Optimization of adaptive optics correction during observations: Algorithms and system parameters identification in closed-loop

Clémentine Béchet, Michel Tallon and Éric Thiébaut

Université Lyon 1, Villeurbanne, F-69622, France
Centre de Recherche Astrophysique de Lyon, Observatoire de Lyon, 9 avenue Charles André, Saint-Genis Laval cedex, F-69561, France
CNRS, UMR 5574; Ecole Normale Supérieure de Lyon, Lyon, France

ABSTRACT

The adaptive optics (AO) on the European Extremely Large Telescope, as well as earlier pathfinders like the Adaptive Optics Facility, at the Very Large Telescope in 2014, will no longer be stationary systems. AO is no longer isolated on a bench; some elements are directly in the optical train of the telescope, suffering environment and constrains changes during the observations. To guarantee good performance at any observing time, we investigate a self-calibration strategy. We focus here on one of the most challenging aspects: the identification of system parameters during closed-loop observations without introducing any additional disturbance. Such problem is known in the identification theory to be difficult to solve. We have recently presented (Béchet *et al.*, AO4ELT2 Conference, 2011) an identification method for this, with promising results obtained in simulations. To consolidate these advances, we come back in the present paper to the equations and provide a theoretical analysis to justify the choice of the algorithm. We highlight the benefit of using incremental data and commands to decorrelate the disturbance. We also present 2 implementations of the method, currently studied at the European Southern Observatory.

Keywords: identification, optimization, adaptive optics, parameters estimation, closed-loop, telemetry, maximum likelihood

1. INTRODUCTION

Current adaptive optics (AO) usually work fine after a calibration of the system which remains valid during a long-exposure or even several nights of observations. This is the case when the AO components, like the deformable mirrors and the wavefront sensors, are gathered on an isolated AO bench. Nevertheless, a new generation of AO system appears which require much more frequent calibrations: the AO systems of the *adaptive telescopes*. By adaptive telescope, one means a telescope in which a deformable mirror is directly settled in the optical train of the telescope, as for instance when a deformable secondary mirror (DSM) is used.

The Multi Mirror Telescope (MMT) and the Large Binocular Telescope (LBT) are two telescopes already observing with such a deformable secondary mirror.[1] The European Extremely Large Telescope (E-ELT), designed by the European Southern Observatory (ESO), is also expected to be an adaptive telescope, with its fourth mirror being a deformable mirror for the AO. And before that, ESO will install the Adaptive Optics Facility (AOF), which includes a DSM, in 2014 at the Very Large Telescope (VLT) in Chile.[2] The VLT-AOF design, without intermediate focus, will prevent calibrating the interaction matrix (IM) of the AO systems using an internal reference source. This leads ESO to develop new calibration strategies, such as the use of a pseudo-synthetic IM and on-sky calibration of its long-term fixed parameters. A report on these preparatory activities is presented in Kolb *et al.*[3]

Further author information: (Send correspondence to Clémentine Béchet)
E-mail: Clementine.Bechet@univ-lyon1.fr, Telephone: +33 (0) 4 78 86 85 37

Adaptive Optics Systems III, edited by Brent L. Ellerbroek, Enrico Marchetti, Jean-Pierre Véran, Proc. of SPIE Vol. 8447, 84472C · © 2012 SPIE · CCC code: 0277-786/12/$18 · doi: 10.1117/12.926574

The VLT-AOF design is a key feature of the GALACSI AO system feeding the multi Unit Spectroscopic Explorer (MUSE), a panoramic integral field spectrograph for the detection of young galaxies in the visible. To reach such a goal, MUSE[4] and GALACSI[2] will accumulate exposures up to hundred of hours, these exposures being cut into one-hour shots. Reaching 5 to 10% of Strehl in the visible with GALACSI for MUSE during one-hour will require a very stable performance of the AO correction. Therefore, in addition to efficient on-sky calibration strategies, we need a constant optimization of the AO interaction matrix in order to track changes in the system parameters. For instance, misregistrations between the deformable mirror and the wavefront sensors of the AO may progressively appear during exposures, induced by gravity or temperature variations. The present paper studies an on-sky identification method of AO interaction matrix parameters during closed-loop correction. To track the changes of the systems, such estimation of parameters needs to be done with approximately a minute-scale period.

The need for identification of interaction matrix parameters, like shifts or rotation misregistrations, has been already enhanced thanks to simulations of AOF-like systems.[5] This study highlighted loss of performance or even instabilities when shifts misregistrations were beyond $\sim$ 15-20% of a subaperture size.[5] A method for identification of the misregistrations parameters such as differential shifts and rotation between the DSM and the wavefront sensor has been presented and successfully used in simulations of a simplified AO system.[5] However, the accuracy of this estimation method has not been clearly analyzed so far. Its accuracy, its convergence speed, its robustness and its computational complexity need to be addressed. These characteristics may depend on the AO signal-to-noise ratio, on the turbulence conditions (*e.g.* Fried parameter r_0, coherence time τ_0, wind specific directions), on the length of the telemetry data used for estimation, on the AO configuration (*e.g.* Ground Layer correction *GLAO*, Laser Tomography AO *LTAO*) and even on the reconstruction and control methods.

The aim of the current paper is to address these dependencies and their impact from a theoretical analysis of the estimation method. This is important because such identification method will be efficiently used on the real AO systems if and only if we understand how it works, when it does not work, and how well it can improve the AO performance. In that purpose, the Section 2 reminds the equations and the notations used to represent the closed-loop system of the AOF. In Section 3, we introduce a first approach for parameters identification based on the measurement equation of the closed-loop system. A theoretical and numerical analysis enhances the strong correlation of the residuals in this approach, and so its difficult use in practice. Therefore, in Sect. 4, we introduce the criterion already used for identification in Béchet *et al.*,[5] based on increments of the measurement equation. Again, a theoretical and numerical analysis shows the benefit of this approach to obtain better decorrelation of the residuals. Finally, in Sect. 5, two different implementations of this method based on increment of measurements are discussed.

2. AO EQUATIONS

The closed-loop AO system is modeled by the set of two equations

$$\boldsymbol{d}_k = \mathbf{S}(w_k) - \mathbf{G} \cdot \boldsymbol{a}_k + \boldsymbol{e}_k \quad (1)$$

$$\boldsymbol{a}_{k+\tau} = \alpha\, \boldsymbol{a}_{k+\tau-1} + \beta\, \boldsymbol{a}_{k+\tau-2} + \gamma\, \mathbf{C} \cdot \boldsymbol{d}_k \quad (2)$$

where

- $\boldsymbol{d}_k$ is the measurement vector (Shack-Hartmann slopes coming from n_s guide stars) at $k-$th loop
- w_k is a continuous representation of the turbulence in the atmosphere volume (tomography is possible) during measurement at $k-$th loop
- $\boldsymbol{a}_k$ is the DM command vector applied during measurement at $k-$th loop
- $\boldsymbol{e}_k$ is the noise vector on measurement at $k-$th loop
- $\mathbf{S}$ is a linear model of the multi guide stars Shack-Hartmann sensor, from a continuous representation of the turbulence to a discrete set of data

- $\mathbf{G}$ is the command-to-measurement linear model, *i.e.* interaction matrix
- $\mathbf{C}$ is the command matrix
- α, β and γ are the scalar parameters of the control law, *i.e.* in case of a pure integrator law : $\alpha = 1$, $\beta = 0$, and γ is the gain
- τ is the AO delay (number of frames)

Note that index k for discrete time vectors is not chosen as commonly used in AO, but this aims at simplifying equations notations below. Hence, a measurement vector $\boldsymbol{d}_k$ has been influenced by the simultaneous application of the command vector $\boldsymbol{a}_k$ to the DM.

The aim of the identification method investigated here is to estimate parameters on which $\mathbf{G}$, the interaction matrix, depends, using a sequence of data $\boldsymbol{d}_k$ and commands $\boldsymbol{a}_k$ recorded during closed-loop AO, *i.e.* from the AO telemetry. If $\mathbf{G}$ can be determined using a model, that is to say a synthetic matrix, only a few parameters have to be identified, to fully determine the matrix $\mathbf{G}$. We could also consider that all the matrix coefficients of $\mathbf{G}$ are parameters to be identified. To keep general notations, and allow both considerations, we denote the interaction matrix as $\mathbf{G}(\boldsymbol{p})$, that is to say as a function of a vector of parameters $\boldsymbol{p}$. These are the parameters we want to estimate hereafter. Note that there is no linearity assumption between the parameters and the interaction matrix coefficients in our study.

3. CRITERION FROM THE MEASUREMENT EQUATION

The measurement equation (1) can also be written

$$\boldsymbol{d}_k = -\mathbf{G}(\boldsymbol{p}) \cdot \boldsymbol{a}_k + \boldsymbol{z}_k \tag{3}$$

where $\boldsymbol{z}_k = \mathbf{S}(w_k) + \boldsymbol{e}_k$ is now considered as a disturbance vector, including both turbulence contribution to the measurements and measurement noise. From this new point of view, in Eq. (3), the data $\boldsymbol{d}_k$ are considered to be noisy measurements of the slopes induced by the command vector $\boldsymbol{a}_k$ applied to the DM.

The disturbance, $\boldsymbol{z}_k$, follows zero-mean Gaussian statistics and its covariance matrix can be written

$$\mathbf{C}_{\mathbf{z_k}} = \left\langle \boldsymbol{z}_k \cdot {\boldsymbol{z}_k}^{\mathrm{T}} \right\rangle = \left\langle \boldsymbol{s}_k \cdot \boldsymbol{s}_k^{\mathrm{T}} \right\rangle + \left\langle \boldsymbol{e}_k \cdot \boldsymbol{e}_k^{\mathrm{T}} \right\rangle = \mathbf{C}_{\mathbf{s_k}} + \mathbf{C}_{\mathbf{e_k}} \tag{4}$$

where $\boldsymbol{s}_k = \mathbf{S}(w_k)$ is the noiseless Shack-Hartmann measurement of the turbulence w_k. Turbulence and noise are considered to be stationary process at the scale of our estimation study, so that the covariance matrices of $\boldsymbol{s}_k$, $\boldsymbol{e}_k$ and $\boldsymbol{z}_k$ are independent of k, and eventually written $\mathbf{C}_\mathbf{s}$, $\mathbf{C}_\mathbf{e}$ and $\mathbf{C}_\mathbf{z}$ respectively, so that Eq. (4) becomes

$$\mathbf{C}_\mathbf{z} = \mathbf{C}_\mathbf{s} + \mathbf{C}_\mathbf{e}\,. \tag{5}$$

3.1 Criterion

From Eq. (3), a maximum likelihood approach to estimate the parameters $\boldsymbol{p}$ of the interaction matrix $\mathbf{G}$ leads to the minimization of the following criterion

$$\chi_1^2(\boldsymbol{p}) = \left(\boldsymbol{d}_k + \mathbf{G}(\boldsymbol{p}) \cdot \boldsymbol{a}_k\right)^{\mathrm{T}} \cdot \mathbf{C}_z^{-1} \cdot \left(\boldsymbol{d}_k + \mathbf{G}(\boldsymbol{p}) \cdot \boldsymbol{a}_k\right), \tag{6}$$

in case of using only one measurement vector $\boldsymbol{d}_k$. The lower indice 1 of the χ_1^2 stands for using only one vector of measurements and commands.

Selecting several sets of measurements separated by a certain number of frames, ΔT, large enough to avoid correlations between the associated $\boldsymbol{z}_k$ and $\boldsymbol{z}_{k+\Delta T}$, more residuals can be added in the χ^2 computation, and the criterion can be changed into

$$\chi_{N,\Delta T}^2(\boldsymbol{p}) = \sum_{k=1}^{k=N} \left(\boldsymbol{d}_{k\,\Delta T} + \mathbf{G}(\boldsymbol{p}) \cdot \boldsymbol{a}_{k\,\Delta T}\right)^{\mathrm{T}} \cdot \mathbf{C}_z^{-1} \cdot \left(\boldsymbol{d}_{k\,\Delta T} + \mathbf{G}(\boldsymbol{p}) \cdot \boldsymbol{a}_{k\,\Delta T}\right) \tag{7}$$

Note that this estimation differs from recent innovations for on-sky calibration[6,7] by the fact that we cannot choose here the applied commands $\boldsymbol{a}_k$. Furthermore, these applied commands $\boldsymbol{a}_k$ act to counter the turbulence w_k. A good AO correction thus means a significant correlation between command, $\boldsymbol{a}_k$, and disturbance, $\boldsymbol{z}_k$. Overall, it is important to note that, contrary to the measurement noise $\boldsymbol{e}_k$ in a conventional AO, the disturbance $\boldsymbol{z}_k$ here is correlated from one subaperture to another (as well as from one guide star measurement to another guide star measurement and from one measurement to the next in time) because it includes the turbulence contribution $\boldsymbol{w}_k$.

The best parameters $\boldsymbol{p}^*$ are thus obtained solving

$$\boldsymbol{p}^* = \arg\,\min_{\boldsymbol{p}} \quad \chi^2_{N,\Delta T}(\boldsymbol{p})\,, \tag{8}$$

taking into account the statistics of $\boldsymbol{z}_k$ in $\mathbf{C}_z$. The purpose of Sec. 3.2, below, is to provide an analysis of this covariance matrix of the disturbance, in order to understand the existing correlations in the residuals of this criterion.

3.2 Covariance expression for $\mathbf{C_z}$

From Eq. (5), $\mathbf{C_z}$ requires the expression of $\mathbf{C_s}$ and $\mathbf{C_e}$. An analytical expression for the covariance of Shack-Hartmann slopes $\mathbf{C_s}$, in square radians, can be deduced from[8] or[9] for instance. For clarity reasons, we only write in this section the details of the mathematical derivation for covariance of slopes issued from the same Shack-Hartmann (same guide star). It can be written

$$\langle \boldsymbol{s}_x(i)\cdot\boldsymbol{s}_x(j)\rangle = \frac{1}{2\,S^2}\,\frac{\partial^2 D_\phi}{\partial x^2}(\boldsymbol{r}_{ij})\otimes\Pi(\boldsymbol{r}_{ij})\otimes\Pi(\boldsymbol{r}_{ij}) = \mathcal{R}_{Sxx}(\boldsymbol{r}_{ij}) \tag{9}$$

$$\langle \boldsymbol{s}_x(i)\cdot\boldsymbol{s}_y(j)\rangle = \frac{1}{2\,S^2}\,\frac{\partial^2 D_\phi}{\partial x\,\partial y}(\boldsymbol{r}_{ij})\otimes\Pi(\boldsymbol{r}_{ij})\otimes\Pi(\boldsymbol{r}_{ij}) = \mathcal{R}_{Sxy}(\boldsymbol{r}_{ij}) \tag{10}$$

$$\langle \boldsymbol{s}_y(i)\cdot\boldsymbol{s}_y(j)\rangle = \frac{1}{2\,S^2}\,\frac{\partial^2 D_\phi}{\partial y^2}(\boldsymbol{r}_{ij})\otimes\Pi(\boldsymbol{r}_{ij})\otimes\Pi(\boldsymbol{r}_{ij}) = \mathcal{R}_{Syy}(\boldsymbol{r}_{ij})\,, \tag{11}$$

where $\boldsymbol{s}_x(i)$ and $\boldsymbol{s}_y(i)$ are the modeled Shack-Hartmann slopes along x and y directions respectively in the ith subaperture (radians), $\boldsymbol{r}_{ij}$ is the vector of distance between the centers of ith and jth subapertures (meters), D_ϕ is the structure function of the phase on the pupil in this guide star direction and Π is the subaperture windowing function, equal to 1 in the $[-\sqrt{S}/2;\sqrt{S}/2]\times[-\sqrt{S}/2;\sqrt{S}/2]$ square and equal to 0 elsewhere, S being the subaperture square area (m^2) and $\otimes$ denoting the convolution product. Such model of the Shack-Hartmann slope measurement takes into account the slope averaging over the subaperture area, such that x-slope in the ith subaperture can be written, in radians, as

$$\boldsymbol{s}_x(i) = \frac{1}{S}\int\int \frac{\partial\phi}{\partial x}(\boldsymbol{u})\,\Pi(\boldsymbol{u}-\boldsymbol{r}_i)\,d\boldsymbol{u}\,, \tag{12}$$

where $\boldsymbol{r}_i$ is the 2D-coordinate of the center of the $i-$th subaperture.

We recall that the structure function of the wavefront D_ϕ, in rad^2, is

$$D_\phi(r) = \begin{cases} 6.88\,(r/r_0)^{5/3} & \text{if} \quad r_0/L_0 = 0 \\ \alpha\,\left(\frac{L_0}{r_0}\right)^{5/3}\left[2^{5/6}\,\Gamma\left(\frac{5}{6}\right) - 2\,\left(2\,\pi\,\frac{r}{L_0}\right)^{5/6} K_{5/6}\left(2\,\pi\,\frac{r}{L_0}\right)\right] & \text{otherwise} \end{cases} \tag{13}$$

where $K_{5/6}$ is the modified Bessel function of the 3rd kind and of order 5/6 (Mac Donald's function) and $\alpha = [12/5\,\Gamma(6/5)]^{5/6}\,\Gamma(11/6)/\pi^{8/3} \simeq 0.0858$. L_0 and r_0 are respectively the outer scale and the Fried parameter of the atmospheric turbulence. The first expression in Eq. (13) stands for Kolmogorov model of turbulence and the second line stands for von Kármán model.

Some approximations for the covariance of the Shack-Hartmann slopes already exist in the literature for Kolmogorov model.[8] Another approximation of the slopes variance can be obtained, for Kolmogorov and von

Kármán turbulence models, following the Fried model for the Shack-Hartmann and based on phase differences between the corners of a subaperture. Such approximation provides slightly different numerical values than the analysis of Equations (9)-(11), due to the approximation of the Fried model neglecting the contribution of higher frequencies than the cut-off frequency of the lenslet array. This is why in the following, we prefer to develop our analysis based on Equations (9)-(11), in order to study the exact properties of the slopes covariance.

Equations (9)-(11) are not easy to manipulate and to analyze, but the convolution product being a product in Fourier domain, the elements of the covariance matrix are more easily studied below in Fourier domain. The details of the derivation of the covariance in the Fourier space are gathered in Appendix A.

From Fourier-Domain formulation of covariance maps, we can come back to $(x, y)-$domain and represent the covariance map of slopes for a Shack-Hartmann. Plots of covariance maps of Shack-Hartmann slopes in

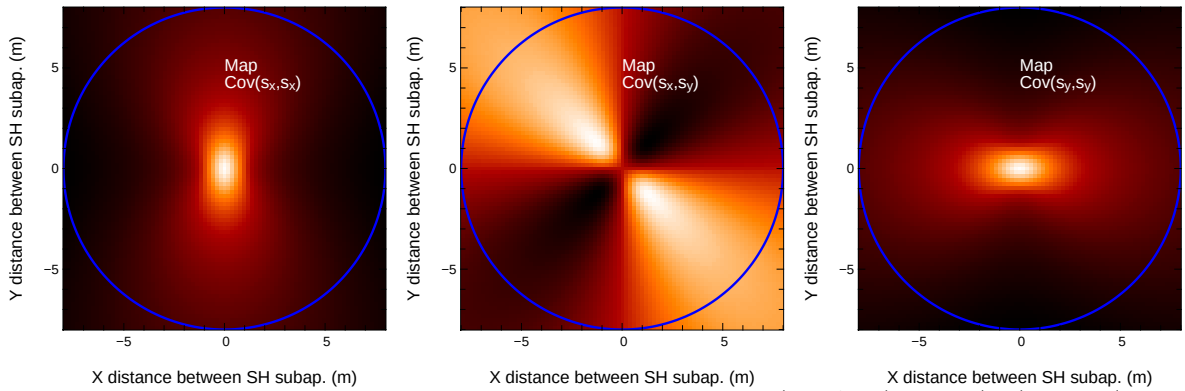

Figure 1. Covariance map of a 40×40 shack-Hartmann on an 8m-telescope (like AOF): $\mathcal{R}_{Sxx}$ (left), $\mathcal{R}_{Sxy}$ (center), and $\mathcal{R}_{Syy}$ (right). The inside of the circle contains the values that are effectively present in a covariance matrix of slopes for a pupil diameter of 8 m.

$(x, y)-$space for a Shack-Hartmann of subaperture size $\sqrt{S} = 0.2$ m over an 8 m-pupil (as in the AOF case) can be numerically obtained, as shown in Fig. 1.

From Eqs. (39)-(41) in Appendix A, and the linearity of the Fourier transform, we can see that $\mathcal{R}_{Sxx}$, $\mathcal{R}_{Sxy}$ and $\mathcal{R}_{Syy}$ are proportional to $r_0^{-5/3}$ via the turbulence spectrum $\mathcal{S}_\phi$. In addition, the dependence on L_0 is only visible at low frequencies ($< 1/L_0$). In order to understand how much the slopes are correlated over the pupil, for every guide star, we represent in Fig. 2 cuts of the left and center maps of Fig. 1, where they exhibit the highest correlation. For the left map of Fig. 1, the cut is made for $x = 0$ and $y \geq 0$ and presented on the left plot of Fig. 2 for various r_0 values. For central map of Fig. 1, the cut is made for $x = -y \geq 0$ and is presented on the right plot of Fig. 2 for the same r_0 values.

The values observed in Fig. 2 allow to estimate the order of magnitude of the $\mathbf{C}_s$ contribution to $\mathbf{C}_z$ matrix, compared to the measurement noise part $\mathbf{C}_e$. In the context of the AOF, with the laser guide stars and the wavefront sensors design, the measurement noise has been estimated in[5] to be of the order of 50 to 150 nm rms between two subaperture sides. This means that $\mathbf{C}_s$ contribution to the total variance of $\mathbf{C}_z$ in more than 50% rms even in the optimistic case of an r_0 value of 16 cm. It also shows that the covariance matrix $\mathbf{C}_z$ could not be reasonably approximated by a diagonal, because this contribution from $\mathbf{C}_s$ to the non diagonal terms of $\mathbf{C}_z$ is large in comparison with the variance value. As a matter of fact, the covariance presented in the left plot of Fig. 2 decreases very slowly with the distance r between 2 subapertures. The covariance is equal to the variance divided by 2 only for r beyond 1.2 m (6 subapertures), and divided by 4 for r beyond 2 m (10 subapertures).

As a conclusion, the covariance matrix $\mathbf{C}_z$ has a complex structure, and is far from being diagonal. A good estimator based on the solving of Eq. (8) would need to take these noise correlations into account. All this leads us to derive another measurement model in Section 4, in order to overcome these difficulties.

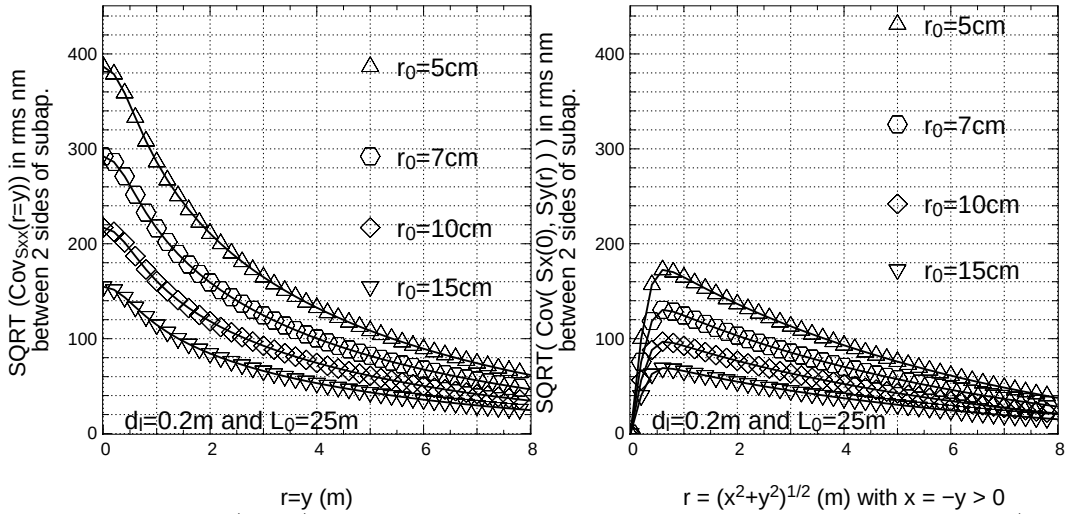

Figure 2. Left plot : cut of $\mathrm{Cov}(\boldsymbol{s}_x, \boldsymbol{s}_x)$ map in Fig. 1 along $x = 0$ and $y \geq 0$. Right plot : cut of $\mathrm{Cov}(\boldsymbol{s}_x, \boldsymbol{s}_y)$ map in Fig. 1 along $x = -y \geq 0$. For both plots, the square root of the covariance is represented instead of the covariance, and it is expressed in nm rms between 2 sides of a subaperture. Various colors are used for different r_0 rescalings. The markers stand for the values effectively appearing in the covariance matrix $\mathbf{C_s}$.

4. CRITERION FROM INCREMENTS OF THE MEASUREMENT EQUATION

Instead of using Eq (3), we use now the difference between two successive measurements,

$$\delta \boldsymbol{d}_k = \boldsymbol{d}_{k+1} - \boldsymbol{d}_k = -\mathbf{G} \cdot \delta \boldsymbol{a}_k + \delta \boldsymbol{z}_k \tag{14}$$

where $\delta \boldsymbol{a}_k = \boldsymbol{a}_{k+1} - \boldsymbol{a}_k$ and $\delta \boldsymbol{z}_k = \boldsymbol{z}_{k+1} - \boldsymbol{z}_k$.

The idea of such approach is to transform the estimation problem keeping its linearity and Gaussian statistics properties, with a disturbance covariance matrix $\mathbf{C}_{\delta z}$ which could be better approximated by a diagonal matrix and decorrelating (approximately) signal ($\mathbf{G} \cdot \delta \boldsymbol{a}_k$) and disturbance ($\delta \boldsymbol{z}_k$).

In the same way as in Sec. 3, we justify this approach with an analysis of the covariance matrix of the disturbance, here $\mathbf{C}_{\delta z}$. The disturbance $\delta \boldsymbol{z}_k$ follows a zero-mean Gaussian statistics of covariance matrix

$$\mathbf{C}_{\delta z} = \left\langle (\boldsymbol{z}_{k+1} - \boldsymbol{z}_k) \cdot (\boldsymbol{z}_{k+1} - \boldsymbol{z}_k)^{\mathrm{T}} \right\rangle = \left\langle \mathbf{S}(\delta w_k) \cdot \mathbf{S}(\delta w_k)^{\mathrm{T}} \right\rangle + \left\langle \delta \boldsymbol{e}_k \cdot \delta \boldsymbol{e}_k^{\mathrm{T}} \right\rangle \tag{15}$$

$$= \mathbf{C}_{\delta s} + 2\,\mathbf{C}_e\,. \tag{16}$$

4.1 Criterion

Using only one set of incremental measurements $\delta \boldsymbol{d}_k$, one can fit the interaction matrix parameters $\boldsymbol{p}$ using maximum likelihood approach, *i.e.* minimizing

$$\chi_1^2(\boldsymbol{p}) = (\delta \boldsymbol{d}_k + \mathbf{G}(\boldsymbol{p}) \cdot \delta \boldsymbol{a}_k)^{\mathrm{T}} \cdot \mathbf{C}_{\delta z}^{-1} \cdot (\delta \boldsymbol{d}_k + \mathbf{G}(\boldsymbol{p}) \cdot \delta \boldsymbol{a}_k)\,, \tag{17}$$

where the lower indice 1 of the χ_1^2 stands for using only one set of incremental measurements and commands. Selecting several sets of incremental measurements separated by a certain number of frames, ΔT, large enough to avoid correlations between the associated $\delta \boldsymbol{z}_k$ and $\delta \boldsymbol{z}_{k+\Delta T}$, more residuals can be added in the χ^2 computation, and the criterion can be changed to

$$\chi_{N,\Delta T}^2(\boldsymbol{p}) = \sum_{k=1}^{k=N} (\delta \boldsymbol{d}_{k\,\Delta T} + \mathbf{G}(\boldsymbol{p}) \cdot \delta \boldsymbol{a}_{k\,\Delta T})^{\mathrm{T}} \cdot \mathbf{C}_{\delta z}^{-1} \cdot (\delta \boldsymbol{d}_{k\,\Delta T} + \mathbf{G}(\boldsymbol{p}) \cdot \delta \boldsymbol{a}_{k\,\Delta T}) \tag{18}$$

In other words, the method now estimates that the best interaction matrix parameters are

$$\boldsymbol{p}^* = \arg\,\min_{\boldsymbol{p}} \quad \chi_{N,\Delta T}^2(\boldsymbol{p})\,. \tag{19}$$

Equations (18) and (19) characterize the identification method successfully studied in simulations.[5] In order for this method to be efficient, we need to :

1. Know, in a good approximation, $\mathbf{C}_{\delta z}$. Here approximated by $2\,\mathbf{C}_e$.
2. Demonstrate that $\boldsymbol{\delta a}_k$ is not correlated to $\boldsymbol{\delta z}_k$.

These 2 points are critical to be able to claim with how much precision we could estimate the system parameters. The theoretical analysis below clarifies the first one of these two points, with the result of Eq. (20). The second point is still not clearly analyzed from the theory.

4.2 Covariance of incremental Shack-Hartmann slopes

As for the covariance map of th Shack-Hartmann slopes in Sec. 3.2, Fourier space analysis is again used to model $\mathbf{C}_{\delta z}$. The details of the mathematical derivation are gathered in Appendix B. Once the Fourier-domain spectral density maps of the incremental slopes are computed, we can come back to $(x, y)-$space, to get the expected covariance maps over the aperture. We obtain maps like in Figs. 3 estimated for typical configuration of the AOF system at the VLT.

This contains exactly the values that are present in the covariance matrix $\mathbf{C}_{\delta s}$ we are searching for. Some $x-$ and $y-$cuts are done over these maps, and are represented in Figs. 4 and 5. Abscissas are in meters, and ordinates in square nanometers between two side of a subaperture. It has been observed that, when r_0, L_0 and τ_0 change, the shape of these cuts stay similar, and the level does not change much, staying in the range of 5 to 50 nm^2. In particular, Fig. 4 enhance that the reduction of the subaperture size from 57 cm, like for NAOS, to 20 cm, like for the AOF, significantly accelerates the decorrelation of $\boldsymbol{\delta z}$ when considering farther subapertures. From Fig. 5, one can observe the increase of the correlation with the WFS exposure time. However, in any case from 0.5 ms to 2ms of exposure, the covariance quickly drops down after 1 meter of distance.

More time would be required to really get a quantitative evaluation of these covariance coefficients as a function of the atmospheric parameters. Nevertheless, one observes that beyond 1 meter of distance the covariance is below a few nm^2, and even at its maximum values it is below the level of expected measurement noise in the AOF system (∼80 nm rms).

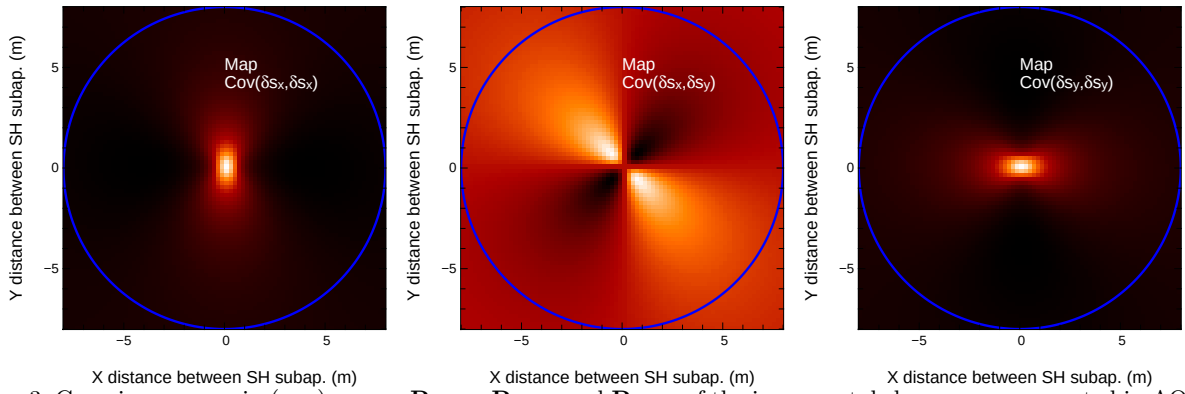

Figure 3. Covariance maps in $(x, y)-$space $\mathbf{R}_{\delta Sxx}$, $\mathbf{R}_{\delta Sxy}$ and $\mathbf{R}_{\delta Syy}$ of the incremental slopes are represented in AOF-like conditions.

These examples of covariance values for the incremental slopes of a Shack-Hartmann provide order of magnitude for the diagonal elements and off-diagonal elements of the covariance matrix $\mathbf{C}_{\delta s}$. It confirms that the contribution of the off-diagonal elements decreases faster than for $\mathbf{C}_s$ with respect to the maximum diagonal ones, when the distance between the subapertures increases. This can be clearly seen in Figs.4, where the decorrelation appears even steeper for subaperture size of $d = 0.2$ m (AOF-like), than for subaperture size of $d = 0.57$ m (NAOS-like). In addition, the amplitude of the diagonal elements of the covariance matrix is also small compared to the expected measurement noise level we mentioned for the AOF system.

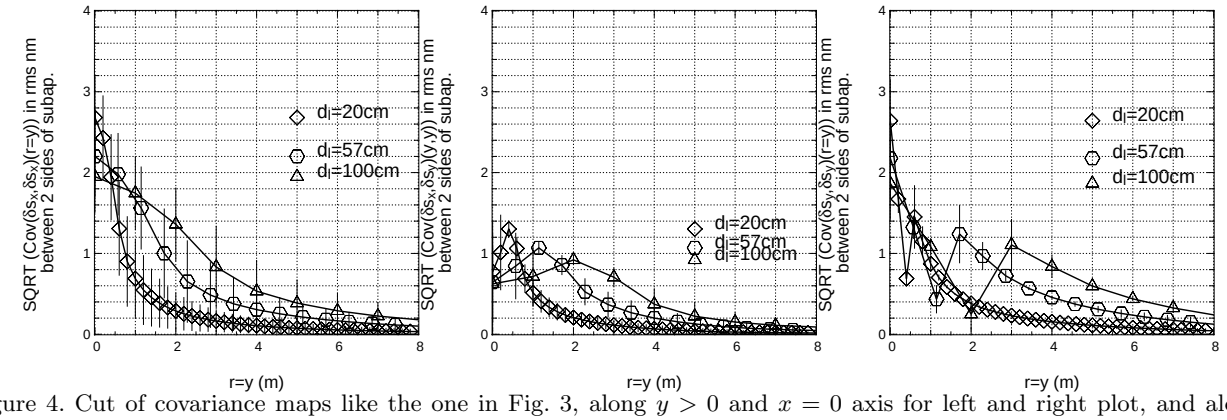

Figure 4. Cut of covariance maps like the one in Fig. 3, along $y > 0$ and $x = 0$ axis for left and right plot, and along $x = y \geq 0$ axis for the plot in the middle. The square root of the Covariance is plotted, in order to express it in nm rms, and to ease the comparison with noise level. WFS exposure time is 1 ms. The curves with different markers stand for different subaperture sizes d_l: 1 m (triangles), 57 cm (hexagons) and 20 cm (diamonds).

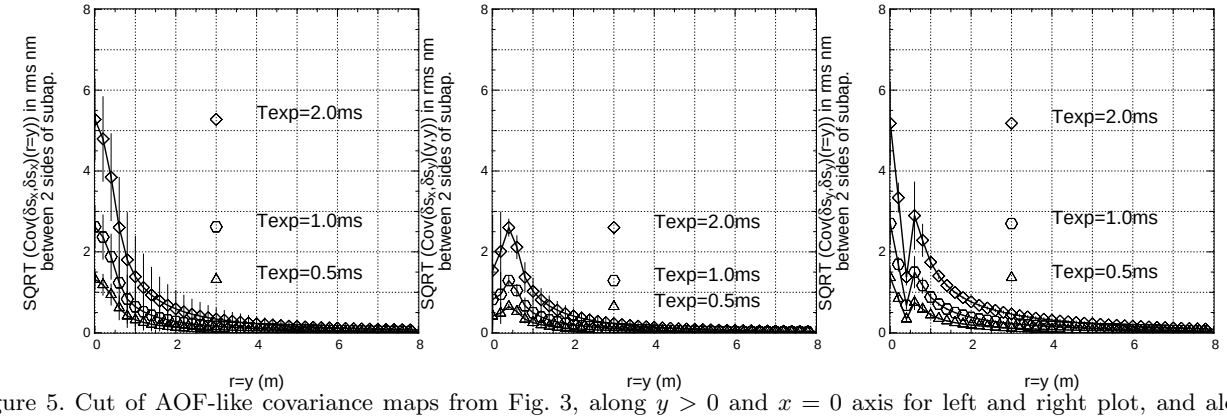

Figure 5. Cut of AOF-like covariance maps from Fig. 3, along $y > 0$ and $x = 0$ axis for left and right plot, and along $x = y \geq 0$ axis for the plot in the middle. The square root of the Covariance is plotted, in order to express it in nm rms, and to ease the comparison with noise level. The curves with different markers stand for different WFS exposure times: 0.5 ms (triangles), 1 ms (hexagons) and 2 ms (diamonds).

In the case of incremental slopes, we can therefore approximate the covariance matrix $\mathbf{C}_{\delta z}$ by twice the diagonal covariance of the measurement noise $\mathbf{C}_e$, *i.e.*

$$\mathbf{C}_{\delta z} \simeq 2\,\mathbf{C}_e\,. \tag{20}$$

Eq. (16) is then replaced by the approximation in Eq. (20). As another example, Fig.6 presents the three matrices $\mathbf{C}_{\delta z}$, $\mathbf{C}_{\delta s}$ and $\mathbf{C}_e$ as numerically estimated from the simulation of a closed-loop AO system. From this simulation, we observe that the non diagonal terms of $\mathbf{C}_{\delta z}$ are about 20 times lower than the diagonal ones. The noise measurement dominates in $\delta \boldsymbol{z}$.

5. IMPLEMENTATIONS

Another theoretical analysis of this identification procedure has been done, which concerns the comparison of different implementations of the identification method introduced in Sec. 4. Various implementations have been used for the AOF study and the aim of this section is to clarify what can be said about these different implementations and the reasons why they can provide identical or different results.

5.1 Direct computation of misregistration parameters

This implementation consists in a model-fitting directly defined by Equations (18) and (19). It is also the implementation used in the simulations presented in Béchet *et al.*[5] The interaction matrix is defined as a

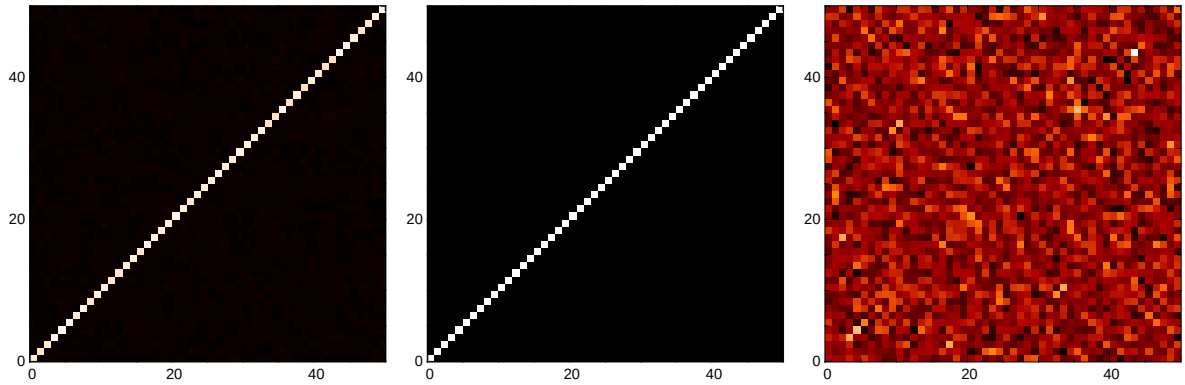

Figure 6. Example of estimation of covariance matrices deduced from an open-loop sequence of Shack-Hartmann measurements of an evolving turbulent atmosphere (only the 50 first rows and columns of the full 2304×2304 matrices on the 40×40 subapertures SH are displayed): $\mathbf{C}_{\delta z}$ (left), $\mathbf{C}_e$ (middle) and $\mathbf{C}_{\delta s}$ (right). The non zero (diagonal) values of $\mathbf{C}_e$ are $\sigma_e^2 = 6800$ mas^2 per subaperture (equivalent to $\sim$80 nm rms between 2 subaperture sides), while the noisy shape of $\mathbf{C}_{\delta s}$ has values of the order of 5% of this, *i.e.* between -300 and $+300$mas^2.The contribution of the right matrix to the covariance $\mathbf{C}_{\delta z}$ on the left is thus quite small.

parametric model $\mathbf{G}(\boldsymbol{p})$, build from what we know about DM influence functions and shack-Hartmann linearized modeling. The criterion of Eq. (18) is directly minimized with respect to the parameters $\boldsymbol{p}$, using a non-linear optimization algorithm, which is a Levenberg-Marquardt algorithm modified to take into account a *trust region*.

5.2 Computation of misregistration parameters after an interaction matrix estimation

The criterion mentioned in Eq. (18) is general, so that it could be applied to some ten parameters like misalignments characteristics for instance, as well as to the full interaction matrix coefficients such that each element of $\boldsymbol{p}$ is a coefficient $\mathbf{G}_{i,j}$.

We could then rewrite the criterion of Eq. (18)

$$\chi^2_{N,\Delta T}(\mathbf{G}) = \frac{1}{2} \sum_{k=1}^{k=N} (\delta \boldsymbol{d}_{k\,\Delta T} + \mathbf{G} \cdot \delta \boldsymbol{a}_{k\,\Delta T})^{\mathrm{T}} \cdot \mathbf{C}_{\delta z}^{-1} \cdot (\delta \boldsymbol{d}_{k\,\Delta T} + \mathbf{G} \cdot \delta \boldsymbol{a}_{k\,\Delta T}) \tag{21}$$

In this last case, the criterion (21) becomes a quadratic function of $\mathbf{G}$, for which the solution

$$\mathbf{G}^* = \arg\,\min_{\mathbf{G}} \quad \chi^2_{N,\Delta T}(\mathbf{G}) \tag{22}$$

is obtained when the first derivative of the criterion is zero, *i.e.*

$$\frac{\partial \chi^2_{N,\Delta T}}{\partial \mathbf{G}} = \mathbf{C}_{\delta z}^{-1} \sum_{k=1}^{k=N} (\delta \boldsymbol{d}_{k\,\Delta T} + \mathbf{G} \cdot \delta \boldsymbol{a}_{k\,\Delta T}) \cdot \delta \boldsymbol{a}_{k\,\Delta T}^{\mathrm{T}} = 0\,. \tag{23}$$

Since $\mathbf{C}_{\delta z}$ is a positive definite covariance matrix (symmetric), its inverse is also symmetric positive definite, and then the solution $\mathbf{G}^*$ verifies

$$\mathbf{G}^* \cdot \sum_{k=1}^{k=N} (\delta \boldsymbol{a}_{k\,\Delta T} \cdot \delta \boldsymbol{a}_{k\,\Delta T}^{\mathrm{T}}) = - \sum_{k=1}^{k=N} (\delta \boldsymbol{d}_{k\,\Delta T} \cdot \delta \boldsymbol{a}_{k\,\Delta T}^{\mathrm{T}})\,. \tag{24}$$

We choose the following notations for approximated estimates of the covariance matrices

$$\mathbf{C}^*_{\delta a,\delta a} = \frac{1}{N} \sum_{k=1}^{N} (\delta \boldsymbol{a}_{k\,\Delta T} \cdot \delta \boldsymbol{a}_{k\,\Delta T}^{\mathrm{T}}) \quad \text{and} \quad \mathbf{C}^*_{\delta d,\delta a} = \frac{1}{N} \sum_{k=1}^{N} (\delta \boldsymbol{d}_{k\,\Delta T} \cdot \delta \boldsymbol{a}_{k\,\Delta T}^{\mathrm{T}})\,. \tag{25}$$

In case $\mathbf{C}^*_{\delta a,\delta a}$ is invertible, then the solution $\mathbf{G}^*$ is

$$\mathbf{G}^* = -\mathbf{C}^*_{\delta d,\delta a} \cdot \mathbf{C}^{*\,-1}_{\delta a,\delta a} \,. \tag{26}$$

In general, it is difficult to ensure that $\mathbf{C}^*_{\delta a,\delta a}$ will be invertible (*e.g.* some actuators may not be commanded or the recorded sequence is not long enough). As a consequence, the inverse above will be replaced by the generalized inverse of $\mathbf{C}^*_{\delta a,\delta a}$, computed by Truncated Singular Value Decomposition (T-SVD). In such case, it is mathematically equivalent to compute the matrix estimate $\mathbf{G}^*$, following the procedure below. This procedure, suggested by J. Kolb,[3] computes the matrix $\mathbf{G}^*$ using the expression

$$\mathbf{G}^* = -\Delta_d \cdot {\Delta_a}^\dagger \tag{27}$$

where the Δ_d and Δ_a matrices consist of concatenation of incremental vectors $\delta\boldsymbol{d}_k$ of measurements and $\delta\boldsymbol{a}_k$ of commands respectively, in N columns, as illustrated below

$$\Delta_d \;=\; [\delta\boldsymbol{d}_{\Delta T} \mid \delta\boldsymbol{d}_{2\Delta T} \mid \delta\boldsymbol{d}_{3\Delta T} \mid ... \mid \delta\boldsymbol{d}_{N\Delta T}] \tag{28}$$

$$\Delta_a \;=\; [\delta\boldsymbol{a}_{\Delta T} \mid \delta\boldsymbol{a}_{2\Delta T} \mid \delta\boldsymbol{a}_{3\Delta T} \mid ... \mid \delta\boldsymbol{a}_{N\Delta T}] \;, \tag{29}$$

and ${\Delta_a}^\dagger$ is a general inverse of Δ_a obtained by Truncated Singular Value Decomposition (T-SVD).

The mathematical equivalence between Eqs. (26) and (27) is proved in AppendixC. In addition, we have checked with our simulations that the two computational approaches provide the same estimate for the interaction matrix $\mathbf{G}^*$.

5.3 Precision on this intermediate interaction matrix estimate

The accuracy of the estimation of the coefficients of the interaction matrix $\mathbf{G}^*$, by the method described above, can be theoretically formulated thanks to the Hessian of the criterion. The coefficients of the inverse of the Hessian define the error bars (diagonal coefficients) and the correlations (off-diagonal coefficients).

More work is required to derive an exact expression for this precision, but in a first approximation it would lead to show that the error bar on coefficient $\mathbf{G}_{i,j}$ is $1/\sigma^2_{\delta aj}$, the inverse of the variance of the incremental commands on the $j-$th actuator.

In other words, the error bars are the same in a given column of the interaction matrix, but it can be different from one column to another. Another remark is that the higher the variance of incremental commands, the better the precision. This explains why when the system tends to become unstable, the estimation of the interaction matrix on sky in closed-loop could be more precise, thanks to some particular highly excited actuators.

6. CONCLUSION

This paper presents the recent advances obtained in the analysis of the identification method for misregistration parameters in AO. Following the promising results previously obtained on simplified simulations,[5] a theoretical analysis of the estimation method, its assumptions and its accuracy needed to be addressed.

Thanks to both analytical and numerical analysis, the paper enhances the benefit of the method based on the *incremental* measurement equation, compared to the one based on the classical measurement equation. Covariances and correlations between signal and disturbance in this case appear easier to compute and to approximate, in order to ensure the use of a reliable estimator.

Two implementations of this identification method are also discussed, and the equivalence between two algorithms to estimate an intermediate estimation of the interaction matrix is proved. This is important since the identification of misregistration parameters requires to be applied in real time, and thus computational aspects must be considered to select an efficient algorithm even for large AO systems.

Finally, the report show perspectives to derive a theoretical analysis of the accuracy through the study of the Hessian matrix. However, more theoretical work is required to reach such a goal.

APPENDIX A. FOURIER-DOMAIN COVARIANCE OF SHACK-HARTMANN SLOPES

We express the Shack-Hartmann slopes covariances of Eqs. (9-11) using their power spectral density in the Fourier domain

$$\mathcal{S}_{Sxx}(\boldsymbol{\kappa}) = \mathcal{F}(\mathcal{R}_{Sxx}) = \frac{1}{2\,S^2}\,\mathcal{F}\left(\frac{\partial^2 D_\phi}{\partial x^2}\right)\left[\mathcal{F}\left(\Pi\right)\right]^2 \tag{30}$$

$$\mathcal{S}_{Sxy}(\boldsymbol{\kappa}) = \mathcal{F}(\mathcal{R}_{Sxy}) = \frac{1}{2\,S^2}\,\mathcal{F}\left(\frac{\partial^2 D_\phi}{\partial x\,\partial y}\right)\left[\mathcal{F}\left(\Pi\right)\right]^2 \tag{31}$$

$$\mathcal{S}_{Syy}(\boldsymbol{\kappa}) = \mathcal{F}(\mathcal{R}_{Syy}) = \frac{1}{2\,S^2}\,\mathcal{F}\left(\frac{\partial^2 D_\phi}{\partial y^2}\right)\left[\mathcal{F}\left(\Pi\right)\right]^2 . \tag{32}$$

We can use the properties of the Fourier Transform to write

$$\mathcal{F}\left(\frac{\partial^2 D_\phi}{\partial x^2}\right) = -4\pi^2\kappa_x^2\mathcal{F}(D_\phi) \tag{33}$$

$$\mathcal{F}\left(\frac{\partial^2 D_\phi}{\partial x\partial y}\right) = -4\pi^2\kappa_x\kappa_y\mathcal{F}(D_\phi) \tag{34}$$

$$\mathcal{F}\left(\frac{\partial^2 D_\phi}{\partial y^2}\right) = -4\pi^2\kappa_y^2\mathcal{F}(D_\phi) \tag{35}$$

$$\mathcal{F}\left(\Pi\right) = S\,\mathrm{sinc}(\pi\sqrt{S}\kappa_x)\,\mathrm{sinc}(\pi\sqrt{S}\kappa_y) \tag{36}$$

$$\mathcal{F}(D_\phi) = 2\delta_{\mathrm{Dirac}}(\boldsymbol{\kappa}) - 2\mathcal{S}_\phi(\boldsymbol{\kappa})\,, \tag{37}$$

where $\mathcal{S}_\phi(\boldsymbol{\kappa})$ is the power spectral density of the phase, in square radians, known as (by Fourier Transform of Eq. (13))

$$\mathcal{S}_\phi(\boldsymbol{\kappa}) = \begin{cases} 0.0229\,r_0^{-5/3}\,\|\boldsymbol{\kappa}\|^{-11/3} & \text{if} \quad r_0/L_0 = 0 \\ 0.0229\,r_0^{-5/3}\left(\|\boldsymbol{\kappa}\|^2 + \frac{1}{L_0^2}\right)^{-11/6} & \text{otherwise} \end{cases} \tag{38}$$

Then Eqs. (30-32) can be simplified into

$$\mathcal{S}_{Sxx}(\boldsymbol{\kappa}) = 4\,\pi^2\kappa_x^2 \quad \mathrm{sinc}^2(\pi\sqrt{S}\kappa_x)\,\mathrm{sinc}^2(\pi\sqrt{S}\kappa_y)\;\mathcal{S}_\phi(\boldsymbol{\kappa}) \tag{39}$$

$$\mathcal{S}_{Sxy}(\boldsymbol{\kappa}) = 4\,\pi^2\kappa_x\kappa_y \quad \mathrm{sinc}^2(\pi\sqrt{S}\kappa_x)\,\mathrm{sinc}^2(\pi\sqrt{S}\kappa_y)\;\mathcal{S}_\phi(\boldsymbol{\kappa}) \tag{40}$$

$$\mathcal{S}_{Syy}(\boldsymbol{\kappa}) = 4\,\pi^2\kappa_y^2 \quad \mathrm{sinc}^2(\pi\sqrt{S}\kappa_x)\,\mathrm{sinc}^2(\pi\sqrt{S}\kappa_y)\;\mathcal{S}_\phi(\boldsymbol{\kappa})\,. \tag{41}$$

These simple expressions are obtained thanks to the fact that the Dirac part of Eq. (37) is never non zero at the same time as the factors κ_x^2, $\kappa_x\,\kappa_y$ and κ_y^2 in Eqs. (39)-(41) .

APPENDIX B. FOURIER-DOMAIN COVARIANCE OF INCREMENTAL SHACK-HARTMANN SLOPES

In the same way we derived slopes covariance maps in the Fourier-Domain in Appendix A, we can express the incremental slopes covariance maps in the Fourier-Domain. In a first step, let consider a single turbulent layer translating horizontally in the pupil plane with wind speed vector $\boldsymbol{V}$, and let us note τ_e the WFS exposure time. Taylor frozen flow assumption is expressed here as

$$\delta\boldsymbol{\phi}_k(\boldsymbol{r}) = \boldsymbol{\phi}_{k+1}(\boldsymbol{r}) - \boldsymbol{\phi}_k(\boldsymbol{r}) = \boldsymbol{\phi}_k(\boldsymbol{r} - \boldsymbol{V}\,\tau_e) - \boldsymbol{\phi}_k(\boldsymbol{r})\,. \tag{42}$$

It is then possible to use similar formulae as Eqs. (30)-(32) but in terms of covariance incremental slopes. Measuring the slope at a posterior instant $\boldsymbol{V}\tau_e$ is equivalent to translate the subaperture center coordinates by $-\boldsymbol{V}\tau_e$. As a consequence,

$$\langle \delta s_x(0)\cdot\delta s_x(\boldsymbol{r})\rangle = 2\,\langle s_x(0)\cdot s_x(\boldsymbol{r})\rangle - \langle s_x(0)\cdot s_x(\boldsymbol{r}-\boldsymbol{V}\tau_e)\rangle - \langle s_x(0)\cdot s_x(\boldsymbol{r}+\boldsymbol{V}\tau_e)\rangle\,. \tag{43}$$

Using the fact that a translation of $\boldsymbol{V}\tau_e$ in $(x,y)-$space is equivalent to a multiplication by $\exp(j\boldsymbol{\kappa}\cdot\boldsymbol{V}\tau_e)$ in Fourier Domain, then

$$\begin{aligned}
\mathcal{S}_{\delta Sxx}(\boldsymbol{\kappa}) &= 4\,\mathcal{S}_{Sxx}(\boldsymbol{\kappa})\,\sin^2(\pi\boldsymbol{\kappa}\cdot\boldsymbol{V}\tau_e) &(44)\\
\mathcal{S}_{\delta Sxy}(\boldsymbol{\kappa}) &= 4\,\mathcal{S}_{Sxy}(\boldsymbol{\kappa})\,\sin^2(\pi\boldsymbol{\kappa}\cdot\boldsymbol{V}\tau_e) &(45)\\
\mathcal{S}_{\delta Syy}(\boldsymbol{\kappa}) &= 4\,\mathcal{S}_{Syy}(\boldsymbol{\kappa})\,\sin^2(\pi\boldsymbol{\kappa}\cdot\boldsymbol{V}\tau_e)\,, &(46)
\end{aligned}$$

where $\mathcal{S}_{Sxx}(\boldsymbol{\kappa})$, $\mathcal{S}_{Sxy}(\boldsymbol{\kappa})$ and $\mathcal{S}_{Syy}(\boldsymbol{\kappa})$ are given by Eqs. (39)-(41). If the wind speed is zero, then obviously the incremental slopes are zero everywhere. The way the Fourier-Domain slopes covariance are weighted by the sine squared in the expression of the Fourier-Domain incremental covariances depends on the wind speed direction $\boldsymbol{V}$. It is a zero factor for frequencies orthogonal to the wind speed, and it increases along with the frequency component collinear to the wind, $\boldsymbol{\kappa}\cdot\boldsymbol{V}$.

The atmosphere can be considered to have several turbulent layers (statistically independent), so that the sum of the contribution of each layer is used to compute the spectral density of the incremental slopes. Assuming n_l layers with the turbulence strength ratio $C^2_{n\,i_l}$, and wind speed $\boldsymbol{V}_{i_l}$ standing for the i_l-th layer and the global atmosphere Fried parameter noted r_0, we can write

$$\mathcal{S}_{\delta Sxx}(\boldsymbol{\kappa}) = 4\,\mathcal{S}_{Sxx}(\boldsymbol{\kappa})\,\sum_{i_l=1}^{n_l}(C^2_{n\,i_l}\,\sin^2(\pi\boldsymbol{\kappa}\cdot\boldsymbol{V}_{i_l}\tau_e))\,, \tag{47}$$

and the same sum factor applies to $\mathcal{S}_{\delta Sxy}(\boldsymbol{\kappa})$ and $\mathcal{S}_{\delta Syy}(\boldsymbol{\kappa})$ in Eqs. (45)-(46). Due to the fact that the wind speed distribution in altitude in the turbulent atmosphere may present various directions, the total sum factor in Eq. (47) is not in general radially symmetric.

APPENDIX C. EQUIVALENCE OF IMPLEMENTATIONS

We demonstrate how to go from Eq. (26) to Eq. (27). From Eq. (26), the $(i,j)-$th element of matrix $\mathbf{G}^*$ can be decomposed as follows:

$$\begin{aligned}
\mathbf{G}^*(i,j) &= -\sum_{k=1}^{n_a}\mathbf{C}^*_{\delta d,\delta a}(i,k)\,\mathbf{C}^{*\,\dagger}_{\delta a,\delta a}(k,j) &(48)\\
&= -\frac{1}{N}\sum_{k=1}^{n_a}\sum_{l=1}^{N}\boldsymbol{\delta d}(i,l)\,\boldsymbol{\delta a}(k,l)\,\mathbf{C}^{*\,\dagger}_{\delta a,\delta a}(k,j) &(49)\\
&= -\frac{1}{N}\sum_{l=1}^{N}\boldsymbol{\delta d}(i,l)\,\left[\sum_{k=1}^{n_a}\boldsymbol{\delta a}(k,l)\,\mathbf{C}^{*\,\dagger}_{\delta a,\delta a}(k,j)\right] &(50)
\end{aligned}$$

In addition, by definition,

$$\mathbf{C}^*_{\delta a,\delta a} = \frac{1}{N}\,(\Delta_a\cdot{\Delta_a}^{\mathrm{T}})\,, \tag{51}$$

so that Eq. (50) becomes

$$\begin{aligned}
\mathbf{G}^*(i,j) &= -\sum_{l=1}^{N}\boldsymbol{\delta d}(i,l)\,\sum_{k=1}^{n_a}\boldsymbol{\delta a}(k,l)\,(\Delta_a\cdot{\Delta_a}^{\mathrm{T}})^{\dagger}(k,j)\,. &(52)\\
&= -\sum_{l=1}^{N}\boldsymbol{\delta d}(i,l)\,\sum_{k=1}^{n_a}\Delta_a(k,l)\,(\Delta_a\cdot{\Delta_a}^{\mathrm{T}})^{\dagger}(k,j)\,, &(53)
\end{aligned}$$

because, with the chosen notations, $\Delta_a(k,l)=\boldsymbol{\delta a}(k,l)$.

Finally, the pseudo-inverse definition allows us to write

$${\Delta_a}^{\dagger} = {\Delta_a}^{\mathrm{T}}\cdot(\Delta_a\cdot{\Delta_a}^{\mathrm{T}})^{\dagger}\,, \tag{54}$$

so that Eq. (53) can also be written

$$\mathbf{G}^*(i,j) = -\sum_{l=1}^{N} \boldsymbol{\delta d}(i,l)\, {\Delta_a}^\dagger(l,j) \tag{55}$$

$$= -\sum_{l=1}^{N} \Delta_d(i,l) {\Delta_a}^\dagger(l,j)\,, \tag{56}$$

which means that

$$\mathbf{G}^* = -\Delta_d \cdot {\Delta_a}^\dagger\,, \tag{57}$$

as exploited in Eq. (27). We have thus demonstrated the equivalence between the two formulae (26) and (27).

ACKNOWLEDGMENTS

C. Béchet particularly thanks Johann Kolb and Pierre-Yves Madec from the European Southern Observatory (Garching, Germany), for fruitful discussions about the presented method. This work has been partially funded by the OPTICON-JRA2 project of the European Commission FP7 programme, under Grant Agreement number 226604.

REFERENCES

[1] Esposito, S., Riccardi, A., Fini, L., Puglisi, A. T., Pinna, E., Xompero, M., Briguglio, R., Quirós-Pacheco, F., Stefanini, P., Guerra, J. C., Busoni, L., Tozzi, A., Pieralli, F., Agapito, G., Brusa-Zappellini, G., Demers, R., Brynnel, J., Arcidiacono, C., and Salinari, P., "First light AO (FLAO) system for LBT: final integration, acceptance test in Europe, and preliminary on-sky commissioning results," in [*Adaptive Optics Systems II*], *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* **7736** (July 2010).

[2] Arsenault, R., Madec, P., Hubin, N., Stroebele, S., Paufique, J., Vernet, E., Hackenberg, W., Pirard, J., Jochum, L., Glindemann, A., Jost, A., Conzelmann, R., Kiekebusch, M., Tordo, S., Lizon, J., Donaldson, R., Fedrigo, E., Soenke, C., Duchateau, M., Bruton, A., Delabre, B., Downing, M., Reyes, J., Kolb, J., Bechet, C., Lelouarn, M., Bonaccini Calia, D., Quattri, M., Guidolin, I., Buzzoni, B., Dupuy, C., Guzman, R., Comin, M., Silber, A., Quentin, J., La Penna, P., Manescau, A., Jolley, P., Heinz, V., Duhoux, P., Argomedo, J., Gallieni, D., Lazzarini, P., Biasi, R., Andrighettoni, M., Angerer, G., Pescoller, D., Stuik, R., and Deep, A., "Manufacturing of the ESO adaptive optics facility," in [*Adaptive Optics Systems II*], *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* **7736** (July 2010).

[3] Kolb, J., Le Louarn, M., Muller, N., and Béchet, C., "Calibration strategy of the adaptive optics facility," in [*Adaptive Optics Systems III*], *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* **7747** (July 2012).

[4] Loupias, M., Bacon, R., Caillier, P., Fleischmann, A., Jarno, A., Kelz, A., Kosmalski, J., Laurent, F., Le Floch, M., Lizon, J. L., Manescau, A., Nicklas, H., Parès, L., Pécontal, A., Reiss, R., Remillieux, A., Renault, E., Roth, M. M., Rupprecht, G., and Stuik, R., "MUSE instrument global performance analysis," in [*Modeling, Systems Engineering, and Project Management for Astronomy IV*], *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* **7738** (July 2010).

[5] Béchet, C., Kolb, J., Madec, P.-Y., Tallon, M., and Thiébaut, E., "Identification of system misregistrations during ao-corrected observations," in [*Adaptive Optics for Extremely Large Telescopes II*], *AO4ELT2 Conference Proceedings* (2011).

[6] Meimon, S., Fusco, T., and Petit, C., "An optimized calibration strategy for high order adaptive optics systems : the Slope-Oriented Hadamard Actuation," in [*Adaptative Optics for Extremely Large Telescopes*], (2010).

[7] Pieralli, F., Puglisi, A., Quiros Pacheco, F., and Esposito, S., "Sinusoidal calibration technique for Large Binocular Telescope system," in [*Adaptive Optics Systems*], *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* **7015** (July 2008).

[8] Roddier, F., [*The Effects of Atmospheric Turbulence in Optical Astronomy*], ch. V, Progress in Optics, North-Holland Publishing Company (1981).
[9] Vidal, F., Gendron, E., and Rousset, G., "Tomography approach for multi-object adaptive optics," *J. Opt. Soc. Am. A* **27**, A253–A264 (Nov 2010).