Results of AO simulations for ELTs

R. Conan, M. Le Louarn, J. Braud, E. Fedrigo and N. Hubin
European Southern Observatory – Instrumentation Division – Adaptive Optics Group
Karl-Schwarzschild-Str. 2, D-85748 Garching bei Muenchen, Germany

ABSTRACT
The design and elaboration of Extremely Large Telescopes (ELT) with primary mirror from 20m to 100m face many challenges: mechanical, optical, computational, etc. To benefit completely of the full potentiality of such facilities, an Adaptive Optics System (AOS) has also to be designed for these telescopes. For whole field-of-view compensation and full sky coverage, the new but promising Multi-Conjugated Adaptive Optics (MCAO) technique has to be envisaged. The first step towards the design of an MCAO system is the numerical simulation. This is the first challenge we have to face. The scale of AO simulations being imposed by the ratio $(D/r_0)^2$, the simulation requirements of a MCAOS for an ELT, in terms of computing power and memory available to store the data, reach and sometimes surpass the capacity of current computers. At ESO, we have evaluated different hardware and software strategies to achieve MCAO simulation goals. Two codes have been developed to simulate MCAOS using an analytical and an end-to-end model. The goals and advantages/limitations of both approaches are shown. The hardware requirements for both methods are also given through the size of their largest matrices. And finally, results of hardware and software tests for MCAO simulations with PC-cluster and paralleled code are presented.

Keywords: adaptive optics, Karhunen-Loeve basis, object-oriented, parallel programming, workstation cluster

1. INTRODUCTION
The design and elaboration of Extremely Large Telescopes (ELT) with primary mirror from 20m to 100m face many challenges: mechanical, optical, computational, etc. To benefit completely of the full potentiality of such facilities, an Adaptive Optics System (AOS) has to be designed for these telescopes. For whole field-of-view compensation and full sky coverage, the new but promising Multi-Conjugated Adaptive Optics (MCAO) technique has to be envisaged. The first step towards the design of an MCAO system is the numerical simulation. This is the first challenge we have to face. The scale of simulations and mainly the array sizes are proportional to the ratio $(D/r_0)^2$ where $D$ is the diameter of the telescope and $r_0$ is the Fried’s parameter. The storage of such arrays becomes rapidly beyond the capacity offered by the memory of current computer. Moreover these simulations will be very demanding in terms of computing time and different solutions have to be investigated, among them: powerful workstations, PC-clusters, super-computers, grid-computing.

Taking the example of OWL, it will be equipped with at least two adaptive optics systems. In order to design this system, thorough simulations are needed. Among the parameters to be constrained, there is, for example:

- the deformable mirror(s) characteristics (number of actuators, actuator stroke), and for MCAO systems number of DMs and conjugation height,
- the effect of given influence functions on AO performances (e.g. if MEMs are used),
- the guide star nature (LGS/NGS), the number of guide stars, the required brightness, the spot elongation correction efficiency if LGSs are used.

Further author information: (Send correspondence to R. Conan)
R. Conan: E-mail: rconan@eso.org. Telephone: +49 (0)89 3200 6541
• the efficiency of the control laws (sparse and non sparse matrix techniques will be tested)

For non-MCAO systems, scaling laws exist, allowing rough dimensioning of an AO system without running a full simulation. Such scaling laws are not yet available for MCAO on a 100m telescope.

The strategies followed at ESO for AO/MCAO simulation for OWL is to build two codes: an analytical and an end-to-end model. The first one is based on a simplified open-loop system and is expected to be developed faster with respect to the end-to-end model. It will be already able to provide relevant informations for system analysis, informations that the end-to-end closed-loop model will be able to re-use for finer and advanced studies.

This paper presents both approaches with first AO results and gives an overview of the hardware chosen to implement the codes.

2. ANALYTICAL MODEL

2.1. Principle

The analytical model consists in the derivation of the optimal command matrix of an MCAO system. The optimization is performed on the residual phase variance after partial phase correction using a minimum-variance reconstruction method. The residual phase variance after an optimal correction in direction $\theta$ can be written\(^1\,^2\):

$$\langle e^2(\theta) \rangle = \langle e_0^2 \rangle - \text{tr} \left[ 2MT^T - RMSM^T \right]$$

(1)

where $\langle e_0^2 \rangle$ is the piston removed phase variance, M the command matrix to derive and T, R and S are, following the denomination in Ref. 2, the target, cross-talk and signal covariance matrices, respectively. The minimization of $\langle e^2(\theta) \rangle$ leads to\(^3\)

$$M = R^{-1}TS^{-1}$$

(2)

Our method follows the one developed in Ref. 2. The only difference is the replacement of the Zernike polynomials basis by a Karhunen–Loeve basis. It has been shown\(^4,\,^5\) that a Karhunen–Loeve basis is the optimal expansion of corrugated phase as it is constituted of the phase fluctuation eigen vectors. Moreover, due to their polynomial origin, Zernike basis are useless to model AO system with a very high order of correction as the highest modes have significant values only on the edges of the telescope pupil. However, the drawback of the Karhunen–Loeve basis is the lack of an analytical formula to generate it. Usually, a Karhunen–Loeve basis is build by expanding each of its vector on another basis, like the Zernike one\(^6\) for example. To avoid the artifacts introduced by another basis, we have used an original method\(^7\) to generate a Karhunen–Loeve basis from scratch. This has permitted to build a very high order basis, up to 40,10\(^3\) modes in a relatively short time, around one hour on a PC workstation. More recently, a Karhunen–Loeve basis of 100,10\(^3\) (Fig. 2) has been computed using a parallel version of the code running simultaneously on six processors.

2.2. Practical implementation

The analytical model is implemented as a library written in C++ following an object-oriented architecture. The heart of the library is constituted by a set of classes called “general components”. These classes with their interdependence are represented on the collaboration diagram in Fig. 1 where full arrows show inheritance and dashed arrows show layering. They are the required modules to model the wavefront phase statistics after propagation through atmospheric turbulence. The turbulence profile parameters are stored into the turbulence_profile class and the optical parameters of the turbulence belong to the atmosphere class which inherits from the turbulence_profile class being also a turbulence profile. There is also the class sources containing the coordinates of the unresolved star from which the wavefront is coming from. The second order statistics of the phase is derived by the phase_stat class which implements both atmosphere and sources classes. To derive telescope pupil-average statistics of the phase, we need the telescope class which contains telescope parameters and the tel_phase_stat class inheriting from phase_stat class and implementing the telescope
2.3. Preliminary results

The development of the code for a full numerical/analytical analysis of an MCAO system is still in progress. However, the already existing routines allow to investigate performances of ELT with and without phase correction with eXtreme Adaptive Optics (XAO).

We will restrict our study to the extreme case of a 100m diameter telescope. The atmospheric model used to derive second orders phase fluctuations statistics is the Von Kármán model for which the spectrum is given by:

\[
W_\varphi(f) = 0.0229 r_0^{-5/3} \left( f^2 + \frac{1}{\mathcal{L}_0^2} \right)^{-11/6}
\]  

with \(0.0229 = \frac{\Gamma^2(11/6)}{2\pi^{1/3}} \left[ \frac{24}{5} \Gamma \left( \frac{6}{5} \right) \right]^{5/6}\), \(r_0\) the Fried’s parameter and \(\mathcal{L}_0\) the spatial coherence outer scale (aka. outer scale).

---


†Fast Fourier Transform in the West, http://www.fftw.org/

‡Hierarchical Data Format version 5, http://hdf.ncsa.uiuc.edu/HDF5/

Figure 2. At the top, the full $100 \times 10^3$ Karhunen-Loève mode computed for a 100m-diameter telescope with a central obscuration ratio of 1/3. The values of the optical parameters of the turbulence used for this computation are 0.8m for the Fried’s parameter at a wavelength of 2.2μm and 25m for the outer scale. The circular regular patterns are Moiré effects. The outlined sector indicates the location of the zoom plotted below showing the real pattern of this mode.
The criteria used for performances evaluation are the Strehl ratio (SR) and the full width at half the maximum (FWHM). Both are derived from the point-spread function (PSF) of the system telescope-atmosphere. The PSF is computed performing a Hankel transform on the optical transfer function (OTF) given by

\[
OTF_{\text{tel.+atm.}} = OTF_{\text{tel.}} \times OTF_{\text{atm.}}.
\]

\[
OTF_{\text{tel.}}(\rho) = \frac{2}{\pi} \left\{ \left( \frac{1}{1-r_i^2} \right) \left[ \arccos \left( \frac{\rho}{D} \right) - \left( \frac{\rho}{D} \right) \sqrt{1 - \left( \frac{\rho}{D} \right)^2} \right] - \left( \frac{r_i^2}{1-r_i^2} \right) \left[ \arccos \left( \frac{\rho}{Dr_i} \right) - \left( \frac{\rho}{Dr_i} \right) \sqrt{1 - \left( \frac{\rho}{Dr_i} \right)^2} \right] \right\}
\]

is the telescope OTF with \(D\) and \(r_i\), the diameter and the central obscuration ratio, respectively.

\[
OTF_{\text{atm.}} = \exp \left( -\frac{1}{2} D_{\varphi}(\rho) \right)
\]

is the atmosphere OTF where \(D_{\varphi}(\rho)\) is the phase fluctuations structure function given by

\[
D_{\varphi}(\rho) = 0.1717 \left( \frac{L_0}{r_0} \right)^{5/3} \left[ \frac{\Gamma(5/6)}{2^{5/6}} - \left( \frac{2\pi \rho}{L_0} \right)^{5/6} K_{5/6} \left( \frac{2\pi \rho}{L_0} \right) \right]
\]

where \(0.1717 = 2 \left[ \frac{24}{5} \frac{\Gamma \left( \frac{6}{5} \right)}{\Gamma \left( \frac{11}{6} \right)} \right]^{5/6} \frac{\Gamma(11/6)}{25/6 \pi^{8/3}} \) and \(K_{\nu}(r)\) is the modified Bessel function of the second kind (aka. McDonald function).

From a computational point of view, the tel_phase_stat class implements both routines for OTF and PSF computation. To obtain SR and FWHM values, the class template perfEval is invoked on an object tel_phase_stat. The perfEval class template defines function members to compute SR and FWHM for any object implementing a PSF function member.

The figure 3 shows SR and FWHM as a function of the telescope diameter for three \(L_0\) values with a \(r_0\) of 1m at 2.2\(\mu\)m. When the diameter becomes much larger than \(L_0\), near diffraction limit images can be obtained but with a poor dynamic range as SR is only of few percent. This can be understood looking at Fig.4 showing the modal spectrum of the phase fluctuations expanded on a Karhunen–Loève basis. In the case of a finite \(L_0\) model, the energy distribution among the modes is affected by any change in the ratio \(D/L_0\). When this ratio increases, the total energy of the phase fluctuations decreases but the energy of the lowest orders modes, and mainly the first of all i.e. the tip-tilt mode, decrease faster compared to the energy of the higher modes. Consequently, for large \(D/L_0\) and \(r_0\), tip-tilt energy becomes so low that almost no image wandering is still visible in the focal plane of the instrument and the long exposure PSF core is maintained within a diffraction limited width.

Such peculiar effects can appear only for favorable seeing conditions. The table 1 shows the occurrence of simultaneous small \(L_0\) and large \(r_0\) in Paranal. These measurements were done during 19 nights in December 1999 by the Generalized Seeing Monitor\(^9\). These data give a snapshot of the site quality for ELTs and longer measurement time is required for a finer analysis. They show that an appropriate site for 30m up to 100m class ELTs should have \(L_0\) and \(r_0\) typically lower and larger than 25m and 15cm, respectively, to reach near diffraction limit images without AO correction during a reasonable amount of time.

3. MONTE-CARLO (END–TO–END) MODEL

3.1. Principle

The principle of this simulation is explained in details in Ref. 10. This simulation is based on phase screens which represent Kolmogorov (or von Karman) turbulence. Any number of infinitely thin screens can be included to represent the atmospheric turbulence.
Figure 3. Strehl ratio (right panel) and full width at half the maximum (left panel) plotted versus the telescope diameter for three outer scale 10m, 15m and 25m and a single seeing of 0.44" in K (2.2μm). On the left panel, the full lines give the diffraction limited and seeing limited FWHM bounds.

\[
\begin{align*}
    r_0 &\geq 15 \text{ cm} & 20 \text{ cm} \\
    L_0 &\leq 10 \text{ m} & 2.6 \% & 0.5 \% \\
    L_0 &\leq 15 \text{ m} & 8.3 \% & 1.8 \%
\end{align*}
\]

Table 1. Occurrence of small \( L_0 \) and large \( r_0 \) in Paranal during the 19 nights measurement campaign of the Generalized Seeing Monitor.

A wavefront sensor module makes a Fourier transform of some portion of the propagated screens and computes the measurements. A reconstructor matrix is then used to get the commands to be sent to the deformable mirror module, whose shape is computed. This shape is then subtracted from the atmospheric phase, which has evolved (according to the Taylor hypothesis) between the wavefront sensing step and the correction step.

3.2. Practical implementation

The implementation described in Ref. 10 was written in IDL (Interactive Data Language), a high level, interpreted language. This implementation is too slow and memory consuming to be able to simulate an ELT. Therefore, the work has begun to port this simulation into C/C++ and parallel processing (using the MPI library).

Two types of computational challenges are faced by the Monte-Carlo simulation:

- Storage of large arrays: these can be the phase screens which need to sample the Fried parameter on a large aperture, or the command matrices which are very large due to the large number of actuators on the ELT-AO system.

- Large number of computations, which are due to the complexity of the AO system and the large number of actuators and wavefront sensing elements.

The use of several computers allows to tackle both these problems. The arrays can be split between several machines, thus reducing the required memory per machine (allowing to use cheaper 32-bit architectures, i.e.
Figure 4. Modal spectrum of the wave front phase fluctuations when this one is expanded of a Karhunen–Loeve basis of 40.10^3 modes. Two kind of telescopes are compared: the VLT with a diameter of 8m and a central obscuration ratio of 0.14 and OWL with a diameter of 100m and a central obscuration ratio of 1/3. Full and dashed lines corresponds respectively to outer scales of 25m and 500m. The seeing has been fixed to 0.44” in K(2.2μm).

PCs, limited to 4 Gbytes of RAM per machine). The computational load per machine can also be reduced by dividing the task into several smaller pieces.

The linear algebra tasks are treated in this simulation using the Scalapack library, which allows standard linear algebra operations (like matrix multiply, matrix inversion and SVD) to be done on large (dense) distributed matrices. The use of a sparse matrix formulation (as proposed in Ref. 11) for the command matrix is possible, although not currently implemented.

The large number of computations can be easily broken up, since most operations are independent:

- Several independent WFSs are used, so they can be processed in parallel.
- All Fourier transforms made in the different wavefront sensors can be made independently (neglecting diffraction effects between sub-apertures).
- All phase screens can be propagated individually, before summation at the wavefront sensor conjugation height.
- The shape of the deformable mirrors can be computed independently.

*http://www.netlib.org/scalapack/
3.3. Preliminary results

As this contribution was being written, only some parts of the porting from IDL to C had been completed, like the parallel command matrix calculation (SVD and matrix-vector multiplication), and the Shack-Hartmann wavefront sensor FFT computation. However, some results were already available, as shown on Fig. 5 and Fig. 6. The first of these figures shows the shape of a $46 \times 46$ sub-aperture Shack-Hartmann wavefront-sensor, on a 30m telescope. In each sub-aperture, a fast Fourier transform has been made, to compute the shape and
position of the spot. Random noise (photon, read-out, dark, sky-background) was added in each sub-aperture. A total of six PC were used to compute the spot shapes shown in this figure, each PC doing a part of the FFTs.

The following figure shows the AO loop closing on a single conjugate, single guide star ("conventional") AO system, at 2.2 \( \mu \text{m} \), on a 30m telescope. Again, the computations were done in parallel, on 6 PCs. Although not scientifically very significant, this shows the first steps in our simulation effort.

4. HARDWARE DEFINITION

To be able to simulate a 30-50m telescope (full numerical simulation) and eventually a 100m telescope (analytical simulation), with an MCAO system in the infrared, we have estimated that a cluster made of about 20 PCs, each running at \( \sim 2 \) MHz should be sufficient.

A dual-CPU master machine serves as a file server and is used to launch and control the parallel application. Linux will be used as operating system on all machines.

To test this approach, a very small cluster made of 5 PC (running at 1.7 GHz) connected with a 100 Mbit/s network and a cheap switch was assembled. Different applications were run on it to get an estimate of the performance: parts of the analytical and numerical simulations, and network tests. This confirmed our estimations (see below) that a faster network was necessary, but that otherwise the idea of distributing for example control matrices to several machines to increase the processable matrix size was functioning.

To be completely sure that a Gigabit network will not introduce additional problems, this test cluster will be upgraded to Gigabit Ethernet. Tests will be run again and after successful completion, the cluster will be upgraded to the full amount of about 20 PCs.

It has been checked (although currently only on a 100 Mbit network) that latency should not dominate the transmission rate, since these arrays are large enough. This system will allow parallelization of the independent tasks (like Fourier Transforms) and distribution of the large arrays (phase screens, command matrices) onto several machines.

Therefore, simulating MCAO in K(2.2\( \mu \text{m} \)) on telescope up to 30-50m diameter should be reachable with this approach. It will allow to scan the parameters space for a large range of telescope size given the first scaling laws. The extrapolation of the latter will give first MCAO performances insights for OWL. For OWL, it will be able also to simulate full AO simulation in K and partial or simpler AO simulation in shorter wavelengths.

4.1. Dimensioning the network of the PC cluster for FFTs

The data are stored into a cube in which each plane is processed via a 2D-Fourier transform. For one plane, the computation time is:

\[
T_{\text{FFT}_{\text{W}}} = \alpha \cdot N \log_2 (N)
\]  

(8)

where \( \alpha \in \mathbb{R} \) and \( N \) is the total number of elements (pixels) in the plane computed. The value of \( \alpha \) has been determined experimentally for our cluster and is equal to \( 5 \cdot 10^{-8} \text{ s Mb}^{-1} \). Due to the independence between each plane of the cube, the 2D-Fourier transforms are distributed among the processors.

With only one processor, the computation time for the whole data cube is:

\[
T_{\text{FFT}_{\text{W}1}} = \alpha \cdot M \cdot N \log_2 (N)
\]

(9)

where \( M \) is the number of planes in the cube.

On several processors, if we suppose a master processor distributes the cube on the others, and gathers the cube after computations, the computation time is increased by the communications time:

\[
T_{\text{FFT}_{\text{W}_p}} = \alpha \cdot \left( \frac{M}{p} \right) \cdot N \log_2 (N) + 2 \cdot (p - 1) \cdot \left( \frac{M}{p} \right) \cdot \left( \frac{64 \cdot N}{D} \right)
\]

(10)
where \( p \) is the number of processors and \( D \) is the measured throughput of the network in Mb·s\(^{-1}\). The first part of the equation is the computation time on each processor done at the same time. The second part represents the communication time between the master processor and the \( p-1 \) slave ones with a throughput of \( D \) Mb·s\(^{-1}\). The matrix to send is an array of double with the size \( 8 \cdot 8 \cdot N \) bytes = 64·\( N \) bytes. There is a factor 2 because it is a send-receive communication between master and slaves.

The goal is to determine the values for the parameters \( p, D, N \) and \( M \) performing the most efficient computation with \( p \) processors:

\[
T_{FFT_W_p} < T_{FFT_W_1}
\]  
(11)

This equation can be written:

\[
\frac{(p-1)}{p} \cdot N \cdot M \left( \frac{128}{D} - \alpha \cdot \log_2(N) \right) < O
\]  
(12)

Amazingly, neither \( M \) nor \( p \) are relevant parameters for parallel computation efficiency. This equation is simply solved in:

\[
2^{\left( \frac{128}{\alpha D} \right)} < N
\]  
(13)

Whatever the throughput is, the computation will be faster in parallel if the planes contain enough elements. In other words, the slower the network is, the bigger the planes have to be in order for the computation time per processor to be much larger than the communication time.

With the current network (100 Mb·s\(^{-1}\)), the size of squared matrix must be at least 7200 x 7200. The typical matrix size for our simulations are 1024 x 1024 (PSF) or 32 x 32 (Shack-Hartmann subaperture spot) which correspond to a network throughput of 128 Mb·s\(^{-1}\) and 256 Mb·s\(^{-1}\) respectively. These results demonstrate the need for a network with a throughput of 1 Gb·s\(^{-1}\).

5. CONCLUSIONS

The state of the development of AO/MCAO simulations for ELTs has been presented and the strategy for OWL performance evaluation established. From a hardware point of view, the solution retained is a set of \( \sim 20 \) PCs forming a cluster inside a dedicated fast network of 1Gb·s\(^{-1}\). Intensive parallel computing will be used with these machines to overcome the problem of limited memory size on individual workstation and to reduce the computation time.

A C++ library has been written for the implementation of a fully analytical code. It is partially parallelized and the full parallelization is expected to be completed during the next months. This library is intended to be of general use for all the codes in development.

As first results, the parallelized routines have given birth to the first full basis of 100,10\(^3\) Karhunen–Loève modes. It has permitted also to show how ELT performances is affected by the outer scale of turbulence even when no AO correction is applied to phase corrugation. It has been shown that for favorable seeing conditions, almost diffraction limited images can be obtained with a 100m telescope without AO support. This is mostly due to the attenuation of the tip/tilt by more than a factor 100 when the outer scale goes from infinity to a value of 25m for a 100m telescope. The favorable seeing condition for such results will be able to appear \( \sim 8\% \) of the time in Paranal.

ACKNOWLEDGMENTS

This research has benefited from the support of the European Commission RTN program: ”Adaptive Optics for the Extremely Large Telescopes”, contract #HPRN-CT-2000-00147.

We would like to thank the ESO IT Department and in particular Giorgio Filippi who provided us with the majority of the computers composing the Beowulf cluster and supported us with his department; Stefano Turolla (ESO-IT) for his help in the cluster installation.
REFERENCES