Electrical Signal Processing Techniques in Long-Haul Fiber-Optic Systems

JACK H. WINTERS, SENIOR MEMBER, IEEE, AND RICHARD D. GITLIN, FELLOW, IEEE

Abstract—**The purpose of this paper is to demonstrate the potential for electrical signal processing to mitigate the effect of intersymbol interference in long-haul fiber-optic systems. Intersymbol interference in long-haul fiber-optic systems can severely degrade performance and consequently limit both the maximum distance and data rate. The sources of intersymbol interference include nonlinearity in the laser transmitter, chromatic dispersion in systems operated at wavelengths other than the dispersion minimum of the fiber, polarization dispersion, bandwidth limitations in the receiver, and mode partition fluctuations. It is expected that longer, repeaterless spans made feasible through the use of optical amplifiers will increase the need for the processing techniques described in this paper.**

In this paper we discuss several techniques for reducing intersymbol interference in single-mode fiber systems with single-frequency lasers, and we show which techniques are appropriate at high data rates in direct and coherent detection systems. In particular, we analyze the performance of linear equalization (tapped delay lines), nonlinear cancellation (variable threshold detection), maximum likelihood detection, coding, and multilevel signaling. Our results, for a simulated binary 8 Gbps system, show that simple techniques can be used to substantially reduce intersymbol interference, increasing system margin by several dB. In particular, a 6-tap linear equalizer increases the dispersion-limited distance (due to chromatic or polarization dispersion) by 20% (or reduces the optical power penalty by as much as a factor of two) in direct detection systems, even when the distortion is nonlinear. A novel, but simple, nonlinear cancellation technique (adjusting the decision threshold in the detector based on previously detected bits) can more than *double* the dispersion-limited distance and/or data rate.

I. Introduction

In this paper we demonstrate the potential for electrical signal processing techniques to mitigate the effect of intersymbol interference (ISI) in long-haul fiber-optic systems. Intersymbol interference is a major impairment in long-haul fiber-optic systems that limits both the transmission distance and data rate. Intersymbol interference can have many sources, including laser nonlinearity, mode partition fluctuations, nonideal receiver frequency response characteristics, and chromatic and polarization dispersion. It is the variation in these impairments from device to device, along with the evolution of data technology, that makes intersymbol interference so difficult to predict in future lightwave systems.[1] For example, the advent of optical amplifiers will make longer repeaterless spans feasible, and the processing techniques we describe may be needed to compensate for the overall ISI on such links.

In [1] we described, in qualitative terms, a wide variety of techniques to reduce intersymbol interference in lightwave systems. These included receiver signal processing techniques, such as linear equalization, nonlinear cancellation, and maximum likelihood detection, which can be made adaptive and thus capable of optimizing system performance over a wide range of impairments and device characteristics. Here, we analyze in detail these receiver signal processing techniques in long-haul fiber-optic systems with single-mode fiber and single-frequency lasers.[2] We propose practical methods for implementing these techniques in high data-rate systems, and we show examples of the substantial performance improvement obtainable with these techniques.

First, we discuss the long-haul fiber-optic system and the sources of intersymbol interference. We then consider several receiver signal processing techniques and describe how these techniques can be implemented. These techniques include linear equalization (via a tapped delay line), nonlinear cancellation (using variable threshold detection), maximum likelihood detection, coding, and multilevel signaling. Finally, we report the performance of the techniques with laser nonlinearity, mode partition fluctuations, nonideal receiver frequency response, chromatic dispersion, and polarization dispersion in a computer simulated 8 Gbps system using measured laser and receiver characteristics. Results for this system show that, in some cases, these techniques can increase the system margin by several dB. In particular, a 6-tap linear equalizer increases the dispersion-limited distance (due to chromatic or polarization dispersion) by 20% (or reduces the optical power penalty by as much as a factor of two) in direct detection systems, even when the distortion is nonlinear. A novel, but simple, nonlinear cancellation technique (adjusting the decision threshold in the detector based on previously detected bits) can more than *double* the dispersion-limited distance and/or data rate. The discussion of these results should provide insight into the performance improvement possible in other systems as well.

In Section II we describe the long-haul system. Techniques for reducing intersymbol interference are described in Section III. In Section IV we present results for the performance improvement with these techniques. Finally, a summary and our conclusions are presented in Section V.

II. System

Fig. 1 shows a block diagram of a long-haul fiber-optic system. The nonreturn-to-zero (NRZ) input data stream $x_1(t)$ is filtered [by the transmit filter[3] with frequency response $H_T(f)$] and the filtered signal, $x_2(t)$, modulates a single-frequency (distributed-feedback, DFB) laser, producing the signal $x(t)$ which modulates the electrical field described by (1). Alternatively, the data stream can be used to externally modulate the laser (to avoid laser nonlinearity). The transmitted electric field (optical signal) is given by

$$E = \left(\boldsymbol{a}_x E_x + \boldsymbol{a}_y E_y e^{j\theta}\right) e^{-j\beta z} x(t) \tag{1}$$

where $\boldsymbol{a}_x$ and $\boldsymbol{a}_y$ are unit vectors in the x and y directions, respectively, θ is the angle that determines the signal polarization, β

Paper approved by the Editor for Optical Transmission Systems of the IEEE Communications Society. Manuscript received April 17, 1989; revised August 14, 1989. This paper was presented at the ICC/SUPERCOM'90, Atlanta, GA, April 16-19, 1990.

The authors are with AT&T Bell Laboratories, Holmdel, NJ 07733.

IEEE Log Number 9037857.

[1] In contrast to the usage of the term in voice-band modems, ISI in fiber-optic systems refers to both linear and nonlinear distortion introduced by data symbols other than the one currently being detected.

[2] Equalization techniques for multimode fiber systems have been studied previously for multimode [2] lasers.

[3] This filter is used to reduce the high bandwidth components of the modulating signal and thereby reduce laser nonlinearity (chirp).

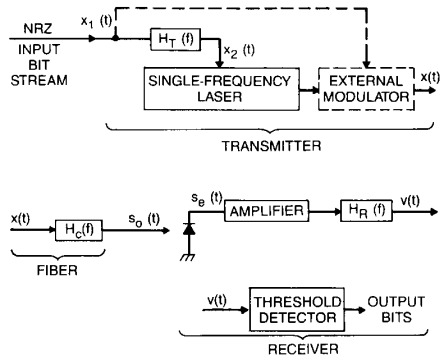


Fig. 1. Block diagram of a long-haul, direct detection, fiber-optic system.

is the propagation constant, and z is the direction of propagation. For direct modulation,

$$x(t) = \sqrt{P(t)}\, e^{j(\phi_c(t) + \angle x_2(t))}\, e^{j\omega_c t} \tag{2}$$

where ω_c is the (radian) lightwave frequency, $P(t)$ is the optical power, and $\angle x_2(t)$ is the phase angle of $x_2(t)$. The phase variation $d\phi_c(t)/dt$ is referred to as chirp, while the amplitude variation in $P(t)$ versus $\sqrt{|x_2(t)|}$ (since the laser output power is proportional to the input current) is referred to as relaxation oscillation. We refer to these two effects as laser nonlinearities. With external modulation, the complex waveform is given by

$$x(t) = e^{j(\omega_c t + \phi(t))} \sqrt{|x_1(t)|}\, e^{j \angle x_1(t)} \tag{3}$$

where $\phi(t)$ is the phase noise of the laser. Note that the laser power, $|x(t)|^2$, is proportional to the electrical signal $x_1(t)$. The transmitted optical signal passes through a fiber with frequency response $H_c(f)$, which may have chromatic [3] and polarization [4] dispersion. Chromatic dispersion will have a significant impact on system performance if the laser frequency is different from the dispersion minimum of the fiber (which is 1.3 μm in a standard fiber and 1.55 μm in a dispersion-shifted fiber). In this case, the dominant chromatic dispersion is linear delay distortion, and the frequency response of the fiber is given by [3]

$$H_c(f) = e^{-j\alpha f^2}, \qquad \alpha = \pi D(\lambda) \frac{\lambda^2}{c} L \tag{4}$$

where L is the fiber length, $D(\lambda)$ is the linear delay coefficient, and λ is the wavelength ($= 2\pi c/\omega_c$). For example, for a 1.55 μm signal in a standard fiber, $D(\lambda) = 17$ ps/km/nm. At the receiver, the optical signal, $s_o(t) = h_c(t) \circledast x(t)$ [where $\circledast$ denotes convolution and $h_c(t)$ is the Fourier transform of $H_c(f)$], is converted to an electrical signal by a photodetector (usually an avalanche photodiode APD).

With direct detection, as shown in Fig. 1, the electrical signal, $s_e(t)$, is proportional to $|s_o(t)|^2$. Alternatively, coherent detection can be used (see [1]), where the received optical signal is mixed with a local oscillator optical signal (at approximately the same frequency as the received signal) to generate an IF electrical signal whose *envelope* is proportional to $s_o(t)$.

Noise in the electrical signal is due to both shot noise in the received signal and thermal noise in the receiver (i.e., the preamplifier). With direct detection, noise in the electrical signal is primarily due to thermal noise in the receiver because of the limitations of present-day detectors. With coherent detection, the shot noise is due to both the received and local oscillator signals. If the local oscillator signal is strong enough, the shot noise dominates the thermal noise [5], and it is known that the high-intensity shot noise can be modeled as additive, Gaussian noise [6]. Thus, we will assume that the only noise in the electrical signal is additive, Gaussian noise (as is true in practical receivers). Consequently, we do not need to consider the received optical signal from a quantum physics point of view, i.e., a system model that contains shot noise is not needed (in contrast to earlier papers on lightwave equalization, e.g., [7]).

Polarization dispersion can be characterized mainly in terms of first [4] and second order (in frequency) effects [8]. The first-order effect is a delay in the signal in one polarization relative to the delay in the signal in the other polarization. Thus, with direct detection and first-order polarization dispersion effects, since signals in orthogonal polarizations add powerwise at the receiver, the electrical signal is given by the *linear* combination of the individually detected signals,

$$s_e(t) = \alpha_1 \left(|s_o(t)|^2 + \alpha |s_o(t+\tau)|^2 \right) \tag{5}$$

where α_1 is the conversion constant between the optical and electrical signals, α is the ratio of the signal strengths in the two polarizations, and τ is the time delay between propagation in the two polarizations. With coherent detection, the receiver is polarization sensitive, and the effect of polarization dispersion depends on the receiver technique used. Because of the wide variety of coherent system receiver techniques (e.g., [9] and [10]), in this paper we note only that the signals in the two different polarizations add *powerwise* (with both direct and coherent detection) and will not further discuss polarization dispersion in coherent systems (see [4] for the effects of first-order polarization dispersion on different modulation formats). The second-order polarization dispersion effects are linear delay distortion that differs in sign in the two polarizations and a cross-coupling of power between the polarizations at the receiver. Since linear delay distortion is considered in this paper (for chromatic dispersion in the fiber), we will only present results for the first-order effects of polarization dispersion in this paper (see Section IV).

Next, the electrical signal is amplified and filtered to increase the signal-to-noise ratio (where the receiver frequency response $H_R(f)$ is due to both the frequency characteristics of the receiver filter and the amplifier), producing an output signal $v(t) = h_R(t) \circledast s_e(t)$. This signal is detected by comparing the signal level, during a short period of time at the peak opening of the eye of the signal (e.g., 20 ps for an 8 Gbps data rate), to a decision threshold.

Intersymbol interference in the detected signal can be caused by laser nonlinearity, chromatic dispersion, polarization dispersion, and nonideal receiver frequency response. The laser nonlinearity and the receiver frequency response will vary among devices, and polarization dispersion will vary slowly over time (e.g., on the order of hours). Chromatic dispersion is reasonably fixed for a given length of fiber, but its effect on system performance depends on laser nonlinearity and receiver frequency response, which can vary. This suggests that adaptive signal processing structures may be required.

In addition, mode partition fluctuations [11], [12] in combination with chromatic dispersion can cause random intersymbol interference. Mode partition fluctuation refers to the random amplitude variation in the side mode (or modes) of a nearly single-frequency laser. With chromatic dispersion, the side mode has a different propagation delay than the main mode, and thus produces intersymbol interference, which may extend over several bits. Higher side mode amplitudes mainly occur at low-to-high transitions in directly modulated lasers, but may also occur while the laser output power is constant. Although the amplitude of the side mode is usually small, occasionally the amplitude may be large enough that the resulting intersymbol interference closes the eye at the receiver, resulting in an error rate floor. The amplitude may be high for just one bit, but in some devices, at Gbps data rates, may remain high for several bits. The amplitude distribution of the side mode is difficult to characterize and varies substantially with laser operating conditions and among devices.

A key issue in the effectiveness of intersymbol interference compensation techniques is the linearity[4] of the intersymbol inter-

[4]That is, whether or not the intersymbol interference in the electrical signal at the symbol detector can be considered as a superposition of intersymbol interference from symbols other than the symbol to be detected.

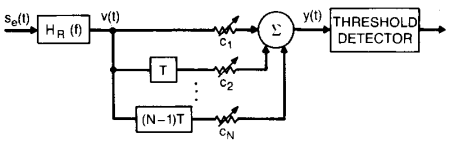


Fig. 2. Tapped delay line for equalization with N taps.

ference. The linearity of the intersymbol interference determines whether linear or nonlinear equalization techniques are more appropriate. Below, we consider the linearity of the various impairments.

Linear distortions include the receiver frequency response and polarization dispersion (first-order effect). Laser nonlinearity is, of course, a nonlinear distortion. Chromatic dispersion (and also the second-order polarization dispersion effect of linear delay distortion) is a linear distortion in the optical fiber; however, whether this distortion is linear at the detector [i.e., $v(t)$], depends on the transmitter and receiver techniques used. Specifically, chromatic dispersion appears as linear distortion in a coherent receiver [13], [14], but as long as the laser linewidth is smaller than the data rate (as is the trend for high data rate systems using state-of-the-art DFB lasers—see [7] for multimode laser systems where chromatic dispersion is a linear distortion), this form of dispersion appears as *nonlinear* distortion in a direct detection system [3]. However, even though the distortion is, in general, nonlinear, linear equalization can still be useful (see [29], Appendix, and Section IV).

III. Compensation Techniques

A. Linear Equalization

To compensate for linear distortion, a linear equalizer (transversal filter) can be used between the receiver filter and the detector. In particular, we will consider the (analog) tapped delay line implementation of the equalizer[5] with N taps with tap weights c_i, $i = 1, N$, as shown in Fig. 2. Note that the equalizer output signal $y(t)$ is given by

$$y(t) = \sum_{j=1}^{N} c_j v\big(t - (j-1)T\big). \tag{6}$$

Since the receiver filter bandwidth is usually much less than the signal bandwidth (to reduce receiver noise), the tap spacing (T) need only be the symbol period, i.e., the equalizer is a synchronous linear equalizer (a fractionally spaced equalizer is not needed in this case to reduce the ISI, although such an equalizer can reduce the equalizer noise enhancement—see Section III-A-2). At high data rates, symbol delays can be implemented by a short transmission line (e.g., less than 4 cm at 8 Gbps), and the weights can be implemented by a variable-gain amplifier. An even simpler structure for the tapped delay line is shown in Fig. 3 (as suggested by Kasper and Mizuhara), where a transmission line is used to produce the delays. Negative weighting is achieved by weighting the inverted input signal, and the weights can be implemented by single transistors rather than amplifiers.

Since, in many cases (e.g., with laser nonlinearities such as chirp and relaxation oscillation), most of the intersymbol interference is due the symbols preceding the detected symbol, there may be more taps for the precursor symbols than for the future symbols. This is discussed further in Section IV. Other performance issues include weight adjustment techniques, the noise enhancement of the equalizer, and the linearity of the distortion. These issues are discussed below.

1) Weight Adjustment Techniques: The weights can be preset in the factory or set by craftsman during installation. With manual adjustment, the weights can be set to minimize the bit error rate or maximize the eye opening (system margin). However, manual adjustment to optimize these parameters may be difficult when there

[5]Such an equalizer has been implemented at 8 Gbps [15].

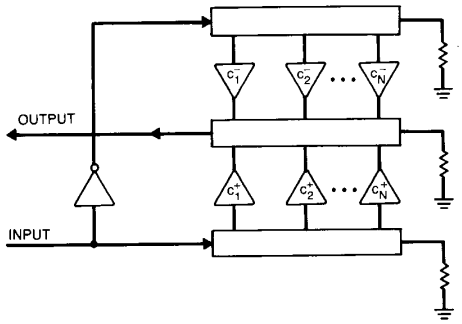


Fig. 3. A possibly simpler implementation of the tapped delay line, where

$$c_i^+ = \begin{cases} c_i & \text{if } c_i > 0 \\ 0 & \text{otherwise} \end{cases}$$

and

$$c_i^- = \begin{cases} -c_i & \text{if } c_i < 0 \\ 0 & \text{otherwise.} \end{cases}$$

are more than a few weights, or when the optimum weights vary significantly among receivers (i.e., with variations in device characteristics). In addition, manual adjustment cannot be used to compensate for variations in devices over time (e.g., with temperature) or to track polarization dispersion. In these cases, some type of adaptive algorithm must be used.

Two celebrated adaptive algorithms are the least-mean-square (LMS) algorithm of Widrow [16] and the zero-forcing algorithm of Lucky [17], [18]. Fig. 4 shows adaptive equalizers using these two algorithms. For the least mean-square-error algorithm, the weight update equation is given by

$$c_j^{k+1} = c_j^k + \Delta \epsilon_k v_{k-j}, \qquad k = 1, 2, \cdots, \qquad j = 1, 2, \cdots, N \tag{7a}$$

or in vector notation

$$\boldsymbol{c}_{k+1} = \boldsymbol{c}_k + \Delta \epsilon_k \boldsymbol{v}_k \tag{7b}$$

where c_j^k is the jth tap weight during the kth symbol period, Δ is a small constant that controls the magnitude of the weight adjustment, v_k is the input signal to the equalizer during the kth symbol interval, and ϵ_k is the error in the equalizer output signal y_k given by

$$\epsilon_k = (I_k - y_k) \tag{8}$$

where I_k is the kth output symbol. In (7b), $\boldsymbol{c}_k$ is the vector of tap weights and $\boldsymbol{v}_k$ is the vector of samples in the delay line at the kth sampling instant. The essential function of the algorithm is to multiply (or correlate) the error signal by the received line samples. Since the error rate at the receiver is assumed to be negligible in long-haul systems ($<10^{-9}$), the detected output symbol is virtually the same as the actual transmitted symbol. In this paper, we assume that it is the same. For the zero-forcing algorithm, the weight update equation is given by

$$c_j^{k+1} = c_j^k + \Delta \epsilon_k I_{k-j}, \qquad k = 1, 2, \cdots \tag{9a}$$

$$\boldsymbol{c}_{k+1} = \boldsymbol{c}_k + \Delta \epsilon_k \boldsymbol{I}_k, \tag{9b}$$

and the error is correlated with the data decisions. Neither of these algorithms necessarily adjusts the weights to minimize the bit error rate or to provide maixmum system margin. The mean-square-error algorithm minimizes the variance of the error signal, while the zero-forcing algorithm minimizes the peak distortion. The zero-forcing algorithm is only effective at minimizing peak distortion when the unequalized eye is open. With modest amounts of linear

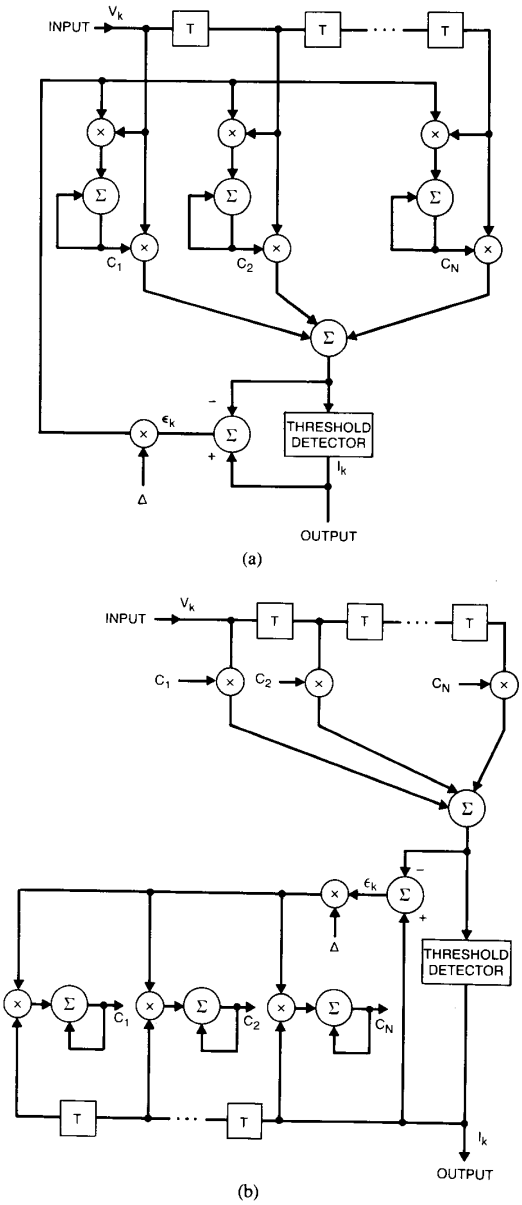


Fig. 4. Adaptive linear equalizer using (a) the mean-square-error algorithm and (b) the zero-forcing algorithm.

ISI and thermal noise, both equalizers produce distortion-free outputs and provide performance close to the optimum.

At the high data rates of long-haul systems, generating analog samples of signals is very costly and, hence, may not be practical.[6] Thus, single bit accuracy samples (± 1) must be used[7] whenever possible. Two algorithms that provide single-bit accuracy are the

[6]Since the analog sample of the error ϵ_k gives the eye opening, it is also useful for determining the system margin and showing degradations before bit errors occur. Thus, analog sampling of the error may be worthwhile, even if it is not required for weight adaptation.

[7]For the remainder of this section, plus Sections III-B and III-C, we will restrict our attention to binary signaling (on–off keying). The analysis can easily be extended to M-level signaling.

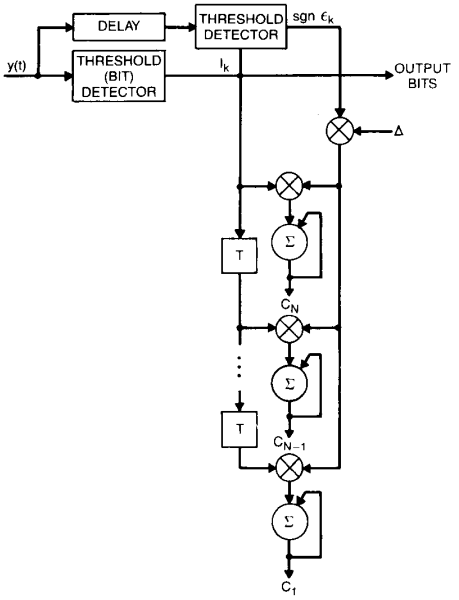


Fig. 5. Implementation of the zero-forcing adaptive algorithm.

sgn-sgn mean-square-error algorithm [17]

$$c_j^{k+1} = c_j^k + \Delta \operatorname{sgn}(\epsilon_k) \operatorname{sgn}(v_{k-j}), \tag{10a}$$

$$\boldsymbol{c}_{k+1} = \boldsymbol{c}_k + \Delta \operatorname{sgn}(\epsilon_k) \operatorname{sgn}(\boldsymbol{v}_k), \tag{10b}$$

and the modified zero-forcing algorithm [17], [19]

$$c_j^{k+1} = c_j^k + \Delta \operatorname{sgn}(\epsilon_k) I_{k-j} \tag{11a}$$

$$\boldsymbol{c}_{k+1} = \boldsymbol{c}_k + \Delta \operatorname{sgn}(\epsilon_k) \boldsymbol{I}_k. \tag{11b}$$

Quantizing the signal samples reduces the rate of convergence of the algorithms (which is not a major concern, since channel impairments should change very slowly with time—on the order of hours or longer). Of more concern, especially for the sgn-sgn LMS algorithm, is the steady-state weights. Since the *direction* of the correction term (sgn $\boldsymbol{v}_k$) is different than for the LMS algorithm ($\boldsymbol{v}_k$), the weights may settle at a different value. However, for the quantized version of the zero-forcing algorithm (11), the direction of the correction term ($\boldsymbol{I}_k$) is the same as for the unquantized algorithm, and the algorithm converges to the same weights as the continuous version (9) (when the eye of the received signal is open [neither algorithm works when the eye is closed]). Quantizing the signal samples does make the equalizer more sensitive to DC offsets in the signals, however. (This is discussed further in Section IV.)

Fig. 5 shows a possible implementation of the quantized (discrete) zero-forcing adaptive algorithm. The detected bits, the I_k's, are used to adjust the threshold of a second detector that compares the received signal samples to the predicted levels for ones and zeros (i.e., determines the signum of the error). An even simpler technique [20] is to set the decision threshold in the second detector to that for the predicted level for a one, and to only use the signum of error when a one is detected. This eliminates the need to vary the decision threshold in the second detector. (However, this equalizer is even more sensitive to dc offsets in the signal, as shown in Section IV). Once the sgn (ϵ_k) and I_k values have been determined, the multiplications (by ± 1, i.e., sign inversion) and additions required for weight adaptation can be done at a much slower rate (e.g., by a microprocessor). Thus, the only costly hardware for the adaptive algorithm is the second detector.

It is well known that the (unquantized) mean-square-error algorithm gives better performance (i.e., lower bit error rate or higher

system margin) than the zero-forcing algorithm because it minimizes both the intersymbol interference and thermal noise, while the zero-forcing algorithm only minimizes the intersymbol interference (the noise enhancement of the equalizer is studied below). Furthermore, unlike the zero-forcing algorithm, when supplied with a training sequence the mean-square-error algorithm also works when the eye of the received signal is closed. However, the implementation practicalities of Gbps operation preclude analog sampling of the signal (as required for weight adaptation by the LMS algorithm) and, as mentioned above, the quantized version of the mean-square-error algorithm (10) is not guaranteed to converge to the same weights as the continuous version (7). In fact, with the eye open, the quantized version of the mean-square-error algorithm (10) is identical to the quantized version of the zero-forcing algorithm (11), since $I_{k-j} = \operatorname{sgn}(v_{k-j})$. Thus, with the eye open, the quantized version of the mean-square-error algorithm has the same performance (in terms of convergence and steady-state error rate) as the quantized zero-forcing algorithm. However, the quantized mean-square-error algorithm requires samples of the received signal (before equalization) in addition to the samples of the error signal required by the zero-forcing algorithm. Thus, the quantized version of the mean-square-error algorithm requires twice as many detectors, yet gives the same performance as the quantized zero-forcing algorithm, when the eye is open. Furthermore, when the eye is closed, the quantized mean-square-error algorithm may not work. Thus, the quantized zero-forcing algorithm is preferred over the quantized mean-square-error algorithm for our application, when we know *a priori* that the eye is open. However, when we know *a priori* that the eye is sometimes closed, the continuous version of the mean-square-error algorithm (7) may be used to track variable impairments. Note that while the mean-square-error algorithm may be difficult (or costly) to implement at high data rates because analog samples are required, the performance results obtained using this algorithm can be used as a lower bound on the performance improvement that can be obtained the manual adjustment of the weights to minimize the bit error rate or maximize system margin, when the impairments are fixed.

The performance criterion we will consider in this paper is the optical signal power penalty due to intersymbol interference (i.e., the increase in received optical signal power required to maintain the same eye opening with intersymbol interference), which can be derived from the minimum eye opening over all input bit sequences. The minimum eye opening is the minimum sampled signal value for a "1" minus the maximum sampled signal value for a "0," with no noise at the receiver. Thus, if the difference between the signal levels for a "1" and a "0" without ISI is Y, the minimum eye opening (in percent) is given by

$$\text{eye opening} = \frac{\min_k (y_k/I_k = 1) - \max_k (y_k/I_k = 0)}{Y} \cdot 100 \tag{12}$$

or

$$\text{eye} = \min_{\substack{k,i \\ I_k=1 \\ I_i=0}} \left(\frac{y_k - y_i}{Y} \right) \cdot 100. \tag{13}$$

The optical power penalty is given by [3]

$$\text{penalty} = \begin{cases} 10\log_{10}(\text{eye}/100)\ \text{dB} & \text{for direct detection} \\ 20\log_{10}(\text{eye}/100)\ \text{dB} & \text{for coherent detection} \end{cases}, \tag{14}$$

since the received current is proportional to the optical power with direct detection and the received current is proportional to the magnitude of the optical field with coherent detection.

2) Noise Enhancement: Since the linear equalizer combines the weighted received signals, the noise level in the output signal may be increased if the signal is amplified over a range of frequencies (to compensate for attenuation over the frequency range). In this paper, we will not consider the absolute level of the noise at the receiver, but only consider the relative increase in the noise level due to equalization (referred to as noise enhancement). If we assume that, at the output of the receiver filter, samples of the noise taken every T seconds are independent, zero-mean, Gaussian random variables with variance σ^2, then the noise enhancement of the equalizer is just $\sum_{j=1}^{N} c_j^2$. A measure of the relative received signal-to-noise ratio (SNR), reflecting both ISI and noise, is given by (in percent), from (13),

$$\rho = \min_{\substack{k,i \\ I_k=1 \\ I_i=0}} \left[\frac{(y_k - y_i)}{Y\sqrt{\sum_{j=1}^{N} c_j^2}} \right] \cdot 100, \tag{15}$$

with the optical power penalty being $10\log_{10}(\rho/100)$ dB for direct detection and $20\log_{10}(\rho/100)$ dB for coherent detection.

The noise enhancement can be reduced or eliminated by using a fractionally spaced equalizer (decision-feedback equalization or linear cancellation can also be used [1]). In general, a tap spacing of $T/2$ is adequate for reducing the noise enhancement as much as possible. However, with chirp, tap spacings may need to be closer than the reciprocal of the chirp bandwidth, i.e., tap spacings of less than $T/5$ may be required to significantly reduce the noise enhancement. Therefore, a fivefold increase in the number of taps could be required to eliminate ISI without noise enhancement. To simplify our analysis and keep the number of taps small, we will only study synchronous (tap spacings of T) equalizers (see Section IV).

B. Nonlinear Cancellation

As shown in Section II, there are cases when the distortion is nonlinear and, therefore, nonlinear techniques must be used to significantly reduce distortions. Here, as in [1], we consider the use of nonlinear cancellation (see also [21]). Nonlinear cancellation can be implemented as follows. Using knowledge of previously detected bits and, perhaps, estimates of bits to be detected, the decision threshold in the detector is adjusted up or down to be halfway between the expected signal levels for each bit to be detected. If the adjustment is a linear sum of the previous and estimated bits, then the technique is just linear cancellation [22], [23] (or decision-feedback equalization if only previously detected bits are used). Otherwise, a lookup table (or explicit computation[8]) may be used, with 2^{N-1} entries for $N-1$ previously detected and estimated bits (a total of N bits are used to determine the data bit) to provide nonlinear cancellation. At the high data rates of the long-haul systems, the size of the lookup table may limit $N-1$ to just a few bits (fortunately, this is typically the extent of the ISI in high-speed lightwave systems).

Since estimating bits to be detected requires an additional detector (or even additional interference reduction techniques if the eye is closed), the most practical technique is to adjust the decision threshold based on previously detected ("decided") bits only, and use an analog tapped delay line (which has been implemented at 8 Gbps [15]) to reduce distortion caused by "future" bits. This canceller is shown in Fig. 6(a) where N_1 decided bits are used for nonlinear cancellation and a tapped delay line with $N_2 + 1$ taps is used for the N_2 future symbols plus the data bit to be detected.[9] Thus, at the detector,

$$y(t) = \sum_{i=1}^{N_2+1} c_i v(t - (i-1)T), \tag{16}$$

[8]In practice, the lookup table could be implemented by a switch with the previously detected and estimated bits controlling which one of 2^{N-1} voltage levels is connected to the threshold.

[9]This is similar to the structure of the RAM-DFE [24] used at much lower data rates, which because of the lower data rate, can use a RAM for the lookup table and add the feedback signal to the signal to be detected.

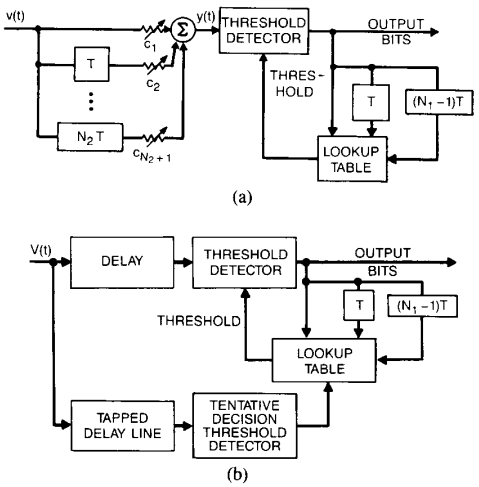


Fig. 6. Nonlinear canceller using N_1 previous bits and (a) a tapped-delay line for reduction of intersymbol interference from N_2 future bits and (b) tentative bit decisions for N_2 future bits.

and the decision rule is

$$I_k = \begin{cases} 1 & \text{if } y(t_0) > y_0 + f(I_{k-1}, \cdots, I_{k-N_1}) \\ 0 & \text{otherwise} \end{cases} \tag{17}$$

where y_0 is the unadjusted decision threshold, t_0 is the sampling time, and $f(\cdot)$ is the output of the look-up table which provides an estimate of the (nonlinear) intersymbol interference. With severe intersymbol interference, it may be possible that, for a given set of N_1 decided bits, the level for a "1" is always (without noise) lower than that for a "0" (the reverse of the usual case). Thus, a more robust nonlinear cancellation technique reverses the decision rule (17) (or inverts the detector output bits) when these sets of N_1 decided bits occur. Note that if, for a given set of N_1 decided bits, the level for a "1" can be (without noise) both above or below that for a "0,"[10] nonlinear cancellation cannot produce error-free output bits even without noise. Now, if the intersymbol interference is from a combination of the interference from several previous bits, with severe ISI, an overlapping of the levels for "1's" and "0's" for a given set of N_1 bits may be more likely to occur than reversed, but nonoverlapping, levels for "1's" and "0's". However, with intersymbol interference predominately from 1 or 2 decided bits (as is often the case in lightwave systems), reversed, but nonoverlapping, levels for "1's" and "0's" with severe ISI may be more likely. This type of nonlinear canceller can only reduce nonlinear distortion that is caused by the N_1 decided bits. However, this is the main form of the distortion with chirp (since laser chirp produces an increase in the fall time of the pulse).

The adaptive algorithms for nonlinear cancellation are a generalization of the algorithms for linear equalizer updating (see Section III-A-1) and [21]).

In the remainder of the paper, we will consider nonlinear cancellation using both future as well as decided bits, rather than using a tapped delay line for the future bits [see Fig. 6(b)]. This simplifies our analysis and results, and gives an upper bound on the performance improvement with a tapped delay line for future bits.

The performance of nonlinear cancellation is given by the minimum distance between sampled signal values for a "1" and a "0," given the same N_1 decided and N_2 future bits. That is, for nonoverlapping levels for "1's" and "0's" for each set of N_2 future and N_1 decided bits, the minimum percent eye in the detected signal is given by

$$\text{eye} = \min_{\substack{k,i \\ I_k=1 \\ I_i=0 \\ \{z_k\}=\{z_i\}}} \left[\frac{|y_k - y_i|}{Y} \right] \cdot 100 \tag{18}$$

where $\{z_j\}$ is the set $\{I_{j-N_1}, \cdots, I_{j-1}, I_{j+1}, \cdots, I_{j+N_2}\}$, which is the set of the $N-1$ $(= N_1 + N_2)$ bits used in the lookup table. Note that if

$$\min_{\substack{k,i \\ I_k=1 \\ I_i=0 \\ \{z_k\}=\{z_i\}}} [y_k - y_i] \le 0$$

for some sets of N_2 future and N_1 decided bits, then the decision regions in the detector are reversed for these sets.[11]

Nonlinear cancellation will fail (i.e., the error rate will be nonzero even without noise at the receiver) if both

$$\min_{\substack{k,i \\ I_k=1 \\ I_i=0 \\ \{z_k\}=\{z_i\}}} [y_k - y_i] \le 0$$

and

$$\max_{\substack{k,i \\ I_k=1 \\ I_i=0 \\ \{z_k\}=\{z_i\}}} [y_k - y_i] \ge 0,$$

i.e., the levels for "1's" and "0's" overlap for some set of N_1 decided and N_2 future bits. To see where this could occur, consider the simple example of direct detection where the sampled values of the optical signal with intersymbol interference are given by (using baseband notation)

$$s_o(kT) = I_k - 0.5I_{k-1}. \tag{19}$$

Thus, neglecting the receiver filter and any noise,

$$y_k = s_e(kT) = |s_o(kT)|^2 = I_k^2 - I_k I_{k-1} + 0.25I_{k-1}^2, \tag{20}$$

or

$$y_k = \begin{cases} 0, & I_{k-1} = 0, \quad I_k = 0 \\ 1, & I_{k-1} = 0, \quad I_k = 1 \\ 0.25, & I_{k-1} = 1, \quad I_k = 0, 1. \end{cases} \tag{21}$$

Note that when $I_{k-1} = 1$, the signal level y_k is independent of I_k, and, therefore, the output levels cannot be separated and nonlinear cancellation will not work. However, we can see that I_k can be determined from y_{k+1} with a 3-level detector, i.e., a reasonable decision rule is

$$I_k = \begin{cases} 1 & \text{if } 0.125 \le y_{k+1} \le 0.625 \\ 0 & \text{otherwise.} \end{cases} \tag{22}$$

This leads us to consider the optimum detection scheme—maximum likelihood detection—in the next subsection.

[10] This can occur if the nonlinear component of the ISI depends on bits in addition to the N_1 decided bits, e.g., the bit to be detected.

[11] For the nonlinear canceller of Fig. 6(a), the relative signal-to-noise ratio ρ is given by (18) divided by

$$\sqrt{\sum_{j=1}^{N_2+1} c_j^2},$$

the noise enhancement of linear equalizer. Note that there is no noise enhancement by the decided bits. The equalizer tap weights can be adjusted by the zero-forcing or mean-square-error algorithm, as described in Section III-A-1).

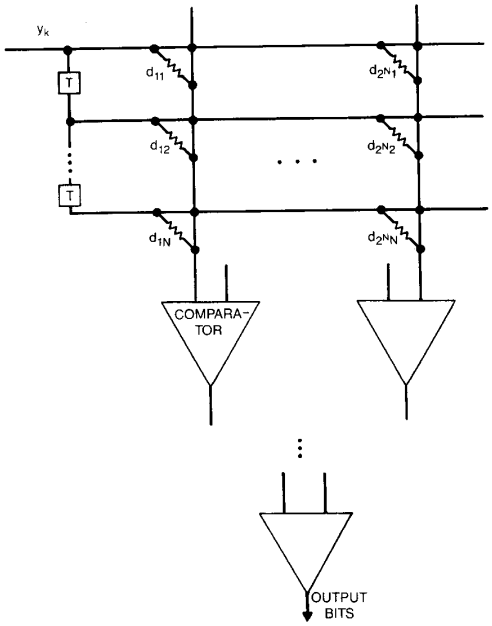


Fig. 7. Maximum likelihood detector using N signal samples to determine blocks of N bits.

C. Maximum Likelihood Detection

Maximum likelihood detection (MLD) is the optimum detection technique in that it minimizes the error probability for determining a bit (or bit sequence), given N received signal samples. It can be complex to implement for large N, although it is useful as a bound on performance. There are techniques to implement simplified versions of maximum likelihood detection, however, that may be practical at high data rates, if N is small. Here we present two such techniques. As before, we assume that the received signal levels and intersymbol interference are deterministic (given the bit pattern), and that the only source of noise is additive Gaussian noise (thermal noise with direct detection and high-intensity shot noise with coherent detection).

In the first technique, which is a block-oriented maximum likelihood detector, we determine the bits in a block of length N, based on N consecutive signal samples where the signal is corrupted by additive Gaussian noise. Specifically, we compare blocks of N consecutive received signal samples to each of 2^N possible (stored) signal sample vectors (corresponding to the 2^N possible bit sequences). With the additive Gaussian noise model, the detected bits correspond to the vector that has the closest Euclidean distance to the received vector. That is, we find the closest signal vector that maximizes the correlation between the received vector and the allowable (transmitted) signal vectors,

$$\max_{l} \left\{ \sum_{j=1}^{N} 2 d_{lj} y_{k-j} - d_{lj}^2 \right\} \tag{23}$$

where l is the set of the 2^N stored signal vectors and d_{lj} is the jth value of the lth signal vector. Fig. 7 shows a block diagram of a neural-network based implementation of this technique, which is similar to that presented in [25] for classification problems and in [26] for Viterbi decoding. The received signal vector is multiplied by each of the 2^N stored signal vectors using resistors, and the summed currents are compared. The bank of $(2^N - 1)$ comparators determines which of the summed currents is the largest, i.e., the closest signal vector, and outputs the corresponding bit stream. (Note that each comparator outputs not only the largest input signal, but the corresponding bit stream as well.) The operation of this

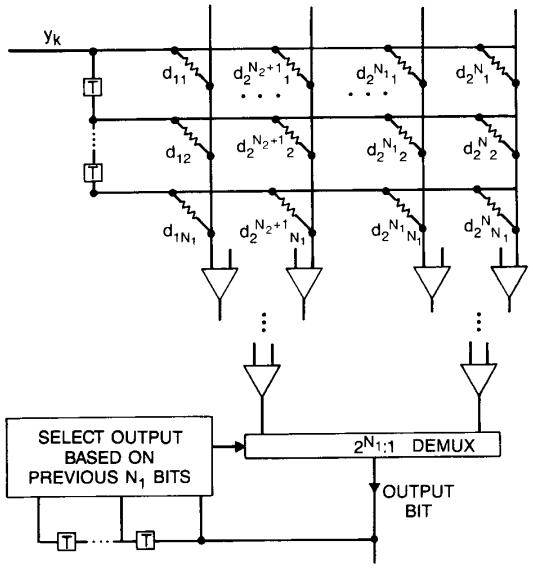


Fig. 8. Maximum likelihood detector using N_1 previous bits and an $N_2 + 1$ bit signal vector to determine a single bit.

detector is at $1/N$ times the data rate. (Note that, for $N = 1$, the technique is just bit-by-bit detection.) The stored signal vectors (the d_{lj}'s) are chosen to minimize the error rate, a task which can become very complicated with large N and severe intersymbol interference.

The performance of this detector can be determined from the minimum Euclidean distance between received vectors with different transmitted bits. That is, the minimum eye opening is

$$\text{eye} = \min_{\substack{k,i \\ \{I_{k-(N-1)} \cdots I_k\} \neq \{I_{i-(N-1)} \cdots I_i\}}} \left[\frac{\sqrt{\sum_{j=1}^{N} (y_{k-j} - y_{i-j})^2}}{Y} \right] \cdot 100. \tag{24}$$

The performance of this detector is degraded by the edge effects of ISI. That is, since the MLD makes decisions on block of N bits from N signal samples of these bits, bits at the edge of the block are more likely to be detected in error because of ISI from bits outside the block.

A second, more complex, technique uses a sliding window, with N_1 previously detected bits used to determine the state (i.e., one of the 2^{N_1} possible cases that exist prior to transmission of the next $N_2 + 1$ signal samples), and $N_2 + 1$ signal samples used to determine the detected bit ($N = N_1 + N_2 + 1$). Specifically, the detector calculates the Euclidean distance between the received signal vector of length $N_2 + 1$ and each of the 2^{N_2+1} stored signal vectors to determine the stored signal vector that is closest to the received signal vector and outputs the bit corresponding to the first bit in that stored signal vector. There is a separate set of 2^{N_2+1} stored signal vectors for each of the possible 2^{N_1} states, for a total of $2^{N_1+N_2+1}$ (or 2^N) signal vectors. Fig. 8 shows a block diagram of this technique. The problem of edge effects of ISI on the detected bit is reduced with this technique, because signal samples for the bits adjacent to the detected bit are used. Thus, the detector of Fig. 8 will have better performance than the detector of Fig. 7. In addition, the detector of Fig. 8 requires fewer comparators and a smaller size for the stored signal vectors. However, this detector

must make bit decisions at the data rate, i.e., the circuitry must be N times faster than that of Fig. 7.

The performance of the sliding window detector can be determined from the minimum Euclidean distance between received signal vectors with the same N_1 decided bits and different first bits, followed by any combination of N_2 future bits. That is, the minimum eye opening is

$$\text{eye} = \min_{\substack{k,i \\ I_k \neq I_i \\ \{I_{k-N_1} \cdots I_{k-1}\} \neq \{I_{i-N_1} \cdots I_{i-1}\}}} \left[\frac{\sqrt{\sum_{j=0}^{N_2+1} (y_{k-j} - y_{i-j})^2}}{Y} \right] \cdot 100. \quad (25)$$

Both the detectors of Figs. 7 and 8 can be made adaptive by channel estimation (when the detected bits are virtually error free). In addition, these maximum likelihood detectors can be reduced in size and made adaptive through the use of neural network techniques, in particular, with a 3-layer feedforward network using backpropagation for adaptation [27].

Results for maximum likelihood detection are presented in Section IV, where only the second version of the detector is analyzed.

D. Coding

Coding can be used to increase system margin by allowing higher raw (before decoding) bit error rates while maintaining the same output bit error rate. With respect to system margin, the performance improvement with coding depends on the decrease in system margin versus the increase in data rate due to coding. Specifically, the increase in data rate (for the binary system we are discussing) due to coding must result in a smaller penalty than the coding gain in order for coding to be useful. Thus, as a first step in studying the effect of coding, we need to determine the increase in penalty with data rate. For example, in [28] it was shown that coding, which increased the data rate by 4%, reduced the error rate to 311 P_e^2, or from 2×10^{-6} to 10^{-9}, corresponding to an increase in system margin of 1.1 dB at a 10^{-9} bit error rate. Therefore, in this example, the decrease in system margin (increase in optical power penalty, e.g., as determined from the increase in the channel impairments with data rate) with a 4% increase in data rate must be less than 1.1 dB for coding to be useful. It should be noted, as in [1], that coding affects the required electrical signal-to-noise ratio, with the optical power penalty decrease only half the electrical signal-to-noise ratio improvement (in dB) for direct detection systems, but equal to the electrical signal-to-noise ratio improvement for coherent detection systems. Thus, we would expect that coding is more likely to be useful in coherent systems.

Coding can also be used to reduce error rate floors (e.g., as caused by mode partition fluctuations). With an error rate floor, the system margin is not important, and the increase in data rate due to coding may not significantly increase the error rate floor. Thus, it is the reduction in error rate with coding that is important. For example, the 4% overhead code [28] discussed above can decrease an error rate floor of 10^{-6} to below 10^{-9}. However, for simple, single-error correction codes (such as in [28]) to be effective, the error rate floor must be caused by random errors. If these codes are implemented by multiplexing several lower speed encoders/decoders, however (as may be required for multigigabit per second data rates), then bursts of several bits in error can still be corrected. Note that since coding corrects errors after bit detection, coding is equally effective in reducing error rate floors with both coherent and direct detection systems.

E. Multilevel Signaling

Multilevel signaling can also be used to increase system margin when intersymbol interference is present. As compared to on–off keying (with 2-levels), multilevel signaling with M levels decreases the symbol rate by a factor of $\log_2 M$, which reduces the amount of dispersion and decreases the noise power in the detector by $\log_2 M$. However, multilevel signaling also decreases the eye opening by a factor of at least $M - 1$ even without intersymbol interference. Thus, without intersymbol interference, for a given symbol error rate, multilevel signaling introduces an optical power penalty of $10 \log_{10}((M-1)/\sqrt{\log_2 M})$ dB in direct detection systems and $20 \log_{10}((M-1)/\sqrt{\log_2 M})$ dB in coherent systems. Thus, multilevel signaling ($M > 2$) can only be useful in reducing the optical power penalty due to intersymbol interference, when the optical power penalty with 2-level signaling exceeds the above values. In these cases, we need to compare the optical power penalty to 2-level signaling to that with M-level signaling at $1/\log_2 M$ times the symbol rate. For example, 4-level signaling could only reduce the optical power penalty of a 2-level signaling system if the penalty exceeded 3.3 and 6.5 dB in direct detection and coherent systems, respectively. Thus, we would expect that multilevel signaling is more likely to be useful in direct detection systems. Note that this is opposite to the conclusion with coding.

TABLE I
LASER PARAMETER VALUES (USED WITH (1-3) OF [29] AND (2) OF [34])

$a = 2.5\times 10^{-16}$ cm^2	$v_g = 7.89\times 10^9$ cm/sec
$\Gamma_v = .4$	$W = 1.5\times 10^{-4}$ cm
$\Gamma_{lateral} \approx 1$	$\epsilon = 1.5\times 10^{-17}$
$\tau_{ph} = 1.78\times 10^{-12}$ sec	$\beta = 3.87\times 10^{-4}$
$\tau_e = 10^{-9}$ sec	$V_{act} = 5.625\times 10^{-11}$ cm^3
$D = 10$ cm^2/sec	$\alpha = -6$
$\lambda = 1.5\times 10^{-4}$ cm	$\eta = .41$
$P_{max(1)} = 8.3$ mW	$P_{min(0)} = 1.7$ mW
$N_t = 10^{18}$ cm	
$\frac{\psi(x)}{\psi(0)} = \begin{cases} 1 & \lvert x\rvert < W/18 \\ .928 & W/18 \le \lvert x\rvert < W/6 \\ .73 & W/6 \le \lvert x\rvert < 5W/18 \\ .47 & 5W/18 \le \lvert x\rvert < 7W/18 \\ .21 & 7W/18 \le \lvert x\rvert \le W/2 \end{cases}$	

IV. RESULTS

The results in this section were generated, via computer simulation, as follows. The data rate was 8 Gbps, with a pulse shape for the signal into the laser or external modulator given by

$$P(t) = \begin{cases} 0.5\left(\dfrac{t+0.85T}{0.35T}\right)^2 & -0.85T \le t < -0.5T \\ 1 - 0.5\left(\dfrac{t+0.15T}{0.35T}\right)^2 & -0.5T \le t < -0.15T \\ 1 & -0.15T \le t < 0.15T \\ 1 - 0.5\left(\dfrac{t-0.15T}{0.35T}\right)^2 & 0.15T \le t < 0.5T \\ 0.5\left(\dfrac{0.85T-t}{0.35T}\right)^2 & 0.5T \le t \le 0.85T. \end{cases} \quad (26)$$

Note that smoothed pulses, rather than rectangular pulses (as in [3]), were used to reduce the effect of chromatic dispersion. With direct modulation, the input pulse stream was filtered by a low-pass RC filter with a 3 dB bandwidth of 4 GHz to account for the transmit filter and laser parasitics. The transmitted waveform out of the laser was generated by the programs (i.e., from the rate equations) described in [29], using the laser characteristics shown in Table I

TABLE II
EFFECT OF CODING

Code	Data Rate with Coding (÷ 8 Gbps)	Coding Gain at 10^{-9} BER (dB)	Penalty (dB)				
			Chromatic Dispersion			Polarization Dispersion	
			DIR, 40 km	DIR 80 km	EXT, 140 km	DIR 60 psec	EXT, 80 psec
None	1	0	4.4	5.3	2.0	2.0	2.1
(224, 216) Hamming Code [26]	1.04	1.1	3.1	6.5	1.4	1.0	1.2
(64, 56) Reed-Solomon Code [32]	1.14	2.4	0.6	12.8	1.4	-0.1	0.6
(256, 192) Reed-Solomon Code [32]	1.33	3.1	-0.9	-	-	-0.2	1.9

(obtained from [30]). The data rate and low-pass filter bandwidth are similar to those studied in [29], but the laser parameters differ from those used in [29], since the parameters values shown in Table I are more typical of DFB lasers. With external modulation, the laser characteristics were changed such that the laser effects were negligible (ideal external modulation was assumed). The program in [29] use a repetitive pseudorandom data stream of length 64 (versus 256 in [3], which did not consider laser nonlinearity), which contains all bit sequences of length 6. Thus, the results should be accurate as long as the intersymbol interference extends over only a few bit periods. As stated in Section I, the results generated using these parameters demonstrate the typical improvement that can be obtained with the techniques of Section III. However, the improvement in other systems (e.g., with different data rates, laser characteristics, etc.) would vary.

We first consider direct detection systems, both with direct laser modulation and with external modulation. We then study coherent detection systems.

A. Direct Detection

1) Receiver Frequency Response: Let us first consider a receiver with a 3-pole Butterworth filter with a bandwidth of 6.24 GHz (70% of the data rate as in [29]). This filter is nearly ideal in that the intersymbol interference after the receiver filter is very small. (Of course, with Nyquist filtering the intersymbol interference is completely eliminated.) Fig. 9 shows the minimum eye opening for the case of external modulation (i.e., no laser nonlinearity and no chromatic or polarization dispersion) with linear equalization, nonlinear cancellation, and maximum likelihood detection. Results are shown versus the number signal samples or bits N used by the techniques. For each technique, N consecutive signal samples (T spaced taps) or bits were used, with the number of decided N_1 and future N_2 samples or bits chosen to maximize the improvement (with $N = N_1 + N_2 + 1$). However, because most of the intersymbol interference is due to the decided bits, the results in Fig. 9 are for equalization (and cancellation) with decided taps (bits) only. A sampling time at the peak eye opening prior to equalization was used. Since the distortion is linear (and the ISI is small), the tapped delay line has the same improvement as nonlinear cancellation and maximum likelihood detection. Note that most of the intersymbol interference is removed with a 3-tap equalizer. (If the sampling time was optimized, a 100% eye opening with a 3-tap equalizer can be obtained.)

Next consider the effect of dc offset on the equalizer performance. Fig. 10 shows the minimum eye opening versus the number of taps for several values of offset. The results show that the continuous version of the zero-forcing algorithm is unaffected by small offset. However, the performance of the discrete version degrades with small offset. Furthermore, the degradation is even worse for the discrete version that uses only the error with "1" data bits. Note that when the offset is on the order of the level of the intersymbol

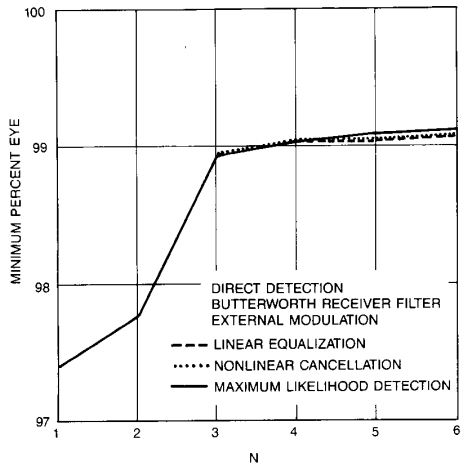


Fig. 9. Minimum eye opening versus the number of signal samples or bits for linear equalization, nonlinear cancellation, and maximum likelihood detection with a Butterworth receiver filter, and direct detection with external modulation.

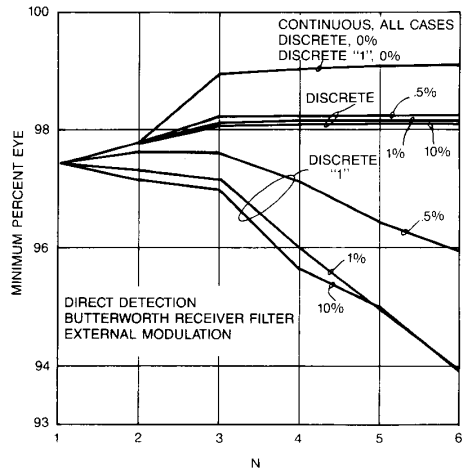


Fig. 10. Minimum eye opening versus the number of taps for linear equalization with different adaptation algorithms and dc offsets. The adaptive algorithms are the continuous zero-forcing algorithm, the quantized (discrete) zero-forcing algorithm, and the discrete zero-forcing algorithm that adapts only when "1's" are detected. Results are shown for dc offsets of 0, 0.5, 1, and 10% of the received minimum eye opening.

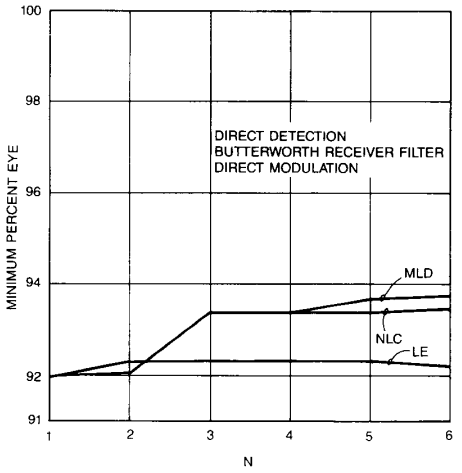


Fig. 11. Effect of equalization techniques with laser nonlinearity and a Butterworth receiver filter. Minimum eye opening versus number of bit samples for linear equalization (LE), nonlinear cancellation (NLC), and maximum likelihood detection (MLD).

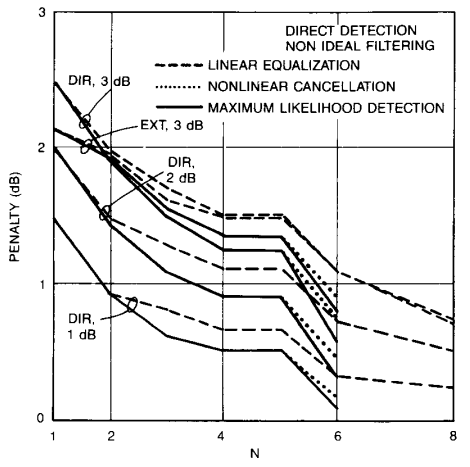


Fig. 12. Effect of equalization techniques with and without laser nonlinearity and a nonideal receiver filter—optical power penalty versus number of bit samples for LE, NLC, and MLD, direct detection, and a receiver filter with 1, 2, or 3 dB of ripple.

interference, equalization with the tap weights adapted by the discrete version using only the error with "1's," can actually decrease the eye opening. For the results in the remainder of this section, the tap weights were adapted using the continuous zero-forcing algorithm when the eye was open, and the mean-square-error algorithm when the eye was closed.

Next, consider the effect of equalization techniques with laser nonlinearity (direct modulation of the laser). In this case, both the optical signal distortion and the optical-to-electrical conversion are nonlinear. Fig. 11 shows the minimum eye opening as a function of the number of samples or bits used by linear equalization (LE), nonlinear cancellation (NLC), and maximum likelihood detection (MLD). Comparing these results to those for external modulation (Fig. 9), we see that, with the Butterworth receiver filter, the laser nonlinearity causes more than twice as much intersymbol interference as the receiver filter alone. Thus, the linear equalizer is not effective in increasing the eye opening, while nonlinear cancellation and maximum likelihood detection increase the eye opening by only 2%.

With current technology, however, typical receivers do not have an ideal 3-pole, 0.64 GHz, Butterworth frequency response. Typical receivers [20] have a higher decrease in gain with frequency across the passband, a higher rolloff above the 3 dB bandwidth, and ripple in the passband. Here, we model such a receiver as a 10 GHz RC filter in cascade with a 6 GHz 8-pole Butterworth filter and a 3 GHz cosine ripple of a few dB. Such a receiver has amplitude and phase characteristics that agree well with the measured receiver frequency response for an 8 Gbps receiver [20].

Fig. 12 shows the effect of equalization techniques with a nonideal receiver filter. The optical power penalty is plotted versus N for LE, NLC, and MLD with the sampling time optimized at each point. Results are shown for direct modulation with 1, 2, and 3 dB of ripple and external modulation with 3 dB of ripple. A 6-tap linear equalizer reduces the penalty by more than half (in dB), while NLC and MLD can reduce the penalty even further. Since most of the penalty is due to the filter, laser nonlinearity does not significantly affect performance.

Since, with this nonideal filter, intersymbol interference is mainly due to decided bits, the best N taps or samples included at most one future sample (as before, results are for the best N consecutive samples or bits). However, some of the samples do not help to improve performance, as shown by the small decrease in penalty with increasing N in some cases (in particular, increasing N from 4 to 5). Thus, for a fixed number of samples, a further decrease in penalty can be obtained with unequally spaced (nonconsecutive) samples. However, such an equalizer may not be as robust against changes in filter characteristics.

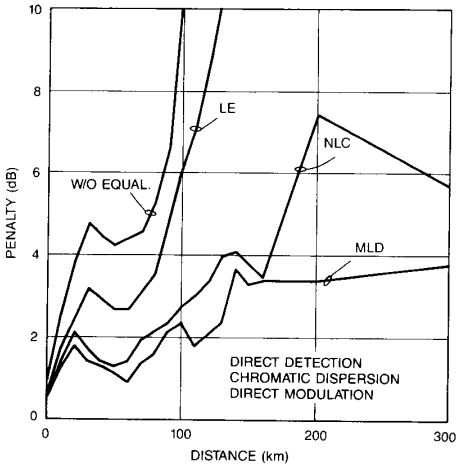


Fig. 13. Effect of equalization techniques with chromatic dispersion and laser nonlinearity—optical power penalty versus distance with LE, NLC, and MLD for $N = 6$.

2) Chromatic Dispersion: Next consider the effect of the various impairment-reducing techniques on chromatic dispersion. Fig. 13 shows the effect of equalization techniques with chromatic dispersion and laser nonlinearity. (All results in Figs. 13–16 and 20 and Table II are for $D(\lambda) = 17$ ps/km/nm.) Results are shown for LE, NLC, and MLD with $N = 6$ and optimum sampling time, and a 3-pole Butterworth (nearly ideal) receiver filter. (With more distortion, we display the results in terms of optical power penalty rather than eye opening.) Note that the combined effect of chromatic dispersion and laser nonlinearity produces a dip in penalty above 40 km, similar to that shown in [29] where different DFB laser characteristics were used. However, the power penalty in this region is much higher than in [29], showing that the DFB lasers with more typical laser characteristics are more affected by chromatic dispersion. Although chromatic dispersion produces nonlinear

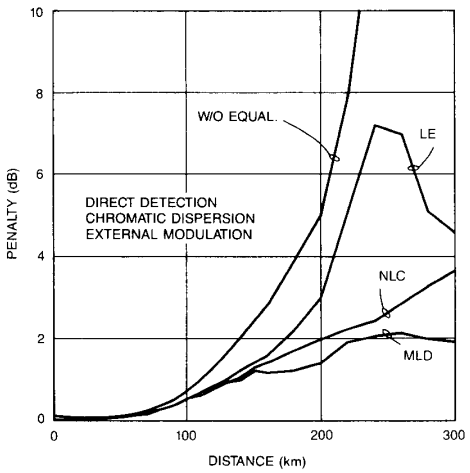


Fig. 14. Effect of equalization techniques with chromatic dispersion without laser nonlinearity in direct detection systems.

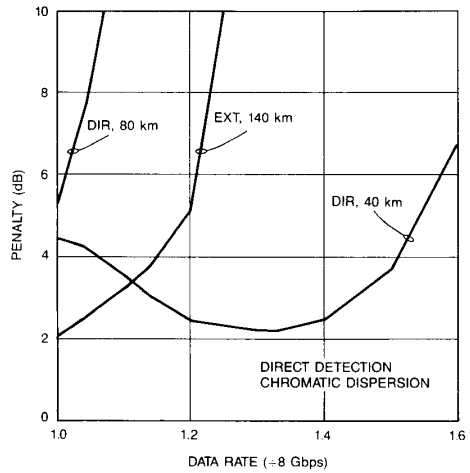


Fig. 15. Effect of coding (i.e., increased data rate) with chromatic dispersion in direct detection systems.

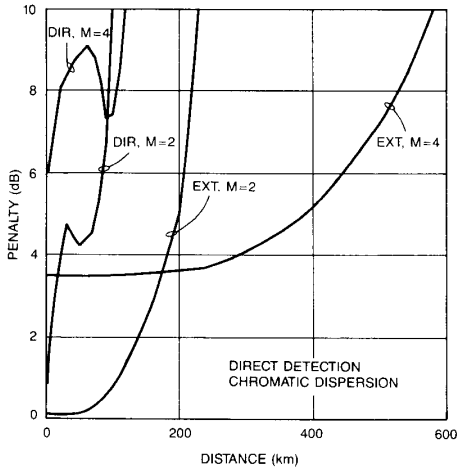


Fig. 16. Effect of multilevel signaling with chromatic dispersion in direct detection systems.

intersymbol interference at the receiver, LE still reduces the power penalty somewhat—by at least 1.5 dB for distances above 40 km (or equivalently increases the maximum distance for a given penalty by about 20%).[12] NLC and MLD, however, decrease the penalty by more than 3 dB for distances above 40 km and substantially increase the dispersion-limited transmission distance (to beyond 300 km).[13] Note that these curves are jagged because the results were generated at discrete distances (at least 10 km between points), and because the penalty fluctuates with the level of intersymbol interference in the signal samples. Here, as with nonideal receiver filtering, the best N consecutive samples included at most 1 future bit sample, which simplifies implementation of NLC.

Fig. 14 shows the effect of equalization techniques with chromatic dispersion [3] when laser nonlinearity is not present (i.e., with external modulation).[14] As before, LE decreases the penalty (by more than 1.5 dB above 160 km) or, alternatively, increases the dispersion-limited distance (by 25% for a 3 dB penalty).[15] Although with LE the penalty dips above 240 km, this effect is highly dependent on the transmit pulse shape and receiver frequency characteristics and, therefore, may not necessarily be present in systems with slightly different characteristics. On the other hand, NLC and MLD do not have such variations. These techniques greatly reduce the penalty above 200 km, increasing the dispersion-limited distance for a 3 dB penalty to 270 and 400 km with NLC and MLD, respectively.

Next, consider the effect of coding with chromatic dispersion. As discussed in Section III, the first step in analyzing the effect of coding is calculating the increase in penalty with increased data rate. Since the dispersion limit of an optical fiber is proportional to the data rate squared times the distance [3], [31] (without any other impairments), we would expect the shape of the penalty versus data rate curves to be similar to that of the penalty versus distance curves (Figs. 13 and 14—see [3] for other modulation formats). This is illustrated in Fig. 15 where the penalty versus data rate is shown at 40 and 80 km for direct modulation and at 140 km for external modulation. Here we assume that the speed (bandwidth) of the electronics increases in proportion to the data rate, but that the laser characteristics are unchanged (i.e., the laser nonlinearity increases with data rate for direct laser modulation). With laser nonlinearity, at 80 km the penalty rapidly increases with data rate, and therefore, coding is not useful in extending the dispersion-limited distance. At 40 km, however, the penalty actually decreases with the data rate (as it does with distance). Thus, coding can be useful in reducing the penalty for distances between 20 to 60 km. With external modulation, as shown at 140 km, coding may slightly decrease the power penalty, but the steep rise in penalty with distances above 140 km means that coding will not significantly increase the dispersion-limited distance. As an illustration of the above conclusions, Table II shows the penalty with three codes.[16] For example, a (256, 192) Reed-Solomon code, which has a coding gain of 3.1 dB [32] (without distortions) but increases the data rate by 33%, would reduce the optical power penalty due to chromatic dispersion with direct laser modulation over 40 km, from 4.4 dB to -0.9 dB (the 5.3 dB decrease with chromatic dispersion is greater than the 3.1 dB coding gain without dispersion because the penalty *decreases* with small increases in data rate at 40 km). Thus, coding is shown to reduce the penalty by up to 5.3 dB. Note that with NLC or MLD,

[12]A fractionally spaced equalizer may do even better.

[13]Distances above 200 km are only feasible because of the advent of optical amplifiers.

[14]The distances without equalization shown in Fig. 14 are about 25% longer than the results calculated for OOK from [3], mainly because rectangular [rather than smoothed (26)] transmit pulses were studied in [3].

[15]Again, a fractionally spaced equalizer may do even better.

[16]Note that these codes have never been implemented at the data rates being considered here, but they merely serve as an example of the performance improvement that is theoretically possible. The coding techniques can, of course, be implemented on lower rate channels that are multiplexed to generate multigigabit per second signals, as in [28], but the number of channels to be multiplexed can become very large.

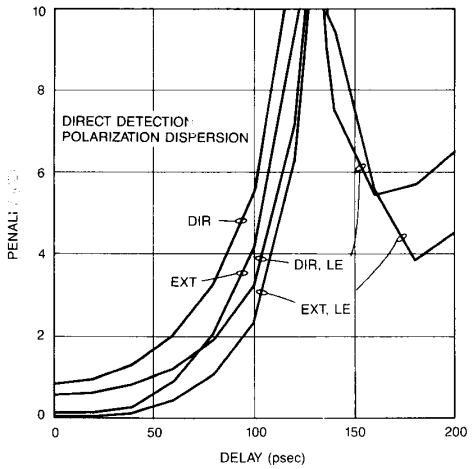


Fig. 17. Effect of linear equalization with polarization dispersion (worst case) in direct detection systems.

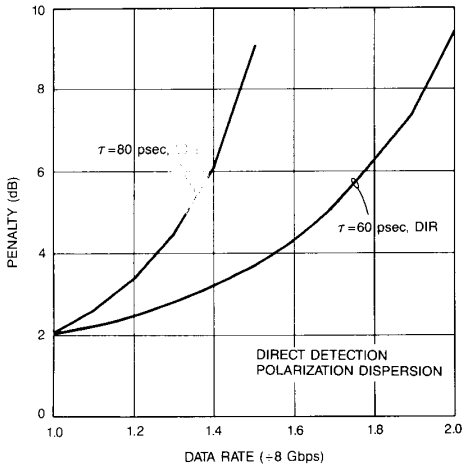


Fig. 18. Effect of coding (i.e., increased data rate) with polarization dispersion (worst case) in direct detection systems.

the penalty increases very slowly with distance (see Figs. 13 and 14) and, thus, coding in combination with NLC or MLD should decrease the penalty even further.

Finally, consider the effect of multilevel signaling. Fig. 16 shows the penalty versus distance with 2- and 4-level signaling and direct and external modulation (at 8 Gbps). Laser nonlinearity is shown to greatly increase the penalty with 4-level signaling, with a 5 dB higher penalty than 2-level signaling for distances less than 80 km. Also, with 4-level signaling, there is no significant increase in the dispersion-limited distance even with large penalties. However, without laser nonlinearity, 4-level signaling has a lower penalty than 2-level signaling for distances greater than 180 km, greatly increasing the transmission distance for penalties greater than 4 dB (e.g., twice the distance at 5 dB penalty). Note that it would be more difficult to generate and detect a multilevel signal at 8 Gbps (e.g., multiple threshold detectors are required) than a 2-level signal. Although such a multilevel system has not been implemented at 8 Gbps, it would appear to be feasible.

3) Polarization Dispersion: Next, let us consider the effect of the various impairment reducing techniques on first-order polarization dispersion effects [4] (second-order effects include linear delay distortion, which was studied in Section IV-A-2). Because of its simple form, first-order polarization dispersion effects can easily be eliminated by NLC and MLD using at most 3 signal samples, with the available adaptive algorithms for NLC cancellation [21]. However, since linear equalization has been shown to be effective against nonideal receiver frequency response and chromatic dispersion, and, unlike NLC and MLD, has been implemented at Gbps rates [15], let us consider linear equalization, which, as shown in Section III, can easily be adaptive. Now, the inverse filter for polarization dispersion is a very simple infinite impulse response (IIR) filter. However, adaptive IIR filters can become unstable (although techniques exist that can be used to maintain stability [33]), and the weights can become trapped in local minimum (that are not the global minimum) of the performance surface during adaptation [33]. These problems are unacceptable given the high reliability usually required in high-speed lightwave systems. These problems do not occur with the feedforward adaptive tapped delay lines considered in this paper, although the tapped delay line may require more taps (delays) than an IIR filter for equal performance.

The effect of polarization dispersion depends on α (the ratio of signal powers in the two polarizations) and τ (the time delay between the two polarizations). The worst case for α occurs when $\alpha = 1$. In this case, the eye is closed when τ is a multiple of the bit period (125 ps), and an LE will not open the eye. However, for all other values of α and τ, LE reduces the penalty. Fig. 17 shows the effect of LE (with 6 taps) for $\alpha = 1$, with direct and external laser modulation.[17] LE reduces the penalty by more than 1 dB for $\tau > 70$ ps and increases the tolerable delay for penalties greater than 3 dB by about 10% (this corresponds to a 20% distance increase since the delay varies as the square root of the distance). Note also that, for delays greater than the bit period, LE reduces the penalty to as low as 4 dB. Of course, for $\alpha \neq 1$, LE reduces the penalty to even lower values.[18]

Next, consider the effect of coding with polarization dispersion. As before, we first need to consider the increase in penalty with increased data rate. Since increasing the data rate should have the same effect as increasing the delay by the same proportion (without any other impairments), we would expect that the penalty versus data rate curves would look similar to the penalty versus delay curves. This is illustrated in Fig. 18 where the penalty versus data rate is shown for direct modulation with $\tau = 60$ ps, and for external modulation with $\tau = 80$ ps. As before, for these results we have assumed that the speed of the electronics increases in proportion to the data rate increase, but that the laser characteristics are unchanged. Table II shows that coding reduces the penalty by as much as 2 dB when laser nonlinearity is present and slightly less than 2 dB when external modulation is used. However, because of the steepness of the penalty versus delay curves as the delay approaches 125 ps, coding is not useful in increasing the tolerable delay with large penalties.

Fig. 19 shows the effect of multilevel signaling with polarization dispersion. As with chromatic dispersion, the results for direct modulation systems show that multilevel signaling does not increase the tolerable delay except for very high penalties (greater than 11 dB with polarization dispersion). However, without laser nonlinearity, 4-level signaling has a lower penalty than 2-level signaling for delays greater than 100 ps and for penalties greater than 4.2 dB. Multilevel signaling increases the tolerable delay by more than 30% for a 6 dB penalty. For given τ, however, the penalty with 2-level signaling is lower for $\alpha \neq 1$, while the penalty with 4-level signaling is at least 3.3 dB. Thus, the improvement with 4-level signaling should be less for $\alpha \neq 1$.

To determine completely the effect of equalization on polarization dispersion, we need to consider both first- and second-order effects together. This will be studied in a future paper.

[17]The results shown in Fig. 17 for external modulation without equalization are in agreement with those of [4] for OOK.

[18]In addition, a fractionally spaced equalizer may reduce the penalty even further.

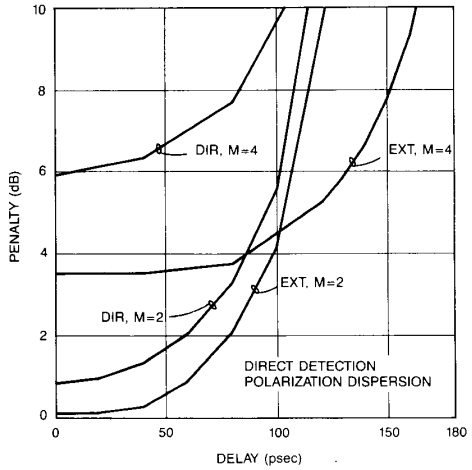


Fig. 19. Effect of multilevel signaling with polarization dispersion (worst case) in direct detection systems.

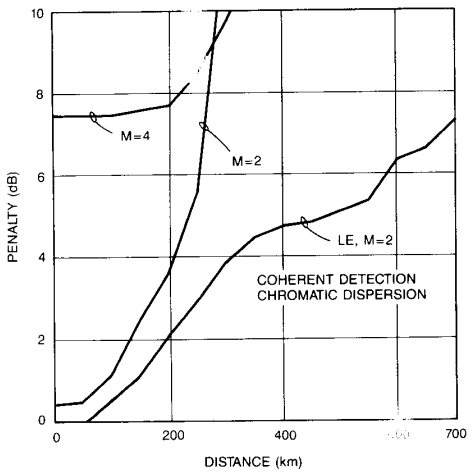


Fig. 20. Effect of linear equalization and multilevel signaling with chromatic dispersion in coherent systems.

4) Mode Partition Fluctuations: Finally, consider the effect of the various impairment reducing techniques on mode partition fluctuations. Since the amplitude of the side mode varies randomly from bit to bit (or at least over a period of several bits), the intersymbol interference is random, and LE, NLC, and MLD will not be effective in reducing the error rate floor. Also, multilevel signaling would only increase the error rate floor, since the error floor would not be reduced significantly if the symbol rate was reduced, while a multilevel signal would be more sensitive to intersymbol interference. Coding, however, can be useful in reducing the error rate floor. Since the side mode amplitude usually varies randomly from bit to bit, the errors occur randomly, and simple single-error correcting block codes can be used. Even error bursts of several bits can be corrected if coding is implemented on lower rate channels multiplexed for multigigabit per second signals. For example, the code of [28] can reduce an error rate floor of 10^{-6} to below 10^{-9} (as noted previously), while the more complicated (64, 56) Reed–Solomon code [32] can reduce an error rate floor of 10^{-4} to below 10^{-11}.

B. Coherent Detection

Let us next briefly consider coherent detection. With coherent detection, the optical power penalty is twice that of direct detection systems (in dB) for the same electrical power penalty. Therefore, we would expect greater penalties than direct detection with the same receiver filtering, chromatic dispersion, and polarization dispersion, and greater reductions in penalty with LE[19] and coding, but lower reductions in penalty with multilevel signaling (except for receiver filtering, where it may be possible to build better filters with lower bandwidth).

As an example, Fig. 20 shows the effect of LE and multilevel signaling with chromatic dispersion.[20] Results show that LE (with 6 taps) reduces the penalty more than 1 dB (for penalties with 2-level signaling greater than 3 dB) and significantly increases the dispersion-limited distance (by more than twice, for penalties greater than 5 dB).[21] Because of the steepness of the slope of the penalty versus distance curve, however, coding is not effective in increasing the transmission distance, although it could be useful in reducing the penalty in combination with LE. Multilevel signaling is not useful in reducing the penalty unless it exceeds 9 dB.

V. Summary and Conclusion

In this paper, we have studied the various impairments in high-speed lightwave systems, presented techniques to reduce these impairments, and analyzed their performance in a typical system. The impairments include laser nonlinearity, mode partition fluctuations, chromatic and polarization dispersion, and nonideal receiver filtering. The techniques include linear equalization, nonlinear cancellation, maximum likelihood detection, coding, and multilevel signaling. Methods for implementing these techniques, including adaptive linear equalization, were presented. Computer simulation results for an 8 Gbps system using measured laser and receiver characteristics showed that in direct detection systems a 6-tap linear equalizer can reduce the penalty due to nonideal receiver filtering in half (in dB), and the penalty due to chromatic and polarization dispersion by more than 1 dB (or increase the dispersion-limited distance by more than 20%). Nonlinear cancellation and maximum likelihood detection can reduce the penalty even further, more than doubling the dispersion-limited distance in some cases. Coding was shown to reduce the penalty by a few dB in those cases where the increase in penalty versus distance or delay was not too large (e.g., direct detection with chromatic dispersion and a 40 km fiber length or polarization dispersion and a delay of 50% of the bit period). Multilevel (4-level) signaling was shown to double the dispersion-limited distance and increase by more than 30% the tolerable delay in systems with external laser modulation. In coherent systems, linear equalization (with 6 taps) can more than double the dispersion-limited distance. Combinations of equalization, coding, and multilevel signaling should result in even larger improvements.

Appendix

Here we consider the effect of linear equalization on quadratic (nonlinear) distortion. In general, linear equalization will not improve performance with quadratic distortion. However, in lightwave

[19]With coherent detection, all the impairments discussed in this paper are linear in the receiver (laser nonlinearity is not present because external modulation is used in coherent systems). Thus, LE should do a very good job of reducing the penalty, and NLC and MLD may not be needed.

[20]The distances without equalization shown in Fig. 20 are about 50% longer than the results calculated for synchronous ASK from [3], mainly because rectangular (rather than smoothed (26)) transmit pulses were studied in [3].

[21]In this case, a fractionally spaced equalizer should be able to completely remove the ISI without noise enhancement, i.e., 0 dB penalty even over hundreds of kilometers. Also, microwave waveguides [13] or microstrip lines [14] can be used to equalize the chromatic dispersion.

systems, if the signaling is binary, on-off keying and the major portion of the ISI is from a single adjacent symbol, then linear equalization can be effective in reducing the ISI.

In the system studied in this paper (a DFB laser at 8 Gbps—see Section IV and [29]), the extinction ratio of the laser modulation is 5, i.e., the level for a "1" is approximately 5 times that for a "0." Thus, with ISI from one adjacent bit, the received electrical signal at time kT can be calculated to be given by

$$\begin{aligned} s_e(kT) &= |s_0(kT)|^2 \approx |0.45 + 0.55I_k + \rho(T)I_{k+1}|^2 \\ &\approx 0.2 + 0.3I_k^2 + 0.5I_k + \rho^2(T)I_{k-1}^2 \\ &\quad + (0.9 + 1.1I_k)I_{k-1}\rho(T), \end{aligned} \tag{A1}$$

where $\rho(t)$ is the received pulse response of the system (assumed to be maximum at $t = 0$). Note that if $\rho(T) = 0$, $s_e(kT)$ is 0.2 for $I_k = 0$ and 1 for $I_k = 1$ (for an extinction ratio of 5). Now, the linear component[22] of the ISI (in terms of the squared data symbols) is $\rho^2(T)I_{k-1}^2$. If the ISI is reasonably small, then this component will be much less than the nonlinear component $(0.9 + 1.1I_k)I_{k-1}\rho(T)$. In this case, cancellation of the linear component (or partial cancellation through linear equalization) will have a negligible impact.

However, for binary, on-off keying, the effect of the nonlinear component *can* be reduced by linear equalization. First, since $I_k^2 = I_k$, (A1) can be expressed as

$$\begin{aligned} s_e(kT) &= 0.2 + 0.8I_k + (\rho^2(T) + 0.9)I_{k-1} \\ &\quad + 1.1\rho(T)I_kI_{k-1}, \end{aligned} \tag{A2}$$

or

$$s_e(kT) = \begin{cases} 0.2 & \text{if } I_{k-1} = 0 \quad \text{and } I_k = 0 \\ 1 & \text{if } I_{k-1} = 0 \quad \text{and } I_k = 1 \\ 0.2 + (\rho^2(T) + 0.9) & \\ \quad \text{if } I_{k-1} = 1 \quad \text{and } I_k = 0 & \\ 1 + (\rho^2(T) + 0.9) + 1.1\rho(T) & \\ \quad \text{if } I_{k-1} = 1 \quad \text{and } I_k = 1. & \end{cases} \tag{A3}$$

The linear component of the ISI is now $\rho^2(T) + 0.9$ and the nonlinear component $1.1\rho(T)I_kI_{k-1}$. Thus, in contrast to (A1), the linear component is always larger (and is much larger with small ISI) than the nonlinear component, and linear equalization can be effective. Furthermore, the effect of the smaller nonlinear term can also be somewhat reduced by linear equalization. When $\rho(T) < 0$, the nonlinear term $1.1\rho(T)I_kI_{k-1}$ reduces the level for a "1" when $I_{k-1} = 1$. (The nonlinear term increases the eye opening for $\rho(T) > 0$.) In this case, when the ISI is small, we can increase the minimum distance to the threshold (set at 0.6) for $\rho(T) < 0$, by using a linear equalizer to add the term $0.55\rho(T)s_e((k-1)T)$ (which is approximately $0.55\rho(T)I_{k-1}$ with small ISI) to $s_e(kT)$.

Acknowledgment

We gratefully acknowledge discussions with T. L. Koch concerning single-frequency lasers and the use of his programs; discussions with D. G. Duff, B. L. Kasper, and O. Mizuhara on equalization techniques; and the programming assistance of C. A. Doubrava.

References

[1] J. H. Winters and R. D. Gitlin, "Techniques for mitigating the effects of transmission impairments in single-mode digital fiber optic systems," submitted to *IEEE Commun. Mag.*

[2] B. L. Kasper, "Equalization of multimode optical fiber systems," *B.S.T.J.*, vol. 61, pp. 1367-1388, Sept. 1982.

[3] A. F. Elrefaie, R. E. Wagner, D. A. Atlas, and D. G. Daut, "Chromatic dispersion limitations in coherent lightwave transmission systems," *IEEE J. Lightwave Technol.*, vol. 6, pp. 704-709, May 1988.

[4] R. E. Wagner and A. F. Elrefaie, "Polarization-dispersion limitations in lightwave systems," in *Tech. Dig. Optical Fiber Commun. Conf.*, New Orleans, LA, Jan. 25-28, 1988, p. 37.

[5] U. Timor and R. A. Linke, "A comparison of sensitivity degradations for optical homodyne vs. direct detection of on-off keyed signals," *IEEE J. Lightwave Technol.*, vol. 6, pp. 1782-1788, Nov. 1988.

[6] E. A. Lee and D. G. Messerschmitt, *Digital Communication*. Boston, MA: Kluwer Academic, 1988, p. 65.

[7] D. G. Messerschmitt, "Minimum MSE equalization of digital fiber optic systems," *IEEE Trans. Commun.*, vol. COM-16, pp. 1110-1118, July 1978.

[8] C. D. Poole and C. R. Giles, "Polarization-dependent pulse compression and broadening due to polarization dispersion in dispersion-shifted fiber," *Opt. Lett.*, vol. 13, pp. 155-157, Feb. 1988.

[9] G. Vannucci, "A coherent phase-diversity optical receiver with double frequency conversion," submitted to *IEEE Trans. Commun.*

[10] L. J. Cimini, Jr., I. M. I. Habbab, S. Yang, A. J. Rustako, Jr., K. Y. Liou, and C. A. Burrus, "A polarization-insensitive coherent lightwave system using wide-deviation FSK and data-induced polarization switching," *Electron. Lett.*, vol. 24, pp. 358-360, Mar. 17, 1988.

[11] R. A. Linke, B. L. Kasper, C. A. Burrus, I. P. Kaminow, J.-S. Ko, and T. P. Lee, "Mode power partition events in nearly single-frequency lasers," *IEEE J. Lightwave Technol.*, vol. 3, pp. 706-712, 1985.

[12] N. Henmi, Y. Koizumim, M. Yamaguchi, M. Shakada, and I. Mito, "The influence of directly modulated DFB LD sub-mode oscillation on long-span transmission system," *IEEE J. Lightwave Technol.*, vol. 6, pp. 636-642, May 1988.

[13] J. H. Winters, "Equalization in coherent lightwave systems using microwave waveguides," *IEEE J. Lightwave Technol.*, vol. 7, pp. 813-815, May 1989.

[14] N. Takachio and K. Iwashita, "Compensation of fibre chromatic dispersion in optical heterodyne detection," *Electron. Lett.*, pp. 108-109, Jan. 21, 1988.

[15] B. L. Kasper, J. C. Campbell, J. P. Talman, A. H. Gnauck, J. E. Bowers, and W. S. Holden, "An APD/FET optical receiver operating at 8 Gbit/s," *J. Lightwave Technol.*, vol. LT-5, pp. 344-347, Mar. 1987.

[16] B. Widrow, "Adaptive filters," in *Aspects of Network and System Thory*, R. E. Kalman and N. De Claris, Eds. New York: Holt, Rinehart, Winston, 1970, pp. 563-587.

[17] J. G. Proakis, *Digital Communications*. New York: McGraw-Hill, 1983.

[18] R. W. Lucky, "Automatic equalization for digital communications," *B.S.T.J.*, vol. 44, pp. 547-588, Apr. 1965.

[19] D. Hirsh and W. J. Wolf, "A simple adaptive equalizer for efficient data transmission," *IEEE Trans. Commun.*, vol. COM-18, pp. 5-12, Feb. 1970.

[20] D. G. Duff, private communication.

[21] E. Biglieri, A. Gersho, R. D. Gitlin, and T. L. Lim, "Adaptive cancellation of nonlinear intersymbol interference for voiceband data transmission," *IEEE J. Select. Areas Commun.*, vol. SAC-2, pp. 765-777, Sept. 1984.

[22] A. Gersho and T. L. Lim, "Adaptive cancellation of intersymbol interference for data transmission," *B.S.T.J.*, vol. 60, pp. 1997-2021, Nov. 1981.

[23] A. Gersho, E. Y. Ho, R. D. Gitlin, V. B. Lawrence, and T. L. Lim, "Interference cancellation apparatus," U.S. Patent 4 412 341.

[24] K. Fisher, J. Cioffi, and C. M. Melas, "An adaptive DFE for storage channels suffering from nonlinear ISI," in *Proc. Int. Conf. Commun.*, Boston, MA, June 11-14, 1989, pp. 53.7.1-53.7.5.

[25] J. H. Winters and C. Rose, "Minimum distance automata in parallel networks for optimum classification," *Neural Networks*, vol. 2, pp. 127-132, Mar. 1989.

[26] H. K. Thapar and J. M. Cioffi, "A block partitioning method for designing high-speed Viterbi detectors," in *Proc. Int. Conf. Commun.*, Boston, MA, June 11-14, 1989, pp. 35.5.1-35.5.5.

[27] R. W. Means and B. Caid, "A backpropagation network error correcting decoder for convolutional and block codes," in *Proc. Ist Annu. INNS Meet.*, Boston, MA, 1988, p. 37.

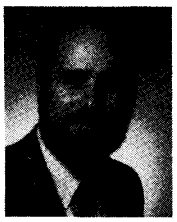
[28] W. D. Grover, "Forward error correction in dispersion-limited lightwave systems," *J. Lightwave Technol.*, vol. 6, pp. 643-654, May 1988.

[29] T. L. Koch and P. J. Corvini, "Semiconductor laser chirping-induced dispersion distortion in high-bit-rate optical fiber communications

[22] In terms of the above notation, the system would be linear, in terms of the squared data symbols, if $s_e(kT) = \sum_j I_{k-j}^2\rho^2(jT)$.

systems," in *Proc. IEEE Int. Conf. Commun. '88*, June 12-15, 1988, pp. 19.4.1-19.4.4.
[30] T. L. Koch, private communication.
[31] P. S. Henry, "Introduction to lightwave transmission," *IEEE Commun. Mag.*, pp. 12-16, May 1985.
[32] E. R. Berlekamp, R. E. Peile, and S. P. Pope, "The application of error control to communications," *IEEE Commun. Mag.*, pp. 44-56, Apr. 1987.
[33] B. Windrow and A. D. Stearns, *Adaptive Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1985, p. 154.
[34] P. J. Corvini and T. L. Koch, "Computer simulation of high-bit-rate optical fiber transmission using single-frequency lasers," *IEEE J. Lightwave Technol.*, vol. LT-5, pp. 1591-1595, Nov. 1987.

★



Jack H. Winters (S'77-M'82-SM'88) was born in Canton, OH, on September 17, 1954. He received the B.S.E.E. degree (summa cum laude) from the University of Cincinnati, Cincinnati, OH, in 1977 and the M.S. and Ph.D. degrees in electrical engineering from The Ohio State University, Columbus, in 1978 and 1981, respectively.

From 1973 to 1976 he was with the Communications Satellite Corporation, Washington, DC, and from 1977 to 1981 with the ElectroScience Laboratory, The Ohio State University, where he studied adaptive antenna arrays, and received The Ohio State University ElectroScience Laboratory's award for the outstanding dissertation of 1981. Since 1981 he has been with AT&T Bell Laboratories, Holmdel, NJ, where he is in the Network Systems Research Department. He has been involved in research on modulation and coding, mobile and indoor radio systems, neural networks, and applications of high-T_c superconductors in communication systems. Currently he is involved in research on signal processing for TV and lightwave systems.

Dr. Winters is a member of Sigma Xi.

★

Richard D. Gitlin (S'67-M'69-SM'76-F'85) was born in Brooklyn, NY, on April 25, 1943. He received the B.E.E. degree (cum laude) from the City College of New York, New York, NY, in 1964 and the M.S. and D.Eng.Sc. degrees from Columbia University, New York, NY, in 1965 and 1969, respectively.

Since 1969 he has been with AT&T Bell Laboratories, Holmdel, NJ. From 1969 to 1979 he did applied research and exploratory development in the field of high-speed voiceband modems, with emphasis on adaptive equalization, bandwidth-efficient modulation, echo cancellation, carrier and timing recovery, and digital signal processing. From 1979 to 1982 he supervised a group doing exploratory and advanced development in these areas. From 1982 to 1987 he was head of a department responsible for systems engineering, exploratory development, and final development of data communications equipment. Currently he is Head of the Network Systems Research Department where he manages research in broadband multimedia networking, packet switching, and data networking. He is the author of more than 50 technical papers, numerous conference papers, and he holds 25 patents in the areas of data communications and digital signal processing. He is co-author of a paper on fractionally spaced adaptive equalization that was selected as the Best Paper in Communications by the *Bell System Technical Journal* in 1982.

Dr. Gitlin is a member of Sigma Xi, Tau Beta Pi, and Eta Kappa Nu. He is a former Chairman of the Communication Theory Committee of the IEEE Communications Society, as well as a member of the COMSOC Awards Board. Previously he was Editor for Communication Theory of the IEEE TRANSACTIONS ON COMMUNICATIONS, and a member of the Editorial Advisory Board of the PROCEEDINGS OF THE IEEE. In 1985 he was elected a Fellow of the IEEE for his contributions to data communications technology, and in 1987 he was named an AT&T Bell Laboratories Fellow.