

Comparison of multiconjugate adaptive optics configurations and control algorithms for the Gemini-South 8-m telescope

Ralf Flicker^{ab}, François Rigaut^b and Brent Ellerbroek^b

^aLund Observatory, Box 43, SE-22100 Lund, Sweden

^bGemini Observatory, 670 N. A'ohoku Place, Hilo HI-96720, USA

ABSTRACT

This paper focuses on two main categories of the multiconjugate adaptive optics (MCAO) parameter space for performance optimization: the geometrical configuration of guide stars and deformable mirrors (DMs), and the wavefront reconstructors. From the first category it is shown how, for a fixed reconstructor and imaging wavelength, the performance metrics with a few important exceptions improve with an increasing number of i) DMs, ii) actuators per DM and iii) guide stars. The metrics are seen to degrade with i) an increasing field of compensation and ii) DM conjugation altitude mismatch with the significant turbulent atmospheric layers. In the second category, this study also compares the performance with a fixed MCAO configuration using the least-square estimator (LSE) and the maximum *a posteriori* estimator (MAP) for wavefront reconstruction. The MAP is shown to perform significantly better than the LSE at low or intermediate signal-to-noise ratios (SNRs), and somewhat better even in the absence of noise due to its *a priori* knowledge of the phase statistics.

Keywords: Multiconjugate adaptive optics, control algorithms, wavefront reconstruction

1. INTRODUCTION

First proposed in 1988,¹ MCAO is today emerging as one of the most promising concepts for field-widening of adaptive optics compensation, which up until now has been able to provide a high degree of correction only within a small angular field about the size of the isoplanatic patch. It is also a key technology for future generations of extremely large telescopes, which will not be feasible without MCAO.² Previous studies have demonstrated the potential of MCAO,³⁻⁷ and in this paper a number of specific configurations and control algorithms are investigated in an effort to understand the fundamental behaviour of MCAO systems and how to optimize their performance.

An MCAO system employs an ensemble of guide stars, which allow for three-dimensional tomography of the atmospheric turbulence, and a number of altitude-conjugated DMs to widen the compensated field of view by correcting the phase distortions *in situ*. The commands applied to the DMs are calculated by a wavefront reconstruction algorithm that is an optimal estimator designed to produce commands that are optimal in, for instance, a least-square or a minimum variance sense. MCAO for 8-m class telescopes must by necessity be supported by laser guide stars (LGS), as, were only natural guide stars (NGS) to be used for probing the atmospheric turbulence over a several arc minute field, the sky coverage would be too poor to justify an MCAO instrument.⁸⁻¹⁰ Due to the inability of LGS to determine the global tip and tilt over the aperture though, a few NGS must still be used for tip/tilt sensing. Since these global modes may be controlled at a lower bandwidth and because the whole telescope aperture is used for a light-bucket, however, the requirements on NGS magnitudes are relaxed by a substantial factor to allow a fair fraction of the sky to be accessible to MCAO.

2. WAVEFRONT RECONSTRUCTORS

The wavefront reconstructor E is the brain of the AO system, estimating the optimal DM commands \mathbf{c} from wavefront sensor (WFS) data \mathbf{s} according to some *a priori* optimization criterion to give $\mathbf{c} = E\mathbf{s}$. Two reconstructors are considered in this study: the LSE and the MAP.

Send correspondence to (E-mail):

Ralf: ralf@astro.lu.se, François: frigaut@gemini.edu, Brent: bellerbroek@gemini.edu

2.1. Least square estimator

In an AO system with an ideal noise-free Shack-Hartmann WFS we assume that there is a linear relationship between the n actuator commands \mathbf{c} and the m slope measurements \mathbf{s} according to

$$\mathbf{s} = G\mathbf{c}, \quad (1)$$

where G is the $m \times n$ *interaction matrix*. To estimate the actuator commands one needs to invert the relationship (1), either by the generalized matrix inverse (*aka* the Moore-Penrose *pseudoinverse*) or by singular value decomposition (SVD). Depending upon the DM/WFS geometry there will be a number of highly singular modes and at the very least two - the piston and waffle modes - that must be filtered out before inversion. Hence the SVD is the preferred method as it explicitly delivers the singular values and subject them to filtering according to some *a priori* threshold. Decomposing G with the SVD, inverting and filtering, involve these three steps:

$$G = U\Lambda V^T \quad (2)$$

$$G^{-1} = V\Lambda^{-1}U^T \quad (3)$$

$$E_{ls} = V\Gamma U^T, \quad (4)$$

where

$$\Lambda = \text{diag}(\{\lambda_i\}_{i=1}^n), \text{ the singular values} \quad (5)$$

$$\Gamma_{ii} = \begin{cases} \lambda_i^{-1} & \text{for valid modes} \\ 0 & \text{for singular modes} \end{cases} \quad (6)$$

are diagonal matrices and U and V are orthonormal matrices. The notation E_{ls} is introduced for the least-square *reconstruction matrix*, or just *reconstructor*, which in this case is just the filtered inverse of G . The least-square optimized actuator commands are thus given by the reconstructed slope measurements according to $\mathbf{c} = E_{ls}\mathbf{s}$. This solution is a least-square estimate in the sense that the squared norm of the measurement error $\varepsilon_s \equiv \|\mathbf{s} - G\mathbf{c}\|^2$ is minimized.

2.2. Maximum a posteriori estimator

Although the ideal LSE performs rather well under a relatively wide range of conditions, it is worthwhile to consider a more realistic case where the WFS measurements are distorted by a purely additive random noise \mathbf{n} such that

$$\mathbf{s} = G\mathbf{c} + \mathbf{n}. \quad (7)$$

In the case of Gaussian statistics for the noise, a noise-weighted least-square optimal estimate (*aka* the Gauss-Markov estimate, GME) of the actuator commands may be derived from a probabilistic approach using the *maximum likelihood* technique. The GME solution is flawed, however, in that another singular inversion presents itself so that G must be pre-filtered anyhow. This solution may be regularized, however, by employing a *maximum a posteriori*¹¹ technique in stead, and as this result contains features attractive to MCAO applications this study will leave the GME and focus on the MAP. Employing the Bayes' formula for conditional probabilities, one may inquire about the probability that a certain set of actuator commands \mathbf{c} produced what was actually measured \mathbf{s} , and attempt to maximize this quantity:

$$P(\mathbf{c}|\mathbf{s}) = \frac{P(\mathbf{c})P(\mathbf{s}|\mathbf{c})}{P(\mathbf{s})}. \quad (8)$$

Assuming Gaussian statistics for the WFS noise and for the phase distortion at any point in the field, the various terms in (8) are readily derived as χ^2 expressions. The equation may then be differentiated with respect to \mathbf{c} and solved for the optimal MAP estimator^{6,12}:

$$E_{map} = (G^T C_n^{-1} G + C_c^{-1})^{-1} G^T C_n^{-1}. \quad (9)$$

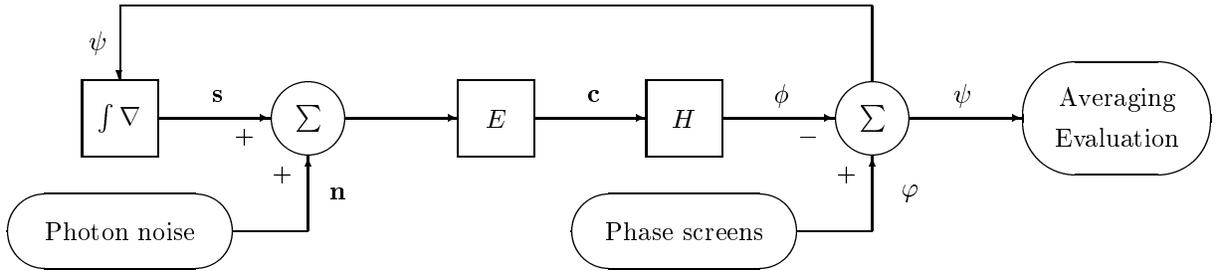
The terms $C_n \equiv \langle \mathbf{n}\mathbf{n}^T \rangle$ and $C_c \equiv \langle \mathbf{c}\mathbf{c}^T \rangle$ are the closed loop covariance matrices of the noise and the actuator command vectors, and it may be shown that the MAP estimate is also optimal in the least-square sense. For low SNRs, the regularization term C_c^{-1} will dominate the expression within parentheses and keep it invertible. What this means, generally speaking, is that when there is a lot of noise in the system, the reconstructor will trust the

measurements less and less and use the *a priori* knowledge of the atmospheric statistics in stead for reconstructing the phase. For high SNR the GME term will dominate, and the MAP solution converges to the GME in the limit of no noise.

For MCAO, the inclusion of phase statistics has another consequence altogether. Although perhaps not the most severe, one of the fundamental limitations to MCAO relying on tomography is still the fact that some portions of the atmosphere (where there is no measurement redundancy) will not be mapped tomographically.¹³ In the outer regions of the field where there is no overlap between the guide star beams, the reconstructor will not be able to position the measured phase distortions vertically, and tomography fails. With some knowledge of the phase statistics though, be it only a generic estimate, the reconstructor may use this information to distribute the correction among the DMs in a fashion that on the average will be more right than wrong, and long-exposure performance will improve.

3. MCAO SIMULATION

The MCAO configurations and reconstructors have been tested in a Monte Carlo simulation that closely models the spatial characteristics such as fitting, anisoplanatism and spatial aliasing errors of a closed loop Shack-Hartmann (SH) based MCAO system. Temporal issues such as servo lag effects and the temporally augmented WFS noise are not modeled, hence the loop may be closed with zero frame delay. The WFS slope vectors are calculated from the average phase gradient over the subapertures, and WFS noise is modeled only with respect to photon shot noise. This means that no account is taken for the effects of subaperture spot deformation due to extended beacons or partially illuminated subapertures. More detailed simulations including these effects^{5,10} give similar results, however, which suggests that this simplified model gives a good first-order estimation of the performance. The algorithm is sketched in block-diagram form below, where the phase maps ϕ , φ and ψ are field dependent quantities and \mathbf{s} and \mathbf{c} refer to multiple WFSs and DMs respectively (*cf.* §3.1).



3.1. Modeling of the system

The simulation computes the pupil plane residual phase distortion upon adaptive correction for atmospheric turbulence, adopting the near-field approximation and neglecting scintillation effects. Taking \mathbf{r} as the pupil plane coordinate and θ as the all-sky field angle, the residual phase ψ may be written:

$$\psi(\mathbf{r}, \theta) = \sum_{l=1}^7 \varphi_l(\gamma \mathbf{r} - h_l \theta) - \sum_{k=1}^{N_{mir}} \phi_k(\gamma \mathbf{r} - h_k \theta), \quad (10)$$

where φ_l is the turbulent phase of layer l at altitude h_l , and ϕ_k is the mirror deformation of DM k conjugated to an altitude h_k . The scale factor γ accounts for the LGS cone effect when (10) is calculated in the WFS measurement algorithm, and it assumes the value

$$\gamma = 1 - \frac{h}{h_{lgs}} \delta(\alpha_{lgs}), \quad \delta(\alpha_{lgs}) = \begin{cases} 1 & \theta = \alpha_{lgs} \\ 0 & \text{otherwise} \end{cases}, \quad (11)$$

where h is the altitude, h_{lgs} is the LGS altitude and α_{lgs} is the vector of LGS directions. The SH-WFS measurement vector is proportional to the subaperture spot centroid, which is calculated from the average gradient of the phase,

$$\mathbf{s}(\theta) \propto \int d\mathbf{r} \nabla \tilde{\psi}(\mathbf{r}, \theta) \xrightarrow{\theta \rightarrow \alpha_i} \mathbf{s}_i \propto \int d\mathbf{r} \nabla \tilde{\psi}(\mathbf{r}, \alpha_i), \quad (12)$$

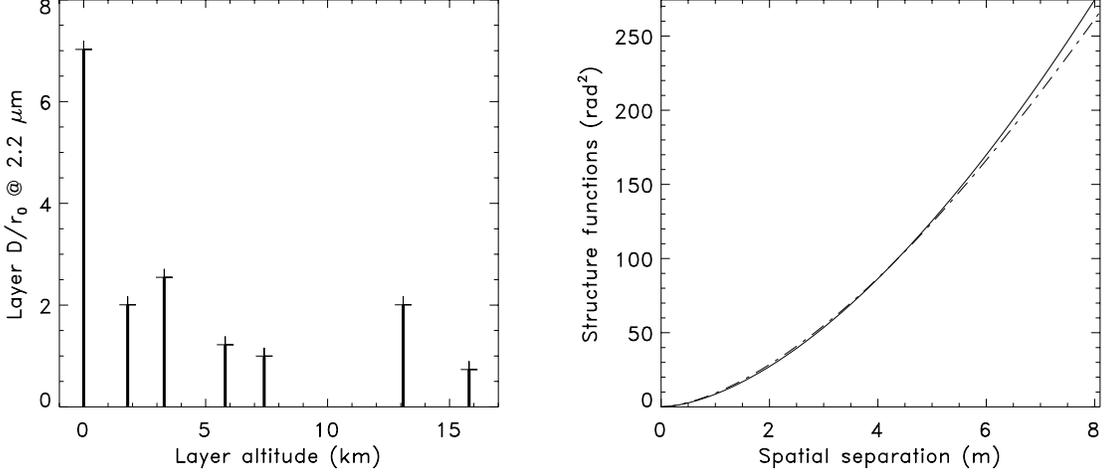


Figure 1. Left panel: layered 50% Cerro Pachon turbulence profile for an 8-m telescope in K band at a 2.7 km elevation, expressed in D/r_0 . Right panel: structure functions of a theoretical Kolmogorov spectrum (solid) and the phase screens used in the simulation.

where the directions for measurements have been restricted to the i guide star directions α_i . In $\tilde{\psi}$ the piston, and for LGS also the tip/tilt components, have been projected out according to

$$\tilde{\psi}(\mathbf{r}, \alpha_i) = \psi(\mathbf{r}, \alpha_i) - \frac{1}{P} \sum_j m_j(\mathbf{r}) \int_P d\mathbf{r}' m_j(\mathbf{r}') \psi(\mathbf{r}', \alpha_i), \quad (13)$$

where $m_j(\mathbf{r})$ are the modes to be removed and P the pupil area. The interaction matrix G is the Jacobian of WFS measurements to DM commands, $G = \partial \mathbf{s} / \partial \mathbf{c}$, which is obtained during the initialization of the system by poking one actuator at a time and measuring the response. The mode matrix H is likewise the Jacobian of mirror deformations to actuator commands, $H = \partial \phi / \partial \mathbf{c}$, with the DM influence functions or any other desirable modes for its columns. Ignoring temporal complications, the process of measurement, reconstruction and building the k mirror deformations can be written

$$\phi_k = H_k \sum_{i=1}^{N_{gs}} E_i \int d\mathbf{r} \nabla \psi(\mathbf{r}, \alpha_i), \quad (14)$$

where N_{gs} is the number of guide stars. The DMs are modeled with the characteristic influence functions of a piezostack DM, where the pupil plane DM is configured in a Fried geometry with respect to the WFS subapertures. The LGS have been fixed to an X-like configuration consisting of five sodium LGS at 90 km, and the four tip/tilt NGS are distributed in-between (*cf.* top left panel of figure 2) in a fashion that fills the field adequately. For global tip/tilt sensing only three NGS are required, but we have used four here merely for the benefit of symmetry. This allows for a faster evaluation, and there is only a marginal increase in performance in going from three to four NGS as long as noise sensitivity is not concerned. The photon shot noise on the beacons was modeled by adding a random vector of Gaussian statistics with standard deviation σ_m to the SH-WFS slope vector, expressed in radians of average phase difference between the subaperture edges. The subaperture variance σ_m^2 was given by⁹

$$\sigma_m^2 = \left(\frac{\lambda_s}{\lambda_{im}} \right)^2 \frac{2\pi}{N_{pde} r_{0_s}^2}, \quad (15)$$

where N_{pde} is the number of photo detection events (PDEs) in the subaperture detector, λ_s and λ_{im} the sensing and imaging wavelengths respectively and $r_{0_s}^2$ the integrated Fried parameter at the sensing wavelength. This is not the best formula for LGS since it neglects effects of LGS spot size. As only LGS are considered for high-order beacons in this study, however, this approximation will result in a systematic error affecting all simulations equally.

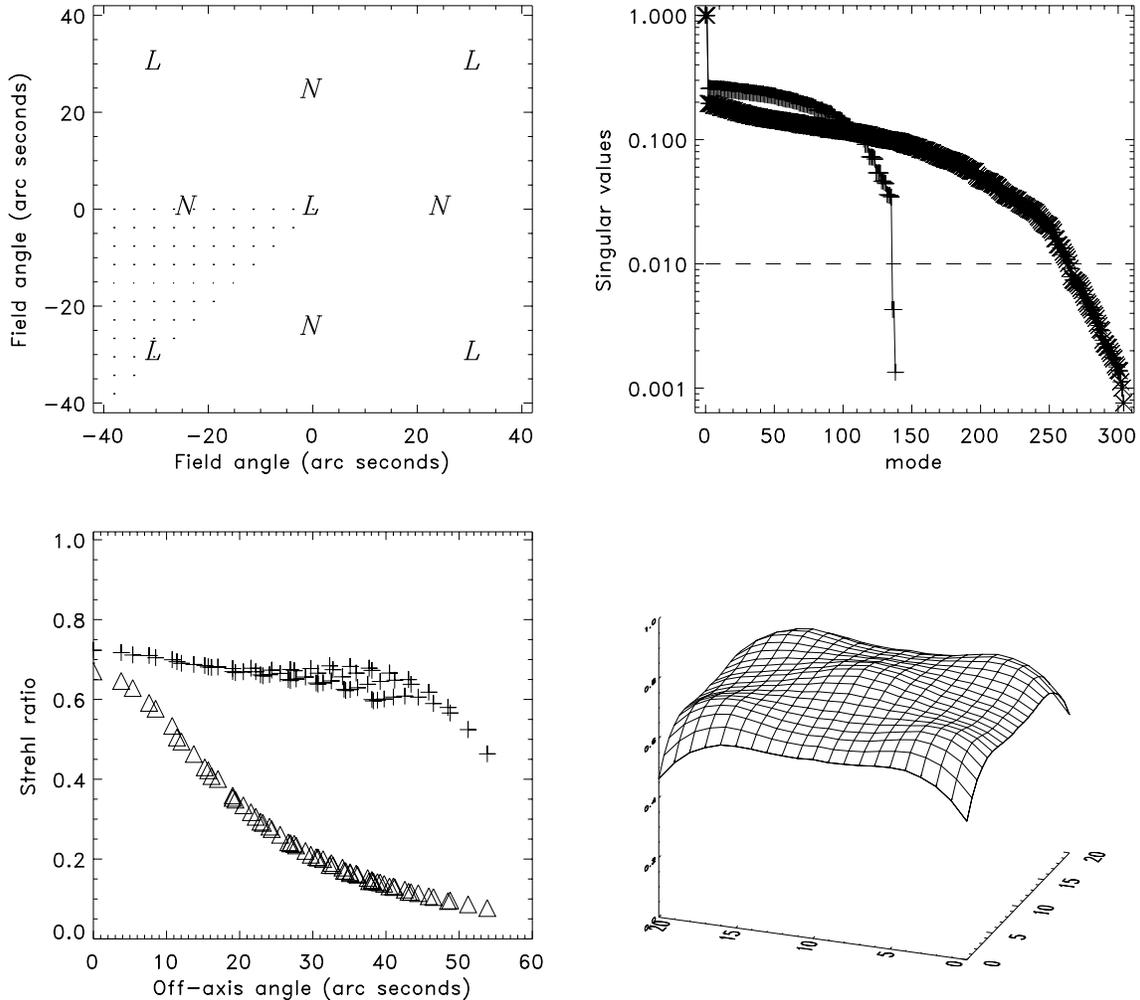


Figure 2. Configuration and data for the generic 3-DM (12×12 WFS subapertures) system described in section 4. Upper left: configuration of natural (N) and laser (L) guide stars, and field evaluation points (dots). Upper right: normalized singular values of the conventional (plus) and multiconjugate (star) AO system. Lower left: Strehl ratio versus field angle for AO/MCAO systems. Lower right: surface plot of the MCAO Strehl ratio.

3.2. Modeling of the atmosphere

The atmospheric turbulence is modeled by a C_n^2 profile obtained from generalized-SCIDAR and balloon data at Cerro Pachon¹⁴ (CP), where the Gemini-South 8-m telescope is being constructed. Adopting the Taylor hypothesis of frozen turbulence, the 50% CP data was compiled and discretized into a 7-layer profile by fitting Gaussians to the real data and assigning a delta function containing the total energy of the Gaussian to their respective offsets. Figure 1 (left) shows the turbulence profile thus obtained, expressed in D/r_0 for an 8-m telescope in K band. The integrated Fried parameter in V band ($0.55 \mu\text{m}$) is $r_0 = 16.6$ cm, which is equivalent to an overall seeing of $\theta_s \approx 0.7''$.

The phase screens used for simulating the turbulent layers were generated from the theoretical Kolmogorov power spectral density (PSD) by the FFT method. The phase map is obtained by adding a uniformly distributed random phase ξ to the square-root of the Kolmogorov PSD Φ_{kol} before applying the Fourier transform: $\varphi = \mathcal{F}^{-1} [\sqrt{\Phi_{kol}} e^{-i\xi}]$. Figure 1 (right) compares the structure functions

$$D_{kol}(\rho) = 6.88 \left(\frac{\rho}{r_0} \right)^{5/3} \quad \text{and} \quad D_{mod}(\rho) = \langle |\varphi(r) - \varphi(r + \rho)|^2 \rangle \quad (16)$$

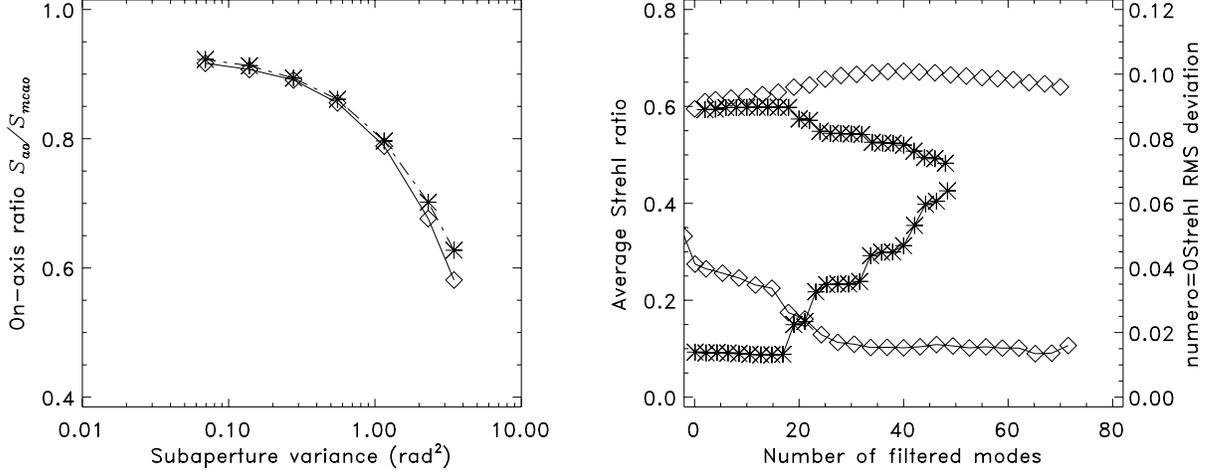


Figure 3. Left: on-axis ratio of AO and MCAO Strehl ratio for LSE (solid) and the MAP (dot-dash) estimator. Right: \bar{S} (symbols) and σ_S (symbols+line) for the generic 3-DM system (diamonds) and the 2-DM system used in §4.1.2 conjugated to [0,8] km (stars), as a function of the number of modes discarded after filtering.

of the theoretical Kolmogorov and the computer generated phase screens. At small separations ρ , the discrepancy is negligible to none, and while the difference grows at larger separations this is not a big concern, as it actually bends off in the right direction. A more realistic PSD (*e.g.* the von Karman spectrum) would show a similar behavior due to the finite outer scale of the turbulence. Due to the periodicity of the Fourier transform, however, the FFT method imposes an outer scale to phase screens generated even from Kolmogorov statistics, but it is typically a much larger scale than would be assumed for a von Karman spectrum. Since the screens were generated in a 2048×2048 format, this outer scale would for an 8-m pupil D be $L_0 = 16384/D_{pix}$ m, where D_{pix} is the number of pixels across D . For the values of D_{pix} that concern the simulations in this paper we find that $L_0 \sim 250 - 500$ m.

4. NUMERICAL RESULTS

The MCAO system compensation quality has been evaluated in a triangular area one eighth of the full field, *cf.* figure 2 (top left panel), with the square area within the LGS chosen as the field relevant to optimization. The main metrics chosen for evaluating the performance are the average \bar{S} and the standard deviation σ_S (or sometimes the relative standard deviation $\delta_S \equiv \sigma_S/\bar{S}$, *aka* uniformity) of the long exposure Strehl ratio $\langle S(\theta) \rangle$ within this field. $S(\theta)$ was estimated from the residual phase variance by the approximation $S(\theta) = \exp[-\sigma_\psi^2(\theta)]$, where

$$\sigma_\psi^2(\theta) = \frac{1}{P} \left[\int_P d\mathbf{r} \psi^2(\mathbf{r}, \theta) - \left(\int_P d\mathbf{r} \psi(\mathbf{r}, \theta) \right)^2 \right], \quad (17)$$

bearing in mind that the exponential approximation is an underestimation below $S \sim 0.2$. Figure 2 shows an example of a typical data output from one simulation, here for a 3-DM system with a 12×12 subaperture SH-WFS and [13, 11, 9] actuators across the diameter of the three DMs conjugated to [0, 4, 12] km. The performance metrics for this system (henceforth referred to as the “generic” system) in K band with no noise are $\bar{S} = 0.67$ and $\delta_S = 2.6\%$. When citing number of actuators across the pupil, the relevant measure in reality is the actuator pitch d , which defines the spatial cut-off frequency $f_c = 1/2d$ of the DM. The pitch is given by $d = W/(N - 1)$, where W is the meta-pupil diameter and N the number of actuators across W . W also depends on the conjugation altitude h , the LGS altitude h_{lgs} and the angle α to the LGS according to

$$W = D \left(1 - \frac{h}{h_{lgs}} \right) + 2\alpha h \quad \Rightarrow \quad d = \frac{1}{N - 1} \left[D \left(1 - \frac{1}{h_{lgs}} \right) + 2\alpha h \right]. \quad (18)$$

This parameter (d) goes into the estimation of the fitting error, and it is seen that d mainly depends upon N , as α and h will be correlated to N .

4.1. Modal filtering

One major hassle that complicates the analysis of least-square type reconstructors is the subjectivity in the choice of conditioning number. Strictly, to obtain the optimal performance for a given configuration, one should test it for *all* choices of conditioning numbers and choose the optimal. Figure 3 (right panel) shows \bar{S} and σ_S for the generic 3-DM system (solid curve) and a 2-DM system of the same order (dot-dashed curve). There is clearly some dependence upon conditioning number, and moreover, this dependence varies between configurations. It seems thus very difficult to know beforehand what a good conditioning number will be for a specific system, and only checking them all would reveal. Unfortunately such an optimization was not possible for the simulations in this paper, and an empirically “sound” conditioning number was chosen for each configuration. This may be expected to produce some spurious effects occasionally, though we believe this to be a second order effect in most (but not all) cases considered here. The fundamental reason for the inadequacy of this naive absolute filtering criterion is that the WFS sensitivity to a specific mode need not necessarily correlate with the atmospheric content in that mode. That is, some modes which are poorly sensed by the WFS - and thus likely to be filtered - may still contain a non-negligible amount of atmospheric turbulence statistics, and vice versa. A more careful filtering might alleviate this effect by projecting the SVD modes on Zernikes to get a measure of the atmospheric content in the mode, but this was not done in the simulations presented here. Another effect, particular to MCAO, is that locally well sensed modes on the individual DMs may combine to form globally singular modes as seen from the WFSs (the DM “null modes”, *cf.* reference [12]). This will result in a disproportional number of modes being filtered out as the number and order of the DMs grow, but it is unclear whether this is liable to cause trouble for high-order multiple-DM configurations to the extent that performance is affected.

4.2. Geometrical configurations

For the purpose of probing some of the geometrical MCAO configuration space, the wavefront reconstructor, the imaging wavelength and the number of PDEs have been kept fixed as the geometrical parameters are varied. Thus all simulations in this subsection were done with the LSE reconstructor in K band ($2.2 \mu\text{m}$) and without any WFS noise.

4.2.1. Field of compensation

The configuration parameter that seems to be the most crucial to the anisoplanatism error budget is the density of high-order reference beacons, *e.g.* the LGS. For a fixed number of LGS (which is the case we are considering) this translates directly into the size of the field one wishes to correct. Invariably, one finds that the quality of compensation decreases with an increasing field (*e.g.* decreasing LGS density). Hence, the size of the science field will ultimately be governed by what kind of penalties MCAO observers are ready to take in performance. The science field was defined as the square marked by the four LGS in the corners of the “X”, and the distance cited as “LGS distance” is half the side of this square. Using the generic intermediate-order 3-DM system described above, the LGS distance was varied between $17'' - 44''$ in steps of $4.5''$, and the results are plotted in the upper half of figure 4. There are two ways of comparing the performance for a given system configuration as the field size varies: either one keeps everything else fixed, or one allows the size of the altitude-conjugated DMs to vary to accommodate the changing size of the meta-pupil at different altitudes. These two alternatives may be expected to give slightly different results as the former experiences a loss of number of effective actuators when the field increases (solid line), and the latter changes the spatial frequency content of the DMs (dot-dashed line). Which case is the more correct or representative may be debated, but whichever one chooses for a norm, the other one will then be contaminated by a spurious effect. But as the plots show, these effects are not very prominent anyhow, making the distinction a non-issue.

4.2.2. DM conjugation altitudes

A key issue in the design of an MCAO system is the conjugation altitude of a particular DM. In most cases there will be a unique location that is optimal in, for instance, average Strehl ratio or uniformity over the field. These metrics need not necessarily be optimal at the same altitude, however, which again introduces a subjectivity in the definition of what should be thought of as optimal performance. To first order, this optimal altitude is purely a property of the atmospheric turbulence distribution, but it is plausible that it is also a function of the size of the science field and the order of the system. As with varying the field size previously, it is again possible to optimize keeping the size of the DMs fixed or allowing them to adapt to the meta-pupils as the conjugation altitude is varied. Given the results from §4.2.1 though, the dynamic version was chosen here. For simplicity, a 2-DM system with one DM fixed at zero

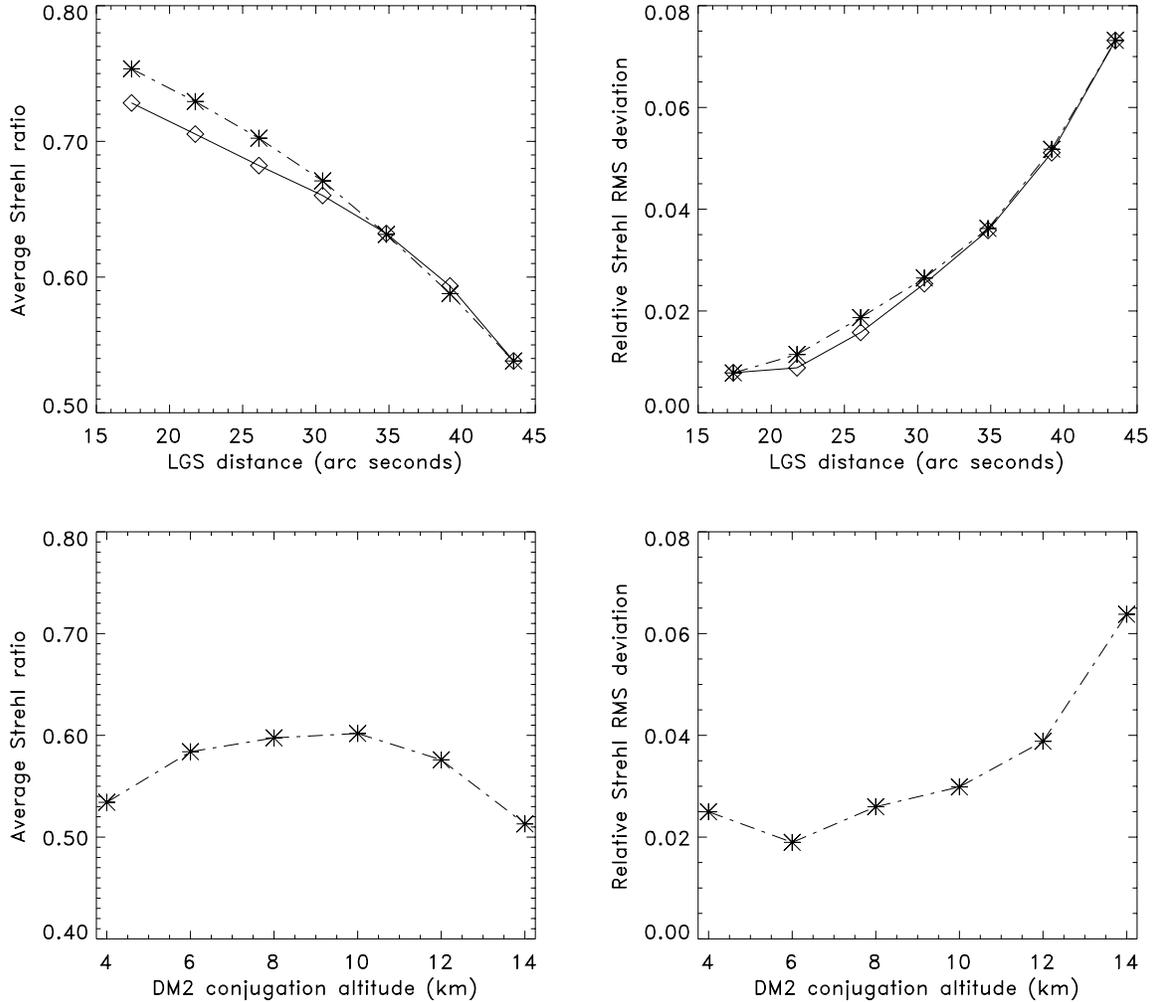


Figure 4. Upper panels: average \bar{S} and relative RMS deviation δ_S of the Strehl ratio for the generic 3-DM system as a function of the separation of the LGS (*cf.* §4.2.1). Lower panels: \bar{S} and δ_S for the 2-DM system described in §4.2.2, as a function of conjugation altitude for DM2.

kilometers and one varying between 4-14 km was used for correcting a $1'$ field in this simulation. As the plots in the lower half of figure 4 show, the average Strehl ratio peaks at approximately 10 km, whereas the uniformity hits a local extremum around 6 km. The somewhat wobbly shape of the δ_S curve could be a conditioning number related effect, as discussed above. Notable is the observation that there is a wide range spanning more than 6 km (DM2 \sim 6-12 km) where \bar{S} has an absolute variation of 2.6%. This suggests that the precise value of the conjugation altitude does not need to be particularly fine tuned. Equally noteworthy is the fact that within the same range, δ_S actually doubles. Some attempts have been made to derive analytical expressions for the optimal conjugation altitudes as functions of turbulence models, but as of yet, an accurate algorithm for this optimal estimation problem has still to be provided (see [12]). For our generic 3-DM case, however, a provisional optimum seems to be close to [0, 4, 12] km, which is in remarkable concordance with the numerical simulations in reference [15].

4.2.3. Number of DMs

Using an intermediate-order 12×12 subaperture WFS, the number of altitude-conjugated DMs was varied from 1-4. Due to the dominant nature of the ground layer turbulence in this atmospheric model, no appreciable field-widening was achieved by conjugating a single DM to some finite altitude. In stead, a decent trade-off between average Strehl and uniformity was found by conjugating the DM to the telescope pupil, leaving the field-widening to tomography

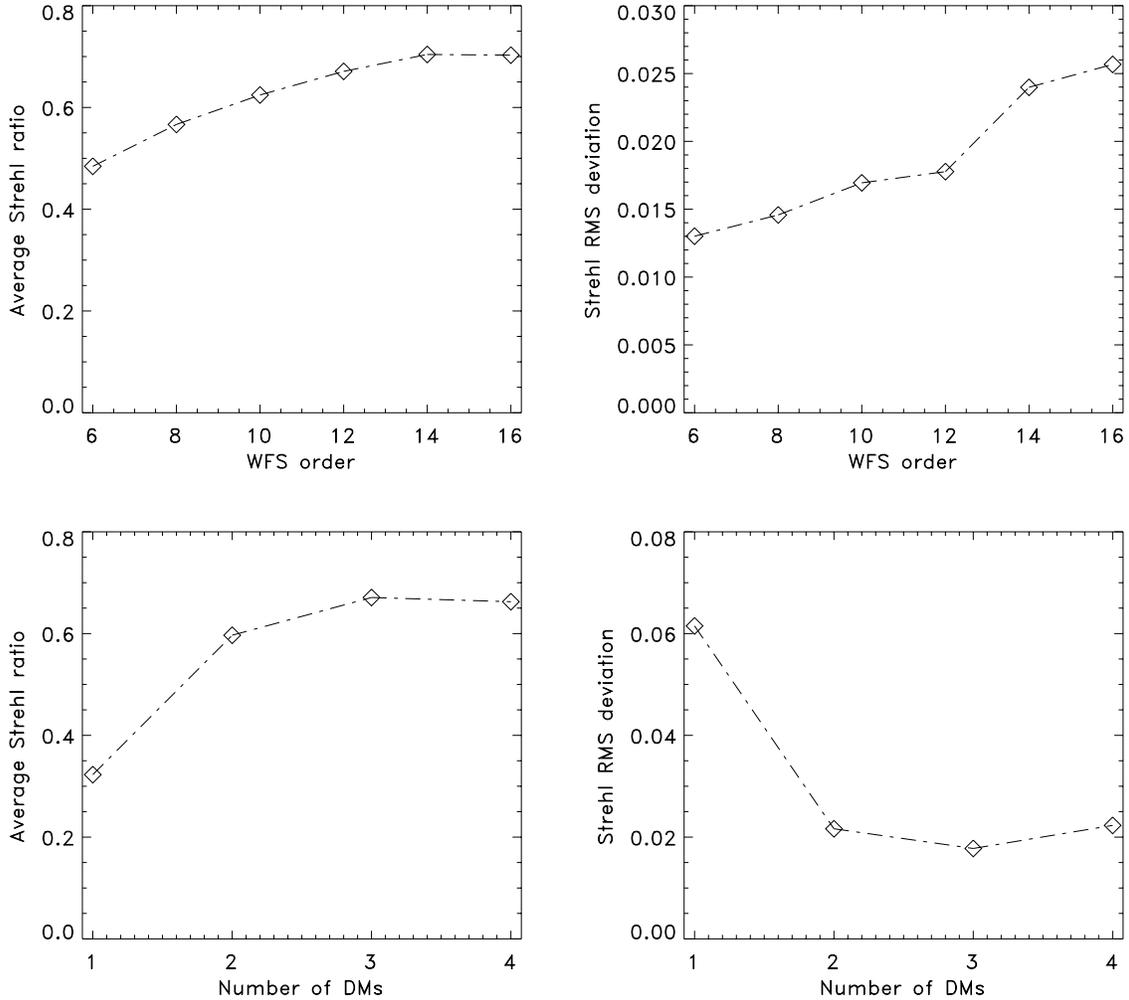


Figure 5. Average and RMS deviation of Strehls as a function of the WFS order for a 3-DM system conjugated to $[0,4,12]$ km (upper panels; cf. §4.2.4) and as a function of the number of DMs with a 12×12 subaperture WFS (lower panels; cf. §4.2.3).

alone. In the 2-4 DM cases, estimated optimal conjugation altitudes were employed, and the four cases run were: 1 DM with 13 actuators across the pupil conjugated to 0 km; 2 DM $[13,10]$ @ $[0,10]$ km; 3 DM $[13,11,9]$ @ $[0,4,12]$ km; 4 DM $[13,11,10,9]$ @ $[0,4,8,12]$ km. The results are plotted in the lower half of figure 5, and the one notable feature worth elaborating on is the fact that the 4-DM case is actually worse than the 3-DM case. This can only be a consequence of the worsened conditioning of the system as the number of modes increases, resulting in a poorer filtering of the singular values. Although this may not be the best possible result with four DMs (it is not), other studies have shown that the gain in going from three to four or five DMs is very small, and that three DMs is very close to optimal performance.⁶

4.2.4. Order of system

The “order” of the system would strictly imply the total number of active actuators in the system (or something proportional thereto). This is a rather degenerate measure, however, as the same order may be attained by distributing the same total number of actuators between the DMs in different proportions, thus producing very different performance characteristics. In analogy with a conventional AO system, we have chosen here the number of WFS subapertures across the pupil as the parameter defining the order - bearing in mind that this number is not linear with respect to the total number of actuators. Using a 3-DM system conjugated to $[0, 4, 12]$ km correcting a $1' \times 1'$

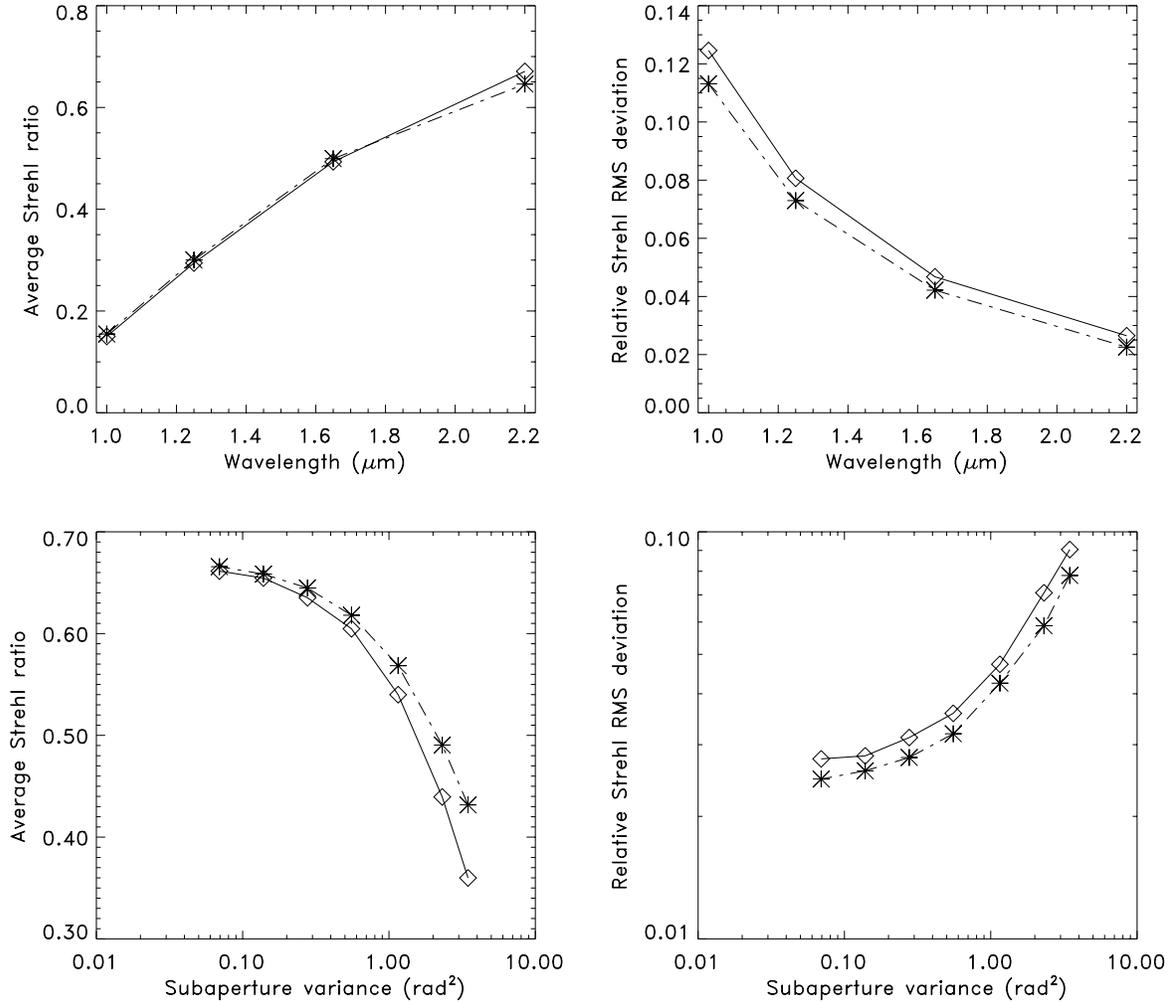


Figure 6. Average Strehl ratio \bar{S} and relative Strehl RMS δ_S as functions of wavelength (upper panels) and K band subaperture variance σ_m^2 (lower panels) for the generic 3-DM system, comparing the LSE (solid) and the MAP (dot-dashed) reconstructors.

field, the order was varied from 6-16 in steps of two. The linear number of actuators per DM were adjusted according to: [7,8,9], [9,9,9], [11,10,9], [13,11,9], [15,13,11], [17,13,11], and the plots of \bar{S} and δ_S are shown in the upper half of figure 5.

The “saturation” observed at the highest order considered here is likely a combination of two different effects: a fitting error level-off, and a filtering problem with the SVD. As the overall fitting error σ_{fit}^2 in an MCAO system may be somewhat hard to define, it is instructive to consider a conventional 1-DM/1-GS AO system as a first-order approximation. In this case the fitting error depends upon the linear number of actuators N like $\sigma_{fit}^2 \propto N^{-5/3}$, which shows that the gain in performance per unit increment of N decreases as N increases. Secondly, as already noted, the higher the order of the system the more problematic the selection of the conditioning number, and the more likely that the one chosen is *not* the optimal. The gradual increase in δ_S may be understood from a modal perspective, regarding the correction per LGS as a superposition of the corrections per mode for that LGS. It is known that for Zernike modes, the isoplanatic angle per mode $\theta_0(j)$ decreases rapidly with increasing radial degree.¹² As increasing the order of the system is equivalent to adding a few high-order modes, this means that the additional correction achieved will contribute increasingly to a smaller and smaller region around the line of sight to the LGS.

4.3. Wavefront reconstructors

For the purpose of comparing the performance of the reconstructors, a fixed geometrical configuration was chosen and the imaging wavelength and the number of PDEs were varied along with the type of reconstructor. The configuration is the generic 3-DM system at conjugation altitudes [0, 4, 12] km with [13, 11, 9] actuators across the diameter and the LGS spanning a 1' field. The imaging wavelength was varied from I - K band (*e.g.* 1.0, 1.25, 1.65 and 2.2 μm), and the photon noise was varied by setting the number of PDEs to [200, 100, 50, 25, 12, 6, 4].

4.3.1. MAP

The main results are plotted in figure 6, and it is seen that the MAP estimator performs consistently better than the LSE when the system is affected by photon noise (lower panels), and the differential also grows with the amplitude of the noise. Even in the absence of noise the MAP shows a somewhat lower δ_S , which may be attributed to its *a priori* knowledge on the residual phase statistics. This knowledge is encoded in the command covariance matrix C_c , and the weight it will have in the reconstruction process is in turn governed by the noise covariance matrix C_n . In this zonal reconstruction scheme we may consider the WFS subapertures as statistically independent with respect to photon noise, so the noise covariance is simply a diagonal matrix with the subaperture noise variance on the diagonal: $C_n = \sigma_m^2 \mathbf{x}I$, where I is the identity and the vector \mathbf{x} is a weighting factor to account for the partial illumination in the edge subapertures. C_c , on the other hand, must be generated in a noise-free closed loop simulation using the LSE. What is interesting though is that, when calculating the C_n matrix, one has the option to intentionally enter a too large or too small value for the variance, in the purpose of forcing the reconstructor to rely on phase statistics to a greater or lesser extent. The net effect of this manipulation is seen mainly in the uniformity δ_S , which is actually promoted by entering a higher noise variance that is expected from the system

Interesting to note is also how the MCAO system is degraded by noise in relation to a conventional AO system. Thus for comparison, the same simulations as in the lower half of figure 6 were done for one 13×13 DM conjugated to zero km with one on-axis LGS. Figure 3 (left panel) shows the ratio of the AO on-axis Strehl ratio to the MCAO ditto, for the MAP (dot-dashed) and the LSE (solid). We see that the AO system degrades faster than does the MCAO system, regardless of the reconstructor. This is a purely geometrical effect, independent of any phase statistics embedded in the reconstructor: due to the redundancy in WFS measurements over a large portion of the sky, the MCAO system attains better noise rejection by WFS tomography alone.

5. CONCLUSIONS

It has become clear, through simulations in this paper and other, that MCAO for 8-m class telescopes is an entirely feasible concept. This study has gone some way towards understanding the spatial characteristics and limitations of MCAO systems for the 50% Cerro Pachon turbulence model at a V band seeing of $0.7''$. Anisoplanatism remains as one of the major limitations to the size of the compensated field of view (as shown in the upper half of figure 4), until it is demonstrated that more than five LGS are feasible on a telescope. It is seen, for instance, that a 3-DM system of the order 12×12 WFS subapertures with a fixed configuration of five high-order LGS, can only sustain a science field of approximately $80''$ with $> 60\%$ average Strehl ratio and $< 5\%$ uniformity in K band. These numbers improve to $\bar{S} \sim 68\%$ and $\delta_S \sim 2.4\%$ for a 1' field. The DM conjugation altitude is seen to be the least sensitive parameter for systems with more than one DM, allowing some 2-3 km mismatch with the optimal position without degrading the performance more than a few percent in both average and uniformity. Using the LSE in a high- or intermediate-order system, there is essentially no gain in going beyond three DMs. The order of the system is an important parameter up to a certain level, around 16×16 WFS subapertures in K band in this study, where the diminishing gain in performance is eaten up by decreasing precision in the SVD filtering. This is a rather severe drawback for the LSE and the MAP, but more elaborate filtering methods may be employed, or one may choose to work with a different set of DM modes, which might improve performance. Future extensions of this work include a minimum variance estimator (MVE) for wavefront reconstruction,^{2,5} intended for this paper but not yet successfully implemented in the Monte Carlo code. A real MCAO system would also benefit in flexibility from a decomposition of the reconstructor into the parts pertaining to LGS and NGS, driven by separate control loops (currently under study). Hence, a decoupled LSE will also be implemented in the near future.

ACKNOWLEDGMENTS

The Gemini 8-m Telescopes Project and Observatory is managed by the Association of Universities for Research in Astronomy, for the National Science Foundation and the Gemini Board, under an international partnership agreement. The authors wish to thank Mark Chun for producing the layered turbulence profile used in the simulations, and Alexander Gontcharov for helpful comments and fruitful discussions.

REFERENCES

1. J. Beckers, "Increasing the size of the isoplanatic patch size with multiconjugate adaptive optics," in *ESO conference on Very large telescopes and their instrumentation*, M.-H. Ulrich, ed., p. 693, 1988.
2. B. L. Ellerbroek, "Scaling multi-conjugate adaptive optics performance estimates to extremely large telescopes," in *Adaptive Optical Systems Technology*, vol. 4007, SPIE, march 2000. These proceedings.
3. J. M. Beckers, "Detailed compensation of atmospheric seeing using multiconjugate adaptive optics," vol. 114, pp. 215–217, SPIE, 1989.
4. D. C. Johnston and B. M. Welsh, "Analysis of multiconjugate adaptive optics," *J. Opt. Soc. Am.* **11**, pp. 394–408, 1994.
5. B. Ellerbroek, "First order performance evaluation of adaptive-optics system for atmospheric turbulence compensation in extended field-of-view astronomical telescope," *J. Opt. Soc. Am. A* **11**(2), pp. 783–805, 1994.
6. 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 and Imaging through the Atmosphere III*, vol. 3762, SPIE, 1999.
7. R. Ragazzoni, E. Marchetti, and G. Valente, "Adaptive-optics corrections available for the whole sky," *Nature* **403**, pp. 54–56, January 2000.
8. B. L. Ellerbroek and D. W. Tyler, "Adaptive optics sky coverage calculations for the gemini-north telescope," *Publ. Astr. Soc. Pac.* **110**, pp. 165–185, 1998.
9. F. Rigaut and E. Gendron, "Laser guide star in adaptive optics: the tilt determination problem," *Astron. Astrophys.* **261**, pp. 677–684, 1992.
10. F. J. Rigaut (editor) *et al.*, "Multi-conjugate adaptive optics: A feasibility study for gemini-south," Internal report RPT-AO-G0091, Gemini Observatory, January 2000. Available upon request or at <http://www.gemini.edu/sciops/instruments/adaptiveOptics/AOIndex.html>.
11. R. J. Sasiela and J. G. Mooney, "An optical phase reconstructor based on using a multiplier-accumulator approach," in *Adaptive Optics*, vol. 551, pp. 170–174, SPIE, 1985.
12. F. J. Roddier, ed., *Adaptive Optics in Astronomy*, Cambridge University Press, 1999.
13. F. J. Rigaut, "Performance and limitations of multiconjugate ao," in *Adaptive Optical Systems Technology*, vol. 4007, SPIE, March 2000. These proceedings.
14. J. Vernin, A. Agabi, R. Avila, M. Azouit, R. Conan, F. Martin, E. Masciadri, L. Sanchez, and A. Ziad, "Gemini cp site characterization report," Internal report RPT-AO-G0094, Gemini Observatory, January 2000. Available upon request or at <http://www.gemini.edu/sciops/instruments/adaptiveOptics/AOIndex.html>.
15. D. L. Fried, "Fundamental limits in field widening for a multiple deformable mirror adaptive optics system," Tech. Rep. TR-1195, TOSC, march 1992.