Efficient Quantification and Simulation of Modal Dynamics in Multimode Fiber Links

Karthik Choutagunta, *Student Member, IEEE*, Ian Roberts, *Student Member, OSA*, and Joseph M. Kahn, *Fellow, IEEE*

(Invited Paper)

***Abstract*—Sudden environmental effects, such as mechanical vibration, wind, and lightning, impart microsecond-timescale changes to the transmission matrix of multimode optical fibers. We introduce a timescale parameter to characterize the rate of channel changes in mode-division-multiplexed (MDM) links. We show how to efficiently generate continuous unitary matrices that form the basis for a simplified, yet general, dynamic channel model incorporating the effects of modal dispersion, mode-dependent loss, and time-varying mode coupling. We show that fast environmental perturbations can be modeled by deriving the unitary matrices from a linear ramp model, where the ramp slopes depend on the timescale parameter. We discuss the implications of channel dynamics on adaptive multiple-input–multiple-output (MIMO) equalization at the receiver in long-haul MDM links using coherent detection and in short-reach MDM links using direct detection. We show that the channel timescale and the rotation rate of the generalized Stokes vector on the generalized Poincaré sphere are both predictive of adaptive MIMO equalizer misadjustment. Our simplified channel model can be applied to optimize adaptive MIMO equalization algorithms for tracking modal dynamics.**

***Index Terms*—Adaptive equalization, dynamic channel modeling, mode-division multiplexing, multimode fiber.**

I. Introduction

MODE-DIVISION multiplexing (MDM) in multimode (MMF) or coupled-core multicore fibers (MCF) offers a means to increase capacity per fiber by transmitting information in multiple spatial modes. When D spatial and polarization modes are used in such a multiple-input multiple-output (MIMO) system, the per-fiber capacity increases in proportion to D. Ignoring nonlinearities, the MDM channel is a linear system that can be succinctly described by a $D \times D$ frequency-dependent transfer matrix that includes the effects of random mode coupling, modal dispersion (MD), and mode-dependent loss and gain (collectively referred to as MDL). Single-mode fiber (SMF) systems, which use two polarizations of a single spatial mode, are a special case of MDM with $D = 2$. Data gathered from field trial experiments and laboratory measurements suggest that SMF systems can experience fast polarization changes on the microsecond timescale [1]–[6]. While extensive field measurements of MMF system dynamics have not yet been performed, initial laboratory measurements suggest that the transfer matrix of MMF systems can change up to an order of magnitude faster than SMF systems [7], [8].

In long-haul systems employing coherent detection, adaptive MIMO equalizers track and invert the MDM channel at the receiver to establish decoupled communication subchannels. MIMO equalizers should ideally have fast tracking capabilities and low complexity, but these are often opposing objectives. Owing to the large modal group delay spread in long-haul MDM links, equalizers must have a long temporal memory, which increases computational complexity. Equalization is most efficiently performed in the frequency domain using fast Fourier transform (FFT)-based digital signal processing after prepending a cyclic prefix (CP) that is longer than the channel impulse response duration. It is necessary to use a long FFT block length N_{FFT} in the frequency-domain equalization (FDE) algorithm to achieve high CP efficiency and maximize information throughput. However, since the MIMO FDE acts on blocks of received N_{FFT} samples, the ability of the receiver to track the channel is reduced by increasing N_{FFT}. Furthermore, since longer links are susceptible to more perturbations and hence faster changes, it is crucial to optimize the adaptive equalization algorithm hyperparameters to accommodate the link's modal dispersion while enabling sufficiently fast tracking.

Dynamic channel models are needed to simulate the temporal behavior of MIMO channels so that adaptive equalization algorithms can be studied and their parameters can be optimized. The earliest dynamic models are based on the "hinge model" of SMF [4], [9]–[17]. These models treat a long SMF as a concatenation of many birefringent segments, some of which act as random polarization rotators that scramble the state of polarization (SOP) over the Poincaré sphere to govern overall polarization dynamics. Such models are simple to generate, and more importantly, easy to fit to experimental observations because both polarization degrees of freedom are fully coupled. However, a straightforward extension of the hinge model to $D > 2$ modes is not trivial because for $D > 2$, pairwise mode-coupling strengths depend on the propagation constant

Manuscript received August 25, 2018; revised December 16, 2018; accepted December 21, 2018. Date of publication December 25, 2018; date of current version April 3, 2019. This work was supported in part by Maxim Integrated and in part by a Stanford Graduate Fellowship. *(Corresponding author: Karthik Choutagunta.)*

The authors are with the E. L. Ginzton Laboratory, Department of Electrical Engineering, Stanford University, Stanford, CA 94305 USA (e-mail: kchoutag@stanford.edu; iroberts@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.2018.2889675

0733-8724 © 2018 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

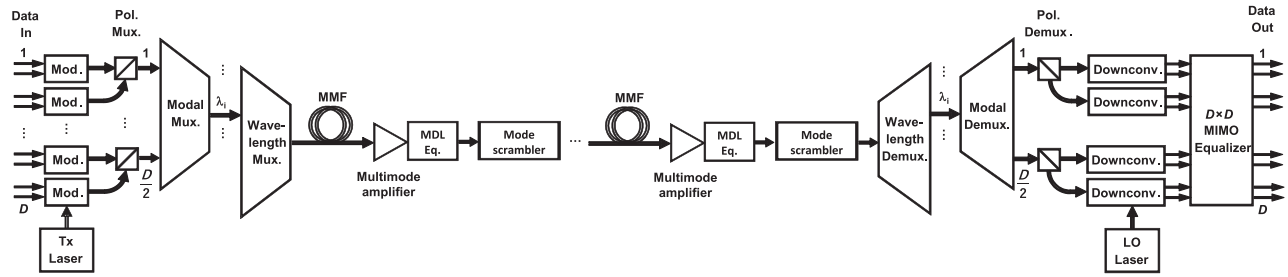

Fig. 1. Block diagram of an exemplary long-haul mode-division multiplexed system using coherent detection. Mod.: modulator, Tx: transmitter, Pol. (De)Mux.: polarization (de)multiplexer, MMF: multimode fiber, MDL Eq.: mode-dependent loss equalizer, Downconv.: downconverter, LO: local oscillator.

mismatch between the interacting modes. Furthermore, the generalized Poincaré sphere, which resides in D^2-1 dimensions, contains non-physical points that have no equivalent generalized Jones vector representations [18]. Models based purely on generalized Stokes-space rotations are unlikely to be correct, and hence, more sophisticated models are required. Recently, theoretical channel models for weakly coupled MCFs have been proposed [19], [20], in which time-varying phase shifts at discrete "phase matching points" contribute to stochastic time evolution of inter-core crosstalk. While these models are useful for modeling channel changes over timescales of seconds to minutes, they cannot readily describe mode coupling transients on the microsecond timescale that challenge adaptive MIMO equalizers.

Our previously proposed dynamic channel model describes the dynamics of mode-coupling in the generalized Jones space using coupled-mode theory [21]. It uses perturbative modelling to describe the impact of various environmental effects on modal propagation in MMFs. Manufacturing imperfections and cabling layout disturb the MMF's ideal refractive index profile and lead to static mode coupling between spatial and polarization modes. These effects are captured in a fixed frequency-dependent mode-coupling coefficient matrix (MCCM) that is held constant across time. Dynamic effects, such as stretching and bending of the fiber, or electromagnetic fields resulting from lightning strikes, modify these coefficients. Numerical integration of the coupled-mode equations yields a time-varying MDM transfer matrix. In theory, any physical perturbation can be modeled by this approach. However, this model is time-consuming to simulate, because a link of length L requires computing $O(L/dl)$ matrix exponentials, where dl is a small length increment. Another drawback of the approach in [21] is that many fiber and environmental parameters need to be specified, some of which may not be readily available.

It is highly desirable to develop channel models that simulate realistic physical responses while abstracting away parameters that do not significantly affect the temporal behavior of the link. We note that simplified channel models are not new in the field of telecommunications; they have historically been successful at explaining complicated phenomena to draw general conclusions. For example, in [22], a simplified channel model for the indoor power line channel was used to show that multipath effects give rise to log-normally distributed fading. In [23], the indoor radio propagation channel was modeled as a linear time-varying filter at each location in three-dimensional space to derive various properties of the signal arrival times and the temporal and spatial variations of the channel. Reference [24] showed that realistic multipath infrared channels can be characterized well by only the optical path loss and rms delay spread, and also developed a computationally efficient model to evaluate the impact of multipath distortion. In [25], [26], the authors show that a simplified theoretical model for polarization-mode dispersion (PMD) in SMF systems can reproduce the measured statistics of the PMD vector of an installed system using just two parameters, and that this model can predict the evolution of outage probability. In this paper, we aim to achieve something similar for optical signal propagation in MMFs and MCFs by providing a simplified method to study the effects of MDM channel dynamics.

This paper is organized as follows. Section II describes the theory behind the construction of our simplified dynamic channel model. We present a simple method to generate time-varying MDM channels that are perturbed by fast or slow perturbations. Section III discusses the effects of channel dynamics on adaptive MIMO equalization in long-haul links using coherent detection. Section IV discusses adaptive MIMO equalization for short links using direct detection. Section V provides conclusions.

II. Simplified Dynamic Channel Model

We present a method to easily construct MIMO channel matrices with desired dynamic mode coupling properties. Subsection II-A reviews multi-section modeling of MDM channels with mode coupling, modal dispersion, and mode-dependent loss. Subsections II-B and II-C show how to simulate the time-evolution of MIMO channels in response to fast and slow environmental perturbations, respectively. In Subsection II-D, we compare the simplified channel model to our previously proposed rigorous model and show how to extract information from the rigorous model to use in the simplified model.

A. Multi-Section Modeling of MDM Channels

We consider an end-to-end MDM link with D spatial and polarization modes, shown in Fig. 1. Following the multi-section approach, we model the link as a concatenation of N independent sections of fiber. The end-to-end propagation operator can be written as a $D \times D$ matrix

$$\mathbf{M}_{tot}\left(\Omega\right) = \prod_{k=1}^{N} \mathbf{M}^{(k)}\left(\Omega\right). \tag{1}$$

A long-haul MDM link can be modeled using (1) by employing a large number of independent sections ($N \gg 1$). For short-reach systems, it is often sufficient use a small number of independent sections ($N \approx 1-2$) because the modes experience weak coupling.

The propagation operator for the kth section at each angular frequency Ω around a center frequency Ω_0 can be written as a $D \times D$ matrix of the form

$$\mathbf{M}^{(k)}(\Omega) = \mathbf{V}^{(k)} \times \cdots \\ \operatorname{diag}\left[\exp\left(\mathbf{g}^{(k)}/2 - j\,(\Omega-\Omega_0)\,\tau^{(k)}\right)\right] \times \mathbf{U}^{(k)H}, \quad (2)$$

where $\mathbf{g}^{(k)}$ is a D-dimensional vector of the uncoupled modal gains in log-power-gain units, $\tau^{(k)}$ is a D-dimensional vector of uncoupled group delays, and $\mathbf{V}^{(k)}$, $\mathbf{U}^{(k)}$ are basis-changing unitary matrices accounting for mode coupling. The superscript ^{H} denotes matrix Hermitian conjugate. We use the short-hand notation $\exp(\cdot)$ to denote element-wise exponentiation, and $\operatorname{diag}(\cdot)$ to represent a square matrix formed by placing a vector on the main diagonal of a $D \times D$ matrix of zeros.

We have chosen to factor $\mathbf{M}^{(k)}(\Omega)$ like a singular value decomposition because the transfer matrix of most passive optical devices (including all fibers) can be linearized over a narrow bandwidth around a center frequency. Therefore, each section of the link is modeled as a multiplication of two unitary matrices with a (possibly) non-unitary diagonal matrix placed between them. Multiplying many sections in the form (2), with phase linearly dependent on frequency, generates higher-order dependence on frequency, provided that the $\mathbf{V}^{(k)}$ and $\mathbf{U}^{(k)}$ provide mode coupling [27].

Mode coupling in each section, described by $\mathbf{V}^{(k)}$ and $\mathbf{U}^{(k)}$, must respect the mode group structure of the MDM system. When multiple mode groups are present (for example, in graded-index MMFs or MCFs), phase-matching conditions cause the modes within the same group to couple strongly and modes from different mode groups to couple weakly [28], [29]. Therefore, these matrices are typically block diagonal, with unitary blocks corresponding to each mode group. Typically, fibers can be considered as unitary devices that do not introduce MDL, and so $\mathbf{g}^{(k)} = \mathbf{0}$ for many sections. Multimode erbium-doped fiber amplifiers (MM-EDFAs), which are periodically inserted in links to boost the signal amplitude, can be modeled by choosing non-zero $\mathbf{g}^{(k)}$ in the appropriate sections. When modeling short-reach systems, we can usually set $\tau^{(1)} = \mathbf{0}$ because modal dispersion can be considered negligible in these links.

Observing the structure of (1) and (2), we can introduce time-dependence in $\mathbf{M}_{tot}(\Omega, t)$ through the uncoupled modal gains, uncoupled group delays, or the unitary mode-coupling matrices. Indeed, in the most general case, environmental perturbations can impart changes in all three quantities. However, it is reasonable to expect that most environmental perturbations only affect local mode-coupling properties of the fiber, while leaving intact the modal gains and modal group velocities.[1] This assumption was also used in our previous, more rigorous channel model [21], where it was assumed that external perturbations lead to mode coupling and phase effects. In the following subsections, we describe methods to generate a propagation operator $\mathbf{M}_{tot}(\Omega, t)$ that evolves at desired timescales. For the purposes of our channel model, we define "timescale" as the smallest time such that the transfer matrix of a channel section becomes decorrelated (see Appendix A).

[1]Since our model is based on concatenating multiple independent matrices for static and dynamic sections, the mixing of terms due to matrix multiplication still allows the overall *coupled* group-delays and modal gains to change with time.

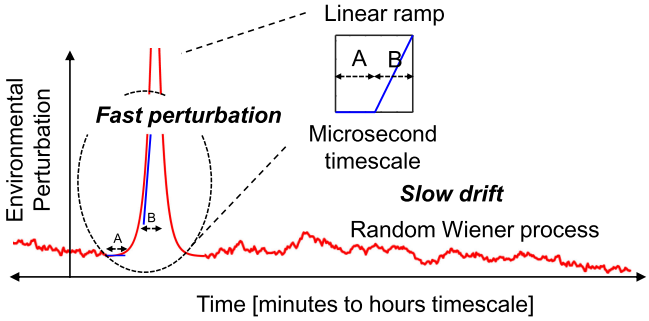

Fig. 2. Modeling the smooth temporal evolution of MDM channels in the presence of fast and slow enviromental perturbations. Fast environmental perturbations impose microsecond timescale changes to the channel. For the purpose of evaluating adaptive equalizer performance, we can model fast perturbations as a linear ramp in the time domain, where the slope of the ramp signifies the worst-case rate of change of the channel. Slow perturbations over the second-to-minute timescales stress the equalizer less and can be modeled as a drift using random Wiener processes.

B. Modeling Fast Perturbations

In this section, we use the technique presented in Appendix A for generating smoothly varying unitary matrices to produce fast-changing MDM channels (timescale $\tau_{env} \approx 1-100\ \mu$s) in response to a sudden perturbation.

Our MMF link comprises N sections, so we must specify a desired timescale of change for each section $\tau_{env}^{(k)}$, $k = 1, \ldots, N$. For well-isolated fiber sections that remain static during a perturbation event, we can assign $\tau_{env}^{(k)} = \infty$. In the following, we will focus on the modeling fast perturbations of a single section in the link.

As shown in Fig. 2, fast perturbations are typically characterized by impulse-like behavior, caused by sudden events such as strong wind gusts [30], lightning strikes [31], [32], or impacts by mechanical tools [2], [3]. During such sudden events, which can last for several seconds, the elements of the MDM channel's transfer matrix fluctuate with large amplitudes and at high frequencies. If the perturbation is elastic and causes no long-term change to fiber properties, the channel eventually relaxes back to its original steady-state condition. If the perturbation permanently changes the fiber properties (e.g., if the fiber cabling is disturbed), then the MDM channel may relax to a different state.

Even though fast external perturbations to the MDM link can last up to a few seconds, we only need to model a few microseconds of the much longer perturbation. This is because we are most interested in the effects of channel dynamics on adaptive MIMO equalization. Adaptive equalizers in MDM systems operate at symbol rates of tens of GHz, and a simulation of several microseconds corresponds to thousands of received symbols, which is sufficient for studying equalizer convergence and

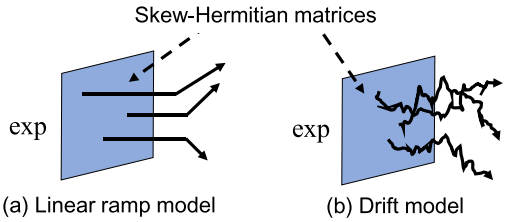

Fig. 3. Fast perturbations using (a) a linear ramp model and slow perturbations using (b) a drift model can be simulated by matrix exponentials of time-varying skew-Hermitian matrices.

tracking behavior. In studying adaptive equalizer performance, it is sufficient to model the worst-case slope of the channel variation, since this places the greatest stress on the adaptive equalizer. Hence, we model fast perturbations using a *linear ramp model* so that we can conveniently study both equalizer convergence (e.g., during time period A in Fig. 2 inset) and equalizer tracking (during time period B).

Modifying (2) to include dynamic effects caused by changes in the unitary mode coupling matrices, the transfer matrix for the kth section of fiber is given by

$$\mathbf{M}^{(k)}(\Omega,t)=\mathbf{V}^{(k)}(t)\times\cdots \\ \operatorname{diag}\left[\exp\left(\mathbf{g}^{(k)}/2-j(\Omega-\Omega_0)\tau^{(k)}\right)\right]\times\mathbf{U}^{(k)H}(t), \tag{3}$$

where $\mathbf{V}^{(k)}(t)$ and $\mathbf{U}^{(k)H}(t)$ are to be chosen so that $\mathbf{M}^{(k)}(\Omega,t)$ changes at a timescale $\tau_{env}^{(k)}$. To model a linear ramp perturbation of the section starting at time t_0, we construct $\mathbf{V}^{(k)}(t)$ and $\mathbf{U}^{(k)}(t)$ as

$$\mathbf{V}^{(k)}(t)=\exp\left(\mathbf{V}_{sh}^{(k)}(t)\right),\quad \mathbf{U}^{(k)}(t)=\exp\left(\mathbf{U}_{sh}^{(k)}(t)\right),$$
$$\mathbf{V}_{sh,ij}^{(k)}(t)=\mathbf{V}_{sh,ij}^{(k)}(t_0)+s_{ij}^{(k)}\cdot r(t-t_0),$$
$$\mathbf{U}_{sh,ij}^{(k)}(t)=\mathbf{U}_{sh,ij}^{(k)}(t_0)+s_{ij}^{(k)}\cdot r(t-t_0),$$
$$\mathrm{Re}s_{ij}^{(k)},\mathrm{Im}s_{ij}^{(k)}\sim\mathcal{N}\left(0,\sigma_{ramp,k}^2\right),\ s_{ij}^{(k)}=-s_{ji}^{(k)*}, \tag{4}$$

where

$$\sigma_{ramp,k}^2=\frac{\kappa_D}{\left(\tau_{env}^{(k)}\right)^2}, \tag{5}$$

is the perturbation strength of the linear ramp, $r(t)=\max(t,0)$ is a unit ramp function and

$$\kappa_D\approx 0.154+\frac{7.361}{D}-\frac{18.255}{D^2}+\frac{34.101}{D^3}, \tag{6}$$

is a scaling factor that depends on the number of modes (see Appendix A). The subscript "sh" in $\mathbf{V}_{sh}$ and $\mathbf{U}_{sh}$ stands for "skew Hermitian" and reminds the reader that $\mathbf{V}_{sh}^{(k)}(t)=-\mathbf{V}_{sh}^{(k)H}(t)$ and $\mathbf{U}_{sh}^{(k)}(t)=-\mathbf{U}_{sh}^{(k)H}(t)$. The matrix exponential of a skew-Hermitian matrix yields a unitary matrix.

When the channel is stable, built-in defects in the MMF cause static mode-coupling effects described by off-diagonal elements in $\mathbf{V}_{sh}^{(k)}(t)$ and $\mathbf{U}_{sh}^{(k)}(t)$. As visualized in Fig. 3(a), the expressions in (4) assign time-domain ramps of random slopes to each of the elements of $\mathbf{V}_{sh}^{(k)}(t)$ and $\mathbf{U}_{sh}^{(k)}(t)$. The slopes of the skew-Hermitian exponent ramps $s_{ij}^{(k)}$ are chosen from a complex Gaussian distribution with variance $\sigma_{ramp,k}^2$ prior to the start of the ramp. The choice of the probability distribution for the slopes is somewhat arbitrary; although we have decided to use the normal distribution here as a proof-of-concept, it is also possible to use other distributions, such as a uniform distribution. Since these random slopes are multiplied by the time variable t, the variance of the random skew-Hermitian perturbation increases quadratically with time. After time offset $\tau_{env}^{(k)}$, $\mathbf{V}_{sh}^{(k)}(t)$ and $\mathbf{U}_{sh}^{(k)}(t)$ are perturbed by random skew-Hermitian matrices whose entries have variance $t^2\times\sigma_{ramp,k}^2\big|_{t=\tau_{env}^{(k)}}=\kappa_D$. Smaller values of $\tau_{env}^{(k)}$ permit larger slopes to be sampled, which results in a faster-changing $\mathbf{M}^{(k)}(t)$. As seen in Fig. 4(a), this scheme produces MDM channels that rapidly change at desired microsecond timescales.

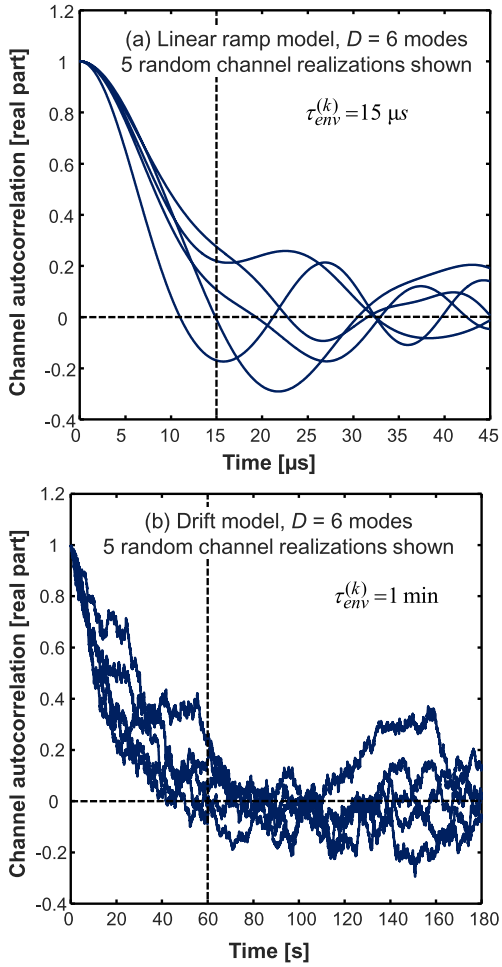

Fig. 4. Channel autocorrelation function for five random dynamic channels with $D=6$ modes. In (a), the channel dynamics are simulated using a linear ramp model with $\tau_{env}^{(k)}=15$ μs,whereas in (b), the channel dynamics are simulated using the drift model with $\tau_{env}^{(k)}=1$ min.

Fig. 5 shows an example simulation of the linear ramp model for $D=6$ modes. The link contains 10 total sections of unitary fiber transfer matrices, of which one section is dynamic and

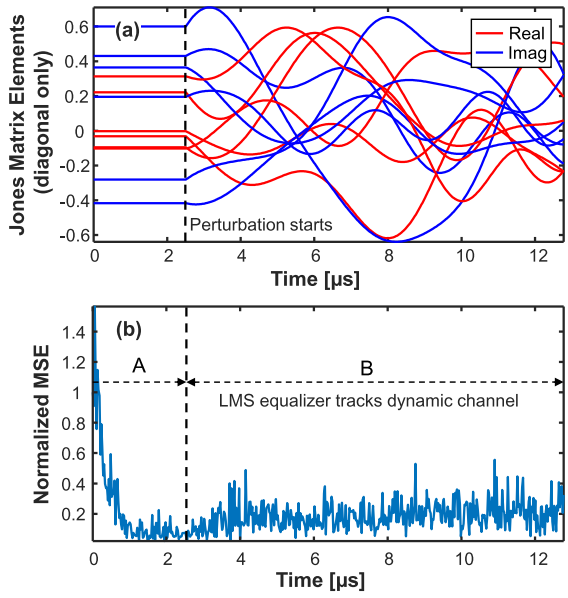

Fig. 5. (a) Temporal evolution of the generalized Jones matrix of a $D = 6$-mode MDM channel using the linear ramp model. Only the diagonal elements of the 6×6 channel matrix are shown for clarity. The link is comprised of 10 sections of fiber, where one section is dynamic and has timescale $\tau_{env} = 5$ μs. (b) An LMS MIMO equalizer with block size $N_{FFT} = 512$ and step size $\mu = 0.05$ tracks the modal dynamics. The taps of the equalizer converge when the channel is static, but misadjust when the channel perturbation is turned on, leading to an excess output mean-squared error.

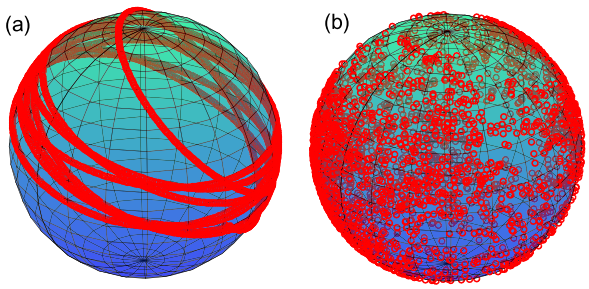

Fig. 6. Movement of the Stokes vector on the Poincaré sphere during a dynamic perturbation produced by (a) the linear ramp model, and (b) the drift model.

has timescale $\tau_{env} = 5$ μs. We chose random group delays of each fiber section such that the root-mean-square (rms) coupled group delay spread of the end-to-end link is 10 ns. We assume for sake of simplicity that there are no sources of MDL in the link (all MM-EDFAs are ideal). We also neglect chromatic dispersion and high-order modal dispersion, as they do not have significant impact on channel dynamics. We further assume that an adaptive MIMO equalizer (discussed in detail in Section III-A) tracks the channel dynamics at the receiver. In Fig. 5(a), the elements of the channel's generalized Jones matrix are stable until $t = 2.5$ μs, when the perturbation starts. Fig. 5(b), shows the equalizer taps converging to yield a low output error, and then slightly misadjusting while trying to track the dynamics of the changing channel.

Fig. 6(a) shows an example simulation of a linear ramp model for the case of $D = 2$ modes, applicable for fast perturbations of dual-polarization SMF systems. A fast ramp-like perturbation causes the three-dimensional Stokes vector to orbit around the Poincaré sphere.

C. Modeling Slow Perturbations

In this section, we discuss how to generate MDM channels that are affected by slow perturbations, e.g., a slow drift of the MMF properties over minutes or hours ($\tau_{env} \approx 1$ m $-$ 1 hr) due to changes in ambient temperature or other slow phenomena. Ambient phenomena such as changing temperature during daytime and nighttime [33] and low-amplitude vibrations from traffic and seismic activity [34] continually perturb the MDM channel. Since the channel changes very little due to such effects over timescales comparable to the symbol rate, the receiver's adaptive MIMO equalizer can easily track changes to the channel and any impact on the equalizer output error is negligible. Nevertheless, for the sake of completeness, we show how our simplified dynamic channel model can describe the slow evolution of the MDM channel.

If many independent perturbation sources are acting upon the MDM channel simultaneously, then it is reasonable to model the evolution of $\mathbf{V}_{sh}^{(k)}(t)$ and $\mathbf{U}_{sh}^{(k)}(t)$ by a random Wiener process. A similar approach was recently used in [16] to model the polarization drift in SMF systems. We can construct a *drift model* as

$$\begin{aligned}
&\mathbf{V}^{(k)}(t) = \exp\left(\mathbf{V}_{sh}^{(k)}(t)\right), \quad \mathbf{U}^{(k)}(t) = \exp\left(\mathbf{U}_{sh}^{(k)}(t)\right), \\
&\mathbf{V}_{sh,ij}^{(k)}(t+\Delta t) = \mathbf{V}_{sh,ij}^{(k)}(t) + s_{ij}^{(k)}\sqrt{\Delta t}, \\
&\mathbf{U}_{sh,ij}^{(k)}(t+\Delta t) = \mathbf{U}_{sh,ij}^{(k)}(t) + s_{ij}^{(k)}\sqrt{\Delta t}, \\
&\mathrm{Re}s_{ij}^{(k)}, \mathrm{Im}s_{ij}^{(k)} \sim \mathcal{N}\left(0, \sigma_{drift,k}^2\right), \; s_{ij}^{(k)} = -s_{ji}^{(k)*},
\end{aligned} \tag{7}$$

where

$$\sigma_{drift,k}^2 = \frac{\kappa_D}{\tau_{env}^{(k)}}, \tag{8}$$

and Δt is a characteristic time step for the evolution of the channel drift. Note the presence of the $\sqrt{\Delta t}$ factors in (7). Since the variance of a random Wiener process increases linearly with time, the $\sqrt{\Delta t}$ scaling in (7) ensures that after each time offset $\tau_{env}^{(k)}$, $\mathbf{V}_{sh}^{(k)}(t)$ and $\mathbf{U}_{sh}^{(k)}(t)$ are perturbed by random skew-Hermitian matrices whose entries have variance $t \times \sigma_{drift,k}^2|_{t=\tau_{env}^{(k)}} = \kappa_D$. As $\tau_{env}^{(k)}$ is decreased, the strength of the drift is increased, and the MDM channel decorrelates faster. As seen in Fig. 4(b), this scheme produces MDM channels that slowly change at the desired seconds-to-hours timescales. When the drift model is applied to SMF systems, the Stokes vector undergoes a random walk on the Poincaré sphere, as seen in Fig. 6(b).

D. Comparing the Simplified and Rigorous Channel Models

In this section, we characterize the modal dynamics of an exemplary aerial fiber link perturbed by a wind gust using two approaches: a rigorous channel model based on coupled

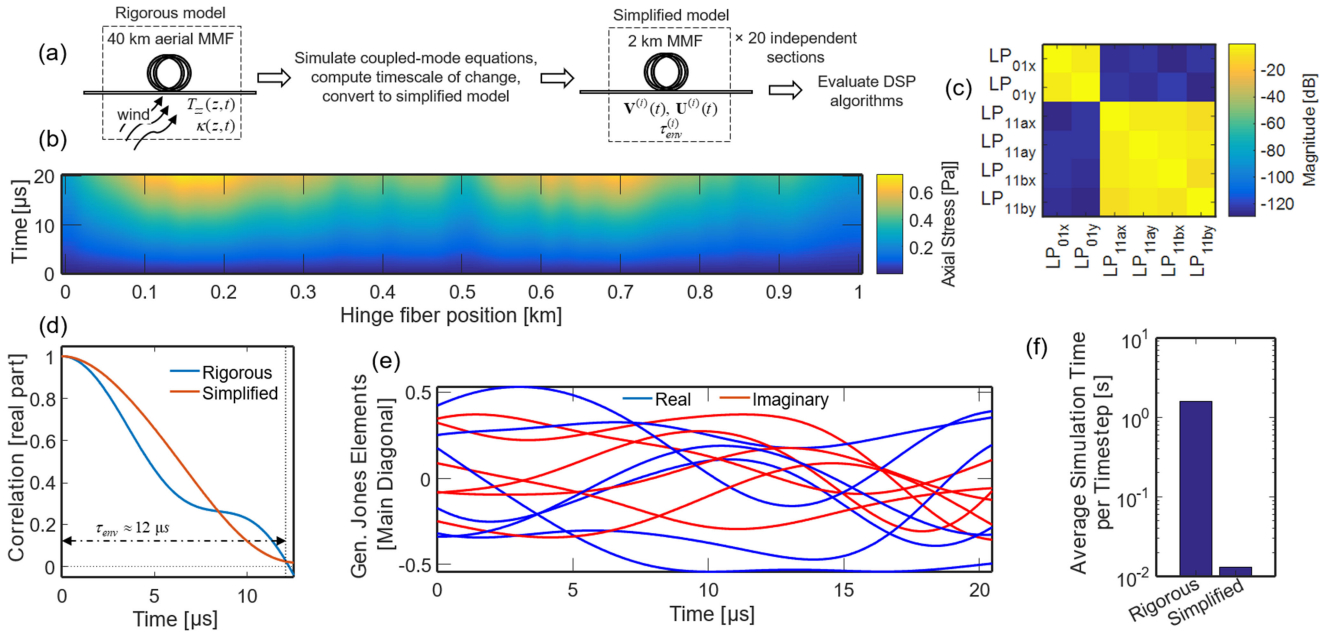

Fig. 7. Converting from a rigorous dynamic MDM channel model to a simpified model, illustrated for a 6-mode aerial fiber blowing in the wind. (a) Schematic of procedure. (b) Axial stress distribution resulting from a wind perturbation, shown for a 1 km section of the aerial fiber. (c) Coupling matrix obtained using the rigorous channel model. (d) Correlation curve for the rigorous model and the simplified model, both of which have a timescale $\tau_{env} = 12$ μs. (e) Evolution of the elements of the 6×6 generalized Jones matrix of the channel during the wind perturbation obtained from the simplified model. Only the main diagonal elements are shown for clarity. (f) Comparison of simulation times per time step for the rigorous and simplified channel models.

mode theory [21], and the simplified model presented in this paper. We show that while the rigorous model can predict the behavior of the link using the wind distribution and other physical attributes, the simplified model can reproduce a similar channel impulse response in less time using fewer parameters. While the rigorous model is well suited for *predicting* the temporal behavior of modes in an MDM link for arbitrary fibers or for arbitrary environmental perturbations, it is advantageous to use the simplified model to *recreate* similar channels to study signal processing algorithms for tracking modal dynamics.

The simulated aerial fiber is a multimode graded-index graded-depressed cladding (GIGDC) fiber supporting $D = 6$ spatial and polarization modes that is optimized for low uncoupled group delay (GD) spread. As shown in Fig. 7(a), a 40 km section of this fiber is exposed a wind perturbation that is smoothly varying on a sub-second to second timescale. We assume that the wind gust affects the MMF in two ways: (a) axial stress is induced inside the fiber, and (b) the fiber bends, causing it to relax from one configuration to another.

To model (a), we assume that a very small fraction of the fiber yield stress ($T_{\mathrm{zz,max}} = 0.5$ GPa [35], [36], defined as the maximum axial stress producing elastic deformation) is induced inside the fiber in a characteristic timescale of 0.1 s. Since a wind gust will be stronger in some locations and weaker in others, we generate random spatially correlated tensile stresses that stretch the fiber using a Langevin process with a correlation length of 20 m. The distribution of the axial stress for a 1 km section of fiber is shown in Fig. 7(b). In the worst case, only 0.00015% of the yield stress is induced in the fiber, and this stress is low enough such that the perturbative modeling described in [21] is valid. To model (b), we assume that the local curvatures of the MMF relax from one distribution to another in the same timescale as stress, because the fiber bends in response to wind. The initial curvatures of the fiber bends, $\kappa_0(z) = 1/|R_0(z)|$, are sampled from the positive part of a normal distribution with a standard deviation of 1/2.5 m^{-1}. The curvatures of the bends 0.1 s later, $\kappa_1(z)$, are independently sampled from the same distribution. Then, the local curvature of the bend at each spatial location is linearly interpolated using $\kappa(z,t) = \kappa_0(z) + (\kappa_1(z) - \kappa_0(z)) \times (t/0.1)$. We assume that stretching and bending the fiber only cause phase and mode-coupling effects, and the channel remains unitary throughout the wind gust. To simulate the dynamics of the 40 km aerial fiber, the link is divided into shorter sections, each 10 m long, and the multisection model is employed to generate the overall transfer matrix. Fig. 7(c) shows the magnitude of the transfer matrix elements at a specific time instant. There is strong intra-group coupling within the LP_{01} and LP_{11} mode groups because bending and stretching the fiber perturbs its refractive index profile and causes modes with closely spaced propagation constants to couple. The intergroup coupling is weak because modes in different mode groups are not phase matched.

In order to simulate the wind gust using the simplified channel model, we must first determine the timescale of the channel produced by the rigorous model. We define a scalar correlation metric (see Appendix A) as

$$C(\tau) = \frac{1}{D}\mathrm{trace}\left(\mathbf{M}_{\mathrm{tot}}(t=0)\mathbf{M}_{\mathrm{tot}}^{H}(t=\tau)\right), \qquad (9)$$

where $C(0) = 1$ (following from the unitarity of the channel), $C(\tau_{env}) \approx 0$, and τ_{env} is the timescale of change of the aerial fiber. The correlation curve of the rigorous model is numerically

computed using (9) and is shown in Fig. 7(d). This suggests that $\tau_{env} \approx 12$ μs. Given τ_{env}, simulating the wind gust using the simplified model is straightforward. We proceed to construct 2×2 and 4×4 block diagonal matrices for the LP_{01} and LP_{11} mode group subspaces by using the linear ramp method described in Subsection II-B. We assume a total of $N = 20$ sections to model the 40 km hinge aerial fiber, and we adjust the timescale $\tau_{env}^{(k)}$ of each section until the timescale of the matrix product of all sections closely matches that of the rigorous model. To produce a channel with an overall timescale of 12 μs, we choose each $\tau_{env}^{(k)}$ to be 80 μs. The evolution of the generalized Jones matrix $\mathbf{M}_{\text{tot}}$ elements produced using the simplified model is shown in Fig. 7(e). Only the diagonal elements are shown for clarity. The simulation speed for both the rigorous and simplified models are shown in Fig. 7(f). Both models are implemented in MATLAB R2017a software and run on a 2015 MacBook Pro computer with a 2.8 GHz Intel Core i7 processor and 16 GB of 1600 MHz DDR3 memory. The simplified model is over an order of magnitude faster to run than the rigorous model because fewer matrix exponentiation and multiplication operations are required.

III. Tracking Modal Dynamics in Long-Haul Coherent Detection Systems

We employ the simplified channel models presented in Section II to study the dynamics of co-propagating modes in long-haul MDM systems employing coherent detection. Subsection III-A discusses the design of adaptive MIMO equalizers and the impact of equalization algorithm hyperparameters and channel dynamics on equalizer performance. Subsection III-B discusses metrics that can be used to quantify the rate of change of MIMO channels, such as the rotation rate of the generalized Stokes vector on the generalized Poincaré sphere.

A. Adaptive MIMO Equalization for Long-Haul Systems

A canonical architecture for long-haul MDM systems employing coherent detection is shown in Fig. 1. In this section, we briefly review key concepts regarding adaptive MIMO equalization for MDM links.

In long-haul systems, the group velocity differences between the propagating modes cause the channel to become frequency-dependent and have a lengthy impulse response. When the overall channel $\mathbf{M}_{tot}(\Omega, t)$ is time-varying, the receiver tracks the mode-coupling dynamics using a frequency-dependent adaptive MIMO equalizer $\mathbf{W}(\Omega, t)$ such that the matrix product $\mathbf{W}(\Omega, t)\mathbf{M}_{tot}(\Omega, t)$ is approximately a $D \times D$ identity matrix.

To ensure tolerable complexity of adapting $\mathbf{W}$, most systems implement equalization in the frequency domain using FFT-based signal processing. FDEs can either be implemented using block convolution (e.g., using overlap-save convolution, as in [37], [38]) or by prepending a cyclic prefix of N_{CP} symbols before each block of N_{FFT} symbols before transmission and processing each block independently of others. The latter approach is preferred because it simplifies the realization of an adaptive FDE that can track MDM channel dynamics even though there is a small throughput loss of $N_{CP}/(N_{CP} + N_{FFT})$ (up to 20% for practical systems [39]) resulting from discarding the cyclic prefix samples.

There exist many candidate algorithms to adapt the equalizer, such as the least mean squares (LMS) algorithm, normalized least mean squares (NMLS) algorithm, and the recursive least squares (RLS) algorithm, among others [39]. LMS is by far the most popular choice due to its computational simplicity. The filter taps of the $D \times D$ MIMO-LMS equalizer $\mathbf{W}$ at each discrete frequency $k = 0, 1, \ldots, N_{FFT} - 1$ are iteratively updated as [21]

$$\mathbf{W}[k] \leftarrow \mathbf{W}[k] + \mu\left(\tilde{\mathbf{x}}[k] - \mathbf{W}[k]\,\tilde{\mathbf{y}}[k]\right)\tilde{\mathbf{y}}^{H}[k], \tag{10}$$

where μ is a scalar step-size parameter,

$$\tilde{\mathbf{x}}[k] = [X_1[k], \ldots, X_D[k]]^T \tag{11}$$

the N_{FFT}-point FFT of a time-domain block $\mathbf{x}[n]$, is a block of known or estimated frequency-domain transmitted data symbols,

$$\begin{aligned}\tilde{\mathbf{y}}[k] &= \mathbf{M}_{tot}[k]\,\tilde{\mathbf{x}}[k] + \tilde{\mathbf{n}}[k] \\ &= [Y_1[k], \ldots, Y_D[k]]^T\end{aligned} \tag{12}$$

the N_{FFT}-point FFT of a time-domain block $\mathbf{y}[n]$, is a block of received frequency-domain symbols at the input of the equalizer, and $\tilde{\mathbf{n}}[k] = [N_1[k], \ldots, N_D[k]]^T$ is a block of noise samples. The step size parameter μ is often chosen empirically in practice to ensure fast convergence and tracking, and low output mean-squared error (MSE). Some MDM systems initially use the MIMO-LMS algorithm in data-aided mode, and after the filter taps have converged, track the channel using blind equalization such as the constant-modulus algorithm or the multi-modulus algorithm, depending on the modulation format [40].

B. Metrics for Rate of Change of the MDM Channel

The equalizer design objectives include fast convergence of the filter taps to the optimal MMSE estimate during initial adaptation, low MSE floor when the channel is stable, and a small MSE when the channel is dynamic. The best choices for the adaptive equalization algorithm parameters, such as μ, N_{FFT}, N_{CP} and the frequency of equalizer updates, all depend on how fast the MDM channel changes. It is necessary to continually monitor the rate of change of the MDM channel to ensure that time-varying modal dispersion and mode coupling are properly compensated [41]. Since a $D \times D$ MDM system will have D^2 frequency-dependent and time-varying elements in its transfer matrix, choosing the proper metric to quantify a channel's rate of change is a non-trivial problem. A good metric for the rate of change of the MDM channel should be correlated closely to the difficulty of tracking by an adaptive algorithm or, equivalently, to the equalizer's instantaneous output MSE.

A *de-facto* standard for measuring channel change is the average rate of rotation of an output generalized Stokes vector on the generalized Poincaré sphere [2]. Alternative metrics for the rate of channel change can also be formed in the generalized Jones space [7], but these metrics are not as widely used in

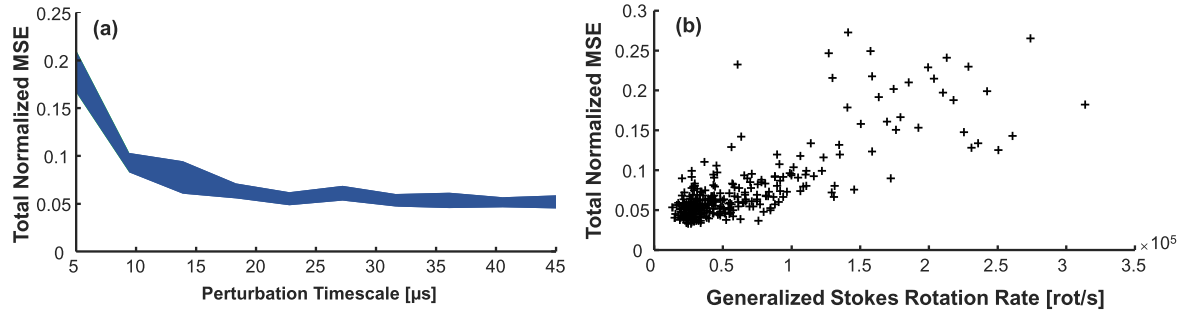

Fig. 8. (a) Steady-state dynamic MSE of the LMS equalizer is inversely proportional to the channel's perturbation timescale. Blue shaded region shows one standard deviation around the mean value. (b) The generalized Stokes vector rotation rate is correlated with the tracking performance of the equalizer.

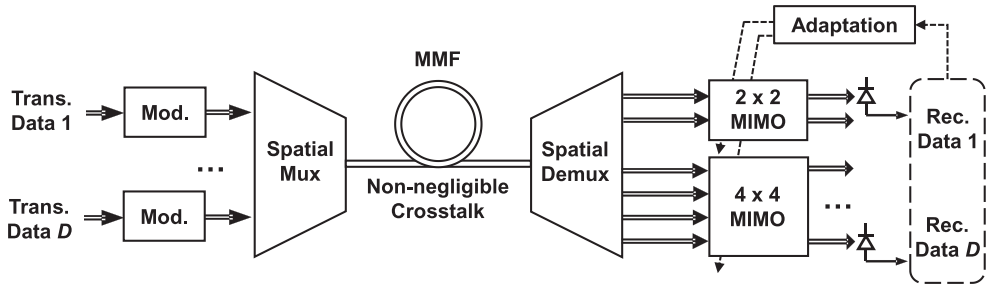

Fig. 9. Architecture for short-reach MDM links employing direct detection. In this example link, $D = 6$ spatial and polarization modes are adapted by optical MIMO equalizers in the 2×2 and 4×4 mode group subspaces.

MDM literature. A D-dimensional and complex-valued generalized Jones vector $\mathbf{j}$ can be converted to a $D^2 - 1$-dimensional real-valued generalized Stokes vector $\mathbf{s}$ as [18], [41]

$$\mathbf{s}_i = \sqrt{\frac{D}{2\,(D-1)}}\mathrm{trace}\left(\sigma_i \mathbf{j}\mathbf{j}^H\right) = \sqrt{\frac{D}{2\,(D-1)}}\,\langle \mathbf{j}|\sigma_i|\mathbf{j}\rangle\,, \tag{13}$$

where σ_i is the ith Hermitian matrix of the generalized Gell-Mann basis. Using (13), the average generalized Stokes vector rate of rotation on the generalized Poincaré sphere can be defined in the time interval $[t, t + \Delta t]$ as

$$r_{Stokes-rotation}(t) = \arccos\left(\mathbf{s}(t) \cdot \mathbf{s}(t + \Delta t)\right) / \left(2\pi \Delta t\right), \tag{14}$$

which is a measure of the solid angle spanned by the generalized Stokes vector $\mathbf{s}$ in duration Δt. The Stokes rotation rate is unaffected by laser phase noise and can be experimentally measured by a polarimeter. In systems with a small number of modes, higher rotation rates are correlated with increased difficulty of tracking by an adaptive equalizer because all degrees of freedom are usually fully coupled. However, even though (14) is mathematically valid for all D, it is not immediately clear if rotation of the generalized Stokes vector on the generalized $D^2 - 1$-dimensional Poincaré sphere is a good indicator of the equalizer's instantaneous MSE. Fig. 8 shows the impact of channel dynamics on equalization performance. As seen in Fig. 8(a), adaptive equalization yields lower output MSE on channels that change more slowly. Since Fig. 8(b) shows that equalizer output MSE is roughly correlated to the rotation rate of the generalized Stokes vector, it can be concluded that the latter is indeed a good metric of channel rate of change.

IV. Tracking Modal Dynamics in Short-Reach Direct Detection Systems

In this section, we consider MIMO equalization for short-reach systems. Subsection IV-A discusses the canonical architecture of short-reach MDM links, including key differences from long-haul systems. Subsection IV-B discusses various methods for tracking modal dynamics using optical MIMO equalizers, which have the challenging task of unscrambling mode coupling in the optical domain, prior to photodetection and any further electronic processing.

A. Direct-Detection MDM System Architecture

The growing demand for bandwidth in intra-datacenter links (~ 0.1 – 2 km) motivates the use of multiple modes to increase information capacity. Since these links must operate under tight cost and power budgets, direct detection (DD) is preferred over coherent detection. The most common approach to scaling capacity in short-reach DD systems is mode-group division multiplexing (MGDM), where the different mode groups are used as channels of communication [42]. MGDM is a suboptimal approach, however, because the degrees of freedom provided by the multiplicity of mode groups $(\mathrm{roughly}\sqrt{D})$ is much smaller than the total number of modes D. Realization of DD-MDM using all available modes to transmit information has been an active area of recent research [43]–[45].

Fig. 9 shows the canonical architecture for DD-MDM systems. Mode-selective multiplexers, such as multiplane light converters (MPLCs) [46], [47], are used to selectively multiplex data signals onto individual modes of an MMF. The MMF is assumed sufficiently short that frequency-dependent effects

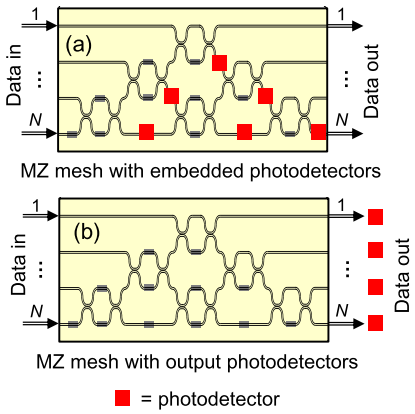

Fig. 10. Physical realization of an $N \times N$ optical MIMO equalizer using a mesh of Mach-Zehnder interferometers arranged in a triangular lattice. (a) shows a mesh with embedded photodetectors, whereas (b) contains photodetectors at the output only.

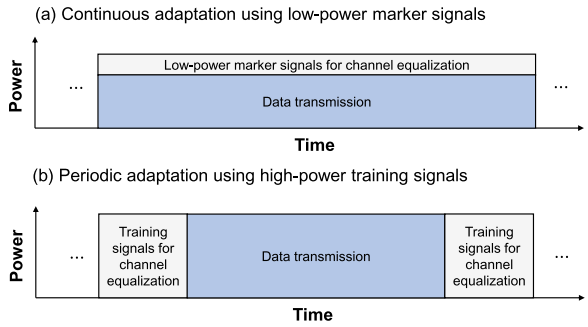

Fig. 11. Adaptive optical MIMO equalizers implemented using a mesh of MZ interferometers can track a dynamic channel either (a) continuously or (b) periodically. In (a), both the high-speed data signal and a superimposed low-amplitude marker signal, which uniquely identifies the mode, are transmitted continuously in each mode. In (b), equalization is performed at periodic intervals, during which data signals are stopped and replaced by high-power training signals.

such as chromatic and modal dispersion are negligible, although bends, twists and other perturbations along the link may introduce mode coupling within the mode groups. After a mode-selective demultiplexer (for example, an MPLC used in reverse configuration) is used to spatially separate the modes, optical MIMO processing in the mode-group subspaces is required to undo modal crosstalk and separate the data signals. Optical MIMO can be physically realized by a mesh of Mach-Zehnder (MZ) interferometers [44] or a cascade of multimode interference couplers [48], either of which effectively implements a unitary matrix multiplication that is controllable by internal phase shifters. The task of determining the required phase shifts, however, is challenging in DD systems because the receiver detects only intensity and does not have directly extract information on the phases of received optical signals. Adapting and tracking channel dynamics must be done based only on a sequence of intensity measurements. Since the channels of short links can vary on millisecond timescales, these optical MIMO equalizers must be able to adapt on sub-millisecond to millisecond time scales. The following subsection presents several methods of adapting optical MIMO equalizers.

B. Methods of Adapting Optical MIMO Equalizers

We focus on the problem of adapting a mesh of MZ interferometers that is arranged in a triangular configuration, as shown in Fig. 10. The mesh optically implements an $N \times N$ unitary matrix multiplication, where $N \leq D$ is the size of a single mode-group subspace. The mesh contains photodetectors, which can be placed either inside the mesh or external to the mesh, to sample the powers of data signals and provide feedback to an adaptive configuration algorithm. The goal is to set phase shifters inside the mesh so interference causes all the power from data signal i to be routed to data output i, for each $1 \leq i \leq N$. When the mesh phase shifts are properly adapted, the matrix product of the channel and mesh transfer matrices is a diagonal unitary matrix, i.e., the diagonal elements are of the form $e^{j\theta_{ii}}$. This is a more relaxed requirement than traditional MIMO channel equalization because we do not require the product to be an identity matrix. The transmitter modulates the intensity of the data symbols and the receiver uses direct detection, which is independent of the value of θ_{ii}.

Equalization methods for DD systems can be grouped into two categories. Those in the first category perform channel estimation using phase retrieval techniques [43]. These methods use known training signals to estimate the channel by solving a sequence of convex optimization problems. Once an estimate of the channel is found, the channel inverse is realized optically by a MZ mesh. However, these methods can work only if the MZs inside the mesh are calibrated precisely across a range of temperatures, potentially making these methods difficult to use in practice. The methods in the second category use self-configuration, proposed by Miller in [49], to undo modal crosstalk by directly setting the required phase shifts without an explicit channel estimation step. In self-configuration, mutually orthogonal marker signals are superimposed onto the data signals, and the phase shifters are optimized sequentially based on measurement of these marker signals at photodetectors inside the mesh. Because the marker signals corresponding to different data signals are mutually orthogonal, self-configuration can determine what fraction of the power of each mode is present at each phase shifter location. These measurements are used to sequentially minimize modal crosstalk inside the mesh so that all the power from signal *i* is correctly routed to data output *i*. In this paper, we consider the self-configuration method and variants derived from it, because they are robust against unknown initial phase shifts in the mesh caused by fabrication tolerances or temperature variations. The convergence speed and error performance of these methods are affected by several design choices, including (i) use of continuous or periodic equalization, (ii) the type of transmitted marker signals, (iii) use of in-mesh photodetectors or output-only photodetectors, and (iv) the use of special optimization methods to speed up tracking. We discuss the effects of these four design choices below.

1) Continuous and Periodic Equalization: Optical MIMO can track channel dynamics either continuously (at the same time as data transmission), or periodically (payload data is turned off and replaced by known training signals). As shown in Fig. 11(a), continuous adaptation of the MZ interferometer mesh requires a small portion of the power of each data signal

to be allocated to a low-speed marker signal. The original self-configuration algorithm [49] works best with in-mesh detectors that sample a small fraction of the signal power, and this fraction must be chosen large enough such that accurate crosstalk measurements can be made, but low enough such that the propagating signals still have high signal-to-noise ratios (SNRs) when they reach the mesh output. Since the marker signals are superimposed onto high-speed data signals, interference from the high-speed data signals can cause noise that disrupts the adaptation process, leading to mesh acquisition times longer than the required millisecond timescale [50]. To prevent interference from high-speed payload data, we propose periodic equalization using training signals, as shown in Fig. 11(b) [50]. In this scheme, the high-speed data is turned off periodically, and the entire signal power is allocated to training symbols, causing a loss in data throughput. The main advantage is that since the training symbols are detected at higher SNRs, they can be processed faster.

2) Type of Transmitted Marker Signals: The original self-configuration algorithm proposed in [49] superimposes sinusoidal pilot tones at different frequencies [51] onto the data signals. We instead propose to use mutually orthogonal codes based on rectangular pulses [50]. Rectangular pulse-based codes maximize the detected marker energy for a given marker signal amplitude. Using rectangular-pulse codes can enable either a more accurate measurement of crosstalk at the mesh photodetectors, or a faster measurement time for the same level of accuracy, which allows tracking faster modal dynamics.

3) In-Mesh Vs Output-Only Photodetectors: As shown in Fig. 10, photodetectors that measure signal powers can be placed either inside the mesh or external to the mesh. When detectors are placed inside the mesh, all phase shifters can be optimized sequentially in a single pass, as described by the original self-configuration algorithm in [49]. When detectors are placed only at the mesh output, it is no longer possible to set the phase shifters in a single pass; an iterative algorithm must be used instead [50]. Our simulations in Fig. 12 indicate that periodic equalization using in-mesh detection and periodic equalization using output-only detection take approximately the same time to converge. The in-mesh detectors sample a small portion of the impinging power, leading to longer measurement times per detector. The output-only detectors sample all the impinging power, achieving a higher SNR in each measurement; however, with output-only detectors, the phase shifters in the MZ mesh must be iteratively optimized many times. While this finding may seem to imply that it is better to use output-only detectors to reduce hardware cost, the benefit of in-mesh detectors becomes evident when improved optimization methods are used, as discussed below.

4) Special Optimization Methods: In the original self-configuration algorithm using in-mesh detection [49], each phase shifter is optimized by tuning the phase shift across a continuous range until signal crosstalk is minimized. However, the adaptation time can be decreased by up to an order of magnitude if minimization is replaced by curve fitting, as shown in Fig. 12. When using in-mesh photodetectors, the detected marker or training signal power is a sinusoidal function of each MZ phase adjustment, assuming correct tuning of MZs in previous stages [50]. In the curve-fitting technique, several power measurements at different phase values are used to fit a sinusoidal function of the mesh phase values. Minimization of the single-period sinusoid yields the optimal phase shift.

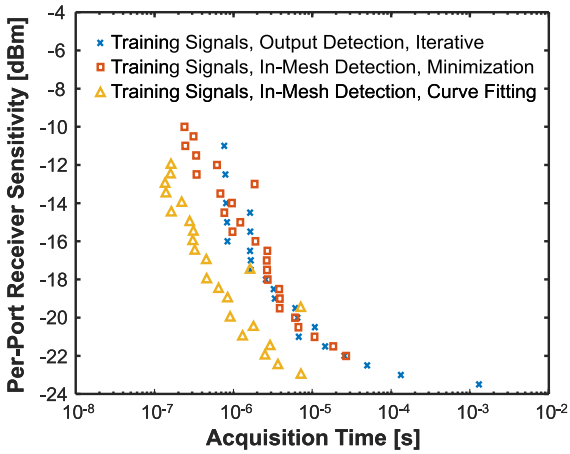

Fig. 12. Comparison of MZ mesh tracking schemes showing the per-port receiver sensitivity achieved for various acquisition times, for a mesh size of $N = 4$ ports. All three schemes use high-power training signals to periodically adjust the mesh phase shifters on a sub-millisecond timescale. For each test point of 25 GBaud on-off keyed transmission, a minimum of 32 averaging runs are performed for different initial mesh and transmission matrix conditions to estimate the SNR and calculate the bit-error rates via enumeration of interference patterns. The mesh photodetectors are avalanche photodiodes (APDs) with parameters $R = 0.74$ A/W, $k_a = 0.18$, $I_{dark} = 40$ nA, low-gain bandwidth of 24 GHz, gain-bandwidth product of 290 GHz, and gain of 10. In-mesh photodetectors can operate at low speed and have transimpedance amplifier (TIA) noise levels of 2.9 pA/$\sqrt{\text{Hz}}$, whereas the output photodetectors are operating at higher speed and have TIA noise levels of 29 pA/$\sqrt{\text{Hz}}$. For output symbol detection, 3 dB of penalty is assumed due to equalization for modulator and receiver limitations. APD shot noise uses the Gaussian model, with signal-dependent variance. The mesh is assumed lossless excepting the 5% tap-off of optical power absorbed by in-mesh photodetectors. In curve fitting, best-fit sinusoidal curves with a fixed 2π period are found using power measurements at different phase values sampled from $[0, 2\pi]$, which are then used to compute the optimal phase shifts.

The most important finding of our analysis is that periodic equalization of MZ meshes using full-power training signals can be achieved on sub-millisecond speeds. Including a 10% or smaller duty cycle for the dedicated training signals, this leaves a two-to-three-order-of-magnitude speedup over previously proposed schemes that continually track the channel by using in-mesh photodetectors to detect low-power marker signals. For a more rigorous and detailed discussion about mesh adaptation strategies in DD MDM systems, the reader is referred to [50].

V. Conclusion

We have proposed a simplified channel model for describing the temporal dynamics of MDM systems that are subject to physical perturbations, such as fast vibrations or slow drifts caused by temperature variations. Our proposed model can describe polarization dynamics of SMF systems and can also be extended naturally to the case of multiple coupled modes, as in coupled MCFs or MMFs. We have also discussed adaptive equalization to track channel dynamics, examining canonical examples from long-haul coherent-detection systems and short-reach direct-detection systems.

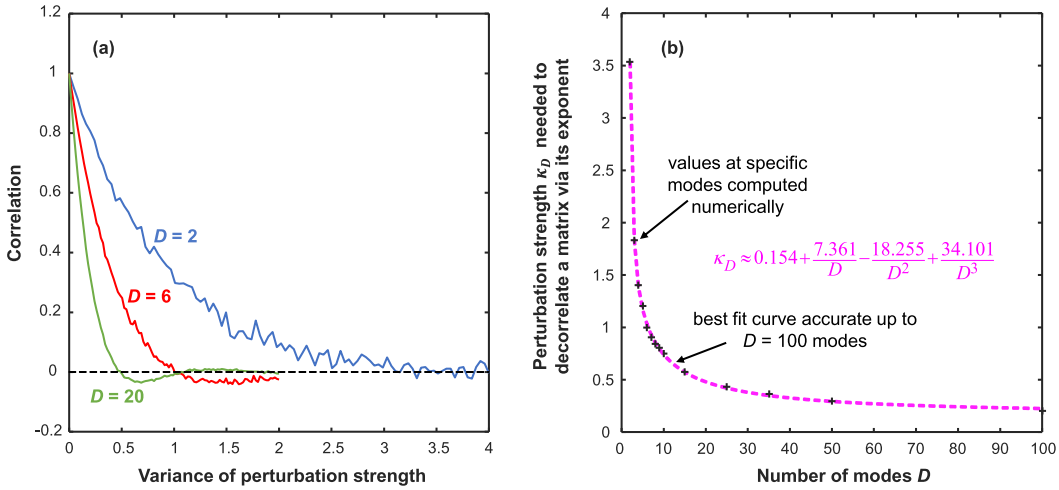

Fig. 13. (a) Decorrelating unitary matrices of size $D \times D$ by perturbing their skew-Hermitian exponents. Plot shows the correlation between unperturbed and perturbed matrices averaged across an ensemble of random unitary matrix realizations. (b) Perturbation strength κ_D required to decorrelate a unitary matrix via its exponent. The plus signs corresponds to values of κ_D that haven been computed numerically at specific values of D, whereas the dotted line is the best-fit curve to the data. The analytical expression for the best-fit curve is accurate up to $D = 100$ modes.

Appendix A
Smoothly Varying Unitary Matrices

In this Appendix, we present a method to generate arbitrary unitary matrices $\mathbf{A}(t)$ of size $D \times D$ that change smoothly on a predefined timescale $\tau_{\mathbf{A}}$. We define $\tau_{\mathbf{A}}$ as the smallest time such that $\mathbf{A}(t)$ and $\mathbf{A}(t+\tau_{\mathbf{A}})$ are uncorrelated:

$$\mathbb{E}\left[\mathbf{A}\left(t\right)\mathbf{A}(t+\tau_{\mathbf{A}})^{H}\right] = \mathbf{0}, \tag{15}$$

where $\mathbb{E}[\cdot]$ is the expectation operator over random unitary matrices $\mathbf{A}(t)$. Instead of working with (15) directly, we write $\mathbf{A}(t)$ as the matrix exponent of a time-varying skew-Hermitian matrix $\mathbf{A}_{sh}(t) = -\mathbf{A}_{sh}(t)^{H}$ of the same size [52]:

$$\mathbf{A}\left(t\right) = \exp\left(\mathbf{A}_{sh}\left(t\right)\right). \tag{16}$$

This result stems from group theory, which states that the special unitary group of degree D, $SU(D)$, is a Lie group of $D \times D$ unitary matrices. Any member of $SU(D)$ can be equivalently represented by the matrix exponential of some skew-Hermitian matrix. Since there is no restriction on the values $\mathbf{A}_{sh}(t)$ can assume to ensure $\mathbf{A}(t)$ is unitary (apart from the condition of skew-Hermiticity), the task of relating $\tau_{\mathbf{A}}$ to the elements of $\mathbf{A}(t)$ is equivalent to relating $\tau_{\mathbf{A}}$ to the perturbation of elements of $\mathbf{A}_{sh}(t)$. In other words, if we combine (15) and (16) as

$$\mathbb{E}\left[\exp\left(\mathbf{A}_{sh}\left(t\right)\right)\exp\left(\mathbf{A}_{sh}\left(t+\tau_{\mathbf{A}}\right)\right)^{H}\right] = \mathbb{E}\left[\exp\left(\mathbf{A}_{sh}\left(t\right)\right)\exp\left(\mathbf{A}_{sh}\left(t\right)+\Delta\mathbf{A}_{sh}\right)^{H}\right] = \mathbf{0}, \tag{17}$$

then any environmental perturbation that causes $\mathbf{A}_{sh}(t)$ to change smoothly by $\Delta\mathbf{A}_{sh}$ in a time $\tau_{\mathbf{A}}$ will also cause $\mathbf{A}(t)$ to decorrelate in the same time.

We continue by making a slight relaxation to (17). Instead of requiring all D^2 elements of the matrix product to equal zero, we use a weaker trace condition, similar to (4) in [53]

$$\mathbb{E}\left[\frac{1}{D}\mathrm{trace}\left(\exp\left(\mathbf{A}_{sh}\left(t\right)\right)\exp\left(\mathbf{A}_{sh}\left(t\right)+\Delta\mathbf{A}_{sh}\right)^{H}\right)\right] = 0. \tag{18}$$

This scalar constraint is easier to implement in simulations than a matrix-valued constraint. Equation (18) can be understood intuitively because if $\Delta\mathbf{A}_{sh} = \mathbf{0}$, then the properties of unitary matrices imply that $\exp(\mathbf{A}_{sh}(t))\exp\left(\mathbf{A}_{sh}(t)\right)^{H} = \mathbf{A}(t)\mathbf{A}(t)^{H} = \mathbf{I}$. Once the elements in $\Delta\mathbf{A}_{sh}$ assume non-negligible values after time $\tau_{\mathbf{A}}$, then $\mathbf{A}(t)$ and $\mathbf{A}(t+\tau_{\mathbf{A}})$ behave like independent matrices, causing the trace to tend to zero.

To determine the required magnitude of the elements of $\Delta\mathbf{A}_{sh}$, we consider an ensemble average

$$\mathbb{E}\left[\frac{1}{D}\mathrm{trace}\left(\exp\left(\mathbf{A}_{sh}\right)\times\ \ldots\right.\right.$$
$$\left.\left.\exp\left(\mathbf{A}_{sh}+\sqrt{\kappa_D}\Delta\mathbf{A}_{sh}\right)^{H}\right)\right] = 0, \tag{19}$$

where the upper triangular elements of $\mathbf{A}_{sh}$ are chosen according to a multivariate Gaussian distribution, $\mathrm{Re}\mathbf{A}_{sh}$, $\mathrm{Im}\mathbf{A}_{sh}$, $\mathrm{Re}\Delta\mathbf{A}_{sh}$, $\mathrm{Im}\ \Delta\mathbf{A}_{sh} \sim \mathcal{N}(\mathbf{0},\mathbf{I})$, and the conditions $\mathbf{A}_{sh} = -\mathbf{A}_{sh}^{H}$, $\Delta\mathbf{A}_{sh} = -\Delta\mathbf{A}_{sh}^{H}$ are numerically enforced. The ensemble is averaged across random realizations of both $\mathbf{A}_{sh}$ and $\Delta\mathbf{A}_{sh}$ using Monte-Carlo simulations. The parameter κ_D represents the variance of each perturbing random matrix element $\Delta\mathbf{A}_{sh,ij}$ that causes $\mathbf{A}(t)$ to be decorrelated. Fig. 13(b) shows values of κ_D that have been numerically computed for several values of D from 2 to 100. A best-fit curve for the data is found to be

$$\kappa_D \approx 0.154 + \frac{7.361}{D} - \frac{18.255}{D^2} + \frac{34.101}{D^3}, \tag{20}$$

which is accurate even for large D. It is interesting to note that (20) implies the amount of perturbation strength required to decorrelate a unitary matrix decreases as the size of the matrix increases. This result agrees qualitatively with experimental

observations of MDM channel dynamics, where the generalized Jones matrix of MMF was shown to be more sensitive to external perturbations than that of SMF [8].

Therefore, to generate a $D \times D$ unitary matrix $\mathbf{A}(t) = \exp(\mathbf{A}_{sh}(t))$ that changes at a desired timescale $\tau_{\mathbf{A}}$, we need to model the smooth temporal evolution of a perturbative skew-Hermitian matrix which changes by an amount $\Delta\mathbf{A}_{sh} \sim \mathcal{N}(\mathbf{0}, \kappa_D \mathbf{I})$ on a timescale $\tau_{\mathbf{A}}$. There is considerable flexibility in choosing the exact form of the temporal evolution. In this paper, we have chosen to model the evolutions as time-domain ramps and random Wiener processes, for reasons explained in Sections II-B and II-C.

REFERENCES

[1] P. M. Krummrich, E.-D. Schmidt, W. Weiershausen, and A. Mattheus, "Field trial results on statistics of fast polarization changes in long haul WDM transmission systems," presented at the *Opt. Fiber Commun. Conf. Expo. Nat. Fiber Opt. Eng. Conf.*, Anaheim, CA, USA, 2005, Paper OThT6.

[2] P. M. Krummrich and K. Kotten, "Extremely fast (microsecond timescale) polarization changes in high speed long haul WDM transmission systems," presented at the *Opt. Fiber Commun. Conf.*, Los Angeles, CA, USA, 2004, Paper FI3.

[3] K. Roberts *et al.*, "Performance of dual-polarization QPSK for optical transport systems," *J. Lightw. Technol.*, vol. 27, no. 16, pp. 3546–3559, Aug. 2009.

[4] M. Boroditsky *et al.*, "Polarization dynamics in installed fiber optic systems," in *Proc. IEEE LEOS Annu. Meet. Conf.*, 2005, pp. 414–415.

[5] M. Brodsky *et al.*, "Physical mechanism for polarization mode dispersion temporal dynamics," *IEEE LEOS Newslett.*, vol. 18, pp. 4–6, Jun. 2004, [Online]. Available :https://www.eng.tau.ac.il/∼tur/pdfs/124.pdf

[6] K. Ogaki, M. Nakada, Y. Nagao, and K. Nishijima, "Fluctuation differences in the principal states of polarization in aerial and buried cables," presented at the *Opt. Fiber Commun. Conf.*, Atlanta, GA, USA, 2003, Paper MF13.

[7] X. Chen *et al.*, "Characterization and analysis of few-mode fiber channel dynamics," *IEEE Photon. Technol. Lett.*, vol. 25, no. 18, pp. 1819–1822, Sep. 2013.

[8] K. Choutagunta *et al.*, "Modal dynamics in spatially multiplexed links," presented at the *Opt. Fiber Commun. Conf.*, San Diego, CA, USA, 2019, Paper W4C.1.

[9] J. P. Gordon and H. Kogelnik, "PMD fundamentals: Polarization mode dispersion in optical fibers," *Proc. Nat. Acad. Sci.*, vol. 97, pp. 4541–4550, 2000.

[10] J. Schuster, Z. Marzec, W. L. Kath, and G. Biondini, "Hybrid hinge model for polarization-mode dispersion in installed fiber transmission systems," *J. Lightw. Technol.*, vol. 32, no. 7, pp. 1412–1419, Apr. 2014.

[11] J. Li, G. Biondini, W. L. Kath, and H. Kogelnik, "Outage statistics in a waveplate hinge model of polarization-mode dispersion," *J. Lightw. Technol.*, vol. 28, no. 13, pp. 1958–1968, Jul. 2010.

[12] J. Li, G. Biondini, W. L. Kath, and H. Kogelnik, "Anisotropic hinge model for polarization-mode dispersion in installed fibers," *Opt. Lett.*, vol. 33, pp. 1924–1926, 2008.

[13] C. Antonelli and A. Mecozzi, "Theoretical characterization and system impact of the hinge model of PMD," *J. Lightw. Technol.*, vol. 24, no. 11, pp. 4064–4074, Nov. 2006.

[14] D. S. Waddy, L. Chen, and X. Bao, "Theoretical and experimental study of the dynamics of polarization-mode dispersion," *IEEE Photon. Technol. Lett.*, vol. 14, no. 4, pp. 468–470, Apr. 2002.

[15] A. Mecozzi, C. Antonelli, M. Boroditsky, and M. Brodsky, "Characterization of the time dependence of polarization mode dispersion," *Opt. Lett.*, vol. 29, pp. 2599–2601, 2004.

[16] C. B. Czegledi, M. Karlsson, E. Agrell, and P. Johannisson, "Polarization drift channel model for coherent fibre-optic systems," *Sci. Rep.*, vol. 6, 2016, Art. no. 21217.

[17] M. Brodsky, N. J. Frigo, M. Boroditsky, and M. Tur, "Polarization mode dispersion of installed fibers," *J. Lightw. Technol.*, vol. 24, no. 12, pp. 4584–4599, Dec. 2006.

[18] C. Antonelli, A. Mecozzi, M. Shtaif, and P. J. Winzer, "Stokes-space analysis of modal dispersion in fibers with multiple mode transmission," *Opt. Express*, vol. 20, pp. 11718–11733, 2012.

[19] T. M. F. Alves and A. V. T. Cartaxo, "Intercore crosstalk in homogeneous multicore fibers: Theoretical characterization of stochastic time evolution," *J. Lightw. Technol.*, vol. 35, no. 21, pp. 4613–4623, Nov. 2017.

[20] R. S. Luís *et al.*, "Time and modulation frequency dependence of crosstalk in homogeneous multi-core fibers," *J. Lightw. Technol.*, vol. 34, no. 2, pp. 441–447, Jan. 2016.

[21] K. Choutagunta and J. M. Kahn, "Dynamic channel modeling for mode-division multiplexing," *J. Lightw. Technol.*, vol. 35, no. 12, pp. 2451–2463, Jun. 2017.

[22] S. Galli, "A simplified model for the indoor power line channel," in *Proc. IEEE Int. Symp. Power Line Commun. Appl.*, Mar. 29–Apr. 1, 2009, pp. 13–19.

[23] H. Hashemi, "The indoor radio propagation channel," *Proc. IEEE*, vol. 81, no. 7, pp. 943–968, Jul. 1993.

[24] J. B. Carruthers and J. M. Kahn, "Modeling of nondirected wireless infrared channels," *IEEE Trans. Commun.*, vol. 45, no. 10, pp. 1260–1268, Oct. 1997.

[25] C. Antonelli, C. Colamarino, A. Mecozzi, and M. Brodsky, "A model for temporal evolution of PMD," *IEEE Photon. Technol. Lett.*, vol. 20, no. 12, pp. 1012–1014, Jun. 2008.

[26] A. Mecozzi, C. Antonelli, and M. Brodsky, "Statistics of the polarization mode dispersion dynamics," *Opt. Lett.*, vol. 32, pp. 3032–3034, 2007.

[27] 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.

[28] T. Hayashi *et al.*, "Coupled-core multi-core fibers: High-spatial-density optical transmission fibers with low differential modal properties," in *Proc. Eur. Conf. Opt. Commun.*, Sep. 27–Oct. 1, 2015, pp. 1–3.

[29] R. Olshansky, "Mode coupling effects in graded-index optical fibers," *Appl. Opt.*, vol. 14, pp. 935–945, 1975.

[30] J. Wuttke, P. M. Krummrich, and J. Rosch, "Polarization oscillations in aerial fiber caused by wind and power-line current," *IEEE Photon. Technol. Lett.*, vol. 15, no. 6, pp. 882–884, Jun. 2003.

[31] S. M. Pietralunga, J. Colombelli, A. Fellegara, and M. Martinelli, "Fast polarization effects in optical aerial cables caused by lightning and impulse current," *IEEE Photon. Technol. Lett.*, vol. 16, no. 11, pp. 2583–2585, Nov. 2004.

[32] M. Kurono, K. Isawa, and M. Kuribara, "Transient state of polarization in optical ground wire caused by lightning and impulse current," *Proc. SPIE*, vol. 12873, pp. 242–245, 1996.

[33] M. Brodsky, P. Magill, and N. J. Frigo, "Polarization-mode dispersion of installed recent vintage fiber as a parametric function of temperature," *IEEE Photon. Technol. Lett.*, vol. 16, no. 1, pp. 209–211, Jan. 2004.

[34] T. G. Gutowski and C. L. Dym, "Propagation of ground vibration: A review," *J. Sound Vib.*, vol. 49, pp. 179–193, 1976.

[35] Corning SMF-28 Optical Fiber, Datasheet, Corning Inc., Corning, NY, USA, 2002.

[36] R. Romaniuk, "Tensile strength of tailored optical fibers," *Opto-Electron. Rev.*, vol. 8, pp. 101–116, 2000.

[37] D. D. Falconer and S. L. Ariyavisitakul, "Broadband wireless using single carrier and frequency domain equalization," in *Proc. 5th Int. Symp. Wireless Pers. Multimedia Commun.*, Oct. 27–30, 2002, vol. 1, pp. 27–36.

[38] S. Haykin, *Adaptive Filter Theory*. Upper Saddle River, NJ, USA: Prentice-Hall, 2001.

[39] S. Ö. Arık, D. Askarov, and J. M. Kahn, "Adaptive frequency-domain equalization in mode-division multiplexing systems," *J. Lightw. Technol.*, vol. 32, no. 10, pp. 1841–1852, May 2014.

[40] Y. Jian, J. Werner, and G. A. Dumont, "The multimodulus blind equalization and its generalized algorithms," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 5, pp. 997–1015, Jun. 2002.

[41] Q. Hu, X. Chen, A. Li, and W. Shieh, "High-dimensional stokes-space analysis for monitoring fast change of mode dispersion in few-mode fibers," presented at the *Opt. Fiber Commun. Conf.*, San Francisco, CA, USA, Mar. 9–13, 2014, Paper W3D.3.

[42] H. Liu *et al.*, "Demonstration of stable 3 × 10 Gb/s mode group-multiplexed transmission over a 20 km few-mode fiber," presented at the *Opt. Fiber Commun. Conf.*, San Francisco, CA, USA, 2018, Paper W4J.2.

[43] S. Ö. Arik and J. M. Kahn, Direct-detection mode-division multiplexing in modal basis using phase retrieval," *Opt. Lett.*, vol. 41, pp. 4265–4268, 2016.

[44] D. A. B. Miller, "Self-aligning universal beam coupler," *Opt. Express*, vol. 21, pp. 6360–6370, 2013.

[45] F. Morichetti *et al.*, "4-channel all-optical MIMO demultiplexing on a silicon chip," presented at the *Opt. Fiber Commun. Conf. Exhib.*, Anaheim, CA, USA, Mar. 20–24, 2016, Paper Th3E.7.

[46] N. Barre *et al.*, "Broadband, mode-selective 15-mode multiplexer based on multi-plane light conversion," presented at the *Opt. Fiber Commun. Conf. Exhib.*, Los Angeles, CA, USA, Mar. 19–23, 2017, Paper Th2A.7.
[47] G. Labroille *et al.*, "Efficient and mode selective spatial mode multiplexer based on multi-plane light conversion," in *Proc. IEEE Photon. Conf.*, Oct. 12–16, 2014, pp. 518–519.
[48] R. Tang, T. Tanemura, and Y. Nakano, "Integrated reconfigurable unitary optical mode converter using MMI couplers," *IEEE Photon. Technol. Lett.*, vol. 29, no. 12, pp. 971–974, Jun. 2017.
[49] D. A. B. Miller, "Self-configuring universal linear optical component [Invited]," *Photon. Res.*, vol. 1, pp. 1–15, 2013.
[50] I. Roberts and J. M. Kahn, "Tracking dynamic mode-mixing in an intensity-modulated communication system using a Mach-Zehnder mesh for channel equalization," *IEEE J. Lightw. Technol.*, submitted for publication.
[51] A. Annoni *et al.*, "Unscrambling light–Automatically undoing strong mixing between modes," *Light, Sci. Appl.*, vol. 6, 2017, Art. no. e17110.
[52] T. A. Loring, "Computing a logarithm of a unitary matrix with general spectrum," arXiv:1203.6151, 2012.
[53] A. Mecozzi, C. Antonelli, and M. Shtaif, "Intensity impulse response of SDM links," *Opt. Express*, vol. 23, pp. 5738–5743, 2015.

Karthik Choutagunta (S'11) received the B.S. degree from The University of Texas at Austin, Austin, TX, USA, in 2014, and the M.S. degree in electrical engineering from Stanford University, Stanford, CA, USA, in 2016, where he is currently working toward the Ph.D. degree. He was with Coherent Logix, Austin, in Summer 2012, with Intel Corp., Austin, from 2012 to 2014, with Apple, Inc., Cupertino, CA, in Fall 2015 and Summer 2016, respectively, with Facebook, Inc., Menlo Park, CA, in Summer 2017, and with Nokia Bell Laboratories, Crawford Hill Laboratory, Holmdel, NJ, USA, in Summer 2018. His main current research interests include various topics in communications, signal processing, machine learning, and optimization.

Ian Roberts received the B.S. degree in electrical engineering from the University of Ottawa, Ottawa, ON, Canada, and the M.S. and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA, USA. He was with Ciena on DSP hardware development for WaveLogic 3 Extreme prior to beginning graduate studies. His research interests include optical communication, optimization, and compensation of nonlinear propagation.

Joseph M. Kahn (M'90–SM'98–F'00) received the A.B., M.A., and Ph.D. degrees in physics from UC Berkeley, Berkeley, CA, USA, in 1981, 1983, and 1986, respectively. From 1987 to 1990, he was with AT&T Bell Laboratories, Crawford Hill Laboratory, Holmdel, NJ, USA. He demonstrated multi-Gb/s coherent optical fiber transmission systems, setting world records for receiver sensitivity. From 1990 to 2003, he was with the Faculty of the Department of Electrical Engineering and Computer Sciences, UC Berkeley, performing research on optical and wireless communications. Since 2003, he has been a Professor of electrical engineering with Stanford University, Stanford, CA, 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. From 1993 to 2000, he was a 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. In 2000, he helped found StrataLight Communications, where he served as the Chief Scientist from 2000 to 2003. StrataLight was acquired by Opnext, Inc., in 2009. He received the National Science Foundation Presidential Young Investigator Award in 1991.