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

Mahdieh B. Shemirani and Joseph M. Kahn, *Fellow, IEEE*

Abstract— **In previous work, a technique was proposed for compensation of modal dispersion in multimode fiber (MMF) systems by using adaptive optics. A spatial light modulator (SLM), with amplitude and phase adjustable over discrete blocks, was used to control the field launched into a MMF. The system impulse response was found as a function of the SLM settings and the field patterns of the MMF principal modes, which were assumed to be known. Using principles of convex optimization, an optimal solution for the SLM settings was obtained, which minimizes intersymbol interference (ISI) subject to physical constraints. Here, we extend the previous method to include control over launched polarization. We propose five adaptive optics configurations, some permitting control only of overall polarization, and others permitting block-wise polarization control. For each configuration, we find optimal adaptive optics settings, which minimize ISI. We show that the best performance is obtained by two configurations that permit independent block-wise control of amplitude, phase and polarization.**

I. INTRODUCTION

Multimode fiber (MMF) is widely used for data communication 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 with different group velocities. MMF imperfections, 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.

In high-bit-rate systems, electrical equalization is commonly used to mitigate ISI caused by modal dispersion [2],[3]. While electrical equalization can extend the achievable 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]. It was then demonstrated experimentally in [7] and later in [8]. 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 [9], 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) [10],[9].

Experiments have demonstrated that the intensity impulse response of a MMF can depend strongly on launched polarization [7]. This polarization dependence was explained by the observation that perturbations in a MMF cause both spatial- and polarization-mode coupling, and was modeled numerically in [11]. The optimization method in [9] did not take account of launched polarization, and hence, cannot obtain optimal performance in fibers with substantial polarization dependence.

In this paper, we extend the work of [9] to include polarization. We find an expression for the MMF impulse response valid to first order in frequency [11],[12], which exhibits its dependence on the PM field patterns and GDs, which are assumed known, and on the amplitude, phase and polarization of the launched field pattern, which is assumed piecewise-constant over discrete blocks. We consider five different adaptive optics configurations, some permitting only control of overall polarization, and others permitting block-wise polarization control. For each configuration, we determine an optimal solution. Each configuration poses different constraints and results in a different optimization problem. These problems are not convex in terms of all optimization variables. However, using certain methods, such as alternating optimization [10] and a multi-dimensional search, we can cast the problem in the format of the convex SOCP problem previously solved in [9]. We prove that the best performance is achieved by two configurations permitting independent block-wise control of amplitude, phase and polarization. We compare the performance of the five adaptive optics configurations with optimized settings through numerical modeling of MMFs

This research was sponsored by National Science Foundation Grant ECCS-0700899.

Mahdieh B. Shemirani and Joseph M. Kahn are with the Department of Electrical Engineering, Stanford University, 350 Serra Mall, Stanford, CA 94305 USA. E-mail: mahdieh@stanford.edu, jmk@ee.stanford.edu.

with both spatial- and polarization-mode coupling, using the model of [11]. Our numerical results confirm our theoretical proof of the optimality of block-wise polarization control.

The remainder of this paper is organized as follows. In Section II, we formulate the problem. Assuming a general adaptive optics system that can control the launched amplitude, phase and polarization over discrete blocks, and assuming a MMF with known input PMs, we derive an expression for the MMF intensity impulse response. We quantify ISI through an objective function to be optimized over the variables. In Section III, we describe five different adaptive optics configurations, and show how to optimize the variables in each configuration. We provide a general theoretical performance comparison of the five configurations. In Section IV, we present numerical simulations comparing their performance on MMFs with spatial- and polarization-mode coupling. We present conclusions in Section V.

II. GENERAL OPTIMIZATION PROBLEM

A. Problem Characterization

Fig. 1 shows the adaptive transmission first described in [6],[8]. 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, polarization controllers, and other components. The beam is partitioned into a set of disjoint blocks, and the adaptive optics system provides block-wise control of its amplitude, phase and polarization. The output beam is imaged 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.

We assume that the collimated beam incident upon the SLM is polarized along the y axis with field pattern $A_0(x,y)\hat{y}$. Assume that the SLM is partitioned into N disjoint blocks (typically squares). We define an indicator function for the jth block [9]:

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

Obviously, the $\{s_j(x,y), j=1,...,N\}$ form an orthogonal set. The imaging function performed by the second lens is described by a linear spatial operator $\mathcal{L}$. The field launched into the MMF can be expressed as:

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

The coefficients u_j and v_j describe the complex reflectance of the jth SLM block in the x and y polarizations, respectively. Defining the contribution to the launched field from the jth block:

$$k_j(x,y)=\mathcal{L}\left[A_0(x,y)s_j(x,y)\right], \quad (3)$$

we can rewrite (2) as:

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

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 τ_i, $i=1,...,2M$. Throughout this study, we assume the PM field patterns and GDs are known. In order to find the first-order impulse response, we project $\vec{A}_{launched}(x,y)$ into these PMs [12]. The ith PM is excited with amplitude:

$$\mu_i=\xi\vec{\Phi}_i(x,y)\cdot\vec{A}_{launched}(x,y) =\xi\left[\Phi_{ix}(x,y)\cdot\sum_{j=1}^{N}u_j k_j(x,y)+\Phi_{iy}(x,y)\cdot\sum_{j=1}^{N}v_j k_j(x,y)\right], \quad (5)$$

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 in the next subsection. The first-order impulse response is:

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

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 our 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:

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

and also the vectors:

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

The impulse response can be expressed as:

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

[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.

where the superscript ^{H} denotes Hermitian conjugate. Expanding the square term in (9), we can express the impulse response as:

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

Following [8],[9], we assume that 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 impulse response is $r(t)$. The receiver output waveform is $g(t) = p(t) * h(t) * r(t)$, which is the 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 the discrete-time system impulse response $g(nT;t_0) = g(t)|_{t=nT}$, $n = \cdots,-1,0,1,\cdots$. We define an objective function quantifying ISI [13]:

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

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 depends on $g(nT;t_0)$ only through f [13]. Defining $q(t) = p(t) * r(t)$, we can write f as:

$$f = \begin{bmatrix} u^H & v^H \end{bmatrix} Q \begin{bmatrix} u \\ v \end{bmatrix}, \quad (12)$$

where:

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

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

Q is a $2N \times 2N$ Hermitian matrix, and, therefore, has N real eigenvalues. The problem we intend to solve is:

$$\text{maximize } \begin{bmatrix} u^H & v^H \end{bmatrix} Q \begin{bmatrix} u \\ v \end{bmatrix} \quad (14)$$

subject to constraints imposed by the system configuration

.

B. Problem Normalization

For convenience, we will normalize the problem so that when there is no loss in the adaptive optical system, the power launched into the MMF is unity. The total launched power is the time integral of the impulse response (6), which corresponds to the sum of the powers carried by the individual PMs:

$$S_{launched} = \sum_{i=1}^{2M} |\mu_i|^2. \quad (15)$$

Expanding $|\mu_i|^2$, we can rewrite (15) as:

$$S_{launched} = |\xi|^2 \left[u^H \left(\sum_{i=1}^{2M} a_i a_i^H \right) u + u^H \left(\sum_{i=1}^{2N} a_i b_i^H \right) v + v^H \left(\sum_{i=1}^{2M} b_i a_i^H \right) u + v^H \left(\sum_{i=1}^{2M} b_i b_i^H \right) v \right]. \quad (16)$$

Using the properties of PMs, (16) can be simplified. We define the $2L \times 2M$ matrix of PMs as:

$$\Phi = \begin{bmatrix} \Phi_{1x} & \cdots & \Phi_{2Mx} \\ \Phi_{1y} & \cdots & \Phi_{2My} \end{bmatrix}. \quad (17)$$

As mentioned above, Φ_{ix} and Φ_{iy} are $L \times 1$ vectors of samples of the PM field components $\Phi_{ix}(x,y)$ and $\Phi_{iy}(x,y)$. Let H be a $L \times 2M$ matrix of the samples of the ideal modes of the MMF, i.e., the eigenmodes in the absence of mode coupling. For example, if the MMF has a parabolic index profile, these are Hermite-Gaussian modes. We can express the PMs by writing:

$$\begin{aligned} \Phi_{ix} &= HP_{ix} \\ \Phi_{iy} &= HP_{iy}, \end{aligned} \quad (18)$$

where P_{ix} and P_{iy} are $2M \times 1$ vectors of PMs represented in the basis of the ideal modes. The orthonormality of the PMs dictates that $\sum_{i=1}^{2M} P_{ix} P_{ix}^H = I$, $\sum_{i=1}^{2M} P_{iy} P_{iy}^H = I$, $\sum_{i=1}^{2M} P_{iy} P_{ix}^H = 0$. Hence:

$$\Phi\Phi^H = \begin{bmatrix} HP_{1x} & \cdots & HP_{2Mx} \\ HP_{1y} & \cdots & HP_{2My} \end{bmatrix} \begin{bmatrix} P_{1x}^H H^H & P_{1y}^H H^H \\ \vdots & \vdots \\ P_{2Mx}^H H^H & P_{2My}^H H^H \end{bmatrix}$$

can be written as:

$$\Phi\Phi^H = \begin{bmatrix} HH^H & 0 \\ 0 & HH^H \end{bmatrix}. \quad (19)$$

Using (7) and (19) we find:

$$\sum_{i=1}^{2M} a_i b_i^H = \sum_{i=1}^{2M} \begin{bmatrix} k_1^H \Phi_{ix} \\ \vdots \\ k_N^H \Phi_{ix} \end{bmatrix} \begin{bmatrix} \Phi_{iy}^H k_1 & \cdots & \Phi_{iy}^H k_N \end{bmatrix} = 0 \quad (20)$$

$$\sum_{i=1}^{2M} b_i a_i^H = 0$$

and, similarly:

$$\sum_{i=1}^{2M} a_i a_i^H = \sum_{i=1}^{2M} b_i b_i^H = \begin{bmatrix} k_1^H HH^H k_1 & \cdots & k_1^H HH^H k_N \\ \vdots & \ddots & \vdots \\ k_N^H HH^H k_1 & \cdots & k_N^H HH^H k_N \end{bmatrix} = K^H HH^H K \quad (21)$$

where $K = \begin{bmatrix} k_1 & \cdots & k_N \end{bmatrix}$ is a $L \times N$ matrix. Using (20) and (21), we can write (16) as:

$$S_{launched} = |\xi|^2 \left(u^H K^H HH^H Ku + v^H K^H HH^H Kv \right) = |\xi|^2 \left(\left\| H^H Ku \right\|^2 + \left\| H^H Kv \right\|^2 \right). \quad (22)$$

For a lossless adaptive optics system we would like to have $S_{launched} = 1$. In order to find the normalization constant ξ, we can consider setting the adaptive optics system so that the light is launched into the MMF in one polarization, say *x* (meaning that $v = 0$), and with the SLMs set to unit reflectivity (meaning that $u = 1_{N\times 1}$, where $1_{N\times 1} = \begin{bmatrix} 1 & \cdots & 1 \end{bmatrix}^T$). We find that the normalization constant must have magnitude

$$|\xi| = \left\| H^H K 1_{N\times 1} \right\|^{-1}. \quad (23)$$

III. OPTIMAL SOLUTIONS FOR PROPOSED CONFIGURATIONS

Here we propose five different configurations for the adaptive optical system, and for each case we show how to obtain the optimal solution by transforming the problem into a previously solved problem called spatial light modulator optimization (SLMO), which was solved in [9].

A. One Amplitude/Phase SLM and Uniform Polarization Controller

Fig. 2(a) shows configuration A, which employs a SLM providing independent, block-wise control of amplitude and phase, followed by a polarization controller, which uniformly transforms the reflected light from a *y*-polarization to an arbitrary elliptical state. In this case:

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \gamma_x \vartheta \\ \gamma_y \vartheta \end{bmatrix}, \quad (24)$$

where ϑ is a complex vector of SLM block reflectances, and γ_x and γ_y are complex variables representing the elliptical polarization, with the constraint $|\gamma_x|^2 + |\gamma_y|^2 = 1$. Hence, the optimization problem is:

$$\begin{aligned} &\text{maximize } f_A = \begin{bmatrix} \gamma_x^* \vartheta^H & \gamma_y^* \vartheta^H \end{bmatrix} Q \begin{bmatrix} \gamma_x \vartheta \\ \gamma_y \vartheta \end{bmatrix} \\ &\text{subject to } |\vartheta_i|^2 \le 1 \quad \forall i = 1,\ldots,N \\ &\qquad\qquad |\gamma_x|^2 + |\gamma_y|^2 = 1 \end{aligned} \quad (25)$$

This problem is not convex in terms of the three variables $(\gamma_x, \gamma_y, \vartheta)$, but a solution can be found using an alternating optimization method [10]. By fixing ϑ, the problem becomes convex in terms of (γ_x, γ_y), and by fixing (γ_x, γ_y), the problem becomes convex in terms of ϑ. By alternately performing these optimizations, the optimal solution is found. Consider $Q = \begin{bmatrix} Q_1 & Q_2 \\ Q_2^H & Q_3 \end{bmatrix}$, where each of the submatrices is a $N \times N$ matrix.

Fixing the polarization, the optimization problem is:

$$\begin{aligned} &\text{maximize } f_A|_{\gamma_x, \gamma_y \text{ constant}} \\ &\quad = \vartheta^H \left(|\gamma_x|^2 Q_1 + \gamma_y^* \gamma_x Q_2^H + \gamma_x^* \gamma_y Q_2 + |\gamma_y|^2 Q_3 \right) \vartheta \\ &\text{subject to } |\vartheta_i|^2 \le 1 \qquad \forall i = 1,\ldots,N \end{aligned} \quad (26)$$

Using arguments similar to those in [9], one can show that since the SLM cannot control polarization, the matrix inside the parentheses has exactly one positive eigenvalue and $N-1$ non-positive eigenvalues. Hence this problem is a SLMO problem, for which the exact solution is known [9].

Fixing ϑ, the problem becomes:

$$\begin{aligned} &\text{maximize } f_A|_{\vartheta \text{ constant}} = = \begin{bmatrix} \gamma_x^* & \gamma_y^* \end{bmatrix} J_A \begin{bmatrix} \gamma_x \\ \gamma_y \end{bmatrix}, \\ &\text{subject to } |\gamma_x|^2 + |\gamma_y|^2 = 1 \end{aligned} \quad (27)$$

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

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

Performing an eigenvalue decomposition $J_A = V \Lambda V^H$, where $\Lambda = \begin{bmatrix} \lambda_{\min} & 0 \\ 0 & \lambda_{\max} \end{bmatrix}$, the optimal polarization can be found as $\begin{bmatrix} \gamma_x^{opt} \\ \gamma_y^{opt} \end{bmatrix} = V \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

B. One Amplitude/Phase/Polarization SLM

Fig. 2(b) shows configuration B, which employs a SLM that can provide independent, block-wise control of amplitude, phase and polarization. In this case:

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \vartheta_x \\ \vartheta_y \end{bmatrix}, \tag{28}$$

where ϑ_x and ϑ_y are complex vectors of SLM block reflectances for the *x* and *y* polarizations, with the constraint $|\vartheta_{xi}|^2 + |\vartheta_{yi}|^2 \le 1$, $i = 1,\ldots,N$. The optimization problem is:

$$\begin{aligned} &\textbf{maximize} \quad f_B = \begin{bmatrix} \vartheta_x^H & \vartheta_y^H \end{bmatrix} Q \begin{bmatrix} \vartheta_x \\ \vartheta_y \end{bmatrix} \\ &\textbf{subject to} \quad |\vartheta_{xi}|^2 + |\vartheta_{yi}|^2 \le 1 \quad \forall i = 1,\ldots,N \end{aligned} . \tag{29}$$

In order to proceed, for definiteness, we assume the adaptive optics system can open the eye using only one of the two available polarizations (either *x* or *y*). In other words, it can open the eye even when a linear polarizer, oriented along *x* or *y*, is placed at its output. Hence the matrix Q has two positive eigenvalues and $N-2$ non-positive eigenvalues. Therefore, we can write Q as:

$$Q = \psi_1 \psi_1^H + \psi_2 \psi_2^H - \Psi_3 \Psi_3^H, \tag{30}$$

where $\psi_2, \psi_1 \in \mathbb{C}^{N\times 1}$ and $\Psi_3 \in \mathbb{C}^{N\times N-2}$, and where $\Psi_3^H \psi_i = 0$ for $i = 1,2$ and $\psi_1^H \psi_2 = 0$. Defining $\vartheta = \begin{bmatrix} \vartheta_x \\ \vartheta_y \end{bmatrix}$, we can rewrite (29) as:

$$\begin{aligned} &\textbf{maximize} \quad f_B = |\psi_1^H \vartheta|^2 + |\psi_2^H \vartheta|^2 - \|\Psi_3^H \vartheta\|^2 \\ &\textbf{subject to} \quad |\vartheta_{xi}|^2 + |\vartheta_{yi}|^2 \le 1 \quad \forall i = 1,\ldots,N \end{aligned} . \tag{31}$$

Because of the two positive terms in the objective function, this problem cannot be cast as a SLMO problem with conical constraint.[2] However, by fixing one of the terms, e.g., $|\psi_2^H \vartheta|^2$, the problem becomes a SLMO. In this case, by performing a two-dimensional search over possible values of the complex variable $\psi_2^H \vartheta$, we can find the optimal solution. Hence (31) can be rewritten as:

$$\begin{aligned} &\textbf{search over } d \qquad \forall d \in \mathbb{C} \\ &\textbf{maximize} \quad |\psi_1^H \vartheta|^2 - \|\Psi_3^H \vartheta\|^2 \\ &\textbf{subject to} \quad |\vartheta_{xi}|^2 + |\vartheta_{yi}|^2 \le 1 \quad \forall i = 1,\ldots,N \\ &\qquad\qquad \mathrm{Re}(\psi_2^H \vartheta) = \mathrm{Re}(d) \\ &\qquad\qquad \mathrm{Im}(\psi_2^H \vartheta) = \mathrm{Im}(d) \\ &\textbf{repeat} \end{aligned} , \tag{32}$$

where d is a complex variable whose magnitude varies from zero to a maximum, which can be found from:

$$\begin{aligned} &\textbf{maximize} \quad |d|^2 = \vartheta^H \psi_2 \psi_2^H \vartheta \\ &\textbf{subject to} \quad |\vartheta_{xi}|^2 + |\vartheta_{yi}|^2 \le 1 \;\; \forall i = 1,\ldots,N \end{aligned} . \tag{33}$$

C. Two Amplitude/Phase SLMs with Equal Amplitude Allocation Between Polarizations

The remaining three configurations use two SLMs, SLM_x and SLM_y, which provide block-wise control of amplitude and phase in the *x* and *y* polarizations, respectively. The configurations differ in the way that amplitude is allocated between the two polarizations. 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 SLM_x (Figs. 2(c)-(e) are drawn assuming the former).

Fig. 2(c) shows configuration C, which uses a uniform, fixed beam splitter to allocate equal amplitudes to the *x* and *y* polarizations in each block. We can write:

$$\begin{bmatrix} u \\ v \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} \vartheta_x \\ \vartheta_y \end{bmatrix}, \tag{34}$$

where ϑ_x and ϑ_y are complex vectors of block reflectances for SLM_x and SLM_y, respectively. The resulting optimization problem is:

$$\begin{aligned} &\text{maximize} \quad f_C = \begin{bmatrix} \vartheta_x^H & \vartheta_y^H \end{bmatrix} \frac{1}{2} Q \begin{bmatrix} \vartheta_x \\ \vartheta_y \end{bmatrix} \\ &\text{subject to} \quad |\vartheta_{xi}|^2 \le 1 \quad \forall i = 1,\ldots,N \\ &\qquad\qquad |\vartheta_{yi}|^2 \le 1 \quad \forall i = 1,\ldots,N \end{aligned} . \tag{35}$$

As in the case of configuration B, for definiteness, we assume the adaptive optics system can open the eye using only one of the two available polarizations. Hence, Q has two positive eigenvalues and $N-2$ non-positive eigenvalues. The solution of this problem is the same as that of (29), but with slightly different constraints. Assuming the vector of reflectances for the two SLMs is $\vartheta = \begin{bmatrix} \vartheta_x \\ \vartheta_y \end{bmatrix}$, the optimization is:

[2] It is also possible that both the *x* and *y* polarizations are required to open the eye. In this case, the matrix Q has one positive eigenvalue and $N-1$ non-positive eigenvalues. The objective function f_B can be optimized using a method similar to the SLMO described in [9].

$$\begin{aligned} &\text{search over } d && \forall d \in \mathbb{C} \\ &\text{maximize} && \frac{1}{2}\left|\psi_1^H \vartheta\right|^2 - \frac{1}{2}\left\|\Psi_3^H \vartheta\right\|^2 \\ &\text{subject to} && \left|\vartheta_i\right|^2 \le 1 \quad \forall i = 1,\ldots,N \\ & && \mathrm{Re}\left(\psi_2^H \vartheta\right) = \mathrm{Re}(d) \\ & && \mathrm{Im}\left(\psi_2^H \vartheta\right) = \mathrm{Im}(d) \\ &\text{repeat} \end{aligned} \quad (36)$$

where the maximum value for magnitude of *d* is found from:

$$\begin{aligned} &\textbf{maximize} && \left|d\right|^2 = \vartheta^H \psi_2 \psi_2^H \vartheta \\ &\textbf{subject to} && \left|\vartheta_i\right|^2 \le 1 \quad \forall i = 1,\ldots,N \end{aligned} \quad (37)$$

which can be solved easily as: $\left|d_{\max}\right|^2 = \lambda_{\max}\left(\psi_2 \psi_2^H\right)$.

D. Two Amplitude/Phase SLMs with Adaptive Uniform Amplitude Allocation Between Polarizations

Figure 2(d) shows configuration D, which uses a uniform, variable beam splitter to allocate amplitude ratios $\sqrt{\varepsilon}$ and $\sqrt{1-\varepsilon}$ to the *x* and *y* polarizations, respectively. In this case:

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \sqrt{\varepsilon}\,\vartheta_x \\ \sqrt{1-\varepsilon}\,\vartheta_y \end{bmatrix}, \quad (38)$$

and the optimization problem is:

$$\begin{aligned} &\text{maximize} && f_D = \begin{bmatrix} \sqrt{\varepsilon}\,\vartheta_x^H & \sqrt{1-\varepsilon}\,\vartheta_y^H \end{bmatrix} Q \begin{bmatrix} \sqrt{\varepsilon}\,\vartheta_x \\ \sqrt{1-\varepsilon}\,\vartheta_y \end{bmatrix} \\ &\text{subject to} && \left|\vartheta_{xi}\right|^2 \le 1 \quad \forall i = 1,\ldots,N \\ & && \left|\vartheta_{yi}\right|^2 \le 1 \end{aligned} \quad . (39)$$

Similar to configuration A, this problem is non-convex in terms of $(\vartheta_x, \vartheta_y, \varepsilon)$, but can be cast as a SLMO by using an alternating optimization. By fixing ε, the problem becomes:

$$\begin{aligned} &\text{maximize} && f_D\big|_{\varepsilon \text{ constant}} = \begin{bmatrix} \vartheta_x^H & \vartheta_y^H \end{bmatrix} U \begin{bmatrix} \vartheta_x \\ \vartheta_y \end{bmatrix} \\ &\text{subject to} && \left|\vartheta_{xi}\right|^2 \le 1 \quad \forall i = 1,\ldots,N \\ & && \left|\vartheta_{yi}\right|^2 \le 1 \end{aligned} \quad , \quad (40)$$

where

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

which can be written as $U = RQR$, where R is a diagonal matrix with elements:

$$R_{ii} = \begin{cases} \sqrt{\varepsilon} & i = 1,\ldots,N \\ \sqrt{1-\varepsilon} & i = N+1,\ldots,2N \end{cases}.$$

Theorem 3.1: For a Hermitian matrix $Q \in \mathbb{C}^{N\times N}$ with p positive eigenvalues and $N-p$ non-positive eigenvalues, and a diagonal positive-definite matrix R, the total number of positive and non-positive eigenvalues for RQR^H are p and $N-p$, respectively.

Proof: This is a special case of the Sylvester Law of Inertia [14]. The theory indicates that for a nonsingular matrix X and a symmetric matrix Z, Z and $X^H Z X$ have the same inertia. The inertia of a symmetric matrix Z is defined as the triplet of integers (m, z, p) with *m*, *z*, *p* being the number of negative, zero and positive eigenvalues of Z. A proof is given in [14], which can be extended easily to Hermitian matrices.■

Hence, U has two positive eigenvalues and $N-2$ non-positive eigenvalues. Therefore, (40) can be solved in the same way as (35).

As for fixing ϑ and finding the optimal ε, one should notice that $\begin{bmatrix} \sqrt{\varepsilon} \\ \sqrt{1-\varepsilon} \end{bmatrix}$ is real and positive. A real solution can be achieved by observing that the objective function f_D is always real, so we can write objective function in (39) as:

$$f_D = \mathrm{Re}\left(\begin{bmatrix} \sqrt{\varepsilon}\,\vartheta_x^H & \sqrt{1-\varepsilon}\,\vartheta_y^H \end{bmatrix} Q \begin{bmatrix} \sqrt{\varepsilon}\,\vartheta_x \\ \sqrt{1-\varepsilon}\,\vartheta_y \end{bmatrix} \right). \quad (41)$$

We can define the symmetric Jones matrix $J_D \in \mathbb{R}^{2\times 2}$:

$$J_D = \mathrm{Re}\left(\begin{bmatrix} \vartheta_x^H Q_1 \vartheta_x & \vartheta_x^H Q_2 \vartheta_y \\ \vartheta_y^H Q_2^H \vartheta_x & \vartheta_y^H Q_3 \vartheta_y \end{bmatrix} \right). \quad (42)$$

Fixing ϑ, the problem of finding the optimal ε becomes:

$$\begin{aligned} &\text{maximize} && f_D\big|_{\vartheta \text{ constant}} = \rho^{\mathrm{T}} J_D \rho \\ &\text{subject to} && \mathrm{norm}(\rho) \le 1 \\ & && \rho \ge 0 \end{aligned} \quad , \quad (43)$$

where $\rho = \begin{bmatrix} \sqrt{\varepsilon} \\ \sqrt{1-\varepsilon} \end{bmatrix} \in \mathbb{R}^{+2}$.

In order to find a solution for (43) we present two lemmas.

Lemma 3.1: $J_{D21} = J_{D12} = \mathrm{Re}\left(\vartheta_x^H Q_2 \vartheta_y\right) \ge 0$ for a vector $\begin{bmatrix} \vartheta_x \\ \vartheta_y \end{bmatrix}$, which is a solution to (40).

Proof: We prove by counterexample. Extending (40), we can write:

$$f_D\big|_{\varepsilon\ \text{constant}} = \varepsilon\,\mathrm{Re}\left(\vartheta_x^H Q_1 \vartheta_x\right) + (1-\varepsilon)\mathrm{Re}\left(\vartheta_y^H Q_3 \vartheta_y\right) + 2\sqrt{\varepsilon}\sqrt{1-\varepsilon}\,\mathrm{Re}\left(\vartheta_x^H Q_1 \vartheta_y\right) \quad (44)$$

Suppose that after maximizing $f_D\big|_{\varepsilon\ \text{constant}}$ over $\begin{bmatrix}\vartheta_x\\ \vartheta_y\end{bmatrix}$, we find $\mathrm{Re}\left(\vartheta_x^H Q_1 \vartheta_y\right) \le 0$. Now we can choose $-\vartheta_y$ as a new vector of reflectances for SLM$_y$. Examining (44), we see that this would increase the objective function $f_D\big|_{\varepsilon\ \text{constant}}$, which is supposed to be maximized. Hence, we should have $\mathrm{Re}\left(\vartheta_x^H Q_1 \vartheta_y\right) \ge 0$. ■

Lemma 3.2: For a solution $\beta \in \mathbb{R}^2$ to:

$$\begin{aligned} &\textbf{maximize} \quad w\big|_{\vartheta\ \textbf{constant}} = \beta^{\mathbf{T}} J_D \beta \\ &\textbf{subject to} \quad \textbf{norm}(\beta) \le 1 \end{aligned} \quad (45)$$

we have $\beta_1\beta_2 \ge 0$.

Proof: Following reasoning similar to the proof of Lemma 3.1 and observing that $J_{D21} = J_{D12} \ge 0$, we conclude that in order to maximize $w\big|_{\vartheta\ \text{constant}}$, we should have $\beta_1\beta_2 \ge 0$. ■

A solution to (43) can be found by solving (45) and then letting $\rho^{opt} = \left|\beta^{opt}\right|$. This is because as $\beta_1\beta_2 \ge 0$, taking the absolute value would not change $w\big|_{\vartheta\ \text{constant}}$, and the resulting ρ^{opt} still satisfies $\textbf{norm}\left(\rho^{opt}\right) \le 1$, which means that ρ^{opt} is also a solution to (45). Now since (45) is a less constrained version of (43), we should have $w^*\big|_{\vartheta\ \text{constant}} \ge f_D^*\big|_{\vartheta\ \text{constant}}$. ρ^{opt} satisfies all of the constraints in (43) and at the same time causes $f_D\big|_{\vartheta\ \text{constant}}$ to assume its maximum possible value $w^*\big|_{\vartheta\ \text{constant}}$. Hence ρ^{opt} is the solution to (43). The solution to (45) can be easily found by eigenvalue decomposition of J_D.

E. Two Amplitude/Phase SLMs with Adaptive Block-Wise Amplitude Allocation Between Polarizations

Fig. 2(e) shows configuration E, which uses a variable beam splitter to vary the amplitude allocation between *x* and *y* polarizations independently in each block. Assuming $\chi_x, \chi_y \in \mathbb{R}^{N\times 1}$ are vectors of the amplitude ratios allocated to *x* and *y* polarizations, with the condition ${\chi_{xi}}^2 + {\chi_{yi}}^2 \le 1$, $i = 1,\ldots,N$, and defining a diagonal matrix W, where

$$W_{ii} = \begin{cases} \chi_{xi} & i = 1,\ldots,N \\ \chi_{xi} & i = N+1,\ldots,2N \end{cases},$$

we can write:

$$\begin{bmatrix} u \\ v \end{bmatrix} = W \begin{bmatrix} \vartheta_x \\ \vartheta_y \end{bmatrix}. \quad (46)$$

We assume the beam splitter is lossy, so that we can cast the constraint in a convex format. The optimization problem can be written as:

$$\begin{aligned} \text{maximize} \quad & f_E = \begin{bmatrix}\vartheta_x^H & \vartheta_y^H\end{bmatrix} WQW \begin{bmatrix}\vartheta_x \\ \vartheta_y\end{bmatrix} \\ \text{subject to} \quad & {\chi_{xi}}^2 + {\chi_{yi}}^2 \le 1 \quad \forall i = 1,\ldots,N \\ & \left|\vartheta_{xi}\right|^2 \le 1 \\ & \left|\vartheta_{yi}\right|^2 \le 1 \end{aligned} \quad (47)$$

At first glance, it may seem that the solution to (47) would require alternately optimizing the beam splitter and SLM reflectances. However, one can show that configurations E and B represent different implementations of the same adaptive optics system. Therefore, in performing the optimization, we can simultaneously adjust corresponding blocks in the two SLMs and in the variable beam splitter. In order to show this, assume that there exists a solution for configuration B, which is denoted as $\begin{bmatrix}\vartheta_x^B \\ \vartheta_y^B\end{bmatrix}$. By choosing:

$$\begin{bmatrix}\chi_x^E \\ \chi_y^E\end{bmatrix} = \left\|\begin{bmatrix}\vartheta_x^B \\ \vartheta_y^B\end{bmatrix}\right\|$$

$$\angle\begin{bmatrix}\vartheta_x^E \\ \vartheta_y^E\end{bmatrix} = \angle\left(\begin{bmatrix}\vartheta_x^B \\ \vartheta_y^B\end{bmatrix}\right),$$

$$\left\|\begin{bmatrix}\vartheta_x^E \\ \vartheta_y^E\end{bmatrix}\right\| = 1_{2N\times 1}$$

all conditions in (47) are satisfied, so we can conclude that $f_E \ge f_B$. Now we assume a solution for configuration E, denoted as $\begin{bmatrix}\chi_x^E \\ \chi_y^E\end{bmatrix}$ and $\begin{bmatrix}\vartheta_x^E \\ \vartheta_y^E\end{bmatrix}$. We can construct a new solution for configuration B, described in the *i*th block by $\vartheta_{xi}^B = \chi_{xi}^E \vartheta_{xi}^E$ and $\vartheta_{yi}^B = \chi_{yi}^E \vartheta_{yi}^E$. Since that solution satisfies the conditions in (47), we can conclude that $\left(\chi_{xi}^E \vartheta_{xi}^E\right)^2 + \left(\chi_{yi}^E \vartheta_{yi}^E\right)^2 \le 1$, which is actually the necessary condition for optimizing configuration B (see (29)). Hence, $f_E \le f_B$. Therefore we can conclude that configurations B and E yield the same optimal solution, and only implementation considerations would favor one configuration or the other.

F. Performance Comparison of the Configurations

In this subsection, we prove some inequalities for the performance of the five configurations.

1) $f_D \geq f_C$. Configuration C is the same as D when $\varepsilon = \frac{1}{2}$, but by optimizing over ε, we can improve the performance for configuration D.

2) $f_B \geq f_D$. For every solution for configuration D of the form $\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \sqrt{\varepsilon}\vartheta_x^D \\ \sqrt{1-\varepsilon}\vartheta_y^D \end{bmatrix}$, satisfying the constraints in (39), we can say that $\left|\sqrt{\varepsilon}\vartheta_x^D\right|^2 + \left|\sqrt{1-\varepsilon}\vartheta_y^D\right|^2 \leq 1$. Hence the vector $\begin{bmatrix} u \\ v \end{bmatrix}$ satisfies the constraint in (29). This means that it can be a solution for configuration B, yet by maximizing the objective function in (29) it is possible to attain a better solution.Therefore $f_B \geq f_D$.

3) $f_E = f_B$. The proof has been provided in Subsection E.

4) $f_D \geq f_A$. Consider a solution for configuration A of the form $\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \gamma_x \vartheta^A \\ \gamma_y \vartheta^A \end{bmatrix}$, satisfying the condition $|\gamma_x|^2 + |\gamma_y|^2 = 1$. We can choose $\begin{bmatrix} \sqrt{\varepsilon} \\ \sqrt{1-\varepsilon} \end{bmatrix} = \begin{bmatrix} |\gamma_x| \\ |\gamma_y| \end{bmatrix}$ and by choosing the SLM reflectances as $\begin{bmatrix} \vartheta_x^D \\ \vartheta_y^D \end{bmatrix} = \begin{bmatrix} \exp(j\angle\gamma_x)\vartheta^A \\ \exp(j\angle\gamma_y)\vartheta^A \end{bmatrix}$, we obtain a solution that satisfies all the conditions in (39). Hence configuration D can perform better than configuration A.

5) In general, the theory does not indicate whether configuration A is better than configuration C or vice versa. To demonstrate this, consider configuration D, which performs better than both A and C. Suppose we have $Q = \begin{bmatrix} Q_1 & 0 \\ 0 & Q_3 \end{bmatrix}$, where Q_1 and Q_3 are diagonal matrices, each with one positive and $N-1$ negative elements. In this case, the matrix Q is diagonal with two positive and $N-2$ non-positive eigenvalues. The optimal solution for D is obtained by solving:

$$\begin{aligned} \text{maximize} \quad & \varepsilon\left(\vartheta_x^D\right)^H Q_1 \vartheta_x^D + (1-\varepsilon)\left(\vartheta_y^D\right)^H Q_3 \vartheta_y^D \\ \text{subject to} \quad & \left|\vartheta_{xi}^D\right|^2 \leq 1 \quad \forall i = 1,\ldots,N \\ & \left|\vartheta_{yi}^D\right|^2 \leq 1 \\ & 0 \leq \varepsilon \leq 1 \end{aligned} \tag{48}$$

Assume that in configuration D, the adaptive optics system can open the eye using only one of the two available polarizations (either *x* or *y*). Hence, we should have $\eta = \left(\vartheta_x^D\right)^H Q_1 \vartheta_x^D > 0$ and $\phi = \left(\vartheta_y^D\right)^H Q_3 \vartheta_y^D > 0$. Solving for ε, we find that the objective function is maximized if $\begin{cases} \varepsilon = 1 & \text{for } \eta \geq \phi \\ \varepsilon = 0 & \text{for } \eta < \phi \end{cases}$, i.e., all available energy should be sent to the polarization that can yield the more open eye. Suppose we have $\eta < \phi$, and hence, we should choose $\varepsilon = 0$. Configuration A can achieve this solution by polarizing the light along *y*, but configuration C will waste energy by unnecessarily allocating half the energy to the *x* polarization. Hence, configuration C would perform worse than configuration A.

Now to prove that configuration A is not always better than configuration C, suppose for a different Q, the optimal solution for configuration D is $\varepsilon = 0.5$ and $\vartheta_x^D \neq \vartheta_y^D$. These settings can be achieved by configuration C, which employs equal amplitude allocation and two SLMs. But configuration A cannot generally achieve these settings, as it employs just one SLM, and hence we cannot always achieve a solution of the form $\gamma_x \vartheta^A = \vartheta_x^D$, $\gamma_y \vartheta^A = \vartheta_y^D$. Therefore, the general theory does not indicate which of configurations A or C would perform better.

IV. NUMERICAL RESULTS

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 [11],[12]. 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 silica MMF of total length L = 1000 m. The fiber has a numerical aperture NA = 0.19. At the wavelength λ = 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 [8], the refractive index exponent is chosen to be α = 2.17. Birefringence, defined as the difference between refractive indices seen by *x*- and *y*-polarized waves, is assumed to be induced by stress due to curvature [15]:

$$n_x - n_y = \delta \frac{C_s}{2k_0}(a\kappa)^2, \tag{49}$$

where κ is the curvature of a fiber section, and C_s / k_0 is referred to as the strain-optical coefficient. For a SMF, $C_s / k_0 = 0.0878 n^3$ [15], and the birefringence scale factor should be set to δ = 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, whose pdf is the positive side of a normal pdf. Simulations are performed for a fiber in low-coupling regime with a curvature standard deviation of $\sigma_\kappa = 2.53$ m^{-1}, and a fiber in medium-coupling regime with a curvature standard deviation of $\sigma_\kappa = 4.62$ m^{-1}. Here, the curvature standard deviation in both regimes is higher than the value employed in previous studies [11],[12], which considered a refractive index exponent $\alpha = 2.00$. Here, the higher exponent $\alpha = 2.17$ tends to reduce spatial-mode coupling, so higher curvature is required to produce enough ISI to close the eye completely before optimization. The first-order model described in [11] is used to calculate the first-order PMs and their GDs.

We model an adaptive transmission system, as shown in Fig. 1, choosing parameters corresponding to previous experimental studies [8]. 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 numerical aperture NA = 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 imaged onto the MMF by a lens of 10.4-mm focal length. Thus, the beam launched into the MMF has a numerical aperture NA = 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, 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 [8] provided phase control only, here, the SLM(s) are assumed to control amplitude and phase (and polarization, the case of configuration B). A phase-only SLM can control amplitude and phase if high-spatial-frequency patterns, such as a checkerboard pattern of anti-phased pixels within each block, are introduced to diffract light away from the core of the MMF [9].

Optimal solutions are computed numerically using CVX, a freely distributed convex optimization library for MATLAB [16].

B. Simulation Results

Table 1 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 2 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 Sections III.A-E. In each fiber, optimized objective function values are positive, indicating that the eye is open. Objective function values are consistent with the comparisons made in Section III.F. In particular, configuration A yields the worst performance, configurations C and D yield intermediate performance, and configurations B and E (whose optimal solutions are equivalent) yield the best performance.

In what follows, we present results only for the fiber in the medium-coupling regime.

Fig. 3 considers configuration A, which uses a single SLM providing block-wise control of amplitude and phase and a controller of overall polarization. The figure shows initial and optimized polarizations (on the Poincaré sphere), optimized SLM amplitude and phase, initial and optimized impulse responses and initial and optimized eye diagrams. Note that the optimized SLM amplitudes are unity in most, but not all, blocks, indicating that the SLM must discard some signal energy.

Fig. 4 considers configuration B, which uses a single SLM providing block-wise control of amplitude, phase and polarization. The figure shows optimized amplitudes and phases for the two polarizations on the SLM, initial and optimized impulse responses and initial and optimized eye diagrams. Note that in each SLM block, the squares of the optimized amplitudes sum to unity, indicating that the SLM does not discard any signal energy. Obviously, the eye diagram is cleaner and more open than for configuration A.

Figs. 5-7 consider configurations C-E, which employ two SLMs, SLM_x and SLM_y, which provide block-wise control of amplitude and phase in the x and y polarizations, respectively. The three configurations differ in terms of how signal amplitude is allocated between the two SLMs. All three figures show the amplitudes allocated to the two SLMs, optimized amplitudes and phases on the two SLMs, initial and optimized impulse responses and initial and optimized eye diagrams.

Fig. 5 considers configuration C, which uses a uniform, fixed beam splitter to allocate equal amplitudes of = to the two SLMs. In each block, the squares of the amplitudes reflected from the two SLMs[3] do not sum to unity, indicating that some signal energy is discarded by this suboptimal adaptive optics configuration.

[3] The amplitude reflected from an SLM in a given block is the product of the amplitude allocated by the beam splitter and the amplitude setting of the SLM block.

Fig. 6 considers configuration D, which uses a uniform, variable beam splitter to allocate amplitudes to the two SLMs. In this example, the optimized amplitude allocations are $\begin{bmatrix} \sqrt{\varepsilon} \\ \sqrt{1-\varepsilon} \end{bmatrix} = \begin{bmatrix} 0.6247 \\ 0.7808 \end{bmatrix}$. In each block, the squares of the amplitudes reflected from the two SLMs[3] still do not sum to unity, but they are generally higher than for configuration C, and yield slightly better performance than configuration C. Comparing Fig. 6 to Fig. 3, we find that for this particular fiber, configuration D performs better than configuration A, although we showed in Section III.F that this need not be true in general.

Fig. 7 considers configuration E, which uses a variable beam splitter to provide block-wise amplitude allocation to the two SLMs. In each block, amplitudes and phases reflected from the two SLMs are identical to those in configuration B, the squares of the amplitudes reflected from the two SLMs do sum to unity, as in configuration B. The optimized impulse response and eye diagram for configuration E are identical to those obtained using configuration B, and are the best among the five configurations studied.

V. CONCLUSIONS

We have considered polarization as an additional degree of freedom in compensating modal dispersion in MMF systems by using transmitter-based adaptive optics. We have proposed five adaptive optical system configurations, each providing some control of the amplitude, phase and polarization of the launched signal over discrete blocks. We have studied how to optimize these variables, assuming *a priori* knowledge of the fiber's PMs and their GDs. We have shown that for each configuration, optimization of the variables is non-convex. We have used an alternating optimization method and, in some cases, a multi-dimensional search, to convert these problems to a SOCP, which was solved with low computational complexity in previous work. We have performed simulations to evaluate the performance of the five configurations with typical MMFs, showing that all five can mitigate ISI and open an otherwise closed eye, and all five yield better performance than previous methods that did not control polarization. We have found that, among the five configurations studied, the best performance is obtained by two configurations that permit independent block-wise control of amplitude, phase and polarization.

Future work will address development and experimental evaluation of algorithms for adaptively optimizing the adaptive optics system for a MMF whose PMs are not known *a priori*.

REFERENCES

[1] G. P. Agrawal, *Fiber-Optic Communications Systems*, 3rd ed., John Wiley & Sons, Inc., 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 Photonics Technology Letters*, 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, pp. 2131-2137, Dec. 2003.

[4] J. Proakis, *Digital Communications*, 4th ed. McGraw-Hill, 2001.

[5] S. Fan and J. M. Kahn, "Principal modes in multi-mode waveguides", *Optics Letters*, 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", *Proc. of 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", *Optics Letters*, vol. 30, no. 22, pp. 2985-2987, Nov. 2005.

[8] P. L. Neo, J. P. Freeman, and T. D. Wilkinson, "Modal control of a 50 μm core diameter multimode fiber using a spatial light modulator", in *OFC/NFOEC*, Anaheim, CA, pp. 1-3, 25-29 March 2007.

[9] R. A. Panicker, J. M. Kahn, and S. P. Boyd, "Compensation of multimode fiber dispersion using adaptive optics via convex optimization", *J. of Lightwave Technol.*, vol. 26, no. 10, pp. 1295-1305, May 2008.

[10] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge, UK, 2004.

[11] 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. of Lightwave Technol.*, vol. 27, issue 10, pp. 1248-1261, May 2009.

[12] M. B. Shemirani and J. M. Kahn, "Higher-order modal dispersion in graded-index multimode fiber", *J. of Lightwave Technol.*, vol. 27, no. 23, pp. 5461-5468, December 1, 2009.

[13] J. M. Kahn, W. J. Krause, and J. B. Carruthers, "Experimental characterization of non-directed indoor infrared channels", *IEEE Trans. Communications*, vol. 43, pp. 1613-1623, Apr. 1995.

[14] G. H. Golub and C. F. Van Loan, *Matrix Computations*, The John Hopkins University Press, Baltimore, US, 1983.

[15] S. C. Rashleigh, "Origins and control of polarization effects in single-mode fibers", *J. of Lightwave Technol.*, vol. LT-1, no. 2, pp 312-331, June 1983.

[16] M. Grant, S. Boyd, and Y. Ye, cvx: Matlab Software for Disciplined Convex Programming [Online]. Available: http://www.stanford.edu/~boyd/cvx

Table 1: Objective functions obtained using configuration A with blank SLM and three different polarizations for a 1-km, 50-μm-core graded-index MMF in low- and medium-coupling regimes.

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

Table 2: 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.

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

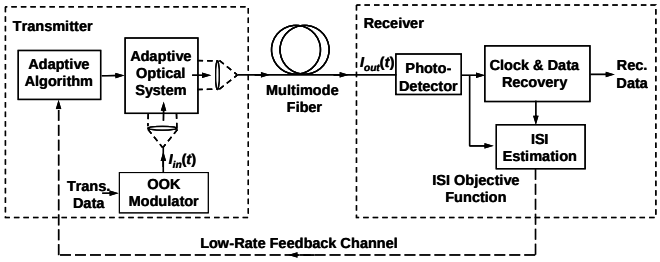

Fig. 1. Adaptive transmission system. A modulated optical signal is passed into an adaptive optical system, whose output is launched into a MMF. At the MMF output, residual ISI is estimated, and the estimate is passed to an adaptive algorithm that controls the adaptive optical system.

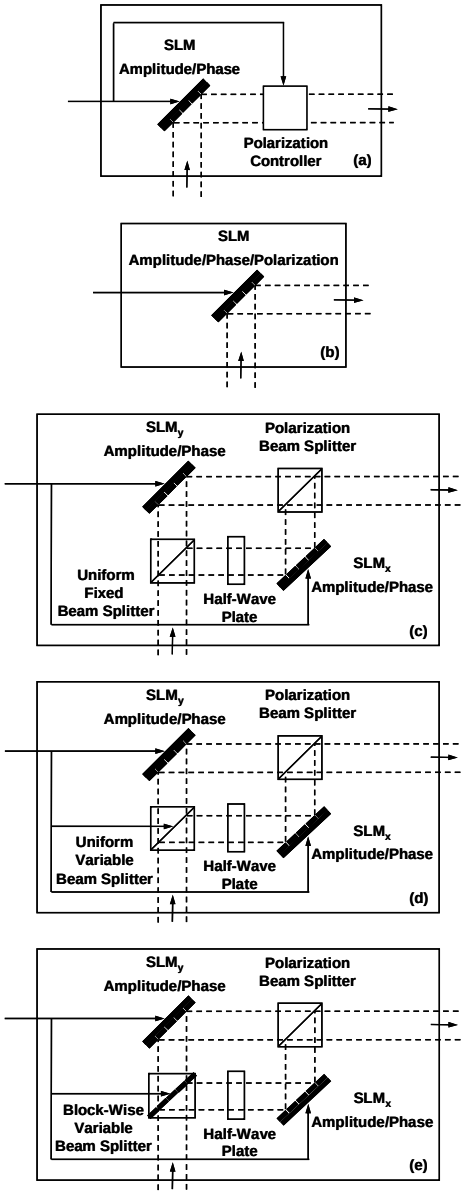

Fig. 2. Adaptive optical system configurations. (a) Configuration A comprises one SLM providing block-wise control of amplitude and phase, and a polarization controller to control the overall polarization. (b) Configuration B comprises one SLM providing block-wise control of amplitude, phase and polarization. (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.

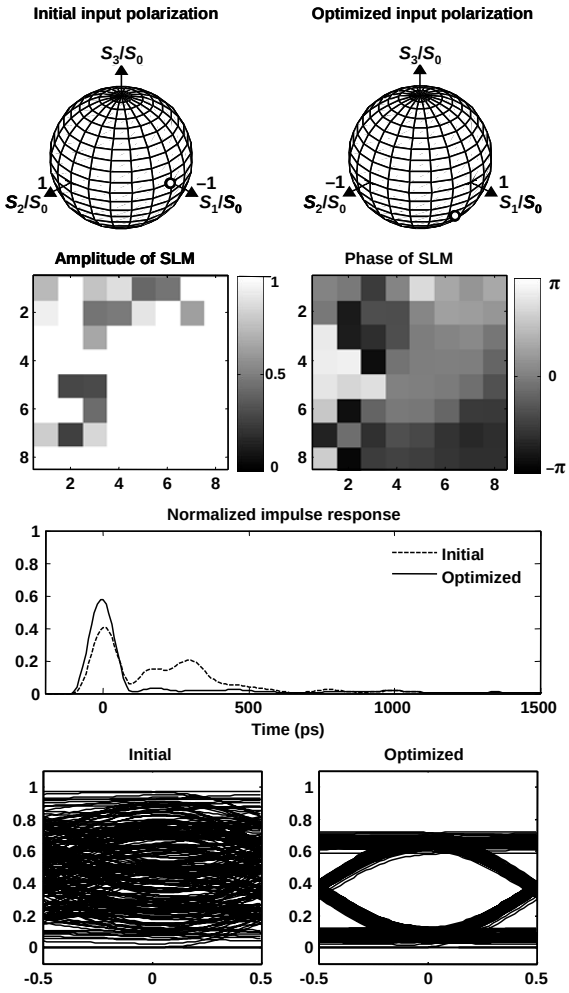

Fig. 3. Configuration A: initial and optimized polarizations, optimized SLM amplitude and phase, initial and optimized impulse responses and eye diagrams for a 1-km, 50-μm-core graded-index MMF in medium-coupling regime.

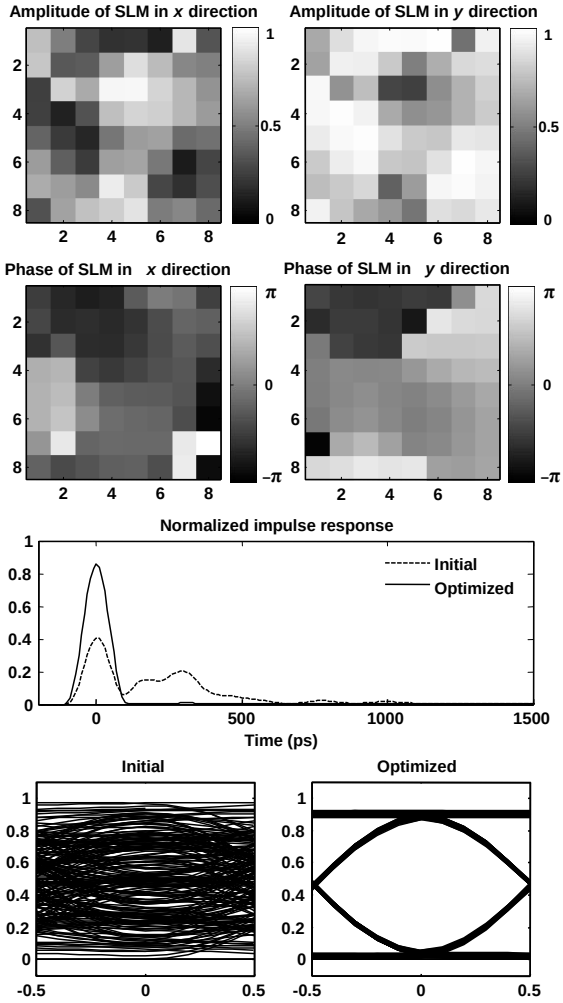

Fig. 4. Configuration B: optimized amplitudes and phases for two polarizations on SLM, initial and optimized impulse responses and eye diagrams for a 1-km, 50-μm-core graded-index MMF in medium-coupling regime**.**

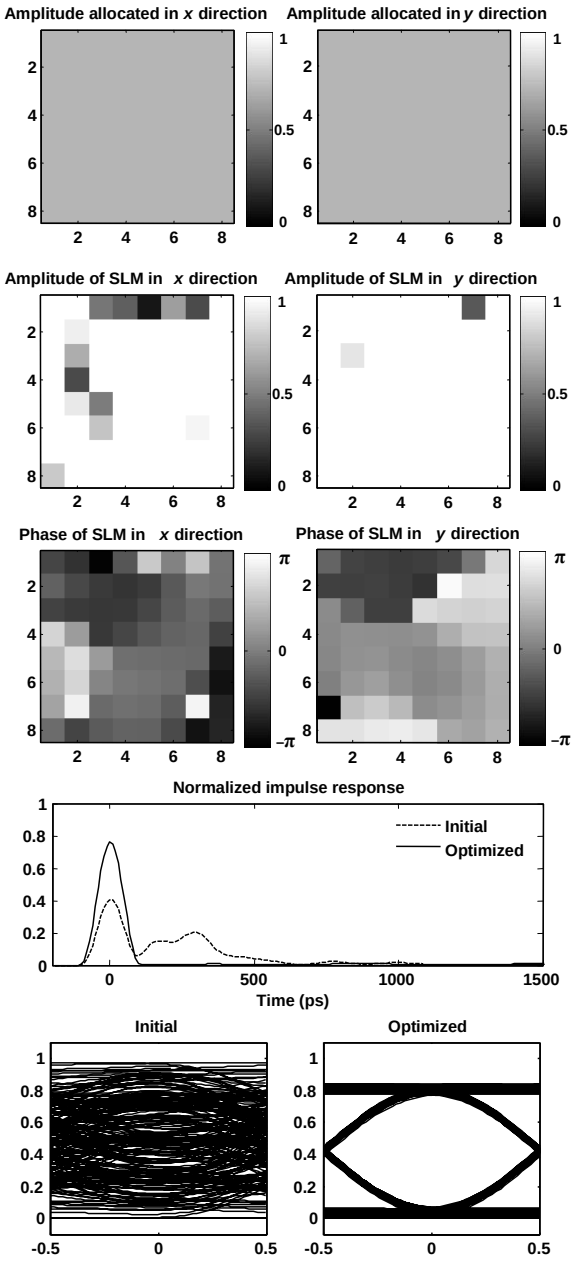

Fig. 5. Configuration C: fixed uniform amplitudes allocated to two SLMs, optimized amplitudes and phases on two SLMs, initial and optimized impulse responses and eye diagrams for a 1-km, 50-μm-core graded-index MMF in medium-coupling regime, $\sigma_\kappa = 4.62$ m^{-1}.

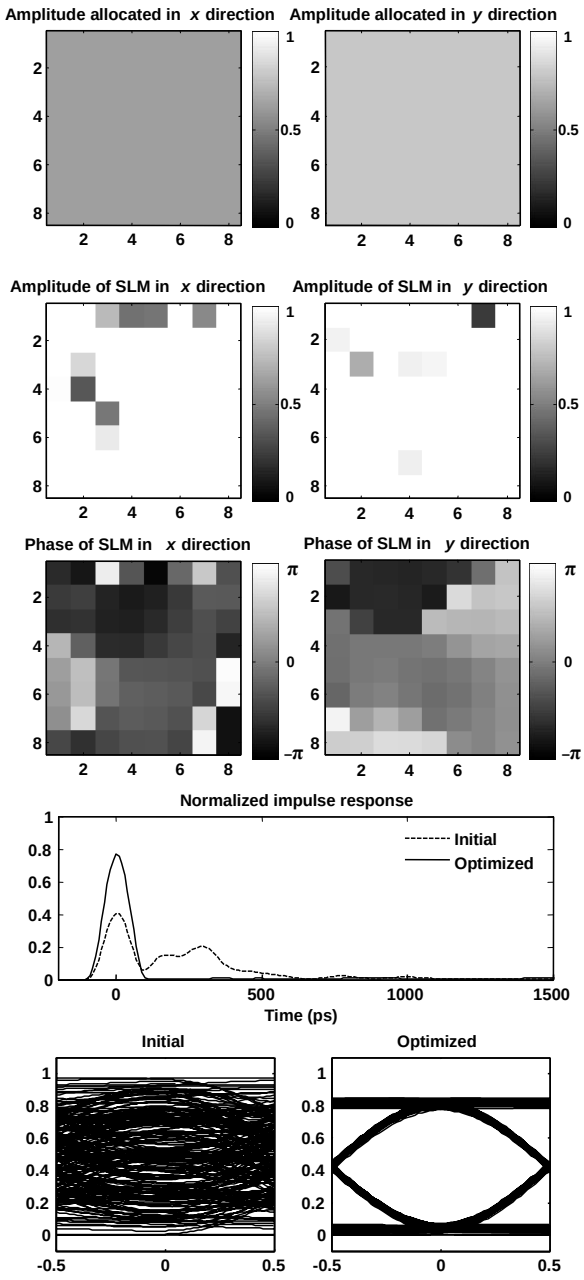

Fig. 6. Configuration D: optimized uniform amplitudes allocated to two SLMs, optimized amplitudes and phases on two SLMs, initial and optimized impulse responses and eye diagrams for a 1-km, 50-μm-core graded-index MMF in medium-coupling regime.

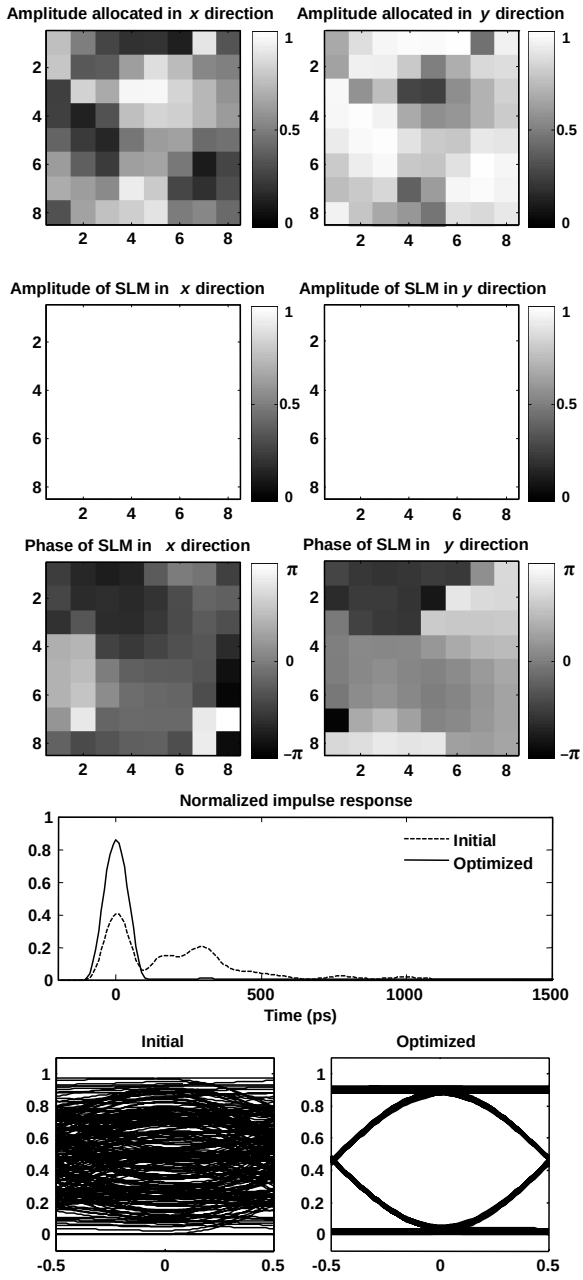

Fig. 7. Configuration E: optimized amplitudes allocated block-wise to two SLMs, optimized amplitudes and phases on two SLMs, initial and optimized impulse responses and eye diagrams for a 1-km, 50-μm-core graded-index MMF in medium-coupling regime.