CHAPTER 5

FADING MITIGATION IN WIRELESS SYSTEMS

5.0 INTRODUCTION

One of the major consequences of fading is the random nature of the wireless signal. Random fluctuations in the received signal power arise primarily from the Rayleigh fading experienced by the signal due to the multipath effect (Pric 1958, Beck 1962, Casa 1990, Stei 1987, Mons 1980, Akki 1994, Ho 1999). Shadowing or lognormal fading (Brau 1991, Hans 1977, Suzu 1977) also contributes to the fluctuations in the received signal. The other consequence of fading is intersymbol interference (ISI) caused by the frequency selectivity (time dispersion) of the channel (Bell 1963a,b). This time dispersion results in the effective channel bandwidth being smaller than the information or message bandwidth (Bell 1963a; Cox 1972, 1975; Bult 1983; Gane 1989a,b; Mela 1986; Rapp 1996b). The problems caused by ISI can be controlled through the use of equalizers (Adac 1988a, Schw 1996, Samp 1997). The equalizer compensates for the various differential delays of the multipath, reducing the effective pulse spread. This leads to an increase in the effective channel bandwidth, reducing or even eliminating the effects of intersymbol interference. Note that the equalizers must be adaptive, since wireless channels are essentially unknown and time varying. We will look at equalizers later in this chapter and concentrate initially on mitigating the effects of fading.

Randomness in the received signal arises from the multipath effect. As seen in Figure 2.5, Rayleigh fading is of a short-term nature while lognormal fading is of a long-term nature. This means that any approaches to mitigate fading must take the short-term (or long-term) nature of fading into account. The randomness of the received signal can be compensated for through diversity techniques that make use of a number receiving antennae instead of a single one. Since Rayleigh fading is of a short-term nature, these receivers can be fairly close; the diversity techniques (Wint 1984, Bren 1959, Gilb 1969, Lee 1971a, Jake 1974, Turk 1990) are commonly referred to as "microscopic." The techniques employed to compensate for long-term fading are referred to as "macroscopic." The terms *microscopic* and *macroscopic* (Bern 1987) are distinguished on the basis of how close the receivers are to each other. We will initially look at the microscopic diversity techniques employed to mitigate the effects of Rayleigh fading.

5.1 EFFECTS OF FADING AND THE CONCEPT OF DIVERSITY

Even though we examined the principle of fading earlier, we need to understand how fading adversely impacts wireless communication systems. Consider the case of

signal transmission through a Rayleigh channel. The received signal, $r(t)$, can be expressed as

$$r(t) = A\exp\,[-j\theta(t)]s(t) + n(t)\,, \tag{5.1}$$

where $s(t)$ is the transmitted signal; A and $\theta(t)$ are the random gain and random phase, respectively, associated with Rayleigh fading; and $n(t)$ is additive white Gaussian noise (Schw 1996). The effect of fading, therefore, is to create a multiplicative noise term, A, in addition to the additive white Gaussian noise. This multiplicative noise (Schw 1996) will cause problems in bit detection. Rewriting eq. (5.1), the output, $z(T)$, at the sampling instant after the matched filter is

$$z(T) = A\sqrt{E} + n\,, \tag{5.2}$$

where E is the bit energy, A is a scaling factor that is Rayleigh distributed, and n is the white Gaussian noise. We can view the first term as the signal and the second term as noise. It is clear that the signal power is now a random variable. Therefore, for a fixed threshold set on the basis of a deterministic signal in the presence of additive white Gaussian noise, we will not be able to correctly identify the bits since the power/energy may go up or down at any given sampling instant. In other words, the signal-to-noise ratio (SNR) is a random variable, and the bit error rates given in Chapter 3 will be conditioned on the random variable A. If we assume that the fading is slow, the average bit error rate in the presence of fading, $p_{\text{fad}}(e)$, can be expressed as

$$p_{\text{fad}}(e) = \int_0^\infty p(e/a)f(a)\,da\,, \tag{5.3}$$

where $p(e/a)$ is the probability of error conditioned on the envelope and $f(a)$ is the Rayleigh density. The error probability under flat fading conditions can be easily evaluated by transforming the random variable A to γ,

$$\gamma = \frac{A^2E}{N_0}\,, \tag{5.4}$$

with a probability density function $p(\gamma)$ given by

$$\text{'}(\gamma) = \frac{1}{\gamma_0}\exp\left(-\frac{\sqrt{\gamma}}{\gamma_0}\right), \quad \gamma \geq 0\,, \tag{5.5}$$

where γ_0 is the average signal-to-nose ratio under fading conditions, given by

$$\gamma_0 = \frac{E}{N_0}\langle A^2\rangle. \tag{5.6}$$

The quantity N_0 is the noise power, discussed in Section 3.2. Using eqs. (5.3) and (5.5) and the expressions for error probability from Chapter 3, the error probability in the presence of fading and in the absence of fading can be expressed in terms of SNR, the average signal-to-noise ratio, as

$$\left.\begin{aligned} p(e) &= \frac{1}{2}\text{erfc}\left(\sqrt{\gamma_0}\right) \\ p_{\text{fad}}(e) &= \frac{1}{2}\left[1 - \sqrt{\frac{\gamma_0}{1+\gamma_0}}\right] \end{aligned}\right\} \quad \text{coherent BPSK} \tag{5.7a}$$

$$\left.\begin{aligned} p(e) &= \frac{1}{2}\exp(-\gamma_0) \\ p_{\text{fad}}(e) &= \frac{1}{2}\left[\frac{1}{1+\gamma_0}\right] \end{aligned}\right\} \quad \text{DPSK} \tag{5.7b}$$

$$\left.\begin{aligned} p(e) &= \operatorname{erfc}\sqrt{\varepsilon\gamma_0} \\ p_{\text{fad}}(e) &= \frac{1}{2}\left[1-\sqrt{\frac{\varepsilon\gamma_0}{1+\varepsilon\gamma_0}}\right] \end{aligned}\right\} \quad \text{coherent GMSK,} \tag{5.7c}$$

where ε is a factor determined by the bandwidth of the Gaussian low-pass filter (eq. (3.136) and γ_0 is the average signal-to-noise ratio.

The error probability plots for some modulation formats are shown in Figures 5.1, 5.2, and 5.3. It is clear that for a given SNR (E/N_0), the bit error rate in a faded channel is much higher than the error rate in a fading-free channel. This slower decline of the probability of error versus SNR is a direct consequence of the multiplicative noise. For a fixed bit error rate, this translates into margin of a few dB for the faded channel, as shown in Figure 5.1 for the case of BPSK. For an error probability of 10^{-3}, the fading margin $M \sim 17$ dB. This phenomenon may be explained as follows. The signal-to-noise ratio in a faded channel can be expressed as

$$\left(\frac{S}{N}\right)_{\text{fad}} = \frac{S}{f(s)+f(N)}, \tag{5.8}$$

where the first term in the denominator, $f(s)$, is due to the fading and the second term, $f(N)$, is the system noise. The signal is denoted by S. Note that the first term in the denominator depends on the signal. At low values of the signal, the signal-to-noise ratio in fading, $(S/N)_{\text{fad}}$, goes up as the signal power goes up since the signal-dependent noise term is negligible. However, at higher values of signal power, the signal-dependent noise term $f(s)$ starts going up, thus reducing $(S/N)_{\text{fad}}$. As the signal-dependent term overtakes the signal-independent noise term, $f(N)$, any further increase in signal power will have less and less impact on the effective $(S/N)_{\text{fad}}$ as well as on the probability of error. In other words, in a faded channel, the probability of error cannot be reduced at the same rate as in the nonfaded case by increasing the signal power.

EXAMPLE 5.1

For a coherent BPSK receiver, what is the fading margin required to maintain a bit error rate of 10^{-3}?

Answer From eq. (5.7a), the signal-to-noise ratio γ_0 required to maintain a bit error rate of 10^{-3} for a Gaussian channel (no fading) is obtained by solving

$$p(e) = 10^{-3} = \frac{1}{2}\operatorname{erfc}(\sqrt{\gamma_0}).$$

Inverting this equation, we get $\gamma_0 = 6.79$ dB.

For a Rayleigh channel, the signal-to-noise ratio is obtained by solving

$$p_{\text{fad}}(e) = 10^{-3} = \frac{1}{2}\left[1 - \sqrt{\frac{\gamma_0}{1+\gamma_0}}\right].$$

Inverting this equation, we get $\gamma_0 = 25$ dB. The fading margin is $25 - 6.79 = 18.21$ dB. ■

One of the ways in which the effects of fading can be reduced is through "diversity" techniques (Jake 1971; Lee 1971a,b; Vaug 1988). Diversity techniques involve the creation of multiple independent versions of the received signal and combining them. Consider, for example, that instead of one receiving antenna, there are M different antennae receiving the same signal. If we assume that the locations of these antennae are such that the signals being received can be considered to be independent of one another, we indeed have created M different versions of the signal. The signal envelopes of these M channels will all be identically distributed Rayleigh-faded signals. This leads to the following scenario. If one examines the signals being received in these M diverse receivers, the likelihood of all of them being severely faded (very weak signals) is very small. For example, if the envelope is Rayleigh distributed, the power will be exponential. Under this condition, the probability that the signal power in any one of the receivers is below the threshold, ζ_T, can be obtained as follows.

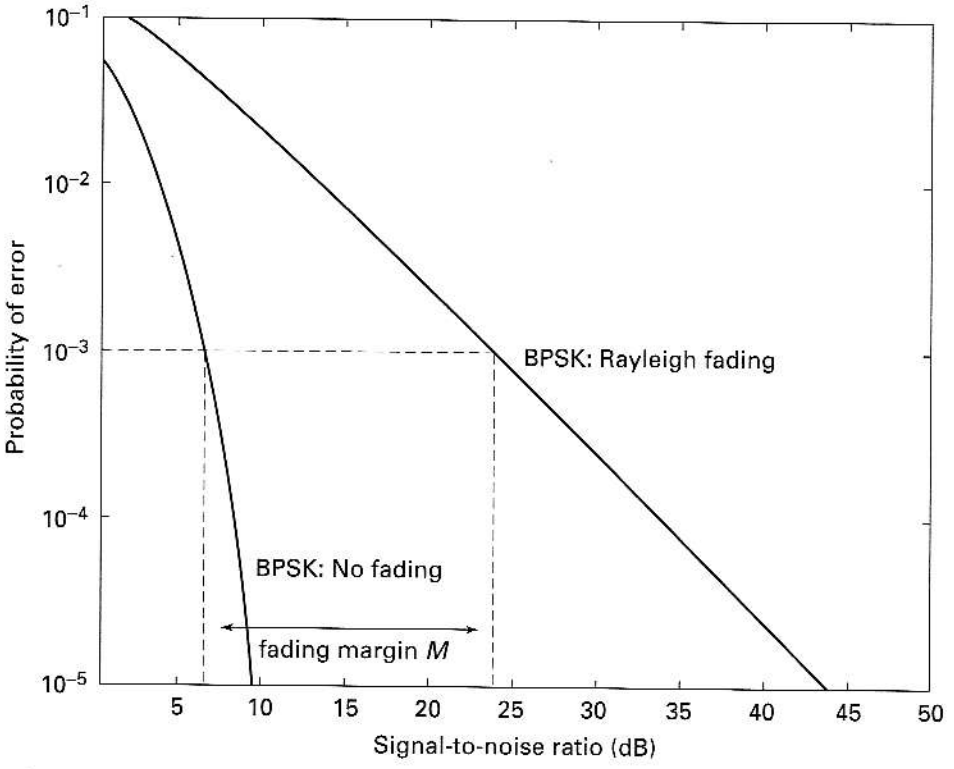

FIGURE 5.1 Probability of error for coherent BPSK in the presence of Rayleigh fading and in the absence of fading.

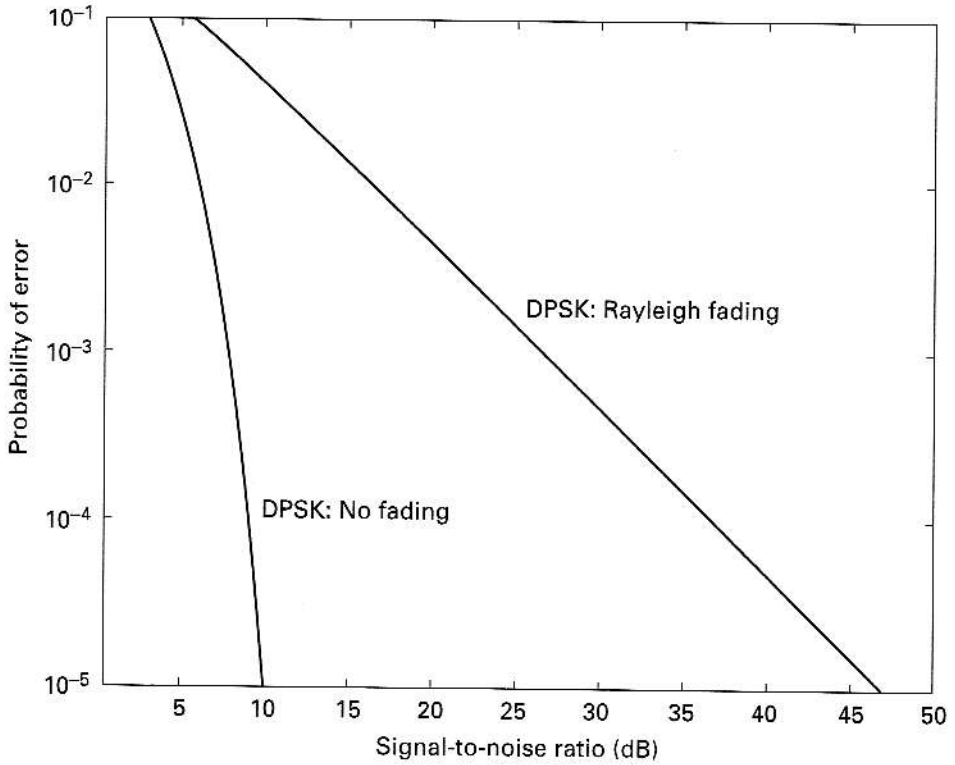

FIGURE 5.2 Probability of error for DPSK in the presence of Rayleigh fading and in the absence of fading.

$$P(\zeta_i \leq \zeta_T) = \int_0^{\zeta_T} p(\zeta_i)\, d\zeta_i = \int_0^{\zeta_T} \frac{1}{\zeta_0} \exp\left(-\frac{\zeta_i}{\zeta_0}\right) d\zeta_i = 1 - \exp\left(-\frac{\zeta_T}{\zeta_0}\right), \quad (5.9)$$

where $p(\zeta_i)$ is the exponential density function of the power ζ_i, and ζ_0 is the average power. The probability, $P_M(\zeta_T)$, that all the receivers simultaneously have signal power less than ζ_T is

$$P_M(\zeta_T) = \left[1 - \exp\left(-\frac{\zeta_T}{\zeta_0}\right)\right]^M. \quad (5.10)$$

The probability that at least one of the receivers has a signal power above the threshold, ζ_T, is

$$P\{\text{at least one receiver has power} \geq \zeta_T\} = 1 - P_M(\zeta_T). \quad (5.11)$$

Equation (5.11) is plotted in Figure 5.4 for two values of ζ_T/ζ_0. As the number of diverse channels increases, the probability that at least one receiver has power $> \zeta_T$ also increases. Thus, as M increases, the likelihood of performance degradation becomes lower and lower.

The use of different receivers at various locations is not the only form of diversity available. A number of different ways exist for the engineer to combine the signals from these diverse "branches" to produce the final signal. Descriptions of various diversity techniques and the processing methods used in combining these diverse components are given in Section 5.3.

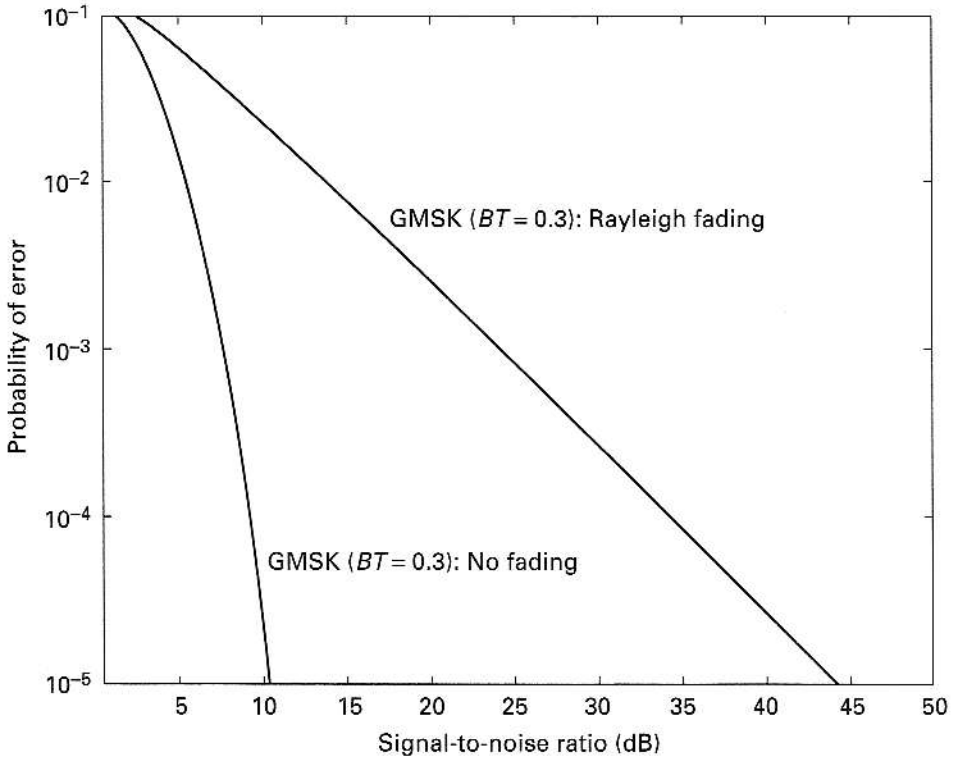

FIGURE 5.3 Probability of error for coherent GMSK in the presence of Rayleigh fading and in the absence of fading.

EXAMPLE 5.2

The average signal-to-noise ratio at the receiver is 14 dB. The receiver goes into outage when the signal-to-noise ratio goes below 10 dB. What is the probability of outage in a Rayleigh channel?

Answer The outage probability is given by eq. (5.9). Note that 14 dB corresponds to a signal-to-noise ratio in absolute units of 25.1. The threshold signal-to-noise ratio is 10 dB, or 10 in absolute units. Outage $= [1 - \exp(-10/25.1)] = 0.3286$. ■

EXAMPLE 5.3

In a five-channel diversity receiver scheme, what is the probability of outage if all channels have the same characteristics as in Example 5.2? Assume that the system goes into outage if all channels simultaneously have a signal-to-noise ratio below 10 dB.

Answer Using eq. (5.10), the outage probability is $[1 - \exp(-10/25.1)]^5 = 0.0038$. ■

5.2 OTHER SIGNAL DEGRADATION EFFECTS

Rayleigh fading is not the only signal degradation effect imposed by the channel. The relative motion of the MU with respect to the base station introduces random phase changes in the received signal. In addition, as described in Chapter 2, the relative difference between the channel bandwidth and the message bandwidth makes the fading

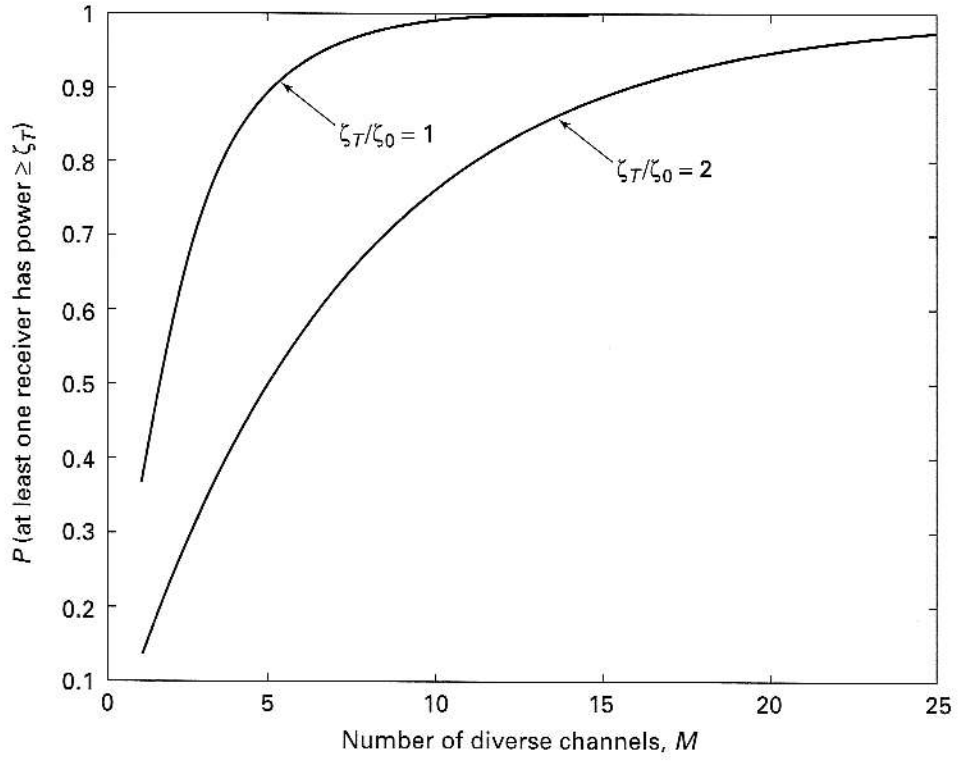

FIGURE 5.4 The number of diverse channels M is plotted against the probability that at least one channel has a power greater than ζ_T.

frequency selective, introducing intersymbol interference. We will briefly review these topics before we explore ways of mitigating the effects of fading.

5.2.1 Effects of Random Frequency Modulation

The preceding discussion is based on the assumption that the only effect of fading is a fluctuation in the power levels due to a Rayleigh-distributed envelope. However, the mobile channels also exhibit fluctuations in the frequency/phase from the motion of the vehicle, or the Doppler effect. In eq. (2.51), the presence of $\cos(\theta_i)$ in the argument of the sine and cosine terms is similar to the frequency or phase modulation seen in eq. (3.20) or (3.24), clearly showing the effect of random FM resulting from the random nature of the phase θ_i. The difficulty with the presence of random fluctuations in the frequency/phase is the fact that any increase in the transmitted power and, hence, in the received power will not help reduce the bit error rates (Adac 1989, 1993; Davi 1971; Jake 1974; Fung 1986; Liu 1991b). Indeed, the error rates reach a low "floor value" as the power increases, and beyond a certain value, any further increase in power has no effect on the bit error rates. Such error floors (Korn 1991) are characteristic of systems suffering random FM, as we have seen in Chapter 3 (Figure 3.30).

The exact error probability equations for all the modulation formats of interest are difficult to express in closed form. However, we will summarize the effects of random FM in the next section along with the effects of frequency-selective fading.

5.2.2 Effects of Frequency-Selective Fading and Co-Channel Interference

As discussed in Chapter 2, frequency-selective fading arises when the coherent bandwidth of the channel is less than the message bandwidth. The error rates vary with the form of modulation and demodulation used. The performance of the modems also depends on the ratio σ_d/T, where σ_d is the rms delay spread (eq. (2.42)) and T is the symbol period. The exact equations governing the error probability, taking frequency- selective fading into account, are again very complex and are beyond the scope of this book. Numerical results are available in a number of research papers.

We will try to explore the effects using the results for the differential detector for MSK. We can express the error probability in terms of the average signal-to noise-ratio (γ_0), average signal-to-CCI ratio (γ_{CCI}), and the Doppler parameter f_dT, where f_d is the maximum Doppler shift and T is the symbol period. The average error probability, $p_{av}(e)$, for MSK with differential detection has been shown (Hira 1979b, Akai 1998) to be

$$p_{av}(e) = \frac{1}{2}\left[1 - \frac{J_0(2\pi f_d T)\gamma_0\gamma_{CCI}}{\sqrt{(\gamma_0\gamma_{CCI} + \gamma_0 + \gamma_{CCI})^2 - \left(\frac{\gamma_0}{\pi}\right)^2 J_0^2(2\pi f_d T)}}\right]. \tag{5.12}$$

The average error probability is plotted in Figure 5.5 for the case where there is no motion (Doppler fading absent, $f_d = 0$). At a low value of signal-to-CCI ratio (10 dB), the average error is quite high and stays high around 0.1. Even when the signal-to-CCI ratio is 60 dB, the effects of error floor are seen, pointing to the fact that any additional increase in the signal power will have no effect on the performance of the system. Note that we still have not included the effects of frequency-selective fading, which will degrade the performance still further.

The average error probability in the absence of CCI (γ_{CCI} = infinite) can be obtained from eq. (5.12) as

$$p_{av}(e)_{\gamma_{CCI}\to\infty} = \frac{1 + \gamma_0[1 - J_0(2\pi f_d T)]}{2(1+\gamma_0)}. \tag{5.13}$$

The average error probability under conditions of no CCI is plotted in Figure 5.6. When the speed of the mobile unit goes up (higher values of f_d), the average error rates go up and appear to converge to an "error floor" value, which changes only marginally as the signal-to-noise ratio goes up. The performance of the system mirrors the case shown in Figure 5.5 in terms of the error floor.

The error floor, $p_{fc}(e)$, for the presence of CCI can be obtained from eq. (5.12) by letting the average signal-to-noise ratio γ_0 approach ∞ while $f_d = 0$ (no Doppler fading). It can be shown that the error floor for CCI becomes

$$p_{fc}(e) = \frac{1}{2}\left[1 - \frac{\gamma_{CCI}}{\sqrt{(\gamma_{CCI}+1)^2 - \left(\frac{1}{\pi}\right)^2}}\right]. \tag{5.14}$$

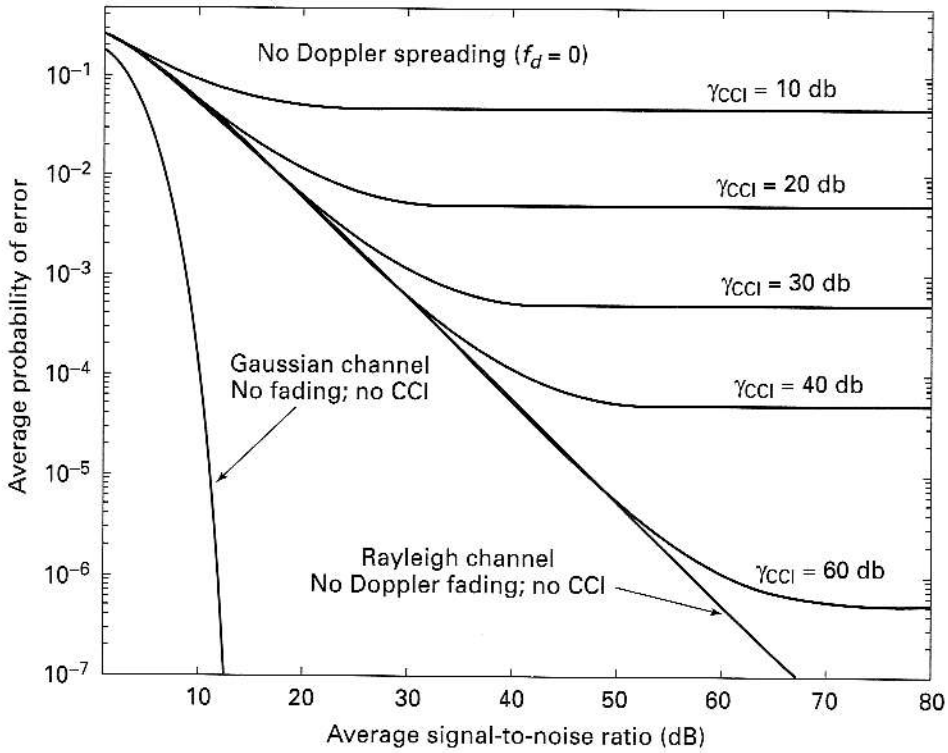

FIGURE 5.5 Error floor behavior in the presence of CCI fading for the case of MSK with a differential detector.

Similarly, the error floor $p_{fd}(e)$ for the case of Doppler fading can be obtained from eq. (5.13) by letting both the average signal-to-noise ratio, γ_0, and the signal-to-CCI ratio, γ_{CCI}, go to ∞. This error floor is

$$p_{fd}(e) = \frac{1}{2}[1 - J_0(2\pi f_d T)]\,. \tag{5.15}$$

The error floor when CCI is present is shown in Figure 5.7. It is clear that at what may be considered an acceptable value of signal-to-CCI ratio of 20 dB, in the absence of Rayleigh fading, the error is still around 0.004, demonstrating the need to keep the CCI as low as possible under fading conditions.

The error floor under the conditions of a mobile unit in motion is shown in Figure 5.8. The effect of the Doppler shift on the error is clearly seen. The error is about 4×10^{-4} when the Doppler fading term $f_d T = 0.01$.

Note that even though the example of MSK was chosen, the performance of MSK with a differential detector is identical to the case of DPSK. As indicated earlier, we can certainly draw conclusions from these results regarding the general behavior of the error probability in the presence of Doppler fading and CCI. The differences among the various modulation formats will be in the values of the error floor and at what values of the signal-to-noise ratio they are reached. The general behavior of the error rates will be similar.

When the channel is fast and frequency selective, the behavior of the error floor will be similar to that seen in Figures 5.5 and 5.6. Note that the channel is flat if σ_d

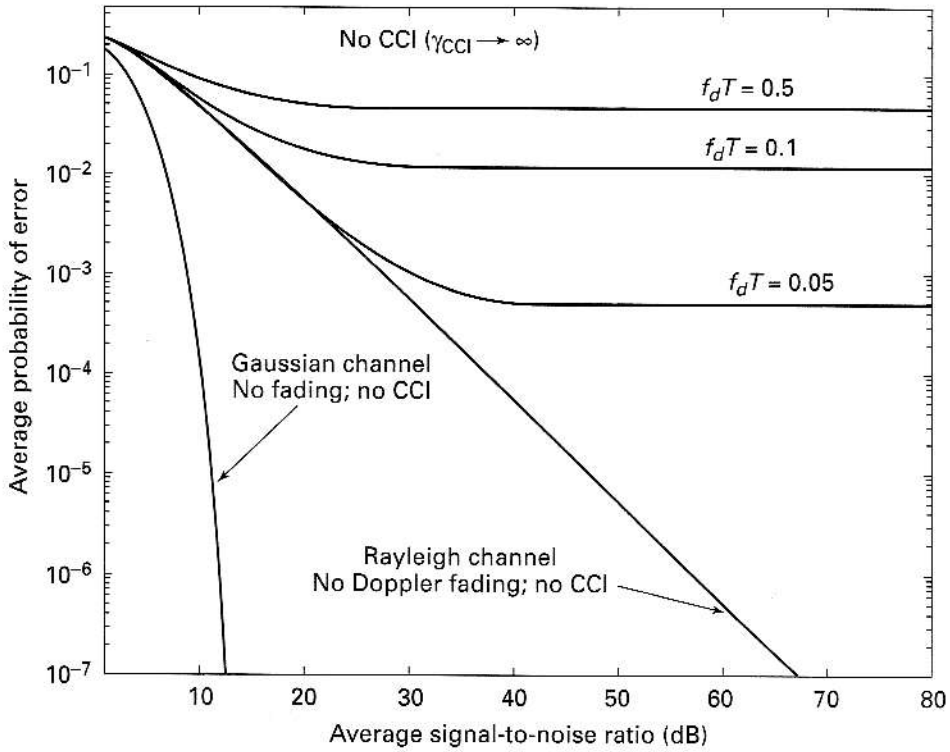

FIGURE 5.6 Error floor behavior in the presence of Doppler fading for the case of MSK with a differential detector.

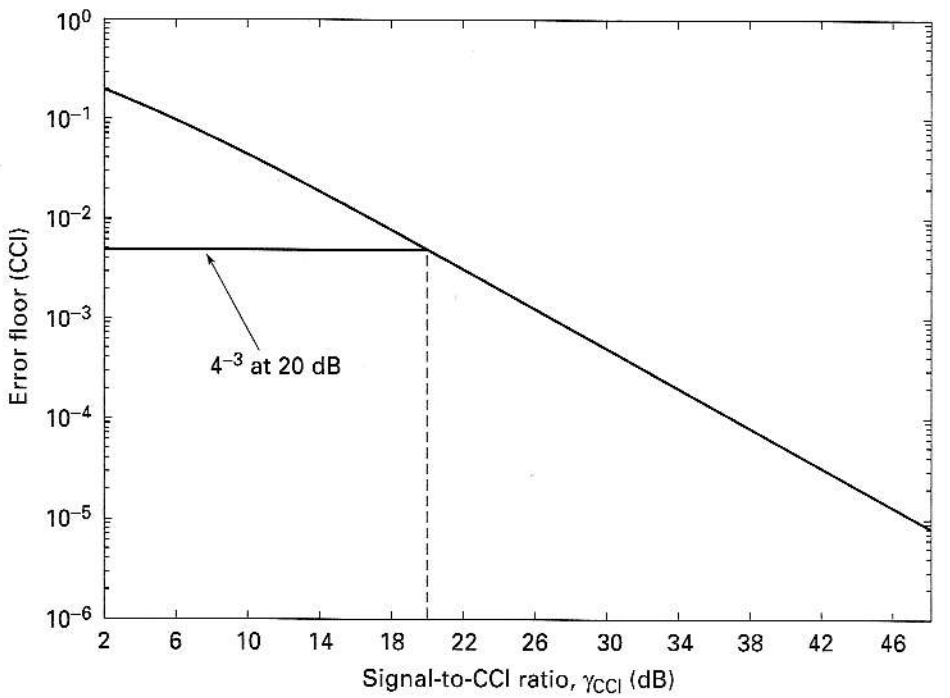

FIGURE 5.7 The error floor plotted as a function of signal-to-CCI ratio.

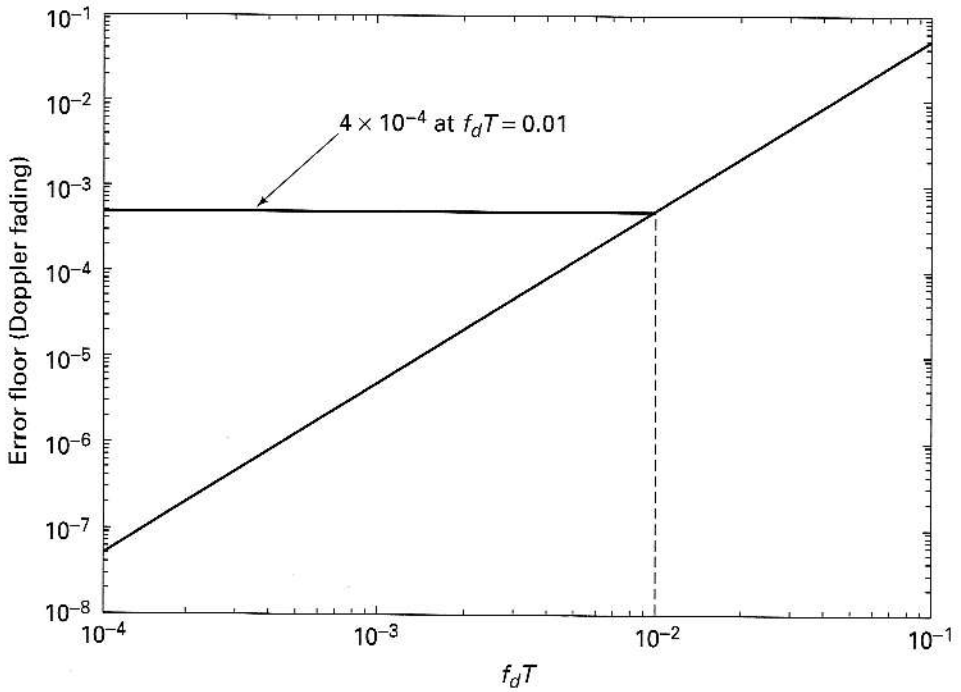

FIGURE 5.8 The error floor plotted as a function of Doppler fading parameter f_dT.

is zero. The results are extensively discussed in a number of research papers and books (Fung 1986, Liu 1991b, Fech 1993, Smit 1994, Fehe 1995).

We will now look at ways of overcoming the problems associated with fading using various diversity schemes.

5.3 FORMS OF DIVERSITY

Multiple signals can be created in the following ways (Pier 1960, Gilb 1969, Schw 1996):

Space diversity (antenna locations physically separated)
Angle diversity (different angles of arrival)
Frequency diversity (different frequencies/frequency bands)
Polarization diversity (different polarizations)
Time diversity
Multipath diversity

All these forms of diversity can produce multiple signals depending on a number of factors, such as available space, frequency, angle, and time, except for polarization diversity, where M is limited to 2. The trade-off in the improvement in performance between M and available space, frequency, angle, etc. will be discussed later. The multiple signals can be combined and processed in a number of ways in addition to picking the strongest of the M signals. These methods are

Selection combining
Maximal ratio combining
Equal gain combining.

We will examine the different forms of diversity before we look at the various methods of diversity combining.

5.3.1 Space or Spatial Diversity

Spatial diversity takes advantage of the random nature of propagation in different directions. Consider, for example, the transmitted signal propagating through a multipath scattering environment as shown in Figure 5.9*a*. Two antenna sites, separated by a distance D, are shown at the receiving location (Schw 1996, Maki 1967, Lee 1971b). It is possible for the signals arriving at the two antenna sites to be independent or at least uncorrelated, depending on the separation between them. This possibility arises from the very nature of multipath fading, specifically the fact that many independent paths exist

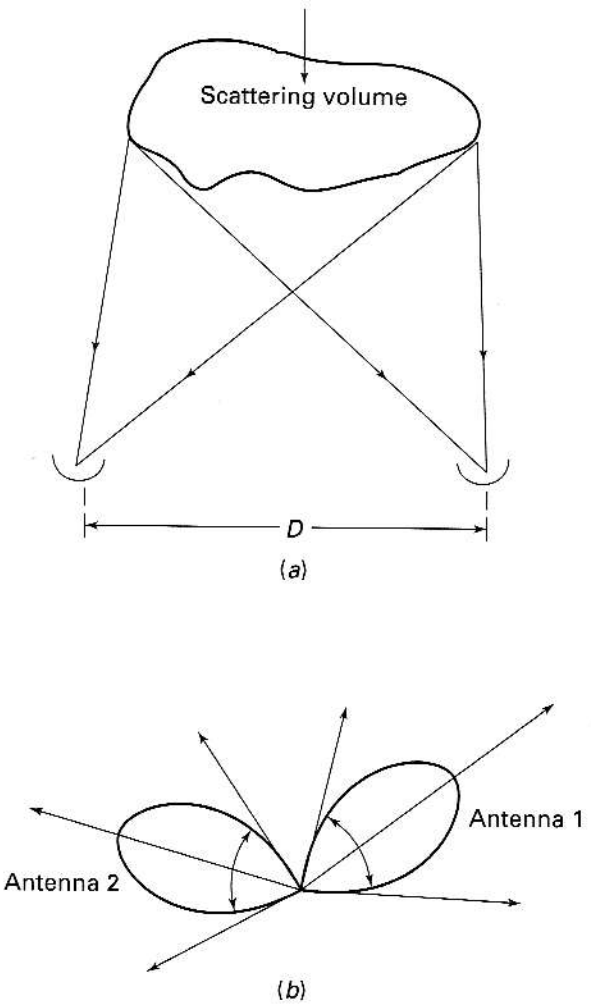

FIGURE 5.9 (*a*) Spatial diversity. D is the antenna separation. (*b*) Angular diversity. Antennae 1 and 2 are two antennae with narrow beam patterns.

at any location, and therefore it is possible for the paths coming to one receiver to be independent of those to the other. It is obvious that if the separation is reduced, we will reach a point where there is a significant overlap between the groups of paths arriving at the two sites. One can now visualize locating a number of such antennae at the receiving site, each antenna separated from its neighbor by the minimum distance necessary to make the signals arriving at the particular antenna independent of (or at least uncorrelated with) the signals arriving at the other antenna sites. Independence of the signals from the different diversity branches is essential for the improvement in performance from multiple observations expressed in eq. (5.11). Note that if the radio frequency components are uncorrelated and Gaussian, they are also independent. Thus, in the case of Rayleigh fading, the requirement of zero correlation (uncorrelated) is sufficient since the Rayleigh density function results from two Gaussian random variables (Appendix A).

5.3.2 Angle Diversity

Angular diversity makes use of antennae having directional properties. Each antenna responds to a received signal propagating in a specific direction or angle, and this faded signal will be uncorrelated with (preferably independent from) other signal components being received by the other directional antennae (Peri 1998, 1999). A schematic of this arrangement is shown in Figure 5.9*b*. A mobile unit is shown with two antennae having narrow beam widths, positioned in different angular directions.

In spatial diversity, the two antennae are separated horizontally and pointed in the same direction as shown in Figure 5.9*a*; in angular diversity, the antennae are co-located, collecting signals coming from different angular directions. Whereas space diversity is implemented at the base station, angular diversity may be implemented at the base station or at the mobile unit (Lee 1997, Kuch 1999, Peri 1999). Improvement in performance is realized when the signals from the two antennae are uncorrelated.

5.3.3 Frequency Diversity

Rather than using several antennae separated by a specific distance, we can transmit the same information at different carrier frequencies. The resulting diversity is referred to as frequency diversity. The carrier frequencies must be separated enough so that the signals corresponding to the different carrier frequencies are at least uncorrelated, and preferably independent. Under these conditions, the signals at these carrier frequencies will result in independently fading signals that can be combined using any of the techniques described later in this chapter. The limitation of the technique is the availability of bandwidth to allow a number of different frequencies and the ability of the receiver to pick up these diverse signals without the need for multiple receivers tuned to different frequencies. In some implementations of cdma2000/IMT 2000, frequency diversity is used to reduce the effects of fading (Knis 1998a,b; Rao 1999). This has been made possible through the allocation of a significant amount of bandwidth for third-generation (3G) wireless communications.

5.3.4 Polarization Diversity

Polarization diversity makes use of the inherent property of most scattering media to depolarize the propagating beam even when the transmitted signal may be vertically (or horizontally) polarized. The depolarized signal arriving at the antenna can be split

into two orthogonal polarizations (vertical and horizontal), producing two independently fading signals (Bren 1959, Vaug 1990, Atan 2000). The only limitation of this technique is the inability to generate more than two diverse signals, whereas it is possible to have many diverse signals in spatial, frequency, or angular diversity.

5.3.5 Time Diversity

If we can transmit the same data stream repeatedly at intervals that exceed the coherence time of the channel, these multiple information streams will be subjected to independent fading. These independently faded components can then be combined using any diversity combining techniques. Time diversity techniques have one major advantage over the space and frequency diversity systems in not requiring multiple antennae and needing only a single antenna at the receiver. Time diversity systems, however, do require storage of the received data streams for processing (Turk 1990, Stei 1987). Time diversity (similar to in-band diversity, since we are using the same frequency characteristics) might require some form of transmit power division penalty since one transmitter is used to transmit the data repeatedly. One can consider the case of error control coding or sending more data than necessary (redundant data transmission) as a form of time diversity. Error control coding is discussed in Section B.3, Appendix B.

5.3.6 Multipath Diversity

Instead of transmitting at different times, it is possible to take advantage of the existing multipath signals present in fading wireless channels (Lehn 1987). If the time duration of the pulses is very small, then the multipath signals corresponding to the different paths will be nonoverlapping, and, therefore, uncorrelated. One such case is shown in Figure 5.10.

The three pulses, 1, 2, and 3, are time-delayed versions of the same pulse and are resolvable. Such pulses can be combined like any other diverse signals by taking into account the amount of delay that exists between the multiple paths, as shown in Figure 5.11. For most data transmission systems, the multipath components are

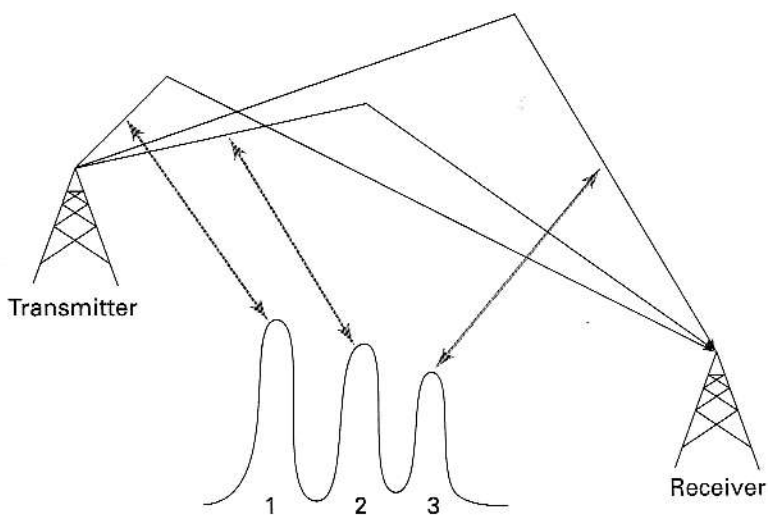

FIGURE 5.10 Multipath diversity.

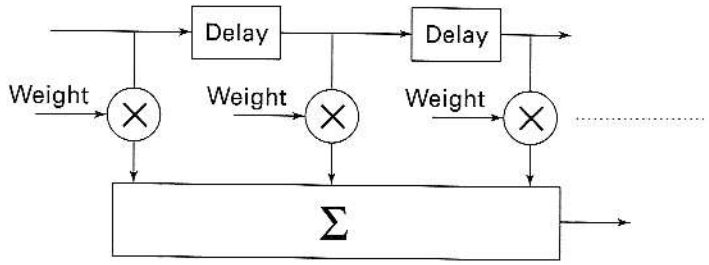

FIGURE 5.11
Combining approach used in multipath diversity.

unresolvable and overlap, making them useless for multipath diversity approaches. On the other hand, in wideband systems employing code division multiple access, since the pulse widths used are extremely small, it is possible to "pick" each delayed version of the same pulse for the purposes of employing multipath diversity. A RAKE receiver collects these time-shifted (or delayed) versions of the transmitted signals for processing. We will look at RAKE receivers in Chapter 6 after we discuss the details of code division multiple-access techniques.

5.4 DIVERSITY COMBINING METHODS

The multiple versions of the signals created using the various diversity techniques must be combined in some fashion to provide improved performance in terms of lowering the minimum signal-to-noise ratio required to maintain an acceptable level of bit error rate. There are three primary means of combining the signals; selection combining, maximal ratio combining, and equal gain combining (Bren 1959; Alhu 1985; Beau 1991; Adac 1989, 1991a,b; Anna 1998, 1999).

5.4.1 Selection Combining

Selection combining is the simplest form of diversity combining techniques. Its effectiveness is based on the simple principle that not all signals coming out of the diversity branches will have low values at the same time. Under these conditions, it is possible to look for the branch having the highest signal-to-noise ratio and use that particular branch as the primary received signal. This is illustrated in Figure 5.12. Five sets (1000 each) of exponentially distributed random numbers corresponding to the power of a Rayleigh-faded signal were generated. Two of these sets (only 40 samples) are shown. Also shown is the largest of the five sets for every sample number. It demonstrates that it is possible to ensure that the signal power will not always be small if we choose the largest of a number of outputs. The processor principle is shown in Figure 5.13.

Consider M independent Rayleigh-fading components produced by any diversity scheme. Since the envelope is Rayleigh distributed, the signal power and therefore the signal-to-noise ratio, γ_n, will be exponentially distributed. If γ_0 is the average signal-to-noise ratio of any one of the diversity components, the pdf, $f(\gamma_n)$, of the instantaneous signal-to-noise ratio, γ_n, of any one of the M components can be expressed as

$$f(\gamma_n) = \frac{1}{\gamma_0}\exp\left(-\frac{\gamma_n}{\gamma_0}\right)U(\gamma_n), \quad n = 1, 2, \ldots, M. \tag{5.16}$$

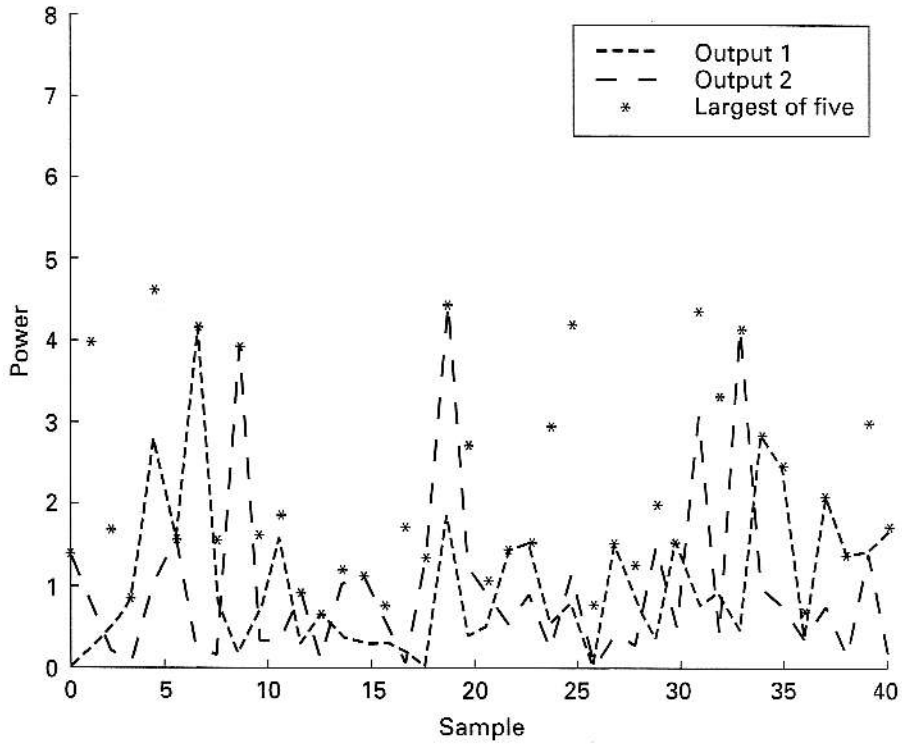

FIGURE 5.12 Principle of selection diversity.

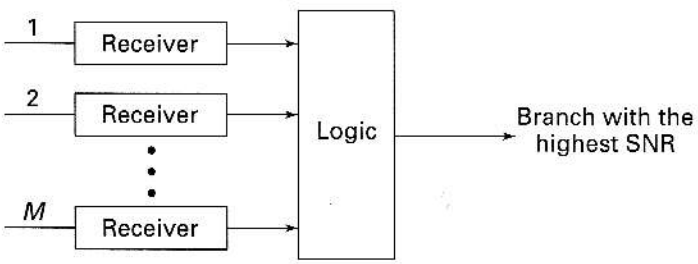

FIGURE 5.13 Signal processor for selection diversity.

The probability that the signal-to-noise ratio in any one of the diversity components will be less than any specific value, γ, is given by

$$P[\gamma_n \leq \gamma] = \int_0^{\gamma} f(\gamma_n)\, d\gamma_n = 1 - \exp\left(-\frac{\gamma}{\gamma_0}\right), \quad n = 1, 2, 3, \ldots, M. \quad (5.17)$$

Since all of the M components or branches are independent, the probability that all of them would have a SNR less than γ is

$$P_r(\gamma_1, \gamma_2, \gamma_3, \ldots, \gamma_M \leq \gamma) = \left[1 - \exp\left(-\frac{\gamma}{\gamma_0}\right)\right]^M. \quad (5.18)$$

The probability, $P_M(\gamma)$, that at least one branch achieves a signal-to-noise ratio greater than γ will now be

$$P_M(\gamma) = 1 - P_r(\gamma_1, \gamma_2, \gamma_3, \ldots, \gamma_M \leq \gamma) = 1 - \left[1 - \exp\left(\frac{\gamma}{\gamma_0}\right)\right]^M. \quad (5.19)$$

This is the probability associated with the selection combining algorithm (Adac 1991a, Fehe 1995, Rapp 1996b, Garg 1996). It provides a measure of the degree of improvement we can expect if we are able to examine the various diversity components and pick the one having the signal-to-noise ratio exceeding the value that has been set. To compute the actual improvement we can expect out of this selection combining, we must calculate the pdf of the output of the selection combining algorithm, namely, the pdf associated with the probability given in eq. (5.19). The pdf of the output of the "selection combiner," $f(\gamma)$, is obtained by differentiating eq. (5.16) with respect to γ, leading to

$$f(\gamma) = \frac{d}{d\gamma}\left\{1 - \left[1 - \exp\left(-\frac{\gamma}{\gamma_0}\right)\right]^M\right\}$$

$$= \frac{M}{\gamma_0}\left[1 - \exp\left(-\frac{\gamma}{\gamma_0}\right)\right]^{M-1} \exp\left(-\frac{\gamma}{\gamma_0}\right). \quad (5.20)$$

Equation (5.20) may also be obtained directly from the results on "order statistics" (Papo 1991). The probability density function is plotted in Figure 5.14 for

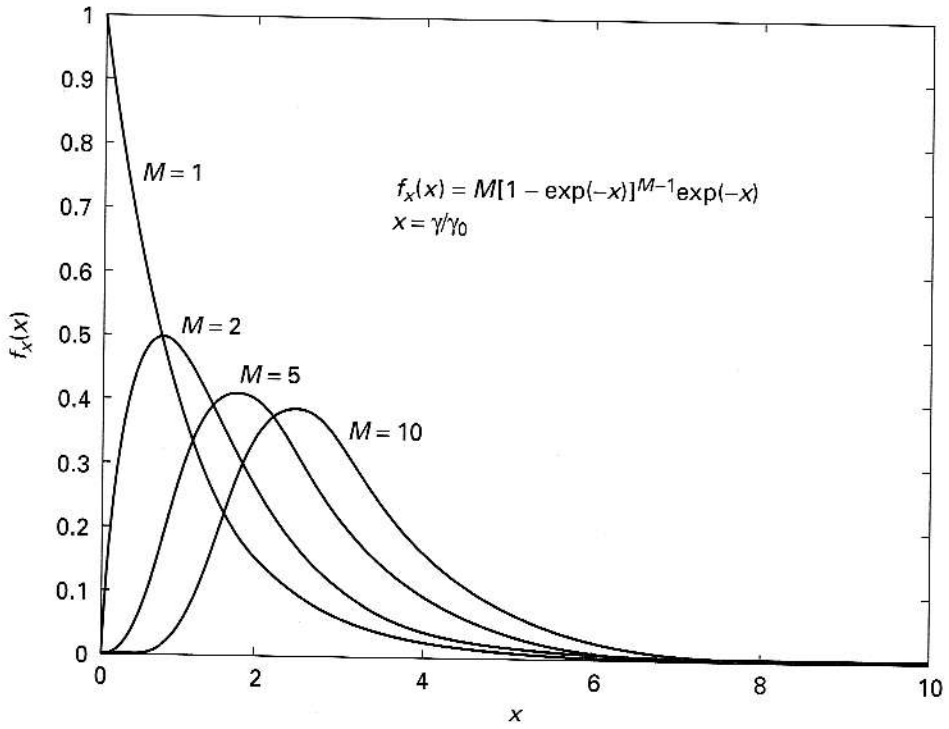

FIGURE 5.14 Probability density function of the selection combiner.

different values of M. For $M = 1$, the pdf is identical to the exponential density function associated with the signal-to-noise ratio of a single branch. It is seen that the peak of the density function moves to the right as M increases. This shift to the right will make the pdf attain more and more symmetry, thus making it closer and closer to being a Gaussian density function. When that happens, the fading channel can be considered to be no different from an ideal channel, i.e., a Gaussian channel. This transformation from Rayleigh to Gaussian is the primary reason for the expected reduction in bit error rates after compounding. We can also draw an analogy with Rician statistics (Chapter 2) and the improvement brought on by the presence of a steady component as we look at this density function as its peak continues to move to the right.

The average signal to noise ratio, γ_{se}, of the output of the selection combiner will be (Jake 1971, 1974)

$$\gamma_{se} = \int_0^\infty \gamma f(\gamma)\, d\gamma = \gamma_0 \sum_{n=1}^{M} \frac{1}{n}. \tag{5.21}$$

The improvement in signal-to-noise ratio obtained through selection combining is

$$\frac{\gamma_{se}}{\gamma_0} = \sum_{n=1}^{M} \frac{1}{n}. \tag{5.22}$$

The actual improvement in performance will depend on the modulation/demodulation format used.

Practical Considerations The implementation of selection combining starts with the examination of the received signal components from the diversity branches, with the one that has the strongest signal being selected. This is a little cumbersome since the process of examination for the strongest signal has to be undertaken continuously. This problem can be overcome if we resort to the "scanning selection combining" method. In this approach, at the beginning of the process of selection combining, the branch with the strongest signal is chosen. This particular branch is used until such time that the SNR of this branch goes below the threshold value. As the SNR value in that branch goes down, a new selection is made, and the process is continued.

EXAMPLE 5.4

Consider a two-channel selection combiner. The outage occurs when the signal-to-noise ratio goes below one-fourth of the average. Show that the outage probability with a two-channel selection combiner is smaller than the outage probability with no selection diversity.

Answer The pdf of the signal-to-noise ratio of a two-channel selection combiner can be written as

$$f(\gamma) = \frac{2}{\gamma_0}\left[1 - \exp\left(-\frac{\gamma}{\gamma_0}\right)\right]\exp\left(-\frac{\gamma}{\gamma_0}\right)U(\gamma).$$

The outage probability for a two-channel selection combiner is

$$\int_0^{\gamma_0/4} f(\gamma)\, d\gamma = 2[0.5 + 0.5 \times \exp(-0.5) - \exp(-0.25)] = 0.0489.$$

The outage probability in the absence of diversity is

$$\int_0^{\gamma_0/4} \frac{1}{\gamma_0} \exp\left(-\frac{\gamma}{\gamma_0}\right) d\gamma = 1 - \exp(-0.25) = 0.2212.$$

The reduction in outage probability because of diversity is clearly seen in the two probabilities obtained. ■

5.4.2 Maximal Ratio Combining

The selection combining approach to the processing of diversity branches, though simple, does not take full advantage of the availability of multiple signals. Maximal ratio combining (Wint 1984, Schw 1996, Akai 1998) creates a new signal that is a linear combination of all the multiple signals with appropriate weighting. The concept is shown in Figure 5.15. The outputs of the different diversity branches are added in phase with a gain factor such that the processed signal, r_{MR}, can be expressed as

$$r_{\text{MR}} = \sum_{n=1}^{M} g_n a_n + \sum_{n=1}^{M} g_n n_n, \tag{5.23}$$

where g_n is the weighting factor for the nth diversity branch and a_n is a multiplicative component that represents the Rayleigh fading. We can rewrite eq. (5.23) as

$$r_{\text{MR}} = r_M + N_M, \tag{5.24}$$

where r_M and N_M, respectively, are the signal and noise terms, given by

$$r_M = \sum_{n=1}^{M} g_n a_n, \quad N_M = \sum_{n=1}^{M} g_n n_n. \tag{5.25}$$

The signal-to-noise ratio of the maximal ratio combiner, γ_{MR}, is

$$\gamma_{\text{MR}} = \frac{r_M^2}{2N_{MP}}, \tag{5.26}$$

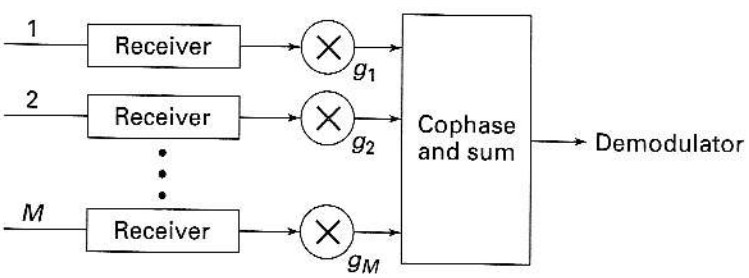

FIGURE 5.15
Signal processing for maximal ratio combining (MRC).

where N_{MP} is the total noise power. This value is obtained by assuming that the noises in all the channels are identical and independent, and therefore

$$N_{MP} = N\sum_{n=1}^{M} g_n^2, \tag{5.27}$$

where N is the average noise power in any channel. It can be shown that the signal-to-noise ratio of the maximal ratio combiner will be maximum when the gain factor, g_n, is chosen to be equal to

$$g_n = \frac{a_n}{N}. \tag{5.28}$$

Rewriting eq. (5.26) for the signal-to-noise ratio of the maximal ratio combiner, γ_{MR} becomes

$$\gamma_{MR} = \frac{1}{2N}\frac{\sum_{n=1}^{M}\left(a_n^2/N\right)^2}{\sum_{n=1}^{M}\left(a_n/N\right)^2} = \sum_{n=1}^{M}\gamma_n, \tag{5.29}$$

where γ_n is the instantaneous signal-noise-ratio given by

$$\gamma_n = \frac{a_n^2}{2N}. \tag{5.30}$$

The SNR of the maximal ratio combiner is the sum of the signal-to-noise ratios of all the diversity branches. Since all of them are independent and identically distributed, the signal-to-noise ratio is

$$\gamma_{MR} = \sum_{n=1}^{M}\gamma_n = M\gamma_0, \tag{5.31}$$

where γ_0 is the signal-to-noise ratio in any one of the branches. It is thus seen that the maximal ratio combiner results in a higher signal-to-noise ratio than the selection combiner.

Once again, the probability density function, $f(\gamma)$, of the signal-to-noise ratio γ at the output of the maximal ratio combiner can be expressed as (Schw 1996)

$$f(\gamma) = \frac{1}{(M-1)!}\frac{\gamma^{M-1}}{\gamma_0^M}\exp\left(-\frac{\gamma}{\gamma_0}\right). \tag{5.32}$$

The density function is plotted in Figure 5.16. Once again, for $M = 1$, the pdf is identical to the density function of the signal-to-noise ratio of a single branch (no diversity). We see that as M increases, the peaks of the density functions move to the right. Just as in the case of selection diversity, the density functions of the compounded signals become more Gaussian-like. The channels after fading thus become Gaussian channels, leading to a reduction in bit error rates and outage probabilities.

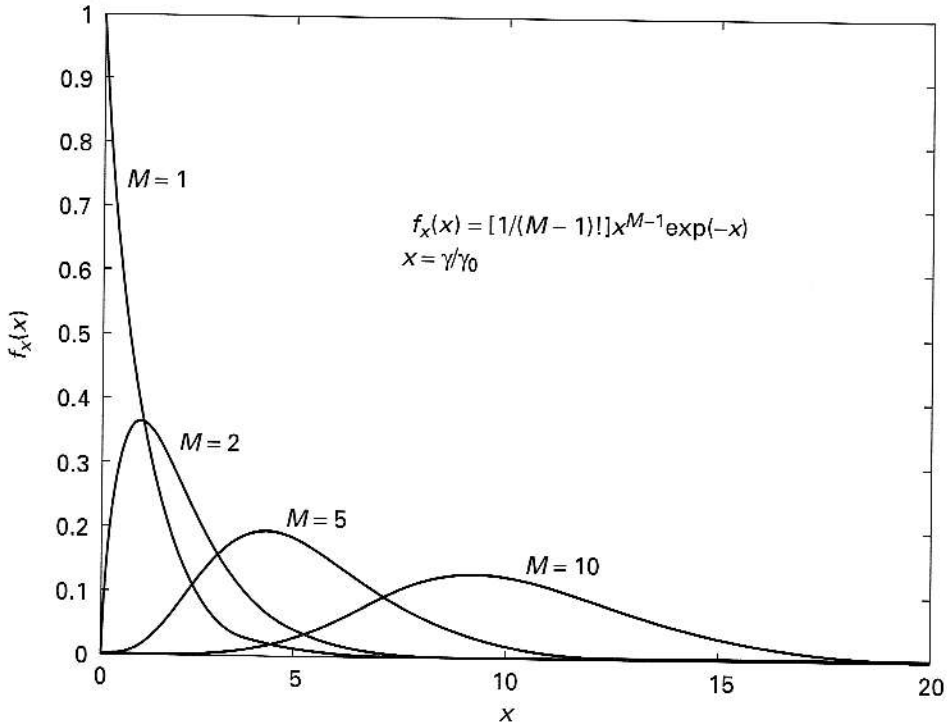

FIGURE 5.16 Probability density function for the maximal ratio combiner.

EXAMPLE 5.5

Consider a two-channel MRC system. The outage occurs when the signal-to-noise ratio goes below one-fourth of the average signal-to-noise ratio. Show that the outage probability with a two-channel selection combiner is smaller than the outage probability with no selection diversity.

Answer The density function for a two-channel MRC receiver is

$$f(\gamma) = \frac{\gamma}{\gamma_0^2}\exp\left(-\frac{\gamma}{\gamma_0}\right)U(\gamma) .$$

The outage probability is obtained as

$$\int_0^{\gamma_0/4} \frac{\gamma}{\gamma_0^2}\exp\left(-\frac{\gamma}{\gamma_0}\right) d\gamma = 1 - \frac{5}{4}\exp(-0.25) = 0.0265$$

The outage probability in the absence of compounding is 0.2212 (from Example 5.4). We once again see the reduction in the outage probabilty as a result of diversity. We also see that the outage with the MRC system is less than the outage for a selection combiner. ■

Practical Considerations The maximal ratio combiner does outperform the selection combiner. However, this improvement is realized at the cost of increased complexity. Considerable signal processing is necessary to achieve the correct weighting factors. This implementation problem naturally leads to the case where the weighting factors can be made equal, as described next.

5.4.3 Equal Gain Combining

The equal gain combiner is a maximal ratio combiner in which all the weights are equal. All the weights are set to unity, and the processing described in the previous section is undertaken. The signal-to-noise ratio for the equal gain combiner, γ_{EC}, can be shown to be equal to

$$\gamma_{EC} = \gamma_0\left[1 + \frac{\pi}{4}(M-1)\right], \tag{5.33}$$

which lies between the SNRs of the maximal ratio combiner and the selection combiner. In fact, as M increases, the difference between equal gain combining and maximal ratio combining becomes much smaller.

Once again, the approximate probability density function, $f(\gamma)$, of the signal-to-noise ratio (energy) γ at the output of the equal gain combiner can be expressed as (for $\gamma \ll \gamma_0$),

$$f(\gamma) = \frac{2^{M-1}M^M}{(2M-1)!}\frac{\gamma^{M-1}}{\gamma_0^M}. \tag{5.34}$$

The performance of the three signal-processing schemes is compared in Figure 5.17, with γ_{av} equal to the signal-to-noise ratio after combining. The performance comparison is also given in tabular form in Table 5.1.

EXAMPLE 5.6

Consider a two-channel equal gain diversity system. Outage occurs when the signal-to-noise ratio goes below one-fourth of the average SNR. Show that the outage probability with a two-channel equal gain combiner is smaller than the outage probability with no diversity.

Answer The density function of the signal-to-noise ratio of the equal gain combiner output is given by eq. (5.34),

$$f(\gamma) = \frac{3}{2}\left(\frac{\gamma}{\gamma_0^2}\right) \text{(approx.)}.$$

The outage probability is

$$\int_0^{\gamma_0/4} f(\gamma)\,d\gamma = 0.0469.$$

The outage in the absence of compounding is 0.2212 (from Example 5.4). We once again see the reduction in outage probability because of diversity. We also see that the outage with the equal gain system is less than that for the case of a selection combiner. However, the outage probability for an equal gain combiner is higher than the outage probability for a MRC system. ■

The performance of the equal gain combiner is midway between the performance levels of selection combining and maximal ratio combining, the latter providing the best performance. The selection combiner is much easier to implement than the maximal ratio combiner.

TABLE 5.1 Signal-to-Noise Ratio Improvement through Diversity

	Signal-to-noise ratio improvement (dB)		
Number of branches, M	Selection combiner $\sum_{K=1}^{M} \frac{1}{K}$	Maximalratio combiner M	Equal gain combiner $1+(M-1)\frac{\pi}{4}$
1	0	0.00	0.00
2	1.761	3.01	2.52
3	2.632	4.77	4.10
4	3.187	6.02	5.26
5	3.585	6.99	6.17
6	3.892	7.78	6.92

Effects of Correlated Signals All of the analysis presented so far has assumed that the multiple signals generated by any one of the diversity techniques are all uncorrelated. Because of incorrect positioning of the receiving antennae (spatial diversity) or insufficient separation between the frequency bands (frequency diversity), the signals from the different branches can be correlated. For correlation in the range of a few percentage points, the degradation in performance will not be significant. Even when the correlation is high (>50% for example), the performance of the diversity-based system will still be better than for nondiversity systems, even though the im-

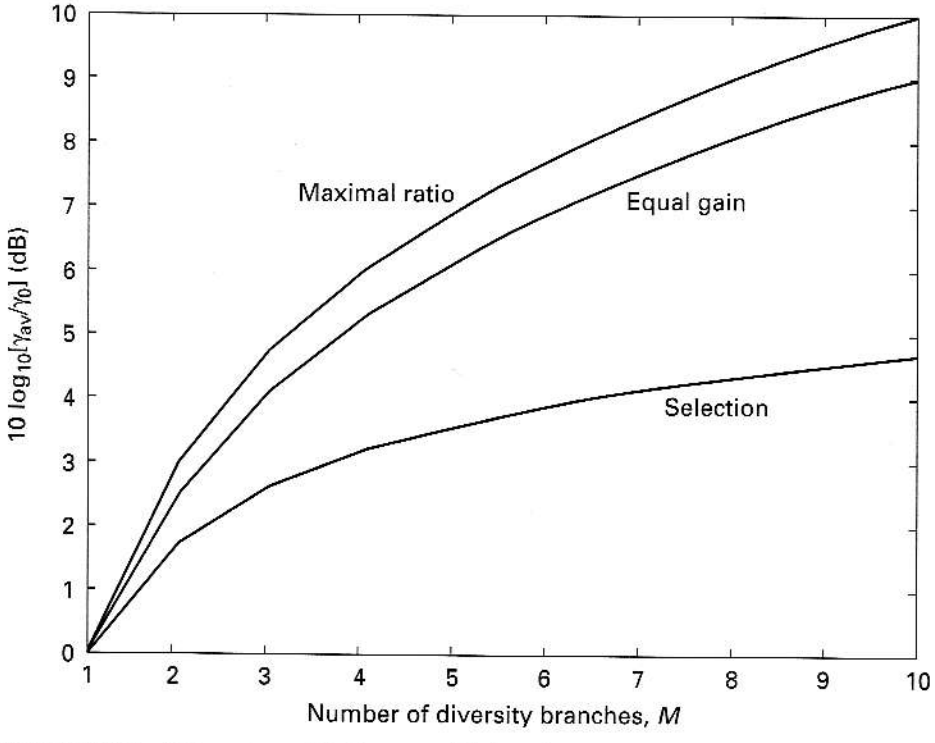

FIGURE 5.17 Comparison of the three combining schemes.

provement will be much less than if the signals were truly uncorrelated. In fact, the improvement (signal-to-noise ratio) in performance due to diversity goes down by a factor of $\sqrt{1-\rho^2}$, where ρ is the correlation coefficient (Schw 1996, Akai 1998).

5.5 PERFORMANCE IMPROVEMENT FROM DIVERSITY IN TERMS OF REDUCED BIT ERROR RATE

The improvement in performance realized through the use of diversity techniques can now be calculated. We will simplify the analysis by expressing the equation for the probability of error in terms of a general expression such as

$$p(e) = \frac{1}{2}\text{erfc}(\sqrt{b\gamma}) \tag{5.35}$$

for coherent modulation techniques, with $b = 1$ for BPSK, $b = 0.5$ for FSK, and $b = 0.85$ or 0.68 for GMSK. The signal-to-noise ratio (energy) is γ. Similarly, the probability of error for a noncoherent receiver can be expressed as

$$p(e) = \frac{1}{2}\exp(-b\gamma)\,. \tag{5.36}$$

The error probability, $P_{\text{av}}(e)$, at the output of the diversity combiner is given by

$$p_{\text{av}}(e) = \int_0^\infty p(e) f(\gamma)\, d\gamma, \tag{5.37}$$

where $f(\gamma)$ is the pdf of the signal-to-noise ratio γ for the specific diversity combining algorithm. The equations governing the bit error rate at the output of the diversity receiver are complex (Schw 1996; Proa 1995; Akai 1998; Samp 1997; Eng 1996; Wint 1984; Bigl 1998; Biya 1995; Bour 1993; Chen 1991, 1995). However, approximate expressions are available. For example, at high signal-to-noise ratios, the bit error rate at the output of a selection combiner system for BPSK can be expressed as

$$p_{\text{avsc}}(e) = \int_0^\infty \frac{1}{2}\text{erfc}\sqrt{\gamma}\, M\left[1-\exp\left(-\frac{\gamma}{\gamma_0}\right)\right]^{M-1}\exp\left(-\frac{\gamma}{\gamma_0}\right)d\gamma\,. \tag{5.38}$$

This equation can be simplified (Schw 1996, Samp 1997, Proa 1995) to

$$p_{\text{avsc}}(e) = \frac{1}{2}M\sum_{k=0}^{M-1}(-1)^k C_k^{M-1}\frac{1}{k+1}\left(1-\frac{1}{\sqrt{1+\dfrac{k+1}{\gamma_0}}}\right). \tag{5.39}$$

Similarly, the average error probability for a maximal ratio combiner can be expressed as

$$p_{\text{avmc}}(e) = \int_0^\infty \frac{1}{2}\text{erfc}\left(\sqrt{\gamma}\right)\frac{1}{(M-1)!}\frac{\gamma^{M-1}}{\gamma_0^M}\exp\left(-\frac{\gamma}{\gamma_0}\right)d\gamma\,. \tag{5.40}$$

This equation can be simplified to

$$p_{\text{avmc}}(e) = 2^{-M}\left(1 - \sqrt{\frac{\gamma_0}{1+\gamma_0}}\right)^M \sum_{k=0}^{M-1} C_k^{(M+k-1)}\left(\frac{1}{2}\right)^k \left(1 + \sqrt{\frac{\gamma_0}{1+\gamma_0}}\right)^k. \quad (5.41)$$

For noncoherent receivers, the average error probability for the cases of selection combining and maximal ratio combining can be expressed as

$$p_{\text{avsn}}(e) = \int_0^\infty \frac{1}{2}\exp(-\gamma)\frac{M}{\gamma_0}\left[1 - \exp\left(-\frac{\gamma}{\gamma_0}\right)\right]^{M-1} \exp\left(-\frac{\gamma}{\gamma_0}\right) d\gamma\,, \quad (5.42)$$

which simplifies to

$$p_{\text{avsn}}(e) = \frac{1}{2}M\sum_{k=0}^{M-1}(-1)^k C_k^{M-1}\frac{1}{k+1}\left(1 - \frac{1}{\gamma_0/(k+1)}\right), \quad (5.43)$$

and

$$p_{\text{avmn}}(e) = \int_0^\infty \frac{1}{2}\exp(-\gamma)\frac{1}{(M-1)!}\frac{\gamma^{M-1}}{\gamma_0^M}\exp\left(-\frac{\gamma}{\gamma_0}\right) d\gamma\,, \quad (5.44)$$

which simplifies to

$$p_{\text{avmn}}(e) = \frac{1}{2}\left[\frac{(M-1)!}{(1+\gamma_0)^M}\right]. \quad (5.45)$$

The bit error rate curves for the different cases are shown in Figures 5.18–5.21.

Comparison of Figures 5.18 and 5.19 shows the usefulness of diversity. As M increases, the performance approaches that of an ideal channel, i.e., a Gaussian one. This convergence to Gaussian behavior was discussed in Sections 5.4.1 and 5.4.2 in connection with the shape of the density functions after compounding. It is also obvious from these figures that MRC diversity is better than selection combining in terms of how quickly Gaussian channel conditions are realized. For example, with the selection combiner, a bit error rate of 10^{-5} is reached at a signal-to-noise ratio of about 13 dB. The same bit error rate is reached at a lower SNR value with the MRC system. The performance of the DPSK system (Figures 5.20 and 5.21) for the two methods of diversity combining is similar to that for the BPSK system.

These results do not include the effects of Doppler fading, CCI, or frequency-selective fading. Indeed, the analytical results of the analysis of GMSK and $\pi/4$-DQPSK systems are not simple to derive. Numerical results are available in a number of publications (Guo 1990; Kaas 1998; Liu 1991a,b; Ng 1993, 1994). We will, however, look at a simple case of a DPSK system employing two-branch diversity using the MRC scheme in the presence of Doppler fading (Jake 1974). The average error probability has been shown (Jake 1974) to be

$$p_{\text{avm}}(e) = \left[\frac{2(\gamma_0+1) + \gamma_0 J_0(2\pi f_d T)}{1+\gamma_0}\right]\left(\frac{1 + \gamma_0[1 - J_0(2\pi f_d T)]}{2(1+\gamma_0)}\right)^2. \quad (5.46)$$

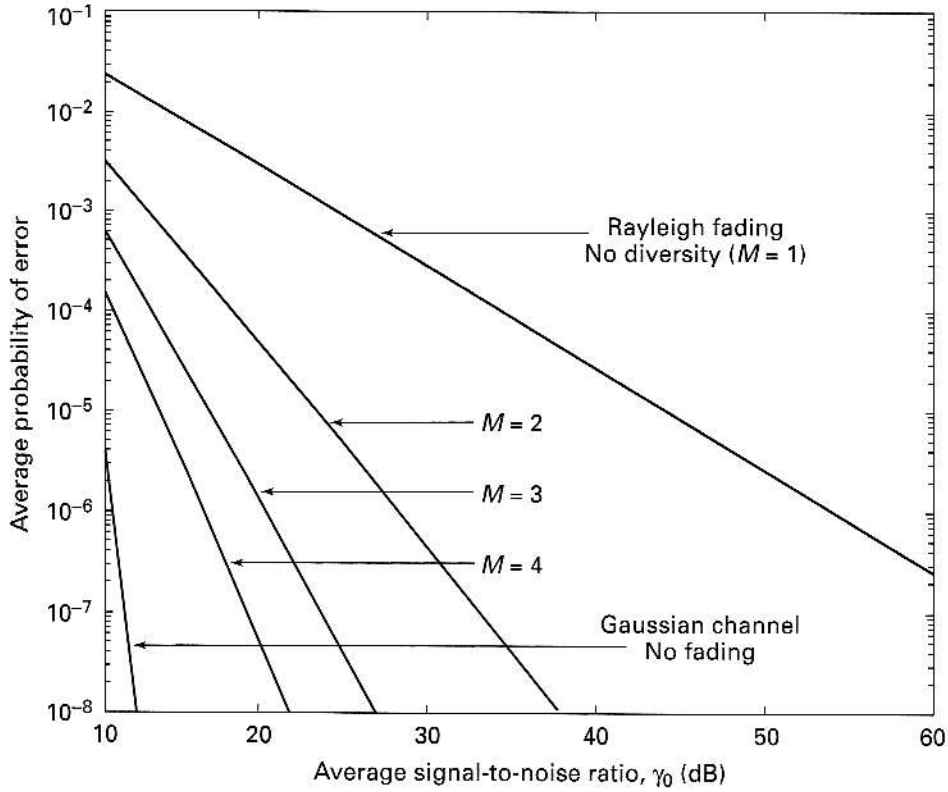

FIGURE 5.18 The average probability of error for the selection combiner for the case of coherent BPSK.

The results in Figure 5.22 show that even for a low value of f_dT (= 0.05), the error floor exists (10^{-2}) and diversity techniques can bring the error floor down to 10^{-3}. Diversity techniques also lead to a steeper drop in error rates as the signal-to-noise ratio values go up. The error floors seen here are typical of communications systems operating in a random FM environment, as we saw in Section 5.2.

The lowering of the error floor through the use of diversity techniques can be seen clearly if we plot the error floor values as a function of the parameter f_dT. These results are shown in Figure 5.23. The steep drop in error floor seen is evidence of the importance of diversity techniques in mobile communication systems.

5.6 MACROSCOPIC DIVERSITY

The diversity techniques described so far are sufficient to mitigate the effects of short-term fading, i.e., Rayleigh fading. The fading in this case occurs over very short distances; in the case of lognormal fading or shadowing, the fading is very slow and occurs over distances of many wavelengths. For the case of lognormal fading, the use of multiple receivers separated by short distances will not produce any difference between the multiple signals. To compensate for lognormal fading, the diversity technique must be implemented on a site-by-site basis, with the sites separated by many wavelengths. This form of diversity, where receivers are located at

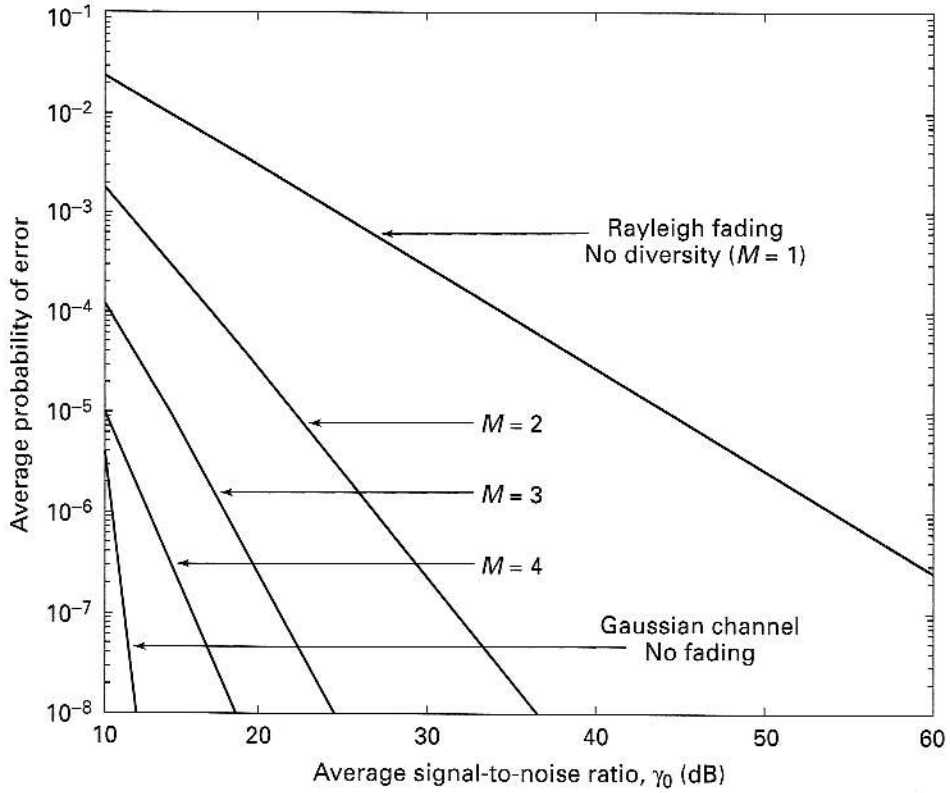

FIGURE 5.19 The average error probability for MRC diversity for the coherent BPSK receiver.

distances orders of magnitude larger than those in the case of microscopic diversity, is known as macroscopic diversity. Macroscopic diversity is also sometimes referred to as site diversity (or multiple-base-station diversity). One can visualize macroscopic diversity as a means by one site of collecting the signal from another site when the primary site is obstructed by a shadowing region. An example of this form of diversity exists in cellular systems: the hand-over between base stations can be treated as a form of macroscopic diversity.

The different combining techniques described in connection with microscopic diversity are also applicable to macroscopic diversity. A detailed description of macroscopic diversity techniques is given in Section C.3, Appendix C.

5.7 EQUALIZATION/FREQUENCY-SELECTIVE FADING

So far we have considered the attempts made to overcome problems associated with flat fading. If the channel is frequency selective, i.e., the coherent bandwidth happens to be less than the information bandwidth, pulse spreading takes place, leading to intersymbol interference, as shown in Figure B.6.2 (see Appendix B). ISI is a serious problem in high-speed data transmission. We could certainly reduce the data rate so as to reduce the pulse spreading, and, thus, reduce the effects of ISI. But this

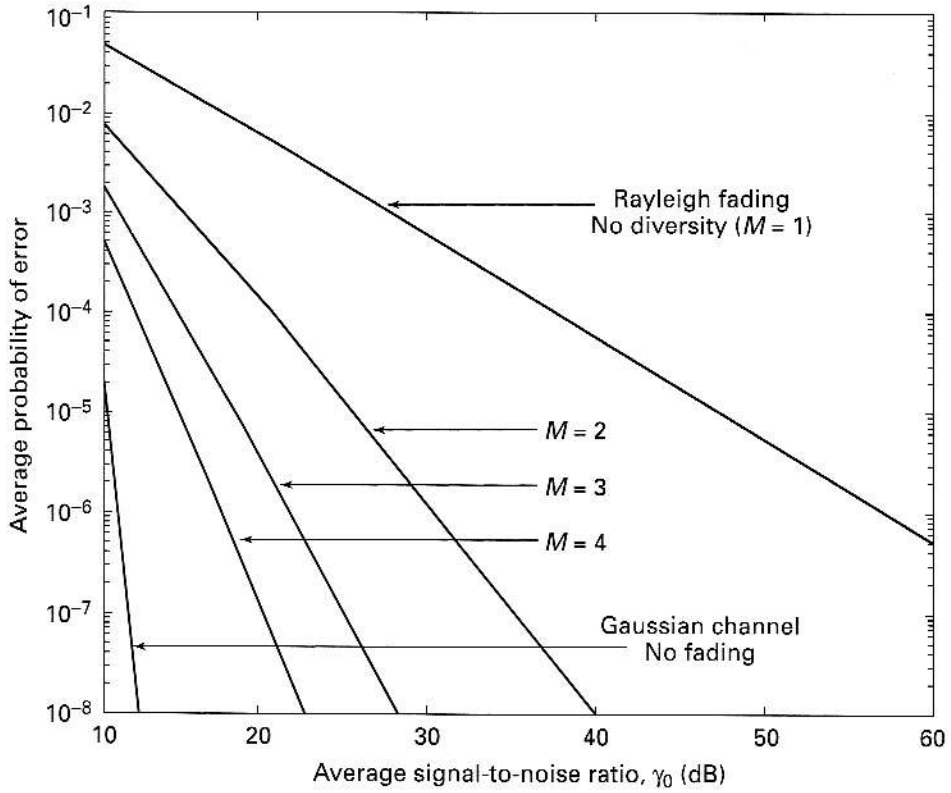

FIGURE 5.20 The average probability of error for selection diversity for the case of DPSK.

is not an option, since the goal is to transmit data at higher rates. One of the ways in which the frequency selectivity of the channel can be reduced or eliminated is through the process known as equalization (Mons 1980, Skla 1988, Proa 1995, Hayk 2001).

The transfer function of the channel, $H_c(f)$, can be expressed as

$$H_c(f) = A(f)\exp\,[\theta(f)], \tag{5.47}$$

where $A(f)$ is the envelope and $\theta(f)$ is the phase response. The envelope delay, $\tau(f)$, is expressed as

$$\tau(f) = -\frac{1}{2\pi}\frac{d\theta(f)}{df}. \tag{5.48}$$

The channel is distortionless if

$$A(f) = A_0 \tag{5.49}$$

and if

$$\tau(f) = \tau_0, \tag{5.50}$$

where A_0 and τ_0 are constants in the frequency range of interest. The latter condition means that the phase $\theta(f)$ must be linear. If $A(f)$ is not constant, we have envelope or amplitude distortion, and if $\tau(f)$ is not constant, we have delay distortion. The consequence of distortion is ISI.

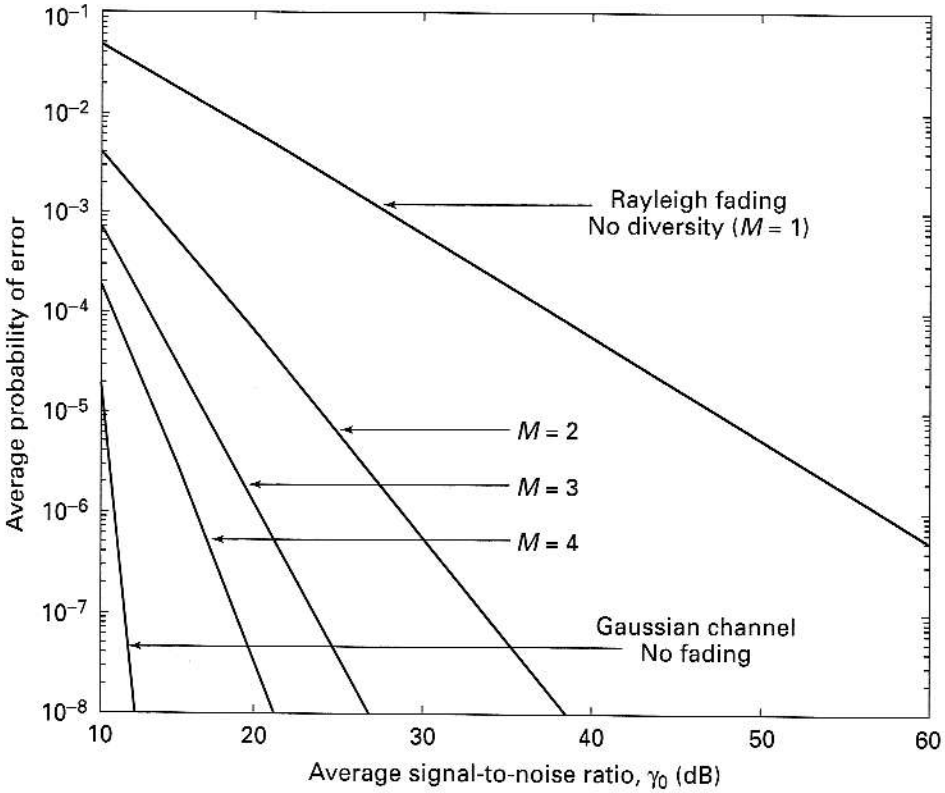

FIGURE 5.21 The average probability of error for MRC diversity for the case of DPSK.

To understand equalization in purely qualitative terms, let us look at the frequency response of a frequency-selective channel, shown in Figure 5.24. The ideal response is almost flat over the complete spectrum. The phase response is also linear for the ideal channel. If the channel impulse response could be altered to look like the ideal response shown in Figure 5.24, i.e., if it is possible to equalize the response at all frequencies, the pulse spreading would be eliminated. Note that pulse distortion also occurs if the phase response of the channel is nonlinear. An equalizer must correct the magnitude as well as the phase response of the system.

A simple conceptual block diagram of an equalizer is shown in Figure 5.25. The transmitting filter includes the modulator. The receiving filter includes the receiver. As explained in the previous paragraph, the ultimate goal of the equalizer is to ensure that the overall impulse response, when the channel and equalizer together are taken into account, is a delta function. This is realized if

$$H_{eq}(f)H_c^*(-f) = 1\,. \tag{5.51}$$

In other words, the equalizer is nothing but an inverse filter, as shown in Figure 5.26.

Even though the concept of equalization appears simple, there are practical problems associated with its implementation owing to the fact that the channel is time varying. This means that an equalizer must be able to adjust or retune to keep track with the changes in the channel characteristics. Thus the idea of a simple filter

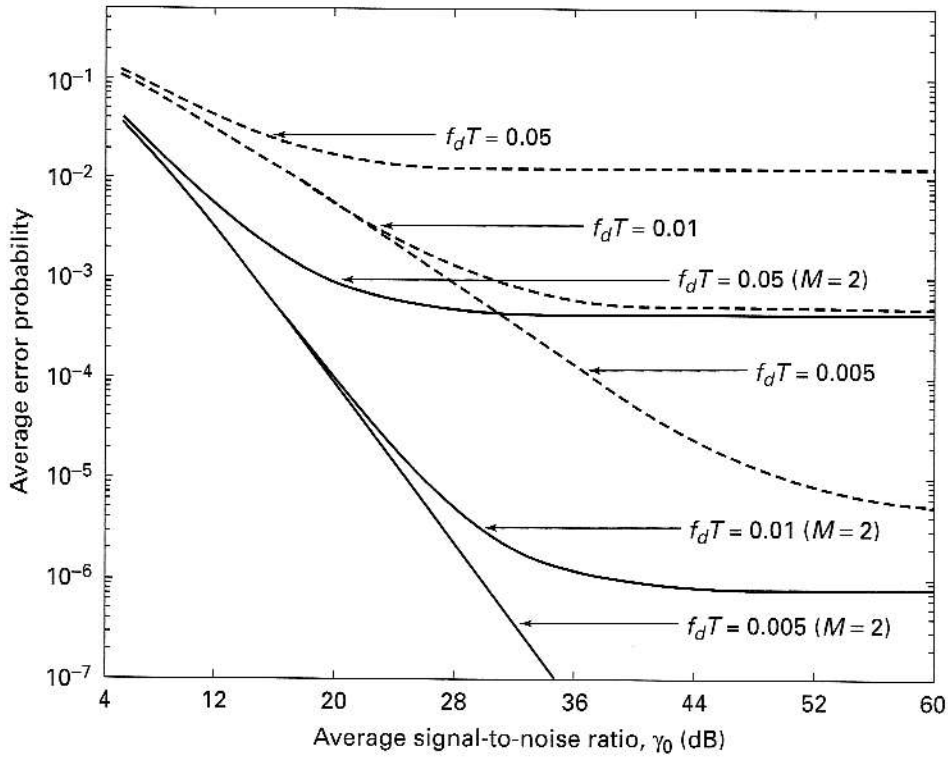

FIGURE 5.22 Results for a two-branch diversity receiver using the MRC scheme for a DPSK system ($M = 2$). Results are also shown for the case of no diversity.

with the required response after the demodulator/receiver filter is inadequate. Therefore, equalizers must be designed to be adaptive. We will briefly review some of the equalization techniques commonly used in wireless communications (Tayl 1998).

5.7.1 Linear Transversal Equalizer

A block diagram depicting a simple linear equalizer, more commonly known as a linear transversal equalizer, is shown in Figure 5.27. It consists of a number of elements, successively introducing a delay of τ. The maximum time delay τ is T. For the case where the delay is the symbol duration (T), the equalizer is known as a symbol-spaced equalizer.

Each delayed version of the input signal is weighted appropriately. The number of taps required is determined by the severity of ISI. If the ISI spreads a given symbol over many symbol periods, the number of required taps increases. It is possible to calculate the weights denoted by c_0, c_1, c_2, ... to ensure that the ISI is zero. The ISI approaches zero when the number of taps is infinite. This equalizer may be considered to be a zero-forcing equalizer (ZF algorithm), since in an ideal case, the inverse filter given in eq. (5.51) is expected to force the ISI to be zero at the sampling instants.

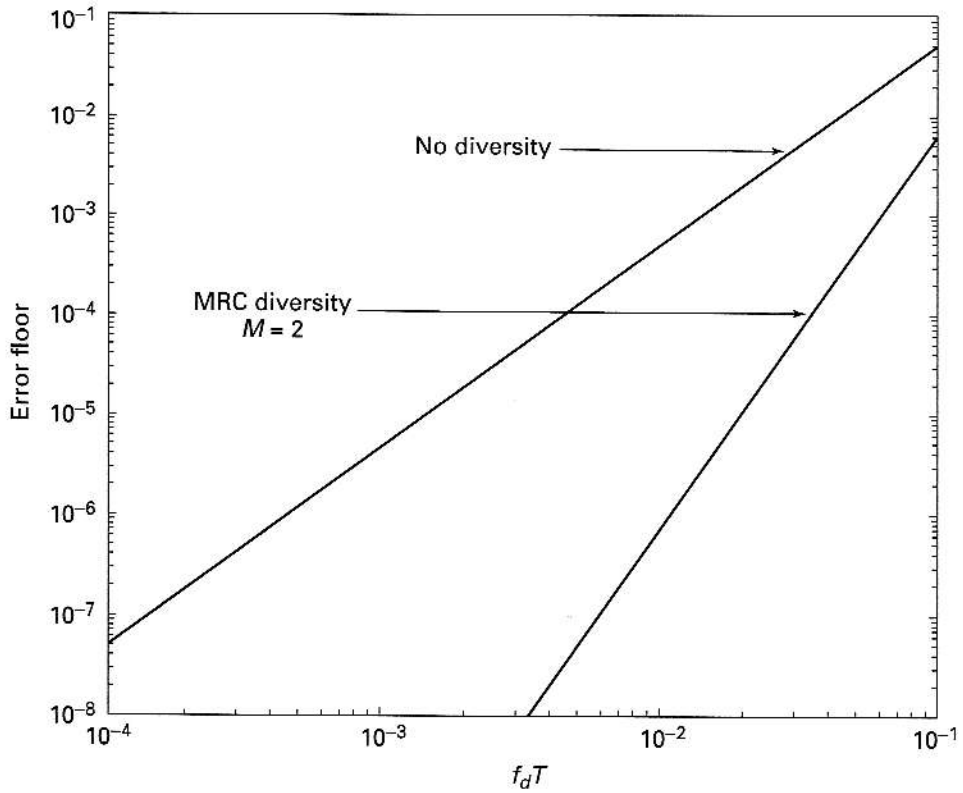

FIGURE 5.23 The error floor for a two-channel MRC diversity system for DPSK compared with the case of no diversity.

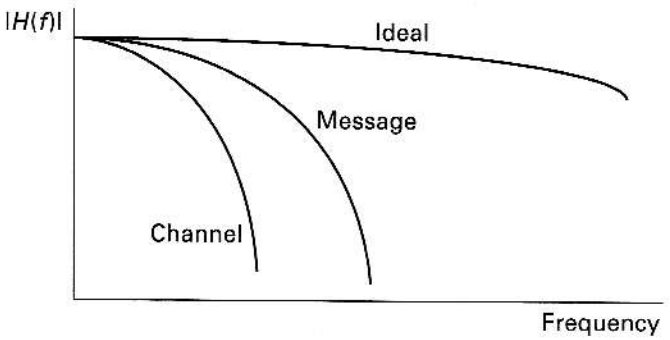

FIGURE 5.24 The ideal low-pass transfer function, spectrum of the message, and the transfer function of the frequency-selective channel.

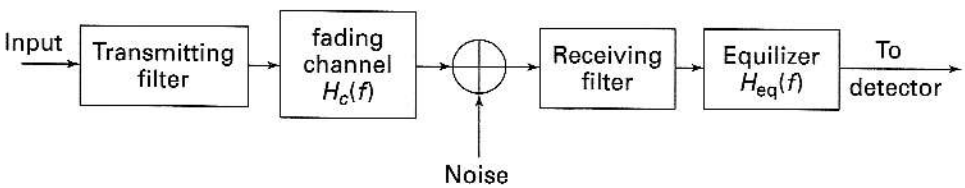

FIGURE 5.25 A conceptual block diagram of an equalizer.

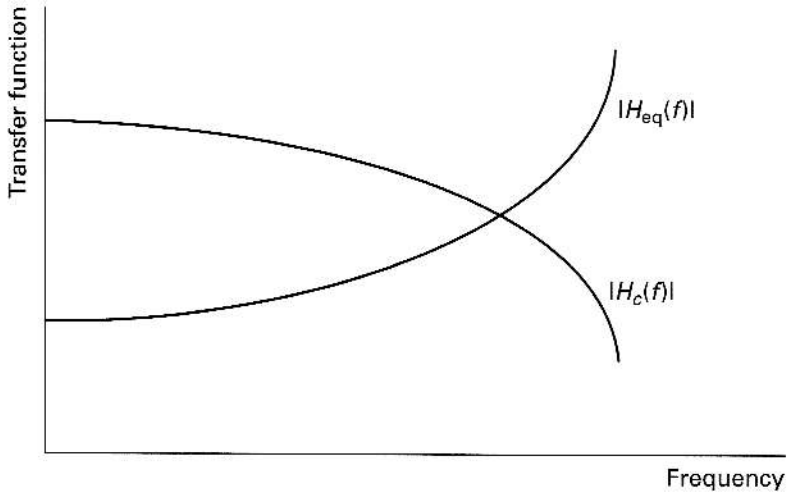

FIGURE 5.26 The overall low-pass transfer function of the channel, $H_c(f)$, in the absence of an equalizer and the transfer function of an equalizer, $H_{eq}(f)$.

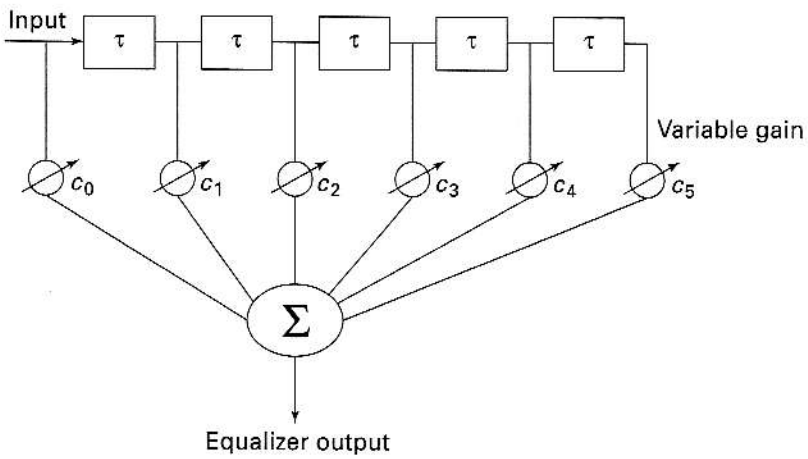

FIGURE 5.27 Linear transversal equalizer.

EXAMPLE 5.7

Consider a mobile channel with a channel response $h_c(t)$ given by

$$h_c(t) = a\delta(t - t_0) + b\delta(t - t_1),$$

where a and b are constants and t_0 and t_1 are time delays. If a delay line with three taps is used as an equalizer, calculate (1) the transfer function of the channel and (2) the transfer function of the "equalizer" and the relationship between a and b given that $b \ll a$ and $t_1 > t_0$. Assume that there is no noise.

Answer Taking the Fourier transform of $h_c(t)$, we obtain the transfer function, $H_c(f)$, of the channel:

$$H_c(f) = a\exp(-j2\pi f t_0) + b\exp(-j2\pi f t_1).$$

The equalizer transfer function, $H_{eq}(f)$, must be such that

$$H_{eq}(f)H_c^*(-f) = 1 .$$

(eq. (5.51)). For a three-tap equalizer that uses delays of T,

$$H_{eq}(f) = c_0 + c_1 \exp(-j2\pi fT) + c_2 \exp(-j2\pi f2T)$$

$$= c_0\left[1 + \frac{c_1}{c_0}\exp(-j2\pi fT) + \frac{c_2}{c_0}\exp(-j2\pi f2T)\right].$$

Using the equation for the "inverse filter,"

$$H_{eq}(f) = \frac{\exp(-j2\pi f\Lambda)}{H_c^*(-f)}$$

$$= \frac{\exp(-j2\pi f\Lambda)}{a\exp(-j2\pi ft_0) + b\exp(-j2\pi ft_1)} = \frac{(1/a)\exp\left[-j2\pi f(\Lambda - t_0)\right]}{1 + (b/a)\exp\left[-j2\pi f(t_1 - t_0)\right]}.$$

Since $b \ll 1$, we can rewrite this equation as

$$H_{eq}(f) = \left(\frac{1}{a}\right)\exp\left[-j2\pi f(\Lambda - t_0)\right] \times \left\{1 - \frac{b}{a}\exp\left[-j2\pi f(t_1 - t_0)\right] + \left(\frac{b}{a}\right)^2 \exp\left[-j4\pi f(t_1 - t_0)\right]\right\}.$$

If we assume that $\Lambda \approx t_0$, comparing the two equations we get

$$c_0 = \frac{1}{a}$$

$$\frac{c_1}{c_0} = -\frac{b}{a}$$

$$\frac{c_2}{c_0} = \left(\frac{b}{a}\right)^2$$

and

$$t_1 - t_0 = T .$$

The coefficients of the taps can now easily be found. ■

In contrast to a symbol-spaced equalizer, the time delay between adjacent taps may be made less than T. This is known as a fractionally spaced equalizer. Quite often the delay is chosen to be $T/2$. In this case, the impulse response of the equalizer is

$$H_{eqf}(f) = \sum_{n=-N}^{N} C_n \exp\left(-j2\pi fn\frac{T}{2}\right), \tag{5.52}$$

where there are $2N + 1$ taps. By choosing taps spaced by delays less than T, the chances of aliasing are eliminated, and hence, ISI is likely to be compensated for more fully. The number of taps is chosen such that it exceeds the number of symbols spanned by ISI, as shown in Figure 5.28.

Note that this type of equalizer is typically operated by sending a training data set to calculate the weights. However, such an approach is not adaptive. It is possible to make the equalizer adaptive by continuously updating the weights. This can be

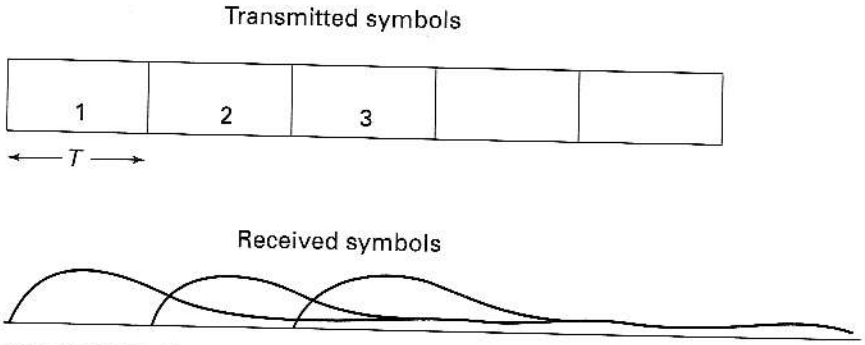

FIGURE 5.28 Transmitted symbols and received symbols, showing the "overflowing" nature of ISI.

accomplished by observing the equalized bits on a block-by-block basis and comparing them to some known or expected signal. The weights are then adjusted to minimize the mean square error (MMSE algorithm).

5.7.2 Nonlinear Equalizer

It is possible that the distortion introduced by the channel is too severe to be compensated for through the use of a linear transversal equalizer. To understand the problem with linear equalizers, let us examine the impulse responses of three channels shown in Figure 5.29. The first one can be considered to correspond to low-ISI channels such as telephone lines, while the other two correspond to the wireless channels. Their transfer functions can be approximated as shown in Figures 5.30*a*, *b*, and *c*, respectively.

The existence of a null (Figure 5.30*c*) and the sharp dropoff (Figure 5.30*b*) show the presence of significant amounts of ISI, and the linear equalizer may not be able to compensate for it; the frequency response in Figure 5.30*a* points to a low amount of ISI, which the linear equalizer should be able to handle. Another drawback of the linear equalizer may be amplification of noise present around the spectral null, since a linear equalizer will introduce a large gain to compensate for the spectral null.

Severe cases of ISI can be handled using nonlinear equalizers (Tayl 1998). Nonlinear techniques include decision feedback equalization (DFE), maximum likelihood symbol detection (MLSD), and maximum likelihood sequence estimation (MLSE). We will briefly examine two of these techniques.

Decision Feedback Equalizer (DFE) To combat severe distortion brought on by frequency-selective fading, DFE uses a feedback mechanism to eliminate the ISI caused by previously detected symbols interfering with the current symbols being detected.

A block diagram of the DFE is shown in Figure 5.31. The input information is passed through a linear transversal filter like the one discussed earlier. This filter is identified as the *feedforward* filter and is normally a fractionally spaced equalizer. The output of this filter is applied to a second filter, the *feedback* filter, which is a symbol-spaced equalizer with adjustable tap gains.

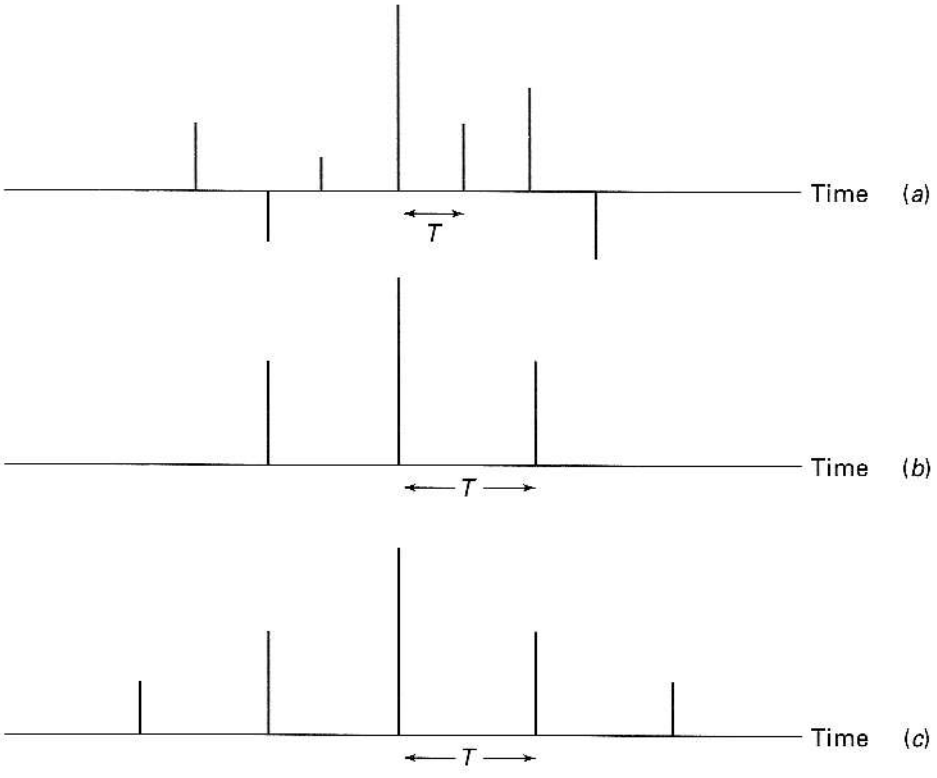

FIGURE 5.29 Impulse responses of three channels.

The input to the feedback filter is the previously detected symbols. The feedback filter subtracts the portion of ISI produced by these symbols. The equalizer is nonlinear because of the existence of the decision device in the feedback loop.

DFE can compensate for moderate to severe amounts of ISI.

Maximum Likelihood Sequence Estimation (MLSE) Even though the decision feedback equalizer is better than the linear transversal equalizer in its ability to mitigate ISI caused by fading, it is not the optimal processor for minimizing the probability of error in the detection of the transmitted symbols from the received symbol sequences. The MLSE equalizer estimates the channel impulse response and uses the Viterbi algorithm (Vite 1971, Forn 1973) to generate estimates of the transmitted symbols. A block diagram of the MLSE equalizer is shown in Figure 5.32. It consists of a channel estimator and the MLSE algorithm (Viterbi algorithm). The channel estimator is typically an adaptive one and estimates the channel response. The MLSE receiver performs a correlation of the received signal with all possible transmitted symbol sequences and chooses the optimal one based on the maximum likelihood procedure. The Viterbi algorithm is a computationally efficient means to accomplish this task.

The MLSE technique is more efficient than DFE; however, MLSE is computationally very exhaustive.

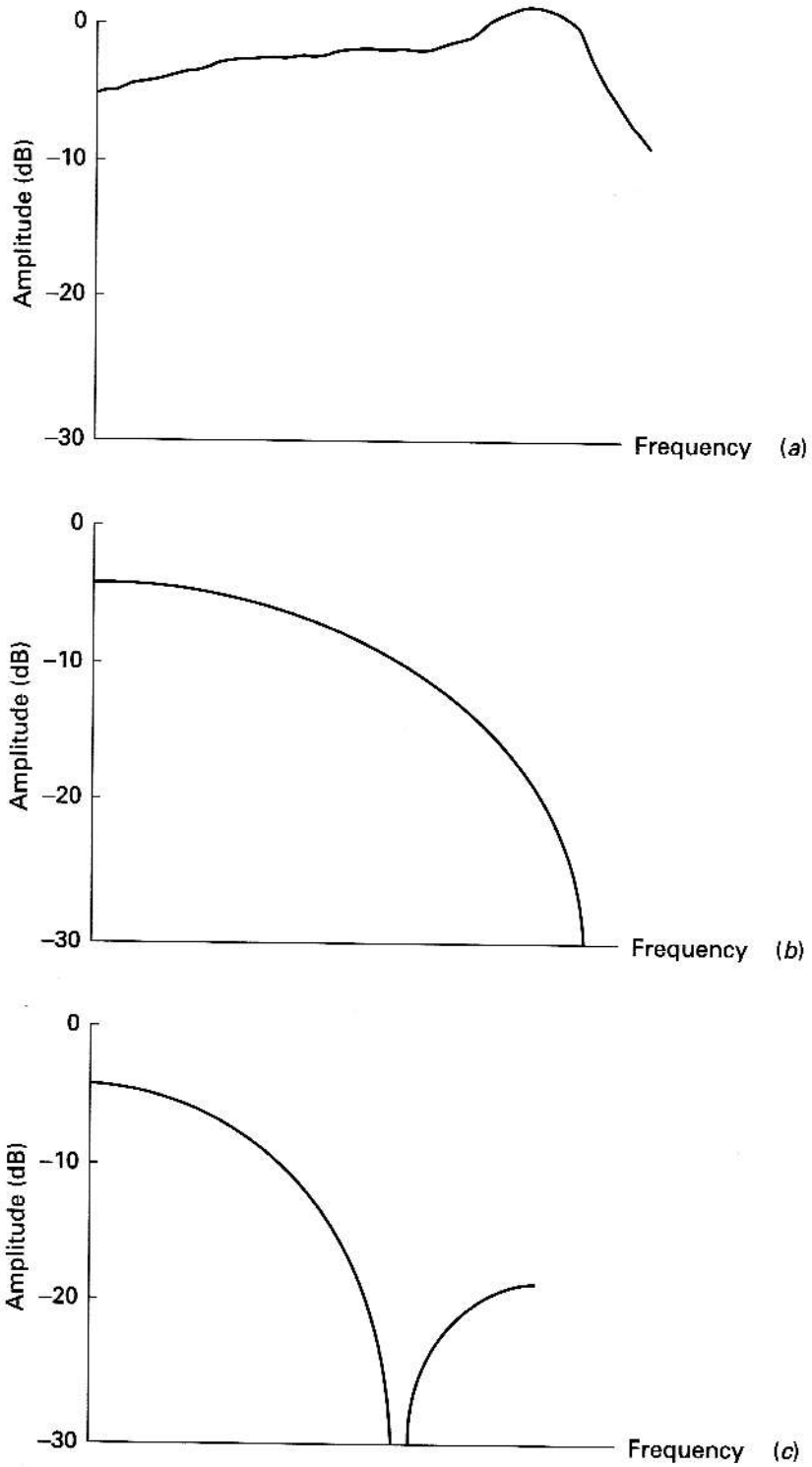

FIGURE 5.30 Transfer functions corresponding to the three impulse responses shown in Figure 5.29.

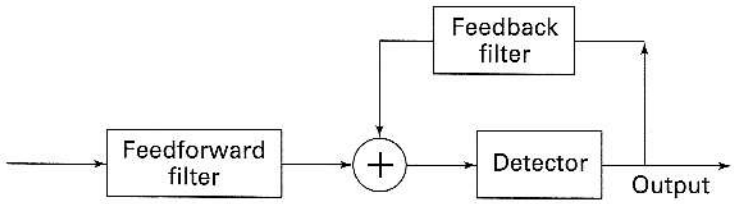

FIGURE 5.31 Block diagram of the decision feedback equalizer.

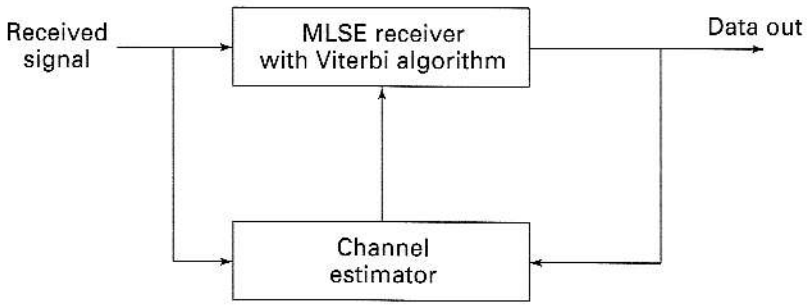

FIGURE 5.32 Block diagram of a MLSE equalizer.

5.8 SUMMARY

The effects of fading on data transmission have been studied in this chapter. The random nature of the received power due to fading increases the probability of error. In frequency-selective fading, the probability of error reaches a "floor" beyond which any increase in signal-to-noise ratio does not seem to lower the bit error rate. Similar effects are observed in the presence of random frequency modulation induced by relative motion of the transmitter with regard to the receiver.

The effects of flat fading may be reduced using diversity techniques. In diversity, multiple versions of the signal are created and then combined using some algorithm to result in a new signal that has an acceptable bit error rate. Other fading mitigation approaches use a multifrequency transmission technique known as orthogonal frequency division multiplexing (OFDM). A brief discussion of OFDM is given in Appendix B, Section B.1.

- The additional signal-to-noise ratio necessary to keep the bit error rate in fading channels the same as in unfaded channels (Gaussian channels) is referred to as the fading margin.
- Fading margins may be reduced using diversity techniques.
- Diverse signals may be created using the following methods: spatial diversity, frequency diversity, polarization diversity, and time diversity.
- These diverse signals must be uncorrelated to result in improved performance.
- Diverse signals may be combined using selection diversity, maximal ratio combining, and equal gain diversity.
- The maximal ratio combining algorithm has the best performance, followed by equal gain diversity and selection diversity.
- Some modulation schemes show better fading mitigation than others.

- Performance improvement goes up with an increase in the number of diverse signals.
- RAKE diversity operates on the basis of combining delayed versions of the signals arriving through a multipath. For this form of diversity to work, the duration of the pulse must be very short so that multiple versions of the signals coming through the various paths are resolvable.
- Macroscopic diversity is implemented to mitigate the effects of long-term fading. In this case, the multiple receivers must be far apart.
- Frequency-selective fading can be mitigated through the use of equalizers.

PROBLEMS

*** *Asterisks refer to problems better suited for graduate-level students.*

1. The instantaneous signal-to-noise ratio in a Rayleigh channel is given by A^2/N_0, where N_0 is the noise power. The average signal-to-noise ratio expected at the receiver is $\langle A^2 \rangle / N_0 = 3$ dB. A is Rayleigh distributed. To mitigate the effects of Rayleigh fading, the service provider is planning to use a six-channel diversity receiver with MRC. If outage occurs when the instantaneous SNR goes below 0 dB, calculate the outage probability with and without use of the diversity system.

2. In Problem 1, if the diversity is instead based on selection combining, calculate the outage probability.

3. If selection combining is to be carried out on the basis of envelopes instead of powers, derive an expression for the pdf of the selection diversity of a system with M diverse channels.

4. Plot the density functions of the selection combiner of Problem 3 for $M = 1, 3$, and 6.

5. Derive the expression for the selection combiner using order statistics.

6. Using MATLAB, generate 10 sets of Rayleigh-distributed random variables. Plot the samples and show that selection diversity will improve the performance. If the performance is defined in terms of U = mean/std. dev., compute U with and without diversity.

7. Using MATLAB, plot the bit error rates for the Rayleigh-faded and unfaded cases for coherent BPSK. For error rates equal to 10^{-3} and 10^{-4}, calculate the power margin.

8. Consider a transmission system where the received signal $r(t)$ is expressed as

$$r(t) = As(t) + n(t),$$

where $s(t)$ is the transmitted signal and $n(t)$ is white Gaussian noise of power spectral density $N_0/2$. The parameter A is a scaling factor that is a random variable (similar to fading) having the following pdf:

$$f(a) = 0.5, \quad 0 \le a \le 2.$$

Calculate the average probability of error if a matched filter is used at the receiver. Assume that the modulation is BPSK.

9. The bit error rate for BPSK in the presence of phase mismatch is given by $\frac{1}{2}\text{erfc}(\sqrt{z}\cos\phi)$ (see Chapter 3). Proceed in a manner similar to Problem 8, considering the phase to be random (analogous to fading). If the phase is uniform in the range $(-20°, 20°)$, compare the bit error rates when ϕ is zero and when ϕ is random.

10. Evaluate the average error probability for a BPSK receiver for the case of a Rician-faded channel. Use MATLAB and plot the average BER for values of K ranging from –10 to 20 dB in steps of 5 dB. Comment on your results.

11. Evaluate the average error probability for a DPSK receiver for the case of a Rician-faded channel. Use MATLAB and plot the average BER for values of K ranging from –10 to 20 dB in steps of 5 dB. Comment on your results.

12. In Problem 8, if you are asked to use a diversity receiver to mitigate fading present in the signal, calculate the performance improvement arising from selection diversity.

13. Repeat Problem 8 if the pdf of a is given by

$$f(a) = 0.1\,\delta(a) + 0.5\,\delta(a-1) + 0.4\,\delta(a-2).$$

14. In Problem 13, if you are asked to use a diversity receiver ($M = 2$) to mitigate fading in the signal,

calculate the performance improvement arising from selection diversity. (Proceed as in Problem 12.)

15. In Problem 13, if you are asked to use a diversity receiver ($M = 2$) to mitigate fading in the signal, calculate the performance improvement arising from equal gain diversity. (Proceed as in Problem 12.)

16. Derive the expression for the bit error rate for a DPSK system in the presence of Nakagami fading.

17. Plot the error rate for a BPSK system in the presence of Nakagami fading. Compare the results to those of Problem 10.

18. Derive the expression for the bit error rate in Rayleigh fading for BPSK.

19. Derive the expression for the bit error rate in Rayleigh fading for DPSK.

20. Repeat Problem 6 using a set of exponentially distributed random variables.

21. Compare the outages after diversity combining, using MRC for $M = 3$ and 4. The average signal-to-noise ratio of a single channel is 5 dB. The threshold signal-to-noise ratio required is 0 dB.

22. Repeat Problem 21 if the selection diversity technique is used.

23. Use MATLAB to compare outages for MRC and selection diversity. Use an average SNR of 3 dB and a threshold of 0 dB. Consider M of 1 through 8.

24. Explain the concept of excess power margin in the context of the modulation scheme, and coherent BPSK in the absence of fading and in the presence of fading. Generate a MATLAB plot of bit error rate versus excess power margin (dB).

25. Explain the concept of excess power margin in the context of the modulation scheme, and DPSK in the absence of fading and in the presence of fading. Generate a MATLAB plot of bit error rate versus excess power margin (dB).

26. Explain why the performance of digital communication systems in Rician fading is better than in Rayleigh fading.

27. Examine the relationship between power efficiency and spectral efficiency in the presence of Rayleigh fading using MATLAB (see Figure B.5.1 in Appendix B).***

28. Examine the relationship between efficiency and spectral efficiency in Rayleigh fading when diversity is used to mitigate fading (see Figure B.5.1 in Appendix B).***

29. Examine the relationship between power efficiency and spectral efficiency in the presence of Nakagami fading. Note that the signal-to-noise ratio z in Nakagami fading has the distribution given by

$$f(z) = \left(\frac{m}{z_0}\right)^m \frac{z^{m-1}}{\Gamma(m)} \exp\left(-m\frac{z}{z_0}\right) U(z),$$

where z_0 is the average signal-to-noise ratio and m is the Nakagami parameter. Use $m = 1, 1.5, 2, 5, 50$. Comment on the results and see what happens when $m = 50$.***

30. Based on the results of Problem 29, comment on the relationship between power efficiency and spectral efficiency when fading is Rician.***

31. Use the results of Problems 16 and 17 to show that diversity will improve the performance of the communication system. (Hint: Use the results showing that the sum of two Nakagami random variables, each with a parameter m, is another Nakagami random variable, with a parameter $2m$.)***

32. Generate a BPSK waveform subject to Rayleigh fading and noise. Demodulate the waveform in a manner similar to that used in the problem in Chapter 3. Compare the recovered data stream in conjunction with fading and no fading.***

33. Generate a number (M) of BPSK waveforms subject to fading and noise. Demodulate them. Combine the detected signals by adding them (a form of equal gain diversity) and then averaging the resultant waveform. For $M = 3, 5, 10, 15$, examine the results and compare them with those of Problem 32.***