Polarization Mode Dispersion, Decorrelation, and Diffusion in Optical Fibers with Randomly Varying Birefringence

P. K. A. Wai and C. R. Menyuk

Abstract— **Polarization mode dispersion and the polarization decorrelation and diffusion lengths are calculated in fibers with randomly varying birefringence. Two different physical models in which the birefringence orientation varies arbitrarily are studied and are shown to yield nearly identical results. These models are appropriate for communication fibers. We show that both the length scales for polarization mode dispersion and polarization decorrelation measured with respect to the local axes of birefringence are equal to the fiber autocorrelation length. We also show that the coupled nonlinear Schrödinger equation which describes wave evolution over long length along a communication fiber can be reduced to the Manakov equation. The appropriate averaging length for the linear polarization mode dispersion is just the fiber autocorrelation length but the appropriate averaging length for the nonlinear terms is the diffusion length in the azimuthal direction along the Poincaré sphere which can be different. The implications for the nonlinear evolution are discussed.**

I. Introduction

POLARIZATION mode dispersion plays an important role in modern-day, long-distance communication systems that depend upon erbium-doped fiber amplifiers rather than repeaters to compensate for loss. Just like signal distortion due to chromatic dispersion and nonlinearity accumulate along the length of a communication link, so does the signal distortion due to polarization mode dispersion. Physically, polarization mode dispersion has its origin in the birefringence that is present in any optical fiber. While this birefringence is small in absolute terms in communication fibers, with values of $\Delta n/n \sim 10^{-7}$, the corresponding beat length L_B is only about 10 m—far smaller than the dispersive or nonlinear scale lengths which are typically hundreds of kilometers or more—so that the birefringence should be considered large. This large birefringence would be devastating in communication systems but for the fact that the orientation of the birefringence is randomly varying on a length scale that is on the order of 100 m.

Manuscript received February 15, 1995; revised October 18, 1995. This work was supported by the Department of Energy, the National Science Foundation, and the Advanced Research Projects Agency through the Air Force Office of Scientific Research.

P. K. A. Wai is with the Department of Computer Science and Electrical Engineering, University of Maryland, Baltimore, MD 21228 USA. He is also with the Institute of Plasma Research, University of Maryland, College Park, MD 20742 USA.

C. R. Menyuk is with the Department of Computer Science and Electrical Engineering, University of Maryland, Baltimore, MD 21228 USA.

Publisher Item Identifier S 0733-8724(96)01428-4.

The rapid variation of the birefringence orientation tends to make the effect of the birefringence average out to zero. The residual effect leads to pulse spreading, referred to as polarization mode dispersion. This effect has been the object of extensive experimental and theoretical study [1]–[14]. In early work, Poole and Wagner [1] showed that over a finite length of fiber, in the absence of polarization-dependent loss, there are two orthogonal polarization states that map into two other orthogonal polarization states, independent of frequency to first order. Mathematically, the frequency-derivative of the 2×2 propagation matrix is Hermitian; it follows that its eigenvectors are orthogonal and correspond to the principal states, while its eigenvalues are real and the difference between them corresponds to the differential delay time τ_d. Shortly thereafter, this theoretical result was experimentally verified [2], [3]. Poole [4] derived the expectation for the differential delay time, when averaged over an ensemble of fibers, using a model in which he assumed that the fiber has nearly fixed axes of birefringence, varying only slightly but very rapidly, leading to coupling of the orthogonal eigenmodes. This model is similar to one proposed originally by Kaminow [15] which has been used to investigate polarization-holding in polarization-preserving fiber and which has been shown experimentally to be very useful in that context [16], [17]. Despite its lack of obvious physical relevance to communication fibers in which the orientation of the birefringence can vary arbitrarily, experiments are in good agreement with the predicted behavior [6]–[8]. Moreover, a later prediction based on the same model that the values of τ_d are distributed according to a three-dimensional Maxwellian is also in good agreement with the experiments [7], [8]. In recent work, we provided a theoretical explanation for the success of this model by showing that models in which the birefringence orientation varies randomly are related to Poole's model by a simple mathematical transformation, so that the distribution of the differential delay time will be the same [13]. Indeed, we showed that any sufficiently random model will have the same behavior. We also proved an ergodic theorem, demonstrating the relevance of ensemble-averaged results to a single, long length of optical fiber. We then considered two specific physical models [14]. In the first model, we allowed the birefringence orientation to vary randomly while keeping the birefringence strength fixed; in the second model, we allowed both the birefringence strength and orientation to vary in accordance with a bi-Maxwellian distribution. We carried out Monte Carlo simulations of the

field evolution, and we showed in both cases that we obtain the same variation of τ_d with distance as in Poole's model.

Some time ago, Foschini and Poole [10] revisited the problem of calculating the expected differential time delay analytically for Poole's model using an approach based on the theory of stochastic differential equations. They were thus able to obtain Poole's earlier result, as well as a number of other results, in a very straightforward way. In this paper, we show that a similar approach can be applied to our physical models. We calculate the expectation of τ_d for both models, and we show that the results are consistent with our earlier simulations. The approach used by Foschini and Poole and by ourselves makes use of a master equation for the expected quantities. This equation which is simply given by Foschini and Poole with appropriate references to the mathematics literature is usually introduced with a language that physicists and engineers often find somewhat obscure [18]. This master equation can, however, be derived quite simply at the level of rigor common in the physics and engineering literature, and we include such a derivation in Appendix.

Despite yielding identical expressions for the distribution of τ_d, there is an important distinction between Poole's model and our models in which the birefringence orientation varies randomly. In Poole's model the fiber autocorrelation length h_{fiber} is infinite, where the fiber autocorrelation is the length over which an ensemble of fibers, all of which initially have the same orientation of the axes of birefringence, loses memory of this initial orientation. By contrast, in the physical models that we considered, not only is h_{fiber} finite, but simulations have previously shown that it equals $h_{E,\text{local}}$, the length scale over which the field measured with respect to the local axes of birefringence, loses memory of its orientation with respect to those axes [14]. In this paper, we will obtain this result analytically for our first physical model in which the birefringence strength is fixed. In work by ourselves [19] and others [20], it has been shown that when the coupled nonlinear Schrödinger equation that describes pulse propagation in optical fibers for both NRZ (nonreturn-to-zero) and soliton communication signals is averaged over the length scale of the large but rapidly and randomly varying birefringence, Manakov's equation results. The physical meaning of the averaged fields has been somewhat obscure. In this article, we discuss the meaning of the averaged fields. Furthermore, in the simulations to date, it has been the practice to simply scramble the fields at fixed intervals, while treating their evolution inside the intervals deterministically. We show that this practice is not strictly correct in part because there are two scrambling lengths. The length over which the linear terms average is given by the fiber autocorrelation length which is also the polarization decorrelation length measured with respect to the local axes of birefringence. This length also roughly equals the diffusion length in the equatorial direction along the Poincaré sphere measured with respect to the local axes of birefringence. The length over which the nonlinear terms average is given by the diffusion length in the azimuthal direction the Poincaré sphere. We previously defined the polarization decorrelation length h_E so that, given an ensemble of fibers whose birefringence is randomly varying and Stokes parameters whose initial values are set so that $(S_1, S_2, S_3) = (1, 0, 0)$, h_E is the length over which $\langle S_1 \rangle$ falls to $1/e$ [14]. We also previously defined the polarization diffusion lengths d_i as the lengths over which the variances of the S_i rise to $1/e$ of their asymptotic value of $1/3$ [21]. Physically, the polarization decorrelation length and the first two diffusion lengths d_1 and d_2 are related to the length scale in which the field loses memory of its polarization orientation, while d_3 is related to the length scale on which the field loses memory of the ratio of the major and minor axes of the polarization ellipse. On the Poincaré sphere, d_1 and d_2 correspond to equatorial diffusion, while d_3 corresponds to azimuthal diffusion.

The remainder of this paper is organized as follows. In Section II, we present the two physical models that we will study. We use the master equation that is derived in Appendix to calculate $\langle \tau_d^2 \rangle$ the expectation of τ_d^2 for both models. In Section III, we calculate the polarization decorrelation length and the diffusion lengths with respect to both the local polarization eigenaxes and the initial eigenaxes for the first physical model in which the birefringence strength is fixed. We show that $h_{E,\text{local}} = h_{\text{fiber}}$. In Section IV, we derive the equation of motion for nonlinear pulse propagation in optical fibers. Section V contains the conclusions.

II. Differential Delay Time

In the plane wave approximation, which is an excellent approximation for optical fibers, the light evolution can be described by two complex field amplitudes $E_1(z)$ and $E_2(z)$, where z is distance along the optical fiber. (See, for example, the discussion in [13].) Writing $\mathbf{E} = (E_1, E_2)^t$, we find that the evolution $\mathbf{E}$ is governed by the equation

$$\frac{d\mathbf{E}}{dz} = i\mathrm{K}\mathbf{E} \tag{1}$$

where defining the standard Pauli matrices

$$\begin{aligned} \mathrm{I} &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \\ \sigma_2 &= \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \end{aligned} \tag{2}$$

we may write

$$\mathrm{K} = k_0\mathrm{I} + \kappa_1\sigma_1 + \kappa_2\sigma_2 + \kappa_3\sigma_3. \tag{3}$$

Assuming that there is no polarization-dependent loss, then k_0 may be complex, but all the κ_i must be real. We eliminate k_0 from (1) by making the transformation

$$\mathbf{A} = \mathbf{E}\exp\left[-i\int_0^z k_0(z')\,dz'\right] \tag{4}$$

and we note that communication fibers are nearly linearly birefringent so that we may set $\kappa_2 = 0$. We now rewrite (1) in two alternative forms:

$$\frac{d\mathbf{A}}{dz} = i\big(b\cos\theta\sigma_3 + b\sin\theta\sigma_1\big)\mathbf{A} \tag{5a}$$

$$= i\big(x\sigma_3 + y\sigma_1\big)\mathbf{A} \tag{5b}$$

the first of which will be useful for the model in which the birefringence strength is fixed and the orientation is allowed to vary and the second of which will be useful for the model in which both vary. The quantity $\triangle = 2b$ is the birefringence strength and θ is the orientation angle.

In the first model, we assume that the rate of change of the birefringence orientation is driven by a white noise process $g_\theta(z)$ [14],

$$\frac{d\theta}{dz} = g_\theta(z) \tag{6}$$

where

$$\langle g_\theta(z)\rangle = 0, \qquad \langle g_\theta(z) g_\theta(z')\rangle = \sigma_\theta^2 \delta(z - z'). \tag{7}$$

To determine the relationship between σ_θ^2 and the fiber autocorrelation length h_{fiber}, we calculate the fiber autocorrelation function $\langle \cos[\theta(z) - \theta(0)]\rangle$, and we find

$$\begin{aligned}
&\langle \cos[\theta(z) - \theta(0)]\rangle \\
&\quad = \left\langle \cos\left[\int_0^z dz'\, g_\theta(z')\right]\right\rangle \\
&\quad = 1 - \left\langle \frac{1}{2}\int_0^z dz' \int_0^z dz''\, g_\theta(z') g_\theta(z'')\right\rangle + \cdots \\
&\quad = \exp\left(-\frac{\sigma_\theta^2}{2} z\right)
\end{aligned} \tag{8}$$

which implies that $\sigma_\theta^2 = 2/h_{\text{fiber}}$. In the second model, we assume that x and y are independent Langevin processes

$$\begin{aligned}
\frac{dx}{dz} &= -\alpha x + g_x(z), \\
\frac{dy}{dz} &= -\alpha y + g_y(z)
\end{aligned} \tag{9}$$

where

$$\begin{aligned}
&\langle g_x(z)\rangle = \langle g_y(z)\rangle = 0, \qquad \langle g_x(z) g_y(z')\rangle = 0, \\
&\langle g_x(z) g_x(z')\rangle = \langle g_y(z) g_y(z')\rangle = \sigma^2 \delta(z - z').
\end{aligned} \tag{10}$$

To relate α and σ^2 to the fiber parameters, we first note that (9) integrates to

$$\begin{aligned}
x(z) &= x(0)\exp(-\alpha z) + \int_0^z dz'\, g_x(z') \exp[-\alpha(z - z')], \\
y(z) &= y(0)\exp(-\alpha z) + \int_0^z dz'\, g_y(z') \exp[-\alpha(z - z')].
\end{aligned} \tag{11}$$

Calculating the fiber autocorrelation function $\langle x(0)x(z) + y(0)y(z)\rangle = [x^2(0) + y^2(0)] \times \exp(-\alpha z)$, we conclude that $\alpha = 1/h_{\text{fiber}}$. Demanding that $\langle x^2(z) + y^2(z)\rangle \to \langle b^2\rangle$ as $z \to \infty$ where $\langle b^2\rangle$ is the expected square birefringence strength, we find that $\sigma^2 = \langle b^2\rangle / h_{\text{fiber}}$. From a physical standpoint, it is not required that the δ-correlated variables really change instantaneously, it is only required that they change rapidly compared to the beat length which is the next smallest length scale in the problem. Simulations show that as long as the variables that we consider here to be δ-correlated change on a length scale that is small compared to the beat length, the results do not depend on the exact length scale [14].

It is now useful to switch to the Poincaré representation of the field in which, rather than following the evolution of the fields $A_1(z)$ and $A_2(z)$, one follows the evolution of the three Stokes parameters $S_1 = A_1 A_1^* - A_2 A_2^*$, $S_2 = A_1 A_2^* + A_2 A_1^*$, and $S_3 = -i(A_1 A_2^* - A_2 A_1^*)$. These variables are the natural ones to use when calculating the polarization decorrelation length, the diffusion lengths, and the differential delay time [14], [21]. The equations of motion (5) imply that the Stokes parameters obey the equation,

$$\frac{\partial \mathbf{S}}{\partial z} = \mathbf{W}(z,\omega) \times \mathbf{S} \tag{12}$$

where ω is the frequency, $\mathbf{S}(z,\omega) = (S_1, S_2, S_3)^t(z,\omega)$ is the Stokes vector, and $\mathbf{W}(z,\omega) = (2x, 2y, 0)^t$ represents the local birefringence. The strength of the birefringence is $|\mathbf{W}| = \triangle$ [14]. Equation (12) describes the spatial evolution of the Stokes vector at a fixed frequency. In order to calculate the differential delay time, we must first find the frequency evolution of the Stokes vector at a fixed position. The corresponding equation is

$$\frac{\partial \mathbf{S}}{\partial \omega} = \mathbf{\Omega}(z,\omega) \times \mathbf{S} \tag{13}$$

where $\mathbf{\Omega}(z,\omega)$ is the dispersion vector. Equation (13) is the defining relation for $\mathbf{\Omega}(z,\omega)$. The magnitude of the dispersion vector is the average differential time delay, i.e., [8]

$$\langle \tau_d^2\rangle = \langle \Omega_1^2 + \Omega_2^2 + \Omega_3^2\rangle. \tag{14}$$

From (12) and (13), we obtain the dynamical equation for the dispersion vector $\mathbf{\Omega}(z,\omega)$ [8], [10]

$$\frac{\partial \mathbf{\Omega}(z,\omega)}{\partial z} = \frac{\partial \mathbf{W}(z,\omega)}{\partial \omega} + \mathbf{W}(z,\omega) \times \mathbf{\Omega}(z,\omega). \tag{15}$$

We can determine $\langle \tau_d^2\rangle$ from (14) and (15) directly, but the calculation is simpler if we choose the frame of reference in which the birefringence vector is in the Ω_1-direction. For the first model, we define $\tilde{\mathbf{\Omega}} = R(z)\mathbf{\Omega}$ where

$$\mathrm{R}(z) = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{16}$$

With these definitions, we obtain

$$\frac{\partial}{\partial z}\begin{pmatrix} \tilde{\Omega}_1 \\ \tilde{\Omega}_2 \\ \tilde{\Omega}_3 \end{pmatrix} = \begin{pmatrix} \tilde{\Omega}_2 \\ -\tilde{\Omega}_1 \\ 0 \end{pmatrix} g_\theta + \begin{pmatrix} 2b' \\ -2b\tilde{\Omega}_3 \\ 2b\tilde{\Omega}_2 \end{pmatrix} \tag{17}$$

where $b' = \partial b/\partial\omega$. Since the magnitude of a vector is invariant under a rotation, so that $|\tilde{\mathbf{\Omega}}| = |\mathbf{\Omega}|$, we may use (17) instead of (15) to calculate $\langle \tau_d^2\rangle$. Equation (17) is what is referred to in the mathematic literatures as a stochastic differential equation because g_θ is randomly varying [18]. According to the mathematical theory of these equations, any sufficiently smooth, real function $\psi(\tilde{\mathbf{\Omega}})$ of the process $\tilde{\mathbf{\Omega}}$ obeys the equation

$$\frac{\partial\langle\psi\rangle}{\partial z} = \langle G(\psi)\rangle \tag{18}$$

where the so-called generator G is a second order differential operator. We give a simple derivation of this result in the

Appendix, and we show that G is given by

$$G = \frac{1}{2}\sigma_\theta^2\Big\{\tilde{\Omega}_2^2\frac{\partial^2}{\partial\tilde{\Omega}_1^2} + \tilde{\Omega}_1^2\frac{\partial^2}{\partial\tilde{\Omega}_2^2} - 2\tilde{\Omega}_1\tilde{\Omega}_2\frac{\partial^2}{\partial\tilde{\Omega}_1\partial\tilde{\Omega}_2} - \tilde{\Omega}_1\frac{\partial}{\partial\tilde{\Omega}_1} - \tilde{\Omega}_2\frac{\partial}{\partial\tilde{\Omega}_2}\Big\} + 2b'\frac{\partial}{\partial\tilde{\Omega}_1} - 2b\tilde{\Omega}_3\frac{\partial}{\partial\tilde{\Omega}_2} + 2b\tilde{\Omega}_2\frac{\partial}{\partial\tilde{\Omega}_3} \quad (19)$$

in our cases. [See (A14) and (A16).] Applying (18) and (19) to τ_d^2 defined in (14), we obtain the following:

$$\frac{\partial\langle\tau_d^2\rangle}{\partial z} = 4b'\langle\tilde{\Omega}_1\rangle. \quad (20)$$

To determine $\langle\tilde{\Omega}_1\rangle$, we apply (18) and (19) to $\tilde{\Omega}_1$ and we find

$$\frac{\partial\langle\tilde{\Omega}_1\rangle}{\partial z} = -\frac{1}{2}\sigma_\theta^2\langle\tilde{\Omega}_1\rangle + 2b'. \quad (21)$$

Solving (20) and (21), we conclude

$$\langle\tau_d^2(z)\rangle = 2h_{\text{fiber}}^2\triangle'^2[\exp(-z/h_{\text{fiber}}) + z/h_{\text{fiber}} - 1] \quad (22)$$

where we have used the initial condition $\langle\tau_d^2(0)\rangle = 0$.

In the second model given by (9) and (10), the birefringence $\triangle$ is a function of both frequency and distance. We assume that the orientation of the birefringence axis θ is only a function of z and that the frequency variation of the birefringence strength is separate from the z-variation so that $b(\omega, z) = k(\omega)\tilde{b}(z)$, where $k(\omega)$ is a deterministic function of frequency. These assumptions are physically reasonable for communication signals in optical fibers. It then follows that $f(\omega, z) = k(\omega)\tilde{f}(z)$ where $f(\omega, z)$ stands for any one of the functions x, y, g_x, or g_y. We again rotate (15) using the rotation matrix $\mathrm{R}(z)$. After the rotation, we combine (9) and (15) to obtain the following matrix equation:

$$\frac{\partial}{\partial z}\begin{pmatrix}\tilde{\Omega}_1\\ \tilde{\Omega}_2\\ \tilde{\Omega}_3\\ x\\ y\end{pmatrix} = \frac{\sigma^2}{\tilde{b}^2}\begin{pmatrix}-y\tilde{\Omega}_2 & x\tilde{\Omega}_2\\ y\tilde{\Omega}_1 & -x\tilde{\Omega}_1\\ 0 & 0\\ \tilde{b}^2 & 0\\ 0 & \tilde{b}^2\end{pmatrix}\begin{pmatrix}g_x\\ g_y\end{pmatrix} + \begin{pmatrix}2k'\tilde{b}\\ -2k\tilde{b}\tilde{\Omega}_3\\ 2k\tilde{b}\tilde{\Omega}_2\\ -x/h_{\text{fiber}}\\ -y/h_{\text{fiber}}\end{pmatrix}. \quad (23)$$

Using (A14), the generator for the random process $(\tilde{\mathbf{\Omega}}, x, y)$ is given by

$$G = \frac{1}{2}\frac{\sigma^2}{\tilde{b}^2} \times \Big\{\tilde{\Omega}_2^2\frac{\partial^2}{\partial\tilde{\Omega}_1^2} + \tilde{\Omega}_1^2\frac{\partial^2}{\partial\tilde{\Omega}_2^2} + \tilde{b}^2\frac{\partial^2}{\partial x^2} + \tilde{b}^2\frac{\partial^2}{\partial y^2} - 2\tilde{\Omega}_1\tilde{\Omega}_2\frac{\partial^2}{\partial\tilde{\Omega}_1\partial\tilde{\Omega}_2} - 2y\tilde{\Omega}_2\frac{\partial^2}{\partial x\partial\tilde{\Omega}_1} + 2x\tilde{\Omega}_2\frac{\partial^2}{\partial y\partial\tilde{\Omega}_1} + 2y\tilde{\Omega}_1\frac{\partial^2}{\partial x\partial\tilde{\Omega}_1} - 2x\tilde{\Omega}_1\frac{\partial^2}{\partial y\partial\tilde{\Omega}_1} - \tilde{\Omega}_1\frac{\partial}{\partial\tilde{\Omega}_1} + 2k'\tilde{b}\frac{\partial}{\partial\tilde{\Omega}_1} - \tilde{\Omega}_2\frac{\partial}{\partial\tilde{\Omega}_2}\Big\} - 2k\tilde{b}\tilde{\Omega}_3\frac{\partial}{\partial\tilde{\Omega}_2} + 2k\tilde{b}\tilde{\Omega}_2\frac{\partial}{\partial\tilde{\Omega}_3} - \frac{x}{h_{\text{fiber}}}\frac{\partial}{\partial x} - \frac{y}{h_{\text{fiber}}}\frac{\partial}{\partial y}. \quad (24)$$

From (18) and (24), the equations determining $\langle\tau_d^2(z)\rangle$ are as follows:

$$\frac{\partial\langle\tau_d^2\rangle}{\partial z} = 4k'\langle\tilde{b}\tilde{\Omega}_1\rangle.$$
$$\frac{\partial\langle\tilde{b}\tilde{\Omega}_1\rangle}{\partial z} = -\frac{1}{h_{\text{fiber}}}\langle\tilde{b}\tilde{\Omega}_1\rangle + 2k'\langle\tilde{b}^2\rangle. \quad (25)$$

We assume that the average birefringence $\langle\triangle^2\rangle = 4k^2\langle\tilde{b}^2\rangle$ is independent of z. Solving (25), we obtain

$$\langle\tau_d^2(z)\rangle = 2h_{\text{fiber}}^2\langle\triangle'^2\rangle[\exp(-z/h_{\text{fiber}}) + z/h_{\text{fiber}} - 1]. \quad (26)$$

In both of the models presented here, the differential time delay has the same functional form found by Poole [4] and by Foschini and Poole [10] using the weak coupling model. However, their model assumes that the birefringence orientation is nearly constant so that h_{fiber} is infinite. Using simulations, we have already shown for both models that $h_{E,\text{local}}$, the polarization decorrelation length measured with respect to the local axes of birefringence, is equal to the fiber autocorrelation length h_{fiber} [14]. In Section III of this paper, we will obtain this result once again for the first model. It can also be shown that the parameters appearing in the expression for $\langle\tau_d^2(z)\rangle$ in [4] $(1/2h)$ and [10] $(1/\sigma^2)$ are equal to $h_{E,\text{local}}$ in the weak coupling model. Recently, using an approach similar to that of [4], we showed [13] that when the variation length of $d\theta/dz$ is much shorter than the beat length which in turn is much shorter than the variation length of θ, b, and the field, we again obtain expression (26) for the polarization mode dispersion.

Since the same expression for polarization mode dispersion appears using several very different physical models, we conjecture that the polarization mode dispersion of a fiber with randomly varying birefringence will always have the form given by (26). Furthermore, the characteristic length scale that appears in $\langle\tau_d^2(z)\rangle$ will be $h_{E,\text{local}}$.

III. Polarization Decorrelation and Diffusion

Because of the rapidly and randomly varying birefringence, even if the input polarization is along one of the local eigenaxes, the polarization states of the electric field at large distances will become uniformly distributed on the Poincaré sphere so that $\langle S_i\rangle \to 0$ and $\langle S_i^2\rangle \to 1/3$. In [14], we define the polarization decorrelation length h_E so that it equals the length at which $\langle S_1(z)\rangle$ falls to $1/e$ of its initial value. We start our simulations setting $(S_1, S_2, S_3)(z_0) = (1, 0, 0)$. In [21], we define the diffusion length d_i as the distance at which the variance of S_i rises to $1/e$ of its asymptotic value of $1/3$. The polarization decorrelation length and the diffusion lengths for the first model in which the birefringence is fixed can be calculated from (12) using the same master equations (18), that we used to calculate $\langle\tau_d{}^2(z)\rangle$. It is not possible to use the master equation to calculate the polarization decorrelation length and the diffusion lengths for the second model in which the birefringence strength varies because one obtains an infinite series of cumulants. However, previous results based on Monte Carlo simulations show that the two models produce nearly identical results [14], [21].

There are two physically sensible ways to choose the reference axes for the Stokes vector. We may choose them to equal the polarization eigenstates at the initial point of the fiber, or we may choose them to equal the local polarization eigenstates. As in Section II, we carry out our calculations using the local polarization eigenaxes. Using (6), (12), and (16), we find

$$\frac{\partial}{\partial z}\begin{pmatrix}\tilde{S}_1\\ \tilde{S}_2\\ \tilde{S}_3\\ \theta\end{pmatrix}=\begin{pmatrix}\tilde{S}_2\\ -\tilde{S}_1\\ 0\\ 1\end{pmatrix}g_\theta+\begin{pmatrix}0\\ -2b\tilde{S}_3\\ 2b\tilde{S}_2\\ 0\end{pmatrix} \tag{27}$$

where $\tilde{\mathbf{S}}(z,\omega)=\mathrm{R}(z)\mathbf{S}(z,\omega)$ is the Stokes vector measured with respect to local polarization eigenaxes. From (A14), we find that the generator that corresponds to the process $\{\tilde{\mathbf{S}},\theta\}$ is given by

$$\begin{aligned}G=&\frac{1}{2}\sigma_\theta^2\\ &\times\left\{\tilde{S}_2^2\frac{\partial^2}{\partial\tilde{S}_1^2}+\tilde{S}_1^2\frac{\partial^2}{\partial\tilde{S}_2^2}+\frac{\partial^2}{\partial\theta^2}-2\tilde{S}_1\tilde{S}_2\frac{\partial^2}{\partial\tilde{S}_1\partial\tilde{S}_2}\right.\\ &\left.+2\tilde{S}_2\frac{\partial^2}{\partial\theta\partial\tilde{S}_1}-2\tilde{S}_1\frac{\partial^2}{\partial\theta\partial\tilde{S}_2}-\tilde{S}_1\frac{\partial}{\partial\tilde{S}_1}-\tilde{S}_2\frac{\partial}{\partial\tilde{S}_2}\right\}\\ &-2b\tilde{S}_3\frac{\partial}{\partial\tilde{S}_2}+2b\tilde{S}_2\frac{\partial}{\partial\tilde{S}_3}.\end{aligned} \tag{28}$$

Using (18) and (28), we can calculate the ensemble averages of functions of $\{\tilde{\mathbf{S}},\theta\}$. For example, to determine the fiber autocorrelation length that we have already calculated directly in (8), we apply (18) to $x=b\cos[\theta(z)-\theta(0)]$, and we obtain

$$\frac{d}{dz}\langle x(z)\rangle=-\frac{1}{2}\sigma_\theta^2\langle x(z)\rangle. \tag{29}$$

Solving (29) yields (8).

The evolution of the ensemble-averaged Stokes parameters measured in the local reference frame $\langle\tilde{S}_i(z)\rangle$, $i=1,2,3$ is described by the following equations:

$$\begin{aligned}\frac{d}{dz}\langle\tilde{S}_1\rangle&=-\frac{1}{2}\sigma_\theta^2\langle\tilde{S}_1\rangle,\\ \frac{d}{dz}\langle\tilde{S}_2\rangle&=-\frac{1}{2}\sigma_\theta^2\langle\tilde{S}_2\rangle-2b\langle\tilde{S}_3\rangle,\\ \frac{d}{dz}\langle\tilde{S}_3\rangle&=2b\langle\tilde{S}_2\rangle.\end{aligned} \tag{30}$$

If we assume the input polarization is along one of the local eigenaxes, i.e., $(\tilde{S}_1,\tilde{S}_2,\tilde{S}_3)(z=0)=(1,0,0)$, then $\langle\tilde{S}_1(z)\rangle=\exp(-\sigma_\theta^2 z/2)$, and $\langle\tilde{S}_2(z)\rangle=\langle\tilde{S}_3(z)\rangle=0$. The polarization decorrelation length measured with respect to the local axes is therefore given by $h_{E,\mathrm{local}}=2/\sigma_\theta^2=h_{\mathrm{fiber}}$, which agrees with the results obtained using Monte Carlo simulations [14].

To determine the evolution of the ensemble-averaged Stokes parameters measured in a fixed reference frame, we can either construct a generator for (12) directly or make use of (28) and the transformation $\mathbf{S}=\mathrm{R}^{-1}\tilde{\mathbf{S}}$ which can be written explicitly as,

$$\begin{aligned}S_1&=\tilde{S}_1\cos\theta-\tilde{S}_2\sin\theta,\\ S_2&=\tilde{S}_1\sin\theta+\tilde{S}_2\cos\theta,\\ S_3&=\tilde{S}_3.\end{aligned} \tag{31}$$

Because the input is aligned with the initial polarization eigenaxes, one can show that $\langle S_2(z)\rangle=\langle S_3(z)\rangle=0$. The equations governing the evolution of $\langle S_1(z)\rangle$ are

$$\begin{aligned}\frac{d}{dz}\langle S_1\rangle&=2b\langle\tilde{S}_3\sin\theta\rangle,\\ \frac{d}{dz}\langle\tilde{S}_2\sin\theta\rangle&=-\sigma_\theta^2(\langle S_1\rangle+2\langle\tilde{S}_2\sin\theta\rangle)-2b\langle\tilde{S}_3\sin\theta\rangle,\\ \frac{d}{dz}\langle\tilde{S}_3\sin\theta\rangle&=2b\langle\tilde{S}_2\sin\theta\rangle-\frac{1}{2}\sigma_\theta^2\langle\tilde{S}_3\sin\theta\rangle.\end{aligned} \tag{32}$$

Equation (32) is a third-order ordinary differential equation. It can be solved and analytically the results agree with the simulations results reported in [14]. The explicit analytical expressions for the solutions of (32) are not particularly illustrative, but simple expressions for the polarization decorrelation length can be obtained when the fiber autocorrelation length is much shorter than the beat length L_B or when the fiber autocorrelation length is much longer than the beat length. When $h_{\mathrm{fiber}}\ll L_B$, one can show that

$$\langle S_1(z)\rangle=\exp\left(-4\frac{b^2}{\sigma_\theta^2}z\right). \tag{33}$$

The polarization decorrelation length measured with respect to the fixed axes $h_{E,\mathrm{fixed}}=L_B^2/2\pi^2 h_{\mathrm{fiber}}$ is inversely proportional to the h_{fiber}, which agrees with our simulations [14] and the results in [22]. In the other limit when $h_{\mathrm{fiber}}\gg L_B$, we find

$$\langle S_1(z)\rangle=\exp(-\sigma_\theta^2 z) \tag{34}$$

and $h_{E,\mathrm{fixed}}=h_{\mathrm{fiber}}/2=h_{E,\mathrm{local}}/2$. The polarization decorrelation length with respect to the fixed axes is only half the polarization decorrelation length with respect to the local axes which again is observed in simulations [14].

Next, we determine the evolution of the variances of the Stokes parameters. Using (18) and (28), $\langle\tilde{S}_i^2(z)\rangle$, $i=1,2,3$, are given by the following set of equations:

$$\begin{aligned}\frac{d}{dz}\langle\tilde{S}_1^2\rangle&=-\sigma_\theta^2\big(\langle\tilde{S}_1^2\rangle-\langle\tilde{S}_2^2\rangle\big),\\ \frac{d}{dz}\langle\tilde{S}_2^2\rangle&=\sigma_\theta^2\big(\langle\tilde{S}_1^2\rangle-\langle\tilde{S}_2^2\rangle\big)-4b\langle\tilde{S}_2\tilde{S}_3\rangle,\\ \frac{d}{dz}\langle\tilde{S}_3^2\rangle&=4b\langle\tilde{S}_2\tilde{S}_3\rangle,\\ \frac{d}{dz}\langle\tilde{S}_2\tilde{S}_3\rangle&=2b\big(\langle\tilde{S}_2^2\rangle-\langle\tilde{S}_3^2\rangle\big)-\frac{1}{2}\sigma_\theta^2\tilde{S}_2\tilde{S}_3\end{aligned} \tag{35}$$

where we recall that $\langle\tilde{S}_1^2+\tilde{S}_2^2+\tilde{S}_3^2\rangle=1$. Equation (35) is easily integrated, and the results are consistent with what has been earlier obtained from Monte Carlo simulations [21]. In the limit when the fiber autocorrelation length is much shorter than the beat length, $h_{\mathrm{fiber}}\ll L_B$, we obtain the following:

$$\begin{aligned}\langle\tilde{S}_1^2\rangle&\approx\frac{1}{6}\left[2+\exp\left(-24\frac{b^2}{\sigma_\theta^2}z\right)+3\exp\left(-2\sigma_\theta^2 z\right)\right],\\ \langle\tilde{S}_2^2\rangle&\approx\frac{1}{6}\left[2+\exp\left(-24\frac{b^2}{\sigma_\theta^2}z\right)-3\exp\left(-2\sigma_\theta^2 z\right)\right],\\ \langle\tilde{S}_3^2\rangle&\approx\frac{1}{3}\left[1-\exp\left(-24\frac{b^2}{\sigma_\theta^2}z\right)\right].\end{aligned} \tag{36}$$

From (36), we find that the evolution of $\langle \tilde{S}_1^2 \rangle$ and $\langle \tilde{S}_2^2 \rangle$ depends on two very different length scales, $h_{\text{fiber}}/4$ and $L_B^2/12\pi^2 h_{\text{fiber}}$. When $h_{\text{fiber}}/4 \ll z \ll L_B^2/12\pi^2 h_{\text{fiber}}$, $\langle \tilde{S}_1^2 \rangle = \langle \tilde{S}_2^2 \rangle \approx 1/2$ and $\langle \tilde{S}_3^2 \rangle \approx 0$. Physically, the electric field in the local axes becomes uniformly distributed on the equator of the Poincaré sphere because of the rapid fluctuation of the fiber birefringence, while remaining confined to the equator. When $z \gg L_B^2/12\pi^2 h_{\text{fiber}}$, $\langle \tilde{S}_i^2 \rangle = 1/3$, $i = 1, 2, 3$, which means that the electric field is uniformly distributed on the surface of the Poincaré sphere. The distance at which the electric field is equatorially randomized is much shorter than in the azimuthal direction. When the fiber autocorrelation length is much longer than the beat length, $h_{\text{fiber}} \gg L_B$, we find

$$\begin{aligned} \langle \tilde{S}_1^2 \rangle &\approx \frac{1}{3}\left[1 + 2\exp\left(-\frac{3}{2}\sigma_\theta^2 z\right)\right], \\ \langle \tilde{S}_2^2 \rangle &\approx \frac{1}{3}\left[1 - \exp\left(-\frac{3}{2}\sigma_\theta^2 z\right)\right], \\ \langle \tilde{S}_3^2 \rangle &\approx \frac{1}{3}\left[1 - \exp\left(-\frac{3}{2}\sigma_\theta^2 z\right)\right]. \end{aligned} \tag{37}$$

Both the equatorial and azimuthal diffusion lengths are the same, so that the randomization of the field on the Poincaré sphere is uniform.

The evolution of the variance of the Stokes parameters in the fixed frame can be determined using the following relations:

$$\begin{aligned} S_1^2 &= \frac{1}{2}\left[1 - \tilde{S}_3^2 + \left(\tilde{S}_1^2 - \tilde{S}_2^2\right)\cos 2\theta - 2\tilde{S}_1\tilde{S}_2 \sin 2\theta\right], \\ S_2^2 &= \frac{1}{2}\left[1 - \tilde{S}_3^2 - \left(\tilde{S}_1^2 - \tilde{S}_2^2\right)\cos 2\theta + 2\tilde{S}_1\tilde{S}_2 \sin 2\theta\right]. \end{aligned} \tag{38}$$

The equation governing $\langle \tilde{S}_1^2 \cos 2\theta \rangle$, $\langle \tilde{S}_2^2 \cos 2\theta \rangle$, and $\langle \tilde{S}_1\tilde{S}_2 \sin 2\theta \rangle$ are

$$\begin{aligned} \frac{d}{dz}\langle \tilde{S}_1^2 \cos 2\theta \rangle &= \sigma_\theta^2\big(-3\langle \tilde{S}_1^2 \cos 2\theta \rangle + \langle \tilde{S}_2^2 \cos 2\theta \rangle \\ &\qquad - 4\langle \tilde{S}_1\tilde{S}_2 \sin 2\theta \rangle\big), \\ \frac{d}{dz}\langle \tilde{S}_2^2 \cos 2\theta \rangle &= \sigma_\theta^2\big(\langle \tilde{S}_1^2 \cos 2\theta \rangle - 3\langle \tilde{S}_2^2 \cos 2\theta \rangle \\ &\qquad + 4\langle \tilde{S}_1\tilde{S}_2 \sin 2\theta \rangle\big) \\ &\qquad - 4b\langle \tilde{S}_2\tilde{S}_3 \cos 2\theta \rangle, \\ \frac{d}{dz}\langle \tilde{S}_1\tilde{S}_2 \sin 2\theta \rangle &= \sigma_\theta^2\big(-2\langle \tilde{S}_1^2 \cos 2\theta \rangle + 2\langle \tilde{S}_2^2 \cos 2\theta \rangle \\ &\qquad - 4\langle \tilde{S}_1\tilde{S}_2 \sin 2\theta \rangle\big) \\ &\qquad - 2b\langle \tilde{S}_1\tilde{S}_3 \sin 2\theta \rangle, \\ \frac{d}{dz}\langle \tilde{S}_1\tilde{S}_3 \sin 2\theta \rangle &= 2b\langle \tilde{S}_1\tilde{S}_2 \sin 2\theta \rangle - \sigma_\theta^2 \\ &\qquad \times \left(\frac{5}{2}\langle \tilde{S}_1\tilde{S}_3 \sin 2\theta \rangle - 2\langle \tilde{S}_2\tilde{S}_3 \cos 2\theta \rangle\right), \\ \frac{d}{dz}\langle \tilde{S}_2\tilde{S}_3 \cos 2\theta \rangle &= 2b\big(\langle \tilde{S}_1^2 \cos 2\theta \rangle + 2\langle \tilde{S}_2^2 \cos 2\theta \rangle\big) \\ &\qquad + \sigma_\theta^2\left(2\langle \tilde{S}_1\tilde{S}_3 \sin 2\theta \rangle - \frac{5}{2}\langle \tilde{S}_2\tilde{S}_3 \cos 2\theta \rangle\right) \\ &\qquad - 2b\langle \cos 2\theta \rangle. \end{aligned} \tag{39}$$

When the fiber autocorrelation length is much shorter than the beat length, $h_{\text{fiber}} \ll L_B$, we obtain,

$$\begin{aligned} \langle S_1^2 \rangle &\approx \frac{1}{6}\left[2 + 3\exp\left(-8\frac{b^2}{\sigma_\theta^2}z\right) + \exp\left(-24\frac{b^2}{\sigma_\theta^2}z\right)\right], \\ \langle S_2^2 \rangle &\approx \frac{1}{6}\left[2 - 3\exp\left(-8\frac{b^2}{\sigma_\theta^2}z\right) + \exp\left(-24\frac{b^2}{\sigma_\theta^2}z\right)\right]. \end{aligned} \tag{40}$$

Since $S_3 = \tilde{S}_3$, $\langle S_3{}^2 \rangle = \langle \tilde{S}_3^2 \rangle$. From (36) and (40), the decorrelation lengths for $\langle S_1^2 \rangle$ are $\langle S_2^2 \rangle$ are similar but differ from that of $\langle S_3^2 \rangle$. However, they are all of the same order of magnitude, $\simeq L_B^2/\pi^2 h_{\text{fiber}}$.

When the fiber autocorrelation length is much longer than the beat length, $h_{\text{fiber}} \gg L_B$, we obtain

$$\begin{aligned} \langle S_1^2 \rangle &\approx \frac{1}{6}\left[2 + 3\exp\left(-\frac{7}{2}\sigma_\theta^2 z\right) + \exp\left(-\frac{3}{2}\sigma_\theta^2 z\right)\right], \\ \langle S_2^2 \rangle &\approx \frac{1}{6}\left[2 - 3\exp\left(-\frac{7}{2}\sigma_\theta^2 z\right) + \exp\left(-\frac{3}{2}\sigma_\theta^2 z\right)\right]. \end{aligned} \tag{41}$$

The diffusion lengths in the equatorial and azimuthal directions are also somewhat different in this limit.

IV. Implications for Nonlinear Evolution

In this section, we determine the effect of the rapidly varying birefringence on nonlinear pulse propagation in optical fibers. Using the first model we show analytically that the scrambling length over which the linear terms average is the fiber autocorrelation length, but the scrambling length over which the nonlinear terms average is the diffusion length in the azimuthal direction on the Poincaré sphere, d_3. The equation describing the evolution in an optical fiber of the electric field $\mathbf{A}$ defined in (4) is [23],

$$\begin{aligned} i\frac{\partial \mathbf{A}}{\partial z} + b\Sigma\mathbf{A} + ib'\Sigma\frac{\partial \mathbf{A}}{\partial t} \pm \frac{1}{2}\frac{\partial^2 \mathbf{A}}{\partial t^2} + \tfrac{5}{6}|\mathbf{A}|^2\mathbf{A} \\ + \tfrac{1}{6}(\mathbf{A}^\dagger\sigma_3\mathbf{A})\sigma_3\mathbf{A} + \tfrac{1}{3}\mathbf{B} = 0 \end{aligned} \tag{42}$$

where $\mathbf{A}^\dagger = (A_1^*, A_2^*)$, $\mathbf{B} = (A_1^* A_2^2, A_1^2 A_2^*)^t$, and $\Sigma = \cos\theta\sigma_3 + \sin\theta\sigma_1$. The second term and third term on the left-hand side of (42) are the phase-velocity and group-velocity birefringence. The fourth term describes the chromatic dispersion. The plus sign corresponds to anomalous dispersion and the minus sign corresponds to normal dispersion. The last three terms are due to the Kerr nonlinearity. We transform (42) to the local axes of birefringence with the following transformation:

$$\Psi = \begin{pmatrix} U \\ V \end{pmatrix} = \begin{pmatrix} \cos\theta/2 & \sin\theta/2 \\ -\sin\theta/2 & \cos\theta/2 \end{pmatrix}\begin{pmatrix} A_1 \\ A_2 \end{pmatrix}. \tag{43}$$

Equation (42) becomes

$$\begin{aligned} i\frac{\partial \Psi}{\partial z} + \tilde{\Sigma}\Psi + ib'\sigma_3\frac{\partial \Psi}{\partial t} \pm \frac{1}{2}\frac{\partial^2 \Psi}{\partial t^2} + \tfrac{5}{6}|\Psi|^2\Psi \\ + \tfrac{1}{6}(\Psi^\dagger\sigma_3\Psi)\sigma_3\Psi + \tfrac{1}{3}\mathbf{N} = 0 \end{aligned} \tag{44}$$

where $\mathbf{N} = (U^*V^2, U^2V^*)^t$, and

$$\tilde{\Sigma} = \begin{pmatrix} b & -i\theta_z/2 \\ i\theta_z/2 & -b \end{pmatrix}. \tag{45}$$

The nonlinear terms in (42) are invariant with respect to rotation [23]. If the orientation angle θ is independent of

the distance z, the matrix $\tilde{\Sigma}$ is diagonal and the fields U and V only couple nonlinearly. Equations (42) and (44) are written in dimensionless form. The distance is normalized to the dispersion length scale l_d which is typically hundreds of kilometers. Since the orientation of the birefringence is randomly varying on a length scale of the order of 100 m, the ratio $\epsilon = h_{\text{fiber}}/l_d$ is very small. With this normalization, the phase birefringence Σ is typically very large, of the order ϵ^{-1}. The short fiber autocorrelation length, however, means that Σ changes many times in a distance l_d. If we consider a weak cw wave so that all but the first two terms in (44) can be ignored, we find that the effect of the randomly varying birefringence is the linear randomization of the electric field that we have already discussed in Section III. We can remove the rapid variation of the state of polarization of the electric field by the following transformation:

$$\Psi(z,t) = \mathrm{T}(z)\bar{\Psi}(z,t) \tag{46}$$

where

$$\mathrm{T}(z) \equiv \begin{pmatrix} u_1 & u_2 \\ -u_2^* & u_1^* \end{pmatrix} \tag{47}$$

is a unitary matrix with coefficients u_1 and u_2, so that $|u_1|^2 + |u_2|^2 = 1$ and $\mathrm{T}(z)$ satisfies the following equation:

$$i\frac{\partial \mathrm{T}}{\partial z} + \tilde{\Sigma}\mathrm{T} = 0. \tag{48}$$

Substitution of (46) and (47) into (44), we obtain

$$i\frac{\partial \bar{\Psi}}{\partial z} + ib'\bar{\sigma}\frac{\partial \bar{\Psi}}{\partial t} \pm \frac{1}{2}\frac{\partial^2 \bar{\Psi}}{\partial t^2} + \frac{5}{6}|\bar{\Psi}|^2\bar{\Psi} + \tfrac{1}{6}(\bar{\Psi}^\dagger\sigma_3\overline{\Psi})\sigma_3\bar{\Psi} + \tfrac{1}{3}\widehat{\mathbf{N}} = 0 \tag{49}$$

where letting $\bar{\Psi}(z,t) = (\bar{U},\bar{V})^t(z,t)$,

$$\bar{\sigma} = \mathrm{T}^\dagger\sigma_3\mathrm{T} = \begin{pmatrix} a_1 & a_4^* \\ a_4 & -a_1 \end{pmatrix} \tag{50}$$

and $\widehat{\mathbf{N}} = (\widehat{\mathbf{N}}_1, \widehat{\mathbf{N}}_2)^t$ where

$$\begin{aligned} \widehat{\mathbf{N}}_1 &= a_3^2(2|\bar{V}|^2 - |\bar{U}|^2)\bar{U} - a_3a_6^*(2|\bar{U}|^2 - |\bar{V}|^2)\bar{V} \\ &\quad - a_3a_6\bar{U}^2\bar{V}^* - a_6^{*2}\bar{V}^2\bar{U}^*, \\ \widehat{\mathbf{N}}_2 &= a_3^2(2|\bar{U}|^2 - |\bar{V}|^2)\bar{V} + a_3a_6(2|\bar{V}|^2 - |\bar{U}|^2)\bar{U} \\ &\quad + a_3a_6^*\bar{V}^2\bar{U}^* - a_6^2\bar{U}^2\bar{V}^*. \end{aligned} \tag{51}$$

The coefficients a_i, $i = 1,\cdots,6$ are defined in terms of u_1 and u_2 as

$$\begin{aligned} a_1 &= |u_1|^2 - |u_2|^2, & a_4 &= 2u_1u_2^*, \\ a_2 &= -(u_1u_2 + u_1^*u_2^*), & a_5 &= u_1^2 - {u_2^*}^2, \\ a_3 &= i(u_1u_2 - u_1^*u_2^*), & a_6 &= -i(u_1^2 + {u_2^*}^2). \end{aligned} \tag{52}$$

From the definition, the coefficients a_1, a_2, and a_3 are real while a_4, a_5, and a_6 are complex. One can further show that

$$\begin{aligned} a_1^2 + a_2^2 + a_3^2 &= 1, \\ a_4^2 + a_5^2 + a_6^2 &= 0. \end{aligned} \tag{53}$$

The coefficients $\{a_1, a_2, a_3\}$ satisfy (27), the equation of motion for the Stokes parameters in the local axes if one replaces $\tilde{S}_i$ by a_i, $i = 1,2,3$. Similarly, the coefficients $\{a_4, a_5, a_6\}$ also satisfy (27) with the same white noise source g_θ. Since the coefficients of (27) are real, the real parts and the imaginary parts of $\{a_4, a_5, a_6\}$ satisfy (27) separately.

The random coefficients a_i in (49) are rapidly varying. We separate the spatial average of the coefficients in (49) from its randomly varying part, and we obtain

$$\begin{aligned} &i\frac{\partial \bar{\Psi}}{\partial z} + ib'\langle\bar{\sigma}\rangle\frac{\partial \bar{\Psi}}{\partial t} \pm \frac{1}{2}\frac{\partial^2 \bar{\Psi}}{\partial t^2} \\ &\quad + \frac{5}{6}|\bar{\Psi}|^2\bar{\Psi} + \frac{1}{6}(\bar{\Psi}^\dagger\sigma_3\bar{\Psi})\sigma_3\bar{\Psi} + \frac{1}{3}\langle\widehat{\mathbf{N}}\rangle \\ &\quad = -ib'(\bar{\sigma} - \langle\bar{\sigma}\rangle)\frac{\partial \bar{\Psi}}{\partial t} - \frac{1}{3}(\widehat{\mathbf{N}} - \langle\widehat{\mathbf{N}}\rangle). \end{aligned} \tag{54}$$

We have used the ergodic theorem [13],

$$\langle f\rangle = \lim_{z\to\infty}\frac{1}{z}\int_0^z ds f(s) \tag{55}$$

which states that as $z \to \infty$, spatial averages of a function $f(z)$ are equivalent to ensemble averages. The notation $\langle\widehat{\mathbf{N}}\rangle$ represents replacing the coefficients a_ia_j in $\widehat{\mathbf{N}}$ by their ensemble average. Since $\{a_1, a_2, a_3\}$ and $\{a_4, a_5, a_6\}$ are solutions of (27), the equations governing their means and their variances have been derived in (30) and (35) of Section III. To determine $\langle a_3a_6\rangle$, we construct the generator for the process $\{a_i\}$, $i = 1, \cdots, 6$, and we find that $\langle a_3a_6\rangle$ also satisfies (35) if we use the following substitution:

$$\begin{aligned} \tilde{S}_1^2 &\to a_1a_4, \\ \tilde{S}_2^2 &\to a_2a_5, \\ \tilde{S}_3^2 &\to a_3a_6, \\ \tilde{S}_2\tilde{S}_3 &\to (a_2a_6 + a_3a_5)/2. \end{aligned}$$

From the definition of the transfer matrix T, we find that at $z = 0$, we have $u_1 = 1$ and $u_2 = 0$. Equivalently we have $(a_1, a_2, a_3) = (1, 0, 0)$ and $(a_4, a_5, a_6) = (0, 1, -i)$. From (3), we find that $\langle a_1\rangle = \exp(-\sigma_\theta^2 z/2)$ and $\langle a_4\rangle = 0$. Since (35) is homogeneous and initially $(a_1a_4, a_2a_5, a_3a_6, a_2a_6 + a_3a_5) = (0,0,0,0)$, we conclude that $\langle a_3a_6\rangle = 0$. Similarly, one can show that $\langle \mathrm{Re}(a_i)\,\mathrm{Im}(a_i)\rangle = 0$, $i = 4,5,6$. The ensemble averages of $\bar{\sigma}$ and $\widehat{\mathbf{N}}$ are

$$\begin{aligned} \langle\bar{\sigma}\rangle &= \lim_{z\to\infty}\begin{pmatrix} \exp(-\sigma_\theta^2 z/2) & 0, \\ 0 & \exp(-\sigma_\theta^2 z/2) \end{pmatrix} = 0 \\ \langle\widehat{\mathbf{N}}\rangle &= \lim_{z\to\infty}\begin{pmatrix} -\langle a_3^2\rangle(|\bar{U}|^2 - 2|\bar{V}|^2)\bar{U} - \langle \mathrm{Re}(a_6^{*2})\rangle\bar{U}^*\bar{V}^2 \\ \langle a_3^2\rangle(2|\bar{U}|^2 - |\bar{V}|^2)\bar{V} - \langle \mathrm{Re}(a_6^2)\rangle\bar{U}^2\bar{V}^* \end{pmatrix}. \end{aligned} \tag{56}$$

To calculate $\langle a_3^2\rangle$ and $\langle \mathrm{Re}(a_6^2)\rangle$, we first note that $\langle \mathrm{Re}(a_6^2)\rangle = \langle \mathrm{Re}^2(a_6)\rangle - \langle \mathrm{Im}^2(a_6)\rangle$ and that a_3^2, $\mathrm{Re}^2(a_6)$, and $\mathrm{Im}^2(a_6)$ can be obtained by solving (35) with the initial conditions $(a_1^2, a_2^2, a_3^2) = (1,0,0)$, $[\mathrm{Re}^2(a_4), \mathrm{Re}^2(a_5), \mathrm{Re}^2(a_6)] = (0,1,0)$, and $[\mathrm{Im}^2(a_4), \mathrm{Im}^2(a_5), \mathrm{Im}^2(a_6)] = (0,0,1)$. We conclude that $\langle a_3^2\rangle = 1/3$ and $\langle \mathrm{Re}(a_6^2)\rangle = 0$. Using these results, we find that (54) becomes

$$i\frac{\partial \bar{\Psi}}{\partial z} \pm \frac{1}{2}\frac{\partial^2 \bar{\Psi}}{\partial t^2} + \frac{8}{9}|\overline{\Psi}|^2\bar{\Psi} = -ib'(\bar{\sigma} - \langle\bar{\sigma}\rangle)\frac{\partial \bar{\Psi}}{\partial t} - \frac{1}{3}(\widehat{\mathbf{N}} - \langle\widehat{\mathbf{N}}\rangle). \tag{57}$$

The left-hand side of (57) is known as the Manakov equation. The first term on the right-hand side of (57) describes the usual

linear polarization mode dispersion, and the second term on the right-hand side describes a nonlinear polarization mode dispersion that has not to our knowledge been previously discussed in the literature. The coefficients on the right-hand side of (57) have zero mean when averaged over the Poincaré sphere. The magnitude of the random coefficients a_1, a_4, a_3^2, a_3a_6, and a_6^2 are of order one and they change sign on a length scale given by h_{fiber} which is much shorter than the dispersion length scale l_d. Physically, the pulse envelope is not able to respond to the rapidly varying a_i but rather to their cumulative effects. One can therefore average (57) over an intermediate length scale l which is much longer than h_{fiber} but is much shorter than the chromatic dispersion scale length, the nonlinear scale length, and the polarization mode dispersion scale length. Since the electric field $\bar{\Psi}(z,t)$ does not change significantly on the length scale l, the effect of this averaging is to replace the random coefficients a_i and a_ia_j on the right-hand side of (57) by their integrals $(1/l)\int_z^{z+l} dz' a_i(z')$ and $(1/l)\int_z^{z+l} dz' a_ia_j$.

The variance of these random coefficients is typically proportional to h/l, where h is the decorrelation length of a_i or a_ia_j. Since the decorrelation lengths of a_1 and a_4 approximately equal the equatorial diffusion length d_1, while the decorrelation lengths of a_3^2, a_3a_6, and a_6^2 approximately equal the azimuthal diffusion length d_3, the relative contribution from the linear and the nonlinear polarization mode dispersion in (57) to the pulse deformation will depend on the relative strength of d_1 and d_3. From Section III, in the limit where the fiber autocorrelation length is much shorter than the beat length, $h_{\text{fiber}} \ll L_B$, the azimuthal diffusion length is much longer than the equatorial diffusion length, and the contribution from the nonlinear polarization mode dispersion will dominate.

We derive the equations of motion for the averaged random coefficients on the right-hand side of (57) by constructing the corresponding generators, and we then solve the equations numerically. We find that $\text{var}[\int_0^l dz'(a_1 - \langle a_1\rangle)] = \text{var}[\int_0^l dz' \,\text{Re}(a_4)] = \text{var}[\int_0^l dz' \text{Im}(a_4)]$, $\text{var}[\int_0^l dz'(a_3^2 - 1/3)] = \text{var}\{\int_0^l dz'[\text{Re}^2(a_6) - 1/3]\} = \text{var}\{\int_0^l dz'[\text{Im}^2(a_6) - 1/3]\}$ and $\text{var}[\int_0^l dz' \,\text{Re}(a_3a_6)] = \text{var}[\int_0^l dz' \text{Im}(a_3a_6)] = \text{var}[\int_0^l dz' \,\text{Re}(a_6)\text{Im}(a_6)]$. In Fig. 1, we plot the variance of the random coefficients versus h_{fiber}/L_B at large distance. The solid curve gives $\text{var}[\int_0^l dz'(a_1 - \langle a_1\rangle)]$, the long-dashed curve gives $\text{var}[\int_0^l dz'(a_3^2 - 1/3)]$, and the dotted curve gives $\text{var}[\int_0^l dz' \,\text{Re}(a_3a_6)]$. The variances are normalized by the distance l and the beat length L_B. The short-dashed curve is the azimuthal diffusion length normalized by L_B.

We note that [21] the equatorial diffusion length $d_1 \simeq h_{\text{fiber}}$. From Fig. 1, $\text{var}[\int_0^l dz'(a_1 - \langle a_1\rangle)] \simeq d_1 l$, while $\text{var}[\int_0^l dz'(a_3^2 - 1/3)]$ and $\text{var}[\int_0^l dz' \,\text{Re}(a_3a_6)] \propto d_3 l$. When $h_{\text{fiber}} \ll L_B$ the variance of the coefficients of the nonlinear terms are larger than is the case for the linear terms and thus the nonlinear terms dominate the polarization mode dispersion. When $h_{\text{fiber}} \gtrsim L_B$, the opposite applies. The relative contribution of the polarization mode dispersion to pulse evolution depends on the strength of the birefringence b', the pulse width, and pulse intensity as well as the random coefficients.

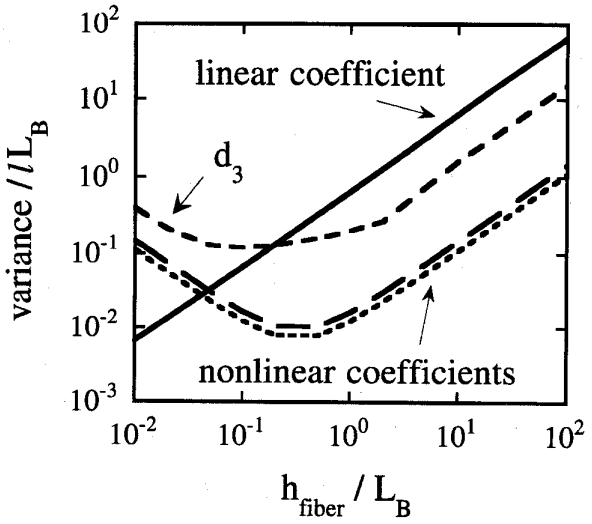

Fig. 1. The variances of the averaged random coefficients of the linear and nonlinear polarization mode dispersion to the Manakov equation versus h_{fiber}/L_B at large distance. The solid curve gives $\text{var}[\int_0^l dz'(a_1 - \langle a_1\rangle)]$, the long-dashed curve gives $\text{var}[\int_0^l dz'(a_3^2 - 1/3)]$, and the dotted curve gives $\text{var}[\int_0^l dz' \,\text{Re}(a_3a_6)]$. The variances are normalized by the distance l and the beat length L_B. The short-dashed curve is the azimuthal diffusion length d_3 normalized by L_B.

In simulations to date, it has been the practice to use a coarse-step approach in which one simply randomizes the electric fields at fixed intervals separated by a length z_{step} that is long compared to h_{fiber} and the d_j, while treating the field evolution inside the intervals deterministically. As long as the dispersive and nonlinear scale lengths are very long compared to z_{step}, this procedure will accurately reproduce the averaging that yields the Manakov equation. However, it will exaggerate the variance of the linear polarization mode dispersion by a factor that approximately equals $z_{\text{step}}/h_{\text{fiber}}$ and will exaggerate the variance of the nonlinear polarization mode dispersion by a factor that approximately equals z_{step}/d_3. In the former case, one can compensate for the exaggeration by appropriately lowering b', but that is not possible in the latter case, and the coarse-step approach will significantly overestimate the contribution of nonlinear polarization mode dispersion when $d_3 \ll z_{\text{step}}$.

V. Conclusion

In the two different physical models presented here, both of which allow for arbitrary orientation of the birefringence, we use an approach based on the theory of stochastic differential equations to show that the polarization mode dispersion has the same expression as that found for polarization-preserving fibers with small random mode coupling by Poole [4]. It is also the same as that found for a model with very general assumptions presented in [13]. We observed that the polarization mode dispersion depends on $h_{E,\text{local}}$, the polarization decorrelation length measured with respect to the local axes. We conjecture that with sufficient randomization, the polarization mode dispersion for fibers with randomly varying birefringence will

have the same functional form given in (26) for any physically reasonable distribution of the birefringence.

For the first model, we derived the polarization decorrelation length and the diffusion lengths measured with respect to both the local and initial eigenaxes. The results agree with those obtained from Monte Carlo simulations. We have shown that the length over which the linear term in the evolution equation of the electric field averages is given by the polarization decorrelation length measured with respect to the local axes of birefringence which is also the fiber autocorrelation length. The length over which the nonlinear term averages is given by the diffusion length in the azimuthal direction on the Poincaré sphere. At distances much larger than the longer of the two lengths, the effect of the random birefringence is that the electric field polarization varies rapidly on the Poincaré sphere, and the pulse envelope evolves according to the Manakov equation with corrections due to polarization mode dispersion. To accurately measure the relative contribution of the linear and nonlinear polarization mode dispersion, one must take into account that their relative contribution depends on the ratio of h_{fiber} and d_3. The usual practice of simply scrambling the electric fields at fixed intervals while treating their evolution inside the intervals deterministically tends to exaggerate the contribution of both linear and nonlinear polarization mode dispersion. While one can compensate for the former by changing the birefringence strength, that is not possible for the latter.

APPENDIX

Consider the stochastic differential equation

$$\frac{d\Omega_i}{dz} = \sum_{k=1}^{m} Q_{ik} g_k + U_i \tag{A1}$$

where $\Omega_i(z)$ and $U_i(z)$ are n-dimensional vectors, and $Q_{ik}(\Omega_i, z)$ is an $n \times m$ matrix. The process $g_i(z)$ is an m-dimensional white noise process with zero mean and variances σ_i^2, i.e.,

$$\langle g_k(z) \rangle = 0, \tag{A2}$$

$$\langle g_k(z) g_l(z') \rangle = \sigma_k^2 \delta_{kl} \delta(z - z') \tag{A3}$$

where δ_{kl} is the Konecker delta function

$$\delta_{kl} = \begin{cases} 1, & k = l \\ 0, & k \neq l \end{cases} \tag{A4}$$

and $\delta(z - z')$ is the Dirac delta function,

$$\delta(z - z') = \begin{cases} \infty, & z = z' \\ 0, & z \neq z' \end{cases}. \tag{A5}$$

To solve this equation, we must find a mathematically well-defined way to integrate (A1). This problem has been carefully studied by mathematicians who have completely solved it [18]; however, the presentations that have been made of this subject use the language of Lebesgue integration and measure theory which makes the results inaccessible to many physicists and engineers.

In this Appendix, we derive the master equation that governs the evolution of ensemble-averaged quantities at the level of rigor that is common in physics and engineering texts. Our starting point is that in any physical system the $g_k(z)$ will not really change instantaneously—but only over some finite but very small interval that we label ζ. We will therefore assume that the $g_k(z)$ are constant over intervals of length ζ and change randomly from interval to interval. With this assumption, the integration of (A1) is well defined. After carrying out this integral and ensemble-averaging, we will let $\zeta \to 0$ and show that the limit is well defined. We note that when one carries out Monte Carlo simulations of this sort of problem, one must use finite values of ζ to carry out the integration and verify that the results are independent of ζ once ζ is made sufficiently small. Thus, the analytical procedure that we are following closely mirrors both our physical understanding of the problems and what is done in computer simulations.

Let us consider any smooth function $\psi(\tilde{\mathbf{\Omega}})$ of the Ω_i. We will assume that the interval ζ is sufficiently small that ψ only changes slightly over the interval ζ so that if let $z = z_0$ be the beginning of an interval over which the $g_k(z)$ are constant, then at the end of that interval

$$\psi(z_0 + \zeta) = \psi(z_0) + \frac{d\psi}{dz}\zeta + \frac{1}{2}\frac{d^2\psi}{dz^2}\zeta^2 + \cdots \tag{A6}$$

where the first two derivatives may be written in accordance with the chain rule as

$$\begin{aligned} \frac{d\psi}{dz} &= \sum_{j=1}^{n} \frac{\partial \psi}{\partial \Omega_j}\frac{d\Omega_j}{dz}, \\ \frac{d^2\psi}{dz^2} &= \sum_{j=1}^{n}\left[\frac{\partial \psi}{\partial \Omega_j}\frac{\partial}{\partial z}\left(\frac{d\Omega_j}{dz}\right)\right] \\ &\quad + \sum_{j=1}^{n}\sum_{k=1}^{n}\left\{\frac{\partial^2 \psi}{\partial \Omega_j \partial \Omega_k}\frac{d\Omega_j}{dz}\frac{d\Omega_k}{dz}\right. \\ &\quad \left. + \frac{\partial \psi}{\partial \Omega_j}\left[\frac{\partial}{\partial \Omega_k}\left(\frac{d\Omega_j}{dz}\right)\right]\frac{d\Omega_k}{dz}\right\}. \end{aligned} \tag{A7}$$

We may now substitute (A1) into (A7) and (A7) into (A6), and we then ensemble average to obtain the ensemble-averaged change in $\psi(z)$. To calculate the ensemble-averaged change, we note that the appropriate means and variances for the g_k in the interval $(z_0, z_0 + \zeta)$ are given by

$$\begin{aligned} \langle g_k \rangle &= 0, \\ \langle g_k g_l \rangle &= \sigma_k^2 \delta_{kl}/\zeta. \end{aligned} \tag{A8}$$

The g_k in different intervals are uncorrelated. This form of the means and variances guarantees that the solutions of (A1) have a well-defined limit as $\zeta \to 0$. For example, if we consider the simple equation

$$\frac{d\Omega}{dz} = g \tag{A9}$$

and we integrate over the interval $(0, Z)$, we obtain

$$\Omega(Z) = \sum_{r=1}^{N} g^{(r)}\zeta \tag{A10}$$

where $N = Z/\zeta$ and $g^{(r)}$ is the value of g in the rth interval. We now find that

$$\langle \Omega(Z) \rangle = \sum_{r=1}^{N} \langle g^{(r)} \rangle \zeta = 0 \tag{A11}$$

and that

$$\begin{aligned} \langle \Omega^2(Z) \rangle &= \sum_{r=1}^{N} \sum_{s=1}^{N} \langle g^{(r)} g^{(s)} \rangle \zeta^2, \\ &= \sum_{r=1}^{N} \langle \left(g^{(r)}\right)^2 \rangle \zeta^2 = \sum_{r=1}^{N} \sigma^2 \zeta, \\ &= \sigma^2 N \zeta = \sigma^2 Z \end{aligned} \tag{A12}$$

which is independent of ζ.

When carrying out the ensemble average, there are two different averages that must be carried out. The first average, designated $\langle\ \rangle_1$, is over the different possible values of the g_k in the interval $(z_0, z_0 + \zeta)$. This average yields

$$\frac{1}{\zeta}[\langle \psi(z_0 + \zeta) \rangle_1 - \psi(z_0)] = G[\psi(z_0)] + \mathcal{O}(\zeta) \tag{A13}$$

where

$$\begin{aligned} G = &\sum_{j=1}^{n} U_j \frac{\partial}{\partial \Omega_j} + \frac{1}{2} \sum_{j=1}^{n} \sum_{k=1}^{n} \sum_{p=1}^{m} \sigma_p^2 \\ &\times \left(Q_{jp} Q_{kp} \frac{\partial^2}{\partial \Omega_j \partial \Omega_k} + Q_{kp} \frac{\partial Q_{jp}}{\partial \Omega_k} \frac{\partial}{\partial \Omega_j} \right) \end{aligned} \tag{A14}$$

and $\mathcal{O}(\zeta)$ indicates contributions that are proportional to ζ and higher powers of ζ and that tend to zero as ζ tends to zero. The second average designated $\langle\ \rangle_2$, is over all possible history of the g_i in all the intervals up to the point $z = z_0$. We note that the compound average $\langle \langle\ \rangle_1 \rangle_2$ is equivalent to the average over all possible histories up to $z = z_0 + \zeta$. Defining $\langle\ \rangle$ as the average over all possible histories up to the point z in the argument of $\psi(z)$, and carrying out the average $\langle\ \rangle_2$ on (A13), we obtain

$$\frac{1}{\zeta}[\langle \psi(z_0 + \zeta) \rangle - \langle \psi(z_0) \rangle] = \langle G[\psi(z_0)] \rangle + \langle \mathcal{O}(\zeta) \rangle. \tag{A15}$$

Finally, allowing $\zeta \to 0$, we obtain the result

$$\frac{\partial \langle \psi \rangle}{\partial z} = \langle G[\psi(z)] \rangle. \tag{A16}$$

In the mathematics literature, the generator $G(\psi)$ given in (A14) is referred to as the Stratonovich generator [18]. In this literature, another form of the generator in which the last term in (A14) is dropped, referred to as the Ito generator, is still widely in use. Historically, the Ito formulation developed first, and it has properties that make it convenient for proving mathematical theorems. However, as we have shown, the Stratonovich formulation appears as the natural limit of physically realistic models in which all variables have nonzero autocorrelation times. As the autocorrelation times of some of the variables tend to zero, the Stratonovich formulation is obtained.

REFERENCES

[1] C. D. Poole and R. E. Wagner, "Phenomenological approach to polarization dispersion in long single-mode fibers," *Electron. Lett.*, vol. 22, pp. 1029–1030, 1986.

[2] D. Andresciani, F. Curti, F. Matera, and B. Daino, "Measurement of the group-delay difference between the principal states of polarization on a low-birefringence terrestrial fiber cable," *Opt. Lett.*, vol. 12, pp. 844–846, 1987.

[3] N. S. Bergano, C. D. Poole, and R. E. Wagner, "Investigation of polarization dispersion in long lengths of single-mode fiber using multilongitudinal mode lasers," *J. Lightwave Technol.*, vol. 5, pp. 1618–1622, 1987.

[4] C. D. Poole, "Statistical treatment of polarization dispersion in single-mode fiber," *Opt. Lett.*, vol. 13, pp. 687–689, 1988.

[5] F. Curti, B. Daino, Q. Mao, F. Matera, and C. G. Someda, "Concatenation of polarization-dispersion in single-mode fibers," *Electron. Lett.*, vol. 25, pp. 290–291, 1989.

[6] C. D. Poole, "Measurement of polarization-mode dispersion in single-mode fibers with random mode coupling," *Opt. Lett.*, vol. 14, pp. 523–525, 1989.

[7] F. Curti, B. Daino, G. De Marchis, and F. Matera, "Statistical treatment of the evolution of the principal states of polarization in single-mode fibers," *J. Lightwave Technol.*, vol. 8, pp. 1162–1165, 1990.

[8] C. D. Poole, J. H. Winters, and J. A. Nagel, "Dynamical equation for polarization dispersion," *Opt. Lett.*, vol. 16, pp. 372–374, 1991.

[9] S. Betti, F. Curti, G. De Marchis, E. Iannone, and F. Matera, "Evolution of the bandwidth of the principal states of polarization in single-mode fibers," *Opt. Lett.*, vol. 16, pp. 467–469, 1991.

[10] G. J. Foschini and C. D. Poole, "Statistical theory of polarization dispersion in single mode fibers," *J. Lightwave Technol.*, vol. 9, pp. 1439–1456, 1991.

[11] Y. Namihara, T. Kawazawa and H. Wakabayashi, "Polarization mode dispersion measurements in 1520 km EDFA system," *Electron. Lett.*, vol. 28, pp. 881–882, 1992.

[12] A. Galtarossa and M. Schiano, "Complete characterization of polarization mode dispersion in erbium doped optical amplifiers," *Electron. Lett.*, vol. 28, pp. 2143–2144, 1992; also in A. Galtarossa, "Comment: Polarization mode dispersion in 1520 EDFA system," *Electron. Lett.*, vol. 29, pp. 564–565, 1993.

[13] C. R. Menyuk and P. K. A. Wai, "Polarization evolution and dispersion in fibers with spatially varying birefringence," *J. Opt. Soc. Am. B*, vol. 11, pp. 1288–1296, 1994.

[14] P. K. A. Wai and C. R. Menyuk, "Polarization decorrelation in fibers with randomly varying birefringence," *Opt. Lett.*, vol. 19, pp. 1517–1519, 1994.

[15] I. P. Kaminow, "Polarization in optical fibers," *IEEE J. Quantum Electron.*, vol. QE-17, pp. 15–22, 1981.

[16] S. C. Rashleigh and R. Ulrich, "Polarization mode dispersion in single-mode fibers," *Opt. Lett.*, vol. 3, pp. 60–62, 1978.

[17] S. C. Rashleigh, W. K. Burns, R. P. Moeller, and R. Ulrich, "Polarization holding in birefringent single-mode fibers," *Opt. Lett.*, vol. 7, pp. 40–42, 1982.

[18] L. Arnold, *Stochastic Differential Equations, Theory and Applications*. New York: Wiley, 1974.

[19] P. K. A. Wai, C. R. Menyuk, and H. H. Chen, "Stability of solitons in randomly varying birefringent fibers," *Opt. Lett.*, vol. 16, pp. 1231–1233, 1991.

[20] S. G. Evangelides, Jr., L. F. Mollenauer, J. P. Gordon, and N. S. Bergano, "Polarization multiplexing with solitons," *J. Lightwave Technol.*, vol. 10, pp. 28–35, 1992.

[21] P. K. A. Wai and C. R. Menyuk, "Anisotropic evolution of the state of polarization in optical fibers with random varying birefringence," to appear in *Opt. Lett.*.

[22] T. Ueda and W. L. Kath, "Dynamics of optical pulses in randomly birefringent fibers," *Physica D*, vol. 55, pp. 166–181, 1992.

[23] C. R. Menyuk, "Pulse propagation in an elliptically birefringent kerr medium," *IEEE J. Quantum Electron.*, vol. 25, pp. 2674–2682, 1989.

P. K. A. Wai, photograph and biography not available at the time of publication.

C. R. Menyuk, photograph and biography not available at the time of publication.