# Fast computation of an optimal controller for large-scale adaptive optics

**Paolo Massioni,[1,2,*] Caroline Kulcsár,[1] Henri-François Raynaud,[1] and Jean-Marc Conan[2]**

[1]*Institut Galilée, L2TI, Université Paris 13, 93430 Villetaneuse, France*
[2]*ONERA, BP 72, 92322 Châtillon Cedex, France*
**Corresponding author: massioni@univ-paris13.fr*

Received May 31, 2011; revised September 12, 2011; accepted September 23, 2011;
posted September 26, 2011 (Doc. ID 148274); published October 19, 2011

The linear quadratic Gaussian regulator provides the minimum-variance control solution for a linear time-invariant system. For adaptive optics (AO) applications, under the hypothesis of a deformable mirror with instantaneous response, such a controller boils down to a minimum-variance phase estimator (a Kalman filter) and a projection onto the mirror space. The Kalman filter gain can be computed by solving an algebraic Riccati matrix equation, whose computational complexity grows very quickly with the size of the telescope aperture. This "curse of dimensionality" makes the standard solvers for Riccati equations very slow in the case of extremely large telescopes. In this article, we propose a way of computing the Kalman gain for AO systems by means of an approximation that considers the turbulence phase screen as the cropped version of an infinite-size screen. We demonstrate the advantages of the methods for both off- and on-line computational time, and we evaluate its performance for classical AO as well as for wide-field tomographic AO with multiple natural guide stars. Simulation results are reported. 

*OCIS codes:* 010.1080, 010.1330, 010.7350.

## 1. INTRODUCTION

Adaptive optics (AO) [1] is a technique that allows the real-time correction of the atmospheric turbulence effects on image formation. AO systems can often be considered as linear, and so they are amenable to linear optimal control, where the criterion to be minimized is the variance of the residual phase. If the temporal correlation of the turbulence is neglected, then the optimal solution is given by the minimum mean square error estimator (MMSE), which has been extensively studied in the literature; see, for example, [2,3] and references therein. On the other hand, if the temporal correlations are considered, then the MMSE estimator is not optimal anymore; it has been used however, for example, in [4], to build a suboptimal scheme adapted to large dimensions. The optimal solution is known to be obtained with the linear quadratic Gaussian (LQG) regulator [5]. The first step in the real-time operations of such a controller is the reconstruction of the wavefront with the use of a Kalman filter [6], which allows estimation of the phase from the measurements while attenuating the effects of measurement noise. The efficiency of the Kalman filter is related to the chosen state-space model, for which different possibilities have been investigated in the literature. A simple but valuable possibility is the autoregressive model of order 1 (AR1) [7], which can be implemented either in a zonal or Zernike basis.

The use of a Kalman filter involves off-line computations, to be executed only once before the beginning of the operations, and on-line computations, to be executed at each time step during operational time. The off-line computations are those involved in the calculation of a matrix, the Kalman gain, which requires solving the associated discrete algebraic Riccati equation (DARE) [6]. The complexity of this off-line computation is proportional to the third power of the number $n$ of states of the system, or $\mathcal{O}(n^3)$. If we consider that for a zonal basis the number of states grows with the square of the diameter $D_t$ of the telescope pupil, then the computational cost of finding the Kalman gain is $\mathcal{O}(D_t^6)$. In this article, we introduce a computationally simple procedure for finding the Kalman gain of AO systems with a large number of degrees of freedom described on a zonal basis. This problem has been addressed in [8] for classical AO systems using a sparse approximation of the turbulence covariance matrix, which avoids the resolution of a Riccati equation. The innovation that we introduce here is that instead of exploiting the sparsity of the matrices involved in the computations, we follow a distributed control approach, a technique that recently is being more and more investigated for application in AO [9,10]. Here we focus on the approach developed in [11], which has already seen application in a number of engineering fields, like flow control [12], paper production plants [13], and active optics on large segmented telescopes [14]. The method is based on a Fourier transform that is used to decompose an infinite-order system into an infinite set of low, finite-order ones.

Following this idea, we introduce a method that approximates the Kalman gain by assuming the telescope pupil as the cropped version of an infinite-size phase screen. The infinite-size Kalman filtering problem is amenable to a special treatment that makes the computational cost of finding the gain not dependent on the number of states, so it is possible to get a quick approximation of this gain in a very short time for any size. This method is significantly different from the Fourier-based optimal control in [15], where the on-line wavefront reconstruction is executed on Fourier-transformed versions of the measurements (which requires extending the circular pupil to a square periodic support). In this work, the Fourier transform (on an infinite, nonperiodic support)

 

is used as an off-line tool that simplifies the computations, but the resulting Kalman filter works directly on the measurements and no Fourier transform is executed on-line. We will discuss the off-line and on-line computational complexity, both for the classical and tomographic AO cases, and we will show through simulations that the losses in terms of the Strehl ratio due to this approximation are quite limited.

The article is organized as follows. Section 2 introduces the modeling of the AO system and the control problem, and Section 3 describes the new procedure that allows the fast computation of the Kalman filter gain. Section 4 shows the extensions of the method to wide-field tomographic AO, to different turbulence models and to the computation of the inverse of the influence matrix. Section 5 contains a discussion on the on-line computational times, while Section 6 presents the results obtained in simulation as well as performance evaluations, both theoretical and empirical. Some conclusions are proposed in Section 7.

## 2. OPTIMAL CONTROL PROBLEM

The LQG controller is model based; therefore, its implementation requires a state-space model of the physical system. For an AO system, the model includes the dynamics of the turbulence and the characteristics of the sensors and the actuators. The turbulence is a phenomenon that has its natural description in continuous time; as in the practice the control system is digital with a finite sampling time $T$, we choose to use a discrete-time model of the AO system. All the discrete-time variables that will appear in this paper have to be considered as the time averages of their continuous-time counterparts; this means that a generic discrete-time variable $x(k)$ is defined by

$$x(k) = \frac{1}{T}\int_{(k-1)T}^{kT} x_c(t)\mathrm{d}t, \tag{1}$$

where $x_c(t)$ is the related continuous-time variable.

Different models of turbulence have been proposed in the literature; for the sake of simplicity, we focus here on discrete-time AR1 on a zonal basis [8,16,17]. The procedure that we introduce for the fast approximation of the Kalman gain matrix is thus applied to an AR1 model, but the procedure itself is not limited to this kind of model and can be extended to any linear model, e.g., autoregressive models of higher order. If we consider a single turbulent layer, and we call $\phi(k)$ the vector containing the phases of the turbulent screen on a number of points in the area of interest (the pupil) at a time $k$, then the evolution of the turbulence is described by the following equation:

$$\phi(k+1) = A\phi(k) + v(k), \tag{2}$$

where $v(k)$ is the process noise and $A$ is a stability matrix (i.e., a matrix whose eigenvalues are inside the unit circle). A common assumption for AR1 models as in Eq. (2) is that $A$ is a diagonal matrix with the same value $a_{\mathrm{tur}}$ in the diagonal [8,16,17], due to the fact that the turbulence is a spatially homogeneous process. According to von Kármán's law, the turbulence is a stationary stochastic process whose statistical properties are known as a function of the Fried parameter $r_0$ and the outer scale $L_0$; namely, the power spectral density of the phase is

$$\sigma_\phi^2(\nu) = 0.023 r_0^{-\frac{5}{3}}\left(|\nu|^2 + \frac{1}{L_0^2}\right)^{-\frac{11}{6}}, \tag{3}$$

where $\nu$ is the bidimensional spatial frequency. From such a model we can compute the covariance matrix $\Sigma_\phi$ of the turbulent phase vector, which is a full matrix. As a consequence of Eq. (2) the covariance matrix $\Sigma_v$ that is consistent with a given $\Sigma_\phi$ is simply

$$\Sigma_v = (1 - a_{\mathrm{tur}}^2)\Sigma_\phi. \tag{4}$$

The goal of the AO control system is to compensate for the phase distortion due to turbulence, by inserting in the optical path a deformable mirror (DM) generating a correction phase $\phi^{(\mathrm{corr})}$ that partially cancels the distortion, making sure that the residual phase

$$\phi^{(\mathrm{res})}(k) = \phi(k) - \phi^{(\mathrm{corr})}(k) \tag{5}$$

has a variance as close as possible to zero. The DM compensates for the distorted wavefront with local displacements of its surface due to an array of distributed actuators. The displacement caused in the surface by applying a given voltage to one actuator can be assumed to be a kind of bell-shaped function, for example, a Gaussian or double Gaussian. We assume that the response of the mirror has no dynamics, so that it achieves the desired shape immediately. In matrix form and discrete time, a set of voltages $u(k)$ generates a phase correction $\phi^{(\mathrm{corr})}(k)$ according to

$$\phi^{(\mathrm{corr})}(k) = Nu(k-1), \tag{6}$$

where $N$ is the influence matrix of the DM.

At last, the model includes the measurement equation, which depends on the kind of wavefront sensor (WFS) used by the AO system. In general the measurement $s(k)$ is given by

$$s(k) = D\phi(k-1) + w(k), \tag{7}$$

where $D$ is a matrix that emulates the effect of the WFS and $w(k)$ is the measurement noise, which is assumed to be uncorrelated for each output channel with variance $\sigma_w^2$, white and Gaussian. Notice the presence of a one-step delay here as well.

Minimum-variance control is obtained using the separation principle [18,19] under the form

$$u(k) = N^\dagger \Pi \hat{\phi}(k+1|\mathcal{S}_k), \tag{8}$$

where the minimum-variance estimate $\hat{\phi}(k+1|\mathcal{S}_k) = E(\phi(k+1)|\mathcal{S}_k)$ is the conditional expectation of the phase at time $k+1$ based on the available information at time $k$, which is the set $\mathcal{S}_k$ of all past measurements, $\mathcal{S}_k = \{s(k), s(k-1), \ldots, s(0)\}$. The control gain matrix $N^\dagger = (N^T N)^{-1} N^T$ is the pseudoinverse of $N$; $\Pi$ is a projection matrix that removes the piston term from the phase, as it is not necessary to compensate for it because it affects neither the image quality nor the measurements.

The minimum-variance estimate $\hat{\phi}(k+1|\mathcal{S}_k)$ is obtained as the output of a Kalman filter, derived from a state-space model of the AO system. Such a model is built here from

the turbulence model in Eq. (2) and the measurement model in Eq. (7), which features a one-step delay that has to be taken into account. Nevertheless, the computation of the Kalman gain, which is the object of this article, is the same regardless of the presence of the delays. So if we define the output $y(k-1)=s(k)$ and $\mathcal{Y}_k=\{y_k, y_{k-1}, \ldots, y_{-1}\}$ then the measurement delay disappears and we can compute the gain in this case; the obtained filter can be used to compute $\hat{\phi}(k|\mathcal{Y}_{k-1})=E(\phi(k)|\mathcal{Y}_{k-1})$ iteratively (with $\mathcal{Y}_{k-1}=\mathcal{S}_k$). Then we can retrieve the optimal prediction $\hat{\phi}(k+1|\mathcal{S}_k)$ using the identity

$$\hat{\phi}(k+1|\mathcal{S}_k)=\hat{\phi}(k+1|\mathcal{Y}_{k-1})=A\hat{\phi}(k|\mathcal{Y}_{k-1}), \tag{9}$$

which is a simple propagation of the estimate $\hat{\phi}(k|\mathcal{Y}_{k-1})$ to the following time step. The recursive formulation of the Kalman filter is then

$$\begin{cases}\hat{\phi}(k+1|\mathcal{Y}_k)=A\hat{\phi}(k|\mathcal{Y}_{k-1})+K(y(k)-\hat{y}(k|\mathcal{Y}_{k-1})),\\ \hat{y}(k|\mathcal{Y}_{k-1})=D\hat{\phi}(k|\mathcal{Y}_{k-1}),\end{cases} \tag{10}$$

where $K$ is the Kalman gain, whose off-line computation is the main subject of this article.

In the remainder of the paper, for any variable $x$ the conditional expectation $\hat{x}(k|\ell)$ will refer to the information $\mathcal{Y}_\ell$, so, for example, $\hat{\phi}(k+1|k)=\hat{\phi}(k+1|\mathcal{Y}_k)$. Using this notation, Eq. (8) can be rewritten as

$$u(k)=N^{\dagger}\Pi\hat{\phi}(k+1|k-1). \tag{11}$$

In this work we have assumed an open-loop situation, where the correction phase term does not appear in the output equation [Eq. (7)], but the whole reasoning is applicable to the closed-loop case as well (as the gain computation is the same in the open or closed loop).

## 3. FAST COMPUTATION OF THE KALMAN FILTER GAIN

Let us now consider a turbulent phase screen of infinite size, sampled at regular intervals of length $\Delta x$ on a square lattice of integer coordinates $x_1, x_2$. We assume that the size of the segments of the phase screen associated to each of the WFS subapertures and the space between each actuator are $\Delta x$ as well. We can then express the phase and all the other spatially distributed variables (noises, outputs) as depending not only on the time $k$ but also on the spatial coordinates. This means that Eq. (2) and (7) will take the following form:

$$\begin{cases}\phi_{x_1,x_2}(k+1)=a_{\text{tur}}\phi_{x_1,x_2}(k)+v_{x_1,x_2}(k),\\ y_{x_1,x_2}(k)=F_{x_1,x_2}(\phi(k))+w_{x_1,x_2}(k),\end{cases} \tag{12}$$

where the dependence of variables from the spatial coordinates is shown in the subscripts, and $F$ is the equivalent of the matrix operator $D$ as a function in the spatial coordinates we have just introduced. For a Shack–Hartmann sensor with Fried geometry (Hudgin or Southwell could be used as well), the measurement is given by the integral of the gradient of the phase over each of the sensor's subapertures. The geometry featured by the sensor is shown in Fig. 1. The output of this WFS is approximated in $F$ by means of finite differences:

$$\begin{aligned}&F_{x_1,x_2}(\phi(k))\\ &\quad=\frac{1}{2}\begin{bmatrix}\phi_{x_1+1,x_2}(k)+\phi_{x_1+1,x_2+1}(k)-\phi_{x_1,x_2}(k)-\phi_{x_1,x_2+1}(k)\\ \phi_{x_1,x_2+1}(k)+\phi_{x_1+1,x_2+1}(k)-\phi_{x_1,x_2}(k)-\phi_{x_1+1,x_2}(k)\end{bmatrix},\end{aligned} \tag{13}$$

where the gradient at each subaperture is given as a linear function of the phase at the four corners. For the remainder of the article, we are going to focus only on this model of Shack–Hartmann sensor, but the procedure can be extended to different models or different sensors, as long as their function $F$ is spatially invariant. The output $y_{x_1,x_2}(k)$ is a two-entry array for each spatial location, and it is computed with finite differences of the phase as in Eq. (13). The model in Eq. (12) describes the evolution of the turbulence at each discretization point; if the dynamics of each point depended only on variables at the same point, then we could write a Kalman filter for reconstructing the phase for each point alone, without the necessity of any cross-talks among points. In fact this is not possible because of two forms of interaction. First, from Eq. (13) we can see the computation of the gradient at one point requires the knowledge of the phase in four points; and second, the process noise variables $v_{x_1,x_2}(k)$ are not uncorrelated at each spatial location; i.e., $E[v_{x_1,x_2}(k)v_{x_1',x_2'}(k)]\neq 0$ for all $x_1, x_2, x_1', x_2'$. We can then look at Eq. (12) as an infinite set of first order cross-coupled state-space systems.

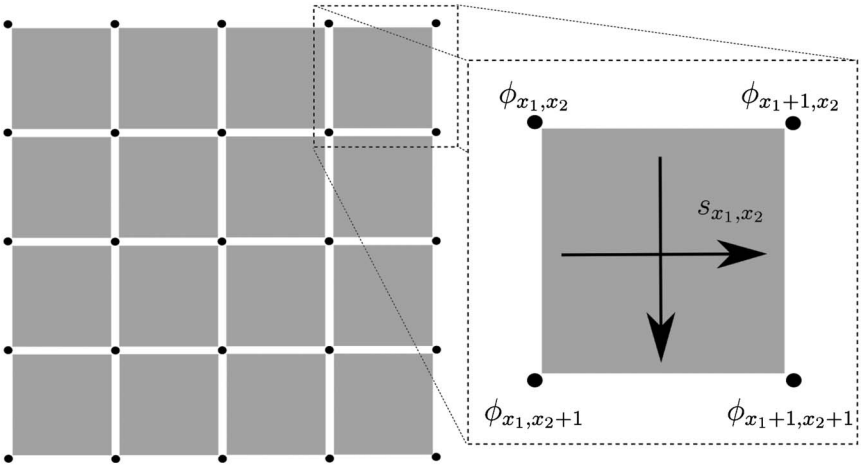


Fig. 1. Fried geometry. The black dots represent the points where the phase is reconstructed, and the gray areas represent the lenslets or subapertures of the sensor. The actuators' positions correspond to the black dots as well.

However, the equations in Eq. (12) feature the property of being *spatially invariant*. This means that the dynamics of the system does not change with respect to spatial shifts of any value $(d_1, d_2)$. This is clear just by inspection of the equations, which are the same at each point and because of the spatial stationarity of the turbulence, which implies that the related correlation functions are spatially invariant: $E[\phi_{x_1,x_2}(k)\phi_{x_1',x_2'}(k)]=E[\phi_{x_1+d_1,x_2+d_2}(k)\phi_{x_1'+d_1,x_2'+d_2}(k)]$. This, with Eq. (4), implies $E[v_{x_1,x_2}(k)v_{x_1',x_2'}(k)] = E[v_{x_1+d_1,x_2+d_2}(k)\, v_{x_1'+d_1,x_2'+d_2}(k)]$. Equations with this property are amenable to a diagonal Fourier decomposition [11], which means that if state, input, and output variables are replaced by their Fourier-transformed version, then the system is equivalent to an infinite set of decoupled state-space models, where each equation can be considered on its own. The transform to be used here is the discrete-time Fourier transform (DTFT) [20,21], which for a function $\gamma(x)$ with $x \in \mathbb{R}$ is defined as

$$\Gamma(\nu) = \sum_{n=-\infty}^{+\infty} \gamma(n\Delta x)\exp(-2\pi j \Delta x n \nu), \tag{14}$$

where $\nu \in [-1/(2\Delta x), 1/(2\Delta x)]$ is a spatial frequency variable, $\Delta x$ is the sampling distance, and $j$ is the imaginary unit. The inverse transform is given by the following formula, which recovers the values of $\gamma$ at the sampling points:

$$\gamma(n\Delta x) = \Delta x \int_{-1/(2\Delta x)}^{1/(2\Delta x)} \Gamma(\nu)\exp(2\pi j \Delta x n \nu)\mathrm{d}\nu. \tag{15}$$

By applying the DTFT for both spatial directions of the infinite pupil, the equations in the spatial variables $x_1, x_2 \in \mathbb{Z}$ become equations in the frequency variables $\nu_1,\ \nu_2 \in [-1/(2\Delta x), 1/(2\Delta x)]$; moreover, shifting the original variable of one unit in the spatial domain is equivalent to a multiplication by the frequency-dependent terms $X_1 = \exp(-2\pi j \Delta x \nu_1),\ X_2 = \exp(-2\pi j \Delta x \nu_2)$. This means, if we indicate the Fourier-transformed version of the variables with capital letters, that

$$\begin{matrix} \phi_{x_1+1,x_2}(k) \overset{\text{DTFT}}{\leftrightarrow} X_1 \Phi_{\nu_1,\nu_2}(k) \\ \phi_{x_1-1,x_2}(k) \overset{\text{DTFT}}{\leftrightarrow} X_1^{-1} \Phi_{\nu_1,\nu_2}(k) \end{matrix}, \tag{16}$$

and the same holds for the other spatial variable. Then Eq. (12) becomes

$$\begin{cases} \Phi_{\nu_1,\nu_2}(k+1) = a_{\text{tur}}\Phi_{\nu_1,\nu_2}(k) + V_{\nu_1,\nu_2}(k), \\ Y_{\nu_1,\nu_2}(k) = \frac{1}{2}\begin{bmatrix} X_1 + X_1X_2 - 1 - X_2 \\ X_2 + X_1X_2 - 1 - X_1 \end{bmatrix}\Phi_{\nu_1,\nu_2}(k) + W_{\nu_1,\nu_2}(k). \end{cases} \tag{17}$$

Each equation in the set depends only on one value of the couple of frequency variables $\nu_1, \nu_2$; moreover, the autocorrelation of $V_{\nu_1,\nu_2}(k)$ is directly given by the von Kármán power spectrum law in Eq. (3) times a correction factor as in Eq. (4):

$$\begin{aligned} E[|V_{\nu_1,\nu_2}(k)|^2] &= (1-a_{\text{tur}}^2)\sigma_\phi^2(\nu_1,\nu_2) \\ &= (1-a_{\text{tur}}^2)\left(0.023 r_0^{-\frac{5}{3}}\left((\nu_1^2+\nu_2^2)+\frac{1}{L_0^2}\right)^{-\frac{11}{6}}\right), \end{aligned} \tag{18}$$

and the covariance $E[V_{\nu_1,\nu_2}(k)V_{\nu_1',\nu_2'}(k)]$ for $\nu_1 \neq \nu_1'$ and $\nu_2 \neq \nu_2'$ is null due to the spatial stationarity of the turbulence model [22]. In writing Eq. (18) we have neglected the spatial aliasing of the continuous spectrum due to spatial discretization. We have thus a formulation of the problem where the dynamics of each frequency-dependent variable is independent from the others; nevertheless, this is not yet a significant achievement, as the real AO system is going to read $y_{x_1,x_2}(k)$ and not $Y_{\nu_1,\nu_2}(k)$ from its WFS. We can remark here again the difference between this approach and the one in [15], which has been already briefly discussed in the introduction. In [15], the wavefront estimation algorithms work on the Fourier-transformed variables, and a discrete Fourier transform (DFT) is used due to the limited spatial extension of the measured wavefront. This leads to an underlying hypothesis of periodicity. Here we use a DTFT, with an underlying hypothesis of infinite supports, and this is only an intermediate step for computing the Kalman filter gain that will work directly on the measurements $y_{x_1,x_2}(k)$ and not on $Y_{\nu_1,\nu_2}(k)$. However, both transforms would formally lead to the same system equations [Eq. (17)], with variables defined in a different way. Avoiding the online use of the DFT allows us to bypass the problems linked to the fact that the support is not square but round, and all the issues concerning how to deal with the boundary conditions and "fill in" the empty points of the grid.

The implementation of a state-feedback control system requires the estimation of the phase from the measurements, for which the Kalman filtering provides the optimal solution in the case of Gaussian linear models. The steady-state Kalman filter gain can be found through the solution of the DARE, which corresponds for the Fourier-transformed model in Eq. (17) to a set of frequency-dependent scalar equations:

$$\begin{aligned} P_{\nu_1,\nu_2} &= a_{\text{tur}}^2 P_{\nu_1,\nu_2} + (1-a_{\text{tur}}^2)\sigma_\phi^2(\nu_1,\nu_2) \\ &\quad - a_{\text{tur}}^2 P_{\nu_1,\nu_2}^2 C_{\nu_1,\nu_2}^H (P_{\nu_1,\nu_2} C_{\nu_1,\nu_2} C_{\nu_1,\nu_2}^H + \sigma_w^2 I_2)^{-1} C_{\nu_1,\nu_2}, \end{aligned} \tag{19}$$

where $P_{\nu_1,\nu_2}$ is the (scalar) unknown, the superscript $H$ indicates the matrix Hermitian (transpose and complex conjugate), $I_2$ is the identity matrix of order 2, and

$$C_{\nu_1,\nu_2} = \frac{1}{2}\begin{bmatrix} X_1 + X_1X_2 - 1 - X_2 \\ X_2 + X_1X_2 - 1 - X_1 \end{bmatrix} \tag{20}$$

is the output matrix in Eq. (17).

With this we can compute the frequency-dependent Kalman gain

$$K_{\nu_1,\nu_2} = a_{\text{tur}} P_{\nu_1,\nu_2} C_{\nu_1,\nu_2}^H (C_{\nu_1,\nu_2} P_{\nu_1,\nu_2} C_{\nu_1,\nu_2}^H + \sigma_w^2 I_2)^{-1} \tag{21}$$

for the following frequency-dependent Kalman filter:

$$\begin{cases} \hat{\Phi}_{\nu_1,\nu_2}(k+1|k) = a_{\text{tur}}\hat{\Phi}_{\nu_1,\nu_2}(k|k-1) + K_{\nu_1,\nu_2}(Y_{\nu_1,\nu_2}(k) - \hat{Y}_{\nu_1,\nu_2}(k|k-1)), \\ \hat{Y}_{\nu_1,\nu_2}(k) = C_{\nu_1,\nu_2}\hat{\Phi}_{\nu_1,\nu_2}(k), \end{cases} \tag{22}$$

where the hat indicates the estimate of a variable as a function of past measurements; e.g., $\hat{\Phi}_{\nu_1,\nu_2}(k|k-1)$ is the estimate of $\Phi_{\nu_1,\nu_2}(k)$ based on all the measurements $Y_{\nu_1,\nu_2}$ until time $k-1$.

As seen before, any spatially invariant operator in the spatial domain corresponds to a diagonal operator in the frequency domain. The converse is true as well, so the diagonal operator given by the Kalman gain in the frequency domain corresponds to a spatially invariant operator in the spatial domain; the Kalman gain $K_{\nu_1,\nu_2}$ corresponds therefore to a gain $k_{n_1,n_2}$ in the spatial domain that depends only on the relative position between the location of the measurement and the phase point that is being reconstructed. The Kalman filter thus takes the form

$$\begin{cases} \hat{\phi}_{x_1,x_2}(k+1|k) = a_{\text{tur}}\hat{\phi}_{x_1,x_2}(k|k-1) + \sum_{x_1'=-\infty}^{+\infty}\sum_{x_2'=-\infty}^{+\infty} k_{x_1-x_1',x_2-x_2'}(y_{x_1',x_2'}(k) - \hat{y}_{x_1',x_2'}(k|k-1)), \\ \hat{y}_{x_1,x_2}(k|k-1) = F_{x_1,x_2}(\hat{\phi}(k|k-1)). \end{cases} \tag{23}$$

The gain can be reconstructed with the inverse Fourier transform [Eq. (15)]

$$k_{n_1,n_2} = \Delta x^2 \int_{-1/(2\Delta x)}^{1/(2\Delta x)}\int_{-1/(2\Delta x)}^{1/(2\Delta x)} K_{\nu_1,\nu_2}\exp(2\pi j\Delta x n_1\nu_1) \times \exp(2\pi j\Delta x n_2\nu_2)\mathrm{d}\nu_1\mathrm{d}\nu_2. \tag{24}$$

In general we do not have a closed form for solving the Riccati equation, which makes it necessary to use an appropriate solver to get the solution. Nevertheless, for the case we are considering now, $P_{\nu_1,\nu_2}$ in Eq. (19) can be computed directly, as the Riccati equation is scalar, and it is equivalent to a second-order algebraic equation, whose solution is

$$P_{\nu_1,\nu_2} = (1-a_{\text{tur}}^2)\left(\sigma_\phi^2(\nu_1,\nu_2) - c_{\nu_1,\nu_2} + \sqrt{(c_{\nu_1,\nu_2}-\sigma_\phi^2(\nu_1,\nu_2))^2 + 4\frac{\sigma_\phi^2(\nu_1,\nu_2)}{1-a_{\text{tur}}^2}c_{\nu_1,\nu_2}}\right), \tag{25}$$

where $c_{\nu_1,\nu_2} = \sigma_w^2/(C_{\nu_1,\nu_2}^H C_{\nu_1,\nu_2})$, and then $K_{\nu_1,\nu_2}$ can be computed with Eq. (21). The integral in Eq. (24) has then to be approximated with finite sums. For this work, we have evaluated the values of $K_{\nu_1,\nu_2}$ in a fixed $100\times100$ regular grid and used a rectangle approximation; this strategy is faster than using a solver with an adaptive sampling grid, as it does not require the recalculation of $K_{\nu_1,\nu_2}$ in different grid points each time $k_{n_1,n_2}$ is computed for different $n_1, n_2$. The rectangle approximation turns Eq. (24) into the following:

$$k_{n_1,n_2} \approx \frac{1}{M^2}\sum_{m_1=-\frac{M}{2}}^{\frac{M}{2}-1}\sum_{m_2=-\frac{M}{2}}^{\frac{M}{2}-1} K_{\nu_1,\nu_2}\exp\left(\frac{2\pi j n_1 m_1}{M}\right)\exp\left(\frac{2\pi j n_2 m_2}{M}\right), \tag{26}$$

where $M$ is the (even) number of integration steps. Notice that this last equation corresponds to an inverse DFT, i.e., the inverse of a Fourier transform which maps a periodic (or "wrap-around") domain to a discrete frequency set. This observation shows that it is also possible to take a different point of view on the problem: with a DTFT, the finite telescope area is approximated as a finite part of an infinite surface; the approximation of the DTFT as a DFT corresponds to approximating the telescope area as a finite part of a bigger (size $M\times M$), wrap-around fictitious surface.

At this point we should remark that the filter cannot be implemented directly in the form of Eq. (23), as it involves a sum with an infinite number of items. We then need to introduce an approximation assuming that $k_{n_1,n_2}$ is different from zero only for a limited "patch," i.e., a small interval of $n_1, n_2$ centered around 0 and of extension $z$. This assumption is justified by the reasoning in [11], which shows that spatially distributed systems with a level of cross-interaction that decays with the distance originate optimal controllers or observers that have the same property. In addition to this, the *a posteriori* observation of the gain confirms the assumption, as the computations show that the gain has significant values centered around zero, and then it decays as the distance from zero increases. In Fig. 2, the gain $k_{n_1,n_2}$ resulting from typical values of the physical parameters is shown. It is apparent that the gain quickly decays to values close to 0 for growing values of $|n_1|$, $|n_2|$. Once the Kalman gain for the infinite case is obtained, it is possible to get its finite version for the pupil of the desired size just by "cutting" the part that is needed and ignoring the rest. To summarize, in real-time the estimate of the phase will be propagated with the following equation:

$$\hat{\phi}_{x_1,x_2}(k+1|k) = a_{\text{tur}}\hat{\phi}_{x_1,x_2}(k|k-1) + \sum_{x_1'=-z}^{z}\sum_{x_2'=-z}^{z} k_{x_1-x_1',x_2-x_2'}(y_{x_1',x_2'}(k) - \hat{y}_{x_1',x_2'}(k|k-1)), \tag{27}$$

where $z$ is an arbitrary integer number that determines the size of the patch or interval of $n_1$, $n_2$ that we consider and $y_{x_1',x_2'}(k)$ has to be taken as equal to 0 if the coordinates $x_1'$, $x_2'$ indicate a point outside of the pupil. Equation (27) is the recursive equation of the Kalman filter, which is fed by a convolution product (the term with the nested sums). Such a Kalman filter can be verified *a posteriori* to be stable, so the estimated phase will converge for all the modes excluding piston. The values of $\hat{\phi}_{x_1,x_2}(k+1|k-1) = \hat{\phi}_{x_1,x_2}(k+1|\mathcal{S}_k)$ that are needed for the control can then be obtained [as in Eq. (9)] from the relation

$$\hat{\phi}_{x_1,x_2}(k+1|k-1) = a_{\text{tur}}\hat{\phi}_{x_1,x_2}(k|k-1). \tag{28}$$

From $\hat{\phi}_{x_1,x_2}(k+1|k-1)$ the control action can be found in turn according to Eq. (11).

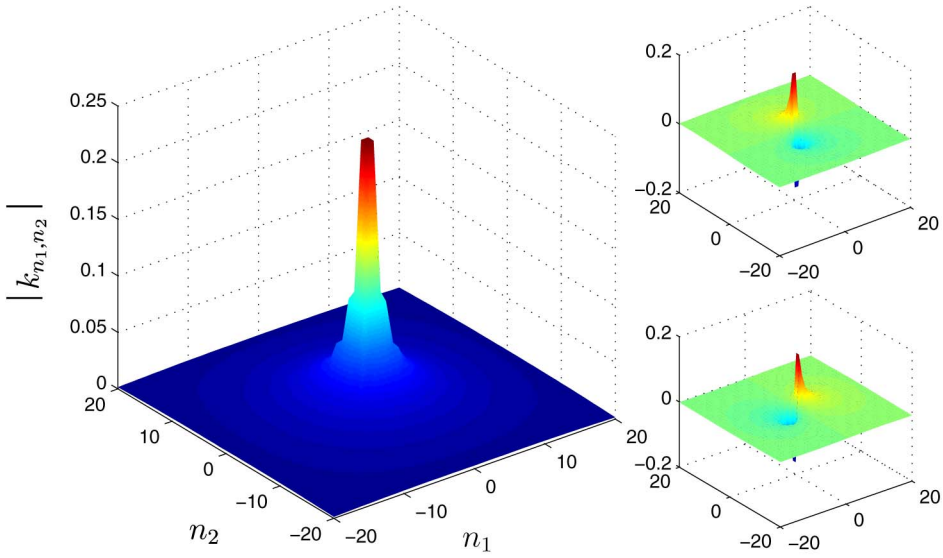


Fig. 2. (Color online) Modulus of the Kalman gain $k_{n_1,n_2}$. The smaller pictures on the right show the real value (not absolute value) of the two components of the gain.

For more discussion about the practice of applying the results of infinite-case computations to the finite case, we invite the interested reader to consult [23,24].

## 4. WIDE-FIELD ADAPTIVE OPTICS AND OTHER EXTENSIONS

### A. Tomographic Wide-Field Adaptive Optics

The wide-field AO case [25] can be treated with the same technique. The state equation will have to consider $n_L$ different layers at different altitudes. Moreover, there will be as many outputs as the number $n_{\text{gs}}$ of guide stars (we consider natural guide stars). We call $h^{(l)}$ the altitude of the $l$th layer, and we call $\alpha_1^{(s)}$ and $\alpha_2^{(s)}$ the angular separations of the $s$th star from the center of the field of view in the two directions. The state-space model will then be given by the following set of equations:

$$\begin{cases} \phi^{(1)}_{x_1,x_2}(k+1) = a^{(1)}_{\text{tur}}\phi^{(1)}_{x_1,x_2}(k) + v^{(1)}_{x_1,x_2}(k), \\ \phi^{(2)}_{x_1,x_2}(k+1) = a^{(2)}_{\text{tur}}\phi^{(2)}_{x_1,x_2}(k) + v^{(2)}_{x_1,x_2}(k), \\ \dots \\ \phi^{(n_L)}_{x_1,x_2}(k+1) = a^{(n_L)}_{\text{tur}}\phi^{(n_L)}_{x_1,x_2}(k) + v^{(n_L)}_{x_1,x_2}(k), \\ y^{(1)}_{x_1,x_2}(k) = \sum_{l=1}^{n_L} F_{x_1+\alpha_1^{(1)}h^{(l)},x_2+\alpha_2^{(1)}h^{(l)}}(\phi^{(l)}(k)) + w^{(1)}_{x_1,x_2}(k), \\ y^{(2)}_{x_1,x_2}(k) = \sum_{l=1}^{n_L} F_{x_1+\alpha_1^{(2)}h^{(l)},x_2+\alpha_2^{(2)}h^{(l)}}(\phi^{(l)}(k)) + w^{(2)}_{x_1,x_2}(k), \\ \dots \\ y^{(n_{gs})}_{x_1,x_2}(k) = \sum_{l=1}^{n_L} F_{x_1+\alpha_1^{(n_{gs})}h^{(l)},x_2+\alpha_2^{(n_{gs})}h^{(l)}}(\phi^{(l)}(k)) + w^{(n_{gs})}_{x_1,x_2}(k). \end{cases} \tag{29}$$

In this formulation we assume that the products $\alpha_1^{(s)}h^{(l)}$ yield an integer for each valid $s$, $l$, but this assumption can be lifted if interpolation techniques are introduced for computing the values of the phase at any spatial position in the grid (not only at the nodes). The procedure shown in the previous section goes through for this case as well, with the only difference that the Riccati equation to be solved for each frequency will not be scalar anymore:

$$\begin{aligned} P_{\nu_1,\nu_2} = {} & AP_{\nu_1,\nu_2}A^T + (I_{n_L} - A^2)\Sigma^2_\phi(\nu_1,\nu_2) \\ & - AP_{\nu_1,\nu_2}C^H_{\nu_1,\nu_2}(C_{\nu_1,\nu_2}P_{\nu_1,\nu_2}C^H_{\nu_1,\nu_2} + \sigma_w^2 I_2)^{-1}P_{\nu_1,\nu_2}AC_{\nu_1,\nu_2}, \end{aligned} \tag{30}$$

where $P_{\nu_1,\nu_2}$ is now an $n_L \times n_L$ symmetric matrix, $A$ is an $n_L \times n_L$ diagonal matrix containing $a^{(1)}_{\text{tur}}, a^{(2)}_{\text{tur}}, \dots a^{(n_L)}_{\text{tur}}$ as diagonal entries, $\Sigma^2_\phi(\nu_1,\nu_2)$ is an $n_L \times n_L$ matrix that appropriately generalizes $\sigma^2_\phi(\nu_1,\nu_2)$, and $I_{n_L}$ is the identity matrix of order $n_L$. In this case the closed form solution in Eq. (25) cannot be used, and the solution must be found with a solver. The frequency depending variable $K_{\nu_1,\nu_2}$ is going to be a $2n_{\text{gs}} \times n_L$ matrix, obtainable through:

$$K_{\nu_1,\nu_2} = AP_{\nu_1,\nu_2}C^H_{\nu_1,\nu_2}(C_{\nu_1,\nu_2}P_{\nu_1,\nu_2}C^H_{\nu_1,\nu_2} + \sigma_w^2 I_2)^{-1}, \tag{31}$$

while the gain $k_{n_1,n_2}$, which will be a $2n_{\text{gs}} \times n_L$ matrix as well, can be obtained as before with Eq. (24).

Once the different turbulent layers have been estimated, in the simulations appearing later on in this paper, we consider the case of a single DM that corrects the image in a single direction of observation. The phase distortion in such direction is computed as the sum of the contributions of each layer and then compensated for in the same way as in the classical AO case. In the multiconjugate adaptive optics (MCAO) case, there would be multiple DMs, each one conjugated to a given altitude.

### B. Other Extensions: Models of Higher Order and Wind Velocity

In this paper we will only show the results of the application of the method to AR1 models, but as said before, the method can handle any linear state-space model with spatial invariance. The first possible extension that we can foresee is the use of turbulence models of higher order, for example autoregressive models of order $n_o$. This would imply, in the general wide-field AO case, a block-diagonal $A$ matrix of size $n_o n_L \times n_o n_L$, with blocks of size $n_o \times n_o$. The second possibility is also the introduction of the wind velocity in the model equations. For example, along the same line as in [15], we can assume a state equation of the form

$$\phi_{x_1,x_2}(k+1) = a_{\text{tur}}\phi_{x_1+v^{(1)}_{\text{wind}},x_2+v^{(2)}_{\text{wind}}}(k) + v_{x_1,x_2}(k), \qquad (32)$$

where $v^{(1)}_{\text{wind}}$ and $v^{(2)}_{\text{wind}}$ depend on the wind velocity, i.e., they are the amount of steps the screen slides at each time step. These values in general are not integer, so finding $\phi_{x_1+v^{(1)}_{\text{wind}},x_2+v^{(2)}_{\text{wind}}}(k)$ might require the use of interpolation techniques. Equation (33) is spatially invariant for any $v^{(1)}_{\text{wind}}$ and $v^{(2)}_{\text{wind}}$, so this makes it possible to diagonalize the system by means of a DTFT and use the proposed technique.

### C. Influence Matrix Inversion

As seen before, the LQG regulator contains a linear state-feedback optimal controller in addition to the Kalman filter. In the case of a mirror with no dynamics, the optimal controller that minimizes the variance of $\phi^{(\text{res})}(k)$ is linear, and its gain is given by the pseudoinverse of $N$ as seen in Eq. (11). In the geometry that we consider, where for each phase sampling point an actuator is associated, $N$ is square and invertible as well. So the controller can be simply expressed as

$$u(k) = N^{-1}\Pi\hat{\phi}(k+1|k-1) = N^{-1}\Pi\hat{\phi}(k+1|\mathcal{S}_k). \qquad (33)$$

The control gain $N^{-1}$ can be approximated with the same technique if the mirror influence functions are spatially invariant (this holds for Gaussian, double Gaussian). We do not go into the details as the general procedure has been already well explained in Section 3 and the steps are basically the same: writing the frequency-decoupled Fourier transform of the function, computing the inverse for each frequency, inverse Fourier transform, and then crop the result. It is also to be noted that the matrix inversion is less computationally demanding than the solution of a Riccati equation, so this approach will be advantageous only at huge sizes, as will be shown later. For the MCAO case, it will be necessary to approximate the inverse of a more complex operator accounting for the DMs' geometry.

It is worth mentioning that other approaches are possible that do not require the explicit computation of the inverse of $N$, such as the conjugate gradient method [26,27], which would find $u(k)$ as the solution of the linear system $Nu(k) = \Pi\hat{\phi}(k+1|k-1)$.

## 5. CONSIDERATIONS ON THE ON-LINE COMPUTATIONAL TIMES

Although the main objective of the method that we presented is to reduce the bulk of the off-line computations, there are advantages at run time as well. Let us focus on the execution of the Kalman filter recursions. For a general Kalman filter, the on-line complexity is determined by the matrix-vector multiplication between the gain and the innovation (the difference between the measurement vector and its estimate). The gain matrix has a number of entries that are proportional to the number of measurements times the number of phase points; the number of measurements is proportional to the square of the diameter $D_t$ times the number of guide stars, while the number of phase points is proportional to the square of the diameter times the number of layers, so the gain matrix has a number of entries that is $\mathcal{O}(n_L n_{gs} D_t^4)$. The computational complexity of each Kalman filter iteration is then $\mathcal{O}(n_L n_{gs} D_t^4)$ as well.

Using sparse approximations of the covariance and replacing the matrix-vector multiplication in the update equation by an iterative resolution scheme as proposed in [8] yields a complexity of $\mathcal{O}(D_t^2 \log D_t)$. Likewise, performing the filtering in the Fourier domain also results in a complexity of $\mathcal{O}(D_t^2 \log D_t)$, in essence, the fast Fourier transform complexity. The proposed method requires a number of operations significantly smaller than a standard Kalman filter, thanks to the limited number of nonzero entries in the spatially distributed gain $k_{n_1,n_2}$. In fact $k_{n_1,n_2}$ contains only $\mathcal{O}(n_L n_{gs} z^2)$ entries. Moreover, because the number of operations to be executed at run time for reconstructing the phase in a single point is thresholded by a number proportional to $z^2$, the on-line computational complexity of each Kalman filter iteration is only $\mathcal{O}(n_L n_{gs} D_t^2 z^2)$. A key issue is whether $z$ needs to be increased with $D_t$ and how, which is investigated in the next section.

## 6. SIMULATION RESULTS: OFF-LINE COMPUTATIONAL TIMES AND SYSTEM PERFORMANCE

The proposed fast method has been implemented in MATLAB, and its speed and performance have been compared with the results that can be obtained with the preimplemented functions, namely `dare` for solving the Riccati equation and `inv` for the matrix inversion. The tests have been run on a personal computer with a 2.67 GHz processor and 8 GB of random access memory.

### A. Classical Adaptive Optics

We have simulated a classical AO system with an increasing telescope diameter with a Shack–Hartmann sensor whose subapertures grow in number but are kept constant in size, each of them corresponding to 0.5 m at the telescope pupil. Together with the number of lenslets in the sensor array, the number of state variables to be reconstructed by the Kalman filter grows as well, as each state variable corresponds to the phase in one point of the grid. The growth of the size of the state with the diameter is shown in Fig. 3.

For the simulation, we consider a von Kármán turbulence with $r_0 = 0.53$ m, $L_0 = 25$ m. Using Taylor's hypothesis, we generate turbulence phase screens as the superimposition of three layers, moving at 7.5, 12, and 15 m/s, with relative strengths of 50%, 17%, and 33%, respectively. The AO system works in an open loop at 250 Hz, there is a two-step delay between measurement and correction, and the WFS has a measurement noise of 0.28 rad$^2$ at the measurement wavelength, in terms of difference of the phase at the edges of

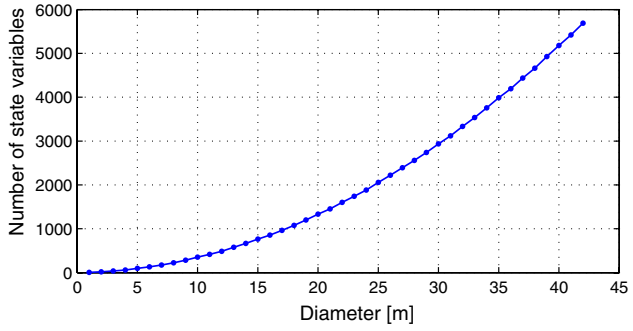


Fig. 3. (Color online) Number of state variables as a function of the telescope diameter.

the subapertures. We assume that the DM has an instantaneous response, and its actuators have a Gaussian influence function; the coupling factor of the actuators is 0.3 (a unitary displacement of one actuator induces a 0.3 displacement of the nearest ones). Whereas the phase is reconstructed only at the actuators' locations, the simulation is performed with phase screens generated on a finer sampling grid, i.e., with 5 times 5 points per each subaperture. For all the diameters considered, the computations of the fast Kalman filter is performed with a value of $z = 20$ and a $100 \times 100$ regular grid for the evaluation of the integral in Eq. (24). This choice of $z$ will be validated later on.

Figure 4 compares the times for computing the Kalman filter gain with the standard MATLAB functions and with the method shown in this article (the "fast" method). Similarly, Fig. 5 shows the same data for the inversion of the influence matrix.

It is clear from Fig. 4 that the new method offers a significant computational advantage even for medium sizes: from $D_t = 9\,\mathrm{m}$ on, the computation of the Kalman gain is quicker than for the standard procedure. If we extrapolate the computational time for $D_t = 42\,\mathrm{m}$, the European Extremely Large Telescope (E-ELT) size, we can roughly estimate 9 h of time for the standard off-line computations, while the new procedure gives an answer in less than 4 s. The computational time for the fast solver is independent of the diameter, as the spatially invariant gain $k_{n_1,n_2}$ is by construction the same for all diameters, as it only depends on the physical parameters of the turbulence, on $z$ and on the type of WFS. From Fig. 5 we can instead see that the proposed procedure is not as advantageous for the inversion of $N$. Even for $D_t = 42\,\mathrm{m}$, the standard MATLAB function takes only 44.7 s, making the "fast" approximation interesting only for sizes beyond the reach of the current technologies. This implies that a good, fast way of computing the LQG controller could be using the fast procedure for the Kalman filter gain and any appropriate inversion method for the influence matrix.

We have just shown that the novel method allows a great saving in terms of computational times; at this point, we have to verify that the savings in terms of times do not correspond to excessive losses of performance. For this reason, we have

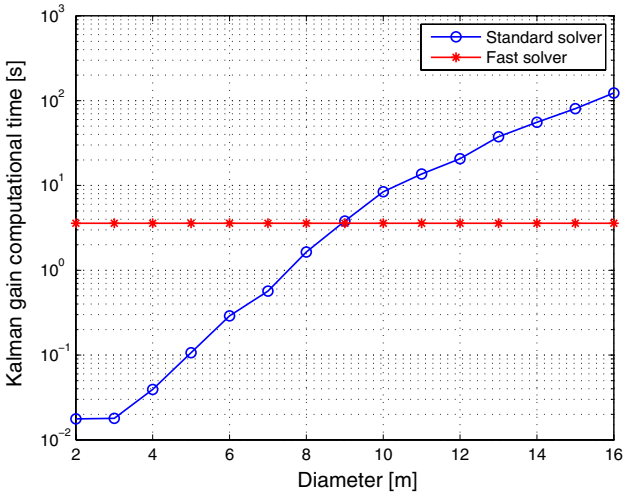


Fig. 4. (Color online) Off-line computational time for the Kalman gain for classical AO system, for different telescope diameters (2.67 GHz processor, 8 GB of RAM).

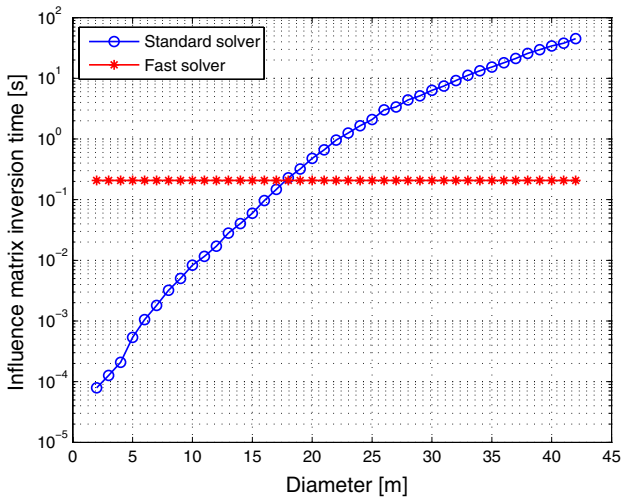


Fig. 5. (Color online) Off-line computational time for inverting the influence matrix, for different telescope diameters (2.67 GHz processor, 8 GB of RAM).

evaluated the performance of the new method in terms of Strehl ratios obtained in Monte Carlo simulations over 5000 time steps. The Strehl ratio $S$ is computed from the empirical covariance $\hat{\sigma}^2$ of the residual phase as

$$S = \exp(-\hat{\sigma}^2), \tag{34}$$

where $\hat{\sigma}^2$ is the average over time of the residual phase empirical variance on the pupil without the piston contribution. The LQG performance obtained with the standard MATLAB solvers is compared to the one with the Kalman filter gain computed in the fast way (we did not use the new procedure for the inversion of $N$). Figure 6 shows the Strehl ratios for the different methods. The figure, besides the variations due to the random nature of the tests, shows that the approximation of the Kalman filter gain does not imply a significant loss of Strehl ratio, for example roughly 2% at $D_t = 16\,\mathrm{m}$. We have executed the same simulation for the 8 m case, for an outer scale $L_0 = 100\,\mathrm{m}$ and for double wind speed, and we have observed the same trends, so we can assume that such a trend is general and not restricted to a limited range of turbulence

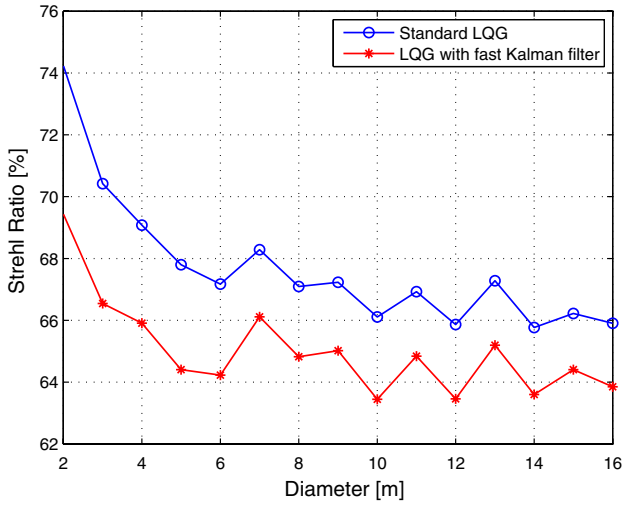


Fig. 6. (Color online) Strehl ratios obtained in simulation with Kalman gains computed in different ways.

parameters. An analysis of the Zernike modes has shown that the losses are more relevant for the modes that have bigger values at the borders of the pupil and relatively small values in the center.

It is also interesting to evaluate the effect of the size $z$ of the patch of nonzero spatially invariant gain matrix $k_{n_1,n_2}$. The evolution of the Strehl ratio as a function of $z$ is shown in Fig. 7 for three different telescope diameters. It can be seen in the figure that the Strehl ratio increases with $z$ up to an asymptotic value, which is lower than the value obtained by the standard LQG (as shown in Fig. 6). The asymptotic value is practically reached for a value of $z > 14$ (this justifies the assumption $z = 20$). Reducing $z$ can help speeding up the computations (as their complexity is proportional to $z^2$). These curves clearly show the two approximations involved in the proposed method; one approximation is due to assuming a finite size $z$ of the patch, and the effects of this approximation can be reduced by increasing $z$. The other approximation is basically due to the fact that we deal with a *finite* surface as if it were an *infinite* one, and this approximation is always present even for very big surface sizes (simply told: a "very big" thing is always different from an "infinite" thing). This is the reason why even for when $z \to \infty$ we would not converge to the exact solution.

It is also possible to estimate theoretically the loss of performance due to the use of the proposed fast Kalman filter. Having in mind the fact that the two-step estimation covariance corresponds to the residual phase covariance, as the matrix $N$ is invertible, the performance loss can be evaluated with the following reasoning. It is known that the solution $P_{\text{st}}$ of the Riccati equation for the standard Kalman filter is the covariance matrix of the estimation error $\hat{\phi}(k+1|k) - \phi(k+1)$. The covariance $P'_{\text{st}}$ of the two-step estimation error $\hat{\phi}(k+1|k-1) - \phi(k+1)$ can be simply obtained by propagating $P_{\text{st}}$ one step ahead, which gives $P'_{\text{st}} = a_{\text{tur}}^2 P_{\text{st}} + (1 - a_{\text{tur}}^2)\Sigma_\phi$, where $\Sigma_\phi$ is the covariance matrix of the turbulence phase. Remembering that $\Pi$ is the projection matrix that removes the piston contribution, then the mean variance of the estimation error is simply the average of the entries in the diagonal of $\Pi P'_{\text{st}} \Pi^T$:

$$\sigma^2_{\text{err,st}} = \text{mean}(\text{diag}(\Pi P'_{\text{st}} \Pi^T)). \tag{35}$$

If a different gain is used, such as the gain that can be obtained with the proposed method, denoted by $K_{\text{new}}$, then the covariance $P_{\text{new}}$ of the estimation error $\hat{\phi}(k+1|k) - \phi(k+1)$ can be obtained by solving the Lyapunov equation

$$(a_{\text{tur}} I - K_{\text{new}} D) P_{\text{new}} (a_{\text{tur}} I - K_{\text{new}} D)^T - P_{\text{new}} + \sigma_w^2 K_{\text{new}} K_{\text{new}}^T + (1 - a_{\text{tur}}^2)\Sigma_\phi = 0, \tag{36}$$

and the two-step estimation error covariance is therefore $P'_{\text{new}} = a_{\text{tur}}^2 P_{\text{new}} + (1 - a_{\text{tur}}^2)\Sigma_\phi$. The mean variance is then

$$\sigma^2_{\text{err,new}} = \text{mean}(\text{diag}(\Pi P'_{\text{new}} \Pi^T)). \tag{37}$$

We can consider the relative difference in the two covariance values

$$f = \frac{\sigma^2_{\text{err,new}} - \sigma^2_{\text{err,st}}}{\sigma^2_{\text{err,st}}} \tag{38}$$

as a number that indicates the loss of performance due to the infinite approximation of the finite pupil. The values of $f$ for the classical AO simulations are shown in Fig. 8, for different values of patch size $z$. Notice that these values exhibit the same trends as the Strehl ratio in Fig. 7, but they have the advantage of being obtained without time-domain Monte Carlo simulations. These curves can be used as a tool for evaluating *a priori* the performance and for deciding the value of $z$. For example, for $D_t = 8$ m, we can see that for an increasing value of $z$, $f$ decreases until it reaches an asymptotic value near to 14%. Losses that are bigger than this 14% can be attributed to the cropping of the patch to a limited size; the asymptotic value can be interpreted as the error caused by approximating the finite pupil with an infinite one.

The performance loss in terms of Strehl ratio depends not only on $f$ but also on the maximum theoretical performance that would be achievable by the Kalman filter (the standard one), which we call $S$. As a consequence of Eq. (34) and of the definition of $f$ in Eq. (38), the loss $S_{\text{lost}}$ of the Strehl ratio can then be obtained as

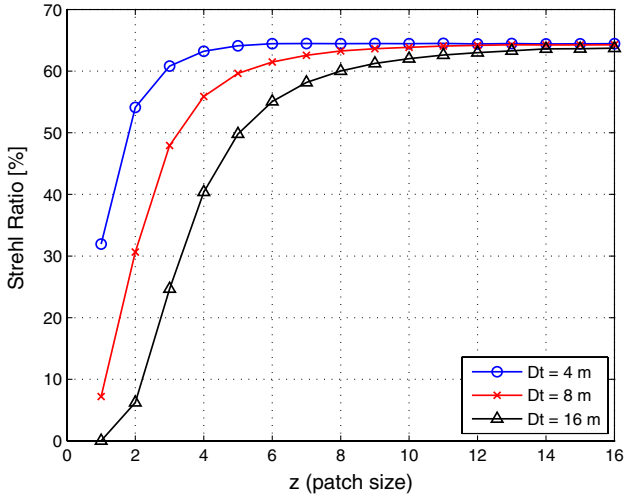


Fig. 7. (Color online) Strehl ratios obtained in simulation assuming different sizes $z$ for the patch. The simulations have been run for the case of $D_t = 4$ m, $D_t = 8$ m, and $D_t = 16$ m.

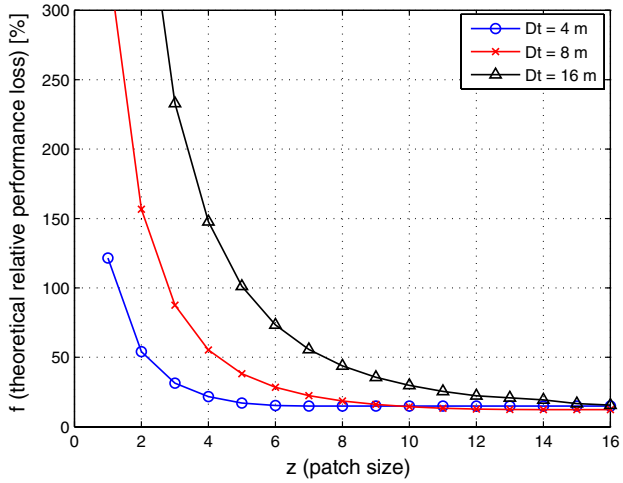


Fig. 8. (Color online) Relative theoretical performance loss as a function of the patch size $z$, for the case of $D_t = 4$ m, $D_t = 8$ m, and $D_t = 16$ m.

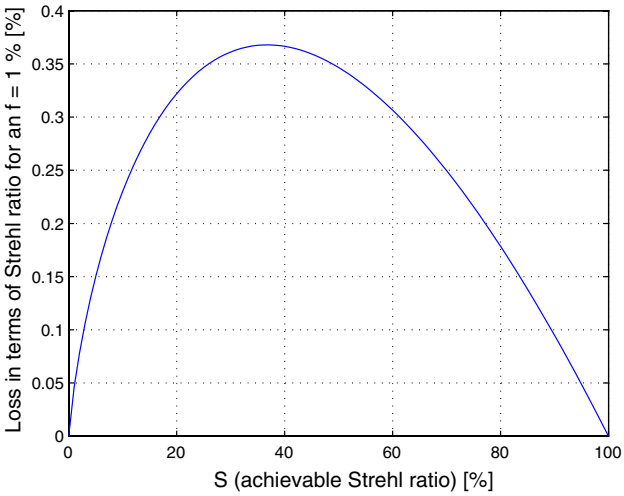


Fig. 9. (Color online) Plot of $-S\log(S)$, the equivalent value of 1% of $f$ in terms of loss of Strehl ratio, as a function of the maximum reachable Strehl ratio $S$.

$$S_{\text{lost}} = S - \exp((f+1)\log(S)) = S - S^{f+1} \approx -fS\log(S), \quad (39)$$

where $S$ has to be expressed as pure relative number (not percentage) when executing the calculation. Figure 9 shows the plot of $-S\log(S)$, which gives (in percent) the equivalence of a 1% of $f$ in terms of loss of Strehl ratio, as a function of the maximal theoretically reachable Strehl ratio $S$. We should remark that these curves are based on an AR1 model, so they provide theoretical results based on the assumption of the model. In practice, as we simulate the turbulence with Taylor's hypothesis, the losses of the Strehl ratio estimated by means of Figs. 8 and 9 tend to be overestimated; for example, for the simulation case with $D_t = 16$ m and $z = 20$, from Fig. 6 we see that $S \approx 0.66$, which in turn gives $-S\log(S) \approx 0.275$ (see Fig. 9). Hence we would estimate a loss of Strehl ratio of 0.15 (the asymptotic value of Fig. 8) times 0.275, which means slightly bigger than 4%; the simulation instead has shown roughly 2% losses (see Fig. 6). Remember that Fig. 8 has been plotted for a specific atmospheric condition, so a change in the atmospheric parameters (e.g., wind speed, $r_0$) would require retracing the curve.

With this technique we can also check the behavior of the method for extremely large telescope sizes: Fig. 10 shows the performance loss for the E-ELT case. From the figure it can be seen that a value of $z = 20$ gives an $f \approx 19\%$ while increasing to $z = 30$ would reduce $f$ to approximately 9%. As a

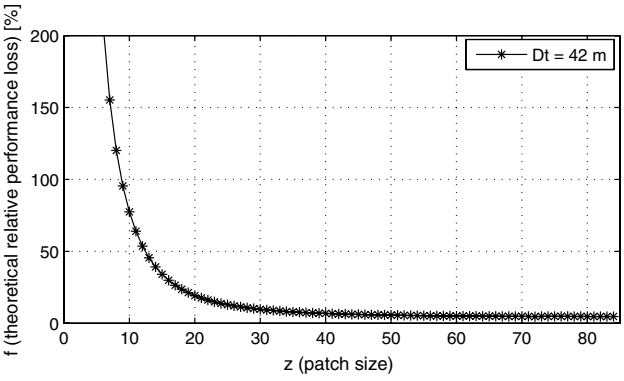


Fig. 10. Relative theoretical performance loss as a function of the patch size $z$, for the case of $D_t = 42$ m.

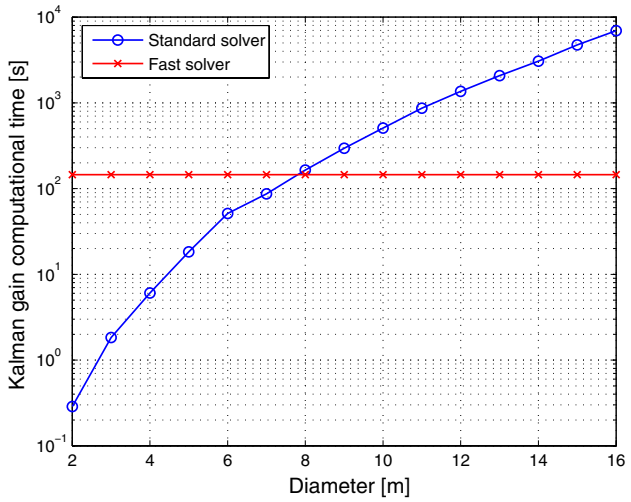


Fig. 11. (Color online) Off-line computational time for the Kalman gain for a tomographic wide-field AO system, for different telescope diameters (2.67 GHz processor, 8 GB of RAM).

comparison, notice that the asymptotic value of $f$ is about 14% for the case of $D_t = 8$ m. This trend reassures us that, even for extremely large diameters, the value of the patch size $z$ can be kept small with a limited performance loss.

### B. Wide-Field Adaptive Optics

We have run tests for the wide-field tomographic case as well, with the same physical parameters as in the previous simulations. We consider three layers at 0, 2000, and 4000 m of altitude, with relative energy levels of, respectively, 50%, 17%, and 33% and four Shack–Hartmann sensors associated to four natural guide stars disposed regularly around the center of the field at a distance of 30″. Figure 11 shows the computational times for the Kalman gain, and Fig. 12 shows the Strehl ratios obtained in the simulation with the standard LQG and the one using the fast-computation Kalman filter gain. Figure 13 shows the theoretical relative loss of performance.

The same kind of observations as for the classical AO case can be done. This time, according to the gross extrapolation of the curve in Fig. 11, the E-ELT case would take more than 230 h to compute, roughly ten days, while the fast procedure only takes a few seconds to output the spatially invariant gain

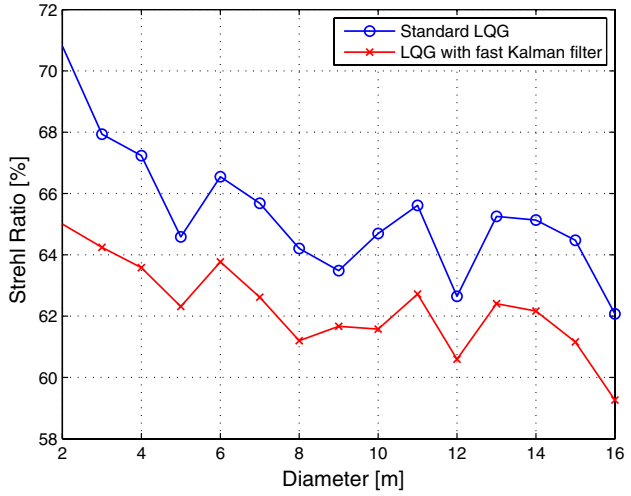


Fig. 12. (Color online) Strehl ratios obtained in simulation with the Kalman gains computed in different ways, for the wide-field case.

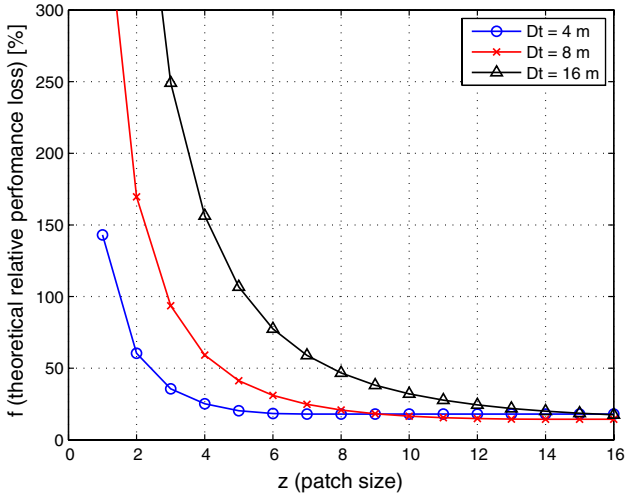


Fig. 13. (Color online) Relative theoretical performance loss as a function of the patch size $z$, for the wide-field case.

(which, again, is the same for all sizes). While 10 days of computation are still affordable (and can be significantly reduced with the use of supercomputers), it is clear that the use of models more complex than the AR1 might make imperative the use of fast methods.

## 7. CONCLUSIONS

We have shown a fast way of approximating the Kalman gain matrices for large-scale AO systems. The advantages of the proposed method in terms of off-line computational time are quite significant, especially for tomographic reconstruction techniques on large telescopes, e.g., of the ELT range. Because of the distributed approach, the complexity of the spatial invariant Kalman gain computation does not increase with the telescope size, but only with the number of turbulent layers and guide stars. As a consequence, the bottlenecks due to the growing off-line computational times necessary for solving the Riccati equation is overcome, making it possible to use the LQG control in AO even at large scales. Moreover, the sparse structure of the obtained Kalman gain proves to be convenient also for what concerns the on-line computations, and computing the overall complexity is straightforward. The cost of the approximation in terms of losses of the Strehl ratio is small, around 2% of the Strehl ratio for the cases we have considered here in simulation (with diameters up to 16 m). For the case of a 42 m diameter telescope, the theoretical tool we have proposed for performance assessment predicts that the losses due to a truncation of the patch can be kept small for limited patch sizes. This allows considering this method a good candidate for controlling AO systems of extremely large telescopes, as the computational complexity of the on-line computations can be limited at little cost.

The method has been applied to an open-loop system, but it is readily applicable to closed-loop systems as well. Moreover, it is possible to extend the method to turbulence models other than the AR1 that has been considered here; in fact, we could consider any other spatially invariant linear model, which might take into account higher model orders or the direction and intensity of the wind. This article has taken into account both classical and wide-field AO. The tomographic wide-field AO case has been demonstrated for natural guide stars; the applicability of the procedure for laser guide stars is the object of future research, as the setting is no more spatially invariant due to the "cone effect" of the laser star light, and due to the presence of tip/tilt uncertainty that has to be confronted. Future research could also include the integration of a filter for compensating the effect due to the vibration of the telescope.

## ACKNOWLEDGMENTS

The authors thank Gaetano Sivo for his help on the simulations, and Carlos Correia for his useful comments. The work of the first author is sponsored by the Netherlands Organisation for Scientific Research (NWO) and the Marie Curie Cofund Action by means of a Rubicon grant. This work has also been supported by the French National Research Agency (ANR) through project CHAPERSOA ANR-09-BLAN-0162-01.

## REFERENCES AND NOTES

1. F. Roddier, ed., *Adaptive Optics in Astronomy* (Cambridge University Press, 1999).
2. C. Béchet, M. Tallon, and E. Thiébaut, "Closed-loop AO performance with FrIM," in *Adaptive Optics: Analysis and Methods/Computational Optical Sensing and Imaging/Information Photonics/Signal Recovery and Synthesis Topical Meetings*, OSA Technical Digest (CD) (Optical Society of America, 2007), paper JTA4.
3. C. Vogel and Q. Yang, "Fast optimal wavefront reconstruction for multi-conjugate adaptive optics using the Fourier domain preconditioned conjugate gradient algorithm," Opt. Express **14**, 7487–7498 (2006).
4. L. Gilles, "Closed-loop stability and performance analysis of least-squares and minimum-variance control algorithms for multiconjugate adaptive optics," Appl. Opt. **44**, 993–1002 (2005).
5. C. Kulcsár, H.-F. Raynaud, C. Petit, J.-M. Conan, and P. Viaris de Lesegno, "Optimal control, observers and integrators in adaptive optics," Opt. Express **14**, 7463–8012 (2006).
6. T. Kailath, A. Sayed, and B. Hassibi, *Linear Estimation* (Prentice-Hall, 2000).
7. B. Le Roux, J.-M. Conan, C. Kulcsár, H.-F. Raynaud, L. Mugnier, and T. Fusco, "Optimal control law for classical and multiconjugate adaptive optics," J. Opt. Soc. Am. A **21**, 1261–1276 (2004).
8. C. Correia, J.-M. Conan, C. Kulcsár, H.-F. Raynaud, and C. Petit, "Adapting optimal LQG methods to ELT-sized AO systems," in *Proceedings of the 1st Conference on Adaptive Optics for Extremely Large Telescopes* (EDP Sciences, 2010).
9. R. Fraanje, P. Massioni, and M. Verhaegen, "A decomposition approach to distributed control of dynamic deformable mirrors," Int. J. Optomechatron. **4**, 269–284 (2010).
10. R. Fraanje, J. Rice, M. Verhaegen, and N. Doelman, "Fast reconstruction and prediction of frozen flow turbulence based on structured Kalman filtering," J. Opt. Soc. Am. A **27**, A235–A245 (2010).
11. B. Bamieh, F. Paganini, and M. Dahleh, "Distributed control of spatially invariant systems," IEEE Trans. Autom. Control **47**, 1091–1107 (2002).
12. J. Kim and T. Bewley, "A linear systems approach to flow control," Annu. Rev. Fluid Mech. **39**, 383–417 (2007).
13. G. Stewart, P. Baker, D. Gorinevsky, and G. Dumont, "An experimental demonstration of recent results for spatially distributed control systems," in *Proceedings of the 2001 American Control Conference* (IEEE, 2001), Vol. 3, pp. 2216–2221.
14. S. Jiang, P. Voulgaris, L. Holloway, and L. Thompson, "Distributed control of large segmented telescopes," in *Proceedings of the 2006 American Control Conference* (IEEE, 2006), pp. 1942–1947.
15. L. Poyneer and J.-P. Véran, "Optimal modal Fourier-transform wavefront control," J. Opt. Soc. Am. A **22**, 1515–1526 (2005).
16. A. Costille, C. Petit, J.-M. Conan, C. Kulcsár, H.-F. Raynaud, and T. Fusco, "Wide field adaptive optics laboratory demonstration with closed-loop tomographic control," J. Opt. Soc. Am. A **27**, 469–483 (2010).

17. D. Looze, "Minimum variance control structure for adaptive optics systems," J. Opt. Soc. Am. A **23**, 603–612 (2006).
18. P. Joseph and J. Tou, "On linear control theory," AIEE Trans. Appl. Indus., **80**, 193–196 (1961).
19. Y. Bar-Shalom and E. Tse, "Dual effect, certainty equivalence, and separation in stochastic control," IEEE Trans. Autom. Control **19**, 494–500 (1974).
20. Notice that even if the usual name of the transform contains the words "discrete time," we actually use it for transformations on spatial variables, so in this case it might be more appropriate to call it the discrete-space Fourier transform.
21. A. Oppenheim and A. Willsky, *Signals & Systems*, 2nd ed. (Prentice-Hall, 1997).
22. S. L. Marple, *Digital Spectral Analysis*, Signal Processing Series (Prentice-Hall, 1987).
23. M. Jovanović and B. Bamieh, "On the ill-posedness of certain vehicular platoon control problems," IEEE Trans. Autom. Control **50**, 1307–1321 (2005).
24. R. Curtain, O. Iftime, and H. Zwart, "A comparison between LQR control for a long string of SISO systems and LQR control of the infinite spatially invariant version," Automatica **46**, 1604–1615 (2010).
25. F. Rigaut, B. Ellerbroek, and R. Flicker, "Principles, limitations, and performance of multiconjugate adaptive optics," Proc. SPIE **4007**, 1022–1031 (2000).
26. A. Quarteroni, R. Sacco, and F. Saleri, *Numerical Mathematics* (Springer, 2000).
27. L. Gilles, B. Ellerbroek, and C. Vogel, "Preconditioned conjugate gradient wave-front reconstructors for multiconjugate adaptive optics," Appl. Opt. **42**, 5233–5250 (2003).