

A Comparative Study of Single-Section Polarization-Mode Dispersion Compensators

Ivan T. Lima, Jr., Aurenice O. Lima, *Student Member, IEEE*, Gino Biondini, Curtis R. Menyuk, *Fellow, IEEE, Fellow, OSA*, and William L. Kath, *Member, OSA*

Abstract—This paper shows how to use multiple importance sampling to study the performance of polarization-mode dispersion (PMD) compensators with a single differential group delay (DGD) element. We compute the eye opening penalty margin for compensated and uncompensated systems with outage probabilities of 10^{-5} or less with a fraction of the computational cost required by standard Monte Carlo methods. This paper shows that the performance of an optimized compensator with a fixed DGD element is comparable to that of a compensator with a variable DGD element. It also shows that the optimal value of the DGD compensator is two to three times larger than the mean DGD of the transmission line averaged over fiber realizations. This technique can be applied to the optimization of any PMD compensator whose dominant sources of residual penalty are both the DGD and the length of the frequency derivative of the polarization-dispersion vector.

Index Terms—Birefringence, compensation, optical communications, optical fiber polarization, polarization-mode dispersion (PMD).

I. INTRODUCTION

THERE is substantial interest in upgrading the current per channel data rates to 10 Gb/s and beyond in terrestrial wavelength-division multiplexed (WDM) systems. Polarization-mode dispersion (PMD) is a significant barrier to achieving this goal. Designers want to ensure that the probability of an eye opening penalty due to PMD beyond some value occurs only a very small fraction of the time. For example, a designer might require that a penalty larger than 1 dB occurs with probability 10^{-5} or less. Therefore, there have been numerous proposals to use optical and electrical PMD compensators to mitigate this problem [1]–[9], and much of this work has focused on compensators with a single differential group delay (DGD) element because they are the simplest to build, to

control, and to analyze. In [9], we used importance sampling in which the DGD alone was biased [10] to show that the average reduction of the eye opening penalty in compensators does not address the issue of greatest practical importance, which is the increase of the penalty margin for a given outage probability. Here, we extend the work in [9] by using multiple importance sampling in which both the DGD and the length of the frequency derivative of the polarization-dispersion vector are biased [11], and, for the first time, we completely describe the implementations of importance sampling that we use. We focus on compensators with a single DGD element due to their simplicity and practical importance [1], [3], [12].

The main idea in any biased Monte Carlo simulation, including simulations based on importance sampling, is to cause the events that contribute to the statistical quantities of interest to occur more frequently and thus to reduce the relative variation in the numerical estimate of those quantities with a fixed number of samples [13]. In our case, the quantities of interest are the eye opening penalties and their probability density functions (pdfs) both before and after compensation. As is desirable with any Monte Carlo simulation, we monitor its effectiveness by calculating the standard deviation divided by the mean in the quantities of interest. We show that biasing the DGD is sufficient to accurately calculate the uncompensated penalties and their pdfs, but it is not sufficient to accurately calculate the compensated penalties and their pdfs. For the compensators that we consider in this paper, we also show that biasing both the DGD and the length of the frequency derivative of the polarization-dispersion vector is sufficient to accurately calculate the compensated penalties and their pdfs.

In Section II, we describe the fiber transmission model that we use. In Section III, we describe the PMD compensators that we study, and describe how we optimize the compensators and how we determine the eye opening penalty. In Section IV, we describe the implementation of importance sampling in which we only bias the DGD. We previously used this implementation in [9]. In Section V, we describe the implementation of multiple importance sampling in which we bias both the DGD and the length of the frequency derivative of the polarization-dispersion vector. In Section VI, we show how we combine the samples from several biasing distributions using importance sampling to obtain the statistical results that we present in this work. Finally, in Section VII, we apply the techniques that we develop in this paper to efficiently compute the outage probability of single-section PMD compensators and to optimize the constant DGD element in fixed-DGD compensators.

Manuscript received May 6, 2003. The work of A. O. Lima and C. R. Menyuk was supported by the Department of Energy and the National Science Foundation. The work of W. L. Kath was supported by Air Force Office of Scientific Research and the National Science Foundation.

I. T. Lima, Jr., was with the Department of Computer Science & Electrical Engineering, University of Maryland Baltimore County. He is now with the Department of Electrical and Computer Engineering, North Dakota State University, ECE 101, Fargo, ND 58105-5285 USA.

A. O. Lima and C. R. Menyuk are with the Department of Computer Science & Electrical Engineering, University of Maryland Baltimore County, TRC 205A, Baltimore, MD 21250 USA (e-mail: alima1@umbc.edu).

G. Biondini is with the Department of Mathematics, Ohio State University, Columbus, OH 43210 USA (e-mail: biondini@math.ohio-state.edu).

W. L. Kath is with the Department of Engineering Sciences and Applied Mathematics, McCormick School of Engineering and Applied Sciences, Northwestern University, Evanston, IL 60208-3125 USA (e-mail: kath@northwestern.edu).

Digital Object Identifier 10.1109/JLT.2004.825895

II. SYSTEM PARAMETERS AND FIBER TRANSMISSION MODEL

We simulate a 10 Gb/s nonreturn-to-zero (NRZ) system with 30 ps of rise time. The pulses are generated by perfect rectangular pulses filtered by a Gaussian shape filter that produces the designed rise time. The results of our simulations can also be applied to 40 Gb/s systems by scaling down the time quantities by a factor of four. The PMD simulation model that we use is based on the representation of the fiber link by a frequency-dependent transfer Jones matrix $\bar{T}(\omega)$, which corresponds to the Müller matrix $T(\omega)$. We do not take into account polarization-dependent loss, chromatic dispersion, or fiber nonlinearity. The complex envelope of the electrical field vector at the end of the fiber link, $\mathbf{E}_{\text{out}}(\omega)$, equals

$$\mathbf{E}_{\text{out}}(\omega) = \bar{T}(\omega) \mathbf{E}_{\text{in}}(\omega) \quad (1)$$

where $\mathbf{E}_{\text{in}}(\omega)$ is the input field vector in the Jones space. Using the coarse step model of a fiber [14], the Jones transfer matrix $\bar{T}(\omega)$ of an optical fiber that consists of N linearly birefringent sections may be written as [14]

$$\bar{T}(\omega) = \prod_{n=1}^N \bar{T}_n(\omega) \quad (2)$$

where

$$\bar{T}_n(\omega) = \bar{P}(\omega) \bar{S}_n \quad (3)$$

is the transmission matrix of the n th fiber section, \bar{S}_n models the random mode coupling of the n th birefringent fiber section in the unbiased PMD simulation model, which is shown in (4) at the bottom of the page, and

$$\bar{P}(\omega) = \begin{pmatrix} \exp\left(\frac{-i\omega\tau_s}{2}\right) & 0 \\ 0 & \exp\left(\frac{i\omega\tau_s}{2}\right) \end{pmatrix} \quad (5)$$

models the propagation of the light through a birefringent section. The parameter τ_s is the DGD in a single section, while ξ_n , ψ_n , and ϕ_n are random variables that are independent at each n and from each other. The pdfs of the angles ψ_n and ϕ_n are uniformly distributed between 0 and 2π , while the pdf of the $\cos \xi_n$ are uniformly distributed between -1 and 1 . The Müller matrix S_n that is equivalent to the Jones matrix \bar{S}_n in (4) is comprised of elementary rotations around the three orthogonal axes [15] of the Poincaré sphere,

$$S_n = R_x(\psi_n) R_y(\xi_n) R_x(\phi_n) \quad (6)$$

which can produce a uniform rotation on the Poincaré sphere [14]. In (6), $R_x(\psi_n)$ is a rotation around the x -axis, and $R_y(\xi_n)$ is a rotation around the y -axis. Likewise, the Müller matrix

$P(\omega)$ equivalent to the Jones matrix $\bar{P}(\omega)$ in (5) is comprised of an elementary rotation around the x -axis,

$$P(\omega) = R_x(-\omega\tau_s). \quad (7)$$

Since the Müller matrix of a section is equal to $P(\omega)S_n$, the polarization-dispersion version of a single section is given by [16]

$$\boldsymbol{\tau}_s = -\tau_s \hat{\mathbf{a}}_x \quad (8)$$

where $\hat{\mathbf{a}}_x$ is a unit vector along the x -axis.

The formulation of (3) is consistent with the one in [16], where the random mode coupling in the n th section occurs prior to the birefringent element of that section. We set τ_s equal to

$$\tau_s = \sqrt{\frac{3\pi}{8N}} \langle \tau \rangle \quad (9)$$

where $\langle \tau \rangle = \langle \tau^{(N)} \rangle$ is the mean DGD of the fiber with N sections [17]. In this work, we emulate an optical fiber with eighty birefringent sections $N = 80$. In [9], we showed that $N = 80$ is sufficient to obtain a Maxwellian distribution of the DGD in the outage probability range of 10^{-5} to 10^{-6} . We assume that the fiber passes ergodically through all possible orientations of the birefringence.

III. PMD COMPENSATOR AND RECEIVER MODEL

In order to compensate for PMD distortions, we use a compensator with an arbitrarily rotatable polarization controller and a single DGD element, which can be fixed [1] or variable [12]. The expression for the polarization-dispersion vector after compensation is similar to the one in (16),

$$\boldsymbol{\tau}_{\text{tot}}(\omega) = \boldsymbol{\tau}_c + T_c(\omega) R_{\text{pc}} \boldsymbol{\tau}_f(\omega) \quad (10)$$

where $\boldsymbol{\tau}_c$ is the polarization-dispersion vector of the compensator, $\boldsymbol{\tau}_f(\omega)$ is the polarization-dispersion vector of the transmission fiber, R_{pc} is the polarization transformation in the Stokes space that is produced by the polarization controller of the compensator, and $T_c(\omega)$ is the polarization transformation in the Stokes space that is produced by the DGD element of the compensator, which is similar to (7). We model the polarization transformation R_{pc} as

$$R_{\text{pc}} = R_x(\phi_{\text{pc}}) R_y(\psi_{\text{pc}}) R_x(-\phi_{\text{pc}}). \quad (11)$$

We note that the two parameters of the polarization controller's angles in (11) are the only free parameters that a compensator with a fixed DGD element possesses, while the value of the DGD element of a variable DGD compensator is an extra free parameter that needs to be adjusted during the operation. In (11), the parameter ϕ_{pc} is the angle that determines the axis of polarization rotation in the $y-z$ plane of the Poincaré sphere,

$$\bar{S}_n = \begin{pmatrix} \cos\left(\frac{\xi_n}{2}\right) \exp\left[i\frac{(\psi_n + \phi_n)}{2}\right] & i \sin\left(\frac{\xi_n}{2}\right) \exp\left[i\frac{(\psi_n - \phi_n)}{2}\right] \\ i \sin\left(\frac{\xi_n}{2}\right) \exp\left[-i\frac{(\psi_n - \phi_n)}{2}\right] & \cos\left(\frac{\xi_n}{2}\right) \exp\left[-i\frac{(\psi_n + \phi_n)}{2}\right] \end{pmatrix} \quad (4)$$

while parameter ψ_{pc} is the angle of rotation around that axis of polarization rotation. An appropriate selection of these two angles will transform an arbitrary input Stokes vector into a given output Stokes vector. While most electronic polarization controllers have two or more parameters to adjust that are different from ϕ_{pc} and ψ_{pc} , it is possible to configure them to operate according to the transformation matrix R_{pc} in (11).

Throughout this paper, we use the eye opening as the feedback parameter for the optimization algorithm unless otherwise stated. We define the eye opening as the difference between the lowest mark and the highest space at the decision time in a noise-free bit string. The eye opening penalty is defined as the ratio between the back-to-back and the PMD-distorted eye opening. Since PMD causes pulse spreading in amplitude-shift keyed modulation formats, the isolated marks and spaces are the ones that suffer the highest penalty [18]. To define the decision time, we recovered the clock using an algorithm based on one described by Trischitta and Varma [19]. We simulated 16 bit strings of the form “0 100 100 101 101 101.” The receiver is modeled by a square-law photodetector followed by a fifth-order electrical Bessel filter with a 3 dB width of 8.6 GHz. After the electronic receiver, we delayed the bit stream by half bit slot and subtracted it from the original stream, which is then squared. As a result a strong tone is produced at 10 GHz. The decision time is set equal to the time at which the phase of the tone is equal to $\pi/2$.

The goal of our study is to determine the performance limit of the compensators. We therefore show the global optimum of the compensated feedback parameter for each fiber realization. To obtain the optimum, we start with 5 evenly spaced initial values for each of the angles ϕ_{pc} and ψ_{pc} in the polarization transformation matrix R_{pc} . If the DGD of the compensator is adjustable, we start the optimization with the DGD of the compensator equal to the DGD of the fiber. We then apply the conjugate gradient algorithm [20] to each of these 25 initial polarization transformations. To ensure that this procedure yields the global optimum, we studied the convergence as the number of initial polarization transformations is increased. We examined 10^4 fiber realizations spread throughout our phase space, and we never found more than 12 local optima in the cases that we examined. In only three of these cases, we missed the global optimum, because several optima were closely clustered, but the penalty difference was small. We therefore concluded that 25 initial polarization transformations were sufficient to obtain the global optimum with sufficient accuracy for our purposes. We observed that the use of the eye opening as the objective function for the conjugate gradient produces multiple optimum values when both the DGD and the length of the frequency derivative of the polarization dispersion vector are very large.

The performance of the compensator depends on how the DGD and the effects of the first- and higher-order frequency derivatives of the polarization-dispersion vector of the transmission fiber interact with the DGD element of the compensator to produce a residual polarization dispersion vector and on how the signal couples with the residual principal states of polarization over the spectrum of the channel. Therefore, the operation of PMD compensators is a compromise between reducing the DGD and setting one principal state of polarization after com-

ensation that is approximately copolarized with the signal. An expression for the pulse spreading due to PMD as a function of the polarization-dispersion vector of the transmission fiber and the polarization state over the spectrum of the signal was given in [21].

IV. IMPORTANCE SAMPLING BIASING THE DGD

When we use importance sampling to bias the DGD, we are taking advantage of the large correlation that exists between the PMD-induced penalty and the DGD. We note that first- and higher-order frequency derivatives of the polarization-dispersion vector are included in the simulations, although this approach does not produce larger values of first- and higher-order frequency derivatives of the polarization-dispersion vector than the moderately large values that are naturally obtained when one biases the DGD.

To apply the multiple importance sampling technique, we first recall that P_I , the probability of an event defined by the indicator function $I(\mathbf{x})$, can be estimated as [22]

$$\hat{P}_I = \sum_{j=1}^J \frac{1}{M_j} \sum_{i=1}^{M_j} I(\mathbf{x}_{j,i}) w_j(\mathbf{x}_{j,i}) L_j(\mathbf{x}_{j,i}) \quad (12)$$

where

$$L_j(\mathbf{x}_{j,i}) = \frac{p(\mathbf{x}_{j,i})}{p_j^*(\mathbf{x}_{j,i})} \quad (13)$$

is the likelihood ratio of the i th sample $\mathbf{x}_{j,i}$ drawn from the j th biasing distribution, and where M_j is the number of samples drawn from the j th biasing distribution $p_j^*(\mathbf{x})$. The term $p(\mathbf{x})$ is the pdf of the unbiased distribution, and J is the number of different biasing distributions. The weights $w_j(\mathbf{x})$ allow one to combine different biasing distributions and are defined in Section VI. The indicator function $I(\mathbf{x})$ in (12) is chosen to compute the probability of having an eye opening penalty within any range, such as a bin in a histogram. Thus, $I(\mathbf{x})$ is defined as 1 inside the desired penalty range and 0 otherwise. A confidence interval for the estimator of the probability P_I of the indicator function $I(\mathbf{x})$ in (12) can be defined from the estimator of the variance of \hat{P}_I , which is given by

$$\hat{\sigma}_{\hat{P}_I}^2 = \sum_{j=1}^J \frac{1}{M_j(M_j-1)} \sum_{i=1}^{M_j} [I(\mathbf{x}_{j,i}) w_j(\mathbf{x}_{j,i}) L_j(\mathbf{x}_{j,i}) - \hat{P}_I]^2 \quad (14)$$

where

$$\hat{P}_{I_j} = \frac{1}{M_j} \sum_{i=1}^{M_j} I(\mathbf{x}_{j,i}) w_j(\mathbf{x}_{j,i}) L_j(\mathbf{x}_{j,i}) \quad (15)$$

is the contribution of the samples drawn from j th biasing distribution to the estimator \hat{P}_I . The confidence interval of the estimator \hat{P}_I equals the range $[\hat{P}_I - \hat{\sigma}_{\hat{P}_I}, \hat{P}_I + \hat{\sigma}_{\hat{P}_I}]$. The relative variation equals $\hat{\sigma}_{\hat{P}_I} / \hat{P}_I$.

The polarization-dispersion vector after n fiber sections is determined by the following concatenation rule [23]:

$$\boldsymbol{\tau}^{(n)} = \boldsymbol{\tau}_n + T_n \boldsymbol{\tau}^{(n-1)} \quad (16)$$

where T_n is the equivalent Müller matrix of the n th section in (3). Biondini *et al.* [10] demonstrated that the appropriate parameters to bias are the angles θ_n between the polarization-dispersion vector of the first $(n-1)$ sections $\boldsymbol{\tau}^{(n-1)}$ and the polarization-dispersion vector of the n th section $\boldsymbol{\tau}_n$ at the center frequency of the channel, such that $\cos \theta_n$ is biased toward one, thereby increasing the probability that the polarization-dispersion vector at that frequency will lengthen after that section. In other words, $\boldsymbol{\tau}_n$ is biased toward a direction $\hat{\mathbf{b}}$ that is equal to the direction of $\boldsymbol{\tau}^{(n-1)}$. In standard Monte Carlo simulations using the coarse step method, the pdf of $\cos \theta_n$, $p(\cos \theta)$, is uniformly distributed. Therefore, we have $p(\cos \theta) = 1/2$, where $\cos \theta \in [-1, 1]$. The angles θ_n are directly determined by the realization of the random mode coupling between the birefringent sections. Thus, the values of $\cos \theta_n$ play the role of the components of the random vector \mathbf{x} in (12). Specifically, we pick $\cos \theta_n$ from the following pdf [24]:

$$p_\alpha^*(\cos \theta) = \frac{\alpha}{1 - e^{-2\alpha}} e^{-\alpha(1 - \cos \theta)} \quad (17)$$

which biases $\cos \theta$ toward 1 when α is positive, and corresponds to standard Monte Carlo simulations in the limit $\alpha = 0$. The likelihood ratio for each value of $\cos \theta$ that is obtained from the biasing pdf in (17) is given by

$$L_\alpha(\cos \theta) = \frac{1 - e^{-2\alpha}}{2\alpha} e^{\alpha(1 - \cos \theta)}. \quad (18)$$

Since the values of $\cos \theta_n$ are independent random variables, the likelihood ratio for a biased realization of the fiber PMD is equal to the product of the likelihood ratios for each of its biased angles

$$L_j(\mathbf{x}_{j,i}) = \prod_{n=1}^N L_{\alpha_j}(\cos \theta_{j,i,n}) \quad (19)$$

where α_j is the amount of bias used in the j th distribution, and $\cos \theta_{j,i,n}$ is $\cos \theta$ for the n th section of the i th sample obtained from the j th distribution.

We must determine the value of the biasing parameter α that enables us to statistically resolve the histogram of the eye opening penalty over a range of large eye opening penalty values whose probability is on the order of a given target probability P_α , such as $P_\alpha = 10^{-5}$. Intuitively, we anticipate that the required value of α is the one for which the target probability P_α is equal to the likelihood ratio of the biased realization of the fiber PMD evaluated at the expected value of the random variable $\cos \theta$ with biasing pdf $p_\alpha^*(\cos \theta)$. That is, the parameter α satisfies the equation

$$P_\alpha = L_\alpha^N[\langle \cos \theta \rangle] = L_\alpha^N \left(\frac{1 + e^{-2\alpha}}{1 - e^{-2\alpha}} - \frac{1}{\alpha} \right) \quad (20)$$

where N is the number of fiber sections, and $\langle \cdot \rangle$ is the expectation operator. Our motivation for this heuristic comes from (12) and from the observation that, over a given range of penalties, the biased samples statistically resolve the histogram of the eye opening penalty when the indicator function for this range has the value 1 for a large proportion of the biased samples. For the

simulation results in this paper in which we only bias the DGD, we chose $\alpha = 0$, which produces unbiased samples, together with $\alpha = 0.5$, and $\alpha = 1.0$. The target probabilities of the biased simulations are $P_{0.5} = 3.9 \times 10^{-2}$ and $P_{1.0} = 5.4 \times 10^{-6}$. We use 10^4 samples for each of the three biases except as noted. As we increase the number of samples with bias parameter α , the size of the interval about P_α for which the histogram of the eye opening penalty is well resolved increases.

In order to bias the angle θ_n between the polarization dispersion vector of the first $(n-1)$ sections $\boldsymbol{\tau}^{(n-1)}$ and the polarization-dispersion vector of the n th section $\boldsymbol{\tau}_n$, the polarization rotation matrix S_n in (6) has to be modified to properly account for the bias. In practice, we bias the direction of the polarization-dispersion of the previous $(n-1)$ sections $\boldsymbol{\tau}^{(n-1)}$ toward $\boldsymbol{\tau}_n$ because the polarization-dispersion vector of the previous sections $\boldsymbol{\tau}^{(n-1)}$ is the vector rotated by the matrix T_n , as shown in (16). The polarization-dispersion vector of any of the sections of the transmission fiber modeled by (3) is given by $\boldsymbol{\tau}_s = -\tau_s \hat{\mathbf{a}}_x$, as in (8), which is independent of S_n . Therefore, the matrix S_n of the n th section must bias the polarization-dispersion vector of the previous sections toward the vector $-\hat{\mathbf{a}}_x$. We obtain this bias by replacing the first two random rotations in (6) by the combination of one random rotation with two deterministic rotations around the y - and the z -axes. The first rotation $R_y(\chi_n)$ eliminates the z -component of the polarization-dispersion vector of the previous $(n-1)$ sections $\boldsymbol{\tau}^{(n-1)}$, which is accomplished by choosing $\chi_n = \arctan(\tau_z/\tau_x)$. The second rotation $R_z(\zeta_n)$ eliminates the y -component of $R_y(\chi_n)\boldsymbol{\tau}^{(n-1)}$, where ζ_n is chosen like χ_n , with the additional constraint that the resultant vector $R_z(\zeta_n)R_y(\chi_n)\boldsymbol{\tau}^{(n-1)}$ should be in the $-\hat{\mathbf{a}}_x$ direction. Then, we chose a random angle θ_n from the biasing pdf in (17) to rotate $R_z(\zeta_n)R_y(\chi_n)\boldsymbol{\tau}^{(n-1)}$ around the z -axis. This rotation can be combined with the previous deterministic rotation around the z -axis by ζ_n to produce a single rotation. Finally, we add a uniformly distributed random rotation ψ_n around the x -axis to obtain the polarization rotation matrix $S_n^{(1)}$, which becomes

$$S_n^{(1)} = R_x(\psi_n)R_z(\theta_n + \zeta_n)R_y(\chi_n). \quad (21)$$

A uniform rotation like S_n in (4) could in principle be added at the end of the fiber so that the direction of $\boldsymbol{\tau}^{(N)}$ is uniformly distributed on the Poincaré sphere. However, our receiver model has no polarization dependence; so, this final rotation is unnecessary.

In Fig. 1, we show the pdf of the normalized DGD, $\tau/\langle \tau \rangle$, of a fiber with 80 birefringent sections and 30 ps of mean DGD $\langle \tau \rangle$, where $\tau = |\boldsymbol{\tau}|$ and the DGD is normalized with respect to $\langle \tau \rangle$. Hence, these results are independent of $\langle \tau \rangle$. The unbiased probability of obtaining normalized DGD values outside the domain [0,4] that we show in Fig. 1 is less than 10^{-8} . We obtained this curve with only 10^4 samples from Monte Carlo simulations for each of the three biasing distributions: $\alpha = 0$, which produces unbiased samples, $\alpha = 0.5$, and $\alpha = 1.0$. We combined the results of the biasing distributions using the balanced heuristic method that we describe in Section VI. The largest relative variation over the domain [0,3,4] is 8%. We observed an excel-

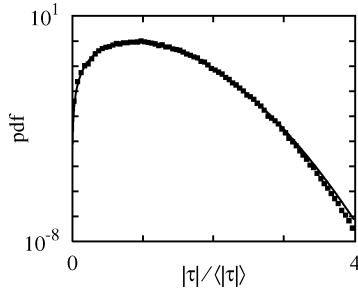


Fig. 1. The pdf of the normalized DGD, $|\tau|/\langle|\tau|\rangle$, plotted on a logarithmic scale with 80 bins. The squares show the results of Monte Carlo simulations with importance sampling with DGD bias using 3×10^4 samples. The solid line shows the Maxwellian distribution with the same mean.

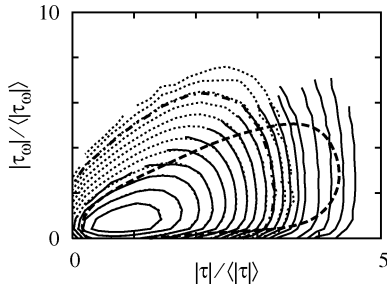


Fig. 2. Joint pdf of the DGD and the length of the frequency derivative of the polarization-dispersion vector with 25×25 bins. The solid lines show the results of Monte Carlo simulations with importance sampling with DGD bias using 3 biases with 6×10^4 samples in each bias. We only show the results with importance sampling where the relative variation does not exceed 25%. The dashed line shows the contour level corresponding to 10% relative variation in the results using importance sampling. We applied the Bezier smoothing algorithm [28] in the contour level of the relative variation. The dotted lines show the results of standard Monte Carlo simulations using 10^8 samples. The dot-dashed line shows the contour level corresponding to 10% relative variation in the results using standard Monte Carlo simulations with 10^8 samples. The contours of the joint pdf are at 3×10^{-1} , 10^{-1} , 3×10^{-2} , 10^{-2} , 3×10^{-3} , 10^{-3} , 3×10^{-4} , 10^{-4} , 3×10^{-5} , 10^{-5} , 3×10^{-6} , 10^{-6} , 3×10^{-7} , 10^{-7} , 10^{-8} , 10^{-9} , 10^{-10} , and 10^{-11} .

lent agreement between the numerically calculated pdf of the DGD obtained with importance sampling and the Maxwellian pdf with the same mean. The slight deviation in the tail between the numerically calculated pdf and the Maxwellian distribution occurs because we use 80 sections rather than a larger number [25].

In Fig. 2, we compare the joint pdf of the DGD and the length of the frequency derivative of the polarization-dispersion vector that is obtained with the implementation of the importance sampling with DGD bias to standard Monte Carlo simulations with 10^8 samples. We had the same configuration as in Fig. 1, except that we used 6×10^4 samples per bias. We observe that the length of the frequency derivative of the polarization-dispersion vectors that are statistically correlated to the DGD that we bias are correctly accounted for. However, this implementation is not efficient in obtaining samples with large lengths of the frequency derivative of the polarization-dispersion vector associated with moderately small values of DGD. Hence, the use of DGD bias is limited to systems where the DGD is the dominant source of penalties, which is the case in uncompensated systems and in systems with limited PMD compensation.

V. IMPORTANCE SAMPLING BIASING BOTH THE DGD AND THE LENGTH OF THE FREQUENCY DERIVATIVE OF THE POLARIZATION-DISPERSION VECTOR

The derivative of the polarization-dispersion vector with respect to the angular frequency ω after n fiber sections $\tau_\omega^{(n)}$, is determined by the following concatenation rule [23],

$$\tau_\omega^{(n)} = \tau_{n\omega} + \tau_n \times \left(T_n \tau_\omega^{(n-1)} \right) + T_n \tau_\omega^{(n-1)} \quad (22)$$

where T_n is the Müller matrix that is equivalent to the Jones matrix in (3), and $\tau_{n\omega}$ is the derivative of the polarization dispersion vector of the n th section with respect to the angular frequency. Note that $|\tau_{n\omega}| = 0$ in this problem, since each section has a constant—frequency independent—polarization-dispersion vector τ_n .

In order to obtain large values of the frequency derivative of the polarization-dispersion vector τ_ω with a relatively small number of Monte Carlo simulations, Fogal *et al.* [11] demonstrated that it is necessary to bias the polarization-dispersion vector of the n th section in a direction $\hat{\mathbf{b}}$ that is different from the direction used in the DGD bias that was described in Section IV. The biasing direction $\hat{\mathbf{b}}$ is located in a plane that contains the vectors $\tau^{(n-1)}$ and $\tau^{(n-1)} \times \tau_\omega^{(n-1)}$, and this direction is chosen so that the angle between the biasing direction $\hat{\mathbf{b}}$ and the polarization-dispersion vector of the previous sections $\tau^{(n-1)}$ varies linearly along the fiber sections from $\beta_1 = 0$ in the first fiber section to $\beta_N = \beta$ in the last fiber section, where the values of β and α in (17) determine the region in the plane formed by the DGD $|\tau|$ and the length of the frequency derivative of the polarization-dispersion vector $|\tau_\omega|$ that one wants to statistically resolve. Specifically [11], we choose

$$\beta_n = \frac{n}{N} \beta \quad (23)$$

where $0 \leq \beta \leq \pi$. Note that the choice $\beta = 0$ produces only DGD bias. However, the parameter α completely determines the target probability (20), since the parameter β simply selects an equiprobable region in the parameter space. For the simulation results in this paper with both the DGD and the length of the frequency derivative of the polarization-dispersion vector bias we chose the following values of (α, β) to bias the distributions with 10^4 samples each: $(0, 0)$, which produces unbiased samples, $(0.5, 0)$, $(0.5, \pi/3)$, $(0.5, 2\pi/3)$, $(0.5, \pi)$, $(1, 0)$, $(1, \pi/3)$, $(1, 2\pi/3)$, $(1, \pi)$, and $(0.7, 0)$.

In order to implement the bias for both the DGD and the length of the frequency derivative of the polarization-dispersion vector, the polarization rotation matrix S_n in (6) has to be modified in a way that is analogous to the derivation of $S_n^{(1)}$ in Section IV. The goal is to choose a set of rotations so that the vector $T_n^{-1} \tau_n$ ends up at angle θ with the biasing direction $\hat{\mathbf{b}}$ in the three-dimensional (3-D) Stokes space. The first step is similar to the one described in Section IV, where two deterministic rotations are obtained to rotate the polarization dispersion vector of the previous $(n-1)$ sections $\tau^{(n-1)}$ to the $-\hat{\mathbf{a}}_x$ direction, $R_z(\zeta_n) R_y(\chi_n) \tau^{(n-1)}$. Then, a rotation around the x -axis $R_x(\nu_n)$ eliminates the y -component of $R_z(\zeta_n) R_y(\chi_n) \tau^{(n-1)}$, while leaving the z -component positive. The next step is to apply a deterministic rotation $R_z(\beta_n)$

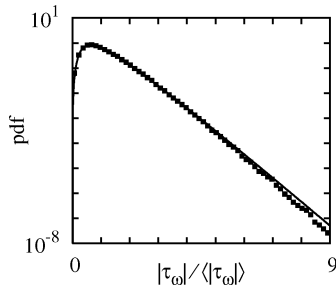


Fig. 3. The pdf of the normalized length of the frequency derivative of the polarization-dispersion vector plotted on a logarithmic scale with 60 bins. The squares show the results of Monte Carlo simulations with importance sampling biasing both the DGD and the length of the frequency derivative of the polarization-dispersion vector using 10^5 samples. The solid line shows the results of the theoretical distribution of the length of the frequency derivative of the polarization-dispersion vector.

so that $[R_z(\beta_n) R_x(\nu_n) R_z(\zeta_n) R_y(\chi_n)]^{-1} \tau_n$ is parallel to the biasing direction \mathbf{b} , where β_n is determined by (23). Then, a uniformly distributed rotation around the x -axis $R_x(\psi_n)$ is added to produce the correct statistical randomization of $S_n^{(2)}$. Finally, a biased rotation around the z -axis, $R_z(\theta_n)$, is applied to obtain an appropriate bias for both the DGD and the length of the frequency derivative of the polarization-dispersion vector, where $\cos \theta_n$ is obtained from the pdf in (17). The matrix $S_n^{(2)}$ in this case becomes

$$S_n^{(2)} = R_z(\theta_n) R_x(\psi_n) R_z(\beta_n) R_x(\nu_n) R_z(\zeta_n) R_y(\chi_n) \quad (24)$$

where ψ_n is a random variable whose pdf is uniformly distributed between 0 and 2π as in (21). Note that none of the angles χ_n , ζ_n , ν_n , and β_n in (24) are random, and that these rotations are not unique; it is possible to produce a bias for both the DGD and the length of the frequency derivative of the polarization-dispersion vector using a different set of elementary rotations. A uniform rotation like S_n in (4) could be added at the end of the fiber model to make sure that $\tau^{(N)}$ is uniformly distributed on the Poincaré sphere. However, this extra rotation is not necessary here, just as in the case of DGD bias alone.

In Fig. 3, we show the pdf of the length of the frequency derivative of the polarization-dispersion vector τ_ω of a fiber with 80 birefringent sections and 10 ps of mean DGD $\langle |\tau| \rangle$. We show the length of τ_ω normalized with respect to its expected value $\langle |\tau_\omega| \rangle$. In Fig. 3, we combine the results of the 10 biasing distributions with 10^4 samples per bias using the balanced heuristic method that we describe in Section VI. The largest relative variation over the domain [0,9] is 17%. We observed an excellent agreement between the numerically calculated pdf of the length of the frequency derivative of the polarization-dispersion vector obtained using importance sampling with the theoretical pdf of the length of the frequency derivative of the polarization-dispersion vector [26].

In Fig. 4, we show the results of the joint pdf of the DGD and the length of the frequency derivative of the polarization-dispersion vector that is obtained with Monte Carlo simulations with our implementation of multiple importance sampling biasing both the DGD and the length of the frequency derivative of the polarization-dispersion vector in comparison with results

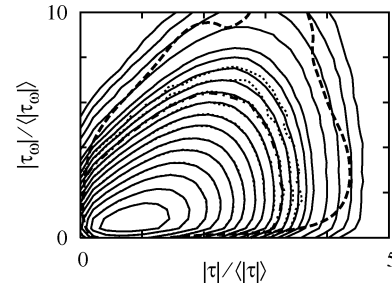


Fig. 4. The joint pdf of the DGD and the length of the frequency derivative of the polarization-dispersion vector with 25×25 bins. The solid lines show the results of Monte Carlo simulations with importance sampling biasing both the DGD and the length of the frequency derivative of the polarization-dispersion vector using 10 biases with 6×10^4 samples in each bias. The dashed line shows the contour level corresponding to 10% relative variation in the results using importance sampling. We applied the Bezier smoothing algorithm [28] in the contour level of the relative variation. This curve is distorted by the limited number of biases. It is farthest out in directions corresponding to the biases and moves inward in between these directions. The dotted lines show the results of standard Monte Carlo simulations using 10^8 samples. The dot-dashed line shows the contour level corresponding to 10% relative variation in the results using standard Monte Carlo simulations with 10^8 samples. In all directions, the 10% relative variation level is farther out for the biased Monte Carlo simulations than for the standard simulations, indicating the effectiveness of the biasing procedure. The contours of the joint pdf are at 3×10^{-1} , 10^{-1} , 3×10^{-2} , 10^{-2} , 3×10^{-3} , 10^{-3} , 3×10^{-4} , 10^{-4} , 3×10^{-5} , 10^{-5} , 3×10^{-6} , 10^{-6} , 3×10^{-7} , 10^{-7} , 10^{-8} , 10^{-9} , 10^{-10} , and 10^{-11} .

of standard Monte Carlo simulations with 10^8 samples. We had the same configuration as in Fig. 3, except that we used 6×10^4 samples per bias. We observed an excellent agreement between these results. We point out that the relative variation in the joint pdf of the DGD and the length of the frequency derivative of the polarization-dispersion vector in the results of standard Monte Carlo simulations depends only on the number of samples used. In addition to the number of samples, the relative variation in the results with importance sampling strongly depends on the set of biases that are combined to produce the numerical joint pdf. As a consequence, the contours of relative variation do not follow the probability contour lines and have a bumpy structure. We show this behavior in Fig. 4 for the 10% relative variation contour. However, we see that the 10% relative variation contour for the biasing simulations lies well outside the 10% relative variation contour for the standard simulations, although the standard simulations have a much larger number of samples. This result supports the conclusion that the biasing procedure is effective.

VI. COMBINATION OF MULTIPLE BIASING DISTRIBUTIONS

In this section, we describe two heuristic methods to combine samples from different biasing distributions in order to accurately determine the eye opening penalty or any other quantity that is dependent on PMD. We observed that one single biasing distribution is insufficient to accurately determine the eye opening penalty over the entire range of interest in the $\tau - |\tau_\omega|$ plane. It is necessary to combine the statistical result from multiple biasing distributions, since each distribution provides accuracy in different regions of the parameter space. However, the biasing distributions that do not have a low variance in a given region in the parameter space can degrade significantly the accuracy of the estimate of the probability in that region if they are inadequately combined with the other distributions.

This effect is particularly evident in the computation of the eye opening penalty because the polarization-dispersion vector $\boldsymbol{\tau}$ and its frequency derivative $\boldsymbol{\tau}_\omega$ are not the only quantities that determine the eye opening penalty. The second- and higher-order frequency derivatives of $\boldsymbol{\tau}$ and the coupling factor between the polarization state of the signal and the principal states of polarization at the central frequency of the channel are other important factors.

In (12), we described how the probability of an event P_I defined by the indicator function $I(\mathbf{x})$ can be estimated with the combination of several Monte Carlo simulations with importance sampling. In addition to the likelihood ratio $L_j(\mathbf{x})$ in (19), a weight $w_j(\mathbf{x})$ is assigned to each sample from each of the J biasing distributions in order to estimate the probability P_I . In [9], a heuristic technique called stratified importance sampling was used to determine the weight $w_j(\mathbf{x})$ of each sample from all the distributions. The goal of this technique is to exclude samples from a biasing distribution that appear in a region with a small number of hits. Then, the statistical result obtained from the biasing distribution with the largest number of hits in that region of the $\tau \times |\boldsymbol{\tau}_\omega|$ plane can be used to statistically resolve that region. This technique requires some *a priori* knowledge of what is the region of the parameter space that each biasing distribution resolves. The weight function for the sample \mathbf{x} is given by

$$w_j(\mathbf{x}) = \begin{cases} 1, & \mathbf{x} \in R_j, \\ 0, & \text{otherwise} \end{cases} \quad (25)$$

where R_j is the j th region in the $\tau \times |\boldsymbol{\tau}_\omega|$ plane, and all regions are nonoverlapping. In the case of [9], the regions correspond to the DGD ranges that produce the greatest number of hits for each bias α . We determined these ranges by experimentation.

A more efficient technique to combine the samples from multiple biasing distributions is the balanced heuristic method [22]. The balanced heuristic weight assigned to the sample \mathbf{x} is given by

$$w_j(\mathbf{x}) = \frac{M_j L_j^{-1}(\mathbf{x})}{\sum_{k=1}^J M_k L_k^{-1}(\mathbf{x})}. \quad (26)$$

The idea behind the balanced heuristic method is that samples are weighted according to the likelihood of each particular distribution producing samples in that region; distributions that are more likely to put samples there are weighted more heavily.

The computation of the balanced heuristic weights for any given sample requires that the likelihood ratio of all the biasing distributions be evaluated for that sample. In other words, the likelihood ratio of all the J biasing distributions have to be evaluated for the i th sample drawn from the j th distribution, even though this sample was obtained using only the biasing pdf of the j th distribution. This process is simple for the DGD bias, since one only has to evaluate the likelihood ratio of all the J distributions using the values of $\cos \theta_n$ from the i th sample that was drawn using the j th biasing pdf. That is so because the biasing direction $\hat{\mathbf{b}}$ is fixed in the direction of $\boldsymbol{\tau}_n$ in the DGD bias. However, the biasing direction $\hat{\mathbf{b}}$ varies linearly along the fiber, proportionally to the biasing parameter β , when both the

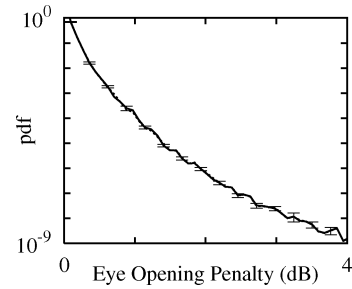


Fig. 5. The pdf of the eye opening penalty of a fiber transmission system with $\langle \tau \rangle = 25$ ps and a PMD compensator with a fixed DGD element $\tau_c = 25$ ps using Monte Carlo simulations with importance sampling biasing the DGD. We divide the domain of the penalty into 46 bins. The solid line shows the results of combining the samples using the stratified importance sampling described in (25) whose confidence interval is shown with error bars for one out of three consecutive bins. The dotted line shows the result of combining the samples with the balanced heuristic method described in (26). The solid and the dotted curves lie on top of each other.

DGD and the length of the frequency derivative of the polarization-dispersion vector are biased, as described in Section V. Consequently, the weight is determined not only by the values of $\cos \theta_n$, but also by the values of ψ_n in (24), since the likelihood ratio in each section of a given distribution depends on the angle between the vector $\mathbf{T}_n^{-1} \boldsymbol{\tau}_n$ and the biasing direction $\hat{\mathbf{b}}$ in that distribution.

In Fig. 5, we show the pdf of the eye opening penalty for a 10-Gb/s NRZ system with 30 ps of rise time, 25 ps of mean DGD, and a PMD compensator with a fixed DGD element $\tau_c = 25$ ps. The fiber transmission model, the PMD compensator and the receiver model were described in Sections II–V. The samples from the various biasing distributions are combined employing both stratified importance sampling and the balanced heuristic method in order to compare them. Their weight functions are given in (25) and (26), respectively. In both cases, we biased the DGD alone using the three biasing distributions described in Section IV. When applying stratified importance sampling, we used the following bounds: $R = [0, 2\langle \tau \rangle)$ for $\alpha = 0$, $[2\langle \tau \rangle, 3.5\langle \tau \rangle)$ for $\alpha = 0.5$, and $[3.5\langle \tau \rangle, \infty)$ for $\alpha = 1.0$. We observed a very good agreement between the two weighting methods. We verified that it is sufficient to bias only the DGD for this set of parameters by biasing both the DGD $|\boldsymbol{\tau}|$ and the length of the frequency derivative of the polarization-dispersion vector $|\boldsymbol{\tau}_\omega|$ with ten biasing distributions using the balanced heuristic method and verified that we obtained the same results. Due to its efficiency and generality, we used the balanced heuristic method in most of the statistical results shown in this work, including all the results when we biased both $|\boldsymbol{\tau}|$ and $|\boldsymbol{\tau}_\omega|$.

VII. CALCULATIONS OF OUTAGE PROBABILITY DUE TO PMD

We now use multiple importance sampling to bias both the DGD and the length of the frequency derivative of the polarization-dispersion vector to determine the eye opening penalty in both compensated and uncompensated 10 Gb/s NRZ systems with 30 ps of rise time. The fiber transmission model, the PMD compensator and the receiver model were described in Sections II–V. We use 10^4 samples in each of the 10 distributions with both the DGD and the length of the frequency derivative of

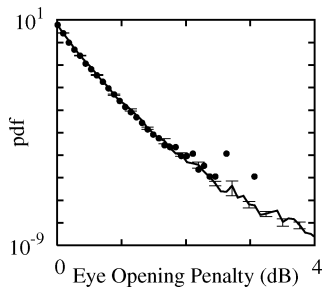


Fig. 6. Compensated pdf of the eye opening penalty due to PMD for a fiber transmission system with $\langle\tau\rangle = 30$ ps. We divide the domain of the penalty into 46 bins. The compensator is comprised of a variable DGD compensator, in which the residual DGD of the system at the center frequency of the channel is canceled after compensation. The solid line are compensated results using importance sampling with 10^5 Monte Carlo simulations whose confidence interval is shown with error bars for one out of three consecutive bins. The dots are compensated results using 1.1×10^7 standard Monte Carlo simulations.

the polarization-dispersion vector bias that we described in Section V, which we combine using the balanced heuristic method, to obtain the results that we show in this section.

In Fig. 6, we show the pdf of the eye opening penalty, and we validate our implementation of importance sampling with both the DGD and the length of the frequency derivative of the polarization-dispersion vector bias using 10^5 samples by comparison with standard Monte Carlo simulations with 1.1×10^7 samples. The fiber transmission system has 30 ps of mean DGD $\langle\tau\rangle$, and the compensator is comprised of a variable DGD element. We observed an excellent agreement between the two techniques, even though the method with importance sampling has a fraction of the computational cost of the standard Monte Carlo method. In this case, we used the residual DGD at the central frequency of the channel after compensation as the feedback parameter for the compensator, rather than the eye opening, because it does not require the use of an optimization procedure in simulations like that described in Section III. Hence, we were able to carry out the large number of standard Monte Carlo simulations that were required for this validation.

In Fig. 7, we show the outage probability as a function of the eye opening penalty margin ρ for compensated and uncompensated system with $\langle\tau\rangle = 25$ ps. The outage probability at an eye opening penalty margin ρ is the complement of the cumulative density function (cdf) of the eye opening penalty ρ , where

$$\text{cdf}(\rho) = \int_{\rho}^{\infty} p(\rho') d\rho' \quad (27)$$

and $p(\rho)$ is the corresponding pdf. Fig. 7(a) shows the results on a linear scale and Fig. 7(b) shows the same results on a logarithmic scale. When the fixed DGD element of the compensator τ_c is set to 25 ps, which is equal to the uncompensated mean DGD $\langle\tau\rangle$, we observe a larger reduction of the average penalty due to PMD than when $\tau_c = 69$ ps. However, we observe that the choice $\tau_c = 69$ ps provides a more significant reduction of the outage probability for penalties larger than 0.4 dB than does $\tau_c = 25$ ps. In Fig. 7, we show that the performance of the compensator with $\tau_c = 69$ ps is close to the performance of the compensator with a variable DGD element whose feedback parameter is also the eye opening, in agreement with [27],

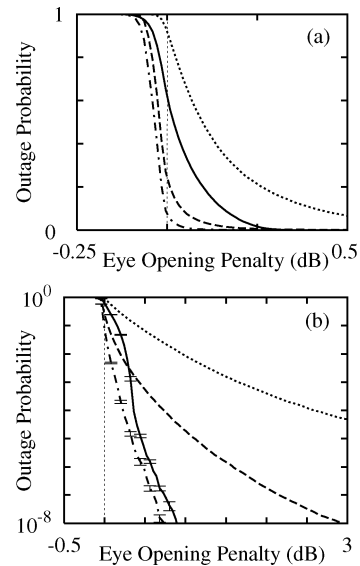


Fig. 7. Compensated and uncompensated cdfc of the eye opening penalty due to PMD for a fiber transmission system with $\langle\tau\rangle = 25$ ps. The dotted lines are uncompensated results. The dashed lines are results for a compensator with $\tau_c = \langle\tau\rangle$. The solid lines are results for $\tau_c = 2.75\langle\tau\rangle$. The dot-dashed lines are results for a compensator with a variable DGD element. The dotted vertical lines show the 0 dB eye opening penalty. The error bars show the confidence interval for the curves that have at least one bin whose relative variation exceeds 10%. For those curves, we show the error bars for one out of three consecutive bins. (a) Results plotted on a linear scale with 63 bins. (b) Results plotted on a logarithmic scale with 88 bins.

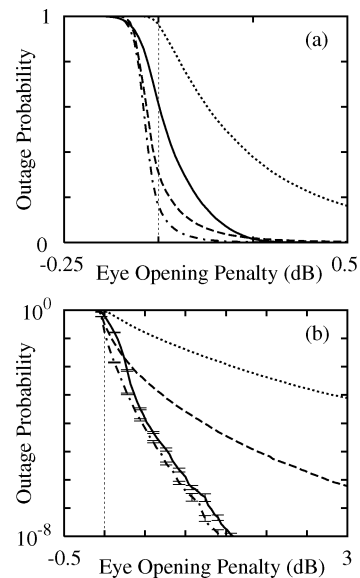


Fig. 8. Compensated and uncompensated cdfc of the eye opening penalty due to PMD for a fiber transmission system with $\langle\tau\rangle = 30$ ps. Dotted lines are uncompensated results. Dashed lines are results for a compensator with $\tau_c = \langle\tau\rangle$. Solid lines are results for $\tau_c = 2.25\langle\tau\rangle$. Dot-dashed lines are results for a compensator with a variable DGD element. The dotted vertical lines show the 0 dB eye opening penalty. The error bars show the confidence interval for the curves that have at least one bin whose relative variation exceeds 10%. For those curves, we show the error bars for one out of three consecutive bins. (a) Results plotted on a linear scale with 63 bins. (b) Results plotted on a logarithmic scale with 88 bins.

despite the difference in the complexity of these compensators. The outage probability does not equal 1 at 0 dB because there is a finite, albeit small probability that the PMD in the transmission line will interact with the DGD in the compensator to

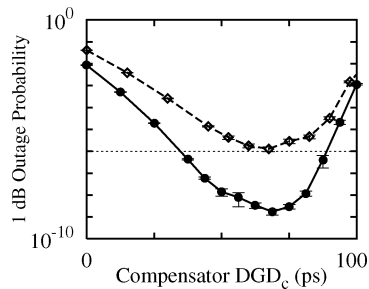


Fig. 9. The outage probability as a function of the fixed DGD element of the compensator, τ_c . The value $\tau_c = 0$ corresponds to the uncompensated case. The solid line with circles are results for $\langle \tau \rangle = 25$ ps. The dashed lines with diamonds are results for $\langle \tau \rangle = 30$ ps. The error bars show the confidence interval for the curves. The dotted line shows the 10^{-6} outage probability level.

compress the signal and reduce the penalty. In Fig. 8, we show results similar to the ones in Fig. 7, except that $\langle \tau \rangle = 30$ ps. This figure is a corrected version of [9, Fig. 2], where only the DGD bias was applied. In Figs. 7 and 8, we plot error bars for the curves that have at least one bin whose relative variation exceeds 10%.

In Fig. 9, we plot the outage probability for a 1-dB penalty as function of τ_c for fibers with $\langle \tau \rangle = 25$ ps and $\langle \tau \rangle = 30$ ps. This figure is a corrected version of [9, Fig. 3], where only the DGD bias was applied. There is an optimum value for τ_c that minimizes the outage probability for both cases. This value is about 69 ps. The reason why the outage probability rises when τ_c becomes larger than this optimum is because large values of τ_c add unacceptable penalties to fiber realizations that could be adequately compensated at lower values of τ_c . The reduction in the outage probability that the fixed DGD compensator can provide in the fiber system with $\langle \tau \rangle = 30$ ps is substantially smaller than when $\langle \tau \rangle = 25$ ps because the number of PMD realizations that the compensator cannot adequately compensate increases rapidly with the average DGD. We have also observed that it is increasingly difficult to find an optimal operating point when τ_c becomes large because the penalty depends more sensitively on the polarization controller's orientation. Thus, it is preferable to operate with the smallest possible τ_c that produces an acceptable outage probability.

VIII. CONCLUSION

This paper showed how to use multiple importance sampling, in which we bias both the DGD and the length of the frequency derivative of the polarization-dispersion vector, and it showed how to combine the samples from several biasing distributions to study PMD compensators with a single DGD element. This paper demonstrated that fixed DGD compensators can reduce the outage probability by several orders of magnitude for NRZ signals that are transmitted in optical fibers with PMD. It showed that the optimal value of the fixed DGD element of the compensator for realistic penalties of 1 dB is two to three times larger than the mean DGD of the line. The optimized fixed DGD compensator can provide a performance that is close to the one provided by a variable DGD compensator, despite the difference in the complexity of these compensators. This paper's results show that it is not sufficient to determine the impact of PMD compensators on the average penalty when

designing realistic systems because the average penalty is not directly related to the outage probability, which is the most important design parameter. It is, therefore, crucial to accurately model the tail of the probability density function of the eye opening penalty, and importance sampling is a technique that makes this study feasible.

REFERENCES

- [1] T. Takahashi, T. Imai, and M. Aiki, "Automatic compensation technique for timewise fluctuation polarization mode dispersion in in-line amplifier systems," *Electron. Lett.*, vol. 30, no. 4, pp. 348–349, 1994.
- [2] R. Noé, D. Sandel, M. Yoshida-Dierolf, S. Hinz, C. Glingener, C. Scheerer, A. Schopflin, and G. Fisher, "Polarization mode dispersion compensation at 20 Gbit/s with fiber-based distributed equalizer," *Electron. Lett.*, vol. 34, no. 25, pp. 2421–2422, 1998.
- [3] D. Mahgerefteh and C. R. Menyuk, "Effect of first-order PMD compensation on the statistics of pulse broadening in a fiber with randomly varying birefringence," *IEEE Photon. Technol. Lett.*, vol. 11, pp. 340–342, 1999.
- [4] H. Bülow, "Limitation of optical first-order PMD compensation," in *Proc. ECOC 1999*, 1999, WE1, pp. 74–76.
- [5] C. Francia, F. Bruyère, J. P. Thiéry, and D. Penninckx, "Simple dynamic polarization mode dispersion compensator," *Electron. Lett.*, vol. 35, no. 5, pp. 414–415, 1999.
- [6] H. Rosenfeldt, R. Ulrich, U. Feiste, R. Ludwig, H. G. Weber, and A. Ehrhardt, "PMD compensation in 10 Gbit/s NRZ field experiment using polarimetric error signal," *Electron. Lett.*, vol. 36, no. 5, pp. 448–450, 2000.
- [7] H. Sunnerud, C. Xie, M. Karlsson, and P. A. Andrekson, "Outage probabilities in PMD compensated transmission systems," in *Proc. ECOC 2001*, Tu.A.3.1, pp. 204–205.
- [8] A. O. Lima, I. T. Lima, Jr., T. Adali, and C. R. Menyuk, "A novel polarization diversity receiver for PMD mitigation," *IEEE Photon. Technol. Lett.*, vol. 14, pp. 465–467, 2002.
- [9] I. T. Lima, Jr., G. Biondini, B. S. Marks, W. L. Kath, and C. R. Menyuk, "Analysis of PMD compensators with fixed DGD using importance sampling," *IEEE Photon. Technol. Lett.*, vol. 14, pp. 627–629, 2002.
- [10] G. Biondini, W. L. Kath, and C. R. Menyuk, "Importance sampling for polarization-mode dispersion," *IEEE Photon. Technol. Lett.*, vol. 14, pp. 310–312, 2002.
- [11] S. L. Fogal, G. Biondini, and W. L. Kath, "Multiple importance sampling for first- and second-order polarization-mode dispersion," *IEEE Photon. Technol. Lett.*, vol. 14, pp. 1273–1275, 2002.
- [12] F. Heismann, "Automatic compensation of first-order polarization mode dispersion in a 10 Gbit/s transmission system," in *Proc. ECOC 1998*, 1998, pp. 329–330.
- [13] R. Y. Rubinstein, *Simulation and the Monte Carlo Method*. New York: Wiley, 1981.
- [14] D. Marcuse, C. R. Menyuk, and P. K. A. Wai, "Application of the Manakov-PMD equation to studies of signal propagation in optical fibers with randomly varying birefringence," *J. Lightwave Technol.*, vol. 15, pp. 1735–1746, 1997.
- [15] H. Goldstein, *Classical Mechanics*. Reading, MA: Addison-Wesley, 1980.
- [16] I. T. Lima, Jr., R. Khosravani, P. Ebrahimi, E. Ibragimov, A. E. Willner, and C. R. Menyuk, "Comparison of polarization mode dispersion emulators," *J. Lightwave Technol.*, vol. 19, pp. 1872–1881, 2001.
- [17] G. Poole and D. L. Favin, "Polarization dispersion measurements based on transmission spectra through a polarizer," *J. Lightwave Technol.*, vol. 12, pp. 917–929, 1994.
- [18] H. Bülow, "System outage probability due to first- and second-order PMD," *IEEE Photon. Technol. Lett.*, vol. 10, pp. 696–698, 1998.
- [19] P. R. Trischitta and E. L. Varma, *Jitter in Digital Transmission Systems*. Boston, MA: Artech House, 1989.
- [20] E. Polak, *Computational Methods in Optimization*. New York: Academic, 1971.
- [21] M. Karlsson, "Polarization mode dispersion-induced pulse broadening in optical fibers," *Opt. Lett.*, vol. 23, pp. 688–690, 1998.
- [22] E. Veach, "Robust Monte Carlo Methods for Light Transport Simulation," Ph.D. dissertation, Stanford University, Palo Alto, CA, 1997.
- [23] J. P. Gordon and H. Kogelnik, "PMD fundamentals: Polarization mode dispersion in optical fibers," *Proc. Nat. Acad. Sci.*, vol. 97, no. 9, pp. 4541–4550, 2000.

- [24] I. T. Lima, Jr., A. O. Lima, J. Zweck, and C. R. Menyuk, "Efficient computation of outage probabilities due to polarization effects in a WDM system using a reduced Stokes model and importance sampling," *IEEE Photon. Technol. Lett.*, vol. 15, pp. 45–47, 2003.
- [25] G. Biondini, W. L. Kath, and C. R. Menyuk, "Non-Maxwellian DGD distributions of PMD emulators," in *Proc. OFC 2001*, 2001, ThA5, pp. 1–3.
- [26] P. Ciprut, N. Gisin, R. Passy, J. P. Von der Weid, F. Prieto, and C. W. Zimmer, "Second-order polarization mode dispersion: Impact on analog and digital transmissions," *J. Lightwave Technol.*, vol. 16, pp. 757–771, 1998.
- [27] A. O. Lima, I. T. Lima, Jr., B. S. Marks, C. R. Menyuk, and W. L. Kath, "Performance analysis of single-section PMD compensators using multiple importance sampling," in *Proc. OFC 2003*, 2003, ThA3, pp. 419–421.
- [28] B. Falcidieno, I. Herman, and C. Pienovi, *Computer Graphics and Mathematics*. New York: Springer-Verlag, 1992.



Ivan T. Lima, Jr., received the B.Sc. degree in electrical engineering from the Federal University of Bahia, Bahia, Brazil, in 1995, the M.Sc. degree in electrical engineering from the State University of Campinas, Campinas, Brazil, in 1998, and the Ph.D. degree in electrical engineering in the field of photonics from the University of Maryland Baltimore County in 2003.

From 1986 to 1996, he was with Banco do Brasil (Bank of Brazil), where he served as the Information Technology Advisor of the State Superintendence of

Bahia. He was a Research Assistant in the Optical Fiber Communications Laboratory at the University of Maryland Baltimore County from 1998 to 2003. He is a tenure-track Assistant Professor in the Department of Electrical and Computer Engineering at North Dakota State University, Fargo. His research interests have been devoted to the modeling of transmission effects and receivers in optical fiber communications systems. He has authored or coauthored 18 archival journal papers, 36 conference contributions, one book chapter, and one U.S. patent.

Dr. Lima received the 2003 IEEE Lasers & Electro-Optics Society (LEOS) Graduate Student Fellowship Award, and he was a corecipient of the Venice Summer School on Polarization Mode Dispersion Award. In 2004, he was co-instructor of the Short Course SC210B: Hands-on Polarization Measurement Workshop, which was offered at the Optical Fiber Communications Conference and Exposition (OFC) 2004, Los Angeles, CA.



Aurenice O. Lima (S'00) received the B.Sc. degree in electrical engineering from the Federal University of Bahia, Bahia, Brazil, in 1996 and the M.Sc. degree in electronics and communications from the State University of Campinas, Campinas, Brazil, in 1998. She is currently pursuing the Ph.D. degree in the Department of Computer Science and Electrical Engineering at the University of Maryland Baltimore County (UMBC).

Since 2000, she has been a Research Assistant in the Optical Fiber Communications Laboratory at UMBC. Her current research interests include modeling and statistical analysis of polarization effects and signal processing for optical fiber communication systems. She has authored or coauthored nine archival journal papers and 20 conference papers.

Ms. Lima is a Student Member of the IEEE Laser and Electro-Optics Society (LEOS) and of the IEEE Women in Engineering Society. She received the Venice Summer School on Polarization Mode Dispersion Award in 2003. In 1996 and 1998, she was awarded graduate scholarships from the Brazilian Ministry of Education and from the Brazilian Ministry of Science and Technology, respectively.



Gino Biondini was born November 18, 1966 in Perugia, Italy. He received the Laurea in physics and the Doctorate in theoretical physics from the University of Perugia, Perugia, Italy, in 1991 and 1997, respectively.

He was a Postdoctoral Research Associate at the University of Colorado from 1997 to 1999, and from 1999 to 2001, he was a Research Assistant Professor at Northwestern University, Evanston, IL, where he started working on the problem of rare events in optical fiber communications. Since 2001, he has been a Zassenhaus Assistant Professor in the Mathematics Department at Ohio State University, Columbus. Starting in fall 2004, he will join the Mathematics Department of the State University of New York (SUNY), Buffalo. He is the author of more than 30 refereed publications and one patent application, and he has given numerous invited presentations at national and international conferences. His research interests include nonlinear wave equations, solitons and integrable systems, nonlinear optics and optical fiber communications, applied probability and stochastic processes, and variance reduction techniques.



Curtis R. Menyuk (F'98) was born March 26, 1954. He received the B.S. and M.S. degrees from the Massachusetts Institute of Technology (MIT), Cambridge, in 1976 and the Ph.D. from the University of California at Los Angeles (UCLA) in 1981.

He has worked as a Research Associate at the University of Maryland, College Park, and at Science Applications International Corporation, McLean, VA. In 1986, he became an Associate Professor in the Department of Electrical Engineering at the University of Maryland Baltimore County (UMBC),

and he was the Founding Member of this department. In 1993, he was promoted to Professor. He was on partial leave from UMBC from fall 1996 until fall 2002. From 1996 to 2001, he worked part-time for the Department of Defense, codirecting the optical networking program at the Department of Defense (DoD) Laboratory for Telecommunications Sciences, Adelphi, MD, from 1999 to 2001. From 2001 to 2002, he was Chief Scientist at PhotonEx Corporation. He has authored or coauthored more than 180 archival journal publications as well as numerous other publications and presentations. He has also edited two books. For the last 17 years, his primary research area has been theoretical and computational studies of fiber-optic communications. The equations and algorithms that he and his research group at UMBC have developed to model optical fiber transmission systems are used extensively in the telecommunications industry.

Dr. Menyuk is a Fellow of the Optical Society of America (OSA). He is a Member of the Society for Industrial and Applied Mathematics and the American Physical Society and a former UMBC Presidential Research Professor.



William L. Kath received the B.S. degree in mathematics from the Massachusetts Institute of Technology (MIT), Cambridge, in 1978 and the Ph.D. degree in applied mathematics from the California Institute of Technology (Caltech), Pasadena, in 1981.

He was with Caltech until 1984, where he was an NSF Postdoctoral Fellow (1981–1982), and a Von Kármán Instructor of Applied Mathematics (1982–1984). He joined the Department of Engineering Sciences and Applied Mathematics,

Northwestern University, Evanston, IL, in 1984, and from 1985 to 1990, he was an NSF Presidential Young Investigator. His research interests include optical fibers and waveguides, solitons, polarization-mode dispersion, parametric amplification, and computational neuroscience. He is the author or coauthor of more than 130 publications.

He is a Member of the Optical Society of America (OSA) and the Society for Industrial and Applied Mathematics.