European Journal of Control (2011)3:237–248

DOI:10.3166/EJC.17.237–248

European Journal of Control

Structure of LQG Controllers Based on a Hybrid Adaptive Optics System Model

Douglas P. Looze

Department of Electrical and Computer Engineering, University of Massachusetts, Amherst, MA, USA

The effects of a controller on the residual wavefront variance in an adaptive optics system can be represented by a discrete-time system. Consequently, the controller design that minimizes the wavefront is given by the solution of a discrete-time Linear-Quadratic-Gaussian (LQG) problem. The purpose of this paper is to analyze the structure of the LQG controller that minimizes the residual wavefront variance. It is shown that the LQG controller is an integral controller when the DM has no dynamics, there is no computational delay (the total loop delay is one frame), and the PSD of the incident wavefront decreases by f^{-2} at all frequencies. Non-zero computational delays result in a multivariable generalization lead element being added to the controller with the zero of the lead element at the origin.

Keywords: Adaptive optics, control, optimal control

1. Introduction

A wavefront from an astronomical object that is incident on a ground-based telescope is distorted by the varying index of refraction of the atmosphere along its optical path. The objective of an adaptive optics (AO) system is to reduce the effects of these atmospheric aberrations on the science instrument. A block diagram of a classical AO system (with co-located guidestar and science object is shown) in Fig. 1. The wavefront of the light from the target star – the incident wavefront (IWF) – is reflected from a deformable mirror (DM) to produce the residual wavefront. This is then supplied to both the science instrument and a wavefront sensor (WFS). The WFS creates a discrete-time measurement by sampling an approximation to the slopes or curvatures of the continuous-time residual wavefront integrated over one frame[1]. The discrete-time controller acts on this measurement to produce a discrete-time control (actuator) signal. This signal is converted to an analog voltage signal and applied to the DM to modify its surface shape. All devices in the AO loop (DM, WFS and controller) can have computational delays in addition to their inherent dynamics. Since the computational delays are assumed to be the same for all signals, they can be regarded (as shown in Fig. 1) as occurring at the input to the DM.

The temporal performance of an AO system measures the quality of the residual wavefront presented to the science instrument. Because the target star in classical astronomic AO is effectively at infinity, the residual wavefront from the target star should be as flat (constant) as possible. One common measure of AO performance is the time-average of the aperture variance of the residual wavefront. Thus, the structure of the feedback loop is a standard hybrid (sampled-data) system with a continuous-time performance signal. The effects of the discrete-time control signals and continuous-time IWF on the continuous-time residual wavefront variance can be represented by

E-mail: looze@ecs.umass.edu

Received 14 January 2010; Accepted 26 January 2011
Recommended by C. Kulcsár, H.-F. Raynaud, J.-M. Conan, D.W. Clarke

[1] This paper assumes that the integration is over the entire frame of the CCD camera. The model used can be suitably modified to handle integrations that are less than a frame.

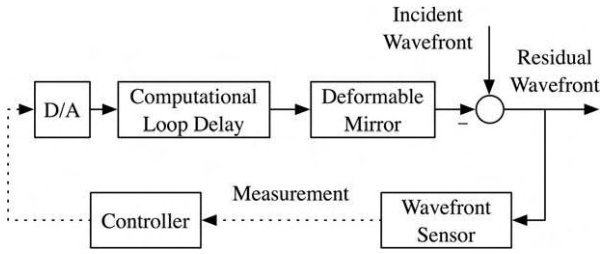

Fig. 1. Organization of a classical AO system. Continuous-time signals are denoted by a solid (——) line, while discrete-time signals are denoted by a dotted line (· · ·).

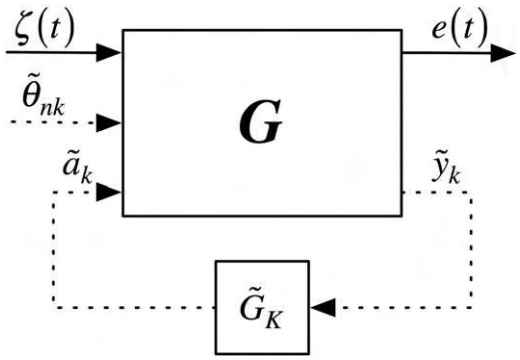

Fig. 2. Generalized feedback architecture for the AO system from Fig. 1.

a discrete-time system with a discrete-time, quadratic performance (c.f. [14], [2], [3], [15]). The controller design can then be optimized via Linear-Quadratic-Gaussian (LQG) control design problem.

Several authors have considered AO control design using the discrete-time LQG methodology. LQG techniques (and equivalents) were first applied to AO systems by Paschall and Anderson [16] with a controller design based on a system that modeled the incident wavefront (IWF) using the first 14 Zernike modes (ignoring the piston mode) with spectra generated by first order, independent Markov models. Looze *et al*. [12] designed an LQG modal controller based on a model of the spatial modes and the measured modal spectra of the IWF. Le Roux *et al.* [11] developed and simulated LQG designs for both classical and multiconjugate AO systems. An LQG design of a static controller was presented by Wiberg *et al.* [23]. Looze [13] analyzed the structure of an LQG controller based on a discrete-time model with the assumption that the intra-frame response of the residual was approximately constant. The efficiency of an LQG controller was first demonstrated in a laboratory experiment for both classical AO and MCAO systems by Petit *et al.* ([17], [18]). The classical AO experiment illustrated the ability of an LQG controller to reject vibrations. Hinnen *et al.* [6] used measurement data to specify the (LQG) parameters of an AO controller, and demonstrated performance improvement in a laboratory experiment. The structure of an LQG controller and its relation to integral controllers was investigated by Kulcsár *et al.* in [10].

The purpose of this paper is to give an analysis of the structure of the LQG controller that minimizes the residual wavefront variance when the effects of intra-frame signal behavior is included in the model. The paper first shows that the LQG control is the projection of the predicted IWF estimate on the actuators. When a discrete-time IWF model is assumed, this result has been shown for delays that are integral multiples of the frame period [10]. It has also been speculated that it is true for all delays (c.f. [19]). This paper demonstrates that this structure is valid for arbitrary delays and a continuous-time IWF model.

Reference [13] showed that when there is no computational delay, the DM dynamics can be neglected, and when the IWF PSD is proportional to f^{-2}, the standard integral AO controller is the solution to the LQG control design problem. However, this analysis assumed that the intra-frame response of the residual was approximately constant. When the intra-frame behavior of the IWF is represented, the resulting discrete-time model differs from that used in [13]. However, as shown in this paper, the same result holds (albeit with a different gain).

This paper also presents an analysis of the controller when the computational delay is non-zero but the other conditions (on DM dynamics and IWF PSD) remain. The structure of the resulting controller is an integral type compensator with additional filtering being added – a multivariable generalization of a lead filter. The controller form is computed in terms of the Kalman gain when the computational delay is one frame. The roles of the signal-to-noise (SNR) ratio and computational delay are illustrated in terms of a modal controller.

Section 2 presents the hybrid model of a classical AO system without isoplanatism, and the resulting performance-equivalent discrete-time model. The general LQG solution using this discrete-time model is presented in Section 3. Section 4 inspects the structure of the general LQG solution imposed by the AO system. In particular, the effects of the computational delay and atmospheric aberration model on the solution are examined.

2. Discrete-Time Model of the Adaptive optics System

2.1. Continuous-Time Models and Performance

The AO feedback system shown in Fig. 1 is a single sample rate feedback system that can be represented as shown in Fig. 2. The generalized plant model $\boldsymbol{G}$ is shown in Fig. 3. In these two figures, $e(t)$ is the residual wavefront, $\tilde{y}_k$ is the measurement, and $\tilde{a}_k$ is the control. The discrete-time

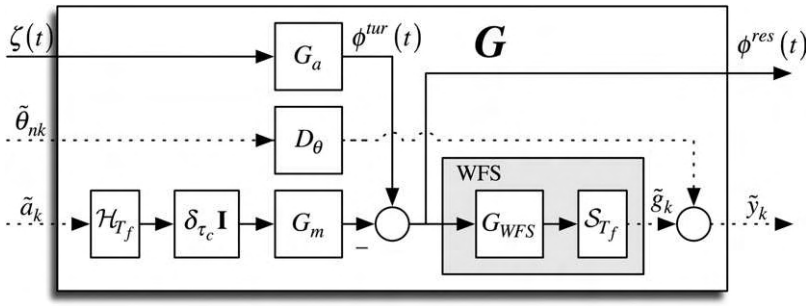

Fig. 3. Standard adaptive optics system with no anisoplanatism. To emphasize the sampled-data system, continuous-time signals are shown with solid lines, while discrete-time signals are shown with dotted lines.

control is applied to the zero-order hold (ZOH) $\mathcal{H}_{T_f}$ as soon as it is available. Consequently, the sample time and the hold time are not physically simultaneous.

Throughout this paper, the computational delay, denoted by τ_c, will be distinguished from the total loop delay. The computational loop delay includes the readout time of the CCD, the time needed to compute the gradients from the CCD readout, and the time needed to compute the control. It also includes the effective delay due to the asynchronous operation of the WFS and ZOH. It does not include the integration of the WFS or the effects of the sampling and ZOH. When the latter effects are included, loop delay or latency will be referred to as the total loop delay.

Because the computational delay is assumed to be equal for all signals, it can be represented as occurring immediately following the ZOH, and is represented in Fig. 3 by the operator $\delta_{\tau_c} I$. The delayed output of the ZOH is applied to the DM whose transfer function is $G_m(s)$. The DM surface position is the output of the DM dynamics represented by the transfer function $G_m(s)$.

The IWF is modeled as the output $\phi^{tur}(t)$ of a linear system with transfer function $G_a(s)$ whose input is a white noise process $\zeta(t)$. The residual wavefront is the difference between the IWF and the DM surface and is both passed to the science instrument and applied to the WFS The WFS is represented by a continuous-time transfer function $G_{WFS}(s)$ followed by a sampler $\mathcal{S}_{T_f}$. The measurement is the output of the WFS with additive discrete-time white noise $\theta_k = D_\theta \theta_{nk}$ (where θ_{nk} is a discrete-time, unit white noise process).

2.1.1. Wavefronts

The effects of the varying index of refraction of the atmosphere on a wavefront have been studied extensively ([5], [21]). For the weak scattering appropriate in astronomical AO, either the Kolmogorov [9] or von Karman model [7] is used to represent the spatial PSD of these variations. The temporal variations of the atmospheric index of refraction are modeled by dividing the atmosphere into thin layers and invoking the Taylor frozen flow approximation [22]. That is, the shape index of refraction remains unchanged and the only variation is due to horizontal translation by the wind velocity of the layer. This approach was used in [1] to compute illustrations of spectra for a single layer.

The IWF will be denoted by the vector signal $\phi^{tur}(t) \in \mathbb{R}^q$, and can be either a vector of the IWF phase values at points in the pupil (often at the actuator locations) or a vector of coefficients of a spatial modal decomposition of the IWF. The IWF is modeled as the output of a finite dimensional, strictly causal, LTI system whose input $\zeta(t) \in \mathbb{R}^{m_a}$ is a zero mean, white stochastic process with unit power spectral density:

$$\begin{aligned} \dot{x}_a(t) &= A_a x_a(t) + B_a \zeta(t) \\ \phi^{tur}(t) &= C_a x_a(t) \end{aligned} \tag{1}$$

The realization (1) determines the IWF transfer function $G_a(s)$.

Appropriate selection of the realization matrices A_a, B_a and C_a can approximate the true PSD of the arbitrarily closely [8]. The model (1) also incorporates spatial correlations. For example, if $\phi^{tur}(t)$ is a vector of IWF values at points in the aperture, then the cross correlation between two components of $\phi^{tur}(t)$ is the spatial correlation between the points represented by those components at the time t.

The use of Fourier modes (c.f. [19]) has been proposed as a way to exploit the frozen flow process. The IWF model used in [19] is the discrete-time version of (1) with the generalization that the system matrices can be complex. Since Fourier modes are a form of modal decomposition of the IWF, (1) can also be applied the model in [19].

The signal $s_m(t) \in \mathbb{R}^q$ represents the surface of the DM using the same definition of components as IWF vector $\phi^{tur}(t)$. The residual wavefront $\phi^{res}(t)$ is formed by reflecting the light from the incident wavefront from the DM:

$$\phi^{res}(t) = \phi^{tur}(t) - s_m(t) \in \mathbb{R}^q \tag{2}$$

The residual wavefront is sent to both the science instrument and the wavefront sensor. Hence, $\phi^{res}(t)$ is also the performance signal of the system.

2.1.2. Deformable Mirror and Computational Delay

Although any finite dimensional linear, time-invariant (LTI) model can be used to represent the DM dynamics, this paper will use only models with no dynamics ($G_m(s) = D_m$). The actuator voltages for the DM $a(t) \in \mathbb{R}^m$ are the output of a ZOH and delayed by the loop computational delay τ_c:

$$a = \delta_{\tau_c}\mathcal{H}_{T_f}\tilde{a} : a(t-\tau_c) = \tilde{a}_k \quad \text{for } kT_f \le t < (k+1)T_f \tag{3}$$

It will be assumed throughout this paper that the DM is rigid with no dynamics. Thus, the DM surface is determined by the instantaneous actuator value:

$$s_m(t) = D_m a(t) \tag{4}$$

That is, the DM transfer function is $G_m(s) = D_m$.

2.1.3. Wavefront Sensor Model

The WFS integrates the residual over one frame of the CCD camera and then samples the result to produce the WFS measurement. The measurements $\tilde{y}_k \in \mathbb{R}^p$ produced by the WFS approximate the residual wavefront gradients (or curvatures) $\tilde{g}_k \in \mathbb{R}^p$ with additive measurement noise $\tilde{\theta}_k \in \mathbb{R}^p$. The measurement noise is a zero mean, white noise sequence with covariance $\Theta \in \mathbb{R}^{p\times p}$. It is normalized by the square root of the covariance to obtain a zero mean, white noise sequence with unit variance:

$$\tilde{\theta}_k \triangleq D_\theta \tilde{\theta}_{nk} \quad D_\theta = \sqrt{\Theta} \in \mathbb{R}^{p\times p} \tag{5}$$

The gradients $g_k \in \mathbb{R}^p$ are approximated by a linear combination of the components of the residual wavefront vector (represented by multiplication by the geometry matrix $H \in \mathbb{R}^{p\times q}$). The WFS approximates the gradient by integrating the residual wavefront gradients over one frame and sampling the result. The WFS measurement can be modeled as:

$$\tilde{y}_k = \tilde{g}_k + \tilde{\theta}_k = \tilde{g}_k + D_\theta \tilde{\theta}_{nk} \tag{6}$$

$$\tilde{g} = \mathcal{S}_{T_f} g : \tilde{g}_k = g(kT_f) \tag{7}$$

$$g = G_{WFS}\phi^{res} : g(t) = \frac{1}{T_f} H \int_{t-T_f}^{t} \phi^{res}(\sigma)d\sigma \tag{8}$$

The linear operators in (8) are LTI and have Laplace transforms

$$G_{WFS}(s) = \frac{1}{T_f} H\left(1 - e^{-sT_f}\right) \quad \delta_{\tau_c}(s) = e^{-s\tau_c} I \tag{9}$$

2.1.4. Performance

A common measure the quality of an image produced by a science instrument (and hence of adaptive optics system performance) is the Strehl ratio ([5], [20], [21]). Assuming the extended Maréchal approximation is valid, the Strehl ratio is a strictly decreasing function of the residual (error) wavefront variance. The steady state, frame-averaged residual variance J is an appropriate measure of the performance of an AO system:

$$J = \lim_{W\to\infty} \frac{1}{W} \int_0^W \mathcal{E}\left\{\phi^{res}(\sigma)^T \phi^{res}(\sigma)\right\} d\sigma \tag{10}$$

2.2. Discrete-Time Model

2.2.1. Model

This section will present an equivalent discrete-time model for the generalized adaptive optics system in Fig. 3 (see [15]).

The discrete-time model for the AO system (Fig. 3) presented in [15] can be used to determine the performance of the hybrid AO system for any controller. When the DM is rigid (4), the discrete-time model is:

$$\begin{aligned} \tilde{\mathbf{x}}_{k+1} &= \mathbf{A}\tilde{\mathbf{x}}_k + \mathbf{B}_a\tilde{\zeta}_k + \mathbf{B}_m\tilde{a}_k & \tilde{\mathbf{x}}_k &\in \mathbb{R}^{n_{aug}} \\ \tilde{\phi}_k &= \mathbf{C}_p\tilde{\mathbf{x}}_k - D_{pm}\tilde{a}_k & \tilde{\phi}_k &\in \mathbb{R}^{p_e} \\ \tilde{y}_k &= \mathbf{C}_m\tilde{\mathbf{x}}_k + D_\theta\tilde{\theta}_{nk} & \tilde{\zeta}_k &\in \mathbb{R}^{m_\zeta} \end{aligned} \tag{11}$$

where $\tilde{\zeta}_k$ is a unit variance, white noise sequence, and $\tilde{\phi}_k$ is the discrete-time performance signal. The dimension is $n_{aug} = 2m + n_a + q$, and the dimensions m_ζ and p_e are defined by the Cholesky factorizations below. The state vector $\tilde{\mathbf{x}}_k$ in (11) can be partitioned as

$$\begin{gathered} \tilde{\mathbf{x}}^T = \begin{bmatrix} \tilde{x}_{m1}^T & \tilde{x}_{m2}^T & | & \tilde{x}_{a1}^T & \tilde{x}_{a2}^T \end{bmatrix} \\ \tilde{x}_{m1}, \tilde{x}_{m2} \in \mathbb{R}^m;\ \tilde{x}_{a1} \in \mathbb{R}^{n_a};\ \tilde{x}_{a1} \in \mathbb{R}^q \end{gathered} \tag{12}$$

Define the discrete-time cost functional as:

$$J_d = \lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N} \mathcal{E}\left\{\tilde{\phi}_k^T \tilde{\phi}_k\right\} \tag{13}$$

The matrices that define the discrete-time system (11) can be partitioned to conform to the state. The system matrix (for a rigid DM) is:

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{md} & 0 \\ 0 & \mathbf{A}_{ad} \end{bmatrix} = \left[\begin{array}{cc|cc} 0 & I & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & A_{ad} & 0 \\ 0 & 0 & C_{ad} & 0 \end{array}\right] \tag{14}$$

$$A_{ad} = e^{A_a T_f} \quad C_{ad} = \int_0^{T_f} C_a e^{A_a v} dv \tag{15}$$

The control and incident wavefront input matrices are:

$$\mathbf{B}_m = \left[\begin{array}{c} 0 \\ \hline I \\ \hline 0 \\ 0 \end{array}\right] \in \mathbb{R}^{n_{aug} \times m} \quad \mathbf{B}_a = \left[\begin{array}{c} 0 \\ \hline \mathbf{B}_{ad} \end{array}\right] \in \mathbb{R}^{n_{aug} \times m_\zeta};$$

$$\mathbf{M} = \mathbf{B}_{ad}\mathbf{B}_{ad}^T = \begin{bmatrix} M_{11} & M_{12} \\ M_{12}^T & M_{22} \end{bmatrix} \in \mathbb{R}^{(n_a+p)\times(n_a+p)} \tag{16}$$

where

$$\begin{aligned} \Gamma_a(t) &= C_a \int_0^t e^{A_a \sigma} d\sigma \\ M_{11} &= \int_0^{T_f} e^{A_a \gamma} B_a B_a^T e^{A_a^T \gamma} d\gamma \\ M_{12} &= \int_0^{T_f} e^{A_a \gamma} B_a B_a^T \Gamma_a^T(\gamma) d\gamma \\ M_{22} &= \int_0^{T_f} \Gamma_a^(\gamma) B_a B_a^T \Gamma_a^T(\gamma) d\gamma \end{aligned} \tag{17}$$

Let $\beta = T_f - \tau_c$ and $\overline{H} = (1/T_f)H$. The output matrices in (11) are:

$$\begin{aligned} \mathbf{C}_p &= \left[0 \quad -C_{pm} \;\middle|\; C_{pa} \quad 0\right] \in \mathbb{R}^{p_e \times n_{aug}} \\ \mathbf{C}_m &= \bar{H}\left[-\tau_c D_m \quad -\beta D_m \;\middle|\; 0 \quad I\right] \in \mathrm{R}^{q \times n_{aug}} \end{aligned} \tag{18}$$

The submatrices C_{pm}, C_{pa}, and D_{pm} are:

$$\begin{aligned} &\begin{bmatrix} -C_{pm}^T \\ C_{pa}^T \\ -D_{pm}^T \end{bmatrix} \begin{bmatrix} -C_{pm} & C_{pa} & -D_{pm} \end{bmatrix} \\ &= \mathbf{N} = \begin{bmatrix} N_{11} & -N_{12} & N_{13} \\ -N_{12}^T & N_{22} & -N_{23} \\ N_{13}^T & -N_{23}^T & N_{33} \end{bmatrix} \end{aligned} \tag{19}$$

The submatrices of N in (19) are:

$$\begin{aligned} N_{11} &= \tau_c D_m^T D_m & N_{33} &= \beta D_m^T D_m \\ N_{22} &= \int_0^{T_f} e^{A_a^T \sigma} C_a^T C_a e^{A_a \sigma} d\sigma & N_{13} &= 0 \\ N_{12} &= D_m^T \int_0^{\tau_c} C_a e^{A_a \sigma} d\sigma & N_{23} &= \int_{\tau_c}^{T_f} e^{A_a^T \sigma} C_a^T d\sigma D_m \end{aligned} \tag{20}$$

For a specified compensator, the closed-loop performance is the sum of the closed-loop performance of the discrete-time model (11)-(13) and a term that is independent of the controller. The performance is:

$$J = J_0 + J_d \quad J_0 = \mathrm{tr}\left\{ B_a^T \int_0^{T_f} \int_0^{\gamma} e^{A_a^T \sigma} C_a^T C_a e^{A_a \sigma} d\sigma B_a \right\} \tag{21}$$

Each of the integrals that define the model (11)–(21) can be computed from appropriate matrix exponentials (see [15]).

2.2.2. *Interpretation*

Although many researchers have based their designs on discrete-time models ([6], [10], [11], [12], [13], [15], [16], [17], [18], [19], [23]), the discrete-time model given by (11)–(21) differs by incorporating the effects of the intra-fame residual behavior by appropriately modeling the discrete-time IWF and the discrete-time performance signal. Unless these effects are correctly incorporated, formal control design procedures can give poor results.

The components of the state in $\tilde{x}_{a2,k}$ are the outputs of the WFS due to the IWF before multiplication by the geometry matrix H at time $(k-1)T_f$. The components of the state in $\tilde{x}_{m2,k}$ are the actuators at time $(k-1)T_f$ (delayed by one frame) while the components of the state in $\tilde{x}_{m1,k}$ are the actuators at time $(k-2)T_f$ (delayed by two frames). Thus, the measurement is the difference between the frame-averaged wavefront and the DM surface as commanded but delayed by τ_c. The intra-frame statistics of the IWF are incorporated by defining $\mathbf{B}_{ad}$ appropriately (16)–(17).

The performance signal accounts for both the intra-frame behavior of the IWF and the non-constant DM surface during the frame. These effects are incorporated

through the matrices C_{pa}, C_{pm} and D_{pm} (18)–(20). In particular, the performance now is a direct function of the actuator commands $\tilde{a}_k$.

Note that representation of the WFS measurement requires two delays of the actuator commands and one delay of the average IWF. One loop delay of the actuator commands and average IWF is required by the interaction of the WFS and ZOH – with no computational delay, the loop delay is still one frame. The second actuator delay is required to reflect the added effects of the computational delay.

3. LQG Solution

Selecting the controller to minimize the discrete-time quadratic cost (13) will also minimize the continuous-time performance (10) since the term J_0 is independent of the controller (21). It will also maximize the Strehl ratio of the science signal (if the extended Maréchal approximation is valid [5], [21]). The minimizing controller is given by the LQG solution if the exogenous noise sources $\tilde{\zeta}$ are assumed Gaussian, or if the noises are assumed non-Gaussian but the controller is constrained to be LTI.

The hybrid LQG controller is defined by the solutions of two algebraic Riccati equations (ARE). The AREs use the quadratic cost ($\mathbf{Q}$, $\mathbf{N}_p$ and $\mathbf{R}$) and noise PSD (Ξ and Θ) matrices:

$$\mathbf{Q} = \mathbf{C}_p^T\mathbf{C}_p = \left[\begin{array}{cc:cc} 0 & 0 & 0 & 0 \\ 0 & N_{11} & -N_{12} & 0 \\ \hdashline 0 & -N_{12}^T & N_{22} & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

$$\mathbf{N}_p = -\mathbf{C}_p^T D_{pm} = \left[\begin{array}{c} 0 \\ 0 \\ -N_{23} \\ 0 \end{array}\right] \quad \mathbf{R} = D_{pm}^T D_{pm} \tag{22}$$

$$\Xi = \mathbf{B}_{ad}\mathbf{B}_{ad}^T \quad \Theta = D_\theta D_\theta^T$$

The minimizing (LQG) controller is:

$$\begin{aligned} \hat{\mathbf{x}}_{k+1} &= (\mathbf{A} - \mathbf{L}\mathbf{C}_m)\hat{\mathbf{x}}_k + \mathbf{B}_m\tilde{a}_k + \mathbf{L}\tilde{y}_k \\ \hat{\mathbf{x}}_{k|k} &= (I - \mathbf{L}_f\mathbf{C}_m)\hat{\mathbf{x}}_k + \mathbf{L}_f\tilde{y}_k \\ \tilde{a}_k &= \mathbf{K}\hat{\mathbf{x}}_{k|k} \end{aligned} \tag{23}$$

where $\hat{\mathbf{x}}_k$ is the 1-step prediction estimate of $\mathbf{x}_k$ (using measurements through sample $k-1$), and $\hat{\mathbf{x}}_{k|k}$ is the current estimate of $\mathbf{x}_k$ (using measurements through k).

The feedback gain $\mathbf{K}$ is found by solving the control ARE for the symmetric, stabilizing, positive semi-definite cost matrix $\mathbf{S}$:

$$\begin{aligned} \mathbf{S} &= \mathbf{A}^T\mathbf{S}\mathbf{A} + \mathbf{Q} - \left(\mathbf{A}^T\mathbf{S}\mathbf{B}_m + \mathbf{N}_p\right) \\ &\quad \times \left(\mathbf{B}_m^T\mathbf{S}\mathbf{B}_m + \mathbf{R}\right)^{-1}\left(\mathbf{A}^T\mathbf{S}\mathbf{B}_m + N_p\right)^T \\ \mathbf{K} &= -\left(\mathbf{B}_m^T\mathbf{S}\mathbf{B}_m + \mathbf{R}\right)^{-1}\left(\mathbf{A}^T\mathbf{S}\mathbf{B}_m + \mathbf{N}_p\right)^T \end{aligned} \tag{24}$$

The estimation $\mathbf{L}$ and innovations $\mathbf{L}_f$ gains are found by solving the filter ARE for the symmetric, stabilizing, positive semi-definite (one-step prediction) error covariance matrix Σ:

$$\begin{aligned} \Sigma &= \mathbf{A}\Sigma\mathbf{A}^T + \Xi - \mathbf{A}\Sigma\mathbf{C}_m^T\left(\mathbf{C}_m\Sigma\mathbf{C}_m^T + \Theta\right)^{-1}\mathbf{C}_m\Sigma\mathbf{A}^T \\ \mathbf{L}_f &= \Sigma\mathbf{C}_m^T\left(\mathbf{C}_m\Sigma\mathbf{C}_m^T + \Theta\right)^{-1} \quad \mathbf{L} = \mathbf{A}\mathbf{L}_f \end{aligned} \tag{25}$$

The solution to the estimation ARE Σ is the one-step prediction error covariance.

The controller states $\hat{\mathbf{x}}_k$ are the Kalman-Bucy filter (KBF) estimates of states of the discrete-time model (11) using measurements through sample $k-1$.

4. Structure of the LQG Solution

4.1. General

4.1.1. Control

Because the 1st and 4th blocks of states are not penalized in the objective functional, and cannot affect the objective functional through the model dynamics, these states are not observable through the objective functional. Hence, the blocks in the Riccati solution corresponding to these states are zero:

$$\mathbf{S} = \left[\begin{array}{cccc} 0 & 0 & 0 & 0 \\ 0 & \mathbf{S}_{22} & \mathbf{S}_{23} & 0 \\ 0 & \mathbf{S}_{23}^T & \mathbf{S}_{33} & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \tag{26}$$

Using the structure of the system matrix (14), the mirror input matrix (16), the Riccati solution (26), and the definitions of the cost matrices (22) and (19)–(20), the (2,2) and (2,3) blocks of the Riccati equation (24) are:

$$\begin{aligned} \mathbf{S}_{22} &= N_{11} = \tau_c D_m^T D_m \\ \mathbf{S}_{23} &= -N_{12} = -D_m^T \int_0^{\tau_c} C_a e^{A_a\sigma}\, d\sigma \end{aligned} \tag{27}$$

The gain $\mathbf{K}$ is then computed (independent of $\mathbf{S}_{33}$) as:

$$\begin{aligned}\mathbf{K} &= \begin{bmatrix}0 & 0 & \underbrace{K_3}_{\boldsymbol{\kappa}_a} & 0\end{bmatrix}\\ K_3 &= \tfrac{1}{T_f}\left(D_m^T D_m\right)^{-1} D_m^T C_{ad} e^{A_a \tau_c}\end{aligned} \tag{28}$$

4.1.2. Estimation

Because the 1st and 2nd blocks of states (the DM states) are not affected by the process noise (the IWF model), these states are not controllable by process noise. The corresponding blocks in the Riccati equation (25) are zero. Let $\tilde{\mathbf{x}}_a^T = \begin{bmatrix}\tilde{x}_{a1}^T & \tilde{x}_{a2}^T\end{bmatrix}$. Then:

$$\Sigma = \begin{bmatrix}0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & \Sigma_a\end{bmatrix} \quad \mathbf{L} = \begin{bmatrix}0\\ 0\\ \mathbf{A}_{da}\mathbf{L}_{fa}\end{bmatrix} \quad \mathbf{L}_f = \begin{bmatrix}0\\ 0\\ \mathbf{L}_{fa}\end{bmatrix} \tag{29}$$

The non-zero matrix Σ_a in (29) is the error covariance of the discrete-time IWF states (i.e., the current discrete-time IWF states $\tilde{x}_{a1}^T$ and the one-step delayed discrete-time IWF).

The non-zero matrices in (29) satisfy standard KBF filter equations:

$$\begin{aligned}\mathbf{C}_a &= \bar{H}\begin{bmatrix}0 & I\end{bmatrix}\\ \Sigma_a &= \mathbf{A}_{ad}\Sigma_a\mathbf{A}_{ad}^T + \mathbf{B}_{da}\mathbf{B}_{da}^T\\ &\quad - \mathbf{A}_{ad}\Sigma_a\mathbf{C}_a^T\left(\mathbf{C}_a\Sigma_a\mathbf{C}_a^T + \Theta\right)^{-1}\mathbf{C}_a\Sigma_a\mathbf{A}_{ad}^T\\ \mathbf{L}_{fa} &= \Sigma_a\mathbf{C}_a^T\left(\mathbf{C}_a\Sigma_a\mathbf{C}_a^T + \Theta\right)^{-1}\end{aligned} \tag{30}$$

Thus, only the IWF states need to be estimated by the KBF.

4.1.3. Combined

The LQG control is the product of the LQ state feedback gain with the KBF estimate of the state vector. The LQ gain only multiplies the current IWF state estimate. It projects the estimate of the frame-integrated IWF using the pseudo-inverse of the DM:

$$a_k = \frac{1}{T_f}\left(D_m^T D_m\right)^{-1} D_m^T C_{ad} e^{A_a\tau_c}\hat{x}_{a1}[k|k] \tag{31}$$

where $\hat{x}_{a1}[k|k]$ denotes the 3rd ($a1$) block of the current estimate vector $\hat{\mathbf{x}}_{k|k}$. The predicted estimated average IWF at the DM surface is determined by the control (31) is $\frac{1}{T_f}C_{ad}e^{A_a\tau_c}\hat{x}_{a1}[k|k]$. This prediction is then projected on the DM. Note that this produces the same result as in [10] when the delay is an integral number of frames.

In terms of the original partitioning of the discrete-time system, the Kalman gains (29)are:

$$\mathbf{L}_f = \begin{bmatrix}0\\ 0\\ L_{f3}\\ L_{f4}\end{bmatrix} \quad \mathbf{L} = \begin{bmatrix}0\\ 0\\ A_{ad}L_{f3}\\ C_{ad}L_{f3}\end{bmatrix} \tag{32}$$

Using the definitions of the system matrix $\mathbf{A}$ (14), the input matrix $\mathbf{B}_m$ (16), the measurement matrix (18) $\mathbf{C}_m$, the LQ gain $\mathbf{K}$ (28), and the Kalman gains $\mathbf{L}_f$ and $\mathbf{L}$ (32) in the controller (23) gives:

$$\begin{aligned}\hat{\mathbf{x}}_{k+1} &= \mathbf{A}_{c0}\hat{\mathbf{x}}_k + \mathbf{B}_{c0}\tilde{y}_k\\ \tilde{a}_k &= \mathbf{C}_{c0}\hat{\mathbf{x}}_k + \mathbf{D}_{c0}\tilde{y}_k\end{aligned} \tag{33}$$

$$\mathbf{A}_{co} = \begin{bmatrix}0 & I & 0 & 0\\ \tau_c D_m K_3 L_{f3}\bar{H} & \beta D_m K_3 L_{f3}\bar{H} & D_m K_3 & -D_m K_3 L_{f3}\bar{H}\\ \tau_c A_{ad} L_{f3}\bar{H} & \beta A_{ad} L_{f3}\bar{H} & A_{ad} & -A_{ad} L_{f3}\bar{H}\\ \tau_c C_{ad} L_{f3}\bar{H} & \beta C_{ad} L_{f3}\bar{H} & C_{ad} & -C_{ad} L_{f3}\bar{H}\end{bmatrix}$$

$$\mathbf{B}_{c0} = \begin{bmatrix}0\\ D_m K_3 L_{f3}\\ A_{ad} L_{f3}\\ C_{ad} L_{f3}\end{bmatrix}$$

$$\mathbf{C}_{co} = K_3\begin{bmatrix}\tau_c L_{f3}\bar{H} & \beta L_{f3}\bar{H} & I & -L_{f3}\bar{H}\end{bmatrix}$$

$$\mathbf{D}_{c0} = K_3 L_{f3}$$

This controller has two blocks of unobservable states. This becomes evident with the transformation:

$$\hat{\mathbf{x}}_k = \begin{bmatrix}I & 0 & 0 & 0\\ 0 & 0 & I & 0\\ 0 & L_{f3}\bar{H} & 0 & I\\ \tau I & I & 0 & 0\end{bmatrix}\hat{\mathbf{z}}_{a,k} = \mathbf{P}_1\hat{\mathbf{z}}_{a,k};$$

$$\hat{\mathbf{z}}_{a,k} = \begin{bmatrix}\hat{\mathbf{z}}_{a1,k}\\ \hat{\mathbf{z}}_{a2,k}\\ \hat{\mathbf{z}}_{a3,k}\\ \hat{\mathbf{z}}_{a4,k}\end{bmatrix} \tag{34}$$

When this is applied to the realization (33), the transformed controller is:

$$\begin{aligned}\hat{\mathbf{z}}_{a,k+1} &= \mathbf{A}_{c1}\hat{\mathbf{z}}_{a,k} + \mathbf{B}_{c1}\tilde{y}_k\\ \tilde{a}_k &= \mathbf{C}_{c1}\hat{\mathbf{z}}_{a,k} + \mathbf{D}_{c1}\tilde{y}_k\end{aligned} \tag{35}$$

where

$$\mathbf{A}_{c1} = \mathbf{P}_1^{-1}\mathbf{A}_{c0}\mathbf{P}_1$$

$$= \begin{bmatrix} 0 & 0 & I & 0 \\ 0 & 0 & \beta C_{ad}L_{f3}\bar{H} - \tau_c I & C_{ad} \\ 0 & 0 & \beta D_m K_3 L_{f3}\bar{H} & D_m K_3 \\ 0 & 0 & \beta\left(A_{ad} - L_{f3}\bar{H}C_{ad} + \frac{\tau_c}{\beta}I\right)L_{f3}\bar{H} & A_{ad} - L_{f3}\bar{H}C_{ad} \end{bmatrix}$$

$$\mathbf{B}_{c1} = \mathbf{P}_1^{-1}\mathbf{B}_{c0} = \begin{bmatrix} 0 \\ C_{ad}L_{f3} \\ D_m K_3 L_{f3} \\ \left(A_{ad} - L_{f3}\bar{H}C_{ad}\right)L_{f3} \end{bmatrix}$$

$$\mathbf{C}_{c1} = \mathbf{C}_{c0}\mathbf{P}_1 = K_3\begin{bmatrix} 0 & 0 & \beta L_{f3}\bar{H} & I \end{bmatrix}$$

$$\mathbf{D}_{c1} = \mathbf{D}_{c0} = K_3 L_{f3}\tilde{y}_k$$

The two unobservable blocks of states can be removed to give the following description for the LQG controller with state $\hat{\mathbf{z}}_k^T = \begin{bmatrix}\hat{\mathbf{z}}_{a3k}^T & \hat{\mathbf{z}}_{a4k}^T\end{bmatrix}$:

$$\begin{aligned} \hat{\mathbf{z}}_{k+1} &= \mathbf{A}_c\hat{\mathbf{z}}_k + \mathbf{B}_c\tilde{y}_k \\ \tilde{a}_k &= \mathbf{C}_c\hat{\mathbf{z}}_k + \mathbf{D}_c\tilde{y}_k \end{aligned} \tag{36}$$

where

$$\mathbf{A}_c = \begin{bmatrix} \beta D_m K_3 L_{f3}\bar{H} & D_m K_3 \\ \beta\left(A_{ad} - L_{f3}\bar{H}C_{ad} + \frac{\tau_c}{\beta}I\right)L_{f3}\bar{H} & A_{ad} - L_{f3}\bar{H}C_{ad} \end{bmatrix}$$

$$\mathbf{B}_c = \begin{bmatrix} D_m K_3 L_{f3} \\ \left(A_{ad} - L_{f3}\bar{H}C_{ad}\right)L_{f3} \end{bmatrix}$$

$$\mathbf{C}_c = K_3\begin{bmatrix} \beta L_{f3}\bar{H} & I \end{bmatrix}$$

$$\mathbf{D}_c = K_3 L_{f3}$$

Note that this realization of the controller has $(m + n_a)$ states – the number of actuators plus the number of IWF states. The original controller model (33) has $(2m + n_a + q)$ states – an additional number of actuators plus the number of measurements.

4.2. Analysis for No Computational Delay

A commonly used controller for AO systems applies as its control the integration (summation in discrete-time) of the reconstructed IWF. A main purpose of this subsection is to show that the LQG controller is the commonly used integral controller under certain conditions: (1) the DM has no dynamics; (2) the computational delay is zero ($\tau_c = 0$); and (3) the IWF temporal dynamics have a PSD that is proportional to f^{-2}. The first of these conditions has been assumed throughout the paper. The second two conditions will be assumed throughout this subsection.

These conditions can be reasonable approximations. The computational delay will be approximately zero when dim guidestars are being used by the AO system (T_f will be large). The assumption regarding the IWF PSD is a reasonable approximation (c.f. [15]) to an atmosphere with Kolmogorov statistics (the theoretical dependence on frequency is $f^{-8/3}$ [1]).

Let α_a be a positive scalar. Assume the IWF is the output of the following system:

$$\begin{aligned} \dot{x}_a(t) &= -\alpha_a x_a(t) + B_a\zeta(t) \\ \phi^{tur}(t) &= \alpha_a x_a(t) \end{aligned} \Rightarrow A_a = -\alpha_a I \;\; C_a = \alpha_a I \tag{37}$$

This model of the IWF has uniform temporal dynamics with a PSD that behaves like f^{-2} at all frequencies greater than $\sim \alpha_a$. The discrete-time matrices (15) are:

$$A_{ad} = e^{-a_a T_f} I \quad C_{ad} = \frac{1}{\alpha_a}\left(1 - e^{-\alpha_a T_f}\right) I \tag{38}$$

The analog of a Kolmogorov atmosphere would have $\alpha_a = 0$ corresponding to an infinite outer scale and infinite energy. In this case, the solution to the filter Riccati equation (25) does not exist. However, the Riccati equation (25) has a solution for all non-zero values of α_a, the LQG controller (36) is well defined, the closed-loop system is input-output stable, and the performance (21) is finite. As α_a decreases to zero, the performance of the AO system is bounded, and the controller remains valid. At the limit $\alpha_a = 0$, the controller modeled by (36) minimizes the frame-averaged residual wavefront variance.

Note that for $\alpha_a = 0$, $C_{ad} = T_f I$ and hence $C_{ad}\bar{H} = H$. With $\tau_c = 0$, $\beta = T_f$ and $\beta\bar{H} = H$. The LQG controller (36) becomes

$$\begin{aligned} \hat{\mathbf{z}}_{k+1} &= \begin{bmatrix} D_m K_3 L_{f3} H & D_m K_3 \\ \left(I - L_{f3}H\right)L_{f3}H & I - L_{f3}H \end{bmatrix}\hat{\mathbf{z}}_k \\ &\quad + \begin{bmatrix} D_m K_3 L_{f3} \\ \left(I - L_{f3}H\right)L_{f3} \end{bmatrix}\tilde{y}_k \\ \tilde{a}_k &= K_3\begin{bmatrix} L_{f3}H & I \end{bmatrix}\hat{\mathbf{z}}_k + K_3 L_{f3}\tilde{y}_k \end{aligned} \tag{39}$$

This model again contains one block of unobservable states. Define the transformation

$$\hat{\mathbf{z}}_k = \begin{bmatrix} I & 0 \\ -L_{f3}H & I \end{bmatrix}\bar{\mathbf{z}}_{ak}; \quad \bar{\mathbf{z}}_{ak} = \begin{bmatrix} \bar{\mathbf{z}}_{a1k} \\ \bar{\mathbf{z}}_{a2k} \end{bmatrix} \tag{40}$$

When this is applied to the realization (39), the transformed controller is:

$$\begin{aligned}\bar{\mathbf{z}}_{a,k+1} &= \begin{bmatrix} 0 & D_m K_3 \\ 0 & I - L_{f3} H\,(I - D_m K_3) \end{bmatrix} \bar{\mathbf{z}}_{ak} \\ &\quad + \begin{bmatrix} D_m K_3 L_{f3} \\ \left(I - L_{f3} H\,(I - D_m K_3)\right) L_{f3} \end{bmatrix} \tilde{y}_k \\ \tilde{a}_k &= \begin{bmatrix} 0 & K_3 \end{bmatrix} \bar{\mathbf{z}}_{ak} + K_3 L_{f3} \tilde{y}_k \end{aligned} \tag{41}$$

The unobservable block of states can be removed to give the following description for the LQG controller with state $\bar{\mathbf{z}}_k = \bar{\mathbf{z}}_{a2k}$:

$$\begin{aligned}\bar{\mathbf{z}}_{k+1} &= \left(I - L_{f3} H\,(I - D_m K_3)\right) \hat{\mathbf{z}}_k \\ &\quad + \left(I - L_{f3} H\,(I - D_m K_3)\right) L_{f3} \tilde{y}_k \\ \tilde{a}_k &= K_3 \bar{\mathbf{z}}_k + K_3 L_{f3} \tilde{y}_k \end{aligned} \tag{42}$$

The Z-Transform of this controller is

$$\begin{aligned}\tilde{G}_K(z) &= K_3 z (zI - \Gamma)^{-1} L_{f3} \\ \Gamma &= I - L_{f3} H\,(I - D_m K_3) \end{aligned} \tag{43}$$

The measurement from the WFS is multiplied by the matrix L_{f3} to "reconstruct" the wavefront. The reconstruction is then fed to the dynamics of the LQG controller governed by the system matrix

$$\begin{aligned}\Gamma &= I - L_{f3} H\,(I - D_m K_3) \\ &= I - L_{f3} H \left(I - D_m \left(D_m^T D_m \right)^{-1} D_m^T \right) \end{aligned} \tag{44}$$

because for $C_{ad} = T_f I$, K_3 is the pseudo-inverse of D_m (28). Thus, if the reconstruction lies in the range of D_m (the actuators can apply this signal to the DM), it is integrated. Conversely, if the reconstruction is orthogonal to the range of D_m, the controller dynamics are those of the IWF estimator (and thus are stable). The result is projected onto the DM surface by K_3.

The latter point is emphasized by assuming that the actuator effectiveness matrix is invertible. Then, $I - D_m K_3 = 0$ and the controller (43) becomes

$$\tilde{G}_K(z) = \frac{z}{z-1} D_m^{-1} L_{f3} \tag{45}$$

Thus, the controller is the discrete-time integration of a reconstructed wavefront, with the reconstructor given by $D_m^{-1} L_{f3}$.

This subsection has shown that the commonly used integral-type (discrete-time summation) controller is an LQG controller when the computational delay is zero and when the IWF PSD is proportional to f^{-2}. Violations of these assumptions would lead to an LQG controller of a different form. In particular, if the IWF has a finite outer scale ($\alpha_a > 0$), the LQG controller will no longer have a pole at 1, and will not act as an integrator. However, representation of other errors (such as telescope static aberrations) can restore the integrator.

4.3. Analysis for One Frame Computational Delay

It will now be assumed that the computational delay is one frame: $\tau_c = T_f$ and $\beta = 0$. As in the last section, it will be assumed that the IWF has temporal dynamics that are uniform with a PSD that is proportional to f^{-2}. Consequently, (38) holds for $\alpha_a \to 0$ with the result that $A_{ad} = I$, $C_{ad} = T_f I$ and $C_{ad} \bar{H} = H$. The LQG controller (36) is:

$$\begin{aligned}\hat{\mathbf{z}}_{k+1} &= \mathbf{A}_{cf} \hat{\mathbf{z}}_k + \mathbf{B}_{cf} \tilde{y}_k \\ \tilde{a}_k &= \mathbf{C}_{cf} \hat{\mathbf{z}}_k + \mathbf{D}_{cf} \tilde{y}_k \end{aligned} \tag{46}$$

where

$$\mathbf{A}_{cf} = \begin{bmatrix} 0 & D_m K_3 \\ L_{f3} H & I - L_{f3} H \end{bmatrix} \quad \mathbf{B}_{cf} = \begin{bmatrix} D_m K_3 L_{f3} \\ \left(I - L_{f3} H\right) L_{f3} \end{bmatrix} \tilde{y}_k$$

$$\mathbf{C}_{cf} = K_3 \begin{bmatrix} 0 & I \end{bmatrix} \qquad \mathbf{D}_{cf} = K_3 L_{f3} \tilde{y}_k$$

When the reconstructed wavefront is orthogonal to the actuators, the controller cannot affect the performance of the AO system. The structure of the controller is similar to that described in the preceding section. To avoid a redundant discussion, it will also be assumed that the actuator effectiveness matrix D_m is invertible. Define the state transformation

$$\hat{\mathbf{z}}_k = \begin{bmatrix} I & 0 \\ -L_{f3} H & I \end{bmatrix} \bar{\mathbf{z}}_k \tag{47}$$

Application of (47) to the LQG controller (46) yields:

$$\begin{aligned}\bar{\mathbf{z}}_{k+1} &= \mathbf{A}_{cf0} \bar{\mathbf{z}}_k + \mathbf{B}_{cf0} L_{f3} \tilde{y}_k \\ \tilde{a}_k &= \mathbf{C}_{cf0} \hat{\mathbf{z}}_k + \mathbf{D}_{cf0} \tilde{y}_k \end{aligned} \tag{48}$$

where

$$\mathbf{A}_{cf0} = \begin{bmatrix} -L_{f3} H & 0 \\ L_{f3} H & I \end{bmatrix} \quad \mathbf{B}_{cfo} = \begin{bmatrix} L_{f3} H \\ I - L_{f3} H \end{bmatrix} L_{f3}$$

$$\mathbf{C}_{cf0} = D_m^{-1} \begin{bmatrix} 0 & I \end{bmatrix} \qquad \mathbf{D}_{cf0} = D_m^{-1} L_{f3}$$

The transfer function for the controller is

$$\begin{aligned}\tilde{G}_K(z) &= \mathbf{C}_{cf0}\left(zI-\mathbf{A}_{cf0}\right)^{-1}\mathbf{B}_{cf0}+\mathbf{D}_{cf0}\\ &= D_m^{-1}\left(\begin{bmatrix}0 & I\end{bmatrix}\begin{bmatrix}\left(zI+L_{f3}H\right)^{-1} & 0\\ \dfrac{1}{z-1}L_fH\left(zI+L_{f3}H\right)^{-1} & \dfrac{1}{z-1}I\end{bmatrix}\right.\\ &\quad\left.\times\begin{bmatrix}L_{f3}H\\ I-L_{f3}H\end{bmatrix}+I\right)L_{f3}\end{aligned} \tag{49}$$

A bit of algebra yields

$$\tilde{G}_K = D_m^{-1}\left(\frac{z}{z-1}\right)\left(z\left(zI+L_{f3}H\right)^{-1}\right)L_{f3} \tag{50}$$

The controller transfer function shows that the Kalman gain L_{f3} again serves to reconstructed the wavefront from the outputs of the WFS. The reconstructed wavefront is then filtered with a bi-proper transfer function that has m zeros at the origin and m poles at the eigenvalues of $-L_{f3}H$ – a multivariable generalization of a lead filter. The reconstructed and filtered wavefront is integrated and applied to the DM.

4.4. SISO Analysis

A single-input, single output (SISO) control design results when the component models are of a single mode of an AO system (c.f. [4] and [19]). SISO models of the AO components can also be used when it can be assumed that the incident wavefront and DM are uniform across the aperture. It will also be assumed that the IWF dynamics can be represented by a first-order system. The DM model is

$$s_m(t) = a(t) \Rightarrow D_m = 1 \tag{51}$$

Then, the incident wavefront model and geometry matrix are is:

$$\begin{aligned}&\dot{x}_a(t) = -\alpha_a x_a(t) + B_a\zeta(t)\\ &\phi^{tur}(t) = x_a(t)\end{aligned} \Rightarrow \begin{aligned}&A_a = -\alpha_a,\ C_a = 1,\\ &\mathbf{B}_a = B_a\end{aligned}$$
$$H = 1 \tag{52}$$

That is, the IWF transfer function is

$$G_a(s) = \frac{B_a}{s+\alpha_a} \tag{53}$$

The LQ gain (28) becomes:

$$\mathbf{K} = \left[\begin{array}{cc|cc}0 & 0 & \gamma_d & 0\end{array}\right] \tag{54}$$

where $\alpha_d = e^{-\alpha_a T_f}$ and $\gamma_d = (1-\alpha_d)/\left(T_f\alpha_a\right)$. The innovations and Kalman gains (29) become:

$$\mathbf{L} = \begin{bmatrix}0\\0\\ \alpha_d l_3\\ \gamma_d l_3\end{bmatrix} \quad \mathbf{L}_f = \begin{bmatrix}0\\0\\ l_3\\ l_4\end{bmatrix} \tag{55}$$

Define the normalized delay $\tau_n = \tau/T_f$ and its normalized complement $\beta_n = \beta/T_f = 1-\tau_n$. The LQG controller (36) is:

$$\begin{aligned}\hat{\mathbf{z}}_k &= \mathbf{A}_c\hat{\mathbf{z}}_k + \mathbf{B}_c\tilde{y}_k\\ \tilde{a}_k &= \mathbf{C}_c\hat{\mathbf{z}}_k + \mathbf{D}_c\tilde{y}_k\end{aligned} \tag{56}$$

where

$$\begin{aligned}\mathbf{A}_c &= \begin{bmatrix}\gamma_d\beta_n l_3 & \gamma_d\\ \left(\beta_n\left(\alpha_d-\gamma_d l_3\right)+\tau_n\right)l_3 & \alpha_d-\gamma_d l_3\end{bmatrix}\\ \mathbf{B}_c &= \begin{bmatrix}\gamma_d l_3\\ \left(\alpha_d-\gamma_d l_3\right)l_3\end{bmatrix}\\ \mathbf{C}_c &= \begin{bmatrix}\gamma_d\beta_n l_3 & \gamma_d\end{bmatrix}\\ \mathbf{D}_c &= \gamma_d l_3\end{aligned} \tag{57}$$

The transfer function of the controller can be computed as:

$$\begin{aligned}\tilde{G}_K(z) &= \mathbf{C}_c\left(z\mathbf{I}-\mathbf{A}_c\right)^{-1}\mathbf{B}_c+\mathbf{D}_c\\ &= \frac{\gamma_d l_3 z^2}{z^2+\left(\alpha_d-\gamma_d l_3\tau_n\right)z-\gamma_d l_3\tau_n}\end{aligned} \tag{58}$$

This second order system will have one positive pole and one negative pole on the real axis. The positive pole will be approximately at the pole of the IWF model. The second origin zero and the negative pole provide lead to compensate for the computational delay.

When the computational delay is zero, the controller is a 1st order lag with dynamics equal to the IWF model:

$$\tilde{G}_K(z) = \gamma_d l_3\left(\frac{z}{z-\alpha_d}\right) \tag{59}$$

When the IWF dynamic model can be approximated by an integrator ($\alpha_a T_f \ll 1$), the transfer function in (59) can be approximated by a discrete-time integrator ($\alpha_a - 0$):

$$\tilde{G}_K(z) = l_3\left(\frac{z}{z-1}\right) \tag{60}$$

For non-zero computational delays, and $\alpha_a = 0$ the transfer function in (58) is:

$$\tilde{G}_K(z) = l_3\left(\frac{z}{z-1}\right)\left(\frac{z}{z+l_3\tau_n}\right) \tag{61}$$

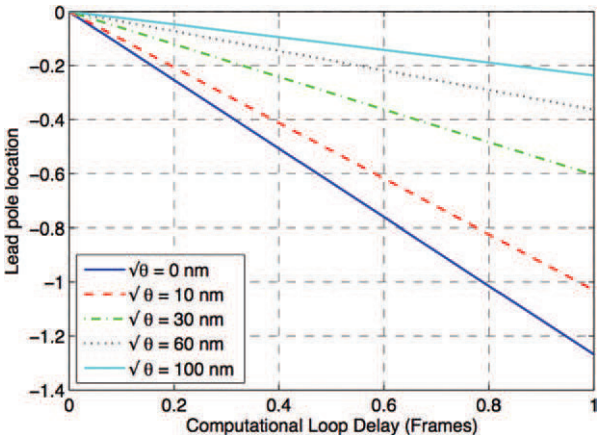

Fig. 4. Lead pole location as a function of computational delay and measurement noise.

This LQG controller can be viewed as a product of three terms. The first term is the KBF gain for the first incident wavefront state. The second term is a discrete-time integrator. The first two terms constitute the commonly used integral control law.

The third term is a lead filter that compensates for the lag of the computational delay. The filter consists of a zero at the origin for all delays and measurement noises. As the delay increases, the pole of the lead filter moves away from the origin along the negative real axis. The motion of the pole is linear is governed by the gain of the controller. The relationship between the lead pole, the computational delay, and the signal-to-noise ratio (governed here by the measurement noise) is illustrated in Fig. 4. Note that for low signal-to-noise ratios and for delays that are a significant portion of the frame, the magnitude of the lead controller pole can become greater than unity. Thus, the LQG controller is unstable, although the closed-loop system remains stable.

5. Summary

The LQG controller for an AO system is the controller that minimizes frame-averaged the residual wavefront variance. When the DM has no continuous-time dynamics, the LQG controller is given by (31) and is the projection $\left(\left(D_m^T D_m\right)^{-1} D_m^T\right)$ of the average IWF that is predicted to occur at the time the control is applied $\left(\frac{1}{T} C_{ad} e^{A_a \tau_c} \hat{x}_3[k|k]\right)$.

For zero computational delay and when the PSD of the incident wavefront is proportional to f^{-2} at all significant frequencies ($\alpha_a T_f \ll 1$), the LQG controller is the commonly used integral controller. The amplitude of the IWF disturbance and the measurement noise determine the gain of the controller. In many cases of interest in astronomical applications, the computational delay is a significant fraction of the frame time. The corresponding LQG controller produces improved performance by additional filtering of the reconstructed wavefront. When the AO system can be represented as SISO modules, this additional filtering is a traditional lead filter consisting of a zero at the origin and a pole on the negative real axis.

References

1. Conan J-M, Rousset G, Madec P-Y. Wave-front temporal spectra in high-resolution imaging through turbulence. *JOSA A* 1995; 12: 1559–1570.
2. Correia C, Raynaud H-F, Kulcsár C, Conan J-M. Globally optimal minimum-variance control in adaptive optics systems with mirror's dynamics. *SPIE Astron Instrum* Marseille 2008; 7015.
3. Correia C, Raynaud H-F, Kulcsár C, Conan J-M. Accounting for mirror dynamics in optical adaptive optics control. *ECC'09* Budapest, 2009.
4. Ellerbroek BL, Van Loan C, Pitsianis NP, Plemmons RJ. Optimizing closed-loop adaptive-optics performance with use of multiple control bandwidths. *JOSA A* 1994; 11: 2871–1886.
5. Hardy JW. *Adaptive Optics for Astronomical Telescopes.* Oxford, New York, 1998.
6. Hinnen K, Verhaegen M, Doelman N. Exploiting the spatio-temporal correlation in adaptive optics using data-driven H2 – optimal control. *JOSA A* 2007; 24: 1714–1725.
7. Ishimaru A. *Wave Propagation and Scattering in Random Fields.* Academic Press, London, 1978.
8. Krantz SG. Runge's Theorem. In *Handbook of Complex Variables.* Birkhäuser, Boston, 1999, pp. 143–144.
9. Kolmogorov AN. The local structure of turbulence in incompressible viscous fluids for very large Reynolds' numbers. In *Turbulence, Classical Papers on Statistical Theory*, Friedlander SK, Topper L (eds.). Wiley-Interscience, 1961, pp. 151–155.
10. Kulcsár C, Raynaud H-F, Petit C, Conan J-M, de Lesegno PV. Optimal control, observers, and integrators in adaptive optics. *Opt Exp* 2006; 14: 7464–7476.
11. Le Roux B, Conan J-M, Kulcsár C, Raynaud H-F, Mugnier LM, Fusco T. Optimal control law for multi-conjugate adaptive optics. *JOSA A* 2004; 21: 1261–1276.
12. Looze DP, Kasper M, Hippler S, Beker O, Weiss R. Optimal compensation and implementation for adaptive optics systems. *Exp Astron* 2003; 15: 67–88.
13. Looze DP. Minimum variance control structure for adaptive optics systems. *JOSA A* 2006; 23: 603–612.
14. Looze DP. Discrete-time model of an adaptive optics system. *JOSA A* 2007; 24: 2850-2863.
15. Looze DP. Discrete-time model for an adaptive optics system with input delay. *Int J Control* 2010; 83(6): 1217–1231.
16. Paschall RN, Anderson DJ. Linear quadratic Gaussian control of a deformable mirror adaptive optics system with time-delayed measurements. *Appl Opt* 1993; 32: 6347–6358.
17. Petit C, Conan J-M, Kulcsár C, Raynaud H-F, Fusco T, Montri J, Rabaud D. Optimal control for multi-conjugate adaptive optics. *C R Acad Sci I Math* 2005; 6: 1059–1069.
18. Petit C, Conan J-M, Kulcsár C, Raynaud H-F, Fusco T, Montri J, Rabaud D. First laboratory demonstration of closed-loop Kalman based optimal control for vibration

filtering and simplified MCAO. *Proc SPIE* 2006; 6272: 62721T-1–62721T-12.
19. Poyneer LA, Véran J-P. Optimal modal Fourier-transform wavefront control. *JOSA A* 2005; 22: 1515-1526.
20. Roddier F. *Adaptive Optics in Astronomy.* Cambridge University Press, Cambridge, 1999.
21. Roggemann MC, Welsh BM. *Imaging Through Turbulence.* CRC Press, Boca Raton, 1996.
22. Taylor GI. The spectrum of turbulence. *Proc R Soc Lond A* 1938; 164: 476–490.
23. Wiberg DM, Max CE, Gavel DT. A spatial non-dynamic LQG controller: Part II, theory. *Proceedings of 2004 IEEE Conference on Decision and Control.* Atlantis, Paradise Island, Bahamas, 2004, pp. 3333–3338.