ELECTRONIC EQUALIZATION IN OPTICAL FIBER COMMUNICATIONS

Tülay Adalı, Wei Wang, and Aurenice O. Lima

Department of Computer Science and Electrical Engineering
University of Maryland Baltimore County, Baltimore, MD, 21250

ABSTRACT

Electronic equalizers, which have been used widely in wireless and wireline communications, have recently been recognized as effective solutions for mitigating the impairments in the optical communications channel as well. Now with the increasing availability of voltage-tunable integrated circuits for high speed operation, equalizers, in particular those based on the minimum mean-square error (MMSE) criterion have emerged as practical and cost-effective solutions. Certain properties of the optical domain, however, are different than other communications systems where these equalizers have been used. We study the effects of these properties on the performance of the MMSE equalizers through eigenanalysis of the input autocorrelation matrix.

1. INTRODUCTION

Physical impairments in the optical fiber, in particular, chromatic dispersion, fiber nonlinearities, polarization effects, and amplified spontaneous emission noise from the amplifiers, all interact, limiting the data rate and/or the transmission distances. Solutions for mitigating effects of these impairments are traditionally based on techniques in the *optical* domain, *i.e.*, before the detection. The primary reason for this trend has been the background of researchers working in the field, who are mostly device physicists. Optical compensators, however, rely on adaptive optics and are usually slow in responding to the system degradation, and are expensive and bulky devices. Electrical domain approaches based on signal processing, on the other hand, offer great flexibility in design and can be integrated within the chip sets at the receiver, reducing bulkiness. Also, they can potentially operate after the optical signal has been partially demultiplexed so that electrical processing is done at a lower rate, hence substantially lowering the costs.

The promise of signal processing approaches for optical communications has been noted more than a decade ago [12], but their successful demonstrations for high-speed optical communications have appeared more recently (see e.g. [2] and the references therein). In particular, it has been shown that equalizers based on the MMSE criterion are effective in reducing the penalty due to polarization mode dispersion (PMD), the primary source of inter-symbol and inter-carrier interference (ISI and ICI) in installed terrestrial fiber systems [3]. Experimental results at 10 Gbit/s use SiGe and GaAs technology and implement the adaptive filter in the analog domain through weighted tapped-delay lines using feedforward and decision feedback filter structures [2]. The coefficients of the filters are adapted using gradient-descent type minimization techniques through a control signal such as eye opening or an error monitor [2], or by the practical least mean squares (LMS) algorithm [11].

Research supported in part by the National Science Foundation award, NSF CCR-0123409.

However, certain properties of the optical domain are different than transmission media such as wireless and wireline where the MMSE equalizers have widely been used. In optical systems, bipolar signal transmission formats are seldom used, hence the transmitted signal is non-zero mean and the use of photodetectors that act as square-law devices at the receiver introduces a noncentral and signal-dependent noise into the received signal, the input of the equalizer. In this paper, we present eigenanalysis of the input autocorrelation matrix when the input is a direct-detected nonzero mean signal. We discuss how the properties of this signal affects the eigenvalues of the signal when compared to the zero mean additive noise input that is typically encountered in other communications applications. The results of the eigenanalysis are then applied to PMD equalization, an area where electronic equalizers are shown to be particularly useful and practical. We then show how eye opening penalty can effectively be reduced using equalization with an all-order PMD model and a realistic receiver structure. We use importance sampling to evaluate the performance for very low values of outage probabilities efficiently. Issues in the implementation of MMSE equalizers for PMD mitigation are considered as well as ways to improve the performance, such as by centralizing the input.

2. EIGENANALYSIS OF THE INPUT CORRELATION MATRIX

Autocorrelation matrix of the input describes the behavior of MMSE equalizers, in terms of their misadjustment, and convergence characteristics when they rely on gradient optimization. In this analysis, we are interested in the deviation in the performance of the adaptive algorithm in the presence of noncentral and multiplicative noise as well as noncentral signal. The notation used in the analysis is discrete as even when the filter is implemented in the analog domain, its implementation is through a tapped delay line structure and it is the samples taken at the sampling rate that determine the behavior of the adaptive algorithm used.

Let the output of the photodetector be written as $u(n) = (s(n) + \eta(n))^2$ where $s(n)$ is the transmitted signal, $\eta(n)$ is white amplifier noise, distributed $\mathcal{N}(0, \sigma_0^2)$ in the optical domain, prior to detection. Even though the nonlinearities present in the fiber introduces correlation to the amplifier noise in an optical system, which is the dominant noise term, it can be assumed to be white in the band of interest. Define $v(n) \equiv s^2(n)$, $\omega(n) \equiv \eta^2(n)$, and the signal-dependent noise term $\zeta(n) \equiv 2s(n)\eta(n)$. Define the two autocorrelation matrices $\mathbf{R} = E[\mathbf{u}(n)\mathbf{u}^T(n)]$ and $\mathbf{R}_0 = E[\mathbf{v}(n)\mathbf{v}^T(n)]$ where the sample vectors are defined such that they contain the last M samples, e.g., $\mathbf{v}(n) = [v(n), v(n-1), \ldots, v(n-M+1)]^T$. We can establish the relationship between these two matrices by:

$$\mathbf{R} = \mathbf{R}_0 + \alpha\mathbf{I} + \beta\mathbf{1}\mathbf{1}^T \tag{1}$$

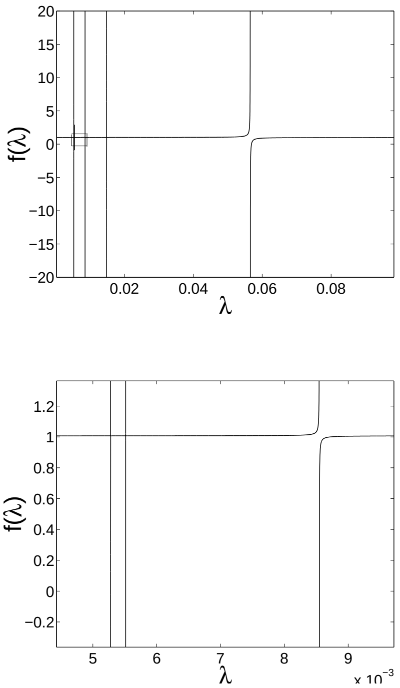

Fig. 1. (a) An example of the characteristics of $f(\lambda)$ for a 5×5 matrix; (b) Close-up of the range shown in the top figure

where $\alpha = \sigma_\omega^2 + 4\mu_v\mu_w$, $\beta = \mu_\omega^2 + 2\mu_v\mu_\omega$, and $\mathbf{1}$ is the $M \times 1$ vector of all 1s. Note that $\mu_\omega = \sigma_0^2$ and the variance of $\omega(n)$ is given by $2\sigma_0^4$.

The noiseless correlation matrix $\mathbf{R}_0$ is a symmetric positive definite matrix and can be written as $\mathbf{R}_0 = \mathbf{Q}\mathbf{\Lambda}_0\mathbf{Q}^T$, where $\mathbf{\Lambda}_0$ is a diagonal matrix with all positive and real eigenvalues $\lambda_1^0 > \lambda_2^0 > \cdots > \lambda_M^0$ on its diagonal, and $\mathbf{Q}$ is the orthogonal matrix with its columns as the corresponding eigenvectors. Define

$$\begin{aligned} \mathbf{A} &\equiv \mathbf{Q}^T\mathbf{R}\mathbf{Q} \\ &= \mathbf{\Lambda}_0 + \alpha\mathbf{I} + \beta\mathbf{Q}^T\mathbf{1}\mathbf{1}^T\mathbf{Q} \equiv \mathbf{G} + \beta\mathbf{q}\mathbf{q}^T \end{aligned} \quad (2)$$

where the diagonals of $\mathbf{G}$ are given by $g_i = \lambda_i^0 + \alpha$, and $\mathbf{q} \equiv \mathbf{Q}^T\mathbf{1} = [q_1, q_2, \ldots, q_M]^T$, *i.e.*, q_i is the sum of the entries of the ith eigenvector. Note that $\mathbf{A}$ and $\mathbf{R}$ are similar matrices, hence they have the same eigenvalues.

The eigenvalues of $\mathbf{A}$ are given by the roots of the *secular equation* [4]:

$$\det(\mathbf{A} - \lambda\mathbf{I}) \equiv f(\lambda) = 1 + \beta\sum_{i=1}^{M}\frac{q_i^2}{g_i - \lambda} \quad (3)$$

The characteristics of $f(\lambda)$ is shown in Figure 1 for a first-order autoregressive (AR) process, with AR coefficient of 0.7 and 24 dB signal-to-noise ratio (SNR) defined as $E[v^2(n)]/E[(\zeta(n) + \omega(n))^2]$. The proximity of zero crossings to the poles in $f(\lambda)$ will allow a useful approximation as we show next. As observed in the figure, roots of $f(\lambda)$ are bounded on the right by g_is, except the maximum one. To obtain an approximation for the maximum eigenvalue, rewrite equation (3) as:

$$f(\lambda) = 1 + \beta\sum_{i=2}^{M}\frac{q_i^2}{g_i - \lambda} + \frac{\beta q_1^2}{g_1 - \lambda}$$

Since $f(\lambda)$ has a pole at g_1, the last term dominates the behavior of $f(\lambda)$ for $\lambda > g_1$, and the second term, the sum of terms associated with the other eigenvalues is almost constant in this region. Thus we can write

$$f(\lambda_{\max}) \approx 1 + \beta\sum_{i=2}^{M}\frac{q_i^2}{g_i - g_1} + \frac{\beta q_1^2}{g_1 - \lambda_{\max}}.$$

Solution of $f(\lambda_{\max}) = 0$ yields

$$\tilde{\lambda}_{\max} = \lambda_{\max}^0 + \alpha + \frac{\beta q_1^2}{1 - \beta\sum_{i=2}^{M}\frac{q_i^2}{g_1 - g_i}}$$

where we have replaced g_1 by its definition given in (2). Note that $\tilde{\lambda}_{\max} > \lambda_{\max}$ where $\lambda_{\max}$ is the *true* maximum eigenvalue of $\mathbf{R}$.

We can make use of a further approximation when the input autocorrelation matrix $\mathbf{R_0}$ is positive. This will be the case when the input is nonnegative, the typical case in optical communications systems. Then, we can use Perron's theorem [5] that states that the maximum eigenvalue $\lambda_{\max}^0$ of a positive matrix is simple and positive, and the eigenvector associated with this eigenvalue has all positive elements. This implies that all other eigenvectors of $\mathbf{R_0}$ have alternating negative and positive entries because of the orthogonality condition among the eigenvectors. Hence, $q_1^2 \gg q_i^2, i = 2, 3, \ldots, M$. Thus, the maximum eigenvalue can further be approximated as

$$\tilde{\lambda}_{\max} \approx \lambda_{\max}^0 + \alpha + \beta q_1^2.$$

Also notice that even when matrix $\mathbf{R_0}$ is not positive, β will typically be small compared to the eigenvalues and such an approximation might be plausible for these cases as well.

From the characteristics shown in Figure 1, note that all roots of equation (3), except the maximum one, are very close to the corresponding poles. This suggests the use of similar arguments to approximate the other eigenvalues of the correlation matrix $\mathbf{R}$. For example, the approximate minimum eigenvalue of $\mathbf{R}$ can be written as

$$\tilde{\lambda}_{\min} = \lambda_{\min}^0 + \alpha + \frac{\beta q_M^2}{1 + \beta\sum_{i=1}^{M-1}\frac{q_i^2}{\lambda_i - \lambda_{\min}}} \approx \lambda_{\min}^0 + \alpha. \quad (4)$$

Thus we can obtain the eigenvalue spread of $\mathbf{R}$ as:

$$\chi(\mathbf{R}) = \lambda_{\max}/\lambda_{\min} \approx \frac{\lambda_{\max}^0 + \alpha + \beta q_1^2}{\lambda_{\min}^0 + \alpha}. \quad (5)$$

Note the effect of the two terms in the autocorrelation matrix given in (1) on the final eigenvalue spread: the bias on the diagonal entries, α decreasing, and the bias on all coefficients, the term β, increasing the eigenvalue spread. We discuss the effect of different cases of noise and signal statistics in the simulations section for the output of a PMD channel as well as the implications for the performance of an MMSE equalizer.

3. EQUALIZATION FOR POLARIZATION MODE DISPERSION

Polarization mode dispersion (PMD) is the primary barrier to achieving single-channel data rates at 40 Gbit/s and beyond in installed terrestrial fiber systems. Electronic equalizers have emerged as one of the most promising solutions for its mitigation. Here, we first briefly discuss the properties of PMD and then the properties of equalizers used for mitigating its effects.

Perturbations that cause loss of circular symmetry in the core and cladding of the fiber lead to birefringence and hence to PMD. The PMD-induced distortion can be considered to be a stationary process in a system whose bit rate is on the order of Gbit/s, and can be characterized by two principal states of polarization (PSP) at a given frequency, which propagate through the fiber with different group velocity. The propagation delay between the two PSP is defined as the differential group delay (DGD), τ. This difference in the arrival times of the two polarization states leads to pulse broadening, *i.e.*, to ISI.

Given a fixed input polarization, the output polarization of the fiber undergoes a rate of rotation on the Poincaré sphere with respect to the frequency that can be characterized by [8]: $\mathrm{d}\mathbf{s}/\mathrm{d}\omega = \boldsymbol{\Omega} \times \mathbf{s}$, where $\mathbf{s}$ is the unit Stokes vector describing the output polarization state and $\boldsymbol{\Omega}$ is the polarization dispersion vector of the fiber. The magnitude of the polarization dispersion vector is equal to the DGD between the two PSP, $|\boldsymbol{\Omega}| = \tau$, while its direction determines the direction of the two orthogonal PSP, $\pm\boldsymbol{\Omega}/|\boldsymbol{\Omega}|$. The higher-order PMD distortion is due to the frequency dependence of the polarization dispersion vector $\boldsymbol{\Omega}$.

The effects of first-order PMD can be represented by the channel response

$$h(t) = \gamma\delta(t) + (1-\gamma)\delta(t-\tau) \qquad (6)$$

where γ, a random variable uniformly distributed in $[0, 1]$, represents the distribution of power between the PSP pair, and τ is the DGD value that is Maxwellian distributed. Hence, the continuous time frequency response has zeros at the frequencies $f_k = \frac{2k+1}{2\tau}$, $k \in Z$ when $\gamma = 0.5$ and its location depends on the value of the DGD, moving into the signal spectrum and introducing severe distortion when the DGD is large (for $\tau > 50$ps in a 10 Gbit/s system). This property exhibits itself as a penalty pole for the PMD equalizers at $\gamma = 0.5$ as observed experimentally in [2]. Hence, the feedforward equalizer will place a large gain in the vicinity of the spectral null to compensate for the distortion, amplifying the additive noise in the system, observed for the optical noise dominated system in [2]. The decision feedback equalizer (DFE), on the other hand, can compensate for the distortion without significant noise amplification when the PMD is severe. Also, important to note is that the pulse shape and the sampling instant determines the response of the discrete channel response that is actually being equalized.

Based on the first-order PMD model, it is also easy to see that high PMD distortion (when γ is close to 0.5 and the DGD is large) introduces significant correlation into the signal, slowing down the convergence of gradient descent type minimization schemes, such as the LMS. Also, as we have shown in Section 2, the noncentral and multiplicative noise characteristics contribute to the increased eigenvalue spread of the equalizer input. They also introduce a bias to the optimal MMSE coefficients [10]. Processing the signal prior to the equalizer by subtracting its mean is one way to improve the conditioning of the input signal, and hence of the gradient descent-based estimation of the filter coefficients.

The selection of modulation formats and the receiver filter characteristics also play an important role in the performance of adaptive equalizers. PMD distortion introduces significant ISI into non return-to-zero (NRZ) pulses while for return-to-zero (RZ) pulses, the amount of ISI introduced is usually limited except in the case of RZ pulses with larger full-width-half-maximum (FWHM) values. The adaptive filter optimizes its coefficients to minimize the ISI at its output and hence works more effectively for NRZ signals, and typically needs fractionally-spaced samples for RZ signals, especially when the pulses are narrow. Thus, as the interplay between the noise and additional ISI introduced into the system is taken into account for receiver design, the implications for the following processing steps, such as adaptive filters, should be carefully accounted for as well.

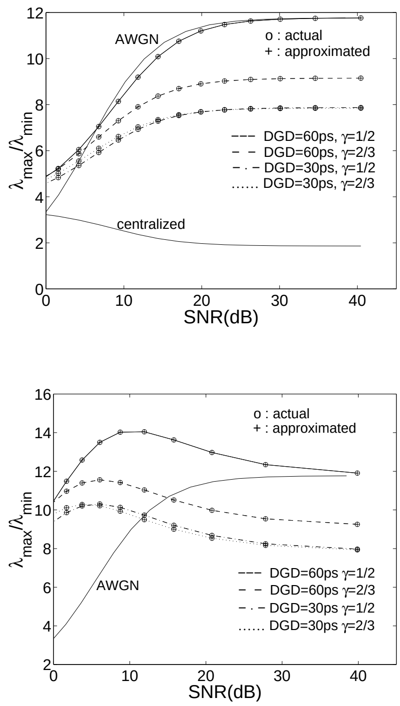

Fig. 2. Eigenvalue spread of $\mathbf{R}$ for a 10 Gbit/s RZ system at the output of a first order PMD channel with (i) both additive and multiplicative noise; (ii) only additive noise.

4. SIMULATIONS

4.1. Eigenanalysis

A first-order PMD channel as given in (6) is simulated for transmission of 10 Gbit/s RZ Gaussian pulses with 50 ps FWHM and the eigenvalue spread at its input is shown in Figure 2. To isolate the effect of multiplicative noise, the true and approximated values of the eigenvalue spread are drawn with and without the multiplicative noise term. As observed in the figure, the given approximations are very close to the actual values for all cases that we considered. An interesting observation is on the role of the signal-dependent noise component. This term adds a positive bias to the diagonal entries of the autocorrelation matrix when the input is nonzero mean thus reducing the eigenvalue spread. As shown on the top plot for a single case (DGD $= 60$ ps and $\gamma = 0.5$), the eigenvalue spread then is close to that of an additive white Gaussian noise (AWGN) channel, *i.e.*, a channel with zero mean noise and the same, noncentral signal. Also shown on the same plot is the eigenvalue spread for the same case after preprocessing the input by subtracting its mean. As shown, this simple procedure that can be implemented with a simple circuitry to block the DC component at the photodetector effectively decreases the eigenvalue spread. Hence, this simple preprocesing will improve the convergence rate of the gradient-descent based algorithm, such as the LMS and its misadjustment, as also suggested in [7]. In the next section, we show that it also decreases the outage probability even in the absence of noise. Note that the decrease in the eigenvalue spread will not be as significant when centralizing an AR type process.

Also important to note is the observation that the noncentral noise alone (no multiplicative noise term), increases the eigenvalue spread. For this case, the spread increases with increasing noise levels but there is a point after which the spread starts to decrease as then, noise starts to dominate in the statistics improving the conditioning of the matrix.

4.2. Outage Probability Analysis

We also perform the outage probability analysis of equalizers for a practical optical transmission system in the presence of all order PMD as in [6], and use importance sampling to evaluate the performance of equalizers in the range that is of interest to systems designers, such as 10^{-6}. The use of importance sampling applied to PMD [1] allows one to efficiently study events that have low probability. In our case, the important rare events are the large DGD values, those in the tails of the Maxwellian pdf of the DGD. Importance sampling biases the Monte-Carlo simulations so that these large DGD configurations occur more frequently than they normally would.

Outage probability is the measure most typically used to evaluate PMD sensitivity. Designers specify a penalty margin for the PMD (typically 2 or 3 dB), and they want to ensure that the outage probability, *i.e.*, the probability that that the actual penalty exceeds this margin is very low. The eye opening is defined as the difference between the the minimum value for a mark (bit 1) I_1 and the maximum value for a space (bit 0) of the electrical signal at the sampling instant, I_0. The eye opening penalty y is the ratio of the eye opening without PMD to that with PMD, and hence the outage probability of β in dBs is given by $\mathrm{Prob}\,(y \geq \beta) = \mathrm{Prob}\left(I_1 - I_0 \leq 10^{-\beta/10}\right)$.

The pulses generated by the modulator are standard NRZ with raise time of 30 ps, optical carrier wavelength of 1532nm and peak power of 1 mW. We use an optical fiber as the source of PMD and assume that the fiber passes ergodically through all possible polarization states with the same PMD. Thus, we analyze ensembles of fibers with the same PMD in order to compute the outage probability for each case. We use 10,000 fiber realizations each 100 km in length and model the fiber using 80 sections of birefringent elements with the coarse step method [9], which reproduces first and higher-order PMD distortions. As the optical filter we use a Gaussian bandpass filter with 60 GHz of full width of half maximum and as an electrical filter a lowpass Bessel filter of 5-th order that is placed after the photodetector. The filters are used primarily for noise suppression. The feedforward filter structures are chosen as symmetric, *i.e.*, the input vector is given by $[u(n-L)\ldots u(n)\ldots u(n+L)]$ where $M = 2L+1$ to account for the fact that the pulses are likely to interact with both the preceding and the following pulses. In Figure 3, we show the outage probability caused by PMD in an NRZ system with average DGD value of $\langle\tau\rangle = 25$ ps, and the performance of two types of

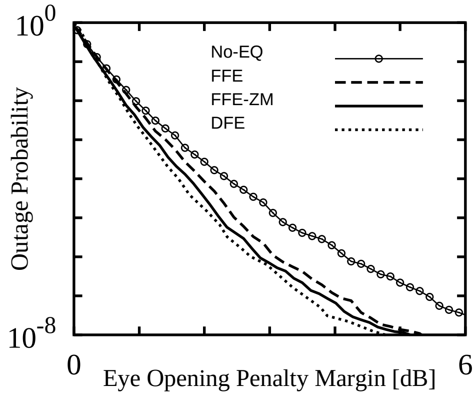

Fig. 3. Outage Probability (probability that the eye opening penalty exceeds the value on the horizontal axis) for direct detection (i) No-EQ: without equalization; (ii) FFE; (iii) FFE-ZM: FFE with centralized input; (iv) DFE.

MMSE equalizers: a feed-forward equalizer (FFE) with 7 taps and a DFE with 5 forward and 3 feedback taps. The use of an equalizer reduces the eye opening penalty by about 1.5 dB at an outage probability of 10^{-7}, and since the system tested here is noise-free, the DFE does not lead to significant performance gain. We also evaluate the performance of the FFE when it is employed after subtraction of the input mean As shown, subtraction of the mean results in some improvement on the performance of the equalizer in reducing the outage probability. As is obvious from equation (5) and the examples shown in Figure 2, this procedure will also significantly improve the convergence of gradient descent learning.

5. REFERENCES

[1] G. Biondini, W. L. Kath, and C. R. Menyuk, "Importance Sampling for Polarization Mode Dispersion," *IEEE Photon. Technol. Lett.*, vol. 14, no. 2, pp. 310–312, Feb. 2002.

[2] H. Bülow, "Electronic equalization of transmission impairments," in *Proc. OFC 2002*, Anaheim, CA, TuF.

[3] H. Bülow, *et al.*,"PMD mitigation at 10 Gbit/s using linear and nonlinear integrated electronic equalizer circuits," *Electron. Letts.*, vol. 36, no. 2, pp. 163–164, Jan. 2000.

[4] J. W. Demmel, Applied Numerical Linear Algebra, SIAM 1997.

[5] R. Bellman, Introduction to Matrix Analysis, McGraw-Hill 1970.

[6] A. O. Lima, T. Adalı, I. T. Lima, Jr., and C. R. Menyuk, "Polarization diversity for PMD mitigation in optical communication systems," in *Proc. ICASSP*, Orlando, FL, 2002.

[7] J.-N. Lin, R. Unbehauen,"Adaptive filtering incorporating a local mean estimation substructure", *IEE Proceedings G*, vol. 140, no. 1, pp. 16–22, Feb. 1993

[8] C. D. Poole and J. Nagel, Optical Fiber Telecommunication, Academic Press, III-A; San Diego, CA, 1997.

[9] D. Marcuse and C. R. Menyuk and P. K. A. Wai, "Application of the Manakov-PMD equation to studies of signal propagation in optical fibers with randomly varying birefringence," *J. Lightwave Tech.*, vol. 15, pp. 1735–1746, 1997.

[10] W. Wang, T. Adalı, W. Xi, and C. R. Menyuk, "On properties of adaptive filters for optical domain," to appear in *Proc. LEOS 2002*, Glasgow, UK.

[11] B. Wedding, A. Chiarotto, W. Kuebart, and H. Bülow, "Fast adaptive control for electronic equalization of PMD," in *Proc. OFC 2001*, Anaheim, CA, TuP4, pp. 1–3.

[12] J. H. Winters and R. D. Gitlin, "Electrical signal processing techniques in long-haul fiber-optic systems," *IEEE Trans. Comm.*, vol. 38, no. 9, pp. 1439–1453, 1990.