Algorithms for Compensation of Multimode Fiber Dispersion Using Adaptive Optics

Rahul Alex Panicker and Joseph M. Kahn, *Fellow, IEEE*

Abstract—We propose adaptive algorithms for mitigating inter-symbol interference (ISI) in multimode fiber (MMF) systems using a spatial light modulator (SLM). Minimizing ISI in MMF systems using an SLM has previously been posed as a convex optimization problem. Based on these results, we propose a range of algorithms for adapting the SLM settings. Some of these are shown to converge to the global optimum in the absence of noise. We then propose modified versions of these algorithms to improve resilience to noise and speed of convergence. Simulation results are included, showing that these algorithms open an otherwise closed eye pattern.

Index Terms—Adaptive optics, algorithms, optical fiber dispersion, spatial light modulators.

I. INTRODUCTION

MULTIMODE FIBER (MMF) is the dominant type of fiber used for data communications in current local-area networks. In achieving higher signalling rates, the dominant limiting factor is the inter-symbol interference (ISI) caused by modal dispersion [1]. Light propagates in a MMF in modes, with each mode propagating at its group velocity. The set of modes excited depends on launch conditions at the input of the fiber, and on mode coupling within the fiber. Thus, a pulse of light that excites many modes in the fiber arrives as several pulses at the output of the fiber—a phenomenon known as modal dispersion. This effect is analogous to multipath in wireless.

In the past, electrical equalization [2], [3] has been used to mitigate ISI caused by modal dispersion. However, this can lead to noise enhancement, and, therefore, degradation of achievable bit-error ratio (BER) [4].

As an alternative to electrical equalization, the use of adaptive optical compensation was proposed in [5]. This involves shaping the spatial profile of the electric field at the input end of the fiber using a spatial light modulator (SLM) to excite only desired desired principal modes (PMs). SLM is a two-dimensional (2-D) array of pixels, capable of modifying the local phase and/or amplitude of an incident electric field. PMs are a complete set of orthogonal modes propagating in the fiber, such that a pulse launched into a PM at the input of the fiber emerges as a single pulse at the output, even in the presence of mode

coupling [6]. This technique leads to no noise enhancement. Experiments [7], [8] have shown that this approach is capable of realizing high bit rates.

A comprehensive theoretical framework for analyzing this system was developed in [9]. It was shown that minimizing ISI by changing the SLM settings can be posed as a convex optimization problem. In particular, maximizing eye opening, subject to constraints on the SLM, was cast as a second-order cone program (SOCP). Based on this, globally optimal solutions were computed.

In practical systems, it is difficult to compute the optimal SLM settings directly as the solution of an optimization problem. This is because the solution is a function of a system matrix that captures information about the fiber parameters and exact details of the mode coupling, which in turn depend on every bend, twist, and refractive index imperfection in the fiber. Therefore, the optimal settings can only be found either by explicitly estimating this matrix, or by an adaptive algorithm that implicitly learns these parameters through a feedback loop, and cause the SLM settings to converge to the optimum.

In this paper, we propose adaptive algorithms for setting the SLM, for mitigation of ISI. These are shown to have guaranteed convergence to the global optimum in the absence of noise. We also develop a range of suboptimal algorithms, that trade off suboptimality of the final solution for speed of convergence. We then introduce modified versions of these algorithms to improve resilience to noise.

The remainder of this paper is organized as follows. In Section II, we introduce the transmission scheme, and give the theoretical framework for analysis. This includes the impulse response of the system as a function of SLM settings, and the optimization problem posed in [9] that we wish to solve adaptively. In Section III, we propose an adaptive algorithm that is guaranteed to converge to the global optimum in the absence of noise. We then introduce a suboptimal algorithm, that does not have guaranteed convergence to the global optimum, but performs close to optimal in simulation, and better than currently employed techniques. In Section IV, we introduce modified algorithms with increased robustness to noise. In Section V, we give simulation results, using parameters of commercially available components. We show that our algorithms clean up the impulse response and can open an otherwise closed eye pattern. We compare performance of these algorithms for various levels of noise, degrees of mode coupling, and SLM resolutions.

II. ADAPTIVE TRANSMISSION SYSTEM

We first describe a physical system that enables ISI mitigation using an SLM. Next, we write down the impulse response of the

R. A. Panicker was with the Department of Electrical Engineering, Stanford University, Stanford, CA 94305 USA. He is now with Embrace, Richmond Town, Bangalore 560025, India (e-mail: rahul@embraceglobal.org).

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

system as a function of SLM settings. We then link system performance—the eye opening—to the impulse response, and pose minimization of ISI as an optimization problem. We also look at estimation of the objective function so that this optimization problem may be solved using an adaptive algorithm.

A. Transmission Scheme

The system configuration is shown in Fig. 1, and was originally proposed in [5]. It consists of a transmitter, the MMF itself, and a receiver in which received light pulses are detected and decoded. At the transmitter, we have an SLM followed by an imaging system that focuses the light from the SLM into the input end of the MMF. The SLM is a two-dimensional (2-D) array of pixels, and is used to manipulate the phase and/or amplitude of the electric field incident on the MMF, in the spatial domain. The ISI at the receiver depends on the excitation pattern of modes in the MMF. We, therefore, try to set the SLM pixels to optimally shape the light field incident on the MMF, and selectively excite modes in the MMF. At the receiver, light from the MMF is detected by a photodetector, and decoded by a recovery circuit. The detected signal is also used for ISI estimation, which drives an adaptive algorithm that controls the SLM.

Unlike traditional electrical equalization, the energy of the light signal can be refocused into the desired modes of the fiber, eliminating ISI without amplifying the fixed noise in the system. This can, of course, only be achieved with the proper SLM pixel settings.

B. Impulse Response

With a given SLM setting, a pulse of light entering the MMF has a certain spatial profile. This in turn governs the set of PMs excited, and the distribution of energy across the PMs. Since each PM propagates with a well defined group delay, we get a sequence of pulses at the output. For a unit impulse at the input, this sequence of output pulses is the impulse response. Thus, the impulse response of the system is a function of the SLM settings.

A detailed derivation was presented in [9]. Here, we reproduce some of the key results

$$h(t) = e^{-\alpha l} \sum_{i=1}^{2M} \left| \int \left[\vec{\mathbf{E}}_{\text{fiberIn}}(x,y) \times \vec{\mathbf{H}}_{\text{PM},i}^*(x,y) \right] \right. $$
$$\left. \cdot \hat{z} \, dx \, dy \right|^2 \delta(t - \tau_i) \quad (1)$$

where α is the fiber loss coefficient (approximated as mode independent), l is the fiber length, $2M$ is the number of propagating PMs in both polarizations, $\vec{\mathbf{E}}_{\text{fiberIn}}(x,y)$ is the electric field at the input face of the fiber, $\vec{\mathbf{H}}_{\text{PM},i}(x,y)$ is the normalized magnetic field corresponding to the ith input PM, and τ_i is the group delay of the ith PM. $\vec{\mathbf{E}}_{\text{fiberIn}}(x,y)$ is a function of the SLM reflectances, an exact expression for which is given in [9]. This leads to a final expression of the impulse response of the following form:

$$h(t) = \mathbf{v}^H \left[\sum_{i=1}^{2M} \mathbf{u}_i \mathbf{u}_i^H \delta(t - \tau_i) \right] \mathbf{v} \quad (2)$$
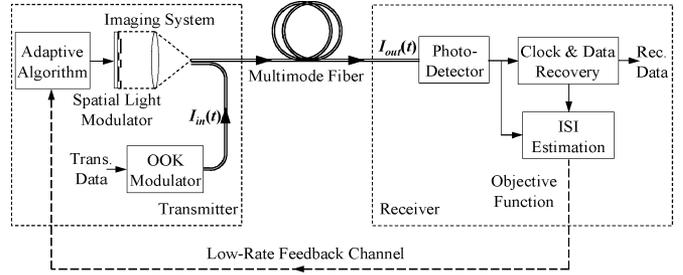

Fig. 1. Adaptive transmission system: output of the single mode fiber is imaged onto the SLM, SLM output is imaged onto MMF, MMF output is detected, and fed to an adaptive algorithm that controls the SLM.

where $\mathbf{v} = [v_1 \, v_2 \ldots v_N]^T, v_j \in \mathbb{C}$ is the reflectance of the jth SLM block, N is the number of SLM blocks, and $\mathbf{u}_i \in \mathbb{C}^N$, $i = 1, \ldots, 2M$ are vectors that capture the spatial structure of the $2M$ PMs [9]. Note that the impulse response is a quadratic function of the SLM reflectances.

C. Objective Function

The pulse response of the system was shown [9] to be

$$g(t) = h(t) * q(t)$$
$$= \mathbf{v}^H \left[\sum_{m=1}^{2M} \mathbf{u}_m \mathbf{u}_m^H q(t - \tau_m) \right] \mathbf{v} \quad (3)$$

where $q(t) = p(t) * r(t)$, with $p(t)$ being the transmit pulse, and $r(t)$ the receiver impulse response. It was also shown that minimization of ISI was equivalent to maximization of the eye opening in the eye diagram generated by the above pulse response. The eye opening is given by

$$F = g(0; t_0) - \sum_{n \neq 0} g(nT; t_0) \quad (4)$$

where T is the bit duration, t_0 is an initial sampling offset, and n is an integer [10], with $n = 0$ corresponding to the desired peak. F is directly proportional to the eye opening, with $F < 0$ when the eye is closed, and $F > 0$ when the eye is open. Thus, minimization of ISI is equivalent to maximization of F.

D. Optimization Problem

It was also shown in [9] that F can be simplified to the form $F = \mathbf{v}^H \mathbf{P} \mathbf{v}$, where $\mathbf{P} \in \mathbb{C}^{N \times N}$ is a Hermitian matrix, with one positive, and $N - 1$ non-positive eigenvalues. $\mathbf{P}$ characterizes the system, and is fixed for a given system. Also, $|v_i| \leq 1$, $i = 1, \ldots, N$, because the SLM is a passive reflective device, and can, therefore, not amplify electric field locally. This leads to an optimization problem of the form

$$\begin{aligned} \text{maximize} \quad & \mathbf{v}^H \mathbf{P} \mathbf{v} \\ \text{subject to} \quad & |v_i| \leq 1, i = 1, \ldots, N \end{aligned} \quad (5)$$

where $\mathbf{v}$, the vector of SLM reflectances, is the variable over which we optimize, and $\mathbf{P}$ is a property of the system, and may not be known explicitly. This is the problem we wish to solve using our adaptive algorithms. Noting that $\mathbf{P}$ may be factorized

as $\mathbf{p}_1\mathbf{p}_1^H - \mathbf{P}_2\mathbf{P}_2^H$, where $\mathbf{p}_1 \in \mathbb{C}^N$, $\mathbf{P}_2 \in \mathbb{C}^{N \times N-1}$, and $\mathbf{P}_2^H\mathbf{p}_1 = 0$, leads to the SLMO problem as posed in [9]

$$\begin{aligned} \text{maximize} \quad & \left|\mathbf{p}_1^H\mathbf{v}\right|^2 - \left\|\mathbf{P}_2^H\mathbf{v}\right\|^2 \\ \text{subject to} \quad & |v_i|^2 \le 1 \quad i = 1, \dots, N. \end{aligned} \tag{6}$$

This problem is not in any standard form. For example, the objective function is not convex. However, in [9], it was shown that SLMO is equivalent to the following convex optimization problem:

$$\begin{aligned} \text{maximize} \quad & y \\ \text{subject to} \quad & \sqrt{\mathbf{x}^H\mathbf{x} + y^2} \le t \\ & \mathbf{p}_1^H\mathbf{v} = t \\ & \mathbf{P}_2^H\mathbf{v} = \mathbf{x} \\ & |v_i|^2 \le 1 \quad i = 1, \dots, N. \end{aligned} \tag{7}$$

Here, $\mathbf{v} \in \mathbb{C}^N$, $\mathbf{x} \in \mathbb{C}^{N-1}$, $t \in \mathbb{R}$ and $y \in \mathbb{R}$ are the optimization variables, and $\mathbf{p}_1$ and $\mathbf{P}_2$ are problem data. This problem is a second-order cone program (SOCP), which can be globally solved with great efficiency [11]. The computational cost of solving this problem is N^3, the same order as solving a set of N linear equations. We will refer to the SOCP (7) as convex-form SLMO. Many publically distributed software packages can be used to solve convex form SLMO, e.g., CVX or SeDuMi.

E. Adaptive Algorithms

We note that in real fiber systems, $\mathbf{P}$ is not known explicitly, since it captures the details of the mode-coupling and the optical system, and depends on the exact refractive profile, every bend and twist in the fiber, and so on. However, if for every SLM setting $\mathbf{v}$, we can observe $\mathbf{v}^H\mathbf{P}\mathbf{v}$, we can still find the optimal solution. Since $\mathbf{v}^H\mathbf{P}\mathbf{v}$ is the eye opening, it can be measured directly.

While it may appear from (4) that estimating F when the eye is closed requires many samples of the impulse response, in practice, it can be estimated from a few samples of the step response. For example, in [7], the following procedure was used. When the eye is closed, a periodic training sequence comprising a long string of 0-bits followed by an equal number of 1–bits is transmitted, as shown in Fig. 2. The largest positive and negative excursions of the received waveform over any interval of duration T is used to estimate $g(0; t_0)$, and the total excursion of the received waveform is used to estimate $\sum g(nT; t_0)$. F is then estimated as

$$F = (y_0 - y_{-1}) + (y_{L-1} - y_L) - (y_{\max} - y_{\min})$$

with $L = 64$ in [7]. When the eye is open, the eye opening itself may be used as an estimate of F.

III. ADAPTIVE ALGORITHMS: NOISELESS SYSTEMS

In this section, we introduce an algorithm involving adaptation of the amplitude and phase of the SLM for solving the optimization problem posed in (5), and prove its convergence to the global optimum. Next, we introduce a suboptimal algorithm involving phase-only adaptation, keeping a uniform unit amplitude over the SLM. We also give other algorithms currently in
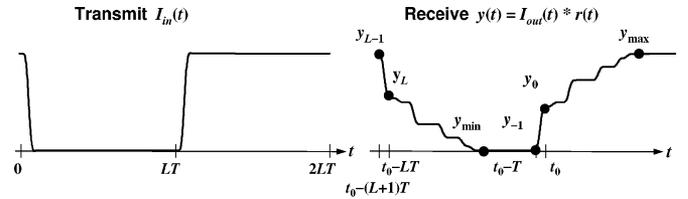

Fig. 2. Estimation of the objective function using six samples from the response to a step training sequence.

use, for comparison. All algorithms that we consider in this section are sequential-coordinate-ascent (SCA) algorithms. In this section, we also assume that the system has no noise.

A. Amplitude-and-Phase SCA

Algorithm 1 Amplitude-and-Phase SCA

1: $v_i := 1, \quad i = 1, \dots, N.$
2: $i := 1$
3: **repeat**
4:     **for** $m = 1$ to 3 **do**
5:       $v_i := \exp(j2\pi(m-1)/(3))$
6:       Estimate F_m
7:     **end for**
8:     $v_i := 0$
9:     Estimate F_4
10:    $a := (F_1 + F_2 + F_3)/3 - F_4$
11:    $b := ((2F_1 - F_2 - F_3) + j\sqrt{3}(F_2 - F_1)/3)$
12:    $c := F_4$
13:    $\tilde{v} := -(b/2a)$
14:    **if** $a < 0$ and $|\tilde{v}| < 1$ **then**
15:       $v_{\mathrm{opt}} := \tilde{v}$
16:    **else**
17:       $v_{\mathrm{opt}} := (\tilde{b}/|\tilde{b}|)$
18:    **end if**
19:    $v_i := v_{\mathrm{opt}}$
20:    $i := (i) \bmod(N) + 1$
21: **until** Termination

In amplitude-and-phase SCA (APSCA) (Algorithm 1), a single SLM block is chosen, and its complex reflectance is adjusted to maximize the eye opening, which is the objective function. Next, another SLM block is chosen, and the process is repeated. (Note that amplitude-and-phase adaptation can be achieved using a phase-only SLM if SLM pixels are grouped into blocks, and if high-spatial-frequency (pixel-scale) phase gratings are introduced to diffract light away from the fiber core.) Since at every step, the objective function can only increase, the sequence of objective function values is monotonically non-decreasing. It is also bounded above by the global maximum of the SLMO problem. Therefore, this sequence converges.

Theorem 3.1: The SLM reflectance, $\mathbf{v}$, in APSCA converges to the global optimum.

     Proof: Refer to Appendix. ∎

Maximizing the objective function with respect to a single SLM block is also an optimization problem. Since F is a quadratic function of $\mathbf{v}$, it is also a quadratic function of each v_i. It can, therefore, be written as

$$F = a|v_i|^2 + \Re(b^* v_i) + c \tag{8}$$

where $a, c \in \mathbb{R}$, and $b \in \mathbb{C}$. Note that $a, b,$ and c are functions of v_k, $k = 1, \ldots, N$, $k \neq i$. In order to find the optimum v_i, $a, b,$ and c are first estimated. This is done using measurements of F at four different values of v_i, since $a, b,$ and c involve four real parameters. Once these are estimated, the optimal v_i, subject to the constraint that $|v_i| \leq 1$, may be computed analytically. Note that four measurements of the eye opening are required to optimize a single SLM block.

B. Continuous-Phase SCA

Algorithm 2 Continuous-Phase SCA

1: $v_i := 1, \quad i = 1, \ldots, N.$
2: $i := 1$
3: **repeat**
4:     **for** $m = 1$ to 3 **do**
5:       $v_i := \exp(j 2\pi (m - 1/3))$
6:       Estimate F_m
7:     **end for**
8:     $b := ((2F_1 - F_2 - F_3) + j\sqrt{3}(F_2 - F_1)/3)$
9:     $v_{\text{opt}} := \exp(j \arg(b))$
10:    $v_i := v_{\text{opt}}$
11:    $i := (i) \bmod (N) + 1$
12: **until** Termination

Continuous-phase SCA (CPSCA) (Algorithm 2) is similar to APSCA, except that only the phase of each SLM block is optimized, while the amplitude is held constant at 1. A single SLM block is chosen, and the amplitude of its reflectance is set to unity, while its phase is optimized to maximize the eye opening. Next, another SLM block is chosen, and the process is repeated. This algorithm is suboptimal, because every SLM block has unit amplitude, while the global optimum may involve SLM blocks with amplitudes less than 1. Optimization of the SLM phases to maximize eye opening is not a convex problem. Therefore, there may be many local optima that are not global optima. The final solution that the algorithm converges to will, in general, depend on the initial settings of the SLM. However, simulations show that the CPSCA converges to within 1% of the value of the objective function achieved by APSCA.

Maximizing the objective function with respect to the phase of a single SLM block is, again, an optimization problem. Since $|v_i| = 1$, from (8), we see that

$$F = \Re\{b^* \exp(j\phi_i)\} + d \tag{9}$$

where $\phi_i = \arg(v_i), d \in \mathbb{R}$, and $d = a + c$. The optimal ϕ_i is given by $\arg(b)$, where b is estimated using measurements of F at three different values of ϕ_i. Note that three measurements of the eye opening are required to optimize a single SLM block.

C. 4-Phase and 2-Phase SCA

Algorithm 3 4-Phase SCA

1: $v_i := 1, \quad i = 1, \ldots, N.$
2: $i := 1$
3: **repeat**
4:     **for** $m = 1$ to 4 **do**
5:       $v_i := \exp(j 2\pi (m - 1/4))$
6:       Estimate F_m
7:     **end for**
8:     $m_{\text{opt}} := \operatorname{argmax}(\mathrm{F}_1, \mathrm{F}_2, \mathrm{F}_3, \mathrm{F}_4)$
9:     $v_{\text{opt}} := \exp(j 2\pi (m_{\text{opt}} - 1/4))$
10:    $v_i := v_{\text{opt}}$
11:    $i := (i) \bmod (N) + 1$
12: **until** Termination

In 4-phase SCA (4PSCA) (Algorithm 3), a single SLM block is chosen, and the amplitude of its reflectance is set to unity, while its phase is optimized over the set $\{0, \pi/2, \pi, 3\pi/2\}$ to maximize the eye opening. Next, another SLM block is chosen, and the process is repeated. This algorithm is suboptimal with no guarantee of convergence to the global maximum, and in general performs poorer than APSCA and CPSCA. Like CPSCA, the final solution that the algorithm converges to will, in general, depend on the initial settings of the SLM.

We also mention 2-phase SCA (2PSCA), which is the algorithm used in [7]. 2PSCA is similar to 4PSCA, except that the phase of an SLM block is optimized over $\{0, \pi\}$. The performance of other algorithms will be compared against this.

IV. ADAPTIVE ALGORITHMS: NOISY SYSTEMS

In Section III, the estimates of the eye opening were assumed to be without noise. In practical systems, this is not the case. In this section, we look at modifications to the algorithms presented in the previous section, enabling trade-off between speed of convergence and resilience to noise. We first look at a modified version of APSCA to make it more robust to noise. We then look at a noise-resilient version of CPSCA.

A. Noise-Resilient Amplitude-and-Phase SCA

Algorithm 4 noise-resilient Amplitude-and-Phase SCA

1: $v_i := 1, \quad i = 1, \ldots, N$
2: $\mathbf{g} := [\underbrace{1 \ e^{(j2\pi/P)} \ldots e^{(j2\pi(P-1)/P)}}_{P} \underbrace{0 \ldots 0}_{Q}]^T$
    $\{\mathbf{g} = [g_1 \ldots g_{P+Q}]^T, g_i \in \mathbb{C}\}$
3: Compute $\mathbf{G} \in \mathbb{C}^{(P+Q) \times 4}$, where the kth row of
    $\mathbf{G} = [|g_k|^2 \Re(g_k) \Im(g_k) 1]$
4: $i := 1$
5: **repeat**
6:     **for** $m = 1$ to $P + Q$
7:       $v_i := g_m$
8:       Estimate F_m
9:     **end for**
10:    $\mathbf{f} := [F_1 \ldots F_{P+Q}]^T$
11:    $[a \ b_{\text{Re}} \ b_{\text{Im}} \ c]^T := \mathbf{G}^\dagger \mathbf{f}$
12:    $b := b_{\text{Re}} + j b_{\text{Im}}$
13:    $\tilde{v} := -(b/2a)$

14: **if** $a < 0$ and $|\tilde{v}| < 1$ **then**
15:     $v_{\text{opt}} := \tilde{v}$
16: **else**
17:     $v_{\text{opt}} := (\tilde{b}/|\tilde{b}|)$
18: **end if**
19:   $v_i := v_{\text{opt}}$
20:   $i := (i) \bmod (N) + 1$
21: **until** Termination

In noise-resilient APSCA (NRAPSCA), for each iteration of the algorithm, objective function values are observed for $(P + Q)$ values of the SLM block reflectance v_i. When $P = 3, Q = 1$, this algorithm is identical to the previous version of APSCA. As P and Q increase, the noise resilience of the algorithm increases. The $(P+Q)$ values of v_i at which the objective function is observed are given by **g**. From these, minimum mean-square error (MMSE) estimates of a, b, and c (8) are obtained. In presence of Gaussian noise, these are the maximum likelihood (ML) estimates of a, b, and c. These are in turn used to compute the optimal v_i. Note that $\mathbf{G}^{\dagger}$ is the pseudo-inverse of $\mathbf{G}$.

Let the noise in each observation of F be i.i.d., with zero mean and variance σ^2. We observe that v_{opt} depends only on a and b.

Theorem 4.1: At each step of the algorithm, the errors in the estimates of a, b_{Re}, and b_{Im} have zero mean, and variances $\sigma^2(1/P + 1/Q), 2\sigma^2/P$, and $2\sigma^2/P$ respectively.

   Proof: Refer to Appendix. ∎

Theorem 4.1 means that the errors in a and b due to noise can be made arbitrarily low by increasing P and Q. However, this increased noise resilience comes at the expense of slowing down the algorithm.

B. Noise-Resilient Continuous-Phase SCA

Algorithm 5 noise-resilient Continuous-Phase SCA

1: $v_i := 1, \quad i = 1, \ldots, N$.
2: $\mathbf{g} := [1 \; e^{(j2\pi/P)} \ldots e^{(j2\pi(P-1)/P)}]^T$
   $\{\mathbf{g} = [g_1 \ldots g_P]^T, g_i \in \mathbb{C}\}$
3: Compute $\mathbf{G} \in \mathbb{C}^{P \times 3}$, where the kth row of
   $\mathbf{G} = [\Re(g_k) \; \Im(g_k) \, 1]$
4: $i := 1$
5: **repeat**
6:   **for** $m = 1$ to P **do**
7:     $v_i := g_m$
8:     Estimate F_m
9:   **end for**
10:   $\mathbf{f} := [F_1 \ldots F_P]^T$
11:   $[b_{\text{Re}} \; b_{\text{Im}} \, d]^T := \mathbf{G}^{\dagger}\mathbf{f}$
12:   $b := b_{\text{Re}} + jb_{\text{Im}}$
13:   $v_{\text{opt}} := \exp(j \arg(b))$
14:   $v_i := v_{\text{opt}}$
15:   $i := (i) \bmod (N) + 1$
16: **until** Termination

In noise-resilient CPSCA (NRCPSCA), for each iteration of the algorithm, objective function values are observed for P values of the SLM block reflectance v_i. When $P = 3$, this

algorithm is identical to the previous version of CPSCA. As P increases, the noise resilience of the algorithm increases. The P values of v_i at which the objective function is observed are given by **g**. From these, MMSE estimates of b and d (9) are obtained. In presence of Gaussian noise, these are the ML estimates of b and d. These are in turn used to compute the optimal v_i.

Let the noise in each observation of F be i.i.d., with zero mean and variance σ^2. We observe that v_{opt} depends only on b.

Theorem 4.2: At each step of the algorithm, the errors in the estimates of b_{Re} and b_{Im} have zero mean, and variance $2\sigma^2/P$.

   Proof: Refer to Appendix. ∎

Theorem 4.2 means that the errors in b due to noise can be made arbitrarily low by increasing P. Again, this increased noise resilience comes at the expense of slowing down the algorithm.

V. SIMULATION RESULTS

A. Fiber and System Parameters

For simulation, we use parameters from the experimental setup used in [7]. We use a 50-μm-core graded-index multimode fiber, 1 km in length. We use the infinite-core approximation, and therefore approximate the IMs as Hermite-Gaussian modes. Mode shapes and propagation constants (β_i) are computed analytically. The group delays of the IMs are scaled up by a factor of 10, so as to make them comparable to modal delays observed in experiment. This is because the differential group delay in a fiber with perfectly quadratic refractive index profile is less than that observed in most real fibers by about an order of magnitude. We operate at a wavelength of 1550 nm. We use a bit rate of 10 Gb/s. At a wavelength of 1550 nm, the fiber supports 55 modes in each polarization.

Light from an SMF with 10.4-μm mode field diameter is imaged onto the SLM through a 10.4-mm focal-length lens. The guided mode of the SMF is approximated as Gaussian. Light reflected by the SLM is imaged onto the MMF through a 10.4-mm focal-length lens. The input faces of both fibers are in the focal planes of their respective lenses.

Although the SLM in [7] provided phase control only, here, the SLM is assumed to control both amplitude and phase with a 128×128 array of pixels, covering a region containing 95% of the incident power. Each pixel is $18 \times 18 \; \mu\text{m}^2$. These pixels are grouped into larger square blocks during operation. A typical block size is 16×16 pixels, to have a 2-D array of 8×8 blocks on the SLM, though block size is part of the parameter space explored in this simulation. The total area of the SLM is always kept constant. Grouping a large number of pixels into blocks means that even if, in practice, the SLM has only phase control, both amplitude and phase control may be achieved at the block level by introducing high spatial frequency (pixel-level) phase gratings to diffract light away from the fiber core. Simulations show that a block size of 4×4 pixels is sufficient to mimic amplitude-and-phase control with a phase-only SLM for the imaging system, wavelength, and fiber parameters used in this simulation. We assume a switching time of 10 μs, which is typical for a micro-electro-mechanical system SLM [12].
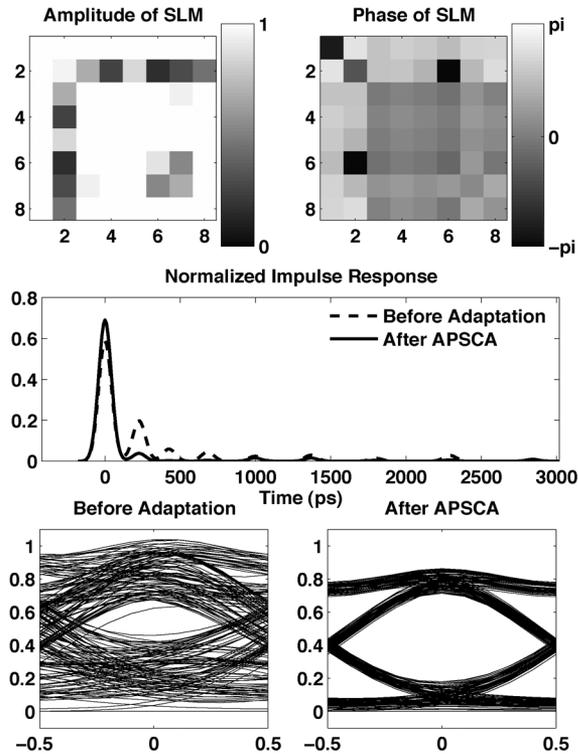
Fig. 3. Amplitude and phase of SLM reflectance, impulse responses before and after convergence, and eye diagrams, with APSCA.
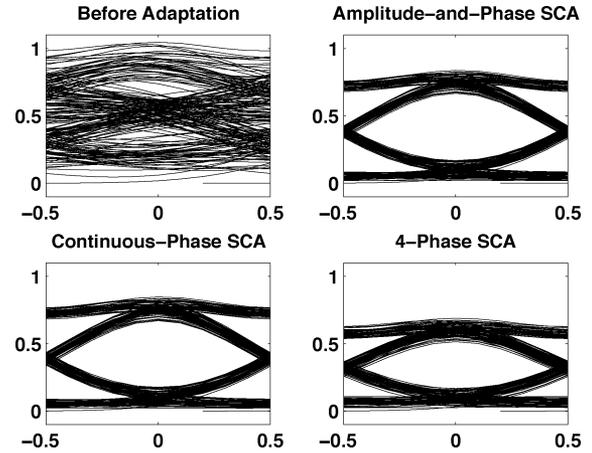

Fig. 4. Eye diagrams before adaptation, and after APSCA, CPSCA, and 4PSCA.

TABLE I
PERFORMANCE COMPARISON OF APSCA, CPSCA, 4PSCA, AND 2PSCA, FOR MMF WITH RANDOM MODE COUPLING AND CENTER LAUNCH, AND FOR MMF WITH NO MODE COUPLING AND OFFSET LAUNCH, 8×8 BLOCK SLM

	Objective function				
	Before	APSCA	CPSCA	4PSCA	2PSCA
Random mode coupling					
Low coupling	0.2176	0.5373	0.5317	0.4826	0.3813
High coupling	-0.1594	0.4210	0.4155	0.3385	0.1493
Offset launch					
2 μm	0.4710	0.6429	0.6389	0.5449	0.5093
5 μm	-0.0950	0.6289	0.6249	0.4350	0.1231
10 μm	-0.8795	0.5836	0.5791	0.4875	-0.0834

Mode coupling is simulated by coupling the IMs with randomly generated complex unitary matrices. PMs are thus unitary combinations of IMs. Unitarity ensures conservation of energy. The high-mode-coupling regime is simulated using unitary matrices that have large off-diagonal terms. The fiber loss coefficient α may be assumed to be zero without loss of generality. Correspondence between simulated impulse responses and those observed in experiment suggests that this is a reasonable model.

Our figure of merit is the objective function $\mathbf{v}^H \mathbf{P} \mathbf{v}$. $\mathbf{P}$ is normalized so that, when the SLM blocks are all set to amplitude 1 and phase 0, ($v_i = 1, i = 1, \ldots, N$) the total power in all modes excited in the MMF is unity. Note that a negative value of the objective function indicates a closed eye.

The optimal solution is computed by solving using CVX, a freely distributed convex optimization library for MATLAB [13].

B. Simulation Results: Noiseless Systems

In this section, we compare performance of APSCA, CPSCA, 4PSCA and 2PSCA. Fig. 3 shows the impulse response of the system before and after adaptation of the SLM with APSCA. The SLM is a 2-D array of 8×8 blocks, each of 16×16 pixels. The fiber has random mode coupling, simulated as described in the previous section. We see that, after adaptation, higher-order PMs have smaller amplitudes. This leads to reduction in ISI and a larger eye opening, as is evident from the eye diagrams before and after adaptation. Fig. 3 also shows the SLM settings achieved after convergence. In all simulations, the SLM starts from a blank state, with all blocks set to amplitude 1 and phase 0, ($v_i = 1, i = 1, \ldots, N$).

Fig. 4 compares the eye openings after APSCA, CPSCA, and 4PSCA. We see that 4PSCA gives a smaller eye than APSCA or CPSCA. Table I gives a compares the value of the objective function achieved by these algorithms under various impairments.

Fig. 5 shows typical adaptation curves for the four algorithms. Here, the fiber has random mode coupling, and the SLM uses 8×8 blocks, each of 16×16 pixels. Since one step in APSCA and 4PSCA require four SLM block flips, while CPSCA requires only three, they take different numbers of block flips for a single pass over the SLM. We see that, because of this, CPSCA converges faster than APSCA, while achieving a marginally suboptimal value of the objective function. This makes CPSCA a very attractive algorithm, as it requires only a phase-only SLM.

Table II gives a comparison of performance across various blocks sizes, in a system with random mode coupling. A smaller sized block translates to higher resolution in spatial frequency. However, this also means a larger number of blocks, and therefore slower convergence. We see that an SLM with 8×8 blocks, each of 16×16 pixels, gives good performance, while keeping convergence time small. We also see that, as SLM resolution increases, the difference between the final values of the objective function achieved by CPSCA and APSCA decreases.

Figs. 6 and 7 show adaptation curves for various SLM configurations. Note that at higher spatial resolution of the SLM, differences in performance between the algorithms decreases.
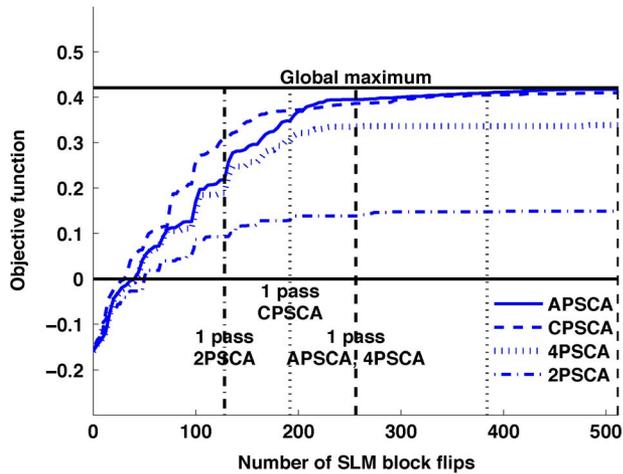
Fig. 5. Convergence curves: APSCA, CPSCA, 4PSCA, and 2PSCA; 8×8 SLM blocks (each of 16×16 pixels); noiseless system; vertical lines indicate completion of a pass over all SLM blocks.

TABLE II
OBJECTIVE FUNCTION FOR VARIOUS SLM RESOLUTIONS: TOTAL ACTIVE AREA OF THE SLM IS KEPT CONSTANT AT 128×128 PIXELS

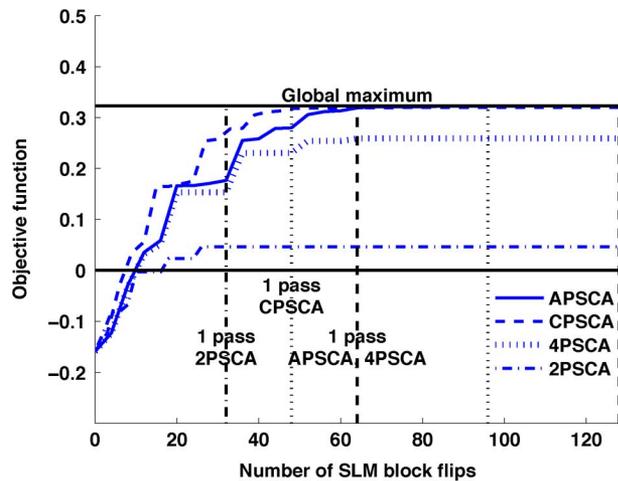| Number of SLM blocks | Objective function | | | | |
|---|---|---|---|---|
| | Before | APSCA | CPSCA | 4PSCA | 2PSCA |
| 4×4 | -0.1593 | 0.3229 | 0.3200 | 0.2587 | 0.04605 |
| 8×8 | -0.1593 | 0.4209 | 0.4154 | 0.3384 | 0.1492 |
| 16×16 | -0.1593 | 0.4629 | 0.4628 | 0.4142 | 0.2335 |
| 32×32 | -0.1593 | 0.4802 | 0.4790 | 0.4317 | 0.2482 |

Fig. 6. Convergence curves: APSCA, CPSCA, 4PSCA, and 2PSCA; noiseless system; 4×4 SLM blocks (each of 32×32 pixels); vertical lines indicate completion of a pass over all SLM blocks.

C. Simulation Results: Noisy Systems

In this section, we compare performance of algorithms NRAPSCA, NRCPSCA, 4PSCA, and 2PSCA. It is assumed that adaptation is done using the periodic square-wave training sequence in Section II.E, with $L = 64$, leading to a training sequence 128 bits long.

Fig. 8 shows the convergence of NRAPSCA, NRCPSCA, 4PSCA, and 2PSCA, for 500 random realizations of noise sequences. The center line shows the median value of the objective function, with the error bars indicating the 5th and 95th percentiles. The SLM employs 8×8 blocks, each of 16×16 pixels.
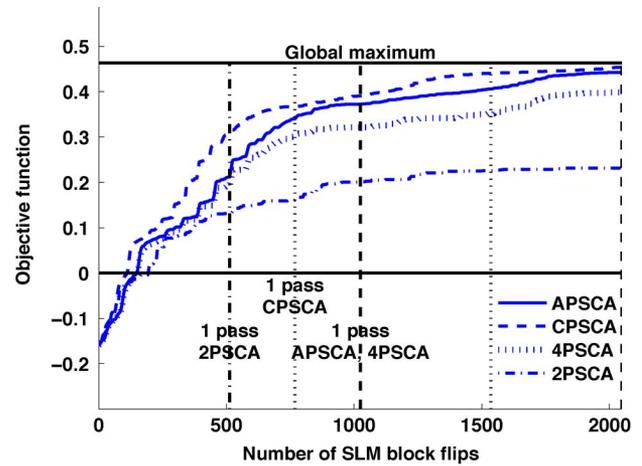

Fig. 7. Convergence curves: APSCA, CPSCA, 4PSCA, and 2PSCA; noiseless system; 16×16 SLM blocks (each of 8×8 pixels); vertical lines indicate completion of a pass over all SLM blocks.
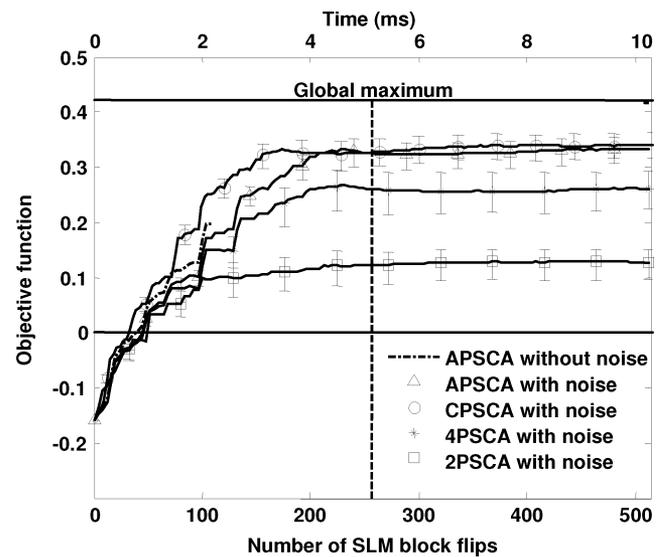

Fig. 8. Convergence in presence of noise: APSCA, CPSCA, 4PSCA, and 2PSCA; 8×8 SLM blocks (each of 16×16 pixels); 10 μs SLM switching time; 10 μs noise averaging; vertical line indicates one pass over all SLM blocks by APSCA.

Noise is assumed to be i.i.d. Gaussian, with noise power chosen so that NRCPSCA achieves a BER less than 10^{-3} in more than 95% of the simulated cases. A BER of 10^{-3} is a typical forward error correction (FEC) threshold. A lower BER requirement (e.g., 10^{-6} or 10^{-9}) will only set the noise power lower, thus making conditions more favorable for the algorithms. This is because the maximum eye opening in the absence of noise is fixed at the global maximum indicated. Therefore, setting a lower BER threshold necessitates a system with higher SNR, and thus lower noise power. Since the SLM switching time is 10 μs, and bit period is 0.1 ns, we use noise averaging over 10 μs, with multiple copies of the training sequence being sent over this period. Here, $P = 3, Q = 1$ for NRAPSCA, and $P = 3$ for NRCPSCA. From Section IV, we know that the effect of increasing P and Q is equivalent to increased noise averaging.
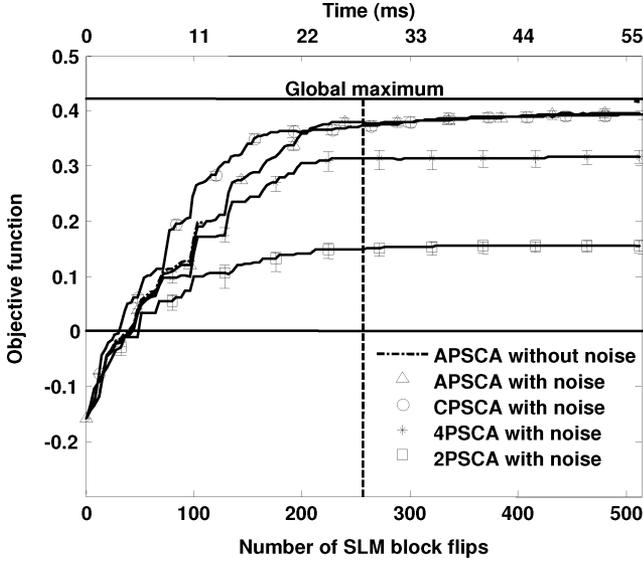
Fig. 9. Convergence in presence of noise: APSCA, CPSCA, 4PSCA, and 2PSCA; 8×8 SLM blocks (each of 16×16 pixels); 10 μs SLM switching time; 100 μs noise averaging; vertical line indicates one pass over all SLM blocks by APSCA.

Fig. 9 shows the convergence of the various algorithms with noise averaging over 100 μs. We see that with more noise averaging, the performance is closer to what is obtained in the absence of noise.

D. Multiple Modes in a Single Bit Period

Note that in our analysis and simulations, the modal delay differences between different mode groups are assumed to be large enough so that only one spatial mode (both polarizations) falls into the first bit period. Modes falling into other bit periods, and thus causing ISI, need not be discernible in time. However, in a short fiber, or at low symbol rates, multiple mode groups may fall into the first bit period. This means that, instead of exciting just one mode group and suppressing all others, we wish to excite all mode groups that fall within the first bit period, and to suppress the rest. In this case, the underlying optimization problem cannot be posed as a convex problem as presented here. However, simulations show that our algorithms work well even in the scenario. In fact, the performance of our algorithms in these cases is lower bounded by the case where only one mode group lies in the first bit period. This means that performance in these cases can only be better than what is seen in the simulations here.

VI. CONCLUSION

In this paper, we have proposed adaptive algorithms to ISI in MMF systems using an SLM. One of the algorithms, APSCA, is shown to converge to the global optimum in the absence of noise. We also propose suboptimal algorithms with faster convergence, and lesser hardware complexity. We also develop modified versions of these algorithms to enable resilience to noise. We present simulation results, using parameters of commercially available fibers and SLMs, and show that these techniques mitigate ISI and open up an otherwise closed eye

pattern. Simulations show that CPSCA appears to be the most attractive algorithm in terms of speed of convergence, performance in presence of noise, and the fact that a phase-only SLM is sufficient to implement this.

Future work will involve experimental verification of the performance of these algorithms.

APPENDIX

Proof of Theorem 3.1: Our proof proceeds as follows: we first obtain necessary conditions satisfied by $\mathbf{v}$ upon convergence of APSCA; we then show that these imply optimality (KKT) conditions of convex form SLMO; finally, we conclude that the $\mathbf{v}$ that APSCA converges to must be the optimal solution to convex form SLMO, and therefore to SLMO.

We start by observing that

$$
\begin{aligned}
F = \mathbf{v}^H \mathbf{P} \mathbf{v} &= [v_i^* \ \bar{\mathbf{v}}_i^H] \begin{bmatrix} P_{ii} & \mathbf{q}_i^H \\ \mathbf{q}_i & \mathbf{R}_i \end{bmatrix} \begin{bmatrix} v_i \\ \bar{\mathbf{v}}_i \end{bmatrix} \\
&= P_{ii}|v_i|^2 + 2\Re(\bar{\mathbf{v}}_i^H \mathbf{q}_i v_i) + \bar{\mathbf{v}}_i^H \mathbf{R}_i \bar{\mathbf{v}}_i \quad (10)
\end{aligned}
$$

where $\bar{\mathbf{v}}_i = [v_1 \ldots v_{i-1} v_{i+1} \ldots v_N]^T$, and $P_{ii} \in \mathbb{R}, \mathbf{q}_i \in \mathbb{C}^{N-1}$, and $\mathbf{R}_i \in \mathbb{C}^{(N-1) \times (N-1)}$ are appropriate sub-matrices of $\mathbf{P}$ (for e.g., P_{ii} is the ith diagonal element of $\mathbf{P}$).

Comparing with (8), we see that $a = P_{ii}, b = 2\mathbf{q}_i^H \bar{\mathbf{v}}_i$, and $c = \bar{\mathbf{v}}_i^H \mathbf{R}_i \bar{\mathbf{v}}_i$. Therefore, in Algorithm 1, steps III.A–III.A are equivalent to:

if $(P_{ii} < 0$ and $|\mathbf{q}_i^H \bar{\mathbf{v}}_i / P_{ii}| < 1)$ **then**

$v_i := -\mathbf{q}_i^H \bar{\mathbf{v}}_i / P_{ii}$

else

$v_i := \mathbf{q}_i^H \bar{\mathbf{v}}_i / |\mathbf{q}_i^H \bar{\mathbf{v}}_i|$

end if

Therefore, $\mathbf{v}$ when APSCA converges, defined as $\mathbf{v}^\star$, will be such that

$$
v_i^\star = \begin{cases} -\mathbf{q}_i^H \bar{\mathbf{v}}_i^\star / P_{ii} & P_{ii} < 0 \quad \text{and} \ |\mathbf{q}_i^H \bar{\mathbf{v}}_i^\star / P_{ii}| < 1 \\ \mathbf{q}_i^H \bar{\mathbf{v}}_i^\star / |\mathbf{q}_i^H \bar{\mathbf{v}}_i^\star| & \text{otherwise} \end{cases}
$$

$$(11)$$

for $i = 1, \ldots, N$. We now show that $\mathbf{v}^\star$ satisfies KKT conditions of convex form SLMO, and must therefore be globally optimal.

In SLMO, $\mathbf{P} = \mathbf{p}_1 \mathbf{p}_1^H - \mathbf{P}_2 \mathbf{P}_2^H$. Here, $\mathbf{p}_1$ is unique only to within an overall phase. Therefore, we choose the phase of $\mathbf{p}_1$ such that $\mathbf{p}_1^H \mathbf{v}^\star \in \Re$ and $\mathbf{p}_1^H \mathbf{v}^\star \geq 0$.

The Lagrangian to convex form SLMO, $L(\mathbf{v}, t, \mathbf{x}, y, \lambda, \nu)$, may be written as

$$
\begin{aligned}
L = -y &+ \lambda_0(\sqrt{\mathbf{x}^H \mathbf{x} + y^2} - t) + \sum_{i=1}^N \lambda_i(|v_i|^2 - 1) \\
&+ \Re[\nu_1^*(\mathbf{p}_1^H \mathbf{v} - t)] + \Re[\nu_2^H (\mathbf{P}_2^H \mathbf{v} - \mathbf{x})]. \quad (12)
\end{aligned}
$$

We define

$$t^\star = \mathbf{p}_1^H \mathbf{v}^\star \quad (13)$$

$$\mathbf{x}^\star = \mathbf{P}_2^H \mathbf{v}^\star \quad (14)$$

$$y^\star = \sqrt{\mathbf{v}^{\star H} \mathbf{P} \mathbf{v}^\star} \quad (15)$$

$$\lambda_0^\star = \frac{\mathbf{p}_1^H \mathbf{v}^\star}{\sqrt{\mathbf{v}^{\star H} \mathbf{P} \mathbf{v}^\star}} \tag{16}$$

$$\lambda_i^\star = \frac{[\mathbf{P}]_i \mathbf{v}^\star}{2 v_i \sqrt{\mathbf{v}^{\star H} \mathbf{P} \mathbf{v}^\star}} \quad i = 1, \ldots, N \tag{17}$$

$$\nu_1^\star = -\frac{\mathbf{p}_1^H \mathbf{v}^\star}{\sqrt{\mathbf{v}^{\star H} \mathbf{P} \mathbf{v}^\star}} \tag{18}$$

$$\nu_2^\star = \frac{\mathbf{P}_2^H \mathbf{v}^\star}{\sqrt{\mathbf{v}^{\star H} \mathbf{P} \mathbf{v}^\star}} \tag{19}$$

where $[\mathbf{P}]_i$ is the ith row of $\mathbf{P}$. Therefore,

$$[\mathbf{P}]_i \mathbf{v}^\star = P_{ii} v_i^\star + \mathbf{q}_i^H \bar{\mathbf{v}}_i^\star. \tag{20}$$

We proceed to show that $(\mathbf{v}^\star, t^\star, \mathbf{x}^\star, y^\star, \lambda^\star, \nu^\star)$ satisfy KKT conditions to convex form SLMO, and are therefore the primal and dual optimal variables.

The KKT conditions to convex form SLMO are

$$\sqrt{\mathbf{x}^{\star H} \mathbf{x}^\star + y^{\star 2}} \leq t^\star \tag{21}$$

$$|v_i^\star|^2 \leq 1 \quad i = 1, \ldots, N \tag{22}$$

$$\mathbf{p}_1^H \mathbf{v}^\star = t^\star \tag{23}$$

$$\mathbf{P}_2^H \mathbf{v}^\star = \mathbf{x}^\star \tag{24}$$

$$\lambda_i^\star \geq 0 \quad i = 0, \ldots, N \tag{25}$$

$$\lambda_0^\star \left(\sqrt{\mathbf{x}^{\star H} \mathbf{x}^\star + y^{\star 2}} - t^\star \right) = 0 \tag{26}$$

$$\lambda_i^\star \left(|v_i^\star|^2 - 1 \right) = 0 \quad i = 1, \ldots, N \tag{27}$$

$$\left. \frac{\partial L}{\partial \{y, \mathbf{x}, t, \mathbf{v}\}} \right|_{y^\star, \mathbf{x}^\star, t^\star, \mathbf{v}^\star, \lambda^\star, \boldsymbol{\nu}^\star} = 0. \tag{28}$$

Before proving that (11), (13)–(19) satisfy (21)–(28), we note that $\mathbf{p}_1^H \mathbf{v}^\star \geq 0$, and $\sqrt{\mathbf{x}^{\star H} \mathbf{x}^\star + y^{\star 2}} = \mathbf{p}_1^H \mathbf{v}^\star$. Now, from direct substitution, (21)–(24) are seen to be satisfied, with (21) being satisfied with equality. For $i = 0$, by definition, (25) is satisfied. For $i = 1, \ldots, N$, from (11) and (20), we see that (25) is equivalent to (29), shown at the bottom of the page. Equation (26) follows because (21) is satisfied with equality. (27) is satisfied because, when $P_{ii} < 0$ and $|\mathbf{q}_i^H \bar{\mathbf{v}}_i^\star / P_{ii}| < 1$, $\lambda_i = 0$,

and otherwise, from (11), $|v_i^\star|^2 = 1$. To prove (28), we observe that

$$\frac{\partial L}{\partial y} = -1 + \frac{\lambda_0^\star y^\star}{\sqrt{\mathbf{x}^{\star H} \mathbf{x}^\star + y^{\star 2}}} = 0 \tag{30}$$

$$\frac{\partial L}{\partial t} = -\lambda_0^\star - \nu_1^\star = 0 \tag{31}$$

$$\frac{\partial L}{\partial \mathbf{x}} = \frac{\lambda_0^\star \mathbf{x}^\star}{\sqrt{\mathbf{x}^{\star H} \mathbf{x}^\star + y^{\star 2}}} - \nu_2^\star = 0 \tag{32}$$

$$\frac{\partial L}{\partial \mathbf{v}} = 2 [\lambda_i^\star v_i^\star] + \mathbf{p}_1 \nu_1^\star + \mathbf{P}_2 \nu_2^\star$$
$$= \left(\mathbf{P} \mathbf{v}^\star - \mathbf{p}_1 \mathbf{p}_1^H \mathbf{v}^\star + \mathbf{P}_2 \mathbf{P}_2^H \mathbf{v}^\star \right) / \sqrt{\mathbf{v}^{\star H} \mathbf{P} \mathbf{v}^\star}$$
$$= 0 \tag{33}$$

where $[\lambda_i^\star v_i^\star] \in \mathbb{C}^N$, with the ith element being $\lambda_i^\star v_i^\star$.

Thus, $\mathbf{v}^\star$, the $\mathbf{v}$ that APSCA converges to, satisfies the KKT conditions to convex form SLMO, and must therefore be the globally optimal solution.

Proof of Theorems 4.1 and 4.2: In Algorithm 4

$$F_m = a|g_m|^2 + b_{\mathrm{Re}} \Re(g_m) + b_{\mathrm{Im}} \Im(g_m) + c + n_m \tag{34}$$

where n_m is the noise corrupting F_m. Therefore,

$$\mathbf{f} = \mathbf{G}[a \, b_{\mathrm{Re}} \, b_{\mathrm{Im}} \, c]^T + \mathbf{n} \tag{35}$$

where $\mathbf{n} = [n_1 \ldots n_{P+Q}]^T$, $n_m, m = 1, \ldots, P + Q$, are i.i.d. zero mean, and $\mathbf{G}$ is defined in Algorithm 4. Therefore, the MMSE estimate of $[a \, b_{\mathrm{Re}} \, b_{\mathrm{Im}} \, c]^T$ is given by $\mathbf{G}^\dagger \mathbf{f}$. This will be the ML estimate when noise is i.i.d. $N(0, \sigma^2)$. Also, the mean error is given by $\mathbf{E}(\mathbf{G}^\dagger \mathbf{n}) = \mathbf{0}$, and the covariance matrix is given by $\mathrm{Cov} \left([a \, b_{\mathrm{Re}} \, b_{\mathrm{Im}} \, c]^T \right) = \sigma^2 \mathbf{G}^\dagger \mathbf{G}^{\dagger T}$. Therefore, the variances of $a, b_{\mathrm{Re}}, b_{\mathrm{Im}}$, and c are given by the diagonal elements of $\sigma^2 \mathbf{G}^\dagger \mathbf{G}^{\dagger T}$. Also, we see that (36), shown at the bottom of the page. Therefore, the diagonal elements of $\sigma^2 \mathbf{G}^\dagger \mathbf{G}^{\dagger T}$ are given by $\sigma^2 (1/P + 1/Q), 2\sigma^2/P, 2\sigma^2/P$, and σ^2/Q. Theorem 4.1 follows.

An analogous proof holds for Theorem 4.2. Here, the MMSE estimate of $[b_{\mathrm{Re}} \, b_{\mathrm{Im}} \, d]^T$ is given by $\mathbf{G}^\dagger \mathbf{f}$, with $\mathbf{G}$ as defined in Algorithm 5. This will be the ML estimate when n_m are i.i.d.

$$
\lambda_i^\star =
\begin{cases}
0 & P_{ii} < 0, |\mathbf{q}_i^H \bar{\mathbf{v}}_i^\star / P_{ii}| < 1 \\
\frac{P_{ii} \mathbf{q}_i^H \bar{\mathbf{v}}_i^\star + \mathbf{q}_i^H \bar{\mathbf{v}}_i^\star |\mathbf{q}_i^H \bar{\mathbf{v}}_i^\star|}{2 \mathbf{q}_i^H \bar{\mathbf{v}}_i^\star \sqrt{\mathbf{v}^{\star H} \mathbf{P} \mathbf{v}^\star}} & \text{otherwise}
\end{cases}
$$
$$
=
\begin{cases}
0 & P_{ii} < 0, |\mathbf{q}_i^H \bar{\mathbf{v}}_i^\star / P_{ii}| < 1 \\
\frac{|\mathbf{q}_i^H \bar{\mathbf{v}}_i^\star|}{2\sqrt{\mathbf{v}^{\star H} \mathbf{P} \mathbf{v}^\star}} \left(\frac{P_{ii}}{|\mathbf{q}_i^H \bar{\mathbf{v}}_i^\star|} + 1 \right) & \text{otherwise}
\end{cases}
$$
$$
\geq 0. \tag{29}
$$

$$
\mathbf{G}^\dagger =
\begin{bmatrix}
\overbrace{1/P \ldots 1/P}^{P} & \overbrace{-1/Q \ldots -1/Q}^{Q} \\
\frac{2}{P} \cos(\frac{2\pi \cdot 0}{P}) \ldots \frac{2}{P} \cos(\frac{2\pi(P-1)}{P}) & 0 \ldots 0 \\
\frac{2}{P} \sin(\frac{2\pi \cdot 0}{P}) \ldots \frac{2}{P} \sin(\frac{2\pi(P-1)}{P}) & 0 \ldots 0 \\
0 \ldots 0 & 1/Q \ldots 1/Q
\end{bmatrix}
\tag{36}
$$

$N(0, \sigma^2)$. Again, the mean error is given by $\mathbf{E}(\mathbf{G}^\dagger \mathbf{n}) = \mathbf{0}$, and the covariance matrix is given by $\text{Cov}\left([b_{\text{Re}}\ b_{\text{Im}}\ d]^T\right) = \sigma^2 \mathbf{G}^\dagger \mathbf{G}^{\dagger T}$. Therefore, the variances of $b_{\text{Re}}, b_{\text{Im}},$ and d are given by the diagonal elements of $\sigma^2 \mathbf{G}^\dagger \mathbf{G}^{\dagger T}$. Here,

$$
\mathbf{G}^\dagger = \begin{bmatrix} \frac{2}{P}\cos\left(\frac{2\pi \cdot 0}{P}\right) & \cdots & \frac{2}{P}\cos\left(\frac{2\pi(P-1)}{P}\right) \\ \frac{2}{P}\sin\left(\frac{2\pi \cdot 0}{P}\right) & \cdots & \frac{2}{P}\sin\left(\frac{2\pi(P-1)}{P}\right) \\ 1/P & \cdots & 1/P \end{bmatrix}. \quad (37)
$$

Therefore, the diagonal elements of $\sigma^2 \mathbf{G}^\dagger \mathbf{G}^{\dagger T}$ are given by $2\sigma^2/P, 2\sigma^2/P,$ and σ^2/P. Theorem 4.2 follows.

ACKNOWLEDGMENT

The authors would like to thank S. Boyd, J. Wilde, and A. Lau for many useful discussions.

REFERENCES

[1] G. P. Agrawal, *Fiber-Optic Communications 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] 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.

[6] S. Fan and J. M. Kahn, "Principal modes in multi-mode waveguides," *Opt. Lett.*, vol. 30, no. 2, pp. 135–137, Jan. 15, 2005.

[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. 15, 2005.

[8] R. A. Panicker, J. P. Wilde, J. M. Kahn, D. F. Welch, and I. Lyubomirsky, "10 × 10 Gb/s DWDM transmission through 2.2 km multimode fiber using adaptive optics," *IEEE Photon. Technol. Lett.*, vol. 19, no. 15, pp. 1154–1156, Aug. 1, 2007.

[9] R. A. Panicker, J. M. Kahn, and S. P. Boyd, "Compensation of multimode fiber dispersion using adaptive optics via convex optimization," *J. Lightw. Technol.*, submitted to.

[10] J. M. Kahn, W. J. Krause, and J. B. Carruthers, "Experimental characterization of non-directed indoor infrared channels," *IEEE Trans. Commun.*, vol. 43, no. 8, pp. 1613–1323, Apr. 1995.

[11] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2003.

[12] S. Cornelissen, T. G. Bifano, and P. A. Bierden, M. E. Motamedi and R. Goering, Eds., "MEMS spatial light modulators with integrated electronics," in *Proc. SPIE*, vol. 4561, pp. 28–34, MOEMS and Miniaturized Systems II.

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

Rahul Alex Panicker received the B.Tech. degree from the Indian Institute of Technology (IIT), Madras, India, in 2002, and the M.S. and Ph.D. degrees from Stanford University, Stanford, CA, in 2004 and 2007, respectively, all in electrical engineering.

He spent 2008 working with the Optical Systems Group at Infinera Corporation. His research interests included applying convex-optimization techniques to optical communications systems and designing next-generation systems based on Infinera's photonic-integrated-circuit technology. He currently works as President of Rural Products at Embrace, Bangalore, India, a company trying to save premature babies in developing countries through a low-cost infant warmer that can work without electricity.

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

From 1987-1990, he was at AT&T Bell Laboratories, Crawford Hill Laboratory, Holmdel, NJ. He demonstrated multi-Gb/s coherent optical fiber transmission systems, setting world records for receiver sensitivity. From 1990-2003, he was on the faculty of the Department of Electrical Engineering and Computer Sciences at U.C. Berkeley, performing research on optical and wireless communications. Since 2003, he has been a Professor of Electrical Engineering at Stanford University. His current research interests include single- and multi-mode optical fiber communications, free-space optical communications, and MEMS for optical communications.

Prof. Kahn received the National Science Foundation Presidential Young Investigator Award in 1991. From 1993-2000, he served as a Technical Editor of IEEE *Personal Communications Magazine*. Since 2009, he has been an Associate Editor of IEEE/OSA JOURNAL OF OPTICAL COMMUNICATIONS AND NETWORKING. In 2000, he helped found StrataLight Communications (now Opnext Subsystems), where he served as Chief Scientist from 2000-2003.