A practical approach to modal basis selection and wavefront estimation

M. E. Kasper[a], D. P. Looze[b], S. Hippler[a], M. Feldt[a], A. R. Weiß[a], A. Glindemann[c], R. Davies[d]

[a]Max-Planck-Institut für Astronomie, Heidelberg, Germany
[b]University of Massachusetts, Amherst, USA
[c]European Southern Observatory, Garching, Germany
[d]Max-Planck-Institut für Extraterrestrische Physik, Garching, Germany

ABSTRACT

The MPIA/MPE adaptive optics with a laser guide star system ALFA works excellent with natural guide stars up to 13th magnitude in R-band. Using fainter natural guide stars or the extended laser guide star, ALFA's performance does not entirely satisfy our expectations. We describe our efforts in optimizing the wavefront estimation process.
Starting with a detailed system analysis, this paper will show how to construct a modal basis set which efficiently uses Shack-Hartmann measurements while keeping a certain number of low order modes close to analytical basis sets like Zernikes or Karhunen-Loève functions. We will also introduce various phase estimators (least squares, weighted least squares, maximum a posteriori) and show how these can be applied to the ALFA AO. A first test done at the Calar Alto 3.5-m-telescope will be discussed.

Keywords: Modal control, Shack-Hartmann sensor, wavefront reconstruction

1. THE ALFA CONTROL SYSTEM

ALFA uses feedback or closed-loop control to flatten distorted wavefronts in real-time. In a feedback system, the variable being controlled is measured by a sensor, and the information is fed back into the process to influence the controlled variable. The key components of an AO system are the wavefront sensor which measures the state of the incident wavefront (controlled variable), the control computer which interprets the measurement and calculates the feedback signal, and the deformable mirror which affects reflected wavefronts by continuously adjusting its shape. In the following we describe the specific types implemented in ALFA.

Deformable Mirror

The DM was purchased from Xinetics Inc., USA. Its 97 PMN (lead magnesium niobate) ceramic piezo actuators are capable of delivering a 2 μm inter-actuator stroke, and they reach their commanded positions in well below 1 ms. The actuators are operated in their linear response region between 35 and 105 Volts.

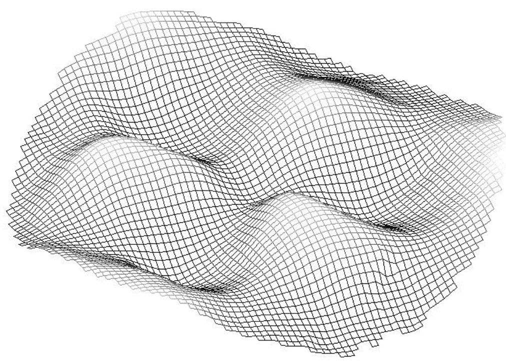

Figure 1. DM surface with a Karhunen-Loève mode of 5th radial order applied.

Approximately 15 actuators lie too far outside the pupil annulus to influence the reflected wave. The DM has a diameter of 80 mm, and the distance between two adjacent actuators is 7 mm corresponding to 35 cm on the telescope's primary.
The rms aberration of the reflected wave is about 600 nm when the same voltage is applied to all actuators. The actuators show very little hysteresis, and the configuration remains the same when the temperature is stable. An automated procedure which uses a Twyman-Green interferometer to monitor the DM surface flattens the mirror to approximately 50 nm rms within a few minutes.

Shack-Hartmann waveront sensor
In order to react on varying seeing conditions and guide star magnitudes, 4 different lenslet arrays are mounted on a motorized stage (Figure 2, one array is not shown). The *D3 array* (3 subapertures per pupil diameter) provides 6 subapertures to sense the wavefront slopes, with the central lens obscured by the secondary telescope mirror. This array is used for guide stars fainter than 13th magnitude as well as a faint LGS. Two *D5 arrays* with different focal lengths provide 18 subapertures each. The shorter focal length one produces smaller images and spot motions on the CCD and is used in poorer seeing conditions. The D5 arrays are used for guide stars around 12th magnitude and a bright LGS. *The D7 array* provides 30 subapertures and is the choice for guide stars of 11th magnitude and brighter.

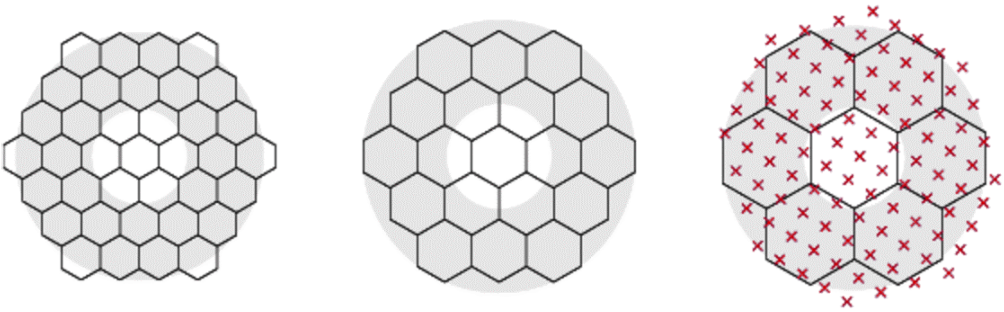

Figure 2. Microlens configurations of the ALFA Shack-Hartmann sensor. The gray annulus resembles the telescope pupil of the 3.5-m-telescope. From left to right: D7, D5, and D3 array. The crosses are the positions of the 97 actuators

The camera detector is a 64 × 64 pixels thinned Lincoln Labs CCD that is used in frame transfer mode. The maximum frame rate is 1206 Hz with a pixel clock of 1.8 MHz and a read-out noise of about 4 e^-. The quantum efficiency between 600 nm and 750 nm exceeds 80%. The camera has a two stage thermo-electric cooler providing -35°C which reduces dark currents. Thermal energy is transported out of the camera via a heat-pipe. Read-out electronics and digital interface to the control system were built by AOA[1].

Control Computer
All matrix operations, such as determination of the spot locations and reconstruction of the modal coefficients, are implemented in the VME-bus based real-time computer (RTC) system using five digital signal processor (DSP) boards. Each board holds four DSP chips for a total of 20 DSPs between which the required calculations are distributed. This implementation is capable of projecting 50 wavefront control modes in 833 μs execution time (corresponding to the maximum loop frequency of 1206 Hz). A more detailed presentation of the computer system is given by Hippler et al. (1998).

2. SYSTEM ANALYSIS

Although the number of actuators is fixed, the number of subapertures driving them may be varied from 6 to 30 giving a maximum of 60 Shack-Hartmann gradients. This is less than the number of actuators, and thus limits ALFA's control space. The coarse spatial sampling of the wavefront by the microlens arrays does not allow us to sense the phase at the positions of all actuators (zonal control). Thus, some interpolation scheme must be used to derive the actuator commands.
In the *modal control* approach, the wavefront phase $\phi(x,y,t)$ is approximated by a superposition of a collection of shapes that the DM can form, hereafter called the *control modes* $M_i(x,y)$, i.e.

$$\phi(x,y,t) \approx \sum_{i=1}^{n} a_i(t) M_i(x,y). \tag{1}$$

[1] Adaptive Optics Associates, Cambridge, Massachusetts, USA

Equality in this equation could only be reached for an infinite number of control modes ($n = \infty$).
The estimation of the modal coefficients **a** based on the Shack-Hartmann gradients (wavefront slopes) **g** is the subject of wavefront reconstruction. The measurement process can be described by the following linear relationship

$$\mathbf{g} = \mathbf{D}\mathbf{a} + \mathbf{D}_\perp \mathbf{a}_\perp + \mathbf{n}. \tag{2}$$

The columns of the interaction matrix **D** contain the gradients originating from the control modes and are obtained during the calibration procedure (for a complete description see Kasper et al., 2000). **g** is further affected by gradients from the modes which complement the control modes to infinite dimension as described their interaction matrix $\mathbf{D}_\perp$, weighted by their coefficients $\mathbf{a}_\perp$. Finally, the measurement noise **n** is added to the gradients.
Given a reconstruction matrix **R**, the estimate for the control mode coefficients can be written as

$$\hat{\mathbf{a}} = \mathbf{R}\mathbf{g} = \mathbf{R}\mathbf{D}\mathbf{a} + \mathbf{R}\mathbf{D}_\perp \mathbf{a}_\perp + \mathbf{R}\mathbf{n}. \tag{3}$$

Using the most common least-squares estimate (**R** is the pseudo-inverse of **D, RD** = **I**d), and assuming uncorrelated noise, the reconstruction error variance for the *i*th mode can be written as (Dai, 1996)

$$\left\langle \left| a_i - \hat{a}_i \right|^2 \right\rangle = \left(\mathbf{C} \left\langle \mathbf{a}_\perp \mathbf{a}_\perp^t \right\rangle \mathbf{C}^t \right)_{ii} + \left(\mathbf{R} \left\langle \mathbf{n}\mathbf{n}^t \right\rangle \mathbf{R}^t \right)_{ii}, \tag{4}$$

with $\mathbf{C} := \mathbf{R}\,\mathbf{D}_\perp$, and <.> denoting the ensemble average.

Modal cross-talk
Cross-talk occurs, if some of the column vectors of **D** are almost colinear, or very small compared to others (Herrmann, 1981). The gradient produced by one of these mode can then easily be shifted into another by adding small **n**, which dramatically increases the sensitivity of phase estimation to noise. The liability of **D** to cross-talk is given by its condition number, which is defined as the ratio of the largest to the smallest singular value. Figure 3 shows the condition numbers of **D** for Zernike polynomials and Karhunen-Loève (KL) functions measured by the ALFA microlens arrays. In general, more pure KL than Zernike modes can be controlled. Our experience with ALFA is such that the condition numbers should be kept below 5 in order to ensure loop stability.

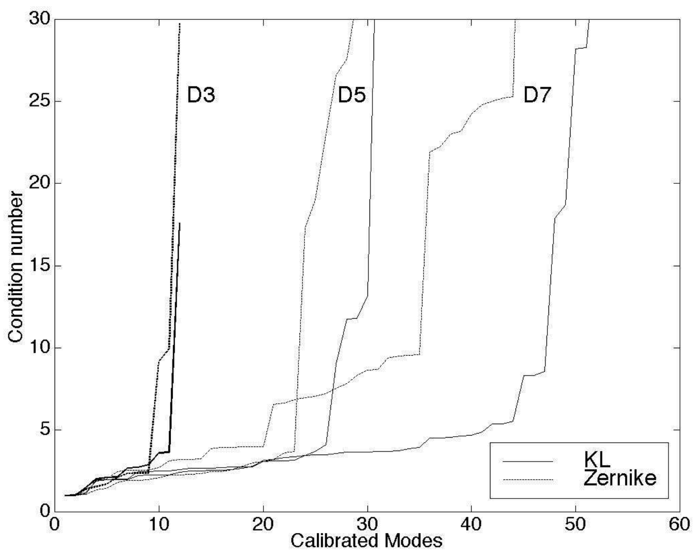

Figure 3. Condition numbers of interaction matrices. Solid: Karhunen-Loève functions. Dashed: Zernike polynomials. Thick-Medium-Thin: D3, D5 and D7 microlens arrays.

Modal aliasing
Aliasing originates from the second term on the right side of eq. (2). It is due to the finite spatial sampling by the subapertures and has got its name from an analogy to spectrum estimation. Neglecting noise, the resulting reconstruction error is given by eq. (4). To be able to investigate the effects of modal aliasing for ALFA's microlens / control modes configurations, $\mathbf{D}_\perp$ has to be known. It cannot be measured since the high order modes cannot be reproduced by the DM, but it can be simulated with a detailed computer model of the ALFA wavefront sensor. This model reproduces the calibrated

slope vectors of the first 30 Karhunen-Loève modes applied to the D7 array with a mean correlation coefficient of 0.97 (minimum 0.94). The simulations were performed using 275 modes, which is a close approximation to infinity because the relative residual fitting error is less than one per mille.
Figure 4 shows the impact of high order modes on the least-squares estimation of the first 30 Zernikes with the D7 array. Aliasing and atmospheric signal are comparable for some modes (especially for the 2nd and 3rd order spherical aberrations), so these should be excluded from the control modes.

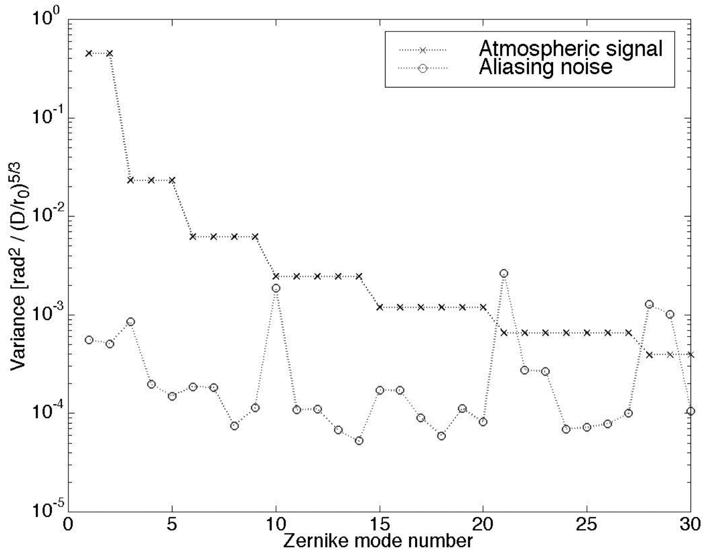

Figure 4. Aliasing error variance for least-squares reconstruction of Zernike polynomials with the D7 microlens array.

Figure 5 shows the ratio of the aliasing- to the fitting error (cumulative variance of uncontrolled modes) for KL and Zernikes. The drop at 7 Zernike- and 9 KL modes respectively is due to the inclusion of coma aberrations, preventing them from being confused with tip-tilt (Glindemann & Rees, 1993). For 30 controlled Zernike modes aliasing- and fitting error are comparable.

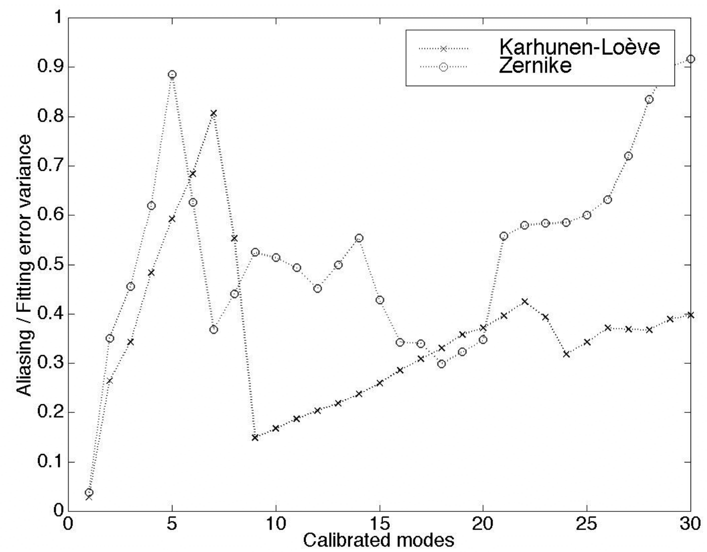

Figure 5. Ratio of aliasing to fitting error for least-squares reconstruction of a varying number of control modes with the D7 microlens array.

3. OPTIMIZED MODAL BASIS SET

Modal cross-talk limits the maximum number of pure Zernike or KL control modes to approximately 8, 22, and 32 modes when measured with the D3, D5 and D7 arrays respectively. In principle, the maximum number of control modes is equal to the number of measurements, i.e. 12, 36, and 60 modes for the different arrays.

It should now be possible to find linear combinations of high-order modes that produce gradient vectors which are large in the space which is not yet used by the pure control modes. These combinations can fill up the measurement space while keeping a certain number of pure control modes, which are the biggest contributors in the decomposition of atmospheric turbulence. This problem can be solved using the technique of Singular Value Decomposition (SVD).

Let [**U**, **S**, **V**] be the SVD of **D** (interaction matrix of pure modes, number of modes < number of gradients), i.e.

$$\mathbf{D} = \left[\mathbf{U}_{\parallel}\ \mathbf{U}_{\perp}\right]\begin{bmatrix}\mathbf{S}\\ 0\end{bmatrix}\mathbf{V}^{t}. \tag{5}$$

$\mathbf{U}_{\perp}$ is the left nullspace of **D**. Its column vectors are an orthonormal basis set for the space orthogonal to that used by the pure control mode gradients.

($\mathbf{U}_{\perp}\,\mathbf{U}_{\perp}^{t}$) **G** is then the projection of some interaction matrix **G** onto the not yet used sensor space. Let **G** be the interaction matrix produced by the DM influence functions (DM surface after pushing one specific actuator), since we are only interested in modes that the DM can produce. Let now [$\mathbf{U}_r$, $\mathbf{S}_r$, $\mathbf{V}_r$] be the SVD of ($\mathbf{U}_{\perp}\,\mathbf{U}_{\perp}^{t}$) **G**, i.e.

$$\mathbf{U}_{\perp}\mathbf{U}_{\perp}^{t}\mathbf{G} = \mathbf{U}_{r}\left[\mathbf{S}_{r}\quad 0\right]\begin{bmatrix}\mathbf{V}_{r\perp}^{t}\\ \mathbf{V}_{r\parallel}^{t}\end{bmatrix} \tag{6}$$

$\mathbf{V}_{r\perp}$ is the rowspace of ($\mathbf{U}_{\perp}\,\mathbf{U}_{\perp}^{t}$) **G**. The column vectors of $\mathbf{V}_{r\perp}$ contain linear combinations of the influence functions which produce a gradient in $\mathbf{U}_{\perp}$, and SVD orders them such that the ones with the largest gradients are placed first. These combinations resemble the desired modes which complete the available measurement space.

As an example, consider the D7 array. With 60 gradients available, the first 34 KL modes can be calibrated (Figure 3) while keeping the condition number low. With the method described above, we can find 20 additional modes without an increase of the the condition number which results in a total of 54 control modes. We labeled these enhanced modal basis sets *sensor modes*, since the pure modes are completed by modes which the DM can produce and which fill up the sensor space.

Comparison to experiment

The data was taken in 0.8″ K-band seeing on the V=7 star SAO 65123 with ALFA operating at a framerate of 300 Hz. Strehl ratios were computed for each individual image, and averaged over 10 images to provide a data point. The experiment was a complete randomized design covering two bases (Karhunen-Loève and the Sensor-modes) and three microlens arrays (D3, D5, and D7).

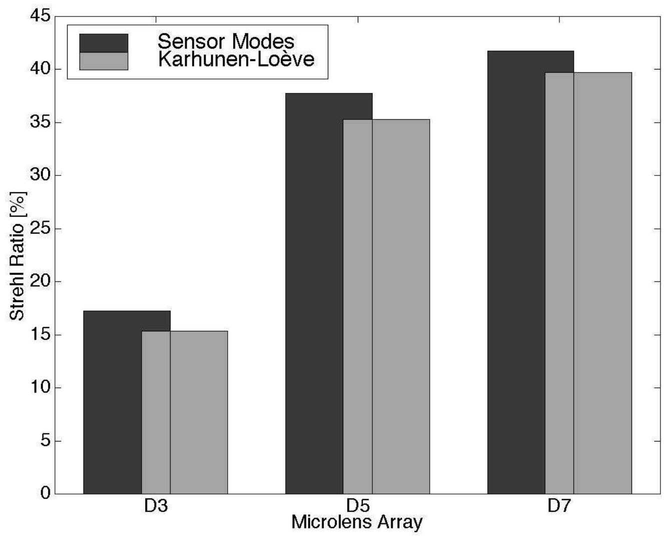

Figure 6. Performance of KL and sensor modal basis sets using the ALFA microlens arrays.

The effects of the basis were compared using both a pooled and a paired t analysis for each microlens array individually. Additionally, a paired analysis for all arrays was performed. As Table 1 shows, the sensor mode basis performed better than the KL basis, with confidence levels ranging from 79% to 93%. When all mode sets are combined, the confidence level rises to 98%. This result is not too surprising since the sensor mode basis contain the KL basis, but further corrects for some additional modes. It should also be noted, that using the additional modes will be most effective in the bright guide star case, because they are of a high spatial order, and thus of a low signal to noise ratio (Gendron, 1995).

Table 1. Comparison of Karhunen-Loève and Sensor Mode Bases

Microlens Array	Pooled Significance	Paired Significance
D3	0.88	0.86
D5	0.84	0.84
D7	0.79	0.93
All	*	0.98

4. PHASE ESTIMATION

After having calculated the gradients (for a comparative study of centroiding algorithms see Kasper et al., 1999) various estimation methods can be used to reconstruct the incident wavefront (e.g. Wallner, 1983, Solomon, 1995, Law, 1996).
The estimation of modal coefficients from Shack-Hartmann gradients can be considered the general problem of estimating a gaussian vector **a** in the presence of gaussian noise **e** (uncorrelated with **a**). Using the linear model from eq. (2)

$$\mathbf{g} = \mathbf{D}\mathbf{a} + \mathbf{D}_\perp \mathbf{a}_\perp + \mathbf{n} = \mathbf{D}\mathbf{a} + \mathbf{e}, \tag{7}$$

we can define an effective error **e** as the sum of aliasing- and measurement error. If the Karhunen-Loève basis set is used, i.e. if the control modes are statistically independent, **a** and **e** will be statistically independent. Further, both will have gaussian statistics.
The optimum estimator for this class of estimation problem is shown to be (e.g. Melsa & Cohn, pp. 211)

$$\begin{aligned} &\hat{\mathbf{a}} = \mathbf{R}\cdot\mathbf{g} = (\mathbf{D}^t\mathbf{V}_\mathbf{e}^{-1}\mathbf{D} + \mathbf{V}_\mathbf{a}^{-1})^{-1}\mathbf{D}^t\mathbf{V}_\mathbf{e}^{-1}\cdot\mathbf{g} \\ &\mathbf{V}_\mathbf{e} := \langle \mathbf{e}\mathbf{e}^t \rangle, \quad \mathbf{V}_\mathbf{a} := \langle \mathbf{a}\mathbf{a}^t \rangle \end{aligned} \tag{8}$$

with $\mathbf{V}_\mathbf{e}$ and $\mathbf{V}_\mathbf{a}$ denoting the covariance matrices of noise and estimation parameters (modal coefficients). The mean values of the estimation parameters (static aberrations) are considered zero, which can always be reached by applying modal offsets to the DM. Eq. (8) is also known as the maximum a posteriori (MAP) estimate, because prior information on the statistics of the estimation parameters is used. Neglecting this information, i.e. assuming infinite modal covariance ($\mathbf{V}_\mathrm{a}^{-1} = 0$), eq. (8) reduces to the weighted least squares estimate. If the gradient noise is further assumed to be uniform and uncorrelated, the $\mathbf{V}_\mathbf{e}$s vanish, and eq. (10) reduces to the least squares estimate.

Determination of $\mathbf{V}_\mathrm{e}$.
Using eq. (7) and the given assumptions, the effective noise covariance can be written as

$$\mathbf{V}_\mathbf{e} = \mathbf{D}_\perp \langle \mathbf{a}_\perp \mathbf{a}_\perp^t \rangle \mathbf{D}_\perp^t + \langle \mathbf{e}\mathbf{e}^t \rangle. \tag{9}$$

The first term, i.e. the aliasing noise, is obtained from simulations and has to be scaled with (D/r_0)^(-5/3). The Fried parameter r_0 can be estimated from open loop gradients, as well as from closed loop data (Veran, 1997).
The second term can be measured using a method described by Gendron (1995). This method uses the fact that the autocorrelation function of a gradient time-series shows a step at the origin which is produced by the uncorrelated measurement noise. This noise can then be estimated as the difference between the first two values of the autocorrelation function. Figure 7 shows a measurement for the ALFA D5 array which rules out the assumption of uniform noise. Possible reasons for this observation are unequally illuminated subapertures (Figure 2) or static aberrations of the microlenses.

Determination of $\mathbf{V}_\mathrm{a}$.
Using the eqs. (7, 9) and the pseudo-inverse of the interaction matrix $\mathbf{D}^+$, the covariance of the modal coefficients is

$$\mathbf{V}_\mathbf{a} = \mathbf{D}^+ \langle \mathbf{g}\mathbf{g}^t \rangle \mathbf{D}^{+t} - \mathbf{D}^+\mathbf{V}_\mathbf{e}\mathbf{D}^{+t}. \tag{10}$$

The performance of the estimator in eq. (8) depends only weakly on the accuracy of $\mathbf{V}_a$. Results obtained from simulations showed that an relative error of 20% in the modal coefficient covariance does not significantly change the estimator performance (Figure 8). This accuracy can easily be reached by careful measurements.

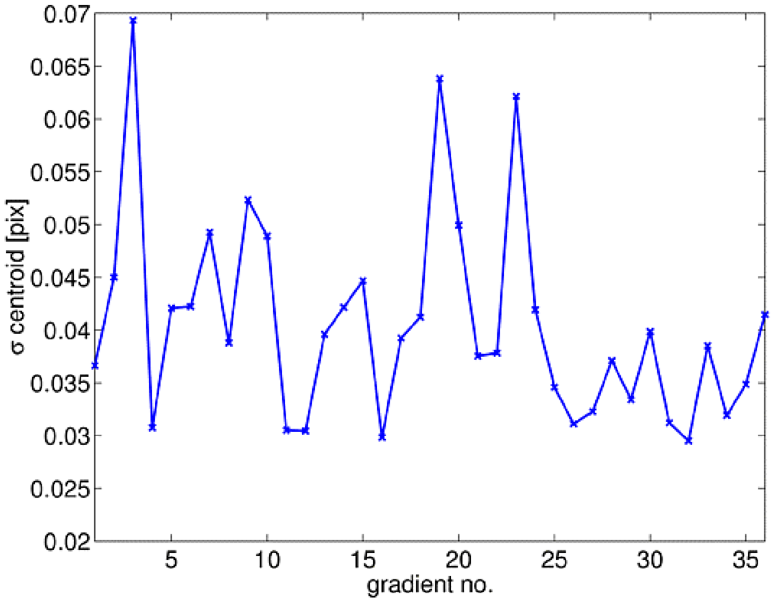

Figure 7. Measured noise for the 36 gradients of the D5 array.

Figure 8 compares the simulated performance of least-squares, weighted least-squares and minimum variance estimators. The minimum variance estimator performed better in any case, even if the assumed $\mathbf{V}_a$ is uncertain within 20%. Its estimation error variance is bounded even at very low light levels while that of the other estimators is not.

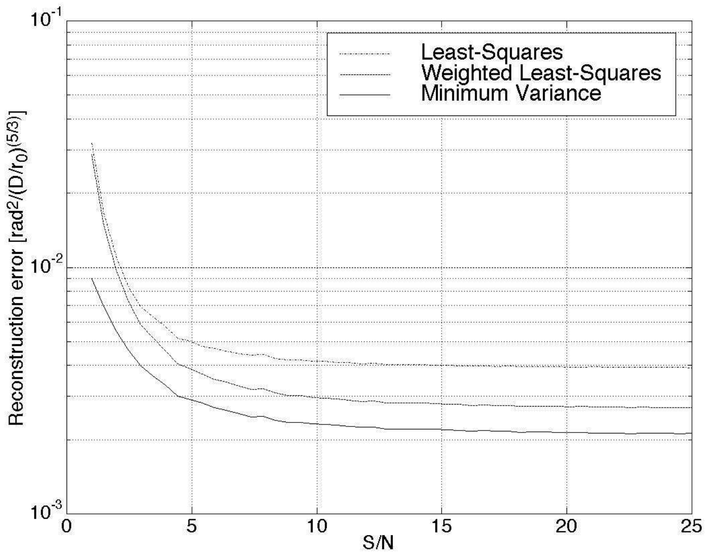

Figure 8. Simulated error variance for the reconstruction of 30 KL modes with the D7 array for different phase estimators. Closed loop variance of the modal coefficients is uncertain within 20%.

Comparison to experiment

An comparison of least-squares and weighted least-squares estimation was made, but under unfavourable seeing conditions (1.3" in K and quite variable). A V = 11 star was observed using the D5 array. Noise and closed loop variance was measured on this star and the reconstruction matrices as introduced above have been calculated. The obtained Strehl-ratios were used as a performance indicator, because the Strehl ratio is a function of the residual wavefront error variance. The weighted least-squares estimator performed better than the least squares fit, but the significance level (74%) was not high enough to conclude this with certainty.

A repetition of this experiment has been done in February 2000 including also the minimum variance estimator. The increased data basis raised the confidence levels, so that we can now conclude that both the weighted least-squares and the minimum variance estimator have been superior to the least-squares fit with a confidence level of 95 %, but no significant difference was found between the two.

5. CONCLUSION

Modal wavefront reconstruction with a Shack-Hartmann sensor depends on the geometric interaction between the microlens array and the modal basis set. A system analysis of ALFA's Shack-Hartmann sensor showed, that it is not possible to independently identify more pure Zernike or KL modes than roughly two thirds the number of measured gradients. However, we showed how to construct a modal basis set which fills up the available measurement space while keeping the pure low order modes. The new modal basis set was tested, and was shown to produce better results than the pure basis sets. We further showed how the modal estimation problem in adaptive optics can be treated as a standard gaussian estimation problem, and simulated expected performances for ALFA's microlens arrays. The comparison to a first experiment did not produce a striking result, but after a repetition we have been able conclude that the advanced estimators performed better in a test at the telescope. Since these estimators depend on the atmospheric conditions during the observation which will change with time, we plan to continuously measure noise and r_0 in closed loop and frequently update the estimator.

6. REFERENCES

1. Cannon, R. C., *Optimal bases for wave-front simulation and reconstruction on annular apertures*, JOSA A **13**, 862-867, 1996
2. Dai, G.-M., *Modal wavefront reconstruction with Zernike polynomials and Karhunen-Loève functions*, JOSA A **13**, 1218-1225, 1996
3. Gendron E., Léna P., *Experimental results of an optimized modal control*, A&A Suppl. Ser. **111**, 153-167, 1995
4. Glindemann, A., Rees, N. P., *UKIRT 5-axis tip-tilt secondary, wavefront sensor simulations*, ICO-16 Satellite Conference on Active and Adaptive Optics, Ed. F. Merkle, 273-278, Garching, 1993
5. Herrmann, J., *Cross coupling and aliasing in modal wave-front estimation*, JOSA A **71**, 989-992, 1981
6. Hippler, S., et al., ALFA: *The MPIA/MPE Adaptive Optics with a Laser for Astronomy Project*, SPIE Vol. 3353, 44-55, 1998
7. Kasper, M. E., et al., *ALFA: Adaptive Optics for the Calar Alto Observatory*, Experimental Astronomy accepted
8. Kasper, M. E., et al., *Increasing the sensitivity of a Shack-Hartmann sensor*, Proc. Wavefront sensing (download from http://www.ukc.ac.uk/physical-sciences/main/School/submit.htm), Canterbury, 1999
9. Law, N. F., Lane, R. G., *Wavefront estimation at low light levels*, Opt. Comm. **126**, 19-24, 1996
10. Melsa, J. L., Cohn, D. L., *Decision and Estimation Theory*, McGraw-Hill Book Company
11. Solomon, C. J., Dainty, J. C., Wooder, N. J., *Bayesian Estimation of Atmospherically Distorted Wavefronts Using Shack-Hartmann Sensors*, Opt. Rev. **2**, 217-220, 1995
12. Veran, J.-P., et al., *Estimation of the adaptive optics long-exposure point-spread function using control loop data*, JOSA A **14**, 3057, 1997
13. Wallner, E. P., *Optimal wave-front correction using slope measurements*, JOSA A **73**, 1771-1776, 1983