Correlation theory of polarization mode dispersion in optical fibers

Qiang Lin and Govind P. Agrawal

Institute of Optics, University of Rochester, Rochester, New York 14627

Received May 9, 2002; revised manuscript received August 12, 2002

A general theory is used to describe the correlation properties of polarization mode dispersion (PMD) in a birefringent, linear, dispersive medium such as optical fibers. The theory includes the effects of frequency dependence of birefringence on all orders, and it is capable of providing statistical information about second- and higher-order correlations among the polarization and PMD vectors. We apply the general theory to study pulse broadening induced by different-order PMD and PMD-induced pulse distortion through the third- and fourth-order temporal moments (related to skewness and flatness, respectively). Our analytic results are in good agreement with numerical simulations. © 2003 Optical Society of America

OCIS codes: 060.2330, 060.4510, 060.5530, 260.5430, 260.2030.

1. INTRODUCTION

The phenomenon of polarization mode dispersion (PMD) becomes of considerable concern as the bit rate of each channel in a light-wave system increases beyond 10 Gbit/s.[1,2] It is anticipated that the performance of long-haul 40-Gbit/s systems would be limited by PMD, and the situation would become worse as the bit rate increases toward 80 Gbit/s or more. It is well known that the effects of higher-order PMD should be included for ultrashort pulses used in designing such high-bit-rate systems.[3] In a linear dispersive medium, the PMD effects are well described by the PMD vector $\mathbf{\Omega}$ whose direction is along a principal state of polarization (PSP), and whose magnitude corresponds to the differential group delay (DGD) between the two PSPs.[4] When the bandwidth of optical pulses is small, the two PSPs can be assumed to remain constant over the whole pulse spectrum. This approximation is referred to as the first-order PMD. The magnitude of the PMD vector is known to follow the Maxwellian statistics in this case.[5]

For short optical pulses, the bandwidth increases enough that higher-order effects should be included. The statistics of the second-order PMD vector defined as $d\mathbf{\Omega}/d\omega$ has been thoroughly investigated in Refs. 5–8. However, it seems rather difficult to find the probability distributions of higher-order PMD vectors and the joint probability of multiple PMD vectors of different frequencies. The correlation functions between the polarization and the PMD vectors are likely to play an important role because considerable information about the statistics of pulse propagation can be obtained through them even when the joint probability distribution is unknown. For example, the correlation between two PMD vectors of different frequencies is useful to understand the pulse broadening induced by PMD effects.[9,10] However, if we consider other effects such as PMD-induced pulse distortion, we would need higher-order correlation functions. In this paper we develop a general theory that can be used to describe the correlations among PMD vectors and polarizations of an arbitrary order and apply it to discuss pulse distortion induced by PMD.

2. GENERAL FORMALISM

The propagation of an optical pulse in a linear, birefringent, dispersive medium can be studied if we consider each frequency component of the pulse through the Fourier transform

$$\tilde{\mathcal{A}}(z,\omega)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\mathcal{A}(z,t)\exp(i\omega t)\mathrm{d}t, \tag{1}$$

where $\mathcal{A}(z,t)$ is the Jones vector of the optical field and $\tilde{\mathcal{A}}(z,\omega)=a(z,\omega)\exp[i\theta(z,\omega)]\mathcal{S}(z,\omega)$ is its spectrum. Here, a, θ, and $\mathcal{S}$ are the amplitude, phase, and state of polarization (SOP) of the pulse spectrum, and $\mathcal{S}^{\dagger}\mathcal{S}=1$. The effects of fiber birefringence in the Jones matrix formalism are governed by[11]

$$\frac{\partial\tilde{\mathcal{A}}(z,\omega)}{\partial z}\equiv\frac{\partial}{\partial z}\begin{pmatrix}\tilde{A}_x\\ \tilde{A}_y\end{pmatrix}=i\begin{pmatrix}\beta_0-b_1/2 & -b_2/2+ib_3/2\\ -b_2/2-ib_3/2 & \beta_0+b_1/2\end{pmatrix}\begin{pmatrix}\tilde{A}_x\\ \tilde{A}_y\end{pmatrix}. \tag{2}$$

This equation can be rewritten in matrix form as

$$\frac{\partial\tilde{\mathcal{A}}(z,\omega)}{\partial z}\equiv i\left[\beta_0(z,\omega)-\frac{1}{2}\mathbf{b}(z,\omega)\cdot\boldsymbol{\sigma}\right]\tilde{\mathcal{A}}(z,\omega), \tag{3}$$

where $\beta_0(z,\omega)$ is the propagation constant and $\mathbf{b}(z,\omega)=b_1\hat{e}_1+b_2\hat{e}_2+b_3\hat{e}_3$ is the local birefringence vector defined in the Stokes space. The vector $\boldsymbol{\sigma}$ is formed by use of the Pauli matrices and is given by $\boldsymbol{\sigma}=\sigma_1\hat{e}_1+\sigma_2\hat{e}_2+\sigma_3\hat{e}_3$, where $\hat{e}_1$, $\hat{e}_2$, and $\hat{e}_3$ are unit vectors in the Stokes space and

$$\sigma_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \tag{4}$$

The PMD vector appears when one considers the frequency dependence of polarization $\mathcal{S}(z, \omega)$ through[11]

$$\frac{\partial \mathcal{S}(z, \omega)}{\partial \omega} = i\left[\tau(z, \omega) - \frac{1}{2}\mathbf{\Omega}(z, \omega) \cdot \boldsymbol{\sigma}\right]\mathcal{S}(z, \omega), \tag{5}$$

where τ is the common group delay induced by the dispersion, defined as $\tau = \int_0^z (\partial\beta_0/\partial\omega)\mathrm{d}z$, and $\mathbf{\Omega}$ is the PMD vector that points along the fast PSP. Note that Eq. (5) differs slightly from Eq. (6.9) in Ref. 11 because of a different convention that was used for the Fourier transform.[12]

In the absence of polarization-dependent losses, the dynamics of polarization and PMD vectors are governed by[11,13]

$$\frac{\partial \mathbf{S}(z, \omega)}{\partial z} = \mathbf{b}(z, \omega) \times \mathbf{S}(z, \omega), \tag{6}$$

$$\frac{\partial \mathbf{\Omega}(z, \omega)}{\partial z} = \mathbf{b}(z, \omega) \times \mathbf{\Omega}(z, \omega) + \mathbf{b}_\omega(z, \omega), \tag{7}$$

where $\mathbf{S}$ is the normalized Stokes vector of the polarization defined as $\mathbf{S} = \mathcal{S}^\dagger \boldsymbol{\sigma} \mathcal{S}$. It is normalized such that $\mathbf{S}\cdot\mathbf{S}=1$. The subscript ω denotes the frequency derivative. It is evident that the frequency dependence of PMD vector $\mathbf{\Omega}$ is governed by that of the birefringence vector $\mathbf{b}$. PMD is a consequence of the randomly varying birefringence along the fiber length. It is important to note that both the orientation of the principal axes and the magnitude of birefringence are random.

Physically, it is reasonable to view the fiber as an array of concatenated sections whose lengths are small enough that their birefringence can be considered constant inside each section.[11–15] Within each section, the frequency dependence of birefringence is deterministic. Random coupling occurs only between different fiber sections because of the random rotations of the principal axes. It is expected that the frequency dependence of $\mathbf{b}$ will not change from section to section. It is thus reasonable to assume that $\mathbf{b}(z, \omega)$ can be factored as

$$\mathbf{b}(z, \omega) = f(\omega)\mathbf{b}(z), \tag{8}$$

where $f(\omega)$ represents the frequency dependence and is a deterministic function, whereas $\mathbf{b}(z)$ is a Gaussian random process with a correlation length of ~0.1 km. In practice, one is interested in the PMD effects at distances much larger than the correlation length. For such large distances, $\mathbf{b}(z)$ can be considered a Markovian Gaussian process (white noise) whose first two moments are given by

$$\langle \mathbf{b}(z) \rangle = 0, \quad \langle \mathbf{b}(z_1)\mathbf{b}(z_2) \rangle = \eta^2 \overleftrightarrow{\mathbf{I}} \delta(z_2 - z_1), \tag{9}$$

where the angle brackets denote an ensemble average, $\overleftrightarrow{\mathbf{I}}$ is a unit tensor, and η represents the rms value of birefringence $\mathbf{b}(z)$.

Since the birefringence vector $\mathbf{b}(z, \omega)$ is a three-dimensional vector that moves uniformly in the Stokes space and its amplitude is frequency dependent on any order (i.e., the local DGD is frequency dependent on any order), the model includes the random rotations of principal axes as well as the random phase additions, while it maintains the frequency dependence of birefringence on all orders. Using Eq. (8) in Eqs. (6) and (7), the dynamic equations for the polarization and the PMD vectors become

$$\frac{\partial \mathbf{S}(z, \omega)}{\partial z} = f(\omega)\mathbf{b}(z) \times \mathbf{S}(z, \omega), \tag{10}$$

$$\frac{\partial \mathbf{\Omega}(z, \omega)}{\partial z} = f(\omega)\mathbf{b}(z) \times \mathbf{\Omega}(z, \omega) + f_\omega(\omega)\mathbf{b}(z). \tag{11}$$

For each frequency, the polarization and the PMD vectors evolve randomly along the fiber. The polarizations and the PMD vectors of different frequencies follow the same dynamics but with different random phase shifts and PSP rotations at each fiber section because of the frequency dependence of birefringence. Mathematically, Eqs. (10) and (11) constitute a set of six nonlinear Langevin equations. The moments of polarization and PMD vectors can be calculated by use of the well-known technique based on the Stratonovich calculus.[16–18] Although there is some ambiguity between the Ito and the Stratonovich interpretations, the Stratonovich calculus is considered the natural limit of real physical models.[16,18] For this reason we used it to investigate the statistical behavior of the polarization and the PMD vectors and their correlations for different frequencies. In Appendix A we provide the mathematical details and show how to determine the Stratonovich generator for the PMD and the polarization vectors of different frequencies.

3. SECOND-ORDER CORRELATION FUNCTIONS

Generator G in Eq. (A10) of Appendix A can be used to calculate the correlation between the PMD vectors and the polarization vectors at n different frequencies. If we consider only the correlation between two PMD vectors of different frequencies, $\mathbf{\Omega}(z, \omega_1)$ and $\mathbf{\Omega}(z, \omega_2)$, the generator reduces to

$$\begin{aligned} G = \frac{\eta^2}{2}\{&-2{f_1}^2\mathbf{\Omega}_1 \cdot \nabla_1 - 2{f_2}^2\mathbf{\Omega}_2 \cdot \nabla_2 \\ &+ [({f_{1\omega}}^2 + {f_1}^2{\Omega_1}^2)\overleftrightarrow{\mathbf{I}} - {f_1}^2\mathbf{\Omega}_1\mathbf{\Omega}_1]:\nabla_1\nabla_1 \\ &+ [({f_{2\omega}}^2 + {f_2}^2{\Omega_2}^2)\overleftrightarrow{\mathbf{I}} - {f_2}^2\mathbf{\Omega}_2\mathbf{\Omega}_2]:\nabla_2\nabla_2 \\ &+ 2[f_{1\omega}f_{2\omega}\overleftrightarrow{\mathbf{I}} + f_1f_2[(\mathbf{\Omega}_1 \cdot \mathbf{\Omega}_2)\overleftrightarrow{\mathbf{I}} - \mathbf{\Omega}_2\mathbf{\Omega}_1]]:\nabla_2\nabla_1 \\ &+ 2(f_2f_{1\omega}\mathbf{\Omega}_2 - f_1f_{2\omega}\mathbf{\Omega}_1) \cdot (\nabla_2 \times \nabla_1)\}, \end{aligned} \tag{12}$$

where subscripts 1 and 2 denote the frequency components at ω_1 and ω_2, respectively. For example, $f_1 = f(\omega_1)$, $f_{1\omega} = f_\omega(\omega_1)$, and ∇_1 operates on $\Omega(z, \omega_1)$. The dynamic equations for the correlation between $\mathbf{\Omega}(z, \omega_1)$ and $\mathbf{\Omega}(z, \omega_2)$ are

$$\frac{\mathrm{d}\langle \mathbf{\Omega}_1 \cdot \mathbf{\Omega}_2\rangle}{\mathrm{d}z} = \langle G(\mathbf{\Omega}_1 \cdot \mathbf{\Omega}_2)\rangle = -\eta^2(f_1 - f_2)^2\langle \mathbf{\Omega}_1 \cdot \mathbf{\Omega}_2\rangle + 3\eta^2 f_{1\omega} f_{2\omega}, \tag{13}$$

$$\frac{\mathrm{d}\langle \mathbf{\Omega}_1\mathbf{\Omega}_2\rangle}{\mathrm{d}z} = -\eta^2({f_1}^2 + {f_2}^2)\langle \mathbf{\Omega}_1\mathbf{\Omega}_2\rangle - \eta^2 f_1 f_2\langle \mathbf{\Omega}_2\mathbf{\Omega}_1\rangle + \eta^2[f_{1\omega}f_{2\omega} + f_1 f_2\langle \mathbf{\Omega}_1 \cdot \mathbf{\Omega}_2\rangle]\vec{\vec{\mathbf{I}}}. \tag{14}$$

The dynamic equation for $\langle \mathbf{\Omega}_2\mathbf{\Omega}_1\rangle$ can be written by interchanging subscripts 1 and 2 in Eq. (14). These equations can be easily integrated to yield the following analytic expressions:

$$\langle \mathbf{\Omega}_1 \cdot \mathbf{\Omega}_2\rangle = \frac{3f_{1\omega}f_{2\omega}}{(f_1 - f_2)^2}\{1 - \exp[-\eta^2 z(f_1 - f_2)^2]\}, \tag{15}$$

$$\langle \mathbf{\Omega}_1\mathbf{\Omega}_2\rangle = \frac{f_{1\omega}f_{2\omega}\vec{\vec{\mathbf{I}}}}{(f_1 - f_2)^2}\{1 - \exp[-\eta^2 z(f_1 - f_2)^2]\}. \tag{16}$$

Equation (16) shows that the three components of the PMD vector become uncorrelated when the propagation distance is long and the variance of the three components is the same. Physically speaking, random rotation of the PMD vector fills the Poincare sphere uniformly in the Stokes space, which means that the three components of the PMD vectors are uncorrelated and identical. From Eq. (15), when the frequency difference between the two PMD vectors vanishes, the correlation reduces to the variance of the magnitude of the PMD vector, $\langle \Omega^2\rangle = 3\eta^2 {f_\omega}^2(\omega)z$. Clearly, if we include the frequency-dependent birefringence beyond the DGD, variance of the PMD vector itself becomes frequency dependent. This frequency dependence could be small because local DGD is dominant for typical pulse bandwidths. Using the generator method, one can easily find the probability distribution of the PMD vector at a single frequency. It turns out that the probability distribution of the PMD vector remains Gaussian even when higher-order frequency-dependent birefringence is considered, but the variance becomes frequency dependent as found above. When the frequency difference becomes large, the difference between functions $f(\omega_1)$ and $f(\omega_2)$ becomes large and the correlation tends to zero.

A similar approach can be used for correlation between the polarization vectors at two different frequencies. Through the generator in Appendix A, the correlation function was found to be

$$\langle \mathbf{S}_1 \cdot \mathbf{S}_2\rangle = \langle \mathbf{S}_{01} \cdot \mathbf{S}_{02}\rangle \exp[-\eta^2 z(f_1 - f_2)^2], \tag{17}$$

where $\mathbf{S}_{01}$ and $\mathbf{S}_{02}$ is the input polarization vectors at the two frequencies. Random birefringence causes the SOP to diffuse uniformly in the Stokes space and also causes the polarization vectors at different frequencies to become increasingly uncorrelated as the propagation distance increases. A comparison of Eq. (17) with Eq. (15) shows that the diffusion rate of polarization is determined by variance of the PMD.

Equations (15)–(17) are valid for an arbitrary form of frequency dependence. In practice, $f(\omega)$ is expanded in a Taylor series around the central frequency of the pulse. When the higher-order terms are negligible, birefringence is dominated by the first-order effects. Using $f(\omega_2) - f(\omega_1) \approx f_\omega(\omega_1)(\omega_2 - \omega_1)$ and $f_\omega(\omega_1) \approx f_\omega(\omega_2)$, the correlation function in Eqs. (15) and (17) reduces to the well-known results[9,10,19]

$$\langle \mathbf{S}_1 \cdot \mathbf{S}_2\rangle \approx \langle \mathbf{S}_{01} \cdot \mathbf{S}_{02}\rangle \exp[-\langle \Omega^2\rangle(\omega_2 - \omega_1)^2/3], \tag{18}$$

$$\langle \mathbf{\Omega}_1 \cdot \mathbf{\Omega}_2\rangle \approx \frac{3}{(\omega_2 - \omega_1)^2}\{1 - \exp[-\langle \Omega^2\rangle(\omega_2 - \omega_1)^2/3]\}, \tag{19}$$

where $\langle \Omega^2\rangle$ is constant over the pulse spectrum. Approximations (18) and (19) were first derived in Refs. 9 and 19 by use of different methods. Here, they appear naturally from our general theory within the appropriate limit.

The statistics of the input PMD vector is also of interest because it is directly related to properties such as pulse broadening and distortion. The input PMD vector, $\mathbf{\Omega}_0(z, \omega)$, is obtained as the projection of PMD vector $\mathbf{\Omega}(z, \omega)$ by use of $\mathbf{\Omega}(z, \omega) = R(z, \omega)\mathbf{\Omega}_0(z, \omega)$, where $R(z, \omega)$ is the rotation matrix of the whole fiber in the Stokes space.[11,20] Through the generator given in Appendix A, it is simple to show that

$$\langle (\mathbf{S}_1 \cdot \mathbf{\Omega}_1)(\mathbf{S}_2 \cdot \mathbf{\Omega}_2)\rangle = \frac{\langle \mathbf{S}_{01} \cdot \mathbf{S}_{02}\rangle f_{1\omega}f_{2\omega}}{(f_1 - f_2)^2} \times \{1 - \exp[-\eta^2 z(f_1 - f_2)^2]\}. \tag{20}$$

Since $\langle (\mathbf{S}_1 \cdot \mathbf{\Omega}_1)(\mathbf{\Omega}_2 \cdot \mathbf{S}_2)\rangle = \langle \mathbf{S}_{01} \cdot \mathbf{\Omega}_{01}\mathbf{\Omega}_{02} \cdot \mathbf{S}_{02}\rangle$ and the input polarization is arbitrary, we conclude that $\langle \mathbf{\Omega}_0(z, \omega_1)\mathbf{\Omega}_0(z, \omega_2)\rangle = \langle \mathbf{\Omega}(z, \omega_1)\mathbf{\Omega}(z, \omega_2)\rangle$. Thus, the correlation function for the input PMD vector is exactly the same as that for the output PMD vector.

The results in this section show the power of the generator technique for finding the correlation functions of PMD vectors and polarizations. In Section 4 we apply this technique to calculate several higher-order correlation functions between the PMD and the polarization vectors of different frequencies and relate them to pulse distortion for investigating the PMD effects beyond pulse broadening. Since the pulse bandwidth is usually not large enough to make significant higher-order birefringence effects, we assume that the local DGD is constant and use the approximation $f(\omega_2) - f(\omega_1) \approx (\omega_2 - \omega_1)f_\omega$, where f_ω is constant over the pulse bandwidth.

4. PULSE DISTORTION AND BROADENING

The pulse broadening is governed by the rms width of the pulse defined as $\sigma^2 = \widehat{t^2} - (\hat{t})^2$, where

$$\hat{t} = \int_{-\infty}^{+\infty} t\mathcal{A}^\dagger(z, t)\mathcal{A}(z, t)\mathrm{d}t = -i\int_{-\infty}^{+\infty} \tilde{\mathcal{A}}^\dagger(z, \omega)\tilde{\mathcal{A}}_\omega(z, \omega)\mathrm{d}\omega, \tag{21}$$

$$\widehat{t^2} = \int_{-\infty}^{+\infty} t^2 \mathcal{A}^\dagger(z, t)\mathcal{A}(z, t)\mathrm{d}t = \int_{-\infty}^{+\infty} \tilde{\mathcal{A}}_\omega{}^\dagger(z, \omega)\tilde{\mathcal{A}}_\omega(z, \omega)\mathrm{d}\omega, \quad (22)$$

and the circumflex denotes the average over the pulse intensity profile. We have assumed that the optical field $\mathcal{A}(z, t)$ and its spectrum $\tilde{\mathcal{A}}(z, \omega)$ are normalized such that

$$\int_{-\infty}^{+\infty} \mathcal{A}^\dagger(z, t)\mathcal{A}(z, t)\mathrm{d}t = \int_{-\infty}^{+\infty} \tilde{\mathcal{A}}^\dagger(z, \omega)\tilde{\mathcal{A}}(z, \omega)\mathrm{d}\omega = 1. \quad (23)$$

The dagger represents the Hermitian conjugate.

Consider the input pulse spectrum to be $\tilde{\mathcal{A}}(0, \omega) = a(\omega)\exp[i\theta(\omega)]\mathcal{S}_0$, where $a(\omega)$, $\theta(\omega)$, and $\mathcal{S}_0$ are the amplitude, phase, and polarization of the initial pulse spectrum; $\mathcal{S}_0{}^\dagger\mathcal{S}_0 = 1$. We assume that the input polarization state is fixed for all frequency components, and thus $\mathcal{S}_0$ is constant over the whole pulse spectrum. Since variation of the optical field with frequency is governed by Eq. (5), we substitute it into Eqs. (21) and (22) to obtain pulse broadening in the form

$$\sigma^2 = \sigma_0{}^2 + \sigma_{\mathrm{disp}}{}^2 + \sigma_{\mathrm{PMD}}{}^2, \quad (24)$$

where σ_0 is the rms width of the input pulse, $\sigma_{\mathrm{disp}}{}^2 = [\overline{\tau^2} - (\overline{\tau})^2] + 2[\overline{\tau\theta_\omega} - \overline{\tau}\,\overline{\theta_\omega}]$ is the broadening induced by fiber dispersion, and the PMD-induced broadening σ_{PMD} is given by[11,20]

$$\sigma_{\mathrm{PMD}}{}^2 = \frac{1}{4}[\overline{\Omega_0{}^2} - (\overline{\boldsymbol{\Omega}_0 \cdot \mathbf{S}_0})^2] - [\overline{(\tau + \theta_\omega)(\boldsymbol{\Omega}_0 \cdot \mathbf{S}_0)} - (\overline{\tau} + \overline{\theta_\omega})\,\overline{(\boldsymbol{\Omega}_0 \cdot \mathbf{S}_0)}]. \quad (25)$$

Here the overbar denotes the average over the input spectrum, i.e., $\bar{g} = \int_{-\infty}^{+\infty} g(\omega)\tilde{\mathcal{A}}^\dagger(0, \omega)\tilde{\mathcal{A}}(0, \omega)\mathrm{d}\omega$, $\mathbf{S}_0 = \mathcal{S}_0{}^\dagger\boldsymbol{\sigma}\mathcal{S}_0$ is the normalized Stokes vector of the initial field, $\theta_\omega = \mathrm{d}\theta/\mathrm{d}\omega$ is related to the initial chirp, and $\boldsymbol{\Omega}_0(z, \omega)$ is the input PMD vector.

It is difficult to find an analytic expression for the probability distribution of PMD-induced pulse broadening. However, the ensemble-averaged values provide an estimate of average pulse broadening. Since $\langle\boldsymbol{\Omega}_0(z, \omega)\rangle = 0$, the average broadening induced by PMD becomes

$$\langle\sigma_{\mathrm{PMD}}{}^2\rangle = \frac{1}{4}[\langle\overline{\Omega_0{}^2}\rangle - \langle(\overline{\boldsymbol{\Omega}_0 \cdot \mathbf{S}_0})^2\rangle]. \quad (26)$$

Karlsson and Brentel[9] and Sunnerud *et al.*[21] obtained this expression for unchirped Gaussian pulses after neglecting higher-order frequency-dependent birefringence. Our analysis shows that Eq. (26) can also be used directly for pulses of arbitrary shape and arbitrary initial chirp. It includes higher-order frequency-dependent effects of birefringence as well.

To investigate the higher-order PMD effects, we expand the input PMD vector into a Taylor series as

$$\boldsymbol{\Omega}_0(z, \omega) = \sum_{n=0}^{+\infty} \frac{1}{n!}\boldsymbol{\Omega}_0{}^{(n)}\omega^n,$$

where

$$\boldsymbol{\Omega}_0{}^{(n)} = \partial^n\boldsymbol{\Omega}_0/\partial\omega^n\big|_{\omega=\omega_0}$$

is the nth-order frequency derivative of the input PMD vector at the central frequency ω_0 of the pulse spectrum, which we refer to as the $(n + 1)$th-order PMD vector so that it corresponds to the usual definition of the second-order PMD vector for $n = 1$. Substituting $\Omega_0(z, \omega)$ into Eq. (26), the PMD-induced broadening becomes

$$\langle\sigma_{\mathrm{PMD}}{}^2\rangle = \frac{1}{4}\sum_{N=0}^{+\infty}\left\{\overline{\omega^N}\left[\sum_{m=0}^{N}\frac{\langle\boldsymbol{\Omega}_0{}^{(m)} \cdot \boldsymbol{\Omega}_0{}^{(N-m)}\rangle}{m!(N-m)!}\right] - \sum_{m=0}^{N}\frac{\overline{\omega^m}\;\overline{\omega^{N-m}}}{m!(N-m)!}\mathbf{S}_0 \cdot \langle\boldsymbol{\Omega}_0{}^{(m)}\boldsymbol{\Omega}_0{}^{(N-m)}\rangle \cdot \mathbf{S}_0\right\}, \quad (27)$$

where the overbar means average over the input spectrum as before. Since the correlation between the PMD vectors of two frequencies, as shown in Eq. (19), is stationary in the frequency domain when we neglect the higher-order frequency-dependent birefringence, it is easy to show that

$$\langle\boldsymbol{\Omega}_0{}^{(m)} \cdot \boldsymbol{\Omega}_0{}^{(N-m)}\rangle = (-1)^{N-m}\langle\boldsymbol{\Omega}_0{}^{(N)} \cdot \boldsymbol{\Omega}_0{}^{(0)}\rangle, \quad (28)$$

$$\sum_{m=0}^{N}\frac{\langle\boldsymbol{\Omega}_0{}^{(m)} \cdot \boldsymbol{\Omega}_0{}^{(N-m)}\rangle}{m!(N-m)!} = \langle\boldsymbol{\Omega}_0{}^{(N)} \cdot \boldsymbol{\Omega}_0{}^{(0)}\rangle\sum_{m=0}^{N}\frac{(-1)^{N-m}}{m!(N-m)!} = 0 \quad (29)$$

for all $N \geqslant 1$. Moreover, the input power spectrum is usually symmetric, resulting in $\overline{\omega^{2m+1}} = 0$ for all integers $m \geqslant 0$. With these simplifications, the PMD-induced pulse broadening becomes

$$\langle\sigma_{\mathrm{PMD}}{}^2\rangle = \frac{1}{4}\left\{\langle(\boldsymbol{\Omega}_0{}^{(0)})^2\rangle - \sum_{N=0}^{+\infty}(\mathbf{S}_0 \cdot \langle\boldsymbol{\Omega}_0{}^{(2N)}\boldsymbol{\Omega}_0{}^{(0)}\rangle \cdot \mathbf{S}_0) \times \sum_{m=0}^{N}\frac{\overline{\omega^{2m}}\;\overline{\omega^{(2N-2m)}}}{(2m)!(2N-2m)!}\right\}. \quad (30)$$

Equation (30) constitutes one of our main results. It shows that all the even-order PMD effects (which correspond to odd-order frequency derivative of the input PMD vector) do not participate in the production of average pulse broadening. If the first-order PMD is compensated, the dominant PMD effect on the average pulse broadening is of third order. Note that this does not mean that the even-order PMD effects are irrelevant for pulse broadening. They still affect instantaneous pulse broadening and thus affect the bit error rate and the outage probability.[22] In other words, the tail of the probability distribution of the broadening, which is related to the system outage, is affected by the even-order PMD effects, although the rms width of the distribution does not. Therefore, the ensemble-averaged pulse broadening is not sufficient to judge the PMD effects. The ensemble-

averaged pulse broadening is often used to compare different PMD compensation techniques. Our analysis shows that one should be careful if this quantity is used because it does not include the effects of all even-order PMD.

PMD not only induces pulse broadening but also leads to pulse distortion. In particular, PMD transfers energy between the two polarization components of an optical pulse and pulls them apart randomly. Thus, PMD can make the pulse shape asymmetric and flat even for a symmetric and sharp input pulse. Similar to the coma and spherical aberrations that occur in an imaging system, we apply the concepts of skewness and flatness to describe the extent of pulse distortion. Skewness and flatness are related to the third- and fourth-order central moments of the pulse intensity profile defined as

$$\mu_3 \equiv \int_{-\infty}^{+\infty} (t-\hat{t})^3 \mathcal{A}^{\dagger}(z,t)\mathcal{A}(z,t)\mathrm{d}t = \hat{t}^3 - 3\hat{t}\hat{t}^2 + 2(\hat{t})^3, \tag{31}$$

$$\mu_4 \equiv \int_{-\infty}^{+\infty} (t-\hat{t})^4 \mathcal{A}^{\dagger}(z,t)\mathcal{A}(z,t)\mathrm{d}t = \hat{t}^4 - 4\hat{t}\hat{t}^3 + 6\hat{t}^2(\hat{t})^2 - 3(\hat{t})^4, \tag{32}$$

where

$$\hat{t}^3 = \int_{-\infty}^{+\infty} t^3 \mathcal{A}^{\dagger}(z,t)\mathcal{A}(z,t)\mathrm{d}t = -i\int_{-\infty}^{+\infty} \tilde{\mathcal{A}}_{\omega}{}^{\dagger}(z,\omega)\tilde{\mathcal{A}}_{\omega\omega}(z,\omega)\mathrm{d}\omega, \tag{33}$$

$$\hat{t}^4 = \int_{-\infty}^{+\infty} t^4 \mathcal{A}^{\dagger}(z,t)\mathcal{A}(z,t)\mathrm{d}t = \int_{-\infty}^{+\infty} \tilde{\mathcal{A}}_{\omega\omega}{}^{\dagger}(z,\omega)\tilde{\mathcal{A}}_{\omega\omega}(z,\omega)\mathrm{d}\omega. \tag{34}$$

The complete expressions of skewness and flatness including all-order effects of fiber dispersion, initial chirp, and PMD are too complicated to present here. If we focus only on the PMD effects and assume that the input pulse is symmetric and unchirped and that the fiber dispersion has been totally compensated, the skewness and flatness become (see Appendix B)

$$\mu_3 = \frac{1}{8}\{2\overline{(\mathbf{\Omega}\times\mathbf{\Omega}_{\omega})\cdot\mathbf{S}} - \overline{\Omega^2(\mathbf{\Omega}\cdot\mathbf{S})} + 3\overline{\Omega^2}\,\overline{(\mathbf{\Omega}\cdot\mathbf{S})} - 2\overline{(\mathbf{\Omega}\cdot\mathbf{S})}^3\} - \frac{1}{2}\left\{\int_{-\infty}^{+\infty}[(a_{\omega})^2 - 2aa_{\omega\omega}] \times (\mathbf{\Omega}\cdot\mathbf{S})\mathrm{d}\omega - \overline{(\mathbf{\Omega}\cdot\mathbf{S})}\int_{-\infty}^{+\infty}(a_{\omega})^2\mathrm{d}\omega\right\}, \tag{35}$$

$$\mu_4 = \frac{1}{16}\{\overline{\Omega^4} + 4\overline{\Omega_{\omega}{}^2} + 8\overline{(\mathbf{\Omega}\cdot\mathbf{S})}\,\overline{(\mathbf{\Omega}\times\mathbf{\Omega}_{\omega})\cdot\mathbf{S}} - 4\overline{\Omega^2(\mathbf{\Omega}\cdot\mathbf{S})}\,\overline{(\mathbf{\Omega}\cdot\mathbf{S})} + 6\overline{\Omega^2}\,\overline{(\mathbf{\Omega}\cdot\mathbf{S})}^2 - 3\overline{(\mathbf{\Omega}\cdot\mathbf{S})}^4\} + \frac{1}{2}\left\{\int_{-\infty}^{+\infty}\Omega^2[(a_{\omega})^2 - 2aa_{\omega\omega}]\mathrm{d}\omega + 3\overline{(\mathbf{\Omega}\cdot\mathbf{S})}^2\int_{-\infty}^{+\infty}(a_{\omega})^2\mathrm{d}\omega - 2\overline{(\mathbf{\Omega}\cdot\mathbf{S})}\int_{-\infty}^{+\infty}(\mathbf{\Omega}\cdot\mathbf{S}) \times [(a_{\omega})^2 - 2aa_{\omega\omega}]\mathrm{d}\omega\right\} + \mu_{40}, \tag{36}$$

where $\mathbf{\Omega}_{\omega}$ means the first-order frequency derivative of the PMD vector $\mathbf{\Omega}$, a_{ω} and $a_{\omega\omega}$ are the first- and second-order frequency derivatives of the input pulse spectrum, and μ_{40} is the flatness of the input pulse. For a Gaussian pulse, $\mu_{40} = 3\sigma_0{}^4$. We write the skewness and flatness in the form of output PMD vectors and polarizations because they are easier to deal with than the input PMD vector through the generator given in Appendix A.

An approach similar to that used in Section 3 provides all the correlation functions that appear in the above expressions. It can be shown that the third-order correlation $\langle(\mathbf{\Omega}_1\cdot\mathbf{S}_1)(\mathbf{\Omega}_2\cdot\mathbf{S}_2)(\mathbf{\Omega}_3\cdot\mathbf{S}_3)\rangle = 0$. The PMD vector and its square are uncorrelated, $\langle\Omega_1{}^2(\mathbf{\Omega}_2\cdot\mathbf{S}_2)\rangle = 0$. The correlation of the cross product of the PMD vectors of two frequencies is zero as well, $\langle(\mathbf{\Omega}_1\times\mathbf{\Omega}_2)\cdot\mathbf{S}_1\rangle = 0$. A direct implication is that the ensemble average of the cross product of the PMD vector and its frequency derivative is zero,

$$\langle(\mathbf{\Omega}\times\mathbf{\Omega}_{\omega})\cdot\mathbf{S}\rangle = \lim_{\omega_2\to\omega_1}\frac{\partial}{\partial\omega_2}\langle(\mathbf{\Omega}_1\times\mathbf{\Omega}_2)\cdot\mathbf{S}_1\rangle = 0.$$

Using these results in Eq. (35), we found the ensemble average of the skewness to be zero, i.e., $\langle\mu_s\rangle = 0$. Thus, the PMD effects are equally likely to make the pulse asymmetric on the leading or the trailing side.

When the first-order PMD effects dominate and we can neglect the frequency dependence of the PMD vector, the main contribution to skewness comes from the first three cubic terms in Eq. (35), and this equation can be approximated as

$$\mu_3 \approx \frac{1}{8}\{3\Omega^2(\mathbf{\Omega}\cdot\mathbf{S}) - \Omega^2(\mathbf{\Omega}\cdot\mathbf{S}) - 2(\mathbf{\Omega}\cdot\mathbf{S})^3\} = \frac{1}{4}\Omega_{0x}(\Omega_{0y}{}^2 + \Omega_{0z}{}^2), \tag{37}$$

where we have written the skewness in terms of the input PMD vector, $\mathbf{\Omega}_0 = \Omega_{0x}\hat{e}_1 + \Omega_{0y}\hat{e}_2 + \Omega_{0z}\hat{e}_3$. Since the three components of the input PMD vector are independent Gaussian variables, the probability distribution of this approximated skewness can be in the following semianalytic form:

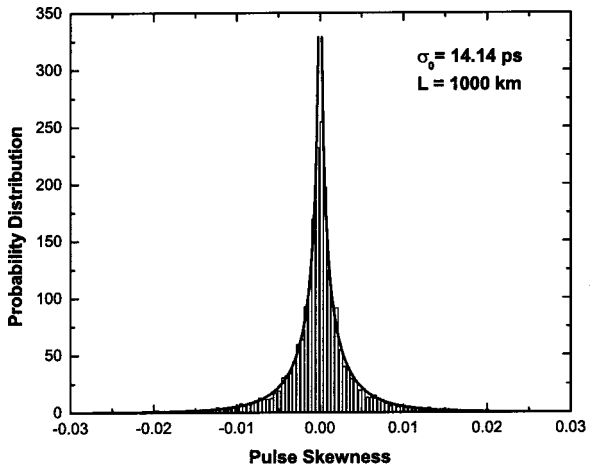

Fig. 1. Probability distribution of skewness at a distance of 1000 km. The skewness was normalized to $\sigma_0{}^3$. The rms width of the input Gaussian is $\sigma_0 = 14.14$ ps. The PMD parameter of the fiber is 0.1414 ps/$\sqrt{\text{km}}$. We realized 5000 random rounds. The solid curve represents the analytic results and the histogram represents numerical simulations.

$$P(\mu_3) = \sqrt{\frac{2}{\pi}}\left(\frac{3}{\langle\Omega^2\rangle}\right)^{3/2}\int_0^{+\infty}\frac{dy}{y} \times \exp\left\{-\frac{3}{2\langle\Omega^2\rangle}\left[y + \left(\frac{4\mu_3}{y}\right)^2\right]\right\}. \quad (38)$$

To validate the general theory presented in this paper, we solved Eq. (2) numerically. The PMD parameter D_p is assumed to be 0.1414 ps/$\sqrt{\text{km}}$ and the correlation length is assumed to be 1 km. The fiber is divided into many sections of 1-km length. Inside each section, birefringence is kept constant. At the end of each section, random rotation and random phase shift are induced. Fiber dispersion is assumed to be zero. The input pulse is taken to be an unchirped Gaussian pulse such that $\mathcal{A}(0,t) = A_0\exp(-t^2/4\sigma_0{}^2)\mathcal{S}_0$. Two different cases are considered here. In one case, a long input pulse with a full width at half-maximum of 33.3 ps ($\sigma_0 = 14.14$ ps) is propagated over a distance of 1000 km. Since $\sqrt{\langle\Omega^2\rangle}/\sigma_0 = 31.6\%$, PMD effects are relatively small. In the second case, a short input pulse with a full width at half-maximum of 8.33 ps ($\sigma_0 = 3.54$ ps) propagates over a distance of 5000 km. Since $\sqrt{\langle\Omega^2\rangle}/\sigma_0 = 2.83$, PMD effects are quite large. In each case, we solve Eq. (2) repeatedly 5000 times to collect the statistics for various moments. Figures 1 and 2 show the probability distributions for the skewness for the two cases by normalizing μ_3 with $\sigma_0{}^3$. The solid curve shows the analytic results from Eq. (38). In the case of large pulse width and small PMD, the theory agrees well with the numerical simulations because the first-order PMD effects dominate. In the second case of a small pulse width and large PMD, the theory and numerical simulations agree reasonably well. However, when the pulse width is reduced to less than 5 ps, the PMD becomes large enough and higher-order PMD effects become important enough that Eq. (38) begins to deviate from numerical simulations.

The probability distribution of the flatness can be considered by use of an approach similar to that used for the skewness. When the first-order PMD dominates, the main contributions to the flatness come from the fourth-order terms of the PMD vector in Eq. (36) and this equation becomes

$$\begin{aligned}\mu_4 &\approx \mu_{40} + \frac{1}{16}\{\Omega^4 - 4\Omega^2(\mathbf{\Omega}\cdot\mathbf{S})(\mathbf{\Omega}\cdot\mathbf{S}) \\ &\quad + 6\Omega^2(\mathbf{\Omega}\cdot\mathbf{S})^2 - 3(\mathbf{\Omega}\cdot\mathbf{S})^4\} \\ &= \mu_{40} + \frac{1}{16}(\Omega_{0y}{}^2 + \Omega_{0z}{}^2)(4\Omega_{0x}{}^2 + \Omega_{0y}{}^2 + \Omega_{0z}{}^2).\end{aligned} \quad (39)$$

Again using the fact that the three components of the input PMD vector are independent Gaussian random variables, we obtain the probability distribution of this approximated flatness:

$$P(\mu_4) = \sqrt{\frac{2}{\pi}}\left(\frac{3}{\langle\Omega^2\rangle}\right)^{3/2}\int_0^{4\sqrt{\mu_4-\mu_{40}}} \times \frac{dy\,\exp\{-3[16(\mu_4-\mu_{40}) + 3y^2]/(8y\langle\Omega^2\rangle)\}}{\sqrt{y[16(\mu_4-\mu_{40}) - y^2]}}. \quad (40)$$

Figure 3 shows the numerical simulated probability distribution for the PMD-induced flatness in the case of a long pulse ($\sigma_0 = 14.14$ ps) that propagates over 1000 km. The solid curve again shows the analytic prediction of our theory and agrees well with the results of numerical simulations. However, when higher-order PMD effects become significant, the numerical probability distribution of the flatness deviates significantly from Eq. (40). Numerical simulations show that, when pulses shorter than 5 ps propagate over a distance of 5000 km ($D_p = 0.1414$ ps), the usefulness of Eq. (40) becomes questionable.

Although the exact probability distribution of the PMD-induced flatness is difficult to find, the exact ensemble-

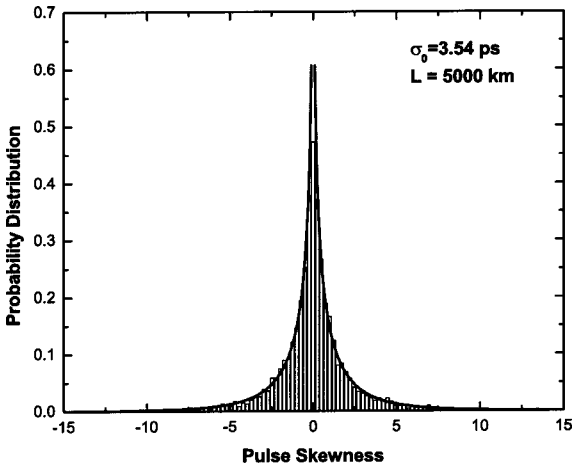

Fig. 2. Probability distribution of skewness at a distance of 5000 km. The skewness was normalized to $\sigma_0{}^3$. The rms width of the input Gaussian is $\sigma_0 = 3.54$ ps. The PMD parameter is the same as in Fig. 1. We realized 5000 random rounds. The solid curve represents the analytic results and the histogram represents numerical simulations.

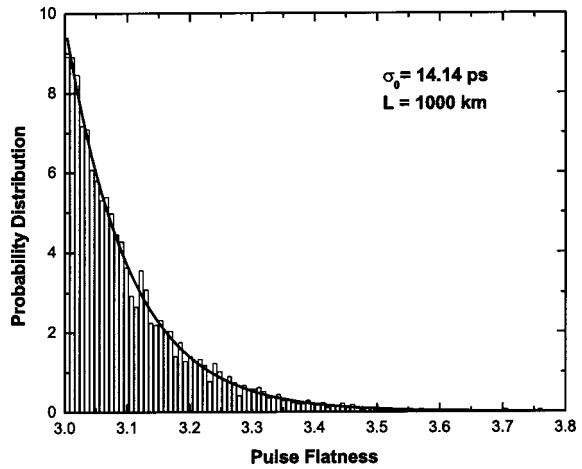

Fig. 3. Probability distribution of pulse flatness at a distance of 1000 km. The flatness was normalized to $\sigma_0{}^4$. The rms width of the input Gaussian pulse is $\sigma_0 = 14.14$ ps. The PMD parameter is the same as in Fig. 1. The solid curve represents the analytic results and the histogram represents numerical simulations.

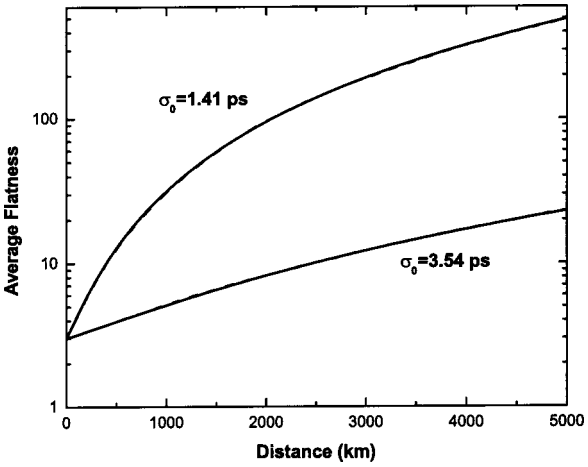

Fig. 4. Ensemble-averaged pulse flatness changes with propagation distance. The flatness is normalized to $\sigma_0{}^4$. The PMD parameter is the same as in Fig. 1. The solid curve represents the analytic theoretical results and the dashed curve represents the numerical simulations.

averaged value of the flatness can be obtained through the technique outlined in Appendix A. Using this approach, we can obtain the ensemble average of all the terms in Eq. (36). Appendix C provides more details. The average value of the flatness is found to be

$$\langle \mu_4 \rangle = \sigma_0{}^4 \left\{ \frac{27x^2}{16} + \frac{21x}{2} - (2x + 15)\sqrt{x+1} + \frac{3}{100}(5x - 4)\sqrt{10x+4} + \frac{16}{\sqrt{3}}\tan^{-1}(\sqrt{3x+3}) + \frac{456}{25} - \frac{16\pi}{3\sqrt{3}} \right\}, \tag{41}$$

where $x = \langle \Omega^2 \rangle / 3\sigma_0{}^2 = D_p{}^2 z / 3\sigma_0{}^2$ and z is the propagation distance.

Figure 4 shows the comparison between the analytic ensemble-averaged flatness in Eq. (41) and the results of numerical simulations for $\sigma_0 = 3.54$ ps and $\sigma_0 = \sqrt{2}$ ps. In each case, the optical pulse is assumed to propagate over a distance of 5000 km. The analytic theory indeed agrees well with the results of numerical simulations even for short pulses. The flatness grows almost quadratically with the propagation distance, which can be seen from Eq. (41) for $x \gg 0$. When the PMD effects are large, the main contribution to flatness comes from $\langle \overline{\Omega^4} \rangle$ and $\langle \overline{\Omega_\omega{}^2} \rangle$ in Eq. (36).

5. CONCLUSIONS

We have developed a general theory for describing the effects of PMD on an optical pulse that propagates inside a long fiber link. The theory includes the frequency dependence of the birefringence and the PMD vector to all orders. We use the Stratonovich calculus for calculating the correlations of the polarization and PMD vectors at two or more different frequencies. We first applied the general formalism to calculate the second-order correlation functions of the polarization vectors and that of the PMD vectors and showed that our results reduce to those obtained previously when only the first-order effects were retained.

To study the effect of PMD on the performance of an optical communication system, we focused on pulse broadening and the extent of pulse distortion characterized through the skewness and flatness related to the third- and fourth-order temporal moments of the pulse. We proved analytically that the even-order PMD effects do not affect the average value of the rms pulse width. We have calculated the probability distribution for the skewness and flatness and compared them with those obtained numerically. Our analytic results are in good agreement with numerical simulations.

APPENDIX A

Here we use the results from Ref. 18 to find the Stratonovich generator that allows us to calculate the average of an arbitrary function of the PMD and polarization vectors. Consider the following set of n differential equations satisfied by n random vectors:

$$\frac{d\mathbf{R}_i}{dz} = \vec{\vec{\mathbf{Q}}}_i \mathbf{g} + \mathbf{U}_i, \qquad i = 1, 2, \ldots, n, \tag{A1}$$

where $\vec{\vec{\mathbf{Q}}}_i$ is a 3×3 tensor and $\mathbf{U}_i$ is a three-dimensional vector. Both $\vec{\vec{\mathbf{Q}}}_i$ and $\mathbf{U}_i$ can be functions of $\mathbf{R}_i$. The vector $\mathbf{g}$ describes a three-dimensional Markovian–Gaussian stochastic process whose first two moments are given by

$$\langle \mathbf{g}(z) \rangle = 0, \qquad \langle \mathbf{g}(z_1)\mathbf{g}(z_2) \rangle = \eta^2 \vec{\vec{\mathbf{I}}} \delta(z_2 - z_1). \tag{A2}$$

Consider a smooth enough arbitrary function $\Psi(\{\mathbf{R}_i\})$ in a small interval $(z, z + \delta z)$. Expand $\Psi(z + \delta z)$ in a Taylor series to obtain (the dependence of Ψ on $\mathbf{R}_1, \mathbf{R}_2, \ldots$ is not shown explicitly)

$$\Psi(z+\delta z)-\Psi(z)=\frac{\mathrm{d}\Psi(z)}{\mathrm{d}z}\delta z+\frac{1}{2}\frac{\mathrm{d}^2\Psi(z)}{\mathrm{d}z^2}(\delta z)^2+O[(\delta z)^3], \tag{A3}$$

where

$$\frac{\mathrm{d}\Psi(z)}{\mathrm{d}z}=\sum_{i=1}^{n}\frac{\mathrm{d}\mathbf{R}_i}{\mathrm{d}z}\cdot\nabla_i\Psi(z), \tag{A4}$$

$$\frac{\mathrm{d}^2\Psi(z)}{\mathrm{d}z^2}=\sum_{i=1}^{n}\left[\frac{\mathrm{d}\mathbf{R}_i}{\mathrm{d}z}\cdot\nabla_i\left(\frac{\mathrm{d}\mathbf{R}_i}{\mathrm{d}z}\right)\cdot\nabla_i\Psi(z)+\frac{\partial}{\partial z}\left(\frac{\mathrm{d}\mathbf{R}_i}{\mathrm{d}z}\right)\cdot\nabla_i\Psi(z)\right]+\sum_{i,j=1}^{n}\frac{\mathrm{d}\mathbf{R}_j}{\mathrm{d}z}\frac{\mathrm{d}\mathbf{R}_i}{\mathrm{d}z}:\nabla_i\nabla_j\Psi(z), \tag{A5}$$

and the gradient is over $\mathbf{R}_i$ ($\nabla_i=\mathrm{d}/\mathrm{d}\mathbf{R}_i$). Since $\mathbf{g}(z)$ is random inside the interval $(z, z+\delta z)$, and $\Psi(z)$ is determined by the history within $(0, z)$, we need to average Eq. (A3) twice, first over the random variable $\mathbf{g}$ inside $(z, z+\delta z)$ and then over the history in the range $(0, z)$,[18] i.e.,

$$\frac{\mathrm{d}\langle\Psi(z)\rangle}{\mathrm{d}z}=\lim_{\delta z\to 0}\left\langle\frac{\langle\Psi(z+\delta z)\rangle_g-\Psi(z)}{\delta z}\right\rangle, \tag{A6}$$

where subscript g denotes averaging over $\mathbf{g}$ inside $(z, z+\delta z)$. Using Eqs. (A2)–(A5) and noting that all higher-order moments can be written in terms of the second-order moment for a Gaussian process, we obtain

$$\frac{\mathrm{d}\langle\Psi(z)\rangle}{\mathrm{d}z}=\langle G\{\Psi(z)\}\rangle, \tag{A7}$$

where the generator is given by

$$G=\lim_{\delta z\to 0}\left\langle\sum_{i=1}^{n}\frac{\mathrm{d}\mathbf{R}_i}{\mathrm{d}z}\cdot\nabla_i+\frac{\delta z}{2}\left\{\sum_{i=1}^{n}\left[\frac{\mathrm{d}\mathbf{R}_i}{\mathrm{d}z}\cdot\nabla_i\left(\frac{\mathrm{d}\mathbf{R}_i}{\mathrm{d}z}\right)\cdot\nabla_i+\frac{\partial}{\partial z}\left(\frac{\mathrm{d}\mathbf{R}_i}{\mathrm{d}z}\right)\cdot\nabla_i\right]+\sum_{i,j=1}^{n}\frac{\mathrm{d}\mathbf{R}_j}{\mathrm{d}z}\frac{\mathrm{d}\mathbf{R}_i}{\mathrm{d}z}:\nabla_i\nabla_j\right\}\right\rangle_g. \tag{A8}$$

We now apply this technique to the set of stochastic equations (6) and (7). The two equations can be rewritten as a single equation in the form of Eq. (A1) if we identify $\mathbf{R}_i$ as a six-dimensional vector $[\mathbf{S}_i;\mathbf{\Omega}_i]$, where $\mathbf{S}_i=\mathbf{S}(z,\omega_i)$ and $\mathbf{\Omega}_i=\Omega(z,\omega_i)$. Futhermore, $\mathbf{g}$ is a six-dimensional random vector with three components related to $\mathbf{b}$ and the others being zero, $\mathbf{g}(z)=[\mathbf{b}(z);0]$. $\mathbf{U}_i=0$ and $\vec{\vec{\mathbf{Q}}}_i$ is a 6×6 matrix defined as

$$\vec{\vec{\mathbf{Q}}}_i=\begin{pmatrix}-f(\omega_i)\mathbf{S}_i\times & 0\\ -f(\omega_i)\mathbf{\Omega}_i\times+f_\omega(\omega_i)\vec{\vec{\mathbf{I}}} & 0\end{pmatrix}. \tag{A9}$$

The cross-product operator $\times$ is defined in Ref. 11. We substituted Eq. (A1) into Eq. (A8) and averaged over $\mathbf{b}(z)$ in the interval $(z, z+\delta z)$ to obtain the following expression for the Stratonovich generator:

$$\begin{aligned}G=&-\eta^2\sum_{i=1}^{n}f_i^2(\mathbf{S}_i\cdot\nabla_{S_i}+\mathbf{\Omega}_i\cdot\nabla_{\Omega_i})\\&+\frac{\eta^2}{2}\sum_{i,j=1}^{n}\{f_if_j[(\mathbf{S}_i\cdot\mathbf{S}_j)\vec{\vec{\mathbf{I}}}-\mathbf{S}_i\mathbf{S}_j]:\nabla_{S_i}\nabla_{S_j}\\&+f_if_j[(\mathbf{\Omega}_i\cdot\mathbf{\Omega}_j)\vec{\vec{\mathbf{I}}}-\mathbf{\Omega}_i\mathbf{\Omega}_j]:\nabla_{\Omega_i}\nabla_{\Omega_j}\\&+f_{i\omega}f_{j\omega}\nabla_{\Omega_i}\cdot\nabla_{\Omega_j}+[f_if_{j\omega}\mathbf{\Omega}_i-f_jf_{i\omega}\mathbf{\Omega}_j]\cdot(\nabla_{\Omega_i}\times\nabla_{\Omega_j})\\&+f_if_j[(\mathbf{S}_i\cdot\mathbf{\Omega}_j)\vec{\vec{\mathbf{I}}}-\mathbf{S}_i\mathbf{\Omega}_j]:\nabla_{S_i}\nabla_{\Omega_j}\\&+f_if_j[(\mathbf{\Omega}_i\cdot\mathbf{S}_j)\vec{\vec{\mathbf{I}}}-\mathbf{\Omega}_i\mathbf{S}_j]:\nabla_{\Omega_i}\nabla_{S_j}\\&+f_if_{j\omega}\mathbf{S}_i\cdot(\nabla_{S_i}\times\nabla_{\Omega_j})-f_jf_{i\omega}\mathbf{S}_j\cdot(\nabla_{\Omega_i}\times\nabla_{S_j})\},\end{aligned} \tag{A10}$$

where subscripts S_i and Ω_i denote the gradient on $\mathbf{S}_i$ and $\mathbf{\Omega}_i$, respectively.

APPENDIX B

We now give details about the derivation of Eqs. (35) and (36) for the skewness and flatness used in this paper. The propagation of optical fields can be described by use of the Jones matrix as $\mathcal{A}(z,\omega)=\mathbf{T}(z,\omega)\mathcal{A}(0,\omega)$, where $\mathbf{T}$ is the transfer matrix and is unitary for a lossless media. Consider an unchirped pulse so that $\mathcal{A}(0,\omega)=a(\omega)\mathcal{S}_0$, where $\mathcal{S}_0$ is the input polarization and is frequency independent. Assume there is no fiber dispersion, so that $\tau(z,\omega)=0$ in Eq. (5). The frequency derivative of $\mathbf{T}$ relates to the PMD vector directly[11] as shown in Eq. (5):

$$\mathbf{T}_\omega=-\frac{i}{2}[\mathbf{\Omega}(z,\omega)\cdot\boldsymbol{\sigma}]\mathbf{T}, \tag{B1}$$

$$\mathbf{T}_{\omega\omega}=\frac{1}{4}\{-\Omega^2(z,\omega)-2i\mathbf{\Omega}_\omega(z,\omega)\cdot\boldsymbol{\sigma}\}\mathbf{T}. \tag{B2}$$

Since $\mathcal{A}_\omega(z,\omega)=\mathbf{T}_\omega\mathcal{A}(0,\omega)+\mathbf{T}\mathcal{A}_\omega(0,\omega)$ and $\mathcal{A}_{\omega\omega}(z,\omega)=\mathbf{T}_{\omega\omega}\mathcal{A}(0,\omega)+2\mathbf{T}_\omega\mathcal{A}_\omega(0,\omega)+\mathbf{T}\mathcal{A}_{\omega\omega}(0,\omega)$, by substituting Eqs. (B1) and (B2) into Eqs. (33) and (34), we can get the analytic expressions for the skewness and flatness in terms of the output PMD vector $\mathbf{\Omega}(z,\omega)$ and its frequency derivative $\mathbf{\Omega}_\omega(z,\omega)$. The final result is given in Eqs. (35) and (36).

APPENDIX C

We now provide more details about the derivation of Eq. (41) from Eq. (36).

Using the generator in Appendix A, we can find the dynamic equations for the correlation functions $\langle\Omega^4\rangle$, $\langle{\Omega_1}^2(\mathbf{\Omega}_2\cdot\mathbf{S}_2)(\mathbf{\Omega}_3\cdot\mathbf{S}_3)\rangle$, and

$$\langle(\mathbf{\Omega}_1\cdot\mathbf{S}_1)(\mathbf{\Omega}_2\times\mathbf{\Omega}_{2\omega})\cdot\mathbf{S}_2\rangle=\lim_{\omega_3\to\omega_2}\frac{\partial}{\partial\omega_3}\langle(\mathbf{\Omega}_1\cdot\mathbf{S}_1)(\mathbf{\Omega}_2\times\mathbf{\Omega}_3)\cdot\mathbf{S}_2\rangle. \tag{C1}$$

After averaging over the input pulse spectrum, we obtain

$$\langle \overline{\Omega^4} \rangle = \int_{-\infty}^{+\infty} \langle \Omega^4(z,\ \omega) \rangle a^2(\omega) \mathrm{d}\omega = 15\sigma_0{}^4 x^2, \quad \text{(C2)}$$

$$\langle \overline{\Omega_\omega{}^2} \rangle = \int_{-\infty}^{+\infty} \mathrm{d}\omega_1 a^2(\omega_1) \lim_{\omega_2 \to \omega_1} \partial^2 \langle \mathbf{\Omega}_1 \cdot \mathbf{\Omega}_2 \rangle / \partial\omega_1 \partial\omega_2 = 3\sigma_0{}^4 x^2, \quad \text{(C3)}$$

where $x = \langle \Omega^2 \rangle / 3\sigma_0{}^2$. The other correlations are more complicated and are given by

$$\langle (\overline{\mathbf{\Omega}\cdot\mathbf{S}})(\overline{\mathbf{\Omega}\times\mathbf{\Omega}_\omega}) \cdot \mathbf{S} \rangle = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \mathrm{d}\omega_1 \mathrm{d}\omega_2 a^2(\omega_1) a^2(\omega_2) \langle (\mathbf{\Omega}_1 \cdot \mathbf{S}_1)(\mathbf{\Omega}_2 \times \mathbf{\Omega}_{2\omega}) \cdot \mathbf{S}_2 \rangle = \frac{2\sigma_0{}^4 x}{\sqrt{x+1}} (2\sqrt{x+1} - x - 2), \quad \text{(C4)}$$

$$\langle \overline{\Omega^2(\mathbf{\Omega}\cdot\mathbf{S})}(\overline{\mathbf{\Omega}\cdot\mathbf{S}}) \rangle = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \mathrm{d}\omega_1 \mathrm{d}\omega_2 a^2(\omega_1) a^2(\omega_2) \langle \Omega_1{}^2 (\mathbf{\Omega}_1 \cdot \mathbf{S}_1)(\mathbf{\Omega}_2 \cdot \mathbf{S}_2) \rangle = 10\sigma_0{}^4 x(\sqrt{x+1} - 1), \quad \text{(C5)}$$

$$\langle \overline{\Omega^2}(\overline{\mathbf{\Omega}\cdot\mathbf{S}})^2 \rangle = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \mathrm{d}\omega_1 \mathrm{d}\omega_2 \mathrm{d}\omega_3 a^2(\omega_1) a^2(\omega_2) a^2(\omega_3) \langle \Omega_1{}^2 (\mathbf{\Omega}_2 \cdot \mathbf{S}_2)(\mathbf{\Omega}_3 \cdot \mathbf{S}_3) \rangle$$
$$= \sigma_0{}^4 \left\{ \frac{2948}{75} - \frac{128\pi}{9\sqrt{3}} + 10x + \left(4x - \frac{116}{3}\right)\sqrt{x+1} + \frac{2}{25}(5x-4)\sqrt{10x+4} + \frac{128}{3\sqrt{3}} \tan^{-1}[\sqrt{3(x+1)}] \right\}. \quad \text{(C6)}$$

Several other integrals appear in Eq. (36), all of which can be evaluated analytically and were found to be

$$\int_{-\infty}^{+\infty} \langle \Omega^2 \rangle [(a_\omega)^2 - 2a a_{\omega\omega}] \mathrm{d}\omega = 9\sigma_0{}^4 x, \quad \text{(C7)}$$

$$\langle (\overline{\mathbf{\Omega}\cdot\mathbf{S}})^2 \rangle \int_{-\infty}^{+\infty} (a_\omega)^2 \mathrm{d}\omega = 2\sigma_0{}^4 (\sqrt{x+1} - 1), \quad \text{(C8)}$$

$$\left\langle (\overline{\mathbf{\Omega}\cdot\mathbf{S}}) \int_{-\infty}^{+\infty} (\mathbf{\Omega}\cdot\mathbf{S})[(a_\omega)^2 - 2a a_{\omega\omega}] \mathrm{d}\omega \right\rangle = \sigma_0{}^4 \left(\frac{7x+8}{\sqrt{x+1}} - 8 \right). \quad \text{(C9)}$$

The term $\langle (\overline{\mathbf{\Omega}\cdot\mathbf{S}})^4 \rangle$ is determined by the correlation function $\langle (\mathbf{\Omega}_1 \cdot \mathbf{S}_1)(\mathbf{\Omega}_2 \cdot \mathbf{S}_2)(\mathbf{\Omega}_3 \cdot \mathbf{S}_3)(\mathbf{\Omega}_4 \cdot \mathbf{S}_4) \rangle$. Although the dynamic equation associated with this correlation function can be obtained, it can be solved only numerically because of its complexity. The moment theorem shows that, if a random process is Gaussian, its derivative should also be a Gaussian random process.[23] The non-Gaussian sech probability distribution of the second-order PMD vector[5–8] implies that the PMD vector is not necessarily a Gaussian random process in frequency domain when higher-order effects are included, although the probability distribution of the PMD vector of each frequency is exactly a three-dimensional independent Gaussian. For the same reason, the term $\langle (\overline{\mathbf{\Omega}\cdot\mathbf{S}})(\overline{\mathbf{\Omega}\times\mathbf{\Omega}_\omega}) \cdot \mathbf{S} \rangle$ can be nonzero. However, numerical simulations show that the moment theorem holds fairly well for the spectrum-averaged correlations. For example, $\langle (\overline{\mathbf{\Omega}\cdot\mathbf{S}})^4 \rangle \approx 3\langle (\overline{\mathbf{\Omega}\cdot\mathbf{S}})^2 \rangle^2 = 12\sigma_0{}^4 (\sqrt{x+1} - 1)^2$. We used this approximation in Eq. (41).

ACKNOWLEDGMENTS

This research is supported by the National Science Foundation under grants ECS-9903580 and DMS-0073923.

Q. Lin's e-mail address is linq@optics.rochester.edu.

REFERENCES

1. C. D. Poole, "Polarization effects in lightwave systems," in *Optical Fiber Telecommunications III A*, I. P. Kaminow and T. L. Koch, eds. (Academic, San Diego, Calif., 1997), Chap. 6.
2. H. Kogelnik, R. Jopson, and L. Nelson, "Polarization-mode dispersion," in *Optical Fiber Telecommunications IV B*, I. P. Kaminow and T. L. Koch, eds. (Academic, San Diego, Calif., 2002), Chap. 15.
3. P. Ciprut, B. Gisin, N. Gisin, R. Passy, J. P. Von der Weid, F. Prieto, and C. W. Zimmer, "Second-order polarization mode dispersion: impact on analog and digital transmissions," J. Lightwave Technol. **16**, 757–771 (1998).
4. C. D. Poole and R. E. Wagner, "Phenomenological approach to polarisation dispersion in long single-mode fibers," Electron. Lett. **22**, 1029–1030 (1986).
5. G. J. Foschini and C. D. Poole, "Statistical theory of polarization dispersion in single mode fibers," J. Lightwave Technol. **9**, 1439–1456 (1991).
6. G. J. Foschini, R. M. Jopson, L. E. Nelson, and H. Kogelnik, "The statistics of PMD-induced chromatic fiber dispersion," J. Lightwave Technol. **17**, 1560–1565 (1999).
7. G. J. Foschini, L. E. Nelson, R. M. Jopson, and H. Kogelnik, "Probability densities of second-order polarization mode dispersion including polarization dependent chromatic fiber dispersion," IEEE Photon. Technol. Lett. **12**, 293–295 (2000).
8. G. J. Foschini, L. E. Nelson, R. M. Jopson, and H. Kogelnik, "Statistics of second-order PMD depolarization," J. Lightwave Technol. **19**, 1882–1886 (2001).
9. M. Karlsson and J. Brentel, "Autocorrelation function of the polarization-mode dispersion vector," Opt. Lett. **24**, 939–941 (1999).
10. M. Shtaif, A. Mecozzi, and J. A. Nagel, "Mean-square magnitude of all orders of polarization mode dispersion and the relation with the bandwidth of the principal states," IEEE Photon. Technol. Lett. **12**, 53–55 (2000).
11. J. P. Gordon and H. Kogelnik, "PMD fundamentals: polarization mode dispersion in optical fibers," Proc. Natl. Acad. Sci. USA **97**, 4541–4550 (2000).

12. G. P. Agrawal, *Fiber Optic Communication Systems*, 3rd ed. (Wiley, New York, 2002), Chap. 2.
13. C. D. Poole, J. H. Winters, and J. A. Nagel, "Dynamical equation for polarization dispersion," Opt. Lett. **16**, 372–374 (1991).
14. 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. **8**, 1162–1166 (1990).
15. N. Gisin and J. P. Pellaux, "Polarization mode dispersion: time versus frequency domains," Opt. Commun. **89**, 316–323 (1992).
16. H. Risken, *The Fokker–Planck Equation*, 2nd ed. (Springer-Verlag, New York, 1989).
17. C. W. Gardiner, *Handbook of Stochastic Methods*, 2nd ed. (Springer-Verlag, New York, 1985).
18. P. K. A. Wai and C. R. Menyuk, "Polarization mode dispersion, decorrelation, and diffusion in optical fibers with randomly varying birefringence," J. Lightwave Technol. **14**, 148–157 (1996).
19. M. Karlsson, J. Brentel, and P. A. Andrekson, "Long-term measurement of PMD and polarization drift in installed fibers," J. Lightwave Technol. **18**, 941–951 (2000).
20. M. Karlsson, "Polarization mode dispersion-induced pulse broadening in optical fibers," Opt. Lett. **23**, 688–690 (1998).
21. H. Sunnerud, M. Karlsson, and P. A. Andrekson, "Analytical theory for PMD-compensation," IEEE Photon. Technol. Lett. **12**, 50–52 (2000).
22. H. Bülow, "System outage probability due to first- and second-order PMD," IEEE Photon. Technol. Lett. **10**, 696–698 (1998).
23. M. C. Wang and G. E. Uhlenbeck, "On the theory of the Brownian motion II," in *Selected Papers on Noise and Stochastic Processes*, N. Wax, ed. (Dover, New York, 1954), pp. 113–132.