Astronomy & Astrophysics manuscript no.
(will be inserted by hand later)

Optimized modal tomography in Adaptive Optics

A. Tokovinin[1,2], M. Le Louarn[3,2], E. Viard[4,2], N. Hubin[2], and R. Conan[2]

[1] Cerro Tololo Inter-American Observatory, Casilla 603, La Serena Chile
e-mail: `atokovinin@ctio.noao.edu`
[2] European Southern Observatory, Karl-Schwarzschild Str. 2, D-85748 Garching bei München Germany
e-mail: `nhubin@eso.org; rconan@eso.org`
[3] Center for Adaptive Optics, Kerr Hall, University of California, Santa Cruz, CA, 95064 USA
e-mail: `lelouarn@ucolick.org`
[4] Astronomical Observatory of Padova, vicolo dell'Osservatorio 5 I-35122 Padova Italy
e-mail: `viard@pd.astro.it`

Received ; accepted

Abstract. The performance of modal Multi-Conjugate Adaptive Optics systems correcting a finite number of Zernike modes is studied using a second-order statistical analysis. Both natural and laser guide stars (GS) are considered. Optimized command matrix is computed from the covariances of atmospheric signals and noise, to minimize the residual phase variance averaged over the field of view. An efficient way to calculate atmospheric covariances of Zernike modes and their projections is found. The modal covariance code is shown to reproduce the known results on anisoplanatism and cone effect with single GS. It is then used to study the error of wave-front estimation from several off-axis GSs (tomography). With increasing radius of the GS constellation Θ, the tomographic error increases quadratically at small Θ, then linearly at larger Θ when incomplete overlap of GS beams in the upper atmospheric layers provides the major contribution to this error, especially on low-order modes. It is demonstrated that the quality of turbulence correction with two deformable mirrors is practically independent of the conjugation altitude of the second mirror, as long as command matrix is optimized for each configuration.

Key words. Instrumentation: adaptive optics – Atmospheric effects – telescopes

1. Introduction: AO, MCAO, and Tomography

The purpose of Adaptive Optics (AO) is to correct atmospheric turbulence in real time, improving the image quality of ground-based astronomical telescopes (Roddier 1999). Multi-Conjugate Adaptive Optics (MCAO) is a further development of this concept, where correction is made by several Deformable Mirrors (DMs) conjugated to different altitudes. In this way a wider Field of View (FOV) can be corrected and better correction uniformity can be achieved (Johnston & Welsh 1994, Ellerbroek 1994, Fusco et al. 1999).

Atmospheric tomography was conceived as a method to measure the instantaneous 3-dimensional phase perturbations in the atmosphere, in order to get the control signals for MCAO (Tallon & Foy 1990). Light from several Guide Stars (GSs) is used to probe the 3-dimensional perturbations, which are retrieved by solving an inverse problem as in medical tomography. It has been realized later that the signals to control one or more DMs can be derived directly from the information brought by GSs, without a need to reconstruct the whole turbulent volume. We use the term "tomography" in this sense, probably not quite rigorously. An MCAO system can not work without tomography. However, with only one DM (no MCAO) the tomography can be employed to correct the image of a selected scientific target using the signals from surrounding GSs – either natural stars (NGSs) or artificial Laser Guide Stars (LGSs). The interest of such solution lies in the possibility to correct the so-called *cone effect* (Tyler 1994) which degrades the performance of LGS-based AO systems, or to widen the field where NGS can be found, improving the sky coverage of AO.

In Fig. 1 a scheme of an MCAO system is given. Each Wave-Front Sensor (WFS) measures the atmospheric perturbations for its own GS. Here we suppose that the perturbed wave-fronts are described by a finite number of Zernike modes. Of course, WFS signals also contain some measurement noise. Those signals (e.g. the coefficients of wave-front expansion on Zernike basis) are multiplied by some *command matrix* M to control the shape of the DMs.

Send offprint requests to: A. Tokovinin

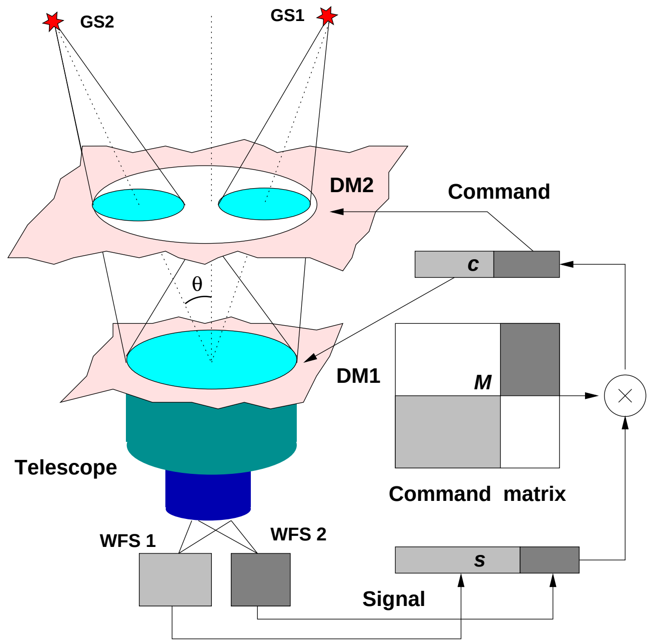

Fig. 1. Scheme of an MCAO with modal control. Atmospheric layers matching the conjugation altitudes of the 2 DMs are shown. Both the WFS signal vector s and the DM command vector c are specified as the coefficients of wavefront expansion on Zernike modes. For DM2, this expansion is defined on a meta-pupil of the diameter $D + 2\theta H_2$, where D is the telescope diameter, θ is the radius of the FOV, H_2 is the conjugation altitude of DM2.

This shape is also specified in terms of Zernike modes, and for this reason such an MCAO system can be called *modal*, as opposed to zonal systems where both WFS and DM signals are specified as local parameters (wavefront slopes or actuator voltages).

Although Zernike modes are slightly sub-optimal for turbulence correction (Roddier 1999), Zernike basis plays an important role in theoretical studies of AO and gives reasonable performance predictions. We extend these studies to 3-dimensional turbulence correction and present an analitycal tool for fast performance estimates which permits to explore rapidly the vast space of system parameters in search for best configurations. Further refined analysis of selected configurations must be done by Monte-Carlo simulation which will take into account many additional details.

In an MCAO system the shape of DMs is driven in such a way as to obtain the best possible correction in closed loop; the WFSs measure then the remaining residual wave-front aberrations. Here we consider an open-loop MCAO system, where the WFSs measure the uncorrected perturbations, not residual aberrations. The DMs correct the turbulence in another (scientific) beam. As shown by Ellerbroek (1994), this approximation can describe a real (closed-loop) AO system under certain assumptions. We neglect the temporal aspects of MCAO operation by supposing that all measurements and corrections are done instantaneously. Thus, our attention is focused on the spatial aspects of turbulence tomography and correction, which are indeed new and specific to MCAO.

The command matrix plays a role of a "magic" black box that receives the WFS signals and produces the reconstructed wave-fronts at the output. How should the command matrix be selected to obtain the best possible correction? What is "the best possible"? How good is the correction finally? Which parameters of MCAO need to be optimized, and how? These are the questions addressed in the present work.

In Sect. 2 we briefly outline possible correction strategies and the method to estimate the first-order MCAO performance based on modal covariances (mathematical details of the derivation are given in Appendix B). Our numerical code is presented in Sect. 3. In Sect. 4 the results on tomography are given, and in Sect. 5 a low-order MCAO system is considered. The conclusions are summarized in Sect. 6.

2. Principles of optimized modal tomography

2.1. Inverse control

The concept of modal MCAO was discussed in the paper of Ragazzoni et al. (1999, hereafter RMR99). It was shown that the deformations of several DMs can be measured (and hence controlled) by measuring a certain number of Zernike modes on natural or laser GSs.

In a classical AO there is a one-to-one correspondence between the wavefront and DM modes, but in MCAO this relation is not so simple. As can be seen in Fig. 1, for a finite Field of View (FOV) of a radius θ, the diameter of the second DM (called *meta-pupil*) must be larger than the telescope pupil D by at least $2\theta H_2$, where H_2 is the conjugation height of the second DM. A beam of some GS illuminates only a portion of the meta-pupil: the beam footprint diameter is smaller than D for an LGS (as shown in the Figure) or equal to D for an NGS. In addition, the footprints are displaced from the center of the meta-pupil as soon as the GSs are not on-axis. The wavefront is decomposed into Zernike modes on a small circle (footprint), and DM commands are decomposed on a larger circle (meta-pupil). The relation between those two sets of modes is called *mode projection*, and is given by a *mode projection matrix*. It is discussed in RMR99 and in Appendix A.

Using mode projections, it is possible to express by a matrix the reaction of all WFSs to a given Zernike mode on a meta-pupil. This *interaction matrix* A relates the vector of WFS signals s to the control signals (commands) c, as in any AO system, and its specific form (projections) is only a consequence of the modal representation of both signals that we choose. As noted in RMR99, the interaction matrix can be inverted, to give a command matrix $M_{\rm inv}$:

$$s = Ac \quad \text{and} \quad c = M_{\rm inv} s, \quad \text{where} \quad M_{\rm inv} = A^{-1*}. \qquad (1)$$

Such a control law is called *inverse* control. In fact the interaction matrix A is not always invertible, and Singular Value Decomposition (SVD) (Press et al. 1992) is typically used to obtain $M_{\rm inv}$ by truncation of some singular values below a pre-selected threshold (A^{-1*} is such pseudo-inverse matrix, and not the inverse matrix A^{-1}). Threshold selection is done by trial and error, it depends on the amount of noise in the WFS signals. In the simplest case of a classical modal AO system with only one WFS and one DM, both interaction matrix and $M_{\rm inv}$ are identity matrices.

2.2. Residual variance

Given some command matrix M and some model of atmospheric turbulence, what is the average variance of the uncompensated (residual) phase $\epsilon^2(\boldsymbol{\theta})$ for a given viewing direction $\boldsymbol{\theta}$? How does it depend on the location in FOV? To answer these questions, second-order statistical quantities (covariances) must be computed. The result (Appendix B) depends on the three matrices S, T, and R:

$$\epsilon^2(\boldsymbol{\theta}) = \epsilon_0^2 - \mathrm{Trace}(2MT^T - RMSM^T) = \epsilon_J^2 + \epsilon_{\rm fit}^2. \qquad (2)$$

Here ϵ_0^2 is the uncorrected atmospheric phase variance on the telescope pupil, superscript T means matrix transpose, and the operator Trace gives the sum of the diagonal elements. It is convenient to sub-divide $\epsilon^2(\boldsymbol{\theta})$ into a sum of the variance ϵ_J^2 of the first corrected J Zernike modes (piston excluded) and the variance of higher uncorrected modes, or fitting error $\epsilon_{\rm fit}^2$.

The *signal covariance* matrix $S_{ij} = \langle s_i s_j \rangle$ is a statistical covariance of all pairs of WFS signals (modes). It depends on the turbulence vertical profile, GS geometry, and noise.

The *target matrix* T depends on the observed target location $\boldsymbol{\theta}$, but also on the DM parameters and atmospheric parameters. It is computed as a product of the system projection matrix and a covariance matrix between Zernike modes of the target and those of GSs.

The *cross-talk* matrix R is direction-dependent and describes the mutual influence of DM commands. It does not depend either on the GSs nor on the atmospheric model. When R is singular, it means that two (or more) combinations of commands produce the same effect, and hence are redundant. For example, R is singular when tip and tilt modes are not removed from all DMs except the first one. The redundant modes are usually removed by SVD filtering of R.

2.3. Optimized control

We might use Eq. 2 to derive such a control matrix that minimizes $\epsilon^2(\boldsymbol{\theta})$. But for which direction $\boldsymbol{\theta}$? A standard approach (Ellerbroek 1994) consists in choosing some set of directions with suitable weights W_θ and in minimizing the FOV-weighted $\bar{\epsilon}^2$. To do this, matrices T and R in Eq. 2 must be replaced by the matrices $\bar{T}$ and $\bar{R}$ which are averaged over the FOV and are direction-independent. The optimized command matrix $M_{\rm opt}$ is then found:

$$M_{\rm opt} = \bar{R}^{-1*}\bar{T}S^{-1}, \qquad (3)$$

a result known for a long time in classical AO (Wallner 1983).

At first look, inverse and optimized command matrices are computed from completely different formulas. Are they indeed so different? The optimized control matrix is a specific case of a Wiener filter and takes into account the noise and turbulence statistics. It does not control the modes for which there is not enough information. At high signal to noise ratio it tends to something resembling the inverse control matrix. The relation between inverse control and optimized control looks like a relation between simple inverse filtering and least-squares (Wiener) filtering in a linear image restoration problem.

When the matrix $\bar{R}$ is singular, it is inverted by SVD by selecting a threshold for weak modes rejection. Of course, both $\bar{R}$ and $M_{\rm opt}$ depend on the weights W_θ. On the other hand, S is always non-singular when the WFS noise is taken into account.

3. Covariance code

A code has been developed which implements the above principles and permits to compute an optimized or inverse command matrix and to estimate the residual variance in selected directions for a modal MCAO system.

Computation of the atmospheric covariance of Zernike modes is the most essential and computer-intensive part of the code. The formulas of Whiteley et al. (1998) express these covariances through the integrals of the triple products of Bessel functions. This calculation was carefully optimized using some term re-arrangement, pre-computing, etc., gaining 10-100 times in computing speed compared to the straightforward integration (Appendix C). Mode projection is another important part of the code (Appendix A).

A similar approach to modal MCAO is being developed by Fusco et al. (1999, 2000, 2001). They compute modal covariances using the analytical Noll formulas and projection relations (thus not accounting for a finite turbulence outer scale). An optimized command matrix is then used in Monte-Carlo simulation to find the residual variance, while we compute it directly from the covariances, without recourse to system simulation.

In the rest of this section we describe the model of the MCAO system used in our simulation and the tests of the code.

3.1. System parameters

The atmosphere is modeled here as several discrete layers. Throughout this paper a 7-layer model for Cerro Pachon is used (Ellerbroek & Rigaut 2000) with $r_0 = 0.15$ m at $\lambda_0 = 0.5$ μm and isoplanatic patch $\theta_0 = 2.47''$. All phase

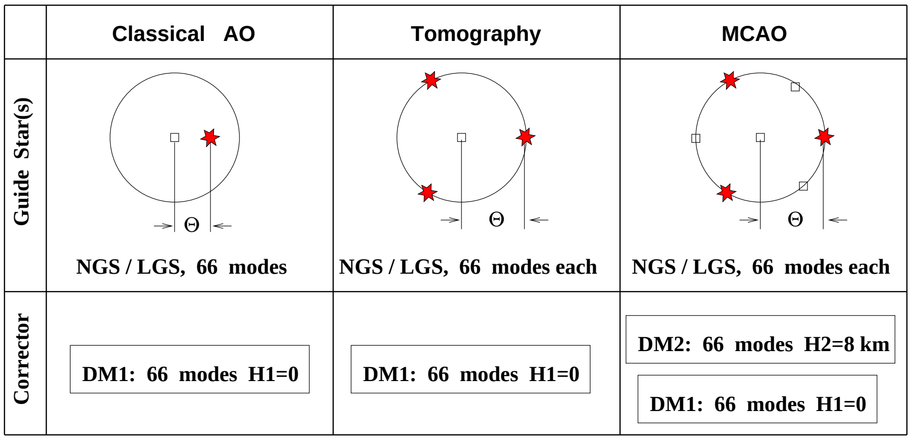

Fig. 2. Representative configurations of AO systems on a 8 m telescope considered in this paper. For a classical system (left), the Zernike modes 2–66 are supposed to be measured and corrected on a natural or laser guide star, possibly displaced from the scientific target (small square). In tomography (center), the wave-front is still corrected by one DM, but wavefronts from 3 off-axis NGSs or LGSs are measured. Finally, an MCAO system attemps to correct a larger field with 2 DMs, minimizing the wavefront errors averaged over several points in the field (squares).

variances are computed for the wavelength λ_0. The turbulence outer scale $L_0 = 25$ m is assumed (Martin et al. 2000).

We apply the code to the case of an 8 m telescope with low-order MCAO system correcting Zernike modes number 2–66 (radial order 10), as appropriate for near-IR imagery. For our atmospheric model, the uncorrected variance of those modes is 170.2 rad^2 (tip-tilt included). For comparison, with infinite outer scale the variance is 822 rad^2. The fitting error (variance of modes 67 and higher) is $\epsilon^2_{\rm fit} = 5.91$ rad^2. To compute the Strehl ratio $SR(\lambda)$ at some imaging wavelength λ, the variance of corrected modes ϵ^2_J must be added to $\epsilon^2_{\rm fit}$ and re-scaled: $SR(\lambda) = \exp[-(\epsilon^2_J + \epsilon^2_{\rm fit})(\lambda_0/\lambda)^2]$.

The noise in the WFSs (assuming that they are of Shack-Hartmann type) is simulated by the method of Rigaut & Gendron (1992). The correspondence between noise level and stellar magnitude is affected by system-dependent WFS parameters. In this paper we focus our attention on the compensation of atmospheric errors and consider only relatively bright GSs, hence imprecise noise modeling is not important.

In Fig. 2 the representative AO systems considered in the rest of the paper are shown schematically. We begin by calculating the anisoplanatic and cone-effect errors for a classical AO system with single GS and one DM (subsections 3.2 and 3.3). Then in Section 4 the possibility to correct an on-axis target with one DM and several off-axis GSs (tomography) is investigated. Full MCAO systems correcting a large field with several DMs are studied in Section 5.

3.2. Modal anisoplanatism

Chassat (1989, 1992) has derived fast-converging series to compute modal covariances for the case of NGS and infinite outer scale using Mellin transform. In the general case (LGS, finite outer scale) such series do not exist (Sasiela 1994) and covariances must be calculated by integration. In Fig. 3 the rms difference between on-axis and off-axis modes from 2 to 66, as computed with Chassat series for Cerro Pachon atmospheric model, is plotted as asterisks.

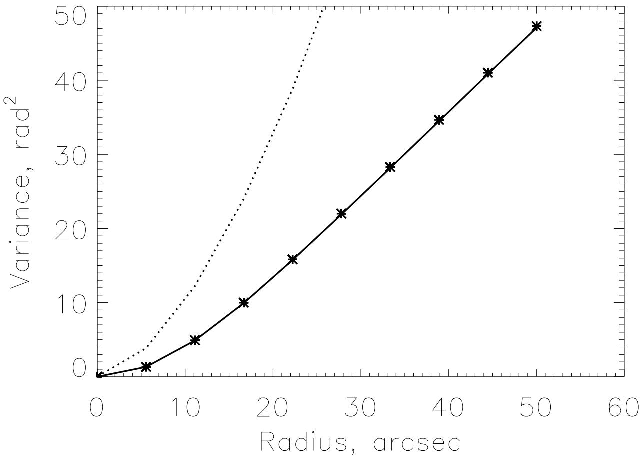

Fig. 3. Residual variance ϵ^2_{66} of the corrected modes 2-66 in function of the NGS off-axis angle, as given by the covariance code (full line) and by the Chassat series (asterisks). Dotted line shows the $(\theta/\theta_0)^{5/3}$ law.

We simulated with our code a conventional AO system with identity command matrix correcting Zernike modes 2-66 with an off-axis NGS (an infinite outer scale was assumed in this case). The full line shows the residual variance as a function of the NGS off-axis distance. A very good agreement with Chassat attests that covariances are computed correctly.

The dotted line in Fig. 3 shows the $(\theta/\theta_0)^{5/3}$ "Fried" anisoplanatic phase variance (Fried 1982). It is superior to the modal anisoplanatism, as noted by many authors (Chassat 1989, Stone et al. 1994, Chun 1998). At small angles the Fried variance increases mostly due to small-scale perturbations which are not included in the 66 corrected modes. At larger angles a contribution of differential piston term to the Fried variance becomes significant. So, the corrected FOV of a low-order AO system is 1.5-2 times larger than the atmospheric isoplanatic patch θ_0.

3.3. Cone effect

The beam of an LGS is converging, and this difference from the stellar beam provokes an error in the corrected

Table 1. Phase variance caused by the cone effect (rad^2 at 0.5 μm)

$H_{\rm LGS}$, km	d_0, m	$(D/d_0)^{5/3}$	ϵ^2_{66}	ϵ^2_{231}
90	4.81	2.33	1.67	2.16
60	3.45	4.07	3.40	4.15
30	2.08	9.46	9.87	11.03

wave-front, known as *cone effect*. The theory predicts this error to be $(D/d_0)^{5/3}$ (Roddier 1999), where d_0 depends on the LGS altitude and on the turbulence profile. For the Pachon atmospheric model the d_0 was computed with formulas of Tyler (1994) for three LGS altitudes $H_{\rm LGS}$. In Table 1 the residual variances are given for a classical AO system at 8 m telescope with an identity command matrix correcting modes up to 66 or up to 231 (radial orders 10 and 20, respectively).

The cone effect for 66 modes is less than $(D/d_0)^{5/3}$ for the same reason as the modal anisoplanatism: low-order modes are better correlated. With increased correction order the cone effect estimated by the covariance code approaches the analytical predictions. Given the approximate nature of the latter, the agreement can be judged satisfactory.

The tests described above show that our code reproduces the known analytical results for classical AO systems.

4. Tomography

In this section an AO system with several GSs and one DM is studied in order to test modal tomography. First, the influence of the angular size of the GS constellation on tomographic reconstruction is evaluated. It is shown that the tomographic error is mostly caused by incomplete overlap of GS beams in the higher layers, exceeding significantly the tomographic error for infinite beam size. The tomographic correction of the cone effect inherent to LGSs is addressed as well; the best correction is achieved when LGSs are projected at a distance of 4 m from the optical axis, providing a good coverage of the 8 m aperture. In this case the tip and tilt are supposed to be perfectly corrected; because of that, the tomographic error at large LGS separations is smaller than for the NGSs, meaning that most of the tomographic error resides in the low-order modes. Finally, still in the LGS context, we study the possibility to estimate atmospheric tilt not from one, but from several more widely spaced NGSs, and show that "tilt tomography" brings a potential gain in the sky coverage.

4.1. Dependence of tomographic error on the constellation size

Four combinations of guide stars with varying radius Θ are considered here: i) a symmetric configuration of 3 NGSs; ii) same with 5 NGSs; iii) 3 sodium LGSs (altitude 90 km) with unknown tip and tilt and one NGS at the FOV center measuring tip and tilt only; iv) same for infinite-altitude NGSs. All sources are bright because we want to compute the atmospheric error. The target is supposed to be on-axis, with modes up to 66 corrected by a single DM using a command matrix optimized for each GS configuration. The residual variance of the corrected modes ϵ^2_{66} is plotted in Fig. 4 against Θ.

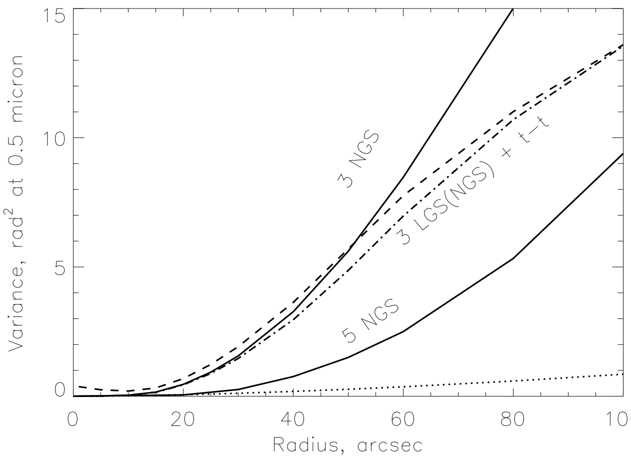

Fig. 4. Residual variance ϵ^2_{66} at the center of FOV for the symmetric configurations of 3 and 5 NGSs at varying distance from the FOV center. Dashed line corresponds to a case of 3 LGSs (unknown tip and tilt) plus an NGS at center measuring modes 2,3. If LGSs are replaced by NGSs in this same configuration, the results are similar (dash-dot line). Dotted line is an atmospheric tomography limit for 3 NGSs.

The tomographic error for a telescope with infinitely large aperture was calculated in Tokovinin & Viard (2001) as $(\Theta/\gamma_K)^{5/3}$, where the tomographic patch size γ_K depends on the number of guide stars K and on the turbulence profile. We estimate $\gamma_3 \approx 110''$ and $\gamma_5 \approx 150''$ for the Cerro Pachon atmosphere. It means that the ultimate atmospheric tomography limit is a way below the modal tomography error at the 8 m telescope! This indicates that for a 8-m class telescope working in the near infrared the major component of tomographic error arises from incomplete overlap of GS beams, and not from the de-correlation between atmospheric layers.

Comparing Fig. 4 with Fig. 3, we see that modal tomography significantly widens the FOV. If the FOV size is arbitrarily defined as a field radius corresponding to a 2-fold reduction of the Strehl ratio $SR(K)$ (13.4 rad^2 error at $\lambda_0 = 0.5$ μm), then it is equal to $20''$, $75''$, and $120''$ when 1, 3, and 5 NGSs are used.

4.2. Correction of the cone effect

The dashed line in Fig. 4 shows results for tomography with 3 bright sodium LGSs which measure all modes from 4 to 66. It is known that when tip and tilt are not constrained by LGSs, either the modes 2-6 need to be measured on a single NGS (Le Louarn & Tallon 2000), or several tip-tilt NGSs are needed (Ellerbroek & Rigaut 2000) in an MCAO system. However, in the case of tomography (one DM only) this complication is not required, and we suppose that tip and tilt only are measured by an additional NGS at the FOV center.

For very small Θ LGSs give worse results, which can be explained by the cone effect (see above). When the radius of LGSs reaches $9''$, their distance from the optical axis is 4 m (at 90 km), which means that LGS beams cover the whole 8-m aperture and tomography is able to reduce the cone effect (the residual error of the first 66 modes is 0.20 rad^2). If additional 3 modes (astigmatisms and defocus) are measured on the central NGS, the error becomes 0.04 rad^2. We did not yet explore the degree of cone effect removal in a higher-order system needed for turbulence correction in the visible.

Remarkably, with increasing Θ the 3-LGS curve falls below the 3-NGS curve. It means that the contribution of modes 2 and 3 (retrieved from tomography for 3 NGSs or measured directly on tip-tilt star for LGSs) is non-negligible. Still better results are obtained with LGSs if more modes (up to 6) are measured on the central NGS, as required for MCAO. When the LGSs are placed at an infinite altitude (dash-dotted line in Fig. 4) the cone effect disappears, but otherwise the results turn out to be similar.

At $\Theta \approx 30''$ the 3-NGS curve changes from quadratic to almost linear growth. This angle corresponds to the beam separation of 0.8 m at a characteristic altitude of 6 km – of the order of the spatial size of the smallest perturbations corrected with 66 modes (radial order 10) on a 8 m aperture. At this same angle the curves for the 3-NGS and 3-LGS cases begin to diverge, signalling that for 3-NGS the contribution of uncorrelated tip and tilt becomes important at larger Θ. The surface of non-overlapping beam portions is roughly proportional to Θ, hence the amount of uncorrelated atmospheric noise and modal anisoplanatism both grow linearly with Θ. This finding supports the crude analytic modeling of the MCAO anisoplanatism suggested by Fusco et al. (2000).

4.3. Tilt tomography

In this sub-section the problem of measuring tip and tilt from an off-axis NGS (as needed for an LGS AO system) is addressed. The variance of modes 2 and 3 is plotted by a dashed line in Fig. 5 against the off-axis distance. The full line shows the variance obtained when 3 off-axis NGSs are used for estimation of the on-axis tilt ("tilt tomography"). Better performance is achieved because the apriori information on turbulence vertical profile is used here in the form of tilt covariances.

For a variance of 1 rad^2 the 3 NGSs can be ~ 3 times further away from the FOV center than a single NGS. The area on the sky increases by a factor of 10, and the probability to find 3 NGSs in a larger field can exceed the single-NGS probability. The sky coverage of an LGS-based AO system can thus be improved by using several NGSs for tilt measurement. However, for larger field size and larger variance, as typical for K-band imagery, the gain in FOV is less ($\sim$ 2 times), making increased sky coverage less likely. This reduced performance is explained again by the incomplete overlap of 8-m beams; for larger apertures the gain in sky coverage brought by tilt tomography will be non-negligible at all wavelengths.

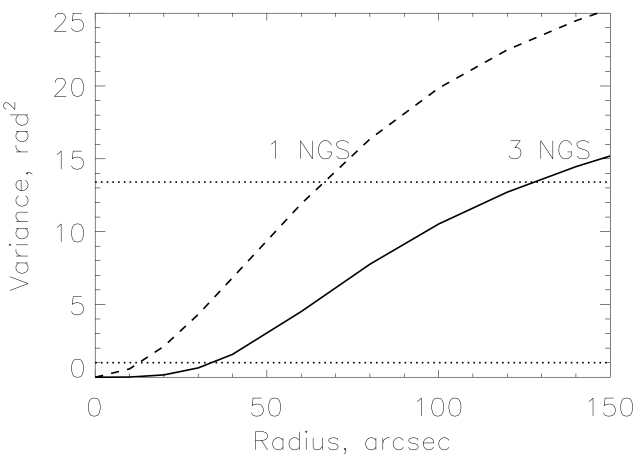

Fig. 5. Residual combined variance of tip and tilt when they are measured on one off-axis NGS (dashed line) and when they are estimated from the 3 off-axis NGSs (full line). Dotted lines show the variance levels of 1 and 13.4 rad^2 (at 0.5 μm).

5. MCAO

In this section, various aspects of a full MCAO modal system are studied. First, we test mode projections which are specific to MCAO and show that projection relations between Zernike modes do permit to control one or more over-sized DMs by measuring the wavefronts on the telescope pupil. Our simulations are compared with the results of two other groups, demonstrating that the modal code predicts similar MCAO performance. Then the selection of the optimum conjugation height H_2 of the second DM in a double-conjugate AO system is addressed: we show that if command matrix is re-adjusted to changing H_2, the MCAO performance remains practically constant. Thus, there is no need to adapt the optical conjugation of DM2 to the changing altitude of turbulent layers, but rather to take this into account in the control. Finally, the superiority of the optimized control over inverse control is illustrated.

5.1. Mode projections

In a conventional AO system with a bright single NGS and one DM, ϵ^2_{66} is zero because all modes up to 66 are perfectly corrected. To test the mode projections, let the DM be conjugated not to telescope pupil, but to $H =$ 8 km, with a meta-pupil size of 9.55 m. Only the central 8 m of the pupil is measured and corrected, although the DM commands are now defined on a larger circle.

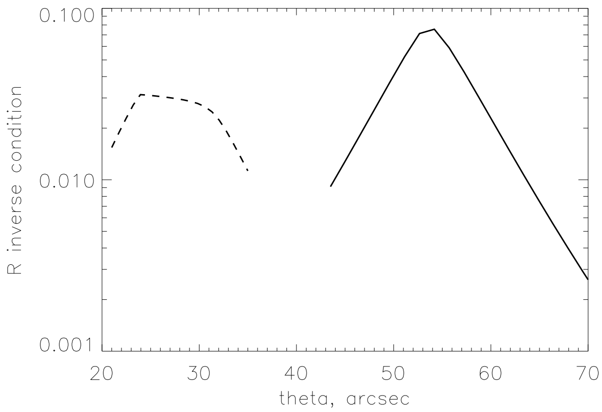

Fig. 6. The inverse condition number of the cross-talk matrix $\bar{R}$ of a 2-DM MCAO system, as a function of the FOV radius θ which defines meta-pupil diameter $D + 2\theta H_2$. Six points of equal weight at the vertices of a hexagon of radius Θ are used as W_θ optimization function. Full curve corresponds to $\Theta = 60''$, dashed curve to $\Theta = 30''$.

Using the inverse command matrix $M_{\rm inv}$, we still obtain $\epsilon^2_{66} = 0$. The system is able to reproduce *exactly* the small-pupil modes by the large-pupil ones. However, $M_{\rm inv}$ has now elements as large as 52 (it was an identity matrix when DM was conjugated to pupil), and the variance of the DM signal is 2600 rad^2, instead of the 170.2 rad^2 atmospheric variance! Of course, this excessive variance corresponds to the unused periphery of the meta-pupil. With optimized control, the command matrix elements are smaller, the DM variance is less (245 rad^2), and $\epsilon^2_{66} = 0.77$ rad^2.

We next consider the problem of the $\bar{R}$-matrix inversion. This matrix depends only on the DM conjugation heights, meta-pupil diameter, and FOV weighting W_θ. A double-conjugate MCAO system correcting Zernike modes up to 66 with DM2 conjugated to $H_2 = 5$ km is chosen. The W_θ consists of 6 equal-weight points at the vertices of a hexagon of radius Θ. Changing the diameter of meta-pupil $D + 2\theta H_2$ (or, equivalently, the unvigneted FOV radius θ) changes the modal basis of the meta-pupil, and hence the cross-talk matrix. As seen in Fig. 6, the inverse condition number of $\bar{R}$ (ratio of the smallest and largest eigenvalues) is a strong function of the radius θ, having a maximum at θ slightly less than Θ. Other weighting patterns were tried (e.g. a triangle with the 4-th heavily weighted point at FOV center), with similar results. It is thus strongly recommended to select meta-pupil size corresponding to $\theta = \Theta$ or slightly less.

5.2. Comparison with other codes

The predictions of the modal covariance code were compared to the results of the simulation of GEMINI MCAO system as given by Ellerbroek & Rigaut (2000). The 3 DMs are conjugated to altitudes 0, 4.5, and 9 km and correct modes up to 66. The 5 LGSs are used, of which 4 are located at the corners of a 60$''$ square field, and the fifth at the FOV center. Four NGSs for tip-tilt measurements are in the middle of the square sides. The 9 weighting directions include center (weight 16/36), NGS positions (weight 4/36) and LGS positions (weight 1/36).

Our code gives weighted residual of $\epsilon^2_{66} = 1.46$ rad^2 (0.66 rad^2 at FOV center). With a threshold of 10^{-3}, 16 modes out of 188 are rejected in the inversion of the $\bar{R}$ matrix. In Fig. 5a of Ellerbroek & Rigaut (2000) the rms OPD of 0.12 μm is given for the similar case, corresponding to 2.3 rad^2 phase variance. The results are quite close, despite the difference between modal basis, optimization procedures, etc.

Fusco et al. (2001) describe the strategy of optimized MCAO control which is very similar (if not identical) to our approach, although presented in a different way. In their Fig. 6 they give $SR(K) = 0.5$ at the center of FOV for a 4 m telescope with 2 DMs, 3 NGSs at a 40.2$''$ radius and for a 4-layer turbulence profile. We computed the performance for the same configuration and found $SR(K) = 0.54$ at the center. Both codes thus give identical results.

5.3. Is the conjugation altitude important?

Now let us study a simple MCAO system with only two DMs to investigate the optimum conjugation height H_2 of the second DM. Three NGSs of magnitude 11 in an equilateral triangle at a distance of 30$''$ from the FOV center are used. One of the 4 FOV weighting points is located at center and the remaining three are configured in a 30$''$-radius triangle rotated by 60° with respect to NGSs. Meta-pupil size at DM2 is equal to $D + 2\theta H_2$, with $\theta = 30''$. The threshold for $\bar{R}$-matrix inversion is 10^{-3}. The results are presented in the Table 2.

Table 2. Residual variance and the variance of DM signals (rad^2 at 0.5 μm) as a function of the conjugation altitude H_2.

H_2, km	ϵ^2_{66}	DM1	DM2	$N_{\rm rej.}$
1	4.67	738.8	826.4	0
4	4.24	155.7	67.6	0
10	3.81	135.4	23.8	3
16	4.29	140.2	19.5	6
20	4.81	143.4	22.7	6
30	4.93	150.2	38.5	9

Surprisingly, the weighted mean variance depends very little on the DM2 conjugation altitude. It corresponds to a $SR(K) \approx 0.59$ (fitting error included). However, as shown in the second and the third columns, the variance of the signals applied to DMs does depend on H_2. When bright NGSs are considered and weak modes are not rejected, the much increased variance of DM signals is obtained, reaching 856 and 936 rad^2 at DM1 and DM2, respectively, for $H_2 = 1$ km. At the same time the correction quality is improved and becomes completely independent of the DM2 conjugation altitude.

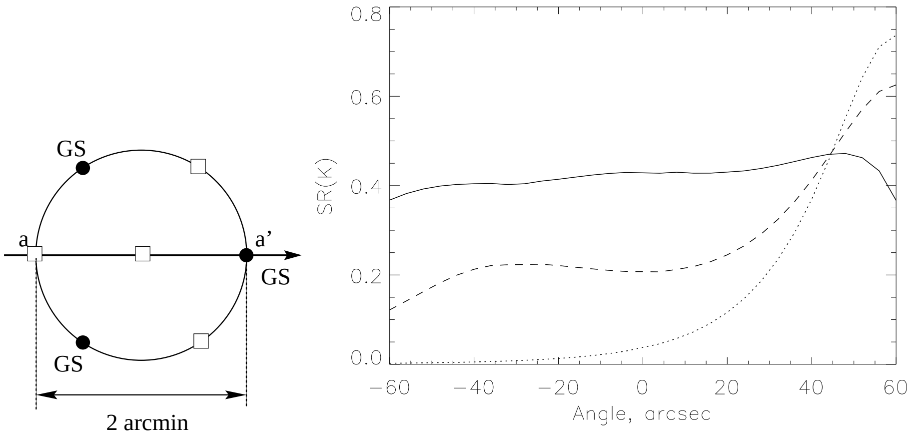

Fig. 7. Strehl ratio in the K band along the aa' cut in the FOV for a 2-DM MCAO system with 3 NGSs in a $120''$ diameter circular field. Left: the GS layout, the squares indicate the optimization points in the FOV. Right: the plot of SR variation on the FOV: full line – optimized command matrix, dashed line – inverse command matrix, dotted line – classical AO with only one NGS.

We offer the following explanation to this behavior. With only 2 DMs it is possible to model the phase in the telescope aperture and its gradient over the FOV. The phase gradient is controlled by the difference of DM shapes, while on-axis phase is defined by their sum. When two DMs are conjugated to 0 and 1 km, large signals of opposite sign must be applied to reproduce the gradient; the variance is hence large and almost equal on both DMs. With increasing H_2, less deformation on DM2 is required to model the gradient. Increasing H_2 further, we obtain quasi-singular $\bar{R}$ matrix and some modes are rejected, reducing the number of controlled parameters and slightly degrading both average residual variance and the correction uniformity over the FOV.

What happens if we conjugate the second DM to a negative altitude? For the altitudes of -4 and -10 km the residual variance remains small, 5.52 and 6.33 rad^2, which is not surprising given the above reasoning. Thus, contrary to the "common sense", DM2 can be conjugated to a negative altitude as long as this is taken into account by optimized command matrix.

This result seems to be of practical importance. Even if the turbulence profile was constant, the distance of turbulent layers from the telescope aperture changes for observations at different zenith angles. The distance variation can be fully compensated by a command matrix optimization, and there is no need to change the DM conjugation altitude optically during the operation of an MCAO system.

5.4. Comparing optimized and inverse control

It is of interest to see how the performance of an MCAO system depends on the type of command matrix. In Fig. 7 the Strehl ratio in the FOV is plotted for a 2-DM case (both with 66 modes) with three 11-magnitude NGSs at $60''$ radius. It was found that the selection of a threshold in computing the inverse command matrix changes the results considerably; the best threshold of 0.15 was used here (85 rejected modes). The SR for a classical AO is plotted for comparison. The gain in performance brought by the optimization is apparent. The lowering of Strehl ratio at the position of the GS with optimized control (as opposed to inverse control, where it is maximum at GS) is explained by the chosen weighting points. If the GS positions are included among the optimisation points, the best Strehl ratio is obtained at these positions, too, at the expense of slightly worse correction over the rest of the field.

6. Conclusions

The main results of this work are:

1. The theory of optimized tomography and MCAO based on Zernike modes is outlined.
2. The modal covariance code applicable to both natural and laser guide stars is presented. Its original features include fast computing of atmospheric covariances and projections for Zernike modes. The code is shown to reproduce the known analytical results on modal anisoplanatism and cone effect.
3. It is shown that for a low-order MCAO system on an 8 m telescope the tomographic error and the associated FOV limitation is related to incomplete beam overlap in upper atmospheric layers, leading to the anisoplanatism of low-order modes. On Extremely Large Telescopes the tomographic FOV will be larger, and will be eventually limited by the turbulence profile, independently of telescope diameter (Tokovinin & Viard 2001).
4. It is suggested to use several NGSs for tilt estimation in an LGS-based AO system. The sky coverage will be increased as compared to the standard 1-NGS case.
5. In an MCAO system with only two DMs the conjugation altitude of the second DM has little effect on the performance, as long as the command matrix is optimized.

Acknowledgements. This work was done while all authors were at the European Southern Observatory. This research has benefited from the support of the European Commission TMR and RTN programs (respectively, "Laser Guide Star for 8 m telescopes", contract ERBFMRXCT960094, and "Adaptive Optics for the Extremely Large Telescopes", contract HPRN-CT-2000-00147). Thanks go to the Referee, C. Dainty, for useful comments.

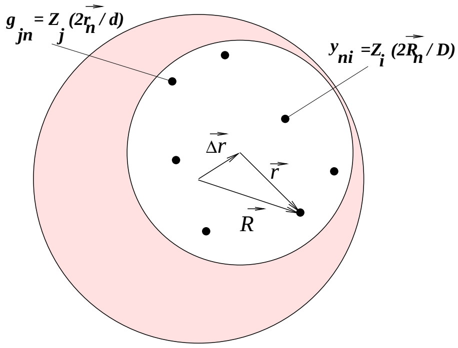

Fig. A.1. Computation of Zernike mode projections. A circular portion of some mode $Z_i(2\boldsymbol{R}/D)$ defined on a large pupil (diameter D) can be expressed as a linear combination of modes $Z_j(2\boldsymbol{r}/d)$ on a small pupil (diameter d) shifted by $\Delta\boldsymbol{r}$. The coefficients are found from the solution of a linear system for a set of grid points on the small pupil.

Appendix A: Projections of Zernike modes

It is well known that a circular portion of any wavefront $W(\boldsymbol{r})$ can be de-composed on the basis of Zernike polynomials. In our context, the "wavefront" is defined on the DM (called also *meta-pupil*) of diameter D by a set of its Zernike coefficients $\{A_j\}$. We want to represent a circular portion of this wavefront (the beam footprint) of a smaller diameter d as a sum of Zernike polynomials on this small circle with coefficients $\{a_j\}$. Let $\boldsymbol{r}$ be the coordinate of a point relative to the center of a small circle, and $\boldsymbol{R} = \boldsymbol{r} + \Delta\boldsymbol{r}$ – coordinate relative to the center of a big circle shifted by $\Delta\boldsymbol{r}$ (Fig. A.1). Then

$$W(\boldsymbol{r}) = \sum_{i=1}^{J} A_i Z_i \left(\frac{2\boldsymbol{R}}{D}\right) = \sum_{j=1}^{J} a_j Z_j \left(\frac{2\boldsymbol{r}}{d}\right), \tag{A.1}$$

where $Z_j(\boldsymbol{x})$ are the Zernike polynomials defined for $|\boldsymbol{x}| \leq 1$.

The relation between the two sets of coefficients exists (cf. RMR99) and is given by the *mode projection matrix* $\mathcal{P}$ which permits to represent a *portion* of a large-circle mode by a sum of small-circle modes:

$$Z_i \left(\frac{2\boldsymbol{R}}{D}\right) = \sum_{j=1}^{J} \mathcal{P}_{ij} Z_j \left(\frac{2\boldsymbol{r}}{d}\right). \tag{A.2}$$

Putting this into Eq. A.1 leads to

$$a_j = \sum_{i=1}^{J} \mathcal{P}_{ij} A_i \quad \text{or} \quad a = \mathcal{P}^T A. \tag{A.3}$$

As noted in RMR99, the Zernike polynomials expressed in rectangular coordinates x, y are nothing but generic monomials of the type $\sum c_{kl} x^k y^l$. The computation of c_{kl} is described in Ch. 13.5 of Malacara (1992). Although Zernike polynomials are defined only inside a unit circle (so that the small circle in Fig. A.1 must always lie within the large one), this constraint can be dropped because the monomials can be computed for arbitrary coordinates. Thus, unlike RMR99, we also consider the cases when the beam footprint spills outside the meta-pupil.

Suppose that the matrix $\mathcal{P}$ up to some radial order of Zernike polynomials Q must be found. A straightforward approach would be to select some i, to compute the mode $Z_i(2\boldsymbol{R}/D)$ on a portion of the large circle and then to find the elements of $\mathcal{P}$ as scalar products, using the fact that Zernike basis is ortho-normal:

$$\mathcal{P}_{ij} = \pi^{-1} \int_{|\boldsymbol{x}|<1} \mathrm{d}\boldsymbol{x}\; Z_i \left(\frac{d}{D}\boldsymbol{x} + \frac{2}{D}\Delta\boldsymbol{r}\right) Z_j(\boldsymbol{x}). \tag{A.4}$$

Thus, matrix $\mathcal{P}$ is constructed column by column by selecting all i-s. This approach, however, is not efficient in terms of computing time. A much better solution is obtained when we note that the coefficients of a polynomial are defined by a set of its values at a sufficient number of points. This problem is solved each time when some data are approximated by a polynomial. In our case, the "data" are the values of a truncated mode $Z_i(2\boldsymbol{R}/D)$ at a set of points (grid), the approximating polynomials are $Z_j(2\boldsymbol{r}/d)$, and we know in advance that the exact approximation exists (RMR99).

A grid of N points $\boldsymbol{x}_n$ is selected inside a unit circle, and the matrix G of the size $J = (Q+1)(Q+2)/2$ by N is computed, $g_{jn} = Z_j(\boldsymbol{x}_n)$. For a given displacement and scaling of the small circle, the grid points are transformed to new positions $\boldsymbol{X}_n = (d/D)\boldsymbol{x}_n + (2/D)\Delta\boldsymbol{r}$ and a portion of the Zernike polynomial number i is computed at these points, giving the matrix $y_{in} = Z_i(\boldsymbol{X}_n)$. The function $y(\boldsymbol{x})$ is then "approximated" on a small circle by the sum of Zernike modes $Z_j(\boldsymbol{x})$ with coefficients a_j which are, by definition, the elements of the i-th column of the projection matrix. Using Eq. A.2, we write

$$y_{in} = \sum_{j=1}^{J} \mathcal{P}_{ij} g_{jn}, \quad \text{or} \quad Y = \mathcal{P}G. \tag{A.5}$$

Now matrix G can be inverted, to yield $\mathcal{P}$ by a single multiplication:

$$\mathcal{P} = YG^{-1}. \tag{A.6}$$

The tricky point of this method is to achieve a good inversion of G, which depends entirely on the grid. We have found that better results are obtained with the number of grid points about twice the number of Zernike modes, $N = 2J$. As for the point locations, a "spiral" grid with polar coordinates $r_n = (n/N)^{0.3}$ and $\phi_n = \pi^2 N r_n$ was found to give a good condition number of the matrix G. This grid resembles a quasi-random distribution of points in a circle, with a slightly enhanced density near the border. The computation and SVD-inversion of the matrix G need to be done just once for a given problem (given Q). Then, for varying geometrical pupil transformations, only the vector Y has to be re-computed. This method reduces the computation time significantly with respect to the direct computation of scalar products.

The behavior of matrix $\mathcal{P}$ is noteworthy. All its elements have absolute values less than one if a small pupil is cut from a larger one (interpolation), but may take high absolute values (with both signs) in the opposite case (extrapolation to a large pupil). All elements above the diagonal are always zero: any given mode projects to a set of modes up to the same number (and not only up to the same radial order, as stated in RMR99). For small displacements and scaling, $\mathcal{P}$ has diagonal elements close to 1. Finally, for a large ratio of pupil diameters the elements near diagonal are small, and only the first columns of $\mathcal{P}$ are important. It means that a small part of any i-th mode is well represented by a sum of lower-order modes; the wavefront decomposition in a small part of a pupil is thus effectively of lower order than on the whole pupil.

Knowing $\mathcal{P}$, it is easy to compute the interaction matrix A which enters into the Eq. 1. Both commands and signals are numbered sequentially, grouping the Zernike coefficients in single vectors (Fig. 1). So, each element A_{lk} corresponds to the j-th mode measured on a particular GS when an i-th mode of unit amplitude is applied to one of the DMs. It is hence equal to $\mathcal{P}_{ij}$, where $\mathcal{P}$ is calculated for a particular GS-DM combination and takes into account the shift of the GS beam and the ratio of its diameter to the diameter of DM. For the first DM conjugated to pupil there are neither shifts nor scaling, and $\mathcal{P}$ is an identity matrix.

Appendix B: Modal covariance theory

As noted above, we consider a modal MCAO system where the WFS signals and the DM commands are specified as Zernike coefficients. The number of modes measured on each of guide stars may be different. Let the total number of measurements be L. The WFS data are then an L-element vector s. The number of controlled parameters (commands) I is equal to the total number of controlled modes on all DMs. The command c is an I-element vector of the Zernike coefficients corresponding to DMs. Our purpose is to find the *command matrix* M that would give the best imaging quality for a linear control, i.e.

$$c = Ms. \tag{B.1}$$

Matrix M has the size IxL.

The DMs are conjugated to different heights. Although their shape does not depend on the viewing direction $\boldsymbol{\theta}$, the combined effect of the corrections does. Following Ellerbroek (1994), the *influence functions* $r_i(\boldsymbol{x},\boldsymbol{\theta})$ are introduced which depend on $\boldsymbol{\theta}$ and describe the wavefront shape on the telescope pupil when a unit command c_i (i.e. a Zernike mode) is applied to one DM (not to be confused with the influence functions of DM actuators in classical AO). For a given command vector c the wavefront correction is hence equal to the sum of $c_i r_i(\boldsymbol{x},\boldsymbol{\theta})$ over i. Using mode projections, an influence function can be expressed through the *system projection matrix* P:

$$r_i(\boldsymbol{x},\boldsymbol{\theta}) = \sum_{j=2}^{J} P_{ij}(\boldsymbol{\theta}) Z_j\left(\frac{2\boldsymbol{x}}{D}\right), \tag{B.2}$$

where Z_j are the Zernike polynomials, J is the number of the highest mode considered in the whole problem. The system projection matrix P has the size Ix$(J-1)$. It is constructed in the following way. For each of DMs, the mode projectiom matrix $\mathcal{P}$ is computed and its lines corresponding to the modes controlled on this particular DM (e.g. excluding piston, tip, tilt) are included into the corresponding sub-section of P. Strictly speaking, the sum in (B.2) must start with $j=1$ (piston mode), but this term is irrelevant for imaging.

The phase residual ϵ (after correction, piston-removed) can be written as

$$\epsilon(\boldsymbol{x},\boldsymbol{\theta}) = \sum_{j=2}^{\infty}\left[a_j(\boldsymbol{\theta}) - \sum_{i=1}^{I} c_i P_{ij}(\boldsymbol{\theta})\right] Z_j\left(\frac{2\boldsymbol{x}}{D}\right). \tag{B.3}$$

Here $a_j(\boldsymbol{\theta})$ are the Zernike coefficients of the wavefront coming from the direction $\boldsymbol{\theta}$. Using the ortho-normality of Zernike modes on a circular telescope pupil, the aperture-averaged and ensemble-averaged r.m.s residual $\langle\epsilon^2\rangle$ is written as a sum of the squared coefficients, leading to $\langle\epsilon^2\rangle = \epsilon^2_{\rm fit} + \epsilon^2_J$, where $\epsilon^2_{\rm fit}$ is the variance of uncorrected high-order modes (fitting error) and ϵ^2_J is the variance of the first J modes:

$$\begin{aligned}\epsilon^2_{\rm fit} &= \sum_{j=J+1}^{\infty}\langle a_j^2\rangle, \qquad\qquad (B.4)\\ \epsilon^2_J &= \sum_{j=2}^{J}\left[\langle a_j^2\rangle - 2\sum_{i=1}^{I} P_{ij}\langle a_j c_i\rangle + \sum_{i=1}^{I}\sum_{i'=1}^{I} P_{ij}P_{i'j}\langle c_i c_{i'}\rangle\right].\end{aligned}$$

Our aim is to express the residual through the covariances of Zernike coefficients and WFS data. This is achieved by putting equation (B.1) for the correction signals (commands) c_i into (B.4). Resulting quantities, apart from M, are described by three matrices.

The data covariance matrix S (independent of $\boldsymbol{\theta}$, symmetric, size LxL):

$$S_{ll'} = \langle s_l s_{l'}\rangle. \tag{B.5}$$

Matrix S is computed as a sum of the covariances of atmospheric Zernike coefficients (see Appendix C) and noise covariances. These latter are estimated for the case of a Shack-Hartmann sensor by the method of Rigaut & Gendron (1992).

The cross-talk matrix R ($\boldsymbol{\theta}$-dependent, symmetric, size IxI) is defined as a product of influence functions integrated over the pupil:

$$R_{ii'}(\boldsymbol{\theta}) = \int_{\rm pupil} {\rm d}\boldsymbol{x}\;\; r_i(\boldsymbol{x},\boldsymbol{\theta}) r_{i'}(\boldsymbol{x},\boldsymbol{\theta}) = PP^T. \tag{B.6}$$

This relation (B.6) between R and P follows from the ortho-normality of Zernike modes.

The target matrix T ($\boldsymbol{\theta}$-dependent, size IxL):

$$T_{il} = \sum_{j=2}^{J} P_{ij} \langle a_j s_l \rangle. \tag{B.7}$$

With these definitions, the r.m.s. residual of the first J modes is

$$\begin{aligned}\langle \epsilon_J^2 \rangle &= \epsilon_{0J}^2 - 2\sum_{i=1}^{I}\sum_{l=1}^{L} T_{il} M_{il} \\ &+ \sum_{i=1}^{I}\sum_{i'=1}^{I}\sum_{l=1}^{L}\sum_{l'=1}^{L} R_{ii'} M_{il} M_{i'l'} S_{ll'} \\ &= \epsilon_{0J}^2 - \mathrm{Trace}(2MT^T - RMSM^T),\end{aligned} \tag{B.8}$$

ϵ_{0J}^2 being the uncorrected variance of these modes.

This r.m.s. residual still depends on $\boldsymbol{\theta}$, so if we use Eq. B.8 to optimize the control, the resulting command matrix will also depend on $\boldsymbol{\theta}$. Moreover, with several DMs the cross-talk matrix R is singular, because for any given viewing direction the same correction can be obtained by different combinations of DM signals. The standard approach consists in averaging the residual the FOV (over $\boldsymbol{\theta}$) with an arbitrary weighting function W_θ which describes the optimization goal. The function W_θ is defined to have a unit integral over $\boldsymbol{\theta}$. Denoting the θ-averaged quantities by bar, we have, for example,

$$\overline{\langle \epsilon_J^2 \rangle} = \int W_\theta(\boldsymbol{\theta}) \langle \epsilon_J^2 \rangle \mathrm{d}\boldsymbol{\theta}. \tag{B.9}$$

So, we must replace T and R in Eq. B.8 with $\bar{T}$ and $\bar{R}$ to obtain the final expression for $\overline{\langle \epsilon_J^2 \rangle}$.

The optimized (in the minimum variance sense) control matrix is found by differentiating Eq. B.8 over M_{il} and solving the linear system of equations. Referring to Ellerbroek (1994), this would be called unconstrained optimum control. The result is Eq. 3.

This solves the problem. In order to compute the optimal control law, we must know the projection operators P for all modes of all DMs (they depend only on the geometry of the MCAO system) and the data covariances $S_{ll'}$ for both natural and laser guide stars. The covariances involving the observed object which are needed for T do not require a separate technique for their estimation. Matrix S depends on the turbulence profile and on the WFS noise, as well as on the configuration and number of guide stars, but it is independent of $\boldsymbol{\theta}$ and of the DM geometry.

In deriving Eq. 3, it is supposed that the "weak" modes corresponding to the small eigenvalues of $\bar{R}$ are removed from the control parameters (hence pseudo-inverse instead of inverse). We do not provide here more detailed discussion of weak mode removal, which differs Eq. 3 from the classical result of Wallner (1983).

Appendix C: Computation of Zernike mode covariances

Formulas for the computation of covariances of Zernike coefficients for a natural guide star, a laser guide star and a mixed case were given by Molodij & Rousset (1997). However, they considered only covariances of the polynomials of the same number. More general formulae for the covariances between *different* modes were later derived by Whiteley et al. (1998). Basically, for each layer the covariances are expressed through a one-dimensional integral I involving triple products of Bessel functions $J_n(x)$ combined with powers:

$$I = \int_0^\infty \mathrm{d}x \; B_1(x) B_2(a_1 x, n_1) B_2(a_2 x, n_2) B_3(s_l x, m), \tag{C.1}$$

where

$$\begin{aligned} B_1(x) &= x^{-1} \, (x^2 + x_0^2)^{-11/6}, \\ B_2(ax, n) &= a^{-1} J_{n+1}(ax), \\ B_3(s_l x, m) &= J_m(s_l x). \end{aligned} \tag{C.2}$$

The dimensionless integration variable x is actually a product of the spatial frequency f and aperture radius, $x = fD/2$. The first term describes the turbulence spatial spectrum (the parameter $x_0 = \pi D / L_0$ takes into account the turbulence outer scale L_0, a non-negligible correction for 8-m class telescopes). The two second terms depend on the radial orders n_1, n_2 of the corresponding Zernike modes and also depend on the layer altitude h through the factors $a_i = 1 - h/H_i$, where H_i are the guide star altitudes. Only the third term depends on the angular distance between guide stars α which enters into the dimensionless beam separation $s_l = 2\alpha h/D$. Finally, m is a combination of the azimuthal orders, $m = m_1 + m_2$ or $m = |m_1 - m_2|$.

Several tricks are used to speed up the covariance computation, which is the most computer-intensive part of the code. First, instead of evaluating I for each turbulent layer and computing a weighted average, we change the order of summation and integration, computing the weighted average of the integrand and then integrating it only once. Secondly, we notice that the integrand is a smoothly oscillating function of x which goes to zero both at $x = 0$ and at $x > 50$ (for radial orders not exceeding 20). So, precise enough integrals are estimated with the Newton-Rapson method from the 200 values of the integrand tabulated with a step of $\Delta x = 0.25$. Both B_1 and B_2 are pre-calculated (the latter – for all orders n and all layers, as a 3-dimensional array). Pre-calculation of B_2 is done for a single value of the LGS altitude H, as encountered in practical problems, and also for the NGS ($a = 1$). Then, for a given set of angles α the B_3 only need to be evaluated. It is multiplied by the two appropriately selected B_2-s and by B_1, averaged over layers and integrated.

Further reduction of computing time is possible for symmetric guide star configurations. In this case the number of distinct angular separations α_i can be small, even though α_i must be counted separately for NGS-NGS, NGS-LGS, and LGS-LGS pairs. For each of these 3 combinations, the I is pre-computed for all α_i and all possible indices n_1, n_2 and m (m goes from 0 to $n_1 + n_2$). For a given radial order Q the number of such combinations ($\propto Q^3$) is less than the number of pairwise mode combinations ($\propto Q^4$), thus leading to faster covariance calculation.

When a procedure of covariance computing is called, it checks first whether the angle α is found among the table of pre-computed integrals (otherwise, the integrals are done directly). Then it extracts the two values of I corresponding to two m-s and combines them with appropriate factors to get the covariance.

References

Chassat, F. 1989, J. Optics (Paris), 20, 13
Chassat, F. 1992, PhD Thesis, Univ. Paris XI, Orsay
Chun, M. 1998, PASP, 110, 317
Ellerbroek, B.L. 1994, J. Opt. Soc. Amer. (A), A11, 924
Ellerbroek, B. L. & Rigaut, F. 2000, in Adaptive Optical Systems Technology, ed. P. L. Wizinovich, Proc. SPIE, 4007, 1088
Fried, D. L. 1982, J. Opt. Soc. Amer., 72, 52
Fusco, T., Conan, J.-M., Michau, V. et al. 1999, in Propagation through the Atmosphere III, ed. M. C. Roggemann & L. R. Bissonnette, Proc. SPIE, 3763, 125
Fusco, T., Conan, J.-M., Michau, V. et al. 2000, in Adaptive Optical Systems Technology, ed. P. L. Wizinovich, Proc. SPIE, 4007, 1032
Fusco, T., Conan, J.-M., Rousset, G. et al. J. Opt. Soc. Amer. (A), 2001 (submitted)
Johnston, D. C. & Welsh, B. M. 1994, J. Opt. Soc. Amer. (A), A11, 394
Le Louarn, M. & Tallon, M. 2000, in Adaptive Optical Systems Technology, ed. P. L. Wizinovich, Proc. SPIE, 4007, 1066
Malacara, D. 1992, Optical Shop Testing (Wiley, Ney York), Ch. 13.5.
Martin, F., Conan, R., Tokovinin, A. et al. 2000, A&AS, 144, 39
Molodij, G. & Rousset, G. 1997, J. Opt. Soc. Amer. (A), A14, 1949
Press, W.H., Teukolsky, S.A., Vetterling, W.T., & Flannery B.P. 1992, Numerical Recipes in C (Cambridge Univ. Press, Cambridge), Ch. 2
Ragazzoni, R., Marchetti, E., & Rigaut, F. 1999, A&A, 342, L53 (RMR99)
Rigaut, G. & Gendron, E. 1992, A&A, 261, 677
Roddier, F., ed. 1999, Adaptive optics in astronomy (Cambridge Univ. Press, Cambridge)
Sasiela, R. J. 1994, Electromagnetic wave propagation in turbulence, (Springer-Verlag, Berlin), p. 236 Eq. 12.7.
Stone, J., Hu, P.H., Mills, S.P., & Ma, S. 1994, J. Opt. Soc. Amer. (A), A11, 347
Tallon, M. & Foy, R., 1990, A&A, 235, 549
Tokovinin, A. & Viard, E. 2001, J. Opt. Soc. Amer. (A), A18, 873
Tyler, G. A. 1994, J. Opt. Soc. Amer. (A), A11, 325
Wallner, E. 1983, J. Opt. Soc. Amer., 73, 1771
Whiteley, M. R., Roggemann, M. C., & Welsh, B. M. 1998, J. Opt. Soc. Amer. (A), A15, 993