On Turbo Code Decoder Performance in Optical-Fiber Communication Systems With Dominating ASE Noise

Yi Cai, *Member, IEEE*, Joel M. Morris, *Senior Member, IEEE*, Tülay Adalı, *Senior Member, IEEE*, and Curtis R. Menyuk, *Fellow, IEEE, Fellow, OSA*

***Abstract*—In this paper, we study the effects of different ASE noise models on the performance of turbo code (TC) decoders. A soft-decoding algorithm, the Bahl, Cocke, Jelinek, and Raviv (BCJR) decoding algorithm [1], is generally used in the TC decoders. The BCJR algorithm is a maximum *a posteriori* probability (MAP) algorithm, and is very sensitive to the noise statistics. The Gaussian approximation of the ASE noise is widely used in the study of optical-fiber communication systems [2]–[8], and there exist standard TCs for additive white Gaussian noise (AWGN) channels. We show that using a MAP decoding algorithm based on the Gaussian noise assumptions, however, may significantly degrade the TC decoder performance in an optical-fiber channel with non-Gaussian ASE noise. To take full advantage of TC, the accurate noise statistics in optical-fiber transmissions should be used in the MAP decoding algorithm.**

***Index Terms*—Amplified spontaneous emission noise, forward error correction, MAP decoding, optical-fiber communication.**

I. Introduction

TURBO codes (TCs) based on soft-decision iterative decoding have been shown to be a very powerful of forward error correction (FEC) code achieving near-Shannon limit performance [9]. As FEC codes have become a practical solution in improving system capacity in fiber communications, the application of TC in fiber transmissions has begun to attract research interest [4], [5], [10].

A soft-decision decoding algorithm, the Bahl, Cocke, Jelinek, and Raviv (BCJR) algorithm [1], is generally used in the TC decoders. The BCJR algorithm is a maximum *a posteriori* probability (MAP) algorithm and requires prior knowledge of the noise statistics in the communication channels and, hence, is very sensitive to the accuracy of the noise statistics in the channel model.

In both undersea and terrestrial systems, the optical amplifiers are critical components, and amplified spontaneous emission (ASE) noise in the optical amplifiers is the major source of noise in optical-fiber channels. ASE noise has an asymmetric statistical nature, and the chi-square distribution model is currently a commonly used model of the ASE noise statistics in the receiver after passing through a square law photodetector and a narrow-band filter [6], [7]. For simplicity, the chi-square distributions are often approximated with Gaussian distributions in characterizing optical-fiber channels [2]–[5]. Moreover, most existing FEC codes are developed and evaluated with the additive white Gaussian noise (AWGN) assumption. Thus, the previous applications and performance evaluations of FEC codes in optical-fiber transmission systems are mostly based on the Gaussian noise approximation or AWGN assumption with little effort to use a more accurate model of the optical-fiber channels.

It has been shown that the theoretical and simulation results of the RS code performance in [3], using BSC and AWGN assumptions, agree well with the experimental measurements. However, RS codes or any other FEC codes using hard-decision algebraic decoding are not sensitive to the exact noise statistics. Because the *a priori* knowledge of the channel noise statistics is not used in algebraic decoding, as long as the channel model assumption gives a good estimate of the uncoded bit error rate (BER), it also gives a good estimate of the algebraic block coded BER.

By contrast, *a priori* knowledge of the channel noise statistics is essential for soft-decision FEC codes that use probabilistic decoding algorithms such as the BCJR algorithm (MAP probability). We show in this paper that the Gaussian approximation of the ASE noise distributions after passing through the photodetector and filter or the AWGN assumption may significantly degrade the performance of TC in optical-fiber channels when used with the BCJR decoding algorithm.

In the following section, we describe and compare three different channel models for the optical-fiber channels with dominating ASE noise. Section III describes the modifications of the BCJR algorithm according to the chi-square noise distribution. Simulation results for the performance of TC decoders based on the different channel models are shown and discussed in Section IV. Finally, Section V concludes the paper.

II. Optical-Fiber Channel Models

ASE noise in optical amplifiers is the major source of errors in optical-fiber systems with a low signal-to-noise ratio (SNR). Generally, the ASE noise leads to asymmetric distributions of marks and spaces after passing through the receiver in the sense that they have different variances and their probability density functions (pdf) are asymmetric. Because of the asymmetric nature of the noise statistics, the Q-factor, which is defined as [8]

$$Q = (I_1 - I_0)/(\sigma_1 + \sigma_0) \tag{2.1}$$

Manuscript received March 7, 2002; revised December 2, 2002.

The authors are with the Department of Computer Science and Electrical Engineering, University of Maryland at Baltimore County, Baltimore, MD 21250 USA (e-mail: ycai@csee.umbc.edu; morris@umbc.edu; adali@umbc.edu; menyuk@umbc.edu).

Digital Object Identifier 10.1109/JLT.2003.809543

is widely used as a SNR measure in optical-fiber channels, where I_1, I_0, σ_1, and σ_0 represent means and variances of the marks and the spaces, respectively. In the following, we describe and compare three different channel models: the chi-square, the asymmetric Gaussian, and the AWGN channel models for ASE noise channels.

A. Chi-Square Channel Model

The pdf of the detected signal I is a function of the energy E of the transmitted optical pulse as well as the power spectral density, N_0, of the ASE noise, as described in [6], [7]. The received marks and spaces have different pdfs that are given by [7]

$$p_1(I) = \frac{1}{N_0}\left(\frac{I}{E}\right)^{(M-1)/2} \exp\left(-\frac{I+E}{N_0}\right) I_{M-1}\left(2\frac{\sqrt{IE}}{N_0}\right) \tag{2.2}$$

$$p_0(I) = \frac{1}{N_0}\,\frac{(I/N_0)^{M-1}\exp(-I/N_0)}{(M-1)!} \tag{2.3}$$

where $M = B_o/B_e$ is the number of modes per polarization state in the received optical spectrum, B_o and B_e are the optical bandwidth and the electrical bandwidth, respectively, of the system at the detector, and I_{M-1} denotes the $(M-1)$th modified Bessel function of the first kind. The means and variances of the received marks and spaces can be derived from the pdfs given in (2.2) and (2.3) $I_1 = MN_0 + E$, $\sigma_1^2 = MN_0^2 + 2EN_0$, $I_0 = MN_0$, $\sigma_0^2 = MN_0^2$, respectively [7]. We can also obtain $\sigma_1^2 = 2(I_1 I_0 - I_0^2)/M + \sigma_0^2$ from the above formulae for I_1, σ_1, I_0, and σ_0 [7]. With these results and the definition of Q in (2.1), we can evaluate I_1, I_0, σ_1, and σ_0 as functions of the system parameters B_o, B_e, and Q, as [11]

$$\begin{aligned} I_1 &= 2Q\sqrt{\frac{B_o}{B_e}} + 2Q^2 + \frac{B_o}{B_e}, \qquad I_0 = \frac{B_o}{B_e} \\ \sigma_1 &= \sqrt{\frac{B_o}{B_e}} + 2Q, \qquad \sigma_0 = \sqrt{\frac{B_o}{B_e}} \end{aligned} \tag{2.4}$$

where N_0 is normalized to 1.

We see that the marks have a noncentral chi-square distribution, the spaces have a central chi-square distribution, and both are asymmetric pdfs with $2M$ degrees of freedom [7]. Thus, we call this model the chi-square channel model. We note that this model does not take into account signal distortion due to optical-fiber transmission, and it assumes an ideal integrate and dump receiver. It is possible to obtain a more accurate characterization of the noise distributions by using more accurate models of the transmission and the receiver [12], [13]. However, numerical methods must be used to generate the pdfs. By contrast, the chi-square assumption yields a simple analytical form for the pdfs of the marks and the spaces and is a substantially better approximation to the actual noise statistics than is the often-used Gaussian approximation. Thus, this model is sufficient to establish the deficiencies in the Gaussian approximation when used in the turbo decoder, which is the main point of our paper. Moreover, it is a useful starting point for carrying out more sophisticated analyses based on the considerations of [12] and [13].

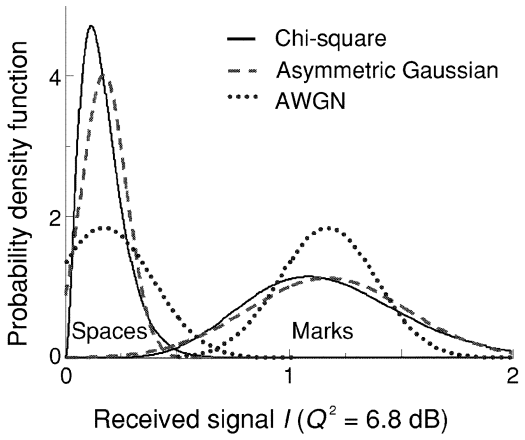

Fig. 1. Probability density functions of the chi-square, asymmetric Gaussian, and AWGN for $M = 3$, $Q^2 = 6.8$ dB, where $(I_1 - I_0)$ is normalized to 1.

Thus, our work in this paper is based on the assumption that the chi-square model is accurate.

B. Asymmetric Gaussian Channel Model

For simplicity of analytical studies of the noise and the induced error probability, Gaussian pdfs with the same means and variances as the chi-square distributions are commonly used. The Gaussian approximation is given by

$$p_1(I) = \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left(\frac{-(I-I_1)^2}{2\sigma_1^2}\right) \tag{2.5}$$

$$p_0(I) = \frac{1}{\sqrt{2\pi\sigma_0^2}} \exp\left(\frac{-(I-I_0)^2}{2\sigma_0^2}\right). \tag{2.6}$$

Note that the detected signal I, as shown in [6] and [7], is a sum of $2M$ independent random variables. From the central limit theorem, the Gaussian approximation can be a good model for both $p_1(I)$ and $p_0(I)$ for large M. However, for small M, as is the case for DWDM systems, and at low Q, the Gaussian distribution is not a good approximation of the chi-square distribution as shown in Fig. 1. Because the marks and spaces have different variances in this model, we call it the asymmetric Gaussian channel model to distinguish it from the AWGN channel model.

C. AWGN Channel Model

The AWGN channel model is the most widely used channel model in error correction code analysis and development. In AWGN channels, marks and spaces have Gaussian distributions with the same variance. This property can significantly simplify the log likelihood ratio (LLR) calculations in the MAP or other soft-decision decoding algorithms. In optical-fiber channels, however, the computational simplification due to the AWGN assumption for the noise diminishes the accuracy of the channel model and, thus, degrades the decoder performance.

To relate a given chi-square channel to an AWGN channel, we first calculate the bit error rate (BER), P_{chi}, of the chi-square channel with optimal hard-decision detection. Then, we construct an AWGN model with the same means and BER as that

of the chi-square channel. Thus, the standard deviation σ_{AWG} in the corresponding AWGN model is given by

$$\sigma_{\mathrm{AWG}} = \frac{\sqrt{E}}{2\mathrm{erfc}^{-1}(2P_{\mathrm{chi}})} \tag{2.7}$$

where E is the energy of the transmitted optical pulse and $\mathrm{erfc}^{-1}(x)$ is the inverse complementary error function.

D. Comparison of the Three Models

As described above, the asymmetric Gaussian and AWGN models are actually Gaussian approximations for the chi-square noise distributions, while the first one keeps the same signal means and variances as the chi-square distribution, the second one keeps the same means and hard-decision BER as the chi-square distribution.

Fig. 1 plots the pdfs of the chi-square, the asymmetric Gaussian, and the AWGN distributions in an optical-fiber transmission system with $Q^2 = 6.8$ dB and $M = 3$. It shows that the asymmetric Gaussian pdfs look more similar to the chi-square pdfs than the AWGN pdfs. But we still see that the central chi-square pdf of the spaces is quite different from the asymmetric Gaussian approximation, even in the central part of the pdfs. The difference between the pdfs of the marks, although not as significant as that between the pdfs of the spaces, is clearly observed. Because the optical detector is a square-law device and its output is thus always a positive electrical voltage, the probability of a negative output voltage is zero. The chi-square pdfs have zero probability density for an output voltage less than zero. By contrast, the Gaussian approximation yields distributions that are not zero when the output voltage is less than zero.

Fig. 1 also clearly shows the asymmetric distribution of the marks and spaces. For both the chi-square and the asymmetric Gaussian pdfs, the variance of the marks are much larger than that of the spaces. The difference between the variances comes from the signal/noise beat term [6], [7].

Fig. 2 compares the three channel models from another point of view, the LLR of the received signal defined as

$$L(I) \equiv \log\left(\frac{p(u=1|I)}{p(u=0|I)}\right) \tag{2.8}$$

where u represents the transmitted signal and I represents the received signal. For equally likely transmitted marks and spaces, i.e., $p(u=1) = p(u=0)$, $L(u)$ can be expressed as the ratio of the pdfs as

$$L(I) = \log\left(\frac{p_1(I)}{p_0(I)}\right). \tag{2.9}$$

The MAP decoder performance highly depends on the accuracy of the LLR of received signals. Therefore, the accuracy of the approximated LLR is more critical than the accuracy of the approximated pdf in choosing a proper channel model used in MAP decoding.

In Fig. 2, we show hatched areas in which the Gaussian approximations have different LLR signs from that of the chi-square model. The size of the shadowed area can be intuitively used as a measure of how close the Gaussian models approximate the chi-square model in terms of the LLR of the received signal. The area hatched with vertical lines corresponds to the asymmetric Gaussian approximation, and the area hatched with oblique lines corresponds to the AWGN approximation. In Fig. 2(a), the system has a $Q^2 = 6.8$ dB. We find that the area hatched with oblique lines is significantly larger than the area hatched with vertical lines. Hence, we predict that the asymmetric Gaussian model approximates the chi-square channel better than the AWGN model at a Q^2 of about 6.8 dB and, hence, use of the asymmetric Gaussian model will lead to better MAP decoder performance. However, Fig. 2(b) shows that in a system with $Q^2 = 0$ dB, the two Gaussian models have almost identical hatched areas. Outside the hatched areas, the LLRs of the AWGN model are more similar to that of the chi-square channel and, hence, use of the AWGN model will result in better MAP decoder performance than will use of the asymmetric Gaussian model.

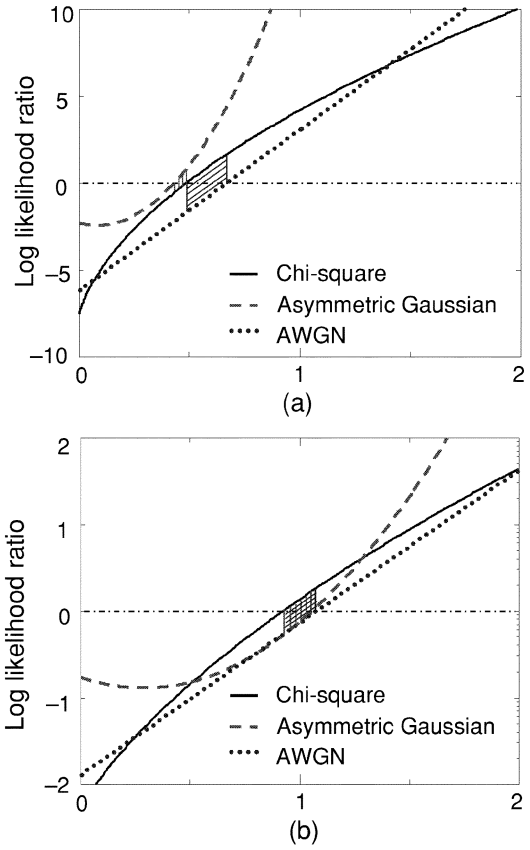

Fig. 2. LLRs based on the chi-square, asymmetric Gaussian, and AWGN for $M = 3$: (a) $Q^2 = 6.8$ dB and (b) $Q^2 = 0$ dB, where $(I_1 - I_0)$ is normalized to 1.

The above LLR comparisons of the two Gaussian models show that neither one is always better than the other one in approximating a chi-square channel, and which approximation is better depends on the Q-factor. Moreover, we observe in Fig. 2 that the use of either Gaussian approximation may significantly skew the ASE noise statistics in MAP decoding and, thus, degrade the TC performance.

III. BCJR Algorithm With Non-AWGN Distributions

A. Standard Rate 1/3 TC

The BCJR algorithm is a recursive algorithm for the MAP decoding of the received noisy codeword $\mathbf{Y} = (y_1^s, \ldots, y_N^s, y_1^{ip}, \ldots, y_N^{ip}, \ldots)$ [1], [14], where y_k^s represents a received information bit corresponding to the transmitted information bit u_k, and y_k^{ip} represents a received parity-check bit corresponding to the transmitted parity-check bit x_k^{ip} generated by the ith constituent encoder. We note that $i = 1, 2$ for our rate 1/3 TC, where each constituent convolutional encoder has rate 1/2. In the ith constituent MAP decoder for TC, the information bit u_k in the transmitted codeword $\mathbf{X} = (u_1, \ldots, u_N, x_1^{ip}, \ldots, x_N^{ip}, \ldots)$ is estimated based on the received noisy codeword $\mathbf{Y}$ by

$$\hat{u}_k = \begin{cases} 1, & \text{if } L(u_k) > 0 \\ 0, & \text{if } L(u_k) < 0 \end{cases} \tag{3.1}$$

where $L(u_k)$ is the LLR given the received codeword $\mathbf{Y}$, defined as

$$L(u_k) \equiv \log\left(\frac{P(u_k = 1|\mathbf{Y})}{P(u_k = 0|\mathbf{Y})}\right), \qquad 0 \le k \le N. \tag{3.2}$$

The key to the BCJR algorithm is to decompose the *a posteriori* probability into three factors α_{k-1}, γ_k, and β_k (we refer to the subscribe k as "time k" in the following discussions). These factors relate the decision on u_k to the previous, current, and future observations, respectively, as

$$P(u_k = u \text{ causing state transition } s' \text{ to } s|\mathbf{Y}) = \frac{1}{p(\mathbf{Y})} \sum_{s',\, s\in S} \alpha_{k-1}(s')\gamma_k(s',\, s)\beta_k(s). \tag{3.3}$$

Here: 1) $S = \{s_1, \ldots, s_k, \ldots, s_N\}$ is the set of all constituent encoder states, the state pair $(s',\, s)$ represents a state transition from $(s_{k-1} = s')$ to $(s_k = s)$; 2) $\alpha_{k-1}(s') = p[s_{k-1} = s', (y_1^s, \ldots, y_{k-1}^s, y_1^{ip}, \ldots, y_{k-1}^{ip})]$ is a probability measure for state s' at time $k-1$ that depends only on the past observations, i.e., the received information and parity-check bits before time k; 3) $\beta_k(s) = p[(y_{k+1}^s, \ldots, y_N^s, y_{k+1}^{ip}, \ldots, y_N^{ip})\,|s_k = s]$ is a probability measure for state s at time k that depends only on the future observations, i.e., the received information and parity-check bits after time k; and 4) $\gamma_k(s',\, s)$ is a probability measure connecting state s' at time $k-1$ to state s at time k that depends only on the present observation $(y_k^s,\, y_k^{ip})$. The $\gamma_k(s',\, s)$ can be written as

$$\gamma_k(s',\, s) = P(u_k)p(y_k^s,\, y_k^p|u_k) \cong P(u_k)p(y_k^s|u_k)p(y_k^p|x_k^p) \tag{3.4}$$

and $\alpha_{k-1}(s')$ and $\beta_k(s)$ can be computed recursively as functions of $\gamma_k(s',\, s)$ given by

$$\alpha_k(s) = \sum_{s'\in S} \alpha_{k-1}(s')\gamma_k(s',\, s) \tag{3.5}$$

and

$$\beta_{k-1}(s') = \sum_{s\in S} \beta_k(s)\gamma_k(s',\, s) \tag{3.6}$$

respectively [14].

We observe that $\gamma_k(s',\, s)$ depends on the conditional pdfs of the received signals and is the key factor in the BCJR algorithm. Hence, the performance of the BCJR algorithm depends strongly on the accuracy of the noise model.

As shown in Fig. 1, the differences between the pdfs of the noise with the chi-square distribution and the Gaussian approximations are not negligible, especially at low Q as in the case of $Q^2 = 6.8$ dB. An obvious question is, therefore, can better TC performance be achieved by modifying the standard formula of $\gamma_k(s',\, s)$, which uses the AWGN model, to a new formula using the more accurate chi-square distribution model given by (2.2) and (2.3), rewritten here as

$$p(y_k|x_k = 1) = \frac{1}{N_0}\left(\frac{y_k}{E}\right)^{(M-1)/2} \exp\left(-\frac{y_k + E}{N_0}\right) I_{M-1}\left(2\frac{\sqrt{y_k E}}{N_0}\right), \qquad y_k \ge 0 \tag{3.7}$$

$$p(y_k|x_k = 0) = \frac{1}{N_0}\frac{(y_k/N_0)^{M-1}\exp(-y_k/N_0)}{(M-1)!}, \qquad y_k \ge 0 \tag{3.8}$$

where y_k represents y_k^s or y_k^p, x_k represents u_k or x_k^p, E is the transmitted signal energy, $N_0/2$ is the two-sided power spectral density of the ASE noise, and $2M$ is the dimensionality of the optical signal space. When we substitute (3.7) and (3.8) into (3.6), we obtain (3.9), as shown at the bottom of the next page. Defining

$$L^e(u_k) \equiv \log\left(\frac{P(u_k = 1)}{P(u_k = 0)}\right) \tag{3.10}$$

we may write

$$P(u_k) = \left(\frac{\exp[-L^e(u_k)/2]}{1 + \exp[-L^e(u_k)]}\right)\exp\left(\frac{(2u_k - 1)L^e(u_k)}{2}\right). \tag{3.11}$$

Note that (3.9) and (3.11) can be substituted into (3.2) and (3.3) to calculate the LLR. Thus, all the common terms in the four cases in (3.9) can be removed to simplify the calculations. Then, the $\gamma_k(s',\, s)$ can be calculated with

$$\gamma_k(s',\, s) \cong \begin{cases} c_1 \exp(L^e(u_k))\, I_{M-1}\left(a\sqrt{y_k^s}\right) I_{M-1}\left(a\sqrt{y_k^p}\right), & u_k = 1,\, x_k^p = 1 \\ \exp(L^e(u_k))\, I_{M-1}\left(a\sqrt{y_k^s}\right)(y_k^p)^b, & u_k = 1,\, x_k^p = 0 \\ I_{M-1}\left(a\sqrt{y_k^p}\right)(y_k^s)^b, & u_k = 0,\, x_k^p = 1 \\ c_0(y_k^s y_k^p)^b, & u_k = 0,\, x_k^p = 0. \end{cases} \tag{3.12}$$

where a, b, c_0, and c_1 are constants given by

$$a = 2\frac{\sqrt{E}}{N_0}, \qquad b = \frac{M-1}{2}$$

$$c_1 = (M-1)!\left(\frac{N_0}{\sqrt{E}}\right)^{M-1}\exp\left(-\frac{E}{N_0}\right)$$

$$c_0 = \left(\sqrt{E}\Big/ N_0\right)^{M-1}\exp\left(\frac{E}{N_0}\right)\Big/ (M-1)!.$$

Defining

$$\gamma_k^e(s',\, s) \equiv \begin{cases} c_1^{u_k} I_{M-1}\left(a\sqrt{y_k^p}\right), & x_k^p = 1 \\ c_0^{1-u_k}(y_k^p)^b, & x_k^p = 0 \end{cases} \tag{3.13}$$

the LLR can be calculated iteratively to yield

$$L(u_k) = \log\left(\frac{I_{M-1}\left(a\sqrt{y_k^s}\right)}{(y_k^s)^b}\right) + L^e(u_k) + \log\left(\frac{\sum\limits_{S^+}\tilde{\alpha}_{k-1}(s')\gamma_k^e(s',\, s)\tilde{\beta}_k(s)}{\sum\limits_{S^-}\tilde{\alpha}_{k-1}(s')\gamma_k^e(s',\, s)\tilde{\beta}_k(s)}\right) \tag{3.14}$$

where S^+ is the set of $(s',\, s)$ caused by $u_k = 1$, and S^- is similarly defined for $u_k = 0$. The first term on the right side of (3.14), which depends on the currently observed information bit and the channel SNR, is sometimes called the channel value. The second term $L^e(u_k)$ represents any *a priori* information provided (extrinsic information received) by the other decoder, and the third term represents extrinsic information passed to the other decoder.

B. Punctured TC

Punctured TC is more practical than the standard TC in optical-fiber transmission systems because of the higher code rates that can be obtained from lower code rate codes. Puncturing can be implemented by deleting some parity and/or information bits at the output of the encoder [14], [15]. At the input of the decoder, the signals corresponding to the punctured bits are set to the same value as the optimal hard-decision threshold $I_{\rm opt}$ [14]. The reason follows. If we assume that the pdfs of the spaces and the marks, $p(x|0)$ and $p(x|1)$, cross at the point $(I_{\rm cross}, p_{\rm cross})$ and satisfy the conditions

$$\begin{aligned} &\text{(a) } p(x|0) > p(x|1), && \text{for all } x < I_{\rm cross} \\ &\text{(b) } p(x|0) < p(x|1), && \text{for all } x > I_{\rm cross} \end{aligned}$$

then (a) and (b) are sufficient conditions to imply that $I_{\rm opt}$, which is optimal in the sense of yielding the minimum hard-decision detection error probability, is the crossover point of the two pdf curves, i.e., $I_{\rm opt} = I_{\rm cross}$.

Suppose $I_{\rm opt} \neq I_{\rm cross}$, there are only two possible cases, $I_{\rm opt} < I_{\rm cross}$ or $I_{\rm opt} > I_{\rm cross}$. First, consider the case when $I_{\rm opt} < I_{\rm cross}$. With condition (a), we have $p(x|0) > p(x|1)$ for $x \in [I_{\rm opt}, I_{\rm cross})$. Then, the minimum hard-decision detection error probability can be expressed as

$$\begin{aligned} P_{\min} &= \int_{-\infty}^{I_{\rm opt}} p(I|1)\, dI + \int_{I_{\rm opt}}^{\infty} p(I|0) dI \\ &= \int_{-\infty}^{I_{\rm cross}} p(I|1)\, dI + \int_{I_{\rm cross}}^{\infty} p(I|0)\, dI \\ &\quad + \int_{I_{\rm opt}}^{I_{\rm cross}} [p(I|0) - p(I|1)] dI \\ &= P_{\rm cross} + P_{\rm ext} \end{aligned} \tag{3.15}$$

where $P_{\rm cross}$ is actually the hard-decision detection error probability with $I_{\rm cross}$ as the decision threshold, and $P_{\rm ext} > 0$. Thus, we obtain $P_{\min} > P_{\rm cross}$, which contradicts the definition of $P_{\min}$. Hence, $I_{\rm opt} < I_{\rm cross}$ is not possible. Similarly, with condition (b), we can prove that $I_{\rm opt} > I_{\rm cross}$ is also not possible and, hence, $I_{\rm opt} = I_{\rm cross}$.

This proof leads to the straightforward likelihood ratio result, i.e., if we set punctured bits to the same value as the optimal hard-decision threshold $I_{\rm opt}$, then

$$\begin{aligned} &\frac{p(\text{preset signal value for punctured bit}|\text{punctured bit} = 0)}{p(\text{preset signal value for punctured bit}|\text{punctured bit} = 1)} \\ &\qquad = \frac{p(I_{\rm opt}|0)}{p(I_{\rm opt}|1)} = \frac{p(I_{\rm cross}|0)}{p(I_{\rm cross}|1)} = 1. \end{aligned} \tag{3.16}$$

Obviously, a likelihood ratio equal to 1 (and LLR $= 0$) is the best guess for the punctured bits in the sense of achieving minimum error probability. Hence, $I_{\rm opt}$ is the best value to use for those virtual signals corresponding to the punctured bits. Note that the chi-square, the asymmetric Gaussian, and the AWGN distributions all satisfy the two conditions mentioned above and, hence, the proof and statements made above are valid for them.

$$\gamma_k(s',\, s) \cong \begin{cases} P(u_k)\dfrac{1}{N_0^2}\left(\dfrac{y_k^s y_k^p}{E^2}\right)^{(M-1)/2} \exp\left(-\dfrac{y_k^s + y_k^p + 2E}{N_0}\right) I_{M-1}\left(2\dfrac{\sqrt{y_k^s E}}{N_0}\right) I_{M-1}\left(2\dfrac{\sqrt{y_k^p E}}{N_0}\right), & u_k = 1,\, x_k^p = 1 \\ P(u_k)\dfrac{1}{N_0^2}\left(\dfrac{y_k^s y_k^p}{E}\right)^{(M-1)/2} \exp\left(-\dfrac{y_k^s + y_k^p + E}{N_0}\right) I_{M-1}\left(2\dfrac{\sqrt{y_k^s E}}{N_0}\right) \dfrac{\left(\sqrt{y_k^p}/N_0\right)^{M-1}}{(M-1)!}, & u_k = 1,\, x_k^p = 0 \\ P(u_k)\dfrac{1}{N_0^2}\left(\dfrac{y_k^s y_k^p}{E}\right)^{(M-1)/2} \exp\left(-\dfrac{y_k^s + y_k^p + E}{N_0}\right) I_{M-1}\left(2\dfrac{\sqrt{y_k^p E}}{N_0}\right) \dfrac{\left(\sqrt{y_k^s}/N_0\right)^{M-1}}{(M-1)!}, & u_k = 0,\, x_k^p = 1 \\ P(u_k)\dfrac{1}{N_0^2}\dfrac{\left(y_k^s y_k^p/N_0^2\right)^{M-1} \exp\left(-\frac{y_k^s + y_k^p}{N_0}\right)}{[(M-1)!]^2}, & u_k = 0,\, x_k^p = 0. \end{cases} \tag{3.9}$$

For the AWGN model, we find $I_{\rm opt} = (I_0 + I_1)/2$. For the chi-square and asymmetric Gaussian models, we find that $I_{\rm opt}$ is given by the solutions to the transcendental equations

$$\left(\frac{I_{\rm opt}}{E}\right)^{M-1/2} \exp\left(-\frac{E}{N_0}\right) I_{M-1}\left(2\,\frac{\sqrt{I_{\rm opt}E}}{N_0}\right) = \frac{(I_{\rm opt}/N_0)^{M-1}}{(M-1)!} \tag{3.17}$$

$$\left(\frac{I_1 - I_{\rm opt}}{\sigma_1}\right)^2 + \ln\sigma_1^2 = \left(\frac{I_{\rm opt} - I_0}{\sigma_0}\right)^2 + \ln\sigma_0^2 \tag{3.18}$$

respectively. We see that for the chi-square distribution, there is no closed-form formula in evaluating the optimal hard-decision threshold. For the asymmetric Gaussian case, the solution is quite complex. Therefore, in addition to the Gaussian approximation of the ASE noise distribution, the hard-decision threshold is customarily set so that the two transition probabilities are equal, which implies a binary symmetric channel assumption. Then, the hard-decision threshold can be simplified to [6]

$$I_{\rm th} = \frac{\sigma_0 I_1 + \sigma_1 I_0}{\sigma_0 + \sigma_1}. \tag{3.19}$$

However, it has been shown in [16] and [17] that evaluating $I_{\rm opt}$ with (3.19) in the asymmetric Gaussian model may significantly degrade the performance of punctured TCs. Hence, in the punctured TC simulations discussed in Section IV, we used the accurate $I_{\rm opt}$ for each of the three channel models.

IV. Simulation Results

In this section, we use simulations to show the result of using each of the three models on the TC decoder performance in optical-fiber channels. We assume chi-square distributed noise at the output of the optical-fiber transmission system, including the receiver, and compare the performance of the TC decoders based on different channel models.

We use a (31, 27, 400) parallel-concatenated-convolutional TC with the encoder and decoder structure as depicted in Fig. 3. The (31, 27, 400) TC is a rate 1/3 code, where the first two parameters, 31 and 27, are octal numbers representing the structure of the constituent encoders. If we transform the octal numbers 31 and 27 into binary numbers 11 001 and 10111, then the digits of the binary numbers represent the coefficients of the parity-check generator polynomials $1+D+D^4$ and $1+D^2+D^3+D^4$. As depicted in Fig. 3(a), "31/27" corresponds to the recursive parity-check generator polynomial $(1 + D + D^4)/(1 + D^2 + D^3 + D^4)$.

A 400-bit interleaver is used between the two constituent encoders shown in Fig. 3(a). The major motivations for using an interleaver are [15]: 1) to generate a long block code from small memory length convolutional codes and 2) to decorrelate the two parity check sequences so that an iterative suboptimal decoding algorithm based on information exchange between the two constituent decoders can be applied.

In the turbo encoder, for each input original information bit u_k, there are two parity check bits, x_k^{1p} and x_k^{2p}, generated by the two parallel concatenated convolutional encoders, respectively. Thus, we have a code rate of 1/3. To achieve higher code rates, a puncturer can be added at the output of the turbo encoder. The puncturing operation can be represented by a puncturing matrix,

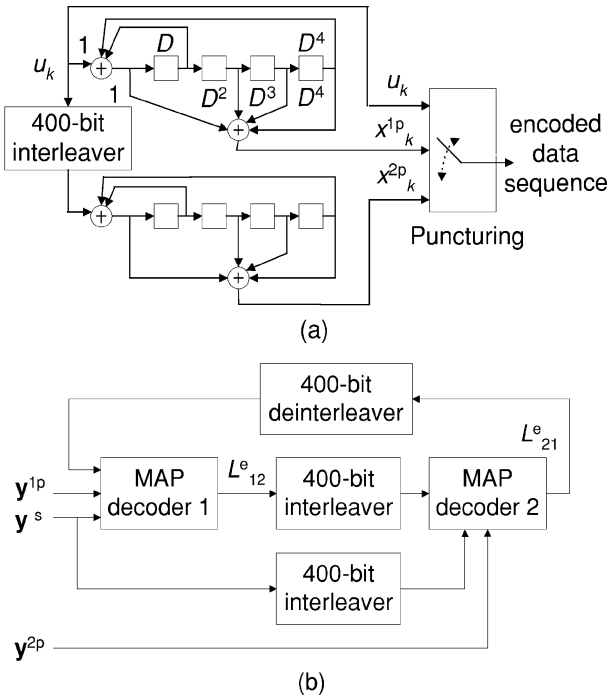

Fig. 3. (31, 27, 400) Turbo code: (a) encoder and (b) decoder structure.

in which each column represents an output block with the element in the first row corresponding to the information bit and the other elements corresponding to the parity check bits. A "0" element in the puncturing matrix means that the corresponding information bit or parity check bit is deleted according to the puncturing mechanism. Similarly, a "1" means that the corresponding bit is transmitted. The puncturing matrices for the rate 1/2 and rate 3/4 punctured TCs are shown in (4.1) and (4.2), respectively

$$\text{Puncturing Matrix (rate 1/3 to rate 1/2)} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \tag{4.1}$$

$$\text{Puncturing Matrix (rate 1/3 to rate 3/4)} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{4.2}$$

As shown in Fig. 3(b), the iterative turbo decoder consists of two serially concatenated constituent decoders, between which there is a 400-bit interleaver identical to the one in the turbo encoder of Fig. 3(a). The first decoder uses MAP decoding on the received information sequence $\mathbf{y}^s$ and parity check sequence $\mathbf{y}^{1p}$ generated by the first encoder, and passes the soft extrinsic information L_{12}^e to the second MAP decoder via the interleaver. Then, the second decoder uses MAP decoding on the interleaved information sequence and the parity check sequence $\mathbf{y}^{2p}$ generated by the second encoder, with an improved estimate of the *a priori* probabilities of the information sequence. The soft extrinsic information L_{21}^e produced by the second MAP decoder is then transferred to the first decoder as improved *a priori* knowledge of the information sequence. Thus, iterative MAP decoding is realized via the information exchange between the two constituent MAP decoders.

We simulate the performance of the TC with BCJR (MAP) decoding algorithms designed based on the chi-square, asymmetric Gaussian, and AWGN models of the optical-fiber channel. In the simulations, the chi-square distributed ASE noise is added to the

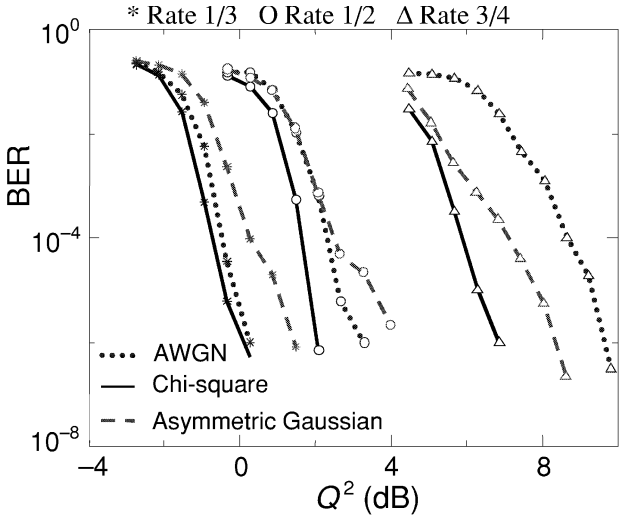

Fig. 4. Output BER of the turbo code (31, 27, 400) decoder based on the chi-square (solid), asymmetric Gaussian (dashed), and AWGN (dotted) models of the ASE noise. The rate 1/2 (circles) and rate 3/4 (triangles) codes are punctured versions of the rate 1/3 (stars) turbo code.

optical-fiber transmission line. We repeat the simulations for different code rates by puncturing the 1/3 turbo code.

Fig. 4 plots the decoded BER with TCs based on different channel models as a function of the Q factor. In all the simulations, the Q factor is evaluated based on the encoded data sequence instead of the original uncoded data sequence, i.e., the penalty caused by lower code rate is not taken into account in the evaluations (which turns the performance curves of the l/3 TCs into a region with negative values of Q^2 in decibels). We only use Fig. 4 to compare the relative performance, at a given code rate, of the TC decoders that are based on different channel models, which is the major concern in this paper. The results show that the TC decoder based on the chi-square model always performs better than the decoders based on the Gaussian approximations. For the rate 3/4 punctured TC, the chi-square model, when used in the BCJR decoding, provides about 1.5- and 3-dB coding gain over the asymmetric Gaussian and AWGN models, respectively, at BER around 10^{-6}.

Comparing the two Gaussian models, we see that the performance of the decoder based on the AWGN model is better at rate 1/3 (low Q), but worse at rate 3/4 (relatively high Q) than that of the asymmetric Gaussian model. This observation agrees with the predictions made in Section II about the accuracy of the two Gaussian approximations for chi-square distributions. Compared to the chi-square model, the AWGN model results in similar decoder performance for the regular rate 1/3 TC where the operating Q-factor is around 0 dB. This result is also consistent with the LLR comparison of the two models shown in Fig. 2(b).

V. Conclusion

In this paper, we discussed the effects of three different noise models: the chi-square, asymmetric Gaussian, and AWGN, on the performance of a TC decoder. We compared the three channel models from two different points of view, the pdf and the LLR of received signals. We showed that the asymmetric Gaussian model is better than the AWGN model in approximating chi-square pdfs, but is not always better than the AWGN model in approximating the LLR of chi-square distributions. We applied the BCJR algorithm to non-AWGN noise distributions. We simulated the performance of TC decoders assuming the chi-square, asymmetric Gaussian, and AWGN distributions, respectively, for a channel that actually has a chi-square noise distribution. We showed that the Gaussian approximations of the chi-square noise distribution might significantly degrade the TC decoder performance. Specifically, the performance degradation for the punctured rate 3/4 TC can be more than 2 dB in Q^2 at 10^{-6} BER. We also showed that the decoder using the AWGN model outperforms the one using the asymmetric Gaussian model at Q^2 around 0 dB, but the latter one outperforms the former one at Q^2 around 7 dB.

Based on these results, we conclude that using accurate channel noise statistics in the iterative MAP decoding algorithm is critical to achieve the expected coding gain from a turbo code. Therefore, we should take into account the accuracy of channel models in soft-decision FEC system designs. To achieve the best possible code performance, accurate noise statistics is required. On the other hand, if accurate statistical model cannot be obtained in a practical implementation, we need to consider a design margin for an approximated or assumed noise statistics.

In Section II of this paper, we proposed an intuitive way to measure how close the Gaussian models approximate the chi-square model in terms of the LLR of received signals. A more complete and accurate measure of LLR approximation should be developed in future research. Moreover, more accurate channel models taking into account signal distortion during transmission and the effects of a realistic low pass filter in the receiver [12], [13] should be incorporated in FEC code studies.

References

[1] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 284–287, Mar. 1974.

[2] A. Puc, F. Kerfoot, A. Simons, and D. L. Wilson, "Concatenated FEC experiment over 5000 km long straight line WDM test bed," in *OFC/IOOC'99 Tech. Dig.*, San Diego, CA, Feb. 1999, pp. ThQ6-1–THQ6-3.

[3] H. Kidorf, N. Ramanujam, I. Hayee, M. Nissov, J. Cai, B. Pedersen, A. Puc, and C. Rivers, "Performance improvement in high capacity, ultra-long distance, WDM systems using forward error correction codes," in *OFC/IOOC'00 Tech. Dig.*, Baltimore, MD, Mar. 2000, pp. ThS3-1–ThS3-3.

[4] O. Ait Sab and V. Lemaire, "Block turbo code performances for long-haul DWDM optical transmission systems," in *OFC/IOOC'00 Tech. Dig.*, Baltimore, MD, Mar. 2000, pp. ThS5-1–ThS5-3.

[5] O. Ait Sab, "FEC techniques in submarine transmission systems," in *OFC/IOOC'01 Tech. Dig.*, Anaheim, CA, Mar. 2001, pp. TuF1-1–TuF1-3.

[6] D. Marcuse, "Derivation of analytical expressions for the bit-error probability in lightwave systems with optical amplifiers," *J. Lightwave Technol.*, vol. 8, pp. 1816–1823, Dec. 1990.

[7] P. A. Humblet and M. Azizoglu, "On the bit error rate of lightwave systems with optical amplifiers," *J. Lightwave Technol.*, vol. 9, pp. 1576–1582, Nov. 1991.

[8] N. S. Bergano, F. W. Kerfoot, and C. R. Davidson, "Margin measurements in optical amplifier systems," *IEEE Photon. Technol. Lett.*, vol. 5, pp. 304–306, Mar. 1993.

[9] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding," in *Proc. IEEE Int. Conf. Communications*, Geneva, Switzerland, May 1993, pp. 1064–1070.

[10] Y. Cai and J. M. Morris, "On performance bounds for linear codes in optical fiber communications systems with asymmetric amplified spontaneous emission noise," in *Proc. Conf. Information Sciences and Systems*, Baltimore, MD, Mar. 2001.

[11] Y. Cai, N. Ramanujam, J. M. Morris, T. Adali, G. Lenner, A. B. Puc, and A. Pilipetskii, "Performance limit of forward error correction codes in optical fiber communications," in *OFC/IOOC'01 Tech. Dig.*, Anaheim, CA, Mar. 2001, pp. TuF2-1–TuF2-3.

[12] R. Holzlöhner, V. S. Grigoryan, C. R. Menyuk, and W. L. Kath, "Accurate calculation of eye diagrams and bit error rates in optical transmission systems using linearization," *J. Lightwave Technol.*, vol. 20, pp. 389–400, Mar. 2002.
[13] P. J. Winzer, M. Pfennigbauer, M. M. Strasser, and W. R. Leeb, "Optimum filter bandwidths for optically preamplified RZ and NRZ receivers," *J. Lightwave Technol.*, vol. 19, pp. 1263–1273, Nov. 2001.
[14] W. E. Ryan. A turbo code tutorial. [Online]. Available: http://www.ece.arizona.edu/~ryan/turbo2c.pdf.
[15] B. Vucetic and J. Yuan, *Turbo Codes Principles and Applications*. Norwell, MA: Kluwer, 2000.
[16] Y. Cai, J. M. Morris, T. Adalı, and C. R. Menyuk, "On the effects of ASE noise models on turbo code decoder performance in optical fiber transmissions," in *CLEO/QELS' 01 Tech. Digest*, Baltimore, MD, May 2001, paper CThO5.
[17] Y. Cai, "Forward error correction codes and line-coding schemes in optical fiber communications," Ph.D. dissertation, Univ. of Maryland Baltimore County, Baltimore, May 2001.

Yi Cai (S'98–M'01) was born on May 12, 1970. He received the B.S. degree in optical engineering from the Beijing Institute of Technology, Beijing, China, in 1992, the M.S. degree in signal and information processing from the Shanghai Institute of Technical Physics, Chinese Academy of Sciences, Shanghai, China, in 1998, and the Ph.D. degree in electrical engineering from the University of Maryland Baltimore County, in 2001.

He was a Development Engineer at Central China Display Labs, Zhengzhou, China, from 1992 to 1995. He is currently a Senior Member of Technical Staff at Tyco Telecommunications Laboratories, Eatontown, NJ. His research interests include forward-error correction, line coding, adaptive signal processing, and their applications in optical fiber communications.

Joel M. Morris (S'64–M'66–SM'91) received the B.S. degree from Howard University, Washington, DC, the M.S. degree from Polytechnic Institute of Brooklyn, Brooklyn, NY, and the Ph.D. degree from The Johns Hopkins University, Baltimore, MD, all in electrical engineering.

He began his professional career in 1966 at Bell Telephone Laboratories, Holmdel, NJ, as a Member of the Technical Staff. Prior to his doctoral studies, he also worked for the Bendix Environmental Sciences Division, Baltimore, MD, as an Electrical Engineer. Since then, he has worked for the Naval Research Laboratory, Communication Sciences Division, Washington, DC, as an Electrical Engineer; the Office of Naval Research, Electronic and Solid State Sciences Program, Arlington, VA, as a Scientific Officer; and the JHU/Applied Physics Laboratory, in both the Submarine Technology Division and Fleet Systems Department, Laurel, MD, as a Senior Staff Engineer. He has also taught part time for many years at Howard University and The Johns Hopkins University Evening College. From 1983 to 1988, he was a Professor in the Electrical Engineering Department of Howard University. From 1988 to 1995, he was with the Electrical Engineering Department, University of Maryland, Baltimore County (UMBC) Campus, developing and contributing to their graduate communications and signal processing program and serving as Chairman during the 1994-1995 academic year. From 1995 to 1997, he served as Chairman of the combined Computer Science and Electrical Engineering Department at UMBC. He is currently a Professor of Electrical Engineering at UMBC. His current research interests are communications, error-control coding, detection, and estimation, and joint-domain representations and techniques, all with application to communication and general signal processing problems. His recent efforts have focused on error-control coding techniques for wireless, wireline, and optical fiber communications. He has served on a number of Federal Government program review and selection panels, and has consulted for or collaborated with the JHU/Applied Physics Laboratory, Atlantic Aerospace Electronics Corporation, BayNetworks, Inc., the Department of Defense, Lawrence Livermore National Laboratory, Nortel Networks, Inc., and Science and Technology Corporation, in the areas of communications, signal processing, and detectability studies.

Dr. Morris is currently the Faculty Advisor to the IEEE Student Branch at UMBC, and a member of the IEEE Communication Society of Baltimore. He is a member of Tau Beta Pi, Sigma Xi, SPIE, and the U.S. National Committee of the International Union of Radio Science (URSI/USNC), where he served as Commission E Chairman from 1984 to 1987.

Tülay Adalı (S'87–M'92–SM'98) received the B.S. degree from Middle East Technical University, Ankara, Turkey, in 1987 and the M.S. and Ph.D. degrees from North Carolina State University, Raleigh, in 1988 and 1992, respectively, all in electrical engineering.

In 1992, she joined the Department of Electrical Engineering at the University of Maryland Baltimore County, Baltimore, where she is currently an Associate Professor. Her research interests are in the areas of adaptive signal processing, estimation theory, neural computation, and their applications in optical communications, biomedical image analysis, channel equalization, and time-series prediction.

Dr. Adalı is an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING and the *Journal of VLSI Signal Processing Systems*. She has worked in the organization of a number of international conference and workshops, including the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) and the IEEE International Workshop on Neural Networks for Signal Processing (NNSP). She has been the General Co-Chair for the NNSP workshops during 2001–2003. She is currently the Chair for the IEEE Neural Networks for Signal Processing Technical Committee and is serving on the IEEE Signal Processing Society conference board. She is the recipient of a 1997 National Science Foundation CAREER Award and the provost's research faculty fellowship for the 1997–1998 academic year.

Curtis R. Menyuk (SM'88–F'98) was born on March 26, 1954. He received the B.S. and M.S. degrees from Massachusetts Institute of Technology, Cambridge, in 1976 and the Ph.D. degree from the University of California at Los Angeles (UCLA) in 1981.

He was a Research Associate at the University of Maryland, College Park, and at Science Applications International Corporation, McLean, VA. In 1986, he became an Associate Professor in the Department of Electrical Engineering at the University of Maryland, Baltimore County (UMBC), and was the founding member of this department. In 1993, he was promoted to Professor. He was on partial leave from UMBC during 1996 to 2002. He is also a former UMBC Presidential Research Professor. From 1996 to 2001, he worked part time for the Department of Defense (DoD), co-directing the Optical Networking Program at the DoD Laboratory for Telecommunications Sciences, Adelphi, MD, from 1999 to 2001. During 2001–2002, he was Chief Scientist at PhotonEx Corporation. For the last 15 years, his primary research area has been theoretical and computational studies of fiber-optic communications. He has authored or co-authored more than 160 archival journal publications, as well as numerous other publications and presentations. He has also edited two books. The equations and algorithms that he and his research group at UMBC have developed to model optical-fiber transmission systems are used extensively in the telecommunications industry.

Dr. Menyuk is a member of the Society for Industrial and Applied Mathematics and the American Physical Society. He is a Fellow of the Optical Society of America (OSA).