

On the estimation of atmospheric turbulence layers for AO systems

Alessandro Beghi

Angelo Cenedese

Andrea Masiero

Abstract—In current and next generation of ground telescopes, Adaptive Optics (AO) are employed to overcome the detrimental effects induced by the presence of atmospheric turbulence, that strongly affects the quality of data transmission and therefore limits the actual resolution of the overall system. The analysis as well as the prediction of the turbulent phase affecting the light wavefront is therefore of paramount importance to guarantee the effective performance of the AO solution. In this work, a layered model of turbulence is proposed, based on the definition of a Markov-Random-Field whose parameters are determined according to the turbulence statistics. The problem of turbulence estimation is formalized within the stochastic framework and conditions for the identifiability of the turbulence structure (numbers of layers, energies and velocities) are stated. Finally, an algorithm to allow the layer detection and characterization from measurements is designed. Numerical simulations are used to assess the proposed procedure and validate the results, confirming the validity of the approach and the accuracy of the detection.

I. INTRODUCTION

Modern ground based telescopes are usually equipped with an Adaptive Optics (AO) system [9] with the specific aim of compensating or at least alleviating the detrimental effects to imaging introduced by atmospheric turbulence. The AO system, in fact, commands a set of correction mirrors (or deformable mirrors) to adapt their shapes to the opposite of 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 sequence of the AO system's working procedure can be summarized as follows:

- 1) estimation of the current turbulent phase;
- 2) prediction of the future turbulent phase;
- 3) computation of the proper control input for the set of deformable mirrors to allow compensation.

Some remarks are now in order. Firstly, since the control action is commonly delayed with respect to the related input because of the time needed for image acquisition and phase reconstruction [6] [7], the prediction step is of fundamental importance for the performances of the AO system. Secondly, the atmosphere is modeled as a linear combination of layers translating, at different altitudes, over the telescope pupil. Moreover, in a Multi Conjugated Adaptive Optics (MCAO) system, used to achieve a larger sky coverage with respect

to simpler AO systems [5], the atmosphere structure is completely reconstructed, and each mirror corrects the turbulent phase associated to one of the atmospheric layers. Hence, to make the MCAO system effective, a good reconstruction of the turbulence has to be performed. Finally, the AO (or MCAO) system shall exploit the knowledge of the turbulence's characteristics to improve the performances of the prediction step.

An example of a turbulence realization is shown in Fig. 1, where a sequence of the so-called phase screens is reported.

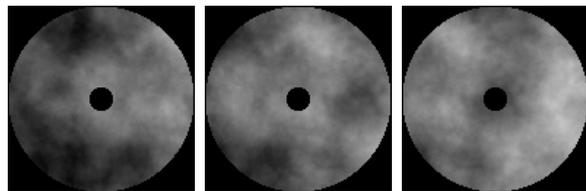


Fig. 1. A sequence of phase screens obtained as the output of a stochastic realization. The picture provides intuition on how the phase screen actually simulates the presence of a “turbulence pattern” over the telescope pupil.

Given this context, the contribution of this work regards an innovative technique to estimate the turbulence's structure. In particular, we model the atmospheric turbulence as a multilayer Markov Random Field and we aim at estimating the number of significant layers, their energy, and their velocities from the measurement of the turbulent phases. Integrated with the information about the layer altitudes, this procedure can improve the performance of the MCAO (or AO) system.

More in general, the developed identification procedure can be of interest for other applications where a similar model applies.

II. PROCESS CHARACTERIZATION

We assume to have a 2D signal transmitted through a turbulent medium, received by a finite set of sensors disposed on a discrete domain \mathbb{L} , which is supposed to be a grid as in Fig. 2(a), i.e. a sensor is placed at each node of the grid. For the sake of simplicity we assume regularity of the grid, therefore the closest neighbors of each sensor (both along the horizontal and the vertical directions) are placed at the same distance p_s . All sensors are assumed to take measurements at a sampling frequency $f_s = 1/T$ (T is the sampling period).

Let \mathbf{u} and \mathbf{v} be two unit vectors indicating two orthogonal spatial directions, as in Fig. 2(a), and let $\phi(u, v, t)$ represent the total amount of turbulent phase arriving 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} .

A.Beghi, A.Cenedese and A.Masiero are with the Department of Information Engineering, University of Padova, via Gradenigo 6/B, 35131 Padova, Italy {beghi,masiero}@dei.unipd.it, angelo.cenedese@unipd.it.

This technical report is an extended version of the European Control Conference 2013 (ECC13) paper. If needed, please cite the ECC paper (same authors and title).

We assume the measured signal ϕ to be zero-mean, stationary and spatially homogeneous. The covariance between two values of the signal, $\phi(u, v, t)$ and $\phi(u', v', t)$, depends only on the distance, r , between the two points:

$$C_\phi(r) = \mathbf{E}[\phi(u, v, t)\phi(u', v', t)], \quad (1)$$

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

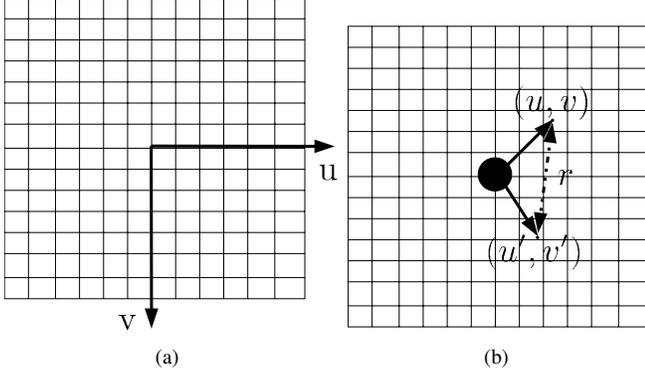


Fig. 2. (a) Coordinates on the domain \mathbb{L} . (b) Two points, (u, v) and (u', v') , separated by a distance r on the domain \mathbb{L} .

In astronomical applications, the spatial statistical characteristics of the atmospheric turbulent phase ϕ are generally described by means of the structure function, which measures the averaged difference between the phase at two points which are separated by a distance r on the aperture plane,

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

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

$$D_\phi(r) = 2(\sigma_\phi^2 - C_\phi(r)),$$

where σ_ϕ^2 is the phase variance.

From the relation between the structure function and the covariance¹, the spatial covariance of the phase between two points at distance r results

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

In order to describe its temporal characteristics, the turbulence is generally modeled as the superposition of a finite number l of layers, where we indicate with $\psi_i(u, v, t)$ the value of the i^{th} layer at point (u, v) and time t : The i^{th} layer models the atmosphere from h_{i-1} to h_i meter high,

¹According to the Von Karman theory, the phase structure function evaluated at distance r is the following (see [2]):

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

where $K(\cdot)$ 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 [3]), and the constant c is:

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

where $h_l \geq \dots \geq h_i \geq h_{i-1} \geq \dots \geq h_0 = 0$. Then, $\phi(u, v, t)$ is given by

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

where $\{\gamma_i\}$ are suitable time invariant coefficients. Without loss of generality we assume that $\sum_{i=1}^l \gamma_i^2 = 1$. Interestingly, the coefficients $\{\gamma_i^2\}$ have the physical meaning of strengths (or normalized energies) of the layers.

The layers are assumed to be zero-mean, stationary, characterized by similar spatial characteristics (i.e. all the layers are spatially described by the same structure function), i.e.

$$\mathbf{E}[\psi_i(u, v, t)\psi_i(u', v', t)] = C_\phi(r), \quad i = 1, \dots, l. \quad (4)$$

Furthermore, they are assumed to be independent, hence

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

A commonly agreed assumption considers that each layer translates in front of the telescope pupil with constant velocity v_i (Taylor approximation [11] [9]), thus (5) holds.

$$\psi_i(u, v, t + kT) = \psi_i(u - v_{i,u}kT, v - v_{i,v}kT, t), \\ i = 1, \dots, l \quad (5)$$

where $v_i = v_{i,u}\mathbf{u} + v_{i,v}\mathbf{v}$, $i = 1, \dots, l$, and kT is a delay multiple of T . In addition, we assume the velocity vectors to be different for different layers, i.e. $v_i \neq v_j$ if $i \neq j$. Finally, according with [8], we assume that the turbulent phase has Gaussian statistics.

The aim of the following sections will be that of developing a strategy for the estimation of l , the numbers of layers, and their characteristics, i.e. their velocities $\{v_i\}$ and their energies $\{\gamma_i^2\}$.

III. A MARKOV RANDOM FIELD SPATIAL MODEL OF THE PROCESS

In this Section we introduce a (discrete) Markov Random Field (MRF) spatial model for ϕ . Since we are interested in a spatial model, which is assumed to be time invariant, here we consider the time as fixed at a constant value $t = \bar{t}$. Also, we assume that ϕ is a scalar random process, and at each sampling time we observe only a limited range of its values, which are in general affected by a zero-mean white-noise process w .

Some observations are now in order. First of all, since the spatial statistics of all the layers $\{\psi_i\}$ are assumed to be characterized by (1) as in (4), then each layer ψ_i can be modeled as an isotropic homogeneous random field. Moreover, being the domain \mathbb{L} actually discrete, the process can be spatially modeled as a discrete random field.

We recall that a spatial process y is a MRF if and only if the Markov property holds for y , that is: Let (\bar{u}, \bar{v}) be a point on the grid \mathbb{L} and let $\mathbb{N}(\bar{u}, \bar{v})$ be the set of points of \mathbb{L} that are in the neighborhood of (\bar{u}, \bar{v}) , defined as follows:

$$\mathbb{N}(\bar{u}, \bar{v}) = \left\{ (u, v) \in \mathbb{L} \mid 0 < \sqrt{(u - \bar{u})^2 + (v - \bar{v})^2} \leq \bar{d} \right\},$$

where \bar{d} is a suitable value. Also, let $y(\bar{u}, \bar{v}, t)$ be the value of the spatial process y on the point (\bar{u}, \bar{v}) at time t . The Markov property for y states that $y(\bar{u}, \bar{v}, t)$ is independent on $y(u', v', t)$, $(u', v') \in \mathbb{L} - \mathbb{N}(\bar{u}, \bar{v})$ given $y(u, v, t) \forall (u, v) \in \mathbb{N}(\bar{u}, \bar{v})$.

Since it is a reasonable assumption that the covariance (1) vanishes fast as r increases, then the layers' random field models can in fact be well approximated by *Markov Random Fields* (see [4]). Furthermore, we assume also the MRFs to be characterized by Gaussian statistics, that is to be Gaussian MRFs.

To simplify the study and the notation, hereafter we consider a 1D case, i.e. the process of interest (ψ_i) is described (and moves) only along a line. The generalization to the 2D case is immediate, leading just to a more complex formulation of the equations. In the 1D case \mathbb{L} reduces to an interval, whose size is m : $\mathbb{L} = [1, \dots, m]$. Given the position $\bar{u} \in \mathbb{L}$, the neighborhood $\mathbb{N}(\bar{u})$ is redefined simply as:

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

and the interval border $\mathbb{L}_{bor} = \{\bar{u} \in \mathbb{L} \mid \bar{u} \leq \bar{d} \text{ or } \bar{u} > m - \bar{d}\}$ and the internal set $\mathbb{L}_{in} = \{\bar{u} \in \mathbb{L} \mid \bar{d} < \bar{u} \leq m - \bar{d}\}$ are also introduced.

Since the layers have the same spatial statistical characterization (i.e. zero-mean Gaussian random processes with covariance (1)), each layer can be described using a similar MRF representation. Let $\psi_i(\bar{u}, t)$ be the value of the i^{th} layer at the position $\bar{u} \in \mathbb{L}$ at time t . Then, if the sensor array is perfect (i.e. measurements are not affected by noise), at time t the measurement vector provides $[\psi_i(1, t), \psi_i(2, t), \dots, \psi_i(m, t)]^T$. As shown in [12], the value of the MRF at the generic point $\bar{u} \in \mathbb{L}_{in}$ at time t , $\psi_i(\bar{u}, t)$, can be expressed as the best linear prediction of $\psi_i(\bar{u}, t)$ given the values of its neighbors $\mathbb{N}(\bar{u})$ plus an "innovation" process $e_i(\bar{u}, t)$:

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

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, [10] for the computation of the coefficients of the best linear predictor). Interestingly, these coefficients show values lower than unity, so that the maximum for the covariance function is reached in $\bar{u} = u$, center of the neighborhood.

Furthermore, since the layers are independent on each other

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

if $i \neq j$, $\forall(\bar{t}, t)$, while

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

with \bar{u} , $u \in \mathbb{L}_{in}$.

For the global turbulence ϕ , (3), it follows that

$$\begin{aligned} \phi(\bar{u}, t) &= \sum_{i=1}^l \gamma_i \left(\sum_{u \in \mathbb{N}(\bar{u})} a_{|\bar{u}-u|} \psi_i(u, t) + e_i(\bar{u}, t) \right) \\ &= \sum_{u \in \mathbb{N}(\bar{u})} a_{|\bar{u}-u|} \phi(u, t) + e(\bar{u}, t) , \end{aligned} \quad (9)$$

where the global innovation $e(\bar{u}, t)$ is expressed as a linear combination of $e_i(\bar{u}, t)$, $i = 1, \dots, l$, that is

$$e(\bar{u}, t) = \sum_{i=1}^l \gamma_i e_i(\bar{u}, t) . \quad (10)$$

In the remainder of the work, to simplify the study and avoid border effects, we will assume that the discrete domain \mathbb{L} actually extends indefinitely, thus neglecting \mathbb{L}_{bor} (i.e. $\mathbb{L}_{bor} \rightarrow 0$ and $\mathbb{L}_{in} \approx \mathbb{L}$). This assumption is well posed if the domain \mathbb{L} results to be much larger than the dimension of neighborhood \mathbb{N} (or equally $m \gg 2\bar{d}$).

IV. LAYER DETECTION

The aim of this section is the derivation of an algorithm for the estimation of the number l of layers, of the linear combination coefficients $\{\gamma_i\}$, and of the layer characteristics, in particular the velocities $\{v_i\}$: for the 1D case, the parameters to be estimated are then $(l, \gamma_1, \dots, \gamma_l, v_1, \dots, v_l)$, whereas the procedure still applies to the more general 2D case referring to Eqs. (3) and (5), just adding a more complex notation to the procedure (in this case the parameter set is $(l, \gamma_1, \dots, \gamma_l, v_{1,u}, \dots, v_{l,u}, v_{1,v}, \dots, v_{l,v})$).

The rationale behind the detection procedure is that when the layers move with different velocities the net effect after some time is to decouple each other neighborhood of influence related to the markovian property, thus allowing the distinguishability of the single layers and their parameter identification. Conversely, if two or more layers show the same dynamic characteristics (same direction and speed) they are not singularly observable.

Equations (6), (7), (8) and (10) allow to formulate the algorithm for the detection of the layers. Three observations are now in order:

- Since the coefficients $\{a_i\}$ can be computed as those of the best linear predictor, given the second order spatial statistical characteristics of the process described by (1), then $e(u, t)$, $u \in \mathbb{L}_{in}$, can be computed for all t from (9).
- From (7) and (10), then

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

- Since layers are moving with constant velocity, it follows that

$$\begin{aligned} \mathbf{E}[e_i(\bar{u}, t)e_i(u + hv_i T, t + hT)] &= \\ &= \mathbf{E}[e_i(\bar{u}, t)e_i(u, t)] , \end{aligned} \quad (12)$$

for $i = 1, \dots, l$, $h \in \mathbb{Z}$. Thus the above expression is different from zero only if $\bar{u} \in \mathbb{N}(u)$.

Since in (11) the space-time correlations of e are expressed as a linear combination of those of the $\{e_i\}_{i=1, \dots, l}$, also the relation (12) can be extended to characterize the space-time correlations of e .

First of all, from the process homogeneity and the Taylor assumption [11], it follows that $\mathbf{E}[e(\bar{u}, \bar{t})e(u, t)]$ depends only on $\zeta = u - \bar{u}$ and $\tau = t - \bar{t}$. Thus, without loss of generality, the initial time and position can be chosen as $\bar{u} = 0$ and $\bar{t} = 0$: be $c(\zeta, \tau)$

$$c(\zeta, \tau) = \mathbf{E}[e(0, 0)e(\zeta, \tau)],$$

and similarly, be $c^i(\zeta, \tau)$

$$c^i(\zeta, \tau) = \mathbf{E}[e_i(0, 0)e_i(\zeta, \tau)].$$

According to the discrete MRF assumption, the probability distribution of the turbulence at $\bar{u} \in \mathbb{L}$ given the values of $u \in \mathbb{N}(\bar{u})$ does not depend on $u \in \mathbb{L} - \mathbb{N}(\bar{u})$: In other words,

$$c(\zeta, 0) = \mathbf{E}[e(0, 0)e(\zeta, 0)] = 0, \text{ if } |\zeta| \geq \bar{d}.$$

Nonetheless, in practical applications, when the MRF model is an approximation of a real world random field, this relation is no longer valid. By introducing

$$\epsilon = \sup |c(\zeta, 0)|, \text{ if } |\zeta| \geq \bar{d}, \quad (13)$$

it typically results $\epsilon \geq 0$, while, in the ideal MRF case, it naturally follows that $\epsilon = 0$. Since all layers share the same statistical description, this result also holds for the i -th layer. From aforementioned considerations, $c^i(0, 0) > c^i(\zeta, 0)$ for every $\zeta \neq 0$, thus $c(0, 0)^i$ is a maximum for the correlation function: $c^i(0, 0) = \sigma_e^2 > \epsilon$.

Proposition 1: Be $\zeta \in \mathbb{L} - \bigcup_i \mathbb{N}(v_i\tau)$ then $|c(\zeta, \tau)| \leq \epsilon$.

Proof: Since $\zeta \notin \mathbb{N}(v_i\tau), \forall i$, ζ is a point outside the neighborhood of any layer. From (12), $c^i(\zeta, \tau) = c^i(\zeta - v_i\tau, 0)$, and from (13) the bound $|c^i(\zeta - v_i\tau, 0)| \leq \epsilon, \forall i$ is set. Hence,

$$\begin{aligned} |c(\zeta, \tau)| &= \left| \sum_i \gamma_i^2 c^i(\zeta, \tau) \right| = \left| \sum_i \gamma_i^2 c(\zeta - v_i\tau, 0) \right| \\ &\leq \epsilon \sum_i \gamma_i^2 = \epsilon. \end{aligned}$$

In practical situations the correlation functions $c^i(\zeta, \tau)$ appear to be strongly peaked and $|c^i(\zeta, \tau)| \gg \epsilon$ for some $\zeta \in \mathbb{N}(v_i\tau)$. We introduce formally the concept with the following:

Definition 1: Given the neighborhood interval $\mathbb{N}(v_i\tau)$ and a centered subset $\bar{\mathbb{N}}(v_i\tau) \subseteq \mathbb{N}(v_i\tau)$, the correlation function $c^i(\zeta, \tau)$ is said to be *strongly peaked* if:

- $|c^i(\zeta, \tau)| > \epsilon \frac{2}{\gamma_i^2}$ if $\zeta \in \bar{\mathbb{N}}(v_i\tau)$;
- $|c^i(\zeta, \tau)| \leq \epsilon$ if $\zeta \in \mathbb{L} - \mathbb{N}(v_i\tau)$.

An instance of strongly peaked correlation function is shown in Fig. 3.

Definition 2: Layer i is said to be *disjoint* from layer j , at time τ if $|v_i\tau - v_j\tau| > 2\bar{d}$ for every $j \neq i$.

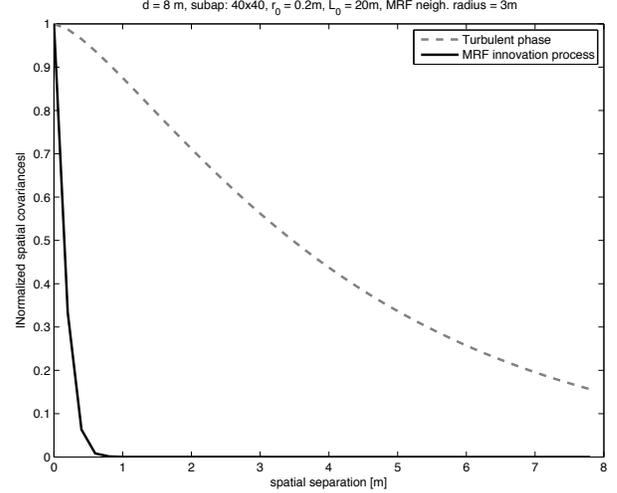


Fig. 3. Example of strongly peaked correlation function. The turbulence phase covariance function (dashed line) is compared to the spatial innovation correlation function (solid line) to highlight the strong peakedness of the latter.

Lemma 1: $\{c^i(\zeta, \tau)\}$ are strongly peaked. Consider layer i : If $\zeta \in \bar{\mathbb{N}}(v_i\tau)$ and layer i is disjoint from all others, then $|c(\zeta, \tau)| > \epsilon$.

Proof: Be $\zeta \in \bar{\mathbb{N}}(v_i\tau)$ and layer i disjoint from its siblings. It follows:

$$\begin{aligned} |c(\zeta, \tau)| &= \left| \sum_i \gamma_i^2 c^i(\zeta, \tau) \right| = \left| \gamma_i^2 c^i(\zeta, \tau) + \sum_{j \neq i} \gamma_j^2 c^j(\zeta, \tau) \right| \\ &\geq \gamma_i^2 |c^i(\zeta, \tau)| - \left| \sum_{j \neq i} \gamma_j^2 c^j(\zeta, \tau) \right| \\ &\geq \gamma_i^2 |c^i(\zeta, \tau)| - \sum_{j \neq i} \gamma_j^2 |c^j(\zeta, \tau)| \\ &> \gamma_i^2 \frac{2\epsilon}{\gamma_i^2} - \epsilon = \epsilon. \end{aligned}$$

Let $\|\bar{\mathbb{N}}\|$ be the width of the interval of strongly-peakedness, and $\bar{d}_{\bar{\mathbb{N}}} = \bar{d} - \|\bar{\mathbb{N}}\|/2$.

Definition 3: Layer i is said to be *strongly disjoint* from layer j , at time τ if $|\zeta_i - \zeta_j| > 3\bar{d}_{\bar{\mathbb{N}}}$ for every pair $(\zeta_i \in \bar{\mathbb{N}}(v_i\tau), \zeta_j \in \bar{\mathbb{N}}(v_j\tau))$.

Remark: In the case of strongly disjoint layers, the distance between the related neighborhoods is always larger than $\bar{d}_{\bar{\mathbb{N}}}$.

Proposition 2: Let \bar{t} be such that all layers are disjoint at time \bar{t} . Then, for every $\tau \geq \bar{t}$, if $|c(\zeta, \tau)| > \epsilon$, $\exists!$ i such that $\zeta \in \bar{\mathbb{N}}(v_i\tau)$.

Proof: Since all layers are disjoint, $\mathbb{N}(v_i\tau) \cap \mathbb{N}(v_j\tau) = \emptyset$ for every pair of layers $\{i, j\}$. At the same time, Prop. 1 ensures that in $\mathbb{L} - \bigcup_i \mathbb{N}(v_i\tau)$ the value of the covariance function is bounded by ϵ . Then, to have $|c(\zeta, \tau)| > \epsilon$, ζ must belong to one and only one $\mathbb{N}(v_i\tau)$.

By definition $c(\zeta, \tau)$ can be expressed as a linear combination of $c^i(\zeta, \tau)$. Remarkably, from the previous observations

and the fact that the distance among layers $|v_i\tau - v_j\tau|$ increases with time, it follows that for sufficiently long τ the elements in the linear combination (11) acts separately. The influence of each layer on $c(\zeta, \tau)$ is visible in a location spatially separated from the others, which makes the turbulent layers potentially distinguishable from each other. These considerations allow us to define the interval $\bar{\mathbb{N}}(v_i\tau)$ as the *influence interval* of layer i at time τ . Viceversa, $|c(\zeta, \tau)|$ is greater than ϵ only if there is at least a layer, i , such that $\zeta \in \mathbb{N}(hv_iT)$ for some integer number h . Thus, since the velocities of different layers are assumed to be different, the time covariances of the spatial prediction error e , provides a simple method to detect the layers.

The detection problem is now to obtain $\{v_i\}$ starting from values of $c(hv_iT, hT)$, and the solution follows a two-step approach:

- first, to distinguish the influence intervals associated to the layers;
- then, to estimate $\{v_i\}$ as that corresponding to the maximum $|c(\zeta, \tau)|$ in the influence intervals of layer i (when separated from those associated to other layers).

Lemma 2: If layers are strongly peaked and layer i is strongly disjoint from the others at time $\bar{\tau}$, then it is possible to determine a closed interval $[\zeta_a, \zeta_b]$ such that $\bar{\mathbb{N}}(v_i\tau) \subseteq [\zeta_a, \zeta_b]$.

Proof: From the strong-peakedness it follows that:

$$\max \{ \zeta_2 - \zeta_1 \mid \zeta_j \in \mathbb{N}(v_i\bar{\tau}), c(\zeta_j, \bar{\tau}) > \epsilon, j = \{1, 2\}; \\ c(\zeta, \bar{\tau}) \leq \epsilon, \forall \zeta \text{ such that } \zeta_1 < \zeta < \zeta_2 \} \leq \bar{d}_{\bar{\mathbb{N}}}.$$

Since layer i is strongly disjoint then for any $j \neq i$: $|\zeta_i - \zeta_j| > \bar{d}_{\bar{\mathbb{N}}}$ for every pair $(\zeta_i \in \mathbb{N}(v_i\tau), \zeta_j \in \mathbb{N}(v_j\tau))$.

Thus ζ_a and ζ_b can be obtained as follows:

$$\zeta_b = \min \{ \zeta \mid \zeta \geq v_i\tau, c(\zeta, \bar{\tau}) > \epsilon, \bar{\zeta} - \zeta > \bar{d}_{\bar{\mathbb{N}}}, \\ \forall \bar{\zeta} > \zeta \text{ such that } c(\bar{\zeta}, \bar{\tau}) > \epsilon \},$$

and

$$\zeta_a = \max \{ \zeta \mid \zeta \leq v_i\tau, c(\zeta, \bar{\tau}) > \epsilon, \zeta - \bar{\zeta} > \bar{d}_{\bar{\mathbb{N}}}, \\ \forall \bar{\zeta} < \zeta \text{ such that } c(\bar{\zeta}, \bar{\tau}) > \epsilon \}.$$

We introduce in this context the notion of *distinguishability* as follows.

Definition 4: Layer i is said to be *distinguishable* from layer j , at time \bar{t} if its velocity v_i and energy γ_i can be estimated with a suitable upper bound on the estimation errors.

The following proposition gives sufficient conditions for the detection of layers.

Proposition 3: If layers are strongly peaked and strongly disjoint from each other at time $\bar{\tau}$, then all the layers can be detected. For each layer, the estimated velocity \hat{v}_i satisfies the following:

$$|\hat{v}_i - v_i| \leq \bar{d}/\bar{\tau}. \quad (14)$$

Furthermore, if $\bar{\tau}v_i \in \mathbb{L}$ then layer i is distinguishable from all the other layers and the following inequality holds for the estimated energy $\hat{\gamma}_i^2$:

$$|\hat{\gamma}_i^2 - \gamma_i^2| \leq \epsilon/\sigma_e^2. \quad (15)$$

Proof: Since layers are strongly disjoint, at time $\bar{\tau}$ we can distinguish $\hat{l} = l$ intervals where the correlation function shows values greater than ϵ , and associate each of these intervals to one layer. Consider the i -th interval and be the local maximum value of $|c(\zeta, \bar{\tau})|$ attained at $\bar{\zeta}$. This is mainly due to the contribution of the i -th layer, that is $\bar{\zeta} = \hat{v}_i\bar{\tau}$. It follows that:

$$\hat{v}_i = \bar{\zeta}/\bar{\tau}.$$

On the other hand, it may happen that due to asynchronous sampling the maximum error in detecting $\bar{\zeta}$ is \bar{d} . Hence, the error bound on the velocity estimation is given by:

$$|\hat{v}_i - v_i| \leq \bar{d}/\bar{\tau}.$$

As for the energy, it stands:

$$|c(\bar{\zeta}, \bar{\tau})| = \left| \gamma_i^2 c^i(\bar{\zeta}, \bar{\tau}) + \sum_{j \neq i} \gamma_j^2 c^j(\bar{\zeta}, \bar{\tau}) \right|.$$

If $\bar{\tau}v_i \in \mathbb{L}$, then the maximum for $|c^i(\bar{\zeta}, \bar{\tau})|$ is σ_e^2 , while there is an upper bound on the second term sum given by ϵ . Therefore, the estimated value is:

$$\hat{\gamma}_i^2 = |c(\bar{\zeta}, \bar{\tau})|/\sigma_e^2,$$

and

$$|\hat{\gamma}_i^2 - \gamma_i^2| \leq \epsilon/\sigma_e^2.$$

Assume to be provided of n_τ covariances of e , that is $\{c(\zeta, \tau)\}$, $\tau \in \{\tau_{\min}, \tau_{\min}+T, \dots, \tau_{\max}\}$. Then, the previous considerations and propositions suggest an algorithm for the layer detection:

Algorithm 1 Detection of the layers

- Be \mathcal{S} the set of the indexes of the already detected layers. At $t = 0$, $\mathcal{S} = \emptyset$.
- At $\tau = \bar{\tau}$, evaluate $\{c(\zeta, \bar{\tau})\}$ for each $\zeta \in \mathbb{L}$.
- If $\exists \bar{\zeta}$ and $\exists! j \in \mathcal{S}$ such that:
 - $\bar{\zeta} \in [-\bar{d} + \hat{v}_j T, \bar{d} + \hat{v}_j T]$,
 - $c(\bar{\zeta}, \bar{\tau})$ is the maximum in $[-\bar{d} + \bar{\zeta}, \bar{d} + \bar{\zeta}]$
 - $c(\bar{\zeta}, \bar{\tau})/\sigma_e^2 > \hat{\gamma}_j^2$,

then update already detected layers with $\hat{v}_j = \bar{\zeta}/\bar{\tau}$, $\hat{\gamma}_j^2 = c(\bar{\zeta}, \bar{\tau})/\sigma_e^2$.

- Else if $\exists \bar{\zeta}$ such that:

- $c(\bar{\zeta}, \bar{\tau}) > \epsilon$,
- $|\bar{\zeta} - \hat{v}_j \bar{\tau}| \geq 2\bar{d} + 1, \forall j \in \mathcal{S}$
- $c(\bar{\zeta}, \bar{\tau})$ is the maximum in $[-\bar{d} + \bar{\zeta}, \bar{d} + \bar{\zeta}]$

then add a new layer i to \mathcal{S} with $\hat{v}_i = \bar{\zeta}/\bar{\tau}$, $\hat{\gamma}_i^2 = c(\bar{\zeta}, \bar{\tau})/\sigma_e^2$.

Remark: Due to the finite base approximation inherent in the detection procedure and data noise, normalization of

the learnt energy values may be needed after detection of all layers in order to ensure that $\sum_i \hat{\gamma}_i^2 = 1$. In actual fact, the net effect of noise is to alter the threshold induced by ϵ , therefore to guarantee the same accuracy in the parameter detection a larger number of samples is required.

V. SIMULATIONS

Simulations are organized as follows. First, in Subsection V-A Algorithm 1 is applied to a 1D case. Then, a more realistic situation is considered in Subsection V-B where it is presented the estimation of the layers characteristics of a 2D atmospheric turbulence for the AO system.

The turbulence phase has been simulated using the algorithm described in [1].

A. Detection of the layers: 1D simulations

Even if the main goal is that of applying the proposed algorithm to 2D signals, we first present a 1D example to give some intuition of the obtained results.

We consider a 4-layer turbulence case (i.e. $l = 4$), with the following values for the parameters: $v_1 = -3.125$, $v_2 = -5.750$, $v_3 = -7.375$, $v_4 = -8.143$, $\gamma_1^2 = 0.310$, $\gamma_2^2 = 0.300$, $\gamma_3^2 = 0.200$, $\gamma_4^2 = 0.190$. The system is simulated for $N = 5000$ temporal instants. Fig. 4 shows the estimated $\hat{c}(\zeta, \tau)$, $\tau = \{1, 3, 6\}$. In Table I we summarize the results: v_i and γ_i corresponds to the true values of the parameters, \hat{v}_i and $\hat{\gamma}_i$ are the estimated ones. Notice that the algorithm first allowed the correct estimation of the number of layers, i.e. $\hat{l} = l$. The velocities are written in [pixels/frame].

TABLE I
DETECTION OF 1D LAYERS.

	layer 1	layer 2	layer 3	layer 4
v_i	-3.125	-5.750	-7.375	-8.143
\hat{v}_i	-3.120	-5.750	-7.380	-8.140
γ_i^2	0.310	0.300	0.200	0.190
$\hat{\gamma}_i^2$	0.319	0.282	0.183	0.216

B. Detection of the turbulent layers: 2D simulations

Since the layers usually move slowly over the telescope pupil, here we consider three cases 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 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 the simulations are reported in Table II, III and IV: $v_{i,u}$, $v_{i,v}$ 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 expressed in [subaperture/frame], and we remind that the subaperture size is $0.4[m]$.

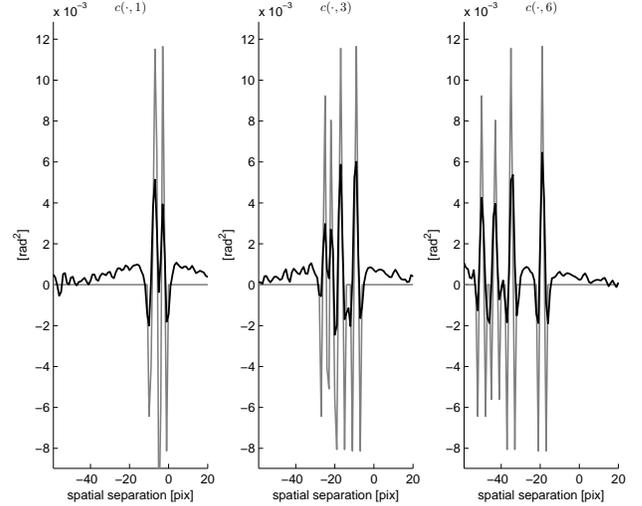


Fig. 4. In black line $\hat{c}(\zeta, \tau)$, $\tau = \{1, 3, 6\}$ estimated by the sample covariances. In gray line the detected layers.

TABLE II
DETECTION OF 2D LAYERS. CASE I.

	layer 1	layer 2	layer 3	layer 4
$v_{i,u}$	0.2160	0.3910	0.6120	0.7950
$\hat{v}_{i,u}$	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	0.3100	0.3000	0.2000	0.1900
$\hat{\gamma}_i^2$	0.3112	0.3007	0.2005	0.1876

Also, in Figs. 5–6, the identification error related to layer velocities and energies respectively, is compared with the corresponding theoretical bound as from Prop. 3.

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

VI. DISCUSSION AND CONCLUSION

In this work, we consider the turbulence modeling problem for AO system in ground telescopes. The problem of turbulence estimation is formalized within the context of stochastic dynamical systems, and a multilayer MRF model

TABLE III
DETECTION OF 2D LAYERS. CASE II.

	layer 1	layer 2	layer 3	layer 4
$v_{i,u}$	0.2160	-0.1910	0	0
$\hat{v}_{i,u}$	0.2162	-0.1905	0	0
$v_{i,v}$	0	0	0.1100	0.2870
$\hat{v}_{i,v}$	0	0	0.1111	0.2881
γ_i^2	0.4100	0.2500	0.2000	0.1400
$\hat{\gamma}_i^2$	0.4137	0.2495	0.1973	0.1395

TABLE IV
DETECTION OF 2D LAYERS. CASE III.

	layer 1	layer 2	layer 3	layer 4	layer 5
$v_{i,u}$	0.1300	-0.0710	0	0	0.0800
$\hat{v}_{i,u}$	0.1321	-0.0714	0	0	0.0800
$v_{i,v}$	0	0	0.0560	-0.0870	0
$\hat{v}_{i,v}$	0	0	0.0556	-0.0877	0
γ_i^2	0.2700	0.2300	0.2000	0.1600	0.1400
$\hat{\gamma}_i^2$	0.2505	0.2332	0.1968	0.1620	0.1576

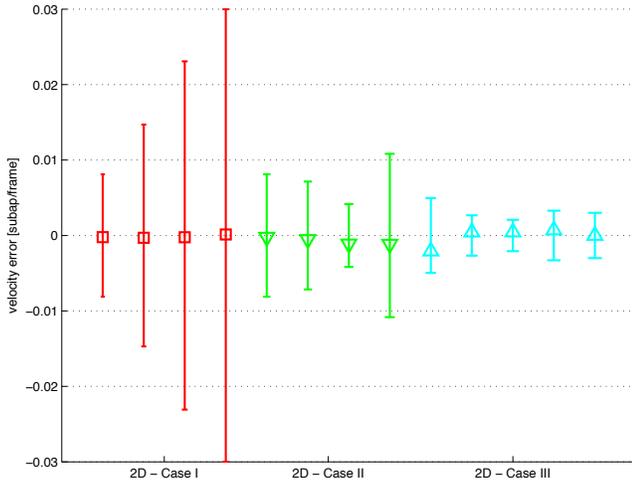


Fig. 5. Layer velocity identification error vs velocity error bound, for the presented 2D cases (highlighted with different markers and colors).

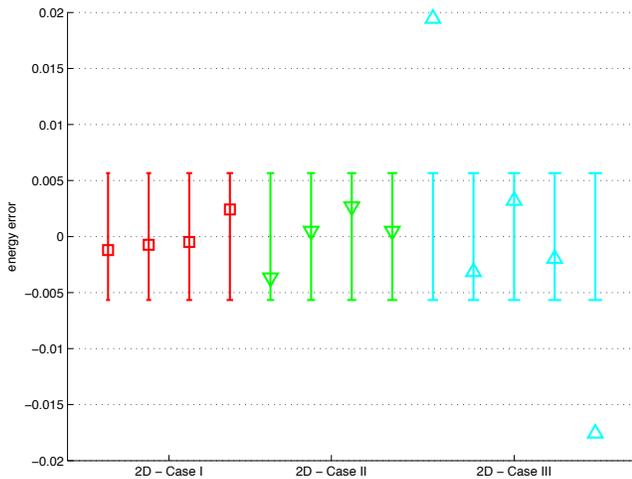


Fig. 6. Layer energy identification error vs energy error bound, for the presented 2D cases (highlighted with different markers and colors).

is derived, whose parameters are determined according to the turbulence statistics. Also, the conditions for the identifiability of the turbulence structure (in terms of numbers of layers, energies and velocities) are stated. Given these results, a procedure to identify the model main parameters is designed, in order to perform the layer detection from a set of acquired data. Numerical simulations are used to assess the proposed algorithm and confirm the validity of the method and the accuracy of the results.

The approach is formally developed through propositions and proofs that support its general validity. In fact, although motivated by the AO application, the methodology can be of interest for the estimation of the properties of other phenomena that can be modeled by a multilayer MRF.

As for the future developments,, it is of interest to design a strategy to detect layers online, when they appear to be distinguishable. The rationale is as follows: As soon as a peak emerges, $|c(\zeta, \tau)| > \epsilon$, a new layer i is set, and its velocity v_i and energy γ_i are provisionally detected; if the peak remain consistent as time evolves, the detection is refined with further observations. On the other hand, though, it may happen that the peak under assessment is generated by the superposition of two or more layers, that will appear to be distinguished in the future: In this case, the procedure will provide correct detection only at a later stage, after gathering further information. The formalization of this approach, as well as the relaxation of the strongly-peakedness hypothesis, are currently under investigation, in order to state sufficient conditions for a correct layer detection. Finally, Activity to employ the algorithm with real large telescope data is under consideration.

REFERENCES

- [1] A. Beghi, A. Cenedese, and A. Masiero. Stochastic realization approach to the efficient simulation of phase screens. *J. Opt. Soc. Am. A*, 25(2):515–525, 2008.
- [2] 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.
- [3] D.L. Fried. Statistics of a geometric representation of wavefront distortion. *J. Opt. Soc. Am.*, 55:1427–1435, 1965.
- [4] R. Kinderman and J.L. Snell. *Markov Random Fields and their Applications*. American Mathematical Society, 1980.
- [5] B. Le Roux. *Commande optimale en optique adaptative classique et multiconjuguee*. PhD thesis, Université de Nice Sophia Antipolis, 2003.
- [6] B. Le Roux, J.M. Conan, C. Kulcsar, H.F. Raynaud, L.M. Mugnier, and T. Fusco. Optimal control law for classical and multiconjugate adaptive optics. *J. Opti. Soc. Am. A*, 21(7):1261–1276, 2004.
- [7] L.A. Poyneer, B.A. Macintosh, and J.P. Véran. Fourier transform wavefront control with adaptive prediction of the atmosphere. *J. Opt. Soc. Am. A*, 24(9):2645–2660, 2007.
- [8] F. Roddier. The effects of atmospheric turbulence in optical astronomy. *Progress in Optics*, 19:281–376, 1981.
- [9] F. Roddier. *Adaptive optics in astronomy*. Cambridge University Press, 1999.
- [10] T. Soderstrom. *Discrete-time stochastic systems*. Springer, 1994.
- [11] G. I. Taylor. Statistical theory of turbulence. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 151(873):421–444, 1935.
- [12] J.W. Woods. Two-dimensional discrete markovian fields. *IEEE Trans. on Information Theory*, 18(2):232–240, 1972.