Available online at www.sciencedirect.com
SCIENCE DIRECT®
C. R. Physique 6 (2005) 1070–1080
COMPTES RENDUS PHYSIQUE
http://france.elsevier.com/direct/COMREN/

MCAO for very large telescopes/OAMC pour les très grands télescopes

Simulation of MCAO on (extremely) large telescopes

Miska Le Louarn *, Christophe Verinaud, Visa Korkiakoski

ESO, Karl Schwarzschild Strasse 2, 85748 Garching, Germany

Available online 9 January 2006

Abstract

In this article, we describe the different methods to simulate Multi-Conjugate Adaptive Optics (MCAO) systems. First, analytical (error-budget type) and semi-analytical (Fourier) methods are described. We then describe the different modules required to make end-to-end (Monte Carlo) simulations of these systems. Finally, we present some of the computational challenges associated with the simulation of MCAO on Extremely Large Telescopes (ELTs). ***To cite this article: M. Le Louarn et al., C. R. Physique 6 (2005).***

Résumé

Simulation d'optique adaptative multi-conjuguée sur les (très) grands télescopes. Dans cet article, nous décrivons les différentes méthodes pour simuler des systèmes d'optique adaptative (multi-conjuguée, MCAO). Tout d'abord, nous présentons les méthodes analytiques (budgets d'erreur) et semi-analytiques (Fourier). Nous décrivons ensuite les différents modules nécessaires au développement d'une simulation complète (Monte Carlo) d'un tel système. Enfin, nous aborderons les problèmes lié à la simulation de la MCAO sur les télescopes extrèmement grands (ELTs). ***Pour citer cet article : M. Le Louarn et al., C. R. Physique 6 (2005).***

Keywords: Adaptive optics; Wavefront sensing; Simulations

Mots-clés : Optique adaptative ; Analyse de surface d'onde ; Simulations

1. Introduction

Before building an Adaptive Optics (AO) system, it is important to model it, in order to dimension its components to correspond to the requirements of the astronomers, and to understand its performance and limitations. Once the system is built, simulations can provide insight into the observed performance, and can help to discover and analyze bottlenecks in performance.

In this article, different levels of simulation are presented, going from the simplest, order of magnitude estimates provided by simple analytical formulae, to the semi-analytical algorithms providing PSF estimation to finish by the full numerical simulations, which allow to model very subtle effects by staying close to the physics of the atmosphere and instrument.

Finally, a section is devoted to the difficulties in making AO simulations for Extremely Large Telescopes (ELTs).

* Corresponding author.
E-mail addresses: lelouarn@eso.org (M. Le Louarn), cverinau@eso.org (C. Verinaud), vkorkiak@eso.org (V. Korkiakoski).

1631-0705/$ – see front matter
doi:10.1016/j.crhy.2005.10.004

2. Analytical modeling

2.1. Error budget

Error budget based tools provide crude estimates to start the dimensioning of an (MC)AO system. They allow, through simple analytical formulae, to estimate the contribution of various error sources. The variance of different errors (i) are added independently, and a Strehl ratio (S) estimate is obtained through the Maréchal criterion, valid for Strehl ratios down to ∼10%:

$$S = \mathrm{e}^{-\sum_i \sigma_i^2} \tag{1}$$

The main error sources for conventional AO (for a conventional Shack–Hartmann (SH) sensor) are detailed here. For a more exhaustive error budget analysis, see for example [1]:

- Aliasing and fitting errors, which are due to the finite spatial sampling of the Wavefront Sensor (WFS) measurement and deformable mirror (DM) correction. $\sigma^2 \propto (d_s/r_0)^{(5/3)}$, where d_s is the sub-aperture and actuator spacing, r_0 the Fried parameter (e.g. [2]).
- Photon noise error, due to the finite amount of photons used to compute the WFS measurement, $\sigma^2 \propto (1/N_{ph})\alpha^2$, where $\alpha = d_s/r_0$ if $d_s > r_0$, 1 otherwise [3].
- Read-out noise, which arises from the electronic noise of the WFS detector, $\sigma^2 \propto (\sigma_e^2/N_{ph}^2)\alpha^4$, where σ_e is the r.m.s read-out noise of the detector [3].
- The temporal error is due to the limited sampling frequency τ of the AO system, compared to the correlation time τ_0 of the atmosphere: $\sigma^2 \propto (\tau/\tau_0)^{(5/3)}$ (e.g. [4]).
- An anisoplanatism error arises if an off-axis guide star (located at an angle θ) is used to correct an on-axis science object. $\sigma^2 \propto (\theta/\theta_0)^{(5/3)}$, where θ_0 is the isoplanatic angle of the atmosphere [5].
- In addition, if a laser guide star (LGS) is used, the cone effect [6] reduces the performance, $\sigma^2 \propto (d/d_0)^{(5/3)}$, where d_0 is an atmospheric parameter describing the intensity of the cone effect [7].

Some of these formulae can be applied for pyramid and curvature sensors, whereas others are directly related to a slope sensor.

In multi-conjugate AO (MCAO), the anisoplanatism error can be decomposed in two contributions, one due to the finite number M of DMs and the other due to the finite number of guide stars. The DMs are unable to fully correct the atmospheric turbulence (composed of many thousands of layers), which this produces residual anisoplanatism (also called generalized fitting error, [8]): $\theta_M = 2.905(2\pi/\lambda)\int \mathrm{d}h\, C_n^2(h) F_M(h, H_1, H_2, \ldots, H_M)^{-3/5}$, where H_i are the conjugation heights of the DMs, and F_M are weighting functions. See [9] for more details.

The second error specific to MCAO is the measurement error, which arises because a finite number of guide stars does not allow to perfectly measure the atmospheric turbulence volume over the field of view which is to be corrected: $\langle\sigma^2(\theta')\rangle = e_k(\theta')(\Theta/\gamma_k)^{(5/3)}$, with $\gamma_k = (r_0/\delta_k)$, where δ_k is the effective thickness of turbulence for a given GS configuration, γ_k the radius of GS position for which the variance reaches 1 rad^2, Θ the distance between reference star and science object, $\theta' = \theta/\Theta$ and $e_k(\theta')$ represents the error variation in the corrected field of view (FOV, $e_k(\theta') = 1$) at the FOV center. For more details see [10].

These error budgets allow us to get an idea of what FOV and limiting magnitudes will be available with a given (MC)AO system, depending on the atmospheric conditions and system parameters. However, the PSF remains unaccessible. The following, semi-analytical methods allow us to also get an estimate of it.

2.2. Fourier methods

The Fourier (or frequency domain) models are similar to PSF reconstruction algorithms [11], in that they model the residual power spectral density (PSD, $S_{\phi c}(k,\theta)$) of the corrected AO phase:

$$S_{\phi c}(k,\theta) = S_{\text{fitting}}(k) + S_{\text{aliasing}}(k) + S_{\text{servo-lag}}(k) + S_{\text{noise}}(k) + S_{\text{anisop}}(k,\theta) \tag{2}$$

Each error source has it's own residual power spectrum. Once these errors have been computed (see [12] and [13], the structure function of the corrected phase can be written:

$$D_\phi(\rho) = 2\iint \mathrm{d}k\, S_{\phi c}\left[1 - \cos(2\pi \mathrm{i} k\rho)\right] \tag{3}$$

From there, the modulation transfer function is given by:

$$MTF\left(\frac{\rho}{\lambda}\right) = T\left(\frac{\rho}{\lambda}\right)\cdot \mathrm{e}^{-\frac{1}{2}D_\phi(\rho)} \tag{4}$$

where $T(\rho/\lambda)$ is the Modulation Transfer Function of the telescope without atmospheric aberrations.

This approach has been recently extended to work also for Ground Layer adaptive optics (GLAO) [14].

More recently, another Fourier based approach has been developed, which allows the modeling of MCAO, and also takes into account the correlations between the different error sources. This approach is described in detail in [15]. A freely available software package which implements this method, Cibola, exists (see Section 5).

Note that these methods assume an infinite telescope aperture for the estimation of some of the error sources. This amounts to neglecting the beam overlap effects and also has some impact on the aliasing estimation, for example. Also the use of LGSs is not presently supported.

One particular application of these Fourier methods is to compute the statistics of the AO residuals and to generate residual phase screens, following this statistics (similar to the method to generate atmospheric phase screens, see Section 3.1). One can then inject these screens into coronograph simulations for example, without having to run a full AO simulation. This approach is presented for example in [16]. Another way to use these simulations is to obtain an analytical PSF. It has the advantage of being a true (infinitely) long exposure PSF, without any residual speckles. This can be useful for example in the analysis of extreme (very high number of degrees of freedom, aimed at finding extra-solar planets) AO systems. Indeed, the speckle noise can then be added independently of the AO simulation, to take into account a finite integration time (which can be significantly longer than that of a numerical AO simulation).

3. Numerical modeling

Numerical (also called Monte Carlo or end-to-end) simulations of AO allow to model very subtle errors and to model very precisely what happens in an AO system. Almost any effect can be taken into account, because one models the propagation of the electric field (phase and intensity) through the atmosphere and then through the AO system. The drawback is that the time to simulate 1 second of integration time can take hours or even days, depending on the system and telescope size.

3.1. Atmosphere, propagation

To model atmospheric turbulence, one usually makes the approximation that the turbulence is concentrated in a few (from 1 or 2 to about 10) infinitely thin layers. In reality, several thousands of individual layers are present in the atmosphere, but this approximation is precise enough, provided there are more simulated layers than atmospheric correctors. These 'phase screens' can be generated using several different methods, of which the most simple is described here. For a more comprehensive discussion on this subject, one can read for example [17]. The 'standard' method is based on [18], where a screen of random numbers is taken and filtered according to the Kolmogorov (or von-Karman) spectrum and then inverse Fourier transformed.

The pros of this method are:

- It is fast, since only one Fourier transform is needed.
- It is simple to implement
- It produces 'circular' phase screens, which means that one edge of the screen is connect with the edge opposite to it. This allows to not need to worry about reaching the end of a phase screen, when temporal evolution is applied (see Section 3.6).

The cons are:

- Because of the periodicity property of the Fast Fourier Transform, the power spectrum of the turbulence is Kolmogorov (or von-Karman) only on small spatial scales. At the level of the whole screen, the wrapping reduces the amplitude of low order modes (like tip-tilt).
- To overcome the lack of low spatial frequencies, the phase screens need to be made significantly larger that what might needed for the simulation itself (roughly set by the pupil size in pixels).

Note that some of these cons can be solved by using more advanced phase screen generation algorithms (as proposed in [17]).

The choice of the number of pixels on the phase screen across the telescope's diameter is crucial. Indeed, this sampling determines the maximum field of view of the simulated wavefront sensor(s) and Point Spread Function(s). This field of view (θ_s) is given by:

$$\theta_s = \lambda \frac{D_{\text{pix}}}{D} \tag{5}$$

where D_{pix} is the diameter of the telescope in phase screen pixels and D is the telescope diameter [m]. λ is the wavelength at which the image is computed.

As an example, for a Shack–Hartmann WFS working at a wavelength of 0.7 μm, with a sensing field of view of 2.3″, and a sub-aperture size of 1 m, the minimum required number of pixels across the sub-aperture is 16. So for an 8 m telescope, at least 128 pixels across the pupil should be used.

Once the phase screens have been generated, one has to propagate these screens from their associated altitudes to the telescope aperture (or more precisely to the conjugation height of the WFSs). In the geometric optics approximation, a screen i is shifted by S_i (i.e. the number of pixels corresponding to S_i meters) and zoomed by a factor γ_i (if they are observed with Laser Guide Stars) according to the following:

$$S_i = \theta h_i \tag{6}$$

$$\gamma_i = \frac{H_{\mathrm{LGS}}}{H_{\mathrm{LGS}} - h_i} \tag{7}$$

where θ is the angular position of the guide star through which the observation is done.

Finally, all the phase screens associated to each wavefront sensor are added. Note that this procedure is changed when considering layer oriented wavefront sensing (see Section 3.3).

A more rigorous way of simulating propagation through the atmosphere is to use Fresnel propagation. Using Fresnel propagation allows to take scintillation effects into account. These are usually small at astronomical sites (for more details, see [19]). However, some differential scintillation effects can be non-negligible on extended objects like the sun [20].

3.2. Wavefront sensor(s) – star-oriented WFSs

Once the wavefront has been propagated to the wavefront sensor, the phase (and possibly the intensity) can be analyzed.

3.2.1. Shack–Hartmann WFS

The incoming complex amplitude screen is cut into small squares representing the Shack–Hartmann lenslets. It is advisable to have an integer number of simulation pixels per sub-aperture, in order to avoid extra interpolation steps.

The amplitude is then embedded in a larger support, and an FFT is made on this array, allowing to obtain the PSF of one lenslet. Several points are worth mentioning:

- The size of the support should be at least two times larger than the complex amplitude array, in order to ensure that the spot in the SH are Nyquist sampled in simulation pixels.
- The field of view in the resulting lenslet PSF is solely determined by Eq. (5).
- Usually, the FFT centers the peak of the PSF on a single pixels. This is not the optimal configuration for a SH, where one would like to have the spot centered on the intersection of 4 pixels. To achieve this, a tilt can be applied to the complex amplitude before the FFT.
- In the calibration phase (when building the interaction matrix), when only small perturbations are observed by the SH-WFS, the SH spot is (nearly) diffraction limited, which might not be the case when the WFS observes the atmosphere (even in closed loop, if the wavelength of wavefront sensing is in the visible). In order to distribute the light of the lenslet PSF onto several pixels (to avoid non-linearities in the calibration phase), one should simulate the use of a calibration fiber. This can be done by multiplying the input complex amplitude (before the FFT) by a Gaussian of adequate width (corresponding roughly to the FWHM of the SH-spots during the closed loop operation).

Now, the simulation pixels must be converted into CCD pixels. The easiest way of doing this is to have the lenslet PSF image be an integer number times larger than the expected CCD image. This allows to build the CCD image by averaging an integer number of simulation pixels.

To summarize, the constraints on the size (in simulation pixels) of the sub-aperture are:

$$d_{\mathrm{pix}} >= \theta_s \frac{d}{\lambda} \tag{8}$$

$$d_{\mathrm{support}} = N d_{\mathrm{pix}} \tag{9}$$

$$d_{\mathrm{support}} = N' N_{\mathrm{CCD}} \tag{10}$$

where N is an integer greater or equal to 2, N' is an integer greater or equal to 1. d_{support} is the support size (in pixels) in which the array of complex amplitudes is embedded.

An additional constrain can be set on d_{support} to accelerate the computation of the Fourier transform (e.g. it must be a power of 2). Since d_{support} has several constrains, usually one takes the largest pixel number of all the constraints.

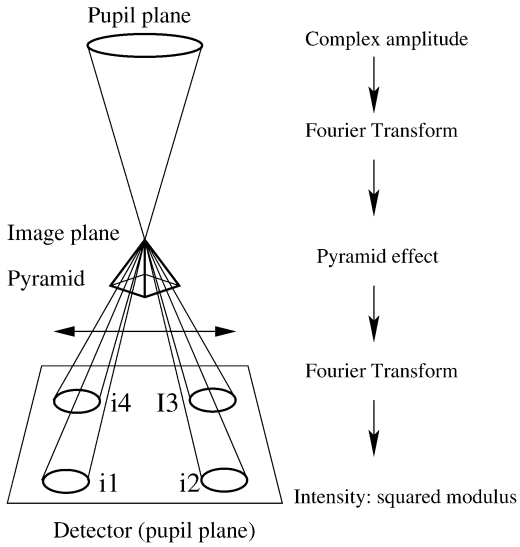

Fig. 1. Main steps for the Pyramid sensor simulations.

After the image in CCD pixels has been computed, noise has to be added. This can be photon noise (Poisson statistics) and/or read-out noise (Gaussian statistics). One can also simulate other instrumental effects (like non-uniform flat field). After the noise has been added, one can apply a thresholding algorithm before computing the centroids, as is done in real life WFSs.

The final stage is the computation of the measurement, using either a simple center of gravity, or weighted center of gravity [21]. This is done on the sensor image as one would do it on real images.

The advantage of such a simulation method for the WFS is that almost any kind of effect (scintillation, non-uniform flat field etc.) can be simulated. Since it is close to 'reality', new methods (like spatial filtering [22] can be easily implemented, by filtering the complex amplitude before injection in the WFS module. The drawback is that the method is computationally more demanding (use of the FFT), than just computing the slope as phase differences between sub-aperture edges.

3.2.2. Pyramid WFS

The pyramid wave-front sensor [23] is based on a generalization of the Foucault test for optical surfaces. The main difference with respect to the original Foucault test is the replacement of the knife edge by a pyramidal refractive optical element similar in its function to two knife-edges in two perpendicular directions. A dynamic modulation of the beam can also be applied by means, for example, of a tip-tilt mirror placed in a pupil plane. Beam modulation permits to linearize the response of the sensor and to increase the dynamical range for the low order modes.

Two approaches for modeling the pyramid wave-front sensor have been developed so far. End-to-end simulations using diffraction optics permits to simulate accurately the diffraction effects that are quite important in the Pyramid sensor. The diffractive model requires two Fourier transforms to compute the intensity distribution in the pupil plane as described in Fig. 1. Once the intensity distribution is obtained the signal is computed for each sub-aperture with the formula given in Eq. (11).

$$\begin{aligned} S_x(x,y) &= \big[\big(I_1(x,y)+I_2(x,y)\big)-\big(I_3(x,y)+I_4(x,y)\big)\big]/I_0 \\ S_y(x,y) &= \big[\big(I_1(x,y)+I_4(x,y)\big)-\big(I_2(x,y)+I_3(x,y)\big)\big]/I_0 \end{aligned} \tag{11}$$

where $I_i(x,y)$ is the intensity in the sub-aperture located at (x,y) in the quadrant i, integrated during a modulation cycle; I_0 is the average intensity per sub-aperture of the incoming beam.

The effect of the pyramidal prism on the focal plane electromagnetic field can be modeled in two ways, whether interference between the four pupils is taken into account or not. A complete description of the methods for end-to-end simulations of the Pyramid sensor can be found in [24] and the parallelized version of the algorithms has been described in [25]. This approach permitted to highlight several interesting new properties such as the possible application of pyramid wave-front sensing to interferometry as a differential piston sensor [26]. The simulation code developed in [24] has been used during the concept study of the Large Binocular Telescope single conjugate adaptive optics system [27]. The gain in terms of sensitivity to photon noise provided by the pyramid sensor has also been confirmed by these simulations. End-to-end modeling also permitted to show that even in the bright end regime the pyramid sensor gives better results than the classical Shack–Hartmann sensor. This effect has been explained using an analytical model developed in [28]. Analytical modeling is in fact the second approach

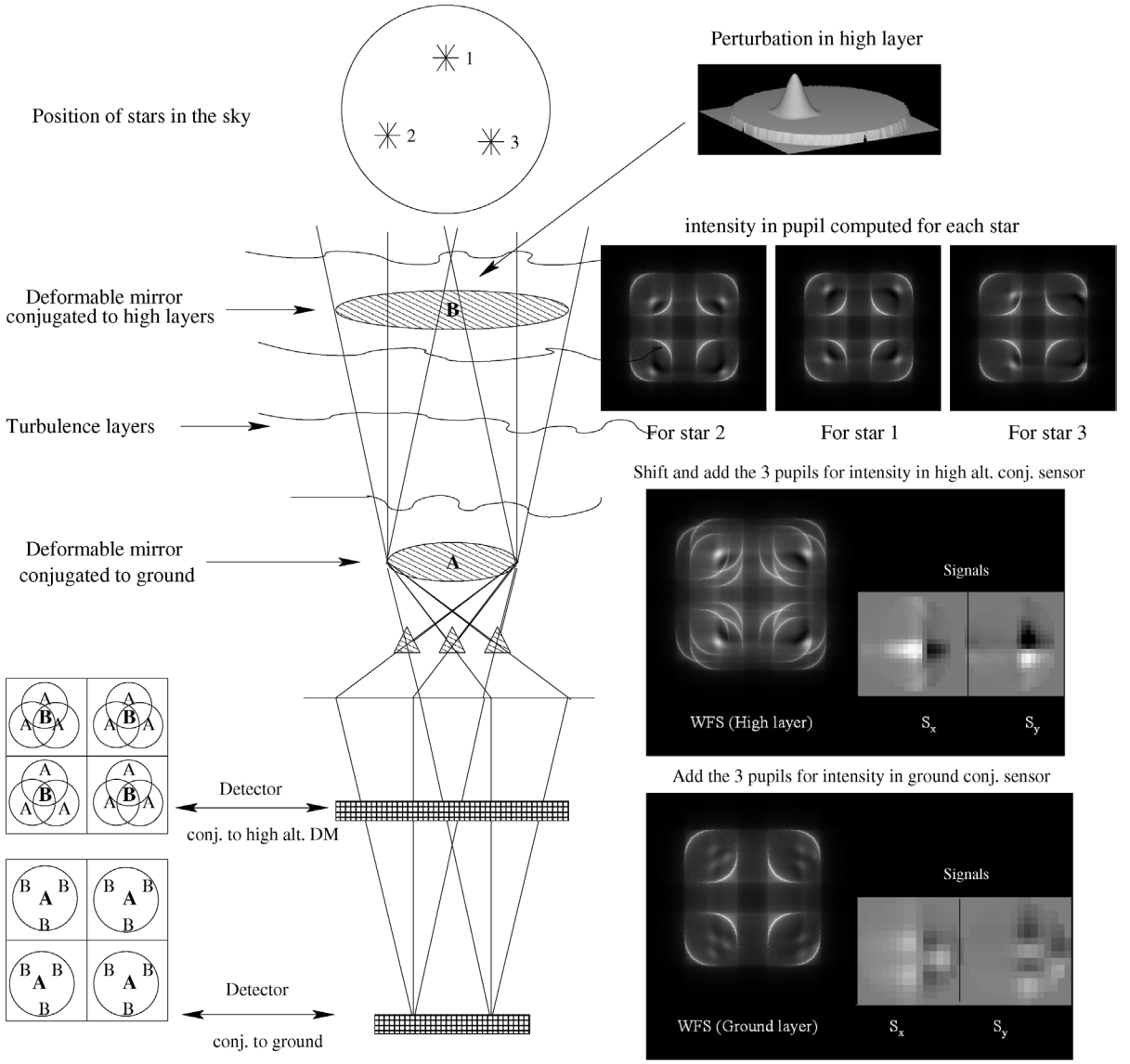

Fig. 2. The Layer-Oriented wave-front sensing concept based on refractive pyramids and method for simulation.

that has been developed. In [28] it has been shown that the Pyramid sensor, in the diffraction limited regime, is very different from a slope sensor and inherits actually properties from direct phase sensors. The main result of this model is that the noise propagation spectrum of the Pyramid sensor in the Fourier domain favors the correction of the low spatial frequencies with respect to a Shack–Hartmann. This is a very interesting property for the detection of exo-planets at small separations, especially with ELTs (the highest spatial frequencies show up as components far from the PSF core, and therefore to detect exoplanets close to the PSF core, a very good correction of the low orders is necessary). The analytical model of [28] has been used to predict the performance in halo rejection for an extreme adaptive optics system in the visible on a 100-m Extremely Large Telescope [29]. In this study, it has been shown that an order of magnitude in integration time can be gained for the detection of Earth-like planets in the habitable zone around a solar-type star at more than 20 pc distance.

3.3. Layer-oriented WFS

The Pyramid sensor performs the measurements in a pupil plane. This property has been used to define the concept of Layer oriented (LO) wave-front sensing [30]. The LO concept is illustrated in Fig. 2. LO methods allow us to reduce the effect of

read-out noise in the WFS, since the light of several stars can be optically co-added onto the same detector. Moreover, more stars can be used for WFS, since one does not need to add (noisy) detectors for each additional star. The method for end-to-end simulation of the Layer-Oriented scheme is very simple. First, the intensity distribution in the pupil plane is computed for the wave-front error seen by each individual guide star following the usual method described in Fig. 1. We can see in Fig. 2 an example of simulated intensity distributions when a perturbation occurs in the plane conjugated to the high altitude deformable mirror. This perturbation is seen at different location in the pupil by each pyramid, simply because the footprints on the high altitude plane are shifted one to each other. Then, in order to obtain the intensity distribution on the ground conjugated detector, these individual 4-pupil images are added up, taking into account the real brightness of the stars. For the high altitude conjugated detector, before adding up the 4-pupil images, these are shifted to take into account the footprints displacement when not in a pupil plane. Finally the signals are computed using Eq. (11). We can clearly see how the perturbation is seen in focus by the high altitude detector and out of focus for the ground conjugated detector. The method described here has been used to simulate the Layer-Oriented sensor [31] of the ESO Multi-conjugate Adaptive optics Demonstrator (MAD) [32].

3.4. Deformable mirrors

The deformable mirror can be described as an additional phase screen. If the order in which the phase screens does not matter (geometric propagation), then the DM can be placed directly in the atmosphere. In case Fresnel propagation is done, one can also propagate the shape of the DM shape to it's conjugation altitude. See [33] for more details on the order of conjugation of DMs in MCAO.

The shape of the DM is usually modeled with influence functions (IFs), which represent the way the DM deforms under the movement of a single actuator. Several functional representations of the IFs have been proposed:

- Linear splines, of the type $f(x, y) = (1 - \frac{x}{x_0})(1 - \frac{y}{y_0})$. The advantage of this type of function is that perfect piston (when all actuators of the DM are pushed simultaneously) and perfect tilt can be produced. These functions are however 'sharp', i.e. have high spatial frequency components. These can produce unwanted artifacts (spikes) on the simulated AO PSF.
- More realistic influence functions based on the fit of existing DM influence function. For example, Rigaut proposes to use:

$$f(x, y) = \left(1 - \left(\frac{x}{x_0}\right)^{3.805} + 3.74 \log\left(\frac{x}{x_0}\right)\left(\frac{x}{x_0}\right)^{2.451}\right) \cdot \left(1 - \left(\frac{y}{y_0}\right)^{3.805} + 3.74 \log\left(\frac{y}{y_0}\right)\left(\frac{y}{y_0}\right)^{2.451}\right)$$

  (F. Rigaut, private communication). The advantage of these is that they represent fairly accurately the shape of a piezo-stack deformable mirror IF. However, pushing all actuators produces a rippled surface (contrary to real DMs), which could produce artifacts.
- B. Ellerbroek (private communication) uses splines, $H(x, y) = h(x)h(y)$:

$$h(x) = \begin{cases} h(|x|) & \text{for } x < 0, \\ 1 + (4c - 2.5)x^2 + (-3c + 1.5)x^3 & \text{for } 0 \leqslant x \leqslant 1, \\ (2c - 0.5)(2 - x)^2 + (-c + 0.5)(2 - x)^3 & \text{for } 1 \leqslant x \leqslant 2, 0 \text{ for } 2 \leqslant x \end{cases}$$

  where c is the interactuator coupling coefficient. These produce a perfect piston when pushed simultaneously. Also, a perfect global tip-tilt can be produced (some IF models cannot do this, and produce an 'wavy' tilt, contrary to a true DM).

In general, all IFs of the same DM are assumed to be identical.

It should be noted that for the simulation of large DMs, the computation time of the shape of a DM can become large.

3.5. Interaction matrix generation

The interaction matrix allows to describe the response of the WFS(s) to a known shape of the DM. There are two ways to compute this matrix, which can be generalized to MCAO through the use of a proper propagation routine (either geometrical or Fresnel). To generate such a matrix, one pushes a mode on the DM, and records a vector of measurements on the WFS(s). These modes can either be the IFs (in which case the interaction matrix is 'zonal') or they can be global modes on the DM (Zernikes, Karhunen–Loeve polynomials, etc., 'modal' IM).

In the case of MCAO, the procedure is repeated for each star/WFS and each DM. Propagation should be done as it will be done in the closed loop phase. Each measurement vector is then stored in a matrix of size $N_{\text{act}} \times N_{\text{meas}}$ where N_{act} is the **total** number of actuators (i.e., the sum of all active actuators on each DM concatenated in a single vector) and N_{meas} the **total** number of measurements (i.e., the sum of the measurements of all WFSs, concatenated into a single vector).

Once this interaction matrix is filled, it can be inverted through SVD filtering, or a more complex reconstruction method can be used (see J.M. Conan et al., in this volume).

It should be noted that the generation of the IM can be numerically intensive, and it is possible to compute synthetic IMs, by computing analytically the measurement obtained from some DM shape. This is done, for example, by pre-computing the derivative of the IFs, in case a slope sensor is used.

3.6. Temporal evolution, control

Once the control matrix has been obtained, almost all elements are available to close the AO loop. The sequence is as follows:

- Initialization
  - Generate WFS, DM geometries.
  - Generate atmospheric phase screens.
  - Load or generate the control matrix.
- Propagation
  - Propagate phase screens to the WFS conjugation plane, taking into account the location and height of each guide star.
- AO model
  - Compute the measurement vector for each WFS.
  - Multiply by control matrix to generate command vector.
  - Generate DM shape according the command vector.
- Performance estimation
  - Compute the instantaneous (short exposure) PSF in all evaluation directions, for all wavelengths.
  - Compute short exposure Strehl ratio, to see how system evolves.
  - Add short exposure PSF to the previously obtained PSFs, to get the long exposure PSF.
  - Compute long exposure Strehl ratio, to see how system evolves.
- Temporal evolution
  - Apply DM shape (with appropriate temporal delay).
  - Shift phase screens according to wind speeds. The shift should be done by a non-integer number of pixels to avoid jumps of turbulence between two iterations.
  - Return to propagation phase for next iteration.

The temporal delay in this approach is simulated by not applying the DM shape in the same iteration as the commands were computed. Instead, this shape is stored in a stack and applied at a later iteration (say 1 or 2 iterations later to simulate a 1 or 2 frames delay). Between iterations, the phase screens are shifted, to create a lag. If a delay of a non-integer multiple of the integration time is to be simulated, a finer simulation clock beat is required. In that case, not all the operations above are applied during one iteration.

This method allows us to very finely simulate the dynamical evolution of the AO loop, and to study for example how stable it is. This is more difficult to perform in an analytic approach, since only a very crude model of the loop is usually used.

4. Simulation challenges for ELTs

Many challenges are met when simulating AO on 30–100 m ELTs. Most of these are related to the computational complexity of algorithms involved in the simulation.

Several bottlenecks can be identified in the simulation of MCAO systems.

4.1. Matrix operations

The generation of the command matrix is an $O(n^6)$ process, where n is the number of controlled actuators. Applying the command matrix to a measurement vector is an $O(n^4)$ process. These scaling laws show that going from a 10 m-call telescope AO simulation to a 100 m is a huge jump in the computational requirements. To solve this, several approaches have been suggested.

For the matrix operations, a combination of several approaches seems to be necessary:

- Parallel computing allows to process massive amounts of data quickly. The matrix operations (matrix multiplications, SVD, etc.) parallelize efficiently, and therefore the gain in using several computers simultaneously is large.

Table 1
Pros and cons of the different simulation tools

Method	Pros	Cons
Analytic	Immediate results Scaling laws easily seen	No PSF Difficulty to add new effects/methods Outer scale effects not always taken into account Error coupling not taken into account
Semi-analytic	Fast Provides PSF Infinite integration time (good for XAO) Scaling laws are 'visible' Residual phase screens can be made Error coupling can be taken into account	AO 'inperfections' difficult to add LGSs not taken into account (yet)
Numerical	Very accurate Easy to add new effects and methods Easier comparison of real and simulated data Short exposure data available (speckles) Residual phase screens can be made Error coupling taken into account HW (alignments, ...) effects on loop can be analyzed	Slow Require supercomputers for ELT simulations No obvious scaling laws for results

- Sparse matrix methods allow to reduce the computational complexity by optimizing the way the control matrix is computed. The reasoning, is that the effect of an actuator is local, and the associated response is also local (at least in the Shack–Hartmann WFS). Therefore, the interaction matrix is sparse, i.e. most of its elements are zero. If LGSs are used, the tilt removal causes a de-sparcification of the matrix, if special care is not taken (see [34]).
- With sparse systems, it is advantageous to not compute the inverse matrix, but to solve, iteratively, the direct equation $Ac = s$, where A is the interaction matrix and c and s are the command and measurement vectors, respectively. This approach has been demonstrated in [35].
- FFT reconstructor allow to compute, from a set of Shack–Hartmann slopes, the command vector to be applied [36]. Recently, it has been shown that applying modal optimization to this method is also possible [37]. Unfortunately, applying this method to MCAO systems has not been demonstrated.

4.2. Simulation improvements

Several approaches have been used to improve the performance of AO simulations. The re-coding of existing simulation tools from problem solving environments (like IDL or Matlab) into low level languages (C, C++) allows us to gain significant amounts of computing time. This approach is however very time demanding, due to the cumbersome programming and debugging involved. Hybrid approaches, where only parts of the routines are re-written have shown some success.

Another avenue is to parallelize most of the routines, in order to share the computing load between several CPUs, and to also distribute the large arrays, to reduce the memory requirement per machine (until the advent of low cost 64 bit PC-type machines, accessing more memory than $\sim$3 GB was based on fairly expensive hardware). Some AO tasks are easily parallelize. In addition to matrix operations, the generation of phase screens (either by distributing different layers on different machines or by computing the required large FFTs in parallel), the FFTs in the wavefront sensing process, in the Fresnel propagation process and the PSF computations can easily be parallelized. Particular care must be taken in the overall design of such a large parallel simulation tool, in order to assure maintainability, but also to provide optimum performance.

5. Conclusions

It has been shown in this article how one can simulate (MC)AO system, using different approaches. The pros and cons of each method are summarized in Table 1.

In the course of the life of an AO system, all the different simulation tools are likely to be used. First, at the very early stages, the analytical formulae allow to get a rough estimate of what kind of system is possible (considering the performance requirements and technological limitations), and gives an estimate on how many sub-apertures, actuators, deformable mirrors are required to fulfill the specifications. Then, the semi-analytical tools are used to refine these estimates, and make a trade-off

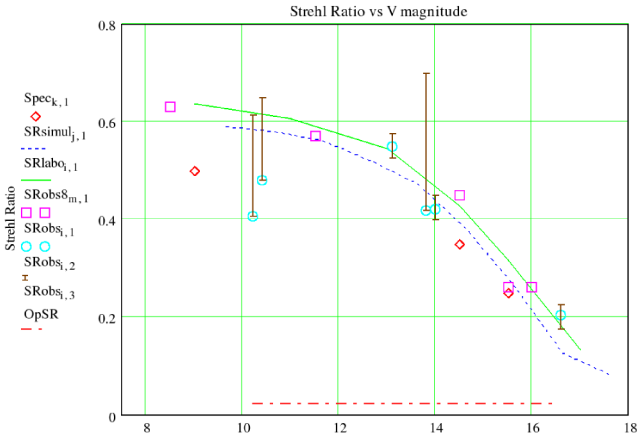

Fig. 3. Comparison between the performance of MACAO in simulations and on the sky. Strehl ratio (K-band) vs. guide star magnitude. The top line is the curve obtained in the lab, the dashed line is the simulation, squares and circles are results obtained at the telescope, diamonds are the specifications, the dot-dashed line is the open loop seeing. Courtesy R. Arsenault and the MACAO team.

analysis. Some PSFs can be generated and provided to the instrument builders and astronomers to evaluate if a given AO system provides the performance they need. Also, some residual phase screens can be generated and given to instrument builders (for example, to study the performance of a coronographs or a spectrograph). Finally, the end-to-end (numerical) software is used to specify for example the tolerances on the alignments, to verify how new algorithms perform, and to get a final performance estimate against which the system will be compared during commissioning. When the system has been integrated in the lab, turbulence phase screens can be placed in the optical setup to produce for the AO system the same conditions than those used in the simulation. Finally the performance at the telescope can be compared with the simulation results, once for example the telescope vibration and such 'real-life' effects have been characterized. Because the end-to-end tool is so close to the true hardware, such effects can be added fairly easily.

The comparison with real systems and simulations has been carried out for example for NACO, the VLT's Shack–Hartmann AO system, and the results can be found in [38]. These show that when all the additional errors encountered at the telescope are taken into account, a close agreement between simulation and reality can be obtained. For MACAO, another ESO AO system, a sample comparison is shown in Fig. 3. We can see that the agreement between simulation, lab results and on-sky tests is very good. For MCAO, experiments (both in the lab and on the sky) arebeing made at various institutes, and one of their goals is to validate simulation tools for this new AO mode (see, for example, the contribution of E. Marchetti et al. on MAD, in this issue).

Freely available tools can be used to do one's own AO simulations. Here are a few examples:

- Analytical tools
  - PAOLA: ftp://ftp.hia.nrc.ca/pub/staff/lj/softwares/paola.
  - Cibola: http://cfao.ucolick.org/software/cibola.php.
- Numerical tools
  - CAOS: http://www.arcetri.astro.it/caos/ [24].
  - Yao: http://www.maumae.net/yao/aosimul.html.
  - Aroyo: http://eraserhead.caltech.edu/arroyo/arroyo.html.

Acknowledgements

The authors wish to thank the Organizers of the JSO for this very productive workshop, and the anonymous referee who helped improve the quality of this paper.

References

[1] R.R. Parenti, R.J. Sasiela, J. Opt. Soc. Am. A 11 (1994) 288.
[2] D.L. Fried, J. Opt. Soc. Am. 67 (1977) 370.
[3] G. Rousset, in: D.M. Alloin, J.-M. Mariotti (Eds.), Adaptive Optics for Astronomy, Kluwer Academic Publisher, 1994, pp. 115–137.

[4] D.L. Fried, in: D.M. Alloin, J.M. Mariotti (Eds.), Atmospheric Turbulence Optical Effects: Understanding the AO Implications, Kluwer Academic Publisher, 1994, pp. 25–57.
[5] D.L. Fried, J. Opt. Soc. Am. 72 (1982) 52.
[6] R. Foy, A. Labeyrie, Astron. Astrophys. 152 (1985) L29.
[7] G.A. Tyler, J. Opt. Soc. Am. A 11 (1994) 325.
[8] F. Rigaut, B.L. Ellerbroek, R. Flicker, in: P.L. Wizinowich (Ed.), Adaptive Optical Systems Technology, vol. 4007, 2000, pp. 1022–1031.
[9] A. Tokovinin, M. Le Louarn, M. Sarazin, J. Opt. Soc. Am. A 17 (2000) 1819.
[10] A. Tokovinin, E. Vernet, J. Opt. Soc. Am. A 18 (2001) 873.
[11] J.-P. Véran, F. Rigaut, H. Maitre, J. Opt. Soc. Am. A (1997) 257.
[12] F.J. Rigaut, J.-P. Veran, O. Lai, in: Proc. SPIE, vol. 3353, 1998, p. 1038.
[13] L. Jolissaint, J. Veran, in: E. Vernet, R. Ragazzoni, S. Esposito, N. Hubin (Eds.), Beyond Conventional Adaptive Optics: A Conference Devoted to the Development of Adaptive Optics for Extremely Large Telescopes, Proceedings of the Topical Meeting Held May 7–10, 2001, Venice, Italy, 2002. 2002 ESO Conference and Workshop Proceedings, vol. 58, European Southern Observatory, Garching, Germany ISBN 3923524617, p. 201.
[14] J.A. Stoesz, L. Jolissaint, J. Veran, J. LeDue, in: D.B. Calia, B.L. Ellerbroek, R. Ragazzoni (Eds.), Advancements in Adaptive Optics, in: Proc. SPIE, vol. 5490, 2004, pp. 713–720.
[15] B.L. Ellerbroek, J. Opt. Soc. Am. A 22 (2005) 310.
[16] R. Conan, T. Fusco, G. Rousset, D. Mouillet, J.-L. Beuzit, in: Advancements in Adaptive Optics, vol. 5490, 2004.
[17] C. Harding, R.A. Johnston, R.G. Lane, Appl. Opt. 38 (1999) 2161.
[18] B.L. McGlamery, in: Proc. SPIE, vol. 74, 1976, p. 225.
[19] C. Vérinaud, M. Carbillet, in: EAS Publ. Series, 2003, pp. 209–219.
[20] C. Robert, J.-M. Conan, V. Michau, T. Fusco, N. Vedrenne, J. Opt. Soc. Am. A (2005).
[21] M. Nicolle, T. Fusco, V. Michau, G. Rousset, in: Advancements in Adaptive Optics, in: Proc. SPIE, vol. 5490, 2004.
[22] L.A. Poyneer, B. Macintosh, J. Opt. Soc. Am. A 21 (2004) 810.
[23] R. Ragazzoni, J. Mod. Opt. 43 (1996) 289.
[24] M. Carbillet, C. Vérinaud, B. Femenía, A. Riccardi, L. Fini, Mon. Not. R. Astron. Soc. 356 (2005) 1263.
[25] M. Le Louarn, C. Verinaud, E. Fedrigo, V. Korkiakoski, N. Hubin, in: Advancements in Adaptive Optics, in: Proc. SPIE, vol. 5490, 2004, pp. 705–712.
[26] C. Verinaud, S. Esposito, Opt. Lett. 27 (2002) 470.
[27] M. Carbillet, C. Vérinaud, S. Esposito, et al., in: P.L. Wizinowich, D. Bonaccini (Eds.), Adaptive Optical System Technologies II, in: Proc. SPIE, vol. 4839, 2003, pp. 131–139.
[28] C. Verinaud, Opt. Comm. 233 (2004) 27.
[29] C. Vérinaud, M. Le Louarn, V. Korkiakoski, M. Carbillet, Mon. Not. R. Astron. Soc. 357 (2005) L26.
[30] R. Ragazzoni, E. Diolaiti, J. Farinato, et al., Astron. Astrophys. 396 (2002) 731.
[31] C. Vérinaud, C. Arcidiacono, M. Carbillet, et al., in: P.L. Wizinowich, D. Bonaccini (Eds.), Adaptive Optical System Technologies II, in: Proc. SPIE, vol. 4839, 2003, pp. 524–535.
[32] E. Marchetti, R. Brast, B. Delabre, et al., in: Advancements in Adaptive Optics, in: Proc. SPIE, vol. 5490, 2004, pp. 236–247.
[33] R. Flicker, Opt. Lett. 26 (2001) 1743.
[34] B.L. Ellerbroek, J. Opt. Soc. Am. A 19 (2002) 1803.
[35] L. Gilles, C.R. Vogel, B.L. Ellerbroek, J. Opt. Soc. Am. A 19 (2002) 1817.
[36] L.A. Poyneer, D. Gavel, J.M. Brase, J. Opt. Soc. Am. A 19 (2002) 2100.
[37] L.A. Poyneer, J.P. Veran, J. Opt. Soc. Am. (2005), submitted for publication.
[38] Y. Clenet, M.E. Kasper, N. Ageorges, et al., in: D.B. Calia, B.L. Ellerbroek, R. Ragazzoni (Eds.), Advancements in Adaptive Optics, in: Proc. SPIE, vol. 5490, 2004, pp. 107–117.