16th Mediterranean Conference on Control and Automation
Congress Centre, Ajaccio, France
June 25-27, 2008

A Markov-Random-Field-based approach to modeling and prediction of atmospheric turbulence

Alessandro Beghi, Angelo Cenedese and Andrea Masiero

Abstract— **Nowadays, the adaptive optics (AO) system is of fundamental importance to improve the real resolution of ground-based telescopes. In practical applications the telescope resolution is limited by the atmospheric turbulence. The aim of the AO system is that of estimating the atmospheric turbulence and computing a suitable input for a set of deformable mirrors to reduce the turbulence effect. A commonly accepted assumption is that of considering the turbulence as formed by a discrete set of layers moving over the telescope lens. In this paper, we first propose a method for estimating the number of layers and their characteristics. Then, we exploit the information on the turbulence layers to construct a linear predictor of the turbulent phase. Performance of the proposed method is shown by means of simulations.**

I. INTRODUCTION

The wavefront signal from a star object to a ground-based telescope is distorted along the light path proportionally to the length of the optic path, and depending on the encountered refraction index. Actually, changes in the local temperature and the presence of wind make the refraction index of the atmosphere change quite fast both spatially and temporally. As a consequence, the light beams are delayed of a different phase. Therefore, a flat wavefront surface coming from a star is no longer flat when it is detected on the telescope pupil, thus significantly reducing the real resolution of the telescope. Hereafter, we will call *turbulent phase* the set of phase delays of the beams arriving on the telescope pupil.

To reduce the problems due to the presence of the atmospheric turbulence, telescopes are usually provided with an adaptive optics (AO) system [1], that commands a set of correction mirrors (or deformable mirrors) to adapt their shapes so as to compensate for the current value of the turbulent phase. Thus the beams arriving on the telescope pupil, after passing through the deformable mirrors, have a residual turbulent phase as close to zero as possible. A cycle of the AO system's working procedure can be summarized into three steps, namely 1) estimating the current turbulent phase, 2) predicting the new one, and 3) computing the proper control input for the set of deformable mirrors. Notice that the control is commonly delayed of two sample periods, that is the time needed for image acquisition and phase reconstruction (see [2], [3]). This fact makes the prediction step fundamental to yield good performance of the AO system.

The atmospheric turbulence is commonly described as a set of layers moving at different altitudes over the telescope lens. The total effect of these layers is usually computed as a linear combination of the values of the single layers. In this paper we describe a method to estimate the characteristics of the layers (energies and velocities) from the turbulent phase measurements. Then, by exploiting the estimated characteristics of the turbulence, we temporally predict the turbulence by means of a linear dynamical system.

One of the main advantages of the proposed approach is that it can be generalized to the Multi-Conjugated Adaptive Optics (MCAO) case, that grants larger sky coverage with respect to standard AO systems [3]. To achieve this, in a MCAO system the atmosphere structure is completely reconstructed and different mirrors are used to correct the turbulent phase associated to different turbulent layers. This fact is reflected in the structure of the linear predictor that we propose: The state vector is partitioned in blocks, each of them describing the current phase values of one of the turbulence layers. Hence, this model provides a complete reconstruction of the turbulence as required by a MCAO system.

The paper is organized as follows. In Section II the common turbulence statistical model is briefly described. Then, in Section III we introduce a Markov Random Field spatial representation for the atmospheric turbulence and we exploits it to estimate the atmospheric turbulence structure. In Section IV we propose a linear system for turbulence prediction based on the turbulence structure estimated in Section III. We conclude in Section V with showing the results of some simulations.

II. TURBULENCE PHYSICAL MODEL

The spatial statistical characteristics of the turbulent phase ϕ are generally 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, which are separated by a distance r on the aperture plane (Fig. 1),

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

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

Alessandro Beghi is Dipartimento di Ingegneria dell'Informazione, Università di Padova, via Gradenigo 6/B, 35131 Padova, Italy. beghi@dei.unipd.it

Angelo 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

Andrea Masiero is Dipartimento di Ingegneria dell'Informazione, Università di Padova, via Gradenigo 6/B, 35131 Padova, Italy. masiero@dei.unipd.it

978-1-4244-2505-1/08/$20.00 ©2008 IEEE

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

$$D_\phi(r) = 2\left(\sigma_\phi^2 - C_\phi(r)\right), \tag{1}$$

where σ_ϕ^2 is the phase variance.

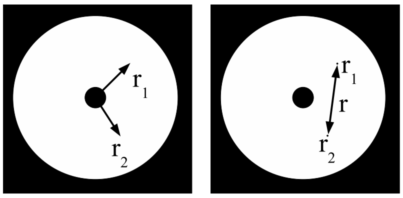


Fig. 1. Telescope image domain: Two points, r_1 and r_2, separated by a distance r on the telescope aperture plane.

According to the Von Karman theory, the phase structure function evaluated at distance r is the following (see [4]):

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

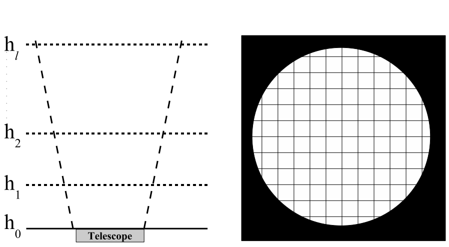
where $K_\cdot(\cdot)$ is the MacDonald function (modified Bessel function of the third type), Γ is the Gamma function, L_0 is the outer scale, r_0 is a characteristic parameter called the Fried parameter (see [5]), and the constant c is:

$$c = \frac{2^{1/6}\Gamma(11/6)}{\pi^{8/3}}\left[\frac{24}{5}\Gamma(6/5)\right]^{5/6}.$$

From the relation between the structure function and the covariance (1), the spatial covariance of the phase between two points at distance r results

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

Notice that 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 of Fig. 2(b), 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. We denote with $\phi(u, v, t)$ the value of the turbulent phase on the point $(u, v) \in \mathbb{L}$ at time t.



(a) (b)

Fig. 2. (a) Atmospheric turbulence is modeled as a superposition of l layers. (b) Discrete domain $\mathbb{L}$.

In order to describe its temporal characteristics, the turbulence is generally modeled as the superposition of a finite number l of layers. The i^{th} layer models the atmosphere from an altitude of h_{i-1} to h_i meters, where $h_l \geq \cdots \geq h_i \geq h_{i-1} \geq \cdots \geq h_0 = 0$ (Fig. 2(a)). Let $\psi_i(u, v, t)$ be the value of the i^{th} layer at point (u, v) on telescope aperture and at time t. Then the total turbulent phase at (u, v) and at time t is

$$\phi(u, v, t) = \sum_{i=1}^{l} \gamma_i \psi_i(u, v, t), \tag{3}$$

where γ_i are suitable coefficients. 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 the same spatial characteristics, i.e. all the layers are spatially described by the same structure function. Furthermore they are assumed to be independent, hence

$$\mathbf{E}[\psi_i(u, v, t)\psi_j(u', v', t')] = 0, \; 1 \leq i \leq l, 1 \leq j \leq l,$$
$$j \neq i, 1 \leq u \leq m, 1 \leq v \leq m, 1 \leq u' \leq m, 1 \leq v' \leq m.$$

A commonly agreed assumption considers that each layer translates in front of the telescope pupil with constant velocity v_i (Taylor approximation [1]), thus

$$\psi_i(u, v, t+kT) = \psi_i(u - v_{i,u}kT, v - v_{i,v}kT, t), \; i = 1, \ldots, l \tag{4}$$

where $v_{i,u}$ and $v_{i,v}$ are the projections of the velocity vector v_i along respectively the horizontal and the vertical axis, while kT is a delay multiple of the sampling period T. The velocity vectors are assumed to be different for different layers, i.e. $v_i \neq v_j$ if $i \neq j$.

III. Detection of layers: Speed and energy

The aim of this section is the estimation of the turbulence parameters $(l, \gamma_1, \ldots, \gamma_l, v_{1,u}, \ldots, v_{l,u}, v_{1,v}, \ldots, v_{l,v})$ in the model given by Eqs. (3) and (4). To simplify the notation, in this Section we consider the 1D case. Then the parameters to be estimated are $(l, \gamma_1, \ldots, \gamma_l, v_1, \ldots, v_l)$, where v_i is the velocity of the ith layer. The generalization of the procedure to the 2D case is immediate, leading just to a more complex formulation of the equations.

We then assume that the turbulence is a scalar random process, and at each sampling time we observe only a limited range of its values, which is generally affected by a zero-mean white-noise process w. However, in this Section for simplicity of exposition we will assume $w = 0$.

In this case $\mathbb{L}$ reduces to a 1D interval: Let m be the interval size, then $\mathbb{L} = [1, \ldots, m]$. Let $y(u, t)$ be the value of the measured phase turbulence on the spatial position $u \in \mathbb{L}$ at time t. Similarly, let $\psi_i(u, t)$ be value of the i^{th} turbulent layer on the position $u \in \mathbb{L}$ at time t. Then at time t we measure the vector $[y(1, t), \; y(2, t), \; \ldots, \; y(m, t)]^T$.

Since the spatial correlation of the turbulence between two points at distance r decreases quite fast when r becomes larger, we spatially model the turbulent phase as a *Markov random field* (MRF). In accordance with the physical model of the turbulence of Section II, the MRF describing the spatial characteristics of the turbulent phase is assumed to be homogeneous and isotropic.

As shown in [6], $y(\bar{u},t)$, the value of the MRF on the generic point $\bar{u}$ at time t, can be expressed as the best linear prediction of $y(\bar{u},t)$ given the values of its neighbors $\mathbb{N}(\bar{u})$ at time t plus an "innovation" process $e(\bar{u},t)$. Where the neighborhood $\mathbb{N}(\bar{u})$ of the point $\bar{u} \in \mathbb{L}$ is defined (in 1D) as follows:

$$\mathbb{N}(\bar{u}) = \left\{ u \in \mathbb{L} \mid 0 < |u - \bar{u}| \leq \bar{d} \right\} ,$$

where $\bar{d}$ is a suitable distance.

According with [7], we assume that the turbulent phase has Gaussian statistics, therefore the best (spatial) linear prediction operator $\hat{\mathbf{E}}[\cdot]$ corresponds to the expectation operator $\mathbf{E}[\cdot]$, that is

$$y(\bar{u},t) = \sum_{u \in \mathbb{N}(\bar{u})} a_{|\bar{u}-u|} y(u,t) + e(\bar{u},t) , \quad (5)$$

where $\{a_i\}$ are suitable coefficients which yield the best (spatial) linear prediction of $y(\bar{u},t)$ given the values of its neighbors (see, for example, [8] for the computation of the coefficients of the best linear predictor).

Furthermore,

$$\mathbf{E}[y(\bar{u},t)e(u,t)] = \sigma_e^2 \delta_{\bar{u}-u}$$

and

$$\mathbf{E}[e(\bar{u},t)e(u,t)] = \begin{cases} \sigma_e^2 & \bar{u} = u \\ -a_{|\bar{u}-u|}\sigma_e^2 & u \in \mathbb{N}(\bar{u}) \\ 0 & \text{otherwise} \end{cases} .$$

In the above equations we denoted the Kronecker's delta with δ_u, that is

$$\delta_u = \begin{cases} 1 & u = 0 \\ 0 & \text{otherwise} \end{cases} .$$

We stress the fact that the MRF representation provides only a statistical approximation of the real process. However, as long as $\bar{d}$ is chosen sufficiently large, this can be considered a good approximation.

Since the layers have the same spatial statistical characterization, each layer can be described using a similar MRF representation of (5), that is

$$\psi_i(\bar{u},t) = \sum_{u \in \mathbb{N}(\bar{u})} a_{|\bar{u}-u|} \psi_i(u,t) + e_i(\bar{u},t) , \quad (6)$$

for $i = 1, \ldots, l$. Since the layers are independent then

$$\mathbf{E}[e_i(\bar{u},\bar{t})e_j(u,t)] = 0 , \quad (7)$$

if $i \neq j$, while

$$\mathbf{E}[e_i(\bar{u},t)e_i(u,t)] = \begin{cases} \sigma_e^2 & \bar{u} = u \\ -a_{|\bar{u}-u|}\sigma_e^2 & u \in \mathbb{N}(\bar{u}) \\ 0 & \text{otherwise} \end{cases} . \quad (8)$$

Notice that $e(u,t)$, referred to the global measured phase y (5), can be expressed as a liner combination of $e_i(u,t),\ i = 1, \ldots, l$, that is

$$e(u,t) = \sum_{i=1}^{l} \gamma_i e_i(u,t) . \quad (9)$$

Equations (5),(7),(8) and (9) allow us to formulate the algorithm for the detection of turbulence layers. First, let us assume that the translations of each layer during a sample period (that is the velocities per frame) are rational multiples of the pixel size. Then, there exists an integer number k_i, such that

$$\psi_i(u,t) = \psi_i(u + k_i v_i T, t + k_i T),\ u = 1, \ldots, m , \quad (10)$$

where v_i is the velocity of the i^{th} layer. Three observations are now in order:

- Since the coefficients $\{a_i\}$ can be easily computed from the spatial theoretical description of the turbulence (the second order spatial statistical characteristics of the turbulence are described by (2), and the coefficients $\{a_i\}$ can be computed as those of the best linear predictor), then $e(u,t),\ u = \bar{d}+1, \ldots, m-d$ can be computed for all t from (5).
- From (7) and (9), then

$$\mathbf{E}[e(\bar{u},\bar{t})e(u,t)] = \sum_{i=1}^{l} \gamma_i^2 \mathbf{E}[e_i(\bar{u},\bar{t})e_i(u,t)] . \quad (11)$$

- From (10) and (8), it follows that

$$\mathbf{E}[e_i(u,t)e_i(u + k_i v_i T, t + k_i T)] = \sigma_e^2,\ i = 1, \ldots, l, \quad (12)$$

while

$$\mathbf{E}[e_i(u,t)e_i(\bar{u}, t + k_i T)] = 0,\ i = 1, \ldots, l, \quad (13)$$

if $\bar{u} \notin \mathbb{N}(u + k_i v_i)$.

Hence, (11) is different from zero only if there is at least a layer, i, such that $\bar{u} \in \mathbb{N}(u + k_i v_i T)$, where $k_i = |\bar{t} - t|/T$. Thus, since the velocities of different layers are assumed to be different, the temporal covariances of the spatial prediction error e, provides us with a simple method to detect the turbulence layers.

Let us define $c_{\tau,j}$, r_τ and $\bar{r}$ as follows

$$c_{\tau,j} = \mathbf{E}[e(u,t)e(u+j, t+\tau)] ,$$

$$r_\tau(t) = \begin{cases} c_{\tau,t} & -m+1 \leq t \leq m-1 \\ 0 & \text{otherwise} \end{cases} ,$$

$$\bar{r}(t) = \sum_{\tau=0}^{\bar{T}} r_\tau(t - \tau(2m-1) - m) \delta_{|t-\tau(2m-1)|<m} .$$

Furthermore, let the following two definitions hold.

Definition 1: Two velocities v_i and v_j are said to be *distinguishable* in $\bar{t}$ temporal instants on the domain $\mathbb{L}$ if there exist $\bar{t}_i, \bar{t}_j$, with $1 \leq \bar{t}_i, \bar{t}_j \leq \bar{t}$, such that $\bar{t}_i v_i \in \mathbb{L}$, $\bar{t}_i |v_i - v_j| \geq 2\bar{d} + 1$ and $\bar{t}_j v_j \in \mathbb{L}$, $\bar{t}_j |v_i - v_j| \geq 2\bar{d} + 1$. This definition can be easily generalized as follows.

Definition 2: The velocities $(v_1, \ldots, v_l)$ are said to be *distinguishable* in $\bar{t}$ temporal instants on the domain $\mathbb{L}$ if for each (i,j), with $i \neq j$, (v_i, v_j) are distinguishable.

The resulting detection algorithm is as follows.

Algorithm 1: Detection of the layers

Step 1: Rough estimation of the velocities

$\hat{\mathcal{S}}_{\hat{l}} = \emptyset$;

for $\tau = \bar{T} : 1$
  for $t = -m+1 : m-1$
    $v = t/\tau$;
    if $(r_\tau(t) \neq 0 \wedge$
      $\wedge\ r_\tau(t) = \max\left(r_\tau(t-\bar{d}), \ldots, r_\tau(t+\bar{d})\right)\ \wedge$
        $\wedge\ \left(\hat{\mathcal{S}}_{\hat{l}} \cap [v - \bar{d}/\tau, v + \bar{d}/\tau] = \emptyset\right)$
      $\hat{\mathcal{S}}_{\hat{l}} = \hat{\mathcal{S}}_{\hat{l}} \cup v$;
    end
  end
end
$\hat{l} = |\hat{\mathcal{S}}_{\hat{l}}|$;
$\{\hat{v}_1, \ldots, \hat{v}_{\hat{l}}\} = \hat{\mathcal{S}}_{\hat{l}}$;
Step 2: velocities update and weights estimation for $i = 1 : \hat{l}$
  $\hat{\gamma}_i = 0$;
  for $\tau = 1 : \bar{T}$
    if $\left(\left(\hat{\mathcal{S}}_{\hat{l}} \cap [\hat{v}_i - \bar{d}/\tau, \hat{v}_i + \bar{d}/\tau] = \hat{v}_i\right)\ \wedge\right.$
      $\left.\wedge\ \left(\max(r_\tau(\hat{v}_i\tau - \bar{d}), \ldots, r_\tau(\hat{v}_i\tau + \bar{d})) > \hat{\gamma}_i^2\right)\right)$
      $\hat{\gamma}_i = (\max(r_\tau(\hat{v}_i\tau - \bar{d}), \ldots, r_\tau(\hat{v}_i\tau + \bar{d}))^{1/2}$;
      $\hat{v}_i = (\arg\max(r_\tau(\hat{v}_i\tau - \bar{d}), \ldots, r_\tau(\hat{v}_i\tau + \bar{d})))/\tau$;
    end
  end
end

where $\hat{\mathcal{S}}_{\hat{l}}$ is the set of the detected velocities. Then the following proposition holds.

Proposition 1: Let $(l, v_1, \ldots, v_l, \gamma_1, \ldots, \gamma_l)$ be the true turbulence parameters and $(\hat{l}, \hat{v}_1, \ldots, \hat{v}_{\hat{l}}, \hat{\gamma}_1, \ldots, \hat{\gamma}_{\hat{l}})$ those learnt with the proposed algorithm. If the velocities $(v_1, \ldots, v_l)$ are distinguishable in $\bar{t}$ temporal instants on the domain $\mathbb{L}$, then $\hat{l} = l$ and $\hat{v}_i = v_i$, $\hat{\gamma}_i = \gamma_i$, $i = 1, \ldots, l$.

We stress the fact that in this algorithm we have directly used the "true" temporal correlations of the spatial innovation e. However, in real applications these will not be directly accessible. Hence the above algorithm has to be slightly modified to take in account of the use of sample covariances, e.g. estimated from N temporal samples. This case, that is only formally more complicated, has been considered in [9].

IV. Turbulence temporal prediction

Notice that an AO system typically requires more than one sample period to complete the reconstruction procedure and the computation of the control law. Hence, to make its effort effective, it is of fundamental importance the introduction of a prediction step.

In this Section we consider a linear dynamic system for the temporal prediction of the turbulent phase. In the computation of the system parameters we exploit the results on layers estimation described in the previous Section.

Let $\phi(t)$ be the vector containing the values of the turbulent phase over the telescope pupil at time t. Notice that the AO system does not take into account the phase translation over the entire telescope aperture. Thus we can neglect the current mean of the signal, and we consider φ, defined as follows

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

where $\mathbf{1}$ is a $|\mathbb{L}| \times |\mathbb{L}|$ matrix of ones.

Let us consider the following linear dynamic system to model the temporal dynamic of the turbulent phase (where with an abuse of notation we use again the symbol y with a slightly different meaning):

$$\begin{cases} x(t+1) = Ax(t) + \eta(t) \\ y(t) = Cx(t) + \xi(t) \end{cases} \tag{14}$$

where $\eta(t)$ and $\xi(t)$ are zero-mean white Gaussian noises, with $\eta(t) \sim \mathcal{N}(0, Q)$ and $\xi(t) \sim \mathcal{N}(0, R)$ and η and ξ are assumed to be orthogonal. The system matrices $\{A, C, Q, R\}$ and the processes y and x are defined more precisely in the following paragraphs.

To reduce the complexity of the control computation, $\varphi(t)$ is projected on a set of spatial bases. In particular here we consider the set of spatial bases U provided by the Principal Component Analysis. In practice we consider only the n (with $n \ll |\mathbb{L}|$) bases associated to the first n principal components of φ, which are those containing most of its energy. As shown in [10], this is a well suited set of bases for turbulent phase representation. Let $z(t)$ be the projection of $\varphi(t)$ on the set of spatial bases U, $z(t) = U^T\varphi(t)$, then $\varphi(t) = Uz(t) + \epsilon(t) \approx Uz(t)$, where $\epsilon(t)$ is the representation error, i.e. the error due to the use of a small finite number of bases, n.

Furthermore, we assume that the measurement process is affected by a zero-mean Gaussian white noise w, with variance Σ_w. In our examples $\Sigma_w = \sigma_w^2 I$. Then the output process y, can be written as $y(t) = \varphi(t) + w(t) = Uz(t) + \epsilon(t) + w(t) = Uz(t) + \xi(t)$, where $\xi(t) = \epsilon(t) + w(t)$.

With an abuse of notation, we call $\psi_i(t)$ the vector containing the values of the ith turbulent layer over the telescope pupil at time t. Then, by construction, we decompose the state vector $x(t)$ in $\hat{l}$ blocks,

$$x(t) = [x_1(t)^T\ \ldots\ x_{\hat{l}}(t)^T]^T\ ,$$

with $\psi_i(t) \approx Ux_i(t)$. Thus $x_i(t)$ takes in account of the contribution of the ith layer to the current value of the turbulent phase. Moreover

$$z(t) = \sum_{i=1}^{\hat{l}} \hat{\gamma}_i x_i(t) = \begin{bmatrix}\hat{\gamma}_1 I & \hat{\gamma}_2 I & \ldots & \hat{\gamma}_{\hat{l}} I\end{bmatrix} x(t)\ .$$

From the block structure of x, it follows a similar structure also for $A \in \mathbb{R}^{(n\cdot\hat{l})\times(n\cdot\hat{l})}$ and $C \in \mathbb{R}^{|\mathbb{L}|\times(n\cdot\hat{l})}$. In particular C is defined as follows

$$C = \begin{bmatrix}\hat{\gamma}_1 U & \hat{\gamma}_2 U & \ldots & \hat{\gamma}_{\hat{l}} U\end{bmatrix}\ ,$$

while, since the layers are assumed to be independent, then

$$A = \begin{bmatrix} \hat{A}_1 & 0 & \ddots & 0 \\ 0 & \hat{A}_2 & \ddots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \ddots & \hat{A}_{\hat{l}} \end{bmatrix} ,$$

where

$$\hat{A}_i = \mathbf{E}\left[\psi_i(t+1)\psi_i(t)^T\right]\mathbf{E}\left[\psi_i(t)\psi_i(t)^T\right]^{-1} \ , i = 1, \ldots, \hat{l}.$$

In practice $\hat{A}_i$ can be easily computed from the turbulence spatial covariance (2) and the estimated velocity $\hat{v}_i$.

From (14) follows immediately that

$$R = \Sigma_y - C\Sigma_x C^T + \Sigma_w$$

and

$$Q = \Sigma_x - A\Sigma_x A^T \ ,$$

where $\Sigma_y = \mathbf{E}[y(t)y(t)^T]$ can be computed from (2). Notice that the layers are independent to each other, hence $\Sigma_x = \mathbf{E}[x(t)x(t)^T]$ and Q are a block diagonal matrices, e.g.

$$\Sigma_x = \begin{bmatrix} \hat{\gamma}_1^2\Sigma_{\psi_1} & 0 & \ddots & 0 \\ 0 & \hat{\gamma}_2^2\Sigma_{\psi_2} & \ddots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \ddots & \hat{\gamma}_{\hat{l}}^2\Sigma_{\psi_{\hat{l}}} \end{bmatrix} ,$$

where, actually, $\Sigma_{\psi_i} = \Sigma_y, \ i = 1, \ldots, \hat{l}$.

Then, the temporal prediction of the turbulent phase is made using a steady-state Kalman filter, i.e., to reduce the computational complexity the asymptotic Kalman gain is used [8].

Finally, let $\hat{x}(t+2|t)$ be the prediction of $x(t+2)$ given the measurements until time t provided by the Kalman filter applied on the linear system (14), then we compute the energy of the 2-step prediction error $\varepsilon_2(t) = (Uz(t) - C\hat{x}(t|t-2))$ as

$$\mathbf{E}\|\varepsilon_2\|^2 = \text{trace } (\Sigma_{\varepsilon_2})$$

where Σ_{ε_2} is the variance of ε_2. In the simulations we performed, the variance of Σ_{ε_2} is approximated with that provided by the steady-state Kalman filter. Let P_∞ be the covariance of the state prediction error provided by the steady-state Kalman filter, then

$$\Sigma_{\varepsilon_2} \approx C(AP_\infty A^T + Q)C^T \ .$$

Furthermore, we normalize the energy of ε_2 with respect to the energy of $U(z(t) - z(t-2))$, which is the same of the energy of $(z(t)-z(t-2))$ since U is part of a unitary matrix.

Then, we consider as performance index

$$100 \cdot \mathbf{E}\|\varepsilon_2\|^2 / \mathbf{E}\|z(t) - z(t-2)\|^2 \ . \qquad (15)$$

We stress the fact that this is not an absolute index of the system performances, since the energy of the error due to the use of a finite number n of bases is not included. However, (15) says how much (in percent energy) of the variation of the signal can be predicted by the considered system. Thus it evaluates only the prediction performances of the system. For example, if (15) is equal to 100, then, from the prediction point of view, the corresponding system is equivalent to use $z(t-2)$ as a prediction of $z(t)$.

Notice that, even if (15) sometimes increases when n becomes larger, the overall (representation and prediction) error of the system usually becomes lower when the number of bases n becomes larger.

V. Simulations

We distinguish two steps in our simulations: First, in sub-section V-A, we estimate the layers characteristics using algorithm 1, then, in sub-section V-B, we use the linear system (14) to predict the turbulence.

A. Estimation of the layers

Since usually the layers move slowly over the telescope pupil here we consider three examples of layer detection where the layers translate less than a pixel per frame. To make this possible we have simulated the layers at a sub-pixel scale: A 10×10 matrix of sub-pixels has been used to simulate each pixel in $\mathbb{L}$. In the simulations of this sub-section we set the simulation parameters to $\sigma_w = 0$ and the number of samples, N, used to estimate the temporal covariances used in algorithm 1 to $N = 5000$. When σ_w is different from zero, a larger number of samples N is needed to obtain results comparable with those reported in this sub-section.

The results of our simulations are reported in Table I,II: $v_{i,u}$, $v_{i,v}$ and γ_i corresponds to the true values of the parameters, $\hat{v}_{i,u}$, $\hat{v}_{i,v}$ and $\hat{\gamma}_i$ are the estimated ones. The velocities are written in [pixels/frame].

The results obtained with the proposed method are quite encouraging: Indeed in all of the examples the number of layers has been correctly estimated, i.e. $\hat{l} = l$, and the values of the estimated parameters are quite close to the true ones.

TABLE I
Detection of the layers.

	1^{st} layer	2^{nd} layer	3^{rd} layer	4^{th} layer
$v_{i,u}$	0.216	0.391	0.612	0.795
$\hat{v}_{i,u}$	0.2162	0.3913	0.6122	0.7949
$v_{i,v}$	0	0	0	0
$\hat{v}_{i,v}$	0	0	0	0
γ_i^2	0.31	0.3	0.2	0.19
$\hat{\gamma}_i^2$	0.3112	0.3007	0.2005	0.1876

B. Temporal prediction of the turbulent phase

In the simulations considered in this sub-section, we use the results of sub-section V-A to compute the parameters of (14) as described in Section IV. Then, for each of the examples of Tables I,II we evaluate the 2-step prediction performances of the linear system (14), computed as in (15). The results are shown respectively in Figs. 3,4.

TABLE II

DETECTION OF THE LAYERS.

	1^{st} layer	2^{nd} layer	3^{rd} layer	4^{th} layer
$v_{i,u}$	0.216	-0.191	0	0
$\hat{v}_{i,u}$	0.2162	-0.1905	0	0
$v_{i,v}$	0	0	0.11	0.287
$\hat{v}_{i,v}$	0	0	0.1111	0.2881
γ_i^2	0.41	0.25	0.2	0.14
$\hat{\gamma}_i^2$	0.4137	0.2495	0.1973	0.1395

For comparison we report also the prediction performances of a linear system with the parameters estimated as described in [11].

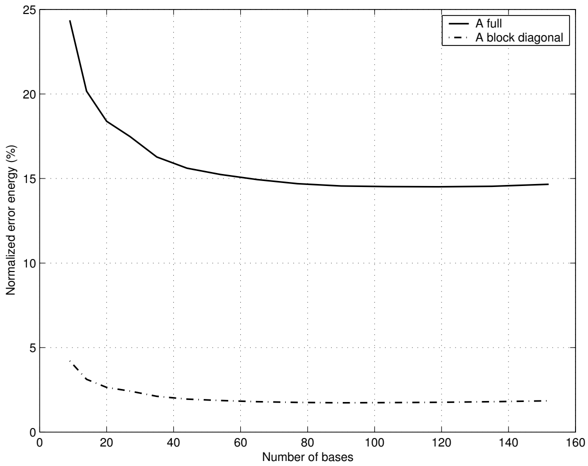


Fig. 3. In solid line the 2-step prediction error obtained using the method described in [11]. In dashed-dotted line those obtained with the system described in Section IV. The error is normalized as in (15). The characteristics of the turbulence layers are those reported in Table I.

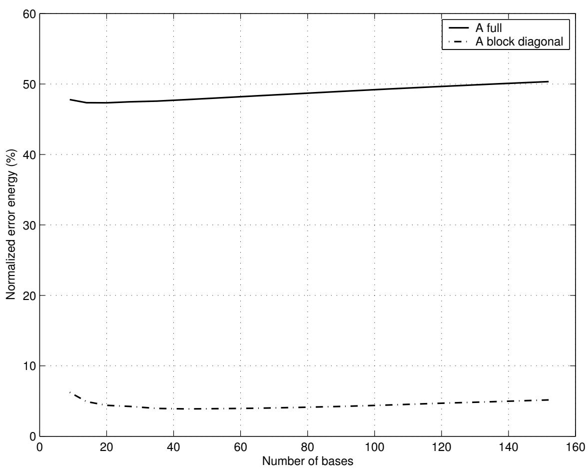


Fig. 4. In solid line the 2-step prediction error obtained using the method described in [11]. In dashed-dotted line those obtained with the system described in Section IV. The error is normalized as in (15). The characteristics of the turbulence layers are those reported in Table II.

As shown in Figs. 3,4, the use of the information on the turbulence structure (as described in Section IV) allows the system to provide better results than those of the system of [11] in all our simulations.

VI. CONCLUSIONS

In this paper we have considered the problem of atmospheric turbulence prediction from the point of view of a telescope's adaptive optics system.

First, we have described a procedure to estimate the characteristics of the turbulence layers. The proposed algorithm is based on a MRF representation of the turbulent phase. The algorithm properly estimates energies and velocities of the layers from the temporal correlations of the spatial innovation of the turbulent phase (computed using the MRF model of the turbulence).

Then, we have proposed a linear dynamic system which, using the estimated information on the turbulence structure, provides a good temporal description of the turbulence. The prediction performances of this system has been tested in some simulations.

We stress the fact that this approach can be considered as the generalization of those proposed in [11] to the case where the 3D structure of the turbulence is estimated. This makes it particularly interesting for future applications in MCAO systems.

VII. 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, and for many valuable and enjoyable discussions on the subject.

REFERENCES

[1] F. Roddier, *Adaptive optics in astronomy*, Cambridge university press; 1999.

[2] B. Le Roux and J.M. Conan and C. Kulcsar and H.F. Raynaud and L.M. Mugnier and T. Fusco, Statistics of a Geometric Representation of Wavefront Distortion, *J. Optical Society of America A*, vol. 21(7), 2004, pp 1261-1276.

[3] B. Le Roux, *Commande optimale en optique adaptative classique et multiconjuguee*, PhD thesis, Université de Nice Sophia Antipolis; 2003.

[4] 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.

[5] D.L. Fried, Statistics of a Geometric Representation of Wavefront Distortion, *J. Opt. Soc. Am.*, vol. 55, 1965, pp 1427-1435.

[6] J.W. Woods, Two-dimensional discrete Markovian fields, *IEEE Trans. on Information Theory*, vol. 18(2), 1972, pp 232-240.

[7] F. Roddier, The effects of atmospheric turbulence in optical astronomy, *Progress in Optics*, vol. 19, 1981, pp 281-376.

[8] T. Soderstrom, *Discrete-time stochastic systems*, Spriger; 1994.

[9] A.Beghi, A.Cenedese and A.Masiero, "On the estimation of atmospheric turbulence layers", in *Proceedings of the 2008 IFAC World Congress*, Seoul, Korea, 2008. To appear.

[10] A.Beghi, A.Cenedese and A.Masiero, "A comparison between Zernike and PCA representation of atmospheric turbulence", *in Conference on Decision and Control*, New Orleans, USA, 2007.

[11] A.Beghi, A.Cenedese, A.Masiero, "Atmospheric turbulence prediction: a PCA approach", Conference on Decision and Control 2007, New Orleans, USA, December 2007.