

Fourier transform wavefront control with adaptive prediction of the atmosphere

Lisa A. Poyneer,^{1,*} Bruce A. Macintosh,¹ and Jean-Pierre Véran²

¹Lawrence Livermore National Laboratory, 7000 East Avenue, Livermore, California 94550, USA

²Herzberg Institute of Astrophysics, 5071 West Saanich Road, Victoria, British Columbia, Canada V9E2E7

*Corresponding author: poyneer1@llnl.gov

Received November 8, 2006; revised March 21, 2007; accepted April 18, 2007;
posted April 27, 2007 (Doc. ID 76885); published July 30, 2007

Predictive Fourier control is a temporal power spectral density-based adaptive method for adaptive optics that predicts the atmosphere under the assumption of frozen flow. The predictive controller is based on Kalman filtering and a Fourier decomposition of atmospheric turbulence using the Fourier transform reconstructor. It provides a stable way to compensate for arbitrary numbers of atmospheric layers. For each Fourier mode, efficient and accurate algorithms estimate the necessary atmospheric parameters from closed-loop telemetry and determine the predictive filter, adjusting as conditions change. This prediction improves atmospheric rejection, leading to significant improvements in system performance. For a 48×48 actuator system operating at 2 kHz, five-layer prediction for all modes is achievable in under 2×10^9 floating-point operations/s.

OCIS codes: 010.1080, 010.1330, 070.6020.

1. OPTIMAL AND PREDICTIVE WAVEFRONT CONTROL

A. Background and Requirements

In adaptive optics (AO), a time-varying phase aberration in the optical system is compensated in real time, generally under closed-loop control. Most AO systems use a fixed temporal control law to implement this feedback control, e.g., an integral controller. In recent years substantial work has been done in developing and implementing more advanced wavefront control schemes that provide better performance but require more computation.

Notable among these schemes is modal gain optimization, which was developed by Gendron and Léna [1]. Modal control uses a decomposition of the wavefront into a modal basis set, which is a set of spatial signals (such as Zernike modes) that are orthogonal. In this framework, each mode of the control system is controlled independently. The gain of the integral controller is optimized using closed-loop system telemetry for each mode. This approach is used in both the Altair system [2] at Gemini and the European Southern Observatory's (ESO's) NAOS [3]. Recently, the method of Fourier transform wavefront reconstruction (FTR) [4] has been adapted into a modal form. Termed optimized-gain Fourier control (OFC) [5], this method combines the computationally efficient wavefront reconstruction of FTR with gain optimization of the Fourier modes. This method is efficient enough to be implemented with off-the-shelf hardware in a 64×64 AO system operating at 2.5 kHz [5]. The optimization of signal and noise components results in reduction of the residual error for all controllable spatial frequencies of the AO system. The exact reduction varies with spatial frequency, but it can be in the range of 50% to 80% under reasonable conditions in simulations [6].

It is particularly important to minimize temporal errors for "extreme" adaptive optics (ExAO) systems, which are designed for extrasolar planet detection. In this regime, the figure of merit is not Strehl ratio, but contrast—the amount of scattered light at a given location in the point-spread function (PSF). Wavefront sensor (WFS) measurement noise and control-loop time-lag errors are the main source of scattered light within the "dark hole" defined by the controllable spatial frequencies of the system, and control algorithms that reduce these errors can substantially increase the science capabilities of ExAO. A representative ExAO system is the Gemini Planet Imager (GPI) [7]. Designed for the 8 m Gemini South telescope, this system has 44 subapertures across the primary mirror and is currently designed to operate at 2.5 kHz to minimize temporal errors, with a 128×128 pixel WFS operated in a quad-cell mode. In addition, particular attention is paid to removing residual static wavefront errors, which produce quasi-static speckles that can otherwise mask a planet. ESO's SPHERE ExAO system [8] is also currently being developed and has similar fundamental characteristics.

OFC is the baseline wavefront control method for GPI, but we are highly interested in further improving dynamic wavefront correction. The primary requirement for any new advanced controller is that it never perform worse than OFC. Flowing down from this are three secondary requirements. First, the method will not rely exclusively on *a priori* models. For example, if the method assumes a translation by the wind, it must measure the wind velocity and direction, not rely on a model such as "at this site, the wind is normally 10 m/s." Second, since the true atmospheric conditions will change during observation, the system must update itself as these conditions change. As a baseline requirement the system should up-

date itself at least every 10 s, though experimental data as to actual on-sky performance will inform whether this requirement needs to be loosened or tightened. Finally, the new method must be computationally efficient enough to run on off-the-shelf hardware. This means a peak computational load that is under 3×10^9 floating-point operations per second (FLOPs/s).

B. Possible Solutions

Two methods that provide further sophistication have been developed, both of which are focused on using prediction to compensate for the inherent temporal lags in an AO system. The first, by Dessenne *et al.* [9], extends the modal gain optimization framework to arbitrary control laws. Though substantial attention is paid to the temporal power spectral density (PSD) approach, Dessenne's final predictor strategy uses a time-domain algorithm similar to recursive least squares (RLS). This method uses open-loop telemetry to estimate the coefficients of the predictive temporal filter, optimizing each mode independently. This method does not explicitly use a model for the atmosphere in the estimation process. A particular insight of Dessenne's, to which we will return, is that the predictor's temporal frequency response should approximate as closely as possible the temporal spectral signal-to-noise ratio (SNR) of the system. However, there are drawbacks to this approach as it is proposed. The time-domain iterations that are used to obtain the filter coefficients are done on open-loop telemetry, which we do not have direct access to during closed-loop operation. More important, RLS-type algorithms can have poor convergence, both in terms of large variation in parameters around the final value and many hundreds (or thousands) of iterations to converge to this final value. We have experienced such difficulties when implementing this method in the Fourier mode-based GPI simulation. In addition there are stability issues; see Section 6.

The second method under consideration for prediction is the Kalman filter. The Kalman filter is both well established and widely used in signal processing. It requires a state space model for the signal to be estimated (e.g., the phase aberration in AO) and a control system. Given this model, the Kalman filter provides the minimum mean-squared-error (MSE) estimate of the phase in the future, using the statistical information in the state space model. Gavel and Wiberg [10] have applied the Kalman filtering concept to optimal AO control, using the terminology of the conditional mean of the phase defined on the actuator grid. Initial work has been done by Wiberg *et al.* [11] in applying this conditional mean framework to frozen flow. Le Roux *et al.* [12] directly applied the Kalman concept to AO and multiconjugate AO using the Zernike modes. That technique has been experimentally validated for tip-tilt vibration reduction by Petit *et al.* [13].

Kalman filtering is justly favored in signal processing. But as it has been proposed to be used in AO, there are two weaknesses that prevent it from being effective in a high-order real-time system with dynamic changes in the atmosphere. The first of these is the fact there is no proposed method for estimating the state space model parameters from real-time observations. Gavel and Wiberg assume that the parameters are known. Le Roux assumes

a first-order autoregressive model for the Zernike modes. Even though this model is not an accurate match to the temporal PSDs, Le Roux asserts that in the multiple layer case, it is adequate to provide improved performance. However, the better the model structure and the parameters, the better the predictor performance.

The second major hurdle is that Kalman filtering can be computationally expensive. The approaches of both Le Roux and Gavel and Wiberg require state space models that contain either all modes or all actuators. Keeping track of and multiplying the necessary matrices is computationally prohibitive for AO systems with hundreds (let alone thousands) of actuators or controlled modes. The computational costs can be evaluated in terms of the number of modes n_{modes} . Using the Kalman iterations directly requires the multiplication of large matrices, resulting in $\mathcal{O}(n_{\text{modes}}^3)$ operations. As such, it is usually assumed (see [8,11,13]) that any Kalman filter will be precomputed offline based on fixed statistical priors. Given this steady-state result, prediction is accomplished in a single update equation with several vector matrix multiplies of cost $\mathcal{O}(n_{\text{modes}}^2)$. In the Le Roux case, the largest matrix is of size $3n_{\text{modes}} \times 3n_{\text{modes}}$, which makes prediction several times more expensive than the vector matrix multiply (VMM) for phase reconstruction.

The best way to reduce this computational load is to find a basis set where the modes can be dealt with independently. For L degrees of freedom for each of n_{modes} , the all-modes-at-once approach requires manipulations of matrices of size $Ln_{\text{modes}} \times Ln_{\text{modes}}$. Dealing with each mode individually results instead in n_{modes} matrices, each of size $L \times L$. This independence must be fulfilled in both the atmospheric model and the AO system model. Le Roux assumes that the WFS produces statistically correlated measurements of the Zernikes, particularly in the multiconjugate case, so such a mode-by-mode strategy is not possible. As we shall see, the Fourier modes, particularly in the high-contrast imaging case, work very well and can be controlled on a mode-by-mode basis.

C. New Approach

In this paper we present predictive Fourier control (PFC), which has been designed to meet our requirements. This method is based on the assumption of frozen flow turbulence. PFC can be viewed as an extension of OFC. As in OFC, closed-loop temporal PSDs will be generated from telemetry. Instead of using the PSDs to optimize the modal gain, PFC will use them to identify components that are caused by frozen flow. The location and power level of these components, plus the noise level, will then be used to produce the predictive control law, which is derived from Kalman filtering.

To develop PFC, we will first discuss the specific advantages of using the Fourier modes as the basis set for wavefront control. Then we will show how the Kalman filter reduces in the case of Fourier modes and frozen flow atmosphere to an easy-to-optimize, stable, and effective predictive controller. The process of estimating layer frequencies in closed-loop and then efficiently solving for a Kalman filter will then be discussed in detail. The computational costs of PFC will be shown to be higher than OFC but still achievable with off-the-shelf hardware. This

analysis will be supplemented by detailed simulations that show the performance improvements possible with PFC in closed-loop AO operation.

2. FOURIER MODES FOR WAVEFRONT CONTROL

When the Fourier basis set is used for modal phase reconstruction and control, the discrete Fourier transform (DFT) is used to calculate the modal coefficients for the slopes and phase. The actuator/WFS grid provides a $N \times N$ domain for the Fourier transforms. On this grid indexed spatially by variables x and y , a complex-valued Fourier mode $[k, l]$ is defined as

$$F_{k,l}[x,y] = \exp\left(\frac{j2\pi[kx + ly]}{N}\right) = \cos\left(\frac{j2\pi[kx + ly]}{N}\right) + j \sin\left(\frac{j2\pi[kx + ly]}{N}\right). \quad (1)$$

Because efficient implementations of the DFT exist (e.g., FFTW3 program [14]) significant computational efficiency is obtained versus traditional matrix-multiply methods. Not only is the wavefront reconstruction efficient, FTR makes the modal coefficients directly available during the phase reconstruction process. Forward DFTs convert the slopes into Fourier modal coefficients. The reconstruction filter is applied. Then an inverse DFT converts back to actuator space. In a typical matrix-based system, the modal coefficients are never explicitly computed. In a system such as Altair, the modal coefficients that are used in gain optimization must be recomputed from the phase estimates, at the cost of another full matrix multiply [2].

Another savings comes from the fact that the phase reconstruction step is accomplished as a discrete step in the reconstruction process. The reconstruction is done by multiplying the slopes by a complex-valued filter. Changing the reconstructor on the fly is very cheap—only the N^2 filter coefficients need be updated. For modal gain optimization this means the modal gains are multiplied into the reconstructor. Again for comparison, in a matrix-based system the entire reconstruction matrix must be recalculated with the modal gains incorporated into it. Use of the Fourier modes provides substantial computational advantages, in the phase reconstruction at each time step and in the modal coefficient-based control-law optimization process.

The Fourier modes are also quite useful as a basis set for describing the PSF and the atmosphere. In high-Strehl situations, the halo of the PSF is well approximated by the spatial PSD of the residual phase [15,16]. The residual error power at a specific Fourier mode directly translates to scattered light at a specific spatial location in the PSF. This framework allows optimization of the PSF by location, which is particularly useful in the ExAO high-contrast imaging scenario.

A. Frozen Flow and Fourier Modes

Fourier modes are particularly useful when it comes to describing the behavior of the atmosphere under the frozen flow model of atmospheric turbulence. In this model,

the phase aberration caused by the atmosphere is assumed to be divided up into distinct layers, each of which translates with a specific velocity vector. The applicability of the frozen flow model is widely assumed, though the number of layers varies. How well the frozen flow hypothesis holds for actual AO operation is of definite interest, but it is beyond the scope of this paper to address that question.

The Fourier basis is particularly amenable to translations—a translation is obtained simply by multiplying the Fourier coefficients by the correct linear-phase complex exponential. Consider the highly idealized case of a pure real-valued Fourier mode phase aberration translating through the atmosphere and hence across the pupil. This mode evolves with time due to translation as

$$\phi(x,y,t) = \cos\left(\frac{2\pi}{Nd}[k(x - v_x t) + l(y - v_y t)]\right), \quad (2)$$

where $[k, l]$ is the Fourier mode frequency, $[v_x, v_y]$ is the velocity vector, and d is the subaperture size in meters. We assume that the interactuator spacing is equivalent to the subaperture size.

The complex-valued Fourier modal coefficient is simply

$$\Phi_{k,l}(t) = \frac{1}{2} \exp\left(-j2\pi t \frac{[kv_x + lv_y]}{Nd}\right). \quad (3)$$

This modal coefficient has a single temporal-frequency component $f_t = -(kv_x + lv_y)/(Nd)$. As the atmosphere evolves, the real (cosine) and imaginary (sine) parts of the modal coefficient alternate.

The sign of this frequency f_t can be either positive or negative. A two-dimensional visualization of the cosine appears as a ripple, with the frequency vector $[k, l]$ giving the direction of propagation, which is orthogonal to the lines of equal value along the ripples. When this cosine translates, the velocity vector of the wind $[v_x, v_y]$ is projected onto this frequency vector. The sign of the frequency f_t indicates which way along this vector the cosine appears to move.

A single layer of frozen flow atmospheric turbulence produces only a single temporal-frequency component at f_t for the complex-valued Fourier modal coefficient. In the past [5] we have chosen to define the Fourier modal basis set as the separate real and imaginary parts of the Fourier coefficient. This was appropriate for OFC, as the integral controller was real valued. However, under the model given by Eq. (3), this is not the best solution. If the Fourier modal coefficient is broken into real-valued sine and cosine components, the resulting temporal PSDs will have even symmetry. Now the power of the layer frequency will show up at f_t and $-f_t$. This will require a second-order model with real-valued coefficients. Though the sine and cosine modes at a specific spatial frequency $[k, l]$ are orthogonal, they are temporally correlated under the model of Eq. (3).

Keeping the modal coefficient in complex form will allow this temporal correlation to be fully captured. A simple first-order model using complex coefficients fully captures the behavior of a Fourier mode. As such, we will now define the Fourier modal basis as $N^2/2 - 2$ complex-valued modes, instead of N^2 real-valued sines and co-

sines. (Note there are still four purely real-valued Fourier modes, including piston and waffle.)

In order to detect this true layer frequency f_l from system telemetry, it is essential that the spatially filtered WFS [17] is used. The spatial filter prevents the aliasing of higher-spatial-frequency modes into the controllable range of the AO system. Each higher-order mode has its own layer frequencies, and aliasing will introduce many different higher-order modes to a given controllable $[k, l]$, producing many false spikes in the temporal PSD.

The equations above assume that a pure cosine is translating across the aperture. In reality the amplitude of the modal coefficient changes through time as different portions of a larger phase screen translate across the aperture. The temporal PSD will not be purely a single frequency. To assess how much this content change affects the temporal PSD, we have used our end-to-end simulation. (For details of the simulation, see Section 8.) The AO system was run in an open-loop observation mode where the modal coefficients of the uncorrected atmosphere were recorded for a 4.096 s run at 2 kHz. The atmospheric model had five layers, which are specified in Table 1.

The temporal PSD of Fourier mode $N=48$, $k=12$, $l=12$ is shown in Fig. 1. (For details on how to calculate an estimate of the temporal PSD from system telemetry, see Section 4.) The five circles on the plot indicate the known layer frequencies as determined by Eq. (3). The pupil diameter D is 8 m. Depending on the layer velocity, between 1.6 and 11.6 distinct portions of each atmospheric layer passed over the aperture during this observation. During this observation, tip and tilt are removed from the slopes and controlled separately. If they are not, tip and tilt project into many Fourier modes and add low-temporal-frequency power to the temporal PSD of many modes. This removal is what will happen during regular AO operation, as the average slope will be removed before phase reconstruction, eliminating tip and tilt from the Fourier modal coefficients. The average slopes will be used to control the tip-tilt mirror directly.

The concentration of power at the layer frequencies can be effectively modeled as suggested by Wiberg [11]. The single-frequency temporal PSD is modeled as an autoregressive process of first order, termed AR(1), with the parameter α a complex number. The magnitude of α controls the strength of the autoregression. The phase of α is $-2\pi(kv_x + lv_y)T/(Nd)$, where T is the sampling interval in seconds. This phase directly encodes the temporal frequency of the Fourier mode for that atmospheric layer. For now we will assume the default $|\alpha|=0.999$, which gives a layer that translates with little change in content. This makes the process $a[t]$ have the standard AR(1) form of

Table 1. Layer Strengths and Velocity Vectors for a Five-Layer Test Case with a Total r_0 of 16 cm

Layer	r_0 (cm)	v (m/s)	angle (deg)
1	38.9	22.7	246
2	44.7	3.28	71
3	45.4	16.6	294
4	38.8	5.89	150
5	43.6	19.8	14

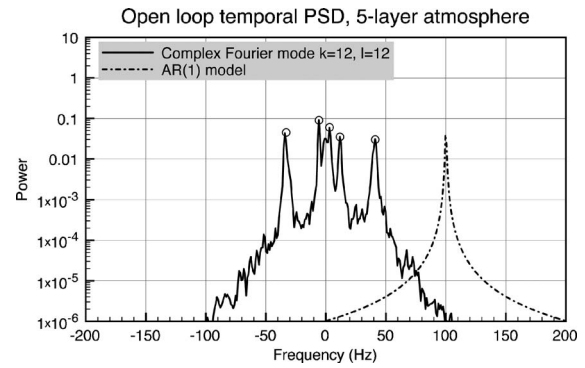


Fig. 1. Temporal PSD of the complex Fourier modal coefficient $k=12$, $l=12$, estimated from spatially filtered, tip-tilt removed open-loop telemetry. The five circles represent the characteristic frequencies due to layer motion. The other temporal-frequency content is due to the fact that the phase content changes slowly as the layer moves across the aperture.

$$a[t] = \alpha a[t-1] + w[t], \quad (4)$$

where it is driven by temporally white noise $w[t]$. The corresponding temporal PSD of this layer component is

$$P_a(\omega) = \frac{\sigma_a^2}{|1 - \alpha \exp(-j\omega)|^2}, \quad (5)$$

where σ_a^2 is the variance of $a[t]$. An illustration of this shape for a higher temporal frequency is also shown in Fig. 1.

For the Fourier modes under frozen flow, the power of the atmosphere is concentrated at a few temporal frequencies. The contribution of each layer to the temporal PSD is well modeled as AR(1). Such a concentration of power does not occur in either the Zernike modal basis or the actuator modal basis under frozen flow. In the Zernike basis with many layers, the temporal PSD is relatively flat until a cutoff frequency, after which it decays with a $f^{-17/3}$ law [18]. In the actuator basis for these simulation parameters, the temporal PSD is dominated by the $f^{-8/3}$ Kolmogorov asymptote [19]. Figure 2 shows representative temporal PSDs from this same data set for an actuator near the center of the aperture and Zernike mode $n=4$, $m=2$. In neither of these cases does a simple model

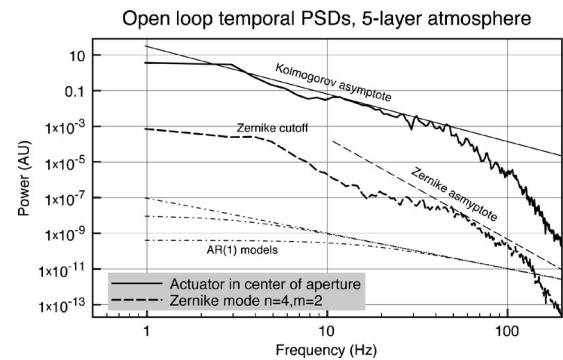


Fig. 2. For comparison with Fig. 1, temporal PSDs for an actuator and for Zernike mode $n=4$, $m=2$ are shown. Simulation agrees with theory as to the power-law structure and cutoff frequencies. Unlike the Fourier modes, under frozen flow neither the actuators nor the Zernike coefficients have concentrated temporal PSDs. (Power normalized to display both curves on the same axis.)

such as AR(1) provide a good approximation. Three AR(1) PSDs (for varying $|\alpha|$) are also shown in the figure to illustrate this.

B. Fourier Modal Independence

As discussed above, one hurdle to previous Kalman formulations is the need to control all the modes (or actuators) together, which leads to a very computationally expensive controller. The Fourier modes enable wavefront control on a mode-by-mode basis because the Fourier modes under frozen flow are nearly uncorrelated both temporally and spatially. This can be determined in two steps using results from random process theory [20]. First, at any instant in time, the Fourier modes are spatially uncorrelated. When a discrete-valued wide-sense stationary process is analyzed with the DFT, the Fourier modal coefficients are uncorrelated. For the atmosphere, the WFS measurements are derived from samples of a continuous process, i.e., atmospheric turbulence. As explained in [20], p. 510, when a continuous-valued random process is represented by Fourier components with a fine resolution in spatial frequency, the Fourier modal coefficients are nearly uncorrelated. Given that the Fourier modal coefficients are nearly uncorrelated, the second step follows: Under frozen flow the temporal cross covariance of any two Fourier modes is nearly zero for all time lags. This is in contrast to a modal basis set like the Zernikes, where the modes are spatially uncorrelated but temporally correlated under frozen flow. For example, a low-spatial-frequency cosine of phase would alternate among focus, tilt, negative focus, and negative tilt as it translated, making those modes correlated.

Use of the Fourier modes provides three major advantages for a predictive controller. First, the modal coefficients are directly available during reconstruction and are generated with a computationally efficient algorithm. Second, under frozen flow the atmospheric phase aberration is compactly represented in the Fourier modes. Third, these Fourier modes are uncorrelated enough that a mode-by-mode optimization of the wavefront can be effectively done, which substantially reduces the basic computational burden from the strategy of optimizing all modes together. All of these properties can be exploited to make a simple and effective predictive control law.

3. PREDICTIVE CONTROL LAW

Before we can design a predictive control law, we must have a model of the AO control system. The AO system will control Fourier modes, and since these modes are orthogonal we can effectively control them independently [5], resulting in a control loop for each Fourier mode. This two-step delay control system for a single mode is shown in block diagram form in Fig. 3. The atmospheric modal coefficient ϕ is controlled in closed-loop. The AO control system only has access to y , the noisy measurement of the residual wavefront at the previous time step. This measurement is actually the result of FTR based on the WFS slopes. It is assumed that the AO system is using an FTR reconstruction filter that fully compensates for the WFS response and the linear-systems effects of the deformable mirror (DM) influence function [5]. The control law $C(z)$

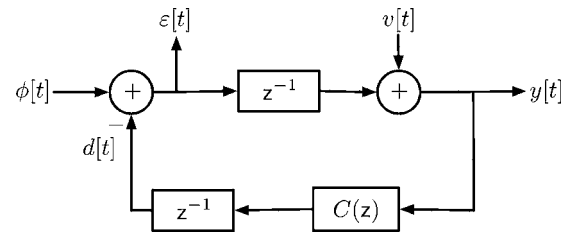


Fig. 3. Block diagram of the two-step delay AO control system. All variables are complex-valued time series of a specific Fourier mode. The input atmosphere ϕ is compensated in a closed loop. The AO system has access only to y , the noisy measurement of the residual error. The control law $C(z)$ will be optimized.

produces the DM commands that will be applied in the next time step. In OFC, this control law is $g/(1-cz^{-1})$, where the modal gain g is optimized for each mode. We normally use an integrator weight c to ensure the stability of the control law itself and reduce mode buildup; many systems use $c=1$. For PFC, we will use the Kalman filtering framework to derive the best $C(z)$ to predict the wavefront.

Given that the temporal PSD of the phase aberration will be composed of distinct components at frequencies corresponding to layer motion, the task of obtaining the predictive controller is reduced to finding a control law that best approximates this shape in the presence of noise. One straightforward strategy is to just treat this as a filter design problem with multiple variables that need to be optimized. In practice, however, this method is computationally expensive. A multivariable optimization must be done, which is generally substantially more expensive than the single-variable optimization done by OFC. In addition, stability is very hard to enforce, and bounding the search domain is quite difficult. For these two reasons, direct filter design and optimization are not viable. The Dessen RLS algorithm was considered but also not chosen for the reasons described above involving convergence and measurement availability.

Instead, the Kalman filter approach has provided the best path. In presenting our Kalman framework, we will build on the formulation of Le Roux. In addition, we will use the autoregressive model for frozen flow that was suggested by Wiberg [11]. Given this model for the evolution and measurement of Fourier modal coefficients, we will derive the Kalman filter and show that it has a particularly simple and easy-to-optimize form.

A. State Space Model

The Kalman filter begins with a state space model of the random process of interest and the observations of it. As noted above, all variables and constants are complex valued. In this situation the matrix transpose is given by superscript T , the matrix conjugation by $*$, and the conjugate transpose by superscript H . The value of a vector at time t is given by $\mathbf{x}[t]$, and the value of an estimated vector at time t based on information up to time $t-1$ is given by $\hat{\mathbf{x}}[t|t-1]$.

The state of the model is given by the vector \mathbf{x} . It contains $L+6$ variables, where L is the number of atmospheric layers. As shown above, each atmospheric layer component is modeled as a first-order autoregressive pro-

cess. Static errors will also exist in the system, and it is essential that they be corrected adequately. To model this, a virtual layer with a real-valued autoregression coefficient α_0 is used. The exact value of $|\alpha_0|$ here determines the rejection of static errors in the final controller and hence will be set for a real system based on the necessary rejection. For the static term and the L atmospheric layers, these variables and parameters will be given by

$$\mathbf{a} = (a_0, a_1, \dots, a_L), \quad (6)$$

$$\mathbf{R} = \text{Diag}(\alpha_0, \alpha_1, \dots, \alpha_L). \quad (7)$$

The power levels of the layers are controlled by the variance of the noise inputs, as given in the covariance matrix

$$\mathbf{P}_w = \text{Diag}(\sigma_{a_0}^2, \sigma_{a_1}^2, \dots, \sigma_{a_L}^2). \quad (8)$$

The state vector has five other elements. Three of these are the modal coefficients of the phase aberration, given by $\phi[t+1]$, $\phi[t]$, and $\phi[t-1]$. The two previous DM commands (in modal coefficient form) $d[t-1]$ and $d[t-2]$ are also necessary. The complete state vector is

$$\mathbf{x}[t] = (\mathbf{a}[t], \phi[t+1], \phi[t], \phi[t-1], d[t-1], d[t-2])^T. \quad (9)$$

The state is updated at each time step as a linear combination of the previous state, plus a constant term and a driving noise input. This update equation has the form $\mathbf{x}[t+1] = \mathbf{A}\mathbf{x}[t] + \mathbf{G}d[t] + \mathbf{B}\mathbf{w}[t]$. The state transition matrix is

$$\mathbf{A} = \begin{bmatrix} \mathbf{R} & \mathbf{0}_{1 \times (L+1)} & \mathbf{0}_{1 \times (L+1)} & \mathbf{0}_{1 \times (L+1)} & \mathbf{0}_{1 \times (L+1)} & \mathbf{0}_{1 \times (L+1)} \\ \mathbf{1}_{(L+1) \times 1} & 0 & 0 & 0 & 0 & 0 \\ \mathbf{0}_{(L+1) \times 1} & 1 & 0 & 0 & 0 & 0 \\ \mathbf{0}_{(L+1) \times 1} & 0 & 1 & 0 & 0 & 0 \\ \mathbf{0}_{(L+1) \times 1} & 0 & 0 & 0 & 0 & 0 \\ \mathbf{0}_{(L+1) \times 1} & 0 & 0 & 0 & 1 & 0 \end{bmatrix}. \quad (10)$$

The $\mathbf{1}_{(L+1) \times 1}$ directly below \mathbf{R} models the fact that the phase at time $t+1$ is just the sum of all the layers at time $t+1$. The three other nonzero elements simply update variables from one time step to the next. The constant term matrix

$$\mathbf{G} = (\mathbf{0}_{(L+1) \times 1}, 0, 0, 0, 1, 0)^T \quad (11)$$

updates the DM variable in the state from the known value. The noise matrix that drives the autoregression is simply identity for the layers and zero elsewhere:

$$\mathbf{B} = \begin{bmatrix} \mathbf{I}_{(L+1) \times (L+1)} \\ \mathbf{0}_{(L+1) \times 5} \end{bmatrix}. \quad (12)$$

The second equation of the state space model is the measurement or observation equation. In our closed-loop control we have access to a single noisy measurement of the residual error. This is captured by the matrix equation $y[t] = \mathbf{C}\mathbf{x}[t] + v[t]$, where the white measurement noise $v[t]$ has variance $\mathbf{P}_v = (\sigma_v^2)$. The measurement matrix that captures this subtraction of the DM shape from the phase aberration is

$$\mathbf{C} = (\mathbf{0}_{(L+1) \times 1}, 0, 0, 1, -1, 0). \quad (13)$$

B. Filter Equations

The Kalman filter is classically implemented in an iterative fashion. At each time step the update and prediction equations are executed, carrying along both the Kalman gain matrix \mathbf{K} and the covariance matrix of the error of estimation \mathbf{P}_s . Implementation of these equations requires many full matrix-matrix multiplies. For this reason, as well as the fact that we do not want the period of substandard performance while the filter converges dur-

ing actual closed-loop operation, the Kalman filter will not be implemented in this iterative fashion directly in the control loop.

Instead, the steady-state solution to the Kalman filter will be used. This steady-state solution is given by the algebraic riccati equation (ARE), which describes the steady-state error covariance matrix \mathbf{P}_s . This equation is

$$\mathbf{P}_s = \mathbf{A}\mathbf{P}_s\mathbf{A}^H + \mathbf{B}\mathbf{P}_w\mathbf{B}^H - \mathbf{A}\mathbf{P}_s\mathbf{C}^H(\mathbf{C}\mathbf{P}_s\mathbf{C}^H + \mathbf{P}_v)^{-1}\mathbf{C}\mathbf{P}_s\mathbf{A}^H. \quad (14)$$

Given this steady-state matrix, the steady-state Kalman gains \mathbf{K}_s can be obtained directly

$$\mathbf{K}_s = \mathbf{P}_s\mathbf{C}^H(\mathbf{C}\mathbf{P}_s\mathbf{C}^H + \mathbf{P}_v)^{-1}. \quad (15)$$

Using this steady-state form, the estimation equation becomes

$$\hat{\mathbf{x}}[t|t] = (\mathbf{I} - \mathbf{K}_s\mathbf{C})\mathbf{A}\hat{\mathbf{x}}[t-1|t-1] + (\mathbf{I} - \mathbf{K}_s\mathbf{C})\mathbf{G}d[t-1] + \mathbf{K}_s y[t]. \quad (16)$$

This update equation can be used in steady state to predict the wavefront.

Another implementation possibility is to derive the control law $C(z)$ from these three equations. This temporal filter $C(z)$ requires only the previous and past residual measurements $y[t]$ and past phase estimates $\hat{\phi}[t]$.

C. General Solution

The solution to the ARE for an arbitrary number of layers L is obtained by using the Z -transform. To do this we will use the elements of \mathbf{P}_s , which are indexed by column then row, e.g., $p_{1,0}$ is the second element in the topmost row

vector. This form is inserted into Eqs. (14) and (15). These are then substituted into Eq. (16) to obtain the recursions for the filter.

For clarity, we will use dummy variables for the state variables. The output of the filter is given by $m[t] = \hat{\phi}[t+1]$, which is the $L+2$ th element of the state vector. The estimate of the previous phase is the $L+3$ th element, given by $e[t] = \hat{\phi}[t]$. Recall that the input of the filter is $y[t]$. For each layer α_i we obtain

$$A_i z^2 = A_i \alpha_i z + \frac{p_{L+3,i}}{p_{L+3,L+3} + \sigma_v^2} (-Ez^{-1} + Mz^{-2} + Y). \quad (17)$$

The output has its equation

$$M = z \sum_{i=0}^L A_i + \frac{p_{L+3,L+1}}{p_{L+3,L+3} + \sigma_v^2} (-Ez^{-1} + Mz^{-2} + Y). \quad (18)$$

For the previous estimate the equation is

$$E = Mz^{-1} + \frac{p_{L+3,L+2}}{p_{L+3,L+3} + \sigma_v^2} (-Ez^{-1} + Mz^{-2} + Y). \quad (19)$$

Examination of the ARE for our model reveals that

$$p_{L+3,L+1} = \sum_{i=0}^L \frac{p_{L+3,i}}{\alpha_i}, \quad (20)$$

$$p_{L+3,L+2} = \sum_{i=0}^L \frac{p_{L+3,i}}{\alpha_i^2}, \quad (21)$$

$$p_{L+3,L+3} = \sum_{i=0}^L \frac{p_{L+3,i}}{\alpha_i^3}. \quad (22)$$

Combining these equations and solving for the control law $M(z)/Y(z)$ produces the result

$$C(z) = \frac{Q^{-1}}{(1 + p_{L+3,L+2} Q^{-1} z^{-1})} \sum_{i=0}^L \frac{\alpha_i^{-1} p_{L+3,i}}{(1 - \alpha_i z^{-1})}, \quad (23)$$

where

$$Q = p_{L+3,L+3} + \sigma_v^2. \quad (24)$$

This equation gives the general form for correcting a static aberration and L layers of atmospheric turbulence. A specific predictor will be named “DC+ L ” to indicate the number of layers that the predictor is correcting, where DC stands for direct current, i.e., the static error.

The structure for implementing a DC+ L predictor is shown in Fig. 4. First, the predictor processes all the layers in parallel. For each layer (and DC) a modified integral controller is used. In a regular integral controller (like OFC) the input is scaled by a gain. The last output is scaled by the integrator weight and is then added to the scaled input. For the predictor, the gain and integrator weight are complex valued. Recall that multiplication by α_i shifts the Fourier mode by the amount of translation in one time step. Each layer goes through an integrator, which shifts the previous output one time step to account for translation. The gain $Q^{-1} \alpha_i^{-1} p_{L+3,i}$ for each mode serves two functions, which become apparent once the values of Q and $p_{L+3,i}$ are numerically solved for in a specific case. This term acts like a modal gain optimizer for each mode, with the scaling given by the relative powers in the atmospheric components and in the WFS noise. The term also applies a translation of the Fourier mode to compensate for two time steps of translation. As such each layer goes through a full translation compensation, shifting the residual input by two time steps and the past output by one, which is the correct amount to compensate for system lags.

The predictor has translation-compensating integral controllers for each layer. The outputs of these are then combined, and fed through a first-order system. This filter is normally a high-pass filter, and it stabilizes all the integrations that were done for the layers. An example error transfer function for a predictor is given in Fig. 5. Each pole α_i produces a notch in frequency response. Each notch corresponds to the prediction of a particular layer frequency.

4. STATE PARAMETER ESTIMATION FROM TELEMETRY

In order for the predictor to be of value, the state model parameters on which it is based must accurately reflect

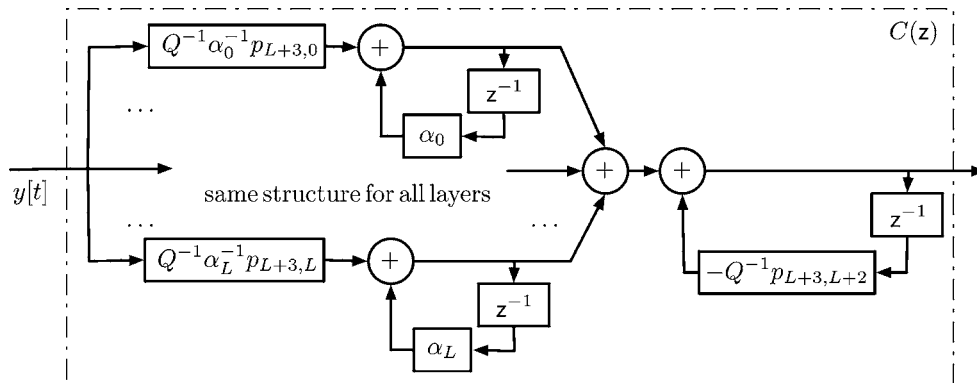


Fig. 4. Structure of implementation of predictive control law $C(z)$. The residual error is integrated separately for each layer, with compensation for each layer's translation. The weighted sum is then high-pass filtered to produce the estimated phase one time step ahead. Use of this structure is necessary for stable implementation of the predictor in the AO control system (see Section 6).

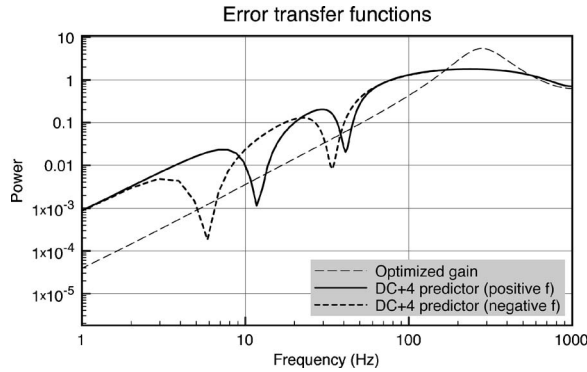


Fig. 5. Magnitude squared of error transfer functions for $k=12$, $l=12$ for OFC and PFC with DC+4 layers. Since the response is different for positive and negative f , the reversed response is also shown. Each identified layer that is predicted results in a notch. In contrast to OFC, PFC has much lower overshoot, but a lower bandwidth.

the characteristics of the atmosphere. When the layer frequencies and power levels are estimated correctly, the temporal PSD specified by the model will be a good fit to the actual data. An example of this is shown in Fig. 6 for Fourier mode $k=12$, $l=12$ in the magnitude $I=6$, $f_{ao}=2000$ Hz case.

The methodology for determining this fit is straightforward, but there are several details of implementation that must be dealt with correctly in order for the parameter estimation to be robust and accurate in closed-loop operation. In essence, the method involves finding the closed-loop temporal PSD, using it to estimate the open-loop PSD, and then identifying peaks.

The process begins with the temporal PSD of the closed-loop residual error, as measured by the WFS. During closed-loop operation with a control law $C_0(z)$ for the given Fourier mode, the AO system records modal coefficients of the residual error. Given atmospheric PSD $A(\omega)$ and the WFS noise PSD $N(\omega)$, the temporal PSD of the residual measurement is

$$P_{cl}(\omega) = [A(\omega) + N(\omega)] \left| \frac{1}{1 + z^{-2}C_0(z)} \right|^2. \quad (25)$$

The estimate of this PSD, $\hat{P}_{cl}(\omega)$, is obtained as in OFC using the averaged modified periodogram technique [21]

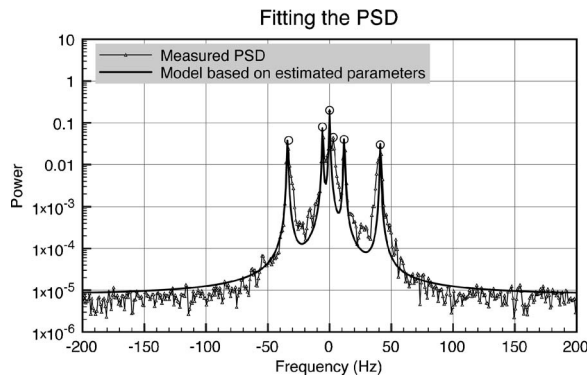


Fig. 6. Estimated joint open-loop temporal PSDs from closed-loop data, along with the Kalman layer model fit to it, DC+4. Layer peak frequencies and power levels, as well as the noise level, are all correctly fit.

on modal coefficient telemetry. Given the windowing function $b[t]$, the modified periodogram is simply the magnitude squared of the DFT of the windowed time series $b[t]x[t]$. The length of the time series in number of samples is s . The averaging comes from taking a different (usually adjacent, but sometimes overlapping) series and averaging the individual periodograms. This provides a consistent (and as the number of averages increases, low-noise) estimate of the true PSD, convolved by the Fourier transform of the windowing function.

This estimate of the closed-loop measurement PSD can be inverted into a joint estimate of the open-loop atmosphere and noise PSDs by multiplication by the known control-law response as

$$\hat{P}_{ol}(\omega) = \hat{P}_{cl}(\omega) |1 + z^{-2}C_0(z)|^2. \quad (26)$$

A. Identification of Peaks

Each layer in the atmosphere produces a peak in $\hat{P}_{ol}(\omega)$. When many layers are present, many peaks will exist. At times it may be necessary to detrend $\hat{P}_{ol}(\omega)$. If a significant trend exists in the PSD, the true peaks will not be found.

Identification of peak locations is done by correlation of the PSD with a reference shape. The correlation function is searched for the strongest peak. The temporal frequency of this peak can be quickly determined to subpixel accuracy by a fast parabolic interpolation of the values of the correlation peak. After a peak is identified, it is windowed out of the correlation function and further peaks are searched for. Peaks are searched for and found until the largest remaining peak is of too low a level. Once the layer frequencies are found, they are used to estimate the power levels of the atmospheric layers.

B. Estimation of Power Levels

The power levels can be directly estimated from the open-loop joint PSD $\hat{P}_{ol}(\omega)$. The WFS noise is temporally white. It dominates at the highest temporal frequencies. An estimate of the noise floor can easily be obtained from the temporal PSD by taking the mean (or median) of the PSD at the highest spatial frequencies. Given a periodogram of length s samples, the estimated noise variance is

$$\hat{\sigma}_v^2 = \text{Mean}[\hat{P}_{ol}(\omega_{hsf})]. \quad (27)$$

The temporal PSD structure of the layer has already been given [see Eq. (5)] and has a simple form that depends on α . Since α will have been estimated already, the power due to this component can be found through projection

$$\hat{\sigma}_a^2 = \frac{\langle P_a(\omega), \hat{P}_{ol}(\omega) \rangle}{\langle P_a(\omega), P_a(\omega) \rangle}. \quad (28)$$

Because $P_a(\omega)$ drops off rapidly, this inner product can be computed over a small set of sampled values of ω .

C. Dealing with Numerical Limits of Periodogram

The averaged modified periodogram technique for temporal PSD estimation from time series data is a very powerful technique. However, there are numerical limitations

that must be accounted for or else the PSD estimates will cause difficulties in the layer-finding process.

The lowest temporal frequency (excluding $f=0$) that can be estimated depends on the length of the time series of data that is used. For an interval of s samples at frame rate f_{ao} , the lowest frequency measured is f_{ao}/s . For example, if an interval of 1024 samples at 2000 Hz (0.512 s) is used for each periodogram, the lowest temporal frequency that is measured is 1.95 Hz. As such, the detection of slowly moving atmospheric layers drives s to be longer. In contrast to this, the desire to use many averages a and to update as atmospheric conditions change drives s to shorter lengths.

Because the model already assumes a static aberration (which is at frequency 0, or DC), very low temporal frequency layer components will be well corrected anyway, obviating the need to search for them. Since we will always correct DC, the layer-finding algorithm will be constrained to not search for peaks in the correlation function below a lower limit. For $F=2000$ we set this limit to 2 Hz.

The “modified” portion of the method derives its name from the fact that the time series of data are multiplied by a windowing function before the DFT is taken. This windowing prevents leakage in the temporal PSD and allows accurate measurement of large variations (greater than 100 times) in power between neighboring frequencies. This is very important in our application, as peaks due to uncorrected atmosphere and notches due to the performance of a predictor must be accurately estimated.

However, the windowing function also has a broad central core in the frequency domain, which smears out the peaks and makes them wider. This will reduce the temporal resolution between neighboring peaks and fill in notches caused by a predictor. To minimize this loss of resolution while maximizing the dynamic range, we have chosen the Hanning window, where $b[t]=0.5-0.5\cos(2\pi t/(s-1))$. It has an approximate main lobe width of $8\pi/s$ rad in the ω domain.

This lobe width sets the minimum size for the mask that is used to block out a peak that has just been found in the correlation function. Even a pure temporal-frequency spike will have width $8\pi/s$ in the periodogram because of the window. The entirety of this peak must be masked so as to not find a “peak” nearby due to the smearing. A more important effect of this smearing occurs when the AO system is already using a predictor.

D. Tracking of Layers with Predictor in Place

As already shown in Fig. 5, the error transfer function of the predictor has notches at the layer frequencies. This means that during closed-loop operation with the predictor, the PSD of the residual $\hat{P}_{ol}(\omega)$ may have no peaks and potentially will also have notches. If these notches exist, they must not be obscured by artifacts in the periodogram.

This is especially important because the closed-loop residual is divided by the error transfer function to produce $\hat{P}_{ol}(\omega)$. Each notch in the transfer function results in a multiplication of $\hat{P}_{cl}(\omega)$ by a very large number. The worst consequence of this is that if the peak location changes

during operation, the inversion process will result in a large and false peak in $\hat{P}_{ol}(\omega)$.

A handy way to mitigate this is to limit the depth of the notch. The depth of the notch is set by $|\alpha|$ and the effective gain for that layer, not by the angle of α . With our framework that DC is always well corrected, this means that a layer at a high temporal frequency can produce a notch well over 10^3 as deep as the surrounding response. It is not clear that such a great depth is truly needed when correcting the atmosphere. Instead, the magnitude of α will be modified as the layer frequency increases. If ω_0 is the frequency of the layer, then the magnitude of the autoregression coefficient is set to be

$$|\alpha| = 1 - |\omega_0|/20. \quad (29)$$

In addition, the maximum value of $|\alpha|$ is set at 0.999. This ensures that the default depth of the notch will be constant relative to the surrounding response, independent of the temporal frequency of the layer. The actual depth of the notch will be adjusted in filter optimization based on the power levels, which ensures that a powerful layer component will still be corrected very well. So while the angle of α is estimated from the telemetry, the magnitude is not estimated. It is instead set so as to ensure robust filter performance.

Limiting the depth of the notch by setting the magnitude of α as a function of layer frequency allows robust peak detection during continuous closed-loop operation with the predictor. However, there are other possible strategies for tracking layers during predictor usage. If the attenuation provided by the sliding- $|\alpha|$ scheme is not sufficient, another approach could be taken. For example, the predictor can be blinked off after an interval of operation if the number of found layers becomes too large. A single iteration of $(a+2)s$ time steps will allow for a clean temporal PSD estimate and a fresh and accurate estimation of the layer locations. The system would, however, spend a fraction of time without prediction and at lower performance levels. In some observing scenarios this effect could be ameliorated by synchronizing this on-off behavior with science detector readout. Another possibility would be to also track the temporal PSDs of the DM commands, which avoids somewhat the false peak problem. This option may require substantial extra computation, however.

E. Layer Finding for Each Mode and All Modes

The end result for a specific Fourier mode is that algorithms find all significant peaks in the temporal PSD estimate $\hat{P}_{ol}(\omega)$. For each peak the temporal frequency and the power level are identified. For a given Fourier mode this is done without a prior assumption of how many layers exist. There will not necessarily be five found peaks if five layers exist in the atmosphere. If the layer frequencies are too low, they will not be detected due to the limitations discussed above. Since these low-temporal-frequency components will be very well corrected by the DC term in the predictor, this is not a problem. For many modes the different layers give rise to layer frequencies that are very close together. These will be detected as a single peak, again due to limitations discussed above. But

again, this is not a problem, as the notch in the error transfer function (see Fig. 5) will correct all the layers within a range of 5 to 10 Hz.

To illustrate this, consider the five-layer atmosphere as specified in Table 1. For Fourier mode $k=3$, $l=0$, the layer frequencies are 3.1, -0.37, -2.3, 1.8, and -6.6 Hz. For this Fourier mode only the two components at 3.1 and -6.6 Hz would be detected. The three other components have too low a temporal frequency. For Fourier mode $k=8$, $l=26$ the layer frequencies are -44.1, 6.8, -44.3, 12.0, and -5.5 Hz. For this mode the first and third layers produce very similar layer frequencies: -44.1 and -44.3 Hz. Four layer frequencies in total will be detected, but the components of all five layers will be corrected.

This means that for a given atmosphere different Fourier modes will end up operating with different DC+ L predictors. It is worth emphasizing that as currently formulated there is no attempt to use information from all of the Fourier modes to globally determine the exact layers in the atmosphere. Such an algorithm would attempt to fit linear terms across spatial-frequency space to the layer frequencies, in essence finding the shift filters for each layer. Though there may be some advantages to this, the problem of developing a robust algorithm to do this is left to further research.

F. Summary

The basic methods used in estimation of the atmospheric parameters are quite simple. However, several numerical issues must be overcome to ensure robust performance. In particular, the main lobe width of the periodogram cannot be ignored, and the magnitude of α must be modified as layer frequencies increase to avoid detection of false peaks.

The use of a temporal PSD method is motivated by the fact that PFC should perform at least as well as OFC. If difficulty is encountered in the process of finding layers, the control computer can revert back to OFC, since the necessary temporal PSDs are already available.

Though the algorithms presented detect layers very well, in future research it will be worth exploring whether other model fitting approaches, such as an efficient non-linear fitting to the PSD, provide any advantages in performance. Another possible modification is the incorporation of telemetry data from the DM commands to obtain the open-loop PSDs without division by the transfer function.

5. EFFICIENT FILTER COEFFICIENT CALCULATION

Once the state parameters are estimated, the filter coefficients must be calculated. There is not an analytic formula for the necessary coefficients used in Eq. (23). There exists substantial literature on efficient solving techniques for the Kalman error covariance matrix. From the many options, we have chosen the doubling algorithm (see [22] for a complete description). This algorithm is an iterative method for directly calculating the coefficients of the error covariance matrix. In this case subscripts will be used to denote iterations of the algorithm. Adapted for our

Table 2. Iterations for Solution of ARE^a

L	Average	Trials Requiring 14 Iterations (%)
1	9.4	0.02
2	10.2	0.16
3	10.7	0.36
4	11.1	0.64
5	11.3	0.88
6	11.5	1.3
7	11.7	1.4
8	11.8	1.8
9	11.9	2.3
10	12.0	3.1

^aNumber of iterations that were necessary for convergence of the solution of ARE. Results determined with Monte Carlo simulation for many different sets of layer frequencies. For each number of layers L , the average number of iterations for convergence is given. The percentage of trials that took 14 iterations (the maximum necessary) is given.

state space model and conditions, it uses three matrices, as defined as follows, in terms of the state space model matrices:

$$\mathbf{F}_0 = \mathbf{A}, \quad (30)$$

$$\mathbf{P}_0 = \mathbf{B} \mathbf{P}_w \mathbf{B}^H, \quad (31)$$

$$\mathbf{O}_0 = \mathbf{C}^H \mathbf{P}_v^{-1} \mathbf{C}. \quad (32)$$

Given these initial values, the three matrices are updated as

$$\mathbf{T}_i = (\mathbf{I} + \mathbf{P}_i \mathbf{O}_i)^{-1}, \quad (33)$$

$$\mathbf{F}_{i+1} = \mathbf{F}_i \mathbf{T}_i \mathbf{F}_i, \quad (34)$$

$$\mathbf{P}_{i+1} = \mathbf{P}_i + \mathbf{F}_i \mathbf{P}_i \mathbf{T}_i^H \mathbf{F}_i^H, \quad (35)$$

$$\mathbf{O}_{i+1} = \mathbf{O}_i + \mathbf{F}_i^H \mathbf{O}_i \mathbf{T}_i \mathbf{F}_i. \quad (36)$$

The matrix \mathbf{P}_i converges to the steady-state error covariance matrix \mathbf{P}_s , which satisfies the ARE. Then the elements necessary for the filter [see Eq. (23)] are simply extracted and used.

The three matrices are all $(L+6) \times (L+6)$, so this method on the surface may not seem particularly efficient. It does, however, converge very rapidly. Our stopping criterion for the algorithm is that all entries of interest for the filter ($P_{L+3,*}$) must have converged to 0.001 relative error compared with the last iteration. A Monte Carlo analysis of the convergence behavior was conducted for $L=1$ to $L=10$. The layer frequencies and power levels were chosen from a distribution based on simulation results; 5000 trials were conducted for each L . By $L=10$, the average number of iterations necessary reaches 12, but no single case (of the 50,000 that were run) took more than 14 iterations. Full results are given in Table 2.

6. STABILITY

A potential pitfall of using complicated temporal filters is that system stability will be compromised. In PFC this

can arise from either inherent instability in the predictor or from numerical errors in the implementation of the predictor in the AO control system computer.

A. Inherent Stability of Controller

The controller of the system must be stable, as determined by the locations of the poles of the closed-loop error transfer function. If all of the poles are inside the unit circle, the system will be bounded input, bounded output stable. For OFC in this two-step delay system, this sets the limits on the range of gains to be $0 \leq g < 1$.

For the Kalman filter, there is a more fundamental result ensuring the stability (see [22] for a thorough discussion). In our formulation, as long as the matrix \mathbf{A} has eigenvalues inside the unit circle, the steady-state Kalman filter that satisfies the ARE will be stable. Since $\mathbf{A} - \lambda \mathbf{I}$ is lower triangular, it is simple to determine that the eigenvalues of \mathbf{A} are 0 (five times) and the layer autoregression coefficients α_0 to α_L . Since the α 's are constrained to be inside the unit circle, the stability of the control loop is guaranteed.

Even if the system is stable as defined above, during operation nonlinearities and perturbations in the system can make it become unstable. The robustness of the system is given by an analysis of its margins. The gain margin and phase margin reflect, respectively, how much external perturbation of gain or phase can be introduced into the feedback loop before the system becomes unstable. A typical phase margin requirement is 45 deg, which means that the system will remain stable even if the feedback loop is advanced in phase by $\pi/4$. The gain and phase margins are analyzed by examining the behavior of $1 + z^{-2}C(z)$. (They can also be analyzed in the continuous Laplace transform domain, but the results are equivalent.) For the integral controller of OFC, the gain and phase margins shrink as the gain is increased. In high-SNR situations the maximum gain must be limited, or the margins become too small. For OFC, we have set a maximum gain of 0.65, which has a phase margin of 33 deg.

In contrast, the predictor has satisfactory gain and phase margins for high SNRs, without a necessary constraint on the filter. For a wide range of layer frequencies and high SNRs, the gain margin is at least 1.7 and the phase margin is at least 45 deg. A graphical interpretation is straightforward given a plot of magnitude versus phase, as shown in Fig. 7 for a DC+4 predictor at high SNR of 40.1 and a DC+3 predictor at low SNR of 1.01. In the high-SNR case, the DC+4 predictor has a gain margin of 2.63 and a phase margin of 54 deg. In the low-SNR case, the gain margin is 13.7 and the phase margin is 82 deg. Each of the loops in the plot reveals the presence of a pole in the predictor that removes a specific layer.

Another consideration is how many layers can be corrected while preserving the stability of the predictive controller. Since the stability is set by the atmospheric model, as long as the layer parameters all have magnitude less than 1, the controller will be stable. So an arbitrary number of layers can be stably corrected.

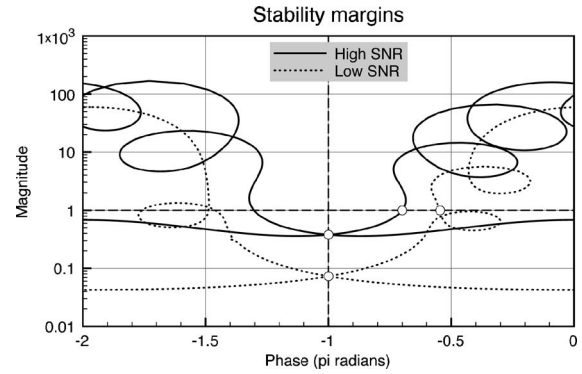


Fig. 7. Magnitude-phase plot for two predictors at high and low SNRs. The phase and gain margins are indicated by the circles. Even at very high SNRs, the predictors exhibit robust phase margins greater than 45 deg.

B. Numerical Implementation and Stability

The form in which the predictor is implemented in the AO control system also affects system stability. This is because the numerical representation of the coefficients (or of the poles and zeros) in the computer is only approximate. This is particularly important given the two possible representations of the control law. The general predictor [Eq. (23)] is written as the product and sum of first-order terms. The poles of the predictor due to the layers (α_i 's) are clustered, especially when the layer frequencies are low.

The predictor can also be written in what is called direct form, as an explicit infinite impulse response (IIR) filter

$$C(z) = \frac{\sum_{i=0}^L n_i z^{-i}}{1 - \sum_{i=1}^{L+2} d_i z^{-i}}, \quad (37)$$

where the coefficients d_i and n_i are determined from the general equation. It has been shown (see [21] for a detailed discussion) that filters with clustered poles are highly susceptible to quantization error when represented in direct form. This means that the very small errors of the values of d_i and n_i translate into large errors in α_i , which can easily make the system unstable. In our simulation, these quantization errors can cause instability in the closed-loop system in PFC for cases as small as DC + 2 when the layer frequencies are close together. Note that the predictor of Dessenne is explicitly derived in direct form, making it susceptible to these errors as well, if it were used to derive the predictor for many layers.

The best way to avoid the problem of instability caused by numerical quantization of filter coefficients is to implement the predictor not in direct form, but in the form derived from pole-zero representation. This representation has already been given, as shown in Fig. 4. It is essential for system stability that PFC be implemented in this manner, as opposed to direct form.

7. COMPUTATIONAL COSTS

The computational costs of PFC can be evaluated in terms of some fundamental system parameters. The grid size on which the phase reconstruction is done is $N \times N$, where N is typically a small number more than the number of subapertures across the pupil and a desirable size for a fast DFT. For example, in the GPI, there are 44 subapertures across the pupil and $N=48$. In this complex modal form, there are $N^2/2 - 2$ complex-valued Fourier modes to control, along with two real-valued Fourier modes and tip and tilt. For simplicity we will do the analysis as if $N^2/2$ complex-valued modes are controlled.

We assume that complex-valued signals are stored as real and imaginary parts and that each complex add takes 2 FLOPs and each multiply takes 6 FLOPs. For DFTs, we assume that when the time/spatial domain signal is real valued, the cost is $2.5K \log K$ FLOPs, where K is the total number of elements. When the signal is complex valued, the cost is $5K \log K$ FLOPs [14]. Figure 8 presents in graphical form the steps and timings of the necessary computation.

A. Operations Done Each Time Step

At each time step the phase is reconstructed and the predictive control law is applied. The reconstruction is dominated by the two-dimensional DFTs. Including reconstruction filter application, the entire reconstruction process takes $15N^2 \log N + 10.8N^2$ FLOPs.

The application of the predictive controller is done in the frequency domain to the complex-valued modal coefficients. This is done before the inverse DFT of FTR. To ensure numerical accuracy, the predictor is implemented as shown in Fig. 4. This requires storage of $L+1$ intermediate signals from the previous iteration, as well as the last output. Each mode requires $2L+3$ complex multiplies and $2L+2$ adds. For all modes, the total cost is $(8L+11)N^2$ FLOPs.

The total computational burden for phase reconstruction and temporal filtering at each time step is then $15N^2 \log N + (8L+21.8)N^2$ FLOPs per time step.

B. Temporal PSD Estimation

The temporal PSD estimation is very similar to that in OFC. Complex-valued modal coefficients are used, of course, but this means that only half as many PSDs need to be calculated (at twice the individual cost). The main difference stems from the long baseline necessary. Because averaging is needed for the periodogram to have reasonable SNR and to approach the correct answer, the method of half-overlapped window will be used to en-

hance the performance [21]. As given in Section 4, intervals of s samples will be used. Instead of taking the starting entries at time steps $0, s, 2s, \dots$, the starting points are spaced at $s/2$ such that periodogram windows begin at $0, s/2, s, 3s/2, \dots$.

At each time step, the complex modal coefficients are stored in a buffer. Initially s steps must be accumulated, but thereafter periodograms begin every $s/2$ steps. We assume that each periodogram must be completed in $s/2$ time steps. Accounting for the windowing, DFTs, and accumulations, the amortized cost is $(6+5 \log s)N^2$ FLOPs per time step.

C. State Parameter Estimation and Filter Coefficient Calculation

The algorithms described above to find the layer parameters from the real-valued temporal PSDs take much less computation than the Riccati solve. As such, they will not be gone over in detail. The total cost of finding the state parameters for a Fourier mode depends on the number of layers that are found and the number of samples in the PSD.

The entries in the steady-state error covariance matrix \mathbf{P}_s are calculated with the matrix-multiply doubling algorithm. Each iteration involves calculating a matrix inverse and executing many matrix multiplies. The costs for the matrix multiplies and additions (all matrices are complex valued) is $72(L+6)^3 - 12(L+6)^2$ FLOPs. The matrix inversion is accomplished using the method of LU (lower-upper triangular) decomposition. The LU algorithm requires multiplies, additions, comparisons, and absolute values. Assuming each comparison takes 2 FLOPs and each absolute value 4 FLOPs, the matrix inversion costs $26(L+6)^3/3 + 15(L+6)^2 + 25(L+6)/3 - 48$. The total cost of solving for the error covariance matrix using I_t iterations is approximately $80.67I_t(L+6)^3$.

To fit in the timeline shown in Fig. 8, an interval of $3s/2$ steps is available for the optimization. The matrix solving dominates the filter coefficients, producing an amortized cost per time step of approximately $80.67N^2I_t(L+6)^3/(3s)$ FLOPs.

D. Housekeeping

In addition to these time critical tasks, there are a small number of things that need to be done but that are inexpensive. These include generating the transfer function used to convert from $\hat{P}_{cl}(\omega)$ to $\hat{P}_{ol}(\omega)$ and $\hat{P}_{leaky}(\omega)$.

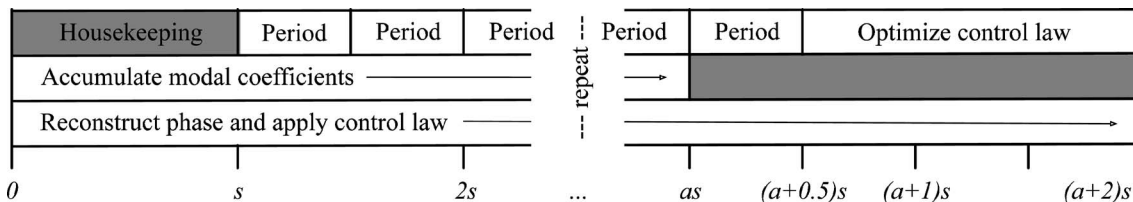


Fig. 8. Timeline of the PFC process. At every time step the phase is reconstructed and the control law is applied to each mode. Accumulation of modal coefficients into buffers begins immediately. After s time steps, the half-overlapped windowed periodograms begin. After $(a+0.5)s$ steps, periodogram estimation is complete. The optimization process then occurs to produce the new predictive control law for each mode. This involves finding layer frequencies and strengths and solving the ARE.

E. Total Amortized Costs

At every time step the reconstruction and control-law application must occur. This computational load is $15N^2 \log N + (8L + 21.8)N^2$. At most time steps, some part of the optimization process is being executed. The periodogram estimation cost is $(6 + 5 \log s)N^2$. The layer finding and filter determination is dominated for reasonable parameters by the solving for the covariance matrix at a cost of $80.67N^2 I_t (L + 6)^3 / (3s)$. The peak load on the processor per second in FLOPs is then $f_{ao} [15N^2 \log N + (8L + 21.8)N^2]$ plus the maximum of $f_{ao}(6 + 5 \log s)N^2$ and $f_{ao} 80.67N^2 I_t (L + 6)^3 / (3s)$.

For a system such as GPI, the amount of computation is still reasonable. The number of allowed iterations I_t for the covariance matrix solve plays a major role in computational costs. Based on the results of Section 5 and Table 2 an upper limit of $I = 15$ is assumed. This level provides a computational cushion in case convergence takes longer than average for many modes. For the case of $f_{ao} = 2000$ and $s = 2048$, the peak load for DC+5 GPI is a challenging system. Assuming $N = 24$, $f_{ao} = 1000$, and $s = 1024$, DC+5 prediction is possible in under 0.4 GFLOPs/s, making PFC on a next-generation general-purpose AO system easily implementable.

8. SIMULATIONS

Our end-to-end Fourier optics AO simulation code has been described in detail elsewhere; please consult [6] for a more in-depth discussion. The atmospheric phase aberration is made up of independent layers. For each layer a large enough screen to contain the necessary translations for the exposure time is created using the spectral factor method [23]. At each time step, each layer is shifted with subpixel accuracy to the appropriate position over the aperture. The total phase aberration is obtained by summation of the layers. No Fresnel propagation is conducted, so scintillation is ignored.

All of the predictors that are used in the simulation are generated using the algorithms described in Sections 4 and 5. Telemetry from a 4.096 s run of the simulation is used for this. This provides a way to test not only the ability of the predictors to correct frozen flow but also the ability of the parameter estimation algorithms to find the layers.

9. PERFORMANCE: HIGH-CONTRAST IMAGING

As mentioned above, the development of PFC has been motivated by the high-contrast imaging requirements of GPI. In this high-order, high-Strehl AO system, uncorrected atmosphere is the dominant term in the residual error budget. The non-common-path error in GPI is much lower than in present systems and hence is not limiting.

For GPI, $N = 48$, $D/d = 44$, and $D = 8$ m. The WFS SNR is determined by photometry, WFS wavelength range, and system throughput. The natural guide star (NGS) magnitude I_{mag} is calibrated to Vega for a band 300 nm wide centered at 806 nm. The WFS band is 700 to 900 nm. For a given I_{mag} and AO system frame rate f_{ao} , the number of received electrons on the WFS CCD per subaperture per

time step is $E = 1.4626 f_{ao}^{-1} 10^{8 - I_{mag}/2.5}$. The WFS CCD RMS read noise in electrons per pixel is 8. Given the 2×2 quad-cell WFS configuration, the resulting WFS SNR is $E(E + 256)^{-1/2}$.

For each NGS I_{mag} from 0 to 9, the simulation was run with three different controllers: the “uniform gain” case with a gain of 0.3 for all modes, the “optimized gain” case with OFC based on 4.096 s of telemetry, and the “predictor” case with DC+4 predictors based on the same telemetry. For all runs at different magnitude, the atmospheric realization was kept the same to allow direct comparison of speckle intensities. That particular atmospheric realization has the same parameters but is a different realization than the one that was used to generate the 4.096 s of telemetry.

A. PSF Intensity in the Dark Hole

For the specific case of $I_{mag} = 6$, 12 different atmospheric realizations were used to generate a composite 12 s exposure. This helps average out the speckles and produces smoother PSF and PSD estimates. Figure 9 shows the three PSFs (uniform gain, optimized gain, and predictor) for this case on the same log color scale from $1e-4$ to $1e-6$ intensity. These PSFs were generated using a gray-scale Blackman apodization function to suppress diffraction, which has an effect similar to that of the coronagraph in GPI. The Blackman apodization is applied to the field in the pupil plane.

The quality of correction is measured in terms of the in-band error, which is the error in the range of spatial frequencies that the system can correct. For the uniform gain case, the in-band root-mean-square (RMS) error is 37.2 nm. For the optimized gain case it is 29.1, and for the Predictor it is 21.0 nm. This reduction in RMS error may not seem significant, but the difference in the PSFs are clearly apparent. Both the uniform gain and optimized gain cases have an ellipsoidal shape due to uncorrected atmospheric errors from the five layers. This shape is a composite the five differently oriented “butterfly” or “hourglass” shapes that each layer gives rise to. Use of PFC restores the radial symmetry of the PSF. The faint radial spikes in the predictor PSF are caused when two layers produce layer frequencies that are very close to each other. Since the implementation of PFC cannot differentiate between very closely spaced layer frequencies, neither layer is corrected quite as well as it could be. Slices through the PSFs (as shown in Fig. 10) quantify the improvement. Use of OFC over no optimization reduces the PSF intensity along this slice by 1.8 times. Use of PFC further improves matters, with 3 times less intensity than OFC.

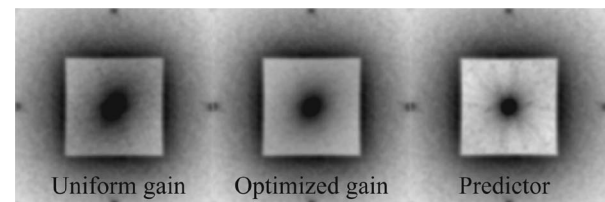


Fig. 9. Composite 12 s PSFs for $I_{mag} = 6$, 2000 Hz GPI case, five-layer atmosphere. White-black log color scale ranges from $1e-4$ to $1e-6$.

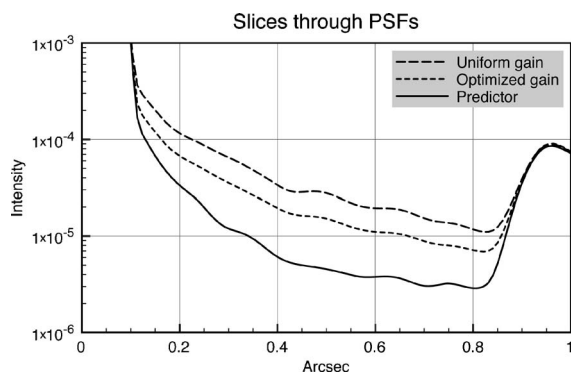


Fig. 10. Slices along the y axis through the center of the three PSFs shown in Fig. 9. The PSF intensity from the uniform gain controller is 1.8 times that provided by OFC. This OFC intensity, in turn, is 3 times higher than that provided by PFC. Use of the predictor significantly reduces scattered light in the PSF.

The predictor is able to correct the layers well regardless of temporal frequency, as shown by the notches in the error transfer function. Figure 5 shows the error transfer functions for Fourier mode $k=12$, $l=12$ for the DC+4 predictor and for the optimized integrator of OFC. This clearly shows why PFC does a better job correcting the atmospheric components than OFC does. The only way that OFC can better correct a high-temporal-frequency phase error is by changing the level of the entire transfer function. In order to better compensate the layer components, the gain in OFC is raised. This lowers the 20 dB/decade rejection, but at the price of increased overshoot in the signal and noise transfer functions. This noise propagation is shown in Fig. 11, along with that for PFC. PFC is better able to trade off signal and noise because the signal correction is very local. Given these transfer functions for this Fourier mode, PFC has 2.77 times less scattered light (and residual error power) than OFC does for the case shown above.

However, the 20 dB/decade rejection of this particular DC+4 predictor is about 30 times less than that exhibited by OFC. This is because the power level at DC is low compared with the total power in the layers and in the WFS noise. The amount of loss in rejection varies greatly with Fourier mode and atmosphere. Many modes have nearly equivalent rejection to OFC. This large difference in 20 dB/decade rejection could become a problem for PFC if

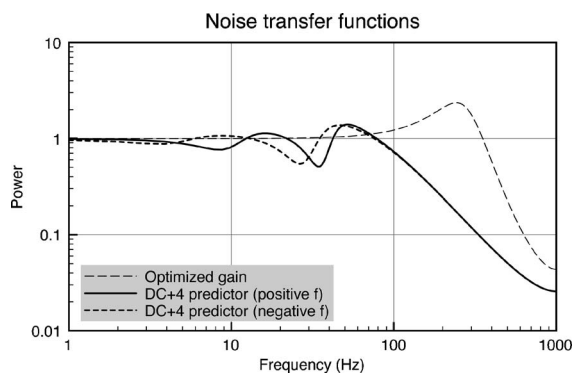


Fig. 11. Magnitude squared of noise transfer functions for $k=12$, $l=12$ for OFC and PFC with DC+4 layers. PFC provides better correction of the layer components (see Fig. 5) and propagates less WFS noise than OFC does.

the atmosphere changes too quickly for the predictor to stay up to date. If the wind velocity vectors change such that the layer frequencies of this mode are outside the notches, PFC will perform significantly worse than OFC in correcting the atmosphere.

B. Performance as a Function of Guide Star Magnitude

PFC has the potential to improve the quality of the observation of a particular star. But in the context of a survey of many stars, it has the potential to expand the population that can be examined. As such, improvement in performance as a function of NGS magnitude is worth examining.

The primary error metric is residual MSE in the controllable spatial-frequency band. Figure 12 shows the performance of the three different controllers in terms of residual MSE. For the various NGS magnitudes, the results are plotted as a function of the WFS SNR. Optimized gain is always as good as or better than uniform gain, because of the natural range in optimal gains as a function of spatial frequency that is due to atmospheric content. At bright stars OFC provides the most improvement, because the default modal gain of 0.3 is much too conservative in those cases. This improvement, however, is limited by the maximum modal gain constraint to ensure an adequate phase margin.

This total in-band error metric shown in Fig. 12 is dominated by the fact that most of the residual power is at very low spatial frequencies. This is simply because most of the power in the atmosphere (as given by the Kolmogorov power law) is at low spatial frequencies. This poses a problem for PFC, however, since these lowest spatial frequencies have correspondingly low layer temporal frequencies and cannot be notched out. Outside the inner working distance (IWD), the layer frequencies are higher, and PFC is able to correct much better than OFC can.

To assess how much the low spatial frequencies dominate the residual error, the spatial frequencies that scatter light inside the IWD are excluded from the total. This is a valid approach in the planet detection case, as a system with a coronagraph will have a minimum IWD set by the coronagraph design. Planets will not be detectable inside the IWD, so we are not as concerned with error there. For the GPI simulation, the IWD was assumed to be $5\lambda/D$. Figure 13 shows the residual MSE in the discovery region in the dark hole. The improvement provided by

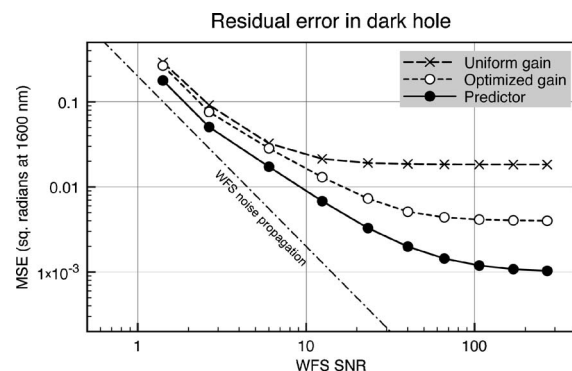


Fig. 12. Residual MSE in the controllable range of spatial frequencies for the three different control laws. Data points from simulations at magnitudes $I_{\text{mag}}=0$ to $I_{\text{mag}}=9$.

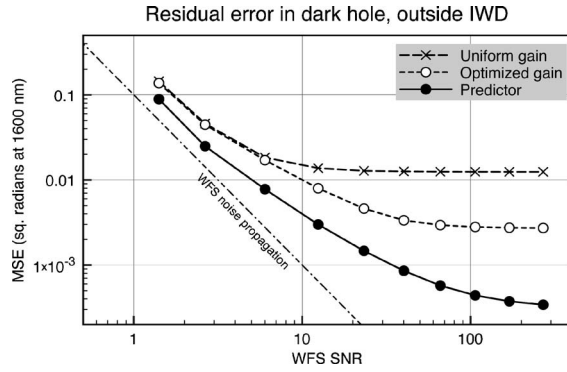


Fig. 13. Predictor improvement is particularly noticeable outside the inner working distance (IWD) of $5\lambda/D$. (Compare with Fig. 12.).

prediction is now greater than before. In this specific atmospheric case, the performance of the predictor at $I_{\text{mag}} = 7$ (SNR of 6.01) is the same as optimized gain for $I_{\text{mag}} = 6$ (SNR of 12.44). The high-SNR performance shows that the benefit of PFC continues, as it does not hit a stability limitation like OFC.

However, most scientifically interesting targets for GPI will be at the lower end of the SNR range. For dimmer stars and lower SNRs, the performance advantage decreases slightly. There are two separate reasons for this. First, as the WFS SNR is lowered, it becomes more difficult to detect slowly moving layers. This is due to a combination of the increased noise in the temporal PSD estimates and the inversion by $C(z)$ at low temporal frequencies. In the $I_{\text{mag}} = 4$ case, PFC detects at least three layers for 86% for the Fourier modes, and it finds four layers for 32% of them. For the dim NGS of $I_{\text{mag}} = 9$, three layers are detected for only 53% of the modes, and four layers for just 14%. This means that at low SNRs, fewer low-temporal-frequency-layer components are detected and corrected, limiting the performance improvement.

The second reason is that as the noise level increases, the notches in the error transfer function are not as relatively deep. Consider the case of $k = 12$, $l = 12$ and specific layer frequency of 11.7 Hz. When $I_{\text{mag}} = 4$, the DC+4 predictor notch is 100 times deeper than the surrounding response. When $I_{\text{mag}} = 9$, the notch is still there, but it is only 10 times deeper than the surrounding response. Whereas the DC rejection of the $I_{\text{mag}} = 9$ predictor for this mode is 7 times less strong, the rejection of this particular layer is 15 times less strong. This could be ameliorated by modification of $|\alpha|$, but such a refinement is left for further research.

C. Importance of Model Fit

The better the predictor's model of the atmosphere, the better it will do at compensating the dynamic atmosphere. If the model is not a good one, performance will not improve from the OFC level and may very well be worse.

For the cases discussed above, the DC+4 predictor was used. In this five-layer case, no more than four layer frequencies were ever found for a given Fourier mode. This is a combination of the fact that two of the layers are moving very slowly and that due to the varying angles of the

wind vectors, in many regions the layer frequencies from different layers are very close together. Since no more than four layers were found, there is no advantage in using a predictor of order greater than DC+4.

As a check, predictors were created with lower orders, e.g., DC+1. In the case of DC+1, performance with PFC was worse than performance with OFC. At DC+2, performance was slightly better than OFC, with performance improving through DC+3 to DC+4. As the predictor order is increased, the model quality improves and so does the correction of the atmosphere. For the design of an AO system with PFC, this indicates that the computational capacity of the controller should be greater than the number of layers that are usually detectable in the atmosphere.

10. CONCLUSIONS AND FURTHER RESEARCH DIRECTIONS

PFC is enabled by the fast modal wavefront reconstruction of FTR. FTR not only makes modal coefficients directly available for control, but its Fourier modal set is particularly useful for describing the dynamic wavefront under the frozen flow assumption. This compactness of the atmospheric aberration in the temporal PSD results in a simple predictive controller that can stably correct for multiple atmospheric layers per Fourier mode. Though care must be taken, the atmospheric parameters that identify layers can be accurately estimated during closed-loop operation. These parameters are used in an iterative, computationally efficient algorithm to determine the predictive control law. All of this can be done with a substantial, but not unreasonable, computational burden. For GPI, correction of static errors plus five atmospheric layers requires under 2 GFLOPs/s.

As formulated, PFC corrects atmospheric components much better than the modal gain optimization of OFC can. The structure of the control law allows specific layers to be selectively compensated. This significantly reduces the intensity of scattered light in the PSF, as demonstrated by our simulations. As discussed in prior work [6], speckles caused by uncorrected atmosphere dominate contrast in long-exposure images. The use of PFC allows these to be atmospheric speckles to be attenuated, leading to improved high-contrast imaging.

Much work remains to be done in the application and testing of PFC for GPI. The essential first step is a study of the behavior of the atmosphere as measured by a spatially filtered WFS with fine spatial resolution. The extent to which frozen flow holds must be ascertained, as well as the time scales on which the layer velocities change. If necessary the atmospheric model can be modified to account for non-frozen flow behavior (e.g., boiling) by adding either more parameters to the model of layer evolution or by adding in another term with the temporal structure of the boiling. As long as these parameters can be estimated from the temporal PSD, the PFC framework can be applied. Given this framework, it is reasonable to envision a control system that uses models tuned to specific modes, such as correcting vibrations in certain modes, correcting frozen flow layers in others, and using OFC in the rest. The ability of a predictive controller to

improve performance depends on how well its model of the phase aberration fits reality, so some flexibility will be required.

Other research topics will include studying the impact of the response of the DM (e.g., nonlinearities that make it not compensate exactly the phase that it is commanded with) on the stability and quality of the predictor. As mentioned in Section 4, directly estimating the linear phase terms for each layer for all modes at once will also be studied.

ACKNOWLEDGMENTS

This work was performed under the auspices of the U.S. Department of Energy by the University of California, Lawrence Livermore National Laboratory under contract W-7405-Eng-48. The document number is UCRL-JRNL-225647. 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 AST-9876783.

REFERENCES

1. E. Gendron and P. Léna, "Astronomical adaptive optics I. Modal control optimization," *Astron. Astrophys.* **291**, 337–347 (1994).
2. J.-P. Véran, "Altair's optimiser," Tech. Rep. (Herzberg Institute of Astrophysics, 1998).
3. G. Rousset, F. Lacombe, P. Puget, N. N. Hubin, E. Gendron, T. Fusco, R. Arsenault, J. Charton, P. Feautrier, P. Gigan, P. Y. Kern, A. Lagrange, P. Madec, D. Mouillet, D. Rabaud, P. Rabou, E. Stadler, and G. Zins, "NAOS, the first AO system of the VLT: on-sky performance," *Proc. SPIE* **4839**, 140–149 (2003).
4. L. A. Poyneer, D. T. Gavel, and J. M. Brase, "Fast wave-front reconstruction in large adaptive optics systems with use of the Fourier transform," *J. Opt. Soc. Am. A* **19**, 2100–2111 (2002).
5. L. A. Poyneer and J.-P. Véran, "Optimal modal Fourier transform wave-front control," *J. Opt. Soc. Am. A* **22**, 1515–1526 (2005).
6. L. A. Poyneer and B. A. Macintosh, "Optimal Fourier control performance and speckle behavior in high-contrast imaging with adaptive optics," *Opt. Express* **17**, 7499–7514 (2006).
7. B. Macintosh, J. Graham, D. Palmer, R. Doyon, D. Gavel, J. Larkin, B. Oppenheimer, L. Saddlemyer, J. K. Wallace, B. Bauman, J. Evans, D. Erikson, K. Morzinski, D. Phillion, L. Poyneer, A. Sivaramakrishnan, R. Soummer, S. Thibault, and J.-P. Véran, "The Gemini Planet Imager," *Proc. SPIE* **6272**, 62720L (2006).
8. T. Fusco, C. Petit, G. Rousset, J.-F. Sauvage, K. Dohlen, D. Mouillet, J. Charton, P. Baudoz, M. Kasper, E. Fedrigo, P. Rabou, P. Feautrier, M. Downing, P. Gigan, J.-M. Conan, J.-L. Beuzit, N. Hubin, F. Wildi, and P. Puget, "Design of the extreme AO system for SPHERE, the planet finder instrument of the VLT," *Proc. SPIE* **6272**, 62720K (2006).
9. C. Dessenne, P.-Y. Madec, and G. Rousset, "Optimization of a predictive controller for closed-loop adaptive optics," *Appl. Opt.* **37**, 4623–4633 (1998).
10. D. T. Gavel and D. Wiberg, "Towards Strehl-optimizing adaptive optics controllers," *Proc. SPIE* **4839**, 890–901 (2002).
11. D. Wiberg, L. Johnson, and D. Gavel, "Adaptive optics control of wind blown turbulence via translation and prediction," *Proc. SPIE* **6272**, 62722X (2006).
12. B. Le Roux, J.-M. Conan, C. Kulcsar, H.-F. Raynaud, L. M. Mugnier, and T. Fusco, "Optimal control law for classical and multiconjugate adaptive optics," *J. Opt. Soc. Am. A* **21**, 1261–1276 (2004).
13. C. Petit, J.-M. Conan, C. Kulcsar, H.-F. Raynaud, T. Fusco, J. Montri, and D. Rabaud, "First laboratory demonstration of closed-loop Kalman based optimal control for vibration filtering and simplified MCAO," *Proc. SPIE* **6272**, 62721T (2006).
14. M. Frigo and S. G. Johnson, "The design and implementation of FFTW3," *Proc. IEEE* **93**, 216–231 (2005).
15. A. Sivaramakrishnan, J. P. Lloyd, P. Hodge, and B. A. Macintosh, "Speckle decorrelation and dynamic range in speckle noise-limited imaging," *Astrophys. J.* **581**, L59–62 (2002).
16. M. D. Perrin, A. Sivaramakrishnan, R. B. Makidon, B. R. Oppenheimer, and J. R. Graham, "The structure of high Strehl ratio point-spread functions," *Astrophys. J.* **596**, 702–712 (2003).
17. L. A. Poyneer and B. Macintosh, "Spatially filtered wave-front sensor for high-order adaptive optics," *J. Opt. Soc. Am. A* **21**, 810–819 (2004).
18. J.-M. Conan, G. Rousset, and P.-Y. Madec, "Wave-front temporal spectra in high-resolution imaging through turbulence," *J. Opt. Soc. Am. A* **12**, 1559–1570 (1995).
19. J. W. Hardy, *Adaptive Optics for Astronomical Telescopes* (Oxford U. Press, 1999).
20. A. Papoulis and S. U. Pillai, *Probability, Random Variables and Stochastic Processes*, 4th ed. (McGraw-Hill, 2002).
21. A. V. Oppenheim, R. W. Schaffer, and J. R. Buck, *Discrete-Time Signal Processing* (Prentice Hall, 1999).
22. T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation* (Prentice Hall, 2000).
23. E. M. Johansson and D. T. Gavel, "Simulation of stellar speckle imaging," *Proc. SPIE* **1237**, pp. 372–383 (1994).