Effect of Mode Coupling on Signal Processing Complexity in Mode-Division Multiplexing

Sercan Ö. Arık, Daulet Askarov, and Joseph M. Kahn, *Fellow, IEEE*

***Abstract*—Mode-division multiplexing systems employ multi-input multi-output (MIMO) equalization to compensate for chromatic dispersion (CD), modal dispersion (MD) and modal crosstalk. The computational complexity of MIMO equalization depends on the number of modes and on the group delay (GD) spread arising from CD and MD. Assuming the strong-coupling regime, in which the total system length far exceeds the correlation length of modal fields, we quantify the GD spread arising from MD, showing that it can be reduced significantly by mode coupling. We evaluate the computational complexity of various MIMO single-carrier equalizers, considering separate or combined equalization of CD and MD, in the time or frequency domain. We present numerical examples for the optimally designed graded-index depressed-cladding fibers supporting $D = 6$, 12, 20 or 30 modes in two polarizations. Assuming a 2000-km system length, a 1-km correlation length, and a combined CD+MD frequency-domain equalizer, the complexity (in complex multiplications per two-dimensional symbol) is a factor 1.4, 1.7, 2.2, 2.8 times higher for $D = 6$, 12, 20, 30 than for polarization-multiplexed systems in standard single-mode fiber ($D = 2$).**

***Index Terms*—DSP complexity, equalization, MIMO, mode coupling, mode-division multiplexing, multimode coherent receiver, multimode fiber, receiver signal processing.**

I. Introduction

THE continued growth of data traffic has motivated research on increasing long-haul optical transmission system capacity [1]. Promising approaches to move beyond the limits of single-mode fibers include spatial multiplexing in multi-core fibers or mode-division multiplexing (MDM) in multi-mode fibers (MMFs), a form of multi-input multi-output (MIMO) transmission [2]–[4]. In MDM, system capacity ideally scales in proportion to the number of modes employed [3], [4]. At present, MMF is widely used for short-range optical links because of relaxed connector alignment tolerances and reduced transceiver component costs [5].

Current long-haul systems already use multiplexing in the two polarization modes of a single-mode fiber (SMF), enabled by coherent detection and digital signal processing (DSP) to compensate for chromatic dispersion (CD), polarization-mode dispersion (PMD) and polarization crosstalk [6], [7]. The latter effects are compensated using a 2×2 MIMO equalizer. Increasing the number of spatial modes beyond one tends to increase DSP complexity, because of the increased dimensionality of the MIMO equalizer, and because of the group delay (GD) spread associated with modal dispersion (MD), which can exceed the GD spread associated with PMD or CD. One of the primary goals of MDM is reducing energy consumption per transmitted bit [8]. Hence, the computational complexity and resulting energy consumption of the MIMO equalizer are of critical importance.

There are two general approaches to reducing MIMO equalizer complexity. The first approach is to design the entire transmission system to minimize mode coupling crosstalk [9]–[11], with goal of using a sparse MIMO equalizer. While possible in principle, this approach requires low crosstalk in all system components, including modal (de)multiplexers, transmission fibers, amplifiers, and optical switches, and may be problematic in practice. The second approach is to design the transmission system for low GD spread, with goal of using a temporally short MIMO equalizer. Several recent MDM experiments [12]–[15] used fibers supporting six modes (two polarization modes and three spatial modes, the latter comprising two degenerate mode groups) with the index profile optimized for low uncoupled GD spread. Achieving low uncoupled GD spread is likely to be more challenging when the number of modes exceeds six (the number of degenerate mode groups exceeds two); some approaches are discussed in [16], [17].

Strong mode coupling can help reduce the GD spread in MMF [18], [19]. This approach is synergistic with a low uncoupled GD spread, since the latter tends to enhance mode coupling. Unintentional mode coupling can arise from random index perturbations, bends, twists or crosstalk at modal (de)multiplexers and other components. Furthermore, mode coupling might be enhanced via mode couplers or scramblers [20] or by intentional perturbation of the MMF [21]. The latter would be analogous to the spinning of SMF employed to reduce the GD spread due to PMD [22]. In the presence of mode coupling, modal fields [23] in a MMF are correlated over a correlation length, which is a generalization of the polarization correlation length in SMF [24], [25]. In the strong-coupling regime, when the total system length far exceeds the correlation length, the GD spread arising from MD is reduced significantly, and scales in proportion to the square-root of the total system length or equivalently the number of independent sections [18], [19].

Aside from reducing GD spread and thus MIMO equalization complexity, strong mode coupling can also mitigate the effects of mode-dependent loss or gain (collectively referred to as MDL), increasing average channel capacity [3], [26]. In the

Manuscript received June 12, 2012; revised November 15, 2012; accepted December 07, 2012. Date of publication December 28, 2012; date of current version January 02, 2013. This research was supported by National Science Foundation Grant Number ECCS-1101905, Corning, Inc., and in part by a Stanford Graduate Fellowship.

The authors are with E. L. Ginzton Laboratory, Department of Electrical Engineering, Stanford University, Stanford, CA 94305 USA (e-mail: soarik@stanford.edu; daskarov@stanford.edu; 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.2012.2234083

presence of MDL, channel capacity becomes a random variable, and system outage occurs with finite probability. Strong mode coupling, in conjunction with MD, can lead to frequency diversity, significantly reducing outage probability [27].

In this paper, we study the effect of strong mode coupling on MIMO equalization complexity. Using the equivalence of strongly coupled MD to a Gaussian unitary ensemble [19], we quantify the GD spread arising from MD for MMFs supporting various numbers of modes. We consider various MIMO equalizers using separate or combined equalization of CD and MD in the time or frequency domain. We study a model long-haul system using MMFs designed for low uncoupled GD spread, and evaluate the computational complexity of the various MIMO equalizers as a function of the number of modes and the correlation length. Quantifying complexity in terms of complex multiplications per two-dimensional symbol, we find that complexity increases with an increase in the number of propagating modes, and decreases with a decrease of the correlation length. Nevertheless, with sufficiently short correlation length and optimized MIMO equalizer structures (combined frequency-domain equalization of CD and MD), the complexity per symbol for MDM systems using $D = 6$, 12, 20, 30 modes (in two polarizations) can be only a factor of about 1.3 to 2.9 times higher than for polarization-multiplexed systems in standard SMF ($D = 2$).

The remainder of the paper is organized as follows. Section II reviews models for MMF with strong mode coupling and quantifies the GD spread arising from MD. Section III describes various MIMO single-carrier equalizer structures and their computational complexities. Section IV presents a long-haul system example, and evaluates the complexity of different MIMO equalizer structures as a function of the number of modes and the correlation length. Sections V and VI provide discussion and conclusions, respectively.

II. Temporal Memory of Modal Dispersion

A. Multi-Section Propagation Model

We consider a fiber supporting a total of D propagating modes (including spatial and polarization degrees of freedom). In a basis of the ideal uncoupled eigenmodes of an ideal fiber, a propagating field can be represented by a $D \times 1$ vector. Neglecting additive amplifier noise and nonlinear effects, the fields at the fiber input and output, at angular frequency ω, can be related by a linear time-invariant propagation operator (or $D \times D$ matrix):

$$\mathbf{M}(\omega) = \mathbf{M}_{CD}(\omega)\mathbf{M}_{MD}(\omega), \tag{1}$$

where $\mathbf{M}_{CD}(\omega)$ describes CD (or group velocity dispersion) operator and $\mathbf{M}_{MD}(\omega)$ describes MD operator.

For analytical convenience, without loss of generality, $\mathbf{M}_{CD}(\omega)$ can be written as:

$$\mathbf{M}_{CD}(\omega) = \exp\left(\frac{j\beta_{2,av}L_{tot}\omega^2}{2}\right)\mathbf{I}, \tag{2}$$

where $\beta_{2,av}$ is the mode-averaged CD constant, L_{tot} is the total fiber length, and $\mathbf{I}$ is a $D \times D$ identity matrix.

Assuming the strong-coupling regime, the fiber can be described in terms of K independent equal-length sections, where the section length is equal to or slightly greater than the correlation length. As the total fiber length far exceeds the correlation length, $K \gg 1$. The modal dispersion operator can be expressed as [19], [26]:

$$\mathbf{M}_{MD}(\omega) = \prod_{k=1}^{K} \mathbf{V}^{(k)}\mathbf{\Lambda}^{(k)}(\omega)\mathbf{U}^{(k)H}. \tag{3}$$

Here, $\mathbf{V}^{(k)}$ and $\mathbf{U}^{(k)}$ are frequency-independent $D \times D$ random unitary matrices representing random mode coupling, which are independently distributed over k in the strong-coupling regime, and ^{H} denotes Hermitian conjugate. The diagonal matrix

$$\mathbf{\Lambda}^{(k)}(\omega) = diag[e^{-j\omega\tau_1}, e^{-j\omega\tau_2} \ldots e^{-j\omega\tau_D}] \tag{4}$$

describes propagation of uncoupled modes, assuming mode-dependent gain/loss is negligible. The uncoupled GDs per section are $(\tau_1, \ldots, \tau_D)$, which are assumed to be identical in each section, with zero mean and variance σ_τ^2. The statistics of strongly coupled GDs are known to depend on the uncoupled GDs only through σ_τ^2 [19]. Mode-dependent CD, if present, can be incorporated in the $\mathbf{\Lambda}^{(k)}(\omega)$, and may alter the statistics of the coupled GDs [19]. In the numerical examples in Section IV, mode-dependent CD is found to be small.

B. Group Delay Spread Statistics

Given the modal dispersion operator, following [19], [28], we define a GD operator:

$$\mathbf{G} = j\frac{\partial \mathbf{M}_{MD}(\omega)}{\partial \omega}\mathbf{M}_{MD}(\omega)^H, \tag{5}$$

whose eigenvalues are the coupled GDs $(T_1, \ldots, T_D)$, where we assume the ordering $T_1 \leq T_2 \leq \cdots \leq T_D$. Assuming a random coupling model for the dispersion operator such as (3), the coupled GDs are random variables. As explained in [19], for $K \gg 1$, the GD operator is equivalent to a zero-trace Gaussian unitary ensemble, so the statistics of the coupled GDs are given by the statistics of the eigenvalues of the ensemble. In particular, for $D = 2$, the probability density function (pdf) of unordered coupled GDs is Maxwellian, as is known from the study of PMD [29], [30]. In the limit of many modes, $D \to \infty$, the pdf approaches a semicircle distribution.

For all values of D, the root-mean-square (rms) coupled GD spread after K independent sections is:

$$\sigma_{gd} = \sqrt{K}\sigma_\tau. \tag{6}$$

In a MIMO equalizer, the temporal memory required for compensating MD is the peak-to-peak coupled GD spread, $T_D - T_1$, for the model (3). Normalizing by the rms coupled GD spread, we define the random variable:

$$\mu_D = \frac{T_D - T_1}{\sigma_{gd}}, \tag{7}$$

which has pdf $f_{\mu_D}(x)$. For $D = 2$ and 3, analytical expressions for the pdf's $f_{\mu_2}(x)$ and $f_{\mu_3}(x)$ are given in [19]. As $D \to \infty$, $\mu_D \to 4$, since an infinite number of GD values are confined in

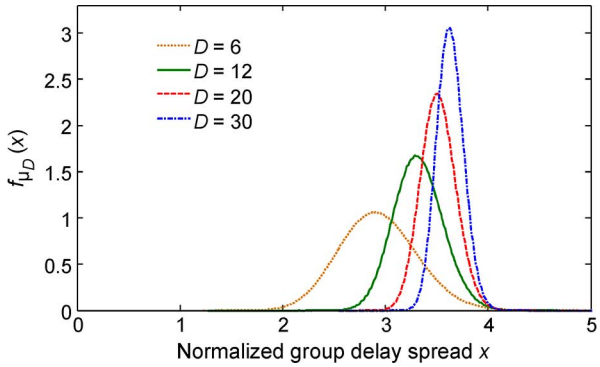

Fig. 1. Empirical probability density function of the normalized peak-to-peak group delay spread μ_D for $D = 6$, 12, 20 and 30 modes.

TABLE I
VALUES OF $u_D(p)$

	p=10^{-2}	p=10^{-3}	p=10^{-4}	p=10^{-5}	p=10^{-6}
D = 6	3.85	4.18	4.47	4.72	4.94
D = 12	3.91	4.12	4.30	4.45	4.59
D = 20	3.93	4.09	4.21	4.32	4.42
D = 30	3.95	4.07	4.16	4.25	4.32

the finite support of the semicircle distribution [19]. For intermediate values of D, the pdf $f_{\mu_D}(x)$ is not known analytically, so we have estimated $f_{\mu_D}(x)$ numerically using 10^7 realizations of the zero-trace Gaussian unitary ensemble, which are generated following the procedure described in [26]. Empirical pdfs for $D = 6$, 12, 20 and 30 are shown in Fig. 1. We have verified the accuracy of this technique by comparison to the known analytical expressions for unordered coupled GD distributions given in [19].

As the required temporal memory is a random variable, we will compute the temporal memory required to equalize MD with high probability. We define $u_D(p)$ as the temporal memory, normalized by the rms coupled GD spread, required to equalize with probability $1 - p$. This is defined by:

$$\int_0^{u_D(p)} f_{\mu_D}(x)dx = 1 - p. \tag{8}$$

Table I gives values of $u_D(p)$ for $D = 6$, 12, 20 and 30. For $p = 10^{-2}$, values of $u_D(p)$ are just below 4 for these values of D, while for the smaller values of p, values of $u_D(p)$ lie between 4 and 5 for these values of D. For a given p, the required temporal memory length in time units becomes:

$$\Delta T_{MD} = \sigma_{gd} u_D(p) = \sqrt{K}\sigma_\tau u_D(p). \tag{9}$$

C. Comparison of Group Delay Spread for Uncoupled and Strongly Coupled Propagation

In Section IV, exact GD values are computed for the fiber type used in numerical examples. Here, in order to illustrate the effect of strong mode coupling on the required temporal memory length, we consider a simple uncoupled GD distribution, as shown in Fig. 2. Uncoupled GDs form degenerate groups of $2, 4, 6, \ldots, \sqrt{4D+1} - 1$ modes, which add up to

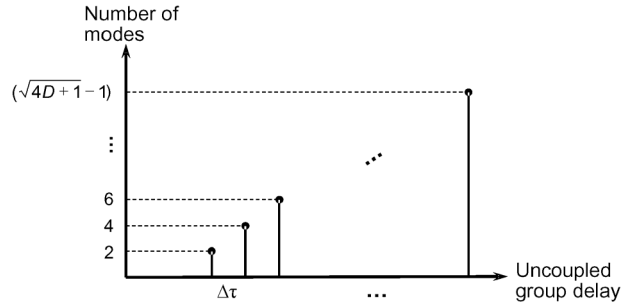

Fig. 2. Histogram of uncoupled GDs per section of a D-mode fiber, assuming a constant spacing, $\Delta\tau$, between mode groups.

$D = 2, 6, 12, 20, 30, \ldots$ modes in two polarizations, as in any circularly symmetric fiber in the weak-guidance limit [31]. Adjacent mode groups are assumed to differ in GD by a constant spacing $\Delta\tau$, as found approximately in a wave-optics analysis of step-index fibers [32], and observed experimentally in several different fiber types [17], [33], [34]. For the uncoupled GDs in Fig. 2, the uncoupled peak-to-peak GD spread per section is $(\sqrt{4D+1} - 3)\Delta\tau/2$, while the uncoupled rms GD spread per section is $\sigma_\tau = \sqrt{(D-2)/18}\Delta\tau$. Including K sections, in the absence of mode coupling, the required temporal memory is K times the uncoupled peak-to-peak GD spread per section, whereas with strong mode coupling, the required temporal memory is $\sqrt{K}u_D(p)$ times the uncoupled rms GD spread per section. Taking the ratio between these last two quantities, the reduction of required temporal memory length by strong mode coupling is approximately $\sqrt{K/5}$ for $D = 6$ and $p = 10^{-5}$, and approaches $\sqrt{9K/8}$ as $D \to \infty$ (independent of p).

This example illustrates that strong mode coupling and a decreased correlation length can substantially reduce the temporal memory required in the MIMO equalizer.

III. COMPUTATIONAL COMPLEXITY OF EQUALIZATION

A. Multi-Input Multi-Output Equalizers and Temporal Memory Lengths

In a coherent MDM receiver, after modal demultiplexing, a set of D dual-quadrature homodyne downconverters yields a set of D complex baseband signals. These are sampled at a rate $r_{os}R_s$, where R_s is the symbol rate and r_{os} is an oversampling ratio. In practice, a ratio $r_{os} = 2$ is usually employed to provide immunity to sampling time errors and increase CD tolerance [22]. Assuming single-carrier modulation, after sampling, DSP operations include timing recovery, interpolation, MIMO equalization, frequency and phase estimation, symbol decisions and error-correction decoding [6], [7], [36], [37]. We focus here on the computational complexity of MIMO equalization, rather than that of other DSP operations. Although the MIMO equalizer is an *adaptive filter*, we focus here on the complexity of the *filtering*, rather than that of the *adaptation*. The computational complexity and convergence time for adaptation are also critical design issues, and are important topics for future research.

Possible structures for the single-carrier MIMO equalizer are shown in Fig. 3. In Fig. 3(a), CD is compensated by a separate single-input single-output (SISO) equalizer for each of the D complex baseband signals, and MD and mode coupling crosstalk are compensated by a $D \times D$ MIMO equalizer. In

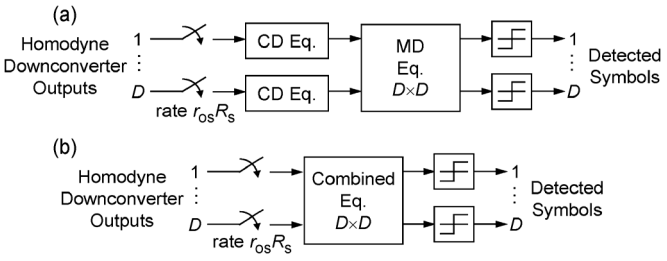

Fig. 3. Single-carrier MIMO equalizer structures: (a) separate equalizers for chromatic dispersion (CD) and modal dispersion (MD), (b) combined equalizer for CD and MD. All inputs and outputs are complex-valued.

each CD equalizer, the number of filter taps (temporal memory measured in sampling intervals) required for compensation with very low power penalty [37], [38] is given by:

$$N_{CD} = \lceil 2\pi |\beta_{2,av}| L_{tot} (r_{os} R_s)^2 \rceil \tag{10}$$

where $\lceil x \rceil$ denote smallest integer greater than or equal to x. In (10), the quadratic dependence on r_{os} arises because the GD spread of the discrete-time impulse response describing CD is proportional to the bandwidth of the signal being sampled, which scales in proportion to r_{os} in a well-designed system. The MD equalizer requires an array of $D \times D$ filters, each with temporal memory equal to ΔT_{MD}, given by (9). In terms of filter taps, each requires a temporal memory approximately given by:

$$N_{MD} = \lceil \Delta T_{MD} r_{os} R_s \rceil. \tag{11}$$

In Fig. 3(b), a $D \times D$ MIMO equalizer compensates for CD, MD and modal crosstalk with temporal memory, measured in filter taps given by $N_{CD} + N_{MD}$, the sum of (10) and (11). This follows from the observation that the convolution of two continuous-time impulse responses has a support equal to the sum of the individual supports.

In the following two subsections, we analyze the computational complexity of the equalizer structures shown in Fig. 3(a), (b). As in [39]–[41], we quantify complexity in terms of complex multiplications per two-dimensional data symbol, which is proportional to the complexity per information bit, assuming a fixed-size signal constellation.

B. Frequency-Domain Equalization Complexity

For long impulse response durations, digital filtering can be implemented most efficiently by frequency-domain equalization (FDE) using the Fast Fourier Transform (FFT) algorithm and either the overlap-save or the overlap-add method [42]. Alternatively, transmitting a cyclic prefix allows one to avoid overlap-save or overlap-add, but decreases net throughput [43].

In Fig. 3(a), in each of the D equalizers for CD, processing a block of N_{FFT} samples requires an FFT, frequency-domain filtering with N_{FFT} complex multiplications and an inverse FFT. Here, for simplicity, we restrict attention to radix-2 FFT algorithms, although FFTs of different radices can be computed efficiently [44], [45], albeit at the cost of an increased number of additions and a more complex implementation. In processing each block, these operations equalize $(N_{FFT} - N_{CD} + 1)/r_{os}$ symbols, so the number of complex multiplications per symbol is [42], [46]:

$$CM_{CD,FDE} = r_{os} \frac{N_{FFT} \log_2(N_{FFT}) + N_{FFT}}{(N_{FFT} - N_{CD} + 1)}, \tag{12}$$

which is independent of D. For equalization of MD, since the equalizer size is $D \times D$, the number of complex multiplications per symbol is:

$$CM_{MD,FDE} = r_{os} \frac{N_{FFT} \log_2(N_{FFT}) D + N_{FFT} D^2}{(N_{FFT} - N_{MD} + 1) D}. \tag{13}$$

In Fig. 3(b), the combined equalizer for CD and MD is analogous to the MD equalizer in Fig. 3(a), but with temporal memory increased and the number of complex multiplications per symbol is:

$$CM_{comb,FDE} = r_{os} \frac{N_{FFT} \log_2(N_{FFT}) D + N_{FFT} D^2}{(N_{FFT} - N_{MD} - N_{CD} + 1) D}. \tag{14}$$

In each frequency-domain equalizer, the FFT block length N_{FFT} can be any integer of power of 2 that is longer than the corresponding temporal memory length (N_{CD}, N_{MD} or $N_{CD} + N_{MD}$) and can be optimized to minimize computational complexity given by (12), (13) or (14). It is shown that when using an optimized FFT block length, the computational complexity is relatively insensitive to the temporal memory length. Unfortunately, as discussed in Section V, using long FFT block lengths can increase circuit complexity and processing latency.

C. Time-Domain Equalization Complexity

Digital filtering can also be implemented using direct time-domain convolution.

In Fig. 3(a), in each of the D equalizers for CD, processing a block of N_{CD} samples requires N_{CD}^2 complex multiplications and equalizes N_{CD}/r_{os} symbols, so the number of complex multiplications per symbol is:

$$CM_{CD,TDE} = r_{os} N_{CD}, \tag{15}$$

which is independent of D. In Fig. 3(a), in the $D \times D$ equalizer for MD, processing D blocks of N_{MD} samples requires $N_{MD}^2 D^2$ complex multiplications and equalizes $N_{MD} D/r_{os}$ symbols, so the number of complex multiplications per symbol for MD compensation is:

$$CM_{MD,TDE} = r_{os} N_{MD} D, \tag{16}$$

which scales linearly with D, unlike (15).

In Fig. 3(b), the combined equalizer for CD and MD is analogous to the MD equalizer in Fig. 3(a), but with temporal memory increased, so the number of complex multiplications per symbol is:

$$CM_{comb,TDE} = r_{os} (N_{MD} + N_{CD}) D, \tag{17}$$

which scales linearly with D, like (16).

In time-domain equalization (TDE), increasing the temporal memory length causes the computational complexity (15), (16) or (17) to increase in direct proportion to the memory length.

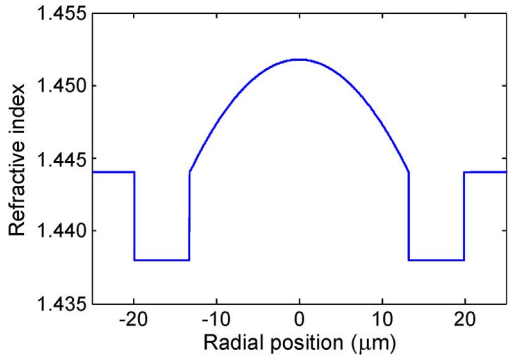

Fig. 4. Index profile of optimized graded-index depressed cladding (GIDC) fiber for $NA = 0.15$ and $D = 12$ modes.

TABLE II
UNCOUPLED FIBER PARAMETERS. VALUES FOR D = 2 ARE FOR STANDARD STEP-INDEX SINGLE-MODE FIBER. VALUES FOR D > 2 ARE FOR GRADED-INDEX DEPRESSED-CLADDING MULTI-MODE FIBERS WITH NA = 0.15

	$D = 2$	$D = 6$	$D = 12$	$D = 20$	$D = 30$
Core radius (μm)	4.1	10.0	13.2	16.6	19.8
Effective area A_{eff} (μm^2)	63.6	166	271	419	580
Mode-averaged chromatic dispersion $\beta_{2,av}$ (ps^2/km)	–22.5	–28.0	–28.4	–28.6	–28.7
Mode-dependent chromatic dispersion, rms $\Delta\beta_{2,rms}$ (ps^2/km)	–	1.3	1.9	2.3	2.7
Modal dispersion per unit length, rms $\Delta\tau_{L,rms}$ (ps/km)	0	277	383	415	451

IV. LONG-HAUL SYSTEM EXAMPLE

In this section, we describe a family of fibers and their uncoupled GDs. We then consider a long-haul system with mode coupling and study how the fiber GD spread and the equalizer memory length and computational complexity scale with the number of modes and the correlation length.

A. Graded-Index Depressed-Cladding Fibers

We consider a family of fibers supporting $D = 6$, 12, 20, 30 modes in two polarizations. We have chosen a graded-index depressed-cladding (GIDC) index profile and numerical aperture $NA = 0.15$, as these enable low GD spreads over a wide range of D and, in conjunction with optimized erbium doping profiles, enable optical amplification with low MDL [47]. At a given D, the core radius has been chosen so that over the C band, exactly D modes propagate and, at the shortest wavelength of 1530 nm, the next higher-order mode is just barely cut off. An index profile for $D = 12$ is shown in Fig. 4. We have used numerical methods to solve for the exact modes [47], [48], without assuming weak guidance or linearly polarized modes.

Table II summarizes key properties of the GIDC fibers. As D varies from 6 to 30, the effective area A_{eff} ranges from 166 to 580 μm^2, in proportion to $D^{0.78}$. For all values of D from 6 to 30, the mode-averaged CD $\beta_{2,av}$ is close to -28.4 ps^2/km, and the rms mode-dependent CD $\Delta\beta_{2,rms}$ is small, in the range of 5–10% of $\beta_{2,av}$, so mode-dependent CD is neglected here. The uncoupled rms modal GD spread per unit length $\Delta\tau_{L,rms}$ varies from 277 to 451 ps/km, in proportion to $D^{0.30}$ from $D = 6$ to 30, a much weaker dependence on D than for the model in Section II-A, where it scales with $\sqrt{D-2}$. It might be possible to further optimize these index profiles to reduce the GD spread and its dependence on D, which is an important topic for future research.

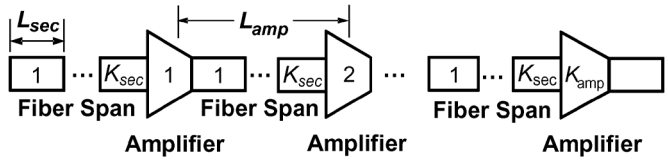

Fig. 5. Multi-section model of long-haul system with inline amplifiers.

Table II also summarizes key properties of standard step-index SMF [49] for comparison. We neglect here the GD spread due to PMD, which is much smaller than that due to MD [35], [50]–[52].

B. Multi-Section System Model

We consider the model long-haul system shown in Fig. 5, comprising $K_{amp} = 20$ spans separated by optical amplifiers, each of length $L_{amp} = 100$ km, a total system length $L_{tot} = 2000$ km. Each span is modeled as comprising K_{sec} sections with independent modal fields, so each section has length $L_{sec} = L_{amp}/K_{sec}$. The total system comprises $K = K_{amp}K_{sec} = L_{tot}/L_{sec}$ independent sections.

The section length L_{sec} is assumed just slightly longer than the correlation length, so each individual section is in the weak-coupling regime, and the uncoupled rms GD spread per section is $\sigma_\tau = L_{sec}\Delta\tau_{L,rms}$. Using (6), the end-to-end system has a coupled rms GD spread:

$$\sigma_{gd} = \sqrt{L_{sec}L_{tot}}\Delta\tau_{L,rms} = \sqrt{\frac{L_{amp}L_{tot}}{K_{sec}}}\Delta\tau_{L,rms}. \quad (18)$$

By holding L_{amp} and L_{tot} constant and varying K_{sec} or L_{sec}, the coupled GD spread can be varied. Given a fiber with known uncoupled rms GD spread per unit length $\Delta\tau_{L,rms}$ in a system with known coupled rms GD spread σ_{gd}, the section length L_{sec} can be chosen empirically so that (18) yields the observed value of σ_{gd}. Computation of the correlation length and the section length from fundamental fiber properties, such as modal field distributions and propagation constants, and perturbations inducing mode coupling, requires further study.

C. Memory Length and Computational Complexity

We consider the family of GIDC fibers supporting $D = 6$, 12, 20, 30 modes described in Section IV-A in the multi-span transmission system described in Section IV-B. We assume single-carrier modulation at a symbol rate $R_s = 32$ Gbaud and an oversampling ratio $r_{os} = 2$. In computing $u_D(p)$ and the temporal memory length of MD, we conservatively assume $p = 10^{-5}$. All results are shown as a function of K_{sec} and L_{sec}.

Fig. 6 shows the temporal memory lengths for CD and MD, N_{CD} and N_{MD} for various D. The CD memory length, N_{CD}, is equal to 1155 samples for SMF ($D = 2$) and 1476 samples for MMF ($D = 6$, 12, 20, 30). The MD memory length in MMF, N_{MD} increases by only about 1.5-fold as D increases from 6 to

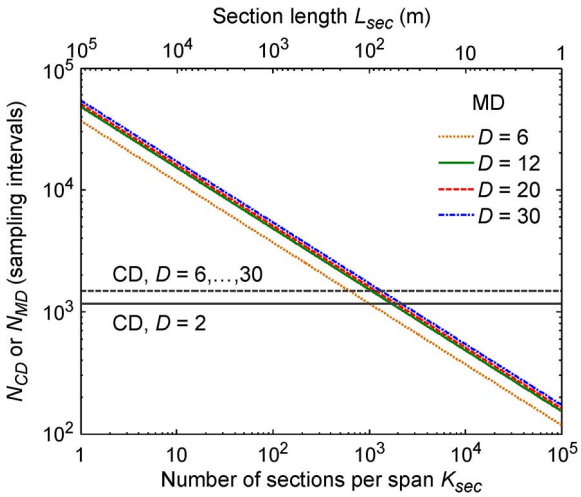

Fig. 6. Memory length (in sampling intervals) required to compensate CD or MD, N_{CD} or N_{MD}, vs. number of sections per span, K_{sec}, in systems with 20 spans, each 100 km long, of GIDC fiber.

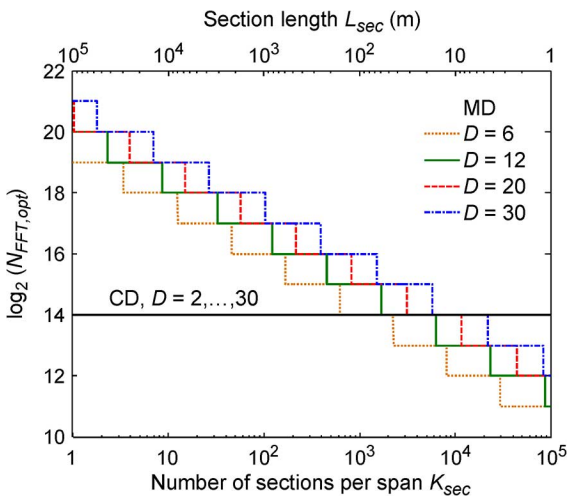

Fig. 7. Optimized FFT block length, $N_{FFT,opt}$, for separate frequency-domain equalization (FDE) of CD and MD vs. number of sections per span, K_{sec}, in systems with 20 spans, each 100 km long, of GIDC fiber.

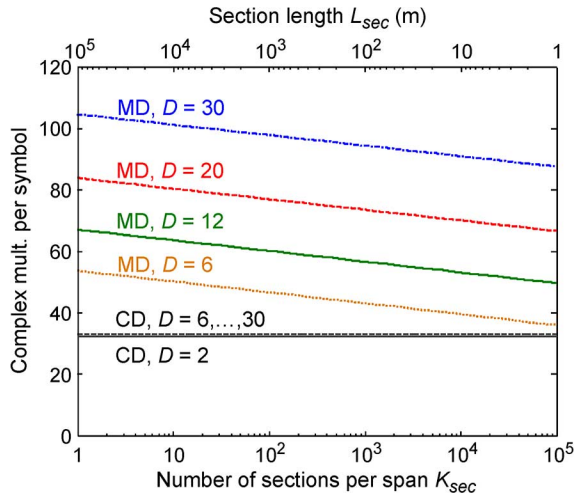

Fig. 8. Complex multiplications per two-dimensional symbol for separate FDE of CD and MD, $CM_{CD,FDE}$ and $CM_{MD,FDE}$, vs. number of sections per span, K_{sec}, in systems with 20 spans, each 100 km long, of GIDC fiber.

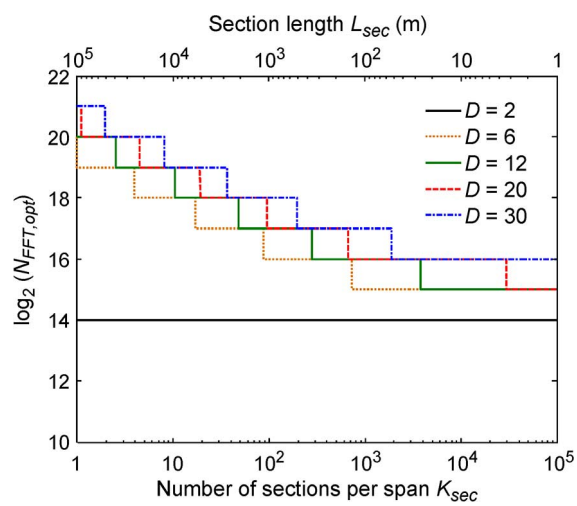

Fig. 9. Optimized FFT block length, $N_{\mathrm{FFT,opt}}$, for combined FDE of CD and MD vs. number of sections per span, K_{sec}, in systems with 20 spans, each 100 km long, of GIDC fiber.

30, because a 1.6-fold increase in $\Delta\tau_{L,rms}$ is partly offset by a 0.9-fold decrease in $u_D(p)$ Most significantly, for all values of D, as K_{sec} and L_{sec} vary, N_{MD} decreases from about 5×10^4 to about 150, demonstrating the dramatic impact of mode coupling on temporal memory length.

Figs. 7 and 8 concern separate FDE of CD and MD. Fig. 7 shows values of the FFT block length optimized to minimize computational complexity, $N_{FFT,opt}$. For equalizing CD, in SMF and all MMF, $N_{FFT,opt} = 2^{14}$, which is equal to about 14 or 11 times the respective values of N_{CD}. For equalizing MD in MMF, $N_{FFT,opt}$ is about 20 to 30 times the respective value of N_{MD}. Fig. 8 shows the computational complexities for separate FDE of CD and MD. The complexity for CD, $CM_{CD,FDE}$, is almost identical for SMF or MMF, about 32–33 complex multiplications/symbol. The complexity for MD in MMF, $CM_{MD,FDE}$, increases modestly with increasing D and decreases modestly with increasing K_{sec}.

Figs. 9 and 10 concern combined FDE of CD and MD. Fig. 9 shows values of the optimized FFT block length, $N_{FFT,opt}$. For

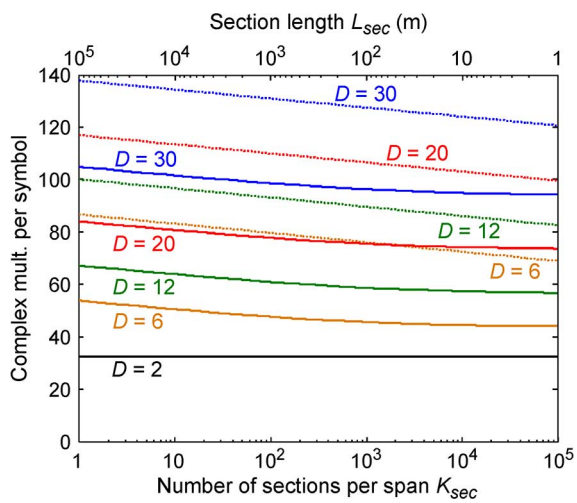

Fig. 10. Complex multiplications per two-dimensional symbol for FDE vs. number of sections per span K_{sec}, in systems with 20 spans, each 100 km long, of GIDC fiber. For $D = 6,\ldots,30$, solid lines denote multiplications for combined equalizer, $CM_{comb,FDE}$, while dotted lines denote sum of multiplications for separate CD and MD equalizers, $CM_{CD,FDE} + CM_{MD,FDE}$.

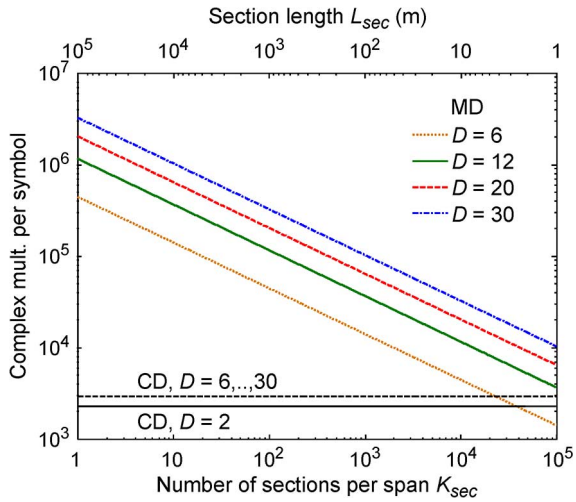

Fig. 11. Complex multiplications per two-dimensional symbol for separate time-domain equalization (TDE) of CD and MD, $CM_{CD,TDE}$ and $CM_{MD,TDE}$, vs. number of sections per span, K_{sec}, in systems with 20 spans, each 100 km long, of GIDC fiber.

SMF, independent of K_{sec}, $N_{FFT,opt} = 2^{14}$, which is about 14 times N_{CD}. For MMF, $N_{FFT,opt}$ is about 20 to 40 times the respective value of $N_{CD} + N_{MD}$.

Fig. 10 shows (using solid lines) the computational complexities for combined FDE of CD and MD, $CM_{comb,FDE}$. In MMF, $CM_{comb,FDE}$ increases modestly with increasing D and decreases modestly with increasing K_{sec}. As K_{sec} increases, for $D = 6$, $CM_{comb,FDE}$ ranges from about 1.6 to 1.3 times the value for $D = 2$, while for $D = 30$, $CM_{comb,FDE}$ ranges from about 3.2 to 2.9 times the value for $D = 2$. Fig. 10 also shows (using dotted lines) the overall computational complexity for separate FDE of CD and MD in MMF, which is the sum of the complexities, $CM_{CD,FDE} + CM_{MD,FDE}$. In terms of its dependence on D or K_{sec}, this exhibits the same general trends as $CM_{comb,FDE}$. But at any K_{sec}, $CM_{CD,FDE} + CM_{MD,FDE}$ is *higher* than $CM_{comb,FDE}$, by a factor of about 1.6 for $D = 6$, and by a factor of about 1.3 for $D = 30$.

Fig. 11 shows the computational complexity for separate TDE of CD and MD. The complexity for CD, $CM_{CD,TDE}$, is about 2300 and 3000 complex multiplications per symbol for SMF and MMF, respectively, which are about 72 to 90 times higher than for FDE. In MMF, the complexity for MD, $CM_{MD,TDE}$, increases by about 7.3 as D increases from 6, to 30, and decreases by nearly 300 as K_{sec} increases from 1 to 10^5.

In MMF, the complexity of separate TDE is always *lower* than that of combined TDE $(CM_{CD,TDE} + CM_{MD,TDE})$, in contrast to FDE, which can be traced to the factor of D that appears in (16) and (17), but not in (15). For small K_{sec}, $N_{MD} \gg N_{CD}$, and $CM_{CD,TDE} + CM_{MD,TDE} \approx CM_{MD,TDE}$. For large K_{sec}, N_{MD} is comparable to N_{CD}, and $CM_{CD,TDE} + CM_{MD,TDE}$ is modestly larger than $CM_{MD,TDE}$. In the limit of very large K_{sec}, $N_{MD} \ll N_{CD}$, and $CM_{CD,TDE} + CM_{MD,TDE} \rightarrow CM_{CD,TDE}$. However, the overall complexity of TDE, whether in combined or separate form, is so high to render it practical, even with strong mode coupling.

V. Discussion

Our results clearly demonstrate the large difference in computational complexity between TDE and FDE in coherent MDM systems, as shown in [52] under different assumptions about propagation. For MDM using $D = 6$, 12, 20, or 30 modes, the complexity of combined FDE, assuming an optimized FFT block length, can be comparable to that for polarization-multiplexed transmission in SMF $(D = 2)$. The complexity depends mainly on D, and depends only weakly on the strength of mode coupling, as quantified by the section length L_{sec}.

In FDE, the optimum FFT block length $N_{FFT,opt}$ is proportional to the corresponding channel memory length, so $N_{FFT,opt}$ depends strongly on the section length L_{sec}. This has major implications for the MIMO FDE, both in terms of DSP circuit complexity and adaptation speed. As a rough measure of overall DSP circuit complexity, we will consider the number of complex multiplications per D blocks of N_{FFT} samples. For SMF $(D = 2)$, the complexity is dominated by CD equalization, and FDE requires $N_{FFT}D[\log_2(N_{FFT}) + 1]$ complex multiplication per DN_{FFT} samples. Assuming $N_{FFT,opt} = 2^{14}$, this corresponds to 4.9×10^5 complex multiplications. For MMF, the separate FDE for MD or the combined FDE for CD+MD requires $N_{FFT}D[\log_2(N_{FFT}) + D]$ complex multiplication per DN_{FFT} samples. With strong mode coupling, such that $L_{sec} = 100$ m, $N_{MD} \sim N_{CD} \sim 1000$. Assuming $N_{FFT,opt} = 2^{16}$, $N_{FFT}D[\log_2(N_{FFT}) + D] = 8.7 \times 10^6$ and 9.0×10^7 for $D = 6$ and 30, respectively. These values are 18 and 184 times the value for $D = 2$, so it may be possible to realize a single-chip MIMO FDE. On the other hand, with weaker mode coupling, such that $L_{sec} = L_{amp} = 10^5$ m, $N_{MD} \sim 5 \times 10^4$. Assuming $N_{FFT,opt} = 2^{21}$, $N_{FFT}D[\log_2(N_{FFT}) + D] = 3.4 \times 10^8$ and 3.2×10^9 for $D = 6$ and 30, respectively. These values are 690 and 6500 times the value for $D = 2$, so it likely would be difficult to realize a single-chip MIMO FDE.

One of the drawbacks of frequency-domain equalization is the block processing latency, which also affects the adaptation time. An FFT block length N_{FFT} implies a block processing latency of $N_{FFT}/(r_{os}R_s)$ seconds. Assuming $r_{os} = 2$ and $R_s = 32$ Gbaud, with strong mode coupling, such that $L_{sec} = 100$ m and $N_{FFT,opt} = 2^{16}$, the latency for equalizing one block is 1 μs. With weaker mode coupling, such that $L_{sec} = 10^5$ m and $N_{FFT,opt} = 2^{21}$, the latency is 33 μs.

Adaptation algorithms are not analyzed in detail in this paper. Exploiting the fact that CD exhibits minimal variation over time, one may choose to employ fixed FDE for CD and adaptive FDE for MD. While not minimizing the computational complexity of equalization, this approach may minimize the computational complexity of adaptation and facilitate faster adaptation. Two major approaches for data-aided adaptive equalization (initially with training sequences, then either with training sequences or decision-directed) are MIMO least mean squares (MIMO-LMS) and MIMO recursive least squares (MIMO-RLS) algorithms [53]. As compared to MIMO-LMS, MIMO-RLS provides faster convergence rates, but has higher complexity, as it requires matrix inversions [54]. The number of adaptation steps required for MIMO-RLS to converge is roughly of the order of D^2 [54]. Since each adaptation step requires equalization of D blocks

of N_{FFT} samples, the convergence time for MIMO-RLS is roughly $D^2 N_{FFT}/(r_{os} R_s)$. Using the latencies estimated in the previous paragraph, for strong mode coupling ($L_{sec} = 100$ m, $N_{FFT,opt} = 2^{16}$) the convergence times are roughly 36 μs and 900 μs for $D = 6$ and 30, respectively. For weaker mode coupling ($L_{sec} = 10^5$ m, $N_{FFT,opt} = 2^{21}$) the convergence times are roughly 1.2 ms and 30 ms for $D = 6$ and 30, respectively. Convergence times for MIMO-LMS are expected to be at least an order of magnitude longer than these. Clearly, mode coupling strength can have a strong impact on adaptive FDE convergence time. Measurements of modal field temporal variations in deployed long-haul MMFs may be needed to determine adaptive FDE convergence time requirements.

Although only single-carrier transmission is considered in this paper, orthogonal frequency division multiplexing (OFDM) might also be applicable in MDM systems. The computational complexity of OFDM is identical to that of the FDE considered here. In OFDM, the optimized block length $N_{FFT,opt}$ corresponds to the number of orthogonal carriers (including carriers used for data transmission and zero carriers used for oversampling).

The impact of MDL is not considered in detail in this paper. MDL can reduce average capacity and increase the variance of capacity, potentially causing outage [26]. Strong mode coupling, in combination with MD, enables frequency diversity, which can reduce the variance of capacity and thus reduce outage probability [27]. The capacity variance is reduced by a frequency diversity order [27] $F_D \sim R_s \sigma_{gd} = N_{MD}/[r_{os} u_D(p)]$. Introducing sufficient mode coupling such that $N_{MD} \sim N_{CD} \sim 1000$ corresponds to a frequency diversity order $F_D \sim 100$ for the parameters considered in this paper, which should be sufficient to make outage probability negligible, assuming low-to-moderate MDL values [26], [27]. Aside from reducing average capacity and potentially causing outage, MDL also affects the statistics of the coupled GDs. Our simulations have shown that MDL typically narrows the GD spread. Hence, values of the temporal memory N_{MD} and the computational complexity presented here should be considered upper bounds for systems with MDL.

Methods to achieve strong mode coupling without introducing substantial MDL, whether by using mode couplers or scramblers, or by intentional perturbation of the MMF, are important subjects for future research.

VI. Conclusion

In MMF, the temporal memory length of MD scales with the square-root of a section length over which modal fields are correlated, and is thus reduced significantly by strong mode coupling. In coherent MDM systems, considering realistic correlation lengths, the computational complexity of MIMO TDE is extremely high. By contrast, using MIMO FDE and an optimized FFT block length, the computational complexity (measured in complex multiplications per two-dimensional symbol) can be only 1.3 to 2.9 times higher than for FDE in SMF, and depends only weakly on the temporal memory length. Nevertheless, the optimized FFT block length is proportional to the temporal memory length of MD. Hence, introducing strong mode coupling to reduce the temporal memory of MD can greatly reduce signal processing circuit complexity, and can greatly reduce block processing latency, thereby reducing MIMO FDE convergence time.

Acknowledgment

The authors acknowledge useful discussions with K.-P. Ho.

References

[1] 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.

[2] T. Morioka, Y. Awaji, R. Ryf, P. J. Winzer, D. Richardson, and F. Poletti, "Enhancing optical communications with brand new fibers," *IEEE Commun. Mag.*, vol. 50, no. 2, pp. s31–s42, Feb. 2012.

[3] P. J. Winzer and G. J. Foschini, "MIMO capacities and outage probabilities in spatially multiplexed optical transport systems," *Opt. Exp.*, vol. 19, no. 17, pp. 16680–16696, Aug. 2011.

[4] R. Essiambre and R. W. Tkach, "Capacity trends and limits of optical communication networks," *Proc. IEEE*, vol. 100, no. 5, pp. 1035–1055, May 2012.

[5] R. E. Freund, C.-A. Bunge, N. N. Ledentsov, D. Molin, and C. Casper, "High speed transmission in multi-mode fibers," in *Proc. Opt. Fiber Commun. Conf*, San Diego, CA, Mar. 2009, Paper OMS1.

[6] E. Ip, A. P. T. Lau, D. J. F. Barros, and J. M. Kahn, "Coherent detection in optical fiber systems," *Opt. Exp.*, vol. 16, no. 2, pp. 753–791, Jan. 2008.

[7] E. Ip and J. M. Kahn, "Fiber impairment compensation using coherent detection and digital signal processing," *J. Lightw. Technol.*, vol. 28, no. 4, pp. 502–519, Feb. 2010.

[8] P. J. Winzer, "Energy-efficient optical transport capacity scaling through spatial multiplexing," *IEEE Photon. Technol. Lett.*, vol. 23, no. 13, pp. 851–853, Jul. 2011.

[9] M. Salsi *et al.*, "Mode-division multiplexing of 2 100 Gb/s channels using an LCOS-based spatial modulator," *J. Lightw. Technol.*, vol. 30, no. 4, pp. 618–623, Feb. 2012.

[10] C. Koebele *et al.*, "40 km transmission of five mode division multiplexed data streams at 100 Gb/s with low MIMO-DSP complexity," presented at the Proc. Eur. Conf. Opt. Commun., Geneva, Switzerland, Sep. 2011, Paper Th.13.C.3, unpublished.

[11] M. Salsi *et al.*, "Mode division multiplexed transmission with a weakly coupled few mode fiber," presented at the Opt. Fiber Comm. Conf., Los Angeles, CA, Mar. 2012, Paper OTu2C.5.

[12] S. Randel *et al.*, "Mode-multiplexed 6 × 20-GBd QPSK transmission over 1200-km DGD-compensated few mode fiber," presented at the Opt. Fiber Comm. Conf., Los Angeles, CA, Mar. 2012, Paper PDP5C.5.

[13] R. Ryf *et al.*, "Mode-division multiplexing over 96 km of few-mode fiber using coherent 6 × 6 MIMO processing," *J. Lightw. Technol.*, vol. 30, no. 4, pp. 521–531, Feb. 2012.

[14] N. Bai *et al.*, "Mode-division multiplexed transmission with inline few-mode fiber amplifier," *Opt. Exp.*, vol. 20, no. 3, pp. 2668–2680, Jan. 2012.

[15] L. Nielsen *et al.*, "Few mode transmission fiber with low DGD, low mode coupling and low loss," presented at the Opt. Fiber Comm. Conf., Los Angeles, CA, Mar. 2012, Paper PDP5A.1.

[16] D. Đonlagic, "Opportunities to enhance multimode fiber links by application of overfilled launch," *J. Lightw. Technol.*, vol. 23, no. 11, pp. 3526–3540, Nov. 2005.

[17] R. E. Freund, C.-A. Bunge, N. N. Ledentsov, D. Molin, and C. R. E. Caspar, "High-speed transmission in multimode fibers," *J. Lightw. Technol.*, vol. 28, no. 4, pp. 569–586, Feb. 2010.

[18] 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 15, 2009.

[19] K.-P. Ho and J. M. Kahn, "Statistics of group delays in multimode fiber with strong mode coupling," *J. Lightw. Technol.*, vol. 29, no. 21, pp. 3119–3128, Nov. 2011.

[20] M. Ikeda, Y. Murakami, and K. Kitayama, "Mode scrambler in optical fibers," *Appl. Opt.*, vol. 16, pp. 1045–1049, 1977.

[21] K. S. Lee and T. Erdogan, "Fiber mode conversion with tilted gratings in optical fiber," *J. Opt. Soc. Amer. A*, vol. 18, pp. 1176–1185, May 2001.

[22] D. A. Nolan and M. J. Li, "Fiber spin-profile designs for producing fibers with low polarization mode dispersion," *Opt. Lett.*, vol. 23, no. 21, pp. 1659–1661, Nov. 1998.
[23] D. Marcuse, *Theory of Dielectric Optical Waveguides*, 2nd ed. New York: Academic, 1991.
[24] C. D. Poole and J. 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.
[25] F. Corsi, A. Galtarossa, and L. Palmieri, "Analytical treatment of polarization dispersion in single-mode fibers by means of backscattering signal," *J. Opt. Soc. Amer. A*, vol. 16, pp. 574–583, 1999.
[26] K.-P. Ho and J. M. Kahn, "Mode-dependent loss and gain: Statistics and effect on mode-division multiplexing," *Opt. Exp.*, vol. 19, no. 17, pp. 16612–16635, Aug. 2011.
[27] K.-P. Ho and J. M. Kahn, "Frequency diversity in mode-division multiplexing systems," *J. Lightw. Technol.*, vol. 29, no. 24, pp. 3719–3726, Dec. 2011.
[28] S. Fan and J. M. Kahn, "Principal modes in multi-mode waveguides," *Opt. Lett.*, vol. 30, no. 2, pp. 135–137, Jan. 2005.
[29] G. J. Foschini and C. D. Poole, "Statistical theory of polarization dispersion in single-mode fibers," *J. Lightw. Technol.*, vol. 9, pp. 1439–1456, 1991.
[30] 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.
[31] A. W. Snyder and J. D. Love, *Optical Waveguide Theory*. New York: Chapman and Hall, 1983.
[32] D. Gloge and E. J. Marcatili, "Multimode theory of graded core fibers," *Bell Syst. Tech. J.*, vol. 52, pp. 1563–1578, Mar. 1973.
[33] K. Takahashi, T. Ishigure, and Y. Koike, "Index profile design for high-bandwidth W-shaped plastic optical fiber," *J. Lightw. Technol.*, vol. 24, no. 7, pp. 2867–2876, July 2006.
[34] T. Mori *et al.*, "Wideband WDM coherent optical MIMO transmission over GI-MMF by using selective mode excitation," presented at the Opt. Fiber Comm. Conf., Los Angeles, CA, Mar. 2012, Paper OTu2C.3.
[35] E. Ip and J. M. Kahn, "Digital equalization of chromatic dispersion and polarization mode dispersion," *J. Lightw. Technol.*, vol. 25, no. 8, pp. 2033–2043, Aug. 2007.
[36] M. Kuschnerov, F. N. Hauske, K. Piyawanno, B. Spinnler, M. S. Alfiad, A. Napoli, and B. Lankl, "DSP for coherent single-carrier receivers," *J. Lightw. Technol.*, vol. 27, pp. 3614–3622, Aug. 2009.
[37] S. J. Savory, "Digital coherent optical receivers: Algorithms and subsystems," *IEEE Sel. Topics Quantum. Electron.*, vol. 16, no. 5, pp. 1164–1179, Oct. 2010.
[38] S. J. Savory, "Digital filters for coherent optical receivers," *Opt. Exp.*, vol. 16, no. 2, pp. 804–817, Jan. 2008.
[39] B. Spinnler, "Equalizer design and complexity for digital coherent receivers," *IEEE J. Quantum Electron.*, vol. 16, no. 5, pp. 1180–1192, Oct. 2010.
[40] B. Spinnler, "Complexity of algorithms for digital coherent receivers," presented at the ECOC, Vienna, Austria, 2009, Paper 7.3.6.
[41] M. S. Faruk and K. Kikuchi, "Adaptive frequency-domain equalization in digital coherent optical receivers," *Opt. Exp.*, vol. 19, no. 13, pp. 12789–12798, Jun. 2011.
[42] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*, 2nd ed. Englewood Cliffs, NJ: Prentice Hall, 1999.
[43] D. Falconer, S. L. Ariyavisitakul, A. Benyamin-Seeyar, and B. Eidson, "Frequency domain equalization for single-carrier broadband wireless systems," *IEEE Commun. Mag.*, vol. 40, no. 4, pp. 58–66, Apr. 2002.
[44] I. J. Good, "The interaction algorithm and practical Fourier analysis," *J. Royal Stat. Soc. B*, vol. 20, pp. 361–372, 1958.
[45] S. Winograd, "On computing the discrete Fourier transform," *Math. Comput.*, vol. 32, no. 141, pp. 175–199, 1978.
[46] B. Inan *et al.*, "DSP complexity of mode-division multiplexed receivers," *Opt. Exp.*, vol. 20, no. 9, pp. 10859–10869, Apr. 2011.
[47] D. Askarov and J. M. Kahn, "Design of transmission fibers and doped fiber amplifiers for mode-division multiplexing," *IEEE Photon. Technol. Lett.*, vol. 24, no. 21, pp. 1945–1948, Nov. 2012.
[48] S. R. A. Dods, "Fiber vector modesolver—Improvements to the efficient 4 × 4 matrix method," in *Proc. Integrated Photonics Research and Applications, IPRA 2006*, Uncasville, CT, Apr. 2006.
[49] Corning, *SMF-28 Optical Fiber* Product Information Sheet., Dec. 2001.
[50] C. R. Fludger *et al.*, "Coherent equalization and POLMUX-RZ-DQPSK for robust 100-GE transmission," *J. Lightw. Technol,*, vol. 26, no. 1, pp. 64–72, Jan. 2008.
[51] N. Mantzoukis *et al.*, "Electronic equalization of polarization mode dispersion in Coherent POL-MUX QPSK systems," presented at the Eur. Conf. Opt. Commun., Vienna, Austria, Sep. 2009, Paper P4.15.
[52] P. Serena, N. Rossi, O. Bertran-Pardo, J. Renaudier, A. Vannucci, and A. Bononi, "Intra-versus inter-channel PMD in linearly compensated coherent PDM-PSK nonlinear transmissions," *J. Lightw. Technol*, vol. 29, no. 11, pp. 1691–1700, Apr. 2011.
[53] J. Proakis, *Digital Communications*, 4th ed. New York: McGraw-Hill.
[54] J. M. Wang and B. Daneshrad, "A comparative study of MIMO detection algorithms for wideband spatial multiplexing systems," in *Proc. IEEE Conf. Wireless Commun. Netw.*, Mar. 2005, vol. 1, pp. 408–413.

Sercan Ö. Arık was born in Turkey in 1990. He received the B.S. degree in electrical and electronics engineering, from Bilkent University, Ankara, Turkey, in 2011. He is currently working toward the M.S. and Ph.D. degrees in electrical engineering at Stanford University.

He has authored papers on discrete processing of optical systems, fundamental information optics, image analysis and classification. In summer 2010, he worked at EPFL Signal Processing Labs, Lausanne, Switzerland and in summer 2012, he worked at Google, Mountain View, CA. His current research interests include high-speed long-haul transmission in multi-mode systems, MIMO digital signal processing algorithms for coherent multi-mode transmission, mode coupling modeling and enhancement, and information theoretical limits of multi-mode systems.

Daulet Askarov received the B.S. degree in applied mathematics and physics from Moscow Institute of Physics and Technology (MIPT) in 2009, and the M.S. degree in electrical engineering from Stanford University in 2011, where he is currently working toward the Ph.D. degree in electrical engineering.

His current research interests include various topics in Multi-Mode Fiber Communications.

Joseph M. Kahn (M'90–SM'98–F'00) received the A.B., M.A., and Ph.D. degrees in Physics from U.C. Berkeley in 1981, 1983, and 1986, respectively.

From 1987–1990, he was at AT&T Bell Laboratories, Crawford Hill Laboratory, Holmdel, NJ. He demonstrated multi-Gbit/s coherent optical fiber transmission systems, setting world records for receiver sensitivity. From 1990–2003, he was on the faculty of the Department of Electrical Engineering and Computer Sciences at UC Berkeley, performing research on optical and wireless communications. Since 2003, he has been a Professor of Electrical Engineering at Stanford University, where he heads the Optical Communications Group. His current research interests include: fiber-based imaging, spatial multiplexing, rate-adaptive and spectrally efficient modulation and coding methods, coherent detection and associated digital signal processing algorithms, digital compensation of fiber nonlinearity, and free-space systems. In 2000, he helped found StrataLight Communications, where he served as Chief Scientist from 2000–2003. StrataLight was acquired by Opnext, Inc. in 2009.

Prof. Kahn received the National Science Foundation Presidential Young Investigator Award in 1991. From 1993–2000, he served as a Technical Editor of IEEE *Personal Communications Magazine*. Since 2009, he has been an Associate Editor of IEEE/OSA JOURNAL OF OPTICAL COMMUNICATIONS AND NETWORKING.