

Hybrid Hinge Model for Polarization-Mode Dispersion in Installed Fiber Transmission Systems

Jonathan Schuster, Zachary Marzec, William L. Kath, and Gino Biondini

Abstract—A hybrid hinge model is proposed to characterize polarization-mode dispersion (PMD) in installed optical fiber communication systems. The model reduces to previously considered PMD generation mechanisms in special cases, but can also simulate more general hinge behavior. A combination of importance sampling and the cross entropy method is used to calculate the probability density function of the differential group delay and the outage probability for individual wavelength bands, and results are averaged over all wavelength bands to obtain the non-compliant capacity ratio (NCR). These results suggest that, for practical outage specifications, the NCR of a system lies between upper and lower bounds provided by the isotropic hinge model and the waveplate hinge model, respectively.

Index Terms—Cross entropy (CE) method, hinge model, importance sampling, non-compliant capacity ratio (NCR), optical fiber commutations, outage probability, polarization-mode dispersion (PMD).

I. INTRODUCTION

IN the hinge model of polarization-mode dispersion (PMD), optical fiber transmission systems are represented by a concatenation of long, stable fiber sections joined by short, unprotected sections, called hinges, that are subject to environmental effects [1]. The differential group delay (DGD) of each stable section is assumed to follow a Maxwellian distribution with respect to wavelength, but is frozen in time. As a result, not all channels are statistically equivalent, and outage probabilities are in general wavelength-dependent. Accordingly, the non-compliant capacity ratio (NCR) was introduced in [2] to quantify the percentage of wavelength bands that do not meet a predetermined outage specification.

A common assumption in most theoretical studies of the hinge model is that the hinges act as polarization controllers that scat-

ter the PMD vector uniformly across the Poincaré sphere [1], [3]. We refer to this as the *isotropic* hinge model. The system outages under this assumption have been well studied, and analytical expressions of the outage probability have also been obtained [4]. Experimental [5] and theoretical studies [6], however, suggest that the isotropic assumption does not accurately represent the actual hinge behavior. Based on this evidence, an anisotropic “waveplate” hinge model was proposed [7], [8], in which the hinges are assumed to produce a random rotation about a static axis. On the other hand, this last assumption may also be too restrictive to accurately characterize the physical mechanisms of PMD production. Here we first introduce a new, hybrid hinge model. In special cases, the hinge behavior in this model reduces to the two models previously considered in the literature; but the model is more general, and can generate more complicated PMD statistics. We then compare the PMD-induced transmission penalties of all of the various hinge models.

II. HYBRID HINGE MODEL OF PMD

At each wavelength, the PMD of a finite number N of sections is given by the PMD concatenation equation [7]–[9]

$$\vec{\tau}_{n+1} = \mathbf{R}_{n+1} \mathbf{H}_n \vec{\tau}_n + \Delta \vec{\tau}_{n+1} \quad (1)$$

where $\vec{\tau}_n$ is the total PMD vector after the n th section and $\Delta \vec{\tau}_n$ is the (fixed) PMD vector of the n th section. \mathbf{R}_{n+1} is the (fixed) rotation matrix of the $(n+1)$ th section and \mathbf{H}_n is the (rapidly varying) rotation matrix of the n th hinge, positioned between the n th and the $(n+1)$ th section

$$\mathbf{R}_n = \cos \theta_n \mathbf{I} + (1 - \cos \theta_n) \hat{r}_n \hat{r}_n^T + \sin \theta_n \hat{r}_n \times \quad (2a)$$

$$\mathbf{H}_n = \cos \phi_n \mathbf{I} + (1 - \cos \phi_n) \hat{h}_n \hat{h}_n^T + \sin \phi_n \hat{h}_n \times \quad (2b)$$

where \hat{r}_n and \hat{h}_n are the respective rotation axes. The PMD vector $\Delta \vec{\tau}_n$ of each section is taken to be uniformly distributed on the Poincaré sphere with respect to wavelength but is frozen (over the temporal scales of interest), and the DGD $\Delta \tau_n = |\Delta \vec{\tau}_n|$ of each section obeys a Maxwellian distribution with respect to wavelength. In the isotropic hinge model, the hinge rotation axis \hat{h}_n and rotation angle ϕ_n are assumed to vary in whatever way ensures that the output is randomized uniformly over the Poincaré sphere. In the waveplate hinge model studied in [7] and [8], \hat{h}_n is fixed and is uniformly distributed on the Poincaré sphere with respect to wavelength, while ϕ_n is random and dynamically varying (and assumed to be uniform in $[0, 2\pi]$). Neither of these models correctly describes the actual hinge output [1], however. On one hand, the uniform coverage assumed in the isotropic model is unrealistic (see [8] and also below). Conversely, while any birefringence

Manuscript received January 23, 2014; accepted February 5, 2014. Date of publication February 11, 2014; date of current version March 7, 2014. This work was supported in part by the National Science Foundation under Grant DMS-0908399. The work of J. Schuster and Z. Marzec was supported in part by the “URGE to Compute” Program at SUNY Buffalo through NSF’s CSUMS Grant 0802994.

J. Schuster is with the Department of Mathematics, State University of New York at Buffalo, Buffalo, NY 14260 USA, and also with the Department of Electrical and Computer Engineering, Boston University, Boston, MA 02215 USA (e-mail: js382@bu.edu).

Z. Marzec and G. Biondini are with the Department of Mathematics, State University of New York at Buffalo, Buffalo, NY 14260 USA (e-mail: zmarzec@gmail.com; biondini@buffalo.edu).

W. L. Kath is with the Department of Engineering Sciences and Applied Mathematics, Northwestern University, Evanston, IL 60208 USA (e-mail: kath@northwestern.edu).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/JLT.2014.2305975

variation is described by a generalized waveplate to first order [6], the higher-order terms produce further scattering of the hinge output on the Poincaré sphere that is not reproduced by the waveplate hinge model (for example, see the experimental output presented in [1, Fig. 13]). Here, we therefore consider a more general hinge model, in which each hinge rotation axis \hat{h}_n varies dynamically, but is preferentially distributed around a fixed reference axis \hat{h}_n^{ref} . Different limiting cases are obtained by considering different values of the parameter c , as described below. We refer to this as the *hybrid* hinge model of PMD, and to the constant c as the *concentration parameter*.

More precisely, in the hybrid hinge model one takes each \hat{h}_n to be preferentially close to a reference axis \hat{h}_n^{ref} in such a way that the angle $\alpha_n = \arccos(\hat{h}_n \cdot \hat{h}_n^{\text{ref}})$ is distributed according to the probability density function (PDF)

$$p_{\cos \alpha}(\cos \alpha) = (c/2) [(\cos \alpha + 1)/2]^{c-1} \quad (3)$$

while the azimuthal angle δ_n between \hat{h}_n and \hat{h}_n^{ref} is uniformly distributed in $[0, 2\pi]$. We refer to the above as the *natural* distribution of the hinge axes, and to the corresponding PDF as the *concentration* PDF. The fixed parameter $c \geq 1$ determines how narrowly \hat{h}_n is concentrated around \hat{h}_n^{ref} . The value $c = 1$ yields a purely random rotator; namely, a rotator in which \hat{h}_n is uniform on the Poincaré sphere and ϕ_n is uniform in $[0, 2\pi]$. Conversely, the limit $c \rightarrow \infty$ recovers the waveplate hinge model considered in [7], i.e., \hat{h}_n fixed and identically equal to \hat{h}_n^{ref} . (The PDF in (3) can be generated by choosing $\cos \alpha = 2\xi^{1/c} - 1$, with ξ uniform in $[0, 1]$). Plots of the concentration PDF in (3) for different values of the concentration parameter can be found in [10, Fig. 2]. Of course (3) is a heuristic choice, and a precise characterization of the hinge action is likely to result in a more accurate (and more complicated) expression. Nonetheless, Equation (3) is a simple and reasonable model given our knowledge of the hinge action. Most importantly, the output distribution of each hinge in the hybrid model does mimic the experimental hinge output (again, see [1, Fig. 13]).

Fig. 1 compares the distribution of the output state of polarization (SOP) $\hat{s}^{\text{out}} = \mathbf{H} \hat{s}^{\text{in}}$ produced by different hinge models acting on the fixed input SOP $\hat{s}^{\text{in}} = (0, 1, 0)^T$. As expected, increasing the value of the concentration parameter c gradually causes the output distribution to change from that produced by a random rotator to that of a generalized waveplate. Note that a random rotator (i.e., a device that produces a uniform rotation about a uniformly distributed axis) does *not* produce a perfectly uniform output, as seen in the top left plot of Fig. 1. This result, which may seem surprising at first, is nonetheless well known [8], and provides additional evidence that the assumption of an isotropic output is likely to be unrealistic.

III. METHODS FOR RARE PMD EVENTS

To determine outage probabilities and the system impact of PMD, one needs accurate rare-event DGD statistics. These can be obtained by combining importance sampling (IS) [10], [11] and the cross-entropy (CE) method [12]. (Of course other variance reduction techniques could be used instead, such as the

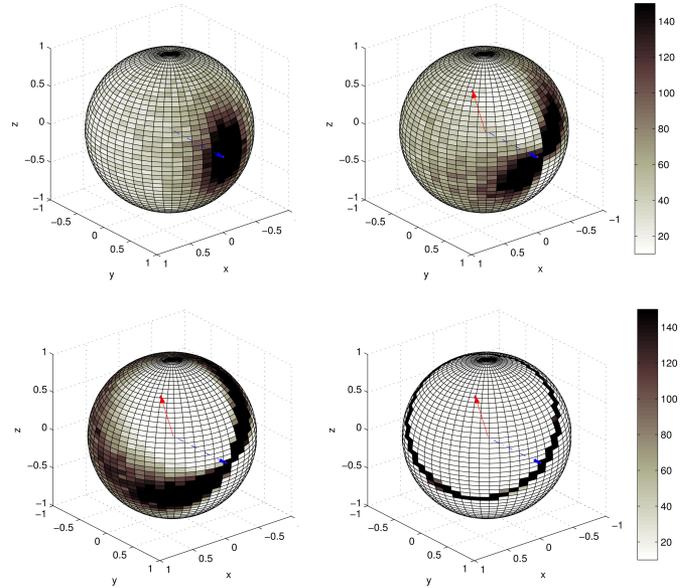


Fig. 1. Distribution of the output SOP on the Poincaré sphere after passing through hinges with different concentration parameter c , as constructed from 100 000 Monte Carlo samples. Top left: random rotator ($c = 1$); top right: hybrid, $c = 10$; bottom left: hybrid, $c = 50$; bottom right: waveplate ($c = \infty$). Dashed blue arrow: input SOP. The preferential hinge axis (solid red arrow) for $c \neq 1$ was $\hat{h}^{\text{ref}} = (0.487, 0.382, 0.786)^T$.

multi-canonical Monte Carlo (MC) method [13] or Markov-chain MC methods [14].)

An introduction to IS and its use for PMD was given in [10]. Recall that, with IS, the MC samples are biased so as to produce large DGD values more frequently, while correcting for this bias with likelihood ratios and weighting [8], [10]. The implementation of the IS-CE method for the hybrid hinge models is similar to that for the waveplate hinge model [15]. Namely, the method comprises two steps: for each random realization of hinge axes, (i) one first finds the biasing point (i.e., the configuration of hinge parameters that yields the largest DGD) then (ii) one iteratively finds numerically the optimal biasing strengths that produce increasingly larger values of DGD. For the waveplate hinge model, the first step must be performed through a numerical, iterative optimization procedure. In the hybrid hinge model, however, this step can be done analytically, as we discuss next.

A. Importance Sampling for the Hybrid Model

As usual, in order for IS to be effective, one must bias the MC samples towards the most likely configurations that produce the events of interest. Even though the PMD dynamics in the hybrid hinge model are in general more complicated than in the isotropic or waveplate hinge model, the most likely configurations that produce the maximum DGD can be determined analytically for all finite values of the concentration parameter c .

For this purpose it is convenient to introduce the unit vectors

$$\hat{t}_n^{\text{in}} = \Delta \vec{\tau}_n / |\Delta \tau_n|, \quad \hat{t}_n^{\text{out}} = \mathbf{R}_{n+1}^{-1} \Delta \vec{\tau}_{n+1} / |\Delta \tau_{n+1}|. \quad (4)$$

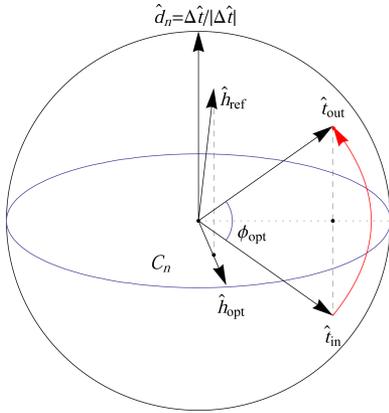


Fig. 2. Diagram showing the “optimal” hinge axis \hat{h}^{opt} that rotates a given \hat{t}^{in} into a given \hat{t}^{out} , together with the corresponding rotation angle ϕ^{opt} .

It should be clear from (1) that any collection of hinge matrices \mathbf{H}_n that rotate each \hat{t}_n^{in} into \hat{t}_n^{out} produces a configuration in which the total DGD τ_N equals its maximum value, $\tau^{\max} = \sum_{n=1}^N \Delta \tau_n$. As illustrated in Fig. 2, such configurations are obtained whenever the following two conditions are met: (i) \hat{h}_n lies in the great circle C_n perpendicular to the line with direction vector $\Delta \vec{t}_n = \hat{t}_n^{out} - \hat{t}_n^{in}$, and (ii) ϕ_n is simply the angle between the perpendicular components of \hat{t}_n^{in} and \hat{t}_n^{out} with respect to \hat{h}_n . One such rotation is shown by the red arc in Fig. 2.

Since \hat{h}_n is preferentially aligned with \hat{h}_n^{ref} , however, not all axes in C_n are equally likely. The most likely such rotation matrix is that for which the hinge axis is closest to \hat{h}_n^{ref} . This choice is readily found as the (normalized) perpendicular projection of \hat{h}_n^{ref} with respect to $\Delta \vec{t}_n$. We denote with \hat{h}_n^{opt} and ϕ_n^{opt} this “optimal” axis and the corresponding rotation angle, respectively. Explicitly,

$$\hat{h}_n^{\text{opt}} = \hat{h}_{n,\perp}^{\text{ref}} / |\hat{h}_{n,\perp}^{\text{ref}}| \quad (5)$$

$$\cos \phi_n^{\text{opt}} = (\hat{t}_{n,\perp}^{\text{out}} \cdot \hat{t}_{n,\perp}^{\text{in}}) / (|\hat{t}_{n,\perp}^{\text{out}}| |\hat{t}_{n,\perp}^{\text{in}}|) \quad (6)$$

where $\vec{a}_\perp = \vec{a} - (\vec{a} \cdot \hat{d}_n) \hat{d}_n$ is the perpendicular part of a vector \vec{a} with respect to the direction $\hat{d}_n = \Delta \vec{t}_n / |\Delta \vec{t}_n|$.

The IS-MC samples can be biased towards this configuration by: (a) biasing \hat{h}_n toward \hat{h}_n^{opt} by taking $\beta_n = \arccos(\hat{h}_n \cdot \hat{h}_n^{\text{opt}})$ to be distributed according to

$$p_{\cos \beta}^*(\cos \beta_n) = (b/2) [(\cos \beta_n + 1)/2]^{b-1} \quad (7)$$

(b) taking the corresponding azimuthal angle, which we denote by ψ_n , to be uniform in $[0, 2\pi]$, and (c) taking the hinge rotation angle ϕ_n to be distributed according to

$$p_\phi^*(\phi_n) = \frac{b}{2\pi^b} |\phi_n - \phi_n^{\text{opt}} - \pi|^{b-1}. \quad (8)$$

Equation (7) is realized similarly as (3), while (8) is realized by taking $\phi_n = \phi_n^{\text{opt}} + \pi(1 + \text{sign}(\zeta)|\zeta|^{1/b})$, with ζ uniform in $[0, 1]$.

Fig. 3 illustrates the angles between the various axes. For the “natural” distribution, the relevant angles are the polar angle $\alpha_n \in [0, \pi]$ and the azimuthal angle $\delta_n \in [0, 2\pi]$ between \hat{h}_n^{ref}

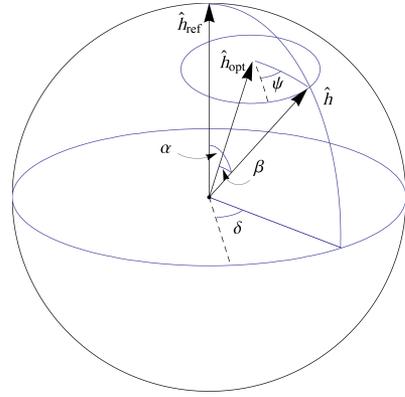


Fig. 3. Diagram showing the angles α and δ for the “natural” distribution for each section, as well as the angles β and ψ for the hybrid biasing distribution. (Note that the orientation of the Poincaré sphere is different from that in Fig. 2.)

and \hat{h}_n . For the hybrid biasing distribution, these are the polar angle $\beta_n \in [0, \pi]$ and the azimuthal angle $\psi_n \in [0, 2\pi]$ between \hat{h}_n^{opt} and \hat{h}_n .

As usual, the value of the biasing strength parameter b determines how heavily the samples are biased around the “optimal” configuration. When $b = 1$, the unbiased distribution is recovered, whereas increasing values $b > 1$ concentrate the MC samples more narrowly around the biasing point. Using several values of b (chosen using the CE method, as described in the next section) in conjunction with a multiple IS simulation (i.e., the method of combining several different biasing distributions [16], [17]) one can accurately produce a whole range of DGDs with correct statistics. We found this biasing method to be an effective way to produce probabilities down to 10^{-15} for values of the concentration parameter c between 3 and 30.

It is important to realize, however, that biasing towards the configuration that produces the largest DGD is not always the most probable way to generate moderately large DGDs. Most importantly, in order for IS to be effective, one needs the latter, not the former. For large values of c , the \hat{h}_n in the natural distribution concentrate so narrowly near \hat{h}_n^{ref} that the above “optimal” configuration is very unlikely, and biasing towards it is not an effective way to generate moderately large DGDs. In this case we use the natural distribution of hinge axes, and we bias the hinge rotation angles $\phi = (\phi_1, \dots, \phi_{N-1})$ towards the collection of values $\Phi = (\Phi_1, \dots, \Phi_{N-1})$ that yields the largest DGD, which can be found as in the waveplate hinge model [8]. Namely, as in [8], we use the biasing PDF (8) for the angles ϕ_n , with ϕ_n^{opt} replaced by Φ_n . Even though this biasing choice is clearly not asymptotically efficient, we nonetheless found it to be effective down to probabilities of 10^{-15} for values of c larger than 20.

Conversely, for $c = 1$, the natural distribution of the hinge axes has no preferred orientation, so all \hat{h}_n in C_n are equally likely. We then bias \hat{h}_n to be close to C_n , by taking

$$p_{\cos \eta}^*(\cos \eta) = 1/(2b |\cos \eta|^{1-1/b}) \quad (9)$$

where $\eta_n = \arccos(\hat{h}_n \cdot \hat{d}_n)$ (as shown in Fig. 4), with the corresponding azimuthal angle uniform in $[0, 2\pi]$, and we bias ϕ_n

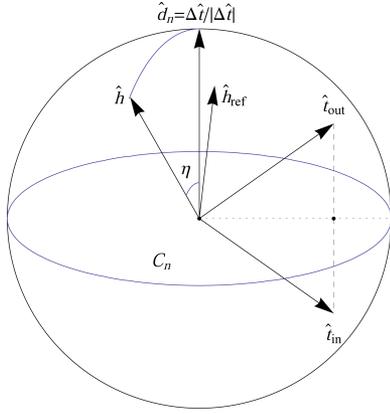


Fig. 4. Diagram showing the biasing angle η for the random rotator biasing.

towards the corresponding angle that produces the largest DGD, as before. (The above distribution for η can be realized by choosing $\cos \eta = \text{sign}(\xi)|\xi|^b$, with ξ uniform in $[-1, 1]$.) We found this biasing choice to be effective down to probabilities of 10^{-15} for values of c up to 5.

B. Cross-Entropy for the Hybrid Model

In order to accurately compute outage probabilities, one must perform IS-MC simulations with a range of values of b chosen so as to effectively cover the whole range of possible DGDs. Such values, however, depend on the specific fiber realization through the sectional PMD vectors and hinge axes, and no simple criterion exists for automatically choosing the biasing strengths. Fortunately, a set of appropriate biasing strengths can be computed using the CE method, as we discuss next.

An introduction to CE and its detailed implementation for the waveplate hinge model was given in Refs. [8], [15]. Recall that the optimal value of b in order to reach a region R of interest in sample space can be obtained by solving numerically the following optimization problem:

$$\sum_{m=1}^M I_R(\tau(\mathbf{y}^{(m)}))L(\mathbf{y}^{(m)}; b) \frac{d}{db} \log[p^*(\mathbf{y}^{(m)}; b)] = 0 \quad (10)$$

where the vector \mathbf{y} collects the relevant angles for all the hinge rotation axes, $\mathbf{y}^{(m)}$ is its value for the m th MC sample, $\tau(\mathbf{y}^{(m)})$ is the corresponding total DGD, $p^*(\mathbf{y}; b)$ is the desired optimal biasing PDF,

$$L(\mathbf{y}; b) = p(\mathbf{y})/p^*(\mathbf{y}; b) \quad (11)$$

is the corresponding IS likelihood ratio and $I_R(\cdot)$ is the indicator function of R . In practice, Equation (10) is used iteratively to find a collection of values of b that cover the desired range of DGDs. That is, one starts with the biasing strength $b_1 = 1$, corresponding to the unbiased PDF. Then, at the j th step, the next value b_{j+1} is computed by using the biasing PDF with $b = b_j$ to generate the MC samples, and the region R_j is defined by a given percentile of the MC samples with the largest DGD in the j th iteration.

Corresponding to the three regimes discussed above, three different biasing strategies will be used: (i) hybrid biasing for

c between 3 and 30, (ii) random rotator biasing for $c < 5$ and (iii) waveplate biasing for $c > 20$. Accordingly, the biasing angles are: for the hybrid biasing, $\mathbf{y} = (\beta_1, \phi_1, \dots, \beta_{N-1}, \phi_{N-1})$; for the random rotator biasing, $\mathbf{y} = (\eta_1, \phi_1, \dots, \eta_{N-1}, \phi_{N-1})$; for the waveplate biasing, $\mathbf{y} = (\phi_1, \dots, \phi_{N-1})$. In each case, Equation (10) can be solved analytically, as we show next.

For all three biasing methods, the overall biasing PDF is the product of the individual biasing PDFs for each hinge. In particular, for the hybrid biasing

$$p^*(\mathbf{y}; b) = \prod_{n=1}^{N-1} p_\beta^*(\beta_n) p_\phi^*(\phi_n) \quad (12)$$

where $p_\beta^*(\beta_n)$ and $p_\phi^*(\phi_n)$ are given by (7) and (8). Inserting these expressions into (10) and solving for b yields the biasing strength parameters for the hybrid biasing (see the Appendix for details)

$$b_{j+1} = \left[\frac{2(N-1) \sum_{m=1}^M IL_j(\mathbf{y}^{(m)})}{\sum_{m=1}^M IL_j(\mathbf{y}^{(m)}) Y(\mathbf{y}^{(m)})} \right]^{1+k} \quad (13)$$

where for brevity we introduced the shorthand notations

$$IL_j(\mathbf{y}) = I_{R_j}(\tau(\mathbf{y}))L(\mathbf{y}; b_j) \quad (14)$$

$$Y(\mathbf{y}) = \sum_{n=1}^{N-1} \ln \left[\left(\frac{\cos \beta_n + 1}{2} \right) \left| \frac{\phi_n - \phi_n^{\text{opt}} - \pi}{\pi} \right| \right]. \quad (15)$$

Actually, the solution of (10) yields (13) with $k = 0$. The additional parameter k is a convergence factor that was artificially added to the expression for b_{j+1} in order to speed up the efficiency of the CE iteration in reaching large values of DGD. The specific values of k used in the simulations were determined empirically, and are given in the next section.

Since a different biasing PDF is used for the random rotator biasing, the expression for the optimal biasing strengths is also different. Specifically, $p^*(\mathbf{y}; b)$ is given by an expression similar to (12)

$$p^*(\mathbf{y}; b) = \prod_{n=1}^{N-1} p_{\cos \eta}^*(\cos \eta_n) p_\phi^*(\phi_n) \quad (16)$$

where the individual PDFs are now given by (8) and (9), and the likelihood ratios change accordingly. Correspondingly, the biasing strength parameters for the random rotator biasing are (see the Appendix for details):

$$b_{j+1} = \left[\frac{\sum_{m=1}^M IL_j(\mathbf{y}^{(m)}) \sum_{n=1}^{N-1} \log |\cos \beta_n|}{\sum_{m=1}^M IL_j(\mathbf{y}^{(m)}) \sum_{n=1}^{N-1} \log \left| \frac{\phi_n - \phi_n^{\text{opt}} - \pi}{\pi} \right|} \right]^{1/2+k}. \quad (17)$$

Again, Equation (10) yields (17) with $k = 0$, and the additional factor k was introduced to speed up the convergence of the CE iteration in reaching large values of DGD.

The CE step for the waveplate biasing is very similar to that for the waveplate model, as detailed in [8]. In other words, the major difference compared to the hybrid biasing and the random rotator biasing is that the hinge rotation axes are no longer biased in this method. Instead they are taken according to the natural

distribution, given by (3). (Of course the waveplated model is also different since in that case the hinge rotation axes are fixed, corresponding to the limit $c \rightarrow \infty$.) The biasing PDF is now given by

$$p^*(\mathbf{y}; b) = \prod_{n=1}^{N-1} p^*(\phi_n) \quad (18)$$

with $p^*(\phi_n)$ given by (8) with ϕ_n^{opt} replaced by Φ_n . That is, the biasing PDF is now solely dependent on ϕ_n , with no dependence on β_n or η_n . The likelihood ratios for the waveplate change accordingly. Solving (10) for b then yields the biasing strength parameters for the waveplate biasing

$$b_{j+1} = -\frac{(N-1) \sum_{m=1}^M IL_j(\mathbf{y}^{(m)})}{\sum_{m=1}^M IL_j(\mathbf{y}^{(m)}) \sum_{n=1}^{N-1} \ln \left| \frac{\phi_n^{(m)} - \Phi_n - \pi}{\pi} \right|}. \quad (19)$$

No convergence factor was added in this case. Equation (19) is identical to the expression for the biasing strength that was given in [8] for the waveplate hinge model—as was to be expected, since the hinge axes are not biased in this case.

Due to the need to perform an iterative search for the optimal biasing strengths, the overall computational complexity of the above combined IS-CE method is comparable to that of adaptive methods such as the multi-canonical MC method of Markov-chain MC methods. As usual, however, the benefit of IS is that it allows one to efficiently compute numerical variances for all the quantities of interest.

IV. SYSTEM PERFORMANCE

We used the above-described methods to numerically investigate the statistical properties of PMD in the various hinge models considered.

Fig. 5 compares the PDF of the total DGD for different hinge models for a system of 10 sections with a specific, randomly chosen realization of reference hinge axes. The smoothness of the numerically reconstructed PDFs serves as a confirmation of the validity of each biasing method, which results in a very small variance in each bin. For the specific fiber realization in the figure, the statistics of the random rotator model are not too dissimilar from those of the isotropic model, while the statistics of the hybrid model vary between those of the random rotator model and the waveplate model depending on the value of the concentration parameter c . It is important to realize, however, that even when the sectional PMD vectors and sectional rotation matrices are kept the same, each choice of reference hinge axes results in a different DGD distribution in the hybrid hinge model. No general pattern seems to exist regarding the relationship between the PDFs corresponding to different hinge models, and in this respect the situation is similar to that in [8]. Nonetheless, the system impact of the various hinge models differ significantly, as we discuss next.

Recall that outage probabilities are required to be very small, typically a minute per year (that is, 10^{-6}) or less. The probability P_{out} of a system outage can be computed using the outage map

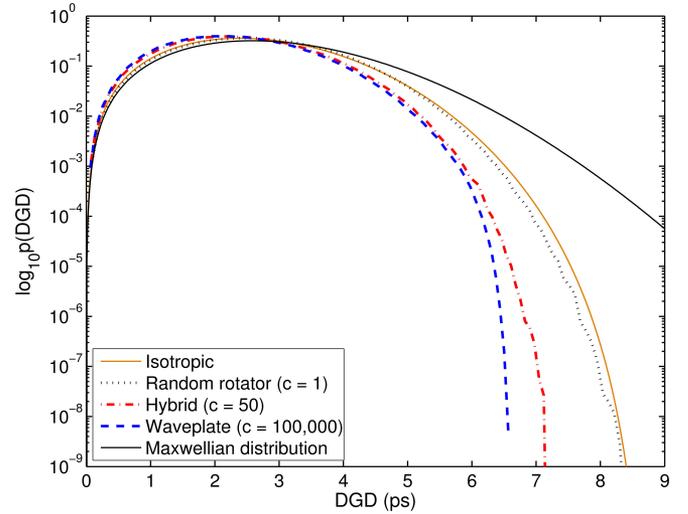


Fig. 5. The PDF of the DGD in logarithmic scale for various hinge models of PMD: isotropic model (solid orange line), random rotator model (dotted black line), hybrid model with $c = 50$ (dot-dashed red line) and waveplate model (dashed blue line). A Maxwellian distribution (solid black line) is also included for comparison. The sectional PMD vectors, Müller matrix parameters and hinge rotation axes are the same as in [8], except for the angles θ_n , whose values are, respectively, 2.3145, 1.7642, 2.4515, 2.7849, 2.5679, 1.8978, 2.5621, 2.3798, 2.8755, and 3.1361.

approach [2], [18]. As in [8]

$$P_{\text{out}}(\mathbf{z}) = \iint I_{\text{out}}(\tau, \gamma) p(\tau, \gamma; \mathbf{z}) d\gamma d\tau \quad (20)$$

where $p(\tau, \gamma; \mathbf{z})$ is the joint PDF of the total DGD τ and the power splitting ratio (PSR) γ , and I_{out} is the indicator function of the outage region (which depends on the outage specifications, transmission format and receiver design): $I_{\text{out}} = 1$ at all combinations of γ and τ for which an outage occurs, and $I_{\text{out}} = 0$ otherwise. As in [8], \mathbf{z} is a multi-component random vector that collects all the “frozen” degrees of freedom in each wavelength band; namely, the sectional PMD vectors, rotation axes and angles of all the sectional Müller matrices, and all the hinge reference axes.

We considered systems with either $N = 6$ or $N = 10$ sections, with mean DGD of 3.0 or 2.5 ps using the non-return-to-zero format at 40 Gb/s, with an allocated power margin ϵ of 1 dB and corresponding outage map parameters as in [7]. We should point out, however, that the outage map is specific to the receiver considered, and different maps yield different outage statistics in general. (In particular, a different outage map was used in [8].)

Also, while outage maps have been extensively used for intensity-modulation and direct-detection systems, newer systems use polarization-division multiplexing (PDM) and are based on coherent receivers where linear impairments are compensated for via digital signal processing, which provides significant tolerance to PMD (e.g., see [19], [20]). Since we are not aware of recent work extending the concept of outage maps to coherent PDM systems, we also used an alternative, simpler outage criterion. That is, we assumed that an outage occurs if the instantaneous DGD exceeds a threshold value, as in [21]

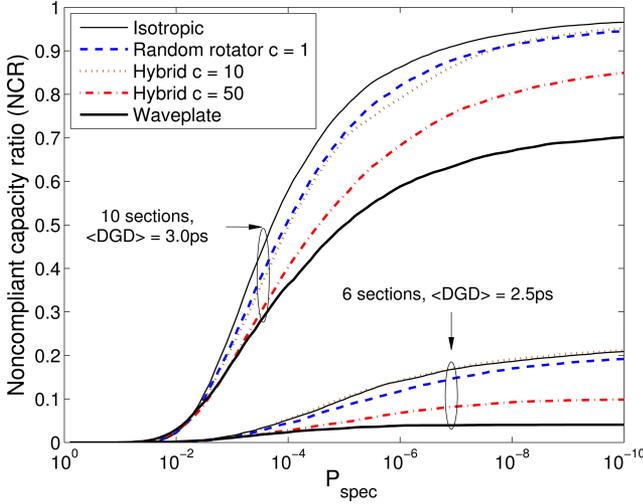


Fig. 6. The NCR as a function of P_{spec} for different hinge models, with outage probabilities computed via the outage map approach as in [8]: isotropic (thin solid line), hybrid with $c = 1$ (dashed line), hybrid with $c = 10$ (dotted line) and $c = 50$ (dot-dashed line) and waveplate (thick solid line). The top curves are for a system with ten sections and mean DGD 3.0 ps; the bottom curves are with six sections and mean DGD 2.5 ps.

and [22]. In this case, the outage probability is still given by (20), but where now $I_{\text{out}} = 1$ if $\tau > \tau_{\text{out}}$ and $I_{\text{out}} = 0$ otherwise, where τ_{out} is a fixed value encoding the tolerance of the given system to PMD. (The degree of tolerance of course depends on the system considered and also on whether or not electronic equalization is applied.)

Regardless of how the outage probability is computed, the expected value of the fraction of wavelength bands that are not compliant with a given outage specification P_{spec} then gives the NCR [2]. That is,

$$\text{NCR} = \mathbb{P}[P_{\text{out}}(\mathbf{z}) > P_{\text{spec}}] = \mathbb{E}_{\mathbf{z}}[I_{\text{NC}}(P_{\text{out}}(\mathbf{z}))] \quad (21)$$

where $I_{\text{NC}} = 1$ if $P_{\text{out}}(\mathbf{z}) > P_{\text{spec}}$ and $I_{\text{NC}} = 0$ otherwise.

For each wavelength band, we used the CE method as described earlier to obtain algorithmically a set of optimal biasing strengths that allow one to efficiently cover the full range of possible DGD values. Recall that (13) and (17) for the CE step in the hybrid biasing and random rotator biasing are both computed using a convergence factor k . The specific values of convergence factor used in the simulations were $k = 0.20$ for the random rotator biasing and $k = 0.25$ for the hybrid biasing. These values were manually selected to ensure that the CE algorithm converged toward high values of DGD sufficiently fast that the total number of biasing strengths thus obtained was not too large (which would have added significant overhead to the numerical simulations due to the need to keep track of all cross-likelihood ratios in multiple IS), but at the same time slowly enough to guarantee that for each wavelength band at least five biasing strengths were present (so that each range of values of DGD was adequately sampled).

We then performed IS-CE simulations, with 100 000 samples for each biasing strength, to reconstruct the joint PDF of the total DGD and the PSR, and we used the results to numerically estimate the outage probability in each wavelength band, using

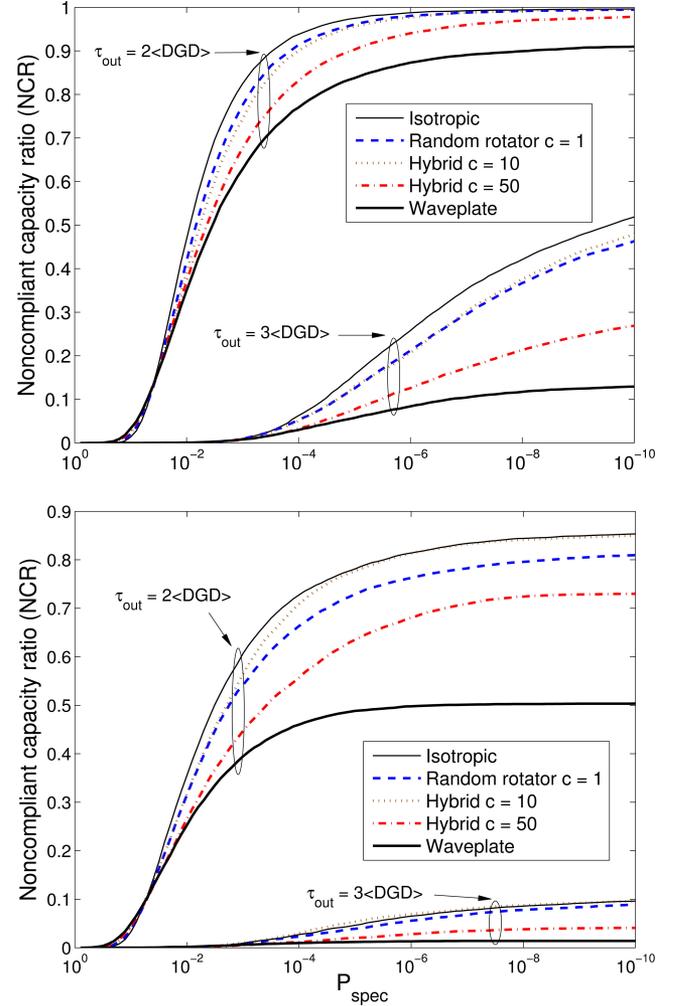


Fig. 7. The NCR as a function of P_{spec} for different hinge models, with outage probabilities computed with the simpler outage criterion as in Refs. [21], [22], for $\tau_{\text{out}} = 2\langle \tau \rangle$ and $\tau_{\text{out}} = 3\langle \tau \rangle$. Line styles and colors are as in Fig. 6. Top figure: a system with ten sections and mean DGD 3.0 ps. Bottom figure: a system with six sections and mean DGD 2.5 ps.

either the outage map approach or the alternative outage criterion discussed above. Finally, we randomized over the frozen degrees of freedom and we averaged over 10 000 wavelength bands to obtain the NCR.

Fig. 6 shows the NCR for each hinge model considered, with outage probabilities calculated via the outage map approach. (Similar results were obtained when the same outage map as in [8] was used in the simulations.) Alternatively, Fig. 7 shows the NCR with outage probabilities computed with the simpler, alternative approach, for $\tau_{\text{out}} = 2\langle \tau \rangle$ and $\tau_{\text{out}} = 3\langle \tau \rangle$. (Numerical simulations were also performed with $\tau_{\text{out}} = 4\langle \tau \rangle$, but the resulting NCR values were very small, and the corresponding curves for each model lied at the very bottoms of the plot and were therefore indistinguishable from each other.)

The curves in Fig. 7(a) should be compared to the upper curves in Fig. 6, and those in Fig. 7(b) to the lower curves in Fig. 6. The two methods of computing outage probabilities yield different values for the NCR. This is to be expected, of course, since they apply to different kinds of systems. Nonetheless, the relative

TABLE I
EMPIRICAL FIT TO NCR CURVES FOR SYSTEMS WITH TEN SECTIONS AND MEAN DGD OF 3 PS

	Outage map			$\tau_{\text{out}} = 2\langle\text{DGD}\rangle$			$\tau_{\text{out}} = 3\langle\text{DGD}\rangle$		
	NCR _o	x_o	n	NCR _o	x_o	n	NCR _o	x_o	n
Isotropic	1.0041	3.2834	3.0759	1.0084	1.8111	3.1320	0.8859	1.8907	2.1090
$c = 1$	1.0051	3.4454	2.7852	1.0132	1.8857	2.8373	0.8539	1.9701	1.9162
$c = 10$	1.0259	3.5287	2.5949	1.0198	1.9342	2.6908	0.9004	2.0063	1.9893
$c = 50$	0.9576	3.7336	2.2559	1.0169	1.9852	2.2233	0.8194	2.1433	1.5644
Waveplate	0.7782	3.5193	2.2985	0.9464	1.9701	2.1797	0.5245	1.6694	2.1557

TABLE II
EMPIRICAL FIT TO NCR CURVES FOR SYSTEMS WITH SIX SECTIONS AND MEAN DGD OF 2.5 PS

	Outage map			$\tau_{\text{out}} = 2\langle\text{DGD}\rangle$			$\tau_{\text{out}} = 3\langle\text{DGD}\rangle$		
	NCR _o	x_o	n	NCR _o	x_o	n	NCR _o	x_o	n
Isotropic	0.2499	4.7879	2.3856	0.7233	6.0559	2.2001	0.1205	4.8176	2.0958
$c = 1$	0.2553	5.2474	2.0466	0.7556	6.7750	1.8966	0.1240	5.3240	1.8075
$c = 10$	0.2538	4.7780	2.3982	0.8292	7.0851	1.8188	0.1108	4.3914	2.4022
$c = 50$	0.1288	4.8770	2.0737	0.4482	6.8215	1.8064	0.0550	4.9025	1.9201
Waveplate	0.0432	3.3172	3.1210	0.1587	4.9499	2.4062	0.0151	2.8564	2.4573

differences between different hinge models are consistent with both methods. Namely, similarly to [7], [8], the NCR for the isotropic hinge model is higher than that of all other models in all cases considered, suggesting that the isotropic hinge model may overestimate the system impact of PMD. The NCR for the random rotator model is close to that of the isotropic model, while the NCR for the waveplate model is much smaller for practical values of P_{spec} (i.e., $P_{\text{spec}} \leq 10^{-3}$). The NCR for the hybrid model tends to vary between these two extremes depending on the value of the concentration parameter.

A simple but accurate empirical fit to the above NCR curves can be obtained in the form

$$\text{NCR} = \text{NCR}_o \exp[-(x_o/x)^n] \quad (22)$$

where $x = \log_{10}(1/P_{\text{spec}})$, NCR_o determines the limiting value of the NCR as P_{spec} tends to zero, and x_o and n determine the shape of the NCR curve. Suitable values of NCR_o , x_o and n can be obtained, for example, using readily available Matlab fitting routines. The parameter values for the different systems, PMD models and outage criteria considered in Figs. 6–7 are provided in Tables I–II.

V. DISCUSSION

It should be clear that the hybrid hinge model is substantially more general than previously studied PMD models such as the isotropic hinge model or the waveplate hinge model. As such, it is better suited to characterize a variety of real devices. One can anticipate that each of the devices whose behavior can be characterized as a hinge (such as amplifiers, dispersion-compensating modules, etc.) will be associated to a different value of the concentration parameter. A precise characterization of each device with regard to the statistical properties of its output remains an important open issue. But this problem is best addressed by experiment, and as such it is outside the scope of this paper.

A natural question is whether the choice of the concentration PDF affects the results, and if so in what way. To definitively answer this question one would need to perform simulations to compare different choices, which is outside the scope of this pa-

per. On the other hand, some insight can be gained by realizing that the concentration PDF is effectively a biasing distribution for the hinge axes. Generically speaking, the choice of the specific biasing distribution does not play a significant role as long as the biasing point is chosen correctly [10], [23]. Since in this case the reference hinge axes (which define the biasing point) are randomized when averaging over different wavelength bands, it stands to reason that the choice of concentration PDF should not be of much consequence, except possibly for a rescaling of the concentration parameter. (Of course the choice of concentration PDF affects the derivation of the optimal biasing strengths in Section III. While the calculations do depend in an essential way on the choice of the biasing PDF, however, they do not depend crucially on the particular form of the concentration PDF, and they could therefore be generalized to other concentration PDFs in a straightforward way.)

At the system level we emphasize that, even though for simplicity all the simulations presented here used the same value of c for all the hinges in the system, one can just as easily simulate systems in which each hinge has a different value of c , depending on the properties of the individual elements that are acting as hinges in that system. Indeed, all the numerical methods discussed in this paper can be trivially extended to that more generalized case. Conversely, one can also envision a scenario in which each system is associated to an overall *effective* value of c .

Finally, based on the simulation results, one could expect the NCR to lie somewhere between the extremes represented by the isotropic (or the random rotator) model and the waveplate model. These two models therefore appear to yield upper and lower bounds for the value of the NCR in any given system. On the other hand, in the absence of more extensive data, this conclusion can not be reached with certainty. Therefore, the validity of this conjecture remains another open issue that will need to be settled via comparison with experimental data.

APPENDIX

Here we provide a few additional details about the derivation of (13) and (18).

Recall that the optimal value of b in order to reach a region R of probability space is given by the solution of (10). For the hybrid biasing method, Equation (12) implies

$$L(\mathbf{y}; b) = \prod_{n=1}^{N-1} \frac{2^{b-c}}{b^2} \left| \frac{\phi_n - \phi_n^{\text{opt}} - \pi}{\pi} \right|^{1-b} \frac{(\cos \alpha_n + 1)^{c-1}}{(\cos \beta_n + 1)^{b-1}}$$

$$\frac{d}{db} \log [p^*(\mathbf{y}; b)] = \frac{2(N-1)}{b} + Y(\mathbf{y})$$

where $Y(\mathbf{y})$ was defined in (15). With these expressions, Equation (10) becomes

$$\sum_{m=1}^M I_R(\tau(\mathbf{y}^{(m)})) L(\mathbf{y}^{(m)}; b) \left(\frac{2(N-1)}{b} + Y(\mathbf{y}^{(m)}) \right) = 0$$

whose solution is (13) with $k = 0$.

Similarly, for the random rotator biasing method, Equation (16) implies

$$L(\mathbf{y}; b) = \prod_{n=1}^{N-1} \left| \frac{\phi_n - \phi_n^{\text{opt}} - \pi}{\pi} \right|^{1-b} |\cos \beta_n|^{1-1/b}$$

$$\frac{d}{db} \log [p^*(\mathbf{y}; b)] = \sum_{n=1}^{N-1} \log \left| \frac{\phi_n - \phi_n^{\text{opt}} - \pi}{\pi} \right|$$

$$- \frac{1}{b^2} \sum_{n=1}^{N-1} \log |\cos \beta_n|.$$

Equation (10) then becomes

$$\sum_{m=1}^M I_R(\tau(\mathbf{y}^{(m)})) L(\mathbf{y}^{(m)}; b) \left[\frac{1}{b^2} \sum_{n=1}^{N-1} \log |\cos \beta_n^{(m)}| \right. \\ \left. - \sum_{n=1}^{N-1} \log \left| \frac{\phi_n^{(m)} - \phi_n^{\text{opt}} - \pi}{\pi} \right| \right] = 0$$

whose solution is (18) with $k = 0$.

ACKNOWLEDGMENT

The authors would like to thank A. Galtarossa, H. Kogelnik, J. Li, and P. Winzer for many insightful discussions.

REFERENCES

- [1] M. Brodsky, N. J. Frigo, M. Boroditsky, and M. Tur, "Polarization mode dispersion of installed fibers," *J. Lightw. Technol.*, vol. 24, no. 12, pp. 4584–4599, Dec. 2006.
- [2] H. Kogelnik, P. J. Winzer, L. E. Nelson, R. M. Jopson, M. Boroditsky, and M. Brodsky, "First-order PMD outage for the hinge model," *IEEE Photon. Technol. Lett.*, vol. 17, no. 6, pp. 1208–1210, Jun. 2005.
- [3] C. Antonelli and A. Mecozzi, "Theoretical characterization and system impact of the hinge model of PMD," *J. Lightw. Technol.*, vol. 24, no. 11, pp. 4064–4074, Nov. 2006.
- [4] C. Antonelli and A. Mecozzi, "Statistics of the DGD in PMD emulators," *IEEE Photon. Technol. Lett.*, vol. 16, no. 8, pp. 1840–1842, Aug. 2004.
- [5] M. Brodsky, J. C. Martinez, N. J. Frigo, and A. Sirenk, "Dispersion compensation module as a polarization hinge," in *Proc. 31st Eur. Conf. Opt. Commun.*, 2005, vol. 3, pp. 335–336.
- [6] W. Shieh and H. Kogelnik, "Dynamic eigenstates of polarization," *IEEE Photon. Technol. Lett.*, vol. 13, no. 1, pp. 40–42, Jan. 2001.
- [7] J. Li, G. Biondini, W. L. Kath, and H. Kogelnik, "Anisotropic hinge model for polarization-mode dispersion in installed fibers," *Opt. Lett.*, vol. 33, pp. 1924–1926, 2008.
- [8] J. Li, G. Biondini, W. L. Kath, and H. Kogelnik, "Outage statistics in a wave plate hinge model of polarization-mode dispersion," *J. Lightw. Technol.*, vol. 28, no. 13, pp. 1958–1968, Jul. 2010.
- [9] J. P. Gordon and H. Kogelnik, "PMD fundamentals: Polarization-mode dispersion in optical fibers," *Proc. Nat. Acad. Sci.*, vol. 97, pp. 4541–4550, 2000.
- [10] G. Biondini, W. L. Kath, and C. R. Menyuk, "Importance sampling for polarization mode dispersion: Techniques and applications," *J. Lightw. Technol.*, vol. 22, no. 4, pp. 1201–1215, Apr. 2004.
- [11] J. A. Bucklew, *Introduction to Rare Event Simulation*. New York, NY, USA: Springer-Verlag, 2004.
- [12] P. T. De Boer, D. P. Kroese, S. Mannor, and R. Y. Rubinstein, "A tutorial on the cross-entropy method," *Ann. Oper. Res.*, vol. 134, pp. 19–67, 2005.
- [13] D. Yevick, "Multicanonical communication system modeling—Application to PMD statistics," *IEEE Photon. Technol. Lett.*, vol. 14, no. 11, pp. 1512–1514, Nov. 2002.
- [14] M. Secondini and E. Forestieri, "All-order PMD outage probability evaluation by Markov chain Monte Carlo simulations," *IEEE Photon. Technol. Lett.*, vol. 17, no. 7, pp. 1417–1419, Jul. 2005.
- [15] Z. Marzec, J. Schuster, and G. Biondini, "On the efficiency of importance sampling techniques for polarization-mode dispersion in optical fiber transmission systems," *SIAM J. Appl. Math.*, vol. 73, pp. 155–17, 2013.
- [16] E. Veach, "Robust Monte Carlo methods for light transport simulation," Ph.D. dissertation, Dept. Comput. Sci., Stanford Univ., Stanford, CA, USA, 1997.
- [17] A. Owen and Y. Zhou, "Safe and effective importance sampling," *J. Amer. Statist. Assoc.*, vol. 95, pp. 135–143, 2000.
- [18] P. J. Winzer, H. Kogelnik, and K. Ramanan, "Precise outage specifications for first-order PMD," *IEEE Photon. Technol. Lett.*, vol. 16, no. 2, pp. 449–451, Feb. 2004.
- [19] J. Renaudier, G. Charlet, M. Salsi, O. B. Pardo, H. Mardoyan, P. Tran, and S. Bigo, "Linear fiber impairments mitigation of 40-Gbit/s polarization-multiplexed QPSK by digital processing in a coherent receiver," *J. Lightw. Technol.*, vol. 26, no. 1, pp. 36–42, Jan. 2008.
- [20] C. Xie, "Polarization-mode-dispersion impairments in 112 Gb/s PDM-QPSK coherent systems," presented at the 36th Eur. Conf. Exhib. Opt. Commun., Torino, Italy, Paper Th.10.E.6
- [21] M. Boroditsky, M. Brodsky, N. Frigo, P. Magill, A. Mecozzi, and C. Antonelli, "Outage probabilities for fiber routes with finite number of degrees of freedom," *IEEE Photon. Technol. Lett.*, vol. 17, no. 2, pp. 345–347, Feb. 2005.
- [22] A. Mecozzi, C. Antonelli, M. Boroditsky, and M. Brodsky, "Characterization of the time dependence of polarization mode dispersion," *Opt. Lett.*, vol. 29, no. 22, pp. 2599–2601, Nov. 2004.
- [23] R. O. Moore, G. Biondini, and W. L. Kath, "A method for computing statistics of large noise-induced perturbations of nonlinear Schrodinger solitons," *SIAM Rev.*, vol. 50, pp. 523–549, 2008.

Authors' biographies not available at the time of publication.