

INAUGURAL-DISSERTATION

**zur
Erlangung der Doktorwürde
der
Naturwissenschaftlich-Mathematischen
Gesamtfakultät
der
Ruprecht-Karls-Universität
Heidelberg**

**vorgelegt von
Dipl.-Phys. Markus Erdmann Kasper
aus Bad Nauheim**

Tag der mündlichen Prüfung: 27. Oktober 2000

Optimierung einer adaptiven Optik und ihre Anwendung in der orts aufgelösten Spektroskopie von T Tauri

Gutachter: Priv.-Doz. Dr. Andreas Glindemann
Prof. Dr. Immo Appenzeller

Dissertation
submitted to the
Combined Faculties for the Natural Sciences and for Mathematics
of the Rupertus Carola University of
Heidelberg, Germany
for the degree of
Doctor of Natural Sciences

**Optimization of an adaptive optics system and its application to high-resolution
imaging spectroscopy of T Tauri**

presented by
Diplom-Physicist: Markus Erdmann Kasper
born in: Bad Nauheim

Heidelberg, 27.10.2000

Referees: Priv.-Doz. Dr. Andreas Glindemann
Prof. Dr. Immo Appenzeller

Zusammenfassung

Die Max-Planck-Institute für Astronomie und für Extraterrestrische Physik betreiben seit wenigen Jahren die Adaptive Optik (AO) ALFA an der Calar Alto Sternwarte. ALFA basiert auf einem Shack-Hartmann-Wellenfrontsensor (SHS) und ist eines der wenigen Systeme weltweit, die mit einem Laserleitstern ausgerüstet sind. Dieser Laserleitstern kann überall am Himmel erzeugt werden und ermöglicht so eine nahezu komplette Himmelsabdeckung, die mit natürlichen Leitsternen alleine nicht möglich wäre. Der aktuelle Stand der Technik erlaubt jedoch nur die Erzeugung eines relativ leuchtschwachen Laserleitsterns, der zudem durch das zweifache Durchlaufen der Atmosphäre räumlich ausgedehnt ist. Beides führt zu einem schwachen Signal im Wellenfrontsensor. Die möglichst effektive Verwendung natürlicher Leitsterne als auch des Laserleitsterns erfordert daher eine Optimierung der Wellenfront-Rekonstruktion, die über Standardalgorithmen hinausgeht.

Der Messprozess beginnt mit der Bestimmung der Intensitätsschwerpunkte im Bildmuster des SHS. Der allgemein verwendete Schwerpunktsalgorithmus wurde durch Ausblendung unwesentlicher Pixel modifiziert. Eine weitere Leistungssteigerung wurde durch die Gewichtung der Intensitäten entsprechend ihrem Signal-Rausch-Verhältnis erreicht. Die Rekonstruktion der Wellenfront aus den SHS Gradienten erfolgt dann in der Regel über die Methode der kleinsten quadratischen Abweichung. Dabei werden diejenigen Moden des deformierbaren Spiegels (DM) gesucht, deren Überlagerung die gemessenen Gradienten am besten reproduziert. Je nach verwendetem Linsenarray wurde der effektivste Modensatz für die Korrektur mit dem DM ermittelt. Darüber hinaus wurde der Maximum a Posteriori (MAP) Schätzer als Alternative zur Methode der kleinsten Quadrate untersucht. Dieser wägt die statistische Wahrscheinlichkeit der Modenkombinationen gegen das Rauschniveau ab und produziert so den kleinstmöglichen Rekonstruktionsfehler. Es wurden Verfahren entwickelt, um die für den MAP Schätzer benötigten Statistiken der Moden und des Rauschens zu bestimmen. Weitere Verbesserungen betreffen ein Mikrolinsenarray, welches optimal an die ringförmige Teleskoppupille angepasst ist, und ein Verfahren zur Bestimmung der besten Bildrate des Regelkreises.

Die angesprochenen Modifikationen wurden im Computer simuliert, in ALFA implementiert und am Teleskop getestet. In ihrer Kombination ergaben sie eine Steigerung der Grenzhelligkeit der benutzbaren Leitsterne von ca. $m_V = 11$ auf $m_V = 14$ bei sehr guten Beobachtungsbedingungen. Die Analyse zeigt, daß eine weitere Steigerung um etwa eine halbe Magnitude durch einen rauschfreien Wellenfrontsensor möglich wäre.

Des Weiteren wurden neue astronomische Beobachtungsmethoden in Verbindung mit AO verwendet, um das junge Sternsystem T Tauri zu untersuchen. Durch räumlich hochaufgelöste Spektroskopie gelang es, die Nahinfrarotspektren der Komponenten T Tau N und T Tau S individuell zu bestimmen. Der Calar Alto ist momentan die weltweit einzige Sternwarte, die diese Möglichkeit bietet. Dabei zeigt sich, daß beide Sterne ein ähnliches Emissionslinienspektrum besitzen, und daß T Tau S keine für stellare Photosphären üblichen Merkmale aufweist. Mit Hilfe der AO-kompensierten Speckle-Interferometrie konnte die Existenz des kürzlich entdeckten Begleiters von T Tau S im Abstand von nur $0.08''$ bestätigt werden. Unter Verwendung bereits publizierter Daten konnten die Bahnparameter weiter eingengt werden, die auf eine relativ hohe Masse von zwei Sonnenmassen für T Tau S hindeuten.

Abstract

The Max-Planck institutes for astronomy and for extraterrestrial physics run the adaptive optics (AO) system ALFA since a few years now. ALFA incorporates a Shack-Hartmann wavefront sensor (SHS) and a laser guide star (LGS) facility. It is one of very few systems worldwide that are able to project an LGS wherever it is needed and (potentially) achieve a nearly complete sky-coverage. The currently available technology only permits relatively faint LGS, which are furthermore inherently extended due to the roundtrip of the laser beam through the atmosphere. This leads to a noisy wavefront sensor measurement. Hence, an effective use of natural guide stars as well as the LGS requires an optimum wavefront estimation process that outperforms standard algorithms.

Wavefront reconstruction with an SHS starts with the calculation of spot centroids. The common weighted pixel average (WPA) algorithm was modified by pixel masking methods. A further improvement was achieved by weighting the pixel intensities with their signal to noise ratio before the WPA is performed. Then, the wavefront is generally reconstructed by least-squares (LS) fitting the gradients of the control modes to the measured gradients. Depending on the actually used microlens array, the most effective modal basis set for a LS reconstruction was determined. In addition, the maximum a posteriori (MAP) estimation method was investigated as an alternative to the LS method. The MAP balances the probability of modes to occur in turbulent wavefronts with the measurement noise, and comes up with the most accurate linear estimate possible. Techniques were developed to determine the statistics of modal basis sets and noise which are required for the MAP calculation. Further enhancements of the system concern the design of microlens arrays perfectly adapted to the annular telescope pupil, and a simple algorithm to determine the optimum framerate of the control loop.

Most of the modifications were first simulated in the computer, then implemented in ALFA, and finally tested at the telescope. The combined effect is an improvement in limiting guide star brightness from about $m_V = 11$ to $m_V = 14$ under the very best observing conditions. The analysis shows that another half of a magnitude could be gained by using a detector with virtually zero read-noise.

In addition, the young stellar system T Tauri was observed with AO assisted imaging and long-slit spectroscopy. In this way, the two stars T Tau N and T Tau S could be separated, and their near-infrared spectra were obtained individually. Only the Calar Alto observatory currently permits those observations. It turned out that both stars exhibit similar hydrogen recombination-line spectra, but only T Tau N shows features that arise in a stellar photosphere. By using speckle methods on AO compensated short exposures, the recently discovered $0.08''$ companion to T Tau S could be confirmed. In combination with published data, the parameter space of its orbit could be narrowed, and indication for a rather large mass of about 2 solar masses was found for T Tau S.

Natur- und Kunstwerke lernt man nicht kennen, wenn sie fertig sind;
man muss sie im Entstehen aufhaschen, um sie einigermaßen zu begreifen.

Goethe an Zelter, 4. August 1803

Contents

1	Introduction	1
2	Wavefront sensing in adaptive optics	7
2.1	Modal decomposition of atmospheric turbulence	8
2.1.1	Turbulent wavefronts	8
2.1.2	Modal decomposition	10
2.1.3	Zernike polynomials	11
2.1.4	Annular Zernike polynomials	13
2.1.5	Karhunen-Loève functions	14
2.1.6	Mirror modes	16
2.2	Wavefront sensing methods	16
2.2.1	Shack-Hartmann sensor	18
2.2.2	Pyramid sensor	20
2.2.3	Curvature sensor	21
2.2.4	Phase diversity	23
2.2.5	Measurement error	24
2.3	Phase reconstruction with the Shack-Hartmann sensor	25
2.3.1	Sensor measurement model	26
2.3.2	Modal cross-talk	27
2.3.3	Sensor mode extension	28
2.3.4	Modal reconstruction error	29
2.3.5	Modal covariance estimation	31
2.3.6	Phase estimators	33
2.4	Laser guide star peculiarities	37
2.4.1	Tip-tilt sensing	38
2.4.2	Focus sensing	39
2.4.3	LGS modes	40
3	Application to the ALFA AO system	41
3.1	System description	41
3.1.1	Optics	42
3.1.2	Control components	43
3.1.3	Modal control	48
3.2	Centroid estimation	54
3.2.1	Monte-Carlo simulation	54
3.2.2	SHS measurement noise determination	56

3.2.3	Impact of pixel-scale	57
3.2.4	Comparison of algorithms by simulations	58
3.2.5	Comparison of algorithms by experiment	59
3.2.6	Quad-cells	60
3.3	Model of ALFA's Shack-Hartmann sensor	61
3.4	Modal basis sets by comparison	62
3.4.1	Cross-talk	62
3.4.2	Reconstruction error	63
3.4.3	Remaining error	67
3.5	Performance at the telescope	68
3.5.1	Zernike polynomials and K-L functions	69
3.5.2	Sensor mode extension	71
3.5.3	Phase estimators	72
3.6	Summary	73
4	Adaptive Optics imaging and spectroscopy of T Tauri	75
4.1	T Tauri - an unusual Pre-Main-Sequence system	75
4.2	Observations and data reduction	77
4.2.1	ALFA + 3D: Imaging spectroscopy	77
4.2.2	ALFA + Omega-Cass	78
4.3	Results	80
4.3.1	H- and K-band photometry	80
4.3.2	The companion to T Tau S	81
4.3.3	H- and K-band spectroscopy of T Tauri	83
4.4	A Model of T Tau S	92
4.5	Summary	95
A	Notations	97
B	Acronyms	99
C	Singular Value Decomposition	101
D	Gaussian Estimation	103

Chapter 1

Introduction

Do we need to observe at the diffraction limit?

The spatial resolution of ground-based telescopes suffers from atmospheric turbulence which disturbs the incoming flat wavefront of astronomical objects. Even the best astronomical sites do not allow of optical and near-infrared resolutions much better than the seeing limit of one arcsecond. Modern large telescopes would be capable of delivering images that are orders of magnitudes sharper than the seeing. As an example, a 10-m telescope operating at optical wavelengths can in principle resolve structures that are separated by an angle of only 10 milli-arcseconds (mas) which is about a hundred times better than the seeing. The consequences of a diffraction limited performance for astronomical observations are diverse and can be split up in two main categories:

Spatial resolution The diffraction limited resolution grows proportionally to the telescope diameter, while it does not depend on telescope size in the seeing limited case. The spatial resolution drives the progress in understanding objects with a complex morphology. Popular examples in the context of star formation research are any sort of disk- and outflow-phenomena, stellar clusters, and of course the direct detection of extra-solar planets. Cosmologists instead may be excited by the structure of quasar host galaxies or gravitational lenses.

Sensitivity The diffraction limited “detectivity” of point sources, defined as the time needed to reach a certain signal-to-noise ratio, grows as the 4th power of the telescope diameter, whereas it grows only as the 2nd power in the seeing limited case. To put this into numbers: In the case of background limited performance (BLIP), a diffraction limited 10-m telescope operating at visible wavelengths has about 80000 times the detectivity of a seeing limited 3.5-m telescope. With the improved sensitivity, it would be possible to exactly measure large distances by detecting Cepheids, measure cosmological parameters by detecting supernovae at high-redshifts, or again finding extra-solar planets.

A side-effect of diffraction limited performance is that it keeps the instrumentation “comparatively” simple. Future giant telescopes will have very large focal lengths resulting in very large magnifications. In the diffraction limited case, however, the lateral size of a point source will be independent of the aperture diameter, making slit widths and beam sizes comparable to those of present instruments.

The two main approaches to reach the diffraction limit are either to correct the atmospheric effects, or simply to avoid them by observing from space. *Space telescopes* naturally reach the image quality allowed by their optics. In addition, they do not suffer from atmospheric background radiation, and they are not restricted to the wavelength-windows where the atmospheric transmission permits

observing. There are, however, considerable drawbacks related to space telescope facilities: They are very expensive (the Hubble Space Telescope (HST) was almost three billion dollars, and costs of 500 million dollars are expected for the Next Generation Space Telescope (NGST) which will be launched in 2008), and they are difficult to maintain.

Without much effort, *speckle interferometry* allows of diffraction limited images from the ground, making this method a formidable tool to detect close binary stars. In terms of sensitivity, however, speckle imaging is even worse than conventional seeing limited imaging. Due to the required short exposures and the poor signal in the single speckle images, speckle interferometry is restricted to bright objects of about 11th magnitude almost independently of the aperture diameter. Further, speckle interferometry does not provide real-time narrow images preventing it from being applied to spectroscopy.

Adaptive Optics (AO) is a technique that flattens distorted wavefronts with a deformable mirror (DM) in a feedback loop in order to provide diffraction limited performance with ground-based telescopes. Mainly developed for military applications, this technique has only recently been introduced to astronomy. Nevertheless, AO has already demonstrated that images with sharpness rivaling those of the HST can be obtained from the ground.

AO history. Although already proposed by Babcock (1953), it took more than 20 years to build the first AO system able to sharpen images (Hardy et al., 1977). From that on, AO systems were widely developed for military applications mainly for imaging satellites. Most satellites are rather bright, whereas most of the exciting astronomical objects are much fainter than the stars that are visible to the naked eye. The development of more sensitive detectors and the realization that the demands are less restrictive in the near-infrared led to the first AO systems dedicated for astronomical observations.

The European Southern Observatory (ESO) first built the ‘COME-ON’-system which used a DM with 19 piezo-electric actuators and a Shack-Hartmann sensor (SHS) at the 1.52-m telescope of the Observatoire de Haute Provence (Rousset et al., 1990). In 1992, this system was upgraded by a new DM with 52 actuators and a larger correction bandwidth and installed as ‘COME-ON-PLUS’ at the 3.6-m telescope on La Silla, Chile. After a concept change in 1994, which mainly involved the control computer and the software to make the system simple and effective to use, it was called ADONIS and became a user instrument facility. In the meantime, a novel technique was conceived at the US National Optical Astronomy Observatories (NOAO). The new curvature wavefront sensor (Roddier, 1988) and the new bimorph mirror were the basic elements for an experimental AO system. This system possessed 13 degrees of freedom and was successfully tested for the first time in December 1993 at the Canada-France-Hawaii telescope (CFHT) on Mauna Kea. In 1996, the 19 actuator system PUEO became an user instrument at the CFHT. Nowadays, each of the large telescope projects plans the installation of or already has installed high order AO systems that uses either Shack-Hartmann sensors (Keck AO system, ESO NAOS) or curvature sensors (Gemini Hokupa’a, ESO MACAO).

Limited sky coverage. Although AO can provide diffraction limited performance at any telescope, the guide star problem limits its applicability. Since atmospheric turbulence evolves temporally and spatially on rather small scales, the guide star must be bright enough to be sampled at a high frequency, and it must be located close to the astronomical object. Natural guide stars (NGS) that serve these purposes have to be brighter than 16th magnitude even for the most sensitive curvature systems working in the near-infrared under the most friendly observing conditions (see Figure 1.1), and they have to be within about half an arcminute around the object (the angular size of the patch over which the wavefront is fairly uniform is called *isoplanatic angle*). Unfortunately, these requirements imply that only a few percent of the sky are observable with an NGS AO system.

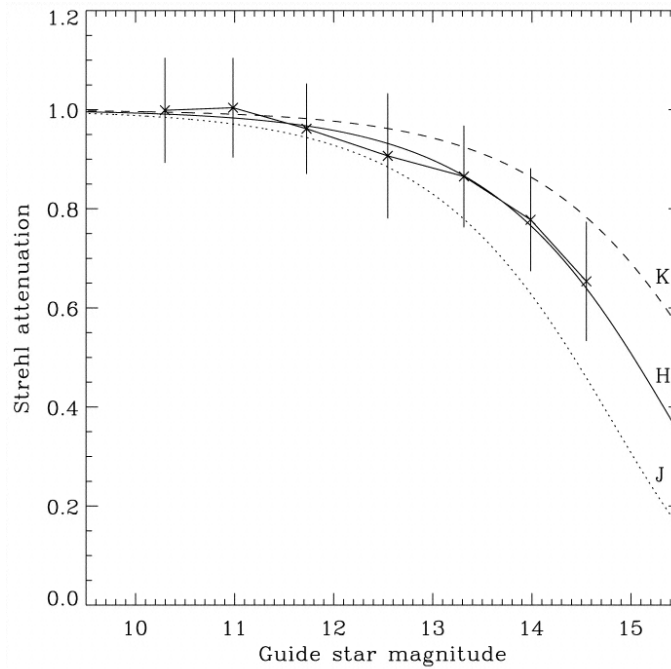


Figure 1.1: Strehl ratio of the CFHT curvature AO system as a function of the guide star magnitude in the visible obtained from simulations and observations.

Laser guide stars. Foy and Labeyrie (1985) proposed the use of laser beacons to create artificial sources with light back-scattered by the atmosphere. The two mechanisms considered so far are Rayleigh scattering up to 30 km above the ground, and resonance fluorescence of sodium atoms concentrated in a layer at about 90 km height. These Laser Guide Stars (LGS) can be created wherever the telescope points, but they also possess considerable drawbacks: First, due to the finite distance between telescope and LGS, the backscattered beam does not sample the full aperture at the height of the turbulent layers. This *cone effect* or *focal anisoplanatism* (see Figure 2.17) is more severe for larger apertures and higher turbulent layers. Second, an LGS cannot be used to measure tip and tilt of the wavefront. An NGS with reduced demands on its brightness and distance to the object is still required for this purpose. Finally, LGSs are (currently) faint and extended which results in a poor signal to noise ratio of wavefront sensor measurements, and the roundtrip through the atmosphere makes them extremely sensitive to seeing and transparency changes.

Beside the military system at the Starfire Optical Range (SOR) 1.5-m telescope, no other LGS project has managed to steadily produce satisfactory results yet. It has still to be demonstrated that LGSs can be bright and stable enough to let astronomical AO systems provide fully corrected images.

Hence, there is an ongoing research on the analysis of poorly and partially corrected images, and the improvement of wavefront sensors in terms of sensitivity and accuracy. Very low read-noise CCDs in combination with new wavefront sensor types promise a gain of at least 1-2 magnitudes for the limiting guide star brightness (Ragazzoni and Farinato, 1999). Advanced algorithms for the reconstruction of wavefronts and the compensation in a feedback-loop can push the limit even further towards faint and/or extended guide stars like the LGS.

The future of AO. There is an obvious trend towards very large and extremely large telescopes, for which the cone effect of a single LGS cannot be accepted. The use of multiple LGSs, however, offers

a solution: They can be arranged such that the cross sections of their beams fully sample the aperture at each height. As one can imagine, this approach requires sophisticated algorithms to reconstruct a wavefront from sensor signals.

In the next step, multiple guide stars could be used to extend the FOV over the isoplanatic patch. The *Multi-Conjugate Adaptive Optics* (MCAO) approach explores how signals from several guide stars can be used to disentangle the turbulent layers in the atmosphere. Placing multiple DMs conjugate to these layers would result in a local correction of a wavefront just at the layers and therefore an enhanced FOV of about 1 arcminute for a configuration with 5 sodium LGSs (Berkefeld, 1998).

Interestingly, as the telescope size increases, the allowable separation of the guide stars for MCAO also increases. The astonishing conclusion is that telescopes with apertures of 100 m may not even need LGSs, because a complete sky-coverage can be obtained with multiple NGSs (Ragazzoni, 1999). However, for the currently existing telescopes with moderate aperture diameters, there is no alternative to LGSs if complete sky-coverage is desired.

ALFA. The Max-Planck-Institutes for Astronomy (MPIA) and for Extraterrestrial Physics (MPE) operate the ALFA (Adaptive Optics with a Laser for Astronomy) AO system at the Calar Alto observatory. This system is one among three non-military AO systems world-wide that are currently able to project an LGS onto the sky - the other two being operated at the Lick observatory's Shane telescope and at the Steward observatory's Multi Mirror Telescope (MMT).



Figure 1.2: The adaptive optics system ALFA attached to the Cassegrain focus of the 3.5-m telescope on Calar Alto. The golden tube below ALFA is the infrared camera Omega-Cass.

Figure 1.2 shows ALFA attached to the Cassegrain focus of the 3.5-m telescope on Calar Alto. The golden tube below the breadboard is the infrared science camera Omega-Cass (Lenzen et al., 1998). The main features of the AO bench which comes under the authority of MPIA are a 97 discrete actuator DM with a continuous facesheet, and interchangeable microlens arrays for the SHS with a

varying number of subapertures. The original idea was to operate ALFA in the visible also, requiring the option of having a microlens array with more than 100 subapertures over the aperture. Nowadays, the ambitions are more modest and the currently used arrays provide a maximum of 30 subapertures which is well-suited for the near infrared. The laser system comes under the authority of MPE. It consists of a dye ring laser with about 3 Watts of continuous wave output power. A total of 8 mirrors, most of them movable, are required to relay the laser beam to the launch telescope which is attached to the side of the main telescope. An extensive description of the laser system can be found in a special issue of *Experimental Astronomy* (Davies et al., 2000, Ott et al., 2000, Rabien et al., 2000).

The design goal of ALFA was to achieve diffraction limited imaging in the K-band and to have access to almost every part of the sky by use of an LGS. When the system was installed at the telescope by the end of 1996, it became clear pretty soon that this ambitious project will keep instrumentalists busy for years. The original software provided only very standard algorithms and virtually no diagnostic tools which could assist in finding the right combination of all the system parameters for different observing conditions. Many modifications and extensions were developed with the aim to make the system more sensitive and easier to use. In the summer of 1998, a breakthrough was obtained by an accurate re-alignment of the optics, and by a better calibration procedure which removes static aberrations in the optical path. Since then, ALFA can be used for scientific observations, and its performance to be expected under different observing conditions is summarized in Table 1.1. Under the best observing conditions, the current limiting magnitude is about $m_V = 14$. A great success was the first diffraction limited performance using the LGS in June, 1999 (Hippler et al., 2000), but still LGS operation is far from being considered “routine”.

Table 1.1: Current performance of ALFA in the K-band. Guide star brightness and seeing indicate observing conditions. *Strehl ratio* and *FWHM* are the common performance metrics for image quality, and *loop robustness* measures the feasibility of the observations (+: routine, \pm : challenging, -: questionable)

Guide star	V-band Seeing [$''$]	Strehl [%]	FWHM [$''$]	Loop robustness
Bright NGS, $m_V \leq 9$	≤ 0.9	≈ 50	0.13	+
	1 – 1.5	≈ 25	≤ 0.2	+
Faint NGS, $m_V \approx 10 - 11$	≤ 0.9	≈ 30	0.15	+
	1 – 1.5	≈ 15	≈ 0.2	+
Very faint NGS, $m_V \geq 12$	≤ 0.9	≈ 15	≈ 0.2	+
	1 – 1.5	$\approx 5 - 10$	≤ 0.4	\pm
LGS	≤ 0.9	≈ 10	≈ 0.3	–
	1 – 1.5	–	–	–

The aim of this work was the development of algorithms that provide the most effective operation of the ALFA Shack-Hartmann system. To achieve this goal, the wavefront estimation process was studied and possible upgrades were found which apply to a variety of observing conditions. These upgrades include effective centroid estimation and subsequent wavefront reconstruction algorithms, diagnostic tools required to choose the optimal operational parameters, and also the design of a most effective microlens geometry for annular pupils.

Observing with AO. AO systems do not have the capability to entirely restore the wavefront and give fully diffraction limited images. A good knowledge of the compensated point spread function (PSF) is therefore required for the data calibration. The AO PSF consists of a diffraction limited core and a halo the size of the seeing disk. The degree of correction is given by the amount of energy

transferred from the halo into the core. Faint structures in the vicinity of an object can still be hidden in the seeing halo. To further increase the complexity of the problem, the PSF varies with time (due to variable seeing) and with angular distance to the guide star. The performance of an AO system depends on the observing wavelength, the intrinsic seeing conditions and the angular distance to the guide star as well as its brightness. As a rule of thumb, the longer the wavelength, the better the intrinsic seeing, and the closer and brighter the guide star, the better the corrected image will be.

The K-band image of the Trapezium-Cluster in Orion in Figure 1.3 shows the decrease of the corrected Strehl ratios with distance to the guide star. The enlarged parts demonstrate the high resolution of the image indicated by the resolved $0.2''$ -binary, and the visible diffraction ring of the PSF.

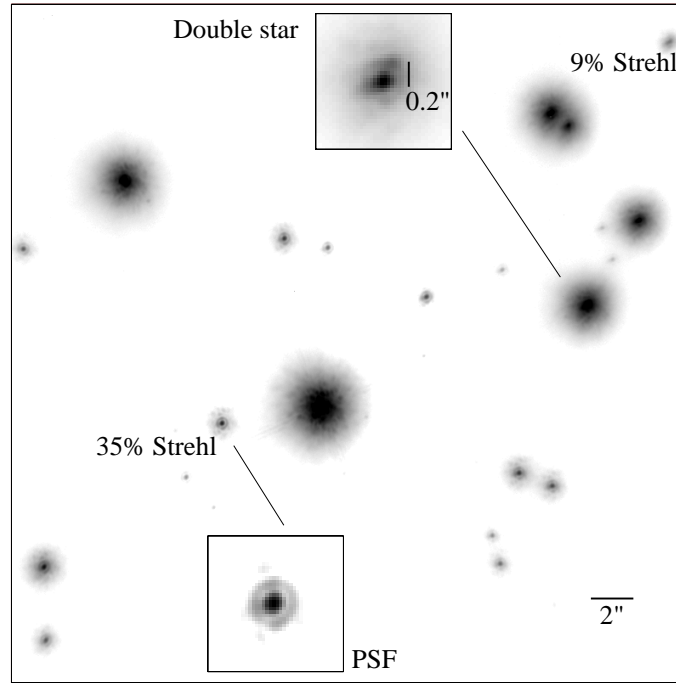


Figure 1.3: Trapezium cluster in Orion imaged with ALFA in the K-band. The bright star close to the center was used as a reference for the AO.

90 % of all refereed AO publications today are based on direct imaging applications (Close, 2000). Chapter 4 presents some new applications of AO for astronomical observations of the triple system T Tauri which is the prototype of a whole class of objects. Although probably being one of the most observed objects in astronomy, T Tauri is only poorly understood still.

T Tauri was observed with AO assisted *imaging spectroscopy*. In this way, the two brighter stars T Tau N and T Tau S could be resolved and individual spectra were obtained simultaneously. These spectra allow us for the first time to spectroscopically study the two components without blending effects. Imaging spectroscopy with AO is also the most powerful tool to construct line-maps at a telescope's diffraction limit.

The other technique was *compensated speckle interferometry* to detect binary stars even slightly beyond the diffraction limit of the telescope. The higher dynamic range of this technique when compared to standard speckle interferometry or deconvolution of compensated images (Roggemann and Welsh, 1996) led to the confirmation of a companion to T Tau S at a separation of only $0.08''$ recently discovered by Koresko (2000).

Chapter 2

Wavefront sensing in adaptive optics

Wavefront sensing in adaptive optics (AO) deals mainly with the interaction of a deformable mirror (DM) with wavefront sensors. Good teamwork of these two components is essential for a high-performance AO system. The DM characteristics define the correctable wavefront shapes, and the wavefront sensor has to be designed such that it can effectively measure these shapes. Constraints are set by the turbulent atmosphere which distributes most of its energy at low spatial frequencies.

The first section of this chapter reviews the fundamental properties of turbulent wavefronts and introduces basis sets of functions commonly used in the modal control approach. In this context, I have calculated the Annular Zernike basis set and its statistical properties. In contrast to zonal control which represents a wavefront by its values in an array of independent subapertures (or zones), modal control represents the wavefront as a series of whole aperture functions of increasing complexity. The advantage of modal control is that the AO performance can be optimized by filtering in both the spatial and temporal domains.

Section 2.2 introduces the most popular wavefront sensing methods in AO. Most of these methods work on a zonal basis in or near the pupil plane. Only Phase Diversity works in the image plane; it is also the only method which inherently works on a modal basis. The pupil plane methods measure either the first or second derivatives of a wavefront's phase. This leads to differences in noise propagation properties, and it will be shown that noise propagation for modal and zonal control is equivalent. Especially interesting for the closed-loop properties is the question, whether diffraction occurs at the full telescope aperture or at the subapertures of the wavefront sensor.

Although most wavefront sensors and correctors work on a zonal basis, modal phase reconstruction algorithms may be used. Section 2.3 explains the topic of modal reconstruction for Shack-Hartmann sensors (SHS), most of the analysis being valid for other sensing methods also. The problem of noise propagation connected to the least-squares reconstruction method is analyzed. The basic result is that the number of natural modes like Zernike polynomials has to be truncated in order to avoid noise amplification. However, we found a method to calculate modes which avoid strong noise amplification when used as an extension to a truncated natural basis set. Further, aliasing noise for SHSs is analyzed, based on which it is shown how the covariance matrix of the modal coefficients can be estimated in open- and closed-loop. The knowledge of the modal covariance is essential for the calculation of optimized phase estimation algorithms presented at the end of the section.

Finally, Section 2.4 discusses the peculiarities of an artificial guide star that is generated by projecting a laser onto the mesospheric sodium layer at about 90 km altitude. Beside well known drawbacks of this method, such as extended spot size and focal anisoplanatism, the effects of sporadic height variations of the sodium layer and the elongation of the spots are investigated.

2.1 Modal decomposition of atmospheric turbulence

This section starts with a repetition of fundamental results of the theory of optical aberrations introduced by refractive index fluctuations in the turbulent atmosphere. An extensive discussion of this subject has been presented by Roddier (1981), or can be found in a more compact form in my diploma thesis (Kasper, 1997). In the context of this work, the modal decomposition of atmospheric distorted wavefronts is of paramount interest. The commonly used basis set is given by the Zernike polynomials, which are discussed in Section 2.1.3. Their popularity is based on their analogy to classical optical aberrations, and their relative ease of mathematical handling. The Annular Zernike polynomials presented in Section 2.1.4 are modified to be orthogonal not on a circle, but on an annulus resembling a telescope pupil with its central obscuration. The Karhunen-Loève functions presented in Section 2.1.5 differ from the other sets in that they represent a mathematical concept rather than a static set of basis functions. They can be calculated for a specific stochastic process, and are optimum in the sense that the number of modes needed to well-approximate this process (in our case a wavefront) tends to be minimal. Finally, in Section 2.1.6, the concept of the *mirror modes*, that depend on the characteristics of the deformable mirror used in an adaptive optics system, is discussed.

2.1.1 Turbulent wavefronts

The *phase structure function* $D_\phi(\mathbf{r})$ of light having travelled through the atmosphere, which is assumed to obey Kolmogorov turbulence, is given by

$$\begin{aligned} D_\phi(\mathbf{r}) &= \langle |\phi(\mathbf{r}' + \mathbf{r}) - \phi(\mathbf{r}')|^2 \rangle \\ &= 2.91 \left(\frac{2\pi}{\lambda} \right)^2 (\cos \gamma)^{-1} |\mathbf{r}|^{5/3} \int_0^\infty C_n^2(z) dz \\ &= 6.88 \left(\frac{|\mathbf{r}|}{r_0} \right)^{5/3}, \end{aligned} \quad (2.1)$$

where \mathbf{r}, \mathbf{r}' are lateral displacement vectors, $\phi(\mathbf{r})$ is the wavefront phase, $\langle \dots \rangle$ denotes the ensemble average, λ is the light's wavelength, γ is the angle of observation measured from zenith, $C_n^2(z)$ is the *refractive index structure constant* which is a function of the height z above ground, and r_0 is the fundamental *Fried-parameter* (Fried, 1965) defined by

$$r_0 = \left[\frac{2.91}{6.88} \left(\frac{2\pi}{\lambda} \right)^2 (\cos \gamma)^{-1} \int_0^\infty C_n^2(z) dz \right]^{-3/5}. \quad (2.2)$$

r_0 has two physical interpretations: First, it is the aperture over which there is approximately one radian of root mean square (rms) phase aberration. Second, it is the aperture which has the “same resolution” (as defined by Fried) as a diffraction-limited aperture in the absence of turbulence. Since λ/D is the diffraction limited resolution of an aperture with diameter D , the maximum angular resolution of a long exposed image through the atmosphere is λ/r_0 , independent of telescope size. This last quantity is commonly known as the astronomical *seeing*.

Using Equation 2.1, Noll (1976) calculated the spatial power spectrum of phase fluctuations due to Kolmogorov turbulence:

$$\Phi(\mathbf{k}) = \frac{0.023}{r_0^{5/3}} |\mathbf{k}|^{-11/3}, \quad (2.3)$$

where \mathbf{k} is a two-dimensional spatial frequency vector. The infinite amount of energy in this spectrum due to the divergence at the origin is not physical. In reality, low order frequencies are truncated by

the *outer scale* of turbulence (defined as the scale of the largest turbulent structures), as modelled by the more realistic *von Kármán* spectrum.

The basic model to study the dynamical effects of turbulence is that the atmosphere is assumed to consist of several distinct turbulent layers at different altitudes, each one moving at some velocity $|\mathbf{v}_\perp|$ parallel to the ground. A further simplification is usually introduced by the “frozen flow” or *Taylor hypothesis*, which assumes that the layers move without changing their refractive index distribution and thus the overall phase screen moves rigidly across the pupil. Conan et al. (1995) have written a comprehensive paper about the temporal behavior of a number of quantities related to atmospheric turbulence. Some useful results of his work are cited here:

While the spatial power spectral density (PSD) of the phase as given in Equation 2.3 is proportional to $|\mathbf{k}|^{-11/3}$, its temporal PSD has the form

$$\Phi(f) \propto C_n^2(z) dz \frac{1}{|\mathbf{v}_\perp|} \left(\frac{f}{|\mathbf{v}_\perp|} \right)^{-8/3}, \quad (2.4)$$

where f is the temporal frequency. The spectral behavior of the Zernike modal basis set is characterized by two distinct domains. At low frequencies, the power spectral density of a mode is constant and beyond a certain cutoff frequency, it drops as $f^{-17/3}$ as displayed in Figure 2.1. The cutoff frequency increases with the radial degree n of the specific mode approximately as

$$f_c(n) \sim 0.3(n+1) \frac{|\mathbf{v}_\perp|}{D}, \quad (2.5)$$

where D is the diameter of the telescope pupil. At very high temporal frequencies, the power of atmospheric turbulence is almost equally distributed among the modes. This has the practical consequence that low order modes have to be corrected with the same bandwidth as high order modes to achieve a good compensation. It should be noted that although Conan discussed only Zernike polynomials, the asymptotic behavior is generally valid for arbitrary modal basis sets. This is due to the fact that all modal basis sets can be transformed into each other. In other words, no selection of modes can simplify the bandwidth problem.

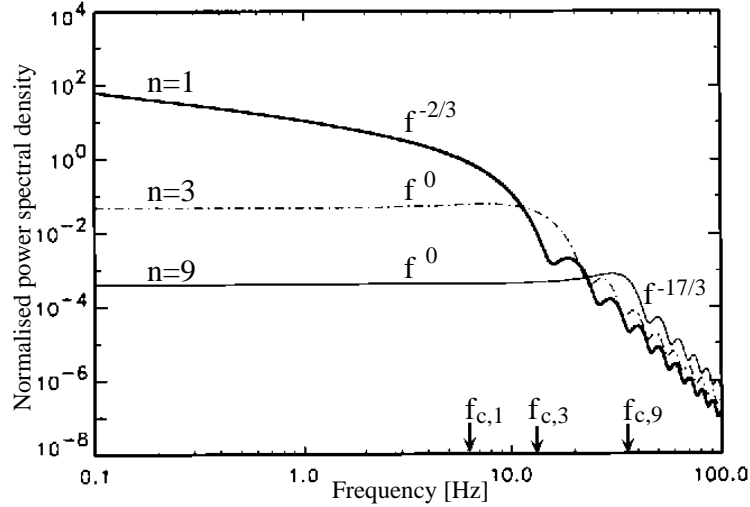


Figure 2.1: Zernike polynomial mean temporal power spectra in a given radial degree n for $n = 1, 3, 9$. The asymptotic power laws and the cutoff frequencies are indicated. From Conan et al. (1995).

To get an idea of how fast the atmospheric turbulence has to be corrected, Greenwood (1977) calculated the residual error of a wavefront which is corrected by a simple servo. Since spatial correction was assumed to be perfect, an error is only introduced by the finite servo bandwidth. The *Greenwood frequency* f_G , defined as the bandwidth at which the residual rms wavefront error is one radian, is given by

$$f_G = \left[0.102 \left(\frac{2\pi}{\lambda} \right)^2 (\cos \gamma)^{-1} \int_0^\infty C_n^2(z) |\mathbf{v}_\perp(z)|^{5/3} dz \right]^{-3/5}. \quad (2.6)$$

The value of f_G is therefore determined by the turbulence and wind profiles of the atmosphere. In the special case of a single turbulent layer, f_G may be expressed as

$$f_G = 0.427 \frac{|\mathbf{v}_\perp|}{r_0}. \quad (2.7)$$

To put this into numbers: in a V-band (550 nm) seeing of $1''$, r_0 is 11 cm in the V- and 60 cm in the K-band ($2.2 \mu\text{m}$). If a turbulent layer moves at a speed of 10 m/s, the closed-loop bandwidths should be around 39 Hz in the V- and 7 Hz in the K-band.

Optical effect of the turbulence. The optical transfer function (OTF) for ground-based observations is the product of the telescope OTF $T(\mathbf{r})$ and the atmospheric OTF $A(\mathbf{r})$. For an optically perfect telescope, $T(\mathbf{r})$ is just the autocorrelation function of the aperture. $A(\mathbf{r})$ is the average autocorrelation function of the phase exponential that can be written as (Roddier, 1981):

$$\begin{aligned} A(\mathbf{r}) &= \langle \exp(i[\phi(\mathbf{r}') - \phi(\mathbf{r}' + \mathbf{r})]) \rangle \\ &= \exp\left[-\frac{1}{2}D_\phi(\mathbf{r})\right]. \end{aligned} \quad (2.8)$$

For partially compensated wavefronts, $D_\phi(\mathbf{r})$ starts growing as in the uncorrected case, but eventually levels off to a constant value $2\sigma^2$, where σ^2 is the variance of the remaining phase distortions. Hence, $A(\mathbf{r})$ can be considered as the sum of a seeing OTF scaled with $1 - \exp(-\sigma^2)$ and a constant term scaled with $\exp(-\sigma^2)$.

The Point Spread Function (PSF) is the Fourier transform of the OTF (see Roggemann and Welsh, 1996). Thus, also the PSF consists of a “seeing” part and a diffraction limited part. The relative amount of energy in these two components is given by $\exp(-\sigma^2)$. The smaller σ^2 is, the more the PSF resembles an Airy function which is the goal of an AO system.

2.1.2 Modal decomposition

The starting point of *modal control* is the decomposition of the wavefront phase $\phi(\mathbf{r}, t)$ into a basis set of functions $\{F_i(\mathbf{r})\}$ which are orthogonal and normalized on a unit circular or annular aperture with inner radius r_i , i.e.

$$\int d^2\mathbf{r} W(\mathbf{r}) F_i(\mathbf{r}) F_j(\mathbf{r}) = \delta_{ij}, \quad (2.9)$$

where δ_{ij} is the Kronecker delta symbol, and $W(\mathbf{r})$ is the pupil weighting function defined as

$$W(\mathbf{r}) = \begin{cases} \frac{1}{\pi(1-r_i^2)} & (|\mathbf{r}| \leq 1) \\ 0 & (|\mathbf{r}| > 1) \end{cases} \quad (2.10)$$

Then, the modal decomposition of the wavefront phase is given by

$$\phi(\mathbf{r}, t) = \sum_{i=1}^{\infty} a_i(t) F_i(\mathbf{r}). \quad (2.11)$$

The basis functions $F_i(\mathbf{r})$ describe the wavefront's spatial characteristics, while their weighting coefficients $a_i(t)$ describe the dynamics. In the following, the time dependencies will be dropped, since those are subject to dynamic compensation and of no particular interest in the static wavefront estimation process.

Using the orthogonality property of the set $\{F_i(\mathbf{r})\}$, the a_i are given by the projection

$$a_i = \int d^2\mathbf{r} W(\mathbf{r}) \phi(\mathbf{r}) F_i(\mathbf{r}). \quad (2.12)$$

Using Equations 2.9 and 2.11, the variance of the wavefront phase over the telescope aperture can be written as

$$\begin{aligned} \int d^2\mathbf{r} W(\mathbf{r}) \langle \phi(\mathbf{r})^2 \rangle &= \sum_{i,j} \langle a_i a_j \rangle \cdot \int d^2\mathbf{r} W(\mathbf{r}) F_i(\mathbf{r}) F_j(\mathbf{r}) \\ &= \sum_i \langle a_i^2 \rangle. \end{aligned} \quad (2.13)$$

The quantities $\langle a_i a_j \rangle$ are the elements of the *modal covariance matrix* $\mathbf{V}_a := \langle \mathbf{a} \mathbf{a}^t \rangle$. The diagonal elements of \mathbf{V}_a are the variances of the modes and describe their contribution to the atmospheric turbulence.

2.1.3 Zernike polynomials

Zernike polynomials are widely used to describe optical aberrations. They are mathematically well defined, and its low order terms correspond to classical aberrations like defocus, astigmatism, and coma. In his seminal paper, Noll (1976) discussed the properties of Zernike polynomials as a basis set for turbulent wavefronts. Zernike polynomials form a set of orthogonal polynomials on the unit circle and are most conveniently expressed in polar coordinates ρ and θ :

$$\begin{aligned} Z_{j_{\text{even}}} &= \sqrt{n+1} R_n^m(\rho) \sqrt{2} \cos(m\theta) & (m \neq 0) \\ Z_{j_{\text{odd}}} &= \sqrt{n+1} R_n^m(\rho) \sqrt{2} \sin(m\theta) & (m \neq 0) \\ Z_j &= \sqrt{n+1} R_n^0(\rho) & (m = 0), \end{aligned} \quad (2.14)$$

where

$$R_n^m(\rho) = \sum_{s=0}^{\frac{n-m}{2}} \frac{(-1)^s (n-s)!}{s! (\frac{n+m}{2} - s)! (\frac{n-m}{2} - s)!} \rho^{n-2s}. \quad (2.15)$$

The values of n and m satisfy $m \leq n$, and $n - m = \text{even}$. The index j is a mode ordering number derived from n and m . There is one mode of radial order $n = 0$, two of $n = 1$, three of $n = 2$, and so on. The total number of modes up to a given radial order is therefore

$$j_n = \frac{(n+1)(n+2)}{2}. \quad (2.16)$$

Table 2.1 shows the low order Zernike polynomials, where the columns and rows indicate azimuthal and radial order respectively, and Figure 2.3 displays them graphically.

Table 2.1: Zernike polynomials Z_j for $j = 1$ to 11. n is the radial order and m the azimuthal order. Even j correspond to symmetric modes ($\cos m\theta$) and odd j to antisymmetric modes ($\sin m\theta$).

n	$m = 0$	1	2	3
0	$Z_1 = 1$ (piston)			
1		$Z_2 = 2\rho \cos \theta$ $Z_3 = 2\rho \sin \theta$ (tip and tilt)		
2	$Z_4 = \sqrt{3}(2\rho^2 - 1)$ (focus)		$Z_5 = \sqrt{6}\rho^2 \sin 2\theta$ $Z_6 = \sqrt{6}\rho^2 \cos 2\theta$ (astigmatism)	
3		$Z_7 = \sqrt{8}(3\rho^3 - 2\rho) \sin \theta$ $Z_8 = \sqrt{8}(3\rho^3 - 2\rho) \cos \theta$ (coma)		$Z_9 = \sqrt{8}\rho^3 \sin 3\theta$ $Z_{10} = \sqrt{8}\rho^3 \cos 3\theta$ (trifoil)
4	$Z_{11} = \sqrt{5}(6\rho^4 - 6\rho^2 + 1)$ (spherical aberration)			

Noll (1976) has also derived the covariance matrix of the Zernike modal coefficients for Kolmogorov turbulence, whose first elements are displayed in Table 2.2 (Wang and Markey (1978) corrected a minor error in Noll's calculation). It is not purely diagonal, which implies some redundancy in the Zernike decomposition. Only modes of the same azimuthal order and symmetry (even or odd, see Table 2.1) are correlated. This redundancy, however, is small and can be completely removed by the use of Karhunen-Loève functions, which will be discussed in Section 2.1.5.

Table 2.2: Covariance matrix of the first 11 Zernike modes (excluding piston) calculated via a formula given by Wang and Markey (1978). The values are normalized to $(D/r_0)^{5/3} = 1$.

$\begin{pmatrix}$	0.448	0	0	0	0	0	-0.0141	0	0	0	0
	0	0.448	0	0	0	-0.0141	0	0	0	0	0
	0	0	0.0232	0	0	0	0	0	0	-0.0039	0
	0	0	0	0.0232	0	0	0	0	0	0	0
	0	0	0	0	0.0232	0	0	0	0	0	-0.0039
	0	-0.0141	0	0	0	0.0062	0	0	0	0	0
	-0.0141	0	0	0	0	0	0.0062	0	0	0	0
	0	0	0	0	0	0	0	0.0062	0	0	0
	0	0	0	0	0	0	0	0	0.0062	0	0
	0	0	-0.0039	0	0	0	0	0	0	0.0024	0
	0	0	0	0	-0.0039	0	0	0	0	0	0.0024
$\left. \vphantom{\begin{pmatrix}} \right)$											

Noll further showed that the infinite amount of energy in the Kolmogorov spectrum (cf. Equation 2.3) lies solely in the piston term. The variance of the piston-subtracted, i.e. mean-subtracted, wavefront over the unit circle is $1.03 (D/r_0)^{5/3} \text{rad}^2$ which is the trace of the Zernike covariance matrix. A least-squares fit to the diagonal elements yields

$$\langle a_j^2 \rangle \approx 0.294 j^{-1.9} (D/r_0)^{5/3} [\text{rad}^2]. \quad (2.17)$$

The residual phase variance after removing j lowest order Zernike modes is shown in Table 2.3. One can see that the first 10 Zernike modes carry about 96 % of the total wavefront disturbance. Noll (1976) showed that for large j this residual phase variance varies asymptotically with

$$\sigma^2(j) \approx 0.2944 j^{-\sqrt{3}/2} (D/r_0)^{5/3} [\text{rad}^2]. \quad (2.18)$$

Table 2.3: Residual variance Δ_j of the turbulent wavefront phase on a circular pupil after correction of j Zernike modes (normalized to $(D/r_0)^{5/3} = 1$). $\Delta_j - \Delta_{j-1}$ is the variance of mode j , i.e. the j th element on the diagonal of the covariance matrix.

j	1	2	3	4	5	6	7	8	9	10
Δ_j	1.030	0.582	0.134	0.111	0.088	0.068	0.0587	0.0525	0.0463	0.0401
$\Delta_j - \Delta_{j-1}$	∞	0.448	0.448	0.023	0.023	0.023	0.0062	0.0062	0.0062	0.0062

Strehl ratio. The corrected phase variance measures the performance of an AO system. Usually, it cannot be obtained directly, so the commonly used performance metric is the Strehl ratio S of an image. S is defined as the peak intensity of the aberrated PSF normalized to the peak of the perfect PSF, and is a function of the phase variance σ^2 . The Maréchal approximation (Born and Wolf, 1970), which is valid only for small variances, states that

$$S \approx \exp(-\sigma^2). \quad (2.19)$$

As a numerical example, let us assume that an AO system working for a 3.5-m-telescope perfectly corrects the first 10 Zernike modes. If r_0 in the K-band is 60 cm, corresponding to $1''$ seeing in the V-band, the variance of the corrected phase would be $0.0401(D/r_0)^{5/3} = 0.76 \text{ rad}^2$, and S would be 47%.

2.1.4 Annular Zernike polynomials

Most astronomical telescopes have annular rather than circular apertures, because the secondary mirror screens the central part of the primary. Zernike polynomials, as described in Section 2.1.3, are no longer orthonormal on such a support. Annular Zernike polynomials (see Figure 2.3) can be constructed from Circular Zernike polynomials by a Gram-Schmidt orthonormalization process. More precisely, since the annulus preserves the azimuthal symmetry, only the radial functions of the same azimuthal order are orthonormalized with respect to each other.

I have written a *Mathematica* code to generate Annular Zernike polynomials, using an algorithm described by Mahajan (1994). One also obtains a sparse matrix \mathbf{T} that transforms the Zernike basis set $\{Z_i(\mathbf{r})\}$ into the Annular Zernike basis set $\{A_i(\mathbf{r})\}$, i.e.

$$\begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_{jmax} \end{pmatrix} = \mathbf{T} \cdot \begin{pmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_{jmax} \end{pmatrix}. \quad (2.20)$$

Hence, modal coefficients of Annular Zernike polynomials $a_{az,i}$ can be derived from Circular Zernike coefficients $a_{z,i}$ via (see Bronstein et al., 1999, transformation of coordinate systems)

$$\mathbf{a}_{az} = \underbrace{(\mathbf{T}^{-1})^t}_{\mathbf{M}} \cdot \mathbf{a}_z. \quad (2.21)$$

The covariance matrix $\langle \mathbf{a}_{az} \mathbf{a}_{az}^t \rangle$ of Annular Zernike polynomials then follows from the covariance matrix $\langle \mathbf{a}_z \mathbf{a}_z^t \rangle$ of the circular Zernike polynomials by

$$\begin{aligned} \langle \mathbf{a}_{az} \mathbf{a}_{az}^t \rangle &= \langle (\mathbf{M} \mathbf{a}_z) (\mathbf{M} \mathbf{a}_z)^t \rangle \\ &= \mathbf{M} \langle \mathbf{a}_z \mathbf{a}_z^t \rangle \mathbf{M}^t. \end{aligned} \quad (2.22)$$

Table 2.4 shows the covariances of low order Annular Zernike polynomials. Obviously, the correlation pattern between the modes is similar to those of Circular Zernike polynomials displayed by Table 2.2.

Table 2.4: Covariance matrix of the first 11 Annular Zernike modes for a central obscuration of 0.39 calculated via Equation 2.22

$$\begin{pmatrix}
 0.5144 & 0 & 0 & 0 & 0 & 0 & -0.01417 & 0 & 0 & 0 & 0 \\
 0 & 0.5144 & 0 & 0 & 0 & -0.0147 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0.0167 & 0 & 0 & 0 & 0 & 0 & 0 & 0.0057 & 0 \\
 0 & 0 & 0 & 0.0271 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0.0271 & 0 & 0 & 0 & 0 & 0 & 0.0057 \\
 0 & -0.0147 & 0 & 0 & 0 & 0 & 0.0057 & 0 & 0 & 0 & 0 \\
 -0.0147 & 0 & 0 & 0 & 0 & 0 & 0 & 0.0057 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.0073 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.0073 & 0 \\
 0 & 0 & 0.0057 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.0015 \\
 0 & 0 & 0 & 0 & 0.0057 & 0 & 0 & 0 & 0 & 0 & 0.0027
 \end{pmatrix}$$

The secondary mirror of the Calar Alto 3.5-m-telescope has a diameter of 1.37 m, resulting in a central obscuration of $r_i = 0.39$ aperture diameters. Table 2.5 shows the residual phase variance as a function of the number of low order modes that have been removed. It also displays the mean contribution of these modes to the overall wavefront. When compared to Table 2.3, the uncorrected phase variance $\Delta_1 = 1.164$ is obviously larger when the central obscuration is considered. This effect is expected, because a piston removed wavefront ($\langle \phi(x, y) \rangle = 0$) is small in the middle and will raise towards the edges of the pupil. Further, the variances of modes of the same radial order are no longer degenerate and modes with a higher azimuthal order have larger variances.

Table 2.5: Normalized residual variance Δ_j of the turbulent wavefront phase on an annular pupil with $r_i = 0.39$ after correction of the first j Annular Zernike. $\Delta_j - \Delta_{j-1}$ is the variance of the mode j , i.e. the j th element on the diagonal of the covariance matrix.

j	1	2	3	4	5	6	7	8	9	10
Δ_j	1.164	0.650	0.135	0.1187	0.0916	0.0644	0.0588	0.0531	0.0458	0.0386
$\Delta_j - \Delta_{j-1}$	∞	0.514	0.514	0.0167	0.0271	0.0271	0.0057	0.0057	0.0073	0.0073

2.1.5 Karhunen-Loève functions

Karhunen-Loève functions (K-L) were introduced into the context of modal compensation of atmospheric turbulence phase distortion by Wang and Markey (1978). In the following, the basic approach to the K-L expansion is summarized.

For a complex orthonormal set of functions $\{F_i(\mathbf{r})\}$, a wavefront can be expressed as

$$\phi(\mathbf{r}) = \sum_i a_i F_i(\mathbf{r}), \quad (2.23)$$

where the coefficients of the series a_i are stochastic quantities (see Section 2.1.2). The requirement that these coefficients be statistically independent implies the following constraint:

$$\langle a_i a_j^* \rangle = \Lambda_i^2 \delta_{ij}, \quad (2.24)$$

where Λ_i^2 is the variance of the i th mode. Based on the orthogonality property of the set $\{F_i(\mathbf{r})\}$,

$$\int d^2\mathbf{r} W(\mathbf{r}) F_i^*(\mathbf{r}) F_j(\mathbf{r}) = \delta_{ij}, \quad (2.25)$$

it follows with Equation 2.23 that

$$a_i = \int d^2\mathbf{r} W(\mathbf{r}) F_i^*(\mathbf{r}) \phi(\mathbf{r}). \quad (2.26)$$

Multiplying both sides of Equation 2.26 by $\phi(\mathbf{r})$, taking an ensemble average, and using Equations 2.23 and 2.24, one obtains:

$$\int d^2\mathbf{r}' W(\mathbf{r}') F_i^*(\mathbf{r}') \langle \phi(\mathbf{r}') \phi(\mathbf{r}) \rangle = \Lambda_i^2 F_i(\mathbf{r}). \quad (2.27)$$

This is a two-dimensional K-L integral equation that defines the K-L eigenfunctions $\{F_i(\mathbf{r})\}$ for a general covariance of a stochastic process. It is important to note that K-L functions are not a fixed basis set on a circular support as in the case of Zernike polynomials. They differ by the underlying statistics of the stochastic process and the support for which they are computed.

The actual evaluation of K-L functions for Kolmogorov statistics can be found in Wang and Markey (1978) for circular pupils, and in Cannon (1996) for an annulus. Robert Cannon also provides a software package written in *IDL*. This package is a formidable tool for the simulation of turbulent wavefronts, and was used for the simulations presented in this work. Since K-L functions are statistically independent, a wavefront is simply a random combination of K-L modes weighted by their contributions Λ_i .

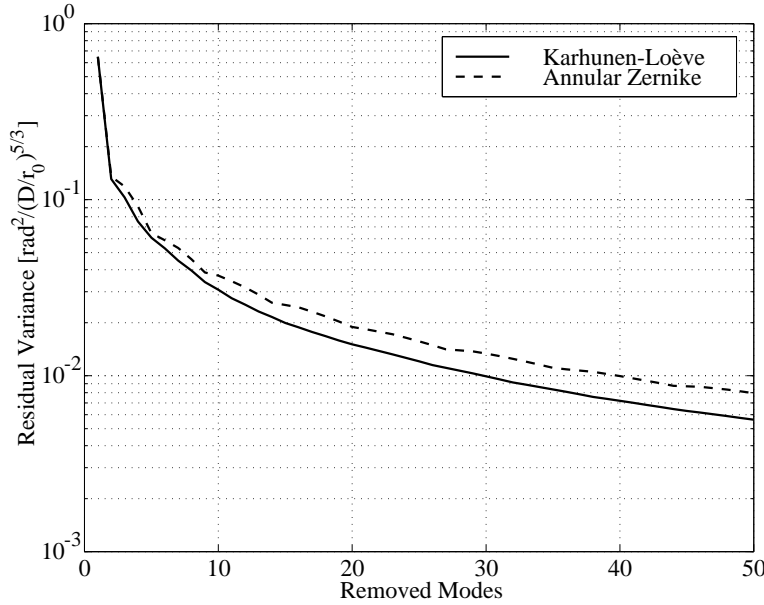


Figure 2.2: Normalized variance of the residual wavefront for Kolmogorov turbulence conditions (D : aperture diameter, r_0 : Fried parameter) after correction of a certain number of modes.

K-L functions carry the maximum amount of information about a wavefront, if a limited number of modes are used only. Thus, they are superior to Zernikes polynomials in removing turbulent wavefront aberrations. This is demonstrated in Figure 2.2, which displays the residual phase variance as a function of the number of corrected K-L or Annular Zernike modes. The variances of low order K-L

Table 2.6: Residual variance Δ_j of the turbulent wavefront phase on an annular pupil with $r_i = 0.39$ after correction of j K-L modes. $\Delta_j - \Delta_{j-1} = \Lambda_j^2$ is the variance of mode j .

j	1	2	3	4	5	6	7	8	9	10
Δ_j	1.164	0.649	0.133	0.105	0.077	0.063	0.055	0.047	0.042	0.036
$\Delta_j - \Delta_{j-1}$	∞	0.515	0.515	0.0279	0.0279	0.0144	0.0078	0.0078	0.0055	0.0055

modes for Kolmogorov turbulence are shown in Table 2.6. It should be noted that deviations from the Kolmogorov model annihilate this advantage. As will be shown in Section 3.4, there are other more important effects which favor K-L functions as a basis set of control modes.

Low order K-L functions differ only marginally from Zernike polynomials as can be seen in Figure 2.3. This is expected, because the covariance matrix of Zernike polynomials is almost orthogonal and thus Zernikes are almost optimal. The differences become more obvious at higher radial degrees, where Zernikes tend to remain small in the inner part and steeply rise at the edges of the pupil, while K-L functions vary more smoothly.

2.1.6 Mirror modes

Mirror modes form a basis set for the part of the phase space that can be affected by the DM. They depend on the mechanical characteristics of the device actually employed. Here, the analysis is restricted to Discrete-Actuator DMs, but the results can easily be applied to other types.

The basic structure of a Discrete-Actuator DM consists of a stiff baseplate, a flexible facesheet with the reflective coating, and the connecting actuators between the two. The stress and strain of the mirror materials result in structural coupling between the components and usually limits the utility of a particular DM.

The actual shape of a DM surface, pushed by a single actuator, is called *influence function* of the actuator. Depending on the stiffness of the facesheet compared to the stiffness of the actuators, there will be some coupling between the actuators, i.e. pushing a specific actuator will result in some deformation of the facesheet at the positions of adjacent actuators. This deformation is usually below 10 % of the actuator stroke, and the influence functions are quite well approximated by Gaussian functions (see Tyson, 1991, p. 203).

Let now n_{act} be the number of actuators, and \mathbf{M} be a $n_{act} \times n_{act}$ matrix whose columns represent the influence functions of all actuators sampled on the actuator grid. The set of eigenvectors $\{\mathbf{v}_i\}$ of \mathbf{M} are the eigenmodes of the DM, or shortly the mirror modes, i.e.

$$\mathbf{M}\mathbf{v}_i = \lambda_i\mathbf{v}_i, \quad (2.28)$$

where the λ_i s are the eigenvalues of \mathbf{M} . If there is no coupling between the actuators, \mathbf{M} is the identity matrix, and any actuator pattern is an eigenvector of \mathbf{M} , i.e. a mirror mode. All analyses in this work neglect a possible coupling between DM actuators, and thus do not consider the mechanical characteristics of the actual DM used in ALFA.

2.2 Wavefront sensing methods

Four different types of high order wave-front sensors are discussed in this section. The SHS is presented first, because it is the most common device for AO applications and the type which is implemented in ALFA. Its attractiveness is based on its conceptual simplicity and the long experience

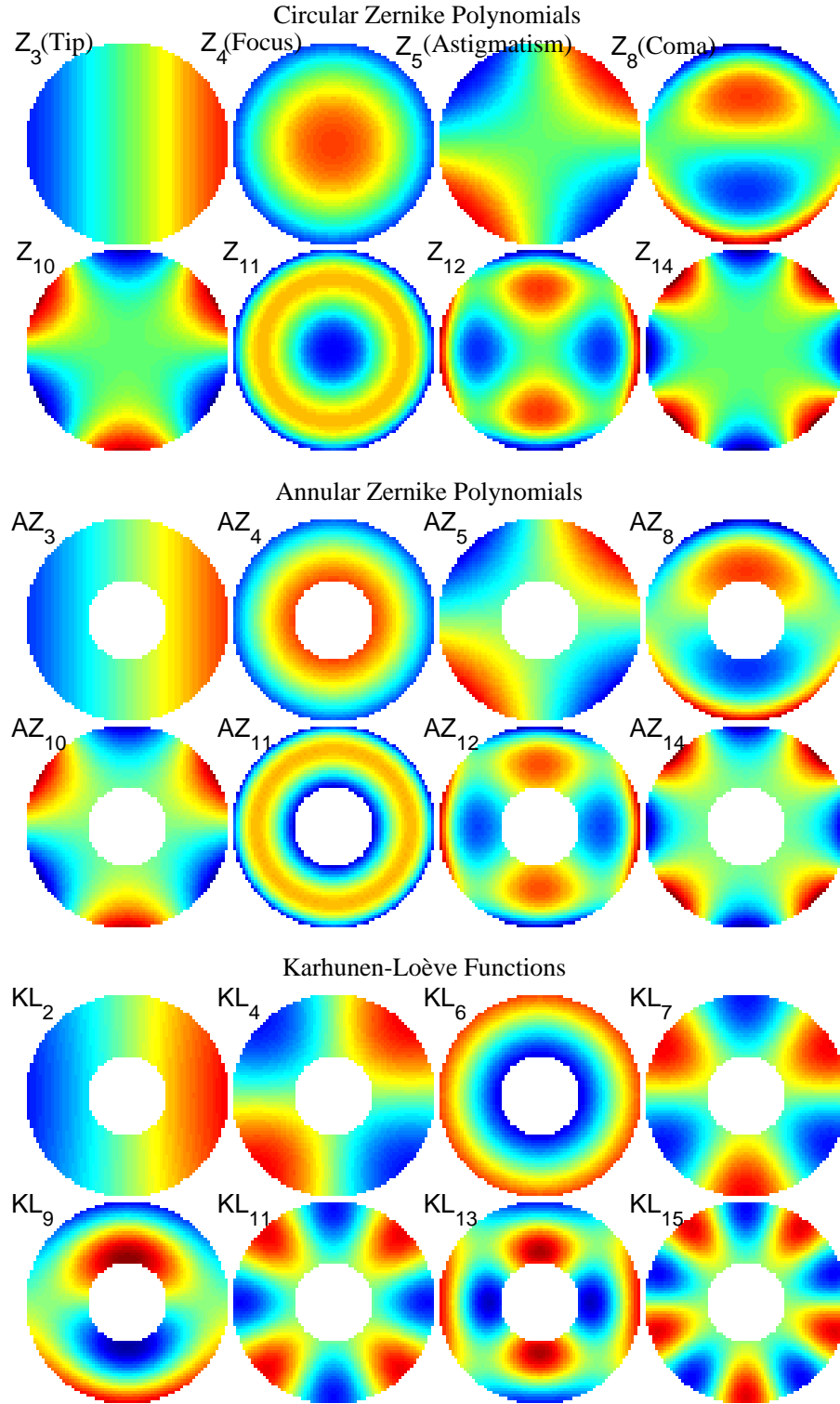


Figure 2.3: Low order Zernike polynomials, Annular Zernike polynomials, and K-L functions for Kolmogorov turbulence, the latter two calculated for $r_i = 0.39$. Piston is excluded, and only one mode of pairs with azimuthal order above zero is plotted.

with this “classic” device. The Pyramid- or Foucault-like wavefront sensor (PWS) described in Section 2.2.2 is a new idea (Ragazzoni, 1996), inspired by the well known Foucault knife-edge test. Like the SHS, the PWS measures first derivatives of the wavefront, i.e. wavefront slopes. The Curvature Wavefront Sensor (CWS) described in Section 2.2.3 measures second derivatives. Curvature sensing is also a technique that was developed recently (Rodier, 1988). The CWS is now almost as popular as the SHS for use in astronomical AO systems, and it has demonstrated its excellent performance in low order systems. The last technique to be discussed briefly is Phase Diversity (PD) (Paxman et al., 1992). PD uses the information content of two images, differing by a known aberration, to obtain the wavefront distortion and object spectrum via an optimization scheme. This technique has not yet been used in AO systems because of its high computational effort, but this may change with the increased speed of modern computers.

2.2.1 Shack-Hartmann sensor

The Hartmann method (1900) for testing a lens employs an opaque mask with holes placed behind the lens. Each of the holes acts as an aperture, and the light passing through the lens eventually produces an image. The positions of these images represent local wavefront slopes corresponding to the holes. Shack’s idea was to use small lenses which are arranged in an array to measure wavefront aberrations. The SHS subdivides the wavefront into an array of small subapertures, and a microlens array forms multiple images (see Figure 2.4). The angular displacement α_x of each of these subimages with respect to a pre-defined reference position is an estimate for the average wavefront slope in x -direction over the subaperture (Primot et al., 1990), i.e.

$$\alpha_x = \frac{c_x}{fF} = \frac{\lambda}{2\pi A_{sa}} \int_{sa} \frac{\partial \phi(x,y)}{\partial x} dx dy, \quad (2.29)$$

where A_{sa} denotes the area of the subaperture, c_x the spot displacement, f the focal length of the microlens, and F the magnification between the microlens plane and the telescope entrance plane. Equation 2.29 shows that in principle an SHS always responds linearly to variations of the slope vectors. Since the SHS measures the first derivative of the wavefront, it is insensitive to a constant offset of the wavefront (piston). This is, however, irrelevant for single apertures, because constant offsets do not affect image quality.

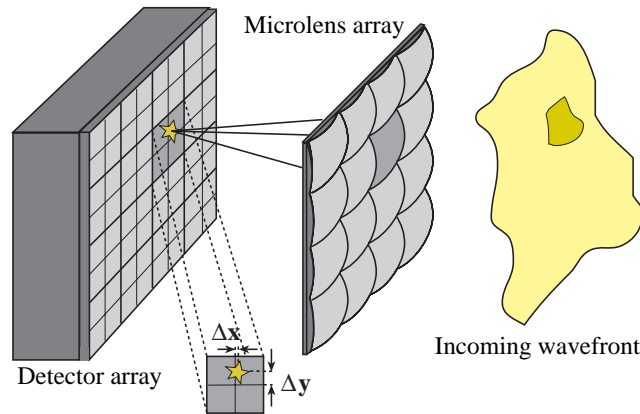


Figure 2.4: Measurement principle of a Shack-Hartmann sensor. The incoming wavefront is subdivided by the microlens array, and image centroids are shifted according to average wavefront slopes.

Various algorithms can be used to derive a spot's position $\mathbf{c} = (c_x, c_y)$, some of which are studied in Section 3.2. When a detector array like a CCD is used, these algorithms are usually modifications of a simple center of mass estimation

$$c_x = \frac{\sum_i x_i I_i}{\sum_i I_i}, \quad (2.30)$$

where x_i is the x-coordinate of the i th pixel position and I_i is its intensity, c_y being calculated in the same way. Fast fitting algorithms require a priori knowledge of the shape of the PSF (parabolic, Gaussian, ...), sometimes being a poor approximation. Iterative algorithms are generally too slow to be used in closed-loop.

If a quad-cell detector is used instead of an array, the measured angular displacement of the spot can be expressed by (see Roddier, 1999, p. 100):

$$\alpha_x = \frac{\theta_*}{2} \frac{I_1 + I_2 - I_3 - I_4}{I_1 + I_2 + I_3 + I_4}, \quad (2.31)$$

where θ_* is the angular spot size, i.e. θ_* equals λ/r_0 in the seeing-limited case and λ/d in the diffraction-limited case. I_1, \dots, I_4 are the intensities detected in the four quadrants. Equation 2.31 shows that the sensor signal, which is the fraction involving only intensity sums, does not depend on the focal length, because it is given by the ratio of the two angles α_x and θ_* . Thus, the SHS lacks flexibility in changing its gain or dynamic range.

The one-axis rms fluctuation σ_x of turbulent wavefront slopes over a subaperture of diameter d is the average strength of SHS signals. An accurate approximation can be found by using Equation 2.1:

$$\begin{aligned} \sigma_x &= \frac{\lambda}{2\pi} \frac{\sqrt{D_\phi(d)}}{d} \\ &= 0.42 \frac{\lambda}{r_0} \left(\frac{d}{r_0} \right)^{-\frac{1}{6}}. \end{aligned} \quad (2.32)$$

σ_x depends only weakly on the size of the subapertures; under typical conditions, it is of the order of a few tens of an arcsecond. An important consequence of Equation 2.32 is that the SHS is achromatic, since σ_x is independent of λ (remember that $r_0 \propto \lambda^{6/5}$, Equation 2.2). Therefore, it is possible to sense wavefronts over a large spectral bandwidth, e.g. at wavelengths not used for imaging.

The spatial PSD of wavefront slopes is proportional to $k^{-5/3}$. In the temporal domain, wavefront slopes exhibit a power spectrum which slowly decreases with $f^{-2/3}$ up to a breakpoint and afterwards rolls off rapidly with $f^{-11/3}$ (Conan et al., 1995). As demonstrated by Equation 2.5, this breakpoint depends on windspeed and aperture size. At high frequencies, spatial filtering effects by aperture averaging dominate.

Applicability. The SHS works perfectly with extended sources. Even the solar surface can be used for wavefront sensing, if a field stop and image correlation techniques are used to measure spot displacements. The use of quad-cells limits the linear range of an SHS to about $\pm(\theta_*/2)$ (see Hardy, 1998, p. 144), and extended or multiple sources have to be used with caution. On the other hand, it is demonstrated by Equation 2.32 that rms image motion is usually smaller than $(\theta_*/2)$, if the subapertures are larger than r_0 . In closed-loop, where the SHS only measures residual image motion after correction, it will certainly be much smaller than $(\theta_*/2)$.

2.2.2 Pyramid sensor

The PWS was proposed by Ragazzoni (1996) for the first time. Like the SHS, it provides a simple linear relationship between measurement signal and wavefront slopes. The PWS can adjust its gain during operation in order to optimize the behavior of the sensor, depending on observing conditions. The principle of operation is sketched in Figure 2.5.

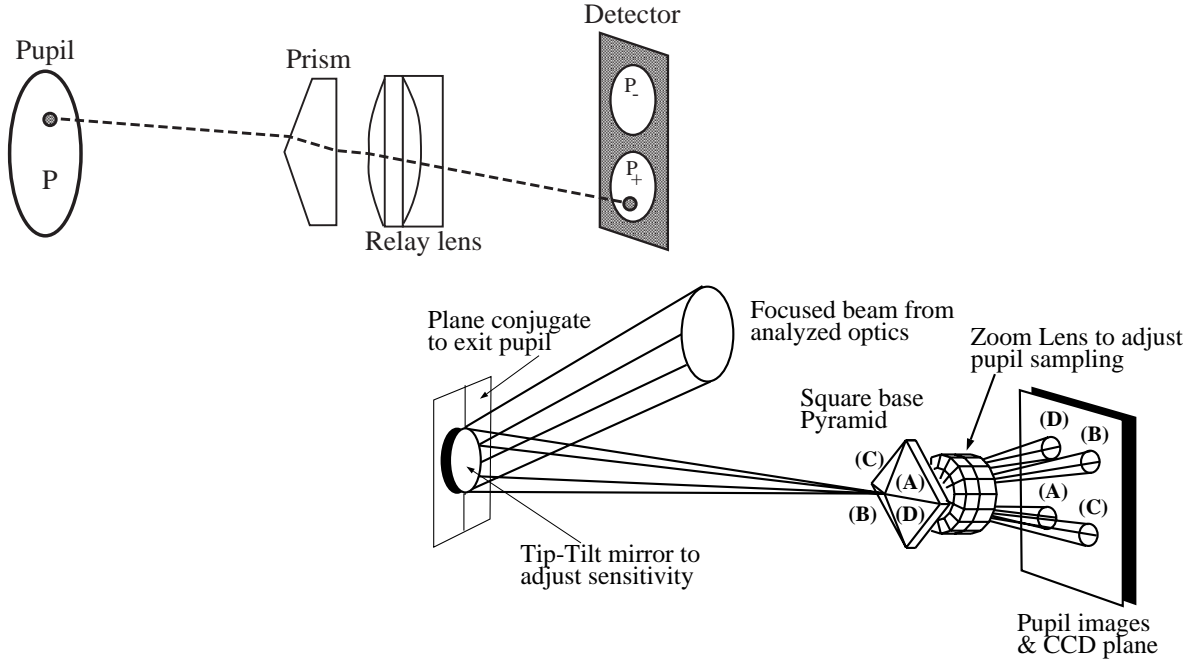


Figure 2.5: *Top:* Principle of the PWS. The focal plane prism and the relay lens create multiple pupil image. *Bottom:* Practical Realization (courtesy of Esposito et al. (2000)). The pupil sampling can be varied with the magnification provided by a zoom lens. The oscillation is done by a steering mirror in the pupil plane.

Assuming that the geometrical optics approximation is valid, an aberrated ray hits the prism on either side of its tip and appears only in one of the multiple pupils as displayed by the upper drawing. The intensity distributions in the multiple pupil images are therefore a measure for the sign of the ray's slope. If the prism oscillates, or equivalently a tip-tilt mirror in the pupil plane moves the image as shown in the lower drawing, the ray will appear in either of the pupil images depending on the modulus of its slope. Thus, the intensity distribution integrated over a couple of oscillations measures wavefront slopes in the pupil. Assuming a linear angular modulation of the tilt (ramp modulation), the following linear relation between the sensor signals ($I_+(x, y)$, $I_-(x, y)$) and the wavefront slopes follows (Riccardi et al., 1998):

$$\frac{\partial \phi}{\partial y}(x, y) = A \frac{I_+(x, y) - I_-(x, y)}{I_+(x, y) + I_-(x, y)}, \quad (2.33)$$

where A is the amplitude of the oscillation.

The PWS principle is highly flexible - the sampling can easily be adjusted by a zoom lens, and the amplitude of the oscillation controls the sensor gain. It is achromatic and works with extended sources. The PWS has not been used in an AO system yet, but it has been investigated and tested in the laboratory with promising results (Riccardi et al., 1998, Esposito et al., 2000). Although the PWS

senses wavefront slopes like the SHS, the ad-hoc assumption that both exhibit a similar error behavior only holds in the open-loop case. Once the loop is closed, and the reference object becomes more narrow, the PWS error propagation is vastly superior (Ragazzoni and Farinato, 1999) at least for very high order systems. This will be discussed more deeply in Section 2.2.5.

2.2.3 Curvature sensor

The curvature wavefront sensor (CWS) was invented by François Roddier (Roddier, 1988). His idea was that intensity distributions in two different planes on either side of the focus are a measure for a wavefront's curvature, i.e. its second derivative, in the telescope pupil. To readily reconstruct a wavefront, the Poisson equation has to be solved. To accomplish this, von Neumann type boundary conditions are required, which are imposed by radial wavefront tilts on the pupil edge.

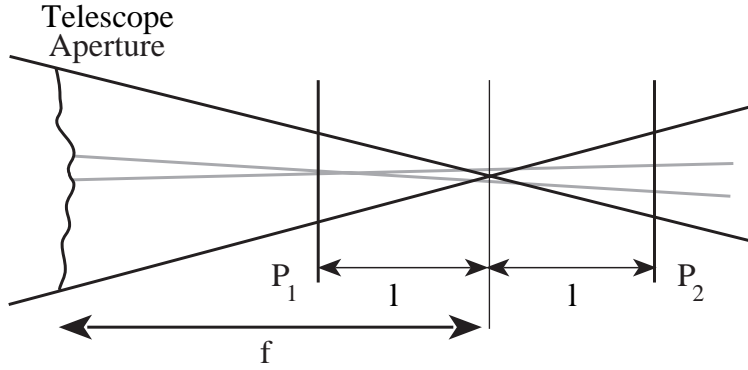


Figure 2.6: Principle of the curvature sensor. The gray lines show the rays from a curved part of the wavefront that forms a focus before the focal plane, leading to a local increase in intensity in plane P_1 and a decrease in P_2 .

The principle is sketched in Figure 2.6. Two intensity distributions are recorded in the planes P_1 and P_2 , at a distance l before and behind the telescope focal plane. These intensity distributions can be considered as defocused images of the telescope pupil. The figure displays the effect of a local curvature of the wavefront: the curved wavefront leads to a local excess of illumination in plane P_1 and to a lack of illumination at the corresponding position in P_2 as the rays are diverging. On the pupil edge, the same difference between the two illuminations provides a measure of the radial tilt.

Figure 2.7 shows a representative sampling geometry for a low order system. The intensities are measured in each of the seven segments and, additionally, in the six areas that are cut out from the segments by the dashed line. The latter measurements provide the edge slopes, which are required to reconstruct the zero curvature modes tip-tilt and astigmatism.

In the geometrical optics approximation, the normalized difference between the two intensity distributions is (Roddier, 1988):

$$c(\mathbf{r}) = \frac{I_1(\mathbf{r}) - I_2(\mathbf{r})}{I_1(\mathbf{r}) + I_2(\mathbf{r})} = \frac{f(f-l)}{2l} \left[\frac{\partial}{\partial \rho} \phi(\rho, \theta) \delta_c - \nabla^2 \phi(\rho, \theta) \right], \quad (2.34)$$

where δ_c is a linear impulse distribution around the pupil edge, I_1 and I_2 are intensities in the two planes, and ∇^2 is the Laplacian. The CWS signal $c(\mathbf{r})$ contains all the information needed to reconstruct the wavefront, and it can be amplified by reducing l , which makes gain adjustments possible. The CWS is achromatic and works within limits with extended sources.

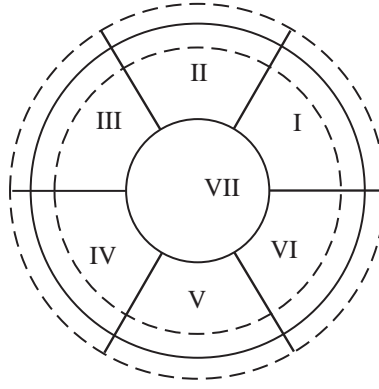


Figure 2.7: Sampling geometry for a low order curvature sensor. The illumination is integrated over each of the 7 segments. The dashed lines display the width of the impulse distribution δ_c (cf. Equation 2.34).

Equation 2.34 is only valid within the geometrical optics approximation, which holds as long as

$$\theta_b(f - l) < \frac{dl}{f}, \quad (2.35)$$

where θ_b is called the blur angle, and d is the size of a subaperture. Physically, the blur angle describes the angular uncertainty of light-rays in a subaperture, caused either by seeing ($\theta_b = \lambda/r_0$) or an extended object ($\theta_b = \text{FWHM}_{obj}$). Thus, the illumination in for example P_1 can be considered as a blurred pupil image with a lateral blur size of $\theta_b(f - l)$. This blur has to be small compared to the size of the wavefront region for which one would like to measure the curvature, i.e. the size of a subaperture in P_1 ($d \cdot l/f$). θ_b is generally much smaller than d/f , so that Equation 2.35 can be approximated by

$$l > \theta_b \frac{f^2}{d}. \quad (2.36)$$

To get a feeling for the involved quantities, the CWS AO system PUEO operated at the CFHT on Mauna Kea with an effective focal length of 216 m uses extra focal distances between 5 cm and 40 cm (O. Lai, priv. comm.).

If a CWS is operated in the visible where no major improvements of the PSF in closed-loop can be expected, θ_b will not be much smaller than the seeing. Therefore, the sensor gain can only be increased to a limit set by the minimum l . A further enhancement could be achieved at the expense of a coarser wavefront sampling (larger d). l must be increased for high order systems (small d) or extended objects, reducing sensor signal and sensitivity of the CWS.

If a CWS performs extremely well in closed-loop, such that θ_b shrinks even in the visible, it has the potential to further increase the sensor signals, because l can then be reduced. In practise, it is sufficient to know what size of θ_b can be expected in closed-loop and setting l accordingly. The feedback, even though initially extremely non-linear in some cases, ensures that the loop always closes (O. Lai, priv. comm.). This flexibility of the CWS is a big advantage compared to the SHS.

Local wavefront curvatures exhibit a spatial PSD that is proportional to $k^{-1/3}$, and are therefore almost decorrelated at two different points in the pupil (Roddier, 1988). This led Roddier to speculate that curvature measurements are less redundant than the correlated slope measurements, which are proportional to $k^{-5/3}$. The curvature power law also indicates aliasing of high spatial frequencies. On the other hand, high frequencies are attenuated by diffraction blur and averaging over the subapertures. In a comparative study of CWS and SHS, Rigaut et al. (1997) showed that potentially both sensors

perform equally well at all light levels. The operational CWS systems have the advantage of measuring only intensities in the subapertures. This work is perfectly done by a single Avalanche Photo Diode (APD) per subaperture, which has the same sensitivity as a CCD, but virtually zero read-out noise.

2.2.4 Phase diversity

The last technique to be discussed aims at deriving the wavefront phase from the intensity distribution in the image. This at first obvious approach imposes some difficulties, because the PSF is given by the Fourier transform of the autocorrelation function of the phase in the pupil. The retrieval of the phase from its autocorrelation function is a non-trivial problem.

Phase estimation using a point source in conjunction with knowledge of the pupil function is generally called *phase retrieval*. This approach was applied to images of stars obtained with the HST to diagnose errors in its mirror figure. The basic phase retrieval algorithm uses the modal expansion of a wavefront. An initial choice of the modal coefficients (usually all zero) allows one to calculate the corresponding OTF. This trial OTF is then compared to the measured OTF to derive some error metric, e.g. the mean square difference. Based on this difference a new guess for the modal coefficients is calculated, and the final solution is found by an iterative optimization process.

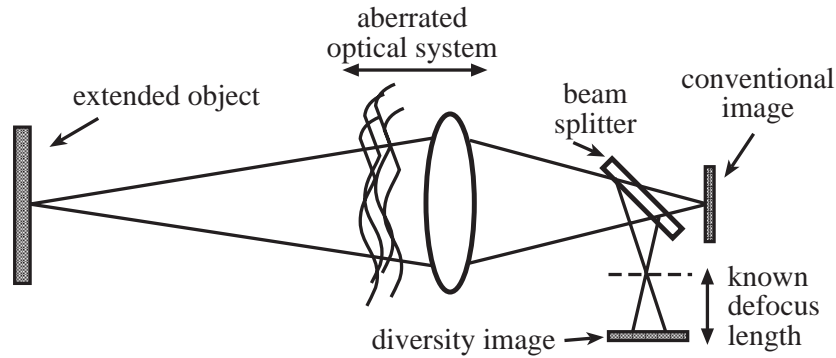


Figure 2.8: Measurement principle of Phase Diversity technique. Two images of the same object with a known phase diversity (e.g. focus difference) between the two are taken.

This method does not work with extended objects, because in this case the Fourier transform of an image does not resemble the OTF any more (it actually happens quite often, that an apparent point source turns out to be a binary when the AO loop is closed). Gonsalves (1982), however, found an error metric whose minimization yields estimates for the phase and the object simultaneously. Figure 2.8 sketches this concept which is called *phase diversity* (PD): two images are taken instantaneously with a known phase diversity between them. Usually, one of these images is the conventional focal-plane image, while the other is taken with a known amount of defocus resulting in a known quadratic phase difference between the two. Expressed in Fourier space, Gonsalves' error metric is

$$E = \sum_k [|I_1(k) - \hat{O}(k)\hat{P}_1(k)|^2 + |I_2(k) - \hat{O}(k)\hat{P}_2(k)|^2], \quad (2.37)$$

where $I_1(k)$, $I_2(k)$ are the spectra of the recorded images, $\hat{O}(k)$ is an estimate of the object spectrum, and $\hat{P}_1(k)$, $\hat{P}_2(k)$ are estimates for the two OTFs. Gonsalves further presented an analytical solution for $\hat{O}(k)$ to minimize E , while the proof was given by Paxman et al. (1992). Since the two OTFs

differ by a known amount, the minimization of E by a proper choice of the OTF through the modal coefficients resembles the phase retrieval problem.

The applicability of Phase Diversity for closed-loop AO has been investigated by several authors (Löfdahl et al., 1998, Carrano et al., 1998) theoretically and in the laboratory, but has not yet been tested at a telescope. Drawbacks are the high computational demands of the optimization procedure and the small spectral bandwidth, required to ensure that the PSF is not altered by chromatic effects. Also it is not assured that the optimization scheme converges towards the correct minimum, especially in the case of large phase aberrations. Advantages of a PD wavefront sensor are its very simple optical setup, and easy calibration.

2.2.5 Measurement error

The two fundamental and unavoidable measurement noise sources are signal photon noise and sky background photon noise. The sky background usually plays a minor role for wavefront sensing in the visible, but it can become important in the near-infrared (NIR). Rousset et al. (1993) demonstrated that in general all wavefront sensors have a similar behavior in terms of signal photon noise:

$$\sigma_s^2 \propto \frac{1}{n_{ph}} \left(\frac{\theta d}{\lambda} \right)^2. \quad (2.38)$$

σ_s^2 is the phase measurement error variance in rad^2 , θ is the angular size of the reference object (due to the object's inherent size, seeing, or diffraction), and n_{ph} is the number of photoelectrons per subaperture and exposure. Generally, when d and λ are fixed, the best sensing accuracy is obtained on bright and narrow reference objects. It is interesting to note that according to Equation 2.38 wavefront measurements under low signal conditions cannot be improved by using larger subapertures, since $n_{ph} \propto d^2$.

Propagation of measurement error. In contrast to the measurement error, error propagation onto reconstructed wavefronts varies with sensor type and reconstruction algorithm (see Section 2.3.6).

For the SHS, the total reconstruction error for the zonal approach, using a least-squares algorithm, scales with $\ln N \cdot \sigma_s^2$, where N is the number of subapertures (Fried, 1977, Hudgins, 1977). For the modal control approach, Rigaut and Gendron (1992) calculated the asymptotic behavior of the error propagation coefficients p_j for Zernike polynomials. The p_j describe how the reconstruction of the j th mode is affected by the measurement noise. Using Noll's formulation of the Zernike derivatives (Noll, 1976), they obtained

$$\begin{aligned} p_j &= 0.295(n_j + 1)^{-2.05} & (n_j = m_j) \\ p_j &= 0.174(n_j + 1)^{-2} & (n_j \neq m_j), \end{aligned} \quad (2.39)$$

where n_j, m_j are radial and azimuthal orders of the polynomial. Recalling that the actual ordering number j of some mode with radial degree n_j is of the order $j \sim n_j^2/2$ (Equation 2.16), we see that the p_j 's are asymptotically proportional to j^{-1} . The total reconstruction error variance σ_t^2 is therefore given by

$$\sigma_t^2 = \sigma_s^2 \sum_{j=1}^N p_j \propto \sigma_s^2 \sum_{j=1}^N \frac{1}{j}. \quad (2.40)$$

The last sum is the *harmonic number* H_N , which is asymptotically proportional to $\ln N$. Hence, the modal analysis confirms the zonal result of Fried and Hudgins. The modal approach, however, has

the big advantage that, depending on observing conditions (guide star brightness, seeing), modes can be filtered and a minimum error propagation can be found. It should be noted that low order modes are the strongest contributors to atmospheric turbulence, although they have a larger error propagation coefficient (see Equation 2.17). The net effect is that their signal to noise ratio (SNR) is still better than the one of high order modes (Gendron and Lena, 1994).

For the CWS, error propagation increases proportionally to N (Rodier et al., 1988). Therefore, the SHS appears to be better suited for very high order systems. Rodier (1995) instead showed, that this drawback can be more than compensated in a closed-loop system. In this case, the compensated wavefront leads to a better image of the reference object and, according to Equation 2.38, to a smaller measurement error. For the CWS, only diffraction effects at the telescope aperture edges are of interest (the subapertures of the CWS do not form an image, they just collect photons). Thus, the initial blur angle θ of the order λ/r_0 approaches λ/D when the loop is successfully closed, i.e. σ_s^2 is smaller by a factor $(D/r_0)^2 = N$. The SHS instead has to deal with diffraction at the subapertures. Therefore, the spot-size is limited to $\theta = \lambda/d$, where d is of the order r_0 for diffraction limited performance.

Like the SHS, the PWS measures wavefront slopes, but it works on the whole pupil like the CWS. Therefore the PWS combines the advantages of SHS and CWS, and is potentially the best choice for high order AO systems. A very high order PWS system for an extremely large telescope can be more sensitive than an SHS system by about 2.5 magnitudes in guide star brightness (Ragazzoni and Farinato, 1999).

Remarks. The potential superiority of PWS and CWS for closed-loop application does not hold for an LGS which is intrinsically extended and will not shrink much once the loop is closed. It is also an open question whether the CWS or PWS will be able to capture the diffraction limited reference from the initially seeing limited reference, and how robust the situation is for periods of worse compensation. Finally, spot shrinking in closed-loop can only be achieved under superior observing conditions in the visible, but it is gorgeous for NIR wavefront sensing.

2.3 Phase reconstruction with the Shack-Hartmann sensor

In the preceding section, generic properties of different wavefront sensors were evaluated by studying their asymptotic behavior for the case of very high order systems and standard reconstruction algorithms. The issue of this section is a detailed analysis of an SHS with a finite number of subapertures leading to ambiguities in the relationship between modes and slope vectors. Although the modal basis sets are orthonormal, their slope vectors are neither orthogonal nor normalized as demonstrated by Figure 2.9. The sensor model as formulated in Section 2.3.1 uses the concept of the *interaction matrix* which maps from wavefront to SHS measurement space to analyze cross-talk of control modes in Section 2.3.2. Cross-talk forces the truncation of the number of low order K-L or Zernike modes used for compensation. The *sensor modes* described in Section 2.3.3 are an extension to these low order modes. The criterion for their selection is minimization of cross-talk.

Uncontrolled wavefront modes introduce another fundamental error source to low-order systems, the *aliasing noise* investigated in Section 2.3.4. By consideration of aliasing, Section 2.3.5 shows how the modal covariance matrix can be estimated from SHS data. The prior knowledge of modal covariance and measurement noise is required to calculate advanced estimators for the wavefront modes. The properties of these estimators are studied in Section 2.3.6.

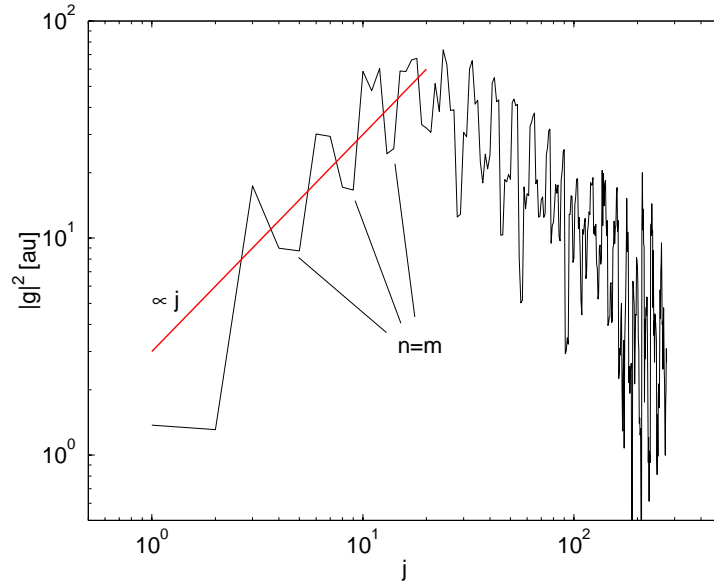


Figure 2.9: Squared length of Zernike slope vectors measured with a 56 subaperture SHS. It grows linearly with the mode number, but reaches a maximum at roughly half the number of subapertures and rolls off afterwards. Modes with identical radial and azimuthal order ($n = m$) produce smaller SHS signals.

2.3.1 Sensor measurement model

The SHS samples the wavefront discretely, i.e. it averages the wavefront over some region (the “subapertures”) to provide a slope vector \mathbf{g} of finite dimension. More precisely, the dimension of \mathbf{g} is twice the number of subapertures, because each subaperture contributes an x- and a y-slope. The *interaction matrix* \mathbf{D}_∞ is defined as the matrix given by the slope vectors of the modal basis set $\{F_i(\mathbf{r})\}$ arranged in its columns. Since the slope vectors depend linearly on the modal coefficients \mathbf{a} (see Section 2.2.1), the measurement process is modelled by the linear relation:

$$\begin{aligned}\mathbf{g} &= \mathbf{D}_\infty \mathbf{a}_\infty + \mathbf{n} \\ &= \mathbf{D} \mathbf{a} + \mathbf{D}_\perp \mathbf{a}_\perp + \mathbf{n},\end{aligned}\tag{2.41}$$

Slope vectors, noise $\mathbf{g}, \mathbf{n} \in R^{n \times 1}$

Modal coefficients $\mathbf{a}_\infty = \begin{bmatrix} \mathbf{a} \\ \mathbf{a}_\perp \end{bmatrix} \in R^{m \times 1}$
 $\mathbf{a}_\perp \in R^{\infty \times 1}$

Interaction matrix $\mathbf{D}_\infty = [\mathbf{D} \ \mathbf{D}_\perp]$, $\mathbf{D} \in R^{n \times m}$, $\mathbf{D}_\perp \in R^{n \times \infty}$.

The modes are split into two groups: One is formed by the finite number m of *control modes*, and the other is formed by the infinite number of complementary modes. \mathbf{D} is the sensor response to the control modes and \mathbf{a} are their weighting coefficients; \mathbf{D}_\perp is the sensor response to the uncontrolled complementary modes weighted by \mathbf{a}_\perp . n is the number of slopes, and \mathbf{n} the SHS measurement noise generally modelled as a Gaussian uncorrelated signal.

To reconstruct a wavefront represented by the modal coefficients \mathbf{a} , the inverse problem to Equation 2.41 has to be solved. This means that a matrix \mathbf{R} has to be found to derive the estimate

$$\mathbf{a} \approx \hat{\mathbf{a}} = \mathbf{R} \mathbf{g}.\tag{2.42}$$

The issue of how \mathbf{R} should actually be selected is closely examined in Section 2.3.6.

2.3.2 Modal cross-talk

Cross-talk occurs, if the column vectors of \mathbf{D} lacks orthogonality (Herrmann, 1981) (note that Herrmann refers to cross-talk as aliasing, while he calls the aliasing phenomenon cross-talk). The slope vector of one mode can then easily be shifted into another mode's vector by small noise vectors \mathbf{n} , dramatically increasing noise propagation through the reconstruction process. Figure 2.10 demonstrates this for Zernike polynomials measured by hexagonal microlenses on an annular support. Some combination of high order spherical aberrations does closely resemble the focus slope vector. Obviously, this situation is not desired, because the SHS loses its ability to effectively measure focus - one of the strongest contributors to atmospheric turbulence.

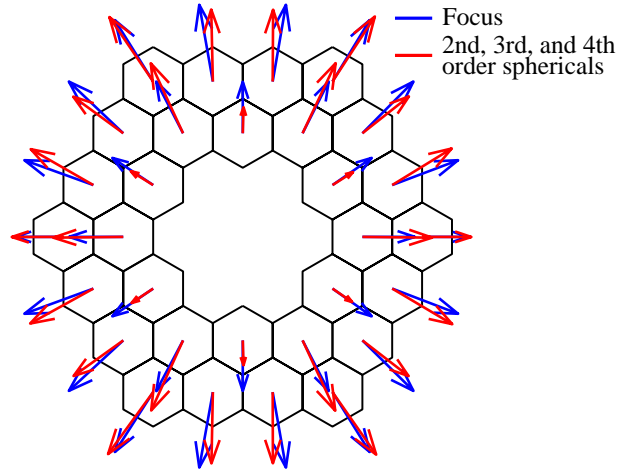


Figure 2.10: When measured with this hexagonal microlens geometry, the focus slope vector (*blue*) is similar to those produced by a combination of 2nd, 3rd, and 4th order spherical aberrations. This leads to large cross-talk between the modes by the addition of small noise vectors.

The mathematical framework in which the liability of \mathbf{D} to cross-talk is best studied is the *Singular Value Decomposition* (SVD), whose fundamental properties are summarized in Appendix C. The SVD of \mathbf{D} is given by

$$\mathbf{D} = \mathbf{U}\mathbf{S}\mathbf{V}^t, \quad (2.43)$$

where \mathbf{U} , \mathbf{V} are orthogonal matrices, and \mathbf{S} is a diagonal matrix whose elements S_{ii} are called *singular values* of \mathbf{D} . The pseudo-inverse (see Appendix C) of \mathbf{D} is then given by

$$\mathbf{D}^+ = \mathbf{V}\mathbf{S}^+\mathbf{U}^t, \quad (2.44)$$

where \mathbf{S}^+ is a diagonal matrix with the reciprocals of the singular values S_{ii}^{-1} . \mathbf{D}^+ is a possible reconstruction matrix (cf. Equation 2.42) and corresponds to the *least-squares* (LS) estimator. If some S_{ii} are small compared to others, their reciprocals will be larger. Thus, small variations of \mathbf{g} originated by noise can produce large variations in the reconstructed modal coefficients $\hat{\mathbf{a}}$.

The *condition number* of a matrix (see Appendix C) is defined as the ratio of its largest to its smallest singular value. It is a measure for the maximum amount of noise amplification through the inversion process in Equation 2.42. Figure 2.11 plots the condition number of the interaction matrix as a function of the number of low order modes included in that matrix. The interaction matrix was calculated for Zernike polynomials and the microlens geometry displayed at the lower right. Note the big step in noise amplification between the modes 35 and 36. The reason for this step is exactly the

situation displayed in Figure 2.10. The 36th mode is the 4th order spherical which, together with 3rd and 2nd order spherical aberrations, can now be used to fit focus under the presence of noise.

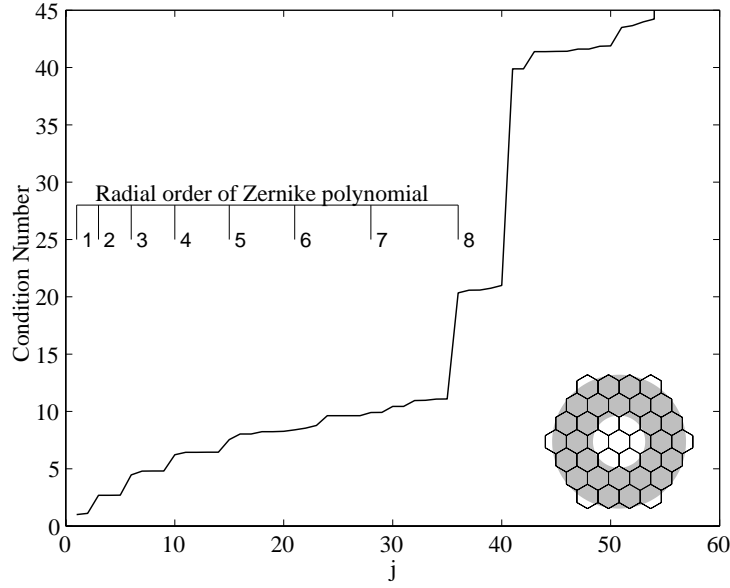


Figure 2.11: Condition number of the interaction matrix as a function of the number of low order modes included in that matrix. The interaction matrix was calculated for Zernike polynomials and the microlens geometry displayed at the lower right.

A set of control modes should always be selected to keep the condition number low, while on the other hand the strongest modes in the decomposition of turbulence should be used. Therefore, low order Zernike or K-L modes truncated at a number that ensures a well-conditioned \mathbf{D} are an effective basis set for AO. A detailed analysis of the interaction of different sets of modes with different microlens geometries is given in Section 3.4. One important result is that cross-talk limits the number of usable modes to about two thirds the number of slopes.

2.3.3 Sensor mode extension

The previous section explained how modal cross-talk limits the number of usable modes of any natural basis set like Zernike polynomials or K-L functions. In principle, however, the number of modes that can be reconstructed equals the number of slopes provided by an SHS. As long as the DM possesses remaining degrees of freedom, it should be possible to find mirror modes which can be measured in addition to a truncated set of natural modes. These additional modes, hereafter *sensor modes*, depend on the microlens geometry and the underlying set of natural modes.

The first step is to find an orthogonal basis set for the SHS measurement space complementary to the one used by the truncated number of natural modes and given by their slope vectors in \mathbf{D} . Instead of using a Gram-Schmidt orthonormalization algorithm on the natural modes' slope vectors, a much quicker and numerically more robust way is to calculate the SVD of \mathbf{D} given by

$$\mathbf{D} = [\mathbf{U}_{\parallel} \ \mathbf{U}_{\perp}] \begin{bmatrix} \mathbf{S} \\ \mathbf{0} \end{bmatrix} \mathbf{V}^t, \quad (2.45)$$

with m, n denoting the number of natural modes and SHS measurements, \mathbf{D} is rectangular ($n \times m$), \mathbf{U}_{\parallel} is ($n \times m$), and \mathbf{U}_{\perp} is ($n \times n - m$). The matrices \mathbf{S} and \mathbf{V} are of minor interest here. One property of

the SVD is that the columns of \mathbf{U}_{\parallel} are an basis set for the SHS measurement space used by the natural modes, and the columns of \mathbf{U}_{\perp} are a basis set for the desired complementary space. Strictly speaking, \mathbf{U}_{\perp} is called the left nullspace of \mathbf{D} .

The next step is to find mirror modes that produce large slope vectors in \mathbf{U}_{\perp} . These slope vectors will be far from being colinear with the columns of \mathbf{D} and thus only marginally affect the condition of \mathbf{D} . Let \mathbf{G} be the interaction matrix of the mirror modes containing all possible slope vectors that can be created by the DM. The projection of \mathbf{G} onto \mathbf{U}_{\perp} tells us how big these slope vectors are in \mathbf{U}_{\perp} . With \mathbf{M} being an orthogonal matrix, Golub and Van Loan (1996) show that the matrix $\mathbf{M}\mathbf{M}^t$ describes the unique orthogonal projection onto the subspace spanned by the columns of \mathbf{M} . Thus, the projection of the slope vectors of the mirror modes onto \mathbf{U}_{\perp} is given by

$$\mathbf{G}_{\perp} = (\mathbf{U}_{\perp}\mathbf{U}_{\perp}^t)\mathbf{G}. \quad (2.46)$$

The linear combinations of the mirror modes that produce the largest slope vectors in \mathbf{U}_{\perp} can now be found by an SVD of \mathbf{G}_{\perp} :

$$\mathbf{G}_{\perp} = \mathbf{U}_r[\mathbf{S}_r \mathbf{0}] \begin{bmatrix} \mathbf{V}_{r\perp}^t \\ \mathbf{V}_{r\parallel}^t \end{bmatrix}. \quad (2.47)$$

The columns of $\mathbf{V}_{r\perp}$ contain the desired sensor modes as linear combinations of the mirror modes. Further, the SVD algorithm orders them such that the sensor modes with the largest SHS signals in \mathbf{U}_{\perp} are placed first.

Remarks. The sensor modes can be used in addition to a set of low order natural modes without considerably increasing modal cross-talk. They are orthogonal to each other, and they are also orthogonal to the natural modes, if those are modelled as a superposition of mirror modes. The effective usage of sensor modes, however, is constrained to bright guide stars, because sensor modes are of high spatial order and therefore of low SNR. The sensor mode extension was tested at the telescope as discussed in Section 3.4.

2.3.4 Modal reconstruction error

A turbulent wavefront is a spatially distributed signal and therefore of infinite dimension, i.e. it is a superposition of an infinite number of modes. The SHS signals of high spatial order modes contaminate those of low order modes which are to be estimated. For convenience, the SHS measurement model from Equation 2.41 is repeated here:

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

The term $\mathbf{D}_{\perp}\mathbf{a}_{\perp}$ is the net contribution from high order modes and can be considered as an additional error source which is generally called *aliasing noise* in analogy to spectrum estimation of time signals. Aliasing noise is a fundamental error that is always present, even for very bright guide stars.

Now, the modal reconstruction error in the presence of aliasing- and measurement noise will be calculated. With the reconstruction matrix \mathbf{R} , the estimation of modal coefficients can formally be written as

$$\begin{aligned} \hat{\mathbf{a}} &= \mathbf{R}\mathbf{g} \\ &= \mathbf{R}\mathbf{D}\mathbf{a} + \mathbf{R}\mathbf{D}_{\perp}\mathbf{a}_{\perp} + \mathbf{R}\mathbf{n}. \end{aligned} \quad (2.48)$$

Using the LS estimate $\mathbf{R} \equiv \mathbf{D}^+$ which is the pseudo-inverse of \mathbf{D} ($\mathbf{R}\mathbf{D} = \mathbf{Id}$, see Appendix C), and assuming uncorrelated measurement noise, the covariance matrix of the reconstruction error can be written as

$$\langle (\mathbf{a} - \hat{\mathbf{a}})(\mathbf{a} - \hat{\mathbf{a}})^t \rangle = \mathbf{C} \langle \mathbf{a}_\perp \mathbf{a}_\perp^t \rangle \mathbf{C}^t + \mathbf{D}^+ \langle \mathbf{nn}^t \rangle \mathbf{D}^{+t}, \quad (2.49)$$

with the two terms describing propagated aliasing and measurement noise respectively, and the *cross-talk matrix* (Dai, 1996) $\mathbf{C} = \mathbf{D}^+ \mathbf{D}_\perp$.

The reconstruction error of an individual mode is

$$\begin{aligned} \sigma_{r,j}^2 &= \langle |a_j - \hat{a}_j|^2 \rangle \\ &= (\mathbf{C} \langle \mathbf{a}_\perp \mathbf{a}_\perp^t \rangle \mathbf{C}^t + \mathbf{D}^+ \langle \mathbf{nn}^t \rangle \mathbf{D}^{+t})_{jj}. \end{aligned} \quad (2.50)$$

Using Equation 2.13 for orthonormal basis sets, the total phase reconstruction error variance can be written as

$$\begin{aligned} \sigma_t^2 &= \sum_{j=1}^{\infty} \langle |a_j - \hat{a}_j|^2 \rangle \\ &= \sum_{j=1}^m \sigma_{r,j}^2 + \sum_{j=m+1}^{\infty} \langle a_j^2 \rangle, \end{aligned} \quad (2.51)$$

where m denotes the number of reconstructed modes. The second term in Equation 2.51 is the *fitting error* σ_{fit}^2 being the cumulative variance of the unreconstructed modes (cf. Figure 2.2).

When measurement noise is negligible, Equation 2.50 reduces to the propagated aliasing error:

$$\sigma_{al,j}^2 = (\mathbf{C} \langle \mathbf{a}_\perp \mathbf{a}_\perp^t \rangle \mathbf{C}^t)_{jj}. \quad (2.52)$$

The total error of the phase reconstruction due to aliasing is then

$$\sigma_{al}^2 = \text{Tr}(\mathbf{C} \langle \mathbf{a}_\perp \mathbf{a}_\perp^t \rangle \mathbf{C}^t). \quad (2.53)$$

Dai (1996) called this the *remaining error*, because it always remains even for the brightest guide stars. Dai was also the first who realized the importance of the remaining error for the interpretation of wavefront reconstruction simulations, while earlier works misinterpreted reconstruction errors in the absence of measurement noise as numerical round-off errors.

Figure 2.12 plots $\sigma_{al,j}^2$ for the reconstruction of 30 Zernike polynomials with the microlens geometry displayed to the upper right. Aliasing noise propagates mainly onto the spherical aberrations $j=10$ and $j=21$, preventing them from being reconstructed accurately. If $\sigma_{al,j}^2$ of a mode is similar to its atmospheric signal $\langle a_j^2 \rangle$, its compensation in the feedback loop cannot reduce the overall wavefront error. It should be noted that the aliasing error is amplified to some extent by using the LS reconstruction; for *maximum a posteriori* (MAP) reconstruction, the error never exceeds the atmospheric signal.

Figure 2.13 shows the ratio of σ_{al}^2 to σ_{fit}^2 as a function of the number of reconstructed modes. Again, Zernike polynomials are reconstructed with the microlens geometry displayed to the lower right. The ratio generally increases with the number of compensated modes; the fitting error decreases, and aliasing increases, because the high order energy can propagate onto a larger measurement space. However, the curve shows an inflection point when the coma aberrations are included in the reconstruction. The reason for this behavior is that coma also shifts the centroid of an image and therefore strongly affects the tip-tilt measurement. The misinterpretation stops and aliasing drops, once the system knows the coma slope vectors ($j=6$ and $j=7$). In general, aliasing and fitting errors are comparable in magnitude.

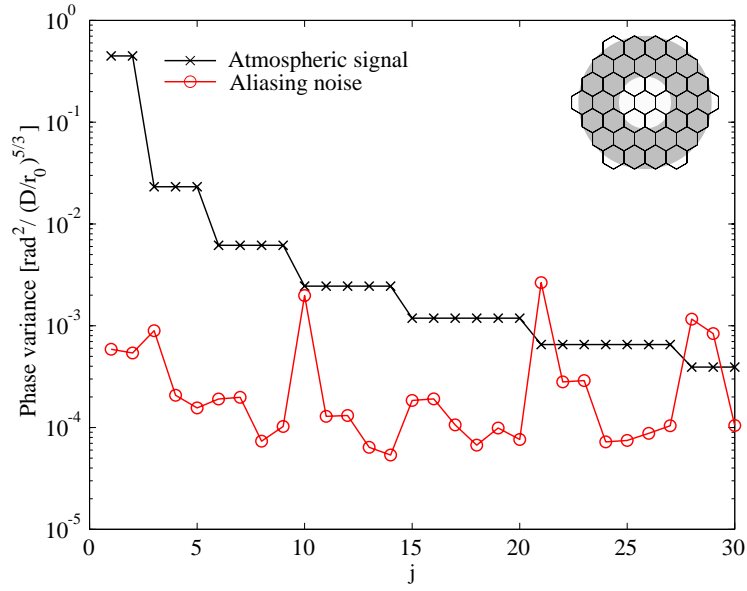


Figure 2.12: Aliasing error variance for LS reconstruction of Zernike polynomials with given microlens geometry. Aliasing noise propagates especially onto the spherical aberrations ($j=10$, $j=21$).

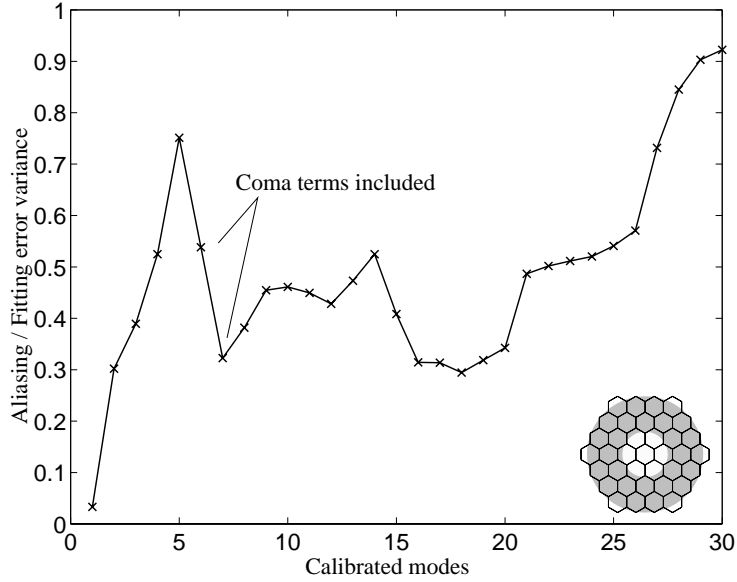


Figure 2.13: Ratio of aliasing to fitting error for LS reconstruction as a function of the number of modes which are to be estimated.

2.3.5 Modal covariance estimation

The covariance matrix of a wavefront's modal decomposition is a complete description of its stochastic properties, because the piston removed wavefront phase is Gaussian and zero mean. It carries all the information present in the turbulence spectrum (cf. Equation 2.3). The knowledge of the covariance matrix is very important, because it can improve phase estimation algorithms. Further, the open-loop covariance matrix provides an accurate estimate of r_0 . Note that in a feedback system, we have to deal with the closed-loop covariance matrix, i.e. the residual aberrations after compensation by the DM.

The covariance matrix of the modal coefficients $\langle \mathbf{a}\mathbf{a}^t \rangle$ can be estimated from slope vectors by using the LS estimate in Equation 2.48 to achieve

$$\hat{\mathbf{a}} = \mathbf{a} + \mathbf{C}\mathbf{a}_\perp + \mathbf{D}^+ \mathbf{n}. \quad (2.54)$$

Since the measurement error \mathbf{n} was assumed to be uncorrelated and will not propagate through the loop without a delay, there is no correlation between \mathbf{n} and \mathbf{a} , or \mathbf{a}_\perp . Thus, multiplying \mathbf{a}_\perp with its transpose and taking the ensemble average yields the modal covariance matrix

$$\mathbf{V}_\mathbf{a} := \langle \mathbf{a}\mathbf{a}^t \rangle = \langle \hat{\mathbf{a}}\hat{\mathbf{a}}^t \rangle - \mathbf{D}^+ \langle \mathbf{n}\mathbf{n}^t \rangle \mathbf{D}^{+t} - \mathbf{C} \langle \mathbf{a}_\perp \mathbf{a}_\perp^t \rangle \mathbf{C}^t - \langle \mathbf{a}\mathbf{a}_\perp^t \rangle \mathbf{C}^t - \mathbf{C} \langle \mathbf{a}_\perp \mathbf{a}^t \rangle. \quad (2.55)$$

Obviously, the two last terms in Equation 2.55 will be very small for Zernike polynomials and zero for K-L functions, because the modes are independent (cf. Section 2.1), i.e.

$$\langle \mathbf{a}\mathbf{a}^t \rangle_{op} \approx \langle \hat{\mathbf{a}}\hat{\mathbf{a}}^t \rangle - \mathbf{D}^+ \langle \mathbf{n}\mathbf{n}^t \rangle \mathbf{D}^{+t} - \mathbf{C} \langle \mathbf{a}_\perp \mathbf{a}_\perp^t \rangle \mathbf{C}^t. \quad (2.56)$$

In closed-loop, Veran et al. (1997) showed that the modal covariance is approximately

$$\langle \mathbf{a}\mathbf{a}^t \rangle_{cl} \approx \langle \hat{\mathbf{a}}\hat{\mathbf{a}}^t \rangle - \mathbf{D}^+ \langle \mathbf{n}\mathbf{n}^t \rangle \mathbf{D}^{+t} + \mathbf{C} \langle \mathbf{a}_\perp \mathbf{a}_\perp^t \rangle \mathbf{C}^t. \quad (2.57)$$

To understand why $\mathbf{C} \langle \mathbf{a}_\perp \mathbf{a}_\perp^t \rangle \mathbf{C}^t$ changed its sign, the response of the feedback system to aliasing noise must be considered. In open-loop, energy at high spatial frequencies is misinterpreted as energy at low frequencies. In closed-loop, aliasing actually leads to more energy at low order modes as those are introduced by the DM trying to compensate for aliasing noise. To first order, the aliasing noise vectors disappear by this compensation. The conclusion is that open-loop aliasing leads to an overestimation of the covariance of low order modes, whereas closed-loop aliasing leads to an underestimation.

Required Quantities. $\hat{\mathbf{a}}$ is derived directly from the slope vectors, and \mathbf{n} is measurable as described in Section 3.2. The interaction matrix of high order modes \mathbf{D}_\perp is required to calculate \mathbf{C} but unfortunately hard to measure, because the DM cannot reproduce them (possibly holographic plates as described by Otsubo et al. (1997) could serve for this purpose). For this work, \mathbf{D}_\perp was determined by a computer model of the ALFA SHS described in Section 2.3.1. The covariance matrix of the high order modes $\langle \mathbf{a}_\perp \mathbf{a}_\perp^t \rangle$ cannot be measured at all, so it is generally assumed to follow Kolmogorov statistics. This requires the measurement of r_0 , because modal covariance scales with $(D/r_0)^{5/3}$.

Measurement of r_0 . Using the Kolmogorov model, Noll (1976) calculated the variances of Zernike modes displayed in Table 2.3, all of them scaling with $(D/r_0)^{5/3}$. Thus, each measured modal variance provides an independent estimate of r_0 , and their average greatly improves the accuracy.

Now, we face the problem that we need to know the variances in order to estimate r_0 , whereas the knowledge of r_0 is required to estimate the variances using Equation 2.56. However, the following recursive scheme was found to converge within a few iterations:

1. Assume $r_0 = \infty$.
2. Calculate $\langle \mathbf{a}\mathbf{a}^t \rangle$ via Equation 2.56.
3. Estimate r_0 from the quantities.
4. Start over with step 2 until r_0 converges.

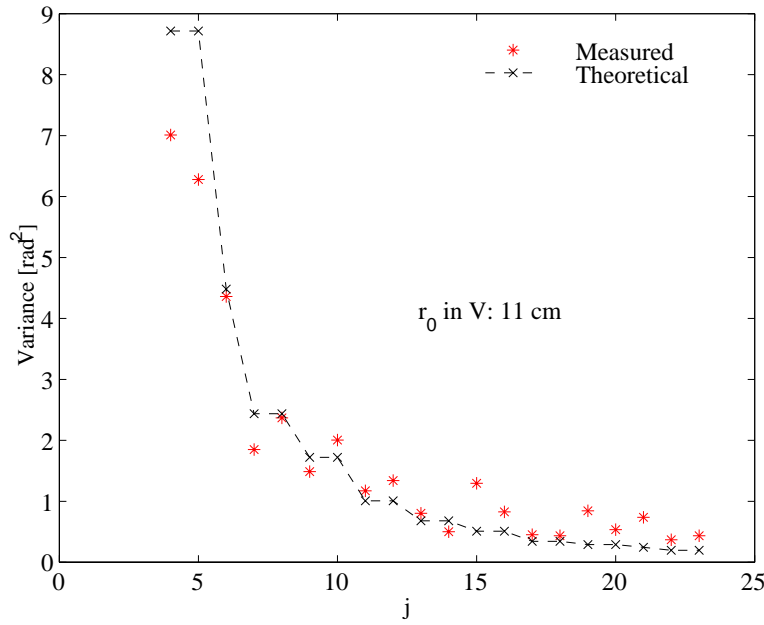


Figure 2.14: Measured variances of low order K-L modes (tip-tilt is excluded), and the theoretical values multiplied by $(D/r_0)^{5/3}$. The best fit is achieved for $r_0 = 11$ cm in the V-band, corresponding to about $1''$ seeing.

Figure 2.14 shows a representative measurement of r_0 and the modal variances. The variance of the K-L modal coefficients is plotted against the mode number. The Kolmogorov model usually fits the measurements quite well. The attenuation of very low order modes due to the finite outer scale of turbulence is always observed. Hence, the most affected modes have not been used in the estimation process of r_0 .

The estimation of r_0 from closed-loop data is more difficult, because the modes are corrected and do not appear in SHS data. Veran et al. (1997), however, showed that it can be determined from mirror signals as the DM roughly follows the motion of low order atmospheric modes. Another possibility would be the simultaneous observation of an off-axis star which is picked up before the wavefront is reflected at the DM.

2.3.6 Phase estimators

Only LS estimation of modal coefficients was considered so far, because this is the standard algorithm used by current AO systems. However, it is well known in estimation theory that prior knowledge of noise and estimation parameter statistics can significantly improve an estimate. The Kolmogorov turbulence and noise models provide this additional information. Wallner (1983) found the optimal wavefront estimator for an SHS based AO system, but used a strictly zonal formalism making his method is hard to adapt to modal control. However, for the linear measurement model:

$$\mathbf{g} = \mathbf{D}\mathbf{a} + \mathbf{e}, \quad (2.58)$$

the optimal estimator for parameters \mathbf{a} with Gaussian statistics under the presence of Gaussian noise \mathbf{e} is well known. By defining an effective measurement noise \mathbf{e} , it is now possible to match our

measurement model in Equation 2.41 with Equation 2.58, i.e.

$$\mathbf{g} = \mathbf{D}\mathbf{a} + \underbrace{\mathbf{D}_\perp \mathbf{a}_\perp}_{\mathbf{e}} + \mathbf{n}. \quad (2.59)$$

Since \mathbf{n} is Gaussian, and the modal coefficients \mathbf{a} are in fact normally distributed with zero mean (Rodier, 1981), \mathbf{e} is also Gaussian. For this problem, the optimal *maximum a posteriori* (MAP) estimate derived in Appendix D is (Melsa and Cohn, 1978)

$$\hat{\mathbf{a}} = (\mathbf{D}'\mathbf{V}_e^{-1}\mathbf{D} + \mathbf{V}_a^{-1})^{-1}\mathbf{D}'\mathbf{V}_e^{-1} \cdot \mathbf{g}. \quad (2.60)$$

with $\mathbf{V}_e := \langle \mathbf{e}\mathbf{e}' \rangle$, and $\mathbf{V}_a := \langle \mathbf{a}\mathbf{a}' \rangle$. The MAP estimate $\hat{\mathbf{a}}$ is the most likely to produce the measurement \mathbf{g} when wavefront and noise statistics are considered. Law and Lane (1996) have demonstrated that Equation 2.60 is equivalent to Wallner's result.

Further, in Appendix D it is shown that the covariance of the estimation error is:

$$\langle (\mathbf{a} - \hat{\mathbf{a}})(\mathbf{a} - \hat{\mathbf{a}})' \rangle = (\mathbf{D}'\mathbf{V}_e^{-1}\mathbf{D} + \mathbf{V}_a^{-1})^{-1}. \quad (2.61)$$

It should be noted that this result is equivalent to Equation 2.49, if the prior knowledge on \mathbf{V}_a is neglected (the term \mathbf{V}_a^{-1} is zero in this case, since the modal coefficients can take arbitrary values and thus are of infinite variance).

Properties of the MAP estimate. Other estimators follow directly from the MAP estimate under special assumptions. Equation 2.60 resembles the *weighted least squares* (WLS) estimate when prior information on the modal coefficients is neglected ($\mathbf{V}_a^{-1} = 0$), i.e.

$$\hat{\mathbf{a}} = (\mathbf{D}'\mathbf{V}_e^{-1}\mathbf{D})^{-1}\mathbf{D}'\mathbf{V}_e^{-1} \cdot \mathbf{g}, \quad (2.62)$$

Equation 2.60 further reduces to the LS estimate, when the measurement error is assumed to be uniform and uncorrelated ($\mathbf{V}_e = \sigma_s^2 \cdot \mathbf{Id}$). Then, the corresponding terms in Equation 2.62 annihilate each other such that

$$\hat{\mathbf{a}} = (\mathbf{D}'\mathbf{D})^{-1}\mathbf{D}' \cdot \mathbf{g}. \quad (2.63)$$

The LS estimator is most commonly used in AO, because it is easy to calculate and does not require any prior assumptions and modelling.

Equation 2.61 shows that as the noise increases ($\mathbf{V}_e \rightarrow \infty$), an upper bound for the MAP estimation error covariance is set by \mathbf{V}_a . Inspecting Equation 2.60, we see that for a very large \mathbf{V}_e the MAP estimate is essentially zero. In contrast, the LS and WLS estimation error (Equation 2.61 with $\mathbf{V}_a^{-1} = 0$) grows without bounds, because these estimators will fit the noise regardless of the likelihood of the fitted modes to arise in atmospheric turbulence. Even with negligible measurement noise, the remaining aliasing noise assures that the MAP performs better than WLS, and LS.

Finally, the MAP estimator allows us to estimate any number of control modes, since the well-conditioned \mathbf{V}_a usually assures that the term $(\mathbf{D}'\mathbf{V}_e^{-1}\mathbf{D} + \mathbf{V}_a^{-1})$ in Equation 2.60 can be inverted. This is especially interesting for applications like *deconvolution from wavefront sensing* (Primot et al., 1990) where the whole wavefront has to be estimated. Whereas in AO, we are already satisfied with the best estimate of the finite number of modes that can be controlled by the DM.

Determination of \mathbf{V}_e and \mathbf{V}_a . The main drawback of the MAP estimation method is that the needed covariance matrices vary with observing conditions. \mathbf{V}_a will vary with seeing and/or correction quality, while \mathbf{V}_e is most sensitive to variations of the atmospheric transparency but also depends on the seeing. Therefore, these quantities have to be measured frequently in order to update the estimator in case of unstable observing conditions.

\mathbf{V}_e follows from Equation 2.59 and the usual assumption of \mathbf{n} being independent of \mathbf{a}_\perp :

$$\mathbf{V}_e = \mathbf{D}_\perp \langle \mathbf{a}_\perp \mathbf{a}_\perp^t \rangle \mathbf{D}_\perp^t + \langle \mathbf{n} \mathbf{n}^t \rangle. \quad (2.64)$$

As already explained in Section 2.3.5, the first term resembling aliasing noise can be obtained from simulations and the measurement of r_0 . The second term resembling measurement noise can be determined by a method described in Section 3.2.

\mathbf{V}_a can be derived in open- or closed-loop respectively by Equations 2.56 and 2.57 as described in the previous section.

Performance of the MAP estimate. The simulated results in Figures 2.15 and 2.16 demonstrate the superior quality of the MAP estimates of K-L modal coefficients for negligible measurement noise.

The simulations were performed using the ALFA-SHS model described in Section 3.3. Hereby, the measurement process is modelled by calculating slope vectors for arbitrary phase-screens and microlens geometries. Phase-screens resembling K-L modes were used to determine the interaction matrix \mathbf{D}_∞ . Since arbitrarily large mode numbers can be used in the computer, the model provides the matrix \mathbf{D}_\perp required for the calculation of the aliasing error.

Kolmogorov phase-screens were created by superposing K-L functions (Cannon, 1996), where the individual modes were weighted by Gaussian random numbers of their respective atmospheric signal variances (cf. Table 2.6). Hence, the K-L decomposition of the phase-screens was exactly known. The ALFA-SHS model could then be used to calculate slope vectors resulting from these phase-screens. Subsequently, Gaussian random vectors were added to simulate measurement noise.

Using the known modal coefficient- and error covariance matrices, MAP and WLS estimators were calculated and used to reconstruct the modal coefficients from the noisy measurements. Finally, the reconstructed coefficients were compared with the exact ones that had been used to simulate the phase-screens. A large number of 1000 phase-screens was used to ensure the significance the results.

Figure 2.15 shows that the MAP estimate is most accurate for all individual modes. Further, it provides effective estimates for more than 100 modes although only 30 subapertures were used. WLS, and LS estimation were truncated to 30 modes to avoid strong noise amplification (see Section 2.3.2). For the first 30 modes, MAP and WLS estimators produce similar results, because the prior knowledge of \mathbf{V}_a is negligible provided \mathbf{V}_e is small. The aliasing noise in \mathbf{V}_e , however, is still large enough to deteriorate the LS estimates.

Figure 2.16 shows the total reconstruction error variance of the first 30 K-L modes as a function of SNR (defined as the ratio of the mean lengths of wavefront slope and noise vectors). As expected the MAP is superior to LS, and WLS estimates at all light levels. A careful inspection of the graph reveals that the MAP needs only about 85 % the SNR of the LS in order to reach a given σ_t^2 . Using Equation 2.38, this means that the MAP needs only 0.85² i.e. 72 % of the flux translating into a gain of about 0.35 magnitudes in guide star brightness requirements when compared to the LS.

Robustness of the MAP. An interesting question is how the uncertainty of the determination of \mathbf{V}_e , and \mathbf{V}_a affects the quality of the MAP estimate. Therefore, the simulations were repeated with errors in the calculation of the MAP estimator resulting from uncertain values of r_0 , $\langle \mathbf{a}_\perp \mathbf{a}_\perp^t \rangle$, and $\langle \mathbf{n} \mathbf{n}^t \rangle$.

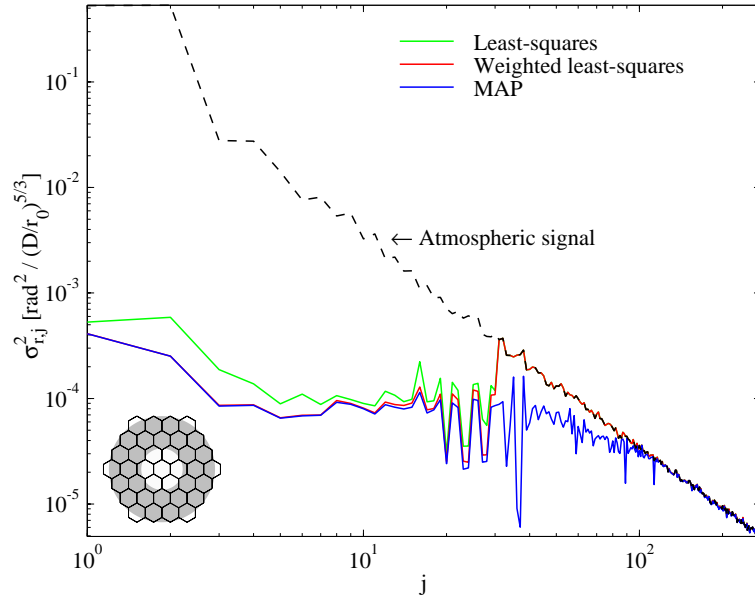


Figure 2.15: Reconstruction error $\sigma_{r,j}^2 = \langle (a_j - \hat{a}_j)^2 \rangle$ of K-L modal coefficients as a function of the mode number j for negligible measurement noise. The simulations used the microlens geometry displayed to the lower left.

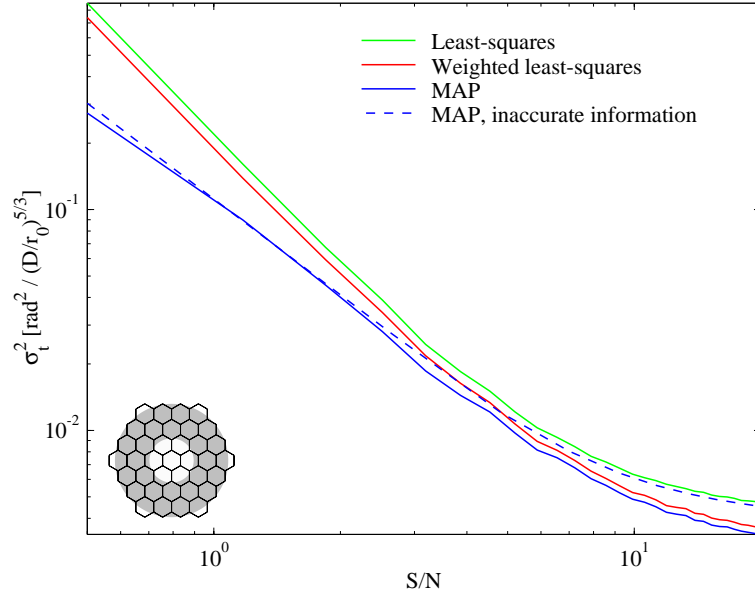


Figure 2.16: Total reconstruction error variance $\sigma_t^2 = \sum_{j=1}^{30} \langle (a_j - \hat{a}_j)^2 \rangle$ of 30 K-L modes with different estimation algorithms as a function of SNR. The dashed line represents a MAP estimate with uncertainties of 30 % for r_0 , 300 % for $\langle \mathbf{a}_\perp \mathbf{a}_\perp^t \rangle$, and 20 % for $\langle \mathbf{n} \mathbf{n}^t \rangle$.

The results show that the MAP estimate is remarkably robust; it performed better than the LS estimate at all light levels even when all quantities were simultaneously uncertain at a level of 75%. This value is overly pessimistic, since r_0 and $\langle \mathbf{n} \mathbf{n}^t \rangle$ can be measured with a much better accuracy. On the other hand, $\langle \mathbf{a}_\perp \mathbf{a}_\perp^t \rangle$ can only be derived from a turbulence model. The Kolmogorov model is usually a very good description of turbulence in the free atmosphere, but this does not include dome-

seeing or outer scale effects. Therefore, the assumed $\langle \mathbf{a}_\perp \mathbf{a}_\perp^t \rangle$ potentially bears the largest uncertainty.

A rather pessimistic test is represented by the dashed line in Figure 2.16. In this case, the assumed errors were 30 % for r_0 , 300 % for $\langle \mathbf{a}_\perp \mathbf{a}_\perp^t \rangle$, and 20 % for $\langle \mathbf{nn}^t \rangle$. Even then the MAP performs better than the LS estimator.

2.4 Laser guide star peculiarities

The sky-coverage of conventional AO systems using an NGS is limited to about 1 % in the K-band, while using an LGS increases this value to about 50 % (Le Louarn et al., 1998). This is derived from the number of NGS of the required brightness multiplied by the isoplanatic patch size which is the correctable area around each NGS. With an LGS, the limits are set by the number of NGS suitable for tip and tilt correction (see Section 2.4.1).

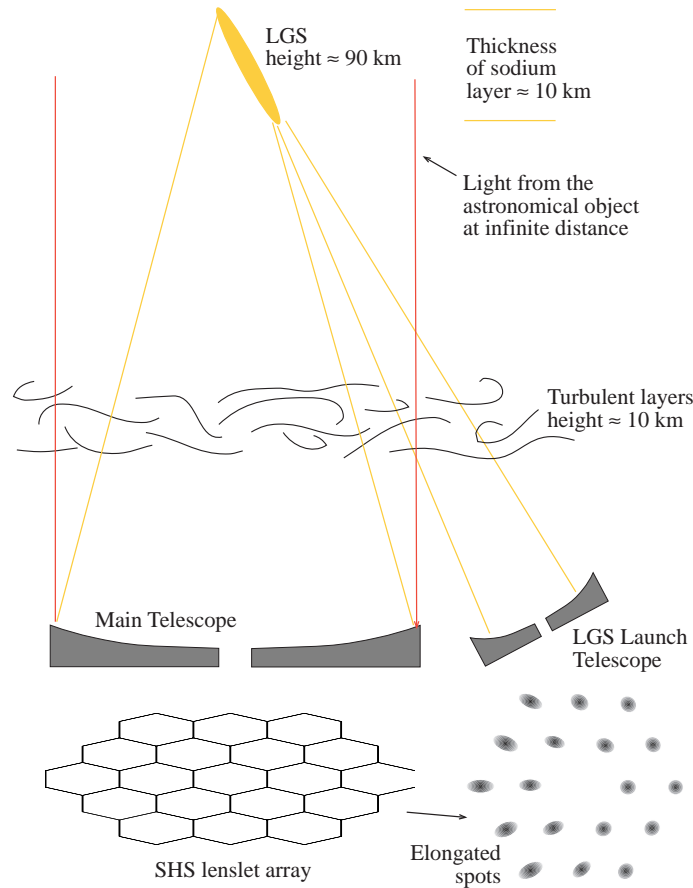


Figure 2.17: Outline of an LGS projected through a launch telescope next to the main telescope. The narrow upward beam of the LGS passes through other turbulence than the downward beam. The LGS cone does not sample exactly the same turbulence as the cylindric beam from an astronomical object at infinite distance (*cone effect*). The LGS spot shape depends on the perspective of the SHS microlenses.

Figure 2.17 outlines a laser beam being projected through a launch telescope next to the main telescope. The narrow upward beam of the LGS passes through other turbulence cells than the downward beam. Since the width of the upward beam is usually chosen close to r_0 , Equation 2.32 can be used to show that the LGS jitters on the sky over an area about half as big as the seeing λ/r_0 . The jitter

resulting from the downward motion is somewhat smaller due to the larger beam diameter.

Additionally can be seen that the LGS cone does not sample exactly the same turbulence as the cylindric beam from an astronomical object at infinite distance. This so called *cone effect* leads to an phase estimation error of (Fried and Belsher, 1994)

$$\sigma_{cone}^2 = (D/d_0)^{5/3}, \quad (2.65)$$

where the quantity d_0 may be considered as the diameter of the aperture over which the wavefront error due to the cone effect is 1 rad^2 . Like r_0 , d_0 depends on observing conditions; it is of the order of 5 to 10 m in the NIR. The cone effect can only be overcome by using multiple LGS which must be arranged such that all turbulent layers are properly sampled. It should be mentioned that a second effect of using multiple guide stars is that the corrected field of view (FOV) can be enlarged by resolving the turbulent layers and using multiple DM conjugate to these layers; this method is called *multi-conjugate adaptive optics* (MCAO). The disentanglement of the contributions of distinct turbulent layers from wavefront sensor signals is a currently investigated topic (Berkefeld et al., 2000, Tallon and Foy, 1990, Ragazzoni et al., 1999).

Finally, the figure demonstrates how the *spot shape* of the LGS depends on the perspective at which it is observed by the SHS microlenses. In addition to its inherent size θ_{in} of at least $\sqrt{2}$ times the seeing (the upward beam also has to travel through the turbulent layers, only the center of the sodium layer is in focus), the LGS is elongated in one axis due to the thickness of the sodium layer. Simple geometry shows that the elongation θ_{el} depends on the radial distance x between subaperture and launch telescope as

$$\theta_{el} = \frac{x\Delta h}{h^2 + h\Delta h} \approx \frac{x\Delta h}{h^2}, \quad (2.66)$$

where the bottom edge of the sodium layer is at an altitude h , and its thickness is Δh . For a distance of $x = 4 \text{ m}$ this corresponds to an elongation of approximately $\theta_{el} \approx 1''$. For the short exposures required in AO, the motion of the LGS does not affect its size thus given by θ_{el} convolved with θ_{in} . We have never observed LGS spots which are smaller than about two times the seeing. Therefore, according to Equation 2.38, the LGS has to be at least 4 times brighter than an NGS in order to achieve the same wavefront sensing accuracy.

2.4.1 Tip-tilt sensing

The overall tip-tilt of a wavefront is hardly measurable with an LGS, as a consequence of the round trip of the laser beam through the atmosphere leaving its absolute position on the sky uncertain. There exist some ideas how this drawback might be overcome by use of auxiliary telescopes which can be directed to measure the upward jitter (Ragazzoni and Esposito, 1999), or by use of a laser that fluoresces at two wavelengths so that tip-tilt can be measured from differential dispersion in the atmosphere (Foy et al., 1995). Neither of these approaches is likely to lead to a breakthrough in the near future, because the laser power required for a sufficient SNR is still too demanding.

Rigaut and Gendron (1992) extensively investigated the tip-tilt determination problem for AO with an LGS. They emphasize that the slow decorrelation time and the fact that only a single aperture is required for tip-tilt sensing is misleading. Even at high frequencies the contribution of tip and tilt is as big as that of other modes (cf. Figure 2.1). Further, Equation 2.38 showed that a single aperture does not provide a more accurate tip-tilt estimate than a number of small subapertures. Therefore, the necessity of an NGS for tip-tilt sensing imposes a stronger limitation on sky-coverage of AO systems using an LGS than one might think.

However, when compared to the brightness of NGS required for full correction, the tip-tilt NGS can be fainter, because measurement noise only propagates onto the tip and tilt modes. It can also be slightly further away due to the slow decorrelation of tip and tilt with angular distance. Simulations show that this gain should be high enough to increase the sky-coverage to about 50 % in the K-band (Le Louarn et al., 1998) for single LGS at a 3.6-m-telescope, and about 75 % at an 8-m-telescope.

2.4.2 Focus sensing

Unlike tip-tilt sensing, focus sensing with an LGS should be feasible but nevertheless requires attention. First of all, the distance to the sodium layer changes with airmass. Depending on the viewing angle γ , the distance to the LGS is $h(\gamma) = h_0 \cdot \cos \gamma^{-1}$, where h_0 is the altitude of the center of the sodium layer. Assuming that h_0 is known (it can be measured with a light detection and ranging (LIDAR) device, see Butler et al. (2000)), this systematic change can be compensated for by adding offsets to the measured focus coefficient according to the viewing angle in an open-loop.

O’Sullivan et al. (2000) monitored the altitude of the sodium layer at Calar Alto over several nights. The measurements revealed centroid-height variations that occur on rather short timescales of seconds and are dominated by the formation of dense sporadic layers. The most dramatic feature was a change in centroid height of 400 m in only 1-2 minutes. The focus error resulting from such variations depends on the size of the telescope and can be calculated as follows:

The aperture D cuts out an parabolic arc of the spherical light wave from the LGS. The peak to valley (P-V) distance of this arc is

$$p = h_0 - \sqrt{h_0^2 - \left(\frac{D}{2}\right)^2} \approx \frac{D^2}{8h_0}. \quad (2.67)$$

One unit of focus has a P-V of $2\sqrt{3} \text{ rad}^2$ (see Table 2.1). Therefore, the wavefront originating from an LGS centered at the height h_0 is aberrated by $(\pi D^2)/(8\sqrt{3}h_0\lambda)$ units of focus.

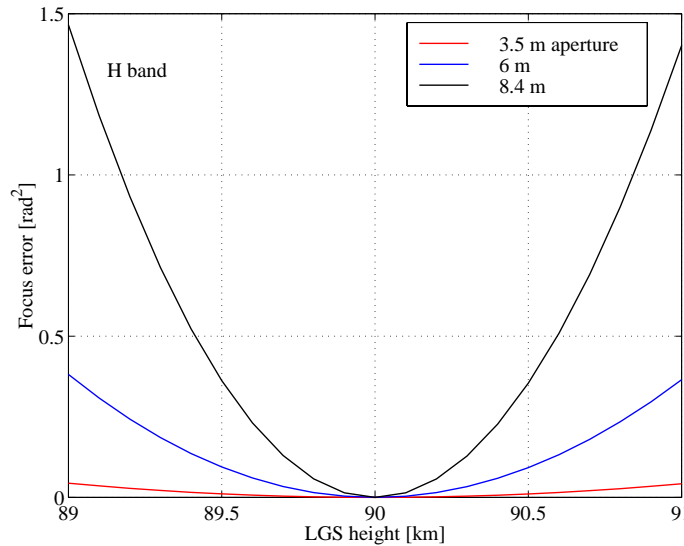


Figure 2.18: Focus error in the H-band resulting from altitude variations of the sodium layer. The effect is negligible for small telescopes, but becomes significant for the 10-m class.

Figure 2.18 shows the focus error in the H-band that results from a deviation of the sodium layer altitude from the nominal value of 90 km. Sporadic variations are negligible for small telescopes, but

may become significant for the 10-m class, especially when observing at shorter wavelengths. In order to compensate for sporadic variations, the height of the LGS should be monitored simultaneously to the observations. Alternatively, a floating average of reconstructed focus coefficients could be calculated and subsequently subtracted from the individual reconstructed coefficients.

2.4.3 LGS modes

As emphasized before, the LGS spots are elongated depending on the perspective of the microlenses. This results in a lower accuracy of the SHS measurement in the respective axis. Another effect of the elongation, which has not been considered yet, is a lateral shift of the centroid resulting in a bias on the measured slopes.

As illustrated by Figure 2.19, the observed centroid of an LGS plume of length $2a$ is shifted by δ when viewed at an angle γ with respect to the real centroid. Geometric consideration lead to the following expressions for the angles α and β :

$$\alpha, \beta = \arccos \sqrt{\frac{(b \pm a \cos \gamma(x))^2}{a^2 + b^2 \pm 2ab \cos \gamma(x)}}, \quad (2.68)$$

where $\gamma(x) \approx x/(h_0 + \Delta h)$ for small distances x . The shift angle is then simply given by $\delta = (\beta - \alpha)/2$. The actual shift values for an assumed uniform LGS plume of 10 km thickness centered at 95 km are of the order of a few hundreds of an arcsecond for current telescope sizes. The shifts grow linearly with the off-axis distance x . Since the shifts produce offsets of the measured slopes, they generate static aberrations. Defocus produces slopes that grow linearly in radial direction with distance to the aperture center. Thus, an LGS projected from behind the secondary telescope mirror will generate a static defocus aberration, while when projected from the side of the main mirror, it will generate a combination of tip-tilt and defocus. The error variance introduced by this effect becomes comparable to 1 rad^2 for 50 m class telescopes, but it is negligible for current aperture sizes.

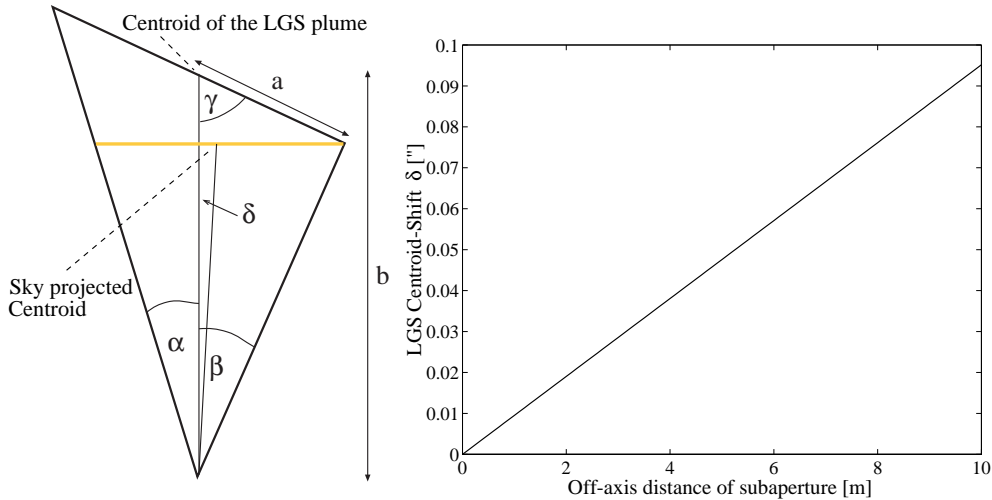


Figure 2.19: *Left:* Centroid shift of the LGS due to its elongation. The centroid of the LGS plume of length $2a$ is shifted by δ when viewed at an angle of γ . *Right:* Centroid shifts calculated for an uniform LGS plume of 10 km thickness centered at 95 km.

Chapter 3

Application to the ALFA AO system

While the last chapter explained the theoretical foundations required to optimize wavefront reconstruction with an SHS, this chapter describes the application of such concepts to the ALFA AO system. The starting point is an introduction to ALFA's technical specifications including its optical design, implemented control components, and modal control architecture. A new type of microlens array for SHS using a keystone design is discussed. In the context of feedback compensation, a simple scheme is sketched that can be used to determine the optimum framerate to operate the system.

The optimization of the SHS measurement process naturally starts with an investigation of centroid estimation. Beside accuracy, computational speed is an issue for appropriate algorithms to provide the high framerates required for astronomical AO. Section 3.2 investigates centroid estimation by use of computer simulations and presents the performance of different algorithms at the telescope.

Section 3.3 describes the computer model used to simulate ALFA's SHS. This model allows us to determine the interaction of high order wavefront modes with the SHS, i.e. the aliasing component, which is crucial for calculating advanced MAP and WLS phase estimators.

Section 3.4 compares the performance of different modal basis sets together with different microlens geometries with respect to their cross-talk and aliasing characteristics. Depending on the actual microlens array, optimum numbers of control modes giving a minimum LS reconstruction error are calculated. Using the MAP estimator, the situation is more simple - the higher the number of reconstructed modes, the lower the reconstruction error.

Eventually, the performance of basis sets and estimators must be confirmed at a telescope. Section 3.5 discusses experiments that were performed using ALFA and aiming at a comparison of K-L, Zernike, and sensor modes as well as LS, WLS, and MAP phase estimators in closed-loop operation.

Based on the simulated and experimental results, Section 3.6 summarizes and draws a guide that will lead to an optimized system operation under different observing conditions. An outlook on possible modifications to further enhance ALFA's performance is given at the end.

3.1 System description

We start with a short description of ALFA's optical design, which is very similar to the one of the PUEO curvature AO system (Rigaut et al., 1998a). A detailed description of the ALFA system is given by Kasper et al. (2000b) or Glindemann et al. (1997a). ALFA's advanced control components are a 97 actuator DM, a very low read-out noise CCD, and a control-computer that is fast enough to project 50 modes at a sampling rate in excess of 1 kHz. Unlike other AO systems, the current design incorporates multiple hexagonal microlens arrays ranging from 6 to 30 subapertures. A novelty are

the keystone shaped microlens arrays, which have not yet been used for SHS in astronomy. Finally, the logic of the modal control system is explained. Hereby, a focus is put on a simple scheme that allows us to determine the best framerate depending on the guide star brightness.

3.1.1 Optics

Two off-axis paraboloids are the main imaging elements of the optical system. The first paraboloid images the telescope pupil onto the DM, and the second paraboloid re-images the telescope focus on the infrared camera. The design goal was to have the re-imaged focus at the same position as the telescope's Cassegrain focus without any optical elements from ALFA. The telescope can thus easily be used without ALFA just by sliding two mirrors out of the beam.

Since the image quality of a high order AO system is always close to the diffraction limit, the pixel-scale of the infrared camera should be chosen accordingly. Without ALFA, large pixels are more convenient. Therefore, the default $f/10$ telescope focus is converted to an $f/24$ focus by the relay optics when using ALFA.

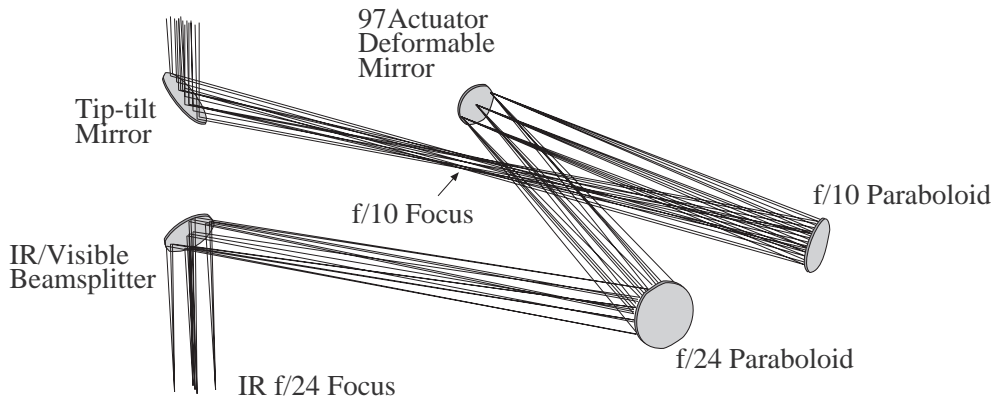


Figure 3.1: The main optical elements of ALFA. Explanations can be found in the text.

The diameter of the deformable mirror determines the basic parameters of the optical system displayed in Figure 3.1. With the $f/10$ telescope focus in the front focal plane of the $f/10$ paraboloid, its focal length has to be 662 mm in order to image the telescope pupil onto the (tilted) DM with an effective diameter of 66 mm. After reflection at the DM, the now collimated beam is intercepted by the $f/24$ paraboloid with a focal length of 1594 mm to re-image the telescope focus on the infrared camera. The first folding mirror serves as a tip-tilt mirror, and the last one acts as a beam-splitter, reflecting the infrared light downwards on the science camera and transmitting the visible light to the wavefront sensor arm.

The ALFA system is mounted on a breadboard as presented in Figure 3.2. Additional components other than those described above fall into five groups: the FISBA interferometer (FISBA Optik, St. Gallen, Switzerland), the $f/10$ reference fiber, the TV guider, the SHS, and the tip-tilt sensor with the sodium/visible beam-splitter to separate LGS light from that of the tip-tilt NGS.

The *f/10 reference fiber* is a monomode fiber with a core diameter of $3.6\mu\text{m}$ that serves as a “perfect” point source for alignment and calibration. It can be remotely moved in and out of the beam.

The *Twyman-Green interferometer* from FISBA is used to monitor the surface of the DM. This turned out to be an absolute necessity, because the microlens arrays are too coarse to sample the DM actuators and their deformations properly.

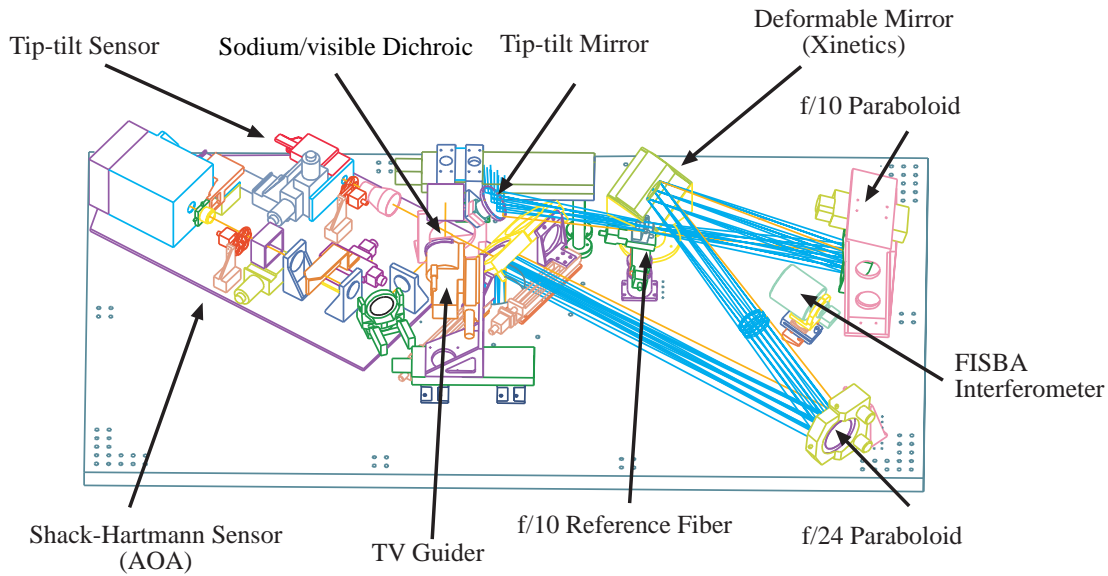


Figure 3.2: A CAD drawing of the ALFA breadboard with all opto-mechanical elements. The optical path in Figure 3.1 is extracted from this drawing. Explanations are found in the text.

The *TV guider* is used for object acquisition. As the field of view (FOV) of the infrared camera Omega-Cass (see Section 4.2.2) in its very high resolution mode ($0.04''/\text{pixel}$) is only 40 arcseconds, it is very helpful to have a sensitive camera with a 4 arcminutes FOV to find astronomical objects and maintain the pointing of the LGS.

3.1.2 Control components

Like other current AO systems, 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 to the process to adjust the controlled variable.

Figure 3.3 shows the principle of AO. The momentary atmospherically distorted wavefront (disturbance) is added to the momentary DM surface, and the resulting wavefront is divided. One part, in general the infrared light, propagates to the output science camera to form an image and the other, the visible light, is transmitted to the wavefront sensor. This works because the atmospheric dispersion is weak resulting in a wavefront that looks similar at NIR and visible wavelengths. The SHS (see Section 2.2.1) subdivides the pupil into subapertures and images each one separately. The difference between the locations of these subimages and some reference locations represent wavefront slopes. The control-computer estimates the shape of the residual wavefront from these slopes and derives an appropriate compensation which is sent to the DM actuators.

It is interesting to note that the full aperture tip-tilt is controlled by a separate loop which drives a fast steering mirror. This is necessary, since the small dynamical range of the DM prevents from compensating the large tip-tilt aberrations. Additionally, LGS operation requires tip-tilt control using an NGS (see Section 2.4.1). Details on the design and operation of the tip-tilt control loop are presented by Glindemann et al. (1997b).

The three essential components of the AO system displayed in Figure 3.3 are the SHS, the DM, and the control-computer. In the following, the specific types implemented in ALFA are described.

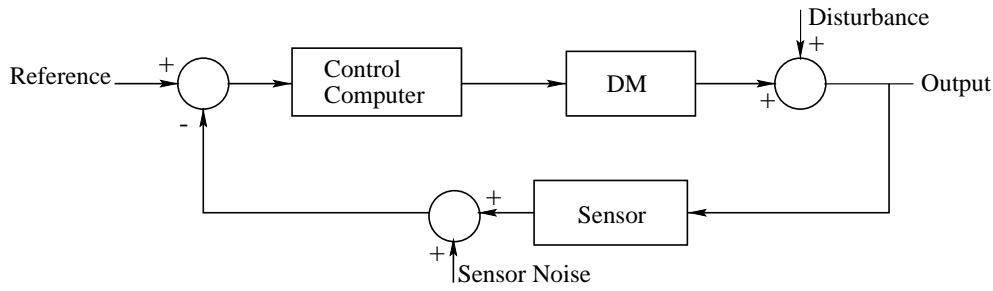


Figure 3.3: Principle of the AO closed-loop system.

ALFA SHS

The SHS used in ALFA was built by AOA (Adaptive Optics Associates, Cambridge, USA). The close-up view in Figure 3.4 depicts the individual components.

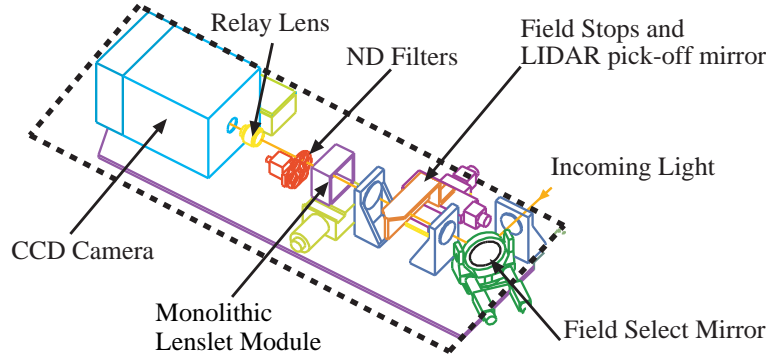


Figure 3.4: The components of the Shack-Hartmann sensor (see Figure 3.2 for the complete system). Explanations can be found in the text.

The *field select mirror* is conjugate to the telescope pupil plane, so that tilting this mirror results in shifting the image on the SHS-CCD. This mirror allows one to place a star onto the SHS within a 30'' radius from the optical axis.

Three *field stops* with different diameters (10'', 5'', and 2.5'') as well as a pick-off mirror for LIDAR observations can be inserted into an re-imaged focus. The field stops are required if the guide star is located in a dense stellar cluster leading to cross talk between subimages, or if the LGS is used and Rayleigh scattered light has to be blocked.

The *monolithic microlens module* (MLM) is also conjugate to the telescope pupil. It is mounted on a motorized stage such that the impressed microlens arrays can easily be interchanged at will. The re-imaged telescope pupil in the MLM plane has a diameter of only 5 mm.

The *relay lens* focusses the resulting spot pattern onto the CCD detector. The *CCD camera* and the relay lens are mounted on motorized stages such that the pixel-scale (usually 0.5''/pixel) can be adjusted, and the focus difference between an NGS at infinite distance and the LGS at about 90 km height (approximately 13 mm for the f/10 beam) can be accommodated.

The essential elements of the SHS, being the microlens arrays and the CCD camera, are discussed in greater detail in the following paragraphs.

Hexagonal microlens geometries. ALFA's current MLM is a circular plate made of BK7 that contains 4 hexagonal microlens arrays with different subaperture sizes. These arrays are coated with MgF to achieve maximum transmission. Figure 3.5 displays the hexagonal arrays and their positions with respect to the telescope pupil.

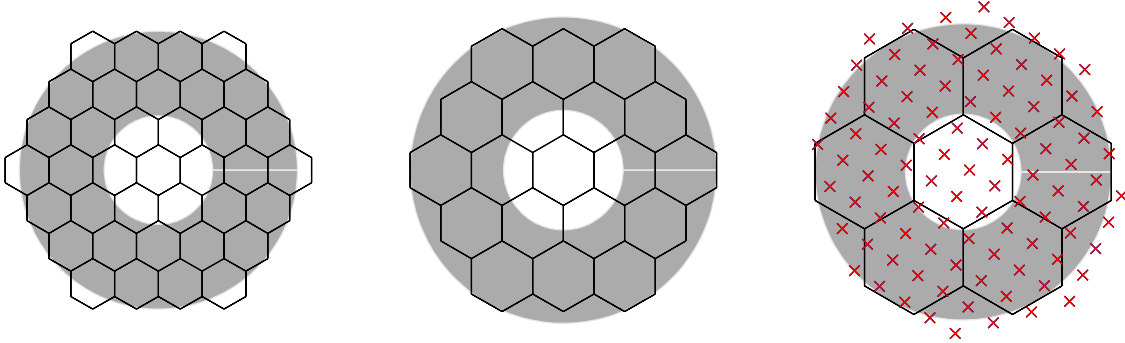


Figure 3.5: Hexagonal microlens configurations. The gray annuli represent the pupil of the 3.5-m telescope. From left to right: D7, D5 and D3 array. The red crosses denote positions of the 97 DM actuators.

The *D7 array* (D7 stands for 7 subapertures per pupil diameter) effectively covers the 5 mm pupil with 30 subapertures, the 7 innermost lenses being obscured by the secondary telescope mirror. The focal length of the microlenses is 45 mm, and the single hexagons have side lengths of $450\ \mu\text{m}$. In the telescope aperture, this translates into a photon collecting area of about $0.26\ \text{m}^2$ for each subaperture. The D7 is mostly used for NGS brighter than $m_V = 11$.

Two *D5 arrays* with different focal lengths (45 mm and 30 mm) provide 18 subapertures each. The 30 mm array was initially thought to produce smaller spots and slope vectors, making it well suited for mediocre seeing conditions and the extended LGS. It turned out that the 45 mm array can serve the same purpose when the magnification is altered by adjusting relay lens and CCD camera. The hexagons have side lengths of $578\ \mu\text{m}$ giving them a collecting area in the telescope aperture of $0.43\ \text{m}^2$ each. The D5 is mostly used for guide stars fainter than $m_V = 11$ as well as for the LGS.

The *D3 array* provides 6 subapertures. The focal length of the microlenses is 45 mm, and the single hexagons have side lengths of $962\ \mu\text{m}$ resulting in a collecting area in the telescope aperture of $1.18\ \text{m}^2$ each. The D3 only coarsely samples the wavefront, and reliable estimates can only be obtained for a maximum of 9 modes. Therefore, it should only be used when the guide star is very faint, i.e. fainter than $m_V = 13$.

Beside the microlens arrays, the MLM also carries a single lens useful for LGS diagnostics (Rabien et al., 2000).

Keystone microlens geometries. Starting from August, 2000, ALFA will be the first SHS AO system that uses a keystone-shaped microlens configuration. This geometry allows of an optimum filling of the annular telescope aperture without the expense of partially illuminated subapertures as for the hexagonal design. Additionally, if the subapertures in the different rings are designed such that all are of equal area, the spots are equally bright, and the measurement noise is uniform.

The keystone microlenses were designed such that their height equals the length of the center arc as illustrated by the dashed lines in the left drawing of Figure 3.6. The center of the lens is put at the intersection between them. The KS28 array covers the telescope aperture using 28 microlenses with an effective size of $0.29\ \text{m}^2$ each, while the KS7 provides 7 microlenses with $1.17\ \text{m}^2$. Both arrays have focal lengths of 45 mm.

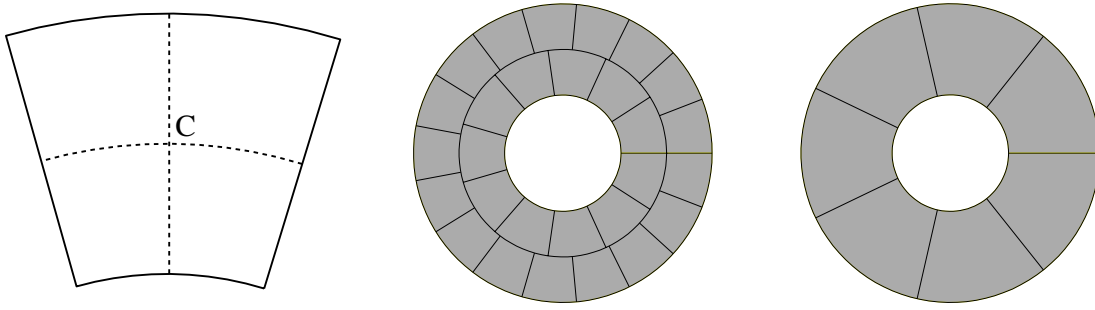


Figure 3.6: Keystone microlens configurations. The gray annuli represent the pupil of the 3.5-m telescope. *Left:* A single keystone lens with the center C. *Middle:* KS28 array. *Right:* KS7 array.

Wavefront sensor CCDs. The SHS uses a Lincoln Labs CCD (LLCCD) that is a thinned 64×64 pixel detector operated in a 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 electrons. The quantum efficiency between 600 nm and 750 nm exceeds 80 % (see Figure 3.7). The camera has a two stage thermo-electric cooler providing -35°C to reduce dark currents. Thermal energy is transported outside the camera via a heat-pipe. Read-out electronics and digital interface to the control system were built by AOA.

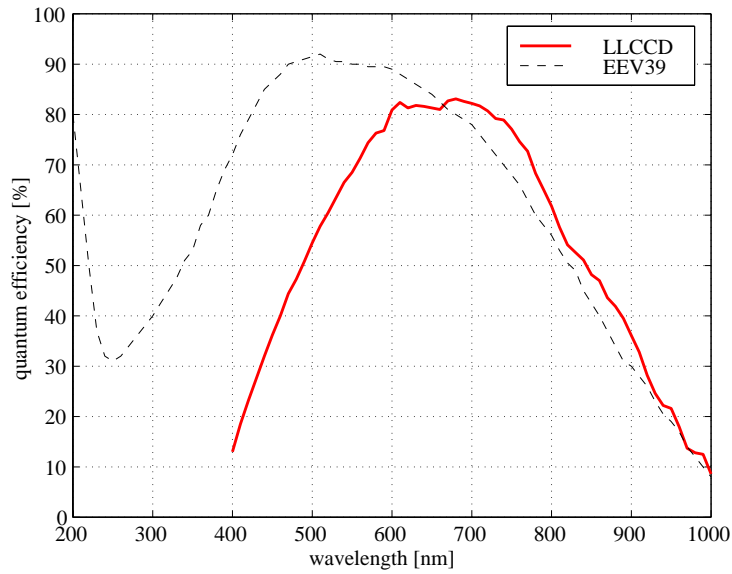


Figure 3.7: Quantum efficiencies of the SHS CCD (LLCCD) and the tip-tilt tracker CCD (EEV39).

The *tip-tilt sensor* is based on an AstroCam (Cambridge, UK) CCD camera with a 4201 controller; it shows less than 3.5 electrons read-out noise. The quantum efficiency of the thinned detector EEV CCD39 is around 90 % between 470 and 600 nm (see Figure 3.7). As an upgrade, first tests with an Avalanche Photodiode (APD) quad-cell (Bonaccini et al., 2000, Hippler et al., 2000) are underway. This device uses the APDs in photon counting mode with zero read-out noise. Since the quantum efficiency of the APDs will be higher than 70 % between 600 and 750 nm, we hope that the very low detector noise helps to reduce image motion considerably on guide stars as faint as $m_V = 17$, an improvement of one magnitude over the EEV CCD39.

Deformable Mirror

The DM was purchased from Xinetics Inc., USA. Its 97 PMN (lead magnesium niobate) ceramic piezo actuators are capable of delivering a $2\mu\text{m}$ inter-actuator stroke. From the manufacturer specifications as well as laboratory measurements, the piezos were found to reach their commanded positions within 0.2 ms. There are 11 actuators on a diameter of the telescope aperture, so 6 actuators cover a radius. Hence, the DM is capable to properly reproduce polynomials up to the 6th radial order, which according to Equation 2.16 includes the first 28 Zernike polynomials. Figure 3.8 shows the DM surface while a Karhunen-Loève mode of 4th radial order is applied. The cross-talk between actuators, i.e. the influence of an actuator command on the DM surface deformation at the position of its nearest neighbor, is below 10 % (Wirth et al., 1998).

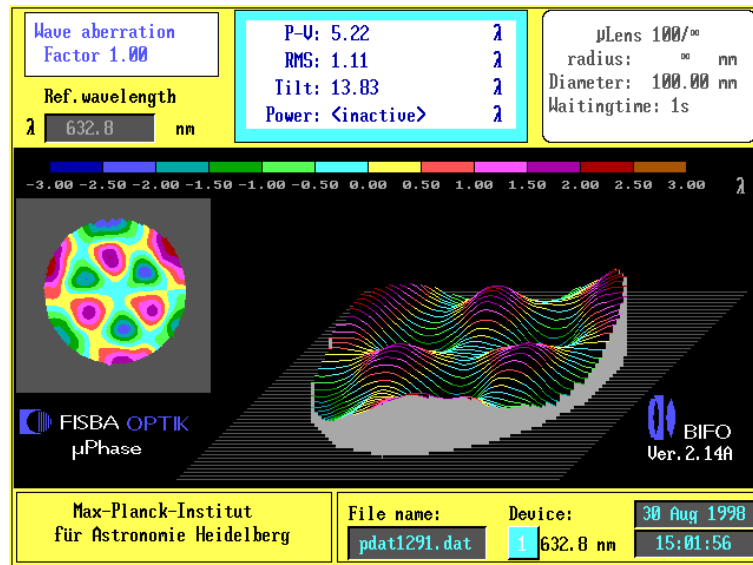


Figure 3.8: The Twyman-Greene interferometer measures the DM surface while a Karhunen-Loève mode of 4th radial order is applied.

The actuators are operated in their linear region between 35 and 105 Volts. Approximately 15 actuators lie too far outside the pupil annulus to influence the reflected wave (see Figure 3.5). 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 primary.

The rms aberration of the reflected wave is about 600 nm when the same voltage is applied to all actuators. The FISBA interferometer (see Section 3.1.1) can be used to flatten the surface by adjusting the individual actuators. This process is now fully automated; it only takes about one minute until a flatness of less than 60 nm rms is reached. The actuators show very little hysteresis, and their configuration remains the same when the temperature in the dome is stable. Thermal expansion of the piezos, however, causes the DM flatness quality to deteriorate, if the temperature changes, requiring a re-flattening of the DM.

In December 1998, a new electronic subsystem replaced the old actuator driver electronics from Xinetics. The new electronics is much more compact and can be attached to the telescope mounting several meters apart from the DM and the main mirror cell. It is now also possible to read back the actuator voltages and detect malfunctions and error conditions like disconnected cables between the electronic chassis and the DM.

Control Computer

The hardware consists of a VME-bus based real-time computer (RTC) system running VxWorks (VME system A). All matrix operations, such as the determination of spot locations and the reconstruction of modal coefficients, are implemented using five digital signal processor (DSP) boards. Each board contains 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 $833\mu\text{s}$ execution time resulting in a maximum loop frequency of 1206 Hz. The mathematical operations are scalable up to 349 actuators by adding more DSP boards. The VME/VxWorks host is connected to workstations running the graphical user interface (GUI) via Ethernet. The Experimental Physics and Industrial Control System (EPICS) database maintains the status of the entire system and serves as the logical link between the GUI and the RTC. The motorized stages on the ALFA-breadboard receive their commands via an RS-232 link from the VME host. A detailed presentation of the computer system is given by Hippler et al. (1998).

3.1.3 Modal control

As described in the preceding section, there are many fewer SHS slope measurements n_s than DM actuators n_a with all the microlens arrays (a maximum of 60 slopes is available with the D7 array to drive 97 actuators). This coarse spatial sampling of the wavefront does not permit the determination of the phase at all actuator positions (*zonal control*). Therefore, a modal compensation approach had to be chosen for ALFA, because some interpolation scheme needs to be used to derive actuator commands. The concept of modal control was described in Section 2.1.2.

Although the wavefront error is a spatially distributed signal (i.e. infinite dimensional), only a n_a -dimensional subspace of the wavefront (*mirror space*) can be affected by the actuators, only a n_s -dimensional subspace of the wavefront can be observed by the SHS, and only the intersection of the two subspaces can be controlled in a feedback loop. If there are no redundant actuators, the influence functions form a basis set of the mirror space. This set is not necessarily orthogonal, and Section 2.1.6 described how an orthogonal set can be found. For the ALFA DM, it is assumed that the cross-talk between the single actuators is negligible as long as the actuator commands produce a relatively smooth surface, making the influence functions an orthogonal basis set for the mirror space. Hence, the DM representation of an arbitrary mode is given by a 97-dimensional *injection vector* which contains the values of the mode sampled at the actuator positions.

Section 2.3.3 explained our strategy for finding a basis set for the intersection between the SHS- and DM subspaces by truncating K-L functions at a number that keeps cross-talk low and occasionally, for bright guide stars, fill up the measurement space with sensor modes.

Control Architecture

The data flow in ALFA's modal control architecture is shown in Figure 3.9. The box encloses a block diagram representation of the open-loop AO system. The compensated wavefront $\phi(x, y, t)$ is the sum of the DM surface ϕ_s (negative sign) and the input wavefront ϕ_d disturbed by atmospheric turbulence. The reconstruction matrix \mathbf{R} is selected to estimate the modal coefficients $\hat{\mathbf{a}}(t)$ from the digitized SHS signal vector $\mathbf{g}(t)$. This estimate is then to be subtracted from the reference values \mathbf{a}_r (usually zero) to obtain an error signal. This signal is then used by the compensator $K(z)$ to determine the modal drive signals $\mathbf{u}(t)$ that are appropriate for an effective compensation. The actual actuator drives $\mathbf{u}_d(t)$ that affect the wavefront shape by modifying the shape of the DM are calculated by multiplying $\mathbf{u}(t)$ by

the injection matrix \mathbf{I} whose columns are the injection vectors, i.e. the actuator representation of the control modes.

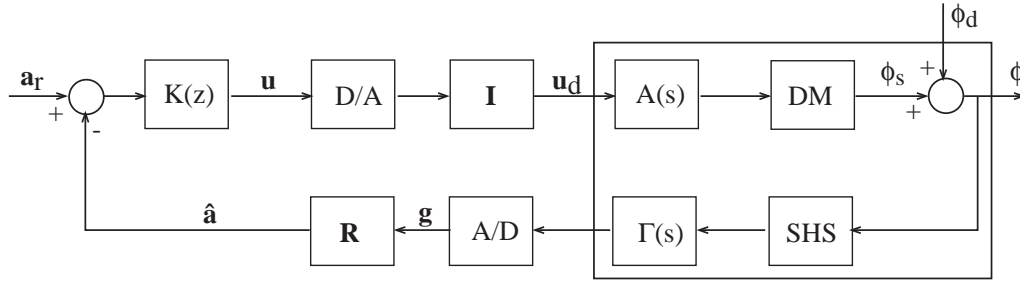


Figure 3.9: Modal control system architecture. The box encloses the open-loop block diagram. Explanations can be found in the text.

The dynamics that affect the system are the actuator transfer function $A(s)$ and the transfer function $\Gamma(s)$ representing the integration of incoming light over one frame of the SHS and the delays in the loop due to camera readout, slope computation, and wavefront reconstruction. At frequencies of 300 Hz and below, the actuator dynamics can be neglected, and the system's dynamics are reasonably modeled by a single delay. At higher frequencies, more complicated dynamical models are required to represent the system (Looze et al., 1999).

Calibration

The aim of the calibration procedure is to define a reference wavefront and to establish the interaction between DM control modes and the SHS. Although it is possible to derive analytical relationships that relate the DM modes to the corresponding SHS measurements, these relationships would be sensitive to modeling assumptions and to the alignment of the SHS microlenses with respect to the DM. ALFA bases its modal reconstruction on a static identification procedure to determine the SHS signal produced by a given control mode.

Reference. The first calibration step is to define a reference wavefront and subsequently the default spot locations. The ultimate criterion for a “good” reference wavefront at the SHS is a high image quality on the science camera. This does not necessarily coincide with a flat wavefront, because there are a number of potentially aberrated elements between the IR/visible beam-splitter (where the wavefront should be reasonable flat if the science camera has a good optical quality) and the SHS (see Figure 3.4). To achieve this goal, a white light reference fiber is positioned in the f/10 focus and imaged with the science camera. Low order static aberrations are removed by applying static offsets to the DM. The dominating aberration is about one wave of astigmatism (at $2.2\mu\text{m}$) which we believe is due to the off-axis parabola glued in at a slightly false rotation angle.

Once the DM cancels the static aberrations, such that a satisfactory high K-band Strehl ratio of close to 90 % is obtained on the fiber, the reference is defined. Figure 3.10 shows a science camera K-band image of the f/10 reference point source after some static aberrations are compensated by the DM. The Strehl ratio in this case is about 85 %.

Identification of the control modes. After defining the reference, the next step is the identification of the control modes. The reference wavefront gives us the reference spot positions. Now, the control modes are sequentially applied to the DM with an unit coefficient ($a_j = 1$), and the corresponding

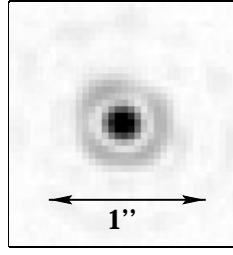


Figure 3.10: f/10 reference fiber imaged with the science camera in K band.

slope vectors \mathbf{g}_j are calculated by subtracting the reference positions from the recorded spot positions. The interaction matrix \mathbf{D} is then given by the \mathbf{g}_j arranged in its columns. Assuming a linear relationship between mirror shape and SHS measurements, the n slopes that result from an arbitrary superposition of the control modes can be computed from

$$\begin{pmatrix} g_{1x} \\ g_{1y} \\ g_{2x} \\ \vdots \\ g_{\frac{n}{2}y} \end{pmatrix} = \mathbf{D} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix},$$

or in a more compact form

$$\mathbf{g} = \mathbf{D}\mathbf{a}. \quad (3.1)$$

When calibrating for LGS operation, it turned out to be important to use the sodium/visible dichroic (see Figure 3.2) instead of the full reflective mirror, because chromatic aberrations in the lenses of the SHS have a significant impact on the calibrated focus term.

Reconstruction

The aim of the reconstruction algorithm is to estimate the modal coefficients $\hat{\mathbf{a}}$ from the slopes. Given a linear estimator \mathbf{R} the reconstruction can be written as

$$\hat{\mathbf{a}} = \mathbf{R}\mathbf{g}. \quad (3.2)$$

Since $\hat{\mathbf{a}}$ and \mathbf{g} are vectors, \mathbf{R} is a matrix. Usually \mathbf{R} is chosen as the pseudo-inverse of \mathbf{D} , i.e.

$$\mathbf{R} \equiv \mathbf{D}^+ = (\mathbf{D}'\mathbf{D})^{-1}\mathbf{D}'. \quad (3.3)$$

As already discussed in Section 2.3.2 the condition number of \mathbf{D} has to be kept low, and it is required that $m \leq 2n$ to be able to invert the term $(\mathbf{D}'\mathbf{D})$. An analysis of cross-talk for Zernike- and K-L basis sets used together with the ALFA's microlens arrays is presented in Section 3.4.

As described in Section 2.3.6, it is possible to make use of a priori knowledge on the statistics of measurement noise and modal coefficients in order to calculate an estimate superior to \mathbf{D}^+ in the sense of a lower estimation error variance. These advanced estimators, namely the MAP and WLS, are now available for ALFA. The calculation, however, currently requires that open-loop (to calculate r_0) and closed-loop SHS data (to estimate the covariance of the corrected control modes and measurement noise) are taken before the observations.

MAP and WLS have to be used with caution, since their performance depends on how well the assumed statistics match with reality. They may perform worse than the LS in a strongly non-Kolmogorov atmosphere or in unstable conditions with large fluctuations of r_0 and/or transparency. Very soon, unstable conditions will be handled by frequently updating the current values of r_0 , and the covariance matrices in closed-loop. The implementation of the algorithms that are required to obtain this information is part of a current PhD work at MPIA. A test of the MAP and WLS estimator is presented in Section 3.5.

Compensation

The dynamic compensation of distorted wavefronts is a subject of control theory and beyond the scope of this work. In the following, the impact of closed-loop control on AO is nevertheless shortly described and possible optimization strategies are presented.

The theoretical behavior of temporal modal power spectral densities (PSD) as examined by Conan et al. (1995) was briefly reviewed in Section 2.1.1. Figure 3.11 shows the measured PSDs of three K-L modes of 2nd, 4th, and 6th radial order. The general behavior agrees remarkably well with the theoretical curves sketched in Figure 2.1. The theoretical spectra are constant at low frequencies and roll off steeply with $f^{-17/3}$ at frequencies higher than a cut-off which is a function of a mode's radial order. The measured open-loop spectra, however, rarely show the expected steep decrease in the high frequency regime. On Calar Alto, we usually observe roll-offs proportional to about f^{-3} . A possible explanation for this observation is that the combination of different layers with different wind-speeds and hence different cut-off frequencies flattens the steep decrease of the PSD finally levelling off in a flat noise floor. Furthermore, it can be seen that the noise floor is lower for high order modes as already discussed in Section 2.2.5.

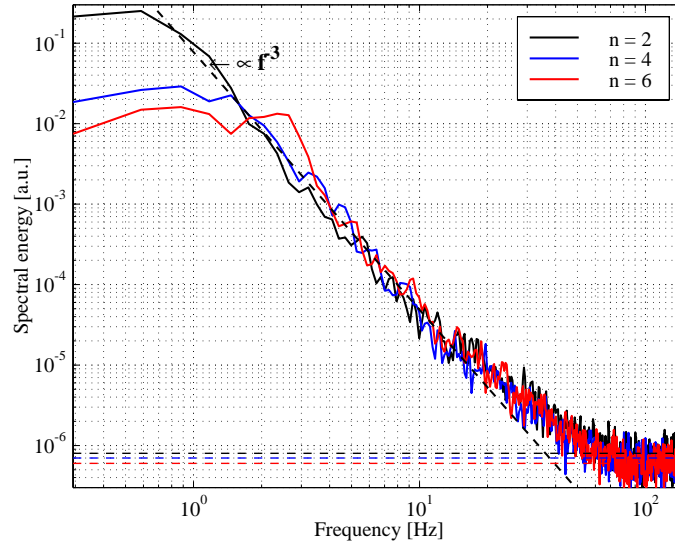


Figure 3.11: Estimated open-loop PSDs for three modes of 2nd, 4th, and 6th radial order. Also drawn are the estimated measurement noise floors which are smaller for high radial degrees and a least-squares fit to the high frequency part of the open-loop spectrum.

The common performance metric for AO systems is the residual phase variance over the telescope aperture which is the sum of the variances of the single modal coefficients (cf. Equation 2.13). Thus, the objective of the control loop is to minimize the residual variance of each individual control mode.

In the Fourier domain, the variance of a mode is given by the integral over its PSD. Therefore, the closed-loop performance of the control system is determined by the residual PSD $P_{res}(f)$ of the compensated mode, which in turn is driven by the atmospheric signal PSD $P_{atm}(f)$ and the noise PSD $P_n(f)$. Specifically, $P_{res}(f)$ is given by (Gendron and Lena, 1994):

$$P_{res}(f) = H_{cor}(f)P_{atm}(f) + H_n(f)P_n(f), \quad (3.4)$$

where the $H_{cor}(f)$ and $H_n(f)$ are the squared moduli of the correction transfer function $h_{cor}(f)$ and the noise transfer function $h_n(f)$ of the control system which are related by

$$h_{cor}(f) \approx 1 - h_n(f). \quad (3.5)$$

Figure 3.12 depicts $h_{cor}(f)$ and $h_n(f)$ for different gain factors¹. At low frequencies, where atmospheric disturbance is strong, $h_{cor}(f)$ is small and the disturbance is compensated, whereas $h_n(f)$ is near unity and the measurement noise appears unattenuated in the residual signal. For frequencies above the correction bandwidth (frequency at which $h_{cor}(f)$ equals unity), the control loop amplifies the disturbance to some extent as demonstrated in Figure 3.13, while the noise no longer propagates through the loop. The conclusion is that the first term of Equation 3.4 decreases with correction bandwidth, while the second term increases. Thus, it should be possible to find a minimum by adjusting the bandwidth which can be altered by either changing the loop gain or the framerate.

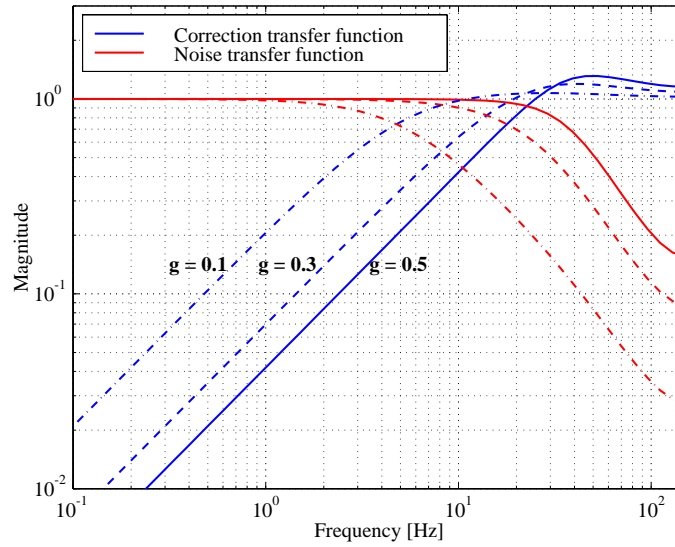


Figure 3.12: Modulus of correction- and noise transfer functions of the ALFA standard controller for different gains. The framerate is 300 Hz.

Framerate optimization. Gendron and Lena (1994) developed a method to calculate the optimum gain providing a bandwidth which minimizes the residual modal variances, i.e. the integral over Equation 3.4 depending on measurement noise. In most cases, this optimum bandwidth turned out to be close to the intersection of the turbulent spectrum with the noise floor f_i as indicated in Figure 3.13. Knowing this fact and the temporal behavior of the modes (which was studied by Conan et al. (1995)

¹The gain is a filtering factor applied to the measurement that regulates loop stability. High gain factors make a control loop fast but less stable.

after Gendron had published his work), a quite simple scheme can be drawn to determine the optimum bandwidth for all modes by simply adjusting the framerate:

The control loop should always use the highest gain that ensures stability. The standard control parameters used with ALFA are selected to provide a correction bandwidth of approximately 1/15th the framerate F (Looze et al., 1999), i.e.

$$f_b(F) \approx \frac{1}{15}F. \quad (3.6)$$

According to Equation 2.38, the noise floor is proportional to n_{ph}^{-1} and thereby also proportional to F . Thus, the frequency at which the noise floor intersects the modal PSD decreasing with f^{-x} is

$$f_i(F) = f_i(F_0) \left(\frac{F}{F_0} \right)^{-1/x}, \quad (3.7)$$

where F_0 is some initial framerate at which the noise floor intersects the modal spectrum at $f_i(F_0)$. Then, the optimum framerate is approximately at the point where $f_b(F)$ equals $f_i(F)$, i.e.

$$F_{opt} = \left(15 \cdot f_i(F_0) \cdot F_0^{1/x} \right)^{x/(x+1)}. \quad (3.8)$$

The nice thing about this number is that it roughly applies to all control modes, because under usual conditions their respective PSDs intersect the noise floors at approximately the same frequency (see Figure 3.11). It was already noted by Gendron and Lena (1995) that in many cases the optimum correction bandwidth is similar for all control modes. However, in the case of a very faint guide star and/or fast seeing due to a combination of a small r_0 and high windspeed, the constant low-frequency power of high order modes may lie beneath their noise floor. Then, it should not be attempted to compensate for these modes at all. It is worth emphasizing that lowering the framerate is superior to lowering the gain, because both reduce the bandwidth but only lower framerates increase the flux and thus reduce measurement noise.

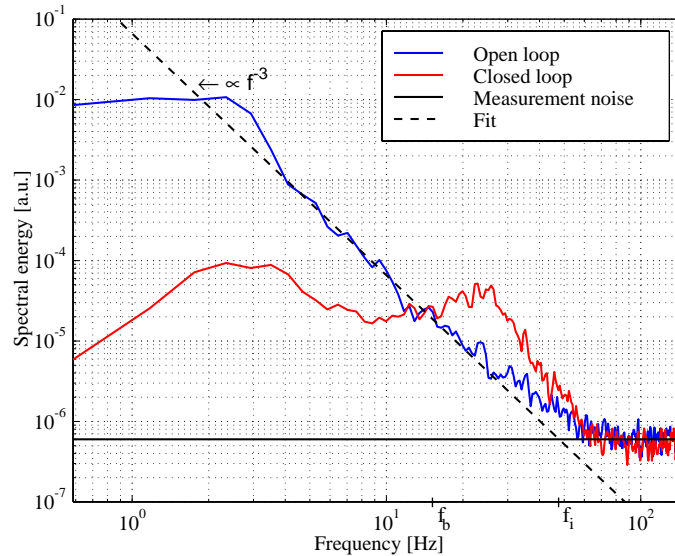


Figure 3.13: Representative estimated PSDs in open- and closed-loop. Also drawn are the estimated measurement noise floor and a least-squares fit to the high frequency part of the open-loop spectrum.

An illustrative example is presented by Figure 3.13: F_0 was 300 Hz, f_b about 15 Hz which is 1/20th of F_0 , $f_i(F_0)$ about 47 Hz, and the modal PSD decreased with f^{-3} . According to Equation 3.8 the framerate that should provide the best performance is $F \approx 700$ Hz.

Optimum Power Control. The above strategy will deliver good results under regular observing conditions, but it is not optimum. The basic procedure for implementing the optimal AO compensator is to gather a set of open-loop data, identify the dynamics of each single mode, and compute the set of control parameters (not only the bandwidth) that provides the maximum amount of disturbance rejection for each mode. In other words, the objective is to minimize the residual variance for each single mode using its PSD to design h_{cor} . This method has been developed based on Kalman filtering algorithms and explored for the use with ALFA (Looze et al., 1999). The actual implementation for routine operation, however, has still to be done.

3.2 Centroid estimation

A crucial limit to the performance of an SHS is optimal estimation of the centroid of a low-resolution spot on the detector array. Equation 2.38 describes the general dependence of measurement noise on spot size and flux. This equation does not take into account measurement related issues like sampling with different pixel-scales, or the influence of detector read-out noise at very low fluxes. Further, it does not explore how different centroiding algorithms affect the measurement accuracy. Irwan and Lane (1999) studied the noise variance for a perfect (no read-out noise, no effects due to finite pixel size) but truncated (finite size) CCD. Their main conclusion is that depending on the flux, an optimum array size exists that balances the accuracy deterioration caused by noise from far-off pixels amplified by the lever effect, and the accuracy improvement caused by sampling the outer parts of the spots.

Here, a simple Monte-Carlo approach is used which allows us to empirically study different centroiding issues by changing the model parameters detector size, read-out noise, pixel-scale, and the power to which the intensities are raised. Based on the simulations, the optimum pixel-scale is found as well as the best “simple” centroiding algorithm which, surprisingly, involves manipulation of the intensities by raising them to the 1.5th power. In this context, “simple” means that the estimators are based on the center of gravity approach; no attempt was made to use a MAP centroiding algorithm that considers correlation between individual spot motions and noise variance (Sallberg et al., 1997). Fast fitting algorithms were not studied as well, because their performance under realistic conditions is doubtful. Parameter-fits usually require a model for the underlying intensity distribution. In reality, the morphology of shortly exposed spots varies with time and with the subapertures rejecting the assumption of a Gaussian or parabolic shape. Iterative procedures are in general too slow to be used in a closed-loop. Most of the material in this section was presented at a conference on wavefront sensing in Canterbury (Kasper et al., 1999).

3.2.1 Monte-Carlo simulation

Simulation process.

1. First, the simulation code constructs a data cube consisting of n detector arrays of the desired size, where n is also the number of different Monte-Carlo realizations. For the upcoming results, $n = 500$ was used in all cases. Detector sizes can be for example 2×2 pixels for a quad-cell configuration or 8×8 pixels to resemble the usual ALFA configuration.

2. The data cube is filled with Poisson-distributed random numbers to simulate detector read-out noise.
3. Arbitrary spot positions x_1, y_1, \dots, y_n are created by $2n$ random numbers that are uniformly distributed between -0.5 and 0.5 pixels around the center of a detector array.
4. These positions represent centers of Gaussian spots of a given size θ . The intensities in the individual pixels are calculated by integrating a Gaussian profile shifted by the correct distance over the pixel area. Since the Gaussian profile is normalized, the intensities are multiplied by the total flux in the subaperture.
5. Photon noise obeys Poisson statistics. However, for all but the lowest intensities Gaussian statistics are a very good approximation and easier to compute. Hence, photon noise was simulated by adding Gaussian random numbers to the individual pixels whose standard deviation equals the square-root of the respective pixel intensity.
6. The conversion of photo-electrons into detector counts is considered. The ALFA SHS camera has two gains providing 1 count per 2.5 photo-electrons or 1 count per 5 photo-electrons. It is assumed that each photon produces an electron.
7. The final image cube is now used together with the desired centroiding algorithm to estimate spot positions $\hat{x}_1, \hat{y}_1, \dots, \hat{y}_n$.
8. Finally, the centroiding accuracy is calculated by the rms deviation of the estimated positions from the initial positions.

Centroid estimation algorithms. Four different centroid estimation algorithms were investigated; all are based on the center of gravity or weighted pixel average (WPA) algorithm:

$$\mathbf{r}_c = \frac{\sum_i \mathbf{r}_i \cdot I_i}{\sum_i I_i}, \quad (3.9)$$

where \mathbf{r}_i denotes the x- and y-coordinates of the i th pixel, and I_i its intensity. The specifications of the different algorithms are listed below. They all aim to limit the impact of read-out and/or photon noise.

The constrained WPA (CWPA) algorithm searches the brightest pixel in the frame and uses only a limited area surrounding this pixel. This excludes weakly illuminated pixels from the centroid estimation. It should be noted that according to Equation 3.9, the distance between a pixel and the centroid has a lever effect on the WPA algorithm. Thus, far-off pixels should always be excluded.

The thresholded WPA (THRESH) algorithm is quite similar to the CWPA. Here, all pixels with intensities below a certain threshold are excluded from the calculation. The THRESH is faster than the CWPA, because it has to pass through the array only once. It also handles extended or multiple reference sources better than the CWPA.

The weighted WPA (WWPA) algorithm raises all pixel values to the 1.5th power before the calculation. After an idea by Ric Davies, the motivation for this algorithm was to weight each pixel with its SNR that is proportional to square-root of its intensity in the photon limited case. The WWPA can easily be combined with pixel masking techniques.

3.2.2 SHS measurement noise determination

A method to determine SHS measurement noise was originally proposed by Rousset (1992) and described as method (b) by Gendron and Lena (1995). It uses the fact that the autocorrelation function of measured wavefront slopes consists of two features as displayed in Figure 3.14: A smoothly varying component due to the temporally correlated turbulence, and a spike at the origin due to uncorrelated measurement noise. The *initial value theorem* (Franklin et al., 1994) can be used to show that the first derivative of the slope autocorrelation function is zero at the origin. Therefore, fitting a parabola is well-suited to extrapolate the turbulent signal to the origin. The noise variance is then just given by the difference between the origin of the autocorrelation function and the extrapolated turbulent signal.

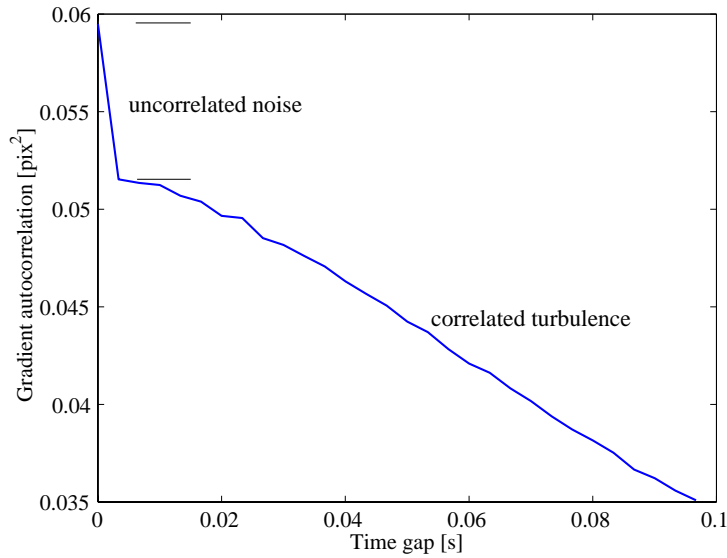


Figure 3.14: Autocorrelation function of a measured wavefront slope time-series near the origin.

It should be noted that the correlation method is not sensitive to aliasing noise, because aliasing is produced by high order modes which are temporally correlated. The correlation time of these modes is usually much longer than the sampling interval, and hence can not be confused with the measurement noise spike. Further, the correlation method works in closed-loop also, because the feedback system does not affect the behavior of wavefronts on small time-scales comparable to the sampling interval.

An example of noise measured for the D5 microlens array is depicted in Figure 3.15. The plot to the right shows the 36 slopes ordered such that x- and y- slopes for the respective subapertures are successive. Apparently, the x-slopes are noisier than the y-slopes. This is commonly observed with ALFA and must be due to the spots being elongated in the x-direction. This elongation could either be produced by coma aberrations of the SHS optics, or by a CCD read-out effect called *charge transfer inefficiency*, where the fast read-out register in x-direction causes a smearing of charge. The left drawing displays the rms measurement noise in the individual subapertures, illustrating another effect: Partially illuminated subapertures gather less light and thus exhibit a larger measurement noise.

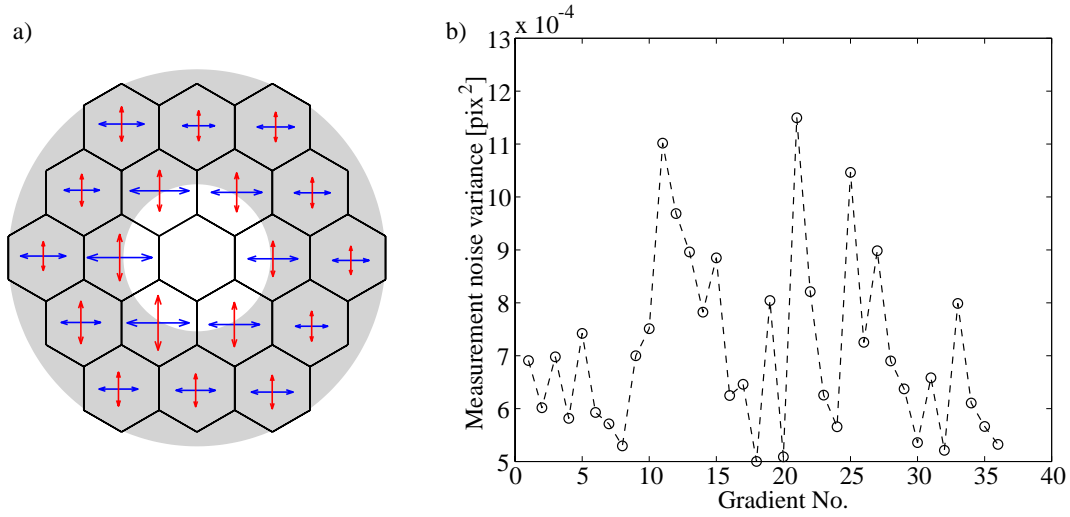


Figure 3.15: Measurement noise variance as determined with the autocorrelation method for the D5 microlens array. Figure a) illustrates the larger noise vectors of partially illuminated subapertures.

3.2.3 Impact of pixel-scale

The most easily changed operational parameter of ALFA's SHS is its pixel-scale. In general, large pixels reduce the effects of read-out and photon noise, but do not have the dynamic range and accuracy of small pixels.

Figure 3.16 plots the centroid error variance against the flux level, expressed in photo-electrons per subaperture, for different pixel-scales. The simulations were performed with $4 e^-$ read-out noise (LLCCD) and the WPA algorithm. The size of the detector array with respect to the spot size was kept constant in order to avoid variable truncation effects. Specifically, the array was four times bigger than a spot's full width at half maximum (FWHM) in all cases.

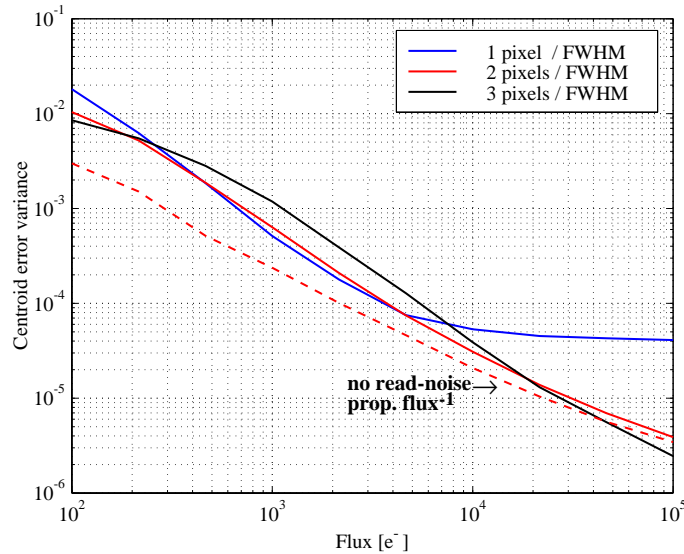


Figure 3.16: Influence of the pixel-scale on the centroiding error for a $4 e^-$ read-out noise CCD. Array sizes are approximately equal, the standard WPA without masking was used. The dashed line represents the no read-out noise case.

In general, sampling a spot with 2 pixels over its FWHM is a good compromise between low-flux sensitivity and high-flux accuracy. Smaller pixels are not quite as accurate at low flux levels, while larger ones are considerably worse at high flux levels. The dashed line in the figure represents the zero read-out noise case for 2 pixels per FWHM sampling. In good agreement with Equation 2.38, the error variance for such a detector is inversely proportional to the flux level. At low to medium flux-levels, a zero read-out noise detector needs less than half the flux of a $4 e^-$ read-out noise one to achieve a given accuracy, yielding a gain of about one magnitude in guide star brightness.

Note that an $m_V = 15$ guide star delivers 100 photo-electrons per spot for a perfectly transmitting SHS with 60 cm subapertures running at a framerate of 200 Hz. A flux level of about 10^5 photo-electrons per spot would just saturate ALFA's LLCCD.

3.2.4 Comparison of algorithms by simulations

Using a pixel-scale of 2 pixels per FWHM and $4 e^-$ read-out noise, the performance of WPA, CWPA, THRESH, and WWPA algorithms was simulated. Hereby, also combined modifications such as intensity weighting together with setting a threshold were investigated. The results of the extensive simulations can be summarized in some major statements:

1. Narrowing the detector area to 6×6 pixels generally delivers the best results. Larger arrays are slightly better at very high flux levels but considerably worse at lower ones. Smaller arrays tend to truncate the moving spots at the array edges leading to large centroid estimation errors.
2. The WWPA is the most accurate algorithm at all flux levels. Hereby, weighting the intensities to the 1.5th power yields the best results. Bigger exponents reduce the performance at high flux levels, smaller ones reduce it at low levels without improving the performance in the respective other regimes.
3. A Threshold of about 3 times the read-out noise should be set when using a WPA without intensity weighting, where the exact value does not critically affect the performance. Setting a threshold has no obvious effect when used in combination with the WWPA.

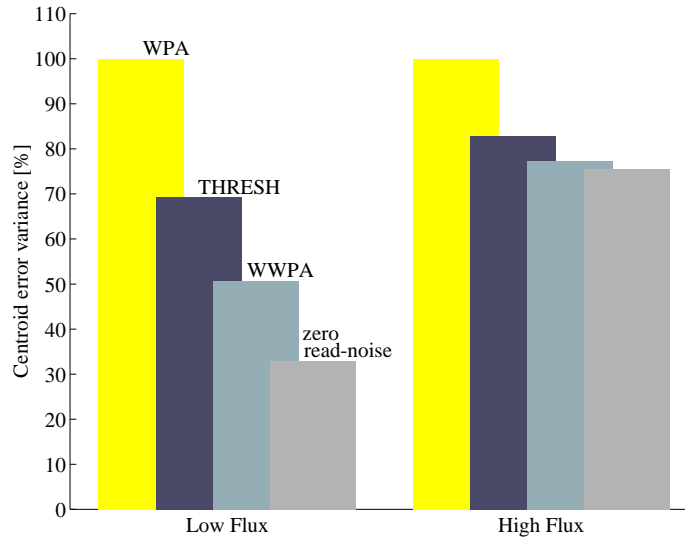


Figure 3.17: Centroid algorithms compared by simulations at signal levels below and above $2500 e^-$.

Figure 3.17 compares the performance of WPA, THRESH with a threshold of 3 times the read-out noise, WWPA, and a zero read-out noise detector at flux levels below and above 2500 e^- corresponding to an $m_V = 11.5$ star. 5 distinct flux levels in the respective regimes were averaged to provide the data points. The centroid error variances were further normalized to the one obtained with the WPA.

At low flux levels, the WWPA provides centroid estimates with only half the error variance of those obtained by the WPA yielding a gain of almost one magnitude in guide star brightness. If we could further purchase a zero read-out noise detector, another half of a magnitude could be gained. At high flux levels, the differences between the algorithms are smaller and at a much lower absolute level.

3.2.5 Comparison of algorithms by experiment

CWPA, THRESH, and WWPA algorithms were implemented in the ALFA software, and their performance was tested at the telescope. The data-sets used throughout the experiment consist of 4000 slope vectors provided by the D5 array. The system was running at a framerate of 300 Hz. An array size of 8×8 pixels was used in all cases, and the CWPA was set up to use 5×5 pixels around the brightest one. The THRESH used a threshold value of three detector counts roughly corresponding to three times the read-out noise, and the WWPA used an exponent of 1.5. As well as different algorithms, three different light levels were investigated using neutral density filters. The ND0.04 filter is almost transparent, the ND1.0 has a light attenuation factor of 10, and the ND1.5 has one of $10^{1.5} \approx 31$. The noise variance was determined by the correlation method.

In order to eliminate effects that could be introduced by variable observing conditions, 96 data sets were taken in a completely randomized order. The 96 sets consist of 8 sets for each of the 12 combinations of 4 algorithms and 3 light levels.

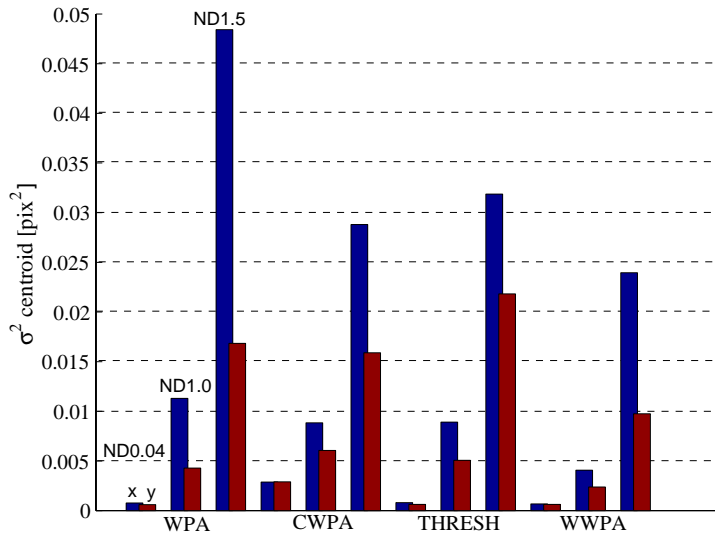


Figure 3.18: Measurement noise variance for different centroid estimation algorithms and light levels measured at the telescope.

The results from the experiment are shown in Figure 3.18 and nicely confirm the results obtained from the simulations. The WWPA performs almost twice as good as the WPA at all light levels, and the THRESH lies between the two. Again, the x-slopes are noisier than the y-slopes as already discussed for Figure 3.15. A statistical analysis was performed based on the Student t-test on the

mean (Frieden, 1991). The WPPA is clearly superior over all other algorithms with a confidence level greater than 99 %. The WPA and the CWPA performances are most sensitive to the light level, with the WPA performing better at high levels and the CWPA better at reduced levels.

3.2.6 Quad-cells

A quad-cell is the minimum pixel configuration required to measure x- and y-shifts of a spot. It only needs 2 pixels per measurement value minimizing the effects of measurement noise. However, a quad-cell also possesses considerable drawbacks: The linear range of the slope measurements is limited by the spot size θ ; in the diffraction limited case ($\theta \approx \lambda/d$), only one wave of tilt over a subaperture is measurable. Furthermore, the measurement signal is a function of θ (see Equation 2.31) making the quad-cell quite sensitive to a spot's shape. Hence, the use of quad-cells for LGS wavefront sensing is questionable, because the LGS neither has a well defined shape nor a fixed position on the sky (see Section 2.4). However, in closed-loop the small dynamic range of a quad-cell should be sufficient to provide linear measurements of the residual corrected slopes when using a point source as a reference.

Figure 3.19 compares the centroid error variance of detector arrays and quad-cells for two detector read-out noise levels. The WPPA algorithm was used to estimate centroids with the $4 e^-$ read-out noise CCD to ensure that the curves resemble the optimum achievable performance in each case.

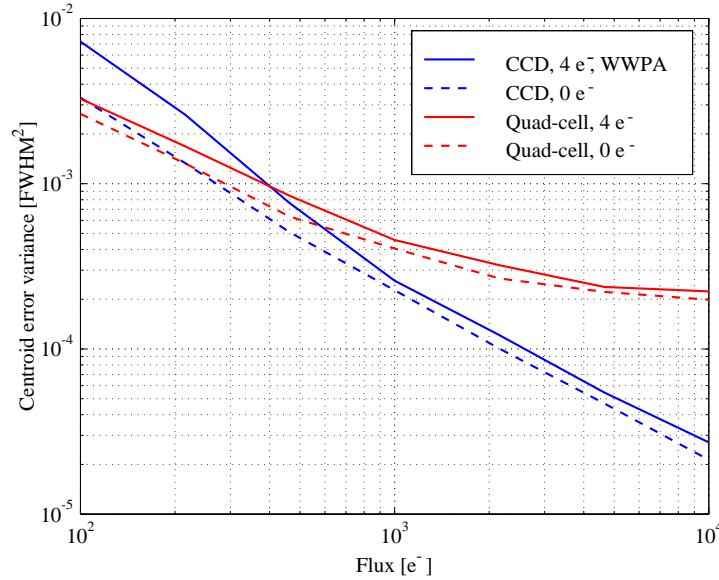


Figure 3.19: Centroid estimation accuracy for Quad-cells and CCD arrays with different levels of read-out noise. For the CCD with $4 e^-$ read-out noise the optimum algorithm was used to give a fair comparison.

Surprisingly, quad-cells have an advantage only at the lowest flux levels where centroid estimation is inaccurate anyway. They produce significantly worse estimates at higher flux levels, and thus do not seem to be well-suited for high-performance wavefront sensing. The zero read-out noise quad-cell (e.g. APDs) performs about 25 % better than the $4 e^-$ read-out noise one (e.g. high performance CCDs). At all flux levels, the best centroid estimations are delivered by a very low read-out noise detector array.

3.3 Model of ALFA's Shack-Hartmann sensor

Many works analyzed the topic of wavefront reconstruction by an SHS analytically (Fried, 1977, Hudgins, 1977, Rigaut and Gendron, 1992, Rigaut et al., 1998b) or numerically by use of specific array geometries together with specific modal basis sets (Herrmann, 1981, Lane and Tallon, 1992, Takato et al., 1994, Dai, 1996). The analytical works mostly provide asymptotically valid expressions, e.g. for modal error propagation or aliasing noise. The numerical works instead derive exact numbers, but the results hold for the assumed geometries only. In order to investigate modal cross-talk and aliasing with the aim to study and improve a specific system, a model of just this system is required.

Realization. As already mentioned, an SHS subdivides the wavefront in the telescope aperture and images each subaperture individually. Therefore, the starting point for modelling a specific SHS is a parametrization of the used microlens geometries. The hexagonal arrays (cf. Figure 3.5) available for ALFA as well as the keystone arrays (cf. Figure 3.6) were modeled. Using this parametrization, any incident wavefront can be cut into parts corresponding to areas covered by the respective microlenses. Given a phase-screen $\phi_j(x, y)$ over the j th subaperture, the imaging process creating a spot is given by (Roggemann and Welsh, 1996):

$$PSF_j = |FT\{\exp i\phi_j\}|^2, \quad (3.10)$$

where $FT\{\dots\}$ denotes the Fourier transform.

The computer model determines the interaction matrix of a specific modal basis exactly like ALFA, whose calibration scheme was described in Section 3.1.3. Reference positions are determined by imaging a flat wavefront through the microlenses, and subsequently calculating the centroids using a WPA algorithm. This procedure is repeated for all individual modes of a given basis set, and the corresponding spot positions are recorded. Finally, the reference positions are subtracted from the spot positions to obtain the wavefront slopes. Arranged in columns, the slope vectors provide the interaction matrix \mathbf{D}_∞ (cf. Equation 2.41) of some microlens array / modal basis set combination.

Accuracy of the model. Throughout this work, interaction matrices were calculated for a fixed number of 275 natural modes. This is a reasonable approximation of infinity, because Kolmogorov wavefronts are well modeled by 275 modes leaving only about two parts per thousand of the total wavefront variance unconsidered (cf. Equation 2.18).

To test the relevance of the computer model in reproducing measurements of ALFA's SHS, the correlation between measured and simulated slope vectors of low order modes was calculated. Hereby, the accuracy of the model depends on a number of factors like orientation of microlens arrays with respect to modal basis sets, their size with respect to the aperture, and the ability of the DM to reproduce low order K-L or Zernike modes. As discussed in Section 3.1.2, ALFA's DM should properly reproduce polynomials up to 6th radial order corresponding to about 30 low order modes.

The correlation coefficients in Table 3.1 generally show a better agreement between the computer model and the measurements for K-L functions. This may be due to the fact, that K-L functions vary more smoothly over the aperture, while Zernike polynomials tend to vary the most over a small region at the aperture edges (see Figure 2.3) making them harder to be reproduced by the DM. The coefficients further show a better agreement for arrays with a large number of subapertures than for the low number ones. However, the fair modelling of high order modes with the D3 array is not critical, since with this array only a low number of modes is reconstructed anyway.

Table 3.1: Correlation between the simulated and the measured slope vectors for the hexagonal arrays, different modal basis sets, and different numbering degrees.

j	D7			D5			D3		
	1-10	11-20	21-30	1-10	11-20	21-30	1-10	11-20	21-30
Karhunen-Loève	0.97	0.97	0.96	0.97	0.96	0.92	0.97	0.91	0.5
Zernike	0.98	0.96	0.84	0.97	0.92	0.87	0.96	0.59	0.65

3.4 Modal basis sets by comparison

As described in Section 2.1, low order K-L modes are the strongest in the decomposition of atmospheric turbulence and hence the desired basis set for a modal control AO system. The analysis in this section will show that there are additional advantages involved with the use of K-L functions that are due to their interaction with ALFA's microlens arrays.

Section 3.4.1 presents the cross-talk properties of the K-L and Zernike basis sets with the different microlens geometries. In general, K-L functions exhibit less cross-talk and a larger number of natural modes can be used.

Section 3.4.2 discusses the total reconstruction error of uncompensated Kolmogorov wavefronts for the all microlens arrays (hexagonal and keystone). For each array, an optimum number of control modes is found when the LS estimator is used.

Section 3.4.3 shows how the total reconstruction error at the optimum number of modes is distributed among the individual modes. This analysis reveals the modes that are only poorly reconstructed and offers an explanation for low order static aberrations commonly observed by AO systems in closed-loop.

3.4.1 Cross-talk

As already mentioned in Section 2.3.2, the SHS measurements of the modes are neither orthogonal nor normalized. If the slope vectors of a given set of modes to be reconstructed are close to linear dependency, the LS estimator may attempt to fit very small measurement vectors by a combination of large modal coefficients leading to strong noise amplification. This implies that the interaction matrix should be far from numerical singularity and thus should have a low condition number. Let me emphasize that using the MAP estimator bypasses this problem, because it will reject unlikely large combinations of modal coefficients.

Figure 3.20 shows the condition numbers of the interaction matrix \mathbf{D} as a function of the number of slope vectors included in \mathbf{D} for hexagonal and keystone microlens arrays. The computer model was used instead of measurements in order to study the behavior for up to 60 natural modes. When 30 low order modes are considered only, the DM is able to reproduce the modes, and the measured curves coincide nicely with the simulated curves.

Since the condition number of a matrix cannot decrease when columns are added, the individual curves are non-decreasing. In general, they rise smoothly up to a point where the slope vector of the next mode is close to a linear combination of the existing vectors. At this point the condition number jumps, and noise is heavily amplified by the LS reconstruction. Then, already a modest noise component such as the one produced by aliasing prevents the control loop from closing. In practice, we found that the condition numbers should not exceed a value of about 5 in order to assure loop stability.

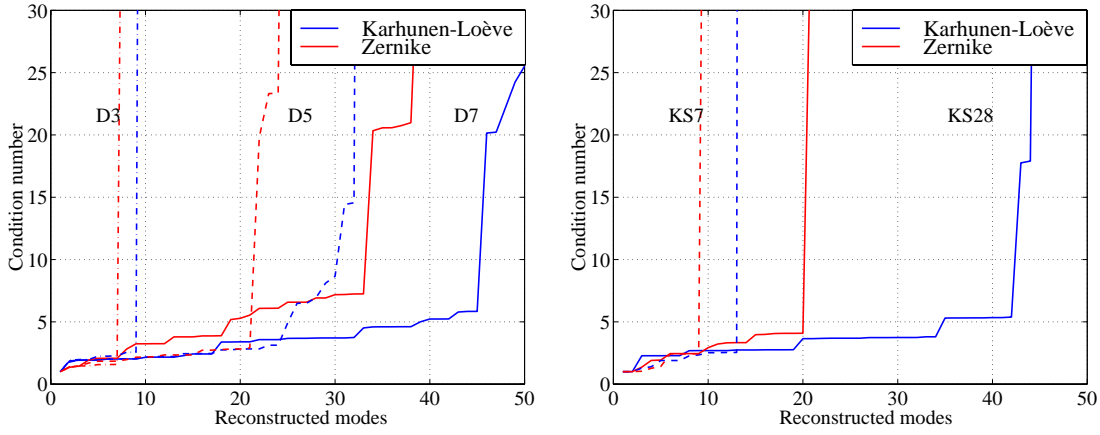


Figure 3.20: Condition numbers of interaction matrices as a function of the number of K-L or Zernike modes to be reconstructed. The labels denote the respective microlens arrays. *Left:* Hexagonal arrays, *Right:* Keystone arrays.

The left plot shows that using the K-L instead of the Zernike basis set increases the maximum number of usable modes by about 10 for the D7 array, by 4 for the D5 array, and by 2 for the D3 array. This is the major advantage of the K-L over the Zernike basis set.

The right plot shows the condition numbers for the keystone microlens arrays. Apparently, the keystone arrays are not well-suited to measure Zernike polynomials. With the 56 subapertures of the KS28, a maximum number of 20 Zernike polynomials can be reconstructed. Including mode 21 (3rd order spherical aberration) puts the interaction matrix close to numerical singularity. The behavior of K-L functions is more favorable; the KS28 allows us to reconstruct a similar number of modes as the D7, but has the advantage that its subapertures are slightly larger and equally illuminated. The KS7 allows us to reconstruct two more K-L modes than the D3, although both arrays have equally large subapertures.

3.4.2 Reconstruction error

The condition number analysis from the previous section gives an idea of how many modes can safely be reconstructed with the LS method depending on a given basis set / microlens array configuration. Reconstructing a larger number of modes increases noise amplification, but simultaneously provides more degrees of freedom for a better fit to a real wavefront. Hence, depending on the noise level, there will be an optimal number of modes providing a minimum LS reconstruction error.

In order to calculate this optimum number, the total reconstruction error has to be calculated as a function of the number of reconstructed modes m . For convenience, the formula for the total reconstruction error presented in Equation 2.51 is repeated here:

$$\sigma_{r,m}^2 = \sum_{j=1}^m \sigma_{r,j}^2 + \sum_{j=m+1}^{\infty} \langle a_j^2 \rangle,$$

where the second term describes the fitting error, and the first term describes the sum over the individual reconstruction errors as given by Equation 2.49 for the LS method and by Equation 2.61 for the MAP method:

$$\begin{aligned} \sigma_{r,j,ls}^2 &= (\mathbf{C}(\mathbf{a}_\perp \mathbf{a}_\perp^t) \mathbf{C}^t + \mathbf{D}^+ \langle \mathbf{m} \mathbf{m}^t \rangle \mathbf{D}^{+t})_{jj} \\ \sigma_{r,j,map}^2 &= (\mathbf{D}' \mathbf{V}_e^{-1} \mathbf{D} + \mathbf{V}_a^{-1})_{jj}^{-1}. \end{aligned}$$

The calculation of σ_t^2 requires knowledge of the modes' interaction with the respective microlens arrays, their turbulence statistics, and the measurement noise. The upcoming analysis considers the reconstruction of K-L and Zernike modes with all ALFA microlens geometries. The interaction matrices were derived from the computer model, and Kolmogorov turbulence statistics were assumed for the modal coefficients. Two different levels of measurement noise were considered: a very bright guide star with negligible measurement noise, and a faint guide star giving SNRs of 5 for the D3 and KS7 arrays, 3 for the D5 array, and 2.3 for the D7 and KS28 arrays. The SNR is defined as the ratio of the turbulent slope rms variation over the noise rms variation. The SNRs for the different arrays were derived from the sizes of the respective microlenses yielding different flux levels. The effect of partially illuminated subapertures for the D5 and D7 arrays was taken into account, too.

Hexagonal microlens arrays. Figure 3.21 displays σ_t^2 as a function of the number of reconstructed modes for the hexagonal microlens arrays. Four cases are shown: LS reconstruction of K-L and Zernike modal basis sets, MAP reconstruction of K-L modes, and the hypothetical case of a perfect reconstruction where σ_t^2 is given by the fitting error only. To calculate the variances under real observing conditions, the values have to be multiplied by the factor $(D/r_0)^{5/3}$. Under typical conditions $((D/r_0)^{5/3} \approx 20$ in the K-band for a 3.5-m telescope) the D3 array does not allow us to LS reconstruct the wavefront more accurately than $\sigma_{t,K}^2 = 1.06 \text{ rad}^2$ independent of the guide star brightness. Hence, even a perfect feedback loop could not provide compensated images with a Strehl ratio higher than 35 % (cf. Equation 2.19).

The K-L basis set provides a more accurate LS reconstruction of wavefronts than the Zernike basis set. This is primarily due to the lower cross-talk of low order K-L modes, and only secondarily to their larger contribution to the turbulence. The optimum numbers of modes N_{opt} and the corresponding reconstruction errors are given in the Table 3.2. For Zernike polynomials, N_{opt} is about half the number of slopes, whereas it is about 2/3 the number of slopes for K-L functions (Kasper et al., 2000a). It is interesting to see that, although the condition number analysis suggested that more than 40 K-L modes can safely be reconstructed with the D7 array, the optimum accuracy is obtained for 34 modes. In the faint guide star case, N_{opt} is slightly smaller, because noise dominates the measurement of high order modes. By simulating the interaction of Zernike polynomials with very high order SHS arrays, Takato et al. (1994) found that in this case N_{opt} is only about 1/8 of the number of slopes. Therefore, the better behavior of the K-L functions could provide a clear advantage for high order modal control SHS systems at large telescopes.

Table 3.2: Minimum LS reconstruction error variance (normalized) at an optimum number of modes to be reconstructed with the hexagonal arrays. The MAP reconstruction improves continuously with the number of modes, 30 modes resemble the number reproducible by ALFA's DM. The high SNR resembles the case of virtually zero measurement noise, whereas at low SNR measurement noise dominates.

	D7		D5		D3	
	K-L	Z	K-L	Z	K-L	Z
$\sigma_{t,min}^2(N_{opt})$, high SNR	0.013 (34)	0.02 (27)	0.022 (26)	0.033 (16)	0.053 (9)	0.056 (9)
$\sigma_{t,min}^2(N_{opt})$, low SNR	0.053 (28)	0.058 (20)	0.057 (17)	0.066 (14)	0.078 (9)	0.077 (9)
$\sigma_{t,map}^2$ (30 modes), high SNR	0.012		0.016		0.043	
$\sigma_{t,map}^2$ (30 modes), low SNR	0.046		0.047		0.067	

Figure 3.21 further demonstrates that the coarser microlens arrays are worse than higher order

ones even for faint guide stars. This behavior was expected and already discussed in Section 2.2.5. At very low light levels the performances become similar for all arrays since only the lowest order modes can be reconstructed anyway. However, the calculations were performed assuming photon-noise only. As noted by Takato et al. (1994), slipping into the detector noise limited case must be avoided, because then the measurement noise variance increases with the inverse squared of the flux. Therefore, coarse arrays may perform better at the lowest fluxes, where detector noise would dominate over photon noise for high order arrays.

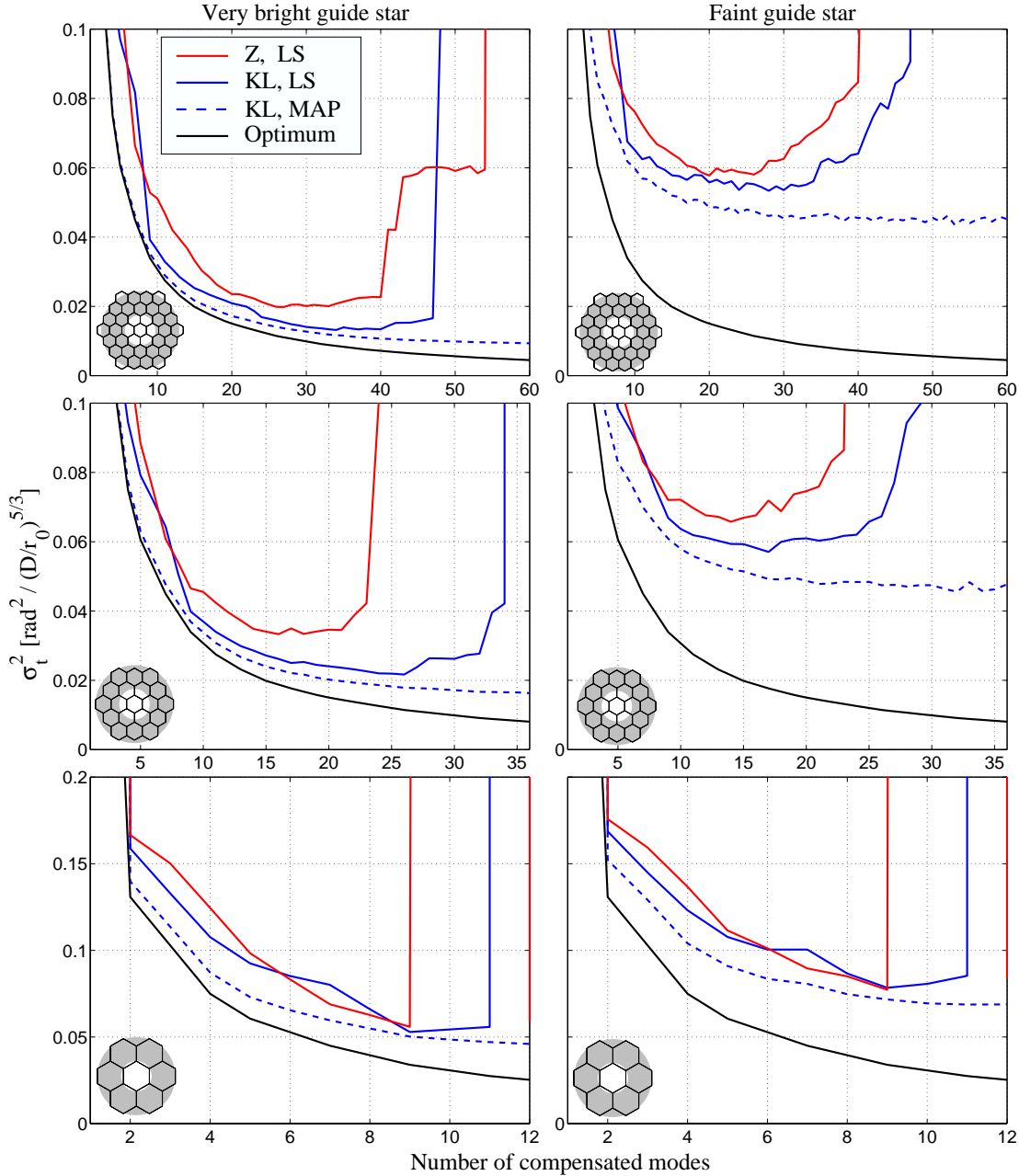


Figure 3.21: Total reconstruction error versus the number of reconstructed modes for the hexagonal microlens arrays. The left figures resemble the case of an infinitely bright guide star (aliasing noise only) and the right figures the case of a faint guide star where measurement noise dominates.

As mentioned in Section 2.3.6, the MAP can reconstruct more modes than available wavefront slopes. Further, $\sigma_{t,map}^2$ steadily decreases with the number of estimated modes, because $\sigma_{r,j,map}^2$ cannot exceed $\langle a_j^2 \rangle$ (cf. Equation 2.61). The numbers in Table 3.2 represent the case of 30 reconstructed K-L modes with all arrays, which is the number of K-L modes that ALFA's DM is able to reproduce properly. The absolute limits of $\sigma_{t,map}^2$ for 275 reconstructed K-L modes are 0.0081 rad^2 for the D7 array, 0.0143 rad^2 for the D5 array, and 0.0424 rad^2 for the D3 array. In comparison, a perfect reconstruction of 275 modes would leave a residual error of about 0.002 rad^2 .

Keystone microlens arrays Figure 3.22 displays σ_t^2 for the keystone microlens arrays. In comparison to the results in Figure 3.21, the bad performance in combination with Zernike polynomials is striking. This behavior is due to the spherical symmetry of the keystone arrays leading to a heavy cross-talk of Zernike spherical aberrations (when observed with the KS28, the 3rd order spherical ($j = 21$) looks similar to a combination of focus and a 2nd order spherical).

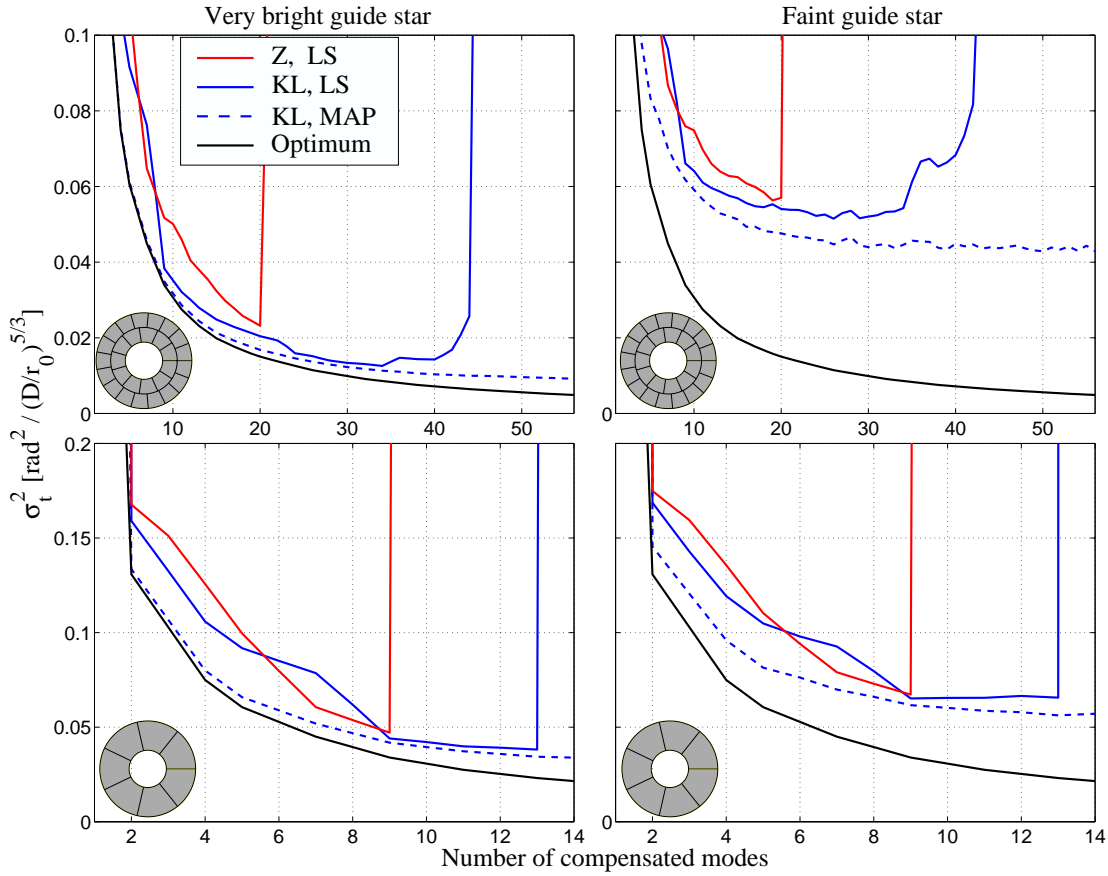


Figure 3.22: Total reconstruction error versus the number of reconstructed modes for the keystone microlens arrays. The left figures represents the case of an infinitely bright guide star (aliasing noise only) and the right figures the case of a faint guide star where measurement noise dominates.

For K-L functions, the keystone arrays perform better than the hexagonal ones. Although possessing two subapertures less than the D7 array, the KS28 allows us to reconstruct the same number of modes as the D7. Simultaneously, its subapertures are bigger by about 10 % and equally illuminated. As can be seen from Table 3.3, this results in a similar bright guide star performance of both arrays. Using a faint guide star, LS and MAP reconstruction errors are about 8 % smaller for the KS28.

The KS7 is significantly better than the D3 array, since it provides one more subaperture while the subaperture area remains the same. It allows us to reconstruct 13 K-L modes instead of 9 with the D3. In terms of reconstruction error variance, the KS7 is superior to the D3 by 40 % and 20% for bright and faint guide star LS reconstruction respectively, and by 25 % for the faint guide star MAP reconstruction. Therefore, the KS7 is well-suited for very low flux wavefront sensing where detector noise would dominate the measurement by higher order arrays.

Table 3.3: Minimum LS reconstruction error variance for the optimum number of modes to be reconstructed with the keystone arrays. The MAP reconstruction improves continuously with the number of modes, 30 modes resemble the number reproducible by ALFA's DM. The high *SNR* resembles the case of virtually zero measurement noise, whereas at low *SNR* the measurement noise dominates.

	KS28		KS7	
	K-L	Z	K-L	Z
$\sigma_{t,min}^2(N_{opt}), \text{high SNR}$	0.013 (34)	0.023 (20)	0.038 (13)	0.047 (9)
$\sigma_{t,min}^2(N_{opt}), \text{low SNR}$	0.049 (29)	0.056 (20)	0.065 (9)	0.067 (9)
$\sigma_{t,map}^2(30 \text{ modes}), \text{high SNR}$	0.012		0.032	
$\sigma_{t,map}^2(30 \text{ modes}), \text{low SNR}$	0.043		0.053	

3.4.3 Remaining error

ALFA commonly exhibits a trefoil shaped PSF when K-L modes are corrected, whereas it usually suffers from spherical aberration when Zernike modes are used. It is thus interesting to see how σ_t^2 is distributed amongst the individual modes depending on the microlens geometry, i.e. to identify poorly reconstructed modes given by their respective reconstruction error $\sigma_{r,j}^2$.

The Figures 3.23 and 3.24 plot $\sigma_{r,j}^2$ against the mode number j for K-L modes together with hexagonal and keystone arrays respectively. LS as well as MAP estimates were calculated. Hereby, the LS reconstructs the optimum number of modes given in the previous section, while the MAP attempts to estimate all modes.

The remaining error looks rather uniform for all the arrays. The D7 array exhibits a slight weakness in measuring mode 16 which is a high order trefoil mode², the D5 array exhibits a stronger weakness in measuring higher order astigmatism (modes 14, 15, and 18), and the D3 array is not well-suited to measure the first trefoil (mode 7). Interestingly, the arrays also show a special affinity to certain modes such as mode 19 (D3 array), or mode 26 (KS7 array). It was already illustrated by Figure 2.12 that the D7 array does not sense Zernike spherical aberrations very well.

It is interesting to see that the performances of LS and MAP estimates are comparable for the lowest order modes, while the LS accuracy deteriorates for higher order modes and less flux. The general pattern, however, is the same, i.e. the MAP cannot overcome a mode's poor signal, but it does a good job in filtering noise effects.

For the faint guide star case, the reconstruction quality of the first couple of modes is similar for all microlens array geometries. This supports the result that low flux performance in the photon-limited case does not depend on the number of subapertures.

²Informally, the Zernike notation of the modes is also used for the corresponding K-L functions. This is not precisely correct, but justified by the similarity of low order K-L and Zernike functions.

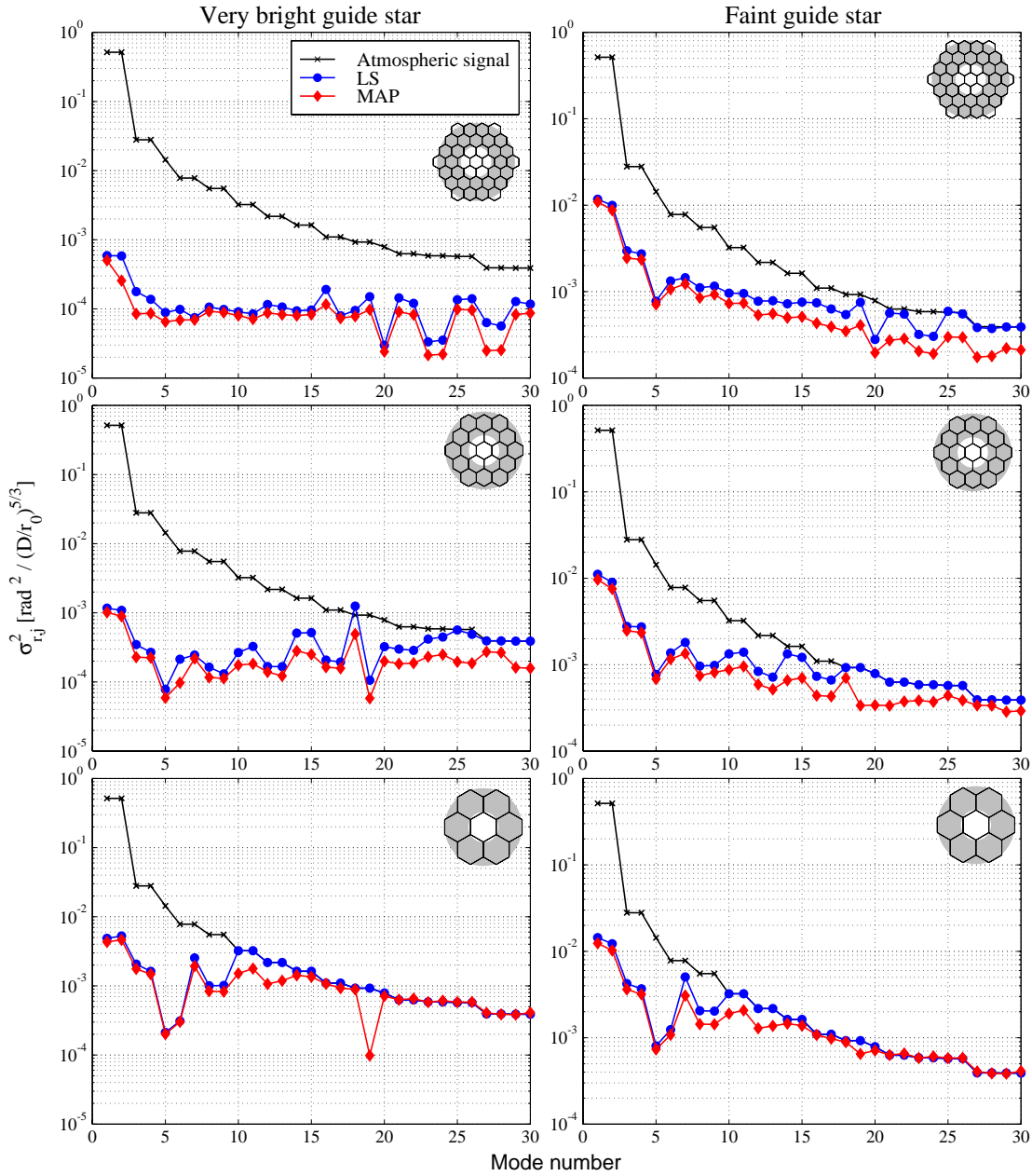


Figure 3.23: Remaining error of the individual low order K-L modes measured with hexagonal microlens arrays as a function of the mode number. The uncorrected variance is shown together with the results obtained from LS and MAP reconstruction at different noise levels.

3.5 Performance at the telescope

The analyses of the preceding sections led to a better understanding of ALFA's SHS, and hence to the ability to choose the various parameters in a more optimized fashion. This section will present results from experiments that were performed at ALFA and presented by Kasper et al. (2000a). The experiments were designed such that they are able to compare the performance of Zernike and K-L basis sets, sensor modes as well as LS, WLS, and MAP phase estimators.

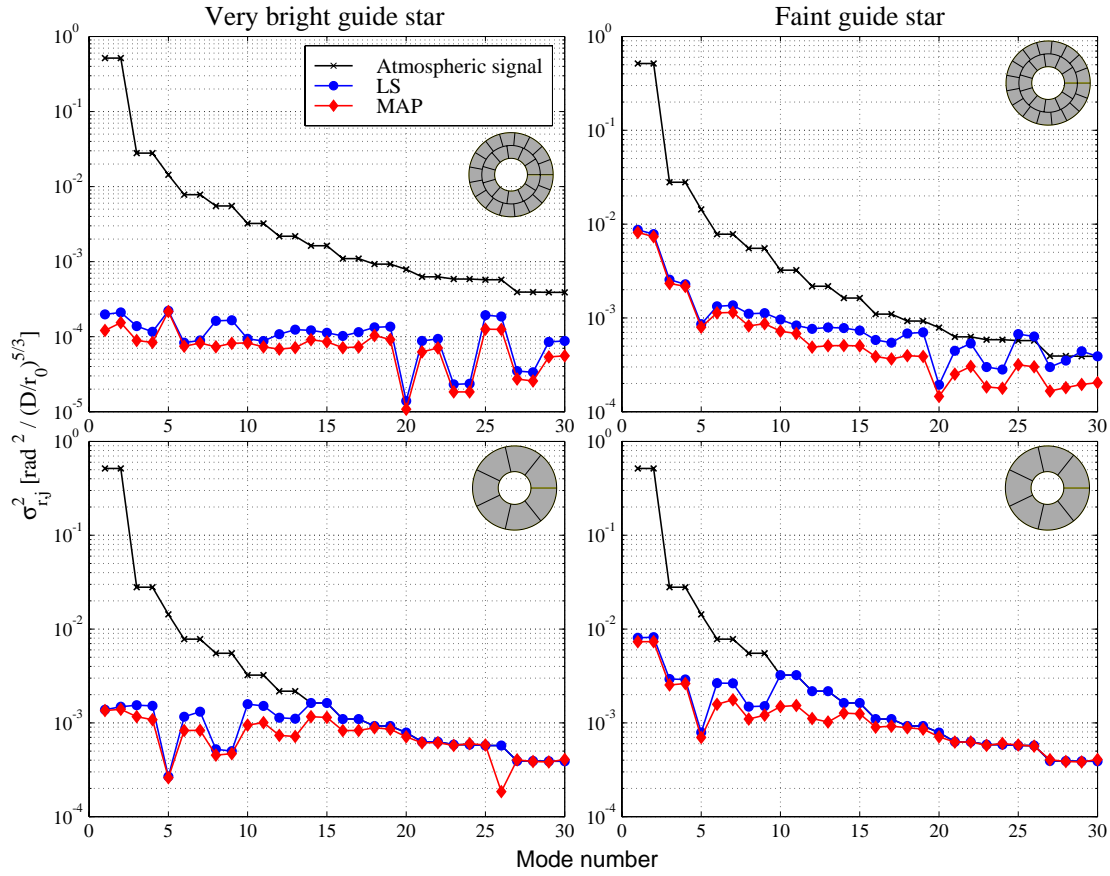


Figure 3.24: Remaining error of the individual low order K-L modes measured with keystone microlens arrays as a function of the mode number. The uncorrected variance is shown together with the results obtained from LS and MAP reconstruction at different noise levels.

3.5.1 Zernike polynomials and K-L functions

To compare the effect of a given modal basis set on the achieved image quality, the D7 array and a bright NGS (SAO 65213 with $m_V = 7$) were used. Beside the basis sets, the number of compensated modes was also varied. 9, 14, 20, and 32 modes were compensated respectively, corresponding to Zernike polynomials up to 3rd, 4th, and 5th order as well as the optimum number of modes to be corrected with the LS method. During the experiment, ALFA was operated at a framerate of 300 Hz.

Sets of 10 K-band exposures of 2 seconds each were taken with Omega-Cass for the different configurations, and Strehl ratios were computed for the plain sums of these images resembling long exposed images of 20 seconds. In order to eliminate effects that could be introduced by variable observing conditions, 96 of these sets were taken in a completely randomized order. The 96 image sets consist of 12 sets for each of the 8 configurations of 2 basis sets and 4 different numbers of compensated modes. Additionally, 12 sets with tip-tilt only correction were recorded intermittently to monitor the observing conditions.

Throughout the experiment, the tip-tilt corrected seeing was very stable at a superb value of about $0.25''$ in the K-band. The left image in Figure 3.25 shows the plain sum of a tip-tilt corrected image set, i.e. a long exposure. The PSF nicely shows the coma aberration that is produced by the telescope. This aberration actually imposes a strong limitation on ALFA's performance by severely limiting the dynamic range of the DM actuators. The right image shows the best long exposure that was obtained

with 32 K-L modes corrected. If the Strehl ratio of 72 % is normalized to the Strehl that was obtained on the reference fiber the value increases to outstanding 85 %.

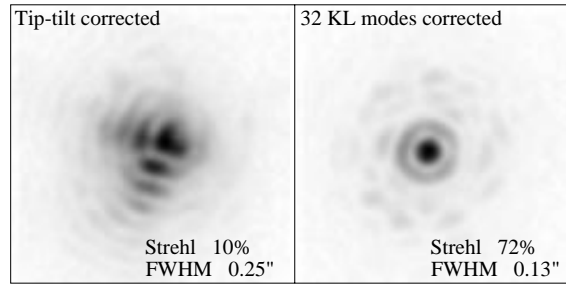


Figure 3.25: Long-exposure PSFs from June 24, 2000. *Left:* tip-tilt only correction. *Right:* 32 K-L modes corrected.

Results. The effects of the chosen basis set were compared using both a pooled-t and a paired-t analysis for each set of modes individually. A paired analysis for all mode sets was also performed. As Table 3.4 shows, the K-L basis performed better than the Zernike basis, with confidence levels ranging from 66 % to 95 %. When all mode sets are combined, the confidence level rises to 97 %. It was verified that the normality assumption of the t-test is justified, so that the overall conclusion of the K-L basis set being superior (with a confidence level of 97 %) is valid.

Table 3.4: Significance levels at which the K-L basis set is superior to the Zernike basis set.

Number of modes	9	14	20	32	all
Pooled significance	0.79	0.66	0.73	0.95	*
Paired significance	0.81	0.66	0.73	0.95	0.97

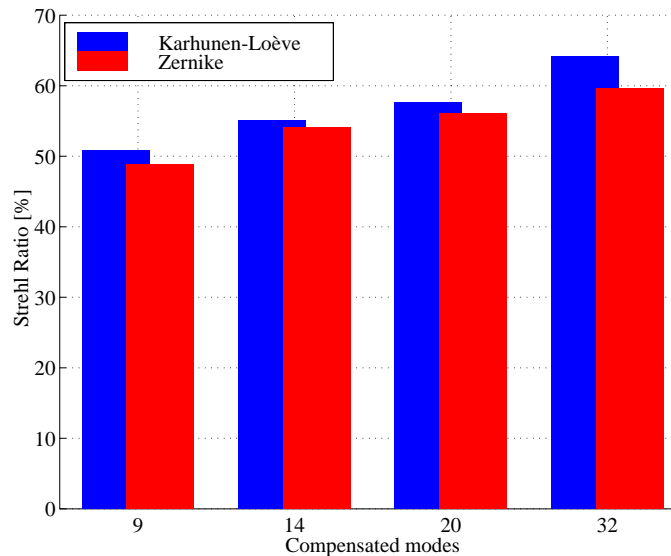


Figure 3.26: Mean values of the Strehl ratios that were measured for different modal basis sets and numbers of compensated modes.

A plot of the mean Strehl ratios versus the number of modes for both the Karhunen-Loève and Zernike bases is shown in Figure 3.26. The graph shows an apparent increase in performance as the number of modes increases for each basis. An Analysis of Variance (ANOVA) test was performed to confirm this for each basis independently, giving probabilities of about 10^{-5} for the K-L and $2 \cdot 10^{-3}$ for the Zernike basis set for the null hypothesis that the number of compensated modes does not affect the performance. As expected from observing the graph, the increase in Strehl ratio with the number of modes is confirmed with virtual certainty.

3.5.2 Sensor mode extension

The setup for this experiment was similar to that of the K-L / Zernike comparison. Again, SAO 65213 was used as a reference, and the framerate was 300 Hz. For each of the hexagonal arrays, the sensor mode extension was tested against a lower number of natural K-L modes. This lower number was not the optimum for the LS reconstruction as calculated in Section 3.4.2, because the calculation had not been done at this time. However, the chosen numbers are close to the optimum, and the test whether the concept of the sensor modes works is still valid. The actual compared configurations were:

9 K-L modes against 9 K-L + 3 sensor modes for the D3 array.

20 K-L modes against 20 K-L + 8 sensor modes for the D5 array.

32 K-L modes against 32 K-L + 20 sensor modes for the D7 array.

Sets of 12 K-band exposures of 2 seconds each were taken with Omega-Cass for the different configurations, and Strehl ratios were computed for the plain sums of these images resembling long exposures of 20 seconds. 24 of these sets were taken for each array, i.e. 72 sets altogether, in a completely randomized order. The 72 image sets consist of 12 sets for each of the 6 configurations of 2 basis sets and 3 different microlens arrays. The seeing was checked frequently and found to remain stable at a level of $0.7''$ in the K-band.

Results. The effects of the basis sets were compared using both a pooled-t and a paired-t analysis for each set of modes individually. A paired analysis for all mode sets was also performed. As Table 3.5 shows, the sensor mode extension produced higher Strehl ratios than the K-L basis alone, with confidence levels ranging from 79 % to 93 %. When all mode sets are combined, the confidence level rises to 98 %. It was verified that the normality assumption of the t-test is justified, so that the overall conclusion that the sensor mode extension works (with a confidence level of 98 %) is valid.

Table 3.5: Significance levels at which the sensor mode extension is superior to the K-L basis set alone.

Microlens array	D3	D5	D7	all
Pooled significance	0.88	0.84	0.79	★
Paired significance	0.86	0.84	0.93	0.98

A plot of the mean Strehl ratios versus the used microlens array for both basis sets is shown in Figure 3.27. Clearly, there is a strong argument for high order microlens arrays. The achieved performances also roughly agree with the reconstruction errors presented in Table 3.2. The sensor mode extension raises the obtained Strehl ratios by about 10 % over the values obtained with the K-L modes alone.

It should be noted, however, that using the sensor mode extension will be most effective for bright guide stars. Since the sensor modes are of a high spatial order, they have a low SNR (cf. Section 2.2.5) and hence will not considerably improve the faint guide star performance.

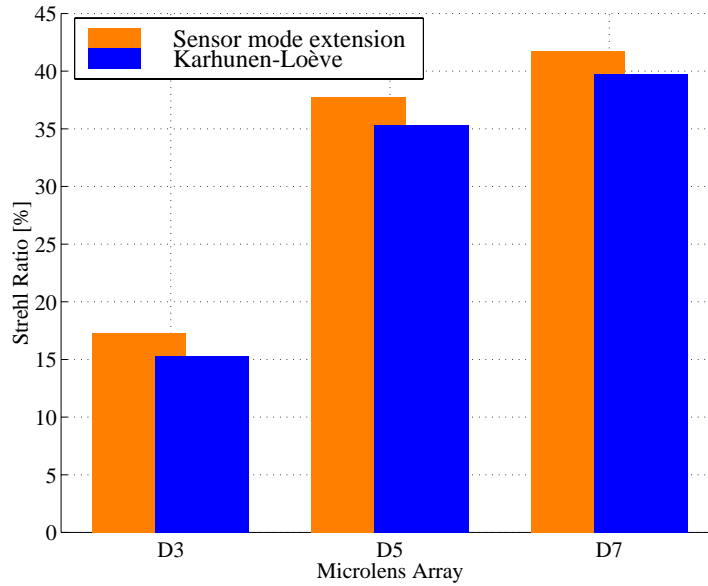


Figure 3.27: Mean values of the Strehl ratios that were measured for the different microlens arrays and the K-L basis set with and without sensor mode extension.

3.5.3 Phase estimators

LS, WLS, and MAP phase estimators were tested in closed-loop operation with ALFA. The covariance matrices required for the WLS and the MAP calculation and given by the Equations 2.57 and 2.64 were determined from SHS data taken just before the experiment. Again, the computer model was used to derive aliasing noise. The $m_V = 9.2$ star HD 75681 was used as a reference for the D7 array, and ALFA was operating at a framerate of 150 Hz.

Sets of 10 K-band exposures of 2 seconds each were taken with Omega-Cass for the different estimators, and Strehl ratios were computed for the plain sums of these images resembling long exposures of 20 seconds. In order to eliminate effects that could be introduced by variable observing conditions, 60 such sets were taken in a completely randomized order. The 60 image sets consist of 20 sets for each of the 3 estimators. The observing conditions were mediocre with a variable seeing between $0.7''$ and $1.1''$ in the K-band.

Results. The mean K-band Strehl ratios over 20 long exposures were 39 % for the LS, 42 % for the WLS, and 43 % for the MAP. The effects of the estimator were compared using multiple pooled t-tests. The MAP is superior to the LS estimator at a significance level of 95 %, but no significant difference was found between MAP and WLS.

Comparing the WLS and MAP reconstruction matrices afterwards, it turned out that they differed only marginally so a big difference in the performance could not be expected. The reason for this was a too high SNR resulting in a secondary impact of the a priori knowledge of the modal statistics. As shown by the simulated results in Figure 2.16, the MAP reveals its superiority at low SNR when compared to the WLS.

Variable observing conditions also limited the performance of WLS and MAP, because the statistics measured beforehand may not have been a good description throughout the experiment. The MAP will reveal its full strength once the statistics can be measured in closed-loop and can be used to update the estimator without interrupting AO operation.

3.6 Summary

The aim of this work was to improve the ALFA system by optimizing the wavefront reconstruction process. Optimized wavefront reconstruction is highly desirable, because a more sensitive sensor will work better with faint guide stars as well as the LGS. Hence, the sky-coverage is enlarged, and that is what it is all about in AO today.

Centroid estimation based on the WPA algorithm was investigated by computer simulations. Beside well-known modifications like pixel masking by setting a threshold (THRESH) or dynamically selecting small areas (CWPA), a new method was found which provided a gain of almost one magnitude in guide star brightness in the simulations. This WWPA method weighs the intensities with their SNR, i.e. it raises them to the 1.5th power. The exponent of 1.5 was further found to be optimum, because lower and higher values worsened the accuracy again. A test at the telescope confirmed the superior performance of this algorithm. Simulations further showed that another half of a magnitude could be gained by using a detector that possesses virtually zero read-out noise. The accuracy of quad-cell centroid estimation was found to be insufficient for very high performance wavefront sensing.

Modal reconstruction of wavefronts from SHS measurements depends on the geometric interaction between microlens array and modal basis set. The analysis for ALFA's microlens arrays showed that cross-talk limits the number of effectively usable K-L modes to about two thirds the number of measured slopes when the LS reconstructor is used. This number is even smaller for Zernike polynomials. This is the major advantage of K-L modes, easily outweighing their larger contribution in the decomposition of Kolmogorov wavefronts. Even for virtually zero measurement noise, i.e. very bright guide stars, strong noise amplification cannot be accepted because of the remaining aliasing noise. The optimum number of control modes and the corresponding reconstruction errors were calculated for all ALFA microlens arrays. This optimum number lies within the range that can be reproduced by the ALFA DM. Hence, the analyses which have been done without taking DM characteristics into account are valid. K-L modes, truncated at the optimum number, are a close to optimum basis set. However, both the DM and the SHS possess additional degrees of freedom which can be exploited for a further improvement of accuracy. Taking the DM characteristics into account, a method was developed to find mirror modes which could be used in addition to K-L modes without increasing cross-talk. Due to their high spatial order and resulting low SNR, these sensor modes will be most effective for bright guide stars. A test at the telescope confirmed that the method worked with increasing the Strehl ratios by about 10% compared to the K-L basis set only.

The LS reconstructor linearly propagates measurement noise. It will fit the noise regardless of the a priori probability of this fit. The MAP estimator considers these probabilities and is superior to the LS for each mode individually. The performance of the MAP estimator was simulated using a detailed model of ALFA's microlens arrays. The MAP clearly outperformed the LS reconstructor and was remarkably robust against imprecise prior knowledge. Since the MAP requires prior knowledge on modal covariance as well as noise variance, algorithms to determine these quantities were developed. It was shown that the MAP potentially increases the sensitivity by almost half a magnitude. The MAP could further be of great advantage for the feedback loop by avoiding unlikely large combinations of modes that may lead to loop crashes. A test at the telescope confirmed the superiority of the MAP over the LS estimate.

The analysis of the reconstruction error further indicated that the new keystone microlens arrays should be significantly superior to the hexagonal arrays by delivering a better performance for equally bright guide stars. The experimental verification of this expectation is scheduled for September, 2000 on Calar Alto.

Optimum wavefront reconstruction also requires some analysis tools which in turn can be used to convert ALFA from an “expert system” into a common user instrument. This involves the determination of quantities like r_0 , measurement noise, and the interaction of high order modes with the microlens array. In this way, a simple scheme was found that provides an estimate for the framerate which maximizes the system performance by balancing the effects of disturbance rejection and noise propagation. It is now possible to draw a guide how the various system parameters should be selected to achieve an optimum performance depending on observing conditions:

ALFA should always be calibrated with the WWP algorithm using 6 to 7 pixel-wide rectangles, and the optimum number of K-L modes. The optimum framerate should be set as obtained from the analysis of the measured PSDs. For bright guide stars, the LS reconstructor does a good job, and the sensor mode extension provides a further improvement for very bright stars. For faint guide stars, the MAP reconstructor should be used, because it prevents noise propagation onto high-order, low-SNR modes. The drawback of the MAP reconstructor is the additional effort for the calculation of the required statistics. The microlens geometry should not be changed to a coarser array until the measurement becomes limited by detector noise with the actual array due to a very faint guide star.

Outlook. Currently, the limiting magnitude of ALFA under the best observing conditions (V-band seeing below $0.7''$, good transparency) is around $m_V = 14$ for an NGS. However, a few possibilities remain to push this limit further towards fainter guide stars. Of course, a zero read-out noise detector would help, but it is doubtful whether such a detector will be available and affordable in the near future. Another hardware upgrade concerns the infrared/visible beam-splitter which will soon be replaced by a better transmitting one. This new beam-splitter promises about 20% more photons in the sensitivity range of the LLCCD. Together with algorithms that update the MAP reconstructor in closed-loop, and the adaptation of the compensation transfer functions to the dynamics of the individual control modes, an effective usage of ALFA on $m_V = 14$ guide stars can be achieved. The limiting magnitude under superb conditions should approach $m_V = 15$.

From my point of view, the reason for the modest results with the LGS are a combination of effects which prevent the wavefronts from being reconstructed accurately. First, the LGS requires both excellent seeing and transparency of the atmosphere. If the transparency is bad, the LGS will suffer twice from it and thus will be very faint. Even in the rare cases where the transparency allows of a relatively bright LGS, its inherent size in mediocre seeing conditions prevents the wavefront from being reconstructed accurately. A second drawback was the LGS jitter. The ALFA software requires pre-defined areas on the LLCCD to measure spot locations. The collective motion of the LGS spots in the SHS was such that they eventually touched the borders of the pre-defined areas spoiling the centroid measurements. It turned out, however, that a large part of the beam-jitter was produced close to the laser lab, and the problem was solved recently by a beam stabilization loop.

Another interesting possibility to get access to a larger number of astronomical objects is the installation of a Pyramid Wavefront Sensor working in the infrared. The analysis of the infrared SHS for NAOS³ predicts a limiting magnitude of the order of 10. The PWS should outperform the SHS in the infrared, because it heavily benefits from the smaller PSF in closed-loop. Therefore, an infrared PWS would allow ALFA to observe a large number of objects which are out of reach of other AO systems in the northern hemisphere.

³<http://www-laog.obs.ujf-grenoble.fr/activites/hra/naos/>

Chapter 4

Adaptive Optics imaging and spectroscopy of T Tauri

4.1 T Tauri - an unusual Pre-Main-Sequence system

T Tauri is an unusually complex system of multiple (at least three) stellar components, jets, and outflows unlike that seen in any other T Tauri star.

The optically visible primary, T Tau N, is a K1 star (Cohen and Kuhi, 1979), which has a companion at $0.7''$ separation to the south found by speckle interferometry in the near-infrared (Dyck et al., 1982). This infrared companion, T Tau S, exhibits a spectral energy distribution (SED) suggestive of a very young embedded source with a luminosity of about $11 L_{\odot}$, making it somewhat more luminous than T Tau N (Koresko et al., 1997). It shows an apparent lack of cold dust emission, and underwent a brightness flare of 2 magnitudes at various wavelengths between 1989 and 1991 (Ghez et al., 1991, Kobayashi et al., 1994). Orbital motion studies using speckle interferometry (Ghez et al., 1995) and adaptive optics (Roddier et al., 2000) imply a minimum mass of $3.3 M_{\odot}$ for the whole T Tauri system. Recent speckle holography observations revealed a companion to T Tau S which is separated by only $0.05''$ (Koresko, 2000). Other authors found no sign of binarity, but an extended envelope around T Tau S (Roddier et al., 2000). At $\lambda = 6$ cm Ray et al. (1997) reported the presence of another source, possibly a star, located $0.3''$ south of T Tau N, while Nisenson et al. (1985) found a source $0.3''$ to the north of T Tau N using optical speckle interferometry. Neither of these two sources has been confirmed by other investigators.

The near environment of T Tauri is a source of surprisingly strong extended ro-vibrational emission of molecular hydrogen (Beckwith et al., 1978). Herbst et al. (1997) used Fabry-Perot images and long-slit echelle spectra to measure velocity distributions, excitation temperatures, and the mass of molecular hydrogen in the T Tauri environment. The spatial distribution of this excited H_2 (see Figure 4.1) shows two separate outflows which are perpendicular to each other in projection (van Langevelde et al., 1994, Herbst et al., 1996). Recent optical spectroscopy (Solf and Böhm, 1999) indicated that these outflows are in fact almost perpendicular with the north-south component originating from T Tau S at 79° inclination, and the east-west component originating from T Tau N corresponding to the HH 155 jet. This jet is seen at an inclination angle of 23° (Eisloffel and Mundt, 1998). T Tau S exhibits more 6 cm flux than T Tau N (Schwartz et al., 1986) due perhaps to free-free radiation produced in a collisionally excited wind. However, previous near-infrared (NIR) imaging spectroscopy (Herbst et al., 1996) showed the Bry line emission to be confined to the primary. Optical HST observations revealed an arc of reflection luminosity which opens $5''$ to the northwest and $2''$ to

the southwest of T Tau N (Stapelfeldt et al., 1998a). This arc corresponds closely in position to the H_2 distribution presented by Herbst et al. (1996). From millimeter observations, Hogerheijde et al. (1997) suggested a mass of $10^{-2} M_{\odot}$ for the circumstellar disk of T Tau N, and an upper limit of 10^{-3} to $10^{-5} M_{\odot}$ for the one of T Tau S. This work also presents a model in which T Tau S is intrinsically similar to T Tau N, but obscured by the outer parts of T Tau N's disk.

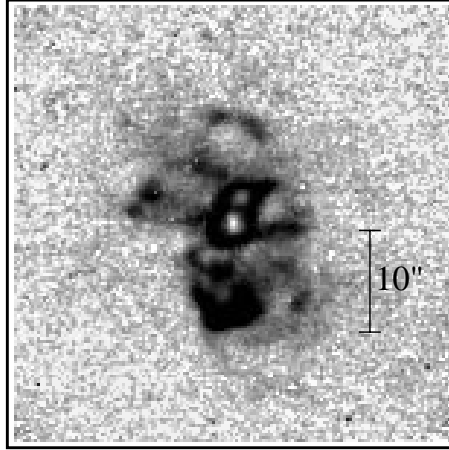


Figure 4.1: Continuum subtracted Fabry-Perot images of the environment of T Tauri. The H_2 emission from the unresolved stars in the middle and from T Tau NW a little more to the upper right is shown with an inverted grayscale to achieve a better dynamic range in the image. Courtesy T. M. Herbst.

Open questions. Even though T Tauri is a frequently observed object in astronomy, significant open questions remain.

First of all, the nature of T Tau S is not yet understood. Its near- to mid infrared SED can be fit by a proto-stellar model of a reddened ($A_V \approx 5$ mag) blackbody with a temperature of the order of 850 K and a radius greater than $100 R_{\odot}$ (Ghez et al., 1991, Herbst et al., 1997), while T Tau N is a more evolved star. The fragmentation model of star formation as well as observational results, however, favor a coeval formation scenario for the components of binary systems. Brandner and Zinnecker (1997), for example, found that all 8 binary pairs they could place on an H-R diagram appeared to be coeval within the observational errors. Also, the origin of the sporadic brightness flares of T Tau S remains unclear. The currently favored explanation for this behavior is that T Tau S undergoes *EXor* outbursts (Herbig, 1977) believed to occur in young stellar objects (YSOs) in the transition from FU Ori to T Tauri phase (see Figure 4.2). These outbursts are believed to be powered by accretion events, so observational evidence of active accretion during a flare of T Tau S should help to clarify the situation.

Secondly, the complex outflow phenomena in the T Tauri system are not yet understood in detail. H_2 emission is rarely observed in YSOs. Greene and Lada (1996) found it in only one of their 24 observed class II objects, but in 30 % of the younger class I objects, which generally have strong molecular outflows. Stellar outflows are believed to be powered by the gravitational energy released in the accretion process, and to be fed by disk material (Königl and Pudritz, 2000). In this context, it is surprising that the environment of T Tau S shows no signs of a massive dust component. Furthermore, the perpendicular outflows of T Tau N and T Tau S indicate misaligned disks, which also seem to be rare, at least for wide binaries with separations above 140 AU (Jensen et al., 2000, Monin et al., 1998).

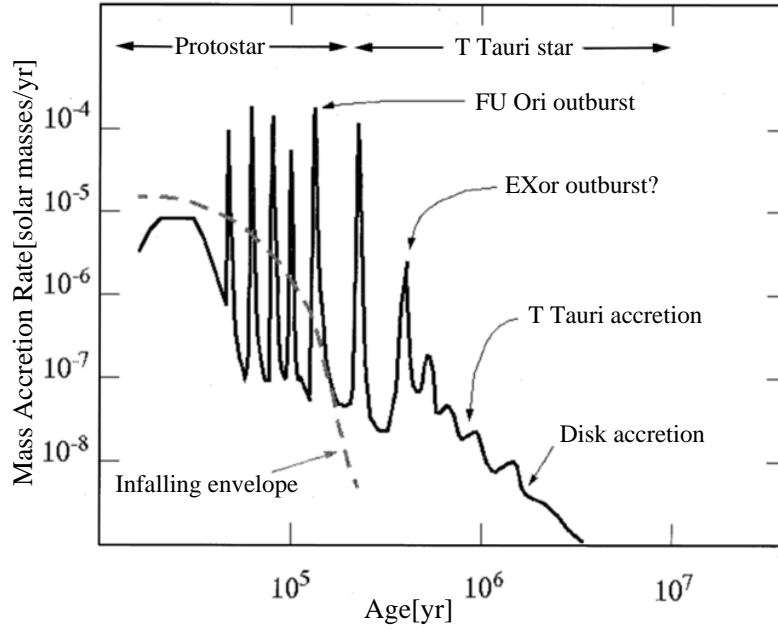


Figure 4.2: Outline of estimated mass accretion rates during the formation of a typical low-mass (solar-type) star. Taken from Hartmann (1998, p. 14).

4.2 Observations and data reduction

In the observing period 1999/2000, we observed T Tauri several times using different instruments (3D, Omega-Cass) and methods (imaging, spectroscopy, imaging spectroscopy, speckle interferometry) as summarized in Table 4.1. All observations benefited from closed-loop wavefront compensation by ALFA.

Table 4.1: Date and corresponding instrument for the T Tauri observations

Date	Instrument	Wavelength	Method
September 28, 1999	3D, small field	K-band	Imaging spectroscopy
October 1, 1999	3D, small field	H-band	Imaging spectroscopy
January 13, 2000	Omega-Cass	K-band	Long-slit spectroscopy
February 18, 2000	Omega-Cass	K-band	H ₂ narrow band imaging
February 20, 2000	Omega-Cass	H-band	Speckle interferometry
February 23, 2000	3D, large field	K-band	Imaging spectroscopy

4.2.1 ALFA + 3D: Imaging spectroscopy

3D is an integrate field spectrometer built at MPE, which can obtain H- and K-band spectra of an entire 16×16 pixel field at resolutions of $R \approx 1000$ to $R \approx 2100$ ($\lambda/\Delta\lambda$) (Weitzel et al., 1996). An aperture interchange module (Anders et al., 1998) allows it to be used with ALFA. Image scales of $0.07''$ and $0.25''$ per pixel give a small field of view (FOV) of about 1×1 arcseconds, and a larger field of 4×4 arcseconds respectively. Additionally, this module can switch between on-axis and off-axis

fields in order to observe sky frames without opening the AO loop. Due to the small FOVs, sky frames have to be observed off-axis, effectively doubling the observing time.

The small field centered on T Tauri was observed with 3D in H- and K-band in September / October 1999 with a spectral resolution of $R \approx 2100$ in H- and $R \approx 1100$ in K-band. ALFA used T Tau N ($m_V = 9.6$) as a natural guide star, delivering angular resolutions of $0.14''$ in H- and $0.16''$ in K-band. The total integration time was 14 and 18 minutes on source, respectively.

Standard techniques of sky subtraction and flat fielding removed background emission, dark current and pixel-to-pixel gain variations. The G2V star HD 26710 was observed at a similar airmass to T Tauri and served as a photometric, spectroscopic, and PSF reference. Division by the spectroscopic standard and subsequent multiplication by a solar spectrum of the same spectral resolution eliminated telluric features, photospheric features of the G2V star, and produced the correct spectral slope across the band. As depicted in Figure 4.3, the final data structure consists of an image cube with 600 separate wavelength slices, each containing spatial information over the FOV. The possibility of creating line maps in the NIR by subtracting the continuum contribution from a slice centered on a line is a unique feature of imaging spectrographs. The 3D line maps do not share the uncertainty of Fabry-Perot or narrow band images due to the PSF having changed between single exposures and thereby producing artifacts by the subtraction.

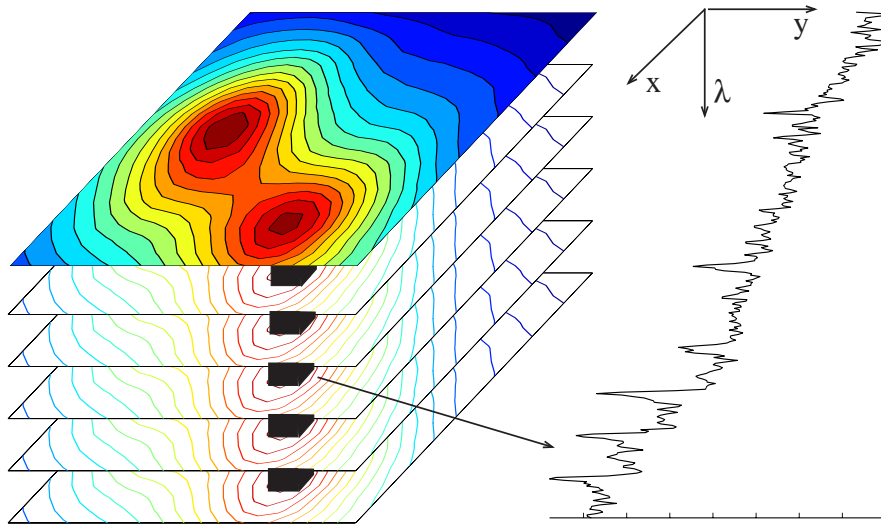


Figure 4.3: Image cube delivered by 3D. The extraction of a spectrum from a part of the image is illustrated.

In February 2000, the large field was observed in K-band with a poorer spatial resolution of about $0.4''$ FWHM. The spectral resolution was again $R \approx 1100$. The total integration time was 13 minutes on source. Data reduction followed the scheme described above, with the exception that the A1 star HD 27749 was observed as a spectroscopic reference this time. K-band spectra of early A-type stars are featureless except for the Br γ line in absorption. Hence, a simple interpolation over the Br γ line removes photospheric features reasonably well.

4.2.2 ALFA + Omega-Cass

The standard instrument for observing with ALFA is the Omega-Cass infrared camera (Lenzen et al., 1998). Three different re-imaging cameras provide image scales of $0.04''$, $0.08''$, and $0.12''$ per pixel.

With the Rockwell 1024 pixel HAWAII detector, the resulting FOVs are 40, 80, and 120 arcseconds square, respectively. The detector is sensitive in the NIR wavelength regime between $1\ \mu\text{m}$ and $2.5\ \mu\text{m}$, for which Omega-Cass provides a variety of broad- and narrow band filters. Long-slit spectroscopy is accomplished using a set of grisms and slit-widths of $50\ \mu\text{m}$, $100\ \mu\text{m}$, or $150\ \mu\text{m}$. This results in achievable spectral resolutions ranging from $R \approx 400$ to about $R \approx 10000$. The slits can be aligned to any position angle on the sky using a camera rotation device recently built at MPIA and available since August 2000.

Long slit spectroscopy

Similar to direct imaging, the benefits of using AO for spectroscopy are twofold. The increased spatial resolution allows one to obtain separated spectra of narrow structures like close binaries. The increased peak intensity of the PSF allows of narrower slit widths and thus higher spectral resolutions without sacrificing light. Hereby, small pixel-scales are useful to reduce background flux and therefore extend the integration-time before saturating the detector.

T Tauri was observed on January 13, 2000 with Grism 3 and the $100\ \mu\text{m}$ ($\approx 0.25''$) wide slit, giving a spectral resolution of $R \approx 1000$. For sky subtraction, the object was observed at two positions shifted along the slit by $10''$. The total integration time was 20 minutes. An A0 star was observed immediately afterwards in the same fashion to serve as a telluric standard star. Tip-tilt only correction was applied to this star, so its light is spread out over the angular size of the fully corrected T Tauri system ($\approx 1''$). This allows us to calibrate the spectral fringes (see Figure 4.4) whose pattern also depends on the object's position. These fringes appear in all the NIR spectra taken with Omega-Cass or 3D and are believed to originate from interference fringes in the coating of the science detector. The Br γ absorption feature of the reference star was removed from the spectrum by fitting a Gaussian profile to the line and subsequently dividing the spectrum by the normalized fit. This leaves the fringe pattern of the telluric calibrator intact such that the fringes of T Tauri mostly disappear when its spectrum is divided by the calibrator.

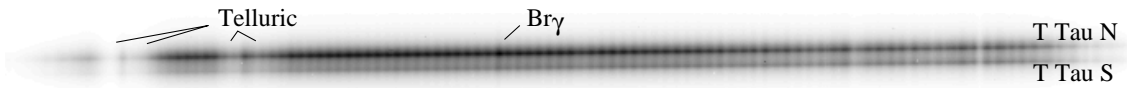


Figure 4.4: Sky subtracted K-band spectra of T Tauri obtained with Omega-Cass and ALFA. Although the spectrum is dominated by fringes and telluric absorption features, the Br γ line is readily visible for both stars.

A dome flat-field was created by obtaining a set of images of the illuminated dome with the same grism/slit configuration. The two-dimensional dome flat was then normalized. The sky-subtracted, two-dimensional data was divided by the dome flats to remove pixel-to-pixel sensitivity variations of the detector. The resulting two-dimensional spectra were collapsed along the cross-dispersed direction. Finally, the telluric standard spectrum is divided into the target spectrum in order to leave only spectral features of the target.

Speckle interferometry

Compensated speckle imaging is the term used to describe the application of speckle techniques to images partially compensated by an AO system. The advantage of using partially compensated wavefronts is that even a rather poor correction results in improvements of the SNR for the unbiased

speckle interferometry estimator, which is no longer bound to unity, and the phase spectrum estimate (Roggemann and Welsh, 1996, p. 278).

The observations were carried out on February 20, 2000 with a corrected image resolution of $0.12''$ FWHM. The uncorrected H-band seeing was about $1.2''$. The H-band was chosen to take advantage of the smaller diffraction limit compared to the K-band. Omega-Cass was operated in its “very high resolution” mode providing a pixel scale of $0.04''/\text{pixel}$. The subarray read-out mode was used to read subarrays of 256×256 pixels with the shortest integration time possible, i.e. 0.281 seconds. Some 1200 individual images were recorded.

The images were calibrated in the standard way by subtracting mean sky frames and replacing bad pixels by the median of adjacent good pixels. As an example, Figure 4.5 shows the sum of a set of 200 resulting images. Subframes 60 pixels wide and 20 pixels high around the primary and secondary star were extracted. The size of the subframes is a compromise between truncation of the PSF and contamination of the images of the secondary by light from the primary. The power spectrum analysis of Labeyrie (1970) was applied to the images, using the primary as a PSF reference for the secondary.

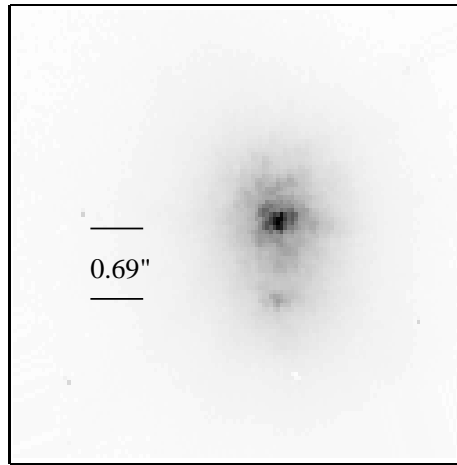


Figure 4.5: Sum of 200 images of the data after sky subtraction and bad pixel correction. A squareroot scaling was chosen to make T Tau S visible.

4.3 Results

4.3.1 H- and K-band photometry

Our observations, although not especially designed for this application, provide estimates of T Tau S’s brightness in the photometric bands H ($1.65\mu\text{m}$) and K ($2.2\mu\text{m}$).

The luminosities from October 1999 were derived from 3D observations. Because of the very small FOV, fluxes were estimated using apertures of only $0.25''$ in diameter. This works reasonably well, because the AO compensated PSFs of T Tau N, T Tau S, and the calibrator star are similar, yielding the same fractional amount of energy in the respective cores of the PSFs. The calibrator star was observed using identical loop parameters while its brightness on the SHS was dimmed by neutral density filters in order to match the one of T Tauri. This technique results in a similar SHS measurement noise for object and calibrator star, and hence a similar correction quality. The small contribution if T Tau N’s seeing-halo at the position of T Tau S was subtracted.

The luminosities from February 2000 were obtained from Omega-Cass observations. More specifically, the K-band fluxes were derived from H₂ narrow band imaging data, and the H-band fluxes from the speckle data. In both cases, fluxes were measured in a 3'' aperture around T Tauri and the speckle calibrator in H-band. The individual luminosities of T Tau N and T Tau S were then derived from the total flux by the relative intensities in their respective PSF cores.

Unfortunately, the exact H- and K-band luminosities of the calibrator stars HD 26710 (3D) and HD 27406 (speckle interferometry) are unknown. Hence, their luminosities were estimated from available V-band photometric data using the intrinsic colors of the actual spectral types as given by Cox (2000). No calibrator star was observed during the K-band observations in February 2000. In this case, the brightness of T Tau S was estimated from its difference to T Tau N, making use of the fact that the latter varies only slightly in brightness (Kobayashi et al., 1994). To generate Figure 4.6, this procedure was also applied to the relative magnitudes of T Tau N and T Tau S reported by Roddier et al. (2000).

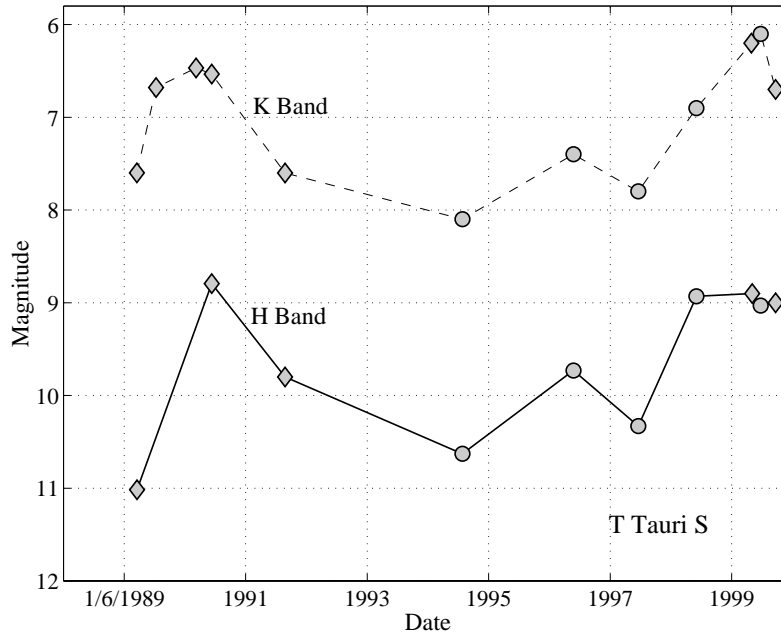


Figure 4.6: H- and K-band luminosities of T Tau S demonstrating its variable brightness. Datapoints from before 1999 were taken from the literature (Ghez et al., 1991, Kobayashi et al., 1994, Herbst et al., 1997, Roddier et al., 2000). The circle marks datapoints using T Tau N as a photometric reference.

Figure 4.6 draws H- and K-band light curves of T Tau S over the last 11 years. The luminosities do not represent accurate photometric data, but nevertheless illustrate the general behavior. Apparently, the 1990/1991 outburst was not a singular event as speculated by Kobayashi et al. (1994). After a quiet phase of about 7 years, another outburst started in late 1997 and lasts until today. It should be noted that the 1990/1991 outburst was nearly colorless between $1.6\mu\text{m}$ and $10\mu\text{m}$ (Ghez et al., 1991), while for the recent outburst no photometric data at wavelengths above $2.2\mu\text{m}$ is available.

4.3.2 The companion to T Tau S

In 1997, a close companion to T Tau S, hereafter called T Tau Sb, with a sky-projected separation of 53 mas, a position angle of 225 degree, and 10 % of the luminosity of T Tau S in the K-band was

found (Koresko, 2000). We took H-band speckle data (Köhler et al., 2000) with the aim to confirm the detection and possibly obtain orbital estimates.

Figure 4.7 shows the visibility of T Tau S being inconsistent with a point source. A multidimensional least-squares fit using the *amoeba* algorithm (Press et al., 1994) was used to compute the parameters of the binary. The reduction program tries to minimize the difference between the power spectrum computed from a model binary and the observational data by varying angular separation, position angle, and brightness ratio of the model. Fits to different subsets of the data yield an estimate for the uncertainty of the binary parameters.

The results of this procedure are the following parameters for T Tau Sb: position angle $(253 \pm 2)^\circ$, separation (79 ± 2) mas corresponding to 11 AU at a distance of 140 pc, and H-band flux ratio between T Tau Sb and T Tau S (0.259 ± 0.011) . Because of the symmetry of the power spectrum, there is an 180° ambiguity in the position angle, i.e. it can also be 73° . However, the phase of the Fourier-transformed image was computed using the Knox-Thompson algorithm (Knox and Thompson, 1974) giving indications that the angle is in fact 253° .

In the K-band, Koresko (2000) measured a flux ratio of 0.09 ± 0.02 between T Tau Sb and T Tau S. Assuming a mean H-K color of 2.5 magnitudes for T Tau S, an H-K color of 1.5 magnitudes is deduced for T Tau Sb, also highly reddened but bluer than T Tau S. This implies that a significant fraction of the extinction towards T Tau S arises in its immediate environment.

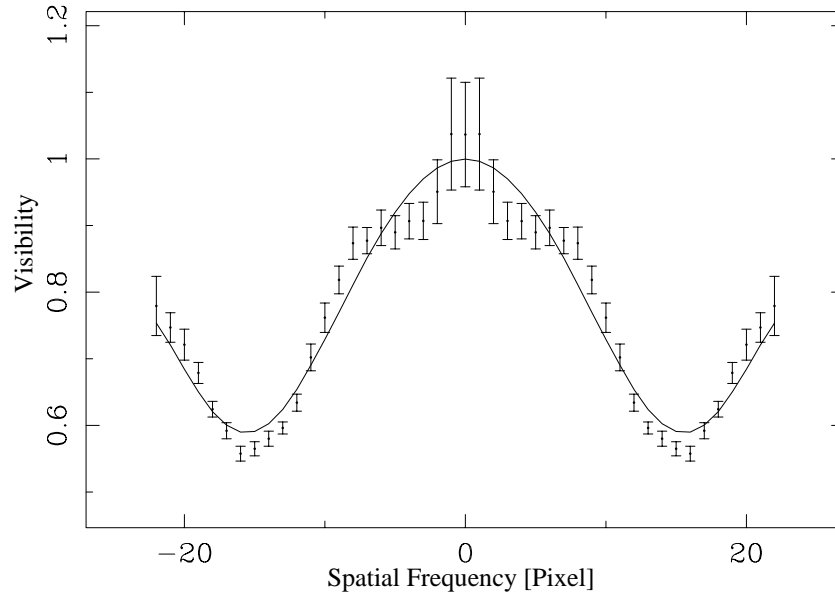


Figure 4.7: One-dimensional power spectrum of T Tau S, obtained by averaging the two-dimensional power spectrum along lines with position angle 163° . The points and error bars denote the measured data, the line represents the binary model.

Based on the data alone, there is no convincing evidence that T Tau S is indeed a binary, because the separation of $0.08''$ is below the H-band diffraction limit of the 3.5-m telescope of $0.1''$. Only the first minimum of the stripe pattern in the power spectrum is visible. Translated to “normal” space, this corresponds to two bright pixels next to each other without a darker pixel in between. It can certainly be stated, however, that the source is extended in one direction only. Perpendicular to this direction, it is point-like on the pixel scale of the camera, i.e. 40 mas. Combined with the results of Koresko (2000), the evidence for a binary star is convincing; its position angle has changed by 28° within 2 years.

Orbit. Together with the orbital position presented by Koresko (2000), we are able to give rough estimates for some orbital parameters of T Tau Sb. The law of cosines yields an estimate for the minimum linear velocity; a value of $v = 12.4 \pm 4$ km/s is derived at a distance of 140 pc.

Assuming a circular orbit, the centrifugal force equals the gravitational force, yielding the velocity

$$v = \sqrt{\frac{GM}{a}}, \quad (4.1)$$

where G is the gravitational constant, M the total mass in the system, and a the orbital radius. The minimum radius of the orbit is set by our observations to 11 AU. Adopting this value, we can derive a minimum mass of $1.9 M_{\odot}$ for the T Tau S system. In this case, the orbit's inclination would be of the order of -23° roughly in good agreement with an close to edge-on scenario, and its resulting orbital period would be of the order of 25 years.

The effect of a close binary on possible circumstellar and circumbinary disks is to produce gaps by tidal interaction (e.g. Lubow and Artymowicz, 2000). The circumpriary disk in such a system is typically truncated at radii between $0.35a$ and $0.5a$ depending on a number of factors including eccentricity, disk turbulent viscosity, and sound speed. In the case of T Tau S, this means that its disk should be truncated at a radius of the order of 5 AU, offering an explanation for the very low disk mass inferred by Hogerheijde et al. (1997). For comparison, typical sizes and masses of disks around T Tauri stars are of the order 100 AU and $10^{-2} M_{\odot}$ (Hartmann, 1998).

4.3.3 H- and K-band spectroscopy of T Tauri

Viewing the 3D data cube signal in a direction perpendicular to the image planes produces conventional spectra. The traces in Figures 4.9 and 4.10 represent the total flux in a $0.25''$ square box centered on T Tau N and T Tau S, respectively. The vertical bars and labels denote diagnostic lines in the actual photometric band. The detection limit for the lines can be estimated by measuring *equivalent widths* (EWs) on the noisy continuum. For the spectra in Figures 4.9 and 4.10, such EWs are smaller than 0.15 \AA suggesting that line features with EWs greater than about 0.4 \AA are safely detected.

Emission lines

Table 4.2 lists the line fluxes and uncertainties based on least-squares fits of Gaussian profiles to the lines drawn from the spectra. Figure 4.8 shows two representative Gaussian fits. The model includes five parameters: the central wavelength λ_0 , the amplitude A , the standard deviation of the Gaussian σ , and the two parameters m and b giving the slope and intercept of the linear baseline, respectively. With the model

$$F_{\lambda} = b + m\lambda + A \exp -\frac{(\lambda - \lambda_0)^2}{2\sigma^2}, \quad (4.2)$$

the flux is given by $\sqrt{2\pi}A\sigma$, and the formal uncertainty is $\Delta F_{\lambda} = [2\pi(\sigma^2\Delta A^2 + A^2\Delta\sigma^2)]^{1/2}$, where ΔA and $\Delta\sigma$ are those changes in the A and σ parameters which increase the χ^2 per degree of freedom by 1 (Bevington and Robinson, 1992). It should be noted that a number of additional systematic effects which are not included in the uncertainties have an impact on the line measurements. These effects are imperfect telluric calibration, uncertainty in the brightness of the photometric standard, and differences between the K magnitudes synthesized from a spectrum and the “true” filter values. The photospheric spectrum of a K1 star like T Tau N usually shows weak Brackett absorption lines (see e.g. Meyer et al., 1998, Wallace and Hinkle, 1997). Since the Brackett emitting regions cannot be spatially separated from the stellar photosphere, the emission lines may be diluted by the additional

absorption features to some extent. The spectral type of T Tau S is unknown, since its spectrum is featureless.

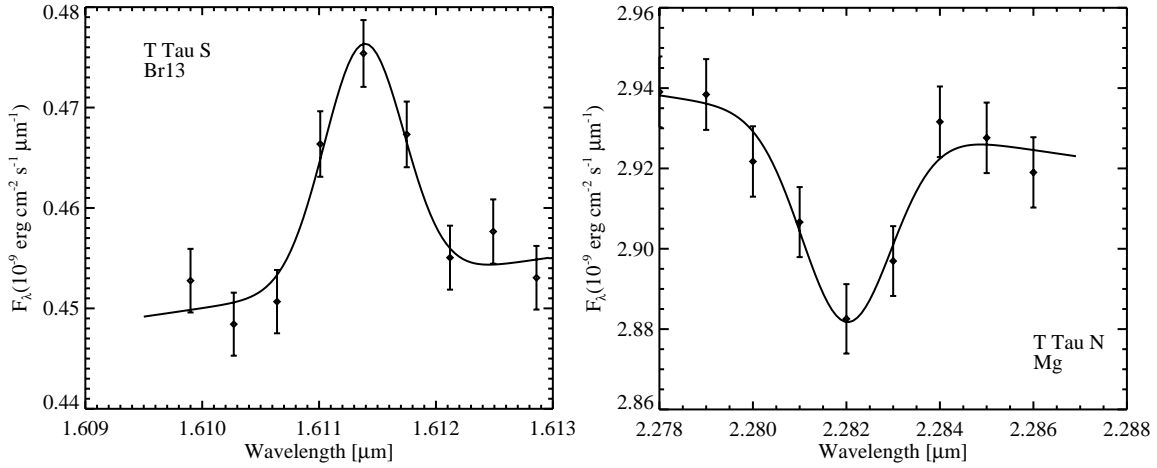


Figure 4.8: Representative line fits. See text for fitting procedure.

Table 4.2: H- and K-band emission line fluxes (in 10^{-13} erg cm $^{-2}$ s $^{-1}$) and EWs (in Å) of T Tau N and T Tau S.

Line	Wavelength [μ m]	T Tau N		T Tau S	
		Flux	EW	Flux	EW
HI					
Br13	1.611	2.2 ± 1.0	0.54 ± 0.25	0.2 ± 0.05	0.45 ± 0.1
Br12	1.641	0.6 ± 0.2	0.16 ± 0.03	0.2 ± 0.05	0.45 ± 0.1
Br11	1.681	0.9 ± 0.6	0.22 ± 0.16	0.4 ± 0.08	0.79 ± 0.16
Br10	1.736	3.0 ± 0.6	0.79 ± 0.15	0.8 ± 0.09	1.42 ± 0.16
Br γ	2.166	9.7 ± 1.1	3.2 ± 0.36	2.6 ± 0.29	2.16 ± 0.25
H $_2$					
1-0 S(1)	2.121

Molecular hydrogen. The H $_2$ lines are mainly important to analyze high temperature regions occurring behind shock waves in molecular clouds and the ion-neutral interface at the edge of ionized regions. H $_2$ line ratios can help to discriminate between the two mechanisms, and Herbst et al. (1996) found that H $_2$ emission in T Tauri is most likely produced by shock heating.

Our data does not show any H $_2$ emission within an arcsecond of the stars. This is a somewhat surprising result, because Herbst et al. (1996) found strong H $_2$ emission centered at the positions of T Tau N and T Tau S. Hence, also the emission line luminosity of T Tau S is variable, suggesting a common cause with its variable brightness.

Atomic hydrogen. Brackett emission lines of atomic hydrogen dominate the H- and K-band spectra of both, T Tau N and T Tau S. These transitions usually trace excitation conditions that involve either radiation energetic enough to ionize hydrogen, temperatures well above 5000 K in a relatively dense

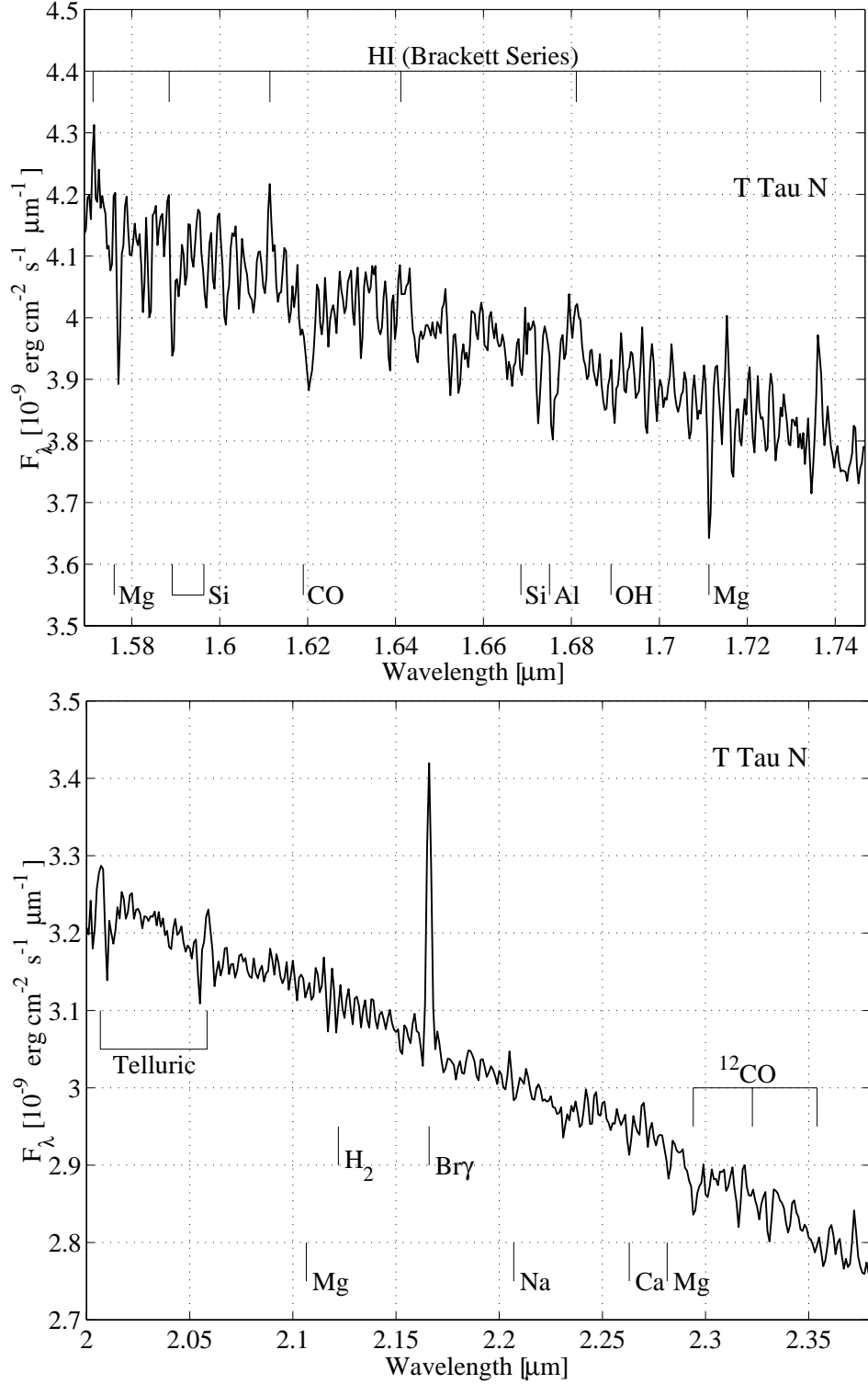


Figure 4.9: H- (*top*) and K-band (*bottom*) spectra of T Tau N obtained with 3D. See text for details of the synthetic apertures and photometric calibration.

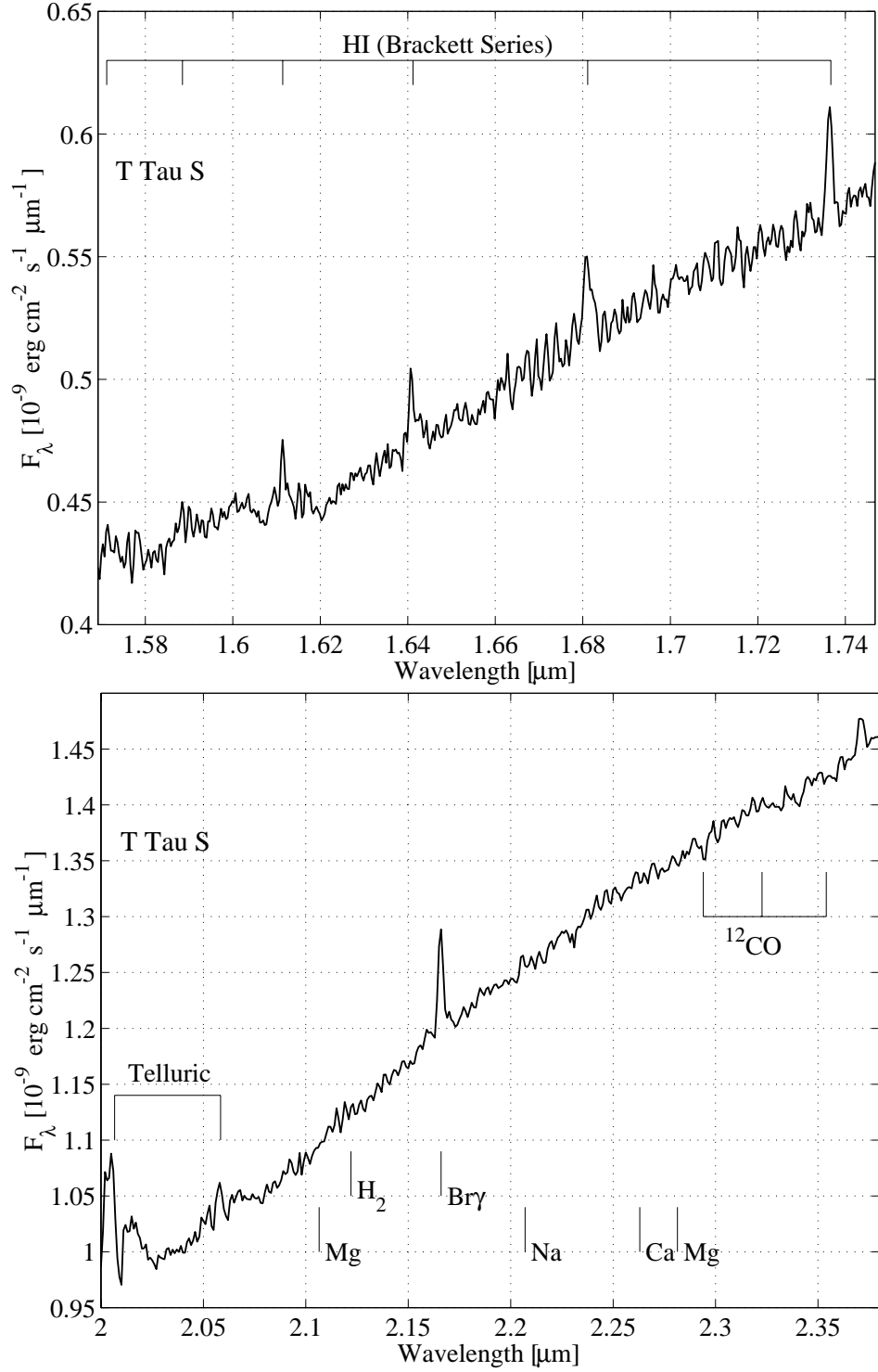


Figure 4.10: H-(*top*) and K-band (*bottom*) spectra of T Tau S obtained with 3D. See text for details of the synthetic apertures and photometric calibration.

medium to collisionally excite the hydrogen atoms, or a combination of both. In the current paradigm for T Tauri stars, energetic radiation is produced in the accretion process of circumstellar material onto the young star. This could either be emission from a boundary layer between the accretion disk and the star, or more likely emission produced in a shock when material falling in along magnetic field lines comes to rest on the stellar surface (Hartmann, 1998). The energetic radiation produced by the latter process is immediately re-absorbed and locally heats the stellar surface on one side and the accretion column on the other side (Calvet and Gullbring, 1998). The UV radiation from these hot areas, as well as from the star itself can then ionize collisionally excited hydrogen atoms in a high density neutral wind. By this process, Brackett recombination lines are produced within a few tenths of an AU around the star (Natta et al., 1988).

Outflow signatures are commonly observed in T Tauri stars, and they seem to be well correlated with infall and accretion (Hartigan et al., 1995). If the model of these outflows being powered by gravitational energy released in the accretion process is valid, then the hydrogen recombination lines indirectly result from accretion. Muzerolle et al. (1998) support this inference by finding a tight correlation between the Br γ and Pa β line luminosities, and the hot continuum excess in the U-band which is considered as a direct measure for accretion luminosity. A least-squares fit to the relation between accretion luminosity L_{acc} and Br γ line luminosity $L_{Br\gamma}$ yielded (Muzerolle et al., 1998):

$$\log(L_{acc}/L_{\odot}) = (1.26 \pm 0.19) \log(L_{Br\gamma}/L_{\odot}) + (4.43 \pm 0.79), \quad (4.3)$$

which is almost consistent with the ad-hoc assumption of $L_{Br\gamma}$ being directly proportional to L_{acc} .

Assuming a distance of 140 pc, absolute luminosities can be inferred from Table 4.2, yielding $L_{Br\gamma} = (6.16 \pm 0.69) 10^{22}$ W for T Tau S and $L_{Br\gamma} = (2.3 \pm 0.26) 10^{23}$ W for T Tau N. Equation 4.3 then gives accretion luminosities of $L_{acc} = (0.44 \pm 0.09) L_{\odot}$ for T Tau S and $L_{acc} = (2.33 \pm 1.04) L_{\odot}$ for T Tau N. The uncertainties of these values include our measurement errors as well as those in Equation 4.3.

L_{acc} can be related to the mass accretion rate \dot{M} via the virial theorem, if we assume the material is initially in an orbit around the star:

$$L_{acc} = \frac{1}{2} \frac{GM_* \dot{M}}{R_*}, \quad (4.4)$$

where G is the gravitational constant, M_* is the stellar mass, and R_* the stellar radius. If we adopt masses of $2 M_{\odot}$ for T Tau N and $1.5 M_{\odot}$ for T Tau S, and choose $3 R_{\odot}$ as a typical value for YSO radii in each case, Equation 4.4 gives accretion rates of $(2.2 \pm 0.98) 10^{-7} M_{\odot} \text{ yr}^{-1}$ for T Tau N and $(5.6 \pm 1.1) 10^{-8} M_{\odot} \text{ yr}^{-1}$ for T Tau S.

The Brackett line fluxes of the two stars and thus the derived accretion luminosities and rates are typical for T Tauri stars (Greene and Lada, 1996, Muzerolle et al., 1998). Observations from January 1995 (Herbst et al., 1996), which took place at a low brightness phase of T Tau S (see Figure 4.6), did not show any apparent Br γ emission from T Tau S. The simultaneous occurrence of a brightness flare and Br γ emission in our data suggests that the outbursts of T Tau S may be related to accretion events. The derived accretion luminosity of $0.44 L_{\odot}$, however, cannot explain a flaring of more than one magnitude for an object more luminous than $10 L_{\odot}$ like T Tau S.

Extinction. So far, the effects of extinction were not considered. The visible component, T Tau N, is obscured by $A_V = 1.39$ (Cohen and Kuhl, 1979). Using the reddening law of Mathis (1990), the light attenuation toward T Tau N is of the order of 12 % in the K-band and 18 % in the H-band. The extinction towards T Tau S is unknown. Ghez et al. (1991) derived $A_V \approx 4.5$ from the $10 \mu\text{m}$ silicate

absorption feature by modelling T Tau S as an 850 K blackbody with a radius of $100 R_{\odot}$. For other models like a central star + disk observed at some inclination angle, the depth of silicate absorption will depend on many parameters and may even appear in emission as it does in T Tau N (Ghez et al., 1991). From their HST data, Stapelfeldt et al. (1998a) were able to set a lower limit of $A_V = 7$. This corresponds to the extinction required to push even the reddest stars ($\log T_{\text{eff}} = 3.5$, K-band veiling $r_K = 4$) beneath their optical detection limit of $m_V = 19.7$ for T Tau S. Since the Br γ emission is believed to arise very close to the star, it should also be obscured by at least $A_V = 7$. The actual extinction, however, is unknown and so are the absolute line fluxes.

Brackett line ratios. It is remarkable that the line ratios between Br γ and Br10 are about 3.25 for both T Tau N and T Tau S, consistent with the value expected for optically thin emission from a gas in local thermal equilibrium and obeying Menzel Case B recombination (see Osterbrock, 1989, p. 84). However, measuring Br α and Br γ luminosities of YSOs, Evans et al. (1987) found that H I recombination line ratios are usually not consistent with Menzel Case B, and Natta et al. (1988) showed that relative intensities of lines created in a mostly neutral wind are a complex function of wind parameters and stellar radiation field.

The unknown intrinsic line-ratio between Br γ and Br10 does not allow us to calculate the extinction towards the emitting region directly from the data. Let us, however, suppose that the recombination line ratios of T Tau N and T Tau S are similar, without regarding the actual generation process. Since the extinction towards T Tau N is negligible in the NIR, an intrinsic $F_{\text{Br}\gamma}/F_{\text{Br}10} = 3.25$ is assumed. With the NIR extinction law from Mathis (1990), $A(\lambda) \propto \lambda^{-1.7}$, it follows that $A_{\text{Br}10}/A_{\text{Br}\gamma} = 1.46$. Hence, if x_0 denotes the intrinsic line ratio, the observed line ratio x as a function of extinction is given by

$$x = x_0 \cdot 10^{-0.4(A_{\text{Br}\gamma}(1-1.46))}. \quad (4.5)$$

If the extinction towards the recombination line emitting region around T Tau S would be as large as for the star itself ($A_V > 7$), an intrinsic $x_0 = 3.25$ would be observed as $x > 4.4$, barely compatible with our observations (see Table 4.2). The conclusion is that in contrast to the star, the Br γ emitting region around T Tau S does not seem to be heavily obscured. However, the wavelength baseline between Br γ and Br10 is rather small, so it would certainly be interesting to additionally observe the Pa β luminosity of T Tau S. In this case an intrinsic ratio of $F_{\text{Pa}\beta}/F_{\text{Br}\gamma} = 6.2 \pm 2.8$ (Greene and Lada, 1996) would be observed as a ratio greater than 17.6 for extinctions $A_V > 7$.

Brackett line maps. The neutral wind model predicts that most of the hydrogen recombination line flux is produced within a few tenths of an AU around the central star (Natta et al., 1988). The near environment of a star that accretes and produces outflows should be too dense to produce a fully ionized region of considerable size. For example, if we assume a typical density of 10^{11} H-atoms per cm^{-3} in the near environment of an YSO, even a strong high-energy radiation of one L_{\odot} could only ionize a region of a few hundredths of an AU.

Radio observations at 6 cm wavelength showed T Tau S as two circularly polarized lobes approximately 0.15 arcseconds apart and revealed a third source, T Tau R, located midway between T Tau S and T Tau N (Ray et al., 1997). The continuum subtracted line maps displayed in Figure 4.11 allow us to find extended Brackett emission when compared to images in the adjacent continuum. The spatial resolution of the H-band line maps from October 1999 is about $0.14''$ FWHM. Assuming that a change in FWHM of half a pixel is significant, structures larger than 10 AU are detectable as extended from our data. As expected, the Brackett series emitting regions around T Tau N and T Tau S are point-like, i.e. smaller than $0.07''$ on the sky. There is also no hint of a third source between T Tau N and T Tau S.

The apparent mismatch between the appearance of T Tauri in radio and Brackett line maps is a strong argument against fully ionized regions as the source of the respective emission.

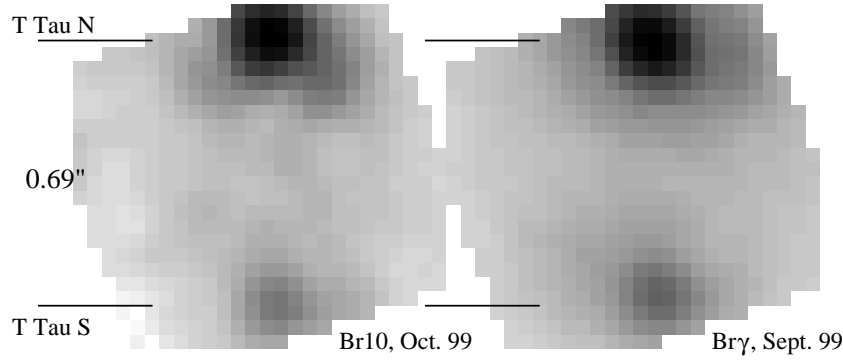


Figure 4.11: Continuum subtracted line maps of Br10 (left) and Br γ (right). The images display the whole FOV of 3D with the small image scale. A squareroot scaling was chosen to enhance the fainter T Tau S.

Figure 4.12 shows a Br γ line map obtained in February 2000. The observing conditions were obviously worse than in October 1999, but nonetheless the brightness ratio between T Tau N and T Tau S clearly has changed. The calculated EWs are $2.0 \pm 0.6 \text{ \AA}$ for T Tau N, and $2.3 \pm 0.6 \text{ \AA}$ for T Tau S. Compared to the results in Table 4.2, the Br γ emission from T Tau N has obviously faded, while it did not change for T Tau S.

There is a new source about $0.7''$ to the NW of T Tau N which is not detected in the continuum images and does not coincide with the H $_2$ source T Tau NW, located about $3''$ NW of T Tau N (Herbst et al., 1996). Unfortunately, it is at a position just outside the small field of 3D so it could not be detected in October 1999. Because of the complexity of the 3D instrument and its data reduction schemes, an artificial origin of this source cannot be precluded. On the other hand, there have been many reports (Schwartz et al., 1986, Nisenson et al., 1985, Ray et al., 1997) of additional sources in the vicinity of T Tau N, so follow-up observations are required to clarify the situation.

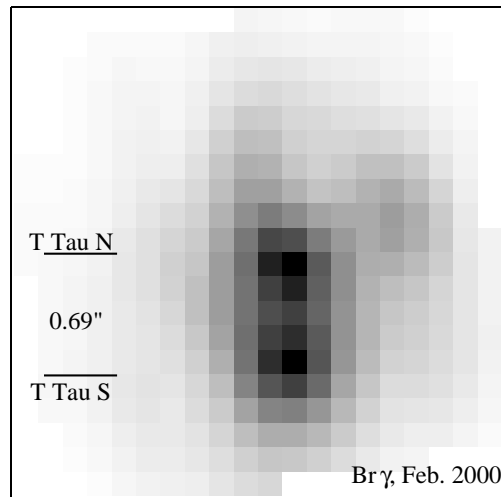


Figure 4.12: Continuum subtracted Br γ line map (linear scaling) of T Tauri obtained with 3D in February 2000 using the large image scale. Note the additional source to the NW of T Tau N.

Long slit spectra Figure 4.13 shows K-band spectra of T Tau N and T Tau S obtained with Omega-Cass in January 2000. The data were taken under mediocre seeing conditions, and therefore there is some blending of the emission from the two stars.

Additionally, the strong fringe pattern which appears in the uncalibrated data (see Figure 4.4) could not be removed entirely by the spectral calibration. However, the spectral region around the Br γ line could be reduced with satisfactory quality. The absolute flux density was not calibrated, because the telluric standard star was observed without AO correction, and therefore an unknown fraction of its flux was blocked by the slit.

The Br γ EWs derived from the reduced spectra were $3.4 \pm 0.6 \text{ \AA}$ for T Tau N and $2.7 \pm 0.6 \text{ \AA}$ for T Tau S, in good agreement with the 3D results given in Table 4.2. The spectrum of T Tau S shows a small enhancement at the position of the H $_2$ S(1) line at $2.12 \mu\text{m}$ which was not observed in September 1999. This enhancement, however, is not a clear detection when seen in relation to other non-real features in the spectrum such as the one at $2.135 \mu\text{m}$. If the H $_2$ emission is real, this would support the idea of a variable emission line spectrum of T Tau S.

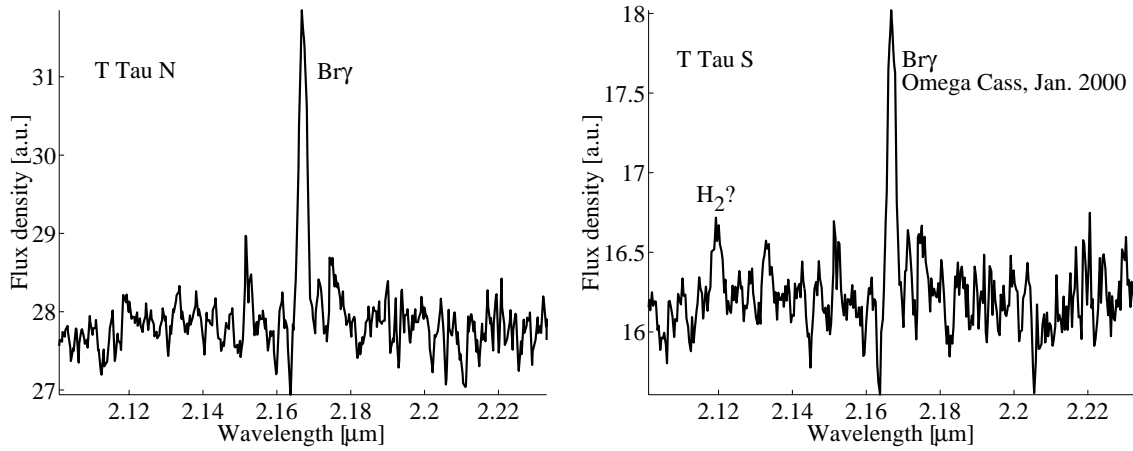


Figure 4.13: Reduced long-slit spectra showing the region around the Br γ line of T Tau N and T Tau S. Note that the spectral slope is not calibrated.

Absorption lines

Most NIR absorption lines appear as doublets or triplets (e.g. Mg, Al) to which Gaussian profiles cannot be fitted. In these cases the line shape was interpolated using a spline fitting routine, and the continuum was derived by a linear fit to the datapoints outside the line. The EWs were measured by summing up the ratio between line flux $F_{l,\lambda} - F_{c,\lambda}$ and continuum flux $F_{c,\lambda}$ multiplied by the stepsize $\Delta\lambda$, i.e.

$$EW = \Delta\lambda \sum_{\lambda} \frac{F_{l,\lambda} - F_{c,\lambda}}{F_{c,\lambda}}. \quad (4.6)$$

Uncertainties on the EWs were estimated from the minimum and maximum acceptable continuum values next to the lines (Duchêne et al., 1999). Figure 4.14 demonstrates the result of this procedure for the Al triplet at $1.674 \mu\text{m}$.

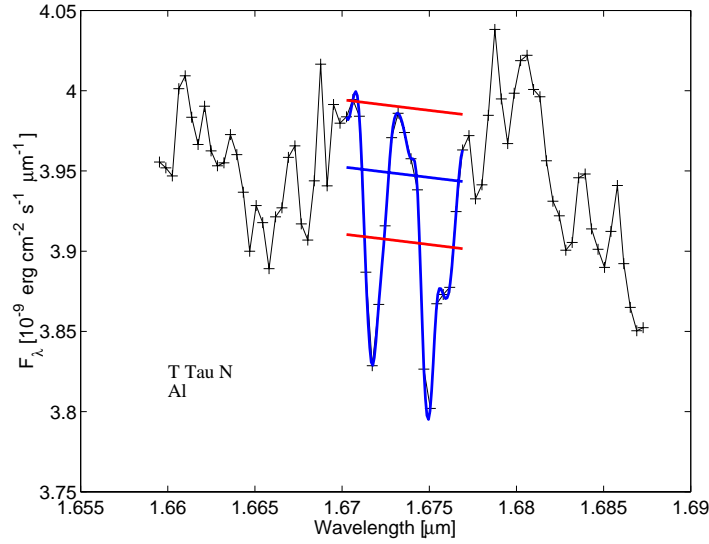


Figure 4.14: Representative spline fit to the Al-triplet at 1.6724, 1.6755, and 1.6768 μm via the Spline method. See text for fitting procedure and Table 4.3 for derived EWs.

T Tau N. Any excess emission r from warm circumstellar material, e.g. the inner part of a circumstellar disk, dilutes the EW of photospheric lines such that

$$EW' = \frac{EW}{1+r}, \quad (4.7)$$

is measured. r is called *veiling factor* and a subscript usually denotes the respective wavelength.

Comparing the measured EWs of T Tau N, given in Table 4.3, with published data from standard star spectra (Greene and Lada, 1996, Meyer et al., 1998) suggests that its NIR spectrum is heavily veiled. The EWs of the photospheric lines given in Table 4.3 are smaller by a factor of about 3 in H- and 7 in K-band than those of an unveiled K1 star with luminosity class III to V typical for pre-main sequence T Tauri stars (Greene and Meyer, 1995).

Table 4.3: Equivalent widths of atmospheric features of T Tau N

	Mg			Al	OH	Ca
λ_0 [μm]	1.575	1.711	2.282	1.674	1.688	2.263
EW [\AA]	0.84 ± 0.45	0.55 ± 0.21	0.33 ± 0.12	0.75 ± 0.44	0.51 ± 0.33	0.27 ± 0.18

Using Equation 4.7, the EWs of T Tau N's photospheric features seem to be diluted by excess emission of $r_k \approx 6$ in K-band, and $r_h \approx 2$ in H-band. These veiling factors are much larger than those expected for class II sources, and are reminiscent of flat spectrum sources for which $r_k \geq 3$ (Greene and Lada, 1996). The ratio of r_k/r_h agrees with the result of Greene and Meyer (1995) who deduced a mean value of 3 for the YSOs of their sample.

Also remarkable are the relatively strong molecular OH (1.688 μm) and CO (1.619 μm) features which are not expected to arise in K1 V or K1 IV photospheres. These features dominate the spectra of very late type stars, and are significantly stronger in stars of lower surface gravity at a respective spectral type (Meyer et al., 1998). Molecular absorption features could also be contributed by a cooler circumstellar disk photosphere (Calvet et al., 1991). Additionally, the silicate emission at 10 μm from

T Tau N (Ghez et al., 1991) might be evidence for a disk atmosphere (Meyer et al., 1999). It should be noted that the OH feature is unlikely to be produced by imperfect telluric calibration, since it does not appear in the spectrum of T Tau S derived from simultaneous observations and identical data reduction procedures.

T Tau S. Apart from the Brackett emission lines, the H- and K-band spectra of T Tau S are featureless. In comparison to our detection limit of about 0.4 \AA , typical EWs of atomic features for late type standard stars are of the order of a few Angströms (Greene and Meyer, 1995, Meyer et al., 1998). This implies that the stellar photosphere is heavily veiled, if visible at all, meaning that we basically see hot dust when observing T Tau S.

4.4 A Model of T Tau S

A proto-stellar or Class I model with an infalling envelope seems unlikely for T Tau S, mostly because of a lack of cold dust in its immediate environment and the presence of its close companion T Tau Sb. Based on the orientation of T Tau S's jet almost perpendicular to the line of sight and the detection of T Tau Sb, Koresko (2000) proposed a scenario in which T Tau S and T Tau Sb are both surrounded by small disks seen nearly edge-on. This could easily provide the required local extinctions without large amounts of dust and could further explain the brightness variations by changes in a disk which processes stellar photons via a combination of extinction and scattering.

What can our data add to this picture? The lack of photospheric light deduced from the H- and K-band spectra rules out a pure extinction scenario and even makes scattering of stellar light at diffuse dust in the upper regions of a disk, such as observed for HK Tau (Koresko, 1998, Stapelfeldt et al., 1998b), an unlikely source of the NIR flux. The relatively low extinction towards the Brackett-series emitting region indicated by our data further suggests a geometry, where the disk is observed at an inclination such that it obscures the star while leaving the line of sight towards the polar regions comparatively unobscured. NIR flux could be contributed by the hot inner parts of the disk itself, if the optical depths in the disk are not too large. Figure 4.15 sketches the model and denotes the regions from where the observed radiation may originate.

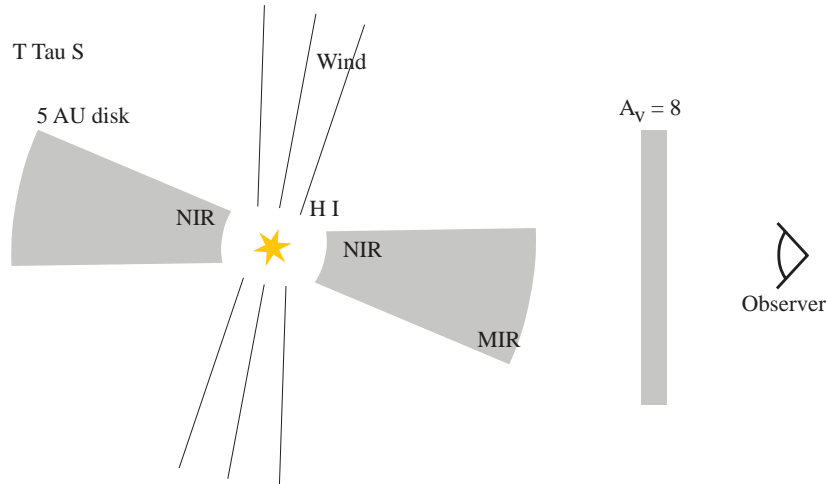


Figure 4.15: Model of T Tau S. The star is obscured by the small truncated disk emitting from near- to mid-infrared. The hydrogen recombination line emitting region is only partly obscured by the disk.

Calculations. To test the scenario described above, we used a two dimensional radiative transfer code developed by Manske and Henning (1998).

This code is based on a method given by and described at length in Men'shchikov and Henning (1997). It relies on a ray-tracing technique to solve the frequency-dependent radiative transfer equations (including isotropic scattering) for flared disks. The main approximation is that the density distribution in the disk depends on the radial coordinate only. The disk is passively heated by stellar radiation. Mean intensities and temperatures are self-consistently calculated for points in the disk's midplane and at its upper and lower conical surfaces only. The flared disk itself is part of a sphere with removed polar cones, i.e. it has a constant opening angle. Extensive explanations of the code and its strategy for solving the radiative transfer problem can be found in the above mentioned papers.

We used common dust grain opacities given by Draine and Lee (1984) as well as a power-law grain size distribution $N(a) \sim a^{-3.5}$ given by Mathis et al. (1977). The actual luminosity of the central star in T Tau S is poorly known. If one adopts spherical geometry, then the luminosity can be obtained from the integrated SED, yielding $11 L_{\odot}$ (Koresko et al., 1997). Such estimates may be off by an order of magnitude, since the fluxes in a non-spherical geometry depend sensitively on a number of parameters including the viewing direction (Men'shchikov and Henning, 1997). Hence, we treated the luminosity of the central star as a free parameter. Its unknown effective temperature is not a critical parameter in the model and was arbitrarily set to 5000 K. The disk was modeled with an opening angle of 40° and an outer radius of 4 AU. The inner radius of the dust is determined by the code and given by the radial distance at which the dust temperature falls below assumed melting temperatures of 1600 K for silicate and 1300 K for carbon particles. In this case, it was 0.5 AU. The radial density distribution of the dust was chosen to be constant, and we adopted a standard dust-to-gas mass ratio of 1 % with a total disk mass of $3.5 \times 10^{-6} M_{\odot}$. The optical depth in the disk's midplane is $\tau_V \approx 30$. The viewing angle was chosen such that the disk's surface is *just* obscuring the central star at an inclination of -18 deg from the line of sight. An additional fore-ground extinction of $A_V = 8$ was assumed, which is required to suppress visible light from the star scattered from the backside of the disk. The basic input data of the model is tabulated below:

Table 4.4: Parameters used in the modelling of T Tau S

<i>Central Object</i>	
Luminosity:	$40 L_{\odot}$
Temperature:	5000 K
Distance:	140 pc
<i>Dust Distribution</i>	
Density distr.:	$\rho(r) \sim const.$
Outer radius:	4 AU
Inner radius:	0.5 AU
<i>Dust Properties</i>	
Size distribution:	$a = 0.005 - 0.250 \mu\text{m}$ $N(a) \sim a^{-3.5}$ (Mathis et al., 1977)
Silicate / Carbon ratio:	Si:C = 6:4
Optical data:	Draine and Lee (1984)
Optical depth of disk:	$\tau_V \approx 30$
Foreground extinction:	$A_V = 8$
Total dust mass:	$3.5 \times 10^{-8} M_{\odot}$
<i>Disk Geometry</i>	
Opening angle:	40°
Inclination:	-18°

Figure 4.16 shows the SED calculated from the model together with photometric data obtained during a high- and low luminosity phase of T Tau S (Ghez et al., 1991, Herbst et al., 1997). A careful inspection of the two photometric data sets suggests that variable extinction is partly responsible for the flaring of T Tau S, explaining the larger differences at short wavelengths and around the silicate absorption feature at $10\mu\text{m}$. The observations, performed with 4-5 m class telescopes, were not able to resolve T Tau N and T Tau S at wavelengths longer than $10\mu\text{m}$. Thus, the apparent lack of mid-infrared luminosity produced by the model can be attributed to an additional unresolved component of dust with moderate temperature ($\approx 150 - 200\text{K}$) contributing to the observations. A special disk geometry with warped edges could also introduce such a dust component.

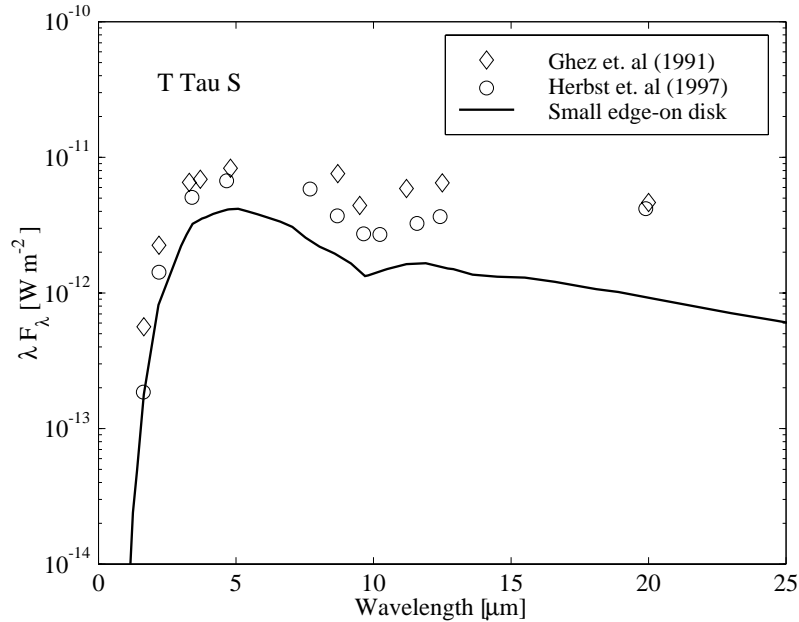


Figure 4.16: Near- and mid-infrared photometry of T Tau S. Observational data from Ghez et al. (1991) and Herbst et al. (1997). The line traces the SED of the small edge-on disk model with a foreground extinction of $A_V = 8$.

While the model is certainly not perfect, it does demonstrate that the assumed geometry of a lightweight, moderately inclined disk can reproduce the basic features of the SED. Stellar light is effectively obscured due to the disk's optical depth of $\tau_V \approx 30$. Most of the NIR luminosity is contributed by the hot ($> 1000\text{K}$) inner parts of the disk explaining the observed lack of photospheric features, while the mid-infrared luminosity originates from outer parts. Additional foreground extinction is a necessity, because straylight from the star scattered at diffuse dust in the polar regions or at the backside of the disk is not observed. This assumption appears reasonable considering that a very tenuous disk has to be replenished by some dust reservoir surrounding T Tau S and T Tau Sb. Otherwise, it would be disrupted on timescales of only a few hundreds years by even modest accretion rates. The edge-on model requires a large stellar luminosity, because most of the light emerges towards the poles of the system without heating the disk. A larger opening angle of the disk could in principle increase the luminosity, but this conflicts with the assumed disk inclination of about -11° (perpendicular to the jet inclined by about 79°) and viewing geometry of the disk just obscuring the star. The assumed value of $40L_\odot$, however, lies within the range of luminosities calculated by D'Antona and Mazzitelli (1994) for young stars with masses around $2M_\odot$ corresponding to the lower limit indicated by the orbit of T Tau Sb.

4.5 Summary

We investigated the stellar components of the T Tauri system using a variety of observational techniques. AO compensated speckle interferometry appeared to be an extremely powerful technique to find binary stars at the telescope's diffraction limit and beyond, allowing us to detect the $0.08''$ companion to T Tau S in the H-band. Together with published data, a projected distance to T Tau S of 11 AU and a linear velocity of 12.4 ± 4 km/s were found for T Tau Sb. Assuming a circular orbit, this implies a mass of at least $1.9 M_{\odot}$ for the T Tau S system.

Using AO compensated imaging spectroscopy, we were able to spatially resolve T Tau N and T Tau S and to obtain individual NIR spectra of the two components for the first time. It was, however, not possible to disentangle the contributions of T Tau S and T Tau Sb. H_2 emission was not detected in the immediate vicinity of the stars, whereas previously published data showed H_2 emission from T Tau N and T Tau S at a lower spatial resolution. T Tau N and T Tau S both exhibit the Brackett series in emission with comparable equivalent widths, while $Br\gamma$ emission seemed to be confined to T Tau N in 1996 when the luminosity of T Tau S was low. At the moment, T Tau S undergoes a high luminosity phase maybe suggesting a common cause of variable emission line spectrum and luminosity. The accretion luminosities deduced from the $Br\gamma$ fluxes are typical for pre-main sequence stars and do not suggest very high accretion rates for both stars. The ratio between $Br\gamma$ and $Br10$ luminosity is similar for T Tau N and T Tau S. Perceiving the low extinction for T Tau N and assuming a similar intrinsic ratio for both stars, the Brackett series emitting region around T Tau S does not appear heavily obscured as compared to the star. Brackett line maps with about $0.15''$ resolution show T Tau N and T Tau S as point sources, no hint of additional components being found between the two.

Photospheric absorption features were only detected for T Tau N. The EWs of these features are weak when compared to published standard star data, implying that the NIR luminosity of T Tau N is dominated by excess emission of hot dust. T Tau S does not exhibit any photospheric features ruling out a pure extinction scenario as an explanation for its unusual SED.

The close companion and lack of cold dust in the environment of T Tau S preclude a spherical infall scenario resembling a Class I source. Given that we do not see a stellar photosphere, but do see the Brackett series emitting region seemingly unobscured, a small edge-on disk may be consistent with the observational results when viewed at an angle such that the star is just obscured. We carried out radiative transfer calculations using a very simple model and found that a small edge-on disk can indeed roughly explain the SED of T Tau S.

In the future, we will follow the orbit of T Tau Sb with the aim to readily derive the mass of T Tau S. Theoretical pre-main sequence tracks allow one to calculate luminosity and effective temperature of a star from its age and mass, and vice versa. T Tau S's luminosity and effective temperature are not observable, but the knowledge of its mass together with the age of T Tau N (coevality assumed) would readily identify this mysterious object. Furthermore, $Pa\beta$ line luminosities should allow us to decide whether the hydrogen recombination line emitting regions are in fact relatively unobscured.

Appendix A

Notations

Quantity	Signification
Scalars	
r_0	Fried-parameter of atmospheric turbulence
D	Aperture diameter of the telescope
d	Subaperture diameter
r_i	Radius of the central obscuration
λ	Wavelength of light
θ	Spot diameter
$\phi(x, y, t)$	Wavefront phase deformation
x, y	Cartesian coordinates in the pupil plane
$\{F_i(\mathbf{r})\}$	Basis set of functions
ρ, θ	Polar coordinates in the pupil plane
F	Framerate of the control loop
\mathbf{v}_\perp	Horizontal wind velocity
σ^2	Phase error variance
σ_s^2	Measurement noise error variance
σ_t^2	Total Phase reconstruction error variance
$\sigma_{r,j}^2$	Reconstruction error variance of mode j
m, n	Number of control modes and number of gradients
j	Mode numbering index
Matrices and Vectors	
\mathbf{a}	Modal coefficients of control modes
\mathbf{a}_\perp	Modal coefficients of complementary modes
\mathbf{g}	SHS gradients
\mathbf{n}	Gradient measurement noise
\mathbf{e}	Effective gradient noise including aliasing
\mathbf{D}	Interaction matrix of control modes
\mathbf{D}_\perp	Interaction matrix of complementary modes
\mathbf{D}^+	Pseudo-Inverse of \mathbf{D}

R	Reconstruction matrix
C	Cross-talk matrix $\mathbf{D}^+ \mathbf{D}_\perp$
Id	Identity matrix
\mathbf{V}_a	Covariance matrix of \mathbf{a} , i.e. $\langle \mathbf{a} \mathbf{a}^t \rangle$
\mathbf{V}_e	Covariance matrix of \mathbf{e} , i.e. $\langle \mathbf{e} \mathbf{e}^t \rangle$
Astronomy	
R	Spectral resolution
M_\odot	Solar mass ($2 \cdot 10^{33}$ g)
L_\odot	Solar luminosity ($3.85 \cdot 10^{33}$ erg/s)
R_\odot	Solar radius ($7 \cdot 10^{10}$ cm)
a	Semimajor axis
F_λ	Flux density [$\text{erg cm}^{-2} \text{s}^{-1} \mu\text{m}^{-1}$]
L_{acc}	Accretion luminosity
$L_{\text{Br}\gamma}$	$\text{Br}\gamma$ line luminosity
A_V	Extinction in the V-band
Misc	
$\hat{\dots}$	Estimate of ...
$FT\{\dots\}$	Fourier transform of ...
$\langle \dots \rangle$	Ensemble average of ...

Appendix B

Acronyms

Acronym	Expression
ALFA	Adaptive Optics with a Laser for Astronomy
AO	Adaptive Optics
APD	Avalanche photo diode
AU	Astronomical unit, distance from sun to earth
CCD	Charge coupled device
CWS	Curvature wavefront sensor
DM	Deformable mirror
EW	Equivalent width
FWHM	Full width at half maximum
K-L	Karhunen-Loève
LGS	Laser guide star
LS	Least-squares
MAP	Maximum a posteriori
mas	milli-arcsecond
NGS	Natural guide star
NIR	Near-infrared
OTF	Optical transfer function
PD	Phase diversity
PSD	Power spectral density
PSF	Point spread function
PWS	Pyramid wavefront sensor
rms	Root mean square
S	Strehl ratio
SED	Spectral energy distribution
SHS	Shack-Hartmann sensor
SNR	Signal to noise ratio
WLS	Weighted least-squares
YSO	Young stellar object

Appendix C

Singular Value Decomposition

The Singular Value Decomposition (SVD) is one of the most important decompositions in matrix computations. It is terrific for the analysis of rank deficiency of a matrix and numerically stable computations with sets of equations or matrices. Based on the SVD, the so called *pseudo-inverse* of a matrix can be defined which is also the least-squares solution of a linear system. Further, it is possible to define the *condition number* of a matrix, which is a measure for the maximum noise amplification of a linear system described by that matrix. The proof of the SVD will be omitted here, it can be found in the book of Golub and Van Loan (1996).

SVD: Any n by m matrix \mathbf{A} of rank r can be factored into

$$\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^t, \quad (\text{C.1})$$

where \mathbf{U} ($n \times n$), and \mathbf{V} ($m \times m$) are orthogonal matrices and \mathbf{S} ($n \times m$) only has the r singular values $\sigma_1, \dots, \sigma_r$ on the diagonal.

The SVD is closely related to the *eigenvector-eigenvalue factorization* of the square matrices $\mathbf{A}\mathbf{A}^t$ and $\mathbf{A}^t\mathbf{A}$. Considering Equation C.1 and the orthogonality property of \mathbf{U} ($\mathbf{U}^t = \mathbf{U}^{-1}$), it is

$$\mathbf{A}^t\mathbf{A} = \mathbf{V}\mathbf{S}^t\mathbf{U}^t\mathbf{U}\mathbf{S}\mathbf{V}^t = \mathbf{V}(\mathbf{S}^t\mathbf{S})\mathbf{V}^t. \quad (\text{C.2})$$

Thus, the columns of \mathbf{V} are the eigenvectors of $\mathbf{A}^t\mathbf{A}$ and the singular values are the square roots of its eigenvalues. Accordingly, the columns of \mathbf{U} are the eigenvectors of $\mathbf{A}\mathbf{A}^t$.

A very important property of \mathbf{U} and \mathbf{V} is that its columns are orthonormal bases for all four *fundamental subspaces* associated with \mathbf{A} :

first r columns of \mathbf{U} :	column space of \mathbf{A}
last $n - r$ columns of \mathbf{U} :	left nullspace of \mathbf{A}
first r columns of \mathbf{V} :	row space of \mathbf{A}
last $m - r$ columns of \mathbf{V} :	nullspace of \mathbf{A}

Consider the rectangular linear system $\mathbf{A}\mathbf{x} = \mathbf{b}$. The column space of \mathbf{A} is the image space of the linear system, i.e. it contains all possible \mathbf{b} vectors given arbitrary \mathbf{x} vectors. The left nullspace of \mathbf{A} then contains the complementary \mathbf{b} vectors which cannot be reached. The row space of \mathbf{A} is mapped onto the column space, and its nullspace (also known as the core of the linear system) is mapped onto a null-vector.

The SVD chooses those bases in an extremely special way. If \mathbf{A} multiplies a column of \mathbf{V} , it produces a multiple of a column of \mathbf{U} , because $\mathbf{AV} = \mathbf{US}$. Because the SVD orders the singular values largest first, the first column of \mathbf{V} is the vector \mathbf{x} that produces the largest \mathbf{b} (first column of \mathbf{U}).

Pseudo-Inverse Suppose the SVD of \mathbf{A} is $\mathbf{A} = \mathbf{USV}^t$. Then, the pseudo-inverse of \mathbf{A} is defined as

$$\mathbf{A}^+ = \mathbf{VS}^+\mathbf{U}^t \quad (\text{C.3})$$

The singular values $\sigma_1, \dots, \sigma_r$ are on the diagonal of \mathbf{S} ($n \times m$), and the reciprocals $1/\sigma_1, \dots, 1/\sigma_r$ are on the diagonal of \mathbf{S}^+ ($m \times n$). The pseudo-inverse of \mathbf{A}^+ is $\mathbf{A}^{++} = \mathbf{A}$. As \mathbf{A} maps its row space onto its column space, and its nullspace onto a null-vector, \mathbf{A}^+ maps the column space onto the row space, and the left nullspace onto a null-vector.

The minimum length least-squares solution to $\mathbf{Ax} = \mathbf{b}$ is $\mathbf{x}^+ = \mathbf{A}^+\mathbf{b}$.

The proof to this statement uses the fact that multiplication by the orthogonal matrix \mathbf{U}^t does not alter the length of a vector, so

$$\|\mathbf{Ax} - \mathbf{b}\| = \|\mathbf{USV}^t\mathbf{x} - \mathbf{b}\| = \|\mathbf{SV}^t\mathbf{x} - \mathbf{U}^t\mathbf{b}\|.$$

Now, let us introduce the new vectors $\mathbf{y} = \mathbf{V}^t\mathbf{x}$ and $\mathbf{p} = \mathbf{U}^t\mathbf{b}$ which have the same lengths as \mathbf{x} and \mathbf{b} . Then minimizing $\|\mathbf{Ax} - \mathbf{b}\|$ is the same as minimizing $\|\mathbf{Sy} - \mathbf{p}\|$. It is easy to see that the shortest \mathbf{y} that minimizes this expression is $\mathbf{y}^+ = \mathbf{S}^+\mathbf{p}$, so the minimum length solution is just

$$\mathbf{x}^+ = \mathbf{Vy}^+ = \mathbf{VS}^+\mathbf{U}^t\mathbf{b}. \quad (\text{C.4})$$

In adaptive optics, one usually encounters the problem $\mathbf{Ax} = \mathbf{b}$, with an interaction matrix \mathbf{A} ($n \times m$, number of modes $m <$ number of gradients n) of full column rank ($r = m$). The least-squares estimate in this case can be obtained from the pseudo-inverse via:

$$\mathbf{VS}^+\mathbf{U}^t = \mathbf{V}(\mathbf{S}^t\mathbf{S})^{-1}\mathbf{S}^t\mathbf{U}^t \quad (\text{C.5})$$

$$= \mathbf{V}(\mathbf{S}^t\mathbf{S})^{-1}\mathbf{V}^t\mathbf{VS}^t\mathbf{U}^t \quad (\text{C.6})$$

$$= (\mathbf{A}^t\mathbf{A})^{-1}\mathbf{A}^t. \quad (\text{C.7})$$

For the first term, the properties of \mathbf{A} given above have been used. For the second, the identity matrix $\mathbf{Id} = \mathbf{V}^t\mathbf{V}$ was inserted, and for the third Equation C.2 was used.

Condition number The sensitivity of a non-singular linear system $\mathbf{Ax} = \mathbf{b}$ to slight perturbations in either \mathbf{A} or \mathbf{b} with respect to inversion can be measured by the so-called condition number $\kappa(\mathbf{A})$. It is defined as the ratio between the largest and the smallest singular value of \mathbf{A} , i.e.

$$\kappa(\mathbf{A}) = \sigma_1/\sigma_r. \quad (\text{C.8})$$

Golub and Van Loan (1996) derived the following inequality:

$$\frac{\Delta\mathbf{x}}{\mathbf{x}} \leq \kappa(\mathbf{A})\left(\frac{\Delta\mathbf{A}}{\mathbf{A}} + \frac{\Delta\mathbf{b}}{\mathbf{b}}\right). \quad (\text{C.9})$$

Thus, the relative error in \mathbf{x} can be $\kappa(\mathbf{A})$ times the relative errors in \mathbf{A} and \mathbf{b} .

Appendix D

Gaussian Estimation

Melsa and Cohn (1978) have written an excellent textbook on estimation theory. Here, a section which is of major interest within the scope of this work is summarized.

The estimation problem considered is the case of a Gaussian message which is corrupted by Gaussian noise. In the context of this work, the message would be an aberrated wavefront which is known to obey Gaussian statistics (Roddier, 1981) and the noise is given by the sensor measurement noise \mathbf{n} . In the most general case, the message will be a K -dimensional vector and the observation will be an I -dimensional vector. Further, the observation \mathbf{z} corresponding to a SHS slope vector is considered a linear operation on the message \mathbf{a} , such that

$$\mathbf{z} = \mathbf{H}\mathbf{a} + \mathbf{n}. \quad (\text{D.1})$$

The objective of this section will be to form the maximum a posteriori (MAP) estimate (which is identical with the minimum variance estimate in the Gaussian case) of \mathbf{a} based on the observation \mathbf{z} . The first step is to write the conditional probability density $p(\mathbf{a}|\mathbf{z})$ of \mathbf{a} as a function of \mathbf{z} , the estimate would then be the maximum of this density. It can easily be obtained from *Bayes rule*:

$$p(\mathbf{a}|\mathbf{z}) = \frac{p(\mathbf{z}|\mathbf{a})p(\mathbf{a})}{p(\mathbf{z})}. \quad (\text{D.2})$$

Since all densities on the right side of Equation D.2 are Gaussian, they are completely determined by their mean vectors and their variance matrices. For simplicity, it will be assumed that all quantities are zero mean which will be the case if no static aberrations are present. The variance matrices $\langle \mathbf{a}\mathbf{a}^t \rangle$ and $\langle \mathbf{n}\mathbf{n}^t \rangle$ will be denoted \mathbf{V}_0 and \mathbf{V}_n .

Now, the density function of the K -dimensional message is just

$$p(\mathbf{a}) = \frac{1}{(2\pi)^{K/2}(\det \mathbf{V}_0)^{1/2}} \exp \left\{ -\frac{1}{2} \mathbf{a}^t \mathbf{V}_0^{-1} \mathbf{a} \right\}. \quad (\text{D.3})$$

When \mathbf{a} is known, \mathbf{z} is just the sum of a constant $\mathbf{H}\mathbf{a}$ and \mathbf{n} . Therefore, its mean is just this constant, and its variance matrix is \mathbf{V}_n yielding the conditional density of \mathbf{z} :

$$p(\mathbf{z}|\mathbf{a}) = \frac{1}{(2\pi)^{I/2}(\det \mathbf{V}_n)^{1/2}} \exp \left\{ -\frac{1}{2} (\mathbf{z} - \mathbf{H}\mathbf{a})^t \mathbf{V}_n^{-1} (\mathbf{z} - \mathbf{H}\mathbf{a}) \right\}. \quad (\text{D.4})$$

Using the Equations D.2, D.3, and D.4, taking logarithms and discarding terms that do not vary with \mathbf{a} (like $p(\mathbf{z})$), we find that the vector \mathbf{a} that maximizes $p(\mathbf{a}|\mathbf{z})$ should be selected to satisfy

$$\frac{\partial}{\partial \mathbf{a}} \left[\mathbf{a}^t \mathbf{V}_0^{-1} \mathbf{a} + (\mathbf{z} - \mathbf{H}\mathbf{a})^t \mathbf{V}_n^{-1} (\mathbf{z} - \mathbf{H}\mathbf{a}) \right] |_{\mathbf{a}=\hat{\mathbf{a}}} = \mathbf{0}. \quad (\text{D.5})$$

With the relations

$$\begin{aligned}\frac{\partial}{\partial \mathbf{x}} \mathbf{x}^t \mathbf{A} \mathbf{z} &= \mathbf{A} \mathbf{z} \\ \frac{\partial}{\partial \mathbf{x}} \mathbf{z}^t \mathbf{A} \mathbf{x} &= \mathbf{A}^t \mathbf{z} \\ \frac{\partial}{\partial \mathbf{x}} \mathbf{x}^t \mathbf{A} \mathbf{x} &= (\mathbf{A} + \mathbf{A}^t) \mathbf{x},\end{aligned}\tag{D.6}$$

the derivative in Equation D.5 can be written as

$$-\mathbf{H}^t \mathbf{V}_n^{-1} (\mathbf{z} - \mathbf{H} \hat{\mathbf{a}}) + \mathbf{V}_0^{-1} \hat{\mathbf{a}} = \mathbf{0}.\tag{D.7}$$

Solving this equation for $\hat{\mathbf{a}}$ yields the desired MAP estimate of \mathbf{a} :

$$\hat{\mathbf{a}} = (\mathbf{H}^t \mathbf{V}_n^{-1} \mathbf{H} + \mathbf{V}_0^{-1})^{-1} \mathbf{H}^t \mathbf{V}_n^{-1} \mathbf{z}.\tag{D.8}$$

Since all quantities have been treated as Gaussian random variables, this is not only the MAP estimate but the optimal estimate for a broad class of Bayes cost criteria.

Given the estimator in Equation D.8, the *estimation error* can be written as

$$\begin{aligned}\tilde{\mathbf{a}} &= \mathbf{a} - \hat{\mathbf{a}} \\ &= (\mathbf{H}^t \mathbf{V}_n^{-1} \mathbf{H} + \mathbf{V}_0^{-1})^{-1} [(\mathbf{H}^t \mathbf{V}_n^{-1} \mathbf{H} + \mathbf{V}_0^{-1}) \mathbf{a} - \mathbf{H}^t \mathbf{V}_n^{-1} \mathbf{z}] \\ &= (\mathbf{H}^t \mathbf{V}_n^{-1} \mathbf{H} + \mathbf{V}_0^{-1})^{-1} [\mathbf{V}_0^{-1} \mathbf{a} - \mathbf{H}^t \mathbf{V}_n^{-1} \mathbf{n}].\end{aligned}\tag{D.9}$$

For the last step, Equation D.1 was used and some terms have been canceled. Now, the estimation error variance $\langle \tilde{\mathbf{a}} \tilde{\mathbf{a}}^t \rangle$ is just the sum of the two independent terms in Equation D.9. Using this fact and noting that the variance matrices are symmetric, it is

$$\begin{aligned}\langle \tilde{\mathbf{a}} \tilde{\mathbf{a}}^t \rangle &= (\mathbf{H}^t \mathbf{V}_n^{-1} \mathbf{H} + \mathbf{V}_0^{-1})^{-1} [\mathbf{V}_0^{-1} \mathbf{V}_0 \mathbf{V}_0^{-1} + \mathbf{H}^t \mathbf{V}_n^{-1} \mathbf{V}_n \mathbf{V}_n^{-1} \mathbf{H}] (\mathbf{H}^t \mathbf{V}_n^{-1} \mathbf{H} + \mathbf{V}_0^{-1})^{-1} \\ &= (\mathbf{H}^t \mathbf{V}_n^{-1} \mathbf{H} + \mathbf{V}_0^{-1})^{-1}.\end{aligned}\tag{D.10}$$

Bibliography

- Anders, S. W., Maiolino, R., Thatte, N. A., and Genzel, R. (1998). Aperture interchange module (AIM) diffraction limited NIR spectroscopy with 3D and ALFA. *Proc. SPIE*, 3354:222–231.
- Babcock, H. W. (1953). The Possibility of Compensating Astronomical Seeing. *PASP*, 65:229+.
- Beckwith, S., Gatley, I., Matthews, K., and Neugebauer, G. (1978). Molecular hydrogen emission from T Tauri stars. *ApJ Letters*, 223:L41–L43.
- Berkefeld, T. (1998). *Untersuchungen zur Messung und Korrektur einzelner Schichten der Erdatmosphäre*. Dissertation, Universität Heidelberg.
- Berkefeld, T., Glindemann, A., and Hippler, S. (2000). Anisoplanatism and Turbulence Separation in Multi-conjugate Adaptive Optics. *submitted to Exp. Astronomy*.
- Bevington, P. R. and Robinson, D. K. (1992). *Data Reduction and Error Analysis for the Physical Sciences*. McGraw Hill.
- Bonaccini, D., Farinato, J., Comin, M., and Andrichettoni, M. (2000). ESO STRAP units. *Proc. SPIE*, 4007.
- Born, M. and Wolf, E. (1970). *Principles of Optics*. Pergamon Press, Oxford.
- Brandner, W. and Zinnecker, H. (1997). Physical properties of 90AU to 250AU pre-main-sequence binaries. *A&A*, 321:220–228.
- Bronstein, I. N., Semendjajew, K. A., and Musiol, G. (1999). *Taschenbuch der Mathematik*. Harri Deutsch, FfM.
- Butler, D. J., Davies, R. I., Hayden, F., Redfern, M., Ageorges, N., Hackenberg, W., Rohloff, R.-R., Rabien, S., Ott, T., and Hippler, S. (2000). Sodium layer monitoring at Calar Alto by LIDAR. *Proc. SPIE*, 4007.
- Calvet, N. and Gullbring, E. (1998). The Structure and Emission of the Accretion Shock in T Tauri Stars. *ApJ*, 509:802–818.
- Calvet, N., Hartmann, L., and Kenyon, S. J. (1991). On the near-infrared spectrum of FU Orionis. *ApJ*, 383:752–756.
- Cannon, R. C. (1996). Optimal bases for wave-front simulation and reconstruction on annular apertures. *JOSA A*, 13:862–867.

- Carrano, C. J., Olivier, S. S., Brase, J. M., Macintosh, B. A., and An, J. R. (1998). Phase retrieval techniques for adaptive optics. *Proc. SPIE*, 3353:658–667.
- Close, L. M. (2000). A Review of Published Galactic and Solar System Science: A Bright Future for Adaptive Optics Science. *Proc. SPIE*, 4007.
- Cohen, M. and Kuhl, L. V. (1979). Observational studies of pre-main-sequence evolution. *ApJ suppl. ser.*, 41:743–843.
- Conan, J.-M., Rousset, G., and Madec, P.-Y. (1995). Wave-front temporal spectra in high-resolution imaging through turbulence. *JOSA A*, 12:1559–1570.
- Cox, A. (2000). *Astrophysical Quantities*. AIP Press.
- Dai, G.-m. (1996). Modal wave-front reconstruction with Zernike polynomials and Karhunen-Loève functions. *JOSA A*, 13:1218–1225.
- D’Antona, F. and Mazzitelli, I. (1994). New pre-main-sequence tracks for M less than or equal to 2.5 solar mass as tests of opacities and convection model. *ApJ suppl. ser.*, 90:467–500.
- Davies, R., Eckart, A., Hackenberg, W., Ott, T., Butler, D., Kasper, M., and Quirrenbach, A. (2000). The ALFA laser guide star: operation and results. *Experimental Astronomy*, 10:103–121.
- Draine, B. T. and Lee, H. M. (1984). Optical properties of interstellar graphite and silicate grains. *ApJ*, 285:89–108.
- Duchêne, G., Monin, J., Bouvier, J., and Ménard, F. (1999). Accretion in Taurus PMS binaries: a spectroscopic study. *A&A*, 351:954–962.
- Dyck, H. M., Simon, T., and Zuckerman, B. (1982). Discovery of an infrared companion to T Tauri. *ApJ Letters*, 255:L103–L106.
- Eisloffel, J. and Mundt, R. (1998). Imaging and Kinematic Studies of Young Stellar Object Jets in Taurus. *Astron. Journal*, 115:1554–1575.
- Esposito, S., Feeney, O., and Riccardi, A. (2000). Laboratory test of a Pyramid Wavefront Sensor. *Proc. SPIE*, 4007.
- Evans, N. J., I., Levreault, R. M., Beckwith, S., and Skrutskie, M. (1987). Observations of infrared emission lines and radio continuum emission from pre-main-sequence objects. *ApJ*, 320:364–375.
- Foy, R. and Labeyrie, A. (1985). Feasibility of adaptive telescope with laser probe. *A&A*, 152:L29–L31.
- Foy, R., Migus, A., Biraben, F., Grynberg, G., McCullough, P. R., and Tallon, M. (1995). The polychromatic artificial sodium star: a new concept for correcting the atmospheric tilt. *A&A suppl.*, 111:569+.
- Franklin, G. F., Powell, J. D., and Emami-Naeini, A. (1994). *Feedback Control of Dynamic Systems*. Addison-Wesley.
- Fried, D. (1965). Statistics of a geometric representation of wavefront distortion. *J Opt Soc Am*, 55:1427–1435.

- Fried, D. L. (1977). Least-squares fitting a wave-front distortion estimate to an array of phase-difference measurements. *J Opt Soc Am*, 67:370–375.
- Fried, D. L. and Belsher, J. F. (1994). Analysis of fundamental limits to artificial guide star adaptive optics system performances for astronomical imaging. *JOSA A*, 11:277–287.
- Frieden, B. R. (1991). *Probability, Statistical Optics, and Data Testing*. Springer Verlag.
- Gendron, E. and Lena, P. (1994). Astronomical adaptive optics. 1: Modal control optimization. *A&A*, 291:337–347.
- Gendron, E. and Lena, P. (1995). Astronomical adaptive optics. 2: Experimental results of an optimized modal control. *A&A suppl.*, 111:153+.
- Ghez, A. M., Neugebauer, G., Gorham, P. W., Haniff, C. A., Kulkarni, S. R., Matthews, K., Koresko, C., and Beckwith, S. (1991). Diffraction limited infrared images of the binary star T Tauri. *Astron. Journal*, 102:2066–2072.
- Ghez, A. M., Weinberger, A. J., Neugebauer, G., Matthews, K., and McCarthy, D. W. (1995). Speckle Imaging Measurements of the Relative Tangential Velocities of the Components of T Tauri Binary Stars. *Astron. Journal*, 110:753+.
- Glindemann, A., Hamilton, D., Hippler, S., Rohloff, R.-R., and Wagner, K. (1997a). ALFA - The Laser Guide Star Adaptive Optics System for the Calar Alto 3.5-m Telescope. *ESO workshop on Laser Technology for Laser Guide Star Adaptive Optics Astronomy*, Ed. N. Hubin, pages 120–125.
- Glindemann, A., McCaughrean, M. J., Hippler, S., Birk, C., Wagner, K., and Rohloff, R.-R. (1997b). CHARM - A Tip-Tilt Tertiary System for the Calar Alto 3.5m Telescope. *PASP*, 109:688–696.
- Golub, G. H. and Van Loan, C. F. (1996). *Matrix Computations*. John Hopkins University Press.
- Gonsalves, R. A. (1982). Phase retrieval and diversity in adaptive optics. *Opt. eng.*, 21:829–832.
- Greene, T. P. and Lada, C. J. (1996). Near-Infrared Spectra and the Evolutionary Status of Young Stellar Objects: Results of a 1.1-2.4 micron Survey. *Astron. Journal*, 112:2184+.
- Greene, T. P. and Meyer, M. R. (1995). An Infrared Spectroscopic Survey of the rho Ophiuchi Young Stellar Cluster: Masses and Ages from the H-R Diagram. *ApJ*, 450:233+.
- Greenwood, D. P. (1977). Bandwidth specification for adaptive optics systems. *J. Opt. Soc. Am.*, 67:174–176.
- Hardy, J. W. (1998). *Adaptive Optics for Astronomical Telescopes*. Oxford University Press.
- Hardy, J. W., Lefebvre, J. E., and Koliopoulos, C. L. (1977). Real-time atmospheric compensation. *J Opt Soc Am*, 67:360–369.
- Hartigan, P., Edwards, S., and Ghandour, L. (1995). Disk Accretion and Mass Loss from Young Stars. *ApJ*, 452:736+.
- Hartmann, L. (1998). *Accretion Processes in Star Formation*. Cambridge University Press.
- Herbig, G. H. (1977). Eruptive phenomena in early stellar evolution. *ApJ*, 217:693–715.

- Herbst, T. M., Beckwith, S. V. W., Glindemann, A., Tacconi-Garman, L. E., Kroker, H., and Krabbe, A. (1996). A Near-Infrared Spectral Imaging Study of T Tau. *Astron. Journal*, 111:2403+.
- Herbst, T. M., Robberto, M., and Beckwith, S. V. W. (1997). Wind-Disk-Ambient Cloud Interactions in the Near Environment of T Tauri. *Astron. Journal*, 114:744+.
- Herrmann, J. (1981). Cross coupling and aliasing in modal wave-front estimation. *JOSA A*, 71:989–992.
- Hippler, S., Glindemann, A., Kasper, M., Kalas, P., Rohloff, R., Wagner, K., Looze, D. P., and Hackenberg, W. K. (1998). ALFA: the MPIA/MPE adaptive optics with a laser for astronomy project. *Proc. SPIE*, 3353:44–55.
- Hippler, S., Kasper, M. E., Feldt, M., Weiß, A. R., Looze, D. P., Montoya, L., and Aceituno, J. (2000). ALFA: three years of experience with adaptive optics with a laser guide star. *Proc. SPIE*, 4007.
- Hogerheijde, M. R., van Langevelde, H. J., Mundy, L. G., Blake, G. A., and van Dishoeck, E. F. (1997). Subarcsecond Imaging at 267 GHz of a Young Binary System: Detection of a Dust Disk of Radius Less than 70 AU around T Tauri N. *ApJ Letters*, 490:L99+.
- Hudgins, R. H. (1977). Wave-front reconstruction for compensated imaging. *J Opt Soc Am*, 67:375–378.
- Irwan, R. and Lane, R. G. (1999). Analysis of optimal centroid estimation applied to Shack-Hartmann sensing. *Applied Optics*, 38:6737–6743.
- Jensen, E. L. N., Donar, A. X., and Mathieu, R. D. (2000). Aligned Disks in Pre-Main Sequence Binaries. In *Proc. IAU Symposium No. 200*, eds. B. Reipurth and H. Zinnecker, pages 85+.
- Kasper, M. E. (1997). *Das astronomische Seeing. Theorie, Methodik und Messungen*. Diploma Thesis, University of Heidelberg.
- Kasper, M. E., Looze, D. P., Hippler, S., Davies, R., and Glindemann, A. (1999). Increasing the sensitivity of a Shack-Hartmann Sensor. *Conf. Wavefront Sensing, Canterbury*.
- Kasper, M. E., Looze, D. P., Hippler, S., Feldt, M., Weiß, A. R., Glindemann, A., and Davies, R. (2000a). A practical approach to modal basis selection and wavefront estimation. *Proc. SPIE*, 4007.
- Kasper, M. E., Looze, D. P., Hippler, S., Herbst, T., Glindemann, A., Ott, T., and Wirth, A. (2000b). ALFA: Optics and Control System. *Experimental Astronomy*, 10:49–73.
- Knox, K. T. and Thompson, B. J. (1974). Recovery of images from atmospherically degraded short-exposure photographs. *ApJ Letters*, 193:L45–L48.
- Kobayashi, N., Nagata, T., Hodapp, K., and Hora, J. L. (1994). Small outburst on the infrared companion of T Tauri finished. *PASJ*, 46:L183–L186.
- Köhler, R., Kasper, M. E., and Herbst, T. M. (2000). T Tauri S confirmed as a binary Star. In *Proc. IAU Symposium No. 200*, eds. B. Reipurth and H. Zinnecker, pages 63+.
- Königl, A. and Pudritz, R. E. (2000). in *Protostars and Planets IV*, eds. Mannings, V., Boss, A., Russell, S. University of Arizona Press.

- Koresko, C. D. (1998). A Circumstellar Disk in a Pre-main-sequence Binary Star. *ApJ Letters*, 507:L145–L148.
- Koresko, C. D. (2000). A Third Star in the T Tauri System. *ApJ Letters*, 531:L147–L149.
- Koresko, C. D., Herbst, T. M., and Leinert, C. (1997). The Infrared Companions of T Tauri Stars. *ApJ*, 480:741+.
- Labeyrie, A. (1970). Attainment of Diffraction Limited Resolution in Large Telescopes by Fourier Analysing Speckle Patterns in Star Images. *A&A*, 6:85+.
- Lane, R. G. and Tallon, M. (1992). Wave-front reconstruction using a Shack-Hartmann sensor. *Applied Optics*, 31:6902–6908.
- Law, N. F. and Lane, R. G. (1996). Wavefront estimation at low light levels. *Opt. comm.*, 126:19–24.
- Le Louarn, M., Foy, R., Hubin, N., and Tallon, M. (1998). Laser Guide Star for 3.6- and 8-m telescopes: Performance and astrophysical implications. *MNRAS*, 295:756–768.
- Lenzen, R., Bizenberger, P., Salm, N., and Storz, C. (1998). Omega Cass: a new multimode NIR-imager/spectrometer for the Calar Alto Observatory. *Proc. SPIE*, 3354:493–499.
- Löfdahl, M. G., Duncan, A. L., and Scharmer, G. B. (1998). Fast phase diversity wavefront sensing for mirror control. *Proc. SPIE*, 3353:952–963.
- Looze, D. P., Beker, O., Kasper, M., and Hippler, S. (1999). Compensator Design for the ALFA Adaptive Optics System. In *Conf. on Control Appl., Hawaii*.
- Lubow, S. and Artymowicz, P. (2000). in *Protostars and Planets IV*, eds. Mannings, V., Boss, A., Russell, S. University of Arizona Press.
- Mahajan, V. N. (1994). Zernike Annular Polynomials and optical aberrations of Systems with Annular Pupils. *Applied Optics Suppl.*, 5:8125–8127.
- Manske, V. and Henning, T. (1998). Two-dimensional radiative transfer with transiently heated particles: methods and applications. *A&A*, 337:85–95.
- Mathis, J. S. (1990). Interstellar dust and extinction. *Ann. Rev A&A*, 28:37–70.
- Mathis, J. S., Rumpl, W., and Nordsieck, K. H. (1977). The size distribution of interstellar grains. *ApJ*, 217:425–433.
- Melsa, J. L. and Cohn, D. L. (1978). *Decision and Estimation Theory*. McGraw-Hill.
- Men'shchikov, A. B. and Henning, T. (1997). Radiation transfer in circumstellar disks. *A&A*, 318:879–907.
- Meyer, M. R., Edwards, S., Hinkle, K. H., and Strom, S. E. (1998). Near-Infrared Classification Spectroscopy: H-Band Spectra of Fundamental MK Standards. *ApJ*, 508:397–409.
- Meyer, M. R., Natta, A., and Beckwith, S. V. W. (1999). Silicate Emission in T Tauri Stars: Evidence for Disk Atmospheres? In *American Astronomical Society Meeting*, volume 195, pages 13508+.

- Monin, J., Menard, F., and Duchene, G. (1998). Using polarimetry to check rotation alignment in PMS binary stars. Principles of the method and first results. *A&A*, 339:113–122.
- Muzerolle, J., Hartmann, L., and Calvet, N. (1998). A Bragg Probe of Disk Accretion in T Tauri Stars and Embedded Young Stellar Objects. *Astron. Journal*, 116:2965–2974.
- Natta, A., Giovanardi, C., and Palla, F. (1988). Ionization structure and emission of winds from low-luminosity pre-main-sequence stars. *ApJ*, 332:921–939.
- Nisenson, P., Stachnik, R. V., Karovska, M., and Noyes, R. (1985). A new optical source associated with T Tauri. *ApJ Letters*, 297:L17–L20.
- Noll, R. J. (1976). Zernike Polynomials and atmospheric turbulence. *J Opt Soc Am*, 66:207–211.
- Osterbrock, D. E. (1989). *Astrophysics of Gaseous Nebulae and Active Galactic Nuclei*. University Science Books, Mill Valley, California.
- O’Sullivan, C., Redfern, M., Ageorges, N., Holstenberg, H.-C., Rabien, S., Ott, T., Davies, R., and Eckart, A. (2000). Short Timescale variability of the mesospheric sodium layer. *Experimental Astronomy*, 10:147–156.
- Otsubo, M., Takami, H., and Iye, M. (1997). Holographic Atmospheric Turbulence Simulator for Testing Adaptive Optics Systems. *PASP*, 109:1057–1061.
- Ott, T., Hackenberg, W., Rabien, S., Eckart, A., and Hippler, S. (2000). The ALFA laser: beam relay and control system. *Experimental Astronomy*, 10:89–101.
- Paxman, R. G., Schulz, T. J., and R., F. J. (1992). Joint Estimation of object and aberrations by using phase diversity. *JOSA A*, 9:1072–1085.
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (1994). *Numerical Recipes in C*. Cambridge University Press.
- Primot, J., Rousset, G., and C., F. J. (1990). Deconvolution from wave-front sensing: a new technique for compensating turbulence-degraded images. *JOSA A*, 7:1598–1608.
- Rabien, S., Ott, T., Hackenberg, W., Eckart, A., Davies, R., Kasper, M., and A., Q. (2000). The ALFA laser and analysis tools. *Experimental Astronomy*, 10:75–88.
- Ragazzoni, R. (1996). Pupil plane wavefront sensing with an oscillating prism. *J. Modern Opt.*, 43:289–293.
- Ragazzoni, R. (1999). No Laser Guide Stars for adaptive optics in giant telescopes? *A&A suppl.*, 136:205–209.
- Ragazzoni, R. and Esposito, S. (1999). Laser guide star absolute tilt recovery using a single auxiliary telescope. *MNRAS*, 307:55–57.
- Ragazzoni, R. and Farinato, J. (1999). Sensitivity of a pyramidal Wave Front sensor in closed loop Adaptive Optics. *A&A*, 350:L23–L26.
- Ragazzoni, R., Marchetti, E., and Rigaut, F. (1999). Modal tomography for adaptive optics. *A&A*, 342:L53–L56.

- Ray, T. P., Muxlow, T. W. B., Axon, D. J., Brown, A., Corcoran, D., Dyson, J., and Mundt, R. (1997). Large-scale magnetic fields in the outflow from the young stellar object T Tauri S. *Nature*, 385:415–417.
- Riccardi, A., Bindi, N., Ragazzoni, R., Esposito, S., and Stefanini, P. (1998). Laboratory characterization of a Foucault-like wavefront sensor for adaptive optics. *Proc. SPIE*, 3353:941–951.
- Rigaut, F., Ellerbroek, B. L., and Northcott, M. J. (1997). Comparison of curvature-based and Shack-Hartmann-based adaptive optics for the Gemini telescope. *Applied Optics*, 36:2856–2867.
- Rigaut, F. and Gendron, E. (1992). Laser guide star in adaptive optics - The tilt determination problem. *A&A*, 261:677–684.
- Rigaut, F., Salmon, D., Arsenault, R., Thomas, J., Lai, O., Rouan, D., Véran, J. P., Gigan, P., Cramp-ton, D., Fletcher, J. M., Stilburn, J., Boyer, C., and Jagourel, P. (1998a). Performance of the Canada-France-Hawaii Telescope Adaptive Optics Bonnette. *PASP*, 110:152–164.
- Rigaut, F. J., Veran, J., and Lai, O. (1998b). An analytical model for Shack-Hartmann-based adaptive optics systems. *Proc. SPIE*, 3353:1038–1048.
- Roddier, F. (1981). The effects of atmospheric turbulence in optical astronomy. In *Progress in optics, North-Holland Publishing Co., Amsterdam*, volume 19, pages 281–376.
- Roddier, F. (1988). Curvature sensing and compensation: a new concept in adaptive optics. *Applied Optics*, 27:1223–1225.
- Roddier, F. (1995). Error propagation in a closed-loop adaptive optics system: a comparison between Shack-Hartmann and curvature wave-front sensors. *Opt. comm.*, 113:357–359.
- Roddier, F. (1999). *Adaptive Optics in Astronomy*. Cambridge University Press.
- Roddier, F., Roddier, C., Brandner, W., and Charissoux, D. (2000). Five-Years of Adaptive Optics Observations of T Tau South. In *Proc. IAU Symposium No. 200, eds. B. Reipurth and H. Zinnecker*, pages 60+.
- Roddier, F., Roddier, C., and Roddier, N. (1988). Curvature sensing: a new wave-front sensing method. *Proc. SPIE*, 976:203–209.
- Roggemann, M. C. and Welsh, B. (1996). *Imaging Through Turbulence*. CRC Press.
- Rousset, G., Fontanella, J. C., Kern, P., Gigan, P., and Rigaut, F. (1990). First diffraction-limited astronomical images with adaptive optics. *A&A*, 230:L29–L32.
- Rousset, G., Fontanella, J. C., Kern, P., Lena, P., Gigan, P., and Rigaut, F. (1993). Wave-front sensing. In *NATO ASI Series, Kluwer, Dordrecht*, volume 423, pages 336–344.
- Sallberg, S. A., Welsh, B. M., and C., R. M. (1997). Maximum a posteriori estimation of wave-front slopes using a Shack-Hartmann wave-front sensor. *JOSA A*, 14:1347–1353.
- Schwartz, P. R., Simon, T., and Campbell, R. (1986). The T Tauri radio source. II - The winds of T Tauri. *ApJ*, 303:233–238.

- Solf, J. and Böhm, K. H. (1999). A Deep Long-Slit Spectroscopic Study of the Two Bipolar Outflows from the T Tauri Binary System. *ApJ*, 523:709–720.
- Stapelfeldt, K. R., Burrows, C. J., Krist, J. E., Watson, A. M., Ballester, G. E., Clarke, J. T., Crisp, D., Evans, R. W., Gallagher, J. S., Griffiths, R. E., Hester, J. J., Hoessel, J. G., Holtzman, J. A., Mould, J. R., Scowen, P. A., Trauger, J. T., and Westphal, J. A. (1998a). Hubble Space Telescope Imaging of the Circumstellar Nebulosity of T Tauri. *ApJ*, 508:736–743.
- Stapelfeldt, K. R., Krist, J. E., Menard, F., Bouvier, J., Padgett, D. L., and Burrows, C. J. (1998b). An Edge-On Circumstellar Disk in the Young Binary System HK Tauri. *ApJ Letters*, 502:L65–+.
- Takato, N., Iye, M., and Yamaguchi, I. (1994). Wavefront reconstruction error of Shack-Hartmann wavefront sensors. *PASP*, 106:182–188.
- Tallon, M. and Foy, R. (1990). Adaptive telescope with laser probe - Isoplanatism and cone effect. *A&A*, 235:549–557.
- Tyson, R. K. (1991). *Principles of Adaptive Optics*. Academic Press.
- van Langevelde, H. J., van Dishoeck, E. F., van Der Werf, P. P., and Blake, G. A. (1994). The spatial distribution of excited H₂ in T Tau: A molecular outflow in a young binary system. *A&A*, 287:L25–L28.
- Veran, J.-P., Rigaut, F., Maitre, H., and Rouan, D. (1997). Estimation of the adaptive optics long-exposure point-spread function using control loop data. *JOSA A*, 14:3057+.
- Wallace, L. and Hinkle, K. (1997). Medium-Resolution Spectra of Normal Stars in the K Band. *ApJ suppl. ser.*, 111:445+.
- Wallner, E. P. (1983). Optimal wave-front correction using slope measurements. *JOSA A*, 73:1771–1776.
- Wang, J. Y. and Markey, J. K. (1978). Modal compensation of atmospheric turbulence phase distortion. *J Opt Soc Am*, 68:78–87.
- Weitzel, L., Krabbe, A., Kroker, H., Thatte, N., Tacconi-Garman, L. E., Cameron, M., and Genzel, R. (1996). 3D: The next generation near-infrared imaging spectrometer. *A&A suppl.*, 119:531–546.
- Wirth, A., Navetta, J., Looze, D. P., Hippler, S., Glindemann, A., and Hamilton, D. (1998). Real-time modal control implementation for adaptive optics. *Applied Optics*, 37:4586–4597.

DANKE AN . . .

Hier möchte ich mich herzlich bei allen Personen bedanken, die zum Gelingen dieser Arbeit beigetragen haben. Besonders hervorheben möchte ich den Dank an

Stefan Hippler, der mir als ALFA Projektleiter alle erdenkliche Unterstützung zukommen ließ. Als Ratgeber und Motivator in schwierigen Situationen und auf entlegenen Gipfeln hatte er sicherlich den größten Anteil am Gelingen meiner Arbeit;

Andreas Glindemann für die wissenschaftlich Betreuung und die Idee der Dissertation. Trotz der räumlichen Distanz, die eine enge Zusammenarbeit ausschloß, war er durch seine fachliche und soziale Kompetenz immer ein wertvoller Rückhalt;

die Direktoren des MPIA, Prof. Rix und Prof. Appenzeller, für neuen Schwung und ihre Unterstützung für das ALFA Projekt;

Tom Herbst für astronomischen Beistand und die Möglichkeit mich an seinen Projekten beteiligen. Sein Sachverstand auf allen Gebieten der Instrumentierung und Beobachtung war für mich von großem Nutzen.

meine Kollegen Markus Feldt und Robert Weiß für die große Unterstützung zum Ende meiner Arbeit und die vielen frohen Stunden abends und morgens nach Feierabend;

Douglas P. Looze von dem ich sehr viel lernte über Regelschleifen, Signalverarbeitung und Baseball;

R.-R. Rohloff, Sam Wagner und Peter Bizenberger für die mechanische, elektronische und optische Instandhaltung von ALFA;

meine Münchner ALFA-Kollegen Ric Davies, Thomas Ott, Sebastian Rabien und Wolfgang Hackenberg für viele durchgemachte Nächte. Die Erzeugung des LGS, auch bei klirrender Kälte oder ätzenden Dämpfen, war sicherlich der schwierigste Teil des Projekts;

die früheren Alfisti Andreas Eckart, Paul Kalas und Chris Holstenberg für gute Zusammenarbeit;

Luzma Montoya und Jesus Aceituno für die Betreuung von ALFA auf dem Calar Alto und die Versorgung mit Ferien-Lodges und Olivenöl;

meine Zimmergenossen Sadegh "Diego" Khochfar, Chris "Krass" Wolf und Richie Idle, die sich immer sehr bemühten, mir eine positive Arbeitsumgebung zu schaffen;

die freundlichen Doktoranden Torsten Naab, Steffi Phleps und Olaf Kessel für Spaß bei Tisch und beim Weizenmann;

Jens Woitas mit dem ich irgendwann einmal einen Beobachtungsantrag durchbringen will und Rainer Köhler fürs gemeinsame Speckeln;

Carla für ihr Verständnis und ihre Hilfe in schwierigen Situation auch außerhalb des Berufs;

meine Eltern, die mir mein Studium ermöglichten, mir die nötige Freiheit liessen, und mir stets mit wertvollem Rat zur Seite standen.