Decoupled stochastic parallel gradient descent optimization for adaptive optics: integrated approach for wave-front sensor information fusion

Mikhail A. Vorontsov

U.S. Army Research Laboratory, Computational and Information Sciences and Technology Directorate, Intelligent Optics Laboratory, Adelphi, Maryland 20783

Received March 30, 2001; revised manuscript received June 14, 2001; accepted July 13, 2001

A new adaptive wave-front control technique and system architectures that offer fast adaptation convergence even for high-resolution adaptive optics is described. This technique is referred to as decoupled stochastic parallel gradient descent (D-SPGD). D-SPGD is based on stochastic parallel gradient descent optimization of performance metrics that depend on wave-front sensor data. The fast convergence rate is achieved through partial decoupling of the adaptive system's control channels by incorporating spatially distributed information from a wave-front sensor into the model-free optimization technique. D-SPGD wave-front phase control can be applied to a general class of adaptive optical systems. The efficiency of this approach is analyzed numerically by considering compensation of atmospheric-turbulence-induced phase distortions with use of both low-resolution (127 control channels) and high-resolution (256×256 control channels) adaptive systems. Results demonstrate that phase distortion compensation can be achieved during only 10–20 iterations. The efficiency of adaptive wave-front correction with D-SPGD is practically independent of system resolution. © 2002 Optical Society of America

OCIS codes: 010.0110; 010.1080; 110.0110.

1. INTRODUCTION

Instantaneous correction of phase distortions emerging from wave propagation through optically inhomogeneous media is a key adaptive optics objective.[1–3] This correction is performed by means of active optical elements (wave-front correctors): electrically addressed deformable mirrors and mirror arrays, multielement liquid crystal phase spatial light modulators, and micromechanical mirrors. These active optical elements are capable of changing the wave-front phase spatial distribution $u(\mathbf{r}, t)$ at each point $\mathbf{r} = \{x, y\}$ on the corrector aperture by applying control signals (controls) $\{u_l(t)\}$, $l = 1, \ldots, N$ to its electrodes. The number of controls (control channels) N determines the adaptive system spatial resolution and hence spatial-frequency bandwidth for phase distortion correction. Increasing spatial resolution without sacrificing adaptive system operational speed is an important avenue for adaptive optics development. In most cases high-resolution phase aberration correction is considered not as a replacement for but rather as an important addition to existing adaptive optics techniques and is used in combination as a high-resolution secondary feedback loop system.[4,5]

Far-reaching increases in spatial resolution (from hundreds to thousands of control channels) cannot be achieved by simply replacing old-fashioned deformable mirrors with large-scale micromechanical mirror arrays[6,7] or with high-resolution LC phase spatial light modulators.[8] Transition to high-resolution adaptive optics is a nontrivial problem that may require radical changes of the entire adaptive system architecture. These changes first of all should affect wave-front aberration sensing devices (wave-front sensors). The most commonly used adaptive optics wave-front sensors, such as the lateral shearing interferometer and Shack–Hartmann and curvature sensors, are based on computation of control signals from wave-front slopes or wave-front-phase second derivatives (slope-type sensors).[1] Those involving matrix algebra computations—known as wavefront phase/control reconstruction—are efficient only for relatively low-resolution systems ($N \leqslant 200$–300). Estimations show that the number of operations (multiplications and additions) required to calculate the controls from "slope-type" wave-front sensor measurements increase as N^2 (Ref. 9). This makes a hundredfold increase in N a challenging task even with parallel architecture computers specially designed to perform matrix operations.

A quite different type of problem emerges for adaptive optics with wave-front control based on gradient descent optimization, often referred to as model-free optimization adaptive optics.[10–13] In these systems all information about the distorted wave front is reduced (collapsed) into a single signal (performance metric J) typically measured by a single photodetector. By applying small perturbations of the controls $\{\delta u_l\}$ ($l = 1, \ldots, N$) one can compute approximations to the metric gradient and use this approximation in a gradient descent type iterative procedure for wave-front corrector control voltage update. The way this gradient approximation is performed is dependent on the technique used: multidithering,[13] sequential perturbations,[10,11] and stochastic parallel gradient descent.[12,14] Typically computation of the gradient approximation is quite simple when compared with corresponding computations of controls from wave-front local slopes. The major problem with model-free optimization adaptive optics is slow convergence. The convergence rate, defined as the number of iterations N_c required to approach the vicinity of the performance metric extremum, typically increases as N increases. In one of the most efficient gradient descent optimization techniques—stochastic parallel gradient descent (SPGD) (also called the simultaneous perturbation stochastic approximation[15,16])—the convergence rate N_c increases at least as $N^{1/2}$ (Ref. 17). A typical number of iterations N_c for an adaptive system with $N \approx 100$ control channels is on the order of $N_c \approx 10^2$ (Ref. 18). A hundredfold increase in resolution would cause the adaptive system to be unacceptably slow for most applications related to the compensation of dynamically changing phase distortions.

The question raised in this paper is how to merge the advantages of wave-front sensor based adaptive optics with model-free optimization adaptive optics and eliminate their inherent problems that limit the transition to high-resolution wave-front control. The limitations imposed by the conventional adaptive wave-front control techniques mentioned above can be overcome by using a new wave-front control technique (decoupled SPGD, or D-SPGD) described here. This technique incorporates features of both wave-front sensor based adaptive optics and model-free gradient descent optimization wave-front control.

In Section 2 we consider briefly two major wave-front control principles used in conventional adaptive systems: wave-front phase conjugation and model-free optimization. The D-SPGD technique is introduced in Section 3. Adaptive system architectures based on the D-SPGD wave-front control are described in Section 4. Numerical analysis of both low-resolution (127 control channels) and high-resolution (256×256 control channels) D-SPGD adaptive systems in the presence of atmospheric-turbulence-induced phase distortions is presented in Sections 4 and 5. Results include high-resolution adaptive wave-front control for both uniform and randomly modulated input wave intensities. The D-SPGD control algorithm demonstrates exceptionally fast convergence rates for both low- and high-resolution adaptive systems even in the presence of strong intensity scintillations. The reason for such fast convergence is the partial decoupling of control channels through incorporation of spatially distributed information from the wave-front sensor into the conventional SPGD wave-front control technique.

2. WAVE-FRONT INFORMATION PROCESSING IN ADAPTIVE OPTICS SYSTEMS

A. Wave-Front Phase Conjugation Adaptive Systems

Consider the two primary adaptive optics system architectures in Fig. 1 from a general information processing and control viewpoint. In both systems wave-front phase update is performed by applying control signals $\{u_l(t)\}$ to wave-front corrector electrodes. The phase $u(\mathbf{r}, t) = \Sigma_{l=1}^{N} u_l(t) S_l(\mathbf{r})$ introduced by the corrector is dependent on both the control parameters $\{u_l(t)\}$ and the influence functions $\{S_l(\mathbf{r})\}$. In the phase conjugation type adaptive system shown in Fig. 1a, phase correction is typically performed by sensing an uncompensated phase $\delta(\mathbf{r}, t) = u(\mathbf{r}, t) + \varphi(\mathbf{r}, t)$ (residual phase error) by using a wave-front sensor. The wave-front sensor transforms the phase error $\delta(\mathbf{r}, t)$ into an intensity distribution $I_\delta(\mathbf{r}, t) = F[\delta(\mathbf{r}, t)]$ that can be directly measured. The transformation $I_\delta = F[\delta]$ is dependent on the wave-front sensor type, but it is always nonlinear. As a result, additional reconstruction of the wave-front phase from measurements is required. This phase reconstruction can be replaced by the computationally more efficient calculation of phase error signals $\{\delta_l(t)\}$ $[\delta(\mathbf{r}, t) = \Sigma_{l=1}^{N} \delta_l(t) S_l(\mathbf{r})]$. The closed-loop feedback controller is designed to null these error signals.[1,19,20] For high-resolution control systems this strategy leads to phase conjugation: $\delta(\mathbf{r}, t) = 0$ or $u(\mathbf{r}, t) = -\varphi(\mathbf{r}, t)$.

B. Model-Free Optimization Adaptive Systems

A different signal-processing concept is used in model-free optimization adaptive optics systems (Fig. 1b). The only

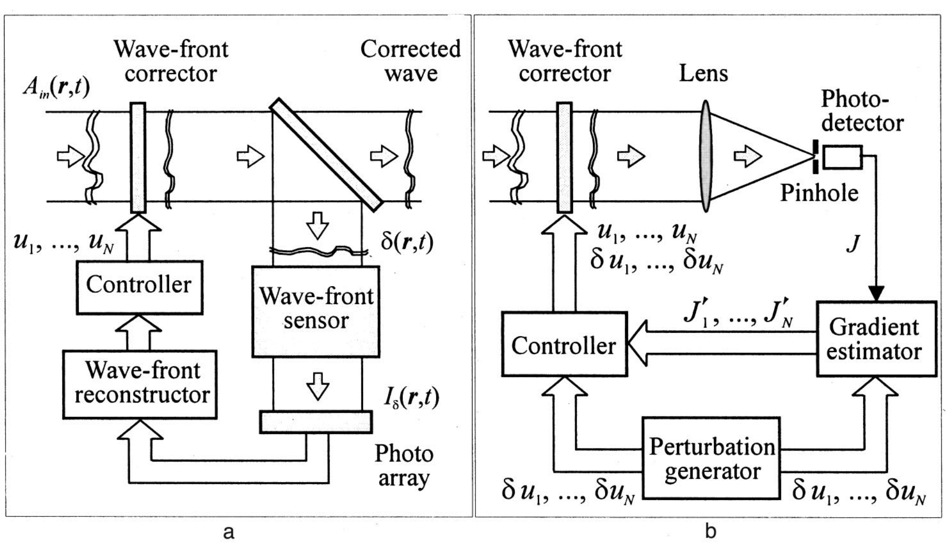


Fig. 1. Schematics for (a) wave-front phase conjugation and (b) model-free optimization adaptive system types.

information available for control is a performance metric $J(t)$. Depending on the adaptive system type, the metric J can be directly measured by a photodetector (e.g., Strehl ratio in Fig. 1b) or calculated by using the system's output intensity distribution $I_\delta(\mathbf{r}, t)$, for example, the sharpness function[21]:

$$J_1^s = \int I_\delta^2(\mathbf{r}, t)\mathrm{d}^2\mathbf{r}. \tag{1}$$

Although traditionally in the sharpness function (1) $I_\delta(\mathbf{r}, t)$ is defined as an intensity distribution in the image plane, we consider here a generalized definition assuming that $I_\delta(\mathbf{r}, t)$ in Eq. (1) represents the output intensity of a wave-front sensor. The wave-front sensor is envisioned here in general terms as an optical system transforming wave-front phase modulation $\delta(\mathbf{r}, t)$ into an intensity distribution $I_\delta(\mathbf{r}, t)$. This can be intensity in a lens image plane, intensity distribution of an interferometric or phase contrast sensor, intensity in a lenslet array focal plane, etc.

Perhaps the most efficient implementation of the model-free optimization technique for adaptive optics is the SPGD algorithm.[12,14] This algorithm allows on-chip implementation as a parallel very-large-scale-integration (VLSI) microprocessor.[12,22] In the SPGD technique, performance metric optimization is achieved by applying small random perturbations (vector $\delta\mathbf{u} = \{\delta u_l\}$, $l = 1,\ldots, N$) to wave-front corrector electrodes and measuring the corresponding perturbation of the metric δJ. The products $\{\delta J\,\delta u_l\}$ are used as an approximation $\{\hat{J}_l'\}$ of the true gradient components $\{J_l'\} = \{\partial J/\partial u_l\}$: $\{\hat{J}_l'\} = \{\delta J\,\delta u_l\}$. The feedback controller updates the control parameters in accordance with the following standard gradient descent procedures:

$$\frac{\mathrm{d}u_l(t)}{\mathrm{d}t} = -\gamma\hat{J}_l'(t), \tag{2a}$$

for continuous time and

$$u_l^{(n+1)} = u_l^{(n)} - \gamma\hat{J}_l'^{(n)} \qquad (l = 1,\ldots, N) \tag{2b}$$

for iterative techniques. The update coefficient γ in Eqs. (2) is positive for system performance metric minimization and negative otherwise.

This SPGD approach was recently applied for high-resolution wave-front control.[14] For a high-resolution piston-type wave-front corrector and a geometrically matched wave-front sensor the high-resolution SPGD controller can be described by the following spatially distributed model:

$$u^{(n+1)}(\mathbf{r}) = u^{(n)}(\mathbf{r}) - \gamma\delta J^{(n)}\delta u^{(n)}(\mathbf{r}), \qquad n = 0,\ 1,\ldots, \tag{2c}$$

where both the wave-front perturbation $\delta u^{(n)}(\mathbf{r})$ and the gradient estimation $\hat{J}'^{(n)}(\mathbf{r}) = \delta J^{(n)}\delta u^{(n)}(\mathbf{r})$ (gradient map) are spatially distributed functions.[14]

As mentioned in Section 1, the main drawback of adaptive optics techniques based on system performance metric optimization is the relatively slow convergence. The low convergence rate for SPGD results from strong cross coupling of the control parameters occurring in Eqs. (2) through the metric perturbation δJ. As described below, decoupling of the control equations (2) can significantly increase the adaptation process convergence rate. Complete decoupling of the control channels means that the wave-front control system can be considered as a set of N independent subsystems operating in parallel. The convergence rate of the entire control system is dependent on the convergence rate of the slowest subsystem but not on the control channel number N as in the case of the conventional gradient descent technique. In principle, the same goal—control channel decoupling and consequent convergence rate increase—can also be achieved by orthogonalization of the control equations (2). Indeed, representing the gradient components

$$\{\hat{J}_l'(t)\} \cong \sum_{k=1}^{N} \frac{\partial\hat{J}_l'}{\partial u_k}\xi_k(t) \equiv \sum_{k=1}^{N} A_{k,l}\xi_k(t)$$

in the form of a Taylor series expansion within a small vicinity of a stationary state solution $\{u_l^{(0)}\}$ of Eq. (2a) we obtain

$$\frac{\mathrm{d}\xi_l(t)}{\mathrm{d}t} = -\gamma\sum_{k=1}^{N} A_{k,l}\xi_k(t), \tag{2d}$$

where $\{\xi_l\} = \{u_l - u_l^{(0)}\}$. Orthogonalization (diagonalization) of the matrix $\{A_{k,l}\}$ in Eq. (2d) results in decoupling of the control equations. The main problem with this technique is that it requires knowledge of the matrix coefficients $\{A_{k,l}\}$. This information is typically not available. In Section 3 we show that in some practically important cases, control system decoupling (partial decoupling) can be realized on the basis of a different idea—the use of a spatially distributed (vector) system performance metric dependent on wave-front sensor data.

3. DECOUPLED STOCHASTIC PARALLEL GRADIENT DESCENT OPTIMIZATION

The metric perturbation δJ in the SPGD control technique describes the integrated system response to the control parameter perturbation $\delta\mathbf{u} = \{\delta u_l\}$. In accordance with the SPGD control rule (2) this integrated response (scalar value δJ) is passed to all individual channels in the form of the products $\{\delta J\,\delta u_l\}$ (gradient estimation). In many cases it is possible to obtain more accurate "distributed" or vector information about system response to the wave-front phase perturbation. The problem is how to obtain and utilize this distributed (vector) information in a gradient descent type optimization procedure. As shown below, use of a distributed performance metric (performance metric map) obtained from the wave-front sensor data can lead to control equation decoupling and result in a substantial improvement of the SPGD adaptive optics convergence rate.

A. Performance Metric Map and Metric Vector

Consider an adaptive optical system performance metric J that can be represented in the following cumulative form:

$$J(t) = \sum_{l=1}^{M} c_l j_l(t), \qquad j_l(t) = \int j(\mathbf{r}, t)Z_l(\mathbf{r})\mathrm{d}^2\mathbf{r}. \tag{3}$$

Expression (3) introduces definitions of the performance metric map $j(\mathbf{r}, t)$ and performance metric vector $\mathbf{j}(t) = \{j_1(t), \ldots, j_M(t)\}$. Both are dependent on the sensor's intensity distribution $I_\delta(\mathbf{r}, t)$. In expression (3), $\mathbf{c} = \{c_l\}$ is a weighting coefficient vector ($c_l \geqslant 0$ and $|\mathbf{c}| = 1$), and $\{Z_l(\mathbf{r})\}$ is a set of functions describing signal processing applied to the sensor data (sensor signal-processing functions). In most examples considered here we assume that the functions $\{Z_l(\mathbf{r})\}$ coincide with the wave-front corrector influence functions $\{S_l(\mathbf{r})\}$. The term "matched wave-front corrector and sensor" corresponds to $\{Z_l(\mathbf{r})\} = \{S_l(\mathbf{r})\}$.

Consider an adaptive system with piston-type wave-front corrector. The stepwise influence functions $\{S_l(\mathbf{r})\} = \{S_0(\mathbf{r} - \mathbf{r}_l)\}$ are defined inside nonoverlapping element areas Ω_l (subapertures) centered at grid points $\mathbf{r}_l$ [$S_0(\mathbf{r} - \mathbf{r}_l) = 1$ for $\mathbf{r}_l \in \Omega_l$ and $S_0(\mathbf{r} - \mathbf{r}_l) = 0$ otherwise]. For the matched corrector and sensor instead of expression (3) we obtain

$$J(t) = \sum_{l=1}^{N} c_l j_l(t), \qquad j_l(t) = \int_{\Omega_l} j(\mathbf{r}, t)\mathrm{d}^2\mathbf{r}. \tag{4}$$

The sharpness function (1) gives an example of the cumulative-type performance metric with the metric map: $j_l^s(\mathbf{r}, t) = I_\delta^2(\mathbf{r}, t)$.

B. Distributed (Decoupled) Stochastic Parallel Gradient Descent Controller

For optimization of the cumulative performance metric (3), consider the iterative SPGD procedure (2b). Represent the lth gradient component approximation $\tilde{J}_1' = \delta J \delta u_l$ in the SPGD algorithm in the form

$$\tilde{J}_l' = c_l \delta u_l \delta j_l(t) + \sum_{k \neq l}^{N} c_k \delta u_l \delta j_k(t).$$

Correspondingly, instead of Eq. (2b) we obtain

$$u_l^{(n+1)} = u_l^{(n)} - \gamma c_l \delta u_l \delta j_l^{(n)} - \gamma \sum_{k \neq l}^{N} c_k \delta u_l \delta j_k^{(n)}, \qquad (l = 1, \ldots, N). \tag{5}$$

As mentioned above, the slow convergence of the SPGD optimization technique is related to control channel cross coupling described by the last term in Eq. (5). If the cross-coupling term is small, the SPGD controller (5) can be described by the following system of equations:

$$u_l^{(n+1)} = u_l^{(n)} - \gamma c_l \delta j_l^{(n)} \delta u_l, \qquad (l = 1, \ldots, N). \tag{6}$$

Define the iterative procedure (6) for control parameter update as the SPGD controller with decoupled metric perturbation. The controller (6) might also be called a distributed SPGD controller, as it is based on performance metric vector information. The abbreviation D-SPGD controller covers both definitions. For neglect of the cross-coupling terms in Eq. (5), the following decoupling conditions should be fulfilled:

$$\left| \sum_{k \neq l}^{N} c_k \delta u_l \delta j_k \right| \ll |c_l \delta u_l \delta j_l| \qquad \text{for all } l = 1, \ldots, N. \tag{7}$$

Condition (7) suggests that the sum of cross-coupling terms for each control channel is small and the gradient estimation is dependent only on "local" information for each channel. Note that if condition (7) is fulfilled, both conventional SPGD [Eq. (2)] and D-SPGD [Eq. (6)] controllers perform optimization of the same system performance metric J but using different control parameter update rules.

C. Spatially Distributed (SD) Wave-Front Control: SD-SPGD Controller

A similar idea for control channel decoupling can be applied to high-resolution adaptive systems. For a geometrically matched high-resolution piston-type wave-front corrector and wave-front sensor, the high-resolution feedback controller can be described by the following spatially distributed iterative gradient descent procedure (SD-SPGD controller):

$$u^{(n+1)}(\mathbf{r}) = u^{(n)}(\mathbf{r}) - \gamma \delta j^{(n)}(\mathbf{r}) \delta u^{(n)}(\mathbf{r}), \qquad n = 0, 1, \ldots. \tag{8}$$

In the SD-SPGD adaptive system architectures, the wave-front phase update is based on the metric map perturbation $\delta j(\mathbf{r})$ resulting from the random wave-front perturbation $\delta u(\mathbf{r})$.

D. Decoupling Condition

Consider decoupling conditions (7)—the backbone of the D-SPGD wave-front-control technique. Assume that the wave-front sensor provides point-to-point mapping of wave-front phase into intensity as described by a function $I_\delta(\mathbf{r}) = I_\delta[u(\mathbf{r})]$. Assume also that the metric map $j(\mathbf{r})$ is a function of intensity $I_\delta : j = j(I_\delta)$. Taking into account the cumulative form of the system performance metric (3), the perturbations $\{\delta j_l\}$ in Eqs. (5)–(7) can be represented in the following form:

$$\delta j_l = \int \partial_I j(\mathbf{r}) \partial_{\mathbf{u}} I_\delta(\mathbf{r}) Z_l(\mathbf{r}) \delta u(\mathbf{r}) \mathrm{d}^2\mathbf{r} = \sum_{k=1}^{N} a_{k,l} \delta u_k, \tag{9}$$

where

$$a_{k,l} = \int \partial_I j(\mathbf{r}) \partial_{\mathbf{u}} I_\delta(\mathbf{r}) S_k(\mathbf{r}) Z_l(\mathbf{r}) \mathrm{d}^2\mathbf{r}. \tag{10}$$

Here $\partial_I j(\mathbf{r})$ and $\partial_u I_\delta(\mathbf{r})$ are the first variations of $j(\mathbf{r}) = j[I_\delta]$ and $I_\delta(\mathbf{r}) = I_\delta[u]$, and $\delta u(\mathbf{r}) = \Sigma_{k=1}^N \delta u_k S_k(\mathbf{r})$. The matrix $\{a_{k,l}\}$ in Eq. (10) (coupling matrix) describes cross coupling between control parameter perturbations and the corresponding perturbations of the metric vector components. When we take into account the representation (9), (10) for $\{\delta j_l\}$, the SPGD control rule (5) reads

$$u_l^{(n+1)} = u_l^{(n)} - \gamma c_l \sum_{k=1}^{N} a_{k,j} \delta u_l \delta u_k - \gamma \sum_{k=1}^{N} \sum_{n \neq l}^{N} c_n a_{k,n} \delta u_l \delta u_k, \quad (l = 1, \ldots, N). \tag{11}$$

In the conventional SPGD technique the perturbations $\{\delta u_l\}$ are statistically independent random variables having zero mean and equal variances: $\langle \delta u_l \delta u_k \rangle = \sigma^2 \delta_{l,k}$, where $\delta_{l,k}$ is the Kronecker symbol ($\delta_{l,k} = 1$ for $k = l$ and 0 otherwise). Recall that the sum of the last two terms in Eq. (11) is proportional to the gradient component approximation: $\hat{J}_l' = \delta J \delta u_l$. Hence for the expectation $\langle \delta J \delta u_l \rangle$ we obtain

$$\langle \delta J \delta u_l \rangle = \langle (\delta u_l)^2 \rangle \left(c_l a_{l,l} - \sum_{k \neq l}^{N} c_k a_{l,k} \right),$$

$$(l = 1, \ldots, N). \quad (12)$$

Thus, in the statistical sense, the decoupling condition (7) means that

$$\left| \sum_{k \neq l}^{N} a_{l,k} \right| \ll |a_{l,l}|, \qquad (l = 1, \ldots, N). \quad (13)$$

Consider an adaptive system with matched piston-type wave-front corrector and sensor: $\{Z_l(\mathbf{r})\} = \{S_l(\mathbf{r})\} = \{S_0(\mathbf{r} - \mathbf{r}_l)\}$, where $\{S_0(\mathbf{r} - \mathbf{r}_l)\}$ are stepwise functions. In this case the coupling matrix (10) is diagonal ($a_{l,k} = a_{l,l}\delta_{l,k}$) and the decoupling conditions in the form (13) are always satisfied ($\Sigma_{k \neq l}^{N} a_{l,k} = 0$). This means that instead of the conventional SPGD update rule (2b) for the cumulative performance metric (3) optimization, the D-SPGD controller in the form (6) can be used. Note that this result is valid for the assumption that the wave-front sensor provides point-to-point mapping $I_\delta(\mathbf{r}) = I_\delta[u(\mathbf{r})]$ and for a matched corrector and sensor having nonoverlapping stepwise functions. In most wave-front sensor types, phase intensity mapping is described by an operator $I_\delta = G[u]$ that may not preserve point-to-point correspondence between I_δ and u. In this case the intensity I_δ at a fixed point $\mathbf{r}_1$ may be dependent on the phase modulation belonging to a local area of $\mathbf{r}_1$ (local mapping) or even on phase modulation belonging to the entire wave-front sensor aperture (global mapping). The decoupling condition for these cases requires a separate analysis.

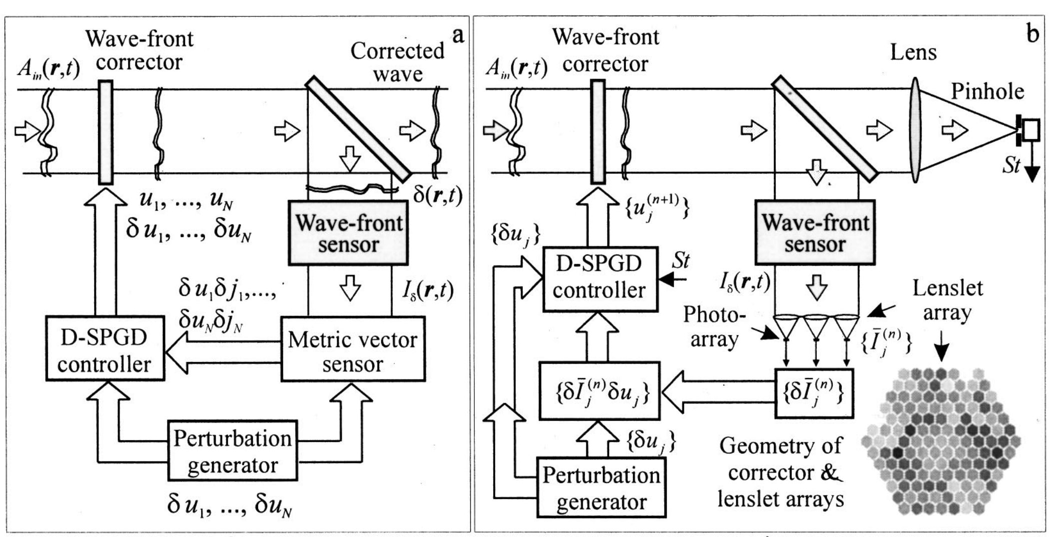
4. DECOUPLED STOCHASTIC PARALLEL GRADIENT DESCENT ADAPTIVE SYSTEM ARCHITECTURES

The schematic of the adaptive system based on SPGD optimization with distributed (decoupled) metric is shown in Fig. 2a. The D-SPGD adaptive optics system consists of the following major components: (a) wave-front sensor for transforming residual wave-front phase modulation $\delta(\mathbf{r}, t)$ into an intensity distribution $I_\delta(\mathbf{r}, t)$, (b) metric vector sensor performing measurements and calculations of the components $\{j_l\}$ and perturbations $\{\delta j_l\}$, (c) generator for the random perturbations $\{\delta u_l\}$ applied to the wave-front corrector electrodes and used for calculation of the update signals $\delta j_l \delta u_l$, and (d) the D-SPGD controller for computing the control parameters $\{u_l\}$ in accordance with the iterative procedure (6) and for supplying the perturbations $\{\delta u_l\}$ and controls $\{u_l\}$ to the wave-front corrector electrodes.

A. Requirements for Wave-Front Sensing Devices

In different types of adaptive systems, wave-front sensor information is utilized in different ways. In phase-conjugation adaptive systems, wave-front sensor data [intensity $I_\delta(\mathbf{r}, t)$] are used for phase (controls) reconstruction. Correspondingly, the choice of wave-front sensor is dictated by the convenience and simplicity of solving the phase reconstruction problem.

For the adaptive wave-front phase distortion compensation technique (D-SPGD controller) introduced here, the criteria for wave-front sensor choice are different. In D-SPGD-type adaptive systems the wave-front sensor should provide a relatively simple measurement (calculation) of the metric map $j(\mathbf{r}, t)$ and/or metric vector $\mathbf{j}(t)$ corresponding to a cumulative system performance metric J. The second important constraint is the decoupling condition that should be fulfilled for the selected performance metric J, sensor type, and wave-front corrector type.



a b

Fig. 2. D-SPGD adaptive system architectures: (a) general schematic, (b) adaptive system with J_3 controller (16a). The geometry of the matched wave-front corrector and 127 subaperture sensor is shown at bottom right.

B. Performance Metrics and Control Algorithms

Consider performance metrics and wave-front sensor types that can be used in D-SPGD adaptive wave-front control systems. For all examples given in this section we assume that wave-front phase aberration correction is performed by using a piston-type wave-front corrector described by stepwise influence functions $\{S_0(\mathbf{r}-\mathbf{r}_l)\}$ and defined over nonoverlapping areas $\{\Omega_l\}$, e.g., having rectangular or hexagonal geometry. We also assume that the wave-front sensor and corrector are matched. Consider the following cumulative performance metrics dependent on the wave-front sensor output intensity distribution $I_\delta(\mathbf{r}, t)$:

$$J_1 = \sum_{l=1}^{N} j_l^{(1)} \equiv \sum_{l=1}^{N} \int_{\Omega_l} I_\delta^2(\mathbf{r})\mathrm{d}^2\mathbf{r}; \tag{14a}$$

$$J_2 = \sum_{l=1}^{N} j_l^{(2)} \equiv \sum_{l=1}^{N} (\bar{I}_l)^2, \qquad \bar{I}_l = \int_{\Omega_l} I_\delta(\mathbf{r})\mathrm{d}^2\mathbf{r}, \tag{14b}$$

$$J_3 = \sum_{l=1}^{N} j_l^{(3)} \equiv \sum_{l=1}^{N} \bar{I}_l. \tag{14c}$$

Here $\bar{I}_l$ is the sensor intensity averaged over the subaperture area Ω_l. The metric vector components $\{j_l\}$ in performance metrics (14) are obtained by integrating the metric maps over the subaperture areas $\{\Omega_l\}$ corresponding to wave-front corrector elements. Performance metric (14a) coincides with the sharpness function (1). Computation of the metric vector $\{j_l^{(1)}\}$ requires squaring the intensity data before integration over the subaperture areas $\{\Omega_l\}$. From a computational viewpoint the performance metric (14b) is more efficient. In it the wave-front-sensor output intensity distribution is first integrated over the subaperture areas, and the integration results are then squared. Note that Eqs. (6) describing the D-SPGD controller depend only on the perturbations $\{\delta j_l\}$ but not on the metric vector components $\{j_l\}$ themselves. For the metric J_2 [Eq. (14b)], the perturbation $\delta j_l^{(2)}$ can be represented in the following equivalent form: $\delta j_l^{(2)} = 2\bar{I}_l\delta\bar{I}_l$. Correspondingly, for this metric the D-SPGD controller (6) is given by

$$u_l^{(n+1)} = u_l^{(n)} - \gamma\bar{I}_l^{(n)}\delta\bar{I}_l^{(n)}\delta u_l, \qquad (l = 1, \ldots, N), \tag{15a}$$

where the factor of 2 is included in the coefficient γ. Note that for high-resolution systems, metrics (14a) and (14b) coincide. The spatially distributed model of the D-SPGD controller optimizing sharpness functions (14a) or (14b) reads

$$u^{(n+1)}(\mathbf{r}) = u^{(n)}(\mathbf{r}) - \gamma I_\delta^{(n)}(\mathbf{r})\delta I_\delta^{(n)}(\mathbf{r})\delta u^{(n)}(\mathbf{r}), \qquad (n = 0, 1, \ldots,). \tag{15b}$$

The performance metric J_3 [Eq. (14c)] is proportional to the sensor's output wave power. The D-SPGD controller for the metric J_3 has the simplest form:

$$u_l^{(n+1)} = u_l^{(n)} - \gamma\delta\bar{I}_l^{(n)}\delta u_l, \qquad (l = 1, \ldots, N). \tag{16a}$$

Correspondingly, the spatially distributed model can be represented by the following equation:

$$u^{(n+1)}(\mathbf{r}) = u^{(n)}(\mathbf{r}) - \gamma\delta I_\delta^{(n)}(\mathbf{r})\delta u^{(n)}(\mathbf{r}), \qquad (n = 0, 1, \ldots,). \tag{16b}$$

A schematic for the adaptive system with D-SPGD controller (16a) [referred to as the J_3 controller] is shown in Fig. 2b. It includes a geometrically matched wave-front corrector with a photoreceiver system composed of a lenslet array and a photoarray placed in its focal plane. The photoarray output signals are proportional to the intensities $\{\bar{I}_l\}$ averaged over the subapertures. The D-SPGD J_3-controller operation at the nth step of the iteration process (16a) include (a) measurement of metric vector components $\{j_l^{(n)}\} = \{\bar{I}_l^{(n)}\}$, (b) generation of the random (pseudorandom) perturbations $\{\delta u_l\}$ applied to the wave-front-corrector electrodes, (c) measurement of the metric vector components $\{j_l^{(n+)}\} = \{\bar{I}_l^{(n+)}\}$ corresponding to the perturbed control parameters $\{u_l^{(n)} + \delta u_l\}$, (d) calculation of the corresponding metric vector perturbations $\{\delta j_l^{(n)}\} = \{\delta\bar{I}_l^{(n)}\} = \{\bar{I}_l^{(n+)} - \bar{I}_l^{(n)}\}$, and (e) computation of the products $\delta\bar{I}_l^{(n)}\delta u_l$ and control parameter update.

The D-SPGD technique can also be implemented as a continuous-time controller. In this case control-parameter perturbations $\{\delta u_l\} = \{\alpha_l(t)\sin\omega t\}$ in the form of harmonic signals with small random amplitudes $\{\alpha_l(t)\}$ and the same dithering frequency ω are applied simultaneously to all control channels. The perturbed metric vector components $\{\delta j_l\} = \{\beta_l(t)\sin\omega t + \chi_l(t)\sin 2\omega t + \ldots\}$ are demodulated by synchronous detectors to obtain modulation amplitudes $\{\beta_l\}$ used in the continuous-time controller:

$$\tau\frac{\mathrm{d}u_l(t)}{\mathrm{d}t} = -\gamma\alpha_l(t)b_l(t). \tag{16c}$$

This approach is similar to the multidithering technique well known in adaptive optics.[10,13] The important difference is that here only a single dithering frequency is used.

C. Interferometer with a Reference Wave

The conventional interferometer presents an example of a wave-front sensor that provides two-dimensional point-to-point mapping of wave-front phase to the sensor's intensity distributions $I_\delta(\mathbf{r}, t) = F[\delta(\mathbf{r}, t)]$. For the interferometer with a reference wave we have

$$I_\delta(\mathbf{r}, t) = I_0(\mathbf{r}) + I_{\mathrm{in}}(\mathbf{r}) + 2\mu(\mathbf{r})\cos[\delta(\mathbf{r}, t) + \Delta], \tag{17}$$

where $I_0(\mathbf{r})$ and $I_{\mathrm{in}}(\mathbf{r})$ are reference and input wave intensities, $\mu(\mathbf{r})$ is the visibility of the interference pattern, and Δ is a constant phase shift. Assume for simplicity uniform intensities $I_0(\mathbf{r}) = I_0$ and $I_{\mathrm{in}}(\mathbf{r}) = I_{\mathrm{in}}$. It can be easily shown that each of the introduced performance metrics (14) has a global minimum corresponding to wave-front phase distortion compensation: $\delta(\mathbf{r}, t) + \Delta = \text{constant}$ or $u(\mathbf{r}, t) = -\varphi(\mathbf{r}, t) + \text{constant}$. Assume that the photoreceiver system of the interferometric wave-front sensor is composed of a lenslet array with subapertures matched to piston-type wave-front corrector elements providing direct measurement of the subaperture averaged intensities $\{\bar{I}_l\}$, as shown in Fig. 2b. For this corrector and sensor, the decoupling condition (13) is fulfilled. Thus minimization of the performance metrics (14) can be performed by using the D-SPGD controller (16).

An interferometer with a reference wave is an ideal sensor for D-SPGD adaptive optics because it provides point-to-point phase and intensity mapping. Unfortunately, this sensor is not practical for most adaptive optics applications because it requires the presence of an undistorted reference wave.

D. Phase-Contrast Sensors

Consider wave-front sensors based on the phase contrast technique: the Zernike filter (ZF) and point diffraction interferometer (PDI) shown in Fig. 3.[23–25] These sensors do not require a reference wave but are vulnerable to large-amplitude wave-front tilts. Both wave-front sensors consist of two lenses with a phase changing (ZF) or an absorbing (PDI) element (spatial filter) placed in the lenses' common focal plane. The spatial filter has a small circular region (a dot) in the middle of the focal plane. The dot introduces either a phase shift θ near $\pi/2$ rad (Zernike sensor) or attenuation by a factor $\gamma < 1$ (PDI sensor) applied only to input wave low-order spatial spectral components. For a simplified model corresponding to a focal plane filter affecting only the zero-order spatial spectral component, the wave-front sensor's output intensity is given by

$$I_\delta(\mathbf{r}) = I_{\text{in}}(\mathbf{r}) + 2(2\pi F)^2 I_F(0) - 4\pi F I_{\text{in}}^{1/2}(\mathbf{r}) I_F^{1/2}(0) \times \{\cos[\delta(\mathbf{r}) - \Delta] - \sin[\delta(\mathbf{r}) - \Delta]\} \tag{18a}$$

for the Zernike wave-front sensor and

$$I_\delta(\mathbf{r}) = I_{\text{in}}(\mathbf{r}) + (2\pi F)^2 I_F(0) - 4\pi F I_{\text{in}}^{1/2}(\mathbf{r}) I_F^{1/2}(0) \cos[\delta(\mathbf{r}) - \Delta] \tag{18b}$$

for the PDI with $\gamma = 0$.[25] Here F is the lens focal length normalized by the diffraction parameter kR^2 (where $k = 2\pi/\lambda$ is the wave number and R is the lens aperture radius). In Eqs. (18) $I_F(0)$ and Δ are the intensity and phase of the zero-order spatial spectral component $\mathbf{q} = 0$ ($\mathbf{q}$ is a wave vector) defined as[25]

$$I_F(\mathbf{q} = 0) = \left| \int I_{\text{in}}^{1/2}(\mathbf{r}) \exp[i\delta(\mathbf{r})] \mathrm{d}^2\mathbf{r} \right|^2,$$

$$\Delta = \arg\left[\int I_{\text{in}}^{1/2}(\mathbf{r}) \exp[i\delta(\mathbf{r})] \mathrm{d}^2\mathbf{r} \right]. \tag{19}$$

The normalized value of $I_F(0)$ is known as the Strehl ratio, $\mathrm{St} = I_F(0)/I_F^0$, where I_F^0 is the intensity of the zero-order component in the absence of phase aberrations. The output intensity distributions (18) are similar to the interference pattern in the conventional interferometer with reference wave (17). The important difference is that both phase-contrast sensors do not preserve point-to-point phase intensity mapping as does a conventional interferometer. The output intensities (18) contain quantities $I_F(0)$ and Δ [see Eq. (19)] dependent on the phase modulation $\delta(\mathbf{r})$ over the entire input sensor aperture, resulting in global phase intensity coupling. In fact, in the phase contrast sensors both point-to-point [through sin and cos terms in Eqs. (18)] and global [through $I_F(0)$ and Δ] phase intensity mapping types are present. Results of the numerical analysis described in Sections 5 and 6 show that despite the presence of this global coupling both the ZF the PDI can be used as wave-front sensors in D-SPGD adaptive systems.

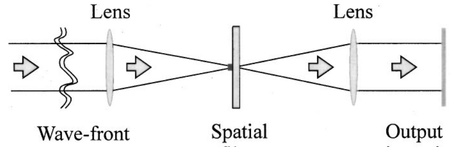


Fig. 3. Phase contrast wave-front sensor schematic.

In the following two sections we consider numerical analysis of two D-SPGD adaptive systems having significantly different spatial resolutions for wave-front distortion correction: a low-resolution adaptive system with 127 control channels and a high-resolution spatially distributed system with 256×256 control channels.

5. LOW RESOLUTION DECOUPLED STOCHASTIC PARALLEL GRADIENT DESCENT ADAPTIVE SYSTEM: PERFORMANCE ANALYSIS

A. System Model

Consider a low-resolution D-SPGD adaptive system (Fig. 2b) with a wave-front corrector composed of 127 hexagonally shaped elements with piston-type influence functions. Geometry of the corrector elements (mirrors or LC cells) is shown in the bottom-right corner of Fig. 2b. For the 256×256 pixel numerical grid the wave-front-corrector fill factor was near 88.3%. The adaptive system receiver is composed of a lens array containing similar hexagonally shaped subapertures geometrically matched with the subapertures of the wave-front corrector. The lenslet array focuses the sensor's output wave onto a photodetector array placed in its focal plane. Thus the receiver system provides measurement of the signals $\bar{I}_l$ ($l = 1, \ldots, 127$), proportional to the intensities averaged over the lens subapertures. The three wave-front sensors described above were considered: interferometer with reference wave, ZF, and PDI.

B. Phase Aberrations

D-SPGD adaptive system performance was analyzed by using two types of phase distortion: phase distortions that can be completely removed by the adaptive system corrector ("corrector friendly" phase aberrations) and phase distortions modeling atmospheric-turbulence-induced phase fluctuation ("atmospheric-like" phase aberrations). For corrector friendly phase distortions we know that ideal compensation can potentially be achieved and that the compensation level, as well as the adaptation convergence rate, characterizes only the control algorithm efficiency. Atmospheric-like phase aberrations were introduced in order to examine D-SPGD adaptive system efficiency in the presence of phase distortion components that low-resolution adaptive systems cannot completely compensate. This adaptation situation characterizes robustness of the control algorithm in the presence of wave-front phase noise.

Digital realizations of a statistically homogeneous and isotropic random function $\varphi(\mathbf{r})$ with zero mean and Andrews power spectrum model were used for generation of both corrector friendly and atmospheric-like phase distortions. The Andrews power spectrum model reads[26]

$$G_A(q) = 2\pi 0.033(1.68/r_0)^{5/3}(q^2 + q_A^2)^{-11/6}\exp(-q^2/q_a^2) \times [1 + 1.802(q/q_a) - 0.254(q/q_a)^{7/6}]. \qquad (20)$$

Here r_0 is the Fried parameter,[27] $q_A = 2\pi/l_{\mathrm{out}}$, and $q_a = 2\pi/l_{\mathrm{in}}$, where l_{out} and l_{in} are the outer and inner scales of turbulence.

An example of an atmospheric-like phase aberration realization (phase screen) superimposed with wave-front corrector and -sensor subaperture elements is shown in Fig. 4a. Random realizations of the corrector friendly aberrations $\varphi^h(\mathbf{r})$ were obtained from $\varphi(\mathbf{r})$ by approximating the functions $\varphi(\mathbf{r})$ by stepwise influence functions $\{S_0(\mathbf{r} - \mathbf{r}_l)\}$ defined over hexagonal shaped subapertures $\{\Omega_l\}$ $(l = 1,..., 127)$. An example of a corrector-friendly aberration is shown in Fig. 4b. Strength of the input phase aberration was characterized by the standard deviation of the phase fluctuations averaged over the wave-front corrector aperture $\sigma_\varphi = [\int \varphi^2(\mathbf{r}) d^2\mathbf{r}]^{1/2}$ and by the Strehl ratio St. The amplitude of the introduced phase distortions was varied by changing the value of the Fried parameter r_0 in Eq. (20). For each fixed value r_0, 50 phase screens were generated. Simulations were performed for an input wave with a uniform intensity distribution and phase aberrations $\varphi(\mathbf{r})$ or $\varphi^h(\mathbf{r})$ located in the wave-front corrector pupil plane.

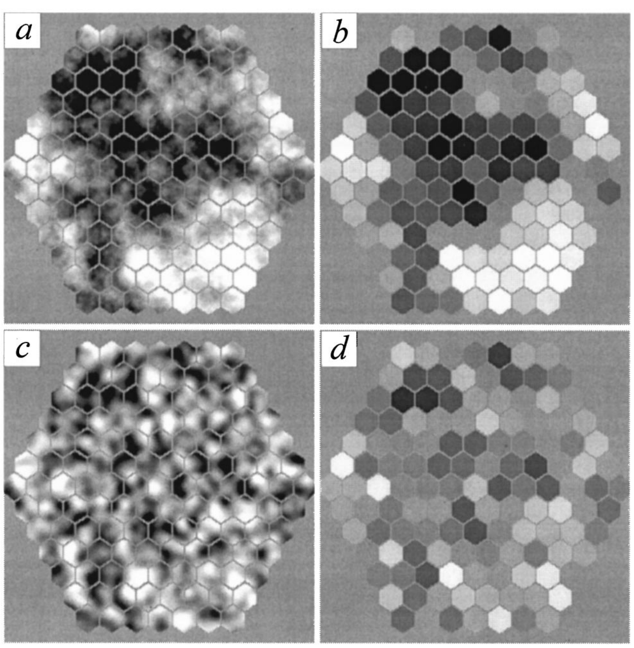


Fig. 4. Input-wave phase distortion/perturbation realizations: (a) atmosphericlike phase aberration pattern superimposed with hexagonal grid of wave-front corrector/sensor subapertures, (b) corrector friendly aberrations $\varphi^h(\mathbf{r})$ corresponding to $\varphi(\mathbf{r})$, (c) phase perturbation $\delta u(\mathbf{r})$ (Gaussian spectrum), (d) corrector-friendly wave-front perturbation $\delta u^h(\mathbf{r})$ corresponding to $\delta u(\mathbf{r})$.

C. Wave-Front Perturbations

To generate corrector friendly wave-front phase perturbations $\delta u^h(\mathbf{r}) = \Sigma_{l=1}^{127} \delta u_l S_l(\mathbf{r})$, the realizations of a statistically homogeneous and isotropic random function $\delta u(\mathbf{r})$ with zero mean and Gaussian power spectrum $G(q) = G_0 \exp(-2q^2/q_\varphi^2)$ were used. Here the characteristic spatial bandwidth $q_\varphi = 2\pi/l_p$, where l_p is the spatial-correlation radius. In all calculations the correlation radius was fixed: $l_p = 1.0d$, where d is the wave-front-corrector subaperture radius. Control parameter random perturbations $\{\delta u_l\}$ were obtained by decomposition of $\delta u(\mathbf{r})$ over the influence functions $\{S_0(\mathbf{r} - \mathbf{r}_l)\}$. Examples of a phase perturbation $\delta u(\mathbf{r})$ corresponding to a Gaussian spectrum and the corresponding corrector-friendly wave-front perturbation $\delta u^h(\mathbf{r})$ are shown in Figs. 4c and 4d, respectively.

D. Feedback Control

Adaptive wave-front control was performed by using the D-SPGD control algorithms (15a) and (16a) corresponding to J_2 and J_3 metric [Eq. (14)] optimization (J_2 and J_3 controllers). D-SPGD adaptive system efficiency was compared with an adaptive system based on Strehl ratio optimization by using the conventional SPGD algorithm (2b) (SPGD controller). To improve the adaptation process convergence rate, the update coefficient γ in all of the control algorithms used here was a function of the Strehl ratio St. The following empirical rule was used to adjust the coefficient γ during the adaptation process: $\gamma = \gamma_0(1 + \mathrm{St}^{-1})$, where γ_0 is a constant. At the beginning of the adaptation process when the Strehl ratio St is relatively small the parameter γ is large, providing for significant control-parameter change during a single iteration. Decreasing γ (following the Strehl ratio increase) prevents the system from oscillatory instability in the vicinity of the metric extremum. Using an adaptive γ led to approximately a 20% improvement in the convergence rate for all examined control algorithms.

E. Decoupled Stochastic Parallel Gradient Descent Adaptive System with Interferometric Sensor

Consider first the numerical analysis results for the D-SPGD systems with an interferometric wave-front sensor. Both turbulence-like and corrector friendly phase aberration realizations $\varphi(\mathbf{r})$ and $\varphi^h(\mathbf{r})$ (phase screens) were generated for a fixed value of the Fried parameter r_0. In all numerical experiments described here the same 50 phase screen realizations were used. A fixed number of 100 iterations corresponding to the D-SPGD wave-front control algorithms (15a) and (16a) were performed for each realization of the phase distorted input field. Adaptive system performance was evaluated by using J_2 and J_3 metric values as well as the Strehl ratio St calculated at each iteration $n = 1,..., 100$. The adaptation process was repeated for each phase distortion realization, and the obtained dependences $J_2(n)$, $J_3(n)$, and $\mathrm{St}(n)$ (adaptation evolution curves) were averaged. For the corrector friendly aberrations $\varphi^h(\mathbf{r})$ the averaged metric values $\langle J_2\rangle$, $\langle J_3\rangle$ and the corresponding Strehl ratios $\langle \mathrm{St}\rangle$ are presented in Fig. 5a as functions of the iteration number n ($\langle\,\rangle$ denotes ensemble averaging over the phase distortion realizations). For the set of phase

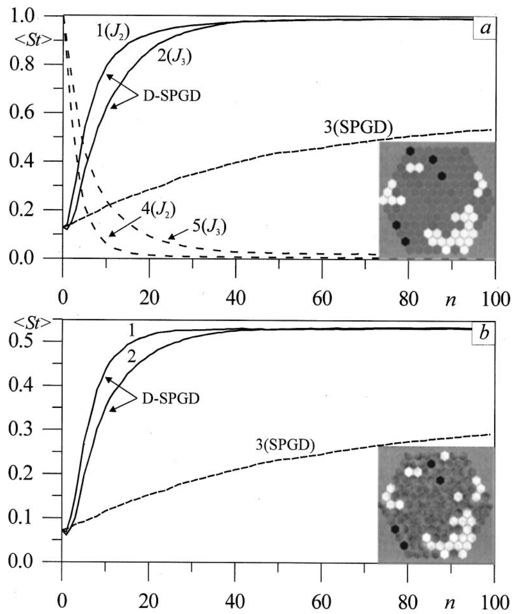


Fig. 5. Simulation results for the low-resolution D-SPGD adaptive system with interferometric wave-front sensor: (a) for corrector friendly aberrations $\varphi^h(\mathbf{r})$, (b) for atmospheric-like phase distortions $\varphi(\mathbf{r})$. Averaged Strehl ratio adaptation curves 1 and 2 and averaged metric curves 4 and 5 corresponding to the D-SPGD algorithms (15a) (J_2 curves) and (16a) (J_3 curves): 1 and 4 for controller (15a) and 2 and 5 for (16a). The averaged adaptation curves 3 correspond to the conventional SPGD controller (2a). At the bottom-right corner are shown residual phase patterns: (a) $\delta^h(\mathbf{r})$, (b) $\delta(\mathbf{r})$. The patterns' gray-scale dynamical range is 4π rad.

screens used, the ensemble-averaged standard deviation for input-field phase fluctuations $\sigma_{\text{in}} = \langle \sigma_\varphi \rangle$ was $\sigma_{\text{in}} = 1.67$ rad and the corresponding Strehl ratio $\langle \text{St} \rangle = 0.13$.

The adaptive system with D-SPGD controller (curves 1 and 2) demonstrated significantly faster convergence for both (15a) and (16a) controllers than did the conventional SPGD control algorithm (curve 3). Although adaptive wave-front control in the D-SPGD-type systems discussed here is based on minimization of the performance metrics J_2 and J_3, for the sake of convenience in comparing results, adaptation efficiency is characterized principally by the Strehl ratio. In Fig. 5a both normalized averaged performance metrics $\langle J_2 \rangle$, $\langle J_3 \rangle$ (curves 4 and 5) and the corresponding averaged Strehl ratio evolution curves (1 and 2) are presented together. As seen from the adaptation curves the D-SPGD feedback control algorithms provided for Strehl ratio increase up to a level of 80% during the first $N_{80\%} \sim 10$–15 iterations and to the 95% level during $N_{95\%} \sim 50$ iterations. With the conventional SPGD control algorithm the corresponding convergence rate was on the order of $N_{80\%} \sim 250$–300 iterations. The fastest convergence was achieved with the D-SPGD controller (15a) (curve 1) minimizing metric J_2. Nevertheless, the computationally more efficient control algorithm (16a) corresponding to minimization of the metric J_3 (wave-front-sensor output power) resulted in only about a 10–15% increase in the convergence rate at the 80% Strehl ratio level and had practically the same convergence rate for the 95% Strehl ratio level (compare curves 1 and 2).

Atmosphericlike phase-distortion compensation results for the D-SPGD adaptive system with interferometric wave-front sensor are shown in Fig. 5b. The Strehl ratio achieved here is approximately half that achieved during the corrector friendly aberration compensation in Fig. 5a. This decrease in the Strehl ratio is caused by the presence of phase distortion components that low-resolution adaptive systems cannot compensate. Nevertheless, the convergence rate remains approximately the same as for corrector friendly aberrations (compare adaptation curves in Figs. 5a and 5b). An example of the residual phase patterns $\delta(\mathbf{r})$ and $\delta^h(\mathbf{r})$ achieved after 100 iterations of the adaptation process corresponding to atmospheric-like (Fig. 4a) and corrector friendly (Fig. 4b) phase aberrations are shown in Fig. 5. Both residual phase patterns have 2π phase cuts (phase jumps). Despite the presence of 2π phase cuts the achieved adaptation level for corrector friendly aberrations was near $\langle \text{St} \rangle \approx 0.995$.

F. Decoupled Stochastic Parallel Gradient Descent System with Phase Contrast Sensors

As has been pointed out, the interferometer is an ideal sensor for a D-SPGD-type adaptive controller because it provides point-to-point intensity phase mapping and hence control channel decoupling. On the other hand, the interferometric sensor cannot be used in most adaptive optics systems, as it requires an undistorted reference wave. Consider the more practical D-SPGD system configurations based on the phase contrast sensors described in Section 4. Numerical simulations were performed for both PDI and ZF wave-front sensors. For each sensor we used two D-SPGD controllers: Eqs. (15a) (J_2 controller) and (16a) (J_3 controller). The ensemble-averaged adaptation evolution curves $\langle \text{St} \rangle$ for each of the examined systems are shown in Fig. 6a for corrector friendly and in Fig. 6b for atmospheric-like aberrations. For input field phase aberrations we used the same set of phase screens as for the D-SPGD system with interferometric sensor. For this reason the efficiency of the phase contrast sensor D-SPGD systems can be directly compared with the efficiency of the adaptive system with interferometric sensor given by curves 1 in Fig. 6. For systems with a phase contrast sensor, as seen from the adaptation curves in Fig. 6a the fastest convergence was exhibited by the D-SPGD system with J_2 controller and Zernike sensor (curve 2). A comparable convergence rate was observed for the D-SPGD system with PDI sensor (curves 3 and 4). The J_3 controller with Zernike wave-front sensor (curves 5) had the worst convergence rate. This result is quite expected. Recall that metric J_3 is proportional to the wave-front sensor output field power and that the ZF affects only the wave's zero spatial spectral component phase. For a Zernike sensor with unlimited aperture the output power is a constant value, and hence metric J_3 is independent of wave-front phase change. In the adaptive system that we analyzed, the ZF had a limited aperture. For this reason the dependence of J_3 on phase aberrations is weak, resulting in relatively slow adaptive system convergence.

As seen from Fig. 6 the D-SPGD controllers with phase-contrast sensors had approximately ten times faster convergence than did the SPGD controller (compare curves 2,

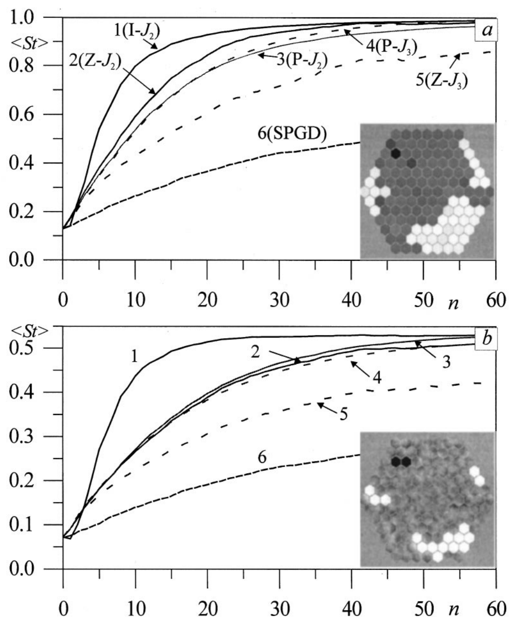


Fig. 6. Averaged Strehl ratio versus iteration number for the low-resolution D-SPGD adaptive system with different wave-front sensor types: (a) for corrector friendly, (b) for atmospheric-like phase distortions. The adaptation curves 1–5 correspond to 1, interferometer with reference wave; 2 and 5, ZF; 3 and 4, PDI. Labels in parentheses correspond to wave-front sensor type (I is interferometer, etc.) and controller [J_2 for (15a) and J_3 for (16a)]. The curves 6 correspond to the SPGD controller. Gray-scale images show residual phase patterns corresponding to the D-SPGD system with PDI (dynamical range is 4π rad.).

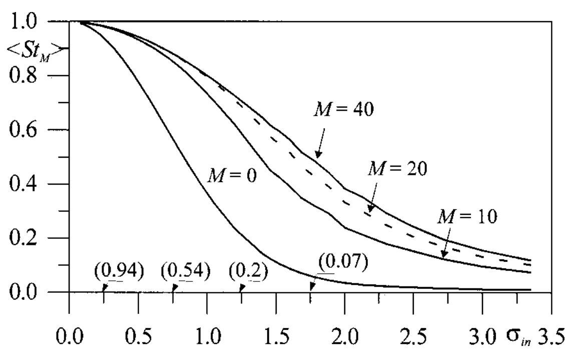


Fig. 7. The averaged Strehl ratios $\langle \mathrm{St}_M \rangle$ achieved after M iterations of the adaptation process versus input phase standard deviation σ_{in} for the low-resolution D-SPGD adaptive system with PDI and control algorithm (16a) (J_3 controller). Numbers in parentheses correspond to Strehl ratio values for σ_{in}.

3, and 4 in Fig. 6 with SPGD curves 6). The residual phase patterns in Fig. 6 show the existence of 2π phase cuts similar to the 2π phase cuts observed with the interferometric sensor (Fig. 5).

Consider numerical analysis results of the D-SPGD system with PDI (J_3 controller) performed for different values of the Fried parameter r_0. Adaptive system performance was evaluated by using a Strehl ratio St_M achieved after M iterations. For fixed r_0 the adaptation process was repeated with 50 phase distortion realizations, and the obtained values St were averaged. The averaged Strehl ratios $\langle \mathrm{St}_M \rangle$ are presented in Fig. 7 as functions of the ensemble-averaged phase fluctuation standard deviation of the input field $\sigma_{\mathrm{in}} = \langle \sigma_\varphi \rangle$ for different iteration numbers M. The results in Fig. 7 show that efficient phase distortion compensation can be achieved over a wide range of phase distortion amplitudes σ_{in} even with a relatively low-resolution adaptive system. The small differences between curves corresponding to $M = 20$ and $M = 40$ iterations suggests that saturation of the adaptation process occurs during the first 20–40 iterations.

G. Decoupled Stochastic Parallel Gradient Descent System with Continuously Deformed Mirror

In the D-SPGD adaptive optics system analyzed above, we assumed a geometrical match between the wave-front corrector and sensor: $\{Z_l(\mathbf{r})\} = \{S_l(\mathbf{r})\}$. This can be easily achieved in the case of a piston-type wave-front corrector having stepwise influence functions $\{S_0(\mathbf{r} - \mathbf{r}_l)\}$. Consider now a wave-front corrector having a continuously deformed surface with the Gaussian-type influence functions $S(\mathbf{r} - \mathbf{r}_l) = \exp(-|\mathbf{r} - \mathbf{r}_l|^2/w^2)$ ($l = 1, \ldots, 127$), where w is influence function width. Assume that corrector and sensor are not matched but still have similar geometries: vectors $\{\mathbf{r}_l\}$ coincide with centers of hexagonally shaped sensor subapertures. Sensor subapertures are described by stepwise functions $\{Z_0(\mathbf{r} - \mathbf{r}_l)\}$ defined over nonoverlapping hexagonal areas $\{\Omega_l\}$ of radius d. A numerical analysis was performed for the D-SPGD adaptive system with PDI and J_3 controller. For a fixed value w of the influence function width 50 corrector friendly phase screens $\varphi^{\mathrm{g}}(\mathbf{r})$ were generated. An example of corrector friendly aberrations $\varphi^{\mathrm{g}}(\mathbf{r})$ composed by Gaussian influence functions $\{S(\mathbf{r} - \mathbf{r}_l)\}$ with $w = d$ together with sensor subaperture geometry is shown in Fig. 8a. Patterns of the residual

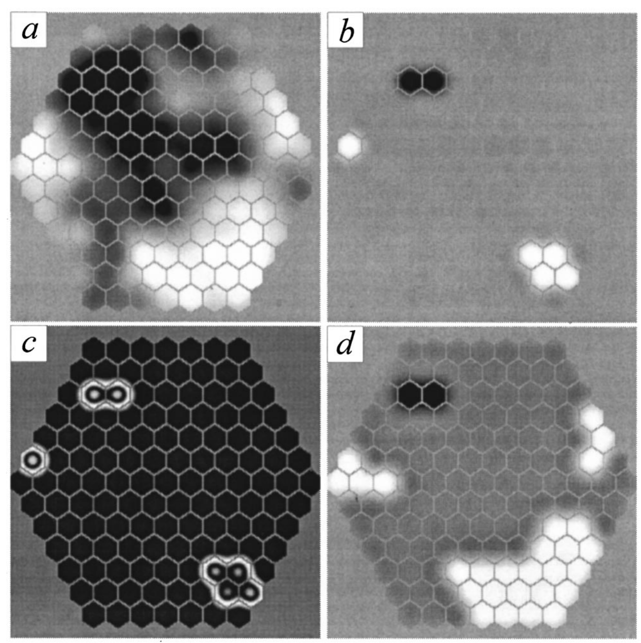


Fig. 8. Adaptive system with "mismatched" wave-front corrector and sensor: (a) sensor friendly aberration $\varphi^g(\mathbf{r})$ superimposed with a hexagonal grid of wave-front sensor subapertures for $w = d$, (b) pattern of the residual phase aberration $\delta^g(\mathbf{r})$, (c) sensor output intensity $I_\delta(\mathbf{r})$; both (b) and (c) are after 60 iterations of the D-SPGD controller (16a); (d) the same as (b) for $w = 1.2d$.

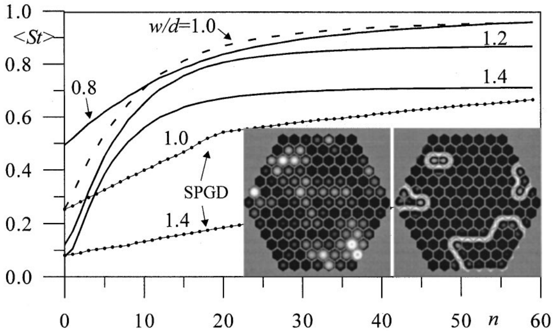


Fig. 9. Averaged Strehl ratio adaptation curves for corrector friendly aberration compensation with the D-SPGD system and "mismatched" wave-front corrector and sensor. Curves are labeled by the mismatch parameter w/d. Evolution curves for the SPGD controller are indicated by dots. Gray-scale images correspond to the intensity patterns $I_\delta(\mathbf{r})$ after 60 iterations: left, for $w/d = 0.8$; right, for $w/d = 1.2$.

phase aberration $\delta^{\mathrm{g}}(\mathbf{r})$ and the sensor's output intensity $I_\delta(\mathbf{r})$ are shown in Figs. 8b and 8c. The residual aberration $\delta^{\mathrm{g}}(\mathbf{r})$ for $w = 1.2d$ is shown in Fig. 8d. From the $\delta^{\mathrm{g}}(\mathbf{r})$ and $I_\delta(\mathbf{r})$ patterns in Fig. 8 one can see that border lines of 2π phase cuts on the wave-front corrector always correspond to border lines of the receiver subapertures.

Strehl ratio adaptation curves for corrector friendly aberration compensation using the D-SPGD system with a wave-front corrector having different normalized width values w/d (mismatch parameter) are presented in Fig. 9. As expected, the most efficient phase distortion compensation occurs for $w \leqslant d$ (curves $w/d = 0.8$ and $w/d = 1.0$). In this case the cross talk between control parameter perturbations $\{\delta u_l\}$ and the corresponding perturbation of the metric vector components $\{\delta \bar{I}_l^{(n)}\}$ is small. With increase of w/d, cross talk increases and correspondingly the efficiency of D-SPGD control declines (compare curves $w/d = 1$ and $w/d = 1.4$ in Fig. 9). The best convergence rates occurred in the system with matched sensor/receiver subaperture sizes (curve $w/d = 1$). Note that the efficiency of the conventional SPGD algorithm also declines when the ratio w/d increases (compare the two SPGD curves in Fig. 9). The intensity patterns $I_\delta(\mathbf{r})$ shown Fig. 9 are quite different for different mismatch parameter values w/d. The border lines in the patterns $I_\delta(\mathbf{r})$ for $w/d = 1.2$ correspond to the 2π phase cut areas in Fig. 8d.

6. HIGH-RESOLUTION DECOUPLED STOCHASTIC PARALLEL GRADIENT DESCENT ADAPTIVE WAVE-FRONT CONTROL

A. Numerical Model

In the high-resolution (spatially distributed) D-SPGD adaptive system model both the piston-type pixelated wave-front corrector and the wave-front sensor had circular apertures of diameter $D = 0.85\,Na$, where a is the corrector element size and $N = 256$ is the numerical grid size. Input wave phase aberrations $\varphi(\mathbf{r})$ as well as wave-front perturbations $\delta u(\mathbf{r})$ were modeled by using the same realizations of a statistically homogeneous and isotropic random function as in the case described above for the low-resolution adaptive system: atmospheric-like phase screens with Andrews power spectrum for $\varphi(\mathbf{r})$ and random function with zero mean and Gaussian power spectrum for $\delta u(\mathbf{r})$.

In addition to the input beam with uniform intensity distribution we also considered an adaptive system with nonuniform intensity. In this case the input wave complex amplitude $A_{\mathrm{in}}(\mathbf{r}) = A_0(\mathbf{r})\exp[i\varphi(\mathbf{r})]$ had both random amplitude $A_0(\mathbf{r})$ and random phase $\varphi(\mathbf{r})$. To model input wave intensity scintillations the following technique was used: a numerical representation of a wave with complex amplitude $A(\mathbf{r}) = A_0 \exp[i\zeta(\mathbf{r})]$ having uniform intensity $I_0 = |A_0|^2$ and random phase $\zeta(\mathbf{r})$ was generated. Free-space propagation of this wave over a distance $z = L$ was modeled by using a fast Fourier transform routine. The complex amplitude modulus at the distance $z = L$ was used as an input wave amplitude: $A_0(\mathbf{r}) = |A_{z=L}(\mathbf{r})|$. For both random phases $\varphi(\mathbf{r})$ and $\zeta(\mathbf{r})$ the same phase screen realizations with Andrews power spectrum were used. This technique allows us to take into account changes in the intensity scintillations accompanying variations in the phase distortion amplitudes (for example, due to change of the Fried parameter). The strength of the input wave intensity fluctuations was controlled by changing the propagation distance L. Intensity scintillations were characterized by the normalized standard deviation of the intensity fluctuations aver-

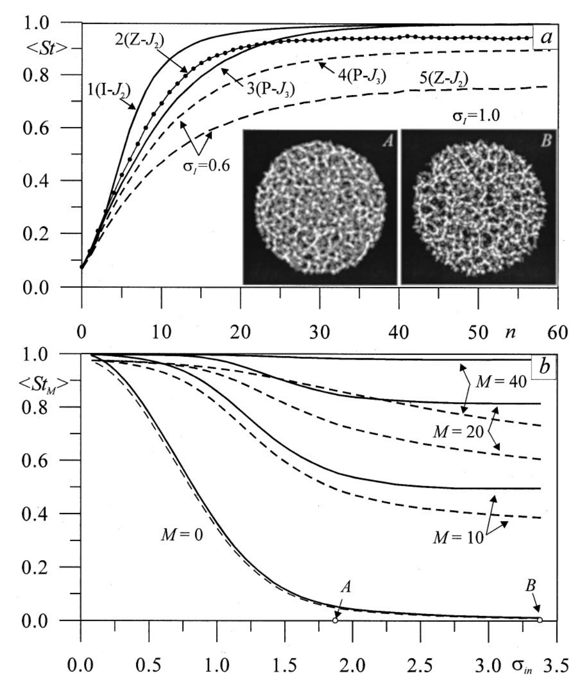


Fig. 10. High-resolution D-SPGD adaptive system adaptation process efficiency for uniform (solid curves) and random (dashed curves) input wave intensity distributions: (a) averaged Strehl ratio $\langle \mathrm{St} \rangle$ versus iteration number, (b) averaged Strehl ratio achieved after M iterations of the adaptation process $\langle \mathrm{St}_M \rangle$ versus input phase standard deviation σ_{in}. Adaptation curves 1–5 correspond to 1, interferometer with reference wave; 2 and 5, ZF; 3 and 4, PDI. Labels in parentheses are the same as in Fig. 5. The standard deviation for intensity scintillations in (a) was $\sigma_1 = 0.6$. The D-SPGD system in (b) corresponds to the J_3 controller with PDI. A and B correspond to gray-scale images of the input wave intensity patterns in (a) for point $\sigma_{\mathrm{in}} = 1.7$ (A) and (b) for point $\sigma_{\mathrm{in}} = 3.4$ (B).

aged over the aperture σ_I, defined as

$$\sigma_I^2 = \int[|A_0(\mathbf{r})|^2 - \bar{I}]^2 \mathrm{d}^2\mathbf{r}/\bar{I}^2,$$

where $\bar{I}$ is aperture-averaged input intensity. In Fig. 10 random patterns (A and B) of the input wave intensity modulations corresponding to phase screens characterized by different phase-fluctuation values σ_φ are shown (the propagation distance L was the same for both intensity patterns in Fig. 10).

B. Numerical Results

Adaptive phase distortion compensation efficiency was examined for adaptive system configurations with the following wave-front sensors: interferometer with reference wave, ZF, and PDI. Wave-front phase control was based on D-SPGD algorithms (15b) and (16b) aimed at minimization of the cumulative metrics (14a) and (14c). Recall that for the high-resolution systems metric, (14a) coincides with both (14b) and sharpness function (1). Results of the numerical simulations are presented in Fig. 10 for both uniform (solid curves) and random (dashed curves) intensity distributions. As seen from the Strehl ratio adaptation evolution curves in Fig. 10a, all considered high-resolution D-SPGD adaptive system types demonstrated exceptionally fast convergence. For the interferometric and Zernike sensors with J_2 controller the Strehl ratio increased up to the 80% level within the first $N_{80\%} \sim 10$–15 iterations. For the conventional SPGD control algorithm the corresponding convergence rate for the high-resolution system was on the order of $N_{80\%} \sim 10^3$–10^4 iterations.[14] In the case of the Zernike sensor the D-SPGD system did not completely remove the phase distortions: the achieved averaged Strehl ratio was near $\langle \mathrm{St} \rangle \approx 0.9$ (curve 2 in Fig. 10). In contrast, in the D-SPGD systems based on the interferometer and the PDI (J_3-controller), input wave phase aberrations were completely removed ($\langle \mathrm{St} \rangle \approx 0.99$) during the first 40 iterations.

The average Strehl ratio $\langle \mathrm{St}_M \rangle$ achieved during the first M iterations as a function of the input phase distortion standard deviation σ_{in} is shown in Fig. 10b for the D-SPGD system with PDI. For the given phase distortion range the Strehl ratio $\langle \mathrm{St} \rangle = 0.8$ was achieved after 20 iterations and $\langle \mathrm{St} \rangle = 0.98$ after 40 iterations.

The presence of intensity scintillations resulted in slower adaptation convergence. The compensation level achieved was also decreased: for the J_3 controller with PDI, $\langle \mathrm{St} \rangle$ was near 0.99 for a uniform intensity and $\langle \mathrm{St} \rangle \approx 0.87$ for intensity scintillations with $\sigma_I = 0.6$ (compare curves 3 and 4 in Fig. 10a). In the model considered here, input wave intensity scintillations decreased when σ_{in} was decreased: from $\sigma_I = 1.0$ for $\sigma_{\mathrm{in}} = 3.4$ to $\sigma_I = 0.6$ for $\sigma_{\mathrm{in}} = 1.7$, and further up to $\sigma_I = 0$ for $\sigma_{\mathrm{in}} = 0$ (see gray-scale images A and B in Fig. 10a). Despite the strong intensity scintillations, the adaptive system with D-SPGD controller was able to effectively suppress phase distortions over a wide range of input phase distortions (dashed curves in Fig. 10b).

C. Adaptive Perturbations

In all simulations considered above, a set of realizations of the random function $\delta u(\mathbf{r})$ with Gaussian power spectrum and fixed spatial correlation radius $l_p = 0.07D$ were used as wave-front phase perturbations (see the perturbations pattern in Fig. 4c). In fact, the correlation radius l_p is an additional parameter that can be used for further adaptation process convergence rate improvement. Figure 11 shows averaged adaptation evolution curves for the D-SPGD controller with PDI for two different values of the correlation radius l_p (curves 1 and 2). Decreasing l_p resulted in faster convergence (curve 2), but the adaptation level achieved was lower. This suggests that the phase perturbation spatial statistics should be adaptively changed during the adaptation process to match the continuing changes in the residual phase distortions. This adaptive perturbation technique was analyzed for the case of the conventional SPGD controller and resulted in noticeable convergence rate improvement.[14] The adaptive perturbation technique applied here resulted in nearly a 10–15% decrease in the convergence rate. A more substantial improvement in the convergence rate occurred with the use of "mixed" perturbations $\delta u(\mathbf{r})$ composed of a random function realization with Gaussian power spectrum and fixed spatial correlation radius $\delta u^{\mathrm{g}}(\mathbf{r})$ and a component proportional to the sensor output intensity distribution $I_\delta(\mathbf{r})$: $\delta u(\mathbf{r}) = \kappa \delta u^{\mathrm{g}}(\mathbf{r}) + (1-\kappa) I_\delta(\mathbf{r})$, where $0 < \kappa < 1$ is a weighting coefficient. As seen in Fig. 11 (curve 3) the use of mixed perturbations in the D-SPGD system with PDI resulted in adaptation process convergence during the first ten iterations. Owing to the dependence of $I_\delta(\mathbf{r})$ on the residual phase $\delta(\mathbf{r})$, the presence of the component $I_\delta(\mathbf{r})$ automatically provides the "right" spatial scale for the perturbations at the beginning of the adaptation process. As $\delta(\mathbf{r})$ decreases after the first few iterations, the relative influence of the random-perturbation component $\delta u^{\mathrm{g}}(\mathbf{r})$ increases, resulting in efficient removal of the remaining high-spatial-frequency components of $\delta(\mathbf{r})$. An example of mixed perturbation patterns $\delta u(\mathbf{r})$ at the beginning and at the end of the adaptation process are shown in Fig. 11.

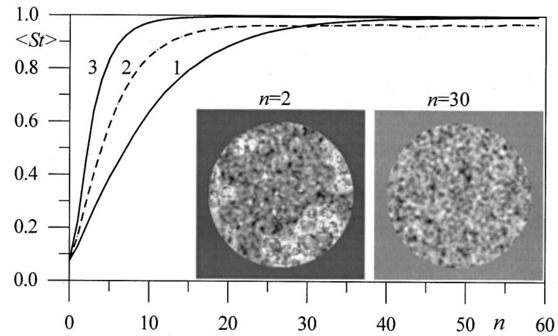


Fig. 11. Averaged Strehl ratio evolution curves for the D-SPGD controller with PDI, perturbations with Gaussian power spectrum, and different values of the correlation radius l_p (curve 1 and 2) and for "mixed" perturbations (curve 3): 1, $l_p = 0.07D$; 2, $l_p = 0.035D$; 3, $\kappa = 0.6$, $l_p = 0.035D$. Gray-scale images represent mixed perturbation patterns $\delta u(\mathbf{r})$ at $n = 2$ (left) and $n = 30$ (right).

7. CONCLUSION

The D-SPGD adaptive optics architectures presented here offer an attractive alternative to both wave-front phase conjugation and conventional gradient descent optimization based adaptive optics. The key advantages of the D-SPGD approach are as follows: (1) fast convergence practically independent of system spatial resolution, (2) the parallel nature of the control algorithm, (3) robustness with respect to input wave intensity scintillations, (4) and flexibility in choice of wave-front sensor. The important next step in the development of this technique is implementation of the D-SPGD controller as a high-resolution, parallel, low-power, VLSI microelectronic system. Recent successful demonstration of adaptive wave-front control using VLSI implementation of the conventional stochastic parallel gradient descent optimization technique[12,22] shows its technical viability for solving this challenging problem.

ACKNOWLEDGMENTS

The author thanks Gary Carhart for assisting with the computer code and J. C. Ricklin for technical and editorial comments. This work was performed at the U.S. Army Research Laboratory's Intelligent Optics Laboratory.

Address correspondence to M. A. Vorontsov at Army Research Laboratory, AMSRL-CI-C, Adelphi, Maryland 20783; e-mail, mvorontsov@arl.army.mil; phone, 301-394-0214; fax, 301-394-0225.

REFERENCES

1. *Adaptive Optics in Astronomy*, F. Roddier, ed. (Cambridge U. Press, New York, 1999), pp. 91–130.
2. R. K. Tyson, *Principles of Adaptive Optics* (Academic, Boston, 1991).
3. M. C. Roggemann and B. M. Welsh, *Imaging through Turbulence* (CRC Press, Boca Raton, Fla., 1996).
4. V. P. Sivokon and M. A. Vorontsov, "High-resolution adaptive phase distortion suppression based solely on intensity information," J. Opt. Soc. Am. A **15**, 234–247 (1998).
5. M. A. Vorontsov, "High-resolution adaptive phase distortion compensation using a diffractive-feedback system: experimental results," J. Opt. Soc. Am. A **16**, 2567–2573 (1999).
6. T. G. Bifano, J. Perrault, R. Krishnamoorthy Mali, and M. N. Horenstein, "Microelectromechanical deformable mirrors," IEEE J. Sel. Top. Quantum Electron. **5**, 83–89 (1999).
7. M. K. Lee, W. D. Cowan, B. M. Welsh, V. M. Bright, and M. C. Roggemann, "Aberration-correction results from a segmented microelectromechanical deformable mirror and refractive lenslet array," Opt. Lett. **23**, 645–647 (1998).
8. S. Serati, G. Sharp, R. Serati, D. McKnight, and J. Stookley, "128×128 analog liquid crystal spatial light modulator," in *Optical Pattern Recognition VI*, D. P. Casasent; and T. Chao, eds., Proc. SPIE **2490**, 378–387 (1995).
9. P. Madec, "Control Techniques," in *Adaptive Optics in Astronomy*, F. Roddier, ed. (Cambridge U. Press, New York, 1999), pp. 131–154.
10. J. W. Hardy, "Active optics: a new technology for the control of light," Proc. IEEE **66**, 651–697 (1978).
11. M. A. Vorontsov, G. W. Carhart, D. V. Pruidze, J. C. Ricklin, and D. G. Voelz, "Adaptive imaging system for phase-distorted extended source/multiple distance objects," Appl. Opt. **36**, 3319–3328 (1997).
12. M. A. Vorontsov, G. W. Carhart, M. Cohen, and G. Cauwenberghs, "Adaptive optics based on analog parallel stochastic optimization: analysis and experimental demonstration," J. Opt. Soc. Am. A **17**, 1440–1453 (2000).
13. T. R. O'Meara, "The multi-dither principle in adaptive optics," J. Opt. Soc. Am. **67**, 306–315 (1977).
14. M. A. Vorontsov and V. P. Sivokon, "Stochastic parallel-gradient-descent technique for high-resolution wave-front phase-distortion correction," J. Opt. Soc. Am. A **15**, 2745–2758 (1998).
15. J. C. Spall, "A stochastic approximation technique for generating maximum likelihood parameter estimates," in *Proceedings of the American Control Conference* (IEEE Press, Piscataway, N.J., 1987), pp. 1161–1167.
16. J. C. Spall, "Multivariate stochastic approximation using a simultaneous perturbation gradient approximation," IEEE Trans. Autom. Control **37**, 332–341 (1992).
17. G. Cauwenberghs, "A fast stochastic error-descent algorithm for supervised learning and optimization," in *Advances in Neural Information Processing Systems* (Morgan Kaufman, Los Altos, Calif., 1993), Vol. 5, pp. 244–251.
18. M. A. Vorontsov, G. W. Carhart, and J. C. Ricklin, "Adaptive phase-distortion correction based on parallel gradient-descent optimization," Opt. Lett. **22**, 907–909 (1997).
19. W. H. Southwell, "Wave-front estimation from slope measurements," J. Opt. Soc. Am. **70**, 998–1006 (1980).
20. B. L. Ellerbroek, C. Van Loan, N. P. Pitsianis, and R. J. Plemmons, "Optimizing closed-loop adaptive-optics performance with use of multiple control bandwidths," J. Opt. Soc. Am. A **11**, 2871–2886 (1994).
21. R. A. Muller and A. Buffington, "Real-time correction of atmospherically degraded telescope images through image sharpening," J. Opt. Soc. Am. **64**, 1200–1210 (1974).
22. R. T. Edwards, M. Cohen, G. Cauwenberghs, M. A. Vorontsov, and G. W. Carhart, "Analog VLSI parallel stochastic optimization for adaptive optics," in *Learning on Silicon* G. Cauwenberghs and Magdy A. Bayoumi, eds. (Kluwer Academic, Boston, Mass., 1999), Chap. 1, pp. 359–382.
23. F. Zernike, "How I discovered phase contrast," Science **121**, 345–349 (1955).
24. R. N. Smartt and W. H. Steel, "Theory and application of point-diffraction interferometers," Jpn. J. Appl. Phys. **14**, 351–356 (1975).
25. M. A. Vorontsov, E. W. Justh, and L. A. Beresnev, "Adaptive optics with advanced phase-contrast technique. I High-resolution wave-front sensing," J. Opt. Soc. Am. A **18**, 1289–1299 (2001).
26. L. C. Andrews, "An analytic model for the refractive index power spectrum and its application to optical scintillations in the atmosphere," J. Mod. Opt. **39**, 1849–1853 (1992).
27. D. L. Fried, "Statistics of a geometric representation of wave-front distortion," J. Opt. Soc. Am. **55**, 1427–1435 (1965).