ON AN ADAPTIVE CONTROL ALGORITHM FOR ADAPTIVE OPTICS APPLICATIONS

MOODY T. CHU*

Abstract. The wavefront aberrations induced by atmospheric turbulence can severely degrade the performance of an optical imaging system. Adaptive optics refers to the process of removing unwanted wave front distortions with the use of a phase corrector before the image is formed. The basic idea in adaptive optics is to control the position of the surface of a deformable mirror in such a way as to approximately cancel the atmospheric turbulence effects on the phase of the incoming light wave front. A critical component in the AO system is the inverse problem of phase computation that transforms the output of a wavefront sensor into a set of drive signals that control the shape of a deformable mirror. This paper addresses two issues pertaining to this inverse problem: 1) It embodies some of the basic principles for an adaptive optics system in simple mathematical expressions and demonstrates the fundamentals by highlighting one such adaptive control algorithm when temporary latency delay is present in the system; 2) It offers one theoretical framework for analyzing the effect of anisoplanatism that limits the performance of AO compensation when light from the wavefront sensor beacon and light from the object of interest sample different volumes of atmospheric turbulence.

Key words. optical image processing, adaptive control, adaptive optics, atmospheric turbulence compensation, deformable mirror, phase profile reconstruction, wavefront sensor, latency delay, anisoplanatism,

1. Introduction. It is known that diffraction effects impose a serious limitation on the resolving power of all optical instruments. In free space the limiting angular resolution θ_R, also known as the Rayleigh criterion, of an optical system is given by the simple yet precisely defined rule:

$$\theta_R = 1.22\frac{\lambda}{D}, \tag{1.1}$$

where D is the aperture diameter of the optical lens and λ is the wavelength. One therefore may think that using telescopes with larger and larger diameters will improve the image resolution. Unfortunately, the limiting angular resolution is not the only consideration in operating an optical system. The blurring of the image by atmospheric turbulence is also a problem, particularly with larger telescopes. Telescopes with larger aperture tend to capture disturbence easier in the imaging process. Indeed, in the absence of any correction to the turbulence it has been observed, both theortically and experimentally, that no design or optical quantity of a telescope can improve the degraded image [11, 19]. When observing visible light, for example, the resolution of an optical system is limited to approximately 1 arcsec, i.e., 5 μrad, regardless of the aperture diameter.

*Department of Mathematics, North Carolina State University, Raleigh, NC 27695-8205. This research was supported in part by the National Science Foundation under grant DMS-9803759.

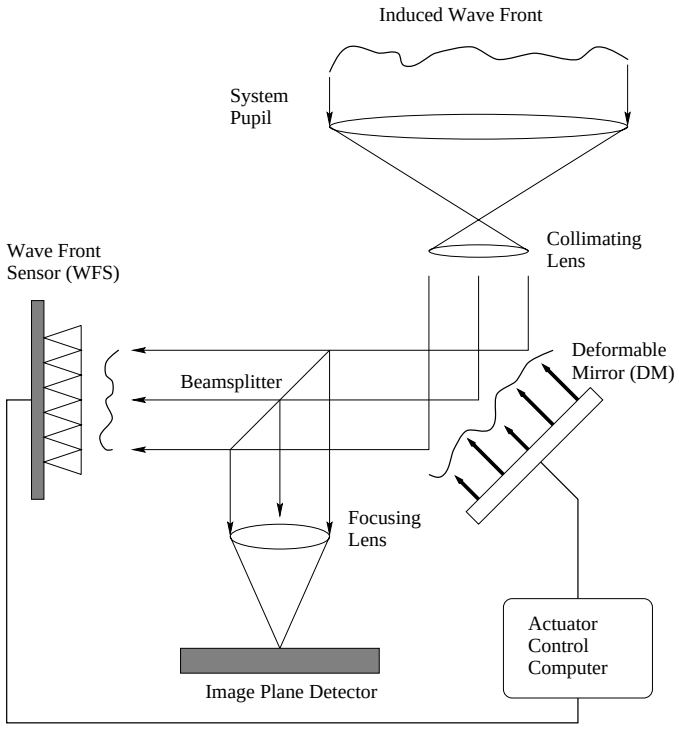

FIG. 1.1. *A simplified closed-loop AO system with main components.*

On the other hand, imaging through turbulence is an inevitable task because of its significant impacts on many important applications, including defense, engineering, and science [6, 20]. Various efforts have been studied to overcome the degradation of astronomical image quality caused by atmospheric turbulence. Among these, one interesting approach is to perform turbulence compensation using mechanical means, now generally referred to as *adaptive optics* (AO) [6, 11, 13, 14, 16, 20, 21, 22]. The basic idea of an AO system is as follows: Light in a narrow spectral band approaching the atmosphere from a distant light source, such as star, is usually modeled by a plane wave. When traveling through the atmosphere that does not have a uniform index of refraction, light waves are aberrated and no longer planar. In the closed-loop AO system depicted in Figure 1.1, this aberrated light is first reflected from the deformable mirror (DM). Some of this light is focused to form an image, and some is diverted to the wavefront sensor (WFS) that measures the wave front phase deformations. These WFS measurements are then fed to the actuator command computer that maps them into real time control commands for the DM. These control commands are used to adjust the DM actuators so as to compensate the wavefront distortions.

A closed-loop AO system contains three basic components — the deformable mirror, the wavefront sensor, and the actuator command computer. One of the major computational tasks in adpative optics is to somehow translate the atmospheric mea-

surements to the actuator controls of the deformable mirrors. The algorithm responsible for transforming the output from a WFS into a sequence of control signals that change the shape of a DM is referred to as a *reconstructor* and offen appears in the form of a matrix-vector multiplication. Different types of reconstructors and their performances can be found, for example, in [7, 14]. See also [5] for a simple, unified mathematical framework that describes the basic relationships among the main components of an AO system.

In practice, it is impossible for an AO system to fully restore the imaging performance. Many factors including inaccurate WFS measurement, physical DM constraints, temporal latency in response, or poor phrase reconstruction techniques could be attributed to the so called *anisoplanatic effect* that degrade the performance of an AO system [14, 20, 24, 25]. Anisoplanatism arises, for example, when the object of interest does not provide enough level of light for the WFS and a nearby beacon must be used to obtain the phase aberration information. Since the reference source used for WFS measurement is displaced from the object being observed, the turbulence sampled by the WFS is different from that in the desirable imaging path. Inaccurate wavefront sensing leads to inaccurate wavefront reconstruction that, in turn, degrades the AO system performance. In this paper, we want to extend the framework in [5] to include the anisoplanatism. Our objective is to set forth some basic principles in mathematical terms for the adpative control in adaptive optics application.

It should be pointed that, regardless of the anisoplanatic effect, the partially compensated images after the AO adjustment usually is followed by a second stage off-line post-processing step to improve the image quality. This inverse problem, usually ill-posed and large-scaled, is generally solved by deconvolution techniques [2, 6, 20]. Some interesting approaches include techniques from regularization, total variation, phase diversity, blind deconvolution, and so on. The study of post-processing image restoration needed for this second stage of reconstructing optical images is itself an area full of exciting research activities.

The paper is organized as follows: Following the notions described in [5], we introduce some background information concerning an AO system in section 2. These mathematical expressions essentially connect the various basic components of an AO system together and show the resulting dynamics. More technical details can be found in [14, 20]. The framework proposed herein can be generalized to include anisoplanatism. In section 3, we discuss ideas for updating the actuator control system by currently measured atmospheric information. Assuming that the measurement comes from a guided star which is not the original intended object, we study in section 4 the expected effect of these DM controls on the feedback WFS measuement and the residual phase error. A simple computer simulation is provided to illustrate the limiting effect of a 2-cyle delay adaptive control scheme. Finally, in section 5 we discuss how an additive anisoplanatism combined with an adaptive control algorithm affect the AO compensation.

The operation of an AO system is a very complicated procedure. In order to quickly manifest the basic principles, many sophisticated engineering details are rehashed and integrated in single mathematical expressions. It is hoped that our discussion in this paper, though with many simplified assumptions, offers a framework that is generic enough for helping to further the understanding of the AO system.

2. Basic Relationship. For convenience, we shall denote the turbulence-induced phase profile at position $\vec{x}$ in the telescope aperture plane, determined by the primary mirror, at time t by $\phi(\vec{x}, t)$. Likewise, the deformable mirror command issued at time t for the ith DM actuator is denoted by $a_i(t)$. The wavefront slope sensor measurement obtained from the kth subaperture of the WFS with *no* correction at time t is denoted by $s_k(t)$. The goal in positioning the DM surface via commands $a_i(t)$ is to represent an approximate conjugate of the turbulence-induced field $\phi(\vec{x}, t)$ so that the field reflected from the DM will have the aberration somewhat canceled and more closely approximate the field when no atmosphere turbulence is present. In this section we discuss the mathematical models representing these quantities.

2.1. Open-Loop Relationship. The mirror surface is controlled by a number of actuators that basically push and pull on the mirror surface to cause it to deform. Assuming that there are m actuators and that the actuators response linearly to the commands, the DM surface can be modeled by

$$\hat{\phi}(\vec{x}, t) = \sum_{i=1}^{m} a_i(t) r_i(\vec{x}), \tag{2.1}$$

where $r_i(\vec{x})$, called the influence function on the DM surface at position $\vec{x}$, denotes the response of the ith actuator to a unit adjustment. Suppose we sample the DM surface at n positions $\vec{x}_j$, $j = 1, \ldots n$, then the relationship between the surface position and the actuator command can be described as

$$\hat{\phi}(t) = H a(t). \tag{2.2}$$

In the above, the n dimensional vector $\hat{\phi}(t) = [\hat{\phi}(\vec{x}_1, t), \ldots, \hat{\phi}(\vec{x}_n, t)]^T$ represents the discrete corrected phase profile at time t. The $n \times m$ *DM configuration matrix* H, whose ith column is the vector $[r_i(\vec{x}_1), \ldots, r_i(\vec{x}_n)]^T$, is independent of time.

The wavefront sensors usually do not measure the wavefront phase $\phi(t)$ directly. Instead, the spatial gradient of $\phi(t)$, commonly referred to as the *wavefront slope*, is estimated. Without given specific details, we shall use the Hartman WFS (H-WFS) in this discussion. Readers are referred to [20] for more details on the physical configuration a H-WFS. In brief, the H-WFS spatially segments the incident wavefront with an array of ℓ small regions in the telescope pupil. Each array element, referred to as a *subaperture*, focuses a spot onto an array of detectors in the focal plane. The average

wavefront slope associated with the kth subaperture given by

$$\begin{aligned} s_k(t) &= \int d\vec{x} W_k(\vec{x}) \nabla \phi(\vec{x}, t) \\ &= -\int d\vec{x} \nabla W_k(\vec{x}) \phi(\vec{x}, t), \end{aligned} \tag{2.3}$$

where $W_k(\vec{x})$ is the kth subaperture weighting function, accounts for the H-WFS slope measurement. Upon approximating the integral in (2.3) by some quadrature rules at designated positions $\vec{x}_j$, $j = 1, \ldots n$, together with possible measurement noises (with mean zero), a linear relationship between wavefront phase and the H-WFS slops measurement can be described as

$$s(t) = W\phi(t) + \epsilon(t). \tag{2.4}$$

In the above, the turbulence-induced phase profile $\phi(\vec{x}, t)$ is spatially discretized at the DM surface positions $\vec{x}_j$, $j = 1, \ldots, n$ and is denoted by the vector $\phi(t) = [\phi(\vec{x}_1, t), \ldots, \phi(\vec{x}_n, t)]^T$ at time t. The entry w_{kj} in $W = [w_{kj}] \in R^{\ell \times n}$ denotes the jth quadrature weight for the integral (2.3) at abscissa $\vec{x}_j$. The quantity $\epsilon(t)$ accounts for any measurement error or noise.

Likewise, the corresponding H-WFS slope measurement of the corrected wave front phase $\hat{\phi}(t)$ can be measured as follows:

$$\hat{s}_k(t) = \sum_{i=1}^{m} \underbrace{\left(-\int d\vec{x} (\nabla W_k(\vec{x}) r_i(\vec{x}) \right)}_{G_{ki}} a_i(t).$$

Once again, upon discretization, we can write

$$\hat{s}(t) = Ga(t) \tag{2.5}$$

where the matrix $G = [G_{ki}] \in R^{\ell \times m}$ must satisfy the relationship

$$WH = G. \tag{2.6}$$

It should be noted that the DM actuators are *not* capable of producing the exact wavefront phase $\phi(\vec{x}, t)$ due to their finiteness of degrees of freedom. So $\hat{s} = Ga$ is never an exact measurement in practice.

Let $\mathcal{A}$ stand for the *vector space of actuator commands*, $\mathcal{S}$ stand for the *vector space of* H-WFS *slope measurements*, and Φ stand for the *vector space of phase profiles*. The mutual relationships among the three main components of an AO system are summarized in Figure 2.1 where the dotted-lines represent some kind of inverse transformations that will be discussed subsequently.

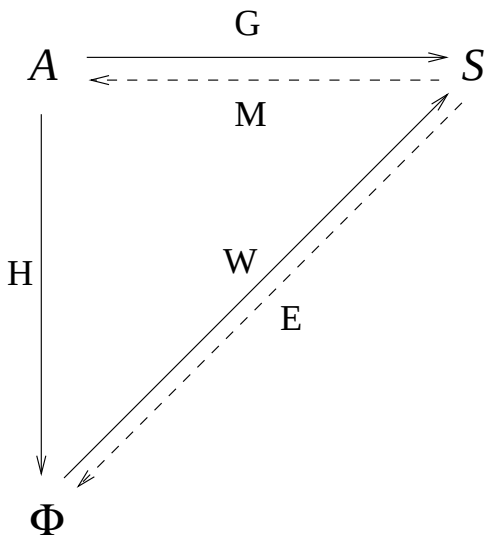

FIG. 2.1. *Diagram of mutual relationship*

2.2. Closed-Loop Relationship. In a closed-loop AO system such as the one demonstrated in Figure 1.1, the wavefront that arrives at either the H-WFS or the image plane detector is the one that has been reflected from the DM. Thus the information obtained at the image plane detector is actually the residual phase error

$$\Delta\phi(t) := \phi(t) - Ha(t). \tag{2.7}$$

That is, after the AO correction, $\Delta\phi(t)$ is the "observable" instantaneous wavefront distortion at time t. Likewise, the information available at the H-WFS is the feedback applied to $s(t)$ by DM actuator adjustment

$$\Delta s(t) := s(t) - Ga(t). \tag{2.8}$$

This is the observable H-WFS slope measurement at time t. Given the relationship (2.4), it is easy to see that an identical linear relationship

$$\Delta s(t) = W\Delta\phi(t) + \epsilon(t) \tag{2.9}$$

holds between the residual phase error $\Delta\phi(t)$ and and the feedback H-WFS slope measurement $\Delta s(t)$. Thus the diagram depicted in Figure 2.1 remains valid if the maps from $\mathcal{A}$ to Φ via H and from $\mathcal{A}$ to $\mathcal{S}$ via G are understood in a closed-loop AO system to mean $a \mapsto \phi - Ha$ and $a \mapsto s - Ga$, respectively.

When a beacon, instead of the intended object, is used to provide the light for the WFS measurement, what is really being reckoned in the WFS and is used to generate the DM controls is not the quantity $\Delta s(t)$ of the original object. Rather, it is the feedback measurement

$$\Delta s'(t) := s'(t) - Ga(t), \tag{2.10}$$

that is being observed from the WFS whereas $s'(t)$ is the wavefront slope corresponding to the phase deformation $\phi'(t)$ of the beacon. A linear relationship

$$\Delta s'(t) = W\Delta\phi'(t) + \epsilon'(t) \tag{2.11}$$

continues to hold in this case. However, note that optical paths from the beacon and the object traverse different regions of the atmosphere, resulting distinct wavefront perturbations. Thus if we attempt to estimate the turbulence-induced wavefront $\phi(t)$ via the information $\Delta s'(t)$ associated with the wavefront $\phi'(t)$ that is related to but not exactly the same as $\phi(t)$, the decreased correlation between $\phi(t)$ and $\phi'(t)$ will result in a degradation of the ability of the AO system to correct the object wavefront. We say that the AO system is suffering from the anisoplanatism effect.

The main idea in the development of the AO technology is that the actuator command $a(t)$ would drive the DM into a surface phase conjugate of the atmospheric wavefront distortion. How much turbulence can be compensated for during the cycle of computation depends on the actuator control used. In the presence of anisoplanatism, the best turbulence information one can hope to retrieve is only about $\phi'(t)$, not $\phi(t)$. Two questions therefore arise: The first, referred to as the *convergence issue*, concerns how the DM command $a(t)$ is determined and how it affects the measurement of $\phi'(t)$. The second, referred to as the *anisoplanatism issue*, concerns how the anisoplanatism between $\phi(t)$ and $\phi'(t)$ affects the AO performance. We shall address these two questions in the sequel. To measure the effectiveness of any scheme, we shall assume henceforth that the vector space Φ of all phase profiles is a Hilbert space where for any $f, g \in \Phi$ the inner product is defined to be

$$\langle f, g\rangle := f^T\Omega g, \tag{2.12}$$

with a specified symmetric and positive definite weighting matrix Ω.

3. Actuator Control. In this section, we outline a mathematical concept on how an actuator command should be issued for a closed-loop AO system. We first point out an ideal situation. From there we then introduce an inverse map from the space $\mathcal{S}$ of WFS measurements to the space $\mathcal{A}$ of actuator adjustments. Taking into account the temporal latency in response of an AO system, we assume a scenario of 2-cycle delay and propose an adaptive control scheme for actuator commands.

3.1. An Ideal Control. Recall that $\Delta\phi'$ represents the residual error after the correction by the current DM command, denoted by a_c. Assume that the phase profile ϕ' has been stationary. We expect that the new DM command, denoted by a_+, should help to reduce the residual error. Ideally, we would like to see that the equality

$$Ha_+ \stackrel{?}{=} \phi' \tag{3.1}$$

is satisfied. We adopt the notation "$\stackrel{?}{=}$" to indicate that the equality might not be materialized in reality. Define

$$\Delta a := a_+ - a_c. \tag{3.2}$$

Then the ideal control (3.1) is equivalent to

$$H\Delta a \stackrel{?}{=} \Delta\phi', \tag{3.3}$$

where $\Delta\phi'$ is defined in a similar way as (2.7). If we have perfect knowledge of $\Delta\phi'$, then the new command a_+ that solves (3.3) in the least squares sense is given by [12, 18]

$$a_+ = a_c + \underbrace{(H^T\Omega H)^{-1}H^T\Omega}_{H_\Omega^\dagger}\Delta\phi'. \tag{3.4}$$

Note that $H_\Omega^\dagger$ is the usual Moore-Penrose generalized inverse $H^\dagger$ of H if $\Omega = I$.

In practice, however, the answer is not as straightforward as indicated above. There are at least two reasons that make the actuator control a much harder task. One reason is that in most AO systems, the residual phase error $\Delta\phi'$ is not available directly. Such a quantity is either not observable from an image plane detector or that its relationship to the image being seen is highly nonlinear and cannot be inverted in real time. Often the residual error wavefront has to be "reconstructed" from the feedback measurement $\Delta s'$ that is available from the WFS. This leads to an inverse problem that we shall address in the next subsection. The other reason is that the atmospheric turbulence is random in nature. The quantities we are about to measure, including $\phi', s', \Delta\phi', \Delta s'$, and even the control command a or Δa, are finite samples of random variables from some unknown (and often time-variant) distribution functions. The inverse estimation therefore should be considered with stochastic viewpoint in mind.

3.2. An Inverse Problem. Based on (2.11), it is reasonable to estimate the residual phase error $\Delta\phi'$ by using a linear model such as

$$\Delta\hat{\phi}' = E\Delta s', \tag{3.5}$$

where E is some reconstructor matrix. Note that all quantities in (2.11) except for W are random variables. For the estimate (3.5) to remain unbiased, it is necessary to require

$$\mathcal{E}[\Delta\hat{\phi}'] = EW\mathcal{E}[\Delta\hat{\phi}'],$$

where $\mathcal{E}[x]$ denotes the expected value of a random variable x. This suggests an one-sided inverse relationship between E and W,

$$EW \stackrel{?}{=} I, \tag{3.6}$$

as is depicted in Figure 2.1. Obviously, W must satisfy a certain condition, such as that of full column rank, in order for the equality in (3.6) to hold. We shall not address

the technicalities in that regard in this paper. Several different ways to construct the matrix E have been proposed in the literature. See, for example, [5, 8, 14, 22]. Notably, the linear minimum-variance estimate of $\Delta\phi'$ based on (2.11) is of particular interest [5, 17, 18]. Replacing the right-hand side of (3.3) by (3.5), we are motivated to consider a direct link between Δa and $\Delta s'$ in the form

$$\Delta a = M\Delta s', \tag{3.7}$$

where M is a linear transformation (also called a reconstructor) from $\mathcal{S}$ to $\mathcal{A}$. Again, there are several ways to construct the matrix M [5, 8, 9, 22]. The choice of M affects the actuator command a and hence the behavior of the DM surface that, in turn, affects the performance of the AO system (even as compensation for the distortion ϕ' alone).

One observation is worth mentioning. We are reminded by the diagram in Figure 2.1 that M should be an inverse of G in some sense. The rationale for this notion goes as follows: Upon substitution, we see from (3.3) that

$$HM\Delta s' \stackrel{?}{=} E\Delta s'. \tag{3.8}$$

It thus suggests that

$$HM \stackrel{?}{=} E \tag{3.9}$$

and hence, by using (2.6) and (3.6), that

$$HMG \stackrel{?}{=} H, \tag{3.10}$$

or that

$$MG \stackrel{?}{=} I. \tag{3.11}$$

We stress that all equalities marked by $\stackrel{?}{=}$ above might never be materialized in reality, but an approximate relationship like (3.11) does suggest that M should be an inverse to G in some sense. Based on (3.3) and (3.5), one possible choice for M is the matrix given by

$$M = H_\Omega^\dagger E, \tag{3.12}$$

if E is available. Another conventional choice, based on (3.11), is

$$M = G_\Omega^\dagger, \tag{3.13}$$

if G is available. Some preliminary discussion on the effect of different choices of M can be found in [5, 8].

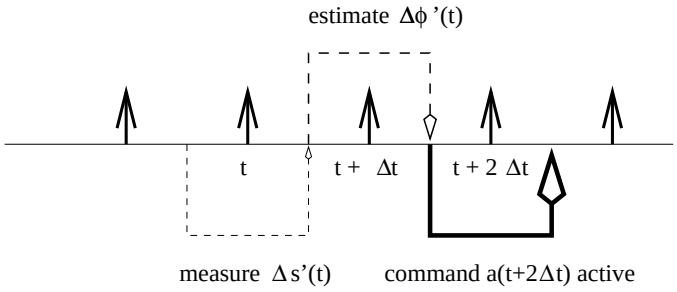

FIG. 3.1. *Time line for a 2-cycle delay AO system*

3.3. Temporary Latency. Even after the reconstructor M has been selected, there is one more concern in the actuator control. Thus far, we have been assuming that $\Delta s'$ is available at any moment it is needed and that a_+ can be issued instantaneously. In reality, there is always a latency delay due to the finite bandwidth of the control loop. The delay is inevitable, for example, due to the concern that the WFS information needs time to be read out as slowly as possible in order to minimize the detector read noise. To illustrate how an AO system operates under this circumstance, we shall now detail the case of a 2-cycle delay more carefully. The discussion can be generalized to other cases.

Let Δt denote the time between successive WFS measurements. This is also the time between successive adjustments to the DM actuator commands. The following is a possible sequence of events taking place in an AO system [10]:

1. The photons which measure $\Delta s'(t)$ are integrated on the WFS system over the interval $[t - 0.5\Delta t, t + 0.5\Delta t]$.
2. The calculation of estimating $\Delta\phi'(t)$ from $\Delta s'(t)$ begins as soon as the first pixels are digitized shortly after $t + 0.5\Delta$, but cannot be completed until the entire information has been read out just before $t + 1.5\Delta t$.
3. At that point the command $a(t+2\Delta t)$ is computed, sent to the DM, and remains in effect until before $t + 2.5\Delta t$.

The time line for these series of events is depicted in Figure 3.1. Since the reconstructor M in (3.7) is not precise, one common practice is to compute the new command $a(t + 2\Delta t)$ via the iterative process,

$$a(t + 2\Delta t) := \sum_{k=0}^{p} c_k a(t + (1 - k)\Delta t) + \sum_{j=0}^{q} b_j M_j \Delta s'(t - j\Delta t), \tag{3.14}$$

with the hope that some of the noise in the WFS measurements will be filtered out and that the stability of the control loop in the presence of latency and modeling errors will be improved. Note that the indices in (3.14) match the 2-cycle of latency discussed above. In particular, $a(t + 2\Delta t)$ is using information from and prior to $\Delta s'(t)$ only. The scheme (3.14) is analogous to the autoregression moving average (ARMA) filter commonly used in time series analysis.

4. Convergence. In this section, we address the convergence issue of the control scheme (3.14) for a closed-loop AO system. We shall analyze the convergence behavior of the sequence of the actuator commands and its effect.

4.1. Effect on Feedback WFS Measurement. To simplify the discussion, we shall assume that the phase profile $\phi'(t)$ and the WFS slope measurement $s'(t)$ stay stationary during the entire course of computation. For convenience of analysis, we shall rewrite (3.14) as

$$a^{(r+2)} = \sum_{k=0}^{p} c_k a^{(r+1-k)} + \sum_{j=0}^{q} b_j M_j (s' - G a^{(r-j)}), \quad r = 0, 1, 2, \ldots, \tag{4.1}$$

where $r = 0$ corresponds to the beginning time of the computation cycle. Note that realistically only the feedback information $\Delta s'$, not the open-loop information s', is available. Thus the scheme (4.1) is used only for analysis purpose. To be consistent with the ideal event (though most unlikely) where *no* turbulence or noise is present, the condition

$$\sum_{k=0}^{p} c_k = 1 \tag{4.2}$$

should be imposed on the scheme.

It is important to distinguish the two meanings that (4.1) represents. On one hand, the randomness of variables $a^{(r-j)}$, $j = -1, 0, \ldots \max\{p-1, q\}$, on the right-hand side of (4.1) implies that the new control command $a^{(r+2)}$ in (4.1) is also a random variable. On the other hand, each vector in (4.1) can also be considered as one *realization* of the corresponding random variable. Given starting values that are random samples of the variables $a^{(j)}$, $j = 1, 0, -1, \ldots, -\max\{p-1, q\}$, and a random but fixed sample of s', the sequence $\{a^{(r)}\}_{r=2}^{\infty}$ is completely determined by the finite difference equation (4.1). In this sense, the scheme becomes a traditional deterministic iteration. The resulting vector from the right-hand side of (4.1) represents a realization of the random variable $a^{(r+2)}$. If we repeat this experiment z times independently, then we will have collected z random samples for the variable $a^{(r+2)}$. This sampling procedure can still be nicely summarized by using (4.1) in which $a^{(r+j)}$, $r \geq 0$, $j = 2, 1, 0, -1, \ldots, -\max\{p-1, q\}$, is interpreted as an $m \times z$ matrix whose columns represent z independent samples of the variable $a^{(r+j)}$ whereas s' is interpreted as an $\ell \times z$ matrix whose columns represent z independent samples of the variable s'. Obviously, the stochastic independence among columns within the same matrix is necessary to assure that we do have random samples. If we can prove that sequence of matrices $\{a^{(r)}\}$ from the matrix finite difference equation (4.1) converge to a fixed matrix a^* in a matrix norm, then columns of the resulting matrix a^* represent z independent samples of a certain random variable a^* that is the *almost sure* [1] limit of the sequence of random variables $\{a^{(r)}\}$.

[1] Suppose $\{X_n\}$ is a sequence of measurable functions (random variables) on a probability space

Before we discuss the almost sure convergence of scheme (4.1) in more details, it might be instructive to first examine the *expected* effect of $a^{(n)}$ on the AO system. We make the following observation.

THEOREM 4.1. *Suppose that the expected value of wavefront slope measurement $s'(t)$ is independent of time t throughout the cycle of computation. Suppose also that the $m \times \ell$ constant matrix $\sum_{j=0}^{q} b_j M_j$ is of full column rank. Let $\{a^{(n)}\}$ denote the sequence of actuator commands generated by scheme (3.14). Then the expected effect of these commands is that the WFS feedback measurement $\Delta s_n^{'}$ (Recall that $\Delta s_n^{'} := s' - Ga^{(n)}$ by (2.8)) is eventually nullified by the actuators. That is,*

$$\mathcal{E}[s'] = G \lim_{n\to\infty} \mathcal{E}[a^{(n)}], \tag{4.3}$$

if the limit on the right-hand side of (4.3) exists. In this case, the expected residual phase error $\Delta\phi_n^{'}$ is inversely related to the expected WFS measurement noise ϵ' via the equation

$$0 = W \lim_{n\to\infty} \mathcal{E}[\Delta\phi_n^{'}] + \mathcal{E}[\epsilon']. \tag{4.4}$$

Proof. Upon taking expected values on both sides of equation (3.14) and assuming the limit exists, we obtain the fixed-point equation

$$\lim_{n\to\infty} \mathcal{E}[a^{(n)}] = \sum_{k=0}^{p} c_k \lim_{n\to\infty} \mathcal{E}[a^{(n)}] + \sum_{j=0}^{q} b_j M_j (\mathcal{E}[s'] - G \lim_{n\to\infty} \mathcal{E}[a^{(n)}]). \tag{4.5}$$

Equation (4.3) follows from the consistency condition (4.2). Equation (4.4) follows from (2.11). □

It is interesting to compare the desired effect mentioned in (3.1) with the limiting effect attained by the delay control scheme (3.14). In the former the command a_+ is chosen so that $\langle \Delta\phi', \Delta\phi' \rangle$ is nullified or at least minimized if the equality in (3.1) is not realizable at all. In the latter the commands has the effect that the average of $\Delta\phi_n^{'}$ eventually satisfies the equation (4.4). Note, however, that even if $\mathcal{E}[\epsilon'] = 0$, the relationship (4.4) implies at most that $\lim_{n\to\infty} \mathcal{E}[\Delta\phi_n^{'}]$ lies in the null space of W. Furthermore, this inverse relationship through W in (4.4) does not necessarily guarantee small variance $\mathcal{E}[\| \lim_{n\to\infty} \Delta\phi_n^{'} \|^2]$ in any means. Numerical simulation in the next subsection clearly illustrates this important difference.

It is also important to point out that the hypothesis that ϕ' (and hence s') stays stationary is not realistic. The atmospheric turbulence changes rapidly, if not continually. We can at most assume a stationary statistics of $\phi'(t)$ for a short period of time. It is therefore desirable, other than the mere convergence, that the convergence should shape up as quick as possible.

$(\mathcal{S}, B, P)$ where $\mathcal{S}$ is a measure space with a σ-algebra B on its subsets and a measure P with $P(\mathcal{S}) = 1$. We say X_n converges almost surely to X if $X_n \longrightarrow X$ almost everywhere, that is, if $P(\{\omega \in B \mid \lim_{n\to\infty} \|X_n(\omega) - X(\omega)\| = 0\}) = 1$.

We now turn to the issue of convergence for the scheme (4.1). For simplicity, we shall assume henceforth $p = q + 1$. The discussion can be extended to other cases by a similar argument. Without causing ambiguity, we shall use the same notation $a^{(r+j)}$ to denote any random sample (a column vector) of the the random variable $a^{(r+j)}$ in (4.1). It suffices to consider the finite difference equation (4.1) for one generic sequence of vectors $\{a^{(r+j)}\}$, although z samples are to be taken independently.

Define

$$\mathbf{a}_{r+2} := [a^{(r+2)}, a^{(r+1)}, \ldots a^{(r-q+1)}]^T, \quad r = 0, 1, \ldots \tag{4.6}$$

$$\mathbf{b}(s') := [\sum_{j=0}^{q} b_j M_j G s', 0, \ldots, 0]^T. \tag{4.7}$$

Then the scheme (4.1) can be written as

$$\mathbf{a}_{r+2} = A\mathbf{a}_{r+1} + \mathbf{b}(s') \tag{4.8}$$

where A is the $m(q+2) \times m(q+2)$ matrix

$$A := \begin{bmatrix} c_0 I_m & c_1 I_m - b_0 M_1 G & \ldots & c_{q+1} I_m - b_q M_q G \\ I_m & 0 & & 0 \\ 0 & I_m & & \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots \; I_m & 0 \end{bmatrix} \tag{4.9}$$

with I_m denoting the $m \times m$ identity matrix. It is now clear that the following theorem concerning the convergence of scheme (4.8) follows from the classical results for stationary iteration analysis [12, 23].

THEOREM 4.2. *The sequence of vectors $\{a^{(r+2)}\}_{r=0}^{\infty}$ computed from (4.1) with any starting values $a^{(1)}, a^{(0)}, \ldots a^{(-q)}$ converges to a limit point if and only the spectral radius $\rho(A)$ of the matrix A is less than one.*

We note that the limit point $a^* = a^*(s')$ depends upon the initial random yet fixed sample vector s' of the WFS slope measurement, but not on the starting values $a^{(1)}, a^{(0)}, \ldots a^{(-q)}$. A conventional notion to measure how fast the scheme (4.1) converges (already known to converge linearly) is via the *asymptotic convergence factor* α defined by

$$\alpha := \sup_{a_1, a_0, \ldots, a_{-q}} \inf_k \sup_{r \geq k} \|a^{(r)} - a^*\|^{1/r}. \tag{4.10}$$

Again, the following is a classical result [23].

THEOREM 4.3. *Suppose $\rho(A) < 1$. Then the iterative scheme (4.1) has asymptotic convergence factor $\alpha = \rho(A)$.*

It is desired that the convergence will happen as quick as possible. The coefficients in the scheme (4.1) therefore should be selected so that the spectral radius $\rho(A)$ is as small as possible.

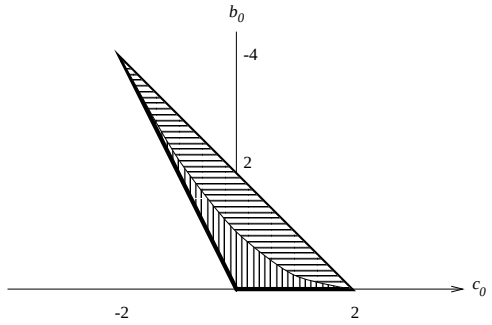

FIG. 4.1. *Diagram of mutual relationship*

The convergence analysis for (4.1) would be even simpler if we could assume from (3.11) that $M_jG = I_m$ for all j. In this case, the finite difference equation is reduced to scale coefficients and the eigenvalues of A are the same as the roots of the associated *characteristic polynomial* $p(r)$,

$$p(r) := r^{q+2} - c_0 r^{q+1} - \sum_{j=0}^{q} (c_{j+1} - b_j) r^{q-j}. \tag{4.11}$$

We conclude this section by an illustration on how the selection of coefficients affects the scheme. Considering the following simple example where $p = 1$ and $q = 0$. The control scheme becomes

$$a^{(r+2)} = c_0 a^{(r+1)} + c_1 a^{(r)} + b_0 M_0 \Delta s_r. \tag{4.12}$$

Assuming $M_0G = I_m$, the characteristic polynomial is given by

$$p(r) = r^2 - c_0 r - (1 - c_0 - b_0), \tag{4.13}$$

where we have assumed $c_0 + c_1 = 1$. The region of (c_0, b_0) for which both roots of (4.13) have modulus no great than one is plotted in Figure 4.1. In the horizontally shaded region, the roots are complex conjugate numbers. In the vertically shaded region, the roots are real numbers. It is easy to conclude that the choice $c_0 = 0$, $c_1 = 1$, and $b_0 = 1$ gives the smallest modulus of roots, i.e., zero. However, it should be noted that with this choice of coefficients we will reach the steady-state control $a^{(r+2)} \equiv M_0 s'$ for all r in one iteration. This control can be interpreted as being consistent with the ideal control (3.1) in the following sense: Pre-multiplying both side of the ideal control (3.1)

by W and using (2.4) and (2.6), we find that the control a_+ should approximately satisfy

$$Ga_+ \stackrel{?}{=} s', \tag{4.14}$$

where the equality might not hold due to the presence of noise ϵ'. If it really happens as we have assumed that $M_0 G = I_m$, then the ideal control should be given by

$$a_+ \stackrel{?}{=} M_0 s'$$

as we have expected.

In practice, G may not have a left inverse at all, e.g., when $\ell << m$. The assumption $M_J G = I_m$ may not be true. However, as long as $\rho(A) < 1$, the 2-cycle delay scheme (3.14) converges.

THEOREM 4.4. *Suppose that coefficients c_k, b_j, and reconstructor matrices M_j are selected so that $\rho(A) < 1$. Suppose also that the matrix $\sum_{j=0}^{q} b_j M_j$ is of full column rank. Then the 2-cycle delay iterative scheme (3.14) converges almost surely to a limit point a_+ that satisfies the equality in (4.14).*

It is interesting to compare the almost sure result (4.14) with the average result (4.3). Once we have the theory of convergence for (3.14) established as above, we may repeatedly use (3.14) to generate z samples with independent starting values. The limit point of each sampling process represents a realization of the limiting distribution of the sequence of random variables $\{a^{(r+2)}\}$. With the sample size z large enough, we should be able to grasp a good understanding of the limiting distribution of the actuator commands.

4.2. Effect on Residual Phase Error. The following result follows from Theorem 4.4. It describes the effect of the adaptive control algorithm (3.14) on the residual phase error.

COROLLARY 4.5. *If the scheme (3.14) converges almost surely, then the residual phase error (of the beacon) will eventually satisfies the equality*

$$W\Delta\phi' + \epsilon' = 0 \tag{4.15}$$

as the limiting actuator command.

We should point out that (4.15) alone is not sufficient to determine the residual phase error $\Delta\phi'$ if W has a nontrivial null space $\mathcal{N}(W)$. Suppose that W is of rank ρ. Let

$$W = U\Sigma V^T \tag{4.16}$$

be the reduced singular value decomposition of W so that columns of $V \in R^{n\times\rho}$ represent an orthonormal basis for the space $\mathcal{N}(W)^\perp$ and $\Sigma \in R^{\rho\times\rho}$ is nonsingular. Then (4.15) is reduced to

$$VV^T\Delta\phi' + W^\dagger\epsilon' = 0. \tag{4.17}$$

surface positions n	=	5
number of actuators m	=	4
number of subapertures ℓ	=	3
size of random samples z	=	2500
H	=	$rand(n, m)$
W	=	$rand(\ell, n)$
G	=	WH
$L_{\phi'}$	=	$rand(n, n)$
$L_{\epsilon'}$	=	$diag(rand(\ell, 1))$
$\mu_{\phi'}$	=	$zeros(n, 1)$
$\mu_{\epsilon'}$	=	$zeros(\ell, 1)$

TABLE 4.1
Parameters used in simulation

Note that the quantity $V^T\Delta\phi'$ in (4.17) stands for the component of the project of $\Delta\phi'$ onto the subspace $\mathcal{N}(W)^\perp$ and is uniquely determined. It follows that at the limiting distribution of actuator command, the covariance matrix of the projected $\Delta\phi'$ is completely determined by that of the WFS measurement noise ϵ', i.e., we have the relationship

$$\text{cov}(VV^T\Delta\phi') = \text{cov}(W^\dagger\epsilon') \tag{4.18}$$

between the two covariance matrices.

To demonstrate the effect of control scheme (3.14) on $\Delta\phi'$, we consider a numerical simulation based on the parameters listed in Table 4.1. For convenience, we use the MATLAB syntax $ones(n, m)$ and $rand(n, m)$ to denote, respectively, $n \times m$ matrices with all entries ones and random entries chosen from a uniform distribution on the interval $(0, 1)$. Likewise, let $randn$ denote a normally distributed random variable with mean zero and variance one. Define

$$\phi' = \mu_{\phi'} * ones(1, z) + L_{\phi'} * randn(n, z),$$

where $*$ denotes standard matrix to matrix multiplication. Then we obtain z random samples for the multivariate normally distributed wavefront phase profile ϕ' with mean $\mu_{\phi'}$ and covariance matrix $V_{\phi'} = L_{\phi'}L_{\phi'}^T$. In a similar way, the WFS measurement noises are simulated by

$$\epsilon' = \mu_{\epsilon'} * ones(1, z) + L_{\epsilon'} * randn(\ell, z).$$

We perform a simulation using the scheme (4.12) with constructor matrix $M_0 = H^\dagger W^\dagger$ and $b_0 = 0.5$. Two initial actuator commands are randomly given. Figure 4.2 represents a typical simulation results. The left two graphs in Figure 4.2 show the dynamical behavior of means and variances of $\Delta\phi'$ throughout 20 iterations. In contrast, the right two drawings in Figure 4.2 show the corresponding dynamics when $\Delta\phi'$ is projected to

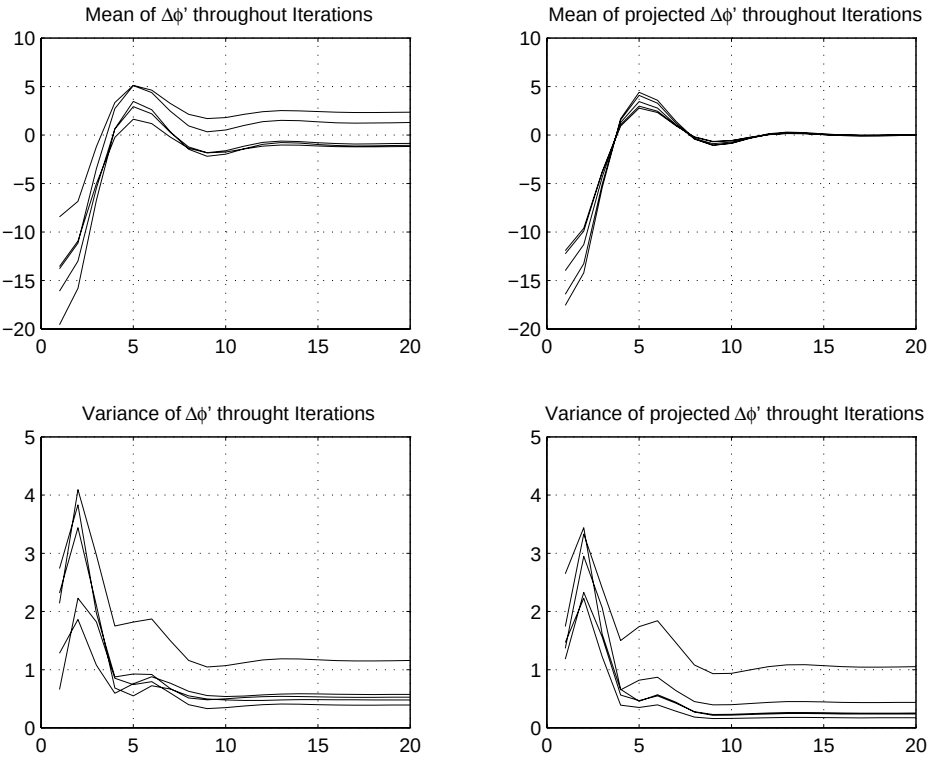

FIG. 4.2. *Convergence of $\Delta\phi'$ with 2-cycle delay control (4.12)*

$\mathcal{N}(W)$. Assuming that $\mathcal{E}[\epsilon'] = 0$, we see that the projection component, i.e., $V^T\mathcal{E}[\Delta\phi']$ becomes zero. On the other hand, we see that the variance of $V^T\Delta\phi'$ is generally smaller than that of $\Delta\phi'$. Both variances are well contained throughout the iterations and reduced to some constants that depend on $L_{\phi'}$ and $L_{\epsilon'}$.

5. Effect of Anisoplanatism. In the preceding section, we have set forth some preliminary consideration on how the DM commands, whether with or without latency, should be determined. In particular, we have studied the effect of a 2-cycle delay control scheme (3.14). That scheme utilizes only discrete information from closed-loop WFS slope measurement $\Delta s'$. Under some mild assumptions we have shown that the scheme will converge almost surely to a DM command that has the limiting effect of nullifying the WFS slope measurement. In this section we want to study how this limiting DM command affects the turbulence-induced phase profile measurement of the original object.

Recall that the optical path taken by the beacon often is differs from that taken by the object that is of interest, causing anisoplanatism effect. Anisoplanatism can arise under many different circumstances. For example, in an AO system for an optical interferometer two beams are displaced and parallel to each other, producing *displace-*

ment anisoplanatism [24]. In an AO system using laser guide star, two beams might propagate at slightly different angles with respect to each other, producing *angular* anisoplanatism or *focal* anisoplanatism. The beams might also be of slightly different wavelengths that experience different amounts of turbulence-induced aberration, introducing *chromatic* anisoplanatism [22]. Also, the wind moving through the atmosphere might induce a time delay between the propagation of two beams, causing *temporal* anisoplanatism. Because of the variety of causes, it is difficult to quantify the anisoplanatism effect in a single and general mathematical term. Still, we shall consider anisoplanatism collectively as a coupling of spatial and temporal effects. To simplify the matter further, we recall the Taylor frozen flow hypothesis[2] and shall assume an additive relationship

$$\phi(t) = \phi'(t) + \delta\phi(t) \tag{5.1}$$

between phase profiles $\phi(t)$ and $\phi'(t)$ for some variable $\delta\phi(t)$. The exact determination of $\delta\phi(t)$, as indicated above, is a very complicated issue which we could not address in this discussion. Under a very loose term we might at least think of the right-hand side of (5.1) as a linear approximation to $\phi(t)$ based on the Taylor hypothesis. Thus (5.1) might be a reasonable assumption.

Our goal is to describe $\Delta\phi$ in terms of $\delta\phi$ under the actuator commands determined by control scheme (3.14). It is clear due to the linearity of the relationship (2.7) that the additivity is passed onto the residual phaser errors, i.e.,

$$\Delta\phi(t) = \Delta\phi'(t) + \delta\phi(t). \tag{5.2}$$

Thus the best possible residual phase error $\Delta\phi(t)$ in the presence of anisoplanatism would inherit the original amount of anisoplanatic error $\delta\phi(t)$ inside. To see the effect of the adaptive controls, we premultiply (5.2) by W and substitute in (4.15) to obtain

$$W\Delta\phi = W\delta\phi - \epsilon'. \tag{5.3}$$

Assuming that the anisoplanatic error $\delta\phi$ and the WFS measurement error ϵ' are stochastically independent, we see that

$$\mathcal{E}[VV^T\Delta\phi] = \mathcal{E}[VV^T\delta\phi] - \mathcal{E}[W^\dagger\epsilon'] \tag{5.4}$$

$$\mathrm{cov}(VV^T\Delta\phi) = \mathrm{cov}(VV^T\delta\phi) + \mathrm{cov}(W^\dagger\epsilon'). \tag{5.5}$$

Recall that $VV^T\Delta\phi$ represents the projection $\Delta\phi$ onto the subspace $\mathcal{N}(W)^\perp$. For Gaussian distribution, the above two moments complete determine the nature of the projection of $\Delta\phi$ in $\mathcal{N}(N)^\perp$. It is interesting to compare the (4.18) with (5.5). The combined effect of anisoplanatism and the actuator command scheme (3.14) is now clear.

[2]Roughly speaking, Taylor hypothesizes that measurements at different points in time correspond to measurements taken along different paths through the turbulence [20].

If $\phi(t)$ and $\phi'(t)$ are related in some ways other than additivity (5.1), the framework of discussion outlined above can be modified properly to study the corresponding new effect.

6. Concluding Remarks. The randomness and time evolution of the atmospheric inhomogeneities make imaging through turbulence a difficult and challenging problem. Adaptive optics techniques afford a mechanical means of sensing and correcting for turbulence effects as they occur. From a simple mathematical framework connecting the major components of an AO system, we set forth some basic concepts of adaptive control for adaptive optics. Hoping to manifest the essential principles behind the complicated AO mechanism, we tried to characterize, sometimes under much simplified assumptions, the operations of an AO system in terms of mathematical expressions. We illustrated an analysis technique for a particular 2-cycle delay control scheme and studied its effect on the AO performance with and without the presence of the anisoplanatism.

Acknowledgments. The author would like to thank Dr. Brent Ellerbroek and Professor Robert Plemmons for providing motivation, encouragement, and very helpful advice for this study.

REFERENCES

[1] H. Andrews and B. Hunt, Digital Image Restoration, Prentice-Hall, Inc., New Jersey, 1977.
[2] M. Banham and A. Katsaggelos, *Digital Image Restoration*, IEEE Signal Processing Magazine, March 1997, pp. 24–41.
[3] T. Bell, *Electronics and the Stars*, IEEE Spectrum 32, no. 8 (1995), pp. 16–24 .
[4] K. Castleman, Digital Image Processing, Prentice-Hall, New Jersey, 1996.
[5] M. T. Chu, V. P. Pauca, R. J. Plemmons, and X. Sun, A mathematical framework for the linear reconstructor problem in adaptive optics, Linear Alg. Appl., to appear.
[6] P. Davis, Mathematical tools power the quest for perfect images, SIAM News, 30(1997).
[7] B. L. Ellerbroek, First-order performance evaluation of adaptive-optics systems for atmospheric-turbulence compensation in extended-field-of-view astronomical telescopes, J. Opt. Soc. Am. A., 11(1994), 783-805.
[8] 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(1994), 2871-2886.
[9] B. L Ellerbroek and T. A. Rhoadarmer, Real-time adaptive optimization of wave-front reconstruction algorithms for closed-loop adaptive-optical systems, in Proceedings of SPIE Conference on Adaptive Optical System Technologies, 3355(1998), 1174-1185.
[10] B. L. Ellerbroek, private communication, 1998.
[11] D. L. Fried, Optical resolution through a randomly inhomogeneously medium for very long and very short exposures", J. Opt. Sco. Am., 56(1966), 1372-1379.
[12] G. H. Golub and C. F. Van Loan, Matrix Computations, 3rd ed., Johns Hopkins University Press, Baltimore, MD, 1996.
[13] J. W. Hardy, Active optics - a progress review, in Proc. SPIE on Active and Adaptive Optical Systems, 1542(1991), 2-17.

[14] J. W. Hardy, Adaptive Optics for Astronomical Telescopes, Oxford Press, New York, 1998.
[15] S. S. Haykin, Adaptive Filter Theory, 3rd ed., Prentice-Hall, Inc., New Jersey, 1996.
[16] A. K. Katsaggelos, ed., Digital Image Restoration, Springer-Verlag, New York, 1991.
[17] D. G. Luenberger, Optimization by Vector Space Methods, John Wiley & Sons, New York, 1968.
[18] J. L. Melsa and D. L. Cohn, Decision and Estimation Theory, McGraw-Hill, New York, 1978.
[19] J. Nelson, *Reinventing the Telescope*, Popular Science, 85 (1995), pp. 57–59.
[20] M. C. Roggemann and B. Welsh, Imaging through Turbulence, CRC Press, New York, 1996.
[21] R. K. Tyson, Adaptive optics system performance approximations for atmospheric turbulence correction, Optical Engi., 29(1990), 1165-1173.
[22] R. K. Tyson, Principles of Adaptive Optics, 2nd Edition, Academic Press, MA, 1998.
[23] R. S. Varga, Matrix Iterative Analysis, Prentice-Hall, Englewood Cliffs, NJ, 1962.
[24] M. R. Whiteley, Optimal Atmospheric Compensation for Anisoplanatism in Adaptive-Optical Systems, Ph.D. Dissertation, Air Force Institute of Technology, 1998.
[25] R. W. Wilson and C. R. Jenkins, Adaptive optics for astronomy: theoretical performance and limitations, Monthly Notice the Royal Astronomical Society, 278(1996), 39-61.