Adaptive Compensation of Multimode Fiber Dispersion by Control of Launched Amplitude, Phase, and Polarization

Mahdieh B. Shemirani, Jeffrey P. Wilde, *Member, IEEE*, and Joseph M. Kahn, *Fellow, IEEE*

Abstract—In previous work, we studied the compensation of modal dispersion in multimode fiber (MMF) using several different configurations of optical systems that can control the amplitude, phase and polarization of the launched field. In that work, we assumed knowledge of a fiber's principal modes (PMs) and their group delays (GDs), enabling us to compute the optimal settings of the optical system. In practice, however, we do not have prior knowledge of the PMs and their GDs. In this paper, for three of the configurations, we propose algorithms for setting the optical system adaptively, based upon measurements of the eye opening. We present simulations showing that in the absence of noise, the performance of the adaptive solution approaches that of the optimal solution, and characterizing the algorithms' convergence speed and tolerance to noise. We present experiments using a particular configuration and adaptive algorithm, demonstrating their effectiveness in 10-Gb/s transmission through up to 2000 m of 50-μm-core graded-index MMF.

Index Terms—Adaptive algorithms, adaptive optics, optical fiber communication, optical fiber dispersion, optimization methods, quadratic programming.

I. INTRODUCTION

MULTIMODE FIBER (MMF) is widely used for data transmission in local-area networks. Modal dispersion causes intersymbol interference (ISI), which can limit the achievable bit rate-distance product [1]. Modal dispersion arises because different modes propagate at different group velocities. Imperfections in a MMF, such as index inhomogeneity, core ellipticity and eccentricity, and bends, introduce coupling between modes. Because of mode coupling, even if a light pulse is launched into a single mode, it tends to couple to other modes, leading to a superposition of several pulses at the output of the MMF.

Electrical equalization is commonly used in high-bit-rate systems to mitigate ISI caused by modal dispersion [2], [3]. While electrical equalization can extend the bit rate-distance product, it is ultimately limited by noise enhancement [4].

It has been shown that even in the presence of mode coupling, there exists a complete set of orthonormal modes, called principal modes (PMs), such that a pulse launched into a PM at the input of a MMF arrives as a single pulse at the output [5]. The PM field patterns and their group delays (GDs) depend on mode coupling, and may change over time as the mode coupling is altered by temperature changes, mechanical vibrations and other perturbations of the MMF.

The use of adaptive optics to control modal dispersion without noise enhancement was proposed in [6] and demonstrated in [7]. In this technique, a spatial light modulator (SLM) is used to control the electric field launched into the MMF, thereby controlling the PMs excited, and thus, the MMF impulse response. Optimized compensation of modal dispersion by adaptive optics was studied in [8], where it was shown that, assuming the PM field patterns and GDs are known, the optimal SLM settings can be obtained by solving an equivalent convex optimization problem, which is a second-order cone program (SOCP) [8], [9]. In practice, we do not have *a priori* knowledge of fiber PMs and their GDs. Hence, we need an adaptive algorithm to adjust the SLM. Adaptive algorithms for optimizing the SLM in real time were proposed in [10]. Amplitude-and-phase sequential coordinate ascent (APSCA), which optimizes amplitude and phase sequentially block-by-block, was shown to converge to the global optimum in the absence of noise. In addition, a suboptimal algorithm, continuous-phase sequential coordinate ascent (CPSCA), which optimizes only a continuous-valued phase block-by-block, was shown to exhibit faster convergence toward the global optimum, while using a simpler SLM. The CPSCA algorithm was demonstrated experimentally in [11].

Experiments have demonstrated that the intensity impulse response of a MMF can depend strongly on launched polarization [7]. We explained this polarization dependence by observing that perturbations in a MMF cause both spatial- and polarization-mode coupling, and modeled these effects numerically in [12]. In [13], we extended the adaptive optical compensation of modal dispersion to include control of launched polarization. We considered five different adaptive optics configurations, permitting various degrees of control of amplitude, phase and polarization, assumed piecewise constant over discrete blocks. Assuming the fiber's PMs and their GDs are known, we showed how to optimize the variables in each of the five configurations. Using a model for propagation in MMF including the effects of both spatial- and polarization-mode coupling [12], we compared the performance of the five configurations. We found that the best performance is achieved by two configurations permitting independent block-wise control of amplitude, phase, and polarization.
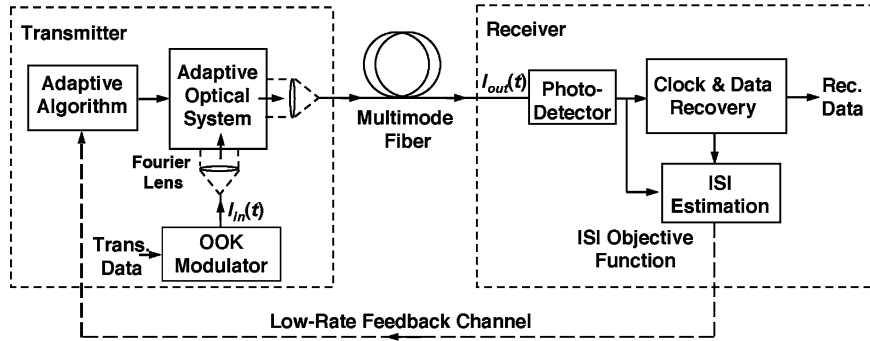
Fig. 1. Adaptive transmission system. A modulated optical signal is passed into an adaptive optical system, whose output is launched into a MMF. At the MMF output, residual ISI is estimated, and the estimate is passed to an adaptive algorithm that controls the adaptive optical system.

In this paper, we extend the work of [13] to the situation in which the fiber's PMs and their GDs are not known *a priori*. For three of the five adaptive optics configuration considered in [13], we develop adaptive algorithms for adjustment of the amplitude, phase, and polarization. We also develop a range of suboptimal algorithms that trade off suboptimality of the final solution for speed of convergence. Finally, we demonstrate experimentally one of the algorithms, which uses one SLM and one polarization controller (PC), showing that it enables transmission at 10 Gb/s through up to 2000 m of graded-index MMF.

The remainder of this paper is organized as follows. In Section II, we describe the adaptive transmission scheme and several different adaptive optics configurations permitting control of launched amplitude, phase and polarization. In Section III, we describe adaptive algorithms for three configurations. In Section IV, we present numerical simulations comparing the performance, convergence speed and noise resilience of the algorithms in fibers with spatial- and polarization-mode coupling. In Section V, we present experimental results obtained for one of the algorithms. We present conclusions in Section VI.

II. TRANSMISSION SYSTEM

Fig. 1 shows the adaptive transmission first described in [6], [7], and later extended in [13] to include polarization control. At the transmitter, a modulated optical signal is collimated by a lens and the beam is input to an adaptive optics system, which comprises some combination of SLMs, fixed or variable beam splitters (BSs), PCs, and other components. The beam is partitioned into a set of disjoint blocks, and the adaptive optics provides block-wise control of its amplitude, phase, and polarization. The output beam is focused by a second lens and launched into a MMF. At the receiver, after signal detection, the residual ISI is quantified in terms of an objective function, which may be fed back to the transmitter to aid in adjustment of the adaptive optics system.

In [13] we proposed five different configurations for the adaptive optics system, which are shown in Fig. 2. For each configuration, assuming prior knowledge of a fiber's PMs and GDs, we showed how to obtain optimal settings of the adaptive optics system by transforming the optimization into a previously solved problem called spatial light modulator optimization (SLMO), which was solved in [8].

Configuration A, shown in Fig. 2(a), employs one SLM providing independent, block-wise control of amplitude and phase, followed by a PC, which uniformly transforms the light reflected from the SLM from a y-polarization state to an arbitrary elliptical state. Configuration B, shown in Fig. 2(b), employs one SLM that can provide independent, block-wise control of amplitude, phase, and polarization. Hence, there is no need for a PC. The remaining three configurations, shown in Fig. 2(c), (d), and (e), utilize two SLMs, $\mathrm{SLM_x}$ and $\mathrm{SLM_y}$, which provide block-wise control of amplitude and phase in the x and y polarizations, respectively. Each of the three configurations uses a half-wave plate to transform y-polarized light to x-polarized light. Depending on implementation, the half-wave plate may be placed before or after $\mathrm{SLM_x}$ (we assume the former). The configurations C, D, and E differ in the way that amplitude is allocated between two polarizations. Configuration C, shown in Fig. 2(c), uses a uniform fixed BS to allocate equal amplitudes to the x and y polarizations in each block. Configuration D, shown in Fig. 2(d), uses a uniform variable BS. Configuration E, shown in Fig. 2(e), uses a variable BS to vary the amplitude allocation between the x and y polarizations independently in each block. We proved in [13] that configurations B and E are equivalent and yield the same optimal solution, and only implementation considerations would favor one over the other.

Following [7], [8], [13], we assume the SLM is a reflective device partitioned into N disjoint blocks (typically squares) in two dimensions. We define an indicator function for the jth block

$$s_j(x,y) = \begin{cases} 1, & (x,y) \text{ in the interior of the } j\text{th block} \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

Obviously, the $\{s_j(x,y), j = 1,\ldots,N\}$ form an orthogonal set. The focusing function performed by the second lens shown in Fig. 1 is described by a linear spatial operator L (often well-approximated by an appropriately scaled spatial Fourier transform). The field launched into the MMF can be expressed as

$$\vec{A}_{\text{launched}}(x,y) = L\left[\sum_{j=1}^{N} u_j A_0(x,y)s_j(x,y)\,\widehat{x}\right.$$
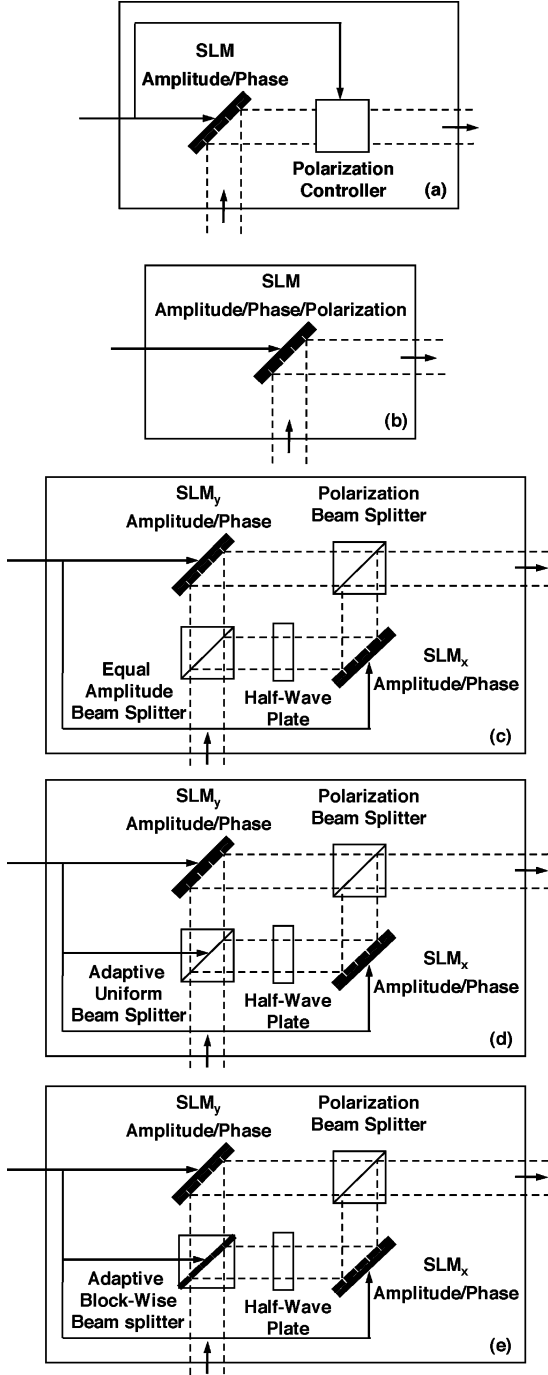$$\left. + v_j A_0(x,y)s_j(x,y)\,\widehat{y}\right]. \tag{2}$$

Fig. 2. Adaptive optical system configurations. (a) Configuration A comprises one SLM providing block-wise control of amplitude and phase, and a polarization controller to control the overall polarization. (b) Configuration B comprises one SLM providing block-wise control of amplitude, phase and polarization. (c) Configuration C comprises two SLMs providing block-wise control of amplitude and phase in the x and y polarizations, respectively. A uniform fixed beam splitter allocates equal amplitudes to the x and y polarizations. (d) Configuration D is the same as configuration C, except that a uniform variable beam splitter allocates variable amplitudes to the x and y polarizations. (e) Configuration E is the same as configuration C, except that a variable beam splitter provides block-wise amplitude allocation to the x and y polarizations.

Here, $A_0(x, y)$ is the complex optical field distribution incident on the SLM. The coefficients u_j and v_j describe the complex reflectance of the j th SLM block in the x and y polarizations, re-

spectively. Defining the contribution to the launched field from the jth block

$$k_j(x, y) = L[A_0(x, y)s_j(x, y)] \tag{3}$$

we can rewrite (2) as

$$\vec{A}_{\text{launched}}(x, y) = \sum_{j=1}^{N} u_j k_j(x, y) \, \widehat{x} + v_j k_j(x, y) \, \widehat{y}. \tag{4}$$

For a MMF with M modes in each polarization, there exist $2M$ first-order PMs, which have field patterns $\vec{\Phi}_i(x, y) = \Phi_{ix}(x, y)\hat{x} + \Phi_{iy}(x, y)\hat{y}$ and GDs $\tau_i, i = 1, \ldots, 2M$. For the remainder of this section, we assume the PM field patterns and GDs are known. In order to find the first-order impulse response, we project $\vec{A}_{\text{launched}}(x, y)$ into the PMs [10]. The ith PM is excited with amplitude

$$\mu_i = \xi \vec{\Phi}_i(x, y) \cdot \vec{A}_{\text{launched}}(x, y)$$

$$= \xi \left[\Phi_{ix}(x, y) \cdot \sum_{j=1}^{N} u_j k_j(x, y) \right.$$

$$\left. + \Phi_{iy}(x, y) \cdot \sum_{j=1}^{N} v_j k_j(x, y) \right] \tag{5}$$

where the dot product indicates an overlap integral over the cross section of the fiber and ξ is a normalization factor that will be found below. The first-order impulse response is

$$h(t) = \sum_{i=1}^{2M} |\mu_i|^2 \delta(t - \tau_i). \tag{6}$$

The $k_j(x, y), \Phi_{ix}(x, y)$ and $\Phi_{iy}(x, y)$ are defined over the (x, y) plane at the input to the MMF. To simplify the problem, we sample these functions over the (x, y) plane and express the samples as $L \times 1$ vectors[1] $k_j, \Phi_{ix}, \Phi_{iy}$. We define the vectors

$$a_i = \begin{bmatrix} k_1^H \Phi_{ix} \\ \vdots \\ k_N^H \Phi_{ix} \end{bmatrix} \quad b_i = \begin{bmatrix} k_1^H \Phi_{iy} \\ \vdots \\ k_N^H \Phi_{iy} \end{bmatrix} \tag{7}$$

and also the vectors

$$u = \begin{bmatrix} u_1 \\ \vdots \\ u_N \end{bmatrix} \quad v = \begin{bmatrix} v_1 \\ \vdots \\ v_N \end{bmatrix}. \tag{8}$$

The impulse response can be expressed as

$$h(t) = \sum_{i=1}^{2M} |\xi|^2 \left| a_i^H u + b_i^H v \right|^2 \delta(t - \tau_i) \tag{9}$$

where the superscript ^{H} denotes Hermitian conjugate. Expanding the square term in (9), we can express the impulse response as given by (10), shown at the bottom of the next page.

---

[1]These functions are sampled over a two-dimensional grid. The number of samples should be sufficient to capture the spatial variations of the highest-order mode that can propagate. The samples are rearranged into $L \times 1$ column vectors.

We assume the transmitted signal is modulated by on-off keying with bit interval T and bit rate $1/T$. Assume that an isolated 1 bit is transmitted, described by an input intensity waveform $p(t)$, and assume the receiver has an impulse response $r(t)$. The receiver output waveform is $g(t) = p(t) * h(t) * r(t)$, which is a continuous-time system impulse response. Assume the receiver output is sampled with a timing offset t_0. The effect of ISI on receiver performance is characterized fully by a discrete-time system impulse response $g(nT; t_0) = g(t)|_{t=nT}, n = \ldots, -1, 0, 1, \ldots$.

We define an objective function quantifying ISI [15]

$$f = g(0; t_0) - \sum_{n \neq 0} g(nT; t_0). \tag{11}$$

f is the "eye opening", with $f < 0$ when the eye is closed, and $f > 0$ when the eye is open. At high signal-to-noise ratio, the bit-error ratio (BER) depends on $g(nT; t_0)$ only through f [15]. Defining $q(t) = p(t) * r(t)$, we can write f as

$$f = [u^H \quad v^H] Q \begin{bmatrix} u \\ v \end{bmatrix} \tag{12}$$

where Q is defined in (13), shown at the bottom of this page. Q is a $2N \times 2N$ Hermitian matrix, and, therefore, has N real eigenvalues. For a lossless adaptive optics system, we want the total launched power to be $S_{\text{launched}} = \sum_{i=1}^{2M} |\mu_i|^2 = 1$, which yields the normalization factor $|\xi|^2$ as [13]

$$|\xi| = \|H^H K 1_{N \times 1}\|^{-1} \tag{14}$$

where $K = [k_1 \quad \cdots \quad k_N]$ is a $L \times N$ matrix, H is a $L \times 2M$ matrix of samples of the ideal modes[2], and $1_{N \times 1} = [1 \quad \cdots \quad 1]^T$. The problem we intend to solve is

$$\text{maximize } [u^H \quad v^H] Q \begin{bmatrix} u \\ v \end{bmatrix} \tag{15}$$

subject to constraints imposed by problem configuration.

III. ADAPTIVE ALGORITHMS

In practice, we do not have prior knowledge of a fiber's PM field patterns and their GDs, which is equivalent to saying we

[2]These are the eigenmodes in the absence of mode coupling. For example, if the MMF has a parabolic index profile, these are Hermite-Gaussian modes.

do not have prior knowledge of the matrix Q. We note, however, from (12) that the objective function (eye opening) is a quadratic function of the variables of the adaptive optics system. Hence, we can approach the optimal solution by using an adaptive algorithm that adjusts these variables to maximize the eye opening. Reference [10] described the APSCA and CPSCA algorithms for systems not using polarization control. Here, we generalize them to systems controlling the amplitude, phase, and polarization of the launched field. As shown in [13], configurations B and E yield the same optimized performance, and configuration C is a special case of configuration D. Hence, here, we study adaptive algorithms only for configurations A, B, and D. In the next three subsections, we derive algorithms for these three configurations in the absence of noise. Then, in a following subsection, we describe their use in the presence of noise.

A. One Amplitude/Phase SLM and Uniform Polarization Controller

Configuration A, shown in Fig. 2(a), comprises one SLM and one PC. As in [13], the block-wise SLM variables are described by

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \gamma_x \vartheta \\ \gamma_y \vartheta \end{bmatrix} \tag{16}$$

where ϑ is the complex vector of SLM block reflectances, and γ_x and γ_y are complex variables representing a uniform elliptical polarization, with constraint $|\gamma_x|^2 + |\gamma_y|^2 = 1$. In [13], we showed that for configuration A, the problem of optimizing over phase, amplitude, and polarization of the launched field is not convex. However, using an alternating optimization method, which alternately optimizes over SLM block reflectances and polarization, we can cast the problem as two separate convex problems. Consider $Q = \begin{bmatrix} Q_1 & Q_2 \\ Q_2^H & Q_3 \end{bmatrix}$ (each of the submatrices is a $N \times N$ matrix). For a fixed polarization, the optimization problem is [13]

$$\begin{aligned} \text{maximize } & f_A|_{\gamma_x, \gamma_y \text{constant}} \\ &= \vartheta^H \Psi \vartheta \\ \text{subject to } & |\vartheta_i|^2 \leq 1 \quad \forall i = 1, \ldots, N \end{aligned} \tag{17}$$

$$h(t) = [u^H \quad v^H] \begin{bmatrix} \sum_{i=1}^{2M} |\xi|^2 a_i a_i^H \delta(t - \tau_i) & \sum_{i=1}^{2M} |\xi|^2 a_i b_i^H \delta(t - \tau_i) \\ \sum_{i=1}^{2M} |\xi|^2 b_i a_i^H \delta(t - \tau_i) & \sum_{i=M+1}^{2M} |\xi|^2 b_i b_i^H \delta(t - \tau_i) \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} \tag{10}$$

$$Q = |\xi|^2 \begin{bmatrix} \sum_{i=1}^{2M} a_i a_i^H q(t_0 - \tau_i) & \sum_{i=1}^{2M} a_i b_i^H q(t_0 - \tau_i) \\ \sum_{i=1}^{2M} b_i a_i^H q(t_0 - \tau_i) & \sum_{i=M+1}^{2M} b_i b_i^H q(t_0 - \tau_i) \end{bmatrix}$$
$$- \sum_{n \neq 0} |\xi|^2 \begin{bmatrix} \sum_{i=1}^{2M} a_i a_i^H q(t_0 + nT - \tau_i) & \sum_{i=1}^{2M} a_i b_i^H q(t_0 + nT - \tau_i) \\ \sum_{i=1}^{2M} b_i a_i^H q(t_0 + nT - \tau_i) & \sum_{i=M+1}^{2M} b_i b_i^H q(t_0 + nT - \tau_i) \end{bmatrix} \tag{13}$$

TABLE I
JONES PARAMETERS FOR FOUR POINTS ON THE POINCARÉ SPHERE WITH
MAXIMAL SEPARATION, LYING ON THE CORNERS OF A TETRAHEDRON

γ_x	γ_y
$0.6739 - 0.2142\,i$	$0.6739 + 0.2142\,i$
$0.2142 - 0.6739\,i$	$0.2142 + 0.6739\,i$
$-0.6739 - 0.6739\,i$	$-0.2142 + 0.2142\,i$
$-0.2142 - 0.2142\,i$	$-0.6739 + 0.6739\,i$

where $\Psi = |\gamma_x|^2 Q_1 + \gamma_y^* \gamma_x Q_2^H + \gamma_x^* \gamma_y Q_2 + |\gamma_y|^2 Q_3$. For fixed SLM reflectances, the optimization problem is [13]

$$\text{maximize } f_A|_{\vartheta\text{constant}} == [\gamma_x^* \quad \gamma_y^*] J_A \begin{bmatrix} \gamma_x \\ \gamma_y \end{bmatrix}$$
$$\text{subject to } |\gamma_x|^2 + |\gamma_y|^2 = 1 \tag{18}$$

where J_A is a 2×2 Jones matrix, given by

$$J_A = \begin{bmatrix} \vartheta^H Q_1 \vartheta & \vartheta^H Q_2 \vartheta \\ \vartheta^H Q_2^H \vartheta & \vartheta^H Q_3 \vartheta \end{bmatrix}.$$

Considering that Ψ has exactly one positive eigenvalue and $N - 1$ non-positive eigenvalues [13], the first problem is an SLMO problem, for which adaptive algorithms were derived in [10]. Here, we provide a simple algorithm to solve the optimization in (18). Once the four independent real variables in J_A are calculated, we can perform an eigenvalue decomposition on J_A to find the best launched polarization, which maximizes $f_A|_{\vartheta\text{constant}}$. In order to perform these measurements optimally, we choose four polarizations that are maximally separated on the Poincaré sphere, lying at the corners of a tetrahedron. The Jones parameters of these four points are given in Table I.

Suppose we can measure the objective function for these four polarizations, described by γ_{xi} and γ_{yi} for $i = 1, \ldots, 4$. The adaptive algorithm is then as follows.

Algorithm 1. Polarization adaptation for configuration A.

1: $m := 4$
2: $G \in \mathcal{C}^{m \times m}$, **where the kth row of**

$$G = [|\gamma_{xk}|^2 \quad 2\text{Re}(\gamma_{yk}\gamma_{xk}^*) \quad -2\text{Im}(\gamma_{yk}\gamma_{xk}^*) \quad |\gamma_{yk}|^2]$$

3: **for** $i = 1$ **to** m **do**
4:    **Estimate** f_{Ai} **for** γ_{xi} **and** γ_{yi}
5: **end for**
6: $F_A := [f_{A1} \quad \cdots \quad f_{Am}]^T$
7: $[J_{A11} \quad \text{Re}(J_{A12}) \quad \text{Im}(J_{A12}) \quad J_{A22}]^T := G^\dagger F_A$
8: $J_A :=$
   $\begin{bmatrix} J_{A11} & \text{Re}(J_{A12}) + j\text{Im}(J_{A12}) \\ \text{Re}(J_{A12}) - j\text{Im}(J_{A12}) & J_{A22} \end{bmatrix}$
9: **Eigenvalue decomposition of** J_A **yields to eigenvalues** λ_0 **and** λ_1 **where** $\lambda_0 \geq \lambda_1$ **and eigenvectors** V_0 **and** V_1 **respectively**.
10: $\begin{bmatrix} \gamma_x^{\text{opt}} \\ \gamma_y^{\text{opt}} \end{bmatrix} := V_0$

In this algorithm, $\dagger$ denotes the pseudo-inverse, which can be taken as an inverse in systems with no noise.

B. One Amplitude/Phase/Polarization SLM

Configuration B, shown in Fig. 2(b), comprises a single SLM that is capable of controlling amplitude, phase and polarization of the launched light independently in each block. Considering ϑ_x and ϑ_y, the vectors of reflectances for x and y polarizations, respectively, the optimization problem is [13]

$$\text{maximize } f_B = [\vartheta_x^H \quad \vartheta_y^H] Q \begin{bmatrix} \vartheta_x \\ \vartheta_y \end{bmatrix}$$
$$\text{subject to } |\vartheta_{xi}|^2 + |\vartheta_{yi}|^2 \leq 1 \quad \forall i = 1, \ldots, N. \tag{19}$$

Comparing (19) to (17), we notice that in (19), the elements of $\begin{bmatrix} \vartheta_x \\ \vartheta_y \end{bmatrix}$ are connected through the constraint. Hence, the adaptive algorithms proposed in [10] cannot be applied to this problem. Nonetheless, the objective function can still be written as a quadratic form

$$f_B = \chi_1 |\vartheta_{ix}|^2 + \chi_2 |\vartheta_{iy}|^2$$
$$+ \text{Re}(\chi_3 \vartheta_{ix}^* \vartheta_{iy}) + \text{Re}(\chi_4^* \vartheta_{ix}) + \text{Re}(\chi_5^* \vartheta_{iy}) + \chi_6. \tag{20}$$

Here $\chi_1, \chi_2, \chi_6 \in \mathcal{R}$ and $\chi_3, \chi_4, \chi_5 \in \mathcal{C}$ are the variables that need to be measured (a total of nine real variables). Writing (20) in matrix format, we have

$$f_B = [\vartheta_{ix}^* \quad \vartheta_{iy}^*] \text{X}_{\text{complex}} \begin{bmatrix} \vartheta_{ix} \\ \vartheta_{iy} \end{bmatrix} + w_{\text{complex}}^H \begin{bmatrix} \vartheta_{ix} \\ \vartheta_{iy} \end{bmatrix}$$
$$+ [\vartheta_{ix}^* \quad \vartheta_{iy}^*] w_{\text{complex}} + \chi_6, \tag{21}$$

where $\text{X}_{\text{complex}} \in \mathcal{C}^{2\times2}$ and $w_{\text{complex}} \in \mathcal{C}^{2\times1}$ are defined as

$$\text{X}_{\text{complex}} = \begin{bmatrix} \chi_1 & \chi_3/2 \\ \chi_3^*/2 & \chi_2 \end{bmatrix} \tag{22}$$

$$w_{\text{complex}} = 1/2 \begin{bmatrix} \chi_4 \\ \chi_5 \end{bmatrix}. \tag{23}$$

During the adaptation process, each time the $\chi_1, \ldots, \chi_6$ are measured, the $\text{X}_{\text{complex}}$ and w_{complex} are computed.

In order to maximize f_B in (19) with respect to the ith block reflectance $\begin{bmatrix} \vartheta_{ix} \\ \vartheta_{iy} \end{bmatrix}$, which is in complex form, it is convenient to transform the problem to a real form. We define real matrices $\text{X}_{\text{real}} \in \mathcal{R}^{4\times4}$ and $w_{\text{real}} \in \mathcal{R}^{4\times1}$

$$\text{X}_{\text{real}} = \begin{bmatrix} \text{Re}(\text{X}_{\text{complex}}) & \text{Im}(\text{X}_{\text{complex}})^T \\ \text{Im}(\text{X}_{\text{complex}}) & \text{Re}(\text{X}_{\text{complex}}) \end{bmatrix} \tag{24}$$

$$w_{\text{real}} = \begin{bmatrix} \text{Re}(w_{\text{complex}}) \\ \text{Im}(w_{\text{complex}}) \end{bmatrix}. \tag{25}$$

Using (24) and (25) with (21), the block-wise adaptation problem, which maximizes f_B with respect to i th block of the SLM, becomes

$$\text{minimize } -f_B = -\eta^T \text{X}_{\text{real}} \eta - 2 w_{\text{real}}^T \eta$$
$$\text{subject to norm}(\eta) \leq 1, \tag{26}$$

where $\eta \in \mathcal{R}^{4\times1}$ is

$$\eta = [\text{Re}(\vartheta_{ix}) \quad \text{Re}(\vartheta_{iy}) \quad \text{Im}(\vartheta_{ix}) \quad \text{Im}(\vartheta_{iy})]^T. \tag{27}$$

Solving (26), we obtain the optimized reflectance for one SLM block, and by repeating the optimization for all blocks, we obtain an adaptive solution for (19). In general, $-X_{\text{real}} \not\succ 0$, so (26) is non-convex. Although non-convex, (26) obeys the strong duality rule [9]. Hence, (26) can be solved by solving the dual problem [9]

$$\text{minimize } -t - \nu$$
$$\text{subject to } \begin{bmatrix} \nu I_{4\times 4} - X_{\text{real}} & -w_{\text{real}} \\ -w_{\text{real}}^T & t \end{bmatrix} \geq 0 \qquad (28)$$

and deriving η^{opt} from

$$\eta^{\text{opt}} = (\nu^{\text{opt}} I_{4\times 4} - X_{\text{real}})^\dagger w_{\text{real}}. \qquad (29)$$

Once η^{opt} has been found, it is easy to construct $\begin{bmatrix} \vartheta_{ix} \\ \vartheta_{iy} \end{bmatrix}$ from (27). These steps have been incorporated in Algorithm 2, which we refer to as amplitude-phase-and-polarization SCA (APPSCA).

An alternative optimization method, which is suboptimal but potentially faster, is to optimize only the phase of each SLM block, holding the amplitude at unity, as in CPSCA [10]. In this case, the objective function can be written as

$$f_B = \text{Re}(\varsigma_1 \exp(j\angle\vartheta_{iy} - j\angle\vartheta_{ix})) + \text{Re}(\varsigma_2^* \exp(j\angle\vartheta_{ix}))$$
$$+ \text{Re}(\varsigma_3^* \exp(j\angle\vartheta_{iy})) + \varsigma_4 \quad (30)$$

where $\varsigma_1, \varsigma_2, \varsigma_3 \in \mathscr{C}$ and $\varsigma_4 \in \mathscr{R}$ are the variables that must be measured. The total number of measurements has been reduced from nine to seven. Using (30) we can write the optimization problem at each step as

$$\text{maximize } f_B = \varphi^H R_{\text{complex}} \varphi + z_{\text{complex}}^H \varphi$$
$$+ \varphi^H z_{\text{complex}} + \varsigma_4$$
$$\text{subject to } |\varphi_i| = 1 \quad \forall i = 1, 2 \qquad (31)$$

where

$$\varphi = \begin{bmatrix} \exp(j\angle\vartheta_{ix}) \\ \exp(j\angle\vartheta_{iy}) \end{bmatrix} \qquad (32)$$

and $R_{\text{complex}} \in \mathscr{C}^{2\times 2}, z_{\text{complex}} \in \mathscr{C}^{2\times 1}$ are defined as

$$R_{\text{complex}} = \begin{bmatrix} 0 & \varsigma_1/2 \\ \varsigma_1^*/2 & 0 \end{bmatrix} \qquad (33)$$

$$z_{\text{complex}} = 1/2 \begin{bmatrix} \varsigma_2 \\ \varsigma_3 \end{bmatrix}. \qquad (34)$$

It can be observed that the constraint in (31) is not in convex format. Hence, (31) cannot be solved using the methods we have discussed thus far. We have not derived an adaptive phase-and-polarization algorithm for configuration B, leaving the problem open for future work.

Algorithm 2. Amplitude-phase-and-polarization SCA for configuration B.

1:   $m := 9$
2:   $\vartheta_{ix} := (1/\sqrt{2}), \vartheta_{iy} := (1/\sqrt{2}), i = 1, \ldots, N$
3:   $\pi_x := [1/\sqrt{2} \quad 1/\sqrt{2} \exp(j\pi/2) \quad 1 \quad \exp(j2\pi/3)$
     $\exp(j4\pi/3) \quad 0 \quad 0 \quad 0 \quad 0]$
4:   $\pi_y :=$
     $[1/\sqrt{2} \quad 1/\sqrt{2} \quad 0 \quad 0 \quad 0 \quad 1 \quad \exp(j2\pi/3)$
     $\exp(j4\pi/3) \quad 0]$
5:   $G \in \mathscr{R}^{m\times m}$, where the kth row of $G =$
     $[|\pi_{xk}|^2 \quad |\pi_{yk}|^2 \quad \text{Re}(\pi_{xk}^*\pi_{yk}) \quad -\text{Im}(\pi_{xk}^*\pi_{yk})$
     $\text{Re}(\pi_{xk}) \quad \text{Im}(\pi_{xk}) \quad \text{Re}(\pi_{yk}) \quad \text{Im}(\pi_{yk}) \quad 1]$
6:   $i := 1$
7:   **repeat**
8:     Estimate f_{B0}
9:     $\vartheta_{ix} := \pi_{x1}$
10:    Estimate f_{B1}
11:    $\vartheta_{iy} := 0$
12:    **for** $m = 1$ to 3 **do**
13:      $\vartheta_{ix} := \pi_{x(m+1)}$
14:      Estimate $f_{B(m+1)}$
15:    **end for**
16:    $\vartheta_{ix} := 0$
17:    **for** $m = 1$ to 3 **do**
18:      $\vartheta_{iy} := \pi_{x(m+4)}$
19:      Estimate $f_{B(m+4)}$
20:    **end for**
21:    $\vartheta_{ix} := 0, \vartheta_{iy} := 0$
22:    Estimate f_{B8}
23:    $F_B := [f_{B0} \cdots f_{B9}]^T$
24:    $[\chi_1 \; \chi_2 \; \chi_{3Re} \; \chi_{3Im} \quad \chi_{4Re} \; \chi_{4Im} \; \chi_{5Re} \; \chi_{5Im} \; \chi_6]^T :=$
     $G^\dagger F_B$
25:   $X_{\text{complex}} :=$
     $\begin{bmatrix} \chi_1 & (\chi_{3Re} + j\chi_{3Im})/2 \\ (\chi_{3Re} - j\chi_{3Im})/2 & \chi_2 \end{bmatrix}$
26:   $w_{\text{complex}} := 1/2 \begin{bmatrix} \chi_{4Re} + j\chi_{4Im} \\ \chi_{5Re} + j\chi_{5Im} \end{bmatrix}$
27:    $X_{\text{real}} := \begin{bmatrix} \text{Re}(X_{\text{complex}}) & \text{Im}(X_{\text{complex}})^T \\ \text{Im}(X_{\text{complex}}) & \text{Re}(X_{\text{complex}}) \end{bmatrix}$
28:    $w_{\text{real}} = \begin{bmatrix} \text{Re}(w_{\text{complex}}) \\ \text{Im}(w_{\text{complex}}) \end{bmatrix}$
29:

$$\text{minimize } -t - \nu$$
$$\text{subject to } \begin{bmatrix} \nu I_{4\times 4} - X_{\text{real}} & -w_{\text{real}} \\ -w_{\text{real}}^T & t \end{bmatrix} \geq 0$$

30:    $[\vartheta_{ix\text{Re}} \; \vartheta_{iy\text{Re}} \; \vartheta_{ix\text{Im}} \; \vartheta_{iy\text{Im}}]^T :=$
     $(\nu^{\text{opt}} I_{4\times 4} - X_{\text{real}})^\dagger w_{\text{real}}$
31:    $\vartheta_{ix} := \vartheta_{ix\text{Re}} + j\vartheta_{ix\text{Im}}$
32:    $\vartheta_{iy} := \vartheta_{ix\text{Re}} + j\vartheta_{iy\text{Im}}$
33:    $i := (i)\text{mod}(N) + 1$
34: **until** Termination

C. Two Amplitude/Phase SLMs With Adaptive Uniform Amplitude Allocation Between Polarizations

Configuration D, shown in Fig. 2(d), uses separate SLMs for the x and y polarizations, SLM_x and SLM_y, and uses a uniform variable BS to allocate amplitude ratios $\sqrt{\varepsilon}$ and $\sqrt{1-\varepsilon}$, respectively, to the two SLMs. In [13], it was shown that, much like configuration A, this configuration can be optimized by alternating between two optimizations. The first optimization is over SLM reflectances, keeping the BS ratio constant

$$\text{maximize } f_D|_{\varepsilon\text{constant}} = \begin{bmatrix} \vartheta_x^H & \vartheta_y^H \end{bmatrix} U \begin{bmatrix} \vartheta_x \\ \vartheta_y \end{bmatrix}$$

$$\text{subject to } |\vartheta_{xi}|^2 \leq 1 \forall i = 1, \dots, N$$

$$|\vartheta_{yi}|^2 \leq 1 \quad (35)$$

where

$$U = \begin{bmatrix} \varepsilon Q_1 & \sqrt{\varepsilon}\sqrt{1-\varepsilon}Q_2 \\ \sqrt{\varepsilon}\sqrt{1-\varepsilon}Q_2^H & (1-\varepsilon)Q_3 \end{bmatrix}. \quad (36)$$

The second optimization is over the BS ratio, keeping the SLM reflectances constant:

$$\text{maximize } f_D|_{\vartheta\text{constant}} = \rho^T J_D \rho$$

$$\text{subject to norm}(\rho) \leq 1,$$

$$\rho \geq 0 \quad (37)$$

where

$$\rho = \begin{bmatrix} \sqrt{\varepsilon} \\ \sqrt{1-\varepsilon} \end{bmatrix} \in R^{+2} \quad (38)$$

and $J_D \in R^{2\times2}$ is

$$J_D = \text{Re}\left(\begin{bmatrix} \vartheta_x^H Q_1 \vartheta_x & \vartheta_x^H Q_2 \vartheta_y \\ \vartheta_y^H Q_2^H \vartheta_x & \vartheta_y^H Q_3 \vartheta_y \end{bmatrix}\right). \quad (39)$$

Adaptive algorithms for the first optimization (35) (both APCSA and CPSCA) are given in [10]. In [13], it was shown that in order to solve the second optimization (37), we can solve

$$\text{maximize } w|_{\vartheta\text{constant}} = \beta^T J_D \beta$$

$$\text{subject to norm}(\beta) \leq 1 \quad (40)$$

and obtain $\rho^{\text{opt}} = |\beta^{\text{opt}}|$ as an optimal solution to (37). The solution to (40) can be obtained by eigenvalue decomposition of J_D. As J_D is a real matrix it has three independent elements that need to be calculated. Hence, three measurements are needed. All of these steps for optimizing the BS amplitude allocation are combined into Algorithm 3.

Algorithm 3. Beam splitter amplitude allocation optimization for configuration C.

1: $m := 3$
2: $e := \begin{bmatrix} 0 & 0.5 & 1 \end{bmatrix}$
3: $G \in R^{m\times m}$, **where the kth row of**
   $G = \begin{bmatrix} e_k & 2\sqrt{e_k(1-e_k)} & (1-e_k) \end{bmatrix}$

TABLE II
NUMBER OF OBJECTIVE FUNCTION MEASUREMENTS PER ADAPTATION PASS.
IN EACH CASE, N REPRESENTS THE NUMBER OF SLM BLOCKS

Variables	Number of measurements				
	Config. A	Config. B	Config. C	Config. D	Config. E
Amplitude, phase and polarization	$4N+4$	$9N$	$8N$	$8N+3$	$9N$
Phase and polarization	$3N+4$	$7N$	$6N$	$6N+3$	$7N$

4: **for** $i = 1$ **to** m **do**
5: $\quad \rho := \begin{bmatrix} \sqrt{e_i} & \sqrt{1-e_i} \end{bmatrix}^T$
6: $\quad$ **Estimate** f_{Di} **for** ρ
7: **end for**
8: $F_D := \begin{bmatrix} f_{D1} & \cdots & f_{Dm} \end{bmatrix}^T$
9: $\begin{bmatrix} J_{D11} & J_{D12} & J_{D22} \end{bmatrix}^T := G^\dagger F_D$
10: $J_D := \begin{bmatrix} J_{D11} & J_{D12} \\ J_{D12} & J_{D22} \end{bmatrix}$
11: **Eigenvalue decomposition of** J_D **yields to eigenvalues** λ_0 **and** λ_1 **where** $\lambda_0 \geq \lambda_1$ **and eigenvectors** V_0 **and** V_1 **respectively**.
12: $\rho^{\text{opt}} := |V_0|$

D. Adaptation in the Presence of Noise

When adapting in the presence of noise, it can become necessary to average each objective function measurement over a substantial time interval in order to reduce its uncertainty. The total time required for adaptation can depend strongly on the number of objective function measurements per pass over the adaptive optics system. Table II compares the number of measurements for each of the five configurations, considering adaptation of amplitude, phase and polarization or of phase and polarization.

Configuration A requires four measurements per block when adapting amplitude and phase, and requires three measurements per block when adapting phase only. After SLM adaptation, four measurements are needed to find the best polarization. Configurations B and E require nine measurements per block when adapting amplitude, phase and polarization, which reduces to seven measurements per block when adapting only phase and polarization. Neither configuration requires separate polarization adaptation. Configuration C and D have two SLMs, and for each SLM, they require four or three measurements per block for adapting amplitude and phase or only phase, respectively. Configuration D requires an additional three measurements for BS adaptation.

Comparing these results, we notice that configurations B and E require the largest number of measurements, however, as we will see, these configurations achieve better performance than configurations A or D. Configuration A requires the smallest number of measurements, which may make it attractive for systems that change rapidly over time, however, this configuration is known to yield the poorest performance among the five [13].

As an alternative or adjunct to noise averaging, the adaptive algorithms described above can be modified to increase their tolerance to noise [10]. In a noise-resilient algorithm, the number of measurements performed exceeds the number of variables that must be measured, and the adaptive optics settings are computed as minimum mean-square estimates given the measurements. Further details are given in [10].

IV. NUMERICAL SIMULATIONS

A. Fiber and System Parameters

In order to model propagation in MMF, including both spatial- and polarization-mode coupling, we use the approach described in detail in [12], [14]. A MMF is modeled by concatenating many curved sections, each section lying in a plane, with the plane of one section rotated with respect to the previous section. We model a 50-μm-core graded-index MMF of total length $L = 1000$ m. The fiber has a numerical aperture (NA) of 0.19. At the wavelength $\lambda = 1550$ nm, there are 55 propagating modes in each polarization and the refractive index at the center of the core is $n_0 = 1.444$. In order to best reproduce the experimental results in [7], the refractive index exponent is chosen to be $\alpha = 2.17$. Birefringence, defined as the difference between refractive indexes seen by x- and y-polarized waves, is assumed to be induced by stress due to curvature [16]:

$$n_x - n_y = \delta \frac{C_s}{2k_0}(a\kappa)^2, \tag{41}$$

where κ is the curvature of a fiber section, and C_s/k_0 is referred to as the strain-optical coefficient. For an SMF, $C_s/k_0 = 0.0878n^3$ [16], and the birefringence scale factor should be set to $\delta = 1$. In MMF, birefringence and spatial-mode coupling do not necessarily have the same physical origins. In our model, both effects are induced by curvature; hence, in order to yield sufficient polarization-dependent spatial-mode coupling, the birefringence scale factor is set to $\delta = 2100$. The fiber is divided into 10^4 sections, each 0.1 m long. Each section is rotated with respect to the previous one by an independent, identically distributed (i.i.d.) angle θ, whose probability density function (pdf) is normal with variance $\sigma_\theta^2 = 0.36$ rad^2. The curvature of each section is an i.i.d. random variable κ_i, whose pdf is the positive side of a normal pdf. Simulations are performed for a fiber in the low-coupling regime with a curvature standard deviation $\sigma_\kappa = 2.53$ m^{-1}, and a fiber in the medium-coupling regime with $\sigma_\kappa = 4.62$ m^{-1}. The model described in [12] is used to calculate the first-order PMs and their GDs.

We model the adaptive transmission system, as shown in Fig. 1, choosing parameters corresponding to previous experiments [7]. On-off keying is performed at a bit rate of 10 Gb/s. In the absence of ISI, an isolated 1 bit is described by a pulse shape $q(t)$, which is modeled by a Gaussian pulse with full-width at half-maximum of 60 ps. The modulator output is conveyed in a SMF with a NA of 0.11, whose output is collimated by a lens of 10.4-mm focal length, and is input to the adaptive optics system. The output of the adaptive optics system is focused onto the MMF by a lens of 10.4-mm focal length. Thus, the beam launched into the MMF has a NA of 0.11 (enclosing 95% of the total power), assuming the adaptive optics system is set to unit reflectivity in all blocks. Within the adaptive optics system,

Fiber regime	Objective function		
	Random polarization	Best polarization	Worst polarization
Low-coupling	0.1775	0.2599	−0.2666
Medium-coupling	−0.1857	−0.0904	−0.3937

the SLM(s) are operated with an 8×8 array of square blocks covering 95% of the beam's total power. Although the SLM in [7] provided phase control only, here, the SLM(s) are assumed to control both amplitude and phase. The noise is assumed to be i.i.d. Gaussian, with the noise power chosen so that CPSCA achieves a BER less than 10^{-3} in at least 95% of the cases simulated, as in [10]. A BER of 10^{-3} is a typical forward error correction code threshold. Estimation of the objective function is done using a periodic training sequence comprising a string of 0 bits followed by a string of 1 bits [10]. The total duration is 128 bits, corresponding to 12.8 ns. The SLM switching time is assumed to be 10 μs, and the PC adaptation time is assumed to be 30 μs. We use a noise averaging over 10 μs, with multiple copies of the training sequence being sent over this period.

Whenever an optimization is required, we have used CVX, a freely distributed convex optimization library for MATLAB [17].

B. Simulation Result: Noiseless Systems

Table III presents objective function values obtained for fibers in the low- and medium-coupling regimes, using configuration A with a blank SLM, optimizing polarization only. Three polarizations are considered: a random polarization, the best polarization (which maximizes the objective function) and the worst polarization (which minimizes the objective function). In the fiber with low coupling, optimizing the polarization alone can open the eye, but in the fiber with medium coupling, optimizing polarization alone is not sufficient to open the eye. These results illustrate the importance of optimizing both spatial and polarization degrees of freedom.

Table IV presents objective function values obtained for fibers in the low- and medium-coupling regimes, for the five adaptive optics configurations, after optimization by the methods described in [13]. For each fiber, optimized objective function values are positive, indicating that the eye is open. Comparing globally optimized objective function values obtained using configurations A, B and D, we observe that for each fiber $f_B > f_D > f_A$, consistent with the analysis in [13].

As discussed previously, the problem of optimizing the launched field over amplitude, phase and polarization is non-convex. Hence, adaptation starting from different initial conditions may yield different locally optimal solutions or different convergence rates. In order to illustrate these effects, for a given fiber, we adapt the system three times, starting from the best polarization, the worst polarization, and a random polarization, comparing the results. Fig. 3(a) shows adaptation for a fiber in the low-coupling regime, with $\sigma_\kappa = 2.53$ m^{-1}. Starting from all three initial polarizations, the objective

TABLE IV
OBJECTIVE FUNCTIONS OBTAINED AFTER OPTIMIZATION USING FIVE
DIFFERENT ADAPTIVE OPTICS CONFIGURATIONS FOR A 1-KM, 50-μM-CORE
GRADED-INDEX MMF IN LOW- AND MEDIUM-COUPLING REGIMES [13]

Fiber regime	Objective function				
	Config. A	Config. B	Config. C	Config. D	Config. E
Low-coupling	0.5208	0.8514	0.6803	0.7373	0.8514
Medium-coupling	0.4185	0.7968	0.6741	0.6791	0.7968
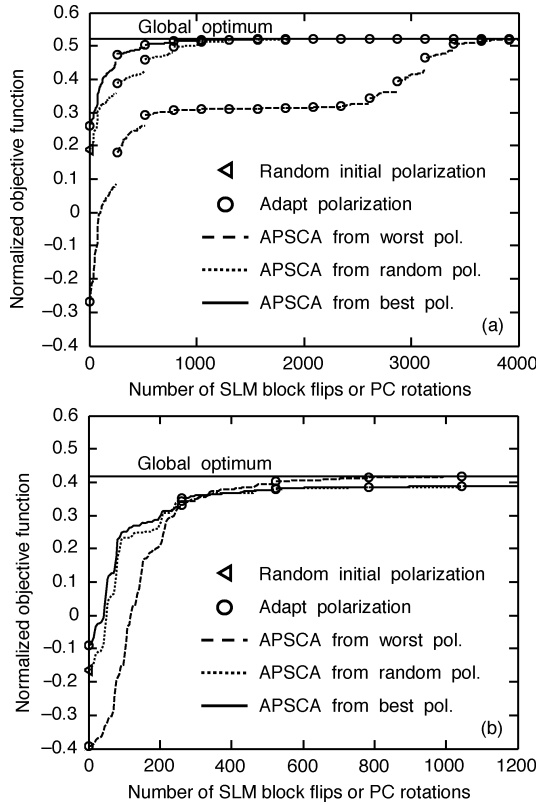
Fig. 3. Convergence for configuration A with APSCA using 8 × 8 SLM blocks in the absence of noise (a) Fiber in low-coupling regime with curvature standard deviation $\sigma_\kappa = 2.53$ m^{-1} and $\alpha = 2.17$. (b) Fiber in medium-coupling regime with $\sigma_\kappa = 4.62$ m^{-1} and index exponent $\alpha = 2.17$. SLM: spatial light modulator, PC: polarization controller.

Fig. 4. Convergence with or without noise for fiber in medium-coupling regime with curvature standard deviation $\sigma_\kappa = 4.62$ m^{-1} and index exponent $\alpha = 2.17$, using 8 × 8 SLM blocks, 10-μs SLM switching time, 30-μs polarization switching time, 10-μs noise averaging. (a) Configuration A, APSCA and CPSCA. (b) Configuration B, APPSCA. (c) Configuration D, APSCA and CPSCA. SLM: spatial light modulator, PC: polarization controller, BS: beam splitter.

function eventually converges to the global optimum given in Table IV. On the other hand, the convergence rate is much slower when starting from the worst polarization. Fig. 3(b) shows adaptation for a fiber in the medium-coupling regime, with $\sigma_\kappa = 4.62$ m^{-1}. In this case, starting from the best polarization, although the adaptation converges rapidly, the algorithm is caught in a local optimum, and the objective function does not converge to the global optimum. Similar behavior is observed when starting from a random polarization. Starting from the worst polarization, however, the objective function ultimately converges to the global optimum given in Table IV. This example shows that optimizing the initial condition does not always result in the best final result.
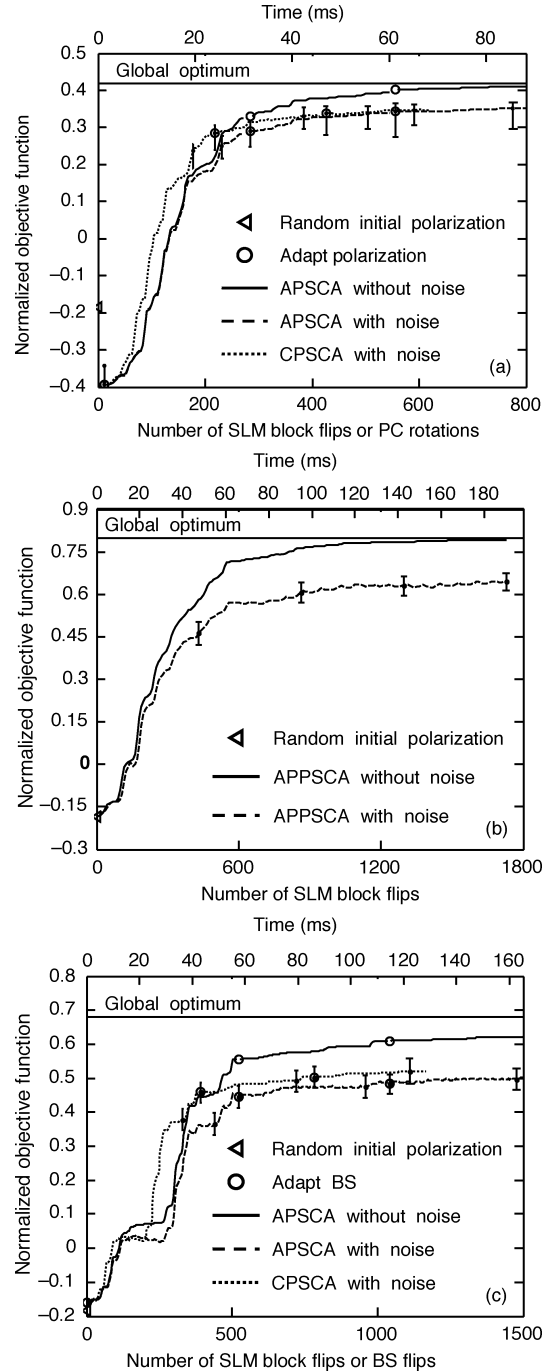
C. Simulation Results: Noisy Systems

In this section, we study the performance of our adaptive algorithms for configurations A, B, and D. Each simulation makes three complete passes over all 64 blocks in the adaptive optical system. Simulation results are shown in Fig. 4(a)–(c). The bottom scales indicate the number of individual variable optimizations (block flips, PC rotations or BS flips), which differ
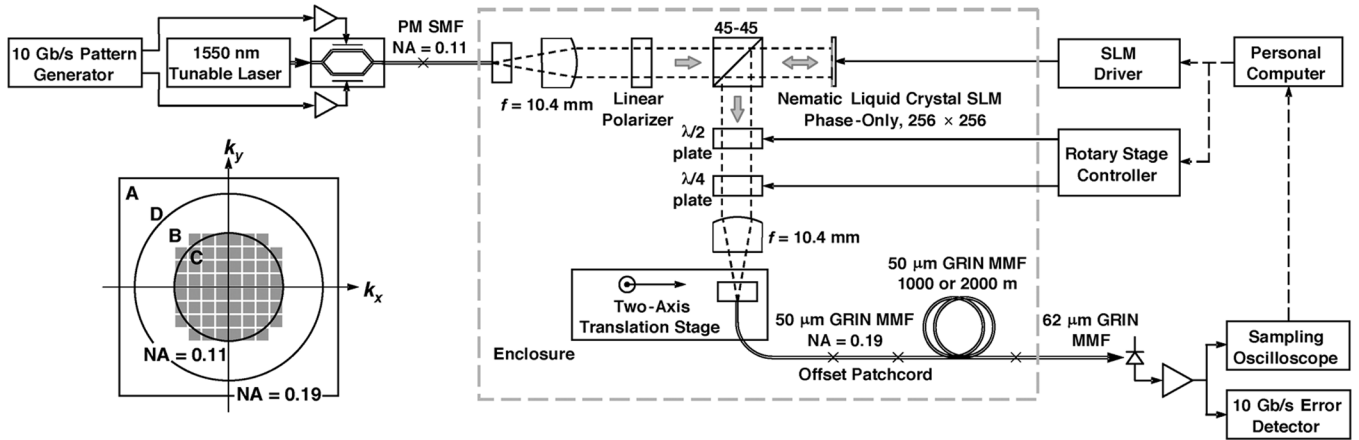
Fig. 5. Setup of 10-Gb/s transmission experiments. Inset: spatial light modulator (SLM) in Fourier plane of the input face of multimode fiber (MMF). A: boundary of SLM active region. B: SLM pattern with 60 blocks. C: Circle enclosing 95% of power of beam incident on SLM, corresponding to a numerical aperture (NA) of 0.11 at the MMF input. D: Circle corresponding to the MMF NA of 0.19.

depending on the configuration and which variables are adapted (amplitude, phase, polarization). The top scales indicate real time, which scales with the number of variable optimizations, and also depends on the number of objective function measurements per variable optimization. We only perform simulations for one random realization of a fiber in the medium-coupling regime with $\sigma_\kappa = 4.62 \text{ m}^{-1}$. When noise is included, we repeat the simulation for 500 random realizations of noise sequences, plotting the median results and using error bars to indicate the 5th and 95th percentiles.

Fig. 4(a) considers configuration A, with the SLM adapted by APSCA or CPSCA as in [10] and the PC adapted using Algorithm 1. In all cases, initially the PC is adapted to the worst-case polarization, because this enables a larger objective function value to be reached ultimately, as illustrated in Fig. 3(b). In subsequent adaptation cycles, the PC is adapted to the best polarization. Without noise, APSCA converges to the global optimum given in Table IV (the lowest among configurations A, B and D). With noise, CPSCA adapts faster than APSCA, as observed in [10]. After three passes through the adaptive optics system, both algorithms yield similar objective function values, which are well below the global optimum. Nonetheless, configuration A with CPSCA is attractive for its hardware simplicity and fast adaptation.

Fig. 4(b) considers configuration B, with the SLM adapted by APPSCA using Algorithm 2. Without noise, APPSCA converges to the global optimum given in Table IV (the highest among configurations A, B and D). With noise, APPSCA yields an objective function value well below the optimum, but much higher than that using configuration A with APSCA or CPSCA.

Fig. 4(c) considers configuration D, with the SLMs adapted by APSCA or CPSCA [10] and the uniform variable BS adapted using Algorithm 3. Without noise, APSCA does converge to the global optimum given in Table IV (intermediate between configurations A and B), but this requires more than three passes through the adaptive optics system. With noise, CPSCA adapts faster than APSCA, as in [10], yielding a higher objective function value after three passes. Eventually, both APSCA and CPSCA converge to similar objective function values, well

below the global optimum, and intermediate between the values reached by configurations A and B.

The adaptation speed of these algorithms may restrict their applicability to systems in which the MMF is not subject to sudden large perturbations. As an example, consider configuration A using CPSCA, depicted in Fig. 4(a). Starting from a blank SLM and the worst polarization, on average, the objective function reaches 90% of its final value in 25 ms. We expect that this algorithm could track arbitrarily large but slow perturbations caused by thermal drift or manual movement of the MMF. At the other extreme, clearly the algorithm could not track large sub-millisecond perturbations, such as mode jumps in a multi-transverse-mode laser source. Further study would be required to evaluate tracking of perturbations at intermediate time scales, such as mechanical vibrations caused by motors.

V. EXPERIMENTS

In this section, we describe 10 experiments using configuration A, adapting the phase-only SLM using CPSCA, and adapting the PC using Algorithm 1.

A. Experimental Setup

The experimental system is shown in Fig. 5. A C-band laser can be tuned to a comb of 100 channels, spaced by 50 GHz, covering a 5.0-THz bandwidth over the 1527–1567 nm wavelength range. A Mach–Zehnder modulator encodes a chirp-free 10-Gb/s non-return-to-zero signal having +6-dBm average power, output into a polarization-maintaining SMF (PM-SMF) having a NA of 0.11. Inline with the PM-SMF, a PM variable optical attenuator (not shown in Fig. 5), which has a 0.6-dB insertion loss, provides an attenuation continuously variable from 0 to 45 dB. The PM-SMF output, in the LP_{01} mode, is collimated by an $f = 10.4$-mm plano-convex lens, passed through a polarizer, and illuminates a liquid crystal SLM. The linear polarizer aligns the PM-SMF output to the linear state of polarization (SOP) required at the SLM.

The nematic liquid crystal, phase-only SLM has 256×256 pixels, each 18 μm in size, each independently controllable with 5–6 bit resolution over the range 0–2π, with switching times

of 10 μs $(0 \leftrightarrow 2\pi, 10\% \leftrightarrow 90\%)$. The SLM surface is depicted in the inset of Fig. 5. "A" denotes the boundary of the 256×256-pixel active region. To reduce adaptation time, pixels are grouped into 16×16-pixel blocks. A set of 60 such blocks is used, denoted by "B". This set covers a circle enclosing 95% of the energy of the incident LP_{01} mode of the PM-SMF, denoted by "C", which corresponds to an NA of 0.11 in the MMF. The circle "D" denotes the MMF NA of 0.19. After reflecting from the SLM and passing through a 45%-45% polarization-independent BS, the beam goes through a PC consisting of half-wave and quarter-wave plates, mounted on motorized rotation stages. The PC allows the launched signal to be adjusted to an arbitrary elliptical SOP. The beam is coupled into a MMF using an $f = 10.4$-mm plano-convex lens, resulting in a launched NA of 0.11. The beam spot is nominally center-launched to excite lower-order PMs preferentially. The free-space optical system has a loss of about 8.5 dB, and an average power of about -3.1 dBm is launched in the MMF (for a blank SLM) [11].

Test fibers are spooled, plastic-jacketed, 50-μm graded-index MMF, having power-law index profiles with exponents between 2.00 and 2.03, and NA of 0.19 at C-band wavelengths. An offset-spliced patchcord is inserted before the test fiber to simulate the effect of connector offsets, which are often encountered in deployed systems. Experiments are performed on two fiber configurations, a 2000-m OM3 fiber with a 4-μm offset patchcord, and a 1000-m OM2 fiber with a 2-μm offset patchcord. BER performance is measured at a bit rate of 10 Gb/s using a $2^{31} - 1$ psuedorandom bit sequence, and a gating period of 3 s. The test fiber output is connected via a 62.5-μm MMF pigtail to a commercial receiver comprising an InGaAs p-i-n photodiode and transimpedence preamplifier, and having d.c.-9.5 GHz bandwidth (at -3 dB). The receiver sensitivity and overload power are -20 dBm and $+2$ dBm, respectively, both at 10^{-10} BER [11].

Receiver electrical output waveforms are acquired by a sampling oscilloscope and sent via a general-purpose instrument bus (GPIB) interface to a personal computer, which estimates ISI, performs computations associated with the adaptive algorithms, and controls the SLM and the PC. A single estimation of the objective function f requires about 0.9 s, which is dominated by waveform sampling and GPIB interface latency. In a single iteration of the adaptive algorithm, the 60 SLM blocks are optimized one-by-one, proceeding outward from the center in concentric rings, which takes about 3 min. using the current setup. Further details about objective function estimation and SLM optimization are given in [11].

B. Experimental Results

Fig. 6 shows eye diagrams and BER-vs.-attenuation performance for a 2000-m OM3 fiber with 4-μm offset splice. Although the fiber itself is near-ideal, the offset introduces mode coupling and polarization sensitivity. Fig. 6(a) shows results for a blank SLM after adaptation to the best polarization, in which relatively little coupling to higher-order PMs occurs. The eye is somewhat open, and 10^{-9} BER is achieved at an attenuation of about 3.5 dB. After polarization adaptation, the SLM is adapted once using CPSCA, yielding results shown in Fig. 6(b). The eye is fully open, and 10^{-9} BER is achieved at an attenuation of
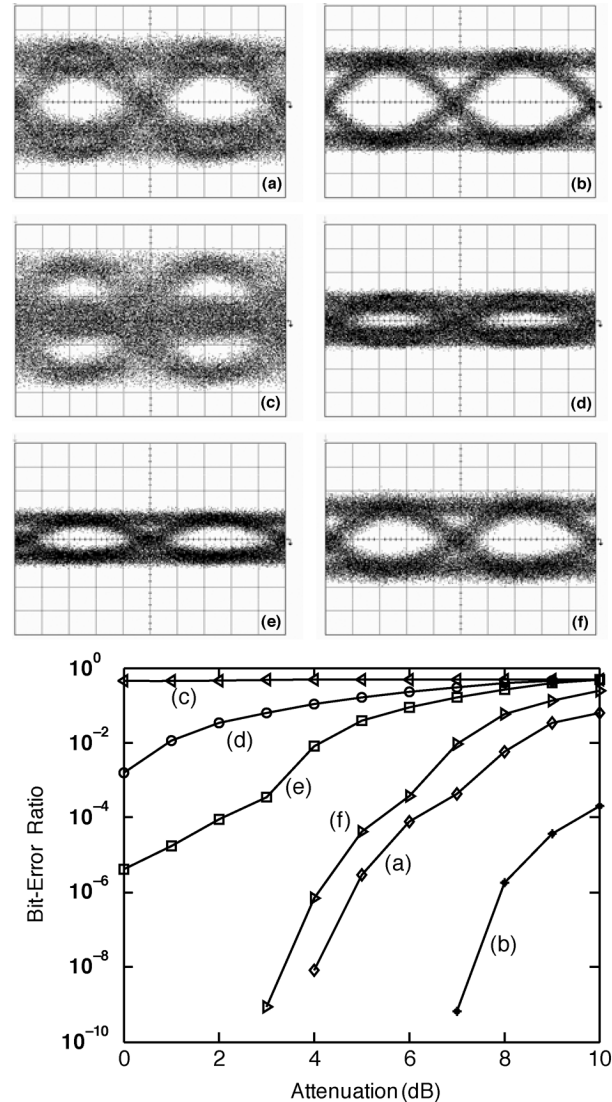

Fig. 6. 10-Gb/s transmission using configuration A for a 2000-m OM3 fiber with 4-μm offset splice, showing eye diagrams (for 0-dB attenuation) and BER versus attenuation: (a) best polarization; (b) best polarization, CPCSA; (c) worst polarization; (d) worst polarization, CPSCA; (e) worst polarization, CPSCA, adapt polarization; (f) worst polarization, CPSCA, adapt polarization, CPSCA.

about 7 dB, representing about a 3.5-dB improvement over polarization adaptation alone.

Fig. 6(c) shows results for a blank SLM after adaptation to the worst polarization, in which significant coupling to higher-order PMs occurs. The eye is completely closed, and a BER below 0.5 cannot be achieved. Remaining in the worst polarization, the SLM is adapted once using CPSCA, yielding Fig. 6(d). The SLM is able to reduce some of the power coupled to higher-order PMs, but also reduces the power coupled to the desired lowest-order PM.[3] The eye is slightly open, but a BER below 10^{-3} cannot be achieved with the power available. A comparison of Fig. 6(b) and (d) illustrates clearly the importance of polarization adaptation in fibers that exhibit strong polarization sensitivity.

---

[3]Although the SLM is phase-only, the introduction of high-spatial-frequency patterns, such as adjacent anti-phased blocks, can diffract light away from the core of the MMF.
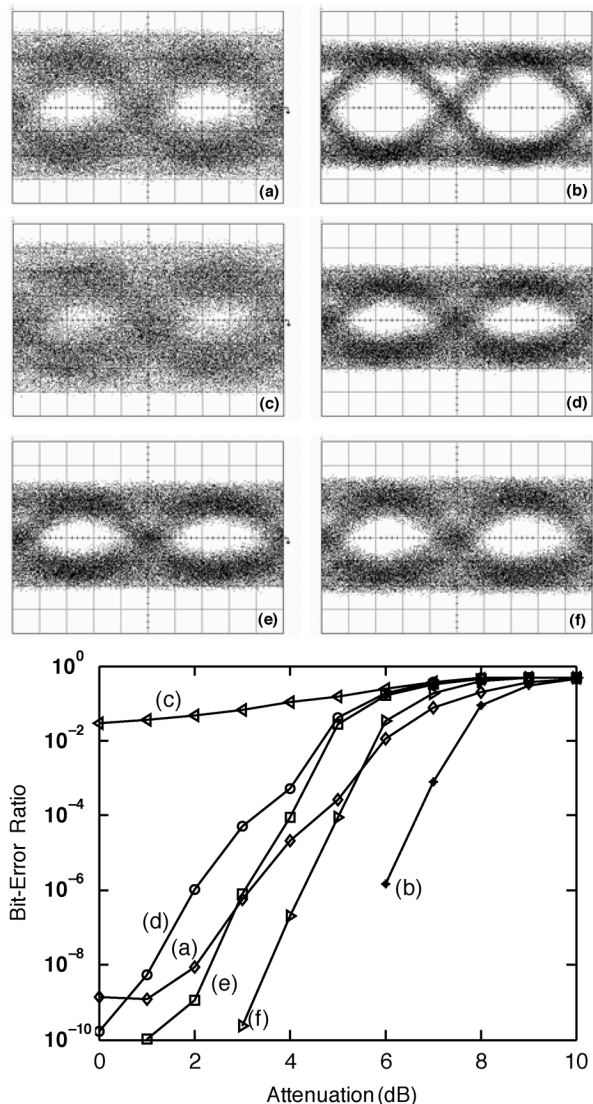
Fig. 8. 10-Gb/s transmission using configuration A for a 2000-m OM3 fiber with 4-μm offset splice, showing BER (for 2-dB attenuation) versus channel number, after adaptation in channel 58. The channel spacing is 50 GHz. (a) Best polarization; (b) best polarization, CPSCA; (c) worst polarization, CPSCA, adapt polarization; (d) worst polarization, CPSCA, adapt polarization, CPSCA.

Fig. 9. 10-Gb/s transmission using configuration A for a 1000-m OM2 fiber with 2-μm offset splice, showing BER (for 2-dB attenuation) versus channel number, after adaptation in channel 58. The channel spacing is 50 GHz. (a) Best polarization, CPSCA; (b) worst polarization, CPSCA, adapt polarization, CPSCA.

Fig. 7. 10-Gb/s transmission using configuration A for a 1000-m OM2 fiber with 2-μm offset splice, showing eye diagrams (for 0-dB attenuation) and BER versus attenuation: (a) best polarization; (b) best polarization, CPCSA; (c) worst polarization; (d) worst polarization, CPSCA; (e) worst polarization, CPSCA, adapt polarization; (f) worst polarization, CPSCA, adapt polarization, CPSCA.

After adapting to the worst polarization and adapting the SLM using CPSCA, the polarization is adapted in an attempt to minimize ISI, yielding results shown in Fig. 6(e). The PC cannot increase the total power coupled into propagating modes. The eye opening is increased slightly, but a BER below 4×10^{-6} cannot be achieved. However, this PC adaptation sets the stage for a second adaptation of the SLM, yielding results shown in Fig. 6(f). The eye is somewhat open, and 10^{-9} BER is achieved at an attenuation of about 3 dB. This result is still slightly worse than Fig. 6(a), obtained with a blank SLM after polarization adaptation alone. Several more cycles of polarization and SLM adaptation are required to approach the result shown in Fig. 6(b). Comparing Fig. 6(a)–(b) and (c)–(f) illustrates how starting the adaptation process from the best polarization can lead to much faster convergence than starting from the worst polarization. These results are in good agreement with the simulations shown in Fig. 3(a).
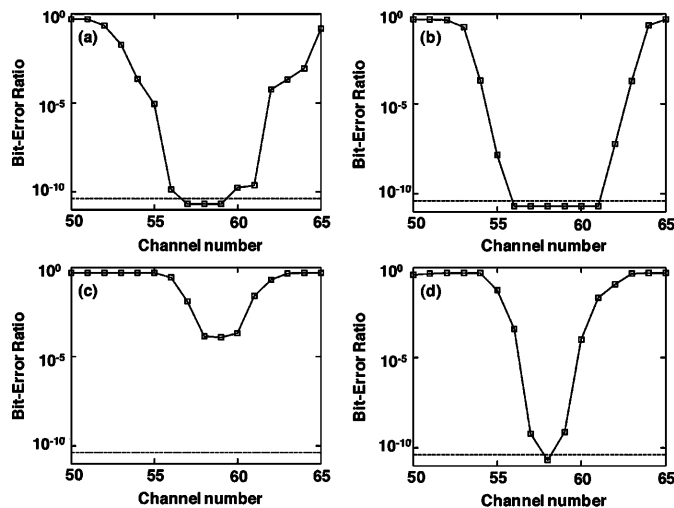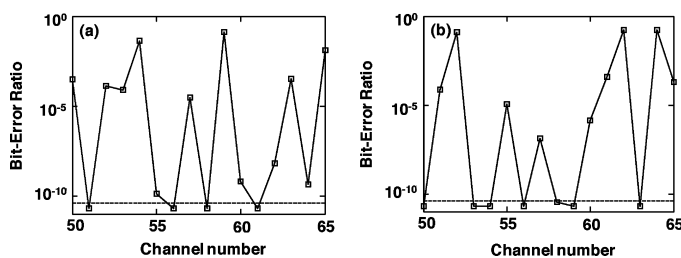
Fig. 7 shows eye diagrams and BER-vs.-attenuation performance for a 1000-m OM2 fiber with 2-μm offset splice. For this less-ideal fiber type, a shorter length and smaller offset lead to significant mode coupling and polarization sensitivity. Adapting first to the best polarization, as shown in Fig. 7(a)–(b), yields results similar to, although worse than, the previous fiber, shown in Fig. 6(a)–(b). Adapting first to the worst polarization, as shown in Figs. 7(c)–(f), yields results similar to, or slightly better than, the previous fiber, shown in Fig. 6(c)–(f).

Figs. 8–9 investigate the possibility of using a single SLM and PC to compensate modal dispersion in several wavelength-division-multiplexed channels. In each experiment, adaptation is first performed (as described below) in channel 58, corresponding to a 1550-nm wavelength. Then, holding the PC and SLM fixed, and setting the attenuation at 2 dB, the laser frequency is swept in 50-GHz steps and the BER is measured.

Fig. 8 illustrates a 2000-m OM3 fiber with 4-μm offset splice. Adaptation to the best polarization with a blank SLM [Fig. 8(a)] yields low BERs in three contiguous channels, while subsequent SLM adaptation by CPSCA [Fig. 8(b)] yields low BERs in six contiguous channels. Adaptation to the worst polarization, followed by SLM adaptation and another polarization adaptation

[Fig. 8(c)] does not yield low BER in any channels, while a subsequent second SLM adaptation [Fig. 8(d)] yields a low BER in only one channel.

Fig. 9 illustrates a 1000-m OM2 fiber with 2-μm offset splice. The best performance in channel 58 is obtained by adapting to the best polarization and then adapting the SLM using CPSCA [Fig. 9(a)], but even in this case, a low BER is not obtained in any other contiguous channels. Similar results are observed after adapting to the worst polarization, adapting the SLM, readapting polarization, and readapting the SLM [Fig. 9(b)]. The erratic frequency dependence of the OM2 fiber (Fig. 9) contrasts with the smooth frequency dependence of the OM3 fiber (Fig. 8), but the difference can be attributed qualitatively to the stronger mode coupling present in the OM2 fiber.

VI. CONCLUSION

We have proposed adaptive algorithms to compensate modal dispersion in MMF using three different configurations of adaptive optical systems that can control the amplitude, phase and polarization of the launched field. These algorithms do not require prior knowledge of a fiber's PMs and their GDs, instead setting the optical system adaptively, based upon measurements of the eye opening. Using simulation, we studied the performance of these algorithms without and with additive noise. For all three configurations, an amplitude-and-phase sequential coordinate ascent algorithm (generalized to include polarization for one configuration) converges to the globally optimal solution in the absence of noise. In two configurations, using phase-only SLM(s) with a continuous-phase sequential coordinate ascent algorithm can reduce hardware complexity and speed up convergence, while achieving near-optimal performance. We performed experiments using a phase-only SLM and a PC with continuous-phase sequential coordinate ascent, demonstrating 10-Gb/s transmission through up to 2000 m of 50-μm-core graded-index MMF, even in the presence of splices offset by up to 4 μm.

REFERENCES

[1] G. P. Agrawal, *Fiber-Optic Communication Systems*, 3rd ed. New York: Wiley, 2002.
[2] X. Zhao and F. S. Choa, "Demonstration of 10-Gb/s transmissions over 1.5-km-long multimode fiber using equalization techniques," *IEEE Photon. Technol. Lett.*, vol. 14, no. 8, pp. 1187–1189, Aug. 2002.
[3] H. Wu, J. A. Tierno, P. Pepeljugoski, J. Schaub, S. Gowda, J. A. Kash, and A. Hajimiri, "Integrated transversal equalizers in high-speed fiber-optic systems," *IEEE J. Solid-State Circuits*, vol. 38, no. 12, pp. 2131–2137, Dec. 2003.
[4] J. Proakis, *Digital Communications*, 4th ed. New York: McGraw-Hill, 2001.
[5] S. Fan and J. M. Kahn, "Principal modes in multi-mode waveguides," *Opt. Lett.*, vol. 30, no. 2, pp. 135–137, Jan. 2005.
[6] E. Alon, V. Stojanovic, J. M. Kahn, S. P. Boyd, and M. A. Horowitz, "Equalization of modal dispersion in multimode fiber using spatial light modulators," in *Proc. IEEE Global Telecommun. Conf.*, Dallas, TX, Nov. 29–Dec. 3 2004.
[7] X. Shen, J. M. Kahn, and M. A. Horowitz, "Compensation for multimode fiber dispersion by adaptive optics," *Opt. Lett.*, vol. 30, no. 22, pp. 2985–2987, Nov. 2005.
[8] R. A. Panicker, J. M. Kahn, and S. P. Boyd, "Compensation of multimode fiber dispersion using adaptive optics via convex optimization," *J. Lightw. Technol.*, vol. 26, no. 10, pp. 1295–1305, May 2008.
[9] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2003.
[10] R. A. Panicker and J. M. Kahn, "Algorithms for compensation of multimode fiber dispersion using adaptive optics," *J. Lightw. Technol.*, vol. 27, no. 24, pp. 5790–5799, Dec. 2009.
[11] R. A. Panicker, A. P. T. Lau, J. P. Wilde, and J. M. Kahn, "Experimental compensation of adaptive optics algorithms in 10-Gb/s multimode fiber systems," *J. Lightw. Technol.*, vol. 27, no. 24, pp. 5783–5789, Dec. 2009.
[12] M. B. Shemirani, W. Mao, R. A. Panicker, and J. M. Kahn, "Principal modes in graded-index multimode fiber in presence of spatial- and polarization-mode coupling," *J. Lightw. Technol.*, vol. 27, no. 10, pp. 1248–1261, May 2009.
[13] M. B. Shemirani and J. M. Kahn, "Compensation of multimode fiber dispersion by optimization of launched amplitude, phase, and polarization," *J. Lightw. Technol.*, submitted for publication.
[14] 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.
[15] J. M. Kahn, W. J. Krause, and J. B. Carruthers, "Experimental characterization of non-directed indoor infrared channels," *IEEE Trans. Commun.*, vol. 43, no. 4, pp. 1613–1623, Apr. 1995.
[16] S. C. Rashleigh, "Origins and control of polarization effects in single-mode fibers," *J. Lightw. Technol.*, vol. LT-1, no. 2, pp. 312–331, Jun. 1983.
[17] M. Grant, S. Boyd, and Y. Ye, "CVX: Matlab Software for Disciplined Convex Programming," [Online]. Available: http://www.stanford.edu/~boyd/cvx

Mahdieh B. Shemirani received the B.Sc. degree from Sharif University of Technology, Tehran, Iran, in 2004, and the M.Sc. and Ph.D. degrees from Stanford University, Stanford, CA, in 2006 and 2010, respectively, all in electrical engineering.

She joined Intel Corporation in March 2010, where she is currently a Senior Software Engineer in computational lithography.

Jeffrey P. Wilde (M'96) received the M.S. and Ph.D. degrees in applied physics from Stanford University, Stanford, CA, in 1989 and 1992, respectively. His thesis work involved the growth and characterization of photorefractive crystals for multiplex holography.

He has cofounded three companies having core technologies centered around optics and photonics. In 1994, he started 3-D Technology Labs to pursue materials development for holographic data storage. In 1996 he cofounded Quinta Corporation, a data-storage company based on a new type of high-capacity magneto-optic disk drive technology that utilized flying optical heads with single-mode fiber for light transmission and MEMS mirrors for high-bandwidth track following. Quinta was acquired by Seagate Technology in 1997, and he subsequently served as Director of Research West for Seagate. He left Seagate in 2000 and started Capella Photonics to develop wavelength switching technology for the optical telecommunications industry. Capella offers reconfigurable optical add-drop products for both metro and long-haul applications. Since 2003, he has served as a consultant and advisor for a range of companies. He is currently a Visiting Scholar with the Electrical Engineering Department, Stanford University, where he is working on high-speed fiber communication technology.

Dr. Wilde is a member of the Optical Society of America.

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

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

Prof. Kahn was the recipient of the National Science Foundation Presidential Young Investigator Award in 1991. From 1993 to 2000, he served as a Technical Editor of *IEEE Personal Communications Magazine*.