Statistics of Group Delays in Multimode Fiber With Strong Mode Coupling

Keang-Po Ho, *Senior Member, IEEE*, and Joseph M. Kahn, *Fellow, IEEE*

***Abstract*—The modal group delays (GDs) are a key property governing the dispersion of signals propagating in a multimode fiber (MMF). An MMF is in the strong-coupling regime when the total length of the MMF is much greater than the correlation length over which local principal modes can be considered constant. In this regime, the GDs can be described as the eigenvalues of zero-trace Gaussian unitary ensemble, and the probability density function (pdf) of the GDs is the eigenvalue distribution of the ensemble. For fibers with two to seven modes, the marginal pdf of the GDs is derived analytically. For fibers with a large number of modes, this pdf is shown to approach a semicircle distribution. In the strong-coupling regime, the delay spread is proportional to the square root of the number of independent sections, or the square root of the overall fiber length.**

***Index Terms*—Modal dispersion, mode-division multiplexing, multimode fiber, semicircle distribution.**

I. Introduction

MULTIMODE fiber (MMF) is widely used in short-range optical links [1]–[3], where it is often favored over single-mode fiber (SMF) because of relaxed connector alignment tolerances and reduced transceiver component costs. MMF supports propagation of multiple spatial modes having different group velocities, and thus different group delays (GDs), an effect called modal dispersion [4], [5]. Even if a signal is launched into one spatial mode, bends, index imperfections, and other perturbations cause the signal to couple into multiple modes [4], [6]–[8], making the signal subject to modal dispersion. Modal dispersion limits current commercial MMF links to 10 Gb/s per fiber up to about 300 m long [1], [9], and next-generation 100 Gb/s Ethernet MMF systems use ten fibers per link [10]. Techniques to increase the bit rate per fiber are desired.

SMF, which is free from modal dispersion, is the dominant medium for longer transmission distances. Emerging long-haul systems use dual-polarization quaternary phase-shift keying and coherent detection to achieve a spectral efficiency of 2 bits/s/Hz [11], [12]. Usage of higher order modulation formats [13] can at least double the spectral efficiency, but further increases are expected to become increasingly difficult [14], because of limits posed by optical amplifier noise and fiber nonlinearity [15], [16]. Techniques to further increase spectral efficiency are desired.

Mode-division multiplexing (MDM) in an MMF [17], [18], a form of multi-input, multi-output transmission, is a potential means to increase transmission capacity in both short- and long-distance optical networks. Like multipath propagation in wireless systems, the plurality of modes in an MMF was long viewed as a strictly negative, bandwidth-limiting effect requiring mitigation, but is now seen as creating additional degrees of freedom in which to transmit information [19]–[23]. Modal dispersion in an MMF typically leads to a larger GD spread than that caused by chromatic dispersion. This GD spread determines the required cyclic prefix length in MDM systems using orthogonal frequency-division multiplexing [21] or the required number of equalizer taps in MDM systems using single-carrier modulation [23]. In other words, receiver complexity increases in proportion to the GD spread caused by modal dispersion.

Effective mitigation of modal dispersion or optimal use of MDM requires a detailed understanding of modal dispersion, especially the effect of mode coupling on the modal GDs. Models for mode coupling were developed more than 30 years ago [4], [5], [7], when MMF links used spatially and temporally incoherent light-emitting diodes. Virtually all the models ignore phase effects, and consider only power coupling between modes. Power coupling models are able to qualitatively explain some observations, such as the scaling of delay spread with fiber length. Delay spread scales linearly with fiber length in the weak-coupling regime (e.g., short glass MMF), and with the square root of fiber length in the strong-coupling regime (e.g., plastic MMF) [24]. However, most modern MMF systems use spatially and temporally coherent laser sources, and power coupling models are not able to explain certain observations, such as a sensitivity of the impulse response to launched polarization [25].

SMF supports propagation in two polarizations, and polarization-mode dispersion (PMD) has long been modeled using electric field coupling models [26], [27]. Field coupling models have been used to demonstrate the existence of principal states of polarization (PSPs), which have well-defined GDs to first order in frequency [27]–[29]. In long SMFs, polarization modes are strongly coupled. In this regime, the differential GD (DGD) between the PSPs scales with the square root of fiber length and follows a Maxwellian distribution [28]–[31]. PSPs form the basis of techniques for optical PMD compensation in direct detection systems.

Manuscript received April 22, 2011; revised July 01, 2011; accepted August 08, 2011. Date of publication August 18, 2011; date of current version October 19, 2011. The work of J. M. Kahn was supported in part by the National Science Foundation under Grant ECCS-1101905 and in part by Corning, Inc., Corning, NY.

K.-P. Ho is with Silicon Image, Sunnyvale, CA 94085 USA (e-mail: kpho@ieee.org).

J. M. Kahn is with the Department of Electrical Engineering, Edward L. Ginzton Laboratory, Stanford University, Stanford, CA 94305 USA (e-mail: jmk@ee.stanford.edu).

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

Digital Object Identifier 10.1109/JLT.2011.2165316

Recently, field coupling models have been extended to MMF [32], [33]. These models explain the polarization sensitivity of mode coupling and demonstrate the existence of principal modes (PMs). The PMs have well-defined GDs to first order in frequency, and form the basis for optical techniques to compensate modal dispersion [34]. The GD differences between PMs scale linearly with fiber length in the weak-coupling regime, and with the square root of fiber length in the strong-coupling regime [33]. To date, however, the statistical properties of the GDs, which are of particular interest in the strong-coupling regime, have not been studied.

Here, the statistics of the GDs are derived analytically for MMF in the strong-coupling regime, considering a number of modes ranging from two to infinity[1]. In this regime, regardless of the number of modes or the GDs in the absence of coupling, the GDs scale with the square root of fiber length or the square root of the number of independent fiber sections, similar to PMD in SMF in the strong-coupling regime [28]. End-to-end modal dispersion effects are described, at each single frequency, by a random complex Gaussian Hermitian matrix or Gaussian unitary ensemble [35]. From such a model, the joint probability density function (pdf) of GDs can be derived analytically. Here, closed-form expressions for the GD distributions are derived for small number of modes. For a large number of modes, the GD distribution asymptotically approaches a semicircle distribution with a radius or upper limit equal to twice its standard deviation.

The remaining parts of this paper are organized as follows. Section II describes the random matrix model for MMF propagation. Section III provides closed-form analyses of the GD distribution in fibers with two to seven modes. Section IV presents asymptotic expressions for the GD distribution in the limit of a large number of modes. Sections V and VI presents discussion and conclusions, respectively.

II. Random Matrix Model for MMFs

The propagation characteristics of an MMF, in particular, the local PMs and their GDs, can be considered invariant over a certain correlation length. Because bends, mechanical stresses, and manufacturing tolerances induce mode coupling, the local PMs in sections separated by distance longer than the correlation length can be considered independent of each other. Throughout this paper, we consider the regime of strong mode coupling, where the total length of the MMF far exceeds the correlation length. The theory presented here is valid regardless of the actual correlation length or the modal GD profile within the correlation length.

An MMF may be divided into K sections, with propagation in each section modeled as a random matrix. The length of each section should be at least slightly longer than the correlation length, so that the local PMs in the different sections can be considered independent. Although the approach used here is applicable even if each section has different properties, for convenience, we assume that all sections are statistically equivalent.

In an MMF with D modes, modal propagation in the kth section can be modeled as a $D \times D$ matrix $\mathcal{M}^{(k)}(\omega)$ as a function of frequency ω. Here, we are only interested in the statistical properties of the modal GD, so for simplicity, we ignore any mode-dependent gain or loss. With strong mode coupling, the kth section may be represented by the product of three $D \times D$ matrices

$$\mathcal{M}^{(k)}(\omega) = \mathcal{V}^{(k)}\Lambda^{(k)}(\omega)\mathcal{U}^{(k)\dagger}, \quad k = 1, \ldots, K \tag{1}$$

where † denotes Hermitian transpose, $\mathcal{U}^{(k)}$ and $\mathcal{V}^{(k)}$ are random unitary matrices representing the mode coupling at the input and output, respectively, and $\Lambda^{(k)}(\omega)$ is the diagonal matrix describing the uncoupled modal GDs, i.e.,

$$\Lambda^{(k)}(\omega) = \mathrm{diag}\left[e^{-j\omega\tau_1^{(k)}}, e^{-j\omega\tau_2^{(k)}}, \ldots, e^{-j\omega\tau_D^{(k)}}\right] \tag{2}$$

where $\tau_i^{(k)}$, $k = 1, \ldots, K$, $i = 1, \ldots, D$, are the uncoupled GDs in the sections.

In the absence of mode-dependent gain or loss $\mathcal{M}^{(k)}$, $\mathcal{U}^{(k)}$, and $\mathcal{V}^{(k)}$ are all unitary matrices, such that $\mathcal{M}^{(k)}(\omega)\mathcal{M}^{(k)\dagger}(\omega) = I$, where I is the identity matrix. With strong random coupling, both $\mathcal{U}^{(k)}$ and $\mathcal{V}^{(k)}$ can be assumed to be independent random unitary matrices, such that both input and output are randomly oriented. The model (1) is similar to the matrix model of PMD described in [31] and [36].

The model here is valid regardless of whether or not the vectors $\vec{\tau}^{(k)} = (\tau_1^{(k)}, \tau_2^{(k)}, \ldots, \tau_D^{(k)})$, $k = 1, \ldots, K$, have the same statistical properties. The vector $\vec{\tau}^{(k)}$ may even be a deterministic vector, identical for each section. For convenience and without loss of generality, we assume that $\sum_i \tau_i^{(k)} = 0$, i.e., we ignore the mode-averaged delay of each section, as it does not lead to modal dispersion.

Using $|s_k\rangle$ and $|t_k\rangle$ to denote the input and output modes, respectively, we have

$$|t_k\rangle = \mathcal{M}^{(k)}(\omega)|s_k\rangle \tag{3}$$

and $|s_k\rangle = \mathcal{M}^{(k)\dagger}(\omega)|t_k\rangle$. Similar to the analysis of PMD [27], [29], [36], the GDs correspond to the eigenvalues of $j\mathcal{M}_\omega^{(k)}\mathcal{M}^{(k)\dagger}$ where $\mathcal{M}_\omega^{(k)} = \partial\mathcal{M}^{(k)}(\omega)/\partial\omega$. With only a single section, we may verify that

$$j\mathcal{M}_\omega^{(k)}\mathcal{M}^{(k)\dagger} = \mathcal{V}^{(k)}\mathcal{T}^{(k)}\mathcal{V}^{(k)\dagger} \tag{4}$$

with $\mathcal{V}^{(k)}$ as the local PMs in the kth section, and where

$$\mathcal{T}^{(k)} = \mathrm{diag}\left[\tau_1^{(k)}, \tau_2^{(k)}, \ldots, \tau_D^{(k)}\right] \tag{5}$$

is a diagonal matrix of their GDs in the kth section. With $\sum_i \tau_i^{(k)} = 0$, we have $\mathrm{tr}(\mathcal{T}^{(k)}) = 0$ and

$$\mathrm{tr}\left(j\mathcal{M}_\omega^{(k)}\mathcal{M}^{(k)\dagger}\right) = 0. \tag{6}$$

Physically, the ith local PM experiences an uncoupled GD $\tau_i^{(k)}$ without mixing with other modes. Because the diagonal matrices $\mathcal{T}^{(k)}$ are real matrices, all matrices $j\mathcal{M}_\omega^{(k)}\mathcal{M}^{(k)\dagger}$, $k = 1, \ldots, K$, are Hermitian.

When K sections of MMF are cascaded together, the overall propagation matrix becomes

$$\mathcal{M}^{(t)} = \mathcal{M}^{(K)}\mathcal{M}^{(K-1)}\cdots\mathcal{M}^{(2)}\mathcal{M}^{(1)}. \tag{7}$$

[1]Throughout this paper, "modes" include both polarization and spatial degrees of freedom. For example, the two-mode case can describe the two polarization modes in the SMF.

The overall PMs and their GDs correspond to the eigenvectors and eigenvalues of [32], [33]

$$\mathcal{G} = j\mathcal{M}_{\omega}^{(t)}\mathcal{M}^{(t)\dagger}. \tag{8}$$

Because

$$\begin{aligned}\mathcal{M}_{\omega}^{(t)} &= \mathcal{M}_{\omega}^{(K)}\mathcal{M}^{(K-1)}\cdots\mathcal{M}^{(2)}\mathcal{M}^{(1)} \\ &\quad + \mathcal{M}^{(K)}\mathcal{M}_{\omega}^{(K-1)}\cdots\mathcal{M}^{(2)}\mathcal{M}^{(1)} \\ &\quad + \cdots + \mathcal{M}^{(K)}\mathcal{M}^{(K-1)}\cdots\mathcal{M}^{(2)}\mathcal{M}_{\omega}^{(1)}\end{aligned} \tag{9}$$

we obtain

$$\begin{aligned}\mathcal{G} &= j\mathcal{M}_{\omega}^{(K)}\mathcal{M}^{(K)\dagger} + j\mathcal{M}^{(K)}\mathcal{M}_{\omega}^{(K-1)}\mathcal{M}^{(K-1)\dagger}\mathcal{M}^{(K)\dagger} \\ &\quad + \cdots + j\mathcal{M}^{(K)}\mathcal{M}^{(K-1)}\cdots\mathcal{M}^{(2)}\mathcal{M}_{\omega}^{(1)}\mathcal{M}^{(t)\dagger}.\end{aligned} \tag{10}$$

From (10), the overall matrix $j\mathcal{M}_{\omega}^{(t)}\mathcal{M}^{(t)\dagger}$ is the summation of K random matrices. All those K matrices have eigenvalues statistically identical to those of (4) and (5). However, their eigenvectors are independent of each other. The first matrix $j\mathcal{M}_{\omega}^{(K)}\mathcal{M}^{(K)\dagger}$ has eigenvectors derived from $\mathcal{V}^{(K)}$. The second matrix $j\mathcal{M}^{(K)}\mathcal{M}_{\omega}^{(K-1)}\mathcal{M}^{(K-1)\dagger}\mathcal{M}^{(K)\dagger}$ has eigenvectors derived from $\mathcal{M}^{(K)}\mathcal{V}^{(K-1)}$. Matrices $\mathcal{V}^{(K)}$ and $\mathcal{M}^{(K)}\mathcal{V}^{(K-1)}$ are both unitary matrices and are obviously independent of each other. All the K matrices summed to form $\mathcal{G}$ are independent of each other with eigenvalues given by the vectors $\vec{\tau}^{(k)}$, $k = 1, \ldots, K$. Even for the case that all vectors $\vec{\tau}^{(k)}$ are deterministic and identical, all the K matrices summed to form $\mathcal{G}$ are independent, owing to the different directions of their independent eigenvectors.

The matrix elements of $\mathcal{G}$ should be identically distributed Gaussian random variables from the central limit theorem (CLT). The matrix elements of $\mathcal{G}$, $g_{i,j}$, $i, j = 1, \ldots, D$, are the summation of K identically distributed random variables, as seen from (10). If K is very large, $g_{i,j}$ are Gaussian random variables from the CLT. Because all K component matrices in (10) are Hermitian, $\mathcal{G}$ is a Hermitian matrix. The diagonal elements $g_{i,i}$, $i = 1, \ldots, D$, are all real Gaussian random variables with variance equal to σ_g^2. All nondiagonal elements $g_{i,j}$, $i \neq j$, are complex Gaussian random variables with independent real and imaginary parts, which have variance equal to $\sigma_g^2/2$. Thus, the elements $g_{i,j}$, $i \neq j$ has variance σ_g^2. The value σ_g^2 depends on the number of modes D, the number of sections K, and the variances of the uncoupled GDs described by $\vec{\tau}^{(k)}$. If the D vectors in $\mathcal{V}^{(k)}$ given by (4) are assumed to be independent of each other, it can be shown that

$$\sigma_g^2 = \frac{1}{D^2}\sum_{k=1}^{K}\left\|\vec{\tau}^{(k)}\right\|^2 = \frac{1}{D}\sum_{k=1}^{K}\sigma_{\tau^{(k)}}^2 \tag{11}$$

where $\sigma_{\tau^{(k)}}^2$, $k = 1, \ldots, K$, are the variances of the GDs in the sections. If all K sections have the same modal GD profiles, we have

$$\sigma_g^2 = \frac{K}{D}\sigma_{\tau}^2 \tag{12}$$

where σ_{τ}^2 are the GD variances in all sections.

In random matrix theory, the matrix $\mathcal{G}$ is described as a Gaussian unitary ensemble ([35], Sec. 2.5). Typically, a Gaussian unitary ensemble does not have any constraint aside from the variance of its Gaussian elements. However, in (10), the matrix components have zero trace so that

$$\mathrm{tr}(\mathcal{G}) = 0. \tag{13}$$

In other words, $\mathcal{G}$ is a zero-trace Gaussian unitary ensemble. The GDs in an MMF are statistically described by the eigenvalues of the zero-trace Gaussian unitary ensemble.

The assumption that the diagonal and off-diagonal elements of $\mathcal{G}$ have the same variance σ_g^2 (12) is valid only if the D orthogonal vectors in $\mathcal{V}^{(k)}$ are independent of each other. However, the condition of orthogonality implies that the Dth vector is determined by the other $D-1$ vectors. Using numerical simulation, we have found that all diagonal elements of $\mathcal{G}$ have equal variance of $(1 - D^{-1})\sigma_g^2$, all off-diagonal elements of $\mathcal{G}$ have equal variance of $[1+1/D(D-1)]\sigma_g^2$, and the average variance of all elements of $\mathcal{G}$ is σ_g^2 (12). The theory of ([35], Sec. 14.3) is able to describe a random Hermitian matrix in which different elements have different variances. Numerical simulations of the zero-trace Gaussian unitary ensemble with unequal variances that is considered here show no observable differences from analytical results derived assuming all matrix elements have equal variance ([35], Sec. 3.3).

III. Modal Dispersion in Few-Mode Fibers

In the regime of strong mode coupling, the PMs and their GDs are given by the eigenvectors and eigenvalues of the zero-trace Gaussian unitary ensemble described by (10) and (13). Without loss of generality, after normalization, the elements of $\mathcal{G}$ may be assumed to be zero-mean identically distributed Gaussian random variables with variance $\sigma_g^2 = 1/2$, similar to the classic normalization of Mehta [35][2]. Before normalization, σ_g^2 is given by either (11) or (12). The diagonal elements of $\mathcal{G}$ are real with a variance of $\sigma_g^2 = 1/2$. The off-diagonal elements of $\mathcal{G}$ are complex Gaussian distributed with independent real and imaginary parts, each having a variance of $1/4$. From (12), each section can be taken to have a normalized standard deviation of GD given by $\sigma_{\tau} = \sqrt{D/2K}$. With this normalization, the notation in this section is similar to that in Mehta [35].

A. Joint Probability Density

The joint pdf for a Gaussian unitary ensemble without the zero-trace constraint is well known. The ordered joint pdf of the eigenvalues of a $D \times D$ Gaussian unitary ensemble is ([35], Sec. 3.3) [37]

$$\alpha_D \prod_{D \geq i > j > 0}(\lambda_i - \lambda_j)^2 \exp\left(-\sum_{i=1}^{D}\lambda_i^2\right) \tag{14}$$

where the eigenvalues possess the order constraint $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_D$ and are all real valued, and α_D is a constant such that the joint pdf integrates to unity. The eigenvalues are the normalized GDs. An analytical expression for α_D can be found in ([35], Theorem 3.3.1). Because a permutation with a different

[2]In some mathematical literature, which is relevant to Gaussian orthogonal ensembles but not Gaussian unitary ensembles, the matrix elements are assumed to have unit variance. Proportionality constants are ignored in some of those references.

ordering of eigenvalues is equivalent to any other permutation, the unordered joint pdf is just $1/D!$ of (14) but without the order constraint ([35], ch. 5).

With zero trace

$$\mathrm{tr}(\mathcal{G}) = \lambda_1 + \lambda_2 + \cdots + \lambda_D = 0 \tag{15}$$

the ordered joint pdf of $\lambda_1, \ldots, \lambda_{D-1}$ becomes

$$p_D(\lambda_1, \ldots, \lambda_{D-1}) = \beta_D \prod_{D \geq i > j > 0} (\lambda_i - \lambda_j)^2 \exp\left(-\sum_{i=1}^{D} \lambda_i^2\right) \tag{16}$$

with the order constraint

$$\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_D \tag{17}$$

where the constant β_D is determined by requiring (16) to integrate to unity, but is not the same as the α_D in (14). The unordered joint pdf is the same as (16) but is a factor of $1/D!$ smaller and does not have the order constraint (17). The statistical properties of GD are fully specified by the joint pdf (16) with the constraint (15).

B. Two-Mode Fiber

Two-mode fiber is the simplest case, and may correspond to the two polarization modes in an SMF, i.e., the well-known PMD problem. The purpose here is not to derive new properties of PMD, but to verify that the general random matrix model is applicable to PMD.

With $\lambda_2 = -\lambda_1$, the pdf (16) for $D = 2$ becomes

$$p_2(\lambda_1) = \beta_2 4\lambda_1^2 e^{-2\lambda_1^2}. \tag{18}$$

As in the PMD literature, we define $\lambda_{1,2} = \pm(1/2)\tau$ and find $\beta_2 = \sqrt{2/\pi}$, obtaining

$$p_2(\tau) = \sqrt{\frac{2}{\pi}} \tau^2 e^{-\tau^2/2}, \quad \tau \geq 0 \tag{19}$$

which is the well-known Maxwellian distribution with normalized mean DGD of $\bar{\tau} = 2\sqrt{2/\pi} = 1.60$. Random matrix models specialized to the two-mode case were used to derive the Maxwellian distribution in [29] and [31]. The second moment of both λ_1 and λ_2 is 3/4.

C. Three-Mode Fiber

Three-mode fiber is the next simplest case, with $\lambda_3 = -(\lambda_1 + \lambda_2)$. The joint pdf (16) for $D = 3$ of (17) is

$$p_3(\lambda_1, \lambda_2) = \beta_3 (\lambda_2 - \lambda_1)^2 (2\lambda_1 + \lambda_2)^2 (\lambda_1 + 2\lambda_2)^2 \times e^{-\lambda_1^2 - \lambda_2^2 - (\lambda_1 + \lambda_2)^2}. \tag{20}$$

Without the order constraint (17), the marginal pdf of the GDs is

$$p_{3\lambda}(\lambda) = \frac{1}{6} \int_{-\infty}^{+\infty} p_3(\lambda, \lambda_2) \mathrm{d}\lambda_2. \tag{21}$$

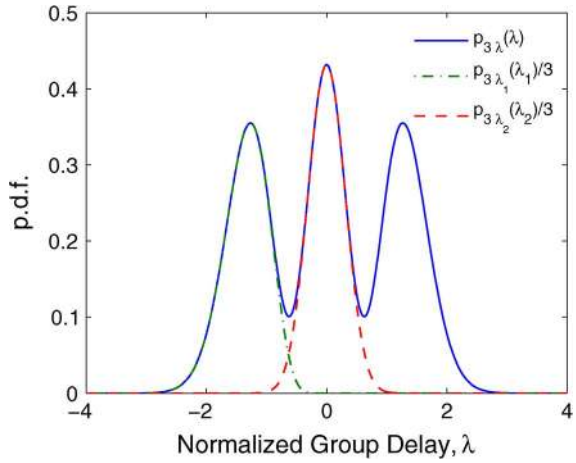

Fig. 1. Statistics of the normalized GDs for three-mode fiber, including the marginal pdf $p_{3\lambda}(\lambda)$, the pdf $p_{3\lambda_1}(\lambda_1)$ of the smallest delay, and the pdf $p_{3\lambda_2}(\lambda_2)$ of the middle delay. To facilitate comparison, both $p_{3\lambda_1}(\lambda_1)$ and $p_{3\lambda_2}(\lambda_2)$ are scaled by a factor of $1/3$.

Some algebra[3] yields the constant $\beta_3 = 4\sqrt{3}/\pi$ and

$$p_{3\lambda}(\lambda) = \frac{1}{16}\sqrt{\frac{6}{\pi}}(27\lambda^4 - 18\lambda^2 + 5)\exp\left(-\frac{3}{2}\lambda^2\right). \tag{22}$$

Fig. 1 plots the marginal pdf $p_{3\lambda}(\lambda)$, which exhibits three peaks, corresponding to the values where λ_1, λ_2, and λ_3 are concentrated. The pdf $p_{3\lambda}(\lambda)$ is symmetrical with respect to $\lambda = 0$ due to the symmetric nature of the three eigenvalues for the 3×3 random matrix $\mathcal{G}$. The middle eigenvalue λ_2 is concentrated near zero. The variance of λ is $\sigma_\lambda^2 = 4/3$.

The pdf of the eigenvalue λ_1, corresponding to the smallest delay, can also be found using (16) with the order constraint (17). The condition $\lambda_1 \leq 0$ is required in order to conform to the zero-trace constraint (15). The pdf of the smallest eigenvalue is

$$p_{3\lambda_1}(\lambda_1) = \int_{\lambda_1}^{-\frac{\lambda_1}{2}} p_3(\lambda_1, \lambda_2)\mathrm{d}\lambda_2 \tag{23}$$

or

$$p_{3\lambda_1}(\lambda_1) = 3p_{3\lambda}(\lambda_1)\mathrm{erf}\left(\frac{3|\lambda_1|}{\sqrt{2}}\right) - \frac{9\sqrt{3}\lambda_1(3\lambda_1^2 - 5)}{8\pi}\exp(-6\lambda_1^2), \quad \lambda_1 \leq 0. \tag{24}$$

Due to symmetric nature of λ_1 and λ_3, we have $p_{3\lambda_3}(\lambda_3) = p_{3\lambda_1}(-\lambda_3)$. Fig. 1 also shows $(1/3)p_{3\lambda_1}(\lambda_1)$ where the pdf is scaled by a factor $1/3$ such that $p_{3\lambda_1}(\lambda_1)$ given by (24) is nearly the same as $p_{3\lambda}(\lambda)$ given by (22) near the first peak of $p_{3\lambda}(\lambda)$.

[3]All the calculations performed for $D \geq 3$ require three steps. In the first step, $\lambda_1^2 + \lambda_2^2 + \cdots + \lambda_{D-1}^2 + (\lambda_1 + \lambda_2 + \cdots + \lambda_{D-1})^2$ is linearly transformed to $D/(D-1) \times \lambda_1^2 + x_2^2 + \cdots + x_{D-1}^2$. In the second step, the linear transform is substituted into $\prod_{D \geq i > j > 0}(\lambda_i - \lambda_j)^2$, which may be expanded to a summation of terms in the form of $c_{f_1, f_2, \ldots, f_{D-1}} \lambda_1^{f_1} x_2^{f_2} \cdots x_{D-1}^{f_{D-1}}$, where f_i are the exponents and $c_{f_1, f_2, \ldots, f_{D-1}}$ are the corresponding coefficients. The last step is the integration over $x_2, x_3, \ldots, x_{D-1}$ using ([38], Sec. 3.461). These calculations are tedious, but can be performed using symbolic mathematical software, such as Maple or MuPAD.

TABLE I
STATISTICAL PARAMETERS FOR THE NORMALIZED GDS OF A THREE-MODE FIBER

p.d.f.	Mean	Variance
$p_{3\lambda}(\lambda)$	0	$\frac{4}{3}$
$p_{3\lambda_1}(\lambda_1)$	$-\frac{27}{16}\sqrt{\frac{2}{\pi}}$	$\frac{4}{3}+\frac{9\sqrt{3}}{8\pi}-\frac{729}{128\pi}=0.1407$
$p_{3\lambda_2}(\lambda_2)$	0	$\frac{4}{3}-\frac{9\sqrt{3}}{4\pi}=0.0928$
$p_{3\lambda_3}(\lambda_3)$	$\frac{27}{16}\sqrt{\frac{2}{\pi}}$	0.1407
$p_{3(\lambda_3-\lambda_1)}(\lambda)$	$\frac{27}{8}\sqrt{\frac{2}{\pi}}$	$4-\frac{729}{32\pi}+\frac{27\sqrt{3}}{4\pi}=0.4700$

Similarly, the pdf of the middle eigenvalue λ_2 can be found by

$$p_{3\lambda_2}(\lambda_2)=\int_{-\infty}^{-|\lambda_2|} p_3(\lambda_1,\lambda_2)\mathrm{d}\lambda_1 \tag{25}$$

or

$$p_{3\lambda_2}(\lambda_2)=3p_{3\lambda}(\lambda_2)\mathrm{erfc}\left(\frac{3|\lambda_2|}{\sqrt{2}}\right)-\frac{9\sqrt{3}|\lambda_2|(3\lambda_2^2-5)}{8\pi}\exp(-6\lambda_2^2). \tag{26}$$

Comparing (22), (24), and (26), the marginal pdf $p_{3\lambda}(\lambda)$ (22) is found to be the combination of $p_{3\lambda_1}(\lambda_1)$, $p_{3\lambda_2}(\lambda_2)$, and $p_{3\lambda_3}(\lambda_3)$

$$p_{3\lambda}(\lambda)=\frac{1}{3}\left[p_{3\lambda_1}(\lambda)+p_{3\lambda_2}(\lambda)+p_{3\lambda_3}(\lambda)\right]. \tag{27}$$

Fig. 1 also shows $(1/3)p_{3\lambda_1}(\lambda_1)$ and $(1/3)p_{3\lambda_2}(\lambda_2)$ that are concentrated near the corresponding peaks of the marginal pdf $p_{3\lambda}(\lambda)$, confirming that each peak of the marginal pdf $p_{3\lambda}(\lambda)$ corresponds to an individual eigenvalue. The statistical parameters of the normalized GDs of a three-mode fiber are presented in Table I. In Fig. 1, the peak for the middle delay λ_2 is narrower than those for the maximum and minimum delays λ_1 and λ_3. In Table I, the variance of λ_2 is smaller than the variance of λ_1 or λ_3.

The difference between the maximum and minimum eigenvalues is the normalized delay spread of the MMF. The pdf of the delay spread for three-mode MMF is

$$p_{3(\lambda_3-\lambda_1)}(\lambda)=\frac{1}{2}\int_{-\lambda/3}^{\lambda/3} p_3\left(\frac{\lambda-\lambda_2}{2},\lambda_2\right)\mathrm{d}\lambda_2 \tag{28}$$

or

$$p_{3(\lambda_3-\lambda_1)}(\lambda)=\frac{\sqrt{3}}{4\pi}(\lambda^5-9\lambda^3)\exp\left(-\frac{2}{3}\lambda^2\right)+\frac{1}{8}\sqrt{\frac{2}{\pi}}\left(\lambda^6-6\lambda^4+27\lambda^2\right)\times\exp\left(-\frac{1}{2}\lambda^2\right)\mathrm{erf}\left(\frac{\lambda}{\sqrt{6}}\right),\quad \lambda>0. \tag{29}$$

The statistical parameters of the delay spread are also given in Table I. For three-mode fiber, many properties of the eigenvalues or normalized GDs can be computed analytically in closed form.

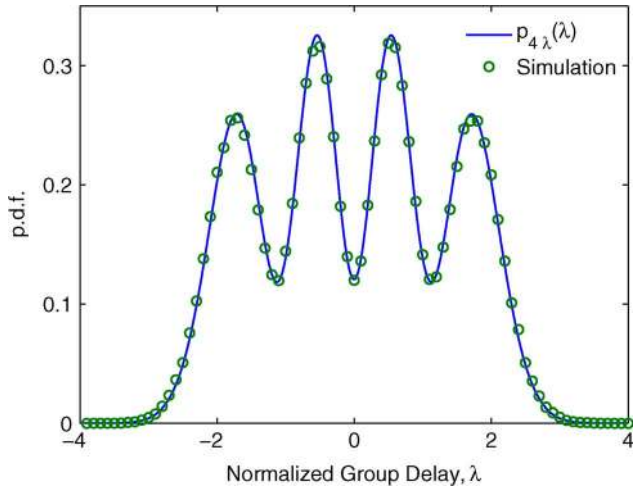

Fig. 2. Marginal pdf of the normalized GDs for a four-mode fiber, comparing analysis and simulation.

D. Four-Mode Fiber

Four-mode fiber is a particularly simple case beyond the two-mode fiber, as it represents a fiber with two spatial modes and two polarizations. Fibers with two spatial modes have been used for dispersion compensation [39] or fiber sensors [40]. We note, however, that a weakly guiding fiber with circular core cannot support exactly two spatial modes [41].

The pdf (16) for $D=4$ becomes

$$p_4(\lambda_1,\lambda_2,\lambda_3)=\beta_4\prod_{4\geq i,j\geq 1}(\lambda_i-\lambda_j)^2 e^{-\lambda_1^2-\lambda_2^2-\lambda_3^2-\lambda_4^2} \tag{30}$$

with the zero trace constraint $\lambda_4=-\lambda_1-\lambda_2-\lambda_3$.

Using the unordered joint pdf, the marginal pdf of λ is

$$p_{4\lambda}(\lambda)=\frac{1}{4!}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} p_4(\lambda,\lambda_2,\lambda_3)d\,\lambda_2 d\lambda_3. \tag{31}$$

After some calculations, we obtain

$$p_{4\lambda}(\lambda)=\frac{2\sqrt{3}e^{-4/3\lambda^2}}{\sqrt{\pi}}\times\left(\frac{4096}{6561}\lambda^6-\frac{1024}{729}\lambda^4+\frac{80}{81}\lambda^2+\frac{5}{81}\right) \tag{32}$$

which has a variance of 15/8.

Fig. 2 shows the marginal pdf of the normalized GDs in a four-mode fiber given by (32). The marginal pdf has four peaks, corresponding to the GD of four different PMs.

To verify the marginal pdf in Fig. 2, the modal dispersion of a four-mode MMF has been simulated. The fiber has $K=256$ independent sections. In each section, the four modes are chosen to have deterministic delays of $+\tau,+\tau,-\tau$, and $-\tau$ where $\tau=\sqrt{2/K}$ to ensure that the elements of $\mathcal{G}$ have a variance of $\sigma_g^2=1/2$. This particular choice could describe a fiber where, in each section, the DGD between the two polarization modes is negligible compared to that between the two spatial modes. The random unitary matrices $\mathcal{U}^{(k)}$ and $\mathcal{V}^{(k)}$, $k=1,\ldots,K$, are first initialized by 4×4 random complex Gaussian matrices and then converted to unitary matrices using the Gram–Schmidt process

([42], Sec. 5.2.8). All sections have independent matrices $\mathcal{U}^{(k)}$ and $\mathcal{V}^{(k)}$. A total of 400 000 eigenvalues are used in the curve shown in Fig. 2.

In Fig. 2, the simulation results show excellent agreement with the analytical pdf (32). Although the modes in each section have only two GDs, with strong mode coupling, a pdf having four peaks is obtained. In the strong-coupling regime, similar results would be obtained using any uncoupled GDs in each section, provided that the four GDs sum to zero and have a variance of $2/K$. For example, the four modes may have modal delays of $+\tau_1, +\tau_1, +\tau_1$, and $-3\tau_1$ with $\tau_1 = \sqrt{2/3K}$. Similar results may also be obtained if the GDs in each section are, for example, $+\tau_2, +\tau_2, -\tau_2, -\tau_2$, where τ_2 follows a statistical distribution with second moment $2/K$.

The marginal pdf of the smallest or largest eigenvalues λ_1 and λ_4 may be found by suitable integration of (30). Unlike the case of a three-mode fiber, it does not seem possible to obtain closed-form expressions for the individual marginal pdf's of the ordered eigenvalues.

E. Other Few-Mode Fibers

The marginal pdf's of the GDs in an MMF with larger number of modes may also be obtained analytically. Following the aforementioned procedure, the marginal pdf's of fiber with five, six, and seven modes are

$$p_{5\lambda}(\lambda) = \frac{\sqrt{5}e^{-5/4\,\lambda^2}}{\sqrt{\pi}}\left(\frac{78125}{196608}\lambda^8 - \frac{15625}{8192}\lambda^6 + \frac{24375}{8192}\lambda^4 - \frac{1975}{2048}\lambda^2 + \frac{903}{4096}\right) \tag{33}$$

$$p_{6\lambda}(\lambda) = \frac{\sqrt{30}e^{-6/5\lambda^2}}{\sqrt{\pi}}\left(\frac{13436928}{244140625}\lambda^{10} - \frac{4478976}{9765625}\lambda^8 + \frac{2581632}{1953125}\lambda^6 - \frac{102816}{78125}\lambda^4 + \frac{7812}{15625}\lambda^2 + \frac{644}{15625}\right) \tag{34}$$

and

$$p_{7\lambda}(\lambda) = \frac{\sqrt{42}e^{-7/6\lambda^2}}{\sqrt{\pi}} \times \left(\frac{1977326743}{146932807680}\lambda^{12} - \frac{282475249}{1632586752}\lambda^{10} + \frac{98001617}{120932352}\lambda^8 - \frac{47883143}{30233088}\lambda^6 + \frac{17707375}{13436928}\lambda^4 - \frac{212219}{746496}\lambda^2 + \frac{88175}{1492992}\right) \tag{35}$$

respectively. The variances of these distributions are 12/5, 35/12, and 24/7, respectively. For all the cases of D from two to seven modes, the variances are given by $(1/2)(D - D^{-1})$, a reduction by a factor of $1 - D^{-2}$ compared with the case without the zero-trace constraint.

Fig. 3 shows the marginal pdf of the normalized GDs of fibers with five, six, and seven modes. The number of peaks in the marginal pdf is the same as the number of modes. In general, the peaks closer to $\lambda = 0$ are both higher and narrower than those farther from the origin. Those peaks cause ripples to appear in the marginal pdf, and the ripples are still significant in a seven-mode MMF. As the number of modes increases, the peaks in the marginal pdf move closer together and merge. As the peaks become indistinguishable for fibers having many modes, the marginal pdf should approach a limiting distribution.

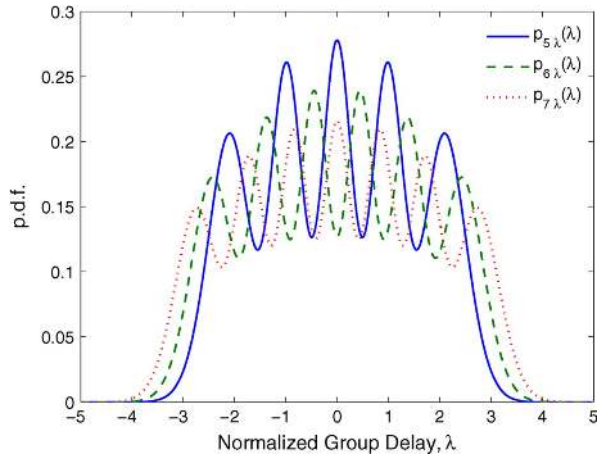

Fig. 3. Marginal pdf of normalized GDs for fibers with five, six, and seven modes.

Numerical simulations, similar to those in Fig. 2, have been used to verify the analytical pdf's in Fig. 3. In all the cases, simulation and theory match with each other.

Numerical simulations have been conducted to further verify the variance reduction factor of $1 - D^{-2}$, which is most significant for $D \leq 3$. Random realizations of zero-trace matrices $\mathcal{G}$ of the form (10) exhibit no observable variance reduction, although the empirically estimated pdf's, when scaled by this reduction factor, are found to match with (19) and (22).

The reduction of variance may be seen as related to degrees of freedom. A random Hermitian matrix without zero-trace constraint has D^2 degrees of freedom, corresponding to $D(D-1)/2$ complex off-diagonal elements and D real diagonal elements. The zero-trace constraint reduces the degrees of freedom by one, proportionally affecting a fraction $1/D^2$ of matrix elements. The zero-trace constraint from (14)–(16) reduces D degrees of freedom to $D - 1$ degrees of freedom, proportionally affecting a fraction $1/D$ of matrix elements. While analytical results scaled to the same variance are consistent with numerical simulations, the variance reduction factor of $1 - D^{-2}$ requires further study.

In the simplest case of $D = 2$, a zero-trace Gaussian unitary ensemble can be generated numerically by three methods. The first method is based on $\mathcal{G}$ given by the summation (10), for example, with $\sigma_g^2 = 1/2$. The second method is based on a random 2×2 Hermitian matrix $\mathcal{A}$ but with $a_{2,2}$ replaced by $-a_{1,1}$. The third method is based on generating random 2×2 Hermitian matrix $\mathcal{A}$, finding its eigenvalues λ_1 and λ_2, and selecting those with $|\lambda_1 + \lambda_2|$ smaller than a small number. In the second and third methods, the elements of the matrix $\mathcal{A}$ have variance of 1/2. The first two methods give eigenvalues with the same variance but the third method gives eigenvalues with a variance 3/4 time smaller than the first two methods. In the third method, the variance of the diagonal elements is 1/4 and

the off-diagonal elements have variance of 1/2. The zero-trace constraint on the eigenvalues (15) reduces the variance of the matrix elements but the pdf maintains the same shape.

The variance reduction of the diagonal elements is due to the zero-trace constraint (15), which selects those matrices with smaller diagonal elements (in the general case, a factor of $1 - D^{-1}$ smaller). The average variance among all elements is a factor $1 - D^{-2}$ smaller, the same as the reduction factor for the variance of the eigenvalues.

As the average variance for all matrix elements of $\mathcal{G}$ is σ_g^2 given by (12), for all MMF studies in this section, and without normalization, the variance of the GD is

$$\sigma_{\mathrm{gd}}^2 = K\sigma_\tau^2 \tag{36}$$

with σ_τ^2 defined by (12). If the GD is characterized by its standard deviation σ_{gd}, it is always proportional to the square root of the number of independent MMF sections.

IV. Modal Dispersion in Many-Mode Fibers

With a large number of modes, a Gaussian unitary ensemble without the zero-trace constraint is described by a semicircle distribution with radius $\sqrt{2D}$ ([35], Sec. 4.2). With the normalization used in Section III, the variance of the eigenvalues is $D/2$. This semicircle law was first derived by Wigner for large random matrices [43], [44]. The Wigner semicircle law is universally valid for many different types of large random matrices [45], [46]. A Gaussian unitary ensemble, even with the zero-trace constraint (15), should follow the semicircle distribution. As an alternative to considering $\mathcal{G}$ as a Gaussian unitary ensemble, a more straightforward derivation using the CLT for free random variables is given in a later part of this section.

In free probability theory, free random variables are equivalent to statistically independent large random matrices [47], [48]. The CLT for the summation of free random variables gives the semicircle distribution [47], [49]. The matrix $\mathcal{G}$ (10) is the summation of many independent random matrices. The CLT for free random variables states the following: Let $\mathcal{X}_k$, $k = 1, \ldots, K$, be identically distributed independent zero-mean free random variables with unit variance. The summation

$$\mathcal{Y}_K = \frac{\mathcal{X}_1 + \mathcal{X}_2 + \cdots + \mathcal{X}_K}{\sqrt{K}} \tag{37}$$

is described by semicircle distribution with radius of two and unit variance

$$p_Y(r) = \begin{cases} \frac{1}{2\pi}\sqrt{4 - r^2} & |r| < 2 \\ 0 & \text{otherwise} \end{cases} \tag{38}$$

as K approaches infinity.

In the aforementioned theorem, when free random variables are represented by large random matrices, the distribution of a free random variable is equivalent to the distribution for the eigenvalues of the random matrices. When the CLT of free random variables is applied to $\mathcal{G}$ given by (10), if the variance of the zero-mean GD per section is σ_τ^2 for all K sections, the eigenvalues of $\mathcal{G}$ are described by a semicircle distribution with radius $2\sqrt{K}\sigma_\tau$ and variance $K\sigma_\tau^2$. Equivalently, the GD of the MMF has a semicircle distribution with variance $K\sigma_\tau^2$. Note that the normalization used in this section based on the eigenvalues of $\mathcal{X}_k$ and $\mathcal{Y}_K$ in (37) is customary in free probability theory. However, the normalization used in Section III is based on the matrix elements of $\mathcal{G}$, similar to that in Mehta [35].

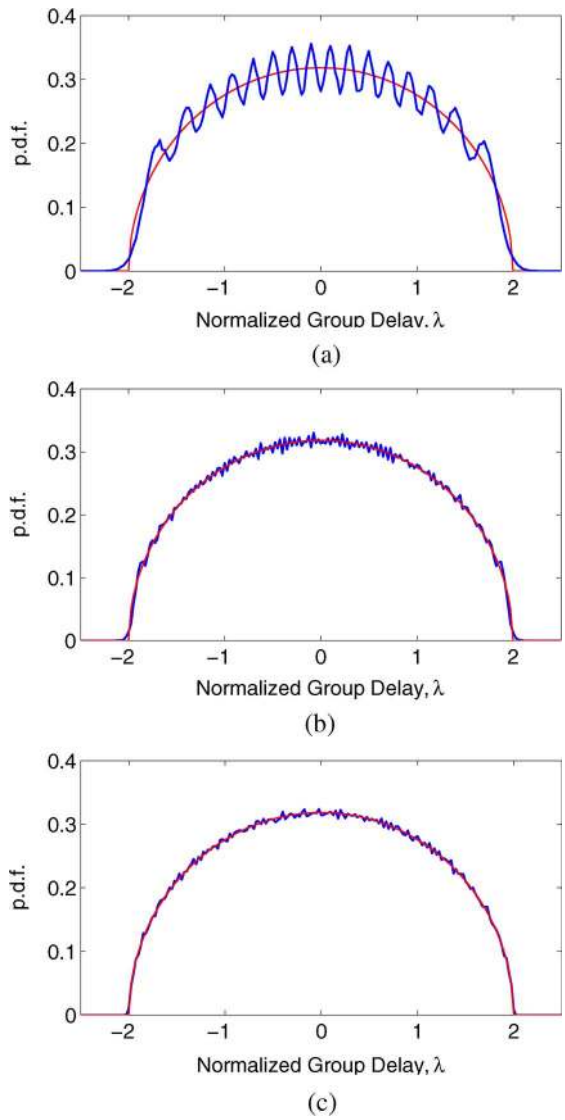

Fig. 4. Simulated marginal pdf of normalized GDs (ripply blue curves) compared with semicircle distribution (smooth red curves) in an MMF with (a) 16, (b) 64, and (c) 512 modes.

Fig. 4 compares the simulated marginal pdf of MMFs having $D = 16, 64$, and 512 modes to the semicircle distribution. Each MMF is comprised of $K = 256$ sections. In each section, the GDs are deterministic, with the first $D/2$ modes with a delay of τ and the other $D/2$ modes with a delay of $-\tau$. The simulated curves are obtained from 1 600 000, 640 000, and 102 400 eigenvalues for $D = 16, 64$, and 512, respectively, using a step size of 0.025 along the λ-axis. The normalization $\tau = 1/\sqrt{K}$ is made to facilitate comparison with a semicircle distribution with radius of 2. The model here is valid as long as the component matrices in (10) may be modeled as free random variables [47].

In Fig. 4, the simulated distributions match the semicircle distribution well for $D = 64$ and 512 modes. For a fiber having

$D = 16$ modes, the distribution is close to a semicircle distribution, but has an obvious periodic structure with 16 peaks. The ripples become less obvious as D increases from 16 to 64 to 512. Upon close examination of the curve for $D = 64$, the ripples seem periodic, similar to those in an MMF having $D = 16$ modes, but much smaller.

The semicircle distribution, describing the GDs in an MMF with an infinite number of modes, has strict upper and lower limits, and, thus, a strictly bounded GD spread. In designing systems for MMF with a finite but large number of modes D, it will be sufficient to provide a GD tolerance just slightly larger than the maximum GD spread of the semicircle distribution, which is given by $4\sigma_{\mathrm{gd}}$.

The GD relationship (36) remains valid when the number of modes D is vary large. With a large number of modes, the relationship $\sigma_{\mathrm{gd}}^2 = K\sigma_\tau^2$ can be derived directly from free probability theory.

V. Discussion

The scaling of modal dispersion with fiber length in an MMF is similar to the scaling of PMD in an SMF. In an MMF shorter than the correlation length over which the local PMs can be considered constant, the GD increases linearly with fiber length. By contrast, in an MMF much longer than the correlation length, the number of independent sections K is large, and strong mode coupling can be assumed. In the strong-coupling regime, a parameter describing GD per unit length may be defined as $\sigma_{\mathrm{km}} = \sigma_\tau/\sqrt{L_s}$, where L_s is the fiber length per section, measured in kilometers. The overall GD, if characterized by σ_{gd} given by (36), is equal to $\sqrt{L_t}\sigma_{\mathrm{km}}$, where $L_t = KL_s$ is the total fiber length.

In practice, there are advantages of introducing strong mode coupling in order to reduce the modal delay spread. In direct-detection systems, this can reduce intersymbol interference, whereas in systems using coherent detection, this can reduce the temporal memory required in digital compensation of modal dispersion. Recent MDM experiments [21]–[23], performed in short spans of an MMF, were probably not in the strong-coupling regime. Future long-distance systems are likely to be operated in the strong-coupling regime, especially if strong mode coupling is used to reduce the overall GD spread. In an MMF, spatial mode coupling is governed, in part, by mode groups [50]. Typically, coupling between modes in different groups is weak, with coupling length as long as 25 km [6], while coupling between modes in the same group is strong, with coupling length less than 1 km [17]. In order to reduce the GD spread in an MMF, coupling between modes in different groups should be enhanced. In manufacturing of an SMF, spinning is used to reduce the polarization coupling length below 100 m, thereby reducing the DGD due to PMD [26]. Manufacturing processes for an MMF may perhaps be modified to increase spatial mode coupling in order to reduce the GD spread.

As seen in Figs. 3 and 4(a), in the marginal pdf of GD, the number of peaks is the same as the number of modes, and the separation betweens adjacent peaks (relative to the semicircle radius) decreases with an increasing of number of modes. In the absence of the zero-trace constraint (15), ripples can be observed in Gaussian unitary ensembles up to at least 51×51 ([35], Fig. 6.1). With a zero-trace constraint, the ripples are larger than those without the constraint. In Fig. 4(b) with $D = 64$ modes, ripples are observable and seems to be very regular. As the number of modes increases, the ripples becomes narrower, similar to the Gibbs phenomenon [51], [52] for Fourier series.

Higher order modal dispersion effects are outside the scope of this paper. In higher order modal dispersion, the PMs and their GDs can vary with frequency [53]. These effects are analogous to polarization-dependent chromatic dispersion and depolarization observed in an SMF with PMD [54], [55]. In the case of an SMF with PMD, the properties of PMD to arbitrary order depend on a single parameter, which may be taken to be the GD standard deviation. In the case of an MMF with modal dispersion and strong coupling, the higher order properties of modal dispersion depend on the number of modes and a single parameter, which may be taken to be the GD standard deviation σ_{gd} given by (36), and may also depend on chromatic dispersion parameters, when they are spatial-mode-dependent.

In this paper, we have studied the distribution of GDs, but not the impulse response of an MMF. At a single frequency, the impulse response of a D-mode fiber consists of D narrow pulses with GDs described by the distribution (16) [33], and with weights depending on the PMs excited by the transmitter launch conditions. Considering a modulated signal occupying a finite bandwidth, because of higher order effects [53], those D narrow pulses broaden and may merge with each other. The overall duration of the impulse response is described by the duration of the pdf of the GD, as shown in Figs. 1–4. In a fiber with many modes, where the pdf is the semicircle distribution shown in Fig. 4, the impulse response duration is just $4\sigma_{\mathrm{gd}}$.

VI. Conclusion

In the regime of strong mode coupling, an MMF may be modeled as the cascade of *many* independent sections, which are described by statistically independent random matrices. The GDs are given by the eigenvalues of Gaussian unitary ensemble with zero-trace constraint. Marginal pdf's of the GDs in fibers with two to seven modes have been derived analytically. Numerical simulations of the pdf's are in excellent agreement with analytical results. In a fiber with many modes, the GD is shown to follow a semicircle distribution from free probability theory. Numerical simulations have been conducted for fibers having $D = 16, 64$, and 512 modes to compare to the semicircle distribution.

References

[1] *Carrier Sense Multiple Access With Collision Detection (CSMA/CD) Access Method and Physical Layer Specifications*, IEEE 802.3 Standard, 2008.

[2] A. F. Benner, M. Ignatowski, J. A. Kash, D. M. Kuchta, and M. B. Ritter, "Exploitation of optical interconnects in future server architectures," *IBM J. Res. Dev.*, vol. 49, pp. 755–775, 2005.

[3] Y. Koike and S. Takahashi, "Plastic optical fibers: Technologies and communication links," in *Optical Fiber Telecommunications VB: Systems and Networks*, I. P. Kaminow, T. Li, and A. E. Willner, Eds. San Diego, CA: Academic, 2008.

[4] D. Gloge, "Optical power flow in multimode fiber," *Bell Sys. Tech. J.*, vol. 51, no. 8, pp. 1767–1780, 1972.

[5] A. Ghatak and K. Thyagarajan, "Graded index waveguides—A review," *Prog. Opt.*, vol. 18, pp. 1–126, 1980.

[6] K.-I. Kitayama, S. Sikai, and N. Uchida, "Impulse response prediction based on experimental mode-coupling coefficient in a 10-km long graded-index fiber," *J. Quantum Electron.*, vol. QE-16, no. 3, pp. 356–362, 1980.

[7] R. Olshansky, "Mode-coupling effects in graded-index optical fibers," *Appl. Opt.*, vol. 14, no. 4, pp. 935–945, 1975.

[8] L. Raddatz, I. H. White, D. G. Cunningham, and M. C. Nowell, "An experimental and theoretical study of the offset launch technique for the enhancement of the bandwidth of multimode fiber links," *J. Lightw. Technol.*, vol. 16, no. 3, pp. 324–331, Mar. 1998.

[9] *Carrier Sense Multiple Access With Collision Detection (CSMA/CD) Access Method and Physical Layer Specifications Amendment 2: Physical Layer and Management Parameters for 10 Gb/s Operation, Type 10 GBASE-LRM*, IEEE Standard 802.3aq, 2006.

[10] *Carrier Sense Multiple Access With Collision Detection (CSMA/CD) Access Method and Physical Layer Specifications Amendment 4: Media Access Control Parameters, Physical Layers and Management Parameters for 40 Gb/s and 100 Gb/s Operation*, IEEE Standard 802.3ba, 2010.

[11] K. Roberts, D. Beckett, D. Boertjes, J. Berthold, and C. Laperle, "100 G and beyond with digital coherent signal processing," *IEEE Commun. Mag.*, vol. 48, no. 7, pp. 62–69, Jul. 2010.

[12] J. Yu and X. Zhou, "Ultra-high-capacity DWDM transmission system for 100 G and beyond," *IEEE Commun. Mag.*, vol. 48, no. 3, pp. S56–S64, Mar. 2010.

[13] J. M. Kahn and K.-P. Ho, "Spectral efficiency limits and modulation/detection techniques for DWDM systems," *IEEE J. Sel. Topics Quantum Electron.*, vol. 10, no. 2, pp. 259–272, Mar./Apr. 2004.

[14] P. J. Winzer, "Beyond 100 G Ethernet," *IEEE. Commun. Mag.*, vol. 48, no. 7, pp. 26–30, Jul. 2010.

[15] P. P. Mitra and J. B. Stark, "Nonlinear limits to the information capacity of optical fibre communications," *Nature*, vol. 411, pp. 1027–1030, 2001.

[16] R.-J. Essiambre, G. Kramer, P. J. Winzer, G. J. Foschini, and B. Goebel, "Capacity limits of optical fiber networks," *J. Lightw. Technol.*, vol. 28, no. 4, pp. 662–701, Feb. 2010.

[17] S. Berdagué and P. Facq, "Mode division multiplexing in optical fibers," *Appl. Opt.*, vol. 21, pp. 1950–1955, 1982.

[18] H. R. Stuart, "Dispersive multiplexing in multimode optical fiber," *Science*, vol. 289, pp. 281–283, 2000.

[19] A. R. Shah, R. C. J. Hsu, A. Tarighat, A. H. Sayed, and B. Jalali, "Coherent optical MIMO (COMIMO)," *J. Lightw. Technol.*, vol. 23, no. 8, pp. 2410–2419, Aug. 2005.

[20] M. Nazarathy and A. Agmon, "Coherent transmission direct detection MIMO over short-range optical interconnects and passive optical networks," *J. Lightw. Technol.*, vol. 26, no. 14, pp. 2037–2045, Jul. 2008.

[21] A. Li, A. Al Amin, X. Chen, and W. Shieh, "Reception of mode and polarization multiplexed 107-Gb/s CO-OFDM signal over a two-mode fiber," in *Proc. Opt. Fiber Commun.*, 2011, pp. 1–3, Paper PDPB8.

[22] M. Salsi, C. Koebele, D. Sperti, P. Tran, P. Brindel, H. Margoyan, S. Bigo, A. Boutin, F. Verluise, P. Sillard, M. Bigot-Astruc, L. Provost, F. Cerou, and G. Charlet, "Transmission at 2×100 Gb/s, over two modes of 40 km-long prototype few-mode fiber, using LCOS based mode multiplexer and demultiplexer," presented at the Opt. Fiber Commun., Los Angeles, CA, 2011, Paper PDPB9.

[23] R. Ryf, S. Randel, A. H. Gnuack, C. Bolle, R.-J. Essiambre, P. Winzer, D. W. Peckham, A. McCurdy, and R. Lingle, "Space-division multiplexing over 10 km of three-mode fiber using coherent 6×6 MIMO processing," presented at the Opt. Fiber Commun., Los Angeles, CA, 2011, Paper PDPB10.

[24] A. F. Garito, J. Wang, and R. Gao, "Effects of random perturbations in plastic optical fibers," *Science*, vol. 281, pp. 962–967, Aug. 1998.

[25] S. S.-H. Yam, F.-T. An, M. E. Marhic, and L. G. Kazovsky, "Polarization sensitivity of 40 Gb/s transmission over short-reach 62.5 μm multimode fiber," presented at the Opt. Fiber Commun., Los Angeles, CA, 2011, Paper FA-5.

[26] M. J. Li and D. A. Nolan, "Fiber spin-profile designs for producing fibers with low polarization mode dispersion," *Opt. Lett.*, vol. 23, pp. 1659–1661, 1998.

[27] C. D. Poole and R. E. Wagner, "Phenomenological approach to polarization dispersion in long single-mode fibers," *Electron. Lett.*, vol. 22, pp. 1029–1030, 1986.

[28] H. Kogelnik, R. M. Jopson, and L. E. Nelson, "Polarization-mode dispersion," in *Optical Fiber Telecommunications IVB: Systems and Impairments*, I. Kaminow and T. Li, Eds. San Diego, CA: Academic, 2002.

[29] C. D. Poole and J. A. Nagel, "Polarization effects in lightwave systems," in *Optical Fiber Telecommunications IIIA*, I. P. Kaminow and T. L. Koch, Eds. San Diego, CA: Academic, 1997, pp. 114–161.

[30] G. J. Foschini and C. D. Poole, "Statistical theory of polarization dispersion in single mode fibers," *J. Lightw. Technol.*, vol. 9, no. 11, pp. 1439–1456, Nov. 1991.

[31] M. Karlsson, "Probability density functions of the differential group delay in optical fiber communication systems," *J. Lightw. Technol.*, vol. 19, no. 3, pp. 324–331, Mar. 2001.

[32] S. Fan and J. M. Kahn, "Principal modes in multi-mode waveguides," *Opt. Lett.*, vol. 30, no. 2, pp. 135–137, 2005.

[33] M. B. Shemirani, W. Mao, R. A. Panicker, and J. M. Kahn, "Principal modes in graded-index multimode fiber in presence of spatial- and polarization-mode coupling," *J. Lightw. Technol.*, vol. 27, no. 10, pp. 1248–1261, May 2009.

[34] X. Shen, J. M. Kahn, and M. A. Horowitz, "Compensation for multimode fiber dispersion by adaptive optics," *Opt. Lett.*, vol. 30, no. 22, pp. 2985–2987, 2005.

[35] M. L. Mehta, *Random Matrices*, 3rd ed. San Diego, CA: Academic, 2004.

[36] J. P. Gordon and H. Kogelnik, "PMD fundamentals: Polarization mode dispersion in optical fibers," *Proc. Natl. Acad. Sci.*, vol. 97, no. 9, pp. 4541–4550, 2000.

[37] J. Ginibre, "Statistical ensembles of complex, quaternion, and real matrices," *J. Math. Phys.*, vol. 6, pp. 440–450, 1965.

[38] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 7th ed. San Diego, CA: Academic, 2007.

[39] C. D. Poole, J. M. Wiesenfeld, and D. J. DiGiovanni, "Elliptical-core dual-mode fiber dispersion compensator," *IEEE Photon. Technol. Lett.*, vol. 5, no. 2, pp. 194–197, Feb. 1993.

[40] B. Y. Kim, J. N. Blake, S. Y. Huang, and H. J. Shaw, "Use of highly elliptical core fibers for two-mode fiber devices," *Opt. Lett.*, vol. 12, no. 9, pp. 729–731, 1987.

[41] D. Gloge, "Weakly guiding fibers," *Appl. Opt.*, vol. 10, pp. 2252–2258, 1970.

[42] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed. Baltimore, MD: The Johns Hopkins Univ. Press, 1996.

[43] E. Wigner, "Characteristic vectors of bordered matrices with infinite dimensions," *Ann. Math.*, vol. 62, pp. 548–564, 1955.

[44] E. Wigner, "On the distribution of the roots of certain symmetric matrices," *Ann. Math.*, vol. 67, pp. 325–328, 1958.

[45] T. Tao and V. H. Vu, "From the Littlewood–Offord problem to the circular law: Universality of the spectral distribution of random matrices," *Bull. Amer. Math. Soc.*, vol. 46, no. 3, pp. 337–396, 2009.

[46] L. Erdős, J. Ramírez, B. Schlein, T. Tao, V. H. Vu, and H.-T. Yau, "Bulk universality for Wigner hermitian matrices with subexponential decay," *Math. Res. Lett.*, vol. 17, no. 4, pp. 667–674, 2010.

[47] D. Voiculescu, K. Dykema, and A. Nica, *Free Random Variables, CRM Monograph Series*. Providence, RI: American Mathematical Society, 1992, vol. 1.

[48] A. Nica and R. Speicher, *Lectures on the Combinatorics of Free Probability, London Mathematical Society Lecture Note Series*. Cambridge, U.K.: Cambridge Univ. Press, 2006, vol. 335.

[49] D. Voiculescu, "Limit laws for random matrices and free products," *Invent. Math.*, vol. 104, no. 1, pp. 201–220, 1991.

[50] P. Pepeljugoshi, S. E. Golowich, A. J. Ritger, P. Kolesar, and A. Risteski, "Modeling and simulation of next-generation multimode fiber links," *J. Lightw. Technol.*, vol. 21, no. 5, pp. 1242–1255, May 2003.

[51] E. Hewitt and R. E. Hewitt, "The Gibbs–Wilbraham phenomenon: An episode in Fourier analysis," *Arch. History Exact Sci.*, vol. 21, no. 2, pp. 129–160, 1979.

[52] D. Gottlieb and C.-W. Shu, "On the Gibbs phenomenon and its resolution," *SIAM Rev.*, vol. 39, no. 4, pp. 644–668, 1997.

[53] M. B. Shemirani and J. M. Kahn, "Higher-order modal dispersion in graded-index multimode fiber," *J. Lightw. Technol.*, vol. 27, no. 23, pp. 5461–5468, Dec. 2009.

[54] 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.

[55] G. J. Foschini, L. E. Nelson, R. M. Jopson, and H. Kogelnik, "Statistics of second-order PMD depolarization," *J. Lightw. Technol.*, vol. 19, no. 12, pp. 1882–1886, Dec. 2001.

Keang-Po Ho (S'91–M'95–SM'03) received the B.S. degree from National Taiwan University, Taipei, Taiwan, in 1991, and the M.S. and Ph.D. degrees from the University of California, Berkeley, in 1993 and 1995, respectively, all in electrical engineering.

He was involved in research at IBM T. J. Watson Research Center, Hawthorne, NY, on all-optical networks in the summer of 1994. From 1995 to 1997, he was a Research Scientist with Bellcore, currently Telcordia Technologies, Red Bank, NJ. From 1997 to 2001, he was with the Department of Information Engineering, the Chinese University of Hong Kong. He was the Chief Technology Officer and Co-Founder of StrataLight Communications (acquired by OpNext by about US$170M), Campbell, CA. From 2003 to 2006, he was with the Institute of Communication Engineering and the Department of Electrical Engineering, National Taiwan University. Since 2006, he has been with SiBEAM (now Wireless Division of Silicon Image) as a Principal Engineer and a Senior Manager for the Baseband Algorithm Group, and as the Primary Inventor of wirelessHD video area networking using 60 GHz millimeter wave. He has published more than 200 technical articles, and authored the book *Phase-Modulated Optical Communication Systems* (New York: Springer, 2005). His research interests include high-speed communication systems via optical fiber, copper wire, or radio frequency.

Joseph M. Kahn (M'90–SM'98–F'00) received the A.B., M.A., and Ph.D. degrees in physics from the University of California (UC), Berkeley, in 1981, 1983, and 1986, respectively.

From 1987 to 1990, he was at AT&T Bell Laboratories, Crawford Hill Laboratory, Holmdel, NJ. He demonstrated multi-Gb/s coherent optical fiber transmission systems, setting world records for receiver sensitivity. From 1990 to 2003, he was in the Faculty of the Department of Electrical Engineering and Computer Sciences, UC, Berkeley, where he was involved in research on optical and wireless communications. Since 2003, he has been a Professor in the Department of Electrical Engineering, Stanford University, Stanford, CA. In 2000, he helped in establishing StrataLight Communications, where he served as a Chief Scientist from 2000 to 2003. His research interests include single mode and multimode optical fiber communications, free-space optical communications, and microelectromechanical system for optical communications.

Dr. Kahn received the National Science Foundation Presidential Young Investigator Award in 1991. From 1993 to 2000, he served as the Technical Editor of the IEEE Personal Communications Magazine. Since 2009, he has been an Associate Editor of the *IEEE/OSA Journal of Optical Communications and Networking*.