ACADEMIC PRESS

Available online at www.sciencedirect.com
SCIENCE DIRECT®

Optical Fiber Technology 8 (2002) 257–294

Optical Fiber Technology

www.academicpress.com

Is there life beyond the principal states of polarization?

Alberto Bononi * and Armando Vannucci

Dipartimento di Ingegneria dell'Informazione, Università di Parma, Parma, Italy

Received 7 August 2001; revised 25 May 2002

Abstract

In this paper we show that it is possible to describe exactly the field intensity distorted by polarization mode dispersion (PMD) at the output of a single-mode fiber, without resorting to the principal states of polarization (PSPs). Such description is based on the eigenmodes of what we call the *extracted matrices*, for which we establish relationships with the PSPs, e.g., by showing their alignment with the PSPs when these are not depolarized over the signal bandwidth. We also show that the eigenmodes of the extracted matrices are less depolarized than the PSPs, and are thus the most convenient frame of reference to express the output field and intensity for high bit rate transmitted signals. We thoroughly investigate the properties of a recently introduced fiber model, which we refer to as the *rotation model*, based on the rotation of the extracted eigenmodes, which extends the intuition of Bruyère's PSPs rotation model (1996) to the most stable frame of reference, and hence is expected to give the most accurate description of the PMD fiber for large bandwidth signals. We finally provide a novel eye closure penalty (ECP) formula based on the extracted eigenmodes, and show its relation with the currently available ECP formulas. The accuracy of the novel ECP formula is tested against simulation results of transmission through a fiber synthesized by the rotation model.

Keywords: Polarization mode dispersion (PMD); Optical communications

1. Introduction

The literature on polarization mode dispersion (PMD) is dominated by the use of the principal states of polarization (PSPs) introduced by Poole et al. [1], and their associated Stokes vector, called the PMD vector, as the main tool for the analysis of the received field [2], its statistics [3–6], and its intensity as it affects the system bit-error rate [7,8]. The PSP approach has also a great value to the purpose of PMD compensation, because of the simple PSP concatenation rule for a cascade of optical components [9] and its visualization on the Poincaré sphere [3,10]. As the bandwidth of the transmitted signal increases, the PMD vector cannot be treated as a constant, as in a PMD analysis to the first-order in frequency and the so-called higher-order PMD has to be taken into account. Using the field expansion given in [2], and the joint statistics of the PMD vector and

* Corresponding author.
E-mail address: alberto@tlc.unipr.it (A. Bononi).

1068-5200/02/$ – see front matter
PII: S1068-5200(02)00500-X

of its derivative [4], it is possible to numerically compute the system outage probability with second-order PMD [11]. An alternative approach to deal with higher-order PMD is provided by the frequency-averaged PSPs [12,13], which lead to an exact pulse-broadening formula [12]. Such formula allows simple approximate expressions of the system penalty [14] that have a strong appeal to the purpose of compensation. Another approach is that of Bruyère [15], followed by [16,17], based on the empirical observation that the PSPs depolarize by following circular trajectories. Yet another approach to higher-order PMD is that of Eyal and coworkers [18] based on a decomposition of the Jones Matrix as an infinite product of exponential matrices.

All these works heavily rely on the concept of PSPs. To quote an expression by Haus, it seems that "there is little that one can add to the mathematical framework developed by Poole et al." [19].

Our view is that the single-mode fiber affected by PMD is a complex linear dynamical system, for which many alternative descriptions exist, and no single description seems to be capable of capturing all its features. It is thus important to be able to switch among different descriptions, knowing their relations and their value in highlighting specific system features.

Aim of this paper is to show that there exist several alternative *exact* descriptions of the output field and of its intensity, which are not based on the PSPs, and that there exists an alternative frame of reference that is more stable than the PSPs and is thus more suitable to model the fiber over larger bandwidths.

Our approach uses the spectral decomposition of the fiber Jones matrix, i.e., its eigenvalues and eigenmodes. The key tool in all the analysis is the extensive use of the decomposition of the fiber Jones matrix in its Pauli coordinates, and the treatment of the coordinates vector, called the *Pauli vector,* in the Stokes domain. The paper is organized as follows.

In Section 2 we provide three alternative expressions of the fiber Jones matrix along with the corresponding exact expressions of the output field intensity, and give for comparison an approximate expression based on the PSPs.

Unfortunately, when using the eigenmodes of the fiber Jones matrix $U(\omega)$, one finds that they are more depolarized than the PSPs. We recently found that the correlation bandwidth of the fiber eigenmodes is $\sqrt{2/3}$ that of the PSPs [20,21], which implies they decorrelate in frequency a little more quickly than the PSPs on average. However, one can write

$$U(\omega) = U_0 U_{\mathrm{r}}(\omega) \tag{1}$$

as the product of the Jones matrix at the reference frequency U_0 and of the *right-extracted matrix* $U_{\mathrm{r}}(\omega)$. In Section 3 we spectrally decompose such matrix and show that its eigenmodes, in the limit $\omega \to 0$, are aligned with the input PSPs, and depolarize in frequency at *half the speed* of the PSPs. In other words, they "are" the input PSPs when these do not vary with frequency over the signal bandwidth, and are frequency independent over a larger bandwidth than the PSPs: they are thus *the novel candidate frame of reference* to express the output field when the signal has a bandwidth comparable to that of the PSPs. Similarly one can write

$$U(\omega) = U_{\mathrm{l}}(\omega) U_0 \tag{2}$$

and we will see that the eigenmodes of the *left-extracted matrix* $U_{\mathrm{l}}(\omega)$ are aligned in the limit $\omega \to 0$ with the *output* PSPs. As an instructive example, we show in Appendix B that, whenever the extracted eigenmodes are frequency-independent (and thus so are the PSPs), the fiber eigenmodes describe a great circle on the Poincaré sphere as ω varies.

In Section 4 we investigate the properties of a novel fiber model [22,23] that postulates that, as ω varies, the extracted eigenmodes rotate around a fixed axis. We call this the *rotation model*. As in Bruyère's model [15], the idea is here to approximate the extracted eigenmodes' trajectory, traced on the Poincaré sphere as ω varies, with its osculating circle at $\omega = 0$, i.e., that circle intersecting the trajectory at three coincident points. The bandwidth over which such approximation is valid is clearly the range of frequencies centered at $\omega = 0$ for which the osculating circle is close to the actual trajectory. We believe the rotation model to be the model that more closely emulates the frequency behavior of single-mode fibers affected by PMD over bandwidths exceeding the PSPs bandwidth. For such model, using the intensity expressions of Section 2, we provide the explicit waveform of the output intensity for varying model parameters, given an isolated input pulse. We provide plenty of details on the decomposition of the output intensity in its elemental building blocks. We will learn for instance several details on the shape of the well-known overshoot on one edge of the pulse due to higher-order PMD.

In Section 5 we first provide a unifying survey of the available eye-closure penalty (ECP) formulas for first-order PMD, and then extend the method proposed in [8] to get a generalized Chen formula (GCF) that provides the ECP for all-order PMD. We provide a thorough analysis of the ECP versus input state of polarization (SOP) surface for the rotation model, to check the combined effect of the differential group delay (DGD), the rotation speed, the aperture of the eigenmode circle, and the chirp of the pulse that may be present at the input or induced by common-mode group velocity dispersion (GVD). In so doing, we also check the range of model parameters for which our all-order GCF formula is accurate. We finally show that our GCF formula can also be derived from the small-signal intensity-to-intensity modulation (IM/IM) fiber response in [24]. We verify that the ECP versus input SOP surfaces synthesized with the GCF well capture the *shape* of the actual ECP surfaces, even when the absolute accuracy of the ECP prediction is poor.

Section 6 summarizes our findings and contains our conclusions.

On notation. We indicate 2×1 (Jones) complex column vectors with boldface capital letters, e.g., $\mathbf{A} = [A_1, A_2]^{\mathrm{T}} = [A_1; A_2]$, where the symbol T stands for transpose. The symbol † stands for transpose conjugate. A 2×2 complex matrix appears in capital letter, e.g., U. Stokes 3×1 column vectors appear with an arrow, e.g., $\vec{a}$. Unit magnitude vectors appear with a circumphlex, e.g., $\hat{a}$, $\widehat{A}$. Stokes 4×1 column vectors appear with an underscore, e.g., $\underline{u}$. The symbol $\times$ stands for vector cross product, while $\cdot$ for scalar product. The symbols $\Re[z]$ and $\Im[z]$ indicate the real and imaginary parts of a complex number z, and j is the imaginary unit. The symbol $\triangleq$ means "equal by definition." The symbol $\otimes$ denotes convolution. The symbol $\mathcal{F}$ indicates the Fourier transform and $\mathcal{F}^{-1}$ the inverse transform. In this paper we will use the engineering notation for Fourier transforms, and the sign conventions for Jones and Mueller matrices, consistently with the reference paper [25].

2. Expressions of the output intensity

A linear lossless single-mode optical fiber can be described by the 2×2 Jones matrix $T(\omega) = \mathrm{e}^{-j\bar{\phi}(\omega)} U(\omega)$, where ω is the offset from the reference frequency, $\bar{\phi}(\omega)$ is the phase due to common mode GVD, and $U(\omega)$ is a unitary matrix with $\det[U] = 1$ [25].

Assume the input field is completely polarized: $\mathbf{E}_{\mathrm{i}}(t) = E_{\mathrm{i}}(t)\hat{\mathbf{J}}$ where $\hat{\mathbf{J}} = [J_1; J_2]$ is the input SOP (ISOP) Jones vector and $E_{\mathrm{i}}(t)$ is the complex scalar amplitude. Let $\overline{E}_{\mathrm{o}}(t) \triangleq \mathcal{F}^{-1}[\mathrm{e}^{-j\bar{\phi}(\omega)} E_{\mathrm{i}}(\omega)]$ be the common-phase filtered input amplitude. The output

field, in the frequency domain, is $\overline{E}_{\rm o}(\omega)\hat{\mathbf{J}}_{\rm o}(\omega)$, where $\hat{\mathbf{J}}_{\rm o}(\omega) = U(\omega)\hat{\mathbf{J}}$. Starting from three forms of the Jones matrix U, we now provide three alternative exact expressions of the fiber output intensity, i.e., of the received current in an optical transmission system.

2.1. First form

A unitary matrix with unit determinant can in general be expressed through the Cayley/Klein (CK) form [1,25]

$$U = \begin{bmatrix} U_1 & U_2 \\ -U_2^* & U_1^* \end{bmatrix}, \tag{3}$$

where the complex scalar entries are such that $|U_1|^2 + |U_2|^2 = 1$.

The output SOP is straightforwardly obtained from (3) as

$$\hat{\mathbf{J}}_{\rm o}(\omega) = \big[J_1 U_1(\omega) + J_2 U_2(\omega)\big]\hat{\mathbf{e}}_1 + \big[-J_1 U_2^*(\omega) + J_2 U_1^*(\omega)\big]\hat{\mathbf{e}}_2,$$

where $\hat{\mathbf{e}}_{1,2}$ represent the Jones unit reference axes. If we now define the impulse responses $U_1(t) \triangleq \mathcal{F}^{-1}[U_1(\omega)]$ and $U_2(t) \triangleq \mathcal{F}^{-1}[U_2(\omega)]$, then we express the received field intensity as

$$I(t) = \big|\big[J_1 U_1(t) + J_2 U_2(t)\big] \otimes \overline{E_{\rm o}}(t)\big|^2 + \big|\big[-J_1 U_2^*(-t) + J_2 U_1^*(-t)\big] \otimes \overline{E_{\rm o}}(t)\big|^2,$$

where each term is the magnitude of the projections of the field on the two *orthogonal and ω-independent* reference axes. Although such expression accurately predicts the output intensity, it requires that the full frequency dependence of $U(\omega)$ is measured, giving little analytical insight on the impact of first- and higher-order PMD on $I(t)$.

2.2. Second form

A second form of the output SOP is obtained from the spectral decomposition (SD) of U [26]

$$U = B \begin{bmatrix} \mathrm{e}^{-j\Delta\phi/2} & 0 \\ 0 & \mathrm{e}^{j\Delta\phi/2} \end{bmatrix} B^{\dagger}, \tag{4}$$

where $B(\omega) \triangleq [\widehat{\mathbf{B}}_{\rm s}, \widehat{\mathbf{B}}_{\rm f}]$ is a unitary matrix whose columns are the unit-norm orthogonal eigenmodes $\widehat{\mathbf{B}}_{\rm s}(\omega)$ and $\widehat{\mathbf{B}}_{\rm f}(\omega)$, and $\Delta\phi(\omega)$ is the retardation angle, chosen such that $\Delta\phi(0) \geqslant 0$. This choice justifies the subscript s for the *slow* eigenmode and f for the *fast* one.

From (4), the output SOP is

$$\begin{aligned}\hat{\mathbf{J}}_{\rm o}(\omega) &= \big[\widehat{\mathbf{B}}_{\rm s}(\omega), \widehat{\mathbf{B}}_{\rm f}(\omega)\big] \begin{bmatrix} \mathrm{e}^{-j\Delta\phi(\omega)/2} & 0 \\ 0 & \mathrm{e}^{j\Delta\phi(\omega)/2} \end{bmatrix} \begin{bmatrix} c_{\rm s}(\omega) \\ c_{\rm f}(\omega) \end{bmatrix} \\ &= \big(c_{\rm s}(\omega)\,\mathrm{e}^{-j\Delta\phi(\omega)/2}\big)\widehat{\mathbf{B}}_{\rm s}(\omega) + \big(c_{\rm f}(\omega)\,\mathrm{e}^{j\Delta\phi(\omega)/2}\big)\widehat{\mathbf{B}}_{\rm f}(\omega)\end{aligned}$$

being $c_{\rm s} \triangleq \widehat{\mathbf{B}}_{\rm s}^{\dagger}\hat{\mathbf{J}}$ and $c_{\rm f} \triangleq \widehat{\mathbf{B}}_{\rm f}^{\dagger}\hat{\mathbf{J}}$ the projections of the ISOP onto the *ω-dependent* eigenmodes. Hence by defining

$$\begin{aligned}\mathbf{E}_{\rm s}(\omega) &= \big[c_{\rm s}(\omega)\widehat{\mathbf{B}}_{\rm s}(\omega)\,\mathrm{e}^{-j\Delta\phi(\omega)/2}\big]\overline{E}_{\rm o}(\omega), \\ \mathbf{E}_{\rm f}(\omega) &= \big[c_{\rm f}(\omega)\widehat{\mathbf{B}}_{\rm f}(\omega)\,\mathrm{e}^{j\Delta\phi(\omega)/2}\big]\overline{E}_{\rm o}(\omega)\end{aligned} \tag{5}$$

the field components projected onto the fast and slow eigenmodes, finally, we get

$$I(t) = \left|\mathbf{E}_{\mathrm{s}}(t) + \mathbf{E}_{\mathrm{f}}(t)\right|^2 = \left|\mathbf{E}_{\mathrm{s}}(t)\right|^2 + \left|\mathbf{E}_{\mathrm{f}}(t)\right|^2 + 2\Re\left\{\mathbf{E}_{\mathrm{s}}^{\dagger}(t)\mathbf{E}_{\mathrm{f}}(t)\right\}. \tag{6}$$

We note that the two field components are two (vector) filtered versions of the output scalar field $\overline{E}_{\mathrm{o}}(\omega)$, and if $\Delta\phi(\omega) = \omega C_1$ is linear in ω with $\Delta\phi(0) = 0$, they represent a retarded (slow) and an accelerated (fast) filtered version of the input field. If $\Delta\phi(0) \neq 0$, such constant phase term cancels in (6), so that we can neglect it in the analysis of the output current. The existence of a beat term in (6) is evidence that the eigenmodes are ω-dependent. When the eigenmodes are frequency independent, the fields $\mathbf{E}_{\mathrm{s}}(t)$ and $\mathbf{E}_{\mathrm{f}}(t)$ are orthogonal for every t, and from (6) and (5) we easily get

$$I(t) = \gamma\left|\mathcal{F}^{-1}\left[\mathrm{e}^{-j\Delta\phi(\omega)/2}\overline{E}_{\mathrm{o}}(\omega)\right]\right|^2 + (1-\gamma)\left|\mathcal{F}^{-1}\left[\mathrm{e}^{j\Delta\phi(\omega)/2}\overline{E}_{\mathrm{o}}(\omega)\right]\right|^2, \tag{7}$$

where $\gamma \triangleq |c_{\mathrm{s}}|^2$ is the power splitting factor. Note that such result is known from the literature when reasoning in terms of PSPs, when these are frequency independent [27]. Both interpretations are correct. However, it is way more likely that the PSPs are constant over the signal bandwidth, rather than the global eigenmodes. An explanation of such fact is provided in Appendix B. In Section 3 we will see that (7) is also valid when reasoning in terms of the extracted eigenmodes. However, we will prove in Section 3 that the extracted eigenmodes move in frequency along their trajectory at half the speed of the PSPs, and are thus more stable than the PSPs.

This second form is very convenient when the "frame of reference" (be it the global eigenmodes, the PSPs or the extracted eigenmodes) is fixed over the signal bandwidth. It is also useful to highlight some effects connected to GVD, as shown next.

Expanding $\Delta\phi(\omega) = \omega C_1 + \omega^2 C_2/2$ to second order in Taylor series in ω, and letting $\mathbf{H}_{\mathrm{s}}(\omega) \triangleq \widehat{\mathbf{B}}_{\mathrm{s}}^{\dagger}(\omega)\hat{\mathbf{J}}\widehat{\mathbf{B}}_{\mathrm{s}}(\omega)$, $\mathbf{H}_{\mathrm{f}}(\omega) \triangleq \widehat{\mathbf{B}}_{\mathrm{f}}^{\dagger}(\omega)\hat{\mathbf{J}}\widehat{\mathbf{B}}_{\mathrm{f}}(\omega)$, from (5) we get

$$\begin{aligned}
\mathbf{E}_{\mathrm{s}}(\omega) &= \mathbf{H}_{\mathrm{s}}(\omega)\,\mathrm{e}^{-j\omega C_1/2}\left[\mathrm{e}^{-jC_2\omega^2/4}\overline{E}_{\mathrm{o}}(\omega)\right],\\
\mathbf{E}_{\mathrm{f}}(\omega) &= \mathbf{H}_{\mathrm{f}}(\omega)\,\mathrm{e}^{j\omega C_1/2}\left[\mathrm{e}^{+jC_2\omega^2/4}\overline{E}_{\mathrm{o}}(\omega)\right]
\end{aligned} \tag{8}$$

showing that the second-order derivative of $\Delta\phi(\omega)$ has the same effect of common-mode GVD. It *adds* to GVD for one eigenmode, and *subtracts* for the other eigenmode. This well-known effect, giving a different broadening of the fast and slow signal replicas [2], is usually attributed to the DGD derivative $\Delta\tau_{\omega}$ in the literature, and called polarization-dependent chromatic dispersion [28]. However, its exact analytical impact on the output field cannot be derived, as in (8), unless the PSPs are frequency-independent.

2.3. Third form

We next present a third form of the output intensity, which we believe to be the most insightful and useful for the analysis of the received current in an optical communications system. Such representation is the one that best highlights the dependence of the received intensity on the input SOP, and it is the one we will be most often working with in the rest of the paper.

We start with the Pauli decomposition (PD) form [12,25] of the Jones matrix U

$$U = \mathrm{e}^{-j(\Delta\phi/2)(\hat{b}\cdot\vec{\sigma})} = \cos\left(\frac{\Delta\phi}{2}\right)\sigma_0 - j\sin\left(\frac{\Delta\phi}{2}\right)(\hat{b}\cdot\vec{\sigma}), \tag{9}$$

where $\vec{\sigma}$ is the Pauli spin vector [25], σ_0 is the 2×2 identity matrix [29], and $\hat{b} = [b_1, b_2, b_3]^{\mathrm{T}}$ is the Stokes vector associated with the slow eigenmode. The derivation of (9) from (4) is established using the relation $B\sigma_1 B^{\dagger} = (\hat{b} \cdot \vec{\sigma})$. The 4×1 vector of Pauli coordinates of matrix U is defined as

$$\underline{u} = \big[u_0; \vec{u}\big] = \left[\cos\left(\frac{\Delta\phi}{2}\right); -j \sin\left(\frac{\Delta\phi}{2}\right)\hat{b}\right] \tag{10}$$

so that, by defining $\underline{\sigma} \triangleq [\sigma_0; \vec{\sigma}]$, U can be expressed as $U = \underline{u} \cdot \underline{\sigma}$.

The time-domain expression of the output field is

$$\mathbf{E}_{\mathrm{o}}(t) = \mathcal{F}^{-1}\big[\overline{E}_{\mathrm{o}}(\omega)U(\omega)\hat{\mathbf{J}}\big] \triangleq \int_{-\infty}^{\infty} \overline{E}_{\mathrm{o}}(\omega)U(\omega)\hat{\mathbf{J}}\, \mathrm{e}^{j\omega t}\frac{\mathrm{d}\omega}{2\pi};$$

hence, the intensity $I(t) = \mathbf{E}_{\mathrm{o}}^{\dagger}(t)\mathbf{E}_{\mathrm{o}}(t)$ becomes

$$I(t) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \overline{E}_{\mathrm{o}}^{*}(\omega_1)\overline{E}_{\mathrm{o}}(\omega_2)\hat{\mathbf{J}}^{\dagger}U^{\dagger}(\omega_1)U(\omega_2)\hat{\mathbf{J}}\, \mathrm{e}^{-j\omega_1 t}\, \mathrm{e}^{j\omega_2 t}\frac{\mathrm{d}\omega_1}{2\pi}\frac{\mathrm{d}\omega_2}{2\pi}. \tag{11}$$

Now decompose

$$U^{\dagger}(\omega_1)U(\omega_2) = \sum_{k=0}^{3} n_k(\omega_1, \omega_2)\sigma_k \tag{12}$$

in its Pauli coordinates n_0 and $\vec{n} = [n_1; n_2; n_3]$, so that

$$I(t) = \sum_{k=0}^{3} N_k(t)\big(\hat{\mathbf{J}}^{\dagger}\sigma_k\hat{\mathbf{J}}\big) = N_0(t) + \vec{N}(t) \cdot \hat{j}, \tag{13}$$

where $\hat{j} = \hat{\mathbf{J}}^{\dagger}\vec{\sigma}\hat{\mathbf{J}}$ is the Stokes vector associated with the ISOP $\hat{\mathbf{J}}$, and, for $k = 0, \ldots, 3$

$$N_k(t) \triangleq \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} n_k(\omega_1, \omega_2)\overline{E}_{\mathrm{o}}^{*}(\omega_1)\overline{E}_{\mathrm{o}}(\omega_2)\, \mathrm{e}^{-j\omega_1 t}\, \mathrm{e}^{j\omega_2 t}\frac{\mathrm{d}\omega_1}{2\pi}\frac{\mathrm{d}\omega_2}{2\pi}. \tag{14}$$

Expression (13) is the sought third form, which shows that the received intensity is composed of a SOP-independent term N_0, and a SOP-dependent contribution expressed in terms of a scalar product of the ISOP Stokes vector with the vector $\vec{N}$. All we need next is the explicit form of N_0, $\vec{N}$. To this aim, we wish now to find the Pauli coordinates n_k for $k = 0, \ldots, 3$. Let $\underline{u}(\omega) = [u_0(\omega); \vec{u}(\omega)]$ be the Pauli vector of $U(\omega)$. From (10) we immediately have that $\underline{u}^*$ is the Pauli vector of $U^{\dagger}$. Then using (A.3), in Appendix A, we find the desired Pauli coordinates

$$\begin{aligned} &n_0(\omega_1, \omega_2) = \underline{u}(\omega_1)^{\dagger}\underline{u}(\omega_2), \\ &\vec{n}(\omega_1, \omega_2) = u_0^{*}(\omega_1)\vec{u}(\omega_2) + u_0(\omega_2)\vec{u}^{*}(\omega_1) + j\big(\vec{u}^{*}(\omega_1) \times \vec{u}(\omega_2)\big). \end{aligned} \tag{15}$$

Substitution in (14) gives for the first element

$$N_0(t) = \big|\underline{u}(t) \otimes \overline{E}_{\mathrm{o}}(t)\big|^2, \tag{16}$$

where $u_k(t) = \mathcal{F}^{-1}[u_k(\omega)]$ are the Pauli impulse responses. Similarly, the remaining components are

$$\begin{aligned}\vec{N}(t) &= \big(u_0 \otimes \overline{E}_o\big)^*\big(\vec{u} \otimes \overline{E}_o\big) + \big(u_0 \otimes \overline{E}_o\big)\big(\vec{u} \otimes \overline{E}_o\big)^* + j\big(\vec{u} \otimes \overline{E}_o\big)^* \times \big(\vec{u} \otimes \overline{E}_o\big) \\ &= 2\Re\big\{\big(u_0 \otimes \overline{E}_o\big)\big(\vec{u} \otimes \overline{E}_o\big)^*\big\} - j\big(\vec{u} \otimes \overline{E}_o\big) \times \big(\vec{u} \otimes \overline{E}_o\big)^*, \end{aligned} \tag{17}$$

where all functions are in the time domain. Such relations are valid for *any* Jones matrix U, even non-unitary. We can also derive the following useful explicit expressions:

$$\begin{aligned} N_0(t) &= \sum_{i=0}^{3} \big|u_i(t) \otimes \overline{E}_o(t)\big|^2, \\ N_k(t) &= 2\Re\big[\big(u_0(t) \otimes \overline{E}_o(t)\big)\big(u_k(t) \otimes \overline{E}_o(t)\big)^*\big] \\ &\quad + 2\Im\big[\big(u_{k+1}(t) \otimes \overline{E}_o(t)\big)\big(u_{k+2}(t) \otimes \overline{E}_o(t)\big)^*\big] \end{aligned} \tag{18}$$

for $k = 1, 2, 3$, where we interpret the indices as modulo 3 (e.g., for $k = 3$, $k + 1 = (3 + 1)_{\mathrm{mod}\,3} = 1$).

The third form is clearly the most complex of the three presented, but it is also the one that gives most insight in the analytical mechanisms that determine the overall output intensity. Although seemingly rather involved, the expressions in (18) are extremely useful to explain the analytical reasons of the dependence of the output intensity on the ISOP, as we will see in the next sections. We also note the key role of the Pauli impulse responses $u_k(t)$ as key linear filtering blocks of the scalar field $\overline{E}_o(t)$.

2.3.1. Approximate third form

While the third form is an *exact* expression of the intensity, we now provide an *approximate* third form for $I(t)$ based on the PSPs, whose relationship with the Pauli vector is provided in Appendix A. The key equation (11) contains the product $U^\dagger(\omega_1)U(\omega_2)$ which we now express by expanding both matrices in Taylor series around the reference frequency

$$U^\dagger(\omega_1)U(\omega_2) = \sum_{i=0}^{\infty}\sum_{k=0}^{\infty} U_0^{(i)^\dagger}\frac{\omega_1^i}{i!}U_0^{(k)}\frac{\omega_2^k}{k!},$$

where $U_0^{(i)}$ is the ith derivative of U evaluated at $\omega = 0$. The zeroth order term ($k + i = 0$) is

$$U_0^\dagger U_0 = \sigma_0. \tag{19}$$

The first-order term ($k + i = 1$) is

$$U_0^\dagger U_0'\omega_2 + U_0'^\dagger U_0\omega_1 = -\frac{j}{2}\big(\vec{\Omega}_i \cdot \vec{\sigma}\big)(\omega_2 - \omega_1), \tag{20}$$

where we used the defining relation (A.2) for $\vec{\Omega}_i$ at $\omega = 0$. The second-order term ($k + i = 2$) is

$$\begin{aligned} U_0^\dagger U_0''\frac{\omega_2^2}{2} + U_0'^\dagger U_0'\omega_1\omega_2 + U_0''^\dagger U_0\frac{\omega_1^2}{2} &= -\frac{\Delta\tau^2}{8}(\omega_2 - \omega_1)^2\sigma_0 \\ &\quad - \frac{j}{4}\big(\vec{\Omega}_i' \cdot \vec{\sigma}\big)\big(\omega_2^2 - \omega_1^2\big), \end{aligned} \tag{21}$$

where on the right-hand side we used the relations

$$U_0'^{\dagger}U_0' = \frac{\Delta\tau^2}{4}\sigma_0, \qquad U_0^{\dagger}U_0'' = -\frac{\Delta\tau^2}{4}\sigma_0 - \frac{j}{2}\big(\vec{\Omega}_{\rm i}'\cdot\vec{\sigma}\big),$$

which are straightforward implications of (A.3). Using (19)–(21) we get

$$\hat{\mathbf{J}}^{\dagger}U^{\dagger}(\omega_1)U(\omega_2)\hat{\mathbf{J}} = 1 - \frac{\Delta\tau^2}{8}(\omega_2-\omega_1)^2 - \frac{j}{2}\big(\vec{\Omega}_{\rm i}\cdot\hat{j}\big)(\omega_2-\omega_1) - \frac{j}{4}\big(\vec{\Omega}_{\rm i}'\cdot\hat{j}\big)\big(\omega_2^2-\omega_1^2\big) + O\big(\omega^3\big), \tag{22}$$

where $O(\omega^3)$ represents terms of order three or higher. The input PSP vector and its derivative are evaluated at $\omega = 0$. Substituting (22) in (11) we get

$$I(t) \cong \big|\overline{E}_{\rm o}(t)\big|^2 + \frac{\Delta\tau^2}{4}\left[\left|\frac{\partial}{\partial t}\overline{E}_{\rm o}(t)\right|^2 + \Re\left[\overline{E}_{\rm o}^*(t)\frac{\partial^2}{\partial t^2}\overline{E}_{\rm o}(t)\right]\right] - \big(\vec{\Omega}_{\rm i}\cdot\hat{j}\big)\Re\left[\overline{E}_{\rm o}^*(t)\frac{\partial}{\partial t}\overline{E}_{\rm o}(t)\right] - \frac{1}{2}\big(\vec{\Omega}_{\rm i}'\cdot\hat{j}\big)\Im\left[\overline{E}_{\rm o}^*(t)\frac{\partial^2}{\partial t^2}\overline{E}_{\rm o}(t)\right], \tag{23}$$

which involves the output field and its derivatives up to the second order, and the approximation sign $\cong$ is due to dropping the $O(\omega^3)$ terms. In such expression, it is easy to verify that:

(i) $$\Re\left[\overline{E}_{\rm o}^*(t)\frac{\partial}{\partial t}\overline{E}_{\rm o}(t)\right] = \frac{1}{2}\frac{\partial}{\partial t}\big|\overline{E}_{\rm o}(t)\big|^2;$$

(ii) $$\left[\left|\frac{\partial}{\partial t}\overline{E}_{\rm o}(t)\right|^2 + \Re\left[\overline{E}_{\rm o}^*(t)\frac{\partial^2}{\partial t^2}\overline{E}_{\rm o}(t)\right]\right] = \frac{1}{2}\frac{\partial^2}{\partial t^2}\big|\overline{E}_{\rm o}(t)\big|^2.$$

Hence, using (ii), we can rewrite

$$I(t) \cong \left(1 + \frac{\Delta\tau^2}{8}\frac{\partial^2}{\partial t^2}\right)\big|\overline{E}_{\rm o}(t)\big|^2 - \left\{\vec{\Omega}_{\rm i}\Re\left[\overline{E}_{\rm o}^*(t)\frac{\partial}{\partial t}\overline{E}_{\rm o}(t)\right] + \frac{1}{2}\vec{\Omega}_{\rm i}'\Im\left[\overline{E}_{\rm o}^*(t)\frac{\partial^2}{\partial t^2}\overline{E}_{\rm o}(t)\right]\right\}\cdot\hat{j}. \tag{24}$$

The similarity with (13) is evident: the first and second terms are thus second-order approximations of N_0 and $\vec{N}$, and from (22) and the defining relation for $\vec{\Omega}_{\rm i}$ we see that they come from second-order approximations for the coefficients $n_k(\omega_1,\omega_2)$

$$n_0(\omega_1,\omega_2) \cong 1 - \frac{\Delta\tau^2}{8}(\omega_2-\omega_1)^2,$$
$$\vec{n}(\omega_1,\omega_2) \cong -\frac{j}{2}\left[(\omega_2-\omega_1)\vec{\Omega}_{\rm i} + \frac{1}{2}\big(\omega_2^2-\omega_1^2\big)\vec{\Omega}_{\rm i}'\right].$$

For instance, if $\overline{E}_{\rm o}(t)$ is real (chirpless pulse, no GVD), then the imaginary part in (24) is zero and using (i) we get

$$I(t) \cong \left(1 + \frac{\Delta\tau^2}{8}\frac{\partial^2}{\partial t^2} - \frac{1}{2}\big(\vec{\Omega}_{\rm i}\cdot\hat{j}\big)\frac{\partial}{\partial t}\right)\big|\overline{E}_{\rm o}(t)\big|^2. \tag{25}$$

In Section 5.6 we show that this expression is related to the small-signal IM/IM fiber response.

We remark that different approximate expressions of the output field which use the PSP vector are known [2,11]. However, such expressions inconsistently mix delays and derivatives of the input field, since they do not come from truly second order expansions of matrix U.

Although such form of the intensity in terms of the PMD vector is novel, we derived it only to show the relation with Eq. (13), and we will not investigate on its physical meaning in this paper.

3. Extracted matrices and their properties

In (1) and (2) we defined the *right-* and *left-extracted* (unitary) matrices. All the intensity forms seen in Section 2 can be recast in terms of the extracted matrices, instead of U. This is so, since the output field can be written as $\mathbf{E}_{\rm o}(\omega) = \overline{E}_{\rm o}(\omega)U(\omega)\hat{\mathbf{J}} = U_0\mathbf{E}_{\rm r}(\omega) = \mathbf{E}_{\rm l}(\omega)$, where we define $\mathbf{E}_{\rm r}(\omega) \triangleq \overline{E}_{\rm o}(\omega)U_{\rm r}(\omega)\hat{\mathbf{J}}$ and $\mathbf{E}_{\rm l}(\omega) \triangleq \overline{E}_{\rm o}(\omega)U_{\rm l}(\omega)\hat{\mathbf{J}}_0$, with $\hat{\mathbf{J}}_0 \triangleq U_0\hat{\mathbf{J}}$. Thus the output intensity can be expressed as

$$I(t) = \left|\mathbf{E}_{\rm o}(t)\right|^2 = \left|\mathbf{E}_{\rm r}(t)\right|^2 = \left|\mathbf{E}_{\rm l}(t)\right|^2. \tag{26}$$

This section is devoted to investigate the relations among eigenmodes and PSPs of U (already treated in Appendix A) and those of the extracted matrices.

We know that $\mathcal{M} = \mathrm{e}^{\Delta\phi[\hat{b}\times]}$ is the Mueller matrix associated with the fiber Jones matrix U [25]. Similarly, let $\mathcal{M}_0 = \mathrm{e}^{\Delta\phi_0[\hat{b}_0\times]}$, $\mathcal{M}_{\rm l} = \mathrm{e}^{\Delta\phi_{\rm l}[\hat{b}_{\rm l}\times]}$, and $\mathcal{M}_{\rm r} = \mathrm{e}^{\Delta\phi_{\rm r}[\hat{b}_{\rm r}\times]}$ be the Mueller matrices associated with U_0, $U_{\rm l}$, and $U_{\rm r}$, respectively, where we used the corresponding eigenmodes and retardation angles. The most relevant feature is that the extracted matrices equal the identity matrix at the reference frequency: $U_{\rm r}(0) = U_{\rm l}(0) = \sigma_0$, so that from (10) one gets: $\Delta\phi_{\rm r}(0) = \Delta\phi_{\rm l}(0) = 0$. Also, any vector is an eigenvector at $\omega = 0$, but by continuity we define $\hat{b}_{\rm r}(0) \triangleq \lim_{\omega\to 0}\hat{b}_{\rm r}(\omega)$, and similarly for $\hat{b}_{\rm l}(0)$.

Now note that, since a similarity transformation links $U_{\rm l}$ and $U_{\rm r}$

$$U_{\rm l}(\omega) = U_0U_{\rm r}(\omega)U_0^{\dagger}; \tag{27}$$

then the two matrices have the same eigenvalues, i.e., the same retardation angle: $\Delta\phi_{\rm l}(\omega) = \Delta\phi_{\rm r}(\omega)$ for any ω. Relation (27) in the Stokes domain becomes $\mathrm{e}^{\Delta\phi_{\rm l}[\hat{b}_{\rm l}\times]} = \mathcal{M}_0\,\mathrm{e}^{\Delta\phi_{\rm r}[\hat{b}_{\rm r}\times]}\mathcal{M}_0^{-1} = \mathrm{e}^{\Delta\phi_{\rm r}[(\mathcal{M}_0\hat{b}_{\rm r})\times]}$ [30], which shows that the eigenmode $\hat{b}_{\rm l}$ is a rotated version of $\hat{b}_{\rm r}$

$$\hat{b}_{\rm l}(\omega) = \mathcal{M}_0\hat{b}_{\rm r}(\omega). \tag{28}$$

Next we note that

$$U'(\omega)U^{\dagger}(\omega) = U_{\rm l}'(\omega)U_{\rm l}^{\dagger}(\omega) = U_0U_{\rm r}'(\omega)U_{\rm r}^{\dagger}(\omega)U_0^{\dagger}, \tag{29}$$

from which we conclude that the output PMD vector of $U_{\rm l}$ coincides with that of U: $\vec{\Omega}_{\rm lo}(\omega) = \vec{\Omega}_{\rm o}(\omega)$ because of (A.1). For the same reason, we find that the input PMD vector of $U_{\rm r}$ coincides with that of U: $\vec{\Omega}_{\rm ri}(\omega) = \vec{\Omega}_{\rm i}(\omega)$. With calculations similar to those leading to (28) one can also prove that the PMD vector of $U_{\rm l}$ is a rotated version of that of $U_{\rm r}$, namely: $\vec{\Omega}_{\rm lo}(\omega) = \mathcal{M}_0\vec{\Omega}_{\rm ro}(\omega)$.

Since the output PMD vectors of U and U_l coincide, we have

$$\vec{\Omega}_\mathrm{o}(0) \stackrel{\Delta}{=} \Delta\tau(0)\hat{q}_\mathrm{o}(0) = \vec{\Omega}_\mathrm{lo}(0) = \Delta\phi_\mathrm{l}'(0)\hat{b}_\mathrm{l}(0), \tag{30}$$

where the second equality comes from (A.4) at $\omega = 0$. That is: *the trajectories $\hat{b}_\mathrm{l}(\omega)$ and $\hat{q}_\mathrm{o}(\omega)$ coincide at $\omega = 0$*, so that $\hat{b}_\mathrm{l}(\omega) \cong \hat{q}_\mathrm{o}(\omega)$ in a neighborhood of $\omega = 0$, thereby *justifying the approximation of the output PSPs as the eigenmodes of the left-extracted matrix over a suitably small bandwidth around $\omega = 0$.*[1]

Similarly,

$$\vec{\Omega}_\mathrm{i}(0) \stackrel{\Delta}{=} \Delta\tau(0)\hat{q}_\mathrm{i}(0) = \vec{\Omega}_\mathrm{ri}(0) = \Delta\phi_\mathrm{r}'(0)\hat{b}_\mathrm{r}(0), \tag{31}$$

i.e., the *trajectories $\hat{b}_\mathrm{r}(\omega)$ and $\hat{q}_\mathrm{i}(\omega)$ coincide at $\omega = 0$*, so that $\hat{b}_\mathrm{r}(\omega) \cong \hat{q}_\mathrm{i}(\omega)$ in a neighborhood of $\omega = 0$, thereby *justifying the approximation of the input PSPs with the eigenmodes of the right-extracted matrix over a suitably small bandwidth around $\omega = 0$.*

From (30) and (31) we also conclude that

$$\Delta\tau(0) = \Delta\phi_\mathrm{r}'(0) = \Delta\phi_\mathrm{l}'(0), \tag{32}$$

i.e., when using the extracted matrices the constant C_1 in Section 2.2 is exactly the DGD.

Since $\underline{u}_{\mathrm{r/l}}(0) = [1, 0, 0, 0]^\mathrm{T}$, using (A.7) at $\omega = 0$ yields

$$\vec{u}_\mathrm{r}''(0) = -j\vec{\Omega}_\mathrm{i}'(0)/2, \qquad \vec{u}_\mathrm{l}''(0) = -j\vec{\Omega}_\mathrm{o}'(0)/2. \tag{33}$$

Differentiating twice the expression of $\vec{u}$ in (10) and plugging into (33) at $\omega = 0$, we get

$$\begin{aligned}
\vec{\Omega}_\mathrm{i}'(0) &\stackrel{\Delta}{=} \Delta\tau'(0)\hat{q}_\mathrm{i}(0) + \Delta\tau(0)\vec{q}_\mathrm{i}'(0) = \Delta\phi_\mathrm{r}''(0)\hat{b}_\mathrm{r}(0) + \Delta\phi_\mathrm{r}'(0)2\vec{b}_\mathrm{r}'(0),\\
\vec{\Omega}_\mathrm{o}'(0) &\stackrel{\Delta}{=} \Delta\tau'(0)\hat{q}_\mathrm{o}(0) + \Delta\tau(0)\vec{q}_\mathrm{o}'(0) = \Delta\phi_\mathrm{l}''(0)\hat{b}_\mathrm{l}(0) + \Delta\phi_\mathrm{l}'(0)2\vec{b}_\mathrm{l}(0).
\end{aligned} \tag{34}$$

We can clearly spot out corresponding terms on the two sides of the equality. Now, since $\hat{b}_\mathrm{r}$ is orthogonal to its derivative $\vec{b}_\mathrm{r}'$, and similarly for $\hat{b}_\mathrm{l}$ and $\hat{q}_\mathrm{i/o}$, taking the scalar product of both sides of (34) by $\hat{b}_\mathrm{r}(0) = \hat{q}_\mathrm{i}(0)$ (top row), or by $\hat{b}_\mathrm{l}(0) = \hat{q}_\mathrm{o}(0)$ (bottom row), proves that

$$\Delta\tau'(0) = \Delta\phi_\mathrm{r}''(0) = \Delta\phi_\mathrm{l}''(0), \tag{35}$$

which implies that, *when using the extracted matrices, the constant C_2 in* (8) *is exactly the derivative of the DGD.* Note that an instance of the results in (32) and (35) appears in [16]. Hence from (34) we establish the *key* relation

$$\vec{b}_\mathrm{l}'(0) = \vec{q}_\mathrm{o}'(0)/2, \qquad \vec{b}_\mathrm{r}'(0) = \vec{q}_\mathrm{i}'(0)/2, \tag{36}$$

which states that *the eigenmodes of the extracted matrices move at half the velocity of the PSPs in a neighborhood of $\omega = 0$*, thus proving that *they are more stable* (*i.e., less*

[1] Note that there is a formal coincidence between the definition of the difference rotation matrix $\mathcal{M}_\Delta(\omega_0) = \mathcal{M}(\omega_0 + \Delta\omega)\mathcal{M}^{-1}(\omega_0)$ in the Mueller matrix method (MMM) [31] and our definition of the left-extracted Mueller matrix $\mathcal{M}_\mathrm{l}$, where ω_0 is the measurement frequency and $\Delta\omega$ is the small, fixed deviation chosen in the MMM. The purposes of the MMM is to measure $\vec{\Omega}_\mathrm{o}(\omega_0)$ as the eigenmode of $\mathcal{M}_\Delta(\omega_0)$ by varying ω_0, in accord with (30), where $\omega_0 = 0$. The major novelty in our treatment of $\mathcal{M}_\Delta \equiv \mathcal{M}_\mathrm{l}$ is that we fix the reference frequency and let the frequency deviation ω freely vary.

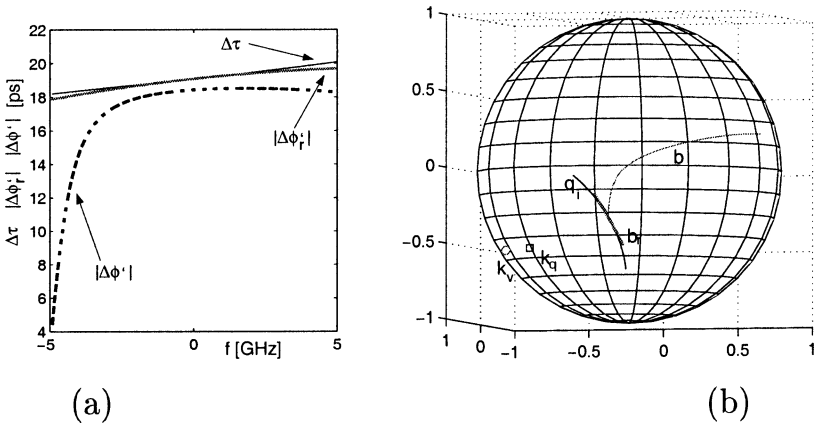

Fig. 1. (a) $\Delta\tau$ (thin solid), $\Delta\phi'$ (dashed), and $\Delta\phi'_{\mathrm{l/r}}$ (thick) versus frequency; (b) global eigenmode $\hat{b}(\omega)$ (thin), right-extracted eigen-mode $\hat{b}_{\mathrm{r}}(\omega)$ (thick), and input PSP $\hat{q}_{\mathrm{i}}(\omega)$ (thin) on the same bandwidth. The rotation axes $\hat{k}_{\mathrm{v}}$ and $\hat{k}_{\mathrm{q}}$ of the osculating circles of $\hat{b}_{\mathrm{r}}(\omega)$ and $\hat{q}_{\mathrm{i}}(\omega)$ at $\omega = 0$ are also shown.

depolarized) *than the PSPs, and that* the trajectories $\hat{b}_{\mathrm{r/l}}(\omega)$ and $\hat{q}_{i/\mathrm{o}}(\omega)$ also have the same tangent at $\omega = 0$.

Figure 1(a) shows a plot of $\Delta\tau$, $\Delta\phi'$, $\Delta\phi'_{\mathrm{l/r}}$, versus frequency for a fiber synthesized with the standard waveplate model [32], with 100 plates and an rms DGD of 20 ps. The figure confirms the coincidence of $\Delta\tau$ with $\Delta\phi'_{\mathrm{l/r}}$ at the reference frequency up to the first derivative. We also note from (A.5) that the gap between $\Delta\tau$ and $\Delta\phi'$ is due to the eigenmodes depolarization, so that the two curves get close when the depolarization is small. Figure 1(b) shows a plot of the trajectories of $\hat{b}_{\mathrm{r}}(\omega)$ and $\hat{q}_{\mathrm{i}}(\omega)$ on the Poincaré sphere, which touch each other at $\omega = 0$, with the same tangent, as well as the rotation axes of their osculating circles at $\omega = 0$. The trajectory $\hat{b}(\omega)$ is also shown, and it is noted that for the selected waveplates realization the global eigenmode $\hat{b}(\omega)$ is way more depolarized than the PMD vector and the extracted eigenmodes vector. This is no coincidence. In fact, in Appendix B we prove that, on the bandwidth for which the extracted eigenmode can be considered frequency independent (which from (36) is larger than the bandwidth over which the PSP is frequency independent) the global eigenmode describes an arc of a *great circle* at non-uniform speed, which in the limit is zero when $\hat{b}_0 \cdot \hat{b}_{\mathrm{r}} = 1$. Incidentally note that Bruyère postulated in his rotation model [15] that *the PSPs* describe great circles, which in general is not true.

For a complete statistical characterization, more fiber realizations of the above 100-waveplate model should be considered. Clearly, when observed on a small bandwidth around $\omega = 0$, both trajectories $\hat{b}_{\mathrm{r}}(\omega)$ and $\hat{q}_{\mathrm{i}}(\omega)$ should for most waveplates realizations coincide with their osculating circles, and from (36) the length of the arc described by $\hat{b}_{\mathrm{r}}(\omega)$ should be half of that described by $\hat{q}_{\mathrm{i}}(\omega)$. This is checked in Fig. 2(a), which shows the length of the trajectory $\hat{b}_{\mathrm{r}}(\omega)$ versus the length of the trajectory $\hat{q}_{\mathrm{i}}(\omega)$, for 1000 fiber samples with an rms DGD of 20 ps. The trajectories are evaluated by sampling $\hat{b}_{\mathrm{r}}(\omega)$ and $\hat{q}_{\mathrm{i}}(\omega)$ at 128 frequency values in a bandwidth of ± 5 GHz around $\omega = 0$, as we did in Fig. 1(b). All lengths are normalized to the length 2π of a great circle. We see that most realizations lie on the straight line with slope $1/2$, as we expect. However, when increasing the rms DGD to 100 ps and observing the trajectories on the same 10 GHz bandwidth, which now equals the inverse of the rms DGD, we note from Fig. 2(b) a much wider spread from the straight line with slope $1/2$, but still with an average slope of 0.58 and with only a few realizations having the PSPs less depolarized than the extracted eigenmodes.

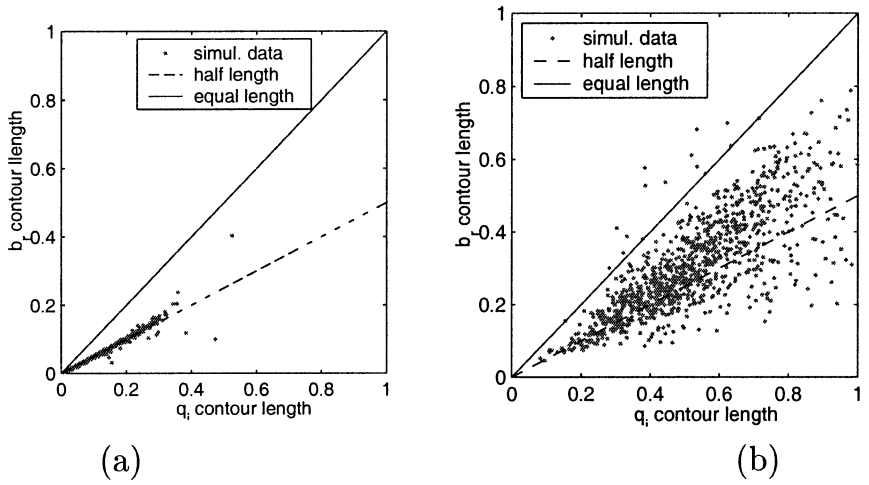

Fig. 2. Contour lengths of $\hat{b}_{\mathrm{r}}(\omega)$ vs. $\hat{q}_{\mathrm{i}}(\omega)$, normalized to 2π, computed on 1000 fiber samples with rms DGD equal to (a) 20 ps, (b) 100 ps, on a 10 GHz bandwidth.

Such deviations are due to higher-order PMD terms that render $\hat{b}_{\mathrm{r}}(\omega) \cong \hat{q}_{\mathrm{i}}(\omega)/2$ no longer a good approximation when the rms DGD is longer than the bit duration.

4. The eigenmodes rotation model

The first-order PMD fiber model with frequency independent PSPs over the transmitted signal bandwidth fails to be accurate with installed single-mode fibers already over terrestrial fiber-optic links and bit rates of 40 Gb/s per channel. The second-order PMD fiber model [4] comes from the Taylor expansion of the PMD vector to first order. The assumed linear variation of the PMD vector with ω gives a questionable prediction of the actual trajectory $\hat{q}_{\mathrm{i}}(\omega)$: a great circle passing through to $\hat{q}_{\mathrm{i}}(\omega)$ at $\omega = 0$. An alternative approach to second- and higher-order PMD was taken by Bruyère [15,33], who thought of more closely approximating the PSP trajectory $\hat{q}_{\mathrm{i}}(\omega)$ by using its osculating circle. However, Bruyère postulated that $\hat{q}_{\mathrm{i}}(\omega)$ always follows great circles, and also confused the PSPs with the fiber global eigenmodes. The confusion was partly clarified in [16], where however the authors kept assuming that their modified PSPs (which actually are our extracted eigenmodes) follow great circles on the Poincaré sphere. The first paper that tacitly used the extracted eigenmodes, and postulated that they follow a circular trajectory, not necessarily a great circle, by rotating around a fixed rotation axis at a constant angular speed was [22]. We will call such fiber model the *rotation model* [23]. The authors in [22] well realized the difference with Bruyère PSP rotation model. A justification of their choice can be traced back to relation (36): since the extracted eigenmode depolarizes at half the speed of the PSP, the rotation model should give an accurate fiber description over a bandwidth larger than that of the model of Bruyère. Experimental and simulation evidence that the rotation model is indeed very accurate comes from the performance of the "all-order" PMD compensator proposed in the same paper [22]. Our simulations also confirm that most fibers synthesized with the retarded waveplate model with a given rms DGD can be fitted by a rotation model that has a very similar frequency response over a bandwidth of the order of the inverse of the rms DGD. However, the focus of this paper is not on the validation of the rotation model, but on the study of its features.

In this section we explicitly derive the rotation model in the light of what we previously learned on the extracted matrices, and then extensively use our third form of the received intensity to describe the characteristics of the received intensity as a function of the rotation model parameters. Our results can be considered as generalizations of the six-pulse model of Francia et al. [33,34] to the more general case of non-equatorial trajectories of the extracted eigenmodes.[2]

Mathematically we describe the rotation model by imposing that the eigenmodes of the right-extracted matrix U_r satisfy the following:

$$\frac{\partial}{\partial\omega}\hat{b}_\mathrm{r}(\omega) = 2k_\mathrm{v}\hat{k}\times\hat{b}_\mathrm{r}(\omega), \quad \hat{b}_\mathrm{r}(0) = \hat{b}_0,$$

where for simplicity of notation we redefine, here and up to the end of the paper, $\hat{b}_0 \triangleq \lim_{\omega\to 0}\hat{b}_\mathrm{r}(\omega)$, which should not be confused with the eigenmode of U_0. The above equation states that the eigenmode rotates at a constant angular speed $2k_\mathrm{v}$ around the fixed axis $\hat{k}$, starting from $\hat{b}_0$ at $\omega = 0$. The rotation is counterclockwise for $\omega > 0$, $2k_\mathrm{v} > 0$. The solution of the motion is

$$\hat{b}_\mathrm{r}(\omega) = \mathrm{e}^{2k_\mathrm{v}\omega[\hat{k}\times]}\hat{b}_0, \tag{37}$$

which can be explicitly written as [25]

$$\hat{b}_\mathrm{r}(\omega) = \cos(2k_\mathrm{v}\omega)\hat{b}_0 + 2\sin^2(k_\mathrm{v}\omega)\big(\hat{k}\cdot\hat{b}_0\big)\hat{k} + \sin(2k_\mathrm{v}\omega)\big[\hat{k}\times\hat{b}_0\big]. \tag{38}$$

It is possible to extend the model to the non-uniform rotation speed simply by replacing in the above expression the term $2k_\mathrm{v}\omega$ with a term $\Phi_\mathrm{v}(\omega) = 2k_\mathrm{v}\omega + \delta\Phi_\mathrm{v}(\omega)$, the last term taking account of higher-order terms in the Taylor expansion of $\Phi_\mathrm{v}(\omega)$.

The Mueller matrix of U_r for the rotation model can be written as [30]

$$\mathcal{M}_\mathrm{r}(\omega) = \mathrm{e}^{\Delta\phi_\mathrm{r}(\omega)[\hat{b}_\mathrm{r}(\omega)\times]} = \big[\mathrm{e}^{2k_\mathrm{v}\omega[\hat{k}\times]}\big]\,\mathrm{e}^{\Delta\phi_\mathrm{r}(\omega)[\hat{b}_0\times]}\big[\mathrm{e}^{-2k_\mathrm{v}\omega[\hat{k}\times]}\big].$$

We interpret $\mathcal{M}_\mathrm{I} \triangleq \mathrm{e}^{\Delta\phi_\mathrm{r}(\omega)[\hat{b}_0\times]}$ as the Mueller matrix associated with the unitary Jones matrix

$$U_\mathrm{I}(\omega) = \mathrm{e}^{-j(\Delta\phi_\mathrm{r}(\omega)/2)(\hat{b}_0\cdot\vec{\sigma})} = B_0\begin{bmatrix}\mathrm{e}^{-j\Delta\phi_\mathrm{r}(\omega)/2} & 0\\ 0 & \mathrm{e}^{j\Delta\phi_\mathrm{r}(\omega)/2}\end{bmatrix}B_0^\dagger$$

corresponding to a first-order model with frequency-independent unitary eigenmode matrix B_0, whose first column is the Jones vector associated with the Stokes vector $\hat{b}_0$. We also interpret $\mathcal{M}_\mathrm{R} \triangleq \mathrm{e}^{2k_\mathrm{v}\omega[\hat{k}\times]}$ as the Mueller matrix associated with the unitary Jones matrix

$$R(\omega) = \mathrm{e}^{-jk_\mathrm{v}\omega(\hat{k}\cdot\vec{\sigma})} = K\begin{bmatrix}\mathrm{e}^{-jk_\mathrm{v}\omega} & 0\\ 0 & \mathrm{e}^{jk_\mathrm{v}\omega}\end{bmatrix}K^\dagger,$$

[2] Note that in [33,34], as well as in [15], the fiber Jones matrix was "spectrally" decomposed by erroneously using the PSPs, instead of the correct decomposition (4) which uses the eigenmodes. Hence all results in [33,34] are correct, provided that the word "PSPs" is changed in "(extracted) eigenmodes."

where K is the unitary matrix whose first column is the Jones vector associated with the Stokes vector $\hat{k}$. Therefore we have

$$U_{\mathrm{r}} = K \begin{bmatrix} \mathrm{e}^{-jk_{\mathrm{v}}\omega} & 0 \\ 0 & \mathrm{e}^{jk_{\mathrm{v}}\omega} \end{bmatrix} F \begin{bmatrix} \mathrm{e}^{-j\Delta\phi_{\mathrm{r}}(\omega)/2} & 0 \\ 0 & \mathrm{e}^{j\Delta\phi_{\mathrm{r}}(\omega)/2} \end{bmatrix} F^{\dagger} \begin{bmatrix} \mathrm{e}^{jk_{\mathrm{v}}\omega} & 0 \\ 0 & \mathrm{e}^{-jk_{\mathrm{v}}\omega} \end{bmatrix} K^{\dagger},$$

where $F \triangleq K^{\dagger} B_0$. Such structure is the foundation of the compensator proposed in [22], which implements the inverse of such global matrix.

4.1. Expression of the output intensity

We now present the exact expression of the received intensity for the rotation model, using our third intensity formula (13). Without loss of generality we assume a reference frame $(\hat{s}_1, \hat{s}_2, \hat{s}_3)$ on the Poincaré sphere in which the rotation axis $\hat{k} \equiv \hat{s}_3$, and $\hat{b}_0$ is in the $(\hat{s}_1, \hat{s}_3)$ plane, and let φ be the circle *aperture angle*, i.e., the angle from $\hat{k}$ to $\hat{b}_0$. We refer to the right-extracted matrix, and make the analysis most general by allowing a non-uniform rotation speed

$$\hat{b}(\omega) = \begin{bmatrix} \cos\Phi_{\mathrm{v}}(\omega)\sin\varphi \\ \sin\Phi_{\mathrm{v}}(\omega)\sin\varphi \\ \cos\varphi \end{bmatrix}. \tag{39}$$

We also account for a general retardation angle where $\Delta\phi(\omega) = \Delta\phi_1\omega + \delta\Delta\phi(\omega)$, where $\delta\Delta\phi(\omega) \triangleq \sum_{i=2}^{\infty} \Delta\phi_{\mathrm{i}}\omega^i / i!$ accounts for all the higher-order terms in a Taylor expansion, and $\Delta\phi_1 = \Delta\tau$ as per (32). Our target is to get explicit expressions for the intensity coefficients $N_0(t), \ldots, N_3(t)$ in (18). To this aim, we now introduce the following functions:

$$p_{\mathrm{f}}(t) \triangleq \mathcal{F}^{-1}\big[\overline{E}_{\mathrm{o}}(\omega)\,\mathrm{e}^{j\delta\Delta\phi(\omega)/2}\big], \qquad p_{\mathrm{s}}(t) \triangleq \mathcal{F}^{-1}\big[\overline{E}_{\mathrm{o}}(\omega)\,\mathrm{e}^{-j\delta\Delta\phi(\omega)/2}\big],$$

and the two functions

$$f_{\pm}(t) \triangleq \frac{1}{2}\left(p_{\mathrm{f}}\left(t + \frac{\Delta\tau}{2}\right) \pm p_{\mathrm{s}}\left(t - \frac{\Delta\tau}{2}\right)\right).$$

Similarly we define

$$\begin{aligned}
p_{\mathrm{ff}}(t) &\triangleq \mathcal{F}^{-1}\big[\overline{E}_{\mathrm{o}}(\omega)\,\mathrm{e}^{j(\delta\Delta\phi(\omega)/2+\delta\Phi_{\mathrm{v}}(\omega))}\big], \\
p_{\mathrm{fs}}(t) &\triangleq \mathcal{F}^{-1}\big[\overline{E}_{\mathrm{o}}(\omega)\,\mathrm{e}^{j(\delta\Delta\phi(\omega)/2-\delta\Phi_{\mathrm{v}}(\omega))}\big], \\
p_{\mathrm{sf}}(t) &\triangleq \mathcal{F}^{-1}\big[\overline{E}_{\mathrm{o}}(\omega)\,\mathrm{e}^{j(-\delta\Delta\phi(\omega)/2+\delta\Phi_{\mathrm{v}}(\omega))}\big], \\
p_{\mathrm{ss}}(t) &\triangleq \mathcal{F}^{-1}\big[\overline{E}_{\mathrm{o}}(\omega)\,\mathrm{e}^{j(-\delta\Delta\phi(\omega)/2-\delta\Phi_{\mathrm{v}}(\omega))}\big]
\end{aligned}$$

and the two functions

$$\begin{aligned}
g_{\mathrm{s}}(t) &\triangleq \frac{1}{2}\left(p_{\mathrm{fs}}\left(t + \frac{\Delta\tau}{2} - 2k_{\mathrm{v}}\right) - p_{\mathrm{ss}}\left(t - \frac{\Delta\tau}{2} - 2k_{\mathrm{v}}\right)\right), \\
g_{\mathrm{f}}(t) &\triangleq \frac{1}{2}\left(p_{\mathrm{ff}}\left(t + \frac{\Delta\tau}{2} + 2k_{\mathrm{v}}\right) - p_{\mathrm{sf}}\left(t - \frac{\Delta\tau}{2} + 2k_{\mathrm{v}}\right)\right).
\end{aligned}$$

From (10) we immediately see that $u_0(t) \otimes \overline{E}_o(t) = f_+(t)$ and

$$\vec{u}(t) \otimes \overline{E}_o(t) = -\left[\frac{\sin\varphi}{2}\big(g_s(t) + g_f(t)\big),\, j\frac{\sin\varphi}{2}\big(g_s(t) - g_f(t)\big),\, \cos\varphi f_-(t)\right]^T.$$

Plug such expressions into (18) to get

$$\begin{aligned} N_0(t) &= \sum_{i=0}^{3}\big|u_i(t) \otimes \overline{E}_o(t)\big|^2 \\ &= |f_+|^2 + \frac{\sin^2\varphi}{4}|g_s + g_f|^2 + \frac{\sin^2\varphi}{4}|g_s - g_f|^2 + \cos^2\varphi|f_-|^2 \\ &= |f_+|^2 + \cos^2\varphi|f_-|^2 + \frac{\sin^2\varphi}{2}\big(|g_s|^2 + |g_f|^2\big) \\ &= \big|f_+(t)\big|^2 + \big|f_-(t)\big|^2 + \sin^2\varphi\left(\frac{|g_s(t)|^2 + |g_f(t)|^2}{2} - \big|f_-(t)\big|^2\right) \geqslant 0 \end{aligned} \tag{40}$$

and such quantity is always positive. Moreover, from (18) we get

$$\begin{aligned} N_1(t) &= \sin\varphi\Re\big[-f_+^*(t)\big(g_s(t) + g_f(t)\big) + \cos\varphi f_-^*(t)\big(g_s(t) - g_f(t)\big)\big], \\ N_2(t) &= \sin\varphi\Im\big[f_+^*(t)\big(g_s(t) - g_f(t)\big) - \cos\varphi f_-^*(t)\big(g_s(t) + g_f(t)\big)\big], \\ N_3(t) &= -2\cos\varphi\Re\big[f_+^*(t)f_-(t)\big] - \frac{\sin^2\varphi}{2}\big(\big|g_s(t)\big|^2 - \big|g_f(t)\big|^2\big). \end{aligned} \tag{41}$$

These are our target expressions. It is useful to note alternative forms of some of their terms

$$\big|f_+(t)\big|^2 + \big|f_-(t)\big|^2 = \frac{1}{2}\left(\left|p_f\left(t + \frac{\Delta\tau}{2}\right)\right|^2 + \left|p_s\left(t - \frac{\Delta\tau}{2}\right)\right|^2\right), \tag{42}$$

$$2\Re\big[f_+^*(t)f_-(t)\big] = \frac{1}{2}\left(\left|p_f\left(t + \frac{\Delta\tau}{2}\right)\right|^2 - \left|p_s\left(t - \frac{\Delta\tau}{2}\right)\right|^2\right). \tag{43}$$

Equations (40) and (41), along with (13), are the generalizations of Francia's six-pulse model, but they refer to the received *intensity*, for a most general non-equatorial trajectory with aperture angle $\varphi \neq 90°$.

Special cases

Case 1 ($\Phi_v(\omega) = 0$: First-order PMD model). Here we have $g_s = g_f = f_-$. Hence from (40)–(43) we get

$$\begin{aligned} N_0(t) &= \tfrac{1}{2}\left(\left|p_f\left(t + \frac{\Delta\tau}{2}\right)\right|^2 + \left|p_s\left(t - \frac{\Delta\tau}{2}\right)\right|^2\right), \\ N_1(t) &= -\frac{\sin\varphi}{2}\left(\left|p_f\left(t + \frac{\Delta\tau}{2}\right)\right|^2 - \left|p_s\left(t - \frac{\Delta\tau}{2}\right)\right|^2\right), \\ N_2(t) &= 0, \\ N_3(t) &= -\tfrac{\cos\varphi}{2}\left(\left|p_f\left(t + \frac{\Delta\tau}{2}\right)\right|^2 - \left|p_s\left(t - \frac{\Delta\tau}{2}\right)\right|^2\right). \end{aligned}$$

Thus, if the ISOP is aligned with $\hat{b}_0 \equiv [\sin\varphi, 0, \cos\varphi]^T$ we get: $I(t) = N_0 + j_1N_1 + j_3N_3 = |p_s(t - \Delta\tau/2)|^2$, as it should.

Case 2 ($\varphi = \pi/2$: Equatorial trajectory for the eigenmode). If also the auxiliary functions $f_\pm$, g_s, g_f are real, then from (40), (41) we get

$$N_0(t) = f_+(t)^2 + \tfrac{1}{2}\big(g_\mathrm{f}(t)^2 + g_\mathrm{s}(t)^2\big), \qquad N_1(t) = -f_+(t)\big(g_\mathrm{f}(t) + g_\mathrm{s}(t)\big),$$
$$N_2(t) = 0, \qquad N_3(t) = \tfrac{1}{2}\big(g_\mathrm{f}(t)^2 - g_\mathrm{s}(t)^2\big). \tag{44}$$

This occurs for instance when the input field is real (chirpless), no GVD is present, and $\Delta\phi_\mathrm{r}(\omega) = \Delta\tau\,\omega$ (linear retardation), $\Phi_\mathrm{v}(\omega) = 2k_\mathrm{v}\omega$ (constant rotation speed). In this case $p_\mathrm{ff} = p_\mathrm{ss} = p_\mathrm{sf} = p_\mathrm{fs} = p_\mathrm{f} = p_\mathrm{s} \triangleq p$ is the input pulse, and $f_\pm(t) = 0.5[p(t + \Delta\tau/2) \pm p(t - \Delta\tau/2)]$, $g_\mathrm{s}(t) = f_-(t - 2k_\mathrm{v})$, $g_\mathrm{f}(t) = f_-(t + 2k_\mathrm{v})$.

Figure 3 shows a graph of f_+, g_s, and g_f for a perfectly rectangular (NRZ) input pulse $p(t)$ of duration T. The graphs are slightly offset to ease the reading of the figure. From such figure and (44) we note that the received intensity can spread at most by $\Delta\tau/2 + 2k_\mathrm{v}$ on each side of the original NRZ pulse, which determines the intersymbol interference (ISI) depth due to such model of PMD. We are in the case $\Delta\tau/2 < 2k_\mathrm{v}$, typical of a small DGD fiber.

For such building functions, Fig. 4, leftmost column, shows the received intensity for three distinct ISOPs:

(i) $\hat{j} = \hat{b}_0 \equiv \hat{s}_1$ corresponding to the case of "transmission over a PSP" [35], i.e., alignment of the ISOP with the input PSP at the reference frequency, or in our terms, alignment with the right-extracted eigenmode at $\omega = 0$. The output intensity (13) in such case is $I(t) = N_0 + N_1$. We note the well-known overshoot and dip in the received pulse [17,33,34], both of duration $\Delta\tau$. From (44) and Fig. 3 we see that, at the overshoot, we have $N_1 = 1\cdot(0.5+0) = 0.5$ and $N_0 = 1^2 + 0.5[0.5^2 + 0^2] = 1.125$, so that at the peak the intensity is $1.125 + 0.5 = 1.625$. Similarly the value at the dip is $1.125 - 0.5 = 0.625$. We note that *the intensity is mostly distorted in this case of transmission over a PSP*, and the overshoot is at the *trailing edge* when the ISOP is aligned with the *slow* eigenmode. Such observations partly explain the results in [17, Fig. 4], although in such reference GVD is also present, whose effects will be discussed in Section 5.5.

(ii) $\hat{j} = \vec{b}'(0)/|\vec{b}'(0)| \equiv \hat{s}_2$ corresponding to an ISOP in the direction of the derivative of $\vec{b}(\omega)$ at $\omega = 0$, i.e., the direction of the tangent to the trajectory at such point. The

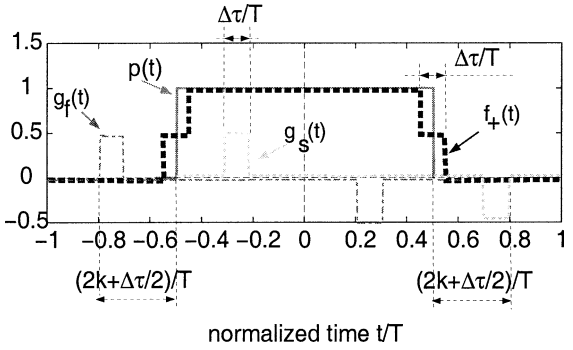

Fig. 3. Building functions $f_+(t)$, $g_\mathrm{s}(t)$, and $g_\mathrm{f}(t)$ for an input NRZ pulse $p(t)$ of duration T, with $\Delta\phi_\mathrm{r}'(0) = \Delta\tau = 0.1T$, $2k_\mathrm{v} = 0.25T$. Graphs are slightly offset for clarity.

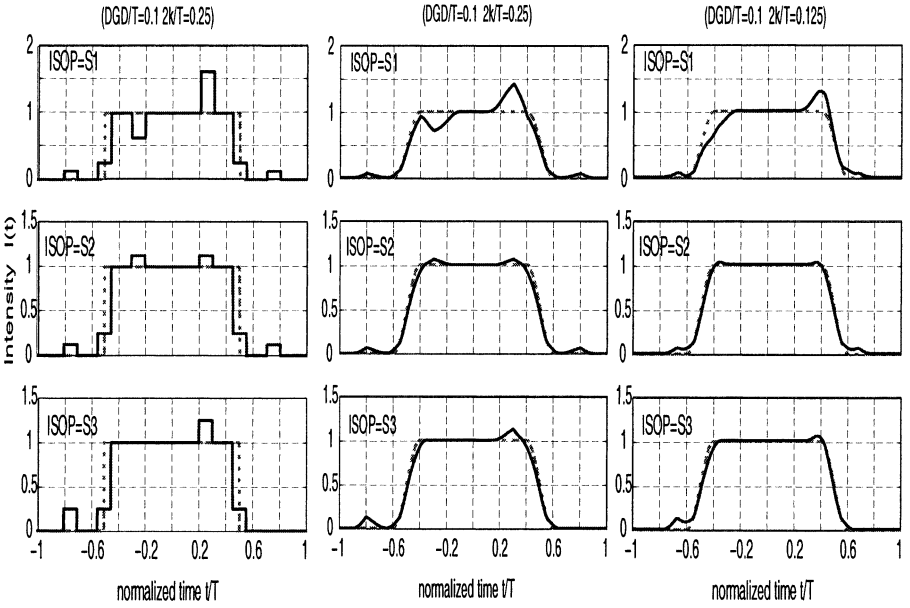

Fig. 4. Rotation model with $\hat{k} = \hat{s}_3$, $\hat{b}_0 = \hat{s}_1$ ($\varphi = \pi/2$), $\Delta\tau/2 < 2k_v$. Left column: intensity for three orthogonal ISOPs, $\Delta\tau/T = 0.1$, roll-off 0.0. Center and right columns: same, but with roll-off 0.2, and $2k_v/T = 0.25$ and 0.125, respectively.

output intensity in such case is $I(t) = N_0(t)$, since $N_2 = 0$, and is an even function of time.

(iii) $\hat{j} = \hat{k} \equiv \hat{s}_3$, corresponding to an ISOP aligned with the rotation axis. Here $I(t) = N_0 + N_3$.

The three plots in the center column in Fig. 4 show the same graphs as those in the leftmost column, but when the input pulse has a raised cosine shape (in power) with roll-off 0.2. The effect of a pulse smoother than a squared NRZ is to smooth out the discontinuities, so that mainly the peaks and dips emerge, and the level of the overshoot is less than 1.625, and gets smaller and smaller as the roll-off increases. The three plots in the rightmost column show the case $\Delta\tau/T = 0.1$, $2k_v/T = 0.125$, i.e., a smaller rotation speed, still with roll-off $\alpha = 0.2$, where we note that the overshoot level is smaller and the dip does not appear, since it falls on the flank of the pulse, being the DGD and the rotation speed too small. Note that long strings of consecutive ones in a digital on off keying (OOK) NRZ transmission look like a single pulse, with a roll-off much smaller than that of the single pulse. To guess the effect of PMD on such strings, one can refer to the single pulse deformation, and rescale the values $\Delta\tau/T$ and $2k_v/T$ by the duration of the string. We can thus see that *smoothing the single pulse by using a larger roll-off has the effect of smoothing the deformations due to PMD on the isolated pulse, the overshoots reappearing on blocks of consecutive marks of sufficient length.* Hence, given a sufficient number of consecutive ones, the overshoot will reach the peak value 1.625.

Finally, here is an argument to explain why the intensity is quite similar in the cases $\hat{j} = \hat{s}_2$ and $\hat{j} = \hat{s}_3$, and is different in the case $\hat{j} = \hat{s}_1$. From Fig. 3 we note that $\frac{1}{2}(g_f(t)^2 + g_s(t)^2)$ and $\frac{1}{2}(g_f(t)^2 - g_s(t)^2)$ are small with respect to $f_+(t)^2$. Hence in (44) we can approximate $N_0(t) \cong f_+^2(t)$, $N_3(t) \cong 0$, so that both when $\hat{j} = \hat{s}_2$ and $\hat{j} = \hat{s}_3$ we have $I(t) \cong N_0(t) \cong f_+^2(t)$. What really makes the difference in the case $\hat{j} = \hat{s}_1$ is the *beat term* $f_+(t)(g_f(t) + g_s(t))$ which is non-negligible since f_+ is large.

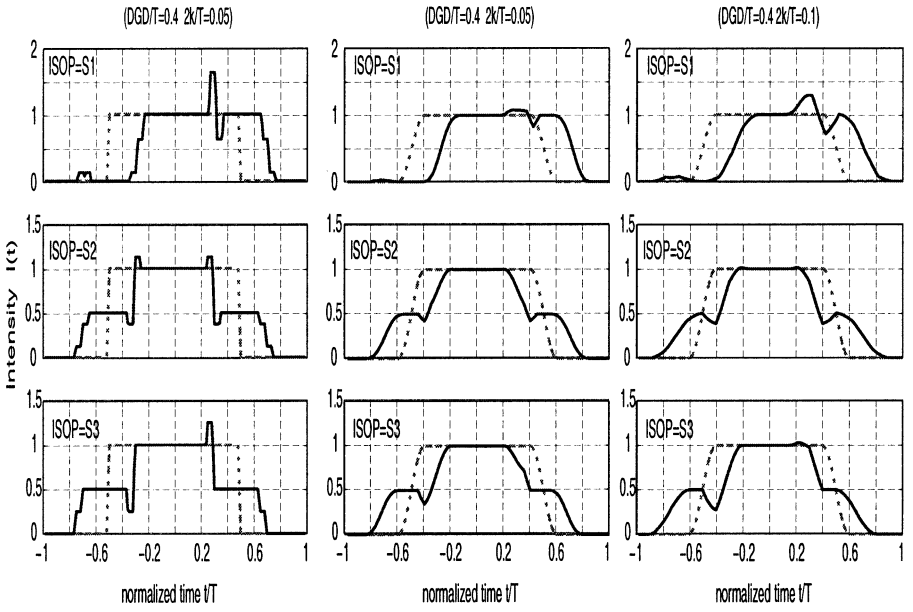

Fig. 5. Same as in Fig. 4, but now $\Delta\tau/2 > 2k_{\mathrm{v}}$.

Let's move to the case $\Delta\tau/2 > 2k_{\mathrm{v}}$, in which DGD dominates over the eigenmode rotation, a typical situation in non-compensated fibers. When $2k_{\mathrm{v}} = 0$ and $\vec{j}$ is aligned with the slow eigenmode we get an undistorted delayed replica of the pulse, as we saw in special Case 1.

Figure 5, leftmost column, shows the received intensity for the same three ISOPs as before.

We see that, when $\hat{j} = \hat{b}_0$, the effect of the rotation $2k_{\mathrm{v}}$ is to create a blip on the flat part of the pulse, and to smear the flanks of the NRZ pulse. The top level of the blip is still 1.625, and the bottom level 0.625. Increasing $2k_{\mathrm{v}}$ further has the effect of enlarging the blip and increasing the smearing of the flanks. We note that *when the ISOP is aligned with the eigenmode $\hat{b}_0$ the pulse in this case is an essentially undistorted, delayed version of the NRZ pulse, while broad pulse smearing is present for the remaining ISOPs, the dominant effect being first-order PMD, i.e., the pulse-splitting effect of the dominant $\Delta\tau$.*

The three plots in the center column in Fig. 5 show the same graphs as those in the leftmost column, but when the input pulse has a raised cosine shape (in power) with roll-off 0.2, where the blips are smoothed out. The three plots in the rightmost column show the effect of doubling $2k_{\mathrm{v}}$, still with roll-off 0.2, and we note that the blip is more marked.

In Fig. 6 we report the four intensity components $N_0, \ldots, N_3$ for the rotation model with $\hat{k} = \hat{s}_3$, $\hat{b}_0 = \hat{s}_1$, $\Delta\tau/T = 0.4$, $2k_{\mathrm{v}}/T = 0.5$, for an OOK modulated input field. The first row reports the output intensity $|\overline{E}_{\mathrm{o}}(t)|^2$ we would get in the scalar filtering case $\mathrm{e}^{-j\bar{\phi}(\omega)}$. The left column shows N_0 through N_3 in the case without GVD ($\bar{\phi}(\omega) = 0$) while the right column gives the values in the presence of GVD, which we quantify in terms of the bit-rate independent factor $\beta_2 L/T^2 = L/L_{\mathrm{D}}$, being L the link length, β_2 the dispersion parameter, and $L_{\mathrm{D}} \triangleq T^2/\beta_2$ the dispersion length [36]. We note the main feature in the absence of GVD, namely that $N_2 = 0$. Such condition implies that the same intensity is obtained when the azimuth of the ISOP changes sign, i.e., when j_2 is changed in $-j_2$.

We note that an even input $E_{\mathrm{i}}(t)$ with GVD gives an even $\overline{E}_{\mathrm{o}}(t)$, so that for an interpretation of the above results we can again use the time symmetry properties of

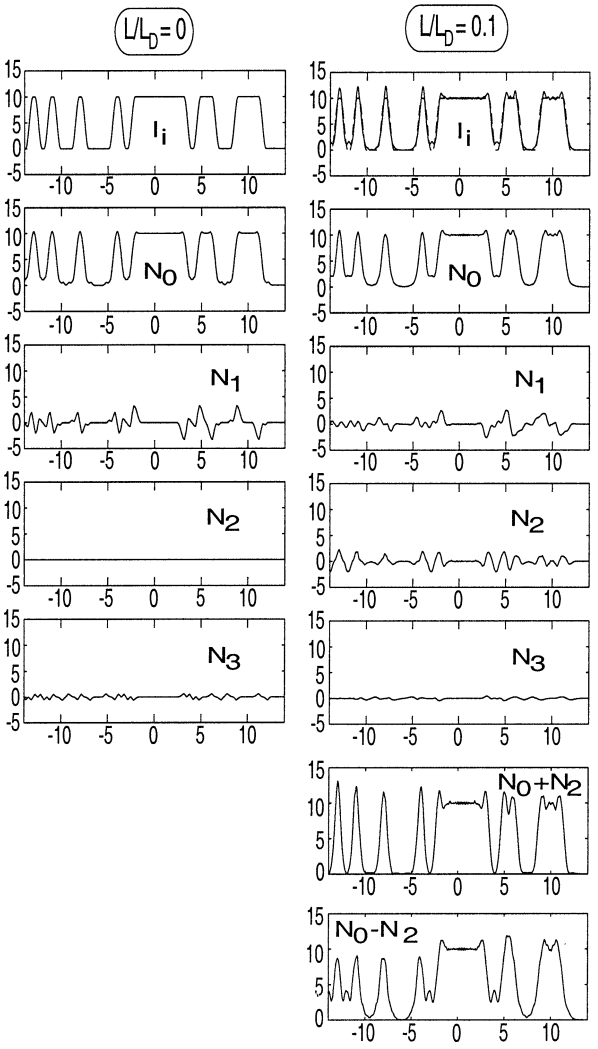

Fig. 6. The four intensity components vs. normalized time, for a rotation model with $\hat{k} = \hat{s}_3$, $\hat{b}_0 = \hat{s}_1$, $\Delta\tau/T = 0.4$, $2k_v/T = 0.5$. The first row reports the intensity $|\overline{E}_o(t)|^2$, while the left column shows N_0 through N_3 in the case without GVD, and the right column for GVD parameter $L/L_D = 0.1$.

$N_0, \ldots, N_3$. For instance, take the 101 pattern. The N_0, N_2 terms contributed by the two marks on the space in the middle tend to add up, since they are even with respect to the position of the marks, while the N_1, N_3 terms tend to cancel, since they are odd. Thus the space level in the 101 sequence is the one determined by the N_0 terms when the ISOP has $j_2 = 0$, while it is mostly enhanced or depressed when $j_2 = \pm 1$ and $j_1 = j_3 = 0$, as evidenced by the bottom two graphs in the right column of Fig. 6. The "critical" ISOPs $\hat{j} = \pm\hat{s}_2$ will be further discussed in Section 5.5.

We remark that all the above results refer to the case $\varphi = \pi/2$, i.e., to a rotation axis at right angles with $\hat{b}_0$, as done in [33,34]. In the general case, we verified that, for given $2k_v$, the pulse distortion is smaller than in the case $\varphi = \pi/2$, where for the same rotation angle the eigenmodes describe a longer arc, and there is thus more polarization dispersion of the spectral components of the signal.

In the next section we will thoroughly explore the ECP performance of an optical transmission over a fiber that follows the rotation model.

5. Eye closure penalty

In this section we derive a simple approximate formula of the ECP applicable to any single-mode fiber whose Jones matrix Pauli vector is known. The formula extends the elegant and simple method presented by Chen [8] to the general case of all-order PMD. Before deriving such generalized Chen's formula (GCF), we first recall Chen's formula for first-order PMD, and link it to other well-known formulas for the ECP available in the literature.

We then test the accuracy of the GCF against simulation results for the fiber rotation model, over a reasonable range of its parameters. In so doing, we test the combined effect of PMD and GVD/chirp of the input pulse.

Finally, we show that the GCF coincides with an ECP formula already known in the literature [24], although the assumptions in the derivation are different.

5.1. Chen's formula for first-order PMD

The method considers only the penalty on the periodic bit sequence ...1010..., which is the one that most contributes to the eye closure penalty in presence of first-order PMD, since long strings of ones and zeros are little affected by DGD.

In absence of PMD, considering that the electrical lowpass filter will pass only the first harmonic of the 1010 sequence, the first harmonic of the received scalar field, taking into account a non-zero extinction ratio, can be expressed as

$$\overline{E}_{\mathrm{o}}(t) = k\big(1 + a\cos(\omega_0 t)\big), \tag{45}$$

where $\omega_0 \triangleq \pi/T$. The photodetected current[3] contribution of this term is [8]

$$I(t) = \left(1 + \frac{1-r}{1+r}\cos(\omega_0 t)\right) P_{\mathrm{avg}},$$

where $P_{\mathrm{avg}} \triangleq k^2(1 + a^2/2)$ and $2a/(1 + a^2/2) = (1 - r)/(1 + r)$, where the extinction ratio r is the ratio of the average power on zeros to the power on ones. In the presence of first-order PMD, the power splits on the two input PSPs (i.e., the frequency-independent right-extracted eigenmodes at $\omega \to 0$), so that the output intensity is

$$I(t) = \left[1 + \frac{1-r}{1+r} Y \cos\big(\omega_0(t-\delta)\big)\right] P_{\mathrm{avg}}, \tag{46}$$

where

$$Y = \sqrt{\cos^2\left(\frac{\pi\Delta\tau}{2T}\right) + \sin^2\left(\frac{\pi\Delta\tau}{2T}\right)[1-2\gamma]^2} = \sqrt{1 - 4\gamma(1-\gamma)\sin^2\left(\frac{\pi\Delta\tau}{2T}\right)} \tag{47}$$

being $\Delta\tau$ the DGD and γ the power splitting factor, and the second form appears in [37,38]. Obviously, the ECP in dB evaluated on the 1010 sequence is: $\mathrm{ECP} \triangleq -10\,\mathrm{Log}\,Y$, while a more complete expression of the Q-factor including Y is given in [8] and allows

[3] We assume for simplicity a responsivity of the photodiode equal to one, so that the field intensity is equal to the photodetected current.

the evaluation of the bit-error rate in the standard Gaussian approximation in the presence of amplified spontaneous emission noise. When $\gamma = 0.5$, we have the worst case penalty

$$\mathrm{ECP} = -10\,\mathrm{Log}\left|\cos\left(\frac{\pi\Delta\tau}{2T}\right)\right|. \tag{48}$$

Moreover, for $\pi\Delta\tau/2T \ll 1$, we can linearize to get: $\mathrm{ECP} \cong A(\Delta\tau/T)^2\gamma(1-\gamma)$, with $A \cong 21.4$ which is quite close to the well-known approximation based on the pulse broadening [7]. It is shown in [39, p. 136] that indeed the coefficient A varies with pulse shape and receiver characteristics.

5.2. Generalized Chen's formula for all-order PMD

We generalize here Chen's formula to all-order PMD. We start from the very general expression (13) of the output intensity, which is based on the Pauli vector $\underline{u}(\omega)$ of $U(\omega)$, whose entries are made explicit in (10). Assume, as in Chen's work, that the common-phase distorted field in formula (13) is the *real* sinusoidal term given in (45). Then

$$u_0 \otimes \overline{E}_\mathrm{o} = k\left[\cos\frac{\Delta\phi(0)}{2} + \frac{a}{2}\,\mathrm{e}^{j\omega_\mathrm{o}t}\cos\frac{\Delta\phi(\omega_0)}{2} + \frac{a}{2}\,\mathrm{e}^{-j\omega_\mathrm{o}t}\cos\frac{\Delta\phi(-\omega_0)}{2}\right] \quad \text{and}$$

$$u_k \otimes \overline{E}_\mathrm{o} = -jk\left[\sin\frac{\Delta\phi(0)}{2}b_k(0) + \frac{a}{2}\,\mathrm{e}^{j\omega_\mathrm{o}t}\sin\frac{\Delta\phi(\omega_0)}{2}b_k(\omega_0) + \frac{a}{2}\,\mathrm{e}^{-j\omega_\mathrm{o}t}\sin\frac{\Delta\phi(-\omega_0)}{2}b_k(-\omega_0)\right], \quad \text{for } k = 1,\ldots,3.$$

As already noted, expression (13) also holds when instead of the Pauli vector $\underline{u}(\omega)$ of $U(\omega)$ we use that $\underline{u}_\mathrm{r}(\omega)$ of $U_\mathrm{r}(\omega)$, since $U^\dagger(\omega_1)U(\omega_2) = U_\mathrm{r}^\dagger(\omega_1)U_\mathrm{r}(\omega_2)$. However, the above terms greatly simplify when using U_r instead of U because of the property $\Delta\phi_\mathrm{r}(0) = 0$. Hence from now on we will work with U_r, although we will omit the subscript for simplicity of notation. Substituting the above terms in (16), (17), and *ignoring the "double-frequency" terms at* $2\omega_0$, we get for the low-frequency (lf) components

$$\left\{N_0(t)\right\}_\mathrm{lf} = k^2\left[1 + \frac{a^2}{2} + a\left(\cos\frac{\Delta\phi(\omega_0)}{2} + \cos\frac{\Delta\phi(-\omega_0)}{2}\right)\cos(\omega_0 t)\right],$$

$$\left\{\vec{N}(t)\right\}_\mathrm{lf} = k^2 a\left(\sin\frac{\Delta\phi(\omega_0)}{2}\hat{b}(\omega_0) - \sin\frac{\Delta\phi(-\omega_0)}{2}\hat{b}(-\omega_0)\right)\sin(\omega_0 t),$$

since the vector product $(\vec{u}\otimes\overline{E}_\mathrm{o}) \times (\vec{u}\otimes\overline{E}_\mathrm{o})^*$ in (17) has only double-frequency components when $\Delta\phi(0) = 0$. Finally, from (13) we get the low-pass filtered photodetected current as in (46), where

$$Y = \left(\left[\frac{\cos(\Delta\phi(\omega_0)/2) + \cos(\Delta\phi(-\omega_0)/2)}{2}\right]^2 + \left[\hat{\jmath}\cdot\frac{\sin(\Delta\phi(\omega_0)/2)\hat{b}(\omega_0) - \sin(\Delta\phi(-\omega_0)/2)\hat{b}(-\omega_0)}{2}\right]^2\right)^{1/2} \tag{49}$$

gives the eye reduction on the 1010 sequence, and is thus the sought generalization of Chen's formula. By defining the four-dimensional vector $\underline{u}_\mathrm{R}(\omega) \triangleq (\underline{u}(\omega) + \underline{u}^*(-\omega))/2$,

representing the Fourier transform of the real part $\underline{u}_{\mathrm{R}}(t)$ of the time-impulse response vector $\underline{u}(t) = \underline{u}_{\mathrm{R}}(t) + j\underline{u}_{\mathrm{I}}(t)$, the GCF formula can be written very compactly as

$$Y(\omega_0) = \left|\underline{u}_{\mathrm{R}}(\omega_0) \cdot \underline{j}\right|. \tag{50}$$

5.3. Sanity checks

5.3.1. First-order model

Assume a first-order model (B.2) for the right-extracted U_{r}. Then from (49) we get

$$Y = \sqrt{\cos^2\left(\frac{\pi\Delta\tau}{2T}\right) + \sin^2\left(\frac{\pi\Delta\tau}{2T}\right)\left(\hat{b}\cdot\hat{j}\right)^2}. \tag{51}$$

If 2θ is the angle between $\hat{b}$ and the ISOP $\hat{j}$, from [27] we know that

$$\gamma \triangleq \left|\widehat{\mathbf{B}}_{\mathrm{s}}^{\dagger}\hat{\mathbf{J}}\right|^2 = \frac{1+(\hat{b}\cdot\hat{j})}{2} = \frac{1+\cos 2\theta}{2} = \cos^2\theta,$$

hence $(\hat{b}\cdot\hat{j})^2 = (2\gamma-1)^2$ and thus the formula for Y checks with Chen's (47).

5.3.2. Rotation model

We specialize here the penalty formula to the rotation model described in Section 4.

From (38), and assuming a linear phase $\Delta\phi = \omega\Delta\tau$ we get the following expression for Y

$$Y = \sqrt{\cos^2\left(\frac{\Delta\tau\omega_0}{2}\right) + \sin^2\left(\frac{\Delta\tau\omega_0}{2}\right)\left[\vec{a}\cdot\hat{j}\right]^2}, \tag{52}$$

where, by defining $\cos\varphi \triangleq (\hat{k}\cdot\hat{b}_0)$, we built the vector

$$\vec{a} \triangleq \frac{\hat{b}(\omega_0)+\hat{b}(-\omega_0)}{2} = \left(\cos(2k_{\mathrm{v}}\omega_0)\hat{b}_0 + \left(1-\cos(2k_{\mathrm{v}}\omega_0)\right)\cos\varphi\hat{k}\right), \tag{53}$$

which is the average of the eigenmodes at $\pm\omega_0$, whose squared magnitude is $a^2 = \sin^2\varphi\cos^2(2k_{\mathrm{v}}\omega_0) + \cos^2\varphi$, and whose direction $\hat{a}$ represents the best ISOP, the one that gives the *minimum* penalty

$$\mathrm{ECP}_{\mathrm{m}} = -5\,\mathrm{Log}\left[\cos^2(\Delta\tau\omega_0/2) + \sin^2(\Delta\tau\omega_0/2)a^2\right].$$

Conversely, the worst ISOPs are those lying on the great circle orthogonal to $\hat{a}$, resulting in a worst-case penalty coinciding with expression (48). This means that, *as far as the* 1010 *sequence is concerned*, a rotation model for the extracted matrix has the worst-case penalty coinciding with that of a first-order PMD model.

The locus of ISOPs that have the same ECP is composed of the two circles on the Poincaré sphere centered at $\hat{a}$, obtained by the intersection of the sphere with the cone having axis $\hat{a}$ and given aperture, as shown in Fig. 7. Hence we can think that the ECP maps on the Poincaré sphere of ISOPs in circles centered at the point $\hat{a}$, with the ECP increasing from the (degenerate) circle at $\hat{a}$, up to a maximum value on the great circle orthogonal to it.

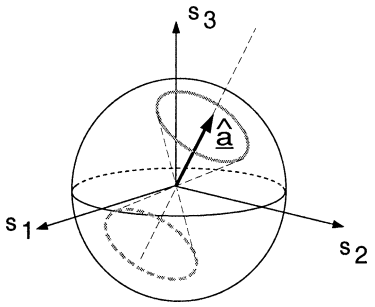

Fig. 7. Locus of ISOPs on Poincaré sphere having the same ECP, according to (52).

We can also find the penalty in the case of PSP transmission [35], i.e., one in which the ISOP is aligned with the right-extracted eigenmode at the central frequency: $\hat{j} = \hat{b}_0$. We find in this case

$$\mathrm{ECP} = -5\,\mathrm{Log}\left[\cos^2\left(\frac{\Delta\tau\omega_0}{2}\right) + \sin^2\left(\frac{\Delta\tau\omega_0}{2}\right)\left[2\sin^2(k_\mathrm{v}\omega_0)\cos^2\varphi + \cos(2k_\mathrm{v}\omega_0)\right]^2\right].$$

The best ISOP $\hat{a}$ is in some sense the point closer to the trajectory $\hat{b}(\omega)$ in the range $[-\omega_0, \omega_0]$. It coincides with $\hat{b}_0$ only when there is no rotation, $k_\mathrm{v} = 0$. When the rotation on the above range is a half circle ($2k_\mathrm{v}\omega_0 = \pi/2$), the best ISOP is aligned with the rotation axis $\hat{k}$, as per (53). When the rotation is an entire circle, the best ISOP is diametrically opposed to $\hat{b}_0$ on the trajectory.

5.4. Numerical verifications

We tested our analytical GCF formula against the exact ECP obtained by propagating a pseudo-random binary sequence (PRBS) of $2^7 - 1$ bits through a fiber synthesized by the rotation model of its right-extracted matrix, with the rotation axis $\hat{k}$ aligned with $\hat{s}_3$. The supporting mark pulses are raised cosine in power, with roll-off 0.8. The extinction ratio is infinite. The received intensity is lowpass filtered with a fourth-order Bessel filter of bandwidth 0.65 the bit-rate.

The analytical GCF is based on the assumption of a purely sinusoidal signal (at the "1010" frequency equal to half the bit rate, e.g., 5 GHz for a 10 Gb/s signal) being transmitted along the fiber. Hence we expect that, whenever the "1010" sequence is the main eye-closing sequence, the GCF gives results close to the correct ECP. Since the exact pulse shape is not taken into account, we expect discrepancies due to such effect.

Initially, we consider an equatorial circle for the extracted eigenmodes, by setting $\hat{b}_0 = \hat{s}_1$, so that $\varphi = \pi/2$ and no common-mode GVD is present.

For a set of given values of relative DGD $\Delta\tau/T$, we swept the values of the (ω-independent) relative angular rotation speed $2k_\mathrm{v}/T$, and recorded the ECP after the lowpass electrical filter. The solid curves in Fig. 8 show the ECP versus $2k_\mathrm{v}/T$ for fixed $\Delta\tau/T = 0.05, 0.1, 0.2, 0.4$, when the ISOP is $\hat{j} = \hat{b}_0$. This corresponds to PSP transmission [35], for which we already observed the maximal pulse deformation caused by the eigenmodes rotation in Section 4. We therefore expect the maximal discrepancy between the analytical ECP and the simulated one. Similar ECP curves were already reported in [34]. It is interesting to observe the oscillatory behavior of the ECP, with

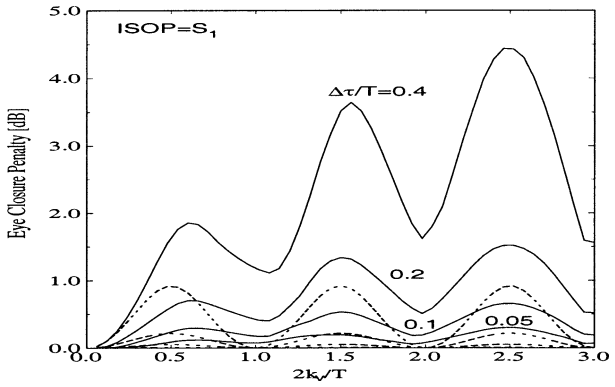

Fig. 8. ECP versus normalized angular rotation speed in rotation model ($\hat{k} = \hat{s}_3, \hat{b}_0 = \hat{s}_1$), for normalized DGD as a parameter (from top to bottom: $\Delta\tau/T = 0.4, 0.2, 0.1, 0.05$). Solid: simulated PRBS; dashed: theoretical GCF. ISOP $\hat{j} = \hat{b}_0$.

a first local maximum around $2k_v/T \cong 0.6$. The local maxima are repeated at intervals of 1 from each other. Note that, since $2k_v\omega_0 = (2k_v/T)\pi$ is the rotation angle between the eigenmode at the reference frequency and the eigenmode at frequency ω_0, a value $2k_v/T = 1$ corresponds to a very large rotation of $\pm\pi$ of the eigenmode at frequency $\pm\omega_0$ with respect to $\hat{b}_0$, i.e., a complete circle is described by the extracted eigenmodes over the bandwidth $[-\omega_0, \omega_0]$. The shown range of $2k_v/T$ is therefore unrealistically large, and in practice it makes sense to observe the ECP only in the range $2k_v/T < 1$. Moreover, care must be taken in reading out the results of the figure, since it has been shown that in standard fibers there is a strong statistical correlation between DGD and PSP rotation speed, so that large rotation speed values correspond usually to very low DGD values and vice versa [17,40].

In the same figure, we reported in dashed lines, and in dB, the analytical GCF (49), which here specializes to the rotation model case (52). We observe a good numerical match only for low values of the rotation speed $2k_v/T < 0.25$. The analytical ECP captures the oscillatory behavior of the ECP, with almost the correct location of maxima and minima, but the actual values of ECP are quite off for large rotation speed $2k_v/T$. In particular, the analytical ECP is bounded by the limiting formula (48), which is also the worst case of the first order model.

The next question is how such ECP curves change when we use the orthogonal ISOP $\hat{j} = -\hat{b}_0$. It can be proven that if we use a time-reversed input $E_i(-t)$ and a reversed ISOP $\hat{j} = -\hat{b}_0$ we obtain a time-reversed output intensity $I(-t)$, and therefore, if the PRBS is sufficiently long, we should obtain the same ECP curves as in Fig. 8. The simulated ECP however slightly differs from the $\hat{j} = \hat{b}_0$ case, by less than half a dB. The reason is that the impulse response of the Bessel filter is not symmetric in time. Thus even with very long PRBS sequences we should not expect exactly the same ECP when reversing the ISOP.

Let's continue to investigate specific ISOPs. Figure 9 shows the smooth transition of the simulated ECP-vs.-$2k_v$ curves when the ISOP is rotated on the equator, starting from $\hat{j} = \hat{s}_2$ up to $\hat{j} = \hat{s}_1$ with azimuth decreasing from 90° to 0° in steps of 15°, as shown in the inset. We observe that $\hat{j} = \hat{s}_2$ gives the smallest penalty when the rotation speed exceeds the critical value $2k_v/T \cong 0.33$ *through which—very surprisingly—essentially all ECP curves pass*. Very similar ECP curves are obtained when rotating the ISOP on the $(\hat{s}_1, \hat{s}_3)$ plane starting from $\hat{s}_3$ and ending on $\hat{s}_1$. Hence we conclude that *for rotation speeds* $2k_v/T < 0.33$ *the best ISOP is the one aligned with* $\pm\hat{b}_0$ (*PSP transmission*), *while*

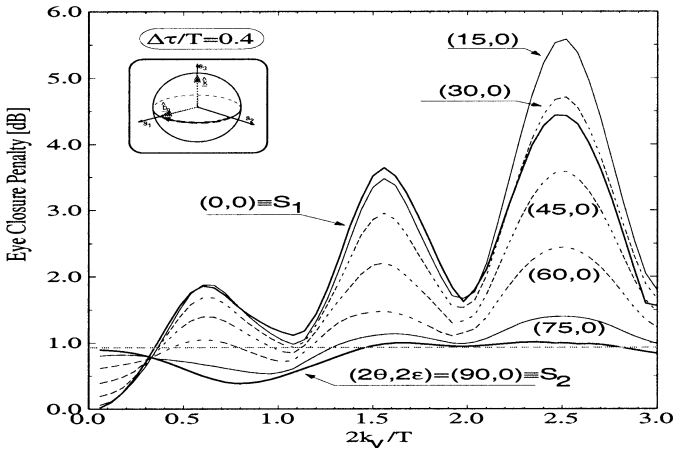

Fig. 9. Simulated ECP vs. angular rotation speed, for rotation model ($\hat{k} = \hat{s}_3$, $\hat{b}_0 = \hat{s}_1$) with fixed $\Delta\tau/T = 0.4$, when the ISOP slides on the equator, moving from $\hat{j} = \hat{s}_2$ to $\hat{j} = \hat{s}_1$ with azimuth decreasing from 90° to 0° in steps of 15°, as shown in the inset.

for rotation speeds exceeding 0.33 *the best ISOPs are those orthogonal* (*on the Poincaré sphere*) *to* $\hat{b}_0$, *i.e., those aligned with* $\pm\vec{b}'(0)$ [17]. Finally, the horizontal dashed line is the theoretical ECP predicted by the GCF (52) for the case $\hat{j} = \hat{s}_2$. We note that the GCF slightly overestimates the ECP for increasing rotation speed up to $2k_v/T \cong 0.6$, but overall the estimation error is much smaller than the case $\hat{j} = \hat{s}_1$.

The above figures give an idea of how the ECP varies with the rotation model parameters for some selected ISOPs. We tested more thoroughly how the simulated ECP varies as a function of the azimuth θ and ellipticity ε of the ISOP in the graphs in Fig. 10, where the shown range of (θ, ε) spans the whole Poincaré sphere. The figure refers to the simulated ECP with PRBS input sequence, for a rotation model with $\hat{k} = \hat{s}_3$, $\hat{b}_0 = \hat{s}_1$ ($\varphi = \pi/2$), with fixed $\Delta\tau/T = 0.4$ and rotation speed increasing from top to bottom graphs. First concentrate on the graphs in the left column. In the top left graph the rotation speed is small. We note one central point of minimum ECP, corresponding to $\hat{j} = \hat{b}_0$, and two side minima corresponding to $\hat{j} = -\hat{b}_0$. The great circle orthogonal to $\hat{b}_0$ maps into the two top rails on the surface at $\theta = \pm\pi/4$, which show maximal ECP, as predicted by the GCF (52). Note also the perfect symmetry in θ, which comes from the fact that $N_2 = 0$ in the absence of GVD, as already noted in (44), and the approximate symmetry in ε, due to the fact that $N_3 \cong 0$ in this case, as already noted in Section 4.1. As the rotation speed increases, the "good" ISOPs $\hat{j} = \pm\hat{b}_0$ deteriorate, while the "bad" ISOPs ($\theta = \pm\pi/4$) become those with lowest ECP, with an intermediate case at $2k_v/T \cong 0.33$ with a global flattening of the ECP surface, which we partly already observed in Fig. 9.

One of the most interesting results coming from the theoretical GCF (52) in the rotation model is the mapping of the ECP on the Poincaré sphere of ISOPs, which shows circular symmetries around the axis determined by the optimal ISOP $\hat{a}$, as already noted in Fig. 7. Clearly, the theoretical GCF only considers what happens to a purely sinusoidal input power. In Fig. 10, *right column*, we tested how such symmetries carry over when a PRBS sequence is applied to the input. We rotated the axes, so that now $\hat{a}$ is aligned with $\hat{s}_3$ and $\hat{k} = \hat{s}_1$, and thus all circles around $\hat{a}$ are the loci of constant elevation on the new (θ, ε) plane. Note that in our case ($\varphi = \pi/2$) we have from (53): $\hat{a} \equiv \hat{b}_0$ for all rotation speeds $2k_v$. We see from the top right plot ($2k_v/T = 0.1$) that indeed the $\wedge$-shaped surface reflects such circular symmetry predicted by (52), and also the ECP values are in agreement with the GCF. As we increase the rotation speed to the special value $2k_v/T = 0.33$, we

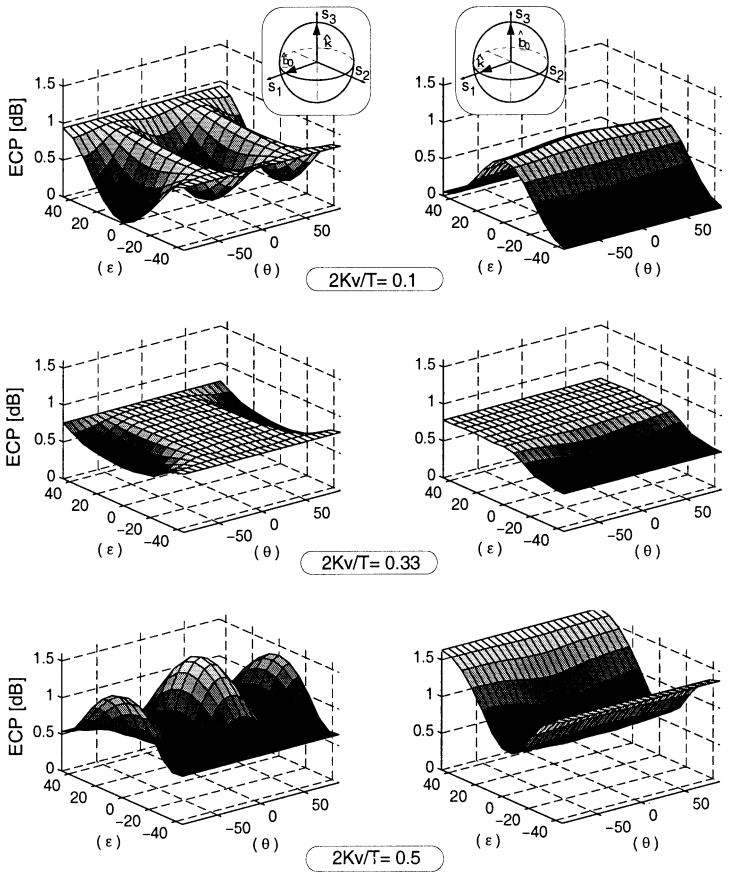

Fig. 10. ECP vs. ISOP azimuth θ and elevation ε for PRBS input sequence. Left column: rotation model ($\hat{k} = \hat{s}_3$, $\hat{b}_0 = \hat{s}_1$) with fixed $\Delta\tau/T = 0.4$, increasing $2k_v/T$ top to bottom. Right column: same model, with axes reoriented so that $\hat{a} \equiv \hat{b}_0 = \hat{s}_3$, $\hat{k} = \hat{s}_1$.

note the ECP "tent" surface flattens out, again indicating that there is circular symmetry around $\hat{a}$, although the numerical ECP values predicted by the theory (52) start to be less accurate. There is a small breakup of the symmetry between the hemisphere centered at $\hat{a}$ and that at $-\hat{a}$ due to the asymmetric electrical filter response, as previously discussed. As we further increase $2k_v$ we see that the $\vee$-shaped ECP surface preserves the circular symmetry around $\hat{a}$, although the numerical predictions of the GCF are completely off.

The important message to get is that, *in the case of $\hat{k}$ at right angles with $\hat{b}_0$, although the numerical predictions of theoretical ECP* (52) *are correct only at small rotation speed, its prediction of circular symmetry around $\hat{a}$ holds surprisingly even at large rotation speed.*

We finally tested the effect of changing the aperture angle φ between $\hat{b}_0$ and $\hat{k}$. In the case of a small aperture $\varphi < 10°$, and by symmetry $> 180° - 10°$, the surface remains close to the $\wedge$-shaped "tent" for all rotation speeds, since the system essentially behaves like a first-order model, which also displays the same "tent" surface. We can intuitively say that in this case all ISOPs $\hat{j}$ on the sphere "see" an essentially fixed eigenmode $\hat{b}_r(\omega)$, except possibly for those ISOPs very close to the trajectory $\hat{b}_r(\omega)$. We verified that the ECP *symmetries around $\hat{a}$ substantially diminish* for increasing rotation speed *when $\hat{b}_0$ is not orthogonal to $\hat{k}$*. Such asymmetries are most evident when $\varphi = 45°$ or $\varphi = 180° - 45°$, and in this case we observed that the symmetry axis for large rotation speed becomes $\hat{b}_0$, *a* fact that is *not* accounted for by the theoretical axis formula (53).

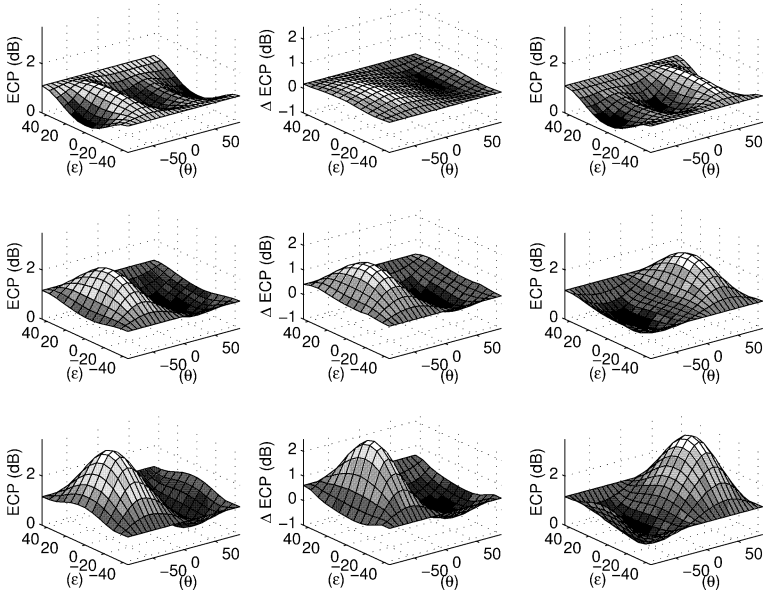

Fig. 11. Rotation model ($\hat{k} = \hat{s}_3$, $\hat{b}_0 = \hat{s}_1$, $\Delta\tau/T = 0.4$) at increasing rotation speed (first row: $2k_v/T = 0.1$; second row: $2k_v/T = 0.33$; third row: $2k_v/T = 0.5$) and $L/L_D = 0.1$ (left column); $L/L_D = -0.1$ (right column). Center column: difference of ECP (dB) between $L/L_D = 0.1$ and $L/L_D = 0$.

5.5. *Effect of GVD*

We now concentrate on the effect of the chirp induced by common-mode GVD on the ECP surfaces. We start with a first-order model, one with zero rotation speed. We numerically verify that in the absence of rotation, GVD preserves the ECP shape, the only effect being an upward shift of the whole ECP surface.

The symmetry in θ is broken more and more by GVD as the rotation speed is increased. The simulated ECP surfaces in Fig. 11 show the case of a rotation model with $\hat{k} = \hat{s}_3$, $\hat{b}_0 = \hat{s}_1$, $\Delta\tau/T = 0.4$ for increasing rotation speed (first row: $2k_v/T = 0.1$; second row: $2k_v/T = 0.33$; third row: $2k_v/T = 0.5$). The first column gives the ECP surfaces for GVD $L/L_D = 0.1$, while the second column gives the ECP difference with respect to the case without GVD. We note that the asymmetry in θ increases with the rotation speed. We also note the presence of a particular ISOP, aligned with $\vec{b}'(0)$, which we denote as $\hat{j} = \hat{b}'_0$ ($= \hat{s}_2$ in this case) for which the penalty (as seen in the central column) is actually reduced with respect to the case without GVD: in this case we have a fruitful interaction between PMD and GVD which opens the eye. For instance, at large rotation speed there is a small region around $\hat{b}'_0$ at which the ECP is negative. While the best ISOP is $\hat{b}'_0$, the worst ISOP is $-\hat{b}'_0$. A plot of the time waveform of the received intensity for both the best and the worst ISOPs with GVD was already shown in Fig. 6, bottom two graphs on the right column, where the most evident feature is the strong closure of the zeros in the 1010 sequences for the worst ISOP, and the "cleaning" of the zeros for the same 1010 sequences for the best ISOP, with respect to the case of zero GVD.

The third column in Fig. 11 shows the ECP in the case $L/L_D = -0.1$. We note that inverting the sign of GVD amounts to a mirror image of the ECP surface with respect to the center line, which corresponds to the great circle in the plane $(\hat{k}, \hat{b}_0)$, which is thus the locus of ISOPs which are unaffected by a change in sign of GVD. The locus of maximal ECP variation is instead the equatorial circle, the one on which the eigenmodes rotate, where the best and worst ISOPs ($\hat{b}'_0$ or $-\hat{b}'_0$ according to the sign of dispersion) lie. From

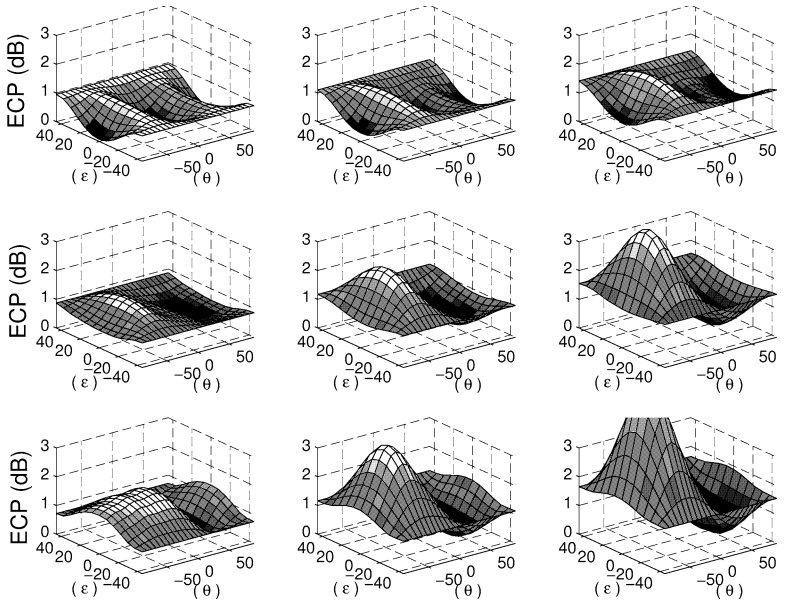

Fig. 12. Rotation model ($\hat{k} = \hat{s}_3$, $\hat{b}_0 = \hat{s}_1$, $\Delta\tau/T = 0.4$) at increasing rotation speed (first row: $2k_v/T = 0.1$; second row: $2k_v/T = 0.33$; third row: $2k_v/T = 0.5$) and (left column) $L/L_D = 0.05$; (center column) $L/L_D = 0.1$; (right column) $L/L_D = 0.1375$.

the ECP surfaces we note that the best ISOP moves from $\hat{b}_0$ at small rotation speed towards $\pm\hat{b}'_0$ as the rotation speed increases. The ISOP domains of best performance centered at the best ISOP, which we here define as the ones for which the ECP remains below 0.5 dBs, is fairly large: $\pm 10°$ in θ and $\pm 30°$ in ε at a large $2k_v/T = 0.5$ rotation speed, and for GVD $L/L_D = 0.1$.

For the same rotation model, Fig. 12 highlights the effect of increasing the amount of GVD from the left to the right column on the shape of the ECP surfaces. In particular we note the change in shape of the ISOP domains of best performance. We see that, while at low rotation speed (first row) such best ISOP domains first shrink and then disappear as GVD increases and the ECP surface shifts upwards, at larger rotation speed they remain almost unchanged for increasing GVD.

5.6. Relation with small-signal baseband frequency response

Noé et al. [24] studied the frequency response of the small-signal IM/IM conversion in a fiber with PMD. If the input field is a CW with a small amplitude modulation, the photodetected intensity modulation at the receiver is linearly related to the amplitude modulation at the transmitter through a (scalar) baseband filter

$$H_{IM}(\omega) = \frac{1}{2}\hat{\mathbf{J}}^{\dagger}\left(e^{-j\delta\bar{\phi}(\omega)}U_0^{\dagger}U(\omega) + e^{j\delta\bar{\phi}(-\omega)}U^{\dagger}(-\omega)U_0\right)\hat{\mathbf{J}}, \tag{54}$$

where as usual $\hat{\mathbf{J}}$ is the ISOP associated with the four-dimensional Stokes vector $\underline{j} = [1; \hat{j}]$, and $\delta\bar{\phi}(\omega) \triangleq \bar{\phi}(\omega) - \bar{\phi}(0)$, being the fiber matrix $T(\omega) = e^{-j\bar{\phi}(\omega)}U(\omega)$. We want to prove here such result, which is given in [24] neglecting the GVD term $\delta\bar{\phi}(\omega)$, and strengthen it by showing that

$$H_{IM}(\omega) = \underline{w}_R(\omega) \cdot \underline{j}, \tag{55}$$

where

$$\underline{w}(\omega) \triangleq \mathrm{e}^{-j\delta\bar{\phi}(\omega)}\underline{u}_{\mathrm{r}}(\omega);$$
$$\underline{w}_{\mathrm{R}}(\omega) \triangleq \frac{\underline{w}(\omega)+\underline{w}^*(-\omega)}{2}, \qquad \underline{w}_{\mathrm{I}}(\omega) \triangleq \frac{\underline{w}(\omega)-\underline{w}^*(-\omega)}{2i} \tag{56}$$

being $\underline{u}_{\mathrm{r}}$ the Pauli vector of the right-extracted matrix U_{r}. Hence, $H_{\mathrm{IM}}(\omega_0)$ is the filter whose magnitude is our $Y(\omega_0)$ function (50).

The proof goes as follows.

Proof. Let the input field be $\mathbf{E}_{\mathrm{i}}(t) = E_{\mathrm{i}}(t)\hat{\mathbf{J}}$, with $E_{\mathrm{i}}(t) = \sqrt{P}(1 + (a/2)\cos\omega_0 t)$ and $|a| \ll 1$. Thus the input intensity is $I_{\mathrm{i}}(t) \cong P(1 + a\cos\omega_0 t) = P(1 + (a/2)(\mathrm{e}^{j\omega_0 t} + \mathrm{e}^{-j\omega_0 t}))$. The output field in the frequency domain is

$$\begin{aligned}\mathbf{E}_{\mathrm{o}}(\omega) &= E_{\mathrm{i}}(\omega)\,\mathrm{e}^{-j\bar{\phi}(\omega)}U(\omega)\hat{\mathbf{J}}\\ &= \sqrt{P}\Big[\delta(\omega)U_0\,\mathrm{e}^{-j\bar{\phi}(0)}\hat{\mathbf{J}} + \frac{a}{4}\big(\delta(\omega-\omega_0)\,\mathrm{e}^{-j\bar{\phi}(\omega_0)}U(\omega_0)\\ &\qquad + \delta(\omega+\omega_0)\,\mathrm{e}^{-j\bar{\phi}(-\omega_0)}U(-\omega_0)\big)\hat{\mathbf{J}}\Big].\end{aligned}$$

Taking the inverse Fourier transform, we build the output intensity $I_{\mathrm{o}}(t) = \mathbf{E}_{\mathrm{o}}^{\dagger}(t)\,\mathbf{E}_{\mathrm{o}}(t)$. Neglecting the small terms in a^2 we get

$$I_{\mathrm{o}} = P\left\{1 + \frac{a}{2}\hat{\mathbf{J}}^{\dagger}\left(\frac{\mathrm{e}^{-j\delta\bar{\phi}(\omega_0)}U_0^{\dagger}U(\omega_0) + \mathrm{e}^{j\delta\bar{\phi}(-\omega_0)}U^{\dagger}(-\omega_0)U_0}{2}\,\mathrm{e}^{j\omega_0 t} + \mathrm{tc}\right)\hat{\mathbf{J}}\right\},$$

where tc is shorthand for the transpose conjugate of the first term in the bracket. In other terms, the response to the input intensity modulation signal $\mathrm{e}^{j\omega_0 t}$ is $H_{\mathrm{IM}}(\omega_0)\,\mathrm{e}^{j\omega_0 t}$ (including the case $\omega_0 = 0$ for which $H_{\mathrm{IM}}(0) = 1$), where the filter expression is given in (54), thus confirming its interpretation as the small-signal IM/IM frequency response.

Now we move to the second part of the proof.

Since for any ω_1, ω_2 it is: $U^{\dagger}(\omega_1)U(\omega_2) = U_{\mathrm{r}}^{\dagger}(\omega_1)U_{\mathrm{r}}(\omega_2)$, we can work with the right extracted matrix instead of U. Hence

$$H_{\mathrm{IM}}(\omega) = \tfrac{1}{2}\hat{\mathbf{J}}^{\dagger}\big(\mathrm{e}^{-j\delta\bar{\phi}(\omega)}U_{\mathrm{r}}(\omega) + \mathrm{e}^{j\delta\bar{\phi}(-\omega)}U_{\mathrm{r}}^{\dagger}(-\omega)\big)\hat{\mathbf{J}} \tag{57}$$

and decomposing U_{r} in its Pauli vector u_{r} we immediately get the sought result (55). □

The fact that $|H_{\mathrm{IM}}(\pi/T)|$ in (55) coincides with our GCF formula (50) (which we could have extended to include GVD) should not surprise. In the GCF, the field modulation depth a can be large, but in the output intensity expression we eliminate the double frequency terms, proportional to a^2, invoking the action of the low-pass electrical filter. Here instead the double frequency terms are neglected on the assumption of small a. The justifications being different, we reach the same result.

However, this last point of view is more helpful in thinking what happens if the input intensity is expanded in Fourier series, i.e., is a superposition of sinusoidal modulating terms. According to the small-signal model, the output is just the superposition of the individual sinusoidal intensity inputs. A similar extension in the GCF does not lead to any insight.

Application to a fiber with linear retardation of the extracted matrix: $\Delta\phi_{\rm r}(\omega) = \Delta\tau\omega$ gives

$$H_{\rm IM}(\omega) = \cos\left(\frac{\Delta\tau}{2}\omega\right)\cos\big(\delta\bar{\phi}(\omega)\big) - j\sin\left(\frac{\Delta\tau}{2}\omega\right)\vec{a}(\omega)\cdot\hat{j},$$

where $\vec{a} \triangleq [e^{-j\delta\bar{\phi}(\omega_0)}\hat{b}_{\rm r}(\omega) + e^{j\delta\bar{\phi}(-\omega)}\hat{b}_{\rm r}(-\omega)]/2$ is in general a complex vector. For the rotation model, the eigenmode $\hat{b}_{\rm r}$ is the sum of an even and an odd real terms: $\hat{b}_{\rm r} = \vec{b}_{\rm e} + \vec{b}_{\rm o}$ where

$$\vec{b}_{\rm e}(\omega) = \cos(2k_{\rm v}\omega)\hat{b}_0 + \big(1 - \cos(2k_{\rm v}\omega)\big)\big(\hat{k}\cdot\hat{b}_0\big)\hat{k},$$
$$\vec{b}_{\rm o}(\omega) = \sin(2k_{\rm v}\omega)\big[\hat{k}\times\hat{b}_0\big];$$

so that, assuming $\delta\bar{\phi}(\omega)$ is even, we can write $\vec{a} = \cos(\delta\bar{\phi}(\omega))\vec{b}_{\rm e}(\omega) - j\sin(\delta\bar{\phi}(\omega))\vec{b}_{\rm o}(\omega)$ and thus

$$H_{\rm IM}(\omega) = \left[\cos\left(\frac{\Delta\tau}{2}\omega\right)\cos\big(\delta\bar{\phi}(\omega)\big) - \sin\left(\frac{\Delta\tau}{2}\omega\right)\sin\big(\delta\bar{\phi}(\omega)\big)\big(\vec{b}_{\rm o}(\omega)\cdot\hat{j}\big)\right] - j\left[\sin\left(\frac{\Delta\tau}{2}\omega\right)\cos\big(\delta\bar{\phi}(\omega)\big)\big(\vec{b}_{\rm e}(\omega)\cdot\hat{j}\big)\right] \tag{58}$$

is the explicit form of the IM/IM filter for the rotation model.

In the absence of GVD ($\delta\bar{\phi}(\omega) = 0$), the above becomes

$$H_{\rm IM}(\omega) = \cos\left(\frac{\Delta\tau\omega}{2}\right) - \frac{j}{2}\sin\left(\frac{\Delta\tau\omega}{2}\right)\big(\hat{b}(\omega) + \hat{b}(-\omega)\big)\cdot\hat{j},$$

and for small ω we can approximate the extracted eigenmode as: $\hat{b}(\omega) \cong \hat{b}_0 + \vec{b}_0'\omega$ and $\hat{b}(-\omega) \cong \hat{b}_0 - \vec{b}_0'\omega$. Linearizing also the sin and cos terms up to second order, and considering that $\hat{b}_0 = \hat{q}_{\rm i}$ and $\vec{\Omega}_{\rm i} = \Delta\tau\hat{q}_{\rm i}$, we get a well-known approximation of the filter in (58)

$$H_{\rm IM}(\omega) \cong \left[1 - \frac{\Delta\tau^2}{8}\omega^2 - j\omega\frac{1}{2}\big(\vec{\Omega}_{\rm i}\cdot\hat{j}\big)\right]$$

reported in [24]. Note the obvious relationship of such filter with the expression in (25) obtained from the approximate intensity expression that uses the PSPs at the reference frequency.

Figure 13 shows the GCF formula $\text{ECP} = -10\,\text{Log}\,|H_{\rm IM}(\omega_0)|$, with $H_{\rm IM}(\omega)$ calculated from (58), for the same rotation model whose ECP was obtained by full numerical simulation in Fig. 12. Comparison of the two sets of figures reveals that the GCF formula is quite effective in predicting the ECP surface in the presence of GVD. The ISOP domains of best performance are well reproduced by the GCF, although the actual ECP values are *overestimated* for most ISOPs. The satisfactory match between simulation and theory confirms that in the presence of GVD/chirping of the input pulse *the* 1010 *sequences become again the main cause of eye degradation, hence of ECP penalty.*

Figure 14 shows the same ECP surfaces as above, when the eigenmode trajectory in the rotation model has an aperture $\varphi = 45°$. We note that the asymmetries in ε found at low rotation speed disappear with large GVD at large rotation speed, where the surfaces are similar to the case $\varphi = 90°$.

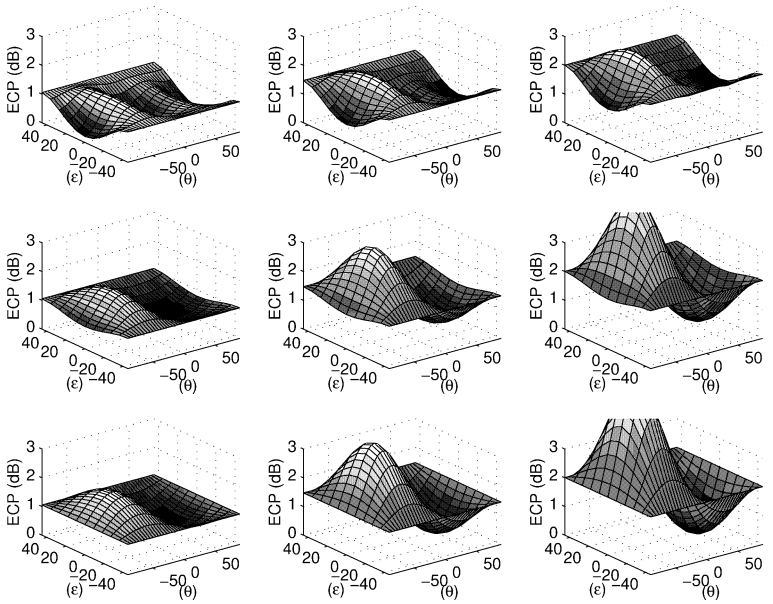

Fig. 13. Theoretical ECP surfaces for the same rotation model simulated in Fig. 12.

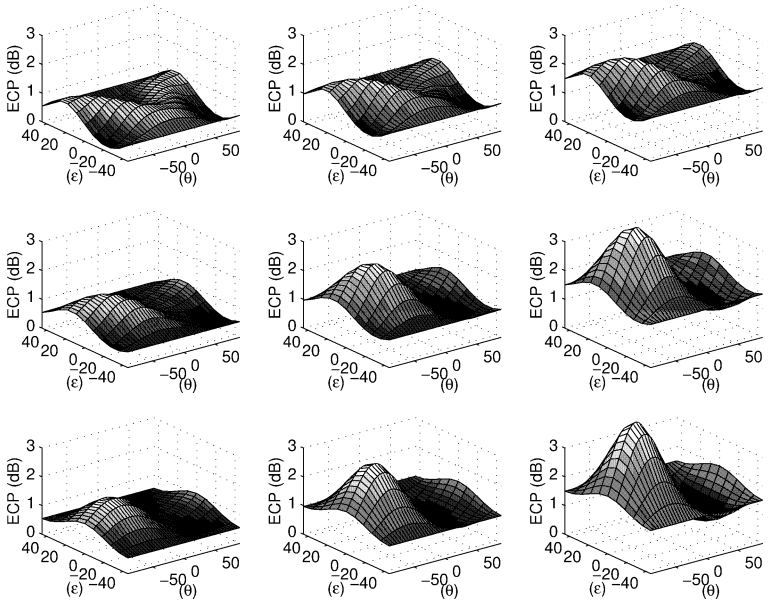

Fig. 14. Theoretical ECP surfaces for the same rotation model of Fig. 13, but now with $\hat{k} = \hat{s}_3$ and $\hat{b}_0$ with zero azimuth and 45° elevation.

In all cases we verify that the ECP versus input SOP surfaces synthesized with the GCF well capture the *shape* of the actual ECP surfaces, even when the absolute accuracy of the ECP prediction is poor, which justifies the use of the eye opening (as per the GCF) as a feedback control signal in PMD compensation [41,42].

6. Conclusions

In this paper we gave evidence that the PSPs are not the most straightforward analytical tool to describe the field intensity at the output of a single-mode fiber affected by PMD.

By introducing the concept of the extracted Jones matrices, we have shown that the eigenmodes of such matrices are more stable in frequency than the PSPs, and can also be used to evaluate exact expressions of the output intensity. In particular, we have given a new expression of the received intensity in terms of the Pauli vector of the right-extracted matrix, which clearly highlights the effect of the input SOP on the received intensity, and allows a detailed study of the signal distortion induced by PMD, which may be useful both to evaluate system performance and to devise electrical compensation algorithms at the receiver.

Working with the extracted eigenmodes, we have also obtained a novel eye closure penalty formula, which we called the GCF, which is the natural extension of well-known ECP formulas for first order PMD. We then showed that such formula can also be obtained by reasoning in terms of the small-signal IM/IM transfer function introduced by Noé et al. [24], thereby providing a comprehensive theoretical framework for the understanding of the ECP formulas in use today, and for their generalization.

With the idea of approximating the extracted eigenmodes trajectory with its osculating circle, we have obtained the rotation model, originally introduced by Mecozzi et al. [22]. We have simulation evidence that the frequency response of most single-mode fibers can be accurately reproduced, over bandwidths of the order of the inverse of the rms DGD, by a fitted rotation model of their right-extracted matrix, although we do not provide details, since it is not the purpose of this paper to validate the rotation model. As a partial validation, we refer the reader to the assessment of the compensator performance in [22].

In its simplest version, the rotation model has linear retardation and constant rotation speed, and is thus completely described by only three parameters: (1) the DGD $\Delta\tau$, (2) the angular rotation speed $2k_{\mathrm{v}}$, (3) the aperture φ of the eigenmodes circle. Once the joint statistics of such parameters are known, our analytical GCF formula allows to quickly find the statistics of the ECP and thus the system outage probability.

Thus the big task that remains to bring our approach to its full conclusion is a thorough statistical analysis of the extracted matrices. We already investigated the statistical properties of the eigenmodes and eigenvalues of global Jones matrix, and studied their relation with the PMD vector [21], and extensions of such work to the extracted matrices are under way.

Acknowledgment

It is a pleasure to acknowledge the contribution of Giorgio Bresciani in deriving many of the numerical results in this paper.

Appendix A. Relations between the PMD vector and the Pauli vector

This appendix describes the connection of matrix U with the output PMD vector $\vec{\Omega}_{\mathrm{o}}$. Such vector is obtained from the defining relation [25]

$$N \triangleq U'U^{\dagger} = -\frac{j}{2}\big[\vec{\Omega}_{\mathrm{o}} \cdot \vec{\sigma}\big], \tag{A.1}$$

a prime indicating derivative with respect to ω. Writing $\vec{\Omega}_{\mathrm{o}} \triangleq \Delta\tau\hat{q}_{\mathrm{o}}$ in terms of its magnitude $\Delta\tau$, the DGD, and its direction vector $\hat{q}_{\mathrm{o}}$, it is apparent from (A.1) that the

Jones vector associated with $\hat{q}_o$ is an eigenvector of the matrix $U'U^\dagger$. Similarly, the input PMD vector of U is defined by the equation

$$A \triangleq U^\dagger U' = -\frac{j}{2}\big[\vec{\Omega}_i \cdot \vec{\sigma}\big]. \tag{A.2}$$

The connection with the eigenmode $\vec{b}$ of U is easily established from the following theorem.

Theorem. *Let $P = H_1 H_2$ be the product of two Jones matrices H_1 and H_2, with Pauli vectors $\underline{h}_1 = [h_{01}; \vec{h}_1]$ and $\underline{h}_2 = [h_{02}; \vec{h}_2]$, respectively. Then the Pauli vector of P is*

$$\underline{p} = \begin{bmatrix} h_{01} & \vec{h}_1^{\mathrm{T}} \\ \vec{h}_1 & h_{01} I + j\big[\vec{h}_1 \times\big] \end{bmatrix} \underline{h}_2 = \begin{bmatrix} h_{02} & \vec{h}_2^{\mathrm{T}} \\ \vec{h}_2 & h_{02} I - j\big[\vec{h}_2 \times\big] \end{bmatrix} \underline{h}_1, \tag{A.3}$$

where the above matrices are block-partitioned, I is the 3×3 identity matrix, and $[\vec{h}\times]$ is the cross-product matrix. □

The theorem expresses the *concatenation rule* for the Pauli vectors in a chain of linear optical elements. Its simple proof is based on Eq. (A1-3) in [29]. It can be shown that (A.3) has a simple geometrical interpretation in terms of spherical trigonometry [43], as reported in Fig. 15. By recalling the definition of Pauli vector (10), we construct a spherical triangle on the Poincaré sphere (i.e., one whose edges are portions of great circles) with vertices $\hat{b}_2$ and $-\hat{b}_1$, and angles $\Delta\phi_2/2$ and $\Delta\phi_1/2$ as reported in the figure. The third vertex and angle thus coincide with the eigenmode and half the retardation of the resulting Jones matrix $P = H_1 H_2$, i.e.,

$$\underline{p} = [p_0; \vec{p}] = \left[\cos\left(\frac{\Delta\phi_p}{2}\right); -j \sin\left(\frac{\Delta\phi_p}{2}\right)\hat{b}_p\right].$$

Thus using (A.3) for U' and $U^\dagger$ one gets [12]

$$\vec{\Omega}_{i/o} = \Delta\phi' \hat{b} + \sin\Delta\phi \vec{b}' \pm (1 - \cos\Delta\phi)\big(\vec{b}' \times \hat{b}\big), \tag{A.4}$$

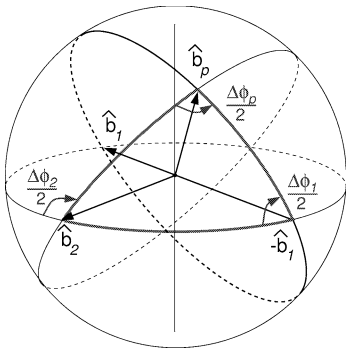

Fig. 15. Spherical triangle constructed from the eigenmodes $\hat{b}_{1,2}$ and retardations $\Delta\phi_{1,2}$ of matrices H_1 and H_2. The eigenmode $\hat{b}_p$ and retardation $\Delta\phi_p$ of matrix $P = H_1 H_2$ result.

where the plus sign applies to the input PMD vector. This is the desired connection between eigenmode vector $\hat{b}$ and PMD vector. Taking the squared magnitude of both sides of (A.4) gives

$$\Delta\tau^2(\omega) = \big(\Delta\phi'(\omega)\big)^2 + 2\big(1-\cos\Delta\phi(\omega)\big)\big|\vec{b}'(\omega)\big|^2, \tag{A.5}$$

from which we learn that $\Delta\phi' = \Delta\tau$ when either $\Delta\phi = 0$ or $\vec{b}' = 0$, i.e., the eigenmode is not depolarized.

The connection between the Pauli vector u and the PMD vector is obtained by writing $U' = NU = UA$ and using (A.3)

$$\underline{u}' = -j\begin{bmatrix} \vec{u}^{\mathrm{T}} \\ u_0 I \pm j\big[\vec{u}\times\big] \end{bmatrix}\frac{\vec{\Omega}_{\mathrm{i/o}}}{2}. \tag{A.6}$$

Such relation shows that, once $\underline{u}$ is known, $\underline{u}'$ is a *linear transformation* of the PMD vector [20,21]. Now differentiate U a second time

$$U'' = N'U + NU' = N'U + N^2U = \left[N' - \frac{|\vec{\Omega}_{\mathrm{o}}|^2}{4}\sigma_0\right]U,$$

where we used the fact that $N^2 = A^2 = -(\Delta\tau^2/4)\sigma_0$, as easily obtained from (A.3). The matrix in square brackets thus has Pauli vector $[-\Delta\tau^2/4; -j\vec{\Omega}'_{\mathrm{o}}/2]$. Using again (A.3) we get

$$\underline{u}'' = \begin{bmatrix} u_0 & \vec{u}^{\mathrm{T}} \\ \vec{u} & u_0 I \pm j\big[\vec{u}\times\big] \end{bmatrix}\begin{bmatrix} -\Delta\tau^2/4 \\ -(j/2)\vec{\Omega}'_{\mathrm{i/o}} \end{bmatrix}, \tag{A.7}$$

where we included the corresponding result for the input PMD vector. The relation states that, once $\underline{u}$ is known, $\underline{u}''$ is a transformation of the pair $[\vec{\Omega}, \vec{\Omega}']$. In other terms, once u is known, the pair $[\underline{u}', \underline{u}'']$ is a transformation of $[\vec{\Omega}, \vec{\Omega}']$ and vice versa. In [21] we proved that for long fibers $\underline{u}$ is statistically independent of $\vec{\Omega}$ and all its derivatives.

Appendix B. On the circular motion of the global eigenmodes

This appendix is devoted to explain why the global eigenmodes are most often more depolarized than the PSPs and the extracted eigenmodes. We already noted in the introduction that the correlation bandwidth of the global eigenmodes is $\sqrt{2/3}$ that of the PSPs, i.e., the PSPs are less depolarized on average than the global eigenmodes. We believe the analytical reason can be explained as follows.

From (A.4) it is obvious that, if the global eigenmode $\hat{b}$ is frequency independent, then the PMD direction vector $\hat{q}$ must also be frequency independent. However, the converse is not true: a frequency-independent $\hat{q}$ implies that the global eigenmode $\hat{b}(\omega)$ describes *great circles* as ω varies.

To show this fact, we start from the relation among the eigenmodes and retardations of matrices U, U_0, and U_{r}

$$\begin{aligned} \cos\frac{\Delta\phi}{2} &= \cos\frac{\Delta\phi_0}{2}\cos\frac{\Delta\phi_{\mathrm{r}}}{2} - \sin\frac{\Delta\phi_0}{2}\sin\frac{\Delta\phi_{\mathrm{r}}}{2}\hat{b}_0\cdot\hat{b}_{\mathrm{r}}, \\ \sin\frac{\Delta\phi}{2}\hat{b} &= \cos\frac{\Delta\phi_0}{2}\sin\frac{\Delta\phi_{\mathrm{r}}}{2}\hat{b}_{\mathrm{r}} + \sin\frac{\Delta\phi_0}{2}\cos\frac{\Delta\phi_{\mathrm{r}}}{2}\hat{b}_0 + \sin\frac{\Delta\phi_0}{2}\sin\frac{\Delta\phi_{\mathrm{r}}}{2}\hat{b}_0\times\hat{b}_{\mathrm{r}}, \end{aligned} \tag{B.1}$$

which is easily obtained by expressing such unitary matrices in their form (10) and then using (A.3).

We now build a global matrix $U(\omega) \triangleq U_0 U_{\rm r}(\omega)$ whose right-extracted matrix $U_{\rm r}$ follows a *first order-model*:

$$\Delta\phi_{\rm r}(\omega) = \Delta\tau\,\omega, \qquad \hat{b}_{\rm r}(\omega) = \hat{b}_{\rm r}, \tag{B.2}$$

with a frequency-independent eigenmode $\hat{b}_{\rm r}$. Then from (A.4) we conclude that the input and output PMD directions $\hat{q}_{\rm ro} = \hat{q}_{\rm ri}$ of $U_{\rm r}$ are ω-independent and equal $\hat{b}_{\rm r}$, and from what we saw in Section 3 that they coincide with the input PMD direction $\hat{q}_{\rm i}$ of U. We have the following results.

Proposition. *If $U_{\rm r}$ follows a first order-model, the eigenmode $\hat{b}(\omega)$ of $U(\omega)$ is ω-independent if and only if $\hat{b}_{\rm r} = \hat{b}_0$.*

Proof. While the "if" part is trivial, the "only if" part is shown as follows. When $\vec{b}'(0) = 0$, Eq. (A.4) gives that $\vec{\Omega}(0)$ is aligned with $\hat{b}_0$, and since $\vec{\Omega}_{\rm r}(\omega)$ is a rotation of $\vec{\Omega}(\omega)$ around $\vec{b}_0$, then $\vec{\Omega}_{\rm r}(0) = \vec{\Omega}(0) = \vec{\Omega}_{\rm l}(0)$. If also $U_{\rm r}$ is first-order, from Eq. (A.4) $\hat{b}_{\rm r}$ is aligned with $\vec{\Omega}_{\rm r}$, and therefore with $\hat{b}_0$. This shows that if there is no motion of $\hat{b}$ at $\omega = 0$ (i.e., $\vec{b}'(0) = 0$), then $\vec{b}$ is frequency independent, i.e., there is no motion at all ω's, and $\hat{b}_{\rm r} = \hat{b}_0$. □

Proposition. *If $\hat{b}_{\rm r}$ is not aligned with $\hat{b}_0$, the eigenmode $\hat{b}(\omega)$ of $U(\omega)$ describes great* (*i.e., maximal*) *circles on the Poincaré sphere as ω varies, in general at non-uniform speed in ω.*

Before providing a general proof of this result, let's check it in the particular case $\hat{b}_{\rm r} \cdot \hat{b}_0 = 0$ and $\Delta\phi_0/2 = \pi/2$. Then from (B.1) we get that $\Delta\phi = \pi$, independent of frequency, and that

$$\hat{b}(\omega) = \cos\frac{\Delta\tau\,\omega}{2}\hat{b}_0 + \sin\frac{\Delta\tau\,\omega}{2}\big(\hat{b}_0 \times \hat{b}_{\rm r}\big),$$

which is clearly the parametric equation of a great circle on the plane orthogonal to $\hat{b}_{\rm r}$ and containing $\hat{b}_0$. Such equatorial trajectory is described at constant angular speed $\Delta\tau/2$. Here is the general proof.

Proof. Let is first visualize the eigenmodes of U on the Poincaré sphere. For fixed ω, the Mueller matrix of U operates a rotation around the axis $\hat{b}(\omega)$. Recalling (1) and (B.2), such rotation is the composition of two successive rotations: a first rotation by an angle $\Delta\tau\,\omega$ around $\hat{b}_{\rm r}$ and a second rotation by $\Delta\phi_0$ around $\hat{b}_0$. Hence *by definition the ISOP coinciding with the eigenmode $\hat{b}$ must be rotated by the operator U_0 so as to undo the rotation of the operator $U_{\rm r}$*. This is shown in Fig. 16(a).[4] Two circles are drawn, one around $\hat{b}_{\rm r}$ and one around $\hat{b}_0$, which intersect at points E_1 and F_1. The ISOP aligned with E_1 is the eigenmode $\hat{b}$ for the specific $\Delta\phi_0$ marked in the figure, and for the special ω corresponding to the angle $\Delta\tau\,\omega$ marked in the figure. The closed loop trajectory from E_1 back to E_1 along the two arcs intercepted on such circles is the effect of the two successive rotations. Once we realize that this *is the basic pattern to recognize the eigenmodes*, we understand that, once $\Delta\phi_0$ is given, the eigenmode $E_{\rm i}$ at the generic frequency $\omega_{\rm i}$ must lay on the *equatorial plane* at an

[4] For $\Delta\phi_0 > 0$, the Mueller matrix $e^{\Delta\phi_0[\hat{b}_0\times]}$ operates a *counter-clockwise* rotation around $\hat{b}_0$. Hence, in Fig. 16 we implicitly assume $\Delta\phi_0 > 0$ and $\omega > 0$, since rotations are all counterclockwise.

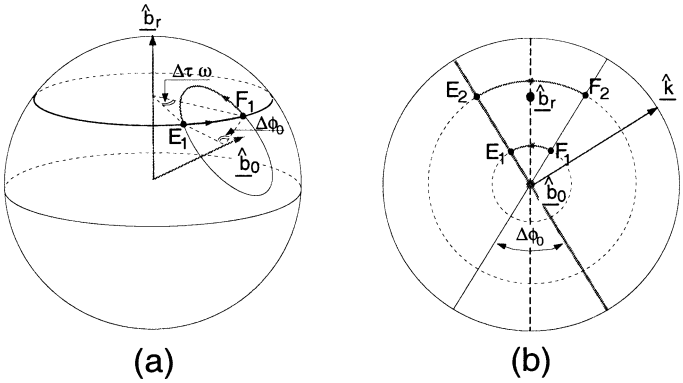

Fig. 16. How to visualize the global eigenmode vector on the Poincaré sphere (see text).

angle $\Delta\phi_0/2$ from the plane that includes $\hat{b}_0$ and $\hat{b}_r$, as shown in Fig. 16(b), which is a view of the Poincaré sphere from the top of $\hat{b}_0$. Here we recognize again the eigenmode E_1 of Fig. 16(a), and we see for instance another eigenmode E_2 corresponding to a larger angle $\Delta\tau\,\omega$. Obviously $\hat{b}_0$ is the eigenmode corresponding to $\Delta\tau\omega = 0, \pm 2\pi, \pm 4\pi, \ldots$. We also note that the points F_i are the eigenmodes when $\Delta\phi_0$ is changed in $2\pi - \Delta\phi_0$. Looking at Fig. 16(b) we realize that, for a fixed plane $(\hat{b}_r, \hat{b}_0)$, the eigenmode great circle (and hence its rotation axis $\hat{k}$, also shown in the figure) is the same for all values of $\cos\phi \triangleq \hat{b}_0 \cdot \hat{b}_r$, which only affects the rotation angular velocity. □

The above shows that when the PSPs/extracted eigenmodes are frequency-independent, the global eigenmode vector describes maximal circles. By observing many fiber realizations obtained by the retarded wave-plate model [32], we found that there are also situations in which the global eigenmode vector is almost fixed, while the PSPs and the extracted eigenmode vector violently rotate. This seems impossible at a superficial look of (A.4), since when the global $\underline{\hat{b}}$ is fixed, so must be the direction of $\underline{\Omega}$. However, according to the PMD vector concatenation rule [25], when the many PSP vectors of the waveplate model concatenate so as to loop back to the origin and give rise to a very small $|\underline{\Omega}(0)| = \Delta\tau(0)$ [40], its is very likely that the rotation speed $2k_v(0)$ is very large, i.e., $\underline{\Omega}(\omega)$ is very depolarized at $\omega = 0$. In the limit of vanishing $\Delta\tau$, it is apparent from the spectral decomposition of $U_r(\omega)$ that $U(\omega) \cong U(0)B(\omega)\sigma_0 B^\dagger(\omega) \cong U(0)$, i.e., we have a global matrix which is almost frequency independent, and so are its eigenmodes, while the extracted eigenmodes rotate at high speed (since so behave the PSPs). In such situation, which corresponds to a vanishingly small $\Delta\tau$ and extremely large $2k_v$, there is no signal distortion.

References

[1] C.D. Poole, R.E. Wagner, Phenomenological approach to polarisation dispersion in long single-mode fibres, IEE Electron. Lett. 22 (1986) 1029–1030.

[2] C.D. Poole, C.R. Giles, Polarization-dependent pulse compression and broadening due to polarization dispersion in dispersion-shifted fiber, Opt. Lett. 13 (1988) 155–157.

[3] F. Curti, B. Daino, G. De Marchis, F. Matera, Statistical treatment of the evolution of the principal states of polarization in single-mode fibers, IEEE J. Lightwave Technol. 8 (1990) 1162–1166.

[4] G.J. Foschini, C.D. Poole, Statistical theory of polarization dispersion in single mode fibers, IEEE J. Lightwave Technol. 9 (1991) 1439–1456.

[5] P. Ciprut, B. Gisin, N. Gisin, R. Passy, J.P. Von der Weid, F. Prieto, C.W. Zimmer, Second-order polarization mode dispersion: Impact on analog and digital transmission, IEEE J. Lightwave Technol. 16 (1998) 757–771.
[6] P.K.A. Wai, C.R. Menyuk, Polarization mode dispersion, decorrelation, and diffusion in optical fibers with randomly varying birefringence, IEEE J. Lightwave Technol. 14 (1996) 148–157.
[7] C.D. Poole, R.W. Tkach, A.R. Chraplyvy, D.A. Fishman, Fading in lightwave systems due to polarization mode dispersion, IEEE Photon. Technol. Lett. 3 (1991) 68–70.
[8] C.-J. Chen, System impairment due to polarization mode dispersion, in: Proc. OFC'99, 1999, pp. 77–79, paper WE2-1.
[9] N. Gisin, J.P. Pellaux, Polarization mode dispersion: time versus frequency domains, Opt. Commun. 89 (1992) 316–323.
[10] M. Karlsson, J. Brentel, Autocorrelation function of the polarization-mode dispersion vector, Opt. Lett. 24 (14) (1999) 939–941.
[11] H. Büllow, System outage probability due to first- and second-order PMD, IEEE Photon. Technol. Lett. 10 (1998) 696–698.
[12] M. Karlsson, Polarization mode dispersion-induced pulse broadening in optical fibers, Opt. Lett. 23 (1998) 688–690.
[13] W. Shieh, Principal states of polarization for an optical pulse, IEEE Photon. Technol. Lett. 11 (1999) 677–679.
[14] W. Shieh, On the second order approximation of PMD, IEEE Photon. Technol. Lett. 12 (2000) 290–292.
[15] F. Bruyere, Impact of first- and second-order PMD in optical digital transmission systems, Optical Fiber Technol. 2 (1996) 269–280.
[16] H. Kogelnik, L.E. Nelson, J.P. Gordon, R.M. Jopson, Jones matrix for second-order polarization mode dispersion, Opt. Lett. 25 (2000) 19–21.
[17] L.E. Nelson, R.M. Jopson, H. Kogelnik, Polarization mode dispersion penalties associated with rotation of principal states of polarization in optical fiber, in: Proc. OFC'2000, 2000, pp. 25–27, paper ThB2.
[18] Y. Li, A. Eyal, A. Yariv, Higher order error of discrete fiber model and asymptotic bound on multistaged PMD compensators, IEEE J. Lightwave Technol. 18 (2000) 1205–1213.
[19] H.A. Haus, Group velocity, energy, and polarization mode dispersion, J. Opt. Soc. Amer. B 16 (1999) 1863–1867.
[20] A. Bononi, A. Vannucci, Statistics of the Jones matrix of fibers affected by polarization mode dispersion, Opt. Lett. 26 (2001) 675–677.
[21] A. Vannucci, A. Bononi, Statistical characterization of the Jones matrix of long fibers affected by PMD, IEEE J. Lightwave Technol. 20 (2002).
[22] A. Mecozzi, M. Shtaif, M. Tur, J. Nagel, A simple compensator for high order polarization mode dispersion effects, in: Proc. OFC'2000, 2000, pp. 192–194, paper WL2.
[23] A. Vannucci, A. Bononi, Sensitivity penalty distribution in fibers with PMD: A Novel semi-analytical technique, in: Proc. OFC'2002, 2002, pp. 54–56, paper TuI6.
[24] R. Noe', D. Sandel, M.Y. Dierolf, S. Hinz, V. Mirvoda, A. Scoepflin, C. Glingener, E. Gottwald, C. Scheerer, G. Fischer, T. Weyrauch, W. Haase, Polarization mode dispersion compensation at 10, 20, and 40 Gb/s with various optical equalizers, IEEE J. Lightwave Technol. 17 (1999) 1602–1615.
[25] J.P. Gordon, H. Kogelnik, PMD fundamentals: Polarization mode dispersion in optical fibers, Proc. Nat. Acad. Sci. 97 (2000) 4541–4550.
[26] K. Hofman, R. Kunze, Linear Algebra, 2nd. Ed., Prentice–Hall, Englewood Cliffs, NJ, 1971.
[27] G. De Marchis, E. Lannone, Polarization dispersion in single-mode optical fibers: A simpler formulation based on pulse envelope propagation, Microwave Optical Technol. Lett. 4 (2) (1991) 75–77.
[28] G.J. Foschini, L.E. Nelson, R.M. Jopson, H. Kogelnik, Probability densities of second-order polarization mode dispersion including polarization dependent chromatic fiber dispersion, IEEE Photon. Technol. Lett. 12 (2000) 293–295.
[29] N.J. Frigo, A generalized geometrical representation of coupled mode theory, IEEE J. Quantum Electron. QE-22 (1986) 2131–2140.
[30] F. Corsi, A. Galtarossa, L. Palmieri, Polarization mode dispersion characterization of single-mode optical fiber using backscattering technique, J. Lightwave Technol. 16 (1998) 1832–1843.
[31] R.M. Jopson, L.E. Nelson, H. Kogelnik, Measurement of second-order polarization-mode dispersion vectors in optical fibers, IEEE Photon. Technol. Lett. 11 (1999) 1153–1155.
[32] A.O. Dal Forno, A. Paradisi, R. Passy, J.P. von der Weid, Experimental and theoretical modeling of polarization mode dispersion in single-mode fibers, IEEE Photon. Technol. Lett. 12 (2000) 296–298.
[33] C. Francia, F. Bruyere, D. Pennickx, M. Chbat, PMD second-order effects on pulse propagation in single-mode optical fibers, Photon. Technol. Lett. 10 (1998) 1739–1741.

[34] C. Francia, D. Pennickx, Polarization mode dispersion in single-mode optical fibers: time impulse response, in: Proc. IEEE ICC'99, Vancouver, Canada, 1999, pp. 1731–1735.
[35] T. Ono, S. Yamazaki, H. Shimizu, H. Emura, Polarization control method for suppressing polarization mode dispersion in optical transmission systems, IEEE J. Lightwave Technol. 12 (1994) 891–898.
[36] G. Agrawal, Nonlinear Fiber Optics, Academic Press, 1989.
[37] G. Ishikawa, H. Ooi, Polarization mode dispersion sensitivity and monitoring in 40 Gbit/s OTDM and 10 Gbit/s NRZ transmission experiments, in: Proc. OFC'98, 1998, pp. 117–119, paper WC5.
[38] H. Ooi, Y. Akyiama, G. Ishikawa, Automatic polarization-mode dispersion compensation in 40 Gbit/s transmission, in: Proc. OFC'99, 1999, pp. 86–88, paper WE5-1.
[39] I.P. Kaminow, T. Koch (Eds.), Optical fiber Telecommunications, Vol. IIIA, Academic Press, 1997.
[40] D. Pennickx, F. Bruyere, Impact of the statistics of second-order polarization mode dispersion on system performance, in: Proc. OFC'98, 1998, pp. 340–341, paper ThR2.
[41] F. Buchali, S. Lanne, J.-P. Thièry, W. Baumert, H. Bülow, Fast eye monitor for 10 Gb/s and its application for optical PMD compensation, in: Proc. OFC'2001, 2001, paper TuP5.
[42] F. Buchali, W. Baumert, H. Bülow, J. Poirrier, S. Lanne, A 40 Gb/s eye monitor and its application to adaptive PMD compensation, in: Proc. OFC'2002, 2002, pp. 202–203, paper WE6.
[43] J.B. Kuipers, Quaternions and Rotation Sequences, Princeton University Press, Princeton, NJ, 1999.