Proceedings of the 17th World Congress
The International Federation of Automatic Control
Seoul, Korea, July 6-11, 2008

On the estimation of atmospheric turbulence layers ⋆

Alessandro Beghi * **Angelo Cenedese** ** **Andrea Masiero** ***

* *Dipartimento di Ingegneria dell'Informazione, Università di Padova via Gradenigo 6/B, 35131 Padova, Italy. (e-mail: beghi@dei.unipd.it)*
** *Dipartimento di Tecnica e Gestione dei Sistemi Industriali, Università di Padova Stradella San Nicola 3, 36100 Vicenza, Italy, (e-mail: angelo.cenedese@dei.unipd.it)*
*** *Dipartimento di Ingegneria dell'Informazione, Università di Padova via Gradenigo 6/B, 35131 Padova, Italy, (Tel: 0039-049-827-7757; e-mail: masiero@dei.unipd.it)*

Random fields; linear prediction.

Abstract: The knowledge of the characteristics of the atmospheric turbulence is of fundamental importance in order to improve the performances of the adaptive optics systems in the next generation of telescopes. Starting from the statistical description of the turbulence, we derive a random field spatial model of the turbulent phase. Then we use this spatial statistical model to compute the spatial innovation. Finally, from the temporal covariances of the spatial innovation we estimate the number of the turbulence layers and their characteristics.

1. INTRODUCTION

Thanks to the technological improvement, the current generation of ground based telescopes house lenses of several meter diameter. However, changes in the atmosphere refraction index influence the beams arriving from a star: Each beam is delayed of a different phase, which depends on its path and on the current atmospheric turbulence. Because of the different values of these delays, a flat wavefront coming from a star is no more flat when it arrives on the telescope lens. This significantly reduces the real resolution of the telescope. The set of values of the phase delays of the beams arriving on the telescope pupil are commonly called *turbulent phase*.

To reduce the problems due to the presence of the atmospheric turbulence, telescopes are usually provided with an Adaptive Optics (AO) system (Roddier (1999)): This 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. The tasks of the AO system can be summarized as estimating the current turbulent phase, predicting the new one, and computing the proper control input for the set of deformable mirrors. Notice that the control is commonly delayed of two sample periods because of the time for image acquisition and phase reconstruction (see Le Roux et al. (2004), Le Roux (2003)): Thus the prediction step is of fundamental importance for the performances of the AO system.

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

A standard AO system can correct the phases only for a small portion of the sky. Multi Conjugated Adaptive Optics (MCAO) can be used to achieve a large sky coverage (Le Roux (2003)). The atmosphere is modeled as a linear combination of layers translating, at different altitudes, over the telescope pupil. In a MCAO system, 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 turbulence reconstruction algorithm is needed. Furthermore the AO (or MCAO) system can exploit the knowledge of the turbulence's characteristics to improve the performances of the prediction step. In this paper we propose an innovative technique to estimate the turbulence structure. In particular we aim at estimating the number of significant layers, their energy, and their velocities, from the measured turbulent phases. The proposed procedure provides estimates of these parameters, which, integrated with the information about the layer altitudes, shall make the overall MCAO (or AO) system more effective.

The paper is organized as follows. In Section 2 the common turbulence statistical model is described. Section 3 introduces a Markov Random Field spatial representation for the atmospheric turbulence. Section 4 presents the core of our procedure for atmospheric structure estimation. We conclude in Section 5 with discussing some simulation results.

2. TURBULENCE PHYSICAL MODEL

The spatial statistical characteristics of the turbulent phase ϕ are typically described by means of the structure function, which measures the averaged difference between

the phase at two points at locations r_1 and r_2 of the wavefront, which are separated by a distance r on the aperture plane (Fig. 1),

$$D_\phi(r) = \mathbf{E}\left[(\phi(r_1) - \phi(r_2))^2\right].$$

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

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

where σ_ϕ^2 is the phase variance.

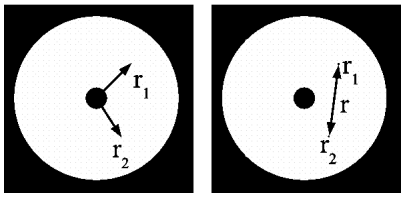


Fig. 1. Telescope image domain: Two points, r_1 and r_2, separated by a distance r on the telescope aperture plane.

According to the Von Karman theory, the phase structure function evaluated at distance r is the following (see Conan (2000)):

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

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 Fried (1965)), and the constant c is:

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

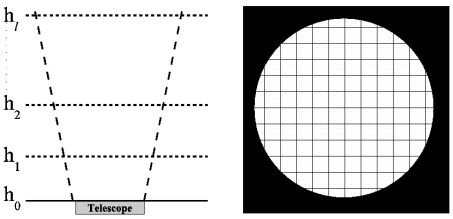
From the relation between the structure function and the covariance (1), the spatial covariance of the phase between two points at distance r results

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

Notice that in real applications only a finite number of sensors is available. These are usually distributed on a grid, thus the turbulent phase is measured only on a discrete domain $\mathbb{L}$, which is that of Fig. 2(b), i.e. a sensor is placed at each node of the grid. Without loss of generality we assume that sensors are uniformly spaced: The closest neighbors of each sensor (both along the horizontal and the vertical directions) are placed at a distance of p_s meters. We denote with $\phi(u, v, t)$ the value of the turbulent phase on the point $(u, v) \in \mathbb{L}$ at time t.

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 h_{i-1} to h_i meter high, where $h_l \geq \cdots \geq h_i \geq h_{i-1} \geq \cdots \geq h_0 = 0$ (Fig. 2(a)). Let $\psi_i(u, v, t)$ be the value of the i^{th} turbulent phase layer at point (u, v) on telescope aperture and at time t. Then the total turbulent phase at (u, v) and at time t is

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



(a) (b)

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

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

The layers are assumed to be stationary and characterized by the same spatial characteristics, i.e. all the layers are spatially described by the same structure function. Furthermore they are assumed to be independent, hence

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

A commonly agreed assumption considers that each layer translates in front of the telescope pupil with constant velocity v_i (Taylor approximation Roddier (1999)), thus

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

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

3. A 2D MARKOV RANDOM FIELD MODEL OF THE TURBULENCE

In this Section we introduce a (discrete) two-dimensional Markov Random Field (MRF) spatial model for the turbulence. Since in this Section we are interested in a spatial model, which is assumed to be time invariant, then here we consider the time as fixed at a constant value $t = \bar{t}$. Furthermore to simplify the notation we will omit $\bar{t}$ from equations.

Some observations are now in order. First, since all the layers are assumed to have the same spatial statistical characterization, then it immediately follows from Eq. (2) that each layer can be modeled as an isotropic homogeneous random field. Since the aperture plane domain is actually discrete, as in Fig. 2(b), then the turbulence can be spatially modeled as a discrete random field.

Furthermore, since the covariance (2) vanishes quite fast when r becomes larger, then the layers' random field model can in fact be well approximated by a *Markov* Random Field (see Kinderman and Snell (1980)).

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

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

where $\bar{d}$ is a suitable distance. Then the Markov property for y can be expressed as follows. Let $y(\bar{u},\bar{v})$ be the value of the spatial process y on the point $(\bar{u},\bar{v})$. Then $y(\bar{u},\bar{v})$ is independent on $y(u',v')$, $(u',v') \in \mathbb{L} - \mathbb{N}(\bar{u},\bar{v})$ given $\{y(u,v),\, \forall (u,v) \in \mathbb{N}(\bar{u},\bar{v})\}$.

In terms of linear prediction, we can summarize the Markov property with the following equation:

$$\hat{\mathbf{E}}[\; y(\bar{u},\bar{v}) \mid \forall y(u,v) \text{ such that } (u,v) \neq (\bar{u},\bar{v}) \;] = \\ = \hat{\mathbf{E}}[\; y(\bar{u},\bar{v}) \mid \forall y(u,v) \text{ such that } (u,v) \in \mathbb{N}(\bar{u},\bar{v}) \;] \;,$$

where $\hat{\mathbf{E}}[y(1)|y(2),\ldots,y(n)]$ stands for the best linear prediction of $y(1)$ given $y(2),\ldots,y(n)$.

We assume that the turbulence, considered only on the discrete domain $\mathbb{L}$, can be modeled as a (scalar) MRF. As proved in Woods (1972), we can express $y(\bar{u},\bar{v})$, the value of the process on a generic point $(\bar{u},\bar{v})$, as the best linear prediction of $y(\bar{u},\bar{v})$ given its neighbors $\mathbb{N}(\bar{u},\bar{v})$ plus an "innovation" process $e(\bar{u},\bar{v})$. According with Roddier (1981), we assume that the turbulent phase has Gaussian statistics: Thus the best linear prediction operator $\hat{\mathbf{E}}[\cdot]$ corresponds to the expectation operator $\mathbf{E}[\cdot]$. That is

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

where

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

and

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

In the above equations we denoted the Kronecker's delta with δ_u. That is

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

Furthermore $\{a_i\}$ are suitable coefficients which lead to the best linear prediction of $y(\bar{u},\bar{v})$ given its neighbors (see Soderstrom (1994)).

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

To conclude, since the layers have the same statistical characterization, each layer can be described using the model of Eq. (5). That is

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

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

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

if $i \neq j$, while

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

if $i = j$.

The algorithm for layer detection which will be presented in the next Sections is based on Eq. (7) and (8).

4. DETECTION OF LAYERS: SPEED AND ENERGY

The aim of this section is the estimation of the turbulence parameters $(l,\gamma_1,\ldots,\gamma_l,v_{1,u},\ldots,v_{l,u},v_{1,v},\ldots,v_{l,v})$ in the model given by Eq. (3) and (4). To simplify the notation, in this Section we consider the 1D case. The generalization of the procedure to the 2D case is immediate: The major drawback is that the equations become more complicated.

Thus we assume that the turbulence is a scalar random process, and at each sampling time we observe only a window of its values, eventually affected by a zero-mean white-noise process w. In this case $\mathbb{L}$ reduces to an 1D interval. Let m be the interval size, then $\mathbb{L} = [1,\ldots,m]$. Let $y(u,t)$ be the value of the turbulence on the spatial position $u \in \mathbb{L}$ at time t. Then at time t we measure

$$z(t) = [y(1,t),\ y(2,t),\ \ldots,\ y(m,t)]^T + w(t) \quad (9)$$

where

$$\mathbf{E}[w(t)w(t)^T] = \sigma_w^2 I \;.$$

We assume that the translations of each layer during a sample period (that is the velocities per frame) are rational multiples of the pixel size. Then, there exists an integer number k_i, such that

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

where v_i and $\psi_i(u,t)$ are respectively the i^{th} layer's velocity and its value on the position $u \in \mathbb{L}$ at time t.

In the 1D case Eq. (6) and (8) become respectively

$$\psi_i(\bar{t}) = \sum_{0<|t-\bar{t}|\leq\bar{d}} a_{|t-\bar{t}|}\psi_i(t) + e_i(\bar{t}),\ i = 1,\ldots,l \;, \quad (10)$$

and

$$\mathbf{E}[e_i(\bar{t})e_j(t)] = \begin{cases} \sigma_e^2 & i=j,\ \bar{t}=t \\ -a_{|t-\bar{t}|}\sigma_e^2 & i=j,\ 0<|t-\bar{t}|\leq\bar{d} \\ 0 & \text{otherwise} \end{cases} . \quad (11)$$

Let $z'(t)$ be a vector containing all the internal measurements of the turbulence at time t, that is the measurements on each node $u \subset \mathbb{L}$ such that $u > \bar{d},\ u < m - \bar{d} + 1$. Then, notice that the spatial prediction, (5) and (10), can be written as follows:

$$z'(t) = Fz(t) + \xi(t) \quad (12)$$

where

$$\xi(t) = \left[\; e(t+\bar{d}+1)\ \ldots e(m-\bar{d})\;\right]^T \;,$$

while

$$F = \begin{bmatrix} & & \ldots & & & & & \\ \ldots & a_{\bar{d}} & \ldots & a_1 & 0 & a_1 & \ldots & a_{\bar{d}} & \ldots \\ & \ldots & a_{\bar{d}} & \ldots & a_1 & 0 & a_1 & \ldots & a_{\bar{d}} & \ldots \\ & & & & \ldots & & & & \end{bmatrix} .$$

In the 1D case the parameters to be estimated are $(l,\gamma_1,\ldots,\gamma_l,v_1,\ldots,v_l)$. To this aim, first we compute $\xi(t)$ from Eq. (12). Then we compute the following covariances:

$$C_{\xi,\tau} = \mathbf{E}[\xi(t)\xi(t+\tau)^T] \;,\ 0 \leq \tau \leq \bar{T} \;.$$

It immediately follows from (11) that the matrix $C_{\xi,\tau}$ is a Toeplitz matrix, i.e.

$$C_{\xi,\tau} = \begin{bmatrix} c_{\tau,0} & c_{\tau,1} & c_{\tau,2} & & \\ c_{\tau,-1} & \ddots & \ddots & \ddots & \\ c_{\tau,-2} & \ddots & \ddots & \ddots & c_{\tau,2} \\ & \ddots & \ddots & \ddots & c_{\tau,1} \\ & & c_{\tau,-2} & c_{\tau,-1} & c_{\tau,0} \end{bmatrix} . \quad (13)$$

When $k_i = 1, \forall i$, there exists a simple expression for the coefficient $c_{\tau,k}$:

$$c_{\tau,k} = \sum_{i=1}^{l} \gamma_i^2 \left[\sigma_e^2 \delta_{v_i\tau-k} + \sum_{j=1}^{\bar{d}} -a_j \sigma_e^2 \delta_{v_i\tau+j-k} + \right. \\ \left. + \sum_{j=1}^{\bar{d}} -a_j \sigma_e^2 \delta_{v_i\tau-j-k} \right] , \ 1 \le \tau \le T . \quad (14)$$

Notice that in the above expression we can distinguish $2\bar{d}+1$ terms associated to the l^{th} layer: Actually the most interesting term is that determined by $\delta_{v_i\tau-k}$ and it is equal to σ_e^2 (that is the maximum of the terms due to the layer).

In the general case, $k_i \neq 0$, $c_{\tau,k}$ can still be written as a linear combination of terms due to the single layers, but the contributes due to a layer have a more complicated expression. However we still have that the term determined by $\delta_{k_i v_i - k}$ is σ_e^2.

If $\sigma_w^2 = 0$, then Eq. (14) is valid for $\tau = 0$ too. Since during the reconstruction process the signal has been de-noised, in this Section we assume that $\sigma_w^2 = 0$.

Let r_τ be the following discrete signal

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

then r_τ summarizes the information contained in $C_{\xi,\tau}$. Thus

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

summarizes the information in $C_{\xi,\tau}$, $\tau = 0, \dots, \bar{T}$.

Exploiting the a priori knowledge on the structure of (13), we propose the following procedure to compute the layers.

We begin with the following two definitions.

Definition 1. Two velocities v_i and v_j are said to be distinguishable in $\bar{t}$ temporal instants on the domain $\mathbb{L}$ if there exist $\bar{t}_i, \bar{t}_j$, with $1 \le \bar{t}_i, \bar{t}_j \le \bar{t}$, such that $\bar{t}_i v_i \in \mathbb{L}$, $\bar{t}_i |v_i - v_j| \ge 2\bar{d}+1$ and $\bar{t}_j v_j \in \mathbb{L}$, $\bar{t}_j |v_i - v_j| \ge 2\bar{d}+1$.

This definition can be easily generalized as follows.

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

A crucial observation here is that

$$r_0(t) = \begin{cases} \sigma_e^2 & t = 0 \\ -a_{|t-\bar{t}|}\sigma_e^2 & 0 < |t| \le \bar{d} \\ 0 & \text{otherwise} \end{cases} ,$$

that is, when $\tau = 0$, the contribution of all the layers is positioned around the coordinate $t = 0$. When $\tau > 0$ the contribution of each layer moves of a quantity which depends on the layer velocity. Since the velocities of different layers are assumed to be different, when τ becomes larger the distance between the layers in r_τ becomes larger too. Thus when τ is sufficiently large the contribute of different layers is separated in r_τ. Exploiting this observation it is sufficient to search in r_τ for separated sums of "bases". We propose the following algorithm:

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

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

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

The above Proposition represents the main result of this paper: In fact it provides a theoretical foundation for the proposed algorithm.

4.1 Dealing with data

In the previous section we have considered $C_{\xi,\tau}$ computed as the expectation $\mathbf{E}[\xi(t)\xi(t+\tau)^T]$. Consequently also the values of $r_\tau(\cdot)$ are the expected ones. Since in practical applications we cannot have directly access to the expected values we have to compute them from data.

We assume to have $N+1$ consecutive measurements $(z(0), \dots, z(N))$. Let $\hat{\xi}(t)$, $0 \le t \le N$ be the sampled value of ξ at time t reconstructed using the measurements. Then assuming the process to be wide-sense ergodic, we have that

$$\lim_{N\longrightarrow+\infty}\frac{1}{N-\tau}\sum_{i=1}^{N-\tau}\hat{\xi}(i)\hat{\xi}(i+\tau)^T=\mathbf{E}[\xi(i)\xi(i+\tau)^T]\ .$$

We first assume that the measurements are not affected by noise. In practice, since we have only a finite number of measurements to estimate the expected values, we have only an estimate $\hat{r}_\tau(t)$ of $r_\tau(t)$. Then we introduce a threshold, $\bar{\sigma}^2$, to distinguish when $\hat{r}_\tau$ has to be considered *practically* zero:

$$\begin{cases} r_\tau(t)=0 & \text{if } \hat{r}_\tau(t)<\bar{\sigma}^2 \\ r_\tau(t)\approx\hat{r}_\tau(t) & \text{otherwise}\end{cases}\ .$$

Notice that in this way we eventually discard layers with energy lower than $\bar{\sigma}^2$. Since we know that $r_0(t)$ should be equal to zero when $\bar{d}<|t|\leq m-1$, then we can estimate σ_0^2, the variance of $\hat{r}_0(t)$ $\bar{d}<|t|\leq m-1$, and set, for example, $\bar{\sigma}^2=4\sigma_0^2$. Otherwise, $\bar{\sigma}^2$ can be manually set to the minimum energy of the layers that we want to detect.

Because of the finite number of samples N, it can happen that Algorithm 1 provides an estimate such that $\sum_{i=1}^{\hat{l}}\hat{\gamma}_i^2>1$. Hence we have to modify the algorithm to take into account this problem. Actually when $\sum_{i=1}^{\hat{l}}\hat{\gamma}_i^2>1$ we reduce the number of layers: Let the layers be sorted in decreasing order of energy and let $\bar{l}$ be such that $\sum_{i=1}^{\bar{l}}\hat{\gamma}_i^2>1$ while $\sum_{i=1}^{\bar{l}-1}\hat{\gamma}_i^2<=1$, then

$$\hat{l}=\arg\min_{l=\{\bar{l},\bar{l}-1\}}|1-\sum_{i=1}^{\bar{l}-1}\hat{\gamma}_i^2|\ .$$

The main drawback of having noisy measurements is that, even if asymptotically $\mathbf{E}[w(t)w(t+\tau)^T]=0$, $\tau\neq0$, this covariance estimated by a finite number of samples is not exactly equal to zero. This in practice implies an increase in the number of samples needed to achieve the same performance of the noise-free case.

Taking in account of these considerations the algorithm becomes:

Algorithm 2. Detection of the layers
Step 1: Rough estimation of the velocities
$\hat{\mathcal{S}}_{\hat{l}}=\emptyset$;
for $\tau=\bar{T}:1$
  for $t=-m+1:m-1$
    $v=t/\tau$;
    if $\big(\hat{r}_\tau(t)>\bar{\sigma}^2\ \wedge$
      $\wedge\ \hat{r}_\tau(t)=\max\big(\hat{r}_\tau(t-\bar{d}),\ldots,\hat{r}_\tau(t+\bar{d})\big)\big)\ \wedge$
        $\wedge\ \Big(\hat{\mathcal{S}}_{\hat{l}}\cap[v-\bar{d}/\tau,v+\bar{d}/\tau]=\emptyset\Big)$
      $\hat{\mathcal{S}}_{\hat{l}}=\hat{\mathcal{S}}_{\hat{l}}\cup v$;
    end
  end
end
$\hat{l}=|\hat{\mathcal{S}}_{\hat{l}}|$;
$\{\hat{v}_1,\ldots,\hat{v}_{\hat{l}}\}=\hat{\mathcal{S}}_{\hat{l}}$;
Step 2: Updating the velocities and estimating the weights
for $i=1:\hat{l}$
  $\hat{\gamma}_i=0$;
  for $\tau=1:\bar{T}$
    if $\Big(\big(\hat{\mathcal{S}}_{\hat{l}}\cap[\hat{v}_i-\bar{d}/\tau,\hat{v}_i+\bar{d}/\tau]=\hat{v}_i\big)\ \wedge$
      $\wedge\ \big(\max(\hat{r}_\tau(\hat{v}_i\tau-\bar{d}),\ldots,\hat{r}_\tau(\hat{v}_i\tau+\bar{d}))>\hat{\gamma}_i^2\big)\Big)$
      $\hat{\gamma}_i=(\max(\hat{r}_\tau(\hat{v}_i\tau-\bar{d}),\ldots,\hat{r}_\tau(\hat{v}_i\tau+\bar{d}))^{1/2}$;
      $\hat{v}_i=(\arg\max(\hat{r}_\tau(\hat{v}_i\tau-\bar{d}),\ldots,\hat{r}_\tau(\hat{v}_i\tau+\bar{d})))/\tau$;
    end
  end
end
$[\hat{\gamma}_1,\ldots,\hat{\gamma}_{\hat{l}},\hat{v}_1,\ldots,\hat{v}_{\hat{l}}]=\text{sort}(\hat{\gamma}_1,\ldots,\hat{\gamma}_{\hat{l}},\hat{v}_1,\ldots,\hat{v}_{\hat{l}})$;
if $\big(\sum_{i=1}^{\bar{l}}\hat{\gamma}_i^2>1\big)$
  $l_{tmp}=1$;
  while $\big(\sum_{i=1}^{\bar{l}_{tmp}}\hat{\gamma}_i^2<1\big)$
    $l_{tmp}=l_{tmp}+1$
  end
  if $(l=1)\vee\big((\sum_{i=1}^{\bar{l}_{tmp}}\hat{\gamma}_i^2-1)<(1-\sum_{i=1}^{\bar{l}_{tmp}-1}\hat{\gamma}_i^2)\big)$
    $\hat{l}=l_{tmp}$;
    $\hat{\mathcal{S}}_{\hat{l}}=\{\hat{v}_1,\ldots,\hat{v}_{l_{tmp}}\}$;
    $\{\hat{\gamma}_1,\ldots,\hat{\gamma}_{\hat{l}}\}=\{\hat{\gamma}_1,\ldots,\hat{\gamma}_{\bar{l}_{tmp}}\}/(\sum_{i=1}^{\bar{l}_{tmp}}\hat{\gamma}_i^2)$;
  else
    $\hat{l}=l_{tmp}$-1;
    $\hat{\mathcal{S}}_{\hat{l}}=\{\hat{v}_1,\ldots,\hat{v}_{l_{tmp}-1}\}$;
  end
end

Where sort$(\cdot)$ is a function which sorts the detected layers in decreasing energy order. Notice that Algorithm 2 reduces to Algorithm 1 when the sample covariances are substituted with the exact covariances.

Notice that the sum of the layer energies can be less than 1, this because the algorithm admits, in the finite number of samples case, that not all the layers have been detected.

5. SIMULATIONS

5.1 1D simulations

Even if our goal is that of applying the proposed algorithm to a $2D$ signal, we first present a $1D$ example to provide some intuition on the obtained results.

We consider the case of the following values for the parameters: $l=4$, $v_1=-3.13$, $v_2=-5.745$, $v_3=-7.42$, $v_4=-8.1$, $\gamma_1^2=0.31$, $\gamma_2^2=0.3$, $\gamma_3^2=0.2$, $\gamma_4^2=0.19$. The system is simulated for $N=5000$ temporal instants. Fig. 3 shows the estimated $\hat{r}_\tau(\cdot)$, $\tau=\{0,3,6\}$. In Table 1 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 has detected the correct number of layers, i.e. $\hat{l}=l$. The velocities are written in [pixels/frame].

Table 1. Detection of the layers.

	1^{st} layer	2^{nd} layer	3^{rd} layer	4^{th} layer
v_i	−3.125	−5.75	−7.375	−8.143
$\hat{v}_i$	−3.12	−5.75	−7.38	−8.14
γ_i^2	0.31	0.3	0.2	0.19
$\hat{\gamma}_i^2$	0.319	0.282	0.183	0.216

5.2 2D simulations

Since usually the layers move slowly over the telescope pupil here we consider three examples of layer detection where the layers translate less than a pixel per frame. To

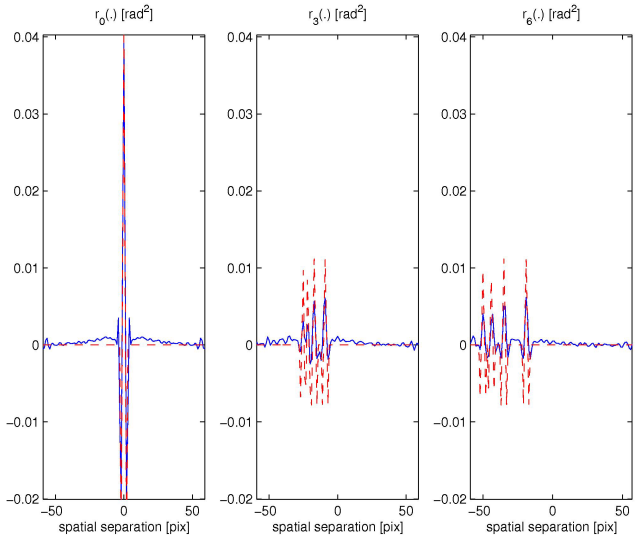


Fig. 3. In blue solid line $\hat{r}_\tau(\cdot)$, $\tau = \{0, 3, 6\}$ estimated by the sample covariances. In red dashed line the detected layers.

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 these simulations we set the simulation parameters to $\sigma_w = 0$ and $N = 5000$.

The results of our simulations are reported in Table 2,3,4: $v_{i,u}$, $v_{i,v}$ and γ_i correspond to the true values of the parameters, $\hat{v}_{i,u}$, $\hat{v}_{i,v}$ and $\hat{\gamma}_i$ are the estimated ones. The velocities are written in [pixels/frame].

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

Table 2. Detection of the layers.

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

Table 3. Detection of the layers.

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

6. CONCLUSIONS

In this paper we have presented a new approach for the detection of atmospheric turbulence layers.

Exploiting a MRF representation of the turbulent phase we have estimated the temporal cross-covariances of the spatial innovation. The proposed algorithm properly analyzes these cross-covariances and extracts the information about the turbulent phase structure, i.e. the number of layers and their characterizing parameters.

Table 4. Detection of the layers.

	1^{st} layer	2^{nd} layer	3^{rd} layer	4^{th} layer	5^{th} layer
$v_{i,u}$	0.13	−0.071	0	0	0.08
$v_{i,v}$	0	0	0.056	−0.087	0
γ_i^2	0.27	0.23	0.2	0.16	0.14
$\hat{v}_{i,u}$	0.1321	−0.0714	0	0	0.08
$\hat{v}_{i,v}$	0	0	0.0556	−0.0877	0
$\hat{\gamma}_i^2$	0.2505	0.2332	0.1968	0.1620	0.1576

With Proposition 1 we claim that asymptotically our algorithm correctly estimates the turbulence parameters. The computational cost of the algorithm is mainly due to the estimation of the covariances: Since the algorithm works properly only with a good estimation of $C_{\xi,\tau}$, $\tau = 1, \ldots, \bar{T}$, a quite large number of samples is needed to estimate them.

The results obtained in this paper can be exploited to improve the performances of the overall system: It should be of particular interest the integration of this procedure in a MCAO system and its use for temporal prediction of the turbulent phase.

ACKNOWLEDGEMENTS

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

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.

D.L. Fried. Statistics of a geometric representation of wavefront distortion. *JOSA*, 55:1427–1435, 1965.

R. Kinderman and J.L. Snell. *Markov Random Fields and their Applications.* American Mathematical Society, 1980.

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

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. *Journal of the Optical Society of America A*, 21(7):1261–1276, 2004.

F. Roddier. *Adaptive optics in astronomy.* Cambridge university press, 1999.

F. Roddier. The effects of atmospheric turbulence in optical astronomy. *Progress in Optics*, 19:281–376, 1981.

T. Soderstrom. *Discrete-time stochastic systems.* Springer, 1994.

J.W. Woods. Two-dimensional discrete markovian fields. *IEEE Trans. on Information Theory*, 18(2):232–240, 1972.