

Proakis, J.G. "Channel Equalization"
Mobile Communications Handbook
Ed. Suthan S. Suthersan
Boca Raton: CRC Press LLC, 1999

Channel Equalization

- 5.1 Characterization of Channel Distortion
- 5.2 Characterization of Intersymbol Interference
- 5.3 Linear Equalizers
 - Adaptive Linear Equalizers
- 5.4 Decision-Feedback Equalizer
- 5.5 Maximum-Likelihood Sequence Detection
- 5.6 Conclusions
- Defining Terms
- References
- Further Information

John G. Proakis
Northeastern University

5.1 Characterization of Channel Distortion

Many communication channels, including telephone channels, and some radio channels, may be generally characterized as band-limited linear filters. Consequently, such channels are described by their frequency response $C(f)$, which may be expressed as

$$C(f) = A(f)e^{j\theta(f)} \quad (5.1)$$

where $A(f)$ is called the *amplitude response* and $\theta(f)$ is called the *phase response*. Another characteristic that is sometimes used in place of the phase response is the *envelope delay* or *group delay*, which is defined as

$$\tau(f) = -\frac{1}{2\pi} \frac{d\theta(f)}{df} \quad (5.2)$$

A channel is said to be nondistorting or ideal if, within the bandwidth W occupied by the transmitted signal, $A(f) = \text{const}$ and $\theta(f)$ is a linear function of frequency [or the envelope delay $\tau(f) = \text{const}$]. On the other hand, if $A(f)$ and $\tau(f)$ are not constant within the bandwidth occupied by the transmitted signal, the channel distorts the signal. If $A(f)$ is not constant, the distortion is called *amplitude distortion* and if $\tau(f)$ is not constant, the distortion on the transmitted signal is called *delay distortion*.

As a result of the amplitude and delay distortion caused by the nonideal channel frequency response characteristic $C(f)$, a succession of pulses transmitted through the channel at rates comparable to the bandwidth W are smeared to the point that they are no longer distinguishable as well-defined pulses at the receiving terminal. Instead, they overlap and, thus, we have **intersymbol interference (ISI)**. As an example of the effect of delay distortion on a transmitted pulse, Fig. 5.1(a) illustrates a band-limited pulse having zeros periodically spaced in time at points labeled $\pm T$, $\pm 2T$, etc. If information

is conveyed by the pulse amplitude, as in pulse amplitude modulation (PAM), for example, then one can transmit a sequence of pulses, each of which has a peak at the periodic zeros of the other pulses. Transmission of the pulse through a channel modeled as having a linear envelope delay characteristic $\tau(f)$ [quadratic phase $\theta(f)$], however, results in the received pulse shown in Fig. 5.1(b) having zero crossings that are no longer periodically spaced. Consequently a sequence of successive pulses would be smeared into one another, and the peaks of the pulses would no longer be distinguishable. Thus, the channel delay distortion results in intersymbol interference. As will be discussed in this chapter, it is possible to compensate for the nonideal frequency response characteristic of the channel by use of a filter or equalizer at the demodulator. Figure 5.1(c) illustrates the output of a linear equalizer that compensates for the linear distortion in the channel.

The extent of the intersymbol interference on a telephone channel can be appreciated by observing a frequency response characteristic of the channel. Figure 5.2 illustrates the measured average amplitude and delay as a function of frequency for a medium-range (180–725 mi) telephone channel of the switched telecommunications network as given by Duffy and Tratcher, 1971. We observe that the usable band of the channel extends from about 300 Hz to about 3000 Hz. The corresponding impulse response of the average channel is shown in Fig. 5.3. Its duration is about 10 ms. In comparison, the transmitted symbol rates on such a channel may be of the order of 2500 pulses or symbols per second. Hence, intersymbol interference might extend over 20–30 symbols.

Besides telephone channels, there are other physical channels that exhibit some form of time dispersion and, thus, introduce intersymbol interference. Radio channels, such as short-wave ionospheric propagation (HF), tropospheric scatter, and mobile cellular radio are three examples of time-dispersive wireless channels. In these channels, time dispersion and, hence, intersymbol interference is the result of multiple propagation paths with different path delays. The number of paths and the relative time delays among the paths vary with time and, for this reason, these radio channels are usually called time-variant multipath channels. The time-variant multipath conditions give rise to a wide variety of frequency response characteristics. Consequently, the frequency response characterization that is used for telephone channels is inappropriate for time-variant multipath channels. Instead, these radio channels are characterized statistically in terms of the scattering function, which, in brief, is a two-dimensional representation of the average received signal power as a function of relative time delay and Doppler frequency (see Proakis [4]).

For illustrative purposes, a scattering function measured on a medium-range (150 mi) tropospheric scatter channel is shown in Fig. 5.4. The total time duration (multipath spread) of the channel response is approximately $0.7 \mu\text{s}$ on the average, and the spread between half-power points in Doppler frequency is a little less than 1 Hz on the strongest path and somewhat larger on the other paths. Typically, if one is transmitting at a rate of 10^7 symbols/s over such a channel, the multipath spread of $0.7 \mu\text{s}$ will result in intersymbol interference that spans about seven symbols.

5.2 Characterization of Intersymbol Interference

In a digital communication system, channel distortion causes intersymbol interference, as illustrated in the preceding section. In this section, we shall present a model that characterizes the ISI. The digital modulation methods to which this treatment applies are PAM, phase-shift keying (PSK) and quadrature amplitude modulation (QAM). The transmitted signal for these three types of modulation may be expressed as

$$\begin{aligned} s(t) &= v_c(t) \cos 2\pi f_c t - v_s(t) \sin 2\pi f_c t \\ &= \text{Re} \left[v(t) e^{j2\pi f_c t} \right] \end{aligned} \quad (5.3)$$

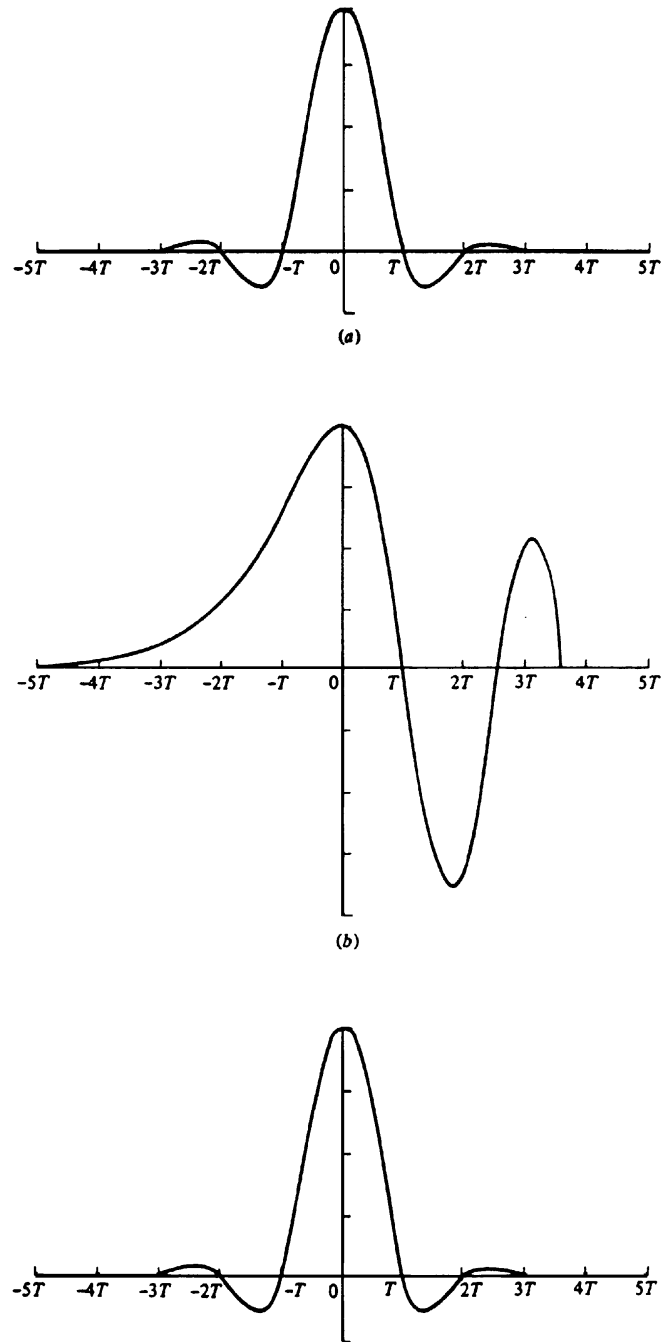


FIGURE 5.1: Effect of channel distortion: (a) channel input, (b) channel output, (c) equalizer output.

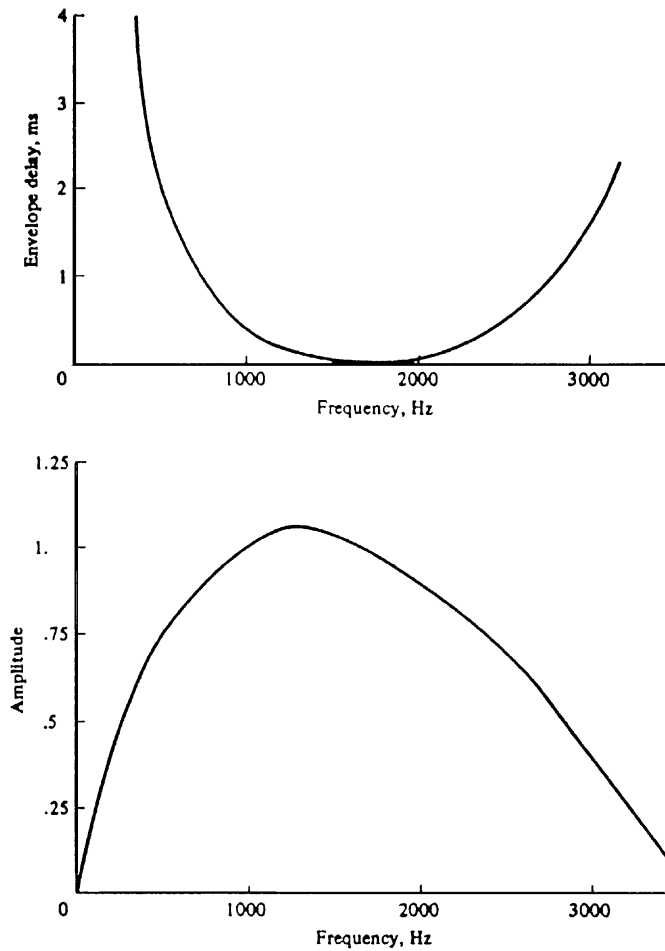


FIGURE 5.2: Average amplitude and delay characteristics of medium-range telephone channel.

where $v(t) = v_c(t) + jv_s(t)$ is called the *equivalent low-pass signal*, f_c is the carrier frequency, and $\text{Re}[\]$ denotes the real part of the quantity in brackets.

In general, the equivalent low-pass signal is expressed as

$$v(t) = \sum_{n=0}^{\infty} I_n g_T(t - nT) \quad (5.4)$$

where $g_T(t)$ is the basic pulse shape that is selected to control the spectral characteristics of the transmitted signal, $\{I_n\}$ the sequence of transmitted information symbols selected from a signal constellation consisting of M points, and T the signal interval ($1/T$ is the symbol rate). For PAM, PSK, and QAM, the values of I_n are points from M -ary signal constellations. Figure 5.5 illustrates the signal constellations for the case of $M = 8$ signal points. Note that for PAM, the signal constellation is one dimensional. Hence, the equivalent low-pass signal $v(t)$ is real valued, i.e., $v_s(t) = 0$ and $v_c(t) = v(t)$. For M -ary ($M > 2$) PSK and QAM, the signal constellations are two dimensional and, hence, $v(t)$ is complex valued.

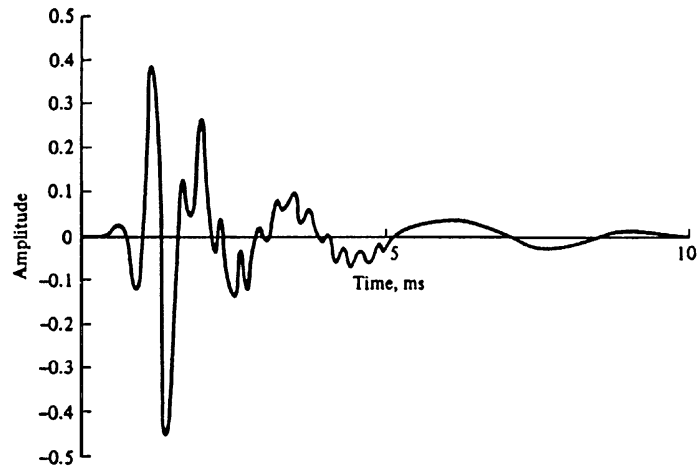


FIGURE 5.3: Impulse response of average channel with amplitude and delay shown in Fig. 5.2.

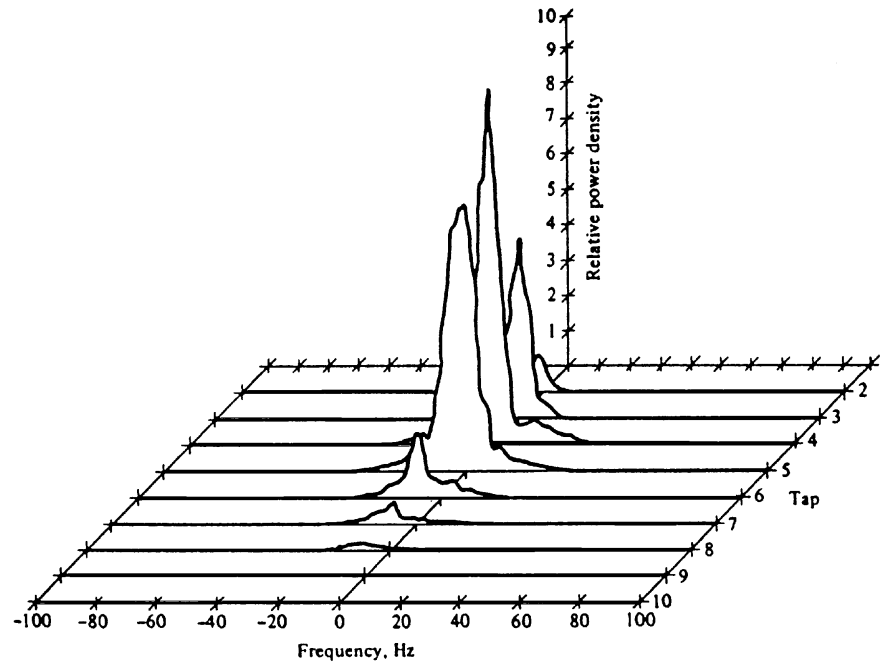
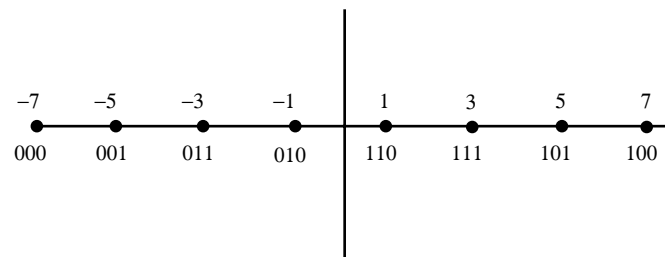


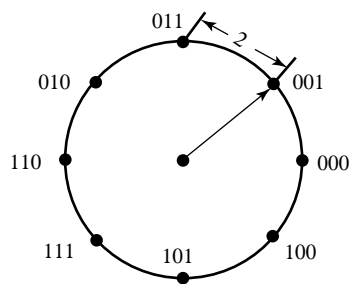
FIGURE 5.4: Scattering function of a medium-range tropospheric scatter channel.

The signal $s(t)$ is transmitted over a bandpass channel that may be characterized by an equivalent low-pass frequency response $C(f)$. Consequently, the equivalent low-pass received signal can be represented as

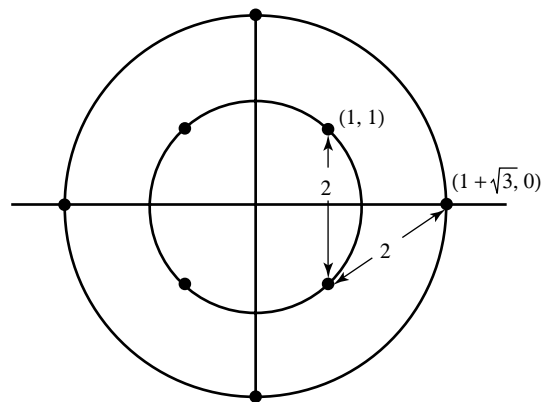
$$r(t) = \sum_{n=0}^{\infty} I_n h(t - nT) + w(t) \quad (5.5)$$



(a) PAM



(b) PSK



(c) QAM

FIGURE 5.5: $M = 8$ signal constellations for PAM, PSK, and QAM.

where $h(t) = g_T(t) * c(t)$, and $c(t)$ is the impulse response of the equivalent low-pass channel, the asterisk denotes convolution, and $w(t)$ represents the additive noise in the channel.

To characterize the ISI, suppose that the received signal is passed through a receiving filter and then sampled at the rate $1/T$ samples/s. In general, the optimum filter at the receiver is matched to the received signal pulse $h(t)$. Hence, the frequency response of this filter is $H^*(f)$. We denote its output as

$$y(t) = \sum_{n=0}^{\infty} I_n x(t - nT) + v(t) \quad (5.6)$$

where $x(t)$ is the signal pulse response of the receiving filter, i.e., $X(f) = H(f)H^*(f) = |H(f)|^2$, and $v(t)$ is the response of the receiving filter to the noise $w(t)$. Now, if $y(t)$ is sampled at times $t = kT$, $k = 0, 1, 2, \dots$, we have

$$\begin{aligned} y(kT) \equiv y_k &= \sum_{n=0}^{\infty} I_n x(kT - nT) + v(kT) \\ &= \sum_{n=0}^{\infty} I_n x_{k-n} + v_k, \quad k = 0, 1, \dots \end{aligned} \quad (5.7)$$

The sample values $\{y_k\}$ can be expressed as

$$y_k = x_0 \left(I_k + \frac{1}{x_0} \sum_{\substack{n=0 \\ n \neq k}}^{\infty} I_n x_{k-n} \right) + v_k, \quad k = 0, 1, \dots \quad (5.8)$$

The term x_0 is an arbitrary scale factor, which we arbitrarily set equal to unity for convenience. Then

$$y_k = I_k + \sum_{\substack{n=0 \\ n \neq k}}^{\infty} I_n x_{k-n} + v_k \quad (5.9)$$

The term I_k represents the desired information symbol at the k th sampling instant, the term

$$\sum_{\substack{n=0 \\ n \neq k}}^{\infty} I_n x_{k-n} \quad (5.10)$$

represents the ISI, and v_k is the additive noise variable at the k th sampling instant.

The amount of ISI, and noise in a digital communications system can be viewed on an oscilloscope. For PAM signals, we can display the received signal $y(t)$ on the vertical input with the horizontal sweep rate set at $1/T$. The resulting oscilloscope display is called an *eye pattern* because of its resemblance to the human eye. For example, Fig. 5.6 illustrates the eye patterns for binary and four-level PAM modulation. The effect of ISI is to cause the eye to close, thereby reducing the margin for additive noise to cause errors. Figure 5.7 graphically illustrates the effect of ISI in reducing the opening of a binary eye. Note that intersymbol interference distorts the position of the zero crossings and causes a reduction in the eye opening. Thus, it causes the system to be more sensitive to a synchronization error.

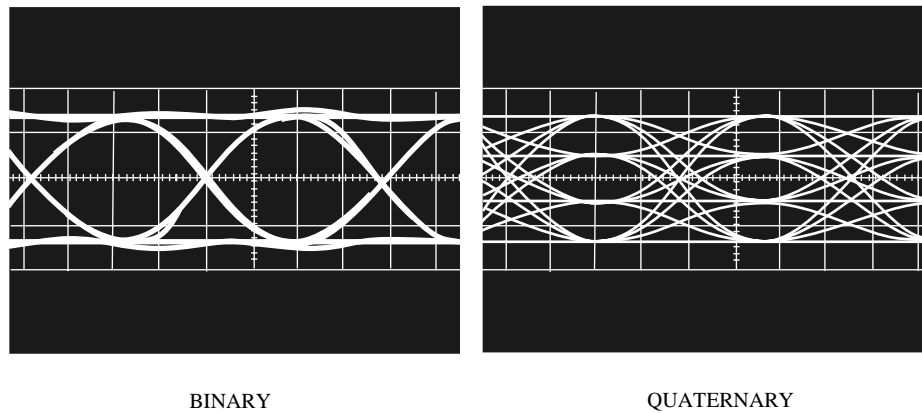


FIGURE 5.6: Examples of eye patterns for binary and quaternary amplitude shift keying (or PAM).

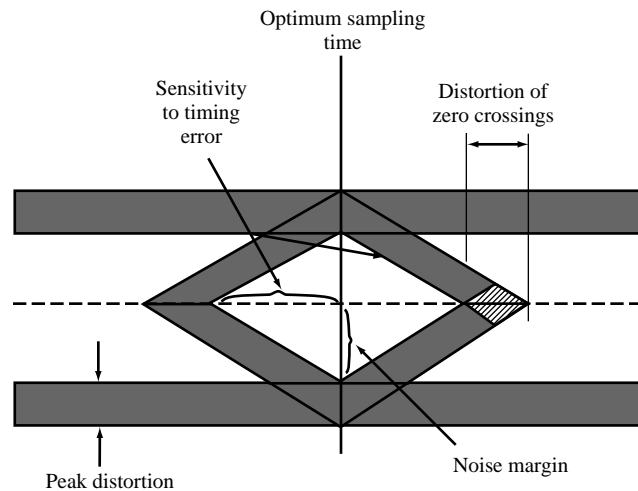


FIGURE 5.7: Effect of intersymbol interference on eye opening.

For PSK and QAM it is customary to display the eye pattern as a two-dimensional scatter diagram illustrating the sampled values $\{y_k\}$ that represent the decision variables at the sampling instants. Figure 5.8 illustrates such an eye pattern for an 8-PSK signal. In the absence of intersymbol interference and noise, the superimposed signals at the sampling instants would result in eight distinct points corresponding to the eight transmitted signal phases. Intersymbol interference and noise result in a deviation of the received samples $\{y_k\}$ from the desired 8-PSK signal. The larger the intersymbol interference and noise, the larger the scattering of the received signal samples relative to the transmitted signal points.

In practice, the transmitter and receiver filters are designed for zero ISI at the desired sampling times $t = kT$. Thus, if $G_T(f)$ is the frequency response of the transmitter filter and $G_R(f)$ is the frequency response of the receiver filter, then the product $G_T(f) G_R(f)$ is designed to yield zero ISI.

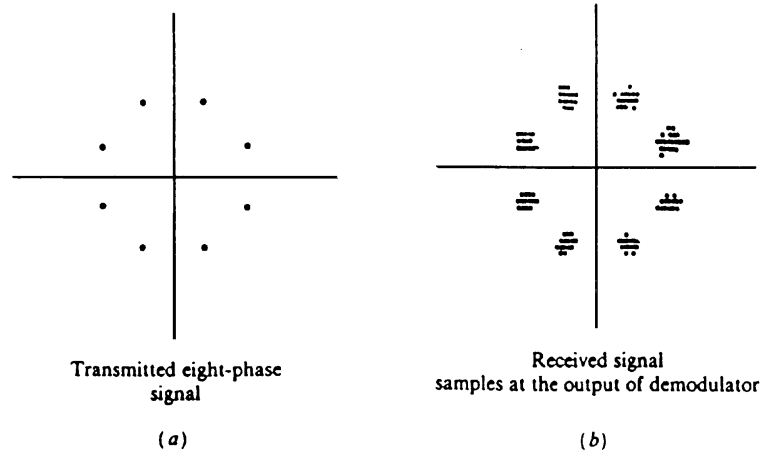


FIGURE 5.8: Two-dimensional digital eye patterns.

For example, the product $G_T(f) G_R(f)$ may be selected as

$$G_T(f)G_R(f) = X_{rc}(f) \quad (5.11)$$

where $X_{rc}(f)$ is the raised-cosine frequency response characteristic, defined as

$$X_{rc}(f) = \begin{cases} T, & 0 \leq |f| \leq (1 - \alpha)/2T \\ \frac{T}{2} \left[1 + \cos \frac{\pi T}{\alpha} \left(|f| - \frac{1 - \alpha}{2T} \right) \right], & \frac{1 - \alpha}{2T} \leq |f| \leq \frac{1 + \alpha}{2T} \\ 0, & |f| > \frac{1 + \alpha}{2T} \end{cases} \quad (5.12)$$

where α is called the *rolloff factor*, which takes values in the range $0 \leq \alpha \leq 1$, and $1/T$ is the symbol rate. The frequency response $X_{rc}(f)$ is illustrated in Fig. 5.9(a) for $\alpha = 0, 1/2$, and 1. Note that when $\alpha = 0$, $X_{rc}(f)$ reduces to an ideal brick wall physically nonrealizable frequency response with bandwidth occupancy $1/2T$. The frequency $1/2T$ is called the *Nyquist frequency*. For $\alpha > 0$, the bandwidth occupied by the desired signal $X_{rc}(f)$ beyond the Nyquist frequency $1/2T$ is called the *excess bandwidth*, and is usually expressed as a percentage of the Nyquist frequency. For example, when $\alpha = 1/2$, the excess bandwidth is 50% and when $\alpha = 1$, the excess bandwidth is 100%. The signal pulse $x_{rc}(t)$ having the raised-cosine spectrum is

$$x_{rc}(t) = \frac{\sin \pi t/T}{\pi t/T} \frac{\cos(\pi \alpha t/T)}{1 - 4\alpha^2 t^2/T^2} \quad (5.13)$$

Figure 5.9(b) illustrates $x_{rc}(t)$ for $\alpha = 0, 1/2$, and 1. Note that $x_{rc}(t) = 1$ at $t = 0$ and $x_{rc}(t) = 0$ at $t = kT$, $k = \pm 1, \pm 2, \dots$. Consequently, at the sampling instants $t = kT$, $k \neq 0$, there is no ISI from adjacent symbols when there is no channel distortion. In the presence of channel distortion, however, the ISI given by Eq. (5.10) is no longer zero, and a **channel equalizer** is needed to minimize its effect on system performance.

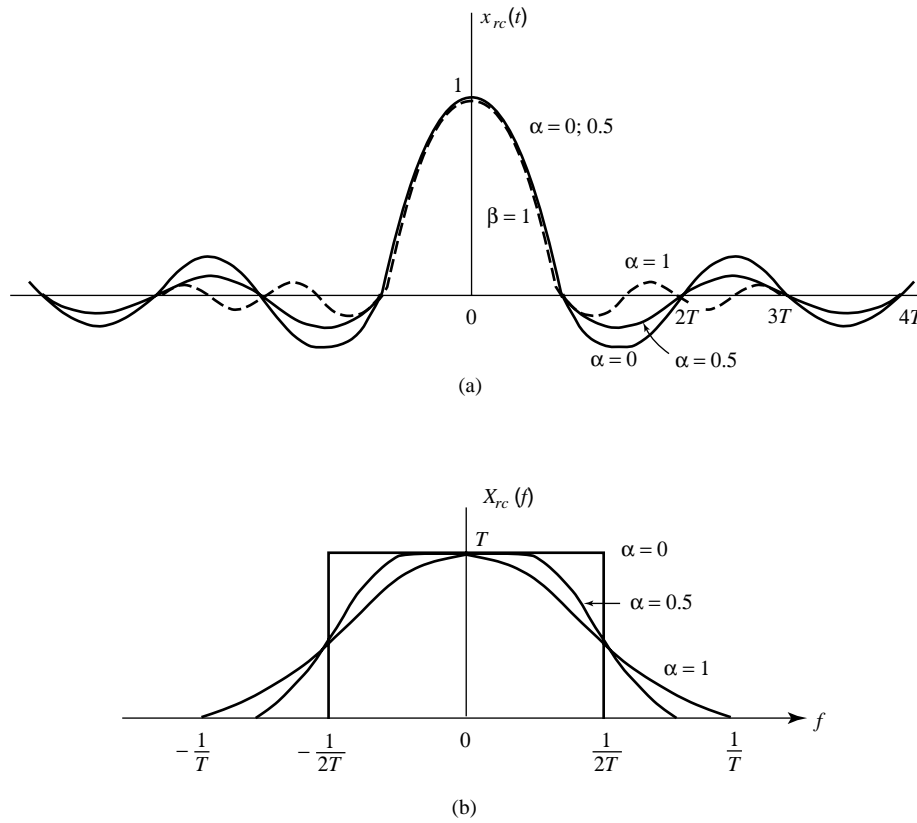


FIGURE 5.9: Pulses having a raised cosine spectrum.

5.3 Linear Equalizers

The most common type of channel equalizer used in practice to reduce SI is a linear transversal filter with adjustable coefficients $\{c_i\}$, as shown in Fig. 5.10.

On channels whose frequency response characteristics are unknown, but time invariant, we may measure the channel characteristics and adjust the parameters of the equalizer; once adjusted, the parameters remain fixed during the transmission of data. Such equalizers are called **preset equalizers**. On the other hand, **adaptive equalizers** update their parameters on a periodic basis during the transmission of data and, thus, they are capable of tracking a slowly time-varying channel response.

First, let us consider the design characteristics for a linear equalizer from a frequency domain viewpoint. Figure 5.11 shows a block diagram of a system that employs a linear filter as a **channel equalizer**.

The demodulator consists of a receiver filter with frequency response $G_R(f)$ in cascade with a channel equalizing filter that has a frequency response $G_E(f)$. As indicated in the preceding section, the receiver filter response $G_R(f)$ is matched to the transmitter response, i.e., $G_R(f) = G_T^*(f)$, and the product $G_R(f)G_T(f)$ is usually designed so that there is zero ISI at the sampling instants as, for example, when $G_R(f)G_T(f) = X_{rc}(f)$.

For the system shown in Fig. 5.11, in which the channel frequency response is not ideal, the desired

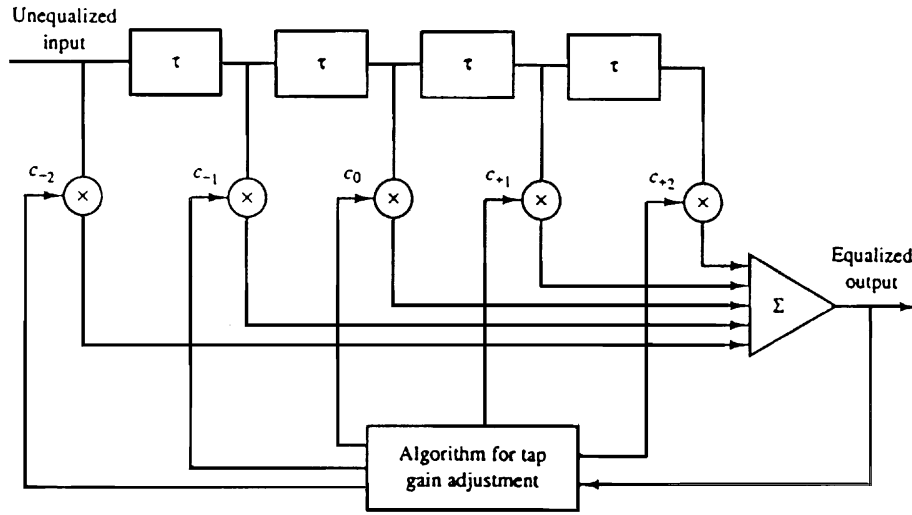


FIGURE 5.10: Linear transversal filter.

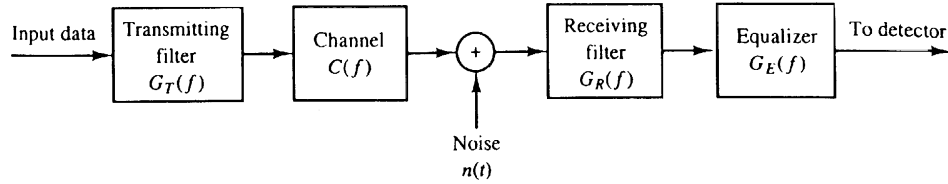


FIGURE 5.11: Block diagram of a system with an equalizer.

condition for zero ISI is

$$G_T(f)C(f)G_R(f)G_E(f) = X_{rc}(f) \quad (5.14)$$

where $X_{rc}(f)$ is the desired raised-cosine spectral characteristic. Since $G_T(f)G_R(f) = X_{rc}(f)$ by design, the frequency response of the equalizer that compensates for the channel distortion is

$$G_E(f) = \frac{1}{C(f)} = \frac{1}{|C(f)|} e^{-j\theta_c(f)} \quad (5.15)$$

Thus, the amplitude response of the equalizer is $|G_E(f)| = 1/|C(f)|$ and its phase response is $\theta_E(f) = -\theta_c(f)$. In this case, the equalizer is said to be the *inverse channel filter* to the channel response.

We note that the inverse channel filter completely eliminates ISI caused by the channel. Since it forces the ISI to be zero at the sampling instants $t = kT$, $k = 0, 1, \dots$, the equalizer is called a **zero-forcing equalizer**. Hence, the input to the detector is simply

$$z_k = I_k + \eta_k, \quad k = 0, 1, \dots \quad (5.16)$$

where η_k represents the additive noise and I_k is the desired symbol.

In practice, the ISI caused by channel distortion is usually limited to a finite number of symbols on either side of the desired symbol. Hence, the number of terms that constitute the ISI in the summation

given by Eq. (5.10) is finite. As a consequence, in practice the channel equalizer is implemented as a finite duration impulse response (FIR) filter, or transversal filter, with adjustable tap coefficients $\{c_n\}$, as illustrated in Fig. 5.10. The time delay τ between adjacent taps may be selected as large as T , the symbol interval, in which case the FIR equalizer is called a **symbol-spaced equalizer**. In this case, the input to the equalizer is the sampled sequence given by Eq. (5.7). We note that when the symbol rate $1/T < 2W$, however, frequencies in the received signal above the folding frequency $1/T$ are aliased into frequencies below $1/T$. In this case, the equalizer compensates for the aliased channel-distorted signal.

On the other hand, when the time delay τ between adjacent taps is selected such that $1/\tau \geq 2W > 1/T$, no aliasing occurs and, hence, the inverse channel equalizer compensates for the true channel distortion. Since $\tau < T$, the channel equalizer is said to have *fractionally spaced taps* and it is called a **fractionally spaced equalizer**. In practice, τ is often selected as $\tau = T/2$. Notice that, in this case, the sampling rate at the output of the filter $G_R(f)$ is $2/T$.

The impulse response of the FIR equalizer is

$$g_E(t) = \sum_{n=-N}^N c_n \delta(t - n\tau) \quad (5.17)$$

and the corresponding frequency response is

$$G_E(f) = \sum_{n=-N}^N c_n e^{-j2\pi f n\tau} \quad (5.18)$$

where $\{c_n\}$ are the $(2N + 1)$ equalizer coefficients and N is chosen sufficiently large so that the equalizer spans the length of the ISI, i.e., $2N + 1 \geq L$, where L is the number of signal samples spanned by the ISI. Since $X(f) = G_T(f)C(f)G_R(f)$ and $x(t)$ is the signal pulse corresponding to $X(f)$, then the equalized output signal pulse is

$$q(t) = \sum_{n=-N}^N c_n x(t - n\tau) \quad (5.19)$$

The zero-forcing condition can now be applied to the samples of $q(t)$ taken at times $t = mT$. These samples are

$$q(mT) = \sum_{n=-N}^N c_n x(mT - n\tau), \quad m = 0, \pm 1, \dots, \pm N \quad (5.20)$$

Since there are $2N + 1$ equalizer coefficients, we can control only $2N + 1$ sampled values of $q(t)$. Specifically, we may force the conditions

$$q(mT) = \sum_{n=-N}^N c_n x(mT - n\tau) = \begin{cases} 1, & m = 0 \\ 0, & m = \pm 1, \pm 2, \dots, \pm N \end{cases} \quad (5.21)$$

which may be expressed in matrix form as $\mathbf{X}\mathbf{c} = \mathbf{q}$, where \mathbf{X} is a $(2N + 1) \times (2N + 1)$ matrix with elements $\{x(mT - n\tau)\}$, \mathbf{c} is the $(2N + 1)$ coefficient vector and \mathbf{q} is the $(2N + 1)$ column vector with one nonzero element. Thus, we obtain a set of $2N + 1$ linear equations for the coefficients of the zero-forcing equalizer.

We should emphasize that the FIR zero-forcing equalizer does not completely eliminate ISI because it has a finite length. As N is increased, however, the residual ISI can be reduced, and in the limit as $N \rightarrow \infty$, the ISI is completely eliminated.

EXAMPLE 5.1:

Consider a channel distorted pulse $x(t)$, at the input to the equalizer, given by the expression

$$x(t) = \frac{1}{1 + \left(\frac{2t}{T}\right)^2}$$

where $1/T$ is the symbol rate. The pulse is sampled at the rate $2/T$ and equalized by a zero-forcing equalizer. Determine the coefficients of a five-tap zero-forcing equalizer.

Solution 5.1 According to Eq. (5.21), the zero-forcing equalizer must satisfy the equations

$$q(mT) = \sum_{n=-2}^2 c_n x(mT - nT/2) = \begin{cases} 1, & m = 0 \\ 0, & m = \pm 1, \pm 2 \end{cases}$$

The matrix X with elements $x(mT - nT/2)$ is given as

$$X = \begin{bmatrix} \frac{1}{5} & \frac{1}{10} & \frac{1}{17} & \frac{1}{26} & \frac{1}{37} \\ 1 & \frac{1}{2} & \frac{1}{5} & \frac{1}{10} & \frac{1}{17} \\ \frac{1}{5} & \frac{1}{2} & 1 & \frac{1}{2} & \frac{1}{5} \\ \frac{1}{17} & \frac{1}{10} & \frac{1}{5} & \frac{1}{2} & 1 \\ \frac{1}{37} & \frac{1}{26} & \frac{1}{17} & \frac{1}{10} & \frac{1}{5} \end{bmatrix} \quad (5.22)$$

The coefficient vector c and the vector q are given as

$$c = \begin{bmatrix} c_{-2} \\ c_{-1} \\ c_0 \\ c_1 \\ c_2 \end{bmatrix} \quad q = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad (5.23)$$

Then, the linear equations $Xc = q$ can be solved by inverting the matrix X . Thus, we obtain

$$c_{\text{opt}} = X^{-1}q = \begin{bmatrix} -2.2 \\ 4.9 \\ -3 \\ 4.9 \\ -2.2 \end{bmatrix} \quad (5.24)$$

One drawback to the zero-forcing equalizer is that it ignores the presence of additive noise. As a consequence, its use may result in significant noise enhancement. This is easily seen by noting that

in a frequency range where $C(f)$ is small, the channel equalizer $G_E(f) = 1/C(f)$ compensates by placing a large gain in that frequency range. Consequently, the noise in that frequency range is greatly enhanced. An alternative is to relax the zero ISI condition and select the channel equalizer characteristic such that the combined power in the residual ISI and the additive noise at the output of the equalizer is minimized. A channel equalizer that is optimized based on the minimum mean square error (MMSE) criterion accomplishes the desired goal.

To elaborate, let us consider the noise corrupted output of the FIR equalizer, which is

$$z(t) = \sum_{n=-N}^N c_n y(t - n\tau) \quad (5.25)$$

where $y(t)$ is the input to the equalizer, given by Eq. (5.6). The equalizer output is sampled at times $t = mT$. Thus, we obtain

$$z(mT) = \sum_{n=-N}^N c_n y(mT - n\tau) \quad (5.26)$$

The desired response at the output of the equalizer at $t = mT$ is the transmitted symbol I_m . The error is defined as the difference between I_m and $z(mT)$. Then, the mean square error (*MSE*) between the actual output sample $z(mT)$ and the desired values I_m is

$$\begin{aligned} MSE &= E |z(mT) - I_m|^2 \\ &= E \left[\left| \sum_{n=-N}^N c_n y(mT - n\tau) - I_m \right|^2 \right] \\ &= \sum_{n=-N}^N \sum_{k=-N}^N c_n c_k R_Y(n - k) \\ &\quad - 2 \sum_{k=-N}^N c_k R_{IY}(k) + E (|I_m|^2) \end{aligned} \quad (5.27)$$

where the correlations are defined as

$$\begin{aligned} R_Y(n - k) &= E [y^*(mT - n\tau) y(mT - k\tau)] \\ R_{IY}(k) &= E [y(mT - k\tau) I_m^*] \end{aligned} \quad (5.28)$$

and the expectation is taken with respect to the random information sequence $\{I_m\}$ and the additive noise.

The minimum *MSE* solution is obtained by differentiating Eq. (5.27) with respect to the equalizer coefficients $\{c_n\}$. Thus, we obtain the necessary conditions for the minimum *MSE* as

$$\sum_{n=-N}^N c_n R_Y(n - k) = R_{IY}(k), \quad k = 0, \pm 1, 2, \dots, \pm N \quad (5.29)$$

These are the $(2N + 1)$ linear equations for the equalizer coefficients. In contrast to the zero-forcing solution already described, these equations depend on the statistical properties (the autocorrelation) of the noise as well as the ISI through the autocorrelation $R_Y(n)$.

In practice, the autocorrelation matrix $R_Y(n)$ and the crosscorrelation vector $R_{IY}(n)$ are unknown a priori. These correlation sequences can be estimated, however, by transmitting a test signal over the channel and using the time-average estimates

$$\begin{aligned}\hat{R}_Y(n) &= \frac{1}{K} \sum_{k=1}^K y^*(kT - n\tau)y(kT) \\ \hat{R}_{IY}(n) &= \frac{1}{K} \sum_{k=1}^K y(kT - n\tau)I_k^*\end{aligned}\tag{5.30}$$

in place of the ensemble averages to solve for the equalizer coefficients given by Eq. (5.29).

5.3.1 Adaptive Linear Equalizers

We have shown that the tap coefficients of a linear equalizer can be determined by solving a set of linear equations. In the zero-forcing optimization criterion, the linear equations are given by Eq. (5.21). On the other hand, if the optimization criterion is based on minimizing the *MSE*, the optimum equalizer coefficients are determined by solving the set of linear equations given by Eq. (5.29).

In both cases, we may express the set of linear equations in the general matrix form

$$\mathbf{B}\mathbf{c} = \mathbf{d}\tag{5.31}$$

where \mathbf{B} is a $(2N + 1) \times (2N + 1)$ matrix, \mathbf{c} is a column vector representing the $2N + 1$ equalizer coefficients, and \mathbf{d} a $(2N + 1)$ -dimensional column vector. The solution of Eq. (5.31) yields

$$\mathbf{c}_{\text{opt}} = \mathbf{B}^{-1}\mathbf{d}\tag{5.32}$$

In practical implementations of equalizers, the solution of Eq. (5.31) for the optimum coefficient vector is usually obtained by an iterative procedure that avoids the explicit computation of the inverse of the matrix \mathbf{B} . The simplest iterative procedure is the method of steepest descent, in which one begins by choosing arbitrarily the coefficient vector \mathbf{c} , say \mathbf{c}_0 . This initial choice of coefficients corresponds to a point on the criterion function that is being optimized. For example, in the case of the *MSE* criterion, the initial guess \mathbf{c}_0 corresponds to a point on the quadratic *MSE* surface in the $(2N + 1)$ -dimensional space of coefficients. The gradient vector, defined as \mathbf{g}_0 , which is the derivative of the *MSE* with respect to the $2N + 1$ filter coefficients, is then computed at this point on the criterion surface, and each tap coefficient is changed in the direction opposite to its corresponding gradient component. The change in the j th tap coefficient is proportional to the size of the j th gradient component.

For example, the gradient vector denoted as \mathbf{g}_k , for the *MSE* criterion, found by taking the derivatives of the *MSE* with respect to each of the $2N + 1$ coefficients, is

$$\mathbf{g}_k = \mathbf{B}\mathbf{c}_k - \mathbf{d}, \quad k = 0, 1, 2, \dots\tag{5.33}$$

Then the coefficient vector \mathbf{c}_k is updated according to the relation

$$\mathbf{c}_{k+1} = \mathbf{c}_k - \Delta \mathbf{g}_k\tag{5.34}$$

where Δ is the *step-size parameter* for the iterative procedure. To ensure convergence of the iterative procedure, Δ is chosen to be a small positive number. In such a case, the gradient vector \mathbf{g}_k converges

toward zero, i.e., $\mathbf{g}_k \rightarrow \mathbf{0}$ as $k \rightarrow \infty$, and the coefficient vector $\mathbf{c}_k \rightarrow \mathbf{c}_{\text{opt}}$ as illustrated in Fig. 5.12 based on two-dimensional optimization. In general, convergence of the equalizer tap coefficients to \mathbf{c}_{opt} cannot be attained in a finite number of iterations with the steepest-descent method. The optimum solution \mathbf{c}_{opt} , however, can be approached as closely as desired in a few hundred iterations. In digital communication systems that employ channel equalizers, each iteration corresponds to a time interval for sending one symbol and, hence, a few hundred iterations to achieve convergence to \mathbf{c}_{opt} corresponds to a fraction of a second.

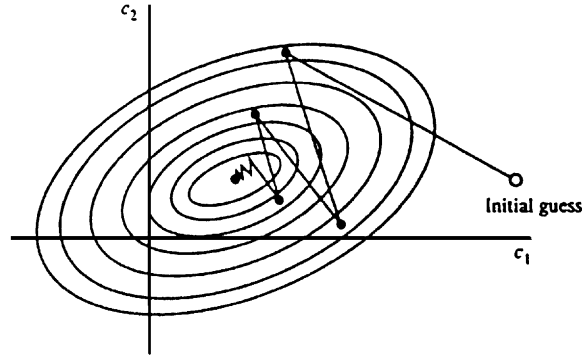


FIGURE 5.12: Examples of convergence characteristics of a gradient algorithm.

Adaptive channel equalization is required for channels whose characteristics change with time. In such a case, the ISI varies with time. The channel equalizer must track such time variations in the channel response and adapt its coefficients to reduce the ISI. In the context of the preceding discussion, the optimum coefficient vector \mathbf{c}_{opt} varies with time due to time variations in the matrix \mathbf{B} and, for the case of the *MSE* criterion, time variations in the vector \mathbf{d} . Under these conditions, the iterative method described can be modified to use estimates of the gradient components. Thus, the algorithm for adjusting the equalizer tap coefficients may be expressed as

$$\hat{\mathbf{c}}_{k+1} = \hat{\mathbf{c}}_k - \Delta \hat{\mathbf{g}}_k \quad (5.35)$$

where $\hat{\mathbf{g}}_k$ denotes an estimate of the gradient vector \mathbf{g}_k and $\hat{\mathbf{c}}_k$ denotes the estimate of the tap coefficient vector.

In the case of the *MSE* criterion, the gradient vector \mathbf{g}_k given by Eq. (5.33) may also be expressed as

$$\mathbf{g}_k = -E(e_k \mathbf{y}_k^*)$$

An estimate $\hat{\mathbf{g}}_k$ of the gradient vector at the k th iteration is computed as

$$\hat{\mathbf{g}}_k = -e_k \mathbf{y}_k^* \quad (5.36)$$

where e_k denotes the difference between the desired output from the equalizer at the k th time instant and the actual output $z(kT)$, and \mathbf{y}_k denotes the column vector of $2N + 1$ received signal values contained in the equalizer at time instant k . The *error signal* e_k is expressed as

$$e_k = I_k - z_k \quad (5.37)$$

where $z_k = z(kT)$ is the equalizer output given by Eq. (5.26) and I_k is the desired symbol. Hence, by substituting Eq. (5.36) into Eq. (5.35), we obtain the adaptive algorithm for optimizing the tap coefficients (based on the *MSE* criterion) as

$$\hat{c}_{k+1} = \hat{c}_k + \Delta e_k y_k^* \quad (5.38)$$

Since an estimate of the gradient vector is used in Eq. (5.38) the algorithm is called a **stochastic gradient algorithm**; it is also known as the **LMS algorithm**.

A block diagram of an adaptive equalizer that adapts its tap coefficients according to Eq. (5.38) is illustrated in Fig. 5.13. Note that the difference between the desired output I_k and the actual output z_k from the equalizer is used to form the error signal e_k . This error is scaled by the step-size parameter Δ , and the scaled error signal Δe_k multiplies the received signal values $\{y(kT - n\tau)\}$ at the $2N + 1$ taps. The products $\Delta e_k y_k^*(kT - n\tau)$ at the $(2N + 1)$ taps are then added to the previous values of the tap coefficients to obtain the updated tap coefficients, according to Eq. (5.38). This computation is repeated as each new symbol is received. Thus, the equalizer coefficients are updated at the symbol rate.

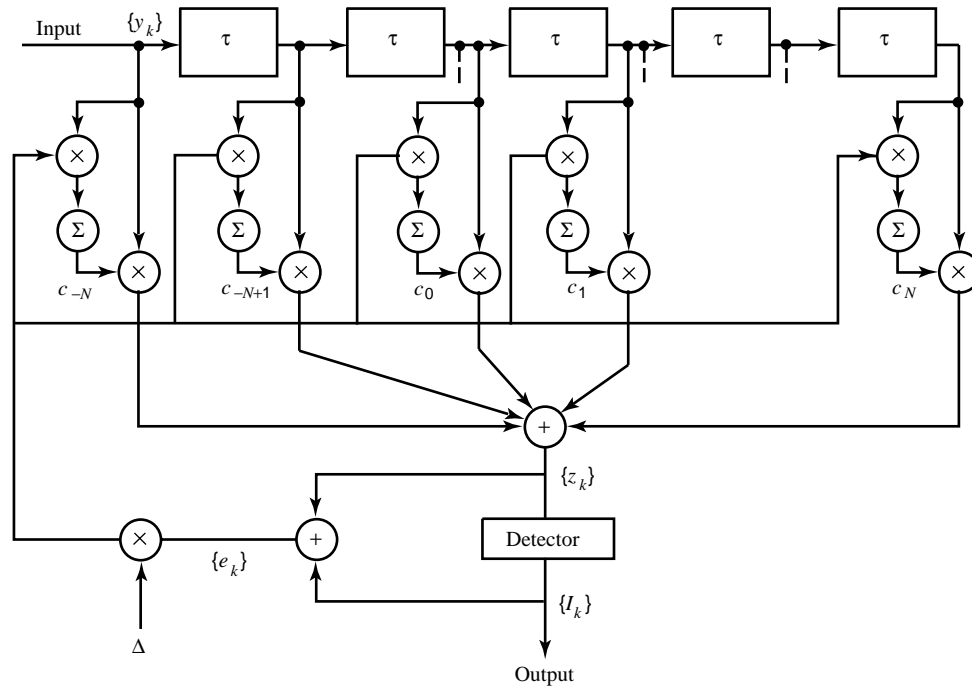


FIGURE 5.13: Linear adaptive equalizer based on the *MSE* criterion.

Initially, the adaptive equalizer is trained by the transmission of a known pseudo-random sequence $\{I_m\}$ over the channel. At the demodulator, the equalizer employs the known sequence to adjust its coefficients. Upon initial adjustment, the adaptive equalizer switches from a **training mode** to a **decision-directed mode**, in which case the decisions at the output of the detector are sufficiently reliable so that the error signal is formed by computing the difference between the detector output

and the equalizer output, i.e.,

$$e_k = \tilde{I}_k - z_k \quad (5.39)$$

where \tilde{I}_k is the output of the detector. In general, decision errors at the output of the detector occur infrequently and, consequently, such errors have little effect on the performance of the tracking algorithm given by Eq. (5.38).

A rule of thumb for selecting the step-size parameter so as to ensure convergence and good tracking capabilities in slowly varying channels is

$$\Delta = \frac{1}{5(2N + 1)P_R} \quad (5.40)$$

where P_R denotes the received signal-plus-noise power, which can be estimated from the received signal (see Proakis [4]).

The convergence characteristic of the stochastic gradient algorithm in Eq. (5.38) is illustrated in Fig. 5.14. These graphs were obtained from a computer simulation of an 11-tap adaptive equalizer operating on a channel with a rather modest amount of ISI. The input signal-plus-noise power P_R was normalized to unity. The rule of thumb given in Eq. (5.40) for selecting the step size gives $\Delta = 0.018$. The effect of making Δ too large is illustrated by the large jumps in *MSE* as shown for $\Delta = 0.115$. As Δ is decreased, the convergence is slowed somewhat, but a lower *MSE* is achieved, indicating that the estimated coefficients are closer to c_{opt} .

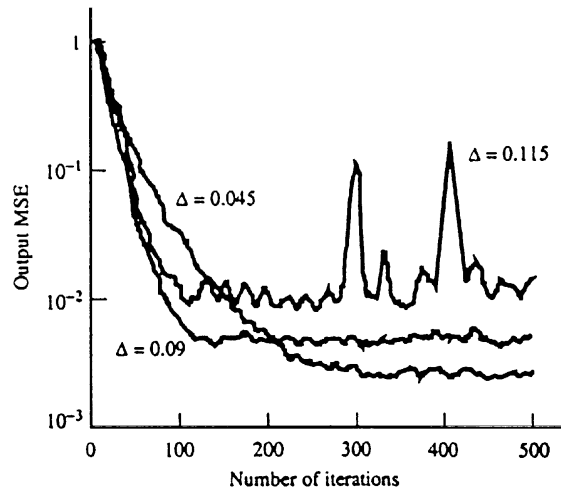


FIGURE 5.14: Initial convergence characteristics of the *LMS* algorithm with different step sizes.

Although we have described in some detail the operation of an adaptive equalizer that is optimized on the basis of the *MSE* criterion, the operation of an adaptive equalizer based on the zero-forcing method is very similar. The major difference lies in the method for estimating the gradient vectors \mathbf{g}_k at each iteration. A block diagram of an adaptive zero-forcing equalizer is shown in Fig. 5.15.

For more details on the tap coefficient update method for a zero-forcing equalizer, the reader is referred to the papers by Lucky [2, 3], and the text by Proakis [4].

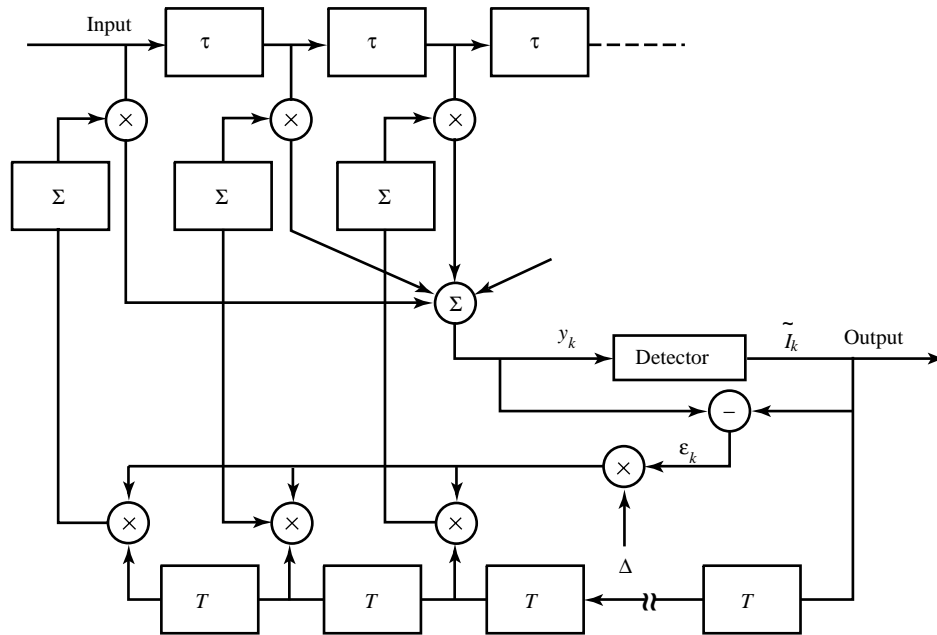


FIGURE 5.15: An adaptive zero-forcing equalizer.

5.4 Decision-Feedback Equalizer

The linear filter equalizers described in the preceding section are very effective on channels, such as wire line telephone channels, where the ISI is not severe. The severity of the ISI is directly related to the spectral characteristics and not necessarily to the time span of the ISI. For example, consider the ISI resulting from the two channels that are illustrated in Fig. 5.16. The time span for the ISI in channel A is 5 symbol intervals on each side of the desired signal component, which has a value of 0.72. On the other hand, the time span for the ISI in channel B is one symbol interval on each side of the desired signal component, which has a value of 0.815. The energy of the total response is normalized to unity for both channels.

In spite of the shorter ISI span, channel B results in more severe ISI. This is evidenced in the frequency response characteristics of these channels, which are shown in Fig. 5.17. We observe that channel B has a spectral null [the frequency response $C(f) = 0$ for some frequencies in the band $|f| \leq W$] at $f = 1/2T$, whereas this does not occur in the case of channel A. Consequently, a linear equalizer will introduce a large gain in its frequency response to compensate for the channel null. Thus, the noise in channel B will be enhanced much more than in channel A. This implies that the performance of the linear equalizer for channel B will be sufficiently poorer than that for channel A. This fact is borne out by the computer simulation results for the performance of the two linear equalizers shown in Fig. 5.18. Hence, the basic limitation of a linear equalizer is that it performs poorly on channels having spectral nulls. Such channels are often encountered in radio communications, such as ionospheric transmission at frequencies below 30 MHz and mobile radio channels, such as those used for cellular radio communications.

A **decision-feedback equalizer (DFE)** is a nonlinear equalizer that employs previous decisions to eliminate the ISI caused by previously detected symbols on the current symbol to be detected. A

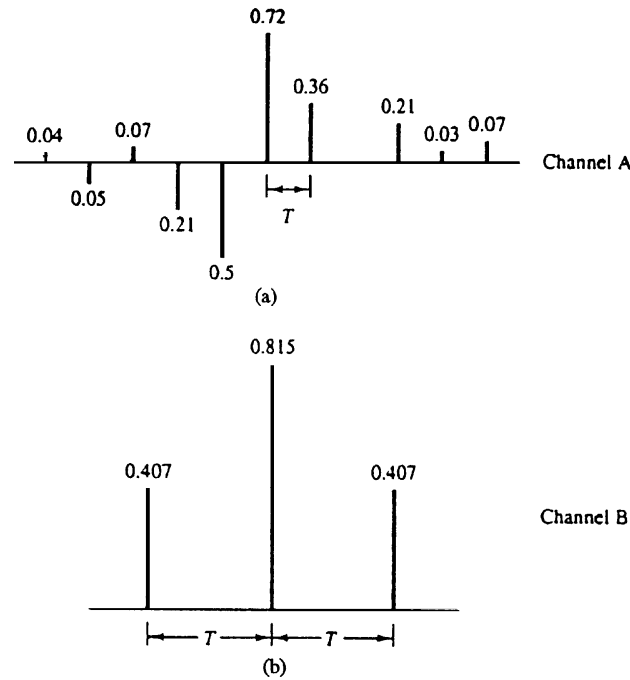


FIGURE 5.16: Two channels with ISI.

simple block diagram for a DFE is shown in Fig. 5.19. The DFE consists of two filters. The first filter is called a *feedforward filter* and it is generally a fractionally spaced FIR filter with adjustable tap coefficients. This filter is identical in form to the linear equalizer already described. Its input is the received filtered signal $y(t)$ sampled at some rate that is a multiple of the symbol rate, e.g., at rate $2/T$. The second filter is a *feedback filter*. It is implemented as an FIR filter with symbol-spaced taps having adjustable coefficients. Its input is the set of previously detected symbols. The output of the feedback filter is subtracted from the output of the feedforward filter to form the input to the detector. Thus, we have

$$z_m = \sum_{n=-N_1}^0 c_n y(mT - n\tau) - \sum_{n=1}^{N_2} b_n \tilde{I}_{m-n} \quad (5.41)$$

where $\{c_n\}$ and $\{b_n\}$ are the adjustable coefficients of the feedforward and feedback filters, respectively, \tilde{I}_{m-n} , $n = 1, 2, \dots, N_2$ are the previously detected symbols, $N_1 + 1$ is the length of the feedforward filter, and N_2 is the length of the feedback filter. Based on the input z_m , the detector determines which of the possible transmitted symbols is closest in distance to the input signal I_m . Thus, it makes its decision and outputs \tilde{I}_m . What makes the DFE nonlinear is the nonlinear characteristic of the detector that provides the input to the feedback filter.

The tap coefficients of the feedforward and feedback filters are selected to optimize some desired performance measure. For mathematical simplicity, the *MSE* criterion is usually applied, and a stochastic gradient algorithm is commonly used to implement an adaptive DFE. Figure 5.20 illustrates the block diagram of an adaptive DFE whose tap coefficients are adjusted by means of the LMS stochastic gradient algorithm. Figure 5.21 illustrates the probability of error performance of the

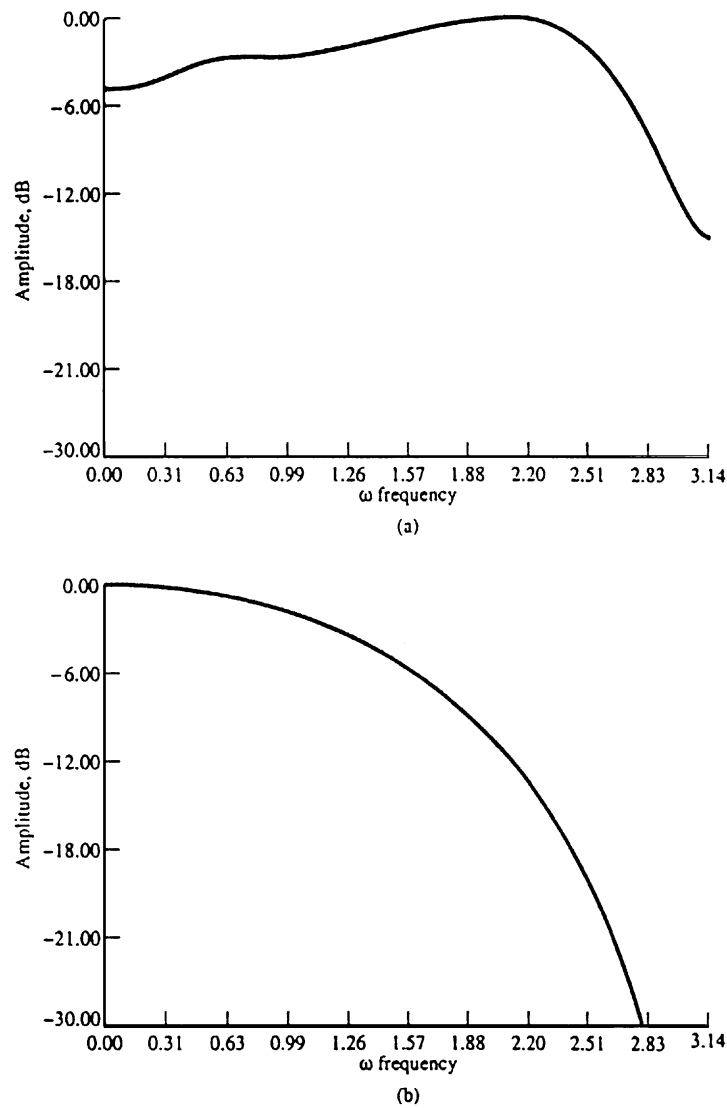


FIGURE 5.17: Amplitude spectra for (a) channel A shown in Fig. 5.16(a) and (b) channel B shown in Fig. 5.16(b).

DFE, obtained by computer simulation, for binary PAM transmission over channel B. The gain in performance relative to that of a linear equalizer is clearly evident.

We should mention that decision errors from the detector that are fed to the feedback filter have a small effect on the performance of the DFE. In general, a small loss in performance of one to two decibels is possible at error rates below 10^{-2} , as illustrated in Fig. 5.21, but the decision errors in the feedback filters are not catastrophic.

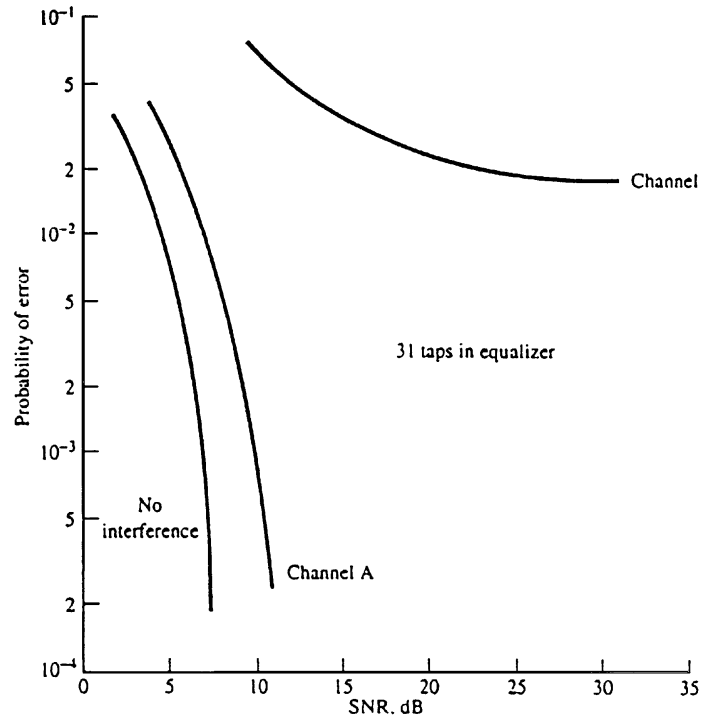


FIGURE 5.18: Error-rate performance of linear *MSE* equalizer.

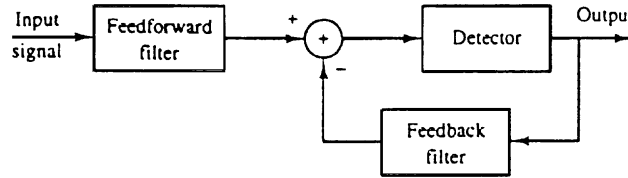


FIGURE 5.19: Block diagram of DFE.

5.5 Maximum-Likelihood Sequence Detection

Although the DFE outperforms a linear equalizer, it is not the optimum equalizer from the viewpoint of minimizing the probability of error in the detection of the information sequence $\{I_k\}$ from the received signal samples $\{y_k\}$ given in Eq. (5.5). In a digital communication system that transmits information over a channel that causes ISI, the optimum detector is a maximum-likelihood symbol sequence detector which produces at its output the most probable symbol sequence $\{\tilde{I}_k\}$ for the given received sampled sequence $\{y_k\}$. That is, the detector finds the sequence $\{\tilde{I}_k\}$ that maximizes the *likelihood function*

$$\Lambda(\{I_k\}) = \ln p(\{y_k\} | \{I_k\}) \quad (5.42)$$

where $p(\{y_k\} | \{I_k\})$ is the joint probability of the received sequence $\{y_k\}$ conditioned on $\{I_k\}$. The sequence of symbols $\{\tilde{I}_k\}$ that maximizes this joint conditional probability is called the **maximum-**

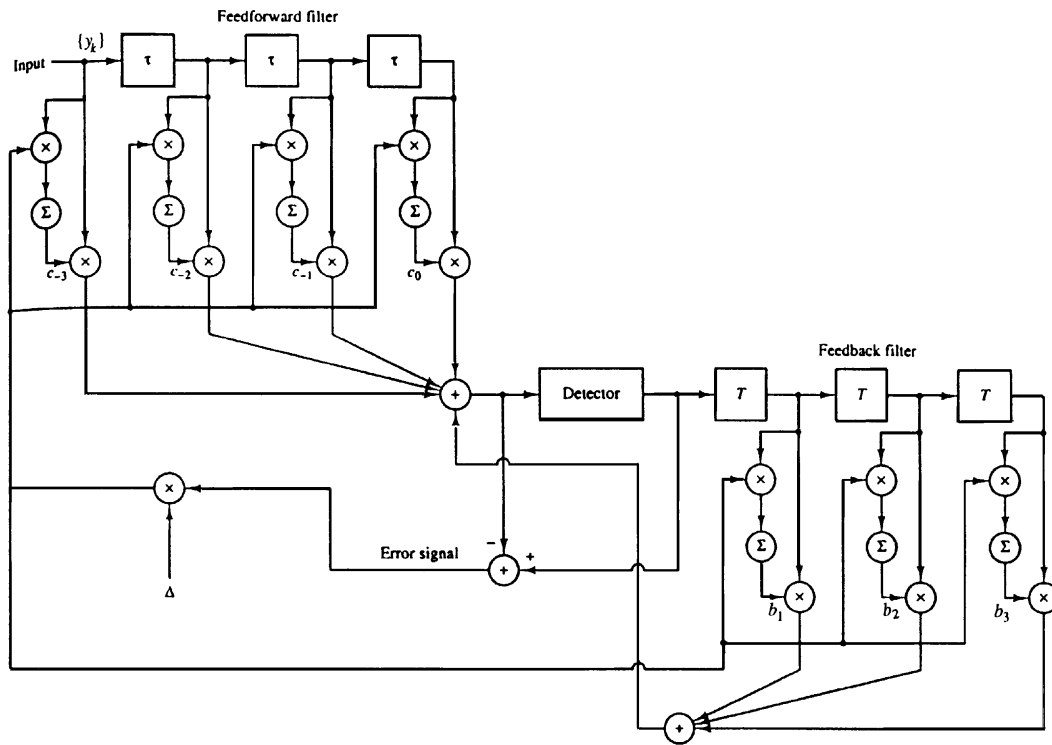


FIGURE 5.20: Adaptive DFE.

likelihood sequence detector.

An algorithm that implements maximum-likelihood sequence detection (MLSD) is the Viterbi algorithm, which was originally devised for decoding convolutional codes. For a description of this algorithm in the context of sequence detection in the presence of ISI, the reader is referred to the paper by Forney [1] and the text by Proakis [4].

The major drawback of MLSD for channels with ISI is the exponential behavior in computational complexity as a function of the span of the ISI. Consequently, MLSD is practical only for channels where the ISI spans only a few symbols and the ISI is severe, in the sense that it causes a severe degradation in the performance of a linear equalizer or a decision-feedback equalizer. For example, Fig. 5.22 illustrates the error probability performance of the Viterbi algorithm for a binary PAM signal transmitted through channel B (see Fig. 5.16). For purposes of comparison, we also illustrate the probability of error for a DFE. Both results were obtained by computer simulation. We observe that the performance of the maximum likelihood sequence detector is about 4.5 dB better than that of the DFE at an error probability of 10^{-4} . Hence, this is one example where the ML sequence detector provides a significant performance gain on a channel with a relatively short ISI span.

5.6 Conclusions

Channel equalizers are widely used in digital communication systems to mitigate the effects of ISI caused by channel distortion. Linear equalizers are widely used for high-speed modems that transmit data over telephone channels. For wireless (radio) transmission, such as in mobile cellular communi-

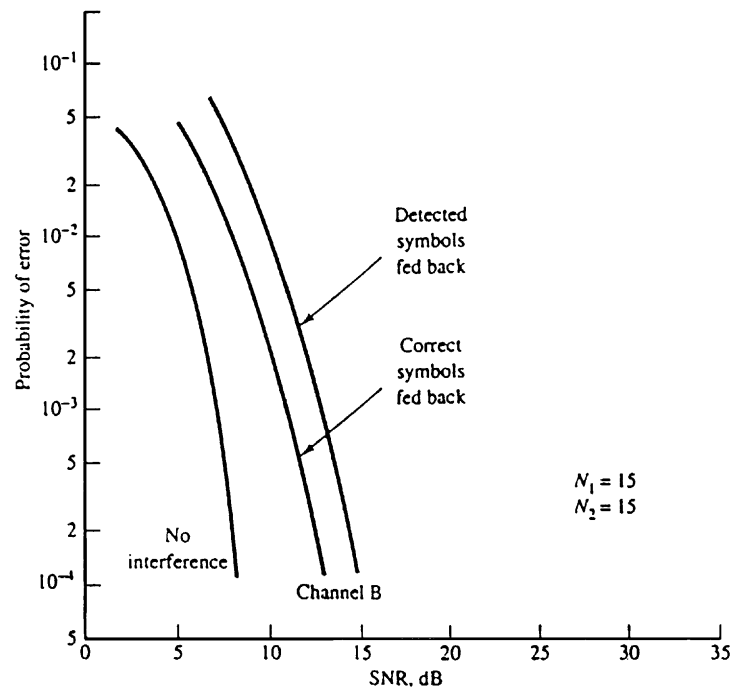


FIGURE 5.21: Performance of DFE with and without error propagation.

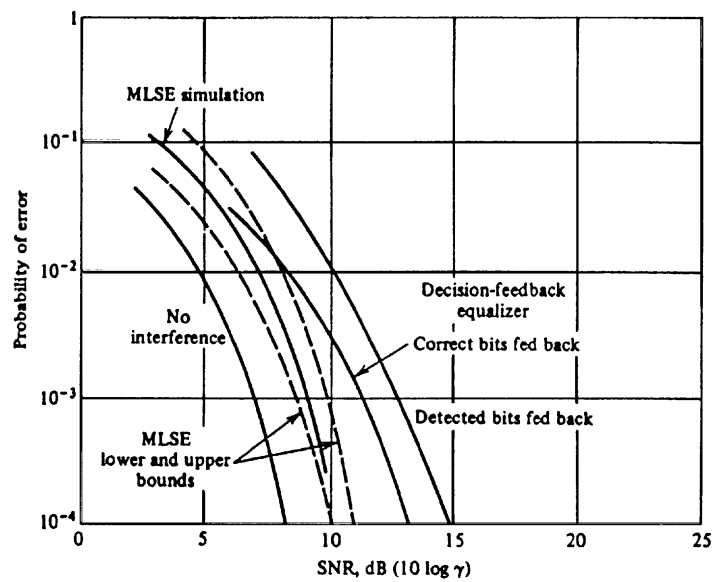


FIGURE 5.22: Comparison of performance between MLSE and decision-feedback equalization for channel B of Fig. 5.16.

cations and interoffice communications, the multipath propagation of the transmitted signal results in severe ISI. Such channels require more powerful equalizers to combat the severe ISI. The decision-feedback equalizer and the MLSD are two nonlinear channel equalizers that are suitable for radio channels with severe ISI.

Defining Terms

Adaptive equalizer: A channel equalizer whose parameters are updated automatically and adaptively during transmission of data.

Channel equalizer: A device that is used to reduce the effects of channel distortion in a received signal.

Decision-directed mode: Mode for adjustment of the equalizer coefficient adaptively based on the use of the detected symbols at the output of the detector.

Decision-feedback equalizer (DFE): An adaptive equalizer that consists of a feedforward filter and a feedback filter, where the latter is fed with previously detected symbols that are used to eliminate the intersymbol interference due to the tail in the channel impulse response.

Fractionally spaced equalizer: A tapped-delay line channel equalizer in which the delay between adjacent taps is less than the duration of a transmitted symbol.

Intersymbol interference: Interference in a received symbol from adjacent (nearby) transmitted symbols caused by channel distortion in data transmission.

LMS algorithm: See stochastic gradient algorithm.

Maximum-likelihood sequence detector: A detector for estimating the most probable sequence of data symbols by maximizing the likelihood function of the received signal.

Preset equalizer: A channel equalizer whose parameters are fixed (time-invariant) during transmission of data.

Stochastic gradient algorithm: An algorithm for adaptively adjusting the coefficients of an equalizer based on the use of (noise-corrupted) estimates of the gradients.

Symbol-spaced equalizer: A tapped-delay line channel equalizer in which the delay between adjacent taps is equal to the duration of a transmitted symbol.

Training mode: Mode for adjustment of the equalizer coefficients based on the transmission of a known sequence of transmitted symbols.

Zero-forcing equalizer: A channel equalizer whose parameters are adjusted to completely eliminate intersymbol interference in a sequence of transmitted data symbols.

References

- [1] Forney, G.D., Jr., Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference. *IEEE Trans. Inform. Theory*, IT-18, 363–378, May 1972.
- [2] Lucky, R.W., Automatic equalization for digital communications. *Bell Syst. Tech. J.*, 44, 547–588, Apr. 1965.
- [3] Lucky, R.W., Techniques for adaptive equalization of digital communication. *Bell Syst. Tech. J.*, 45, 255–286, Feb. 1966.
- [4] Proakis, J.G., *Digital Communications*, 3rd ed., McGraw-Hill, New York, 1995.

Further Information

For a comprehensive treatment of adaptive equalization techniques and their performance characteristics, the reader may refer to the book by Proakis [4]. The two papers by Lucky [2, 3], provide a treatment on linear equalizers based on the zero-forcing criterion. Additional information on decision-feedback equalizers may be found in the journal papers “An Adaptive Decision-Feedback Equalizer” by D.A. George, R.R. Bowen, and J.R. Storey, *IEEE Transactions on Communications Technology*, Vol. COM-19, pp. 281–293, June 1971, and “Feedback Equalization for Fading Dispersive Channels” by P. Monsen, *IEEE Transactions on Information Theory*, Vol. IT-17, pp. 56–64, January 1971. A thorough treatment of channel equalization based on maximum-likelihood sequence detection is given in the paper by Forney [1].