in km ($1 < d < 20$), and

$$a(h_r) = 3.2(\log_{10} 11.75h_r)^2 - 4.97, \quad f \geq 400 \text{ MHz} \quad (14.1-30)$$

Another problem with mobile radio propagation is the effect of shadowing of the signal due to large obstructions, such as large buildings, trees, and hilly terrain between the transmitter and the receiver. Shadowing is usually modeled as a multiplicative and, generally, slowly time varying random process. That is, the received signal may be characterized mathematically as

$$r(t) = A_0 g(t) s(t) \quad (14.1-31)$$

where A_0 represents the mean path loss, $s(t)$ is the transmitted signal, and $g(t)$ is a random process that represents the shadowing effect. At any time instant, the shadowing process is modeled statistically as lognormally distributed. The probability density function for the lognormal distribution is

$$p(g) = \begin{cases} \frac{1}{\sqrt{2\pi\sigma^2}g} e^{-(\ln g - \mu)^2 / 2\sigma^2} & (g \geq 0) \\ 0 & (g < 0) \end{cases} \quad (14.1-32)$$

If we define a new random variable X as $X = \ln g$, then

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2 / 2\sigma^2}, \quad -\infty < x < \infty \quad (14.1-33)$$

The random variable X represents the path loss measured in dB, μ is the mean path loss in dB, and σ is the standard deviation of the path loss in dB. For typical cellular and microcellular environments, σ is in the range of 5–12 dB.

14.2

THE EFFECT OF SIGNAL CHARACTERISTICS ON THE CHOICE OF A CHANNEL MODEL

Having discussed the statistical characterization of time-variant multipath channels generally in terms of the correlation functions describe in Section 14.1, we now consider the effect of signal characteristics on the selection of a channel model that is appropriate for the specified signal. Thus, let $s_i(t)$ be the equivalent low-pass signal transmitted over the channel and let $S_i(f)$ denote its frequency content. Then the equivalent low-pass received signal, exclusive of additive noise, may be expressed either in terms of the time-domain variables $c(\tau; t)$ and $s_i(t)$ as

$$r_i(t) = \int_{-\infty}^{\infty} c(\tau; t) s_i(t - \tau) d\tau \quad (14.2-1)$$

or in terms of the frequency functions $C(f; t)$ and $S_i(f)$ as

$$r_i(t) = \int_{-\infty}^{\infty} C(f; t) S_i(f) e^{j2\pi ft} df \quad (14.2-2)$$

Suppose we are transmitting digital information over the channel by modulating (either in amplitude, or in phase, or both) the basic pulse $s_i(t)$ at a rate $1/T$, where T is the signaling interval. It is apparent from Equation 14.2-2 that the time-variant channel characterized by the transfer function $C(f; t)$ distorts the signal $S_i(f)$. If $S_i(f)$ has a bandwidth W greater than the coherence bandwidth $(\Delta f)_c$ of the channel, $S_i(f)$ is subjected to different gains and phase shifts across the band. In such a case, the channel is said to be *frequency-selective*. Additional distortion is caused by the time variations in $C(f; t)$. This type of distortion is evidenced as a variation in the received signal strength, and has been termed *fading*. It should be emphasized that the frequency selectivity and fading are viewed as two different types of distortion. The former depends on the multipath spread or, equivalently, on the coherence bandwidth of the channel relative to the transmitted signal bandwidth W . The latter depends on the time variations of the channel, which are grossly characterized by the coherence time $(\Delta t)_c$ or, equivalently, by the Doppler spread B_d .

The effect of the channel on the transmitted signal $s_i(t)$ is a function of our choice of signal bandwidth and signal duration. For example, if we select the signaling interval T to satisfy the condition $T \gg T_m$, the channel introduces a negligible amount of intersymbol interference. If the bandwidth of the signal pulse $s_i(t)$ is $W \approx 1/T$, the condition $T \gg T_m$ implies that

$$W \ll \frac{1}{T_m} \approx (\Delta f)_c \quad (14.2-3)$$

That is, the signal bandwidth W is much smaller than the coherence bandwidth of the channel. Hence, the channel is *frequency-nonselective*. In other words, all the frequency components in $S_l(f)$ undergo the same attenuation and phase shift in transmission through the channel. But this implies that, within the bandwidth occupied by $S_l(f)$, the time-variant transfer function $C(f; t)$ of the channel is a complex-valued constant in the frequency variable. Since $S_l(f)$ has its frequency content concentrated in the vicinity of $f = 0$, $C(f; t) = C(0; t)$. Consequently, Equation 14.2-2 reduces to

$$\begin{aligned} r_l(t) &= C(0; t) \int_{-\infty}^{\infty} S_l(f) e^{j2\pi ft} df \\ &= C(0; t) s_l(t) \end{aligned} \quad (14.2-4)$$

Thus, when the signal bandwidth W is much smaller than the coherence bandwidth $(\Delta f)_c$ of the channel, the received signal is simply the transmitted signal multiplied by a complex-valued random process $C(0; t)$, which represents the time-variant characteristics of the channel. In this case, we say that the multipath components in the received are not resolvable because $W \ll (\Delta f)_c$.

The transfer function $C(0; t)$ for a frequency-nonselective channel may be expressed in the form

$$C(0; t) = \alpha(t) e^{-j\phi(t)} \quad (14.2-5)$$

where $\alpha(t)$ represents the envelope and $\phi(t)$ represents the phase of the equivalent low-pass channel. When $C(0; t)$ is modeled as a zero-mean complex-valued Gaussian random process, the envelope $\alpha(t)$ is Rayleigh-distributed for any fixed value of t and $\phi(t)$ is uniformly distributed over the interval $(-\pi, \pi)$. The rapidity of the fading on the frequency-nonselective channel is determined either from the correlation function $\phi_C(\Delta t)$ or from the Doppler power spectrum $S_C(\lambda)$. Alternatively, either of the channel parameters $(\Delta t)_c$ or B_d can be used to characterize the rapidity of the fading.

For example, suppose it is possible to select the signal bandwidth W to satisfy the condition $W \ll (\Delta f)_c$ and the signaling interval T to satisfy the condition $T \ll (\Delta t)_c$. Since T is smaller than the coherence time of the channel, the channel attenuation and phase shift are essentially fixed for the duration of at least one signaling interval. When this condition holds, we call the channel a *slowly fading channel*. Furthermore, when $W \approx 1/T$, the conditions that the channel be frequency-nonselective and slowly fading imply that the product of T_m and B_d must satisfy the condition $T_m B_d < 1$.

The product $T_m B_d$ is called the *spread factor* of the channel. If $T_m B_d < 1$, the channel is said to be *underspread*; otherwise, it is *overspread*. The multipath spread, the Doppler spread, and the spread factor are listed in Table 14.2-1 for several channels. We observe from this table that several radio channels, including the moon when used as a passive reflector, are underspread. Consequently, it is possible to select the signal $s_l(t)$ such that these channels are frequency-nonselective and slowly fading. The slow-fading condition implies



TABLE 14.2-1

Multipath spread, Doppler spread, and spread factor for several time-variant multipath channels

Type of channel	Multipath duration, s	Doppler spread, Hz	Spread factor
Shortwave ionospheric propagation (HF)	10^{-3} – 10^{-2}	10^{-1} –1	10^{-4} – 10^{-2}
Ionospheric propagation under distributed auroral conditions (HF)	10^{-3} – 10^{-2}	10–100	10^{-2} –1
Ionospheric forward scatter (VHF)	10^{-4}	10	10^{-3}
Tropospheric scatter (SHF)	10^{-6}	10	10^{-5}
Orbital scatter (X' band)	10^{-4}	10^3	10^{-1}
Moon at max. libration ($f_0 = 0.4$ kmc)	10^{-2}	10	10^{-1}

that the channel characteristics vary sufficiently slowly that they can be measured.

In Section 14.3, we shall determine the error rate performance for binary signaling over a frequency-nonselctive slowly fading channel. This channel model is, by far, the simplest to analyze. More importantly, it yields insight into the performance characteristics for digital signaling on a fading channel and serves to suggest the type of signal waveforms that are effective in overcoming the fading caused by the channel.

Since the multipath components in the received signal are not resolvable when the signal bandwidth W is less than the coherence bandwidth $(\Delta f)_c$ of the channel, the received signal appears to arrive at the receiver via a single fading path. On the other hand, we may choose $W \gg (\Delta f)_c$, so that the channel becomes frequency-selective. We shall show later that, under this condition, the multipath components in the received signal are resolvable with a resolution in time delay of $1/W$. Thus, we shall illustrate that the frequency-selective channel can be modeled as a tapped delay line (transversal) filter with time-variant tap coefficients. We shall then derive the performance of binary signaling over such a frequency-selective channel model.



14.3

FREQUENCY-NONSELECTIVE, SLOWLY FADING CHANNEL

In this section, we derive the error rate performance of binary PSK and binary FSK when these signals are transmitted over a frequency-nonselctive, slowly fading channel. As described in Section 14.2, the frequency-nonselctive channel results in multiplicative distortion of the transmitted signal $s_f(t)$. Furthermore, the condition that the channel fades slowly implies that the multiplicative process may be regarded as a constant during at least one signaling interval. Consequently, if the transmitted signal is $s_f(t)$, the received equivalent low-pass signal in one signaling interval is

$$r_f(t) = \alpha e^{-j\phi} s_f(t) + z(t), \quad 0 \leq t \leq T \quad (14.3-1)$$

where $z(t)$ represents the complex-valued white Gaussian noise process corrupting the signal.

Let us assume that the channel fading is sufficiently slow that the phase shift ϕ can be estimated from the received signal without error. In that case, we can achieve ideal coherent detection of the received signal. Thus, the received signal can be processed by passing it through a matched filter in the case of binary PSK or through a pair of matched filters in the case of binary FSK. One method that we can use to determine the performance of the binary communication systems is to evaluate the decision variables and from these determine the probability of error. However, we have already done this for a fixed (time-invariant) channel. That is, for a fixed attenuation α , we know the probability of error for binary PSK and binary FSK. From Equation 5.2-5, the expression for the error rate of binary PSK as a function of the received SNR γ_b is

$$P_2(\gamma_b) = Q(\sqrt{2\gamma_b}) \quad (14.3-2)$$

where $\gamma_b = \alpha^2 \mathcal{E}_b / N_0$. The expression for the error rate of binary FSK, detected coherently, is given by Equation 5.2-10 as

$$P_2(\gamma_b) = Q(\sqrt{\gamma_b}) \quad (14.3-3)$$

We view Equation 14.3-2 and 14.3-3 as conditional error probabilities, where the condition is that α is fixed. To obtain the error probabilities when α is random, we must average $P_2(\gamma_b)$, given in Equations 14.3-2 and 14.3-3, over the probability density function of γ_b . That is, we must evaluate the integral

$$P_2 = \int_0^\infty P_2(\gamma_b) p(\gamma_b) d\gamma_b \quad (14.3-4)$$

where $p(\gamma_b)$ is the probability density function of γ_b when α is random.

Rayleigh fading. When α is Rayleigh-distributed, α^2 has a chi-square probability distribution with two degrees of freedom. Consequently, γ_b also is chi-square-distributed. It is easily shown that

$$p(\gamma_b) = \frac{1}{\bar{\gamma}_b} e^{-\gamma_b/\bar{\gamma}_b}, \quad \gamma_b \geq 0 \quad (14.3-5)$$

where $\bar{\gamma}_b$ is the average signal-to-noise ratio, defined as

$$\bar{\gamma}_b = \frac{\mathcal{E}_b}{N_0} E(\alpha^2) \quad (14.3-6)$$

The term $E(\alpha^2)$ is simply the average value of α^2 .

Now we can substitute Equation 14.3-5 into Equation 14.3-4 and carry out the integration for $P_2(\gamma_b)$ as given by Equations 14.3-2 and 14.3-3. The result of this integration for binary PSK is

$$P_2 = \frac{1}{2} \left(1 - \sqrt{\frac{\bar{\gamma}_b}{1 + \bar{\gamma}_b}} \right) \quad (14.3-7)$$

If we repeat the integration with $P_2(\gamma_b)$ given by Equation 14.3-3, we obtain the probability of error for binary FSK, detected coherently, in the form

$$P_2 = \frac{1}{2} \left(1 - \sqrt{\frac{\bar{\gamma}_b}{2 + \bar{\gamma}_b}} \right) \quad (14.3-8)$$

In arriving at the error rate results in Equations 14.3-7 and 14.3-8, we have assumed that the estimate of the channel phase shift, obtained in the presence of slow fading, is noiseless. Such an ideal condition may not hold in practice. In such a case, the expressions in Equations 14.3-7 and 14.3-8 should be viewed as representing the best achievable performance in the presence of Rayleigh fading. In Appendix C we consider the problem of estimating the phase in the presence of noise and we evaluate the error rate performance of binary and multiphase PSK.

On channels for which the fading is sufficiently rapid to preclude the estimation of a stable phase reference by averaging the received signal phase over many signaling intervals, DPSK, is an alternative signaling method. Since DPSK requires phase stability over only two consecutive signaling intervals, this modulation technique is quite robust in the presence of signal fading. In deriving the performance of binary DPSK for a fading channel, we begin again with the error probability for a nonfading channel, which is

$$P_2(\gamma_b) = \frac{1}{2} e^{-\gamma_b} \quad (14.3-9)$$

This expression is substituted into the integral in Equation 14.3-4 along with $p(\gamma_b)$ obtained from Equation 14.3-5. Evaluation of the resulting integral yields the probability of error for binary DPSK, in the form

$$P_2 = \frac{1}{2(1 + \bar{\gamma}_b)} \quad (14.3-10)$$

If we choose not to estimate the channel phase shift at all, but instead employ a noncoherent (envelope or square-law) detector with binary, orthogonal FSK signals, the error probability for a nonfading channel is

$$P_2(\gamma_b) = \frac{1}{2} e^{-\gamma_b/2} \quad (14.3-11)$$

When we average $P_2(\gamma_b)$ over the Rayleigh fading channel attenuation, the resulting error probability is

$$P_2 = \frac{1}{2 + \bar{\gamma}_b} \quad (14.3-12)$$

The error probabilities in Equations 14.3-7, 14.3-8, 14.3-10, and 14.3-12 are illustrated in Figure 14.3-1. In comparing the performance of the four binary signaling systems, we focus our attention on the probabilities of error for large

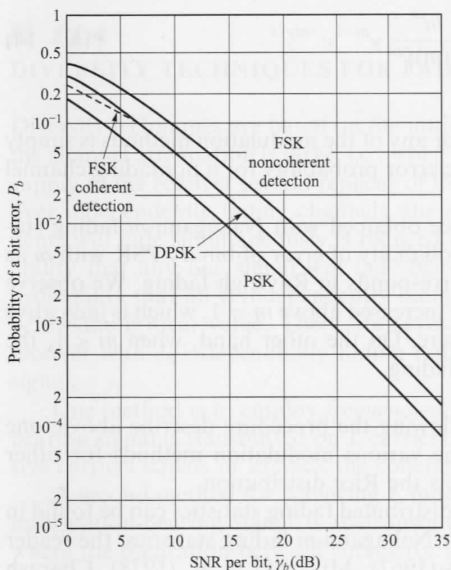


FIGURE 14.3-1
Performance of binary signaling on a Rayleigh fading channel.

SNR, i.e., $\bar{\gamma}_b \gg 1$. Under this condition, the error rates in Equations 14.3-7, 14.3-8, 14.3-10, and 14.3-12 simplify to

$$P_2 \approx \begin{cases} 1/4\bar{\gamma}_b & \text{for coherent PSK} \\ 1/2\bar{\gamma}_b & \text{for coherent, orthogonal FSK} \\ 1/2\bar{\gamma}_b & \text{for DPSK} \\ 1/\bar{\gamma}_b & \text{for noncoherent, orthogonal FSK} \end{cases} \quad (14.3-13)$$

From Equation 14.3-13, we observe that coherent PSK is 3 dB better than DPSK and 6 dB better than noncoherent FSK. More striking, however, is the observation that the error rates decrease only inversely with SNR. In contrast, the decrease in error rate on a nonfading channel is exponential with SNR. This means that, on a fading channel, the transmitter must transmit a large amount of power in order to obtain a low probability of error. In many cases, a large amount of power is not possible, technically and/or economically. An alternative solution to the problem of obtaining acceptable performance on a fading channel is the use of redundancy, which can be obtained by means of diversity techniques, as discussed in Section 14.4.

Nakagami fading. If α is characterized statistically by the Nakagami- m distribution, the random variable $\gamma = \alpha^2 \mathcal{E}_b / N_0$ has the PDF (see Problem 14.15)

$$p(\gamma) = \frac{m^m}{\Gamma(m)\bar{\gamma}^m} \gamma^{m-1} e^{-m\gamma/\bar{\gamma}} \quad (14.3-14)$$

where $\bar{\gamma} = E(\alpha^2)\mathcal{E}/N_0$.

The average probability of error for any of the modulation methods is simply obtained by averaging the appropriate error probability for a nonfading channel over the fading signal statistics.

As an example of the performance obtained with Nakagami- m fading statistics, Figure 14.3-2 illustrates the probability of error of binary PSK with m as a parameter. We recall that $m = 1$ corresponds to Rayleigh fading. We observe that the performance improves as m is increased above $m = 1$, which is indicative of the fact that the fading is less severe. On the other hand, when $m < 1$, the performance is worse than Rayleigh fading.

Other fading signal statistics. Following the procedure describe above, one can determine the performance of the various modulation methods for other types of fading signal statistics, such as the Rice distribution.

Error probability results for Rice-distributed fading statistics can be found in the paper by Lindsey (1964), while for Nakagami- m fading statistics, the reader may refer to the papers by Esposito (1967), Miyagaki et al. (1978), Charash (1979), Al-Hussaini et al. (1985), and Beaulieu and Abu-Dayya (1991).

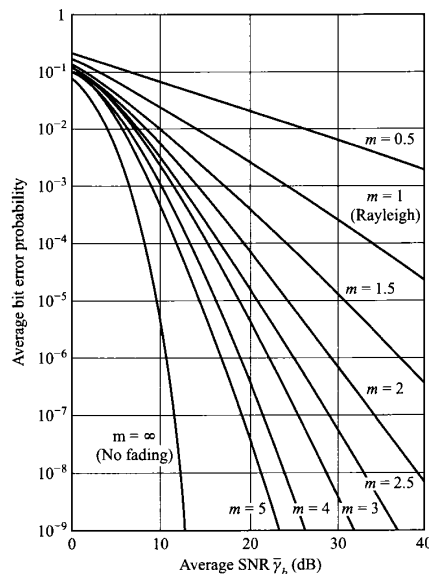


FIGURE 14.3-2

Average error probability for two-phase PSK with Nakagami fading.

■ 14.4

DIVERSITY TECHNIQUES FOR FADING MULTIPATH CHANNELS

Diversity techniques are based on the notion that errors occur in reception when the channel attenuation is large, i.e., when the channel is in a deep fade. If we can supply to the receiver several replicas of the same information signal transmitted over independently fading channels, the probability that all the signal components will fade simultaneously is reduced considerably. That is, if p is the probability that any one signal will fade below some critical value, then p^L is the probability that all L independently fading replicas of the same signal will fade below the critical value. There are several ways in which we can provide the receiver with L independently fading replicas of the same information-bearing signal.

One method is to employ *frequency diversity*. That is, the same information-bearing signal is transmitted on L carriers, where the separation between successive carriers equals or exceeds the coherence bandwidth $(\Delta f)_c$ of the channel.

A second method for achieving L independently fading versions of the same information-bearing signal is to transmit the signal in L different time slots, where the separation between successive time slots equals or exceeds the coherence time $(\Delta t)_c$ of the channel. This method is called *time diversity*.

Note that the fading channel fits the model of a bursty error channel. Furthermore, we may view the transmission of the same information either at different frequencies or in different time slots (or both) as a simple form of repetition coding. The separation of the diversity transmissions in time by $(\Delta t)_c$ or in frequency by $(\Delta f)_c$ is basically a form of block-interleaving the bits in the repetition code in an attempt to break up the error bursts and, thus, to obtain independent errors. Later in the chapter, we shall demonstrate that, in general, repetition coding is wasteful of bandwidth when compared with nontrivial coding.

Another commonly used method for achieving diversity employs multiple antennas. For example, we may employ a single transmitting antenna and multiple receiving antennas. The latter must be spaced sufficiently far apart that the multipath components in the signal have significantly different propagation delays at the antennas. Usually a separation of a few wavelengths is required between two antennas in order to obtain signals that fade independently.

A more sophisticated method for obtaining diversity is based on the use of a signal having a bandwidth much greater than the coherence bandwidth $(\Delta f)_c$ of the channel. Such a signal with bandwidth W will resolve the multipath components and, thus, provide the receiver with several independently fading signal paths. The time resolution is $1/W$. Consequently, with a multipath spread of T_m seconds, there are $T_m W$ resolvable signal components. Since $T_m \approx 1/(\Delta f)_c$, the number of resolvable signal components may also be expressed as $W/(\Delta f)_c$. Thus, the use of a wideband signal may be viewed as just another method for obtaining frequency diversity of order $L \approx W/(\Delta f)_c$. The optimum receiver for processing the wideband signal will be derived in Section 14.5. It is called a

RAKE correlator or a *RAKE matched filter* and was invented by Price and Green (1958).

There are other diversity techniques that have received some consideration in practice, such as angle-of-arrival diversity and polarization diversity. However, these have not been as widely used as those described above.

14.4.1 Binary Signals

We shall now determine the error rate performance for a binary digital communication system with diversity. We begin by describing the mathematical model for the communication system with diversity. First of all, we assume that there are L diversity channels, carrying the same information-bearing signal. Each channel is assumed to be frequency-nonselective and slowly fading with Rayleigh-distributed envelope statistics. The fading processes among the L diversity channels are assumed to be mutually statistically independent. The signal in each channel is corrupted by an additive zero-mean white Gaussian noise process. The noise processes in the L channels are assumed to be mutually statistically independent, with identical autocorrelation functions. Thus, the equivalent low-pass received signals for the L channels can be expressed in the form

$$r_{lk}(t) = \alpha_k e^{-j\phi_k} s_{km}(t) + z_k(t), \quad k = 1, 2, \dots, L, \quad m = 1, 2 \quad (14.4-1)$$

where $\{\alpha_k e^{-j\phi_k}\}$ represent the attenuation factors and phase shifts for the L channels, $s_{km}(t)$ denotes the m th signal transmitted on the k th channel, and $z_k(t)$ denotes the additive white Gaussian noise on the k th channel. All signals in the set $\{s_{km}(t)\}$ have the same energy.

The optimum demodulator for the signal received from the k th channel consists of two matched filters, one having the impulse response

$$b_{k1}(t) = s_{k1}^*(T - t) \quad (14.4-2)$$

and the other having the impulse response

$$b_{k2}(t) = s_{k2}^*(T - t) \quad (14.4-3)$$

Of course, if binary PSK is the modulation method used to transmit the information, then $s_{k1}(t) = -s_{k2}(t)$. Consequently, only a single matched filter is required for binary PSK. Following the matched filters is a combiner that forms the two decision variables. The combiner that achieves the best performance is one in which each matched filter output is multiplied by the corresponding complex-valued (conjugate) channel gain $\alpha_k e^{j\phi_k}$. The effect of this multiplication is to compensate for the phase shift in the channel and to weight the signal by a factor that is proportional to the signal strength. Thus, a strong signal carries a larger weight than a weak signal. After the complex-valued weighting operation is performed, two sums are formed. One consists of the real parts of the weighted outputs from the matched filters corresponding to a transmitted 0. The second consists of the real part of the outputs from the matched filters corresponding to a transmitted 1. This optimum combiner is