Optimal wave-front reconstruction strategies for multiconjugate adaptive optics

Thierry Fusco, Jean-Marc Conan, Gérard Rousset, Laurent Marc Mugnier, and Vincent Michau

Office National d'Études et de Recherches Aérospatiales, Département d'Optique Théorique et Appliquée, BP 72, F-92322 Châtillon cedex, France

We propose an optimal approach for the phase reconstruction in a large field of view (FOV) for multiconjugate adaptive optics. This optimal approach is based on a minimum-mean-square-error estimator that minimizes the mean residual phase variance in the FOV of interest. It accounts for the C_n^2 profile in order to optimally estimate the correction wave front to be applied to each deformable mirror (DM). This optimal approach also accounts for the fact that the number of DMs will always be smaller than the number of turbulent layers, since the C_n^2 profile is a continuous function of the altitude h. Links between this optimal approach and a tomographic reconstruction of the turbulence volume are established. In particular, it is shown that the optimal approach consists of a full tomographic reconstruction of the turbulence volume followed by a projection onto the DMs accounting for the considered FOV of interest. The case where the turbulent layers are assumed to match the mirror positions [model-approximation (MA) approach], which might be a crude approximation, is also considered for comparison. This MA approach will rely on the notion of equivalent turbulent layers. A comparison between the optimal and MA approaches is proposed. It is shown that the optimal approach provides very good performance even with a small number of DMs (typically, one or two). For instance, good Strehl ratios (greater than 20%) are obtained for a 4-m telescope on a 150-arc sec × 150-arc sec FOV by using only three guide stars and two DMs.

OCIS codes: 010.1080, 010.1330, 010.7350.

1. INTRODUCTION

Atmospheric turbulence severely limits the angular resolution of ground-based telescopes. Adaptive optics (AO)[1–3] is a powerful technique to overcome this limitation and to reach the diffraction limit of large telescopes. AO compensates, in real time, for the random fluctuations of wave fronts induced by the turbulent atmosphere. The turbulent wave front is measured by a wave-front sensor (WFS) using a guide star (GS) and optically corrected by a deformable mirror (DM) located in a pupil conjugate plane. This compensation allows the recording of long-exposure images with a resolution close to the diffraction limit. Because of anisoplanatism, the correction is efficient in only a limited field of view (FOV) (the so-called isoplanatic field) around the GS. This effect originates from the fact that turbulence is distributed in the volume above the telescope; then the wave fronts, coming from angularly separated points, are degraded differently. In the visible, the isoplanatic field is approximately a few arc seconds.[4] Beyond this FOV, the correction degrades.[5] Recently, a postprocessing method has been proposed to deal with the spatial variation of such an AO point-spread function.[6] This method gives very good results, but it is limited by the decrease of the correction degree in the FOV, which leads to a decrease of the signal-to-noise ratio (SNR) in the corrected image. Classical AO therefore gives poor high-resolution performance in the case of large FOV. Improved performance is, however, expected with multiconjugate AO (MCAO).[7,8] It consists in using several DMs conjugated at different heights in the atmosphere (see Fig. 1). With such a system, the turbulence effects are corrected not only on the telescope pupil but also in the turbulence volume; hence the increase of the correction field. Generally, several GSs are used to sense the perturbation in different FOV positions and to control these mirrors. The choice of the number of GSs[9–12] and DMs[10,11,13,14] is crucial for the design of such systems. It is related to the turbulence profile $C_n^2(h)$, the telescope diameter, and the observation goals. Note that, in this paper, we consider only natural GSs, but all the theoretical development could be extended to the case of laser GSs provided that all their specificities are taken into account (analysis geometry, cone effect, tip–tilt measurement problems).

We believe that one of the key issues is the phase reconstruction in MCAO. It is linked to the capability of the phase reconstruction algorithm to find the best deformation to apply on each DM from a set of WFS measurements, in order to obtain the best correction in a given FOV of interest. Since the minimization of the residual phase variance maximizes the image quality in the considered direction, we derive a minimum-mean-square-error (MMSE) estimator that minimizes the mean residual phase variance in the FOV of interest.[15] It accounts for the C_n^2 profile in order to optimally estimate the correction wave front to be applied to each DM. This optimal approach also accounts for the fact that the number of DMs will always be smaller than the number of turbulent layers, since the C_n^2 profile is a continuous function of the altitude h. Links between this MMSE approach and a tomographic reconstruction of the turbulence volume are established. In particular, it is shown that the MMSE approach consists of a full tomographic reconstruction of the turbulence volume followed by a projection onto the DMs accounting for the considered FOV of interest.

The case where the turbulent layers are assumed to match the mirror positions, which might be a crude ap-

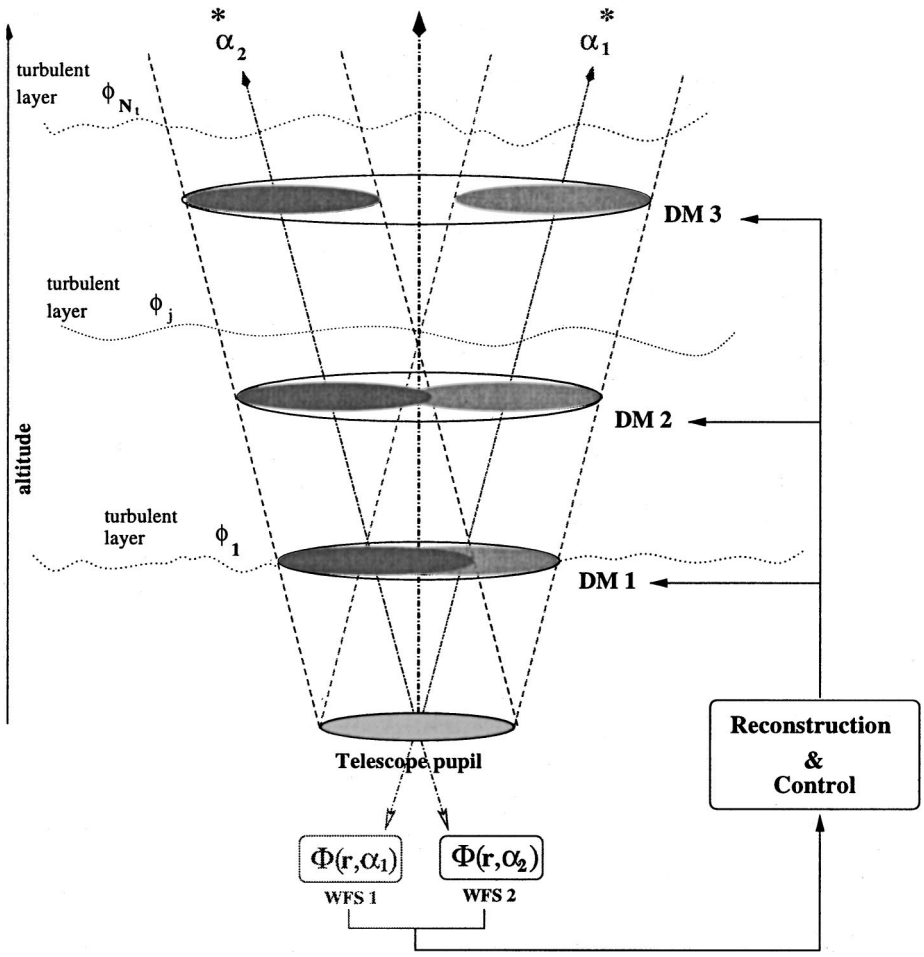

Fig. 1. Concept of a MCAO system. Several DMs are conjugated to different heights in the atmosphere. The wave-front analysis is made on several GSs located in the FOV.

proximation, is also considered for comparison. This model-approximation (MA) approach will rely on the notion of equivalent turbulent layers.[10] In both approaches, the regularization of the ill-posed problem for the phase reconstruction is studied carefully, and the prior knowledge available both on turbulence and on noise statistics has been incorporated into the reconstruction. Note that, in the paper, we consider only open-loop conditions; that is, all the wave-front statistics are derived from the Kolmogorov or the von Kármán theory. No temporal behavior is considered. This open-loop hypothesis is more restrictive than the study performed by Ellerbroek[15] but allows us to obtain simple analytical formulas and to propose physical interpretations of the results.

The theoretical development of the optimal approach for large-FOV phase reconstruction is proposed in Section 2. The crude model approximation where the turbulence is assumed to match the DM position, is considered in Section 3. A comparison of the performance of different phase reconstruction approaches [conventional truncated singular value decomposition (SVD), optimal approximation, and model approximation] is then proposed in Section 4. We study the influence of a well-chosen regularization (Kolmogorov statistics) on the phase reconstruction for a large FOV. The appeal of the optimal phase estimation approach, and then the need of accurate C_n^2 measurements during the observing runs, is demonstrated.

2. OPTIMAL APPROACH FOR LARGE-FIELD-OF-VIEW PHASE RECONSTRUCTION

The concept of MCAO has been studied in the last ten years by many authors.[7–20] The goal of MCAO is to compensate well for the turbulent wave fronts not only in one direction but also in a specified FOV of interest $\{\boldsymbol{\alpha}\}_{\mathrm{FOV}}$ (larger than the classical isoplanatic patch[4]).

Let us assume that the turbulent atmosphere is composed of a discrete sum of thin turbulent layers located at different heights.[21] In the near-field approximation,[21] the resulting phase $\Phi(\mathbf{r}, \boldsymbol{\alpha})$ in the telescope pupil is given, for a sky direction $\boldsymbol{\alpha}$, by

$$\Phi(\mathbf{r}, \boldsymbol{\alpha}) = \sum_{j=1}^{N_t} \phi_j(\mathbf{r} + h_j\boldsymbol{\alpha}), \tag{1}$$

where $\mathbf{r}$ is the pupil coordinate and $\phi_j(\boldsymbol{\rho}_j)$ are the phase perturbations in the jth atmospheric turbulent layer located at the altitude h_j. N_t is the number of turbulent layers.

The wave front is measured in the telescope pupil for the discrete set of GS directions $\{\boldsymbol{\alpha}_i\}_{\mathrm{GS}}$. The correction is computed by using all these measurements $[\{\Phi^m(\mathbf{r}, \boldsymbol{\alpha}_i)\}_{\mathrm{GS}}]$ and considering several DMs located at different heights. Therefore the key points for the design of a MCAO system are the number and the position of DMs and GSs and, of course, the phase reconstruction method that gives the correction phase for the different DMs. In the present paper, we focus on this phase estimation algorithm, since the MCAO performance with respect to the DM and GS number and positions has already been studied.[10–13] The main result of these previous works is that for telescope diameters of 4–8 m and for *K*-band (2.2-μm) imaging, only a small number of DMs and GSs is needed to obtain quasi-uniform correction in a large FOV (typically larger than 1 arc min).

A. Position of the Problem

The goal of our approach is to minimize the residual phase variance in a specified FOV of interest, that is, to derive a MMSE estimator.[15] This phase estimator is defined as the one that minimizes a quadratic distance averaged on the FOV of interest $\{\boldsymbol{\alpha}\}_{\mathrm{FOV}}$ between the resulting true and correction phases:

$$\boldsymbol{\epsilon} = \left\langle \int_{\{\boldsymbol{\alpha}\}_{\mathrm{FOV}}} \|\hat{\Phi}(\mathbf{r}, \boldsymbol{\alpha}) - \Phi(\mathbf{r}, \boldsymbol{\alpha})\|^2 \, \mathrm{d}\boldsymbol{\alpha} \right\rangle_{\Phi,\mathrm{noise}}, \qquad (2)$$

where $\langle \cdot \rangle_{\Phi,\mathrm{noise}}$ stands for a mathematical expectation on both turbulence and WFS noise outcomes, $\| \cdot \|^2$ denotes the spatial variance in the telescope pupil, $\Phi(\mathbf{r}, \boldsymbol{\alpha})$ is the true phase in a given direction $\boldsymbol{\alpha}$ in $\{\boldsymbol{\alpha}\}_{\mathrm{FOV}}$, and $\hat{\Phi}(\mathbf{r}, \boldsymbol{\alpha})$ is the estimated correction phase in that direction $\boldsymbol{\alpha}$. The problem is to estimate $\hat{\Phi}(\mathbf{r}, \boldsymbol{\alpha})$ under the constraint that it will be generated by a finite number of DMs, using not only the WFS measurements but also *a priori* information that we have on the turbulent wave front in the atmospheric volume.

Let us consider that we have N_{GS} GSs, i.e., N_{GS} WFS measurements. For each WFS, we assume that the measured phase can be expressed as

$$\Phi^m(\mathbf{r}, \boldsymbol{\alpha}_i) = \Phi(\mathbf{r}, \boldsymbol{\alpha}_i) + n_i(\mathbf{r}), \qquad (3)$$

where $\boldsymbol{\alpha}_i$ is the angular position of the ith GS. For the sake of simplicity, Eq. (3) assumes that the WFS directly gives phase map measurements and that n_i follows Gaussian statistics (central limit theorem). We suppose here that $\Phi^m(\mathbf{r}, \boldsymbol{\alpha}_i)$ is measured on a basis with an infinite number of modes. The measurements are limited only by the noise. Indeed, this noise on slope measurements given by a Shack–Hartmann (SH) WFS is given by the sum of several pixels in the calculation of the center of gravity and through the reconstruction; the noise on the phase results from a large number of subaperture contributions. To account for the noise propagation through the reconstruction from SH data, we color this Gaussian noise $n_i(\mathbf{r})$ with, in the Fourier domain, a power spectral density following an f^{-2} law.[22] The turbulent phase on the telescope pupil is given by the sum of all the turbulent layer contributions [see Eq. (1)]; then Eq. (3) can be rewritten as

$$\Phi^m(\mathbf{r}, \boldsymbol{\alpha}_i) = \sum_{j=1}^{N_t} \phi_j(\mathbf{r} + h_j \boldsymbol{\alpha}_i) + n_i(\mathbf{r}). \qquad (4)$$

The unknowns of the problem are the correction phases $\hat{\phi}_k$ to be estimated for each DM so as to minimize the criterion defined in Eq. (2). Of course, for practical reasons, the DM number (N_{DM}) will always be smaller than the number of turbulent layers (N_t). In that case, we have, for a given direction $\boldsymbol{\alpha}$,

$$\hat{\Phi}(\mathbf{r}, \boldsymbol{\alpha}) = \sum_{k=1}^{N_{\mathrm{DM}}} \hat{\phi}_k(\mathbf{r} + h_k \boldsymbol{\alpha}). \qquad (5)$$

The DM positions h_k are, for instance, computed as presented in Refs. 10, 11, and 13 by using an average C_n^2 profile. Then Eq. (2) becomes

$$\boldsymbol{\epsilon} = \left\langle \int_{\{\boldsymbol{\alpha}\}_{\mathrm{FOV}}} \left\| \sum_{k=1}^{N_{\mathrm{DM}}} \hat{\phi}_k(\mathbf{r} + h_k \boldsymbol{\alpha}) - \sum_{j=1}^{N_t} \phi_j(\mathbf{r} + h_j \boldsymbol{\alpha}) \right\|^2 \mathrm{d}\boldsymbol{\alpha} \right\rangle_{\phi,\mathrm{noise}}. \qquad (6)$$

For the sake of clarity, let us rewrite all the equations defined above in a matrix form. Equations (1), (4), and (5) become, respectively,

$$\Phi(\mathbf{r}, \boldsymbol{\alpha}) = \mathcal{M}_{\boldsymbol{\alpha}}^{N_t} \boldsymbol{\phi}, \qquad (7)$$

$$\Phi^m(\mathbf{r}, \boldsymbol{\alpha}_i) = \mathcal{M}_{\boldsymbol{\alpha}_i}^{N_t} \boldsymbol{\phi} + n_i, \qquad (8)$$

$$\hat{\Phi}(\mathbf{r}, \boldsymbol{\alpha}) = \mathcal{M}_{\boldsymbol{\alpha}}^{N_{\mathrm{DM}}} \hat{\boldsymbol{\phi}}, \qquad (9)$$

where $\mathcal{M}_{\boldsymbol{\alpha}}^{N_t}$ and $\mathcal{M}_{\boldsymbol{\alpha}}^{N_{\mathrm{DM}}}$ are the matrices that perform the sum of the contributions of each wave front $\phi_j(\boldsymbol{\rho}_j)$ and $\hat{\phi}_k(\boldsymbol{\rho}_k)$ on the telescope pupil for a given direction $\boldsymbol{\alpha}$. $\boldsymbol{\phi}$ and $\hat{\boldsymbol{\phi}}$ are defined as

$$\boldsymbol{\phi} = \begin{pmatrix} \boldsymbol{\phi}_1 \\ \vdots \\ \boldsymbol{\phi}_j \\ \vdots \\ \boldsymbol{\phi}_{N_t} \end{pmatrix}, \qquad \hat{\boldsymbol{\phi}} = \begin{pmatrix} \hat{\boldsymbol{\phi}}_1 \\ \vdots \\ \hat{\boldsymbol{\phi}}_k \\ \vdots \\ \hat{\boldsymbol{\phi}}_{N_{\mathrm{DM}}} \end{pmatrix}. \qquad (10)$$

The criterion to be minimized is then

$$\boldsymbol{\epsilon} = \left\langle \int_{\{\boldsymbol{\alpha}\}_{\mathrm{FOV}}} \|\mathcal{M}_{\boldsymbol{\alpha}}^{N_{\mathrm{DM}}} \hat{\boldsymbol{\phi}} - \mathcal{M}_{\boldsymbol{\alpha}}^{N_t} \boldsymbol{\phi}\|^2 \, \mathrm{d}\boldsymbol{\alpha} \right\rangle_{\phi,\mathrm{noise}}. \qquad (11)$$

B. Optimal Solution

In general, the calculation of the MMSE estimator is not tractable unless the estimator is assumed to be linear with respect to the data (linear MMSE estimator). It is important to note that in the case of joint Gaussian statistics for the noise and the turbulence (which is the case in our problem), this linear estimator is identical to the true MMSE estimator.[23]

We can therefore seek the MMSE solution in the form

$$\hat{\boldsymbol{\phi}} = \mathcal{W} \boldsymbol{\Phi}^m \quad \text{with} \quad \boldsymbol{\Phi}^m = \mathcal{M}_{N_{\mathrm{GS}}}^{N_t} \boldsymbol{\phi} + \mathbf{n}, \qquad (12)$$

where the new unknowns are the elements of the matrix $\mathcal{W}$. $\mathcal{M}_{N_{GS}}^{N_t}$, $\mathbf{\Phi}^m$, and $\mathbf{n}$ are matrices and vectors defined as

$$\mathcal{M}_{N_{GS}}^{N_t} = (\mathcal{M}_{\boldsymbol{\alpha}_1}^{N_t}, \ldots, \mathcal{M}_{\boldsymbol{\alpha}_i}^{N_t}, \ldots, \mathcal{M}_{\boldsymbol{\alpha}_{N_{GS}}}^{N_t}), \tag{13}$$

$$\mathbf{\Phi}^m = \begin{pmatrix} \Phi^m(\mathbf{r}, \boldsymbol{\alpha}_1) \\ \vdots \\ \Phi^m(\mathbf{r}, \boldsymbol{\alpha}_i) \\ \vdots \\ \Phi^m(\mathbf{r}, \boldsymbol{\alpha}_{N_{GS}}) \end{pmatrix}, \qquad \mathbf{n} = \begin{pmatrix} n_1(\mathbf{r}) \\ \vdots \\ n_i(\mathbf{r}) \\ \vdots \\ n_{N_{GS}}(\mathbf{r}) \end{pmatrix}. \tag{14}$$

$\mathcal{W}$ is the reconstruction matrix

$$\mathcal{W} = \begin{pmatrix} \mathcal{W}_1 \\ \vdots \\ \mathcal{W}_j \\ \vdots \\ \mathcal{W}_{N_{DM}} \end{pmatrix},$$

so that

$$\hat{\boldsymbol{\phi}}_j = \mathcal{W}_j \mathbf{\Phi}^m. \tag{15}$$

Putting Eq. (12) in Eq. (11) yields

$$\boldsymbol{\epsilon} = \int_{\{\boldsymbol{\alpha}\}_{FOV}} \langle \| \mathcal{M}_{\boldsymbol{\alpha}}^{N_{DM}} (\mathcal{W} \mathcal{M}_{N_{GS}}^{N_t} \boldsymbol{\phi} + \mathcal{W}\mathbf{n}) - \mathcal{M}_{\boldsymbol{\alpha}}^{N_t} \boldsymbol{\phi} \|^2 \rangle_{\phi,\text{noise}} \, \mathrm{d}\alpha. \tag{16}$$

This equation must be minimized with respect to $\mathcal{W}$. The explicit minimization of Eq. (16) is presented in Appendix A. The final result is

$$\mathcal{W} = \left[\int_{\{\boldsymbol{\alpha}\}_{FOV}} (\mathcal{M}_{\boldsymbol{\alpha}}^{N_{DM}})^{\mathrm{T}} \mathcal{M}_{\boldsymbol{\alpha}}^{N_{DM}} \, \mathrm{d}\boldsymbol{\alpha} \right]^{+} \times \left[\int_{\{\boldsymbol{\alpha}\}_{FOV}} (\mathcal{M}_{\boldsymbol{\alpha}}^{N_{DM}})^{\mathrm{T}} \mathcal{M}_{\boldsymbol{\alpha}}^{N_t} \, \mathrm{d}\boldsymbol{\alpha} \right] \times \mathcal{C}_{\boldsymbol{\phi}} (\mathcal{M}_{N_{GS}}^{N_t})^{\mathrm{T}} [\mathcal{M}_{N_{GS}}^{N_t} \mathcal{C}_{\boldsymbol{\phi}} (\mathcal{M}_{N_{GS}}^{N_t})^{\mathrm{T}} + \mathcal{C}_{\mathbf{n}}]^{-1}, \tag{17}$$

where $\mathcal{C}_{\boldsymbol{\phi}}$ and $\mathcal{C}_{\mathbf{n}}$, defined in Appendix A, are the generalization for several layers and several GSs of the classical turbulence and noise covariance matrices. The T and + superscript symbols denote the transpose and the generalized inverse matrix, respectively. Reintroducing Eq. (17) into Eq. (16), we can easily obtain an analytical expression of the MCAO error as a function of FOV angle.[24] The true model derived above assumes an infinite number of modes to describe the turbulent phase and its correction. For practical reasons and for purposes of limiting the numerical calculations, the number of modes is, however, limited. This undermodeling in the direct problem induces correction errors. It is important to account for these errors, which is easily done by using a Monte Carlo simulation (that is, we simulate turbulent wave fronts, and we apply a correction derived by using $\mathcal{W}$). Furthermore, such a simulation allows us to account for slight discrepancies from the true turbulence model. For instance, the wave fronts are simulated with von-Kármán statistics (finite outer scale), while the reconstruction matrix uses a Kolmogorov regularization.

When $\mathcal{M}_{\boldsymbol{\alpha}}^{N_{DM}} = \mathcal{M}_{\boldsymbol{\alpha}}^{N_t}$, that is, when the DMs are exactly located on the turbulent layers, Eq. (17) simply reads as

$$\mathcal{W}_{N_{DM}=N_t} = \mathcal{C}_{\boldsymbol{\phi}} (\mathcal{M}_{N_{GS}}^{N_t})^{\mathrm{T}} [\mathcal{M}_{N_{GS}}^{N_t} \mathcal{C}_{\boldsymbol{\phi}} (\mathcal{M}_{N_{GS}}^{N_t})^{\mathrm{T}} + \mathcal{C}_{\mathbf{n}}]^{-1}. \tag{18}$$

Equation (11) then becomes

$$\boldsymbol{\epsilon} = \left\langle \int_{\{\boldsymbol{\alpha}\}_{FOV}} \| \mathcal{M}_{\boldsymbol{\alpha}}^{N_t} (\hat{\boldsymbol{\phi}} - \boldsymbol{\phi}) \|^2 \, \mathrm{d}\boldsymbol{\alpha} \right\rangle_{\phi,\text{noise}}. \tag{19}$$

In that case, it can be shown that the minimization of the residual phase variance in the telescope pupil $\boldsymbol{\epsilon}$ is equivalent to the minimization of

$$\boldsymbol{\epsilon}' = \langle \| \hat{\boldsymbol{\phi}} - \boldsymbol{\phi} \|^2 \rangle_{\phi,\text{noise}}, \tag{20}$$

that is, to the minimization of the residual phase variance in each layer (whatever the FOV of interest). Our estimator is therefore equivalent, in that case, to that of a tomographic approach.[19] In particular, there is no dependence on the field angle. Such a DM correction minimizes the phase residual variance whatever the FOV position. But this case is only idealistic. In fact, the number of DMs will always be smaller than the number of turbulent layers. It is, however, interesting to note that in the general case, the solution given in Eq. (17) actually consists of this tomographic reconstruction on all turbulent layers corresponding to Eq. (18) followed by a projection onto the solution space (corresponding to the small number of altitudes where the DMs are located). The projection operator is therefore

$$\mathcal{P}_{N_{DM},N_t} = \left[\int_{\{\boldsymbol{\alpha}\}_{FOV}} (\mathcal{M}_{\boldsymbol{\alpha}}^{N_{DM}})^{\mathrm{T}} \mathcal{M}_{\boldsymbol{\alpha}}^{N_{DM}} \, \mathrm{d}\boldsymbol{\alpha} \right]^{+} \times \left[\int_{\{\boldsymbol{\alpha}\}_{FOV}} (\mathcal{M}_{\boldsymbol{\alpha}}^{N_{DM}})^{\mathrm{T}} \mathcal{M}_{\boldsymbol{\alpha}}^{N_t} \, \mathrm{d}\boldsymbol{\alpha} \right]. \tag{21}$$

This projection matrix is directly linked, through the integral in $\boldsymbol{\alpha}$, to the FOV of interest $\{\boldsymbol{\alpha}\}_{FOV}$, where the correction is optimized. Indeed, when the mirror positions do not match the turbulent layers ($N_{DM} < N_t$), an overall correction in the FOV is not possible. Optimizing for a particular FOV position may degrade the correction in other positions. Trade-offs have to be made for a specified set of FOV positions. The projection $\mathcal{P}_{N_{DM},N_t}$ performs optimally these trade-offs.

One can also show from Eq. (16) that, whatever the position in the FOV, the residual phase variance is minimal for one DM per layer ($N_{DM} = N_t$). However, we will see in Section 4 that this ultimate performance is almost reached with a small number of DMs when considering a reasonable FOV of interest.

Note that there is an analogy between the MCAO correction of the turbulence volume with a finite number of DMs, as presented here, and the correction of the turbulent phase in classical AO with a finite number of actuators, as proposed by Wallner.[25] In both cases, the correction space is smaller than the unknown space (number of DMs smaller than the number of layers or the number of actuators smaller than the number of turbulent modes). This leads to a similar form of the reconstruction: first, a "full" reconstruction, followed by a projection onto the finite-space solution.

With the matrix $\mathcal{W}$, one can compute $\hat{\boldsymbol{\phi}} = \mathcal{W}\,\mathbf{\Phi}^m$, which gives the estimated correction phase on each DM that en-

sures a minimal residual phase variance for all the directions of the specified FOV $\{\boldsymbol{\alpha}\}_{\mathrm{FOV}}$. Of course, the computation of Eq. (17) requires the knowledge of the turbulence profile for the computation of $\mathcal{M}_{\boldsymbol{\alpha}}^{N_t}$ and $\mathcal{C}_{\phi}$. A real-time measurement of the C_n^2 profile can, for instance, be obtained with a generalized SCIDAR.[26] In Section 3, a second approach, based on a cruder turbulence model, is presented.

3. MODEL-APPROXIMATION APPROACH

In this approach, we assume that all the turbulence is located on the DMs. The C_n^2 profile is modeled only by a small number (N_{EL}) of turbulent layers, called equivalent layers (ELs), in which are located the $N_{\mathrm{DM}} = N_{\mathrm{EL}}$ DMs. The computation of the EL position and strength is done by a sampling of the C_n^2 profile into N_{EL} slabs.[10,11]

Using this simplified turbulence model, one can estimate the correction phase with the approach proposed in Subsection 2B. All the equations remain valid, with N_t and N_{DM} replaced by N_{EL}. Therefore the direct problem can be rewritten as

$$\Phi_{\boldsymbol{\alpha}}(\mathbf{r}) \simeq \sum_{j=1}^{N_{\mathrm{EL}}} \phi_j(\mathbf{r} + h_j\boldsymbol{\alpha}),$$

$$\Phi_{\boldsymbol{\alpha}_i}^{m}(\mathbf{r}) \simeq \sum_{j=1}^{N_{\mathrm{EL}}} \phi_j(\mathbf{r} + h_j\boldsymbol{\alpha}_i) + n_i(\mathbf{r}). \tag{22}$$

Consequently, the reconstruction matrix is deduced from Eq. (18):

$$\mathcal{W}_{\mathrm{MA}} = \mathcal{C}_{\boldsymbol{\phi}}(\mathcal{M}_{N_{\mathrm{GS}}}^{N_{\mathrm{EL}}})^{\mathrm{T}}[\mathcal{M}_{N_{\mathrm{GS}}}^{N_{\mathrm{EL}}}\mathcal{C}_{\boldsymbol{\phi}}(\mathcal{M}_{N_{\mathrm{GS}}}^{N_{\mathrm{EL}}})^{\mathrm{T}} + \mathcal{C}_{\mathbf{N}}]^{-1}. \tag{23}$$

The reconstruction phases are therefore given by $\hat{\boldsymbol{\phi}} = \mathcal{W}_{\mathrm{MA}}\boldsymbol{\Phi}^m$.

This MA solution has already been derived in a previous paper[10,11] following a maximum *a posteriori* approach. But MMSE and maximum *a posteriori* estimators are, in any event equivalent[23] here on account of the Gaussian statistics of the noise and the turbulence.

Now let us compare the two approaches and discuss their similarities and differences. They are both derived from the same theoretical development based on the MMSE criterion. The only (but important) difference is that the model approximation uses a simplified direct problem, which leads to a suboptimal solution.

4. RESULTS AND PERFORMANCE

A. Simulation Tool

Let us consider a modal decomposition of the wave fronts onto the Zernike basis. The phase screen on each turbulent layer j becomes

$$\phi_j(\boldsymbol{\rho}_j) = \sum_{l=2}^{\infty} a_{l,j}Z_{l,j}(\boldsymbol{\rho}_j), \tag{24}$$

where $Z_{l,j}(\boldsymbol{\rho})$ is the lth Zernike polynomial defined on a metapupil of diameter D_j depending on the telescope diameter D, the layer altitude h_j, and the maximal FOV angle $\alpha_{\max}$ considered:

$$D_j = D + 2h_j\alpha_{\max}. \tag{25}$$

Of course, all the equations presented above are still valid in this basis. In particular, one can note that in Eqs. (10) the the $\boldsymbol{\phi}_j$ and the $\hat{\boldsymbol{\phi}}_k$ are simply vectors of Zernike coefficients $a_{l,j}$ and $a_{l,k}$; $\mathcal{C}_{\boldsymbol{\phi}}$ is therefore a generalization of the Zernike covariance matrix given by Noll.[27]

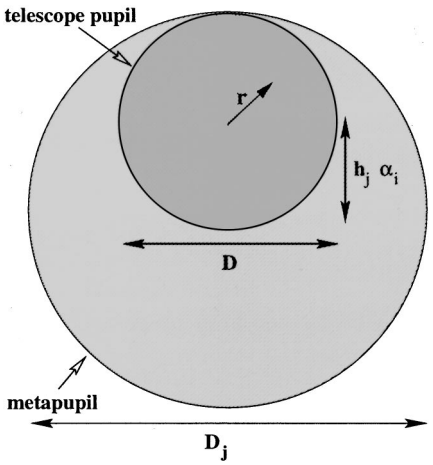

Fig. 2. Decentered part of the metapupil associated with the altitude h_j. The variable vector $\mathbf{r}$ is defined on the telescope pupil. The zone of interest is centered on $h_j\alpha_i$.

The measured phase is also decomposed onto the Zernike polynomial basis. Noise is added on each Zernike coefficient by using a noise covariance matrix $\mathcal{C}_{\mathbf{n}}$. The SH WFS is not really simulated but its characteristic noise propagation is accounted for through the use of $\mathcal{C}_{\mathbf{n}}$, which considers that we measure Zernike polynomial derivatives.[28]

For a direction α_i, only a part of the metapupil associated with the layer j is viewed: a disk of diameter D centered on $h_j\boldsymbol{\alpha}_i$. In this particular basis, $\mathcal{M}_{N_{\mathrm{GS}}}^{N_{\mathrm{EL}}}$, $\mathcal{M}_{N_{\mathrm{GS}}}^{N_t}$, $\mathcal{M}_{\boldsymbol{\alpha}}^{N_t}$, and $\mathcal{M}_{\boldsymbol{\alpha}}^{N_{\mathrm{EL}}}$ are computed as presented in Ref. 11. It consists of the decomposition of each decentered Zernike polynomial $[Z_{l,j}(\mathbf{r} + \boldsymbol{\alpha}h_j)]$ onto a Zernike basis defined on the telescope pupil (see Fig. 2). Ragazzoni *et al.*[20] have shown that the number of modes required for such a decomposition is given by the number l of the metapupil Zernike polynomial. This result is important, since it ensures that the dimension of each matrix $\mathcal{M}$ will be linked only to the number of Zernike polynomials considered in each (turbulent or DM) layer. For example, in the case of a two-DM system, where 66 Zernike modes are sought on the first DM and 135 are sought on the second DM, the dimension of the matrix $\mathcal{M}_{\boldsymbol{\alpha}}^{N_{\mathrm{DM}}}$ will be only $135 \times (66 + 135)$.

Ideally, the number of Zernike modes must be infinite, but for practical reasons (computation time, matrix sizes), only a finite number of Zernike modes is considered both for the measured phases and for the DMs.

B. Simulation Parameters

Let us consider a four-layer profile defined as follows:

No.	Position (km)	Strength (%)
1	0	25
2	2.5	25
3	5	25
4	7.5	25

The phase screens on each turbulent layer are simulated by McGlamery's method.[29] The simulated phase screens are large enough with respect to the telescope pupil to ensure that $L_0/D \simeq 4$, where L_0 is the outer scale of the turbulence. The total r_0 is equal to 0.1 m at 0.5 μm. This leads to an isoplanatic angle θ_0 (defined with the Fried formula[4]) equal to 1.44 arc sec at 0.5 μm and 8.52 arc sec at 2.2 μm. We consider a 4-m telescope, and sev-

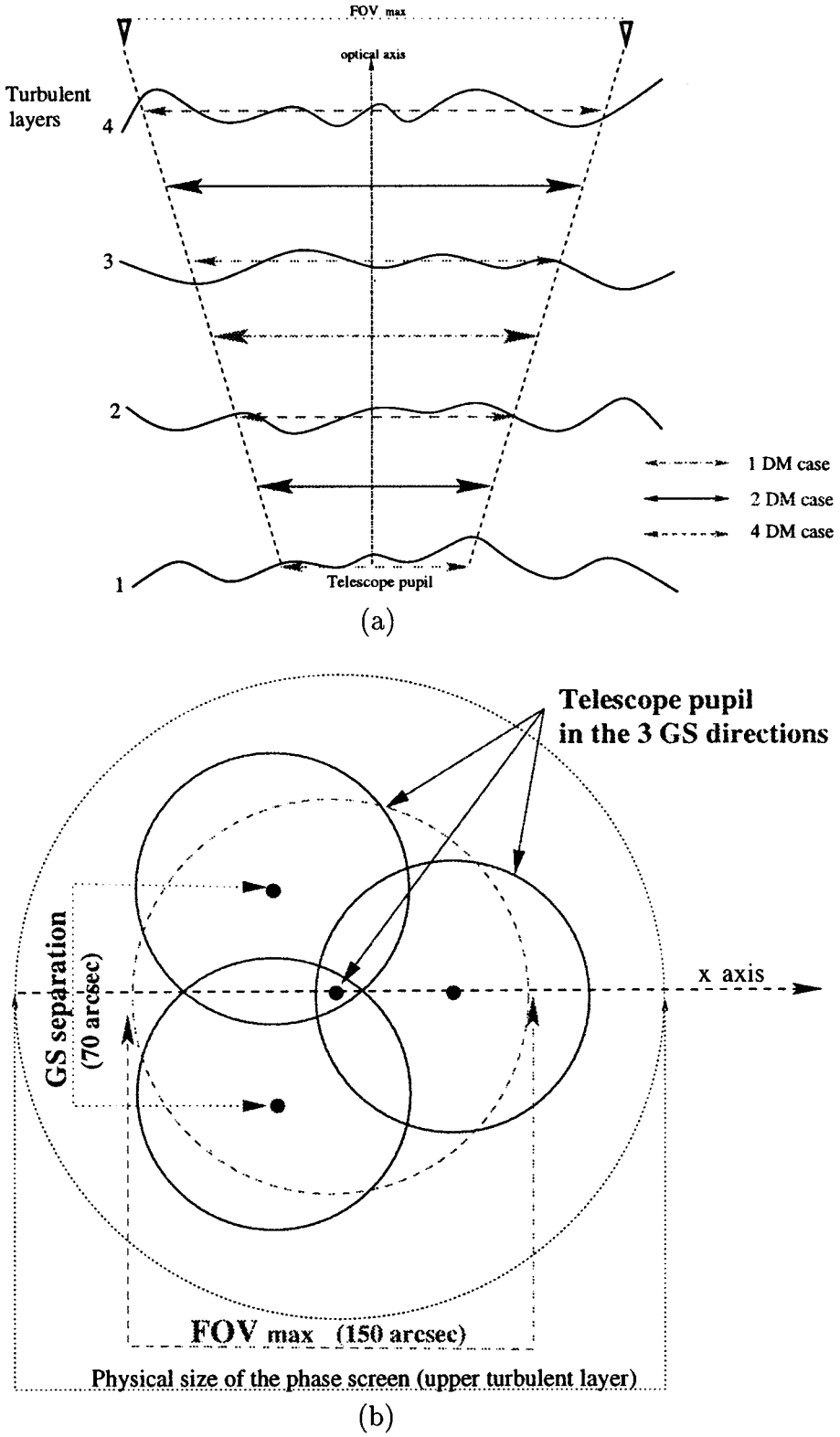

Fig. 3. (a) Turbulence (four layers) and DM (one, two, and four) repartition for the four systems presented in Subsection 4B. (b) Geometrical repartition of the GS pupil projection and FOV on the highest layer ($h = 7.5$ km). The physical size of the DM is equal to the physical size of the layers.

eral MCAO systems are presented in Fig. 3.

The several MCAO systems are constructed as follows:

a. One GS on the optical axis and one DM (with 135 corrected modes) conjugated at 3.75 or 6.5 km (Subsection 4.D.1).

b. One GS on the optical axis and four DMs (with, respectively, 66, 120, 135, and 230 corrected modes) conjugated on the four turbulent layers (tomographic reconstruction) (Subsection 4.D.1).

c. Three GSs located at the vertices of an equilateral triangle with a separation equal to 70 arc sec and two DMs (with, respectively, 66 and 135 corrected modes) conjugated at 1.25 and 6.25 km (Subsections 4.C and 4.D.2).

d. Three GSs located at the vertices of an equilateral triangle with a separation equal to 70 arc sec and four DMs (with, respectively, 66, 120, 135, and 230 corrected modes) conjugated on the four turbulent layers (tomographic reconstruction) (Subsection 4.D.2).

Ideally, a large number of the Zernike modes should be used, but for practical reasons the number of modes per DM considered here is quite reasonable.

First, it is important to note that in cases b and d the number of DMs is equal to the number of true layers; therefore these two cases can be seen as the ultimate performance of cases a and c, as mentioned in Subsection 2B.

The first two cases (a and b) correspond to the first step of a MCAO system, since they are composed of one or several conjugated DMs, but they still only use one GS. Therefore all the information on the off-axis phases is given only by the prior information that we have on the turbulence volume (the C_n^2 repartition and the Kolmogorov statistics of the phase). An example of such a system (case a) is under construction for the 8-m Gemini-North telescope.[30] This AO system (Altair) can be seen as the first order of a MCAO system. In our case, we have considered a 4-m telescope, but all the results can easily be extended to the 8-m case by a simple scaling of the FOV $\{\boldsymbol{\alpha}\}_{\mathrm{FOV}}$ by the diameter ratio and the number of corrected modes by the square of this ratio.

Cases c and d represent more complex systems, since they are composed of both several GS directions and several conjugated DMs.

The $\mathcal{C}_{\mathbf{n}}$ matrix is obtained by considering SH WFSs that measure the wave front in each GS direction. The SNR on each SH (defined as the ratio between the turbulence variance and the noise variance) is computed for a 7×7 subaperture SH and is equal to 10. It roughly corresponds to an 11th-magnitude GS. For each system, the maximal considered FOV (which defines the physical size of each DM; see Fig. 3) is equal to 150 arc sec.

The performance of the different methods is evaluated in terms of a Strehl ratio (SR) approximated by $\exp[-\sigma_{\mathrm{res}}^2(\boldsymbol{\alpha})]$, which is valid for good corrections. $\sigma_{\mathrm{res}}^2(\boldsymbol{\alpha})$ is computed by

$$\sigma_{\mathrm{res}}^2(\boldsymbol{\alpha}) = \langle \|\Phi(\mathbf{r}, \boldsymbol{\alpha}) - \hat{\Phi}(\mathbf{r}, \boldsymbol{\alpha})\|^2 \rangle, \tag{26}$$

where $\langle \cdot \rangle$ is an average on 100 decorrelated simulated phases.

First, a study of the gain brought by the regularization term (Kolmogorov regularization) in Eq. (17) and (18), compared with a more classical approach based on a least-squares minimization, is performed in Subsection 4.C.

C. Influence of the Kolmogorov Regularization

Let us consider the three-GS and two-DM system (case c) presented in Subsection 4.B. The classical approach to inverting the ill-posed problem of the phase correction estimation in each DM is to use a least-squares minimization,[18] that is, to consider a truncated singular value decomposition (SVD). With our notation, this wave-front estimator is therefore given by the following well-known relation:

$$\hat{\boldsymbol{\phi}} = [(\mathcal{M}_{N_{\mathrm{GS}}}^{N_{\mathrm{EL}}})^{\mathrm{T}} \mathcal{M}_{N_{\mathrm{GS}}}^{N_{\mathrm{EL}}}]^{+} (\mathcal{M}_{N_{\mathrm{GS}}}^{N_{\mathrm{EL}}})^{\mathrm{T}} \boldsymbol{\Phi}^m, \tag{27}$$

where $\mathcal{M}_{N_{\mathrm{GS}}}^{N_{\mathrm{EL}}}$ is the interaction matrix between the DMs and the WFSs. Because $(\mathcal{M}_{N_{\mathrm{GS}}}^{N_{\mathrm{EL}}})^{\mathrm{T}} \mathcal{M}_{N_{\mathrm{GS}}}^{N_{\mathrm{EL}}}$ is an ill-conditioned matrix (see Fig. 4), the inversion is made by using a SVD in which the lower-value modes are set to 0 in order to avoid the noise amplification. Of course, this truncation can be seen as a crude regularization, and it is easy to show that this approach is less optimal than the use of a well-chosen regularization term (the Kolmogorov statistics in our case), as shown for various turbulence-related applications in Refs. 10, 23, 31, and 32.

One can see in Fig. 5 that the use of a Kolmogorov regularization [model approximation, Eq. (23)], where $\mathcal{C}_{\boldsymbol{\phi}}$ is computed by assuming that all the turbulence is equally distributed on the two DMs, gives better results than the classical truncated SVD [Eq. (27)] whatever the chosen truncation threshold (note that the optimal choice of this threshold is one of the major problems of the SVD approach).

For the optimal SVD threshold ($\lambda_{\max}/50$ here), only 72 modes are corrected from the available 199 modes of the system (the piston is not considered). This optimal threshold is chosen as the one that gives the minimal mean residual variance in the whole FOV of interest (150 arc sec$\times$150 arc sec).

The great advantage of the well-chosen regularization (as derived from a MMSE approach) is that one does not have to adjust any parameters, since the optimal regularization is directly derived from the noise and turbulence (on each EL) statistics.

It seems clear, in this example, that an adequate regularization of the inverse problem is required in a large FOV to obtain good performance. This Kolmogorov regularization avoids the noise amplification but also allows a good phase extrapolation where the phase is not measured or only partially measured.

Now let us compare the optimal approach defined in Eq. (17) and the MA approach defined in Eq. (23). In Subsection 4.D.1, the one-GS and one-DM system (case a) is studied. In Subsection 4.D.2, the three-GS and two-DM system (case c) is considered. Note that for each GS configuration the best possible performance is obtained, as explained in Section 2, with one DM per turbulent layer, that is, four DMs here (cases b and d).

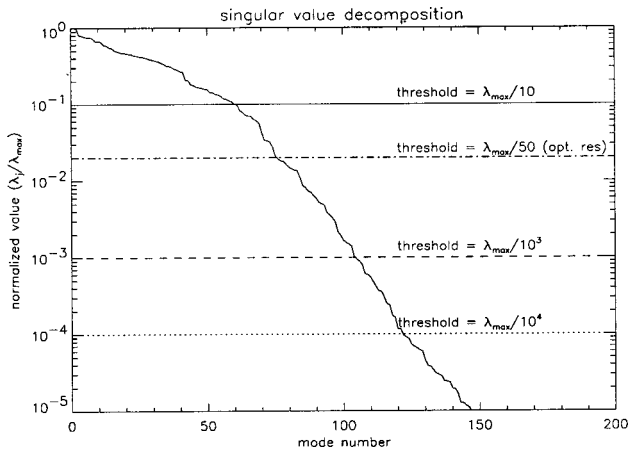

Fig. 4. Singular value of the systems versus mode number. The different considered thresholds are plotted. The optimal threshold (optimal result) is chosen as the one that gives the minimal residual variance in the whole FOV of interest (150 arc sec × 150 arc sec), as shown in Fig. 5.

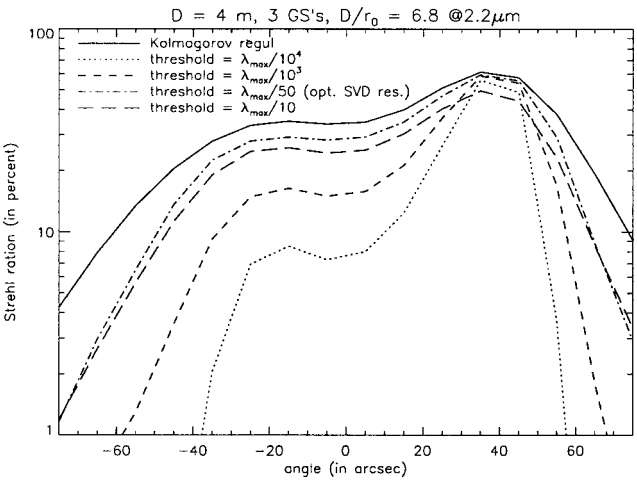

Fig. 5. Comparison between the Kolmogorov regularization (solid curve) and a SVD using different thresholds. The SR versus the FOV position is plotted for the x axis defined in Fig. 3(b).

D. Comparison of the Optimal and Model-Approximation Approaches

We have just shown that a well-chosen Kolmogorov regularization always gives better results than a simple truncated SVD approach. Let us now compare the two phase estimation methods presented in Sections 2 and 3 by using this regularization. More precisely, let us show the gain brought by the optimal phase estimation approach in which the prior information on the turbulence profile is more precise and for which an optimization in a given FOV of interest $\{\boldsymbol{\alpha}\}_{\mathrm{FOV}}$ is performed.

1. Mono-Guide-Star and Mono-Deformable-Mirror System

Let us first consider the simple but illustrative case of a system composed of only one GS and one conjugated DM (case a). In such a configuration, two different cases can be studied.

- Case of a well-placed DM with respect to the turbulence profile (Fig. 6).

The DM is located at 3.75 km (center of gravity of the C_n^2 profile). If we use the MA method (all the turbulence supposed to be on the DM), then because the DM position is well chosen, a good extrapolation is possible. Indeed, as shown in Fig. 6, a quasi-optimal (close to the ultimate four-DM performance) SR is obtained in a 20-arc sec × 20-arc sec FOV. Beyond, if the science object is far from the optical axis (typically, 50 arc sec), the degradation of the extrapolation between one and four DMs becomes important (SR = 3% for one DM and 8% for four DMs).

Now we take into account the knowledge of the true C_n^2 profile for the optimization of the DM correction in the science object direction. In our example, we consider a

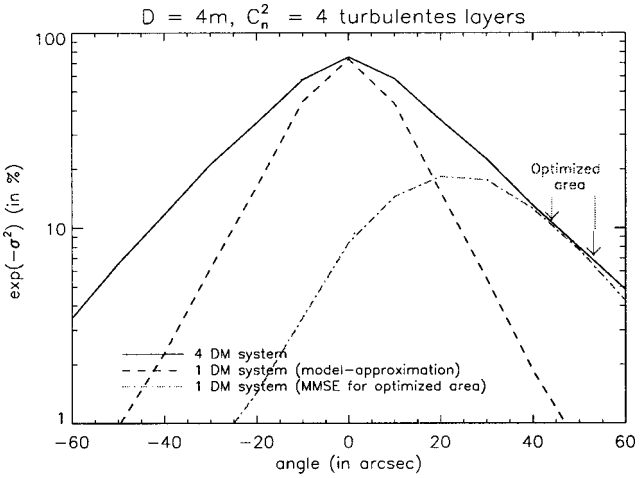

Fig. 6. Comparison of the optimal phase estimation and MA approaches in the case of a one-GS (on the optical axis) and one-DM (conjugated at 3.75 km) system. In each case, an X cut of the FOV is presented. These simulation are made for a four-layer C_n^2 profile and a 4-m telescope. We plot the tomographic reconstruction (four DMs located on each turbulent layer) for comparison.

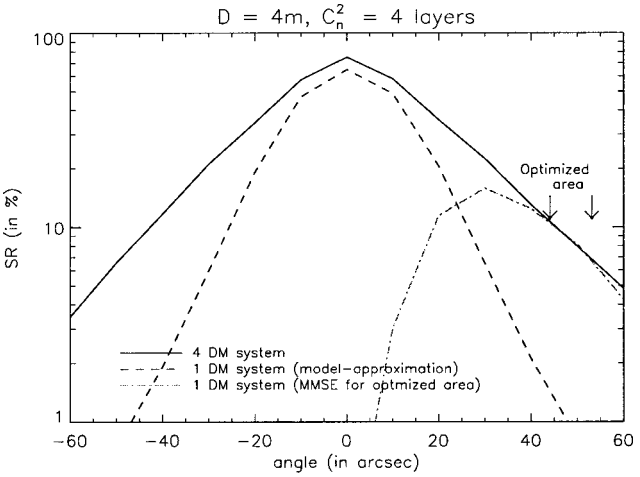

Fig. 7. Comparison of the optimal phase estimation and MA approaches in the case of a one-GS (on the optical axis) and one-DM (conjugated at 6.5 km) system. Note that the DM is misplaced with regard to the C_n^2 profile. In each case, an X cut of the FOV is presented. These simulations are made for a four-layer C_n^2 profile and a 4-m telescope. We plot the tomographic reconstruction (four DMs located on each turbulent layer) for comparison.

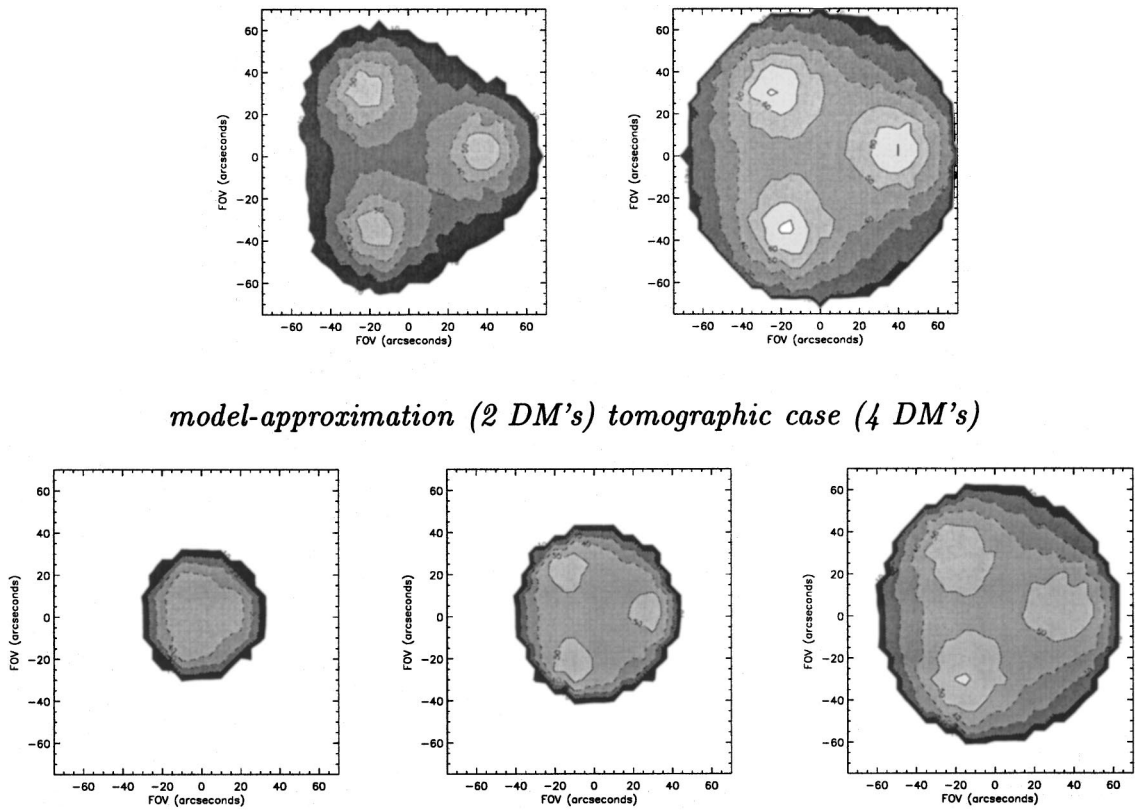

Fig. 8. Upper plots: comparison of iso-SR maps between the MA approach (two DMs) and the tomographic case (four DMs). Lower plots: iso-SR maps computed by using the optimal approach in the case of a two-DM and 3-GS system (case c). Three optimized FOVs are considered: 20, 60, and 120 arc sec. Note that, in all maps, only SR $\geqslant$ 10% are plotted. Black corresponds to 10%, and the succeeding levels (dark to light) are 20%, 30%... .

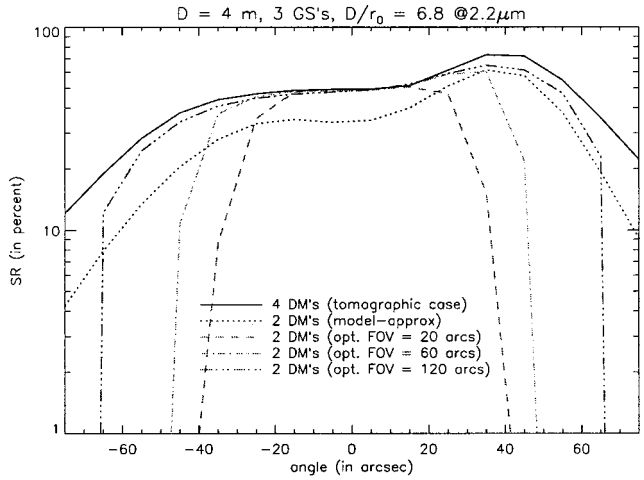

Fig. 9. Comparison of the optimal and MA approaches is for two DMs and a four-layer atmospheric profile. The FOVs of interest are, for the optimal approach, 20, 60, and 120 arc sec. The MA approach (in which the results are independent of a given FOV) is plotted as a dotted curve. The tomographic phase estimation (four DMs in the four turbulent layers) is plotted (solid curve) for comparison. All these curves are, in fact, an X cut (at $Y = 0$) of each corresponding iso-SR map presented in Fig. 8.

5-arc sec $\times$ 5-arc sec FOV around $\alpha = 50$ arc sec. The optimal approach given by Eq. (17) is then used, and we obtain with only one DM nearly the same results as those for four DMs (the difference is only approximately 0.1% in SR) in this particular portion of the FOV.

- Case of a misplaced DM with respect to the turbulence profile (Fig. 7).

The atmospheric conditions are the same as those in the case above; the only difference is in the DM position. Here the DM is misplaced with respect to the turbulence profile (DM located at 6.5 km). Of course, the model approximation gives a poor extrapolation, considering that to regard all the turbulence as being concentrated at 6.5 km is a bad approximation. In Fig. 7, the MA results (dashed curve) are strongly degraded compared with those in Fig. 6. But it is shown that the optimal reconstruction approach still gives good results, similar to the well-placed DM case, in the specified 5-arc sec $\times$ 5-arc sec FOV around $\alpha = 50$ arc sec. Since we have taken into account the true C_n^2 profile, the optimal approach is able to find the best DM deformation to optimize the correction in the direction of the science object, even if the DM position is far from the optimum.

To summarize, even if we have only one GS, the best way to optimize the correction in a large FOV is to have the same number of DMs as the number of turbulent layers. Of course, this is impossible for practical reasons; we have then shown that the use of the information on the true C_n^2 profile in the reconstruction process yields impressive results even if we have only one DM. This approach could therefore be used in conventional AO systems (case of one DM at 0 km; that is, in the pupil plane) to increase their performance in the field when the C_n^2 measurements are available.

However, high and quasi-uniform correction quality in the whole FOV cannot be achieved with only one GS, and in Subsection 4.D.2 three-GS configurations are studied.

2. Multi-Guide-Star and Multi-Deformable-Mirror System

Now let us consider a more complex MCAO system (case *c* of Subsection 4.B) composed of two DMs and three GSs located on the vertices of an equilateral triangle (GS separation = 70 arc sec). For the two-DM system, we plot in Figs. 8 and 9 a comparison between the optimal reconstruction method (for different FOVs of interest: 20, 60, and 120 arc sec) and the MA reconstruction method. The best-performance case, obtained with four DMs conjugated in the four turbulent layers (case d), is also plotted for comparison.

Figures 8 and 9 show the appeal of the optimal phase reconstruction approach, which allows an optimal reconstruction in the FOV of interest $\{\boldsymbol{\alpha}\}_{\mathrm{FOV}}$. For example, let us consider a FOV of interest centered on the optical axis and having a size of 20, 60, or 120 arc sec. With only two DMs, the correction is nearly the same as the best performance obtained with the four-DM system. The correction at the center of the FOV, which is equal only to 31% (in terms of SR) with the MA approach, is equal to 49.3% for the 20-arc sec × 20-arc sec optimized area, 48.5% for the 60-arc sec × 60-arc sec optimized area, and 47% for the 120-arc sec × 120-arc sec optimized area. The SR for the limit case of four DMs is equal to 49.6%. We note the significant increase of the SR when using the true C_n^2 profile in the optimal approach. A very slow decrease of on-axis performance with increasing FOV of interest is observed.

Note that in both the full tomographic and the MA case the best reconstruction is achieved in the GS directions. Of course, it is important that the FOV of interest be well specified, since outside this region of interest the SR decreases quickly. This is a consequence of the tight compromise performed in the optimal estimation.

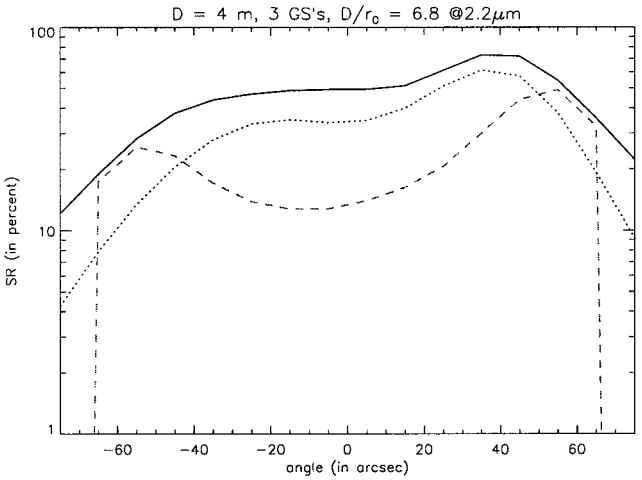

Fig. 10. Comparison of the optimal (dashed curve) and the MA (dotted curve) approach for two DMs and a four-layer atmospheric profile. The FOV of interest is, for the optimal approach, two areas of 5-arc sec diameter located at −60 and 60 arc sec. The tomographic phase estimation (four DMs in the four turbulent layers) is plotted (solid curve) for comparison.

Another example is presented in Fig. 10, where the optimized FOV consists of two areas located at −60 and 60 arc sec (the size of each area is only 5 arc sec in diameter). In that case, for the −60-arc sec position the SR goes from 7.2% (two-DM MA reconstruction) to 23% (two-DM optimal reconstruction), and for the 60-arc sec position the corresponding SR goes from 27% to 44%. Of course, we have globally lost in the 150-arc sec × 150-arc sec FOV, but we have optimized the correction in the two areas of interest.

5. CONCLUSION

In this paper, we have presented an optimal phase reconstruction for MCAO systems. This optimal approach derives from a MMSE estimator that minimizes the mean residual phase variance in the FOV of interest. This optimal approach accounts for the fact that the number of DMs is always smaller than the number of atmospheric layers. It is shown to correspond to a full tomographic reconstruction of the turbulence volume followed by a projection on the DMs. This optimal approach requires a good knowledge of the C_n^2 profile, and therefore a generalized SCIDAR must be coupled to the MCAO system. Even if the mirror positions are not well adapted to the current C_n^2 profile, the knowledge of this profile is incorporated into our estimator, which therefore provides a quasi-ultimate performance in the FOV of interest.

For comparison, we show the result obtained with a cruder approach. An equivalent C_n^2 profile composed of a small number of ELs (equal to the number of DMs) is computed, and we assume that each DM matches an EL. This can be an interesting alternative when only a crude C_n^2 profile knowledge is available.

In both cases, great care has been taken to regularize the inverse problem. This well-chosen regularization brings a nonnegligible gain in the phase reconstruction for a large FOV, compared with that from a classical least-squares estimation using a truncated SVD. To achieve a given performance, a system using an optimized approach will need fewer GSs and hence will gain in terms of sky coverage (if natural GSs are used) or system complexity (if laser GSs are used). Of course, high SRs in a large FOV require a larger number of GSs and a slight increase of DM number. However, even in such demanding conditions, the optimized reconstructor still limits the increase of the system complexity.

In the present paper, we consider only an open-loop case (which allows us to use the turbulent phase covariance matrix), but future work should include a complete closed-loop modeling.

APPENDIX A: MINIMUM-MEAN-SQUARE-ERROR SOLUTION

1. Matrix Differentiation

Let us first recall some important results, used in the paper, on the theory of matrix differentiation.[33] Let us consider two matrices $\mathcal{A}$ and $\mathcal{B}$ and define the first derivative of $\mathcal{A}$ with respect to $\mathcal{B}$ as follows:

$$\frac{\partial \mathcal{A}}{\partial \mathcal{B}} = \begin{bmatrix} \dfrac{\partial \mathcal{A}}{\partial \mathcal{B}_{11}} & \cdots & \dfrac{\partial \mathcal{A}}{\partial \mathcal{B}_{1n}} \\ \vdots & & \vdots \\ \dfrac{\partial \mathcal{A}}{\partial \mathcal{B}_{m1}} & \cdots & \dfrac{\partial \mathcal{A}}{\partial \mathcal{B}_{mn}} \end{bmatrix}. \tag{A1}$$

Now let us suppose that $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$ are real matrices and that $\mathbf{x}$ is a vector. Then the following properties hold:

- $\partial \mathcal{A}/\partial \mathcal{B}^{\mathrm{T}} = (\partial \mathcal{A}^{\mathrm{T}}/\partial \mathcal{B})^{\mathrm{T}}$.
- $(\partial/\partial \mathcal{B})[\mathrm{trace}(\mathcal{B}\mathcal{A})] = (\partial/\partial \mathcal{B})[\mathrm{trace}(\mathcal{A}^{\mathrm{T}}\mathcal{B})] = (\partial/\partial \mathcal{B})[\mathrm{trace}(\mathcal{A}\mathcal{B})] = \mathcal{A}^{\mathrm{T}}$.
- $(\partial/\partial \mathcal{B})[\mathrm{trace}(\mathcal{B}^{\mathrm{T}}\mathcal{A})] = (\partial/\partial \mathcal{B})[\mathrm{trace}(\mathcal{B}\mathcal{A}^{\mathrm{T}}] = \mathcal{A}$.
- $(\partial/\partial \mathcal{B})[\mathrm{trace}(\mathcal{A}\mathcal{B}\mathcal{C})] = \mathcal{A}^{\mathrm{T}}\mathcal{C}^{\mathrm{T}}$.
- $(\partial/\partial \mathcal{B})[\mathrm{trace}(\mathcal{B}\mathcal{A}\mathcal{B})^{\mathrm{T}})] = 2\mathcal{B}\mathcal{A}$.
- $(\partial/\partial \mathcal{B})(\mathbf{x}^{\mathrm{T}}\mathcal{B}\mathcal{A}\mathcal{B}^{\mathrm{T}}\mathbf{x}) = 2\mathbf{x}\mathbf{x}^{\mathrm{T}}\mathcal{B}\mathcal{A}$, which leads to $(\partial/\partial \mathcal{B})[\mathrm{trace}(\mathcal{C}^{\mathrm{T}}\mathcal{B}\mathcal{A}\mathcal{B}^{\mathrm{T}}\mathcal{C})] = 2\mathcal{C}^{\mathrm{T}}\mathcal{C}\mathcal{B}\mathcal{A}$.

2. Optimal Minimum-Mean-Square-Error Solution

The goal is to derive the MMSE criterion defined in Eq. (16) with respect to $\mathcal{W}$. First, let us recall that for a given matrix $\mathcal{A}$ and a given vector $\mathbf{v}$, we have the following relation: $\|\mathcal{A}\mathbf{v}\|^2 = \mathrm{trace}[\mathcal{A}\mathbf{v}(\mathcal{A}\mathbf{v})^{\mathrm{T}}]$. Then, assuming that the noise and turbulent phase statistics are independent, Eq. (16) becomes

$$\begin{aligned} \boldsymbol{\epsilon} = \int_{\{\boldsymbol{\alpha}\}_{\mathrm{FOV}}} \mathrm{trace}[&(\mathcal{M}_{\boldsymbol{\alpha}}^{N_{\mathrm{EL}}}\mathcal{W}\mathcal{M}_{N_{\mathrm{GS}}}^{N_t} - \mathcal{M}_{\boldsymbol{\alpha}}^{N_t}) \\ &\times \langle \boldsymbol{\phi}\boldsymbol{\phi}^{\mathrm{T}}\rangle (\mathcal{M}_{\boldsymbol{\alpha}}^{N_{\mathrm{EL}}}\mathcal{W}\mathcal{M}_{N_{\mathrm{GS}}}^{N_t} - \mathcal{M}_{\boldsymbol{\alpha}}^{N_t})^{\mathrm{T}} \\ &+ \mathcal{M}_{\boldsymbol{\alpha}}^{N_{\mathrm{EL}}}\mathcal{W}\langle \mathbf{n}\mathbf{n}^{\mathrm{T}}\rangle (\mathcal{M}_{\boldsymbol{\alpha}}^{N_t}\mathcal{W})^{\mathrm{T}}]\mathrm{d}\boldsymbol{\alpha}. \end{aligned} \tag{A2}$$

$\langle \boldsymbol{\phi}\boldsymbol{\phi}^{\mathrm{T}}\rangle$ ($\langle \mathbf{n}\mathbf{n}^{\mathrm{T}}\rangle$) are denoted $\mathcal{C}_{\boldsymbol{\phi}}$ ($\mathcal{C}_{\mathbf{n}}$) and defined as

$$\mathcal{C}_{\boldsymbol{\phi}} = \begin{bmatrix} \langle \phi_1\phi_1^{\mathrm{T}}\rangle & 0 & 0 & 0 & 0 \\ 0 & \ddots & 0 & 0 & 0 \\ 0 & 0 & \langle \phi_j\phi_j^{\mathrm{T}}\rangle & 0 & 0 \\ 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 0 & \langle \phi_{N_t}\phi_{N_t}^{\mathrm{T}}\rangle \end{bmatrix},$$

$$\mathcal{C}_{\mathbf{n}} = \begin{bmatrix} \langle n_1 n_1^{\mathrm{T}}\rangle & 0 & 0 & 0 & 0 \\ 0 & \ddots & 0 & 0 & 0 \\ 0 & 0 & \langle n_i n_i^{\mathrm{T}}\rangle & 0 & 0 \\ 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 0 & \langle n_{N_{\mathrm{GS}}} n_{N_{\mathrm{GS}}}^{\mathrm{T}}\rangle \end{bmatrix}, \tag{A3}$$

which can be seen as an N_t-layer (N_{GS}-GS) turbulence (noise) covariance matrix. Note that each $\langle \phi_j\phi_j^{\mathrm{T}}\rangle$ is the Kolmogorov covariance matrix defined for the jth layer. Note that we have assumed that all the turbulent layers are statistically independent,[21] as well as the noise on each GS measurement. Finally, the criterion to be minimized with respect to $\mathcal{W}$ is

$$\begin{aligned} \boldsymbol{\epsilon} = \int_{\{\boldsymbol{\alpha}\}_{\mathrm{FOV}}} \mathrm{trace}[&\mathcal{M}_{\boldsymbol{\alpha}}^{N_t}\mathcal{C}_{\boldsymbol{\phi}}(\mathcal{M}_{\boldsymbol{\alpha}}^{N_t})^{\mathrm{T}} \\ &+ \mathcal{M}_{\boldsymbol{\alpha}}^{N_{\mathrm{EL}}}\mathcal{W}\mathcal{M}_{N_{GS}}^{N_t}\mathcal{C}_{\boldsymbol{\phi}}(\mathcal{M}_{N_{GS}}^{N_t})^{\mathrm{T}}\mathcal{W}^{T}(\mathcal{M}_{\boldsymbol{\alpha}}^{N_{\mathrm{EL}}})^{\mathrm{T}} \\ &- 2\mathcal{M}_{\boldsymbol{\alpha}}^{N_t}\mathcal{C}_{\boldsymbol{\phi}}(\mathcal{M}_{N_{\mathrm{GS}}}^{N_t})^{\mathrm{T}}\mathcal{W}^{\mathrm{T}}(\mathcal{M}_{\boldsymbol{\alpha}}^{N_{\mathrm{EL}}})^{\mathrm{T}} \\ &+ \mathcal{M}_{\boldsymbol{\alpha}}^{N_{\mathrm{EL}}}\mathcal{W}\mathcal{C}_{\mathbf{n}}\mathcal{W}^{\mathrm{T}}(\mathcal{M}_{\boldsymbol{\alpha}}^{N_{\mathrm{EL}}})^{\mathrm{T}}]\mathrm{d}\boldsymbol{\alpha}. \end{aligned} \tag{A4}$$

Using the formulas of the matrix differentiation described in Appendix A.1, we obtain

$$\begin{aligned} \frac{\partial \boldsymbol{\epsilon}}{\partial \mathcal{W}} = \int_{\{\boldsymbol{\alpha}\}_{\mathrm{FOV}}} [&(\mathcal{M}_{\boldsymbol{\alpha}}^{N_{\mathrm{EL}}})^{\mathrm{T}}\mathcal{M}_{\boldsymbol{\alpha}}^{N_{\mathrm{EL}}}\mathcal{W}\mathcal{M}_{N_{\mathrm{GS}}}^{N_t}\mathcal{C}_{\boldsymbol{\phi}}(\mathcal{M}_{N_{GS}}^{N_t})^{\mathrm{T}} \\ &- (\mathcal{M}_{\boldsymbol{\alpha}}^{N_{\mathrm{EL}}})^{\mathrm{T}}\mathcal{M}_{\boldsymbol{\alpha}}^{N_t}\mathcal{C}_{\boldsymbol{\phi}}(\mathcal{M}_{N_{\mathrm{GS}}}^{N_t})^{\mathrm{T}} \\ &+ (\mathcal{M}_{\boldsymbol{\alpha}}^{N_{\mathrm{EL}}})^{\mathrm{T}}\mathcal{M}_{\boldsymbol{\alpha}}^{N_t}\mathcal{W}\mathcal{C}_{\mathbf{n}}]\mathrm{d}\boldsymbol{\alpha} = \mathbf{0}, \end{aligned} \tag{A5}$$

where $\mathbf{0}$ is the null matrix (matrix with all elements equal to 0). Equation (A5) leads to the final result:

$$\begin{aligned} \mathcal{W} = &\left[\int_{\{\boldsymbol{\alpha}\}_{\mathrm{FOV}}} (\mathcal{M}_{\boldsymbol{\alpha}}^{N_{\mathrm{EL}}})^{\mathrm{T}}\mathcal{M}_{\boldsymbol{\alpha}}^{N_{\mathrm{EL}}}\mathrm{d}\boldsymbol{\alpha}\right]^{+}\left[\int_{\{\boldsymbol{\alpha}\}_{\mathrm{FOV}}} (\mathcal{M}_{\boldsymbol{\alpha}}^{N_{\mathrm{EL}}})^{\mathrm{T}}\mathcal{M}_{\boldsymbol{\alpha}}^{N_t}\mathrm{d}\boldsymbol{\alpha}\right] \\ &\times \mathcal{C}_{\boldsymbol{\phi}}(\mathcal{M}_{N_{\mathrm{GS}}}^{N_t})^{\mathrm{T}}[\mathcal{M}_{N_{\mathrm{GS}}}^{N_t}\mathcal{C}_{\boldsymbol{\phi}}(\mathcal{M}_{N_{\mathrm{GS}}}^{N_t})^{\mathrm{T}} + \mathcal{C}_{\mathbf{n}}]^{-1}, \end{aligned} \tag{A6}$$

where $\mathcal{A}^{+}$ denotes the generalized inverse of the matrix $\mathcal{A}$.

Additional author information: All authors: phone, 33-1-46-73-40-40; fax, 33-1-46-73-41-71; e-mail, {last name}@onera.fr. URL: http://www.onera.fr/dota.

REFERENCES

1. J. W. Hardy, J. E. Lefevbre, and C. L. Koliopoulos, "Real time atmospheric compensation," J. Opt. Soc. Am. **67**, 360–369 (1977).
2. G. Rousset, J.-C. Fontanella, P. Kern, P. Gigan, F. Rigaut, P. Léna, C. Boyer, P. Jagourel, J.-P. Gaffard, and F. Merkle, "First diffraction-limited astronomical images with adaptive optics," Astron. Astrophys. **230**, 29–32 (1990).
3. F. Roddier, ed., *Adaptive Optics in Astronomy* (Cambridge U. Press, Cambridge, UK, 1999).
4. D. L. Fried, "Anisoplanatism in adaptive optics," J. Opt. Soc. Am. **72**, 52–61 (1982).
5. F. Chassat, "Calcul du domaine d'isoplanétisme d'un système d'optique adaptative fonctionnant à travers la turbulence atmosphérique," J. Opt. (Paris) **20**, 13–23 (1989).
6. T. Fusco, J.-M. Conan, L. Mugnier, V. Michau, and G. Rousset, "Characterisation of adaptive optics point spread function for anisoplanatic imaging. Application to stellar field deconvolution," Astron. Astrophys. Suppl. Ser. **142**, 149–156 (2000).
7. R. H. Dicke, "Phase-contrast detection of telescope seeing and their correction," Astron. J. **198**, 605–615 (1975).
8. J. M. Beckers, "Increasing the size of the isoplanatic patch with multiconjugate adaptive optics," in *Very Large Telescopes and Their Instrumentation*, M. H. Ulrich, ed. (European Southern Observatory, Garching, Germany, 1988), pp. 693–703.
9. R. Ragazzoni, "No laser guide stars for adaptive optics in giant telescopes?" Astron. Astrophys. Suppl. Ser. **136**, 205–209 (1999).
10. T. Fusco, J.-M. Conan, V. Michau, L. Mugnier, and G. Rousset, "Efficient phase estimation for large field of view adaptive optics," Opt. Lett. **24**, 1472–1474 (1999).
11. T. Fusco, J.-M. Conan, V. Michau, L. M. Mugnier, and G. Rousset, "Phase estimation for large field of view: application to multiconjugate adaptive optics," in *Propagation through the Atmosphere III*, M. C. Roggemann and L. R. Bissonnette, eds., Proc. SPIE **3763**, 125–133 (1999).
12. T. Fusco, J.-M. Conan, V. Michau, G. Rousset, and L.

Mugnier, "Isoplanatic angle and optimal guide star separation for multiconjugate adaptive optics," in *Adaptive Optical Systems Technology*, P. Wizinowich, ed., Proc. SPIE **4007**, 1044–1055 (2000).
13. A. Tokovinin, M. Le Louarn, and M. Sarazin, "Isoplanatism in multiconjugate adaptive optics system," J. Opt. Soc. Am. A **17**, 1819–1827 (2000).
14. M. Le Louarn, N. Hubin, M. Sarazin, and A. Tokovinin, "New challenges for adaptive optics: extremely large telescopes," Mon. Not. R. Astron. Soc. **317**, 535–544 (2000).
15. B. L. Ellerbroek, "First-order performance evaluation of adaptive-optics systems for atmospheric-turbulence compensation in extended-field-of-view astronomical telescopes," J. Opt. Soc. Am. A **11**, 783–805 (1994).
16. M. Tallon, R. Foy, and J. Vernin, "3-D wavefront sensing for multiconjugate adaptive optics," in *Progress in Telescope and Instrumentation Technologies*, M.-H. Ulrich, ed. (European Southern Observatory, Garching, Germany, 1992), pp. 517–521.
17. D. C. Johnston and B. M. Welsh, "Analysis of multiconjugate adaptive optics," J. Opt. Soc. Am. A **11**, 394–408 (1994).
18. R. Flicker, F. Rigaut, and B. Ellerbroek, "Comparison of multiconjugate adaptive optics configurations and control algorithms for the Gemini-South 8-m telescope," in *Adaptive Optical Systems Technology*, P. Wizinovich, ed., Proc. SPIE **4007**, 1032–1043 (2000).
19. M. Tallon and R. Foy, "Adaptive telescope with laser probe: isoplanatism and cone effect," Astron. Astrophys. **235**, 549–557 (1990).
20. R. Ragazzoni, E. Marchetti, and F. Rigaut, "Modal tomography for adaptive optics," Astron. Astrophys. **342**, L53–L56 (1999).
21. F. Roddier, "The effects of atmospherical turbulence in optical astronomy," in *Progress in Optics*, E. Wolf, ed. (North-Holland, Amsterdam, 1981), Vol. XIX, pp. 281–376.
22. G. Rousset, "Wavefront sensing," in *Adaptive Optics for Astronomy*, D. Alloin and J.-M. Mariotti, eds. (Kluwer Academic, Cargèse, France, 1993), pp. 115–137.
23. H. L. Van Trees, *Detection, Estimation, and Modulation Theory* (Wiley, New York, 1968).
24. T. Fusco, J.-M. Conan, V. Michau, G. Rousset, and F. Assémat, "Multiconjugate adaptive optics: comparison of phase reconstruction approaches for large field of view", in *Atmospheric Propagation, Adaptive Systems, and Laser Radar Technology for Remote Sensing*, J. D. Gonglewski, G. W. Kamerman, and A. Kohnle, eds., Proc. SPIE **4167**, 168–179 (2000).
25. E. P. Wallner, "Optimal wave-front correction using slope measurements," J. Opt. Soc. Am. **73**, 1771–1776 (1983).
26. A. Fuchs, M. Tallon, and J. Vernin, "Focusing on a turbulent layer: principle of the 'generalized SCIDAR'," Publ. Astron. Soc. Pac. **110**, 86–91 (1998).
27. R. J. Noll, "Zernike polynomials and atmospheric turbulence," J. Opt. Soc. Am. **66**, 207–211 (1976).
28. F. Rigaut and E. Gendron, "Laser guide star in adaptative optics: the tilt determination problem," Astron. Astrophys. **261**, 677–684 (1992).
29. B. McGlamery, "Computer simulation studies of compensation of turbulence degraded images," in *Image Processing*, J. C. Ulrich, ed., Proc. SPIE **74**, 225–233 (1976).
30. G. Herriot, S. Morris, S. Roberts, M. Fletcher, L. Saddlemyer, J.-P. Singh, G. Véran, and E. Richardson, "Innovations in the Gemini adaptive optics system design," in *Adaptive Optical System Technologies*, D. Bonaccini and R. K. Tyson, eds., Proc. SPIE **3353**, 488–499 (1998).
31. J.-M. Conan, L. M. Mugnier, T. Fusco, V. Michau, and G. Rousset, "Myopic deconvolution of adaptive optics images using object and point spread function power spectra," Appl. Opt. **37**, 4614–4622 (1998).
32. L. M. Mugnier, C. Robert, J.-M. Conan, V. Michau, and S. Salem, "Regularized multiframe myopic deconvolution from wavefront sensing," in *Propagation through the Atmosphere III*, M. C. Roggemann and L. R. Bissonnette, eds., Proc. SPIE **3763**, 134–144 (1999).
33. W. J. Vetter, "Derivative operations on matrices," IEEE Trans. Autom. Control **AC-15**, 241–244 (1970).