

Toward Strehl-Optimizing Adaptive Optics Controllers

D. Gavel, D. Wiberg

This article was submitted to Society of Photo-Optical
Instrumentation Engineers

October 8, 2002

U.S. Department of Energy

Lawrence
Livermore
National
Laboratory

DISCLAIMER

This document was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor the University of California nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or the University of California, and shall not be used for advertising or product endorsement purposes.

This is a preprint of a paper intended for publication in a journal or proceedings. Since changes may be made before publication, this preprint is made available with the understanding that it will not be cited or reproduced without the permission of the author.

Toward Strehl-Optimizing Adaptive Optics Controllers

Donald Gavel^{*a}, Donald Wiberg^{†b}

^aLawrence Livermore National Laboratory; ^bUniversity of California, Santa Cruz

ABSTRACT

A main objective of adaptive optics is to maximize closed-loop Strehl, or, equivalently, minimize the statistical mean-square wavefront residual. Most currently implemented AO wavefront reconstructors and closed-loop control laws do not take into account either the correlation of the Kolmogorov wavefronts over time or the modified statistics of the residual wavefront in closed loop. There have been a number of attempts in the past to generate “predictive” controllers, which utilize wind speed and Cn2 profiles and incorporate one or two previous time steps. We present here a general framework for a dynamic controller/reconstructor design where the goal is to maximize mean closed-loop Strehl ratio over time using all previous data and exploiting the spatial-temporal statistics of the Kolmogorov turbulence and measurement noise.

Keywords: Adaptive optics, optimal control, predictive control

1. INTRODUCTION

We assume that the objective of an adaptive optics system is to maximize the Strehl ratio, or on-axis intensity of a point source relative to that of the diffraction-limit. Meeting this criterion will satisfy many of the stated objectives for diffraction-limited astronomical instruments. For example, maximizing Strehl also maximizes image resolution, detectability of a dim object next to a bright one, contrast of structure in an extended object, and spectrograph spatial and wavelength resolution. Strehl ratio is intimately related to the mean square residual wavefront error, σ_ϵ^2 via the Marechal approximation

$$S = e^{-\sigma_\epsilon^2} \quad (1)$$

Mean square wavefront error is a function of both the atmosphere and the parameters of the AO system. Since the atmosphere is a stochastic process, so too is the wavefront error, thus we speak in terms of its average, or statistical expected value. One ought to be able to minimize the expected residual variance through optimal processing of the data coming from the AO wavefront sensor. Such an optimal processing algorithm would take into account prior knowledge of the statistics of the wavefront, i.e. the Kolmogorov turbulence spectrum, and also take into account the amplitude of wavefront sensor noise.

Optimal statistical processing of AO data in a static wavefront has been approached in the past¹⁻⁴. Accounting for the effects of closed-loop operation³⁻⁴ is somewhat more difficult, since the correlation statistics of the residual wavefront are modified from the uncorrected Kolmogorov statistics by the action of the feedback control loop. Optimum processing has obvious advantages; in addition to maximizing Strehl, it also makes “optimum use of photons” in that it is the absolute best that the AO controller can accomplish given the finite flux of light from the guide star. Thus the greatest benefit of optimal wavefront control may be in low signal-to-noise conditions.

Optimal processing also has its price. One must incorporate a considerable amount of prior information into the algorithm. We will assume that wavefront statistics are Kolmogorov with a known transverse coherence length, r_0 , and outer scale L_0 . Also we will assume the Taylor frozen-flow hypothesis and that we know the wind velocity vector, \mathbf{v} . In a real implementation we might expect to have separate instruments to measure these parameters or, alternatively, use telemetry data from the AO system to provide estimates of these quantities or at least the correlation functions needed by the algorithm which depend on these parameters. We will use simulations to explore the sensitivity of optimal solutions to deviation from the assumed parametric values.

* Lawrence Livermore National Laboratory, L-290, P.O. Box 808, Livermore CA

† Department of Electrical Engineering and Center for Adaptive Optics, UC Santa Cruz, 1156 High St, Santa Cruz CA

2. WAVEFRONT STATISTICS

The phase, $\varphi(x)$, of a Kolmogorov wavefront is Gaussian distributed with point to point variations described by a structure function

$$D_\varphi(x-x') = \langle [\varphi(x) - \varphi(x')]^2 \rangle = 6.88(|x-x'|/r_0)^{5/3} \quad (2)$$

To define a statistical quantity that has a finite variance, we introduce the piston-removed phase, $\phi(x)$:

$$\phi(x) = \varphi(x) - \int \varphi(x') W_A(x') dx' \quad (3)$$

where $W_A(x)$, following the notation of Wallner¹, is a weighting function that is zero outside the aperture. We normalize $W_A(x)$ such that its integral is unity, $\int W_A(x) dx = 1$.

The piston-removed wavefront has the covariance function (equation 21 in Wallner)

$$\langle \phi(x)\phi(x') \rangle = -\frac{1}{2} D_\varphi(x-x') + g(x) + g(x') - a \quad (4)$$

where

$$g(x) = \frac{1}{2} \int D_\varphi(x-x') W_A(x') dx' \quad (5)$$

and

$$a = \frac{1}{2} \int g(x) W_A(x) dx \quad (6)$$

The function $g(x)$ and constant a are “generic” on an aperture in that the $(D/r_0)^{5/3}$ dependence can be factored out. We’ve plotted $g(x)$ for a circular filled aperture in Figure 1 and $a = 0.1498 \times 6.88(D/r_0)^{5/3}$. Figure 2 shows the piston-removed variance $\langle \phi(x)^2 \rangle$ as a function of radial distance from the center of the aperture. Note that the piston-removed phase is not a spatially-stationary process; the statistics depend strongly on the position within the aperture.

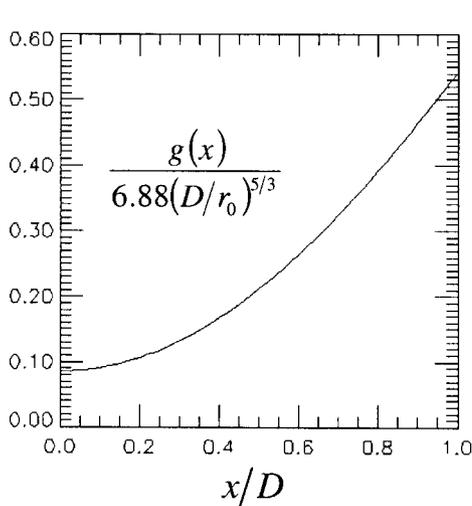


Figure 1. $g(x)$ from equation (5)

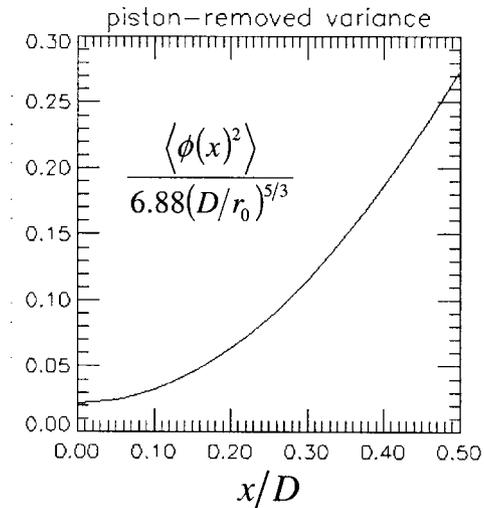


Figure 2. Piston-removed wavefront variance

3. THE CONDITIONAL MEAN WAVEFRONT AND THE SEPARATION PRINCIPLE

To proceed toward developing a Strehl-optimal control law, we first describe the actions of the wavefront sensor and the deformable mirror.

The wavefront sensor produces a vector of measurements, \mathbf{s} , which, for the sake of discussion, we can say are the Hartmann slopes in x and y of the wavefront. These measurements are corrupted by noise, \mathbf{v} , in a way that is consistent with the photon statistics and electronic read noise of the wavefront sensor camera. For simplicity, we assume that the wavefront sensor is linear in its response (in closed-loop operation where the wavefront sensor is operating on residuals, this is very nearly the case)

$$s_i = \int \phi(x) W_i^S(x) dx + v_i \quad (7)$$

where $W_i^S(x)$ (again using Wallner's notation) is the kernel that, in the Hartmann sensor case, takes the subaperture average gradient of wavefront phase.

The deformable mirror produces a wavefront correction

$$\phi_a(x) = \sum_{i=1}^{n_a} a_i r_i(x) \quad (8)$$

where \mathbf{a} is a vector of DM actuator commands, $r_i(x)$ is the response of the mirror surface to each actuator, and n_a is the total number of actuators. In general, the control law is the function $a_i = a_i(\mathbf{s})$ which takes a set of wavefront sensor measurements and maps to a set of actuator commands.

We wish to minimize the residual wavefront error $\varepsilon(x) = \phi(x) - \phi_a(x)$ in a mean-square sense. To restate this formally, the cost-function to be minimized is

$$J = \sigma_\varepsilon^2 = \int \langle \varepsilon(x)^2 \rangle W_A(x) dx \quad (9)$$

Our approach to this problem is to separate it into two components through the introduction of an intermediate variable, the conditional mean wavefront. The first component is associated with optimal *estimation* of the wavefront phase given the measurements from equation (7) and the second is associated with the optimal *control* of the DM surface given the optimal wavefront phase estimate and the action of equation (8). While separation is not strictly necessary in deriving the optimal control for the single measurement (open-loop) case, we shall see it is quite valuable in simplifying the derivation of the optimal control in the multi-measurement (closed-loop) case in Section 4.

The conditional mean wavefront, $\hat{\phi}(x) = \langle \phi(x) | \mathbf{s} \rangle$ is defined via Bayes theorem

$$\hat{\phi}(x) = \langle \phi(x) | \mathbf{s} \rangle \equiv \int \phi P_{\phi|\mathbf{s}}(\phi | \mathbf{s}) d\phi = \int \phi \frac{P_{\phi,\mathbf{s}}(\phi, \mathbf{s})}{P_{\mathbf{s}}(\mathbf{s})} d\phi \quad (10)$$

where $P_{\mathbf{s}}(\mathbf{s})$ is the joint probability distribution function for the measurement vector and $P_{\phi,\mathbf{s}}(\phi, \mathbf{s})$ is the joint probability distribution for the wavefront phase and the measurement vector. The conditional mean has a number of useful properties for our purposes, most importantly the *error* in the conditional mean, $\tilde{\phi}(x) = \phi(x) - \hat{\phi}(x)$ is uncorrelated to the data, \mathbf{s} ⁵. As a consequence, it can easily be shown that the *error* in the conditional mean is unbiased (has zero mean itself), is uncorrelated to the conditional mean (which depends only on the data), and is uncorrelated to the controls \mathbf{a} (since \mathbf{a} depends only on the data).

The cost function can now be written

$$\begin{aligned}
J &= \int \langle [\phi(x) - \phi_a(x)]^2 \rangle W_A(x) dx \\
&= \int \langle [(\phi(x) - \hat{\phi}(x)) - (\phi_a(x) - \hat{\phi}(x))]^2 \rangle W_A(x) dx \\
&= \int \langle \tilde{\phi}^2 \rangle W_A(x) dx - 2 \int \langle \tilde{\phi}(x)(\phi_a(x) - \hat{\phi}(x)) \rangle W_A(x) dx + \int \langle [\phi_a(x) - \hat{\phi}(x)]^2 \rangle W_A(x) dx \\
&= \int \langle \tilde{\phi}^2 \rangle W_A(x) dx + \int \langle [\phi_a(x) - \hat{\phi}(x)]^2 \rangle W_A(x) dx
\end{aligned} \tag{11}$$

The cross term equals zero because the error in the conditional mean is uncorrelated to the actuator commands and also uncorrelated to the conditional mean. Define $J_E = \int \langle \tilde{\phi}(x)^2 \rangle W_A(x) dx$ and $J_C = \int \langle [\phi_a(x) - \hat{\phi}(x)]^2 \rangle W_A(x) dx$. J is minimized if J_E and J_C are separately minimized.

Minimizing J_E is simply a matter of finding the conditional mean error variance, since J_E is its integral by definition. It might be instructive and intuitively appealing to show also that the conditional mean is in fact the one estimate of $\phi(x)$ that minimizes the estimate's error variance. Let $\phi_E(x, \mathbf{s})$ be any estimate of $\phi(x)$ given (only) the data, \mathbf{s} . $\phi_E(x, \mathbf{s})$ differs from $\hat{\phi}(x, \mathbf{s})$ by $\delta\phi(x, \mathbf{s})$, i.e. $\phi_E(x, \mathbf{s}) = \hat{\phi}(x, \mathbf{s}) + \delta\phi(x, \mathbf{s})$. Then (dropping the x dependence; the integrals are over distributions):

$$\begin{aligned}
\langle [\phi - \phi_E(s)]^2 \rangle &= \langle [\tilde{\phi} - \delta\phi]^2 \rangle \\
&= \langle \tilde{\phi}^2 \rangle - 2 \langle \tilde{\phi} \delta\phi \rangle + \langle \delta\phi^2 \rangle \\
&= \langle \tilde{\phi}^2 \rangle - 2 \int \int \tilde{\phi}(\phi, s) \delta\phi(s) P(\phi, s) d\phi ds + \langle \delta\phi^2 \rangle \\
&= \langle \tilde{\phi}^2 \rangle - 2 \int \delta\phi(s) P_S(s) \left[\int \tilde{\phi}(\phi, s) P(\phi | s) d\phi \right] ds + \langle \delta\phi^2 \rangle \\
&= \langle \tilde{\phi}^2 \rangle + \langle \delta\phi^2 \rangle
\end{aligned} \tag{12}$$

where the cross term is zero because the error in the conditional mean is zero-mean. Since the two remaining terms on the right hand side are positive, $J_E = \langle \tilde{\phi}^2 \rangle < \langle (\phi - \phi_E)^2 \rangle$ for any $\delta\phi \neq 0$ and thus $\phi_E(s) = \hat{\phi}(s)$ minimizes J_E .

3.1 Calculating the conditional mean

To calculate the conditional mean given the data we exploit the well known fact⁵ that for Gaussian distributed ϕ and \mathbf{v} , the conditional mean is a linear function of \mathbf{s}

$$\hat{\phi}(s) = \sum_{i=1}^{n_i} k_i(x) s_i = \mathbf{K} \mathbf{s} \tag{13}$$

We cross-correlate both sides with \mathbf{s} and solve for the matrix \mathbf{K}

$$\begin{aligned}
\langle \hat{\phi} \mathbf{s}^T \rangle &= \mathbf{K} \langle \mathbf{s} \mathbf{s}^T \rangle \\
\mathbf{K} &= \langle \hat{\phi} \mathbf{s}^T \rangle \langle \mathbf{s} \mathbf{s}^T \rangle^{-1} = \langle \phi \mathbf{s}^T \rangle \langle \mathbf{s} \mathbf{s}^T \rangle^{-1}
\end{aligned} \tag{14}$$

where we can make the last step because $\langle \tilde{\phi} \mathbf{s} \rangle = 0$. Working out these cross-correlations with the help of equation (7) we have

$$\hat{\phi}(x) = \mathbf{p}^T(x) \mathbf{S}^{-1} \mathbf{s} \tag{15}$$

where

$$\begin{aligned} p_i(x) &= \langle \phi(x) s_i \rangle = \int \langle \phi(x) \phi(x') \rangle W_i^s(x') dx' \\ S_{ij} &= \langle s_i s_j \rangle = \iint W_i^s(x) W_j^s(x') \langle \phi(x) \phi(x') \rangle dx dx' + \langle v_i v_j \rangle \end{aligned} \quad (16)$$

3.2 Calculating the optimal control

Minimizing J_C is the process of finding the vector of commands, \mathbf{a} , that best fit the deformable mirror surface to the conditional mean wavefront, $\hat{\phi}(x)$, under the constraints of equation (8). The best-fit solution is found by substituting equation (8) into the expression for J_C , taking the partial derivative with respect to \mathbf{a} and setting equal to zero. The result is

$$\mathbf{a} = \mathbf{R}^{-1} \mathbf{r} \quad (17)$$

where

$$\begin{aligned} \mathbf{r} &= \left[\int r_j(x) \hat{\phi}(x) W_A(x) dx \right] \\ \mathbf{R} &= \left[\int r_i(x) r_j(x) W_A(x) dx \right] \end{aligned} \quad (18)$$

3.3 Combined optimal estimation and control

Combining equations (15) through (18) yields the following expression for \mathbf{a} :

$$\mathbf{a} = \mathbf{R}^{-1} \mathbf{A} \mathbf{S}^{-1} \mathbf{s} \quad (19)$$

where

$$\mathbf{A} = \int r_i(x) p_j(x) W_A(x) dx \quad (20)$$

which is exactly the same result as that of Wallner. The important difference is that through the use of the separation principle we are able to write the single-measurement solution in terms of the conditional mean wavefront (equation (15)). Using the conditional mean, we will be able to extend this approach to the repeated-measurement, closed loop, control case.

4. THE CLOSED-LOOP CONDITIONAL MEAN WAVEFRONT WITH MOVING TURBULENCE

Closed-loop adaptive optics systems measure and correct the wavefront on time scales that are short compared to the wavefront correlation time. For example, a typical AO system may sample every millisecond, and yet (assuming a Taylor frozen-flow model) it takes the wind clearing time, $\sim D/v$, for turbulence to blow across the aperture. For a typical wind speed of 10 m/sec and a 10 meter aperture, the clearing time is an entire second, so wavefront phase at any point is actually measured 1000 times! We ought to be able to take advantage of the extra sample averaging to significantly improve performance, or, equivalently, significantly increase the limiting guide star magnitude. To do this, we need to condition the wavefront estimate on *all* of the previous data, not just the last measurement:

$$\hat{\phi}(x) = \langle \phi(x) | \mathbf{s}_t, \mathbf{s}_{t-1}, \mathbf{s}_{t-2}, \dots \rangle \quad (21)$$

In the closed-loop formulation, at any sample time we assume we have available to us the present value of the conditional mean wavefront and the task is to incorporate one new wavefront sensor measurement to revise the conditional mean. Furthermore, as time then progresses, we need to progress the conditional mean through time. Since the Taylor frozen-flow model is commonly used, and not a bad approximation, we incorporate it in our developments below. We introduce the following notation to describe the various stages of updating and progressing the conditional mean through time

$$\begin{aligned}
\hat{\phi}_{t-1}^+(x) &= \langle \phi_{t-1}(x) | \mathbf{s}_{t-1}, \mathbf{s}_{t-2}, \dots \rangle \\
\hat{\phi}_t^-(x) &= \langle \phi_t(x) | \mathbf{s}_{t-1}, \mathbf{s}_{t-2}, \dots \rangle \\
\hat{\phi}_t^+(x) &= \langle \phi_t(x) | \mathbf{s}_t, \mathbf{s}_{t-1}, \mathbf{s}_{t-2}, \dots \rangle \\
&\dots
\end{aligned} \tag{22}$$

The + superscript indicates an updated conditional mean given the latest data at the time indicated by the subscript. The – superscript indicates a time-progressed conditional mean given data up to one sample previous. A diagram of this process is shown in Figure 3. This method of incorporating new data and progressing stochastic estimates through time was first presented by Kalman⁶.

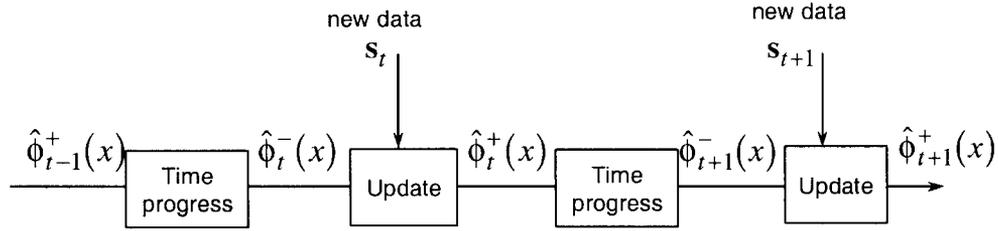


Figure 3. Incorporating new data to update an estimated wavefront over time

There are some difficulties in formulating the optimal closed-loop control law for adaptive optics. First of all, the conditional mean is defined on a continuous domain (x within the aperture) which (in principle) means keeping track of an infinite number of data points, although in practice one would keep track of only a suitably large number of discrete sample points. Secondly, the recalculation of the conditional mean at each step will, as we shall see, require that the conditional mean error covariance $C(x, x') = \langle \tilde{\phi}(x) \tilde{\phi}(x') \rangle$ be available at each step. The covariance is bi-functional i.e. dependent on two points within the aperture. Furthermore, this bi-function must be inverted, that is, a Green's function $G(x, x')$ must be found such that $\int G(x, x'') C(x'', x') dx'' = \delta(x, x')$. Finally, parameters that determine wavefront prior statistics and dynamics, that is r_ϕ and the wind velocity(ies) are assumed known. It is not immediately obvious how the controller will degrade if these parameters are not known precisely, however we explore this issue in the example simulations.

We justify the extra effort however by the following reasoning. The optimal stochastic control law is in itself interesting, if only for off-line simulation and analysis, since we can compare performance to that of the “non-optimal” control law actually implemented. It can tell us how much improvement might be possible. Secondly, we can gain insights into how sensitive the optimal controller is to prior knowledge of operating conditions and compare the degraded performance of the optimal controller to the simpler one. Finally, in light of future large aperture telescopes and laser guide star adaptive optics, the added benefit of multiple measurements of the wavefront may be enough to reduce the laser guide star power requirements and hence the cost of such systems.

4.1. Updating the conditional mean with new data

Given a new measurement of the wavefront, we must combine the information contained in that measurement with information we already have derived from previous measurements. The process of this “statistical addition” is outlined as follows: 1) extract the new information by subtracting from the measurement the *predicted* value of the measurement given the old data. Then 2) add the new information, with the proper scaling, into the wavefront estimate.

Assume we have an estimate based on an old measurement

$$\hat{\phi}^- = \langle \phi | \mathbf{s}_{t-1} \rangle \tag{23}$$

we then take a new measurement and subtract from it the conditional mean of the new measurement to form a residual

$$\mathbf{e}_t = \mathbf{s}_t - \langle \mathbf{s}_t | \mathbf{s}_{t-1} \rangle \quad (24)$$

We should note that since \mathbf{e}_t is an error in a conditional mean, it is uncorrelated to the measurements, $\langle \mathbf{e}_t, \mathbf{s}_t^T \rangle = 0$. Using (14) we calculate the new conditional mean wavefront

$$\begin{aligned} \hat{\phi}^+ &= \langle \phi | \mathbf{s}_t, \mathbf{s}_{t-1} \rangle \\ &= \langle \phi | \mathbf{e}_t, \mathbf{s}_{t-1} \rangle \\ &= \left[\begin{array}{cc} \langle \phi \mathbf{s}_{t-1}^T \rangle & \langle \phi \mathbf{e}_t^T \rangle \end{array} \right] \left[\begin{array}{cc} \langle \mathbf{s}_{t-1} \mathbf{s}_{t-1}^T \rangle & \langle \mathbf{s}_{t-1} \mathbf{e}_t^T \rangle \\ \langle \mathbf{e}_t \mathbf{s}_{t-1}^T \rangle & \langle \mathbf{e}_t \mathbf{e}_t^T \rangle \end{array} \right]^{-1} \begin{bmatrix} \mathbf{s}_{t-1} \\ \mathbf{e}_t \end{bmatrix} \\ &= \langle \phi \mathbf{s}_{t-1}^T \rangle \langle \mathbf{s}_{t-1} \mathbf{s}_{t-1}^T \rangle^{-1} \mathbf{s}_{t-1} + \langle \phi \mathbf{e}_t^T \rangle \langle \mathbf{e}_t \mathbf{e}_t^T \rangle^{-1} \mathbf{e}_t \\ &= \langle \phi | \mathbf{s}_{t-1} \rangle + \langle \tilde{\phi}^- | \mathbf{e}_t \rangle \end{aligned} \quad (25)$$

where the matrix inverse is simplified because the off-diagonal terms are zero. The first term is the old estimate, $\hat{\phi}^-$, and the second term is the “statistical addition” of the new information, \mathbf{e}_t . Again, working out the cross-correlations we get

$$\hat{\phi}_t^+(x) = \hat{\phi}_t^-(x) + \mathbf{p}_t^T(x) \mathbf{S}_t^{-1} (\mathbf{s}_t - \hat{\mathbf{s}}_t) \quad (26)$$

where

$$\begin{aligned} \hat{\mathbf{s}}_t &= \langle \mathbf{s}_t | \mathbf{s}_{t-1}, \mathbf{s}_{t-2}, \dots \rangle = \int \mathbf{W}^s(x) \hat{\phi}_t^-(x) dx \\ \mathbf{p}_t(x) &= \langle \phi(x) \mathbf{e}_t \rangle = \int \langle \tilde{\phi}_t^-(x) \tilde{\phi}_t^-(x') \rangle \mathbf{W}^s(x') dx' \\ \mathbf{S}_t &= \langle \mathbf{e}_t \mathbf{e}_t^T \rangle = \iint \mathbf{W}^s(x) \mathbf{W}^{sT}(x') \langle \tilde{\phi}_t^-(x) \tilde{\phi}_t^-(x') \rangle dx dx' + \langle \mathbf{v} \mathbf{v}^T \rangle \end{aligned} \quad (27)$$

Note that the elements of $\mathbf{p}_t(x)$ and \mathbf{S}_t depend on the covariance of the prior wavefront estimate error and the covariance of the measurement noise. The covariance of the wavefront estimate error is updated by squaring and averaging (26) and using that fact (again) that the error in the conditional mean is uncorrelated to the prior data, $\langle \hat{\phi}_t^-(\mathbf{s}_t - \hat{\mathbf{s}}_t)^T \rangle = 0$:

$$\langle \tilde{\phi}_t^+(x) \tilde{\phi}_t^+(x') \rangle = \langle \tilde{\phi}_t^-(x) \tilde{\phi}_t^-(x') \rangle - \mathbf{p}_t^T(x) \mathbf{S}_t^{-1} \mathbf{p}_t(x') \quad (28)$$

where $\tilde{\phi}_t^\pm(x) \equiv \phi_t(x) - \hat{\phi}_t^\pm(x)$. The initial error covariance at $t = 0$ is simply the a-priori Kolmogorov wavefront covariance, $\langle \tilde{\phi}_0^-(x) \tilde{\phi}_0^-(x') \rangle = \langle \phi(x) \phi(x') \rangle$.

The update process (26) is inherently an iterative, closed-loop “null-seeking” control architecture, since it uses residuals, not direct wavefront measurements. The loop seeks to minimize the variance of these residuals. Figure 4 shows the loop architecture, which bears considerable resemblance to that of the traditional AO controller. There are some important differences however. First of all, the measured wavefront residual, that is, the difference between atmospheric phase and the correction put on the deformable mirror, is not exactly the same as the error in the conditional mean, since the shape on the mirror can not necessarily be made to exactly fit the conditional mean wavefront. So there is a correction term that takes into account the fitting error:

$$\begin{aligned}
\int \mathbf{W}^s(x) \varepsilon(x) dx &= \int \mathbf{W}^s(x) [\phi(x) - \phi_a(x)] dx \\
&= \int \mathbf{W}^s(x) [\phi(x) - \hat{\phi}(x)] - [\phi_a(x) - \hat{\phi}(x)] dx \\
&= (\mathbf{s} - \hat{\mathbf{s}}) - \int \mathbf{W}^s(x) [\phi_a(x) - \hat{\phi}(x)] dx
\end{aligned} \tag{29}$$

Secondly, the ‘‘predictor’’ portion of the control loop, which bears some resemblance to the leaky integrator in standard AO control, is not simply a leaky integrator but a much more complicated operation that progresses the conditional mean through time, which we develop in the following section.

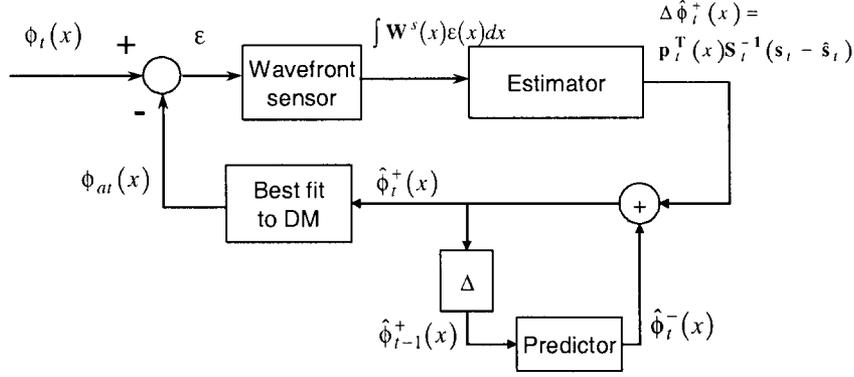


Figure 4. Control loop architecture for the Strehl-Optimal controller.

4.2. Progressing the conditional mean to the next time step

4.2.1. Conditional mean wavefront progression under frozen flow

Given the mean wavefront $\hat{\phi}_{t-1}^+(x)$ conditioned on all the data up to time $t-1$, how do we project ahead to get a new mean wavefront, conditioned on the same data but valid at the next sample time, $\hat{\phi}_t^-(x)$? If, for example, the wavefront were unchanging over time, then $\hat{\phi}_t^-(x) = \hat{\phi}_{t-1}^+(x)$. We will assume, more generally, that the wavefront is translating across the aperture with uniform wind velocity \mathbf{v} :

$$\phi_t(\mathbf{x}) = \phi_{t-1}(\mathbf{x} - \mathbf{v}t) \tag{30}$$

For the wavefront estimate then we would expect that that portion of the wavefront phase screen retained inside the aperture will simply be the same, albeit translated by $\mathbf{v}t$ and the portion of wavefront phase screen that is newly blown in will need to be estimated using previous data. A minor adjustment must be made for the fact that the two apertures will have different constant piston terms subtracted.

Let the set A represent the portion of the atmosphere over the aperture at time t and A' the portion of the atmosphere that was over the aperture at time $t-1$. Define $F(x, x')$ as the operator that satisfies the integral equation

$$\langle \phi_t(x \in A) \phi_{t-1}(x'' \in A') \rangle = \int_{x' \in A'} F(x, x') \langle \phi_{t-1}(x' \in A') \phi_{t-1}(x'' \in A') \rangle dx' \tag{31}$$

(Note, in a discretization of the sets A and A' , equation (31) is converted to a matrix equation, which can be solved by computing a matrix inverse.) Under frozen flow, $F(x, x')$ depends on only a single parameter, \mathbf{v}/D , the wind velocity relative to the aperture size. Now define the quantity $q(x)$ as

$$q_t(x) = \phi_t(x) - \int_{x' \in A'} F(x, x') \phi_{t-1}(x') dx' \tag{32}$$

By analogy with (13) and (14), the second term on the right side of (32) is the mean of $\phi_t(x \in A)$ conditioned on $\phi_{t-1}(x \in A')$ and $q_t(x)$ is the error in this conditional mean.

We need a method of calculating the new conditional mean wavefront in A conditioned on the data, $\{s_{t-1}, s_{t-2}, \dots\}$, given the old conditional mean wavefront in A' which is also conditioned on the same data, $\{s_{t-1}, s_{t-2}, \dots\}$. To begin to do this, we take the conditional mean of both sides of (32) with respect to the data and rearrange to get

$$\hat{\phi}_t^-(x) = \int_{x' \in A'} F(x, x') \hat{\phi}_{t-1}^+(x') dx' + \langle q_t(x) | s_{t-1}, s_{t-2}, \dots \rangle \quad (33)$$

The first term is straightforward to calculate given $\hat{\phi}_{t-1}^+(x)$ and $F(x, x')$. The second term is problematic but it can be dealt with if we make a certain reasonable approximation, which we will describe below.

4.2.2. The near-Markov approximation

We now make a crucial approximation that shall be justified shortly. We assume that

$$\langle \phi_t(x \in A) | \phi_{t-1}(x \in A') \rangle \cong \langle \phi_t(x \in A) | \phi_{t-1}(x \in A'), \phi_{t-2}(x \in A''), \dots \rangle \quad (34)$$

which we'll dub the "near-Markov" approximation. It states that *almost* all the information needed to predict the wavefront phase on A is contained in the wavefront phase on A' and we can neglect the previous history of the phase screen that blew past the aperture at earlier times. A Markov process (by definition) would contain exactly all of the information [ref Pappoulis]. That this is *not* the case in frozen-flow Kolmogorov turbulence is illustrated in Figure 5. Notice that at time t , some points that were in the aperture at time $t-1$ have fallen out of the aperture set, off the downwind side. On the upwind side, a new phase point coming in at time $t+1$ could possibly be better estimated if we didn't neglect the phase estimates that fell off the downwind side. We give a compelling argument here however that, under reasonable wind conditions, these neglected points will contribute very little additional information to that contained in A' and therefore the near-Markov approximation is useful to work with. The advantage of this approximation is that we do not need to keep, in a growing memory, all the prior history of wavefront measurements in order to make a forward prediction, but need only retain the conditional mean wavefront on a finite aperture.

To demonstrate the validity of the approximation, we consider the point in A that is estimated given data in A' . In particular, consider one such point near the windward edge. One of the points in A'' that has been neglected in the near-Markov approximation can also contribute some information about the new point in A , but this is only the new information that is not otherwise contained in A' . That is (letting ϕ_A be shorthand for $\phi(x \in A)$)

$$\langle \phi_A | \phi_{A'}, \phi_{A''} \rangle = \langle \phi_A | \phi_{A'} \rangle + \langle \phi_A | \xi_{A''} \rangle \quad (35)$$

where

$$\xi_{A''} = \phi_{A''} - \langle \phi_{A''} | \phi_{A'} \rangle \quad (36)$$

These conditional means can be calculated generically for Kolmogorov turbulence on a circular aperture and the results are shown in Figure 6. It can be seen that for wind speeds that clear only a fraction of the aperture per sample time, the contribution of the neglected point is orders of magnitude below that of the leading edge point. In fact, additional calculations show clearly that most of the information that is used to form estimates of points in A comes only from points that are located along the leading edge of A' .

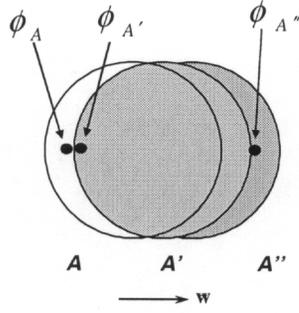


Figure 5. $\phi_{A'}$ is the neglected point in the near-Markov approximation

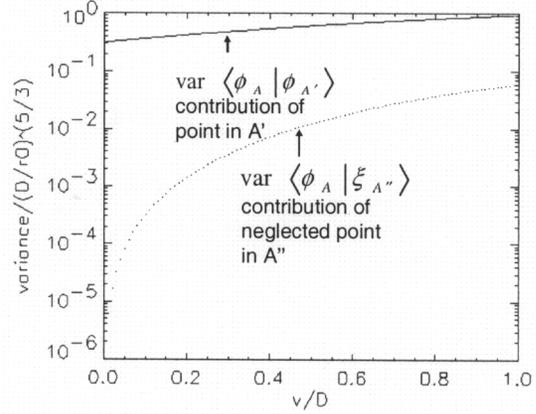


Figure 6. The contribution of the neglected point is negligible at low wind speeds

4.2.3. Progressing the conditional mean with frozen flow under the near-Markov approximation

Now, as a result of (31) and (32) we see that $q_t(x)$ is the error in a conditional mean and so is uncorrelated to the data it is conditioned on, that is, $\langle q_t \phi_{t-1} \rangle = \langle (\phi_t - \langle \phi_t | \phi_{t-1} \rangle) \phi_{t-1} \rangle = 0$. Invoking the near-Markov approximation, we can state $\langle q_t \phi_{t-k} \rangle \cong \langle (\phi_t - \langle \phi_t | \phi_{t-1}, \phi_{t-2}, \dots \rangle) \phi_{t-k} \rangle = 0$ for any $k \geq 1$ as well. Since, according to (7), $s_{t,k}$ is a linear combination of ϕ_{t-k} plus noise that is uncorrelated to ϕ , we can state that $\langle q_t s_{t-k} \rangle \cong 0$, and rewrite (33)

$$\hat{\phi}_t^-(x) \cong \int_{A'} F(x, x') \hat{\phi}_{t-1}^+(x') dx' \quad (37)$$

It is also important to point out that $\langle q_t \tilde{\phi}_{t-k} \rangle \cong 0$ because both $\hat{\phi}_{t-1}^+$ and ϕ_{t-1} consist only of linear combinations of s_{t-k} , $k \geq 1$, plus measurement noise uncorrelated to ϕ .

4.2.4. Wavefront cross-statistics for computing $F(x, x')$

To find the function $F(x, x')$ satisfying (31) we need to know the correlation function for the wavefront within aperture A' (the term inside the integral) and the cross-correlation function between the wavefronts in A and A' (the left hand side term). The correlation function for piston-removed phases within an aperture was given by equations (2) through (6). For the piston-removed phases on different apertures, we have

$$\langle \phi(x \in A) \phi(x' \in A') \rangle = -\frac{1}{2} D_\varphi (|x - x'|) + g(|x - c'|) + g(|x' - c|) - a(|c - c'|) \quad (38)$$

where c and c' are the center points of apertures A and A' respectively and $a(x)$ is

$$a(x) = \frac{1}{2} \int g(x - x') W_A(x') dx \quad (39)$$

Once again, for Kolmogorov turbulence these functions are generic on an aperture, within a $(D/r_0)^{5/3}$ factor.

4.2.5. Progressing the conditional mean error covariance

Squaring (32), substituting (31) to work out the cross terms, we get the covariance of $q_t(x)$ in terms of the Kolmogorov wavefront covariance and the function $F(x, x')$:

$$Q_t(x, x') = \langle q_t(x) q_t(x') \rangle = \langle \phi_t(x) \phi_t(x') \rangle - \iint_{x'', x''' \in A} F(x, x'') \langle \phi_{t-1}(x'') \phi_{t-1}(x''') \rangle F(x', x''') dx'' dx''' \quad (40)$$

Note that since $F(x, x')$ depends only on a single parameter, v/D , and the phase covariance depends only on r_0/D . Thus $Q(x, x')$ depends on both parameters, r_0/D and v/D .

Combining (37) and (32) we see that the error in the conditional mean obeys

$$\tilde{\phi}_t^-(x) = \int_{A'} F(x, x') \tilde{\phi}_{t-1}^+(x') dx' + q_t(x) \quad (41)$$

Taking the mean square, setting the cross-term to zero since $\langle q, \tilde{\phi}_{t-k}^- \rangle \equiv 0$, we get the equation for the progression of conditional mean error covariance

$$\langle \tilde{\phi}_t^-(x) \tilde{\phi}_t^-(x') \rangle = \iint_{x'', x''' \in A'} F(x, x'') \langle \tilde{\phi}_{t-1}^+(x'') \tilde{\phi}_{t-1}^+(x''') \rangle F(x', x''') dx'' dx''' + Q_t(x, x') \quad (42)$$

5. SUMMARY OF PROCESSING STEPS

To implement the Strehl-optimal control law, the conditional mean is updated and progressed according to equations (26), and (37) respectively, using the definitions in (27) and (31). Since these equations require the conditional mean error covariance at each step, the error covariance must be updated and progressed according to equations (28) and (42) respectively. To get the conditional mean residual, \mathbf{e} , (for use in (26)) given the measurement from the wavefront sensor, we must apply the correction given by (29). Finally, given the conditional mean wavefront $\hat{\phi}(x)$, the closed-loop commands to the deformable mirror, \mathbf{a} , are given by (17) and (18).

6. SIMULATIONS

The Strehl-optimizing controller was programmed for a simple example problem in order to compare its behavior to a “fixed matrix” integral controller and also to probe the degradation from optimal performance when the seeing parameters are not precisely known. The problem consists of a Kolmogorov phase screen blowing past a circular aperture. We choose $r_0 = d$ and $D/d = 7$ (D is the aperture diameter and d is the Hartmann subaperture diameter and also the spacing between actuators). The wind speed is $0.1 d$ per sample and the Hartmann slope measurement noise is 1 radian of phase / d , rms. We compared 6 simulation cases: 1) correction with the Strehl-optimizing closed-loop control law, 2) correction with a Strehl-optimizing closed-loop control law that assumes the wind velocity was 50% slower than reality, 3) correction with a Strehl-optimizing closed-loop control law that assumes the wind velocity was 50% faster than reality, 4) correction with an integral controller using Wallner’s control matrix and unity feedback gain, 5) correction with an integral controller using Wallner’s control matrix with feedback gain adjusted for maximum Strehl, and 6) uncorrected. Results are summarized in Table 1. We find that the closed-loop Strehl-optimizing controller performs better than the fixed-matrix integral controllers, even when the wind velocity is uncertain to $\pm 50\%$.

Table 1. Simulation results

Algorithm	Time-average Strehl
Optimal	0.502
Optimal, Wind -50%	0.448
Optimal, Wind +50%	0.492
Wallner, gain = 0.3 (optimized)	0.440
Wallner, gain = 1	0.415
Uncorrected	0.276

7. CONCLUSION

We have presented a method for determining the Strehl-optimizing closed-loop feedback control for adaptive optics. The approach is general, but we’ve worked out a specific solution for the case of frozen-

flow Kolmogorov turbulence. A very mild approximation, the “near-Markov” approximation, allows us to formulate a control law where only the conditional mean wavefront on the finite aperture is needed to carry forward all the statistical information from the past history of wavefront measurements during closed loop operation. Some preliminary simulations have confirmed that improved Strehl performance over traditional integral feedback control methods can be expected. The controller has the desirable null-seeking, feedback control architecture, and has shown in simulations that it is somewhat robust to the seeing parameter assumptions made in its statistical formulation.

ACKNOWLEDGEMENTS

This work was performed under the auspices of the U.S. Department of Energy by University of California, Lawrence Livermore National Laboratory under Contract W-7405-Eng-48. This work has been supported by the National Science Foundation Science and Technology Center for Adaptive Optics, managed by the University of California at Santa Cruz under cooperative agreement No. AST – 9876783.

REFERENCES

1. Wallner, E.P., *Optimal wave-front correction using slope measurements*, **J. Opt. Soc. Am.**, 73, December 1983, p1771.
2. Schwartz, C.; Baum, G.; Ribak, E. N., *Turbulence-degraded wave fronts as fractal surfaces*, **J. Opt. Soc. Am. A**, 11, 1, January 1994, pp.444-451.
3. Wild, W.J., *Predictive optimal estimators for adaptive-optics systems*, **Optics Letters**, 21, 18, September 15, 1996, pp.1433-1435.
4. Wild, W.J., *Innovative wavefront estimators for zonal adaptive optics systems: II*, **Proc. SPIE**, 3353, September 1998, pp. 1164-1173.
5. Papoulis, A., *Probability, Random Variables and Stochastic Processes*, McGraw Hill, 1991.
6. Kalman, R.E., *A New Approach to Linear Filtering and Prediction Problems*, **J. Basic Eng., Trans. ASME**, 82,1, 1960.