Accurate Calculation of Eye Diagrams and Bit Error Rates in Optical Transmission Systems Using Linearization

Ronald Holzlöhner, *Senior Member, IEEE, Member, OSA*, V. S. Grigoryan, *Member, OSA*, C. R. Menyuk, *Fellow, IEEE, Member, OSA*, and W. L. Kath, *Member, OSA*

***Abstract*—We present a novel linearization method to calculate accurate eye diagrams and bit error rates (BERs) for arbitrary optical transmission systems and apply it to a dispersion-managed soliton (DMS) system. In this approach, we calculate the full nonlinear evolution using Monte Carlo methods. However, we analyze the data at the receiver assuming that the nonlinear interaction of the noise with itself in an appropriate basis set is negligible during transmission. Noise–noise beating due to the quadratic nonlinearity in the receiver is kept. We apply this approach to a highly nonlinear DMS system, which is a stringent test of our approach. In this case, we cannot simply use a Fourier basis to linearize, but we must first separate the phase and timing jitters. Once that is done, the remaining Fourier amplitudes of the noise obey a multivariate Gaussian distribution, the timing jitter is Gaussian distributed, and the phase jitter obeys a Jacobi-Θ distribution, which is the periodic analogue of a Gaussian distribution. We have carefully validated the linearization assumption through extensive Monte Carlo simulations. Once the effect of timing jitter is restored at the receiver, we calculate complete eye diagrams and the probability density functions for the marks and spaces. This new method is far more accurate than the currently accepted approach of simply fitting Gaussian curves to the distributions of the marks and spaces. In addition, we present a deterministic solution alternative to the Monte Carlo method.**

***Index Terms*—Amplifier noise, error analysis, Karhunen–Loève transforms, linear approximation, Monte Carlo methods, nonlinearities, optical fiber dispersion, optical fiber theory, simulation.**

I. Introduction

In current optical fiber communications systems, amplifier spontaneous emission (ASE) noise sets the lower limit on the allowed system power [1], [2]. At the receiver, ASE noise causes intensity fluctuations and timing jitter, leading to a deterioration of the bit error rate (BER). The ASE noise spectrum may be considered white at the point that it is contributed [3] because it is always much broader than any channel's spectrum. However, the fiber's Kerr nonlinearity leads to a complex interaction between the signal and the noise, so that the noise spectrum does not remain white during the fiber propagation.

The traditional method of computing the probability distribution function (pdf) of the electrical signal in the receiver, which is based on Monte Carlo simulations, only works for a limited range of BERs, beyond which the BER must be extrapolated [4]. Standard extrapolation methods to date assume that the noise power after narrow-band filtering is Gaussian distributed. This assumption often yields good agreement between simulations and experiments [5], but it is not always reliable. To make progress with the difficult problem of calculating BERs, almost all previous work, aside from Monte Carlo simulations, is based on the linearization assumption. This assumption is that the nonlinear beating of the noise with itself during the fiber transmission can be neglected. We note that is it not generally assumed that nonlinear interactions between the signal and the noise can be ignored. Also, the nonlinear interaction of the noise with itself in the receiver due to the square-law detection is usually kept. An early application of this approach was to use soliton systems. The BER in standard soliton systems is dominated by timing jitter. Gordon and Haus [1] and Haus [2] used linearization to calculate the timing jitter and, hence, the BER. Under the assumption that the amplitude of the marks is constant, that the amplitude of the spaces is strictly zero, that the noise power is white when it arrives at the receiver, that an ideal optical bandpass filter is used right before the receiver, and that an ideal integrate-and-dump circuit follows after a square-law detector in the receiver, Marcuse [6] and Humblet and Azizoğlu [7] showed that the pdf of the power in the marks obeys a noncentral chi-square distribution, whereas the pdf of the spaces obeys a central chi-square distribution. From these pdfs, it is possible to calculate the optimal decision threshold and the BER at any threshold. These assumptions correspond with the assumption that the transmitted signal consists of ideal nonreturn-to-zero (NRZ) pulses, neglecting fiber nonlinearity, and assuming that the net dispersion in the fiber is zero. Because nonlinearity during the transmission is completely neglected, the linearization assumption holds. Despite the restrictive assumptions, this work is highly significant. The central chi-square distribution has an exponential tail and, hence, falls off much more slowly than a Gaussian, indicating that the Gaussian approximation fails even in this highly idealized setting. Similar results can be obtained for any system in which the transmission nonlinearity is neglected. Later work by Hui *et al.* [8], [9] and Carena *et al.* [10] used the linearization assumption, taking into account both the fiber nonlinearity and dispersion, but neglecting the data modulation when calculating the noise power. In effect, they assumed a continuous wave (CW) signal. They found that the power is not uniformly distributed among all of

Manuscript received June 25, 2001; revised November 29, 2001. This work was supported by the Air Force Office of Scientific Research, the Department of Energy, and the National Science Foundation.

The authors are with the Department of Computer Science and Electrical Engineering, University of Maryland Baltimore County, Technological Research Center, Baltimore, MD 21250 USA (e-mail: holzloehner@umbc.edu; web:http://www.photonics.umbc.edu/).

Publisher Item Identifier S 0733-8724(02)02202-8.

the spectral components due to the significant four-wave mixing between the signal and the noise. Later work has extended this result by applying this approach to WDM systems [11] and by adding the noise calculated assuming a CW signal to modulated data [12]–[14].

All of the work just cited used relatively simple analytical forms for the signal when calculating the noise power. In general, however, the signal evolution in realistic optical communications systems is quite complex due to the combined effects of the Kerr nonlinearity and the dispersion in optical fibers. The use of simple analytical approximations for the signal when calculating the noise may not yield sufficient accuracy. Recently, Grigoryan *et al.* [15] showed that it is possible to linearize around a computationally determined signal. They used this approach to calculate the timing and amplitude jitter for NRZ, return-to-zero (RZ), and dispersion-managed soliton (DMS) pulses. They validated their results with comparison to Monte Carlo simulations in all cases and to experiments in the case of DMS pulses. Calculation of the timing and amplitude jitter is, however, not sufficient, in most cases, to determine the BER [16].

In this work, we use the linearization assumption to calculate the full pdf for the marks and spaces in the experimental DMS system described by Mu *et al.* [5], taking into account the full nonlinear and dispersive evolution of the signal and its interaction with noise. We then calculate the optimal decision threshold and the BER as a function of the decision threshold. On the one hand, it is possible to view our work as an extension of previous work on linearization. On the other hand—and perhaps more usefully—it can be viewed as an extension of the Monte Carlo approach. As we noted previously, it is currently the standard practice to simply fit a Gaussian distribution to the rails of a numerically determined eye diagram when calculating the BER. With no additional computational effort, our approach allows the user to take into account the non-Gaussian effects, leading to greatly improved pdfs for the marks and the spaces.

It is possible to criticize the linearization assumption because it is liable to break down just at the point where it is most important to calculate the pdfs—on the tails of the distribution function. Indeed, work by Menyuk [17] and by Georges [18], [19] has shown, in the case of timing jitter for solitons, that the breakdown of the linearization assumption can lead to significant changes in the BER. In response, we first note that there is no experimental evidence that the linearization assumption fails in any practical context, in contrast to the far cruder Gaussian fitting method in common use today, for which there is certainly experimental evidence that it fails [20, Fig. 10.14]. Second, the increased use of forward error correction (FEC) implies that raw error rates in the optical fiber transmission as high as 10^{-3} are often now acceptable, so that systems decreasingly operate far out on the tails of the pdfs where the linearization assumption is expected to be less reliable. Finally, and perhaps most important from a conceptual standpoint, it may be possible to fully account for the nonlinear interaction of the noise with itself by using large deviation techniques [21], such as importance sampling [22], but these techniques *require* the prior solution of the problem using the linearization assumption as a starting point. Applying these techniques in any new situation requires substantial adaptation and is still an unsolved problem in optical fiber communications. From this standpoint, the work presented here is a key step in solving the complete problem of calculating the BERs in optical fiber communications systems.

The DMS system that we study in this paper [5], [16] is simpler than many modern-day communications systems in an important respect. The soliton pulse durations do not change enough to lead to a significant overlap with their neighbors, so that there is no interaction with neighboring bits. Moreover, this system is a single-channel system, so that there is no possibility of interchannel interactions. Because neither intrachannel nor interchannel interactions occur between bits, there are no pattern dependences, and it is sufficient to study the behavior of a single mark and a single space to determine the pdfs for both. Extending our approach to take into account pattern dependences will be a nontrivial undertaking. At the same time, this system is far more nonlinear than are modern-day communications systems and, hence, represents a stringent test of the linearization assumption. In this paper, we will show that the linearization assumption works well and is consistent with experiments [5] over 24 000 km, once phase and timing jitter are properly taken into account. These results bode well for this approach in practical contexts with less nonlinearity.

The utility of the linearization assumption stems from two key mathematical results. The first is the Karhunen–Loève theorem [23], which states that a combination of signal and noise over any finite time can be expanded in an orthonormal basis whose coefficients are independent random variables. When the noise is white, any orthonormal basis will satisfy the Karhunen–Loève theorem. In optical fiber communications systems, the ASE noise is effectively white when it is contributed by the amplifiers, but it only remains white for short distances over which the nonlinear interaction between the signal and the noise can be neglected. Over longer distances, the noise becomes correlated, and the Karhunen–Loève basis becomes unique. The second mathematical result is Doob's theorem [24], which states that when the system is linearizable, each of these independent random variables is Gaussian. Thus, it suffices, in principle, to determine the Karhunen–Loève modes, as well as the mean and variance of its coefficients, to calculate the effective noise pdf. We emphasize that this powerful result allows the signal to interact nonlinearly with itself and with the noise; it only requires that the noise not interact with itself. In practice, one must use an approximate static basis from which to compute the Karhunen–Loève modes. The standard Fourier basis naturally suggests itself, and we will use it with an important caveat. In the case of standard solitons, it has long been known that it is not appropriate to use Fourier modes to linearize around a soliton solution because a perturbation can change the soliton parameters, as well as add to the background [2], [25]. For the DMS system that we are considering, we have found that it is necessary to explicitly calculate the phase and timing jitter of the soliton and to separate them from the remainder of the noise calculation. When the linearization assumption holds, the pdf of the timing jitter is Gaussian, and the pdf of the phase jitter is a Jacobi-Θ function, which is the periodic analogue of a Gaussian distribution. We validate this assumption using extensive Monte Carlo simulations. Thus,

we can account for the effect of the timing jitter in the final eye diagram. The phase jitter does not matter in square-law detectors. The entire description given in this paper will apply to one optical polarization only, which is appropriate for the DMS system that we are using as an example [5]. Moreover, this choice somewhat simplifies the theoretical development. There is no reason to doubt that this formalism can be extended to take into account polarization effects.

After separating the phase and timing jitter, using a procedure that we will describe in the body of this paper, we compute the evolution of the covariance matrix $K_{kl} = \langle a_k a_l^* \rangle$, where $a_k = a_{k,R} + i a_{k,I}$ is the amplitude of the kth Fourier mode, and $a_{k,R}$ and $a_{k,I}$ are its real and imaginary components. When the linearization assumption holds, the $a_{k,R}$ and $a_{k,I}$ obey a multivariate Gaussian pdf that can be determined from the covariance matrix. It is possible to compute the covariance matrix in two ways. First, we can estimate it using Monte Carlo simulations. Second, we derive an ordinary differential equation (ODE) for the K_{kl} from which it can be obtained directly. It is the primary purpose in this paper to demonstrate the validity of the linearization assumption in the highly nonlinear DMS system over 24 000 km. The Monte Carlo simulations allow us to use standard statistical methods to both validate the linearization assumption and estimate K_{kl}. Thus, we focus on the Monte Carlo approach here, although we discuss the ODE approach briefly. The ODE approach offers substantial computational advantages, and a more detailed discussion of it will be the subject of a future publication.

The remainder of this paper is organized as follows. We derive the theory of our linearization approach in Section II. In Section III, we apply this theory to our DMS system. We compute the phase and timing jitter and the covariance matrix $K_{kl}(z)$ from Monte Carlo simulations. We then calculate the pdfs for the marks and spaces. In Section IV, we compare these results to a direct numerical solution of the ODE for the $K_{kl}(z)$. Section V contains the conclusions.

II. Theory

A. Linearization Approach

In our simulation model, we consider the normalized nonlinear Schrödinger equation with a Langevin noise term

$$i\frac{\partial u}{\partial z} + \frac{1}{2}D(z)\frac{\partial^2 u}{\partial t^2} + |u|^2 u = ig(z)u + \hat{F}(z,t). \quad (1)$$

Here, the pulse envelope is normalized as $u = E\sqrt{\gamma L_D}$, where E is the electric field envelope, $\gamma = n_2\omega_0/(A_{\text{eff}}c)$ is the nonlinear coefficient, and L_D is a characteristic length. The quantity $n_2 = 2.6 \times 10^{-16}$ cm^2/W is the Kerr coefficient, $\omega_0 = 2\pi \times 193.23$ THz is the central angular frequency, $A_{\text{eff}} = 47\ \mu$m^2 is the effective fiber core area, and c is the speed of light. The characteristic dispersion length L_D equals $T_0^2/|\beta_0''|$, where T_0 is the characteristic time scale and β_0'' is the scaling dispersion. We choose $T_0 = 5.67$ ps, which roughly equals the root mean square soliton duration, and $\beta_0'' = -0.1$ ps^2/km, which approximately equals the path average dispersion. The distance z is normalized as $z = Z/L_D$, where Z is physical distance. The retarded time is normalized as $t = (t_p - \beta_0' Z)/T_0$, where t_p is the physical time and β_0' is the inverse group velocity. The normalized gain coefficient $g(z)$ is

$$g(z) = \begin{cases} g_m(z), & z_m < z < z_m + \frac{L_{\text{amp}}}{L_D} \\ -\gamma_l, & \text{elsewhere} \end{cases} \quad (2)$$

where g_m represents the normalized gain coefficient inside the mth amplifier, which we assume to begin at $z = z_m$ and to be of length L_{amp}, and γ_l is the normalized fiber loss coefficient. The quantity $\hat{F}$ represents the ASE white noise contribution with zero mean $\langle \hat{F}(z,t) \rangle = 0$ and autocorrelation

$$\left\langle \hat{F}(z,t)\hat{F}^*(z',t') \right\rangle = 2\eta\delta(z-z')\delta(t-t') \quad (3)$$

where $\eta = n_{\text{sp}} g_m \hbar\omega_0 L_D \gamma/T_0$ with Planck's constant $\hbar = 1.0545 \times 10^{-34}$ kg m^2/s, and n_{sp} is the spontaneous emission factor. We assume $n_{\text{sp}} = 1.4$ inside the amplifiers and $n_{\text{sp}} = 0$ in the optical fiber. The angular brackets $\langle \cdot \rangle$ denote the noise ensemble average, and the asterisk denotes complex conjugation. Higher order dispersion, as well as filter terms, are neglected in (1).

Our initial goal is to derive a differential equation that describes the signal-noise beating. We write $u = u_0 + \delta u$ as a sum of a noise-free signal $u_0 = \langle u \rangle$ and accumulated transmitted noise δu. The difference of (1) and the statistical average of (1) then yields the evolution equation for δu

$$i\frac{\partial \delta u}{\partial z} + \frac{D}{2}\frac{\partial^2 \delta u}{\partial t^2} + 2|u_0|^2\delta u + {u_0}^2(\delta u)^* = ig\delta u + \hat{F}. \quad (4)$$

The third and fourth terms on the left-hand side contain the beating of the signal with the noise, whereas the noise–noise beating is quadratic in δu and is omitted. Equation (4) describes the noise growth if $\delta u \ll u_0$ and we neglect any influence of δu on u_0. We can expand u_0 and δu as a Fourier series

$$u_0 = \sum_{n=-N/2}^{N/2-1} A_n(z)\ \exp(i\omega_n t) \quad (5a)$$

$$\delta u = \sum_{n=-N/2}^{N/2-1} a_n(z)\ \exp(i\omega_n t) \quad (5b)$$

where $\omega_n \equiv 2\pi n T_0/T$, and T is the period. After substituting (5a) and (5b) into (4), we find

$$\frac{da_k}{dz} = \left(g - i\frac{D}{2}\omega_k^2\right) a_k + i\sum_{n,l,m=-N/2}^{N/2-1} \left[2A_n A_l{}^* a_m \delta_{n-l,k-m} + A_n A_l a_m^* \delta_{n+l,k+m}\right] - i\Gamma_k(z) \quad (6)$$

where the Γ_k are the Fourier coefficients of the white noise input $\hat{F}$, and δ is Kronecker's delta. The correlation of the Γ_k is $\langle \Gamma_k(z)\Gamma_m^*(z') \rangle = (2\eta T_0/T)\delta(z-z')\delta_{k,m}$, where η is again zero outside of the amplifiers. We define the complex column vectors $\boldsymbol{\alpha} = (a_{-N/2}, \ldots, a_{N/2-1})^T$ and $\boldsymbol{\alpha}^* = (a_{-N/2}^*, \ldots, a_{N/2-1}^*)^T$, as well as $\boldsymbol{\Gamma} = (\Gamma_{-N/2}, \ldots, \Gamma_{N/2-1})^T$, where the

superscript T indicates the transpose operation. Then, we can rewrite (6) in matrix form as

$$\frac{d\boldsymbol{\alpha}}{dz} = \mathsf{B}\boldsymbol{\alpha} + \mathsf{E}\boldsymbol{\alpha}^* - i\boldsymbol{\Gamma} \tag{7}$$

where the complex matrices B and E are defined as[1]

$$B_{km} = \left(g - i\frac{D}{2}\omega_k^2\right)\delta_{km} + 2i\sum_{n,l} A_n A_l^* \delta_{n-l,k-m} \tag{8a}$$

$$E_{km} = i\sum_{n,l} A_n A_l \delta_{n+l,k+m}. \tag{8b}$$

The matrix E is symmetric ($E_{km} = E_{mk}$), whereas the matrix B is anti-Hermitian ($B_{km}^* = -B_{mk}$) if g is zero. The sum in B is circulant and, thus, corresponds to a convolution in the time domain, whereas E can be termed anticirculant. (A matrix M is circulant if there is a vector $\boldsymbol{x}$ with $M_{km} = x_{k-m}$). The number of operations required to evaluate B and E grows like N^3. Equation (6) depends on both a_k and a_k^*, so that the linearized problem becomes non-Hermitian [8], [9]. Our probability space is spanned by the $2N$ real variables $a_{k,R}$ and $a_{k,I}$. It is, therefore, convenient to split (6) into its real and imaginary parts and consider the resulting system of equations. Introducing the real $2N$ vector $\boldsymbol{a} = (\boldsymbol{\alpha}_R, \boldsymbol{\alpha}_I) = (a_{-N/2,R}, \ldots, a_{N/2-1,R}, a_{-N/2,I}, \ldots, a_{N/2-1,I})^T$ as a partitioned vector of the real and imaginary parts of $\boldsymbol{\alpha}$, and similarly $\boldsymbol{w} = (\boldsymbol{\Gamma}_I, -\boldsymbol{\Gamma}_R)^T$, we can rewrite (7) as

$$\begin{aligned} \frac{d\boldsymbol{a}}{dz} &= \mathcal{R}(z)\boldsymbol{a} + \boldsymbol{w}(z) \\ \mathcal{R} &= \begin{bmatrix} \mathsf{B}_R + \mathsf{E}_R & -\mathsf{B}_I + \mathsf{E}_I \\ \mathsf{B}_I + \mathsf{E}_I & \mathsf{B}_R - \mathsf{E}_R \end{bmatrix} \end{aligned} \tag{9}$$

where $\mathcal{R}$ is a real $2N \times 2N$ block matrix and we have used the notation $\mathsf{B} = \mathsf{B}_R + i\mathsf{B}_I$ and $\mathsf{E} = \mathsf{E}_R + i\mathsf{E}_I$. We may formally write the solution to (9) as [26]

$$\boldsymbol{a}(z) = \Psi(z, z_0)\boldsymbol{a}(z_0) + \int_{z_0}^{z} \Psi(z, z')\boldsymbol{w}(z')dz' \tag{10}$$

where $\Psi(z, \zeta)$ is a propagator matrix that obeys the following:

$$\frac{d}{dz}\Psi(z, \zeta) = \mathcal{R}(z)\Psi(z, \zeta) \qquad \Psi(\zeta, \zeta) = \mathcal{I} \quad \forall\, z, \zeta \tag{11}$$

where $\mathcal{I}$ is the identity matrix. Equation (9) describes the spatial evolution of the noise Fourier modes. Neglecting, for now, the necessity of separating the phase and timing jitters, the linearization assumption, along with Doob's theorem [24], implies that the $\boldsymbol{a}(z)$ satisfy a multivariate Gaussian distribution, which is completely described by its first two moments. The mean of $\boldsymbol{a}(z)$ is, by definition, zero, whereas the second moments are given by the covariance matrix $\mathcal{K}$. Hence, its pdf may be written as [27]

$$f_{\boldsymbol{a}}(\boldsymbol{a}, z) = (2\pi)^{-N}\sqrt{\det \mathcal{K}^{-1}(z)}\exp\left[-\frac{1}{2}\boldsymbol{a}^T\mathcal{K}^{-1}(z)\boldsymbol{a}\right] \tag{12}$$

[1]In the rest of this paper, we will use the sans serif font to denote complex $N \times N$ matrices like B, the script font for real $2N \times 2N$ matrices like $\mathcal{R}$, and the bold font for real $2N$ vectors such as $\boldsymbol{a}$. The only exception will be $\boldsymbol{\alpha}$, which is a complex N vector.

with the real symmetric $2N \times 2N$ covariance matrix $\mathcal{K}$

$$\begin{aligned} \mathcal{K} &= \left\langle (\boldsymbol{a} - \langle \boldsymbol{a} \rangle)(\boldsymbol{a}^T - \langle \boldsymbol{a} \rangle^T) \right\rangle \\ &= \left\langle \begin{bmatrix} \boldsymbol{\alpha}_R\boldsymbol{\alpha}_R^T & \boldsymbol{\alpha}_R\boldsymbol{\alpha}_I^T \\ \boldsymbol{\alpha}_I\boldsymbol{\alpha}_R^T & \boldsymbol{\alpha}_I\boldsymbol{\alpha}_I^T \end{bmatrix} \right\rangle \end{aligned} \tag{13}$$

where all vector products above are outer products. This definition of $\mathcal{K}$ embodies the full covariance information in $4N^2$ real numbers of which $N(2N+1)$ are independent, whereas the complex $N \times N$ matrix $\langle a_k a_l^* \rangle = \langle \boldsymbol{\alpha}\boldsymbol{\alpha}^\dagger \rangle_{kl}$, where the $\dagger$ denotes the conjugate transpose, contains only $2N^2$ real numbers of which N^2 are independent, and, thus, $\langle a_k a_l^* \rangle$ lacks some information. From (9) and (10), we now find that $\mathcal{K}$ evolves over distance according to

$$\frac{d}{dz}\mathcal{K}(z) = \mathcal{R}\mathcal{K} + \mathcal{K}\mathcal{R}^T + \frac{\eta I_0}{T}\mathcal{I} \tag{14}$$

where η and T are defined after (3) and (5b), respectively. Equation (14) is a Lyapunov equation [26] and the fundamental linear evolution equation of the covariance matrix that describes the statistical dependence of the a_k. The right-hand side of (14) is symmetric because $(\mathcal{R}\mathcal{K})^T = \mathcal{K}^T\mathcal{R}^T = \mathcal{K}\mathcal{R}^T$, so that $\mathcal{K}$ remains symmetric as it evolves over z. Initially, $\mathcal{K}$ is zero because the launched signal is noise free. The matrix $\mathcal{R}$ is distance dependent and includes amplification–attenuation as well as the beating of the signal with the noise. The last term describes the white noise input and is only nonzero inside the optical amplifiers. Newly added noise only contributes to the diagonal elements K_{kk}. In addition to being symmetric, $\mathcal{K}$ is also positive definite, so that its determinant is positive.

We note that the direct derivation of (14) from (9) is only one possible way of determining the evolution; one can show that the pdf in (12), where the covariance matrix $\mathcal{K}(z)$ is described by (14), represents the exact solution of the Fokker–Planck equation corresponding to the Langevin equation (9). Yet another approach to derive (14) uses Itô's method [28]. All of these methods are, of course, equivalent.

B. Separation of Phase and Timing Jitters

In order to properly account for the phase and timing jitters, we must treat them separately. We do so by removing their contribution to the covariance matrix, keeping track of their magnitude. If these contributions are not treated separately, they distort the distribution functions so that the Fourier magnitudes $a_{k,R}$ and $a_{k,I}$ are no longer Gaussian distributed. The phase jitter makes no contribution to the electrical eye diagrams and, hence, no contribution to the BER in systems with a square-law detector like the one that we are considering. Hence, once the phase jitter is separated, it can be ignored. By contrast, the timing jitter does affect the eye diagrams and the BER and must be taken into account explicitly when calculating these quantities. In this subsection, we will show how to separate the phase and timing jitters from the calculation of the covariance matrix. In Section II-C, we will show how to calculate eye diagrams while accounting for the timing jitter.

The method that we use to separate the phase and timing jitters differs, depending on whether we are calculating the covariance matrix by using Monte Carlo simulations or by directly solving (14).

For the Monte Carlo simulations, we focus on the received signal $u(t) = u_0(t) + \delta u(t)$, where $u_0 = \langle u \rangle$ is the signal average over all noise realizations, and δu is one particular realization of the accumulated, propagated noise at the receiver. The Fourier expansion of u is $u(t) = \sum_{k=-N/2}^{N/2-1} B_k \exp(i\omega_k t)$, where $B_k = A_k + a_k$ and $\omega_k = 2\pi k T_0/T$, conforming to (5a) and (5b). For single-pulse transmission, we apply the nonlinear transformation

$$\begin{aligned} A_k + a_k =& B_k = \bar{B}_k \exp[i(\varphi + \omega_k \tau)] \\ =& (\bar{A}_k + r_k) \exp[i(\varphi + \omega_k \tau)] \end{aligned} \tag{15}$$

where $\bar{A}_k = \langle \bar{B}_k \rangle$ and $\langle r_k \rangle = 0$. For each noise realization, we determine φ and τ by fitting the linear function $\alpha + \beta\omega_k$ to the phase of the B_k using the least-squares criterion

$$H = \min_{\alpha,\beta} \sum_{k=-N/2}^{N/2-1} |B_k|^2 \left[\arctan \frac{B_{k,I}}{B_{k,R}} - \alpha - \beta\omega_k \right]^2 \tag{16}$$

where $B_k = B_{k,R} + iB_{k,I}$, and then setting $\varphi = \alpha$ and $\tau = \beta$. We have found that the linear phase assumption of (16) is good as long as the receiver is placed at the chirp-free maximum pulse compression point of the dispersion map. If the phase and timing jitter are small, $A = \langle B \rangle \approx \langle \bar{B} \rangle = \bar{A}$, and we can decompose the total noise in the new basis, $a_k \exp[-i(\varphi + \omega_k \tau)]$, as

$$\begin{aligned} a_k \exp[-i(\varphi + \omega_k \tau)] =& \bar{A}_k - A_k \exp[-i(\varphi + \omega_k \tau)] + r_k \\ \approx& \bar{A}_k \{1 - \exp[-i(\varphi + \omega_k \tau)]\} + r_k \\ \approx& \varphi c_k + \tau d_k + r_k \end{aligned} \tag{17}$$

where $c_k = i\bar{A}_k$ and $d_k = i\omega_k \bar{A}_k$. The term φc_k is responsible for a phase shift $u(t) = u_0(t)\exp(i\varphi)$, whereas the component τd_k produces the time shift $u(t) = u_0(t + \tau)$ and, thus, leads to timing jitter. The residual noise r_k is orthogonal to the c_k and d_k with the scalar product $(1/2)\sum_k (v_k r_k^* + r_k^* v_k) = 0$, where $v_k = c_k$ or $v_k = d_k$. The vectors c_k and d_k will only be orthogonal to each other if the signal is an even function in time. Our simulation shows that $\langle \varphi \rangle = \langle \tau \rangle = 0$.

We will show in the Appendix that (16) and (17) are consistent for arbitrary pulse shapes. Note that even though c_k and r_k are orthogonal, the quantities φ and r_k are not statistically independent because φ depends on the pulse power due to the nonlinear phase rotation. A noise realization in which the noise increases the pulse power will tend to have both large $|r_k|$ and large φ, leading to a correlation. On the other hand, our simulations indicate that the correlation between τ and r_k is negligible.

We compute the pdfs of φ and τ by averaging over all Monte Carlo noise realizations. With the real partitioned $2N$ vector $\boldsymbol{r} = (r_{-N/2,R}, \ldots, r_{N/2-1,R}, r_{-N/2,I}, \ldots, r_{N/2-1,I})^T$ and, analogously, $\overline{\boldsymbol{B}} = (\bar{B}_{-N/2,R}, \ldots)^T$, we can define a reduced covariance matrix $\mathcal{K}^{(r)}$ as

$$\mathcal{K}^{(r)} = \left\langle \overline{\boldsymbol{B}} \overline{\boldsymbol{B}}^T \right\rangle - \langle \overline{\boldsymbol{B}} \rangle \langle \overline{\boldsymbol{B}} \rangle^T = \langle \boldsymbol{r} \boldsymbol{r}^T \rangle. \tag{18}$$

We will use (18) in Section III-B. The quantity $\boldsymbol{r}$ obeys a multivariate Gaussian distribution, and we replace $\boldsymbol{a}$ by $\boldsymbol{r}$ and $\mathcal{K}$ by $\mathcal{K}^{(r)}$ in (12).

C. Derivation of the Eye Diagram

In this section, we derive the pdf of the filtered output current of a square-law detector. A similar pdf has already been derived by Lee *et al.* [29], Bosco *et al.* [30], and Forestieri [31]. We use this pdf to compute an electrical eye diagram that is not produced, as is traditional in simulations, by overlaying a finite number of traces of "1"s and "0"s with different noise realizations, but displays the continuous probability densities. The inputs we need are the pdf of the timing jitter f_τ, the Fourier modes of the transformed noise-free signal $\bar{A}_k$, and the reduced covariance matrix $\mathcal{K}^{(r)}$.

The receiver first converts the input signal plus noise to an electrical current $I(t)$ in a photodetector. We assume that the photodetector is an ideal square-law detector with $I \propto |u|^2$. We will apply the transformation (15) and, hence, start by describing $I(t)$ in the absence of timing jitter. Using (17), we now obtain

$$\begin{aligned} I(t) =& \kappa |u_0(t) + \delta u(t)|^2 \\ =& \kappa \sum_{k,l=-N/2}^{N/2-1} \left(\bar{A}_k + r_k\right)^* \left(\bar{A}_l + r_l\right) \exp[it(\omega_l - \omega_k)] \end{aligned} \tag{19}$$

where κ is the receiver responsivity and the r_k are the residual noise coefficients. The electrical current $I(t)$ then passes through a low-pass electrical filter. The filter operation can be written in Fourier space as a multiplication with a complex self-adjoint filter matrix F. Hence, we can write the filtered current $y(t)$ as

$$y(t) = \sum_{k,l=-N/2}^{N/2-1} (\bar{A}_k + r_k)^* W_{kl}(t) (\bar{A}_l + r_l) \tag{20}$$

with $W_{kl}(t) = \kappa \exp\left[it(\omega_l - \omega_k)\right] F_{kl}$, so that W is a complex self-adjoint $N \times N$ matrix. We introduce the partitioned vector $\overline{\boldsymbol{A}} = (\bar{A}_{-N/2,R}, \ldots, \bar{A}_{N/2-1,R}, \bar{A}_{-N/2,I}, \ldots, \bar{A}_{N/2-1,I})^T$ to rewrite $y(t)$ as

$$\begin{aligned} y(t) =& (\overline{\boldsymbol{A}} + \boldsymbol{r})^T \mathcal{W} (\overline{\boldsymbol{A}} + \boldsymbol{r}) \\ \mathcal{W} =& \begin{bmatrix} \mathsf{W}_R & -\mathsf{W}_I \\ \mathsf{W}_I & \mathsf{W}_R \end{bmatrix} = \mathcal{W}^T. \end{aligned} \tag{21}$$

The minus sign in the last line appears because W_I is antisymmetric. The right-hand side of (21) is a symmetric bilinear form, but, due to filtering, it is not necessarily positive. In order to obtain an eye diagram, we must derive $f_y(y)$, the pdf of $y(t)$. Our derivation of $f_y(y)$ is a generalization of Marcuse's [6]. However, we must first find a functional basis that diagonalizes both the square-law detection followed by the filtering $\mathcal{W}$ and the inverse covariance matrix ${\mathcal{K}^{(r)}}^{-1}$. Because both $\mathcal{W}$ and ${\mathcal{K}^{(r)}}^{-1}$ are symmetric matrices and, in addition, ${\mathcal{K}^{(r)}}^{-1}$ is positive definite, we can apply the theorem of simultaneous diagonalization [32], which states that there is a real matrix $\mathcal{C}$ satisfying ${\mathcal{K}^{(r)}}^{-1} = \mathcal{C}^T\mathcal{C}$ and $\mathcal{W} = \mathcal{C}^T\Lambda\mathcal{C}$. One procedure that may be used to obtain Λ and $\mathcal{C}$ is to solve the generalized eigenproblem $\mathcal{W}\mathcal{C}^{-1} = {\mathcal{K}^{(r)}}^{-1}\mathcal{C}^{-1}\Lambda$. The matrix Λ is diagonal, and we write it as $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_{2N})$, where the λ_k are real. Note that

if the impulse response of the filter can become negative, as in the case of a Bessel filter, some of the λ_k will be negative. With the definitions $Q_k = C_{kl}\bar{A}_l$ and the $q_k = C_{kl}r_l$ at each offset τ, we simplify (21) to

$$\begin{aligned} y(t) =& (\overline{\boldsymbol{A}}+\boldsymbol{r})^T\mathcal{C}^T\Lambda\mathcal{C}(\overline{\boldsymbol{A}}+\boldsymbol{r}) \\ =& (\boldsymbol{Q}+\boldsymbol{q})^T\Lambda(\boldsymbol{Q}+\boldsymbol{q}) \\ =& \sum_{k=1}^{2N}\lambda_k\left(Q_k^2+2Q_kq_k+q_k^2\right)=\sum_{k=1}^{2N}g_k \end{aligned} \tag{22}$$

where the g_k represent a new set of random variables. The transformation $\mathcal{C}$ yields what may be termed electrical Karhunen–Loève modes, namely the signal modes Q_k and the independent noise modes q_k. The noise pdf (12), with $\boldsymbol{a}$ replaced by $\boldsymbol{r}$, can be factored into independent Gaussian pdfs

$$\begin{aligned} f_{\boldsymbol{r}} =& (2\pi)^{-N}\sqrt{\det\mathcal{K}^{(r)^{-1}}}\exp\left[-\frac{1}{2}\boldsymbol{r}^T\mathcal{C}^T\mathcal{C}\boldsymbol{r}\right] \\ =& (2\pi)^{-N}\sqrt{\det\mathcal{K}^{(r)^{-1}}}\exp\left[-\frac{1}{2}\sum_{k=1}^{2N}q_k^2\right] \\ =& \prod_{k=1}^{2N}f_{q_k}(q_k) \end{aligned} \tag{23}$$

and, hence, $f_{q_k}(q_k) = \exp\left(-q_k^2/2\right)/\sqrt{2\pi} = N(0,1)$ is a normal distribution with zero mean and unit variance. Equation (22) is a sum of $2N$ random variables. The pdf f_y of y, therefore, involves a $2N$-dimensional convolution, but this convolution can be transformed into simple multiplications using characteristic functions. The characteristic function $\Phi_g(\zeta)$ of a random variable g is the expectation $E[\exp(i\zeta g)] = \int_{-\infty}^{\infty} f_g\exp(i\zeta g)dg$ [23]. With the help of the derived distribution identity $f_{g_k}dg_k = f_{q_k}dq_k$, we can write

$$\begin{aligned} &\Phi_{g_k}(\zeta) \\ &= \int_{-\infty}^{\infty}\exp[i\zeta g_k(q_k)]f_{q_k}dq_k \\ &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp\left[-\frac{q_k^2}{2}+i\zeta\lambda_k\left(Q_k^2+2Q_kq_k+q_k^2\right)\right]dq_k \\ &= \frac{1}{\sqrt{1-2i\lambda_k\zeta}}\exp\left(-\frac{2\lambda_k^2Q_k^2\zeta^2}{1-2i\lambda_k\zeta}+i\lambda_kQ_k^2\zeta\right) \end{aligned} \tag{24}$$

at each offset τ. Again, note that the integration variables are all real. The first term in the last exponent describes the influence of the signal–noise beating, while the second term stems from noise–noise beating in the receiver. The characteristic function $\Phi_{y,\tau=0}$ of the electrical current y in the absence of timing jitter satisfies $\Phi_{y,\tau=0} = \prod_{k=1}^{2N}\Phi_{g_k}$.

The relationship of the final pdf f_y that includes timing jitter and $f_{y,\tau=0}$, the pdf with zero timing jitter corresponding to $\Phi_{y,\tau=0}$, is

$$f_y(y,t) = \int_{-\infty}^{\infty} f_{y,\tau=0}(y,\tau)f_\tau(t-\tau)d\tau \tag{25}$$

where f_τ is the pdf of the timing jitter that we will obtain in Section III-A. The integral in (25) is a convolution with respect to τ and, hence, can be expressed as a product of characteristic functions; however, $\Phi_{y,\tau=0}$ is a characteristic function with respect to y, and we cannot use it to simplify (25). The characteristic function corresponding to f_y is

$$\Phi_y(\zeta,t) = \int_{-\infty}^{\infty}\Phi_{y,\tau=0}(\zeta,\tau)f_\tau(t-\tau)d\tau. \tag{26}$$

Because the timing jitter is Gaussian distributed, we use a numerical Gauss–Hermite integration [33] to solve (26). Additive noise sources such as electrical noise can, in principle, be accounted for by multiplying Φ_y with the appropriate characteristic functions [29]. Because y is phase independent, the phase variation does not contribute to Φ_y.

From Φ_y, we obtain f_y [23]

$$\begin{aligned} f_y(y,t) =& \frac{1}{2\pi}\int_{-\infty}^{\infty}\Phi_y(\zeta,t)\exp(-iy\zeta)d\zeta \\ =& \frac{1}{2\pi}\int_{-\infty}^{\infty}d\tau\int_{-\infty}^{\infty}d\zeta f_\tau(t-\tau) \\ & \times\exp\left[-i(y-y_{\rm nf})\zeta-\zeta^2\sum_{k=1}^{2N}\frac{2\lambda_k^2Q_k^2}{1-2i\lambda_k\zeta}\right] \\ & \times\prod_{k=1}^{2N}\frac{1}{\sqrt{1-2i\lambda_k\zeta}} \end{aligned} \tag{27}$$

where $y_{\rm nf} = \sum_{k=1}^{2N}\lambda_kQ_k^2$ is the noise-free electrical current and the quantities λ_k and Q_k are taken at the time τ. If the electrical Karhunen–Loève basis equals the Fourier basis (the matrix $\mathcal{C}$ equals the identity in this case) and all λ_k are equal, (27) is identical to [6, (20)]. Due to the complicated dependence of (27) on ζ, we cannot evaluate the Fourier transform in (27) analytically. However, $\Phi(\zeta)$ can be computed numerically using a discrete Fourier transform.

D. Gain Saturation

In the DMS system, we must include gain saturation in order to obtain good agreement with the experimentally observed evolution [5]. The basic assumption of our linearization approach is that $u_0 = \langle u\rangle$—namely, that the average of the received signal, including the noise, equals the noise-free transmission. However, one must be careful in the presence of saturable amplifiers, because the noise power that the amplifiers add to the signal increases the total power of $u(t) = u_0(t) + \delta u(t)$, according to

$$\left\langle||u||^2\right\rangle = ||u_0||^2 + \left\langle||\delta u||^2\right\rangle \tag{28}$$

where $||f||^2 = (1/T)\int_0^T|f(t)|^2dt$ for any function $f(t)$. The term $\left\langle||\delta u||^2\right\rangle$ is always finite and positive. Saturable amplifiers tend to keep the power of signal plus noise $||u||^2$ constant; hence, when more noise is added to the signal, the signal power decreases. If one attempts to compute u_0 by simply switching off the ASE noise input in the simulation of the saturated erbium-doped fiber amplifiers (EDFAs), the gains and hence the magnitude of the resulting field will be too large. Fortunately, EDFA saturation is a slow process that happens on a time scale of 1 ms, corresponding to 10–40 million bit periods in modern systems. Consequently, the amplifier gains adapt to the constant average power $\left\langle||u||^2\right\rangle$ and cannot follow variations in the noise. Hence, the effect of amplifier saturation is a mere gain renormalization. Setting the amplifiers to match these reduced gains in a static gain model (static gain corresponds to infinite saturation power), we can obtain the correct zeroth-order solution u_0.

III. Monte Carlo Simulation

A. Phase Jitter, Timing Jitter, and Residual Noise

In this section, we will employ Monte Carlo simulations to verify that the phase jitter obeys a Jacobi-Θ distribution, whereas the timing jitter is Gaussian distributed. We will also verify that the real and imaginary parts of the residual noise Fourier coefficients, after the jitter is separated, are multivariate Gaussian distributed. Our simulated transmission line consists of 225 periods of a dispersion map of length 106.7 km [5]. Each map contains a normal span of 4×25 km and dispersion $D_n = -1.03$ ps/nm-km and an anomalous span of 6.7 km and $D_a = 16.7$ ps/nm-km. The average dispersion D_{ave} equals 0.08 ps/nm-km. Third-order dispersion is not relevant in this system [5] and is set to zero. The carrier wavelength is 1551.49 nm, matching the experimental value. The fiber loss is compensated by five EDFAs. One follows each of the four 25-km segments of normal-dispersion fiber, and the fifth follows the segment of anomalous-dispersion fiber. There is a 2.8-nm optical bandpass filter in each map period to reduce the amount of noise. In accordance with Section II-D, the amplifiers are modeled as EDFAs with static gain, as opposed to explicitly including gain saturation. We carefully adjusted the static gains so that they equal the effective gains one would obtain using EDFAs with a saturation time of 1 ms and a saturation power of 10 mW, similar to [5]. The spontaneous emission factor is $n_{\text{sp}} = 1.4$. After each amplifier, we add a random amount of lumped noise separately to the real and imaginary part of the signal in the Fourier domain. This noise input is Gaussian distributed with zero means and variances $\rho_m^2/2$, where $\rho_m^2 = (G_m - 1)\eta$, and $G_m = \int_{z_m}^{z_m + L_{\text{amp}}} g_m dz$ is the gain associated with the mth amplifier [5]. All other quantities are defined in Section II. We chose a Box–Mueller generator [34] to obtain the Gaussian-distributed random variables; the generator takes its inputs from a 48-bit random number generator. The launched pulses have a Gaussian shape with a minimum full-width at half-maximum (FWHM) duration that equals the equilibrium value of 9 ps, and the signal is injected and received in the chirp-free midpoint of the anomalous span. We simulated the two different peak powers $P_{\text{peak}} = 5$ mW and $P_{\text{peak}} = 13$ mW. With the definition of a nonlinear scale length $L_{\text{nl}} = 1/(\gamma P_{\text{peak}})$, the distance of 24 000 km is 250 or 650 times larger than L_{nl}, respectively. We transmit the four-bit sequence 1-0-0-0 in a total simulation time window of $T = 400$ ps; hence, there is no interpulse interaction. The receiver is modeled as an ideal square-law detector with subsequent electrical low-pass fifth-order Bessel filter with a bandwidth of 8.6 GHz. We use the split-step Fourier method to solve the scalar nonlinear Schrödinger equation, which only takes into account one polarization. In the recirculating loop that we are modeling, the polarization dependent loss is large and the polarization controllers are optimized to pass the signal with minimum loss. Consequently, the orthogonal polarization is suppressed.

Using the least-squares method outlined in Section II-B, we now show that the central time offset of the pulses, β, as well as that the $\bar{B}_{n,R}$ and $\bar{B}_{n,I}$ are individually Gaussian distributed, where $\bar{B}_n = \bar{B}_{n,R} + i\bar{B}_{n,I}$, and the $\bar{B}_n$ are the transformed Fourier coefficients defined in (15). Furthermore, we show that

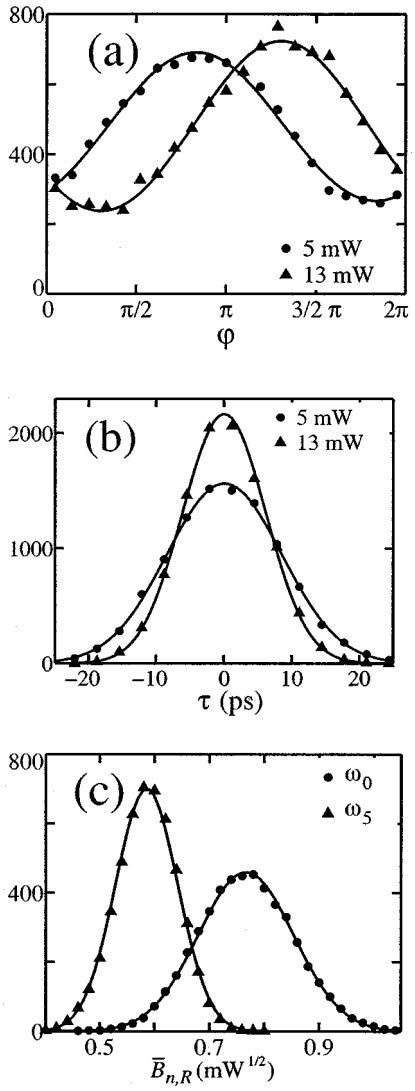

Fig. 1. (a) Histogram of the phase offset φ and (b) histogram of the time offset τ for two different simulations with signal peak powers $P_p = 5$ mW and $P_p = 13$ mW. (c) Histogram of the two real Fourier coefficients $\bar{B}_{0,R}$ and $\bar{B}_{5,R}$ at $\omega_0 = 0$ and $\omega_5 = 2\pi \times 25$ GHz, respectively, after the phase and time offsets are removed ($P_p = 5$ mW). The solid lines are fits of the Jacobi-Θ function in (a) and Gaussians in (b) and (c). The simulation consists of 10 000 Monte Carlo runs.

the distribution of the phase offset α is a Jacobi-Θ function, which is the periodic analogue of a Gaussian [33], with

$$\Theta\left(\mu_\varphi, \sigma_\varphi^2, 2\pi\right) = \sum_{k=-\infty}^{\infty} N\left(\mu_\varphi + 2\pi k, \sigma_\varphi^2\right) \qquad (29)$$

where $N(\mu, \sigma^2)$ is a Gaussian (normal) distribution of mean μ and variance σ^2. The Θ function is the natural choice for the phase fit because α is only determined modulo 2π.

Fig. 1(a) and (b) shows histograms of α and β for two different signal peak powers P_p. The two histograms are approximations to $f_\varphi(\varphi)$ and $f_\tau(\tau)$, respectively. Each simulation consists of 10 000 Monte Carlo runs. The phase distribution $f_\varphi(\varphi)$ converges to the Jacobi-Θ function, and $f_\tau(\tau)$ converges to a Gaussian distribution. We note that the variance of $\langle \tau^2 \rangle$ can be determined more elegantly from a linearized moment approach [15] without the need for Monte Carlo simulations.

The linear phase fit of (16) turns out to be very good even at the large transmission distance of 24 000 km. Pulses are launched at $t_c = 50$ ps. Fig. 1(c) shows histograms of $\bar{B}_{0,R}$ and $\bar{B}_{5,R}$ at the angular frequencies $\omega_0 = 0$ and $\omega_5 = 2\pi \times 25$ GHz, respectively. The simulated data agree very well with the Gaussian fit. The way we removed the linear part of the signal phase causes the imaginary parts of the $\bar{B}_k$ to be close to zero; so, they are not shown here. We verified that all the $\bar{B}_{k,R}$ in our simulation are Gaussian distributed using a chi-square statistical test [35].

B. Computation of the Covariance Matrix From Monte Carlo Simulations

We employ Monte Carlo simulations to compute the reduced covariance matrix $\mathcal{K}^{(r)}$ according to (18). In this section, we present the resulting $\mathcal{K}^{(r)}$ for the DMS system that we are studying, and we compare the resulting pdfs for the marks and the spaces. We then calculate the BER as a function of the decision threshold and determine the BER at the optimal decision point. According to (13), $\mathcal{K}^{(r)}$ is a block matrix consisting of 2×2 blocks of size $N \times N$ each. We will name the four blocks RR, RI, RI, and II, where RR is the block $\langle \boldsymbol{\alpha}_R \boldsymbol{\alpha}_R^T \rangle$, RI is $\langle \boldsymbol{\alpha}_R \boldsymbol{\alpha}_I^T \rangle$, and so on. Because $\mathcal{K}^{(r)}$ is symmetric, $(\mathrm{RI})^T = \mathrm{IR}$. In the following, we will index the elements of the matrix $\mathcal{K}^{(r)}$ by the frequencies whose covariance is contained at each element, so that the upper left matrix element is $K^{(r)}_{-N/2,-N/2}$, the lower right one is $K^{(r)}_{N/2-1,N/2-1}$ and the center element is $K^{(r)}_{0,0}$.

Fig. 2(a) displays the block RR in a three-dimensional (3-D) form. The ridge along the principal diagonal represents the variances $K^{(r)}_{kk}$; all other elements correspond to cross covariances. The CW entry lies at $k = 0$. Fig. 2(b) portrays three slices through the matrix, along the principal diagonal, orthogonal to that (starting from the indices $(-15, 15)$ and ending at $(15, -15)$), and along a parallel to the principal diagonal. The last two slices reveal that the values of $K^{(r)}_{k,k\pm l}$ at small $|k|$ and $|l|$ are actually negative, leading to elongated minima along both sides of the principal diagonal. The shape of the graph of $K^{(r)}_{kk}$ is due to the optical inline filtering in the recirculating loop, as well as signal–noise beating. Because of the inline filters, $K^{(r)}_{kk}$ vanishes for large $|k|$. In the absence of inline filters, $K^{(r)}_{kk}$ would converge for large $|k|$ to the finite value $\sum_m \rho_m^2/2$, where the sum runs over all optical amplifiers. At small $|k|$, the signal–noise beating emerges as a peak of a shape similar to the signal power spectrum $|A_k|^2$.

Fig. 3 shows the block RI whose maximum is about one order of magnitude smaller than that in RR and II. The RI and IR blocks are point symmetric around the center, whereas RR and II obey a mirror symmetry.

We briefly revisit the modeling of gain saturation. In the present DMS system, we compared the average power of the signal and noise $\langle ||u||^2 \rangle$, that we obtain using saturated EDFAs, to the simulated signal power in the absence of ASE noise input $||u_{0,\mathrm{sat}}||^2$, which yields $||u_{0,\mathrm{sat}}||^2 / \langle ||u||^2 \rangle \approx 1.17$. As discussed in Section II-D, $u_{0,\mathrm{sat}}$ is not the proper zeroth-order solution to linearize around. We modeled the amplifiers with static gains rather than employing a saturated EDFA model [5] to compute the covariance matrix $\mathcal{K}$, because we only propagate one pulse in the relatively short time window of 400 ps, and noise power variations would have an exaggerated effect on the solution. Because we are modeling a recirculating loop and the signal passes each amplifier many times, we have to record the static gains for each amplifier and each round trip. This procedure enables us to compute the correct zeroth-order solution u_0.

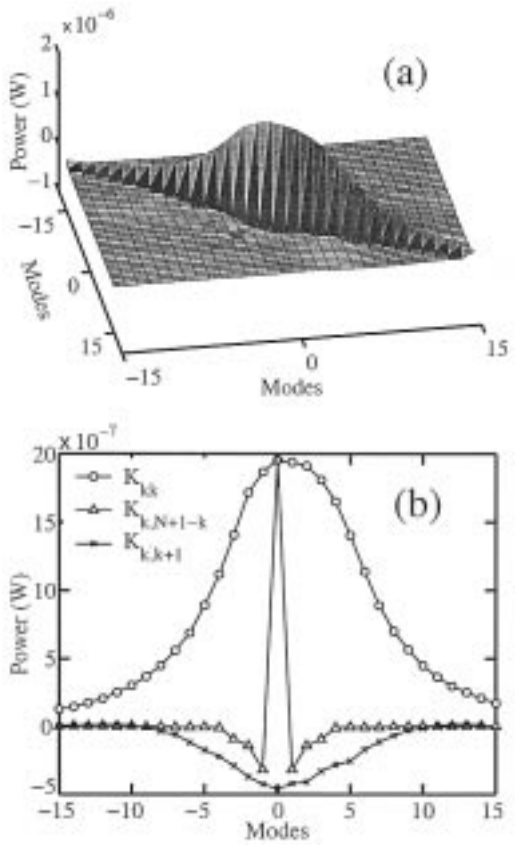

Fig. 2. (a) 3-D plot of the RR block of $\mathcal{K}^{(r)}$ for ω_k in the range $k = [-15, 15]$. (b) Three slices through $\mathcal{K}^{(r)}$. The open circles show the principal diagonal $\left(K^{(r)}_{kk}\right)$ and the triangles the secondary diagonal ($K^{(r)}_{k,N+1-k}$, orthogonal to the principal diagonal). The symbol $\times$ denotes matrix values on a parallel to the principal diagonal ($K_{k,k+1}$). Note the minima in the lower curves where $\mathcal{K}^{(r)}$ becomes negative, corresponding to trenches on both sides of the principal diagonal that are hard to see in (a).

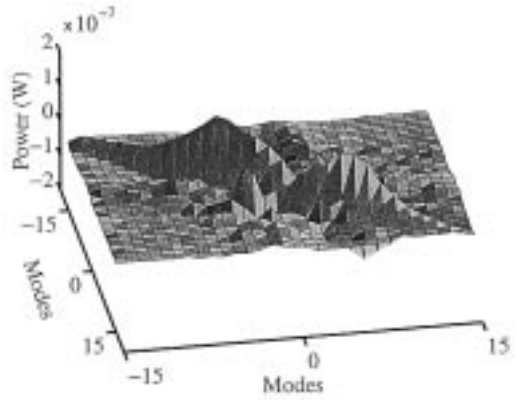

Fig. 3. 3-D plot of the RI block of $\mathcal{K}^{(r)}$, similar to Fig. 2. The values in RI are smaller than those in RR by about an order of magnitude, but not zero. Note the point symmetry of RI, as opposed to the mirror symmetry of RR.

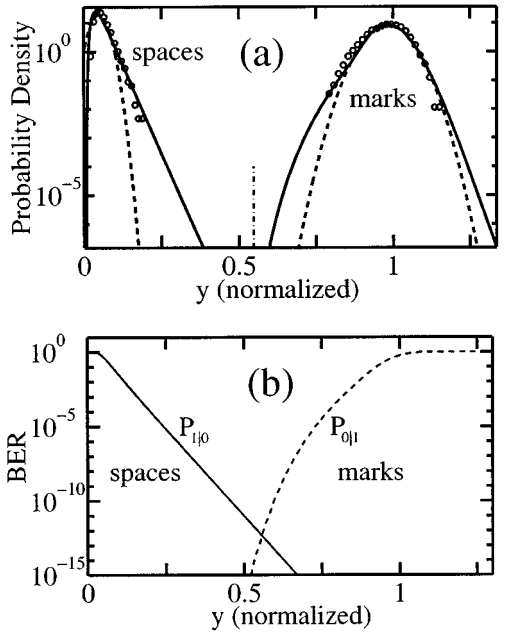

Fig. 4. (a) Probability density function of the filtered current y after square-law detection and an 8.6-GHz Bessel filter. The solid lines are the pdfs calculated using linearization at the center of the bit window of the marks ("1"s) and the spaces ("0"s); the dashed lines are Gaussian fits. Note the bump in the left tail of the marks pdf. The circles are direct results from the Monte Carlo simulation and can be thought of as a slice through an electrical eye diagram. Particularly for the spaces, the agreement with the pdf obtained by linearization is much better than with the Gaussian fit. (b) BER as a function of y. The function $P_{1|0}(y)$ is shown as a solid line, and $P_{0|1}(y)$ is shown dashed. The BER equals half the sum of the two. The optimal decision level lies at $y = 0.55$ near the intersection of the graphs [see the vertical dash–dotted line in (a)] and yields a BER of 5×10^{-13}. From the Gaussian fits, we obtain a Q factor of 13.5, implying an optimal BER of 10^{-41} instead.

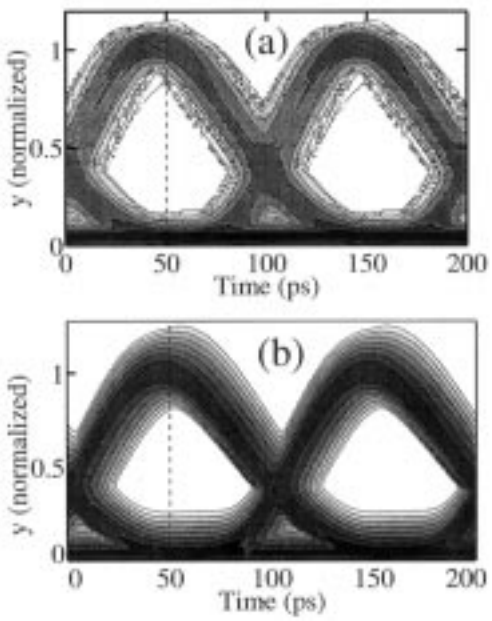

Fig. 5. (a) Eye diagram, taking its input directly from the Monte Carlo simulation of the DMS simulation. The probability density of the current y is displayed as a contour plot. The dashed line at $t = 50$ ps shows the location of the pdf in Fig. 4. The logarithm of the pdf is displayed as different shades of gray. (b) Accurate eye diagram produced by our linearization approach. To obtain a more readable diagram, we only plot probability densities in the range $[10^{-4}, 10^{1}]$. However, our approach allows us to find the probability density at any point (t, y), thereby enabling us to calculate accurate BERs.

C. Calculating the BER

Fig. 4(a) shows the pdf f_y of the filtered current $y(t)$ defined in (21) that corresponds to the output of the electrical receiver. The calculation is performed as explained in Section II-C in two different bit slots, corresponding to the 1 and the central 0 in the 1-0-0-0 pattern that we are simulating, so that we obtain the pdfs for the marks and the spaces separately. The effect of the timing jitter is included in the calculation of the pdfs, as shown in (26). Note that the Gaussian fits are a good approximation over about two orders of magnitude, but deviate strongly at low probability densities. We can show mathematically that the pdf in the spaces can be regarded as a generalization of the *central* chi-square distribution [27]; its exponential decay results from the quadratic noise term in the receiver [6]. The pdf of the marks can be regarded as a generalized *noncentral* chi-square distribution and is controlled by the noncentrality parameter, which is proportional to $u_0(t)$, and the degree of freedom, M. The effective value of M depends on the bandwidth ratio of the optical and the electrical filters and is usually quite small ($M \approx 2$–6). For larger values of M, the noncentral chi-square pdf converges to Gaussian as a consequence of the central limit theorem. At small values of y, the pdf of the marks is dominated by timing jitter, leading to a visible bump. This bump exists because the timing jitter broadens the signal, widening the pdf of the marks inside the eye, toward smaller y. Without this bump, the left tail of the accurate pdf for the marks would cross the Gaussian fit and then run inside, so that the error probability density would be lower there than for a Gaussian pdf, in agreement with [6]. We conclude that the strong nonlinearity in some optical systems can lead to an increased eye penalty due to timing jitter that must be taken into account. Note that the onset of the bump in our system occurs at a low probability density. In the range that can be explored by a standard Monte Carlo simulation, denoted by the circles, the agreement between the pdf and its Gaussian fit is still very good. However, the impact on the optimal BER is large.

The knowledge of the separate pdfs of the marks and spaces allows us to calculate the bit error probabilities. We define the bit error probability at the decision level y as $\mathrm{BER}(y) = [P_{1|0}(y) + P_{0|1}(y)]/2$ [6]. Here, the quantity $P_{1|0}(y)$ is the probability of detecting a mark when a space was transmitted, using the decision level y, and is defined as $P_{1|0}(y) = \int_y^\infty f_y(y', t_1)dy'$. Similarly, the probability of detecting a space when a mark was sent is defined as $P_{0|1}(y) = \int_0^y f_y(y', t_0)dy'$. The quantity $f_y(y, t_1)$ is the pdf of y, taken at the central time in a bit window t_1 when a mark is received, and $f_y(y, t_0)$ is taken at the central time in a bit window when a space is received, t_0. The BER becomes minimal at the optimal decision level y_{opt}. Fig. 4(b) shows $P_{1|0}(y)$ and $P_{0|1}(y)$, as well as the BER as a function of y. We find that at $y_{\mathrm{opt}} = 0.55$, the BER is 5×10^{-13}. From the Gaussian fits, we obtain a Q-factor of 13.5, implying an optimal BER of 10^{-41} instead.

Fig. 5 shows the corresponding eye diagram. It is a contour plot of the pdf for $y(t)$ at each time t. We note that this way of plotting the eye diagram is closer to what is measured experimentally than the current, standard numerical practice of simply superimposing $y(t)$ in the different bit windows. The optimal decision point for our system lies close to one half, normalized

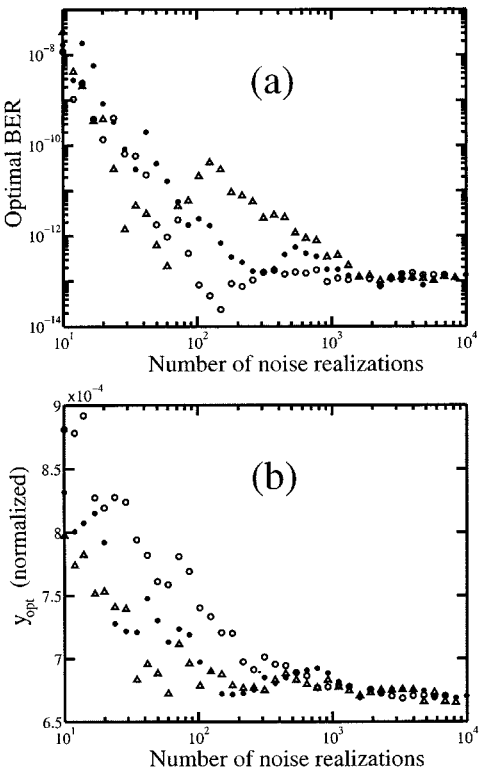

Fig. 6. (a) Convergence of optimal BER values, resulting from $\mathcal{K}^{(r)}$ and A_k obtained from a Monte Carlo simulation, as a function of the number of noise realizations. The symbols ($\triangle$), ($\bullet$), and ($\circ$) pertain to three simulations that were started with different random seeds. (b) Convergence of the optimal decision level $y_{\rm opt}$. Both BER and $y_{\rm opt}$ converge after about 2000 Monte Carlo realizations.

to the location of the peak of the marks pdf. Gaussian extrapolations usually yield optimal decision levels that are much smaller than 0.5 in normalized units [6].

We now consider the off-diagonal elements in $\mathcal{K}^{(r)}$. Although $\mathcal{K}^{(r)}$ is diagonally dominant, we found that the off-diagonal elements have a large impact on the resulting BER. To quantify this, we compared the optimal BER to a computation in which we set the off-diagonal elements to zero. It turns out that the spread in the pdf of the spaces is reduced, as one might expect, whereas the spread in the pdf of the marks is increased, leading to a lower $y_{\rm opt}$ and an optimal BER that deviates from the true value by orders of magnitude. Therefore, we conclude that the nonlinearity can lead to a substantial deviation of the Karhunen–Loève basis from the Fourier basis and it is, in fact, necessary to keep track of off-diagonal elements in K_{kl}.

Next, we turn to the question of the accuracy of the BER and the decision level that we obtain by employing a Monte Carlo simulation to compute $\mathcal{K}^{(r)}$. Fig. 6 shows the convergence of the BER and the optimal decision level $y_{\rm opt}$ as the simulation proceeds. Both $\mathcal{K}^{(r)}$ and $y_{\rm opt}$ converge as we average over more noise realizations. The convergence requires on the order of a few thousand realizations. Looking at the BER is a very stringent test of the accuracy, because a small deviation in the variance of the pdf curves, that might be barely visible in a plot such as Fig. 4(a), leads to a large variation of the BER close to $y_{\rm opt}$. On the other hand, a deviation in the BER from the true value is often tolerable, as long as the BER is well below a given threshold such as 10^{-9}. Note also that the computation of amplitude margins for a given BER requires less accuracy and, in our case, 100 Monte Carlo realizations is sufficient. We note that the values for the optimal BER and the decision level both drop as the averaging proceeds. We find that the statistical fluctuations in $\mathcal{K}^{(r)}$, that vanish as the number of realizations increases, tend to decrease the BER, irrespective of their signs. Thus, the square-law detector with its dependence of $f_y(y)$ on $\mathcal{K}^{(r)}$ corresponds to a biased estimator [36].

IV. Direct Solution of the Noise ODE

As mentioned in Section II-B, a Monte Carlo simulation is only one way to obtain the reduced covariance matrix $\mathcal{K}^{(r)}$. If we find a way to solve (14) in parallel with (1), we can replace the time-consuming Monte Carlo method by a deterministic method. We solve (1) using the standard split-step method. Simultaneously, we must compute $\mathcal{K}(z)$. The propagation of the accumulated noise is governed by (4), which is a linear equation and is homogeneous except at the amplifiers because $\hat{F} = 0$ in the fiber. Its Fourier transform must be linear and homogeneous as well, and we can write it in terms of $\boldsymbol{a}$ as in (9)

$$\frac{d\boldsymbol{a}}{dz} = \mathcal{R}(z)\boldsymbol{a} \tag{30}$$

where we set the ASE term $\boldsymbol{w}(z)$ to zero. We write the solution of (30) as $\boldsymbol{a}(z) = \Psi(z)\boldsymbol{a}(0)$, consistent with (10). We infer that the evolution of $\mathcal{K}$ over one fiber span, followed by an EDFA, is given by

$$\mathcal{K}_{\rm out} = \Psi\mathcal{K}_{\rm in}\Psi^T + \frac{\eta I_0}{T}\mathcal{I} \tag{31}$$

where $\mathcal{K}_{\rm in}$ equals $\mathcal{K}$ at the beginning of the span, $\mathcal{K}_{\rm out}$ equals $\mathcal{K}$ after the amplifier, $\mathcal{I}$ is the identity matrix, T is the period of the Fourier expansion, η is the same as in (3). The last term represents the lumped ASE noise input. We choose a perturbative method to compute Ψ rather than solving (30) directly. Let $u_0(t,0)$ and $u_0(t,L)$ be the noise-free optical field at the beginning and end of a fiber span of length L. We then perturb the incident field in the k-th frequency mode so that $u^{(k)}(t,0) = u_0(t,0) + \Delta\exp(i\omega_k t)$, solve (1) again for the perturbed field $u^{(k)}(t,0)$, and obtain $u^{(k)}(t,L)$. The resulting noise vector $\boldsymbol{a}^{(k)}$ corresponds to $\delta u^{(k)}(t) = u^{(k)}(t,L) - u_0(t,L)$. We then divide $\boldsymbol{a}^{(k)}$ by Δ, which yields the kth column of the matrix Ψ. This perturbative approach corresponds to the Lyapunov method described by Bennetin *et al.* [37]. We find that the value of Ψ is independent of Δ over a wide range and that the entire method is very stable and accurate up to a propagation distance of about 3000 km, as well as very efficient in comparison with the Monte Carlo simulation. Beyond 3000 km, it is necessary to separate the phase and timing jitter continuously from the covariance matrix. Once that is done, it is possible to extend this approach over the full 24 000 km. We will discuss this issue in more detail in a later publication.

We also attempted a direct solution of (14) using the matrix ODE solver package, variable-coefficient ODE solver written in C (CVODE) [38], but the solution was numerically inefficient.

V. Conclusion

We extend the linearization method that describes the noise evolution in an optical transmission system and apply it to a

highly nonlinear DMS system over a transmission distance of 24 000 km. This system is well characterized both theoretically and experimentally [5], [16]. We are able to accurately calculate the optical noise distribution at the receiver and from that obtain accurate eye diagrams and bit error rates (BERs). We compare these results to the standard Monte Carlo simulation technique. To retain the linearity of the system, we find that we must separate the phase and timing jitter from the remainder of the noise. Our simulation confirms that the timing jitter is Gaussian distributed [15], and the phase jitter is distributed according to a Jacobi-Θ function, the periodic analogue of a Gaussian. After separating the jitter, the remaining noisy signal is multivariate Gaussian distributed with finite cross correlations. This result extends and corrects previous work that neglected fiber transmission effects and assumed ideal pulse and filter profiles [6], [7]. We reintroduce the effect of the jitter in the computation of the BER and show that it leads to an additional eye penalty, although this penalty only affects the pdfs at low probability densities and is, therefore, very hard to see using standard Monte Carlo simulations. Neglecting this additional penalty leads to a substantial underestimate of the BER. We note that it is unlikely to be necessary to treat the jitter separately from the rest of the noise in systems that are less nonlinear than ours, as is the case in all commercial systems. The success of the linearization approach in our highly nonlinear system bodes well for its future.

The central statistical quantity in our calculations is the covariance matrix describing the multivariate Gaussian distribution. We present two methods for calculating the evolution of this matrix. One is based on standard Monte Carlo simulations, and the other is based on solving its linear evolution equation. We show that the Monte Carlo method works well over the entire length of 24 000 km, corresponding to the experimental system. We validated the linearization by comparison of its predictions to 10 000 realizations of the Monte Carlo method. We showed that less than 2000 realizations are necessary to accurately calculate the elements of the covariance matrix. It is also possible to directly calculate the covariance matrix by solving the ODE that governs its evolution. However, to do so over the entire length of 24 000 km, it is necessary to continuously separate the time and phase jitter from the rest of the covariance matrix. We will discuss this issue in detail in a future publication.

The main limitation of our current approach is the absence of interpulse interactions since we only propagate a single pulse in a single wavelength channel. This assumption is appropriate for the particular experimental system that we studied, but is not appropriate for most commercial systems. In the future, we will study generalizations of our method to multiple pulses and also apply it to different modulation formats, such as CRZ and NRZ, that are directly relevant to commercial systems.

APPENDIX

At the chirp-free maximum pulse compression point in the loop, we define the average signal phase

$$\alpha_{\text{ave}} = \frac{\sum_{k=-N/2}^{N/2-1} |B_k|^2 \arctan\left(\frac{B_{k,I}}{B_{k,R}}\right)}{\sum_{k=-N/2}^{N/2-1} |B_k|^2} \tag{32}$$

where $B_k = B_{k,R} + iB_{k,I} = A_k + a_k$ are the Fourier modes of $u(t)$. First, we would like to determine the relationship of α_{ave} in (32) and α in (16) to φ in (17), and analogously the relationship of β to τ, for small noise. The Taylor expansion of the signal phase at frequency mode k, using the identity $d\arctan x/dx = (1+x^2)^{-1}$, is

$$\arctan\frac{B_{k,I}}{B_{k,R}} = \alpha_0 + \frac{a_{k,I}A_{k,R} - a_{k,R}A_{k,I}}{|A_k|^2} + O\left(a_k^2\right) \tag{33}$$

where $\alpha_0 = \arctan(A_{k,I}/A_{k,R})$. Using (33), we linearize the least-squares criterion (16) as

$$H = \min_{\alpha,\beta} \sum_{k=-N/2}^{N/2-1} \frac{1}{|A_k|^2} \times \left[a_{k,I}A_{k,R} - a_{k,R}A_{k,I} - (\delta\alpha + \beta\omega_k)|A_k|^2\right]^2 \tag{34}$$

where $\delta\alpha = \alpha - \alpha_0$. The stationary points of the sum in (34) with respect to α and β are given by

$$\frac{\delta H}{\delta\alpha} = -2\sum_k \left[a_{k,I}A_{k,R} - a_{k,R}A_{k,I} - (\delta\alpha + \beta\omega_k)|A_k|^2\right] = 0 \tag{35a}$$

$$\frac{\delta H}{\delta\beta} = -2\sum_k \omega_k \left[a_{k,I}A_{k,R} - a_{k,R}A_{k,I} - (\delta\alpha + \beta\omega_k)|A_k|^2\right] = 0. \tag{35b}$$

We recall $a_k = \varphi c_k + \tau d_k + r_k$ with $c_k = iA_k$ and $d_k = i\omega_k A_k$ from (17) and express the associated orthogonality relations $(\boldsymbol{c},\boldsymbol{r}) = (\boldsymbol{d},\boldsymbol{r}) = 0$ as $\sum_k \left[r_{k,I}A_{k,R} - r_{k,R}A_{k,I}\right] = 0$ and $\sum_k \omega_k \left[r_{k,I}A_{k,R} - r_{k,R}A_{k,I}\right] = 0$, from which we infer

$$\sum_k [a_{k,I}A_{k,R} - a_{k,R}A_{k,I}] = \sum_k (\varphi + \tau\omega_k)|A_k|^2 \tag{36a}$$

$$\sum_k \omega_k [a_{k,I}A_{k,R} - a_{k,R}A_{k,I}] = \sum_k \omega_k(\varphi + \tau\omega_k)|A_k|^2. \tag{36b}$$

Comparing the last two identities with (35a) and (35b), we find that the sums in (35a) and (35b) vanish with the choice $\delta\alpha = \varphi$ and $\beta = \tau$. This result shows that our definition of an average phase (32) is reasonable and consistent with the least-squares fit (16).

ACKNOWLEDGMENT

The authors would like to thank D. Marcuse, J. Zweck, and B. Marks for help with the manuscript and inspiring discussions.

REFERENCES

[1] J. P. Gordon and H. A. Haus, "Random walk of coherently amplified solitons in optical fiber transmission," *Opt. Lett.*, vol. 11, pp. 665–667, 1986.

[2] H. A. Haus, "Quantum noise in solitonlike repeater systems," *J. Opt. Soc. Amer. B*, vol. 8, pp. 1122–1126, 1991.

[3] G. P. Agrawal, *Nonlinear Fiber Optics*, 2nd ed. London, U.K.: Academic, 1995.

[4] 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.

[5] R.-M. Mu, V. S. Grigoryan, C. R. Menyuk, G. M. Carter, and J. M. Jacob, "Comparison of theory and experiment for dispersion-managed solitons in a recirculating fiber loop," *IEEE J. Select. Topics Quantum Electron.*, vol. 6, pp. 248–257, Mar./Apr. 2000.
[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. Azizoğlu, "On the bit error rate of lightwave systems with optical amplifiers," *J. Lightwave Technol.*, vol. 9, pp. 1576–1582, Nov. 1991.
[8] R. Hui, D. Chowdhury, M. Newhouse, M. O'Sullivan, and M. Poettcker, "Nonlinear amplification of noise in fibers with dispersion and its impact in optically amplified systems," *IEEE Photon. Technol. Lett.*, vol. 9, pp. 392–394, Mar. 1997.
[9] R. Hui, M. O'Sullivan, A. Robinson, and M. Taylor, "Modulation instability and its impact in multispan optical amplified IMDD system: Theory and experiments," *J. Lightwave Technol.*, vol. 15, pp. 1071–1081, July 1997.
[10] A. Carena, V. Curri, R. Gaudino, P. Poggiolini, and S. Benedetto, "New analytical results on fiber parametric gain and its effects on ASE noise," *IEEE Photon. Technol. Lett.*, vol. 9, pp. 535–537, Apr. 1997.
[11] G. Bosco, A. Carena, V. Curri, R. Gaudino, P. Poggiolini, and S. Benedetto, "Parametric gain in multiwavelength systems: A new approach to noise enhancement analysis," *IEEE Photon. Technol. Lett.*, vol. 11, pp. 1135–1137, Sept. 1999.
[12] E. A. Golovchenko, A. N. Pilipetskii, N. S. Bergano, C. R. Davidsen, F. I. Khatri, R. M. Kimball, and V. J. Mazurczyk, "Modeling of transoceanic fiber-optic WDM communications systems," *IEEE J. Select. Topics Quantum Electron.*, vol. 6, pp. 337–347, Mar./Apr. 2000.
[13] M. Midrio, "Analytical performance evaluation of nonreturn-to-zero transmission systems operating in normally dispersive nonlinear fibers," *J. Lightwave Technol.*, vol. 15, pp. 2038–2051, Nov. 1997.
[14] A. N. Pilipetskii, V. J. Mazurczyk, and C. J. Chen, "The effect of dispersion compensation on system performance when nonlinearities are important," *IEEE Photon. Technol. Lett.*, vol. 11, pp. 284–286, Feb. 1999.
[15] V. S. Grigoryan, C. R. Menyuk, and R.-M. Mu, "Calculation of timing and amplitude jitter in dispersion-managed optical fiber communications using linearization," *J. Lightwave Technol.*, vol. 17, pp. 1347–1356, Aug. 1999.
[16] J. M. Jacob and G. M. Carter, "Error-free transmission of dispersion-managed solitons at 10 Gbit/s over 24 500 km without frequency sliding," *Electron. Lett.*, vol. 33, pp. 1128–1129, June 1997.
[17] C. R. Menyuk, "Non-Gaussian corrections to the Gordon-Haus distribution resulting from soliton interactions," *Opt. Lett.*, vol. 20, pp. 285–287, 1995.
[18] T. Georges, "Bit-error rate degradation of interacting solitons owing to non-Gaussian statistics," *Electron. Lett.*, vol. 31, pp. 1174–1175, July 1995.
[19] ——, "Study of the non-Gaussian timing jitter statistics induced by soliton interaction and filtering," *Opt. Commun.*, vol. 123, pp. 617–623, 1996.
[20] N. S. Bergano, "Undersea amplified lightwave systems design," in *Optical Fiber Telecommunications*, I. E. Kaminow and T. L. Koch, Eds. San Diego, CA: Academic, 1997, vol. IIIA, ch. 10, pp. 302–335.
[21] A. Dembo and O. Zeitouni, *Large Deviation Techniques and Applications*. Boston, MA: Jones & Barlett, 1993.
[22] P. M. Hahn and M. C. Jeruchim, "Developments in the theory and application of importance sampling," *IEEE Trans. Commun.*, vol. COM-35, pp. 706–714, 1987.
[23] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 3rd ed. New York: McGraw-Hill, 1991.
[24] J. L. Doob, "The Brownian movement and stochastic equations," in *Selected Papers on Noise and Stochastic Processes*, N. Wax, Ed. New York: Dover, 1954.
[25] D. J. Kaup, "Perturbation theory for solitons in optical fibers," *Phys. Rev. A*, vol. 42, pp. 5689–5694, 1990.
[26] L. Dieci and T. Eirola, "Positive definitenness in the numerical solution of Riccati differential equations," *Num. Math.*, vol. 67, pp. 303–313, 1994.
[27] J. G. Proakis, *Digital Communications*. New York: McGraw-Hill, 1995.
[28] L. Arnold, *Stochastic Differential Equations: Theory and Applications*. Melbourne, FL: Krieger, 1992.
[29] J.-S. Lee and C.-S. Shim, "Bit-error-rate analysis of optically preamplified receivers using an eigenfunction expansion method in optical frequency domain," *J. Lightwave Technol.*, vol. 12, pp. 1224–1229, July 1994.
[30] G. Bosco, A. Carena, V. Curri, R. Gaudino, P. Poggiolini, and S. Benedetto, "A novel analytical method for the BER evaluation in optical systems affected by parametric gain," *IEEE Photon. Technol. Lett.*, vol. 12, pp. 152–154, Feb. 2000.
[31] E. Forestieri, "Evaluating the error probability in lightwave systems with chromatic dispersion, arbitrary pulse shape and pre- and postdetection filtering," *J. Lightwave Technol.*, vol. 18, pp. 1493–1503, 2000.
[32] J. Franklin, *Matrix Theory*. New Jersey: Prentice-Hall, 1968.
[33] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*. New York: Dover, 1965.
[34] D. E. Knuth, "The art of computer programming," in *Seminumerical Algorithms*, 3rd ed. Reading, MA: Addison-Wesley, 1998.
[35] H. D. Brunck, *An Introduction to Mathematical Statistics*, 2nd ed. Lexington, MA: Xerox College, 1965.
[36] H. Stark and J. E. Woods, *Probability, Random Processes, and Estimation Theory for Engineers*. New Jersey: Prentice-Hall, 1994.
[37] G. Bennetin, L. Galgani, and J. M. Strelcyn, "Kolgomorov entropy and numerical experiments," *Phys. Rev. A*, vol. 14, pp. 2338–2345, 1976.
[38] S. D. Cohen and A. C. Hindmarsh, *CVODE User Guide*. Livermore, CA: Lawrence Livermore National Laboratory, 1994. Available: http://www.llnl.gov/CASC/PVODE/.

Ronald Holzlöhner (SM'00) was born in Essen, Germany, on December 30, 1970. He received the M.S. degree in physics from the Technical University of Berlin, Berlin, Germany, in 1998. From 1995 to 1996, he studied at the University of California, Santa Barbara, as a Fulbright exchange student. He is currently pursuing the Ph.D. degree at the University of Maryland Baltimore County (UMBC), Baltimore, MD.

He works part time as a consultant for Virtual Photonics, Inc.

Mr. Holzlöhner is a member of the Optical Society of America (OSA).

V. S. Grigoryan, photograph and biography not available at the time of publication.

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

He has worked as 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), Baltimore, and he was the founding member of this department. In 1993, he was promoted to Professor. He has been on partial leave from UMBC since 1996. From 1996 to 2001, he worked part-time for the Department of Defense (DoD), codirecting the Optical Networking Program at the DoD Laboratory for Telecommunications Sciences, Adelphi, MD, from 1999 to 2001. In August 2001, he left the DoD and became Chief Scientist at PhotonEx Corporation, Maynard, MA. For the last 15 years, his primary research area has been theoretical and computational studies of fiber-optic communications. He has authored or coauthored more than 140 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 Fellow of the Optical Society of America (OSA) and a member of the Society for Industrial and Applied Mathematics and the American Physical Society. He is a former UMBC Presidential Research Professor.

W. L. Kath, photograph and biography not available at the time of publication.