Proceedings of the
46th IEEE Conference on Decision and Control
New Orleans, LA, USA, Dec. 12-14, 2007

WeA18.1

Atmospheric turbulence prediction: a PCA approach

Alessandro Beghi, Angelo Cenedese, and Andrea Masiero

Abstract— Current and next generation of telescopes resort to adaptive optics in order to compensate for atmospheric turbulence and correct the incident wavefront, which is affected mainly as far as the phase is concerned, in order to get clearer images of the observed objects. Starting from the statistical description of the atmospheric turbulence, the reconstruction of the turbulent phase is therefore a key problem in astronomical seeing and is central to designing control systems to command the adaptive optic deformable mirror. Moreover, the introduction of a dynamical model able to predict the turbulence helps improving the performance of the control system and its description through a principal component analysis (PCA) approach yields compactness to the representation while preserving the physical insight of a modal decomposition.

I. INTRODUCTION

Current generation of optical telescopes, such as those employed in the Very Large Telescope (VLT) project [1], house lenses of several meters diameter obtained resorting to the segmented mirror technology, that, combined with active and adaptive optics computer based control systems, allows astronomical seeing far beyond that of previous monolithic mirror devices equipped with mechanical control.

In these experiments, the active control system operates with typical one-second timescales on the thin segmented mirror (primary mirror) in order to compensate for environmental factors due for instance to gravity at different telescope inclinations, wind, structure deformations, and so on. Conversely, adaptive systems intervene on an "actively stabilized systems" at higher frequencies (timescales of 1/100th second and less) and are used to correct the distortion introduced by the atmospheric turbulence, by controlling a correction mirror placed along the line of sight or by deforming the telescope secondary mirror.

In this respect, it is of paramount importance, and it will be even more in next generation devices [7] [4], envisaging lenses of thirty to one hundred meters, to develop algorithms that are capable to feed the adaptive control systems with the information on the turbulent phase, which are needed to drive the actuators that modify the shape of the deformable mirrors.

These algorithms, basically referring to the reconstruction and the prediction of the atmospheric turbulence, set a trade-off between the need for accuracy and the need for time

A.Beghi and A.Masiero are with the Dipartimento di Ingegneria dell'Informazione, Università di Padova, via Gradenigo 6/B, 35131 Padova, Italy {beghi,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

efficiency for the real time application, so that the wavefront correction is accurate with respect to the current state of the atmosphere.

To perform the task, the control system calculates the actuator commands from wavefront-sensor measurements, obtained using a reference star or even the observed object itself if it is bright enough and has sufficiently sharp light gradients.

To complete the picture, zonal or modal control methods can be used. In zonal control, each zone or segment of the mirror is controlled independently by wavefront signals that are measured for the sub-aperture corresponding to that zone. In modal control, the wavefront is expressed as the linear combination of modes that best fit the atmospheric perturbations.

The paper is organized as follows: First, Sections II and III introduce the turbulence statistical model and the adaptive optics principles. Then, Sections IV and V present the body of the work, focusing on the reconstruction and the prediction of the turbulent phase. We conclude in Sections VI and VII with discussing some simulations.

II. TURBULENCE PHYSICAL MODEL

A. Turbulence temporal model

The turbulence is generally modeled as the superposition of a finite number l of layers: The ith layer models the atmosphere from h_{i-1} to h_i meter high, where $h_0 = 0$, as in Fig. 1(a). Let r be a point on telescope aperture and $\psi_i(r,t)$ be the value of the ith layer on point r at time t: Thus the total turbulence phase at r is

$$\phi(r,t) = \sum_{i=1}^{l} \gamma_i \psi_i(r,t) \tag{1}$$

where γ_i are suitable coefficients. Without loss of generality we assume that $\sum_{i=1}^{l} \gamma_i^2 = 1$.

Assuming the turbulence to be stationary the spatial characteristics of the turbulence are temporally invariant: the Von Karman model of turbulence spatial characteristics will be described in Section II-B.

The ith layer is assumed to translate in front of the telescope pupil with constant velocity v_i (Taylor approximation [10]), thus

$$\psi_i(r, t + kT) = \psi_i(r - v_i kT, t) \ , \ i = 1, \dots, l \tag{2}$$

where kT is a delay multiple of the sampling period T.

Furthermore the layers are assumed to be independent, hence

$$\mathbf{E}[\psi_i(r_i, t_i)\psi_j(r_j, t_j)] = 0 \ , \ i = 1, \dots, l, \ j \neq i \ .$$

B. Turbulence spatial model

Since the spatial characteristics of the turbulence are assumed to be time invariant, then in this section we consider the time as fixed at a constant value $t = \bar{t}$. Furthermore to simplify the notation we will omit $\bar{t}$ from equations.

Since we omit t from the notation, then $\phi(r_i)$ is the value of the turbulent phase at the point r_i on the aperture plane. Considering the turbulent phase as a spatial process, it is a zero-mean (wide-sense) stationary stochastic process: Thus it is completely characterized by its second order properties. In practice its spatial statistical model is commonly described by means of the structure function, which measures the averaged difference between the phase at two points at locations r_1 and r_2 of the wavefront (see Fig. 1(b)), which are separated by a distance r,

$$D_\phi(r) = \left\langle |\phi(r_1) - \phi(r_2)|^2 \right\rangle.$$

The structure function D_ϕ is related to the covariance function $C_\phi(r) = \langle \phi(r_1), \phi(r_2) \rangle$, as:

$$D_\phi(r) = 2 \left(\sigma_\phi^2 - C_\phi(r) \right), \qquad (3)$$

where σ_ϕ^2 is the phase variance.

According to the Von Karman theory, the phase structure function evaluated at distance r is the following [3]:

$$D_\phi(r) = \left(\frac{L_0}{r_0} \right)^{5/3} c \left[\frac{\Gamma(5/6)}{2^{1/6}} - \left(\frac{2\pi r}{L_0} \right)^{5/6} K_{5/6} \left(\frac{2\pi r}{L_0} \right) \right],$$

where $K_{\cdot}(\cdot)$ is the MacDonald function (modified Bessel function of the third type), Γ is the Gamma function, L_0 is the outer scale, r_0 is a characteristic parameter called the Fried parameter [5], and $c = \frac{2^{1/6}\Gamma(11/6)}{\pi^{8/3}} \left[\frac{24}{5}\Gamma(6/5) \right]^{5/6}$.

From the relation between the structure function and the covariance (3), the spatial covariance of the phase between two points at distance r results

$$C_\phi(r) = \left(\frac{L_0}{r_0} \right)^{5/3} \frac{c}{2} \left(\frac{2\pi r}{L_0} \right)^{5/6} K_{5/6} \left(\frac{2\pi r}{L_0} \right). \qquad (4)$$

A commonly accepted assumption is that all the layers have the same second order statistics: Thus from (4)

$$\mathbf{E}[\psi_i(r_1)\psi_i(r_2)] = C_\phi(|r_1 - r_2|). \qquad (5)$$

C. Description on a discrete domain

Fig. 1(b) shows the domain of images formed on the telescope lens. Since equations derived considering the simplified domain of Fig. 1(c) can easily be extended to the case of Fig. 1(b), we will neglect the central hole of the telescope, thus we will concentrate on the case of Fig. 1(c).

In real applications only a finite number of sensors is available: These are usually distributed on a grid, thus the turbulent phase is measured only on a discrete domain $\mathbb{L}$, which is that in Fig. 1(d), i.e. a sensor is placed at each node of the grid. Without loss of generality we assume that sensors are uniformly spaced: The closest neighbors of each sensor (both along the horizontal and the vertical directions) are placed at a distance of p_s meters.

Let $\phi(t)$ and $\psi_i(t)$ be the column vectors containing respectively $\{\phi(r_j, t), \ r_j \in \mathbb{L}\}$ and $\{\psi_i(r_j, t), \ r_j \in \mathbb{L}\}$, then $\psi_i(t)$ and $\phi(t)$ inherit the statistical properties of $\psi_i(\cdot, \cdot)$ and $\phi(\cdot, \cdot)$: i.e. they are zero-mean (wide-sense) stationary processes, with $C_\psi = \mathbf{E}[\psi_i(t)\psi_i(t)^T] = \mathbf{E}[\phi(t)\phi(t)^T], \ \forall i, t$ and $\mathbf{E}[\psi_i(t)\psi_j(t')^T] = 0 \ \forall i, j, t, t', \ i \neq j$. Each entry of C_ψ can be easily computed from (5) as follows

$$\mathcal{C}_\psi(i, j) = \mathbf{E}\left[\psi(r_i)\psi(r_j)\right] = C_\phi(|r_i - r_j|). \qquad (6)$$

Hence we can rewrite (1) as follows:

$$\phi(t) = \sum_{i=1}^{l} \gamma_i \psi_i(t) \qquad (7)$$

where $\psi_i(t), \ i = 1, \dots, l$ are independent and have the same second order properties. Thus the coefficient γ_i relates to the energy associated to the ith layer.
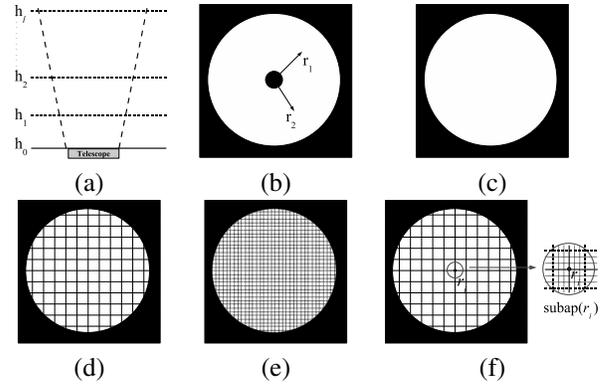
Fig. 1. (a) Atmospheric turbulence is modeled as a superposition of l layers. (b) Telescope image domain and coordinates. (c) Telescope image simplified domain. (d) Continuous line grid: domain $\mathbb{L}$. (e) Dashed line grid: domain $\mathbb{L}_{sp}$. (f) Example of subaperture.

III. ADAPTIVE OPTICS

The adaptive optics system is formed by a wavefront sensor and by a set of deformable mirrors. Its aim is to properly control the deformable mirrors to compensate the signal's phase delays, due to the atmospheric turbulence.

Since the deformable mirrors modify the signal on the telescope aperture, they can be viewed as a feedback. We can summarize the algorithm of the adaptive optics system with the following procedure:

1) estimate the current turbulent phase, as described in Section IV;
2) compute the correction contribution to obtain the compensated phases;
3) control the deformable mirrors, i.e. apply to the system the new correction phases.

The most general relation between the deformable mirrors input $u(t)$, the real turbulent phase $\phi(t)$ and the measurement, $y(t)$, is the following

$$y(t) = \mathcal{H}(\phi(t) + \mathcal{D}(u(t))) + w(t) \qquad (8)$$

where w is a white noise, $\mathcal{H}(\cdot)$ is the measurement function and $D(\cdot)$ the "actuator function". Even if there are several

types of noise which can occur in the measurement process, we shall assume that their global effect, w, is a zero-mean Gaussian noise, i.e. $w \sim \mathcal{N}(0, \Sigma_w)$. Usually $\Sigma_w = I\sigma_w^2$. Moreover we assume $w(t)$ uncorrelated with $\phi(t') \ \forall t'$ and with $w(t') \ \forall t' \neq t$. Usually (8) is well approximated by its linear counterpart [10] [8]:

$$y(t) = H\phi(t) + HDu(t) + w(t) \qquad (9)$$

where H and D are suitable matrices that correspond to $\mathcal{H}(\cdot)$ and $\mathcal{D}(\cdot)$.

The sum of the time to acquire an image and the computational time to compute the new control $u(t)$ is larger than the sampling period T: Thus we assume a $\bar{k}T$ delay in the feedback, i.e. to compute $u(t)$ we can use the measurements $y(t-k), \ k \geq \bar{k}$, while $y(t-k), \ k = 0, \ldots, \bar{k}-1$ cannot be used. To improve performances $u(t)$ is commonly computed exploiting $\hat{\phi}(t|t-\bar{k})$, the prediction of $\phi(t)$ given measurements only till time $t-\bar{k}$. A possible choice for $u(t)$ shall be

$$u(t) = D^\dagger \hat{\phi}(t|t-\bar{k})$$

where $D^\dagger$ is pseudo-inverse of D. A proper choice for $\bar{k}$ shall be $\bar{k} = 2$ (see [8] [9]).

Now let m be the number of actuators, $|\mathbb{L}|$ the number of sensors in $\mathbb{L}$ and p the number of measurements, where commonly $p = 2|\mathbb{L}|$ (in the hypothesis of Shack-Hartmann wavefront sensors, Section IV). Since $|\mathbb{L}|$ can be quite large (in our simulations $|\mathbb{L}| \approx 10^3$), the idea of building a dynamic model on $\phi(t)$ is impracticable. Hence, to reduce the computational time and the influence of noise, $\phi(t)$ is projected on a set of spatial bases $\bar{C} = \begin{bmatrix} c_0 & c_1 & \ldots & c_{|\mathbb{L}|} \end{bmatrix}$.

Astronomers commonly choose the set of Zernike polynomials as bases, however, to exploit the knowledge about the second order statistical properties of the signal, we use the set of bases provided by principal component analysis [6].

First notice that the adaptive optics system doesn't take into consideration the phase translation over the entire telescope aperture: Thus we will neglect the projection of $\phi(t)$ on $c_0 = \begin{bmatrix} 1 & 1 & \ldots & 1 \end{bmatrix}^T$: Instead of $\phi(t)$, we will consider

$$\begin{aligned} \varphi(t) &= \phi(t) - \frac{1}{|\mathbb{L}|} \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \left(\begin{bmatrix} 1 & \ldots & 1 \end{bmatrix} \phi(t) \right) \\ &= \left(I - \frac{1}{|\mathbb{L}|}\mathbf{1} \right) \phi(t) \end{aligned}$$

where $\mathbf{1}$ is a $|\mathbb{L}| \times |\mathbb{L}|$ matrix of ones. Let $\Sigma_\varphi = \mathbf{E}[\varphi(t)\varphi(t)^T]$ then

$$\Sigma_\varphi = \left(I - \frac{1}{|\mathbb{L}|}\mathbf{1} \right) C_\psi \left(I - \frac{1}{|\mathbb{L}|}\mathbf{1} \right)^T .$$

Σ_φ is a covariance matrix, thus there exists a unitary matrix $U = \begin{bmatrix} u_1 & \ldots & u_{|\mathbb{L}|} \end{bmatrix}$, i.e. $UU^T = U^TU = I.$, such that

$$\Sigma_\varphi = U\Lambda U^T \qquad (10)$$

with

$$\Lambda = \text{diag}(\lambda_1, \ \lambda_2, \ \ldots, \ \lambda_{|\mathbb{L}|}), \ \lambda_1 \geq \cdots \geq \lambda_{|\mathbb{L}|} \geq 0 .$$

Since $\varphi(t)$ is orthogonal to c_0, it is simple to prove that $\lambda_{|\mathbb{L}|} = 0$, thus $\{c_0, u_1, \ldots, u_{|\mathbb{L}|-1}\}$ forms a basis of $\mathbb{R}^{|\mathbb{L}|}$.

Define $x' = U^T\varphi$ then $\mathbf{E}[x'x'^T] = \Lambda$. x' is called the vector of principal components of φ, while U is the set of orthogonal bases associated to the principal components.

Principal component analysis provides an optimal dimensionality reduction step: Indeed it is sufficient to consider only the first n principal components to get the minimum variance (of the projection error) approximation of the original vector. Let x be the random vector constructed from φ by means of the first n principal components, that is

$$x(t) = \begin{bmatrix} u_1 & \ldots & u_n \end{bmatrix}^T \varphi(t) . \qquad (11)$$

Then

$$\varphi(t) = Cx(t) + \eta(t) \approx Cx(t)$$

where $C = \begin{bmatrix} u_1 & \ldots & u_n \end{bmatrix}$ and $\eta(t) = \varphi(t) - Cx(t)$. Moreover

$$\mathbf{E}\|\eta\|^2 = \mathbf{E}(\varphi - Cx)^T(\varphi - Cx) = \sum_{i=n+1}^{|\mathbb{L}|-1} \lambda_i. \qquad (12)$$

It is possible to prove that this is the minimum distance between the random vector φ and a vector given by a linear combination of n bases. In the following sections we will consider a linear dynamic model for $x(t)$ instead of $\varphi(t)$: Since usually $n \ll |\mathbb{L}|$, this will remarkably reduce the time to compute predictions and controls. Hereafter we will assume that the effect of the deformable mirrors is exactly known, i.e. D or $\mathcal{D}(\cdot)$, is known. Hence, from the prediction point of view, we can discard it, for example defining a new measurement vector $y'(t) = y(t) - HDu(t)$. This is equal to assume $u(t) \equiv 0, \ \forall t$: Thus we investigate the problem of phase prediction assuming the system in open loop, i.e. $u(t) \equiv 0, \ \forall t$.

IV. PHASE RECONSTRUCTION

In this section we introduce a statistical model for the measurement procedure. To reduce the noise influence on the measurements, sensor at point r_i usually takes some kind of spatial mean of the turbulent phase among its neighborhood. We call *subaperture* corresponding to r_i, the set of points considered by the sensor placed on r_i to take its measurement. To be more precise, first let $\mathbb{L}_{sp}$ be the grid of Fig. 1(e), then $\mathbb{L}_{sp}$ is decomposed in $|\mathbb{L}|$ subsets, which are the subapertures: The subaperture corresponding to $r_i \in \mathbb{L}$ is

$$\text{subap}(r_i) = \left\{ r_j \in \mathbb{L}_{sp} \mid r_i = \arg\min_{r_{i*} \in \mathbb{L}} ||r_{i*} - r_j||^2 \right\} .$$

Accordingly to the above definition subapertures are disjoint sets. An example of subaperture is the set of nodes, appertaining to grid $\mathbb{L}_{sp}$, inside the square in bold dashed line in Fig. 1(f).

Let $\bar{\phi}(t)$ be a vector containing the turbulent phase values on the nodes of $\mathbb{L}_{sp}$ at time t. From the above considerations, (9) should be rewritten as follows:

$$y(t) = H\bar{\phi}(t) + HDu(t) + w(t) \qquad (13)$$

Notice that $\mathbb{L} \subset \mathbb{L}_{sp}$, thus for each i, $1 \le i \le |\mathbb{L}|$, there exists a j, $1 \le j \le |\mathbb{L}_{sp}|$, such that $\phi_i(t)$, the ith component of $\phi(t)$, is equal to $\bar{\phi}_j(t)$, the jth component of $\bar{\phi}(t)$. Thus we can define

$$W(i,j) = \begin{cases} 1 & \text{if } \phi_i(t) = \bar{\phi}_j(t) \\ 0 & \text{otherwise.} \end{cases}$$

such that $\phi(t) = W\bar{\phi}(t)$.

In our simulation we will consider two cases for H:

1) The measurement process of $\phi(r_i',t)$ is modeled as a spatial mean on the subaperture corresponding to r_i' (as in Fig. 1(f) and Fig. 2(a)), that is:

$$\phi(r_i',t) \approx \left(\sum_{r_j \in \text{subap}(r_i')} \bar{\phi}(r_j,t) \right) \frac{1}{|\text{subap}(r_i')|}.$$

Accordingly with the above equation, H is

$$H(i,j) = \begin{cases} \frac{1}{|\text{subap}(r_i')|} & \text{if } r_j \in \text{subap}(r_i') \\ 0 & \text{otherwise.} \end{cases}$$

2) In this case we simulate the *Shack-Hartmann* sensor: It measures the vertical and horizontal slopes of the phase, instead of measuring the phase itself. The measurement procedure is assumed to be *quadcell*-like, [10]. For each $r_i' \in \mathbb{L}$, let the sets $I_1(r_i')$, $I_2(r_i')$, $I_3(r_i')$, $I_4(r_i')$ be defined as follows

$$I_1(r_i') = \{r_j \in \mathbb{L}_{sp} \mid r_j \in \text{subap}(r_i'),\ r_j \text{ is in the top row of subap}(r_i')\}$$

$$I_2(r_i') = \{r_j \in \mathbb{L}_{sp} \mid r_j \in \text{subap}(r_i'),\ r_j \text{ is in the bottom row of subap}(r_i')\}$$

$$I_3(r_i') = \{r_j \in \mathbb{L}_{sp} \mid r_j \in \text{subap}(r_i'),\ r_j \text{ is in the left-border column of subap}(r_i')\}$$

$$I_4(r_i') = \{r_j \in \mathbb{L}_{sp} \mid r_j \in \text{subap}(r_i'),\ r_j \text{ is in the right-border column of subap}(r_i')\}$$

$I_1(r_i')$, $I_2(r_i')$, $I_3(r_i')$, $I_4(r_i')$ are also shown in Fig. 2(b) and Fig. 2(c). Then vertical and horizontal slopes at r_i' are approximated introducing a block matrix $H = \begin{bmatrix} H_1^T & H_2^T \end{bmatrix}^T$ and defining

$$H_1(i,j) = \begin{cases} \frac{1}{|I_1(r_i')|} & \text{if } r_j \in I_1(r_i') \\ -\frac{1}{|I_2(r_i')|} & \text{if } r_j \in I_2(r_i') \\ 0 & \text{otherwise} \end{cases}$$

$$H_2(i,j) = \begin{cases} \frac{1}{|I_3(r_i')|} & \text{if } r_j \in I_3(r_i') \\ -\frac{1}{|I_4(r_i')|} & \text{if } r_j \in I_4(r_i') \\ 0 & \text{otherwise.} \end{cases}$$

Finally we define the (input) signal to noise ratio as follows: $SNR = \text{trace}(H\Sigma_\varphi H^T)/\text{trace}(\Sigma_w)$.

Since we cannot access directly to the value of $\varphi(t)$, we use the measurement vector $y(t)$ to estimate it. The minimum variance linear estimator of $\varphi(t)$ given $y(t)$ is

$$\hat{\varphi}(t|y(t)) = F'y(t)$$
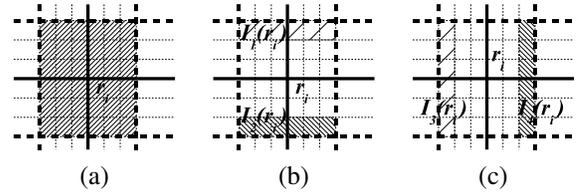
Fig. 2. (a) Spatial mean on the subaperture corresponding to r_i. (b) Shack-Hartmann's vertical slope estimation. (c) Shack-Hartmann's horizontal slope estimation.

where $F' = \left(I - \frac{1}{|\mathbb{L}|}\mathbf{1}\right)W\Sigma_{\bar{\phi}}H^T\left(H\Sigma_{\bar{\phi}}H^T + \Sigma_w\right)^{-1}$ and $\Sigma_{\bar{\phi}} = \mathbf{E}[\bar{\phi}(t)\bar{\phi}(t)^T]$. $\Sigma_{\bar{\phi}}$ can be computed similarly to C_ψ in (6). Recalling the representation of $\varphi(t)$ introduced in Section III we have that $\hat{\varphi}(t|y(t)) \approx C\hat{x}(t|y(t))$ where

$$\hat{x}(t|y(t)) = C^T\hat{\varphi}(t|y(t)) = C^TF'y(t) = Fy(t)$$

where $F = C^TF'$. The above equation provides a "static" estimation of $x(t)$. In the following section we will include it in a dynamic model that shall provide also good predictions of $x(t)$.

V. TURBULENCE PREDICTION

Since the prediction of $\varphi(t)$ is needed to improve the performances of the adaptive optics system, and since $\varphi(t) \approx Cx(t)$, then we model the dynamic of $x(t)$, which hereafter will be called the state of the system. From the Taylor approximation, (2) and (7), the system dynamic seems to be not so complex, hence a linear dynamic shall fit it quite well. We consider the following linear system

$$\begin{cases} x(t+1) = Ax(t) + v(t) \\ z(t) = x(t) + \xi(t) \end{cases} \tag{14}$$

where A is a suitable $n \times n$ matrix, $v(t)$ is a zero-mean white Gaussian noise, $v(t) \sim \mathcal{N}(0,Q)$, $z(t) = Fy(t)$ and $\xi(t)$ is zero-mean white Gaussian noise. The noise processes $v(t)$ and $\xi(t)$ are assumed to be orthogonal.

To reduce the computational load, we use the estimation equation of x as output equation instead of using $y(t)$. $\xi(t)$ is the estimation error of $x(t)$, thus it is orthogonal to $z(t)$. Hence $\xi(t) \sim \mathcal{N}(0,R)$ where $R = C^T\Sigma_\varphi C - F(H\Sigma_{\bar{\phi}}H^T + \Sigma_w)F^T$. We investigate two cases for A and Q:

1) First we consider A and Q diagonal, as already considered in previous works ([8], [9], [2]). In this case $A = \text{diag}(a_1,\ldots,a_n)$ and $Q = \text{diag}(q_1,\ldots,q_n)$. We compute A and Q via least squares from N samples: Let $\bar{x}(1),\ldots,\bar{x}(N)$ be N samples of the state vector and let $X = \begin{bmatrix} \bar{x}(1) & \ldots & \bar{x}(N-1) \end{bmatrix}$, $Y = \begin{bmatrix} \bar{x}(2) & \ldots & \bar{x}(N) \end{bmatrix}$. Let X_i and Y_i be respectively the ith row of X and Y, and let $\bar{x}_i(t)$ be the ith component of $\bar{x}(t)$. Then (assuming $||X_i||_2 \neq 0$)

$$a_i = Y_i^TX_i(X_i^TX_i)^{-1}\ ,\ i = 1,\ldots,n$$

and

$$q_i = \frac{1}{N-2}\sum_{t=1}^{N-1}\bar{v}_i(t)^2$$

where $\bar{v}_i(t) = \bar{x}_i(t+1) - a_i\bar{x}_i(t)$.

2) In this case we consider A and Q full matrices. Again we compute A and Q via least squares from N samples (assuming $|X^T X| \neq 0$):

$$A = Y^T X (X^T X)^{-1} \ , \ i = 1, \ldots, n$$

and

$$Q = \frac{1}{N-2} \sum_{t=1}^{N-1} \bar{v}(t) \bar{v}(t)^T$$

where $\bar{v}(t) = \bar{x}(t+1) - A\bar{x}(t)$.

From (11) and (10), the components of $x(t)$ are uncorrelated: This means that $\mathbf{E}[x(t)x(t)^T]$ is diagonal, however $\mathbf{E}[x(t+1)x(t)^T]$ in general is not diagonal, and thus A is not diagonal too. Hence we expect that the use of a full A in the dynamic model (14) can lead to better performances than an A diagonal.

Let us provide a simple example to prove that A in general can be far to be diagonal: let $\nu(t)$ be a zero-mean Gaussian white-noise process, with $\mathbf{E}\nu(t)^2 = 1$. Let us have two sensors: let the measurement taken by the first be $y_1(t)$ while $y_2(t)$ is that taken by the second, thus $y(t) = [y_1(t) \ y_2(t)]^T$. Moreover let them be perfect sensors: The noise has no influence on the measurements. Let $\nu(t)$ represent the new value of a "turbulence layer" that translates from the first sensor to the second exactly in a sample period, i.e. $y_1(t) = \nu(t)$, $y_2(t) = \nu(t-1)$. In this case $C_\psi = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. To make things simpler we assume here to be interested also on the signal's mean, that is we compute principal components of $\phi(t) = y(t)$ (in this example considering $\varphi(t)$ instead of $\phi(t)$ makes, fruitlessly, computation more complicated). We obtain $U = C = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $x(t) = [x_1(t) \ x_2(t)]^T = [y_1(t) \ y_2(t)]^T$. Then $\mathbf{E}[x(t)x(t)^T] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, while $\mathbf{E}[x(t+1)x(t)^T] = \mathbf{E}[\begin{bmatrix} \nu(t+1) \\ \nu(t) \end{bmatrix} [\nu(t) \ \nu(t-1)]] = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$. Hence $A = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$, which is definitely not diagonal.

This example can be extended also to higher order dynamics and suggests that the common assumption ([8], [2], [9]) of A diagonal (or more in general $\mathbf{E}[x_i(t)x_j(t') \mid x_i(\bar{t}), \ldots, x_i(t-1)] = 0$, with $\bar{t} \leq t' \leq t-1$ and $i \neq j$, $\bar{t}$ is chosen depending on model's order) can sometimes be unrealistic. In Section VI we try to explore how prediction performances of system (14) change taking A diagonal or full.

We use the dynamic model (14) to compute the k-step forward prediction of the state. This is done using the Kalman filter. Assuming A asymptotically stable, it is well known that the Kalman filter will asymptotically converge, i.e. the algebraic Riccati equation (ARE) associated to the Kalman filter will have a unique solution. The ARE associated to the state prediction in system (14) is:

$$P_1 = A(P_1 - P_1(P_1 + R)^{-1}P_1)A^T + Q$$

where P_1 is the asymptotic covariance of the 1-step prediction error. Let $\hat{x}(t+k|t)$ be the prediction of $x(t+k)$ given measurements until time t. Its prediction error is $\epsilon_k(t+k) = x(t+k) - \hat{x}(t+k|t)$. If (14) fits quite well the real system dynamic, then $\Sigma_{\epsilon_k} = \mathbf{E}[\epsilon_k(t+k)\epsilon_k(t+k)^T]$ is well approximated by the asymptotic error covariance computed by the ARE solution: i.e. $\Sigma_{\epsilon_k} \approx P_k$, where P_k is defined as follows

$$P_k = A^{k-1} P_1 (A^T)^{k-1} + \sum_{i=1}^{k-1} A^{k-1-i} Q (A^T)^{k-1-i} \ . \quad (15)$$

On the other hand the prediction error for $\varphi(t+k)$ is $e_k(t+k) = \varphi(t+k) - C\hat{x}(t+k|t) = C(x(t+k) - \hat{x}(t+k|t)) + \eta(t+k)$, thus from (15) and (12), we have

$$\mathbf{E}\|e_k\|^2 = \text{trace}(C\Sigma_{\epsilon_k}C^T) + \sum_{i=n+1}^{|\mathbb{L}|-1} \lambda_i \quad (16)$$

$$\approx \text{trace}(CP_kC^T) + \sum_{i=n+1}^{|\mathbb{L}|-1} \lambda_i \ . \quad (17)$$

We stress the fact that (17) is an asymptotic error, but it depends on the model chosen in (14) and on the learnt parameters, i.e. on how close the identified model is to the true one. Fig. 3 shows a comparison between $\mathbf{E}\|e_k\|^2$ computed with (16) and (17) in a simulation. We set $k = 2$, $L_0 = 16$m, $r_0 = 1.5$m, the telescope aperture diameter $d = 8$m, $p_s = 0.2$m, $|\mathbb{L}_{sp}/\mathbb{L}| = 9$, $SNR = 4$, $N = 2000$. The turbulence is formed by 3 independent layers, which move with different velocities: 7m/s at 0km high, 5.7m/s at 6km, 11m/s at 8.5km. The sample period of the adaptive optics system is set to 32ms.

In practice when A and Q are full matrices (17) provides a good approximation of (16), while when A and Q are diagonal (17) is usually smaller than (16). To make the results of our simulations quite independent on the number of validation samples, in Section VI we report the results obtained using (17) instead of (16). However this is not sufficient to make the results independent on the learning sequence: Since (17) depends on the learnt parameters, it can still take to slight different results in different simulations, e.g. (17) is not strictly decreasing in Fig. 3(b).
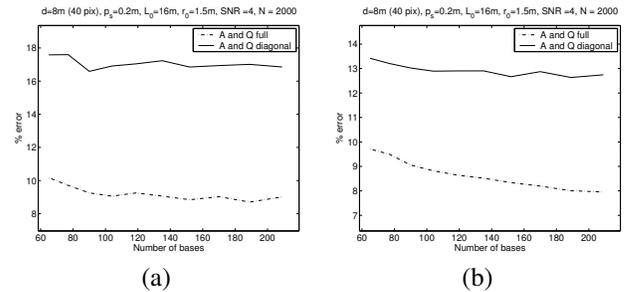
Fig. 3. In solid line the 2-step prediction errors obtained using A and Q diagonals. In dashed line those obtained with A and Q full. Measurements are obtained simulating the Shack-Hartmann sensor. (a) Percent sample error computed with (16). (b) Percent asymptotic error computed with (17).

VI. SIMULATIONS AND DISCUSSION

From (7), the turbulence is a linear combination of the layers: In our simulation we consider turbulence formed by $l = 3$ layers which move with different velocities. We compare prediction performances of the dynamic system (14) with A, Q diagonal and with A, Q full. The comparison is made among two possible turbulence conditions. For each turbulence condition we compute (17) ranging the number of bases from $n = 65$ to $n = 209$. Moreover we compute the performances considering both spatial mean and the Shack-Hartmann models for the measurement process (as described in Section IV). The percent norm of the error is plotted in the figures, i.e. $\mathbf{E}[\|e_k\|^2]/\mathbf{E}[\|\varphi\|^2] \cdot 100$.

Since the delay due to image acquisition and control computation should last approximatively $2T$ [8], [9] we set $k = 2$. The results reported in Fig. 4 are obtained setting the values of the parameters to: $k = 2$, $r_0 = 1.5$m, $d = 8$m, $p_s = 0.2$m, $|\mathbb{L}_{sp}/\mathbb{L}| = 9$, $SNR = 4$, $N = 2000$. Different values for (L_0, SNR) are explored: (a),(b) $L_0 = 12$m, $SNR = 4$; (c),(d) $L_0 = 20$m, $SNR = 4$. The turbulence is formed by 3 independent layers, which move with different velocities: 7m/s at 0km high, 5.7m/s at 6km, 11m/s at 8.5km. Furthermore the sample period of the adaptive optics system is set to 32ms.

As expected, the "full matrices" system outperforms the diagonal one, however at the cost of an increase in the running time: While the time complexity for state prediction is $O(n)$ for the diagonal system, it becomes $O(n^2)$ in the "full matrices" case . Since the system has to work in real-time, it is worth to quantify the increase in complexity and eventually give some suggestions to limit it. The total running time for a cycle of the adaptive optics algorithm is $O(np) + O(nm) + O(n^2)$, where $O(np)$ and $O(nm)$ are respectively the time to estimate $\hat{x}(t|t)$ and to compute the control $u(t + k)$, while $O(n^2)$ is the time to compute the prediction $\hat{x}(t + k|t)$. Notice that by construction $p \gg n$, hence $O(n^2)$ is not the dominant term. Since the "full matrices" system has better performances than the diagonal one even using much less bases, one can use a smaller n to reduce the running time. Finally, since A and Q can be identified off-line, their identification is not to be considered as a relevant factor in the time complexity of the on-line algorithm.

VII. CONCLUSIONS

In this paper we have compared two types of linear models for the prediction of atmospheric turbulent phase.

We have chosen the principal component representation to have the minimum representation error for the signal. Then we have compared the performance on 2-step prediction of a diagonal linear model, as already proposed in literature, and of a "full matrices" dynamic model. In all our simulations the latter improves the performances of the first.

The cost of this improvement is an increase in the time complexity of the prediction step. However in our simulations the performances of the "full matrices" system are better than those of the diagonal one even using much less

(a)   (b)

(c)   (d)
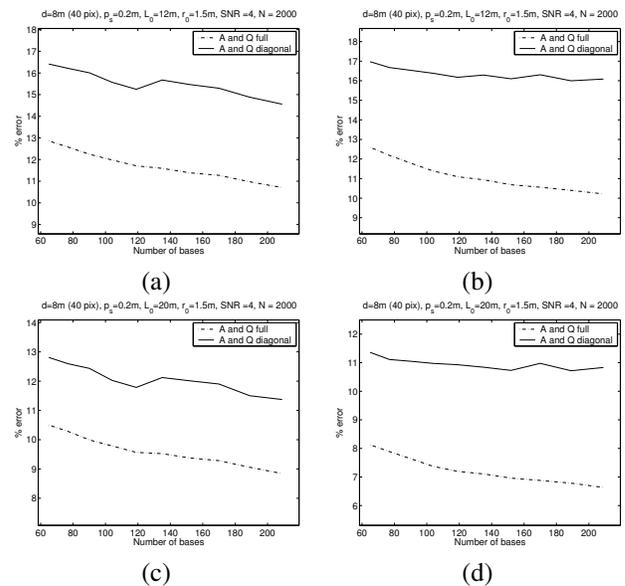
Fig. 4.   In solid line the 2-step prediction errors obtained using A and Q diagonals. In dashed line those obtained with A and Q full. Both spatial mean (in (a), (c)) and the Shack-Hartmann (in (b), (d)) methods of measurement are simulated. Different values for (L_0, SNR) are explored: (a),(b) $L_0 = 12$m, $SNR = 4$; (c),(d) $L_0 = 20$m, $SNR = 4$.

bases: Hence we are confident that the use of a smaller number of bases combined with the use of the "full matrices" dynamic model can lead to better performances without compromising the running time.

VIII. ACKNOWLEDGMENTS

REFERENCES

[1] The very large telescope project. *from the ESO website*, 2006.
[2] J.M. Conan, G. Rousset, and P.Y. Madec. Wave-front temporal spectra in high-resolution imaging through turbulence. *Journal of the Optical Society of America A*, 12(7):1559–1570, 1995.
[3] 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, Université de Nice Sophia Antipolis Faculté des Sciences, 2000.
[4] P. Dierickx, J.L. Beckers, E. Brunetto, and et al. The eye of the beholder: designing the owl. In *Proceedings of SPIE – Future Giant Telescopes*, volume 4840. J. Roger P. Angel, Roberto Gilmozzi, Editors, 2003.
[5] D.L. Fried. Statistics of a geometric representation of wavefront distortion. *J. Opt. Soc. Am.*, 55:1427–1435, 1965.
[6] W. Hardle and L. Simar. *Applied multivariate statistical analysis*. Springer, 2003.
[7] M. Le Louarn, N. Hubin, M. Sarazin, and A. Tokovinin. New challenges for adaptive optics: extremely large telescopes. *Mon. Not. R. Astron. Soc.*, 317, 2000.
[8] B. Le Roux. *Commande optimale en optique adaptative classique et multiconjuguee*. PhD thesis, Université de Nice Sophia Antipolis, 2003.
[9] 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.
[10] F. Roddier. *Adaptive optics in astronomy*. Cambridge university press, 1999.