

Multiple Importance Sampling for First- and Second-Order Polarization-Mode Dispersion

Sarah L. Fogal, Gino Biondini, and William L. Kath

Abstract—A simulation method that targets all possible combinations of first- and second-order polarization-mode dispersion (PMD) is described. Use of this method in importance-sampled Monte Carlo simulations yields a more comprehensive determination of PMD-induced system penalties than first-order biasing alone and significantly speeds up the calculation of outage probabilities, particularly when PMD compensation is employed. The technique is demonstrated by using it to calculate the probability distribution function (pdf) of second-order PMD and the joint pdf of the magnitude of first- and second-order PMD.

Index Terms—Importance sampling, Monte Carlo simulation, optical fiber communications, polarization-mode dispersion.

I. INTRODUCTION

OUTAGE probabilities due to polarization-mode dispersion (PMD) are typically required to be very small. This constraint makes it extremely difficult to use either Monte Carlo simulations or laboratory experiments to determine the outage probability of a system, because of the extremely large number of configurations that must be explored in order to obtain reliable estimates. The rare events where PMD is unusually large are the ones most responsible for generating system outages. Recently, we have applied importance sampling (IS) [2] to numerical simulations of PMD [3]. The method works by biasing Monte Carlo simulations so that the configurations producing large differential group delays (DGDs) are realized much more frequently. This makes it possible to directly calculate the outage probability of systems for which first-order PMD is of primary significance [4].

A natural measure of PMD is the PMD vector τ [1], the magnitude of which is the DGD. The PMD vector is frequency dependent, of course, and therefore the DGD at any given frequency is not the sole determiner of system outages. In particular, second-order PMD, which includes both depolarization and polarization-dependent chromatic dispersion (PCD), is known to produce additional system penalties [5], [6]. Thus, when using IS to calculate outage probability of a system, it is often important to generate large values of the frequency derivative of the PMD vector, τ_ω , which quantifies the second-order PMD. Here we describe an IS technique that employs multiple biasing schemes to generate arbitrary combinations of first- and second-order PMD. As a result, the method generates much more complete PMD statistics than first-order biasing

alone. This is especially important when PMD compensation is applied, since first-order PMD is typically reduced to moderate values (or perfectly cancelled at a particular frequency). Note also that a compensator comprised of a polarization controller and a variable delay line can compensate not only first-order, but also higher order PMD [7]. In these cases, it is possible for second-order PMD to become the primary cause of system penalties [8], [9]. We demonstrate the method by numerically calculating the probability distribution function (pdf) of the magnitude of the second-order PMD vector and the joint pdf of the magnitude of the first- and second-order PMD vectors. Extensions to large PCD events are also possible (but will be present elsewhere).

II. IS FOR BOTH FIRST- AND SECOND-ORDER PMD

Various numerical and experimental PMD generation techniques use a concatenation of birefringent elements (such as high-birefringence fibers or birefringent waveplates), either connected by polarization controllers (i.e., “scramblers”) or rotatable relative to one another. The PMD vector $\tau^{(n+1)}$ and its derivative $\tau_\omega^{(n+1)}$ at a specific frequency after the $(n+1)$ st section are obtained from their respective PMD concatenation equations [1], which for linearly birefringent elements can be written as

$$\tau^{(n+1)} = R_{n+1} (\tau^{(n)} + \Delta\tau_{n+1}) \quad (1)$$

$$\tau_\omega^{(n+1)} = R_{n+1} (\tau_\omega^{(n)} + \Delta\tau_{n+1} \times \tau^{(n)}). \quad (2)$$

Here, $\Delta\tau_{n+1}$ is the differential PMD vector of the $(n+1)$ st section and R_{n+1} is the Müller matrix of the $(n+1)$ st section. For fixed length sections, the magnitude of $\Delta\tau_{n+1}$ is fixed, and only its direction varies. For linearly birefringent elements, $\Delta\tau_{n+1}$ lies on the equatorial plane of the Poincaré sphere.

When polarization controllers are present, an additional rotation matrix Q_{n+1} left multiplies $\tau^{(n)}$ and $\tau_\omega^{(n)}$ in (1), (2). It is possible to factor out Q_{n+1} from the concatenation equations; the resulting equations are formally equivalent to (1), (2) with a new rotation matrix $R'_{n+1} = R_{n+1}Q_{n+1}$, except that the new contributions $\Delta\tau'_{n+1} = Q_{n+1}^T \Delta\tau_{n+1}$ are uniformly distributed on the Poincaré sphere. This property facilitates the implementation of IS with polarization scramblers, the case we describe first. We refer to the situation where there are no polarization controllers as the case of rotatable waveplates.

As shown in [3], the appropriate variables to control when applying IS are the orientations of the individual PMD vectors of each section. IS works by biasing these $\Delta\tau_{n+1}$ toward specific directions $\mathbf{b}^{(n)}$ which maximally increase the particular quantity of interest. We characterize the vector $\mathbf{b}^{(n)}$ relative to the orthonormal frame of reference \mathcal{U} formed by the unit vectors

Manuscript received March 18, 2002; revised May 10, 2002. This work was supported by the National Science Foundation under Grants DMS-0101476 and DGE-9987577.

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

Publisher Item Identifier 10.1109/LPT.2002.801091.

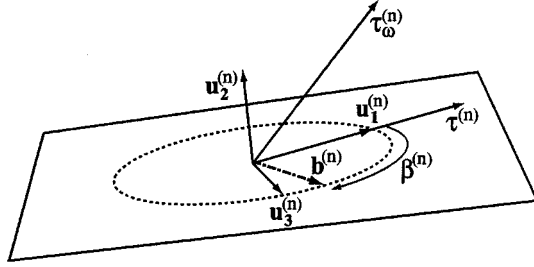


Fig. 1. The unit vectors $\mathbf{u}_1^{(n)}$, $\mathbf{u}_2^{(n)}$, $\mathbf{u}_3^{(n)}$ and the direction $\mathbf{b}^{(n)}$ used to bias the simulations.

$\{\mathbf{u}_1^{(n)}, \mathbf{u}_2^{(n)}, \mathbf{u}_3^{(n)}\}$, where $\mathbf{u}_1^{(n)}$ is parallel to $\boldsymbol{\tau}^{(n)}$, $\mathbf{u}_2^{(n)}$ is parallel to $\boldsymbol{\tau}_{\omega,\perp}^{(n)}$ (the component of $\boldsymbol{\tau}_{\omega}^{(n)}$ perpendicular to $\boldsymbol{\tau}^{(n)}$), and $\mathbf{u}_3^{(n)} = \mathbf{u}_1^{(n)} \times \mathbf{u}_2^{(n)}$, as illustrated in Fig. 1. The magnitudes of $\tau_{\omega,\parallel}$ (the component of $\boldsymbol{\tau}_{\omega}^{(n)}$ parallel to $\boldsymbol{\tau}^{(n)}$) and $\tau_{\omega,\perp}$ quantify the PCD and the depolarization, respectively.

First-order biasing [3] is achieved by choosing $\Delta\boldsymbol{\tau}_{n+1}$ to be preferentially aligned with the previous PMD vector $\boldsymbol{\tau}^{(n)}$, i.e., for $\mathbf{b}^{(n)} = \mathbf{u}_1^{(n)}$. When multiple biasing strengths are used [3], this choice of biasing generates values of $|\boldsymbol{\tau}|$ and $|\boldsymbol{\tau}_{\omega}|$ within region 1 in Fig. 2. (The actual $\Delta\boldsymbol{\tau}_{n+1}$ in each Monte Carlo trial is randomly chosen around $\mathbf{b}^{(n)}$. Increasing the biasing strength has the effect of generating samples that are more narrowly concentrated around $\mathbf{b}^{(n)}$.) First-order biasing yields the largest values of DGD. It does not, however, produce particularly large values of τ_{ω} , because when $\Delta\boldsymbol{\tau}_{n+1}$ is parallel to $\boldsymbol{\tau}_n$ the contribution to $\boldsymbol{\tau}_{\omega}^{(n+1)}$ is zero. Random variations add second-order PMD, but large values are not specifically targeted.

Over a single section, the choice that maximizes the contribution to $\boldsymbol{\tau}_{\omega}^{(n+1)}$ is to align $\Delta\boldsymbol{\tau}_{n+1}$ with $\mathbf{u}_3^{(n)}$. The rate at which $\boldsymbol{\tau}_{\omega}^{(n+1)}$ increases, however, also depends on $\boldsymbol{\tau}^{(n)}$, and thus, when optimizing over many sections, the growth of $\boldsymbol{\tau}^{(n)}$ must also be considered.

When the number of sections in the emulator is large, a continuum approximation can be used to find the deterministic configuration that generates the maximum second-order PMD. Specifically, we let $\lim_{\Delta z \rightarrow 0} \Delta\boldsymbol{\tau}_{n+1}/\Delta z = \mathbf{b}(z)$. The magnitude of $\mathbf{b}(z)$ describes the relative rate at which PMD is added by the birefringent sections. For simplicity, we will present the case $|\mathbf{b}| = b = \text{const.}$, but the calculations can be easily generalized to any prescribed function $b(z)$ (which corresponds to the case of nonequal length sections). In this continuum limit, for polarization scramblers one obtains

$$d\boldsymbol{\tau}/dz = b_1 \quad (3)$$

$$d\tau_{\omega,\parallel}/dz = b_2 \tau_{\omega,\perp}/\tau \quad (4)$$

$$d\tau_{\omega,\perp}/dz = b_3 \tau - b_2 \tau_{\omega,\parallel}/\tau \quad (5)$$

where (b_1, b_2, b_3) are the components of \mathbf{b} in the coordinate frame \mathcal{U} . This system of equations can be solved exactly for any $\mathbf{b}(z)$. The maximum second-order PMD is generated when $b_2 = 0$, i.e., when \mathbf{b} lies in the \mathbf{u}_1 - \mathbf{u}_3 plane. Calculus of variations can then be used to find the optimal choice of b_1 and b_3 . The maximum growth of second-order PMD is found to occur when $\mathbf{b} = b(\mathbf{u}_1 \cos \beta + \mathbf{u}_3 \sin \beta)$, where $\beta(z) = \beta_{\max} z/z_{\max}$, with $\beta_{\max} = \pi/2$. With multiple biasing strengths, this choice produces region 3 in Fig. 2. We refer to this choice, which yields the largest values of $|\boldsymbol{\tau}_{\omega}|$, as optimal second-order biasing.

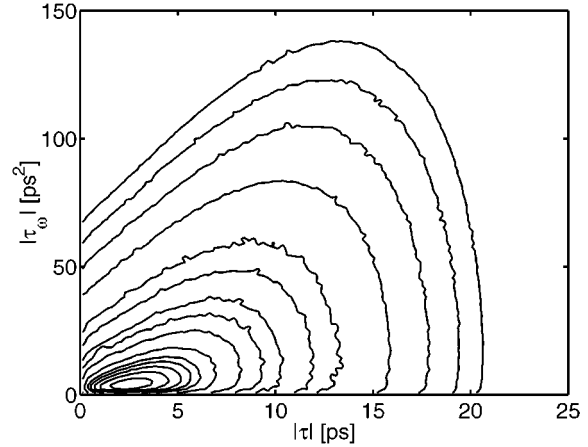


Fig. 2. The regions of the $|\boldsymbol{\tau}|$ - $|\boldsymbol{\tau}_{\omega}|$ plane targeted by the various biasing methods. Region 1 corresponds to first-order biasing ($\beta = 0$), region 3 to pure second-order biasing ($\beta_{\max} = \pi/2$), and regions 2, 4, and 5 to $\beta_{\max} = \pi/4, 3\pi/4$ and π , respectively. The dashed line shows the much smaller region obtained with unbiased samples.

Of course, a more complete coverage of the $|\boldsymbol{\tau}|$ - $|\boldsymbol{\tau}_{\omega}|$ plane is often needed. In this case, intermediate biasing choices must also be used. These intermediate choices can be obtained by using calculus of variations to maximize a linear combination of $|\boldsymbol{\tau}|$ and $|\boldsymbol{\tau}_{\omega}|$. The resulting form of $\mathbf{b}(z)$ is the same as above, except that the value of the final angle β_{\max} ranges between 0 and π , depending upon the specific combination of first- and second-order PMD being maximized. A selection of angles, together with the resulting regions in the $|\boldsymbol{\tau}|$ - $|\boldsymbol{\tau}_{\omega}|$ plane, is illustrated in Fig. 2. The advantage of using multiple biasing—as opposed to just pure first- or second-order biasing or no biasing at all—is evident. Each value of β_{\max} generates samples lying in a region that emanates in a roughly radial fashion from the location where the joint pdf is maximum as the biasing strength is increased. Together, a set of angles β_{\max} covers the entire $|\boldsymbol{\tau}|$ - $|\boldsymbol{\tau}_{\omega}|$ plane.

Once the deterministic directions $\mathbf{b}^{(n)}$ have been selected, the contributions $\Delta\boldsymbol{\tau}_{n+1}$ are randomly chosen to be preferentially aligned with $\mathbf{b}^{(n)}$. This is done by using a biased distribution for the polar angle between $\Delta\boldsymbol{\tau}_{n+1}$ and $\mathbf{b}^{(n)}$ [3], while the remaining angle (describing the azimuthal rotation of $\Delta\boldsymbol{\tau}_{n+1}$ about $\mathbf{b}^{(n)}$) is kept uniform in $[0, 2\pi]$. When birefringent waveplates are used, $\Delta\boldsymbol{\tau}_{n+1}$ must lie the equatorial plane of the Poincaré sphere. In this case, the biasing direction is obtained by projecting the vector $\mathbf{b}^{(n)}$ determined above onto the equatorial plane, and $\Delta\boldsymbol{\tau}_{n+1}$ is biased toward this direction.

The simultaneous use of different biasing methods is called *multiple importance sampling* [10], and several ways exist to combine results coming from the various biasings. Here, we used the *balance heuristic* [11], which is asymptotically close to optimal when the number of realizations is large. Of course, the output of each individual Monte Carlo realization is always adjusted for the biasing using the IS likelihood ratio [2], [3].

III. NUMERICAL RESULTS

As a specific example, we consider a concatenation of 50 sections with 0.5 ps of DGD per section. Fig. 3 shows, for scramblers (squares) and waveplates (circles), the pdfs of first- and second-order PMD, obtained with optimal first- and second-

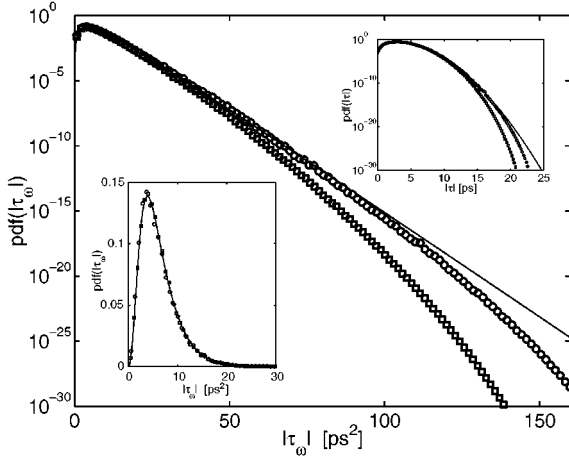


Fig. 3. The pdf of second-order PMD for a concatenation of 50 sections with 0.5-ps DGD each, using scramblers (squares) or waveplates (circles). Bottom-left inset: The pdf on a linear scale. Top-right inset: The pdf of the DGD. The solid lines show the pdfs for real fiber. A total of 2×10^6 Monte Carlo realizations were used.

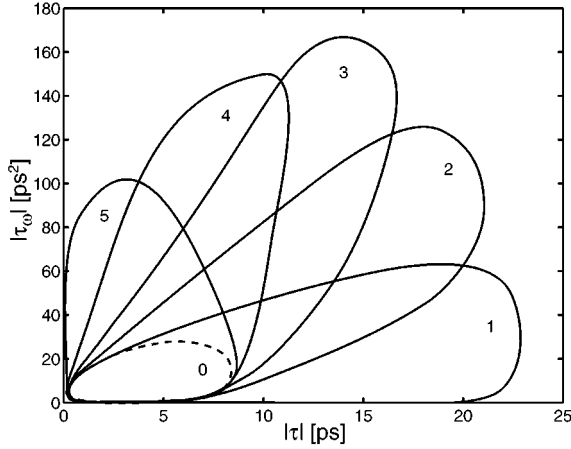


Fig. 4. Contour plots of the joint pdf for a concatenation of 50 sections with 0.5-ps DGD each and polarization scramblers. The contours are at 10^{-n} with $n = 30, 25, 20, 15, 10, 8, 6, 5, 4, 3, 2, 2.5, 2, 1.75$, and 1.5 . A total of 10^6 Monte Carlo realizations were used.

order biasing, respectively. The solid lines show the analytical pdfs for real fiber [12]. As in [3], a relatively small number of realizations allows us to calculate extremely low probabilities.

For both scramblers and waveplates, the pdfs deviate from the real fiber distributions in the tails, since here the PMD is generated by an emulator with a finite number of birefringent sections, which by necessity has finite maximums for both τ and τ_ω . (Similarly, such a device has inaccessible regions in the $|\tau|$ - $|\tau_\omega|$ plane. The emulator should be chosen so that these regions are uninfluential in determining the outage probability of the systems to be tested.) For a device with equal length birefringent sections, an approximate value for the maximum second-order PMD is obtained from the exact solution of the continuum model, evaluated for optimal second-order biasing: $|\tau_\omega|_{\max} = N^2(\Delta\tau)^2/\pi$, where N is the number of sections and $\Delta\tau$ is the DGD of each section. Of course, $|\tau|_{\max} = N\Delta\tau$.

Fig. 4 shows the joint pdf of the magnitude of first- and second-order PMD (a two-dimensional reduction of the full three-dimensional joint pdf of first- and second-order PMD

[13]) for an emulator with polarization scramblers, as calculated with the multiple biasing technique. The characteristic function for PMD in an optical fiber (the limit of an infinite number of infinitesimal birefringent sections) was found in [13], but, to our knowledge, no exact analytical expression exists for the joint pdf. (An approximate pdf was given in [13] for the case where second-order PMD is a small perturbation compared to first-order.) Similarly, the joint pdf for PMD emulators with a finite number of sections is not known analytically.

IV. CONCLUSION

We have presented a simulation method that efficiently generates arbitrary combinations of first- and second-order PMD. Used in conjunction with importance sampling, this technique offers a much more complete coverage of rare PMD events than previously possible, thus enabling a much more accurate determination of system outage probabilities. We have demonstrated the method by calculating the pdf of second-order PMD as well as the joint pdf of the magnitude of first- and second-order PMD. Nontrivial combinations of first- and second-order PMD are expected to be potentially significant in realistic situations.

ACKNOWLEDGMENT

The authors would like to thank J. N. Damask, C. R. Menyuk, and I. T. Lima for many valuable discussions.

REFERENCES

- [1] J. P. Gordon and H. Kogelnik, "PMD fundamentals: Polarization-mode dispersion in optical fibers," in *Proc. Nat. Acad. Sci.*, vol. 97, 2000, p. 4541.
- [2] M. C. Jeruchim, "Techniques for estimating the bit error rate in the simulation of digital communication systems," *IEEE J. Select. Areas Commun.*, vol. SAC-2, pp. 153–170, Jan. 1984.
- [3] G. Biondini, W. L. Kath, and C. R. Menyuk, "Importance sampling for polarization-mode dispersion," *IEEE Photon. Technol. Lett.*, vol. 14, pp. 310–312, Mar. 2002.
- [4] I. T. Lima, G. Biondini, B. S. Marks, W. L. Kath, and C. R. Menyuk, "Analysis of polarization-mode dispersion compensators using importance sampling," *IEEE Photon. Technol. Lett.*, vol. 14, pp. 627–629, May 2002.
- [5] P. Ciprut, B. Gisin, 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, May 1998.
- [6] C. Francia, F. Bruyère, D. Penninckx, and M. Chbat, "PMD second-order effects on pulse propagation in single-mode optical fibers," *IEEE Photon. Technol. Lett.*, vol. 10, pp. 1739–1741, Dec. 1998.
- [7] H. Sunnerud, C. Xie, M. Karlsson, R. Samuelsson, and P. A. Andrekson, "A comparison between different PMD compensation techniques," *J. Lightwave Technol.*, vol. 20, pp. 368–378, Mar. 2002.
- [8] W. Shieh, "On the second-order approximation of PMD," *IEEE Photon. Technol. Lett.*, vol. 12, pp. 290–292, Mar. 2000.
- [9] S. L. Fogal, G. Biondini, and W. L. Kath, "Importance-sampled pulse broadening before and after PMD compensation," presented at the *Opt. Fiber Commun. Conf.*, 2002, paper ThA2.
- [10] A. Owen and Y. Zhou, "Safe and effective importance sampling," *J. Amer. Statist. Assoc.*, vol. 95, pp. 135–143, Mar. 2000.
- [11] E. Veach, "Robust Monte-Carlo methods for light transport simulation," Ph.D. dissertation, Stanford Univ., 1997.
- [12] G. J. Foschini, L. E. Nelson, R. M. Jopson, and H. Kogelnik, "Probability densities of second-order polarization-mode dispersion including polarization-dependent chromatic dispersion," *IEEE Photon. Technol. Lett.*, vol. 12, pp. 293–295, Mar. 2000.
- [13] 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.