18th IEEE International Conference on Control Applications
Part of 2009 IEEE Multi-conference on Systems and Control
Saint Petersburg, Russia, July 8-10, 2009

On the estimation of atmospheric turbulence statistical characteristics

Alessandro Beghi, Angelo Cenedese and Andrea Masiero

***Abstract*— Recent works on adaptive optics (AO) systems have shown that the knowledge of static and dynamic characteristics of the atmospheric turbulence can be exploited to improve the performances of the AO control system. Then, motivated by the importance of an effective control for compensating the atmospheric turbulence effect and thus improving the real resolution of ground-based telescopes, in this paper we investigate the estimation of both the static (i.e. purely spatial) and dynamic characteristics of the turbulence. According with commonly accepted statistical models, spatial second order statistics of the turbulence are characterized by two parameters, namely the outer-scale and the Fried parameter: We propose a data-based estimation procedure for these parameters. Then, from a dynamical point of view, the turbulence is typically assumed to be formed by a discrete set of layers moving over the telescope aperture: We show how a Markov Random Field (MRF) representation, based on the already computed spatial parameters, allow us to estimate the number of layers and their characteristics.**

I. INTRODUCTION

It is well known that the achievable resolution of any telescope theoretically depends on the telescope aperture diameter d: Indeed, because of the diffraction effect, the larger is d, the higher is the telescope theoretical resolution.

However, the real resolution of ground-based telescopes is significantly reduced (with respect to the theoretical one) by the presence of the atmospheric turbulence: 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, which, due to the dynamic evolution of the atmospheric turbulence, may change quite fast both spatially and temporally. Then, the resolution reduction is caused by the fact that flat wavefronts before entering in the atmosphere, arrives corrugated to the ground. The phase delay of beams arriving on the telescope pupil is also called turbulent phase.

The aim of the adaptive optics (AO) system [1], is that of compensating the atmospheric turbulence effect (i.e. reducing, and possibly suppressing, the phase delays). To make this possible, the AO control unit receives measurements of the current turbulent phase by a wavefront sensor and commands a set of deformable mirrors (i.e. actuators) to adapt their shape so as to compensate for the current value of the phase delays. The control is commonly delayed of two sample periods, that is the time needed for image acquisition and phase measurement [2], [3].

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

Because of this delay, a number of Kalman filter based methods have been proposed to improve the control performances [2], [3], [4], [5], [6]. As shown in [6], actually the control performances can be improved if some (even rough) information about the turbulent phase characteristics is available. Then, the aim of this paper, is that of providing some estimates of the parameters characterizing the turbulence statistical model.

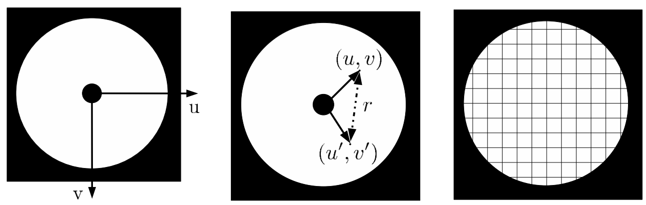
The atmospheric turbulence is commonly assumed to be statistically homogeneous and isotropic. In particular, it has been proved, [7], that its statistics follow a normal distribution. Furthermore, from a dynamic point of view, the turbulence is typically described as a set of layers moving with constant velocities and 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 procedure to estimate the parameters characterizing both spatially and dynamically the atmospheric turbulence: In Section III, we first compute the turbulent phase sample covariances, which in the following are used by a simple randomized algorithm to estimate the parameters characterizing the turbulent phase spatial statistics, e.g. the outer-scale, L_0, and the Fried parameter, r_0 (see [8]). Then, in Section IV, exploiting these estimated values, we compute a "spatially whitening" filter associated to a Markov Random Field (MRF) model of the turbulent phase. Finally, we propose a procedure to estimate the turbulence dynamical characteristics (e.g. energies and velocities of the layers) by means of the spatio-temporal correlations of the "innovation" process resulting from the application of the computed filter on the turbulent phase. This last step is a modified version of the algorithm already proposed in [5].

We conclude in Section V with showing the results of some simulations.

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, $\phi(u, v, t)$ 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 [9]):

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

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 [7], 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 $\phi(u, v, t) = \sum_{i=1}^{l} \gamma_i \psi_i(u, v, t)$, 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

$$\psi_i(u, v, t + kT_s) = \psi_i(u - v_{i,u}kT_s, v - v_{i,v}kT_s, t) , \tag{4}$$

$i = 1, \ldots, l$, 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. Estimation of the spatial parameters

As described in the previous section, the turbulent phase is assumed to be zero-mean and normally distributed. Then its spatial statistical description is characterized by two parameters, $\{L_0, r_0\}$, as shown by (3). The aim of these section is that of estimating these parameters from a set of phase measurements $\{y(1), \ldots, y(T)\}$, with $T \geq 1$. To achieve this task, we will also assume to have some a priori information about the parameter values: We assume to know a (possibly large) range of values which includes L_0, i.e. we assume to know $\{L_{\min}, L_{\max}\}$ such that $L_{\min} \leq L_0 \leq L_{\max}$. Instead, the only restriction imposed on r_0 is to be physically possible, i.e. $r_0 > 0$. Actually, these assumptions are quite realistic in practical applications.

Since the relation between the values of the turbulence and $\{L_0, r_0\}$ is statistically expressed by the structure function, then the estimation procedure is divided in two steps: First, computing the structure function estimates $\{\hat{D}(d_i)\}_{i=1:N}$ on a set of spatial distances $\{d_i\}_{i=1:N}$ from the measurements $\{y(1), \ldots, y(T)\}$. Then, estimating $\{\hat{L}_0, \hat{r}_0\}$ from $\{\hat{D}(d_i)\}$ and (2).

Notice that (1) provides us with a simple way to compute estimates of the structure function, i.e. computing the sample covariances and then obtaining the structure function estimates through (1).

Since ϕ is assumed to be normally distributed, stationary and asymptotically uncorrelated (the wind velocity is assumed to be different from zero and $C_\phi(r)$ vanishes as r becomes large) then it is also ergodic. Then, the ergodicity ensures us that the covariance sample estimates (and thus also the structure function estimates, $\{\hat{D}(d_i)\}_{i=1:N}$) will converge to their true values as T becomes large [10].

As a matter of fact, the convergence results to be quite slow for common choices of the parameter values and in standard Very Large Telescope (VLT) settings. Fig. 2 compares the true structure function values with the estimates obtained from perfectly measured data (no noise) and noisy data. The structure function estimates are computed using $T = 1000$ turbulent phase samples (sample frequency $f_s = 1$ Hz). The results are obtained setting the parameters to the values of case (A) in Section V. It is quite apparent that the values of estimates computed without noise are almost converged, while those computed from noisy data would need more samples to get to the same convergence results. Nevertheless, in Section V we will show that the results obtained using only 1000 samples are sufficient to make the procedure described in the next Section effective. As a consequence of the above considerations, results shown in Section V should get even better using more samples.

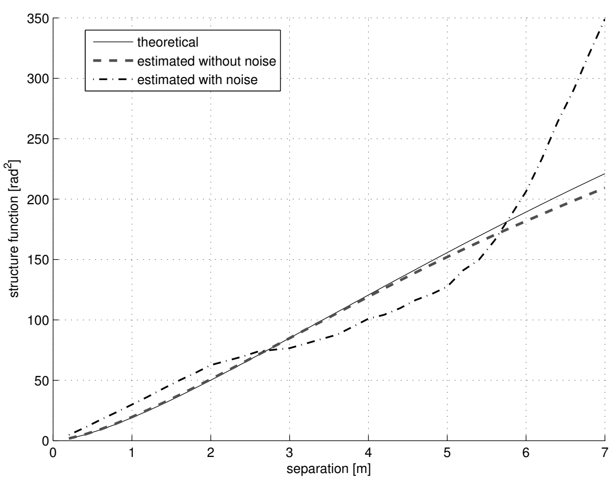


Fig. 2. In solid line the theoretical structure function, with $L_0 = 50$m and $r_0 = 0.4$m, as a function of space distance. In dashed line the structure function values estimated from noise-free samples. In dashed-dotted line the structure function values estimated from noisy SH measurements.

In Section V we will also consider the boundary case $T = 1$, basing the estimation on a purely spatial meaning procedure: In this case the consistency of the estimator, as the domain dimension goes to infinity, can be proved as in [11].

The aim of the second step is computing the estimates of the spatial parameters of the turbulence, $\{\hat{L}_0, \hat{r}_0\}$, from $\{\hat{D}(d_i)\}_{i=1:N}$. Unfortunately, as shown in (2), the relation between $\{\hat{D}(d_i)\}_{i=1:N}$ and $\{L_0, r_0\}$ is nonlinear, hence in general we cannot compute $\{\hat{L}_0, \hat{r}_0\}$ in closed form.

A convenient idea would be that of using a maximum likelihood estimator, however in this specific case its complexity would be quite high. Then, our objective is that of designing an alternative, possibly faster, estimation procedure.

We propose to compute $\{\hat{L}_0, \hat{r}_0\}$ as

$$\{\hat{L}_0, \hat{r}_0\} = \arg\min_{\tilde{L}_0, \tilde{r}_0} \left(\sum_{i=1}^{\bar{N}} T_i \left(\frac{\hat{D}(d_i)}{D_{\phi,\tilde{L}_0,\tilde{r}_0}(d_i)} - 1 \right)^2 \right) \quad (5)$$

where $\bar{N} \leq N$, T_i is the number of samples used to compute $\hat{D}(d_i)$ (the estimation is derived by a spatio-temporal meaning, hence T_i is different for different d_is) and $D_{\phi,\tilde{L}_0,\tilde{r}_0}(d_i)$ is the value at distance d_i of the structure function introducing $\{\tilde{L}_0, \tilde{r}_0\}$ in (2) instead of $\{L_0, r_0\}$ (e.g. $D_\phi(d_i) = D_{\phi,L_0,r_0}(d_i)$). In the following, the function minimized in (5) will be also referred to as a cost function, and its value for $\tilde{L}_0, \tilde{r}_0$ will be indicated as $c(\tilde{L}_0, \tilde{r}_0)$.

In the following we will try to give some intuition on the terms inside of the sum in (5). Consider the ith term: $\hat{D}(d_i)$ has been computed as the mean of T_i variables ω_t^i. If we assume that the underground process is characterized by $D_{\phi,\tilde{L}_0,\tilde{r}_0}(\cdot)$, then $\omega_t^i = (w_t^i)^2$ and $w_t^i \sim \mathcal{N}(0, D_{\phi,\tilde{L}_0,\tilde{r}_0}(d_i))$, $t = 1 \dots T_i$. Thus $\frac{\omega_t^i}{D_{\phi,\tilde{L}_0,\tilde{r}_0}(d_i)} \sim \chi_1^2$. Furthermore if $\{w_t^i\}_t$ are independent, then from the central limit theorem:

$$\frac{1}{\sqrt{2T_i}} \sum_{t=1}^{T_i} \left(\frac{\omega_t^i}{D_{\phi,\tilde{L}_0,\tilde{r}_0}(d_i)} - 1 \right) \sim \mathcal{N}(0,1) \ , \quad (6)$$

for large N_i. Rearranging the terms in (6) and writing $\hat{D}(d_i)$ as the mean of ω_t^i we obtain

$$\frac{\hat{D}(d_i) - D_{\phi,\tilde{L}_0,\tilde{r}_0}(d_i)}{\sqrt{2/T_i}\ D_{\phi,\tilde{L}_0,\tilde{r}_0}(d_i)} \sim \mathcal{N}(0,1) \ ,$$

thus

$$\hat{D}(d_i) \sim \mathcal{N}(D_{\phi,\tilde{L}_0,\tilde{r}_0}(d_i), \frac{2}{T_i} D^2_{\phi,\tilde{L}_0,\tilde{r}_0}(d_i)) \ . \quad (7)$$

Comparing (5) with (7) it is quite apparent that in (5) we are summing local log-likelihood functions (properly scaled).

Of course, actually $\{w_t^i\}_t$ are not independent. However, the central limit theorem can be formulated also for dependent variables, with a slight different expression [12]. Then, actually the cost function in (5) is a quickly computable approximation of the sum of the local log-likelihoods.

The minimization of the cost function in (5) can be computed for instance with standard gradient methods or with Markov Chain Monte Carlo (MCMC) methods. While the first occasionally end on local minima, the latter often are computationally quite expensive.

In this paper we use a simple randomized algorithm to estimate the location of the minimum. Notice that, given the value of L_0 the cost function in (5) is linear in r_0. Thus known the value of L_0, the optimal r_0 can be computed in closed form as solution of a standard least squares problem[1]. For this reason we drop "r_0" from the notation of the cost function, i.e. we indicate with $c(\tilde{L}_0)$ the value of the cost function at $(\tilde{L}_0, \tilde{r}_0)$, where hereafter $\tilde{r}_0$ is assumed to be computed by least squares. Then the optimization algorithm consists of the following two steps:

1) *Initialization*: Compute an equally spaced set of samples $\{L_1, \ldots, L_n\}$ in $[L_{\min}, L_{\max}]$ and let l be the distance between two consecutive samples. Compute the values, $\{c(L_1), \ldots, c(L_n)\}$, of the cost function corresponding to $\{L_1, \ldots, L_n\}$. Define two new sets of variables $\{p(1), \ldots, p(n)\}$ and $\{L^{\max}(1), \ldots, L^{\max}(n)\}$, which are initialized to $p(i) = \exp(-c(L_i)/4)$ and $L^{\max}(i) = L_i$, $i = 1, \ldots, n$. Construct a probability distribution $\pi(\cdot)$ on the interval $[L_{\min}, L_{\max}]$ as follows:
$$\pi(L) = \frac{\sum_{i=1}^{n} p(i)\mathbf{1}(L - L_i - l/2)\mathbf{1}(L_i + l/2 - L)}{l \sum_{j=1}^{n} p(j)} \tag{8}$$
2) *Iteration*: Sample a new value $\bar{L}$ of the outer scale from π. Let i be the index of the closest $\{L_j\}$ to $\tilde{L}$. If $\exp(-c(\tilde{L})/4) > p(i)$ then we update $p(i)$ and $L^{\max}(i)$: $p(i) = \exp(-c(\tilde{L})/4)$, $L^{\max}(i) = \tilde{L}$.

When the number of iterations of 2) becomes large and $\hat{D}(d_i) = D_\phi(d_i)$ (which occurs only asymptotically in T), then $\hat{L}_0 = L^{\max}(\bar{i})$ associated to $\bar{i} = \arg\max(p(i))$ approaches L_0 (and consequently the Fried parameter associated to $\hat{L}_0$ approaches r_0).

Even if simple, the algorithm allow us to quickly find a sufficiently good approximation of $\{L_0, r_0\}$, as shown in Section V.

IV. Detection of layers: Speed and energy

The aim of this section is the estimation of the turbulence parameters $(l, \gamma_1, \ldots, \gamma_l, v_{1,u}, \ldots, v_{l,u}, v_{1,v}, \ldots, v_{l,v})$, and the procedure is a slight modification of that presented in [5]. 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, (4), 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 $c_\phi(u - \bar{u}, v - \bar{v}, \bar{k})$, $k = 1, \ldots, \bar{T}$, which should correspond to translating layers.

[1] Actually we compute r_0 as the solution of a weighted least squares problem. This to allow us the possibility of assigning, in computing r_0, less importance to the values of $\hat{D}(d_i)$ with large d_i, which are estimated using a low number of samples and thus are not so reliable.

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 modify the procedure described above, but maintaining the idea of looking at spatio-temporal correlations. In particular, we take advantage of a MRF spatial representation of the turbulent phase: The goal is that of using this representation to compute a sort of *"spatially whitening filter"*. Once obtained this filter, we use it to compute a spatially "almost white" process e. Then, exploiting the same considerations as above,
$$c(u, v, \bar{k}) = \mathbf{E}[e(u + \bar{u}, v + \bar{v}, t + \bar{k}T_s)e(\bar{u}, \bar{v}, t)]$$
has peaks in locations corresponding to the translating layers. However, thanks to the "almost whiteness" of e, the peaks are much more apparent on $c(\cdot)$ than on $c_\phi(\cdot)$.

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 $\mathbb{N}(\cdot)$,
$$\mathbb{N}(\bar{u}, \bar{v}) = \left\{(u, v) \in \mathbb{L} \mid 0 < |u - \bar{u}, v - \bar{v}| \leq \bar{d}\right\} \ ,$$
where $\bar{d}$ is a suitable distance, and with $|(u, v)|$ we indicate the modulus (Euclidian length) of the 2D vector (u, v).

Then, as shown in [13], $\phi(\bar{u}, \bar{v}, t)$, the value of the MRF on the generic point $(\bar{u}, \bar{v})$ at time t, can be expressed as the best linear prediction of $\phi(\bar{u}, \bar{v}, t)$ given 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
$$\phi(\bar{u}, \bar{v}, t) = \sum_{(u,v)\in\mathbb{N}(\bar{u},\bar{v})} a_{|(\bar{u}-u,\bar{v}-v)|}\phi(u, v, t) + e(\bar{u}, \bar{v}, t) \ , \tag{9}$$
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,
$$\mathbf{E}[e(\bar{u}, \bar{v}, t)e(u, v, t)] = \begin{cases} \sigma_e^2 & (\bar{u}, \bar{v}) = (u, v) \\ -a_{|(\bar{u}-u,\bar{v}-v)|}\sigma_e^2 & (u, v) \in \mathbb{N}(\bar{u}, \bar{v}) \\ 0 & \text{otherwise.} \end{cases} \tag{10}$$
The terms "spatial innovation" and "spatially almost white" for e are motivated by (10).

Assuming to know the covariance values in (3), the coefficients $\{a_i\}$ can be computed as those of the usual best linear predictor [14]. However, in this application the exact covariance values are not available. Instead, we can approximate them introducing the estimated spatial parameters $\{\hat{L}_0, \hat{r}_0\}$ in (3). The derived coefficients $\{\hat{a}_i\}$ will only be an approximation of the true ones, nevertheless (9) can be rewritten as follows:
$$\phi(\bar{u}, \bar{v}, t) = \sum_{(u,v)\in\mathbb{N}(\bar{u},\bar{v})} \hat{a}_{|(\bar{u}-u,\bar{v}-v)|}\phi(u, v, t) + e_{\hat{L}_0,\hat{r}_0}(\bar{u}, \bar{v}, t)$$

where $e_{\hat{L}_0,\hat{r}_0}$ depends on the estimated parameters $\hat{L}_0, \hat{r}_0$. Typically, $e_{\hat{L}_0,\hat{r}_0}$ does not satisfy exactly the requirements to be a "spatial innovation" process for ϕ, i.e. (10) does not hold for $e_{\hat{L}_0,\hat{r}_0}$. However, if the estimates $\hat{L}_0, \hat{r}_0$ are good enough, $e_{\hat{L}_0,\hat{r}_0}$ approximatively satisfy (10) and results to be useful for the detection of moving layers.

Since the above considerations made for ϕ and $c_\phi(\cdot)$ can be repeated for e and $c(\cdot)$, then they approximatively hold also for $e_{\hat{L}_0,\hat{r}_0}$ and $c_{\hat{L}_0,\hat{r}_0}(\cdot)$, where

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

Hence, the idea of the layer detection algorithm is that of associating peaks in $c_{\hat{L}_0,\hat{r}_0}(\cdot)$ to moving layers.

If a peak in $c_{\hat{L}_0,\hat{r}_0}(u,v,\bar{k})$ corresponds to a real layer, then even $c_{\hat{L}_0,\hat{r}_0}(uk/\bar{k}, vk/\bar{k}, k)$ for $k < \bar{k}$ should assume quite large vales. Thus, to reduce further the influence of noise, we integrate $c_{\hat{L}_0,\hat{r}_0}(\cdot)$ over k, that is:

$$c^p_{\hat{L}_0,\hat{r}_0}(u,v,\bar{k}) = \frac{1}{\bar{k}-1}\sum_{k=1}^{\bar{k}-1} c_{\hat{L}_0,\hat{r}_0}(uk/\bar{k}, vk/\bar{k}, k) \ ,$$

for $\bar{k} > 1$.

Finally, let us consider the ith layer, characterized by the velocity vector $(v_{i,u}, v_{i,v})$, and assume to have collected estimations $\hat{c}_{\hat{L}_0,\hat{r}_0}(u,v,k)$ of $c_{\hat{L}_0,\hat{r}_0}(u,v,k)$ for $k = 1,\ldots,\bar{k}$ and $(u,v) \in \mathbb{L}_c$, where $\mathbb{L}_c$ is a finite spatial domain such that $\mathbb{L} \subset \mathbb{L}_c$. Since $(v_{i,u}, v_{i,v}) \neq 0$, its associated peak in $\hat{c}_{\hat{L}_0,\hat{r}_0}(\cdot)$ starts from the origin (where it is at $k = 0$), and gets further as k becomes larger: At time kT_s the peak is in $c_{\hat{L}_0,\hat{r}_0}(v_{i,u}k, v_{i,v}k, k)$, and $|(v_{i,u}k_1, v_{i,v}k_1)| < |(v_{i,u}k_2, v_{i,v}k_2)|$ when $k_1 < k_2$. Hence, $\forall i$ there exists an integer k_i such that $(v_{i,u}k_i, v_{i,v}k_i) \in \mathbb{L}_c$ and $(v_{i,u}(k_i+1), v_{i,v}(k_i+1)) \notin \mathbb{L}_c$. Then, if $\bar{k}$ is large enough, all the peaks corresponding to the real layers are out of $\mathbb{L}_c$.

The algorithm first compute from samples the estimates $\hat{c}^p_{\hat{L}_0,\hat{r}_0}(\cdot)$ of $c^p_{\hat{L}_0,\hat{r}_0}(\cdot)$. Then, it iteratively searches for new layers starting from $k = \bar{k}$ to $k = 1$. The goal is that of detecting layers when they get in $\mathbb{L}_c$: Specifically, at $k = k_i$ for the ith layer.

As a consequence of the above considerations, 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 $k < \bar{k}$, $(u(k+1)/k, v(k+1)/k) \notin \mathbb{L}_c$.

Even if restrictive, 3) is motivated by the fact that $c^p_{\hat{L}_0,\hat{r}_0}(\cdot,\cdot,k+1)$ contains less noise than $c^p_{\hat{L}_0,\hat{r}_0}(\cdot,\cdot,k)$: Thus, if 1) and 2) are satisfied by $c_{\hat{L}_0,\hat{r}_0}(u,v,k)$ and not by $c_{\hat{L}_0,\hat{r}_0}(u(k+1)/k, v(k+1)/k, k+1)$, then the large value in $c_{\hat{L}_0,\hat{r}_0}(u,v,k)$ is probably caused by noise.

See [5] for a detailed description of the previous version of the algorithm described in this Section and its application to turbulent phase prediction.

V. Simulations

We assume to be in VLT-like conditions, e.g. $d = 8$ meter and $n_s = 40$ and we investigate the results of the estimation of turbulence characteristics (L_0, r_0 and layers characteristics) with the following choices for the parameters:

(A) $L_0 = 50$m, $r_0 = 0.4$m, $\sigma_m^2 = 0.6$rad^2, three layers with the following characteristics: $v_1 = 7\mathbf{u}$ m/s, $v_2 = -16\mathbf{u}$ m/s, $v_3 = 30\mathbf{u}$ m/s, $h_1 = 0$ Km, $h_2 = 2$ Km, $h_3 = 8$ Km, $\gamma_1^2 = 0.50$, $\gamma_1^2 = 0.30$, $\gamma_1^2 = 0.20$.

(B) $L_0 = 22$m, $r_0 = 0.2$m, $\sigma_m^2 = 0.6$rad^2 three layers with the following characteristics: $v_1 = 5\mathbf{u}$ m/s, $v_2 = 9\mathbf{u}$ m/s, $v_3 = 24\mathbf{u}$ m/s, $h_1 = 0$ Km, $h_2 = 2$ Km, $h_3 = 8$ Km, $\gamma_1^2 = 0.55$, $\gamma_1^2 = 0.35$, $\gamma_1^2 = 0.10$.

First, we use the procedure described in Section III to estimate the spatial parameters, L_0 and r_0, from T turbulence samples, where T can assume two values: $T = \{1, 1000\}$. To reduce the temporal correlation of the samples, these are sampled at a frequency of 1 Hz (i.e. 1 sample per second). We assume to have a very rough a priori information about L_0: $L_{\min} = 1$ m, $L_{\max} = 200$ m.

The estimated structure functions for the simulations with parameters (A) are shown in Fig. 3, while the estimated parameters in both the considered cases are:

(A)	$T = 1$:	$\hat{L}_0 = 8.6$m, $\hat{r}_0 = 0.23$m.
	$T = 1000$:	$\hat{L}_0 = 15.65$m, $\hat{r}_0 = 0.25$m.
(B)	$T = 1$:	$\hat{L}_0 = 17.2$m, $\hat{r}_0 = 0.19$m.
	$T = 1000$:	$\hat{L}_0 = 14.6$m, $\hat{r}_0 = 0.16$m.

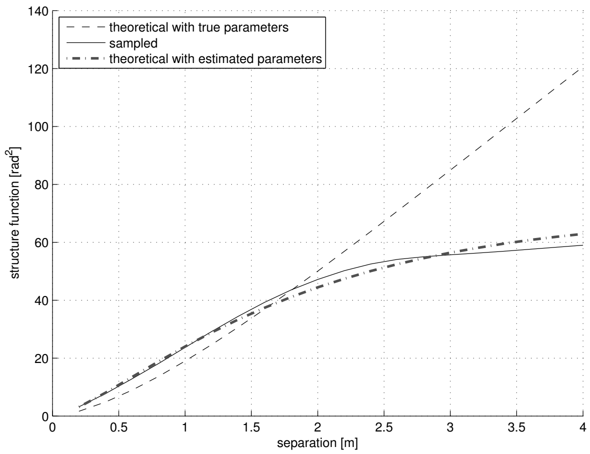
As shown in Fig. 3, the optimization algorithm provides values $\{\hat{L}_0, \hat{r}_0\}$ of the parameters which allow the theoretical structure function $D_{\phi,\hat{L}_0,\hat{r}_0}(\cdot)$ fit very well the sampled structure function. However, due to the influence of noise, the values of the sample structure functions are not so close to the real ones. Consequently, $\{\hat{L}_0, \hat{r}_0\}$ are not so close to their true values.

This would suggest us to use a larger number of samples, T, in the estimation of $\hat{D}(\cdot)$. However, we will see in the following that $\{\hat{L}_0, \hat{r}_0\}$ (estimated using $T = 1000$ samples) are sufficiently good to allow the layer estimation procedure of Section IV work properly.

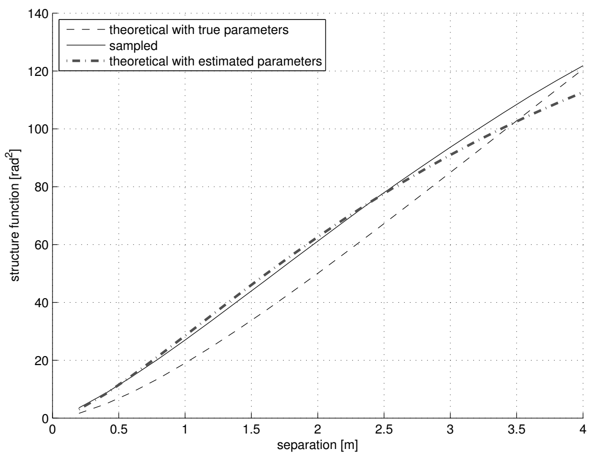
The estimation of the layer characteristics (i.e. velocities and energies) is done using $\bar{T} = 1000$ samples at a sampling frequency of 10 Hz (i.e. with this settings a layer moving at 20 m/s translates of a quarter of the telescope diameter per sample period). The simulations are done with both the two turbulence settings defined in (A) and (B), and the MRF models are computed exploiting the $\hat{L}_0$ and $\hat{r}_0$ computed previously (using $T = 1000$ samples).

The results of the layer estimation are reported in Table I for the (A) settings, and Table II for (B): $v_{i,u}$, $v_{i,v}$ 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.

Even if using approximated values of $\{L_0, r_0\}$, the layer characteristics estimation algorithm still performs well: In all of the examples the number of layers has been correctly



(a)



(b)

Fig. 3. In dashed line the theoretical structure function, with $L_0 = 50$m and $r_0 = 0.4$m, as a function of space distance. In solid line the structure function estimated from samples. In bold dashed-dotted line the structure function reconstructed using the estimated parameters $\{\hat{L}_0, \hat{r}_0\}$. Structure function and $\{\hat{L}_0, \hat{r}_0\}$ are estimated using: (a) $T = 1$ sample, (b) $T = 1000$ samples.

estimated, i.e. $\hat{l} = l$, and the values of the estimated parameters are quite close to the true ones.

VI. Conclusions

In this paper we have presented a procedure for the estimation of the parameter characterizing the turbulent phase spatial statistics and for the detection of atmospheric turbulence layers. First, we have proposed a method which exploits the structure function statistical characterization of the turbulent phase to estimate the outer-scale and the Fried parameter from the sample phase covariances.

Then, exploiting a MRF representation of the turbulent phase we have estimated the spatio-temporal correlations of the "spatial innovation". We have proposed a procedure to properly analyzing these spatio-temporal correlations and extracting the useful information about the turbulent phase

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

structure, i.e. the number of layers and their characterizing parameters.

As shown in [6] and [5], the results obtained in this paper can be exploited to improve the performances of the overall AO control system.

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 and 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.

[6] A.Beghi, A.Cenedese, F.Maran, A.Masiero, "A comparison of Kalman filter based algorithms for turbulent phase control in an adaptive optics system", Conf. on Decision and Control 2008, pp. 1839-1844, Mexico.

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

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

[9] 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.

[10] Yu.A. Rozanov, *Stationary random processes*, Holden-Day; 1967.

[11] A.V. Ivanov and N.N. Leonenko, *Statistical analysis of random fields*, Kluwer academic publishers; 1989.

[12] R. Durrett, *Probability: theory and examples*, Duxbury press; 1996.

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

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