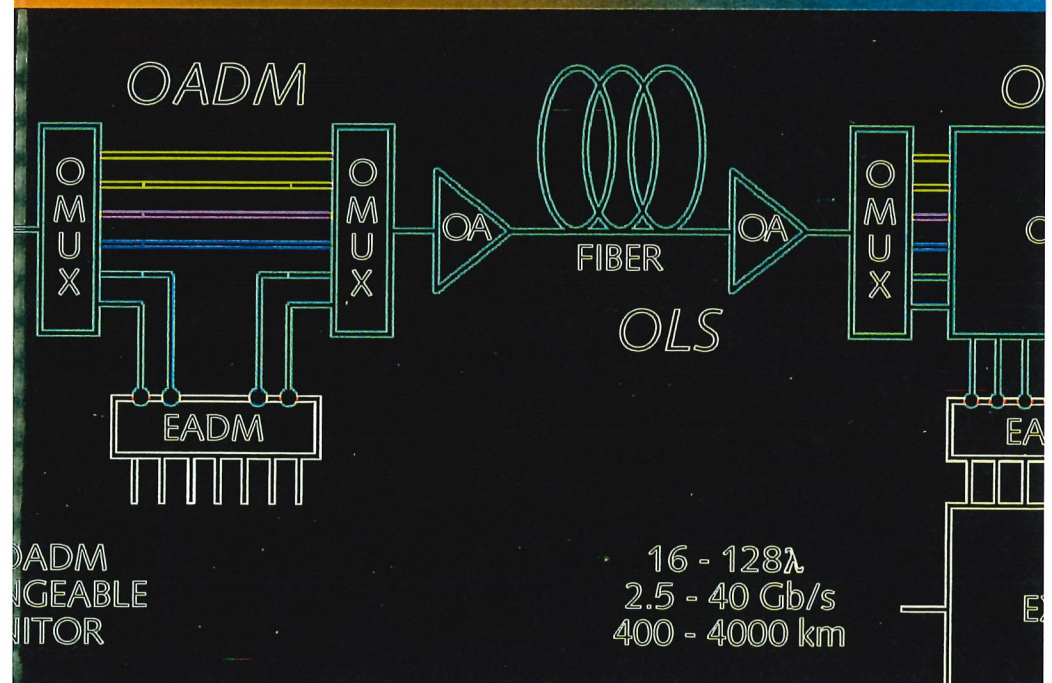


OPTICAL FIBER TELECOMMUNICATIONS SYSTEMS AND IMPAIRMENTS

Edited by

Ivan Kaminow / Tingye Liu



OPTICAL FIBER TELECOMMUNICATIONS IV B SYSTEMS AND IMPAIRMENTS

Edited by

IVAN P. KAMINOW

Bell Laboratories (retired)
Kaminow Lightwave Technology
Holmdel, New Jersey

TINGYE LI

AT&T Labs (retired)
Boulder, Colorado



ACADEMIC PRESS

An Elsevier Science Imprint

San Diego San Francisco New York
Boston London Sydney Tokyo

This book is printed on acid-free paper. (∞)

Copyright © 2002, Elsevier Science (USA). All rights reserved.

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission in writing from the publisher.

The appearance of the code at the bottom of the first page of a chapter in this book indicates the Publisher's consent that copies of the chapter may be made for personal or internal use of specific clients. This consent is given on the condition, however, that the copier pay the stated per copy fee through the Copyright Clearance Center, Inc. (222 Rosewood Drive, Danvers, Massachusetts 01923), for copying beyond that permitted by Sections 107 or 108 of the U.S. Copyright Law. This consent does not extend to other kinds of copying, such as copying for general distribution, for advertising or promotional purposes, for creating new collective works, or for resale. Copy fees for pre-2001 chapters are as shown on the title pages. If no fee code appears on the title page, the copy fee is the same as for current chapters. \$35.00.

Academic Press

An Elsevier Science Imprint

525 B Street, Suite 1900, San Diego, California 92101-4495, USA

<http://www.academicpress.com>

Academic Press

Harcourt Place, 32 Jamestown Road, London NW1 7BY, UK

<http://www.academicpress.com>

Library of Congress Control Number: 2001098830

International Standard Book Number: 0-12-395173-9

PRINTED IN CHINA

02 03 04 05 06 07 RDC 9 8 7 6 5 4 3 2 1

Contents

<i>Contributors</i>	<i>xi</i>
Chapter 1 Overview <i>Ivan P. Kaminow</i>	1
Chapter 2 Growth of the Internet <i>Kerry G. Coffman and Andrew M. Odlyzko</i>	17
Chapter 3 Optical Network Architecture Evolution <i>John Strand</i>	57
Chapter 4 Undersea Communication Systems <i>Neal S. Bergano</i>	154
Chapter 5 High-Capacity, Ultra-Long-Haul Networks <i>John Zysskind, Rick Barry, Graeme Pendock, Michael Cahill, and Jinendra Ranka</i>	198
Chapter 6 Pseudo-Linear Transmission of High-Speed TDM Signals: 40 and 160 Gb/s <i>René-Jean Essiambre, Gregory Raybon, and Benny Mikkelsen</i>	232

Chapter 7	Dispersion-Managed Solitons and Chirped Return to Zero: What Is the Difference?	305
	<i>Curtis R. Menyuk, Gary M. Carter, William L. Kath, and Ruo-Mei Mu</i>	
Chapter 8	Metropolitan Optical Networks	329
	<i>Nasir Ghani, Jin-Yi Pan, and Xin Cheng</i>	
Chapter 9	The Evolution of Cable TV Networks	404
	<i>Xiaolin Lu and Oleh Sneizko</i>	
Chapter 10	Optical Access Networks	438
	<i>Edward Harstead and Pieter H. van Heyningen</i>	
Chapter 11	Beyond Gigabit: Application and Development of High-Speed Ethernet Technology	514
	<i>Cedric F. Lam</i>	
Chapter 12	Photonic Simulation Tools	564
	<i>Arthur J. Lowery</i>	
Chapter 13	Nonlinear Optical Effects in WDM Transmission	611
	<i>Polina Bayvel and Robert Killey</i>	
Chapter 14	Fixed and Tunable Management of Fiber Chromatic Dispersion	642
	<i>Alan E. Willner and Bogdan Hoanca</i>	
Chapter 15	Polarization-Mode Dispersion	725
	<i>Herwig Kogelnik, Robert M. Jopson, and Lynn E. Nelson</i>	

Chapter 16	Bandwidth-Efficient Modulation Formats for Digital Fiber Transmission Systems	862
	<i>Jan Conradi</i>	
Chapter 17	Error-Control Coding Techniques and Applications	902
	<i>P. Vijay Kumar, Moe Z. Win, Hsiao-Feng Lu, and Costas N. Georgiades</i>	
Chapter 18	Equalization Techniques for Mitigating Transmission Impairments	965
	<i>Moe Z. Win, Jack H. Winters, and Giorgio M. Vitetta</i>	
	Index to Volumes IVA and IVB	999

Chapter 18 | Equalization Techniques for Mitigating Transmission Impairments

Moe Z. Win and Jack H. Winters

AT&T Labs-Research, Middletown, New Jersey

Giorgio M. Vitetta

University of Modena and Reggio Emilia, Modena, Italy

18.1 Introduction

Most current multi-gigabit-per-second digital fiber-optic systems use simple modulation and detection techniques such as on-off keying with matched-filter receiver techniques. However, more complex techniques such as equalization, error control coding, and/or multilevel signaling can be used in lightwave systems to significantly increase the data rate and/or reduce the effect of transmission impairments and improve performance, and, in many cases, can be easily implemented [1].

There are numerous sources of transmission impairments in lightwave systems. These include chromatic dispersion and polarization-mode dispersion (PMD) in the fiber, laser and fiber nonlinearities, nonideal receiver response, echo, and distortion caused by semiconductor optical amplifiers. Fortunately, there are also numerous techniques to reduce these impairments, including error control coding, equalization, and modulation techniques with multilevel signaling. These techniques can be used for upgrading existing systems (e.g., reducing chromatic dispersion and PMD in systems with previously installed fiber) or as an alternative to the use of more costly transmitters and receivers (e.g., using coding to permit less stringent laser specifications [2]).

The purpose of this chapter is to provide an overview of equalization techniques, which become increasingly important as device technology matures, in which case substantial increases in performance (especially for already installed systems) may only be achieved through these techniques.¹ We will examine the prospects for the application of digital equalization techniques to overcome these impairments, where the improvement is large and the implementation is relatively simple.

¹ Error control coding techniques are discussed in Chapter 17 of this volume.

18.2 Transmission Impairments in Lightwave Systems

Here we classify the transmission impairments in lightwave systems into three categories: (1) signal distortion with a single signal, (2) signal distortion between multiple signals, and (3) noise. These impairments are discussed in the following text.

Signal distortion (with a single signal) refers to those impairments that distort and broaden the width of pulses, resulting in intersymbol interference (ISI) that limits the maximum bit rate. The key feature of signal distortion is that it is deterministic (can be calculated directly from the impairment), i.e., the distortion is bit-pattern dependent with the distortion for a given fixed pattern (or slowly varying). This distortion results in a narrowing or closing of the received signal eye in the vertical direction (a pattern-dependent signal level at the detector sampling time) and/or the horizontal direction (a pattern-dependent timing jitter). The most extensively analyzed distortion is chromatic dispersion (see, e.g. [3,4]) in long-haul systems. This is material dispersion in the fiber which causes a delay in the received signal spectrum that varies with frequency. The dominant delay distortion is a linear delay of about 17 ps/km/nm at a wavelength of 1.55 μm in a standard fiber. Thus, the delay variation is linear with distance, but is fixed for a given length of fiber, i.e., it doesn't vary significantly with time. A second source of distortion is PMD in long-haul systems (see, e.g. [4–7]). PMD is generated by signal delays that are polarization dependent. These delays increase with distance and also vary slowly with time (due to temperature and other variations) [6,8–10]. At a given frequency, two orthogonal polarizations have different delays. Thus, a pulse with a sufficiently narrow frequency spectrum can be received as two pulses with a time delay between them. For a study of the combined effects of chromatic and polarization dispersion in lightwave systems, see [11]. We group laser nonlinearities and receiver bandwidth limitations as the third source of distortion. Note that this impairment is independent of distance, and varies only due to aging. A fourth possible source of dispersion is a semiconductor optical amplifier [12] (fiber optical amplifiers have negligible distortion), and a fifth source is fiber nonlinearities [13,14].

The previous impairments all increase in severity with the signal bandwidth. This bandwidth is lower bounded by the data rate, but may be much larger than this because of other factors. For a multimode laser, both mode evolution and mode hopping [15,16] increase the bandwidth of the signal to several nanometers. For an analysis of equalization techniques in multimode fiber systems, see [17]. For a single-frequency (mode) laser, the laser linewidth (due to phase noise) and chirping or relaxation oscillation (with direct modulation of the laser) increase the bandwidth of the signal. A single-frequency laser has a nonzero linewidth because of random variations in the phase of the laser (phase noise), as discussed below. Chirp is the variation in carrier frequency

of the laser caused by changes in drive-signal amplitude. Similarly, relaxation oscillation is the variation (undesired oscillation) in amplitude of the laser caused by changes in drive-signal amplitude. Note that chirp (and relaxation oscillation) can be avoided if the laser itself is used in the cw mode with the modulation being done with an external modulator.

The second impairment is interference between multiple signals in different frequency bands on the same fiber. This can be due to nonlinearities in the fiber [13,14] or in semiconductor optical amplifiers [12]. In addition, in duplex systems, echo also can degrade performance. Note that these distortions are deterministic.

The third class of transmission impairment we consider is noise, which varies randomly from bit to bit. Noise in the received signal consists of shot noise, thermal noise, and, with optical amplifiers, amplified spontaneous emission (ASE) noise. Shot noise is the quantum noise due to the fact that the received signal is actually a series of photons. The number of photons received during each symbol interval has a Poisson distribution and, therefore, the received signal level varies randomly from symbol to symbol. Thermal noise is introduced by the receiver preamplifier, and is usually assumed to be additive, white Gaussian noise. ASE is additive Gaussian noise in the optical signal that increases with the gain of the amplifiers. Although ASE is random, with direct detection the electrical signal at the receiver contains a noise times signal component, and thus the ASE noise level in the received electrical signal is signal-level dependent. Without optical amplifiers, thermal noise is the major limitation with direct detection, whereas shot noise is the major limitation with coherent detection if the local oscillator power is large enough. With large local oscillator power, however, the high-intensity shot noise can also be modeled as additive, white Gaussian noise. With optical amplifiers, ASE usually dominates the shot and thermal noise.

Another source of noise is phase noise, which, as discussed previously, is the random variation in phase of the transmitting laser. The main parameter of interest with phase noise is the width of the phase-noise spectrum relative to the data rate. Wider spectra (or linewidths) result in more signal dispersion as discussed earlier. Also, wider linewidths require wider receive filters (if all the signal energy in the received signal is to be detected), which results in higher thermal noise. However, wider linewidths (if wide enough) can have beneficial effects. In particular, with multimode lasers the distortion caused by PMD is fixed, rather than time varying as with single-frequency lasers, where the worst-case dispersion is significantly greater than the fixed value. Also, with multimode lasers, the distortion due to chromatic dispersion is linear in the received electrical signal rather than nonlinear with direct detection of a single-frequency laser signal. Both of these effects can make compensation of the distortion much easier.

18.3 Modeling of Fiber-Optic Communications Systems

In this chapter we focus on direct-detection, long-haul, fiber-optic systems, such as that illustrated in Fig. 18.1, and we provide a quantitative description of some of the impairment affecting them.

In this figure, the non-return-to-zero (NRZ) input data stream, $c(t)$, depending on a symbol sequence $\{c_k\}$, is filtered by a transmit filter having frequency response $H_T(f)$. The filtered data signal $s(t)$ directly modulates a single frequency laser (SFL). Alternatively, the data stream can be used, as in the dashed path, to drive an external modulator (EM) controlling the optical power in order to avoid laser nonlinearity.

For direct modulation the transmitted signal $x(t)$ can be expressed as

$$x(t) = \sqrt{P(t)} \exp[j(2\pi f_c t + \angle s(t) + \phi_c(t))], \quad (18.1)$$

where $P(t)$ is the optical power [depending on $s(t)$], f_c is the lightwave frequency² and $\angle s(t)$ is the phase of $s(t)$. The phase variation $d\phi_c(t)/dt$ is commonly referred to as *chirp*, whereas the amplitude variation of $P(t)$ versus $|s(t)|$ (the laser power is proportional to its input current) is referred to as *relaxation oscillation*. Both effects can be considered as laser nonlinearities [18].

With external modulation, the transmitted signal $x(t)$ is given by

$$x(t) = \sqrt{|c(t)|} \exp[j(2\pi f_c t + \angle c(t) + \phi(t))], \quad (18.2)$$

where $\phi(t)$ is the laser phase noise. It is worth noting that, in this case, the transmitted power $|x(t)|^2$ is proportional to the amplitude of the data signal $c(t)$.

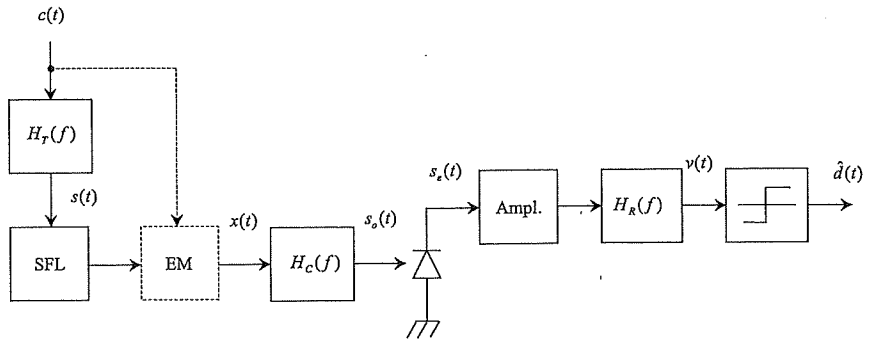


Fig. 18.1 Block diagram of a long-haul, direct-detection fiber-optic system.

² The lightwave frequency is related to the wavelength by $\lambda = c/f$, where c is the speed of light in the fibers.

The transmitted optical signal feeds a fiber, having frequency response $H_C(f)$, which can introduce *chromatic dispersion* and *PMD* [18–20].

Chromatic dispersion will have a significant impact on system performance if the laser frequency is different from that of zero dispersion of the fiber [corresponding to a wavelength equal to $1.3\text{ }\mu\text{m}$ in a standard single-mode fiber (SMF)]. In this case, the main portion of the chromatic dispersion is linear delay distortion, and the frequency response of the fiber is

$$H_C(f) = \exp[j\alpha f^2] \quad (18.3)$$

where

$$\alpha = \pi D \frac{\lambda^2}{c} L, \quad (18.4)$$

where L is the fiber length, D is the *chromatic dispersion parameter*,³ c is the speed of the light in the fiber, and $\lambda(=c/f)$ is the wavelength. Then, in the presence of chromatic dispersion, the optical received signal is expressed by

$$s_o(t) = x(t) \otimes h_C(t), \quad (18.5)$$

where $h_C(t)$ is the channel impulse response of the fiber, and it is the inverse Fourier transform of $H_C(f)$ given in Eq. 18.3; and $\otimes$ denotes convolution operation.

With direct detection, the electrical signal $s_e(t)$ at the photodetector output is related to $s_o(t)$ as

$$s_e(t) = \alpha_{oe} |s_o(t)|^2, \quad (18.6)$$

where α_{oe} is the conversion constant between the optical and the electrical signals.

The PMD can be characterized in terms of first- and higher-order (in frequency) effects. The first-order effect can be simply represented as a delay in the signal in one polarization relative to the delay in the signal in the other one. Thus, with first-order polarization dispersion and direct detection, the electrical signal $s_e(t)$ at the photodetector output is given by

$$s_e(t) = \alpha_{oe} [|s_o(t)|^2 + \alpha_r |s_o(t - \tau)|^2] \quad (18.7)$$

where α_r is the ratio of the signal strengths in the two polarizations and τ is the time delay between propagation in the two polarizations.

³ For example, $D = 17\text{ ps/km/nm}$ for a SMF and $D = 2.6\text{ ps/km/nm}$ for non-zero-dispersion-shifted (NZDS) fiber at $\lambda = 1.55\text{ }\mu\text{m}$.

The electrical signal $s_e(t)$ at the photodetector output is amplified and filtered [the frequency response $H_R(f)$ incorporates the characteristics of frequency selectivity of both the amplifier and the receive filter] to increase the signal-to-noise ratio of the decision variable. This produces the output signal

$$v(t) = s_e(t) \otimes h_R(t), \quad (18.8)$$

where $h_R(t)$ is the inverse Fourier transform of $H_R(f)$. The signal $v(t)$ is detected by comparing the signal level to a decision threshold during a short time interval at the peak opening of the eye in the received digital signal. The quality of data decisions can be seriously affected by ISI, i.e., by interference coming from data transmitted in adjacent bit intervals.

As will become clear in the following sections, a key issue in the effectiveness of equalization techniques against intersymbol interference is the *linearity of the intersymbol interference*, i.e., the linear dependence of the data signal applied to the decision device on the transmitted data. In the communication system of Fig. 18.1, linear distortions include the receiver frequency response $H_R(f)$ and first-order PMD. Chromatic dispersion is a linear distortion in the optical fiber; however, whether this distortion produces a linear dependence on the transmitted data at the detector input [i.e., on $v(t)$] depends on the transmission and detection techniques used. Specifically, chromatic dispersion appears as linear distortion in a coherent receiver, but as long as the laser linewidth is smaller than the data rate, this form of dispersion appears as non-linear in a direct detection system. This is a critical point and we concentrate on this below, for both direct and external modulation systems.

In a direct detection system using external modulation, the transmitted and the received optical signals are expressed by Eqs. 18.2 and 18.5, respectively. Then, after direct detection, the received electrical signal $s_e(t)$ is (see Eqs. 18.5 and 18.6)

$$s_e(t) = \alpha_{oe} |x(t) \otimes h_C(t)|^2, \quad (18.9)$$

and the filtered received signal $v(t)$ is

$$v(t) = \alpha_{oe} |x(t) \otimes h_C(t)|^2 \otimes h_R(t), \quad (18.10)$$

or, in baseband notation,

$$v(t) = \alpha_{oe} \left| \sqrt{|c(t)|} \exp[j(\angle c(t) + \phi(t))] \otimes h_C(t) \right|^2 \otimes h_R(t). \quad (18.11)$$

If we consider long-haul systems operating at a data rate of several Gbps the phase noise bandwidth is typically below 50 MHz, so that the phase noise distortion $\exp[j\phi(t)]$ can be deemed approximately constant over the memory

of the channel impulse response $h_C(t)$. Under this assumption Eq. 18.11 simplifies as

$$v(t) \simeq \alpha_{oe} \left| \sqrt{|c(t)|} \exp[j\angle c(t)] \otimes h_C(t) \right|^2 \otimes h_R(t). \quad (18.12)$$

The last result shows that the ISI in the electrical signal at the input of the decision device is the square of the ISI in the optical signal so that the overall distortion is nonlinear. However, even with this nonlinearity, linear equalization can be partially effective in reducing ISI. In general, linear equalization will not improve performance with quadratic distortion. However, in lightwave systems, if the signaling is binary, on-off keying is used and the major portion of the ISI is due to a single adjacent symbol, a linear equalizer can still be effective. In fact, under these assumptions, it can be shown that the ISI consists of a linear component and a nonlinear component, with the former larger than the latter, if the ISI is reasonably small (see [21], pp. 720–721).

If the laser is directly modulated and the nonlinear amplitude variations in the laser output signal are negligible, the transmitted signal (Eq. 18.1) can be rewritten as

$$x(t) \simeq \sqrt{|s(t)|} \exp[j(2\pi f_c t + \angle s(t) + \phi_c(t))], \quad (18.13)$$

where the phase distortion $\exp[j\phi_c(t)]$ is mainly due to the chirp effect. Substituting Eq. 18.13 into 18.10 produces

$$v(t) \simeq \alpha_{oe} \left| \sqrt{|s(t)|} \exp[j(\angle s(t) + \phi_c(t))] \otimes h_C(t) \right|^2 \otimes h_R(t). \quad (18.14)$$

In this case, the chirp variations are not slow as the chirp bandwidth of lasers is typically larger than that of the data signal [21]. Then the simplification leading from Eq. 18.11 to 18.12 cannot be applied to Eq. 18.14. However, for a rapidly varying $\phi_c(t)$ it can be easily shown that Eq. 18.14 may be approximated as

$$v(t) \cong \alpha_{oe} |s(t)| \otimes |h_C(t)|^2 \otimes h_R(t). \quad (18.15)$$

The last result shows that the ISI in the electrical signal at the input of the decision device is the sum of the squares of the ISI in the optical received signal. The distortion described by Eq. 18.15 may be viewed as due to a *linear* system having impulse response $|h_C(t)|^2$ and fed by $|s(t)|$. The insight provided by Eq. 18.15 is still not accurate. In fact, the chirp effect in the transmitted signal is caused by the signal level changes and, therefore, does not result in fast phase variations over the entire duration of each symbol. In addition, symbols surrounded by identical symbols may not experience chirp. Therefore, in general, in direct modulation systems, the ISI is a combination of linear (Eq. 18.15) and nonlinear distortion (Eq. 18.11). This explains why, in direct detection systems employing direct modulation, linear equalization techniques can only be partially effective in reducing ISI.

18.4 Equalization Algorithms and Their Applications

In this section we delve into the field of equalization algorithms for digital communications and illustrate the application of some of them in the field of optical communications systems. Throughout this section we assume that

1. data decisions are based on the sequence of samples $\{v_k \triangleq v(kT)\}$ (T being the symbol interval equal to the sample interval) of the electrical signal $v(t)$ in Fig. 18.1;
2. the dependence of $\{v_k\}$ on the symbol sequence $\{d_k\}$ is *linear* and the noise⁴ samples $\{n_k\}$ form an additive, white Gaussian noise (AWGN) sequence, i.e.,

$$v_k = \sum_{l=0}^{L-1} c_{k-l} q_l + n_k, \quad (18.16)$$

where $\{q_l\}$ is the discrete-time overall channel impulse response (CIR) and $\{n_k\}$ is a sequence of iid (independent identical distributed) Gaussian random variables having zero mean and variance σ_n^2 . The channel symbols, in general, belong to an M -ary alphabet, but in lightwave systems $M = 2$ is commonly used. Nonetheless, in this scenario multilevel signaling could be used to decrease the symbol rate by a factor $\log_2 M$; this would entail a reduction in the amount of dispersion and decrease the noise power in the detector.

18.4.1 MLSD

It is known from the Equalization theory that, if the CIR is known, maximum likelihood sequence detection (MLSD) is the optimum *sequence* detection technique because it minimizes the error probability in making a decision on the transmitted data sequence. The ML sequence detector for the symbol sequence $\mathbf{c} \triangleq [c_0, c_1, \dots, c_{N-1}]^T$ evaluates the Euclidean distance metric [22]

$$\Lambda(\tilde{\mathbf{c}}) \triangleq \sum_{k=0}^{N-1} |v_k - z_k(\tilde{\mathbf{c}})|^2 \quad (18.17)$$

for each possible trial sequence $\tilde{\mathbf{c}}$, where the quantity

$$z_k(\tilde{\mathbf{c}}) = z_k(\tilde{\mathbf{c}}_{k-L+1}^k) \triangleq \sum_{l=0}^{L-1} \tilde{c}_{k-l} q_l \quad (18.18)$$

⁴ Thermal noise when direct detection is employed.

depends on the data sequence through the vector $\tilde{\mathbf{c}}_{m-L+1}^m \triangleq [\tilde{c}_{m-L+1}, \tilde{c}_{m-L+2}, \dots, \tilde{c}_m]^T$ of L consecutive data symbols only, with the convention that $\tilde{c}_i = 0$ for $i < 0$ or $i > N$. The metric $\Lambda(\tilde{\mathbf{c}})$ in Eq. 18.17 can be also rewritten as

$$\Lambda(\tilde{\mathbf{c}}) = \sum_{k=0}^{N-1} \lambda(x_k, x_{k+1}), \quad (18.19)$$

where

$$\lambda(x_k, x_{k+1}) = |v_k - z_k(\tilde{\mathbf{c}})|^2, \quad (18.20)$$

and x_k is the integer representation of the symbol vector $(\tilde{c}_{k-L+1}, \tilde{c}_{k-L+2}, \dots, \tilde{c}_{k-1})$, consisting of $(L - 1)$ consecutive data symbols. Equations 18.19, 18.20 express the optimal metric as a summation of partial metrics $\{\lambda(x_k, x_{k+1})\}$. The k -th of these terms, $\lambda(x_k, x_{k+1})$, depends on x_k and x_{k+1} , i.e., on the vectors of consecutive trial symbols $(\tilde{c}_{k-L+1}, \tilde{c}_{k-L+2}, \dots, \tilde{c}_{k-1})$ and $(\tilde{c}_{k-L+2}, \tilde{c}_{k-L+3}, \dots, \tilde{c}_k)$, respectively. These considerations suggest a recursive formula for the evaluation of $\Lambda(\tilde{\mathbf{c}})$. In fact, if we define the recursive relation

$$\Lambda(\tilde{\mathbf{c}}_k) = \Lambda(\tilde{\mathbf{c}}_{k-1}) + \lambda(x_k, x_{k+1}), \quad (18.21)$$

with $\tilde{\mathbf{c}}_i \triangleq [\tilde{c}_0, \tilde{c}_1, \dots, \tilde{c}_i]$ ($\tilde{\mathbf{c}}_{N-1} = \tilde{\mathbf{c}}$) and $\Lambda(\tilde{\mathbf{c}}_0) = 0$, then we have

$$\Lambda(\tilde{\mathbf{c}}) = \Lambda(\tilde{\mathbf{c}}_{N-1}) \quad (18.22)$$

after N iterations.

A geometrical representation to the problem of searching over the optimal metric can be given as follows. A trellis diagram with $N_s = M^{L-1}$ states is drawn, as illustrated in Fig. 18.2 for $M = 2$ and $L = 3$. In the k -th interval

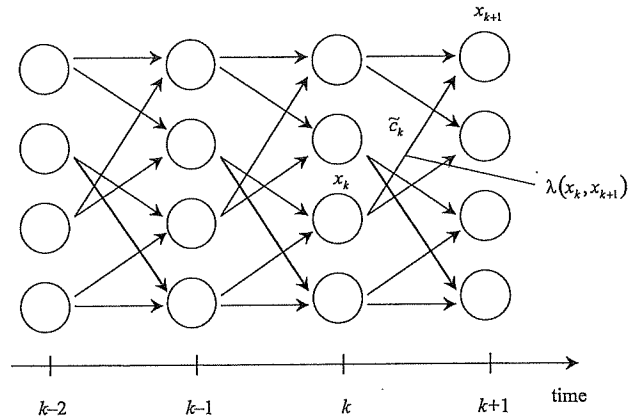


Fig. 18.2 Four-state trellis ($M = 2$ is assumed).

each trellis state represents one of the N_s possible values that x_k can take and is connected via M branches to the next states $\{x_{k+1}\}$. The branch connecting the state that couples x_k and x_{k+1} is labeled by the trial symbol $\tilde{c}_k$ and by the quantity $\lambda(x_k, x_{k+1})$, dubbed the *branch metric*. In this context each trial sequence $\tilde{\mathbf{c}}$ has a one-to-one correspondence with a sequence of states in the trellis diagram, i.e., with a distinct path in the trellis. Moreover, looking for the optimal sequence decision is equivalent to searching the minimum distance path in the trellis. Such a search does not require an exhaustive analysis of all the possible trial sequences, but can be carried out recursively employing the so-called *Viterbi algorithm* (VA) [23–25].

We apply now the VA to implement the MLSD with known CIR. The VA operates on the state trellis, which has already been defined. In the k -th symbol interval, each branch, labeled by the pair of states (x_k, x_{k+1}) , is assigned the corresponding branch metric $\lambda(x_k, x_{k+1})$. The VA's task is to find the path (sequence of branches) through the trellis with smallest metric (i.e., the shortest path). The VA accomplishes this by (see Fig. 18.3):

1. maintaining one *survivor* path per state x_k in the k -th symbol interval;
2. extending these paths one step along all the M branches (labeled by $\tilde{c}_k$) emanating from it;
3. pruning these back by only retaining the path with smallest⁵ metric $\Lambda(\tilde{\mathbf{c}}_k)$ into each state x_{k+1} .

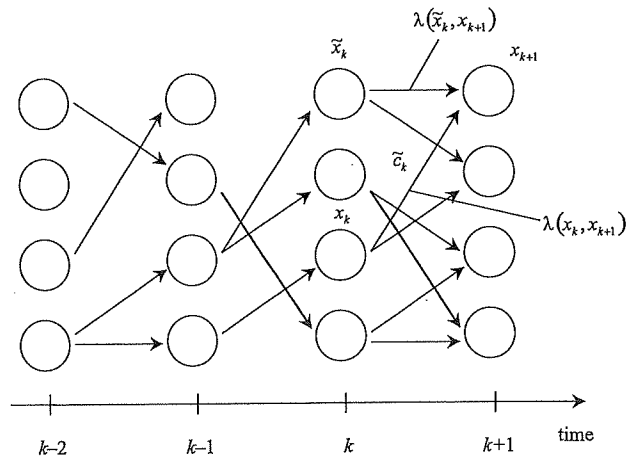


Fig. 18.3 Example of time evolution of the VA.

⁵ When a metric has to be maximized, the VA should be supplied with the negative metric instead.

In the k -th symbol interval, then, the Viterbi algorithm keeps track only of the one path (the so-called *survivor*) leading to each state x_k . Such a path, denoted as $\hat{x}(x_k)$, is the sequence of consecutive states belonging to the path, and it is characterized by an *accumulated metric* $\Lambda(x_k)$.

The VA procedure can be summarized in the following steps (k denotes the time variable):

1. Set

$$k = 0, \quad \hat{x}(x_0) = (x_0), \quad \Lambda(x_0) = 0 \quad (18.23)$$

to initialize the algorithm;

2. Repeat steps 3–7 until $k = N$;
3. Extend path metrics according to Eq. 18.22, that is

$$\Lambda(x_{k+1}) = \Lambda(x_k) + \lambda(x_k, x_{k+1}) \quad (18.24)$$

for all the allowed state transitions $x_k \rightarrow x_{k+1}$;

4. For each destination state x_{k+1} , find the best (minimum metric) incoming path over all the previous states

$$\bar{x}_k = \arg \min_{x_k} \Lambda(x_{k+1}) \quad (18.25)$$

5. Update and store survivor paths as

$$\hat{x}(x_{k+1}) = (\hat{x}(\bar{x}_k), x_{k+1}) \quad (18.26)$$

6. Store the new survivor metrics as

$$\Lambda(x_{k+1}) = \Lambda(\bar{x}_k) + \lambda(\bar{x}_k, x_{k+1}) \quad (18.27)$$

7. Set $k = k + 1$ (increment time counter);
8. Detect the ML decision for the symbol sequence as that associated with the survivor path $\hat{x}(x_N)$ with minimum metric $\Lambda(x_N)$ (termination).

It is worth noting that (a) branch metrics evaluated for state transitions inconsistent with known (i.e., training or pilot) symbols are set to a large value (virtually infinite), and (b) in any real implementation data estimates are generated by the VA with a fixed *decision delay*, K [23] by tracing back from the survivor with instantaneously best metric. In practice, however, there is little degradation if the final decisions are taken after a decision delay of about 5–10 L .

MLSD can be difficult to implement in optical communication systems, even though high-speed VLSI implementations of the VA have been available since the 1990s [26]. The complexity of the VA is governed by the total number of branches, which grows exponentially with the length of CIR. Therefore, for communications channels with ISI over many symbols (i.e., large L) the VA is unreasonably complicated. Various schemes to reduce the VA's complexity have been proposed, and they usually involve searching only part of the trellis [27] or simplifying the trellis. In the latter case delayed decision feedback sequence detection (DDFSD) [28] or reduced state sequence detection (RSSD) [29] techniques can be employed. Both techniques exploit the typical behavior of the CIR, which may have a peak around its middle, referred to as the *cursor*, energy tailing away beforehand, the *precursor*, and energy tailing away afterwards, the *postcursor* (see Fig. 18.4).

In DDFSD, a reduced state trellis is constructed. States in the VA's trellis, $\{x_k = (\tilde{c}_{k-L+1}, \dots, \tilde{c}_{k-1})\}$, are collapsed together if they share the same "older" symbols. In other words, the DDFSD states are defined by the state vector

$$x_k^{RS} = (\tilde{c}_{k-L_{RS}+1}, \dots, \tilde{c}_{k-1}) \quad (18.28)$$

where $L_{RS} \leq L$. The state definition leads to a trellis again, with branches and branch metrics associated with them. Trellis processing closely follows the VA, with survivor paths and survivor path metrics. For each state x_k^{RS} and state transition labeled by the channel symbol $\tilde{c}_k$, the symbols required in the evaluation of the branch metric (Eq. 18.20) are obtained partly from the DDFSD state x_k^{RS} and partly (those unspecified by the state x_k^{RS}) from the corresponding survivor's path $\hat{x}(x_k^{RS})$. This method represents an application of the *decision feedback* concept. As a first approximation, L_{RS} is selected to span the precursor and cursor, and decision feedback is used for the postcursor.

RSSD is an elegant but minor extension of DDFSD. Instead of a full trellis for the precursor and a single decision history per survivor for the postcursor,

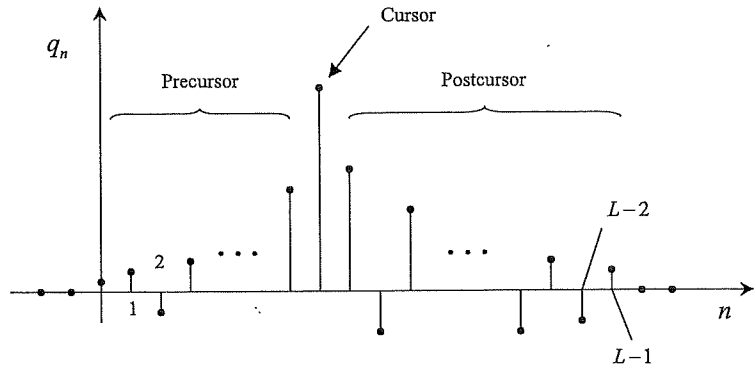


Fig. 18.4 Channel impulse response.

set partitioning principles [30] are applied to steadily reduce the number of hypothesized symbols as more of the received pulse arrives. The general idea behind it is to gracefully degrade performance with decreasing L_{RS} .

Another way of working around a large L is to adaptively prefilter the signal to obtain shorter duration of the overall impulse response by using a linear equalizer [31,32] or a decision feedback equalizer [33–35]. A MLSD based on the VA is then applied to this prefiltered signal. However, the prefilter colors the additive noise, thereby reducing performance if noise correlation is not taken into account.

There are also other techniques to implement simplified, approximate versions of MLSD that may be practical in optical systems operating at high data rates, if the sequence length N is small. One of these is a block-oriented ML detector [36] which aims at determining the bits of a short block of length N , based on N consecutive received signal samples. Specifically, the detector compares blocks of N received signal samples to each of the 2^N ($M = 2$ is assumed) possible (stored) signal sample vectors, each corresponding to one of the possible 2^N bit vectors. The performance of this detector is degraded by the edge effects of ISI. In fact, since ML detection makes decisions on block of N bits from N signal samples referring to these bits, bits at the edge of the block are more likely to be detected in error because of the ISI coming from bits outside the block.

18.4.2 DECISION FEEDBACK EQUALIZATION

The decision feedback equalizer (DFE) can be viewed as the prefilter plus DDFS, where $L_{RS} = 1$ so that x_k^{RS} is an empty state vector. The trellis comprises one state with M branches, each returning to the same state. However, the DFE has a preeminent role in communications due to its nice balance between complexity and performance. Therefore we study it as a separate structure [37–39].

In the DFE terminology, the prefilter is the feedforward filter (FFF), and the decision feedback of DDFS is implemented via a feedback filter (FBF), as illustrated in Fig. 18.5. Since a DFE's trellis has one state, a memoryless quantizer is used as a decision device. As a simplification the FFF in an L_f -tap T -spaced FIR filter reduces precursor ISI, whereas the FBF an L_d -tap T -spaced FIR filter reduces the resulting postcursor ISI.

The impulse responses of the FFF and the FBF are given by

$$\mathbf{f} \triangleq [f_0, f_1, \dots, f_{L_f-1}]^T \quad (18.29)$$

and

$$\mathbf{d} \triangleq [d_0, d_1, \dots, d_{L_d-1}]^T, \quad (18.30)$$

respectively. The feedback filter is always symbol spaced (i.e., the tap spacing is equal to the symbol interval), whereas the feedforward filter may be

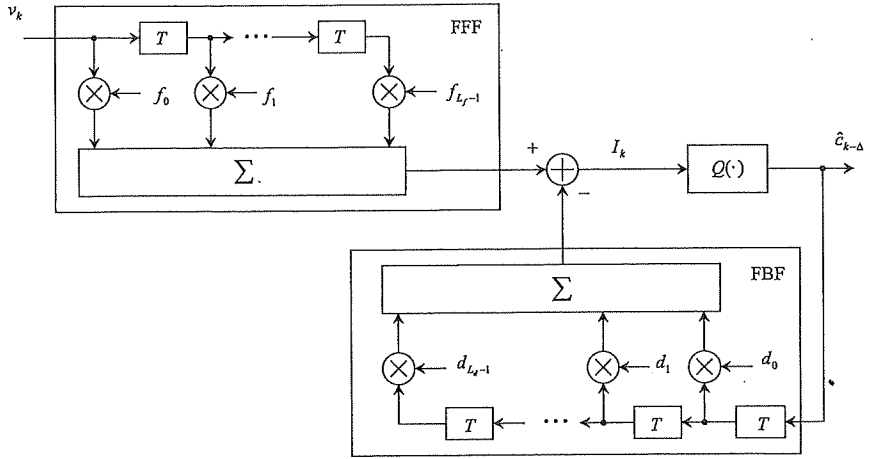


Fig. 18.5 The decision feedback equalizer. The function $Q(\cdot)$ denotes the quantizer.

fractionally spaced (i.e., the tap spacing is a fraction of the symbol interval) [40, 41]. If the FFF is fractionally spaced, η samples per symbol interval are processed by the equalization algorithm. Then the received sample sequence turns into $\{v_k \triangleq v(kT_\eta)\}$, where typically $T_\eta \triangleq T/\eta$, and η is the number of samples per symbol. This increases the BW of the FFF to accommodate the signal whose BW is larger than the symbol rate [42]. In the k -th symbol interval the received samples within the span of the FFF are expressed by (a symbol spaced FFF is assumed for simplicity)

$$\mathbf{y}_k \triangleq [v_{k-L_f+1}, v_{k-L_f+2}, \dots, v_{k-1}]^T.$$

In any DFE, there is always a delay Δ (in symbol periods) between the first received sample containing energy from a symbol and that symbol being detected. Such a parameter is the *decision delay* or *lag*. Therefore, the past decision vector $\hat{\mathbf{c}}_k$ employed in the FBB is given by

$$\hat{\mathbf{c}}_k = [\hat{c}_{k-\Delta-L_d}, \hat{c}_{k+1-\Delta-L_d}, \dots, \hat{c}_{k-\Delta-1}]^T. \quad (18.31)$$

Then from Fig. 18.5, the input to the DFE's quantizer, $Q(\cdot)$, is

$$I_k = \mathbf{f}^H \mathbf{v}_k - \mathbf{d}^H \hat{\mathbf{c}}_k. \quad (18.32)$$

Ideally we would like to have

$$Q(I_k) = c_{k-\Delta}, \quad (18.33)$$

where Δ must satisfy $0 \leq \Delta < L + L_f - 2$.

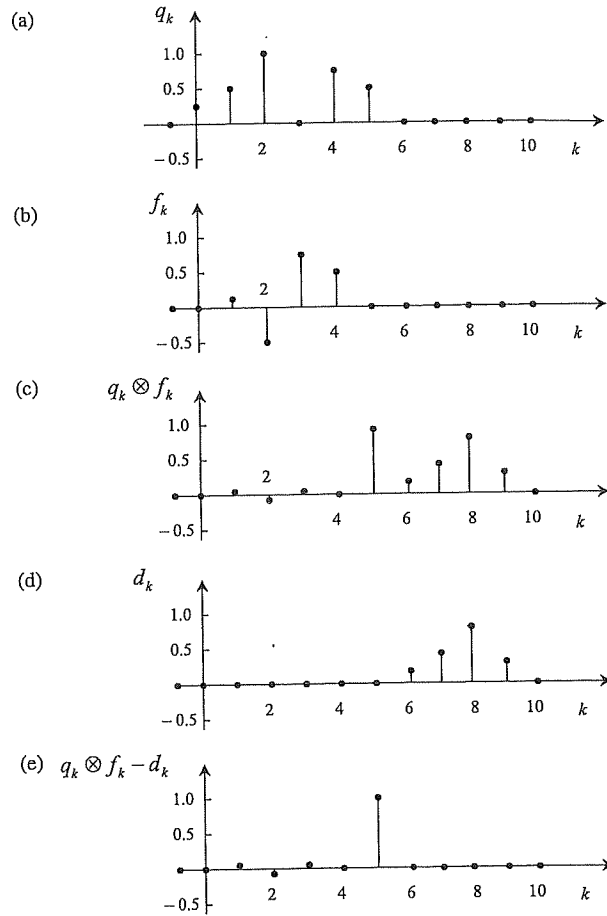


Fig. 18.6 Impulse responses involved in a DFE: (a) channel; (b) feedforward filter; (c) cascade of CIR and FFF; (d) feedback filter; (e) channel and DFE.

The DFE's performance is optimized by appropriately choosing L_f , L_d , Δ , $Q(\cdot)$, $\mathbf{f}$ and $\mathbf{d}$. In known channels, L_f and L_d should be chosen to be as large as possible (i.e., infinite, in which case Δ becomes arbitrary). It is usual to use a hard, nearest-neighbor quantizer (slicer) for $Q(\cdot)$ and then design the DFE's filters so as to minimize the mean square error (MSE) between the quantizer's input I_k and its desired output $c_{k-\Delta}$ under the assumption of correct past decisions, $\hat{\mathbf{c}}_m = \mathbf{c}_m$: the corresponding equalizer is called minimum mean square error–decision feedback (MMSE-DFE) [43,44].

The “correct past decision” assumption has been repeatedly invoked, but it is important to recognize it as a strong assumption. In low SNR, noise may cause a “primary” error, which is fed back into the FBF. Instead of canceling

the postcursor, the FBF can actually enhance it; in turn, this increases the likelihood of subsequent, "secondary" errors [45, 46].

Since the beginning of the 1990s, DFE techniques have been proposed for use in long-haul direct-detection fiber-optic systems where it was shown that they can have more than double the maximum data rate of gigabit-per-second (Gbps) lightwave systems by canceling postcursors' ISI [47]. In this context, DFE can be implemented by the more general technique known as nonlinear cancellation (NLC) [47]. NLC can handle ISI that arises from a variety of impairments, such as transmit laser chirp, chromatic dispersion, PMD, and nonideal receiver frequency response. The first implementation techniques for this equalization strategy are illustrated in [48]. Here, the proposed solution is based on a bank of 2^{L_d} threshold detectors (where L_d is the number of bits fed back), each having a different threshold level that depends on the hypothesized postcursor ISI. The selection of the detector corresponding to the vector of selected symbols occurs after threshold detection. This is accomplished by choosing the output of the detector corresponding to the correct threshold rather than switching analog signals at the data rate in the feedback loop [47]. Since at high data rates, the propagation delay of the detector can exceed one symbol, another advantage of NLC is that the propagation delay of the detector is not an issue. Even if not shown in [48], a nonlinear canceler can also incorporate a feedforward section to mitigate the precursor ISI (see, for instance, [21], p. 713).

Due to the progress in high-speed integrated circuits (ICs) [49], the realization of tunable analog linear and nonlinear filter functions has become feasible for high bit rates [50]. The electronic IC system for decision feedback equalization of 10-Gbps signals is described in [50]. It is based on an eight-tap transversal filter (FFF) and a one-tap feedback section, and its tap weights are adjusted by external tuning voltages. It is interesting to note that in such a device (as in all the other linear and nonlinear electrical equalizers designed until now for optical communications), the FFF is an analog transversal filter, i.e., the input signal of the decision device is

$$i(t) = \sum_{k=0}^7 f_k v(t - kT_\eta), \quad (18.34)$$

where $\{f_k\}$ are the tap coefficients, T_η is the delay introduced by each stage in the tapped delay line (in [50], $T_\eta = 55$ ps, whereas the bit period is $T = 100$ ps). The analog signal $i(t)$ is sampled every T sec (ideally at the instant of maximum eye opening) in the decision circuit and each bit decision is fed back in a unique stage of the FBF. The equalizer, consisting of two ICs (one for the feedforward section and the other one for the feedback section), has been successfully tested in transmission experiments using standard SMF. It has been shown that the power penalties of long-distance links, mainly induced by the chromatic dispersion, can be substantially reduced.

With data rates larger than 10 Gbps (say 40 Gbps), PMD of the fiber becomes a serious impairment, and components and solutions for coping with this distortion become crucial issues. Moreover, since installed fibers are more or less exposed to environmental temperature changes, the actual PMD and, with that, the associated bit error rate (BER), fluctuates with time leading to outage events [51]. Since PMD is an optical property of the fiber, the most straightforward approach to overcome this limitation is the use of optical devices known as *optical PMD compensators*. However, in order to avoid the incorporation of slow bulk optical components, it is useful to move part of the processing from the optical domain to the more flexible electrical domain. For this reason *opto-electronic compensators* based on the use of multiple photo diodes have been proposed [52,53]. A winning alternative to optical and opto-electronic devices is the use of electrical equalizers (as *electronic PMD mitigators*), which offer the potential of a compact integration within the electronics of the optical receiver. Linear [54,55] (see the following section) and, in particular, decision feedback equalizers [56,57], have been proposed for combating PMD. A comparison between different equalization structures is illustrated in [56,58], where it is shown that the DFE (e.g., the concatenation of feedforward transversal filter with a one-tap feedback section) is the technical solution exhibiting the lowest residual penalty in high-data-rate systems. Once again, these results show the potential of postdetection signal processing for the reduction of distortion induced in the optical domain.

18.4.3 LINEAR EQUALIZATION

The linear equalizer (LE) is a linear, normally transversal filter followed by a nearest-neighbor quantizer. The LE is the simplest equalizer and, arguably, the most intuitive as it is an adaptive filter compensating for the channel distortion. Its taps may be T -spaced or fractionally spaced (T_η -spaced) [39]. The T -spaced structure of the LE is illustrated in Fig. 18.7.

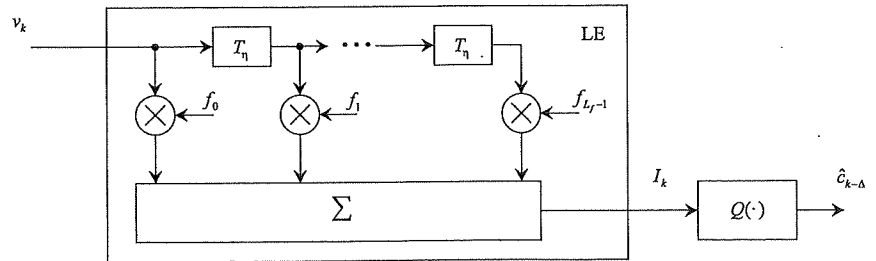


Fig. 18.7 Structure of the symbol-spaced LE.

LEs are also distinguished according to their design criteria. In applications, two criteria are usually considered: zero forcing⁶ (ZF) and minimum mean square error (MMSE). The ZF method neglects the presence of the channel noise and minimizes the ISI contribution in the equalizer time span. This is undesirable since the communication channel may be characterized by multiple, deep notches in frequency. In inverting these, a ZF equalizer inevitably causes undue considerable noise amplification (*noise enhancement*) and a degraded BER. For this reason MMSE is usually a better design criterion, where the combination of noise and ISI are minimized.

In the k -th symbol interval, the received samples within the span of an L_f -tap LE are expressed by (a symbol spaced LE is assumed for simplicity)

$$\mathbf{v}_k \triangleq [v_{k-L_f+1}, v_{k-L_f+2}, \dots, v_{k-1}]^T. \quad (18.35)$$

Then the input to the quantizer of LE is given by

$$I_k \triangleq \mathbf{f}^H \mathbf{v}_k \quad (18.36)$$

where $\mathbf{f}$ is defined in (Eq. 18.29). Then a ZF LE is designed according to

$$(\mathbf{f}, \Delta) = \arg \min_{\tilde{\mathbf{f}}, \tilde{\Delta}} \mathbb{E} \left\{ \left| c_{k-\tilde{\Delta}} - \tilde{\mathbf{f}}^H \mathbf{z}_k \right|^2 \right\}, \quad (18.37)$$

and MMSE LE is designed according to

$$(\mathbf{f}, \Delta) = \arg \min_{\tilde{\mathbf{f}}, \tilde{\Delta}} \mathbb{E} \left\{ \left| c_{k-\tilde{\Delta}} - \tilde{\mathbf{f}}^H \mathbf{v}_k \right|^2 \right\} \quad (18.38)$$

where $\mathbb{E}\{\cdot\}$ denotes expectation operator, $\tilde{\mathbf{f}}$ and $\tilde{\Delta}$ are trial values of $\mathbf{f}$ and Δ , respectively, and $\mathbf{z}_k$ is the useful component of $\mathbf{v}_k$. At high data rates, symbol delays can be implemented by short transmission lines, and the taps can be implemented by variable gains amplifiers. Linear equalizers can be used inserted between the receiver filter and the threshold detector of an optical receiver [42,54,59,60].

In any practical application the coefficients of an LE or a DFE must be adapted to the channel state.

Equalizer tap weights can be adapted using either a single-step technique such as direct matrix inversion (DMI) or a gradient search algorithm (GSA) [21,61]. Although GSAs are much slower to converge than the DMI technique, they are generally suitable for fiber-optic systems since their computational

⁶ The ZF criterion can be also applied in the design of the DFE equalizers. In this case the noise contribution in Eq. 18.32 must be neglected.

complexity is much lower, while the impairment changes with rate much less than the transmission rate.

GSAs include both blind algorithms (see the following section) such as the constant modulus algorithm (CMA) and the least-mean-square (LMS) algorithm [62] or the zero-forcing algorithm [63] in conjunction with a training sequence. It is worth noting that the ZF algorithm minimizes the so-called *peak distortion* (i.e., maximizes the opening of the eye diagram in the absence of thermal noise), whereas the MMSE algorithm minimizes the variance of the error signal. With modest amounts of linear ISI and thermal noise, both equalizers produce distortion-free outputs and provide performance close to the optimum.

To understand the technical problems associated with the use of adaptive LEs in optical systems, let us focus on the celebrated LMS and on the ZF algorithms. They can be expressed as [39] (real signals are assumed here)

$$\mathbf{f}_{k+1}^{LMS} = \mathbf{f}_k^{LMS} - \lambda e_k \mathbf{v}_k \quad (18.39)$$

and

$$\mathbf{f}_{k+1}^{ZF} = \mathbf{f}_k^{ZF} - \lambda e_k \mathbf{c}_k, \quad (18.40)$$

respectively, where $\mathbf{f}_k$ is the tap vector in the k -th signaling interval, λ is the *step size* (controlling the convergence speed and the stability of the algorithm), $\mathbf{c}_k = [c_{k-\Delta-1}, c_{k-\Delta-2}, \dots, c_{k-\Delta-L_f}]^T$ is a past data vector, and

$$e_k \triangleq \tilde{R}_k - c_{k-\Delta} \quad (18.41)$$

is the *error signal*. In order to compute (Eqs. 18.39 and 18.40) the data symbols or very reliable estimates of them have to be available. This is generally achieved by initially using a known sequence to train the equalizer and, after the acquisition period of the equalizer, switch to data decision. In long-haul optical systems operating at high data rates, generating analog samples of the signals may be costly, and therefore may not be desirable. Thus single-bit accuracy samples (corresponding to ± 1) should be used wherever possible. Then an alternative to Eq. 18.39 is provided by the so-called *sgn-sgn least mean square algorithm*,

$$\tilde{\mathbf{f}}_{k+1}^{LMS} = \tilde{\mathbf{f}}_k^{LMS} - \lambda \operatorname{sgn}[e_k] \operatorname{sgn}[\mathbf{v}_k], \quad (18.42)$$

and the *modified zero-forcing algorithm*,

$$\tilde{\mathbf{f}}_{k+1}^{ZF} = \tilde{\mathbf{f}}_k^{ZF} - \lambda \operatorname{sgn}[e_k] \mathbf{c}_k. \quad (18.43)$$

Quantizing the signal samples reduces the rate of convergence of the algorithm. This is not a major concern, since channel impairments are expected to change

very slowly on optical channels. Of more concern, however, especially for the sgn-*sgn* LMS, are the steady-state weights. In fact, since the direction of the correction term $\text{sgn}[\mathbf{v}_k]$ in Eq. 18.42 is different from that for the LMS ($\mathbf{v}_k$ in Eq. 18.39), the equalizer taps may settle at different values in the two cases. On the other hand, the same considerations do not apply to the quantized version of the ZF algorithm (Eq. 18.43). In fact, the direction of the correction term in Eq. 18.43 is the same as for the unquantized algorithm (Eq. 18.40), and the former algorithm converges to the same weights as the latter when the eye diagram of the received signal is open (neither algorithm, however, works when the eye is closed).

As already stated previously, the (unquantized) MMSE algorithm outperforms the ZF one when thermal noise is present and the improvement of MMSE increases with the noise power. Furthermore, unlike the ZF algorithm, when supplied with a training sequence the LMS algorithms can also work when the eye of the receiver is closed. However, the implementation practicalities of Gpbs operation can preclude analog sampling of the signal and, as already mentioned, the quantized version of the LMS is not guaranteed to converge to the correct weights.

In the DFE section it has been pointed out that modern advances in electronic technology have changed the perspective in the implementation of the equalizers. Linear electronic equalization can overcome the link-length limitation imposed by chromatic dispersion over more than 100-km SMF [60, 64–66]. An electronic LE compensating for PMD and chromatic dispersion after 100-km SMF at 10 Gbps is illustrated in [67]. In this case, a three-tap LE is realized with commercially available coaxial components, power splitters, tunable phase shifters, and broadband amplifiers and is preceded by an electrical low-pass filter using the dispersion supported transmission method to compensate for the chromatic dispersion of the fiber [65]. The magnitude of the tap weights in the LE is adjusted by the coaxial attenuator. This seemingly simple solution is shown to reduce the power penalty due to transmission via very high PMD fiber by 8.2 dB. A similar equalization scheme with the same penalty reduction is illustrated in [55].

More recently, advances in SiGe technology [68] have made it possible to implement integrated circuits, that can realize electronic processing functions for 10–40-Gbps optical transmission systems. In particular, it has been possible to implement IC LEs using the continuous time LMS algorithm (Eq. 18.39) for tap adaptation [69]. This solution exhibits excellent properties in terms of speed of convergence as it can accommodate both the slow environmental temperature changes (acting on the PMD properties) and the fast events (within a time scale of a few milliseconds) due to patch-cord moving or mechanical vibrations [57].

Further applications of SiGe tunable electronic equalizers in optical communications over SMF are illustrated in [70].

18.5 Other Equalization Algorithms

In the previous section the most important linear (LE) and nonlinear (MLSD, DFE) equalization techniques for optical communications have been described. In this section we briefly describe other classes of equalization algorithms that might be used in future optical systems.

18.5.1 MAPSD

All previous equalization strategies can be easily incorporated in the receiver structure of Fig. 18.1. Whatever the selected solution, the detector will make a decision on each symbol of the sequence $\{c_k\}$ without providing any information about the reliability of the decision itself to a decoder. Reliability information can be extremely useful if channel coding is employed [71], since it can substantially improve the error performance of the system.

If a vector $\mathbf{v}$ depending on the data vector symbol sequence $\mathbf{c} \triangleq [c_0, c_1, \dots, c_{N-1}]^T$ is processed by the receiver, reliability (or *soft*) information about the transmitted data is produced if the a posteriori probabilities (APPs)

$$P(c_k = \tilde{c} | \mathbf{v}), \quad k = 0, 1, \dots, N-1 \quad (18.44)$$

are evaluated for any possible choice of the symbol $\tilde{c}$. These probabilities can be generated resorting to the so-called maximum a posteriori symbol detection (MAPSD) strategy which represent the optimum *symbol* detection technique since it minimizes the probability of taking a wrong decision on each transmitted symbol.

When the MAP criterion replaces the ML criterion, the MAP forward-backward algorithm (MAP-FBA) replaces the VA [72]. The forward-backward procedure operates on the same state trellis as the VA, but it efficiently calculates the MAP symbol probabilities, instead of just finding the ML sequence. The forward-backward algorithm requires the complete vector $\mathbf{v}$ of received samples to be available before decisions can be made. If a large decision delay is to be avoided, a near-optimal, *fixed-lag* (forward-only) MAP (MAP-FLA) strategy can be adopted [73]. However, both the FB and FL MAP algorithms work with likelihoods rather than log-likelihoods, so that their branch metrics are generally more expensive to compute and their operations are multiplication and addition instead of the addition and minimization characterizing the Viterbi algorithm. Much work on reduced complexity implementations has been undertaken (see [22] and references therein), but this is outside the scope of this chapter.

Let us now briefly describe the MAP algorithms mentioned above.

The Forward-Backward Algorithm

The FBA seeks the probabilities (Eq. 18.44) for any possible trial symbol $\tilde{c}$ transmitted in the k -th signaling interval. As in Section 18.4.1 it is assumed here that the channel memory is not longer than $(L - 1)$ symbol periods, so that any received signal sample depends on L consecutive channel symbols at most. Then the FBA operates on the same M^{L-1} -state trellis as the VA and performs two basic recursions on the stream of received signal samples $\mathbf{v}$ in order to evaluate the probabilities (Eq. 18.44). In the following it is useful to keep in mind that:

1. each node in the trellis has M input branches and M output branches, with a branch corresponding to one of the M data symbols;
2. it is useful to associate the computed values of the FBA with nodes, states, and transitions in the trellis.

The evaluation of Eq. 18.44 requires the computation of intermediate quantities, known as *state transition probabilities*. Given the trellis states x_k and x_{k+1} in the k -th and in the $(k + 1)$ -th symbol interval, respectively, the corresponding *state transition probability* is given by

$$P(x_k, x_{k+1} | \mathbf{v}) \quad (18.45)$$

and equals zero if the states are not connected. This quantity can be evaluated as

$$P(x_k, x_{k+1} | \mathbf{v}) = \frac{\sum_{\tilde{c} \in C(x_k, x_{k+1})} P(\tilde{c} | \mathbf{v})}{\sum_{\tilde{c} \in C} P(\tilde{c} | \mathbf{v})}, \quad (18.46)$$

where $C(x_k, x_{k+1})$ is the subset (of the set $C \triangleq \{\tilde{c}\}$) consisting of all the possible trial sequences that traverse the trellis branch connecting the states x_k and x_{k+1} , as illustrated in Fig. 18.8.

In [22] it is proved that, given the signal model (Eq. 18.16) and under the assumption of independent symbols, Eq. 18.46 can be also rewritten as

$$P(x_k, x_{k+1} | \mathbf{v}) = \frac{\sum_{\tilde{c} \in C(x_k, x_{k+1})} \prod_{k=0}^{N-1} \gamma_k(x_k, x_{k+1})}{\sum_{\tilde{c} \in C} \prod_{k=0}^{N-1} \gamma_k(x_k, x_{k+1})} \quad (18.47)$$

where

$$\gamma_k(x_k, x_{k+1}) \triangleq P(\tilde{c}_k) W_k(x_k, x_{k+1}) \quad (18.48)$$

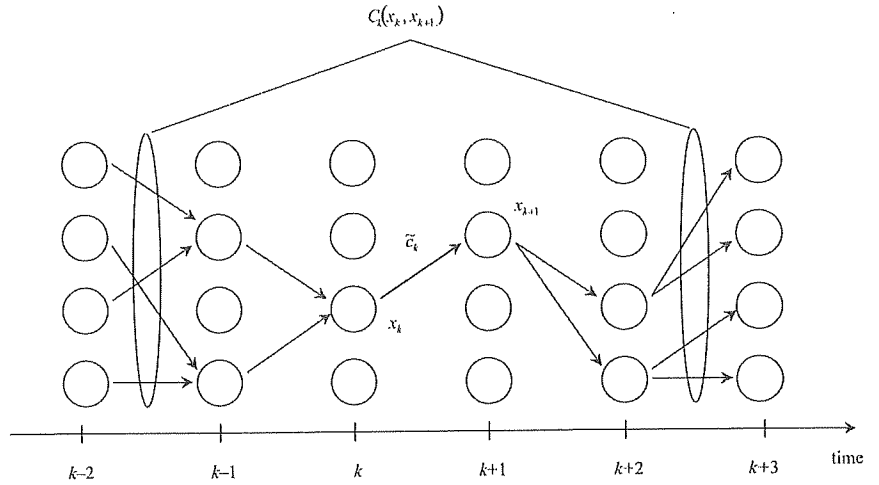


Fig. 18.8 Subset $C_k(x_k, x_{k+1})$ of all the possible trial sequences traversing the trellis branch between the states x_k and x_{k+1} . $M = 2$ and $L = 2$ are assumed.

is a *weight function* depending on trellis branch connecting the states x_k and x_{k+1} , $P(\tilde{c}_k)$ is the probability⁷ of the event $\{c_k = \tilde{c}_k\}$ and

$$W_k(x_k, x_{k+1}) \triangleq p(v_k | \tilde{c}_k, x_k) = \frac{1}{\pi \sigma_n^2} \exp \left[-\frac{1}{\sigma_n^2} \left| r_k - \sum_{l=0}^{L-1} \tilde{c}_{k-l} q_l \right|^2 \right]. \quad (18.49)$$

Equation 18.47 shows that the evaluation of the state transition probabilities requires the computation of (a) the sum of the products of the weights associated with all the paths containing the branch going out of x_k and entering x_{k+1} (see the numerator), and (b) the sum of the products of the weights associated with all the paths in the trellis (see the denominator). A computationally efficient method to solve this problem is the following. Let us define the quantities $\{\alpha_k(x_k)\}$ and $\{\beta_k(x_k)\}$ through the recursive formula

$$\alpha_k(x_k) = \sum_{x_{k-1}} \alpha_{k-1}(x_{k-1}) \cdot \gamma_k(x_{k-1}, x_k) \quad (18.50)$$

with $k = 1, 2, \dots, N-1$, and

$$\beta_k(x_k) = \sum_{x_{k+1}} \beta_{k+1}(x_{k+1}) \cdot \gamma_k(x_k, x_{k+1}) \quad (18.51)$$

with $k = N-2, N-3, \dots, 0$, respectively. Here it is assumed that $\{\alpha_0(x_0)\}$ and $\{\beta_{N-1}(x_{N-1})\}$ are known initial conditions.

⁷ Such a probability is equal to $1/M$ for equiprobable channel symbols.

The quantity $\alpha_k(x_k)$ expresses the sum of the products of the weights along all paths originating from all the possible past initial states $\{x_0\}$ and terminating at x_k in the k -th signaling interval. Similarly, $\beta_k(x_k)$ represents the sum of the products of the weights along all paths ending in all the possible future terminal states $\{x_{N-1}\}$ and originating from x_k in the k -th signaling interval. Then the numerator of Eq. 18.47 can be evaluated as

$$\begin{aligned}\sigma_k(x_k, x_{k+1}) &\triangleq \sum_{\tilde{c} \in C(x_k, x_{k+1})} \prod_{m=0}^{N-1} \gamma_k(x_k, x_{k+1}) \\ &= \alpha_k(x_k) \cdot \gamma_k(x_k, x_{k+1}) \cdot \beta_{k+1}(x_{k+1})\end{aligned}\quad (18.52)$$

whereas its denominator as

$$\sum_{\tilde{c} \in C} \prod_{k=0}^{N-1} \gamma_k(x_k, x_{k+1}) = \sum_{x_k, x_{k+1}} \sigma_k(x_k, x_{k+1}) \quad (18.53)$$

so that

$$P(x_k, x_{k+1} | \mathbf{v}) = \frac{\sigma_k(x_k, x_{k+1})}{\sum_{x_k, x_{k+1}} \sigma_k(x_k, x_{k+1})}. \quad (18.54)$$

The last result shows that all that is needed for the evaluation of the state transition probabilities are the quantities $\{\sigma_k(x_k, x_{k+1})\}$, computed as in Eq. 18.52. This, in turn, requires a forward (Eq. 18.50) and a backward recursion (Eq. 18.51) involving all the trellis states in each signaling interval, as illustrated in Fig. 18.9. It is worth noting that both recursions over the trellis only need to be performed once.

For demodulation purposes, the quantity of interest is $P(\tilde{c}_k | \mathbf{v})$ for any possible value of the channel symbol $\tilde{c}_k$. This probability can be calculated by summing all the state transition probabilities (Eq. 18.54) that correspond to branches associated with the symbol $\tilde{c}_k$. Then if we define the set $S(\tilde{c}_k)$ of all state transitions (x_k, x_{k+1}) , such that the channel symbol labeling the corresponding branch is $\tilde{c}_k$, $P(\tilde{c}_k | \mathbf{v})$ can be evaluated as

$$P(\tilde{c}_k | \mathbf{v}) = \sum_{(x_k, x_{k+1}) \in S(\tilde{c}_k)} P(x_k, x_{k+1} | \mathbf{r}, \tilde{\mathbf{q}}) = \frac{\sum_{(x_k, x_{k+1}) \in S(\tilde{c}_k)} \sigma_k(x_k, x_{k+1})}{\sum_{x_k, x_{k+1}} \sigma_k(x_k, x_{k+1})}. \quad (18.55)$$

Finally, the FBA can be summarized in the following steps:

1. Evaluate the conditional probabilities $\{\gamma_k(x_k, x_{k+1})\}$ for all the trellis branches using Eqs. 18.48 and 18.49;
2. Initialize the forward recursion setting

$$k = 1, \quad \alpha_0(x_0) = 1 \quad (18.56)$$

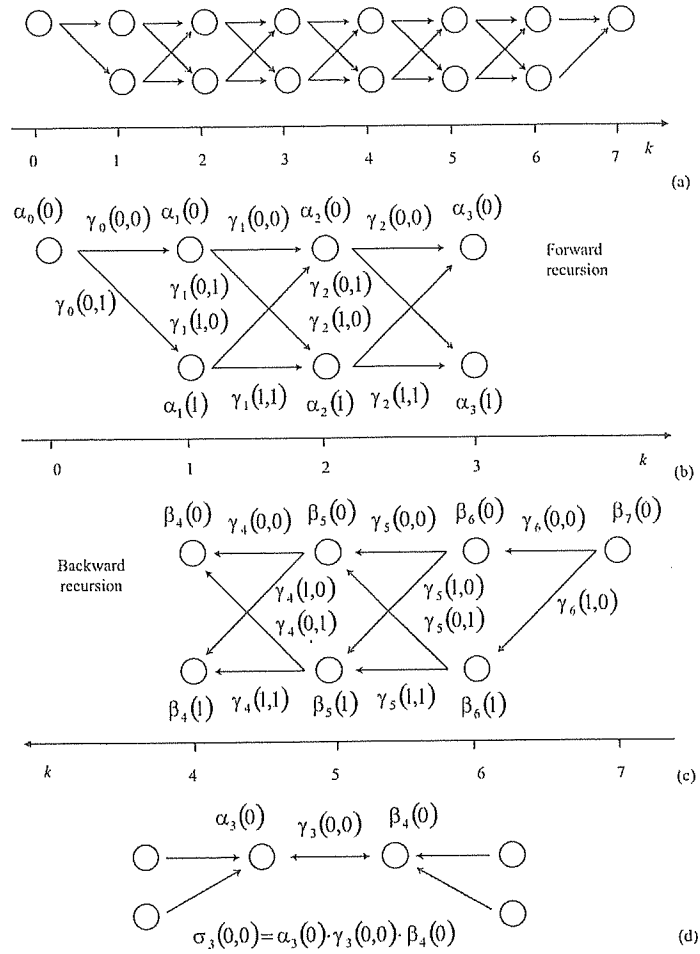


Fig. 18.9 Application of the FBA to a two-state trellis ($M = 2$, $L = 1$). The quantities $\{\alpha_k(x_k)\}$ and $\{\beta_k(x_k)\}$ are computed recursively in the forward (b) and in the backward (c) recursion, respectively, using the branch probabilities $\{\gamma_k(x_k, x_{k+1})\}$. Finally, (d) the state transition probabilities $\{\sigma_k(x_k, x_{k+1})\}$ are evaluated according to Eq. 18.52.

3. Repeat steps 4–5 until $k = N$;
4. Compute and store the forward path probabilities $\{\alpha_k(x_k)\}$ using Eq. 18.50;
5. Set $k = k + 1$;
6. Initialize the backward recursion setting

$$k = N - 2, \quad \beta_0(x_{N-1}) = 1 \quad (18.57)$$

7. Repeat steps 8–9 until $k = 0$;
8. Compute and store the backward path probabilities $\{\beta_k(x_k)\}$ using Eq. 18.51;
9. Set $k = k - 1$;
10. For each trellis branch compute the quantity $\sigma_k(x_k, x_{k+1})$ using Eq. 18.52;
11. Evaluate the a posteriori symbol probabilities $\{P(\tilde{c}_k|\mathbf{v})\}$ by means of Eq. 18.55.

Fixed-Lag MAPSD

The forward-backward algorithm is suited to short burst transmission because, otherwise, its delay and storage requirements are unsatisfactory. The MAP fixed-lag algorithm (MAP-FLA) detector [73,74] makes decisions at a fixed lag of L_{fl} symbols from the current received samples, as

$$\hat{c}_m = \arg \max_{\tilde{c}_k} P(\tilde{c}_k | \mathbf{v}_0^{(L_{fl}+k+1)\eta-1}), \quad (18.58)$$

where the notation $\mathbf{x}_a^b$ denotes the vector $[x_a, x_{a+1}, \dots, x_b]^T$. The parameter L_{fl} is akin to the decision delay in the VA, and so, for effectively optimal performance, it is expected to be as high as $5L$ or $7L$. The decision rule may be implemented via two good algorithms, both requiring a single forward pass only. The newer one, dubbed soft-output algorithm (or OSA) [75], has a smaller computational complexity than the older one [73], since the number of quantities to be stored and recursively updated increases linearly, rather than exponentially, with the decision delay.

A description of the MAP-FLA is not provided here. However, it is important, at this stage, to make some comments about the VA, the MAP-FLA, and the MAP-FBA. Each of these algorithms can be employed for the detection of data sequences in the presence of ISI. Among these algorithms, however, the VA has the smallest complexity and this makes it the most reasonable choice in uncoded communication systems where soft output information is not required. The MAP algorithms have the following relevant disadvantages: (1) they require the knowledge of the noise variance; and (2) they have large memory and computation requirements. In particular, as already stated, MAP algorithms accomplish computations in the probability, instead of the logarithm domain and, consequently, require a large number of multiply and exponential operations. The MAP-FBA detector performs two recursions and, consequently, it operates in block mode. It memorizes all the received signal samples of each block and afterwards processes them. Its memory requirement grows linearly with the sequence length so that only short data sequences should be processed. The MAP-FLA requires only a forward recursion, so that it can operate in continuous-mode. Both the memory and computational

requirements, however, grow exponentially with the decision delay so that this parameter should be kept to minimum. For this reason the FBA may have smaller computational requirements than the FLA.

18.5.2 BLIND TECHNIQUES AND CMA

The LE and the DFE equalization algorithms are commonly derived assuming a known CIR. In practice, such equalization algorithms learn the channel state in the training period during which a known data sequence is transmitted. An alternative to this approach is represented by blind equalization techniques.

An equalizer can compensate for channel distortion even in the absence of known symbols, by using one of a number of *blind* deconvolution algorithms. These may be classified into four classes: algorithms explicitly using higher order statistics (HOS), subspace algorithms, Bussgang algorithms, and joint data detection and channel estimation techniques. Here we shortly discuss only Bussgang algorithms. Such algorithms employ a gradient descent algorithm that attempts to restore some property of the input data that was destroyed by the frequency selective channel [76, 77]. Let us consider a linear equalizer with contents $\mathbf{v}_k$, coefficients $\mathbf{f}_k$, and quantizer input I_k defined by Eqs. 18.29, 18.35 and 18.36, respectively, and update equation (see Eq. 18.39)

$$\mathbf{f}_{k+1} = \mathbf{f}_k - \lambda \mathbf{v}_k e(I_k), \quad (18.59)$$

where λ is the step size and the error factor $e(\cdot)$ is a nonlinear, memoryless mapping that characterizes each Bussgang variant. There are two important variants for $e(\cdot)$ in the family of blind equalizers: (a) the Sato algorithm [76] and (b) the constant modulus algorithm (CMA) [77].

As we have previously shown with LMS-DFE, the error factor is evaluated as

$$e(I_k) \triangleq I_k - Q(I_k), \quad (18.60)$$

where $Q(\cdot)$ is the nearest neighbor quantizer. Given reliable (ideally *correct*) decisions, this choice minimizes the residual mean square error between the input and output of the quantizer, and thus it is well-suited to tracking. It is unsuitable for acquisition since typically many incorrect decisions are fed back initially, and the equalizer may converge to a closed-eye setting (i.e., the cost function has many local minima). For this reason, common practice is to use another blind algorithm until the eye diagram is opened and then operation is switched to decision-directed mode.

Sato's algorithm pioneered the field of blind equalization. Given a real constellation and real received samples, it uses

$$e(I_k) \triangleq I_k - \text{sgn}(I_k). \quad (18.61)$$

Essentially the symbols are quantized as pseudobinary, and this increases the decision reliability when large constellations are employed. It is demonstrably less likely to converge to a local minimum than the decision directed version. However, it is known that it can get hung-up when initializing a DFE.

The algorithm of choice is the CMA [78]. It attempts to ensure that the quantizer input lies as close as possible to a circle with radius $\sqrt{C}$ by minimizing

$$\epsilon_{CMA} \triangleq E \left\{ (|I_k|^2 - C)^2 \right\} \quad (18.62)$$

with

$$C \triangleq \frac{E \{ |I_k|^4 \}}{E \{ |I_k|^2 \}}. \quad (18.63)$$

It may be used for real/complex signals and constant/nonconstant envelope modulations. In recursive form the error factor is calculated as

$$e(I_k) = I_k [|I_k|^2 - C]. \quad (18.64)$$

Given diversity (through fractional sampling for instance), a finite length equalizer adapted by the CMA can be shown to be *globally convergent*, i.e., to converge to the optimum setting whatever its initialization.

References

- [1] J. H. Winters, "Reducing the effects of transmission impairments in digital fiber optic system," *IEEE Comm. Mag.*, pp. 68–76, June 1993.
- [2] W. D. Grover, "Forward error correction in dispersion-limited lightwave systems," *IEEE Journal of Lightwave Tech.*, vol. 6, pp. 643–654, May 1988.
- [3] A. F. Elrefaie, R. E. Wagner, D. A. Atlas, and D. G. Daut, "Chromatic dispersion limitations in coherent lightwave transmission systems," *IEEE Journal of Lightwave Tech.*, vol. 6, p. 704, May 1988.
- [4] H. Bischl and F. Derr, "Chromatic and polarization dispersion limitations in coherent optical bpsk and qpsk systems," *IEEE Journal of Optical Comm.*, vol. 12, pp. 42–46, June 1991.
- [5] R. E. Wagner and A. F. Elrefaie, "Polarization-dispersion limitations in lightwave systems," in *Technical Digest, Optical Fiber Communications Conference* (New Orleans, LA), p. 37, Jan. 1988.
- [6] C. D. Poole and R. E. Wagner, "Phenomenological approach to polarisation dispersion in long single-mode fibres," *IEE Electronics Letters*, vol. 22, pp. 1029–1030, Sept. 1986.
- [7] C. D. Poole and C. R. Giles, "Polarization-dependent pulse compression and broadening due to polarization in dispersion-shifted fiber," *Optics Letters*, vol. 13, pp. 155–157, Feb. 1988.

- [8] J. H. Winters, M. A. Santoro, and Z. Haas, "On the experimental measurement of pmd effects," in *Proc. of SPIE OE/Fibers '92*, p. 1784.05, Sept. 1992.
- [9] C. D. Poole, R. W. Tkach, A. R. Chraplyvy, and D. A. Fishman, "Fading in light-wave systems due to polarization-mode dispersion," *IEEE Photonics Technology Letters*, vol. 3, pp. 68–70, Jan. 1991.
- [10] C. D. Angelis, A. Galtarossa, G. Gianello, F. Matera, and M. Schiano, "Time evolution of polarization mode dispersion in long terrestrial links," *IEEE Journal of Lightwave Tech.*, vol. 10, pp. 552–555, May 1992.
- [11] M. Tsubokawa, M. Ohashi, and Y. Sasaki, "Waveform degradation due to dispersion effects in an optical cfsk system," *IEEE Journal of Lightwave Tech.*, vol. 8, pp. 775–779, May 1990.
- [12] A. A. M. Saleh, R. M. Jobson, and T. E. Darcie, "Compensation of nonlinearity in semiconductor optical amplifiers," *IEEE Electronics Letters*, vol. 15, pp. 950–952, July 1988.
- [13] W. J. Tomlinson and R. H. Stolen, "Nonlinear phenomena in optical fibers," *IEEE Communications Mag.*, vol. 26, pp. 36–44, April 1988.
- [14] A. R. Chraplyvy, "Limitations on lightwave communications imposed by optical-fiber nonlinearities," *IEEE Journal of Lightwave Tech.*, vol. 9, pp. 1548–1557, Oct. 1990.
- [15] R. Giles and J. Conradi, "Laser mode partitioning and mode evolution effects in high-speed fiber-optic transmission," in *Technical Digest of OFC'86*, pp. 110–111, 1986.
- [16] K. Ogawa, "Analysis of mode partition noise in laser transmission system," *IEEE Journal of Quantum Electronics*, vol. QE-18, pp. 849–855, May 1982.
- [17] B. L. Kasper, "Equalization of multimode optical fiber systems," *Bell Syst. Tech. J.*, vol. 61, pp. 1367–1388, Sept. 1982.
- [18] L. Kazosky, S. Benedetto, and A. Willner, *Optical Fiber Communication Systems*. London: Artech House, 1996.
- [19] R. E. Wagner and A. F. Elrefaie, "Polarization-dispersion limitations in coherent lightwave systems," in *Technical Digest, Optical Fiber Communications Conference* (New Orleans, LA), p. 37, Jan. 1988.
- [20] A. F. Elrefaie, R. E. Wagner, D. A. Atlas, and D. G. Daut, "Chromatic dispersion limitations in coherent lightwave systems," *IEEE Journal of Lightwave Tech.*, vol. 6, pp. 704–709, May 1988.
- [21] R. D. Gitlin, J. F. Hayes, and S. B. Weinstein, *Data Communications Principles*. New York: Kluwer Academic/Plenum Publishers, 1992.
- [22] A. F. Molisch, *Wideband Wireless Digital Communications*. Upper Saddle River, New Jersey: Prentice Hall, 2001.
- [23] D. G. Forney, Jr., "Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference," *IEEE Trans. Inf. Theory*, vol. 18, pp. 363–378, May 1972.
- [24] D. G. Forney, Jr., "The Viterbi algorithm," *Proc. IEEE*, vol. 61, pp. 268–278, Mar. 1973.

- [25] J. A. Heller and I. M. Jacobs, "Viterbi decoding for satellite and space communication," *IEEE Trans. Commun. Tech.*, vol. COM-19, pp. 835–848, Oct. 1971.
- [26] G. Fettweis and H. Meyr, "High speed parallel Viterbi decoding: Algorithm and VLSI-architecture," *IEEE Communications Magazine*, vol. 29, pp. 46–55, May 1991.
- [27] J. B. Anderson, "Sequential coding algorithms: A survey and cost analysis," *IEEE Trans. Comm.*, vol. 32, pp. 169–176, Feb. 1984.
- [28] A. D. Hallen and C. Heegard, "Delayed decision-feedback sequence estimation," *IEEE Trans. Comm.*, vol. 37, pp. 428–436, May 1989.
- [29] M. V. Eyuboglu and S. U. H. Qureshi, "Reduced-state sequence estimation with set partitioning and decision feedback," *IEEE Trans. Comm.*, vol. 36, pp. 13–20, Jan. 1988.
- [30] G. Ungerboeck, "Channel coding with Multilevel/Phase signals," *IEEE Trans. Inf. Theory*, vol. 28, pp. 55–67, 1982.
- [31] S. U. H. Qureshi and E. E. Newhall, "An adaptive receiver for data transmission over time-dispersive channels," *IEEE Trans. Inf. Theory*, vol. 19, pp. 448–457, July 1973.
- [32] D. D. Falconer and F. R. Magee, "Adaptive channel memory truncation for maximum-likelihood sequence estimation," *Bell Syst. Tech. J.*, vol. 52, pp. 1541–1562, Nov. 1973.
- [33] W. U. Lee and F. Hill, "A maximum-likelihood sequence estimator with decision-feedback equalisation," *IEEE Trans. Comm.*, vol. 25, pp. 971–979, Sept. 1977.
- [34] Y. Gu and T. Le-Ngoc, "Adaptive combined DFE/MLSE techniques for ISI channels," *IEEE Trans. Comm.*, vol. 44, pp. 847–857, July 1996.
- [35] K. Wesolowski, "An efficient DFE & ML suboptimum receiver for data transmission over dispersive channels using two-dimensional signal constellation," *IEEE Trans. Comm.*, vol. 35, pp. 337–339, March 1987.
- [36] J. H. Winters and S. Kasturia, "Constrained maximum likelihood detection for high-speed fiber-optic systems," in *Proc. Globecom '91*, pp. 1574–1579, Dec. 1991.
- [37] M. Austin, "Decision feedback equalization for digital communication over dispersive channels," *MIT Research Laboratory of Electronics Technical Report 461*, Aug. 1967.
- [38] C. A. Belfiore and J. H. Park, "Decision feedback equalization," *Proc. IEEE*, vol. 67, pp. 1143–1156, Aug. 1979.
- [39] S. U. H. Qureshi, "Adaptive equalization," *Proc. IEEE*, vol. 73, pp. 1349–1387, Sept. 1985.
- [40] R. D. Gitlin and S. B. Weinstein, "Fractionally-spaced equalization: An improved digital transversal equalizer," *Bell Syst. Tech. J.*, vol. 60, pp. 856–864, Feb. 1981.
- [41] G. Ungerboeck, "Fractional tap-spacing equaliser and consequences for clock recovery in data modems," *IEEE Trans. Comm.*, vol. 24, pp. 856–864, Aug. 1976.
- [42] J. H. Winters, "Equalization in coherent lightwave systems using a fractionally-spaced equalizer," *IEEE Journal of Lightwave Tech.*, vol. 8, pp. 1487–1491, Oct. 1990.

- [43] J. Salz, "Optimum mean-square decision feedback equalization," *Bell Syst. Tech. J.*, vol. 52, pp. 1341–1373, Oct. 1973.
- [44] D. D. Falconer and G. J. Foschini, "Theory of minimum mean-square error QAM systems employing decision feedback equalization," *Bell Syst. Tech. J.*, vol. 52, pp. 1821–1849, Dec. 1973.
- [45] D. L. Duttweiler, J. E. Mazo, and D. G. Messerschmitt, "An upper bound on the error probability in decision-feedback equalisation," *IEEE Trans. Inf. Theory*, vol. 20, pp. 490–497, July 1974.
- [46] R. A. Kennedy, B. D. O. Anderson, and R. R. Bitmead, "Tight bounds on the error probabilities of decision feedback equalizers," *IEEE Trans. Comm.*, vol. 35, pp. 1022–1029, Oct. 1987.
- [47] J. H. Winters and R. D. Gitlin, "Electrical signal processing techniques in long-haul fiber-optic systems," *IEEE Trans. Comm.*, vol. 38, pp. 1439–1453, Sept. 1990.
- [48] S. Kasturia and J. H. Winters, "Techniques for high-speed implementation of nonlinear cancellation," *IEEE J. Sel. Areas Comm.*, vol. 9, pp. 711–717, June 1991.
- [49] A. Schüppen, H. Dietrich, U. Seiler, H. V. D. Ropp, and U. Erben, "A SiGe RF technology for mobile communication systems," *Microw. Eng. Europe*, pp. 39–46, 1998.
- [50] F. Buchali, H. Bülow, W. Baumert, R. Ballentin, and T. Wehreu, "Reduction of the chromatic dispersion penalty at 10 Gbit/s by integrated electronic equalisers," in *Optical Fiber Communication Conference, 2000*, vol. 3, pp. 268–270, March 2000.
- [51] H. Bülow, "System outage probability due to first- and second-order PMD," *IEEE Photon. Technol. Lett.*, vol. 10, pp. 696–698, May 1998.
- [52] H. Bülow, "Equalization of bit distortion induced by polarization mode dispersion," in *Proc. NOC'97*, 1997.
- [53] B. W. Hakki, "Polarization mode dispersion compensation by phase diversity detection," *IEEE Photon. Technol. Lett.*, vol. 9, pp. 121–123, Jan. 1997.
- [54] J. H. Winters, "Experimental equalization of polarization dispersion," *IEEE Photon. Technol. Lett.*, vol. 2, pp. 591–593, Oct. 1990.
- [55] H. Bülow, D. Schlump, J. Weber, B. Wedding, and R. Heidemann, "Electronic equalization of fiber PMD-induced distortion at 10 Gbit/s," in *Optical Fiber Communication Conference and Exhibit, 1998 (OFC '98), Technical Digest*, vol. 2, pp. 151–152, Feb. 1998.
- [56] H. Bülow, F. Buchali, W. Baumert, R. Ballentin, and T. Wehren, "PMD mitigation at 10 Gbit/s using linear and nonlinear integrated electronic equaliser circuits," *Electron. Lett.*, vol. 36, pp. 163–164, Jan. 2000.
- [57] H. Bülow, "PMD mitigation techniques and their effectiveness in installed fiber," in *Optical Fiber Communication Conference, 2000*, vol. 3, pp. 110–112, March 2000.
- [58] J. Poirrier, A. Gnauck, and J. Winters, "Experimental nonlinear cancellation of polarization-mode dispersion," in *Optical Fiber Communication Conference, 2000*, vol. 3, pp. 119–121, 2000.

- [59] D. G. Messerschmitt, "Minimum MSE equalization of digital fiber optic systems," *IEEE Trans. Comm.*, vol. 16, pp. 1110–1118, July 1978.
- [60] J. H. Winters and M. A. Santoro, "Experimental equalization of polarization diversity," *IEEE Photon. Technol. Lett.*, vol. 2, pp. 591–593, Aug. 1990.
- [61] J. G. Proakis, *Digital Communications*. New York, New York: McGraw-Hill, Inc., Fourth ed., 2001.
- [62] B. Widrow, J. McCool, M. Larimore, and C. Johnson, "Stationary and non-stationary learning characteristics of the LMS adaptive filter," *Proc. IEEE*, vol. 64, pp. 1151–1162, Aug. 1976.
- [63] R. W. Lucky, "Automatic equalization for digital communications," *Bell Syst. Tech. J.*, vol. 44, pp. 547–588, April 1965.
- [64] D. Schlump, K. Koffers, W. Pohlmann, H. Reichelt, and B. Wedding, "10 Gbit/s dispersion supported transmission field trial over 123 km standard single mode fibre for HDTV studio interconnection," *Electron. Lett.*, vol. 31, pp. 1854–1855, Oct. 1995.
- [65] B. Wedding, B. Franz, and B. Junginger, "10-Gb/s optical transmission up to 253 km via standard single-mode fiber using the method of dispersion-supported transmission," *IEEE Journal of Lightwave Tech.*, vol. 12, pp. 1720–1727, Oct. 1994.
- [66] T. Takahashi, T. Imai, and M. Aiki, "PMD mitigation at 10 Gbit/s using linear and nonlinear integrated electronic equaliser circuits," *Electron. Lett.*, vol. 30, pp. 348–349, Feb. 1994.
- [67] D. Schlump, B. Wedding, and H. Bülow, "Electronic equalisation of PMD and chromatic dispersion induced distortion after 100 km standard fibre at 10 Gbit/s," in *24th European Conference on Optical Communication*, vol. 1, pp. 535–536, Sept. 1998.
- [68] B. Wedding, W. Pohlmann, D. Schlump, E. Schlag, and R. Ballentin, "SiGe circuits for high bit-rate optical transmission systems," in *Proc. of the 1999 IEEE International Symposium on Circuits and Systems (ISCAS '99)*, vol. 2, pp. 492–495, May 1998.
- [69] B. Wedding, A. Chiarotto, W. Kuebart, and H. Bülow, "Fast adaptive control for electronic equalization of PMD," in *Optical Fiber Communication Conference and Exhibit, 2001 (OFC 2001)*, vol. 2, pp. TuP4.1–TuP4.3, March 1998.
- [70] H. Bülow, F. Buchali, and V. Nicolas, "Dispersion mitigation using a fiber-bragg-grating sideband filter and a tunable electronic equalizer," in *Optical Fiber Communication Conference and Exhibit, 2001 (OFC 2001)*, vol. 3, pp. wdd34.1–wdd34.3, Jan. 2001.
- [71] W. D. Grover, "Forward error correction in dispersion-limited lightwave systems," *IEEE Journal of Lightwave Tech.*, vol. 6, pp. 643–654, Oct. 1988.
- [72] R. W. Chang and J. C. Hancock, "On receiver structures for channel having memory," *IEEE Trans. Inf. Theory*, vol. 12, pp. 463–468, Oct. 1966.
- [73] K. Abend and B. D. Fritchman, "Statistical detection for communication channels with intersymbol interference," *Proc. IEEE*, vol. 58, pp. 779–785, May 1970.

- [74] K. Abend, T. J. Hartley, B. D. Fritchman, and C. Gumacos, "On optimum receivers for channels having memory," *IEEE Trans. Inf. Theory*, vol. 14, pp. 152–157, Nov. 1968.
- [75] Y. Li, B. Vucetic, and Y. Sato, "Optimum soft-output detection for channels with intersymbol interference," *IEEE Trans. Inf. Theory*, vol. 41, pp. 704–713, May 1995.
- [76] Y. Sato, "A method of self-recovering equalisation for multilevel amplitude-modulation systems," *IEEE Trans. Comm.*, vol. 23, pp. 679–682, June 1975.
- [77] D. N. Godard, "Self-recovering equalisation and carrier tracking in two-dimensional data communication systems," *IEEE Trans. Comm.*, vol. 28, pp. 1867–1875, Nov. 1980.
- [78] R. Johnson, P. Schniter, T. J. Endres, J. D. Behm, D. R. Brown, and R. A. Casas, "Blind equalization using the constant modulus criterion: A review," *IEEE Proc.*, vol. 10, pp. 1927–1950, Oct. 1998.