

# Evaluation of Polarization Mode Dispersion Statistics using the Multicanonical Monte Carlo Method

Nikolaos Mantzoukis

Department of Electrical and Computer Engineering  
Optical Communications Group  
University of Patras, Greece  
[mantzoukis@ece.upatras.gr](mailto:mantzoukis@ece.upatras.gr)

Ioannis Roudas

Department of Electrical and Computer Engineering  
Optical Communications Group  
University of Patras, Greece  
[roudas@ece.upatras.gr](mailto:roudas@ece.upatras.gr)

Thomas Kamalakis

Department of Informatics and Telematics of the Harokopion  
University of Athens, Greece  
[thkam@hua.gr](mailto:thkam@hua.gr)

Lampros Raptis

Attica Telecommunications,  
Athens, Greece  
[lraptis@hol.net](mailto:lraptis@hol.net)

**Abstract**— We describe, in detail, the principles of the multicanonical Monte Carlo method and its advantages for the evaluation of the probability distribution function (pdf) of a random variable, compared to the conventional Monte Carlo method. We apply the multicanonical Monte Carlo in the estimation of the tails of the pdf of the first- and second-order polarization mode dispersion in fiber-optic communications systems. It is shown that an improvement of several orders of magnitude in the accuracy of the pdf estimation is achieved with the multicanonical Monte Carlo method, compared to its conventional counterpart.

**Index Terms**— Monte Carlo methods, optical fiber theory, optical fiber polarization, polarization-mode dispersion (PMD).

## I. INTRODUCTION

Polarization-mode dispersion (PMD) [1], [2] is one of the most fundamental limitations for the deployment of ultra-high data rate fiber-optic communications systems and networks. PMD stems from the presence of birefringence along optical fibers caused by imperfections in manufacturing, cabling, and fiber deployment. Due to birefringence, different spectral components of the transmitted signal propagate in optical fibers with different phase and group velocities depending on their state of polarization. Consequently, at the receiver, the signal spectral components recombine out of phase and their superposition results in pulses exhibiting time spreading and distortion, which, in turn, can lead to severe power outages.

Fiber birefringence varies as a function of ambient temperature and environmental vibrations and, therefore, PMD-induced outages occur randomly over time. In a well-

designed optical communications system, the probability of occurrence of PMD-induced outages is smaller than  $10^{-5}$ . Therefore, the theoretical evaluation of the performance of fiber-optic communications systems, in the presence of PMD, requires the use of computationally efficient Monte Carlo techniques for fast PMD simulation (e.g., importance sampling [3], multicanonical Monte Carlo method [4], [5]).

The guiding principle behind these techniques is the following: PMD outages, which are of main interest in the simulation, occur for certain rare combinations of the magnitude of local birefringence and orientation of principal axes along the optical fiber. If these rare combinations can be artificially made to occur more frequently, compared to the conventional Monte Carlo method, then the execution time required for the evaluation of the performance of fiber-optic communications systems can be reduced. Hence, the basic methodology is to choose "biased" distributions for the magnitude of local birefringence and orientation of principal axes along the optical fiber, in order to artificially generate more PMD outages. The simulation results are appropriately weighted to correct for the use of the biased input distributions and to ensure that the estimation of the outage probability is unbiased.

This paper reviews the principles of the multicanonical Monte Carlo method and its application in the numerical evaluation of PMD statistics in optical fibers. A salient feature of the multicanonical Monte Carlo method is that the optimal biased distributions of the magnitude of local birefringence and orientation of principal axes along the optical fiber are determined iteratively and need not be calculated a priori. Within each single iteration, the method employs the Metropolis algorithm [6] to produce new

instances of the magnitude of local birefringence and orientation of principal axes along the optical fiber.

The rest of the paper is organized as follows: Section 2 describes the computer model used for PMD simulation, which is based on the representation of the fiber as a concatenation of birefringent waveplates. This section also summarizes key analytical relationships from the previous literature regarding the probability distributions of different PMD orders [7]. In Section 3, the principles of the multicanonical Monte Carlo method are described in detail and they are applied in the evaluation of PMD statistics. Finally, Section 4 presents sample results for the numerical evaluation of the pdf of the first- and second-order polarization mode dispersion. It is shown that the numerical results are in excellent agreement with the respective analytical probability distributions over at least fifteen orders of magnitude.

## II. PMD EMULATION

### A. The PMD emulator model

A widely used computer model of long single mode fibers with PMD [1], is shown in Fig. 1. It is composed of  $NS$  short birefringent fiber segments with constant fast and slow principal axes (called *principal states of polarization* (PSPs)), denoted in Stokes space by  $\hat{p}_{m-}$  and  $\hat{p}_m$ ,  $m=1, \dots, NS$ , respectively. The difference in propagation delay between the principal axes is called *differential group delay* (DGD). Each individual fiber segment is characterized by its PMD vector, defined as  $\Delta\vec{\tau}_m = \Delta\tau_m \cdot \hat{p}_m$ ,  $m=1, \dots, NS$  [2].

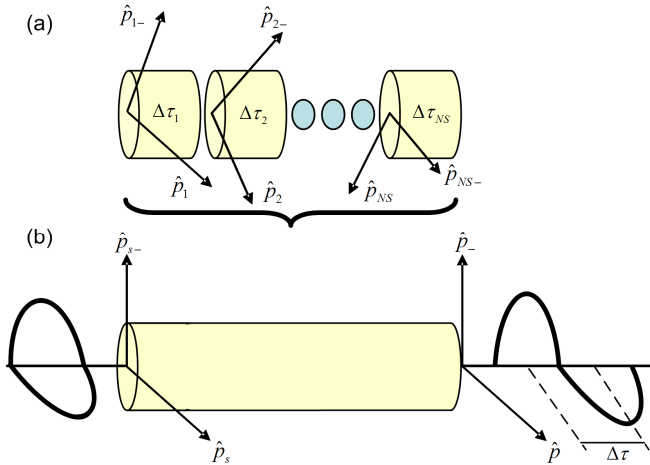


Fig. 1. (a) A long single mode optical fiber can be modeled as a concatenation of  $m=1, \dots, NS$  short birefringent fiber segments, with randomly aligned fast  $\hat{p}_{m-}$  and slow  $\hat{p}_m$  principal axes, respectively, and DGD  $\Delta\tau_m$  following a Gaussian distribution. (b) Propagation of a pulse through such a series of birefringent waveplates yields, at the output of the optical fiber, a multitude of pulses which combine coherently. To a first-order approximation, the fiber can be approximated by a single birefringent waveplate with input PSPs  $\hat{p}_{s-}$ ,  $\hat{p}_s$ , output PSPs  $\hat{p}_-$  and  $\hat{p}$ , and total DGD  $\Delta\tau$ .

We assume that the principal axes of adjacent fiber segments are randomly misaligned. More specifically, we assume that the slow axes are uniformly distributed on the Poincaré sphere, as shown in Fig. 2. The DGDs of individual segments are independent, identically distributed random variables, following a Gaussian distribution with mean value  $\mu$  and standard deviation  $\sigma$ , to be specified below.

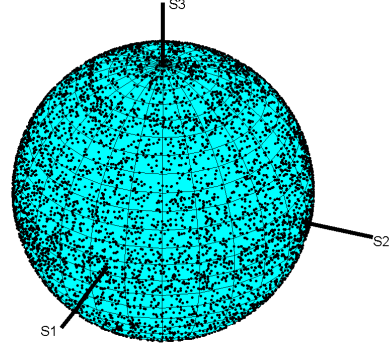


Fig. 2. Uniform distribution of the slow principal axes of  $10^4$  individual fiber segments on the Poincaré sphere. The points at the intersection of the Poincaré sphere with the Stokes axes S1, S2, S3 correspond to h horizontal, linear  $+45^\circ$  and right-circular polarization, respectively.

At the input of each fiber segment, the optical signal is decomposed into two components, traveling at different group velocities, along the fast and slow principal axes of the segment. In a first-order approximation, the optical signal at the fiber output can be considered as the weighted sum of two replicas of the input signal traveling along two orthogonal states of polarization, i.e., called principal states of polarization (PSPs). Mathematically, the PSPs are defined as two eigenmodes whose state of polarization does not change during propagation through the optical fiber to the first order with frequency [2]. The group delay difference between the PSPs is called total DGD and is denoted by  $\Delta\tau$ . We define the total PMD vector  $\vec{\tau}$ , in the right-circular Stokes space, as [2]:

$$\vec{\tau} = \Delta\tau \hat{p}, \quad (1)$$

In (1), both  $\Delta\tau$  and  $\hat{p}$  are functions of the angular frequency deviation  $\omega$  from the signal carrier frequency. Different PMD orders are defined as successive derivatives of the PMD vector with respect to angular frequency deviation  $\omega$ , evaluated at the signal's carrier frequency ( $\omega = 0$ ) [2]:

$$\vec{\tau} = \sum_{n=0}^{\infty} \frac{\vec{\tau}^{(n)}(0)}{n!} \omega^n, \quad (2)$$

$$\text{where } \vec{\tau}^{(n)}(0) = \left. \frac{d^n \vec{\tau}}{d\omega^n} \right|_{\omega=0}.$$

For instance, the first-order PMD vector is defined as

$$\vec{\tau}^{(0)}(0) = \vec{\tau}(0) = \Delta\tau(0) \hat{p}(0). \quad (3)$$

Similarly, the second-order PMD vector is the derivative of the first-order PMD, with respect to angular frequency deviation  $\omega$  and is given by

$$\bar{\tau}^{(1)}(0) = \bar{\tau}_{\omega}^{(0)}(0) = \Delta\tau_{\omega}(0)\hat{p}(0) + \Delta\tau(0)\hat{p}_{\omega}(0), \quad (4)$$

where the subscript  $\omega$  indicates the first derivative with respect to  $\omega$ . The first term, in (4),  $\Delta\tau_{\omega} \cdot \hat{p}$ , is parallel to  $\bar{\tau}$  and causes polarization-dependent chromatic dispersion, while the second term,  $\Delta\tau \cdot \hat{p}_{\omega}$ , is orthogonal to  $\bar{\tau}$  and causes signal depolarization [7].

The total PMD vector  $\bar{\tau}$  of the concatenation of a series of birefringent fiber segments can be computed by adding the PMD vectors  $\Delta\bar{\tau}_m$ ,  $m=1, \dots, NS$  of individual segments using a concatenation rule [2]. The latter is a recursive relation between the output PMD vectors of the  $m$ -th and  $(m+1)$ -th fiber section:

$$\begin{aligned} \bar{\tau}_{m+1} &= \Delta\bar{\tau}_{m+1} + \mathbf{M}_{m+1}\bar{\tau}_m \\ \bar{\tau}_{\omega, m+1} &= \Delta\bar{\tau}_{\omega, m+1} + \mathbf{M}_{m+1}\bar{\tau}_{\omega, m}, \end{aligned} \quad (5)$$

where  $\bar{\tau}_{m+1}$ ,  $\bar{\tau}_{\omega, m+1}$  are the total first- and second-order PMD vector after the  $(m+1)$ -th section,  $\Delta\bar{\tau}_{m+1}$  is the contribution of the  $(m+1)$ -th section, and  $\mathbf{M}_{m+1}$  is the Müller matrix associated with the  $(m+1)$ -th section.

The matrix  $\mathbf{M}_m$  is the product of two terms: (i) a  $3 \times 3$  orthogonal matrix  $\mathbf{R}_m$ , representing a rotation of the state of polarization (SOP) by a polarization controller, and (ii) a diagonal  $3 \times 3$  matrix  $\mathbf{D}_m$ , representing the birefringence of each fiber section. Each rotation matrix  $\mathbf{M}_m$  is written as

$$\mathbf{M}_m = \mathbf{D}_m(\omega, \Delta\tau_m) \mathbf{R}_m(\alpha_m, \varepsilon_m), \quad m=1, \dots, NS \quad (6)$$

where  $\alpha_m$  is the azimuth and  $\varepsilon_m$  is the ellipticity of each fiber segment. In (6),  $\mathbf{D}_m$ ,  $\mathbf{R}_m$  are expressed as [2]

$$\mathbf{D}(\omega, \Delta\tau) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\omega\Delta\tau) & -\sin(\omega\Delta\tau) \\ 0 & \sin(\omega\Delta\tau) & \cos(\omega\Delta\tau) \end{bmatrix} \quad (7)$$

and

$$\mathbf{R}(\alpha, \varepsilon) = \begin{bmatrix} \cos 2\alpha \cos 2\varepsilon & -\sin 2\alpha & -\cos 2\alpha \sin 2\varepsilon \\ \cos 2\varepsilon \sin 2\alpha & \cos 2\alpha & -\sin 2\alpha \sin 2\varepsilon \\ \sin 2\varepsilon & 0 & \cos 2\varepsilon \end{bmatrix}. \quad (8)$$

In (6), each birefringent section's DGD, denoted by  $\Delta\tau_m$ ,  $m=1, \dots, NS$ , follows a Gaussian distribution with mean value  $\mu = \sqrt{3\pi/(8NS)} \langle \Delta\tau \rangle$  and standard deviation  $\sigma = 0.2\mu$ ,

where  $\langle \Delta\tau \rangle$  is the mean total DGD of the PMDE and angled brackets denote ensemble average [3]. The choice of Gaussian DGDs for the individual fiber segments offers significant computational advantages in comparison with fixed-DGD fiber sections, since it leads to a Maxwellian total DGD distribution, for even a small number of fiber segments, in agreement with experimental measurements. For uniformly distributed slow PSPs on the Poincaré sphere, the azimuth  $\alpha$  and the ellipticity  $\varepsilon$  in (6) follow the distributions below [8]

$$p_{\alpha}(\alpha) = \frac{1}{\pi}, \quad |\alpha| \leq \frac{\pi}{2} \quad (9)$$

$$p_{\varepsilon}(\varepsilon) = \cos(2\varepsilon), \quad |\varepsilon| \leq \frac{\pi}{4} \quad (10)$$

In the following, the first- and second-order PMD, at a given frequency, are estimated using the multicanonical Monte Carlo method. In order to validate the model, we compare the simulated and the analytical probability density functions (pdfs) for the first- and second-order PMD vectors' magnitude [7]

$$p_{|\bar{\tau}|}(|\bar{\tau}|) = \frac{32|\bar{\tau}|^2}{\pi^2 \langle \Delta\tau \rangle^3} \exp\left(-\frac{4|\bar{\tau}|^2}{\pi \langle \Delta\tau \rangle^2}\right) \quad (11)$$

$$p_{|\bar{\tau}_{\omega}|}(|\bar{\tau}_{\omega}|) = \frac{32|\bar{\tau}_{\omega}|}{\pi \langle \Delta\tau \rangle^4} \operatorname{sech}\left(\frac{4|\bar{\tau}_{\omega}|}{\langle \Delta\tau \rangle^2}\right) \tanh\left(\frac{4|\bar{\tau}_{\omega}|}{\langle \Delta\tau \rangle^2}\right), \quad (12)$$

and angled brackets denote ensemble average.

### III. THE MULTICANONICAL MONTE CARLO METHOD

#### A. General principles

Let  $f_X(\vec{x})$  be the joint pdf of the input random variables  $\vec{x} = \{x_1, \dots, x_m\}$ , which are defined on a space  $X$ . In our case,  $X$  is the set of all possible configurations of the birefringent sections of the PMD emulator (PMDE). For each configuration, we wish to estimate the total instantaneous DGD of the PMDE (referred to, in the following, as the control parameter). The total instantaneous DGD is a complicated function of the input parameters, denoted by  $DGD = Y(\vec{x})$ . If we define the event  $R_k$  to contain all realizations of  $X$ , for which  $\{y_k < Y < y_k + \Delta y\}$ , then the probability of the event  $R_k$  is simply

$$p_Y(y_k)\Delta y = \int_{R_k} f_X(\vec{x}) d\vec{x}, \quad (13)$$

where  $p_Y(y)$  is the pdf of  $Y$ ,  $k$  is the index of the current bin of the state space of  $Y$ ,  $\Delta y$  is the width of each bin, and  $r_{R_k}(\vec{x})$  is an indicator function of the event  $R_k$ , defined as

$$r_{R_k}(\bar{x}) = \begin{cases} 1, & \text{if } \bar{x} \in R_k \\ 0, & \text{otherwise} \end{cases} \quad (14)$$

The pdf  $p_Y(y_k)$  can be estimated by simulation, using the conventional Monte Carlo (MC) method

$$\hat{p}_Y^{MC}(y_k) = \frac{1}{\Delta y N} \sum_{i=1}^N r_{R_k}(\bar{x}_i), \quad (15)$$

where  $\{\bar{x}_i\}_{i=1}^N$  are  $N$  samples, randomly drawn from space  $X$ , according to the joint pdf  $f_X(\bar{x})$ . Equation (15) indicates that the MC estimator simply evaluates the occurrence frequency of random samples corresponding to the event  $R_k$ .

A reduction of the variance of  $\hat{p}_Y^{MC}(y_k)$  can be achieved by using importance sampling [9]. First, equation (13) is rewritten as

$$p_Y(y_k) \Delta y = \int_X r_{R_k}(\bar{x}) w(\bar{x}) f_X^*(\bar{x}) d\bar{x}, \quad (16)$$

where  $w(\bar{x}) = f_X(\bar{x})/f_X^*(\bar{x})$  is a weight function and  $f_X^*(\bar{x})$  is a biased distribution. The corresponding importance sampling (IS) estimator of the pdf  $p_Y(y_k)$  is given by

$$\hat{p}_Y^{IS}(y_k) = \frac{1}{\Delta y N} \sum_{i=1}^N r_{R_k}(\bar{x}_i) w(\bar{x}_i). \quad (17)$$

In (17), the input samples  $\bar{x}_i$  are drawn from an arbitrary biased distribution  $f_X^*(\bar{x})$ . In order to minimize the variance of the IS estimator, the optimal biased distribution is [11]

$$f_{opt}^*(\bar{x}) = \frac{r_{R_k}(\bar{x}) f_X(\bar{x})}{\int_X r_{R_k}(\bar{x}) f_X(\bar{x}) d\bar{x}} = \frac{r_{R_k}(\bar{x}) f_X(\bar{x})}{p_Y(y_k) \Delta y} \quad (18)$$

Since equation (18) depends on the unknown pdf  $p_Y(y_k)$ , it is impossible to know, a priori, the optimal distribution  $f_{opt}^*(\bar{x})$ . Hence, the major drawback of importance sampling techniques is that they require an approximation of  $f_{opt}^*(\bar{x})$  as a starting point.

The main idea behind the MMC algorithm is to start from a suboptimal initial distribution and, through successive iterations, to converge towards the optimal biased distribution  $f_{opt}^*(\bar{x})$ . During the  $j$ -th iteration, the estimate of the optimal biasing distribution  $f_{opt}^{*(j)}(\bar{x})$  is given by

$$f_{opt}^{*(j)}(\bar{x}) \propto \frac{r_{R_k}(\bar{x}) f_X(\bar{x})}{p_Y^{(j-1)}(y_k) \Delta y}, \quad (19)$$

where  $p_Y^{(j-1)}(y_k)$  is the estimate of the pdf of  $Y$ , obtained in the  $(j-1)$ -th iteration. The suboptimal distribution  $f_{opt}^{*(j)}(\bar{x})$  tends to converge at the optimal biasing distribution  $f_{opt}^*(\bar{x})$ , as the iteration number  $j$  increases. As an initial guess for  $p_Y^{(0)}(y_k)$ , one can use a histogram of the output random variable  $Y$  calculated by the conventional Monte Carlo method, based on the equation (15).

### B. Random Sample Generation

Within the  $j$ -th MMC iteration, the Metropolis algorithm [6] is used to generate samples  $\bar{x}_i$  from  $f_{opt}^{*(j)}(\bar{x})$ . In each Metropolis step, a candidate sample  $\bar{x}_i^*$  is created from the previous accepted sample  $\bar{x}_i$ , by adding a small random perturbation  $\delta \bar{x}$

$$\bar{x}_i^* = \bar{x}_i + \delta \bar{x}, \quad (20)$$

where  $\delta \bar{x}$  is drawn from an arbitrary symmetric pdf. It is worth noting that the appropriate choice of this symmetric distribution is important for fast convergence towards  $f_{opt}^*(\bar{x})$ . Here, we assume that the elements of  $\delta \bar{x}$  follow a uniform distribution. The range of  $\delta \bar{x}$  is defined by perturbation coefficients  $\epsilon$ . These coefficients are set by trial-and-error. A useful criterion for the selection of the perturbation coefficients is the acceptance ratio, defined as the ratio of the accepted Metropolis steps to the total number of Metropolis steps. As a rule of thumb,  $\epsilon$  should be selected in such a way that the acceptance ratio is about 30~50% [10]. The perturbation coefficients play a significant role, both for the number of samples required for each iteration and for the rate of convergence of the estimated pdf. An inappropriate choice of these coefficients can lead to fluctuations of the estimated pdf  $p_Y^{(j)}(y_k)$  or to slow convergence.

The candidate sample  $\bar{x}_i^*$  is accepted with probability

$$\min \left\{ 1, \frac{f_{opt}^{*(j)}(\bar{x}_i^*)}{f_{opt}^{*(j)}(\bar{x}_i)} \right\}. \quad (21)$$

Otherwise, the procedure is carried out again until a sample is accepted.

### C. Evaluation of the output pdf

The final step is the evaluation of the pdf  $p_Y^{(j)}(y_k)$  of the control parameter, at the  $j$ -th iteration. We represent the current value of  $y_k$  as an initial state A, and the newly estimated state,  $y_k^*$ , as a candidate state B. The new variable  $y_k^*$  causes a transition from the old state A to the candidate state B. This transition is accepted with a probability given by

$$\min \left\{ 1, \frac{P_{stateA}^{(j)}}{P_{stateB}^{(j)}} \right\} \quad (22)$$

where  $P_{stateA}^{(j)}$ ,  $P_{stateB}^{(j)}$  are the pdfs evaluated at states A and B, respectively at the  $j$ -th iteration [11], [12].

If the candidate transition is rejected, the old state is employed, instead, in the subsequent step. After a specified number of generated states at the  $j$ -th iteration, a histogram,  $H_k^{(j)}$ , of the control parameter, is constructed. This histogram is used to bias the subsequent iteration [12].

The pdf  $p_Y^{(j)}(y_k) = p_k^{(j)}$  is updated at the end of each iteration, through a recursive relation, so that the number of samples in each bin of the control parameter's histogram becomes approximately equal, as the iteration number increases. The recursion algorithm is given by [12]:

$$p_{k+1}^{(j+1)} = \frac{p_k^{(j+1)} p_{k+1}^{(j)} \left( \frac{H_{k+1}^{(j)}}{H_k^{(j)}} \right)^{\hat{g}_k^{(j)}}}{p_k^{(j)} \left( \frac{H_{k+1}^{(j)}}{H_k^{(j)}} \right)^{\hat{g}_k^{(j)}}}, \quad (23)$$

where  $\hat{g}_k^{(j)}$  is the relative statistical significance of the  $k$ -th bin in the  $j$ -th iteration and is defined as [12]

$$\hat{g}_k^{(j)} = \frac{g_k^{(j)}}{\sum_{h=1}^j g_k^{(h)}}, \text{ with } g_k^{(j)} = \frac{H_{k+1}^{(j)} H_k^{(j)}}{H_{k+1}^{(j)} + H_k^{(j)}}. \quad (24)$$

The relative statistical significance  $\hat{g}_k^{(j)}$  coefficient (or confidence interval) indicates how likely an interval (bin) is to contain the candidate value of the control parameter. If  $H_{k+1}^{(j)} + H_k^{(j)} = 0$  in (23), at a given iteration, then the  $k$ -th bin has no statistical significance in this iteration and we set  $g_k^{(j)} = 0$ . This results in  $\hat{g}_k^{(j)} = 0$ . The exponent  $0 \leq \hat{g}_k^{(j)} \leq 1$  in (23) depends on the previous bins and iterations, introducing a significant coupling between the estimated pdf values  $p_k^{(j)}$  at each MMC iteration. Finally, the  $p_k^{(j)}$  is normalized so that  $\sum_{k=1}^N p_k^{(j)} = 1$ . As the number of iterations increases, the number of hits in each bin will asymptotically be the same ( $H_{k+1}^{(j)} / H_k^{(j)} \rightarrow 1$ ), and the relative statistical significance will asymptotically converge to zero ( $\hat{g}_k^{(j)} \rightarrow 1$ ) [13]. As a consequence, during the last MMC iteration, the estimated pdf will asymptotically converge to the true pdf of the control parameter. The efficiency of the MMC algorithm is based on the appropriate choice of the following parameters:

1. The pdf of the perturbation  $\delta \vec{x}$
2. The perturbation coefficients  $\epsilon$
3. The number of Metropolis steps
4. The number of MMC iterations

5. The number of bins that the sample space of  $Y$  is partitioned.

The selection of these parameters is performed empirically, by trial-and-error. If the parameters are chosen correctly, then the convergence of the estimated pdf to the desired probability distribution is efficiently achieved.

#### D. Application to PMD

In this subsection, we present an implementation of the MMC method for the study of PMD in fiber-optic communication systems using the PMDE model of section IIA. The possible configurations of the PMDE's fiber segments are determined by the azimuth  $\{a_m\}$ , the ellipticity  $\{\epsilon_m\}$ , and the value of each section's DGD  $\{\Delta\tau_m\}$ ,  $m=1, \dots, NS$ . Each Metropolis step within the  $j$ -th MMC iteration is implemented as follows: Assume that the  $l$ -th value of each random parameter is  $\{a_{m,l}\}$ ,  $\{\epsilon_{m,l}\}$ , and  $\{\Delta\tau_{m,l}\}$ ,  $m=1, \dots, NS$  and  $l=1, \dots, N_{samples}$ , where  $N_{samples}$  is the number of samples for each parameter generated in each iteration. Next, we add a small perturbation to these values  $d\alpha_{m,l} \sim \epsilon_\alpha U(-\pi/2, \pi/2)$ ,  $d\epsilon_{m,l} \sim \epsilon_\epsilon U(-\pi/4, \pi/4)$ , and  $d\Delta\tau_{m,l} \sim \epsilon_{\Delta\tau} N(\mu, \sigma)$ , respectively. In the above,  $U(-\lambda, \lambda)$  denotes the uniform distribution in the interval  $[-\lambda, \lambda]$  and  $N(\mu, \sigma)$  denotes the Gaussian distribution with mean  $\mu$  and standard deviation  $\sigma$ . The perturbation coefficients  $\epsilon_\alpha$ ,  $\epsilon_\epsilon$ , and  $\epsilon_{\Delta\tau}$  are selected according to the acceptance ratio rule of thumb, explained in the previous subsection.

Next, we independently accept or reject each one of tentative values  $a_{m,l}^*$ ,  $\epsilon_{m,l}^*$ , and  $\Delta\tau_{m,l}^*$  with probability

$$\min \left\{ 1, \frac{p_\alpha(a_{m,l}^*)}{p_\alpha(a_{m,l})} \right\}, \min \left\{ 1, \frac{p_\epsilon(\epsilon_{m,l}^*)}{p_\epsilon(\epsilon_{m,l})} \right\}, \min \left\{ 1, \frac{p_{\Delta\tau}(\Delta\tau_{m,l}^*)}{p_{\Delta\tau}(\Delta\tau_{m,l})} \right\}, \quad (25)$$

respectively. If each one of these equations (25) is true, then the new value of the random parameters is accepted independently. Then, we calculate the first- and second-order total PMD vectors of the PMDE based on the concatenation rule (5), and we accept the new instantaneous value  $DGD_l^*$  with probability

$$\min \left\{ 1, \frac{p_{DGD}^{(j)}(DGD_l^*)}{p_{DGD}^{(j)}(DGD_l)} \right\}, \quad (26)$$

where  $p_{DGD}^{(j)}(DGD_l^*)$  is the pdf estimated from the histogram of the control parameter for the  $l$ -th sample computed at the  $j$ -th iteration,  $DGD_l$  is the old value and  $DGD_l^*$  is the new proposed value. In our simulation process,

the PMDE model consists of 100 birefringent fiber sections with mean total DGD 10 ps. We have initialized all histogram elements  $H_k^{(0)}$  to unity.

#### IV. RESULTS

Fig. 3 shows the pdfs of the first- and second-order PMD obtained after 8 iterations of the MMC simulation method. The circles correspond to the first iteration, the dotted line corresponds to the 4<sup>th</sup> iteration, and the dashed-dotted line corresponds to the 8<sup>th</sup> iteration. The results of the conventional MC method for  $10^5$  samples (dashed line) and the corresponding theoretical pdfs (black thin solid lines), given by (11) and (12), respectively, are also shown for comparison.

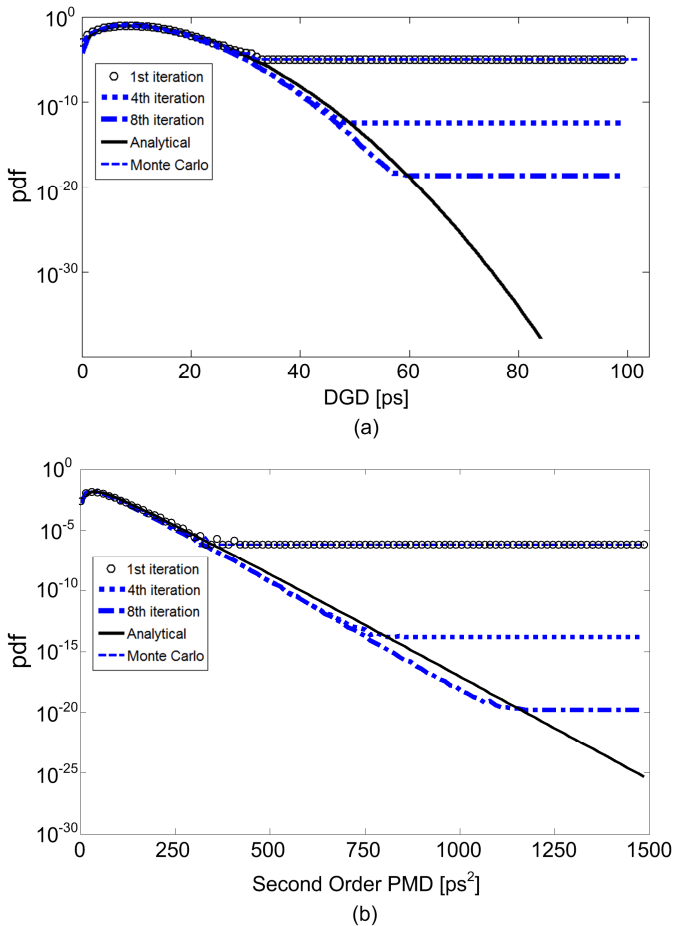


Fig. 3. Estimated pdfs of (a) the first- and (b) the second-order PMD, for a 100-sections PMD emulator, after 8 iterations of MMC method (Symbols: circles: first iteration, dotted line: 4<sup>th</sup> iteration, dashed-dotted line: 8<sup>th</sup> iteration, dashed line: a conventional MC method, and black thin solid lines: analytical relations (11), (12). Conditions:  $NS = 100$ ,  $10^5$  realizations in each iteration).

For each MMC iteration,  $10^5$  realizations are employed and the acceptance ratio is  $\sim 40\%$ . The diagrams indicate that the multicanonical Monte Carlo method greatly increases the accuracy in the tail region estimated pdf, compared the conventional MC technique. The accuracy of

the estimation increases approximately by one order of magnitude for each MMC iteration. An improvement of several orders of magnitude in accuracy is eventually achieved using eight iterations of the MMC method compared to the conventional MC method. These results are in agreement with the previous literature [4], [5].

In Fig. 4, we show contour plots of the joint pdf of the first- and the second-order PMD. The variance of the estimates of the multicanonical Monte Carlo method is almost the same over a wide region of the plot parameters after the 8<sup>th</sup> iteration.

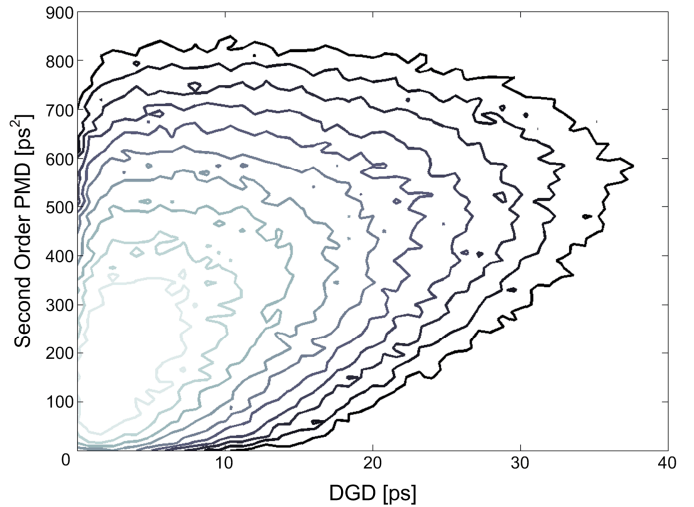


Fig. 4. Probability contours of the joint pdf of the first- and the second-order PMD after the 8<sup>th</sup> iteration of MMC method. (Conditions:  $NS = 100$ ,  $10^5$  realizations in each iteration. The magnitudes of the contours are, from left to right,  $10^{-2d}$  where  $d = 0-9$ ).

#### V. SUMMARY

We reviewed the mathematical principles of the multicanonical Monte Carlo method and applied it in the evaluation of first- and second-order PMD statistics in fiber optic communication systems. We showed that the multicanonical Monte Carlo method can estimate very rare values of the aforementioned effects with a relatively small number of simulation runs, compared to the conventional Monte Carlo method.

#### REFERENCES

- [1] C. D. Poole and J. Nagel, "Polarization effects in lightwave systems," *Optical Fiber Telecommunications*, I. P. Kaminow and T. L. Koch, Eds. San Diego, CA: Academic, 1997, vol. III A.
- [2] J. P. Gordon and H. Kogelnik, "PMD fundamentals: Polarization mode dispersion in optical fibers," *Proc. Nat. Acad. Sci.*, vol. 97, pp. 4541–4550, Apr. 2000.
- [3] G. Biondini, W. L. Kath, and C. Menyuk, "Importance sampling for polarization-mode dispersion: techniques and applications," *J. Lightw. Technol.*, vol. 22, no. 4, pp. 1201–1215, Apr. 2004.
- [4] D. Yevick, "Multicanonical communication system modeling - application to PMD Statistics," *IEEE Photon. Technol. Lett.*, vol. 14, no. 11, pp. 1512–1514, Nov. 2002.

- [5] ———, “The accuracy of multicanonical system models,” *IEEE Photon. Technol. Lett.*, vol. 15, no.2, pp. 224–226, Feb. 2003.
- [6] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, “Equation of state calculations by fast computing machines,” *J. Chem. Phys.*, vol. 21, pp. 1087–1092, 1953.
- [7] G. J. Foschini, L. E. Nelson, R. M. Jopson, and H. Kogelnik, “Probability densities of second-order polarization mode dispersion including polarization dependent chromatic fiber dispersion,” *IEEE Photon. Technol. Lett.*, vol. 12, no. 3, pp. 293–295, Mar. 2000.
- [8] A. Vannucci and A. Bononi, “Statistical characterization of the Jones matrix of long fibers affected by polarization mode dispersion (PMD),” *J. Lightw. Technol.*, vol. 20, no. 5, pp. 811–821, May 2002.
- [9] M. C. Jeruchim, P. Balaban, and K. S. Shanmugan, *Simulation of communications systems: modeling, methodology, and techniques*, 2nd ed. New York: Kluwer Academic/Plenum, 2000.
- [10] M. H. Kalos and P. A. Whitlock, *Monte Carlo methods*. New York, NY: John Wiley and Sons, 1986.
- [11] A. Bilenca, A. Desjardins, B. E. Bouma, and G. J. Tearney, “Multicanonical Monte Carlo simulations of light propagation in biological media,” *Opt. Expr.*, vol. 13, no. 24, pp. 9822–9833, Nov. 2005.
- [12] B. A. Berg, “Algorithmic aspects of multicanonical simulations,” *Nucl. Phys. B (Proc. Suppl.)*, vol. 63A-C, no. 1–3, pp. 982–984, Apr. 1998.
- [13] A. O. Lima, I. T. Lima Jr., and C. R. Menyuk, “Error estimation in multicanonical Monte Carlo simulations with applications to polarization mode dispersion emulators,” *J. Lightw. Technol.*, vol. 23, no. 11, pp. 3781–3789, Nov. 2005.