

Methods for Crosstalk Measurement and Reduction in Dense WDM Systems

Keang-Po Ho and Joseph M. Kahn, *Member, IEEE*

Abstract—We propose a scheme for the monitoring and reduction of crosstalk arising from the limited stop-band rejection of optical bandpass filters in dense WDM systems. The optical carrier at each wavelength is modulated with a subcarrier tone unique to that wavelength. The level of crosstalk from a given channel can be determined by measuring the power of the corresponding tone. Crosstalk from other channels can be cancelled in a linear fashion by weighting and summing the photocurrents of the desired channel and several adjacent interfering channels. Alternatively, in nonlinear crosstalk cancellation, decisions are made on the interfering signals, and these decisions are weighted and summed with the photocurrent of the desired channel. For example, assuming an optical filter having a Gaussian passband, the channel density can be increased from 20 to 30%, depending on the number of adjacent channels detected. The signal-to-interference ratio can be increased by 10–20 dB and the system can achieve a BER $< 10^{-9}$ under conditions where, without interference cancellation, the signal-to-interference ratio would be less than 10 dB.

I. INTRODUCTION

WAVELENGTH-division-multiplexed (WDM) networks can utilize the vast bandwidth provided by a single-mode optical fiber [1]–[3], but the channel density is limited by crosstalk from other adjacent channels arising from limited stop-band rejection of optical filters. Most research to date has sought to characterize the amount of crosstalk and its effects on system performance [4]–[8]. To our knowledge, previous research has not addressed methods to monitor crosstalk in real time. In this paper, we show that by using subcarrier tones to identify channels of different wavelengths, the crosstalk levels can be monitored by measuring the electrical powers of the corresponding identification tones.

Increase in channel density requires some means to reduce crosstalk. Most research to date has emphasized the reduction of crosstalk by improvement of demultiplexer design and fabrication techniques [9]–[12], or by use of Fabry–Perot filters having higher finesse. However, due to the limited stop-band rejection of grating demultiplexers, the number of channels accommodated within the gain bandwidth of

optical amplifiers has been limited. The finesse of a practical single-cavity Fabry–Perot filter is also limited, limiting the number of channels. In this paper, we propose a crosstalk cancellation scheme for use with grating-based demultiplexers [1], [9]–[12]. We propose to use an array of detectors to detect several adjacent channels, weight them, and combine them with the desired channel. Because a grating-based demultiplexer induces only linear crosstalk, our technique can cancel nearly all crosstalk, as shown in this paper. Linear crosstalk cancellation was first proposed in [13] where an adaptive least-mean-square algorithm was used to adjust the weights. The adaptive algorithm requires sampling of the received signal at high speed to generate the error signal [13]. In our scheme, the correlation matrix, and thus the weights, can be determined from the electrical powers of subcarrier tones, and an open-loop algorithm can be employed to calculate the weights directly. An open-loop algorithm is not only faster and more accurate than a closed-loop adaptive algorithm, but also eliminates the potential instability associated with the closed-loop algorithm. Furthermore, in this paper, we also propose and analyze an even more effective nonlinear scheme that uses decision-directed cancellation. In this scheme, decisions are made on the interfering signals, and these decisions are weighted and summed with the photocurrent of the desired channel.

The remainder of this paper is organized as follows. Section II describes the crosstalk monitoring scheme using identification tones. Section III presents a linear canceller in which the tap weights are adjusted using the minimum mean-square error algorithm. Section IV describes a nonlinear canceller using a decision-directed technique. Section V presents numerical results for a grating having a Gaussian-shaped passband. Section V discusses the application of tone identification to monitor the optical power after an optical amplifier. Section VII presents our conclusions.

II. CROSSTALK MONITORING

Several recent papers [14]–[18] have addressed the encoding of packet-routing information in out-of-band subcarrier channels, whereby packets are routed according to header information modulated onto the subcarrier channel [16], [17] or simply according to the frequency of the subcarrier tone [14], [15]. In this section, we propose to encode channels at different wavelengths with subcarrier tones unique to each wavelength. Those tones can be monitored for crosstalk measurements and for routing, as well.

Manuscript received April 17, 1995; revised October 13, 1995. This work was supported by the National Science Foundation Presidential Young Investigator Award ECS-9157089 and The Hughes Aircraft Company.

K.-P. Ho was with Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, CA 94720 USA. He is now with Bellcore, Red Bank, NJ 07701 USA.

J.-M. Kahn is with Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, CA 94720 USA.

Publisher Item Identifier S 0733-8724(96)04077-7.

A. Subcarrier-Multiplexing of Identification Tones

Fig. 1 presents our crosstalk monitoring scheme. At each transmitter, the current modulating the laser includes on-off-keyed data at baseband, multiplexed with a high-frequency tone that identifies the channel. The frequency of the subcarrier tone must be higher than the spectral tail of the data signal; if necessary, the data can be pre-filtered to eliminate crosstalk from the data channel to the subcarrier tone. While the subcarrier channel can carry a header or other control information, for simplicity, we consider the subcarrier channel to be an unmodulated tone. At the receiver, the crosstalk induced by other channels is monitored by measurement of the power of the corresponding identification tones. For example, as shown in the Fig. 1, the crosstalk suffered by the i th channel can be measured from the electrical power at tone frequencies $f_{i\pm 1}, f_{i\pm 2}, \dots$.

Including the high-frequency identification tone, the current modulating the laser of the i th channel is given by

$$I_{m,i}(t) = I_{\text{mod}}\{m \cdot d_i(t) + h[1 + \cos(\omega_{s_i}t + \phi_i)]\} + I_{\text{bias}} \quad (1)$$

where I_{bias} is the bias current to bring the laser above the threshold, $d_i(t) \in \{0, 1\}$ is the data modulation, m and h are the data and subcarrier modulation indices, which are assumed to be identical for all channels, obeying the constraint that $m + 2h = 1$, $\omega_{s_i} = 2\pi f_i$ is the angular frequency of the i th identification tone, and ϕ_i is the initial phase of the i th tone.

We assume that the laser output power is proportional to $I_{m,i} - I_{\text{bias}}$, and we ignore laser chirp, fiber dispersion, fiber and optical amplifier nonlinearity. Those effects can be reduced by using external modulation or a low-chirp laser, dispersion compensation or dispersion-shifted fibers, and limiting the maximum power of each channel, respectively. The optical electric field at the output of the demultiplexer of k th channel can be written as follows:

$$E_k = \sum_{i=1}^N \sqrt{p_{k,i}} \{m \cdot d_i(t) + h[1 + \cos(\omega_{s_i}t + \phi_i)]\}^{1/2} \cdot e^{j(\omega_i t + \delta_i)} \hat{u}_{k,i} + n_k(t) \hat{p} + m_k(t) \hat{q}, \quad k = 1, \dots, N \quad (2)$$

where N is the number of optical channels, ω_i is the optical angular frequency of the i th channel, δ_i is the random phase of the i th channel, $p_{k,i}$ is the optical power of i th channel at the output of k th channel demultiplexer, and $\hat{u}_{k,i}$ is a complex unit vector indicating the light polarization. To include the effect of amplified spontaneous emission (ASE) arising from optical amplifiers, $n_k(t)$ and $m_k(t)$ are broadband, independent, identically distributed Gaussian noise processes traveling in two orthogonal polarizations, identified by the complex unit vectors \hat{p} and \hat{q} , respectively. In (2), we assume that the bandwidth of the demultiplexer is much larger than the bandwidth of each individual channel, so that the filtering effects of each channel can be represented by constant factors $p_{k,i}$ and group-delay dispersion can be ignored. For example, the bandwidth of currently available grating-based demultiplexers is in the range of several Å (corresponding to several tens of GHz at

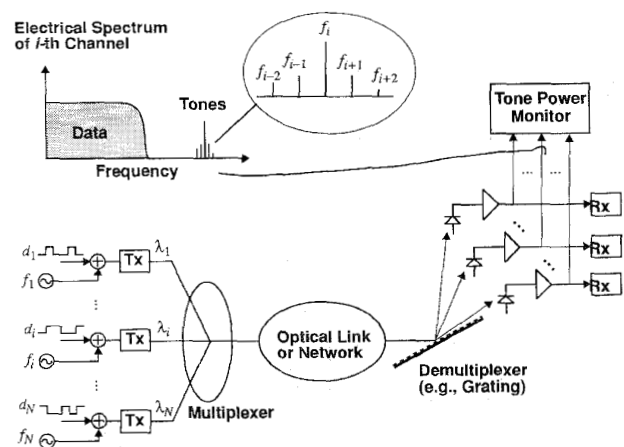


Fig. 1. Crosstalk monitoring scheme using a subcarrier identification tone multiplexed with each individual channel. The level of crosstalk at each wavelength is determined from the power of the corresponding identifying tone in the amplified photocurrent. The electrical spectrum of the i th channel is shown as an example. (Tx: laser transmitter, Rx: receiver).

1.55 μm , as compared with data bandwidths of several GHz, so that our assumption is satisfied.

The photocurrent of the k th channel photodetector is equal to $\rho|E_k|^2$ where ρ is the photodetector responsivity. We assume that each channel has a unique wavelength, so that no crosstalk arises from channels at the same wavelength [7], [19]. Our crosstalk-monitoring method can also be used to measure crosstalk arising from different channels using exactly the same wavelength, as long as different subcarrier frequencies are used. For example, this would be useful in networks that rely upon finite optical loss to permit wavelength reuse. Because the channel separation is much larger than the channel bandwidth, after ignoring the extremely high frequency terms with angular frequencies equal to $\omega_i - \omega_j$, $i \neq j$, the photocurrent plus input-referred noise is

$$i_k(t) = \rho \sum_{i=1}^N p_{k,i} \{m \cdot d_i(t) + h[1 + \cos(\omega_{s_i}t + \phi_i)]\} + i_{\text{ASE}}(t) + i_{\text{shot}}(t) + i_{\text{th}}(t) \quad (3)$$

where $i_{\text{ASE}}(t)$, $i_{\text{shot}}(t)$, and $i_{\text{th}}(t)$ are the Gaussian noise generated by ASE noise (both signal-spontaneous and spontaneous-spontaneous beat noises), shot noise, and thermal noise, respectively. The signal of the k th WDM channel is $\rho p_{k,k} m \cdot d_k(t)$. The crosstalk from i th channel to k th channel is $\rho p_{k,i} m \cdot d_i(t)$, $i \neq k$. The crosstalk can be characterized by the ratio $p_{k,i}/p_{k,k}$, $i \neq k$. All subcarrier tones ω_{s_i} contribute to neither signal nor noise after low-pass filtering at the receiver.

The electrical power of the i th identification tone in the photocurrent of the k th channel demultiplexer is equal to

$$\pi_{k,f_i} = \frac{(\rho p_{k,i} h)^2}{2}, \quad 1 \leq i \leq N. \quad (4)$$

Comparing the electrical power at f_i to that at f_k , we define the crosstalk level in electrical power as

$$\text{XT}_{k,i} = \frac{\pi_{k,f_i}}{\pi_{k,f_k}}, \quad i \neq k. \quad (5)$$

The demultiplexer electrical crosstalk in decibels is $10 \cdot \log_{10} \cdot (XT_{k,i})$ dB, while the optical ratio is $5 \cdot \log_{10} \cdot (XT_{k,i})$ dB. The average electrical signal power of the k th WDM channel is $(1/2)(\rho p_{k,k} m)^2 = (m/h)^2 \pi_{k,f,k}$.

B. Modulation Index of Identification Tones

The crosstalk power of i th channel at the demultiplexer of k th channel can be measured as shown in (5), as long as $\pi_{k,f,i}$ is detectable above the noise. Although we are not concerned here with crosstalk levels less than -30 dB (since they have little effect on system performance), a crosstalk monitoring technique that can monitor very small crosstalk levels is desirable. When $p_{k,i}$ is small, from (4), we see that the identification tone modulation index h must be made sufficiently large for the tone to be detectable. Another factor affecting the tone detectability is the bandwidth of the tone-detection circuit. When the bandwidth is made sufficiently small, a tone can be detected no matter how weak it is.

The required bandwidth depends on how fast we want to detect a change of the crosstalk level. If we want to detect the change of crosstalk within a time τ_c , then the bandwidth of tone detection should satisfy $B_t > 2/\tau_c$. The crosstalk level of most demultiplexers remains fixed for a long time, so that B_t can be as small as 1 kHz, or even smaller. For packet-switched all-optical networks, packets on a wavelength channel are usually transmitted continuously. In most cases, stuff packets or stuff bits may be inserted even if no packet or data is required to be sent. However, in the worst case, when just a single packet lasting for a duration of τ_p is sent, the tone-detection bandwidth must be larger than $B_t > 2/\tau_p$ to avoid the reduction of tone-detection sensitivity.

Assuming that the data occupies a bandwidth of B_d , the signal-to-Gaussian-noise ratio of the k th data channel is

$$\text{SNR}_{dg,k} = \frac{(\rho p_{k,k} m)^2}{2(\zeta_{\text{ASE}} + \zeta_{\text{shot}} + \zeta_{\text{th}}) B_d} \quad (6)$$

where ζ_{ASE} , ζ_{shot} , and ζ_{th} are the power spectral densities arising from ASE noise, shot noise and thermal noise, respectively. The subscript “ dg ” stands for data channel and Gaussian noise. We assume that all of those noises are white, allowing us to express the SNR as (6). The carrier-to-noise ratio (CNR) of the i th identification tone in the k th channel is

$$\text{CNR}_{k,i} = \frac{(\rho p_{k,i} h)^2}{2(\zeta_{\text{ASE}} + \zeta_{\text{shot}} + \zeta_{\text{th}}) B_t} \quad (7)$$

The crosstalk from channels at other wavelengths does not affect the CNR because the subcarrier frequencies of different channels are different. Relating $\text{CNR}_{k,i}$ to $\text{SNR}_{dg,k}$

$$\text{CNR}_{k,i} = \left(\frac{h}{m}\right)^2 \left(\frac{p_{k,i}}{p_{k,k}}\right)^2 \left(\frac{B_d}{B_t}\right) \text{SNR}_{dg,k} \quad (8)$$

If we want to detect a minimum crosstalk level XT_{\min} (ratio in electrical power) with a minimum CNR_{\min} , the minimum h/m is given by

$$(h/m)_{\min} = \sqrt{\frac{B_t \text{CNR}_{\min}}{B_d \text{SNR}_{dg,k} XT_{\min}}} \quad (9)$$

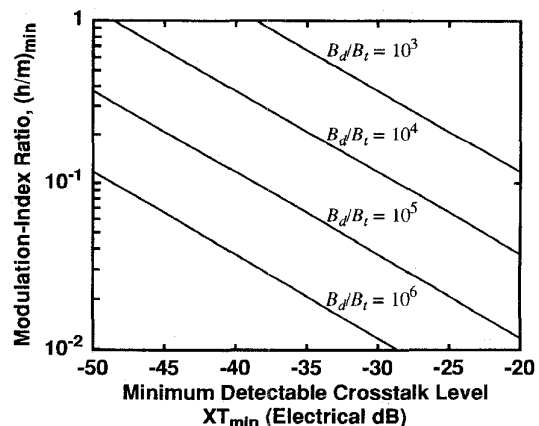


Fig. 2. The minimum modulation-index ratio $(h/m)_{\min}$ required to achieve a minimum detectable crosstalk level XT_{\min} , for different values of B_d/B_t , with $\text{CNR}_{\min} = 10$ dB and $\text{SNR}_{dg,k} = 18.6$ dB.

Assuming the crosstalk is small, so that noise plus crosstalk is approximately Gaussian, the bit-error rate (BER) can be calculated by $\text{BER} = Q(\sqrt{\text{SNR}})$ where SNR is the overall average signal-to-noise ratio, and $Q(x) = (1/2) \text{erfc}(x/\sqrt{2})$ is the Gaussian Q function. For $\text{BER} = 10^{-9}$, $\text{SNR}_{\min} = 18.6$ dB. Due to the effect of crosstalk, we require $\text{SNR}_{dg,k} > 18.6$ dB for $\text{BER} = 10^{-9}$. For a pessimistic estimation of $(h/m)_{\min}$, we can use $\text{SNR}_{dg,k} = 18.6$ dB in (9). Fig. 2 shows $(h/m)_{\min}$ as a function of XT_{\min} , for several values of B_d/B_t . Depending on the system configuration, a wide range of values of B_d/B_t may arise. For example, if $B_d > 1$ GHz, B_t may be as small as 1 kHz and $B_d/B_t > 10^6$. Considering a typical case of $XT_{\min} = -35$ dB, $(h/m)_{\min}$ is in the range of 0.02 to 0.6, depending on B_d/B_t .

III. CROSSTALK CANCELLATION USING LINEAR CANCELLER

We may cancel the crosstalk by using multiple receivers tuned to channels at different wavelengths. Fig. 3 shows a schematic diagram of a linear crosstalk canceller for grating-based systems. The photocurrent from several detected channels are weighted and summed to cancel the crosstalk. However, our method cannot cancel crosstalk from channels having the same wavelength [7], [19]. Based on measured crosstalk levels, the weights can be calculated by a microprocessor that controls the canceller. Each weight tap may be implemented by a multiplier in which one input is the weight given by a digital-to-analog converter. The same technique can be applied using a filter bank [20]. Because the grating spectral response is usually symmetrical with respect to the desired signal channel, we assume here that the number of taps is an odd number.

Assume that n_w taps are used with a weighting vector of $W = (w_1, \dots, w_{n_w})^T$. In order to detect the k th channel, the received photocurrents of n_w receivers $k - (n_w - 1)/2$ to $k + (n_w - 1)/2$ are used. The photocurrent vector is

$$S(t) = Cd(t) + n(t) \quad (10)$$

where

$$S(t) = (i_{k-(n_w-1)/2}(t), \dots, i_k(t), \dots, i_{k+(n_w-1)/2}(t))^T$$

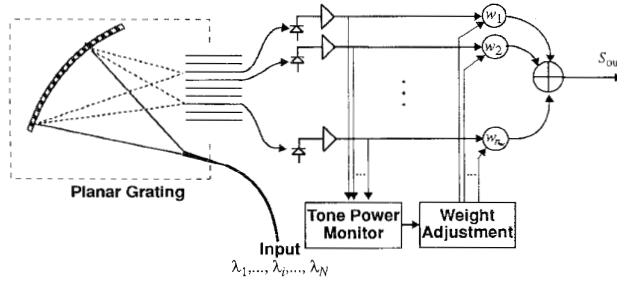


Fig. 3. Linear cancellation of linear crosstalk induced by a planar grating, using a weighted sum of $n_w - 1$ interfering signals with the desired signal. The crosstalk from each interfering channel is monitored by measurement of the electrical power of the corresponding identification tone.

C is an $n_w \times N$ matrix in which $C_{i,j} = \rho m p_{k-(n_w+1)/2+i,j}$ can be determined by the crosstalk monitoring scheme described in the last section, $d(t) = (d_1(t), \dots, d_N(t))^T$ is the transmitted data of all N channels, and $n(t) = (n_1(t), \dots, n_{n_w}(t))^T$ is the input-referred Gaussian noise of each receiver. The received photocurrents are weighted and summed to generate the overall system output $S_{out}(t) = W^T S(t)$. There are several algorithms that can be used to choose the weighting vector.

A. Homogeneous and Nonhomogeneous MSE Minimization

One way to choose the weighting vector is to minimize the mean-squared-error (MSE) of the output signal with respect to the original data. In a homogeneous solution, the goal is to find W for $S_{out}(t) = W^T S(t)$ such that the $MSE \langle e^2 \rangle = \langle [S_{out}(t) - d_k(t)]^2 \rangle$ is minimized. This MSE is given by

$$\langle e^2 \rangle = W^T (C R_{dd} C^T + \sigma_w^2 I) W - 2 W^T C R_{d_k} + \frac{1}{2} \quad (11)$$

where R_{dd} is an $N \times N$ matrix that has $1/2$ on the diagonal and $1/4$ elsewhere, R_{d_k} is a vector that has $1/2$ in the k th position and $1/4$ elsewhere, I is an $N \times N$ identity matrix, and $\sigma_w^2 = (\zeta_{ASE} + \zeta_{shot} + \zeta_{th}) B_d$ is the variance of total Gaussian noise. We assume that all channels experience the same amount of Gaussian noise. The solution of the optimal weighting vector is [21, p. 408]

$$W_o^H = (C R_{dd} C^T + \sigma_w^2 I)^{-1} C R_{d_k} \quad (12)$$

obtained by setting $\nabla_W \langle e^2 \rangle = 0$. Because the matrix C is measured accurately by the crosstalk monitoring scheme, the weights can be calculated deterministically.

Detection of the desired signal $d_k(t)$ can be improved if a constant is added to the sum, i.e., $S_{out}(t) = w_0 + W^T S(t)$, such that the MSE is minimized. In this nonhomogeneous case, the solution for the optimal weighting vector is [21, p. 410]:

$$W_o^{NH} = \frac{C e_k}{2} (C C^T / 4 + \sigma_w^2 I)^{-1} \quad (13)$$

and

$$w_0 = \frac{1}{2} \left(1 - W^T C \sum_{i=1}^N e_i \right), \quad (14)$$

where e_i is a unit vector that has unity at the i th position and zeros elsewhere.

For a given choice of weighting vector, we can now calculate the SNR. The overall Gaussian noise variance of the system is $W^T W \sigma_w^2$, the overall average signal power is $(1/2)(W^T C e_k)^2$, the total variance of the crosstalk signal is $(1/4)W^T C (I - e_k e_k^T) C^T W$. Therefore, the overall SNR is

$$SNR = \frac{(W^T C e_k)^2}{2[W^T W \sigma_w^2 + (1/4)W^T C (I - e_k e_k^T) C^T W]}. \quad (15)$$

The signal-to-crosstalk-interference ratio (SCIR) is $2(W^T C e_k)^2 / (W^T C (I - e_k e_k^T) C^T W)$.

B. Signal-to-Noise Ratio Maximization

Another way to choose the weighting vector W is to maximize the overall SNR in (15). The optimization proceeds by fixing (constraining) the numerator and minimizing the denominator of (15). Using a Lagrange multiplier, the solution is [22, pp. 488–493]

$$W_o^{SNR} = \lambda R_t^{-1} C e_i \quad (16)$$

where $R_t = \sigma_w^2 I + (C(I - e_k e_k^T) C^T) / 4$, λ is a constant multiplier given by $\lambda^{-1} = f^{-1} e_i^T C^T R_t^{-1} C e_i$ where $W^T C e_i = f$ is given by constraining the numerator of (15).

The weighting vector solution, given by (16) is identical to (14) within a scale factor [22, p. 499]. One can show that $W_o^{NH} = c_1 R_t^{-1} C e_i$, where $c_1^{-1} = 1/4 + e_k^T C^T R_t^{-1} C e_k$. Therefore, $W_o^{SNR} = (\lambda/c_1) W_o^{NH}$, i.e., the two weighting vectors are identical within a scale factor, so that identical BER's are provided. Therefore, to save hardware cost, the weight for the desired signal can be set to unity ($w_{(n_w+1)/2} \rightarrow 1$), and the other weights can be scaled correspondingly, i.e., $w_i \rightarrow w_i / w_{(n_w+1)/2}$.

C. Bit-Error-Rate Minimization

The Gaussian approximation overestimates the BER for large crosstalk levels. For a weighting vector of W , we define the crosstalk vector $X = (x_1, \dots, x_N)^T$ as

$$X = C^T W. \quad (17)$$

The output signal can be expressed as

$$S_{out}(t) = x_k d_k(t) + \sum_{i \neq k} x_i d_i(t) + n_o(t) \quad (18)$$

where the first term represents the desired signal, the summation represents crosstalk from other channels, and $n_o(t)$ is Gaussian noise with a variance of $\sigma_n^2 = W^T W \sigma_w^2$. As the crosstalk consists of $N - 1$ channels $d_i(t)$ that can send either "0" or "1", there are a total of 2^{N-1} alternative combinations. Those alternatives can be represented by integers from 0 to $2^{N-1} - 1$, the binary representation of these integers representing whether a given channel transmits "0" or "1". For the l th crosstalk alternative, if $l_i = 1$, the i th channel transmits a "1", and if $l_i = 0$, the i th channel transmits a "0". The integer l is given by

$$l = \sum_{i=1}^{k-1} l_i 2^{i-1} + \sum_{i=k+1}^N l_i 2^{i-2}.$$

Alternatively, the “bits” $l_i, i \neq k$, can be determined from the binary representation of l . The overall BER is

$$\text{BER} = \frac{1}{2^N} \sum_{l=0}^{2^N-1} \left\{ Q\left(\frac{D - x_{T,l}}{\sigma_n}\right) + Q\left(\frac{(I_1 + x_{T,l}) - D}{\sigma_n}\right) \right\} \quad (19)$$

where D is the decision level, $I_1 = x_k$ is the photocurrent when “1” is transmitted, and $x_{T,l} = \sum_{i \neq k} l_i x_i$ is the total crosstalk amplitude for the l th crosstalk alternative. Typically, the decision level D is set to the mean signal-plus-crosstalk amplitude, or $D = \frac{1}{2} \sum_{i=1}^n x_i$; D is also the mean of $S_{\text{out}}(t)$.

In principle, using numerical methods, we can vary the weighting vector W to minimize the BER given by (19). We refer to the weighting vector that yields minimum BER as W_o^{BER} . However, for large number of crosstalk channels, the BER (19) is difficult to evaluate directly. Furthermore, as the number of weight taps increases, the optimization of tap weights becomes extremely difficult. Fortunately, as shown below, the choice of tap weights W_o^{SNR} yields a BER that is indistinguishable from that obtained using W_o^{BER} .

D. Algorithm and Implementation Complexity

In Section II, we discussed the bandwidth required of the tone-detection circuitry to determine the crosstalk matrix C . In this section, we discuss the complexity required to compute the tap weight vectors. We consider a situation that lends itself to very efficient implementation, namely, when all N channels are to be detected at one location. Recall that for each channel, the crosstalk matrix C has n_w rows and N columns. To detect all channels, it might seem necessary to determine $N \times N$ entries from the crosstalk matrices for all N channels, i.e., the N unique rows appearing in the set of crosstalk matrices for the N channels. Often, however, many entries in these crosstalk matrices are very small, because there is little crosstalk between channels that are well-separated in wavelength. Assume that in each row of a crosstalk matrix C , there are only N_C significant entries, i.e., each channel receives significant crosstalk from only $N_C - 1$ other channels. It is obvious that $n_w \leq n_C \leq N$. It is only necessary to determine $N_C \times N$ entries from the crosstalk matrices for the N channels in order to cancel crosstalk in all N channels.

In practice, the algorithm used to determine the weight vectors W should be simple enough to implement in a microprocessor. While the BER-minimization algorithm yields the best performance in theory, this algorithm may be too complex to implement in real time. The MSE-minimization and SNR-maximization algorithms have very nearly the same complexity. We will discuss the complexity of implementing the SNR-maximization algorithm. For a N -channel system using n_w tap weights, after some arithmetic, we find that approximately $(N_C + 2)n_w N$ floating-point operations are needed to find the matrices R_t for all N channels. We note that the matrices R_t of the k th and $(k+1)$ th channels typically have many identical entries, which reduces the number of operations required. Assuming that we do not exploit this dependence

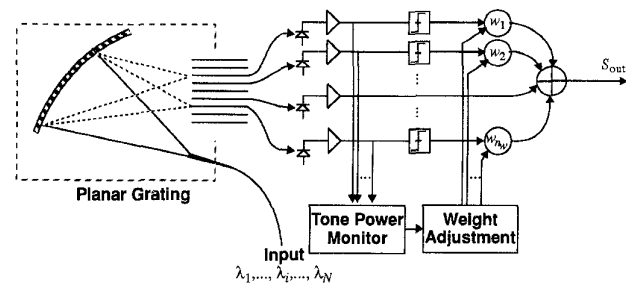


Fig. 4. Nonlinear cancellation of linear crosstalk induced by a planar grating, using weighted sum of $n_w - 1$ detected interfering signals with the desired signal. A decision is made on each interfering signal before the weighted sum is formed.

of the entries in the various matrices R_t , it takes about n_w^3 operations to find the inverses R_t^{-1} for all N channels, and another $n_w (n_w + 1)$ operations to find the weight vectors. Considering a numerical example, if $N = 40, n_w = 5$, and $N_C = 11$, it will require about 9000 operations to find the weight vectors of all 40 channels, or about 225 operations per channel. If we would like to update all weights within 1 ms, a 10 MIPS microprocessor is required. Some digital signal processing (DSP) microprocessors can perform much faster than 10 MIPS.¹ If faster weight updates are needed or there are many more channels, multiple microprocessors may be required.

IV. CROSSTALK CANCELLATION USING NONLINEAR CANCELLER

As an alternative to linear crosstalk cancellation, we can first make decisions on the interfering signals before forming a weighted sum of $n_w - 1$ adjacent channels. This scheme is shown in Fig. 4. Unlike the linear canceller of Fig. 3, at low BER, the nonlinear canceller of Fig. 4 can cancel the crosstalk without enhancing the Gaussian noise. However, at high BER, there may be decision errors in the $n_w - 1$ adjacent channels, leading to incorrect cancellation of crosstalk. Therefore, we would expect the nonlinear canceller to perform better than the linear canceller at low crosstalk levels, but to perform worse at high crosstalk levels. The nonlinear canceller is similar to the decision-feedback equalizer used in digital communication systems [23], but our technique is “decision-directed,” since decisions are not fed back.

In implementing nonlinear crosstalk cancellation, the level of crosstalk can be determined using the crosstalk monitoring scheme described in Section II. In Fig. 4, the photocurrent signal at the k th port is

$$i_k(t) = C_{k,k} d_k(t) + \sum_{i \neq k} C_{k,i} d_i(t) + n_o(t). \quad (20)$$

Assuming that all decisions on the interfering signals are correct, it is obvious that the tap weights should be chosen to completely cancel all crosstalk from the $n_w - 1$ adjacent channels, i.e.,

$$w_i = -C_{k,k-(n_w+1)/2+i} / C_{k,k}, \quad i = 1, \dots, n_w - 1.$$

We make this choice in all that follows.

¹The TMS320C3x and TMS320C4x floating-point DSP microprocessor from Texas Instruments have maximum speeds of 60 MIPS.

The BER performance of the nonlinear canceller is difficult to analyze. To simplify the analysis, we will treat the crosstalk from each of the $n_w - 1$ detected adjacent channels as a binary random variable, but will treat the crosstalk from the remaining channels as Gaussian noise. Since this Gaussian-noise approximation slightly overestimates the BER, it underestimates the improvement using nonlinear canceller. Under this simplifying assumption, the photocurrent (20) can be rewritten as

$$i_k(t) = s_k d_k(t) + \sum_{i=1}^{n_w-1} x_i d_i^*(t) + n_T(t) \quad (21)$$

where $s_k = C_{k,k}$ is the signal amplitude of the k th channel,

$$\begin{aligned} X &= (x_1, \dots, x_{n_w-1}) \\ &= (C_{k,k-(n_w-1)/2}, \dots, C_{k,k-1}, C_{k,k+1}, \dots, \\ &\quad C_{k,k+(n_w-1)/2}), \\ d^* &= (d_1^*, \dots, d_{n_w-1}^*) \\ &= (d_{k,k-(n_w-1)/2}, \dots, d_{k,k-1}, d_{k,k+1}, \dots, \\ &\quad d_{k,k+(n_w-1)/2}) \end{aligned}$$

represent the crosstalk level and input data of the $n_w - 1$ adjacent channels, respectively, and $n_T(t)$, which represents crosstalk from the other adjacent channels plus the receiver noise, is assumed to be Gaussian with variance of

$$\sigma_T^2 = \sigma_w^2 + (1/4) \sum_{i=1}^{k-(n_w+1)/2} C_{k,i}^2 + (1/4) \sum_{i=k+(n_w+1)/2}^N C_{k,i}^2.$$

Let us assume that the $n_w - 1$ detected adjacent channels have BER's of p_{b_i} , $i = 1, \dots, n_w - 1$. Those BER's can be evaluated using (19). There are three possible alternatives for each of these adjacent channels: no decision error, a decision error from "0" to "1," and a decision error from "1" to "0." There are a total of 3^{n_w-1} possible combinations, so it is difficult to evaluate the BER for all of them. Therefore, we will consider the worst case only. When the desired channel transmits a bit of "0" and all adjacent channels transmit "1," if the receiver makes erroneous decisions of "0" on all of the adjacent channels, the crosstalk of the summed output will not be cancelled, and this constitutes one of the worst-case events. Similarly, another worst-case event is when the desired channel transmits a bit of "1," all adjacent channels transmit bits of "0," and the receiver makes decisions of "1" on all of those channels. It can be shown that both worst cases have same BER.

For the worst case, using a decision threshold of $D = s_k/2$, the BER is upper-bounded by

$$\text{BER} < \sum_{l=0}^{2^{n_w-1}-1} P(l) Q\left(\frac{s_k/2 - x_{\max,l}}{\sigma_T}\right) \quad (22)$$

where

$$P(l) = \prod_{i=1}^{n_w-1} (1 - p_{b_i})^{1-l_i} p_{b_i}^{l_i}$$

is the probability that the l th alternative, in which l_i is the i th bit in the binary representation of l , such that the i th channel

is in error if $l_i = 1$ and not in error if $l_i = 0$, $x_{\max,l}$ is the maximum crosstalk in the l th alternative and

$$x_{\max,l} = \sum_{i=1}^{n_w-1} l_i x_i.$$

V. NUMERICAL RESULTS

In the following numerical results, we consider a 40-channel dense WDM system. We assume that all channels arrive at all network nodes with the same optical power, and that all channels are subject to the same amount of loss and amplification, so that the ASE is identical for all channels. For the packet-switching case, we will assume that all channels are transmitting packets. This is the worst case because it has the largest amount of ASE and the largest amount of crosstalk interference. We consider a grating having a Gaussian passband shape [6], [13]. The minimum crosstalk of the grating is -30 dB in terms of optical power, i.e., -60 dB in terms of electrical power. In other words, the spectral response of an output port with center frequency ν_0 and half-width-half-maximum (HWHM) bandwidth $\Delta\nu$ is

$$T(\nu) = \max |\exp\{-\log 2((\nu - \nu_0)/\Delta\nu)^2\}|, 10^{-3}. \quad (23)$$

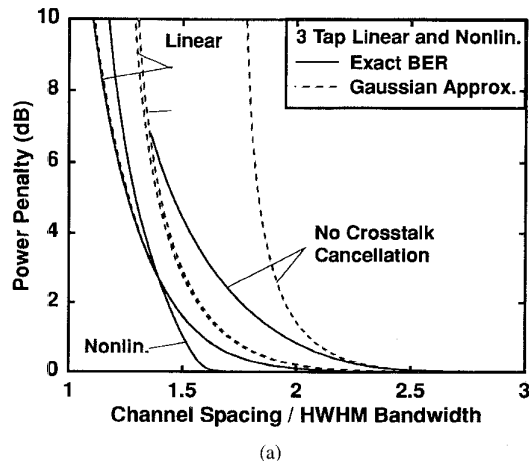
The power penalty is calculated as

$$\delta_e(\text{dB}) = 10 \cdot \log(\text{SNR}_{dg,k}/\text{SNR}_{\min}) \quad (24)$$

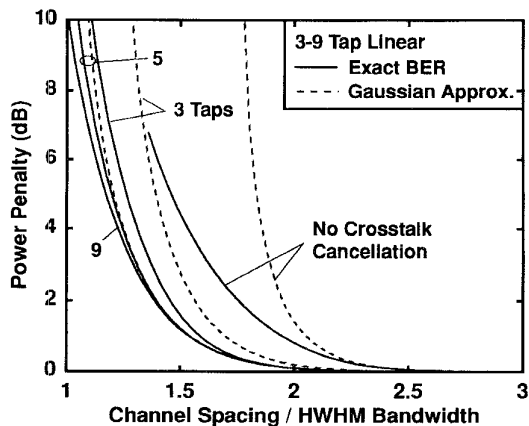
where $\text{SNR}_{\min} = 18.6$ dB is the minimum SNR to achieve $\text{BER} = 10^{-9}$ without crosstalk, and $\text{SNR}_{dg,k}$ is signal-to-Gaussian-noise ratio required to achieve $\text{BER} = 10^{-9}$ in the presence of crosstalk arising from a specific channel spacing. The optical power penalty (in decibels) of the input optical signal is equal to δ_e if the system is dominated by signal-spontaneous beat noise, because then $\text{SNR}_{dg,k}$ is proportional to the input optical power. The optical power penalty of input optical signal is equal to $\delta_e/2$ if the system is dominated by thermal noise or spontaneous-spontaneous beat noise, because then $\text{SNR}_{dg,k}$ is proportional to the square of input optical power.

Fig. 5(a) presents the power penalties for the three-tap linear and nonlinear cancellers, while Fig. 5(b) presents penalties for the three-, five-, and nine-tap linear cancellers. The channel spacings that yield 1 and 3 dB penalties with various schemes are shown in Table I.

First, we compare the performance achieved using three-tap linear cancellers with different choice of the tap weights W . As stated in Section III-C, the optimum choice is W_o^{BER} , which is determined by numerical optimization to minimize the exact BER given by (19). To make this procedure tractable, in evaluation of (19), we have approximated all crosstalk terms having crosstalk levels below 5×10^{-3} as Gaussian noise. We have also used (16) to find W_o^{SNR} , and evaluating the resulting BER using (19), we have found that W_o^{SNR} yields the same BER as W_o^{BER} , i.e., the solution W_o^{SNR} is very nearly optimal. We recall that the choice of W_o^{NH} yields the same performance as W_o^{SNR} (see Section III.B). As shown in Fig. 5(a), the choice of weights W_o^{H} yields a power penalty that is just slightly larger than W_o^{SNR} (or W_o^{NH}).



(a)



(b)

Fig. 5. Power penalty as a function of channel spacing for a 40-channel WDM system using a demultiplexer having a Gaussian-shaped passband and a minimum optical crosstalk of -30 dB. (a) Three-tap linear and nonlinear cancellers. The linear equalizer achieves the same performance using weights W_o^{NH} and W_o^{SNR} , and nearly the same performance using W_o^H . (b) Three-, five- and nine-tap linear equalizers, using weights W_o^{SNR} . With nine taps, the performance predicted by the Gaussian approximation is virtually the same as that calculated using the exact BER formula.

When a three-tap linear canceller with the weights W_o^{SNR} (or W_o^{NH}) is employed, for a 1 dB tolerable power penalty, the exact BER formula (19) indicates that the channel spacing can be decreased from $1.953\Delta v$ to $1.543\Delta v$, corresponding to a 23% increase of channel density. Using the Gaussian approximation to evaluate the BER indicates that for a 1 dB tolerable power penalty, the channel density can be increased by 29%. If the tolerable power penalty is increased to 3 dB, the exact and Gaussian BER expressions indicate that the channel density can be increased by 19 and 34%, respectively.

When a three-tap nonlinear canceller is employed, for 1 and 3 dB permissible power penalties, increases in channel density of 30 and 18% can be obtained. Thus, as expected, the nonlinear canceller performs better than its linear counterpart when crosstalk is weak, but worse when the crosstalk is strong. This is evident in Fig. 5(a) where it can be seen that the nonlinear canceller is superior for channel spacings smaller than about $1.4\Delta v$. When only a very small crosstalk penalty is

TABLE I
CHANNEL SPACING YIELDING 1 AND 3 dB POWER PENALTIES IN A 40 CHANNEL WDM SYSTEM. THE GRATING DEMULTIPLEXER HAS A HALF-WIDTH AT HALF-MAXIMUM BANDWIDTH Δv AND A MINIMUM OPTICAL CROSSTALK OF -30 dB. WHEN ONLY ONE TAP IS USED, THERE IS NO CROSSTALK CANCELLATION

Scheme	Evaluation of BER	Number of Taps	Channel Spacing / Δv for:	
			1-dB Penalty	3-dB Penalty
Linear canceller using W_o^H	Exact	1	1.953	1.629
		3	1.592	1.371
		5	1.546	1.307
	Gaussian approximation	1	2.057	1.894
		3	1.598	1.417
		5	1.547	1.309
Nonlinear canceller	See text	3, 5, 7, 9	1.504	1.380

permissible, the nonlinear canceller permits a channel spacing of about $1.6\Delta v$, while the linear canceller requires a channel spacing of about $2.1\Delta v$. In most cases, the tolerable crosstalk penalty is small, and the nonlinear canceller is more effective than the linear canceller.

We now compare the performance achieved with linear cancellers having three, five, and nine taps. Numerical results show that as number of weights is increased beyond three, the difference in power penalties obtained using W_o^H and W_o^{SNR} (or W_o^{NH}) becomes even smaller. With the choice of W_o^{SNR} , the exact BER expression (19) indicates that for 1 dB tolerable penalties, the three-, five-, and nine-tap linear cancellers yield increases in channel density of 23, 26, and 27%, respectively. For 3 dB tolerable penalties, the corresponding increases are 19, 25, and 26%, respectively. Thus, with the linear canceller, the use of five taps yields a significant improvement over three taps, but little marginal benefit is obtained by using more than five taps. For the case of the nonlinear canceller, our numerical results indicate that there is no significant improvement obtained by increasing the number of taps beyond three.

As mentioned above, typically, the Gaussian approximation overestimates the BER for binary crosstalk, because it approximates the bounded binomial distribution with an unbounded Gaussian distribution. After crosstalk cancellation, the remaining crosstalk is very small, improving the accuracy of Gaussian approximation. In Fig. 5(a), we see that with three taps, the Gaussian approximation substantially overestimates the improvement achieved using crosstalk cancellation. However, in Fig. 5(b) we see that as the number of taps increases, the accuracy of Gaussian approximation improves. For example, for nine taps, there is very little difference between the BER's obtained using the Gaussian approximation and the exact formula (19).

Fig. 6 presents the SCIR versus channel spacing obtained using linear cancellers of one, three, five and nine taps, with tap weights W_o^{SNR} , under conditions where the system power penalties are the same as those in Fig. 5(b). Using crosstalk cancellation, when the channel spacing exceeds approximately $2\Delta v$, crosstalk from other channels is almost

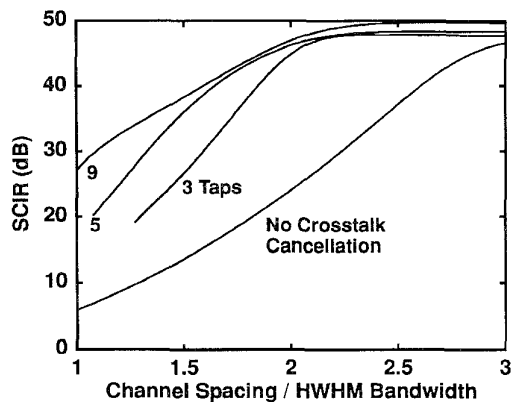


Fig. 6. Signal-to-crosstalk-interference ratio (SCIR) as a function of channel spacing using tap weights W_o^{SNR} , under conditions such that the system power penalties are those shown in Fig. 5(b).

completely cancelled. However, without crosstalk cancellation, when the channel spacing becomes smaller than about $3\Delta\nu$, the crosstalk is dominated by several adjacent channels, and utilization of a canceller having just a few weights can reduce the impact of these dominant interferers, dramatically increasing the SCIR. A three-tap linear canceller can increase the SCIR by more than 20 dB, allowing a BER $< 10^{-9}$ to be achieved even when, in the absence of crosstalk cancellation, the SCIR would be only 10 dB. The nine-tap linear canceller can improve the SCIR by nearly 25 dB, allowing BER $< 10^{-9}$ to be obtained when the SCIR would otherwise be only 6 dB. Additional numerical results indicate that there is no significant difference between the SCIR obtained using W_o^{H} and W_o^{SNR} . We note that Fig. 6 does not show the SCIR of the nonlinear canceller because of the difficulty in defining the SCIR for a nonlinear canceller when decision errors are made in the estimation of crosstalk signals.

VI. OTHER APPLICATIONS OF IDENTIFICATION TONES

In addition to crosstalk monitoring, subcarrier identification tones have a wide range of potential applications in the future all-optical networks. Some well-known applications are for self-routing with and without header encoding [14]–[18] and for contention recovery in all-optical networks [24]. Those applications have been discussed in the literature and will not be discussed further in this paper.

Another application of identification tones is for network management in all-optical networks with optical amplifiers. Optical amplifiers do not have a flat gain spectrum and have potential problems with gain saturation. Although there exist methods to equalize the gain spectrum and place the amplifiers so that all channels are received with approximately equal powers [25], such techniques require monitoring of the received power in each WDM channel, but no simple means has been proposed to perform this monitoring without demultiplexing to individual wavelengths. As shown in Fig. 7, the use of identification tones can be used to enable this monitoring to be performed without optical demultiplexing. A portion of the optical power at the amplifier output is detected using a high-speed photodetector, and the same scheme as

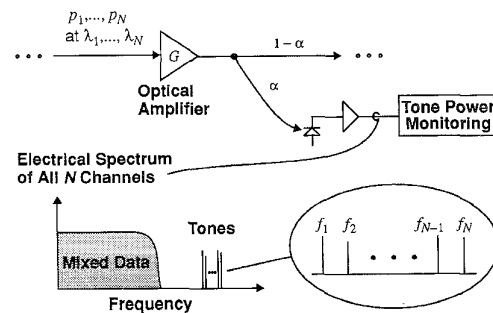


Fig. 7. Use of wavelength-identification subcarrier tones for monitoring power of WDM channels at output of optical amplifier. A fraction of the power is split off and detected. The electrical power of each identification tone indicates the optical power of the corresponding WDM channel.

in Fig. 1 can be used to monitor the tone powers. The proper choice of modulation indices and splitting ratios can be analyzed using the method used in Section II-B. We note that in the present instance, the ASE noise bandwidth and total optical power are approximately N times larger than the corresponding quantities in Section II-B.

VII. CONCLUSION

We propose a scheme for the monitoring and cancellation of crosstalk arising from limited stop-band rejection of optical bandpass filters or gratings in dense WDM systems. The optical carrier at each wavelength is modulated with baseband on-off-keyed data, and also with a high-frequency, out-of-band subcarrier tone unique to that wavelength. At the receiver, the crosstalk from a given wavelength can be determined from measurement of the electrical power of the corresponding tone. We have determined the minimum ratio of the modulation index of the identification tone to the modulation index of digital baseband signal that is required to measure a given crosstalk level.

The crosstalk from other adjacent channels can be cancelled by using multiple receivers to detect different channels, and by forming a weighted sum of the detected signals. This summation may be made linearly or nonlinearly; in the latter case, decisions are made on the detected interfering signals before they are summed with the desired signal. We have described algorithms for determining the tap weights based on measured crosstalk levels. When a linear canceller is used with tap weights that minimize the MSE, maximize the overall SNR or minimize the BER, or when a nonlinear canceller is used, assuming a Gaussian-shaped demultiplexer passband, the channel density can be increased from 20 to 30%, depending on the tolerable crosstalk-induced penalty and the number of tap weights employed. The system can achieve a BER of 10^{-9} even under conditions when, without crosstalk cancellation, the SCIR would be smaller than 10 dB.

REFERENCES

- [1] S. B. Alexander *et al.*, "A precompetitive consortium on wide-band all-optical networks," *J. Lightwave Technol.*, vol. 11, pp. 714–735, 1994.
- [2] C. A. Brackett *et al.*, "A scalable multiwavelength multihop optical network: A proposal for research on all-optical networks," *J. Lightwave Technol.*, vol. 11, pp. 736–753, 1994.

- [3] P. E. Green, Jr. *et al.*, "All-optical packet-switched metropolitan-area network proposal," *J. Lightwave Technol.*, vol. 11, pp. 754–763, 1994.
- [4] A. M. Hill and D. B. Payne, "Linear crosstalk in wavelength-division-multiplexed optical-fiber transmission systems," *J. Lightwave Technol.*, vol. LT-3, pp. 643–651, 1985.
- [5] P. A. Humblet and W. M. Hamdy, "Crosstalk analysis and filter optimization of single- and double-cavity Fabry-Perot filters," *IEEE J. Select. Areas Commun.*, vol. 8, pp. 1095–1107, 1990.
- [6] P. A. Rosher and A. R. Hunwicks, "The analysis of crosstalk in multichannel wavelength division multiplexed optical transmission systems and its impact on multiplexer," *IEEE J. Select. Areas Commun.*, vol. 8, pp. 1108–1114, 1990.
- [7] C. S. Li, C. M. Olsen, and D. G. Messerschmitt, "Analysis of crosstalk penalty in dense optical chip interconnects using single-mode waveguide," *J. Lightwave Technol.*, vol. 9, pp. 1693–1701, 1991.
- [8] P. J. Legg, D. K. Hunter, I. Andonovic, and P. E. Barnsley, "Interchannel crosstalk phenomena in optical time division multiplexed switching networks," *IEEE Photon. Technol. Lett.*, vol. 6, pp. 661–663, 1994.
- [9] P. A. Kirkby, "Multichannel grating demultiplexer receivers for high density wavelength systems," *J. Lightwave Technol.*, vol. 8, pp. 204–211, 1990.
- [10] J. B. D. Soole, A. Scherer, H. P. LeBlanc, N. C. Andreadakis, R. Bhat, and M. A. Koza, "Monolithic InP/In-GaAsP/InP grating spectrometer for the 1.48–1.56 μm wavelength range," *Appl. Phys. Lett.*, vol. 58, pp. 1949–1951, 1991.
- [11] C. Dragone, C. A. Edwards, and R. C. Kistler, "Integrated optics $N \times N$ multiplexer on silicon," *IEEE Photon. Technol. Lett.*, vol. 3, pp. 813–815, 1991.
- [12] H. Takahashi, Y. Hibino, Y. Ohmori, and M. Kawachi, "Polarization-insensitive arrayed-waveguide wavelength multiplexer with birefringence compensating film," *IEEE Photon. Technol. Lett.*, vol. 5, pp. 707–709, 1993.
- [13] M. J. Minardi and M. A. Ingram, "Adaptive crosstalk cancellation in dense wavelength division multiplexed networks," *Electron. Lett.*, vol. 28, pp. 1621–1622, 1992.
- [14] W. I. Way, D. A. Smith, J. J. Johnson, and H. Izadpanah, "A self-routing WDM high-capacity SONET ring network," *IEEE Photon. Technol. Lett.*, vol. 4, pp. 402–405, 1992.
- [15] M. W. Maeda, A. E. Willner, J. R. Wullert, II, J. Patel, and M. Allersma, "Wavelength-division multiple-access network based on centralized common-wavelength control," *IEEE Photon. Technol. Lett.*, vol. 4, pp. 83–85, 1993.
- [16] P. Poggiolini and S. Benedetto, "Performance analysis of multiple subcarrier encoding of packet headers in quasi-optical WDM networks," *IEEE Photon. Technol. Lett.*, vol. 6, pp. 112–114, 1994.
- [17] S. F. Su, A. R. Bugos, V. Lanzisera, and R. Olshansky, "Demonstration of a multiple-access WDM network with subcarrier-multiplexed control channels," *IEEE Photon. Technol. Lett.*, vol. 6, pp. 461–463, 1994.
- [18] S. F. Su and R. Olshansky, "Performance of multiple access WDM networks with subcarrier multiplexed control channels," *J. Lightwave Technol.*, vol. 11, pp. 1028–1033, 1993.
- [19] E. L. Goldstein, L. Eskildsen, A. F. Elrefaie, "Performance implications of component crosstalk in transparent lightwave networks," *IEEE Photon. Technol. Lett.*, vol. 6, pp. 657–660, 1994.
- [20] K.-P. Ho and J. M. Kahn, "Crosstalk cancellation in dense WDM systems using filter-bank receiver," in *OFC '95*, San Diego, CA, 1995, paper TuH7.
- [21] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*. New York: McGraw-Hill, 1984, 2nd ed.
- [22] P. E. Clarkson, *Optimal and Adaptive Signal Processing*. Boca Raton, FL: CRC Press, 1993.
- [23] J. G. Proakis, *Digital Communications*. New York: McGraw-Hill, 1989, pp. 593–601.
- [24] S. F. Su and R. Olshansky, "Use of subcarrier multiplexed acknowledgment tones for contention recovery in WDMA networks," *Electron. Lett.*, vol. 29, pp. 1099–1100, 1993.
- [25] C.-S. Li, F. F.-K. Tong, C. J. Georgiouis, and M. Chen, "Gain equalization in metropolitan and wide area optical networks using optical amplifiers," in *Proc. INFOCOM '94*, Toronto, Ont., Canada, 1994, pp. 130–137.

Keang-Po Ho received the B.S. degree from the National Taiwan University in 1991, and the M.S. and Ph.D. degrees in electrical engineering from the University of California at Berkeley in 1993 and 1995, respectively.

From 1992 to the Spring of 1995, he was a Research Assistant with the Department of Electrical Engineering and Computer Sciences of the University of California at Berkeley, doing research in communication systems. During the Summer of 1994, he was with the IBM T. J. Watson Research Center working on all-optical networks. Currently, he is a Research Scientist at Bellcore, Red Bank, NJ. His current research interests are communication systems, communication theory, signal processing, image and video transmission, and wavelength division multiplexing systems.



Joseph M. Kahn (M'87) received the A.B. degree in physics in 1981, the M.A. degree in physics in 1983, and the Ph.D. degree in physics in 1986 from the University of California, Berkeley. His Ph.D. dissertation was entitled "Hydrogen-Related Acceptor Complexes in Germanium."

He is an Associate Professor with the Department of Electrical Engineering and Computer Sciences at UC Berkeley. From 1987 to 1990, he was a Member of Technical Staff in the Lightwave Communications Research Department of AT&T Bell Laboratories, Crawford Hill Laboratory, Holmdel, NJ, where he performed research on multigigabit-per-second coherent optical fiber transmission systems and related device and subsystem technologies. He demonstrated the first BPSK-homodyne optical fiber transmission system, and achieved world records for receiver sensitivity in multigigabit-per-second systems. He joined the faculty of University of California, Berkeley in 1990, where his research interests include optical fiber communication networks and transmission systems, local-area networks using free-space optical links, and optical interconnects in digital systems.

Dr. Kahn is a recipient of the National Science Foundation Presidential Young Investigator Award, and is a member of the IEEE Communications Society and the IEEE Lasers and Electro-Optics Society. He is serving currently as a Technical Editor of *IEEE Personal Communications Magazine*.