

# A comparison of Kalman filter based algorithms for turbulent phase control in an adaptive optics system

Alessandro Beghi, Angelo Cenedese, Fabio Maran, and Andrea Masiero

**Abstract**— Telescope resolution is theoretically limited by the diffraction effect, and hence it is inversely proportional to the lens diameter. However, the real resolution of images acquired by large ground telescopes is reduced by the atmospheric turbulence effect. For this reason, telescopes are provided with an adaptive optics (AO) system which aims at compensating the turbulence effect. In this paper we consider a control algorithm for the AO system based on a turbulence prediction method. We propose two linear models, both based on a Principal Component Analysis (PCA) spatial representation, to fit the turbulence temporal dynamic and provide its temporal prediction. We assume that some information about the turbulence has already been estimated, and we exploit it in the computation of the model parameters. The first proposed model yields the best performance but at a quite high computational cost, whereas the second model is best suited in the case of high sampling rates. Furthermore, our simulations show that the PCA spatial representation is robust to errors in the parameter estimation.

## I. INTRODUCTION

The real resolution of large ground telescopes is practically limited by the atmospheric turbulence effect. Indeed, the atmospheric refraction index changes quite fast both spatially and temporally, mostly due to temperature oscillations and to the presence of wind. As a consequence, different beams coming from a star are affected by different phase delays, since they follow different optical paths. Hence the wavefront surface arriving on the telescope is far from being flat. Hereafter, we call turbulent phase the set of phase delays of the beams arriving on the telescope pupil.

Adaptive Optics (AO) are used in large telescopes to overcome the resolution limitation caused by atmospheric turbulence, by computing a proper control input to make a set of deformable mirrors adapt their shape so as to compensate for the current value of the turbulent phase.

An AO system is typically composed by one (or more) wavefront sensor, a control unit, and a set of deformable mirrors. In this paper we consider the case of a single conjugated adaptive optics (SCAO) system, i.e.: The system is provided only of a single wavefront sensor and the deformable mirrors are all placed on a single layer.

AO systems work in real time at high sampling rates, e.g. typically 100 Hz or higher. This obviously makes the control input computation quite challenging. Actually, in order to

make the computation feasible, dimensional reduction step is performed, i.e. the turbulent phase is projected on a set of spatial bases. Furthermore, it is commonly assumed that the control computation requires approximatively two sampling periods [8][7][9], hence we consider a 2-step delay in the AO system feedback loop. For this reason a turbulent phase prediction step is often included in the AO control algorithm.

In this paper we assume to be in Very Large Telescope (VLT) operating conditions, and we consider an AO control system based on a temporal model of the turbulence, which, through the use of a Kalman filter, provides predictions of the turbulent phase temporal evolution. In particular, we propose and compare two temporal models of the turbulence based on a Principal Component Analysis (PCA) representation (which has been proved to be effective for this purpose in [2]). While the first model typically fits better the turbulent phase dynamic, the main advantage of the second model is its low computational complexity, which makes it particularly interesting especially for high frequency applications and for extensions to the Extremely Large Telescopes (ELT) case.

In the computation of the temporal models we assume that some information about the turbulence spatial statistics and its structure is available. The latter can be estimated, by using for instance the algorithm recently proposed in [3]. Actually, the models proposed here can be thought as extensions of those considered in [1] to the case when the turbulence structure is known to the algorithm estimating the turbulent phase temporal model.

The paper is organized as follows: First, Sections II introduce the turbulence statistical model and the adaptive optics principles. Then, in Section III we present the main results, focusing on the control computation and on the temporal models of the turbulent phase. We conclude in Section IV with discussing some simulations.

## II. ADAPTIVE OPTICS SYSTEM

The aim of the AO system is that of compensating the turbulence effect on the telescope measurements by properly commanding a set of deformable mirrors.

Since the deformable mirrors modify the signal on the telescope aperture, they can be viewed as performing a feedback control action. A scheme of the AO system working procedure is reported in Fig. 1, where  $e_m(t)$  are the measurements of the current turbulent phase from a wavefront sensor device,  $y(t)$  is the reconstructed current value of the turbulent phase,  $u(t)$  is the computed mirror control input,  $\varphi_d(t)$  are the correction phases obtained through the deformable mirrors.

This work forms part of the ELT Design Study and is supported by the European Commission, within F. P. 6, contract No 011863.

A.Beghi, F.Maran and A.Masiero are with the Dipartimento di Ingegneria dell'Informazione, Università di Padova, via Gradenigo 6/B, 35131 Padova, Italy {beghi, maranfab, masiero}@dei.unipd.it

A.Cenedese is with the Dipartimento di Tecnica e Gestione dei Sistemi Industriali, Università di Padova, Stradella San Nicola 3, 36100 Vicenza, Italy angelo.cenedese@unipd.it

According with [7] [8], in Fig. 1 we have introduced a 2-step delay before the control unit. Let  $\varphi(t)$  be the value of the turbulent phase at time  $t$ . To improve the AO system performances,  $u(t)$  is commonly computed exploiting the prediction of the current value of the turbulence  $\varphi(t)$ , given measurements up to time  $t - 2$  only. Three models for turbulence prediction are considered in Section III.

We define  $e(t)$  as the input signal of the wavefront sensor, i.e.  $e(t)$  is the difference between  $\varphi(t)$  and the correction phases  $\varphi_d(t)$  provided by the deformable mirrors.

In the following Subsections we give a description of the turbulence characteristics and of the AO system.

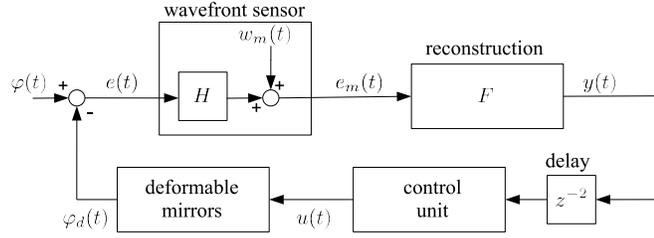


Fig. 1. Scheme of the adaptive optics system.

#### A. Turbulence physical model

Let  $\mathbf{u}$  and  $\mathbf{v}$  be two unit vectors indicating two orthogonal spatial directions, as in Fig. 2, and let  $\phi(u, v, t)$  be the value of the turbulent phase on the point  $(u, v)$  at time  $t$  on the telescope aperture plane, where  $u$  and  $v$  are the coordinates of the point along  $\mathbf{u}$  and  $\mathbf{v}$ . Without loss of generality, we assume that the origin of the coordinate system induced by  $\mathbf{u}$  and  $\mathbf{v}$  be in correspondence with the center of the telescope. The turbulent phase is assumed to be zero-mean stationary and spatially homogeneous, hence the covariance between two values of the turbulence,  $\phi(u, v, t)$  and  $\phi(u', v', t)$ , depends only on the distance,  $r$ , between the two points:  $C_\phi(r) = \mathbf{E}[\phi(u, v, t)\phi(u', v', t)], \forall(u, v, u', v')$ , such that  $r = \sqrt{(u - u')^2 + (v - v')^2}$ .

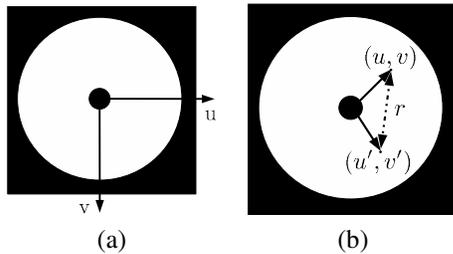


Fig. 2. (a) Coordinates on the telescope image domain. (b) Two points,  $(u, v)$  and  $(u', v')$ , separated by a distance  $r$  on the telescope aperture plane.

According to the Von Karman theory, the spatial covariance  $C_\phi(r)$  is the following (see [5]):

$$C_\phi(r) = \left(\frac{L_0}{r_0}\right)^{5/3} \frac{c}{2} \left(\frac{2\pi r}{L_0}\right)^{5/6} K_{5/6}\left(\frac{2\pi r}{L_0}\right), \quad (1)$$

where  $K(\cdot)$  is the MacDonald function (modified Bessel function of the third type),  $\Gamma$  is the Gamma function,  $L_0$

is the outer scale,  $r_0$  is a characteristic parameter called the Fried parameter (see [10]), and the constant  $c$  is:  $c = \frac{2^{1/6}\Gamma(11/6)}{\pi^{8/3}} \left[\frac{24}{5}\Gamma(6/5)\right]^{5/6}$ .

In order to describe its temporal characteristics, the turbulence is generally modeled as the superposition of a finite number  $l$  of layers. The  $i^{\text{th}}$  layer models the atmosphere from an altitude of  $h_{i-1}$  to  $h_i$  meters, where  $h_l \geq \dots \geq h_1 \geq h_{i-1} \geq \dots \geq h_0 = 0$ . Let  $\psi_i(u, v, t)$  be the value of the  $i^{\text{th}}$  layer at point  $(u, v)$  at time  $t$ . Then the total turbulent phase at  $(u, v)$  and at time  $t$  along the Zenith direction is  $\phi(u, v, t) = \sum_{i=1}^l \gamma_i \psi_i(u, v, t)$ , where  $\gamma_i$  are suitable coefficients associated to the layer energies. Without loss of generality we assume that  $\sum_{i=1}^l \gamma_i^2 = 1$ .

The layers are assumed to be stationary and characterized by similar spatial statistics, i.e. the covariance between two points at distance  $r$  of the  $i$ -th turbulence layer can be written as follows:

$$C_{\psi_i}(r) = \gamma_i^2 \left(\frac{L_0}{r_0}\right)^{5/3} \frac{c}{2} \left(\frac{2\pi r}{L_0}\right)^{5/6} K_{5/6}\left(\frac{2\pi r}{L_0}\right). \quad (2)$$

Furthermore, the layers are assumed to be independent, hence:  $\mathbf{E}[\psi_i(u, v, t)\psi_j(u', v', t')] = 0$ ,  $i \neq j$ .

A commonly agreed assumption considers that each layer translates in front of the telescope pupil with constant velocity  $v_i$  (Taylor approximation [10]), thus

$$\psi_i(u, v, t + kT) = \psi_i(u - v_{i,u}kT, v - v_{i,v}kT, t), \quad (3)$$

$i = 1, \dots, l$ , where  $v_i = v_{i,u}\mathbf{u} + v_{i,v}\mathbf{v}$ , and  $kT$  is a delay multiple of the sampling period  $T$ .

#### B. Wavefront sensor and reconstruction procedure

In this Subsection we introduce a statistical model for the measurement and reconstruction procedure. In real applications only a finite number of sensors is available. These are usually distributed on a grid, thus the turbulent phase is considered only on a discrete domain  $\mathbb{L}$  (Fig. 3(a)).

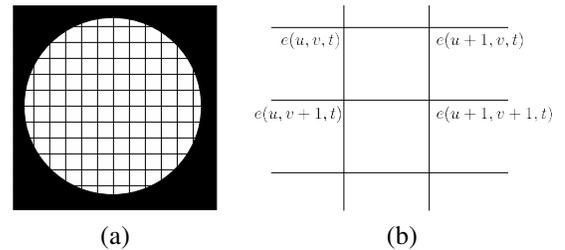


Fig. 3. (a) Discrete domain  $\mathbb{L}$ . (b) Points used by the Shack-Hartmann sensor, according to the Fried geometry, to take the phase slope measurements on the point  $(u, v)$ .

Here we assume that the AO is provided with only one wavefront sensor, which takes measurements on the Zenith direction. Moreover we assume that the wavefront sensor is a Shack-Hartmann (SH) device. The SH sensor measures at each point the phase slope in both the spatial directions  $\mathbf{u}$  and  $\mathbf{v}$ . Let  $d$  be the telescope diameter and  $n_s$  the number of nodes per each side of the grid, then, assuming the nodes uniformly spaced, the distance between two neighbors on the

grid is  $p_s = d/n_s$  meter. By numbering the node coordinates on each axis from 1 to  $n_s$ , we can write

$$\mathbb{L} = \{(u, v) \mid u \in \mathbb{Z}, v \in \mathbb{Z}, 1 \leq u \leq n_s, 1 \leq v \leq n_s, \left\| \begin{array}{c} u - (n_s + 1)/2 \\ v - (n_s + 1)/2 \end{array} \right\|_2 < d/2\} .$$

According with the Fried geometry for the SH device [6], the sensor measurements at point  $(u, v) \in \mathbb{L}$  (with  $1 \leq u \leq n_s - 1$  and  $1 \leq v \leq n_s - 1$ ) can be written as

$$e_{m,\mathbf{u}}(u, v, t) = \frac{1}{2}(e(u+1, v, t) - e(u, v, t)) + \frac{1}{2}(e(u+1, v+1, t) - e(u, v+1, t)) , \quad (4)$$

$$e_{m,\mathbf{v}}(u, v, t) = \frac{1}{2}(e(u, v+1, t) - e(u, v, t)) + \frac{1}{2}(e(u+1, v+1, t) - e(u+1, v, t)) , \quad (5)$$

where  $e_{m,\mathbf{u}}(u, v, t)$  and  $e_{m,\mathbf{v}}(u, v, t)$  corresponds to estimates of the slopes along the  $\mathbf{u}$  and  $\mathbf{v}$  direction in the considered point. Moreover,  $e(u, v, t)$  is the value of the residual phase (i.e. the remaining turbulent phase after the subtraction of  $\varphi_d(t)$ ) on the point  $(u, v) \in \mathbb{L}$  at time  $t$ .

With an abuse of notation, let us denote with  $\varphi(t)$  and  $e(t)$  the vectors containing the values respectively of the turbulence, and of the residual phases after the deformable mirror correction, on the grid  $\mathbb{L}$  at time  $t$ . Notice that, from an optical point of view, the phase translations over the entire telescope aperture can be neglected, hence the AO system does not take into account the projection of turbulent phase on the vector  $u_0 = [1 \ 1 \ \dots \ 1]^T$ . Thus, we consider  $\varphi(t)$  as the vector containing the values of the turbulent phase at time  $t$  neglecting its projection on  $u_0$ .

Furthermore, let us denote with  $e_m(t)$  the vector containing all the SH measurements at time  $t$ . Since (4) and (5) are linear, then  $e_m(t)$  can be written as follows:

$$e_m(t) = He(t) + w_m(t) , \quad (6)$$

where  $H$  is a matrix properly defined to express the relation between  $e(t)$  and  $e_m(t)$  as in (4) and (5).  $w_m$  is a zero-mean white noise, due to the measurement process. Let us call  $Q_m$  the covariance of  $w_m$ . For simplicity of notation, in the following we will assume that  $Q_m = \sigma_m^2 I$ , where  $\sigma_m^2$  is the measurement noise variance.

We define the Signal to Noise Ratio (SNR) as follows:  $SNR = \text{trace}(H\Sigma_\varphi H^T)/\text{trace}(Q_w)$ , where  $\Sigma_\varphi = \mathbf{E}[\varphi(t)\varphi(t)^T]$ , and it can be computed from (1). This definition of SNR corresponds to the signal to noise ratio when the system works in open-loop, i.e. when  $e(t) = \varphi(t)$  in (6).

Let now  $m$  be the number of actuators,  $|\mathbb{L}|$  the number of nodes in  $\mathbb{L}$  and  $p$  the number of measurements, where commonly  $p \approx 2|\mathbb{L}|$ . Since  $|\mathbb{L}|$  can be quite large (in our simulations  $|\mathbb{L}| \approx 10^3$ ), then to reduce the computational load, the turbulent phase is projected on a set of spatial bases. Astronomers commonly choose the set of Zernike polynomials as spatial bases, however, to exploit the knowledge about the second order statistical properties of the signal, in

Section III we consider also models based on the set of bases provided by PCA. We refer the reader to [2] for a comparison between these representations from a static point of view. The dimensionality reduction step is obtained by considering the projections on the first  $n$  spatial bases (neglecting  $u_0$ ) of the considered basis  $U = [u_1 \ \dots \ u_n]$ , which are assumed to be those containing most of the signal energy. Hence

$$\varphi(t) = Ux(t) + e_r(t) , \quad (7)$$

where  $e_r(t) = \varphi(t) - Ux(t)$  is the representation error. Let  $U^\dagger$  be the pseudo-inverse of  $U$ , then  $x(t) = U^\dagger\varphi(t)$ . In the following we call  $x(t)$  the state vector at time  $t$  and we model the dynamic behavior of  $x(t)$  instead of that of  $\varphi(t)$ : Since usually  $n \ll |\mathbb{L}|$ , this remarkably reduces the time to compute turbulence predictions and control actions.

We assume that the reconstruction procedure is performed by premultiplying a proper matrix  $F$  to the measurement vector  $e_m(t)$  (thus this is a vector-matrix-multiply (VMM) reconstructor). To reconstruct the residual phase  $e(t)$ , we should multiply the measurement vector  $e_m(t)$  by  $H^\dagger$ . However, since we are interested in the projection of the phases on the basis  $U$ , the overall reconstructor matrix  $F$  is:  $F = (HU)^\dagger$ . Then, the measurement of the reconstructed phases projected on  $U$  is  $y(t) = Fe_m(t)$ . Let  $\varphi_d(t) = Ux_d(t) + e_d(t)$ , where  $x_d(t) = U^\dagger\varphi_d(t)$ , then

$$y(t) = Fe_m(t) = FH(\varphi(t) - \varphi_d(t)) + Fw_m(t) = FHU(x(t) - x_d(t)) + FH(e_r(t) - e_d(t)) + Fw_m(t) .$$

Since  $(e_r(t) - e_d(t))$  are small and orthogonal to the space generated by the columns of  $U$ , then  $FH(e_r(t) - e_d(t)) \approx 0$ . Furthermore  $FHU \approx I$  (the equality stands when  $(HU)$  is full column-rank). Then,

$$y(t) = (x(t) - x_d(t)) + w(t) , \quad (8)$$

where  $w(t)$  is zero-mean and its covariance is  $R \approx FF^T\sigma_m^2$ . If not differently specified, hereafter we assume

$$R = FF^T\sigma_m^2 . \quad (9)$$

### C. Control unit and deformable mirrors

The aim of the control unit is that of computing a proper input for the deformable mirrors. Computation of the proper correction is a challenging task, mainly because the system usually works at high frequencies. Therefore, the control action has to be sufficiently simple to allow for fast computation, although the presence of the 2-step delay in the system and of the nonlinearities in the behavior of the deformable mirrors has to be carefully taken into account. Further details on the control computation are given in Section III.

It is a common assumption ([10][8]) to take the deformable mirrors transfer function  $\mathcal{D}(u(t))$  as linear and static, i.e.  $\varphi_d(t) = \mathcal{D}(u(t)) = Du(t)$ . In literature, deformable mirrors are usually characterized by the so called *interaction matrix*  $\bar{D}$ , which, using our notation, can be written as  $\bar{D} = HD$ . Since in the control loop the phases are projected on  $U$  and both  $D$  and  $\bar{D}$  are constant, the two approaches are equivalent. In our simulations we assume that  $D = U$ .

#### D. Performances evaluation: Strehl ratio

In this Subsection we first introduce the concept of coherent energy and then we show how to use it to compute the Strehl ratio (SR), which is commonly used by astronomers as a criterion to evaluate the real telescope resolution.

Let us define the coherent energy at time  $t$ ,  $CE(t)$ , as the sample variance of the residual phase, i.e.  $CE = \frac{1}{|\mathbb{L}|} \sum_{(u,v) \in \mathbb{L}} \left( e(u,v,t) - \sum_{(u,v) \in \mathbb{L}} \frac{e(u,v,t)}{|\mathbb{L}|} \right)^2$ , where  $e(u,v,t) = \phi(u,v,t) - \phi_d(u,v,t)$ , and  $\phi_d(u,v,t)$  is the phase correction provided by the deformable mirrors at time  $t$  on the point  $(u,v)$ . Then, the Strehl ratio at time  $t$ , SR, is well approximated by the following

$$SR(t) \approx \exp(-CE(t)) \quad (10)$$

when  $SR(t) > 0.2$ . In our simulations we use the SR as a criterion for AO performances evaluation (a large value of the SR indicates good image quality and viceversa), and we always approximate it as in (10).

### III. CONTROL AND TURBULENCE MODELS

In this Section we describe a method to compute the control for the deformable mirrors based on a turbulence prediction model. The approach is similar to those described in [8],[7],[9]. However, in Subsections III-B, III-C we propose two innovative turbulence temporal models, which, in the simulations of Section IV, improve the results obtained using the model described in [9]. The Petit model, [9], is also summarized in Subsection III-A. Models of Subsections III-B and III-C exploit a PCA representation, which has already been proved in [2] to be a well suited spatial representation for the turbulent phase. Furthermore, they can be considered as extensions of the models presented in [1] whenever the turbulence layer characteristics (the number of layers, their velocities and their strengths) can be assumed to be known. This information can be provided by external sensors or estimated from turbulent phase data, e.g. as described in [3].

First, assume to know the current value of the atmosphere  $\varphi(t)$ . Then, the ideal control input  $u(t)$ , using the state representation, can be written as  $u(t) = D^\dagger U x(t)$ , where  $x(t) = U^\dagger \varphi(t)$ . However, since  $x(t)$  is unknown, we cannot directly apply the above equation to compute the control input  $u(t)$ , but, we can still use it after substituting  $x(t)$  with its estimate  $\hat{x}(t|t-2)$  obtained by using the available measurements up to time  $t-2$ :

$$u(t) = D^\dagger U \hat{x}(t|t-2) . \quad (11)$$

In Subsections III-A, III-C, III-B we describe three linear dynamic models for turbulence representation and prediction: These models are associated to the same strategy for the control of the deformable mirror, i.e. they use a Kalman filter on a properly defined linear dynamic model to estimate  $\hat{x}(t|t-2)$  and then apply the control law (11). Hence the AO system performance largely depends on the ability of the proposed models to describe and predict the temporal evolution of the turbulent phase. In particular we assume

that the turbulence temporal models can be described by the following linear dynamic system

$$\begin{cases} \bar{x}(t+1) = \bar{A}\bar{x}(t) + \begin{bmatrix} 0 \\ u(t+2) \\ 0 \end{bmatrix} + \begin{bmatrix} v(t) \\ 0 \\ 0 \end{bmatrix} \\ y(t) = [ I \ 0 \ -U^\dagger D ] \bar{x}(t) + w(t) \end{cases} \quad (12)$$

where  $\bar{x}(t)$  has the following block structure

$$\bar{x}(t) = \begin{bmatrix} x(t) \\ u(t+1) \\ u(t) \end{bmatrix}, \quad \bar{A} = \begin{bmatrix} A & 0 & 0 \\ 0 & 0 & 0 \\ 0 & I & 0 \end{bmatrix},$$

and  $A$  is a suitable  $n \times n$  matrix. Furthermore,  $y(t)$  is the measurement vector (as defined in (8)),  $\bar{x}(t)$  is the overall state of the dynamic system,  $v(t)$  and  $w(t)$  are assumed to be zero-mean white noise with covariances  $Q$  and  $R$ .

Then, (11) becomes  $u(t) = D^\dagger U A \hat{x}(t-1|t-2)$ , where  $\hat{x}(t-1|t-2)$  is the estimate of  $x(t-1)$  provided by the Kalman filter given the measurements up to time  $t-2$ .

Notice that, to reduce the computational complexity of the algorithm, the time-invariant, asymptotic Kalman filter is used, i.e. the Kalman gain  $\bar{K}$  (computed solving the Algebraic Riccati Equation (ARE)) is constant. Moreover, it is simple to prove that  $\bar{K} = [ K^T \ 0 \ 0 ]^T$ , where

$$K = AP(P+R)^{-1}, \quad (13)$$

and  $P$  is the solution of the following ARE:  $P = A(P - P(P+R)^{-1}PA^T + Q)$ . Since the last two blocks of the state vector can be trivially updated, the only non elementary computation in the Kalman predictor is

$$\hat{x}(t+1|t) = A\hat{x}(t|t-1) + K(y(t) - \hat{x}(t|t-1) + U^\dagger D u(t)) . \quad (14)$$

For this reason, we still call  $x(t)$  the state vector.

Since the models are based on the same control strategy, their performances depends on the particular choices of the parameters  $\{A, U, Q, R\}$  in the dynamical system (12).

#### A. Petit model

The model described by Petit in [9] (a similar model was introduced in [7], [8]) is based on the use of Zernike polynomials as spatial basis  $U$  (see [10], [2]). The turbulence is assumed to be characterized by Kolmogorov statistics (see [10], [7]). The Von Karman model of the turbulence can be reduced to the Kolmogorov one by making  $L_0$  go to infinity.

The matrix  $A$  is assumed to be diagonal, and the  $i$ -th element of its diagonal is associated to the  $i$ -th Zernike polynomial. A radial order is associated to each Zernike polynomial, in particular we denote with  $n(i)$  the radial order associated to the  $i$ -th polynomial. Then, assuming  $l = 1$ , the  $i$ -th element of the diagonal,  $a_{ii}$ , can be computed as follows:  $a_{ii} = \exp(-0.3(n(i)+1)|v_1|T/d)$ .

$Q$  is computed as follows:  $Q = \Sigma_x^{kol} - A \Sigma_x^{kol} A^T$ , where  $\Sigma_x^{kol}$  is the state covariance according to the Kolmogorov statistics [11], [12].  $R$  is the covariance matrix associated to the measurement noise  $w$  and it is computed as follows: First project  $w$  on the first  $n$  Zernike polynomials, then

compute the covariance matrix of the projected signal [7], [9]. Finally,  $K$  is computed from (13). The model can be easily generalized to the multi-layer case [9], [7], [8].

### B. PCA based full-matrices model

In this model we propose a PCA based spatial representation, where  $U$  is composed by the first  $n$  bases provided by the PCA [2]. Define the state covariance as  $\Sigma_x = U^\dagger \Sigma_\varphi (U^\dagger)^T$ , then the matrix  $A$  is computed as follows

$$A = \sum_{i=1}^l \gamma_i^2 U^\dagger \mathbf{E} [\psi_i(t+1) \psi_i(t)^T] (U^\dagger)^T \Sigma_x^{-1}, \quad (15)$$

where  $\mathbf{E}[\psi_i(t+1) \psi_i(t)^T]$ ,  $i = 1, \dots, l$  are computed according to the theoretical spatial covariances (2) and to the Taylor assumption (3). Then  $Q$  is  $Q = \Sigma_x - A \Sigma_x A^T$ , and  $R$  is defined as in (9).  $K$  is computed from (13).

### C. PCA based diagonal-matrices model

The computational load is a stringent design parameter for the choice of the control law, therefore we propose a model particularly convenient from the computational point of view.

Since the only nontrivial computation in the asymptotic Kalman filter is that of (14), we choose both  $A$  and  $K$  to be diagonal, to dramatically reduce the computational complexity of the algorithm.

Let  $U$  be the set of the first  $n$  bases provided by PCA. Then  $A$  is computed taking only the principal diagonal from the  $A$  computed as in (15). Moreover  $Q = \Sigma_x - A \Sigma_x A^T$  and  $R$  is defined as in (9).

Finally, we compute  $K$  from (13), but we set all its elements out of the principal diagonal to 0.

Notice that with this choice of the parameters some numerical problems can occur while solving the ARE to compute  $P$  and  $K$ . Even when this does not happen, we have experimentally observed that the diagonal elements in the computed  $K$  are sometimes too small to make the prediction effective. This is due to the fact that the Kalman gain is computed for system (12) assuming  $A$  diagonal, however, since this is not the real case, the computed  $K$  typically takes to poor performance.

To reduce both of these problems, we propose to artificially enhance the system dynamics by rescaling the matrix  $A$  by a constant factor  $(1-\epsilon)$ , i.e.  $A = (1-\epsilon)A$ , where in our simulations we set  $\epsilon = 10^{-4}$ . As a consequence, although the model is only marginally modified, a larger Kalman gain is obtained.

## IV. SIMULATIONS AND DISCUSSION

In this Section we assume to be in VLT-like conditions, e.g.  $d = 8$  meter and  $n_s = 40$ , and we compare the models of Subsections III-A, III-B, III-C in two simulations. In both the simulations we set the field of view to 58 arcsec and we investigate the case of a quite high sampling rate:  $f_s = 1/T = 1$  KHz.

The turbulence has been simulated, at higher resolution than that of the grid  $\mathbb{L}$ , using the method described in [4], and the number of temporal samples used to estimate the SR has been set to 5000 in all our examples.

### A. Simulation 1: Singular layer

Since most of the turbulent phase energy is concentrated on the ground layer, often the AO system models this layer only. Hence in this example the atmosphere has been considered formed by a unique layer translating over the telescope pupil at 10m/s. Moreover, since the atmosphere parameters are usually known only up to some approximations, we also try to evaluate the robustness of the turbulence models of Subsections III-B and III-C to errors in turbulence parameter estimation. We set the values of the parameters for turbulence simulation to:  $r_0 = 0.4$ m,  $L_0 = 22$ m; and  $\sigma_m^2 = 0.4$ rad<sup>2</sup>.

In Fig. 4 we show the Strehl ratios, obtained using the three models of Section III, where the values of the parameters used to compute models of Subsections III-B and III-C are: (a) the real values, i.e.  $\hat{r}_0 = 0.4$ m,  $\hat{L}_0 = 22$ m,  $\hat{\sigma}_m^2 = 0.4$ rad<sup>2</sup>; (b)  $\hat{r}_0 = 0.2$ m,  $\hat{L}_0 = 100$ m,  $\hat{\sigma}_m^2 = 0.2$ rad<sup>2</sup>.

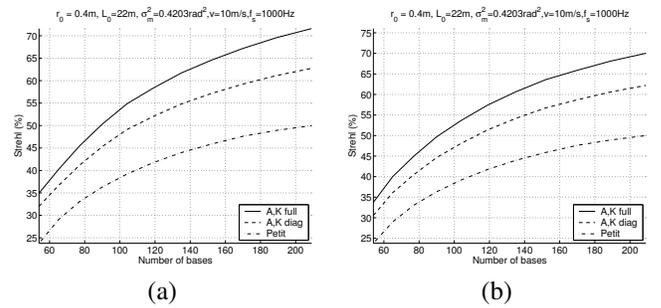


Fig. 4. Strehl ratios obtained using different turbulence models: PCA and  $(A, K)$  full (solid line), PCA and  $(A, K)$  diagonal (dashed line), Petit model (dashed-dotted line). Different values for the parameters used to compute the models of Subsections III-B and III-C are explored: (a)  $\hat{r}_0 = 0.4$ m,  $\hat{L}_0 = 22$ m,  $\hat{\sigma}_m^2 = 0.4$ ; (b)  $\hat{r}_0 = 0.2$ m,  $\hat{L}_0 = 100$ m,  $\hat{\sigma}_m^2 = 0.2$ . Sampling ratio,  $f_s$ , is set to 1KHz.

### B. Simulation 2: A case study

Even if often most of the turbulence energy is concentrated on one layer, the atmosphere is actually formed by several layers. However, in practice only a finite, and usually small, number of layers can be modeled in the adaptive optics system: These usually are those which are assumed to have the largest energies. In order to simulate these operating conditions, in this simulation we consider the atmosphere as formed by three layers, where only the first one is modeled in the models of Section III. Then, the real turbulence parameters are:  $r_0 = 0.2$ m,  $L_0 = 22$ m,  $v_1 = -10\mathbf{u}$  m/s,  $\gamma_1^2 = 0.80$ ,  $h_1 = 0.5$ Km,  $|v_2| = 15$ m/s,  $\gamma_2^2 = 0.15$ ,  $h_2 = 6$ Km,  $v_3 = 30\mathbf{v}$  m/s,  $\gamma_3^2 = 0.05$ ,  $h_3 = 8$ Km. The real measurement noise variance is  $\sigma_m^2 = 0.6$ rad<sup>2</sup> and the sampling rate is  $f_s = 1$ KHz. To compute the models we assume to have (rough) estimations of  $r_0$ ,  $L_0$  and  $\sigma_m^2$ , that is:  $\hat{r}_0 = 0.4$ m,  $\hat{L}_0 = 100$ m,  $\hat{\sigma}_m^2 = 0.2$ rad<sup>2</sup>.

Notice that in this simulation the energy of the unmodeled part of the turbulence is quite large, i.e. it corresponds to the 20 percent of the total signal energy. In this particular condition the model of Subsection III-B is not sufficiently flexible to grant satisfactory performance, i.e. it fits too much the ground layer dynamic, without taking into account the

unmodeled part of the turbulence. To make it more flexible we proceed similarly to Subsection III-C: We scale by  $(1-\epsilon)$  the  $A$  matrix computed as in (15), where we set  $\epsilon = 10^{-4}$ .

In Table I we report the Strehl ratio obtained at Zenith direction, in the above conditions, for the following 4 models: 1) the Petit model described in Subsection III-A, 2) the model of Subsection III-B adapted to the Zernike basis case, with  $(1-\epsilon)$  premultiplied to  $A$ , 3) the model of Subsection III-C, 4) the model of Subsection III-B, with  $(1-\epsilon)$  premultiplied to  $A$ . Where for each model we have used  $n = 405$  spatial bases to represent the turbulence. We also distinguish two possible cases for the direction of the velocity  $v_2$ : Case (A)  $v_2 = -|v_2|\mathbf{u}$ , case (B)  $v_2 = |v_2|\mathbf{v}$ .

TABLE I  
STREHL RATIOS MODELING ONE OF THREE LAYERS

	Petit	Zer; $A, K$ full	PCA; $A, K$ diag	PCA; $A, K$ full
case (A)	18.24	39.92	36.84	45.47
case (B)	18.22	34.92	36.85	39.72

### C. Discussion

In Subsection IV-A we tested, in the ideal case of a single layer turbulence, the performance of the algorithms described in Section III. The results show that, although the use of rough estimates of the turbulence parameters  $r_0$ ,  $L_0$ , influences the computation of the PCA bases, the use of such representation in the prediction algorithms makes them robust to errors in the estimates of  $\hat{r}_0$ ,  $\hat{L}_0$ .

Simulation 2 (Subsection IV-B) shows that the full-matrices model, since it relies on *a priori* information on the turbulent layers characteristics (which can be estimated as described in [3]), is sometimes sensitive to errors in the parameter estimation. In order to make it effective when some unmodeled dynamics are present, we introduced the  $(1-\epsilon)$  scaling of the  $A$  matrix. With this modification the performance is greatly increased, even when there are two unmodeled layers containing the 20 percent of the signal energy moving in an orthogonal direction with respect to that used to compute the model (case (B)).

When the sampling ratio is high, the computational load required by the full-matrices model may be excessive. However, in this case the prediction step is quite simple since the turbulence introduces less variations between two time steps. Hence even simpler models can give good results. As shown in Table I the model of Subsection III-C, exploiting the PCA representation, outperforms the Petit model in both the cases (A) and (B). Moreover, it gives results similar to those of the full-matrices model in case (B), furthermore, since it makes no assumption on the moving direction of the layers, it is not sensitive to errors in the estimation of this parameter.

To conclude, our results suggest that the use of a full-matrices system (Subsection III-B) to model the temporal evolution of the turbulence gives better results than the other considered models. However, when the sampling frequency is high (or the information about the turbulence layers characteristics is not reliable) it is worth to consider simpler

models: In particular the model described in Subsection III-C results to be effective and computationally cheap.

## V. CONCLUSIONS

In this paper we have considered an algorithm for computing the control in an AO system based on turbulence temporal prediction. We have proposed two turbulent phase temporal models and we have compared their closed loop performances in an AO system.

The considered models, which exploit a PCA spatial representation of the turbulent phase, are computed from some information about the turbulence layers characteristics, that has to be previously estimated. The proposed models have been successfully compared also to the Petit model [9].

Our simulations suggest that the PCA representation outperforms the Zernike one even when the PCA bases are computed starting from very rough estimations of the real turbulence parameters.

The use of a full matrices system may take to better results than the other considered models, however, when the sampling ratio is sufficiently high or the available turbulence layers characteristics are not reliable, it is worth to consider the proposed diagonal model, which is interesting both for its performances and its low computational complexity.

## VI. ACKNOWLEDGMENTS

We are pleased to acknowledge the colleagues of the ELT Project, in particular Dr. Michel Tallon at CRAL-Lyon, and Dr. Enrico Fedrigo at ESO-Munich, for their precious help in supporting us with the astronomical view of the problem.

## REFERENCES

- [1] A. Beghi, A. Cenedese, and A. Masiero. Atmospheric turbulence prediction: a pca approach. In *Proc. of the 46th IEEE C. on Dec. and Contr., CDC 2007*, pages 572–577, New Orleans, USA, Dec. 2007.
- [2] A. Beghi, A. Cenedese, and A. Masiero. A comparison between zernike and pca representation of atmospheric turbulence. In *Proc. of the 46th IEEE Conference on Decision and Control, CDC 2007*, pages 561–566, New Orleans, USA, December 2007.
- [3] A. Beghi, A. Cenedese, and A. Masiero. On the estimation of atmospheric turbulence layers. In *Proc. of the 2008 IFAC World Congress*, pages 8984–8989, Seoul, Korea, July 2008.
- [4] A. Beghi, A. Cenedese, and A. Masiero. Stochastic realization approach to the efficient simulation of phase screens. *Journal of the Optical Society of America A*, 25(2):515–525, February 2008.
- [5] R. Conan. *Modelisation des effets de l'echelle externe de coherence spatiale du front d'onde pour l'observation a haute resolution angulaire en astronomie*. PhD thesis, Univ. Nice Sophia Antipolis, 2000.
- [6] D.L. Fried. Least-square fitting a wave-front distortion estimate to an array of phase-difference measurements. *Journal of the Optical Society of America*, 67:370–375, 1977.
- [7] B. Le Roux. *Commande optimale en optique adaptative classique et multiconjuguee*. PhD thesis, Univ. de Nice Sophia Antipolis, 2003.
- [8] B. Le Roux, J.M. Conan, C. Kulcsar, and et al. Optimal control law for classical and multiconjugate adaptive optics. *Journal of the Optical Society of America A*, 21(7):1261–1276, 2004.
- [9] C. Petit. *Étude de la commande optimale en OA et OAMC, validation numérique et expérimentale*. PhD thesis, Université Paris 13, 2006.
- [10] F. Roddier. *Adaptive optics in astronomy*. Cambridge university press, 1999.
- [11] N. Takato and I. Yamaguchi. Spatial correlation of zernike phase-expansion coefficients for atmospheric turbulence with finite outer scale. *J. of the Optical Society of America A*, 12(5):958–963, 1995.
- [12] D.M. Winker. Effect of a finite outer scale on the zernike decomposition of atmospheric optical turbulence. *Journal of the Optical Society of America A*, 8(10):1568–1573, 1991.