A Statistical Treatment of Cross-Polarization Modulation in DWDM Systems

Marcus Winter, *Student Member, IEEE*, Christian-Alexander Bunge, *Member, IEEE*, Dario Setti, and Klaus Petermann, *Fellow, IEEE*

***Abstract*—Starting from a model of random nonlinear polarization rotations for the effect of cross-polarization modulation in DWDM systems, we derive the mean distribution of the time-dependent polarization states at an arbitrary location within an optical link. We show that this distribution is fully parameterized by the degree of polarization of the particular wavelength channel, and we derive expressions to approximate this parameter for general optical links consisting of multiple optically amplified and dispersion-compensated spans, as well as related power thresholds. From the analytical expressions we derive a method to significantly reduce the detrimental effects.**

***Index Terms*—Communication system nonlinearities, optical fiber communication, optical fiber polarization, optical fiber theory, optical Kerr effect.**

I. Introduction

IT has been known for quite some time that the nonlinear coupling of two or more optical signals mediated by the Kerr effect in dielectric media affects not only the phase of these signals, but also their polarization state. This effect enables devices such as Kerr shutters, which combine a pump beam and polarization filter to modulate a probe [1], [2]. For a long time, this nonlinear polarization rotation was neglected in optical systems theory, because the components of these systems were sufficiently independent of the polarization state. However, with increasing relevance of polarization phenomena, especially in the context of polarization-mode dispersion, polarization multiplexing, and coherent reception, it has come under increased scrutiny, both theoretically and experimentally [3]–[12].

Most published work so far has been concerned with the nonlinear polarization rotation itself, its description and demonstration [4], [6], [9], [13]. Previous contributions almost exclusively took a single pump – single probe approach [4]–[6], [9], [13]. Very few authors examined it, in the context of multi-wavelength optical transmission systems, and then only rudimentary [8], [10]. This is the void which the present contribution intends to fill.

We start by reviewing the equations and approximations necessary to derive the nonlinear polarization rotations in randomly birefringent fiber in Section II. From the nature of these rotations we then derive the statistical distribution of the polarization states after nonlinear propagation of a length of fiber, and subsequently an expression to quantify this distribution in Section III. In Section IV we then show how this expression is related to the parameters of the transmission fiber in general, and then make this general expression more precise, starting with simple, short systems in which linear polarization effects such as polarization mode dispersion can be neglected in Section V, and continue with arbitrarily long systems in which these effects play a major role in Section VI. We compare the impact of cross-polarization modulation depending on the properties of the fiber used within the system in terms of the nonlinear power threshold for the interfering copropagating channels, and we also show how systems with residual, uncompensated group velocity dispersion in each span can significantly reduce the magnitude of cross-polarization modulation in Section VII.

Manuscript received February 02, 2009; revised May 25, 2009. First published June 16, 2009; current version published August 12, 2009.

M. Winter and K. Petermann are with the Technische Universität Berlin, FG Hochfrequenztechnik/Photonik (HFT4), 10587 Berlin, Germany (e-mail: winter@hft.ee.tu-berlin.de; petermann@tu-berlin.de).

C.-A. Bunge is with the Hochschule für Telekommunikation Leipzig, FG Photonik, 04277 Leipzig, Germany (e-mail: bunge@hft-leipzig.de).

D. Setti is with Nokia Siemens Networks GmbH & Co. KG, 81541 Munich, Germany (e-mail: dario.setti@nsn.com).

Digital Object Identifier 10.1109/JLT.2009.2025394

II. The SOP Evolution Equation

Several authors have derived various expressions for the nonlinear evolution of a vector field traveling along an optical fiber, including polarization effects [3], [11], [13]–[16]. We will briefly review the derivation of these expressions and at the same time introduce the notation that will be used throughout this article, in order to make our work as self-contained as possible. Those readers familiar with the subject matter may wish to skip ahead to Section III.

Remarks on Notation

To distinguish two-dimensional complex vectors which describe the optical field in Jones space from their three-dimensional brethren in Stokes space, we will use Dirac's bra-ket formalism, e.g., $|a\rangle = (a_x, a_y)^{\mathrm{T}}$, to denote Jones vectors. They are normalized such that $\langle a|a\rangle$ equals the optical power, where $\langle a| = (a_x^*, a_y^*)$ is the conjugate transpose. Stokes vectors are denoted by bold letters, e.g., $\mathbf{S} = (S_1, S_2, S_3)^{\mathrm{T}}$, and unit vectors in Stokes space, describing a state of polarization (SOP), are decorated with a hat, as in $\widehat{\mathbf{S}}$. Complex 2×2 operators or matrices are ornamented with a macron, such as $\bar{\mathbf{T}}$, except for the (permuted) Pauli matrices

$$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \sigma_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \sigma_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_3 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \tag{1}$$

and the Pauli vector $\vec{\sigma} \equiv (\sigma_1, \sigma_2, \sigma_3)^{\mathrm{T}}$, for which we adopt the notation of [17]. We start by writing the DWDM optical

field in phasor form as a field envelope modulated onto a carrier frequency ω_0

$$\big|a(z,t)\big\rangle = \big|v(z,t)\big\rangle \exp\big(i\omega_0 t - i\beta_0 z\big). \tag{2}$$

Propagation in the z-direction along a birefringent fiber can be described with the vector form of the nonlinear Schrödinger equation (NLSE) for this envelope

$$\partial_z\big|v\big\rangle + \frac{\alpha}{2}\big|v\big\rangle + i\Delta\beta_0\bar{\boldsymbol{\sigma}}\big|v\big\rangle + \Big(\beta_1\sigma_0 + \Delta\beta_1\bar{\boldsymbol{\sigma}}\Big)\partial_t\big|v\big\rangle - i\frac{\beta_2}{2}\partial_t^2\big|v\big\rangle - i\gamma\Big[\big\langle v\big|v\big\rangle\big|v\big\rangle - \frac{1}{3}\Big(\big\langle v\big|\sigma_3\big|v\big\rangle\Big)\sigma_3\big|v\big\rangle\Big] = 0. \tag{3}$$

The second term in (3) describes linear attenuation with power attenuation coefficient α. The third term is responsible for birefringence, with $\Delta\beta_0 = \big(\beta_{0s} - \beta_{0f}\big)/2$, in which the propagation constant expansion terms for the slow and fast birefringence axes are defined as

$$\beta_{n(s,f)}(z) = \partial_\omega^n \beta_{(s,f)}(\omega, z)\Big|_{\omega=\omega_0} \tag{4}$$

where ∂_ω^n is short-hand for $\partial^n/\partial\omega^n$. The birefringence axes correspond to the orthogonal eigenvectors of the coupling matrix $\bar{\boldsymbol{\sigma}}$. The fourth term in (3) describes different group velocities along the birefringence axes, with $\Delta\beta_1 = \big(\beta_{1s} - \beta_{1f}\big)/2$, and is the physical cause of polarization mode dispersion (PMD). The fifth term describes group velocity dispersion (GVD), which is generally assumed to be polarization-independent [18]. We neglect higher-order dispersion to keep the expressions simple — it is straightforward to modify the equations to include chromatic dispersion of any order, assuming such higher-order dispersion to also be polarization-independent. The final term in (3) expresses the Kerr nonlinearity in vector form [11], in which γ is the fiber nonlinearity coefficient. All coefficients and the coupling matrix are generally functions of the propagation distance z, even if not explicitly noted to keep the expressions concise.

We now make the transformation

$$\big|u\big\rangle = \big(A\,\bar{\mathbf{T}}\big)^{-1}\big|v\big\rangle \tag{5}$$

with

$$A(z) = \exp\Big[-\frac{\alpha}{2}z\Big] \tag{6}$$

$$\bar{\mathbf{T}}(z) = \exp\Big[-i\int_0^z \Delta\beta_0(\zeta)\,\bar{\boldsymbol{\sigma}}(\zeta)\,d\zeta\Big]. \tag{7}$$

With A we eliminate the attenuation-related term from the NLSE (3), and the unitary transformation $\bar{\mathbf{T}}$ eliminates the birefringence term. The integral expression is a result of the spatially varying nature of the birefringence axes, as expressed by $\bar{\boldsymbol{\sigma}}(z)$ and the birefringence strength $\Delta\beta_0(z)$. It has to be understood in a symbolic fashion, as developed in [19], because the operators $\bar{\boldsymbol{\sigma}}(z)$ do not commute and thus cannot be integrated in a straightforward manner. A formal expression for $\bar{\mathbf{T}}(z)$ can be obtained by expanding the integral in (7) into a Magnus series and integrating the resulting expression [20], [21]. Such a treatment, however, is not required in the scope of this work because we do not care to know the absolute polarization state of the signal in the laboratory reference frame of $\big|v\big\rangle$. In polarization-maintaining fiber, in which $\bar{\boldsymbol{\sigma}}(z)$ is constant and thus commuting, we would obtain the simple expression $\bar{\mathbf{T}}(z) = \exp\big[-i\Delta\beta_0\,\bar{\boldsymbol{\sigma}}z\big]$. The effect of $\bar{\mathbf{T}}$ can be visualized well in Stokes space, in which the birefringence term induces a rapid rotation of the Stokes vector $\mathbf{V}$ associated with $\big|v\big\rangle$, and defined as $\mathbf{V} = \big\langle v\big|\vec{\sigma}\big|v\big\rangle$, around the local birefringence axis with a period corresponding to the beat length, which in regular standard single-mode fiber (SSMF) is on the order of some meters. The transformation in (7) is then equivalent to a counter-rotation of the coordinate system in such a way that this birefringence-related motion is halted for the Stokes vector $\mathbf{U} = \big\langle u\big|\vec{\sigma}\big|u\big\rangle$.

We further introduce the common retarded reference frame with $\tau = t - \beta_1 z$ and $\zeta = z$ in order to eliminate the β_1-term from (3) and obtain the following form of the NLSE for $\big|u\big\rangle$:

$$\partial_\zeta\big|u\big\rangle + \Delta\beta_1\bar{\mathbf{T}}^\dagger\bar{\boldsymbol{\sigma}}\bar{\mathbf{T}}\,\partial_\tau\big|u\big\rangle - i\frac{\beta_2}{2}\partial_\tau^2\big|u\big\rangle - i\exp\big(-\alpha\zeta\big)\gamma \times\Big[\big\langle u\big|u\big\rangle\big|u\big\rangle - \frac{1}{3}\Big(\big\langle u\big|\bar{\mathbf{T}}^\dagger\sigma_2\bar{\mathbf{T}}\big|u\big\rangle\Big)\bar{\mathbf{T}}^\dagger\sigma_2\bar{\mathbf{T}}\big|u\big\rangle\Big] = 0 \tag{8}$$

where † denotes the conjugate transpose matrix and we have used $\partial_t\bar{\mathbf{T}} = 0$ – the birefringence properties are assumed constant in time. In the PMD term, the unitary transform $\bar{\mathbf{T}}$ randomizes the birefringence axes isotropically on the Poincaré sphere. The unit vector $\widehat{\boldsymbol{\Omega}}$ pointing in the direction of the slow birefringence axis in these transformed coordinates is defined via

$$\bar{\mathbf{T}}^\dagger(z)\bar{\boldsymbol{\sigma}}(z)\bar{\mathbf{T}}(z) = \widehat{\boldsymbol{\Omega}}(z)\cdot\vec{\sigma} \tag{9}$$

where $\vec{\sigma}$ is again the Pauli vector. The transformed nonlinear terms are rapidly changing due to $\bar{\mathbf{T}}$, and since the Kerr nonlinearity affects the signal only over much larger distances, these terms may be averaged to arrive at the *Manakov PMD-equation* [3], [11], [16],

$$\partial_\zeta\big|u\big\rangle + \Delta\beta_1\bar{\mathbf{T}}^\dagger\bar{\boldsymbol{\sigma}}\bar{\mathbf{T}}\,\partial_\tau\big|u\big\rangle - i\frac{\beta_2}{2}\partial_\tau^2\big|u\big\rangle - i\exp\big(-\alpha\zeta\big)\frac{8}{9}\gamma\big\langle u\big|u\big\rangle\big|u\big\rangle = 0. \tag{10}$$

We now make the assumption that $\big|u\big\rangle$ comprises a number of narrow-band wavelength channels ν. Of those we will then concentrate on the impairments within a single channel ρ (which may be called the probe channel) which are caused by the other (interfering) channels $\nu \neq \rho$. We can expand the DWDM optical field as

$$\big|u(\zeta,\tau)\big\rangle = \sum_{\nu=1}^{N}\big|u_\nu(\zeta,\tau)\big\rangle\exp\big[i\Delta\omega_\nu\tau\big] \tag{11}$$

in which $\Delta\omega_\nu = \omega_\nu - \omega_0$. The nonlinear distortion in channel ρ can then be determined by expanding the triple product in (10) with (11)

$$\big\langle u\big|u\big\rangle\big|u\big\rangle = \sum_\eta\sum_\mu\sum_\nu\big\langle u_\eta\big|u_\mu\big\rangle\big|u_\nu\big\rangle\exp\big[i\Delta\omega_\rho\tau\big] \tag{12}$$

in which

$$\Delta\omega_\rho = \Delta\omega_\mu - \Delta\omega_\eta + \Delta\omega_\nu. \tag{13}$$

We define ω_0 to be the center frequency of the probe channel, so that $\Delta\omega_\rho = 0$. We further neglect four-wave mixing (FWM), assuming the phase-matching conditions to be poorly fulfilled in

any non-zero dispersion fiber, by demanding that either $\Delta\omega_\rho = \Delta\omega_\nu$ or $\Delta\omega_\rho = \Delta\omega_\mu$, leaving only terms corresponding to degenerate FWM. In order to concentrate on the nonlinear transmission effects, assuming the distortions of linear origin to be recoverable, we choose the probe channel to be unmodulated, continuous wave (CW). This removes the GVD and PMD terms – for the probe channel only – from the Manakov equation (10), and we are left with evolution equation for the probe field $\big|u_\rho\big\rangle$

$$\partial_\zeta\big|u_\rho\big\rangle - i\exp\big(-\alpha\zeta\big)\bar{\gamma} \times\Big[\underbrace{\big\langle u_\rho\big|u_\rho\big\rangle}_{\text{SPM}}+\sum_{\nu\neq\rho}\Big(\underbrace{\frac{3}{2}\,\big\langle u_\nu\big|u_\nu\big\rangle}_{\text{XPM}}+\underbrace{\frac{1}{2}\,\mathbf{U}_\nu\cdot\vec{\sigma}}_{\text{XPolM}}\Big)\Big]\big|u_\rho\big\rangle = 0 \quad (14)$$

in which $\bar{\gamma} = (8/9)\gamma$. The terms corresponding to self-phase modulation (SPM), cross-phase modulation (XPM) and cross-polarization modulation (XPolM) have been identified in (14). It is noteworthy that in the (averaged) Manakov regime, XPM acts with an efficiency of 3/2 relative to SPM. We have used [17, (3.9)] with [17, (3.7)] to transform the XPolM term into this particular product form. In (14), $\mathbf{U}_\nu$ is the Stokes vector corresponding to $\big|u_\nu\big\rangle$. We now transform (14) into Stokes space using the correspondence between [17, (6.4)] and [17, (6.8)] and noting that the SPM and XPM terms cause only phase changes, which are not reflected in Stokes space. We are then left with the concise form

$$\partial_\zeta\widehat{\mathbf{U}}_\rho(\zeta,\tau) = \mathbf{S}_\Sigma(\zeta,\tau)\times\widehat{\mathbf{U}}_\rho(\zeta,\tau) \quad (15)$$

in which

$$\mathbf{S}_\Sigma(\zeta,\tau) = \sum_{\substack{\nu=1\\ \nu\neq\rho}}^{N}\mathbf{S}_\nu(\zeta,\tau) \quad (16)$$

corresponds to the sum of all (scaled) Stokes vectors, into which the nonlinearity and attenuation coefficients have been absorbed:

$$\mathbf{S}_\nu(\zeta,\tau) = \mathbf{U}_\nu(\zeta,\tau)\,\bar{\gamma}\exp\big(-\alpha\zeta\big) \quad (17)$$

$\widehat{\mathbf{U}}_\rho = \mathbf{U}_\rho/|\mathbf{U}_\rho|$ denotes the SOP in Stokes space, and $|\cdot|$ is the vector norm, or length. The evolution (15) describes a rotation in Stokes space of the probe SOP $\widehat{\mathbf{U}}_\rho(\zeta,\tau)$ around the Stokes vector sum $\mathbf{S}_\Sigma(\zeta,\tau)$ with angular velocity, or rate, $\partial_\zeta\varphi = |\mathbf{S}_\Sigma(\zeta,\tau)|$. The $(N-1)$ vectors constituting the sum are random quantities whose length and orientation is determined by data modulation, attenuation, GVD, PMD, and nonlinearities, among others. In general, they may have a time-average value at each position ζ, around which they fluctuate. We can then separate $\mathbf{S}_\Sigma(\zeta,\tau)$ into that time-constant average and a fluctuating part

$$\mathbf{S}_\Sigma(\zeta,\tau) = \mathcal{E}\big[\mathbf{S}_\Sigma(\zeta,\tau)\big] + \Big(\mathbf{S}_\Sigma(\zeta,\tau) - \mathcal{E}\big[\mathbf{S}_\Sigma(\zeta,\tau)\big]\Big) \quad (18)$$

in which $\mathcal{E}\big[\cdot\big]$ is the statistical average over a short amount of time that we consider, some orders of magnitude longer than the symbol duration. During this time the PMD-related quantities $\bar{\boldsymbol{\sigma}}(\zeta)$ and $\Delta\beta_1(\zeta)$ can be assumed time-invariant, as has been implicitly done in all preceding expressions. Equation (15) then describes two separate phenomena: a common (nonlinear) rotation of the probe channel about the mean $\mathcal{E}\big[\mathbf{S}_\Sigma(\zeta,\tau)\big]$ and a rapid random fluctuation which will be different for any $\tau_1 \neq \tau_2$. We assume relevant optical receivers (which may include PMD compensators or polarization demultiplexers) to be able to follow the common polarization rotations which change only on the time scale of PMD, as would also be required of them in the absence of XPolM. We are therefore only interested in the rapidly fluctuating part in parentheses of (18) – this is the source of any detrimental effects due to XPolM, as receivers will not be able to adapt sufficiently fast to these rapid polarization state fluctuations. With this objective in mind, we define a transformed SOP

$$\widehat{\mathbf{X}} = \bar{\mathbf{R}}^{-1}\cdot\widehat{\mathbf{U}}_\rho \quad (19)$$

with the Stokes space rotation transformation

$$\bar{\mathbf{R}}(\zeta) = \exp\Big[\int_0^\zeta \mathcal{E}\big[\mathbf{S}_\Sigma(\xi,\tau)\big]\times\, d\xi\Big] \quad (20)$$

offsetting the nonlinear polarization rotation common to all τ, and where $\mathcal{E}\big[\mathbf{S}_\Sigma\big]\times$ is the crossproduct operator as defined in [17, (4.8)]. Here again, the operator described by the integral is to be understood symbolically, with a formal solution obtained by Magnus series expansion, similar to (7). And again, we are not interested in the absolute SOP in any predefined reference frame, but only the the relative probe SOP $\widehat{\mathbf{X}}$, and therefore do not need to perform the expansion. Inserting (19) into (15) yields the XPolM evolution equation

$$\partial_\zeta\widehat{\mathbf{X}}(\zeta,\tau) = \Big(\mathbf{S}_\Sigma(\zeta,\tau) - \mathcal{E}\big[\mathbf{S}_\Sigma(\zeta,\tau)\big]\Big)\times\widehat{\mathbf{X}}(\zeta,\tau). \quad (21)$$

In the following, we replace ζ by z and τ by t.

III. Statistical Description of SOP Fluctuations

If we consider for a moment the probe SOPs $\widehat{\mathbf{X}}(z,t_n)$ as independent particles n on the unit sphere, then (21) describes a random motion unique for each particle. Because $\big(\mathbf{S}_\Sigma(z,t_n) - \mathcal{E}\big[\mathbf{S}_\Sigma(z,t)\big]\big)$ will vary with z as a result of the various linear and nonlinear propagation effects, their motion will comprise unique sequences of random rotations. The SOPs are thus very similar to Brownian particles undergoing a diffusion process on the Poincaré sphere. Hence, we may call this particular effect of XPolM *SOP diffusion*.

Diffusion processes on the unit sphere $\mathbb{S}^2$ have been examined as early as the 1950s, and the resulting class of distributions is named *Brownian distribution* (or sometimes *Roberts-Ursell distribution* after their discoverers) [22]. Roberts and Ursell modeled a process consisting of a large number of discrete, randomly directed, independent steps to arrive at their limit distribution, which is given in terms of colatitude θ and azimuth ϕ on the unit sphere by the probability density function

$$p(\phi,\theta,V) = \frac{1}{4\pi}\sum_{\mu=0}^{\infty}C(\mu,V)\,LP_\mu(\cos\theta) \quad (22)$$

in which

$$C(\mu, V) = (2\mu + 1) \exp\Big(\frac{-\mu(\mu+1)}{4} V\Big)$$

and $LP_\mu(\cdot)$ denote the Legendre polynomials of integer order μ. A parameter V describes the total variance of the random walk and fully quantifies the distribution. In Section IV we show how it can be determined from the various properties of the transmission fiber. The distribution is normalized such that

$$\int_0^{2\pi} \int_0^{\pi} p(\phi, \theta, V) \, \sin\theta \, d\theta \, d\phi = 1. \tag{23}$$

Since $p(\phi, \theta, V)$ does not depend on the azimuth ϕ, it must be symmetric about the $\theta = 0$ axis. For $V = 0$, we have $p(\phi, \theta, 0) = \delta(\theta)$, i.e., the distribution is concentrated in the "pole", and for $V \to \infty$, the distribution approaches the uniform distribution on $\mathbb{S}^2$. The continuous motion of the SOPs described by (21) can be approximated by the stepwise rotations of Roberts and Ursell by letting the step size approach zero. However, because the vectors constituting the rotation axis do not change arbitrarily fast with propagation along z, many of these infinitesimal SOP rotations will be correlated. We will show that it is the nature of this correlation, and hence the statistical properties of the interfering channels, which determine the variance parameter V.

Also, the Brownian distribution (22) is the distribution in the limit of many independent, isotropic steps. If the length over which the individual rotations are significantly correlated is not much smaller than the system length (as e.g., in short systems or systems with resonant dispersion maps), the convergence of the true SOP distribution to the Brownian may be incomplete. However, due to the complicated expressions arising for the SOP distribution in such systems, we defer a more exact discussion and in the present article assume the validity of the Brownian distribution (22).

In order for the motion of the SOP "particles" to be *randomly directed* on $\mathbb{S}^2$, we need the rotation axis $\big(\mathbf{S}_\Sigma - \mathcal{E}\big[\mathbf{S}_\Sigma\big]\big)$ to be distributed symmetrically about the mean probe SOP everywhere. This is fulfilled for the practically relevant cases of parallel, pairwise-orthogonal, and randomly polarized launch, albeit only in the statistical sense of the *polarization ensemble*. Hence, the Brownian distribution (22) describes the SOP distribution in the sense of the ensemble of all systems with the same statistical properties.

In the context of PMD this ensemble is often defined as the collection of all possible combinations of the PMD parameters $\bar{\boldsymbol{\sigma}}(z)$ and $\Delta\beta_1(z)$ at each location in the fiber. Here, we extend the ensemble to include all possible initial interfering channel SOPs, as well as pseudo-random bit sequences, symbol timings, and relative phases – everything that might alter the evolution of $\mathbf{S}_\Sigma(z)$. The time duration of an ensemble member is the averaging period defined for $\mathcal{E}\big[\cdot\big]$, and to differentiate averages over all ensemble members from time averages, we denote the former by $\big\langle\cdot\big\rangle$.

We now want to relate the somewhat abstract and unmeasurable parameter V to a more accessible quantity, the degree of polarization (DOP). Since our probe is a CW channel, we can approximate the DOP very well by writing

$$\mathcal{E}\big[\hat{\mathbf{X}}(z,t)\big] \approx \mathrm{DOP}(z) \cdot \hat{\mathbf{X}}_0 \tag{24}$$

with $\hat{\mathbf{X}}_0 = \hat{\mathbf{X}}(0,t)$ (the precise expression for modulated channels uses the Stokes vectors instead of the SOPs). The mean probe polarization is fixed at $\hat{\mathbf{X}}_0$ everywhere because of (19). Multiplying both sides by $\hat{\mathbf{X}}_0$ yields

$$\mathrm{DOP}(z) = \mathcal{E}\big[\cos\Theta(z,t)\big] \tag{25}$$

with

$$\cos\Theta(z,t) \equiv \hat{\mathbf{X}}(z,t) \cdot \hat{\mathbf{X}}_0. \tag{26}$$

The random variable Θ corresponds to the displacement angle of the SOP $\hat{\mathbf{X}}$ from the initial SOP. Fig. 1 shows a typical Θ-distribution, which was obtained with numerical field simulations (cf. Table I for an overview of common system parameters used throughout) and is shown with a fit to the Brownian distribution.

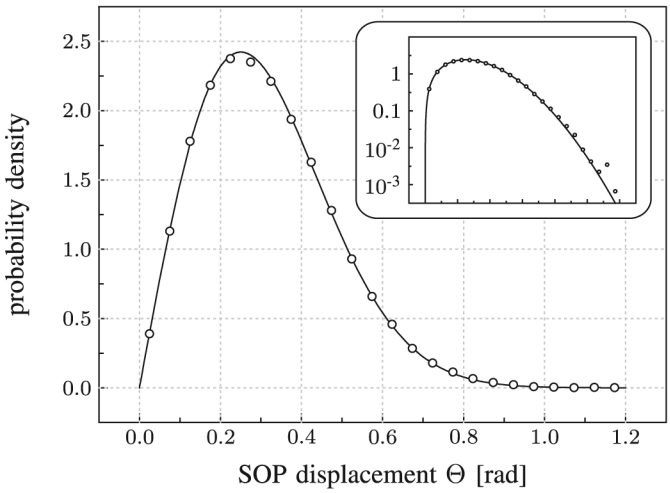

Fig. 1. The distribution of the angle Θ from (26) for 100 randomized simulation runs with 2^{15} samples each (symbols) and the Brownian distribution (lines). Both have an ensemble-averaged DOP of $\mathcal{D} \approx 0.94$. The system setup is shown in Fig. 2 and the system parameters are described in Table I; the dispersion map has a residual dispersion per span equivalent to 25 km transmission fiber. Inset shows logarithmic scale.

TABLE I
COMMON SYSTEM PARAMETERS USED IN SPLIT-STEP SIMULATIONS AND THEORETICAL PREDICTIONS (WHERE APPLICABLE) UNLESS NOTED OTHERWISE

bit rate	10 Gbps
sample rate	2.560 GSps
modulation format	NRZ
MUX filter bandwidth	25 GHz
PRBS order	12
channel spacing	50 GHz
number of channels N	11
average power per channel	4 mW
span length	100 km
GVD parameter	16 ps/nm/km
PMD parameter	0.5 ps/km$^{1/2}$
fiber attenuation α	0.2 dB/km
nonlinearity parameter γ	1.31 (W km)$^{-1}$
number of iterations	100

Since the Brownian (22) describes the distribution of the SOPs in the sense of the polarization ensemble, but (25) only describes the DOP of one member of the ensemble, we must average over the DOPs of all ensemble members, to be able to relate both parameters

$$\mathcal{D}(z) = \Big\langle \mathrm{DOP}(z) \Big\rangle. \tag{27}$$

Combining (27) with (25), we can now equate the expectation value of $\cos\theta$ from the Brownian distribution (22) with the global DOP, and obtain $\mathcal{D}$ in terms of the parameter V:

$$\begin{aligned}\mathcal{D}(z) &= \int_0^{2\pi}\int_0^{\pi} \cos\theta \; p(\phi,\theta,V) \; \sin\theta \; d\theta \; d\phi \\ &= \exp\left(-\frac{V(z)}{2}\right)\end{aligned} \tag{28}$$

where we made use of the orthogonality of the Legendre polynomials to arrive at the concise form of (28). Hence, there exists a bijective relationship between $\mathcal{D}$ and V – we can thus directly compare a physically measurable quantity obtainable in experiments and simulations to the analytical results which are obtained in terms of V. Exactly how to obtain the latter is the concern of the next section.

IV. Estimation of XPolM Magnitude

To relate the discrete theory of Roberts and Ursell to the continuous process described by (21), it becomes necessary to properly unify both theories and notations. This is done in the Appendix , which derives a relation between the parameter V and the variance of the Stokes vector sum (the rotation axis) $\mathbf{S}_\Sigma$ as used in (21):

$$V(z) = \frac{2}{3}\int_0^z\int_0^z \mathcal{C}_\Sigma(z_1,z_2)\, dz_1\, dz_2. \tag{29}$$

Herein, the function

$$\begin{aligned}\mathcal{C}_\Sigma(z_1,z_2) \equiv \Big\langle\, &\mathcal{E}\big[\mathbf{S}_\Sigma(z_1,t)\cdot\mathbf{S}_\Sigma(z_2,t)\big] \\ &- \mathcal{E}\big[\mathbf{S}_\Sigma(z_1,t)\big]\cdot\mathcal{E}\big[\mathbf{S}_\Sigma(z_2,t)\big]\Big\rangle\end{aligned} \tag{30}$$

is the ensemble-averaged autocovariance function (ACovF) of the random vector process $\mathbf{S}_\Sigma$.

This simple form of (29) requires that $\mathbf{S}_\Sigma$ itself be isotropic, and not only symmetric as required previously, as outlined in the Appendix . Strictly, it is therefore only valid when the DWDM channels are launched into the fiber with random polarizations from a uniform distribution on the Poincaré sphere, or have propagated sufficiently far – further than the polarization diffusion length [15] – to have lost memory of their initial polarization. For non-isotropic launch polarizations such as parallel or pairwise orthogonal, the factor 2/3 in (29) becomes a function of z_1 and z_2, which takes its asymptotic value of 2/3 only when z_1 and z_2 sufficiently exceed the polarization diffusion length [15].

To calculate V, we need to find the various stochastic properties of the Stokes vectors $\mathbf{S}_\nu$ contributing to the sum $\mathbf{S}_\Sigma$. To be able to do this in closed form, we must make some simplifying assumptions about the propagation of the interfering channels.

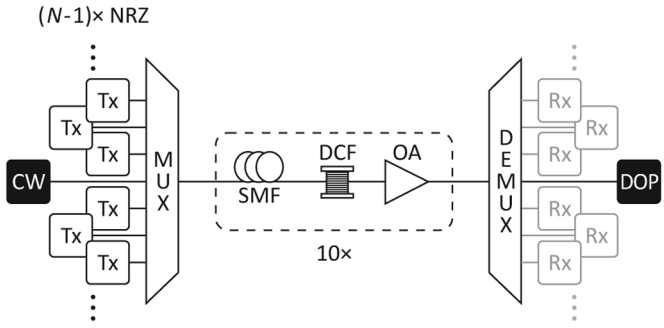

Fig. 2. A schematic of the optical transmission system with 100 km spans in which wavelength channels are arranged on a 50 GHz grid. Accumulated GVD is fully compensated at the end of each span, unless noted otherwise.

First of all, we shall neglect nonlinear distortions in the interfering channels, as the associated changes in the signals build up only over long distances, for which the spatial correlation of the Stokes sum is assumed to be negligible. We then write for the optical fields $\big|u_\nu\big\rangle$ in (11)

$$\big|u_\nu(z,t)\big\rangle = \bar{\mathbf{B}}(\Delta\omega_\nu, z)\; \exp\big[i\Delta\omega_\nu \Delta t_\nu(z)\big]\; \big|u_\nu(0,t)\big\rangle \tag{31}$$

where

$$\Delta t_\nu(z) = \int_0^z \beta_1(\zeta,\Delta\omega_\nu)d\zeta = \Delta\omega_\nu \int_0^z \beta_2(\zeta)d\zeta \tag{32}$$

and

$$\bar{\mathbf{B}}(\Delta\omega_\nu, z) = \exp\left[-i\Delta\omega_\nu \int_0^z \Delta\beta_1(\zeta)\, \widehat{\mathbf{\Omega}}(\zeta)\cdot\vec{\sigma}\, d\zeta\right] \tag{33}$$

is to be understood in a symbolic manner, similar to (7). We can rewrite (31) as

$$\big|u_\nu(z,t)\big\rangle = \bar{\mathbf{B}}(\Delta\omega_\nu, z)\big|u_\nu(0, t+\Delta t_\nu)\big\rangle. \tag{34}$$

We have integrated (10) at the channel carrier $\Delta\omega_\nu$ under the assumption of a signal envelope unaffected by either GVD or PMD. The exponential term in (31) describes a residual group delay which causes a time shift of the signal in the reference frame of channel ρ. The operator $\bar{\mathbf{B}}$ describes a residual birefringence at $\Delta\omega_\nu$ due to PMD. Essentially we argue that the influence of the inter-channel consequences of GVD and PMD is much greater than their intra-channel counterparts (leading to pulse distortion and depolarization).

To verify these assumptions for the class of systems of present interest, we performed numerical simulations of a typical optical transmission system as shown in Fig. 2. A CW probe channel is spectrally surrounded by a number of 10 Gbps NRZ-modulated interfering channels, and is propagated along a number of spans with full inline GVD compensation, in a worst-case scenario for inter-channel nonlinear effects [23]. The system of Fig. 2 will serve as our model system throughout the remainder of this article. Figs. 3 and 4 then show some numerically obtained examples of the ACovFs $\mathcal{C}_\nu(z_1 = 0, z_2)$ for a single interfering channel ν, in a single span and at the beginning of successive spans, respectively. Fig. 3 compares the effects of residual group

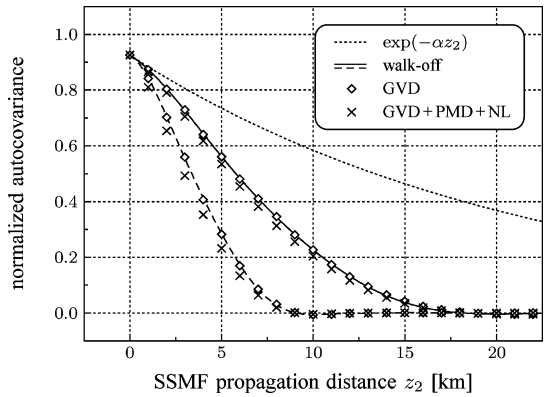

Fig. 3. This figure compares ACovFs $\mathcal{C}_\nu(0, z_2)$ of the Stokes vector of a single interfering channel, which were obtained numerically with different degrees of sophistication (see description in the text) for $\Delta\omega_\nu = 2\pi \cdot 50$ GHz (solid line and associated symbols) and $2\pi \cdot 100$ GHz (dashed line and associated symbols). The results for "GVD + PMD" are indistinguishable from "GVD + PMD + NL". All ACovFs were normalized to the mean optical power $\mathcal{E}[P_\nu]$. Because of the bandlimited, non-square pulse shape (cf. inset in Fig. 5), all ACovFs have an initial value of 0.925, the normalized pulse variance.

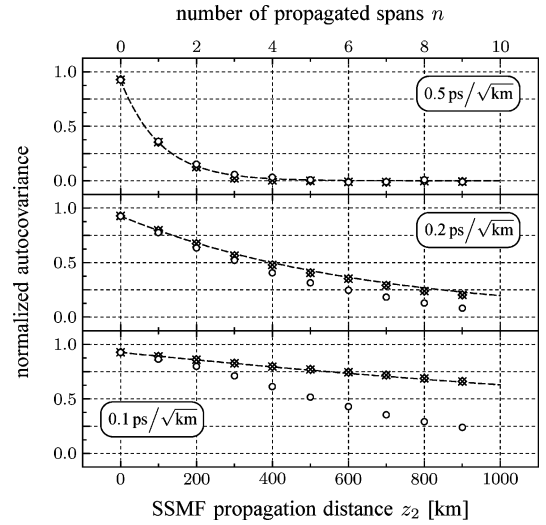

Fig. 4. This figure compares ACovFs $\mathcal{C}_\nu(0, z_2)$ of the Stokes vector of a single interfering channel at $\Delta\omega_\nu = 2\pi \cdot 50$ GHz for different fiber PMD parameters D_{PMD} at the beginning of each of 10 consecutive spans. Crosses and diamonds are averages from 1000 linear simulation runs – evaluating mean channel SOPs only (constant envelope assumption, crosses), and additionally including pulse distortion and depolarization due to PMD (diamonds); circles are average values of 500 nonlinear simulation runs at 6 dBm channel power. ACovFs are normalized to mean optical power $\mathcal{E}[P_\nu]$. Dashed lines additionally show the exponential decorrelation derived in Section VI .

delay ("walk-off") and GVD with pulse distortion ("GVD"), and additionally with PMD and nonlinearities. The exponential attenuation term from (35) is shown for reference. The results for "walk-off" were calculated by simply time-shifting the signals. All other data were obtained by split-step simulations. The plots in which PMD is included are averages of an ensemble of 100 randomly generated systems. The difference between the "walk-off" results and the full-vectorial field simulations is small enough to justify our approximations. For short single-span systems we may even disregard the contribution of residual birefringence as a consequence of PMD.

The systems in Fig. 4 used full (lumped) inline dispersion compensation to cancel the dominating influence of GVD at the end of each span. The span-to-span reduction of the ACovF values in the linear plots of this figure is due to uncompensated PMD alone, which can be split into the inter-channel residual birefringence and the intra-channel PMD-related pulse distortions/depolarization, which are shown as diamonds and crosses, respectively. Since there is virtually no difference between them, we are confirmed in our assumption (34) of a quasi-constant envelope. However, in multi-span systems, residual birefringence can no longer be disregarded. The plot for its analytical description, also shown in the figure, is derived in Section VI.

The circles in Fig. 4 mark the ACovF values when the interferers are allowed to evolve nonlinearly, using 6 dBm launch power per channel. The additional nonlinear ACovF decorrelation is due to nonlinear pulse distortion and XPolM, which in a nonlinear system affects not only to the probe channel but the interferers as well. Even though it can seem very dramatic when the residual birefringence is not large enough to dominate the ACovF, the configurations in which it becomes significant (e.g., high power at low fiber PMD) exhibit very large XPM- and XPolM-related distortions and are unlikely to be used for polarization-sensitive applications in practice. Using (34), we can write the Stokes vector $\widehat{\mathbf{S}}_\nu$ as the product of two stochastic quantities: the normalized optical power $P_\nu = \langle u_\nu | u_\nu \rangle = |\mathbf{U}_\nu|$ and the SOP $\widehat{\mathbf{S}}_\nu = \mathbf{U}_\nu / P_\nu$. We then have

$$\mathbf{S}_\nu(z,t) = \bar{\gamma} \exp\big[-\alpha z\big] P_\nu(t + \Delta t_\nu)\, \widehat{\mathbf{S}}_\nu(z). \tag{35}$$

The P_ν are determined by data modulation with pseudo-random bit sequences. Propagation along z introduces a time shift only, but no waveform distortion. The SOPs $\widehat{\mathbf{S}}_\nu$ evolve as a result of the random residual birefringence (33) in multi-span systems. Thus, both generating processes (modulation/GVD and PMD) are stochastically independent of one another and can be separated as above.

Inserting (16) in the autocovariance function (30) results in an expression with $(N-1)^2$ terms, where N is the total number of channels. However, because the data modulation in different channels is mutually uncorrelated, all cross-channel terms vanish when taking the expectation and we can separate the contributions of the individual channels:

$$\mathcal{C}_\Sigma(z_1, z_2) = \sum_\nu \mathcal{C}_\nu(z_1, z_2) \tag{36}$$

in which the single-channel ACovFs are defined analogously to (30)

$$\mathcal{C}_\nu(z_1, z_2) \equiv \Big\langle \mathcal{E}\big[\mathbf{S}_\nu(z_1,t) \cdot \mathbf{S}_\nu(z_2,t)\big] - \mathcal{E}\big[\mathbf{S}_\nu(z_1,t)\big] \cdot \mathcal{E}\big[\mathbf{S}_\nu(z_2,t)\big] \Big\rangle. \tag{37}$$

We have reduced the task of finding the stochastic properties of a vector sum to one of finding the properties of single-channel data modulation and residual birefringence-related SOP evolution. We will begin by examining short, single-span systems in which we may neglect the latter.

V. Single Fiber Spans

In very short systems or systems comprising a single, unamplified span of fiber, nonlinear distortions occur mainly within the effective length

$$L_{\text{eff}} = \frac{1-\exp(\alpha L)}{\alpha} \approx \frac{1}{\alpha} \tag{38}$$

of a span of length L, for which typical values are on the order of 20 km. Within this short length, polarization effects generally do not play a role (as was shown in Fig. 3) unless the fiber has an exceptionally high PMD coefficient. We can thus assume $\bar{\mathbf{B}} = \bar{\mathbf{I}}$, the identity matrix, and the SOPs $\widehat{\mathbf{S}}_\nu(z)$ of the interfering channels in (35) to be constant, such that $\widehat{\mathbf{S}}_\nu(z_1) \cdot \widehat{\mathbf{S}}_\nu(z_2) = 1$, and upon inserting (35) in (37), we obtain the simplified ACovF

$$\mathcal{C}_\nu(z_1, z_2) \approx \bar{\gamma}^2 \exp\Big[-\alpha\big(z_1 + z_2\big)\Big] \cdot \mathcal{C}_\nu^{\text{WO}}(z_1, z_2) \tag{39}$$

with

$$\mathcal{C}_\nu^{\text{WO}}(z_1, z_2) = \mathcal{E}\big[P_\nu(t)\, P_\nu(t + \Delta t_\nu)\big] - \mathcal{E}\big[P_\nu\big]^2 \tag{40}$$

in which $\Delta t_\nu = \Delta t_\nu(z_1, z_2)$ is the group delay accumulated between z_1 and z_2. Its definition is given in (32) with the integral limits replaced by z_1 and z_2. The label "WO" denotes that this simple ACovF is determined solely by the *walk-off* between the probe ρ and the interferer ν. In writing (40), we have used the property that the P_ν are independent of the polarization properties to replace the ensemble averages $\langle\cdot\rangle$ by their arguments.

Because the data in each channel is statistically random, this ACovF must per definition become zero when the accumulated time shift equals or exceeds the symbol duration (in the absence of pulse overlap) for simple NRZ modulation. The ACovF $\mathcal{C}_\nu^{\text{WO}}$ of an NRZ pulse train can be well approximated by a simple expression valid for rectangular pulses. Inserting this pulse shape into (40) yields a closed-form solution which linearly decreases from 1 to 0:

$$\mathcal{C}_\nu^{\text{WO}} \equiv \mathcal{E}\big[P_\nu\big] \begin{cases} 1 - \frac{|\Delta z|}{L_{\text{WO}\nu}} & |\Delta z| < L_{\text{WO}\nu} \\ 0 & |\Delta z| \geq L_{\text{WO}\nu} \end{cases} \tag{41}$$

in which $L_{\text{WO}\nu}$ is the characteristic *walk-off length* for channel ν

$$L_{\text{WO}\nu} = \frac{T_S}{|\beta_2|\, \Delta\omega_\nu} \tag{42}$$

at which Δt_ν in (32) equals the symbol duration $\pm T_S$. This simple relation can greatly decrease the time required to calculate the integral (29) and has been used throughout the remainder of this work to analytically determine the depolarization in NRZ systems.

By using L_{WO}, we can formulate the $\mathcal{C}_\nu^{\text{WO}}$ independently of frequency separation and symbol rate. The true shape of the function $\mathcal{C}_\nu^{\text{WO}}$ will depend on the shape of the pulses used to modulate the data stream onto the respective carriers. Fig. 5 compares numerically obtained ACovFs for several popular modulation formats. The functions were generated by time-shifting a sample pseudo-random bit sequence sequence of the respective formats and then calculating (40).

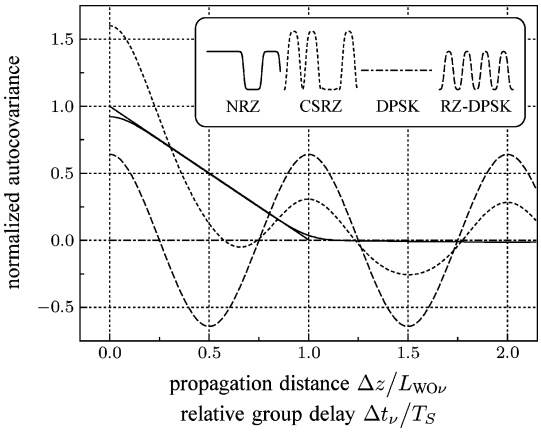

Fig. 5. The numerically obtained ACovF $\mathcal{C}_\nu^{\text{WO}}(z, z + \Delta z)$ related to the walk-off between two wavelength channels for different modulation formats. The solid grey line represents the square NRZ approximation of (41). All signals are normalized to $\mathcal{E}[P_\nu]$; inset shows corresponing signal shapes for a 1101 sequence.

The ACovF of the constant-power differential phase-shift keying (DPSK) pulse train (as obtained with a phase modulator) is identically zero in the constant envelope approximation. This means that DPSK-modulated channels interfere significantly less than NRZ signals – usually only after the signal has become severely distorted, which occurs mostly when the signal power is already attenuated considerably at the end of the span. The periodic pulse shape of RZ-DPSK results in an oscillating ACovF which has a net integral of nearly zero. Thus, for systems which are at least several walk-off lengths long, RZ-DPSK interfering channels will also result in less XPolM than NRZ. Finally, the ACovF of carrier-suppressed RZ (CSRZ) – much like the pulse shape itself – is a combination of the NRZ and RZ-DPSK ACovFs, with a net integral greater than zero. Like all amplitude-modulated formats, it will contribute significantly to XPolM.

For arbitrary pulse shapes, such as signals predistorted to compensate for GVD, we can determine an ACovF matrix numerically by propagating a single channel linearly, regarding only chromatic dispersion (of arbitrary order). This also allows for more precise determination of XPolM in DPSK systems or systems with large accumulated GVD by accounting for the (intra-channel) pulse shape distortions that occur due to GVD. Similarly, pulse shape distortions due to PMD may be accounted for with reasonable effort by performing multiple runs with random PMD and averaging the results, as done in Fig. 3. However, such means are mainly relevant when the walk-off-related ACovF, and thus the nonlinear depolarization, is negligibly small. The ACovF (40) can now be inserted into (39) for each interfering channel. These are then summed over as in (36) and integrated (29) to obtain V, from which we can determine the depolarization according to (28).

By inspecting (40), we realize that $\mathcal{C}_\nu^{\text{WO}}$ scales with the square of the optical power. This power-square dependence is then also true (if the optical power is equally increased in all interfering channels) for the total ACovF $\mathcal{C}_\Sigma$ and thus the variance parameter V, from which the depolarization is

determined. This is generally valid also in multi-span systems, which are examined next.

VI. Multiple Fiber Spans

In multi-span, optically in-line compensated transmission systems, we extend the attenuation-scaled, nonlinear coefficient in (35) to include amplifier gain $g(z)$, and define the power-weighted nonlinearity

$$\Gamma(z) \equiv \bar{\gamma} \exp\left[\int_0^z \Big[g(\zeta) - \alpha(\zeta)\Big]\, d\zeta\right] \tag{43}$$

where we assume α and g to be independent of channel index ν. Should this assumption not hold (e.g., for a non-flat EDFA gain spectrum), one may define a separate $\Gamma_\nu(z)$ for each DWDM channel.

Also, the chromatic dispersion accumulated within a fiber span is usually fully or partially compensated at the end of each span. Hence, the ACovF $\mathcal{C}_\nu^{\mathrm{WO}}(z_1, z_2)$ can remain comparatively large even when z_1 and z_2 describe locations in different fiber spans. We can then no longer neglect the change of the polarization states $\widehat{\mathbf{S}}_\nu(z)$ between z_1 and z_2 due to the residual birefringence when calculating the autocovariance functions $\mathcal{C}_\nu(z_1, z_2)$ and $\mathcal{C}_\Sigma(z_1, z_2)$ of the Stokes vectors. We thus insert the unabridged (35) in (37) to obtain

$$\mathcal{C}_\nu(z_1, z_2) = \Gamma(z_1) \cdot \Gamma(z_2) \cdot \mathcal{C}_\nu^{\mathrm{WO}}(z_1, z_2) \cdot \mathcal{C}_\nu^{\mathrm{SOP}}(z_1, z_2) \tag{44}$$

with

$$\mathcal{C}_\nu^{\mathrm{SOP}}(z_1, z_2) \equiv \Big\langle \widehat{\mathbf{S}}_\nu(z_1) \cdot \widehat{\mathbf{S}}_\nu(z_2) \Big\rangle \tag{45}$$

where we have made use of the statistical independence of P_ν and $\widehat{\mathbf{S}}_\nu$.

The SOP autocorrelation function (ACF) $\mathcal{C}_\nu^{\mathrm{SOP}}$ describes the statistical correlation between the SOP of channel ν at z_1 and z_2. Formally, the reason for a decorrelation is the residual birefringence term $\bar{\mathbf{B}}$ in (34) – it describes a series of rotations around the spatially varying local fiber birefringence vectors $\widehat{\mathbf{\Omega}}$. A characteristic length describing the distance over which the birefringence orientation $\widehat{\mathbf{\Omega}}$ varies is the fiber autocorrelation length, "over which an ensemble of fibers, all of which initially have the same orientation of the axes of birefringence, loses memory of this initial orientation" [15]. This length is usually on the order of only several hundred meters. The SOP $\widehat{\mathbf{S}}_\nu$ thus undergoes a large number of random linear rotations which lead to a certain statistical decorrelation within Δz in the ensemble average.

Karlsson has given an expression from which the ACovF $\mathcal{C}_\nu^{\mathrm{SOP}}$ can in principle be derived [24], and Wai and Menyuk have also derived a corresponding expression by means of stochastic differential equations [15]. We may also obtain it as a limiting case of the analysis in terms of spherical harmonics performed by Ueda and Kath [25]. However, since we have already established a framework to deal with random rotations on the unit sphere within the previous sections, we will show how to apply our model to obtain the SOP decorrelation ACF. First, we identify $\widehat{\mathbf{S}}_\nu(z_1)$ from (45) with the initial SOP $\widehat{\mathbf{X}}_0$ in (24) and $\widehat{\mathbf{S}}_\nu(z_2)$ with the final SOP $\widehat{\mathbf{X}}(z, t)$. Because $\widehat{\mathbf{S}}_\nu(z_2)$ evolves from $\widehat{\mathbf{S}}_\nu(z_1)$ with a large number of random rotations, we deduce that $\widehat{\mathbf{S}}_\nu(z_2)$ will be distributed according to the Brownian distribution (22) centered on $\widehat{\mathbf{S}}_\nu(z_1)$. The expression for the depolarization $\mathcal{D}$ in (27) then corresponds exactly to the ACF $\mathcal{C}_\nu^{\mathrm{SOP}}$ as defined in (45), and we can assign a variance parameter V_ν^{SOP} to this distribution, so that

$$\mathcal{C}_\nu^{\mathrm{SOP}}(z_1, z_2) = \exp\left(-\frac{V_\nu^{\mathrm{SOP}}(z_1, z_2)}{2}\right) \tag{46}$$

analogous to the diffusion of the SOPs $\widehat{\mathbf{X}}$ on the Poincaré sphere due to XPolM in (28). We can determine V_ν^{SOP} from the integral (29) by finding an appropriate expression for $\mathcal{C}_\Sigma$, the ACovF of the random rotation axis. By transforming the PMD-related term of the Manakov-PMD (10) into Stokes space and frequency domain, we obtain the evolution equation for the SOP $\widehat{\mathbf{S}}_\nu(z)$

$$\partial_z \widehat{\mathbf{S}}_\nu(z) = \Delta\omega_\nu\, \Delta\beta_1(z) \widehat{\mathbf{\Omega}}(z) \times \widehat{\mathbf{S}}_\nu(z). \tag{47}$$

This identifies $\Delta\omega_\nu\, \Delta\beta_1 \widehat{\mathbf{\Omega}}$ as the random rotation axis, and we have

$$V_\nu^{\mathrm{SOP}}(z_1, z_2) = \frac{2}{3} \int_{z_1}^{z_2} \int_{z_1}^{z_2} \mathcal{C}_\nu^{\Omega}(\zeta_1, \zeta_2)\, d\zeta_1\, d\zeta_2 \tag{48}$$

with

$$\mathcal{C}_\nu^{\Omega}(\zeta_1, \zeta_2) \equiv \Delta\omega_\nu^2 \Big\langle \Delta\beta_1(\zeta_1) \widehat{\mathbf{\Omega}}(\zeta_1) \cdot \Delta\beta_1(\zeta_2) \widehat{\mathbf{\Omega}}(\zeta_2) \Big\rangle. \tag{49}$$

In [26], Gordon identifies the term in angle brackets with $R\, \delta(\zeta_1 - \zeta_2)$, provided that the fiber length is much larger than the fiber autocorrelation length, where R is a diffusion constant and $\delta(\cdot)$ is the Dirac delta function. Inserting (49) into (48), we then obtain

$$\begin{aligned} V_\nu^{\mathrm{SOP}}(z_1, z_2) &= \frac{2}{3} \int_{z_1}^{z_2} \int_{z_1}^{z_2} \Delta\omega_\nu^2\, R\, \delta(\zeta_1 - \zeta_2)\, d\zeta_1\, d\zeta_2 \\ &= \frac{2}{3} \Delta\omega_\nu^2 \Big\langle \Delta\tau^2(z_1, z_2) \Big\rangle \end{aligned} \tag{50}$$

in which [26]

$$\Big\langle \Delta\tau^2(z_1, z_2) \Big\rangle = R \big| z_1 - z_2 \big| \tag{51}$$

is the mean-square differential group delay (DGD) of the fiber segment $\big[z_1, z_2\big]$. We further have [27]

$$\begin{aligned} \Big\langle \Delta\tau^2(z_1, z_2) \Big\rangle &= \frac{3\pi}{8} \Big\langle \Delta\tau(z_1, z_2) \Big\rangle^2 \\ &= \frac{3\pi}{8}\, D_{\mathrm{PMD}}^2 \big| z_2 - z_1 \big| \end{aligned} \tag{52}$$

in which D_{PMD} is the more common fiber PMD parameter. Defining an SOP decorrelation length of channel ν

$$L_{\mathrm{SOP}\nu} \equiv \frac{8}{\pi \Delta\omega_\nu^2 D_{\mathrm{PMD}}^2} \tag{53}$$

the SOP variance parameter V_ν^{SOP} of channel ν can be expressed as

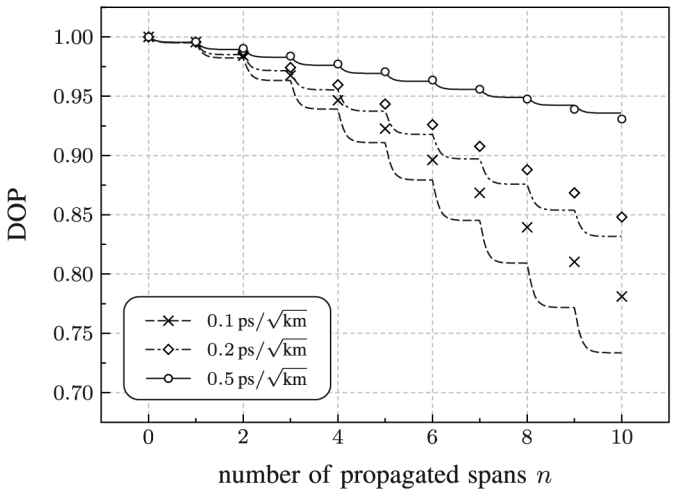

Fig. 6. DOP vs. number of propagated spans for different fiber PMD values for the optical system of Fig. 2 and Table I; symbols denote simulation results and lines denote analytical predictions.

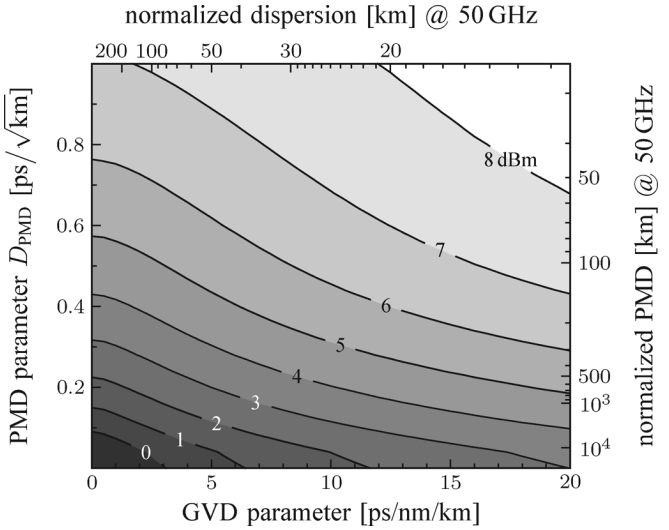

Fig. 7. Nonlinear threshold for 10 Gbps NRZ interfering channels to obtain $\mathcal{D} = 0.90$ in the probe channel as a function of (normalized) dispersion length $T_S/(|\beta_2|\,\Delta\omega)$ and PMD length $8/(\pi\Delta\omega^2 D_{\mathrm{PMD}}^2)$, corresponding to the walk-off and SOP decorrelation lengths for neighboring channels (cf. (42) and (53) in the text), respectively; bottom and left axes denote the related fiber GVD and PMD parameters for channel spacing $\Delta\omega = 2\pi \cdot 50$ GHz and symbol duration $T_S = 100$ ps.

$$V^{\mathrm{SOP}}(z_1, z_2) = 2\,\frac{|z_2 - z_1|}{L_{\mathrm{SOP}\nu}} \tag{54}$$

and we obtain the ACovF $\mathcal{C}_\nu^{\mathrm{SOP}}$ resulting from PMD with the help of (46) as

$$\mathcal{C}_\nu^{\mathrm{SOP}}(z_1, z_2) = \exp\left(-\frac{|z_2 - z_1|}{L_{\mathrm{SOP}\nu}}\right). \tag{55}$$

Using $\mathcal{C}_\nu^{\mathrm{SOP}}$ from (44) should give a good approximation for the ACovF in multi-span systems which can be used to estimate the nonlinear V parameter and the DOP reduction for arbitrary fiber systems.

To test the preceding theory, we compared its predictions to numerical simulation results for a number of different systems (cf. Fig. 2 and Table I). The dashed lines in Fig. 4 show $\mathcal{C}_\nu^{\mathrm{SOP}}(0, z_2)$ for the respective systems, which closely match the linear simulation results, as expected. Fig. 6 shows the corresponding evolution of the probe channel DOP reduction with the number of propagated spans. As expected, XPolM is more pronounced in low-PMD systems due to the slower-declining ACovF for these fibers, leading to larger values of V. The discrepancies in Fig. 6 result mainly from assuming linear propagation of the interfering channels. As mentioned before, configurations in which the difference between theory and simulation becomes significant are unlikely to be met in practical XPolM-sensitive applications, due to the very large SOP distortions described by the correspondingly low DOPs, accompanied by significant phase distortion due to XPM.

VII. Application

We now use our model to estimate the relative impact of XPolM as it depends on various parameters of the transmission fiber. Fig. 7 shows the nonlinear threshold (NLT), defined as the maximum fiber input power of the interfering channels to obtain an average DOP of 0.90 in the probe channel, after 10 spans in a resonant dispersion map versus the GVD and the PMD parameters. As we have noticed previously, V increases with the square of the mean power $\mathcal{E}[P_\nu]$ of the interfering channels. Thus, assuming the mean power to be the same for all channels, the threshold powers given in Fig. 7 (in linear units) can be easily converted to any other DOP value $\mathcal{D}$ by making use of relation (28) to obtain

$$\mathrm{NLT}(\mathcal{D}) = \sqrt{\frac{\ln \mathcal{D}}{\ln 0.90}}\,\mathrm{NLT}(0.90) \tag{56}$$

where $\ln \cdot$ is the natural logarithm and NLT(0.90) is the nonlinear threshold for a mean DOP of 0.90 from Fig. 7. Hence, the NLT values for $\mathcal{D} = 0.97$ will be nearly 3 dB lower than those plotted in the figure. As can also be seen, low-PMD NZDSFs are affected significantly more than older, high-PMD SSMFs and thus have NLTs which are lower by about 5 dB.

From Fig. 7 we also recognize that the NLTs for those configurations of Fig. 4 in which the nonlinear influence on the ACovF is significant, namely low fiber PMD, are sufficiently low to largely remove that influence, improving the analytical estimate.

From the way the walk-off related ACovFs $\mathcal{C}_\nu^{\mathrm{WO}}$ in Fig. 5 depend on the accumulated time shift, we can directly deduce a method to decrease the impact of XPolM in amplitude-modulated multi-span systems: making the average accumulated time shift Δt_ν larger will decrease the corresponding $\mathcal{C}_\nu^{\mathrm{WO}}(z_1, z_2)$. Therefore, avoiding resonant dispersion maps — such that the time shift is zero whenever z_1 and z_2 correspond to fiber locations at the beginning of different spans, where the optical power is high — by leaving some GVD uncompensated in each span (residual dispersion per span, RDPS) can reduce V significantly. As a rule-of-thumb, the RDPS should correspond to about one walk-off length of the interfering channels closest

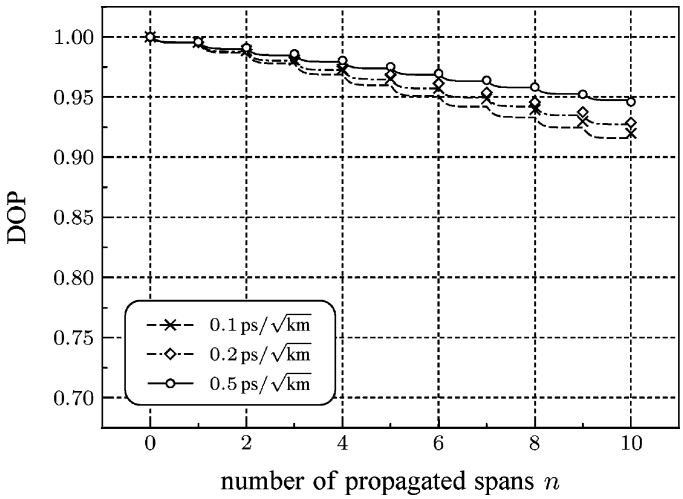

Fig. 8. DOP vs. number of propagated spans for different fiber PMD values; setup is the same as in Fig. 6, but with residual dispersion per span of 400 ps/nm; symbols denote simulation results and lines denote analytical predictions.

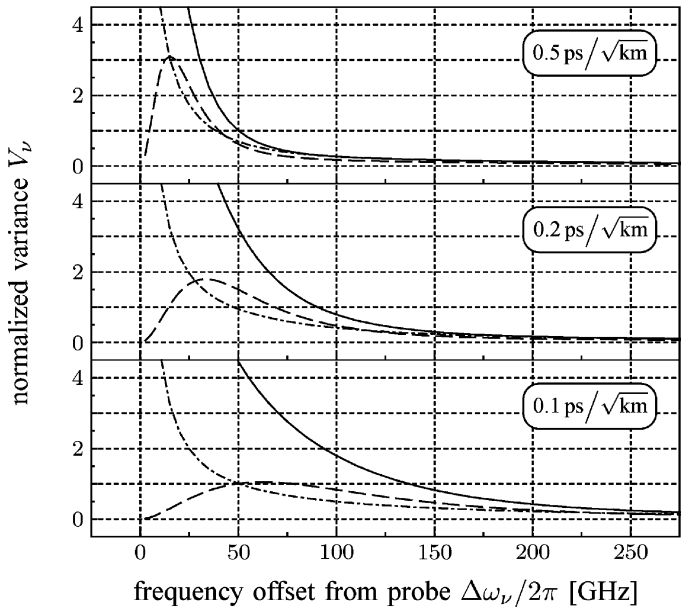

Fig. 9. Relative variance contribution of 10 Gbps NRZ interfering channels after 10 spans, depending on their center frequency separation from the probe for various PMD parameters: resonant dispersion map/random launch polarizations (solid lines), residual dispersion per span of 400 ps/nm/random launch (dash-dotted lines), and resonant map/(anti-)parallel launch (dashed lines); V values are normalized to 0.5 ps/$\sqrt{\text{km}}$, resonant map/random launch at 50 GHz.

to the probe, $L_{\mathrm{WO}(\rho\pm1)}$, or more. This fully decorrelates the Stokes vectors at the beginning of each span and thus reduces the accumulation of variance over several spans. Fig. 8 shows the improvement obtained for the systems of Fig. 6 when leaving an RDPS corresponding to $1.6\ L_{\mathrm{WO}(\rho\pm1)}$. As can be seen in the figure, the decorrelation between spans introduced by RDPS can significantly suppress the XPolM degradation, especially in low-PMD systems in which the ACovF is normally correlated over many spans. Another benefit of the analytical model is an estimation of the relative impact of XPolM caused by the various interfering channels. This impact is compared in Fig. 9 in terms of the parameter V_ν for a single 10 Gbps NRZ interferer, depending on the frequency separation $\Delta\omega_\nu$ between this interferer and the probe. To obtain these curves we made use of the ACovFs (41) and (55). One can see that the immediately neighboring channels will have the largest impact on the total V and that spectrally distant channels will have a comparatively small additional impact. This is shown in a different manner in Fig. 10, which displays the DOP reduction as a function of the number of interferers (filling the spectral slots form the probe outward) as obtained with our theoretical model. We can observe how the impact of XPolM increases quickly for the first few neighbors and then starts to saturate. It would be difficult and time-consuming to obtain similar results with numerical simulations due to the enormous simulation bandwidth required.

From Fig. 9 we can see how the XPolM impact increases with lower fiber PMD, and how RDPS almost equalizes the impact of all fiber types. Also shown is an analytical estimate of the benefit of launching neighboring channels with orthogonal or parallel SOPs – corresponding to (anti-)parallel Stokes vectors. Such launch constellations can suppress spectrally close interferers, especially when fiber PMD is very low, but they become less effective in very long systems in which the launch conditions become less relevant. In systems employing reconfigurable optical add-drop multiplexers (ROADMs), it is also not yet possible to adjust the polarization state of newly added channels to match that of those channels passing through, further reducing any benefit of initial relative launch polarization states.

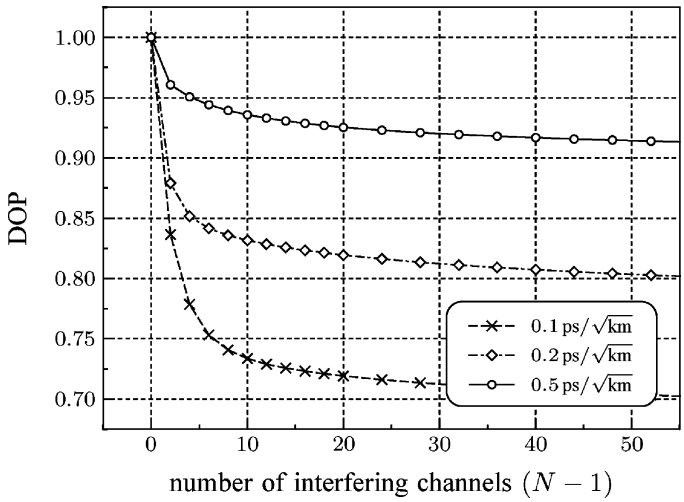

Fig. 10. DOP vs. number of interfering channels for different fiber PMD values for the optical system of Fig. 2 and Table I, obtained analytically.

VIII. Summary and Discussion

In the present contribution we derived the statistical behavior of a probe channel under the influence of cross-polarization modulation from a large number of interfering channels under almost arbitrary conditions. We have made extensive use of the methods of probability theory due to the exceedingly large number of states such a system can be in. Our results were therefore given in terms of probabilities for the polarization properties of the probe signal after traversing an optical link – most importantly the distribution of the polarization states of the probe, in the sense of a distribution average over the

ensemble of polarization parameters (launch polarization, birefringence and PMD properties of the fiber) and parameterized by the (nonlinear) degree of polarization reduction.

We believe that the expressions for the magnitude of XPolM degradation, as described by the nonlinear DOP reduction, are sufficiently straightforward to allow considerable insight into the nonlinear process and its dependence on various system parameters. The integrals in the expressions can be solved numerically sufficiently fast to allow evaluation of a large number of different systems within a reasonable amount of time. We have also demonstrated how to determine the nonlinear threshold power corresponding to a given required minimum DOP value, and how leaving some residual dispersion per span can significantly reduce XPolM degradations in low-PMD fibers.

The knowledge of the distribution of polarization states described by the DOP may subsequently be applied to study error probabilities in polarization-multiplex or polarization-sensitive coherent systems, or consider the usefulness of the DOP as a feedback signal for PMD compensation. Research in this direction will also yield bounds on the permissible nonlinear depolarization, from which realistic limits on the launch power due to XPolM may be derived. The model underlying our calculations may also be applied to the statistics of the PMD vector of a fiber link, which undergoes these nonlinearity-induced rotations as well.

Finally, the present work should be regarded as a theoretical basis for a more thorough investigation of XPolM. While we have worked exclusively with ensemble averages to be able to obtain closed-form solutions, the detailed ensemble statistics will be even more interesting for system operators. As with PMD, depending on their polarization parameters some systems within such an ensemble will be significantly more affected than others, and (outage) probabilities of exceeding a given threshold (e.g., a minimum required DOP) will be necessary to give more meaningful results.

APPENDIX

Roberts and Ursell define V in their distribution as the variance of a corresponding planar random walk. Because they assume all individual rotations to be uncorrelated and randomly directed, this is straightforward – angular displacements Ψ_k on a unit sphere are mapped to step vectors $\mathbf{W}_k$ with $|\mathbf{W}_k| = \Psi_k$ of a planar walk on $\mathbb{R}^2$, and we have

$$V(K) = \sum_{k=1}^{K} \left\langle \left|\mathbf{W}_k\right|^2 \right\rangle \tag{57}$$

for the variance after K steps, where $\langle\cdot\rangle$ is the average over all possible step sizes. If we allow for correlation between steps k and l, we have to extend the expression with the corresponding joint moments,

$$\begin{aligned} V(K) &= \sum_{k=1}^{K} \left\langle \left|\mathbf{W}_k\right|^2 \right\rangle + \sum_{k=1}^{K} \sum_{\substack{l=1 \\ l\neq k}}^{K} \left\langle \mathbf{W}_k \cdot \mathbf{W}_l \right\rangle \\ &= \sum_{k=1}^{K} \sum_{l=1}^{K} \left\langle \mathbf{W}_k \cdot \mathbf{W}_l \right\rangle. \end{aligned} \tag{58}$$

Making the transition to a continuous process over z, i.e., the limit of infinitesimal step size,

$$V(L) = \int_0^L \int_0^L \left\langle d\mathbf{W}(z_1) \cdot d\mathbf{W}(z_2) \right\rangle \tag{59}$$

in which $d\mathbf{W}$ is a differential increment.

We now relate the increment $d\mathbf{W}(z)$ of the planar diffusion to the increment $\partial_z \widehat{\mathbf{X}}(z,t)$ of the motion of a particular probe channel SOP indexed by t, which was given in (21) as

$$\partial_z \widehat{\mathbf{X}}(z,t) = \mathbf{Z}(z,t) \times \widehat{\mathbf{X}}(z,t) \tag{60}$$

with

$$\mathbf{Z}(z,t) = \mathbf{S}_\Sigma(z,t) - \mathcal{E}\left[\mathbf{S}_\Sigma(z,t)\right]. \tag{61}$$

However, $\partial_z \widehat{\mathbf{X}}(z,t)$ is a three-dimensional vector which may not be coplanar for different z and thus does not constitute a planar diffusion (it is in fact the diffusion of $\widehat{\mathbf{X}}$ on the unit sphere that we would like to describe). We can force the motion described by $\partial_z \widehat{\mathbf{X}}$ to be planar by assuming $\widehat{\mathbf{X}}$ in (60) to be constant, $\widehat{\mathbf{X}}(z,t) = \widehat{\mathbf{X}}(0,t) = \widehat{\mathbf{X}}_0$. Our estimate for V will thus become increasingly inaccurate as $\widehat{\mathbf{X}}(z,t)$ deviates considerably from $\widehat{\mathbf{X}}_0$ while $d\mathbf{W}(z_1) = \partial_z \widehat{\mathbf{X}}(z_1,t)$ and $d\mathbf{W}(z_2) = \partial_z \widehat{\mathbf{X}}(z_2,t)$ in (59) are still highly correlated and thus contribute significantly to the integral. For the same reasons as those mentioned in the text, such cases are not practically relevant.

We arbitrarily set $\widehat{\mathbf{X}}_0 = \mathbf{e}_1$, a basis vector of the three-dimensional coordinate system, so that (60) becomes

$$\partial_z \widehat{\mathbf{X}}(z,t) = Z_3(z,t)\, \mathbf{e}_2 - Z_2(z,t)\, \mathbf{e}_3 \tag{62}$$

which depends only on those components of $\mathbf{Z}$ which are orthogonal to $\widehat{\mathbf{X}}_0$. The coordinate system is illustrated in Fig. 11. Inserting (62) into (59) yields

$$\begin{aligned} V(L) = \int_0^L \int_0^L & \mathbf{e}_2 \cdot \mathbf{e}_2 \left\langle \mathcal{E}\left[Z_3(z_1,t)\, Z_3(z_2,t)\right] \right\rangle \\ & -2\mathbf{e}_2 \cdot \mathbf{e}_3 \left\langle \mathcal{E}\left[Z_3(z_1,t)\, Z_2(z_2,t)\right] \right\rangle \\ & +\mathbf{e}_3 \cdot \mathbf{e}_3 \left\langle \mathcal{E}\left[Z_2(z_1,t)\, Z_2(z_2,t)\right] \right\rangle dz_1\, dz_2. \end{aligned} \tag{63}$$

Because $\partial_z \widehat{\mathbf{X}}(z,t)$ describes a separate diffusion process for each t, we must additionally introduce the time average $\mathcal{E}\left[\cdot\right]$ to capture all possible such processes.

Using $\mathbf{e}_2 \cdot \mathbf{e}_2 = \mathbf{e}_3 \cdot \mathbf{e}_3 = 1$, $\mathbf{e}_2 \cdot \mathbf{e}_3 = 0$, and (61) in (63), we obtain

$$V(L) \equiv \int_0^L \int_0^L \mathcal{C}_{\Sigma 2}(z_1,z_2) + \mathcal{C}_{\Sigma 3}(z_1,z_2)\, dz_1\, dz_2 \tag{64}$$

where $\mathcal{C}_{\Sigma n}$ is the autocovariance function of the n-th component of $\mathbf{S}_\Sigma$, defined as

$$\begin{aligned} \mathcal{C}_{\Sigma n}(z_1,z_2) = \Big\langle \mathcal{E}\Big[& \Big(S_{\Sigma n}(z_1,t) - \mathcal{E}\left[S_{\Sigma n}(z_1,t)\right]\Big) \\ & \times \Big(S_{\Sigma n}(z_2,t) - \mathcal{E}\left[S_{\Sigma n}(z_2,t)\right]\Big)\Big] \Big\rangle. \end{aligned} \tag{65}$$

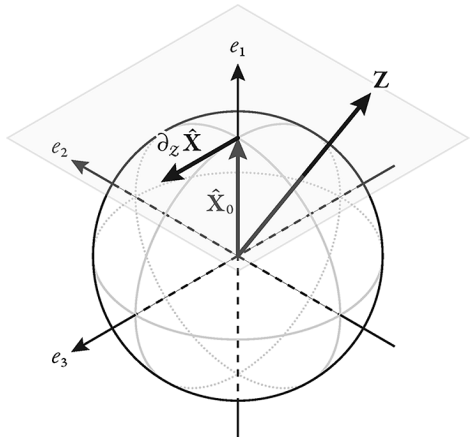

Fig. 11. illustrates the coordinate system, the initial probe SOP $\hat{\mathbf{X}}_0$, the rotation axis $\mathbf{Z} = \mathbf{S}_\Sigma - \mathcal{E}[\mathbf{S}_\Sigma]$, and the infinitesimal increment $\partial_z \hat{\mathbf{X}}$ to the diffusion of $\hat{\mathbf{X}}$.

We will finish with some remarks on the distribution of the rotation axis $\mathbf{Z}$. In order for the $\partial_z \widehat{\mathbf{X}}$ in (60) to be randomly directed in the plane orthogonal to $\widehat{\mathbf{X}}_0$, $\mathbf{Z}$ must be symmetric about $\widehat{\mathbf{X}}_0$. This is of course satisfied when the initial orientations of the Stokes vectors constituting the sum $\mathbf{S}_\Sigma$ are distributed isotropically on the Poincaré sphere. It is also satisfied if all Stokes vectors are initially either parallel or anti-parallel to the probe SOP $\widehat{\mathbf{X}}_0$, because the analysis in Section VI showed that the linear polarization evolution due to PMD is symmetric with respect to the initial SOP. After propagation of some SOP decorrelation lengths, the distribution of all interferer $\mathbf{S}_\nu$ will eventually become isotropic. For all other cases, the simple relation (59) does not hold and V must be calculated regarding the distribution of the initial SOPs.

In the case that the distribution of $\mathbf{S}_\Sigma$ is isotropic then its component autocovariance function cannot depend on the particular component n, and we must have

$$\mathcal{C}_{\Sigma 1} = \mathcal{C}_{\Sigma 2} = \mathcal{C}_{\Sigma 3} = \frac{1}{3}\mathcal{C}_\Sigma \tag{66}$$

in which $\mathcal{C}_\Sigma = \mathcal{C}_{\Sigma 1} + \mathcal{C}_{\Sigma 2} + \mathcal{C}_{\Sigma 3}$ is the vector autocovariance of $\mathbf{S}_\Sigma$, defined analogous to (65). Hence we rewrite (64):

$$V(L) = \frac{2}{3}\int_0^L \int_0^L \mathcal{C}_\Sigma[(z_1, z_2)]\, dz_1\, dz_2\,. \tag{67}$$

We make extensive use of the simple form of (67) under the assumption of isotropy throughout the work.

REFERENCES

[1] M. A. Duguay and J. W. Hansen, "An ultrafast light gate," *Appl. Phys. Lett.*, vol. 15, no. 6, pp. 192–194, 1969.

[2] R. H. Stolen and A. Ashkin, "Optical Kerr effect in glass waveguide," *Appl. Phys. Lett.*, vol. 22, no. 6, pp. 294–296, 1973.

[3] S. G. Evangelides *et al.*, "Polarization multiplexing with solitons," *J. Lightwave Technol.*, vol. 10, pp. 28–35, Jan. 1992.

[4] B. C. Collings and L. Boivin, "Nonlinear polarization evolution induced by cross-phase modulation and its impact on transmission systems," *IEEE Photon. Technol. Lett.*, vol. 12, pp. 1582–1584, Nov. 2000.

[5] C. J. Xie *et al.*, "Effect of cross-phase-modulation-induced polarization scattering on optical polarization mode dispersion compensation in wavelength-division-multiplexed systems," *Opt. Lett.*, vol. 28, no. 23, pp. 2303–2305, Dec. 2003.

[6] A. Bononi *et al.*, "Degree of polarization degradation due to cross-phase modulation and its impact on polarization-mode dispersion compensators," *J. Lightwave Technol.*, vol. 21, pp. 1903–1913, Sep. 2003.

[7] D. van den Borne *et al.*, "Cross phase modulation induced depolarization penalties in 2×10 Gbit/s polarization-multiplexed transmission," in *30th European Conf. Optical Communication (ECOC)*, 2004, Mo4.5.5.

[8] D. van den Borne *et al.*, "Reduction of nonlinear penalties through polarization interleaving in 2×10 Gb/s polarization-multiplexed transmission," *IEEE Photon. Technol. Lett.*, vol. 17, pp. 1337–1339, 2005.

[9] M. R. Phillips and S. L. Woodward, "Cross-polarization modulation: Theory and measurement of a two-channel WDM system," *IEEE Photon. Technol. Lett.*, vol. 17, pp. 2086–2088, Oct. 2005.

[10] M. Karlsson and H. Sunnerud, "Effects of nonlinearities on PMD-induced system impairments," *J. Lightwave Technol.*, vol. 24, pp. 4127–4137, Nov. 2006.

[11] C. R. Menyuk and B. S. Marks, "Interaction of polarization mode dispersion and nonlinearity in optical fiber transmission systems," *J. Lightwave Technol.*, vol. 24, pp. 2806–2826, Jul. 2006.

[12] M. Boroditsky *et al.*, "Effect of nonlinearities on PMD," *J. Lightwave Technol.*, vol. 24, pp. 4100–4107, Nov. 2006.

[13] Q. Lin and G. P. Agrawal, "Vector theory of cross-phase modulation: Role of nonlinear polarization rotation," *IEEE J. Quantum Electron.*, vol. 40, pp. 958–964, Jul. 2004.

[14] D. Marcuse *et al.*, "Application of the Manakov-PMD equation to studies of signal propagation in optical fibers with randomly varying birefringence," *J. Lightwave Technol.*, vol. 15, pp. 1735–1746, Sep. 1997.

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

[16] C. J. McKinstrie *et al.*, "Stokes-space derivations of generalized Schrödinger equations for wave propagation in various fibers," *Opt. Expr.*, vol. 15, no. 17, pp. 10 964–10 983, Aug. 2007.

[17] J. P. Gordon and H. Kogelnik, "PMD fundamentals: Polarization mode dispersion in optical fibers," *Proc. National Academy of Sciences of the United States of America*, vol. 97, no. 9, pp. 4541–4550, Apr. 2000.

[18] C. R. Menyuk, "Application of multiple-length-scale methods to the study of optical fiber transmission," *J. Eng. Math.*, vol. 36, no. 1-2, pp. 113–136, Aug. 1999.

[19] R. P. Feynman, "An operator calculus having applications in quantum electrodynamics," *Phys. Rev.*, vol. 84, no. 1, pp. 108–128, Oct. 1951.

[20] W. Magnus, "On the exponential solution of differential equations for a linear operator," *Commun. Pure and Appl. Math.*, vol. 7, no. 4, pp. 649–673, 1954.

[21] N. Korneev, "Application of Magnus series for polarization evolution in fibers," *Revista mexicana de física*, vol. 48, no. 3, pp. 250–254, Jun. 2002.

[22] P. H. Roberts and H. D. Ursell, "Random walk on a sphere and on a Riemannian manifold," *Philos. Trans. Roy. Soc. London, Ser. A*, vol. 252, no. 1012, pp. 317–356, 1960.

[23] S. Chandrasekhar and X. Liu, "Impact of channel plan and dispersion map on hybrid DWDM transmission of 42.7-Gb/s DQPSK and 10.7-Gb/s OOK on 50-GHz grid," *IEEE Photon. Technol. Lett.*, vol. 19, pp. 1801–1803, 2007.

[24] M. Karlsson *et al.*, "Long-term measurement of PMD and polarization drift in installed fibers," *J. Lightwave Technol.*, vol. 18, pp. 941–951, Jul. 2000.

[25] T. Ueda and W. L. Kath, "Dynamics of optical pulses in randomly birefringent fibers," *Physica D*, vol. 55, no. 1-2, pp. 166–181, Feb. 1992.

[26] J. P. Gordon, "Statistical properties of polarization mode dispersion," in *Polarization Mode Dispersion*, A. Galtarossa and C. R. Menyuk, Eds. New York: Springer, 2005, pp. 52–69.

[27] H. Kogelnik *et al.*, "Polarization-mode dispersion," in *Optical Fiber Telecommunications IV B*, I. P. Kaminov and T. Li, Eds. Boston, MA: Academic Press, 2002, pp. 725–861.

Marcus Winter (S'03) was born in Potsdam, Germany, in 1975. He received the Dipl.-Ing. degree in electrical engineering from Technische Universität Berlin in 2004.

From 2004 until 2006, he was with Hymite GmbH in Berlin, Germany, where he contributed to the development of hermetic packaging for opto-electronic and MEMS components. In 2006, he returned to the Technische Universität Berlin to pursue the Ph.D. degree, where he is currently researching cross-polarization modulation and its effects on polarization-multiplexed transmission in cooperation with NokiaSiemens Networks AG.

Christian-Alexander Bunge (M'03) was born in Berlin, Germany, in 1973. He received the Dipl.-Ing. degree in electrical engineering from University of Technology Berlin (TU Berlin) in 1999. Until 2002, he was with TU Berlin's photonics group of Prof. Petermann, where he received the Ph.D. degree.

From 2002 to 2004, he was with the Polymer Optical Fiber Application Center (POF-AC) in Nuremberg, where he was responsible for fiber modelling and short-haul systems design. In 2004, he he joined the TU Berlin as senior scientist, working on high-speed optical transmission systems, studying nonlinear optics, and modelling of optical components. In 2009, he became Professor at Deutsche Telekom's University for Telecommunication in Leipzig, where his main research interests are short-haul and access transmission techniques, multimode glass and polymer-fiber links, and signal processing.

Dario Setti received the M.Sc. degree in telecommunication engineering from the University of Padova, Italy, in 2000 for a thesis dealing with the modeling of optical fibers affected by polarization mode dispersion.

In 2001, he joined the R&D group of Pirelli Labs, Milan, Italy, where his activity focused on polarization mode dispersion compensation in 40 G optical networks and tunable laser sources for DWDM systems. Since 2006, he has been with Siemens AG (from April 2007 Nokia Siemens Networks GmbH & Co. KG), Munich, Germany, where his responsibilities include the design of erbium doped fiber amplifiers and the impact of nonlinear impairments in 100G polarization-multiplexed transmission with coherent detection.

Klaus Petermann (M'76–SM'85–F'09) was born in Mannheim, Germany, on October 02, 1951. He received the Dipl.-Ing. degree in 1974 and the Dr.-Ing. degree in 1976, both in electrical engineering from the Technische Universität Braunschweig, Germany.

From 1974 to 1976, he was a Research Associate at the Institut für Hochfrequenztechnik, Technische Universität Braunschweig, where he worked on optical waveguide theory. From 1977 to 1983, he was with AEG-Telefunken, Forschungsinstitut Ulm, Germany, where he was engaged in research work on semiconductor lasers, optical fibers, and optical fiber sensors. In 1983 he became a full professor at the Technische Universität Berlin, where his research interests are concerned with optical fiber communications and integrated optics.

In 1993 Dr. Petermann was awarded with the the Leibniz Award from the Deutsche Forschungsgemeinschaft. In 1999/2000 he received the Distinguished Lecturer Award from the Laser and Electro-Optics Society within the IEEE. From 1999 to 2004, he was an associate editor for *IEEE Photonics Technology Letters* and from 1996 to 2004, he was a member of the board of the VDE. From 2004 until 2006, he was Vice President for research at the Technische Universität Berlin, and from 2001 to 2008, he was member of the Senate of the Deutsche Forschungsgemeinschaft (German research council). He is a member of the Berlin-Brandenburg academy of science.