Statistical Characterization of the Jones Matrix of Long Fibers Affected by Polarization Mode Dispersion (PMD)

Armando Vannucci, *Member, IEEE,* and Alberto Bononi

***Abstract*—The unitary transfer matrix of a fiber affected by polarization mode dispersion (PMD) is analyzed using the Stokes representation of its eigenmodes and its retardation angle, or equivalently through its Pauli coordinates. We develop a statistical theory applied to these parameters and relate it to the extensive existing literature on the statistics of the PMD vector $\vec{\Omega}_o(\omega)$. Dynamical equations are established for the Pauli coordinates. Assuming a standard "white Gaussian" model for the local birefringence, and using the tools of stochastic calculus, we derive the distributions of the eigenmodes, the retardation angle, the Pauli coordinates, and of the frequency derivatives of all these parameters. The evolution in space of the Pauli coordinates is also characterized as a standard Brownian motion on the unit sphere in $\Re^4$. An expression for the frequency autocorrelation function of the Pauli coordinates, the eigenmodes and the retardation angle is derived and their coherence bandwidth is compared to that of the PMD vector. All theoretical results are supported by simulation over an ensemble of 10 000 fibers, using the standard retarder plate model.**

***Index Terms*—Optical fiber communication, optical polarization, polarization mode dispersion.**

I. Introduction

THE stochastic analysis of fibers affected by polarization mode dispersion (PMD) has so far concentrated on the statistics of the PMD vector and its frequency derivative [1]–[3]. The interest in the PMD vector is due to its stability in frequency over the bandwidths of interest in optical communications at bit-rates of 10 Gb/s or lower. However, the upgrade to the new 40 Gb/s digital hierarchy standard over installed single mode fibers (SMFs) brings about a strong depolarization of the PMD vector so that its higher order frequency derivatives, the so-called "higher order PMD," become increasingly important. In this new scenario, the analysis of the output field based on the PMD vector may not be any longer the simplest mathematical description [4]. This motivated us to derive a stochastical description of the fiber Jones matrix and its frequency derivatives, highlighting its relationship to the PMD vector description.

In order to describe the Jones transfer matrix in a compact form, we will resort to its decomposition over the basis of Pauli matrices. Our first objective is to derive dynamical equations for the space evolution of the Pauli coordinates, in a similar way to what was done for the dynamical equation of the PMD vector [1], [2]. Once a suitable model for the local birefringence vector has been selected, such dynamical equations are tackled with standard tools of stochastic calculus [5]: i) the Fokker–Planck equation (FPE) is used to derive some useful probability density functions (pdf); ii) the Dynkin formula, also known as the martingale differential equation (MDE), is used to derive some moments of interest, without knowledge of the related pdf; iii) the characterization of canonical Brownian motion on spheres is used to recognize the nature of the spatial evolution of the Pauli coordinates. A rigorous in-depth treatment of such concepts is provided, e.g., in [5]. However, a brief tutorial introduction, in the context of PMD, can be found in the appendices of [1], [3].

On Notation: Throughout the paper, matrices are denoted by capital letters, three-dimensional vectors by an arrow, and unit magnitude vectors by a hat. Four-dimensional real vectors are underlined. A column vector with elements a, b, c is denoted by $[a;\, b;\, c]$. The symbols $\cdot$ and $\times$ denote vector scalar and cross product. I is the 3×3 identity matrix while σ_0 is the 2×2 identity matrix. Ensemble averaging is denoted by $E[\cdot]$.

II. Jones Matrix of a Fiber Affected by PMD

Using the Jones formalism to describe the state of polarization (SOP) of a given optical field, a linear fiber of length z affected by PMD but without polarization dependent loss can be described as a two-input–two-output linear system whose transfer matrix is the product of a scalar function and a unitary matrix $U(z,\, \omega)$ which has unit determinant.

We can compactly express U as a matrix exponential as suggested by Karlsson [6]:

$$U = \mathrm{e}^{-i(\Delta\phi/2)[\hat{b}\cdot\vec{\sigma}]} = \cos\left(\frac{\Delta\phi}{2}\right)\sigma_0 - i\sin\left(\frac{\Delta\phi}{2}\right)\left[\hat{b}\cdot\vec{\sigma}\right] \tag{1}$$

where $\hat{b} = [b_1;\, b_2;\, b_3]$ is the Stokes unit vector associated with the "slow" eigenvector of U, $\Delta\phi$ is the *retardation angle* and $\vec{\sigma} = [\sigma_1;\, \sigma_2;\, \sigma_3]$ is the so called *spin vector*, whose components are the Pauli matrices 1, 2, 3, as defined in [7], so that the scalar product $\hat{b}\cdot\vec{\sigma}$ is shorthand for $\sum_{i=1}^{3} b_i\sigma_i$, and represents a matrix. Equation (1) shows that the decomposition of U on the basis of Pauli matrices has coordinates

$$u_0 \triangleq \cos\left(\frac{\Delta\phi}{2}\right) \qquad i\vec{u} \triangleq -i\sin\left(\frac{\Delta\phi}{2}\right)\hat{b} \tag{2}$$

Manuscript received May 14, 2001; revised December 4, 2001. This work was supported by Alcatel CRC, Marcoussis, France, as part of the Research Project URP/3.99.

The authors are with the Dipartimento di Ingegneria dell'Informazione, Università di Parma, 43100 Parma, Italy.

Publisher Item Identifier S 0733-8724(02)03200-0.

where the *Pauli coordinates vector*[1] $\underline{u} \triangleq [u_0; \vec{u}]$ is a (column) vector in $\Re^4$. Since from (2) $|\underline{u}|^2 = 1$, the coordinates u_k are confined on the surface of a sphere with unit radius in $\Re^4$, that we will indicate with $\mathcal{S}^4$.

The Müller matrix associated with (1) is [6]

$$\begin{aligned}\mathcal{M} &= \mathrm{e}^{\Delta\phi[\hat{b}\times]} \\ &= I + \sin(\Delta\phi)\left[\hat{b}\times\right] + (1-\cos(\Delta\phi))\left[\hat{b}\times\right]\left[\hat{b}\times\right] \quad (3)\end{aligned}$$

where $[\hat{b}\times]$ is the skew-symmetric matrix corresponding to the vector cross product operation. When (3) is applied to a Stokes vector, it causes a counterclockwise rotation of such vector about the $\hat{b}$ axis by an angle $\Delta\phi$, thus giving a geometrical meaning to the eigenmodes and the retardation angle.[2]

Our main motivation for the study of the Pauli coordinates comes from the fact that an exact closed-form expression exists of the photodetected field intensity in an optical communication system. It is proven in [8] that the received intensity can be expressed as

$$I(t) = N_0(t) + \vec{N}(t)\cdot\hat{j} \quad (4)$$

where $\hat{j}$ is the input Stokes SOP, and $N_0(t)$, $\vec{N}(t)$ are calculated using the Pauli coordinates as

$$\begin{aligned}N_0(t) &\triangleq \left|\underline{u}(t)\otimes\overline{E}_o(t)\right|^2 \\ \vec{N}(t) &\triangleq 2\Re\left\{\left(u_0(t)\otimes\overline{E}_o(t)\right)\left(\vec{u}(t)\otimes\overline{E}_o(t)\right)^*\right\} \\ &\quad - i\left(\vec{u}(t)\otimes\overline{E}_o(t)\right)\times\left(\vec{u}(t)\otimes\overline{E}_o(t)\right)^* \quad (5)\end{aligned}$$

where $\otimes$ denotes convolution, $\Re\{\cdot\}$ the real part, * is complex conjugate, and $\overline{E}_o$ is the scalar input field, filtered by common mode chromatic dispersion.

III. The Pauli Coordinates Distribution

The local birefringence vector $\vec{W}(z,\omega)$ governs the spatial evolution of the unitary Jones matrix (see, e.g., [7, eq. (6.4)])

$$\frac{d}{dz}U(z,\omega) = -\frac{i}{2}\left[\vec{W}(z,\omega)\cdot\vec{\sigma}\right]U(z,\omega) \quad (6)$$

where $U(z,\omega) = u_0(z,\omega)\sigma_0 + i\vec{u}(z,\omega)\cdot\vec{\sigma}$ is completely represented by its Pauli vector $\underline{u}(z,\omega) = [u_0(z,\omega); \vec{u}(z,\omega)]$, as per (2). From the properties of Pauli matrices, it can be shown that the Pauli coordinates of the product of matrices U_1 and U_2 with coordinates $\underline{u}_1$ and $\underline{u}_2$ are

$$[u_{01}u_{02} - \vec{u}_1\cdot\vec{u}_2;\ [u_{01}\vec{u}_2 + u_{02}\vec{u}_1 - \vec{u}_1\times\vec{u}_2]]. \quad (7)$$

Applying such result to (6), we get the dynamical equations governing the evolution of $\underline{u}$

$$\begin{cases}\dfrac{du_0}{dz} &= \dfrac{1}{2}\vec{u}\cdot\vec{W} \\ \dfrac{d\vec{u}}{dz} &= -\dfrac{1}{2}u_0\vec{W} - \dfrac{1}{2}\vec{u}\times\vec{W}.\end{cases} \quad (8)$$

[1]For brevity, we will refer to $\underline{u}$ as the *Pauli vector* and to its elements u_k as the *Pauli coordinates*.

[2]Note that one might be tempted to deal with the vector $\vec{\chi} \triangleq \Delta\phi\hat{b}$, by analogy with the PMD vector $\vec{\Omega}_o = \Delta\tau\hat{q}$. Indeed it is possible to derive a dynamical equation for $\vec{\chi}$ similar to (8), which however involves nonlinear functions of its elements, making such equation readily intractable.

There are several models adopted in the literature for $\vec{W}(z,\omega)$, and the choice of each model has nontrivial implications. For analytical simplicity, we choose the model adopted by Gisin *et al.* [2] for analyzing "long" fibers, which assumes $\vec{W}(z,\omega) = \omega\mu\vec{n}(z)$, where $\vec{n}(z)$ is a three-dimensional Gaussian process with independent and identically distributed components, each being a scalar white Gaussian process with zero mean and unit variance. The scalar parameter μ influences the birefringence variance. A discussion of this choice is provided in Appendix A.

With such model, we must be careful to interpret the stochastic differential equation (SDE) system (8) in the proper way: since an unbounded white noise term appears, all the products involving $\vec{n}(z)$ must be interpreted as Stratonovich products [1], [5]. Denoting with $d\vec{\nu}(z) = \vec{n}(z)\,dz$ the differential of standard three-dimensional Brownian motion and applying the Stratonovich rules for the evaluation of products,[3] we recast the dynamical system (8) in its *Ito form*

$$\begin{cases}du_0 &= -\dfrac{3\omega^2\mu^2}{8}u_0\,dz + \dfrac{\omega\mu}{2}\vec{u}\cdot d\vec{\nu} \\ d\vec{u} &= -\dfrac{3\omega^2\mu^2}{8}\vec{u}\,dz - \dfrac{\omega\mu}{2}u_0\,d\vec{\nu} - \dfrac{\omega\mu}{2}\vec{u}\times d\vec{\nu}\end{cases} \quad (9)$$

or, equivalently, in the *canonical form* [5]

$$\frac{d\underline{u}}{dz} = \underline{c}(z,\underline{u}) + V(z,\underline{u})\vec{n}(z) \quad (10)$$

where the *drift* $\underline{c}$ is a 4×1 vector, and the *diffusion* V [5] is a 4×3 matrix, defined as follows:

$$\underline{c} \triangleq -\frac{3\omega^2\mu^2}{8}\underline{u} \qquad V \triangleq -\frac{\omega\mu}{2}\begin{bmatrix}-\vec{u}^T \\ u_0 I + \vec{u}\times\end{bmatrix}. \quad (11)$$

We now apply the Fokker–Planck equation (FPE) [5] to the Ito system (10). This powerful mathematical tool is used to derive the distribution of a vector process, given the SDE governing its evolution, by solving a partial differential equation. It has been used, e.g., in [1] to obtain the distribution of the PMD vector, and is briefly discussed in Appendix A in [1]. In our case, the FPE describes the evolution along z of the joint probability density function (pdf) $p(u_0, u_1, u_2, u_3, z)$ of the processes u_k

$$\begin{aligned}\frac{\partial p(\underline{u},z)}{\partial z} &= -\sum_{k=0}^{3}\frac{\partial(c_k(\underline{u})p(\underline{u},z))}{\partial u_k} \\ &\quad + \frac{1}{2}\sum_{i,j=0}^{3}\frac{\partial^2(a_{ij}(\underline{u})p(\underline{u},z))}{\partial u_i\partial u_j} \quad (12)\end{aligned}$$

where the elements of the drift are $c_k = -(3\omega^2\mu^2/8)u_k$, and the matrix $A = \{a_{ij}\} \triangleq VV^T$ in our case becomes $A = (\omega^2\mu^2/4)(|\underline{u}|^2 I - \underline{u}\,\underline{u}^T)$.

As for any partial differential equation, there is no standard method for solving the FPE. However, we next prove that a pdf $p(\underline{u},z) = p(z)$ which does not depend on $\underline{u}$ can be a solution of (12). Such a pdf corresponds to a uniform distribution for $\underline{u}$ on its domain of definition $\mathcal{S}^4$. For the proof, the FPE can be evaluated using the drift coefficient $\underline{c}(\underline{u})$ and the matrix A,

[3]For a discussion of the Ito–Stratonovich duality in the context of PMD, see [1, Appendix A], [3, Appendix], and [2, Appendix D], reporting the rule for Stratonovich products.

whose diagonal elements a_{ii} do not depend on u_i and whose off-diagonal terms are simply $a_{ij} = -(\omega^2\mu^2/4)u_iu_j$. Equation (12) thus reduces to $\partial p(z)/\partial z = 0$, whose solution is a constant pdf along z, and does not depend on $\underline{u}$. It is then proved that the uniform distribution $p(\underline{u}) = (1/2\pi^2)$, $\underline{u} \in \mathcal{S}^4$ and zero elsewhere, is the FPE solution giving the joint pdf of the Pauli coordinates $\underline{u}$, since $2\pi^2$ is the hyper-surface of $\mathcal{S}^4$, as shown in Appendix B.

Note that the solution does not depend on z, meaning that $\underline{u}$ reaches its steady state uniform distribution immediately after the fiber input. From a mathematical standpoint, this is due to the unbounded white noise term that models the local birefringence (see Appendix A for a discussion). This model is also the cause of the discontinuity in the distribution of $\underline{u}(z, \omega)$ at $z = 0$ (recall that $\underline{u}(0, \omega)$ is deterministic and equal to $[1; 0; 0; 0]$). A transient behavior in z would appear if the birefringence were a colored bounded-variance stochastic process, similarly to the transients for the PMD vector studied in [2, Appendix C]. However, we are interested in *long* single mode fibers, i.e., hundreds of times longer than the correlation length. On such length scales, the variations of the birefringence in z can be suitably modeled as white noise, as the good match between theory and simulations will show.

As shown in Appendix B, the spherical coordinates $(\theta, \varepsilon, \Delta\phi)$ of $\underline{u}$ are found to be independent random variables (RVs). Thus, at fixed z and ω, the eigenmode $\hat{b} = [\cos(2\theta)\cos(2\varepsilon); \sin(2\theta)\cos(2\varepsilon); \sin(2\varepsilon)]$ is independent of $\Delta\phi(z, \omega)$. The marginal densities of the azimuth θ and the ellipticity ε are

$$p(\theta) = \frac{1}{\pi}\left(|\theta| \le \frac{\pi}{2}\right) \qquad p(\varepsilon) = \cos(2\varepsilon)\left(|\varepsilon| \le \frac{\pi}{4}\right) \quad (13)$$

which give a uniform distribution of $\hat{b}$ on the Poincaré sphere, a result already given in [9] without proof. Moreover, we also get the pdf of the fiber retardation angle $\Delta\phi$

$$p(\Delta\phi) = \frac{1 - \cos(\Delta\phi)}{2\pi} \qquad 0 \le \Delta\phi \le 2\pi \quad (14)$$

which is symmetric around the mean value $E[\Delta\phi] = \pi$. We can then state that, "in the average," a long fiber operates a π rotation around a completely random axis on any input SOP. Since the Euclidean coordinates u_k, $k = 0, \ldots, 3$, are a transformation of the RVs $(\theta, \varepsilon, \Delta\phi)$, they are found to have identical marginal pdf

$$p(u_k) = \frac{2}{\pi}\sqrt{1 - u_k^2}\,(-1 \le u_k \le 1) \quad (15)$$

and hence $E[u_k(z, \omega)] = 0$, $E[u_k^2(z, \omega)] = 1/4$, for every $z > 0$, which also implies $E[U] = 0$, i.e., a zero average value for the Jones matrix.

To quote just one possible application of the obtained results, in [10, Sec. V] the authors discuss the coarse step method for solving the Coupled Nonlinear Schrödinger equation, which requires, at a certain step of the method, a uniform scattering of the field on the Poincaré sphere. The concatenation of several simulated Jones matrices [10, Eqs. (34) and (35)] is suggested to produce a uniform *mixing* of the field. Our results suggest instead that such goal is achieved by simulating a single Jones matrix with the distributions of θ, ε and $\Delta\phi$ given in (13) and (14).

A. Pauli Coordinates Incorrelation and Evolution in Length

The Cartesian coordinates u_k are not independent. Nonetheless, we can prove that they are uncorrelated resorting to the Dynkin formula [5, Ch. 7]. This important mathematical tool of stochastic calculus allows the evaluation of a given moment $E[f(\underline{u}(z))]$ of the vector process $\underline{u}(z)$, without knowledge of the related pdf, by solving the differential equation $(\partial E[f(\underline{u}(z))]/\partial z) = E[G\{f(\underline{u}(z))\}]$. Here, $G\{\cdot\}$ is a second-order differential operator called the *Ito generator*, defined on the dynamical system (10), that depends on the drift $\underline{c}$ and on the matrix A defined below (12). The Dynkin formula is briefly introduced in [1, Appendix A] with the alternative name of MDE, along with the Ito generator, in order to evaluate some moments of the PMD vector. Reference [3, Appendix] also discusses the related concept of Stratonovich generator, which is applied to SDEs in the alternative *Stratonovich form* [3, eq. (18)]. In our case, defining $f(\underline{u}) = u_iu_j$, with $i \ne j$, the application of the Ito generator to $f(\underline{u})$ yields $G\{u_iu_j\} = -\omega^2\mu^2u_iu_j$ and the Dynkin formula gives $(\partial E[u_iu_j]/\partial z) = -\omega^2\mu^2E[u_iu_j]$, whose solution is zero: $E[u_i(z, \omega)u_j(z, \omega)] = \mathrm{e}^{-\omega^2\mu^2 z}E[u_i(0, \omega)u_j(0, \omega)] = 0$ for any (z, ω) because for a zero length fiber, $u_0(0, \omega) = 1$, $u_k(0, \omega) = 0 (k = 1, 2, 3)$, thus showing the incorrelation of the Pauli coordinates u_k.

A further important result is that the dynamical equations (10) imply a *Brownian motion on the unit sphere* $\mathcal{S}^4$ for the evolution along z of the fiber coordinates $u_k(z, \omega)$. It is possible to show that the normalized Brownian motion in $\Re^4$ *coincides in law with (i.e., has the same finite-dimensional distributions as* [5]*)* the process $\underline{u}(z)$ of our interest, by applying [5, Th. 8.4.3], which requires the comparison of the drift and diffusion coefficients of the processes. The expressions of $\underline{c}(\underline{u})$ and $V(\underline{u})$ in (11) satisfy the conditions of such theorem, after a suitable change of the z scale which depends on ω, thus proving that the process $\underline{u}(z, \omega)$, for every frequency ω, coincides in law with a four dimensional Brownian motion on the unit sphere and that the motion evolves at different rates, along z, for different frequency components.

B. Check by Simulation

A set of 10 000 fibers was simulated and the parameters $\underline{u}$, $\Delta\phi$ and $\hat{b}$ were measured at the reference frequency. For fiber simulation we used the standard retarder plate model (see, e.g., [11]), where each fiber realization consists of $N = 100$ independent polarization maintaining fiber (PMF) plates, each characterized by local eigenmodes that are uniformly distributed on the equator of the Poincaré sphere. The retardation angle of each nth plate is $\Delta\phi_n(\omega) = \Delta\phi_{n0} + \Delta\phi_{n1}\omega$, where $\Delta\phi_{n0}$ is a RV uniform on $[0, 2\pi]$ and the local birefringence strength $\delta\tau = \Delta\phi_{n1}$ was chosen in order to have the desired root mean square (rms) differential group delay (DGD) for the global fiber: $\sqrt{E[\Delta\tau^2]} = \sqrt{N}\delta\tau$ [12], whose value we chose equal to 10.5 ps for the set of simulated fibers. Possible frequency periodicities for the global Jones matrix resulting from a fixed local $\delta\tau$ [11] are not an issue when performing first order statistics. For frequency-distributed measurements,

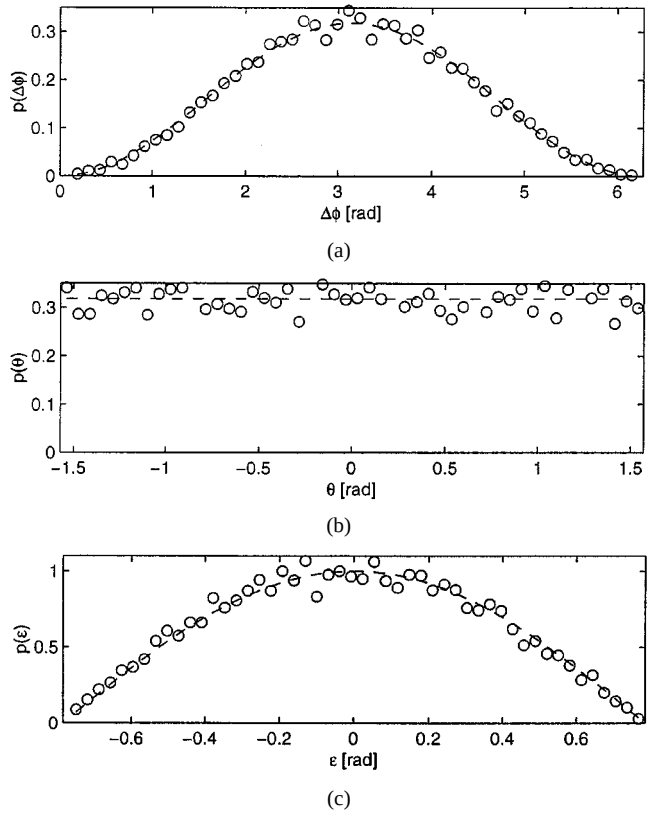

Fig. 1. Probability density of (a) retardation angle $\Delta\phi$; (b) eigenmode azimuth θ_b; (c) eigenmode ellipticity ε_b. Theory (14) and (13): lines; simulations: circles.

performed in Section VII-A, the artifact of frequency periodicities is overcome if the number of fiber plates obeys $\Delta B\sqrt{E[\Delta\tau^2]} \leq \sqrt{N}$, where ΔB is the measurement bandwidth, so that measurements are confined in one period only [13]. Such condition was always satisfied in our measurements.

In Fig. 1, the theoretical pdfs (14) of the retardation angle (top), and (13) of the eigenmode azimuth (center) and ellipticity (bottom) are plotted and compared to simulation results, which represent histograms computed over the fiber realizations at the reference frequency $\omega = 0$, and a good agreement is found. In all the presented figures, lines represent theory and symbols represent simulation. The marginal densities of the cartesian Pauli coordinates u_k are plotted in Fig. 2, using different symbols for $k = 0, 1, 2, 3$, and compared to the expression (15) obtained from theory.

Note that the good agreement between theory and simulation is obtained despite the remarkable differences between the assumptions made in the retarder plate model used for simulations and the white Gaussian theoretical model of birefringence.

IV. On the Independence Between PMD Vector and Pauli Vector

The white Gaussian model adopted for the local birefringence can be inserted in Poole's dynamical equation [1]. The SDE so obtained governs the evolution of the output PMD vector $\vec{\Omega}_o(z, \omega)$ along z: translated in Ito form its expression

$$d\vec{\Omega}_o = -\omega^2\mu^2\vec{\Omega}_o\,dz - \omega\mu\vec{\Omega}_o \times d\vec{\nu} + \mu\,d\vec{\nu} \qquad (16)$$

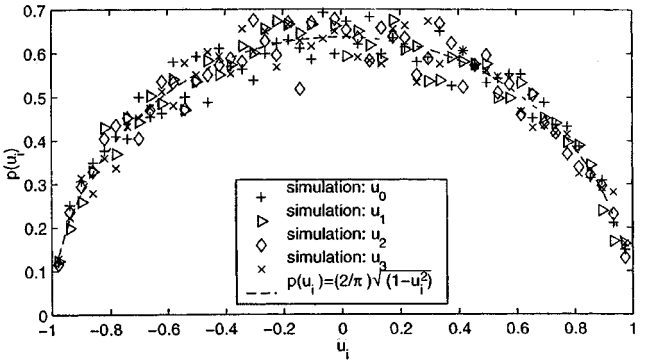

Fig. 2. Probability density of the cartesian Pauli coordinates. Theory (15): line; simulations: symbols.

is different from that obtained in [1], due to the different birefringence models. Equation (16) can be expressed in the canonical form (10): the drift and diffusion coefficients in this case are $\vec{c}_\Omega = -\omega^2\mu^2\vec{\Omega}_o$ and $V_\Omega = [\mu I - \omega\mu\vec{\Omega}_o\times]$, respectively. We can check that the pdf $p(z, \vec{\Omega}_o)$ of the PMD vector that solves the associated FPE is the well-known trivariate Gaussian distribution with zero mean and covariance matrix equal to $\mu^2 z I$. Due to the white Gaussian model of birefringence, which is inherently suitable for "long" fibers only, there is no transient behavior in z for $p(z, \vec{\Omega}_o)$ and the PMD vector reaches immediately its steady state Gaussian distribution. Hence, the rms DGD $\sqrt{E[|\vec{\Omega}_o|^2]} = \sqrt{3\mu^2 z}$, grows as the square root of the fiber length, as typical of "long" birefringent fibers [2].

We are now ready to prove the following.

Theorem: The PMD vector $\vec{\Omega}_o(z, \omega_1)$ and the Pauli vector $\underline{u}(z, \omega_2)$ are independent RVs for any (ω_1, ω_2).

Proof: Let us consider the values taken by the PMD vector and by the Pauli vector at any two (possibly equal) frequencies, $\vec{\Omega}_o(z, \omega_1)$ and $\underline{u}(z, \omega_2)$, at the same position z along the fiber. The joint stochastic differential system obtained from (9), at $\omega = \omega_2$, and (16), at $\omega = \omega_1$, can be associated to a FPE whose solution is the joint pdf $p(z, \vec{\Omega}_o(\omega_1), \underline{u}(\omega_2))$. It is straightforward to show that the solution to such FPE can be factorized into the product of the marginals $p(z, \vec{\Omega}_o(\omega_1))$ and $p(z, \underline{u}(\omega_2))$, which satisfy the respective FPEs, if and only if

$$\sum_{i,l=0}^{3} \frac{\partial^2\Big[a_{xil}\Big(\vec{\Omega}_o(\omega_1), \underline{u}(\omega_2)\Big)p\Big(z, \vec{\Omega}_o(\omega_1)\Big)\,p\left(z, \underline{u}(\omega_2)\right)\Big]}{\partial\Omega_{oi}\partial u_l} = 0 \qquad (17)$$

where $A_x(\vec{\Omega}_o(\omega_1), \underline{u}(\omega_2)) \triangleq V_\Omega(\omega_1, \vec{\Omega}_o)V_u^T(\omega_2, \underline{u})$ is the matrix accounting for the statistical cross-interaction in the joint FPE.

We show next that (17) is always verified for the joint system (9) and (16). Given the structure of V_Ω below (16), and V_u in (11), an explicit evaluation of A_x shows that each of its elements $a_{xil}(\vec{\Omega}_o(\omega_1), \underline{u}(\omega_2))$ does not depend on Ω_{oi} nor on u_l, and hence, can be taken out of the derivative in (17). As a second step toward the verification of (17), we call on the spherical uniformity of the marginal distributions $p(z, \vec{\Omega}_o(\omega_1))$ and $p(z, \underline{u}(\omega_2))$ obtained above. Their derivatives can then be expressed as $(\partial p(\vec{\Omega}_o)/\partial\Omega_{oi}) = (\partial p(\vec{\Omega}_o)/\partial|\vec{\Omega}_o|^2)\cdot 2\Omega_{oi}$

and $(\partial p(\underline{u})/\partial u_l) = (\partial p(\underline{u})/\partial |\underline{u}|^2) \cdot 2u_l$, and the verification of (17) reduces to proving that the double summation $\sum_{i,l=0}^{3} a_{xil}(\vec{\Omega}_o(\omega_1), \underline{u}(\omega_2)) \quad \Omega_{oi}(\omega_1) u_l(\omega_2)$ is zero. Such double summation is a quadratic form

$$\vec{\Omega}_o^T(\omega_1) A_x \left(\vec{\Omega}_o(\omega_1), \underline{u}(\omega_2)\right) \underline{u}(\omega_2) = \left[\vec{\Omega}_o^T(\omega_1) V_\Omega \left(\omega_1, \vec{\Omega}_o\right)\right] \left[\underline{u}^T(\omega_2) V_u \left(\omega_2, \underline{u}\right)\right]^T = 0$$

which is zero since $\underline{u}^T V_u = 0$, for any (z, ω). This last property has a simple geometrical explanation. Since (8) can be written as $(d\underline{u}/dz) = ((1/\omega\mu) V_u)\vec{W}$ for any $\vec{W}$, and since $\underline{u}(z)$ and $(d\underline{u}(z)/dz)$ are orthogonal because $|\underline{u}(z)| = 1$, then $0 = \underline{u}^T(d\underline{u}/dz) = (\underline{u}^T V_u)(\vec{W}/\omega\mu)$ for any $\vec{W}$, which implies $\underline{u}^T V_u = 0$. □

As a consequence of the aforementioned theorem, expanding $\vec{\Omega}_o(\omega) = \sum_i \vec{\Omega}_o^{(i)}(\omega_0)((\omega - \omega_0)^i/i!)$ in Taylor series, such independence implies that the derivatives $\vec{\Omega}_o^{(i)}(\omega_0)$ of any ith order of the PMD vector are independent of $\underline{u}(z, \omega_0)$. In the next section, we will use such independence, for $i = 0$, to statistically characterize the derivative of the Pauli vector in terms of $\vec{\Omega}_o$. The same procedure can be iterated to characterize higher order derivatives $\underline{u}^{(n)}(\omega_0)$ in terms of $(\vec{\Omega}_o(\omega_0), \vec{\Omega}_o^{(1)}(\omega_0) \cdots \vec{\Omega}_o^{(n-1)}(\omega_0))$.

V. The PMD Vector From a New Perspective: The Derivative of the Pauli Vector

The output PMD vector $\vec{\Omega}_o$ is defined by the analogue of (6), where the derivative is taken with respect to (w.r.t.) ω, and $\vec{W}$ is replaced by $\vec{\Omega}_o$ [7, eq. 5.11]. From such an equation, a linear differential system governing the evolution in frequency of the Pauli vector is obtained as in (8)

$$\underline{u}_\omega = -\tfrac{1}{2} \begin{bmatrix} -\vec{u}^T \\ u_0 I + \vec{u}\times \end{bmatrix} \vec{\Omega}_o \tag{18}$$

where subscript ω denotes derivative w.r.t. optical frequency.

Given $\underline{u}$, by the independence of $\vec{\Omega}_o$ and $\underline{u}$, (18) is a linear transformation of a Gaussian random vector, hence $\underline{u}_\omega$ is conditionally Gaussian, with conditional mean $E[\underline{u}_\omega|\underline{u}] = \underline{0}$ and covariance matrix $E[\underline{u}_\omega \underline{u}_\omega{}^T|\underline{u}] = (\mu^2 z/4)[|\underline{u}|^2 I - \underline{u}\,\underline{u}^T]$.

The Pauli vector $\underline{u}$ has unit magnitude, hence its derivative $\underline{u}_\omega$ is always orthogonal to it in $\Re^4$, and consequently, given $\underline{u}$, the probabilistic space in which $\underline{u}_\omega$ is distributed is only 3-D, the fourth dimension being determined by the orthogonality condition. Operating a rotation of the reference system in $\Re^4$ by the 4×4 unitary rotation matrix

$$\begin{bmatrix} u_0 & \vec{u}^T \\ -\vec{u} & u_0 I - \vec{u}\times \end{bmatrix} \tag{19}$$

$\underline{u}$ is rotated onto the 0th canonical axis $[1; 0; 0; 0]$. By applying the same transformation to (18), $\underline{u}_\omega$ is transformed into the vector $[0; -(\vec{\Omega}_o/2)]$. This shows a fundamental result: *the PMD vector can be viewed as -2 times the last three components of $\underline{u}_\omega$, once this is recast in a reference system where $\underline{u}$ coincides with $[1; 0; 0; 0]$, or equivalently, once we identify the hyperplane orthogonal to $\underline{u}$ in $\Re^4$ as the ordinary 3-D Stokes space*. The conditioning on $\underline{u}$ only influences the domain of definition of $\underline{u}_\omega$, i.e., the orientation of its hyperplane of definition in $\Re^4$, but not its distribution.

VI. Distribution of the Derivative of the Jones Matrix

From the previous results, the direction of $\underline{u}_\omega$ is uniformly distributed on $\mathcal{S}^4$ and its magnitude, amounting to $\Delta\tau/2$, is Maxwellian and independent of its direction. Hence, the statistics of the polar coordinates $(|\underline{u}_\omega|, \phi_1, \phi_2, \phi_3)$ of $\underline{u}_\omega$ are completely characterized, since, from geometrical results on the uniform distributions in $\Re^4$ reported in Appendix B, (ϕ_1, ϕ_2, ϕ_3) are independent RVs with marginal pdf's

$$p(\phi_1) = \frac{1}{2\pi}, \quad p(\phi_2) = \frac{\cos\phi_2}{2}, \quad p(\phi_3) = \frac{2\cos^2\phi_3}{\pi} \tag{20}$$

with $-\pi \le \phi_1 \le \pi$ and $-(\pi/2) \le \phi_2, \phi_3 \le (\pi/2)$.

By averaging the conditional covariance matrix of $\underline{u}_\omega$ w.r.t. $\underline{u}$, the cartesian coordinates of $\underline{u}_\omega$ are found to be uncorrelated zero mean RVs, each with variance equal to $E[\Delta\tau^2]/16$. They all have the same pdf, which is obtained from a transformation of the polar coordinates as

$$p(u_{\omega i}) = 4\sqrt{\frac{\beta}{\pi^3}}\, \beta u_{\omega i}^2 \mathrm{e}^{-\beta u_{\omega i}^2} \Gamma\left(\frac{3}{2}\right) \Psi\left(\frac{3}{2}, 2; \beta u_{\omega i}^2\right) \tag{21}$$

where $\beta = 6/E[\Delta\tau^2]$, $\Gamma(z)$ is the Gamma function and $\Psi(\alpha, \gamma; z)$ is the confluent hypergeometric function of second kind [14, eq. (9.211.4)].

We are now interested in evaluating the distributions of the retardation angle derivative $\Delta\phi_\omega$ and the eigenmodes derivative $\vec{b}_\omega$. These two parameters have a precise physical meaning: $\Delta\phi_\omega$ is the differential delay introduced by the Jones matrix for each spectral component of the signal, while $\vec{b}_\omega$ is the *depolarization* vector associated with the eigenmodes: if it is zero, the fiber is a PMF, i.e., we are in the "first-order PMD" case. Otherwise, the *depolarization rate* $|\vec{b}_\omega|$ will cause higher order pulse distortions.[4] Expressing the derivative of the Pauli coordinates in terms of the fiber retardation angle and eigenmodes, from (2) and (18) we have $\Delta\phi_\omega = \hat{b} \cdot \vec{\Omega}_o = \Delta\tau \cos\alpha$, where α is the angle between two independent unit magnitude vectors, uniformly distributed on the Poincaré sphere, with pdf: $p(\alpha) = (\sin\alpha/2)(0 \le \alpha \le \pi)$. Since $\Delta\tau$ is independent of α, $\Delta\phi_\omega$ can be easily shown to be Gaussian distributed: $\Delta\phi_\omega \sim \mathcal{N}(0, E[\Delta\tau^2]/3)$ *and independent of* $\Delta\phi$. From the discussion at the end of Section V, it can be shown that the direction of $\vec{b}_\omega$ is spherically uniform in Stokes space, and, taking the magnitude of both sides of (18), we get $|\vec{b}_\omega| = (\Delta\tau \sin\alpha/2 \sin(\Delta\phi/2))$, with $\Delta\tau$, α and $\Delta\phi$ independent RVs. The pdf of the depolarization rate is then

$$p(|\vec{b}_\omega|) = \frac{4}{\pi}\beta|\vec{b}_\omega|\mathrm{e}^{-\beta|\vec{b}_\omega|^2} B\left(\frac{5}{2}, \frac{1}{2}\right) \Phi\left(\frac{1}{2}, 3; \beta|\vec{b}_\omega|^2\right) \tag{22}$$

where $B(x, y)$ is the Beta function and $\Phi(\alpha, \gamma; z)$ is the confluent hypergeometric function of first kind [14, eq. (9.211.1)]. It is also possible to show that $\sin(\Delta\phi/2)|\vec{b}_\omega|$ is Rayleigh distributed and independent of $\Delta\phi_\omega$.

[4]The only exception to this statement is when the fiber under consideration is made up of a PMF concatenated with a frequency independent polarization controller: in this case the propagating pulse only suffers the DGD effect (pulse splitting) and $|\vec{b}_\omega|$ is not zero [8].

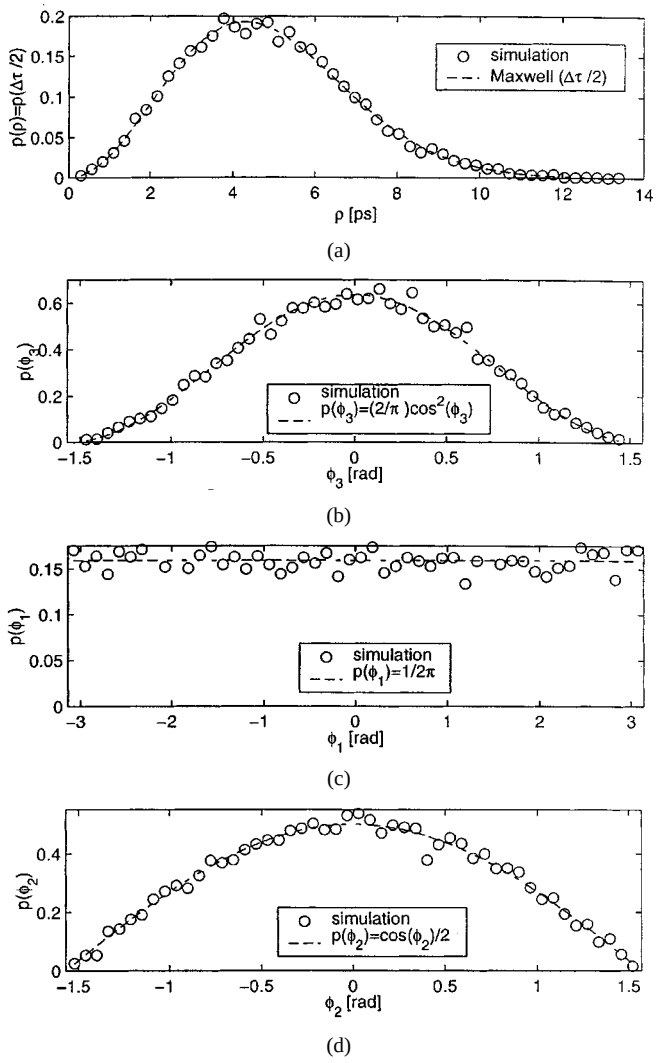

Fig. 3. Probability density of: magnitude of the Pauli vector derivative $|\underline{u}_\omega|$; spherical coordinates ϕ_3, ϕ_1, ϕ_2 of $\underline{u}_\omega$. Theory (20): lines; simulations: circles.

Note that $p(|\vec{b}_\omega|)$ depends through β on the only parameter $E[\Delta\tau^2]$. Since $|\vec{b}_\omega|$ is an indicator of higher order PMD effects, (22) physically quantifies the impact of the average DGD on such indicator.

A. Check by Simulation

The same set of fibers as in Section III-B, simulated with the retarder plate model, was used for computing the numerical distributions of $\underline{u}_\omega$, $\Delta\phi_\omega$ and $\vec{b}_\omega$.

Starting from the polar coordinates of $\underline{u}_\omega$, Fig. 3(a) compares the Maxwellian distribution obtained for the magnitude $|\underline{u}_\omega|$, which corresponds to half the DGD, with the simulation results, which were computed from the Jones matrix frequency derivative by extracting its Pauli vector $\underline{u}_\omega$. The estimated pdfs of the spherical coordinates ϕ_3, ϕ_1, ϕ_2, which identify the orientation of $\underline{u}_\omega$ in $\Re^4$, are also reported in Fig. 3, and compared to theory (20).

The distribution of the derivative $\Delta\phi_\omega$ of the retardation angle $\Delta\phi(\omega)$ is reported in Fig. 4, well matching with the zero mean Gaussian $\mathcal{N}(0,\ (E[\Delta\tau^2]/3))$.

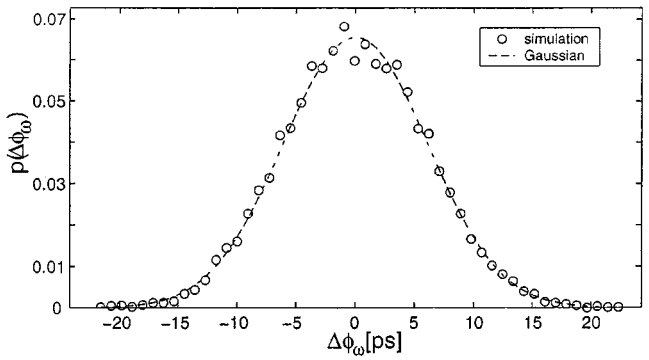

Fig. 4. Probability density of the retardation angle derivative $\Delta\phi_\omega$. Theory $\mathcal{N}(0,\ (E[\Delta\tau^2]/3))$: line; simulations: circles.

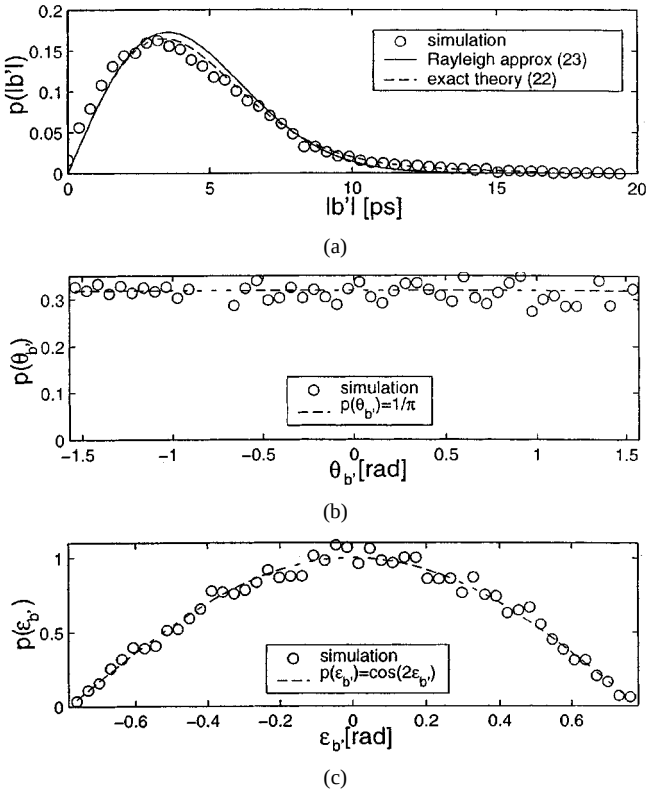

Fig. 5. (a) Probability density of: depolarization rate $|\vec{b}_\omega|$. (b) Depolarization vector azimuth $\theta_{b'}$ (c) Ellipticity $\varepsilon_{b'}$. Theory (22) and approximation (23) (a): lines; simulations: circles.

The theoretical analysis performed in Section VI predicts a uniform orientation of the direction of $\vec{b}_\omega$ in Stokes space. Defining its azimuth $\theta_{b'}$ and ellipticity $\varepsilon_{b'}$ similarly to those of $\hat{b}$ in Section III, their pdfs are the same as in (13): such result is confirmed by the simulation results plotted in Fig. 5 (center and bottom plots). As for the depolarization rate $|\vec{b}_\omega|$, we plot in Fig. 5(a) the theoretical pdf (22) together with the approximating Rayleigh density

$$p(|\vec{b}_\omega|) \simeq \tfrac{3}{2}\,\beta|\vec{b}_\omega|\mathrm{e}^{-(3/4)\beta|\vec{b}_\omega|^2} \tag{23}$$

shown in solid line, with mean value $(\pi/4\sqrt{3})E[\Delta\tau]$. Equation (23) is based on the analytical approximation of the confluent hypergeometric function $\Phi(1/2,\ 3;\ z) \simeq \mathrm{e}^{(z/4)}$ which, though well fitted to the exact solution, tends to underestimate the tail of $p(|\vec{b}_\omega|)$.

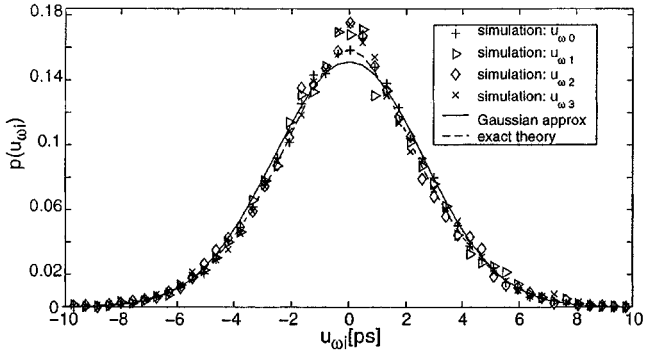

Fig. 6. Probability density of the cartesian coordinates of $\underline{u}_\omega$. Theory: exact pdf (21) in dashed line and Gaussian approximation in solid line; simulations: symbols.

We can provide a simple analytical approximation also for the marginal distributions of the cartesian coordinates of $\underline{u}_\omega$. Approximating the confluent hypergeometric function $\Psi((3/2), 2; z) \simeq (1/z)\mathrm{e}^{(-z/3)}$, we get a normal distribution $\mathcal{N}(0, (E[\Delta\tau^2]/16))$ for $p(u_{\omega i})$. Such approximation is shown in solid line in Fig. 6 and compared to the exact solution (21) in dashed line and to simulation.

VII. Autocorrelation Function of the Pauli Vector

Considering the Ito dynamical system (16) expressed for two different frequency values, we can establish a differential equation for the evolution along z of the frequency autocorrelation function (ACF) of the PMD vector $R_{\vec{\Omega}}(\omega_1, \omega_2) \triangleq E[\vec{\Omega}_o(\omega_1) \cdot \vec{\Omega}_o(\omega_2)]$. Since the differential of the PMD vector involves the unbounded quantity $d\vec{r}$, we must treat $\vec{\Omega}_o(\omega_1) \cdot \vec{\Omega}_o(\omega_2)$ as a Stratonovich product. We obtain the following differential equation for the ACF:

$$\frac{d}{dz} R_{\vec{\Omega}}(\omega_1, \omega_2) = -\Delta\omega^2\mu^2 R_{\vec{\Omega}}(\omega_1, \omega_2) + 3\mu^2$$

where $\Delta\omega = \omega_2 - \omega_1$ is the frequency deviation, and the initial condition is $R_{\vec{\Omega}}(\omega_1, \omega_2) = 0$ at $z = 0$ (a zero length fiber has a zero PMD vector), with solution

$$R_{\vec{\Omega}}(\Delta\omega) = 3\frac{1 - \mathrm{e}^{-\Delta\omega^2\mu^2 z}}{\Delta\omega^2} \tag{24}$$

and peak value $\lim_{\omega\to 0} R_{\vec{\Omega}}(\omega) = E[\Delta\tau^2]$ equal to the mean square DGD. As noted in Section IV, $E[\Delta\tau^2] = 3\mu^2 z$, hence, (24) coincides with the ACF obtained by Karlsson/Shtaif *et al.* [15] and [16] following different approaches.

Similarly, the ACF $R_{\underline{u}}(\omega_1, \omega_2) \triangleq E[\underline{u}(\omega_1) \cdot \underline{u}(\omega_2)]$ of the Pauli vector $\underline{u}(z, \omega)$ can be obtained starting from the dynamical system (9) considered for two different frequency values. The differential equation so obtained is

$$\frac{d}{dz} R_{\underline{u}}(\omega_1, \omega_2) = -\frac{3\Delta\omega^2\mu^2}{8} R_{\underline{u}}(\omega_1, \omega_2)$$

with the initial condition $E[\underline{u}(0, \omega_1) \cdot \underline{u}(0, \omega_2)] = 1$ (a fiber of length zero has the Pauli vector $\underline{u}(0, \omega) = [1; 0; 0; 0]$ of the identity matrix), with solution

$$R_{\underline{u}}(\Delta\omega) = \mathrm{e}^{-(3\Delta\omega^2\mu^2 z/8)} = \mathrm{e}^{-(E[\Delta\tau^2]\Delta\omega^2/8)}. \tag{25}$$

Using the Dynkin formula it is also possible to prove the following result:

$$E[u_i(z, \omega_1)u_j(z, \omega_2)] = \tfrac{1}{4}\,\mathrm{e}^{-(E[\Delta\tau^2]\Delta\omega^2/8)}\delta_{ij} \tag{26}$$

where δ_{ij} is the Kronecker symbol. Equation (26) extends the incorrelation result of Section III-A and shows that each of the four Pauli coordinates gives an equal contribution to the ACF in (25). In the derivation of (26), we used the asymptotic stationarity of the process $\underline{u}(z, \omega)$, and let $\omega_1, \omega_2 \to \infty$ while keeping $\Delta\omega = \omega_2 - \omega_1$ finite, according to the Appendix A. From such result we also get the ACF of $U(\omega)$ as follows. From the product formula (7), the Pauli coordinates of $U^\dagger(\omega_1)U(\omega_2)$ are $[u_0(\omega_1)u_0(\omega_2) + \vec{u}(\omega_1) \cdot \vec{u}(\omega_2);\ [u_0(\omega_1)\vec{u}(\omega_2) - u_0(\omega_2)\vec{u}(\omega_1) + \vec{u}(\omega_1) \times \vec{u}(\omega_2)]]$. Using (26) we see that only the zero-th component has nonzero mean, so that $E[U^\dagger(\omega_1)U(\omega_2)] = R_{\underline{u}}(\Delta\omega)\sigma_0$, extending to the Jones matrix a result already known for the Müller matrix [15], [16].

The importance of the autocorrelation function for the Pauli vector—and for the PMD vector—lies in its interpretation in terms of mutual statistical dependence between the value taken by the vector $\underline{u}(z, \omega)$, and hence by the Jones matrix, at a certain reference frequency and the value $\underline{u}(z, \omega + \Delta\omega)$ it takes at a different frequency. The ACF gives an indication of how quickly, on average, the vector moves in frequency w.r.t. its reference frequency value, but does not give information on the shape of the trajectories the vectors are likely to follow. Nevertheless, we can compute a *correlation bandwidth* for $\vec{\Omega}_o(z, \omega)$ and for $\underline{u}(z, \omega)$, based on (24) and (25), respectively, which gives a measure of the frequency span over which the vectors can be regarded as "nearly constant" (this condition is practically verified only for a fraction of the correlation bandwidth). Defining the correlation bandwidth $\Delta\omega_{Cs}$ for a stochastic process $s(\omega)$ with ACF $R_s(\omega)$ as $\Delta\omega_{Cs} \triangleq \int_{-\infty}^{+\infty} R_s(\omega)d\omega/R_s(0)$, we find $\Delta\omega_{C\Omega} = (4\sqrt{2}/E[\Delta\tau])$, consistently with [15], [16], and $\Delta\omega_{Cu} = (8/\sqrt{3}E[\Delta\tau])$, which is $\sqrt{2/3} \simeq 0.8$ times the correlation bandwidth of the PMD vector. Hence, we conclude that the PMD vector is characterized by a slightly larger stability in the frequency domain, i.e., it undergoes smaller changes in magnitude/orientation, over given frequency bandwidth, as compared to the Pauli vector of the Jones matrix. For signal bandwidths comparable with the correlation bandwidth $\Delta\omega_{C\Omega}$, the PMD vector is no longer sufficient to accurately describe the output field, and the so-called higher order PMD must be taken into account. We defer the discussion of the Jones matrix eigenmodes ACF to the next section, dealing first with numerical results.

A. Check by Simulation

The same set of 10 000 fibers as in Section III-B, was used to validate the results obtained for the ACFs $R_{\vec{\Omega}}(\Delta\omega)$ and $R_{\underline{u}}(\Delta\omega)$ through numerical simulations, with the only difference that an rms DGD of 10 ps was imposed in the simulation (instead of the value 10.5 ps used in Section III-B). Measurements were performed on a two-sided bandwidth of $\Delta B = 250$ GHz, where $\Delta B\sqrt{E[\Delta\tau^2]} \ll \sqrt{N}$ obeys the condition stated in Section III-B [13]. However, such ΔB

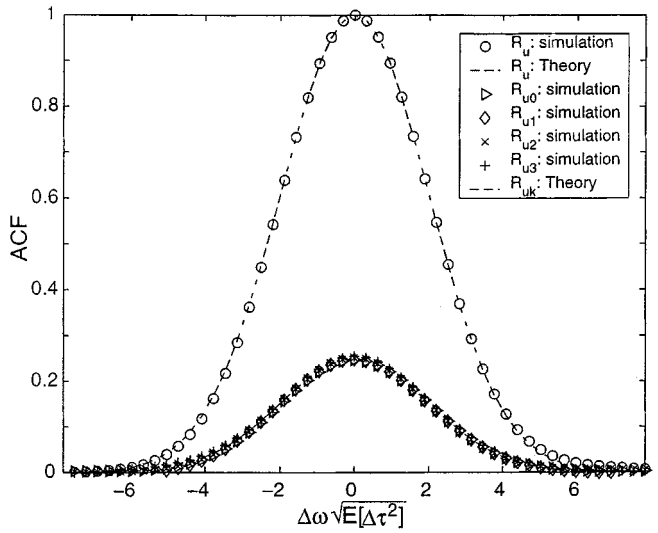

Fig. 7. ACF of the Pauli vector and of its coordinates. Theory (25), (26): line; simulations: symbols.

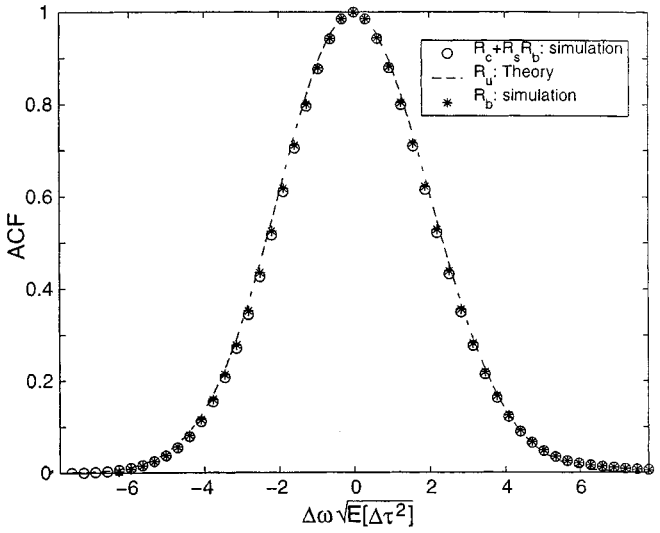

Fig. 8. ACF of the eigenmodes, plotted versus the theoretical ACF of $\underline{u}$. Factorized ACF $R_c(\omega_1, \omega_2) + R_s(\omega_1, \omega_2)R_{\hat{b}}(\omega_1, \omega_2)$ showing the incorrelation of $\hat{b}(z, \omega_1)$ and $\sin(\Delta\phi(z, \omega_2)/2)$.

includes any bandwidth of practical interest for transmission on fibers with the chosen rms DGD.

Since in the dynamical equations (9) the optical frequency ω always appears multiplied by the birefringence standard deviation μ, all the dynamical properties of Pauli coordinates in the frequency domain are expected to scale with μ, and hence, with $E[\Delta\tau]$, as confirmed by (25). Thus, the ACF of $\underline{u}$ and of its four components is plotted in Fig. 7 versus the normalized frequency deviation $\Delta\omega\sqrt{E[\Delta\tau^2]}$, as done in [17] for $R_{|\vec{\Omega}|}(\Delta\omega)$. We note the excellent match of simulations with the theoretically predicted Gaussian ACFs (25), (26).

Information on the ACF of the retardation angle $\Delta\phi$ and the eigenmode $\hat{b}$ is buried in the ACF of $\underline{u}$, and we did not find a way to extract it analytically. However, $R_c(\omega_1, \omega_2) \triangleq R_{u0}(\omega_1, \omega_2) = (1/4)R_{\underline{u}}(\omega_1, \omega_2)$ is the ACF of $\cos(\Delta\phi(z, \omega)/2)$. Defining the ACF of $\sin(\Delta\phi(z, \omega)/2)$ as $R_s(\omega_1, \omega_2)$ and the ACF of the eigenmode as $R_{\hat{b}}(\omega_1, \omega_2) = E[\hat{b}(z, \omega_1) \cdot \hat{b}(z, \omega_2)]$, we would be allowed to factorize $E[\sin(\Delta\phi(z, \omega_1)/2)\sin(\Delta\phi(z, \omega_2)/2)\hat{b}(z, \omega_1) \cdot \hat{b}(z, \omega_2)] = (3/4)R_{\underline{u}}(\omega_1, \omega_2)$ as $R_s(\omega_1, \omega_2)R_{\hat{b}}(\omega_1, \omega_2)$ only under the incorrelation of $\sin(\Delta\phi(z, \omega_1)/2)$ and $\hat{b}(z, \omega_2)$ for any (ω_1, ω_2). We only proved independence for $\omega_1 = \omega_2$, but not for $\omega_1 \neq \omega_2$.

In Fig. 8, we again plot in dashed line the theoretical ACF $R_{\underline{u}}(\omega_1, \omega_2)$, while in circles we plot the function $R_c(\omega_1, \omega_2) + R_s(\omega_1, \omega_2)R_{\hat{b}}(\omega_1, \omega_2)$. Since the two curves are closely matching, we conclude that $\hat{b}(z, \omega_1)$ and $\sin(\Delta\phi(z, \omega_2)/2)$ are practically uncorrelated for any choice of ω_1, ω_2. In Fig. 8 we also report, in stars, the *simulated* ACF $R_{\hat{b}}$ alone and we note that $R_{\hat{b}}(\omega_1, \omega_2) \cong R_{\underline{u}}(\omega_1, \omega_2)$. The significant implication of this simulation result is that the correlation bandwidth of the eigenmode $\hat{b}$, equal to $\Delta\omega_{Cu}$, is $\sqrt{2/3}$ that of the PMD vector $\vec{\Omega}_o$. Hence, it is confirmed that the PSPs possess a slightly greater stability in frequency than the eigenmodes, and are then preferable as a frame of reference for expressing the transmitted field if the signal bandwidth is small enough. But it is also true that the correlation bandwidths

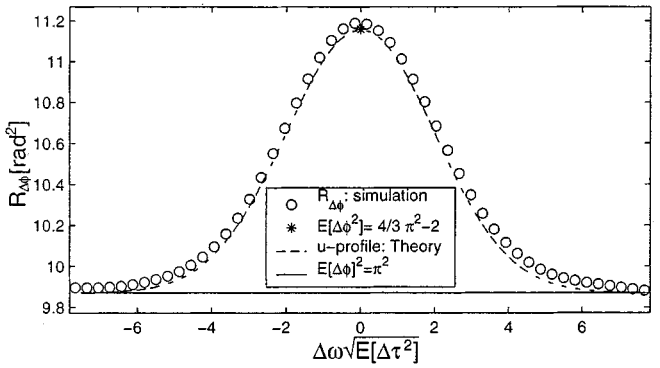

Fig. 9. ACF of the retardation angle $\Delta\phi(z, \omega)$, compared to a Gaussian shape with the same variance of $R_{\underline{u}}$. The analytically computable central value and asymptotic value are also reported.

are comparable ($\Delta\omega_{Cb} \simeq 0.8\Delta\omega_{C\Omega}$) and that if the signal bandwidth is increased, e.g., due to increased transmission rate, and a higher order description is needed, the most appropriate output field description uses the eigenmodes, because of their more regular circular contours, in frequency, on the Poincaré sphere. The "all-order" compensator described in [4] is based on this principle.

An exact coincidence of $R_{\hat{b}}(\omega_1, \omega_2)$ and $R_{\underline{u}}(\omega_1, \omega_2)$ would imply a constant $R_s(\omega_1, \omega_2) = 3/4$. In fact, the peak value $R_s(\omega, \omega) = E[\sin^2(\Delta\phi(z, \omega)/2)] = 3/4$ and the asymptotic value $\lim_{(\omega_2-\omega_1)\to\infty} R_s(\omega_1, \omega_2) = E[\sin(\Delta\phi(z, \omega)/2)]^2 = (64/9\pi^2) \simeq 0.72$, both evaluated analytically from (14), justify that $R_s(\omega_1, \omega_2)$ is always close to $3/4$, as we verified numerically.

Regarding $R_{\Delta\phi}(\omega_1, \omega_2)$, we can compute its peak value[5] $R_{\Delta\phi}(\omega, \omega) = E[\Delta\phi^2(z, \omega)] = (4/3)\pi^2 - 2[rad^2]$ and asymptotic value $\lim_{(\omega_2-\omega_1)\to\infty} R_{\Delta\phi}(\omega_1, \omega_2) = E[\Delta\phi(z, \omega)]^2 = \pi^2$ from (14). Numerical simulations, reported in Fig. 9, show a

[5]We assume wide sense stationarity for the process $\Delta\phi(z, \omega)$. Under this assumption, it is true that the central value $R_{\Delta\phi}(0)$ is the absolute maximum of $|R_{\Delta\phi}(\Delta\omega)|$.

remarkable feature: $R_{\Delta\phi}(\Delta\omega)$ has the same Gaussian shape as $R_{\underline{u}}(\Delta\omega)$, with appropriate scaling. A feature for which we did not find an analytical justification.

From the simulation results, we conclude that

$$R_{\Delta\phi}(\Delta\omega) \simeq \pi^2 + \left(\tfrac{1}{3}\pi^2 - 2\right) \mathrm{e}^{-(E[\Delta\tau^2]\Delta\omega^2/8)} \quad (27)$$

$$R_{\hat{b}}(\Delta\omega) \simeq \mathrm{e}^{-(E[\Delta\tau^2]\Delta\omega^2/8)} \quad (28)$$

can be used with good accuracy for approximating the ACFs of the fiber retardation angle and eigenmodes. An intuitive justification for the Gaussian shape of $R_{\Delta\phi}$ is found by expanding $u_0 = \cos(\Delta\phi/2)$ in Taylor series, to the first order, around $E[\Delta\phi] = \pi$, and calculating $R_{u_0} = (1/4)R_{\underline{u}}$. We get $R_{\Delta\phi}(\Delta\omega) \simeq \pi^2 + R_{\underline{u}}(\Delta\omega)$, that has the same shape as (27), though it slightly underestimates the peak value of $R_{\Delta\phi}$.

We again stress that, in all the cases where theoretical results are available, simulation and theory are in good agreement, notwithstanding the remarkable differences between the retarder plate model used for the simulations and the theoretical model, in which the local birefringence direction has a uniform distribution on the Poincaré sphere, the local birefringence strength has Maxwell distribution, and there is no retardation at the reference frequency.

VIII. Conclusion

Assuming Gisin's "white Gaussian" model for the local birefringence [2, Appendix D], we were able to prove the following.

i) For any given frequency and fiber length, the eigenmodes are a random vector uniformly distributed on the Poincaré sphere and are independent of the retardation angle, whose pdf follows a cosine law.

ii) For any given frequency, the uncorrelated Pauli coordinates evolve in space like a standard Brownian motion on the unit sphere in $\Re^4$.

iii) For any given frequency and fiber length, the Pauli vector is independent of the PMD vector and of its derivatives. The frequency derivative of the Pauli vector is a vector in $\Re^4$ whose direction is uniformly distributed and whose magnitude is Maxwellian, being a linear transformation of the PMD vector.

iv) The frequency derivative of the retardation angle is Gaussian, independent of the retardation angle, and the depolarization rate of the eigenmodes, which depends on the retardation angle but not on the eigenmodes, has a "quasi-Rayleigh" distribution, modified by a confluent hypergeometric function.

v) For fixed fiber length, the Pauli coordinates have identical Gaussian frequency autocorrelation function. The ACF of the eigenmodes and retardation angle is also practically Gaussian-shaped, with coherence bandwidth equal to $\sqrt{2/3}$ that of the PMD vector.

Appendix A
A Critical Comparison of the Different Models for the Local Birefringence

In order to analyze the system of SDE in (8), we need to suitably model the variations of the local birefringence along the fiber length, i.e., the stochastic process $\vec{W}(z,\omega)$. In the literature, mostly two models are adopted for long fibers. The first one, used by Foschini and Poole [1], assumes $\vec{W}(z,\omega) = \omega\Delta'\hat{s}_1 + \vec{\nu}(z)$, where the ω-dependent term is linear in frequency and oriented along the first axis of Stokes space $\hat{s}_1$, while the z-dependent term is a three-dimensional white noise process with independent components, each with variance μ^2. The second model [2, Appendix D] for fibers much longer than the correlation length, assumes $\vec{W}(z,\omega) = \omega\vec{\nu}(z)$, with the same white noise process $\vec{\nu}(z)$ as in [1] (indicated by $\vec{\beta}(z)$ in [2]).

Both models resort to a white noise process, which is convenient for analyzing dynamical equations with the standard tools of stochastic calculus, but are physically very different, leading to some difficulties in the length or frequency domains, respectively. In Poole's model, for $\omega = 0$, the direction of $\vec{W}(z,0) = \vec{\nu}(z)$ is uniformly distributed on the Poincaré sphere and its magnitude, the birefringence strength, is distributed as a Maxwell random variable. As ω increases, the birefringence vector increases along $\hat{s}_1$, and thus, for large ω, the birefringence of the fiber sections has a preferential orientation ($\hat{s}_1$) which prevents the global PMD vector from having a uniformly distributed direction, unless the fiber is long enough. This problem is circumvented in [1, Appendix C] by letting $z \to \infty$, to obtain the equilibrium density for the PMD vector, which is uniform on the sphere. On the other hand, the second model includes fiber segments with a uniformly distributed birefringence direction, for every z, but it neglects the birefringence effects at the reference frequency, since $\vec{W}(z,0) = \vec{0}$ and, consistently with the assumptions in [12], all the fiber segments have an identity transfer matrix at $\omega = 0$. Stated in other words, any global fiber descriptor, which is a random process both in the z and the ω domains, suffers from a stationarity problem since the birefringence is anchored to an initial condition (a preferential orientation dictated by $\hat{s}_1$, for $z \to 0$, in the first model, and the null birefringence strength, for $\omega = 0$, in the second model). Such processes will nevertheless be asymptotically stationary and the problem of initial values is circumvented by letting $z \to \infty$ in the first model (long fibers) as well as by letting $\omega \to \infty$ in the second model (large reference frequencies).

As we see in Section VII, the dependence of some moments of $\underline{u}$ on the absolute frequency reference ω is thus an artifact introduced by the stochastic birefringence model $\vec{W}(z,\omega) = \omega\mu\vec{n}(z)$. When regarding the coordinates $u_k(z,\omega)$ as stochastic processes in the frequency domain, stationarity cannot be expected since their value at $\omega = 0$ is anchored to an initial condition [$u_0(z,0) = 1$ and $u_k(z,0) = 0$, $k = 1, 2, 3$, for any z] because of the absence of a random frequency-independent birefringence term in $\vec{W}(z,\omega)$. Nevertheless, if we regard the fiber transfer matrix as the concatenation of infinitesimal fiber plates, each has a Müller transfer matrix $M_n = \mathrm{e}^{[\vec{W}(z,\omega)\times]dz} = \mathrm{e}^{(\omega_0\mu|\vec{n}|+\overline{\omega}\mu|\vec{n}|)[\hat{n}(z)\times]\,dz}$. For $\omega_0 \to \infty$, the random rotation introduced by the term $\omega_0\mu|\vec{n}|$, when the new frequency variable $\overline{\omega} \triangleq \omega - \omega_0$ goes to zero, makes the global fiber Jones transfer matrix a stationary process in $\overline{\omega}$, provided the phase rotation can be regarded as a uniform variable on $[0, 2\pi]$, which is consistent with the assumptions of the retarder plate model [11], commonly adopted for fiber simulation. Stated in other

words, the Pauli coordinates $u_k(z, \omega)$ are asymptotically stationary random processes in the variable ω and their moments must be evaluated for $\omega \to \infty$ (keeping the frequency deviation $\Delta\omega$ finite, for second order moments) in order to obtain physically meaningful results from the application of the adopted birefringence model.

There is indeed a third model in the literature (the "first model" proposed in [3, p. 150]) which assumes $\vec{W}(z, \omega) = b(\omega)[\cos\theta(z);\ \sin\theta(z);\ 0]$ with birefringence strength $b(\omega)$ fixed along z and $\theta(z)$ a Wiener process. This model seems to best reflect our physical understanding of fiber birefringence, since the birefringence strength, which is mostly induced by the ellipticity of the fiber core, can be reasonably assumed to be constant due to the manufacturing process. The third component of $\vec{W}$ is zero since silica fibers are linearly birefringent. Finally, the rate of change in the orientation θ of the linear axes of birefringence is due to local stresses: while such rotational stresses can be regarded as a scalar white noise process, on a suitably long scale, the birefringence orientation is a z-continuous Wiener process.

Despite possible criticism regarding the neglect of $\Delta\beta_0$, we choose for analytical simplicity the Gisin model. We will see that its predictions agree with simulation results, despite the differences between the theoretical model and the retarder plate model employed for fiber simulation. Thus, in this paper, we assume $\vec{W}(z, \omega) = \omega\mu\vec{n}(z)$, with $\vec{n}(z)$ defined as a 3-D normalized white noise, i.e., with independent components and unit variance per component.

Appendix B
Uniform Distributions on Spheres

The coordinates of a point lying on a sphere of radius ρ in $\Re^3$ are

$$[x_1;\, x_2;\, x_3] = \rho[\cos\theta_2\cos\theta_1;\, \cos\theta_2\sin\theta_1;\, \sin\theta_2]$$

with $|\theta_1| \le \pi$ and $|\theta_2| \le \pi/2$. Note that $[x_1; x_2; x_3]$, with $\rho = 1$, express the coordinates of $\hat{b}$ if θ_1 is interpreted as twice the azimuth and θ_2 as twice the ellipticity of the eigenmode. Extending to the space $\Re^4$, the coordinates of a point lying on a hypersphere of radius ρ are

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \rho \begin{bmatrix} \cos\theta_3\cos\theta_2\cos\theta_1 \\ \cos\theta_3\cos\theta_2\sin\theta_1 \\ \cos\theta_3\sin\theta_2 \\ \sin\theta_3 \end{bmatrix} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_0 \end{bmatrix} \tag{29}$$

with $|\theta_1| \le \pi$, $|\theta_2| \le \pi/2$, and $|\theta_3| \le \pi/2$. Assuming $\rho = 1$, the last equality in (29) relates the coordinates x_j to the Pauli coordinates u_k, provided θ_3 is interpreted as $(\pi/2) - (\Delta\phi/2)$. In particular, given $n = 4$, (29) is a set of n parametric equations. When the set of $n-1$ parameters $[\theta_1; \theta_2; \cdots; \theta_{n-1}]$ varies in a domain $T \in \Re^{n-1}$, (29) defines a region with $n-1$ degrees of freedom, i.e., an *hypersurface* $\Sigma \in \Re^n$ in the Euclidean coordinates x_j. It is a basic result of multidimensional calculus (see, e.g., [18, Sec. 6.1.8]) that, if the parameters θ_i are each varied by an infinitesimal amount $d\theta_i$, the vector $\underline{x} \triangleq [x_1; x_2; \cdots; x_n]$ describes an element of surface $d\Sigma$ with area

$$dA = \sqrt{\sum_{j_1 < \cdots < j_{n-1}} \left(\frac{\partial(x_{j_1}, \ldots, x_{j_{n-1}})}{\partial(\theta_1, \ldots, \theta_{n-1})} \right)^2}\, d\theta_1 \cdots d\theta_{n-1}. \tag{30}$$

The square root term in (30) can be evaluated from the Jacobian matrix $J = \{\partial x_j / \partial\theta_i\}$ by deleting, for $j = 1, 2, \ldots, n$, the jth column from J, and computing the determinant of the resulting $(n-1) \times (n-1)$ matrix. The value of each jth determinant is defined as the jth element l_j of a vector $\underline{l} \in \Re^n$, and (30) can be compactly expressed using the norm of such vector: $dA = |\underline{l}| d\theta_1 \cdots d\theta_{n-1}$. To compute the area of $\mathcal{S}^4$, one can integrate $|\underline{l}| d\theta_1 d\theta_2 d\theta_3$ with $|\theta_1| \le \pi$ and $|\theta_2|$, $|\theta_3| \le \pi/2$, where $|\underline{l}| = \rho^3 \cos^2\theta_3 \cos\theta_2$ is obtained from (29) as described above. Integration, with $\rho = 1$, yields $Area(\mathcal{S}^4) = 2\pi^2$.

We are now ready to evaluate the joint probability density of the angular coordinates $[\theta_1; \theta_2; \ldots; \theta_{n-1}]$ for a uniform distribution on the sphere of unit radius in $\Re^n$. For a uniform distribution, the probability $\Pr[d\Sigma]$ that a point lies in the element of surface $d\Sigma$ is the ratio between the area dA of $d\Sigma$ and the area of the whole sphere. Hence, if $p(\theta_1, \theta_2, \ldots, \theta_{n-1})$ is the joint pdf

$$p(\theta_1, \theta_2, \ldots, \theta_{n-1})\, d\theta_1 d\theta_2 \cdots d\theta_{n-1} = \Pr[d\Sigma] = \frac{dA}{Area(\mathcal{S}^n)}$$

and, for $n = 4$, we get

$$\begin{aligned} p(\theta_1, \theta_2, \theta_3) &= \frac{\cos\theta_2 \cos^2\theta_3}{2\pi^2} \\ &= \left(\frac{1}{2\pi}\right)\left(\frac{\cos\theta_2}{2}\right)\left(\frac{2\cos^2\theta_3}{\pi}\right) \\ &= p(\theta_1)p(\theta_2)p(\theta_3). \end{aligned} \tag{31}$$

The marginal densities $p(\theta_i)$ can be obtained by integration w.r.t. the remaining two variables. The above product form proves that the three angles θ_i are independent random variables.

Acknowledgment

The authors would like to thank the anonymous reviewers, particularly for suggesting possible applications of the results in (13) and (14).

References

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

[2] 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 transmission," *J. Lightwave Technol.*, vol. 16, pp. 757–771, May 1998.

[3] P. K. A. Wai and C. R. Menyuk, "Polarization mode dispersion, decorrelation, and diffusion in optical fibers with randomly varying birefringence," *J. Lightwave Technol.*, vol. 14, pp. 148–157, Feb. 1996.

[4] M. Shtaif, A. Mecozzi, M. Tur, and J. A. Nagel, "A compensator for the effects of high-order polarization mode dispersion in optical fibers," *IEEE Photon. Technol. Lett.*, vol. 12, pp. 434–436, Apr. 2000.

[5] B. Øksendal, *Stochastic Differential Equations*, Fifth ed. New York: Springer-Verlag, 1998.
[6] M. Karlsson, "Polarization mode dispersion-induced pulse broadening in optical fibers," *Opt. Lett.*, vol. 23, pp. 688–690, May 1998.
[7] J. P. Gordon and H. Kogelnik, "PMD fundamentals: Polarization mode dispersion in optical fibers," in *Proc. National Academy Science (PNAS)*, vol. 97, Apr. 2000, pp. 4541–4550.
[8] A. Bononi and A. Vannucci, "Is there life beyond the principal states of polarization?," Optical Fiber Technol., 2001, submitted for publication.
[9] M. O. van Deventer, "Probability density functions of optical polarization states: Theory and applications," *J. Lightwave Technol.*, vol. 12, pp. 2147–2151, Dec. 1994.
[10] D. Marcuse, C. R. Menyuk, and P. K. A. Wai, "Application of the Manakov-PMD equation to studies of signal propagation in optical fibers with randomly varying birefringence," *J. Lightwave Technol.*, vol. 15, pp. 1735–1745, Sept. 1997.
[11] A. O. Dal Forno, A. Paradisi, R. Passy, and J. P. von der Weid, "Experimental and theoretical modeling of polarization-mode dispersion in single-mode fibers," *IEEE Photon. Technol. Lett.*, vol. 12, pp. 296–298, Mar. 2000.
[12] N. Gisin and J. P. Pelleaux, "Polarization mode dispersion: Time versus frequency domains," *Opt. Commun.*, vol. 89, pp. 316–323, 1992.
[13] A. Vannucci and A. Bononi, "Extracting PMD statistics from single emulated fiber sample," *Electron. Lett.*, vol. 37, pp. 884–886, 2001.
[14] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, Fifth ed: Academic, 1994.
[15] M. Karlsson and J. Brentel, "Autocorrelation function of the polarization-mode dispersion vector," *Opt. Lett.*, vol. 24, pp. 939–941, July 1999.
[16] 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.*, vol. 12, pp. 53–55, Jan. 2000.
[17] M. Shtaif and A. Mecozzi, "Study of the frequency autocorrelation of the differential group delay in fibers with polarization mode dispersion," *Opt. Lett.*, vol. 25, pp. 707–709, May 2000.
[18] C. D. Pagani and S. Salsa, *Analisi Matematica*. Milano, Italy: Masson Ed., 1993, vol. 2.

Armando Vannucci (S'95–M'02) was born in Frosinone, Italy, in 1968. He received the degree in electronics engineering (*cum laude*) from the University of Roma "La Sapienza," Italy, and the Ph.D. degree in information engineering from the University of Parma, Italy, in 1993 and 1998, respectively.

Until 1995, he was with the INFO-COM Department, University of Roma, Italy, doing research activity in the field of acoustic phonetics. Since 1995, he has been with the Dipartimento di Ingegneria dell'Informazione, University of Parma, Italy, where his research activity has been in the field of nonlinear radio channels from 1995 to 1998. Since 1999, his research interests have been in the field of optical transmission and optical communication systems.

Alberto Bononi received the Laurea degree in electronics engineering from the University of Pisa, Italy, and the M.A. and Ph.D degrees in electrical engineering from Princeton University, Princeton, NJ, in 1988, 1992, and 1994, respectively.

Currently, he is an Associate Professor of Telecommunications at the School of Engineering, the Universita' di Parma, Italy. He teaches courses in Probability Theory and Stochastic Processes, Telecommunications Networks, and Optical Communications. In 1990, he worked at GEC-Marconi Hirst Research Centre, Wembley, U.K., on a Marconi S.p.A. project on coherent optical systems. From 1994 to 1996, he was an Assistant Professor in the Electrical and Computer Engineering Department at the State University of New York (SUNY), Buffalo, teaching courses in Electric Circuits and Optical Networks. In the summers of 1997 and 1999, he was a Visiting Faculty at the Departement de Genie Electrique, Universite' Laval, QC, Canada, doing research on fiber amplifiers. His present research interests include system design and performance analysis of high-speed all-optical networks, nonlinear fiber transmission for WDM systems, linear and nonlinear polarization mode dispersion, and transient gain dynamics in doped-fiber and Raman amplifiers.