FUTURE VISION BIE

One Stop for All Study Materials

& Lab Programs

Future Vision

By K B Hemanth Raj

Scan the QR Code to Visit the Web Page

Or

Visit : https://hemanthrajhemu.github.io

Gain Access to All Study Materials according to VTU,
CSE – Computer Science Engineering,
ISE – Information Science Engineering,
ECE - Electronics and Communication Engineering
& MORE...

Join Telegram to get Instant Updates: https://bit.ly/VTU_TELEGRAM

Contact: MAIL: futurevisionbie@gmail.com

INSTAGRAM: www.instagram.com/hemanthraj_hemu/

INSTAGRAM: www.instagram.com/futurevisionbie/

WHATSAPP SHARE: https://bit.ly/FVBIESHARE

DIGITAL COMMUNICATION SYSTEMS

Simon Haykin
McMaster University

WILEY

8 Signaling over Band-Limited Channels 445

8.1 Introduction 445
8.2 Error Rate Due to Channel Noise in a Matched-Filter Receiver 446
8.3 Intersymbol Interference 447
8.4 Signal Design for Zero ISI 450
8.5 Ideal Nyquist Pulse for Distortionless Baseband Data Transmission 450
8.6 Raised-Cosine Spectrum 454
8.7 Square-Root Raised-Cosine Spectrum 458
8.8 Post-Processing Techniques: The Eye Pattern 463
8.9 Adaptive Equalization 469
8.10 Broadband Backbone Data Network: Signaling over Multiple Baseband Channels 474
8.11 Digital Subscriber Lines 475
8.12 Capacity of AWGN Channel Revisited 477
8.13 Partitioning Continuous-Time Channel into a Set of Subchannels 478
8.14 Water-Filling Interpretation of the Constrained Optimization Problem 484
8.15 DMT System Using Discrete Fourier Transform 487
8.16 Summary and Discussion 494

9 Signaling over Fading Channels 501

9.1 Introduction 501
9.2 Propagation Effects 502
9.3 Jakes Model 506
9.4 Statistical Characterization of Wideband Wireless Channels 511
9.5 FIR Modeling of Doubly Spread Channels 520
9.6 Comparison of Modulation Schemes: Effects of Flat Fading 525
9.7 Diversity Techniques 527
9.8 "Space Diversity-on-Receive" Systems 528
9.9 "Space Diversity-on-Transmit" Systems 538
9.10 "Multiple-Input, Multiple-Output" Systems: Basic Considerations 546
9.11 MIMO Capacity for Channel Known at the Receiver 551
9.12 Orthogonal Frequency Division Multiplexing 556
9.13 Spread Spectrum Signals 557
9.14 Code-Division Multiple Access 560
9.15 The RAKE Receiver and Multipath Diversity 564
9.16 Summary and Discussion 566

CHAPTER
9

Signaling over Fading Channels

9.1 Introduction

In Chapters 7 and 8 we studied signaling over AWGN and band-limited channels, respectively. In this chapter we go on to study a more complicated communications environment, namely a *fading channel*, which is at the very core of ever-expanding wireless communications. Fading refers to the fact that even though the distance separating a *mobile* receiver from the transmitter is essentially constant, a relatively small movement of the receiver away from the transmitter could result in a significant change in the received power. The physical phenomenon responsible for fading is *multipath*, which means that the transmitted signal reaches the mobile receiver via multiple paths with varying *spatio-temporal characteristics*, hence the challenging nature of the wireless channel for reliable communication.

This chapter consists of three related parts:

First we study signaling over a fading channel by characterizing its statistical behavior in temporal as well as spacial terms. This statistical characterization is carried out from three different perspectives: physical, mathematical, and computational, each of which enriches our understanding of the multipath phenomenon in its own way. This first part of the chapter finishes with:

- BER comparison of different modulation schemes for AWGN and Rayleigh fading channels.
- Graphical display of how different fading channels compare to a corresponding AWGN channel using binary PSK.

This evaluation then prompts the issue of how to combat the degrading effect of multipath and thereby realize reliable communication over a fading channel. Indeed, the second part of the chapter is devoted to this important practical issue. Specifically, we study the use of *space diversity*, which can be one of three kinds:

1. *Diversity-on-receive*, which involves the use of a single transmitter and multiple receivers, with each receiver having its own antenna.
2. *Diversity-on-transmit*, which involves the use of multiple transmitting antennas and a single receiver.
3. *Multiple-input, multiple-output* (MIMO) antenna system, which includes diversity on receive and diversity on transmit in a combined manner.

The use of diversity-on-receive techniques is of long standing in the study of radio communications. On the other hand, diversity-on-transmit and MIMO antenna systems are

of recent origin. The study of diversity is closely related to that of information capacity, the evaluation of which is also given special attention in the latter part of the chapter.

For the third and final part of the chapter, we study spread-spectrum signals, which provide the basis of another novel way of thinking about how to mitigate the degrading effects of the multipath phenomenon. In more specific terms, the use of spread-spectrum signaling leads to the formulation of code-division multiple access, a topic that was covered briefly in the introductory Chapter 1.

9.2 Propagation Effects

The major propagation problems[1] encountered in the use of mobile radio in built-up areas are due to the fact that the antenna of a mobile unit may lie well below the surrounding buildings. Simply put, there is no "line-of-sight" path to the base station. Instead, radio propagation takes place mainly by way of scattering from the surfaces of the surrounding buildings and by diffraction over and/or around them, as illustrated in Figure 9.1. The important point to note from Figure 9.1 is that energy reaches the receiving antenna via more than one path. Accordingly, we speak of a *multipath phenomenon,* in that the various incoming radio waves reach their destination from different directions and with different time delays.

To understand the nature of the multipath phenomenon, consider first a "static" multipath environment involving a stationary receiver and a transmitted signal that consists of a narrowband signal (e.g., unmodulated sinusoidal carrier). Let it be assumed that two attenuated versions of the transmitted signal arrive sequentially at the receiver. The effect of the differential time delay is to introduce a relative phase shift between any two components of the received signal. We may then identify one of two extreme cases that can arise:

- The relative phase shift is zero, in which case the two components add *constructively*, as illustrated in Figure 9.2a.
- The relative phase shift is 180°, in which case the two components add *destructively*, as illustrated in Figure 9.2b.

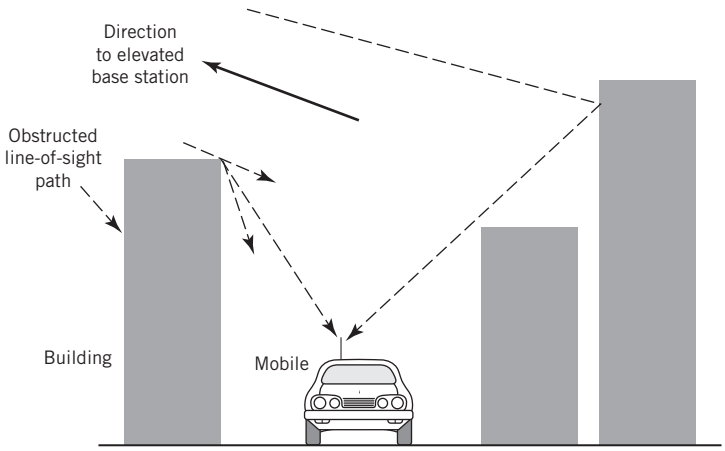


Figure 9.1
Illustrating the mechanism of radio propagation in urban areas.

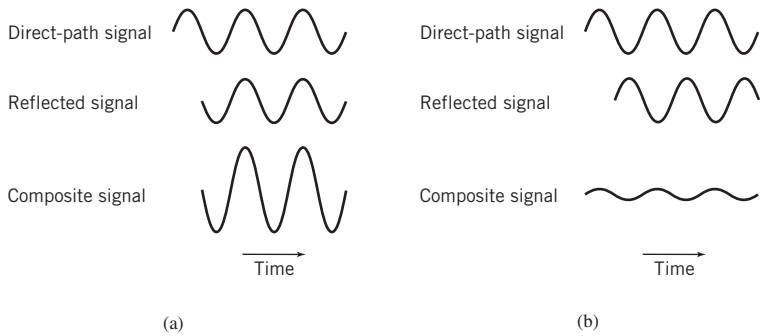


Figure 9.2 (a) Constructive and (b) destructive forms of the multipath phenomenon for sinusoidal signals.

We may also use *phasors* to demonstrate the constructive and destructive effects of multipath, as shown in Figures 9.3a and 9.3b, respectively. Note that, in the static multipath environment described herein, the amplitude of the received signal does not vary with time.

Consider next a "dynamic" multipath environment in which the receiver is in motion and two versions of the transmitted narrowband signal reach the receiver via paths of different

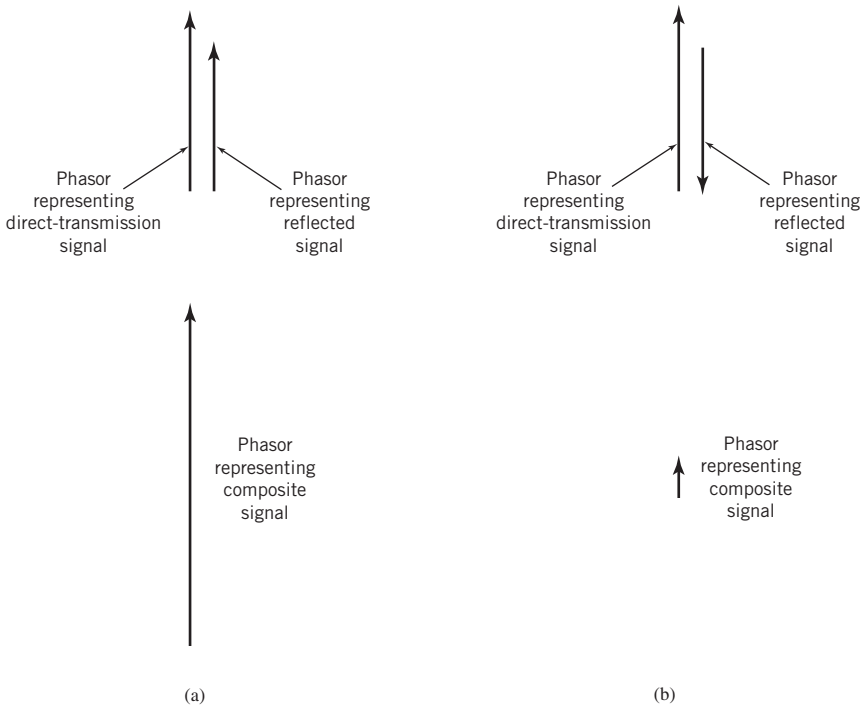


Figure 9.3 Phasor representations of (a) constructive and (b) destructive forms of multipath.

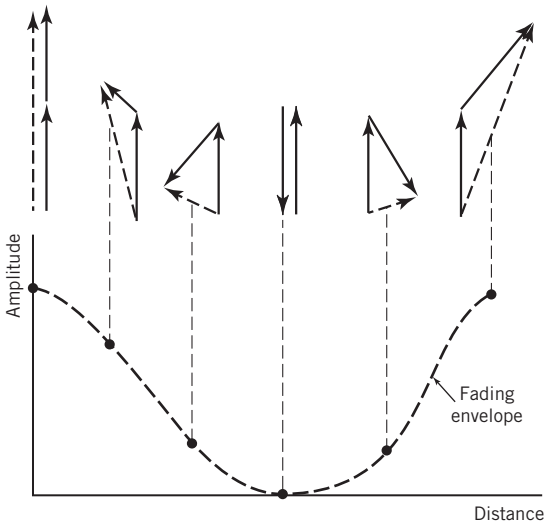


Figure 9.4
Illustrating how the envelope fades as two incoming signals combine with different phases.

lengths. Owing to motion of the receiver, there is a continuous change in the length of each propagation path. Hence, the relative phase shift between the two components of the received signal is a function of spatial location of the receiver. As the receiver moves, we now find that the received amplitude (envelope) is no longer constant, as was the case in a static environment; rather, it varies with distance, as illustrated in Figure 9.4. At the top of this figure, we have also included the phasor relationships for two components of the received signal at various locations of the receiver. Figure 9.4 shows that there is constructive addition at some locations and almost complete cancellation at some other locations. This physical phenomenon is referred to as *fast fading*.

In a mobile radio environment encountered in practice, there may of course be a multitude of propagation paths with different lengths and their contributions to the received signal could combine in a variety of ways. The net result is that the envelope of the received signal varies with location in a complicated fashion, as shown by the experimental record of received signal envelope in an urban area that is presented in Figure 9.5. This figure clearly displays the fading nature of the received signal. The received signal envelope in Figure 9.5 is measured in dBm. The unit dBm is defined as $10 \log_{10}(P/P_0)$, with P denoting the power being measured and $P_0 = 1$ mW as the frame of reference. In the case of Figure 9.5, P is the instantaneous power in the received signal envelope.

Signal fading is essentially a *spatial phenomenon* that manifests itself in the time domain as the receiver moves. These variations can be related to the motion of the receiver as follows. Consider the situation illustrated in Figure 9.6, where the receiver is assumed to be moving along the line AA′ with a constant velocity v. It is also assumed that the received signal is due to a radio wave from a scatterer labelled S. Let Δt denote the time taken for the receiver to move from point A to A′. Using the notation described in Figure 9.6, the incremental change in the path length of the radio wave is deduced to be

$$\begin{aligned} \Delta l &= d \cos \psi \\ &= -v\Delta t \cos \psi \end{aligned} \tag{9.1}$$

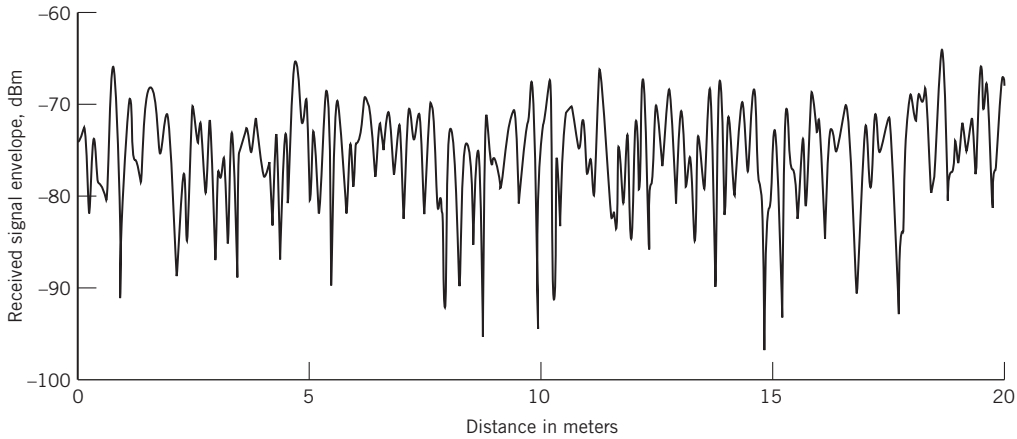


Figure 9.5 Experimental record of received signal envelope in an urban area.

where ψ is the spatial angle subtended between the incoming radio wave and the direction of motion of the receiver. Correspondingly, the change in the phase angle of the received signal at point A′ with respect to that at point A is given by

$$\begin{aligned}\Delta\phi &= \frac{2\pi}{\lambda}\Delta l \\ &= -\frac{2\pi v \Delta t}{\lambda}\cos\psi\end{aligned}$$

where λ is the radio wavelength. The apparent change in frequency, or the *Doppler shift*, is therefore defined by

$$\begin{aligned}\nu &= -\frac{1}{2\pi}\frac{\Delta\phi}{\Delta t} \\ &= \frac{\mathrm{v}}{\lambda}\cos\psi\end{aligned} \tag{9.2}$$

The Doppler shift ν is positive (resulting in an increase in frequency) when the radio waves arrive from ahead of the mobile unit and it is negative when the radio waves arrive from behind the mobile unit.

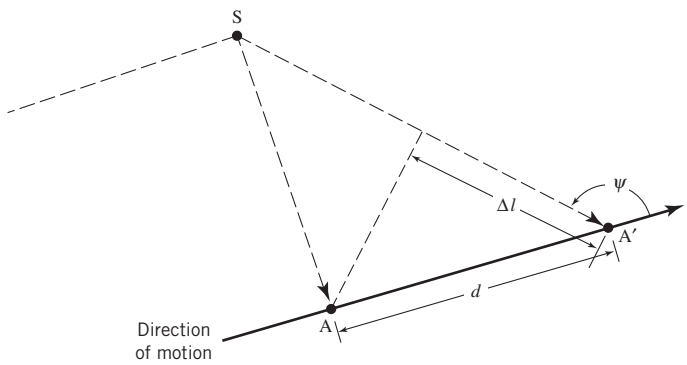


Figure 9.6
Illustrating the calculation of Doppler shift.

9.3 Jakes Model

To illustrate fast fading due to a moving receiver, consider a dynamic multipath environment that involves N *iid fixed scatterers* surrounding such a receiver. Let the transmitted signal be the complex sinusoidal function of unit amplitude and frequency f_c, as shown by

$$s(t) = \exp(j2\pi f_c t)$$

Then, the composite signal observed at the moving receiver, including relative effects of a Doppler shift, is given by

$$x_0(t) = \sum_{n=1}^{N} A_n \exp[j2\pi(f_c + \nu_n)t + j\theta_n]$$

where the amplitude A_n is contributed by the nth scatterer, ν_n is the corresponding Doppler shift, and θ_n is some random phase. The *complex envelope* of the received signal is time varying, as shown by

$$\tilde{x}_0(t) = \sum_{n=1}^{N} A_n \exp[j2\pi \nu_n t + j\theta_n] \tag{9.3}$$

Correspondingly, the autocorrelation function of the complex envelope $\tilde{x}_0(t)$ is defined by

$$R_{\tilde{x}_0}(t) = \mathbb{E}[\tilde{x}_0^*(t)\tilde{x}_0(t+\tau)] \tag{9.4}$$

where $\mathbb{E}$ is the expectation operator with respect to time t and the asterisk in $\tilde{x}_0^*(t)$ denotes complex conjugation. Inserting (9.3) in (9.4) leads to a double summation, one indexed by n and the other indexed by m. Then, simplifying the result under the iid assumption, the autocorrelation function $R_{\tilde{x}_0}(\tau)$ reduces to

$$R_{\tilde{x}_0}(\tau) = \begin{cases} \sum_{n=1}^{N} \mathbb{E}[A_n^2 \exp(j2\pi \nu_n \tau)], & \text{if } m = n \\ 0, & \text{if } m \neq n \end{cases} \tag{9.5}$$

At this point in the discussion, we make two observations:

1. The effects of small changes in distances between the moving receiver and the nth scatterer are small enough for all n for us to write

$$\mathbb{E}[A_n^2 \exp(j2\pi \nu_n \tau)] = \mathbb{E}[A_n^2]\mathbb{E}[\exp(j2\pi \nu_n \tau)] \tag{9.6}$$

   where $n = 1, 2, \ldots, N$.

2. The Doppler shift ν_n is proportional to the cosine of the angle ψ_n subtended between the incoming radio wave from the nth scatterer and the direction of motion of the receiver in Figure 9.6, which follows from (9.2).

We may therefore write

$$\nu_n = \nu_{\max} \cos \psi_n, \qquad n = 1, 2, \ldots, N \tag{9.7}$$

where $\nu_{\max}$ is the *maximum Doppler shift* that occurs when the incoming radio waves propogate in the same direction as the motion of the receiver. Accordingly, using (9.6) and (9.7) in (9.5), we may write

$$R_{\tilde{x}_0}(\tau) = \begin{cases} P_0 \sum_{n=1}^{N} \mathbb{E}[\exp(\mathrm{j}2\pi\nu_{\max}\tau\cos(\psi_n)], & \text{for } m = n \\ 0, & \text{for } m \neq n \end{cases} \tag{9.8}$$

where the multiplying factor

$$P_0 = \sum_{n=1}^{N} A_n^2 \tag{9.9}$$

is the *average signal power* at the receiver input.

We now make two final assumptions:

1. All the radio waves arrive at the receiver from a *horizontal direction* (Clarke, 1968).
2. The multipath is *uniformly distributed* over the range $[-\pi, \pi]$, as shown by the probability density function (Jakes, 1974):

$$f_\Psi(\psi) = \begin{cases} \dfrac{1}{2\pi}, & -\pi \leq \psi \leq \pi \\ 0, & \text{otherwise} \end{cases} \tag{9.10}$$

Under these two assumptions, the remaining expectation in (9.8) becomes independent of n and with it, that equation simplifies further as follows:

$$\begin{aligned} R_{\tilde{x}_0}(\tau) &= P_0\mathbb{E}[\exp(\mathrm{j}2\pi\nu_{\max}\tau\cos\psi)], \qquad -\pi \leq \psi \leq \pi \\ &= P_0\int_{-\pi}^{\pi} f_\Psi(\psi)\exp(\mathrm{j}2\pi\nu_{\max}\tau\cos\psi)\,\mathrm{d}\psi \\ &= P_0\left[\frac{1}{2\pi}\int_{-\pi}^{\pi}\exp(\mathrm{j}2\pi\nu_{\max}\tau\cos\psi)\,\mathrm{d}\psi\right] \end{aligned}$$

The definite integral inside the brackets of this equation is recognized as the *Bessel function of the first kind of order zero*,[2] see Appendix C. By definition, for some argument x, we have

$$J_0(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\exp(\mathrm{j}x\cos\theta)\,\mathrm{d}\theta \tag{9.11}$$

We may therefore express the autocorrelation function of the complex signal $\tilde{x}_0(t)$ at the input of the moving receiver in the compact form

$$R_{\tilde{x}_0}(\tau) = P_0 J_0(2\pi\nu_{\max}\tau) \tag{9.12}$$

The model described by the autocorrelation function of (9.12) is called the *Jakes model*. Figure 9.7a shows a plot of the autocorrelation $R_{\tilde{x}_0}(\tau)$ according to this model.

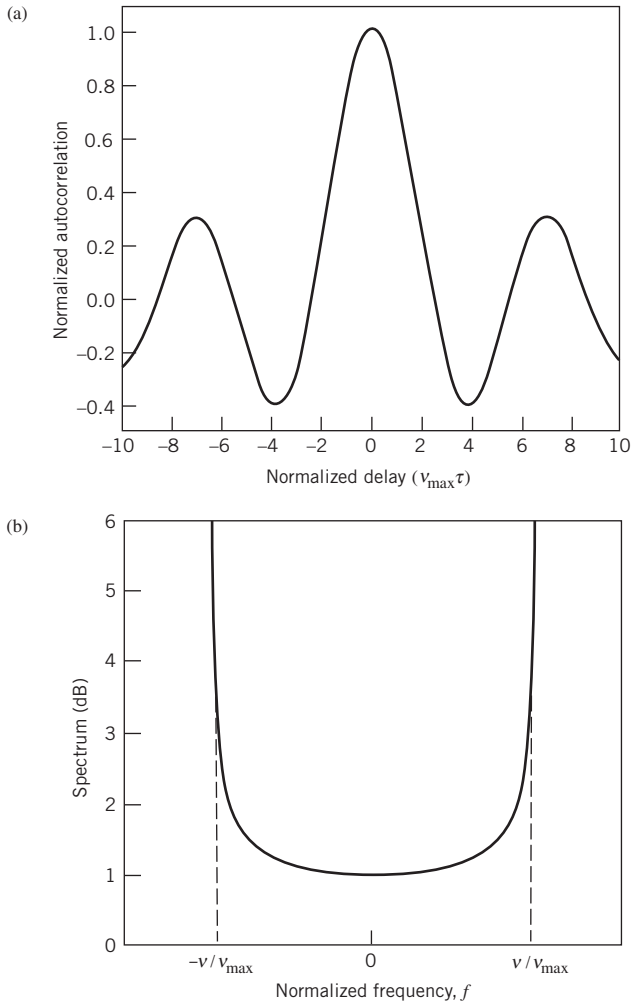


Figure 9.7
(a) Autocorrelation of the complex envelope of the received signal according to the Jakes model. (b) Power spectrum of the fading process for the Jakes model.

According to the Wiener–Khintchine relations for a weakly (wide-sense) stationary process (discussed in Chapter 4), the autocorrelaton function and power spectrum form a Fourier-transform pair. Specifically, we may write

$$S_{\tilde{x}_0}(\nu) = \mathbb{F}[P_0 J_0(2\pi \nu_{\max} \tau)] \tag{9.13}$$

At first sight, it might seem that a closed form solution of this transformation is mathematically intractable; in reality, however, the exact solution is given in (Jakes, 1974):

$$S_{\tilde{x}_0}(\nu) = \begin{cases} \dfrac{P_0}{\sqrt{1-(\nu/\nu_{\max})^2}}, & \text{for } \nu < \nu_{\max} \\ 0, & \text{for } \nu \geq \nu_{\max} \end{cases} \tag{9.14}$$

and with it the model bears his name. Figure 9.7b plots the power spectrum in (9.14) versus the Doppler shift ν for $P_0 = 1$. This idealized graph has the shape of a "bathtub," exhibiting two symmetric integrable singularities at the end points $\nu = \pm\nu_{\max}$.

EXAMPLE 1

Jakes Model Implemented as a FIR Filter

The objective of this example is to compute a FIR (TDL) filter that models the power spectrum of (9.14). To this end, we make use of the following relationships in light of material covered in Chapter 4 on stochastic processes:

1. The autocorrelation function and power spectrum of a weakly stationary process form a Fourier-transform pair, as already mentioned.
2. In terms of stochastic processes, the input–output behavior of a linear system, in the frequency domain, is described by

$$S_Y(f) = |H(f)|^2 S_X(f) \tag{9.15}$$

   where $H(f)$ is the transfer function of the system, $S_X(f)$ is the power spectrum of the input process $X(t)$, and $S_Y(t)$ is the power spectrum of the output process $Y(t)$, both being weakly stationary.
3. If the input process $X(t)$ is Gaussian, then the output process $Y(t)$ is also Gaussian.
4. If the input $X(t)$ is uncorrelated, then the ouput $Y(t)$ will be correlated due to dispersive behavior of the system.

The issue at hand is to find the $H(f)$ required to produce the desired power spectrum of (9.14) using a white noise process of spectral density $N_0/2$ as the input process $X(t)$. Then, given the $S_Y(f)$ and setting the constant $K = N_0/2$, we may solve (9.15) for $H(f)$, obtaining

$$H(f) = \sqrt{\frac{S_Y(f)}{K}} \tag{9.16}$$

In other words, $H(f)$ is proportional to the square root of $S(f)$. (From a practical perspective, the constant K is determined by truncating the power-delay profile, an issue deferred to Section 9.14.)

In light of (9.14) and (9.16), we may now say that the $H(f)$ representing the desired Jakes FIR filter is given by (ignoring the constant K)

$$H(f) = \begin{cases} (1 - f^2)^{-1/4}, & \text{for } -1 \le f \le 1 \\ 0, & \text{otherwise} \end{cases} \tag{9.17}$$

where $f = \nu/\nu_{\max}$. Given this formula, we may then use inverse Fourier transformation to compute the corresponding impulse response of the Jakes FIR filter.

However, before proceeding further, an important aspect of using Jakes model to simulate a fading channel is to pay particular attention to the following point:

> The sampling rate of the input signal applied to the Jakes model and the sampled values of the fading process are highly different.

To be specific, the former is a multiple of the symbol rate and the latter is a multiple of the Doppler bandwidth, $\nu_{\max}$. In other words, the sampling rate is much larger than $\nu_{\max}$. It follows therefore that a *multiple sampling rate with interpolation* must be used in the

simulation; the need for interpolation is to go from a discrete spectrum to its continuous version.

With this point in mind, a 512-point *inverse FFT algorithm* is applied to the transfer function of (9.17) for the following set of specifications:

maximum Doppler shift, $\nu_{max} = 100\,\text{Hz}$

sampling frequency, $f_s = 16\,\nu_{max}$

We thus obtain the discrete-time version of the truncated impulse response h_n of the Jakes FIR filter plotted in Figure 9.8a.

Having computed h_n, we may go on to use the FFT algorithm to compute the corresponding transfer function $H(f)$ of the Jakes FIR filter; the result of this computation is plotted in Figure 9.8b, which has a bathtub-like shape of its own, as expected.

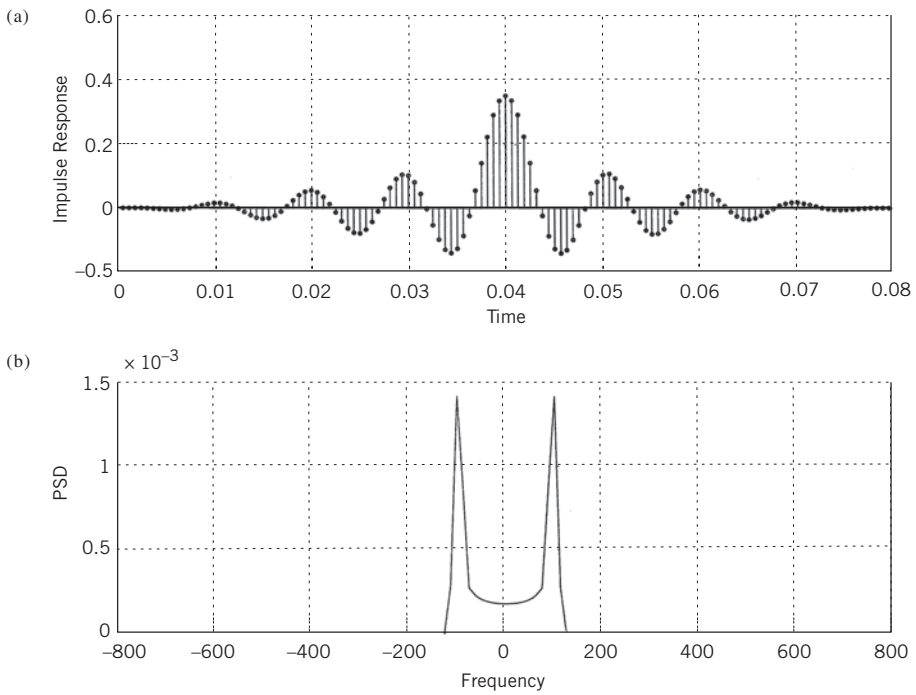


Figure 9.8 Jakes FIR filter. (a) Discrete impulse response. (b) Interpolated power spectral density (PSD).

EXAMPLE 2

Illustrative Generation of Fading Process Using the Jakes FIR Filter

To expand the practical utility of the Jakes FIR filter computed in Example 1 to simulate the fading process, the next thing we do is to pass a complex white noise process through the filter, with the noise having uncorrelated samples. Figure 9.9a displays the power

spectrum of the resulting stochastic process at the filter output. Figure 9.9b shows the envelope of the output process, plotted on a logarithmic scale. This plot is typical of a fading correlated signal.

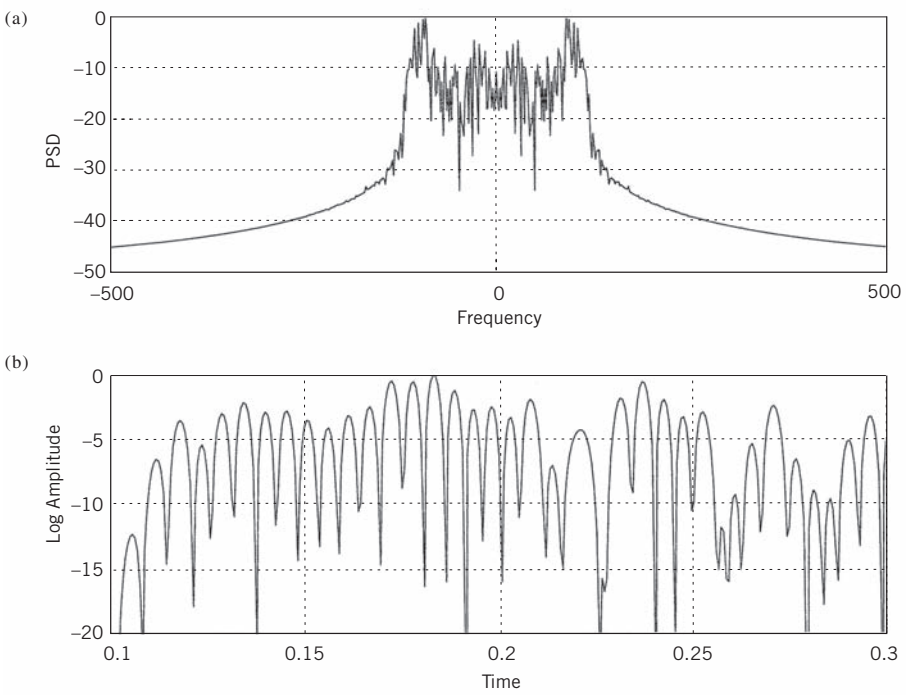


Figure 9.9 Jakes FIR filter driven by white Gaussian noise. (a) Output power spectrum. (b) Envelope of the output process.

9.4 Statistical Characterization of Wideband Wireless Channels

Physical characterization of the multipath environment described in Section 9.3 is appropriate for narrowband mobile radio transmissions where the signal bandwidth is small compared with the reciprocal of the spread in propagation path delays.

However, in real-life situations, we find that the signals radiated in a mobile radio environment occupy a *wide bandwidth*, such that statistical characterization of the wireless channel requires more detailed mathematical considerations, which is the objective of this section. To this end, we follow the complex notations described in Chapter 2 to simplify the analysis.

To be specific, we may express the transmitted band-pass signal as follows:

$$x(t) = \mathrm{Re}[\tilde{x}(t)\exp(\mathrm{j}2\pi f_c t)] \tag{9.18}$$

where $\tilde{x}(t)$ is the complex (low-pass) envelope of $x(t)$ and f_c is the carrier frequency. Since the channel is time varying due to multipath effects, the impulse response of the channel is

delay dependent and, therefore, a time-varying function. Let the impulse response of the channel be expressed as

$$h(\tau;t) = \mathrm{Re}[\tilde{h}(\tau;t)\exp(\mathrm{j}2\pi f_c t)] \tag{9.19}$$

where $\tilde{h}(\tau;t)$ is the complex low-pass impulse response of the channel and τ is a delay variable. The complex low-pass impulse response $\tilde{h}(\tau;t)$ is called the *delay-spread function* of the channel. Correspondingly, the complex low-pass envelope of the channel output, namely $\tilde{y}(t)$, is defined by the convolution integral

$$\tilde{y}(t) = \frac{1}{2}\int_{-\infty}^{\infty} \tilde{h}(\tau;t)\,\tilde{x}(t-\tau)\,\mathrm{d}\tau \tag{9.20}$$

where the scaling factor $1/2$ is the result of using complex notation; see Chapter 2 for details. To be generic, the $\tilde{x}_0(t)$ in Section 9.2 has been changed to $\tilde{x}(t)$.

In general, the behavior of a mobile radio channel can be described only in statistical terms. For analytic purposes and mathematical tractability, the delay-spread function $\tilde{h}(\tau;t)$ is modeled as a *zero-mean complex-valued Gaussian process*. Then, at any time t the envelope $|\tilde{h}(\tau;t)|$ is Rayleigh distributed and the channel is therefore referred to as a *Rayleigh fading channel*. When, however, the mobile radio environment includes *fixed* scatterers, we are no longer justified in using a zero-mean model to describe the delay-spread function $\tilde{h}(\tau;t)$. In such a case, it is more appropriate to use a Rician distribution to describe the envelope $|\tilde{h}(\tau;t)|$ and the channel is referred to as a *Rician fading channel*. The Rayleigh and Rician distributions for a real-valued stochastic process were considered in Chapter 3. In the discussion presented in this chapter we focus largely, but not completely, on a Rayleigh fading channel.

Multipath Correlation Function of the Channel

The *time-varying transfer function* of the channel is defined as the Fourier transform of the delay-spread function $\tilde{h}(\tau;t)$ with respect to the delay variable τ, as shown by

$$\tilde{H}(f;t) = \int_{-\infty}^{\infty} \tilde{h}(\tau;t)\exp(-\mathrm{j}2\pi f\tau)\,\mathrm{d}\tau \tag{9.21}$$

where f denotes the frequency variable. The time-varying transfer function $\tilde{H}(f;t)$ may be viewed as a frequency transmission characteristic of the channel.

For a mathematically tractable statistical characterization of the channel, we make two assumptions motivated by physical considerations; hence the practical importance of the model resulting from these two assumptions.

ASSUMPTION 1 Wide-Sense Stationarity

With interest confined to fast fading in the short term, it is reasonable to assume that the complex impulse response $\tilde{h}(\tau;t)$ *is* wide-sense stationary.

As explained in Chapter 4, a stochastic process is said to be wide-sense (i.e., weakly) stationary if its mean is time independent and its autocorrelation function is dependent only on the difference between two time instants at which the process is observed. In what

follows we use the "wide-sense stationary" terminology because of its common use in the wireless literature.

In the context of the discussion presented herein, this first assumption means that

- The expectation of $\tilde{h}(\tau;t)$ with respect to time t is dependent only on the delay τ.
- Insofar as time t is concerned, the expectation of the product $\tilde{h}^*(\tau_1;t_1) \times h(\tau_2, t_2)$ is dependent only on the time difference $\Delta t = t_2 - t_1$.

Because Fourier transformation is a *linear operation*, it follows that if the complex delay-spread function $\tilde{h}(\tau;t)$ is a zero-mean Gaussian wide-sense stationary process, then the complex time-varying transfer function $\tilde{H}(f;t)$ has similar statistics.

ASSUMPTION 2 Uncorrelated Scattering

The channel is said to be an uncorrelated scattering channel, *when contributions from two or more scatterers with different propagation delays are uncorrelated.*

In other words, the second-order expectation with respect to time t satisfies the requirement

$$\mathbb{E}[\tilde{h}^*(\tau_1;t_1)\tilde{h}(\tau_2;t_2)] = \mathbb{E}[\tilde{h}^*(\tau_1;t_1)\tilde{h}(\tau_1;t_2)]\delta(\tau_1 - \tau_2)$$

where $\delta(\tau_1 - \tau_2)$ is a Dirac-delta function defined in the delay domain. That is, the autocorrelation function of $\tilde{h}(\tau;t)$ is nonzero only when $\tau_2 \neq \tau_1$.

In the literature on statistical characterization of wireless channels, wide-sense stationarity is abbreviated as WSS and uncorrelated scattering is abbreviated as US. Thus, when both Assumptions 1 and 2 are satisfied simultaneously, the resulting channel model is said to be the *WSSUS model*.

Consider then the correlation function[3] of the delay-spread function $\tilde{h}(\tau;t)$. Since $\tilde{h}(\tau;t)$ is complex valued, we use the following definition for the correlation function:

$$R_{\tilde{h}}(\tau_1, t_1;\tau_2, t_2) = \mathbb{E}[\tilde{h}^*(\tau_1;t_1)\tilde{h}(\tau_2;t_2)] \tag{9.22}$$

where $\mathbb{E}$ is the statistical expectation operator, the asterisk denotes complex conjugation, τ_1 and τ_2 are propagation delays of the two paths involved in the calculation, and t_1 and t_2 are the times at which the outputs of the two paths are observed. Under the combined WSSUS channel model, we may reformulate the correlation function in (9.22) as shown by

$$\begin{aligned} R_{\tilde{h}}(\tau_1, \tau_2;\Delta t) &= \mathbb{E}[\tilde{h}^*(\tau_1;t)\tilde{h}(\tau_2;t+\Delta t)] \\ &= r_{\tilde{h}}(\tau_1;\Delta t)\delta(\tau_1 - \tau_2) \end{aligned} \tag{9.23}$$

where Δt is the difference between the observation times t_1 and t_2 and $\delta(\tau_1 - \tau_2)$ is the delta function in the τ-domain. Thus, using τ in place of τ_1 for mathematical convenience, the function in the second line of (9.23) is redefined as

$$r_{\tilde{h}}(\tau;\Delta t) = \mathbb{E}[\tilde{h}^*(\tau;t)\tilde{h}(\tau;t+\Delta t)] \tag{9.24}$$

The function $r_{\tilde{h}}(\tau;\Delta t)$ is called the *multipath correlation profile* of the channel. This new correlation function $r_{\tilde{h}}(\tau;\Delta t)$ provides a statistical measure of the extent to which the signal is *distorted in the time domain* as a result of transmission through the channel.

Spaced-Frequency, Spaced-Time Correlation Function of the Channel

Consider next statistical characterization of the channel in terms of the complex time-varying transfer function $\tilde{H}(f;t)$. Following a formulation similar to that described in (9.22), the correlation function of $\tilde{H}(f;t)$ is defined by

$$R_{\tilde{H}}(f_1, t_1; f_2, t_2) = \mathbb{E}[\tilde{H}^*(f_1;t_1)\tilde{H}(f_2;t_2)] \tag{9.25}$$

where f_1 and f_2 represent two frequencies in the spectrum of the transmitted signal. The correlation function $R_{\tilde{H}}(f_1, t_1; f_2, t_2)$ provides a statistical measure of the extent to which the signal is *distorted in the frequency-domain* by transmission through the channel. From (9.21), (9.22), and (9.25), it is apparent that the correlation functions $R_{\tilde{H}}(f_1, t_1; f_2, t_2)$ and $R_{\tilde{H}}(\tau_1, t_1; \tau_2, t_2)$ form a *two-dimensional Fourier-transform pair*, defined as follows:

$$R_{\tilde{H}}(f_1, t_1; f_2, t_2) \rightleftharpoons \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} R_{\tilde{h}}(\tau_1, t_1; \tau_2, t_2) \exp[-j2\pi(f_1\tau_1 - f_2\tau_2)]\, d\tau_1\, d\tau_2 \tag{9.26}$$

Invoking wide-sense stationarity in the time domain, we may reformulate (9.25) as

$$R_{\tilde{H}}(f_1, f_2; \Delta t) = \mathbb{E}[\tilde{H}^*(f_1;t)\tilde{H}(f_2;t+\Delta t)] \tag{9.27}$$

Equation (9.27) suggests that the correlation function $R_{\tilde{H}}(f_1, f_2; \Delta t)$ may be measured by using pairs of spaced tones to carry out cross-correlation measurements on the resulting channel outputs. Such a measurement presumes stationarity in the time domain. If we also assume stationarity in the frequency domain, we may go one step further and write

$$\begin{aligned} R_{\tilde{H}}(f, f+\Delta f; \Delta t) &= r_{\tilde{H}}(\Delta f; \Delta t) \\ &= \mathbb{E}[\tilde{H}^*(f;t)\tilde{H}(f+\Delta f; t+\Delta t)] \end{aligned} \tag{9.28}$$

The new correlation function $r_{\tilde{H}}(\Delta f; \Delta t)$, introduced in the first line of (9.28), is in fact the Fourier transform of the multipath correlation profile $r_{\tilde{h}}(\tau;\Delta t)$ with respect to the delay-time variable τ, as shown by

$$r_{\tilde{H}}(\Delta f; \Delta t) = \int_{-\infty}^{\infty} r_{\tilde{h}}(\tau;\Delta t) \exp(-j2\pi\tau\Delta f)\, d\tau \tag{9.29}$$

The new function $r_{\tilde{H}}(\Delta f; \Delta t)$ is called the *spaced-frequency, spaced-time correlation function* of the channel, where the double use of "spaced" accounts for Δt and Δf.

Scattering Function of the Channel

Finally, we introduce another new function denoted by $S(\tau; \nu)$ that forms a Fourier-transform pair with the multipath correlation profile $r_{\tilde{h}}(\tau;\Delta t)$ with respect to the variable Δt; that is, by definition, we have

$$S(\tau; \nu) = \int_{-\infty}^{\infty} r_{\tilde{h}}(\tau;\Delta t) \exp(-j2\pi\nu\Delta t)\, d(\Delta t) \tag{9.30}$$

for the Fourier transform and

$$r_{\tilde{h}}(\tau;\Delta t) = \int_{-\infty}^{\infty} S(\tau;\nu) \exp(j2\pi\nu\Delta t)\, d\nu \tag{9.31}$$

for the inverse Fourier transform.

The function $S(\tau;\nu)$ may also be defined in terms of $r_{\tilde{H}}(\Delta f;\Delta t)$ by applying a form of *double Fourier transformation*:

> A Fourier transform with respect to the time variable Δt and an inverse Fourier transform with respect to the frequency variable Δf.

That is to say,

$$S(\tau;\nu) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} r_{\tilde{H}}(\Delta f;\Delta t)\exp(-j2\pi\nu\Delta t)\exp(j2\pi\tau\Delta f)\,d(\Delta t)\,d(\Delta f) \tag{9.32}$$

Figure 9.10 displays the functional relationships between the three important functions: $r_{\tilde{h}}(\tau;\Delta t)$, $r_{\tilde{h}}(\Delta f;\Delta t)$, and $S(\tau;\nu)$ in terms of the Fourier transform and its inverse.

The function $S(\tau;\nu)$ is called the *scattering function* of the channel. For a physical interpretation of it, consider the transmission of a single tone of frequency f' relative to the carrier. The complex envelope of the resulting filter output is

$$\tilde{y}(t) = \exp(j2\pi f' t)\tilde{H}(f';t) \tag{9.33}$$

The correlation function of $\tilde{y}(t)$ is given by

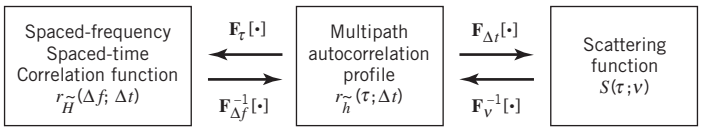
$$\begin{aligned}\mathbb{E}[\tilde{y}^*(t)\tilde{y}(t+\Delta t)] &= \exp(j2\pi f'\Delta t)\mathbb{E}[\tilde{H}^*(f';t)\tilde{H}^*(f';t+\Delta t)]\\ &= \exp(j2\pi f'\Delta t)r_{\tilde{H}}(0;\Delta t)\end{aligned} \tag{9.34}$$

where, in the last line, we made use of (9.28). Putting $\Delta f = 0$ in (9.29) and then using (9.31), we may write

$$\begin{aligned}r_{\tilde{H}}(0;\Delta t) &= \int_{-\infty}^{\infty} r_{\tilde{h}}(\tau;\Delta t)\,d\tau\\ &= \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} S(\tau;\nu)\,d\tau\right]\exp(j2\pi\nu\Delta t)\,d\nu\end{aligned} \tag{9.35}$$

Hence, we may view the integral inside the square brackets in (9.35), namely

$$\int_{-\infty}^{\infty} S(\tau;\nu)\,d\tau$$



$\mathbf{F}_{\tau}[\cdot]$: Fourier transform with respect to delay τ

$\mathbf{F}_{\Delta f}^{-1}[\cdot]$: Inverse Fourier transform with respect to frequency increment Δf

$\mathbf{F}_{\Delta t}[\cdot]$: Fourier transform with respect to time increment Δt

$\mathbf{F}_{\nu}^{-1}[\cdot]$: Inverse Fourier transform with respect to Doppler shift ν

Figure 9.10 Functional relationships between the multipath correlation profile $r_{\tilde{h}}(\tau;\Delta t)$, the spaced-frequency spaced-time correlation function $r_{\tilde{H}}(\Delta f;\Delta t)$, and the scattering function $S(\tau;\nu)$.

as the *power spectral density* of the channel output relative to the frequency f' of the transmitted tone with the Doppler shift ν acting as the frequency variable. Generalizing this result, we may now make the statement:

> The scattering function $S(\tau; \nu)$ provides a statistical measure of the output power of the channel, expressed as a function of the time delay τ and the Doppler shift ν.

Power-Delay Profile

We continue statistical characterization of the wireless channel by putting $\Delta t = 0$ in (9.24) to obtain

$$\begin{aligned} P_{\tilde{h}}(\tau) &= r_{\tilde{h}}(\tau; 0) \\ &= \mathbb{E}[|\tilde{h}(\tau; t)|^2] \end{aligned} \tag{9.36}$$

The function $P_{\tilde{h}}(\tau)$ describes the intensity (averaged over the fading fluctuations) of the scattering process at propagation delay τ for the WSSUS channel. Accordingly, $P_{\tilde{h}}(\tau)$ is called the *power-delay profile* of the channel. In any event, this profile provides an estimate of the average multipath power expressed as a function of the delay variable τ.

The power-delay profile may also be defined in terms of the scattering function $S(\tau; \nu)$ by averaging it over all potentially possible Doppler shifts. Specifically, setting $\Delta t = 0$ in (9.31) and then using the first line of (9.36), we obtain

$$P_{\tilde{h}}(\tau) = \int_{-\infty}^{\infty} S(\tau; \nu)\, d\nu \tag{9.37}$$

Figure 9.11 shows an example of the power-delay profile that depicts a typical plot of the power spectral density versus excess delay;[4] the excess delay is measured with respect to the time delay for the shortest echo path. The "threshold level" K included in Figure 9.11 defines the power level below which the receiver fails to operate satisfactorily.

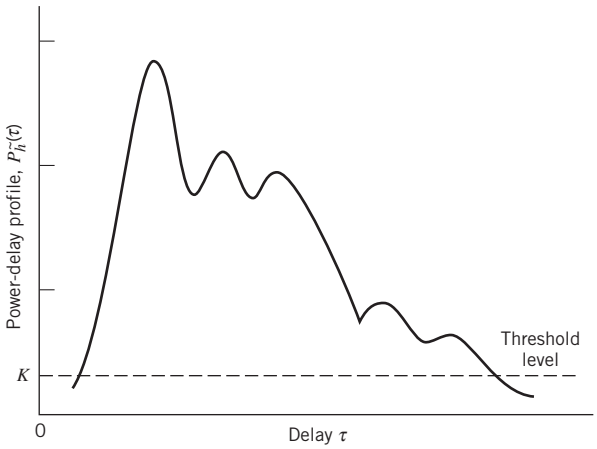


Figure 9.11 Example of a power-delay profile for a mobile radio channel.

Central Moments of $P_{\tilde{h}}(\tau)$

To characterize the power-delay profile of a WSSUS channel in statistical terms, we begin with the moment of order zero; that is, the integrated power averaged over the delay variable τ, as shown by

$$P_{\text{av}} = \int_{-\infty}^{\infty} P_{\tilde{h}}(\tau)\, \mathrm{d}\tau \tag{9.38}$$

The *average delay*, normalized with respect to P_{av}, is defined in terms of the first-order moment by the formula

$$\tau_{\text{av}} = \frac{1}{P_{\text{av}}} \int_{-\infty}^{\infty} \tau P_{\tilde{h}}(\tau)\, \mathrm{d}\tau \tag{9.39}$$

Correspondingly, the second-order central moment, normalized with respect to P_{av}, is defined by the root-mean-square (RMS) formula

$$\sigma_\tau = \left[\frac{1}{P_{\text{av}}} \int_{-\infty}^{\infty} (\tau - \tau_{\text{av}})^2 P_{\tilde{h}}(\tau)\, \mathrm{d}\tau \right]^{1/2} \tag{9.40}$$

The new parameter σ_τ is called the *delay spread*, which has acquired a special stature among the parameters used to characterize the WSSUS channel.

From Chapter 2 on the representation of signals in a linear environment, we recall that the duration of a signal in the time domain is inversely related to the bandwidth of the signal in the frequency domain. Building on this time–frequency relationship, we may define the *coherence bandwidth* $B_{\text{coherence}}$ of a WSSUS channel as follows:

$$B_{\text{coherence}} = \frac{1}{\tau_{\text{av}}} \tag{9.41}$$

In words:

> The coherence bandwidth of the WSSUS channel is that band of frequencies for which the frequency response of the channel is strongly correlated.

This statement is intuitively satisfying.

Doppler Power Spectrum

Consider next the issue of relating Doppler effects to time variations of the channel. In direct contrast to the power-delay profile, this time we set $\Delta f = 0$, which corresponds to the transmission of a single tone (of some appropriate frequency) over the channel. Under this condition, the spaced-frequency, spaced-time correlation function of the channel, described in (9.29), reduces to $r_{\tilde{H}}(0;\Delta t)$. Hence, evaluating the Fourier transform of this function with respect to the time variable Δt, we may write

$$S_{\tilde{H}}(\nu) = \int_{-\infty}^{\infty} r_{\tilde{H}}(0;\Delta t) \exp(-\mathrm{j}2\pi\nu\Delta t)\, \mathrm{d}(\Delta t) \tag{9.42}$$

The function $S_{\tilde{H}}(\nu)$ defines the power spectrum of the channel output expressed as a function of the Doppler shift ν; it is therefore called the *Doppler power spectrum* of the channel.

The Doppler-power spectrum of (9.42) may be interpreted in two insightful ways (Molisch, 2011):

1. The Doppler spectrum describes the frequency dispersion of a wireless channel, which results in the occurrence of transmission errors in narrowband mobile wireless communication systems.
2. The Doppler spectrum provides a measure of temporal variability of the channel, which, in mathematical terms, is described by the channel's correlation function $r_{\tilde{H}}(0;\Delta t)$ for $\Delta f = 0$.

As such, we may view the Doppler-power spectrum as another important statistical characterization of WSSUS channels.

The Doppler power spectrum may also be defined in terms of the scattering function by averaging it over all possible propagation delays, as shown by

$$S_{\tilde{H}}(\nu) = \int_{-\infty}^{\infty} S(\tau;\nu)\,\mathrm{d}\tau \tag{9.43}$$

Typically, the Doppler shift ν assumes positive and negative values with almost equal likelihood. The mean Doppler shift is therefore effectively zero. The square root of the second moment of the Doppler spectrum is thus defined by

$$\sigma_\nu = \left(\frac{\displaystyle\int_{-\infty}^{\infty} \nu^2 S_{\tilde{H}}(\nu)\,\mathrm{d}\nu}{\displaystyle\int_{-\infty}^{\infty} S_{\tilde{H}}(\nu)\,\mathrm{d}\nu} \right)^{1/2} \tag{9.44}$$

The parameter σ_ν provides a measure of the width of the Doppler spectrum; therefore, it is called the *Doppler spread* of the channel.

Another useful parameter that is often used in radio propagation measurements is the *fade rate* of the channel. For a Rayleigh fading channel, the *average fade rate* is related to the Doppler spread σ_ν by the empirical rule:

$$f_{\text{fade rate}} = 1.475\,\sigma_\nu \text{ crossings per second} \tag{9.45}$$

As the name implies, the fade rate provides a measure of the rapidity of the channel fading phenomenon.

Some typical values encountered in a mobile radio environment are as follows:

- the delay spread σ_τ amounts to about 20 μs;
- the Doppler spread σ_ν due to the motion of a vehicle may typically occupy the range 40–100 Hz, but sometimes may well exceed 100 Hz.

One other parameter directly related to the Doppler spread is the *coherence time* of the channel. Here again, as with coherence bandwidth discussed previously, we may invoke the inverse time–frequency relationship to say that the coherence time of a multipath wireless channel is inversely proportional to the Doppler spread, as shown by

$$\begin{aligned} \tau_{\text{coherence}} &= \frac{1}{\sigma_\nu} \\ &\approx \frac{0.3}{2\,\nu_{\max}} \end{aligned} \tag{9.46}$$

where ν_{max} is the maximum Doppler shift due to motion of the mobile unit. In words:

> The coherence time of the channel is that duration for which the time response of the channel is strongly correlated.

Here again, this statement is intuitively satisfying.

Classification of Multipath Channels

The particular form of fading experienced by a multipath channel depends on whether the channel characterization is viewed in the frequency domain or the time domain:

1. When the channel is viewed in the frequency domain, the parameter of concern is the channel's *coherence bandwidth* $B_{coherence}$, which is a measure of the transmission bandwidth for which signal distortion across the channel becomes noticeable. A multipath channel is said to be *frequency selective* if the coherence bandwidth of the channel is small compared with the bandwidth of the transmitted signal. In such a situation, the channel has a filtering effect, in that two sinusoidal components with a frequency separation greater than the channel's coherence bandwidth are treated differently. If, however, the coherence bandwidth of the channel is large compared with the transmitted signal bandwidth, the fading is said to be *frequency nonselective*, or *frequency flat*.
2. When the channel is viewed in the time domain, the parameter of concern is the *coherence time* $\tau_{coherence}$, which provides a measure of the transmitted signal duration for which distortion across the channel becomes noticeable. The fading is said to be *time selective* if the coherence time of the channel is small compared with the duration of the received signal (i.e., the time for which the signal is in flight). For

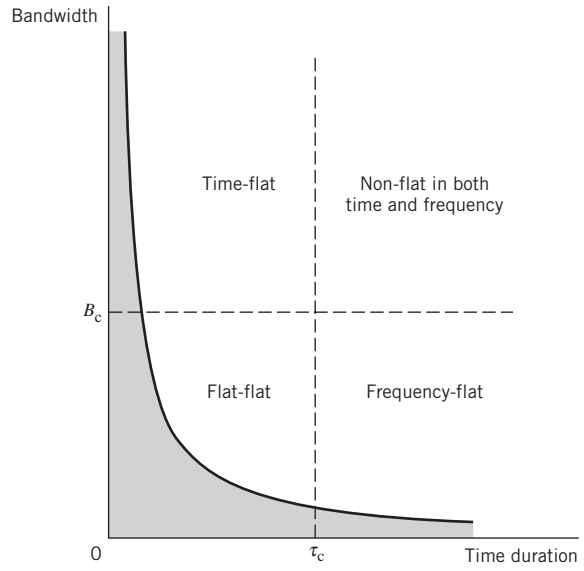


Figure 9.12
Illustrating the four classes of multipath channels: τ_c = coherence time, B_c = coherence bandwidth.

digital transmission, the received signal's duration is taken as the symbol duration plus the channel's delay spread. If, however, the channel's coherence time is large compared with the received signal duration, then the fading is said to be *time nonselective*, or *time flat*, in the sense that the channel appears to the transmitted signal as time invariant.

In light of this discussion, we may classify multipath channels as follows:

- *Flat-flat channel*, which is flat in both frequency and time.
- *Frequency-flat channel*, which is flat in frequency only.
- *Time-flat channel*, which is flat in time only.
- *Completely non-flat channel*, which is flat neither in frequency nor in time; such a channel is also referred to as a *doubly spread channel*.

The classification of multipath channels, based on this approach, is shown in Figure 9.12. The forbidden area, shown shaded in this figure, follows from the inverse relationship that exists between bandwidth and time duration.

9.5 FIR Modeling of Doubly Spread Channels

In Section 9.4, statistical analysis of the doubly spread channel was carried out by focusing on two complex low-pass entities, namely the impulse response $\tilde{h}(\tau;t)$ and the corresponding transfer function $\tilde{H}(f;t)$. Therein, mathematical simplification was accomplished by disposing of the midband frequency f_c of the actual band-pass character of the doubly spread channel. Despite this simplification, the analytic approach used in Section 9.4 is highly demanding in mathematical terms. In this section, we will take an "approximate" approach based on the use of a FIR filter to model the doubly spread channel.[5] From an engineering perspective. this new approach has a great deal of practical merit.

To begin, we use the *convolution integral* to describe the input–output relationship of the system, as shown in (9.20), reproduced here for convenience of presentation

$$\tilde{y}(t) = \frac{1}{2}\int_{-\infty}^{\infty} \tilde{h}(\tau;t)\,\tilde{x}(t-\tau)\,\mathrm{d}\tau \tag{9.47}$$

where $\tilde{x}(t)$ is the complex low-pass input signal applied to the channel and $\tilde{y}(t)$ is the resulting complex low-pass output signal. Although this integral can be formulated in another equivalent way, the choice made in (9.47) befits modeling of a time-varying FIR system, as we will see momentarily. Speaking of the input signal $\tilde{x}(t)$, we assume that its Fourier transform satisfies the condition

$$\tilde{X}(f) = 0 \qquad \text{for } f > W \tag{9.48}$$

where $2W$ denotes the original input band-pass signal's bandwidth centered around the midband frequency f_c. With FIR filtering in mind, it is logical to expand the delayed input signal $\tilde{x}(t-\tau)$ using the *sampling theorem*, discussed in Chapter 6. Specifically, we write

$$\tilde{x}(t-\tau) = \sum_{n=-\infty}^{\infty} \tilde{x}(t-nT)\,\mathrm{sinc}\left(\frac{\tau}{T_s} - n\right) \tag{9.49}$$

where T_s is the sampling period of the FIR filter chosen in accordance with the sampling theorem as follows:

$$\frac{1}{T_s} > 2W \tag{9.50}$$

The sinc function in (9.49) is defined by

$$\text{sinc}\left(\frac{\tau}{T_s} - n\right) = \frac{\sin\left[\pi\left(\frac{\tau}{T_s} - n\right)\right]}{\pi\left(\frac{\tau}{T_s} - n\right)} \tag{9.51}$$

From the standpoint of the sampling theorem we could set $1/T_s = 2W$, but the choice made in (9.50) gives us more practical flexibility.

In (9.49) it is important to note that we have done the following:

- Dependence on the coordinate functions under the summation has been put on the delay variable τ in the sinc function.
- Dependence on the time-varying FIR coefficients has been put on time t.

This separation of variables is the key to the FIR modeling of a linear time-varying system. Note also that the sinc functions under the summation in (9.49) are orthogonal but not normalized.

Thus, substituting (9.49) into (9.47) and interchanging the order of integration and summation, which is permitted as we are dealing with a linear system, we get

$$\tilde{y}(t) = \sum_{n=-\infty}^{\infty} \tilde{x}\left(\frac{t}{T_s} - n\right)\left[\int_{-\infty}^{\infty} \tilde{h}(\tau; t)\,\text{sinc}\left(\frac{t}{T_s} - n\right) d\tau\right] \tag{9.52}$$

To simplify matters, we now introduce the *complex tap-coefficients*[6] $\tilde{c}_n(t)$, defined in terms of the complex impulse response as follows:

$$\tilde{c}_n(t) = \frac{1}{2}\int_{-\infty}^{\infty} \tilde{h}(\tau; t)\,\text{sinc}\left(\frac{t}{T_s} - n\right) d\tau \tag{9.53}$$

Accordingly, we may rewrite (9.52) in the much simplified summation form:

$$\tilde{y}(t) = \sum_{n=-\infty}^{\infty} \tilde{x}\left(\frac{t}{T_s} - n\right)\tilde{c}_n(t) \tag{9.54}$$

Examining (9.54) for insight, we may make our first observation:

> The uniformly sampled functions $\tilde{x}[(t/T) - n]$ are generated as tap-inputs by passing the complex low-pass input signal $\tilde{x}(t)$ through a TDL filter whose taps are spaced T seconds apart.

Turning next to (9.53) for insight, refer to Figure 9.13, where this equation is sketched for three different settings of the function $\text{sinc}[(t/T_s) - n]$; the area shaded in the figure refers to the complex impulse response $\tilde{h}(\tau; t)$ that is assumed to be causal and occupying a finite duration. In light of the three different sketches shown in Figure 9.13, we may make our second observation.

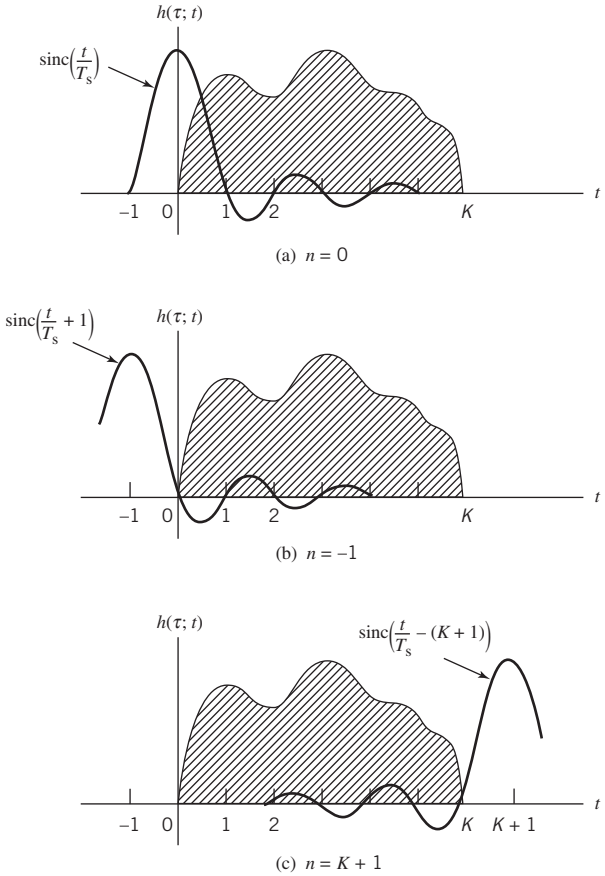


Figure 9.13 Illustrating the way in which location of the sinc weighting function shows up for varying n.

Assuming that the integral in (9.53) is dominated by the mainlobe of the sinc function, the complex time-varying tap-coefficient $\tilde{c}_n(t)$ is essentially zero for negative values of discrete time n and all positive values of n greater than τ/T.

In accordance with these two observations, we may approximate (9.54) as follows:

$$\tilde{y}(t) \approx \sum_{n=0}^{K} \tilde{x}\left(\frac{t}{T_s} - n\right)\tilde{c}_n(t) \tag{9.55}$$

where K is the number of taps.

Equation (9.55) defines a *complex FIR model* for the representation of a complex low-pass time-varying system characterized by the complex impulse response $\tilde{h}(\tau;T)$. Figure 9.14 depicts a block diagram representation of this model, based on (9.55).

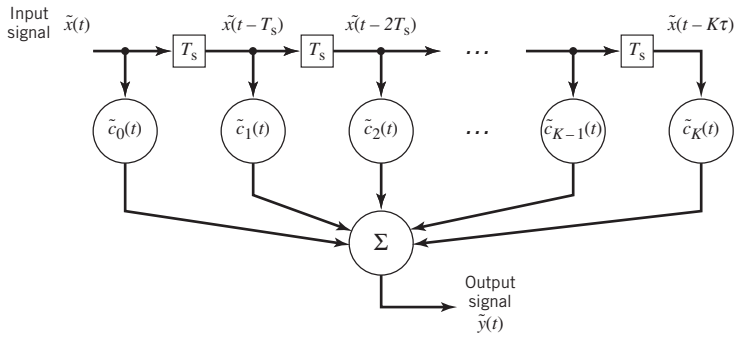


Figure 9.14 Complex FIR model of a complex low-pass time-varying channel.

Some Practical Matters

To model the doubly spread channel by means of a FIR filter in accordance with (9.55), we need to know the sampling rate $1/T_s$ and the number of taps K in this equation. To satisfy these two practical requirements, we offer the following empirical points:

1. The sampling rate of the FIR filter, $1/T_s$, is much higher than the maximum Doppler bandwidth of the channel, ν_{max}; typically, we find that $1/T_s$ is eight to sixteen times ν_{max}. Hence, knowing ν_{max}, we may determine a desirable value of the sampling rate $1/T_s$.
2. The number of taps K in (9.55) may be determined by truncating the power-delay profile $P_{\tilde{h}}(f)$ of the channel. Specifically, given a measurement of this profile, a suitable value of K is determined by choosing a threshold level below which the receiver fails to operate satisfactorily, as illustrated in Figure 9.11.

Generation of the Tap-Coefficients

To generate the tap-coefficients $\tilde{c}_n(t)$, we may use the scheme shown in Figure 9.15 that involves the following (Jeruchim *et al.*, 2000):

1. A complex white Gaussian process of zero mean and unit variance is used as the input.
2. A complex low-pass filter of transfer function $\tilde{H}(f)$ is chosen in such a way that it produces the desired Doppler power spectrum $S_{\tilde{H}}(f)$ where we have used f in place of the Doppler shift ν for convenience of presentation. In other words, we may set

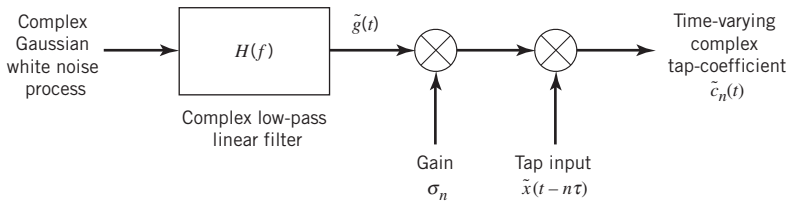


Figure 9.15 Scheme for generating the nth complex weighting coefficients $\tilde{c}_n(t)$ in the FIR model of Figure 9.14.

$$\begin{aligned} S_{\tilde{c}}(f) &= S_{\tilde{H}}(f) \\ &= S_{\tilde{w}}(f)\left|\tilde{H}(f)\right|^2 \\ &= \left|\tilde{H}(f)\right|^2 \end{aligned} \tag{9.56}$$

where, in the second line, $S_{\tilde{w}}(f)$ denotes the power spectral density of the white noise process, which is equal to unity by assumption.

3. The filter is designed in such a way that its output $\tilde{g}(t)$ has a normalized power of unity.
4. The static gain, denoted by σ_n, accounts for different variances of the different tap-coefficients.

EXAMPLE 3

Rayleigh Processes

For complex FIR modeling of a time-varying Rayleigh fading channel, we may use zero-mean complex Gaussian processes to represent the time-varying tap-coefficients $\tilde{c}_n(t)$, which, in turn, means that the complex impulse response of the channel $\tilde{h}(\tau;t)$ is also a zero-mean Gaussian process in the variable t.

Moreover, under the assumption of a *WSSUS channel*, the tap-coefficients $\tilde{c}_n(t)$ for varying n will be uncorrelated. The power spectral density of each tap-coefficient is specified by the Doppler spectrum. In particular, the variance σ_n^2 of the nth weight function is approximately given by

$$\mathbb{E}[\left|\tilde{c}_n(t)\right|^2] \approx T_s^2 p(n\tau) \tag{9.57}$$

where T_s is the sampling period of the FIR and $p(n\tau)$ is a discrete version of the *power-delay profile*, $P_{\tilde{h}}(\tau)$.

EXAMPLE 4

Rician–Jakes Doppler Spectrum Model

The Jakes model, discussed in Example 1, is well suited for describing the Doppler spectrum for a dense-scattering environment, exemplified by an urban area. However, in a *rural environment*, there is a high likelihood for the presence of one strong "direct line-of-sight" path, for which the FIR-based Rician model is an appropriate candidate. In such an environment, we may use the *Rician–Jakes Doppler spectrum* that has the following form (Tranter *et al.*, 2004):

$$\tilde{S}_{\tilde{c}}(f) = \frac{0.41}{\sqrt{1-(f/\nu_{\max})^2}} + 0.91\,\delta(f \pm 0.7\,\nu_{\max}) \tag{9.58}$$

where $\nu_{\max}$ is the maximum magnitude of the Doppler shift. This partially empirical formula, plotted in Figure 9.16, consists of two components: the FIR Jakes filter of Example 1, and two delta functions at $\pm 0.7\,\nu_{\max}$ representing a direct-line-of sight signal received.

Typically, the sequence defined by $p(nT_s)$ decreases with n in an approximate exponential manner, eventually reaching a neglibly small value at some time $T_{\max}$. This exponential approximation of the power-delay profile has been validated experimentally by many measurements; see Note 4. In any event, the number of taps in the FIR filter, K, is

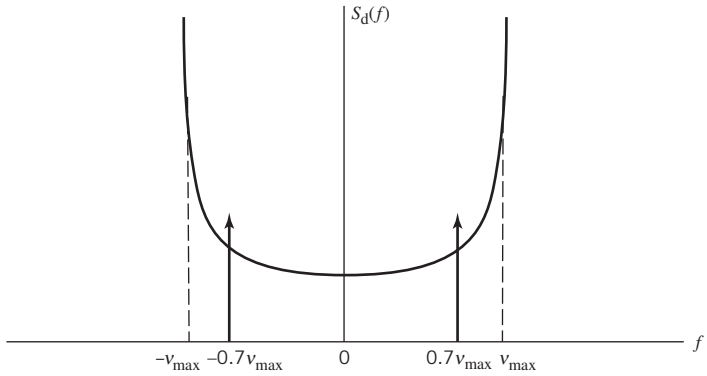


Figure 9.16
Illustrating the Rician–Jakes Doppler spectrum of (9.58).

approximately defined by the ratio T_{max}/T_s. The point made here on the number of taps K substantiates what has been made previously on Jakes model in Example 1 and in point 2 under Some Practical Matters in this section.

9.6 Comparison of Modulation Schemes: Effects of Flat Fading

We bring this first part of the chapter to an end by presenting the effects of flat fading on the behavior of different modulation schemes for wireless communications.

In Chapter 7 we studied the subject of signaling over AWGN channels using different modulation schemes and evaluated their performance under two different receiver conditions: coherence and noncoherence. For the purpose of comparison, we have reproduced the BER for a selected number of those modulation schemes in AWGN in Table 9.1.

Table 9.1 Formulas for the BER of coherent and noncoherent digital receivers

Signaling scheme	BER	
	AWGN channel	Flat Rayleigh fading channel
(a) Binary PSK, QPSK, MSK using coherent detection	$Q\left(\sqrt{\frac{2E_b}{N_0}}\right)$	$\frac{1}{2}\left(1-\sqrt{\frac{\gamma_0}{1+\gamma_0}}\right)$
(b) Binary FSK using coherent detection	$Q\left(\sqrt{\frac{E_b}{N_0}}\right)$	$\frac{1}{2}\left(1-\sqrt{\frac{\gamma_0}{2+\gamma_0}}\right)$
(c) Binary DPSK	$\frac{1}{2}\exp\left(-\frac{E_b}{N_0}\right)$	$\frac{1}{2(1+\gamma_0)}$
(d) Binary FSK using noncoherent detection	$\exp\left(-\sqrt{\frac{E_b}{2N_0}}\right)$	$\frac{1}{2+\gamma_0}$

E_b: transmitted energy per bit; $N_0/2$: power spectral density of channel noise;
γ_0: mean value of the received energy per bit-to-noise spectral density ratio.

Table 9.1 also includes the exact formulas for the BER for a flat Rayleigh fading channel, where the parameter

$$\gamma_0 = \frac{E_b}{N_0}\mathbb{E}[\alpha^2] \tag{9.59}$$

is the *mean value of the received signal energy per bit-to-noise spectral density ratio*. In (9.59), the expectation $\mathbb{E}[\alpha^2]$ is the mean value of the Rayleigh-distributed random variable α characterizing the channel. The derivations of the fading-channel formulas listed in the last column of Table 9.1 are addressed in Problems 9.1 and 9.2.

Comparing the formulas for a flat Rayleigh fading channel with the formulas for their AWGN (i.e., nonfading) channel counterparts, we find that the Rayleigh fading process results in a severe degradation in the noise performance of a wireless communication receiver with the degradation measured in terms of decibels of additional mean SNR spectral density ratio. In particular, the asymptotic decrease in the BER with γ_0 follows an *inverse law*. This form of asymptotic behavior is dramatically different from the case of a nonfading channel, for which the asymptotic decrease in the BER with γ_0 follows an *exponential law*.

In graphical terms, Figure 9.17 plots the formulas under part a of Table 9.1 compared with the BERs of binary PSK over the AWGN and Rayleigh fading channels. The figure also includes corresponding plots for the Rician fading channel with different values of the *Rice factor K*, discussed in Chapter 4. We see that as K increases from zero to infinity, the behavior of the receiver varies all the way from the Rayleigh channel to the AWGN

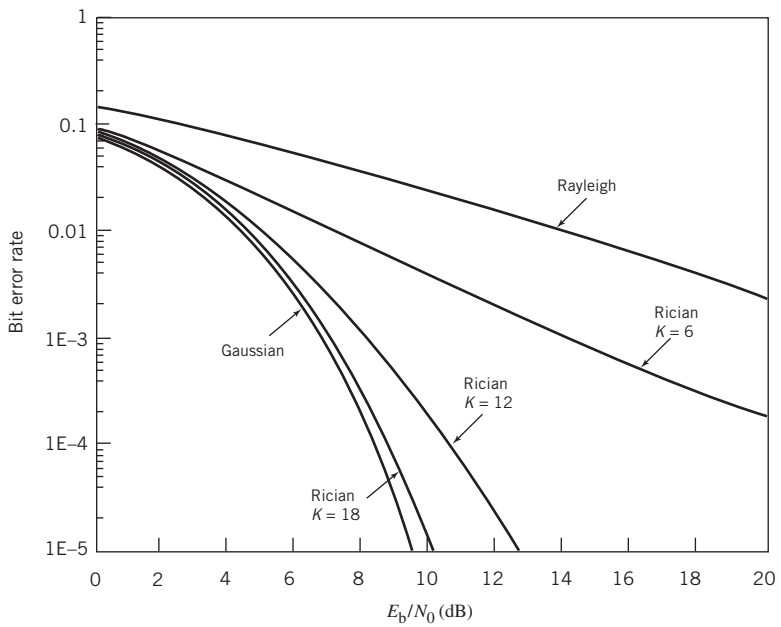


Figure 9.17 Comparison of performance of coherently detected binary PSK over different fading channels.

channel. The results plotted in Figure 9.17 for the Rician channel were obtained using simulations (Haykin and Moher, 2005). From Figure 9.17 we see that, as matters stand, we have a serious problem caused by channel fading. For example, at an SNR of 20 dB and the presence of Rayleigh fading, the use of binary PSK results in a BER of about 3×10^{-2}, which is not good enough for the transmission of speech or digital data over the wireless channel.

9.7 Diversity Techniques

Up to now, we have emphasized the multipath fading phenomenon as an inherent characteristic of a wireless channel, which indeed it is. Given this physical reality, how, then, do we make the communication process across the wireless channel into a *reliable* operation? The answer to this fundamental question lies in the use of *diversity*, which may be viewed as a form of *redundancy* in a spatial context. In particular, if several replicas of the information-bearing signal can be transmitted simultaneously over independently fading channels, then there is a good likelihood that at least one of the received signals will not be severely degraded by channel fading. There are several methods for making such a provision. In the context of the material covered in this book, we identify three approaches to diversity:

1. *Frequency diversity*, in which the information-bearing signal is transmitted using several carriers that are spaced sufficiently apart from each other to provide independently fading versions of the signal. This may be accomplished by choosing a frequency spacing equal to or larger than the coherence bandwidth of the channel.
2. *Time diversity*, in which the same information-bearing signal is transmitted in different time slots, with the interval between successive time slots being equal to or greater than the coherence time of the channel. We can still get some diversity if the interval is less than the coherence time of the channel, but at the expense of degraded performance. In any event, time diversity may be likened to the use of a repetition code for error-control coding.
3. *Space diversity*, in which multiple transmit or receive antennas, or both, are used with the spacing between adjacent antennas being chosen so as to ensure the independence of possible fading events occurring in the channel.

Among these three kinds of diversity, space diversity is the subject of interest in the second part of this chapter. Depending on which end of the wireless link is equipped with multiple antennas, we may identify three different forms of space diversity:

1. *Receive diversity,* which involves the use of a single transmit antenna and multiple receive antennas.
2. *Transmit diversity*, which involves the use of multiple transmit antennas and a single receive antenna.
3. *Diversity on both transmit and receive*, which combines the use of multiple antennas at both the transmitter and receiver.

Receive diversity is the oldest one of the three, with the other two being of more recent origin. In what follows, we will study these three different forms of diversity in this order.

9.8 "Space Diversity-on-Receive" Systems

In "space diversity on receive," multiple receiving antennas are used with the spacing between adjacent antennas being chosen so that their respective outputs are essentially independent of each other. This requirement may be satisfied by spacing the adjacent receiving antennas by as much as 10 to 20 radio wavelengths or less apart from each other. Typically, an elemental spacing of several radio wavelengths is deemed to be adequate for space diversity on receive. The much larger spacing is needed for elevated base stations, for which the angle spread of the incoming radio waves is small; note that the spatial coherence distance is inversely proportional to the angle spread. Through the use of diversity on receive as described here, we create a corresponding set of fading channels that are essentially independent. The issue then becomes that of combining the outputs of these statistically independent fading channels in accordance with a criterion that will provide improved receiver performance. In this section, we describe three different diversity-combining systems that do share a common feature: they all involve the use of linear receivers; hence the relative ease of their mathematical tractability.

Selection Combining

The block diagram of Figure 9.18 depicts a diversity-combining structure that consists of two functional blocks: N_r linear receivers and a logic circuit. This diversity system is said to be of a *selection combining* kind, in that given the N_r receiver outputs produced by a common transmitted signal, the logic circuit *selects* the particular receiver output with the *largest SNR* as the received signal. In conceptual terms, selection combining is the simplest form of space-diversity-on-receive system.

To describe the benefit of selection combining in statistical terms, we assume that the wireless communication channel is described by a *frequency-flat, slowly fading Rayleigh channel.* The implications of this assumption are threefold:

1. The frequency-flat assumption means that all the frequency components constituting the transmitted signal experience the same random attenuation and phase shift.
2. The slow-fading assumption means that fading remains essentially unchanged during the transmission of each symbol.
3. The fading phenomenon is described by the Rayleigh distribution.

Let $\tilde{s}(t)$ denote the complex envelope of the modulated signal transmitted during the symbol interval $0 \le t \le T$. Then, in light of the assumed channel, the complex envelope of the received signal of the kth diversity branch is defined by

$$\tilde{x}_k(t) = \alpha_k \exp(j\theta_k)\tilde{s}(t) + \tilde{w}_k(t), \qquad \begin{matrix} 0 \le t \le T \\ k = 1, 2, \ldots, N_r \end{matrix} \tag{9.60}$$

where, for the kth diversity branch, the fading is represented by the multiplicative term $\alpha_k \exp(j\theta_k)$ and the additive channel noise is denoted by $\tilde{w}_k(t)$. With the fading assumed to be slowly varying relative to the symbol duration T, we should be able to estimate and then remove the unknown phase shift θ_k at each diversity branch with sufficient accuracy, in which case (9.60) simplifies to

$$\tilde{x}_k(t) \approx \alpha_k \tilde{s}(t) + \tilde{w}_k(t), \qquad \begin{matrix} 0 \le t \le T \\ k = 1, 2, \ldots, N_r \end{matrix} \tag{9.61}$$

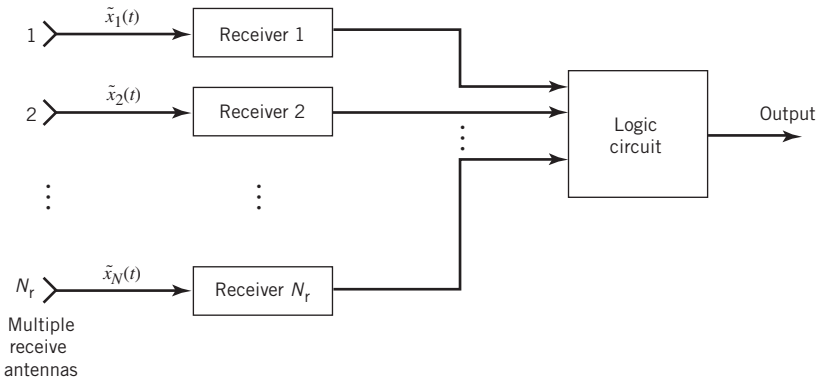


Figure 9.18 Block diagram of selection combiner, using N_r receive antennas.

The signal component of $\tilde{x}_k(t)$ is $\alpha_k\tilde{s}(t)$ and the noise component is $\tilde{w}_k(t)$. The average SNR at the output of the kth receiver is therefore

$$\begin{aligned}(\text{SNR})_k &= \frac{\mathbb{E}[|\alpha_k\tilde{s}(t)|^2]}{\mathbb{E}[|\tilde{w}_k(t)|^2]} \\ &= \left(\frac{\mathbb{E}[|\tilde{s}(t)|^2]}{\mathbb{E}[|\tilde{w}_k(t)|^2]}\right)\mathbb{E}[\alpha_k^2], \qquad k = 1, 2, \ldots, N_r\end{aligned}$$

Ordinarily, the mean-square value of $\tilde{w}_k(t)$ is the same for all k. Accordingly, we may express the $(\text{SNR})_k$ as

$$(\text{SNR})_k = \frac{E}{N_0}\mathbb{E}[\alpha_k^2], \qquad k = 1, 2, \ldots, N_r \tag{9.62}$$

where E is the symbol energy and $N_0/2$ is the noise spectral density. For binary data, E equals the transmitted signal energy per bit E_b.

Let γ_k denote the *instantaneous* SNR measured at the output of the kth receiver during the transmission of a given symbol. Then, replacing the mean-square value $\mathbb{E}[|\alpha_k|^2]$ by the instantaneous value $|\alpha_k|^2$ in (9.62), we may write

$$\gamma_k = \frac{E}{N_0}\alpha_k^2, \qquad k = 1, 2, \ldots, N_r \tag{9.63}$$

Under the assumption that the random amplitude α_k is Rayleigh distributed, the squared amplitude α_k^2 will be *exponentially distributed*[7] (i.e., chi-squared with two degrees of freedom, discussed in Appendix A). If we further assume that the average SNR over the short-term fading is the same, namely γ_{av}, for all the N_r diversity branches, then we may express the probability density functions of the random variables Γ_k pertaining to the individual branches as follows:

$$f_{\Gamma_k}(\gamma_k) = \frac{1}{\gamma_{av}}\exp\left(-\frac{\gamma_k}{\gamma_{av}}\right), \qquad \begin{matrix}\gamma_k \geq 0, \\ k = 1, 2, \ldots, N_r\end{matrix} \tag{9.64}$$

For some SNR γ, the associated cumulative distributions of the individual branches are described by

$$\begin{aligned}\mathbb{P}(\gamma_k \le \gamma) &= \int_{-\infty}^{\gamma} f_{\Gamma_k}(\gamma_k)\,\mathrm{d}\gamma_k \\ &= 1 - \exp\left(-\frac{\gamma}{\gamma_{\mathrm{av}}}\right), \qquad \gamma \ge 0\end{aligned} \tag{9.65}$$

for $k = 1, 2, \ldots, N_{\mathrm{r}}$. Since, by design, the N_{r} diversity branches are essentially statistically independent, the probability that all the diversity branches have an SNR less than the threshold γ is the product of the individual probabilities that $\gamma_k < \gamma$ for all k; thus, using (9.64) in (9.65), we write

$$\begin{aligned}\mathbb{P}(\gamma_k < \gamma) &= \prod_{k=1}^{N_{\mathrm{r}}} \mathbb{P}(\gamma_k < \gamma) \\ &= \prod_{k=1}^{N_{\mathrm{r}}} \left[1 - \exp\left(-\frac{\gamma}{\gamma_{\mathrm{av}}}\right)\right] \\ &= \left[1 - \exp\left(-\frac{\gamma}{\gamma_{\mathrm{av}}}\right)\right]^{N_{\mathrm{r}}}, \qquad \gamma \ge 0\end{aligned} \tag{9.66}$$

for $k = 1, 2, \ldots, N_{\mathrm{r}}$; note that the probability in (9.66) decreases with increasing N_{r}.

The cumulative distribution function of (9.66) is the same as the cumulative distribution function of the random variable Γ_{sc} described by the sample value

$$\gamma_{\mathrm{sc}} = \max\{\gamma_1, \gamma_2, \ldots, \gamma_{N_{\mathrm{r}}}\} \tag{9.67}$$

which is less than the threshold γ if, and only if, the individual SNRs $\gamma_1, \gamma_2, \ldots, \gamma_{N_{\mathrm{r}}}$ are all less than γ. Indeed, the cumulative distribution function of the selection combiner (i.e., the probability that all of the N_{r} diversity branches have an SNR less than γ) is given by

$$F_{\Gamma}(\gamma_{\mathrm{sc}}) = \left[1 - \exp\left(-\frac{\gamma_{\mathrm{sc}}}{\gamma_{\mathrm{av}}}\right)\right]^{N_{\mathrm{r}}}, \qquad \gamma_{\mathrm{sc}} \ge 0 \tag{9.68}$$

By definition, the probability density function $f_{\Gamma}(\gamma_{\mathrm{sc}})$ is the derivative of the cumulative distribution function $F_{\Gamma}(\gamma_{\mathrm{sc}})$ with respect to the argument γ_{sc}. Hence, differentiating (9.68) with respect to γ_{sc} yields

$$\begin{aligned}f_{\Gamma}(\gamma_{\mathrm{sc}}) &= \frac{\mathrm{d}}{\mathrm{d}\gamma_{\mathrm{sc}}} F_{\Gamma}(\gamma_{\mathrm{sc}}) \\ &= \frac{N_{\mathrm{r}}}{\gamma_{\mathrm{av}}} \exp\left(-\frac{\gamma_{\mathrm{sc}}}{\gamma_{\mathrm{av}}}\right)\left[1 - \exp\left(-\frac{\gamma_{\mathrm{sc}}}{\gamma_{\mathrm{av}}}\right)\right]^{N_{\mathrm{r}}-1}, \qquad \gamma_{\mathrm{sc}} \ge 0\end{aligned} \tag{9.69}$$

For convenience of graphical presentation, we use the scaled probability density function

$$f_X(x) = \gamma_{\mathrm{av}} f_{\Gamma_{\mathrm{sc}}}(\gamma_{\mathrm{sc}})$$

where the sample value x of the normalized variable X is defined by

$$x = \gamma_{sc}/\gamma_{av}$$

Figure 9.19 plots $f_X(x)$ versus x for varying number of receive-diversity branches N_r under the assumption that the short-term SNRs for all the N_r branches share the common value γ_{av}. From this figure we make two observations:

1. As the number of diversity branches N_r is increased, the probability density function $f_X(x)$ of the normalized random variable $X = \Gamma/\gamma_{av}$ progressively moves to the right.
2. The probability density function $f_X(x)$ becomes more and more symmetrical and, therefore, Gaussian as N_r is increased.

Stated in another way, a frequency-flat, slowly fading Rayleigh channel is modified through the use of selection combining into a Gaussian channel provided that the number of diversity channels N_r is sufficiently large. Realizing that a Gaussian channel is a *digital communication theorist's dream*, we now see the practical benefit of using selection combining.

According to the theory described herein, the selection-combining procedure requires that we monitor the receiver outputs in a continuous manner and, at each instant of time, select the receiver with the strongest signal (i.e., the largest instantaneous SNR). From a practical perspective, such a selective procedure is rather cumbersome. We may overcome

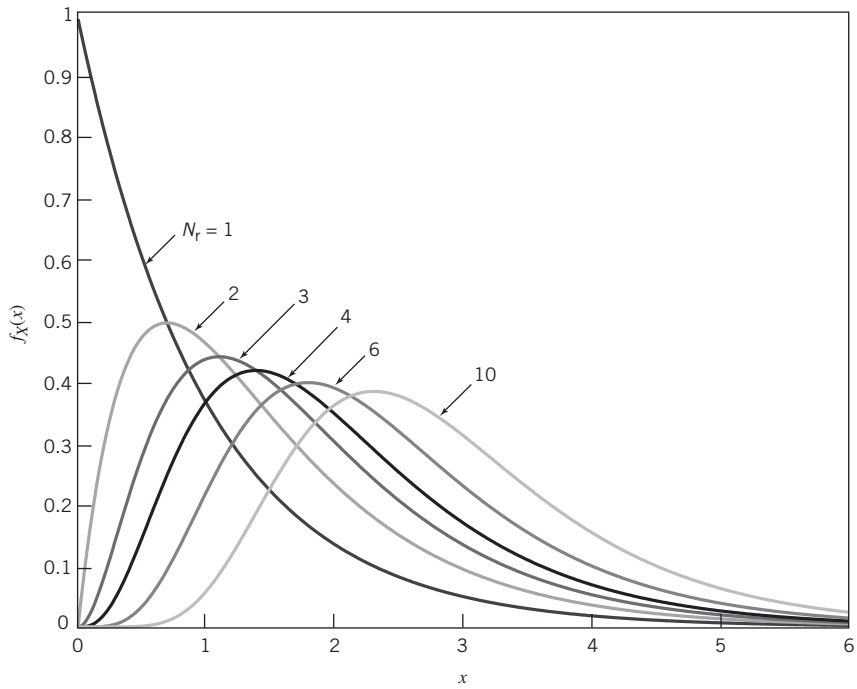


Figure 9.19 Normalized probability density function $f_X(x) = N_r \exp(-x)[1 - \exp(-x)]^{N_r - 1}$ for a varying number N_r of receive antennas.

this practical difficulty by adopting a *scanning* version of the selection-combining procedure:

- Start the procedure by selecting the receiver with the strongest output signal.
- Maintain using the output of this particular receiver as the combiner's output so long as its instantaneous SNR does not drop below a prescribed threshold.
- As soon as the instantaneous SNR of the combiner falls below the threshold, select a new receiver that offers the strongest output signal and continue the procedure.

This technique has a performance very similar to the nonscanning version of selective diversity.

EXAMPLE 5 **Outage Probability of Selection Combiner**

The *outage probability* of a diversity combiner is defined as the *percentage of time the instantaneous output SNR of the combiner is below some prescribed level for a specified number of branches*. Using the cumulative distribution function of (9.68), Figure 9.20 plots the outage curves for the selection combiner with N_r as the running parameter. The horizontal axis of the figure represents the instantaneous output SNR of the combiner relative to 0 dB (i.e., the 50-percentile point for $N_r = 1$) and the vertical axis represents the outage probability, expressed as a percentage. From the figure we observe the following:

The fading depth introduced through the use of space diversity on receive diminishes rapidly with the increase in the number of diversity branches.

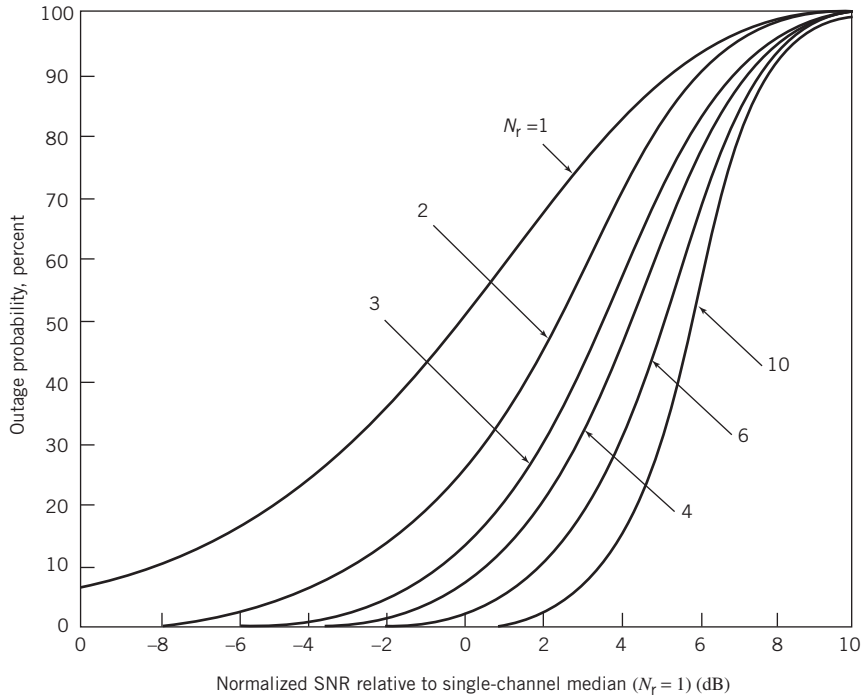


Figure 9.20
Outage probability for selector combining for a varying number N_r of receive antennas.

Maximal-Ratio Combining

The selection-combining technique just described is relatively straightforward to implement. However, from a performance point of view, it is not optimum, in that it ignores the information available from all the diversity branches except for the particular branch that produces the largest instantaneous power of its own demodulated signal.

This limitation of the selection combiner is mitigated by the *maximal-ratio combiner*,[8] the composition of which is described by the block diagram of Figure 9.21 that consists of N_r linear receivers followed by a linear combiner. Using the complex envelope of the received signal at the kth diversity branch given in (9.60), the corresponding complex envelope of the linear combiner output is defined by

$$\begin{aligned}\tilde{y}(t) &= \sum_{k=1}^{N_r} a_k \tilde{x}_k(t) \\ &= \sum_{k=1}^{N_r} a_k [\alpha_k \exp(j\theta_k)\tilde{s}(t) + \tilde{w}_k(t)] \\ &= \tilde{s}(t) \sum_{k=1}^{N_r} a_k \alpha_k \exp(j\theta_k) + \sum_{k=1}^{N_r} a_k \tilde{w}_k(t)\end{aligned} \tag{9.70}$$

where the a_k are *complex weighting parameters* that characterize the linear combiner. These parameters are changed from instant to instant in accordance with signal variations in the N_r diversity branches over the short-term fading process. The requirement is to design the linear combiner so as to maximize the output SNR of the combiner at each instant of time. From (9.70), we note the following two points:

1. The complex envelope of the output signal equals the first expression $\tilde{s}(t) \sum_{k=1}^{N_r} a_k \alpha_k \exp(j\theta_k)$.
2. The complex envelope of the output noise equals the second expression $\sum_{k=1}^{N_r} a_k \tilde{w}_k(t)$.

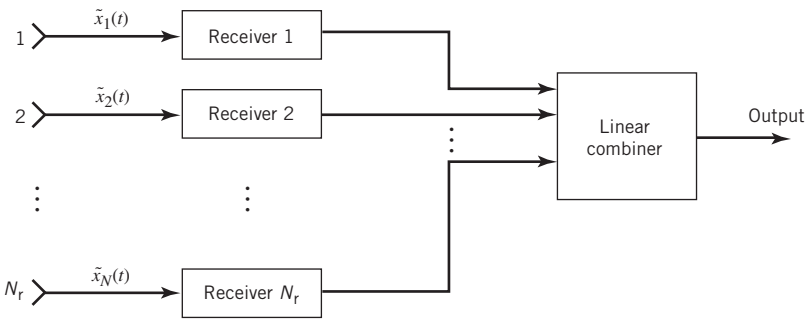


Figure 9.21 Block diagram of maximal-ratio combiner using N_r receive antennas.

Assuming that the $\tilde{w}_k(t)$ are mutually independent for $k = 1, 2, \ldots, N_r$, the output SNR of the linear combiner is therefore given by

$$
\begin{aligned}
(\mathrm{SNR})_c &= \frac{\mathbb{E}\left[\left|\tilde{s}(t)\sum_{k=1}^{N_r} a_k \alpha_k \exp(j\theta_k)\right|^2\right]}{\mathbb{E}\left[\left|\sum_{k=1}^{N_r} a_k \tilde{w}_k(t)\right|^2\right]} \\
&= \frac{\mathbb{E}[|\tilde{s}(t)|^2]}{\mathbb{E}[|\tilde{w}_k(t)|^2]} \frac{\mathbb{E}\left[\left|\sum_{k=1}^{N_r} a_k \alpha_k \exp(j\theta_k)\right|^2\right]}{\mathbb{E}\left[\sum_{k=1}^{N_r} |a_k|^2\right]} \\
&= \frac{E}{N_0} \frac{\mathbb{E}\left[\left|\sum_{k=1}^{N_r} a_k \alpha_k \exp(j\theta_k)\right|^2\right]}{\mathbb{E}\left[\sum_{k=1}^{N_r} |a_k|^2\right]}
\end{aligned}
\tag{9.71}
$$

where E/N_0 is the *symbol energy-to-noise spectral density ratio*.

Let γ_c denote the *instantaneous output SNR* of the linear combiner. Then, using the two terms

$$
\left|\sum_{k=1}^{N_r} a_k \alpha_k \exp(j\theta_k)\right|^2 \quad \text{and} \quad \sum_{k=1}^{N_r} |a_k|^2
$$

as the instantaneous values of the expectations in the numerator and denominator of (9.71), respectively, we may write

$$
\gamma_c = \frac{E}{N_0} \frac{\left|\sum_{k=1}^{N_r} a_k \alpha_k \exp(j\theta_k)\right|^2}{\sum_{k=1}^{N_r} |a_k|^2}
\tag{9.72}
$$

The requirement is to maximize γ_c with respect to the a_k. This maximization may be carried out by following the standard differentiation procedure, recognizing that the weighting parameters a_k are complex. However, we choose to follow a simpler procedure based on the Schwarz inequality, which was discussed in Chapter 7.

Let a_k and b_k denote any two complex numbers for $k = 1, 2, \ldots, N_r$. According to the *Schwarz inequality* for complex parameters, we have

$$\left| \sum_{k=1}^{N_r} a_k b_k \right|^2 \leq \sum_{k=1}^{N_r} |a_k|^2 \sum_{k=1}^{N_r} |b_k|^2 \tag{9.73}$$

which holds with equality for $a_k = c b_k^*$, where c is some arbitrary complex constant and the asterisk denotes complex conjugation.

Thus, applying the Schwarz inequality to the instantaneous output SNR of (9.72), with a_k left intact and b_k set equal to $\alpha_k \exp(j\theta_k)$, we obtain

$$\gamma_c \leq \frac{E}{N_0} \frac{\sum_{k=1}^{N_r} |a_k|^2 \sum_{k=1}^{N_r} |\alpha_k \exp(j\theta_k)|^2}{\sum_{k=1}^{N_r} |a_k|^2}$$

Canceling common terms in the numerator and denominator, we readily obtain

$$\gamma_c \leq \frac{E}{N_0} \sum_{k=1}^{N_r} \alpha_k^2 \tag{9.74}$$

Equation (9.74) proves that, in general, γ_c cannot exceed $\sum_k \gamma_k$, where γ_k is as defined in (9.63). The equality in (9.74) holds for

$$\begin{aligned} a_k &= c[\alpha_k \exp(j\theta_k)]^* \\ &= c\,\alpha_k^* \exp(-j\theta_k), \qquad k = 1, 2, \ldots, N_r \end{aligned} \tag{9.75}$$

where c is some arbitrary complex constant.

Equation (9.75) defines the complex weighting parameters of the maximal-ratio combiner. Based on this equation, we may state that the optimal weighting factor a_k for the kth diversity branch has a magnitude proportional to the signal amplitude α_k and a phase that cancels the signal phase θ_k to within some value that is identical for all the N_r diversity branches. The phase alignment just described has an important implication: it permits the *fully coherent addition* of the N_r receiver outputs by the linear combiner.

Equation (9.74) with the equality sign defines the instantaneous output SNR of the maximal-ratio combiner, which is written as

$$\gamma_{mrc} = \frac{E}{N_0} \sum_{k=1}^{N_r} \alpha_k^2 \tag{9.76}$$

According to (9.62), $(E/N_0)\alpha_k^2$ is the *instantaneous output SNR* of the kth diversity branch. Hence, the maximal-ratio combiner produces an instantaneous output SNR that is the sum of the instantaneous SNRs of the individual branches; that is,

$$\gamma_{mrc} = \sum_{k=1}^{N_r} \gamma_k \tag{9.77}$$

The term "maximal-ratio combiner" has been coined to describe the combiner of Figure 9.21 that produces the optimum result given in (9.77). Indeed, we deduce from this result that the instantaneous output SNR of the maximal-ratio combiner can be large even when the SNRs of the individual branches are small. Since the instantaneous SNR produced by the selection combiner is simply the largest among the N_r terms of (9.77), it follows that:

> The selection combiner is clearly inferior in performance to the maximal-ratio combiner.

The maximal SNR γ_{mrc} is the sample value of a random variable denoted by Γ. According to (9.76), γ_{mrc} is equal to the sum of N_r exponentially distributed random variables for a frequency-flat, slowly fading Rayleigh channel. From Appendix A, the probability density function of such a sum is known to be *chi-square with* $2N_r$ *degrees of freedom*; that is,

$$f_\Gamma(\gamma_{mrc}) = \frac{1}{(N_r-1)!}\frac{\gamma_{mrc}^{N_r-1}}{\gamma_{av}^{N_r}}\exp\left(-\frac{\gamma_{mrc}}{\gamma_{av}}\right) \tag{9.78}$$

Note that for $N_r = 1$, (9.69) and (9.78) assume the same value, which is to be expected.

Figure 9.22 plots the scaled probability density function, $f_X(x) = \gamma_{av} f_\Gamma(\gamma_{mrc})$, versus the normalized variable $x = \gamma_{mrc}/\gamma_{av}$ for varying N_r. Based on this figure, we may make

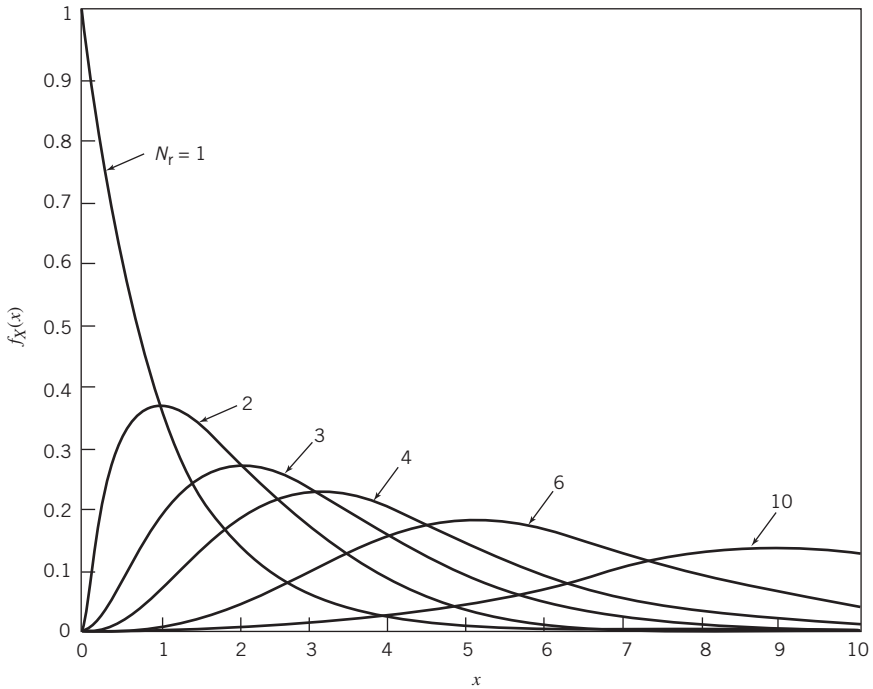


Figure 9.22 Normalized probability density function $f_X(x) = \frac{1}{(N_r-1)}x^{N_r-1}\exp(-x)$ for a varying number of N_r receive antennas.

observations similar to those for the selection combiner, except for the fact that for any N_r we find that the scaled probability density function for the maximal-ratio combiner is radically different from its counterpart for the selection combiner.

EXAMPLE 6

Outage Probability for Maximal-Ratio Combiner

The cumulative distribution function for the maximal-ratio combiner is defined by

$$\begin{aligned}\mathbb{P}(\gamma_{\text{mrc}} < x) &= \int_0^x f_\Gamma(\gamma_{\text{mrc}})\,\mathrm{d}\gamma_{\text{mrc}} \\ &= 1 - \int_x^\infty f_\Gamma(\gamma_{\text{mrc}})\,\mathrm{d}\gamma_{\text{mrc}}\end{aligned} \tag{9.79}$$

where the probability density function $f_\Gamma(\gamma_{\text{mrc}})$ is itself defined by (9.78). Using (9.79), Figure 9.23 plots the outage probability for the maximal-ratio combiner with N_r as a running parameter. Comparing this figure with that of Figure 9.20 for selection combining, we see that the outage-probability curves for these two diversity techniques are superficially similar. The *diversity gain*, defined as the E/N_0 saving at a given BER, provides a measure of the effectiveness of a diversity technique on an outage-probability basis.

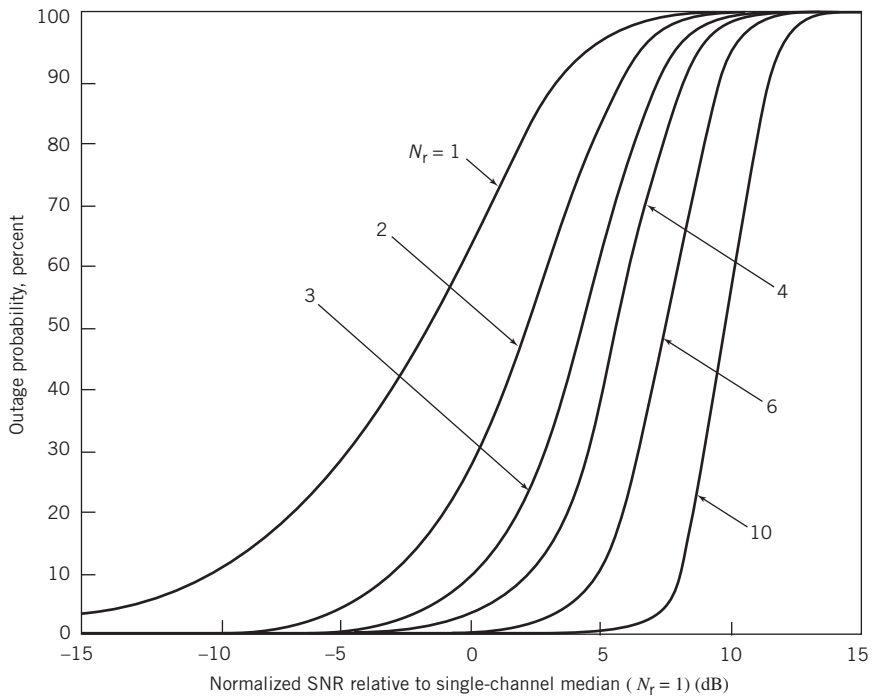


Figure 9.23 Outage probability of maximal-ratio combiner for a varying number N_r of receiver antennas.

Equal-Gain Combining

In a theoretical context, the maximal-ratio combiner is the *optimum* among linear diversity combining techniques, optimum in the sense that it produces the largest possible value of instantaneous output SNR. However, in practical terms, there are three important issues to keep in mind:[9]

1. Significant instrumentation is needed to adjust the complex weighting parameters of the maximal-ratio combiner to their exact values, in accordance with (9.75).
2. The additional improvement in output SNR gained by the maximal-ratio combiner over the selection combiner is not that large, and it is quite likely that the additional improvement in receiver performance is lost in not being able to achieve the exact setting of the maximal-ratio combiner.
3. So long as a linear combiner uses the diversity branch with the strongest signal, then other details of the combiner may result in a minor improvement in overall receiver performance.

Issue 3 points to formulation of the so-called *equal-gain combiner*, in which all the complex weighting parameters a_k have their phase angles set opposite to those of their respective multipath branches in accordance with (9.75). But, unlike the a_k in the maximal-ratio combiner, their magnitudes are set equal to some constant value, unity for convenience of use.

9.9 "Space Diversity-on-Transmit" Systems

In the wireless communications literature, space diversity-on-receive techniques are commonly referred to as *orthogonal space–time block codes* (Tarokh et al., 1999). This terminology is justified on the following grounds:

1. The transmitted symbols form an orthogonal set.
2. The transmission of incoming data streams is carried out on a block-by-block basis.
3. Space and time constitute the coordinates of each transmitted block of symbols.

In a generic sense, Figure 9.24 presents the baseband diagram of a *space–time block encoder*, which consists of two functional units: mapper and block encoder. The *mapper* takes the incoming binary data stream $\{b_k\}$, where $b_k = \pm 1$, and generates a new *sequence of blocks* with each block made up of multiple symbols that are complex. For example, the mapper may be in the form of an *M-ary PSK* or *M-ary QAM message constellation*, which are illustrated for $M = 16$ in the signal-space diagrams of Figure 9.25. All the symbols in a particular column of the transmission matrix are pulse-shaped (in accordance with the

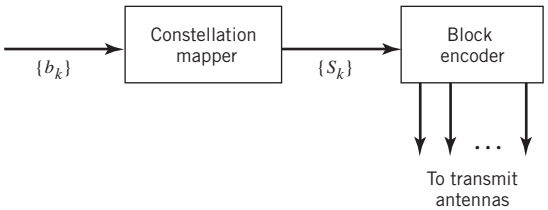


Figure 9.24 Block diagram of orthogonal space–time block encoder.

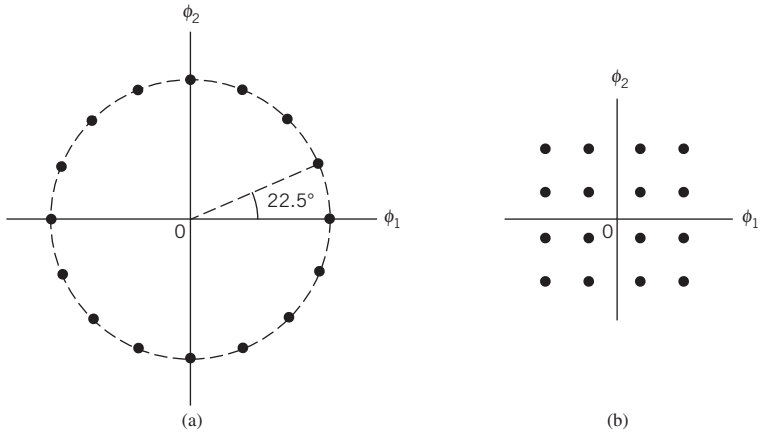


Figure 9.25 (a) Signal constellation of 16-PSK. (b) Signal constellation of 16-QAM.

criteria described in Chapter 8) and then modulated into a form suitable for simultaneous transmission over the channel by the transmit antennas. The pulse shaper and modulator are not shown in Figure 9.24 as the basic issue of interest is that of baseband data transmission with emphasis on the formulation of space–time block codes. The *block encoder* converts each block of complex symbols produced by the mapper into an l-by-N_t transmission matrix $\mathbf{S}$, where l and N_t are respectively the *temporal* dimension and *spatial* dimension of the transmission matrix. The individual elements of the transmission matrix $\mathbf{S}$ are made up of linear combinations of $\tilde{s}_k$ and $\tilde{s}_k^*$, where the $\tilde{s}_k$ are complex symbols and the $\tilde{s}_k^*$ are their complex conjugates.

EXAMPLE 7 **Quadriphase Shift Keying**

As a simple example, consider the map portrayed by the QPSK, $M = 4$. This map is described in Table 9.2, where E is the transmitted signal energy per symbol.

The input dibits (pairs of binary bits) are *Gray encoded*, wherein only one bit is flipped as we move from one symbol to the next. (Gray encoding was discussed in Section 7.6 under "Quadriphase Shift Keying".) The mapped signal points lie on a circle of radius $\sqrt{E}$ centered at the origin of the signal-space diagram.

Table 9.2 **Gray-encoded QPSK mapper**

Dibit: $i = 1, 2, 3, 4$	Coordinates of mapped signal points: s_i, $i = 1, 2, 3, 4$
10	$\sqrt{E/2}(1, -1) = \sqrt{E}\exp(j7\pi/4)$
11	$\sqrt{E/2}(-1, -1) = \sqrt{E}\exp(j5\pi/4)$
01	$\sqrt{E/2}(-1, +1) = \sqrt{E}\exp(j3\pi/4)$
00	$\sqrt{E/2}(+1, +1) = \sqrt{E}\exp(j\pi/4)$

Alamouti Code

Example 6 is illustrative of the *Alamouti code*, which is one of the first space–time block codes involving the use of two transmit antennas and one signal receive antenna (Alamouti, 1998). Figure 9.26 shows a baseband block diagram of this highly popular spatial code.

Let $\tilde{s}_1$ and $\tilde{s}_2$ denote the complex symbols produced by the code's mapper, which are to be transmitted over the multipath wireless channel by two transmit antennas. Signal transmission over the channel proceeds as follows:

1. At some arbitrary time t, antenna 1 transmits $\tilde{s}_1$ and simultaneously antenna 2 transmits $\tilde{s}_2$.
2. At time $t + T$, where T is the symbol duration, signal transmission is switched to $-\tilde{s}_2^*$ transmitted by antenna 1 and simultaneously $\tilde{s}_1^*$ is transmitted by antenna 2.

The resulting two-by-two space–time block code is written in matrix form as follows:

$$\mathbf{S} = \begin{bmatrix} \tilde{s}_1 & \tilde{s}_2 \\ -\tilde{s}_2^* & \tilde{s}_1^* \end{bmatrix} \begin{matrix} \Rightarrow \text{Time} \\ \Downarrow \\ \text{Space} \end{matrix} \tag{9.80}$$

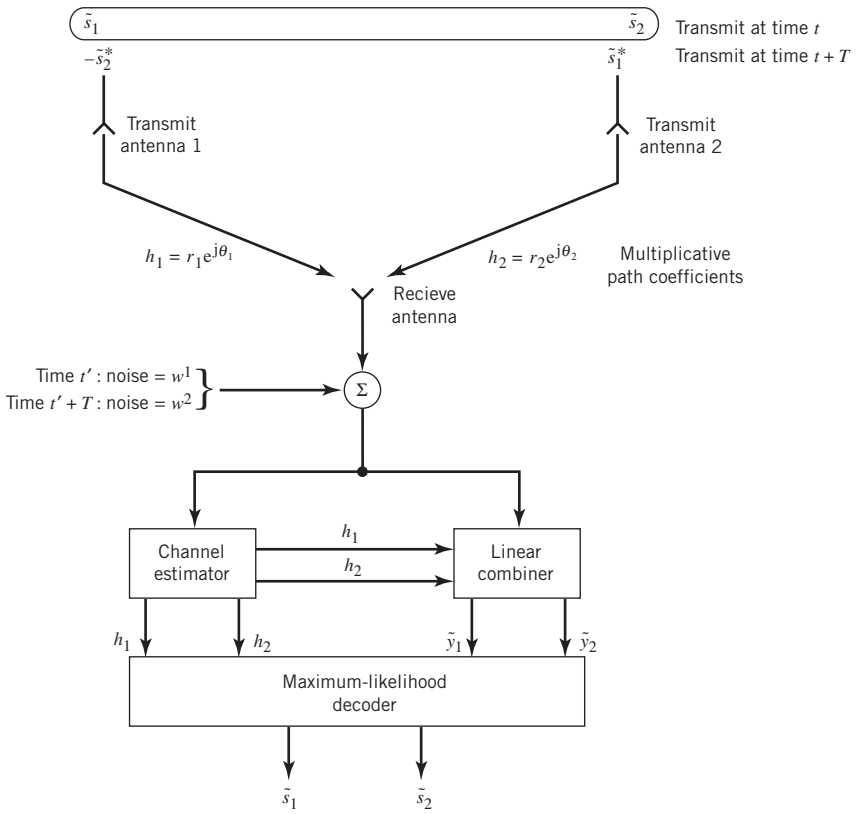


Figure 9.26 Block diagram of the transceiver (transmitter and receiver) for the Alamouti code. Note that $t' > t$ to allow for propagation delay.

This *transmission matrix* is a *complex-orthogonal matrix* (quaternion) in that it satisfies the condition for orthogonality in both the spatial and temporal senses. To demonstrate this important property of the Alamouti, let

$$\mathbf{S}^\dagger = \begin{bmatrix} \tilde{s}_1^* & -\tilde{s}_2 \\ \tilde{s}_2^* & \tilde{s}_1 \end{bmatrix} \Rightarrow \text{Space} \qquad (9.81)$$

$$\Downarrow$$

$$\text{Time}$$

denote the *Hermitian transpose* of $\mathbf{S}$, which involves both transposition and complex conjugation. To demonstrate orthogonality in the spatial sense, we multiply the code matrix $\mathbf{S}$ by its Hermitian transpose $\mathbf{S}^\dagger$ on the right, obtaining

$$\begin{aligned} \mathbf{S}\mathbf{S}^\dagger &= \begin{bmatrix} \tilde{s}_1 & \tilde{s}_2 \\ -\tilde{s}_2^* & \tilde{s}_1^* \end{bmatrix} \begin{bmatrix} \tilde{s}_1^* & -\tilde{s}_2 \\ \tilde{s}_2^* & \tilde{s}_1 \end{bmatrix} \\ &= \begin{bmatrix} |\tilde{s}_1|^2 + |\tilde{s}_2|^2 & -\tilde{s}_1\tilde{s}_2 + \tilde{s}_2\tilde{s}_1 \\ -\tilde{s}_2^*\tilde{s}_1^* + \tilde{s}_1^*\tilde{s}_2^* & |\tilde{s}_2|^2 + |\tilde{s}_1|^2 \end{bmatrix} \\ &= (|\tilde{s}_1|^2 + |\tilde{s}_2|^2)\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \end{aligned} \qquad (9.82)$$

Since the right-hand side of (9.81) is real valued, it follows that the alternative matrix product $\mathbf{S}^\dagger\mathbf{S}$, viewed in the temporal sense, yields exactly the same result. That is,

$$\mathbf{S}\mathbf{S}^\dagger = \mathbf{S}^\dagger\mathbf{S} = (|\tilde{s}_1|^2 + |\tilde{s}_2|^2)\mathbf{I} \qquad (9.83)$$

where $\mathbf{I}$ is the two-by-two identity matrix.

In light of (9.80) and (9.83), we may now summarize three important properties of the Alamouti code:

PROPERTY 1 **Unitarity (Complex Orthogonality)**

The Alamouti code is an orthogonal space–time block code, in that its transmission matrix is a unitary matrix with the sum term $|\tilde{s}_1|^2 + |\tilde{s}_2|^2$ *being merely a scaling factor.*

As a consequence of this property, the Alamouti code achieves full diversity.

PROPERTY 2 **Full-Rate Complex Code**

The Alamouti code (with two transmit antennas) is the only complex space–time block code with a code rate of unity in existence.

Hence, for any signal constellation, full diversity of the code is achieved at the full transmission rate.

PROPERTY 3 **Linearity**

The Alamouti code is linear in the transmitted symbols.

We may therefore expand the transmission matrix **S** of the code as a linear combination of the transmitted symbols and their complex conjugates, as shown by

$$\mathbf{S} = \tilde{s}_1\mathbf{\Gamma}_{11} + \tilde{s}_1^*\mathbf{\Gamma}_{12} + \tilde{s}_2\mathbf{\Gamma}_{21} + \tilde{s}_2^*\mathbf{\Gamma}_{22} \tag{9.84}$$

where the four constituent matrices are themselves defined as follows:

$$\mathbf{\Gamma}_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

$$\mathbf{\Gamma}_{12} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\mathbf{\Gamma}_{21} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

$$\mathbf{\Gamma}_{22} = \begin{bmatrix} 0 & 0 \\ -1 & 0 \end{bmatrix}$$

In words, the Alamouti code is the only two-dimensional space–time code, the transmission matrix of which can be decomposed into the form described in (9.84).

Receiver Considerations of the Alamouti Code

The discussion presented thus far has focused on the Alamouti code viewed from the transmitter's perspective. We turn next to the design of the receiver for decoding the code.

To this end, we assume that the channel is frequency-flat and slowly time varying, such that the complex multiplicative distribution introduced by the channel at time t is essentially the same as that at time $t + T$, where T is the symbol duration. As before, the multiplicative distortion is denoted by $\alpha_k e^{j\theta_k}$ where we now have $k = 1, 2$, as indicated in Figure 9.25. Thus, with the symbols $\tilde{s}_1$ and $\tilde{s}_2$ transmitted simultaneously at time t, the complex received signal at some time $t' > t$, allowing for propagation delay, is described by

$$\tilde{x}_1 = \alpha_1 e^{j\theta_1}\tilde{s}_1 + \alpha_2 e^{j\theta_2}\tilde{s}_2 + \tilde{w}_1 \tag{9.85}$$

where $\tilde{w}_1$ is the complex channel noise at time t'. Next, with the symbols $-\tilde{s}_2^*$ and $\tilde{s}_1^*$ transmitted simultaneously at time $t + T$, the corresponding complex signal received at time $t' + T$ is

$$\tilde{x}_2 = -\alpha_1 e^{j\theta_1}\tilde{s}_2^* + \alpha_2 e^{j\theta_2}\tilde{s}_1^* + \tilde{w}_2 \tag{9.86}$$

where $\tilde{w}_2$ is the second complex channel noise at time $t' + T$. To be more precise, the noise terms $\tilde{w}_1$ and $\tilde{w}_2$ are circularly-symmetric complex-valued uncorrelated Gaussian random variables of zero mean and equal variance.

In the course of time from t' to $t' + T$, the channel estimator in the receiver has sufficient time to produce estimates of the multiplicative distortion represented by $\alpha_k e^{j\theta_k}$

for $k = 1, 2$. Hereafter, we assume that these two estimates are accurate enough for them to be treated as essentially exact; in other words, the receiver has knowledge of both $\alpha_1 e^{j\theta_1}$ and $\alpha_2 e^{j\theta_2}$. Accordingly, we may formulate the combination of two variables, $\tilde{x}_1$ in (9.85) and the complex conjugate of $\tilde{x}_2$ in (9.86), in matrix form as follows:

$$\begin{aligned}\tilde{\mathbf{x}} &= \begin{bmatrix} \tilde{x}_1 \\ \tilde{x}_2^* \end{bmatrix} \\ &= \begin{bmatrix} \alpha_1 e^{j\theta_1} & \alpha_2 e^{j\theta_2} \\ \alpha_2 e^{-j\theta_2} & -\alpha_1 e^{-j\theta_1} \end{bmatrix} \begin{bmatrix} \tilde{s}_1 \\ \tilde{s}_2 \end{bmatrix} + \begin{bmatrix} \tilde{w}_1 \\ \tilde{w}_2 \end{bmatrix}\end{aligned} \tag{9.87}$$

The nice thing about this equation is that the original complex signals s_1 and s_2 appear as the vector of two unknowns. It is with this goal in mind that $\tilde{x}_1$ and $\tilde{x}_2^*$ were used for the elements of the two-by-one received signal vector $\tilde{\mathbf{x}}$, in the manner shown on the right-hand side of (9.87).

According to (9.87), the channel matrix of the transmit diversity in Figure 9.25 is defined by

$$\begin{aligned}\mathbf{H} &= \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} \\ &= \begin{bmatrix} \alpha_1 e^{j\theta_1} & \alpha_2 e^{j\theta_2} \\ \alpha_2 e^{-j\theta_2} & -\alpha_1 e^{-j\theta_1} \end{bmatrix}\end{aligned} \tag{9.88}$$

In a manner similar to the signal-transmission matrix $\tilde{\mathbf{S}}$, we find that the channel matrix $\mathbf{H}$ is also a *unitary matrix*, as shown by

$$\mathbf{H}^\dagger \mathbf{H} = (\alpha_1^2 + \alpha_2^2)\mathbf{I} \tag{9.89}$$

where, as before, $\mathbf{I}$ is the identity matrix and the sum term $\alpha_1^2 + \alpha_2^2$ is merely a scaling factor.

Using the definition of (9.88) for the channel matrix, we may rewrite (9.87) in the compact matrix form

$$\tilde{\mathbf{x}} = \mathbf{H}\tilde{\mathbf{s}} + \tilde{\mathbf{w}} \tag{9.90}$$

where

$$\tilde{\mathbf{s}} = \begin{bmatrix} \tilde{s}_1 \\ \tilde{s}_2 \end{bmatrix} \tag{9.91}$$

is the complex transmitted signal vector and

$$\tilde{\mathbf{w}} = \begin{bmatrix} \tilde{w}_1 \\ \tilde{w}_2 \end{bmatrix} \tag{9.92}$$

is the additive complex channel noise vector. Note that the column vector $\tilde{\mathbf{s}}$ in (9.91) is the same as the first row vector in the matrix $\tilde{\mathbf{S}}$ of (9.80).

We have now reached a point where we have to address the fundamental issue in designing the receiver:

How do we decode the Alamouti code, given the received signal vector $\tilde{\mathbf{x}}$?

To this end, we introduce a new complex two-by-one vector $\tilde{\mathbf{y}}$, defined as the matrix product of the received signal vector $\tilde{\mathbf{x}}$ and the Hermition transpose of the channel matrix $\mathbf{H}$ *normalized* with respect to the reciprocal sum term $\alpha_1^2 + \alpha_2^2$; that is,

$$\begin{aligned}\tilde{\mathbf{y}} &= \begin{bmatrix}\tilde{y}_1\\ \tilde{y}_2\end{bmatrix}\\ &= \left(\frac{1}{\alpha_1^2+\alpha_2^2}\right)\mathbf{H}^{\dagger}\tilde{\mathbf{x}}\end{aligned} \tag{9.93}$$

Substituting (9.90) into (9.93) and then making use of the unitarity property of the channel matrix described in (9.89), we obtain the mathematical basis for decoding of the Alamouti code:

$$\tilde{\mathbf{y}} = \tilde{\mathbf{s}} + \tilde{\mathbf{v}} \tag{9.94}$$

where $\tilde{\mathbf{v}}$ is a modified form of the complex channel noise $\tilde{\mathbf{w}}$, as shown by

$$\tilde{\mathbf{v}} = \left(\frac{1}{\alpha_1^2+\alpha_2^2}\right)\mathbf{H}^{\dagger}\tilde{\mathbf{w}} \tag{9.95}$$

Substituting (9.88) and (9.92) into (9.95), the expanded form of the complex noise vector $\tilde{\mathbf{v}}$ is defined as follows:

$$\begin{bmatrix}\tilde{v}_1\\ \tilde{v}_2\end{bmatrix} = \frac{1}{\alpha_1^2+\alpha_2^2}\begin{bmatrix}\alpha_1 e^{-j\theta_1}\tilde{w}_1 & + & \alpha_2 e^{j\theta_2}\tilde{w}_2^*\\ \alpha_2 e^{-j\theta_2}\tilde{w}_1 & - & \alpha_1 e^{j\theta_1}\tilde{w}_2^*\end{bmatrix} \tag{9.96}$$

Hence, we may go on to simply write

$$\tilde{y}_k = \tilde{s}_k + \tilde{v}_k, \qquad k = 1, 2 \tag{9.97}$$

Examination of (9.97) leads us to make the following statement insofar as the receiver is concerned:

The space–time channel is decoupled into a pair of scalar channels that are statistically independent of each other:

1. The complex symbol $\tilde{s}_k$ at the output of the kth space–time channel is identical to the complex symbol transmitted by the kth antenna for $k = 1, 2$; the decoupling shown clearly in (9.97) is attributed to complex orthogonality of the Alamouti code.
2. Assuming that the original channel noise $\tilde{w}_k$ is white Gaussian, then this statistical characterization is maintained in the modified noise $\tilde{v}_k$ appearing at the output of the kth space–time channel for $k = 1, 2$; this maintenance is attributed to the processing performed in the receiver.

This twofold statement hinges on the premise that the receiver has knowledge of the channel matrix $\mathbf{H}$.

Moreover, with two transmit antennas and one receive antenna, the Alamouti code achieves the same level of diversity as a corresponding system with one transmit antenna and two receive antennas. It is in this sense that a wireless communication system based on the Alamouti code is said to enjoy a *two-level diversity gain*.

Maximum Likelihood Decoding

Figure 9.27 illustrates the signal-space diagram of an Alamouti-encoded system based on the QPSK constellation. The complex Gaussian noise clouds centered on the four signal points and with decreasing intensity illustrate the effects of complex noise term $\tilde{\mathbf{v}}$ on the linear combiner output $\tilde{\mathbf{y}}$.

In effect, the picture portrayed in Figure 9.27 is the graphical representation of (9.94) over two successive symbol transmissions at times t and $t + T$, repeated a large number of times.

Suppose that the two signal constellations in the top half of the signal-space diagram in Figure 9.27 represent the pair of symbols transmitted at time t, for which we write

$$\tilde{\mathbf{s}}_t = \begin{bmatrix} \tilde{s}_1 \\ \tilde{s}_2 \end{bmatrix}$$

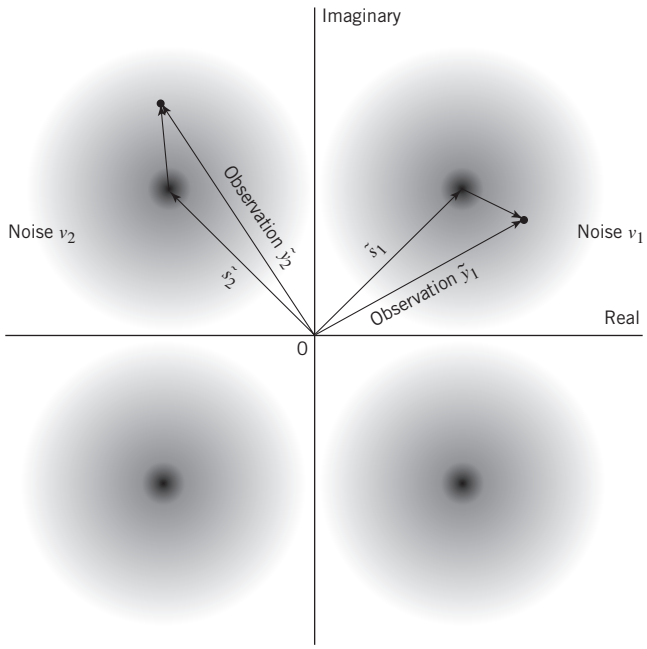


Figure 9.27 Signal-space diagram for Alamouti code, using the QPSK signal constellation. The signal points $\tilde{s}_1$ and $\tilde{s}_2$ and the corresponding linear normalized combiner outputs $\tilde{y}_1$ and $\tilde{y}_2$ are displayed in the top half of the figure.

Then, the remaining two signal constellations positioned in the right half of Figure 9.27 represent the other pair of symbols transmitted at $t + T$, for which we write

$$\tilde{\mathbf{s}}_{t+T} = \begin{bmatrix} -\tilde{s}_2^* \\ +\tilde{s}_1^* \end{bmatrix}$$

On this basis, we may now invoke the *maximum likelihood decoding rule,* discussed in Chapter 7, to make the three-fold statement:

1. Compute the composite *squared Euclidean distance metric*

$$\left\| \tilde{\mathbf{y}}_t - \tilde{\mathbf{s}}_t \right\|^2 + \left\| \tilde{\mathbf{y}}_{t+T} - \tilde{\mathbf{s}}_{t+T} \right\|^2$$

   produced by sending signal vectors $\tilde{\mathbf{s}}_t$ and $\tilde{\mathbf{s}}_{t+T}$, respectively.
2. Do this computation for all four possible signal pairs in the QPSK constellation.
3. Hence, the ML decoder selects the pair of signals for which the metric is the smallest.

The metric's component $\left\| \tilde{\mathbf{y}}_t - \tilde{\mathbf{s}}_t \right\|^2$ in part 1 of this statement is illustrated in Figure 9.27.

9.10 "Multiple-Input, Multiple-Output" Systems: Basic Considerations

In Sections 9.8 and 9.9, we studied space-diversity wireless communication systems employing either multiple receive or multiple transmit antennas to combat the multipath fading problem. In effect, fading was treated as a source that degrades performance, necessitating the use of space diversity on receive or transmit to mitigate it. In this section, we discuss *MIMO wireless communication*, which distinguishes itself in the following ways:[10]

1. The fading phenomenon is viewed not as a nuisance but rather as an environmental source of enrichment to be exploited.
2. Space diversity at both the transmit and receive ends of the wireless communication link may provide the basis for a significant increase in channel capacity.
3. Unlike conventional techniques, the increase in channel capacity is achieved by increasing computational complexity while maintaining the primary communication resources (i.e., total transmit power and channel bandwidth) fixed.

Coantenna Interference

Figure 9.28 shows the block diagram of a MIMO wireless link. The signals transmitted by the N_t transmit antennas over the wireless channel are all chosen to lie inside a common frequency band. Naturally, the transmitted signals are scattered differently by the channel. Moreover, owing to multiple signal transmissions, the system experiences a spatial form of signal-dependent interference, called *coantenna interference* (*CAI*).

Figure 9.29 illustrates the effect of CAI for one, two, and eight simultaneous transmissions and a single receive antenna (i.e., $N_t = 1, 2, 8$ and $N_r = 1$) using binary PSK; the transmitted binary PSK signals used in the simulation resulting in this figure were

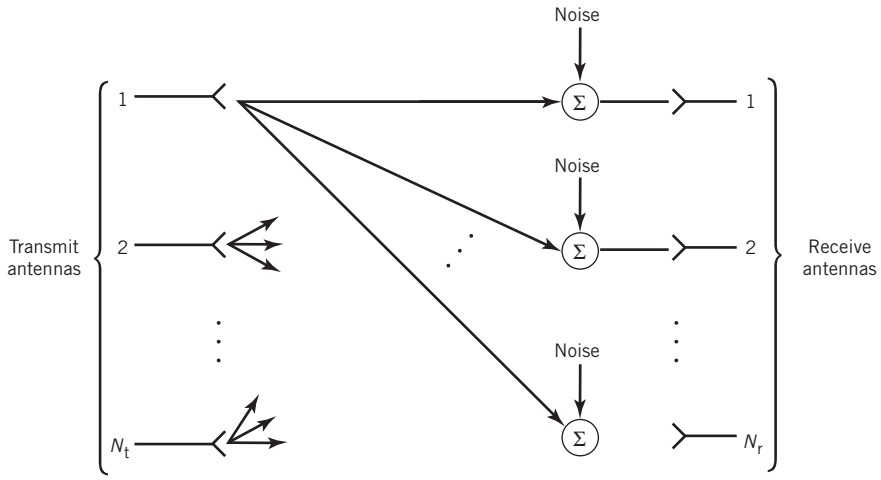


Figure 9.28 Block diagram of MIMO wireless link with N_t transmit antennas and N_r receive antennas.

different but they all had the same average power and occupied the same bandwidth. (Sellathurai and Haykin, 2008). Figure 9.29 clearly shows the difficulty that arises due to CAI when the number of transmit antennas N_t is large. In particular, with eight simultaneous signal transmissions, the *eye pattern* of the received signal is practically closed. The challenge for the receiver is how to mitigate the CAI problem and thereby make it possible to provide increased spectral efficiency.

In a theoretical context, the spectral efficiency of a communication system is intimately linked to the channel capacity of the system. To proceed with evaluation of the channel capacity of MIMO wireless communication, we begin by formulating a baseband channel model for the system as described next.

Basic Baseband Channel Model

Consider a MIMO narrowband wireless communication system built around a flat-fading channel, with N_t transmit antennas and N_r receive antennas. The antenna configuration is hereafter referred to as the pair (N_t, N_r). For a statistical analysis of the MIMO system in what follows, we use baseband representations of the transmitted and received signals as well as the channel. In particular, we introduce the following notation:

- The spatial parameter

$$N = \min\{N_t, N_r\} \tag{9.98}$$

  defines new degrees of freedom introduced into the wireless communication system by using a MIMO channel with N_t transmit antennas and N_r receive antennas.

- The N_t-by-1 vector

$$\tilde{\mathbf{s}}(n) = [\tilde{s}_1(n), \tilde{s}_2(n), \ldots, \tilde{s}_{N_t}(n)]^T \tag{9.99}$$

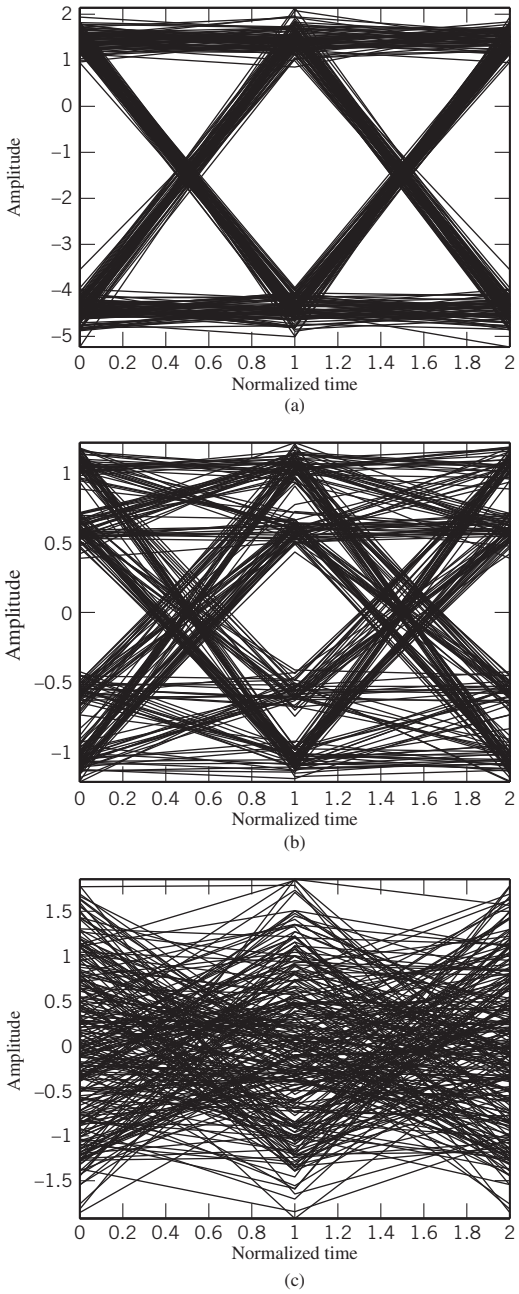


Figure 9.29 Effect of coantenna interference on the eye diagram for one receive antenna and different numbers of transmit antennas. (a) $N_t = 1$, (b) $N_t = 2$, (c) $N_t = 8$.

denotes the complex signal vector transmitted by the N_t antennas at discrete time n. The symbols constituting the vector $\tilde{\mathbf{s}}(n)$ are assumed to have zero mean and common variance σ_s^2. The total transmit power is *fixed* at the value

$$P = N_t \sigma_s^2 \tag{9.100}$$

For P to be maintained constant, the variance σ_s^2 (i.e., power radiated by each transmit antenna) must be inversely proportional to N_t.

- For a flat-fading Rayleigh distributing channel, we may use $\tilde{h}_{ik}(n)$ to denote the sampled complex gain of the channel coupling transmit antenna k to receive antenna i at discrete time n, where $i = 1, 2, \ldots, N_r$ and $k = 1, 2, \ldots, N_t$. We may thus express the N_r-by-N_t *complex channel matrix* as

$$\mathbf{H}(n) = \underbrace{\left.\begin{bmatrix} \tilde{h}_{11}(n) & \tilde{h}_{12}(n) & \cdots & \tilde{h}_{1N_t}(n) \\ \tilde{h}_{21}(n) & \tilde{h}_{22}(n) & \cdots & \tilde{h}_{2N_t}(n) \\ \vdots & \vdots & & \vdots \\ \tilde{h}_{N_r1}(n) & \tilde{h}_{N_r2} & \cdots & h_{N_rN_t}(n) \end{bmatrix}\right\}}_{N_t \text{ transmit antennas}} \left.\begin{array}{l} \\ \\ \\ \\ \end{array}\right\} \begin{array}{l} N_r \\ \text{receive} \\ \text{antennas} \end{array} \tag{9.101}$$

- The system of equations

$$\tilde{x}_i(n) = \sum_{k=1}^{N_t} \tilde{h}_{ik}(n)\tilde{s}_k(n) + \tilde{w}_i(n) \qquad \begin{cases} i = 1, 2, \ldots, N_r \\ k = 1, 2, \ldots, N_t \end{cases} \tag{9.102}$$

defines the complex signal received at the ith antenna due to the transmitted symbol $\tilde{s}_k(n)$ radiated by the kth antenna. The term $\tilde{w}_i(n)$ denotes the additive complex channel noise perturbing $\tilde{x}_i(n)$. Let the N_r-by-1 vector

$$\tilde{\mathbf{x}}(n) = [\tilde{x}_1(n), \tilde{x}_2(n), \ldots, \tilde{x}_{N_r}(n)] \tag{9.103}$$

denote the complex received signal vector and the N_r-by-1 vector

$$\tilde{\mathbf{w}}(n) = [\tilde{w}_1(n), \tilde{w}_2(n), \ldots, \tilde{w}_{N_r}(n)]^T \tag{9.104}$$

denote the complex channel noise vector. We may then rewrite the system of equations (9.102) in the compact matrix form

$$\tilde{\mathbf{x}}(n) = \mathbf{H}(n)\tilde{\mathbf{s}}(n) + \tilde{\mathbf{w}}(n) \tag{9.105}$$

Equation (9.105) describes the *basic complex channel model for MIMO wireless communications*, assuming the use of a flat-fading channel. The equation describes the input–output behavior of the channel at discrete time n. To simplify the exposition, hereafter we suppress the dependence on time n by simply writing

$$\tilde{\mathbf{x}} = \mathbf{H}\tilde{\mathbf{s}} + \tilde{\mathbf{w}} \tag{9.106}$$

where it is understood that all four vector/matrix terms of the equation, **s**, **H**, **w**, and **x**, are in actual fact dependent on the discrete time n. Figure 9.30 shows the basic channel model of (9.106).

For mathematical tractability, we assume a *Gaussian model* made up of three elements:

1. N_t symbols, which constitute the transmitted signal vector $\tilde{\mathbf{s}}$ drawn from a *white complex Gaussian codebook*; that is, the symbols $\tilde{s}_1, \tilde{s}_2, \ldots, \tilde{s}_{N_t}$ are iid complex Gaussian random variables with zero mean and common variance σ_s^2. Hence, the correlation matrix of the transmitted signal vector **s** is defined by

$$\begin{aligned} \mathbf{R}_{\mathbf{s}} &= \mathbb{E}[\mathbf{s}\mathbf{s}^{\dagger}] \\ &= \sigma_s^2 \mathbf{I}_{N_t} \end{aligned} \tag{9.107}$$

   where $\mathbf{I}_{N_t}$ is the N_t-by-N_t identity matrix.

2. $N_t \times N_r$ elements of the channel matrix **H**, which are also drawn from an ensemble of iid complex random variables with zero mean and unit variance, as shown by the complex distribution

$$h_{ik}: \quad \mathcal{N}(0, 1/\sqrt{2}) + \mathrm{j}\mathcal{N}(0, 1/\sqrt{2}) \qquad \begin{cases} i = 1, 2, \ldots, N_r \\ k = 1, 2, \ldots, N_t \end{cases} \tag{9.108}$$

   where $\mathcal{N}(...)$ denotes a real Gaussian distribution. On this basis, we find that the amplitude component h_{ik} is *Rayleigh* distributed. It is in this sense that we sometimes speak of the MIMO channel as a *rich Rayleigh scattering environment*. By the same token, we also find that the squared amplitude component, namely $|h_{ik}|^2$, is a *chi-squared random variable* with the mean

$$\mathbb{E}[|h_{ik}|^2] = 1 \qquad \text{for all } i \text{ and } k \tag{9.109}$$

   (The chi-squared distribution is discussed in Appendix A.)

3. N_r elements of the channel noise vector **w**, which are iid complex Gaussian random variables with zero mean and common variance $\sigma_{\mathbf{w}}^2$; that is, the correlation matrix of the noise vector **w** is given by

$$\begin{aligned} \mathbf{R}_{\mathbf{w}} &= \mathbb{E}[\mathbf{w}\mathbf{w}^{\dagger}] \\ &= \sigma_{\mathbf{w}}^2 \mathbf{I}_{N_r} \end{aligned} \tag{9.110}$$

   where $\mathbf{I}_{N_r}$ is the N_r-by-N_r identity matrix.

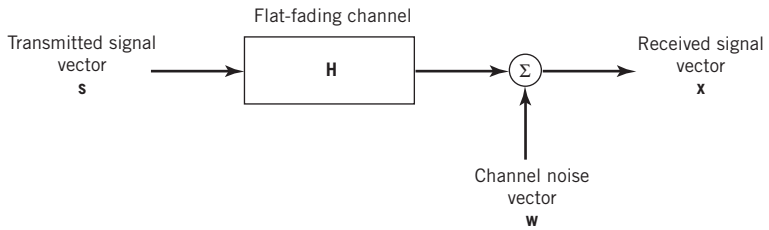


Figure 9.30 Depiction of the basic channel model of (9.106).

In light of (9.100) and the assumption that h_{ik} is a standard Gaussian random variable with zero mean and unit variance, the *average SNR* at each receiver input of the MIMO channel is given by

$$\begin{aligned}\rho &= \frac{P}{\sigma_{\mathbf{w}}^2} \\ &= \frac{N_{\mathrm{t}}\sigma_{\mathbf{s}}^2}{\sigma_{\mathbf{w}}^2}\end{aligned} \tag{9.111}$$

which is, for a prescribed noise variance $\sigma_{\mathbf{w}}^2$, *fixed* once the total transmit power P is fixed. Note also that, first, all the N_{t} transmitted signals occupy a common channel bandwidth and, second, the average SNR ρ is independent of N_{r}.

The idealized Gaussian model just described of a MIMO wireless communication system is applicable to indoor local area networks and other wireless environments, where the extent of user-terminal mobilities is limited.[11]

9.11 MIMO Capacity for Channel Known at the Receiver

With the basic complex channel model of Figure 9.30 at our disposal, we are now ready to focus attention on the primary issue of interest: the channel capacity of a MIMO wireless link. In what follows, two special cases will be considered: the first case, entitled "ergodic capacity," assumes that the MIMO channel is weakly (wide-sense) stationary and, therefore, ergodic. The second case, entitled "outage capacity," considers a nonergodic MIMO channel under the assumption of quasi-stationarity from one burst of data transmission to the next.

Ergodic Capacity

According to *Shannon's information capacity* law discussed in Chapter 5, the capacity of a *real* AWGN channel, subject to the constraint of a fixed transmit power P, is defined by

$$C = B\log_2\left(1 + \frac{P}{\sigma_w^2}\right) \text{ bits/s} \tag{9.112}$$

where B is the channel bandwidth and σ_w^2 is the noise variance measured over the bandwidth B. Given a time-invariant channel, (9.112) defines the maximum data rate that can be transmitted over the channel with an arbitrarily small probability of error being incurred as a result of the transmission. With the channel used K times for the transmission of K symbols in T seconds, the transmission capacity per unit time is K/T times the formula for C given in (9.112). Recognizing that $K = 2BT$ in accordance with the sampling theorem discussed in Chapter 6, we may express the information capacity of the AWGN channel in the equivalent form

$$C = \frac{1}{2}\log_2\left(1 + \frac{P}{\sigma_w^2}\right) \text{ bits/(s Hz)} \tag{9.113}$$

Note that one bit per second per hertz corresponds to one bit per transmission.

With wireless communications as the medium of interest, consider next the case of a *complex* flat-fading channel with the receiver having perfect knowledge of the channel state. The capacity of such a channel is given by

$$C = \mathbb{E}\left[\log_2\left(1 + \frac{|h|^2 P}{\sigma_w^2}\right)\right] \text{ bits/(s Hz)} \tag{9.114}$$

where the expectation is taken over the gain of the channel $|h|^2$ and the channel is assumed to be stationary and ergodic. In recognition of this assumption, C is commonly referred to as the *ergodic capacity* of the flat-fading channel and the channel coding is applied across fading intervals (i.e., over an "ergodic" interval of channel variation with time).

It is important to note that the scaling factor of $1/2$ is missing from the capacity formula of (9.114). The reason for this omission is that this equation refers to a complex baseband channel, whereas (9.113) refers to a real channel. The fading channel covered by (9.114) operates on a complex signal, namely a signal with in-phase and quadrature components. Therefore, such a complex channel is equivalent to two real channels with equal capacities and operating in parallel; hence the result presented in (9.114).

Equation (9.114) applies to the simple case of a *single-input, single-output (SISO) flat-fading channel*. Generalizing this formula to the case of a multiple-input, multiple-output MIMO flat-fading channel governed by the Gaussian model described in Figure 9.30, we find that the ergodic capacity of the MIMO channel is given by the following formula:[12]

$$C = \mathbb{E}\left[\log_2\left\{\frac{\det(\mathbf{R_w} + \mathbf{H}\mathbf{R_s}\mathbf{H}^\dagger)}{\det(\mathbf{R_w})}\right\}\right] \text{ bits/(s Hz)} \tag{9.115}$$

which is subject to the constraint

$$\max_{\mathbf{R_s}} \text{tr}[\mathbf{R_s}] \le P$$

where P is the constant transmit power and $\text{tr}[\cdot]$ denotes the trace of the enclosed matrix. The expectation in (9.115) is over the random channel matrix $\mathbf{H}$, and the superscript dagger notes Hermitian transposition; $\mathbf{R_s}$ and $\mathbf{R_w}$ are respectively the correlation matrices of the transmitted signal vector $\mathbf{s}$ and channel noise vector $\mathbf{w}$. A detailed derivation of (9.115) is presented in Appendix E.

In general, it is difficult to evaluate (9.115) except for a Gaussian model. In particular, substituting (9.107) and (9.110) into (9.115) and simplifying yields

$$C = \mathbb{E}\left[\log_2\left\{\det\left(\mathbf{I}_{N_r} + \frac{\sigma_s^2}{\sigma_w^2}\mathbf{H}\mathbf{H}^\dagger\right)\right\}\right] \text{ bits/(s Hz)} \tag{9.116}$$

Next, invoking the definition of the average SNR ρ introduced in (9.111), we may rewrite (9.116) in the equivalent form

$$C = \mathbb{E}\left[\log_2\left\{\det\left(\mathbf{I}_{N_r} + \frac{\rho}{N_t}\mathbf{H}\mathbf{H}^\dagger\right)\right\}\right] \text{ bits/(s Hz)}, \quad \text{for } N_t \ge N_r \tag{9.117}$$

Equation (9.117), defining the ergodic capacity of a MIMO flat-fading channel, involves the determinant of an N_r-by-N_r sum matrix (inside the braces) followed by the logarithm to base 2. It is for this reason that this equation is referred to as the *log-det capacity formula* for a Gaussian MIMO channel.

As indicated in (9.117), the log-det capacity formula therein assumes that $N_t \geq N_r$ for the matrix product $\mathbf{H}\mathbf{H}^\dagger$ to be of full rank. The alternative case, $N_r \geq N_t$ makes the N_t-by-N_t matrix product $\mathbf{H}^\dagger\mathbf{H}$ to be of full rank, in which case the *log-det capacity formula* of the MIMO link takes the form

$$C = \mathbb{E}\left[\log_2\left\{\det\left(\mathbf{I}_{N_t} + \frac{\rho}{N_r}\mathbf{H}^\dagger\mathbf{H}\right)\right\}\right] \text{ bits/(s Hz)}, \qquad N_r \geq N_t \tag{9.118}$$

where, as before, the expectation is taken over the channel matrix $\mathbf{H}$.

Despite the apparent differences between (9.117) and (9.118), they are equivalent in that either one of them applies to all $\{N_r, N_t\}$ antenna configurations. The two formulas differentiate themselves only when the full-rank issue is of concern.

Clearly, the capacity formula of (9.114), pertaining to a complex, flat-fading link with a single antenna at both ends of the link, is a special case of the log-det capacity formula. Specifically, for $N_t = N_r = 1$ (i.e., no spatial diversity), $\rho = P/\sigma_w^2$, and $\mathbf{H} = h$ (with dependence on discrete-time n suppressed, (9.116) reduces to that of (9.114).

Another insightful result that follows from the log-det capacity formula is that if $N_t = N_r = N$, then, as N approaches infinity, the capacity C defined in (9.117) grows asymptotically (at least) linearly with N; that is,

$$\lim_{N \to \infty} \frac{C}{N} \geq \text{constant} \tag{9.119}$$

In words, the asymptotic formula of (9.119) may be stated as follows:

> The ergodic capacity of a MIMO flat-fading wireless link with an equal number of transmit and receive antennas N grows roughly proportionately with N.

What this statement teaches us is that, by increasing computational complexity resulting from the use of multiple antennas at both the transmit and receive ends of a wireless link, we are able to increase the *spectral efficiency* of the link in a far greater manner than is possible by conventional means (e.g., increasing the transmit SNR). The potential for this very sizable increase in the spectral efficiency of a MIMO wireless communication system is attributed to the key parameter

$$N = \min\{N_t, N_r\}$$

which defines the *number of degrees of freedom* provided by the system.

Two Other Special Cases of the Log-Det Formula: Capacities of Receive and Transmit Diversity Links

Naturally, the log-det capacity formula for the channel capacity of an N_t, N_r wireless link includes the channel capacities of receive and transmit diversity links as special cases:

1. *Diversity-on-receive channel*. The log-det capacity formula (9.118) applies to this case. Specifically, for $N_t = 1$, the channel matrix $\mathbf{H}$ reduces to a column vector and with it (9.118) reduces to

$$C = \mathbb{E}\left[\log_2\left\{\left(1 + \rho \sum_{i=1}^{N_r} |h_i|^2\right)\right\}\right] \text{ bits/(s Hz)} \tag{9.120}$$

Compared with the channel capacity of (9.114), for an SISO fading channel with $\rho = P/\sigma_w^2$, the squared channel gain $|h|^2$ is replaced by the sum of squared magnitudes $|h_i|^2$, $i = 1, 2, \ldots, N_r$. Equation (9.120) expresses the ergodic capacity due to the *linear combination* of the receive-antenna outputs, which is designed to maximize the information contained in the N_r received signals about the transmitted signal. This is simply a restatement of the maximal-ratio combining principle discussed in Section 9.8.

2. *Diversity-on-transmit channel*. The log-det capacity formula of (9.117) applies to this second case. Specifically, for $N_r = 1$, the channel matrix $\mathbf{H}$ reduces to a row vector, and with it (9.117) reduces to

$$C = \mathbb{E}\left[\log_2\left(1 + \frac{\rho}{N_t} \sum_{k=1}^{N_t} |h_k|^2\right)\right] \text{ bits/(s Hz)} \tag{9.121}$$

where the matrix product $\mathbf{HH}^\dagger$ is replaced by the sum of squared magnitudes $|h_k|^2$, $k = 1, 2, \ldots, N_t$. Compared with case 1 on receive diversity, the capacity of the diversity-on-transmit channel is reduced because the total transmit power is being held constant, independent of the number of N_t transmit antennas.

Outage Capacity

To realize the log-det capacity formula of (9.117), the MIMO channel must be described by an ergodic process. In practice, however, the MIMO wireless channel is often nonergodic and the requirement is to operate the channel under *delay constraints*. The issue of interest is then summed up as follows:

> How much information can be transmitted across a nonergodic channel, particularly if the channel code is long enough to see just one random channel matrix?

In the situation described here, the rate of reliable information transmission (i.e., the strict Shannon-sense capacity) is zero, since for any positive rate there exists a nonzero probability that the channel would not support such a rate.

To get around this serious difficulty, the notion of *outage* is introduced into characterization of the MIMO link. (Outage was discussed previously in the context of diversity on receive in Section 9.8.) Specifically, we offer the following definition:

> The outage probability of a MIMO link is defined as the probability for which the link is in a state of outage (i.e., failure) for data transmitted across the link at a certain rate R, measured in bits per second per hertz.

To proceed on this probabilistic basis, it is customary to operate the MIMO link by transmitting data in the form of *bursts* or *frames* and invoke a *quasi-stationary model* governed by four points:

1. The burst is *long* enough to accommodate the transmission of a large number of symbols, which, in turn, permits the use of an idealized *infinite-time horizon* basic to information theory.
2. Yet, the burst is *short* enough to treat the wireless link as *quasi-stationary* during each burst; the slow variation is used to justify the assumption that the receiver has perfect knowledge of the channel state.
3. The channel matrix is permitted to change, from burst k to the next burst $k + 1$, thereby accounting for statistical variations of the link.
4. Different realizations of the transmitted signal vector $\mathbf{s}$ are drawn from a *white Gaussian codebook*; that is, the correlation matrix of $\mathbf{s}$ is defined by (9.107).

Points 1 and 4 pertain to signal transmission, whereas points 2 and 3 pertain to the MIMO channel itself.

To proceed with the evaluation of outage probability under this model, we first note that, in light of the log-det capacity formula (9.117), we may view the random variable

$$C_k = \log_2\left\{\det\left(\mathbf{I}_{N_r} + \frac{\rho}{N_t}\mathbf{H}_k\mathbf{H}_k^{\dagger}\right)\right\} \quad \text{bits/(s Hz) for burst } k \tag{9.122}$$

as the expression for a "sample realization" of the MIMO link. In other words, with the random-channel matrix $\mathbf{H}_k$ varying from one burst to the next, C_k will itself vary in a corresponding way. A consequence of this random behavior is that, occasionally, a sample drawn from the cumulative distribution function of the MIMO link results in a value for C_k that is inadequate to support reliable communication over the link. In this kind of situation the link is said to be in an *outage state*. Correspondingly, for a given transmission strategy, we define the *outage probability at rate R* as

$$P_{\text{outage}}(R) = \mathbb{P}\{C_k < R_k\} \quad \text{for some burst } k \tag{9.123}$$

Equivalently, we may write

$$P_{\text{outage}}(R) = \mathbb{P}\left\{\log_2\left\{\det\left(\mathbf{I}_{N_r} + \frac{\rho}{N_t}\mathbf{H}_k\mathbf{H}_k^{\dagger}\right)\right\} < R \quad \text{for some burst } k\right\} \tag{9.124}$$

On this basis, we may offer the following definition:

> The outage capacity of the MIMO link is the maximum bit rate that can be maintained across the link for all bursts of data transmissions (i.e., all possible channel states) for a prescribed outage probability.

By the very nature of it, the study of outage capacity can only be conducted using Monte Carlo simulation.

Channel Known at the Transmitter

The log-det capacity formula of (9.117) is based on the premise that the transmitter has *no* knowledge of the channel state. Knowledge of the channel state, however, can be made

available to the transmitter by first estimating the channel matrix **H** at the receiver and then sending this estimate to the transmitter via a *feedback channel.* In such a scenario, the capacity is optimized over the correlation matrix of the transmitted signal vector **s**, subject to the power constraint; that is, the trace of this correlation matrix is less than or equal to the constant transmit power *P.* Naturally, formulation of the log-det capacity formula of a MIMO channel for which the channel is known in both the transmitter and receiver is more challenging than when it is only known to the receiver. For details of this formulation, the reader is referred to Appendix E.

9.12 Orthogonal Frequency Division Multiplexing

In Chapter 8 we introduced the DMT method as one discrete form of multichannel modulation for signaling over band-limited channels. *Orthogonal frequency division multiplexing* (OFDM)[13] is another clearly related form of multifrequency modulation.

OFDM is particularly well suited for high data-rate transmission over delay-dispersive channels. In its own way, OFDM solves the problem by following the engineering paradigm of "divide and conquer." Specifically, a large number of closely spaced *orthogonal subcarriers (tones)* is used to support the transmission. Correspondingly, the incoming data stream is divided into a number of low data-rate *substreams*, one for each carrier, with the *subchannels* so formed operating in parallel. For the modulation process, a modulation scheme such as QPSK is used.

What we have just briefly described here is essentially the same as the procedure used in DMT modulation. In other words, the underlying mathematical theory of DMT described in Chapter 8 applies equally well to OFDM, except for the fact that the signal constellation encoder does not include the use of loading for bit allocation. In addition, two other changes have to be made in the implementation of OFDM:

1. In the transmitter, an *upconverter* is included after the digital-to-analog converter to appropriately *translate* the transmitted frequency, so as to facilitate propogation of the transmitted signal over the radio channel.
2. In the receiver, a *downconverter* is included before the analog-to-digital converter to *undo* the frequency translation that was performed by the upconverter in the transmitter.

Figure 9.31 shows the block diagram of an OFDM system, the components of which are configured to accommodate the transmission of a binary data stream at 36 Mbit/s as an illustrative example. Parts a and b of the figure depict the transmitter and receiver of the system, respectively. Specifically, pertinent values of data carrier rates as well as subcarrier frequencies at the various functional blocks are included in part a of the figure dealing with the transmitter. One last comment is in order: the front end of the transmitter and the back end of the receiver are allocated to forward error-correction encoding and decoding, respectively, for improved reliability of the system. (Error-control coding of the forward error-correction variety is discussed in Chapter 10.)

The Peak-to-Average Power Ratio Problem

A compelling practical importance of OFDM to wireless communications is attributed to the computational benefits brought about by the FFT algorithm that plays a key role in its

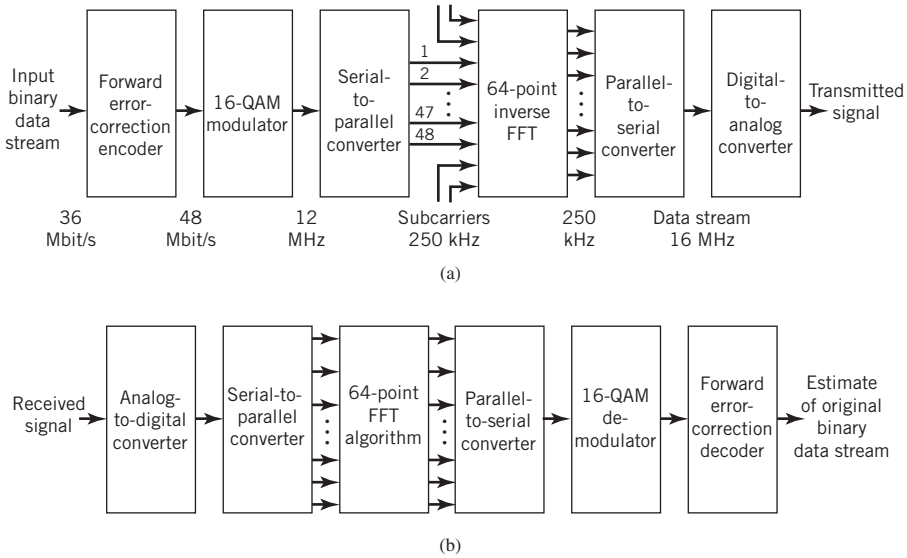


Figure 9.31 Block diagram of the typical implementation of an OFDM, illustrating the transmission of binary data at 36 Mbit/s.

implementation. However, OFDM suffers from the so-called PAPR problem. This problem arises due to the statistical probabilities of a large number of independent subchannels in the OFDM becoming superimposed on each other in some unknown fashion, thereby resulting in high peaks. For a detailed account of the PAPR problem and how to mitigate it, the reader is referred to Appendix G.

9.13 Spread Spectrum Signals

In previous sections of this chapter we described different methods for mitigating the effect of multipath interference in signaling over fading channels. In this section of the chapter, we describe another novel way of thinking about wireless communications, which is based on a class of signals called *spread spectrum signals*.[14]

A signal is said to belong to this class of signals if it satisfies the following two requirements:

1. *Spreading*. Given an information-bearing signal, spreading of the signal is accomplished in the transmitter by means of an independent *spreading signal*, such that the resulting spread spectrum signal occupies a bandwidth much larger than the bandwidth of the original information-bearing signal: the larger the better.
2. *Despreading*. Given a noisy version of the transmitted spread spectrum signal, *despreading* (i.e., recovering the original information-bearing signal) is achieved by *correlating* the received signal with a synchronized replica of the spreading signal in the receiver.

In effect, the information-bearing signal is spread (increased) in bandwidth *before* its transmission over the channel, and the received signal at the channel output is despread (i.e., decreased) in bandwidth by the same amount.

To explain the rationale of spread spectrum signals, consider, first, a scenario where there are no interfering signals at the channel output whatsoever. In this idealized scenario, an exact replica of the original information-bearing signal is reproduced at the receiver output; this recovery follows from the combined action of spreading and despreading, in that order. We may thus say that the receiver performance is *transparent* with respect to the combined spreading–despreading process.

Consider, next, a practical scenario where an additive narrowband interference is introduced at the receiver input. Since the interfering signal is introduced into the communication system *after* transmission of the information-bearing signal, its bandwidth is increased by the spreading signal in the receiver, with the result that its power spectral density is correspondingly reduced. Typically, at its output end, the receiver includes a filter whose bandwidth-occupancy matches that of the information-bearing signal. Consequently, the average power of the interfering signal is reduced, and the output SNR of the receiver is increased; hence, there is practical benefit in improved SNR to be gained from using the spread spectrum technique when there is an interfering signal (e.g., due to multipath) to deal with. Of course, this benefit is obtained at the expense of increased channel bandwidth.

Classification of Spread Spectrum Signals

Depending on how the use of spread spectrum signals is carried out, we may classify them as follows:

1. *Direct Sequence-Spread Spectrum*

   One method of spreading the bandwidth of an information-bearing signal is to use the so-called direct sequence-spread spectrum (DS-SS), wherein a *pseudo-noise* (PN) *sequence* is employed as the spreading sequence (signal). The PN sequence is a periodic binary sequence with noise-like properties, details of which are presented in Appendix J. The *baseband modulated signal*, representative of the DS-SS method, is obtained by multiplying the information-bearing signal by the PN sequence, whereby each information bit is chopped into a number of small time increments, called *chips*. The second stage of modulation is aimed at conversion of the baseband DS-SS signal into a form suitable for transmission over a wireless channel, which is accomplished by using M-ary PSK, discussed in Chapter 7. The family of spread spectrum systems so formed is referred to simply as *DS/MPSK systems,* a distinct characteristic of which is that spreading of the transmission bandwidth takes place *instantaneously.* Moreover, the signal-processing capability of these systems to combat the effect of interferers, commonly referred to as *jammers* be they friendly or unfriendly, is a function of the PN sequence length. Unfortunately, this capability is limited by physical considerations of the PN-sequence generator.

2. *Frequency Hop-Spread Spectrum*

   To overcome the physical limitations of DS/MPSK systems, we may resort to alternative methods. One such method is to force the jammer to occupy a wider

spectrum by *randomly hopping* the input data-modulated carrier from one frequency to the next. In effect, the spectrum of the transmitter signal is *spread sequentially* rather than instantaneously; the term sequentially refers to the pseudo-randomly ordered sequence of frequency hops. This second type of spread spectrum in which the carrier hops randomly from one frequency to another is called *frequency hop-spread spectrum*. A commonly used modulation format used herein is that of M-ary FSK, which was also discussed in Chapter 7. The combination of the two modulation techniques, namely frequency hopping and M-ary FSK, is referred to simply as FH/MFSK. Since frequency-hopping does not cover over the entire spread spectrum instantaneously, we are led to consider the rate at which the hops occur. In this context, we may go on to identify two basic kinds of frequency hopping, which are the converse of each other, as summarized here:

- First, *slow-frequency hopping*, in which the symbol rate of the M-ary FSK signal, denoted by R_s, is an integer multiple of the hop rate, denoted by R_h; that is, several symbols of the input data sequence are transmitted for each frequency hop.
- Second, *fast-frequency hopping*, in which the hop rate R_h is an integer multiple of the M-ary FSK symbol rate R_s; that is, the carrier frequency will change (i.e., hop) several times during the transmission of one input-data symbol.

The spread spectrum technique of the FH variety is particularly attractive for military applications. But, compared with the alternative spread spectrum technique, DS/MPSK, the commercial use of FH/MFSK is insignificant, which is especially so in regard to fast frequency hopping. The limiting factor behind this statement is the expense involved in the employment of frequency synthesizers, which are basic to the implementation of FH/MFSK systems. Accordingly, the FH/MFSK will not be considered further.

Processing Gain of the DS/BPSK

Before closing this section on spread spectrum signals, it is informative to expand on the improvement in SNR gained at the receiver output, mentioned earlier on. To this end, consider the simple case of the DS/BPSK, in which the binary PSK, representing the second stage of modulation in the transmitter, is coherent; that is, the receiver is synchronized with the transmitter in all of its features. In Problem 9.34, it is shown that the processing gain of a spread spectrum signal compared to its unspread version is

$$\text{PG} = \frac{T_b}{T_c} \tag{9.125}$$

where T_b is the bit duration and T_c is the *chip duration*. With PG expressed in decibels, in Problem 9.34 it is also shown that

$$10 \log_{10} (\text{SNR})_O = 10 \log_{10} (\text{SNR})_I + 10 \log_{10} (\text{PG}) \text{ dB} \tag{9.126}$$

where $(\text{SNR})_I$ and $(\text{SNR})_O$ are the input SNR and output SNR, respectively. Furthermore, recognizing that the ratio T_b/T_c is equal to the number of chips contained in a single bit duration, it follows that the processing gain realized by the use of DS/BPSK increases with increasing length of a single period of the PN sequence, which was emphasized previously.

9.14 Code-Division Multiple Access

Modern wireless networks are commonly of a *multiuser* type, in that the multiple communication links within the network are shared among multiple users. Specifically, each individual user is permitted to share the available radio resources (i.e., time and frequency) with other users in the network and do so in an independent manner.

Stated in another way, a *multiple access technique* permits the radio resources to be shared among multiple users seeking to communicate with each other. In the context of time and frequency domains, we recall from Chapter 1 that frequency-division multiple access (FDMA) and time-division multiple access (TDMA) techniques allocate the radio resources of a wireless channel through the use of disjointedness (i.e., orthogonality) in frequency and time, respectively. On the other hand, the *code-division multiple access* (CDMA) technique, building on spread spectrum signals and benefiting from their attributes, provides an alternative to the traditional techniques of FDMA and TDMA; it does so by not requiring the bandwidth allocation of FDMA nor the time synchronization needed in TDMA. Rather, CDMA operates on the following principle:

> The users of a common wireless channel are permitted access to the channel through the assignment of a spreading code to each individual user under the umbrella of spread spectrum modulation.

This statement is testimony to what we said in the first paragraph of Section 9.13, namely that spread spectrum signals provide a novel way of thinking about wireless communications.

To elaborate on the way in which CDMA distinguishes itself from FDMA and TDMA in graphical terms, consider Figure 9.32. Parts a and b of the figure depict the ways in which the radio resources are distributed in FDMA and TDMA, respectively. To be specific:

- In FDMA, the channel bandwidth B is divided equally among a total number of K users, with each user being allotted a subband of width B/K and having the whole time resource T at its disposal.
- In TDMA, the time resource T is divided equally among the K users, with each user having total access to the frequency resource, namely the total channel bandwidth B, but for only T/K in each time frame.

In a way, we may therefore think of FDMA and TDMA as the dual of each other.

Turning next to Figure 9.32c, we see that CDMA operates in a manner entirely different from both FDMA and TDMA. Graphically, we see that each CDMA user has full access to the entire radio resources at every point in time from one frame to the next. Nevertheless, for the full utilization of radio resources to be achievable, it is necessary that the spreading codes assigned to all the K users form an *orthogonal set*.

In other words, orthogonality is a common requirement to the FDMA, TDMA, and CDMA, each in its own specific way. However, this requirement is easier to implement practically in FDMA and TDMA than it is in CDMA.

In an ideal CDMA system, to satisfy the orthogonality requirement, the cross-correlation between any two users of the system must be zero. Correspondingly, for this

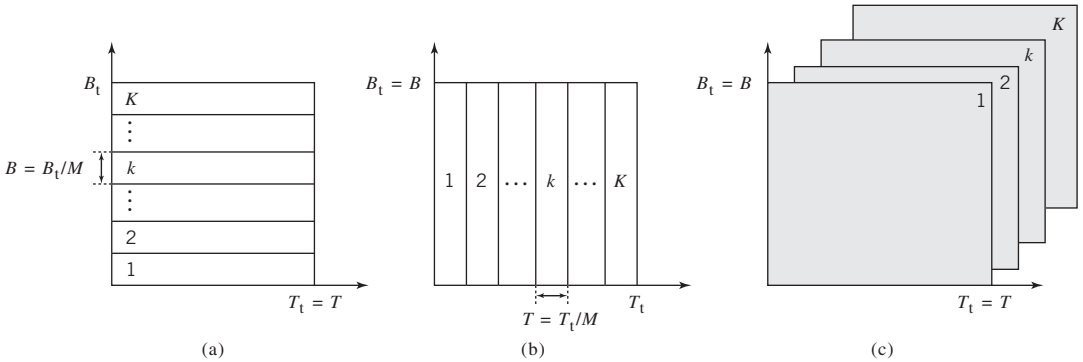


Figure 9.32 Resource distribution in (a) FDMA, (b) TDMA, and (c) CDMA. This figure shows the essence of multiple access as in Figure 1.2 with a difference: Figure 9.32 is quantitative in its description of multiple-access techniques.

ideal condition to be satisfied, we require that the cross-correlation function between the spreading sequences (codes) assigned to any two CDMA users of the system must be zero for all cyclic shifts in time. Unfortunately, ordinary PN sequences do not satisfy the orthogonality requirement because of their relatively poor cross-correlation properties.

Accordingly, we have to look to alternative spreading codes to satisfy the orthogonality requirements. Fortunately, such an endeavor is mathematically feasible, depending on whether synchrony of the CDMA receiver to its transmitter is required or not. In what follows, we describe the use of Walsh–Hadamard sequences for the synchronous case and Gold sequences for the asynchronous case.

Walsh–Hadamard Sequences

Consider the case of a CDMA system, for which synchronization among users of the system is *permissible*. Under this condition, perfect orthogonality of two spreading signals, $c_j(t)$ and $c_k(t)$, respectively assigned to users j and k for different time offsets, namely

$$R_{jk}(\tau) = \int_{-\infty}^{\infty} C_j(t)\, C_k^*(t-\tau)\, dt = 0 \qquad \text{for } j \neq k \tag{9.127}$$

reduces to

$$R_{jk}(0) = \int_{-\infty}^{\infty} C_j(t)\, C_k^*(t)\, dt = 0 \qquad \text{for } j \neq k \text{ and } \tau = 0 \tag{9.128}$$

where the asterisk denotes complex conjugation. It turns out that, for the special case described in (9.128), the orthogonality requirement can be satisfied exactly, and the resulting sequences are known as the *Walsh–Hadamard sequences (codes)*.[15]

To construct a Walsh–Hadamard sequence, we begin with a 2×2 matrix, denoted by $\mathbf{H}_2$, for which the *inner product* of its two rows (or two columns) is zero. For example, we may choose the matrix

$$\mathbf{H}_2 = \begin{bmatrix} +1 & +1 \\ +1 & -1 \end{bmatrix} \tag{9.129}$$

the two rows of which are indeed orthogonal to each other. To go on and construct a Walsh–Hadamard sequence of length 4 using $\mathbf{H}_2$, we construct the *Kronecker product* of $\mathbf{H}_2$ with itself, as shown by

$$\mathbf{H}_4 = \mathbf{H}_2 \otimes \mathbf{H}_2 \tag{9.130}$$

To explain what we mean by the Kronecker product in a generic sense, let $\mathbf{A} = \{a_{jk}\}$ and $\mathbf{B} = (b_{jk}\}$ denote $m \times m$ and $n \times n$ matrices, respectively.[16] Then, we may introduce the following rule:

> The Kronecker product of the two matrices **A** and **B** is made up of an $mn \times mn$ matrix, which is obtained from the matrix **A** by replacing its element a_{jk} in matrix **A** with the scaled matrix $a_{jk}\mathbf{B}$.

EXAMPLE 8 **Construction of Hadamard–Walsh $\mathbf{H}_4$ from $\mathbf{H}_2$**

For the example of (9.129) on matrix $\mathbf{H}_2$, applying the Kronecker product rule, we may express the $\mathbf{H}_4$ of (9.130) as follows:

$$\begin{aligned} \mathbf{H}_4 &= \begin{bmatrix} +1 \times \mathbf{H}_2 & +1 \times \mathbf{H}_2 \\ +1 \times \mathbf{H}_2 & -1 \times \mathbf{H}_2 \end{bmatrix} \\ &= \begin{bmatrix} +1 & +1 & +1 & +1 \\ +1 & -1 & +1 & -1 \\ +1 & +1 & -1 & -1 \\ +1 & -1 & -1 & +1 \end{bmatrix} \end{aligned} \tag{9.131}$$

The four rows (and columns) of $\mathbf{H}_4$ defined in (9.131) are indeed orthogonal to each other.

Carrying on in this manner, we may go on to construct the Hadamard–Walsh sequences $\mathbf{H}_6$, $\mathbf{H}_8$, and so on.

In practical terms, a synchronous CDMA system is achievable provided that a single transmitter (e.g., the base station of a cellular network) transmits individual data streams simultaneously, with each data stream being addressed to a specific CDMA user (e.g., mobile unit).

Gold Sequences

Whereas Walsh–Hadamard sequences are well suited for synchronous CDMA, *Gold sequences*, on the other hand, are well suited for applications in asynchronous CDMA;

therein, time- and phase-shifts between individual user signals, measured with respect to the base station in a cellular network, occur in a random manner; hence the adoption of asynchrony.

Gold sequences constitute a special class of maximal-length sequences, the generation of which is embodied in *Gold's theorem*, stated as follows:[17]

> Let $g_1(X)$ and $g_2(X)$ be a preferred pair of primitive polynomials of degree n whose corresponding linear feedback shift registers generate maximal-length sequences of period $2^n - 1$ and whose cross-correlation function has a magnitude less then or equal to
>
> $$2^{(n+1)/2} + 1 \qquad \text{for } n \text{ odd} \tag{9.132}$$
>
> or
>
> $$2^{(n+1)/2} + 1 \qquad \text{for } n \text{ even and } n \neq 0 \bmod 4 \tag{9.133}$$
>
> Then, the linear feedback shift register corresponding to the product polynomial $g_1(X) \times g_2(X)$ will generage $2^n + 1$ different sequences, with each sequence having a period of $2^n = 1$ and the cross-correlation between any pair of such sequences satisfying the preceding condition.

To understand Gold's theorem, we need to define what we mean by a primitive polynomial. Consider a polynomial $g(X)$ defined over a *binary field* (i.e., a finite set of two elements, 0 and 1, which is governed by the rules of binary arithmetic). The polynomial $g(X)$ is said to be an *irreducible polynomial* if it cannot be factored using any polynomials from the binary field. An irreducible polynomial $g(X)$ of degree m is said to be a *primitive polynomial* if the smallest integer m for which the polynomial $g(X)$ divides the factor $X^n + 1$ is $n = 2^m - 1$. The topic of primitive polynomials is discussed in Chapter 10 on error-control coding.

EXAMPLE 9

Correlation Properties of Gold Codes

As an illustrative example, consider Gold sequences with period $2^7 - 1 = 127$. To generate such a sequence for $n = 7$ we need a preferred pair of PN sequences that satisfy (9.132) (n odd), as shown by

$$2^{(n+1)/2} + 1 = 2^4 + 1 = 17$$

This requirement is satisfied by the Gold-sequence generator shown in Figure 9.33 that involves the modulo-2 addition of these two sequences. According to Gold's theorem, there are a total of

$$2^n + 1 = 2^7 + 1 = 129$$

sequences that satisfy (9.132). The cross-correlation between any pair of such sequences is shown in Figure 9.34, which is indeed in full accord with Gold's theorem. In particular, the magnitude of the cross-correlation is less than or equal to 17.

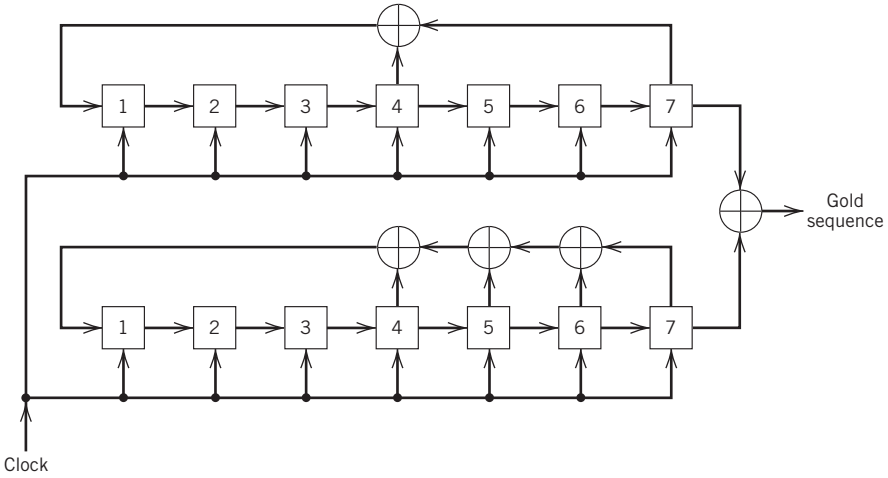


Figure 9.33 Generator for a Gold sequence of period $2^7 - 1 = 127$.

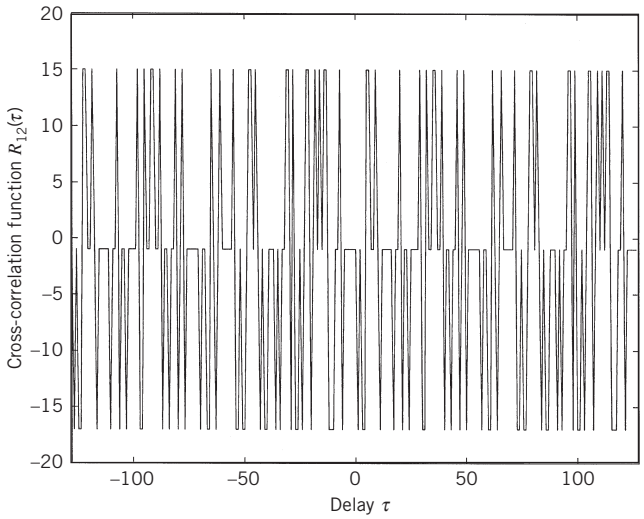


Figure 9.34 Cross-correlation function R_{12} of a pair of Gold sequences based on the two PN sequences [7,4] and [7,6,5,4].

9.15 The RAKE Receiver and Multipath Diversity

A discussion of wireless communications using CDMA would be incomplete without a description of the *RAKE receiver*.[18] The RAKE receiver was originally developed in the 1950s as a *diversity* receiver designed expressly to equalize the effect of multipath. First, and foremost, it is recognized that useful information about the transmitted signal is contained in the multipath component of the received signal. Thus, taking the viewpoint

that multipath may be approximated as a linear combination of differently delayed echoes, as shown in the maximal ratio combiner of Figure 9.21, the RAKE receiver seeks to combat the effect of multipath by using a correlation method to detect the echo signals individually and then adding them algebraically. In this way, intersymbol interference due to multipath is dealt with by reinserting different delays into the detected echoes so that they perform a constructive rather than destructive role.

Figure 9.35 shows the basic idea behind the RAKE receiver. The receiver consists of a number of *correlators* connected in parallel and operating in a synchronous fashion with each other. Each correlator has two inputs: (1) a delayed version of the received signal and (2) a replica of the PN sequence used as the spreading code to generate the spread spectrum-modulated signal at the transmitter. In effect, the PN sequence acts as a *reference signal*. Let the nominal bandwidth of the PN sequence be denoted as $W = 1/T_c$, where T_c is the chip duration. From the discussion on PN sequences presented in Appendix J, we find that the autocorrelation function of a PN sequence has a single peak of width $1/W$, and it disappears toward zero elsewhere inside one period of the PN sequence (i.e., one symbol period). Thus, we need only make the bandwidth W of the PN sequence sufficiently large to identify the significant echoes in the received signal. To be sure that the correlator outputs all add constructively, two other operations are performed in the receiver by the functional blocks labeled "phase and gain adjustors":

1. An appropriate delay is introduced into each correlator output, so that the phase angles of the correlator outputs are in agreement with each other.
2. The correlator outputs are weighted so that the correlators responding to strong paths in the multipath environment have their contributions accentuated, while the correlators not synchronizing with any significant path are correspondingly suppressed.

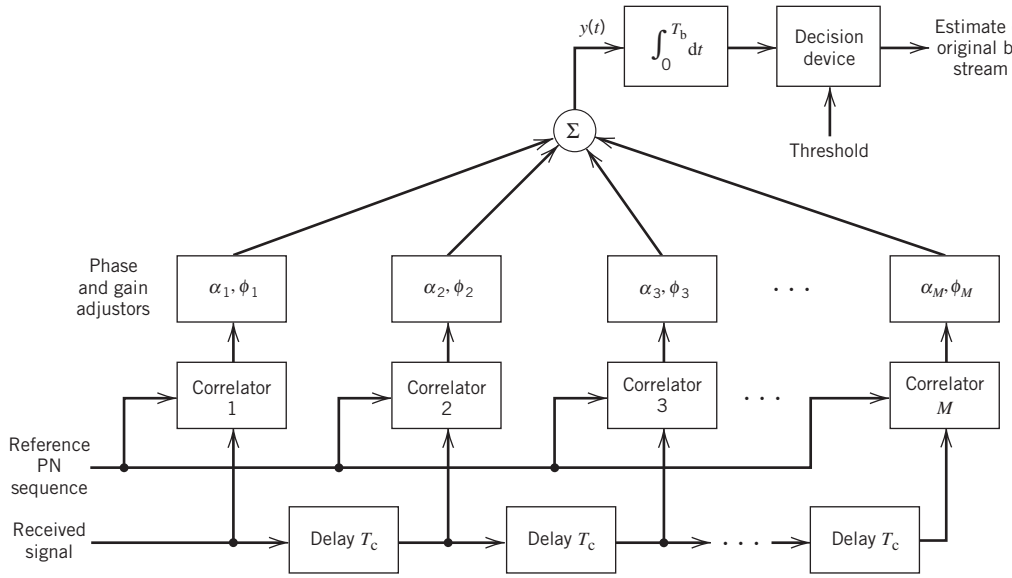


Figure 9.35 Block diagram of the RAKE receiver for CDMA over multipath channels.

The weighting coefficients a_k are computed in accordance with the *maximal ratio combining principle*, discussed in Section 9.8. Specifically, we recall that the SNR of a weighted sum, where each element of the sum consists of a signal plus additive noise of fixed power, is maximized when the amplitude weighting is performed in proportion to the pertinent signal strength. That is, the linear combiner output is

$$y(t) = \sum_{k=1}^{M} a_k z_k(t) \tag{9.134}$$

where $z_k(t)$ is the phase-compensated output of the kth correlator and M is the number of correlators in the receiver. Provided that we use enough correlators in the receiver to span a region of delays sufficiently wide to encompass all the significant echoes that are likely to occur in the multipath environment, the output $y(t)$ behaves essentially as though there was a single propagation path between the transmitter and receiver rather than a series of multiple paths spread in time.

To simplify the presentation, the receiver of Figure 9.35 assumes the use of binary PSK in performing spread spectrum modulation at the transmitter. Thus, the final operation performed in Figure 9.35 is that of integrating the linear combiner output $y(t)$ over the bit duration T_b and then determining whether binary symbol 1 or 0 was transmitted in that bit interval.

The RAKE receiver derives its name from the fact that the bank of parallel correlators has an appearance similar to the fingers of a rake; see Figure 9.36. Because spread spectrum modulation is basic to the operation of CDMA wireless communications, it is natural for the RAKE receiver to be central to the design of the receiver used in this type of multiuser radio communication.

Figure 9.36
Picture of a rake, symbolizing the bank of correlators.

9.16 Summary and Discussion

In this chapter we discussed the topic of signaling over fading channels, which is at the heart of wireless communications. There are three major sources of signal degradation in wireless communications:

- co-channel interference,
- fading, and
- delay spread.

The latter two are by-products of the multipath phenomenon. A common characteristic of these channel impairments is that they are all *signal-dependent phenomena*. As it is with intersymbol interference that characterizes signaling over band-limited channels discussed in Chapter 8, the degrading effects of interference and multipath in wireless communications cannot be combated by simply increasing the transmitted signal, which is what is done when noise is the only source of channel impairment as discussed in Chapter 7.

To combat the effects of multipath and interference, we require the use of specialized techniques that are tailor-made for wireless communications. These specialized techniques include *space diversity*, which occupied much of the material presented in this chapter.

We discussed different forms of space diversity, the main idea behind which is that two or more propagation paths connecting the receiver to the transmitter are better than a single propagation path. In historical terms, the first form of space diversity used to mitigate the multipath fading problem was that of *receive diversity*, involving a single transmit antenna and multiple receive antennas. Under receive diversity, we discussed the selection combiner, maximal-ratio combiner, and equal-gain combiner:

- The selection combiner is the simplest form of receive diversity. It operates on the principle that it is possible to select, among N_r receive-diversity branches, a particular branch with the largest output SNR; the branch so selected defines the desired received signal.
- The maximal-ratio combiner is more powerful than the selection combiner by virtue of the fact that it exploits the full information content of all the N_r receive-diversity branches about the transmitted signal of interest; it is characterized by a set of N_r receive-complex weighting factors that are chosen to maximize the output SNR of the combiner.
- The equal-gain combiner is a simplified version of the maximal-ratio combiner.

We also discussed diversity-on-transmit techniques, which may be viewed as the dual of their respective diversity-on-receive techniques. Much of the discussion here focused on the Alamouti code, which is simple to design, yet powerful in performance, in that it realizes a two-level diversity gain: in other terms of performance, the Alamouti code is equivalent to a linear diversity-on-receive system with a single antenna and two receive antennas.

By far, the most powerful form of space diversity is the use of multiple antennas at both the transmit and receive ends of the wireless link. The resulting configuration is referred to as a MIMO wireless communication system, which includes the receive diversity and transmit diversity as special cases. The novel feature of the MIMO system is that, in a rich scattering environment, it can provide a high spectral efficiency, which may be simply explained as follows. The signals transmitted simultaneously by the transmit antennas arrive at the input of each receive antenna in an uncorrelated manner due to the rich scattering mechanism of the channel. The net result is a spectacular increase in the spectral efficiency of the wireless link. Most importantly, the spectral efficiency increases roughly linearly with the number of transmit or receive antennas, whichever is the smaller one of the two. This important result assumes that the receiver has knowledge of the channel state. The spectral efficiency of the MIMO system can be further enhanced by including a feedback channel from the transmitter to the receiver, whereby the channel state is also made available to the transmitter and with it the transmitter is enabled to exercise control over the transmitted signal.

Multiple Access Considerations

An issue of paramount practical importance in wireless communications is that of multiple access to the wireless channel, in the context of which the following two approaches are considered to be the dominant ones:

1. *Orthogonal frequency division multiple access* (*OFDMA*), which is the multi-user version of OFDM that was discussed in Section 9.12. In OFDMA multiple access is accomplished through the assignment of subchannels (subcarriers) to individual users.

Naturally, OFDMA inherits the distinctive features of OFDM. In particular, OFDMA is well suited for high data-rate transmissions over delay-dispersive channels, realized by exploiting the principle of "divide and conquer." Accordingly, OFDMA is computationally efficient in using the FFT algorithm. Moreover, OFDMA lends itself to the combined use of MIMO, hence the ability to improve spectral efficiency and take advantage of channel flexibility.

2. *Code-division multiple access* (*CDMA*), which distinguishes itself by exploiting the underlying principle of spread spectrum signals, discussed in Section 9.13. To be specific, through the combined process of spectrum spreading in the transmitter and corresponding spectrum despreading in the receiver, a certain amount of processing gain is obtained, hence the ability of CDMA users to occupy the same channel bandwidth. Moreover, CDMA provides a flexible procedure for the allocation of resources (i.e., PN codes) among a multiplicity of active users. Last but by no means least, in using the RAKE, viewed as an adaptive TDL filter, CDMA is enabled to match the receiver input to the channel output by adjusting tap delays as well as tap weights, thereby enhancing receiver performance in the presence of multipath.

To conclude, OFDMA and CDMA provide two different approaches for the multiple access of active users to wireless channels, each one of which builds on its own distinctive features.

Problems

Effect of Flat Fading on the BER of Digital Communications Receivers

9.1 Derive the BER formulas listed in the right-hand side of Table 9.2 for the following signaling schemes over flat fading channels:

a. Binary PSK using coherent detection
b. Binary FSK using coherent detection
c. Binary DPSK
d. Binary FSK using noncoherent detection

9.2 Using the formulas derived in Problem 9.1, plot the BER charts for the schemes described therein.

Selective Channels

9.3 Consider a time-selective channel, for which the modulated received signal is defined by

$$x(t) = \sum_{n=1}^{N} \alpha_n(t) m(t) \cos(2\pi f_c t + \phi(t) + \sigma_n(t))$$

where $m(t)$ is the message signal, $\phi(t)$ is the result of angle modulation; the amplitude $\alpha_n(t)$ and phase $\sigma_n(t)$ are contributed by the nth path, where $n = 1, 2, \ldots, N$.

a. Using complex notation, show that the received signal is described as follows:

$$\tilde{x}(t) = \tilde{\alpha}(t)\tilde{s}(t)$$

where

$$\tilde{\alpha}(t) = \sum_{n=1}^{N} \tilde{\alpha}_n(t)$$

What is the formula for $\tilde{s}(t)$?

b. Show that the delay-spread function of the multipath channel is described by

$$\tilde{h}(\tau;t) = \tilde{\alpha}(t)\delta(\tau)$$

where $\delta(\tau)$ is the Dirac delta function in the τ-domain. Hence, justify the statement that the channel described in this problem is a *time-selective channel*.

c. Let $S_{\tilde{\alpha}}(f)$ and $S_{\tilde{s}}(f)$ denote the Fourier transforms of $\tilde{\alpha}(t)$ and $\tilde{s}(t)$, respectively. What then is the Fourier transform of $\tilde{x}(t)$?

d. Using the result of part c, justify the statement that the multipath channel described herein can be approximately frequency-flat. What is the condition that would satisfy this description?

9.4 In this problem, we consider a multipath channel embodying large-scale effects. Specifically, using complex notation, the received signal at the channel output is described by

$$\tilde{x}(t) = \sum_{l=1}^{L} \tilde{\alpha}_l \tilde{s}(t - \tau_l)$$

where $\tilde{\alpha}_l$ and τ_l denote the amplitude and time delay associated with the lth path in the channel for $l = 1, 2, \ldots, L$. Note that $\tilde{\alpha}_l$ is assumed to be constant for all l.

a. Show that the delay-spread function of the channel is described by

$$\tilde{h}(\tau;t) = \sum_{l=1}^{L} \tilde{\alpha}_l \delta(\tau - \tau_l)$$

where $\delta(\tau)$ is the Dirac delta function expressed in the τ-domain.

b. This channel is said to be time-nonselective. Why?

c. The channel does exhibit a frequency-dependent behavior. To illustrate this behavior, consider the following delay-spread function:

$$\tilde{h}(\tau;t) = \delta(\tau) + \tilde{\alpha}_2 \delta(\tau - \tau_2)$$

where τ_2 is the time delay produced by the second path in the channel. Plot the magnitude (amplitude) response of the channel for the following specifications:

i. $\tilde{\alpha}_2 = 0.5$

ii. $\tilde{\alpha}_2 = \mathrm{j}/2$

iii. $\tilde{\alpha}_2 = -\mathrm{j}$

where $\mathrm{j} = \sqrt{-1}$. Comment on your results.

9.5 Expanding on the multipath channel considered in Problem 9.4, a more interesting case is characterized by the scenario in which the received signal at the channel output is described as follows:

$$\tilde{x}(t) = \sum_{l=1}^{L} \tilde{\alpha}_l(t) \tilde{s}(t - \tau_l(t))$$

where the amplitude $\tilde{\alpha}_l(t)$ and time delay $\tau_l(t)$ for the lth path are both time dependent for $l = 1, 2, \ldots, L$.

a. Show that the delay-spread function of the multipath channel described herein is given by

$$\tilde{h}(\tau;t) = \sum_{l=1}^{L} \tilde{\alpha}_l(t) \delta(\tau - \tau_l(t))$$

where $\delta(\tau)$ is the Dirac delta function in the τ-domain. This channel is said to exhibit both large- and small-scale effects. Why?

b. The channel is also said to be both time selective and frequency selective. Why?

c. To illustrate the point made under b, consider the following channel description:

$$\tilde{h}(\tau;t) = \tilde{\alpha}_1(t)\delta(\tau) + \tilde{\alpha}_2(t)\delta(t-\tau_2)$$

where $\tilde{\alpha}_1(t)$ and $\tilde{\alpha}_2(t)$ are both Rayleigh processes.

For selected $\tilde{\alpha}_1(t)$, $\tilde{\alpha}_2(t)$ and τ_2, do the following:

i. At each time $t = 0$, compute the Fourier transform of $\tilde{h}(\tau;t)$.

ii. Hence, plot the magnitude spectrum of the channel, that is, $|\tilde{H}(f;t)|$, expressed as a function of both time t and frequency f.

Comment on the results so obtained.

9.6 Consider a multipath channel where the delay-spread function is described by

$$\tilde{h}(\tau;t) = \sum_{l=1}^{L} \tilde{\alpha}_l(t)\delta(\tau-\tau_l)$$

where the scattering processes attributed to the time-varying amplitude $\tilde{\alpha}_l(t)$ and fixed delay τ_l are uncorrelated for $l = 1, 2, \ldots, L$.

a. Determine the correlation function of the channel, namely $R_{\tilde{h}}(\tau_1, t_1; \tau_2, t_2)$.

b. With a Jakes model for the scattering process described in (9.12), find the corresponding formula for the correlation function of the channel under part a of the problem.

c. Hence, justify the statement that the multipath channel described in this problem fits a WSSUS model.

9.7 Revisit the Jakes model for a fast fading channel described in (9.12). Let the coherence time be defined as that range of values Δt over which the correlation function defined in (9.12) is greater than 0.5.

For some prescribed maximum Doppler shift $\nu_{\max}$, find the coherence time of the channel.

9.8 Consider a multipath channel for which the delay-spread function is given by

$$\tilde{h}(\tau;t) = \sum_{l=1}^{L} \tilde{\alpha}_l(t)\delta(t-\tau_l)$$

where the amplitude $\tilde{\alpha}_l(t)$ is time varying but the time delay τ_l is fixed. As in Problem 9.4, the scattering processes are described by the Jakes model in (9.12). Determine the power-delay profile of the channel, $P_{\tilde{h}}(\tau)$.

9.9 In real-life situations, the wireless channel is *nonstationary* due to the presence of moving objects of different kinds and other physical elements that can significantly affect radio propagation. Naturally, different types of wireless channels have different degrees of nonstationarity.

Even though many wireless communication channels are indeed highly nonstationary, the WSSUS model described in Section 9.4 still provides a reasonably accurate account of the statistical characteristics of the channel. Elaborate on this statement.

"Space Diversity-on-Receive" Systems

9.10 Following the material presented on Rayleigh fading in Chapter 4, derive the probability density function of (9.64).

9.11 A receive-diversity system uses a selection combiner with two diversity paths. The outage occurs when the instantaneous SNR γ drops below $0.25\gamma_{\text{av}}$, where γ_{av} is the average SNR.

Determine the probability of outage experienced by the receiver.

9.12 The average SNR in a selection combiner is 20 dB. Compute the probability that the instantaneous SNR of the selection combiner drops below $\gamma = 10$ dB for the following number of receive antennas:

a. $N_r = 1$
b. $N_r = 2$
c. $N_r = 3$
d. $N_r = 4$.

Comment on your results.

9.13 Repeat Problem 9.12 for $\gamma = 15$ dB.

9.14 In Section 9.8 we derived the optimum values of (9.75) for complex weighting factors of the maximal-ratio combiner using the Cauchy–Schwartz inequality.

This problem addresses the same issue, but this time we use the standard maximization procedure. To simplify matters, the number of diversity paths N_r is restricted to two, with the complex weighting parameters denoted by a_1 and a_2. Let

$$a_k = x_k + jy_k, \qquad k = 1, 2$$

The complex derivative with respect to a_k is defined by

$$\frac{\partial}{\partial a_k^*} = \frac{1}{2}\left(\frac{\partial}{\partial x_k} + j\frac{\partial}{\partial y_k}\right), \qquad k = 1, 2$$

Applying this formula to the combiner's output SNR γ_c of (9.71), derive the optimum γ_{mrc} in (9.75).

9.15 As discussed in Section 9.8, an *equal-gain combiner* is a special form of the maximal-ratio combiner for which the weighting factors are all equal. For convenience of presentation, the weighting parameters are set to unity.

Assuming that the instantaneous SNR γ is small compared with the average SNR γ_{av}, derive an approximate formula for the probability density function of the random variable Γ represented by the sample γ.

9.16 Compare the performances of the following linear "diversity-on-receive" techniques:

a. Selection combiner.
b. Maximal-ratio combiner.
c. Equal-gain combiner.

Base the comparison on signal-to-noise improvement, expressed in decibels for the following number of diversity branches: $N_r = 2, 3, 4, 5, 6$.

9.17 Show that the maximum-likelihood decision rule for the maximal-ratio combiner may be formulated in the following two equivalent forms:

a. If

$$[(\alpha_1^2 + \alpha_2^2)|s_i|^2 - y_1 s_i^* - y_1^* s_i] < [(\alpha_1^2 + \alpha_2^2)|s_k|^2 - y_1 s_k^* - y_1 s_k], \qquad k \neq i$$

then choose symbol s_i over s_k.

b. If, by the same token,

$$[(\alpha_1^2 + \alpha_2^2 - 1)|s_i|^2 + d^2(y_1, s_i)] < [(\alpha_1^2 + \alpha_2^2 - 1)|s_k|^2 + d^2(y_1, s_k)], \qquad k \neq i$$

then choose symbol s_i over s_k. Here, $d^2(y_1,s_i)$ denotes the squared Euclidean distance between the signal points y_1 and s_i.

9.18 It may be argued that, in a rather loose sense, transmit-diversity and receive-diversity antenna configurations are the dual of each other, as illustrated in Figure P9.18.

a. Taking a general viewpoint, justify the mathematical basis for this duality.

b. However, we may cite the example of frequency-division diplexing (FDD) for which, in a strict sense, we find that the duality depicted in Figure P9.18 is violated. How is it possible for the violation to arise in this example?

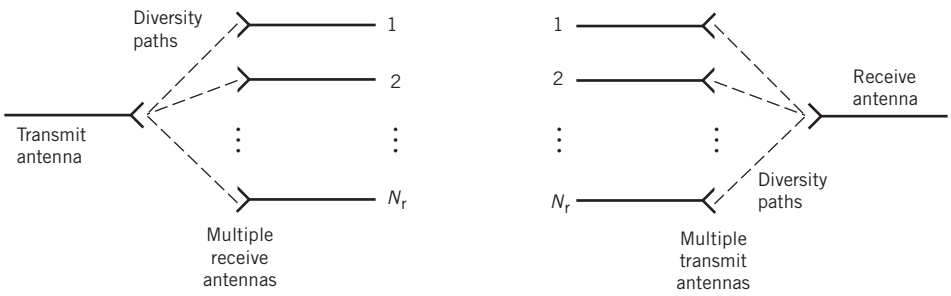


Figure P9.18

"Space Diversity-on-Transmit" Systems

9.19 Show that the two-by-two channel matrix in (9.88), defined in terms of the multiplicative fading factors $\alpha_1 e^{j\theta_1}$ and $\alpha_2 e^{j\theta_2}$, is a unitary matrix, as shown by

$$\begin{bmatrix} \alpha_1 e^{j\theta_1} & \alpha_2 e^{j\theta_2} \\ \alpha_2 e^{-j\theta_2} & -\alpha_1 e^{-j\theta_1} \end{bmatrix}^{\dagger} \begin{bmatrix} \alpha_1 e^{j\theta_1} & \alpha_2 e^{j\theta_2} \\ \alpha_2 e^{-j\theta_2} & -\alpha_1 e^{-j\theta_1} \end{bmatrix} = (\alpha_1^2 + \alpha_2^2) \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

9.20 Derive the formula for the average probability of symbol error incurred by the Alamouti code.

9.21 Figure P9.22 shows the extension of orthogonal space–time codes to the Alamouti code, using two antennas on both transmit and receive. The sequence of signal encoding and transmissions is identical to that of the single-receiver case of Figure 9.18. Part a of the table below defines the channels between the transmit and receive antennas. Part b of the table defines the outputs of the receive antennas at times t' and $t' + T$, where T is the symbol duration.

a. Derive expressions for the received signals $\tilde{x}_1, \tilde{x}_2, \tilde{x}_3$, and $\tilde{x}_4$, including the respective additive noise components expressed in terms of the transmitted symbols.

b. Derive expressions for the line of combined outputs in terms of the received signals.

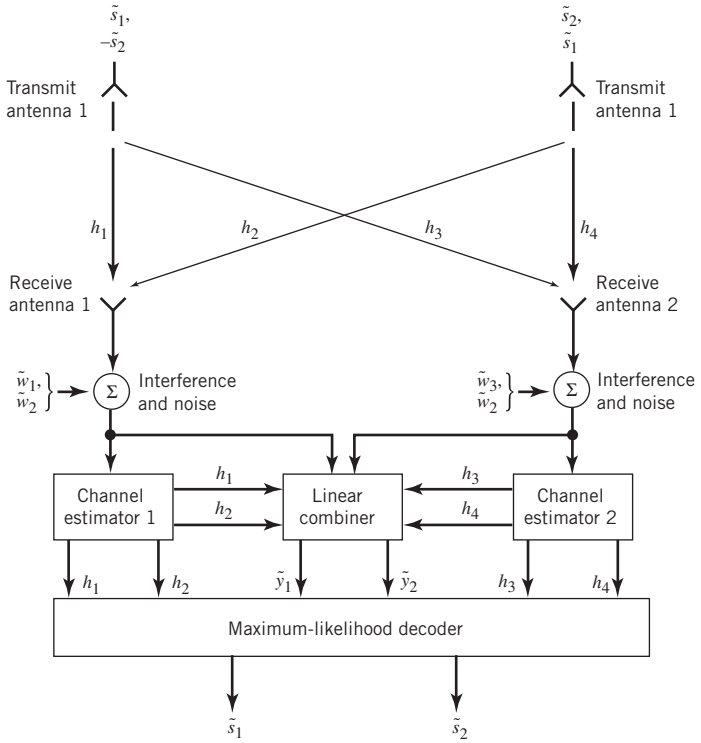
c. Derive the maximum-likelihood decision rule for the estimates $\tilde{s}_1$ and $\tilde{s}_2$.

		Receive antenna 1	Receive antenna 2
a.	Transmit antenna 1	h_1	h_3
	Transmit antenna 2	h_2	h_4
b.	Time t'	$\tilde{x}_1$	$\tilde{x}_3$
	Time $t' + T$	$\tilde{x}_2$	$\tilde{x}_4$

9.22 This problem explores a new interpretation of the Alamouti code. Let

$$\tilde{s}_i = s_i^{(1)} + j s_i^{(2)}, \qquad i = 1, 2$$

Figure P9.22



where $s_i^{(1)}$ and $s_i^{(2)}$ are both real numbers. The complex entry $\tilde{s}_i$ in the 2-by-2 Alamouti code is represented by the 2-by-2 real orthogonal matrix

$$\begin{bmatrix} s_i^{(1)} & s_i^{(2)} \\ -s_i^{(2)} & s_i^{(1)} \end{bmatrix}, \qquad i = 1, 2$$

Likewise, the complex-conjugated entry $\tilde{s}_i^*$ is represented by the 2-by-2 real orthogonal matrix

$$\begin{bmatrix} s_i^{(1)} & -s_i^{(1)} \\ s_i^{(2)} & s_i^{(2)} \end{bmatrix}, \qquad i = 1, 2$$

a. Show that the 2-by-2 complex Alamouti code $\mathbf{S}$ is equivalent to the 4-by-4 *real* transmission matrix

$$\mathbf{S}_4 = \left[\begin{array}{cc:cc} s_1^{(1)} & s_1^{(2)} & s_2^{(1)} & s_2^{(2)} \\ -s_1^{(2)} & s_1^{(1)} & -s_2^{(2)} & s_2^{(1)} \\ \hdashline -s_2^{(1)} & s_2^{(2)} & s_1^{(1)} & -s_1^{(2)} \\ -s_2^{(2)} & -s_2^{(1)} & s_1^{(2)} & s_1^{(1)} \end{array}\right]$$

b. Show that $\mathbf{S}_4$ is an orthogonal matrix.

c. What is the advantage of the complex code $\mathbf{S}$ over the real code $\mathbf{S}_4$?

9.23 For two transmit antennas and simple receive antenna, the Alamouti code is said to be the only optimal space–time block. Using the log-det formula of (9.117), justify this statement.

9.24 Show that the channel capacity of the Alamouti code is equal to the sum of the channel capacities of two SISO systems with each one of them operating at half the original bit rate.

MIMO Wireless Communications

9.25 Show that, at high SNRs, the capacity gain of a MIMO wireless communication system with the channel state known to the receiver is $N = \min\{N_t, N_r\}$ bits per second per hertz for every 3 dB increase in SNR.

9.26 To calculate the outage probability of MIMO systems, we use the complementary cumulative distribution function of the random channel matrix $\mathbf{H}$ rather than the cumulative probability function itself.

Explain this rationale for calculating the outage probability.

9.27 Equation (9.120) defines the formula for the channel capacity of diversity-on-receive channel.

In Section 9.8 we pointed out that the selection combiner is a special case of the maximal-ratio combiner. Using (9.120), formulate an expression for the channel capacity of wireless diversity using the selection combiner.

9.28 For the special case of a MIMO system having $N_t = N_r = N$, show that the ergodic capacity of the system scales linearly, rather than logarithmically, with increasing SNR as N approaches infinity.

9.29 In this problem we continue with the solution to Problem 9.28, namely

$$C \to \left(\frac{\lambda_{av}}{\log_e 2}\right)\rho \qquad \text{as } N \to \infty$$

where $N_t = N_r = N$ and λ_{av} is the average eigenvalue of the matrix produced $\mathbf{H}\mathbf{H}^\dagger = \mathbf{H}^\dagger\mathbf{H}$. What is the value of the constant?

a. Justify the asymptotic result given in (9.119); that is,

$$\frac{C}{N} \geq \text{constant}$$

b. What conclusion can you draw from this asymptotic result?

9.30 Suppose that an additive, temporally stationary, Gaussian interference $\mathbf{v}(t)$ corrupts the basic complex channel model of (9.105). The interference $\mathbf{v}(t)$ has zero mean and correlation matrix $\mathbf{R}_v$. Evaluate the effect of the interference $\mathbf{v}(t)$ on the ergodic capacity of the MIMO link.

9.31 Consider a MIMO link for which the channel may be considered to be essentially "constant for k users of the channel."

a. Starting with the basic channel model of (9.105), formulate the input–output relationship of this link with the input being described by the N_r-by-k matrix

$$\mathbf{S} = [\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_k]$$

b. How is the log-det capacity formula of the link correspondingly modified?

9.32 In a MIMO channel, the ability to exploit space-division multiple-access techniques for spectrally efficient wireless communications is determined by the rank of the complex channel matrix $\mathbf{H}$. (The rank of a matrix is defined by the number of independent columns in the matrix.) For a given (N_t, N_r) antenna configuration, it is desirable that the rank of $\mathbf{H}$ equal the minimum one of N_t transmit and N_r receive antennas, for it is only then that we are able to exploit the full potential of the MIMO antenna

configuration. Under special conditions, however, the rank of the channel matrix **H** is reduced to unity, in which case the scattering (fading) energy flow across the MIMO link is effectively confined to a very narrow pipe, and with it, the channel capacity is severely degraded.

Under the special conditions just described, a physical phenomenon known as the *keyhole channel* or *pinhole channel* is known to arise. Using a propagation layout of the MIMO link, describe how this phenomenon can be explained.

OFDMA and CDMA

9.33 Parts a and b of Figure 9.31 show the block diagrams of the transmitter and receiver of an OFDM system, formulated on the basis of digital signal processing. It is informative to construct an analog interpretation of the OFDM system, which is the objective of this problem.

a. Construct the analog interpretations of parts a and b in Figure 9.31.

b. With this construction at hand, compare the advantages and disadvantages of the digital and analog implementations of OFDM.

9.34 Figure P9.34 depicts the *model* of a DS/BPSK system, where the order of spectrum spreading and BPSK in the actual system has been interchanged; this is feasible because both operations are linear. For system analysis, we build on signal-space theoretic ideas of Chapter 7, using this model and assuming the presence of a jammer at the receiver input. Thus, whereas signal-space representation of the transmitted signal, $x(t)$, is one-dimensional, that of the jammer, $j(t)$, is two-dimensional.

a. Derive the processing gain formula of (9.125).

b. Next, ignoring the benefit gained from coherent detection, derive the SNR formula of (9.126).

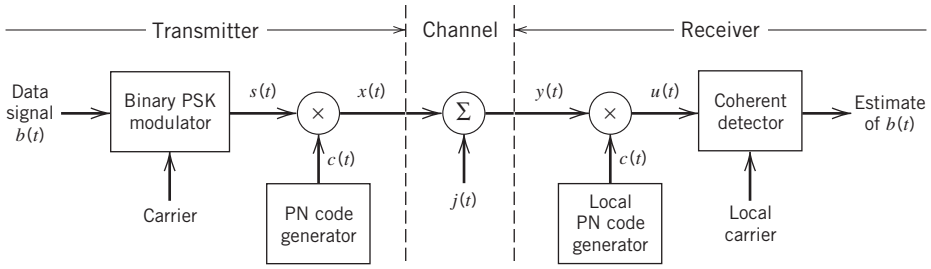


Figure P9.34

Notes

1. Local propagation effects are discussed in Chapter 1 of the classic book by Jakes (1974). For a comprehensive treatment of this subject, see the books by Parsons (2000) and Molisch (2011).

2. Bessel functions are discussed in Appendix C.

3. To be precise, we should use the terminology "autocorrelation" function rather then "correlation" function as we did in Section 9.3. However, to be consistent with the literature, hereafter we use the terminology "correlation function" for the sake of simplicity.

4. On the basis of many measurements, the power-delay profile may be approximated by the one-sided exponential functions (Molisch, 2011):

$$P_{\tilde{h}}(\tau)\begin{cases} = \exp(-\tau/\sigma_\tau), & \text{for } \tau \geq 0 \\ = 0, & \text{otherwise} \end{cases}$$

For a more generic model, the power-delay profile is viewed as the sum of several one-sided exponential functions representing multiple clusters of interacting objects, as shown by

$$P_{\tilde{h}}(\tau) = \sum_i \left(\frac{P_i}{\sigma_{\tau,i}} \right) P_{\tilde{h}}(\tau - \tau_{0,i})$$

where P_i, $\tau_{0,i}$, and $\sigma_{\tau,i}$ are respectively the power, delay, and delay spread of the ith cluster.

5. The approximate approach described in Section 9.5 follows Van Trees (1971).

6. The complex *tap-coefficient* $\tilde{c}_n(t)$ is also referred to as the tap-gain or tap-weight.

7. The chi-squared distribution with two degrees of freedom is described in Appendix A.

8. The term "maximal-ratio combiner" was coined in a classic paper on linear diversity combining techniques by Brennan (1959).

9. The three-point exposition presented in this section on maximal-ratio combining follows the chapter by Stein in Schwartz *et al.* (1966: 653–654).

10. The idea of MIMO for wireless communications was first described in the literature by Foschini (1996). In the same year, Teletar (1996) derived the capacity of multi-antenna Gaussian channels in a technical report.

11. As a result of experimental measurements, the model is known to be decidedly non-Gaussian owing to the *impulsive nature* of human-made electromagnetic interference and natural noise.

12. Detailed derivation of the ergodic capacity in (9.115) is presented in Appendix E.

13. The idea of OFDM has a long history, dating back to Chang (1966). Then, Weinstein and Ebert (1971) used the FFT algorithm and guard intervals for the first digital implementation of OFDM. The first use of OFDM for mobile communications is credited to Cemini (1985).

In the meantime, OFDM has developed into an indispensable tool for broadband wireless communications and digital audio broadcasting.

14. The literature on spread spectrum communications is enormous. For classic papers on spread spectrum communications, see the following two:

- The paper by Scholtz (1982) describes the origins of spread spectrum communications.
- The paper by Pickholtz, *et al.* (1982) addresses the fundamentals of spread spectrum communications.

15. The Walsh–Hadamard sequences (codes) are named in honor of two pioneering contributions:

- Joseph L. Walsh (1923) for finding a new set of orthogonal functions with entries ± 1 .
- Jacques Hadamard (1893) for finding a new set of square matrices also with entries ± 1 , which had all their rows (and columns) orthogonal.

For more detailed treatments of these two papers, see Harmuth (1970), and Seberry and Yamada (1992), respectively.

16. To be rigorous mathematically, we should speak of the matrices **A** and **B** to be over the Galois field, GF(2). To explain, for any *prime p*, there exists a *finite field* of p elements, denoted by GF(P). For any positive integer b, we may expand the finite field GF(p) to a field of p^b elements, which is called an *extension field* of GF(p) and denoted by GF(p^b). Finite fields are also called *Galois fields* in honor of their discoverer.

Thus, for the example of (9.129), we have a Galois field of $p = 2$ and thus write GF(2). Correspondingly, for the $\mathbf{H}_4$ in (9.130) we have the Galois field GF(2^2) = GF(4)

17. The original papers on Gold sequences are Gold (1967, 1968). A detailed discussion of Gold sequences is presented in Holmes (1982).

18. The classic paper on the RAKE receiver is due to Price and Green (1958). For a good treatment of the RAKE receiver, more detailed than that presented in Section 9.15, see Chapter 5 in the book by Haykin and Mohr (2005). For application of the RAKE receiver in CDMA, see the book by Viterbi (1995).