

# Optimal control law for classical and multiconjugate adaptive optics

Brice Le Roux and Jean-Marc Conan

*Office National d'Etudes et de Recherches Aérospatiales, B.P. 72, 92322 Châtillon, France*

Caroline Kulcsár and Henri-François Raynaud

*Laboratoire de Traitement et de Transport de l'Information, Institut Galilée, Université Paris 13, Villetaneuse, France*

Laurent M. Mugnier and Thierry Fusco

*Office National d'Etudes et de Recherches Aérospatiales, B.P. 72, 92322 Châtillon, France*

Received December 16, 2003; accepted February 25, 2004

Classical adaptive optics (AO) is now a widespread technique for high-resolution imaging with astronomical ground-based telescopes. It generally uses simple and efficient control algorithms. Multiconjugate adaptive optics (MCAO) is a more recent and very promising technique that should extend the corrected field of view. This technique has not yet been experimentally validated, but simulations already show its high potential. The importance of MCAO of an optimal reconstruction using turbulence spatial statistics has already been demonstrated through open-loop simulations. We propose an optimal closed-loop control law that accounts for both spatial and temporal statistics. The prior information on the turbulence, as well as on the wave-front sensing noise, is expressed in a state-space model. The optimal phase estimation is then given by a Kalman filter. The equations describing the system are given and the underlying assumptions explained. The control law is then derived. The gain brought by this approach is demonstrated through MCAO numerical simulations representative of astronomical observation on a 8-m-class telescope in the near infrared. We also discuss the application of this control approach to classical AO. Even in classical AO, the technique could be relevant especially for future extreme AO systems. © 2004 Optical Society of America

OCIS codes: 010.1080, 010.1330, 070.6020, 100.3190.

## 1. INTRODUCTION

High-resolution imaging with ground-based telescopes is now possible with adaptive optics (AO). However, classical AO, which uses a single deformable mirror (DM) in the pupil, provides a limited corrected field of view (FOV). Large FOV corrections can be obtained by correcting the turbulence volume above the telescope, with several DMs optically conjugated at various altitudes and with several wave-front sensors (WFSs) looking at guide stars (GSs), a process that corresponds to the concept of multiconjugate adaptive optics (MCAO). This concept was first proposed by Dicke<sup>1</sup> and reintroduced in the early 1990s in the papers of Beckers,<sup>2</sup> Tallon *et al.*,<sup>3</sup> and Ellerbroek.<sup>4</sup> More recently, an impressive number of MCAO papers have been published, all showing the high potential of this technique.<sup>5–11</sup>

MCAO brings new problems in reconstruction and control: It leads to a larger number of degrees of freedom, and it relies on a complex reconstruction process for the phase estimation in the volume. The presence of unseen and so-called badly seen modes (see Section 3) requires a careful treatment of this inverse problem. These modes have to be estimated to ensure a good interpolation of the perturbation between GSs and, in turn, a good performance in the global FOV. In MCAO, such unseen modes can be numerous, they evolve with the GS geometry, and

they are difficult to visualize since they correspond to a phase in the turbulence volume. For all these reasons, MCAO requires us to find systematic ways of deriving an optimal control law that accounts for the system characteristics, including the GS geometry, and for the spatial and temporal priors on turbulence and noise. Besides, the optimization needs to be global, that is, applied to the multivariable servoloop.

Control laws based on a mode-per-mode optimization, and generally used in classical AO,<sup>12–16</sup> could be generalized to MCAO, but their performance is not optimal since they cannot take advantage of the global spatial priors on the turbulence related to Kolmogorov statistics and to the distribution of the turbulence strength in altitude. The need for a global optimization for an efficient estimation of unseen modes has already been demonstrated in ideal open-loop simulations.<sup>6</sup> Future real MCAO systems will, however, operate in closed loop. The objective of this paper, and its most innovative aspect, is to propose a closed-loop control law resulting from a global optimization and to demonstrate its efficiency in the context of MCAO. We propose an approach based on a state-space-model formalism, a Kalman filter, and a feedback control derived from the classical linear estimation theory.<sup>17–19</sup> Global approaches based on Kalman filtering have already been proposed in the literature.<sup>20,21</sup> In our paper we make

some assumptions that simplify the theoretical developments. We assume that the DMs are fast compared with the sampling period. This assumption is often valid for astronomical applications. Paschall and Anderson<sup>20</sup> have shown that the mirror dynamics can be taken into account in the same framework when necessary. We also derive our control law from a simple turbulence-temporal-prior-model, but we show in Subsection 5.C.2 that good performance can still be obtained when one is correcting realistic turbulent screens (Taylor model). Note that Gavel and Wiberg<sup>21</sup> recently proposed a model that could truly account for the Taylor hypothesis. Finally, our paper is, to our knowledge, the first one that applies such an approach to MCAO. Throughout the paper we explain why this approach is particularly relevant in the context of MCAO, often referring to the unseen mode issue.

The performance of the proposed approach is quantified by numerical simulations. The system that we simulated is representative of astronomical observations on a 8-m-class telescope in the near infrared. We begin by simulating a classical AO system. We compare the results with those obtained by the most common control law, the optimized modal gain integrator<sup>12,22</sup> (OMGI). We start with a simplified temporal model and then show that a Kalman approach based on this approximative prior is still able to efficiently correct a realistic Taylor turbulence. In classical AO, the major gain is due to the capability of our approach to make a good prediction. But we verify also that it is able to deal efficiently with the unseen modes. This last point becomes a very important aspect in the MCAO case. Finally, we simulate this more complex case and compare the performance of our approach with that of the Optimized Modal Gain Integrator generalized to MCAO. This results demonstrate the gain brought by the Kalman approach in the MCAO context.

The MCAO system is presented in Section 2. The open-loop optimal reconstruction is briefly recalled in Section 3. Our closed-loop optimal approach is described in Section 4, and its links with previous works are discussed. In Section 5, we present numerical simulations in classical AO and in MCAO.

## 2. MULTICONJUGATE ADAPTIVE OPTICS AND OPTIMAL CONTROL

The anisoplanatism phenomenon comes from the fact that the turbulence is not located in one layer on the ground but in a volume above it. In classical AO, one single DM located on the telescope pupil cannot correct the turbulence in all directions. The concept of MCAO is based on correcting anisoplanatism through the use of DMs optically conjugated at various altitudes in the atmosphere. We present the system and notation in Subsection 2.A and, we explain in Subsection 2.B how the estimation and control problems can be separated.

### A. System Description and Notation

The turbulence volume is modeled by  $N_L$  discrete independent turbulent layers located at altitudes  $\{h_j\}$ . We associate with each layer a turbulence strength  $C_n^2(h_j) \delta h$ , where  $C_n^2(h_j)$  is the index structure constant in layer  $j$

and  $\delta h$  is the thickness of the layer. This turbulence volume is corrected by  $N_{DM}$  DMs optically conjugated at altitudes  $\{h'_j\}$ . The turbulence statistics are assumed to be Kolmogorov for each turbulent layer. The geometry of the system is summarized in Fig. 1.

The measurement is done with several WFSs looking at several stars, the so-called GSs. We consider  $N_{GS}$  GSs in the  $\beta = \{\beta_i\}$  directions. The FOV of interest, where the correction has to be optimized, is discretized into  $K$  angles and denoted  $\alpha = \{\alpha_i\}$ . We note  $\phi_{\alpha_i}^{\text{tur}}$  the turbulent phase propagated onto the pupil in the direction  $\alpha_i$  and  $\phi_{\alpha_i}^{\text{cor}}$  the correction phase on the pupil in the direction  $\alpha_i$ . We note  $\varphi^{\text{tur},j}$  the turbulent phase on layer  $j$  and  $\varphi^{\text{cor},j}$  the correction phase given by the DM number  $j$ .

The turbulent phase arriving on the telescope pupil in the direction  $\beta_i$  is given, in the near-field approximation, by the sum of all the turbulent layers' contributions,

$$\phi_{\beta_i}^{\text{tur}}(\mathbf{r}) = \sum_{j=1}^{N_L} \varphi^{\text{tur},j}(\mathbf{r} + h_j \beta_i), \quad (1)$$

where  $\mathbf{r}$  is the position inside the pupil. The correction phase  $\phi_{\beta_i}^{\text{cor}}$ , generated by the DMs in the direction  $\beta_i$ , is also defined as

$$\phi_{\beta_i}^{\text{cor}}(\mathbf{r}) = \sum_{j=1}^{N_{DM}} \varphi^{\text{cor},j}(\mathbf{r} + h'_j \beta_i). \quad (2)$$

In the rest of this paper, rather than talking about continuous functions of the coordinate  $\mathbf{r}$ , we will use for  $\phi$  and  $\varphi$  a discrete representation based on a modal expansion of the phase, for instance on the Zernike polynomials. The turbulent and correction phases  $\phi^{\text{cor}}$  and  $\phi^{\text{tur}}$  are then represented by vectors of  $N_{\text{modes}}$  coefficients:  $\{\phi_k^{\text{cor}}\}$ ,  $\{\phi_k^{\text{tur}}\}$ . In this representation, we note  $\varphi^{\text{tur}} = \{\varphi^{\text{tur},j}\}$  the volumic turbulent phase in all layers and  $\varphi^{\text{cor}} = \{\varphi^{\text{cor},j}\}$  the volumic correction phase generated by all the DMs.  $\varphi^{\text{tur}}$  is modeled as a stochastic centered variable of Gaussian statistics characterized by its covariance matrix  $\mathbf{C}_\varphi$ . As the turbulent layers are independent,  $\mathbf{C}_\varphi$  is a block matrix that contains all the covariance matrices for all

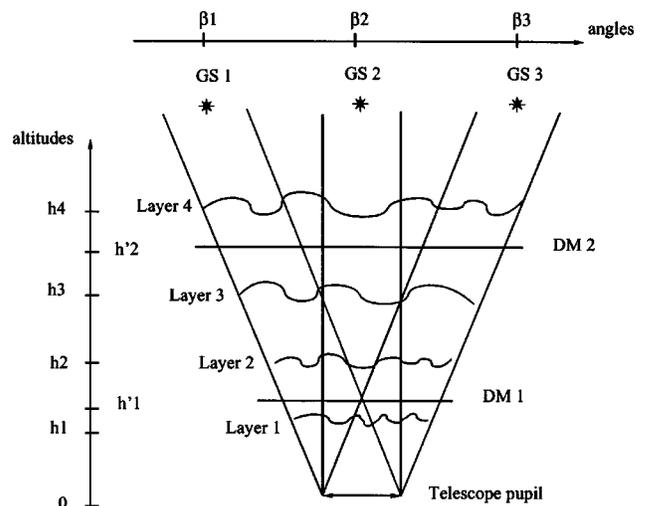


Fig. 1. Illustration of the geometry of an MCAO system. Example with  $N_L = 4$  layers,  $N_{DM} = 2$  DMs, and  $N_{GS} = 3$  guide stars.

layers (Kolmogorov weighted by the turbulence strength  $C_n^2 \delta h$ ) and zeros for the cross correlations between layers.

$\mathbf{M}_{\alpha_i}^L$  is defined as the matrix that performs the sum of the contributions of each turbulent wave front in the direction  $\alpha_i$ .  $\mathbf{M}_{\alpha_i}^{\text{DM}}$  performs the sum of the contributions of each DM in the direction  $\alpha_i$ . Equations (1) and (2) can therefore be written as

$$\phi_{\alpha_i}^{\text{tur}} = \mathbf{M}_{\alpha_i}^L \phi^{\text{tur}}, \quad (3)$$

$$\phi_{\alpha_i}^{\text{cor}} = \mathbf{M}_{\alpha_i}^{\text{DM}} \phi^{\text{cor}}. \quad (4)$$

We define  $\mathbf{M}_\alpha^L$  as  $\mathbf{M}_\alpha^L = [(\mathbf{M}_{\alpha_1}^L)^T, \dots, (\mathbf{M}_{\alpha_i}^L)^T, \dots, (\mathbf{M}_{\alpha_{N_{\text{GS}}}}^L)^T]^T$ ,  $\mathbf{M}_\alpha^L$  being the matrix that performs the sum of the contributions of each turbulent wave front in all the directions  $\alpha = \{\alpha_i\}$ :

$$\Phi_\alpha^{\text{tur}} = \{\phi_{\alpha_i}^{\text{tur}}\} = \mathbf{M}_\alpha^L \phi^{\text{tur}}. \quad (5)$$

In the same way,  $\mathbf{M}_\alpha^{\text{DM}}$  is the matrix that performs the sum of the contributions of each DM in all the directions  $\alpha = \{\alpha_i\}$ , which is written  $\mathbf{M}_\alpha^{\text{DM}} = [(\mathbf{M}_{\alpha_1}^{\text{DM}})^T, \dots, (\mathbf{M}_{\alpha_i}^{\text{DM}})^T, \dots, (\mathbf{M}_{\alpha_{N_{\text{GS}}}}^{\text{DM}})^T]^T$ .

This notation is also defined in the same way for  $\beta_i$  and  $\beta$  so that Eqs. (3)–(5) can be written by using  $\mathbf{M}_{\beta_i}^L$ ,  $\mathbf{M}_{\beta_i}^{\text{DM}}$ ,  $\mathbf{M}_\beta^L$ ,  $\mathbf{M}_\beta^{\text{DM}}$ ,  $\phi_{\beta_i}^{\text{tur}}$ ,  $\Phi_\beta^{\text{tur}}$ ,  $\phi_{\beta_i}^{\text{cor}}$  and  $\Phi_\beta^{\text{cor}}$ .

We assume that the response of a DM to voltages is linear, and we denote by  $\mathbf{N}$  the matrix defining the linear relationship between the voltages  $\mathbf{u}$  applied on the DMs and the generated correction phase  $\phi^{\text{cor}}$ :

$$\phi^{\text{cor}} = \mathbf{N}\mathbf{u}. \quad (6)$$

Each column of  $\mathbf{N}$  corresponds to the modal representation of the deformation of one DM actuator.

## B. Principle of Separation between Estimation and Control

In order to simplify the analytical expressions of the control law, one can use the fact that the mirror dynamics can be neglected. We consider that the DM reacts very fast compared with the sampling period,  $\tau_{\text{DM}} \ll T_{\text{samp}}$ . This assumption is generally valid for the current astronomical AO systems equipped with piezo-stacked DMs<sup>24</sup> and operating at a sampling frequency of a few hundred herz. At higher sampling rates, or for other types of DMs (adaptive secondaries), the dynamics have to be taken into account. Paschall and Anderson<sup>20</sup> showed that this could be done in the state-space framework. They use in this case a linear quadratic Gaussian control law.

With such an assumption, the optimal correction corresponds to the voltages  $\mathbf{u}$  that minimize  $\langle \epsilon'(\phi^{\text{tur}}, \mathbf{u}) \rangle_{\phi, \text{noise}}$ , where the notation  $\langle \cdot \rangle_{\phi, \text{noise}}$  stands for the average over turbulence and measurement noise outcomes and  $\epsilon'(\phi^{\text{tur}}, \mathbf{u})$  is the mean square error, defined as

$$\begin{aligned} \epsilon'(\phi^{\text{tur}}, \mathbf{u}) &= \sum_i \|\phi_{\alpha_i}^{\text{tur}} - \phi_{\alpha_i}^{\text{cor}}\|^2 \\ &= \sum_i \|\mathbf{M}_{\alpha_i}^L \phi^{\text{tur}} - \mathbf{M}_{\alpha_i}^{\text{DM}} \mathbf{N}\mathbf{u}\|^2. \end{aligned} \quad (7)$$

$\|\phi\|^2 = 1/S \int_S \phi(\mathbf{r})^2 d\mathbf{r}$  denotes the so-called spatial variance in the telescope pupil  $S$ .  $\|\phi\|^2 = \sum_i \phi_i^2$  if the base is orthonormalized.

It can be shown that searching  $\mathbf{u}$  that minimizes  $\langle \epsilon'(\phi^{\text{tur}}, \mathbf{u}) \rangle_{\phi, \text{noise}}$  is equivalent to minimizing consecutively the criterion

$$\epsilon'' = \langle \|\phi^{\text{tur}} - \hat{\phi}^{\text{tur}}\|^2 \rangle_{\phi, \text{noise}} \quad (8)$$

with respect to  $\hat{\phi}^{\text{tur}}$  and then to finding  $\mathbf{u}$  that minimizes  $\epsilon'(\hat{\phi}^{\text{tur}}, \mathbf{u})$ .

In Section 4 we will propose an approach that gives the estimate  $\hat{\phi}^{\text{tur}}$  that minimizes Eq. (8). Knowing  $\hat{\phi}^{\text{tur}}$ , we can write the solution of the minimization of the criterion  $\epsilon'(\hat{\phi}^{\text{tur}}, \mathbf{u})$  written as

$$\mathbf{u} = \mathbf{P}_{[\alpha; \text{DM}]} \hat{\phi}^{\text{tur}} \quad (9)$$

with the operator

$$\mathbf{P}_{[\alpha; \text{DM}]} = \left[ \sum_i [\mathbf{M}_{\alpha_i}^{\text{DM}} \mathbf{N}]^T \mathbf{M}_{\alpha_i}^{\text{DM}} \mathbf{N} \right]^+ \left[ \sum_i [\mathbf{M}_{\alpha_i}^{\text{DM}} \mathbf{N}]^T \mathbf{M}_{\alpha_i}^L \right], \quad (10)$$

where  $(\cdot)^+$  denotes the generalized inverse. One can note that this matrix depends only on the number and positions of the DMs with respect to the true layers and on the FOV of interest.

This separation property between estimation and projection is a classical result of linear control theory.<sup>19</sup> It has already been used in the context of MCAO.<sup>6,21</sup>

## 3. OPTIMAL RECONSTRUCTION IN OPEN LOOP

In this section we consider the phase estimation for an open-loop system, which means that the measurements are performed on the uncorrected turbulent phase. The measurement process is presented in Subsection 3.A. The estimator defined in Subsection 3.C gives the best estimate of the turbulent phase for a measurement at a given time. There are no time series here; we consider an instant-by-instant estimation. The correction is also assumed to be instantaneous without delay. The temporal issues therefore do not apply here; thus only the spatial priors are introduced in this section.

### A. Open-Loop Measurement Process

We note  $\phi_{\beta_i}^{\text{meas}}$  the phase analyzed by the  $i$ th WFS (in the direction  $\beta_i$ ). Because the measurements are done before the correction,  $\phi_{\beta_i}^{\text{meas}}$  is the turbulent phase arriving on the telescope pupil in the direction  $\beta_i$ :

$$\phi_{\beta_i}^{\text{meas}} = \phi_{\beta_i}^{\text{tur}}. \quad (11)$$

The WFS measurements are assumed to be linear, so the measurement given by the WFS in the direction  $\beta_i$  can then be expressed as

$$\mathbf{y}_{\beta_i} = \mathbf{D} \phi_{\beta_i}^{\text{meas}} + \mathbf{w}_i, \quad (12)$$

with  $\mathbf{D}$  the interaction matrix that defines the linear relation between the phase and the measurement (each column of  $\mathbf{D}$  corresponds to the measurements given for a single-mode phase). Here  $\mathbf{w}_i$  is the measurement noise for WFS  $i$  of covariance matrix  $\mathbf{C}_{w_i}$ .  $\mathbf{Y} = \{\mathbf{y}_{\beta_i}\}$  is defined

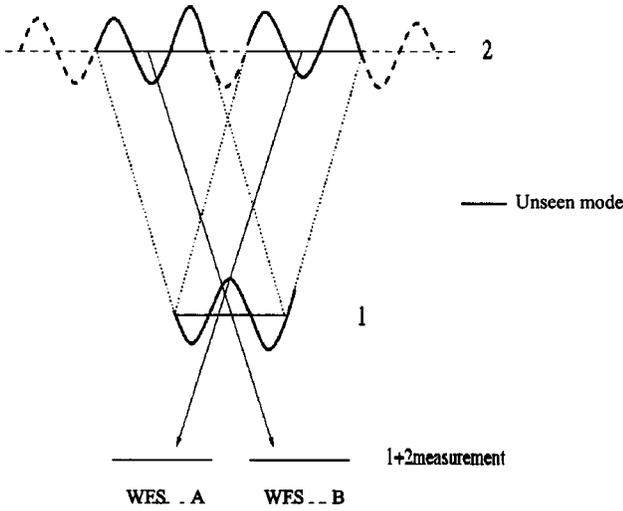


Fig. 2. Illustration of the concept of unseen modes.

as the vector containing all the measurements for all directions  $\beta$ ,  $\mathbf{w}$  is defined as the vector containing the measurement noise for all WFS, and  $\mathbf{C}_w$  is its covariance matrix.

### B. Notion of Unseen Mode

What we call “unseen mode” is a mode that is not measured by the WFSs.<sup>25–27</sup> An example of unseen modes in MCAO is shown in Fig. 2. This figure represents two wave fronts at different altitudes that exactly compensate each other in the GS directions. The WFSs are blind to such a wave-front distribution. This also means that such a mode has no effect on the image quality in the GS directions. However, it is still very important to estimate and then correct this mode since the resulting phase in other observation directions can be nonnegligible. The more distant the GSs, the smaller the spatial frequency of the unseen modes and the larger the turbulent energy contained in these modes. Simply filtering out such high-energy unseen modes would degrade the performance in the FOV, particularly when looking away from the GS positions. Estimating these modes can actually be seen as a way of performing an interpolation of the turbulent perturbation in between the GS directions. Of course since the WFS measurements are noisy, the same problem applies to modes associated to nonzero but small signal-to-noise ratio (SNR), so-called badly seen modes. In the rest of the paper, when we speak of unseen modes, it should actually be understood that the badly seen ones are also considered.

### C. Minimum-Mean-Square-Error Estimator

The presence of unseen modes and their influence on image quality motivates the use of turbulent-phase reconstruction that is able to estimate these modes. In the open-loop case, it has been shown<sup>6</sup> that the use of spatial priors through a minimum-mean-square-error (MMSE) estimator could satisfy this objective. We now recall this approach briefly.

As presented in Subsection 2.B, one should estimate  $\hat{\phi}^{\text{tur}}$  from all the measurements  $\mathbf{Y}$  so that the criterion of

Eq. (8) is minimized. This implies some statistical knowledge of noise and turbulence.

When Gaussian statistics for noise and turbulence are assumed, the minimization of this criterion leads to a solution in  $\hat{\phi}^{\text{tur}}$  that is linear with respect to the wave-front measurements. The solution can be written in this simple form,<sup>6</sup>

$$\hat{\phi}^{\text{tur}} = \mathbf{W}_{\text{tom}} \mathbf{Y}, \quad (13)$$

with

$$\mathbf{W}_{\text{tom}} = \mathbf{C}_\phi [\mathbf{DM}_\beta^L]^T [\mathbf{DM}_\beta^L \mathbf{C}_\phi [\mathbf{DM}_\beta^L]^T + \mathbf{C}_w]^{-1}, \quad (14)$$

with  $\mathbf{C}_\phi$  and  $\mathbf{C}_w$  the covariance matrices of the turbulent phase and the measurement noise.  $\mathbf{C}_\phi$  is a block matrix that contains all the covariance matrices for all layers (Kolmogorov weighted by the  $C_n^2$  profile) and 0 for cross correlations between layers.  $\mathbf{C}_w$  is a block matrix that contains all the noise covariance matrices corresponding to all WFSs’ noise, and zero for the cross correlations between WFSs.

The equation  $\hat{\phi}^{\text{tur}} = \mathbf{W}_{\text{tom}} \mathbf{Y}$  corresponds to an optimal stochastic estimation of the turbulent phase on each layer for single time-sample measurement. This reconstruction is often called tomographic since it gives a reconstruction of the turbulence volume from the projections measured by the WFS. It takes into account the GS geometry, the WFS measurement model, and the noise and turbulence statistics, including the  $C_n^2$  profile. Both noise and turbulence are characterized by their covariance matrices. Note that  $\mathbf{W}_{\text{tom}}$  is not related to the DMs and is independent of the FOV of interest. After this estimation of the turbulent phase, as stated in Subsection 2.B, the optimal control consists of a deterministic “projection” of the tomographic solution onto the DM to obtain the voltages that optimize the correction in the desired FOV. It corresponds to  $\mathbf{P}_{[\{a_i\}; \text{DM}]}$ , as already mentioned in Subsection 2.B.

One can note that if we assume  $\mathbf{N} = Id$  (which is equivalent to saying that each DM can produce any correction phase) and  $\mathbf{D} = Id$  (which means the WFS directly analyze the turbulent modes), Eqs. (10) and (13) exactly correspond to the results given in Ref. 6.

A cruder approach, which is often used to inverse this ill-posed problem of the phase-correction estimation in each DM, consists of using a least-square minimization on measurements  $\|\mathbf{Y} - \mathbf{DM}_\beta^{\text{DM}} \mathbf{N} \mathbf{u}\|^2$ ; see Refs. 28 and 29. With our notation, this wave-front estimator is given by

$$\mathbf{u} = [[\mathbf{DM}_\beta^{\text{DM}} \mathbf{N}]^T [\mathbf{DM}_\beta^{\text{DM}} \mathbf{N}]]^+ [\mathbf{DM}_\beta^{\text{DM}} \mathbf{N}]^T \mathbf{Y}. \quad (15)$$

Because  $[\mathbf{M}_\beta^{\text{DM}}]^T [\mathbf{M}_\beta^{\text{DM}}]$  is generally ill-conditioned, the inversion is made by using a singular-value decomposition in which the lower singular values are set to zero in order to avoid the noise amplification [truncated singular-value decomposition (TSVD)].

It has been shown<sup>29</sup> that the MMSE estimator gives quite a bit better results than the TSVD method, whatever the truncation threshold is.

There are essentially two reasons why the MMSE is better than the TSVD approach, especially in the presence of energetic unseen, or badly seen, modes. The first one is the optimal minimization of noise propagation (no

*ad hoc* threshold adjustment procedure). The second fundamental reason is that the MMSE not only controls the noise amplification on unseen modes but also estimates these modes by using their correlation with better-seen modes. The unseen modes appear in the basis that diagonalizes

$$[\mathbf{DM}_\beta^{\text{DM}}\mathbf{N}]^T[\mathbf{DM}_\beta^{\text{DM}}\mathbf{N}] \quad (16)$$

as the eigenvectors associated with zero, or nearly zero, eigenvalues.

In general,  $[\mathbf{DM}_\beta^{\text{DM}}\mathbf{N}]$  is not square but rectangular. The size of this matrix depends on the number of DMs, the number of actuators, the number of GSs, and the number of measurements by the WFS. For example, in the realistic case that we simulate in Subsection 5.D, there are a total of 482 corrected modes and a total of 312 measured modes (see Subsection 5.D for more details). In that case, there are many unseen modes. In the optimal approach,<sup>29,30</sup> the use of spatial priors allows us to partially recover these unseen modes, whereas a TSVD simply filters them out.

#### 4. OPTIMAL CONTROL LAW FOR CLOSED-LOOP OPERATION

We showed in Section 3 that to estimate the unseen modes with the open-loop MCAO control law, the use of spatial priors is mandatory.

A real AO or MCAO system is not open loop but closed loop. What we call a closed loop is a system in which the WFSs are behind the DMs and analyze the residual phase. This means that the open-loop MMSE approach presented above is not directly implementable in a real system. It turns out that it is necessary to define a closed-loop control law that retains the ability to estimate unseen modes by the use of spatial priors.

The optimal control of a realistic closed-loop system with a temporal delay actually requires that both spatial and temporal priors be accounted for.

We present the system in Subsection 4.A, and in Subsection 4.B we describe the temporal and spatial priors in a state-space model. The estimator proposed in Subsection 4.C gives, for a given temporal sequence of measurements, the best estimate of the turbulent phase in the mean-square sense, knowing the temporal and spatial statistics of turbulence and noise. As explained in Subsection 2.B, once  $\hat{\varphi}^{\text{tur}}$  is known, the optimal voltages are given by Eq. (9).

##### A. Closed-Loop Basic Relationships

The measurements are obtained with an exposure time  $T$ , and the correction  $\varphi^{\text{cor}}(t)$  is constant between  $(n-1)T$  and  $nT$ , where  $n$  corresponds to the frame number. Thus the problem can be discretized by using  $T$  as the sampling period. For any continuous variable  $\mathbf{f}(t)$ , one can associate the discrete quantity  $\mathbf{f}_n$  defined as

$$\mathbf{f}_n = \frac{1}{T} \int_{(n-1)T}^{nT} \mathbf{f}(t) dt. \quad (17)$$

The temporal diagram of the system in Fig. 3 shows how measurements and computations follow one another.

The CCD camera integrates during one sampling period, and it is read out during the following period. Here we assume that the voltage computation is done during the same time period as the CCD readout. The voltages are applied during the following period of time. There is then a two-sampling-period delay between the beginning of the integration and the application of the correction. This corresponds to a rather common case for astronomical systems. Of course, other situations could be accounted for. For example avalanched photodiodes often used in curvature WFSs lead to a significantly reduced readout time. In any case, the total delay cannot be smaller than one sampling period because of the integration time.

The turbulent phase in all layers  $\varphi_n^{\text{tur}}$ , the correction phase in all DMs  $\varphi_n^{\text{cor}}$ , and the residual phase in the pupil in the  $N_{\text{GS}}$  GS directions  $\Phi_n^{\text{res}}$  are linked by

$$\{\phi_{\beta_i}^{\text{res}}\}_n = \Phi_n^{\text{res}} = \mathbf{M}_\beta^L \varphi_n^{\text{tur}} - \mathbf{M}_\beta^{\text{DM}} \varphi_n^{\text{cor}}, \quad (18)$$

where  $\mathbf{M}_\beta^L$  and  $\mathbf{M}_\beta^{\text{DM}}$  are the matrices defined in Subsection 2.A.

Let  $\mathbf{N}$  be the matrix defined in Subsection 2.A and  $\mathbf{u}_{n-1}$  the voltages applied between  $n-1$  and  $n$ ,  $\mathbf{u}_{n-1}$  is linked to the correction phase  $\varphi_n^{\text{cor}}$  induced by the DM between  $n-1$  and  $n$  by the relation

$$\varphi_n^{\text{cor}} = \mathbf{N}\mathbf{u}_{n-1}. \quad (19)$$

It must be noted that Eq. (19) means that the mirror dynamics are neglected, as already mentioned in Subsection 2.B. Vector  $\mathbf{u}_n$ , which is applied between  $n$  and  $n+1$ , should be given from Eq. (9) by the knowledge of  $\hat{\varphi}_{n+1}^{\text{tur}}$ , which minimizes

$$\epsilon''_{n+1} = \langle \|\varphi_{n+1}^{\text{tur}} - \hat{\varphi}_{n+1}^{\text{tur}}\|^2 \rangle_{\varphi, \text{noise}}. \quad (20)$$

Considering the time delay, the measurements that are used to compute  $\mathbf{u}_n$  are

$$\mathbf{Y}_n = \mathbf{D}\Phi_{n-1}^{\text{res}} + \mathbf{w}_n, \quad (21)$$

$\mathbf{D}$  being the matrix defined in Section 3 and  $\mathbf{w}$  a white noise (measurement noise). Its covariance matrix is denoted  $\mathbf{C}_w$ . Then, by using Eqs. (18) and (19), one can get

$$\mathbf{Y}_n = \mathbf{D}(\mathbf{M}_\beta^L \varphi_{n-1}^{\text{tur}} - \mathbf{M}_\beta^{\text{DM}} \mathbf{N}\mathbf{u}_{n-2}) + \mathbf{w}_n. \quad (22)$$

Our prior knowledge of the turbulence temporal evolution can be expressed with an autoregressive (AR) model,

$$\varphi_{n+1}^{\text{tur}} = \mathbf{F}\{\varphi_n^{\text{tur}}, \varphi_{n-1}^{\text{tur}}, \varphi_n^{\text{tur}}, \dots\} + \nu_n, \quad (23)$$

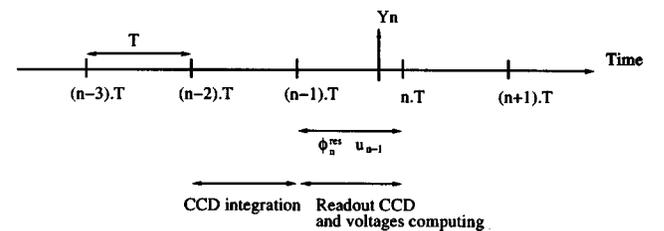


Fig. 3. Temporal diagram showing the different time intervals.  $\varphi_n^{\text{cor}}$ ,  $\varphi_n^{\text{tur}}$ , and  $\Phi_n^{\text{res}}$  are integrated between  $n-1$  and  $n$ , and  $\mathbf{u}_{n-1}$  is applied in the same time interval.

where  $\nu$  is a white noise, of covariance matrix  $\mathbf{C}_\nu$ , and  $\mathbf{F}$  is a linear function. In the rest of the paper, we have chosen to use the following first-order AR prior model:

$$\varphi_{n+1}^{\text{tur}} = \mathbf{A}\varphi_n^{\text{tur}} + \nu_n. \quad (24)$$

In this model,  $\mathbf{C}_\nu$  can be easily determined in order to conserve the global energy of the turbulence; hence

$$\mathbf{C}_\nu = \mathbf{C}_\varphi - \mathbf{A}^T \mathbf{C}_\varphi \mathbf{A}, \quad (25)$$

with  $\mathbf{C}_\varphi$  the covariance matrix of the turbulent phase.

With a first-order-prior model the temporal correlation function decreases exponentially. Real turbulence temporal evolution<sup>31</sup> can be fitted more precisely by using a higher-order model. We will discuss the choice of the prior model in more detail in Subsections 5.A and 5.B.

## B. Linear State-Space Model

A linear state-space model describes the dynamical behavior of a system and its outputs (measurements) by using a state-space vector, whose evolution is given by a linear equation called the state equation.

In our case, the state model based on a state vector  $\mathbf{X}_n$  must summarize the basic relationships of the system into the standard stationary formulation,<sup>19</sup>

$$\mathbf{X}_{n+1} = \mathcal{A}\mathbf{X}_n + \mathcal{B}\mathbf{u}_n + \mathbf{V}_n, \quad (26)$$

$$\mathbf{Y}_n = \mathcal{C}\mathbf{X}_n + \mathbf{w}_n, \quad (27)$$

where  $\mathbf{w}_n$  is the noise defined in Eq. (21) and  $\mathbf{V}_n$  is a Gaussian white noise with covariance matrix  $\mathbf{C}_V$ .

The choice of the state vector is crucial.  $\mathbf{X}$  must contain all the variables necessary for Eqs. (22) and (23) to be summarized into Eqs. (26) and (27) and for the estimation of the voltages  $\mathbf{u}$ . Equation (22) implies then that  $\mathbf{X}_n$  must contain  $\varphi_{n-1}^{\text{tur}}$  and  $\mathbf{u}_{n-2}$ . The voltages  $\mathbf{u}_n$  are determined only through  $\mathbf{X}_n$  so as to correct  $\varphi_{n+1}^{\text{tur}}$  (this corresponds to a prediction). This implies that  $\varphi_{n+1}^{\text{tur}}$  must be in  $\mathbf{X}_n$ .

At this stage,  $\mathbf{X}_n$  is composed of at least  $\varphi_{n+1}^{\text{tur}}$ ,  $\varphi_{n-1}^{\text{tur}}$  and  $\mathbf{u}_{n-2}$ . For the evolution equation for  $\varphi_{n+1}^{\text{tur}}$  to be written,  $\varphi_n^{\text{tur}}$  must be in the state vector, too; and, as  $\mathbf{u}_{n-2}$  must be kept in memory,  $\mathbf{u}_{n-1}$  is also contained in  $\mathbf{X}_n$ .

$\mathbf{X}_n$  needs also to contain all the  $\varphi_{n-i}^{\text{tur}}$  used in Eq. (23). For the first-order prior model considered in this paper, we need only  $\varphi_n^{\text{tur}}$ , which is in any case already included in  $\mathbf{X}_n$ . For a model order less than 3, the state vector is left unchanged, as  $\varphi_n^{\text{tur}}$  and  $\varphi_{n-1}^{\text{tur}}$  are contained in the state vector  $\mathbf{X}_n$ . For a model order higher than 3, the state vector should incorporate a larger number of time steps.

For a first-order-AR prior model, the state vector is then

$$\mathbf{X}_n = \begin{pmatrix} \varphi_{n+1}^{\text{tur}} \\ \varphi_n^{\text{tur}} \\ \varphi_{n-1}^{\text{tur}} \\ \mathbf{u}_{n-1} \\ \mathbf{u}_{n-2} \end{pmatrix}$$

and the state model is

$$\mathbf{X}_{n+1} = \begin{bmatrix} \mathbf{A} & 0 & 0 & 0 & 0 \\ \mathbf{Id} & 0 & 0 & 0 & 0 \\ 0 & \mathbf{Id} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathbf{Id} & 0 \end{bmatrix} \mathbf{X}_n + \begin{bmatrix} 0 \\ 0 \\ 0 \\ \mathbf{Id} \\ 0 \end{bmatrix} \mathbf{u}_n + \begin{pmatrix} \nu_n \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad (28)$$

$$\mathbf{Y}_n = \mathbf{D}[0 \quad 0 \quad \mathbf{M}_\beta^L \quad 0 \quad -\mathbf{M}_\beta^{\text{DM}} \quad \mathbf{N}] \mathbf{X}_n + \mathbf{w}_n. \quad (29)$$

## C. Kalman Filter and Feedback Control

As stated in Subsection 2.B, once the estimation of  $\mathbf{X}_n$  is done in order to minimize Eq. (20), the optimal command is given by Eq. (9).

If a system can be described by a linear state model, the optimal estimation of  $\mathbf{X}_n$ , minimizing a given quadratic criterion, is provided by a Kalman filter,<sup>17</sup>

$$\hat{\mathbf{X}}_{n/n} = \hat{\mathbf{X}}_{n/n-1} + \mathbf{H}_n(\mathbf{Y}_n^{\text{meas}} - \mathcal{C}\hat{\mathbf{X}}_{n/n-1}), \quad (30)$$

where  $\hat{\mathbf{X}}_{n/n}$  is the estimation of  $\mathbf{X}_n$  obtained by using  $\{\mathbf{Y}_0, \dots, \mathbf{Y}_n\}$ . The vector  $\mathbf{Y}_n^{\text{meas}}$  stands for the experimental measurement at  $n$  [as opposed to the measurement model  $\mathbf{Y}_n$  of Eq. (29)]. In a Gaussian framework, this estimate is exactly the posterior mean of  $\mathbf{X}_n$  with knowledge of  $\{\mathbf{Y}_0, \dots, \mathbf{Y}_n\}$ . By taking the conditional mean of both parts in Eq. (26), one obtains the prediction vector  $\hat{\mathbf{X}}_{n+1/n}$  as a linear function of  $\hat{\mathbf{X}}_{n/n}$ . It gives the recursive estimation,

$$\hat{\mathbf{X}}_{n+1/n} = \mathcal{A}\hat{\mathbf{X}}_{n/n-1} + \mathcal{B}\mathbf{u}_n + \mathcal{A}\mathbf{H}_n(\mathbf{Y}_n^{\text{meas}} - \mathcal{C}\hat{\mathbf{X}}_{n/n-1}), \quad (31)$$

where  $\hat{\mathbf{X}}_{n+1/n}$  is the prediction of  $\mathbf{X}_{n+1}$  by use of  $\{\mathbf{Y}_0, \dots, \mathbf{Y}_n\}$ .  $\mathbf{H}_n$  is called the observer gain and is making the trade-off between priors and measurements. It is equal to

$$\mathbf{H}_n = \mathbf{C}_{n/n-1} \mathcal{C}^T (\mathcal{C} \mathbf{C}_{n/n-1} \mathcal{C}^T + \mathbf{C}_w)^{-1}, \quad (32)$$

with  $\mathbf{C}_w$  the matrix covariance of the noise and  $\mathbf{C}_{n/n-1}$  the matrix covariance of the state vector estimation error, predicted for the instant  $n$  at the instant  $n-1$ .  $\mathbf{C}_{n/n-1}$  is computed by solving the Ricatti equation,<sup>17</sup>

$$\begin{aligned} \mathbf{C}_{n+1/n} &= \mathcal{A} \mathbf{C}_{n/n-1} \mathcal{A}^T + \mathbf{C}_V \\ &- \mathcal{A} \mathbf{C}_{n/n-1} \mathcal{C}^T (\mathcal{C} \mathbf{C}_{n/n-1} \mathcal{C}^T + \mathbf{C}_w)^{-1} \mathcal{C} \mathbf{C}_{n/n-1} \mathcal{A}^T. \end{aligned} \quad (33)$$

Practically, the recursive Eq. (31) is the one that has to be implemented. The new measurement is introduced at each step as  $\mathbf{Y}_n^{\text{meas}}$ , and the estimate of  $\hat{\mathbf{X}}_{n+1/n}$  at  $n$  is given by Eq. (31). This means that  $\mathbf{H}_n$  should be estimated at each step, too. The state-model equation does not appear explicitly in the Kalman filter implementation. It is present through matrices  $\mathcal{A}$ ,  $\mathcal{B}$ ,  $\mathcal{C}$ ,  $\mathbf{C}_V$ , and  $\mathbf{C}_w$ . Once initial values for  $\mathbf{X}_0$  and  $\mathbf{C}_0$  have been chosen, Eqs. (31)–(33) are the only ones needed for the estimation iterations.

One must note that, as we have already said, there is a delay between the measurement and the correction. This means that it is necessary to make a prediction of the evolution of the turbulent phase. The approach that we propose makes this prediction by using the equation of

the evolution of the turbulence, induced by Eq. (28). This is why the state vector  $\mathbf{X}_n$  contains  $\varphi_{n+1}^{\text{tur}}$ . The estimation of  $\mathbf{X}_n$  therefore implicitly includes this prediction step.

The optimal voltages are then given by Eq. (9):

$$\mathbf{u}_n = \mathbf{P}_{[\alpha; \text{DM}]} \hat{\varphi}_{n+1/n}^{\text{tur}}, \quad (34)$$

with  $\mathbf{P}_{[\alpha; \text{DM}]}$  the projector given in Subsection 2.B.

It is worth noting that the complete closed-loop system including Kalman filtering is stable as long as the model of Eqs. (28)–(29) is relevant and stable.<sup>19</sup>

#### D. Kalman Filtering in a Classical Adaptive Optics Case

The notation and expressions in Section 3 and 4 are defined for the MCAO case. The classical AO case can of course be seen as a special application of that notation. The classical AO optimal control can be deduced immediately from the MCAO case by replacing  $N_{\text{DM}} = 1$ ,  $N_L = 1$ ,  $N_{\text{GS}} = 1$ ,  $\alpha = \beta$ , the DM altitude  $h' = 0$ , the layer altitude  $h = 0$ , and the matrices  $\mathbf{M}_\alpha^{\text{DM}} = \mathbf{M}_\beta^{\text{DM}} = Id$  and  $\mathbf{M}_\alpha^L = \mathbf{M}_\beta^L = Id$ . This is what we did in the classical AO simulations presented in Subsection 5.C.

#### E. Kalman Filtering and Previous Studies

##### 1. Optimized Modal Gain Integrator

The OMGI was proposed in 1994 by E. Gendron<sup>12,13,22,32</sup> and is used in classical AO systems [the Nasmyth Adaptive Optics System (NAOS), for instance<sup>33</sup>]. This approach minimizes the variance of the residual phase in the context of an integrator control law. It is performed mode by mode, for instance on the eigenmodes of the system (that is, on the basis that diagonalizes  $\mathbf{D}^T \mathbf{D}$ ) or on the Karhunen–Loeve modes. As we will see in Subsections 5.C.3 and 5.D, such a mode-by-mode treatment is suboptimal, especially when dealing with unseen modes. Hence the interest of the multivariable Kalman approach described in this paper.

If the noise and the signal are decorrelated, one can express, in a Laplace transform representation, the variance of the  $i$ th mode as<sup>12,13</sup>

$$\sigma_i^2 = \int |E_i(j\omega)|^2 \langle |\tilde{\varphi}_{\text{tur},i}(j\omega)|^2 \rangle d\omega + \int |H_i(j\omega)|^2 \langle |\tilde{w}_i(j\omega)|^2 \rangle d\omega, \quad (35)$$

where  $\langle |\tilde{\varphi}_{\text{tur},i}|^2 \rangle$  and  $\langle |\tilde{w}_i|^2 \rangle$  are the temporal power spectral density (PSD) of the turbulent phase and of the noise, respectively.  $E_i$  is the so called rejection transfer function, and  $H_i$  is the noise transfer function; they are both defined in the Laplace space and depend on the gain of the integrator.

The higher the gain, the more efficient the correction of the signal, but the more the noise is amplified, too. By minimizing  $\sigma_i^2$  mode by mode with respect to the gain (which appears in  $E_i$  and  $H_i$ ), one can obtain the optimal gain for the modal integrator control law. This gain is applied in the system eigenmode basis. We must note that the gains are generally thresholded for stability reasons. With a total time delay in the loop of two frames, it can be shown that the integrator is strictly stable up to a

unit gain. Yet we will set the maximum gain to 0.5 to respect the standard stability margins.<sup>14,34</sup>

It must be noted here that an integrator control law does not explicitly include a prediction of the turbulence evolution. There is in fact in the integrator an implicit model of evolution that corresponds to a static phase. This model is indeed not very relevant and does not account efficiently for the temporal evolution of turbulence.

##### 2. Closed-Loop Generalization of the Open-Loop Minimum-Mean-Square-Error Approach

The Kalman approach as defined above can be seen as an implementation in closed loop of the open-loop MMSE approach.

Indeed, considering the estimation part, one can first note that the state vector consists only of  $\varphi_n^{\text{tur}}$ , because there is no delay between the measurements and the corrections in the open-loop description used in Section 3 and because there is no time series of turbulent phase and the phase is estimated instant by instant. We also get  $\mathcal{A} = 0$  (which means only that there is no temporal correlation) and an adapted measurement equation ( $\mathcal{C} = \mathbf{DM}_\beta^L$ ). The Kalman estimator then exactly boils down to the open-loop estimator given in Eq. (14): With  $\varphi_n^{\text{tur}}$  as state vector, Eq. (30) becomes

$$\hat{\varphi}_{n/n}^{\text{tur}} = \hat{\varphi}_{n/n-1}^{\text{tur}} + \mathbf{H}_n^1 (\mathbf{Y}_n^{\text{meas}} - \mathbf{D} \hat{\varphi}_{n/n-1}^{\text{tur}}), \quad (36)$$

where  $\mathbf{H}_n^1$  is the new  $\mathbf{H}_n$  adapted to the open-loop system. As  $\mathcal{A} = 0$ , Eq. (33) gives  $\mathbf{C}_{n+1/n} = \mathbf{C}_v$ . Matrix  $\mathbf{H}_n^1$  is then given by Eq. (32) and is equal to

$$\mathbf{H}_n^1 = \mathbf{C}_v (\mathbf{DM}_\beta^L)^T [\mathbf{DM}_\beta^L \mathbf{C}_v (\mathbf{DM}_\beta^L)^T + \mathbf{C}_w]^{-1}, \quad (37)$$

with  $\mathbf{C}_v = \mathbf{C}_\varphi$ . The turbulent phase from one instant to the other is then totally decorrelated, and the prediction for the turbulent phase  $\hat{\varphi}_{n/n-1}^{\text{tur}}$  is 0.

By replacing  $\hat{\varphi}_{n/n-1}^{\text{tur}}$  with 0 in Eq. (36) and  $\mathbf{H}_n^1$  with its expression, one finds indeed the tomographic open-loop optimal expression already given in Eq. (14).

This is quite important because it means that previous results and such conclusions given in open loop with the optimal approach in MCAO,<sup>5,6,29,35</sup> as the optimal number of DMs, number and geometry of GSs, FOV and achievable performance, should also be verified in closed loop.

##### 3. Comparison with Previous Studies on Prediction

In 1998 Dessenne *et al.*<sup>14</sup> proposed a temporal predictor. The global servoloop was assumed to be composed of parallel scalar servoloops applied to some modal coefficients. Each modal control law was then derived from the temporal priors expressed in a frequency representation. The so-called modal predictor was shown to use the temporal priors more efficiently than the OMGI approach.

For one mode, the transfer function of the corrector can be expressed in a  $z$ -transform representation as

$$C(z) = \frac{\sum_{i=0}^{q-1} a_i z^{-i}}{1 + \sum_{i=1}^{p-1} b_i z^{-i}}, \quad (38)$$

where  $(a_i)_{0 < i < q-1}$  and  $(b_i)_{1 < i < p-1}$  are parameters to be adjusted (and  $z$  represents the  $z$ -transform variable);  $p$  and  $q$  define the corrector order.

The corrector can be expressed in a state-model formalism, which uses an implicit model for the evolution of the turbulence. This model can be read for mode  $m$  as

$$\varphi_{n+1}^{\text{tur},m} = \sum_{i=1}^{p-1} b_i \varphi_{n-i}^{\text{tur},m} + \nu_n. \quad (39)$$

The equivalent estimator in the state-model formalism takes the form of Eq. (30) in which the equivalent  $\mathbf{H}_n$  is no longer given by the Kalman filter theory. The equivalent  $\mathbf{H}_n$  is a function of  $(a_i)_{0 < i < q-1}$  and  $(b_i)_{1 < i < p-1}$ , which is no longer optimal. This approach must also deal with the problem of the stability of the control. The parameters  $(a_i)_{0 < i < q-1}$  and  $(b_i)_{1 < i < p-1}$  must then be adjusted under a stability constraint. The Kalman filter, as noted in Subsection 4.C, ensures the stability of the control loop as long as the model used does not diverge, which is ensured here through Eq. (25), and as long as the model is not too far from the system.

Finally, the other main limitation of Dessenne's approach comes from the fact that, like the OMGI, it proposes a mode-by-mode tuning rather than a global optimization of the AO multivariable servoloop.

#### 4. Previous Attempts to Use a Kalman Filter in Classical Adaptive Optics

In 1993, Paschall and Anderson<sup>20</sup> proposed to use a Kalman filter (along with a state-space model) to control a classical AO system. There are several differences between their proposal and ours. First, our control law is adapted to MCAO and not only to the classical AO. Second, our analytical expressions are considerably simplified because we restrict our formulation to the case in which mirror dynamics can be neglected (see the discussion in Subsection 2.B). Finally, we avoid some difficulties by treating the problem as a discrete-time problem from the beginning. Paschall and Anderson were modeling one part of the system as continuous (for instance, the turbulence) and the other as discrete (the WFS), which forces linkage of the two parts later on.

### F. Practical Considerations

#### 1. Number of Operations and Computing Time

Before comparing the performance of a Kalman filter with other types of estimators in different cases, it is interesting to compare the practical operations that are necessary in our approach and in an integrator approach.

The most basic and usually used control law is the so-called integrator law. Each time the system gets a new measurement  $\mathbf{y}_n$ , with an integrator the new command is computed as

$$\mathbf{u}_n = \mathbf{u}_{n-1} + \mathbf{M}_{\text{com}} \mathbf{y}_n^{\text{meas}}, \quad (40)$$

where  $\mathbf{M}_{\text{com}}$  is the command matrix. It is the only parameter that can be adjusted. The integrator control law does not provide any estimation of the turbulent phase.

With a Kalman filter the new command is a projection of one component of the predicted state vector given by Eq. (31).

In fact, in practice, it is possible to decrease the number of stored parameters by dividing the state vector into two parts: the turbulent phases and the voltages. The operation that is actually needed at each step to compute the new phase estimate is given by the first three components of Eq. (31):

$$\begin{pmatrix} \hat{\varphi}_{n+2/n} \\ \hat{\varphi}_{n+1/n} \\ \hat{\varphi}_{n/n} \end{pmatrix} = \mathbf{M}_1 \begin{pmatrix} \hat{\varphi}_{n+1/n-1} \\ \hat{\varphi}_{n/n-1} \\ \hat{\varphi}_{n-1/n-1} \end{pmatrix} + \mathbf{M}_2 (\mathbf{Y}_n^{\text{meas}} - \mathbf{M}_3 \cdot \mathbf{u}_{n-2}). \quad (41)$$

The new command is then deduced from Eq. (41):

$$\mathbf{u}_n = \mathbf{M}_4 \hat{\varphi}_{n+1/n}, \quad (42)$$

where the size of  $\mathbf{u}$  is the number of actuators,  $n_{\text{act}}$ ; the size of  $\varphi_n$  is the number of modes used to describe the turbulent phase,  $n_{\text{mod}}$ ; and the size of  $\mathbf{Y}_n$  is the number of measurements of the WFS,  $n_{\text{WFS}}$ .

On the one hand, for the integrator control law, we need to keep in memory the matrix  $\mathbf{M}_{\text{com}}$ , whose size is  $n_{\text{act}} \times n_{\text{WFS}}$ . On the other hand, for the Kalman and feedback control, we need to keep in memory the matrices  $\mathbf{M}_1$ , whose size is  $3n_{\text{mod}} \times 3n_{\text{mod}}$ ;  $\mathbf{M}_2$ , whose size is  $3n_{\text{mod}} \times n_{\text{WFS}}$ ;  $\mathbf{M}_3$ , whose size is  $n_{\text{WFS}} \times n_{\text{act}}$ ;  $\mathbf{M}_4$ , whose size is  $n_{\text{act}} \times n_{\text{mod}}$ , and the voltages at the current instant and the previous instant  $\mathbf{u}_{n-1}$  and  $\mathbf{u}_n$ . It is also possible to decrease the number of parameters to be kept in memory with sparse matrix considerations. Obviously, the Kalman approach that we suggest still requires us to record more variables than does an integrator approach.

Additionally, it must be noted that  $\mathbf{H}_n$  must be computed at each time step. A solution for dealing with this problem is to use the limit of  $\mathbf{H}_n$  when  $n$  tends toward infinity. This limit can be computed independently, and using it instead of  $\mathbf{H}_n$  limits the performance of the system only during the few first iterations. The major reason that this "trick" is effective is that the Riccati equation [see Eq. (32)] that must be solved to compute  $\mathbf{H}_n$  converges fast.<sup>19,36</sup>

#### 2. Taking into Account the Static Aberrations

Up to now, we have considered that the phase to be corrected was purely a Kolmogorov turbulent phase, evolving in time. In a real system, one must take into account static aberrations, coming essentially from the optical components. It is easy and instructive to examine how to estimate and compensate these aberrations in the Kalman framework. One has only to add a constant phase  $\varphi^{\text{cst}}$  to the turbulent phase  $\varphi^{\text{tur}}$ . The evolution, Eq. (23), then becomes

$$\varphi_{n+1}^{\text{tot}} = \varphi_{n+1}^{\text{cst}} + \varphi_{n+1}^{\text{tur}}, \quad (43)$$

$$\varphi_{n+1}^{\text{tur}} = \mathbf{F}[\varphi_n^{\text{tur}}, \varphi_{n-1}^{\text{tur}}, \varphi_{n-2}^{\text{tur}}, \dots] + \nu_n, \quad (44)$$

$$\varphi_{n+1}^{\text{cst}} = \varphi_n^{\text{cst}}. \quad (45)$$

For a first-order evolution model, Eq. (24), the state vector then becomes

$$\mathbf{X}_n = \begin{pmatrix} \varphi_{n+1}^{\text{cst}} \\ \varphi_{n+1}^{\text{tur}} \\ \varphi_n^{\text{tot}} \\ \varphi_{n-1}^{\text{tot}} \\ \mathbf{u}_{n-1} \\ \mathbf{u}_{n-2} \end{pmatrix}.$$

The state model becomes

$$\begin{aligned} \mathbf{X}_{n+1} = & \begin{bmatrix} \mathbf{Id} & 0 & 0 & 0 & 0 & 0 \\ 0 & \mathbf{A} & 0 & 0 & 0 & 0 \\ \mathbf{Id} & \mathbf{Id} & 0 & 0 & 0 & 0 \\ 0 & 0 & \mathbf{Id} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \mathbf{Id} & 0 \end{bmatrix} \mathbf{X}_n + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \mathbf{Id} \\ 0 \end{bmatrix} \mathbf{u}_n \\ & + \begin{pmatrix} 0 \\ \mathbf{v}_n \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad (46) \\ \mathbf{Y}_n = & \mathbf{D}[0 \quad 0 \quad 0 \quad \mathbf{M}_\beta^L \quad 0 \quad -\mathbf{M}_\beta^{\text{DM}}\mathbf{N}]\mathbf{X}_n + \mathbf{w}_n. \quad (47) \end{aligned}$$

The equations of the Kalman filter [Eqs. (31)–(33)] can then be applied with the new matrices  $\mathcal{A}$ ,  $\mathcal{B}$ ,  $\mathcal{C}$ . Thus, by introducing the static aberrations into the model in this way, it is possible to estimate and compensate them without difficulty.

## 5. SIMULATIONS, RESULTS, AND INTERPRETATIONS

We have presented the theoretical development of an optimal control law based on a Kalman filter, and we have explained why we think it is a promising approach for estimating the turbulent phase in closed loop. We now quantify with numerical simulations the gain brought by this new approach. The idea is first to illustrate the method on the easy and well-understood classical AO case. We then proceed to the more complex and challenging MCAO case. The simulation conditions are representative of typical astronomical observations on a 8-m-class telescope in the near infrared.

We briefly describe in Subsection 5.A the prior models and phase-generation models used in the simulation. Then in Subsection 5.B we discuss the differences between a first-order-AR model and a Taylor model. The classical AO simulations are presented in Subsection 5.C. We compare the Kalman performance with those of the OMGI in Subsection 5.C.2. We first consider a favorable case where both the prior and the phase-generation models follow a first-order AR. We then verify that a Kalman approach based on a simple first-order AR can still conserve its gain over the OMGI when phase generation is based on more representative translating screens (Taylor case). Finally (see Subsection 5.C.3), we artificially introduce unseen modes in the classical AO simulation.

This tutorial case clearly demonstrates the Kalman approach's ability to deal with unseen modes in MCAO.

Finally, we simulate the MCAO case in Subsection 5.D. The performance of the Kalman approach is compared with that of the OMGI generalized to MCAO. We demonstrate the gain of the Kalman approach, particularly in the FOV between the GSs.

### A. Turbulence Models

First of all, one has to understand that there are two different types of turbulence models: on the one hand, the phase-generation model, which is used to generate the time series of the turbulent phase, and on the other hand, the prior model, which is used to build the Kalman estimator as described in Section 4.

In classical AO, the prior model used in all our simulations is a first-order-AR turbulence model presented in Eq. (24). For the phase-generation-model, we use successively two models. The first one is a first-order AR identical to the prior model. The second one is a so-called Taylor model. In this case, we create three phase screens by the McGlamery approach,<sup>37</sup> and we shift them across the pupil with the same wind speed ( $V/D = 2$  Hz) but with different wind directions ( $0^\circ$ ,  $120^\circ$ , and  $240^\circ$ ). We use linear interpolations for translations of a decimal number of pixels.

In MCAO, the prior model and the phase-generation model are identical; they correspond to a first-order-AR model on each layer, with a scaling factor accounting for the turbulence profile  $\mathbf{C}_n^2$ .

Note that in the OMGI there is also an underlying temporal PSD model that is used to optimize the gains [see Eq. (35)]. In all the simulations presented here, the PSD model used for the OMGI is always derived from the phase-generation model itself.

### B. Autoregressive Model

We must now specify in more detail which first-order-AR (AR1) model we used.

The matrix  $\mathbf{A}$  in Eq. (24) has been chosen diagonal, and its elements have been adjusted with respect to the characteristics of the temporal evolution of the turbulence.

More precisely, we enforce a correlation time that decreases with the Zernike radial order. We based this choice on the result, given in Ref. 31, that the cutoff frequency  $f_c$  of the PSD of a Taylor turbulent phase is proportional to  $n + 1$ , where  $n$  is the Zernike radial order. The characteristic time of evolution of the turbulent phase  $\tau_c$  can be approximated by  $\tau_c \approx 1/f_c$  and is then proportional to  $1/(n + 1)$ .

The characteristic time of evolution of the AR-generated turbulence is defined as the correlation time at  $1/e$ . It can be written, for the radial order  $n$ , as  $\tau_c^{\text{AR}} = -1/\log(a_n)$ , where  $a_n$  is the coefficient of matrix  $\mathbf{A}$  for the radial order  $n$  and  $\log$  is the Neperian logarithm.  $\tau_c^{\text{AR}} = \tau_c$  then gives us the relative evolution of  $a_n$  where  $n$  is representative of a turbulence evolution. This approach still leaves us one parameter to adjust, for example,  $a_1$ .

The decorrelation of the first-order-AR turbulence is exponential, this corresponds to a PSD that is constant before a cutoff frequency  $f_c^{\text{AR}}$  and then decreases with a  $f^{-2}$

law. On the other hand, the Taylor PSD also exhibits a cutoff frequency  $f_c$ , but after  $f_c$ , the decrease is much sharper and follows a  $-17/3$  power law.<sup>31,38,39</sup> This means the AR turbulence contains more energy at high temporal frequencies. This difference of behavior between the AR and the Taylor models prompted us to define the concept of equivalent wind speed for the AR-generated turbulence (see Subsection 5.C.1). One must also note that the AR temporal behavior is related to the model order. A higher order could provide a PSD closer to that of a Taylor turbulence. On the other hand, it would also increase the complexity of the filter and could potentially decrease its stability and robustness.

In any case, we will verify in Subsection 5.C.2 that our Kalman control based on a simple first-order-AR prior model can be rather efficient even when operating on a Taylor phase-generation model.

One could also wonder whether our simple first-order-AR prior model truly allows us to perform an efficient global optimization. Indeed,  $\mathbf{A}$  is a diagonal matrix, and  $\mathbf{C}_\varphi$  is quasi-diagonal in the Zernike basis. One must realize, however, that, in general, the measurement equation is not diagonal in the Zernike basis. Despite the simple prior model, the Kalman multivariable loop is therefore far from being equivalent to independent scalar loops. In other words, in the system eigenmode basis, where unseen modes appear, the turbulence covariance matrix is far from diagonal, which means that the eigenmodes are correlated. Our simple prior is sufficient to encode these spatial correlations and give some information about the temporal behavior. The Kalman approach uses all this to recover unseen modes, as will be shown in Subsections 5.C.3 and 5.D. The only case in which the Kalman control is probably close to decoupled scalar loops is the case of Subsection 5.C.2, where, for the sake of simplicity, the WFS is assumed to directly measure Zernike coefficients.

Concerning the temporal prior, it could also be of interest to build a turbulence model that really imposes the Taylor frozen-flow hypothesis. A recent paper<sup>21</sup> has proposed a way to apply such a constraint. In this case both the wind speed and the wind direction can be taken into account. Less informative models, such as our AR description, are still of interest when the wind is not well known or not well defined. That case succeeds when the wave front is the sum of several Taylor contributions in several layers characterized by different wind vectors.

## C. Simulations in Classical Adaptive Optics

### 1. Simulation Conditions

We first describe the turbulence and system simulation conditions. The turbulent phase is composed of the first 13 Zernike radial orders (Zernike polynomials 2–105).  $D/r_0$  is set to 10 at the imaging wavelength. The delay of the loop is two sampling periods (as already mentioned and described in Fig. 3). The WFS directly measures the Zernike polynomials 2–105, meaning that the matrix  $\mathbf{D}$  is assumed to be identity. The measurement noise is Gaussian and representative of a  $9 \times 9$ -microlens Shack–Hartmann WFS. So that the noise would be equivalent to the noise propagation through the recon-

struction from Shack–Hartmann data, the noise was colored<sup>40,41</sup> with a variance proportional to  $(n + 1)^{-2}$ . The SNR was taken between 5 and 50. The SNR is defined as the variance of the slopes on the equivalent SH divided by the measurement noise variance. The sampling frequency  $f_{\text{samp}}$  was taken as  $f_{\text{samp}} = 50$  Hz and  $f_{\text{samp}} = 100$  Hz.

These simulation conditions are representative of near-infrared observation on a 8-m telescope at a good astronomical site. As an example, taking a zero point derived from NAOS<sup>33</sup> for a visible WFS operating at  $0.7 \mu\text{m}$  and considering an imaging wavelength of  $2.2 \mu\text{m}$ , a WFS SNR of 10 is representative of a GS magnitude 15 with a sampling frequency of 100 Hz. Such a sampling frequency is well adapted to such a GS magnitude to provide a good correction. Using a higher frequency would lead to an increased noise level, leading to a reduced loop gain and to no improvement in terms of effective bandwidth.

The DM directly corrects the Zernike polynomials 2–105, which means that  $\mathbf{N} = Id$ . In all the simulation results presented in this paper, we have added the variance generated by the turbulent modes with radial order larger than 13 to the residual variance obtained for the corrected first 13 radial orders. This means that all the variances are computed as<sup>42,43</sup>

$$\sigma_{\text{res}}^2 = \sigma_{\text{res}}^2(n = 1:n_{\text{max}}) + \sum_{n=n_{\text{max}}+1}^{\infty} \sigma_{\text{res}}^2(n), \quad (48)$$

$$\simeq \sigma_{\text{res}}^2(n = 1:n_{\text{max}}) + 0.458(n_{\text{max}} + 1)^{(-5/3)}(D/r_0)^{(5/3)}, \quad (49)$$

where  $n$  represents the radial order and  $n_{\text{max}}$  is the last radial order simulated in the turbulent phase; here  $n_{\text{max}} = 13$ . We still need to give an appropriate definition of an equivalent wind speed. To do so, we simulated an integrator control law (fixed gains 0.5 on each mode) on the classical AO system defined above with a sampling frequency of 100 Hz and a SNR of 50. We estimate the residual variances for the case when this integrator is applied to a Taylor turbulence parameterized by the wind speed  $V$  and to the AR turbulence parameterized by  $a_1$  defined in Subsection 5.A. We say that  $V$  is the equivalent wind speed for a given  $a_1$  if the residual variances are equal. For  $V/D = 2$  Hz and  $f_{\text{samp}} = 100$  Hz, we found  $a_1 = 0.99014$ .

### 2. Comparison of the Kalman Approach with the Optimized Modal Gain Integrator

We mentioned that here  $\mathbf{D}$  was the identity, which means that the eigenmodes of the system are the Zernike polynomials themselves. The gains are therefore optimized on the Zernike polynomials. They are thresholded to 0.5 for stability reasons.

To compare the two approaches, we have defined the factor  $\rho$ , which gives the enhancement factor of a Kalman filter compared with the OMGI. If  $\sigma_{\text{res}}^2$  (Kalman) and  $\sigma_{\text{res}}^2$  (OMGI) are the variances of the residual phases obtained with the two methods, the enhancement factor is defined as

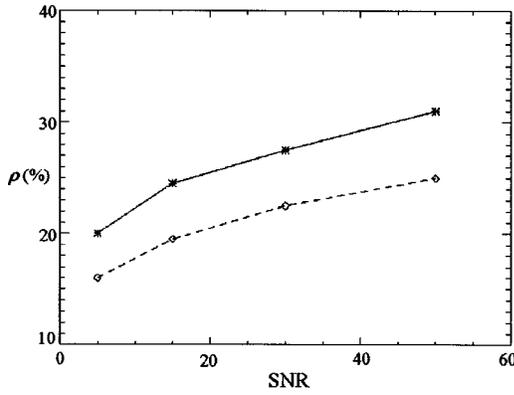


Fig. 4. Enhancement factor of the Kalman filter  $\rho$ , in %, versus SNR for two sampling frequencies, 50 Hz (solid curve) and 100 Hz (dashed curve).

$$\rho = \frac{\sigma_{\text{res}}^2(\text{OMGI}) - \sigma_{\text{res}}^2(\text{Kalman})}{\sigma_{\text{res}}^2(\text{OMGI})}. \quad (50)$$

If  $\sigma_{\text{res}}^2(\text{Kalman})$  is 0,  $\rho$  is 100%. If it is equal to  $\sigma_{\text{res}}^2(\text{OMGI})$ ,  $\rho$  is equal to 0%.

Figure 4 represents the evolution of  $\rho$  with the SNR for these two sampling frequencies. For 100 Hz, the enhancement factor goes from 16% to 25%, and for 50 Hz it goes from 20% to 31% when SNR increases.

In terms of absolute performance, the Strehl ratios, given respectively for Kalman and the OMGI, are 28% and 22% for 100 Hz and SNR = 15 and 37% and 29% for 100 Hz and SNR = 50.

It is easy to understand why the enhancement factor increases when the sampling frequency decreases. The lower the sampling frequency, the more the turbulent phase changes between two measurements and the more we need to make a good prediction. As already said in Subsection 4.C, the Kalman approach provides a prediction of the evolution of the turbulent phase, whereas the OMGI approach does not; hence  $\rho$  increases when the sampling frequency decreases.

Now,  $\rho$  increases with the SNR. Actually, if the SNR tends toward zero, all estimators tend to be equivalent, and if the SNR tends toward infinity, the phase estimation is better and thus provides a better prediction.

Our results can be compared directly with those obtained by Dessenne and co-workers with a modal temporal predictor,<sup>14,34</sup> because we took similar simulation conditions. The two results are very close. This is not surprising, since it corresponds to two different ways of implementing an efficient temporal prediction. The advantage of the Kalman approach is that it is easy to derive and avoids the stability constraints that have to be imposed in Dessenne's approach. It is also more efficient since the optimization is global rather than mode by mode. Here, however, with our simplistic WFS model, and as mentioned in Subsection 5.B, the global control is probably almost equivalent to decoupled scalar loops. With real-world WFS and DM models, the global approach provides an appropriate framework for dealing with the usual waffle modes.

In a second step, as real systems have to compensate multilayer Taylor turbulence, we wanted to apply a Kal-

man filter built on the first-order-AR prior model to a Taylor turbulence and to quantify its performance. We already said in Subsection 5.A that in this case, we simulate the turbulence as three layers translating with a constant speed but in three different directions (0°, 120°, and 240°). A summation of the three screens gives the turbulent phase on the pupil. By projecting the resulting screen on the Zernike basis, we obtain the WFS measurements. One must keep in mind that we consider a sensor that directly measures the Zernike coefficients.

The previous simulations were then made again for two cases: case (1), the Kalman filter and the OMGI applied on the AR phase-generation model, and case (2), the Kalman filter and the OMGI applied on the Taylor phase-generation model. The prior model used for the Kalman approach is always the AR model, whereas, as mentioned in Subsection 5.A, the OMGI PSD model is derived from the phase-generation model itself. The simulations are done under the conditions used for equivalent-wind-speed estimation ( $\nu/D = 2$  Hz,  $f_{\text{samp}} = 100$  Hz, or  $a_1 = 0.99014$ ).

We have observed that the enhancement factor  $\rho$  is not sensitive to the turbulence generator used; i.e.,  $\rho_{\text{AR}} \approx \rho_{\text{Taylor}}$ , where  $\rho_{\text{AR}}$  and  $\rho_{\text{Taylor}}$  are obtained from the Kalman and the OMGI results of case (1) and case (2), respectively. This means that, even if the first-order turbulence model is not optimal, the Kalman filter induced by this model does not lose its advantage over the OMGI. This is a first argument to say that the Kalman filter is robust to model errors. The interest of using the Kalman approach is kept even if the evolution model is approximative. This is an important point, because it justifies *a posteriori* the use of the first-order-AR turbulence model. One should investigate whether a higher-order-AR model could improve the Kalman performance.

### 3. MCAO-like Case, Introduction of Unseen Modes, and Correction of These Modes

The gain given by the Kalman approach in Subsection 5.C.2 is essentially linked to its ability to make a good temporal prediction. We now illustrate its ability to deal with unseen modes.

In Subsection 5.C.2, the same 2–105 Zernike polynomials were used to describe the turbulent phase, the measured phase, and the correction phase. As we said, this means that the interaction matrix  $\mathbf{D}$  between the turbulent phase and the measured phase was the identity matrix.

We now artificially introduce unseen modes in a classical AO case. To do so, we have chosen to use a matrix  $\mathbf{D}$  that mixes two polynomials. More precisely, we consider here that the WFS cannot distinguish  $Z_4$  and  $Z_{17}$  and that it measures only the average of the corresponding Zernike coefficients.

If  $\{z_i\}$  are the coefficients of the turbulent phase on the Zernike basis, the measurement is then, apart from the noise,

$$\mathbf{y}_4 = \frac{z_4 + z_{17}}{2}, \quad (51)$$

$$\mathbf{y}_{17} = \frac{z_4 + z_{17}}{2}, \quad (52)$$

$$\mathbf{y}_i = z_i \quad \text{for } i \neq 4 \text{ and } i \neq 17 \quad (53)$$

instead of  $y_i = z_i$  for all modes as previously. With this matrix  $\mathbf{D}$  we create one unseen mode:  $Z_4 - Z_{17}$ . We chose those two modes because  $Z_4$ , a low-order mode, is very energetic, while  $Z_{17}$ , a high-order mode, is less energetic. Estimating  $z_4$  and  $z_{17}$  correctly is a problem that is then very similar to the MCAO unseen-mode problem presented in Subsection 3.C.

In Fig. 5 we present the variance of the residual phase mode by mode for the Kalman approach and for the OMGI in the presence of the unseen mode. We show as a reference the variances obtained in the case without the unseen mode described in Subsection 5.C.2. In the reference case, the difference between the two methods is essentially due to the prediction step included in the Kalman filter. This is a modal illustration of the results of Subsection 5.C.2.

In the case with unseen modes, the OMGI simply filters out the unseen modes and thus introduces a large estimation error on both  $z_4$  and  $z_{17}$ . A first analytical estimation of this error can be provided by an open-loop reasoning. The OMGI affects half of the measurement to each mode because of the lack of prior information on the relative energy of modes 4 and 17. Thus it overestimates the energy of  $z_{17}$  and underestimates the energy of  $z_4$ . This gives the following expected residual variances:

$$\begin{aligned} \sigma_{\text{res,OMGI},z_4}^2 &\simeq \sigma_{\text{res,OMGI},z_{17}}^2 \simeq \frac{\sigma_{\text{turb},z_4}^2}{4} + \frac{\sigma_{\text{turb},z_{17}}^2}{4} \\ &\simeq \frac{\sigma_{\text{turb},z_4}^2}{4}. \end{aligned} \quad (54)$$

This is what is observed in Fig. 5. The variance of the turbulent phase is also plotted for comparison. Note that  $\sigma_{\text{res,OMGI},z_{17}}^2$  is even larger than the turbulence variance.

The Kalman approach performs a global optimization that makes use of the spatial priors to estimate these two modes. The result corresponds to what was expected:

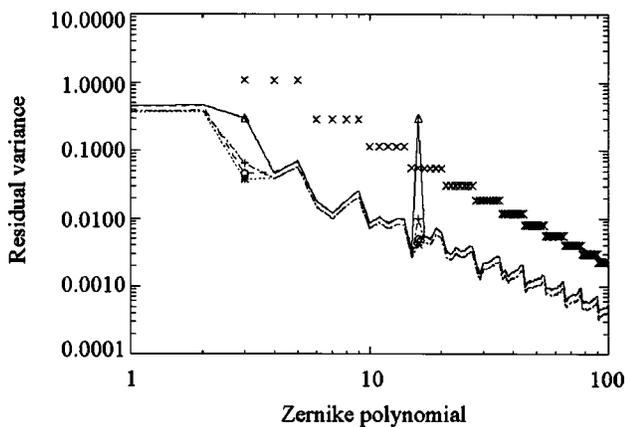


Fig. 5. Variance of the residual phase (in  $\text{rad}^2$ ) with unseen modes for a Kalman estimator (dashed-dotted curve) or the OMGI approach (solid curve) and without unseen modes with the Kalman estimator (dotted curve) or the OMGI approach (dashed curve). The turbulent-phase variance is also plotted for comparison ( $\times$ 's). As the piston mode is not considered here, the  $x$  axis begins with the tilt mode.  $x = 1$  stands for the tilt. The values for  $Z_4$  and  $Z_{17}$  are underlined on the curves by adding the symbols  $+$ ,  $\Delta$ ,  $*$ , and  $\diamond$ .

the loss of performance is much smaller than in the OMGI case. We obtain, for  $z_4$ ,  $\sigma_{\text{res,OMGI},z_4}^2 = 0.3 \text{ rad}^2$  and  $\sigma_{\text{res,Kalman},z_4}^2 = 0.06 \text{ rad}^2$ , and  $\sigma_{\text{turb},z_4}^2 = 1.1 \text{ rad}^2$ . For  $z_{17}$ ,  $\sigma_{\text{res,OMGI},z_{17}}^2 = 0.3 \text{ rad}^2$  and  $\sigma_{\text{res,Kalman},z_{17}}^2 = 0.01 \text{ rad}^2$  while  $\sigma_{\text{turb},z_{17}}^2 = 0.05 \text{ rad}^2$ .

In other words, it is the use of spatial correlations between the unseen modes, here  $Z_4 - Z_{17}$ , and the other modes that allow the Kalman estimator to conserve good performance.

This ability to deal with unseen modes will now be illustrated on MCAO simulations. Note that it could also be of interest in real classical AO systems, since unseen modes, the so-called waffle modes, generally appear with real-world WFS and DM characteristics. The Kalman approach could then avoid the usual *ad hoc* filtering procedures used in this case. It can also provide a better rejection of these modes, which is important for high-dynamic-range AO.

## D. Simulations in Multiconjugate Adaptive Optics

### 1. Simulation Conditions

We first present the turbulence and system simulation conditions. We consider a two-layer atmosphere. The layer altitudes are 500 m and 10 km, and the strength of the turbulence is 80% in the lower layer and 20% in the higher one. The global  $D/r_0$  is set to 9. For an 8-m telescope diameter,  $r_0 = 0.89 \text{ m}$  at  $2.2 \mu\text{m}$  and  $\theta_0 = 8.59''$  at  $2.2 \mu\text{m}$ , which is representative of astronomical sites.

For the maximum spatial frequency to be the same on the two layers, the number of Zernike radial orders in each layer should be proportional to the size of the pupil on the layer, the so-called metapupil. The size of the metapupil is determined by the projections, on the layer, of the telescope pupil in the whole FOV (2 arc min). The metapupil diameters are then respectively 8.3 m and 13.8 m at 500 m and 10 km. We chose a factor 2 between the number of radial orders on the lower and on the higher layers. We then created a turbulent phase composed of 13 radial orders in the first layer and 26 in the second one. The phase is generated layer by layer with an AR process in the same way as in classical AO. The parameters  $a_1^{500 \text{ m}}$  and  $a_1^{10 \text{ km}}$  are both equal to 0.99104.

The results presented here correspond to an 8-m-class telescope observing in the near infrared ( $2.2 \mu\text{m}$ ). For the wave-front sensing we use three GSs located on the vertices of an equilateral triangle inscribed in a FOV of 2 arc min, as shown in Fig. 6. The SNR on Shack-Hartmann measurements<sup>41</sup> is equal to 10.

The sampling frequency is 100 Hz, and the delay of the loop is two sampling periods. The WFS can measure 13 radial orders of Zernike polynomials, and the noise on the measurements is representative of a  $12 \times 12$ -microlens Shack-Hartmann WFS. We use two DMs conjugated at 500 m and 10 km, that is, conjugated on the turbulent layers themselves, mirrors that can correct 13 radial orders for the lower layer and 26 radial orders of Zernike polynomials for the higher layer. The choice of the factor 2 between the number of corrected radial orders on the lower and the higher layers is motivated by the same reasons as previously for the turbulent phase. (Since the

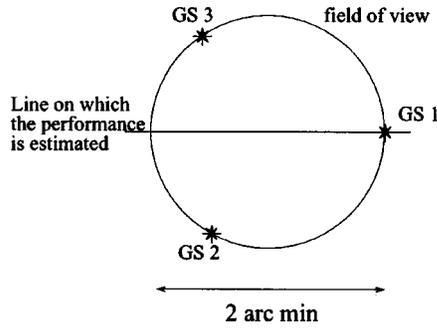


Fig. 6. Illustration of the GSs geometry and of the directions for which the performance is estimated.

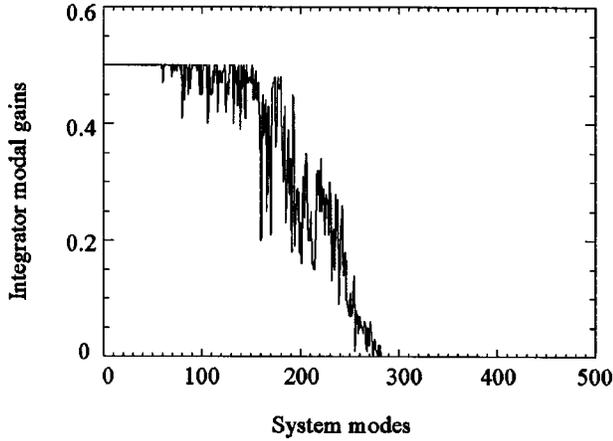


Fig. 7. Optimized model gains in the MCAO case, in the basis of the eigenmodes of  $[\mathbf{DM}_\beta^{\text{DM}}\mathbf{N}]^T[\mathbf{DM}_\beta^{\text{DM}}\mathbf{N}]$ . Note that the effective  $x$ -axis range is the number of eigenmodes, 482 in this case.

DMs are located at the same altitudes as the turbulent layers, the metapupils sizes are also identical.)

We compute the variances of the residual phase in different directions. In this computation, we take into account the Noll residual variance through Eq. (49) with  $n_{\text{max}} = 13$ .

We also present the Strehl ratio, which is approximated as the coherent energy  $\exp(-\sigma_{\text{res}}^2)$ . This is a good approximation for good corrections.

### 2. Results and Interpretations

We compared the Kalman filter performance in the MCAO case with a new approach, the Multiconjugate OMGI (MOMGI), which is a generalization of the MCAO of the OMGI approach in classical AO.<sup>12,13</sup> The gains of the MOMGI are optimized (as explained in Subsection 5.C.2) on the basis of the eigenmodes of  $[\mathbf{DM}_\beta^{\text{DM}}\mathbf{N}]^T[\mathbf{DM}_\beta^{\text{DM}}\mathbf{N}]$ . As in Subsection 5.C.2, the gains of the MOMGI estimator have been thresholded to 0.5 for stability reasons.<sup>14</sup> We recall that  $\mathbf{M}_\beta^L$  is the matrix that performs the summation on the turbulent layers in all directions  $\beta$ . In this basis, the modes whose eigenvalues are low are badly seen, and the modes whose eigenvalues are 0 are unseen. The lower the eigenvalue, the lower the gain on this mode. The optimized gains in this basis are plotted in Fig. 7. One must note that the number of lines of matrix  $\mathbf{DM}_\beta^{\text{DM}}\mathbf{N}$  is the total number of measured modes  $3 \times 104 = 312$  modes (3 GSs, 104 Zernikes per

GS, 13 radial orders without the piston mode). The number of columns is the total number of corrected modes,  $104 + 378 = 482$  modes (104 Zernikes for the first DM, 378 for the second one, corresponding to 15 radial orders). Matrix  $\mathbf{DM}_\beta^{\text{DM}}\mathbf{N}$  is then  $312 \times 482$ . This means that the matrix  $[\mathbf{DM}_\beta^{\text{DM}}\mathbf{N}]^T[\mathbf{DM}_\beta^{\text{DM}}\mathbf{N}]$  is  $482 \times 482$ . There are then 482 eigenmodes and eigenvalues, but the rank of  $[\mathbf{DM}_\beta^{\text{DM}}\mathbf{N}]^T[\mathbf{DM}_\beta^{\text{DM}}\mathbf{N}]$  is at most 312, which means there are at least  $482 - 312 = 170$  zeros in the eigenvalues. This is the reason why there are so many zeros in the gains observed in Fig. 7.

As we can see, the number of unseen modes in MCAO can be large; the number is, in fact, a complex function of the GS number and geometry and of the FOV of interest. In our case, with three GSs, we observe that there are many unseen modes. In classical AO, unseen modes are less of an issue since the systems are usually designed so that the number of correction modes is directly related to the number of measured modes.

As the gains decrease with the eigenvalues, the unseen modes are filtered out by the MOMGI approach, while the Kalman filter estimates them by using spatial *a priori* knowledge. As we know from previous studies,<sup>6</sup> the estimation of unseen modes can be critical for the performance of the system in the FOV between the GSs. We then expect a significant gain for the Kalman approach.

Figure 8 shows the Strehl ratio along a line joining the center of the GS triangle and one of the GSs (as shown in Fig. 6) for the two approaches and for the classical AO OMGI case. The difference of performance observed in the GS direction between the Kalman approach and the others is due to the temporal error. The Kalman approach that we propose provides a prediction. For each case the best performance is obtained on the GS and degrades away from it. Our approach provides a noticeable improvement over the MOMGI and a better interpolation between the GSs. The difference between the two corresponds to a few percent of Strehl ratio on the GS and 10% on the border of the FOV.

In Figs. 9 and 10 we present the variance of the residual phase on the pupil versus the Zernike mode in two directions: the GS direction and the center of the FOV.

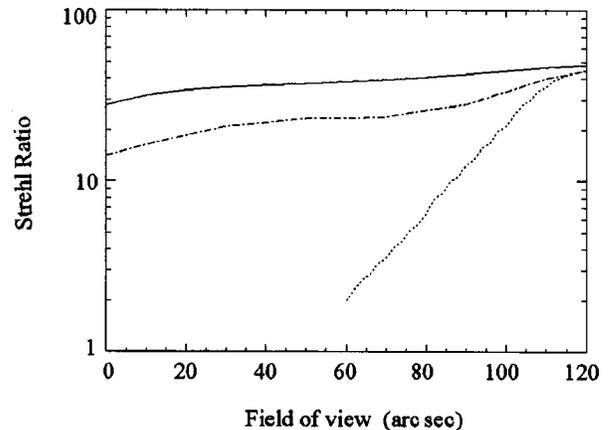


Fig. 8. Comparison between the Kalman approach (solid curve), the MOMGI approach (dashed-dotted curve), and the classical AO case (dotted curve). The Strehl ratio (in %) is plotted versus the position in the FOV in arc seconds.

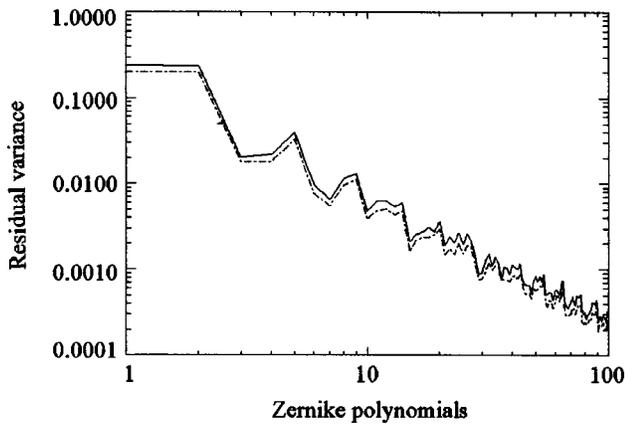


Fig. 9. Residual phase variance (in  $\text{rad}^2$ ) as a function of the Zernike mode for the Kalman approach (dashed-dotted curve) and the MOMGI approach (solid curve) on a GS. As the piston mode is not considered here, the  $x$  axis begins with the tilt.  $x = 1$  stands for the tilt.

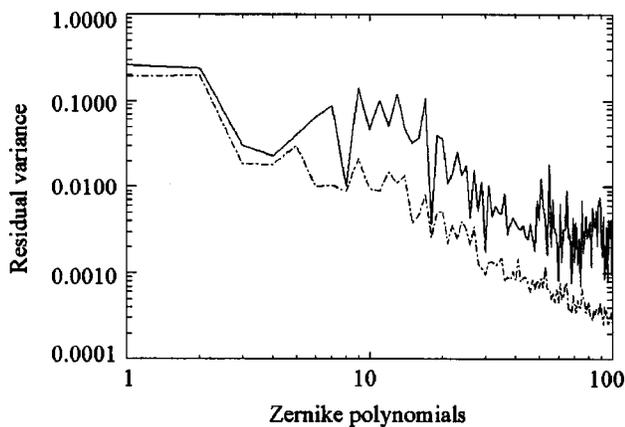


Fig. 10. Residual phase variance (in  $\text{rad}^2$ ) as a function of the Zernike mode for the Kalman approach (dashed-dotted curve) and the MOMGI approach (solid curve) at the center of the FOV. As the piston mode is not considered here, the  $x$  axis begins with the tilt.  $x = 1$  stands for the tilt.

The deterioration of the integrator performance at the center of the FOV is obvious and is due to the presence of unseen or badly seen modes.

It must be noted that we also performed the same MCAO simulation with a slightly irregular GS geometry (distance of the GSs from the center of the FOV equal to 50", 60", 70" instead of 60", 60", 60"). The gain given by the Kalman approach is conserved even with such a geometry. Nevertheless, if the geometry is irregular and the number of GSs is large, unseen modes may be less numerous.<sup>27</sup>

#### E. Discussion: Interest of the Kalman Filter

It has been shown in the previous sections that a Kalman filter provides a significant benefit in the performance of an AO or a MCAO system. In this subsection we stress many practical advantages.

First of all, the state-space approach provides a rigorous framework for the use of spatial and temporal priors that are necessary for dealing with unseen modes, and

the Kalman filter provides a fully optimal estimation with separation between estimation and control.

Second, the question of the stability of the control, which was constraining the control laws previously proposed,<sup>12-14</sup> can be avoided when the model used in the Kalman filtering is relevant and stable itself.<sup>19</sup> It has been explained that the Kalman approach uses a physical description of the system through the state equations, Eqs. (28) and (29). Any kind of control implicitly uses a model for turbulence evolution and measurement. In the approach proposed, we explicitly write this model. This allows us to make physical interpretations of the control law and to understand physically the behavior of the control.

It then becomes easy to introduce into the system model any pertinent parameter or behavior. We have shown this for the problem of static aberrations. To take these aberrations into account, we need only adapt the model equations, which does not change the structure of the control. In the same way, various phenomena that generally limit AO performance can be introduced into the model so that they can be taken into account. For instance, aliasing, vibrations of the telescope, or miscalibrations can be handled efficiently by introducing them into the state equations, Eqs. (28) and (29).

We have already suggested in this paper that it is not necessary to obtain a very precise knowledge about the priors used in the Kalman filter (Subsection 5.C.2). This is a classical issue of regularized estimation, and we have already observed in Ref. 44 that precision of spatial-prior-knowledge is not critical in open-loop MCAO.

We therefore used a temporal turbulence model that does not require us to know the wind direction in each layer, which would be in practice quite difficult to obtain. Imposing the PSD on each mode is informative enough and gives, as we saw in this paper, very good results.

## 6. CONCLUSION

We have presented an optimal closed-loop control law for multiconjugate adaptive optics (MCAO). MCAO control has to deal with a larger number of degrees of freedom than classical AO and needs a more complex process for estimating the turbulent phase. The solution that we propose is a global approach, as opposed to the previous modal approaches, that is based on a linear state-space model with a Kalman estimator. This approach gives an optimal estimation of the turbulence in closed loop. It incorporates both spatial and temporal information on the turbulent phase, as well as information on the system noise statistics, through the so-called state-space model. The temporal priors allow us to make good predictions of turbulence evolution, and the spatial priors allow us to deal with the unseen modes. Furthermore, this approach is flexible enough to allow us to easily take into account various physical parameters or phenomena.

We have shown through a numerical simulation that this approach gives much better results than the usual techniques. The performance has been quantified in classical AO and MCAO and compared with the performance of the optimized modal gain integrator approach.

The major gain comes from the prediction in classical AO and from the estimation of unseen modes in MCAO.

To our knowledge, this is the first time that a Kalman-based approach has been proposed for MCAO. The approach would allow us to optimize future MCAO systems, such as MAD<sup>45–47</sup> or FALCON<sup>48,49</sup> projects, and to relax technical constraints on those projects. This approach should also be quite promising for very-high-Strehl-ratio AO systems, since it has the potential to handle efficiently various effects that generally limit classical AO performance.<sup>50–52</sup>

Corresponding author Brice Le Roux may be reached by e-mail at leroux@arcetri.astro.it.

## REFERENCES

- R. Dicke, "Phase-contrast detection of telescope seeing errors and their correction," *Astrophys. J.* **198**, 605–615 (1975).
- J. Beckers, "Increasing the size of the isoplanatic patch with multiconjugate adaptive optics," in *Very Large Telescopes and Their Instrumentation*, M.-H. Ulrich, ed., Vol. 2 of ESO Conference and Workshop Proceedings (European Southern Observatory, Garching, Germany, 1988), pp. 693–703.
- M. Tallon, R. Foy, and J. Vernin, "3-d wavefront sensing and multiconjugate adaptive optics," in *Progress in Telescope and Instrumentation Technologies* (M.-H. Ulrich, ed., Vol. 42 of ESO Conference and Workshop Proceedings (European Southern Observatory, Garching, Germany, 1992), pp. 517–521.
- B. L. Ellerbroek, "First-order performance evaluation of adaptive-optics systems for atmospheric-turbulence compensation in extended-field-of-view astronomical telescopes," *J. Opt. Soc. Am. A* **11**, 783–805 (1994).
- T. Fusco, J.-M. Conan, V. Michau, L. Mugnier, and G. Rousset, "Efficient phase estimation for large-field-of-view adaptive optics," *Opt. Lett.* **24**, 1472–1474 (1999).
- T. Fusco, J.-M. Conan, G. Rousset, L. Mugnier, and V. Michau, "Optimal wave-front reconstruction strategies for multiconjugate adaptive optics," *J. Opt. Soc. Am. A* **18**, 2527–2538 (2001).
- M. Le Louarn, "Étoiles laser pour les grands télescopes: effet de cône et implications astrophysiques," Ph.D. thesis (Université Lyon I–Claude Bernard, Lyon, France, 2000).
- R. Ragazzoni, E. Marchetti, and F. Rigaut, "Modal tomography for adaptive optics," *Astron. Astrophys.* **342**, 53–56 (1999).
- R. Ragazzoni, "No laser guide stars for adaptive optics in giant telescopes," *Astron. Astrophys., Suppl. Ser.* **136**, 205–209 (1999).
- F. Rigaut, B. Ellerbroek, and R. Flicker, "Principles, limitations and performance of multi-conjugate adaptive optics," in *Adaptive Optical Systems Technology*, P. L. Wizinowich, ed., Proc. SPIE **4007**, 1022–1031 (2000).
- A. Tokovinin, M. Le Louarn, and M. Sarazin, "Isoplanatism in a multiconjugate adaptive optics system," *J. Opt. Soc. Am. A* **17**, 1819–1827 (2000).
- E. Gendron and P. Léna, "Astronomical adaptive optics I. Modal control optimization," *Astron. Astrophys.* **291**, 337–347 (1994).
- E. Gendron and P. Léna, "Astronomical adaptive optics II. Experimental results of an optimize modal control," *Astron. Astrophys.* **111**, 153–167 (1994).
- C. Dessenne, P.-Y. Madec, and G. Rousset, "Optimization of a predictive controller for closed-loop adaptive optics," *Appl. Opt.* **37**, 4623–4633 (1998).
- B. L. Ellerbroek and T. A. Rhoadarmer, "Optimizing the performance of closed-loop adaptive-optics control systems on the basis of experimentally measured performance data," *J. Opt. Soc. Am. A* **14**, 1975–1987 (1997).
- B. L. Ellerbroek and T. A. Rhoadarmer, "Adaptive wave-front control algorithms for closed loop adaptive optics," *Math. Comput. Model.* **33**, 145–158 (2001).
- H. Van Trees, *Detection, Estimation and Modulation Theory*, Part I (Wiley, New York, 1968).
- A. Papoulis, *Probability, Random Variables and Stochastic Processes*, 3rd ed. (McGraw-Hill, New York, 1991).
- B. Anderson and J. Moore, *Optimal Control, Linear Quadratic Methods* (Prentice Hall, London, 1990).
- R. Paschall and D. Anderson, "Linear quadratic Gaussian control of a deformable mirror adaptive optics system with time-delayed measurements," *Appl. Opt.* **32**, 6347–6358 (1993).
- D. Gavel and D. Wiberg, "Toward Strehl-optimizing adaptive optics controllers," in *Adaptive Optical System Technologies II*, P. L. Wizinowich, ed., Proc. SPIE **4839**, 890–901 (2002).
- P.-Y. Madec, "Control techniques," in *Adaptive Optics in Astronomy*, F. Roddier, ed. (Cambridge U. Press, Cambridge, UK, 1999), Chap. 6, pp. 131–154. See also Ref. 23.
- F. Roddier, ed., *Adaptive Optics in Astronomy* (Cambridge U. Press, Cambridge, UK, 1999).
- M. Séchaud, "Wave-front compensation devices," in *Adaptive Optics in Astronomy*, F. Roddier, ed., (Cambridge U. Press, Cambridge, UK, 1999), Chap. 4, pp. 57–90. See also Ref. 23.
- T. Fusco, J.-M. Conan, V. Michau, and G. Rousset, "Noise propagation for multiconjugate adaptive optics systems," in *Optics in Atmospheric Propagation and Adaptive Systems IV*, A. Kohnle, J. D. Gonglewski, and T. J. Schmutge, eds., Proc. SPIE **4538**, 144–155 (2002).
- M. Le Louarn, "Multiconjugate adaptive optics: a PSF study," in *Beyond Conventional Adaptive Optics*, S. Esposito and R. Ragazzoni, eds., Vol. 58 of ESO Conference and Workshop Proceedings (European Southern Observatory, Garching, Germany, 2001), pp. 217–222.
- E. Diolaiti, C. Arcidiacono, R. Ragazzoni, and E. Fedrigo, "Identification and rejection of waffle modes in layer-oriented adaptive optics," in *Adaptive Optical System Technologies II*, P. L. Wizinowich, ed., Proc. SPIE **4839**, 1001–1010 (2003).
- R. Flicker, F. Rigaut, and B. Ellerbroek, "Comparison of multiconjugate adaptive optics configurations and control algorithms for the gemini-south 8-m telescope," in *Adaptive Optical Systems Technology*, W. Bellingham, ed., Vol. 4007 of ESO Conference and Workshop Proceedings (European Southern Observatory, Garching, Germany, 2000), pp. 1032–1043.
- T. Fusco, J.-M. Conan, V. Michau, G. Rousset, and F. Assémat, "Multi-conjugate adaptive optics: comparison of phase reconstruction approaches for large field of view," in *Optics in Atmospheric Propagation and Adaptive Systems VI*, P. L. Wizinowich, ed., Proc. SPIE **4167**, 168–179 (2001).
- T. Fusco, J.-M. Conan, V. Michau, G. Rousset, and L. Mugnier, "Isoplanatic angle and optimal guide star separation for multiconjugate adaptive optics," in *Adaptive Optical Systems Technology*, P. L. Wizinowich, ed., Proc. SPIE **4007**, 1044–1055 (2000).
- J.-M. Conan, G. Rousset, and P.-Y. Madec, "Wave-front temporal spectra in high-resolution imaging through turbulence," *J. Opt. Soc. Am. A* **12**, 1559–1570 (1995).
- E. Gendron, "Optimisation de la commande modale en optique adaptative: application à l'astronomie," Ph.D. thesis (Université Paris VII, Paris, 1995).
- G. Rousset, F. Lacombe, P. Puget, E. Gendron, N. Hubin, G. Zins, E. Stadler, J. Charton, P. Gigan, and P. Feautrier, "Status of the VLT Nasmyth Adaptive Optics System (NAOS)," in *Adaptive Optical Systems Technology*, P. L. Wizinowich, ed., Proc. SPIE **4007**, 72–81 (2000).
- C. Dessenne, "Commande modale et prédictive en optique adaptative classique," Ph.D. thesis (Université Paris VII, Paris, 1998).
- T. Fusco, "Correction partielle et anisoplanetisme en optique adaptative: traitements *a posteriori* et optique adaptative multiconjuguée," Ph.D. thesis (Université de Nice Sophia-Antipolis, Nice, France, 2000).

36. P. S. Maybeck, *Stochastic Models, Estimation and Control, Vol. 3* (Academic, New York, 1982).
37. B. McGlamery, "Computer simulation studies of compensation of turbulence degraded images," in *Image Processing*, J. C. Urbach, ed., Proc. SPIE **74**, 225–233 (1976).
38. F. Roddier, M. J. Northcott, J.-E. Graves, D. L. McKenna, and D. Roddier, "One-dimensional spectra of turbulence-induced Zernike aberrations: time-delay and isoplanicity error in partial adaptive optics," J. Opt. Soc. Am. A **10**, 957–965 (1993).
39. P.-Y. Madec, J.-M. Conan, and G. Rousset, "Temporal characterisation of atmospheric wavefront for adaptive optics," in *Progress in Telescope and Instrumentation Technologies*, M.-H. Ulrich, ed., Vol. 42 of ESO Conference and Workshop Proceedings, (European Southern Observatory, Garching, Germany, 1992), pp. 471–474.
40. G. Rousset, "Wave-front sensors," *Adaptive Optics in Astronomy*, F. Roddier, ed. (Cambridge U. Press, Cambridge, UK, 1999), Chap. 5, pp. 119, para. 5.5.2. See also Ref. 23.
41. F. Rigaut and E. Gendron, "Laser guide star in adaptive optics: the tilt determination problem," *Astron. Astrophys.* **261**, 677–684 (1992).
42. J.-M. Conan, "Etude de la correction partielle en optique adaptative," Ph.D. thesis (Université Paris XI, Paris, 1994).
43. R. Noll, "Zernike polynomials and atmospheric turbulence," J. Opt. Soc. Am. **66**, 207–211 (1976).
44. J.-M. Conan, B. Le Roux, D. Bello, T. Fusco, and G. Rousset, "Multiconjugate adaptive optics: performance with optimal wavefront reconstruction," in *Beyond Conventional Adaptive Optics*, S. Esposito and R. Ragazzoni, eds., Vol. 58 of ESO Conference and Workshop Proceedings (European Southern Observatory, Garching, Germany, 2001), pp. 209–215.
45. N. Hubin, E. Marchetti, E. Fedrigo, R. Conan, R. Ragazzoni, E. Diolaiti, M. Tordi, G. Rousset, T. Fusco, P.-Y. Madec, D. Butler, S. Hippler, and S. Esposito, "The ESO MCAO demonstrator MAD: a European collaboration," in *Beyond Conventional Adaptive Optics*, S. Esposito and R. Ragazzoni, eds., Vol. 58 of ESO Conference and Workshop Proceedings (European Southern Observatory, Garching, Germany, 2001), pp. 27–35.
46. E. Vernet-Viard, R. Ragazzoni, C. Arcidiacono, A. Baruffolo, E. Diolaiti, J. Farinato, E. Fedrigo, E. Marchetti, R. Fialomo, S. Esposito, M. Carbillet, and C. Vérinaud, "Layer oriented wavefront sensor for mad: status and progress," in *Adaptive Optical System Technologies II*, D. Bonaccini and P. L. Wizinowich, eds., Proc. SPIE **4839**, 344–353 (2003).
47. F. Rigaut, M. Chun, B. Ellerbroek, and C. D'Orgeville, "MCAO for Gemini South," in *Conceptual Design Review Documents*, MCAO Conceptual Design Documentation, 2000; <http://www.gemini.edu/sciops/instruments/adaptiveOptics/AOIndex.html>.
48. F. Hammer, F. Sayède, E. Gendron, T. Fusco, D. Burgarella, V. Cayatte, J.-M. Conan, F. Courbin, H. Flores, I. Guinouard, Jocou, A. Lançon, G. Monnet, M. Mouhcine, F. Rigaud, D. Rouan, G. Rousset, V. Buat, and F. Zamkotsian, "The FALCON concept: multi-object spectroscopy combined with MCAO in near-IR," in *Scientific Drivers for ESO Future VLT/VLTI*, Vol. 42 of ESO Conference and Workshop Proceedings (European Southern Observatory, Garching, Germany, 2002), p. 139.
49. F. Assémat, F. Hammer, E. Gendron, F. Sayède, P. Laporte, M. Marteaude, M. Puech, J.-M. Conan, T. Fusco, and F. Zamkotsian, "FALCON: a new-generation spectrograph with adaptive optics for the ESO VLT," in *Optics in Atmospheric Propagation and Adaptive Systems VI*, J. D. Gonglewski and K. Stein, eds., Proc. SPIE **5237**, 211–222 (2004).
50. B. A. Macintosh, S. Olivier, B. Bauman, J. Brase, E. Carr, C. Carrano, D. T. Gavel, C. E. Max, and J. Patience, "Practical high-order adaptive optics systems for extrasolar planet searches," in *Adaptive Optical Systems and Technologies II*, M. C. Roggeman, R. K. Tyson, and D. Bonaccini, eds., Proc. SPIE **4494**, 60–68 (2002).
51. D. Mouillet, A.-M. Lagrange, J.-L. Beuzit, F. Ménard, C. Moutou, T. Fusco, L. Abé, T. Gillot, R. Soummer, and P. Riaud, "VLT planet finder: specifications for a ground-based high contrast imager," in *Scientific Highlights 2002*, F. Combes and D. Barret, eds. (EDP Sciences, Société Française d'Astronomie et d'Astrophysique, Les Ulis, France, 2002).
52. D. Mouillet, T. Fusco, A.-M. Lagrange, and J.-L. Beuzit, "'Planet finder' on the VLT: context, goals and critical specifications, for adaptive optics," in *Astronomy with High Contrast Imaging*, C. Aime and R. Soummer, eds., Vol. 8 of EAS Publications Series (European Astronomical Society, Petit-Lancy 2, Switzerland, 2002), pp. 193–199.