2009 IEEE International Conference on Control and Automation
Christchurch, New Zealand, December 9-11, 2009

System theoretic tools in Adaptive Optics

Alessandro Beghi, Angelo Cenedese and Andrea Masiero

Abstract— **Adaptive Optics (AO) systems provide a real challenge to the control engineer in many respects, the foremost of which are scalability and computational complexity of the control algorithms. On the other hand, systems theoretic tools can be applied to look at several problems under new perspectives. In this paper, we review a recent stochastic realization based method for turbulence simulation. Then, we investigate the estimation of the turbulence structure (i.e. the characteristics of its layers) through the use of a Markov Random Field (MRF) representation. Finally, we present a subspace algorithm for the identification of a dynamic model of the turbulence. The proposed method exploits the previously estimated turbulence characteristics to perform the first step of classical subspace identification procedures (Ho-Kalman's algorithm).**

I. Introduction

Mainly due to the large system dimension and to the real-time constraint on the control algorithms, AO systems are a really challenging application in control engineering. On the other hand, systems and control theory offers a number of well-established tools that can be used to analyze such problems using a methodological-based approach that may offer new solutions. Then, in this paper we present some applications of control theory on AO systems related problems: Some of the results have been discussed in previous works, whereas Section V reports a new approach to the problem of modeling the turbulence temporal evolution.

The aim of AO systems [1], is that of compensating the atmospheric turbulence effect (i.e. reducing, and possibly suppressing, the phase delays due to changes of atmospheric refraction index) properly controlling a set of deformable mirrors. Unfortunately, things are made complicated by several factors, for instance, just to cite some, only quite noisy measurements are available, the system dimension (i.e. the number of phase sensors) is quite large and typically has to be controlled at quite high sampling frequencies.

Since often real data (from the telescope) are not available, simulating the turbulence is of particular interest. Thus, in Section III we first review a recently proposed method for generating new samples of the turbulence (also called turbulent phase) [5]. The method is based on a stochastic realization approach, which will be similarly used also in Section V. The main advantage of this algorithm with respect to standard FFT based methods is that it easily allows for long sequence simulations.

From a temporal point of view, the turbulence is assumed to be formed by a set of layers moving independently on each other. Then, in Section IV we consider the problem of estimating the characteristics of these turbulent phase layers (e.g. their number, their energies and velocities). Such estimation is done taking advantage of a MRF spatial representation of the turbulent phase. The MRF representation allow us to derive a *"spatially whitening filter"*, which is used to compute a spatially "almost white" process e. Provided that the turbulence is moving with constant velocity, because of this almost whiteness of e the presence of the layers is quite apparent in the spatio-temporal correlations of e. The presented algorithm is a slight modification of that presented in [6].

Since the control is commonly delayed of two sample periods (the time needed for image acquisition and phase measurement [2], [3]) a number of prediction based methods have been proposed to improve the control performances [2], [3], [4], [6], [7]. In Section V we consider a model for the temporal evolution of the turbulent phase: We assume to have some information about the turbulence (e.g. estimated similarly to Section IV) which we use to compute the second order statistical description of the system. Then, similarly to Section III, we identify the model parameters using a stochastic realization approach. The resulting algorithm acts similarly to the Ho-Kalman's algorithm, but it computes the system temporal covariances from the statistical description of the turbulence. Thus, in practice the physical model presented in Section II acts like a regularization in the identification procedure. The main advantage of this algorithm with respect to data-based subspace identification algorithms is that it does not require to handle large amounts of data (which in this application may be a big issue).

II. Turbulence physical model

Let $\mathbf{u}$ and $\mathbf{v}$ be two unit vectors indicating two orthogonal spatial directions, as in Fig. 1, 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, $\phi(u, v, t)$

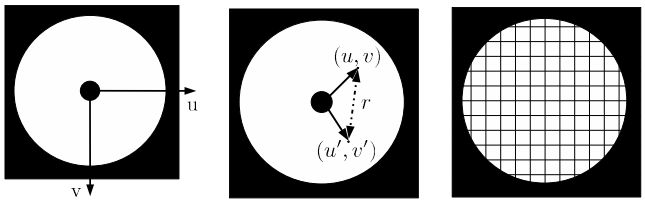
This work forms part of the ELT Design Study and is supported by the European Commission, within Framework Programme 6, contract No 011863.

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

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 $r = \sqrt{(u-u')^2+(v-v')^2}$.



(a) (b) (c)

Fig. 1. (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. (c) Discrete domain $\mathbb{L}$.

Astronomers usually describe the spatial statistical characteristics of the turbulent phase ϕ by means of the structure function, which measures the averaged difference between the phase at two points, (u,v) and (u',v'), of the wavefront separated by a distance r on the aperture plane (Fig. 1),

$$D_\phi(r) = \left\langle |\phi(u,v,t) - \phi(u',v',t)|^2 \right\rangle .$$

The structure function D_ϕ is related to the covariance function $C_\phi(r)$ as:

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

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 (see [10]):

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

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

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

Furthermore, the turbulent phase is supposed to be normally distributed [8], hence the second order statistics are sufficient to completely describe its statistical properties.

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 $h_l \geq \cdots \geq h_i \geq h_{i-1} \geq \cdots \geq h_0 = 0$. Let $\psi_i(u,v,t)$ be the value of the i^{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:

$$\phi(u,v,t) = \sum_{i=1}^{l} \gamma_i \psi_i(u,v,t) \ , \tag{4}$$

where γ_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 $C_{\psi_i}(r) = C_\phi(r)$. 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 [1]), thus

$$\psi_i(u,v,t+kT_s) = \psi_i(u - v_{i,u}kT_s, v - v_{i,v}kT_s, t) \ , \tag{5}$$

$i = 1,\ldots,l$, where $v_i = v_{i,u}\mathbf{u} + v_{i,v}\mathbf{v}$, and kT_s is a delay multiple of the sampling period T_s. The velocity vectors are assumed to be different for different layers, i.e. $v_i \neq v_j$ if $i \neq j$.

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. 1(c), i.e. a sensor is placed at each node of the grid. Then, measurements are taken using a Shack-Hartmann device (which introduces also some noise [1]), and projected on a set of spatial bases (which in this paper we assume to be the Zernike polynomials, as typically chosen by astronomers) both for having a compact turbulence representation and some de-noising. Finally, we call $y(t)$ the vector containing the measured phases at time t on the telescope aperture domain $\mathbb{L}$.

We refer to [1] for a detailed description of adaptive optics systems.

III. Atmospheric turbulence simulation

Standard methods for turbulence simulation are based on the use of Fast Fourier Transform (FFT) to generate samples with reliable statistics. However, such methods generate all the turbulence samples at the same time, thus limiting the maximum number of temporal samples to those allowed by the system memory size, thus making this almost impracticable for long simulations.

Furthermore, the turbulence characteristics typically (slowly) change over time, thus it is worth to investigate the possibility of simulating the performances of AO systems even in different conditions during the same simulation.

A different approach was first proposed in [16]: An elementary dynamic system was considered for generating new turbulent phase samples. This allowed the generation for long sequence simulations (only the current state and the system matrices are necessary to generate the new samples), however it exhibits problems in accurately reproducing the theoretical statistics of the turbulent phase.

Motivated by the above considerations, we proposed a different dynamic model to simulate the turbulence, where the matrices of the model where computed by stochastic realization [5].

Without loss of generality, let the turbulence be formed by one layer which moves along the $\mathbf{u}$ direction: Since the layers are independent, the generalization to l layers can be obtained by simulating them separately and combining them using (4). Then, generating new turbulent phase samples is equivalent to properly generating new columns of the phase screen.

Let $y(t)$ be a vector containing the values of the turbulent phase on a column of the phase screen (e.g. using Matlab notation: $y(t) = \phi(:, t, 1)$ when $t \leq m$). Then, we represent $y(t)$ as the output of the following linear dynamical system in state space form:

$$\begin{cases} x(t+1) = Ax(t) + Ke(t) \\ y(t) = Cx(t) + e(t) \end{cases} \tag{6}$$

where $e(t)$ is a zero mean white noise process with covariance matrix $\Sigma_e = \mathbf{E}\left[e(t)e(t)^T\right] = R \in \mathbb{R}^{m\times m}$. In Eq. (6), the state x and the output y vectors have dimensions respectively n and m, and $A \in \mathbb{R}^{n\times n}$, $K \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{m\times n}$.

Then, pose the computation of $\{A, C, K, R\}$ as a *(partial) stochastic realization problem*: Let Λ_i be the expected value of the product between two output samples $y(t+i)$ and $y(t)$, $\Lambda_i = \mathbf{E}\left[y(t+i)y(t)^T\right]$, $i = 0, \cdots, 2\nu - 1$, where ν is a design parameter in the procedure.

Furthermore, considering $y(t)$ as the output of (6) $\{\Lambda_i\}$ can be written as follows:

$$\begin{cases} \Lambda_1 = CG \\ \Lambda_2 = CAG \\ \vdots \\ \Lambda_{2\nu-1} = CA^{2\nu-2}G \end{cases}$$

where $G = A\Sigma C^T + KR$, and $\Sigma = \mathbf{E}\left[x(t)x(t)^T\right]$.

Exploiting the Taylor approximation it is possible to compute $\{\Lambda_i\}$ from the values of covariance function Eq. (3). Letting η be the distance traveled in a sample period (proportional to the translation velocity), then Λ_i can be obtained as follows:

$$\Lambda_i(h,k) = \mathbf{E}\left[y_h(t+i)y_k(t)\right] = C_\phi\left(\sqrt{(i\eta)^2 + (h-k)^2 p_s^2}\right)$$

where $\Lambda_i(h,k)$ is the element at position (h,k) in Λ_i, $y_k(t)$ is the value of the kth element in $y(t)$, and p_s is the spatial separation between two points in $\mathbb{L}$ along the $\mathbf{v}$ direction.

Then, the the stochastic realization algorithm proceeds as a two steps procedure:

1) Compute $\{\hat{A}, \hat{C}, \hat{G}\}$ in such a way that

$$\begin{cases} \bar{\Lambda}_1 = \hat{C}\hat{G} \approx \Lambda_1 \\ \bar{\Lambda}_2 = \hat{C}\hat{A}\hat{G} \approx \Lambda_2 \\ \vdots \\ \bar{\Lambda}_{2\nu-1} = \hat{C}\hat{A}^{2\nu-2}\hat{G} \approx \Lambda_{2\nu-1} \end{cases}$$

with the constraint that: $\{\bar{\Lambda}_i = \hat{C}\hat{A}^{i-1}\hat{G}\} > 0,\ \forall i \geq 1$, i.e. $\{\bar{\Lambda}_i\}$ has to be a covariance sequence.

2) Compute $\hat{K}, \hat{R}$ from the solution of the previous step.

To compute $\{\hat{A}, \hat{C}, \hat{G}\}$, first we construct the following Hankel matrix:

$$\begin{aligned} H &= \begin{bmatrix} \Lambda_1 & \Lambda_2 & \cdots & \Lambda_\nu \\ \Lambda_2 & \Lambda_3 & \cdots & \Lambda_{\nu+1} \\ \vdots & \vdots & \ddots & \vdots \\ \Lambda_\nu & \Lambda_{\nu+1} & \cdots & \Lambda_{2\nu-1} \end{bmatrix} \\ &= \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{\nu-1} \end{bmatrix} \begin{bmatrix} G & AG & \cdots & A^{\nu-1}G \end{bmatrix}. \end{aligned}$$

Notice that the equation $\Lambda_i = CA^{i-1}G$ stands only if the signal is really generated by a linear dynamic model, hence, when working with real data, the above expression for the Hankel matrix should be considered as an approximation. Anyway, from such expression it is quite apparent that it is possible to compute $\{\hat{A}, \hat{C}, \hat{G}\}$ from a factorization of H (e.g. obtained using Singular Value Decomposition (SVD)). However, to ensure the positivity condition we introduce some intermediate steps.

In particular, define the normalized Hankel matrix as follows:

$$\hat{H} = L^{-1}HL^{-T},$$

where L is a Cholesky factor of the following Toeplitz matrix T,

$$T = \begin{bmatrix} \Lambda_0 & \Lambda_1 & \Lambda_2 & \cdots & \Lambda_{\nu-1} \\ \Lambda_1^T & \Lambda_0 & \Lambda_1 & \ddots & \Lambda_{\nu-2} \\ \Lambda_2^T & \Lambda_1^T & \Lambda_0 & \ddots & \Lambda_{\nu-3} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \Lambda_{\nu-1}^T & \Lambda_{\nu-2}^T & \Lambda_{\nu-3}^T & \cdots & \Lambda_0 \end{bmatrix}.$$

Then, $\hat{H}$ can be factorized according to the SVD algorithm:

$$\hat{H} = USV^T = US^{1/2}S^{1/2}V^T,$$

with U, V unitary matrices, and S diagonal matrix whose elements are the singular values of $\hat{H}$: Since most of the singular values of $\hat{H}$ will be close to zero we consider only the first $\bar{n}$ and set the remaining ones to 0:

$$\hat{H} \approx U_{\bar{n}}S_{\bar{n}}V_{\bar{n}}^T = U_{\bar{n}}S_{\bar{n}}^{1/2}S_{\bar{n}}^{1/2}V_{\bar{n}}^T.$$

Thus $\hat{C}$ and $\hat{G}$ can be computed as follows:

$$\begin{cases} \hat{C} = \rho_1(H)L^{-T}V_{\bar{n}}S_{\bar{n}}^{-1/2} \\ \hat{G} = (\rho_1(H^T)L^{-T}U_{\bar{n}}S_{\bar{n}}^{-1/2})^T \end{cases}$$

where the $\rho_1(\cdot)$ operator selects the first m rows of a matrix.

The computation of $\hat{A}$ can be done similarly by using a shifted version of H:

$$\sigma(H) = \begin{bmatrix} \Lambda_2 & \Lambda_3 & \cdots & \Lambda_{\nu+1} \\ \Lambda_3 & \Lambda_4 & \cdots & \Lambda_{\nu+2} \\ \vdots & \vdots & \ddots & \vdots \\ \Lambda_{\nu+1} & \Lambda_{\nu+2} & \cdots & \Lambda_{2\nu} \end{bmatrix},$$

then,

$$\hat{A} = S_{\bar{n}}^{-1/2} U_{\bar{n}}^T L^{-1} \sigma(H) L^{-T} V_{\bar{n}} S_{\bar{n}}^{-1/2} \ .$$

To compute K, the Kalman gain, the knowledge of the steady-state state covariance Σ is needed. It can be shown that Σ solves the Algebraic Riccati Equation (ARE):

$$\Sigma = \hat{A}\Sigma\hat{A}^T + (\hat{G} - \hat{A}\Sigma\hat{C}^T)(\Lambda_0 - \hat{C}\Sigma\hat{C}^T)^{-1}(\hat{G}^T - \hat{C}\Sigma\hat{A}^T).$$

Finally, $\hat{R} = \Lambda_0 - \hat{C}\Sigma\hat{C}^T$, and the Kalman gain is: $\hat{K} = (\hat{G} - \hat{A}\Sigma\hat{C}^T)\hat{R}^{-1}$.

If ν is sufficiently large, it is possible to prove that computing the parameters $\hat{A}, \hat{C}, \hat{R}, \hat{K}$ obtained as described above, then $\left\{\bar{\Lambda}_0, \bar{\Lambda}_1, \bar{\Lambda}_2, \dots\right\}$ is a positive sequence [17].

Fig.2 shows a comparison between the stochastic realization approach and the method of Assémat et al. on reproducing the correct theoretical statistics of the turbulent phase. For both the methods it has been computed the asymptotic structure function exploiting the invariant density of the dynamical system (6).

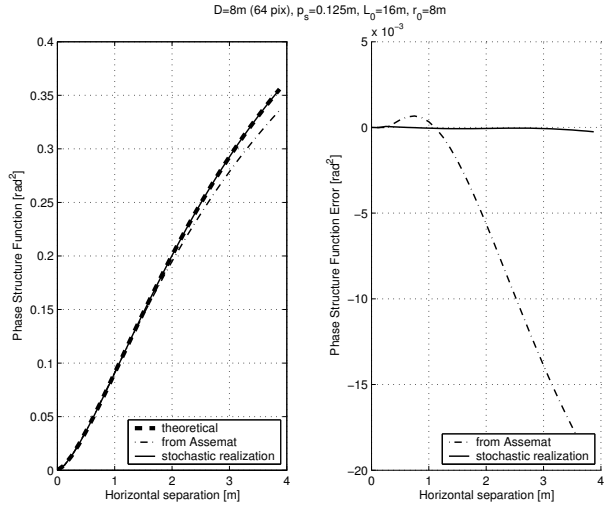


Fig. 2. Phase structure function along the wind direction. A comparison of the theoretical values (dashed line) and those obtained with: (i) the dynamical model computed via stochastic realization (solid line) (ii) the method of Assémat et al. (dash-dotted line). The values of the parameters are set to $L_0 = 20$m, $r_0 = 8$m, $D = 16$m, $p_s = 0.25$m.

IV. Estimation of the atmospheric structure from phase measurements

The knowledge of atmospheric turbulence characteristics plays an important role in the design of the AO system control algorithm, thus it is of fundamental importance. In this section we assume to already have estimates, $\{\hat{L}_0, \hat{r}_0\}$, of the outer-scale and of the Fried parameter and we consider the problem of estimating the characteristics of the layers: $(l, \gamma_1, \dots, \gamma_l, v_{1,u}, \dots, v_{l,u}, v_{1,v}, \dots, v_{l,v})$. The resulting procedure is a modification of that presented in [6].

For simplicity of exposition, in this Section we will assume to use a perfect sensor, i.e. it provides measurements of the turbulent phase without superimposed noise.

First, let us consider the spatio-temporal correlation,

$$c_{\psi,i}(u - \bar{u}, v - \bar{v}, \bar{k}) = \mathbf{E}[\psi_i(u, v, t + \bar{k}T_s)\psi_i(\bar{u}, \bar{v}, t)] \ ,$$

as a function of u and v (we consider $\bar{u}$, $\bar{v}$ and $\bar{k}$ as fixed to constant values). By the Taylor assumption, (5), the layers translate over the telescope aperture with constant velocities, thus $c_{\psi,i}(u - \bar{u}, v - \bar{v}, \bar{k})$ has a peak in correspondence of $u = \bar{u} + v_{i,u}\bar{k}T_s$, $v = \bar{v} + v_{i,v}\bar{k}T_s$.

Since actually the available spatio-temporal correlations are those of ϕ, $c_\phi(u - \bar{u}, v - \bar{v}, \bar{k}) = \mathbf{E}[\phi(u, v, t + \bar{k}T_s)\phi(\bar{u}, \bar{v}, t)]$, and $c_\phi(u, v, k) = \sum_{i=1}^{l} \gamma_i^2 c_{\psi,i}(u, v, k)$, $\forall(u, v, k)$, then the intuitive idea is that of searching for peaks in $c_\phi(u - \bar{u}, v - \bar{v}, \bar{k})$, $k = 1, \dots, \bar{T}$, which should correspond to translating layers.

However, the covariance (3) vanishes not so quickly, hence the peak founded in $c_{\psi,i}(v_{i,u}\bar{k}T_s, v_{i,v}\bar{k}T_s, \bar{k})$ is not so well marked in $c_\phi(v_{i,u}\bar{k}T_s, v_{i,v}\bar{k}T_s, \bar{k})$: In fact, due to noise, finite number of samples used in the estimation of covariances and the combination of elements, $\{c_{\psi,i}(\cdot)\}$, associated to different layers, the peaks may be wrongly detected or not founded at all in $c_\phi(\cdot)$.

To reduce the effect of these unavoidable bad factors, we maintain the idea of looking at spatio-temporal correlations, but we take advantage of a MRF spatial representation of the turbulent phase.

In accordance with the physical model of the turbulence of Section II, we model the turbulent phase ϕ as a homogeneous and isotropic MRF with circular neighborhood $\mathbb{N}(\cdot)$ of radius $\bar{d}$. Then, as shown in [14], $\phi(\bar{u}, \bar{v}, t)$ can be expressed as the best linear prediction of $\phi(\bar{u}, \bar{v}, t)$ based on the values of its neighbors $\mathbb{N}(\bar{u}, \bar{v})$ at time t plus an "innovation" process $e(\bar{u}, \bar{v}, t)$. According with the normal distribution of ϕ, the best (spatial) linear prediction operator $\hat{\mathbf{E}}[\cdot]$ corresponds to the expectation operator $\mathbf{E}[\cdot]$, that is

$$\phi(\bar{u}, \bar{v}, t) = \sum_{(u,v)\in\mathbb{N}(\bar{u},\bar{v})} a_{|(\bar{u}-u,\bar{v}-v)|}\phi(u, v, t) + e(\bar{u}, \bar{v}, t) \ , \tag{7}$$

where $\{a_i\}$ are the coefficients which yield the best (spatial) linear prediction of $\phi(\bar{u}, \bar{v}, t)$ given the values of its neighbors. Furthermore,

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

The terms "spatial innovation" and "spatially almost white" for e are motivated by (8).

The coefficients $\{a_i\}$ can be computed from the covariance values, (3), as those of the usual best linear predictor [15]. In practice, the exact covariance values are not available. Nevertheless, it is possible to approximate $\{a_i\}$ with $\{\hat{a}_i\}$, the values computed using $\{\hat{L}_0, \hat{r}_0\}$ instead of $\{L_0, r_0\}$. Then, as long as the estimates $\{\hat{L}_0, \hat{r}_0\}$ are quite good approximations of $\{L_0, r_0\}$, the process $e_{\hat{L}_0,\hat{r}_0}$ (obtained from (7) using $\{\hat{a}_i\}$ instead of $\{a_i\}$) approximatively satisfy (8) and results to be useful for the detection of moving layers.

The considerations made for ϕ and $c_\phi(\cdot)$ approximatively hold also for $e_{\hat{L}_0,\hat{r}_0}$ and $c_{\hat{L}_0,\hat{r}_0}(\cdot)$, where

$$c_{\hat{L}_0,\hat{r}_0}(u, v, k) = \mathbf{E}[e_{\hat{L}_0,\hat{r}_0}(u+\bar{u}, v+\bar{v}, t+kT_s)e_{\hat{L}_0,\hat{r}_0}(\bar{u}, \bar{v}, t)].$$

Thus the following observations are now in order:

- $c_{\hat{L}_0,\hat{r}_0}(\cdot)$ has peaks in locations corresponding to the translating layers.
- Thanks to the "almost whiteness" of $e_{\hat{L}_0,\hat{r}_0}$, the peaks are much more apparent on $c(\cdot)$ than on $c_\phi(\cdot)$.
- If a peak in $c_{\hat{L}_0,\hat{r}_0}(u,v,\bar{k})$ corresponds to a real layer, then it will be visible also in $c_{\hat{L}_0,\hat{r}_0}(uk/\bar{k}, vk/\bar{k}, k)$ for $k < \bar{k}$.
- Thus,to reduce the influence of noise, we consider also $c^p_{\hat{L}_0,\hat{r}_0}(u,v,\bar{k})$, the integrated version of $c_{\hat{L}_0,\hat{r}_0}(\cdot)$ over k (from $k=1$ to $\bar{k}-1$).

Then, the algorithm iteratively searches for new layers starting from $k=\bar{k}$ to $k=1$: A new layer is detected at step k if:

1) $c_{\hat{L}_0,\hat{r}_0}(u,v,k)$ corresponds to a large peak and $c^p_{\hat{L}_0,\hat{r}_0}(u,v,k)$ is large.
2) $c_{\hat{L}_0,\hat{r}_0}(u,v,k)$ is not close to already detected layers (e.g. disjoint neighborhoods).
3) $(u,v) \in \mathbb{L}_c$ and, for $k<\bar{k}$, $(u(k+1)/k, v(k+1)/k) \notin \mathbb{L}_c$.

In Table I and II we report the detection results ($\{v_{i,u}, v_{i,v}, \gamma_i^2\}$ are the true values of the parameters for each layer, while $\{\hat{v}_{i,u}, \hat{v}_{i,v}, \hat{\gamma}_i^2\}$ are the estimated ones) for two possible choices for the parameters:

(A) $d=8$, $L_0=50$m, $r_0=0.4$m, measurement noise variance $\sigma_m^2 = 0.6\text{rad}^2$, three layers with the characteristics reported in the table, $\hat{L}_0=15.65$m, $\hat{r}_0=0.25$m.

(B) $d=8$, $L_0=22$m, $r_0=0.2$m, noise variance $\sigma_m^2 = 0.6\text{rad}^2$, three layers with the characteristics reported in the table, $\hat{L}_0=14.6$m, $\hat{r}_0=0.16$m.

These results have been obtained using 1000 turbulence samples at a sampling frequency of 1 Hz.

TABLE I

DETECTION OF THE LAYERS.

	1^{st} layer	2^{nd} layer	3^{rd} layer
$v_{i,u}$ [m/s]	7.00	−16.00	30.00
$\hat{v}_{i,u}$ [m/s]	7.05	−16.11	30.00
$v_{i,v}$ [m/s]	0	0	0
$\hat{v}_{i,v}$ [m/s]	0	0	0
γ_i^2	0.50	0.30	0.20
$\hat{\gamma}_i^2$	0.54	0.26	0.20

TABLE II

DETECTION OF THE LAYERS.

	1^{st} layer	2^{nd} layer	3^{rd} layer
$v_{i,u}$ [m/s]	5.00	9.00	24.00
$\hat{v}_{i,u}$ [m/s]	5.00	9.06	24.17
$v_{i,v}$ [m/s]	0	0	0
$\hat{v}_{i,v}$ [m/s]	0	0	0
γ_i^2	0.55	0.35	0.10
$\hat{\gamma}_i^2$	0.48	0.37	0.15

V. SPATIO-TEMPORAL REPRESENTATIONS FOR PREDICTION

The choice of a proper spatio-temporal representation of the turbulence plays a fundamental role in determining the overall AO system performances.

In this Section we present a subspace approach for computing such representation. In particular, with a slight abuse of notation, we will consider the following dynamic system:

$$\begin{cases} x(t+1) = Ax(t) + Ke(t) \\ y(t) = Cx(t) + e(t) \end{cases} \tag{9}$$

where $y(t)$ is a modal representation of the turbulent phase values over the telescope aperture at time t. In fact, considering $y(t)$ instead of the turbulence values allow us to reduce the dimensionality of the output. Hereafter we will assume that such a dimensionality reduction step is obtained considering a Principal Component Analysis (PCA) representation of the signal [18], where the second order statistics of the turbulent phase are computed using (3) and the estimates $\{\hat{L}_0, \hat{r}_0\}$ of the outer-scale and of the Fried parameter. Nevertheless, the following identification algorithm can be used also with whatever other choice of the basis.

Recent works provided proofs of the asymptotic optimality of the state of the art procedures for subspace system identification, [19]. Hence it may be worth to consider a subspace identification method for computing the parameters in (9). However, in such procedures it is not exploited any a priori information about the system to regularize the identification results (they are completely data-based). Furthermore, some computational difficulties may arise when handling large systems (both of computational efficiency and memory requirement).

Motivated by these observations we propose an alternative method which exploits some already estimated information about the turbulence (e.g. the turbulent phase structure estimated in Section IV). The idea is to use the statistical model of the turbulence, (3), (4), (5), and the physical parameters estimated in Section IV to compute the second order spatio-temporal statistical characteristics of the turbulent phase. Then these second order characterization is introduced in a subspace identification algorithm to compute the parameters in (9). Thus, in practice the physical model presented in Section II acts like a regularization in the identification procedure.

Let U be the matrix composed by the first m components (those with larger energies) of the PCA spatial representation. Let $U^\dagger$ be the pseudo-inverse of U and $\phi(t)$ be the vector containing the turbulent phase values at time t. Then $\phi(t) \approx Uy(t)$. Furthermore, redefine Λ_i as $\Lambda_i = \mathbf{E}\left[y(t+i)y(t)^T\right]$. Then the proposed method can be summarized as follows:

- Estimate the turbulent phase characteristics, e.g. the layers structure (Section IV) and $\{L_0, r_0\}$ (for instance as described in [20]).
- Compute $\Lambda_i = U^\dagger \mathbf{E}\left[\phi(t+i)\phi(t)^T\right](U^\dagger)^T$, $i = 1, \cdots, 2\nu-1$, where $\mathbf{E}\left[\phi(t+i)\phi(t)^T\right]$ can be easily computed from (3), (4), (5) and ν is a design parameter.

- Compute $\Lambda_0 = U^\dagger \mathbf{E}\left[\phi(t)\phi(t)^T\right](U^\dagger)^T + R_w$, where R_w is the noise variance matrix projected on the spatial basis U (in fact it is necessary to take into account also the sensor model [7]).
- Solve the partial stochastic realization problem similarly to Section III. This step provides us with the desired estimates of the dynamic model parameters.

Notice that this method is very similar to the early subspace identification algorithms ([21],[22],[23],[24]), where the first step (computing the system covariances) here is performed taking advantage of the physical statistical characterization of the turbulence.

The main advantage of this procedure is that it is not necessary to handle large amounts of turbulent phase temporal samples. In fact phase samples are necessary only to estimate the turbulent phase structure: In our examples of Section IV we used 1000 samples to estimate the layers characteristics, while much more thousands should be used to obtain reliable results in completely data-based subspace identification algorithms.

On the other hand, the effectiveness of this approach depends on the reliability of the statistical model presented in Section II and of the estimates of the turbulence layers characteristics.

In table III we report the AO system closed-loop performances (Strehl ratios, [1],[7], and mean square values of the residual error after the correction obtained with the control) obtained using the identified model to compute the deformable mirror control in a prediction based approach ([2],[7]). The considered cases are those of Section IV, and, specifically, we exploited the estimated parameters of Table I,II. The results are obtained setting the sampling frequency of the system to 1 KHz. Furthermore, since usually quite good estimates $\{\hat{L}_0, \hat{r}_0\}$ are available, here we set them to their correct values.

TABLE III

CLOSED-LOOP PERFORMANCES.

Simulation	m	Strehl ratio (%)	Mean square error [rad^2]
case (A)	500	71.11	0.3410
case (B)	500	54.23	0.6120

VI. CONCLUSIONS

In this paper we have presented the use of some system theoretic tools in AO systems: First, we have summarized a recently proposed method based on stochastic realization to simulate the turbulent phase.

Then, exploiting a MRF representation of the turbulent phase we have proposed a procedure to estimate the turbulent phase structure, i.e. the number of layers and their characterizing parameters.

Finally, taking advantage of the previously estimated turbulent phase characteristics, we have proposed a stochastic realization based method for modeling the temporal dynamic of the turbulence. This may be attractive with respect to other subspace identification algorithms because it provides quite good results using not so large amounts of data.

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 univ. 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, Univ. de Nice Sophia Antipolis; 2003.
[4] C. Petit, *Etude de la commande optimale en OA et OAMC, validation numérique et expérimentale*, PhD thesis, Université Paris 13; 2006.
[5] A. Beghi, A. Cenedese, A. Masiero, "Stochastic realization appoach to the efficient simulation of phase screens", Journal of the Optical Society of America A, **25**, No. 2, 515–525 (2008).
[6] A. Beghi, A. Cenedese, A. Masiero, "A Markov-Random-Field-based approach to modeling and prediction of atmospheric turbulence", in *Proceedings of the 16th Mediterranean Conference on Control and Automation*, pp. 1735-1740, Ajaccio, Corsica, 2008.
[7] A.Beghi, A.Cenedese, F.Maran, A.Masiero, "A comparison of Kalman filter based algorithms for turbulent phase control in an adaptive optics system", Conference on Decision and Control 2008, pp. 1839-1844, Cancun, Mexico, 2008.
[8] F. Roddier, The effects of atmospheric turbulence in optical astronomy, *Progress in Optics*, vol. 19, 1981, pp 281-376.
[9] D.L. Fried, Statistics of a Geometric Representation of Wavefront Distortion, *J. Opt. Soc. Am.*, vol. 55, 1965, pp 1427-1435.
[10] 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.
[11] Yu.A. Rozanov, *Stationary random processes*, Holden-Day; 1967.
[12] A.V. Ivanov and N.N. Leonenko, *Statistical analysis of random fields*, Kluwer academic publishers; 1989.
[13] R. Durrett, *Probability: theory and examples*, Duxbury press; 1996.
[14] J.W. Woods, Two-dimensional discrete Markovian fields, *IEEE Trans. on Information Theory*, vol. 18(2), 1972, pp 232-240.
[15] T. Soderstrom, *Discrete-time stochastic systems*, Spriger; 1994.
[16] F.Assémat, R.W.Wilson, and E.Gendron, "Method for simulating infinitely long and non stationary phase screens with optimized memory storage," Opt. Express **14**, No. 3, 988–999 (2006).
[17] A.Lindquist, and G.Picci, "Canonical correlation analysis, approximate covariance extension, and identification of stationary time series," Automatica **32**(5), 709–733 (1996).
[18] A.Beghi, A.Cenedese and A.Masiero, "A comparison between Zernike and PCA representation of atmospheric turbulence", *in Conference on Decision and Control*, pp. 561-566, New Orleans, USA, 2007.
[19] A.Chiuso, "On the relation between CCA and predictor-based subspace identification", IEEE Trans. on Automatic Control **52**, No. 10, 1795-1812, (2007).
[20] A.Beghi, A.Cenedese and A.Masiero, "On the estimation of atmospheric turbulence statistical characteristics", *MSC 2009 (to appear)*.
[21] M.Aoki, *State Space Modeling of Time Series*, 2nd edition, Springer-Verlag, 1991.
[22] H.P.Zeiger, and A.J.McEwen, "Approximate linear realization of given dimension via Ho's algorithm", IEEE Trans. Automatic Control. AC-19, p. 153, 1974.
[23] B.L.Ho and R.E.Kalman, "Effective construction of linear state-variable models from input/output functions", Regelungstechnik, vol. 14, pp.545–548, 1966.
[24] S.Y.Kung, "A new identification and model reduction algorithm via singular value decomposition". In Proceedings of the 12th Asilomar Conference on Circuits, Systems and Computers, Pacific Grove, CA, USA, pp. 705–714, 1978.