

# Dynamic Calibration of Adaptive Optics Systems: A System Identification Approach

Alessandro Chiuso, Riccardo Muradore and Enrico Marchetti

**Abstract**—Adaptive optics systems are used in astronomy to obtain high resolution images of stars and galaxies with ground telescopes, otherwise blurred by atmospheric turbulence. In order to provide diffraction limited performance with ground-based telescopes, adaptive optics techniques use the measurements of one or more wavefront sensor to flatten distorted wavefronts with one or more deformable mirror in a feedback loop. In this paper we shall report our experience on the problem of building an accurate (dynamical) model of an adaptive optics system. This will be done adapting state-of-the-art system identification and model reduction techniques to the problem at hand. Our results are based on *real* data collected under various operating conditions from demonstrator developed at the European Southern Observatory (ESO). Data have been collected when the system was sitting at the ESO facilities in Garching, before being shipped to the Paranal observatory for the actual test on the field.

**Index Terms**—Adaptive Optics, Closed Loop Identification, Subspace Methods.

## I. INTRODUCTION

Adaptive optics (AO) is a recent engineering discipline arising from the interconnection of optics, electro-optics, electrical engineering, mechanical engineering and chemistry, [40]. All these disciplines are necessary to compensate for a kind of disturbance that is known since Newton time: the optical distortion produced by the Earth's atmosphere. Initially developed for military applications, this technique has been successfully used for several years in astronomy and has demonstrated that images with sharpness rivaling those of the Hubble Space Telescope can be obtained from the ground.

In order to provide diffraction limited performance with ground-based telescopes, AO techniques use the measurements of one or more wavefront sensor (WFS) to flatten distorted wavefronts with one or more deformable mirror (DM, from a control point of view the “actuator”) in a feedback loop [36], [40], [37], [18]. The cascade of these two components (DM + WFS) represents what in this paper will be called the *plant* or *system*.

In this paper we shall work with an adaptive optics system called Multi-Conjugate Adaptive Optics Demonstrator (MAD) [28], [27], which is a demonstrator instrument aiming

A. Chiuso is with the Dipartimento di Tecnica e Gestione dei Sistemi Industriali, Università di Padova (sede di Vicenza), stradella San Nicola, 3 - 36100 Vicenza, Italy. E-mail: chiuso@dei.unipd.it. The work of this author is partially supported by the national project *New techniques and applications of identification and adaptive control* funded by MIUR

R. Muradore is with the European Southern Observatory, Garching bei München, Germany. rmurador@eso.org

E. Marchetti is with the European Southern Observatory, Garching bei München, Germany. emarchet@eso.org

at performing large field of view atmospheric turbulence correction by implementing for the first time an adaptive optics technique called Multi-Conjugate Adaptive Optics (MCAO). MCAO and similar atmospheric turbulence correction techniques have been recognized as strategic both for the 2nd generation Very Large Telescope (VLT) instrumentation and for the European Extremely Large Telescope.

A high performance control system ought to be based on an (hopefully) accurate model of the system to be controlled. State-of-the-art systems, such as MAD considered in this paper, are based on a very simple static approximation of the system. The main motivation for using such a simple approximation of the plant is that the frequency response of the actuators is essentially constant in the frequency band of interest and that, by physical considerations, the wavefront sensor can be modeled as an averaging block on a sampling interval. Further considerations concerning modeling of an adaptive optics system can also be found in [25], [26], [32] where an optimal (minimum variance or LQG) approach to control design is taken. In any case the models considered, besides their static gain, are always based on physical considerations and are not built starting from data.

In this paper we shall report our experience on the problem of building an accurate (dynamical) model of an Adaptive Optics system. This paper can be seen as a continuation of our previous work [13], in which we tackled identification of a SCAO system (MACAO). In that work, however, we used data from a simulator and not from the real telescope.

Instead, in this paper the results are based on real data collected under various operating conditions from MAD. Data have been collected when MAD was sitting at the ESO facilities in Garching, before being shipped to the Paranal observatory (Chile) for the actual test on the field<sup>1</sup>.

In our view, the main contribution of this paper is to show that, indeed, a well designed dynamical model can outperform (in the sense of improving of about one order of magnitude in terms of simulation error variance) state-of-the-art static models which are now being used for control design. We also expect that the same consideration will become more and more important in the future; in fact, the next generation's AO systems will most likely have a control loop operating at higher frequency, making it even less reasonable to model the system with a static gain.

The models have been obtained using very recently de-

<sup>1</sup>Unfortunately it has not been possible to perform further tests on the real system after MAD has been re-mounted in the Paranal observatory.

veloped subspace techniques [11], which have also allowed identification to be performed in closed loop, and hence around the nominal operating conditions.

As a side contribution, we have also implemented a recursive version (see Section IV-A) of the  $PBSID_{opt}$  algorithm in [11], which has allowed to handle the large data sets used in this paper.

Finally, model reduction techniques have been applied to the models obtained from the identification experiment, allowing to obtain reasonably sized models, while maintaining good performance in validation.

Unfortunately we shall not be able to implement a controller based on this model in the near future, at least not until MAD will be shipped back to ESO for further experimentation.

The outline of the paper is as follows: Section II contains an introduction to the basic principles of adaptive optics and describes the main features of MAD; the main two components are the Deformable mirror and of the Wavefront sensor. The cascade of these two systems, as mentioned above, is the system to be modeled. Section III describes state-of-the-art AO controllers and also gives some details on how currently used models for the plant are being built. In Section IV we set the identification problem and describe the algorithm we use. This is followed, in Section IV-A, by the recursive implementation of the method and in Section IV-B by the description of how reduced order models have been obtained. Section V contains the experimental results and Section VI draws some conclusions.

## II. ADAPTIVE OPTICS: SCAO vs. MCAO

In the so called *single conjugated adaptive optics* (SCAO) systems the wavefront distortion is measured by one wavefront sensor (WFS). These measurements are used to control one deformable mirror (DM) which is used to compensate for the phase delays. Figure 1 shows the scheme of a SCAO system.

The field of view which benefits from the real-time atmospheric turbulence correction is very limited, typically few arcseconds for images obtained at infrared wavelengths. This limitation is coming from the fact that the distorted wavefront is estimated by the wavefront sensor only in the direction of a sufficiently bright guide star located at or nearby the observed astronomical object, and it is corrected for the same direction by a deformable mirror. In this configuration only the volume of the atmosphere probed by the light of the observed guide star is efficiently sensed while the atmospheric volumes probed by astronomical objects far from the guide star are only partially seen. The direct consequence of this mis-registration is that images of astronomical objects far from the guide star are only partially corrected, the amount of blurring increasing with the distance from the guide star. This phenomenon is called atmospheric anisoplanatism and a graphical representation is given in Figure 1.

To overcome these inconveniences, systems based on the so called *Multi-Conjugate Adaptive Optics* are being built. This technology allows to correct for atmospheric turbulence

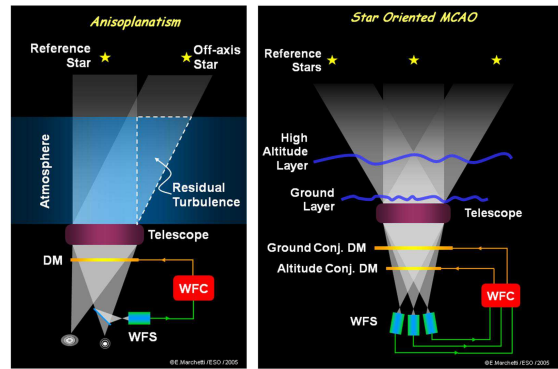


Fig. 1. Left: SCAO block diagram. Right: MCAO block diagram

on a field of view which is much larger than the one typically covered by the existing adaptive optics systems installed on 8-meter class telescopes.

ESO has developed a Multi-Conjugate Adaptive Optics Demonstrator (MAD) which, after a long period of testing at the ESO premises, has been installed on a Very Large Telescope (VLT) in the Paranal observatory on early 2007 for evaluating its correction performance. MAD is a MCAO systems with 2 Deformable mirrors, the first one conjugated at 8.5 km above the telescope and the second at the telescope pupil. Both deformable mirrors are spare units of the ones used in the MACAO family adaptive optics systems installed at VLT (bimorph mirror, 60 elements), [3].

The principle behind MCAO is described pictorially in Figure 1. MAD senses and corrects for the whole atmospheric volume probed by the observed field of view. The process of implementing MCAO correction consists of three main steps. The first one is to measure the deformation of the wavefront due to the atmospheric turbulence along different directions in the field of view. This is performed with several wavefront sensors each one looking at a different natural guide star in the field of view. The second step is called atmospheric tomography and consists in reconstructing the vertical distribution of the atmospheric turbulence at different locations of the field in order to obtain a three-dimensional mapping of the turbulence above the telescope, [34], [35]. The third step is to apply the wavefront correction to the whole field of view and not only at a specified direction. This is achievable by using several deformable mirrors which are optically conjugated at different altitudes in the atmosphere above the telescope.

The system is equipped with Multi Shack-Hartmann WFS consisting of 3 Shack-Hartmann units (SHU) with  $8 \times 8$  subapertures capable to scan the whole 2 arcmin FoV to pick-up the natural guide stars for the wavefront sensing.

The SH sensor provides the gradients of the 2D shape of the wavefront in a number of points along a cartesian pattern. The principle of the SH is extremely simple and is illustrated in figure 2. The main components of the SH sensor are a lenslet array located in a conjugated pupil plane and an array of detectors placed in the focal plane of each lens.

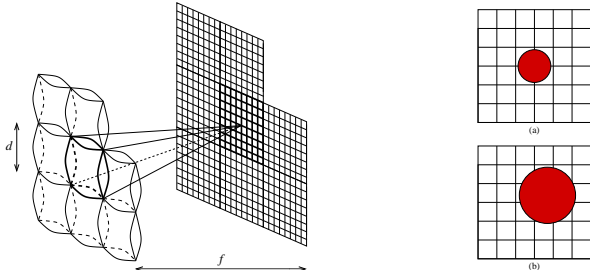


Fig. 2. Shack-Hartmann wave-front sensor principle.

The incoming wave-front is spatially sampled by the lenslet and each subaperture focuses the incident light in the focal plane. Thanks to the CCD detector located in the focal plane it is possible to measure the displacement using an algorithm computing the position of the center of gravity; in our case we have simply

$$s_x = \frac{\sum_{i,j} x_{i,j} I_{i,j}}{\sum_{i,j} I_{i,j}}, \quad s_y = \frac{\sum_{i,j} y_{i,j} I_{i,j}}{\sum_{i,j} I_{i,j}} \quad (1)$$

$x_{i,j}$  and  $y_{i,j}$  are weighting numbers related to the particular position of the pixel  $(i, j)$  and  $I_{i,j}$  is the corresponding intensity.

Besides the spatial sampling there is also a temporal sampling due to the integration time of the signal and the read-out of the CCD pixels (read-out delay). The latter has a duration comparable with the integration time. Since the actuators are very fast, it is easy to understand that the limitation in performance is mainly due to such delays.

Another important factor is the measurement noise which is essentially due to the quantum nature of photodetection. The photon noise can be well modeled with a Poisson statistics and is strictly related to the star magnitude  $m$  and the WFS integration time.

### III. STANDARD CALIBRATION

In this section we shall briefly review the control strategy implemented in MAD. In particular the focus will be on the calibration step and on its connection with the implemented control architecture. For ease of exposition, we will consider a SCAO system schematically described in figure 3. The control loop is composed by one WFS and one DM. The same procedure can be extended to the more general MCAO case, involving several WFSs and several DMs. In this simplified scheme we shall only consider the piezoelectric part of the mirror and assume orientation/tip/tilt have been taken care of.

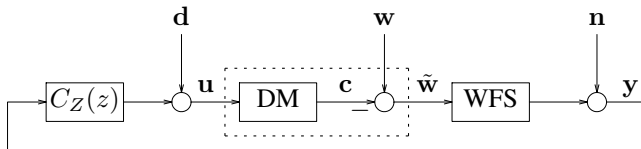


Fig. 3. Adaptive optics block diagram

The symbols in figure 3 have the following meaning:  $w$  is the incoming turbulence,  $c$  is the DM correction,  $\tilde{w}$  is the residual error,  $y$  is the measurement vector (sensed slopes) and  $n$  is the measurement noise. The signal  $u$  is the control input (i.e. the voltages applied to the DM) and  $d$  is an excitation source which is normally zero and has been added only to the purpose of identification, see Section IV.

The dimension of  $y \in \mathbb{R}^p$  is twice the number of illuminated subapertures on which the incoming turbulence is spatially sampled. The dimension of  $u \in \mathbb{R}^m$  is equal to the number of actuators of the deformable mirror. On the other hand, the quantities  $w$ ,  $c$  and  $e$  are continuous signals in time and space, with a finite spatial support equal to the pupil size. The controller  $C_Z(z)$  is usually composed by a set of  $m$  independent PI controllers

$$C_Z(z) = \text{diag} \left\{ K_p + \frac{K_i T_s}{1 - z^{-1}}, \dots, K_p + \frac{K_i T_s}{1 - z^{-1}} \right\} M_C. \quad (2)$$

Their set-points at time  $t$  are given by the product of the so-called control matrix  $M_C$  and the measurements  $y(t)$ . This type of architecture is called *zonal* control because the control matrix  $M_C$  is simply obtained by pseudo-inverting the so-called interaction matrix  $M_I$ , which provides a static model describing the link between applied voltages and sensed slopes. The control matrix  $M_C$  is used to reconstruct the required DM pattern at the next step from the current WFS measurements. The performance of the controller (2) is strictly related to the *quality* of  $M_C$  and hence to that of  $M_I$ .

Finding the matrix  $M_I$  is normally called *calibration*. This problem can be tackled in a number of ways: by pushing each actuator at a time, exciting the mirror so that its shape approximates one of the “mirror modes” or the modes of some pre-specified basis, (e.g. Zernike or Karhunen-Loève). For a deeper discussion about calibration, the reader is referred [31], [30], [17].

However, the most commonly used<sup>2</sup> procedure is based on exciting the mirror with a sequence of inputs which are orthogonal to each other; the input pattern is based on the so-called Hadamard matrix, [22]. The constant input is applied as long as the system reaches steady state. The (temporal) average of the steady state response if taken as output. This has an averaging effect on the noise which, together with the good numerical properties due to the orthogonal inputs, makes the approach very robust to noise. This procedure has been shown to be robust both on fiber and on sky calibration. Figure 4 shows the results obtained using this procedure.

The results of this paper provide an alternative procedure for calibration. In fact, once a dynamical model relating control inputs  $u$  to outputs  $y$  has been obtained via system identification, a static approximation can be found, e.g., by considering the DC gain of the identified model. Of course the static approximation can be seen as the zero-order approximation of the identified model. We shall discuss in Section IV-B the application of model reduction techniques

<sup>2</sup>This is also the one implemented in MAD.

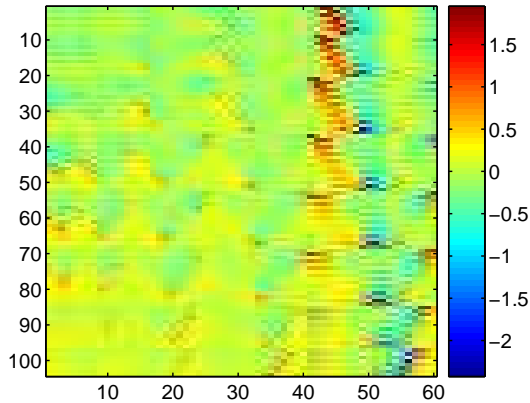


Fig. 4. Interaction Matrix

to this problem; the results are quite interesting, see Section V.

#### IV. CALIBRATION BY SUBSPACE IDENTIFICATION

As discussed previously, MAD has 2 Deformable Mirrors (conjugated to different altitudes) and 3 Wavefront sensors looking at 3 different guide stars. In this paper we address, for the Multi-Conjugate adaptive optics system, the identification of a dynamical model linking the control action to *one* DM to the slope measurements of *one* WFS<sup>3</sup>.

To this purpose the system has been operated “switching off” two WFS and actuating only one DM; the control command to the DM under test is computed using only measurements from the active WFS, as if it was a SCAO system, see Figure 1. Also the validation results reported in Section V are obtained in the same setup.

We shall refer to the scheme in Figure 3. Let us only recall that  $\mathbf{d}$  is a disturbance which has to be injected in order to guarantee identifiability. During normal operation of the plant  $\mathbf{d} = 0$ . It is therefore advisable to keep  $\mathbf{d}$  as small as possible also during identification in order to operate the plant around the nominal working conditions.

Note also that recent results [7] suggest that provided the controller is rich enough (as compared to the system dynamics) it is not necessary to excite all reference signals to obtain identifiability. However, it turns out that in this application the controller *is not* complex enough and in fact we experience problems when the signal  $\mathbf{d}$  is either zero or very small. Suffices here to say that the results from validation experiments on the real plant suggest that  $\mathbf{d}$  should be large enough to excite the systems and small enough not to excite non-linear effects.

One might wonder why one should dare to perform identification in closed loop, having to face the problem of

<sup>3</sup>We could have, of course, tried to model the overall system (i.e. 2 DM commands to 3 WFS measurements, 120 inputs, 312 outputs). However the identification problem is already challenging as it stands. An overall model can be obtained by combining the results of 6 identification experiments from all possible choices of DM’s and WFS’s.

identifiability and other difficulties which this setup implies. There are indeed a number of reasons for doing so, one of which we have just mentioned. In fact, the overall plant can be considered linear only in the vicinity of the operating conditions; in particular the wavefront sensor can be reasonably described by a linear model only provided it does not saturate, i.e. if the sensed slope is small; this is so if the DM shape ( $\mathbf{c}$ ) compensates the incoming phase distortion ( $\mathbf{w}$ ), which can only be achieved in closed loop.

To these considerations we might also add that sometimes closed-loop identification gives better results than open-loop [8] and that it could be desirable to perform identification in closed-loop when the model has to be used for the purpose of control design [19], [21].

The system is MIMO with 60 inputs and 104 outputs; therefore it is quite natural to use a subspace approach. Our aim here is not to provide an in-depth comparison of subspace methods applied to the AO system, but rather show that these methods provide an interesting alternative to the most commonly used calibration procedures [22] which are being normally used in the AO community.

The main issue, besides the presence of feedback, is the large number of inputs (60) and outputs (104); to the authors’ knowledge the  $\text{PBDSI}_{opt}$  algorithm described in [11], [10] is the simplest (in terms of computational load) subspace algorithm working also with closed-loop data. It turns out that the results in [11], [10], [12], [9] imply also that  $\text{PBDSI}_{opt}$  is (asymptotically) optimal within a certain class of subspace algorithms.

Moreover, in order to deal with the large amount of data we have implemented a recursive version of the algorithm which is based on the recursive computation of the VARX<sup>4</sup> parameters. In the next subsection we give some details of this algorithm; we refer the reader to the papers [4], [11], [10] for a thorough discussion of the literature on subspace methods as well as for the properties of the  $\text{PBSID}_{opt}$  algorithm.

Our purpose is to estimate the parameters  $(A, B, C, D, K)$  of a state space model

$$\begin{cases} \mathbf{x}(t+1) &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) + \mathbf{K}\mathbf{e}(t) \\ \mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t) + \mathbf{e}(t) \end{cases} \quad (3)$$

Since in this application it is known that there is a delay between inputs and outputs,  $D = 0$  from now on.

We shall also denote with  $F(z) := C(zI - A)^{-1}B$  the transfer function from  $\mathbf{u}$  (control inputs) to  $\mathbf{y}$  (sensed slopes).

In subspace identification two integers,  $p$  for “past” and  $f$  for “future”, have to be selected by the user. These are the number of block rows in certain block Hankel data matrices we shall encounter later on. The choice of these parameters is not entirely trivial, and in fact this is subject of current research [5], [6], [4], [33], [11], [12], [9]. We shall briefly discuss the choices we have made in Section V.

<sup>4</sup>Short for Vector Autoregressive with exogenous inputs.

### A. Recursive Implementation

In the application we are considering handling large data sets is quite demanding as both computation and memory requirements are concerned. Suffices here to recall that the first step in the PBSID<sub>opt</sub> algorithm is to estimate a VARX model

$$y(t) = \sum_{i=1}^p \Phi_i z(t-i) + e(t) \quad (4)$$

where  $z(t) = [y^\top(t); u^\top(t)]^\top$ . In our situation we have verified, see the discussion in Section V, that a reasonable choice for the “past” ( $p$ ) and “future” ( $f$ ) lengths are respectively  $p = 9$  and  $f = 7$ . Since the number of inputs and outputs is respectively 60 and 104, the number of parameters in the VARX model (4) is  $104 \times (60 + 104)9 = 104 \times 1476$ . This means that the matrix containing the past data has dimension  $1476 \times N$ . This requires an huge amount of memory space to be stored. Therefore we have implemented a recursive estimator for the VARX coefficients (4). Let us note that one has to start with a minimum number of data  $N_0$  so that the sample covariance of the past data is positive definite. In our case this means that  $N_0$  need to be certainly larger than 1476. The recursive estimator of  $\Phi_i$  in (4) is based on the Recursive Least Squares (RLS) algorithm (see for instance [24], [38]) and updates the estimator of the VARX coefficients as new data becomes available.

**Remark IV.1** Even though in our application the algorithm is not meant for on-line identification, the recursive version could be adapted for identification of slowly time varying systems by implementing the RLS algorithm using a forgetting factor [38].  $\diamond$

In order to provide a fully recursive implementation of the PBSID<sub>opt</sub> algorithm, besides using RLS in the VARX estimation step, it is also necessary to implement both (i) the SVD and (ii) the estimation of  $A, B, C, K$  from the state sequences in a recursive manner.

### B. Reduced Order Models

As can be seen from the results reported in Section V, in order to obtain good performance, the identified model order should be in the range [100, 140]. It turns out that this is a rather large model which might impose limitations to the purpose of control design.

Therefore we have explored the possibility of using reduced order models. As is well known, balanced model reduction is based on the state space model of  $\hat{F}(z) = \hat{C}(zI - \hat{A})^{-1}\hat{B}$ . Let us assume that  $\hat{F}(z)$  has Mac Millan degree  $n$  and assume we want to find a reduced order model of Mac Millan degree  $n_r$ . The so called balanced truncation is based on the idea that states have been sorted in order of importance, which is measured by their degree of observability and controllability. Balanced truncation simply neglects states which are at the same time “little observable” and “little controllable”, defining

$$F_r(z) := \hat{C}_1(zI - \hat{A}_{11})^{-1}\hat{B}_1.$$

For more details concerning the errors involved in this approximation we refer the reader to the paper [2].

A first drawback of this approach is that the steady state (DC) gain of the model is not preserved. However a simple modification of balanced truncation allows to preserve the DC gain of the original model. This is obtained by observing that, for constant input and at steady state, the state is constant. Let us consider the partition  $\mathbf{x} := [\mathbf{x}_1^\top, \mathbf{x}_2^\top]^\top$ ; the first component is the one to be retained (the most “observable and controllable”) while the second is to be eliminated.

However, while balanced truncation simply neglects the second component, here we observe that, at steady state and for constant input,  $\mathbf{x}_2(t) = \mathbf{x}_2(t+1)$ . One can therefore “solve” for  $\mathbf{x}_2(t)$

$$\mathbf{x}_2(t) = (I - \hat{A}_{12})^{-1} \left( \hat{A}_{11}\mathbf{x}_1(t) + \hat{B}_2\mathbf{u}(t) \right)$$

and retain only the  $\mathbf{x}_1(t)$  component of the state by substituting the equation above for  $\mathbf{x}_2(t)$ . This guarantees that the DC gain is the same as that of the original model. We shall call this procedure *DC-preserving balanced model reduction*.

Since in this application the DC gain plays a crucial role, and indeed, state-of-the-art control algorithms use a static approximation of the plant, we shall mainly consider this latter technique. However, in our experience, as long as the model order is large (say larger than 60) there are no major differences in the validation results using the two reduction approaches.

Note also that the identified model has, by construction, one delay (i.e.  $\hat{D} = 0$ ). This implies that the input-output relation can be written in the form

$$\mathbf{y}(t) = \hat{F}(z)\mathbf{u}(t) = \hat{F}_1(z)\mathbf{u}(t-1)$$

where  $\hat{F}_1(z)$  is such that  $\hat{F}(z) = z^{-1}\hat{F}_1(z)$ . However, the reduced order model obtained from the DC-preserving balanced model reduction does not have a delay; in fact  $\hat{D}_r := (I - \hat{A}_{12})^{-1}\hat{B}_2 \neq 0$  in general.

The validation results (see in particular Figure 8) suggest that it is indeed advantageous to perform model reduction of  $\hat{F}_1(z)$  and then “add” a delay, so that also the reduced order model has a strict delay. Of course, if we denote with  $\hat{F}_{1r}(z)$  the (order  $n_r$ ) reduced model of  $\hat{F}_1(z)$ , the transfer function  $z^{-1}\hat{F}_{1r}(z)$  has Mc Millan degree  $n_r + m$ , where  $m$  is the number of inputs<sup>5</sup>.

However the delay can simply be implemented with a memory buffer and reflects the presence of a physical delay in the loop. We have indeed verified that this latter strategy allows to obtain reduced order model of significantly lower order<sup>6</sup> with respect to  $\hat{F}_r(z)$ , i.e. the model obtained from direct model reduction of  $\hat{F}(z)$ .

## V. RESULTS

We have carried out several identification experiments using *real data* operating the system in closed-loop, under the

<sup>5</sup>Recall that there are less inputs than outputs.

<sup>6</sup>If one neglects the memory required to implement the physical delay.

so-called *calibration on beacon* setup, i.e. when the mirror is illuminated with a constant light (no incoming atmospheric aberrations).

The identification was carried out using  $\bar{N} = 40000$  data points. The length of the past horizon  $p$  has been chosen using the Akaike (AIC) [1] criterion. It can be shown that (see [16], Remark 3.6, [23], page 93, end of Section 2 and [20]) if  $\hat{p}$  is the minimizer of the AIC criterion,  $p = M\hat{p}$  with  $M > 1$  is a possible choice.

The AIC criterion is reported in Figure 5, from which it is clear that  $\hat{p} = 7$  is a minimizer. For  $M = 2$  this would lead to an order selection rule  $p = 2\hat{p} = 14$ . It turns out, however, that while this choice meets some asymptotic criteria as discussed in [23] it is not necessarily the best choice for finite data length  $N$  (see also Remark 3.6 in [16]). We have indeed verified in our experiments that  $p = 9$  or  $p = 10$  seem to provide the best results while also reducing the computational complexity (w.r.t. the choice  $p = 14$ ).

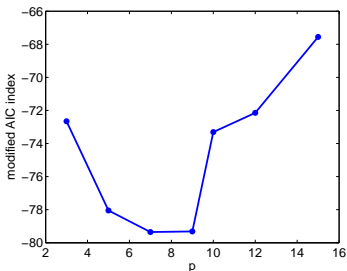


Fig. 5. AIC Criterion

As far as the choice of  $f$ , as discussed in [9], it only makes sense to consider  $f \leq p$ .

Unfortunately, since the plant is controlled using an integral controller, the sharp results presented in [9] (stating that under suitable assumptions the asymptotic variance of the estimators is monotonically decreasing in  $f$ ) cannot be applied. Therefore we have adopted a validation based approach to the choice of  $f$ . Several values of  $f$  (ranging from 3 to 9) have been tested. It turns out that  $f \geq 7$  seemed to provide the best results in terms of simulation error on validation data.

Several models, with orders ranging in the interval  $n \in [60, 140]$ , have been estimated. Also the identification experiments have been repeated with two different levels of input injection (the signal  $\mathbf{d}$  in Figure 3); the signal  $\mathbf{d}$  is white noise with standard deviations, respectively 0.07 and 0.04.

All identified models have been validated, according to the validation measure:

$$J = \frac{1}{N} \sum_{t=1}^N \text{Tr} \left[ (y_s(t) - y(t)) (y_s(t) - y(t))^{\top} \right] \quad (5)$$

where  $y_s$  is the simulated output using the identified model while  $y$  is the output from the real plant. Of course  $J$  has been computed on real data *not used* in the identification experiment.

Also the validation experiments have been carried out using three different levels of input excitation, namely with standard deviation 0.07, 0.04 and 0.016 (see figures 6, and 7). All experiments (identification and validation) have been carried out with a sampling frequency of  $200Hz$ .

As anticipated in Section III, it is a standard practice in AO systems to model the input-output behavior using a static gain (Interaction Matrix, IM) computed as discussed in Section III. This is justified by the fact that the system should be well approximated by a constant gain (and a delay) in the frequency range of interest. The results of this paper seem to suggest that this is not the case; in fact the dynamic model does significantly improve in terms of simulation error.

In our validation experiments we therefore compare the validation results using: a) the dynamic model we have identified, b) the standard static model, i.e. the IM computed using standard techniques, c) the static approximation of the identified model, i.e. its DC gain, d) a reduced order model<sup>7</sup> obtained as discussed in Section IV-B.

The validation results can be summarized as follows:

- The dynamic model outperforms the static model in essentially all situations. In fact the validation results show that an appropriate dynamical model performs about 40 times better (in terms of normalized error variance) w.r.t. the static model; this may suggest that a control design which uses the dynamic model may compare favorably with standard techniques based on static linear approximations (using the IM).
- A reduced order model of order  $n_r$  (obtained using balanced truncation) performs as good as the full order identified model provided  $n_r \geq 60$ . See Figure 8 (left). This may be very useful for the purpose of control design. Note in fact that currently used PI controllers require a memory which has the same dimension as the input space, i.e. 60. Hence a model-based control scheme could be implemented with about the same order of complexity as the standard PI controller.
- The static model obtained from the DC gain of the (dynamic) identified model is comparable in terms of simulation error w.r.t. the static model obtained using the standard approach for the computation of the IM.
- The dynamic models identified with large input excitation (standard deviation 0.07) perform slightly worse than those with standard deviation 0.04, suggesting that, perhaps, non-linear effects come into play. In particular larger order models are necessary to reach the same performance. Compare, for instance, the plots on the bottom left corners in Figures 6 and 7.

A typical pole map of an identified system is reported in Figure 8 (right). The blue crosses refer to an identified system of order 100 while the red crosses are the poles of

<sup>7</sup>Note that a reduced order model of order  $n_r$ , obtained from an identified model of order  $n$ , does not necessarily give the same results as a model of order  $n_r$  obtained directly from identification. It is actually well known that under certain circumstances identifying a large order model and the perform model reduction is to be preferred. See e.g. [39] for the case of  $L_2$  model reduction.

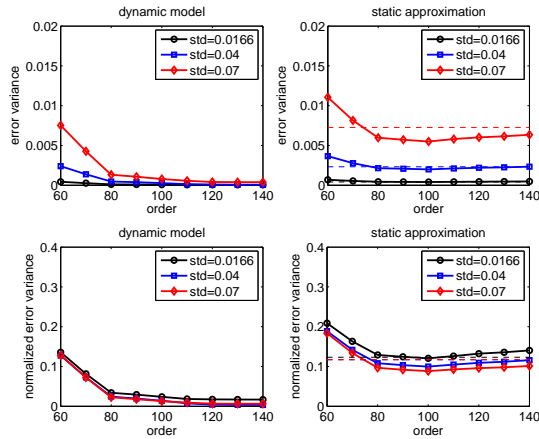


Fig. 6. Validation results (simulation error). Left hand side: using the identified model. Right hand side: using the static approximation (DC gain) and the IM (horizontal line). Injected noise  $\mathbf{d}$  with standard deviation 0.04. The three curves refer to validation data with different level of injected disturbance  $\mathbf{d}$  (0.07, 0.04, 0.016 respectively).

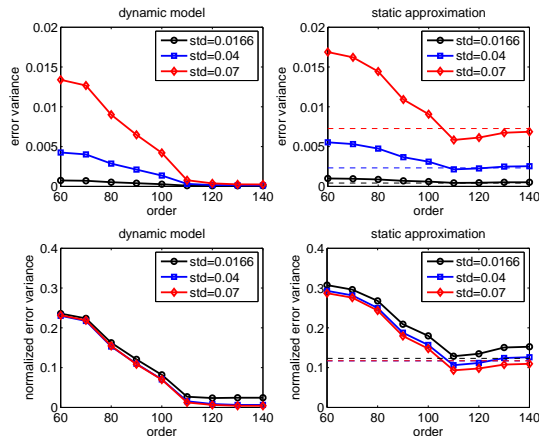


Fig. 7. Validation results (simulation error). Left hand side: using the identified model. Right hand side: using the static approximation (DC gain) and the IM (horizontal line). Injected noise  $\mathbf{d}$  with standard deviation 0.07. The three curves refer to validation data with different level of injected disturbance  $\mathbf{d}$  (0.07, 0.04, 0.016 respectively).

the reduced order system of order  $n_r = 60$  obtained by DC-preserving balanced model reduction of  $\hat{F}_1(z)$  (see Section IV-B).

In order to further support the dynamic model against the static model, we have analyzed the spectral content of the simulation errors  $e_s(t) := y_s(t) - y(t)$  for both the simulation using the static and dynamic models.

The signal  $e_s(t)$  has 104 components (the number of sensed slopes); we have computed the power spectral density (PSD) of its components. For reasons of space we report the power spectrum of the error only for two components; in particular we have selected the two slopes which yield, respectively, the maximum and minimum simulation error (normalized). It turns out the these are respectively slope

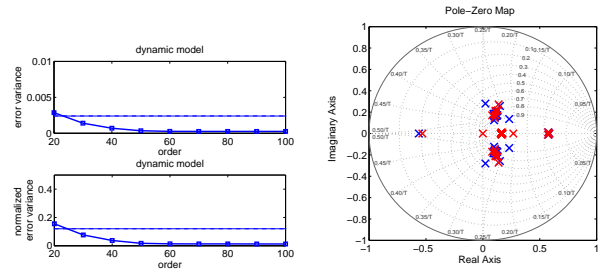


Fig. 8. Left: Simulation error variance. Reduced order model (starting from identified model of order 100) v.s. Interaction Matrix. Right: Pole map. Blue crosses: identified system of order 100. Red crosses: reduced order system of order  $n_r = 60$

#18 and #19.

Some representative results are reported in figure 9; it is clear that the frequency content of the error using the static models (both the IM computed using the standard tools in AO and the DC gain of the identified model) has a sharp peak in high frequency. It is also remarkable that, while the peak is still present using the dynamic model, its height and also its width are considerably reduced.

Note also that both the dynamic model and its static approximation (DC gain) perform better than the IM for low to medium frequency range.

It should also be stressed that these results have been obtained when the input is perturbed with a white noise process. Instead, during normal operating conditions, the “external excitation” is provided by the incoming turbulence. It is a well known fact that this perturbation has a “low-pass” spectrum [29], [15] and hence excites only the systems in the low to medium frequency range. It is therefore quite natural to expect that a dynamic model which performs better in that frequency range is to be preferred. Unfortunately, the data available does not allow to perform these sort of tests and, as mentioned in the introduction, we shall not be able to collect further data until the system will be operating in the Paranal observatory.

## VI. CONCLUSIONS

In this paper we have reported our experience on the problem of modeling an Adaptive Optics system. We have tackled the problem applying state-of-the-art subspace identification techniques to real data taken from a Multi-Conjugate Adaptive Optics Demonstrator developed at ESO. In order to be able to handle the large dimension of the data set we have developed a recursive version of the PBSID<sub>opt</sub> algorithm [11]; the recursive version is interesting per se and, with minor modifications, can also be used to tackle identification of slowly time-varying systems. The validation results are very encouraging, showing that the identified model outperforms currently used (static) models. We have also applied model reduction techniques, which have allowed to obtain reasonably sized models with very little loss in term of performance on validation data; more details can be found in a longer version of this work [14].

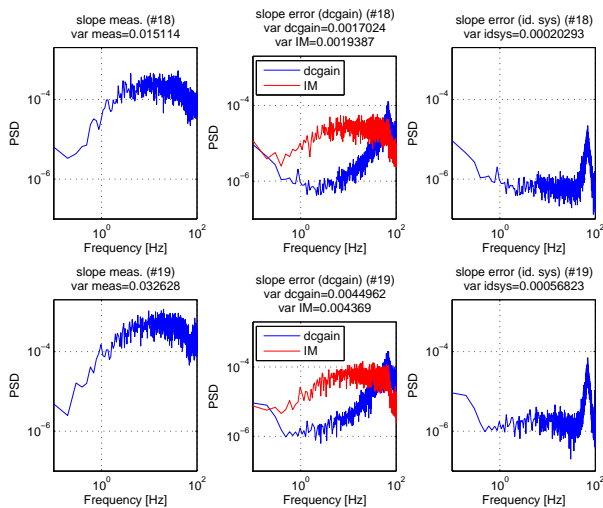


Fig. 9. Validation 200Hz. Power spectral density of the simulation errors, slopes #18 and #19. Left: output signal. Center: error using the static model. Right: error using the dynamic model.

Future work will include exploring the possibility of using such models for the purpose of control design. As mentioned above, it has not been possible to perform such test since, to date, the system is not available for further testing.

## REFERENCES

- [1] H. Akaike. Fitting autoregressive models for prediction. *Annals of the Institute of Statistical Mathematics*, 21:243–247, 1969.
- [2] U.M. Al-Saggaf and G.F. Franklin. An error bound for a discrete reduced order model of linear multivariable system. *IEEE Trans. on Aut. Control*, 32(9):815–819, 1987.
- [3] R. Arsenault, R. Donaldson, C. Dupuy, E. Fedrigo, N.N. Hubin, L. Ivanescu, M.E. Kasper, S. Oberti, J. Paufigue, S. Rossi, et al. MACAO-VLTI adaptive optics systems performance. *Proceedings of SPIE*, 5490:47–58, 2004.
- [4] D. Bauer. Asymptotic properties of subspace estimators. *Automatica*, 41:359–376, 2005.
- [5] D. Bauer and M. Jansson. Analysis of the asymptotic properties of the MOESP type of subspace algorithms. *Automatica*, 36:497–509, 2000.
- [6] D. Bauer and L. Ljung. Some facts about the choice of the weighting matrices in Larimore type of subspace algorithm. *Automatica*, 38:763–773, 2002.
- [7] A.S. Bazanella, M. Gevers, and L. Miskovic. Closed-loop identification of MIMO systems: A new look at identifiability and experiment design. In *Proc. of European Control Conference*, 2007.
- [8] X. Bombois, Gevers M, and G. Scorletti. Open-loop versus closed-loop identification of box-jenkins models: A new variance analysis. In *Proc. 44th IEEE Conf. on Dec. & Control*, Seville, Spain, 2005.
- [9] A. Chiuso. Choosing the future horizon in closed-loop CCA-type subspace algorithms. Technical report, University of Padova, 2007. Submitted to IEEE Trans. on Aut. Control, available at: <http://www.dei.unipd.it/~chiuso>.
- [10] A. Chiuso. On the relation between CCA and predictor-based subspace identification. *IEEE Trans. on Automatic Control*, 52(10):1795–1812, October 2007.
- [11] A. Chiuso. The role of Vector AutoRegressive modeling in predictor based subspace identification. *Automatica*, 43(6):1034–1048, June 2007.
- [12] A. Chiuso. Some insights on the choice of the future horizon in CCA-type subspace algorithms. In *Proc. of ACC*, 2007.
- [13] A. Chiuso, R. Muradore, and E. Fedrigo. Adaptive optics systems: A challenge for closed loop subspace identification. In *Proc. of ACC*, 2007.
- [14] A. Chiuso, R. Muradore, and E. Marchetti. Dynamic calibration of adaptive optics systems: A system identification approach. Technical report, Univ. of Padova, 2008. Submitted to IEEE Trans. on Contr. System Tech., available at [www.dei.unipd.it/~chiuso](http://www.dei.unipd.it/~chiuso).
- [15] J.M. Conan, G. Rousset, and P.Y. Madec. Wave-front temporal spectra in high-resolution imaging through turbulence. *J. Opt. Soc. Am. A*, 12(7):1559–1570, 1995.
- [16] A. Dahlén and W. Scherrer. The relation of CCA subspace method to a balanced reduction of an autoregressive model. *Journal of Econometrics*, 118(1-2):293–312, 2004.
- [17] S. Esposito, R. Tubbs, A. Puglisi, S. Oberti, A. Tozzi, M. Xompero, and D. Zanotti. High SNR measurement of interaction matrix on-sky and in lab. *Proceedings of SPIE*, 6272:62721C, 2006.
- [18] E. Fedrigo, M. Kasper, L. Ivanescu, and H. Bonnet. Real-time control of ESO adaptive optics systems. *Automatisierungstechnik*, 53(10):470–483, 2005.
- [19] M. Gevers and L. Ljung. Optimal experiment design with respect to the intended model application. *Automatica*, 22:543–554, 1986.
- [20] E.J. Hannan and M. Deistler. *The Statistical Theory of Linear Systems*. Wiley, 1988.
- [21] H. Hjalmarsson. From experiment design to closed-loop control. *Automatica*, 41(3):393–438, 2005.
- [22] M. Kasper, E. Fedrigo, D. P. Looze, H. Bonnet, L. Ivanescu, and S. Oberti. Fast calibration of high-order adaptive optics systems. *J. Opt. Soc. Am. A*, 21(6):1004–1008, 2004.
- [23] G.M. Kuersteiner. Automatic inference for infinite order vector autoregressions. *Econometric Theory*, 21:85–115, 2005.
- [24] L. Ljung. *System Identification, Theory for the User*. Prentice Hall, 1997.
- [25] D.P. Looze. Minimum variance control structure for adaptive optics systems. *Journal of the Optical Society of America A*, 23(3):603–612, 2006.
- [26] D.P. Looze, M. Kasper, S. Hippler, O. Beker, and R. Weiss. Optimal Compensation and Implementation for Adaptive Optics Systems. *Experimental Astronomy*, 15(2):67–88, 2003.
- [27] E. Marchetti, R. Brast, B. Delabre, R. Donaldson, E. Fedrigo, C. Frank, N. Hubin, J. Kolb, M. Le Louarn, J.L. Lizon, et al. MAD: practical implementation of MCAO concepts. *Comptes rendus-Physique*, 6(10):1118–1128, 2005.
- [28] E. Marchetti, R. Brast, B. Delabre, R. Donaldson, E. Fedrigo, C. Frank, N. Hubin, J. Kolb, M. Le Louarn, J.L. Lizon, et al. MAD star oriented: laboratory results for ground layer and multi-conjugate adaptive optics. *Proceedings of SPIE*, 6272:627200, 2006.
- [29] R.J. Noll. Zernike polynomials and atmospheric turbulence. *JOSA*, 66(3):207–211, 1976.
- [30] S. Oberti, H. Bonnet, E. Fedrigo, L. Ivanescu, M.E. Kasper, and J. Paufigue. Calibration of a curvature sensor/bimorph mirror AO system: interaction matrix measurement on MACAO systems. *Proceedings of SPIE*, 5490:139–150, 2004.
- [31] S. Oberti, F. Quirós-Pacheco, S. Esposito, R. Muradore, R. Arsenault, E. Fedrigo, M. Kasper, J. Kolb, E. Marchetti, A. Riccardi, et al. Large DM AO systems: synthetic IM or calibration on sky? *Proceedings of SPIE*, 6272:627220, 2006.
- [32] R.N. Paschall and D.J. Anderson. Linear quadratic Gaussian control of a deformable mirror adaptive optics system with time-delayed measurements. *APPLIED OPTICS*, 32(31/1), 1993.
- [33] S.J. Qin and L. Ljung. On the role of future horizon in closed-loop subspace identification. In *Proc. of SYSID 2006*, Newcastle, Australia, 2006.
- [34] R. Ragazzoni, E. Marchetti, and F. Rigaut. Modal tomography for adaptive optics. *Astronomy and Astrophysics*, 342:53–56, 1999.
- [35] R. Ragazzoni, E. Marchetti, and G. Valente. Adaptive-optics corrections available for the whole sky. *Nature*, 403(6765):54–6, 2000.
- [36] F. Roddier. *Adaptive Optics in Astronomy*. Cambridge University Press, 1999.
- [37] M.C. Roggemann and B. Welsh. *Imaging through Turbulence*. CRC Press, 1996.
- [38] T. Söderström and P. Stoica. *System Identification*. Prentice-Hall, 1989.
- [39] F. Tjörnström. Variance analysis of  $L_2$  model reduction when undermodeling: the output error case. *Automatica*, 39:1809–1815, 2003.
- [40] R.K. Tyson. *Principles of Adaptive Optics*. Academic Press, 2nd edition, 1998.