Importance Sampling for Polarization-Mode Dispersion

Gino Biondini, William L. Kath, and Curtis R. Menyuk, *Fellow, IEEE*

***Abstract*—We describe the application of importance sampling to Monte-Carlo simulations of polarization-mode dispersion (PMD) in optical fibers. The method allows rare differential group delay (DGD) events to be simulated much more efficiently than with standard Monte-Carlo methods and, thus, it can be used to assess PMD-induced system outage probabilities at realistic bit-error rates. We demonstrate the technique by accurately calculating the tails of the DGD probability distribution with a relatively small number of Monte-Carlo trials.**

***Index Terms*—Optical fiber communications, polarization-mode dispersion, simulation.**

I. Introduction

POLARIZATION-MODE dispersion (PMD) has become one of the major impairments to upgrading current per channel data rates to 10 Gb/s and beyond in terrestrial wavelength-division-multiplexed (WDM) systems. A key difficulty with PMD is that it is a random phenomenon and, therefore, the penalties it produces change randomly over distance and time as the ambient temperature and other environmental parameters vary. In system design, a maximum power penalty is usually assigned to PMD, and one demands that the outage probability, (that is, the probability of the PMD-induced penalty exceeding this allowed value), is very small, typically 10^{-6} or less. Because of this stringent requirement, it has been impossible to use either Monte-Carlo simulations or laboratory experiments to determine the outage probability of a system, due to the extremely large number of system configurations that must be explored in order to obtain a reliable estimate.

In the absence of effective tools for calculating outage probabilities, system designers have resorted to stopgap techniques. An important measure of PMD is the polarization dispersion vector [1]–[3], the magnitude of which is the differential group delay (DGD). In optical fibers, the DGD is a random variable with a Maxwellian probability distribution function (pdf) [1]. The tails of the DGD distribution are particularly important, since the rare events where the DGD is significantly larger than its mean are the ones most likely to result in system outages. Since calculating probabilities in the tails of the pdf with standard Monte-Carlo techniques is not feasible, one technique used is to produce artificially large DGD values, determine the penalties at these large DGDs, and then weight the results using the Maxwellian distribution. A fundamental problem with this method, however, is that there is no direct relationship between the DGD and the power penalty. In addition, different configurations can give the same DGD but not contribute equally to the penalty, and, therefore, should be weighted differently. An alternative approach is to calculate the average DGD after PMD compensation, and then, assuming that the compensated DGD still obeys a Maxwellian distribution, to calculate the distribution of the power penalties, and thus, obtain the reduction in the outage probability. This approach is also seriously flawed, however, since the DGD distribution in compensated systems is typically far from Maxwellian.

In this letter, we use a technique called importance sampling (IS)[4]–[6] which addresses all of these difficulties and provides a tool that can be used in numerical simulations [7], and in principle in experiments, to accurately estimate PMD-induced system penalties. The technique allows low probability events to be efficiently simulated by enabling one to concentrate Monte-Carlo simulations in the most significant regions of the sample space. Here, we introduce the method and show how it can be applied to the numerical simulation of PMD-induced effects generated by a concatenation of birefringent sections. As a specific example, we calculate the DGD distribution produced by different PMD emulators. Extensions are immediately possible, such as the direct calculation of system penalties and modifications that allow the determination of rare second-order PMD events.

II. Importance Sampling for PMD

A standard technique for simulating PMD effects is the coarse-step method [7], which approximates the continuous birefringence variations present in real fibers with a concatenation of fixed length birefringent sections. Many experimental PMD generation techniques also employ a concatenation of birefringent elements, such as high-birefringence fibers [8] or birefringent waveplates [9]. These can be connected by either polarization scramblers (e.g., polarization controllers [8]) or rotatable connectors [9]. In all cases, the total polarization dispersion vector after the $(n+1)$st section can be obtained from the PMD concatenation equation [3]

$$\mathbf{\Omega}_{n+1} = \Delta\mathbf{\Omega}_{n+1} + \mathtt{R}_{n+1}\mathbf{\Omega}_n. \quad (1)$$

Manuscript received August 14, 2001; revised October 30, 2001. This work was supported in part by the Air Force Office of Scientific Research (AFOSR) under Award F49620-99-1-0016, and by the National Science Foundation (NSF) under Award DMS-9804602.

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

C. R. Menyuk is with the Department of Computer Science and Electrical Engineering, TRC 201A, University of Maryland Baltimore County, Baltimore, MD 21228 USA and is also with PhotonEx Corporation, Maynard, MA 01754 USA.

Publisher Item Identifier S 1041-1135(02)00686-9.

Here, $|\boldsymbol{\Omega}_n|$ is the polarization dispersion vector after n sections and $\Delta\boldsymbol{\Omega}_{n+1}$, which lies in the equatorial plane of the Poincaré sphere for linearly birefringent elements [10], is the individual contribution to the polarization dispersion vector of the $(n+1)$st section. For the case of birefringent waveplates, the Müller matrix $\mathtt{R}_{n+1}$ is given by the expression $\exp[\varphi_{n+1}(\hat{\boldsymbol{r}}_{n+1}\times)]$, representing a rotation through an angle φ_{n+1} about the axis $\hat{\boldsymbol{r}}_{n+1} = \Delta\boldsymbol{\Omega}_{n+1}/|\Delta\boldsymbol{\Omega}_{n+1}|$ [3]. When polarization scramblers are present, $\mathtt{R}_{n+1}$ combines the above rotation with an additional rotation that randomly and uniformly scatters the polarization state and the polarization dispersion vector on the Poincaré sphere.

Here, we apply IS to Monte-Carlo simulations of the PMD concatenation (1). The configurations that lead to large DGD values are the ones in which the individual contributions to the polarization dispersion vector from each section tend to be aligned with each other. Thus, the appropriate variables to control are the angles θ_n between $\boldsymbol{\Omega}_n$ and $\Delta\boldsymbol{\Omega}_{n+1}$. Suppose we are interested in determining the probability P that a random variable which depends upon the angles $(\theta_1, \theta_2, \ldots \theta_N) \equiv \boldsymbol{\theta}$ falls in a given range, where N is the total number of sections. Here, we will study the total DGD, $|\boldsymbol{\Omega}_N|$, but the method can be applied to any random variable, such as the amount of pulse broadening or the power penalty. The probability P can be represented as the expectation value of an indicator function $I(\boldsymbol{\theta})$, where $I = 1$ if the random variable of interest falls in the prescribed range and $I = 0$ otherwise. Using IS, we can write the Monte-Carlo estimate of the above probability as [4]–[6]

$$P = \frac{1}{M}\sum_{k=1}^{M} I(\boldsymbol{\theta}_k)L(\boldsymbol{\theta}_k) \qquad (2)$$

where M is the total number of trials and $L(\boldsymbol{\theta}) = p(\boldsymbol{\theta})/p^*(\boldsymbol{\theta})$ is the IS likelihood ratio [4]–[6]. Here $p(\boldsymbol{\theta})$ is the unbiased joint probability distribution function for the angles, while $p^*(\boldsymbol{\theta})$ is the biased distribution which is actually used to draw the samples $\boldsymbol{\theta}_k$. If $p^*(\boldsymbol{\theta}) \equiv p(\boldsymbol{\theta})$, (2) simply yields the relative number of trials falling in the range of interest. The problem with this choice is that, for low probability events, an exceedingly large number of samples is necessary in order for the desired events to occur. Using a biased probability distribution allows the desired regions of sample space to be visited more frequently. At the same time, the likelihood ratio $L(\boldsymbol{\theta})$ automatically adjusts the results so that all of the different realizations are correctly weighted, thus, contributing properly to the final probability.

When polarization scramblers are present, the length and orientation of the successive differential polarization dispersion vectors $\Delta\boldsymbol{\Omega}_{n+1}$ can be regarded as fixed, while the output of the polarization scramblers varies. Therefore, we bias the simulations by making the scramblers preferentially rotate $\boldsymbol{\Omega}_n$ toward the direction of $\Delta\boldsymbol{\Omega}_{n+1}$. More specifically, we bias the angle θ_n between the polarization dispersion vector at the output of each scrambler and the next differential polarization dispersion vector toward zero. This choice does not uniquely determine the orientation of the polarization dispersion vector at the scrambler output, because $\boldsymbol{\Omega}_n$ can still be rotated by an arbitrary amount about $\Delta\boldsymbol{\Omega}_{n+1}$ while keeping θ_n constant. We assume that this additional rotational angle is uniformly distributed.

In the unbiased case, the angles θ_n are independent random variables, with $\cos\theta_n$ uniformly distributed in [−1,1]. When applying IS, we choose $\cos\theta_n = 2x^{1/\alpha} - 1$, where x is a uniform random variable in [0,1] and $\alpha \geq 1$ is a biasing parameter. The value $\alpha = 1$ reproduces the unbiased case, while increasing values of α bias the configuration toward increasingly large values of DGD. Other choices are possible for the biased distribution of the θ_n; the effectiveness of the method is not very sensitive to the particular distribution used. The above choice yields $L(\boldsymbol{\theta}) = \prod_{n=1}^{N} p_1(\cos\theta_n)/p_\alpha(\cos\theta_n)$, where $p_\alpha(y) = (\alpha/2)[(1+y)/2]^{\alpha-1}$. We emphasize that different configurations with the same DGD can have likelihood ratios that differ by orders of magnitude, and, thus, their relative contribution to the final result can vary substantially. As a consequence, different realizations of the same DGD are expected to give very different contributions to the power penalty.

In the case of rotatable waveplates, the relative orientations between sections are the primary variables determining the total DGD. The biasing toward large DGD values is done by choosing $\Delta\boldsymbol{\Omega}_{n+1}$ to be preferentially aligned with the projection of $\boldsymbol{\Omega}_n$ onto the equatorial plane. Specifically, we choose $\Delta\boldsymbol{\Omega}_{n+1}$ so that the angle θ_n between it and the projection of $\boldsymbol{\Omega}_n$ is distributed as described previously.

When simulating rotatable waveplates, the phase retardations φ_{n+1} in $\mathtt{R}_{n+1}$ must also be specified. Recall that each beat length of a birefringent section generates a 2π retardation. In practice, sections with significant DGDs are many beat lengths long, and unless the section lengths are precise to within a small fraction of a beat length, these retardations will vary from section to section. Therefore, we choose a random set of retardation angles between 0 and 2π. Once these angles are selected, different DGDs are generated solely by rotating sections relative to one another. The results do not depend significantly upon the particular angles used, except for certain clearly pathological cases such as identical angles with φ_{n+1} equal to 0 or π.

III. Numerical Results and Comparisons

As a simple application, we have calculated the pdf of the DGD for a PMD emulator comprised of 15 birefringent sections each with 1 ps of DGD and employing either polarization scramblers or rotatable waveplates. In both cases, the biasing parameters $\alpha = 1, 2, 4, 8,$ and 12 were used, with 20 000 realizations each. As indicated in Fig. 1, each biasing parameter generates DGDs in a particular range, with larger DGDs produced by larger αs. Each value of α, therefore, allows us to sample a different portion of the DGD's pdf, and the outputs obtained from different αs combine to reconstruct the entire curve. The simulations are combined by sorting the DGD values into bins and using the likelihood ratios [12]. The final pdfs are shown in Fig. 2.

When polarization scramblers are present, the evolution of the polarization dispersion vector is equivalent to a three-dimensional (3-D) random walk, and an exact solution is available for the pdf that can be compared against the simulations [11], [13]. The numerically calculated pdf for the case of scramblers agrees extremely well with the exact solution (solid line under the data). An analytic expression for

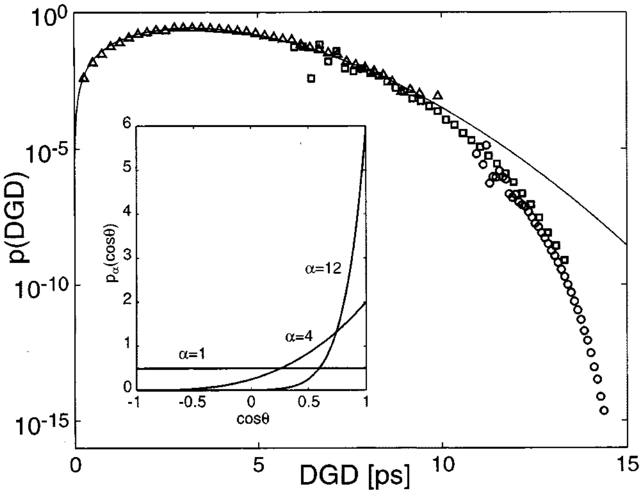

Fig. 1. Segments of the importance-sampled DGD pdf obtained using $\alpha = 1$ (triangles), $\alpha = 4$ (squares) and $\alpha = 12$ (circles) for a 15 section PMD emulator with polarization scramblers, 1 ps of DGD per section, and 20 000 realizations for each α. Solid curve: Maxwellian pdf with $\langle \text{DGD} \rangle = 3.6$ ps. Inset: Angular biasing pdfs, $p_\alpha(\cos\theta)$.

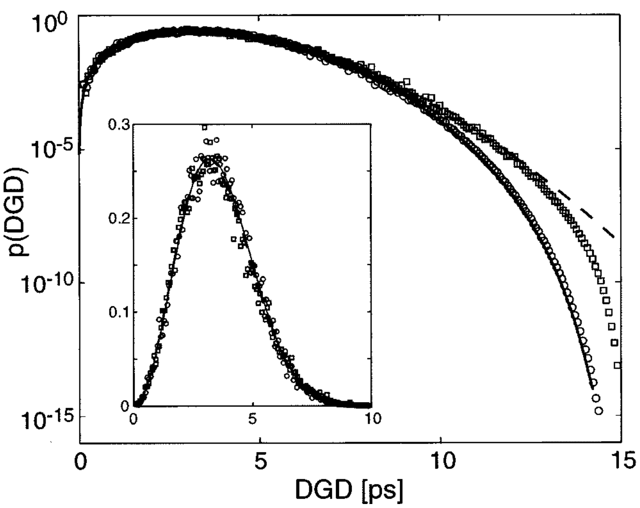

Fig. 2. Importance-sampled pdf for 15 1 ps DGD sections with polarization scramblers (circles) or birefringent waveplates (squares). Dashed curve: Maxwellian distribution with $\langle \text{DGD} \rangle = 3.6$ ps. Solid line: Exact and asymptotic solutions from [12]. Inset: Linear scale.

the pdf in the case of waveplates is not available, but IS allows it to be calculated quite easily. In both cases, the accuracy of the numerically determined pdf improves as the number of simulations (and the number of bins) is increased. It should be noted, however, that very good results are obtained even with the 100 000 Monte-Carlo simulations employed here. In particular, a good approximation is achieved even for probabilities below 10^{-12}. To obtain comparable accuracy with unbiased Monte-Carlo simulations, at least 10^{14} or 10^{15} trials would be required. Thus, for these cases IS provides a speedup of several orders of magnitude.

For moderate values of the DGD r, the pdf $p_N(r)$ is well approximated by a Maxwellian distribution $p_M(r) = (\sqrt{2}\, r^2/\sqrt{\pi}\, \sigma^3)e^{-r^2/2\sigma^2}$, where $\sigma^2 = \langle r^2 \rangle/3$ with $\langle r^2 \rangle = \sum_{n=1}^{N} a_n^2$, and where the a_n are the individual DGDs of each section. The degree of agreement, of course, improves as the number of sections N is increased [13]. The emulator using rotatable sections yields better agreement with the Maxwellian approximation than the emulator with polarization scramblers. Note, however, that a concatenation of equal length rotatable birefringent sections is known to be a poor model for real fibers due to artificial periodicities of the polarization dispersion vector's autocorrelation function in the frequency domain [8], [14]. The frequency behavior is not considered here.

IV. Conclusion

We have shown how importance sampling can be applied to Monte-Carlo simulations of PMD that use birefringent sections connected by either polarization scramblers or rotatable connectors. Importance sampling biases the Monte-Carlo simulations so that large DGD configurations occur more frequently than they would normally; as a result, the method allows rare differential group delay (DGD) events to be simulated much more efficiently than with standard methods. Importance-sampled Monte-Carlo techniques, therefore, provide a natural and effective means to assess PMD-induced impairments in optical transmission systems at realistic bit-error rates.

Acknowledgment

The authors would like to thank B. L. Nelson, D. Chowdhury, and J. Damask for several helpful discussions.

References

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

[2] N. Gisin, R. Passy, J. C. Bishoff, and B. Perny, "Experimental investigations of the statistical properties of polarization mode dispersion in single-mode fibers," *Photon. Technol. Lett.*, vol. 5, p. 819, July 1993.

[3] J. P. Gordon and H. Kogelnik, "PMD fundamentals: polarization-mode dispersion in optical fibers," *Proc. Nat. Acad. Sci.*, vol. 97, p. 4541, 2000.

[4] M. C. Jeruchim, "Techniques for estimating the bit error rate in the simulation of digital communication systems," *IEEE J. Select. Areas Commun.*, vol. 2, p. 153, Jan. 1984.

[5] J.-C. Chen, D. Lu, J. S. Sadowsky, and K. Yao, "On importance sampling in digital communications—Part I: Fundamentals," *IEEE J. Select. Areas Commun.*, vol. 11, p. 289, Apr. 1993.

[6] P. J. Smith, M. Shafi, and H. Gao, "Quick simulation: A review of importance sampling techniques in communications systems," *IEEE J. Select. Areas Commun.*, vol. 15, p. 597, May 1997.

[7] 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, Sept. 1997.

[8] R. Khosravani, Jr., I. T. Lima, P. Ebrahimi, E. Ibragimov, A. E. Willner, and C. R. Menyuk, "Time and frequency domain characteristics of polarization-mode dispersion emulators," *IEEE Photon. Technol. Lett.*, vol. 13, pp. 127–129, Feb. 2001.

[9] J. N. Damask, "A programmable polarization-mode dispersion emulator for systematic testing of 10 Gb/s PMD compensators," in *OSA Optical Fiber Communication Tech. Dig.*, 2000, vol. ThB3.

[10] A. Galtarossa, L. Palmieri, M. Schiano, and T. Tambosso, "Statistical characterization of fiber random birefringence," *Opt. Lett.*, vol. 25, p. 1322, 2000.

[11] B. D. Hughes, *Random Walks and Random Environments*. Oxford, U.K.: Clarendon Press, 1995.

[12] E. Veach, "Robust Monte Carlo methods for light transport simulation," Ph.D. dissertation, Stanford Univ., Stanford, CA, 1997.

[13] M. Karlsson, "Probability density functions of the differential group delay in optical fiber communication systems," *J. Lightwave Technol.*, vol. 19, pp. 324–331, Mar. 2001.

[14] I. T. Lima, R. Khoshravani, P. Ebrahimi, E. Ibragimov, A. E. Willner, and C. R. Menyuk, "Polarization mode dispersion emulator," in *OFC'2000 Technical Dig.* paper ThB4, pp. 31–33.