Chapter 7
Equalization and Interference Cancelation

Chapters 5 and 6 have considered digital signaling on frequency nonselective or flat fading channels. Such channels are typical for low data rate systems that occupy a bandwidth that is smaller than the channel coherence bandwidth. However, as the data rate increases, the bandwidth of the transmitted waveform will typically be larger than the channel coherence bandwidth. Under this condition, the channel is nonideal and will exhibit frequency selectivity or time delay spread. Such time delay spread causes interference between modulated symbols, a phenomenon known as inter-symbol interference (ISI). This chapter concentrates on the modeling of ISI channels and the various signal processing methods for recovering digital information transmitted over such channels.

This chapter begins with a treatment of ISI channel modeling in Sect. 7.1 that includes a vector representation of digital signaling on ISI channels. Section 7.2 then develops the maximum likelihood receiver for ISI channels, leading to an equivalent model of the ISI channel known as the discrete-time white noise channel model. We also consider the effects of using fractional sampling or over-sampling at the receiver, where the sampling rate is an integer multiple of the modulated symbol rate. Section 7.3 provides a treatment of symbol-by-symbol equalizers, including the linear zero-forcing and minimum mean-square-error (MMSE) equalizers, and the nonlinear decision feedback equalizer. Section 7.4 provides a treatment of sequence estimators beginning with maximum likelihood sequence estimation (MLSE) and the Viterbi algorithm. Since the MLSE receiver can have high complexity for channels that have a long impulse response, we consider some reduced complexity sequence estimation techniques such as reduced state sequence estimation (RSSE) and delayed decision feedback sequence estimation (DDFSE). Section 7.5 provides an analysis of the bit error rate performance of MLSE on static ISI channels and multipath fading ISI channels. Section 7.6 considers fractionally spaced MLSE receivers for ISI channels.

Finally, Sect. 7.7 concludes the chapter with a discussion of co-channel demodulation for digital signals on ISI channels. The basic idea is to simultaneously recover the data from multiple users that transmit in the same bandwidth. The problem is formulated as a multiple-input multiple-output (MIMO) channel, where the inputs

G.L. Stüber, *Principles of Mobile Communication*, DOI 10.1007/978-1-4614-0364-7_7,

are the waveforms transmitted by multiple users in the same bandwidth, and the outputs are the signals received at multiple antenna elements. We first develop a vector representation of the received signals, along with the optimum receiver that uses joint maximum likelihood sequence estimation (J-MLSE). Similar to the single-user case, we consider the effects of using fractional sampling as well. Finally, we wrap up with a receiver structure that incorporates a combination of optimal combining as discussed in Chap. 6 and sequence estimation as implemented with the Viterbi algorithm.

7.1 Modeling of ISI Channels

Chapter 4 showed that the complex envelope of any modulated signal can be expressed in the general form

$$\tilde{s}(t) = A\sum_{n} b(t-nT, \mathbf{x}_n). \tag{7.1}$$

This chapter restricts attention to linear full-response modulation schemes where

$$b(t, \mathbf{x}_n) = x_{\mathrm{n}} h_{\mathrm{a}}(t), \tag{7.2}$$

$h_{\mathrm{a}}(t)$ is the amplitude shaping pulse, and $\{x_{\mathrm{n}}\}$ is the complex data symbol sequence.

Suppose that the signal in (7.2) is transmitted over a channel having a *time-invariant* complex low-pass impulse response $g(t)$. The received complex envelope is

$$\tilde{r}(t) = \sum_{n} x_{\mathrm{n}} h(t-nT) + \tilde{n}(t), \tag{7.3}$$

where

$$h(t) = \int_{-\infty}^{\infty} h_{\mathrm{a}}(\tau) g(t-\tau) \mathrm{d}\tau, \tag{7.4}$$

is the received pulse, given by the convolution of the transmitted pulse $h_{\mathrm{a}}(t)$ and the channel impulse response $g(t)$, and $\tilde{n}(t)$ is complex-valued additive white Gaussian noise (AWGN) with a one-sided power spectral density (PSD) of N_{o} W/Hz. Since the transmitted pulse $h_{\mathrm{a}}(t)$ is causal ($h_{\mathrm{a}}(t) = 0,\ t < 0$), the lower limit of integration can be replaced by zero, and since the physical channel is causal ($g(t) = 0,\ t < 0$), the upper limit of integration in (7.4) can be replaced by t, so that

$$h(t) = \int_{0}^{t} h_{\mathrm{a}}(\tau) g(t-\tau) \mathrm{d}\tau\ ,\ \ t \geq 0. \tag{7.5}$$

Finally, the received pulse $h(t)$ is assumed to have a finite duration so that $h(t) = 0$ for $t < 0$ and $t \geq (L+1)T$, where L is some positive integer.

7.1.1 Vector Representation of Received Signals

Using a Gram–Schmidt orthonormalization procedure, the received signal $\tilde{r}(t)$ in (7.3) can be expressed in the form

$$\tilde{r}(t) = \lim_{N\to\infty} \sum_{k=1}^{N} \tilde{r}_k \varphi_k(t), \tag{7.6}$$

where the $\{\varphi_n(t)\}$ are a complete set of complex orthonormal basis functions. Note that the basis functions span over the entire length of the waveform $\tilde{r}(t)$, and for the present purpose it is not necessary to actually generate the basis functions. Also, this set of basis functions should not be confused with the set of basis functions that is used to represent the signal set as in Sect. 5.1. It can be readily shown that

$$\tilde{r}_k = \sum_n x_{\mathrm{n}} h_{n_k} + \tilde{n}_k, \tag{7.7}$$

where

$$\begin{aligned} h_{n_k} &= \int_{-\infty}^{\infty} h(t-nT)\varphi_k^*(t)\mathrm{d}t, \\ \tilde{n}_k &= \int_{-\infty}^{\infty} \tilde{n}(t)\varphi_k^*(t)\mathrm{d}t. \end{aligned} \tag{7.8}$$

The noise samples $\tilde{n}_k$ are zero-mean complex Gaussian random variables with covariance $\phi_{\tilde{n}_k\tilde{n}_m} = \frac{1}{2}\mathrm{E}[\tilde{n}_k^*\tilde{n}_m] = N_{\mathrm{o}}\delta_{km}$. Hence, it follow that the observation vector $\tilde{\mathbf{r}} = (\tilde{r}_1, \tilde{r}_2, \cdots, \tilde{r}_N)$ has the conditional multivariate complex Gaussian distribution

$$p(\tilde{\mathbf{r}}|\mathbf{x},\mathbf{h}) = \prod_{k=1}^{N} \frac{1}{2\pi N_{\mathrm{o}}} \exp\left\{ -\frac{1}{2N_{\mathrm{o}}} \left| \tilde{r}_k - \sum_n x_{\mathrm{n}} h_{n_k} \right|^2 \right\}, \tag{7.9}$$

where $\mathbf{h} = \{h_{n_k}\}$.

7.2 Maximum Likelihood Receiver for ISI Channels with AWGN

The maximum likelihood receiver decides in favor of the symbol sequence $\mathbf{x}$ that maximizes the likelihood function $p(\tilde{\mathbf{r}}|\mathbf{x},\mathbf{h})$ or the log-likelihood function $\log\{p(\tilde{\mathbf{r}}|\mathbf{x},\mathbf{h})\}$, that is,

$$\text{choose } \mathbf{x} \text{ if } \quad \log\{p(\tilde{\mathbf{r}}|\mathbf{x},\mathbf{h})\} > \log\{p(\tilde{\mathbf{r}}|\hat{\mathbf{x}},\mathbf{h})\} \quad \forall\, \hat{\mathbf{x}} \neq \mathbf{x}. \tag{7.10}$$

For an AWGN channel, $p(\tilde{\mathbf{r}}|\mathbf{x},\mathbf{h})$ has the form in (7.9) and the decision rule in (7.10) is equivalent choosing $\mathbf{x}$ to maximize the quantity

$$\begin{aligned}\mu(\mathbf{x}) &= -\sum_{k=1}^{N}\left|\tilde{r}_k - \sum_n x_{\mathrm{n}} h_{n_k}\right|^2 \\ &= -\sum_{k=1}^{N}|\tilde{r}_k|^2 + \sum_{k=1}^{N}\left(\tilde{r}_k^*\sum_n x_{\mathrm{n}} h_{n_k} + \tilde{r}_k\sum_n x_{\mathrm{n}}^* h_{n_k}^*\right) \\ &\quad -\sum_{k=1}^{N}\left(\sum_n x_{\mathrm{n}} h_{n_k}\right)\left(\sum_m x_m^* h_{m_k}^*\right). \end{aligned} \tag{7.11}$$

Since the term $\sum_{k=1}^{N}|\tilde{r}_k|^2$ is independent of $\mathbf{x}$, it may be omitted so that the maximum likelihood receiver chooses $\mathbf{x}$ to maximize

$$\mu(\mathbf{x}) = 2\mathrm{Re}\left\{\sum_n x_{\mathrm{n}}^*\sum_{k=1}^{N}\tilde{r}_k h_{n_k}^*\right\} - \sum_n\sum_m x_{\mathrm{n}} x_m^*\sum_{k=1}^{N} h_{n_k} h_{m_k}^*, \tag{7.12}$$

where $\mathrm{Re}\{z\}$ denotes the real part of z. In the limit as the number of observable random variables N approaches infinity, we define the following:

$$y_n \triangleq \lim_{N\to\infty}\sum_{k=1}^{N}\tilde{r}_k h_{n_k}^* = \int_{-\infty}^{\infty}\tilde{r}(t)h^*(t-nT)\mathrm{d}t, \tag{7.13}$$

$$f_{m-n} \triangleq \lim_{N\to\infty}\sum_{k=1}^{N} h_{n_k} h_{m_k}^* = \int_{-\infty}^{\infty} h(t-nT)h^*(t-mT)\mathrm{d}t. \tag{7.14}$$

Using (7.13) and (7.14) in (7.12), we have the final form

$$\mu(\mathbf{x}) = 2\mathrm{Re}\left\{\sum_n x_{\mathrm{n}}^* y_n\right\} - \sum_n\sum_m x_{\mathrm{n}} x_m^* f_{m-n}. \tag{7.15}$$

The variables $\{y_n\}$ are obtained by passing the received complex low-pass waveform $\tilde{r}(t)$ through the matched filter $h^*(-t)$ and sampling the output. Note that the T-spaced samples at the output of the matched filter must be obtained with the correct timing phase, and in the above development perfect symbol synchronization is implied. Hence, the optimum front-end processing is as shown in Fig. 7.1. Finally by changing the variable of integration, the $\{f_{m-n}\}$ in (7.14) can be rewritten in the form

$$f_\ell = \int_{-\infty}^{\infty} h(t+\ell T)h^*(t)\mathrm{d}t, \tag{7.16}$$

where $\ell = m-n$. From (7.16), it is seen that the $\{f_\ell\}$ represent the sampled autocorrelation function of the received pulse $h(t)$ with sample spacing T, and have the property that $f_n^* = f_{-n}$. Sometimes the $\{f_\ell\}$ are called the ISI coefficients.

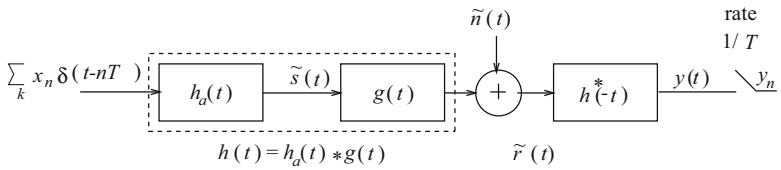

Fig. 7.1 Digital signaling on an ISI channel. The optimum front-end processor implements a filter that is matched to the received pulse $h(t)$ followed by a symbol rate sampler

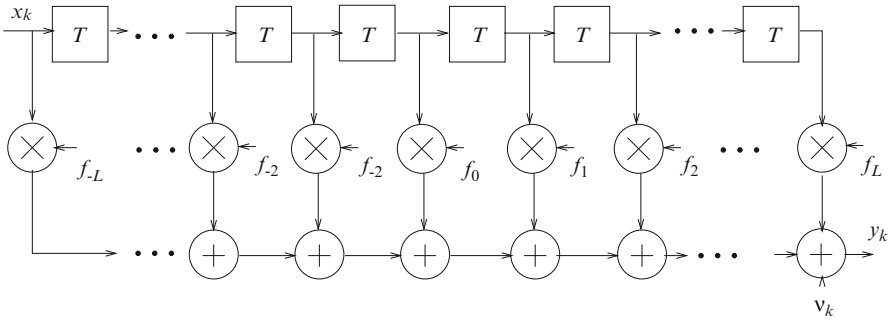

Fig. 7.2 Discrete-time model for digital signaling on an ISI channel

7.2.1 Discrete-Time White Noise Channel Model

Sampling the output of the matched filter $h^*(-t)$ in Fig. 7.1 every T seconds yields the sample sequence $\{y_k\}$, where

$$\begin{aligned} y_k &= \sum_n x_n f_{k-n} + v_k \\ &= \sum_{\ell=-L}^{L} x_{k-\ell} f_\ell + v_k \end{aligned} \tag{7.17}$$

and

$$v_k = \int_{-\infty}^{\infty} \tilde{n}(\tau) h^*(\tau - kT) \mathrm{d}\tau, \tag{7.18}$$

is the noise sample at the output of the matched filter. It follows that the overall discrete-time system in Fig. 7.1 can be represented by a discrete-time transversal filter with coefficients

$$\mathbf{f} = (f_{-L},\ f_{-L+1},\ \ldots,\ f_{-1},\ f_0,\ f_1,\ \ldots,\ f_{L-1},\ f_L). \tag{7.19}$$

This representation is depicted in Fig. 7.2.

As shown in (7.15), the maximum likelihood receiver uses the observation vector $\mathbf{y} = \{y_k\}$ and knowledge of the ISI coefficients $\{f_\ell\}$ to determine the most likely transmitted sequence $\mathbf{x}$. Equivalently, the maximum likelihood receiver decides in favor of the symbol sequence $\mathbf{x}$ that maximizes the likelihood function $p(\mathbf{y}|\mathbf{x},\mathbf{f})$ or the log-likelihood function $\log\{p(\mathbf{y}|\mathbf{x},\mathbf{f})\}$, that is,

$$\text{choose } \mathbf{x} \text{ if } \quad \log\{p(\mathbf{y}|\mathbf{x},\mathbf{f})\} > \log\{p(\mathbf{y}|\hat{\mathbf{x}},\mathbf{f})\} \qquad \forall\, \hat{\mathbf{x}} \neq \mathbf{x}. \tag{7.20}$$

The noise samples at the matched filter output in (7.18), have the discrete autocorrelation function

$$\phi_{vv}(n) = \frac{1}{2}\mathrm{E}[v_{k+n}v_k^*] = N_0 f_n. \tag{7.21}$$

Hence, the noise sequence $\{v_k\}$ will be correlated unless $f_n = 0, n \neq 0$, meaning that the overall pulse

$$f(t) = \int_{-\infty}^{\infty} h(t+\tau)h^*(\tau)\mathrm{d}\tau, \tag{7.22}$$

satisfies the first Nyquist criterion. Such a condition will not be true for ISI channels due to the nonideal channel $g(t)$, and the resulting correlation between the noise samples $\{v_k\}$ results in a log-likelihood function $\log\{p(\mathbf{y}|\mathbf{x},\mathbf{f})\}$ that has a complicated form. This difficulty can be overcome by passing the sample sequence at the output of the matched filter, $\{y_k\}$, through a noise-whitening filter as described below, to whiten the noise samples.

The z-transform of the vector $\mathbf{f}$ is

$$F(z) = \sum_{n=-L}^{L} f_n z^{-n}. \tag{7.23}$$

Using the property $f_n^* = f_{-n}$, we can write

$$F^*(1/z^*) = F(z). \tag{7.24}$$

This means that if z is a root of $F(z)$, then $1/z^*$ is a root of $F(z)$, that is, the roots of $F(z)$ occur in conjugate reciprocal pairs. It follows that $F(z)$ has $2L$ roots with the factorization

$$F(z) = G(z)G^*(1/z^*), \tag{7.25}$$

where $G(z)$ and $G^*(1/z^*)$ are polynomials each of degree L. The roots of $G(z)$ are $z_1, z_2, \ldots, z_L$, while the roots of $G^*(1/z^*)$ are $1/z_1^*, 1/z_2^*, \ldots, 1/z_L^*$. Hence, there are 2^L possible choices for the roots of $G^*(1/z^*)$, and any one will suffice for a noise-whitening filter $1/G^*(1/z^*)$. However, some reduced state equalization techniques such as RSSE and DDFSE require that the polynomial of the overall system $G(z) = F(z) \cdot 1/G^*(1/z^*)$ be minimum-phase, meaning all the poles *and* zeroes of $G(z)$ lie inside the unit circle. . For such cases, the noise-whitening filter $1/G^*(1/z^*)$

will be a stable filter, but it is noncausal since all its poles are outside the unit circle. In practice, such a noncausal noise-whitening filter can be approximated to sufficient accuracy using a long enough filter delay. If the overall response $G(z)$ need not have minimum phase, then we can choose $G^*(1/z^*)$ to have minimum phase, that is, all the poles and zeros of the noise-whitening filter $1/G^*(1/z^*)$ are inside the unit circle. This choice will ensure that the noise-whitening filter is both causal and stable.

If the noise-whitening filter is chosen such that $G(z)$ has minimum phase, then the resulting discrete-time white noise channel satisfies the *minimum energy-delay property*. To explain this further, let $G_{\min}(e^{j2\pi f})$ be the frequency response function corresponding to the $G(z)$ having minimum phase, and let $g_{k\ \min}$ be the corresponding time-domain impulse response. All 2^L choices for $G(z)$ will have the same magnitude response, that is, $|G(e^{j2\pi f})| = |G_{\min}(e^{j2\pi f})|$. Consequently, all impulse responses g_k whose magnitude response $|G(e^{j2\pi f})|$ is equal to $|G_{\min}(e^{j2\pi f})|$ will have the same total energy by Parseval's theorem, that is,

$$\sum_{n=0}^{\infty} |g_k|^2 = \int_{-1/2}^{1/2} |G(e^{j2\pi f})|^2 \mathrm{d}f = \int_{-1/2}^{1/2} |G_{\min}(e^{j2\pi f})|^2 \mathrm{d}f = \sum_{n=0}^{\infty} |g_{k\ \min}|^2. \quad (7.26)$$

If we define the partial energy of the impulse response as

$$E(k) \stackrel{\triangle}{=} \sum_{n=0}^{k} |g_k|^2, \quad (7.27)$$

then it can be shown that [198],

$$E(k) = \sum_{n=0}^{k} |g_k|^2 \leq \sum_{n=0}^{k} |g_{k\ \min}|^2 = E_{\min}(k), \quad (7.28)$$

for all impulse responses g_k that have the same magnitude response. Accordingly, the energy of the system having minimum phase is most concentrated around $k = 0$. This means that the energy of the minimum phase system has the least delay among all systems that have the same magnitude response function. For this reason, the minimum phase system is said to satisfy the minimum energy-delay property.

Example 7.1:

Consider a simple T-spaced two-ray channel where the received pulse is

$$h(t) = h_{\mathrm{a}}(t) + a h_{\mathrm{a}}(t - T)$$

and the transmitted pulse $h_{\mathrm{a}}(t)$ has duration T and is normalized to have unit energy, that is, $\int_{-\infty}^{\infty} h_{\mathrm{a}}^2(t)\mathrm{d}t = 1$. The corresponding ISI coefficients are

$$f_\ell = \int_{-\infty}^{\infty} h^*(t)h(t+\ell T)\mathrm{d}t$$
$$= \begin{cases} 1+|a|^2 & \ell = 0 \\ a & \ell = 1 \\ a^* & \ell = -1 \end{cases}$$

and, hence,

$$F(z) = a^*z + (1+|a|^2) + az^{-1}$$
$$= (az^{-1}+1)(a^*z+1).$$

There are two possible choices for the noise-whitening filter.

Case 1: Under the assumption that $|a| < 1$, suppose that the zero of $G^*(1/z^*)$ is chosen to be outside the unit circle. That is,

$$G(z) = 1 + az^{-1},$$
$$G^*(1/z^*) = 1 + a^*z.$$

In this case, the noise-whitening filter $1/G^*(1/z^*)$ is noncausal yet stable, and the overall system is characterized by the minimum phase polynomial

$$G(z) = 1 + az^{-1}.$$

Note that the zero of $G(z)$ is inside the unit circle, with a pole at the origin.
Case 2: Under the assumption that $|a| < 1$, suppose that the zero of $G^*(1/z^*)$ is instead chosen to be inside the unit circle. That is,

$$G(z) = 1 + a^*z,$$
$$G^*(1/z^*) = 1 + az^{-1}.$$

In this case, the noise-whitening filter $1/G^*(1/z^*)$ is a minimum phase filter that is both stable and causal. However, the overall system $G(z)$ is characterized by the non-minimum phase polynomial

$$G(z) = 1 + a^*z.$$

Note that the zero of $G(z)$ is outside the unit circle with a pole at infinity.

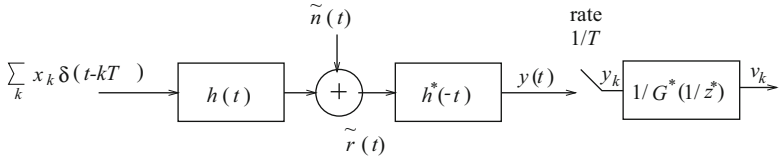

Fig. 7.3 Block diagram of system that implements a filter matched to $h(t)$ followed by a discrete-time noise-whitening filter

We now show that the sequence of noise samples at the output of the noise-whitening filter is indeed uncorrelated. From (7.17), the z-transform of the sample sequence at the output of the noise-whitening filter is

$$\begin{aligned} V(z) &= (X(z)F(z)+\nu(z))\frac{1}{G^*(1/z^*)} \\ &= X(z)G(z)+\nu(z)\frac{1}{G^*(1/z^*)} \\ &= X(z)G(z)+W(z). \end{aligned} \tag{7.29}$$

From (7.21), the psd of the noise samples $\{v_k\}$ at the input to the noise-whitening filter is

$$S_{vv}(e^{j2\pi fT}) = N_o F(e^{j2\pi fT}), \quad |f| \le \frac{1}{2T}. \tag{7.30}$$

Therefore, the psd of the noise samples $\{w_k\}$ at the output of noise-whitening filter $1/G^*(1/z^*)$ is

$$\begin{aligned} S_{ww}(e^{j2\pi fT}) &= N_o \frac{F(e^{j2\pi fT})}{|G^*(e^{j2\pi fT})|^2} \\ &= N_o \frac{G(e^{j2\pi fT})G^*(e^{j2\pi fT})}{G(e^{j2\pi fT})G^*(e^{j2\pi fT})} \\ &= N_o, \quad |f| \le \frac{1}{2T}, \end{aligned} \tag{7.31}$$

which is clearly white.

The above development leads to the system shown in Fig. 7.3, and the discrete-time white noise channel model shown in Fig. 7.4. Sometimes the concatenation of the matched filter and noise-whitening filter in Fig. 7.3 is called a whitened matched filter. The overall system system function $G(z)$ can be viewed as a finite impulse response (FIR) filter with tap coefficients $\{g_n\}$. The discrete-time samples at the output of the noise-whitening filter are

$$v_k = \sum_{n=0}^{L} g_n x_{k-n} + w_k. \tag{7.32}$$

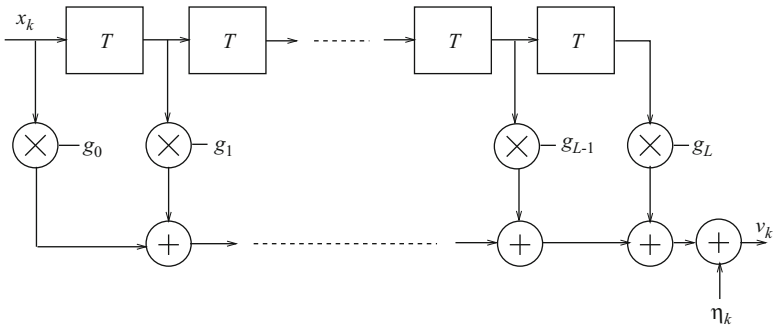

Fig. 7.4 Discrete-time white noise channel model

The maximum likelihood receiver uses the observation vector $\mathbf{v} = \{v_k\}_{k=0}^{L}$ to decide in favor of the symbol sequence $\mathbf{x}$ that maximizes the likelihood function $p(\mathbf{v}|\mathbf{x},\mathbf{g})$ or the log-likelihood function $\log\{p(\mathbf{v}|\mathbf{x},\mathbf{g})\}$, that is,

$$\text{choose } \mathbf{x} \text{ if } \quad \log\{p(\mathbf{v}|\mathbf{x},\mathbf{g})\} > \log\{p(\mathbf{v}|\hat{\mathbf{x}},\mathbf{g})\} \qquad \forall\, \hat{\mathbf{x}} \neq \mathbf{x}, \tag{7.33}$$

where

$$\mathbf{g} = (g_0,\, g_1, \ldots,\, g_L)^{\mathrm{T}}, \tag{7.34}$$

is the overall channel impulse response. Since the noise samples $\{w_k\}$ are white, the likelihood function has the simple product form

$$p(\mathbf{v}|\mathbf{x},\mathbf{g}) = \prod_k \frac{1}{2\pi N_\mathrm{o}} \exp\left\{ -\frac{1}{2N_\mathrm{o}} \left| v_k - \sum_{n=0}^{L} g_n x_{k-n} \right|^2 \right\}. \tag{7.35}$$

The log-likelihood function $\log\{p(\mathbf{v}|\mathbf{x},\mathbf{g})\}$ results in the decision rule

$$\text{choose } \mathbf{x} \text{ if } \quad \mu(\mathbf{x}) > \mu(\hat{\mathbf{x}}) \qquad \forall\, \hat{\mathbf{x}} \neq \mathbf{x}, \tag{7.36}$$

where

$$\mu(\mathbf{x}) = -\sum_k \left| v_k - \sum_{n=0}^{L} g_n x_{k-n} \right|^2. \tag{7.37}$$

An efficient method for finding the sequence $\mathbf{x}$ is the Viterbi algorithm as discussed in Sect. 7.4.1.

Finally, for an ISI channel, the received symbol energy-to-noise ratio is defined as

$$\gamma_s = \frac{\mathrm{E}[|x_k|^2] \sum_{i=0}^{L} |g_i|^2}{2N_\mathrm{o}} = \frac{\mathrm{E}[|x_k|^2] f_0}{2N_\mathrm{o}} = \frac{2\sigma_x^2 E_{h_\mathrm{r}}}{N_\mathrm{o}} = \frac{E}{N_\mathrm{o}}, \tag{7.38}$$

where

$$E_{h_r} = \frac{1}{2}\int_{-\infty}^{\infty} |h(t)|^2 \mathrm{d}t \tag{7.39}$$

is the energy in the received pulse $h(t)$. The bit energy-to-noise ratio is $\gamma_b = \gamma_s / \log_2 M$ where M is the modulation alphabet size.

7.2.1.1 Slowly Fading ISI Channels with Diversity

Consider a fading channel with D-branch receiver diversity. The received pulse on each diversity branch is equal to the convolution

$$h_d(t) = \int_{-\infty}^{\infty} g_d(t,\tau) h_a(t-\tau)\mathrm{d}\tau\ ,\ \ d = 1,\ldots,D, \tag{7.40}$$

where $g_d(t,\tau)$ is the time-variant channel impulse response for branch d. For slow fading, the channel impulse responses $g_d(t,\tau)$ can be assumed to change slowly with respect to the duration of the received pulses. When data is transmitted in short frames, for example, 10–20 ms long, the channel may remain constant over the duration of the frame. This is sometimes called a block fading channel or quasi-static fading channel. In any case, at the kth epoch the received pulses can be accurately approximated as

$$h_{d,k}(t) = \int_{-\infty}^{\infty} g_d(kT,\tau) h_a(t-\tau)\mathrm{d}\tau\ ,\ \ d = 1,\ldots,D. \tag{7.41}$$

The receiver then implements a matched filter on each diversity branch having the impulse response $h^*_{d,k}(-t)$, and samples are taken at the output of the matched filter every T seconds. The samples at the output of each matched filter are passed through a corresponding noise-whitening filter $1/G^*_{d,k}(1/z^*)$. This results in the discrete-time white noise channel model shown in Fig. 7.5. At epoch k, the tap gains associated with diversity branch d are described by the vector

$$\mathbf{g}_d(k) = (g_{0,d}(k),\ g_{1,d}(k),\ \ldots\ g_{L,d}(k))^{\mathrm{T}}. \tag{7.42}$$

The $\{g_{i,d}(k)\}$ are discrete-time complex Gaussian random processes that are generally correlated with the correlation matrix

$$\Phi_{\mathbf{g}_d}(m) = \frac{1}{2}\mathrm{E}[\mathbf{g}_d(k)\mathbf{g}_d^H(k+m)], \tag{7.43}$$

where $\mathbf{x}^H$ is the complex conjugate transpose of the vector $\mathbf{x}$. The received sample on branch d at epoch k is

$$v_{k,d} = \sum_{i=0}^{L} g_{i,d}(k) x_{k-i} + w_{k,d}, \tag{7.44}$$

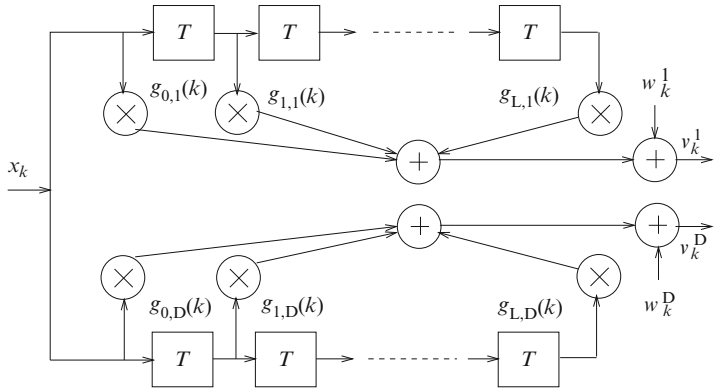

Fig. 7.5 Discrete-time white noise channel model with D-branch diversity

where the $w_{k,d}$ are independent complex zero-mean white Gaussian noise samples with variance $\frac{1}{2}\mathrm{E}[|w_{k,d}|^2] = N_\mathrm{o}$.

For a fading ISI channel, the average received symbol energy-to-noise ratio for branch d is

$$\bar{\gamma}_s^d = \frac{\mathrm{E}[|x_k|^2]\sum_{i=0}^{L}\mathrm{E}[|g_{i,d}|^2]}{2N_\mathrm{o}} = \frac{\mathrm{E}[|x_k|^2]\mathrm{E}[f_{0,d}]}{N_\mathrm{o}} = \frac{2\sigma_x^2\mathrm{E}[E_{h,d}]}{N_\mathrm{o}} = \frac{\bar{E}}{N_\mathrm{o}}. \tag{7.45}$$

In many cases, the branches are balanced so that $\bar{\gamma}_s^d = \bar{\gamma}_s,\ d = 1,\ldots,D$. The averaged received branch bit energy-to-noise ratio is $\bar{\gamma}_s/\log_2 M$.

Note that the matched filter and noise-whitening filter impulse responses change slowly with time due to variations in the underlying channels. This presents a practical difficulty because implementation and adjustment of the matched filter and noise-whitening filter require knowledge of the underlying channel. Later we will show that we can overcome this difficulty by implementing a filter that is matched to the transmitted pulse $h_\mathrm{a}(t)$, over-sampling the output, and processing the output samples with a fractionally spaced noise-whitening filter. First, we consider the effect of over-sampling the matched filter output.

7.2.1.2 *T*/2-Spaced Receiver

In practice, the matched filter outputs are often over-sampled for the purpose of symbol timing synchronization and to mitigate the effects of timing errors. One important example that will be considered at various points in this chapter is when the output of the matched filter is sampled with rate $2/T$. In this case, the overall channel impulse response and sampler can be represented by a discrete-time transversal filter with coefficients

$$\mathbf{f}^{(2)} = \left(f_{-2L}^{(2)},\, f_{-2L+1}^{(2)},\, \cdots,\, f_{-1}^{(2)},\, f_0^{(2)},\, f_1^{(2)},\, \cdots,\, f_{2L-1}^{(2)},\, f_{2L}^{(2)} \right), \tag{7.46}$$

where $(\,\cdot\,)^{(2)}$ indicates rate $2/T$ sampling. If the samples in (7.46) are obtained with the correct timing phase, that is, $f_n^{(2)} = f(nT/2)$, then

$$\begin{aligned} \mathbf{f} &= (\, f_{-L},\, f_{-L+1},\, \cdots,\, f_{-1},\, f_0,\, f_1,\, \cdots,\, f_{L-1},\, f_L \,) \\ &= \left(f_{-2L}^{(2)},\, f_{-2L+2}^{(2)},\, \cdots,\, f_{-2}^{(2)},\, f_0^{(2)},\, f_2^{(2)},\, \cdots,\, f_{2L-2}^{(2)},\, f_{2L}^{(2)} \right), \end{aligned} \tag{7.47}$$

where $f_n^{(2)} = \left(f_{-n}^{(2)} \right)^*$ and $f_n = f_{2n}^{(2)}$. More details on timing phase sensitivity will be provided in Sect. 7.6.3.

The $T/2$-spaced noise samples at the matched filter output have the discrete-time autocorrelation function

$$\phi_{vv}(n) = N_o f_n^{(2)}. \tag{7.48}$$

The z-transform of $\mathbf{f}^{(2)}$, denoted as $F^{(2)}(z)$, has $4L$ roots with the factorization

$$F^{(2)}(z) \;=\; G^{(2)}(z) G^{(2)^*}(1/z^*), \tag{7.49}$$

where $G^{(2)}(z)$ and $G^{(2)^*}(1/z^*)$ are polynomials of degree $2L$ having conjugate reciprocal roots. The correlated noise samples can be whitened using a filter with transfer function $1/G^{(2)^*}(1/z^*)$. Once again, $G^{(2)^*}(1/z^*)$ can be chosen as a noncausal stable filter such that the overall system function $G^{(2)}(z)$ has minimum phase with all its roots inside the unit circle. The output of the noise-whitening filter is

$$v_n^{(2)} = \sum_{k=0}^{2L} g_k^{(2)} x_{n-k}^{(2)} + w_n^{(2)}, \tag{7.50}$$

where $\{w_{\rm n}^{(2)}\}$ is a white Gaussian noise sequence with variance $\frac{1}{2}\mathrm{E}[|w_{\rm n}^{(2)}|^2] = N_o$ and the $\{g_{\rm n}^{(2)}\}$ are the coefficients of a discrete-time transversal filter having a transfer function $G^{(2)}(z)$. The sequence $\{x_{\rm n}^{(2)}\}$ is the corresponding $T/2$-spaced input symbol sequence and is given by

$$x_{\rm n}^{(2)} \;=\; \begin{cases} x_{n/2}\,, & n = 0, 2, 4, \ldots \\ 0\,, & n = 1, 3, 5, \ldots \end{cases} \tag{7.51}$$

Note that each transmitted symbol is padded with a zero. In general, if rate K/T sampling is used, then each input symbol is padded with $K-1$ zeros. The overall system and equivalent discrete-time white noise channel models are shown in Figs. 7.6 and 7.7, respectively.

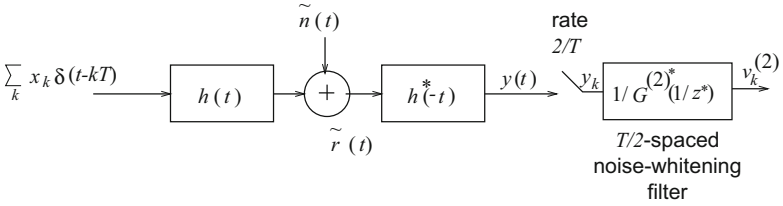

Fig. 7.6 Block diagram of system that implements a filter matched to $h(t)$ followed by a $T/2$-spaced sampler and a discrete-time noise-whitening filter

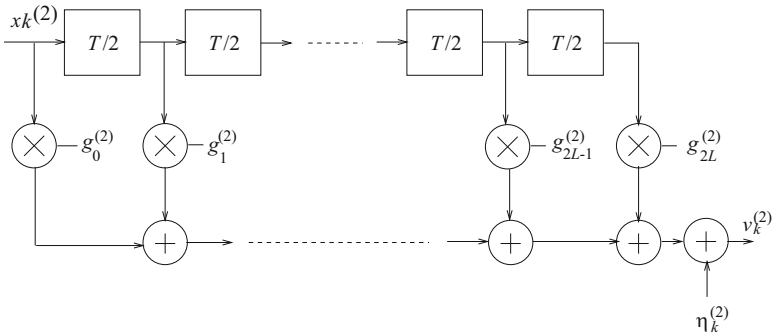

Fig. 7.7 Discrete-time white noise channel model with rate-$2/T$ sampling

Comparing (7.25) and (7.49), we have

$$\sum_{k=0}^{2L} |g_k^{(2)}|^2 = \sum_{k=0}^{L} |g_k|^2 = f_0^{(2)} = f_0. \tag{7.52}$$

Finally, we note that the samples $v_{2n}^{(2)}$ and $v_{2n+1}^{(2)}$ correspond to the nth received baud, where

$$v_{2n}^{(2)} = \sum_{k=0}^{L} g_{2k}^{(2)} x_{n-k} + w_{2n}^{(2)}, \tag{7.53}$$

$$v_{2n+1}^{(2)} = \sum_{i=0}^{L-1} g_{2k+1}^{(2)} x_{n-k} + w_{2n+1}^{(2)}. \tag{7.54}$$

Finally, by comparing (7.32) and (7.54), we note that $v_{2\mathrm{n}}^{(2)}$ is not necessarily equal to v_{n} because a different noise-whitening filter is used to whiten the $T/2$-spaced noise samples than that is used to whiten T-spaced noise samples.

7.3 Symbol-by-Symbol Equalizers

7.3.1 Linear Equalizer

As shown in Fig. 7.8, a linear forward equalizer consists of a transversal filter with adjustable tap coefficients. The tap coefficients of the equalizer are denoted by the column vector

$$\mathbf{c} = (c_0,\ c_1,\ \cdots,\ c_{N-1})^{\mathrm{T}}, \tag{7.55}$$

where N is the number of equalizer taps. Assuming that the equalizer is preceded by a whitened matched filter that outputs the sequence $\{v_{\mathrm{n}}\}$, the output of the equalizer is

$$\tilde{x}_n = \sum_{j=0}^{N-1} c_j v_{n-j}, \tag{7.56}$$

where the v_{n} are given by (7.32). The equalizer output $\tilde{x}_k$ is quantized to the nearest (in Euclidean distance) information symbol to form the decision $\hat{x}_k$.

Observe that the overall discrete-time white noise channel and equalizer can be represented by a single filter having the sampled impulse response

$$q = (q_0, q_1,\ \ldots,\ q_{N+L-1})^{\mathrm{T}}, \tag{7.57}$$

where

$$\begin{aligned} q_n &= \sum_{j=0}^{N-1} c_j g_{n-j}, \\ &= \mathbf{c}^T \mathbf{g}(n) \end{aligned} \tag{7.58}$$

with

$$\mathbf{g}(n) = (g_n, g_{n-1}, g_{n-2}, \ldots, g_{n-N+1})^{\mathrm{T}}, \tag{7.59}$$

and $g_i = 0, i < 0, i > L$. That is, q is the discrete convolution of $\mathbf{g}$ and $\mathbf{c}$.

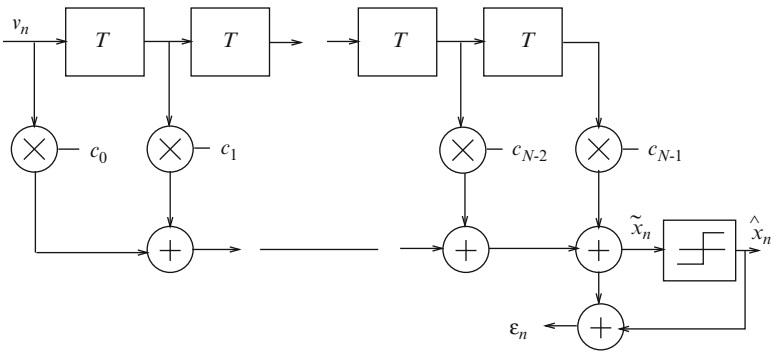

Fig. 7.8 Linear transversal equalizer with adjustable T-spaced taps

If the equalizer is preceded by a noise-whitening filter, then the discrete-time white noise channel has a system function $G(z)$ with minimum phase. Although the minimum phase system function $G(z)$ satisfies the minimum energy-delay property as discussed previously, it is not necessarily true that $|g_0|^2 \geq |g_k|^2, \forall k \geq 1$, that is, the first tap g_0 does not necessarily have the largest magnitude. Let the component of $\mathbf{g}$ of greatest magnitude be denoted by g_{d_1}. Also, let the number of equalizer taps be equal to $N = 2d_2 + 1$ where d_2 is an integer. Perfect equalization means that

$$q = \mathbf{e}_d = (\underbrace{0,\ 0,\ \ldots,\ 0}_{d-1 \text{ zeroes}},\ 1,\ 0,\ \ldots,\ 0,\ 0)^{\mathrm{T}}, \tag{7.60}$$

where $d-1$ zeroes precede the "1" and d is an integer representing the overall delay. Unfortunately, perfect equalization is difficult to achieve and does not always yield the best performance.

7.3.1.1 Zero-Forcing Solution

Lucky [165, 166] was the first to develop an adaptive (linear) equalizer for digital communication systems in the mid-1960s. This equalizer was based on the peak distortion criterion, where the equalizer forces the ISI to zero, and it is called a zero-forcing (ZF) equalizer. With a ZF equalizer, the tap coefficients $\mathbf{c}$ are chosen to minimize the peak distortion of the equalized channel, defined as

$$D_p = \frac{1}{|q_d|} \sum_{\substack{n=0 \\ n \neq d}}^{N+L-1} |q_n - \hat{q}_n|, \tag{7.61}$$

where $\hat{q} = (\hat{q}_0,\ \ldots,\ \hat{q}_{N+L-1})^{\mathrm{T}}$ is the desired equalized channel, and the delay d is a positive integer optimized to have the value $d = d_1 + d_2$ [57]. Lucky showed that if the initial distortion before equalization is less than unity, that is,

$$D = \frac{1}{|g_{d_1}|} \sum_{\substack{n=0 \\ n \neq d_1}}^{L} |g_n| < 1, \tag{7.62}$$

then D_p is minimized by those N tap values which simultaneously cause $q_j = \hat{q}_j$ for $d - d_2 \leq j \leq d + d_2$. However, if the initial distortion before equalization is greater than unity, the ZF criterion is not guaranteed to minimize the peak distortion. For the case when $\hat{q} = \mathbf{e}_d$, the equalized channel is given by

$$q = (q_0, \ldots, q_{d_1-1}, 0, \ldots, 0, 1, 0, \ldots, 0, q_{d_1+N}, \ldots, q_{N+L-1})^{\mathrm{T}}. \tag{7.63}$$

In this case the equalizer forces zeroes into the equalized channel, and hence the name "zero-forcing equalizer."

Equalizer Tap Solution

For a known channel impulse response, the tap gains of the ZF equalizer can be found by the direct solution of a simple set of linear equations [57]. To do so, we form the matrix

$$\mathbf{G} = [\mathbf{g}(d_1), \ldots, \mathbf{g}(d), \ldots, \mathbf{g}(N + d_1 - 1)] \tag{7.64}$$

and the vector

$$\tilde{q} = \left(\hat{q}_{d_1}, \ldots, \hat{q}_d, \ldots, \hat{q}_{N+d_1-1}\right)^{\mathrm{T}}. \tag{7.65}$$

Then the vector of optimal tap gains, $\mathbf{c}_{\mathrm{op}}$, satisfies

$$\mathbf{c}_{\mathrm{op}}^T \mathbf{G} = \tilde{q}^T \longrightarrow \mathbf{c}_{\mathrm{op}} = (\mathbf{G}^{-1})^{\mathrm{T}} \tilde{q}. \tag{7.66}$$

Example 7.2:

Suppose that a system has the channel vector

$$\mathbf{g} = (0.90, -0.15, 0.20, 0.10, -0.05)^{\mathrm{T}},$$

where $g_i = 0, i < 0, i > 4$. The initial distortion before equalization is

$$D = \frac{1}{|g_0|} \sum_{n=1}^{4} |g_n| \; = \; 0.5555$$

and, therefore, the minimum distortion is achieved with the ZF solution. Suppose that we wish to design a three-tap ZF equalizer. Since g_0 is the component of $\mathbf{g}$ having the largest magnitude, $d_1 = 0$ and the equalizer delay is chosen as $d = d_1 + d_2 = 1$. Suppose that the desired response is $\hat{\mathbf{q}}_0 = \mathbf{e}_1^T$, so that $\tilde{\mathbf{q}} = (0,\ 1,\ 0)$. We then construct the matrix

$$\begin{aligned} \mathbf{G} &= [\mathbf{g}(0), \mathbf{g}(1), \mathbf{g}(2)] \\ &= \begin{bmatrix} 0.90 & -0.15 & 0.20 \\ 0.00 & 0.90 & -0.15 \\ 0.00 & 0.00 & 0.90 \end{bmatrix} \end{aligned}$$

and obtain the optimal tap solution

$$\mathbf{c}_{\mathrm{op}} = (\mathbf{G}^{-1})^T \tilde{\mathbf{q}} = (0,\ 1.1111,\ -0.185185)^{\mathrm{T}}.$$

The overall response of the channel and equalizer is

$$q = (0.0,\ 1.0,\ 0,\ 0.194,\ 0.148,\ -0.037,\ -0.009,\ 0,\ \ldots)^{\mathrm{T}}.$$

Finally, the distortion after equalization is

$$D_{\mathrm{p}} = \frac{1}{|q_1|} \sum_{\substack{n=0 \\ n \neq 1}}^{6} |q_n - \hat{q}_n| = 0.388.$$

Adaptive Solution

In practice, the channel impulse response is unknown to the receiver, and a known finite length sequence $\mathbf{x}$ is used to train the equalizer. During this training mode, the equalizer taps can be obtained using the following steepest-descent recursive algorithm:

$$c_j^{n+1} = c_j^n + \alpha \varepsilon_n x_{n-j-d_1}^*, \qquad j = 0, \ldots, N-1, \tag{7.67}$$

where

$$\begin{aligned} \varepsilon_n &= x_{n-d} - \tilde{x}_n \\ &= x_{n-d} - \sum_{i=0}^{N-1} c_i v_{n-i} \end{aligned} \tag{7.68}$$

is the error sequence, $\{c_j^n\}$ is the set of equalizer tap gains at epoch n, and α is an adaptation step-size that can be optimized to trade-off convergence rate and steady-state bit error rate performance. Notice that adaptation rule in (7.67) attempts to force the cross-correlations $\varepsilon_n x_{n-j-d_1}^*$, $j = 0, \ldots, N-1$, to zero. To see that (7.67) leads to the desired solution, we use (7.32) and (7.68), along with the fact that the symbol sequence $\{x_{\mathrm{n}}\}$ is uncorrelated with the noise sequence $\{w_{\mathrm{n}}\}$ to obtain

$$\begin{aligned} \frac{1}{2}\mathrm{E}[\varepsilon_n x_{n-j-d_1}^*] &= \frac{1}{2}\mathrm{E}[x_{n-d} x_{n-j-d_1}^*] - \frac{1}{2} \sum_{i=0}^{N-1} \sum_{\ell=0}^{L} c_i g_\ell \mathrm{E}[x_{n-i-\ell} x_{n-j-d_1}^*] \\ &= \sigma_x^2 \left(\delta_{d_2-j} - \sum_{i=0}^{N-1} c_i g_{j+d_1-i} \right) \\ &= \sigma_x^2 (\delta_{d_2-j} - q_{j+d_1}), \quad j = 0, 1, \ldots, N-1, \end{aligned} \tag{7.69}$$

where $\sigma_x^2 = \frac{1}{2}\mathrm{E}[|x_k|^2]$. Note that the conditions $\frac{1}{2}\mathrm{E}[\varepsilon_n x_{n-j-d_1}^*] = 0$ are satisfied when $q_d = 1$ and $q_i = 0$ for $d - d_2 \leq i < d$ and $d < i \leq d + d_2$, which is the zero forcing solution.

After training the equalizer, a decision-feedback mechanism is typically used where the sequence of symbol decisions $\hat{\mathbf{x}}$ is used to update the tap coefficients.

This mode is called the data mode and allows the equalizer to track variations in the channel vector **g**. In the data mode,

$$c_j^{n+1} = c_j^n + \alpha \varepsilon_n \hat{x}_{n-j-d_1}^*, \quad j = 0, \ldots, N-1, \tag{7.70}$$

where the error term ε_n in (7.68) becomes

$$\varepsilon_n = \hat{x}_{n-d} - \sum_{i=0}^{N-1} c_i v_{n-i}, \tag{7.71}$$

and, again, $\hat{x}_{n-d}$ is the decision on the equalizer output $\tilde{x}_n$ delayed by d samples.

Performance of the ZF Equalizer

If the ZF equalizer has an infinite number of taps, it is possible to select the tap weights so that $D_\mathrm{p} = 0$, that is, $q = \hat{q}$. Assuming that $\hat{q}_\mathrm{n} = \delta_{\mathrm{n}0}$, this condition means that

$$Q(z) = 1 = C(z)G(z). \tag{7.72}$$

Therefore,

$$C(z) = \frac{1}{G(z)} \tag{7.73}$$

and the ideal ZF equalizer has a discrete transfer function that is simply the inverse of overall channel $G(z)$. The cascade of the noise-whitening filter with transfer function

$$W(z) = \sqrt{2E_{h_\mathrm{r}}}/G^*(1/z^*), \tag{7.74}$$

and the ZF equalizer with transfer function $1/G(z)$ results in an equivalent equalizer with transfer function[1]

$$C'(z) = \frac{\sqrt{2E_{h_\mathrm{r}}}}{G^*(1/z^*)G(z)} = \frac{\sqrt{2E_{h_\mathrm{r}}}}{F(z)}. \tag{7.75}$$

Recall from (7.30) that the noise sequence at the input to the equivalent equalizer $C'(z)$ has the discrete autocorrelation function $\phi_{vv}(n) = \frac{N_\mathrm{o}}{2E_{h_\mathrm{r}}} f_\mathrm{n}$ and psd

$$S_{\tilde{n}\tilde{n}}(f) = \frac{N_\mathrm{o}}{2E_{h_\mathrm{r}}} F(\mathrm{e}^{\mathrm{j}2\pi fT}), \quad |f| \le \frac{1}{2T}. \tag{7.76}$$

Therefore, the psd of the noise sequence $\{\zeta_n\}$ at the output of the equalizer is

$$S_{\zeta\zeta}(f) = \frac{N_\mathrm{o}}{F(\mathrm{e}^{\mathrm{j}2\pi fT})}, \quad |f| \le \frac{1}{2T} \tag{7.77}$$

[1]The scaling of the noise-whitening filter gain by $\sqrt{2E_{h_\mathrm{r}}}$ is not necessary in practice and is done here for mathematical convenience.

and the noise samples have variance

$$\sigma_\zeta^2 = T\int_{-1/2T}^{1/2T} S_{\zeta\zeta}(f)\mathrm{d}f$$
$$= T\int_{-1/2T}^{1/2T} \frac{N_\mathrm{o}}{F(\mathrm{e}^{\mathrm{j}2\pi fT})}\mathrm{d}f. \tag{7.78}$$

If $\sigma_x^2 = \frac{1}{2}\mathrm{E}[|x_k|^2]$ and $\hat{q}_\mathrm{n} = \delta_{\mathrm{n}0}$, then the signal-to-noise ratio at the output of the infinite-tap equalizer is

$$\gamma_\infty = \frac{\sigma_x^2}{\sigma_\zeta^2}. \tag{7.79}$$

Finally, we can show that (Problem 7.1)

$$F(\mathrm{e}^{\mathrm{j}2\pi fT}) = F_\Sigma(f), \quad |f| \le \frac{1}{2T}, \tag{7.80}$$

where $F_\Sigma(f)$ is the folded spectrum of $F(f)$ defined by

$$F_\Sigma(f) \stackrel{\triangle}{=} \frac{1}{T}\sum_{n=-\infty}^{\infty} F\left(f+\frac{n}{T}\right) \tag{7.81}$$

and $F(f)$ is the Fourier transform of the pulse $f(t) = h(t) * h^*(-t)$. Hence, the signal-to-noise ratio at the equalizer output can be written in the final form

$$\gamma_\infty = \sigma_x^2\left(TN_\mathrm{o}\int_{-1/2T}^{1/2T}\left(\frac{1}{T}\sum_{n=-\infty}^{\infty} F\left(f+\frac{n}{T}\right)\right)^{-1}\mathrm{d}f\right)^{-1}. \tag{7.82}$$

It is clear from (7.82) that ZF equalizers are unsuitable for channels that have severe ISI, where the folded spectrum $F_\Sigma(f)$ has spectral nulls or very small values. Under these conditions, the equalizer tries to compensate for the nulls in the folded spectrum by introducing infinite gain at their frequencies. Unfortunately, this results in severe noise enhancement at the output of the equalizer at these same frequencies. Land mobile radio channels often exhibit spectral nulls, and, therefore, linear ZF equalizers are not used for land mobile radio applications.

On the other hand, when the overall channel $f(t)$ satisfies the conditions for ISI-free transmissions as discussed in Sect. 4.2, then the ISI coefficients satisfy the property $f_\mathrm{n} = f_0\delta_{n0}$, and the matched filter output is

$$\tilde{r}_k = \sqrt{2E_{h_\mathrm{r}}}x_k + \tilde{n}_k. \tag{7.83}$$

The noise samples $\{\tilde{n}_k\}$ in the case are white due to the fact that the overall pulse $f(t)$ satisfies the first Nyquist criterion. From Sect. 4.2, an equivalent condition in the frequency-domain is that the folded spectrum $F_\Sigma(f)$ is flat, that is,

$$F_\Sigma(f) = \frac{1}{T}\sum_{n=-\infty}^{\infty} F\left(f + \frac{n}{T}\right) = f_0 = 2E_{h_\mathrm{r}}. \tag{7.84}$$

Under this condition, the signal-to-noise ratio in (7.82) reaches its maximum value $\gamma_\infty = 2\sigma_x^2 E_{h_\mathrm{r}}/N_\mathrm{o}$.

7.3.1.2 Minimum Mean-Square-Error Solution

Soon after Lucky introduced the ZF equalizer, Proakis and Miller [216], Lucky et. al. [167], and Gersho [109] developed the linear MMSE equalizer based on the least mean square (LMS) criterion. The MMSE equalizer is more robust and superior to the ZF equalizer in its performance and convergence properties [216, 217, 220]. By defining the vector

$$\mathbf{v}_n = (v_\mathrm{n},\ v_{n-1},\ \ldots,\ v_{n-N+1}), \tag{7.85}$$

where v_k is the output of the whitened matched filter in (7.32), the output of the equalizer in (7.56) can be expressed in the form

$$\tilde{x}_n = \mathbf{c}^T\mathbf{v}_n = \mathbf{v}_n^T\mathbf{c}. \tag{7.86}$$

An MMSE equalizer adjusts the tap coefficients to minimize the mean square error (MSE)

$$\begin{aligned} J &\stackrel{\triangle}{=} \frac{1}{2}\mathrm{E}[|x_{n-d} - \tilde{x}_n|^2] \\ &= \frac{1}{2}\mathrm{E}\left[\mathbf{c}^T\mathbf{v}_n\mathbf{v}_n^H\mathbf{c}^* - 2\mathrm{Re}\{\mathbf{v}_n^H\mathbf{c}^* x_{n-d}\} + |x_{n-d}|^2\right], \end{aligned} \tag{7.87}$$

where, again, d is the equalizer delay assumed here to be chosen as $d = d_1 + d_2$.

Equalizer Tap Solution

If the channel impulse response is known, the optimum equalizer taps can be obtained by direct solution. Define

$$\begin{aligned} \mathbf{M}_v &\stackrel{\triangle}{=} \frac{1}{2}\mathrm{E}[\mathbf{v}_n\mathbf{v}_n^H], \\ \mathbf{v}_x^H &\stackrel{\triangle}{=} \frac{1}{2}\mathrm{E}[\mathbf{v}_n^H x_{n-d}] \end{aligned} \tag{7.88}$$

where $\mathbf{M}_v$ is an $N \times N$ Hermitian matrix (meaning that $\mathbf{M}_v = \mathbf{M}_v^H$) and $\mathbf{v}_x$ is a length N column vector. Using these definitions and assuming that $\frac{1}{2}\mathrm{E}[|x_{n-d}|^2] = \sigma_x^2$, the MSE is

$$J = \mathbf{c}^T\mathbf{M}_v\mathbf{c}^* - 2\mathrm{Re}\{\mathbf{v}_x^H\mathbf{c}^*\} + \sigma_x^2. \tag{7.89}$$

The tap vector $\mathbf{c}$ that minimizes the MSE can obtained by equating the gradient $\nabla_{\mathbf{c}} J$ to zero. It can be shown that (Problem 7.15)

$$\nabla_{\mathbf{c}} J = \left(\frac{\partial J}{\partial c_0}, \cdots, \frac{\partial J}{\partial c_{N-1}} \right) = 2\mathbf{c}^T \mathbf{M}_v - 2\mathbf{v}_x^H. \tag{7.90}$$

Setting $\nabla_{\mathbf{c}} J = 0$ gives the MMSE tap solution

$$\mathbf{c}_{\text{op}} = (\mathbf{M}_v^T)^{-1} \mathbf{v}_x^*. \tag{7.91}$$

By substituting (7.91) into (7.89), using the identity $(\mathbf{A}^{-1})^T = (\mathbf{A}^T)^{-1}$ and the fact that $\mathbf{M}_v$ is Hermitian, the MMSE can be expressed as

$$\begin{aligned} J_{\min} &= \mathbf{c}_{\text{op}}^T \mathbf{M}_v \mathbf{c}_{\text{op}}^* - 2\text{Re}\{\mathbf{v}_x^H \mathbf{c}_{\text{op}}^*\} + \sigma_x^2 \\ &= \sigma_x^2 - \mathbf{v}_x^H \mathbf{M}_v^{-1} \mathbf{v}_x. \end{aligned} \tag{7.92}$$

To proceed further, the ith component of the vector $\mathbf{v}_x^H$ is

$$\frac{1}{2}\text{E}[x_{n-d} v_{n-i}^*] = \sigma_x^2 \sum_{\ell=0}^{L} g_\ell^* \delta_{d-i-\ell} = \sigma_x^2 g_{d-i}^*, \quad i = 0, \ldots, N-1, \tag{7.93}$$

so that

$$\mathbf{v}_x^H = \sigma_x^2 (g_d^*, g_{d-1}^*, \ldots, g_0^*, 0, \ldots, 0)^{\text{T}}. \tag{7.94}$$

Also,

$$\frac{1}{2}\text{E}[v_{k-i} v_{k-j}^*] = \begin{cases} \sigma_x^2 f_{j-i} + N_{\text{o}} \delta_{ij}, & |i-j| \le L \\ 0, & \text{otherwise} \end{cases}, \tag{7.95}$$

where we have used the property

$$f_n = \sum_{k=0}^{L-n} g_k^* g_{k+n}, \quad n = 0, \ldots, L. \tag{7.96}$$

Hence, the $N \times N$ matrix $\mathbf{M}_v$ has the form

$$\mathbf{M}_v = \sigma_x^2 \begin{bmatrix} f_0 + N_{\text{o}}/\sigma_x^2 & f_1 & f_2 & \cdots & f_{N-1} \\ f_1^* & f_0 + N_{\text{o}}/\sigma_x^2 & f_1 & \cdots & f_{N-2} \\ f_2^* & f_1^* & f_0 + N_{\text{o}}/\sigma_x^2 & \cdots & f_{N-3} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ f_{N-1}^* & \cdots & f_2^* & f_1^* & f_0 + N_{\text{o}}/\sigma_x^2 \end{bmatrix}. \tag{7.97}$$

Example 7.3:

Consider a system having the same channel vector $\mathbf{g}$ as in Example 7.2. Suppose that we wish to design a three-tap MMSE equalizer. In this case, $g_{d_1} = 0$ and $N = 2d_2 + 1 = 3$, so that $d = d_1 + d_2 = 1$. Hence,

$$\mathbf{v}_x^H = \sigma_x^2(g_1^*, g_0^*, 0) = \sigma_x^2(-0.15,\ 0.90,\ 0.00)$$

and

$$\mathbf{M}_v = \sigma_x^2 \begin{bmatrix} \beta & -0.1500 & 0.1550 \\ -0.1500 & \beta & -0.1500 \\ 0.1550 & -0.1500 & \beta \end{bmatrix},$$

where $\beta = 0.8850 + N_o/\sigma_x^2$. The inverse of $\mathbf{M}_v$ is

$$\mathbf{M}_v^{-1} = \frac{\text{adj}(\mathbf{M}_v)}{\det(\mathbf{M}_v)},$$

where $\det(\mathbf{M}_v) = (\sigma_x^2)^3(\beta(\beta^2 - 0.069025) + 0.006975)$ and

$$\text{adj}(\mathbf{M}_v) = (\sigma_x^2)^2 \begin{bmatrix} \beta^2 - 0.0225 & 0.15\beta - 0.02325 & 0.0225 - 0.155\beta \\ 0.15\beta - 0.02325 & \beta^2 - 0.024025 & 0.15\beta - 0.02325 \\ 0.0225 - 0.155\beta & 0.15\beta - 0.02325 & \beta^2 - 0.0225 \end{bmatrix}.$$

Hence,

$$\mathbf{c}_{\text{op}} = \frac{(\sigma_x^2)^3}{\det(\mathbf{M}_v)} \begin{pmatrix} -0.15\beta^2 + 0.135\beta - 0.1755 \\ 0.90\beta^2 - 0.0225\beta - 0.018135 \\ 0.15825\beta - 0.0243 \end{pmatrix}.$$

With this tap solution,

$$J_{\min} = \sigma_x^2 \left(1 - \frac{0.8325\beta^2 - 0.013689}{\beta(\beta^2 - 0.069025) + 0.006975}\right)$$

and as $N_o \to 0$, $J_{\min} = 0.001089424\sigma_x^2$.

Adaptive Solution

In practice, the channel impulse response is unknown beforehand so that the MMSE solution cannot be obtained by the matrix inversion in (7.91). However, the equalizer taps can be obtained using the stochastic gradient algorithm

$$c_j^{n+1} = c_j^n + \alpha\varepsilon_n v_{n-j}^*, \quad j = 0, \ldots, N-1, \tag{7.98}$$

where ε_n is given by (7.68). To show that (7.98) leads to the desired solution, note from (7.90) that

$$\begin{aligned}\nabla_{\mathbf{c}} J &= \mathrm{E}[\mathbf{c}^T \mathbf{v}_n \mathbf{v}_n^H - x_{n-d}\mathbf{v}_n^H] \\ &= \mathrm{E}[(\mathbf{c}^T \mathbf{v}_n - x_{n-d})\mathbf{v}_n^H] \\ &= \mathrm{E}[\varepsilon_n \mathbf{v}_n^H] = 0. \end{aligned} \tag{7.99}$$

It follows that

$$\frac{1}{2}\mathrm{E}[\varepsilon_n v_{n-j}^*] = 0, \qquad j = 0, \ldots, N-1, \tag{7.100}$$

and, therefore, the adaptive solution tends to force the cross-correlations $\varepsilon_n v_{n-j}^*$, $j = 0, \ldots, N-1$ to zero.

Performance of the MMSE Equalizer

The performance of an MMSE equalizer having an infinite number of taps provides some useful insight. In this case

$$\begin{aligned}\mathbf{c} &= (c_{-\infty}, \ldots, c_0, \ldots, c_\infty), \\ \mathbf{v}_n &= (v_{n+\infty}, \ldots, v_{\mathrm{n}}, \ldots, v_{n-\infty}).\end{aligned}$$

Since the decision delay d with an infinite-tap equalizer is irrelevant, we can choose $d = 0$ so that

$$\frac{1}{2}\mathrm{E}[x_{\mathrm{n}} v_{n-j}^*] = \begin{cases} \sigma_x^2 g_{-j}^*, & -L \le j \le 0 \\ 0, & \text{otherwise} \end{cases}. \tag{7.101}$$

The equation for the optimal tap gain vector $\mathbf{c}^T \mathbf{M}_v = \mathbf{v}_x^H$ can be written in the form

$$\sum_{i=-\infty}^{\infty} c_i (f_{j-i} + N_{\mathrm{o}}\delta_{ij}) = g_{-j}^*, \qquad -\infty < j < \infty. \tag{7.102}$$

Taking the z-transform of both sides of (7.102) gives

$$C(z)\Big(G(z)G^*(1/z^*) + N_{\mathrm{o}}\Big) = G^*(1/z^*) \tag{7.103}$$

and, therefore,

$$C(z) = \frac{G^*(1/z^*)}{G(z)G^*(1/z^*) + N_{\mathrm{o}}}. \tag{7.104}$$

The equivalent MMSE equalizer $C'(z) = C(z)W(z)$ that includes the noise-whitening filter in (7.74) is

$$C'(z) = \frac{\sqrt{2E_{h_r}}}{G(z)G^*(1/z^*) + N_o} = \frac{\sqrt{2E_{h_r}}}{F(z) + N_o}. \tag{7.105}$$

Notice that $C'(z)$ has the same form as the ZF equalizer in (7.75), except for the noise term N_o in the denominator. Clearly, the ZF and MMSE criterion lead to the same solution in the absence of noise.

The most meaningful measure of performance is the bit error probability. However, for many equalization techniques, the bit error probability is a highly nonlinear function of the equalizer coefficients. Another measure of performance is the MSE. The MMSE of an infinite-length MMSE equalizer is given by [217]

$$J_{\min} = \sigma_x^2 T \int_{-1/2T}^{1/2T} \frac{N_o}{F_\Sigma(f) + N_o} \mathrm{d}f, \tag{7.106}$$

where $\sigma_x^2 = \frac{1}{2}\mathrm{E}[|x_k|^2]$. Note that $0 \le J_{\min} \le \sigma_x^2$, and that $J_{\min} = 0$ when there is no noise and $J_{\min} = \sigma_x^2$ when the folded spectrum $F_\Sigma(f)$ exhibits a spectral null. Furthermore, the relationship between the signal-to-noise ratio at the equalizer output and $J_{\min}$ is

$$\gamma_\infty = \sigma_x^2 \cdot \frac{\sigma_x^2 - J_{\min}}{J_{\min}}. \tag{7.107}$$

When there is no ISI, $F_\Sigma(f) = f_0 = 2E_{h_r}$, we have

$$J_{\min} = \frac{\sigma_x^2 N_o}{2E_{h_r} + N_o} \tag{7.108}$$

and the equalizer reaches its maximum output signal-to-noise ratio $\gamma_\infty = 2\sigma_x^2 E_{h_r}/N_o$. Finally, another useful measure for the effectiveness of linear equalization techniques is the *signal-to-interference-plus-noise ratio* (SINR) defined as

$$\mathrm{SINR} = \frac{2\sigma_x^2 |q_d|^2}{2\sigma_x^2 \sum_{\substack{j=0 \\ j\neq d}}^{N+L-1} |q_j|^2 + N_o \sum_{j=0}^{N-1} |c_j|^2}. \tag{7.109}$$

Although the MMSE equalizer accounts for the effects of noise, satisfactory performance still cannot be achieved for channels with severe ISI or spectral nulls, because of the noise enhancement at the output of the equalizer [93, 217]. Another problem with a linear equalizer is the adaptation of the equalizer during data mode. This problem is especially acute when bandwidth efficient trellis-coded modulation schemes are used with non-iterative receivers. In this case, equalizer-based decisions are unreliable and inferior to those in uncoded systems due to the reduced separation between the points in the enlarged signal constellation. This problem can be partially alleviated using periodic training, where the equalizer taps are allowed to converge in periodic training modes [81].

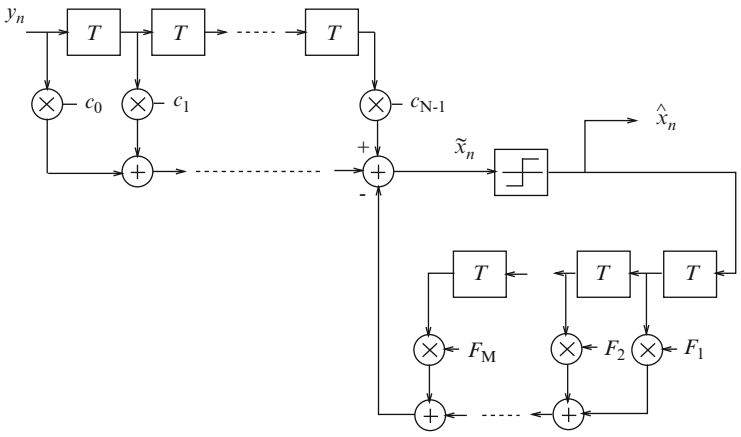

Fig. 7.9 Decision feedback equalizer

7.3.2 Decision Feedback Equalizer

Linear equalizers have the drawback of enhancing channel noise while trying to eliminate ISI, a characteristic known as noise enhancement. As a result, satisfactory performance is unattainable with linear equalizers for channels having severe amplitude distortion. In 1967, Austin [21] proposed the nonlinear decision feedback equalizer (DFE) to mitigate noise enhancement. The DFE consists of two sections; a feedforward section and a feedback section as illustrated in Fig. 7.9. The DFE is nonlinear because the feedback path includes a decision device. The feedforward section has an identical structure to the linear forward equalizer discussed earlier, and its purpose is to reduce the precursor ISI. It has been shown that the optimum tap setting of a zero-forcing DFE having infinite length feedforward and feedback filters is such that the feedforward filter is identical to a noise-whitening filter with system function $1/G^*(1/z^*)$, such that the system function $G(z)$ has minimum phase [220]. Such a filter suppresses the postcursor of the channel response and whitens the noise. The combination of the matched filter, sampler, and feedforward filter yields an equivalent discrete-time white noise channel having the system function $G(z)$.

To eliminate the postcursor ISI, decisions made on the equalizer outputs are propagated through the feedback filter. The optimal coefficients of the feedback filter are the sampled impulse response of the tail of the overall system impulse response that includes the forward part of the DFE. This feedback mechanism introduces error propagation which can degrade the performance of the DFE and complicate its performance analysis.

The output of the DFE shown in Fig. 7.9 is

$$\tilde{x}_n = \sum_{i=0}^{N-1} c_i \tilde{r}_{n-i} - \sum_{i=1}^{M} F_i \hat{x}_{n-i}, \tag{7.110}$$

where $\{c_n\}$ and $\{F_n\}$ are the tap coefficients of the feedforward and feedback filters, respectively, $\{y_n\}$ is the sample sequence at the output of the matched filter, and $\{\hat{x}_n\}$ is the sequence of previously detected data symbols. Recall that the overall channel and feedforward portion of the equalizer can be represented by the sampled impulse response $\mathbf{q}$ in (7.57). Hence, the DFE output can be written as

$$\tilde{x}_n = \sum_{i=0}^{N+L-1} q_i x_{n-i} - \sum_{i=1}^{M} F_i \hat{x}_{n-i} + \tilde{\eta}_n, \tag{7.111}$$

where

$$\tilde{\eta}_n = \sum_{i=0}^{N-1} c_i \tilde{n}_{n-i} \tag{7.112}$$

is the nth noise sample at the output of the feedforward filter. By adding and subtracting terms, the output of the DFE can be rewritten as

$$\begin{aligned}\tilde{x}_n = x_{\mathrm{n}} q_0 + \sum_{i=1}^{M} q_i (x_{n-i} - \tilde{x}_{n-i}) + \sum_{i=1}^{M} (q_i - F_i)\hat{x}_{n-i} \\ + \sum_{i=M+1}^{N+L-1} q_i x_{n-i} + \tilde{\eta}_n.\end{aligned} \tag{7.113}$$

If we choose

$$F_i = q_i = \mathbf{c}^T \mathbf{g}(i), \quad i = 1, 2, \ldots, M \tag{7.114}$$

so that the second summation is zero, and if correct decisions are made so that the first summation is zero, then

$$\tilde{x}_n = x_{\mathrm{n}} q_0 + \sum_{i=M+1}^{N+L-1} q_i x_{n-i} + \tilde{\eta}_n. \tag{7.115}$$

The summation in (7.115) represents the residual ISI that remains from the feedforward filter, which is zero if $M = N + L - 1$.

Equalizer Tap Solution

The coefficients $\{c_i\}$ and $\{F_i\}$ can be adjusted simultaneously to minimize the MSE, resulting in an equalizer that is sometimes called a MMSE-DFE. Define

$$\mathbf{c} = (c_0,\ c_1,\ \ldots,\ c_{N-1})^{\mathrm{T}}, \tag{7.116}$$

$$\tilde{\mathbf{r}}_n = (\tilde{r}_n,\ \tilde{r}_{n-1},\ \ldots,\ \tilde{r}_{n-N})^{\mathrm{T}}, \tag{7.117}$$

$$\hat{\mathbf{x}}_n = (\hat{x}_{n-1},\ \hat{x}_{n-2},\ \ldots,\ \hat{x}_{n-M})^{\mathrm{T}}, \tag{7.118}$$

$$\mathbf{F} = (F_1,\ F_2,\ \ldots,\ F_M)^{\mathrm{T}} \tag{7.119}$$

and define the error at the nth epoch, ε_n, as

$$\begin{aligned}\varepsilon_n &= \tilde{x}_n - x_{\mathrm{n}} \\ &= \mathbf{c}^T\tilde{\mathbf{r}}_n - \mathbf{F}^T\hat{\mathbf{x}}_n - x_{\mathrm{n}}.\end{aligned} \tag{7.120}$$

Now define

$$\tilde{\mathbf{t}} \stackrel{\triangle}{=} \mathrm{vec}(\mathbf{c}^T, \mathbf{F}^T)^{\mathrm{T}}, \tag{7.121}$$

$$\tilde{\mathbf{y}}_n \stackrel{\triangle}{=} \mathrm{vec}(\tilde{\mathbf{r}}_n^T, -\hat{\mathbf{x}}_n^T)^{\mathrm{T}} \tag{7.122}$$

so that $\varepsilon_n = \tilde{\mathbf{t}}^T\tilde{\mathbf{y}}_n - x_{\mathrm{n}}$. Then the MSE can be expressed as

$$\begin{aligned}J &= \frac{1}{2}\mathrm{E}[|\varepsilon_n|^2] \\ &= \frac{1}{2}\mathrm{E}\left[\tilde{\mathbf{t}}^T\tilde{\mathbf{y}}_n\tilde{\mathbf{y}}_n^H\tilde{\mathbf{t}}^* - 2\mathrm{Re}\{\tilde{\mathbf{y}}_n^H\tilde{\mathbf{t}}^* x_{\mathrm{n}}\} + |x_{\mathrm{n}}|^2\right].\end{aligned} \tag{7.123}$$

Notice that (7.123) and (7.87) have the exact same form. Therefore, the MMSE tap solution can be obtained by defining

$$\tilde{\mathbf{M}}_y \stackrel{\triangle}{=} \frac{1}{2}\mathrm{E}[\tilde{\mathbf{y}}_n\tilde{\mathbf{y}}_n^H], \tag{7.124}$$

$$\tilde{\mathbf{y}}_x^H \stackrel{\triangle}{=} \frac{1}{2}\mathrm{E}[\tilde{\mathbf{y}}_n^H x_{\mathrm{n}}]. \tag{7.125}$$

Using the same argument that lead to (7.91), we have the MMSE-DFE tap solution

$$\tilde{\mathbf{t}}_{\mathrm{op}} = (\tilde{\mathbf{M}}_y^T)^{-1}\tilde{\mathbf{y}}_x^*. \tag{7.126}$$

Adaptive Solution

The feedforward taps of the DFE can be adjusted using

$$c_j^{n+1} = c_j^n + \alpha\varepsilon_n v_{n+j}^*, \quad j = 0, \ldots, N-1, \tag{7.127}$$

while the feedback coefficients can be adjusted according to

$$F_j^{n+1} = F_j^n + \alpha\varepsilon_n\hat{x}_{n-j}^*, \quad j = 1, \ldots, M, \tag{7.128}$$

where α is the step size. To see that this leads to the desired solution, we use the same argument that lead to (7.99). Then observe that $\nabla_{\mathbf{c}}J = \mathrm{E}[\varepsilon_n\tilde{\mathbf{y}}_n^H] = \mathbf{0}$ implies that

$$\frac{1}{2}\mathrm{E}[\varepsilon_n v^*_{n+j}] = 0, \quad j = 0, \ldots, N-1, \tag{7.129}$$

$$\frac{1}{2}\mathrm{E}[\varepsilon_n \hat{x}^*_{n-j}] = 0, \quad j = 1, \ldots, M, \tag{7.130}$$

where the second expectation is zero under the assumption that the DFE makes correct decision so that $\hat{x}^*_{n-j} = x^*_{n-j}$.

Performance of the DFE

The performance of a DFE is complicated by the fact that incorrect decisions in the feedback portion of the equalizer result in error propagation. Since the feedback section of the DFE eliminates the postcursor residual ISI at the output of the forward filter, it is apparent that the optimum setting for an infinite-length forward filter is identical to a stable, noncausal, noise-whitening filter that results in an overall system response $G(z)$ having minimum phase [220]. The MMSE for the infinite length DFE is [227]

$$J_{\min} = \sigma_x^2 \exp\left\{T \int_{-1/2T}^{1/2T} \ln\left(\frac{N_0}{F_\Sigma(f) + N_0}\right) \mathrm{d}f\right\}, \tag{7.131}$$

where $0 \le J_{\min} \le \sigma_x^2$. The corresponding signal-to-noise ratio at the output of the DFE is

$$\gamma_\infty = \sigma_x^2 \cdot \frac{\sigma_x^2 - J_{\min}}{J_{\min}}. \tag{7.132}$$

Once again, when there is no ISI $F_\Sigma(f) = f_0 = 2E_{h_r}$ and

$$J_{\min} = \frac{\sigma_x^2 N_0}{2E_{h_r} + N_0}, \tag{7.133}$$

and the equalizer reaches its maximum output signal-to-noise ratio $\gamma_\infty = 2\sigma_x^2 E_{h_r}/N_0$.

7.4 Sequence Estimation

7.4.1 Maximum Likelihood Sequence Estimation

The Viterbi algorithm was originally devised by Viterbi for maximum likelihood decoding of convolutional codes [266,267]. Forney recognized the analogy between an ISI channel and a convolutional encoder, and applied the Viterbi algorithm for the detection of digital signals corrupted by ISI and AWGN [104]. Because of

the efficiency of the Viterbi algorithm, the implementation of the optimum MLSE optimum for detecting ISI-corrupted signals is feasible.

Recall that the overall discrete-time white noise channel with D-branch diversity reception can be modeled by collection of D transversal filters that are T-spaced and have $(L+1)$-taps, as shown in Fig. 7.5. From Fig. 7.5, it can be seen that the channel has a finite number of states defined by contents of the L memory elements in the tapped delay lines. If the size of the signal constellation is 2^n, there are total of $N_S = (2^n)^L$ states. The state at epoch k is

$$\rho_k = (x_{k-1}, x_{k-2}, \cdots, x_{k-L}). \tag{7.134}$$

Example 7.4:

Suppose that the binary sequence $\mathbf{x}$, $x_n \in \{-1, +1\}$, is transmitted over a three-tap static ISI channel with channel vector $\mathbf{g} = (1,\ 1,\ 1)$. In this case, there are four states ($N_S = 4$) and the system can be described the state diagram shown in Fig. 7.10. Note that there are two branches entering and leaving each state since binary modulation is used. In general there are $M = 2^n$ such branches for an M-ary modulation alphabet. The dashed lines correspond to an input symbol equal to "-1," while the solid lines correspond to an input symbol equal to "1."

The system state diagram can be used to construct the trellis diagram shown in Fig. 7.11, where the initial zero state is assumed to be $\rho_0^{(0)} = (-1, -1)$. Again, state transitions with a solid line correspond to an input symbol $+1$, while those with a dashed line correspond to an input symbol -1.

Suppose that the data sequence $\mathbf{x} = (-1,\ 1,\ 1,\ -1,\ 1,\ 1,\ -1,\ -1, \ldots)$ is transmitted over the channel $\mathbf{g}$. Then the state sequence follows the shaded path in Fig. 7.11. The noiseless received sequence is $\mathbf{v} = (v_0,\ v_1,\ v_2,\ v_3,\ v_4, \ldots)$, where

$$\begin{aligned} v_n &= g_0 x_n + g_1 x_{n-1} + g_2 x_{n-2} \\ &= x_n + x_{n-1} + x_{n-2}. \end{aligned}$$

Hence, for the data sequence $\mathbf{x} = (-1,\ 1,\ 1,\ -1,\ 1,\ 1,\ -1,\ -1, \ldots)$ the noiseless received sequence is $\mathbf{v} = (-3,\ -1,\ 1,\ 1,\ 1,\ 1,\ 1,\ -1, \ldots)$.

Assume that k symbols have been transmitted over the channel. Let $\mathbf{V}_n = (v_{n,1},\ v_{n,2},\ \ldots,\ v_{n,D})$ denote the vector of signals received on all D diversity branches at epoch n. After receiving the sequence $\{\mathbf{V}_n\}_{n=1}^k$, the ML receiver decides in favor of the sequence $\{x_n\}_{n=1}^k$ that maximizes the likelihood function

$$p(\mathbf{V}_k,\ \cdots,\ \mathbf{V}_1 |\ x_k,\ \cdots,\ x_1) \tag{7.135}$$

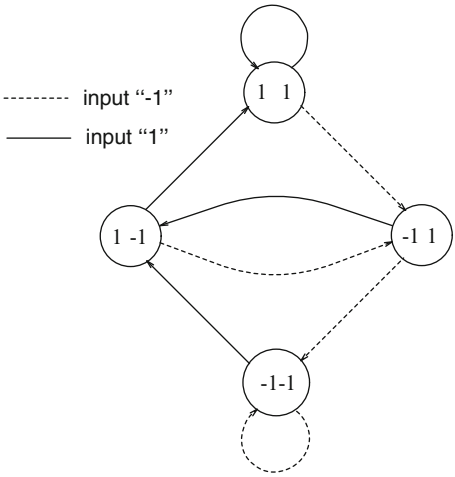

Fig. 7.10 State diagram for binary signaling on a three-tap ISI channel

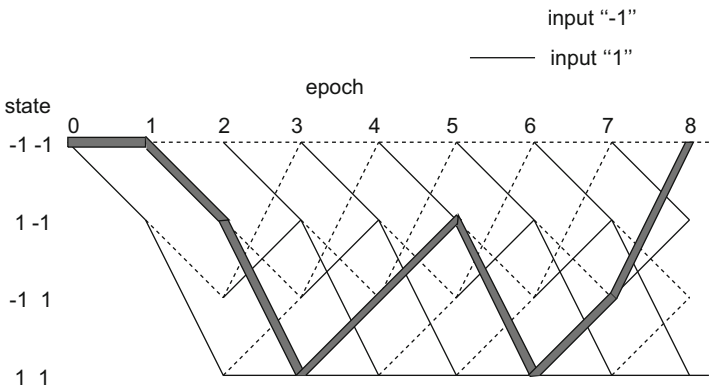

Fig. 7.11 Trellis diagram for binary signaling on a three-tap ISI channel

or, equivalently, the log-likelihood function

$$\log\{p(\mathbf{V}_k, \cdots, \mathbf{V}_1 | \, x_k, \cdots, x_1)\}. \tag{7.136}$$

Since the noise samples $\{w_{n,d}\}$ in (7.32) are mutually independent with respect to the indices n and d, and $\mathbf{V}_n$ depends only on the L most recent transmitted symbols, the log-likelihood function (7.136) can be rewritten as

$$\begin{aligned}&\log\{p(\mathbf{V}_k, \cdots, \mathbf{V}_1 | x_k, \cdots, x_1)\} \\ &\quad = \log\{p(\mathbf{V}_k | x_k, \cdots, x_{k-L})\} + \log\{p(\mathbf{V}_{k-1}, \cdots, \mathbf{V}_1 | x_{k-1}, \cdots, x_1)\}, \end{aligned} \tag{7.137}$$

where $x_{k-L} = 0$ for $k - L \leq 0$. If the second term on the right side of (7.137) has been calculated previously at epoch $k-1$ and stored in memory, then only the first term, called the branch metric, has to be computed for the incoming signal vector $\mathbf{V}_k$ at epoch k.

The model in Fig. 7.5 gives the conditional pdf

$$p(\mathbf{V}_k|x_k, \cdots, x_{k-L}) = \frac{1}{(2\pi N_o)^D} \exp\left\{ -\frac{1}{2N_o} \sum_{d=1}^{D} \left| v_{k,d} - \sum_{i=0}^{L} g_{i,d} x_{k-i} \right|^2 \right\} \quad (7.138)$$

so that $\log p(\mathbf{V}_k|x_k, \cdots, x_{k-L})$ yields the branch metric

$$\mu_k = -\sum_{d=1}^{D} \left| v_{k,d} - \sum_{i=0}^{L} g_{i,d} x_{k-i} \right|^2, \quad (7.139)$$

using the same argument that was used to arrive at (7.37). Note that the receiver requires knowledge of the set of channel vectors $\{\mathbf{g}_d, d = 1, \ldots, D\}$ to compute the branch metric. As discussed later, this can be obtained using a separate channel estimator.

Based on the recursion in (7.137) and the branch metric in (7.139), the well-known Viterbi algorithm [267] can be used to implement the ML receiver by searching through the N_S-state trellis for the most likely transmitted sequence $\mathbf{x} = \{x_k\}$ given the sequence of observation vectors $\mathbf{V} = \{\mathbf{V}_k\}$. This search process is called MLSE. At epoch k, the Viterbi algorithm stores N_S surviving sequences known as *survivors* $\check{\mathbf{x}}(\rho_k^{(i)})$ (paths through the trellis) along with their associated *path metrics* $\Gamma(\rho_k^{(i)})$ (squared Euclidean distances from the received sequence) that terminate at state $\rho_k^{(i)}$, $i = 0, \ldots, N_S - 1$. The path metric is defined as

$$\Gamma(\rho_k^{(i)}) = \sum_{n=1}^{k} \mu_n^{(i)}, \quad i = 0, \ldots, N_S - 1, \quad (7.140)$$

where $\{\mu_n^{(i)}\}$ is the sequence of branch metrics along the surviving path $\check{\mathbf{x}}(\rho_k^{(i)})$. Here, we give an outline of MLSE as implemented by the Viterbi algorithm followed by an example.

The Viterbi algorithm is initialized at time index $k = 0$, by setting all path metrics to zero, that is, $\Gamma(\rho_0^{(i)}) = 0, i = 1, \ldots, N_S - 1$.

1. After the vector $\mathbf{V}_{k+1}$ has been received, compute the set of path metrics $\Gamma(\rho_k^{(i)} \to \rho_{k+1}^{(j)}) = \Gamma(\rho_k^{(i)}) + \mu(\rho_k^{(i)} \to \rho_{k+1}^{(j)})$ for all possible paths through the trellis that terminate in each state $\rho_{k+1}^{(j)}, j = 0, \ldots, N_S - 1$, where $\mu(\rho_k^{(i)} \to \rho_{k+1}^{(j)})$ is the branch metric defined below. For a modulation alphabet of size M, there will be M such paths that terminate in each state $\rho_{k+1}^{(j)}$.

2. Find $\Gamma(\rho_{k+1}^{(j)}) = \max_i \Gamma(\rho_k^{(i)} \rightarrow \rho_{k+1}^{(j)}), j = 0, \cdots, N_S - 1$ where, again, the maximization is over all M possible paths through the trellis that terminate in state $\rho_{k+1}^{(j)}$.
3. Store $\Gamma(\rho_{k+1}^{(j)})$ and its associated surviving sequence $\hat{\mathbf{x}}(\rho_{k+1}^{(j)})$. Drop all other paths that terminate in state $\rho_{k+1}^{(j)}$.
4. Increment the time index k, goto Step 1, and repeat the entire algorithm.

In Step 1 above, $\mu(\rho_k^{(i)} \rightarrow \rho_{k+1}^{(j)})$ is the branch metric associated with the state transition $\rho_k^{(i)} \rightarrow \rho_{k+1}^{(j)}$ and is computed according to the following variation of (7.139)

$$\mu(\rho_k^{(i)} \rightarrow \rho_{k+1}^{(j)}) = -\sum_{d=1}^{D} \left| v_{k,d} - g_{0,d} x_k(\rho_k^{(i)} \rightarrow \rho_{k+1}^{(j)}) - \sum_{m=1}^{L} g_{m,d} x_{k-m}(\rho_k^{(i)}) \right|^2, \tag{7.141}$$

where $x_k(\rho_k^{(i)} \rightarrow \rho_{k+1}^{(j)})$ is a symbol that is uniquely determined by the state transition $\rho_k^{(i)} \rightarrow \rho_{k+1}^{(j)}$, and the L most recent symbols $\{x_{k-m}(\rho_k^{(i)})\}_{m=1}^{L}$ are uniquely specified by the state $\rho_k^{(i)}$.

Example 7.5:

Consider again binary four-state system in Example 7.4. In the presence of noise, the noise received sequence is

$$\begin{aligned} v_n &= g_0 x_n + g_1 x_{n-1} + g_2 x_{n-2} + w_k \\ &= x_n + x_{n-1} + x_{n-2} + w_k, \end{aligned}$$

where the w_k are i.i.d. zero-mean Gaussian random variables with variance N_0. Suppose that due to AWGN, the noisy received sequence is

$$\begin{aligned} \mathbf{v} &= (v_0, v_1, v_2, v_3, v_4, \ldots) \\ &= (-3.2, -1.1, 0.9, 0.1, 1.2, 1.5, 0.7, -1.3, \ldots). \end{aligned}$$

The Viterbi algorithm is initialized with $\Gamma(\rho_0^{(i)}) = 0$ for $i = 0, \ldots, 3$. The initial state is assumed to be $\rho_0^{(0)} = (-1, -1)$. Executing the Viterbi algorithm yields the result shown in Fig. 7.12, where the Xs on the branches in the trellis denote dropped paths (the other path at each state is the survivor), and the numbers in the trellis are the path metrics corresponding to the surviving sequences at each state. The path metric at each state $\rho_k^{(j)}, j = 0, \ldots, 3$, is equal to the

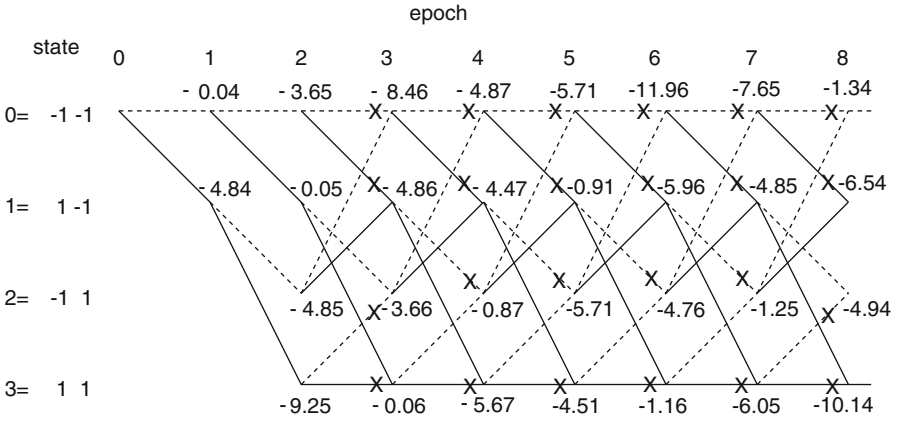

Fig. 7.12 Cumulative path metrics and surviving sequences with the Viterbi algorithm

squared Euclidean distance between the corresponding surviving sequence $\check{\mathbf{x}}(\rho_k^{(j)})$ and the received sequence $\mathbf{v}$ at the output of the noise-whitening filter.

As implied in (7.137), the ML receiver theoretically waits until the entire sequence $\{\mathbf{V}_n\}_{n=1}^{\infty}$ has been received before making a decision. In practice, such a long delay (maybe infinite) is intolerable. One method for solving this problem is to modify the Viterbi algorithm to introduce a fixed decoding delay. Typically, the N_S surviving sequences $\check{\mathbf{x}}(\rho_k^{(i)}), i = 1,\ldots,N_S-1$ at time index k will be identical for bit (or symbol) indices $k-Q$ or less where Q is some sufficiently large number. That is, all the surviving sequences will share a common parent sequence for bit (or symbol) indices $k-Q$ or less. With this in mind, one possibility is to modify the Viterbi algorithm to implement a finite fixed decoding delay, by storing only the Q most recent bits or symbols for each surviving sequence. When the channel vector $\mathbf{V}_k$ is received and the path metrics $\Gamma(\rho_{k+1}^{(j)}), j = 1,\ldots,N_S-1$ are being calculated, the final decision is made on the bit (or symbol) Q branches back in the trellis by deciding in favor of the bit (or symbol) at index $k-Q$ in the surviving sequence $\check{\mathbf{x}}(\rho_k^{(i)})$ having the largest path metric $\Gamma(\rho_k^{(i)})$. It is well known that if $Q > 5L$, the performance degradation caused by the resulting path metric truncation is negligible [267].

Another possibility is to transmit the data in blocks $\mathbf{x} = (x_1, x_2, \ldots, x_N)$ of length N, and use tail symbols to terminate the trellis in a known state. Recall that the state at epoch k is $\rho_k = (x_{k-1}, x_{k-2}, \ldots, x_{k-L})$. Therefore, if each block of N data symbols is appended with a known L-symbol tail sequence, the trellis will terminate in a known state that is uniquely determined by the L-symbol tail sequence. After the received vector $\mathbf{V}_{N+L}$ has been processed, there will be only one surviving sequence left, and this surviving sequence is then used to make a decision on the entire block

of bits (or symbols). This approach suffers a loss in power and data rate by the factor $N/(N+L)$ due to the insertion of tail symbols.

7.4.1.1 Adaptive MLSE Receiver

The Viterbi algorithm requires knowledge of the channel vectors $\mathbf{g}_d, d = 1,\ldots,D$ to compute the branch metrics in (7.139) so that an adaptive channel estimator is needed. Various channel estimators have been proposed in the literature [58,88,170]. Often the LMS algorithm is used for this purpose, because of its good performance, numerical stability, and simplicity of implementation [130, 170]. Another possible adaptation algorithm is the Recursive Least Squares (RLS) or the Kalman algorithm [130]. The RLS algorithm has a very fast convergence rate as compared to the LMS algorithm. However, it is very complicated to implement and it is sensitive to roundoff noise that accumulates due to recursive computations which may cause numerical instabilities in the algorithm [217]. It has also been reported that the tracking properties of the LMS algorithm for the purpose of channel estimation in a fast varying environment are quite similar to those of the RLS algorithm [88, 158, 241]. For these reasons, the LMS algorithm is commonly used during the tracking mode in adaptive MLSE receivers. During the training mode, it is possible that the RLS algorithm could offer better performance than the LMS algorithm.

A straightforward method for adaptive channel estimation with an MLSE receiver is to use the final decisions at the output of the Viterbi algorithm to update the channel estimator during the tracking mode. As mentioned previously, a decision on the data symbol x_{k-Q}, denoted by $\hat{x}_{k-Q}$, is made when the vector $\mathbf{V}_k$ is received and processed. With the LMS algorithm, the tap coefficients are updated according to

$$\hat{g}_{i,d}(k+1) = \hat{g}_{i,d}(k) + \alpha \varepsilon_{k-Q,d} \hat{x}^*_{k-i-Q}, \quad i = 0,\ldots,L \qquad d = 1,\ldots,D, \tag{7.142}$$

where α is the adaptation step size, and

$$\varepsilon_{k-Q,d} = v_{k-Q,d} - \sum_{i=0}^{L} \hat{g}_{i,d}(k) \hat{x}_{k-i-Q}, \tag{7.143}$$

is the error associated with branch d at epoch k. A major problem with this channel estimator is that it lags behind the true channel vector by the decision delay Q that is used in the Viterbi algorithm. To see this, we can write

$$v_{k-Q,d} = \sum_{i=0}^{L} g_{i,d}(k-Q) x_{k-i-Q} + \eta_{k-Q,d}, \tag{7.144}$$

and if correct decisions are made

$$\varepsilon_{k-Q,d} = \sum_{i=0}^{L} \Big(g_{i,d}(k-Q) - \hat{g}_{i,d}(k) \Big) x_{k-i-Q} + \eta_{k-Q,d}. \tag{7.145}$$

Hence, channel time variations over the decision delay Q will cause the terms $\{g_{i,d}(k-Q) - \hat{g}_{i,d}(k)\}_{k=1}^{L}$ to be nonzero, and this will degrade the tracking performance. The decision delay Q could be reduced but this will unfortunately reduce the reliability of the data decisions $\hat{x}_{k-i-Q}$. Since decision errors will also degrade the performance of the channel estimator, the overall performance improvement obtained by reducing Q is minimal at best.

One effective solution to this problem is to use per-survivor processing [157, 221, 237, 238], where *each state* maintains a separate channel estimator to track the channel. After storing the path metrics $\Gamma(\rho_{k+1}^{(j)})$ and the associated surviving sequences $\check{\mathbf{x}}(\rho_{k+1}^{(j)})$ for each state $j = 0, \ldots, N_S - 1$, the channel tap estimates are updated according to

$$\begin{aligned} \hat{g}_{i,d}^{(j)}(k+1) = \hat{g}_{i,d}^{(j)}(k) + \alpha \varepsilon_{k,d}^{(j)} \check{\mathbf{x}}_{k+1-i}^{*}(\rho_{k+1}^{(j)}), \quad & i = 0, \ldots, L \\ & d = 1, \ldots, D \\ & j = 0, \ldots, N_S - 1, \end{aligned} \tag{7.146}$$

where

$$\varepsilon_{k,d}^{(j)} = v_{k,d} - \sum_{i=0}^{L} \hat{g}_{i,d}^{(j)}(k) \check{\mathbf{x}}_{k+1-i}^{*}(\rho_{k+1}^{(j)}) \tag{7.147}$$

and $\check{\mathbf{x}}_{k+1-i}(\rho_{k+1}^{(j)})$ is element $k+1-i$ of the surviving sequence $\check{\mathbf{x}}(\rho_{k+1}^{(j)})$ associated with state $\rho_{k+1}^{(j)}$. Notice that the individual channel estimators for each state use zero-delay symbols in their adaptation algorithm and, therefore, good channel tracking performance is expected. The zero-delay symbols that are used to update the channel tap estimates associated with state $\rho_{k+1}^{(j)}$ are uniquely defined by surviving sequence $\check{\mathbf{x}}(\rho_{k+1}^{(j)})$ that is associated with $\rho_{k+1}^{(j)}$.

7.4.1.2 Fractionally Spaced MLSE Receiver

As mentioned previously, the matched filter outputs are often over-sampled to aid symbol timing synchronization and to mitigate the effects of sample timing errors. Suppose that the matched filter output is sampled at rate $2/T$ and the $T/2$-spaced samples are processed with a $T/2$-spaced noise-whitening filter as shown in Fig. 7.6. Once again, the channel can be modeled as a finite-state machine with the states

defined in (7.134). However, the Viterbi decoder searches for the most likely path in the trellis based on two samples per branch. For each state transition $\rho_k^{(j)} \rightarrow \rho_{k+1}^{(i)}$ at epoch k, the samples $v_{2k}^{(2)}$ and $v_{2k+1}^{(2)}$ are used by the Viterbi algorithm to evaluate the branch metric[2]

$$\mu\left(\rho_k^{(i)} \rightarrow \rho_{k+1}^{(j)}\right) = \left| v_{2k}^{(2)} - g_0^{(2)} x_k\left(\rho_k^{(i)} \rightarrow \rho_{k+1}^{(j)}\right) - \sum_{m=1}^{L} g_{2m}^{(2)} x_{k-m}(\rho_k^{(i)}) \right|^2 + \left| v_{2k+1}^{(2)} - g_1^{(2)} x_k\left(\rho_k^{(i)} \rightarrow \rho_{k+1}^{(j)}\right) - \sum_{m=1}^{L-1} g_{2m+1}^{(2)} x_{k-m}(\rho_k^{(i)}) \right|^2 . \tag{7.148}$$

Other than the change in the branch metric, the Viterbi algorithm proceeds as before. Note also that the adaptive channel estimator must estimate and track two different channel vectors; $\mathbf{g}_e^{(2)} = \{g_0^{(2)}, g_2^{(2)}, \ldots, g_{2L}^{(2)}\}$ and $\mathbf{g}_o^{(2)} = \{g_1^{(2)}, g_3^{(2)}, \ldots, g_{2L-1}^{(2)}\}$.

7.4.2 *Delayed Decision-Feedback Sequence Estimation*

The complexity of the MLSE receiver grows exponentially with the channel memory length. When the channel memory length becomes large, the MLSE receiver quickly becomes impractical. Considerable research has been undertaken to reduce the complexity of MLSE while retaining most of its performance. Duel-Hallen and Heegard [82, 83] proposed DDFSE, a technique that reduces the receiver complexity by truncating the effective channel memory to μ terms, where μ is an integer that can be varied from 0 to L. Thus, a suboptimum receiver is obtained with a complexity that is controlled by the parameter μ.

The system function $G(z)$ of the overall discrete-time white noise channel can be written as

$$G(z) = G_\mu(z) + z^{-(\mu+1)} G^+(z), \tag{7.149}$$

where

$$G_\mu(z) = \sum_{i=0}^{\mu} g_i z^{-i}, \tag{7.150}$$

$$G^+(z) = \sum_{i=0}^{L-\mu-1} g_{i+\mu+1} z^{-i}. \tag{7.151}$$

[2]For notational simplicity we assume $D = 1$.

Let $U(z) = G^+(z)X(z)$, where $X(z)$ is the z-transform of the input sequence. Then

$$u_k = \sum_{i=0}^{L-\mu-1} g_{i+\mu+1} x_{k-i} \tag{7.152}$$

and

$$v_k = \sum_{i=0}^{\mu} g_i x_{k-i} + u_{k-\mu-1} + w_k. \tag{7.153}$$

From (7.152) and (7.153), the system state at epoch k can be decomposed into the state

$$\rho_k^{\mu} = (x_{k-1}, \ldots, x_{k-\mu}) \tag{7.154}$$

and a partial state

$$\kappa_k = (x_{k-\mu-1}, \ldots, x_{k-L}). \tag{7.155}$$

There are $N_\mu = 2^{n\mu}$ states in (7.154).

The DDFSE receiver can be viewed as a Viterbi algorithm with a decision feedback mechanism. For each state transition $\rho_k^{\mu(i)} \to \rho_{k+1}^{\mu(j)}$, the DDFSE receiver uses the branch metric

$$\mu_k\left(\rho_k^{\mu(i)} \to \rho_{k+1}^{\mu(j)}\right) = -\Bigg| v_k - g_0 x_k\left(\rho_k^{\mu(i)} \to \rho_{k+1}^{\mu(j)}\right) - \sum_{\ell=1}^{\mu} g_\ell x_{k-\ell}(\rho_k^{\mu(i)}) - \sum_{\ell=\mu+1}^{L} g_\ell \check{x}_{k-\ell}\left(\rho_k^{\mu(i)}\right)\Bigg|^2, \tag{7.156}$$

where $\check{x}_{k-\ell}(\rho_k^{\mu(i)})$ is element $k-\ell$ of the surviving sequence $\check{\mathbf{x}}(\rho_k^{\mu(i)})$. Since each path uses decision feedback based on its own history, the DDFSE receiver avoids using a single unreliable decision for feedback. Hence, error propagation with a DDFSE receiver is not as severe as with a DFE receiver. When $\mu = 0$, the DDFSE receiver is equivalent to Driscoll's decoder [81], and when $\mu = L$, the DDFSE receiver is equivalent to the MLSE receiver.

Finally, since only the μ most recent symbols are represented by the state in (7.154), it is important to have most of the signal energy contained in these terms. Hence, it is very important that the noise-whitening filter be selected so that the overall channel $G(z)$ has minimum phase. If $G(z)$ does not have minimum phase, DDFSE does not work as well.

Example 7.6:

Consider again the system in Example 7.4, where the received sequence is

$$\begin{aligned}\mathbf{v} &= (v_0, v_1, v_2, v_3, v_4, \ldots) \\ &= (-3.2, -1.1, 0.9, 0.1, 1.2, 1.5, 0.7, -1.3, \ldots)\end{aligned}$$

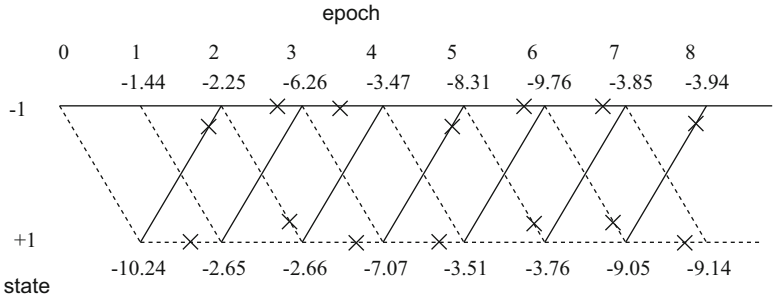

Fig. 7.13 Cumulative path metrics and surviving sequences with DDFSE

Recall that $\rho_k^{(i)} = (x_{k-1},\ x_{k-2})$, so there are four system states. However, we wish to apply DDFSE with the state $\rho_k^{\mu(i)} = x_{k-1},\ i = 0, 1$. The initial state is assumed to be $\rho_0^{\mu(0)} = -1$. Since the channel has finite length, (7.156) gives the branch metric

$$\mu_k\left(\rho_k^{\mu(i)} \to \rho_{k+1}^{\mu(j)}\right) = \left|y_k - x_k\left(\rho_k^{\mu(i)} \to \rho_{k+1}^{\mu(j)}\right) - x_{k-1}(\rho_k^{\mu(i)}) - \check{x}_{k-2}\left(\rho_k^{\mu(i)}\right)\right|^2.$$

Applying DDFSE with the Viterbi algorithm gives the result shown in Fig. 7.13. Once again, the $X's$ on the branches in the trellis denote the dropped paths and the numbers on the trellis nodes denote the path metrics.

7.4.3 Reduced-State Sequence Estimation

For large signal constellations, the number of states with DDFSE, $2^{n\mu}$, is substantial even for small μ. Eyuboğlu and Qureshi [94] proposed a reduced complexity method called RSSE, a technique that is especially useful for systems with large signal constellations. RSSE reduces the number of states using Ungerboeck-like set partitioning principles that were developed for trellis-coded modulation [258]. As described in [94], for each element x_{k-n} in $\rho_k^{\mu} = (x_{k-1}, \ldots, x_{k-\mu})$, a set partitioning $\Omega(n), 1 \le n \le \mu \le L$ is defined where the signal set is partitioned into J_i subsets in a way of increasing intrasubset minimum Euclidean distance.[3]

[3]If $J_1 = J_2 = \cdots = J_\mu = M$ and $\mu < L$, then RSSE becomes DDFSE.

The subset in the partitioning $\Omega(i)$ to which x_{k-i} belongs is denoted by $c_i(x_{k-i})$. The subset partitioning is constrained such that $\Omega(i)$ is a finer partition of $\Omega(i+1)$, $1 \le i \le \mu - 1$ and $J_1 \ge J_2 \cdots \ge J_\mu$. In this case this following subset-state can be defined

$$t_k^\mu = (c_1(x_{k-1}),\, c_2(x_{k-2}),\, \ldots,\, c_\mu(x_{k-\mu})). \tag{7.157}$$

Note that the RSSE subset-state does not completely specify the μ most recent symbols $\{x_{k-i}\}_{i=1}^{\mu}$. Rather, the subset-state only specifies the subset to which these symbols belong.

The constraints on the subset partitioning ensure a properly defined subset-trellis. Given the current subset-state t_k^μ and the subset $c_0(x_k)$ to which the current symbol x_k belongs, the next subset-state t_{k+1}^μ is uniquely determined. Since $c_i(x_{k-i})$ can only assume J_i possible values, there are $\prod_{i=1}^{\mu} J_i$ subset-states which could be much less than $2^{n\mu}$. Note that if $J_1 < 2^n$, there are parallel transitions associated with each subset-transition. The number of the parallel transitions is equal to the number of symbols in the corresponding subset.

The Viterbi algorithm used to search the subset-trellis is the same one used for MLSE, except for a different branch metric and the possibility of parallel transitions associated with the subset-transitions.[4] When there are parallel transitions, the Viterbi algorithm chooses the parallel transition with the maximum branch metric first[5] and then the steps are executed as defined in Sect. 7.4.1.

With RSSE, the branch metric in (7.139) is not uniquely determined by the associated pair of subset-states. This is solved by introducing a decision feedback mechanism for the branch metric calculation [83, 94]. The RSSE branch metric for a particular parallel transition associated with the subset-transition $t_k^{\mu(i)} \to t_{k+1}^{\mu(j)}$ is

$$\mu\left(t_k^{\mu(i)} \to t_{k+1}^{\mu(j)}\right) = -\left| v_k - g_0 x_k\left(t_k^{\mu(i)} \to t_{k+1}^{\mu(j)}\right) - \sum_{\ell=1}^{L} g_\ell \check{x}_{k-\ell}\left(t_k^{\mu(i)}\right)\right|^2, \tag{7.158}$$

where $x_k\left(t_k^{\mu(i)} \to t_{k+1}^{\mu(j)}\right)$ is the data symbol corresponding to the particular parallel transition, and $\check{x}_{k-\ell}(t_k^{\mu(i)})$ is element $k-\ell$ of the surviving sequence $\check{\mathbf{x}}\left(t_k^{\mu(i)}\right)$ associated with the subset-state $t_k^{\mu(i)}$.

[4] With DDFSE there are no parallel transitions.

[5] If the signal constellation has some symmetries, this step can be done using a slicing operation [94].

7.5 Error Probability for MLSE on ISI Channels

We now consider the error probability performance of MLSE on ISI channels. We will see that it is impossible to derive the exact error probability so that bounding techniques, such as the union bound, must be used. Let $\mathbf{x}$ and $\hat{\mathbf{x}}$ be the transmitted and estimated symbol sequences, respectively. For every pair $\mathbf{x}$ and $\hat{\mathbf{x}}$, an error sequence $\varepsilon = \{\varepsilon_i\}$ can be formed by defining $\varepsilon_i = x_i - \hat{x}_i$. We arbitrarily assume that the bit error probability at epoch j_1 is of interest, so that $\varepsilon_{j_1} \neq 0$ for all error sequences that are considered. For each error sequence ε, define the following useful error events:

$\mathscr{E}'(\varepsilon)$: The sequence $\mathbf{x} - \varepsilon$ is the maximum likelihood sequence.
$\mathscr{E}(\varepsilon)$: The sequence $\mathbf{x} - \varepsilon$ has a larger path metric than sequence $\mathbf{x}$.

It is also convenient to define the events

$$\mathscr{E}'_G = \bigcup_{\varepsilon \in G} \mathscr{E}'(\varepsilon) \tag{7.159}$$

and

$$\mathscr{E}_F = \bigcup_{\varepsilon \in F} \mathscr{E}(\varepsilon), \tag{7.160}$$

where G is the set of all possible error sequences having $\varepsilon_{j_1} \neq 0$ and $F \subset G$ is the set of error sequences containing no more than $L-1$ consecutive zeroes amid nonzero elements.

Let $\rho = \{\rho_k\}$ and $\hat{\rho} = \{\hat{\rho}_k\}$ be the system state sequences corresponding to the symbol sequences $\mathbf{x}$ and $\hat{\mathbf{x}}$, respectively. An error event occurs between k_1 and k_2, of length $k_2 - k_1$, if

$$\rho_{k_1} = \hat{\rho}_{k_1},\ \rho_{k_2} = \hat{\rho}_{k_2},\ \text{and } \rho_j \neq \hat{\rho}_j \text{ for } k_1 < j < k_2, \tag{7.161}$$

where $k_1 \leq j_1 \leq k_2$. The symbol error probability at epoch j_1 is

$$\begin{aligned} P_s(j_1) &= \mathrm{P}[x_{j_1} \neq \hat{x}_{j_1}] \\ &= \mathrm{P}[\mathscr{E}'_G] \\ &= \sum_{\varepsilon \in G} \sum_{\mathbf{x} \in \mathscr{X}(\varepsilon)} \mathrm{P}[\mathscr{E}'(\varepsilon)|\mathbf{x}]\mathrm{P}[\mathbf{x}], \end{aligned} \tag{7.162}$$

where $\mathscr{X}(\varepsilon)$ is the set of symbol sequences that can have ε as the error sequence. For different ε, the set $\mathscr{X}(\varepsilon)$ might be different. The third equation in (7.162) is obtained using the property that the events $\mathscr{E}'(\varepsilon)$ are disjoint for $\varepsilon \in G$. Unfortunately, (7.162) does not admit an explicit expression, and hence, upper bounding techniques are needed for the performance evaluation. A union bound on the error probability will be used in our analysis.

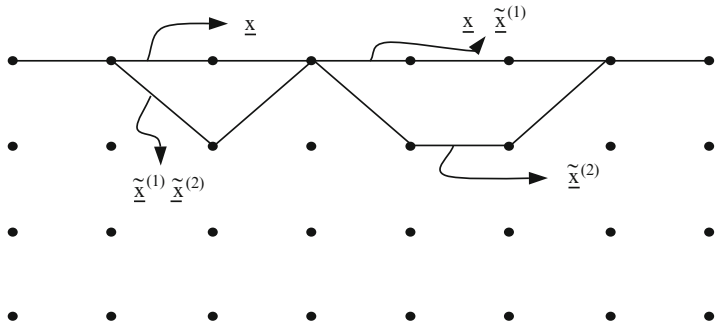

Fig. 7.14 A typical error state trellis diagram

To obtain a tighter union bound, we now prove that the symbol error probability at epoch j_1 is

$$P_s(j_1) = \mathrm{P}[\mathscr{E}_F]. \tag{7.163}$$

Consider the typical trellis diagram as shown in Fig. 7.14, where $\mathbf{x}$ denotes the transmitted symbol sequence, and $\tilde{\mathbf{x}}^{(1)}$ and $\tilde{\mathbf{x}}^{(2)}$ denote two different symbol sequences. It can be seen that the error sequence $\boldsymbol{\varepsilon}^{(1)}$ associated with $\tilde{\mathbf{x}}^{(1)}$ and the error sequence $\boldsymbol{\varepsilon}^{(2)}$ associated with $\tilde{\mathbf{x}}^{(2)}$ belong to sets F and $G \setminus F$, respectively. For every $\boldsymbol{\varepsilon}^{(2)} \in G \setminus F$, there always exists an $\boldsymbol{\varepsilon}^{(1)} \in F$. If the sequence $\mathbf{x} - \boldsymbol{\varepsilon}^{(2)}$ is the ML sequence, that is, the event $\mathscr{E}'_G$ has occurred, then the sequence $\mathbf{x} - \boldsymbol{\varepsilon}^{(1)}$ has a larger path metric than the sequence $\mathbf{x}$, that is, the event $\mathscr{E}_F$ has occurred. This means that $\mathscr{E}'_G$ implies $\mathscr{E}_F$. On the other hand, if $\boldsymbol{\varepsilon}^{(1)} \in F$ and the sequence $\mathbf{x} - \boldsymbol{\varepsilon}^{(1)}$ has a larger path metric than sequence $\mathbf{x}$, then there exists a sequence $\boldsymbol{\varepsilon} \in G$ such that the sequence $\mathbf{x} - \boldsymbol{\varepsilon}$ is the ML sequence. Therefore, $\mathscr{E}_F$ implies $\mathscr{E}'_G$, and (7.163) is proven.

The union bound on (7.163) yields

$$\begin{aligned} P_s(j_1) &\le \sum_{\boldsymbol{\varepsilon} \in F} \mathrm{P}]\mathscr{E}(\boldsymbol{\varepsilon})] \\ &= \sum_{\boldsymbol{\varepsilon} \in F} \sum_{\mathbf{x} \in \mathscr{X}(\boldsymbol{\varepsilon})} \mathrm{P}[\mathscr{E}(\boldsymbol{\varepsilon})|\mathbf{x}]\mathrm{P}[\mathbf{x}] \end{aligned} \tag{7.164}$$

or, equivalently,

$$P_s(j_1) \le \sum_{\boldsymbol{\varepsilon} \in E} w_s(\boldsymbol{\varepsilon}) \sum_{\mathbf{x} \in \mathscr{X}(\boldsymbol{\varepsilon})} \mathrm{P}[\mathscr{E}(\boldsymbol{\varepsilon})|\mathbf{x}]\mathrm{P}[\mathbf{x}], \tag{7.165}$$

where $E \in F$ is the set of error sequences that have the first nonzero element starting at time j_1, and $w_s(\boldsymbol{\varepsilon})$ is the number of symbol errors associated with the error sequence $\boldsymbol{\varepsilon}$. To obtain (7.165), we have used the following observations; (1) there are $w_s(\boldsymbol{\varepsilon})$ places for the error sequence $\boldsymbol{\varepsilon}$ to start such that $\varepsilon_{j_1} \neq 0$, and (2) the

error probability $\mathrm{P}[\mathscr{E}(\boldsymbol{\varepsilon})|\mathbf{x}]$ is independent of the place where the error sequence $\boldsymbol{\varepsilon}$ starts. If the transmitted symbol sequence is long enough, then the symbol error probability $P_s(j_1)$ is independent of the time index j_1 and, therefore, the time index will be omitted hereafter. Finally, for a given transmitted symbol sequence $\mathbf{x}$, the events $\{\mathscr{E}(\boldsymbol{\varepsilon})\}$ for $\boldsymbol{\varepsilon} \in F$ in (7.164) might overlap. The reason is that there may be multiple symbol sequences that simultaneously have a larger path metric than the path metric of the transmitted symbol sequence. When the system is operating at a low SNR, there are more overlapping events $\mathscr{E}(\boldsymbol{\varepsilon})$ and, hence, the union bound (7.164) becomes looser.

From the definition of event $\mathscr{E}(\boldsymbol{\varepsilon})$, the union bound (7.165) becomes

$$P_s \leq \sum_{\boldsymbol{\varepsilon} \in E} w_s(\boldsymbol{\varepsilon}) \sum_{\mathbf{x} \in \mathscr{X}(\boldsymbol{\varepsilon})} \mathrm{P}[\Gamma(\mathbf{x}-\boldsymbol{\varepsilon}) \geq \Gamma(\mathbf{x})|\mathbf{x}]\mathrm{P}[\mathbf{x}], \tag{7.166}$$

where $\Gamma(\mathbf{x})$ is the path metric associated with the input sequence $\mathbf{x}$. To obtain the bit error probability, (7.166) can be easily modified as

$$P_b \leq \frac{1}{n} \sum_{\boldsymbol{\varepsilon} \in E} w_b(\boldsymbol{\varepsilon}) \sum_{\mathbf{x} \in \mathscr{X}(\boldsymbol{\varepsilon})} \mathrm{P}[\Gamma(\mathbf{x}-\boldsymbol{\varepsilon}) \geq \Gamma(\mathbf{x})|\mathbf{x}]\mathrm{P}[\mathbf{x}], \tag{7.167}$$

where n is the number bits transmitted per unit time, and $w_b(\boldsymbol{\varepsilon})$ is the number of bit errors associated with the error sequence $\boldsymbol{\varepsilon}$. The probability

$$\mathrm{P}[\Gamma(\mathbf{x}-\boldsymbol{\varepsilon}) \geq \Gamma(\mathbf{x})|\mathbf{x}] \tag{7.168}$$

is called the pairwise error probability. We will see in the following two sections that the pairwise error probability is independent of the transmitted symbol sequence $\mathbf{x}$. Therefore, the union bounds (7.166) and (7.167) simplify to

$$\mathrm{P}_s \leq \sum_{\boldsymbol{\varepsilon} \in E} w_s(\boldsymbol{\varepsilon}) \mathrm{P}[\Gamma(\mathbf{x}-\boldsymbol{\varepsilon}) \geq \Gamma(\mathbf{x})|\mathscr{X}(\boldsymbol{\varepsilon})]\,\mathrm{P}[\mathscr{X}(\boldsymbol{\varepsilon})] \tag{7.169}$$

and

$$\mathrm{P}_b \leq \frac{1}{n} \sum_{\boldsymbol{\varepsilon} \in E} w_b(\boldsymbol{\varepsilon}) \mathrm{P}[\Gamma(\mathbf{x}-\boldsymbol{\varepsilon}) \geq \Gamma(\mathbf{x})|\mathscr{X}(\boldsymbol{\varepsilon})]\,\mathrm{P}[\mathscr{X}(\boldsymbol{\varepsilon})], \tag{7.170}$$

respectively. The expressions in (7.169) and (7.170) are easier to calculate than those in (7.166) and (7.167), because not all of the symbol sequences need to be considered in the calculation.

7.5.1 Static ISI Channels

The pairwise error probability associated with the error event of length $\ell = k_2 - k_1$ in (7.161) is (Problem 7.16)

$$\mathrm{P}[\Gamma(\mathbf{x}-\boldsymbol{\varepsilon}) \geq \Gamma(\mathbf{x})|\mathbf{x}] = Q\left(\sqrt{\frac{\Delta^2}{4N_{\mathrm{o}}}}\right), \tag{7.171}$$

where

$$\Delta^2 = \sum_{k=k_1}^{k_1+\ell-1} \left| \sum_{i=0}^{L} g_i \varepsilon_{k-i} \right|^2 \tag{7.172}$$

and Δ^2 is the squared Euclidean path distance. At high signal-to-noise ratios, the error event probability is approximately

$$P_e \approx N_{\min} Q\left(\sqrt{\frac{d_{\min}^2}{4N_{\mathrm{o}}}}\right), \tag{7.173}$$

where $d_{\min}^2$ is the minimum value of Δ^2 and $N_{\min}$ denotes the average number of error events at distance $d_{\min}$.

The squared Euclidean path distance in (7.172) can be rewritten as

$$\Delta^2 = \sum_{k=k_1}^{k_1+\ell-1} \Delta_k^2, \tag{7.174}$$

where

$$\Delta_k^2 = \mathbf{g}^H \mathbf{E}_k \mathbf{g} \tag{7.175}$$

is the squared branch distance and

$$\mathbf{E}_k = [(e_{mn})_k] \tag{7.176}$$

is the $(L+1) \times (L+1)$ branch distance matrix with elements $(e_{mn})_k = \varepsilon_{k-m+1}^* \varepsilon_{k-n+1}$. Define the error vector

$$\mathbf{e}_k = (\varepsilon_k, \varepsilon_{k-1}, \ldots, \varepsilon_{k-L})^H. \tag{7.177}$$

It follows that $\mathbf{E}_k = \mathbf{e}_k \mathbf{e}_k^H$ and, hence, $\mathbf{E}_k$ has rank one. Note that $\mathbf{E}_k \mathbf{e}_k = (\mathbf{e}^H \mathbf{e}_k)\mathbf{e}_k$ and, therefore, $\mathbf{e}_k$ is an eigenvector of $\mathbf{E}_k$ and the only eigenvalue of $\mathbf{E}_k$ is $\lambda(k) = \sum_{i=0}^{L} |\varepsilon_{k-i}|^2$. The path distance matrix of the length ℓ error event in (7.161) is defined as

$$\mathbf{E} \stackrel{\triangle}{=} \sum_{k=k_1}^{k_1+\ell-1} \mathbf{E}_k. \tag{7.178}$$

Using (7.134) and (7.161), the elements of $\mathbf{E}$ are

$$e_{mn} = r_\ell(n-m), \tag{7.179}$$

where

$$r_\ell(i) = \begin{cases} \sum_{k=k_1}^{k_1+\ell-(L+1)-i} \varepsilon_k \varepsilon_{k+i}^* & i \geq 0 \\ r_\ell^*(-i) & i < 0 \end{cases}. \tag{7.180}$$

It follows that (7.174) has the Hermitian form $\Delta^2 = \mathbf{g}^H \mathbf{E} \mathbf{g}$. Since $\Delta^2 > 0$, $\mathbf{E}$ is a positive definite matrix with all eigenvalues being real and positive. The matrix $\mathbf{E}$ depends on the signal constellation and the length of the channel $L+1$.

By noting that $\mathbf{g}^H \mathbf{g} = f_0 = 2E_{h_\mathrm{r}}$, the squared Euclidean path distance can be expressed in the form

$$\Delta^2 = 2E_{h_\mathrm{r}} \frac{\mathbf{g}^H \mathbf{E}\, \mathbf{g}}{\mathbf{g}^H \mathbf{g}} = 2E_{h_\mathrm{r}} R(\mathbf{g}), \tag{7.181}$$

where the ratio of the Hermitian form $\mathbf{g}^H \mathbf{E}\, \mathbf{g}$ to the inner product $\mathbf{g}^H \mathbf{g}$ is called the Rayleigh quotient of the vector $\mathbf{g}$ and is denoted as $R(\mathbf{g})$ [130]. The eigenvalues of $\mathbf{E}$ are equal to the Rayleigh quotient of the corresponding eigenvectors. The Rayleigh quotient of $\mathbf{E}$ satisfies $\lambda_{\min} \leq R(\mathbf{g}) \leq \lambda_{\max}$. The minimum value of $R(\mathbf{g})$ corresponds to the smallest possible Δ^2 (or largest pairwise error probability) for a given error sequence $\boldsymbol{\varepsilon}$, and occurs when $\mathbf{g} = \mathbf{v}_{\min}$. Likewise, the maximum value of $R(\mathbf{g})$ corresponds to the largest possible Δ^2 (or smallest pairwise error probability) for a given error sequence $\boldsymbol{\varepsilon}$, and occurs when $\mathbf{g} = \mathbf{v}_{\max}$. Note that the eigenvalues of $\mathbf{E}$ are only a function of the error sequence $\boldsymbol{\varepsilon}$ and do not depend on the channel vector $\mathbf{g}$. While the eigenvalues of $\mathbf{E}$ can be calculated exactly for a given error sequence $\boldsymbol{\varepsilon}$, the eigenvalues of $\mathbf{E}$ are bounded by [130]

$$\lambda_{\max} \leq \sum_{i=0}^{L} |r_\ell(i)| \quad \text{and} \quad \lambda_{\min} \geq r_\ell(0) - \sum_{i=1}^{L} |r_\ell(i)|. \tag{7.182}$$

The upper and lower bounds on the eigenvalues are easier to calculate than the eigenvalues themselves. Finally, the condition number of $\mathbf{E}$ is defined as $c(\mathbf{E}) \triangleq \lambda_{\max}/\lambda_{\min}$, and $c(\mathbf{E}) \geq 1$. If the condition number of a particular error sequence $\boldsymbol{\varepsilon}$ is large (small), then the corresponding pairwise error probability will have a large (small) variation with the channel vector $\mathbf{g}$.

7.5.2 *Fading ISI Channels*

Consider the case of a fading ISI channel with D-branch diversity reception and maximal ratio combining. The pairwise error probability is still given by (7.171), but the squared Euclidean path distance associated with an error event of length ℓ is [240]

$$\Delta^2 = \sum_{d=1}^{D} \Delta_d^2, \tag{7.183}$$

where

$$\Delta_d^2 = \sum_{k=k_1}^{k_1+\ell-1} \left| \sum_{i=0}^{L} g_{i,d}(k)\varepsilon_{k-i} \right|^2. \tag{7.184}$$

The above expression can be written in the form

$$\Delta_d^2 = \sum_{k=k_1}^{k_1+\ell-1} \mathbf{g}_d^H(k)\mathbf{E}_k\mathbf{g}_d(k). \tag{7.185}$$

In general, the correlation matrix $\mathbf{\Phi}_{\mathbf{g}_d}(0)$, where $\mathbf{\Phi}_{\mathbf{g}_d}(m)$ is defined in (7.43) is not diagonal. Although a nondiagonal $\mathbf{\Phi}_{\mathbf{g}_d}(0)$ matrix is more realistic, it leads to considerable analytical difficulty and loss of insight. Therefore, we restrict our attention to the case where $\mathbf{\Phi}_{\mathbf{g}_d}(0)$ is a diagonal matrix, meaning that the channel tap gains are all mutually uncorrelated (assuming they have zero mean). In such a case, a normalized channel vector $\hat{\mathbf{g}}_d(k)$ can be defined such that $\mathbf{\Phi}_{\hat{\mathbf{g}}_d}(0) = \mathbf{I}_{L+1}$. As a result, (7.185) can be rewritten as

$$\Delta_d^2 = \sum_{k=k_1}^{k_1+\ell-1} \hat{\mathbf{g}}_d^H(k)\mathbf{A}_{k,d}\hat{\mathbf{g}}_d(k), \tag{7.186}$$

where

$$\mathbf{A}_{k,d} = \mathbf{\Sigma}_d\mathbf{E}_k\mathbf{\Sigma}_d \tag{7.187}$$

and

$$\mathbf{\Sigma}_d = \text{diag}[\sigma_{0,d},\ \sigma_{1,d},\ \ldots,\ \sigma_{L,d}] \tag{7.188}$$

with $\sigma_{i,d}^2 = \frac{1}{2}\text{E}[|g_{i,d}|^2]$. It follows that $\mathbf{A}_{k,d} = \mathbf{u}_{k,d}\mathbf{u}_{k,d}^H$ where $\mathbf{u}_{k,d} = \mathbf{\Sigma}_d\mathbf{e}_k$ and, hence, $\mathbf{A}_{k,d}$ is a rank one matrix and $\mathbf{u}_{k,d}$ is an eigenvector of $\mathbf{A}_{k,d}$. The only nonzero eigenvalue of $\mathbf{A}_{k,d}$ is $\lambda_d = \sum_{i=0}^{L}\sigma_{i,d}^2|\varepsilon_{k-i}|^2$.

For slowly time-variant channels, it is reasonable to assume that $\mathbf{g}_d(k)$ remains constant over the length of the dominant error events, that is, $\mathbf{g}_d(k) \equiv \mathbf{g}_d$. This assumption holds even for relatively large Doppler frequencies and error event lengths. For example, if the channel exhibits 2D isotropic scattering and $f_\text{m}T = 0.0025$, then error events up to length 20 have $J_0(2\pi f_\text{m}|k|T) \geq J_0(2\pi f_\text{m}20T) = 0.9755 \approx 1$. Using the above assumption, (7.186) can be written as

$$\Delta_d^2 = \hat{\mathbf{g}}_d^H\mathbf{A}_d\hat{\mathbf{g}}_d, \tag{7.189}$$

where

$$\begin{aligned}\mathbf{A}_d &= \sum_{k=k_1}^{k_1+\ell-1} \mathbf{A}_{k,d} \\ &= \mathbf{\Sigma}_d\mathbf{E}\mathbf{\Sigma}_d.\end{aligned} \tag{7.190}$$

The matrix $\mathbf{A}_d$ is also positive definite with all its eigenvalues real and positive. The elements of $\mathbf{A}_d$ are given by $[(a_{mn})]_d = \sigma_{m-1,d}\ \sigma_{n-1,d}\ r_\ell(n-m)$ where $r_\ell(i)$ is given by (7.180). The trace of the matrix $\mathbf{A}_d$ is

$$\mathrm{tr}(\mathbf{A}_d) = \sum_{i=0}^{L} \lambda_{i,d} = (\bar{E}/\sigma_x^2) r_\ell(0), \tag{7.191}$$

where the $\lambda_{i,d},\ i = 0,\ldots,L$ are the eigenvalues of $\mathbf{A}_d$. The last equality in (7.191) is obtained using (7.45). Since $\mathbf{A}_d$ is Hermitian, there exists a diagonalization $\mathbf{A}_d = \mathbf{U}_d \Lambda_d \mathbf{U}_d^H$ such that $\mathbf{U}_d$ is a unitary matrix and Λ_d is a diagonal matrix consisting of the eigenvalues of $\mathbf{A}_d$. Let $\omega_d = \mathbf{U}_d^H \hat{\mathbf{g}}_d$ be the corresponding diagonal transformation. Hence,

$$\Delta_d^2 = \omega_d^H \Lambda_d \omega_d = \sum_{i=0}^{L} \lambda_{i,d} |\omega_{i,d}|^2, \tag{7.192}$$

where $\frac{1}{2}\mathrm{E}[\omega_d \omega_d^H] = \mathbf{I}_{L+1}$ so that the $\{\omega_{i,d}\}$ are independent zero-mean unit-variance Gaussian random variables. Using (7.183) and (7.192) gives

$$\Delta^2 = \sum_{d=1}^{D} \sum_{i=0}^{L} \alpha_{i,d}, \tag{7.193}$$

where $\alpha_{i,d} = \lambda_{i,d} |\omega_{i,d}|^2$. The $\alpha_{i,d}$ are chi-square distributed with two degrees of freedom and, therefore, the characteristic function of Δ^2 is

$$\psi_{\Delta^2}(z) = \prod_{d=1}^{D} \prod_{i=0}^{L} \frac{1}{1 - \overline{\alpha}_{i,d} z}, \tag{7.194}$$

where $\overline{\alpha}_{i,d} = 2\lambda_{i,d}$. Finally, the pairwise error probability is

$$\mathrm{P}[\Gamma(\mathbf{x} - \varepsilon) \geq \Gamma(\mathbf{x}) | \mathbf{x}] = \int_0^\infty Q\left(\sqrt{2x}\right) f_{\Delta^2}(x) \mathrm{d}x, \tag{7.195}$$

where $f_{\Delta^2}(x)$ is the probability density function of Δ^2. Note that if some of the eigenvalues $\lambda_{i,d}$ are the same, then there will be repeated poles in the characteristic function in (7.194). This will be the case for balanced diversity branches and will also be the case if the channel vectors $\mathbf{g}_d$ have some equal strength taps. Consider the case where balanced D-branch diversity is used, but where the channel taps in each diversity branch are not of equal strength. In this case, $\lambda_{i,d} \equiv \lambda_i,\ d = 1,\ldots,D$ and the characteristic function in (7.194) has the form

$$\begin{aligned} \psi_{\Delta^2}(z) &= \prod_{i=0}^{L} \frac{1}{(1 - z\bar{\alpha}_i)^D} \\ &= \sum_{i=0}^{L} \sum_{d=1}^{D} \frac{A_{id}}{(1 - z\bar{\alpha}_i)^d}, \end{aligned} \tag{7.196}$$

where

$$A_{id} = \frac{1}{(D-d)!(-\bar{\alpha}_i)^{D-d}} \left\{ \frac{\mathrm{d}^{D-d}}{\mathrm{d}z^{D-d}} \ (1 - z\bar{\alpha}_i)^D \ \psi_{\Delta^2}(z) \right\}_{z = 1/\bar{\alpha}_i} \tag{7.197}$$

and $\bar{\alpha}_i = 2\lambda_i$. The pdf of Δ^2 is

$$f_{\Delta^2}(x) = \sum_{i=0}^{L} \sum_{d=1}^{D} A_{id} \ \frac{1}{(d-1)!(\bar{\alpha}_i)^d} \ x^{d-1} \ \mathrm{e}^{-x/\bar{\alpha}_i} \ , \ \ x \geq 0. \tag{7.198}$$

From (7.195) and (7.198), the exact pairwise error probability is

$$\mathrm{P}[\Gamma(\mathbf{x} - \boldsymbol{\varepsilon}) \geq \Gamma(\mathbf{x}) | \mathbf{x}] = \sum_{i=0}^{L} \sum_{d=1}^{D} A_{id} \left(\frac{1-\mu_i}{2} \right)^d \sum_{m=0}^{d-1} \binom{d-1+m}{m} \left(\frac{1+\mu_i}{2} \right)^m , \tag{7.199}$$

where

$$\mu_i = \sqrt{\frac{\bar{\alpha}_i}{1+\bar{\alpha}_i}}. \tag{7.200}$$

From (7.191), the $\overline{\alpha}_{i,d}$ have the sum value constraint

$$\sum_{i=0}^{L} \overline{\alpha}_{i,d} = 2 \sum_{i=0}^{L} \lambda_{i,d} = (2\bar{E}/\sigma_x^2) r_\ell(0). \tag{7.201}$$

Define $S \subseteq R^{L+1}$ as the set of all $(L+1)$-component vectors $\{\bar{\alpha}_d : \sum_{i=0}^{L} \bar{\alpha}_{i,d} = (2\bar{E}/\sigma_x^2) r_\ell(0)\}$. The set S is convex, since for any pair of vectors $\bar{\alpha}_{d,k}$ and $\bar{\alpha}_{d,j}$ the convex combination $\theta\bar{\alpha}_{d,k} + (1-\theta)\bar{\alpha}_{d,j}$ is contained in S for any $0 \leq \theta \leq 1$. If the pairwise error probability is treated as a mapping from S to R, then it is a convex function of $\bar{\alpha}_d$ and, hence, has a unique minimum. For example, Fig. 7.15 shows the pairwise error probability for a three-tap channel ($L = 2$, $D = 1$) with equal strength taps ($\bar{\alpha}_0 = \bar{\alpha}_1 = \bar{\alpha}_2$). Note that the value of $\bar{\alpha}_2$ is determined uniquely by the values of $\bar{\alpha}_0$ and $\bar{\alpha}_1$, and that is why a three-dimensional graph is used. Using variational calculus, it can be shown (Appendix 1) that the pairwise error probability is minimized when the $\bar{\alpha}_{i,d}$ are all equal, that is, $\lambda_{i,d} = \lambda = (\bar{E}/\sigma_x^2) r_\ell(0)/(L+1)$, resulting in the minimum pairwise error probability

$$P_{\min} = \left(\frac{1-\mu}{2} \right)^{D(L+1)} \sum_{m=0}^{D(L+1)-1} \binom{D(L+1)-1+m}{m} \left(\frac{1+\mu}{2} \right)^m , \tag{7.202}$$

where

$$\mu = \sqrt{\frac{\lambda/4N_\mathrm{o}}{1+\lambda/4N_\mathrm{o}}}. \tag{7.203}$$

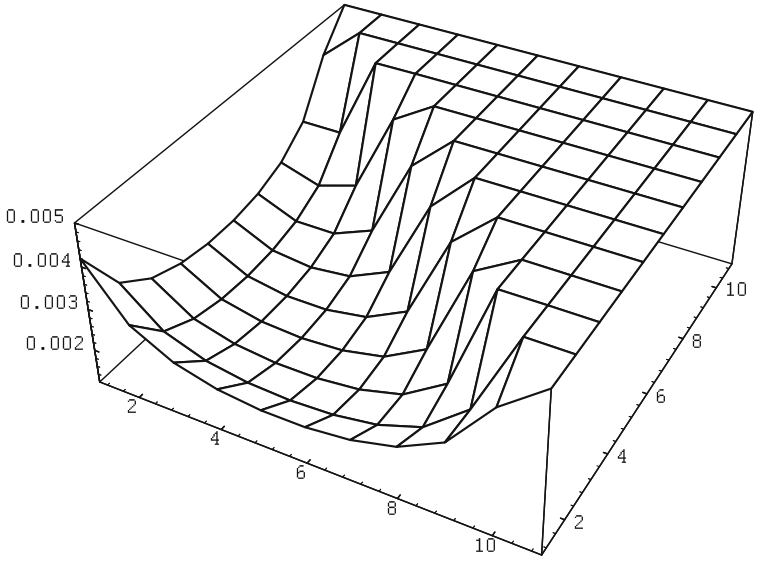

Fig. 7.15 Pairwise error probability for a three-tap fading ISI channel

For a given error event, the pairwise error probability is minimized when $\mathbf{A}_d$ is perfectly conditioned, that is, $c(\mathbf{A}_d) = 1$. Recall $c(\mathbf{A}_d) = c(\mathbf{\Sigma}_d \mathbf{E} \mathbf{\Sigma}_d) \leq (c(\mathbf{\Sigma}_d))^2 c(\mathbf{E})$, where $(c(\mathbf{\Sigma}_d))^2$ represents the ratio of the maximum and minimum channel tap variances $(\sigma_d^2)_{\max}/(\sigma_d^2)_{\min}$ for diversity branch d. We have seen that $\mathbf{E}$ depends only the signal constellation being used and the channel vector length $L+1$. However, $\mathbf{A}_d$ has information about the signal constellation and power distribution of the fading ISI channel. It follows that $c(\mathbf{A}_d) \leq c(\mathbf{E})$ with equality if and only if the channel has equal strength taps. This means that any system has the best performance when the fading ISI channel has equal strength taps.

7.6 Error Probability for $T/2$-Spaced MLSE Receiver

Referring to Fig. 7.6, let $X(z)$, $V(z)$, and $V^{(2)}(z)$ be the z-transforms of the input sequence $\mathbf{x}$, the T-spaced received sequence $\mathbf{v}$ and the $T/2$-spaced received sequence $\mathbf{v}^{(2)}$, respectively. The mappings from $X(z)$ to $V(z)$ and from $X(z)$ to $V^{(2)}(z)$ are one-to-one, and both the T-spaced and $T/2$-spaced MLSE receivers operate on received sequences that are corrupted by noise samples with variance N_o. Therefore, we only need to compare the Euclidean distances between allowed sequences of channel outputs to determine the relative performance of the T- and $T/2$-spaced receivers.

7.6.1 T-Spaced MLSE Receiver

From the definition of the error event in (7.161), the z-transform of the length-ℓ error sequence $\boldsymbol{\varepsilon} = \{\varepsilon_{k_1}, \varepsilon_{k_1+1}, \ldots, \varepsilon_{k_1+\ell-1}\}$ is

$$\mathscr{E}(z) = \varepsilon_{k_1} + \varepsilon_{k_1+1} z^{-1} + \cdots + \varepsilon_{k_1+\ell-1} z^{-\ell+1}, \tag{7.204}$$

where $\varepsilon_k = x_k - \hat{x}_k$. The z-transform of the received signal error sequence associated with the length-ℓ error sequence $\boldsymbol{\varepsilon}$ is

$$\begin{aligned} \mathscr{E}_v(z) = {} & (v_{k_1} - \hat{v}_{k_1}) + (v_{k_1+1} - \hat{v}_{k_1+1}) z^{-1} + \ldots \\ & \cdots + (v_{k_1+\ell+L-1} - \hat{v}_{k_1+\ell+L-1}) z^{-(\ell+L-1)} \end{aligned} \tag{7.205}$$

and we have

$$\mathscr{E}_v(z) = \mathscr{E}(z) G(z). \tag{7.206}$$

From (7.172), the squared Euclidean distance Δ^2 of the error event in (7.161) is [104]

$$\begin{aligned} \Delta^2 &= \sum_{k=k_1}^{k_1+\ell-1} \left| \sum_{i=0}^{L} g_i \varepsilon_{k-i} \right|^2 \\ &= [\, \mathscr{E}_v(z) \mathscr{E}_v^*(1/z^*) \,]_0 \\ &= [\, \mathscr{E}(z) F(z) \mathscr{E}^*(1/z^*) \,]_0, \end{aligned} \tag{7.207}$$

where $[\,\cdot\,]_0$ is the coefficient of z^0.

7.6.2 T/2-Spaced MLSE Receiver

For the same error event described in (7.161), the corresponding $T/2$-spaced error sequence is $\boldsymbol{\varepsilon}^{(2)} = \{\varepsilon_{k_1}^{(2)}, \varepsilon_{k_1+1}^{(2)}, \ldots, \varepsilon_{k_1+2\ell-1}^{(2)}\}$, where

$$\varepsilon_{k_1+i}^{(2)} = \begin{cases} \varepsilon_{k_1+i/2}, & i \text{ even} \\ 0, & i \text{ odd} \end{cases}. \tag{7.208}$$

The z-transform of $\boldsymbol{\varepsilon}^{(2)}$ is

$$\mathscr{E}^{(2)}(z) = \varepsilon_{k_1} + \varepsilon_{k_1+1} z^{-2} + \cdots + \varepsilon_{k_1+\ell+L-1} z^{-2(\ell+L-1)} \tag{7.209}$$

and because $\boldsymbol{\varepsilon}^{(2)}$ is zero in the odd coordinates $\mathscr{E}^{(2)}(z) = \mathscr{E}(z^2)$. The corresponding z-transform of the $T/2$-spaced received signal error sequence associated with the error sequence $\boldsymbol{\varepsilon}^{(2)}$ is

$$\mathscr{E}_v^{(2)}(z) = \mathscr{E}^{(2)}(z)G^{(2)}(z). \tag{7.210}$$

From (7.172), the squared Euclidean distance of the error event in (7.161) is

$$\begin{aligned}
\left(\Delta^{(2)}\right)^2 &= \sum_{k=k_1}^{k_1+2\ell-1} \left| \sum_{i=0}^{2L} g_i^{(2)} \varepsilon_{k-i}^{(2)} \right|^2 \\
&= \left[\mathscr{E}_v^{(2)}(z) \mathscr{E}_v^{(2)^*}(1/z^*) \right]_0 \\
&= \left[\mathscr{E}^{(2)}(z) F^{(2)}(z) \mathscr{E}^{(2)^*}(1/z^*) \right]_0 \\
&= \left[\mathscr{E}(z^2) F^{(2)}(z) \mathscr{E}^*(1/z^{*2}) \right]_0,
\end{aligned} \tag{7.211}$$

where $F^{(2)}(z) = G^{(2)}(z)(G^{(2)}(1/z^*))^*$. Note that the polynomial $\mathscr{E}(z^2)\ \mathscr{E}^*(1/z^{*2})$ has the property that the odd powers of z have zero coefficients. Therefore, the contributions to the coefficient $[\mathscr{E}(z^2)F^{(2)}(z)\mathscr{E}^*(1/z^{*2})]_0$ arise only from the coefficients of $F^{(2)}(z)$ associated with even powers of z. Note also from (7.46) and (7.47) that the coefficients $f_{2k}^{(2)}$ of $F^{(2)}(z)$ associated with even powers of z are equal to the coefficients f_k of $F(z)$, that is, $f_{2k}^{(2)} = f_k$. Therefore,

$$\left(\Delta^{(2)}\right)^2 = [\ \mathscr{E}(z^2)F^{(2)}(z)\mathscr{E}^*(1/z^{*2})\ ]_0 = [\ \mathscr{E}(z)F(z)\mathscr{E}^*(1/z^*)\ ]_0 = \Delta^2. \tag{7.212}$$

Consequently, the T- and $T/2$-spaced MLSE receivers have identical error probability performance.

Example 7.9:
Let

$$\begin{aligned}
\mathscr{E}(z) &= \varepsilon_0 + \varepsilon_1 z^{-1} + \varepsilon_2 z^{-2}, \\
F^{(2)}(z) &= f_{-2}^{(2)} z^2 + f_{-1}^{(2)} z + f_0^{(2)} + f_1^{(2)} z^{-1} + f_2^{(2)} z^{-2}.
\end{aligned}$$

Then

$$\begin{aligned}
F(z) &= f_{-1}z + f_0 + f_1 z^{-1} \ = \ f_{-2}^{(2)} z + f_0^{(2)} + f_2^{(2)} z^{-1}, \\
\mathscr{E}^{(2)}(z) &= \varepsilon_0 + \varepsilon_1 z^{-2} + \varepsilon_2 z^{-4}.
\end{aligned}$$

Therefore,

$$[\,\mathscr{E}(z)F(z)\mathscr{E}^*(1/z^*)\,]_0 = \left(|\varepsilon_0|^2+|\varepsilon_1|^2+|\varepsilon_2|^2\right)^2 f_0 + \varepsilon_0\varepsilon_1^* f_1 + \varepsilon_1\varepsilon_0^* f_{-1}$$

and

$$[\,\mathscr{E}^{(2)}(z)F^{(2)}(z)\mathscr{E}^{(2)^*}(1/z^*)\,]_0 = \left(|\varepsilon_0|^2+|\varepsilon_1|^2+|\varepsilon_2|^2\right)^2 f_0^{(2)} + \varepsilon_0\varepsilon_1^* f_2^{(2)} + \varepsilon_1\varepsilon_0^* f_{-2}^{(2)}.$$

Hence, $\Delta^2 = (\Delta^{(2)})^2$.

7.6.3 Timing Phase Sensitivity

The conventional MLSE receiver based on T-spaced sampling at the output of the matched filter suffers from sensitivity to the sampler timing phase [220]. We now show that a $T/2$-spaced MLSE receiver is insensitive to the sampler timing phase. This can be readily seen by examining the sampled spectrum at the output of the matched filter. Suppose that the received pulse has less than 100% excess bandwidth, that is, $H(f) = 0$ for $|f| > 1/T$. With T-spaced sampling and a timing offset of t_o, the sampled spectrum is equal to

$$F_\Sigma(f) = \frac{1}{T}\sum_n F\left(f+\frac{n}{T}\right)\mathrm{e}^{\mathrm{j}2\pi(f+n/T)t_o}. \tag{7.213}$$

This spectrum exhibits aliasing for any received pulse having a bandwidth greater than the Nyquist frequency of $1/(2T)$ Hz. Consider, for example, the sampled spectrum at the Nyquist frequency $f = 1/(2T)$ Hz. If there is no timing offset, then

$$F_\Sigma(1/(2T)) = \frac{1}{T}\left(F(1/(2T)) + F(-1/(2T))\right).$$

In this case, the sampled spectrum at frequency $f = 1/(2T)$ Hz adds constructively. Now suppose there is a timing offset $t_o = T/2$. Then,

$$F_\Sigma(1/(2T)) = \frac{1}{T}\left(F(1/(2T))\mathrm{e}^{\mathrm{j}\pi/2} + F(-1/(2T))\mathrm{e}^{-\mathrm{j}\pi/2}\right) = 0.$$

In this case, the sampled spectrum at frequency $f = 1/(2T)$ Hz adds destructively and there is a null. Hence, the folded spectrum is sensitive to the timing offset with T-spaced sampling. With $T/2$-spaced sampling, on the other hand,

$$F_{\Sigma}(f) = \frac{2}{T}\sum_{n} F\left(f + \frac{2n}{T}\right) e^{j2\pi(f+2n/T)t_o}. \tag{7.214}$$

Notice that the sampled spectrum is a function of the timing offset t_o, but does not exhibit aliasing since $H(f)$ has no more than 100% excess bandwidth.

For a given a timing offset t_0, the sampled impulse response at the output of the matched filter is represented by the vector $\mathbf{f}_{t_0}^{(2)}$, where $f_{t_0,k}^{(2)} = f(kT' \pm t_0)$ and $T' = T/2$. Note that $f_{t_0,n}^{(2)} \neq \left(f_{t_0,-n}^{(2)}\right)^*$ when a timing offset is present. The discrete-time Fourier transform (DTFT) of $\mathbf{f}_{t_0}^{(2)}$ is

$$F_{t_0}^{(2)}(e^{j\omega}) = F^{(2)}(e^{j\omega})e^{\pm j\omega\tau_0}, \tag{7.215}$$

where $\tau_0 = t_0/T'$. If the sampler phase is known, then a discrete-time filter with a frequency response function $e^{\mp j\omega\tau_0}$ after the sampler will give the sampled response $\mathbf{f}^{(2)}$ at its output. However, as we now show, there is no need to correct this phase offset before performing MLSE equalization when $T/2$-spaced sampling is used.

The power spectrum of the noise at the output of the matched filter is independent of the timing offset t_0 and is given by

$$S_{vv}(f) = N_o F^{(2)}(e^{j\omega}). \tag{7.216}$$

Since the DTFT of the noise-whitening filter is

$$1/(G^{(2)}(1/z^*))^*|_{z=e^{j\omega}} = 1/(G^{(2)}(e^{j\omega}))^* \tag{7.217}$$

and we have

$$F^{(2)}(e^{j\omega}) \;=\; G^{(2)}(e^{j\omega})\,(G^{(2)}(e^{j\omega}))^* \;=\; |G^{(2)}(e^{j\omega})|^2, \tag{7.218}$$

it follows that the noise samples are white at the output of the noise-whitening filter with variance N_o. The DTFT of the message signal at the output of the noise-whitening filter is

$$X^{(2)}(e^{j\omega})G_{t_0}^{(2)}(e^{j\omega}) = X^{(2)}(e^{j\omega})G^{(2)}(e^{j\omega})e^{\pm j\omega\tau_0} \tag{7.219}$$

and we have

$$\sum_i |g_i^{(2)}|^2 = \sum_i |g_{t_0,i}^{(2)}|^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} |G^{(2)}(e^{j\omega})|^2 d\omega, \tag{7.220}$$

according to Parseval's theorem. Also

$$G_{t_0}^{(2)}(z)(G_{t_0}^{(2)}(1/z^*))^* = F^{(2)}(z). \tag{7.221}$$

Hence, the distances between allowed sequences of channel outputs with the $T/2$-spaced MLSE receiver in (7.211) remain the same and the performance is insensitive to the sampler phase $e^{\pm j\omega\tau_0}$. The MLSE receiver just needs to estimate the set of tap coefficients $\{g^{(2)}_{t_0,i}\}$ instead of $\{g^{(2)}_i\}$ for use in the branch metric calculations.

7.6.4 Practical $T/2$-Spaced MLSE Receiver

The receivers in Figs. 7.3 and 7.6 use a filter that is matched to the received pulse $h^*(-t)$. Since this filter requires knowledge of the unknown channel impulse response, it is impractical. One solution is to implement an "ideal" low-pass filter with a cutoff frequency of $1/T$ and sample the output at rate $2/T$. The noise samples at the output of this filter will be uncorrelated and, therefore, the $T/2$-spaced MLSE receiver can be implemented. Vachula and Hill [260] showed that this receiver is optimum; however, it has some drawbacks. First, it is not suitable for bandwidth efficient systems that use frequency division multiplexing, because the cutoff frequency of the low-pass filter will extend significantly into the adjacent band and pass large amounts of adjacent channel interference. Second, the ideal low-pass filter is noncausal and difficult to approximate. Another solution is to use a receiver filter that is matched to the transmitted pulse $h_{\rm a}(t)$ followed by a $T/2$-spaced sampler and a noise-whitening filter [54, 126]. If the transmitted signals are strictly bandlimited, for example, $h_{\rm a}(t)$ is a root raised cosine pulse, such that the received pulse $h(t)$ has at most 100% excess bandwidth,[6] then rate-$2/T$ sampling satisfies the Shannon sampling theorem and the $T/2$-spaced samples will provide sufficient statistics as we now show.

For simplicity, consider a time-invariant channel with impulse response $g(t)$. Let $H^{(2)}_a(z)$, $G^{(2)}(z)$, and $H^{(2)}(z)$ be the z-transforms of the $T/2$-spaced discrete-time signals corresponding to $h_{\rm a}(t)$, $g(t)$, and $h(t)$, respectively. The z-transform of the autocorrelation function of the noise samples at the output of the receive filter $h^*_{\rm a}(-t)$ is $N_{\rm o}F^{(2)}_h(z)$, where $F^{(2)}_h(z) = H^{(2)}_a(z)\left(H^{(2)}_a(1/z^*)\right)^*$. Using the factorization

$$F^{(2)}_h(z) = G^{(2)}_h(z)\left(G^{(2)}_h(1/z^*)\right)^*, \tag{7.222}$$

the $T/2$-spaced noise sequence can be whitened using a filter with transfer function $1/\left(G^{(2)}_h(1/z^*)\right)^*$ such that $G^{(2)}_h(z)$ has minimum phase. The resulting system is shown in Fig. 7.16. We now show that the receivers in Figs. 7.6 and 7.16 yield identical performance.

The z-transform of the $T/2$-spaced discrete-time white noise channel in Fig. 7.16 is

[6] The received pulse will a larger bandwidth due to Doppler spreading.

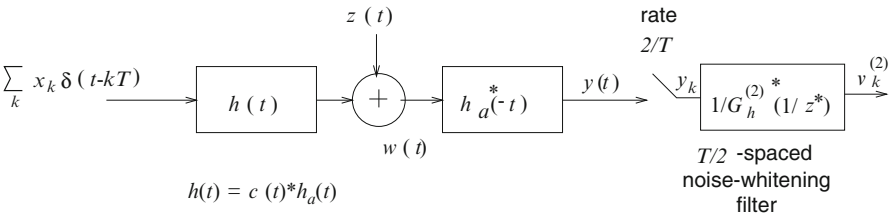

Fig. 7.16 Block diagram of system that implements a filter matched to $h_a(t)$ followed by a $T/2$-spaced noise-whitening filter. The structure of the noise-whitening filter depends only on the pulse $h_a(t)$

$$G_{\text{eq}}^{(2)}(z) = H_a^{(2)}(z)G^{(2)}(z)\left(H_a^{(2)}(1/z^*)\right)^* / \left(G_h^{(2)}(1/z^*)\right)^*$$
$$= G_h^{(2)}(z)G^{(2)}(z). \tag{7.223}$$

On the other hand, referring to the conventional system shown in Fig. 7.6, we have

$$H^{(2)}(z) = H_a^{(2)}(z)\, G^{(2)}(z) \tag{7.224}$$

and

$$F^{(2)}(z) = H_a^{(2)}(z)\left(H_a^{(2)}(1/z^*)\right)^* G^{(2)}(z)\left(G^{(2)}(1/z^*)\right)^*. \tag{7.225}$$

Let

$$G^{(2)}(z)\left(G^{(2)}(1/z^*)\right)^* = G_c^{(2)}(z)\left(G_c^{(2)}(1/z^*)\right)^*, \tag{7.226}$$

be a factorization of $G^{(2)}(z)\left(G^{(2)}(1/z^*)\right)^*$ such that $G_c^{(2)}(z)$ has minimum phase. Using (7.222), (7.225), and (7.226) yields

$$F^{(2)}(z) = G_h^{(2)}(z)\left(G_h^{(2)}(1/z^*)\right)^* G_c^{(2)}(z)\left(G_c^{(2)}(1/z^*)\right)^*. \tag{7.227}$$

If the system function of the noise-whitening filter is chosen as

$$1/\left(\left(G_h^{(2)}(1/z^*)\right)^*\left(G_c^{(2)}(1/z^*)\right)^*\right), \tag{7.228}$$

then the system function in Fig. 7.6 is

$$G^{(2)}(z) = G_h^{(2)}(z)G_c^{(2)}(z). \tag{7.229}$$

The system responses $G_{\text{eq}}^{(2)}(\text{e}^{\text{j}\omega})$ and $G^{(2)}(\text{e}^{\text{j}\omega})$ have the same amplitude response but a different phase response. Also

$$G_{\text{eq}}^{(2)}(z)\left(G_{\text{eq}}^{(2)}(1/z^*)\right)^* = F^{(2)}(z). \tag{7.230}$$

Therefore, the Euclidean distance between sequences of noise-whitening filter outputs for the T-spaced and $T/2$-spaced systems is the same. It follows that the systems shown in Figs. 7.6 and 7.16 yield identical performance. The main advantage of the system in Fig. 7.16 is that the noise-whitening filter does not depend on the unknown channel and has a fixed structure. Moreover, channel estimation can be performed after the noise-whitening filter and the Viterbi algorithm can be implemented using the metric in (7.148). Although the number of computations needed in the $T/2$-spaced MLSE receiver is twice that of a T-spaced receiver, the latter cannot be implemented for unknown channels. Moreover, a T-spaced MLSE receiver has poor performance when it is implemented with a matched filter that is derived from an inaccurate channel estimate [196].

7.7 Co-channel Demodulation

By extending Forney's maximum likelihood receiver [104], Van Etten [92] proposed MLSE for the joint detection of co-channel signals. In this section, we derive optimum and suboptimum MLSE receivers for co-channel demodulation of digital signals corrupted by ISI. By modeling the overall system as a discrete-time MIMO channel, the optimum MIMO joint MLSE (J-MLSE) receiver is derived. By following a parallel argument used for single-input single-output (SISO) channels, a $T/2$-spaced MIMO J-MLSE receiver is shown to have the same performance as the T-spaced receiver, but with insensitivity to timing phase errors. The optimality of a practical $T/2$-spaced receiver is shown, which consists of a filter that is matched to the transmitted pulse, followed by a rate-$2/T$ sampler, a $T/2$-spaced noise-whitening filter, and a Viterbi algorithm. The optimum MIMO J-MLSE receiver requires complete knowledge of the complex channel gains associated with all co-channel signals. In many cases, such complete knowledge is impractical or even infeasible to obtain. For such cases, we discuss an interference rejection combining MLSE (IRC-MLSE) receiver that only requires knowledge of the complex channel gain for the desired signal but not the co-channel interferers.

7.7.1 System and Channel Model

Consider a system where the signals from K co-channel signals are received by J antenna elements. This system can be modeled by a $K \times J$ MIMO channel, where the MIMO channel inputs are the symbol sequences from the K co-channel users, and the MIMO channel outputs are the combination of the signals that are received from the co-channel users at each of the J receiver antenna elements. The co-channel detection problem is mathematically very similar to the CDMA multiuser detection problem. However, while the users in a CDMA system are distinguished by different spreading sequences as discussed in Chap. 9, the K co-channel users will all use the

same transmitter pulse shaping filter $h_{\mathrm{a}}(t)$. The impulse response of the channel between the kth user and the jth antenna element is denoted by $g^{(k,j)}(t,\tau)$, where we have assumed that the channels are modeled as time-variant linear filters. While the channel introduces frequency and time selectivity into the received signals, it also allows the co-channel signals to be distinguished at the receiver, since the received pulses $h^{(k,j)}(t) = g^{(k,j)}(t,\tau) * h_{\mathrm{a}}(t)$ are all distinct due to the different channel impulse responses. The received pulses $h^{(k,j)}(t)$ are all assumed to have a length of at most $L^{(k,j)}T$, that is, $h^{(k,j)}(t) = 0$ for $t \leq 0$ and $t \geq (L^{(k,j)}+1)T$. The received complex envelope at the jth antenna element is

$$\tilde{r}^{(j)}(t) = \sum_{k=1}^{K} \sum_{\ell} x_{\ell}^{(k)} h^{(k,j)}(t - \ell T - \tau_k) + \tilde{n}^{(j)}(t), \tag{7.231}$$

where τ_k, $0 \leq \tau_k \leq T$ is the random transmission delay due to symbol asynchronous users, and $\tilde{n}^{(j)}(t)$ is AWGN assumed to be independent on the different antenna branches.

7.7.2 Joint Maximum Likelihood Co-channel Receiver

The J-MLSE receiver processes the total received vector

$$\tilde{\mathbf{r}}(t) = \left(\tilde{r}^{(1)}(t), \tilde{r}^{(2)}(t), \ldots, \tilde{r}^{(J)}(t)\right), \tag{7.232}$$

to generate a ML estimate of the information sequence

$$\mathbf{x} = (\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(K)}), \tag{7.233}$$

where $\mathbf{x}^{(k)} = \{x_{\mathrm{n}}^{(k)}\}$. To derive the structure of the joint ML receiver, we follow the same approach used in Sect. 7.1.1. Let $\{\varphi_n(t)\}$ denote a complete set of complex orthonormal basis functions that span the entire duration of the observation vector $\tilde{\mathbf{r}}(t)$. Then

$$\tilde{r}^{(j)}(t) = \sum_{n=1}^{N} \tilde{r}_n^{(j)} \varphi_n(t), \tag{7.234}$$

where

$$\tilde{r}_n^{(j)} = \sum_{k=1}^{K} \sum_{\ell} x_{\ell}^{(k)} h_{\ell_n}^{(k,j)} + \tilde{n}_n^{(j)} \tag{7.235}$$

and

$$h_{\ell_n}^{(k,j)} = \int_{-\infty}^{\infty} h^{(k,j)}(t - \ell T - \tau_k) \varphi_n^*(t) \mathrm{d}t, \tag{7.236}$$

$$\tilde{n}_n^{(j)} = \int_{-\infty}^{\infty} \tilde{n}^{(j)}(t) \varphi_n^*(t) \mathrm{d}t. \tag{7.237}$$

Define the received vector

$$\tilde{\mathbf{r}} = \text{vec}(\tilde{\mathbf{r}}^{(1)}, \tilde{\mathbf{r}}^{(2)}, \ldots, \mathbf{r}^{(J)}), \tag{7.238}$$

where $\tilde{\mathbf{r}}^{(j)} = \{\tilde{r}_n^{(j)}\}$. Since the noise components $\tilde{n}_n^{(j)}$ associated with the J antenna elements are uncorrelated zero-mean complex Gaussian random variables with variance $\frac{1}{2}\text{E}[|\tilde{n}_n^{(j)}|^2] = N_\text{o}$, the received vector $\tilde{\mathbf{r}}$ has the complex multivariate Gaussian density

$$p(\tilde{\mathbf{r}}|\mathbf{x},\mathbf{h}) = \prod_{n=1}^{N} \frac{1}{2\pi N_\text{o}} \exp\left\{ -\frac{1}{2N_\text{o}} \sum_{j=1}^{J} \left| \tilde{r}_n^{(j)} - \sum_{k=1}^{K} \sum_{\ell} x_\ell^{(k)} h_{\ell_n}^{(k,j)} \right|^2 \right\}, \tag{7.239}$$

where $\mathbf{h} = \{h_{\ell_n}^{(k,j)}\}$.

The optimum receiver chooses $\mathbf{x}$ to maximize $p(\tilde{\mathbf{r}}|\mathbf{x},\mathbf{h})$ or, equivalently, the metric

$$\mu(\mathbf{x}) = -\sum_{n=1}^{N} \sum_{j=1}^{J} \left| \tilde{r}_n^{(j)} - \sum_{k=1}^{K} \sum_{\ell} x_\ell^{(k)} h_{\ell_n}^{(k,j)} \right|^2. \tag{7.240}$$

Since $\sum_{n=1}^{N} \sum_{j=1}^{J} |\tilde{r}_n^{(j)}|^2$ is independent of $\mathbf{x}$, maximizing (7.240) is equivalent to maximizing

$$\begin{aligned} \mu(\mathbf{x}) = \sum_{j=1}^{J} \Bigg(& 2\text{Re}\left\{ \sum_{k=1}^{K} \sum_{\ell} x_\ell^{(k)*} \sum_{n=1}^{N} \tilde{r}_n^{(j)} h_{\ell_n}^{(k,j)*} \right\} \\ & - \sum_{k=1}^{K} \sum_{k'=1}^{K} \sum_{\ell} \sum_{\ell'} x_\ell^{(k)} x_{\ell'}^{(k')*} \sum_{n=1}^{N} h_{\ell_n}^{(k,j)} h_{\ell'_n}^{(k',j)*} \Bigg). \end{aligned} \tag{7.241}$$

In the limit as the number of observable random variables N approaches infinity, we define

$$\begin{aligned} y_\ell^{(k,j)} &\triangleq \lim_{N\to\infty} \sum_{n=1}^{N} \tilde{r}_n^{(j)} h_{\ell_n}^{(k,j)*} \\ &= \int_{-\infty}^{\infty} \tilde{r}^{(j)}(t) h^{(k,j)*}(t - \ell T - \tau_k)\text{d}t, \end{aligned} \tag{7.242}$$

$$\begin{aligned} f_{\ell'-\ell}^{(k,k',j)} &\triangleq \lim_{N\to\infty} \sum_{n=1}^{N} h_{\ell_n}^{(k,j)} h_{\ell'_n}^{(k',j)*} \\ &= \int_{-\infty}^{\infty} h^{(k,j)}(t - \ell T - \tau_k) h^{(k',j)*}(t - \ell' T - \tau_{k'})\text{d}t. \end{aligned} \tag{7.243}$$

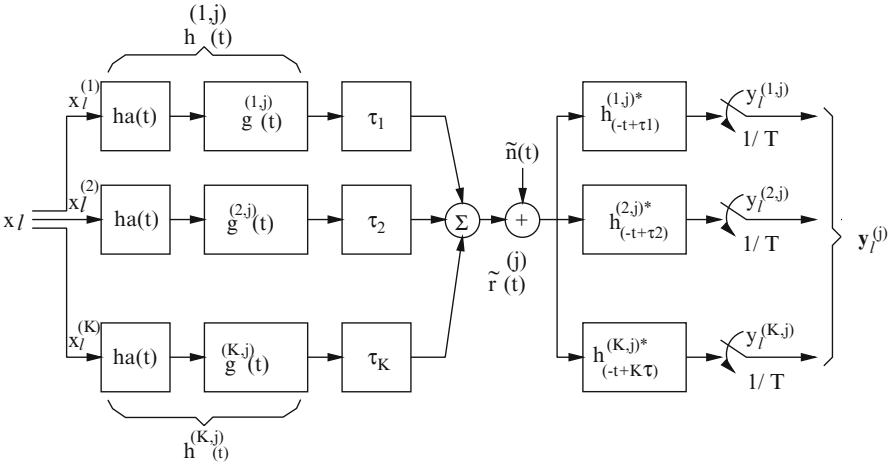

Fig. 7.17 Overall MIMO system model for co-channel demodulation

Using (7.242) and (7.243) in (7.241), we arrive at the final form

$$\mu(\mathbf{x}) = \sum_{j=1}^{J} \left\{ 2\mathrm{Re}\left(\sum_{k=1}^{K} \sum_{\ell} x_{\ell}^{(k)*} y_{\ell}^{(k,j)} \right) - \sum_{k=1}^{K} \sum_{k'=1}^{K} \sum_{\ell} \sum_{\ell'} x_{\ell}^{(k)} x_{\ell'}^{(k')*} f_{\ell'-\ell}^{(k,k',j)} \right\}. \tag{7.244}$$

The variables $\{y_{\ell}^{(k,j)}\}$ in (7.242) are obtained by passing the received complex envelope $\tilde{r}_n^{(j)}(t)$ through the matched filter $h^{(k,j)*}(-t-\tau_k)$ and sampling the output.[7] Note that there are K such matched filters on each of the J receiver antenna elements, and the samples must be taken with the timing delay $\tau_k, k = 1, \ldots, K$. The above development leads to the overall system model shown in Fig. 7.17. Finally, by changing the variable of integration, the $\{f_{\ell'-\ell}^{(k,k',j)}\}$ in (7.243) can be rewritten in the form

$$f_n^{(k,k',j)} = \int_{-\infty}^{\infty} h^{(k,j)}(t - nT - \tau_k) h^{(k',j)*}(t - \tau_{k'}) \mathrm{d}t. \tag{7.245}$$

where $n = \ell' - \ell$. Using (7.245), it can be shown that the ISI coefficients have the symmetric property

$$f_n^{(k,k',j)} = f_{-n}^{(k',k,j)*}. \tag{7.246}$$

[7] It is assumed that the differential delay across the multiple receiver antenna elements is negligible.

7.7.3 Discrete-Time MIMO Channel Model

Sampling the outputs of the matched filter $h^{(k,j)^*}(-t-\tau_k)$ in Fig. 7.17 every T seconds yields the sample sequence $\{y_\ell^{(k,j)}\}$, where

$$\begin{aligned} y_\ell^{(k,j)} &= \sum_{k'=1}^{K} \sum_{n} x_{\mathrm{n}}^{(k')} f_{\ell-n}^{(k',k,j)} + \nu_\ell^{(k,j)} \\ &= \sum_{k'=1}^{K} \sum_{m=-L_j-1}^{L_j+1} x_{\ell-m}^{(k')} f_m^{(k',k,j)} + \nu_\ell^{(k,j)} \end{aligned} \tag{7.247}$$

and

$$\nu_\ell^{(k,j)} = \int_{-\infty}^{\infty} \tilde{n}^{(j)}(t) h^{(k,j)^*}(t-\ell T-\tau_k)\mathrm{d}t, \tag{7.248}$$

and where $L_j = \max_k L^{(k,j)}$ is the maximum discrete-time channel length. This is an extension of the result in (7.17) for SISO ISI channels. However, for MIMO ISI channels, the range of summation must be expanded from $(-L_j, L_j)$ to $(-L_j - 1, L_j + 1)$ to account for the random user delays in the case of symbol asynchronous users. If we define

$$\mathbf{y}_\ell^{(j)} \triangleq \left(y_\ell^{(1,j)}, \ldots, y_\ell^{(K,j)}\right)^{\mathrm{T}}, \tag{7.249}$$

$$\mathbf{x}_\ell \triangleq \left(x_\ell^{(1)}, \ldots, x_\ell^{(K)}\right)^{\mathrm{T}}, \tag{7.250}$$

$$\nu_\ell^{(j)} \triangleq \left(\nu_\ell^{(1,j)}, \ldots, \nu_\ell^{(K,j)}\right)^{\mathrm{T}}, \tag{7.251}$$

$$\mathbf{F}_m^{(j)} \triangleq \left[f_{\mathrm{m}}^{(k',k,j)}\right]_{K\times K}, \tag{7.252}$$

then (7.247) leads to the convenient matrix representation

$$\mathbf{y}_\ell^{(j)} = \sum_{m=-L_j-1}^{L_j+1} \mathbf{F}_m^{(j)} \mathbf{x}_{\ell-m} + \nu_\ell^{(j)}. \tag{7.253}$$

From (7.244) and (7.253), it follows that the joint maximum likelihood receiver uses the entire set of J observation vectors $\mathbf{y} = \{\mathbf{y}_\ell^{(j)}, j = 1, \ldots, J\}$ along with knowledge of the entire set of ISI coefficient matrices $\mathbf{F} = \{\mathbf{F}_m^{(j)}, m = -L_j - 1, \ldots, L_j + 1, j = 1, \ldots, J\}$ to determine the most likely transmitted sequence $\mathbf{x} = \{\mathbf{x}_\ell\}$. Hence, the joint maximum likelihood receiver decides in favor of the symbol sequence $\mathbf{x}$ that maximizes the joint likelihood function $p(\mathbf{y}|\mathbf{x}, \mathbf{F})$ or the log-likelihood function $\log\{p(\mathbf{y}|\mathbf{x}, \mathbf{F})\}$.

The noise samples in (7.248) are zero-mean complex Gaussian variables that are correlated by the matched filters $h^{(k,j)^*}(-t-\tau_k), k = 1, \ldots, K$. The vector of zero mean noise samples $\nu_\ell^{(j)}$ has the covariance matrix

$$\Phi_{\boldsymbol{\nu}^{(j)}\boldsymbol{\nu}^{(j)}}(n) = \frac{1}{2}\mathrm{E}\left[\nu_{\ell}^{(j)}\nu_{\ell+n}^{(j)^H}\right] = N_{\mathrm{o}}\mathbf{F}_n^{(j)}. \tag{7.254}$$

As is the case with SISO ISI channels, the correlation of the noise samples complicates the implementation of various equalization schemes. However, a *matrix* noise-whitening filter can be used to whiten the vector of noise samples $\nu_{\ell}^{(j)}$. The z-transform of the noise covariance matrix for the jth antenna element can be defined as $\Phi_{\boldsymbol{\nu}^{(j)}\boldsymbol{\nu}^{(j)}}(z) = N_{\mathrm{o}}\mathbf{F}^{(j)}(z)$, where [84]

$$\mathbf{F}^{(j)}(z) = \sum_{n=-L_j-1}^{L_j+1} \mathbf{F}_n^{(j)} z^{-n} \tag{7.255}$$

and, once again, $L_j = \max_k L^{(k,j)}$ is the maximum discrete-time channel length. This is a straightforward extension of the SISO ISI channels where $F(z) = \sum_{n=-L}^{L} f_n z^{-n}$. However, an MIMO channel is described by a covariance matrix, and the range of summation in (7.255) must be expanded from $(-L_j, L_j)$ to $(-L_j - 1, L_j + 1)$ to account for the random user delays in the case of symbol asynchronous users.

Using the symmetric property $f_n^{(k,m,j)} = f_{-n}^{(m,k,j)^*}$ of the ISI coefficients, it follows that $\mathbf{F}_n^{(j)} = \mathbf{F}_{-n}^{(j)^H}$. Therefore, $\mathbf{F}^{(j)}(z)$ has the symmetric form

$$\mathbf{F}^{(j)}(z) = \mathbf{F}^{(j)^H}(1/z^*). \tag{7.256}$$

It follows that the matrix $\mathbf{F}^{(j)}(z)$ can be factored as

$$\mathbf{F}^{(j)}(z) = \mathbf{G}^{(j)}(z)\mathbf{G}^{(j)^H}(1/z^*) \tag{7.257}$$

and the filter $[\mathbf{G}^{(j)^H}(1/z^*)]^{-1}$ will serve as a noise-whitening matrix filter.

Example 7.10:

Consider a two user system with a single receiver antenna. Since $J = 1$, we can omit the index (j). Let $L = 1$ and $\mathbf{F}(z)$ be

$$\begin{aligned}
\mathbf{F}(z) &= \mathbf{F}_1^H z + \mathbf{F}_0 + \mathbf{F}_1 z^{-1} \\
&= \begin{bmatrix} 1 & 0.48 + 0.48z^{-1} \\ 0.48 + 0.48z & 1 \end{bmatrix} \\
&= \begin{bmatrix} 0 & 0 \\ 0.48 & 0 \end{bmatrix} z + \begin{bmatrix} 1 & 0.48 \\ 0.48 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0.48 \\ 0 & 0 \end{bmatrix} z^{-1}.
\end{aligned}$$

The matrix spectral factorization of $\mathbf{F}(z)$ has the form

$$\begin{aligned}\mathbf{F}(z) &= \mathbf{G}^H(1/z^*)\mathbf{G}(z)\\ &= \left[\mathbf{G}_0 + \mathbf{G}_1 z^*\right]^H \left[\mathbf{G}_0 + \mathbf{G}_1 z^{-1}\right].\end{aligned}$$

$\mathbf{F}_0$ and $\mathbf{F}_1$ can be represented by

$$\begin{aligned}\mathbf{F}_0 &= \mathbf{G}_0^H\mathbf{G}_0 + \mathbf{G}_1^H\mathbf{G}_1,\\ \mathbf{F}_1 &= \mathbf{G}_0^H\mathbf{G}_1,\end{aligned}$$

where $\mathbf{G}_0$ is lower triangular and $\mathbf{F}_1$ is upper triangular with zero diagonal. In turn, $\mathbf{G}_1$ must be upper triangular with zero diagonal. This results in the spectral factorization

$$\mathbf{F}(z) = \left[\begin{bmatrix}0.8 & 0\\ 0.6 & 0.8\end{bmatrix} + \begin{bmatrix}0 & 0.6\\ 0 & 0\end{bmatrix} z^*\right]^H \left[\begin{bmatrix}0.8 & 0\\ 0.6 & 0.8\end{bmatrix} + \begin{bmatrix}0 & 0.6\\ 0 & 0\end{bmatrix} z^{-1}\right].$$

The matrix noise-whitening filter $[\mathbf{G}^H(1/z^*)]^{-1}$ is noncausal and stable with an infinite length.

$$\begin{aligned}[\mathbf{G}^H(1/z^*)]^{-1} &= \begin{bmatrix}0.8 & 0.6\\ 0.6z & 0.8\end{bmatrix}^{-1}\\ &= \frac{1}{0.64 - 0.36z}\begin{bmatrix}0.8 & -0.6\\ -0.6z & 0.8\end{bmatrix}.\end{aligned}$$

In practice, the filter $[\mathbf{G}^H(1/z^*)]^{-1}$ can be approximated with a finite length filter using a sufficient filter delay. Finally, the overall discrete-time white noise matrix channel has transfer function

$$\mathbf{G}(z) = \begin{bmatrix}0.8 & 0\\ 0.6 & 0.8\end{bmatrix} + \begin{bmatrix}0 & 0.6\\ 0 & 0\end{bmatrix} z^{-1}.$$

Let $\mathbf{x}(z) = (x^{(1)}(z), x^{(2)}(z), \ldots, x^{(K)}(z))^{\mathrm{T}}$ be the z-transform of the input sequence $\{\mathbf{x}_n\}$ to the channel, where $\mathbf{x}_n = (x_{\mathrm{n}}^{(1)}, x_{\mathrm{n}}^{(2)}, \ldots, x_{\mathrm{n}}^{(K)})^{\mathrm{T}}$, and let $\mathbf{v}^{(j)}(z) = (v^{(1,j)}(z), v^{(2,j)}(z), \ldots, v^{(K,j)}(z))^{\mathrm{T}}$ be the z-transform of the output sequence $\{\mathbf{v}_n^{(j)}\}$ from the noise-whitening filter on antenna j, where $\mathbf{v}_n^{(j)} = (v_{\mathrm{n}}^{(1,j)}, v_{\mathrm{n}}^{(2,j)}, \ldots, v_{\mathrm{n}}^{(K,j)})^{\mathrm{T}}$. Then

$$\mathbf{v}^{(j)}(z) = \mathbf{G}^{(j)}(z)\mathbf{x}(z) + \mathbf{w}^{(j)}(z), \tag{7.258}$$

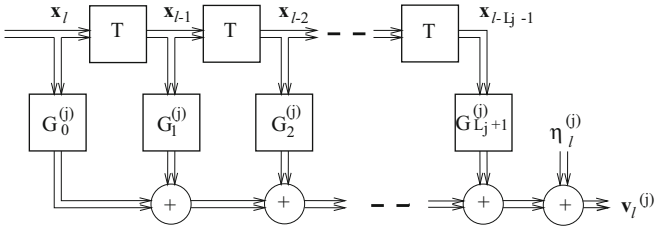

Fig. 7.18 Discrete-time white noise MIMO channel model

where $\mathbf{w}^{(j)}(z) = (w^{(1,j)}(z), w^{(2,j)}(z), \ldots, w^{(K,j)}(z))^T$ is the z-transform of the sequence of white Gaussian noise vectors $\mathbf{w}_{\mathrm{n}}^{(j)} = (w_{\mathrm{n}}^{(1,j)}, w_{\mathrm{n}}^{(2,j)}, \ldots, w_{\mathrm{n}}^{(K,j)})^T$. In the time-domain

$$\mathbf{v}_n^{(j)} = \sum_{k=0}^{L_j} \mathbf{G}_k^{(j)} \mathbf{x}_{n-k} + \mathbf{w}_{\mathrm{n}}^{(j)}, \tag{7.259}$$

where the zero mean noise vector $\mathbf{w}_{\mathrm{n}}^{(j)}$ has covariance matrix

$$\Phi_{\mathbf{w}_{\mathrm{n}}^{(j)}\mathbf{w}_{\mathrm{n}}^{(j)}}(n) = \frac{1}{2}\mathrm{E}\left[\tilde{\mathbf{w}}_{\ell}^{(j)} \tilde{\mathbf{w}}_{\ell+n}^{(j)H}\right] = N_{\mathrm{o}}\delta_n \mathbf{I}. \tag{7.260}$$

The optimum receiver consists of a bank of K matched filters at the output of each antenna element, followed by a baud-rate sampler and a $K \times K$ matrix noise-whitening filter. With J-branch diversity reception, the overall matrix channel consisting of the transmit filters, channels, matched filters, samplers, and matrix noise-whitening filters can be modeled as a parallel collection of J T-spaced matrix filters with independent white noise sequences as shown in Fig. 7.18. To determine the number of states in the overall channel model, we first define $L_k = \max_j L^{(k,j)}$ as the length of the channel for the kth input. Then there are $2^{n\sum_{k=1}^{K} L_k}$ states, where 2^n is the size of the signal constellation.

7.7.4 The Viterbi Algorithm

Suppose that m symbols from each of the K transmitters have been transmitted over the channel. Let $\mathbf{V}_n = \mathrm{vec}(\mathbf{v}_n^{(1)^{\mathrm{T}}}, \mathbf{v}_n^{(2)^{\mathrm{T}}}, \ldots, \mathbf{v}_n^{(J)^{\mathrm{T}}})^{\mathrm{T}}$, where $\mathbf{v}_n^{(j)} = (v_{\mathrm{n}}^{(1,j)}, \ldots, v_{\mathrm{n}}^{(K,j)})^{\mathrm{T}}$ denote the collection of vectors at the outputs of the matrix noise-whitening filters for the J antenna branches at epoch n. After receiving the output sequence $\{\mathbf{V}_n\}_{n=1}^m$, the ML receiver decides in favor of the sequence of input vectors $\{\mathbf{x}_n\}_{n=1}^m$ that maximizes the log-likelihood function

$$\begin{aligned}&\log\{p(\mathbf{V}_m,\ldots,\mathbf{V}_1|\mathbf{x}_m,\ldots,\mathbf{x}_1)\}\\&\quad=\log\left\{p\left(\mathbf{V}_m|x_m^{(1)},\ldots,x_{m-L_1}^{(1)};x_m^{(2)},\ldots,x_{m-L_2}^{(2)};\ldots;x_m^{(K)},\ldots,x_{m-L_K}^{(K)}\right)\right\}\\&\qquad+\log\{p(\mathbf{V}_{m-1},\ldots,\mathbf{V}_1|\mathbf{x}_{m-1},\ldots,\mathbf{x}_1)\}.\end{aligned}\tag{7.261}$$

The first term on the right-hand side of (7.261) is the *branch metric* used in the Viterbi algorithm. The discrete-time white noise matrix channel model leads to the conditional density function

$$\begin{aligned}&\log\left\{p\left(\mathbf{V}_m|x_m^{(1)},\ldots,x_{m-L_1}^{(1)};x_m^{(2)},\ldots,x_{m-L_2}^{(2)};\ldots;x_m^{(K)},\ldots,x_{m-L_K}^{(K)}\right)\right\}\\&\quad=\frac{1}{(2\pi N_\mathrm{o})^{KJ}}\exp\left\{-\frac{1}{2N_\mathrm{o}}\sum_{j=1}^{J}\|\mathbf{v}_m^{(j)}-\sum_{n=0}^{L}\mathbf{G}_n^{(j)}\mathbf{x}_{m-n}\|^2\right\},\end{aligned}\tag{7.262}$$

where $L=\max_k L_k$. Note that some elements in the matrix $\mathbf{G}_n^{(j)}$ may be zero if $L_k\neq L$, $k=1,\ldots,K$ in which case the branch metric computation can be somewhat simplified. The density in (7.262) leads to the branch metric

$$\mu_m=-\sum_{j=1}^{J}\left\|\mathbf{v}_m^{(j)}-\sum_{n=0}^{L}\mathbf{G}_n^{(j)}\mathbf{x}_{m-n}\right\|^2.\tag{7.263}$$

7.7.5 Pairwise Error Probability

Let $\mathbf{x}$ and $\hat{\mathbf{x}}$ be the transmitted and estimated symbol sequences, respectively, and define the error sequence $\varepsilon=\mathbf{x}-\hat{\mathbf{x}}$. The pairwise error probability is the probability that the receiver decides in favor of sequence $\hat{\mathbf{x}}$ when sequence $\mathbf{x}$ was transmitted, and is equal to

$$\mathrm{P}[\Gamma(\mathbf{x}-\varepsilon)\geq\Gamma(\mathbf{x})|\mathbf{x}]=\mathrm{P}[\Gamma(\hat{\mathbf{x}})>\Gamma(\mathbf{x})],\tag{7.264}$$

where $\Gamma(\mathbf{x})=\sum_m\mu_m$ is the path metric associated with the input sequence $\mathbf{x}$ with the branch metric μ_m defined in (7.263). From (7.263),

$$\begin{aligned}\mathrm{P}[\Gamma(\mathbf{x}-\varepsilon)\geq\Gamma(\mathbf{x})|\mathbf{x}]&=\mathrm{P}\left[\sum_m\sum_{j=1}^{J}\|\sum_{n=0}^{L}\mathbf{G}_n^{(j)}\varepsilon_{m-n}+\mathbf{w}_m^{(j)}\|^2<\sum_m\sum_{j=1}^{J}\|\mathbf{w}_m^{(j)}\|^2\right]\\&=\mathrm{P}\left[\sum_m\sum_{j=1}^{J}2\mathrm{Re}\left\{\sum_{n=0}^{L}\mathbf{w}_m^{(j)H}\mathbf{G}_n^{(j)}\varepsilon_{m-n}\right\}\right.\\&\qquad\left.<-\sum_m\sum_{j=1}^{J}\|\sum_{n=0}^{L}\mathbf{G}_n^{(j)}\varepsilon_{m-n}\|^2\right].\end{aligned}\tag{7.265}$$

Define

$$\Delta^2 \stackrel{\triangle}{=} \sum_m \sum_{j=1}^{J} \| \sum_{n=0}^{L} \mathbf{G}_n^{(j)} \boldsymbol{\varepsilon}_{n-m} \|^2, \tag{7.266}$$

$$\chi \stackrel{\triangle}{=} \sum_m \sum_{j=1}^{J} 2\mathrm{Re} \left\{ \sum_{n=0}^{L} \mathbf{w}_m^{(j)H} \mathbf{G}_n^{(j)} \boldsymbol{\varepsilon}_{m-n} \right\}. \tag{7.267}$$

It can be shown that χ is a zero-mean Gaussian random variable with variance $4N_o\Delta^2$. Therefore, the pairwise error probability becomes

$$\mathrm{P}[\Gamma(\mathbf{x}-\boldsymbol{\varepsilon}) \geq \Gamma(\mathbf{x})|\mathbf{x}] = Q\left(\sqrt{\frac{\Delta^2}{4N_o}}\right). \tag{7.268}$$

7.7.6 T/2-Spaced MIMO J-MLSE Receiver

Suppose that the matched filter outputs $h^{(k,j)^*}(-t-\tau_k), k=1,\ldots,K$ are sampled every $T/2$ seconds at the correct timing phase. Again, the zero mean noise samples at the output of the matched filters are correlated due to the matched filtering, and their covariance matrix is

$$\tilde{\boldsymbol{\Phi}}_{\tilde{\boldsymbol{\nu}}^{(j)}\tilde{\boldsymbol{\nu}}^{(j)}}(n) = \frac{1}{2}\mathrm{E}\left[\tilde{\boldsymbol{\nu}}_\ell^{(j)} \tilde{\boldsymbol{\nu}}_{\ell+n}^{(j)H}\right] = N_o \tilde{\mathbf{F}}_n^{(j)}, \tag{7.269}$$

where the elements of the matrix $\tilde{\mathbf{F}}_n^{(j)}$ are $\tilde{f}_n^{(k,k',j)} = \tilde{f}^{(k,k',j)}(nT/2)$, the function $f^{(k,k',j)}(t)$ is defined in (7.245), and where the tilde denotes rate $2/T$ sampling. It follows that the z-transform of the noise covariance matrix for the jth antenna element can be defined as $\tilde{\boldsymbol{\Phi}}_{\tilde{\boldsymbol{\nu}}^{(j)}\tilde{\boldsymbol{\nu}}^{(j)}}(z) = N_o \tilde{\mathbf{F}}^{(j)}(z)$, where

$$\tilde{\mathbf{F}}^{(j)}(z) = \sum_{n=-2L_j-1}^{2L_j+1} \tilde{\mathbf{F}}_n^{(j)} z^{-n}. \tag{7.270}$$

Since $\tilde{f}_\ell^{(k,k',j)} = \tilde{f}_{-\ell}^{(k',k,j)^*}$, it follows that $\tilde{\mathbf{F}}^{(j)}(z)$ has the factorization

$$\tilde{\mathbf{F}}^{(j)}(z) = \tilde{\mathbf{G}}^{(j)}(z)\tilde{\mathbf{G}}^{(j)H}(1/z^*). \tag{7.271}$$

As with baud-rate sampling, the $T/2$-spaced noise samples can be whitened using a stable noncausal matrix noise-whitening filter with the transfer function $\left[\tilde{\mathbf{G}}^{(j)H}(1/z^*)\right]^{-1}$. Analogous to (7.258), the z-transform of the output of the matrix

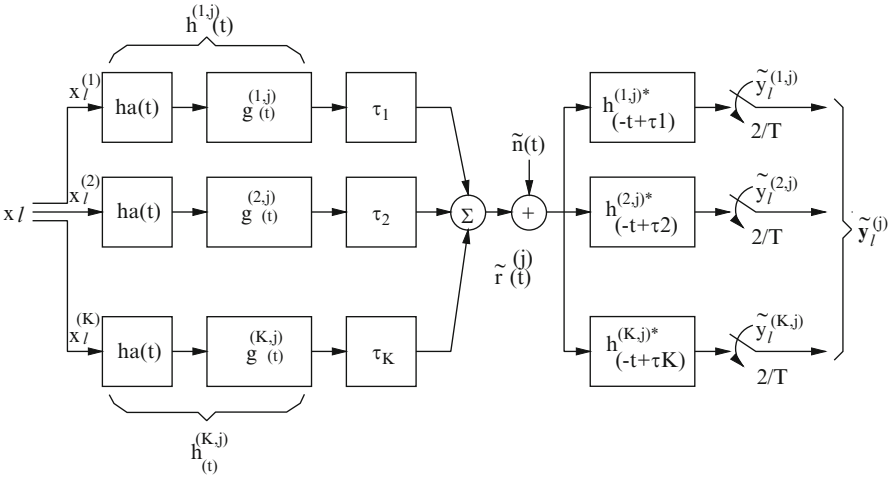

Fig. 7.19 Overall MIMO system model with $T/2$-spaced sampling

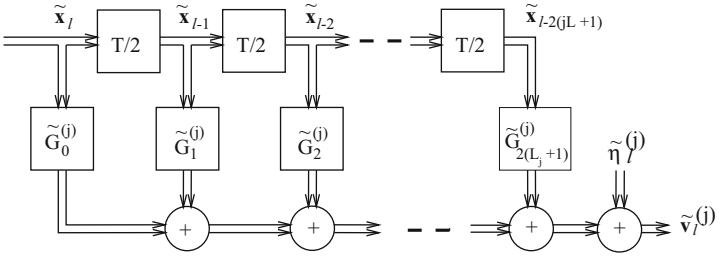

Fig. 7.20 Overall discrete-time white noise MIMO channel model with $T/2$-spaced sampling

noise-whitening filter is

$$\tilde{\mathbf{v}}^{(j)}(z) = \tilde{\mathbf{G}}^{(j)}(z)\tilde{\mathbf{x}}(z) + \tilde{\mathbf{w}}^{(j)}(z) \tag{7.272}$$

or in the time-domain

$$\tilde{\mathbf{v}}_\ell^{(j)} = \sum_{n=0}^{2L_j} \tilde{\mathbf{G}}_n^{(j)} \tilde{\mathbf{x}}_{\ell-n} + \tilde{\mathbf{w}}_\ell^{(j)}, \tag{7.273}$$

where $\{\tilde{\mathbf{w}}_\ell^{(j)}\}$ is a $T/2$-spaced white noise sequence with power spectrum $S_{\tilde{\mathbf{w}}\tilde{\mathbf{w}}}(f) = N_\mathrm{o}\mathbf{I}$. The sequence $\{\tilde{\mathbf{x}}_n\}$ is the corresponding $T/2$-spaced input symbol sequence and is given by

$$\tilde{\mathbf{x}}_n = \begin{cases} \mathbf{x}_{n/2}, & n = 0,2,4,\ldots \\ 0, & n = 1,3,5,\ldots \end{cases} \tag{7.274}$$

The overall system and equivalent discrete-time white noise models are shown in Figs. 7.19 and 7.20, respectively.

Note that the vector samples $\tilde{\mathbf{v}}_{2\ell}^{(j)}$ and $\tilde{\mathbf{v}}_{2\ell+1}^{(j)}$ correspond to the ℓth received baud, such that

$$\tilde{\mathbf{v}}_{2\ell}^{(j)} = \sum_{n=0}^{L_j} \tilde{\mathbf{G}}_{2n}^{(j)} \mathbf{x}_{\ell-n} + \tilde{\mathbf{w}}_{2\ell}^{(j)} \tag{7.275}$$

$$\tilde{\mathbf{v}}_{2\ell+1}^{(j)} = \sum_{n=0}^{L_j-1} \tilde{\mathbf{G}}_{2n+1}^{(j)} \mathbf{x}_{\ell-n} + \tilde{\mathbf{w}}_{2\ell+1}^{(j)}. \tag{7.276}$$

With $T/2$-spaced fractional sampling, there are two samples per baud, and the branch metric becomes

$$\mu_m = -\sum_{j=1}^{J} \left(\|\tilde{\mathbf{v}}_{2m}^{(j)} - \sum_{n=0}^{L} \tilde{\mathbf{G}}_{2n}^{(j)} \mathbf{x}_{m-n}\|^2 + \|\tilde{\mathbf{v}}_{2m+1}^{(j)} - \sum_{n=0}^{L-1} \tilde{\mathbf{G}}_{2n+1}^{(j)} \mathbf{x}_{m-n}\|^2 \right). \tag{7.277}$$

Once again, if $L_k \neq L, k = 1, \ldots, K$ then some of the $\tilde{\mathbf{G}}_{2n}^{(j)}$ and $\tilde{\mathbf{G}}_{2n+1}^{(j)}$ may be zero. Notice that $T/2$-spaced fractional sampling doubles the number of computations in forming the branch metrics as compared to T-spaced sampling.

7.7.6.1 Error Probability

We now generalize the result for SISO channels and show that the T-spaced and $T/2$-spaced MIMO J-MLSE receivers for co-channel demodulation have identical performance. For T-spaced sampling, define

$$\mathbf{E}(z) \stackrel{\triangle}{=} \sum_n \boldsymbol{\varepsilon}_n z^{-n}, \tag{7.278}$$

$$\mathbf{E}_v^{(j)}(z) \stackrel{\triangle}{=} \sum_n \mathbf{v}_n^{(j)} z^{-n}. \tag{7.279}$$

Then

$$\mathbf{E}_v^{(j)}(z) = \mathbf{E}(z)\mathbf{G}^{(j)}(z) \tag{7.280}$$

and

$$\Delta^2 = \sum_m \sum_{j=1}^{J} \left\| \sum_{n=0}^{L} \mathbf{G}_n^{(j)} \boldsymbol{\varepsilon}_{n-m} \right\|^2 \tag{7.281}$$

$$= \sum_{j=1}^{J} \left[\mathbf{E}_v^{(j)}(z) \mathbf{E}_v^{(j)H}(1/z^*) \right]_0$$

$$= \sum_{j=1}^{J} \left[\mathbf{E}(z)\mathbf{G}^{(j)}(z)\mathbf{G}^{(j)^H}(1/z^*)\mathbf{E}^*(1/z^*) \right]_0$$

$$= \sum_{j=1}^{J} \left[\mathbf{E}(z)\mathbf{F}^{(j)}(z)\mathbf{E}^*(1/z^*) \right]_0. \tag{7.282}$$

For $T/2$-spaced sampling, define

$$\tilde{\mathbf{E}}(z) \stackrel{\triangle}{=} \sum_n \tilde{\varepsilon}_n z^{-2n}. \tag{7.283}$$

Since $\tilde{\varepsilon}_n = \mathbf{x}_n - \hat{\mathbf{x}}_n$ is zero for even k, we have $\tilde{\mathbf{E}}(z) = \mathbf{E}(z^2)$. Also,

$$\tilde{\mathbf{E}}_v^{(j)}(z) = \tilde{\mathbf{E}}(z)\tilde{\mathbf{G}}^{(j)}(z). \tag{7.284}$$

Therefore,

$$\tilde{\Delta}^2 = \sum_m \sum_{j=1}^{J} \left\| \sum_{n=0}^{L} \tilde{\mathbf{G}}_n^{(j)} \tilde{\varepsilon}_{n-m} \right\|^2 \tag{7.285}$$

$$= \sum_{j=1}^{J} \left[\tilde{\mathbf{E}}_v^{(j)}(z)\tilde{\mathbf{E}}_v^{(j)H}(1/z^*) \right]_0$$

$$= \sum_{j=1}^{J} \left[\tilde{\mathbf{E}}(z)\mathbf{G}^{(j)}(z)\mathbf{G}^{(j)^H}(1/z^*)\tilde{\mathbf{E}}^*(1/z^*) \right]_0$$

$$= \sum_{j=1}^{J} \left[\tilde{\mathbf{E}}(z)\mathbf{F}^{(j)}(z)\tilde{\mathbf{E}}^*(1/z^*) \right]_0$$

$$= \sum_{j=1}^{J} \left[\mathbf{E}(z^2)\mathbf{F}^{(j)}(z)\mathbf{E}^*(1/z^{*^2}) \right]_0, \tag{7.286}$$

where $[\,\cdot\,]_0$ is the coefficient of z^0. Since the odd powers of $\mathbf{E}(z^2)\mathbf{E}^*(1/z^{*^2})$ are zero and $\mathbf{F}_\ell = \tilde{\mathbf{F}}_{2\ell}$, we have $\tilde{\Delta}^2 = \Delta^2$. Therefore, the T-spaced and $T/2$-spaced receivers have identical bit error probability performance.

7.7.6.2 Timing Phase Sensitivity

The T-spaced MIMO J-MLSE receiver is sensitive to sampler timing phase due to aliasing, similar to the SISO MLSE receiver. As a result, the T-spaced receiver must accurately estimate the delays $\tau_k, k = 1, \ldots, K$. A $T/2$-spaced MIMO J-MLSE receiver will avoid aliasing (assuming that the received pulse exhibits less than

100% excess bandwidth) and is insensitive to sampler timing phase as we now show. Suppose that the timing phase offset for the kth sampler and the jth antenna branch is $t^{(k,j)}$ seconds. The $T/2$-spaced sampled pulse $f^{(k,k',j)}(t)$ at the output of the matched filter $h^{(k,j)^*}(-t-\tau_k)$ is denoted as $\tilde{f}_{\ell,t}^{(k,k',j)} = \tilde{f}^{(k,k',j)}(\ell T' \pm t^{(k,j)})$, where $T' = T/2$. Due to the timing phase offset, the ISI coefficients are not symmetric, that is, $\tilde{f}_{\ell,t}^{(k,k',j)} \neq \tilde{f}_{-\ell,t}^{(k,k',j)^*}$. Define the matrices

$$\tilde{\mathbf{F}}_n^{(j)} \stackrel{\triangle}{=} \left[\tilde{f}_n^{(k,k',j)}\right]_{K\times K}, \tag{7.287}$$

$$\tilde{\mathbf{F}}_{n,\mathbf{t}}^{(j)} \stackrel{\triangle}{=} \left[\tilde{f}_{n,t}^{(k,k',j)}\right]_{K\times K}. \tag{7.288}$$

The DTFT of $\tilde{\mathbf{F}}_{n,\mathbf{t}}^{(j)}$ is

$$\begin{aligned}\tilde{\mathbf{F}}_{\mathbf{t}}^{(j)}(\mathrm{e}^{\mathrm{j}\omega}) &= \sum_{n=-2L_j-1}^{2L_j+1} \tilde{\mathbf{F}}_{n,\mathbf{t}}^{(j)}(\mathrm{e}^{\mathrm{j}\omega})^{-n} \\ &= \mathrm{e}^{\pm j\omega\boldsymbol{\tau}^{(j)}}\tilde{\mathbf{F}}^{(j)}(\mathrm{e}^{\mathrm{j}\omega}),\end{aligned} \tag{7.289}$$

where $\mathrm{e}^{\pm\mathrm{j}\omega\boldsymbol{\tau}^{(j)}} = (\mathrm{e}^{\pm\mathrm{j}\omega\tau^{(1,j)}},\ldots,\mathrm{e}^{\pm\mathrm{j}\omega\tau^{(K,j)}})$ and $\boldsymbol{\tau}^{(k,j)} = t^{(k,j)}/T'$.

Since the noise is circularly symmetric, the psd of the noise at the output of the jth matched filter is independent of the timing offset vector $\mathbf{t}^{(j)} = (t^{(1,j)},\ldots,t^{(K,j)})$ and is given by

$$S_{\tilde{\mathbf{w}}^{(j)}\tilde{\mathbf{w}}^{(j)}}(f) = N_0\tilde{\mathbf{F}}^{(j)}(\mathrm{e}^{\mathrm{j}\omega}). \tag{7.290}$$

The DTFT of the noise-whitening matrix filter is

$$\left[\tilde{\mathbf{G}}^{(j)^H}(1/z^*)\right]^{-1}_{z=\mathrm{e}^{\mathrm{j}\omega}} = \left[\tilde{\mathbf{G}}^{(j)^H}(\mathrm{e}^{\mathrm{j}\omega})\right]^{-1}, \tag{7.291}$$

and we have

$$\tilde{\mathbf{F}}^{(j)}(\mathrm{e}^{\mathrm{j}\omega}) = \tilde{\mathbf{G}}^{(j)}(\mathrm{e}^{\mathrm{j}\omega})\tilde{\mathbf{G}}^{(j)^H}(\mathrm{e}^{\mathrm{j}\omega}) = |\tilde{\mathbf{G}}^{(j)^H}(\mathrm{e}^{\mathrm{j}\omega})|^2. \tag{7.292}$$

Hence, the noise at the output of the matrix noise-whitening filter is white. The DTFT of the message vector at the output of the noise-whitening filter is

$$\tilde{\mathbf{x}}(\mathrm{e}^{\mathrm{j}\omega})\tilde{\mathbf{G}}_{\mathbf{t}}^{(j)}(\mathrm{e}^{\mathrm{j}\omega}) = \tilde{\mathbf{x}}(\mathrm{e}^{\mathrm{j}\omega})\mathrm{e}^{\pm j\omega\boldsymbol{\tau}^{(j)}}\tilde{\mathbf{G}}^{(j)}(\mathrm{e}^{\mathrm{j}\omega}), \tag{7.293}$$

and we have

$$\sum_n \tilde{\mathbf{G}}_n^{(j)}\tilde{\mathbf{G}}_n^{(j)^H} = \sum_n \tilde{\mathbf{G}}_{n,\mathbf{t}}^{(j)}\tilde{\mathbf{G}}_{n,\mathbf{t}}^{(j)^H} = \frac{1}{2\pi}\int_{-\pi}^{\pi} |\tilde{\mathbf{G}}^{(j)}(\mathrm{e}^{\mathrm{j}\omega})|^2 \mathrm{d}\omega \tag{7.294}$$

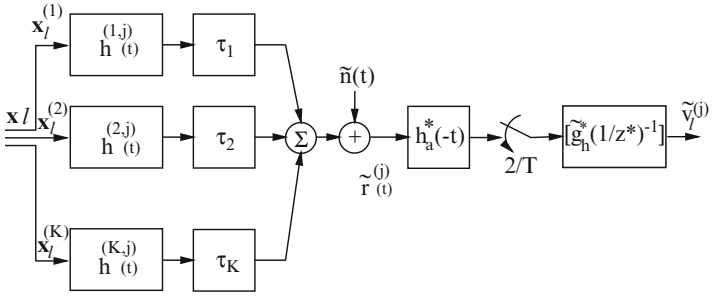

Fig. 7.21 Practical MIMO system with $T/2$-spaced sampling

by Parseval's theorem. Also

$$\tilde{\mathbf{G}}_{\mathbf{t}}^{(j)}(z)\tilde{\mathbf{G}}_{\mathbf{t}}^{(j)^H}(1/z^*) = \tilde{\mathbf{F}}^{(j)}(z). \tag{7.295}$$

It follows that the distances between allowed sequences of channel outputs in (7.286) remains the same and the $T/2$-spaced J-MLSE receiver is not sensitive to the sampler phase $e^{\pm j\omega\tau^{(j)}}$. The J-MLSE receiver just needs to estimate the $\{\tilde{\mathbf{G}}_{n,\mathbf{t}}^{(j)}\}$ rather than the $\{\tilde{\mathbf{G}}_n^{(j)}\}$ to calculate the branch metrics in the Viterbi algorithm. Finally, we note that (7.294) does not hold for the T-spaced receiver due to aliasing of the signal spectrum.

7.7.6.3 Practical Receiver

Section 7.6 showed that the optimal front-end processing for a SISO ISI channel can be realized by a receiver filter that is matched to the transmitted pulse $h_{\mathrm{a}}(t)$ followed by a rate-$2/T$ sampler and a noise-whitening filter. Here we generalize this concept to MIMO ISI channels. For an MIMO system where all co-channel waveforms are transmitted with the same pulse shape $h_{\mathrm{a}}(t)$, a significant complexity reduction is realized using this receiver since a matched filter bank is no longer required at each antenna element. As shown in Fig. 7.21, the receiver simply consists of a single matched filter for each antenna element followed by a rate-$2/T$ sampler and a noise-whitening filter. Although the T-spaced samples at the output of the filter $h_{\mathrm{a}}^*(-t)$ are white since the pulse $p(t) = h_{\mathrm{a}}(t) * h_{\mathrm{a}}(-t)$ satisfies the first Nyquist criterion, the $T/2$-spaced samples are not and, therefore, the noise-whitening filter is necessary. However, the structure of the noise-whitening filter is completely known because it only depends on the known filter $h_{\mathrm{a}}^*(-t)$.

We now establish that the systems shown in Figs. 7.19 and 7.21 yield identical performance. Assuming that rate-$2/T$ sampling satisfies the sampling theorem, the two systems can be completely represented by their $T/2$-spaced discrete-time signals. This is achieved, for example, using root raised cosine pulse shaping with

less than 100% excess bandwidth. Once again, assume that the channels are time-invariant and have impulse responses $g^{(k,j)}(t), k = 1, \ldots K$. Define

$$\mathbf{g}^{(j)}(t) \stackrel{\triangle}{=} (g^{(1,j)}(t), \ldots, g^{(K,j)}(t))^{\mathrm{T}}, \tag{7.296}$$

$$\mathbf{h}^{(j)}(t) \stackrel{\triangle}{=} (h^{(1,j)}(t), \ldots, h^{(K,j)}(t))^{\mathrm{T}}, \tag{7.297}$$

where $h^{(k,j)}(t) = h_{\mathrm{a}}(t) * g^{(k,j)}(t)$. Let $\tilde{H}_a(z)$, $\tilde{\mathbf{G}}^{(j)}(z)$, and $\tilde{\mathbf{H}}^{(j)}(z)$ be the z-transforms of the $T/2$-spaced samples of the waveforms $h_{\mathrm{a}}(t)$, and waveform vectors $\mathbf{g}^{(j)}(t)$ and $\mathbf{h}^{(j)}(t)$, respectively, where the tilde indicates $T/2$-spaced sampling. The z-transform of the autocorrelation function of the noise samples at the output of the receive filter $h_{\mathrm{a}}^*(-t)$ is $N_{\mathrm{o}}\tilde{F}_h(z)$, where $\tilde{F}_h(z) = \tilde{H}_a(z)\tilde{H}_a^*(1/z^*)$. Using the factorization

$$\tilde{F}_h(z) = \tilde{G}_h(z)\tilde{G}_h^*(1/z^*), \tag{7.298}$$

the $T/2$-spaced noise sequence at the output of the matched filter $h_{\mathrm{a}}^*(-t)$ can be whitened using a filter having the transfer function $1/\tilde{G}_h^*(1/z^*)$ as shown in Fig. 7.21. Note that the noise-whitening filter is not a matrix filter, but just a scalar filter.

The z-transform of the overall $T/2$-spaced discrete-time channel in Fig. 7.21 that includes the noise-whitening filter is

$$\begin{aligned}\tilde{\mathbf{G}}_{\mathrm{eq}}^{(j)}(z) &= \tilde{H}_a(z)\tilde{\mathbf{G}}^{(j)}(z)\tilde{H}_a^*(1/z^*)/\tilde{G}_h^*(1/z^*) \\ &= \tilde{\mathbf{G}}^{(j)}(z)\tilde{G}_h(z).\end{aligned} \tag{7.299}$$

On the other hand, referring to the conventional system shown in Fig. 7.19, we have

$$\tilde{\mathbf{H}}^{(j)}(z) = \tilde{H}_a(z)\tilde{\mathbf{G}}^{(j)}(z) \tag{7.300}$$

and

$$\begin{aligned}\tilde{\mathbf{F}}^{(j)}(z) &= \tilde{H}_a(z)\tilde{\mathbf{G}}^{(j)}(z)(\tilde{H}_a(1/z^*)\mathbf{G}^{(j)}(1/z^*))^H \\ &= \tilde{H}_a(z)\tilde{\mathbf{G}}^{(j)}(z)\tilde{\mathbf{G}}^{(j)^H}(1/z^*)\tilde{H}_a^*(1/z^*).\end{aligned} \tag{7.301}$$

Let

$$\tilde{\mathbf{G}}_c^{(j)}(z)\mathbf{G}_c^{(j)^H}(1/z^*) = \tilde{\mathbf{G}}^{(j)}(z)\tilde{\mathbf{G}}^{(j)^H}(1/z^*), \tag{7.302}$$

be a factorization of the matrix $\tilde{\mathbf{G}}^{(j)}(z)\tilde{\mathbf{G}}^{(j)^H}(1/z^*)$ such that $\tilde{\mathbf{G}}_c^{(j)}(z)$ has minimum phase. Combining (7.298), (7.301), and (7.302) gives

$$\tilde{\mathbf{F}}^{(j)}(z) = \tilde{G}_h(z)\tilde{\mathbf{G}}_c^{(j)^H}(z)\tilde{\mathbf{G}}_c^{(j)^H}(1/z^*)\tilde{G}_h^*(1/z^*). \tag{7.303}$$

The transfer function of the matrix noise-whitening filter is chosen as

$$\left[\tilde{\mathbf{G}}_c^{(j)^H}(1/z^*)\tilde{G}_h^*(1/z^*)\right]^{-1}. \tag{7.304}$$

Therefore, the overall transfer function at the output of the matrix noise-whitening filter is

$$\tilde{\mathbf{G}}^{(j)}(z) = \tilde{G}_h(z)\tilde{\mathbf{G}}_c^{(j)}(z). \tag{7.305}$$

Finally, we note that

$$\tilde{\mathbf{G}}_{\text{eq}}^{(j)}(z)\tilde{\mathbf{G}}_{\text{eq}}^{(j)^H}(1/z^*) = \tilde{\mathbf{F}}^{(j)}(z) = \tilde{\mathbf{G}}^{(j)}(z)\tilde{\mathbf{G}}^{(j)^H}(1/z^*). \tag{7.306}$$

Therefore, the Euclidean distance between sequences of noise-whitening filter outputs for the MIMO receiver in Fig. 7.21 is the same as those for the $T/2$-spaced MIMO J-MLSE receiver in Fig. 7.19. Consequently, the system shown in Fig. 7.21 achieves ML performance. The main advantage of the system in Fig. 7.21 is that the noise-whitening filter does not depend on the unknown channel and has a fixed structure.

The receiver shown in Fig. 7.21 has a scalar output, while the receiver in Fig. 7.19 has a vector output, and furthermore, $\tilde{\mathbf{G}}_{\text{eq},n}^{(j)}$ is a vector while $\tilde{\mathbf{G}}_n^{(j)}$ is a matrix. As a result, the branch metric used in the Viterbi algorithm needs to be modified accordingly. From (7.277)

$$\mu_m = -\sum_{j=1}^{J}\left(\left|\tilde{v}_{2m}^{(j)} - \sum_{n=0}^{L_j}\tilde{\mathbf{G}}_{\text{eq},2n}^{(j)^{\text{T}}}\mathbf{x}_{m-n}\right|^2 + \left|\tilde{v}_{2m+1}^{(j)} - \sum_{n=0}^{L_j-1}\tilde{\mathbf{G}}_{\text{eq},2n+1}^{(j)^{\text{T}}}\mathbf{x}_{m-n}\right|^2\right). \tag{7.307}$$

Although the $T/2$-spaced receiver is optimum, there are several key issues that must be resolved before it can be implemented. First, the receiver must be trained to derive an initial estimate of the channel vectors $\{\tilde{\mathbf{G}}_{\text{eq},n}^{(j)}\}$. This synchronization and training problem is particularly challenging for an asynchronous co-channel waveforms where the training sequences are not coincident. Second, the receiver must be able to track the channel vectors $\{\tilde{\mathbf{G}}_{\text{eq},n}^{(j)}\}$ during data demodulation if necessary. Perhaps, a per-survivor processing approach such as the one suggested in Sect. 7.4.1.1 could be used for such cases.

7.7.7 Interference Rejection Combining MLSE Receiver

In many cases, the structure of the co-channel interferers is often unknown. In other cases, it is only necessary to recover the data from a single desired co-channel user. Bottomley and Jamal [39] developed a scheme called the IRC-MLSE that combines adaptive antenna arrays and MLSE equalization. Co-channel interference

cancelation is performed in the Viterbi metric, and the receiver is equivalent to Winter's optimum linear combiner [283, 285] under flat fading channel conditions.

Once again, we assume that the receiver filter on each antenna element is matched to the transmitted pulse and followed by a rate $2/T$ sampler. Since the co-channel waveforms are assumed to have an unknown structure, a matched filter is only required for the desired signal. The overall pulse response consisting of the transmit filter, channel, and receiver filter is $f^{(j)}(t) = h_{\mathrm{a}}(t) * g^{(1,j)}(t) * h_{\mathrm{a}}^{*}(-t)$. The vector of matched filter outputs from the J receiver antenna elements is

$$\tilde{\mathbf{y}}(t) = \sum_{\ell=0}^{L} x_{\ell}\mathbf{f}(t-\ell T) + \tilde{\mathbf{z}}(t), \tag{7.308}$$

where

$$\begin{aligned}
\tilde{\mathbf{y}}(t) &= \left(y^{(1)}(t), \ldots, y^{(J)}(t)\right)^{\mathrm{T}}, \\
\mathbf{f}(t) &= \left(f^{(1)}(t), \ldots, f^{(J)}(t)\right)^{\mathrm{T}}, \\
\tilde{\mathbf{z}}(t) &= \left(z^{(1)}(t), \ldots, z^{(J)}(t)\right)^{\mathrm{T}}
\end{aligned} \tag{7.309}$$

and where LT is the length of the pulse $f^{(j)}(t)$. The impairment vector $\tilde{\mathbf{z}}(t)$ at the output of the matched filter is due to the $K-1$ interfering co-channel signals plus AWGN, and has the form

$$\tilde{\mathbf{z}}(t) = \sum_{k=1}^{K-1} \tilde{\mathbf{I}}_k(t) + \tilde{\nu}(t), \tag{7.310}$$

where

$$\begin{aligned}
\tilde{\mathbf{I}}_k(t) &= \left(I_k^{(1)}(t), \ldots, I_k^{(J)}(t)\right)^{\mathrm{T}}, \\
\tilde{\nu}(t) &= \left(\nu^{(1)}(t), \ldots, \nu^{(J)}(t)\right)^{\mathrm{T}}.
\end{aligned} \tag{7.311}$$

The matched filter outputs are sampled at rate $2/T$ and passed to a noise-whitening filter. The noise-whitening filter is suboptimum in the presence of co-channel interference, since the co-channel interference at the input to the receiver filter can be viewed as colored noise. However, the noise-whitening filter ensures maximum likelihood performance in the absence of co-channel interference. The noise-whitening filter is obtained using the same procedure leading to the overall $T/2$-spaced discrete-time channel with the transfer function defined in (7.299). It follows that the overall channel consisting of the transmit filter, channel, receiver filter, $T/2$-spaced sampler, and noise-whitening filter can be modeled as a $T/2$-spaced tapped delay line with tap coefficients

$$\tilde{\mathbf{g}}^{(j)} = \left(\tilde{g}_0^{(j)}, \ldots, \tilde{g}_{2L}^{(j)}\right)^{\mathrm{T}}.$$

Define

$$\tilde{\mathbf{g}}_\ell \stackrel{\triangle}{=} \left(\tilde{g}_\ell^{(1)}, \ldots \tilde{g}_\ell^{(J)}\right)^{\mathrm{T}}. \tag{7.312}$$

Then the vectors $\tilde{\mathbf{v}}_{2k} = (\tilde{v}_{2k}^{(1)}, \ldots, \tilde{v}_{2k}^{(J)})^{\mathrm{T}}$ and $\tilde{\mathbf{v}}_{2k+1} = (\tilde{v}_{2k+1}^{(1)}, \ldots, \tilde{v}_{2k+1}^{(J)})^{\mathrm{T}}$ at the output of the noise-whitening filter corresponding to the kth received baud are

$$\tilde{\mathbf{v}}_{2k} = \sum_{\ell=0}^{L_h} x(k-\ell)\tilde{\mathbf{g}}_{2\ell} + \tilde{\mathbf{n}}_{2k}, \tag{7.313}$$

$$\tilde{\mathbf{v}}_{2k+1} = \sum_{\ell=0}^{L_h-1} x(k-\ell)\tilde{\mathbf{g}}_{2\ell+1} + \tilde{\mathbf{n}}_{2k+1}, \tag{7.314}$$

where $\tilde{\mathbf{n}}_{2k} = (\tilde{n}_{2k}^{(1)}, \ldots, \tilde{n}_{2k}^{(J)})^{\mathrm{T}}$ and $\tilde{\mathbf{n}}_{2k+1} = (\tilde{n}_{2k}^{(1)}, \ldots, \tilde{n}_{2k}^{(J)})^{\mathrm{T}}$. Once again, we emphasize that the noise vectors $\tilde{\mathbf{n}}_{2k}$ and $\tilde{\mathbf{n}}_{2k+1}$ are not white due to the presence of the co-channel signals.

To derive a feasible receiver structure, we now *assume* that the sampled impairment vector $\tilde{\mathbf{n}}_\ell$ at the output of the noise-whitening filter at epoch ℓ is a vector of J correlated zero-mean complex Gaussian random variables having the joint pdf

$$p(\tilde{\mathbf{n}}_\ell) = \frac{1}{(2\pi)^J |\mathbf{R}_\ell|} \exp\left\{-\frac{1}{2}\tilde{\mathbf{n}}_\ell^H \mathbf{R}_\ell^{-1} \tilde{\mathbf{n}}_\ell\right\}, \tag{7.315}$$

where $|\mathbf{R}_\ell|$ is the determinant of $\mathbf{R}_\ell$ and

$$\mathbf{R}_\ell = \frac{1}{2}\mathrm{E}[\tilde{\mathbf{n}}_\ell \tilde{\mathbf{n}}_\ell^H] \tag{7.316}$$

is the impairment correlation matrix. Assuming an MLSE-like algorithm, the branch metric should be related to the likelihood of the impairment vector. For each state transition $\rho_k^{(j)} \rightarrow \rho_{k+1}^{(i)}$ at epoch k, the samples $\tilde{\mathbf{v}}_{2k}$ and $\tilde{\mathbf{v}}_{2k+1}$ are used by the Viterbi algorithm to evaluate the branch metric

$$\begin{aligned}\mu_k = {} & (\tilde{\mathbf{v}}_{2k} - \hat{\mathbf{u}}_{2k}^1)^H \mathbf{R}_{2k}^{-1} (\tilde{\mathbf{v}}_{2k} - \hat{\mathbf{u}}_{2k}^1) \\ & + (\tilde{\mathbf{v}}_{2k+1} - \hat{\mathbf{u}}_{2k+1}^2)^H \mathbf{R}_{2k+1}^{-1} (\tilde{\mathbf{v}}_{2k+1} - \hat{\mathbf{u}}_{2k+1}^2),\end{aligned} \tag{7.317}$$

where

$$\hat{\mathbf{u}}_{2k}^1 = \sum_{\ell=0}^{L_h} \hat{x}(k-\ell)\tilde{\mathbf{g}}_{2\ell}, \tag{7.318}$$

$$\hat{\mathbf{u}}_{2k+1}^2 = \sum_{\ell=0}^{L_h-1} \hat{x}(k-\ell)\tilde{\mathbf{g}}_{2\ell+1}. \tag{7.319}$$

Notice that the metric calculation requires the impairment correlation matrix $\mathbf{R}_k$ and its inverse, and the channel impulse vectors

$$\tilde{\mathbf{g}}^1 = (\tilde{\mathbf{g}}_0, \tilde{\mathbf{g}}_2, \ldots, \tilde{\mathbf{g}}_{2L_h})^{\mathrm{T}} \tag{7.320}$$

$$\tilde{\mathbf{g}}^2 = (\tilde{\mathbf{g}}_1, \tilde{\mathbf{g}}_3, \ldots, \tilde{\mathbf{g}}_{2L_h-1})^{\mathrm{T}}. \tag{7.321}$$

Computing the inverse of the $J \times J$ matrix $\mathbf{R}_k$ can be computationally intensive for large J, the number of computations required being proportional to J^3. However, when $J = 2$ (two receiver antenna elements) the inverse can be obtained using direct matrix inversion (DMI), that is, the inverse of the matrix $\mathbf{R}_k$ is

$$\mathbf{R}_k^{-1} = \frac{\mathrm{adj}(\mathbf{R}_k)}{|\mathbf{R}_k|} = \frac{1}{r_{k_{11}} r_{k_{22}} - r_{k_{12}} r_{k_{21}}} \begin{bmatrix} r_{k_{22}} & -r_{k_{12}} \\ -r_{k_{21}} & r_{k_{11}} \end{bmatrix}. \tag{7.322}$$

Division by the determinant $|\mathbf{R}_k|$ is unnecessary, provided that $\mathbf{R}_k$ remains constant over the decision delay in the Viterbi algorithm, since the determinant will just scale all the path metrics. For such cases, the Viterbi algorithm can use the simplified branch metric

$$\begin{aligned} \mu_k = {} & (\tilde{\mathbf{v}}_{2k} - \hat{\mathbf{u}}_{2k}^1)^H \mathrm{adj}(\mathbf{R}_{2k})(\tilde{\mathbf{v}}_{2k} - \hat{\mathbf{u}}_{2k}^1) \\ & + (\tilde{\mathbf{v}}_{2k+1} - \hat{\mathbf{u}}_{2k+1}^2)^H \mathrm{adj}\mathbf{R}_{2k+1})(\mathbf{v}_{2k+1} - \hat{\mathbf{u}}_{2k+1}^2), \end{aligned} \tag{7.323}$$

which only requires multiplications and additions.

Finally, a metric combining MLSE (MC-MLSE) receiver is one that zeroes the off diagonal elements of the matrix $\mathbf{R}_k$. The metric combining receiver is equivalent to maximal ratio combining when the additive impairment is white Gaussian noise.

7.7.8 Examples

The performances of the J-MLSE, IRC-MLSE, and MC-MLSE receivers discussed in the previous sections are now compared and contrasted. For this purpose, an EDGE burst format is assumed. The EGDE burst format is the same as the GSM burst format described in Fig. 1.2. However, instead of the GMSK modulation used in GSM, EDGE uses eight-PSK modulation with square-root raised cosine pulse shaping with a roll-off factor of $\beta = 0.5$. For illustrative purposes, a T-spaced, two equal ray, model is assumed for the desired signal. The interference impairment consists of a single flat faded EDGE interferer. In all cases, the receiver front-end consists of a receiver filter that is matched with the transmitted pulse followed by a rate-$2/T$ sampler and a noise-whitening filter. The J-MLSE receiver has 512 states, as defined by two symbols for the desired signal and one symbol for the co-channel interferer. The MC/IRC-MLSE receivers have 64 states, as defined by two symbols for the desired signal. Each simulation run consists of 3,000 frames of 142 eight-PSK symbols.

Figure 7.22 shows the $E_\mathrm{b}/N_\mathrm{o}$ performance of the three receivers for a fixed carrier-to-interference ratio $\mathrm{C/I} = 30\,\mathrm{dB}$. Under this condition, the interference

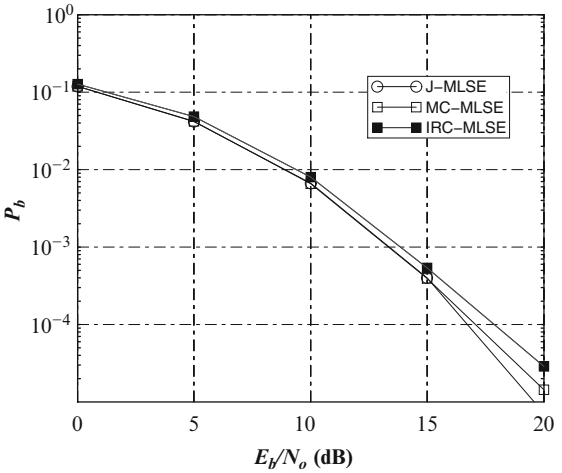

Fig. 7.22 Relative E_b/N_o performance of the J-MLSE, MC-MLSE and IRC-MLSE receivers; $C/I = 30\,dB$

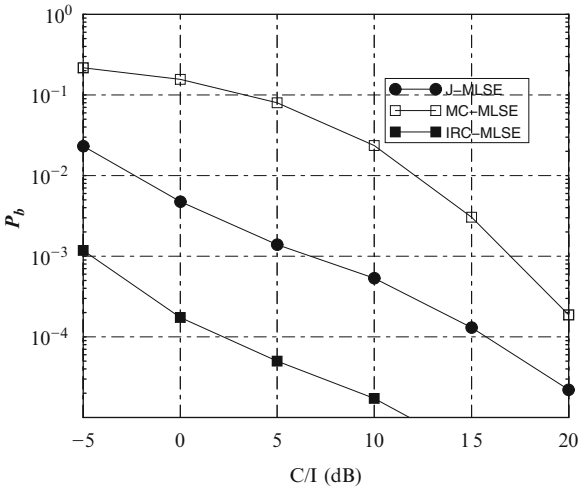

Fig. 7.23 Relative C/I performance of the J-MLSE, MC-MLSE and IRC-MLSE receivers; $E_b/N_o = 30\,dB$

is very light. Perfect channel knowledge is assumed. The J-MLSE receiver is the optimum receiver in the maximum likelihood sense and achieves the best possible performance in AWGN. The MC-MLSE receiver is also optimum for AWGN channels, but exhibits some degradation at higher E_b/N_o due to the co-channel interference that is present. The IRC-MLSE receiver gives the worst E_b/N_o performance.

Figure 7.23 shows the C/I performance of the three receivers for $E_b/N_o = 30\,dB$. Observe that the MC-MLSE receiver gives the worst performance, while the

J-MLSE receiver and IRC-MLSE receivers offer huge C/I performance gains. The best performance is realized with the IRC-MLSE receiver. Hence, the IRC-MLSE receiver sacrifices a small amount of E_b/N_o performance for a large gain in C/I performance.

It is curious that the IRC-MLSE receiver outperforms the J-MLSE receiver. First, the J-MLSE receiver that we have implemented does not have a sufficient number of receiver states due to pulse truncation effects. Hence, there is some residual ISI that is significant at low C/I. Second, the overall signal constellation produced by the combination of the desired signal and the co-channel signal may degenerate when the signal constellation points overlap. In this case, errors can occur even for large E_b/N_o values.

Appendix 1: Derivation of (7.202)

Consider the case where $D = 1$ and suppose that the characteristic function in (7.194) has $L+1$ different poles $\bar{\alpha} = (\bar{\alpha}_0, \bar{\alpha}_1, \ldots, \bar{\alpha}_L)$. Then the pairwise error probability is equal to

$$P(\bar{\alpha}) = \sum_{i=0}^{L} \left(\left(\frac{1}{2} - \frac{1}{2}\sqrt{\frac{\bar{\alpha}_i}{1+\bar{\alpha}_i}} \right) \prod_{j \neq i} \left(1 - \frac{\bar{\alpha}_j}{\bar{\alpha}_i} \right)^{-1} \right). \tag{7.324}$$

Define the function $\phi(\bar{\alpha}) = \sum_{i=0}^{L} \bar{\alpha}_i - C = 0$, where C is a constant. The method of Lagrange multipliers suggests that

$$\frac{\partial P(\bar{\alpha})}{\partial \bar{\alpha}_i} + \lambda \frac{\partial \phi}{\partial \bar{\alpha}_i} = 0 \quad i = 1, \ldots, L \tag{7.325}$$

for any real number λ. It can be shown by induction that

$$\begin{aligned} \frac{\partial P(\bar{\alpha})}{\partial \bar{\alpha}_k} = &- \left(\frac{1}{2} - \frac{1}{2}\sqrt{\frac{\bar{\alpha}_k}{1+\bar{\alpha}_k}} \right) \sum_{i \neq k} \left(\frac{\bar{\alpha}_i}{\bar{\alpha}_k^2} \left(1 - \frac{\bar{\alpha}_i}{\bar{\alpha}_k} \right)^{-2} \prod_{j \neq i,k} \left(1 - \frac{\bar{\alpha}_j}{\bar{\alpha}_k} \right)^{-1} \right) \\ &+ \sum_{i \neq k} \left(\frac{1}{\bar{\alpha}_i} \left(1 - \frac{\bar{\alpha}_k}{\bar{\alpha}_i} \right)^{-2} \left(\frac{1}{2} - \frac{1}{2}\sqrt{\frac{\bar{\alpha}_i}{1+\bar{\alpha}_i}} \right) \prod_{j \neq i,k} \left(1 - \frac{\bar{\alpha}_j}{\bar{\alpha}_i} \right)^{-1} \right) \\ &- \left(\frac{1}{4\bar{\alpha}_k^{1/2}} \frac{1}{(1+\bar{\alpha}_k)^{3/2}} \right) \prod_{j \neq k} \left(1 - \frac{\bar{\alpha}_j}{\bar{\alpha}_k} \right)^{-1}. \end{aligned} \tag{7.326}$$

By solving (7.325) and observing the symmetry of $P(\bar{\alpha})$ and the derivative (7.326) with respect to the permutations of $\bar{\alpha}$, it is apparent that the minimum of $P(\bar{\alpha})$ is achieved when $\bar{\alpha}_0 = \bar{\alpha}_1 = \cdots = \bar{\alpha}_L$.

Problems

7.1. Starting with

$$f_k = \int_{-\infty}^{\infty} h^*(\tau)h(\tau + kT)\mathrm{d}\tau,$$

show that

$$F(\mathrm{e}^{\mathrm{j}2\pi fT}) = F_\Sigma(f).$$

7.2. Suppose that the impulse response of an overall channel consisting of the transmit filter, channel, and receive filter is

$$F(f) = \begin{cases} 1, & |f| \le f_\ell \\ \frac{f_u - |f|}{f_u - f_\ell}, & f_\ell \le |f| \le f_u \end{cases}.$$

(a) Find the overall impulse response $f(t)$.
(b) Is it possible to transmit data without ISI?
(c) How do the magnitudes of the tails of the overall impulse response decay with large values of t?
(d) Suppose that binary signaling is used with this pulse shape so that the noiseless signal at the output of the receive filter is

$$y(t) = \sum_n x_\mathrm{n} f(t - nT),$$

where $x_\mathrm{n} \in \{-1, +1\}$. What is the maximum possible magnitude that $y(t)$ can achieve?

7.3. Suppose a digital communication system operates over an "ideal channel" and uses an overall pulse $p(t)$ that has the Gaussian-shaped form

$$p(t) = \mathrm{e}^{-\pi a^2 t^2}.$$

(a) Explain why $p(t)$ does not admit ISI-free transmission.
(b) To reduce the level of ISI to a relatively small amount, we impose the condition that $p(T) = 0.01$, where T is the symbol interval. The bandwidth W of the pulse $p(t)$ is defined as that value of W for which $P(W)/P(0) = 0.01$, where $P(f)$ is the Fourier transform of $p(t)$. Determine the value W and compare this value to that of a raised cosine spectrum with 100% roll-off.

7.4. Show that the ISI coefficients $\{f_n\}$ may be expressed in terms of the channel vector coefficients $\{g_n\}$ as

$$f_n = \sum_{k=0}^{L-n} g_k^* g_{k+n} \qquad n = 0, 1, 2, \ldots, L.$$

7.5. Suppose that BPSK is used on a static ISI channel. The complex envelope has the form

$$\tilde{s}(t) = A \sum_{k=-\infty}^{\infty} x_k h_a(t - kT),$$

where $x_k \in \{-1, +1\}$ and $h_a(t)$ is the amplitude shaping pulse. The full response rectangular pulse $h_a(t) = u_T(t)$ is used and the impulse response of the channel is

$$g(t) = g_0 \delta(t) - g_1 \delta(t - \tau),$$

where g_0 and g_1 are complex numbers and $0 < \tau < T$.

(a) Find the received pulse $h(t)$.
(b) What is the filter matched to $h(t)$?
(c) What are the ISI coefficients $\{f_i\}$?

7.6. Suppose that BPSK signaling is used on a static ISI channel having impulse response

$$g(t) = \delta(t) + 0.1\delta(t - T).$$

The receiver employs a filter that is matched to the transmitted pulse $h_a(t)$, and the sampled outputs of the matched filter are

$$y_k = x_k q_0 + \sum_{n \neq k} x_n q_{k-n} + \nu_k,$$

where $x_n \in \{-1, +1\}$. Decisions are made on the $\{y_k\}$ without any equalization.

(a) What is the variance of noise term ν_k?
(b) What are the values of the $\{q_n\}$?
(c) What is the probability of error in terms of the average received bit-energy-to-noise ratio?

7.7. A typical receiver for digital signaling on an ISI channel consists of a matched filter followed by an equalizer. The matched filter is designed to minimize the effect of random noise, while the equalizer is designed to minimize the effect of ISI. Using mathematical arguments, show that (1) the matched filter tends to accentuate the effect of ISI, and (2) the equalizer tends to accentuate the effect of random noise.

7.8. Consider an ISI channel, where $f_n = 0$ for $|n| > 1$. Suppose that the receiver uses a filter matched to the received pulse $h(t) = h_a(t) * g(t)$, and the T-spaced samples at the output of the matched filter, $\{y_k\}$, are filtered as shown in Fig. 7.24. The values of g_0 and g_1 are chosen to satisfy

$$\begin{aligned} |g_0|^2 + |g_1|^2 &= f_0, \\ g_0 g_1^* &= f_1. \end{aligned}$$

Find an expression for the filter output v_k in terms of g_0, g_1, x_k, x_{k-1}, and the noise component at the output of the digital filter, w_k.

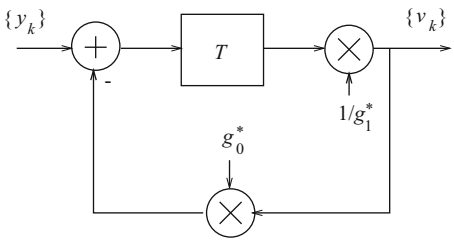

Fig. 7.24 Digital filter for Problem 7.8

7.9. The z-transform of the channel vector $\mathbf{g}$ of a communication system is equal to

$$G(z) = 0.1 + 1.0z^{-1} - 0.1z^{-2}.$$

A binary sequence $\mathbf{x}$ is transmitted, where $x_n = \in \{-1, +1\}$. The received samples at the output of the noise whitening filter are

$$v_k = \sum_{n=0}^{2} g_n x_{k-n} + w_k,$$

where $\{w_k\}$ is a white Gaussian noise sequence with variance $\sigma_w^2 = N_o$.

a) Evaluate the probability of error if the demodulator ignores ISI.
b) Design a 3-tap zero-forcing equalizer for this system.
c) What is the response $\{v_k\}$ for the input sequence

$$\{x_n\} = (-1)^k,\ k = 0, 1, 2, 3?$$

What is the response at the output of the equalizer?
d) Evaluate the probability of error for the equalized channel.

7.10. Suppose that a system is characterized by the received pulse

$$h(t) = \sqrt{2a}\mathrm{e}^{-at}, \qquad 0 \le t \le \infty.$$

A receiver implements a filter matched to $h(t)$ and generates T-spaced samples at the output of the filter. Note that the matched filter is actually noncausal.

(a) Find the ISI coefficients f_i.
(b) What is the transfer function of the noise-whitening filter that yields a system having an overall minimum phase response?
(c) Find the transfer function of the equivalent zero-forcing equalizer $C'(z)$.
(d) Find the noise power at the output of the zero-forcing equalizer, and find the condition when the noise power becomes infinite.

7.11. Consider M-PAM on a static ISI channel, where the receiver uses a filter that is matched to the received pulse. The sampled outputs of the matched filter are

$$y_n = x_n f_0 + \sum_{k \neq n} x_k f_{n-k} + \nu_n,$$

where the source symbols are from the set $\{\pm 1, \pm 3, \ldots, \pm(M-1)\}$. Decisions are made on the $\{y_n\}$ without any equalization using a threshold detector. The ℓth ISI pattern can be written as

$$D(\ell) = \sum_{k \neq n} x_{\ell,k} f_{n-k},$$

and $D(\ell)$ is maximum when $\text{sgn}(x_{\ell,k}) = \text{sgn}(f_{n-k})$ and each of the $x_{\ell,k}$ takes on the maximum signaling level, that is, $x_{\ell,k} = (M-1)d$ for M even. The *maximum distortion* is defined as

$$D_{\max} = \frac{1}{f_0} \sum_{n \neq 0} |f_n|.$$

(a) Discuss and compare error performance M-ary signaling ($M > 2$) with binary signaling ($M = 2$), using $D_{\max}$ as a parameter.
(b) Suppose that the channel has ISI coefficients

$$\begin{aligned} f_i &= 0.0, \quad |i| \geq 3, \\ f_2 &= f_{-2} = 0.1, \\ f_1 &= f_{-1} = -0.2, \\ f_0 &= 1.0. \end{aligned}$$

Plot the probability of error against the signal-to-noise ratio and compare with the ideal channel case, that is, $f_0 = \delta_{n0}$. Show your results for $M = 2$ and 4.

7.12. Consider a linear MSE equalizer and suppose that the tap gain vector $\mathbf{c}$ satisfies

$$\mathbf{c} = \mathbf{c}_{\text{op}} + \mathbf{c}_{\text{e}},$$

where $\mathbf{c}_{\text{e}}$ is the tap gain error vector. Show that the MSE that is achieved with the tap gain vector $\mathbf{c}$ is

$$J = J_{\min} + \mathbf{c}_{\text{e}}^{\text{T}} \mathbf{M}_v \mathbf{c}_{\text{e}}^{*}.$$

7.13. The matrix $\mathbf{M}_v$ has an eigenvalue λ_k and eigenvector $\mathbf{x}_k$ if

$$\mathbf{x}_k \mathbf{M}_v = \lambda_k \mathbf{x}_k, \quad k = 1, \ldots, N.$$

Prove that the eigenvectors are orthogonal, that is, $\mathbf{x}_i \mathbf{x}_j^T = \delta_{ij}$.

7.14. Show that the relationship between the output SNR and $J_{\min}$ for an infinite-tap mean-square error linear equalizer is

$$\gamma_\infty = \frac{1 - J_{\min}}{J_{\min}},$$

where γ_∞ indicates that the equalizer has an infinite number of taps. Note that this relationship between γ_∞ and $J_{\min}$ holds when there is residual ISI in addition to the noise.

7.15. In this question, we will show in steps that

$$\nabla_{\mathbf{c}} J = 2\mathbf{c}^T \mathbf{M}_v - 2\mathbf{v}_x^H.$$

Define

$$\begin{aligned}
\mathbf{M}_v &\triangleq \mathbf{M}_{v_R} + j\mathbf{M}_{v_I}, \\
\mathbf{c} &\triangleq \mathbf{c}_R + j\mathbf{c}_I, \\
\mathbf{v}_x &\triangleq \mathbf{v}_{x_R} + j\mathbf{v}_{x_I}.
\end{aligned}$$

(a) Using the Hermitian property $\mathbf{M}_v = \mathbf{M}_v^H$, show that

$$\mathbf{M}_{v_R} = \mathbf{M}_{v_R}^T \text{ and } \mathbf{M}_{v_I} = -\mathbf{M}_{v_I}^{\mathrm{T}}.$$

(b) Show that

$$\begin{aligned}
\nabla_{\mathbf{c}_R} \mathrm{Re}\{\mathbf{v}_x^H \mathbf{c}^*\} &= \mathbf{v}_{x_R}^{\mathrm{T}}, \\
\nabla_{\mathbf{c}_I} \mathrm{Re}\{\mathbf{v}_x^H \mathbf{c}^*\} &= -\mathbf{v}_{x_I}^{\mathrm{T}}, \\
\nabla_{\mathbf{c}_R} \mathbf{c}^T \mathbf{M}_v \mathbf{c}^* &= 2\mathbf{c}_R^T \mathbf{M}_{v_R} - 2\mathbf{c}_I^T \mathbf{M}_{v_I}, \\
\nabla_{\mathbf{c}_I} \mathbf{c}^T \mathbf{M}_v \mathbf{c}^* &= 2\mathbf{c}_I^T \mathbf{M}_{v_R} + 2\mathbf{c}_R^T \mathbf{M}_{v_I},
\end{aligned}$$

where $\nabla_{\mathbf{x}}$ is the gradient with respect to vector $\mathbf{x}$.

(c) If we define the gradient of a real-valued function with respect to a complex vector $\mathbf{c}$ as

$$\nabla_{\mathbf{c}} = \nabla_{\mathbf{c}_R} + j\nabla_{\mathbf{c}_I},$$

show that

$$\begin{aligned}
\nabla_{\mathbf{c}} \mathrm{Re}\{\mathbf{v}_x^H \mathbf{c}^*\} &= \mathbf{v}_x^H, \\
\nabla_{\mathbf{c}} \mathbf{c}^H \mathbf{M}_v \mathbf{c}^* &= 2\mathbf{c}^T \mathbf{M}_v.
\end{aligned}$$

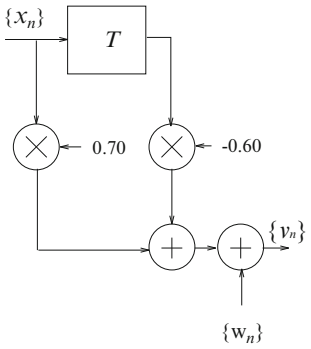

Fig. 7.25 Discrete-time white noise channel model for Problem 7.17

7.16. Show that the pairwise error probability for digital signaling on an ISI channel is given by (7.171).

7.17. Consider the transmission of the binary sequence $\mathbf{x}$, $x_n \in \{-1, +1\}$ over the equivalent discrete-time white noise channel model shown in Fig. 7.25. The received sequence is

$$\begin{aligned} v_0 &= .70x_0 + w_1 \\ v_1 &= .70x_1 - .60x_0 + w_2 \\ v_2 &= .70x_2 - .60x_1 + w_3 \\ &\vdots \\ v_k &= .70x_k - .60x_{k-1} + w_k. \end{aligned}$$

(a) Draw the state diagram for this system.
(b) Draw the trellis diagram.
(c) Suppose that the received sequence is

$$\{v_i\}_{i=0}^{6} = \{1.0, -1.5, 0.0, 1.5, 0.0, -1.5, 1.0\}.$$

Show the surviving paths and their associated path metrics after v_6 has been received.

7.18. Suppose that BPSK signaling is used on a frequency selective fading channel. The discrete-time system consisting of the transmit filter, channel, receiver filter, and baud-rate sampler can be described by the polynomial

$$F(z) = \frac{5}{16} - \frac{1}{8}z^{-1} - \frac{1}{8}z.$$

The samples at the output of the receiver filter are processed by a noise-whitening filter such that the overall discrete-time white noise channel model $G(z)$ has minimum phase.

(a) Find $G(z)$.
(b) Draw the state diagram and the trellis diagram for the discrete-time white noise channel model.
(c) A block of ten symbols $\mathbf{x} = \{x_i\}_{i=0}^{9}$ is transmitted over the channel and it is known that $x_9 = -1$. Assume that $x_i = 0, i < 0$ and the suppose that the sampled sequence at the output of the noise-whitening filter is

$$\begin{aligned}\mathbf{v} &= \{v_0, v_1, v_2, v_3, \ldots v_9\} \\ &= \{1/2, 1/4, -3/4, 3/4, -3/4, -1/4, 3/4, -3/4, -1/4, -1/4\}.\end{aligned}$$

What sequence $\mathbf{x}$ was most likely transmitted?