ABSTRACT

LONG EXPOSURE POINT SPREAD FUNCTION ESTIMATION FROM SOLAR ADAPTIVE OPTICS LOOP DATA

by
Jose Marino

Adaptive optics (AO) systems provide partial correction to wavefront distortions introduced by the Earth's atmosphere. They have become an essential tool to obtain diffraction-limited observations from ground-based telescopes. However, the AO correction is only partial and post-processing with a good estimate of the point spread function (PSF) is required.

PSF estimates are impossible to measure directly during solar observations due to the lack of point sources in the field-of-view. Moreover, the highly variable day-time seeing conditions require the estimated PSF to be simultaneous with the captured image. A method is presented to estimate the long-exposure PSF of AO-corrected solar observations using the AO control loop data. The wavefront sensor and the deformable mirror data produced by the AO system during normal operation provide enough information to estimate the long-exposure PSF. Using this method, each individual AO-corrected image can be deconvolved with its own estimated PSF.

An attempt was made to verify the accuracy of the method by observing the star Sirius. The AO system successfully produced AO-corrected star images, which provide direct PSF measurements that can be compared to the estimated PSF. However, the poor performance of the AO system under low light levels, for which it was not designed, led to large uncertainties in the estimated PSFs.

The PSF estimation method was tested on real solar observations, where an estimation of the AO-corrected PSF is normally difficult. The observations were deconvolved with the estimated PSFs, producing significantly improved quantitative measurements and scientific data. A measurement of the performance of the solar AO during different seeing conditions was obtained for the first time from these observations.

LONG EXPOSURE POINT SPREAD FUNCTION ESTIMATION FROM SOLAR ADAPTIVE OPTICS LOOP DATA

by
Jose Marino

A Dissertation
Submitted to the Faculty of
New Jersey Institute of Technology and
Rutgers The State University of New Jersey - Newark
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Applied Physics

Federated Physics Department

May 2007

APPROVAL PAGE

LONG EXPOSURE POINT SPREAD FUNCTION ESTIMATION FROM SOLAR ADAPTIVE OPTICS LOOP DATA

Jose Marino

Dr. Thomas Rimmele, Dissertation Co-Advisor — Date
Astronomer, National Solar Observatory, New Mexico, USA

Dr. Haimin Wang, Dissertation Co-Advisor — Date
Distinguished Professor of Physics, New Jersey Institute of Technology

---

Dr. Philip R. Goode, Committee Member — Date
Distinguished Professor of Physics, New Jersey Institute of Technology

---

Dr. Carsten Denker, Committee Member — Date
Assistant Professor of Physics, New Jersey Institute of Technology

---

Dr. Zhen Wu, Committee Member — Date
Professor of Physics, Rutgers

BIOGRAPHICAL SKETCH

Author: Jose Marino

Degree: Doctor of Philosophy

Date: May 2007

Undergraduate and Graduate Education:

- Doctor of Philosophy in Applied Physics,
  New Jersey Institute of Technology, Newark, New Jersey, 2007
- Master of Science in Opto-Electronics and Optical Information Processing,
  Queen's University of Belfast, Belfast, UK, 2000
- Degree in Physics,
  University of Barcelona, Barcelona, Spain, 1999

Major: Applied Physics

Presentations and Publications:

T. Rimmele and J. Marino, "The Evershed Flow: Flow Geometry and Its Temporal Evolution", Ap. J. **646**, 593-604 (2006).

J. Marino, T. Rimmele, and J. Christou, "Long exposure point spread function estimation from adaptive optics loop data: validation and results", in *Proceedings of SPIE* **6272**, 62723W (2006).

J. Marino, T. Rimmele, and J. Christou, "Long Exposure Point Spread Function Estimation from Adaptive Optics Loop Data", in *Proceedings of AGU Spring Meeting* (2005).

J. Marino, T. R. Rimmele, and J. C. Christou, "Long-exposure point spread function estimation from adaptive optics loop data", in *Proceedings of SPIE* **5490**, 184-194 (2004).

L. Jolissaint, J.-P. Véran, and J. Marino, "OPERA, an automatic PSF reconstruction software for Shack-Hartmann AO systems: application to Altair", in *Proceedings of SPIE* **5490**, 151-163 (2004).

T. R. Rimmele, K. Richards, S. Hegwer, et al., "First results from the NSO/NJIT solar adaptive optics system", in *Proceedings of SPIE* **5171**, 179-186 (2004).

T. R. Rimmele, K. Richards, S. L. Hegwer, et al., "Solar adaptive optics: a progress report", in *Proceedings of SPIE* **4839**, 635-646 (2003).

J. Marino, T. Rimmele, and E. Tatulli, "Long Exposure Point Spread Function Estimation using Wavefront Sensor Data", in *Current Theoretical Models and Future High Resolution Solar Observations: Preparing for ATST, ASP Conference Series* **286**, 69-76 (2003).

This dissertation is dedicated
to Leah Simon

ACKNOWLEDGMENT

I would like to thank the members of my thesis committee Dr. Thomas Rimmele, Dr. Haimin Wang, Dr. Carsten Denker, Dr. Phillip Goode and Dr. Zhen Wu, and everybody at the Center for Solar-Terrestrial Research at the New Jersey Institute of Technology.

I am very grateful to the entire NSO staff in Sunspot, New Mexico. The AO system engineer Kit Richards and software engineer Chris Berst who helped me implement the modifications to the AO system and camera system required by my work. Thanks to the telescope operators Doug Gilliam, Joe Elrod and Mike Bradford for assisting me during my observing runs. And thanks to Ramona Elrod, Rebecca Coleman and Jackie Diehl for always being there when I needed help.

Thanks to the scientific staff in Sunspot for teaching me so much: Dr. Steve Keil, Dr. Han Uitenbroek, Dr. K.S. Balasubramaniam, Dr. Richard Radick, Dr. Alexei Pevtsov, Dr. Alexandra Tritschler and Dr. Friedrich Wöger.

Also I would like to thank Craig Gullixson and the rest of the volleyball team for all those nice and sometimes intense volleyball games. And, in general, all the people living in Sunspot for always treating me just like another member of their families.

And I would also like to give special thanks to my supervisor Dr. Thomas Rimmele. My sincere gratitude goes to him and to his wonderful family: Elli, Lukas and Simon, who always welcomed me in their home and kept me well fed throughout all these years living in Sunspot.

This work was supported by NSF grant AST-0079482.

TABLE OF CONTENTS

TABLE OF CONTENTS
(Continued)

Chapter **Page**

TABLE OF CONTENTS
(Continued)

Chapter **Page**

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES
(Continued)

Figure **Page**

LIST OF FIGURES
(Continued)

Figure **Page**

LIST OF FIGURES
(Continued)

Figure **Page**

LIST OF FIGURES
(Continued)

Figure **Page**

LIST OF FIGURES
(Continued)

Figure **Page**

LIST OF FIGURES
(Continued)

Figure **Page**

LIST OF FIGURES
(Continued)

Figure **Page**

CHAPTER 1

INTRODUCTION

1.1 Overview

The Earth's atmosphere plays an essential role in ground-based astronomical observations [33, 67, 97]. Its effects can severely distort the light coming from astronomical objects, even though it constitutes just a minuscule layer when compared to the vast distances traveled by this light. The obvious solution to this problem is to capture the light before it enters the atmosphere by building orbital observatories or placing them on high altitude balloons. However, the technological difficulties, inflexible setups and prohibitive costs involved in these approaches make ground-based astronomy a much more feasible alternative.

Adaptive optics (AO) is a very successful technique that attempts to eliminate the distortions introduced by the Earth's atmosphere in real-time. It has become a key technology for ground-based astronomical observations, greatly improving the quality and resolution of the observations [61]. However, due to the physical and technological limitations inherent to any AO system, residual distortions remain after correction and impact the quality of the captured images skewing any quantitative measurements produced by the observations [33].

The correction provided by the AO system, while not complete, is sufficient to preserve high spatial frequency information that would have been otherwise lost [33, 100]. This allows further restoration of the image by using post-processing techniques, such as deconvolution with a known point spread function (PSF). The PSF characterizes the performance of the optical system and provides a measure of the image quality obtained [33, 101]. With a known AO-corrected PSF, the amplitudes of the spatial frequency components of

the AO-corrected image can be restored to their diffraction-limited levels improving image quality and quantitative measurements.

The latest developments in solar physics require high-resolution observations that produce accurate quantitative measurements (see Section 1.3). AO correction plus post-processing with an estimation of the AO-corrected PSF are vital to produce accurate scientific measurements. In night-time observations, the PSF can be measured directly from the image of a star. However, a direct measurement of the PSF from solar observations is impossible due to the lack of point sources in the field-of-view. Moreover, the high variability of the seeing conditions, in particular during the day, require the PSF to be estimated simultaneously while the image is being captured. The data generated internally by the AO system during normal operation contain enough information to produce an estimate of the long-exposure PSF. Since the AO data are captured while the image is being exposed, the estimated PSF characterizes the actual conditions that affected the image. A long-exposure is defined as an exposure that includes a large number of configurations of atmospheric distortions, which are required to produce good statistical measurements.

A method to obtain an estimation of the long-exposure PSF from the AO telemetry data has been developed and implemented. The method uses the control data produced by a solar AO system during normal correction to compute an estimate of the long-exposure PSF. The results of applying the method to solar and stellar observations are shown. The accuracy of the PSFs estimated from the AO telemetry data must be quantified by comparison to PSFs obtained through an independent approach. AO-corrected observations of an unresolved star provided such an independent estimation of the PSF and proved to be a valuable tool to fine tune the performance of the AO system.

Chapter 1 provides an overview of the issues and motivation for this work.

A complete understanding of the AO system and its components is required by the PSF estimation method. An introduction to AO systems is presented in Chapter 2. This

chapter describes the technical characteristics of the high-order solar adaptive optics system at the Dunn Solar Telescope (DST).

The PSF estimation method makes some assumptions about the properties of atmospheric turbulence. Chapter 3 discusses the effects and properties of atmospheric turbulence. The Kolmogorov model is used to derive the statistical properties of wavefront distortions induced by atmospheric turbulence.

Chapter 4 presents a method to estimate the AO-corrected long-exposure PSF from internal AO telemetry data. The method was developed and implemented to be used for AO-corrected solar observations.

The results produced by the observations of the star Sirius are discussed in Chapter 5.

Chapter 6 discusses the results obtained from applying the PSF estimation method to solar observations.

The scientific results obtained from applying the PSF estimation method to a time series of velocity measurements of a sunspot are discussed in Chapter 7.

1.2 Historical Overview

There is significant archaeological evidence that suggests that mankind has been intrigued by the stars, the Sun and other astronomical bodies since ancient times. The origin of Astronomy as a science can be traced back to ancient Babylon before the year 600 B.C.E. [33]. Babylonian astronomers observed and kept detailed records of the motions of the Sun, the Moon and the planets. This legacy was passed on to the Greeks who greatly advanced astronomy through many theoretical and observational breakthroughs. However, the effects of the Earth's atmosphere on astronomical observations were almost entirely unknown during this time.

The first reference to atmospheric refraction is credited to Cleomedes during the first century A.D. While observing an eclipse, Cleomedes deduced that the Earth's atmosphere bends the light rays traveling through it [33]. The first recorded account of the distortions

introduced by the atmosphere on astronomical observations dates back to the end of the 10th century. Arabian physicist Ibn al-Haitham described the shape changes of the Moon and the Sun as they move closer to the horizon, and how atmospheric refraction alters the apparent positions of the stars [33].

Many ingenious astronomical instruments were used to measure the movement and positions of astronomical bodies, but ultimately all observations were performed with the naked eye. The full severity of atmospheric distortions is only realized when using optical systems with aperture sizes larger than a certain size defined by the turbulence, which typically is on the order of a few tens of centimeters. The small pupil size of the eye, of just a few millimeters, limits the nature of the atmospheric effects that can be observed. The full effects of atmospheric turbulence, commonly referred to as *seeing effects*, became apparent with the introduction of the telescope by Galileo in the early 17th century [33].

The impact of atmospheric turbulence on images captured with a telescope was first recorded by Isaac Newton [51, 97] in his book *Opticks* in 1730. The twinkling of the stars at night and the loss of resolution when looking through telescopes of a certain size were already well known effects at the time. Newton correctly attributed these effects to the propagation of light through the Earth's atmosphere: "For the air through which we look upon the Stars, is in perpetual Tremor" [51, 69]. He realized how these effects would always be present regardless of the optical quality of the telescope used. And he went as far as to suggest that the only solution to alleviate them would be to locate telescopes on top of high mountains. This was the accepted view for the next 200 years. Seeing effects were considered as unavoidable and observatories were built on top of high mountains in an attempt to minimize them.

One of the first breakthroughs came during the 1950s when film technology became advanced enough to capture short-exposures, on the order of milliseconds, of bright objects. These short-exposures are called *speckle images* due to the speckled pattern they present when imaging point sources. The exposure time of a speckle image is comparable to the

evolution time of atmospheric turbulence. Hence, a speckle image can be assumed to be affected by a *frozen* configuration of the atmosphere, which introduces a particular set of static distortions. These distortions affect some of the spatial frequency components of the image while leaving others undistorted. Different speckle images captured at slightly different times contain unique undistorted spatial frequency components. By considering several speckle images, the undistorted components from each one can be merged to form a reconstruction of an aberration free image. This powerful post-facto technique is called *speckle reconstruction* and is still used today [18].

Around the same time additional methods for reducing seeing effects began to develop. In 1953, Babcock [2] proposed using an adaptive optical element and a wavefront sensor to compensate for atmospheric distortions in real-time. This technique is nowadays called *adaptive optics*. However, it could not be fully implemented at the time due to technological limitations.

Over time, technological advances gradually allowed for the implementation of AO. Much successful work on AO was pioneered by the United States Air Force, mostly as a tool to improve imaging of artificial satellites from Earth-based telescopes. This work was classified by the Air Force and was unavailable to the astronomical community until its declassification in 1992. During the 1980s, several scientific organizations were already working on the development of adaptive optics for astronomical observations, such as the National Optical Astronomy Observatory (NOAO), the European Southern Observatory (ESO) and the Office National d'Etudes et de Reserches Aerospatiales (ONERA) in France.

1.3 Scientific Motivation

Many basic processes on the Sun take place on small scales ($<1''$). Observations from the ground and from space missions such as SoHO, Yohkoh and TRACE have revealed that the Sun's atmosphere is highly dynamic. Most solar phenomena and nearly all of the variability of the Sun are a direct result of the dynamic magnetic field continuously emerging

from the solar interior on scales from granulation to active regions. The Sun's magnetic field produces coronal mass ejections, flares, and the solar wind. Unfortunately, the understanding of these phenomena is still very limited. For example, it is not certain what causes solar variability and the solar cycle; there is no complete explanation for the existence of sunspots; why does the Sun have a corona or what triggers a coronal mass ejection and what mechanisms are responsible for variations in the spectral and total irradiance of the Sun.

According to Cattaneo [90] any three-dimensional solar turbulent flow is likely to behave as a dynamo. As a consequence, each scale of turbulent motion will produce its own "magnetic scale" and magnetic fields should be observed virtually everywhere on the Sun. Since the magnetic fields are "frozen in" to the plasma, granular flows will concentrate the magnetic structures into the intergranular lanes. It is likely that several dynamos are at work on the Sun and other stars: a global dynamo and a turbulent dynamo. The local dynamo action may produce the magnetic flux tubes, organized as a magnetic carpet, recently observed to cover the entire Sun. Understanding how large and small scale structures behave may play an important role in the understanding of the physics of the Sun.

Numerical simulations have helped tremendously to improve the understanding of the nature of convection. However, in many cases observations of sufficient resolution to verify model predictions are still missing. Progress in answering such critical questions requires studying the interaction between the magnetic field and convection with sufficient resolution to observe scales fundamental to these processes. Thus, it is very important to eliminate seeing effects to achieve diffraction-limited observations that will lead to a deeper understanding of solar processes.

1.4 Numerical Simulations

To illustrate the effects of seeing and the impact of AO in a quantitative manner, numerical simulations of solar granulation are used. The simulated object is obtained from a

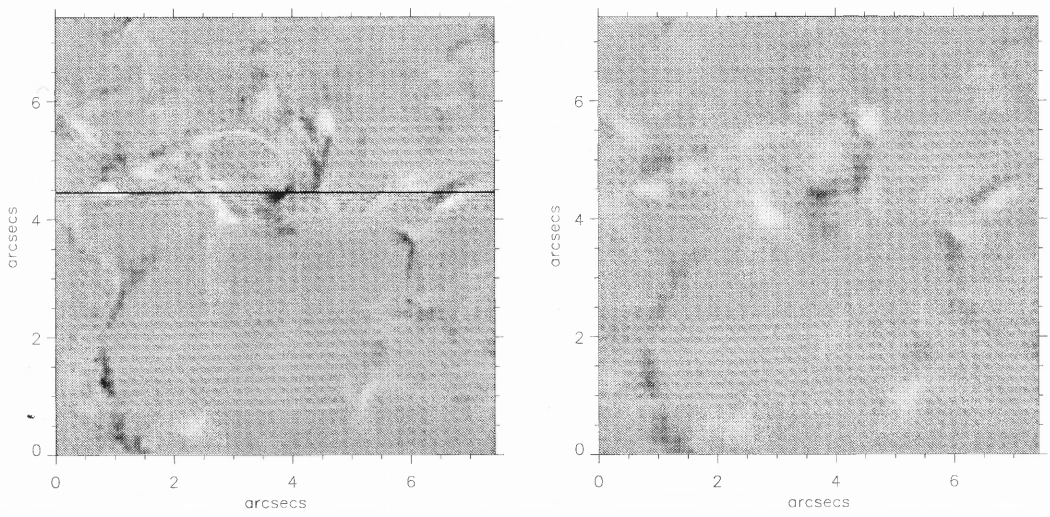

Figure 1.1 Simulated Stokes V profile "observed" with two telescopes of different aperture sizes (left: 2 m and right: 76 cm). Data provided by Keller et al. [40]. The dark line on the left image indicates the location where a cross-section was obtained.

magneto-hydrodynamic (MHD) simulation of solar granulation provided by Keller, Stein and Nordlund [40]. The simulation provides an intensity image of solar granulation and its corresponding Q, U, and V Stokes profiles that have not been altered by any imaging system.

A fundamental limitation of any optical system is diffraction. Diffraction of light on the aperture of the telescope limits the maximum resolution achievable by the telescope. The images that two telescopes of 2 m and 76 cm aperture size would produce when imaging the simulated Stokes V profile are shown in Figure 1.1.

The image captured by the 2 m telescope appears much sharper than the image captured by the 76 cm telescope. A 2 m telescope can achieve a spatial resolution more than two times higher than the 76 cm telescope. The increase in spatial resolution is clearly illustrated in Figure 1.2, which plots a cross-section of the Stokes V profiles captured with both telescopes.

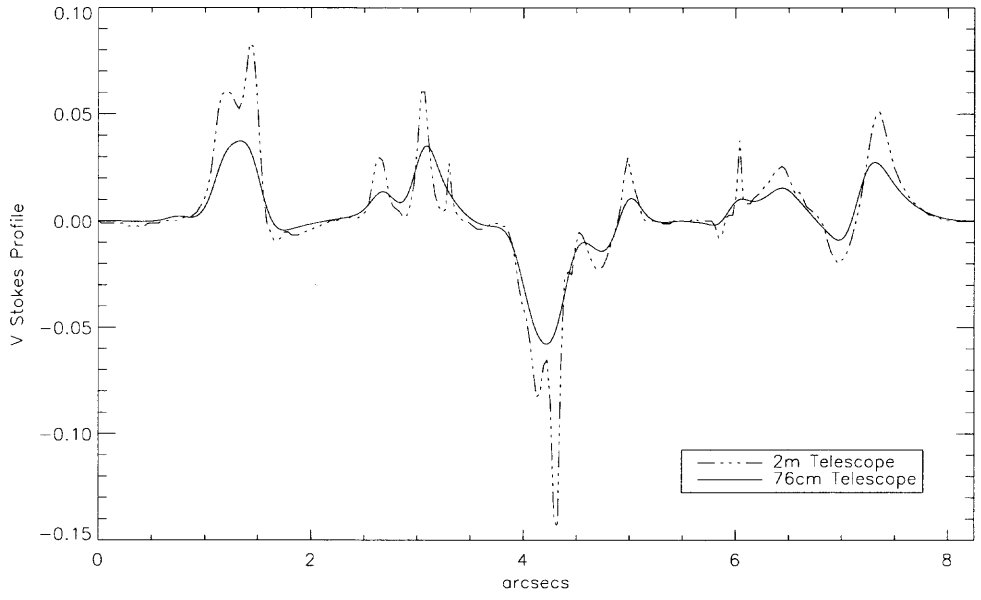

Figure 1.2 Comparison of a cross-section from the simulated Stokes V profile "observed" by two telescopes with different aperture sizes, shown in Figure 1.1.

1.5 Adaptive Optics Correction

During regular ground-based observations, the effects of atmospheric turbulence have a much more damaging effect on the spatial resolution of the captured images. While the aperture size sets an upper limit to the spatial resolution achievable, the actual resolution achieved is determined by the aberrations introduced by atmospheric turbulence. Atmospheric distortions effectively reduce the spatial resolution of an image to that of an image captured by a telescope with an aperture size of just a few tens of centimeters.

The development of AO made the partial correction of atmospheric distortions in real-time possible. AO correction aims to preserve the spatial resolution of the images up to the diffraction limit. However, even the most sophisticated AO systems are not able to deliver perfect correction and some residual aberrations always remain. These residual aberrations introduce distortions on the quantitative measurements obtained from the AO-corrected images, which compromise the scientific information that can be extracted from

an AO-corrected image. Such images could easily produce inaccurate or even meaningless results.

The benefits and limitations of the AO correction on scientific measurements can again be illustrated with the help of the MHD simulations of solar granulation described before. The simulated data were degraded to simulate the images produced by a telescope with an aperture size of 76 cm under three different conditions: aberration free, distorted by medium seeing with no AO correction and with AO correction. The images are shown in Figure 1.3.

The scientific measurements obtained from the seeing distorted images, both AO-corrected and uncorrected, can be compared to the results that would be obtained in the aberration free case, i.e., the diffraction-limited case. The top row of Figure 1.3 shows the aberration free, i.e., diffraction-limited, granulation intensity image and the corresponding stokes V profile, while the middle and bottom rows show the uncorrected and AO-corrected seeing distorted images, respectively. The images from Figure 1.3 illustrate the importance of AO correction in long-exposure images.

The residual aberrations left after AO correction reduce the accuracy of any quantitative measurements taken from the image. Figure 1.4 compares cross-sections taken from the Stokes V profiles from Figure 1.3 and illustrates the drastic improvement provided by the partial AO correction. The aberration free and AO-corrected cross-sections indicate that both images present more or less the same resolution, but the quantitative information is quite different. AO correction by itself is not sufficient to provide reliable quantitative data that accounts for all the physics that should be observed. However, AO is essential to preserve diffraction-limited information in the image, which is lost beyond recovery without correction. Post-processing reconstruction of the AO-corrected images with a good estimate of the AO-corrected PSF further restores image quality and increases the accuracy of any quantitative measurements.

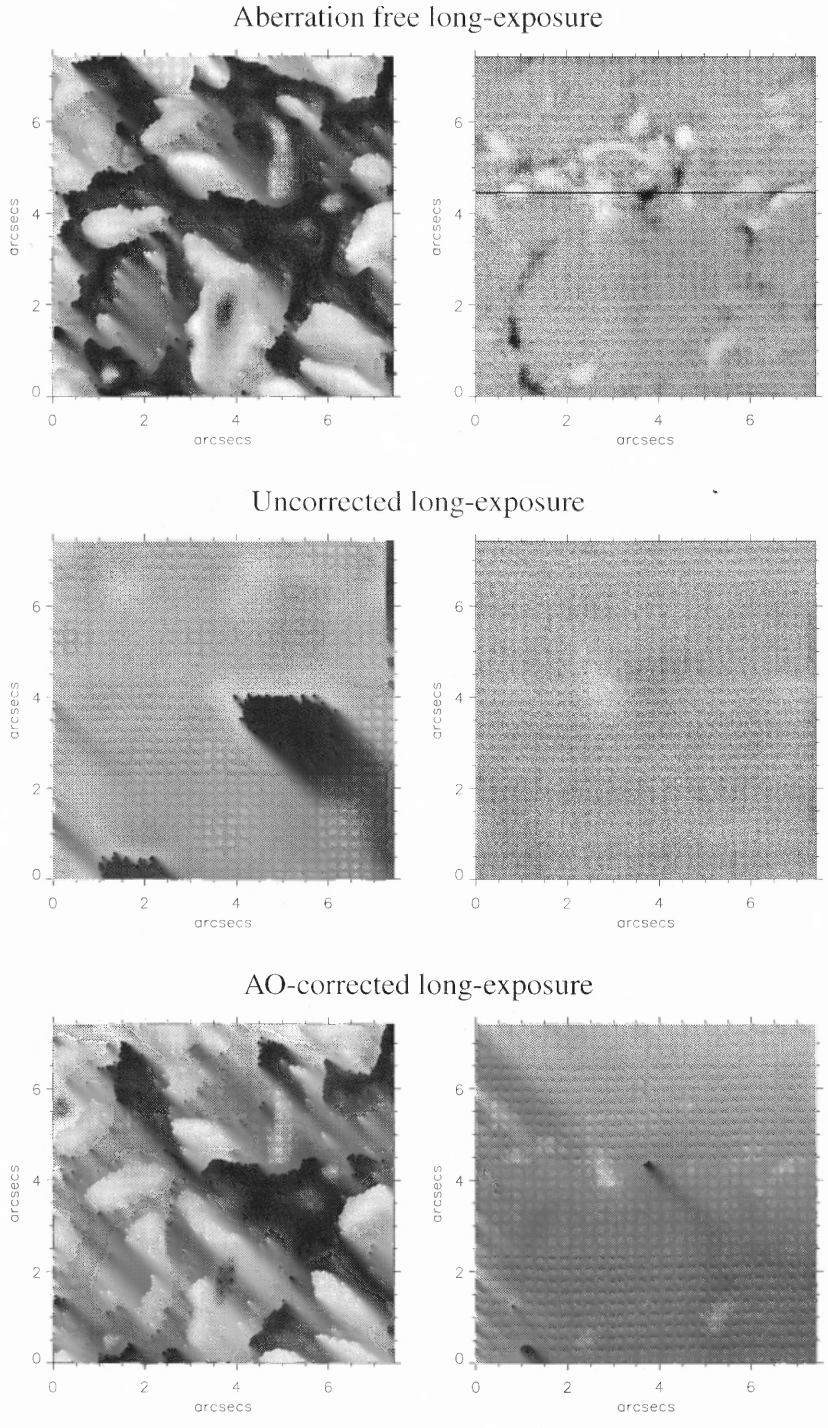

Figure 1.3 Granulation intensity and its Stokes V profile obtained from a simulation by Keller et al. [40]. The effects of the telescope aperture, atmospheric seeing and AO correction are simulated. Images are scaled individually. The dark line on the top-left image indicates the location where a cross-section was obtained.

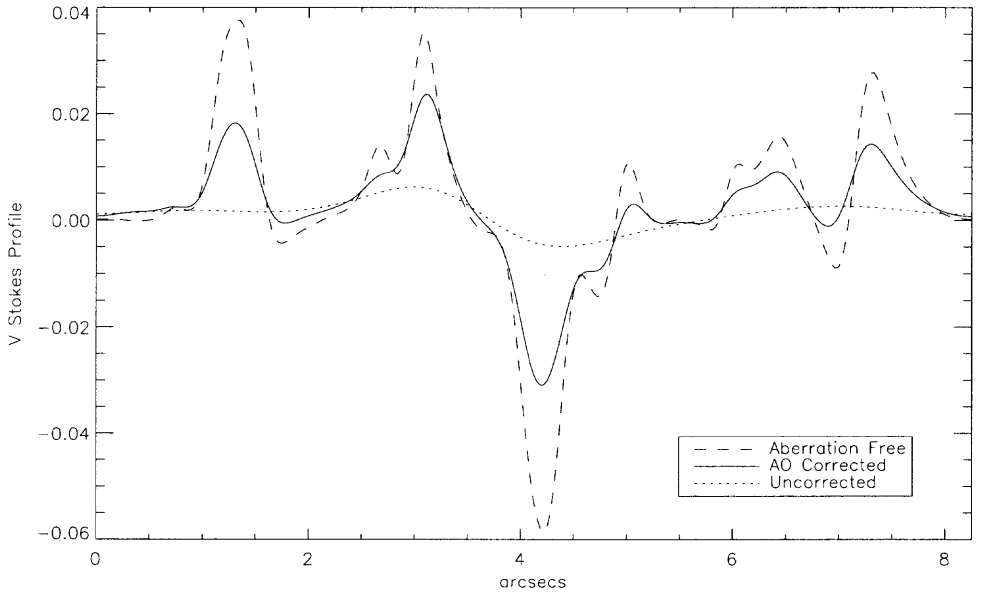

Figure 1.4 Stokes V profile cross-section for the cases: aberration free (dashed line) and affected by atmospheric distortions with (solid line) and without (dotted line) AO correction (telescope aperture: 76 cm, r_0: 10 cm).

1.6 Point Spread Function Estimation

The PSF completely characterizes the optical quality of an optical system. The frequency response of the optical system is given by the Fourier transform of the PSF, called the optical transfer function (OTF). The OTF is defined in the Fourier space and it is in general a complex quantity. The modulus of the OTF is known as the modulation transfer function (MTF) and its phase is known as the phase transfer function (PTF). In the case of a long-exposure with incoherent illumination, the OTF is a real function [33] and therefore equal to the MTF. Figure 1.5 shows the azimuthal average of theoretical [33] PSFs and OTFs computed for the cases of: aberration free, i.e., diffraction-limited, and atmospheric distortions with and without AO correction.

The resolution and image quality produced by an optical system are completely characterized by its PSF. An ideal optical system would be a system that images a point from an object as a point in the image, producing an identical representation of the object. Its

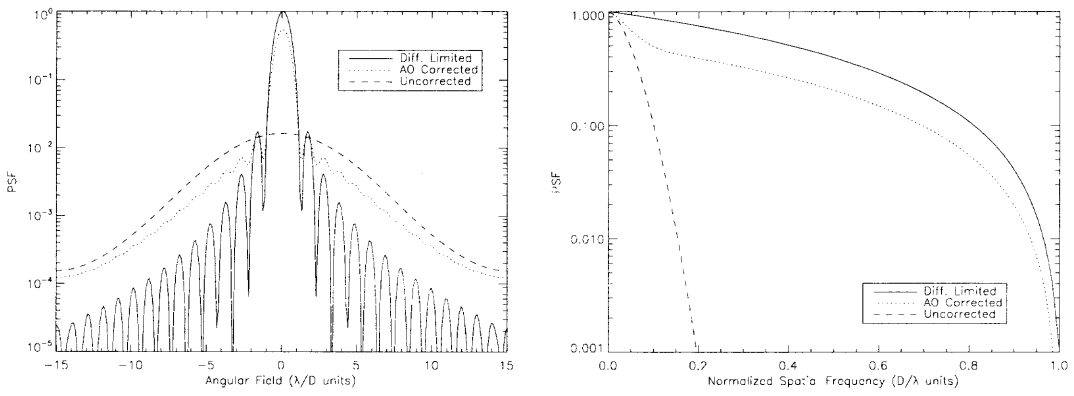

Figure 1.5 Azimuthal average of simulated PSFs and OTFs for the cases of: diffraction-limited (solid line), atmospheric distortions (dashed line) with and without AO correction (dotted line).

PSF would be described as an infinitesimal point. However, in reality, light diffracted by the entrance pupil of an optical system produces a PSF with a certain width determined by the size of the aperture. Thus, in a real optical system, the image of a point is never a point but the diffraction pattern produced by the entrance pupil. The width of the diffraction pattern sets a fundamental limit to the maximum resolution of the optical system. The diffraction-limited PSF produced by a perfect circular aperture is known as the Airy function and consists of a high narrow central peak, the diffraction-limited core, followed by diffraction rings, as illustrated in the left panel of Figure 1.5.

The width of the uncorrected PSF is determined by the seeing conditions and it is much wider than the diffraction-limited PSF. The PSF energy is spread over a wide area, known as the seeing halo, which translates into a much lower image resolution.

The AO-corrected PSF presents a high diffraction-limited core superimposed over a wide seeing halo [14, 20, 33]. This characteristic shape is produced by the partial AO correction, which restores most of the PSF energy from the seeing halo into the diffraction-limited core. Since the correction is not perfect, some of the energy is still left in the seeing halo. With adequate AO correction, the central core of the AO-corrected PSF should

have the same width as the diffraction-limited PSF core, therefore AO correction preserves diffraction-limited information in the image.

The degree of correction provided by the AO system is usually quantified by the Strehl ratio. It is defined as the ratio of the intensity contained in the diffraction-limited core of the PSF versus the intensity contained in the seeing halo [33, 67]. Or alternatively, as the ratio of the peak intensity of the PSF versus the peak intensity of the aberration free (diffraction-limited) PSF. Thus, the Strehl ratio of an aberration free image is one. The Strehl ratio can be calculated from the variance of the wavefront phase σ_{φ}^{2} with the following approximation:

$$S \approx e^{-\sigma_{\varphi}^{2}} \tag{1.1}$$

This relation is known as the Marechal approximation [33, 67] and it is an accurate approximation for small wavefront phase errors, up to a value of $\sigma_{\varphi} \sim 2$ rad.

The frequency response of these PSFs is illustrated by the OTFs shown in the right panel of Figure 1.5. The high spatial frequency information is preserved by the AO correction, but the amplitude of this spatial frequency information is not at the same level as in the diffraction-limited case. For the uncorrected case, most of this high spatial frequency information is lost in the background noise, making its recovery impossible. Thus, AO correction preserves high spatial frequency information in the image, up to the diffraction-limited cut-off frequency, which can be further restored with post-processing techniques.

An estimation of the PSF that affected an exposure is critical for both night-time and solar astronomy. During night-time observations, an estimation of the PSF can be obtained from a star, which can be considered as a point source. An image of such a point-like object provides a direct measurement of the PSF. However, most astronomical science images do not have any suitable point source in the field-of-view that could provide direct PSF information. A common solution to this problem during night-time observations is

to routinely point the telescope away from the region of scientific interest to capture an image of a star. This has the inconvenience that valuable observing time is lost during the capture of the PSF. Moreover, seeing conditions can change during this time, producing an inaccurate PSF.

In the case of solar astronomy, there are no point sources in the field-of-view. Furthermore, the variable seeing conditions during the day may cause two images taken at different times to be affected by different seeing conditions, which requires a separate PSF to be estimated for each individual image. This is especially important when those images are combined to generate some indirect quantitative measurements. For example, magnetic field maps (magnetograms) can be constructed by subtracting images produced by left and right circular polarized light (Zeeman effect). Line of sight velocity maps (or dopplergrams) can be obtained by subtracting images formed by light from the blue and red wings of a spectral line (Doppler effect). These images are usually not taken simultaneously and variations in the PSF between them result in spurious magnetic or velocity signals.

Many science applications require time sequences of consistent high-resolution images or spectra in order to study the highly dynamic solar atmosphere. The high variability of day time seeing conditions, which typically change on time scales similar to the exposure times (seconds), cause large variations in the PSF from image to image and make the interpretation of the time sequence very difficult.

The data produced internally by the AO system while correcting long-exposures contain enough information to produce an estimation of the long-exposure PSF. By using the AO telemetry data no extra observing time is required and the estimated PSF characterizes the exact conditions present during the image acquisition. Such a method has been successfully developed for night-time observations using curvature sensing AO systems by Véran et al. [101]. This method is adapted to a solar Shack-Hartmann based AO system and applied to solar observations.

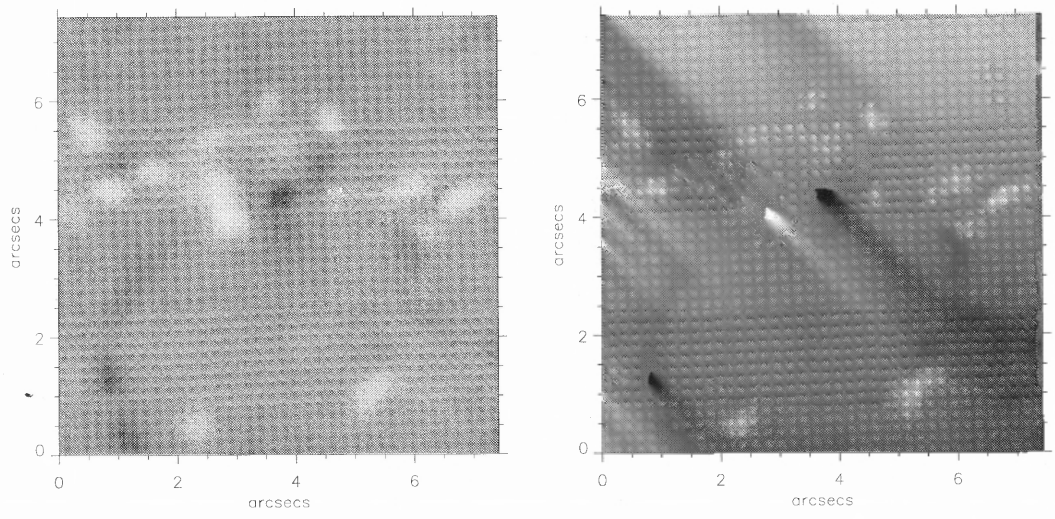

Figure 1.6 AO-corrected Stokes V profiles for different seeing conditions (Strehl ratios of 0.09 and 0.5). Data provided by Keller et al. [40].

1.7 Limitations

The estimated AO-corrected PSF can be used to efficiently increase the science output of AO-corrected images by applying post-facto image reconstruction techniques. However, some minimum seeing conditions are required to produce good reconstructions. The level of correction that the AO system can provide depends critically on the severity of the seeing conditions, characterized by the Fried parameter r_0.

Figure 1.6 shows two AO-corrected Stokes V magnetograms affected by very different seeing conditions. The left and right AO-corrected images present Strehl ratios of 0.09 and 0.5, respectively. The AO-corrected image with a Strehl ratio of 0.09 represents the case of marginal AO correction, where the seeing conditions were so severe that most of the energy of the PSF ($> 90\%$) remains in the seeing halo. The AO-corrected image with a Strehl ratio of 0.5 contains 50% of the PSF energy in the diffraction-limited core, which constitutes a good AO correction.

Figure 1.7 shows the correlation between the pixel intensity of the AO-corrected Stokes V profile and the diffraction limited Stokes V profile. Left and right panels correspond to the AO-corrected images with Strehl ratios of 0.09 and 0.5 shown in Figure 1.6,

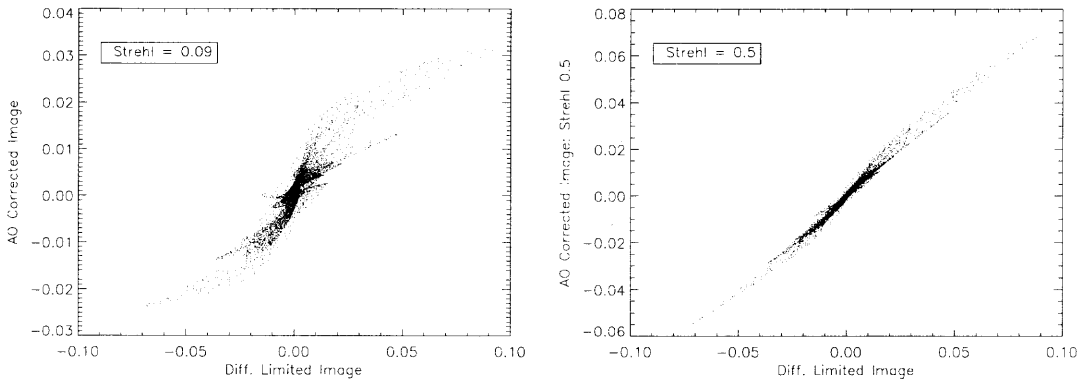

Figure 1.7 Correlation between a diffraction-limited Stokes V profile and two AO-corrected profiles captured during different seeing conditions (Strehl ratios of 0.09 and 0.5). The correlation coefficients are r=0.90 and r=0.99 for the left and right panels, respectively.

respectively. A perfect linear correlation would indicate that the diffraction-limited information is perfectly preserved in the AO-corrected image.

The data shown in the left panel of Figure 1.7 presents a correlation coefficient of r=0.90. This indicates that some of the diffraction limited information is lost in the AO-corrected Stokes V profile with a Strehl ratio of 0.09. In the case of the AO-corrected Stokes V profile with a Strehl ratio of 0.5, the correlation coefficient is r=0.99 indicating a strong correlation between the AO-corrected profile and the diffraction limited profile. Thus, the diffraction limited information contained in an AO-corrected image depends on the Strehl ratio.

Figure 6.8 illustrates the dependency of the AO-corrected Strehl ratio on seeing conditions. The performance of the AO system is characterized by the Strehl ratio of the AO-corrected image. The figure indicates how the quality of the correction increases with good seeing conditions. However, it never reaches a value of one, which indicates perfect diffraction-limited performance, even for very good seeing conditions. The figure clearly illustrates the need for deconvolution with a good PSF estimation even for the case of very good seeing conditions.

The effects that seeing conditions have on AO performance are very significant. A good understanding of the properties and effects of atmospheric turbulence is necessary to quantify and optimize the performance of the AO system. Chapter 3 describes the characteristics of atmospheric turbulence and its effects on the light propagating through the atmosphere.

The correction the AO system is able to provide also depends on the characteristics of the system components. Their technical specifications place a limit on the maximum correction the system can provide. A description of the AO system, its components, and the error sources that affect the correction, are described in Chapter 2.

The PSF estimation method requires knowledge about the residuals not corrected by the AO system and about the original seeing conditions to estimate the AO-corrected PSF. This information is extracted from the AO telemetry data considering the characteristics of the components of the AO system, and the properties of atmospheric turbulence. A description of the PSF estimation method from AO telemetry data is given in Chapter 4.

CHAPTER 2

THE ADAPTIVE OPTICS SYSTEM

The term *adaptive optics* refers to a group of techniques that attempt to correct optical aberrations in real-time. Any AO system consists of a wavefront sensing element and a correcting element. The correction is performed by means of an optical adaptive component, which can alter its optical properties to apply phase changes to the wavefront in real-time. The sensing element measures the wavefront after correction to obtain information about the residuals. The state of the adaptive optical element is then updated in order to minimize the measured residuals.

AO systems are used for many applications, such as: retinal imaging, high power laser applications, point to point optical communications and ground-based astronomical and solar observations. All of these applications share a common element: a turbulent medium that introduces distortions in the light propagating through it. The characteristics and design of an AO system are dictated by the requirements of each particular application. For example, the distortions produced by the vitreous humor, the fluid inside the eye, change very slowly with time and do not require an AO system with a fast loop frequency. These systems usually work at frequencies of just a few Hz. In contrast, the distortions produced by atmospheric turbulence evolve very rapidly, on the order of milliseconds. AO systems designed for ground-based astronomical applications require fast loop frequencies to successfully correct the atmospheric distortions. Loop frequencies of current astronomical AO systems range from hundreds of Hz to a few kHz.

The design requirements of AO systems for night-time and solar astronomical observations are different. Night-time observations rely on the light coming from a bright star to perform wavefront sensing. Such a star is commonly referred to as a natural guide star (NGS)[67]. Occasionally the star used for wavefront sensing is very faint, which forces the

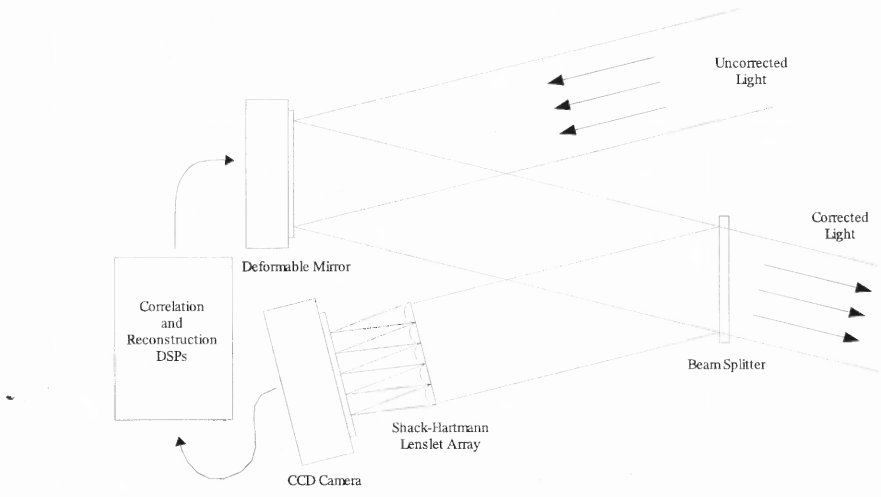

Figure 2.1 Schematic description of the individual subsystems of an AO system.

AO system to run slower to obtain wavefront measurements with a sufficient signal-to-noise ratio (SNR). Thus, the apparent magnitude of the chosen NGS introduces a fundamental limit on the loop frequency of the system. Usually, there are no bright stars close enough to the object of interest. In these cases, a laser is used to generate an artificial star, called a laser guide star (LGS).

During solar observations, there are no point sources in the field of view. Solar AO systems use structures that form on the surface of the Sun to perform wavefront sensing. The light levels are always high enough [61] and do not constitute a limiting factor. Hence, the loop frequency of a solar AO system is not limited by light input, it is only limited by the signal processing power. However, the contrast of the particular structure being used for wavefront sensing and the seeing conditions limit the sensitivity of the measuring process and the quality of the correction. During good seeing conditions, solar granulation can be used to produce reliable wavefront measurements, whereas during worse seeing conditions, larger structures with more contrast, such as sunspots and pores, must be used.

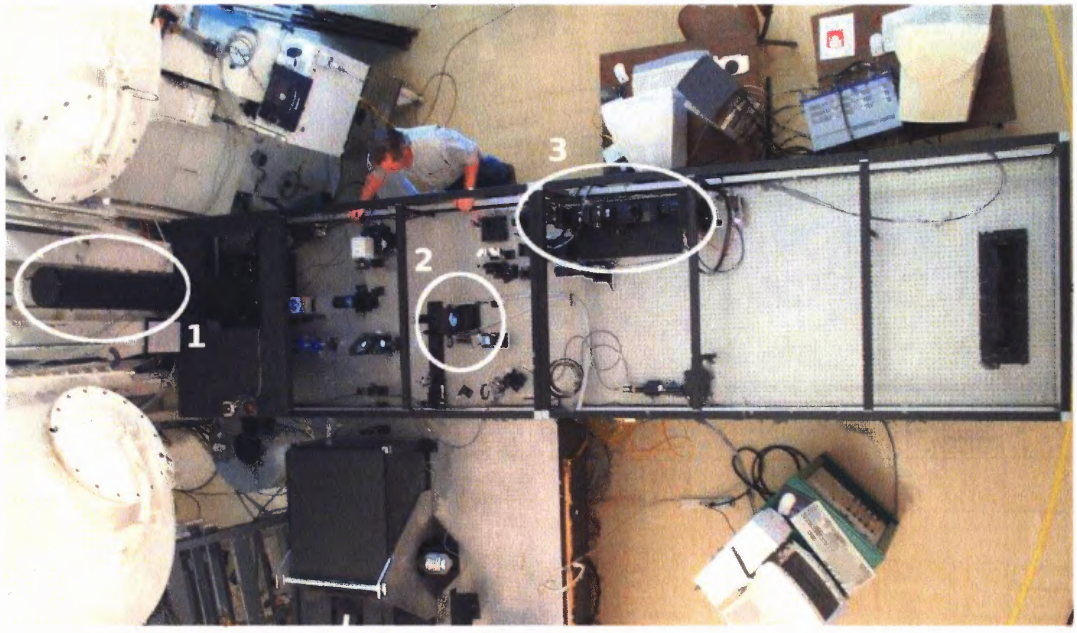

Figure 2.2 Top view of the AO system at the Dunn Solar Telescope. The picture shows the location of the tip-tilt mirror (1), the deformable mirror (2) and the wavefront sensor (3).

2.1 Solar Adaptive Optics

The DST (DST) at Sacramento Peak (Sunspot, New Mexico) is equipped with a high-order solar AO system (Figure 2.2), which is routinely used in observations. It is one of two twin AO systems developed at NSO [56]. The second AO system is located at Big Bear Solar Observatory in Big Bear City, California.

Figure 2.3 shows images captured during a test of the high-order AO system at the DST in April 2003. The wavefront sensing system consists of a 76 subaperture cross-correlating Shack-Hartmann wavefront sensor (WFS). The correction is applied by a 97 actuator continuous face plate deformable mirror (DM) manufactured by Xinetics. The control system that updates the shape of the DM from the WFS measurements is based on off-the-shelf digital signal processors (DSPs). The system runs at a frequency of 2.5 kHz achieving a correcting bandwidth, defined as the 0 dB cross over of the error rejection curve, of approximately $\sim$ 130 Hz. The layout of the DST AO system is illustrated in Figure 2.2.

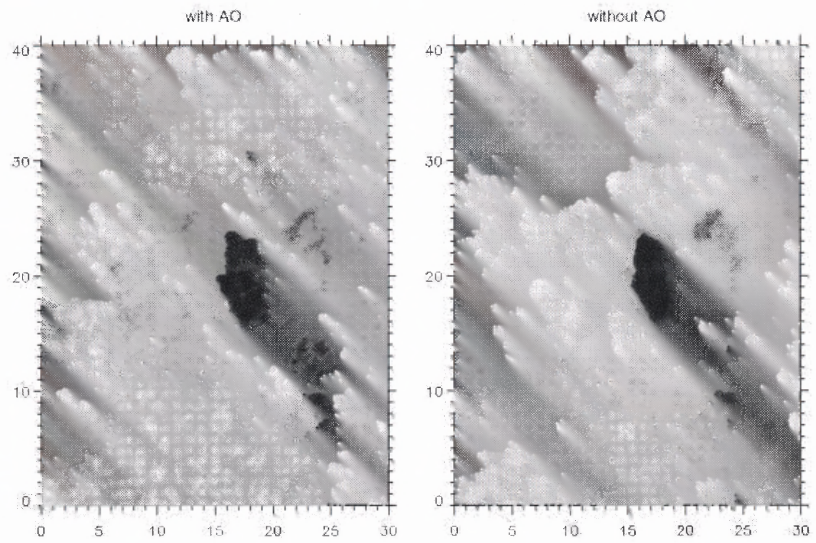

Figure 2.3 Images taken during the first light of the high order AO system at the Dunn Solar Telescope in 2003. The left image is AO corrected while the right image is uncorrected (the tick marks are given in arc seconds).

2.1.1 Cross-Correlating Shack-Hartmann Wavefront Sensor

A small fraction of the light after reflection on the DM is used for wavefront sensing. A beam splitter separates 5% of the light and sends it to the WFS while the other 95% continues to the science path to be used for the observations. The light used for wavefront sensing is filtered by a wide-band filter located in front of the WFS to reduce the bandwidth of the light used for wavefront sensing.

The WFS of the AO system is a cross-correlating Shack-Hartmann WFS [61]. It measures the first derivative of the wavefront at several points in the pupil. An array of small lenslets, placed in a pupil plane, sample the wavefront at the entrance pupil of the telescope, dividing it into small subpupils or subapertures. Each lenslet images its portion of the wavefront, producing a subaperture image, which is shifted proportionally to the mean wavefront slope inside that subaperture, as shown schematically in Figure 2.4. A map of the first derivative of the wavefront phase can be constructed by measuring the

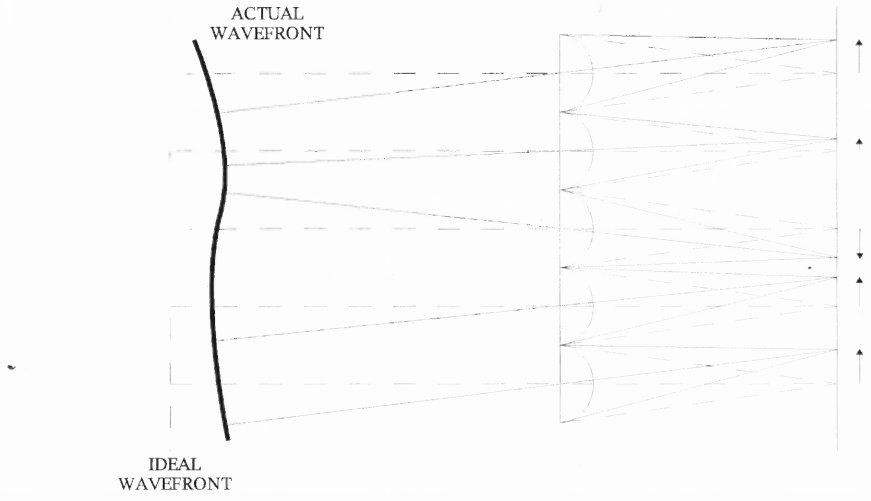

Figure 2.4 Schematic drawing illustrating the principle of a Shack-Hartmann Wavefront Sensor.

shift of each individual subaperture image. The actual wavefront phase (plus an arbitrary constant) can be estimated from its first derivative.

The measured wavefront is reconstructed as a linear combination of basis functions, such as Zernike polynomials [52] or Karhunen-Loève modes [102, 15, 16]. This reconstruction process is described in Appendix C and illustrated in Figure 2.5.

Each subaperture image is cross-correlated with a preselected reference subaperture. Each cross-correlation produces a displaced central peak. The position of the central cross-correlation peak delivers the displacement of the subaperture image with respect to the reference image, which is proportional to the wavefront tilt at the subaperture. Thus, the position shifts of the cross-correlation peaks produce a measurement proportional to the tilt of the wavefront sampled at each subaperture.

Since the measured shifts are relative to the reference subaperture image, the measured wavefront can only be known to within a constant global tilt. This unknown global tilt introduces an uncertainty in the measured absolute position in the sky. Fortunately, this is not an issue for solar applications since solar observations do not require absolute astrometric measurements.

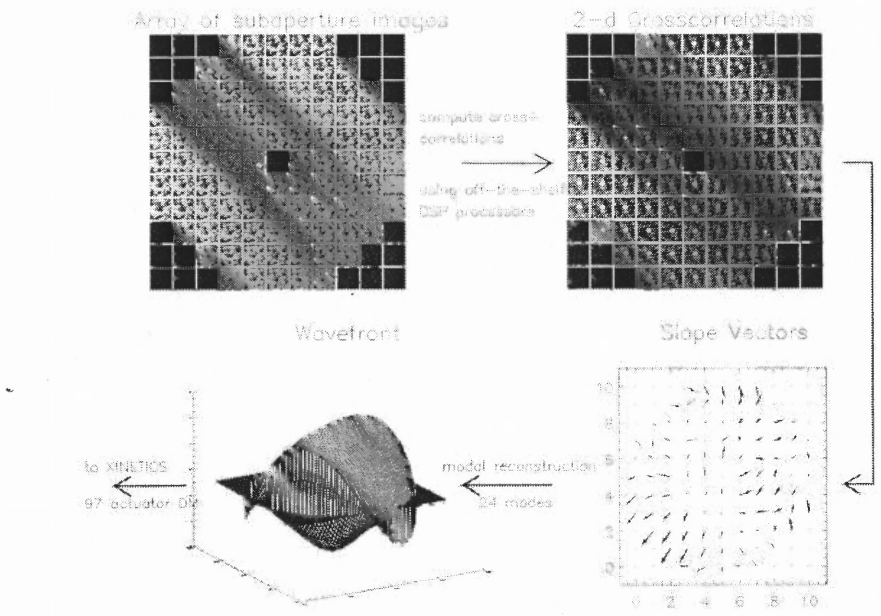

Figure 2.5 Schematic drawing illustrating the principle of a cross-correlating Shack-Hartmann WFS. Image by Rimmele & Radick [64].

2.1.2 Control Loop

The DM is a continuous face plate deformable mirror, which consists of a thin mirror plate attached over an array of piezoelectric actuators. The length of the piezoelectric actuators can be changed by applying a voltage, pushing and pulling on the thin mirror and thus modifying its shape. The AO control system sets the shape of the DM by sending actuator commands to the DM controller. These commands are specified in *counts* units, which are proportional to the distance pushed by the actuator. The specifications provided by the manufacturer indicate that an actuator value of 1000 counts produces a linear push of 1.2 micrometers.

At each loop iteration during normal AO operation, the shape of the DM is updated by the AO system controller to minimize the residual wavefront measured by the WFS. The AO system controller evaluates the pixel shifts measured by the WFS and computes the optimal shape of the DM as a list of actuator commands. The pixel shifts measured by the WFS are transformed into coefficients of some modal basis (KL or Zernike) that

describe the wavefront shape. This shape is then mapped to the DM actuators to produce a list of actuator commands. The process is described mathematically by two matrix multiplications, which can be optimized and combined into a single matrix multiplication, as described in Appendix C.

2.2 Adaptive Optics Error Sources

The correcting performance achieved by an AO system is determined by the amount of residual wavefront errors still present after correction. Several sources contribute to these residual errors that depend on the technical characteristics of the individual components of the system, and also on the seeing conditions. Possible sources of error include WFS noise error, aliasing, angular anisoplanatism error, wavefront fitting error and correction bandwidth error. Each of these individual sources of error adds a certain wavefront variance to the final residual wavefront variance after correction.

The WFS noise error arises from the detector noise and the contrast of the small images in the subapertures of the WFS. The detector noise σ_b^2 contains the readout noise and photon noise from the WFS camera. In solar observations this contribution is dominated by the photon noise due to the high levels of light available for wavefront sensing. The spatial variance of the subaperture images σ_i^2, i.e., their contrast, affects the cross-correlation results between the subapertures. The error variance of the pixel shifts measured by the WFS can be expressed as [48, 47]:

$$\sigma_x^2 = \frac{5m^2\sigma_b^2}{8n_r^2\sigma_i^2} \tag{2.1}$$

where m is the width of the reference subimage autocorrelation function (in pixels) and n_r the width of the subaperture imge (in pixels). As discussed in Appendix D, the high light levels available for wavefront sensing during solar observations make the ratio σ_b^2/σ_i^2 very small and the contribution of the WFS noise error can be safely neglected.

The limited spatial sampling of the WFS introduces aliasing error. The WFS measures high spatial frequencies components of the wavefront as low spatial frequencies. The aliasing error can be considered as 30% to 40% of the wavefront fitting error [21, 25], discussed below.

$$\sigma_A^2 = 0.07 \left(\frac{\Lambda}{r_0} \right)^{5/3} \tag{2.2}$$

The angular anisoplanatism error translates into differential correction across the field-of-view. The wavefront measured by the WFS is measured at a particular field position. Other positions in the field receive light coming from a different angle in the sky and hence travel through different volumes of the atmosphere that may present different seeing conditions. Thus, the correction produced by the AO system is only adequate for a certain angular distance around the location where the wavefront is measured. This angular distance θ_0 is referred to as the isoplanatic patch (see Chapter 3). The mean-square wavefront error introduced by anisoplanatism at a certain angle θ is quantified [33] by:

$$\sigma_\theta^2 = \left(\frac{\theta}{\theta_0} \right)^{5/3} \tag{2.3}$$

The wavefront fitting error arises from the limited ability of the correcting element of the AO system to flatten the wavefront. The limited number of degrees of freedom of the correcting element, 97 actuators in this case, introduces an additional wavefront variance to the overall residual variance of the corrected wavefront. Hardy [33] provides an expression for the wavefront variance introduced by the fitting error (σ_F^2) in the case of a continuous face plate DM with actuator spacing Λ.

$$\sigma_F^2 = 0.28 \left(\frac{\Lambda}{r_0} \right)^{5/3} \tag{2.4}$$

The limited correcting bandwidth of the AO system introduces some residual errors in the corrected wavefront. These bandwidth errors are proportional to the ratio between the bandwidth of the turbulence and the bandwidth of the AO correction. The additional wavefront variance introduced by this error was obtained by Greenwood [30, 33] and is expressed as:

$$\sigma_{BW}^2 = \left(\frac{f_G}{f_S}\right)^{5/3} \tag{2.5}$$

where f_S is the bandwidth of the AO servo loop and f_G is a characteristic frequency of the atmospheric turbulence called the Greenwood frequency. In the special case of a single turbulent layer moving at a speed v, the Greenwood frequency f_G can be written as [33, 97]:

$$f_G = 0.427\frac{v}{r_0} \tag{2.6}$$

Figure 6.8 illustrates the effects of these residual error sources. It plots the Strehl ratio produced by AO correction for different seeing conditions. The error terms described here become more severe for bad seeing conditions (low r_0 values), and increase the residual wavefront variance, thus, reducing the Strehl ratio (see Equation (1.1) for an expression relating Strehl ratio to wavefront variance).

CHAPTER 3

ATMOSPHERIC TURBULENCE

The properties of atmospheric turbulence must be well understood to allow AO systems to provide an optimal correction of the distortions that the turbulence introduces. Similarly, the PSF estimation algorithm makes certain assumptions about the properties of the turbulence that require a good understanding of its mechanisms and properties. This section provides a review of the properties of atmospheric turbulence from knowledge found elsewhere in the literature. It introduces required terminology and tools that provide a foundation for the derivation of the PSF estimation algorithm.

The study of fluid dynamics reveals the highly turbulent nature of Earth's atmosphere. The properties of fluid flows are characterized by the Reynolds number $Re = v_{\text{avg}} L/k_v$, where v_{avg} is the average fluid velocity, L a characteristic length scale, and k_v the kinematic viscosity of the fluid. The Reynolds number of a fluid is a measure of how easily a perturbation in a laminar flow will be dissipated [38]. It is defined as the ratio of two quantities: the rate at which the perturbation kinetic energy is generated and the rate at which the perturbation energy is dissipated. Turbulence will only appear in the case where the kinetic energy associated with a perturbation is created much faster than it is dissipated. Otherwise the perturbation will be damped down and disappear. A fluid flow will be turbulent if its Reynolds number is larger than some critical value defined by the geometrical structure of the flow [66].

For the air, the fluid viscosity is $k_v = 1.5 \times 10^{-6}$ m^2/s and assuming flows with wind speeds of a few m/s and length scales from several meters to hundreds of meters the Reynolds number can be estimated [66, 69, 54] to be $Re^{\text{air}} \geq 10^6$, which implies that air flows in the atmosphere will in general present fully developed turbulence.

Turbulence affects the light traveling through the atmosphere by producing small inhomogeneities in the temperature of the air as different temperature pockets mix. Small temperature variations ($\sim$ 1 K) produce air density changes and thus refractive index variations. These inhomogeneities change with time in an aleatory way, introducing random time-dependent phase differences in the wavefront of the light propagating through the atmosphere. When this light is imaged by an optical system, such as a telescope, the phase changes in the wavefront translate into optical aberrations at the focal plane of the optical system, distorting the captured image.

3.1 The Kolmogorov Model

During the 1940's, Kolmogorov [41] developed a model to describe the structure of fully developed turbulence in fluids with high Reynolds numbers. The model describes how energy is distributed on turbulent cells of different sizes called *eddies*.

Atmospheric turbulence arises from the heating and cooling of the surface of the Earth by the Sun. Air masses with different temperatures are mixed by wind shears, producing large scale disturbances that feed energy into the turbulence. As the turbulence develops, large eddies produce smaller eddies by inertial interaction and kinetic energy is transfered from larger to smaller eddies [37, 54]. Some of the energy is dissipated by viscosity effects, which become more important as the eddie size decreases, down to a certain size l_0, for which most of the kinetic energy is dissipated through molecular friction into heat.

Once the turbulence is in a fully developed state, eddies smaller than a certain size L_0 can statistically be assumed to be in a steady state. This means that the amount of kinetic energy that is dissipated, plus the energy passed down to smaller eddies, must be equal to the kinetic energy received from larger eddies. Furthermore, the rate of energy change for these eddies can be assumed to be negligible. In this situation, these eddies are considered to be in statistical equilibrium with one another.

Eddies of sizes between l_0 and L_0 get their kinetic energy through inertial interaction from larger eddies and thus are not affected by the external conditions that inject the energy into the turbulence. This assumption, which has been verified by experimental observation [37], implies that fully developed turbulence in this range of eddie sizes is isotropic and homogeneous, i.e., statistically invariant under rotation and spatial translation, respectively. This range of eddie sizes between l_0, the *inner scale*, and L_0, the *outer scale*, is commonly referred to as the *inertial range* of the turbulence. Typical values of the outer scale L_0 range from just a few meters to tens of meters [69]. The inner scale l_0 is typically in the order of a few millimeters.

Kolmogorov postulated [24] that in the *inertial range* all small scale statistical properties of the turbulence are determined solely by the scale size l and the mean energy dissipation ϵ. Following dimensional considerations [89], a proportionality expression for the velocity variations V of the turbulence in a scale size l can be found. Kinetic energy fluctuations E, which are proportional to the square of the velocity, are similarly derived.

$$V \propto \epsilon^{1/3} l^{1/3} \tag{3.1}$$

$$E \propto \epsilon^{2/3} l^{2/3} \tag{3.2}$$

A similar relation can be derived from Equation (3.2) for the power spectral density of the turbulence in terms of the scale size. The power spectral density is normally expressed in terms of the spatial wave number $\kappa = 2\pi/l$, instead of the eddie scale size l. The energy of the turbulence in an interval $d\kappa$ is written in terms of the power spectral density: $\Phi(\kappa)d\kappa \propto \kappa^{-2/3}$. Therefore, the power spectral density of the turbulence Φ is proportional to:

$$\Phi(\kappa) \propto \kappa^{-5/3} \tag{3.3}$$

This corresponds to the one dimensional Kolmogorov power law for the turbulence power spectral density. It should be noted that the previous relation is only valid for scales inside the inertial range: $L_0 > l > l_0$.

Kolmogorov introduced the *structure function* as a statistical tool to characterize non-stationary random functions. The structure function of a random function $f(x)$ in a homogeneous and isotropic medium is defined as:

$$D_f(\rho) = \langle |f(x) - f(x+\rho)|^2 \rangle \tag{3.4}$$

where the angled brackets $\langle \ldots \rangle$ indicate an ensemble average. A relation for the velocity structure function can be derived from the proportionality relation in Equation (3.1).

$$D_V(\rho) = \langle |V(\vec{r}) - V(\vec{r}+\vec{\rho})|^2 \rangle = C_V^2 \rho^{2/3} \tag{3.5}$$

where the quantity C_V^2 is called the *velocity structure constant* and relates to the strength of the fluctuations.

The Kolmogorov theory provides useful relations describing the mechanical properties of the turbulence. However, it does not directly address any of its optical properties. On the other hand, the turbulence mixes eddies with different temperatures, which have different densities and therefore different refractive indexes. It is this distribution of pockets with different refractive indexes that introduces distortions in the optical path of the light traveling through them and gives rise to wavefront distortions. Obukhov and Yaglom [33] showed that the results from Kolmogorov's theory also apply to the concentration of any passive and conservative additive, i.e., it does not affect the dynamics of the turbulence and it does not disappear by chemical reaction, such as the mixing of air with different temperatures. Thus, an expression for the temperature structure function D_T can be directly obtained from Equation (3.5).

$$D_T(\rho) = \langle |T(\vec{r}) - T(\vec{r}+\vec{\rho})|^2 \rangle = C_T^2 \rho^{2/3} \tag{3.6}$$

where C_T^2 is the temperature structure constant. The refractive index of air n can be derived from the temperature and pressure of the air and a structure function describing the refractive index variations can be found. However, n is always very close to one. It is more convenient to introduce [55] another quantity called *refractive modulus* or *refractivity* N defined [33] as: $N = (n-1)\,10^6$. The air refractivity can be expressed as a function of pressure P and temperature T by the following approximation:

$$N = (n-1) \times 10^6 \approx 77.6 \frac{P}{T} \tag{3.7}$$

The pressure variations caused by atmospheric turbulence are negligible compared to the temperature variations [33]. Therefore, the pressure P can be considered a constant in the previous expression for the refractivity and the refractive index structure function can be derived from Equation (3.6):

$$D_N(\rho) = \langle |N(\vec{r}) - N(\vec{r}+\vec{\rho})|^2 \rangle = C_N^2 \rho^{2/3} \tag{3.8}$$

where the refractive index structure constant C_N^2 is related to the temperature structure constant C_T^2 by: $C_N = \delta N / \delta T \; C_T$.

$$C_N(\rho) = \left(7.8 \times 10^{-5} \frac{P[\text{mbar}]}{T^2[\text{K}]} \right) C_T \tag{3.9}$$

Here, it is convenient to introduce the *autocorrelation* function of a random variable, such as the air refractivity N:

$$\Gamma_N(\rho) = \int_{-\infty}^{\infty} N(x)N(x+\rho)\,dx \tag{3.10}$$

According to their definitions in Equation (3.4) and Equation (3.10) the structure function and the autocorrelation function are related [28, 33, 54] by the following expression:

$$D_N(\rho) = 2\left[\Gamma_N(0) - \Gamma_N(\rho)\right] \tag{3.11}$$

An important property of the autocorrelation function is described by the Wiener-Khinchin theorem [28], which states that the autocorrelation function Γ_N and the power spectral density Φ_N form a Fourier transform pair:

$$\Gamma_N(\rho) = \int_{-\infty}^{\infty} \Phi_N(\kappa) e^{-i\kappa\rho}\,d\kappa \tag{3.12}$$

where $\kappa = 2\pi/l$ is the spatial wave number as defined previously. Thus, the structure function and the power spectral density can be related through the Wiener-Khinchin theorem and Equation (3.11):

$$D_N(\rho) = 2\int_{-\infty}^{\infty} \left[1 - \cos(\kappa\rho)\right] \Phi_N(\kappa)\,d\kappa \tag{3.13}$$

Tatarski [33, 89] used Equation (3.13) and Equation (3.8) to compute an expression for the one dimensional refractive index power spectral density.

$$\begin{aligned} \Phi_N(\kappa) &= \frac{\Gamma\left(\frac{5}{3}\right)\sin\left(\frac{\pi}{3}\right)}{(2\pi)^{5/3}} C_N^2\,\kappa^{-5/3} \\ &= 0.0365\,C_N^2\kappa^{-5/3} \end{aligned} \tag{3.14}$$

This represents the same $\kappa^{-5/3}$ power law shown in Equation (3.3). The one dimensional case describes the case where there is turbulence in a longitudinal velocity flow. However, turbulence in the Earth's atmosphere is a three dimensional phenomenon. Calculations can be extended to the three dimensional case by changing the integral in Equation (3.13) to a three dimensional integral with volume element $4\pi\kappa^2 d\kappa$, which produces a power spectrum that follows a $\kappa^{-11/3}$ power law [69, 54].

$$\begin{aligned} {}^{3D}\Phi_N(\kappa) &= \frac{\Gamma\left(\frac{8}{3}\right)\sin\left(\frac{\pi}{3}\right)}{4\pi^2} C_N^2\,\kappa^{-11/3} \\ &= 0.033\,C_N^2\,\kappa^{-11/3} \end{aligned} \tag{3.15}$$

The expression for the power spectrum predicted by the Kolmogorov model is only valid inside the inertial range, i.e., for κ in the range $2\pi/L_0 < \kappa < 2\pi/l_0$. In reality, for scales with wave numbers $\kappa > 2\pi/l_0$ the turbulence disappears by dissipating its energy through molecular friction into heat. Also, turbulent eddies with a size larger than L_0 are not believed to be homogeneous. The Kolmogorov model does not provide valid results for the power spectrum outside these limits. Moreover, the expressions given by the Kolmogorov model present a divergent behavior in the case where $\kappa \to 0$.

Several semi-empiric models exist in order to take into account the limitations of the Kolmogorov model. A widely accepted model that agrees well with empirical measurements is the Von Karman model [38, 69]. It eliminates the singularity at $\kappa = 0$ and takes into account the effects of the inner and outer scale, l_0 and L_0. The expression for the three dimensional power spectrum of the turbulence provided by the Von Karman model is shown in Figure 3.1 and given by:

$${}^{3D}\Phi_V(\kappa) = 0.0330 C_N^2 (\kappa^2 + \kappa_0^2)^{-11/6} \exp\left(-\frac{\kappa^2}{\kappa_m^2}\right) \tag{3.16}$$

where $\kappa_0 = 2\pi/L_0$ and $\kappa_m = 2\pi/l_0$.

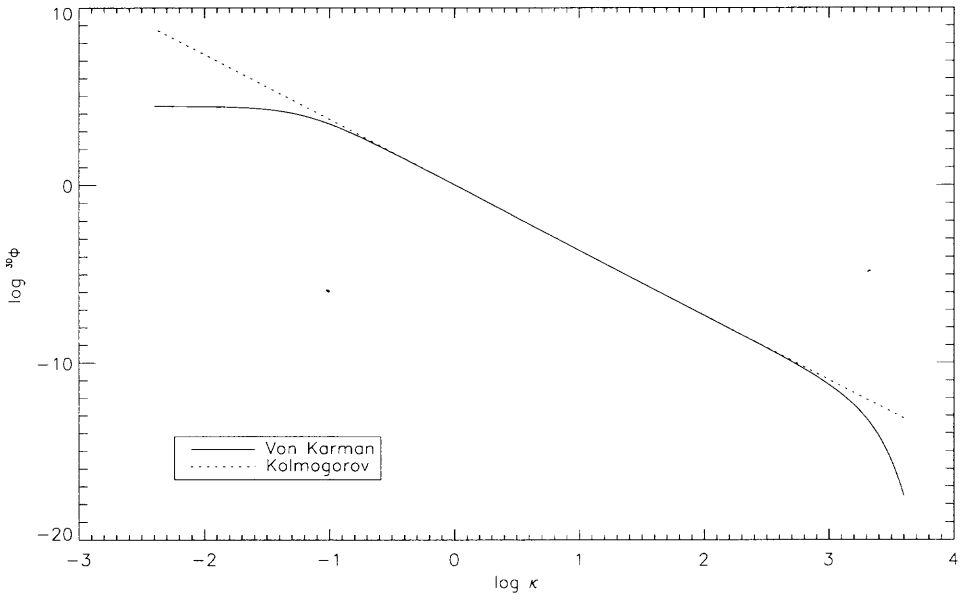

Figure 3.1 Three dimensional atmospheric phase power spectra derived from the Kolmogorov and the Von Karman turbulence models.

The Von Karman model provides a more general description of the turbulence and produces more physically accurate results. However, both models, Kolmogorov and Von Karman, are identical inside the inertial range, shown in Figure 3.1. In cases where the telescope aperture size is much smaller than the outer scale of the turbulence ($D_{\text{tel}} \ll L_0$), the Kolmogorov model is a very good approximation that allows an easier derivation of additional quantities [100, 101]. Aperture sizes of current solar telescopes are small compared with usual values of L_0 and the Kolmogorov model constitutes a good approximation.

3.2 Light Propagation Through Turbulence

Light traveling through the Earth's atmosphere suffers random distortions due to the effects of turbulence. Fluctuations in the refractive index of the air introduce random variations of amplitude and phase in the wavefront of the incoming light. When this light is collected and imaged by a telescope, these distortions translate into image aberrations that degrade image quality.

The variations of the amplitude of the wavefront are referred to as *scintillation* [19, 33, 55]. They arise from Fresnel diffraction as the light propagates through the non-uniform refractive index atmosphere. The variations in refractive index of the air act as a grating diffracting the incoming light, which causes interference with the un-diffracted (zero order) wave, creating a pattern of intensity variations. In general, the refractive index variations are small and the diffracted waves have to travel a long distance before the interference with the un-diffracted zero order wave is noticeable. The intensity fluctuations produced by scintillation are averaged by any telescope with an aperture larger than a few tens of centimeters and can be safely neglected.

When imaging extended objects, as is the case for solar observations, the intensity variations produced by atmospheric turbulence are extremely faint. Under good seeing conditions, intensity variations are estimated [78] to be around $\Delta I/\langle I\rangle \approx 5 \times 10^{-4}$. This fact, combined with the effects of the aperture size, further reduce the significance of scintillation during solar observations. Therefore, it is safe to assume that for solar observations the effects of scintillation can be neglected and the amplitude of the wavefront that reaches the telescope is constant.

As shown before, the statistical properties of the refractive index fluctuations in the atmosphere can be derived from the Kolmogorov model. Wave propagation translates fluctuations of refractive index into distortions in the wavefront phase. The wavefront phase distortions produced by a turbulent layer of thickness δh at a height h that presents a distribution of refractive index $n(x, z)$ is given by:

$$\varphi(x, h) = k \int_{h}^{h+\delta h} n(x, z)\, dz \tag{3.17}$$

where $k = 2\pi/\lambda$ is the wavenumber. Using this equation, an expression for the autocorrelation function of the wavefront phase Γ_{φ} can be found:

$$\begin{aligned}\Gamma_{\varphi}(\rho,h) &= \int \varphi(x)\varphi(x+\rho)\,dx \\ &= k^2 \int_h^{h+\delta h} dz' \int_h^{h+\delta h} dz'' \int n(x,z')\,n(x+\rho,z'')\,dx \\ &= k^2 \int_h^{h+\delta h} dz' \int_{h-z'}^{h+\delta h-z'} \Gamma_N(\rho,z)\,dz \end{aligned} \tag{3.18}$$

For a layer thickness δh much larger than the correlation scales of the refractive index fluctuations, the integration can be extended to $\pm\infty$ producing:

$$\Gamma_{\varphi}(\rho,h) = k^2 \delta h \int_{-\infty}^{\infty} \Gamma_N(\rho,z)\,dz \tag{3.19}$$

Considering Equation (3.11) and Equation (3.19), an expression for the phase structure function can be obtained:

$$\begin{aligned} D_{\varphi}(\rho,h) &= 2\left[\Gamma_{\varphi}(0) - \Gamma_{\varphi}(\rho)\right] \\ &= 2k^2\delta h \int_{-\infty}^{\infty} \left[\Gamma_N(0,z) - \Gamma_N(\rho,z)\right] \\ &= k^2\delta h \int_{-\infty}^{\infty} \left[D_N(\rho,z) - D_N(0,z)\right] \end{aligned}$$

The integral can be computed by inserting D_N from Equation (3.8). The phase structure function produced by a layer of thickness δh is given by [33, 69, 54]:

$$D_{\varphi}(\rho,h) = 2.914\,k^2\delta h\,C_{N_i}^2(h)\,\rho^{5/3} \tag{3.20}$$

The global phase structure function D_{φ} is found by integrating Equation (3.20) along the line of sight over the whole atmosphere. The zenith angle θ_z is introduced to account for changes in the length of the line of sight travel path with observing angle.

$$D_{\varphi}(\rho) = 2.914\, k^2 (\sec\theta_z) \rho^{5/3} \int C_N^2(h)\, dh \tag{3.21}$$

Fried [23] wrote Equation (3.21) as a function of a single parameter called *seeing cell size*. Such a parameter is commonly referred to as the Fried parameter r_0 and is defined by:

$$r_0 \equiv \left[0.423\, k^2 (\sec\theta_z) \int C_N^2(h)\, dh \right]^{-3/5} \tag{3.22}$$

The value of the Fried parameter has a dependency on wavelength inside the wave number k in Equation (3.22), namely $r_0 \propto \lambda^{6/5}$. As a convention, the value of the Fried parameter is always given as defined at a wavelength of 550 nm.

The phase structure function can be expressed in terms of the Fried parameter [23], resulting in a much simpler form:

$$D_{\varphi}(\rho) = 6.88 \left(\frac{\rho}{r_0} \right)^{5/3} \tag{3.23}$$

Similarly, the power spectral density of the wavefront phase distortions Φ_{φ} can be obtained [33, 100] and expressed in terms of the Fried parameter.

$$\Phi_{\varphi}(\kappa) = \frac{0.023}{r_0^{5/3}} \kappa^{-11/3} \tag{3.24}$$

The Fried parameter is referred to as the coherence length of the turbulence. It characterizes the strength of the phase distortions introduced by the turbulence. For example, a circular section with diameter r_0 of a wavefront distorted by turbulence presents a phase variance of around 1 rad^2 (more precisely [52] $\sigma_{\varphi}^2 = 1.03\text{rad}^2$). Hence, the larger the value of r_0, the weaker the phase distortions introduced by the turbulence. It can be assumed that

distortions inside a region of diameter r_0 are negligible. Thus, if the aperture diameter of the optical system collecting this wavefront is smaller than r_0, the effects of the distortions in the captured image can be neglected. And consequently, for larger aperture sizes, the maximum spatial resolution achievable by the optical system is limited by the value of r_0. In general, $r_0 < D$ and images captured with a telescope of diameter D have the same resolution as images taken with a telescope of aperture r_0. This clearly illustrates the severe consequences that atmospheric turbulence has on astronomical observations. The wavelength dependence of r_0 makes its value larger at longer wavelengths, which means that for a given telescope aperture D it is much easier to achieve diffraction-limited performance at longer wavelengths.

3.3 Time Evolution of Turbulence

The spatial structure of the turbulence changes on time scales much larger than the time it takes the wind to blow the turbulence across the telescope aperture. This is the *Taylor hypothesis of frozen turbulence* and it states that the effects caused by a turbulent layer can be modeled by a "frozen" pattern that is moved across the aperture of the telescope [54].

The overall turbulence over the whole height of the atmosphere can be caused by the joint effect of several turbulent layers causing its time evolution to be quite complicated. However, it can globally be characterized [54, 69] by a single parameter, i.e., the correlation time τ_0. It is defined as the time it takes the turbulence to travel a distance r_0.

$$\tau_0 \equiv r_0/v, \tag{3.25}$$

where v is the wind speed of the dominant turbulent layer. Typical values for r_0 ~10 cm and v ~10 m/s produce a value for $\tau_0 \sim 10$ ms. Observations with exposure times $t \gg \tau_0$ will integrate over several random configurations of the atmosphere. These kind of observations are called *long-exposures*.

Observations with exposure times $t \sim \tau_0$ are called *short-exposures* or *speckle images* and their distortions are produced by frozen realizations of atmospheric turbulence. Since they are produced by a frozen configuration of the atmosphere, they are not affected by image motion caused by overall tilt and contain partial information in a wide range of spatial frequencies up to the diffraction-limited frequency of the optical system. Long-exposure images are formed as the result of integrating several of these speckle images, each of which contains information in different spatial frequency ranges. Spatial frequency information that is common to many individual speckle images is reinforced by the integration process and appears in the final long-exposure. Any other underrepresented spatial frequency components are effectively damped down in the final long-exposure image. The integration process translates into a loss of high spatial frequency information in the long-exposure image (as illustrated in Figure 4.3).

Each individual speckle image contains a partial representation of the original spatial frequency information. With a large set of individual speckle images, it is possible to reconstruct the undistorted original frequency information and produce an estimate of a perfect diffraction-limited image. This technique is called *speckle reconstruction* and it is currently widely used in some astronomical and solar applications.

The parameter τ_0 has a very important role in the design of an AO system. The value of τ_0 characterizes how fast atmospheric distortions change with time and sets a requirement on the correcting bandwidth of the AO system. The AO system must have a bandwidth larger than $1/\tau_0$ to be able to reject the changing atmospheric distortions.

3.4 Anisoplanatism

Each position in the field-of-view of a telescope receives light coming from different angles in the sky. The light coming from each of these angles travels through a particular volume of the atmosphere, which may present its own seeing conditions. Hence, different points in

the field-of-view may have been affected by different seeing conditions and the image may contain distortions that change across the field-of-view. This is called anisoplanatism.

The implications of this problem can be studied from a geometrical optics approach. Light coming from two directions separated by an angle θ travels through two different points of a turbulent layer of height h. These two points are separated by a distance ρ at the layer height. At zenith angle θ_z, the distance d from the telescope to where the rays cross the turbulent layer is $d = h \sec\theta_z$. Thus, ρ is related to d and h by $\rho = \theta d$ and $\rho = \theta h \sec\theta_z$.

These relations can be inserted in Equation (3.21) to compute the phase structure function in terms of the angular separation θ. The definition of the structure function in Equation (3.4) indicates that $D_\varphi(\rho)$ is the mean-square wavefront phase difference between two points separated a distance ρ. Thus, the mean phase variance [33, 54] between two light rays separated by an angle θ is:

$$\langle\sigma_\varphi^2\rangle = D_\varphi(\theta) = 2.914\, k^2(\sec\theta_z) \int C_N^2(h)(\theta h \sec z)^{5/3} dh \tag{3.26}$$

Following the same approach as with the definition of the Fried parameter r_0, the *isoplanatic angle* θ_0 is defined as:

$$\theta_0 = \left[2.914\, k^2(\sec\theta_z)^{8/3} \int C_N^2(h)\, h^{5/3} dh\right]^{-3/5} \tag{3.27}$$

The isoplanatic angle θ_0 is the separation angle for which the relative phase variance is $\sim$1 rad^2. Equation (3.26) can be written in a much simpler form:

$$\langle\sigma_\varphi^2\rangle = \left(\frac{\theta}{\theta_0}\right)^{5/3} \tag{3.28}$$

The definition of θ_0 is similar to the definition of the Fried parameter r_0 in Equation (3.22). The isoplanatic angle θ_0 can be expressed in terms of r_0:

$$\theta_0 = 0.314(\cos\theta_z)\frac{r_0}{H} \tag{3.29}$$

where H is the *mean effective turbulent height*:

$$H = \left(\frac{\int C_N^2(h)\, h^{5/3} dh}{\int C_N^2(h)\, dh}\right)^{3/5} \tag{3.30}$$

Equation (3.29) indicates that a situation with large r_0 (weak turbulence) and small H (low-altitude turbulence) is the most favorable for a large isoplanatic angle. In other words, anisoplanatism arises mostly from high-altitude turbulent layers. The isoplanatic angle θ_0 also contains a dependency on wavelength, as shown in Equation (3.27). It scales with $\lambda^{6/5}$, making its effects less critical for longer wavelengths. Typical values (in seconds of arc) for θ_0 range from $10''$ in the visible to around $40''$ in the infrared (IR).

The effects of anisoplanatism are very important whenever image correction is to be applied. With aberrations being different for different parts of the field-of-view, anisoplanatism sets a limit to the correctable field area [5]. AO systems perform wavefront sensing in a small region of the field around the so called lock point. In night-time AO, the lock point is either a NGS or a LGS. In solar AO, it is any kind of solar structure with enough contrast, such as a sunspot or granulation. Aberrations are measured for that region of the field and the applied correction is only valid in an area defined by θ_0 around that region, called the *isoplanatic patch.*

Anisoplanatism translates into different areas of the field-of-view having different PSFs. A PSF estimated from the AO loop data will only be valid inside the isoplanatic patch. Therefore, the results of a deconvolution with that estimated PSF will only be accurate inside the isoplanatic patch.

CHAPTER 4

POINT SPREAD FUNCTION ESTIMATION METHOD

The method to estimate the long-exposure PSF from solar AO loop data is based on work by Veran et al. [101, 100] to implement a similar method for a night-time curvature sensor based AO system at the CFHT telescope in Hawaii. The method obtains information about the AO-corrected residuals and the atmospheric seeing conditions from the AO loop data produced by the AO system during normal correction.

Figure 4.1 schematically illustrates the structure of the PSF estimation method. The long-exposure OTF, i.e., the Fourier transform of the PSF, is expressed as the product of three independents OTF components: a component derived from the measured AO corrected residuals (obtained from the WFS measurements), a component derived from the atmospheric seeing conditions (obtained from the DM commands) and a component derived from the telescope.

4.1 Image Formation

An optical system, such as a telescope, captures light from an object and forms an image, which ideally should be a perfectly identical representation of the object. However, in reality, any optical system behaves like a low band pass filter with respect to spatial frequency. The spatial frequency components of the original object are modulated by the optical system and spatial frequencies larger than a cut-off frequency are destroyed. In a perfect, i.e., aberration free optical system, the maximum spatial frequency that can pass through the system is determined by diffraction. If aberrations are present in the optical system, the cut-off frequency can be considerably reduced, further degrading the quality of the image.

The resulting image $i(\vec{\rho})$ can be described as a convolution between the object $o(\vec{\rho})$ and the characteristic function of the optical system, i.e., the PSF. An object can be de-

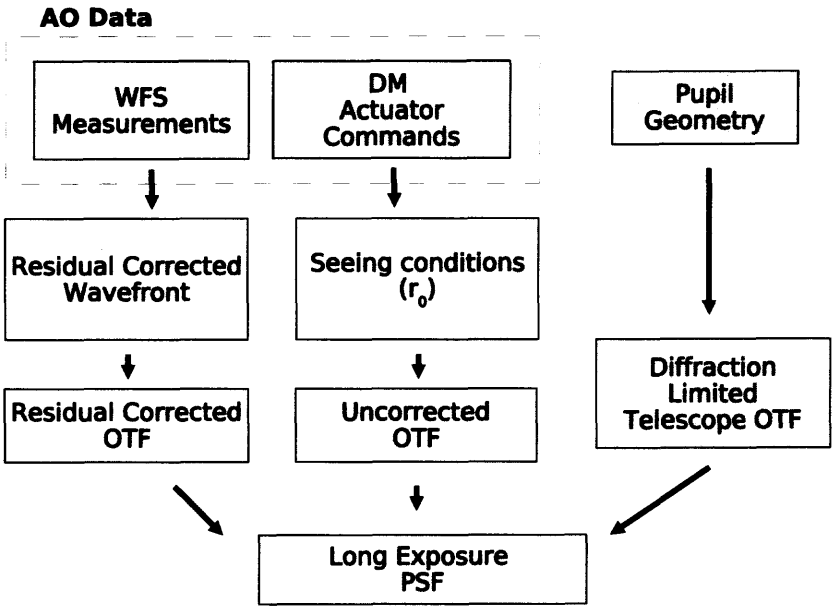

Figure 4.1 Schematic block diagram describing the method to estimate the long-exposure PSF from solar AO loop data.

scribed as formed by a distribution of individual points, each one of which is imaged as a PSF by the optical system to produce the final image. The image is the result of the convolution of the object with the PSF of the optical system [85, 14]. It should be noted that this convolution scheme assumes isoplanatism, i.e., the PSF is the same over the entire field-of-view. The convolution is expressed as:

$$i(\vec{\rho}) = PSF(\vec{\rho}) * o(\vec{\rho}) + n(\vec{\rho}) \tag{4.1}$$

$$= \int_{-\infty}^{\infty} PSF(\vec{x})\, o(\vec{\rho} - \vec{x})\, d\vec{x} + n(\vec{\rho}) \tag{4.2}$$

The term $n(\vec{\rho})$ accounts for any additive noise present in the image. The PSF can be visualized as the blurring function that degrades the quality of the captured image with respect to the true object. If the PSF is known, it fully describes the distortions introduced by the optical system and may make it possible to restore a better estimation of the actual object.

A convolution can be computed more easily in the Fourier domain, where it transforms into a regular product:

$$\begin{aligned} \mathcal{F}[i](\vec{\nu}) &= \mathcal{F}[PSF * o + n](\vec{\nu}) \\ &= \mathcal{F}[PSF](\vec{\nu})\,\mathcal{F}[o](\vec{\nu}) + \mathcal{F}[n](\vec{\nu}) \qquad (4.3) \\ I(\vec{\nu}) &= OTF(\vec{\nu})\,O(\vec{\nu}) + N(\vec{\nu}) \qquad (4.4) \end{aligned}$$

where I and O are the Fourier transforms of the image (i) and the object (o), respectively. The OTF is the Fourier transform of the PSF, i.e., the spatial frequency response of the optical system.

4.2 Aberration Free Optical Transfer Function

According to diffraction theory [9, 28], the light distribution at the focal plane of an optical system arises from Fraunhofer diffraction at the entrance pupil. The optical system acts as a Fourier operator so the wavefront at the focal plane is proportional to the Fourier transform of the wavefront at the entrance pupil. A telescope forms an image of a very distant object on its focal plane. Hence, the wavefront distribution at the image plane is the Fourier transform of the wavefront on the pupil. Considering monochromatic light of wavelength λ:

$$\phi_{\mathrm{im}}(\vec{\rho}, t) \propto \int \phi_{\mathrm{pup}}(\vec{x}, t)\, P(\vec{x}) \exp\left[-i\frac{2\pi}{\lambda}\vec{x}\vec{\rho}\right] d\vec{x} \qquad (4.5)$$

The function $P(\vec{x})$ is the pupil function, i.e., a function that is 1 inside the pupil and 0 outside. The product $\phi_{\mathrm{pup}}(\vec{x}, t)\, P(\vec{x})$ indicates the wavefront that falls inside the pupil of the telescope. In the case where ϕ_{pup} is the wavefront coming from a distant point source, the resulting intensity in the focal plane is exactly the PSF of the optical system. Note that the intensity of the image is the modulus square of the phase:

$$PSF(\vec{\rho}, t) \propto |\mathcal{F}[\phi_{\text{pup}} P]|^2 \tag{4.6}$$

Equation (4.6) indicates that the PSF is the power spectral density of the wavefront in the pupil of the optical system. The Wiener-Khinchin theorem (see Equation (3.12)), states that the power spectral density and the autocorrelation function constitute a Fourier transform pair. The Fourier transform of the PSF, i.e., the OTF, can be written as a normalized autocorrelation of the wavefront on the pupil [28, 100].

$$OTF(\vec{\rho}/\lambda, t) = \frac{1}{S} \int \phi_{\text{pup}}(\vec{x}, t) \phi^*_{\text{pup}}(\vec{x} + \vec{\rho}, t) P(\vec{x}) P^*(\vec{x} + \vec{\rho})\, d\vec{x} \tag{4.7}$$

where S is the surface area of the pupil. The normalization factor $1/S$ ensures that the energy contained by the PSF is normalized to unity. Any static aberrations that may affect the optical system may be included in the OTF of the system by projecting the distortions they produce into ϕ_{pup}.

An aberration free wavefront produced by a very distant point object can be considered to be flat. In this case, the wavefront ϕ_{pup} in Equation (4.7) is constant and can be taken out of the integral so that the OTF is solely determined by light diffracted by the pupil of the optical system: the OTF of the system is said to be diffraction-limited. From Equation (4.7), it follows that the diffraction-limited OTF of an optical system OTF_{diff} is the autocorrelation of the pupil function $P(\vec{x})$. A cross-section of a diffraction-limited OTF of a telescope with a circular pupil is shown in Figure 4.2.

$$OTF_{\text{diff}}(\vec{\rho}/\lambda) = \frac{1}{S} \int P(\vec{x}) P^*(\vec{x} + \vec{\rho}) d\vec{x} \tag{4.8}$$

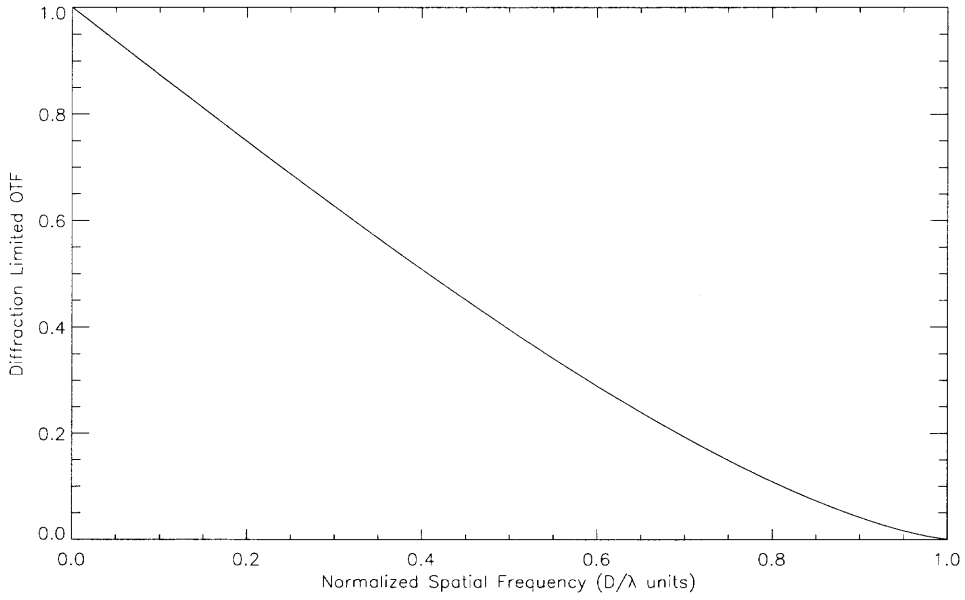

Figure 4.2 Azimuthal average of the diffraction-limited OTF_{diff}.

4.3 Atmospheric Optical Transfer Function

The optical distortions introduced by the atmospheric turbulence must be taken into account when computing the OTF of the optical system. The effective OTF is the OTF of the combined optical system formed by the Earth's atmosphere and the telescope. The distorted wavefront at the pupil of the telescope is written as:

$$\phi_{\text{pup}}(\vec{x}, t) = A \exp\left(-i\varphi(\vec{x}, t)\right) \tag{4.9}$$

The amplitude variations of the wavefront (scintillation) are neglected as discussed in Section 3.2. Equation (4.7) refers to the instantaneous OTF of the system at a time t produced by a wavefront phase $\varphi(\vec{x}, t)$ at the pupil. When exposure times are longer than just a few tens of milliseconds, several of these instantaneous configurations add up to produce an integrated result. This is normally the case during solar observations, since exposure times on the order of several hundreds of milliseconds are common.

The long-exposure OTF is defined as the ensemble average of the instantaneous OTFs over the exposure time period.

$$\begin{aligned} OTF(\vec{\rho}/\lambda) &= \langle OTF(\vec{\rho}/\lambda, t)\rangle = \\ &= \frac{1}{S}\int P(\vec{x})P^*(\vec{x}+\vec{\rho})\langle \exp\left[i\varphi(\vec{x},t) - i\varphi(\vec{x}+\vec{\rho},t)\right]\rangle d\vec{x} \end{aligned} \quad (4.10)$$

The phase function $\varphi(\vec{x}, t)$ represents a random variable with gaussian statistics and zero mean. A useful property of gaussian random variables with zero mean [66] can be applied here to simplify the expression for the long-exposure OTF:

$$\langle \exp(\alpha u)\rangle = \exp\left(\frac{1}{2}\alpha^2\langle u^2\rangle\right) \quad (4.11)$$

Considering Equation (4.11) and the definition of the structure function shown in Equation (3.4), the long-exposure OTF, in Equation (4.10), can be rewritten in a more simplified form as:

$$OTF(\vec{\rho}/\lambda) = \frac{1}{S}\int P(\vec{x})P^*(\vec{x}+\vec{\rho})\exp\left[-\frac{1}{2}D_\varphi(\vec{x},\vec{\rho})\right]d\vec{x} \quad (4.12)$$

where the phase structure function D_φ is defined according to Equation (3.4) as:

$$D_\varphi(\vec{x},\vec{\rho}) = \langle|\varphi(\vec{x},t) - \varphi(\vec{x}+\vec{\rho},t)|^2\rangle \quad (4.13)$$

The effects of a fully developed atmospheric turbulence on the propagating wavefront are described by the Kolmogorov model. As previously discussed, the turbulence is assumed to be homogeneous and isotropic, which translates into a phase structure function that does not depend on the position in the pupil. It is only a function of the distance be-

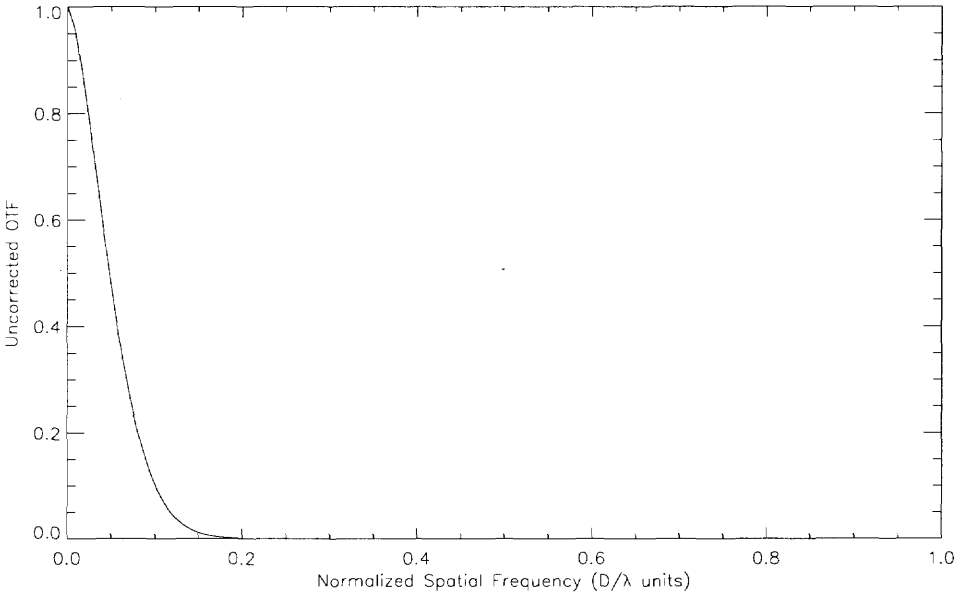

Figure 4.3 Azimuthal average of an example uncorrected OTF caused by uncorrected atmospheric turbulence (OTF_{atm})

tween two points ρ. This allows the term containing the phase structure function to come out of the integral:

$$OTF(\vec{\rho}/\lambda) = \left[\exp\left(-\frac{1}{2}D_{\varphi}(\vec{\rho})\right)\right]\left[\frac{1}{S}\int P(\vec{x})P^{*}(\vec{x}+\vec{\rho})d\vec{x}\right] \tag{4.14}$$

$$OTF(\vec{\rho}/\lambda) = OTF_{\mathrm{atm}}(\vec{\rho}/\lambda)\,OTF_{\mathrm{tel}}(\vec{\rho}/\lambda) \tag{4.15}$$

Fried [23] obtained an analytic expression for the phase structure function of Kolmogorov turbulence, Equation (3.23), which can be substituted in Equation (4.15) to produce:

$$OTF_{\mathrm{atm}}(\vec{\rho}/\lambda) = \exp\left[-3.44\left(\frac{\rho}{r_0}\right)^{5/3}\right] \tag{4.16}$$

The distortions introduced in the wavefront by the atmosphere effectively limit the performance of the telescope. Figure 4.3 shows the azimuthal average of an uncorrected

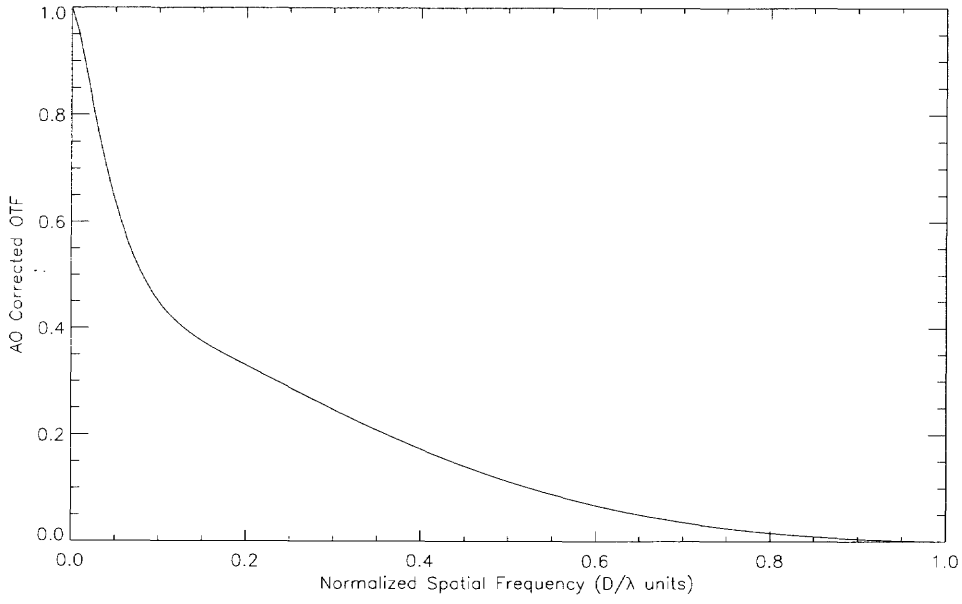

Figure 4.4 Azimuthal average of an example of AO-corrected OTF. It illustrates the effects of AO correction, specially for the high spatial frequencies.

long-exposure OTF distorted by atmospheric turbulence. This OTF produces an image where most of the high spatial frequency information is lost. In this situation, knowing the OTF of the optical system does not provide any advantage, since the high spatial frequency information cannot be recovered.

AO correction preserves the high spatial frequency components in the image up to the diffraction-limited cut-off frequency, as illustrated in Figure 4.4. With knowledge of the AO-corrected OTF, it is possible to use post-processing techniques to restore the amplitudes of all spatial frequencies to their diffraction-limited levels.

4.4 Optical Transfer Function after Adaptive Optics Correction

Once the wavefront distortions are corrected by the AO system, their characteristics are no longer described by the Kolmogorov model of atmospheric turbulence. The results from Section 4.1 apply to an uncorrected wavefront distorted by atmospheric turbulence φ_{atm} and cannot be directly applied to the residual wavefront phase after AO correction φ_{ϵ}.

As seen previously, the turbulence described by the Kolmogorov model is homogeneous and isotropic, making the phase structure function a function of just ρ, i.e., the distance between two points. In the AO-corrected case, this assumption is no longer valid because the phase distortions have been partially corrected by the AO system. The AO-corrected phase structure function is in general a function of position in the pupil as well as distance between two points ($\vec{x}$ and $\vec{\rho}$, respectively) $D_{\varphi_\epsilon}(\vec{x}, \vec{\rho})$.

Computing this structure function involves the averaging of four-dimensional functions, which makes it computationally impractical for these purposes. However, it has been suggested [13, 100] that an average of the phase structure function over the variable $\vec{x}$ inside the pupil is a good approximation for the AO-corrected structure function. This approximation represents the case of having a long exposure time where a large number of phase configurations are included.

$$\bar{D}_{\varphi_\epsilon}(\vec{\rho}) = \frac{\int P(\vec{x})P(\vec{x}+\vec{\rho})D_{\varphi_\epsilon}(\vec{x},\vec{\rho})d\vec{x}}{\int P(\vec{x})P(\vec{x}+\vec{\rho})d\vec{x}} \tag{4.17}$$

Using the mean phase structure function, which is a function of just $\vec{\rho}$, the same approximations that were applied in the derivation of the uncorrected OTF can be applied here. Hence, the AO-corrected OTF can be written as a product of individual OTFs:

$$OTF(\vec{\rho}/\lambda) = OTF_{\varphi_\epsilon}(\vec{\rho}/\lambda)\, OTF_{\text{tel}}(\vec{\rho}/\lambda) \tag{4.18}$$

The previous expression can be further simplified by considering the characteristics of the residual phase φ_ϵ after AO correction. The AO system corrects the phase distorted by atmospheric turbulence φ_{atm} by adapting the shape of the DM to introduce a correcting phase φ_m. The phase present after reflection on the DM φ_ϵ is the residual phase after correction.

$$\varphi_\epsilon(\vec{x},t) = \varphi_{\text{atm}}(\vec{x},t) - \varphi_m(\vec{x},t) \tag{4.19}$$

However, the possible phase configurations of the DM are limited by its physical design characteristics, such as its limited number of actuators, actuator stroke and inter-actuator separation. The separation between actuators is a direct consequence of the total number of actuators, while the actuator stroke limits the severity of correctable distortions.

With a number m of actuators, the mirror can only reproduce phase representations that belong to an m-dimensional vector space of wavefront phases. The space containing all possible phase configurations $\mathcal{E}$ can be divided [101] into two orthogonal subspaces $\mathcal{E} = \mathcal{M}^{\parallel} \oplus \mathcal{M}^{\perp}$, where the operator $\oplus$ represents the direct sum of two subspaces. The subspace $\mathcal{M}^{\parallel}$ is generated by the m actuators of the DM while the subspace $\mathcal{M}^{\perp}$ is the orthogonal complement to $\mathcal{M}^{\parallel}$ in $\mathcal{E}$, which contains higher spatial frequency components that are not corrected by the DM. Thus, the incoming aberrated phase can be projected onto these two subspaces producing:

$$\varphi_{\text{atm}}(\vec{x},t) = \varphi_{\text{atm}\parallel}(\vec{x},t) + \varphi_{\text{atm}\perp}(\vec{x},t) \tag{4.20}$$

Since the DM is only able to produce phase configurations that are contained within the $\mathcal{M}^{\parallel}$ space, the residual phase after correction will still contain all the wavefront components that are part of $\mathcal{M}^{\perp}$, i.e., the phase component $\varphi_{\text{atm}\perp}$. After AO correction, the residual wavefront can be written as:

$$\begin{aligned} \varphi_\epsilon(\vec{x},t) &= \underbrace{\left[\varphi_{\text{atm}\parallel}(\vec{x},t) - \varphi_m(\vec{x},t)\right]}_{\varphi_{\epsilon\parallel}(\vec{x},t)} + \varphi_{\text{atm}\perp}(\vec{x},t) \\ &= \varphi_{\epsilon\parallel}(\vec{x},t) + \varphi_{\text{atm}\perp}(\vec{x},t) \end{aligned} \tag{4.21}$$

The residual phase after correction is separated into two components: the *parallel* component, which contains the residuals after correction of $\varphi_{\text{atm}\parallel}$ ($\varphi_{\epsilon\parallel} \in \mathcal{M}^{\parallel}$), and the *perpendicular* uncorrected phase component ($\varphi_{\text{atm}\perp} \in \mathcal{M}^{\perp}$), which the AO system cannot correct and traveled unaffected through the system.

Combining Equation (4.21), the definition of the phase structure function from Equation (4.13), and the mean phase structure function from Equation (4.17) results in:

$$\bar{D}_{\varphi_\epsilon}(\vec{\rho}) = \bar{D}_{\varphi_{\epsilon\parallel}}(\vec{\rho}) + \bar{D}_{\varphi_{\text{atm}\perp}}(\vec{\rho}) + \bar{D}_{\varphi_{\epsilon\parallel}\varphi_\perp}(\vec{\rho}) \tag{4.22}$$

$$\bar{D}_{\varphi_\epsilon}(\vec{\rho}) \approx \bar{D}_{\varphi_{\epsilon\parallel}}(\vec{\rho}) + \bar{D}_{\varphi_{\text{atm}\perp}}(\vec{\rho}) \ , \tag{4.23}$$

where the term $\bar{D}_{\varphi_{\epsilon\parallel}\varphi_\perp}(\vec{\rho})$ is a cross term arising from correlations between $\varphi_{\epsilon\parallel}$ and $\varphi_{\text{atm}\perp}$, which can be assumed to be negligible [101]. It should be noted that the perpendicular structure function $D_{\varphi_{\text{atm}\perp}}$ can be assumed as a function of only the distance between two points ($\vec{\rho}$), since it is computed from uncorrected atmospheric phase that is homogeneous and isotropic. Hence, $D_{\varphi_{\text{atm}\perp}}$ comes out of the integral in Equation (4.17) producing $\bar{D}_{\varphi_{\text{atm}\perp}} \equiv D_{\varphi_{\text{atm}\perp}}$.

Substituting the mean phase structure function as the phase structure function in Equation (4.14), the total AO-corrected long-exposure OTF can be written as a product of three independent components:

$$OTF(\vec{\rho}/\lambda) = OTF_{\epsilon\parallel}(\vec{\rho}/\lambda)\ OTF_{\text{atm}\perp}(\vec{\rho}/\lambda)\ OTF_{\text{tel}}(\vec{\rho}/\lambda) \tag{4.24}$$

This equation is similar to Equation 4.15 for the uncorrected case. In the AO-corrected case, most of the incoming aberrations have been corrected by the AO system, and the term OTF_{atm} from Equation (4.15) is now split in two components: $OTF_{\epsilon\parallel}$, which arises from the measurable phase residuals after AO correction $\varphi_{\epsilon\parallel}$; and $OTF_{\text{atm}\perp}$ from

the atmospheric phase distortions that traveled undetected and uncorrected through the AO system $\varphi_{\mathrm{atm}\perp}$.

$$OTF_{\epsilon\parallel}(\vec{\rho}/\lambda) = \exp\left[-\frac{1}{2}\bar{D}_{\varphi_{\epsilon\parallel}}(\vec{\rho})\right] \tag{4.25}$$

$$OTF_{\mathrm{atm}\perp}(\vec{\rho}/\lambda) = \exp\left[-\frac{1}{2}D_{\varphi_{\mathrm{atm}\perp}}(\vec{\rho})\right] \tag{4.26}$$

The computation of the AO-corrected OTF (and consequently the AO-corrected PSF) requires the estimation of both residual phase components: $\varphi_{\epsilon\parallel}$ and $\varphi_{\mathrm{atm}\perp}$. The AO loop data produced by the AO system during normal operation contain sufficient information to estimate the statistical properties of both phase components.

The residual corrected phase component $\varphi_{\epsilon\parallel}$ can be extracted from the WFS measurements, since the WFS directly measures the residual wavefront after reflection on the DM. The uncorrected orthogonal phase component $\varphi_{\mathrm{atm}\perp}$ travels undisturbed through the AO system. Hence, its characteristics are not directly measured by the system. Details about the orthogonal component must be derived from a model that describes the wavefront distortions introduced by atmospheric turbulence, i.e., the Kolmogorov model for atmospheric turbulence [41]. The Kolmogorov model characterizes the strength of the turbulence with just one parameter, i.e., the Fried parameter r_0. If r_0 is known, the Kolmogorov model can be used to quantify the characteristics of $\varphi_{\mathrm{atm}\perp}$.

The residual wavefront measured by the WFS $\varphi_{\epsilon\parallel}$ has already been corrected by the AO system and does not contain any information about the original seeing conditions. The orthogonal wavefront component $\varphi_{\mathrm{atm}\perp}$ contains high spatial frequencies that are not measurable by the WFS. Thus, the characteristics of $\varphi_{\mathrm{atm}\perp}$ cannot be obtained from the WFS measurements.

Information about the atmospheric seeing conditions can be extracted from the shape of the DM during correction, which can be reconstructed from the actuator commands sent by the AO system to the DM controller. The DM shape replicates the wavefront distortions

measured by the WFS in a least-squares sense, and can be used to extract information about the $\varphi_{\text{atm}\parallel}$ component of the atmospheric phase distortions. The measured properties of the atmospheric wavefront component $\varphi_{\text{atm}\parallel}$ can be fitted to the Kolmogorov model to obtain a value for the Fried parameter r_0. Finally, the properties of the $\varphi_{\text{atm}\perp}$ component can be quantified from the Kolmogorov model using the fitted value of r_0 (see Section 4.4.2). Thus, both the WFS measurements and the commands sent to the DM are needed in order to completely determine the residual wavefront after AO correction.

4.4.1 Calculation of $\text{OTF}_{\epsilon\parallel}$

The corrected component of the residual phase $\varphi_{\epsilon\parallel}$ is directly measured by the WFS. The cross-correlating Shack-Hartmann WFS measures the image shift at each subaperture, which is proportional to the mean wavefront slope inside the subaperture. The measured shifts are pixel shifts, since they are obtained from the images formed on the camera chip. They must be translated into physical values that relate to the actual value of the wavefront slope at that position using a calibration factor, which transforms pixels in the WFS camera to radians on the wavefront k_{wfs} (see Appendix B).

The measured wavefront is reconstructed from the slopes measured by WFS and represented as a linear combination of a modal basis. KL functions were chosen over Zernike polynomials as the modal basis because they are a better match to describe the phase distortions produced by atmospheric turbulence (see Appendix A). Moreover, the Shack-Hartmann WFS is more sensitive to KL modes than to Zernike polynomials. The reconstruction process is carried out by a matrix multiplication with the reconstruction matrix $\mathbf{D}^+$. The calculation of the reconstruction matrix is described in detail in Appendix C.

The KL coefficients $\hat{\epsilon}$ that describe the measured residual wavefront are obtained by multiplying the reconstruction matrix $\mathbf{D}^+$ with the measured WFS shifts $\vec{w}$. The calibration factor k_{wfs} transforms the measured pixel shifts to actual wavefront slopes, as shown in Equation (4.27).

$$\vec{\hat{\epsilon}}(t) = \mathbf{D}^{+}\ [k_{\text{wfs}}\,\vec{w}(t)] \tag{4.27}$$

where the residual WFS measurements $\vec{w}(t)$ are the measured pixel shifts placed in column vector form. The hat on top of $\hat{\epsilon}_i$ indicates that these coefficients are obtained from real measurements and as such are affected by noise. The calibration factor k_{wfs} should be incorporated into $\mathbf{D}^{+}$ in order to reduce the number of multiplications: $\mathbf{D}^{+} \rightarrow \mathbf{D}^{+}\,k_{\text{wfs}}$. The phase components from Eqs. (4.20) and (4.21) can be expressed in terms of KL modes K_i:

$$\varphi_{\text{atm}}(\vec{x},t) = \sum_{i=1}^{\infty} a_i(t)K_i(\vec{x}) \tag{4.28}$$

$$\varphi_{\text{atm}\parallel}(\vec{x},t) = \sum_{i=1}^{N} a_i(t)K_i(\vec{x}) \tag{4.29}$$

$$\varphi_{\text{atm}\perp}(\vec{x},t) = \sum_{i=N+1}^{\infty} a_i(t)K_i(\vec{x}) \tag{4.30}$$

$$\varphi_{m}(\vec{x},t) = \sum_{i=1}^{N} k_i(t)K_i(\vec{x}) \tag{4.31}$$

The a_i are the KL coefficients that describe the uncorrected phase produced by Kolmogorov turbulence and k_i are the KL coefficients that describe the shape of the DM, i.e., the applied correction. Therefore, the measurable residual wavefront after correction $\varphi_{\epsilon\parallel}$ introduced in Equation (4.21) can be also expressed as:

$$\begin{aligned}\varphi_{\epsilon\parallel}(\vec{x},t) &= \sum_{i=1}^{N} \left[a_i(t) - k_i(t)\right] K_i(\vec{x}) \\ &= \sum_{i=1}^{N} \epsilon_i(t)K_i(\vec{x})\end{aligned} \tag{4.32}$$

The ϵ_i coefficients are noise free residual KL coefficients and must be related to the noisy measured residual coefficients $\hat{\epsilon}_i$ from Equation (4.27). The noise sources that affect the measured coefficients $\hat{\epsilon}_i$ are described by:

$$\hat{\epsilon}_i(t) = \epsilon_i(t) + n_i(t) + r_i(t) \tag{4.33}$$

The $n_i(t)$ term arises from detection noise, i.e., any process that affects the computation of the subaperture image shifts. In the case of a cross-correlating Shack-Hartmann WFS used for solar observations, this term is dominated by the error in the determination of the cross-correlation peaks. An analytical expression for this source of noise is derived in Appendix D.

The term $r_i(t)$ arises from aliasing and cross-coupling errors [36, 16]. The limited number and finite physical size of the subapertures in the WFS limit the spatial sampling of the WFS and introduce aliasing when measuring a wavefront distorted by atmospheric turbulence. Aliasing folds high spatial frequency components inside the sampling range of the WFS, making them appear to the WFS as lower frequency components (as discussed in Appendix E). The fact that the WFS measures the first derivative of the wavefront and not the wavefront itself, introduces cross-coupling between different spatial wavefront modes. Even though these modes are linearly independent, their derivatives are not necessarily so.

From the definition of the phase structure function in Equation (4.13) it follows that the phase structure function of the residual corrected phase component is written as:

$$D_{\varphi_{\epsilon\|}}(\vec{x}, \vec{\rho}) = \left\langle \left| \varphi_{\epsilon\|}(\vec{x}, t) - \varphi_{\epsilon\|}(\vec{x} + \vec{\rho}, t) \right|^2 \right\rangle \tag{4.34}$$

The KL decomposition of the AO-corrected phase component from Equation (4.32) is substituted into the previous relation:

$$D_{\varphi_{\epsilon\|}}(\vec{x},\vec{\rho}) = \sum_{i=1}^{N}\sum_{j=1}^{N}\langle\epsilon_i\epsilon_j\rangle\left[K_i(\vec{x}) - K_i(\vec{x}+\vec{\rho})\right]\left[K_j(\vec{x}) - K_j(\vec{x}+\vec{\rho})\right] \tag{4.35}$$

The term $\langle\epsilon_i\epsilon_j\rangle$ is the covariance of the noise free KL coefficients that describe the residual wavefront. The covariance of two random variables is defined below together with some useful properties:

$$\begin{aligned}
\operatorname{cov}(X,Y) &= \langle XY\rangle - \langle X\rangle\langle Y\rangle & (4.36)\\
\operatorname{cov}(X,Y) &= \operatorname{cov}(Y,X) & (4.37)\\
\operatorname{cov}(X+Z,Y) &= \operatorname{cov}(X,Y) + \operatorname{cov}(Z,Y) & (4.38)
\end{aligned}$$

The phase distortions introduced by atmospheric turbulence constitute a random gaussian variable with zero mean, and $\operatorname{cov}(\epsilon_i,\epsilon_j) = \langle\epsilon_i\epsilon_j\rangle$. These covariance values for all values of i and j form the elements of a matrix called the covariance matrix designated by $\mathbf{C}_{\epsilon\epsilon}$.

The measured KL coefficients $\hat{\epsilon}_i$ are related to the noise free ϵ_i according to Equation (4.33). The covariance of that relation can be obtained by considering the properties described above. The detection noise n_i is not correlated with the aliasing error r_i, since the aliasing error only depends on the orthogonal component of the wavefront phase $\varphi_{\mathrm{atm}\perp}$. Moreover, n_i is also not correlated with the noise free KL coefficients ϵ_i. The detection noise n_i is propagated through the AO loop. Thus the n_i measured on the current loop iteration is correlated to the ϵ_i from the previous iteration. This delay does not affect the estimates [100], since the noise n_i can be characterized as white noise. Thus, taking the covariance of Equation (4.33) produces:

$$\langle \hat{\epsilon}_i \hat{\epsilon}_j \rangle = \langle \epsilon_i \epsilon_j \rangle + \langle n_i n_j \rangle + \langle r_i r_j \rangle + 2\langle \epsilon_i r_j \rangle \tag{4.39}$$

$$\langle \epsilon_i \epsilon_j \rangle = \langle \hat{\epsilon}_i \hat{\epsilon}_j \rangle - \langle n_i n_j \rangle - \langle r_i r_j \rangle - 2\langle \epsilon_i r_j \rangle \tag{4.40}$$

The term $\langle \epsilon_i r_j \rangle$ is problematic because of its dependency on ϵ_i. However, it can be shown [100] that if the bandwidth of the AO system is large enough it can be approximated by:

$$\langle \epsilon_i r_j \rangle \approx -\langle r_i r_j \rangle \tag{4.41}$$

which produces:

$$\langle \epsilon_i \epsilon_j \rangle = \langle \hat{\epsilon}_i \hat{\epsilon}_j \rangle - \langle n_i n_j \rangle + \langle r_i r_j \rangle \tag{4.42}$$

The solar AO system has a working frequency of 2.5kHz that produces a correcting 0 dB bandwidth of ~130Hz, which fulfills the assumption of large bandwidth. An analytical expression for the covariance of the detection noise is obtained in Appendix D. It can also be directly measured from open-loop data captured by the AO system (as described in Section D.5). Because of the high levels of light available for wavefront sensing during solar observations, the noise covariance term $\langle n_i n_j \rangle$ can be safely neglected from Equation (4.42), as discussed in Appendix D. An expression for the covariance of the aliasing error $\langle r_i r_j \rangle$ is obtained in Appendix E.

The mean phase structure function of the residual parallel phase component $\bar{D}_{\varphi_{\epsilon\parallel}}$ is computed as:

$$\bar{D}_{\varphi_{\epsilon\parallel}}(\vec{\rho}) = \frac{\int P(\vec{x}) P(\vec{x}+\vec{\rho}) D_{\varphi_{\epsilon\parallel}}(\vec{x},\vec{\rho}) d\vec{x}}{\int P(\vec{x}) P(\vec{x}+\vec{\rho}) d\vec{x}}, \tag{4.43}$$

where P is the pupil function introduced in Section 4.2. Equation (4.35) is substituted into the previous expression producing:

$$\bar{D}_{\varphi_{\epsilon\parallel}}(\vec{\rho}) = \sum_{i=1}^{N}\sum_{j=1}^{N}\langle\epsilon_i\epsilon_j\rangle U_{ij}(\vec{\rho}) \tag{4.44}$$

where the U_{ij} functions are defined by:

$$U_{ij}(\vec{\rho}) = \frac{\int P(\vec{x})P(\vec{x}+\vec{\rho})\left[K_i(\vec{x}) - K_i(\vec{x}+\vec{\rho})\right]\left[K_j(\vec{x}) - K_j(\vec{x}+\vec{\rho})\right]d\vec{x}}{\int P(\vec{x})P(\vec{x}+\vec{\rho})d\vec{x}} \tag{4.45}$$

The U_{ij} functions can be more easily calculated by noting the definition of the cross-correlation of two functions:

$$\Gamma\left[f,g\right](\vec{\rho}) = \int f(\vec{x})g(\vec{x}+\vec{\rho})d\vec{x} \tag{4.46}$$

Note that Equation (3.10) presents the particular case of a cross-correlation of the same function, i.e., the autocorrelation $\Gamma\left[N,N\right] \equiv \Gamma_N$. The U_{ij} functions can be rewritten in terms of cross-correlation functions producing a much simpler expression:

$$U_{ij}(\vec{\rho}) = \frac{-\Gamma\left[K_iP,K_jP\right] - \Gamma\left[K_jP,K_iP\right] + \Gamma\left[K_iK_jP,P\right] + \Gamma\left[P,K_iK_jP\right]}{\Gamma\left[P,P\right]} \tag{4.47}$$

The cross-correlation functions are easily calculated in the Fourier domain where a correlation transforms into a regular product:

$$\mathcal{F}\left[\Gamma\left[f,g\right]\right] = \mathcal{F}\left[f\right]\mathcal{F}^{\star}\left[g\right] \tag{4.48}$$

The bulk of the computation of $\bar{D}_{\varphi_{\epsilon\parallel}}$ occurs in the calculation of the U_{ij} functions. However, the U_{ij} are functions of the telescope pupil geometry and the modal basis used in

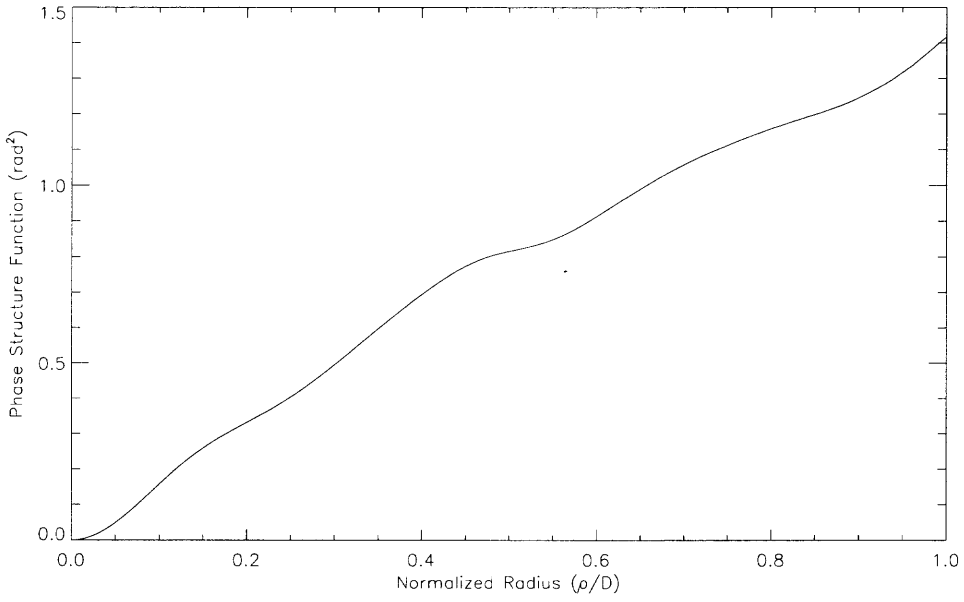

Figure 4.5 Azimuthal average of an example mean phase structure function of the parallel component $\bar{D}_{\varphi_{\epsilon\|}}$.

the phase decomposition. These are fixed parameters and thus, the U_{ij} functions only have to be computed once. Once calculated, they can be stored on disk and read whenever they are needed, which significantly speeds up the PSF estimation process.

The function $\bar{D}_{\varphi_{\epsilon\|}}$ is computed from the covariance of the residual KL coefficients $\langle \epsilon_i \epsilon_j \rangle$ and the U_{ij} functions. Figure 4.5 shows an example of $\bar{D}_{\varphi_{\epsilon\|}}$ obtained from real AO data. The residual corrected phase $OTF_{\epsilon\|}$ is then computed as:

$$OTF_{\epsilon\|}(\vec{\rho}/\lambda) = \exp\left[-\frac{1}{2}\bar{D}_{\varphi_{\epsilon\|}}(\vec{\rho})\right] \tag{4.49}$$

An example of the parallel residual $OTF_{\epsilon\|}$ calculated from real AO loop data are shown in Figure 4.6. The AO loop data were obtained during actual solar observations.

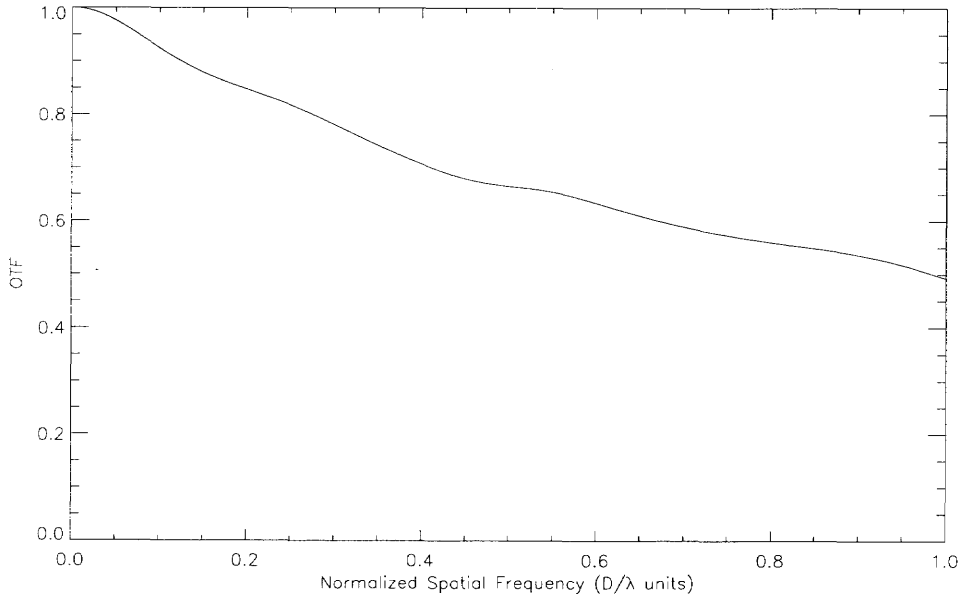

Figure 4.6 Azimuthal average of an example OTF from parallel phase component $OTF_{\epsilon\|}$.

4.4.2 Calculation of $\mathrm{OTF}_{\mathrm{atm}\perp}$

The orthogonal component of the residual wavefront $\varphi_{\mathrm{atm}\perp}$ travels through the AO system unaffected. Its contribution to the long-exposure OTF is characterized by the factor $OTF_{\mathrm{atm}\perp}$ in Equation (4.24). This phase component consists of high spatial frequency components that cannot be directly measured by the AO system. However, an estimation of the parallel component of the original phase distortions $\varphi_{\mathrm{atm}\|}$ can be obtained from the DM commands. The original atmospheric seeing conditions, characterized by the Fried parameter r_0, can be derived from $\varphi_{\mathrm{atm}\|}$ and used to extrapolate the properties of $\varphi_{\mathrm{atm}\perp}$.

Noll [52] applied the Kolmogorov model to derive an analytical expression for the covariance matrix of the Zernike coefficients C_Z that describe a wavefront distorted by atmospheric turbulence. This Zernike covariance matrix can be transformed into a covariance matrix of KL coefficients C_{KL}, since KL modes are just linear combinations of Zernike modes (see Appendix A). The transformation of a Zernike covariance matrix into a KL covariance matrix is derived from Equation (A.2), namely $\mathrm{C}_{KL} = \mathrm{U}^T \mathrm{C}_Z \mathrm{U}$.

The DM corrects the incoming wavefront phase distortions $\varphi_{\mathrm{atm}\|}$ by adapting its shape to reproduce the incoming phase distortions in a least squares sense [33]. The shape of the DM can be reconstructed from the DM commands and expressed in terms of KL coefficients. Its covariance matrix $\hat{\mathrm{C}}_{KL}$ can then be computed and fitted to the theoretical one C_{KL} given by Noll to obtain the value of r_0. Note that the DM shape replicates just half the amplitude of the phase distortions because the correction is applied by reflection on the DM. This factor of two must be accounted for when computing $\hat{\mathrm{C}}_{KL}$.

The commands sent to the DM are translated into voltages to be applied to the actuators, which push the DM surface proportionally to the voltage received. The shape of the mirror is approximated as a linear combination of the individual shapes of each actuator, called *influence functions*. Thus, the influence functions constitute a basis of the mirror space, i.e., the space of all the shapes the mirror can reproduce. From Equation (4.19) the wavefront correction introduced by the mirror, i.e., the DM shape, can be expressed as:

$$\varphi_m(\vec{x},t) = \sum_{i=1}^{A} m_i(t) I_i(\vec{x}) \tag{4.50}$$

where the coefficients $m_i(t)$ are proportional to the voltage received by each actuator in the DM and $I_i(\vec{x})$ are their influence functions. The influence functions were directly measured from the AO system by poking each actuator individually and measuring the resulting shape with an interferometer (see Figure 4.7). However, the shapes extracted from the interferometric measurements were affected by some noise so it was necessary to construct some synthetic influence functions with the same characteristics as the measured ones.

Two different types of synthetic influence functions were tested against the measured ones. Figure 4.8 shows a comparison between the synthetic and measured influence functions. The first type of synthetic influence function is a Gaussian fit while the second type is the result of combining four smaller Gaussians. Details about both types of synthetic influence functions can be found in Appendix B. The second type of synthetic influence

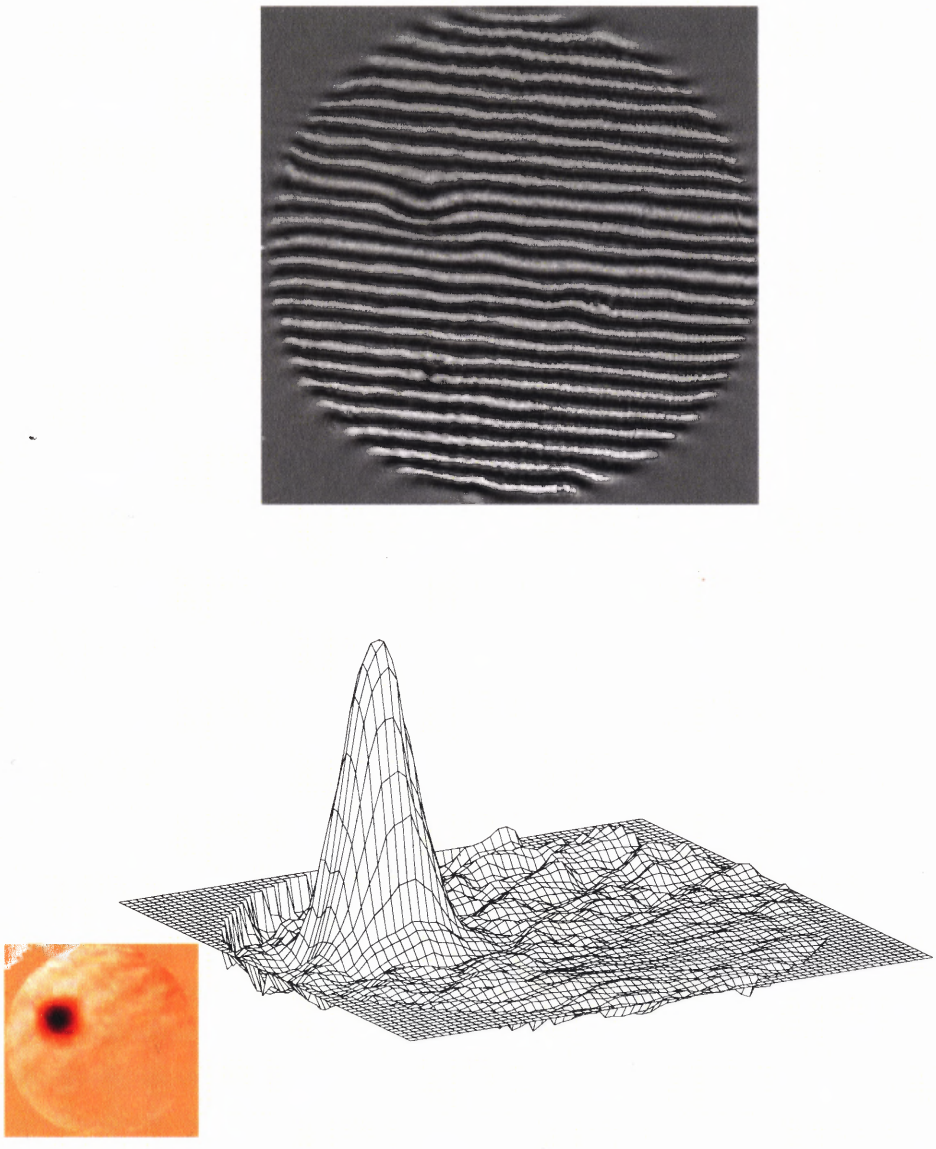

Figure 4.7 Influence function of a DM actuator measured with an interferometer. The top image shows the captured interferogram; the bottom image shows the shape of the influence function extracted from the phase encoded in the interferogram.

functions presents a slight square shape that is closer to the actual shape of the measured influence function. It also produces lower cross-talk between two contiguous influence functions, as illustrated in Figure B.5. The four Gaussian synthetic influence functions constitute a slightly better fit to the measured influence functions in a least-squares sense.

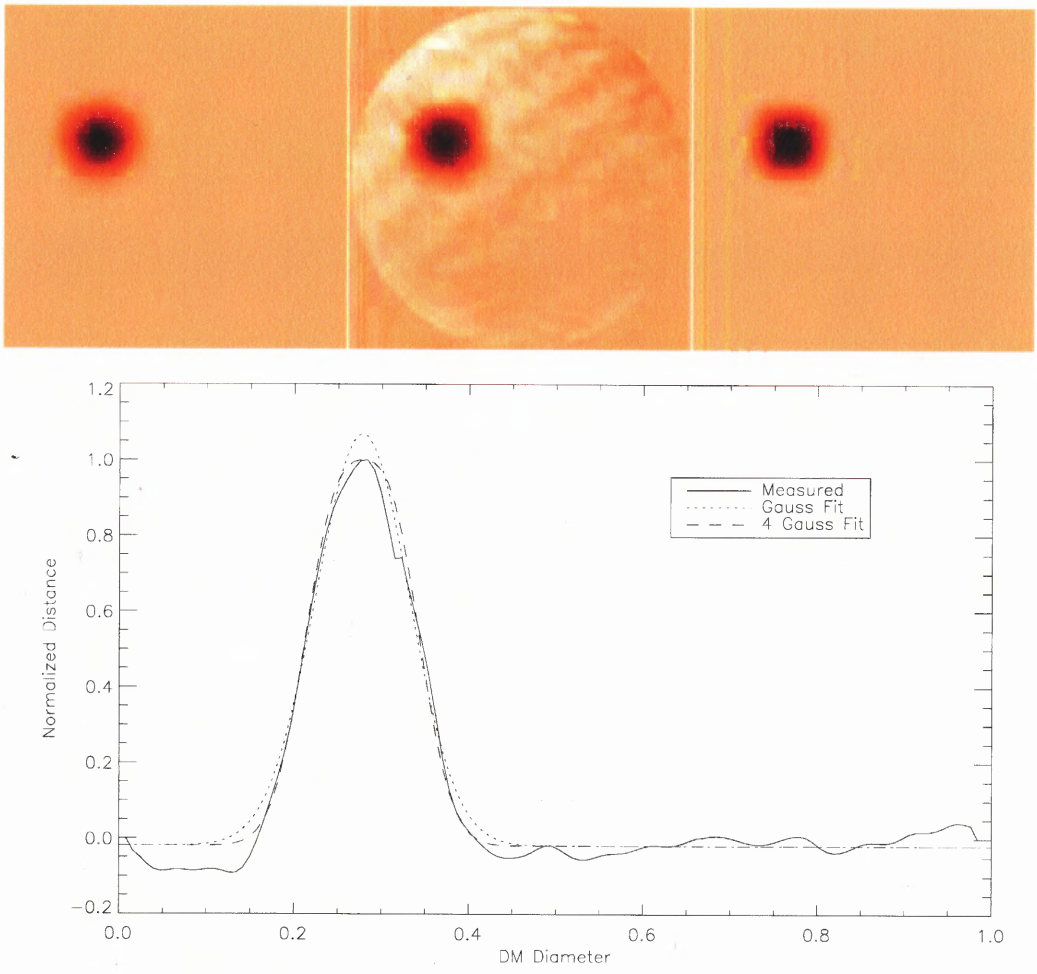

Figure 4.8 Synthetic and measured influence functions. Top: Gaussian fit (left), measured (middle) and four Gaussians fit (right). Bottom: horizontal cross-sections of all three influence functions.

The variance of the difference between two images provides a measurement of the degree of similarity between them:

$$D_{\text{image}} \equiv \text{image}_1 - \text{image}_2 \tag{4.51}$$

$$\sigma^2_{\text{image}} = \frac{1}{N-1}\sum^{N}\left[D_{\text{image}} - \bar{D}_{\text{image}}\right]^2 \tag{4.52}$$

The σ^2 values obtained from both sets of synthetic versus measured influence functions are: $\sigma^2_{\text{gauss}} = 0.0017$; $\sigma^2_{4\,\text{gauss}} = 0.0016$.

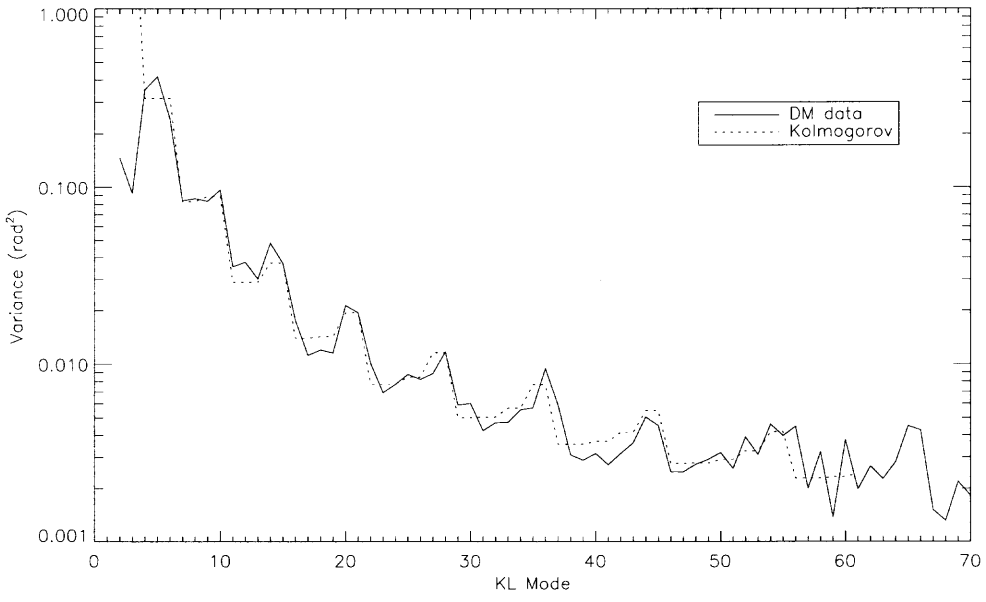

Figure 4.9 KL coefficient variance obtained from DM commands and fitted to the theoretical variance given by the Kolmogorov model ($r_0 = 16.1$ cm).

The DM commands stored by the AO system are the values that were passed to the DM controller. These values are measured in "counts" units and are proportional to the actual voltage applied to each actuator and to the deformation that the actuator produces. The values in "counts" must be related to the actual wavefront in radians through a calibration factor $k_{\rm dm}$ (see Appendix B).

The transformation from influence function coefficients m_i into KL mode coefficients k_i is computed as a change of basis operation. The DM commands constitute the coordinate representation of φ_m in the influence function basis. They must be transformed from their influence function representation into a KL representation. The change of basis is carried out with a projection matrix (B), which can be readily computed when both basis are known (see Appendix C). The KL coefficients k_i are obtained from the m_i in the following way: $\vec{k} = \mathbf{B}\,(k_{\rm dm}\,\vec{m})$. Again the calibration factor should be incorporated inside the matrix to reduce the number of multiplications: $\mathbf{B} \rightarrow \mathbf{B}\,k_{\rm dm}$.

$$\varphi_m(\vec{x},t) = \sum_{i=1}^{N} k_i K_i(\vec{x}) = \sum_{i=1}^{N} \left[\mathbf{B}\,\vec{m}(t)\right]_i K_i(\vec{x}) \tag{4.53}$$

The covariance matrix $\hat{\mathbf{C}}_{KL}$ of the KL coefficients k_i that describe the mirror shape can now be calculated from the DM commands m_i:

$$\hat{\mathbf{C}}_{KL} = cov\left[\mathbf{B}\,\vec{m}(t)\right] \tag{4.54}$$

The variance of the KL coefficients obtained from the mirror commands, i.e., the diagonal elements of the covariance matrix $\hat{\mathbf{C}}_{KL}$, is fitted to the variance derived from the Kolmogorov model, i.e., the diagonal elements of $\mathbf{C}_{KL}$ as given by Noll [52]. The value of the Fried parameter r_0 is obtained from the fit, as discussed later in this Section. Figure 4.9 shows an example of such a fit obtained from real AO data. The figure plots the KL coefficients variance calculated from the DM commands and its Kolmogorov model fit. It should be noted that the spikes of the Kolmogorov variance shown in the plot arise from the particular ordering of KL modes chosen, as described in Appendix A. It is also important to note that the variance obtained from the DM for modes 2 and 3 (tip and tilt), cannot be used in the fitting, since these modes are corrected by a separate mirror, the tip-tilt mirror, and therefore do not appear on the DM. Moreover, a finite outer scale L_0 introduces deviations from the results provided by the Kolmogorov model, which assumes an infinite outer scale. The variances of the global tip and tilt components are greatly reduced [33] by the effects of a finite outer scale, as illustrated by the power spectral density derived from the Von Karman model shown in Figure 3.1. Therefore, tip and tilt variances should not be considered in the fit to avoid the effects of a finite outer scale L_0.

The value of r_0 that was fitted from the DM commands stored by the AO system provides information about the characteristics of the seeing conditions during the exposure, see Section 3.2. This information can now be applied to obtain $OTF_{\mathrm{atm}\perp}$ by a similar

approach to the one taken in the calculation of $OTF_{\epsilon\|}$. The residual uncorrected phase, which is identical to the one produced by the turbulence, must be expressed in terms of KL modes:

$$\varphi_{\text{atm}\perp}(\vec{x},t) = \sum_{i=N+1}^{N_\infty} a_i(t) K_i(\vec{x}) \tag{4.55}$$

The previous sum should extend to infinity. However, for practical reasons the sum is truncated at an upper limit of N_∞. Using the definition of the phase structure function from Equation (4.13) and Equation (4.55), the residual uncorrected phase structure function can be expressed as:

$$\begin{aligned} D_{\varphi_\perp}(\vec{x},\vec{\rho}) &= \langle |\varphi_{\text{atm}\perp}(\vec{x},t) - \varphi_{\text{atm}\perp}(\vec{x}+\vec{\rho},t)|^2 \rangle \\ &= \sum_{i,j=N+1}^{N_\infty} \langle a_i a_j \rangle \left[K_i(\vec{x}) - K_i(\vec{x}+\vec{\rho})\right] \left[K_j(\vec{x}) - K_j(\vec{x}+\vec{\rho})\right] \end{aligned} \tag{4.56}$$

The covariance values $\langle a_i a_j \rangle$ are obtained from the Kolmogorov KL covariance matrix $\mathbf{C}_{KL}$ as described above. The Zernike covariance matrix $\mathbf{C}_Z$ given by Noll follows a scaling law with D/r_0, where D is the aperture size of the telescope. According to Equation (A.2), this same scaling law also applies to the KL covariance matrix:

$$\mathbf{C}_Z = [\mathbf{C}_Z]_{(D/r_0)=1} \left(\frac{D}{r_0}\right)^{5/3} \tag{4.57}$$

$$\Rightarrow \mathbf{C}_{KL} = [\mathbf{C}_{KL}]_{(D/r_0)=1} \left(\frac{D}{r_0}\right)^{5/3} \tag{4.58}$$

This scaling law provides the functional dependency of the KL covariance matrix with respect to the Fried parameter r_0, and can be used to fit r_0 from the $\hat{\mathbf{C}}_{KL}$ computed from the DM actuator commands. The scaling law in Equation (4.58) is applied to the

covariance elements from Equation (4.56) to produce the same scaling relation for $D_{\varphi_\perp}$ and $\bar{D}_{\varphi_\perp}$.

$$D_{\varphi_\perp}(\vec{\rho}) = [D_{\varphi_\perp}(\vec{\rho})]_{(D/r_0)=1} \left(\frac{D}{r_0}\right)^{5/3} \tag{4.59}$$

This implies that the mean structure function can be computed for the general case of $D/r_0 = 1$ and then be scaled for the correct value of D/r_0 obtained from the DM fit.

Several methods to compute $D_{\varphi_\perp}$ are available. Some methods attempt to generate simulated phase screens [100] that recreate the phase distortions produced by atmospheric turbulence. The lower modes in the modal decomposition of the phase screens are neglected, since they are assumed to be corrected by the AO system. What is left in each phase screen is a theoretical simulation of a possible instance of $\varphi_{\text{atm}\perp}$. A large set of different phase screens must be used to accurately simulate the statistics of real atmospheric turbulence and produce a reliable $D_{\varphi_\perp}$. These methods produce good results but require a very large number of phase screens and may need long computation times. The resulting structure function presents a sharp increase for small values of ρ. It oscillates around a saturation value for larger ρ values and, finally, it presents a sharp overshoot for very high ρ values close to the edge of the pupil [39, 100]. The presence of very high edges in the orthogonal structure function is a very characteristic result from these methods. The influence of these large edges is deemed to be negligible because they mostly affect spatial frequencies close to the diffraction-limited cut-off frequency where they are dampened by the amplitude of the diffraction-limited OTF. The large computation times of this method makes it not practical for these purposes.

Another suggested method [31, 33, 39] is based on an expression by Tatarski [89], which relates the phase structure function to the integral of the phase power spectrum. The phase power spectrum for uncorrected turbulence is derived directly from the Kolmogorov model (see Chapter 3). AO correction eliminates power from the low spatial frequency

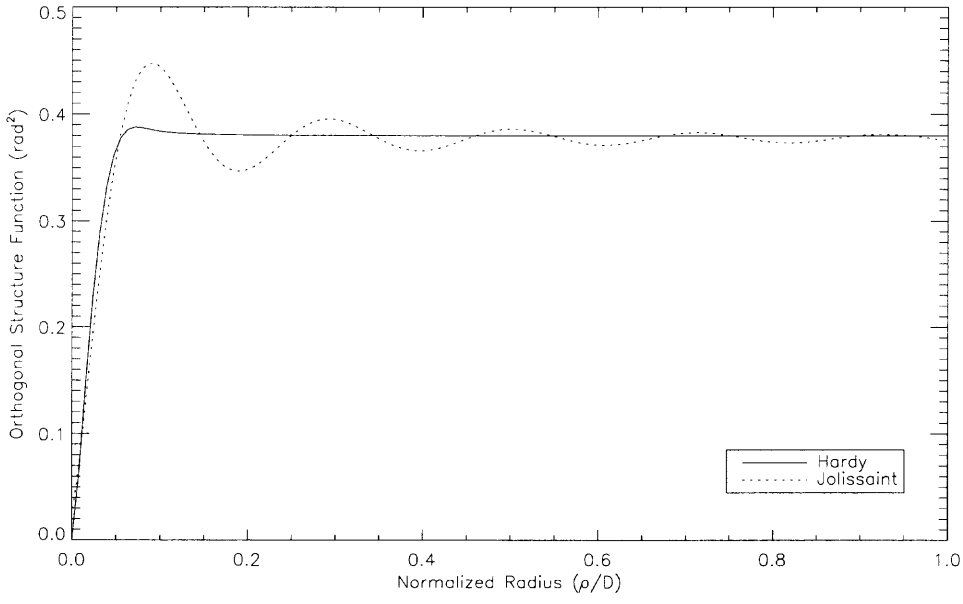

Figure 4.10 Azimuthal average of examples of orthogonal phase structure functions computed by two different algorithms given by Hardy [33] and Jolissaint [39].

components of the wavefront distortions. Thus, the effects of AO correction on the phase power spectrum are equivalent to the effects of a high-pass filter. This method requires very short computational times and completely avoids the edge effects that arise from the phase screens methods.

Two different high-pass filter functions used to compute the AO-corrected phase power spectrum are described by Hardy [33] and Jolissaint [39]. Jolissaint's algorithm produces oscillations in the structure function that replicate the oscillations produced by the phase screens method discussed earlier. A comparison of the orthogonal structure functions produced by both algorithms is shown in Figure 4.10.

The orthogonal phase structure function presents two asymptotic behaviors [33, 100]: for $\rho \ll \Lambda$ (where Λ is the mean actuator spacing) it behaves similarly to the uncorrected case $D_{\varphi_\perp} \propto \rho^{5/3}$; and for $\rho \gg \Lambda$ the structure function saturates at a value of $2\sigma_\perp^2$, where $\sigma_\perp^2$ is the variance of the residual uncorrected phase $\varphi_{\mathrm{atm}\perp}$. The phase variance $\sigma_\perp^2$ can be calculated by integrating the high-pass filtered atmospheric power spectrum [100]:

$$\begin{aligned}\sigma_{\perp}^{2} &= 0.023\,\frac{6\pi}{5}\,2^{5/3}\left(\frac{\Lambda}{r_0}\right)^{5/3} \\ &= 0.275\left(\frac{\Lambda}{r_0}\right)^{5/3} \end{aligned} \tag{4.60}$$

The orthogonal mean phase structure functions produced by both algorithms saturate at the correct value for high values of ρ. However, the one provided by Hardy's algorithm does not present oscillations around the saturation value. These oscillations are present in the orthogonal structure functions computed with the phase screen method. They seem to generate the typical sharp boundary between the halo and the core of the AO-corrected PSFs.

The computed orthogonal phase structure function $D_{\varphi\perp}$ can be inserted into Equation (4.61) to compute the orthogonal OTF component:

$$OTF_{\mathrm{atm}\perp}(\vec{\rho}/\lambda) = \exp\left[-\frac{1}{2}D_{\varphi_{\perp}}(\vec{\rho})\right] \tag{4.61}$$

Figure 4.11 shows an example of a $OTF_{\mathrm{atm}\perp}$ calculated from real AO loop data stored during solar observations.

4.4.3 Final Adaptive Optics Corrected OTF

The final OTF is the product of the three individual OTF components shown in Equation (4.24): the parallel component $OTF_{\epsilon\|}$, the orthogonal component $OTF_{\mathrm{atm}\perp}$ and the telescope OTF_{tel}. An azimuthal average of the final OTF calculated from real AO data using this method is shown in Figure 4.12.

The estimated long-exposure OTF shown in Figure 4.12 presents the characteristics to be expected in an AO-corrected OTF. This OTF can be compared with the uncorrected OTF shown in Figure 4.3, which presents very low amplitude at high spatial frequencies. The correction provided by the AO system preserves high spatial frequency components

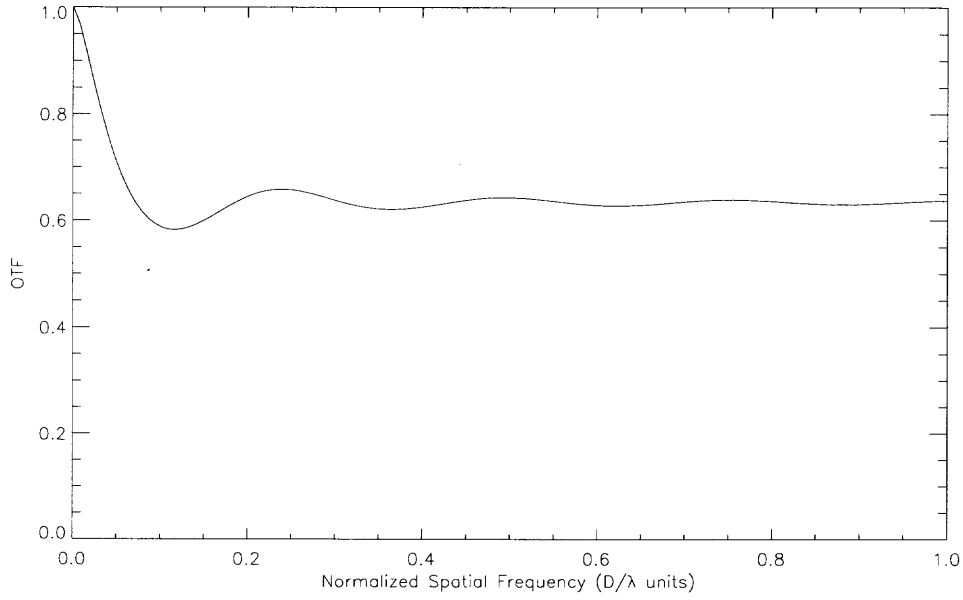

Figure 4.11 Azimuthal average of OTF from orthogonal phase component $OTF_{\text{atm}\perp}$.

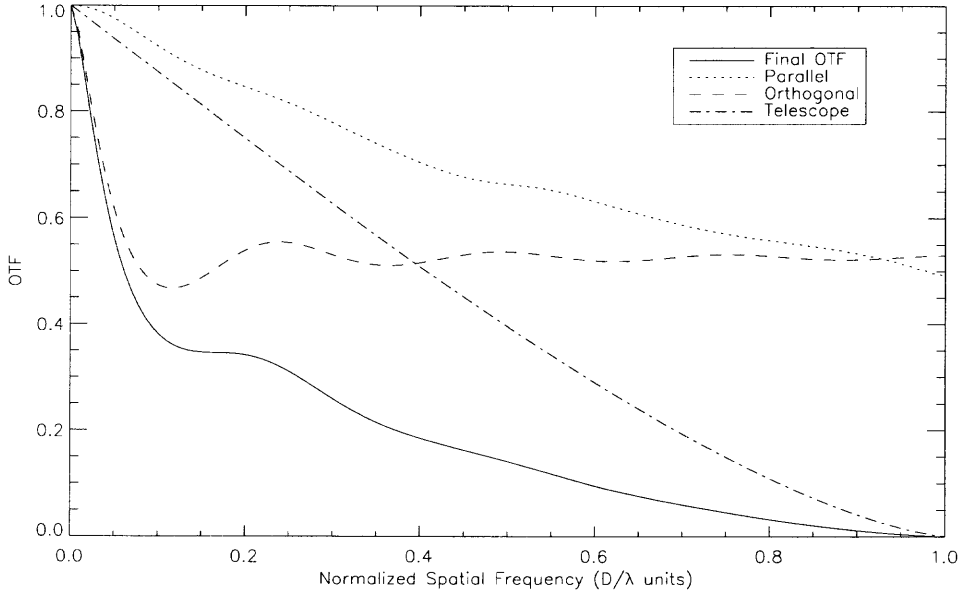

Figure 4.12 Azimuthal average of final AO-corrected OTF together with individual components. Computed from real data captured on August 2004 ($r_0 = 5.9$ cm).

in the image, which translates into an OTF with larger amplitude for high spatial frequen-

cies. The estimated long-exposure AO-corrected OTF from Figure 4.12 presents larger amplitudes over all frequencies up to the diffraction-limited cut-off frequency.

4.5 AO Loop Data Storage

The computation of the individual OTF components requires the loop data from the AO system. By default the AO system stores the complete time series of AO loop data produced during each exposure, WFS measurements, and DM commands. At a working frequency of 2500 frames per second, this can translate into large data files that can take a long time to be written to disk and can impact the cadence at which images can be taken. In particular, a 1s exposure produces $2500 \times 2 \times 76$ real numbers to account for the $x-$ and $y-$shifts measured by the 76 subapertures of the WFS plus 2500×97 real numbers required to store the actuator commands of the 97 actuator DM. Storing a real number in 4 bytes, the complete time series of AO data produces a file of size ~ 2.4 MB.

The long-exposure PSF estimation method described in this chapter only requires the covariance matrices of the AO loop data, not the complete time series. For this reason, the AO system controller was updated to compute the covariance matrices of both the WFS measurements and the DM commands in real-time. The system can be instructed to store only these covariance matrices, producing much smaller data files. For example, a 1 s exposure produces a covariance matrix of WFS measurements that can be stored in $(152\times152+152)/2$ real numbers plus $(97\times97+97)/2$ real numbers to store the covariance matrix of the DM commands. These estimations take into account that a covariance matrix is symmetric and only half the matrix needs to be stored. The data can be stored in a file of size ~ 65 kB, i.e., around 38 times smaller than the previous value.

It should be noted that the PSF estimation method requires the covariance matrices of the KL coefficients that describe the measured residual wavefront and the shape of the DM. The AO controller computes in real-time the covariance matrices of the pixel shifts

measured by the WFS $\hat{\mathrm{C}}_{ww}$ and the actuator commands $\hat{\mathrm{C}}_{mm}$ that are sent to the DM. These raw matrices must be transformed to the correct KL coefficients covariance matrices.

Equation (4.27) transforms pixel shifts measured by the WFS $\vec{w}$ into KL coefficients $\vec{\epsilon}$ that describe the residual wavefront using the reconstruction matrix (D^{+}). The reconstruction matrix can also be used to compute the covariance matrix of KL coefficients $\hat{\mathrm{C}}_{\epsilon\epsilon}$ from the WFS shifts covariance matrix $\hat{\mathrm{C}}_{ww}$:

$$\hat{\mathrm{C}}_{\epsilon\epsilon} = \mathrm{D}^{+}\,\hat{\mathrm{C}}_{ww}\,\mathrm{D}^{+T} \tag{4.62}$$

The DM commands stored by the AO system are transformed into KL coefficients that describe the shape of the DM by means of a change of basis matrix B, as shown in Equation (4.53). The covariance matrix of KL coefficients $\hat{\mathrm{C}}_{KL}$ is calculated from the covariance matrix of DM commands $\hat{\mathrm{C}}_{mm}$ as:

$$\hat{\mathrm{C}}_{KL} = \mathrm{B}\,\hat{\mathrm{C}}_{mm}\,\mathrm{B}^{T} \tag{4.63}$$

The AO loop data can be more efficiently captured by instructing the AO system to directly store the covariance matrices of the pixel WFS measurements and DM commands. These can be later transformed into the covariance matrices of KL coefficients that the PSF estimation method requires.

CHAPTER 5

SIRIUS OBSERVATIONS

Stars are very distant objects that cannot be resolved by the DST. They are effectively ideal point sources, which can be used to provide direct measurements of the PSF of the telescope. Night-time astronomers routinely point their telescopes away from the object of interest towards an isolated star to capture measurements of the PSF. This same approach can be employed to validate the accuracy of the PSFs obtained with the PSF estimation method. By capturing AO-corrected images of a star, the PSF estimated from the AO loop data can be compared against a direct measurement of the AO-corrected PSF.

Observations of the star Sirius were performed after sunrise, during the morning, at the DST on July 2005. The star images were corrected by the AO system and the AO loop data corresponding to each individual image were stored and used to estimate the AO-corrected PSFs. An interference filter with a bandpass of 7 nm FWHM centered at 616 nm was placed in front of the science camera.

The DST was designed for solar observations, which present very high light levels. It has a long f-ratio (f/72) with an aperture size of 76.2 cm, which make it less than ideal for stellar observations. Moreover, the AO system was designed for solar observations and some modifications were required to increase the light input on the wavefront sensor. To improve the chances of successful AO correction and increase the signal to noise ratio of the observations, a bright star was chosen as the target. The star Sirius [26] was chosen as the target for these observations because it is the brightest star in the night sky, with an apparent visual magnitude of -1.46. Furthermore, Sirius was above the horizon at the time of the observations. Sirius was observed in the morning after sunrise instead of during the night. This mode of operation was preferred because of the complications of operating the DST during the night, and because the atmospheric conditions would be similar to the

conditions during regular solar observations. The effects of an increased sky brightness in the observations were not expected to be significant.

The main difficulty that the AO system encounters when operating with low light levels is wavefront sensing. The WFS camera is a CMOS camera, with low quantum efficiency and high read out noise, which is not at all suited for stellar observations. During normal solar observations, the light levels are high and only a small fraction of the light is sufficient to provide adequate wavefront sensing. A beam splitter separates 5% of the light for wavefront sensing while allowing the remaining 95% to continue into the science path. This setup does not work with low light conditions as is the case with the stellar observations. For these observations, the output from the beam splitter was reversed, 95% of the light was sent on the WFS path while the remaining 5% was sent to the science camera.

To further increase the light available for wavefront sensing, the loop frequency of the AO system can be reduced to allow longer integration times for the WFS. However, this also reduces the correcting bandwidth that can be achieved by the AO system. During normal solar observations, the light levels are high and the AO system can run at its maximum loop frequency of 2.5 kHz. In general, for solar observations, the running frequency of the AO system is not limited by the light levels but by the technical characteristics of the AO system components. Under low light conditions, the AO loop frequency must be reduced to provide the WFS with longer integration times and increase the signal to noise ratio of the WFS measurements.

For the Sirius observations, the AO loop frequency was reduced in an attempt to improve the sensitivity of the WFS at the expense of correction bandwidth. Different loop frequencies were tested depending on the seeing conditions ranging from 800 MHz to 1.2 kHz. Better seeing conditions produced sharper star images and WFS measurements with a higher signal to noise ratio, which allowed for higher loop frequencies. A loop frequency of 1.2 kHz provided acceptable uninterrupted correction during good seeing con-

ditions, while more severe seeing conditions required a slower loop frequency and the AO loop frequency was lowered to 800 MHz to improve performance.

The lack of accurate absolute pointing of the DST made the initial finding of the star Sirius challenging. The tracking of the star Sirius during the observations was also problematic because the DST is not equipped with a star tracking system. The continuous adjustment of telescope pointing required to track the star had to be performed by hand by the telescope operator. The tip and tilt correction values displayed by the AO system while correcting in closed-loop were used as visual cues to assist the tracking.

5.1 Initial Results

The AO system was able to successfully lock on the star Sirius and provide continuous stable correction. AO-corrected images of Sirius with exposure times of 2.5 s were captured together with their corresponding AO loop data. Figure 5.1 shows two of the early results obtained from the observing run. The image with full AO correction (right panel in Figure 5.1) presents an elongated shape. This elongation is present in all of the captured AO-corrected images, which seems to indicate oscillations of the separate tip-tilt correction. It was discovered that the gain settings for the tip-tilt mirror were not adequate for stellar observations and needed to be readjusted.

The elongation present in the corrected images was corrected by adequately adjusting the gains of the tip-tilt mirror controller. The resulting images appeared round with no significant elongation along a particular axis, as shown in Figure 5.2. However, under closer inspection the width of the images was found to be larger than the diffraction-limited width that should be expected. While this could be a consequence of the lower correction achievable under these extreme conditions, i.e., low light level and reduced system frequency, it could also indicate a problem with the tip-tilt correction sub-system while operating on a star.

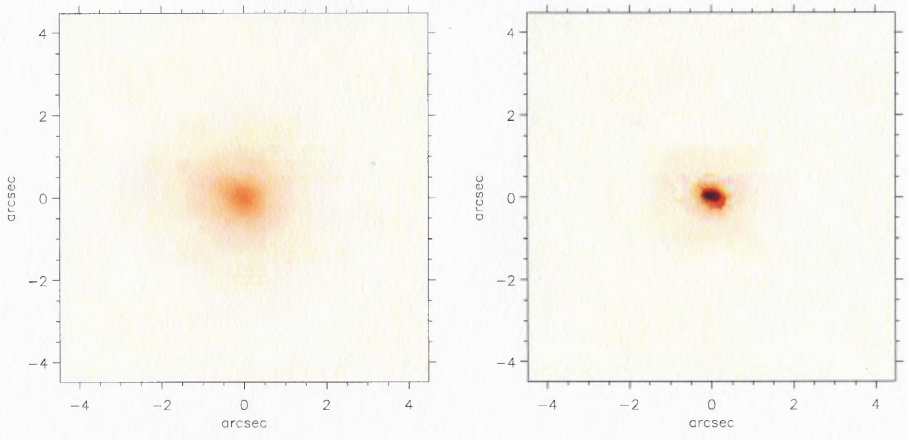

Figure 5.1 Captured AO-corrected images of Sirius. Tip-Tilt correction only (left) and full AO correction (right). AO-corrected image presents clear elongation.

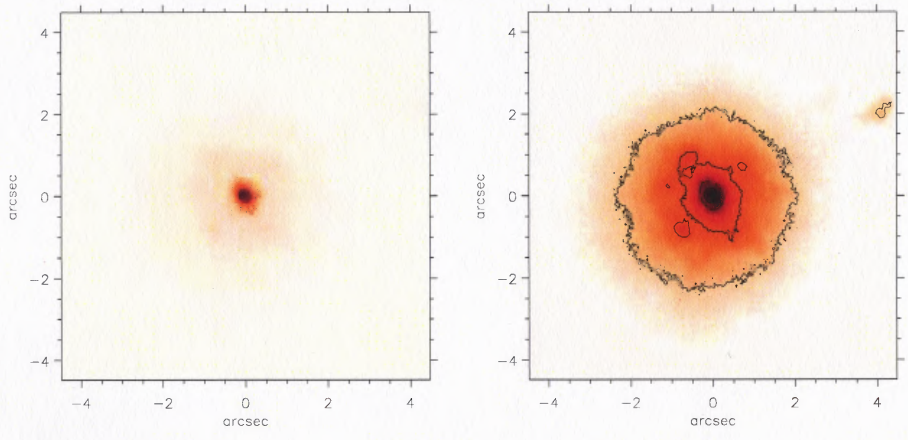

Figure 5.2 AO-corrected image of Sirius after elongation issues were corrected showing a typical example of a waffle pattern. Linear scale (left) and logarithmic stretch (right).

Figure 5.2 illustrates another very important correction issue that was detected during the stellar observations. The images from Figure 5.2 display a static pattern of four small peaks distributed on the vertices of a square around the central star image. This pattern is called *waffle pattern* and it is evidence for the presence of waffle modes introduced by the DM. The pattern is significant enough to be visible in both the linear and logarithmic stretch images.

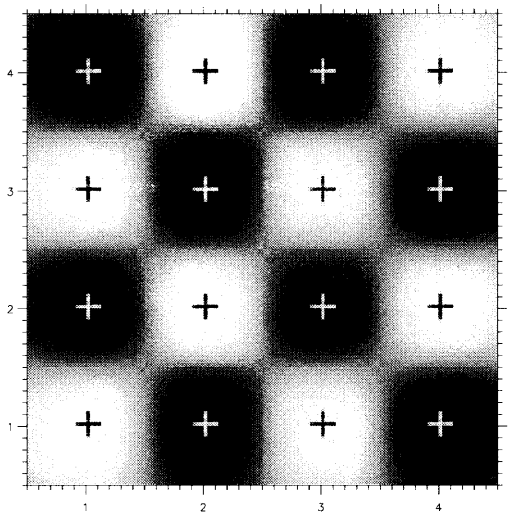

Figure 5.3 Schematic representation of the fundamental waffle mode. Actuator positions are marked with crosses.

Waffle modes are DM configurations that cannot be detected by a Shack-Hartmann WFS with a square array of subapertures. They typically resemble a checkerboard pattern, which produces a zero mean wavefront slope inside each WFS subaperture. A complete list of waffle modes can be found in Gavel 2003 [27]. The simplest waffle mode, schematically illustrated in Figure 5.3, is created by the alternating pushing and pulling of contiguous DM actuators. In Figure 5.3, white and black represent positive and negative wavefront distortion, and the crosses indicate the position of the DM actuators, which coincide with the corners of the WFS subapertures. Figure 5.3 shows that the mean slope produced by the waffle mode inside each WFS subaperture is zero. Hence, this mode is invisible to the WFS.

The distribution of light at the focal plane of an optical system is given by the Fourier transform of the wavefront at the pupil plane [28]. Therefore, the characteristic waffle pattern displayed in Figure 5.2 is the result of the Fourier transform of the periodic checkerboard structure of the waffle modes [45].

The waffle modes must be actively removed or attenuated from the reconstructed wavefront by the AO control system. If they are not adequately filtered, noise from the measured residual wavefront may be reconstructed as waffle modes that are sent to the DM. Since they are not visible to the WFS, they may continue to build up over time dominating the shape of the DM introducing wavefront errors and reducing the correcting range of the DM. The control matrix used by the AO control system must be conditioned to filter and suppress any invisible modes from the reconstructed DM shapes.

The WFS measurements during the stellar observations present a much higher level of noise than during solar observations, which translates into a faster build up and a greater incidence of waffle errors. This is illustrated by the fact that most of the AO-corrected star images display some amount of waffle pattern. The time evolution of the waffle patterns observed can be summarized as: a particular waffle pattern that builds up quickly and remains more or less static for several consecutive images until it suddenly disappears and allows a new slightly different waffle pattern to emerge. This temporal evolution of the waffle error seems to arise from the AO system being stopped between exposures. Every time the AO system is stopped, the DM is flattened and the AO controller is reset causing any waffle mode to disappear. When the loop is closed again, new invisible waffle modes build up in the DM producing different patterns. The frequent AO system stops were dictated by the extreme conditions of the observations.

As a positive side effect, the waffle pattern can be used to verify the measured plate scale of the science camera. As indicated above, the simplest waffle pattern arises from the DM adopting a checker-board pattern with a period of 2Λ, where Λ is the DM inter-actuator spacing and the WFS subaperture size. The waffle mode of the DM produces a characteristic waffle pattern [45] in the focal plane. Replicas of the central PSF appear at the intersections of grid lines located a distance $n\lambda/2\Lambda$ from the optical axis (where $n = 1, 3, 5 \ldots$). The star image shown in Figure 5.2 displays four small peaks around the main central PSF that define four of such grid lines. Thus, the distance between each pair

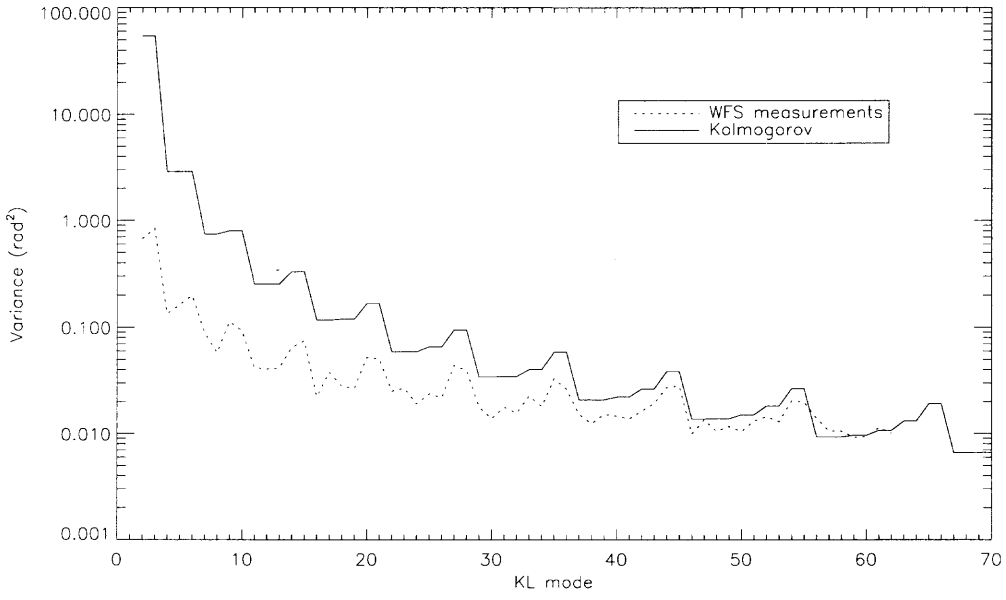

Figure 5.4 KL coefficient variance from the residual WFS measurements and from the atmospheric distortions as given by the Kolmogorov model for a value of $r_0 \sim 4$ cm.

of grid lines is λ/Λ. The comparison between the measured distance from the image, in pixels, and the theoretical distance λ/Λ was used to corroborate the value of the camera image scale, of 0.″042 /pixel.

During normal solar observations, the waffle pattern is very difficult to observe in extended field images and has gone undetected in previous operations. A lower incidence of waffling is expected during solar observations due to the much lower (several orders of magnitude, see Figure 5.5) noise levels of the WFS. However, stellar observations proved very useful for detecting performance issues of the AO system and more stellar observations will be performed in the future to evaluate the performance of control matrices and the AO system.

Figure 5.4 shows the residual modal variances computed from the WFS measurements together with the atmospheric modal variances. Figure 5.4 illustrates the correction provided by the AO system. The seeing conditions were not very good ($r_0 \simeq 4$ cm). The total variance contained in the first 60 modes of the atmospheric variance is 123.5 rad^2

while the total residual variance in the first 60 modes is 3.6 rad^2. The total residual variance measured by the WFS can be used in Equation (1.1) to obtain an estimate of the Strehl ratio of the AO-corrected image, which results in less than 0.02. This very small Strehl ratio indicates that the correction provided by the AO system was not very good. This is to be expected because of the challenges that the AO system faces when correcting the star Sirius, such as the low signal to noise ratio of the WFS measurements and the reduced correcting bandwidth of the system.

The noise affecting the WFS measurements during stellar observations is quite significant and must be taken into account in the PSF estimation method. The WFS noise is several orders of magnitude larger than the WFS noise observed during solar observations. These larger noise levels are caused by the low light levels available for wavefront sensing. Figure 5.5 shows two power spectra computed from the time series of pixel shifts in the x direction measured by a single WFS subaperture. The solid line was obtained from closed-loop WFS data captured during solar observations while the dotted line corresponds to open-loop WFS data captured during stellar observations. The difference in frequency range between both plots is due to the AO system running at a lower loop frequency (1.2 kHz) during the stellar observations. The high frequency tail of the power spectra provides information about the variance of the white noise affecting the measurements (see Section D.5). The standard deviation of the wavefront error introduced by the noise levels obtained from this particular data set, are measured at 3 nm for the solar data versus 114 nm for the star data (standard deviation). This corresponds to a noise $\sim$40 times larger in the stellar observations than in the solar observations. As mentioned previously, for solar observations the noise affecting WFS measurements can be safely neglected in the PSF estimation algorithm. This is clearly not the case for stellar observations. The noise in the WFS measurements is much larger and must be taken into account. The PSF estimation code used for stellar observations was modified accordingly to estimate and subtract the noise covariance that affects the WFS measurements, as described by Equation (4.42).

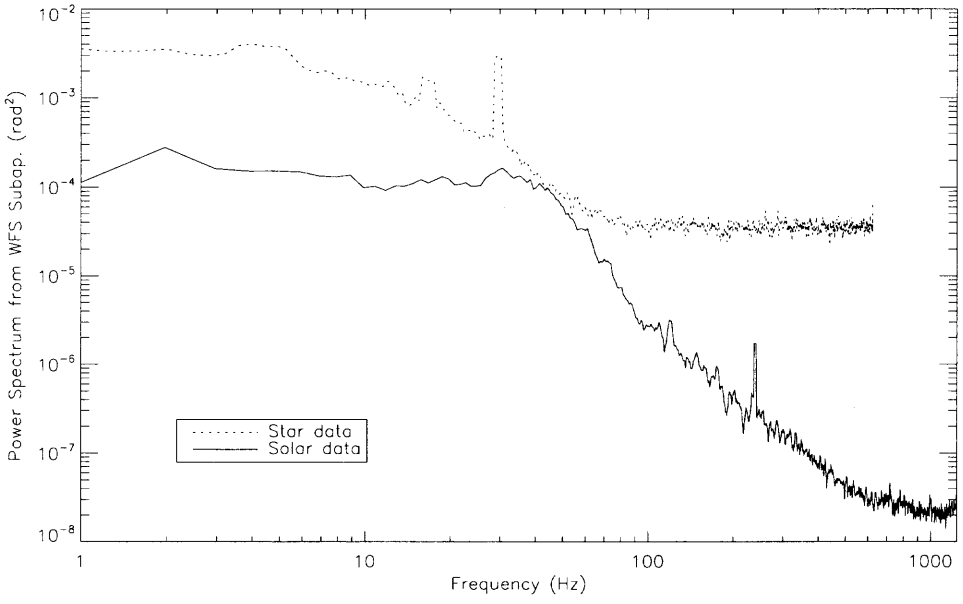

Figure 5.5 Power spectrum of WFS x-shifts measurements. The solid line corresponds to closed-loop data captured during solar observations and the dotted line to open-loop stellar data. During stellar observations, the system's loop frequency was reduced to 1.2 kHz.

5.2 PSF Estimation Challenges

The AO system was synchronized with the DST camera control system to capture data at the same time. Exposures of 2.5 seconds were taken at a cadence of 3 seconds. This setup only allowed 0.5 seconds in between exposures for the data to be written to disk. While this interval was sufficient for the DST camera control system to store each image, it proved too short for the AO control computer. The AO control computer could successfully store AO data for only one out of every two exposures. Thus, for every two Sirius images stored by the DST camera control system only one AO loop data file was available. This problem could have been easily solved but it went unnoticed while the observations were taking place. Missing half of the AO loop data files effectively halved the amount of useful data obtained during the observations and introduced uncertainty about the correct synchronization of the data.

5.2.1 Impact of Limited Bandwidth

The estimation of r_0 from the DM commands was affected by the reduced working frequency of the AO system. With a reduced loop frequency, the AO system is unable to drive the DM fast enough to completely follow the changing distortions introduced by atmospheric turbulence. The atmospheric distortions changed before the DM had time to reach the shape required to correct them. This translates to a smaller variance of the DM commands that would produce overestimated values of r_0. A simple simulation was constructed to illustrate this point. A sinusoidal wave with a frequency similar to the frequency of the turbulence (around 40 Hz) is used to simulate the changing conditions of atmospheric turbulence. This signal is followed by a systems that emulates the behavior of the AO system working at a particular frequency. At each iteration, the system attempts to keep its state as close as possible to the value of the signal by updating its state proportionally to the difference between the value of the atmospheric signal and its current state. A gain coefficient of $g = 0.25$ is used as a proportionality factor. Thus, the state of the systems at a time $t + 1$ is computed as: $S_{t+1} = S_t + g\,(A_t - S_t)$, where A_t represents the value of the atmospheric signal at a time t and S_t is the state of the system. Figure 5.6 shows the results obtained when running the simulated system at frequencies of 2.5 kHz and 800 Hz.

When estimating r_0 from the DM actuator commands, the assumption is that the variances of the modal coefficients that describe the shape of the DM are the same as the atmospheric ones. The variance of the original atmospheric signal in the simulation is 0.5, while the variance measured by the system running at 800 Hz is 0.23. The system running at 2.5 kHz produced a much better result with a variance of 0.45. By running the system at 800 Hz with a gain of 0.25 the variance of the signal was underestimated by 54%. Thus, in the case of an AO system running at 800 Hz, the value of r_0 estimated from the DM actuator commands will be overestimated. The value will be overestimated by a certain factor that depends on the gain settings, the ratio between the frequency of the atmospheric distortions and the working frequency of the AO system.

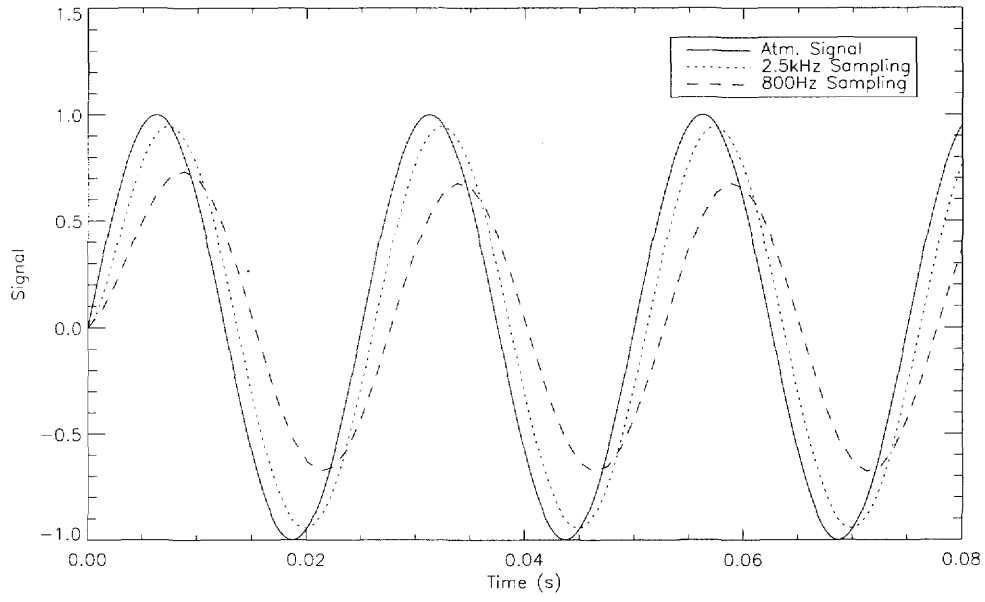

Figure 5.6 Simulated AO system following a sinusoidal signal at different loop frequencies.

Another consequence of the reduced correcting bandwidth affects the assumption of large bandwidth from Equation (4.41). Running at less than a third the normal frequency and with very high noise levels invalidates this assumption and the term $\langle \epsilon_i r_j \rangle$ cannot be easily computed [29]. The full expression of this term is given by Véran [100].

5.2.2 Calibration

The calibration factors for the WFS and the DM measurements are very important for the correct estimation of the residual wavefront and the seeing conditions. These calibration factors relate the stored raw AO system data, such as WFS measurements in pixels and DM commands in counts, to actual wavefront values. A set of theoretical calibration factors (shown in Table 5.1) can be computed from the characteristics of the AO system components (see Appendix B).

More reliable calibration factors could ideally be measured directly from the AO system. Two attempts to do this were performed in January and August 2005. A known

Table 5.1 Adaptive Optics Calibration Factors

Type	*WFS Factor*		*DM Factor*	
Measured January 2005	17.7 18.9 18.3	Mean: 18.3	0.022 0.017 0.016	Mean: 0.018
Measured August 2005	20.4 21.8 17.9 22.5	Mean: 20.6	0.023 0.025 0.020 0.024	Mean: 0.023
Theoretical	21.1		0.027	

amount of tilt was introduced into the AO system and the response of the WFS and the DM to this known tilt was recorded. This procedure is described in detail in Appendix B. Several samples of calibration measurements were obtained during both attempts. However, the results varied significantly across all samples for both attempts, as shown in Table 5.1. This indicates some issues with the measuring procedure such as the possibility of tilt saturation in the AO system or some other problems not yet identified. A more careful study of the procedure or a completely new approach will be necessary in the future.

5.3 Comparison of Sirius Images to Estimated PSFs

A PSF was obtained from the AO loop data for each individual star image. The star images and the estimated PSFs are both normalized to a peak value of one to eliminate any intensity dependencies from the comparison. The seeing conditions and the correction performance of the AO were not very good for the whole data set. Figure 5.7 shows one of the best images and its estimated PSF. The Strehl ratio of the star image was measured as 0.24 using a software package provided by M. van Dam [65]. The calibration factors were the

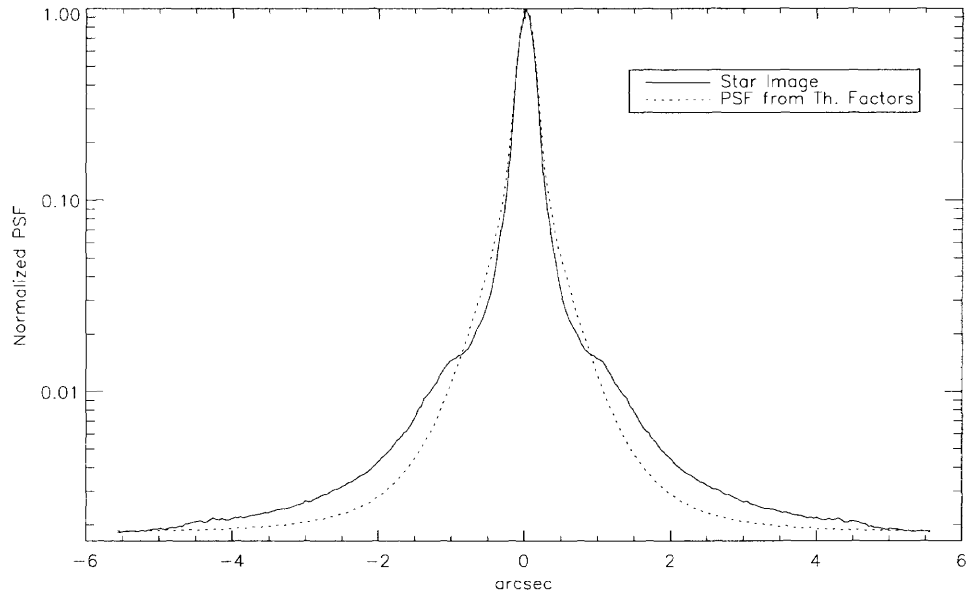

Figure 5.7 Azimuthal average of an AO-corrected image of the star Sirius and its estimated PSFs. The estimated PSF was calculated using theoretical calibration factors.

theoretical factors from Table 5.1. The value of r_0 was estimated to be between 4 and 7 cm and the Strehl ratio of the estimated PSF is around 0.20.

The width of the central core of the estimated PSF corresponds very well with the width of the star image. The FWHM of the star in the image is 0.″266 and the FWHM of the estimated PSF is 0.″269. The width of the central core is controlled by the residual aberrations measured by the WFS. Hence, the theoretical WFS calibration factor constitutes a good estimate. The halo of the PSF does not provide a good fit to the halo of the star image.

Figure 5.8 shows the same star image with two PSFs estimated using the calibration factors measured from the system in January and August 2005, given in Table 5.1. The Strehl ratios of the PSFs estimated using the measured calibration factors (Figure 5.8) are 0.25 and 0.17, respectively. The estimated r_0 values are 12.2 cm and 9.1 cm.

The PSFs produced by these two sets of measured calibration factors do not provide a good fit to the star image. It is to be expected that the measured calibration factors would

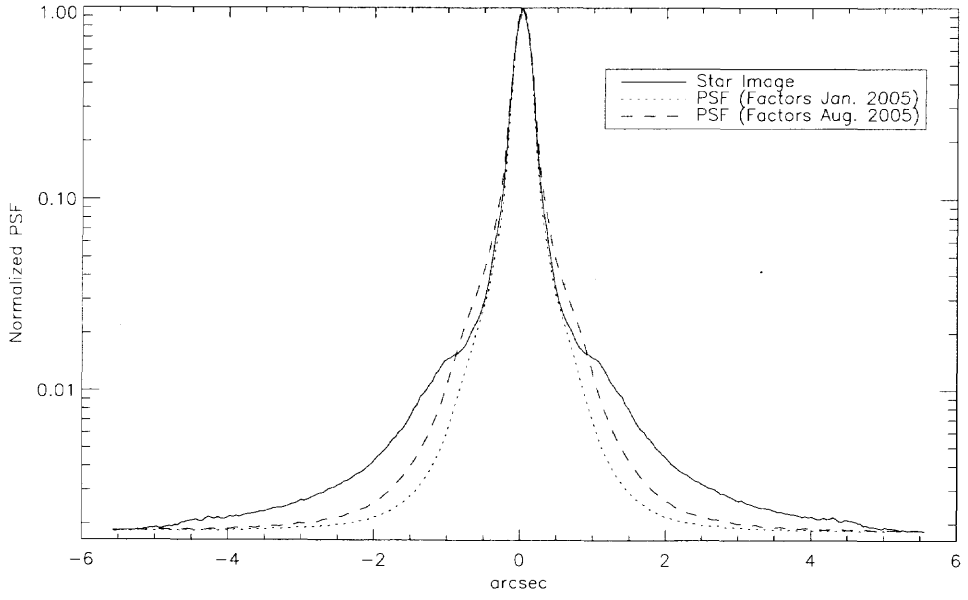

Figure 5.8 Azimuthal average of an AO-corrected image of the star Sirius and two different estimated PSFs calculated using different calibration factors that were measured directly from the AO system on January 2005 and August 2005.

not produce good results, judging by the large spread between the successive measurements for each attempt, as seen in Table 5.1.

As discussed in the Section 5.2, the extreme conditions under which the AO system was operating, produced an overestimation of the value of r_0 and invalidated the assumption of large bandwidth. The PSF estimation code was modified to compensate and account for these effects in the covariances of the WFS measurements and the variance of the DM commands. The DM calibration factor was artificially increased by a factor of 1.5 and a fixed amount was subtracted from the computed WFS measurements covariance. The results of these artificial corrections of estimated PSFs using the theoretical calibration factors are shown in Figure 5.9. The Strehl ratio of this new PSF is 0.16. Its FWHM was measured as 0.″259 and the value of r_0 was estimated as 4.5 cm.

The modifications to the PSF estimation code, which attempt to account for the effects of the lower bandwidth, seem to provide a better fit to the star image. However,

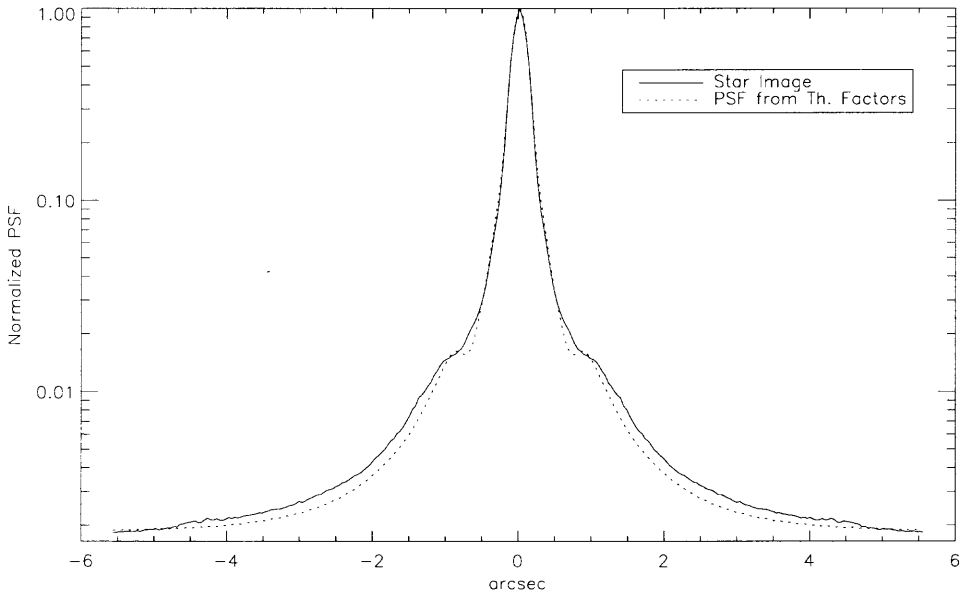

Figure 5.9 Azimuthal average of an AO-corrected image of the star Sirius and its estimated PSFs. The estimated PSF was calculated using theoretical calibration factors. The PSF estimation algorithm was adapted to compensate for the lower correcting bandwidth of the AO system during the stellar observations.

when these same modifications were tested on another star image, the estimated PSF did not produce an adequate fit to the stellar observations, as illustrated by Figure 5.10. These modifications provided a closer match between some of the star images and the estimated PSFs. However, more studies are required to completely understand the effects of low light levels and low bandwidth on the PSF estimation method.

5.4 Conclusions

The observations took place during the day, in the mid-morning, and the seeing conditions were not very good during the entire observing run. Despite the adverse seeing conditions, the low light levels and the lower bandwidth, the AO system was able to lock on the star Sirius and provide continuous correction. However, the Strehl ratios of the AO-corrected images were extremely low. The highest Strehl in the complete time series was 0.20.

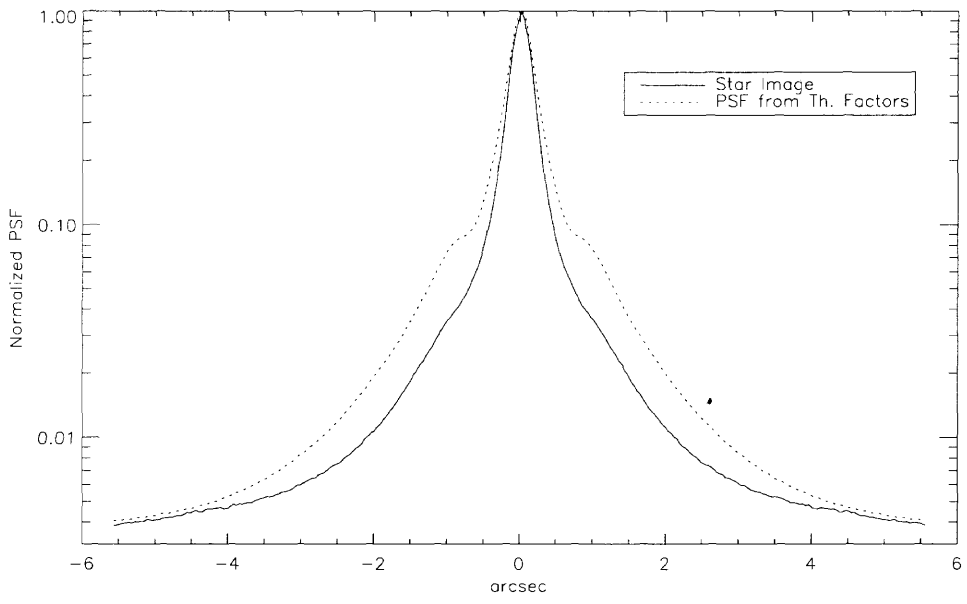

Figure 5.10 Azimuthal average of an AO-corrected image of the star Sirius and its estimated PSFs. The estimated PSF was calculated using theoretical calibration factors. The PSF estimation algorithm was adapted in the same way as in Figure 5.9.

The chip in the WFS camera of the AO system is a CMOS chip with a very low quantum efficiency (less than 0.20) and a very high read out noise (50-70 electrons). This camera performs well during solar observations, under which the WFS camera receives a very large number of photons and provides a signal with a very good signal to noise ratio. However, it was poorly suited for the stellar observations. The noise in the WFS measurements, which was several orders of magnitude larger than in the case of solar observations, reduced the performance of the AO system.

The AO system's residual errors (discussed in Section 2.2) were large during the stellar observations, which reduced the performance of the AO correction. In particular, the bandwidth error was significantly increased by the reduced working frequency of the AO system. During normal operation, the AO system runs at 2.5 kHz, while during the stellar observations, the frequency was reduced to 800 Hz most of the time to allow for

longer integration periods for the WFS camera. This reduction in frequency lowered the correcting bandwidth and increased the bandwidth error by a factor of ten.

The PSF estimation algorithm was unable to produce adequate PSF estimates from the AO loop data that fit the stellar observations. Some critical assumptions made by the PSF estimation method were violated by the extreme conditions of the observations. Further studies are required to completely understand the effect of the low Strehl ratios produced by the bad seeing, the high background light, the low light conditions and the low system bandwidth on the PSF estimation method. It may be worth repeating the observations of a star under more favorable conditions. For example, the observations could take place during the night, i.e., with much lower background light levels and seeing conditions that are usually better. Also, using a more sensitive CCD camera with low read-out noise as the WFS detector would significantly reduce the WFS noise. Alternatively, a different independent method to estimate the PSF, such as phase diversity or speckle reconstruction, could be attempted in the future.

CHAPTER 6

SUNSPOT OBSERVATIONS

The long-exposure PSF estimation method was applied to solar observations. The object of the observations was to obtain high-resolution velocity measurements to study the characteristics of the Evershed effect [22] in the penumbra of a sunspot. The scientific results of this study were published by Rimmele & Marino [57] and discussed in Chapter 7. The loop data produced by the AO system while correcting each of the captured images were stored and used to estimate its long-exposure PSF. Each image was deconvolved with its own long-exposure AO-corrected PSF.

6.1 Experimental Setup

The observations were performed on May 2004 at the 76 cm DST in Sunspot, New Mexico. A relatively round sunspot (NOAA 0605) positioned at 12 South, 23 West at a position angle of $\cos\Theta = 0.86$ was observed. The tunable Universal Birefringent Filter (UBF) was used to obtain a time sequence of dopplergrams and narrow band filtergrams. The UBF [3, 86] was tuned into the red and blue wings (offset ± 0.01 nm) of the spectral line Fe I 557.6 nm, which in the penumbra forms at an altitude of about 240km above the solar surface. The "non-magnetic" Fe I 557.6 nm line (effective Landé-factor $g = 0$) provides a Doppler signal without any cross-talk from the magnetic field. At this wavelength, the UBF has a passband of about 0.015 nm. Two additional cameras were set up to record images simultaneously in two additional wavelengths: G-band (430.5 nm) and Ca K-line (393.3 nm). An schematic drawing of the experimental setup is shown in Figure 6.1 and a picture showing the UBF optical path and the AO bench is found in Figure 6.2.

UBF filtergrams were recorded with a 1k$\times$1k CCD camera manufactured by Spectral Instruments. The optical layout provided an image scale of $0.''025$/pixel, i.e., the images are

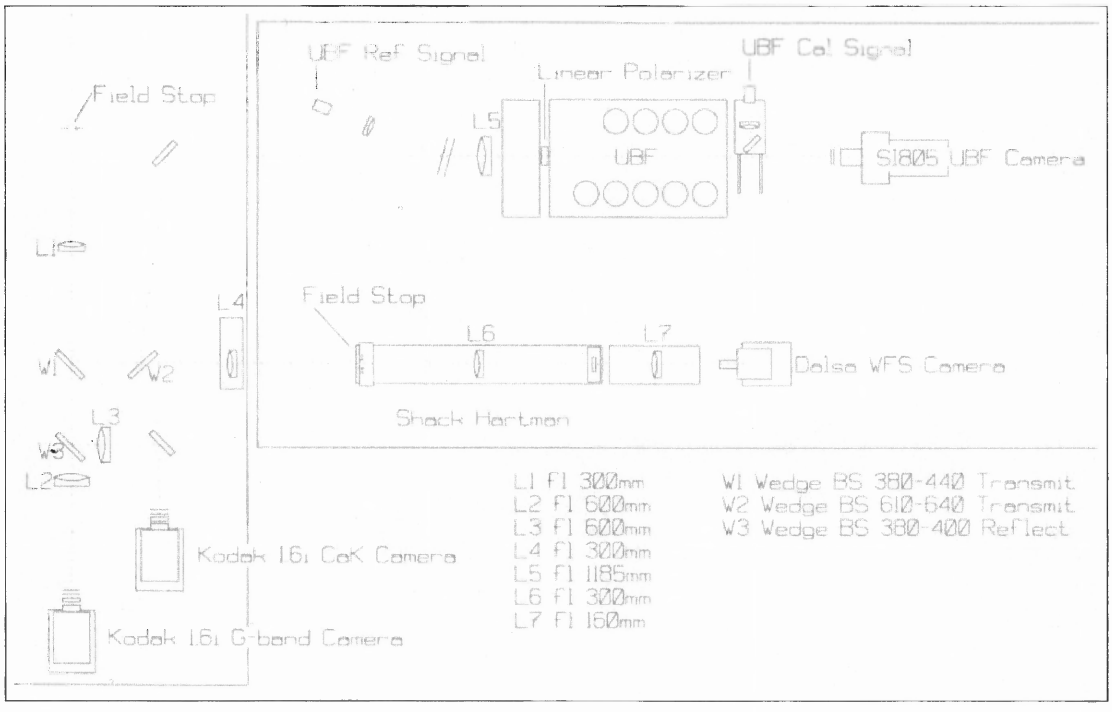

Figure 6.1 Drawing showing an schematic layout of the experimental setup.

highly oversampled. The diffraction limit at 557.6 nm is $\lambda/D = 0.''15$. The field-of-view of the UBF was approximately 25″.

The G-band and K-line images were captured with two Kodak MegaPlus 1.6i CCD cameras. The optical layout for both cameras produce a pixel scale of $0.''034$/pixel.The G-band images were obtained using a 1 nm wide filter centered at 430.5 nm. The filter used for the K-line images was a 0.3 nm wide filter centered at 393.3 nm.

The high-order AO system was deployed during the observations to correct atmospheric distortions in real-time. The AO loop data produced by the AO system during the capture of each image were stored to be used for PSF estimation.

The images were captured using the DST camera control system, which acts as a centralized controller for all the subsystems involved in the observations and coordinated the operation of the UBF and the science cameras. The AO system is set up as an independent system and at the time of the observations there were no mechanisms in place to synchronize the AO system with the DST camera control system. The AO system con-

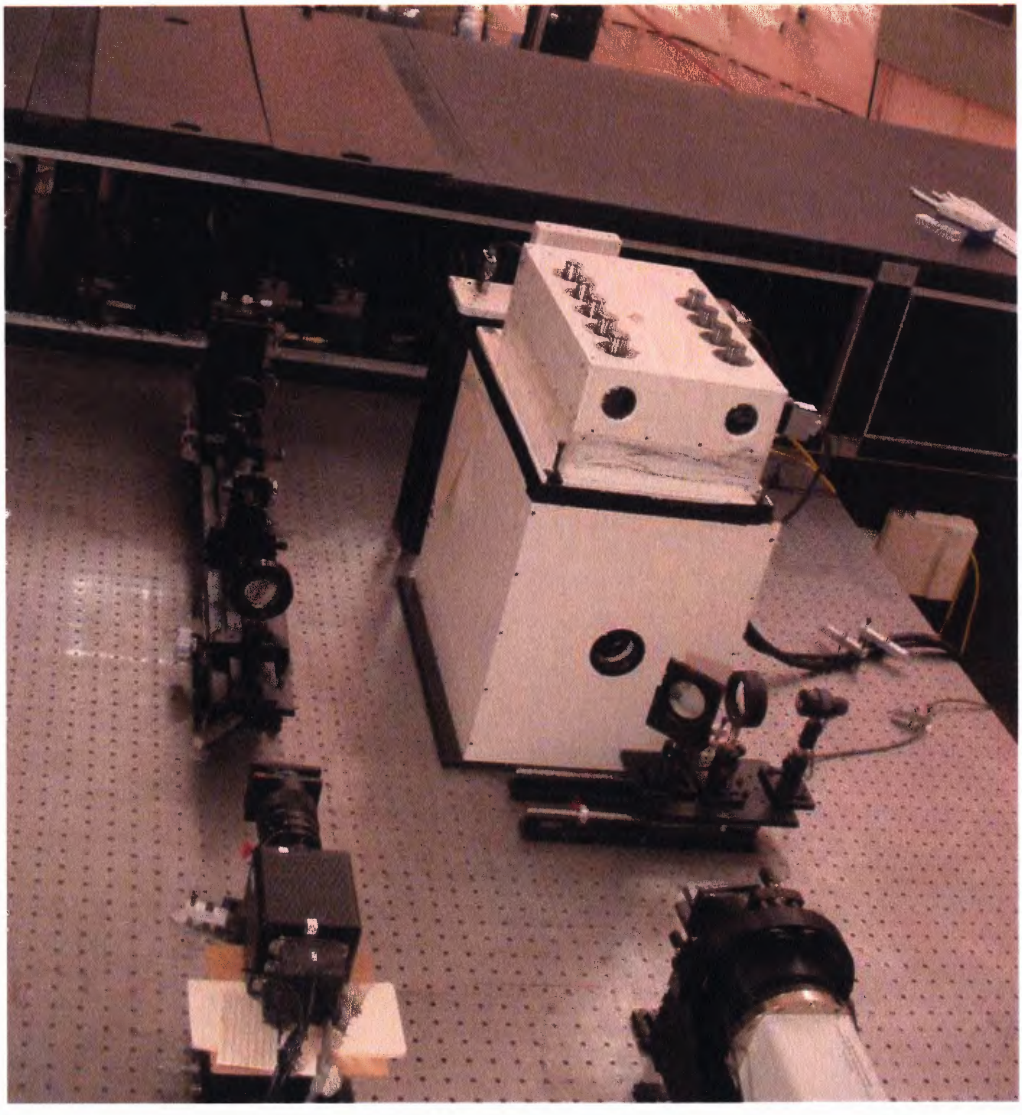

Figure 6.2 Picture of the experimental setup at the DST. The picture shows the UBF and the science camera located at the output from the AO system.

troller was upgraded to accept a TTL signal that would trigger the recording of loop data. The capability of sending such a signal was added to the camera control system creating a mechanism of synchronization between the DST camera control system and the AO system. The camera control system raises the TTL signal when an exposure is initiated, and lowers it at the end of the exposure. The AO system receives this signal and starts storing

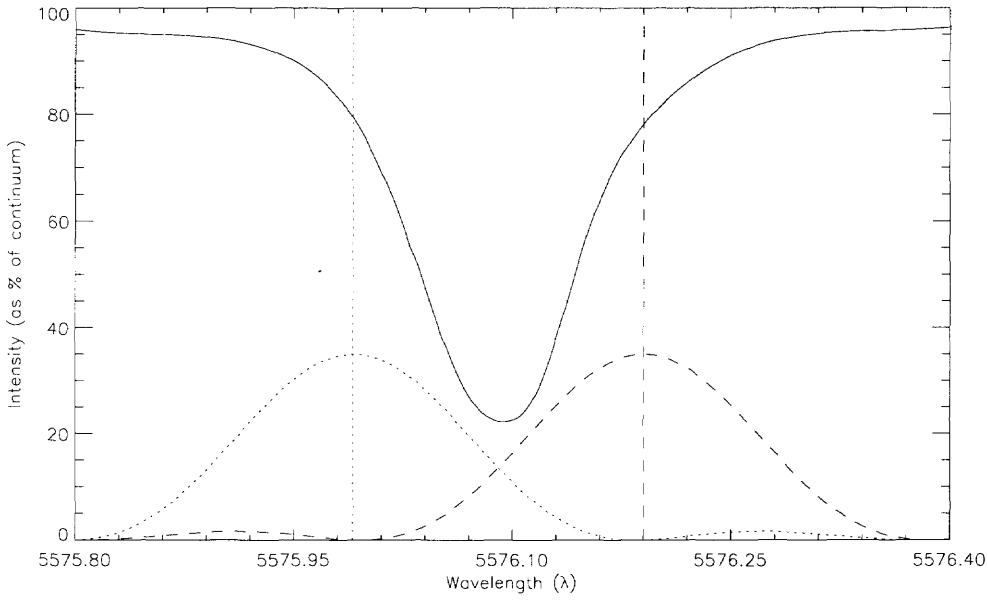

Figure 6.3 Solar spectrum around Fe I line (557.609 nm) as given by the Liège atlas of the solar spectrum (solid line). The band-pass of the UBF on both sides of the wing is shown (dotted and dashed lines).

data until the signal is lowered. The TTL signal sent by the DST camera control system ensures that the AO data are captured strictly simultaneously with the image exposure.

The steps to be performed during the observations were specified with a script that the DST camera control system executed. The camera control system tuned the UBF to 557.619 nm and captured an image together with a G-band and K-line image. Then, the UBF was tuned to 557.599 nm and another image was captured together with a G-band and K-line image. This process was repeated for as long as the seeing conditions allowed, producing an uninterrupted time series of measurements. The AO system stored the AO loop data produced during each exposure.

Figure 6.3 shows the solar spectrum around the Fe I absorption line as given by the Liège solar spectral atlas [17]. Superimposed on the image are the theoretical band-passes of the UBF filter tuned on the sides of the wing. The wavelengths 557.599 nm and

557.619 nm are located on the blue and red wings of the Fe I line centered at 557.609 nm, as shown in Figure 6.3.

The Fe I 557.6 nm is an absorption line formed in the solar photosphere by neutral iron atoms. The wavelength at which the absorption takes place changes according to the Doppler effect. The subtraction of the intensity recorded from the red and blue wing images produces a signal proportional to the line of sight velocity of the absorbing material.

A dopplergram encodes velocity information as intensity information by subtracting the red wing image from the blue wing image, as shown in Equation (6.1). A plasma with a line of sight velocity coming towards the observer shifts the absorption line towards the blue, making the red wing image brighter than the blue wing image. This situation produces a negative signal in the dopplergram. Negative values in the dopplergram are displayed as dark grays and black, indicate gas that is moving towards the observer (blue shift). While positive values, displayed as light grays and white, indicate gas that is moving away from the observer (red shift). The values in the raw dopplergram are normalized intensity differences. They must be calibrated to reflect velocity measurements in meters per second. This is done with the calibration factor c_v. This factor is computed in the following way: the theoretical shift produced by a known line of sight velocity is calculated. This shift is applied to the solar spectrum obtained from the Liège atlas. The intensity produced by each wing image is integrated from the product of the theoretical bandpass of the UBF and the shifted spectrum; the factor c_v is obtained by comparing the difference between both normalized wing intensities and the value of the known velocity.

$$\text{velocity} = c_v \times \frac{I_{\text{blue}} - I_{\text{red}}}{I_{\text{blue}} + I_{\text{red}}} \tag{6.1}$$

Each pair of wing images is not captured simultaneously. There is an interval of a few seconds between consecutive wing exposures. In general, the seeing conditions during the capture of each wing image may be different, producing exposures with different image

quality, which is commonly the case during day time observations. Since the wing images in a pair are subtracted to obtain velocity measurements it is very important that the quality of each wing image is of equivalent magnitude. Thus, PSF estimation for each individual image and subsequent image restoration are critical for the computation of a dopplergram.

The AO system correction and wavefront sensing are conjugated to the entrance pupil of the telescope because that is where the most important turbulent layer is located during the day time. Other turbulent layers also introduce distortions in the image, which are not corrected by the AO system. These higher altitude distortions will be different in both wing images because they are captured at different times. The differential distortions between both wing images are corrected using a de-stretch algorithm before they are subtracted. This de-stretching is made possible by the very similar solar structures present in both wing images due to the large width of the UBF bandpass.

6.2 Point Spread Function Estimation

Exposure times ranging from 1.0 to 1.5 seconds were used during these observations. Figure 6.4 shows an example of raw AO loop data captured during a 1s exposure with Strehl ratio of 0.38 and seeing conditions with $r_0 = 9.4$ cm. The plot shows a time series of pixel shifts in the x direction measured by a particular WFS subaperture and a time series of the commands sent to a particular actuator of the DM.

The WFS measurements are affected by detection noise as described in Section 4.4.1. However, during solar observations, the large number of photons available for wavefront sensing produce WFS measurements with very high signal to noise ratio, as discussed in Appendix D and illustrated by Figure 5.5. Thus, for solar observations the noise covariance term $\langle n_i n_j \rangle$ from Equation (4.42) can be safely neglected in the PSF estimation algorithm.

The PSF for each individual exposure is estimated from their corresponding AO loop data. The raw AO loop data are translated into real wavefront data with the help of the calibration factors that relate pixel shifts in the WFS and counts in the DM to a wavefront

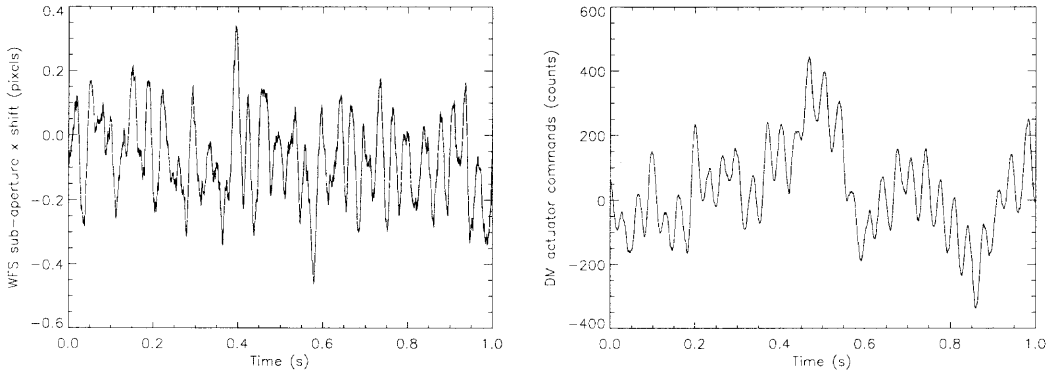

Figure 6.4 Time series of x-shifts measured by one WFS subaperture (left) and commands sent to one DM actuator (right).

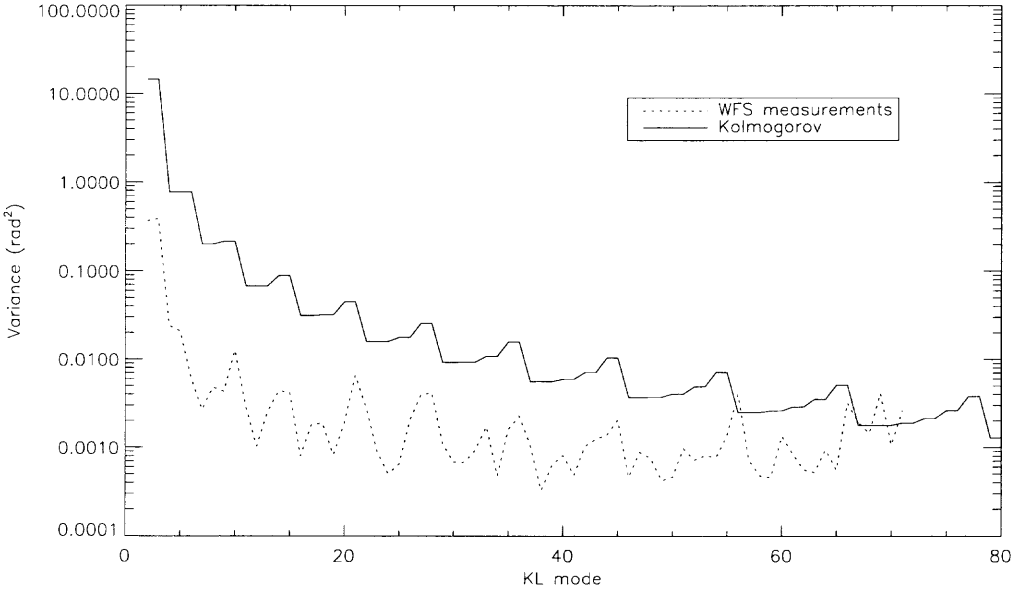

Figure 6.5 Residual KL coefficient variance obtained from WFS measurements. Compared to theoretical atmospheric variance given by the Kolmogorov model. The Fried parameter for these data is $r_0 = 9.5$ cm.

in the telescope's pupil plane. As discussed in Chapter 5, the calibration factors greatly influence the final estimated PSF and must be accurately measured from the system.

The covariances of the raw AO system loop measurements are transformed into modal covariances with the help of the reconstruction matrix D^+ and the projection ma-

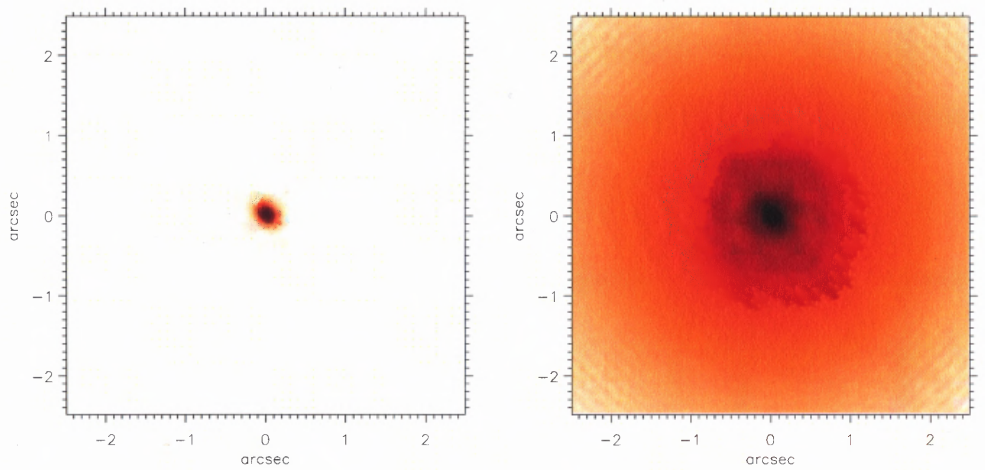

Figure 6.6 PSF estimated from real AO loop data captured on May 6th 2004. Left and right images show the same PSF displayed on a linear and logarithmic scale, respectively.

trix B, described in Equations (4.62) and (4.63). Figure 6.5 shows an example of modal variance of residual WFS measurements compared to the theoretical atmospheric variance before correction. These data were stored during seeing conditions of r_0=9.5cm (fitted from the DM commands), which resulted in a corrected image with a Strehl ratio of 0.4.

Figure 6.5 illustrates the correction provided by the AO system, which according to the plot, extends to KL mode 65. The AO correction provides a significant reduction in wavefront phase variance. The total wavefront phase variance contained in the first 70 KL modes, including tip and tilt, before correction is 33.16 rad^2, while after correction is reduced to 0.92 rad^2.

The estimated long-exposure PSF is obtained from the inverse Fourier transform of the OTF, which is expressed as the product of three independent components, as described in Equation (4.24). An example of a PSF estimated from AO data is displayed in linear (left panel) and logarithm (right panel) scale in Figure 6.6. The azimuthal average is shown in Figure 6.7. This PSF is computed from the AO data displayed in Figure 6.5. It illustrates the typical structure of an AO-corrected PSF: a narrow central core superposed on a wide seeing limited halo. With perfect correction, the full width at half maximum (FWHM) of

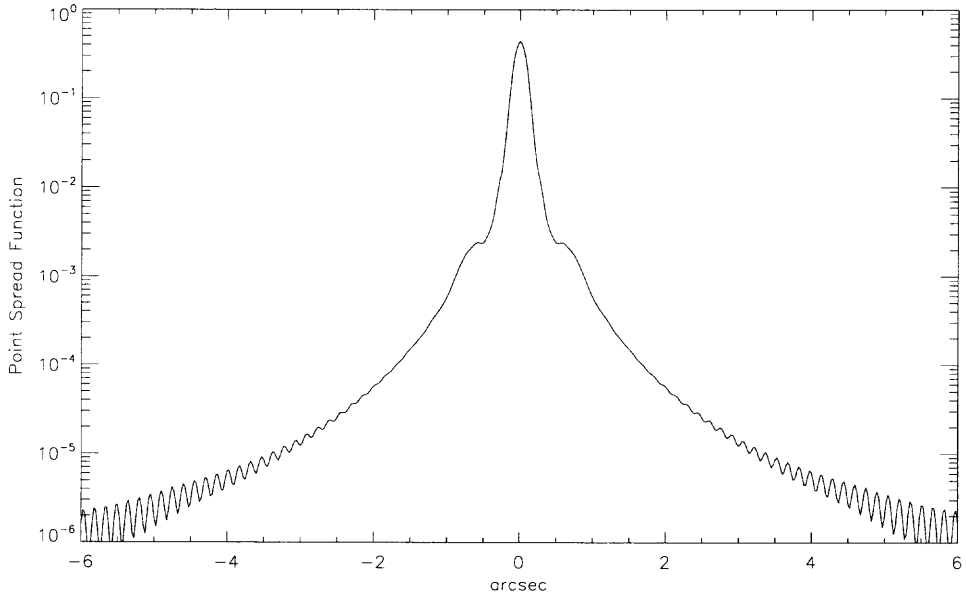

Figure 6.7 Azimuthal average of PSF shown in Figure 6.6.

the central core is the same as the FWHM of the diffraction-limited PSF, which at this wavelength (557.619 nm) is $0.''151$. However, the correction is not perfect and some residual aberrations remain after correction that increase the FWHM of the PSF core. The FWHM of the central core of the estimated PSF in Figure 6.7 is $0.''194$.

Hardy [33] provides an expression to calculate the FWHM of the central core of the AO-corrected PSF a_c from the variance of the residual tip-tilt σ_α^2:

$$a_c = \frac{\lambda}{D}\sqrt{1 + 5.17\left(\frac{D}{\lambda}\right)^2 \sigma_\alpha^2} \tag{6.2}$$

The variance σ_α^2 refers to the variance of the angle α shown in Figure B.1, while the variances of the residual WFS measurements, shown in Figure 6.5, are KL coefficient variances that describe the phase of the wavefront. The variance of the modal coefficient for tilt $\sigma_{k_2}^2$ is related to the tilt variance σ_α^2 with the following expression:

$$\sigma_{k_2}^2 = \left(\frac{2\pi}{\lambda}\frac{D}{A_t}\right)^2 \sigma_\alpha^2 \tag{6.3}$$

The factor A_t represents the amplitude (peak to valley) of the tilt mode, which for the KL tilt mode is 3.83. This expression can be substituted in the expression given by Hardy producing:

$$a_c = \frac{\lambda}{D}\sqrt{1 + 5.17\left(\frac{A_t}{2\pi}\right)^2 \sigma_{k_2}^2} \tag{6.4}$$

The tilt residual variance $\sigma_{k_2}^2$ can be obtained from the WFS measurements, shown in Figure 6.5, and can be used in the previous equation to produce an estimation of the FWHM of the central core of the PSF $a_c = 0.''198$. This result is comparable to the FWHM measured from the PSF estimated from AO data.

The level of correction that the AO system provides depends strongly on the seeing conditions. Adverse seeing conditions produce more severe wavefront distortions, which stress the performance of the AO system and reduce the correction provided. The dependency of AO correction on seeing conditions is illustrated by Figure 6.8, which plots the Strehl ratio of AO-corrected images versus the Fried parameter r_0. The data from the plot were obtained from a long times series of solar observations.

The residual wavefront variance after correction arises from several error sources discussed in Section 2.2. The residual wavefront error introduced by these sources depends on the design characteristics of the AO system and on the seeing conditions. Bad seeing conditions (low values of r_0) will produce higher wavefront variances from these sources and increase the overall residual wavefront variance after correction. This higher residual wavefront variance lowers the Strehl ratio of the image, see Equation (1.1), and reduces the quality of the correction.

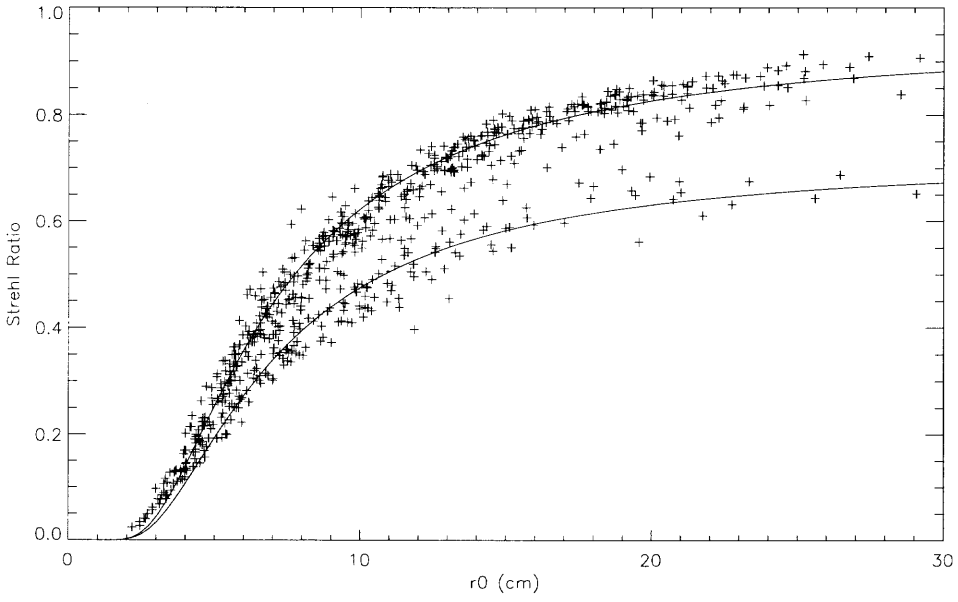

Figure 6.8 AO-corrected Strehl ratio obtained from the estimated PSFs versus the Fried parameter r_0.

The relationship between the Strehl ratio and r_0 in Figure 6.8 presents two distinct trends that correspond to different AO performance regimes. The higher trend in the figure indicates a higher AO performance that is able to provide Strehl ratios of 0.9 with very good seeing conditions ($r_0 \sim 30$ cm). The lower trend corresponds to a lower AO performance that can only provide Strehl ratios of 0.65 with the same very good seeing conditions. Two different theoretical Strehl profiles were fitted to the data shown in Figure 6.8. Both lines were obtained from Equation (1.1), substituting the wavefront variances arising from the different AO system error sources (see Section 2.2). The only difference between both fitted lines is a different estimated WFS noise variance added to the global residual variance. The top fit in the figure was obtained by adding a WFS noise variance of 0.06 rad^2, while the bottom fit was obtained by adding a variance of 0.33 rad^2. The appearance of two different additive variances affecting the residual wavefront variance after correction is not yet understood. A possible explanation involves the wavefront sensing errors in-

troduced by anisoplanatism [99] arising from high altitude turbulent layers, which change independently of seeing conditions at the ground level.

6.3 Image Post-Processing

Each individual captured image is deconvolved with its long-exposure PSF estimated from the AO loop data. The deconvolution attempts to restore the amplitude of the spatial frequency components of the image to their diffraction-limited levels. Equation (4.1) describes the process of image formation as the convolution of the original object with the PSF of the optical system. Deconvolution is the inverse problem where the original object is computed from the image and the PSF. However, in the presence of noise [85], and due to the cut-off spatial frequency of the imaging system, the deconvolution problem turns into a difficult problem, which generally does not have a unique and stable solution. From the expression of a convolution in the Fourier domain, shown in Equation (4.4), a very simplistic deconvolution approach can be attempted by dividing the equation by the OTF.

$$\hat{O}(\vec{\nu}) = \frac{I(\vec{\nu})}{OTF(\vec{\nu})} = O(\vec{\nu}) + \frac{N(\vec{\nu})}{OTF(\vec{\nu})} \tag{6.5}$$

However, the presence of noise indicated by the term $N(\vec{\nu})$ makes this simple approach impossible. Around the spatial frequency cut-off of the PSF, the values of the OTF are very small, and when inverted, they cause noise amplification at high spatial frequencies, i.e., the term $N(\vec{\nu})/OTF(\vec{\nu})$ dominates Equation (6.5). A tradeoff must be introduced to allow for image reconstruction while limiting noise amplification. This is achieved by the introduction of constraints derived from a priori information about the characteristics of the object, i.e., the image and the noise. The use of such constraints is called *regularization*.

There are many iterative and non-iterative deconvolution algorithms available. A simple non-iterative Wiener filter approach was chosen to deconvolve because of its speed and simplicity. A Wiener filter is a linear regularization method where the constraints are

included in a filtering function W that depends on the characteristics of the image and the noise.

$$\hat{O}(\vec{\nu}) = W(\vec{\nu})\frac{I(\vec{\nu})}{OTF(\vec{\nu})} \tag{6.6}$$

The particular Wiener filter chosen in this analysis assumes that the image is affected by additive white noise, i.e., the noise that affects each pixel in the image is independent from the noise affecting its neighboring pixels, which is a valid assumption for the noise sources that affect images captured with a CCD. The long flat noise tail of the power spectrum of the image before deconvolution in Figure 6.9 demonstrates that the noise affecting the images is white noise. Thus, the Wiener filter used is expressed as:

$$W(\vec{\nu}) = \frac{|OTF(\vec{\nu})|^2}{|OTF(\vec{\nu})|^2 + N_f(\vec{\nu})} \tag{6.7}$$

The term N_f acts as a frequency dependent noise filter that controls the tradeoff between image reconstruction and noise amplification. It is computed from the inverse of the image power spectrum. Thus, for spatial frequencies where the power spectrum is low, containing mostly noise, the term N_f is large and dominates the denominator of Equation (6.7), dampening the power of the noise in the deconvolved image.

While linear regularized deconvolution methods present a number of problems [85] they are extremely fast, because only a few Fourier transforms are required. More advanced deconvolution algorithms [14, 85] will be considered in the future, such as Richardson-Lucy algorithm [85], maximum entropy and wavelet-based deconvolution [85].

The quantitative data obtained from the observations benefit from deconvolving each image with its corresponding PSF. This is possible because the amplitudes of the high spatial frequencies components of the image were preserved by AO correction. Deconvolution

restores these amplitudes to their diffraction-limited levels, consequently improving any quantitative measurements obtained from the image.

6.4 Results

6.4.1 Wing Images Affected by Similar Seeing Conditions

Figure 6.10 shows an example of AO-corrected blue and red wing images captured with the UBF. The images shown in the figure are 1 s exposures recorded at successive times, about 6 seconds apart. Both wing images were exposed during similar seeing conditions, with r_0 around 13 cm, and received comparable levels of AO correction producing images of equivalent image quality, as illustrated by their almost identical power spectra shown in Figure 6.9. The spatial frequency cut-off of the image is around 5.5 cycles/arcsec according to the power spectra plot. This is smaller than the diffraction-limited cut-off frequency at this wavelength, which is $D/\lambda = 6.6$ cycles/arcsec. As discussed previously, this is due to the partial correction provided by the AO system, which produces a PSF with a central core width slightly larger than the diffraction-limited width. Both wing images are expected to present similar RMS contrast values because of the large width of the UBF bandpass compared to the width of the spectral line. The similar RMS contrast values of both wing images (see Table 6.1) also indicate images with comparable image quality.

Each wing image was deconvolved with its corresponding estimated PSF, producing the results shown in Figure 6.11. The deconvolved wing images have higher contrasts with sharper features. The RMS contrast enhancement introduced by the deconvolution for both wing images is shown in Table 6.1. The RMS contrast of quiet sun granulation obtained after deconvolution of the wing images (around 9%) disagrees with the results produced by current MHD simulations, which predict a contrast of >20%. However, this result is consistent with the results obtained by other image reconstruction techniques, such as speckle reconstruction and multi-frame blind deconvolution [106], which are unable to reproduce such high contrasts.

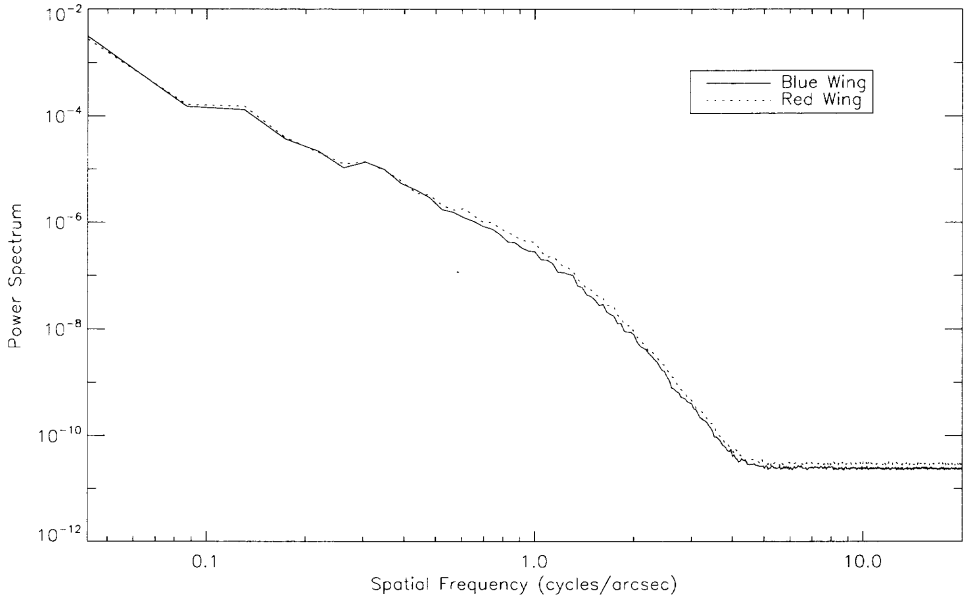

Figure 6.9 Azimuthal averages of power spectra of wing images shown in Figure 6.10, captured during similar median seeing conditions.

Figure 6.12 shows the power spectrum of the blue wing image before and after deconvolution. The figure illustrates the effects of the deconvolution: an increase of the amplitude of the spatial frequency components and a reduction of the noise. The red wing image was similarly affected by the deconvolution suffering a similar amplitude increase and noise filtering.

The SNR of an image can be computed from its power spectrum, i.e., the power content of the image at different spatial frequencies. The CCD noise in the image is white noise that produces the flat high spatial frequency tail seen in Figures 6.9 and 6.12. White noise presents the same power over all spatial frequencies. Hence its power can be extracted from the power of the high frequency tail and integrated over all the spatial frequencies in the image to produce a measurement of the noise variance. Similarly, the integral of the power spectrum over all spatial frequencies in the image provides the variance of the signal. The SNR of the image is computed as the square root of the ratio of image variance

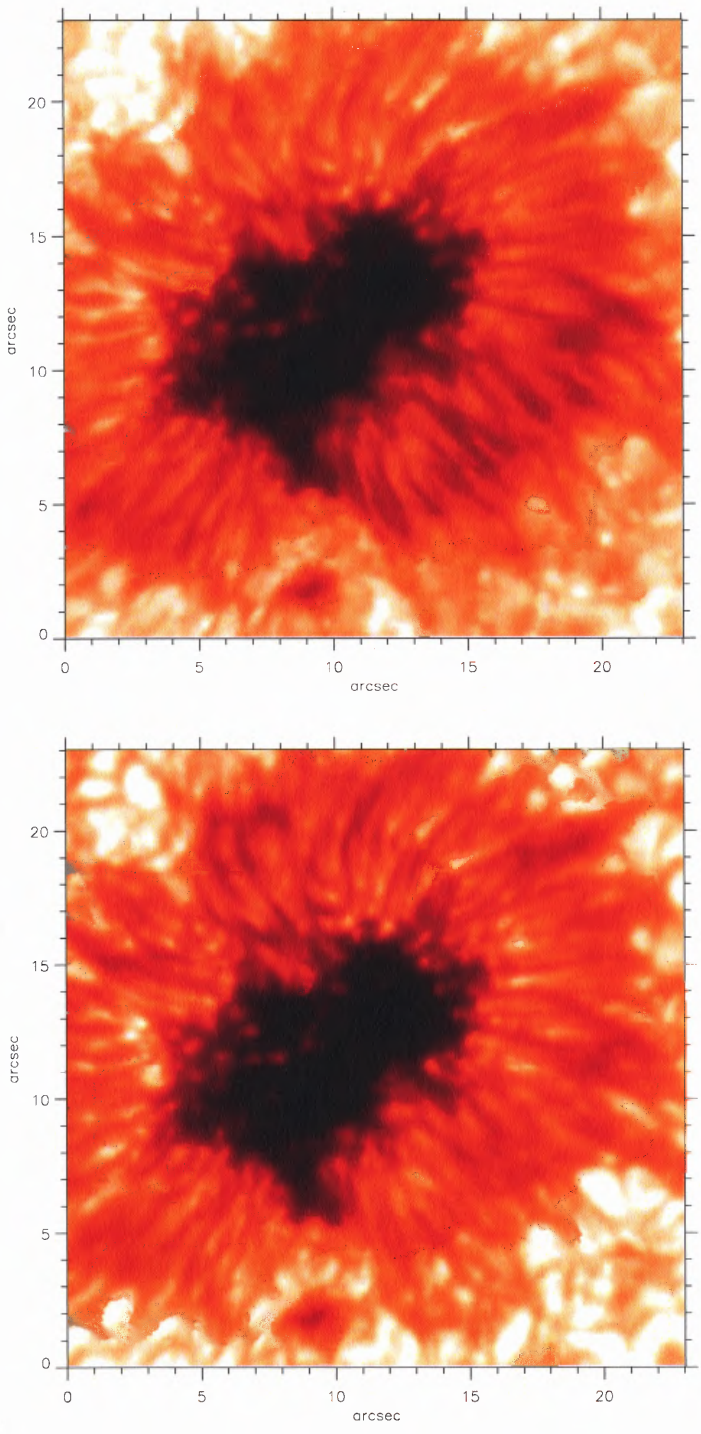

Figure 6.10 AO-corrected blue (top) and red (bottom) wing images. Both wing images were affected by similar seeing conditions ($r_0 \sim 9$ cm).

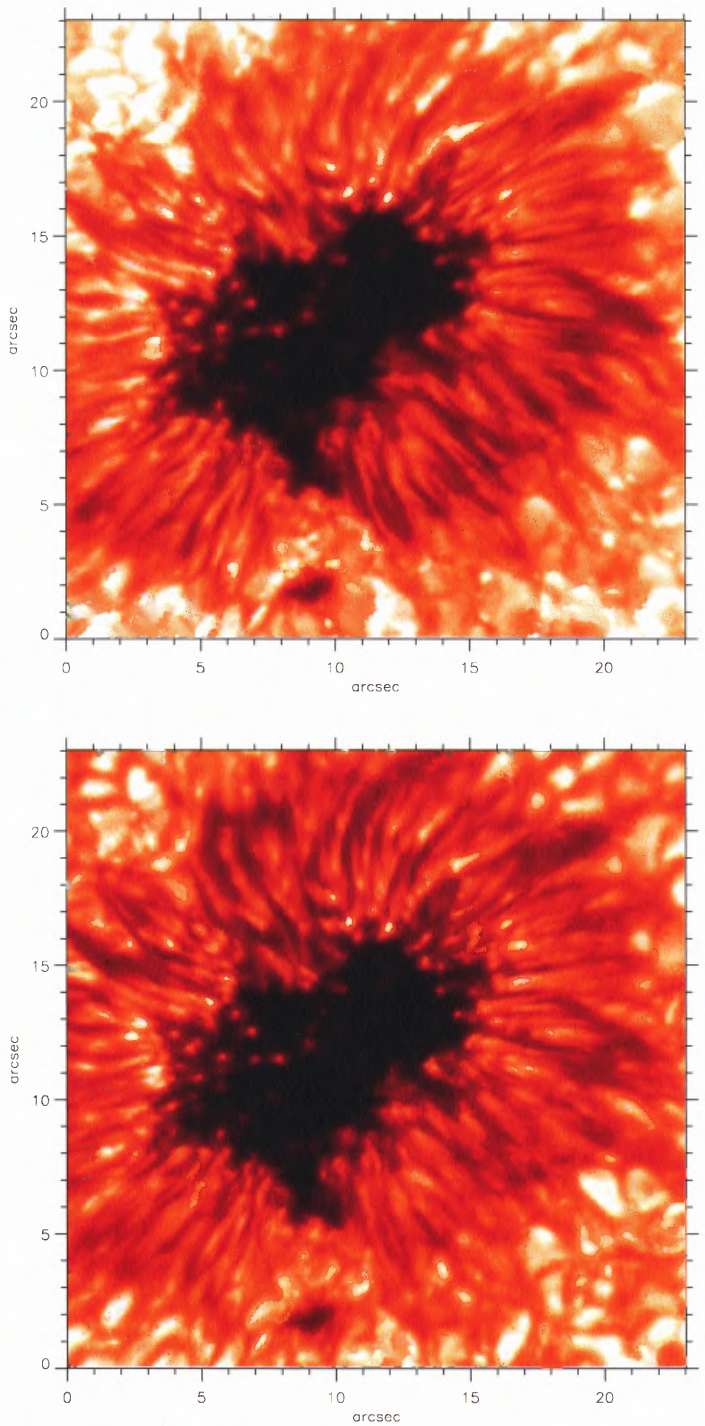

Figure 6.11 Deconvolved AO-corrected blue (top) and red (bottom) wing images. Each image was deconvolved with its corresponding estimated PSF. Images were captured during similar median seeing conditions ($r_0 \sim 9$ cm).

Table 6.1 Wing Images Captured During Similar Seeing Conditions

Image	r_0 *(cm)*	*Strehl Ratio*	*Original RMS*	*Deconvolved RMS*
Blue Wing	8.8	51.2%	0.2472	0.2648
Red Wing	9.2	50.5%	0.2474	0.2677

RMS Contrast						
	Umbra		*Penumbra*		*Quiet Sun*	
Image	*Original*	*Deconv.*	*Original*	*Deconv.*	*Original*	*Deconv.*
Blue Wing	0.1424	0.1990	0.1129	0.1367	0.0598	0.0746
Red Wing	0.1464	0.2080	0.1041	0.1378	0.0682	0.0902

Dopplergram Standard Deviation (m/s)							
Full Image		*Umbra*		*Penumbra*		*Quiet Sun*	
Original	*Deconv.*	*Original*	*Deconv.*	*Original*	*Deconv.*	*Original*	*Deconv.*
493.8	564.8	100.8	159.8	562.9	627.2	398.4	482.4

to noise variance. The SNR of the images shown in Figure 6.10 was about 190, which is a reasonably high SNR.

The velocity information is encoded in the small intensity differences between both wing images. The dopplergrams are computed according to Equation (6.1) by subtracting the red wing image from the blue wing image and dividing the result by their sum. The dopplergrams computed from the AO-corrected wing images (Figure 6.10) and from the AO-corrected deconvolved wing images (Figure 6.11) are shown in Figure 6.13.

Convection inside the umbra of the sunspot is suppressed by the large magnetic fields [4] and, therefore, points inside the umbra should show zero velocity. The mean velocity measured inside the umbra of the sunspot is subtracted from the entire dopplergram to eliminate undesired global velocity offsets [60].

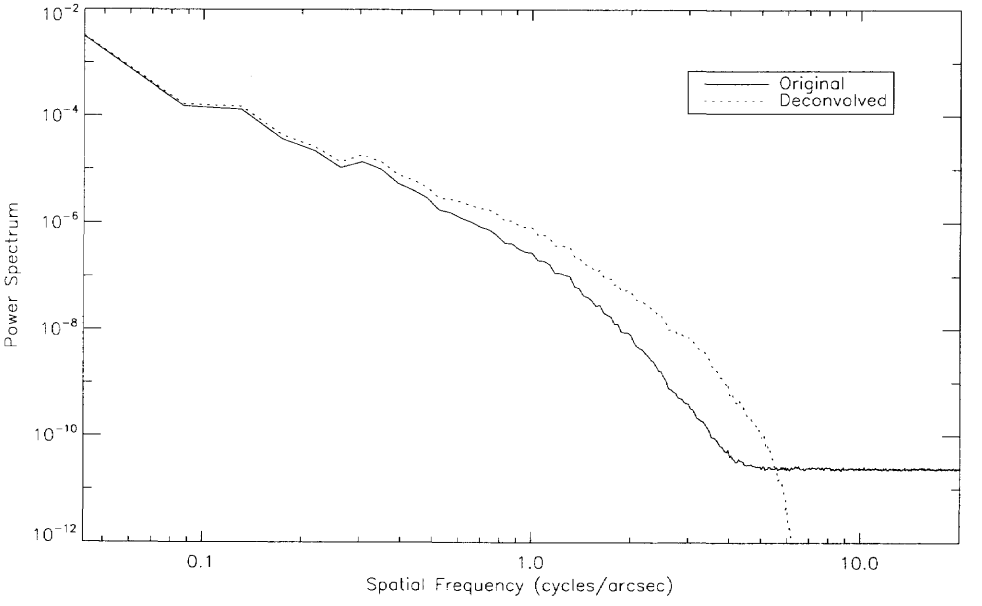

Figure 6.12 Azimuthal average of power spectra of blue wing image before and after deconvolution. Blue wing image was captured during seeing conditions with $r_0 \sim 9$ cm.

The measurements obtained from the original and deconvolved wing images produce slightly different quantitative data. Figure 6.14 illustrates the difference in the quantitative information contained in both wing images before and after deconvolution. The plots in Figure 6.14 show a cross-section, before and after deconvolution, taken across several filaments of the penumbra, as illustrated by the white line in the top panel of Figure 6.13. The left and right panels in Figure 6.14 show data from the blue and red wing images, respectively. The intensity displayed in the plots is normalized to the mean intensity of the surrounding quiet Sun to eliminate global intensity dependencies from the images. The plots further illustrate the increase of RMS contrast and noise filtering produced by the deconvolution. The noise suppression of the Wiener filter is controlled by the noise filter term N_f in Equation (6.7), which must be estimated for each image individually.

A cross-section taken from the calibrated dopplergrams computed from the original and deconvolved wing images is shown in Figure 6.15. The velocity amplitudes are enhanced by the deconvolution of the wing images, particularly for small spatial frequencies.

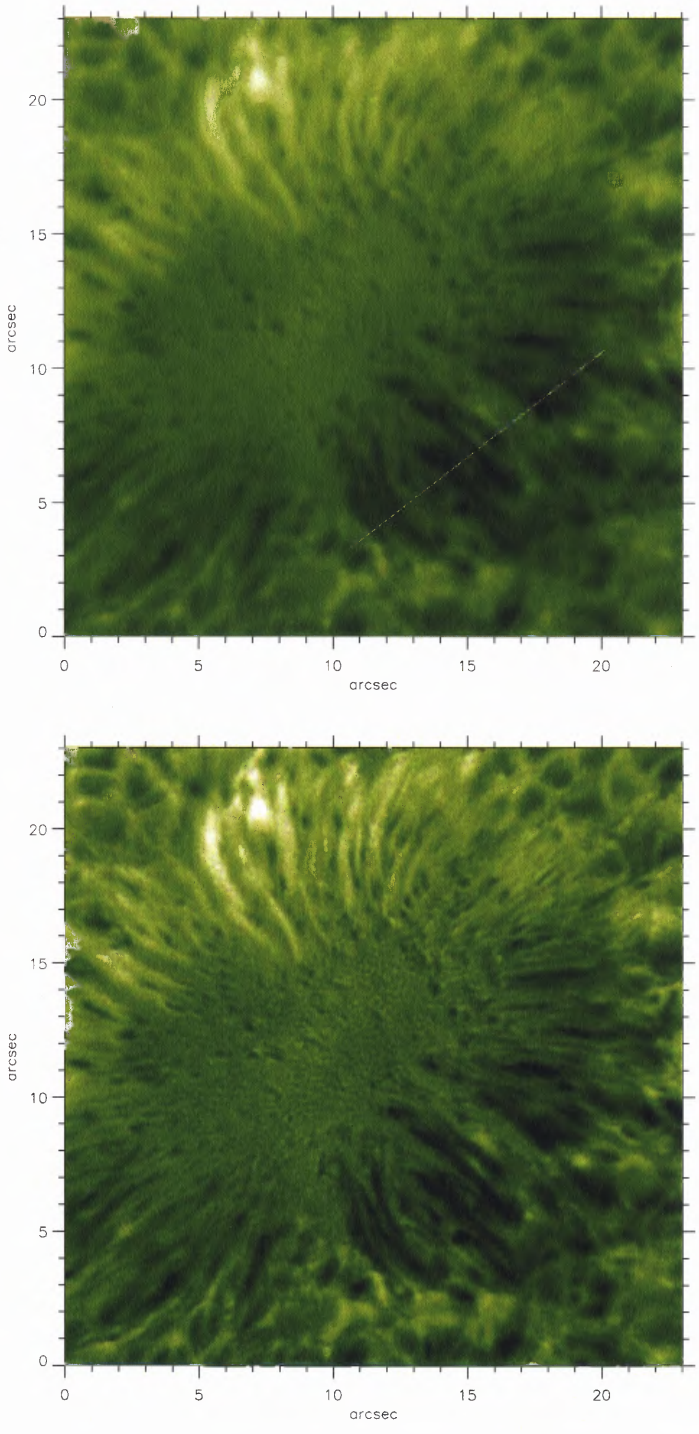

Figure 6.13 Dopplergrams obtained from subtracting the wing images. Before deconvolution (top) and after deconvolution (bottom). The white line on the top panel indicates the location where the cross-sections were obtained. Wing images were captured during seeing conditions with $r_0 \sim 9$ cm.

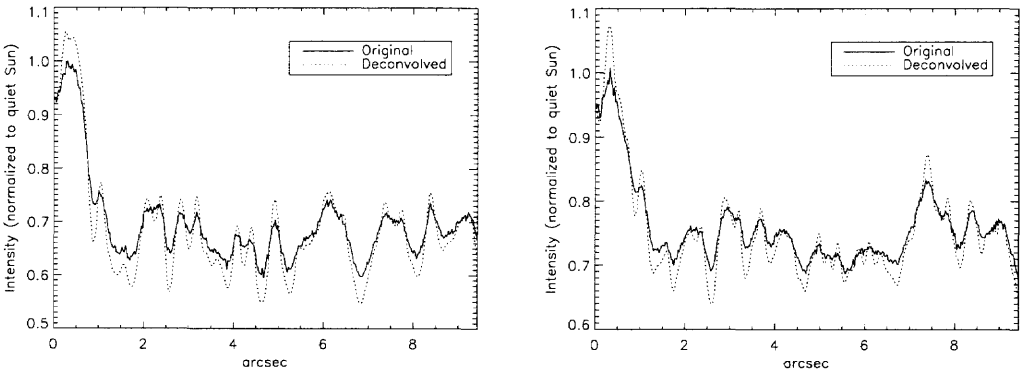

Figure 6.14 Cross-section from AO-corrected blue wing (left) and red wing (right) before and after deconvolution. Cross-sections are taken from location indicated in Figure 6.13.

Table 6.2 Dopplergram Difference: Similar Seeing Conditions

	Full Field	***Umbra***	***Penumbra***
Standard Deviation (m/s)	143.3	120.1	145.0
Average (m/s)	40.2	-2.3	40.2

The very high spatial frequency components beyond the frequency cut-off of the telescope, which are dominated by noise, are damped by the deconvolution. Figure 6.15 illustrates the enhancement of the quantitative data and noise reduction produced by the deconvolution, including a significant global red shift of the dopplergram computed after deconvolution with respect to the dopplergram obtained before deconvolution. The average velocity over the whole field of the dopplergram before deconvolution is about -82 m/s while the mean velocity after deconvolution is about 166 m/s indicating a global red shift of the absolute velocity zero point by about 248 m/s caused by the deconvolution. This global red shift arises from the subtraction of the mean velocity inside the umbra, which has been reduced by deconvolution. The difference between the dopplergram before and the dopplergram after deconvolution was computed. The standard deviation and mean values of this difference image, shown in Table 6.2, provide a measure of the correction that deconvolution provides to the quantitative velocity measurements in the dopplergram. The correlation coefficient

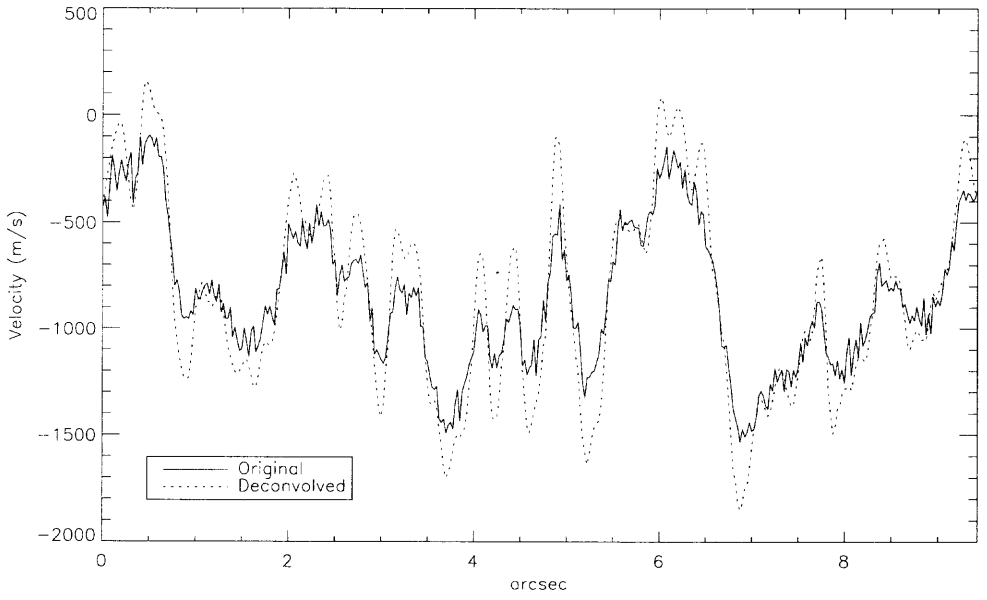

Figure 6.15 Cross-section from the dopplergrams before and after deconvolution shown in Figure 6.13. Cross-sections are taken from location indicated in Figure 6.13.

between the velocity cross-sections before and after deconvolution is 0.93, which indicates a high correlation between the velocity signals. This implies that the improvement provided by the deconvolution mostly affected the quantitative values of the signal and not its overall structure.

6.4.2 Wing Images Affected by Different Seeing Conditions

In the examples discussed above, both wing images were affected by similar seeing conditions. Another more common case to consider is the case where the seeing conditions change drastically between the capture of the blue and red wing exposures. In this case, the image quality and RMS contrast of each wing image can be significantly different. The dopplergram obtained as the difference of such a pair of wing images may contain many artifacts and completely skewed quantitative values. Figure 6.16 shows an example of such a pair of wing images that were exposed during very different seeing conditions, as indicated in Table 6.3.

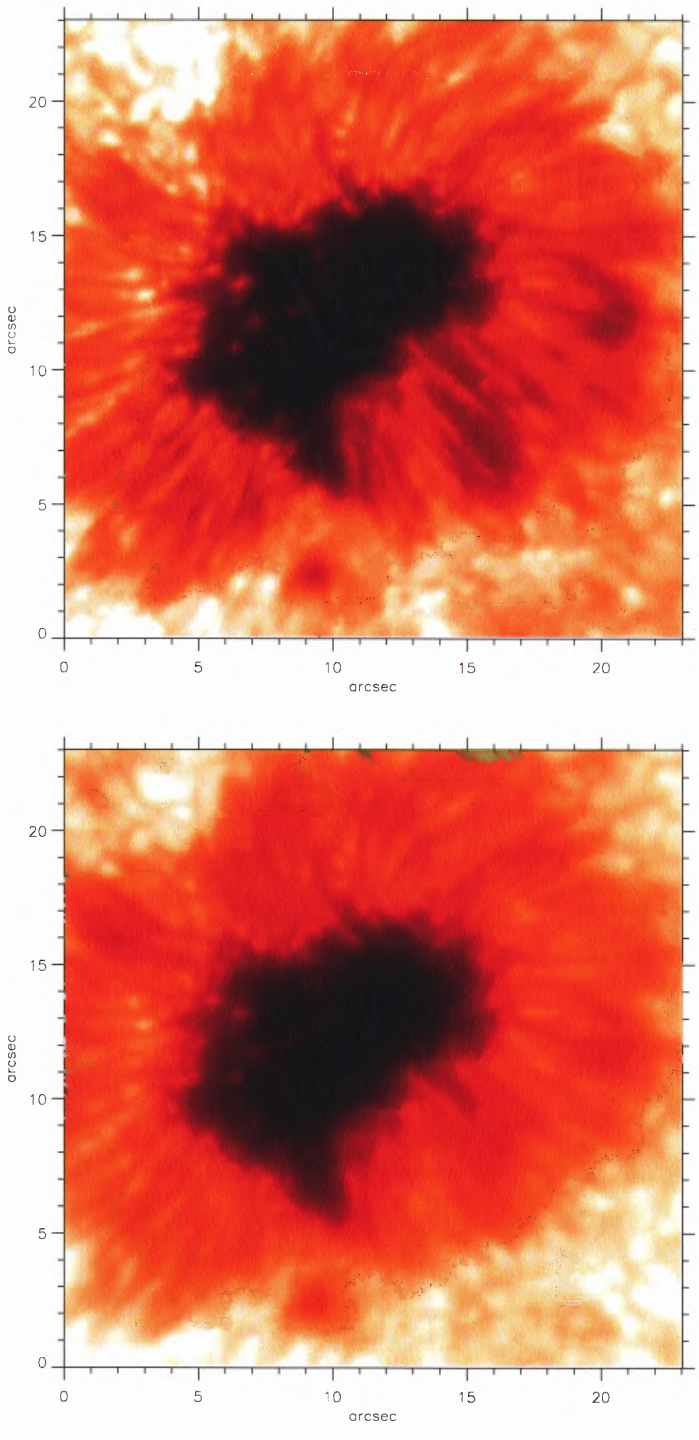

Figure 6.16 AO-corrected blue (top) and red (bottom) wing images affected by very different seeing conditions. The Fried parameters r_0 for the blue and red wing images are $r_0 \simeq 14$ cm and $r0 \simeq 6$ cm, respectively.

Table 6.3 Wing Images Captured During Different Seeing Conditions

Image	r_0 *(cm)*	*Strehl Ratio*	*Original RMS*	*Deconvolved RMS*
Blue Wing	9.4	61.3%	0.2438	0.2555
Red Wing	4.3	20.4%	0.2244	0.2495

RMS Contrast						
	Umbra		*Penumbra*		*Quiet Sun*	
Image	*Original*	*Deconv.*	*Original*	*Deconv.*	*Origignal*	*Deconv.*
Blue Wing	0.0958	0.1252	0.1003	0.1144	0.0538	0.0636
Red Wing	0.0695	0.1145	0.0703	0.1010	0.0376	0.0640

Dopplergram Standard Deviation (m/s)							
Full Image		*Umbra*		*Penumbra*		*Quiet Sun*	
Original	*Deconv.*	*Original*	*Deconv.*	*Original*	*Deconv.*	*Original*	*Deconv.*
493.1	507.0	274.0	192.5	508.4	543.9	388.4	453.1

The blue wing image was recorded during seeing conditions with a Fried parameter of $r_0 \simeq 9$ cm, while the red wing image was affected by much more severe seeing conditions with $r_0 \simeq 4$ cm. The RMS contrast of the red wing image is significantly lower than the RMS contrast of the blue wing image (see Table 6.3). Deconvolution of each wing image with its corresponding estimated PSF attempts to restore both images to an equivalent image quality level. The deconvolved wing images are shown in Figure 6.17.

Figure 6.18 shows dopplergrams calculated from the wing images before and after deconvolution. The dopplergram computed before deconvolution (top panel in the figure) shows significant intensity-velocity cross-talk, especially in the top left area of the dopplergram and in the umbra of the sunspot. The characteristic velocity signals produced by the Evershed flow in the penumbra of the sunspot are severely distorted and almost non-

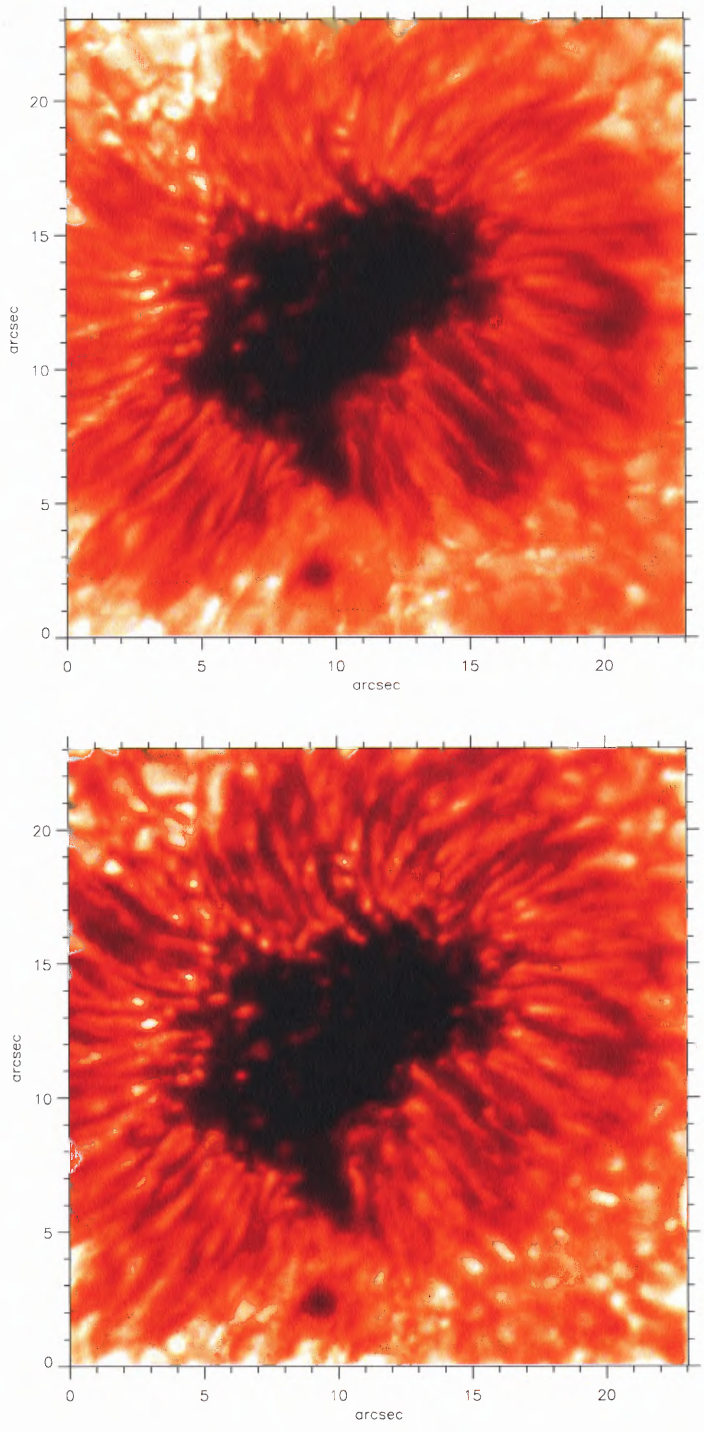

Figure 6.17 Deconvolved AO-corrected blue (top) and red (bottom) wing images affected by very different seeing conditions.

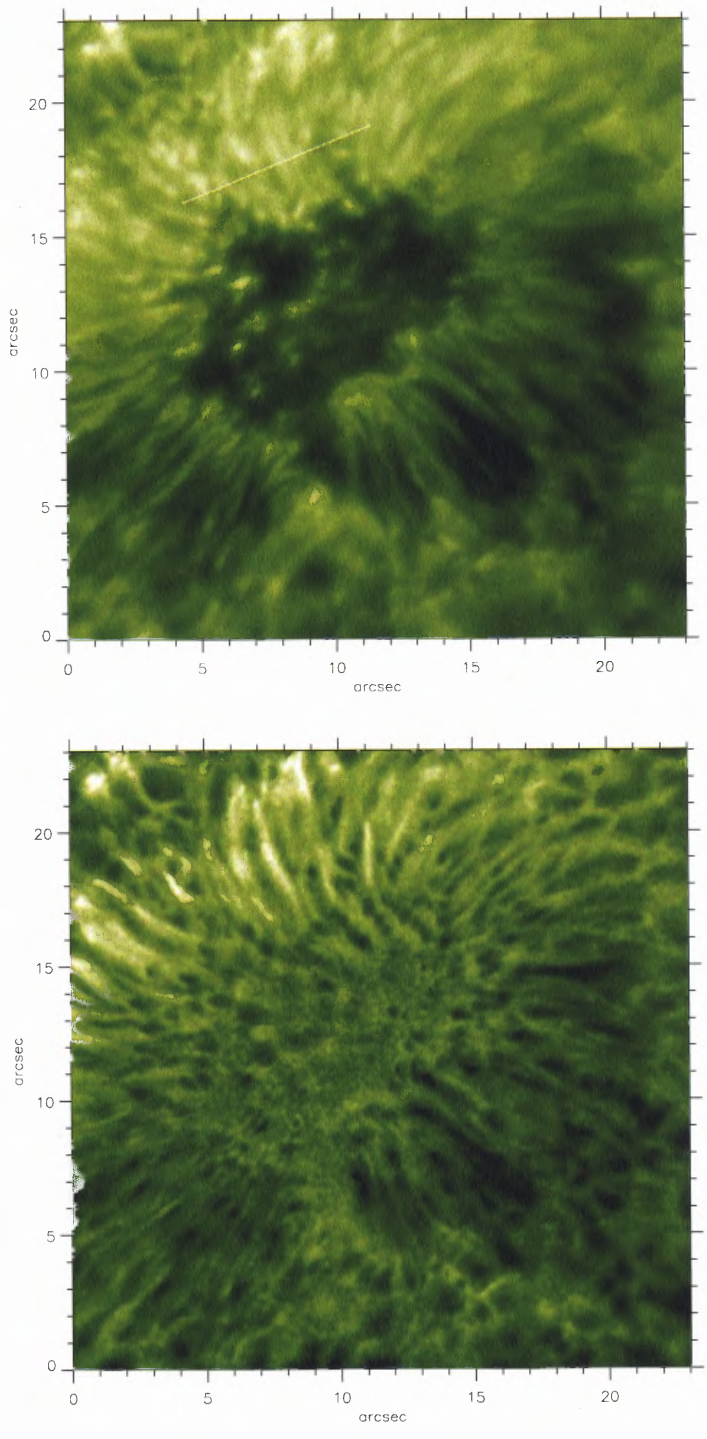

Figure 6.18 Dopplergrams computed from the original (top) and deconvolved (bottom) wing images. Each wing image was affected by very different seeing conditions. The white line on the top panel indicates the location where the cross-sections were obtained.

Table 6.4 Dopplergram Difference: Different Seeing Conditions

	Full Field	***Umbra***	***Penumbra***
Standard Deviation (m/s)	235.2	204.7	214.9
Average (m/s)	-267.1	-6.1	-299.2

existent in some areas of the dopplergram. The Evershed effect appears as radial filament shaped structures, which are clearly visible in other dopplergrams captured during better seeing conditions, such as the ones shown in Figures 6.13 and 6.26.

The velocity filament structures in the penumbra are considerably recovered in the dopplergram computed after deconvolution of the wing images (bottom panel in Figure 6.18). Radial filament structures can be seen in the top left area of the penumbra, appearing as long thin regions of red shifted velocity. However, the dopplergram still presents some residual intensity-velocity cross-talk after deconvolution, which indicates that the deconvolved wing images were not restored to exactly the same image quality.

The very different quantitative velocity measurements obtained from the dopplergrams before and after deconvolution are shown in Figure 6.19. Figure 6.19 shows cross-sections taken along the white line in Figure 6.18 where the intensity-velocity cross-talk is most pronounced. The velocity signal obtained from the dopplergram before (dotted line) and after (solid line) deconvolution are significantly different.

Figure 6.19 also shows the velocity signal obtained from another dopplergram (dashed line) computed from wing images captured shortly after during good seeing conditions (bottom panel of Figure 6.21), which is assumed to represent the "ground-truth". The new good seeing dopplergram was computed from data captured 1 min and 20 s later during seeing conditions characterized by $r_0 \simeq 18$ cm. The structures in the umbra and penumbra of the sunspot do not noticeably evolve [57] during the time interval between both data sets. Hence, both data sets refer to the same object and should produce identical results. The only change to be expected between the two data sets is in the granulation patterns, which

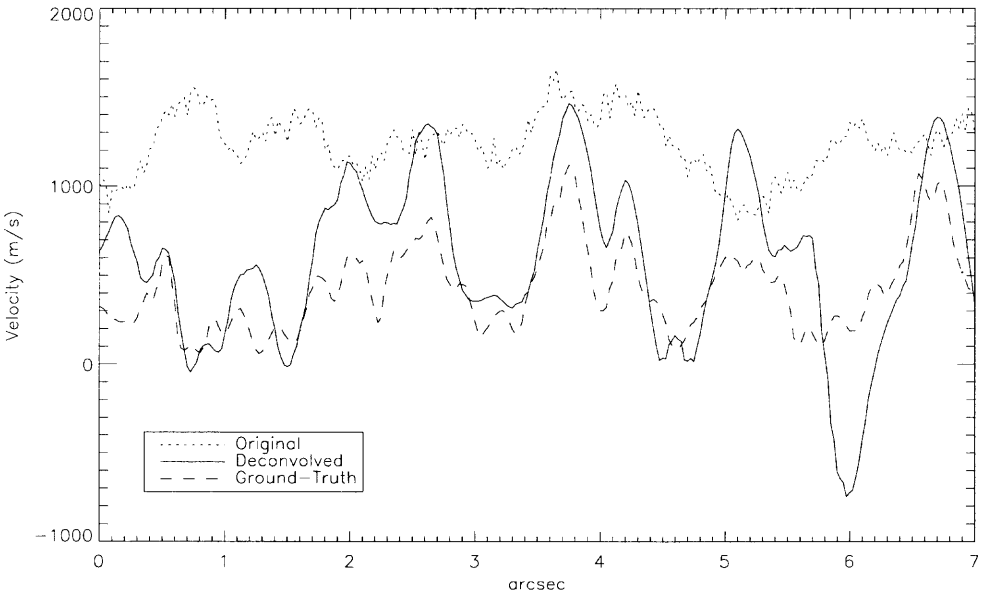

Figure 6.19 Cross-section of velocities obtained from AO-corrected wing images and from deconvolved wing images. Wing images were affected by very different seeing conditions. The cross-sections were obtained from the location indicated in Figure 6.18.

evolve much quicker and may have changed in the time between both sets of exposures. The wing images obtained during the very good seeing conditions yield Strehl ratios around 0.8 and, therefore, are very close to perfect diffraction-limited images. The dopplergrams they produce provide reliable quantitative velocity measurements, i.e., the "ground-truth", that can be contrasted to the measurements obtained from the dopplergrams computed from wing images affected by different seeing conditions (Figure 6.18).

The correlation coefficients between the velocities before and after deconvolution and the velocities from the good seeing dopplergram indicate the considerable improvement produced by deconvolution. The correlation coefficient between the velocity signal before deconvolution and the good seeing dopplergram is 0.13, while for the signal after deconvolution is 0.74. The small correlation coefficient (0.13) obtained from the data before deconvolution indicates that both signals contain completely different information, i.e., it is evidence of the large intensity-velocity cross-talk. For the deconvolved data, the

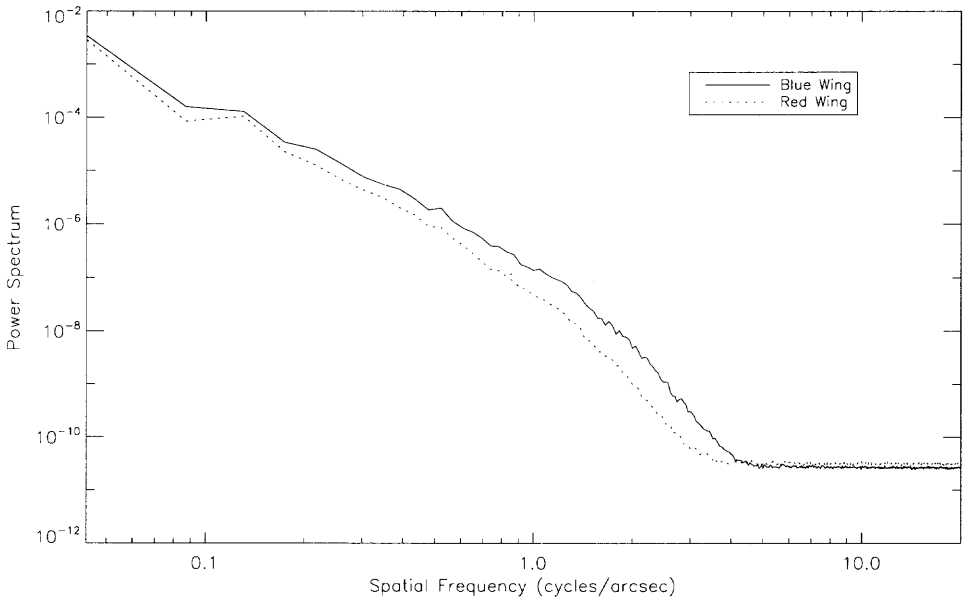

Figure 6.20 Azimuthally averaged power spectra of blue and red wing images captured during very different seeing conditions. Conditions were worse during the exposure of the red wing image.

coefficient is considerably higher (0.74) but still not high enough to indicate a good correlation. This indicates that even though deconvolution significantly reduced the amount of intensity-velocity cross-talk, it did not completely eliminate it.

The dopplergrams before and after deconvolution are subtracted to obtain a measure of the correction provided by deconvolution. The standard deviation and average values of this difference are shown in Table 6.4.

The RMS contrast values of the wing images before deconvolution, shown in Table 6.3, provide evidence of the different image quality of both wing images. After deconvolution the RMS contrast values are higher for both wing images, but they are still not brought to the same level. This can be further explored by computing the power spectrum of the wing images. Figure 6.20 shows the azimuthally averaged power spectra of both wing images before deconvolution. The power spectrum of the red wing image presents a consistently lower power over the whole spectral range, indicating its lower image quality

and RMS contrast. The differences between the power spectra indicate that the AO system did not provide the same level of correction for both wing images. The cut-off spatial frequency, i.e., the frequency at which the power of the image is lost in the noise, is also lower for the red wing image than the blue wing image. Deconvolution will not be able to restore both wing images equally because they contain different high spatial frequency information. This is also shown by the different RMS contrast values of both wing images after deconvolution (see Table 6.3).

The severe seeing conditions during the exposure of the red wing image reduced the level of correction provided by the AO system, producing an image with a much lower Strehl ratio than the blue wing image. Some of the high spatial frequency information may not be adequately preserved by the lower correction provided by the AO system and may be lost in the noise beyond recovery. This loss of high spatial frequency information is illustrated by the power spectra shown in Figure 6.20. Hence, due to the different spatial frequency contents of both wing images, it is to be expected that deconvolution will not be able to equally restore both wing images and produce an accurate dopplergram, as shown in Figure 6.19.

Deconvolution of the wing images significantly reduced the amount of intensity-velocity cross-talk from the dopplergram but some residual cross-talk can still be observed in the deconvolved dopplergram. As discussed earlier, another dopplergram was computed from a pair of wing images captured shortly after, during very good seeing conditions, and can be used for comparison here. Figure 6.21 displays the deconvolved dopplergrams from the variable seeing data set (top panel) and the very good seeing data captured short after (bottom panel). The presence of intensity-velocity cross-talk in the dopplergram for variable seeing is made evident when displayed alongside the dopplergram captured under very good seeing.

Figure 6.22 displays the difference between the dopplergrams from variable seeing data and the dopplergrams captured during good seeing conditions. The top image in Fig-

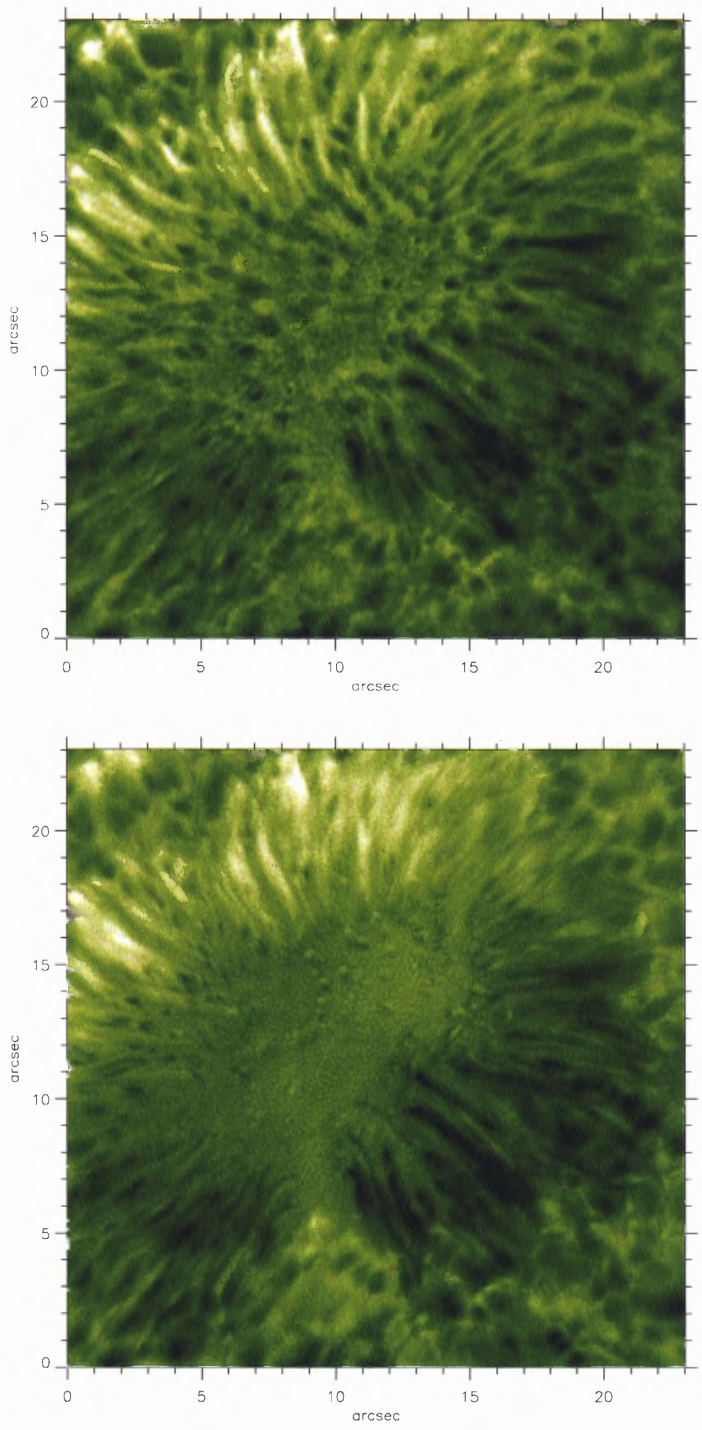

Figure 6.21 Deconvolved dopplergrams from variable seeing data (top) and from data captured during good seeing conditions (bottom).

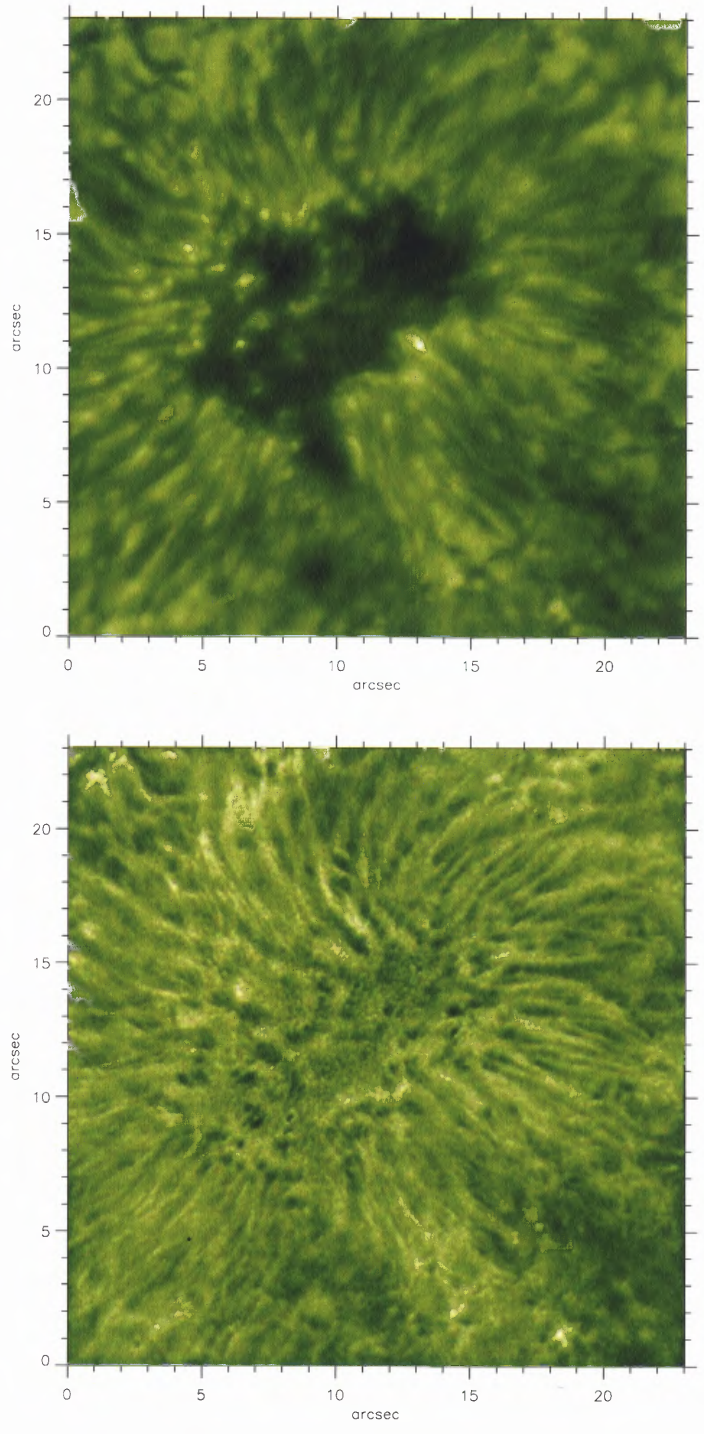

Figure 6.22 Difference image taken between dopplergrams from bad seeing data and very good seeing data. Before deconvolution (top) and after deconvolution (bottom).

Table 6.5 Difference Between Good and Bad Seeing Dopplergrams

	Standard Deviation		
	Full Field	***Sunspot Umbra***	***Sunspot Penumbra***
Before Deconvolution	350.5	303.3	245.0
After Deconvolution	274.9	231.3	191.4

ure 6.22 shows the difference taken before deconvolution while the bottom image shows the difference taken after deconvolution. Ideally, the difference between both sets of dopplergrams should be very small, since they refer to the same object and should provide the same measurements. The standard deviation taken in several areas of the dopplergram difference images illustrates the great improvement provided by deconvolution (see Table 6.5).

The dopplergram difference image computed before deconvolution from Figure 6.22 is dominated by intensity signal with almost no discernible velocity signal. The dopplergram difference image computed after deconvolution shows much less detail with lower intensity cross-talk. The dopplergram difference image computed before deconvolution has very little structure and is uniformly gray, as opposed to the dopplergram difference image computed after deconvolution. This demonstrates that deconvolution of the wing images with their estimated PSFs produces significantly improved scientific data and that the PSFs produced by the method seem to provide good estimations of the AO-corrected PSF.

6.4.3 Very Good Seeing Conditions

Under very good seeing conditions, the AO correction provides excellent image quality that produces very good quantitative velocity measurements. However, even in excellent seeing conditions the AO correction remains only partial, allowing some high-order aberrations to affect the image and reduce its quality. Deconvolution can help to further improve the quality of these images by further restoring their Fourier amplitudes to their diffraction-

Table 6.6 Wing Images Captured During Very Good Seeing Conditions

Image	r_0 *(cm)*	*Strehl Ratio*	*Original RMS*	*Deconvolved RMS*
Blue Wing	18.7	84.4%	0.2523	0.2633
Red Wing	18.7	82.1%	0.2591	0.2726

RMS Contrast						
	Umbra		*Penumbra*		*Quiet Sun*	
Image	*Original*	*Deconv.*	*Original*	*Deconv.*	*Original*	*Deconv.*
Blue Wing	0.1717	0.2012	0.1234	0.1394	0.0636	0.0741
Red Wing	0.1721	0.2045	0.1159	0.1375	0.0728	0.0878

Dopplergram Standard Deviation (m/s)							
Full Image		*Umbra*		*Penumbra*		*Quiet Sun*	
Original	*Deconv.*	*Original*	*Deconv.*	*Original*	*Deconv.*	*Original*	*Deconv.*
533.2	581.8	111.8	154.2	560.7	606.1	479.4	541.4

limited levels. Figure 6.23 shows an example of AO-corrected wing images captured during very good seeing conditions, with r_0 values around 18 cm. Each wing image is deconvolved with its corresponding estimated PSFs, producing the results shown in Figure 6.24.

Figure 6.25 illustrates this improvement by comparing the power spectrum of the blue wing before and after deconvolution. The power spectra in the Figure 6.25 indicate that the deconvolution affects the high frequencies the most, as would be expected for such a very good seeing image. The improvement is also seen in the increase of RMS contrast shown in Table 6.6, which results in an increase of 4.5%.

The dopplergrams obtained from the wing images before and after deconvolution are shown in Figure 6.26 and velocity cross-sections before and after deconvolution are found in Figure 6.27. The cross-sections were taken across several velocity filaments of

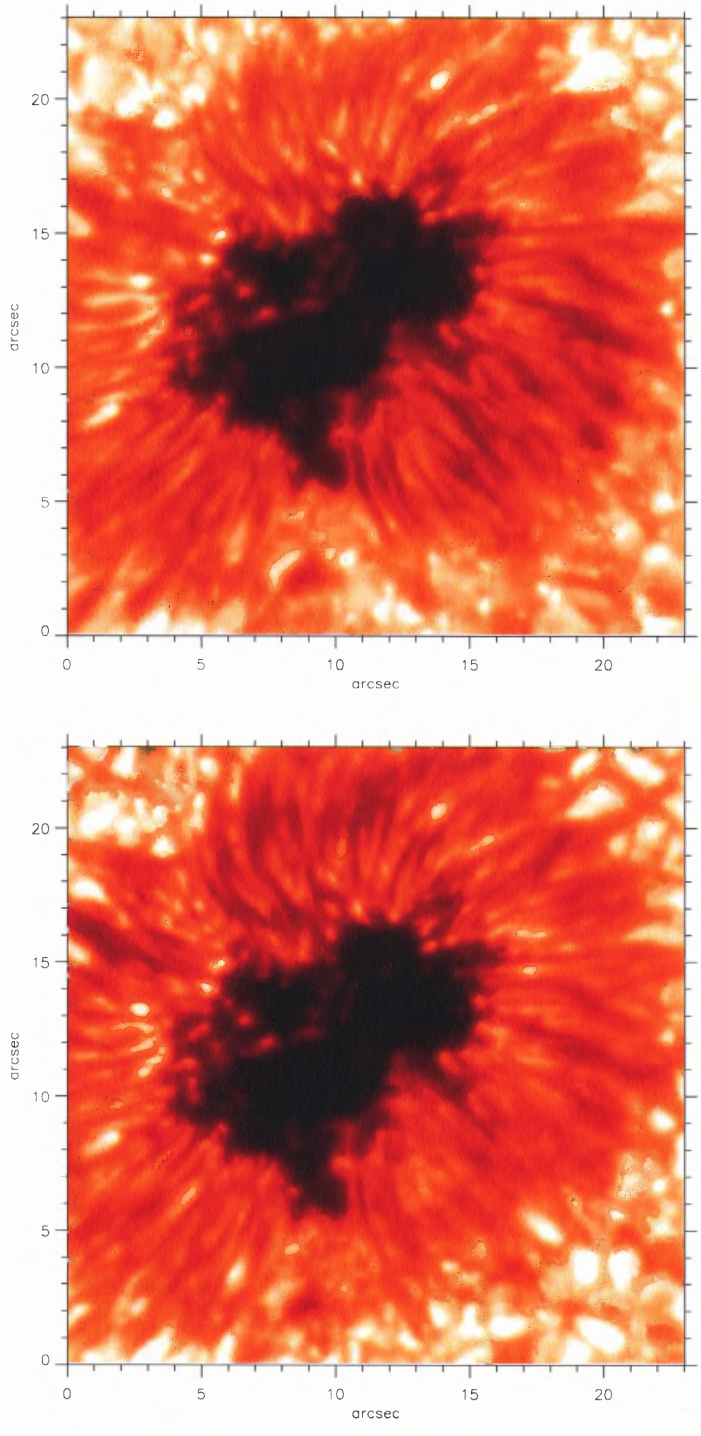

Figure 6.23 AO-corrected blue (top) and red (bottom) wing images captured during very good seeing conditions: $r_0 \simeq 18$ cm.

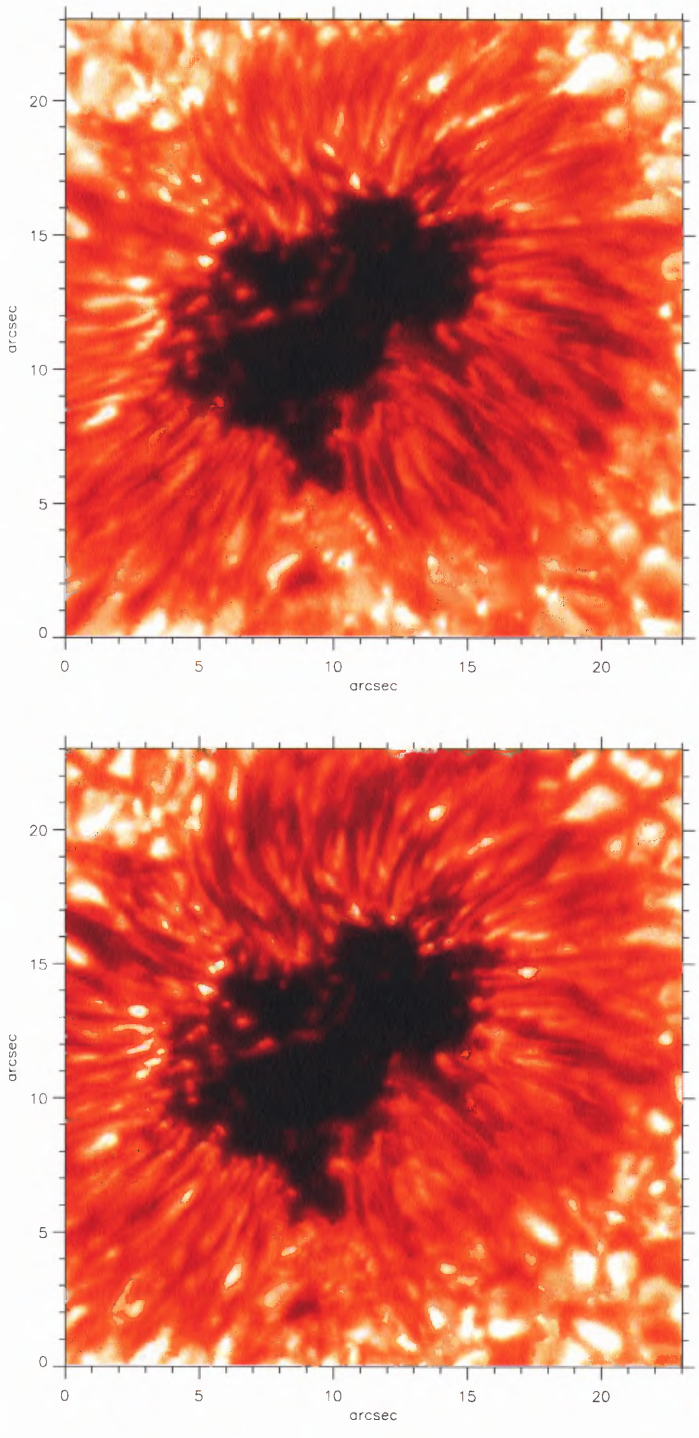

Figure 6.24 Deconvolved AO-corrected blue (top) and red (bottom) wing images captured during very good seeing conditions: $r_0 \simeq 18$ cm.

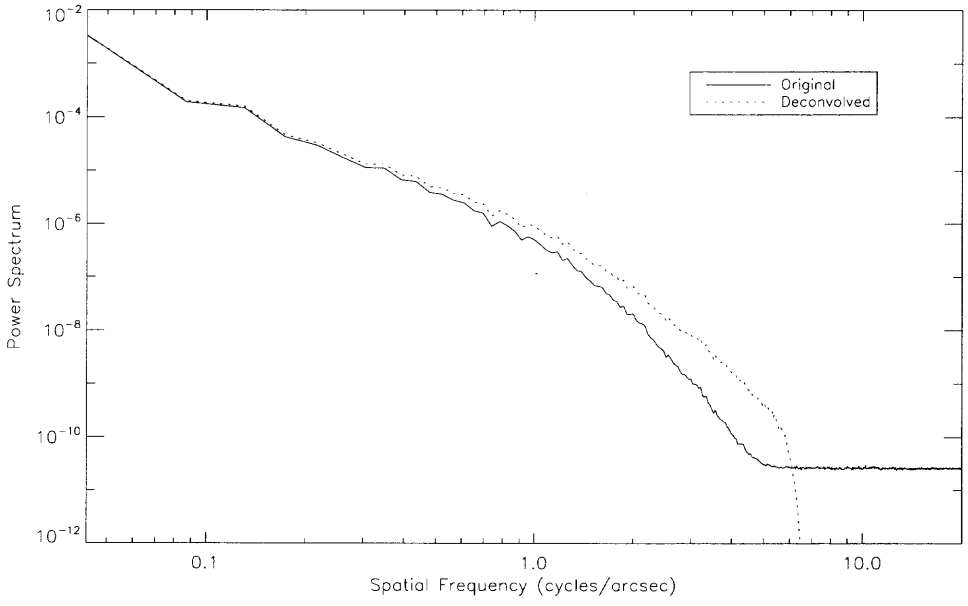

Figure 6.25 Azimuthally averaged power spectrum of blue wing image captured during very good seeing conditions before and after deconvolution.

the dopplergram, as indicated by the white line in the top panel of Figure 6.26. The RMS contrast values shown in Table 6.6 and the cross-sections from Figure 6.27 further illustrate the possibility of improvement even of data captured during very good seeing conditions.

The difference between the dopplergram before and the dopplergram after deconvolution provides a measurement of the improvements provided by deconvolution. Table 6.7 shows the standard deviation and average values taken from the dopplergram difference before and after deconvolution. The comparison of Table 6.7 with Tables 6.2 and 6.4 indicates that the quantitative improvement produced by deconvolution is much more pronounced for data captured during bad or variable seeing conditions. While good seeing data are also enhanced by deconvolution, the changes introduced are not as significant as in the case of mediocre seeing conditions.

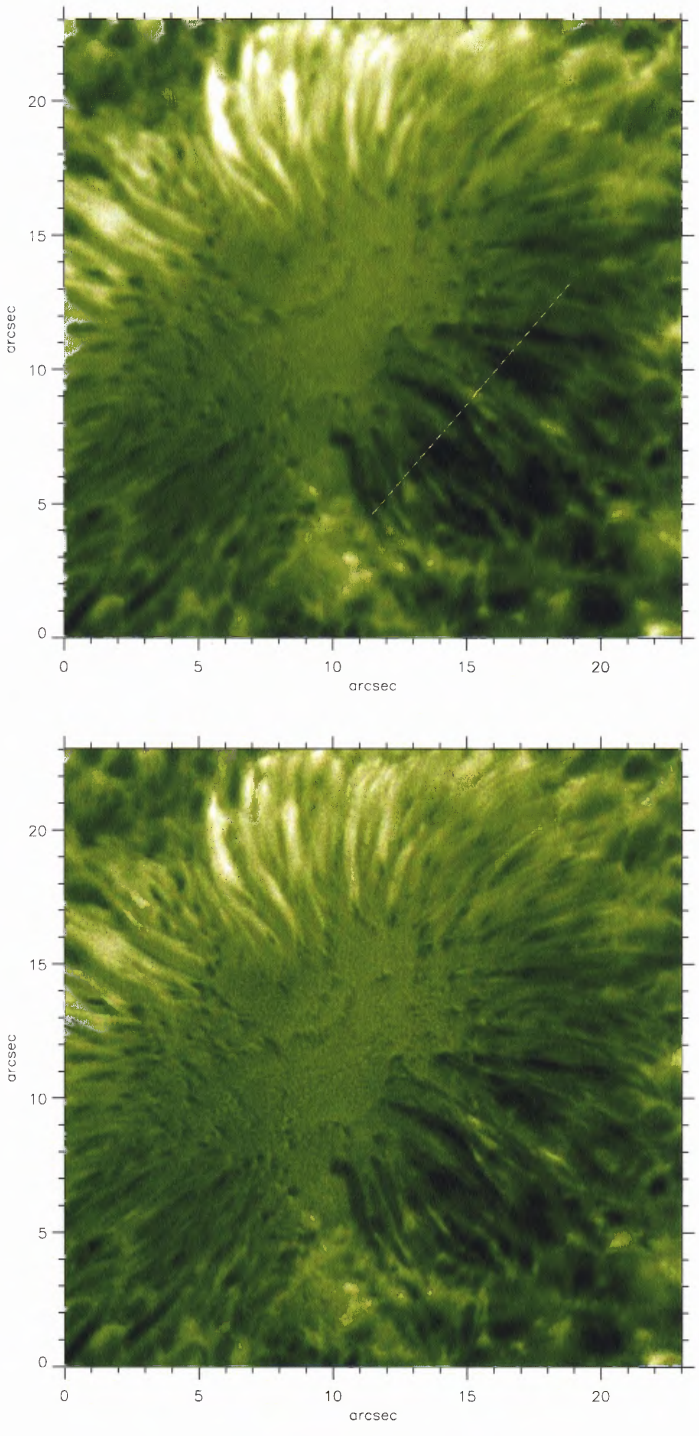

Figure 6.26 Dopplergrams computed from the original (top) and deconvolved (bottom) wing images. Both wing images were captured during very good seeing conditions. The white line in the top panel indicates the location where the cross-sections were obtained.

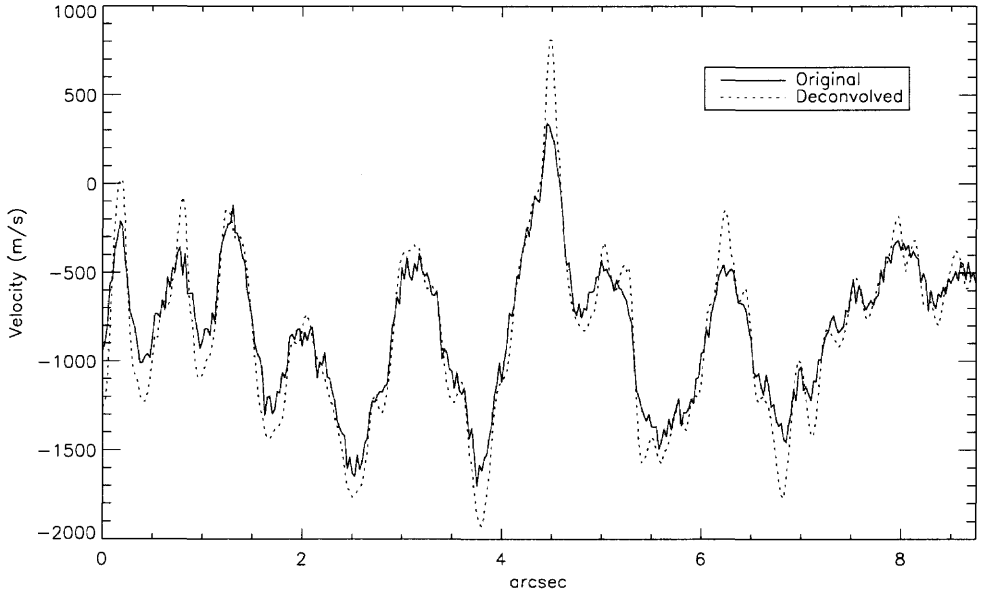

Figure 6.27 Cross-section of velocities obtained from AO-corrected wing images and from deconvolved wing images. Wing images were captured during very good seeing conditions: $r_0 \simeq 18$ cm. Cross-sections were obtained from the location indicated in Figure 6.26

Table 6.7 Dopplergram Difference: Very Good Seeing Conditions

	Full Field	***Umbra***	***Penumbra***
Standard Deviation (m/s)	112.1	109.5	113.3
Average (m/s)	−8.4	−0.5	−10.9

6.5 Accuracy of the Estimated PSF

The improved quantitative measurements produced by deconvolution have been demonstrated for data captured during different seeing conditions. An accurate estimation of the PSF is very important to produce good deconvolution results. An incorrect PSF estimate may affect the quantitative measurements obtained and even introduce artifacts in the image. With solar observations as the ones described in this chapter, it is very difficult to quantify the accuracy of the estimated PSF. An independent parallel method is required to obtain reliable PSFs that can then be compared to the ones produced by the PSF estimation

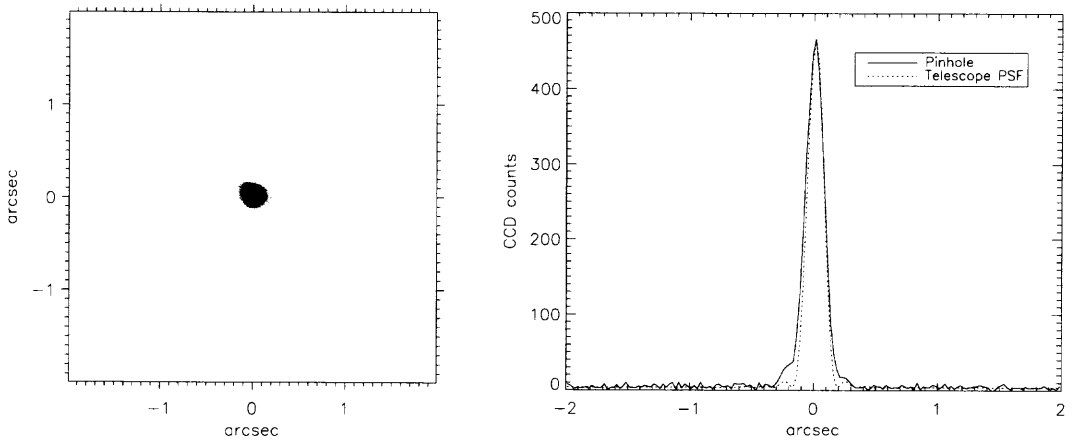

Figure 6.28 Image of the 30 μm pinhole in a logarithmic display (left) and its cross-section compared to the diffraction-limited PSF of the telescope (right).

algorithm. An attempt to obtain a direct parallel measurement of the PSF by observing the bright star Sirius was already discussed in Chapter 5.

Several factors, which are not considered in the current PSF estimation method, can affect the accuracy of the PSFs estimated from the AO loop data, such as non-common path aberrations, anisoplanatism effects and scattered light. Non-common path aberrations are aberrations introduced by the optical elements between the AO system and the science camera. Hence, these aberrations appear in the final image but are not sensed by the WFS of the AO system. To evaluate the severity of these aberrations a very small pinhole (30 μm) was placed at the focus of the telescope. This small pinhole cannot be resolved and behaves as an ideal point source. An AO-corrected image of this pinhole represents the PSF of the optics in the non-common path and characterizes its aberrations.

Figure 6.28 shows an image of the pinhole and its cross-section compared to the diffraction-limited PSF of the telescope. The image of the pinhole is shown in a logarithmic display (left panel) and it clearly shows traces of a diffraction ring around the central peak. The cross-sections from Figure 6.28 show a small difference between the widths of the pinhole image and the diffraction-limited PSF. The FWHM of the pinhole image is measured at $0.''166$ while the FWHM of the diffraction-limited PSF is $0.''151$. The slight

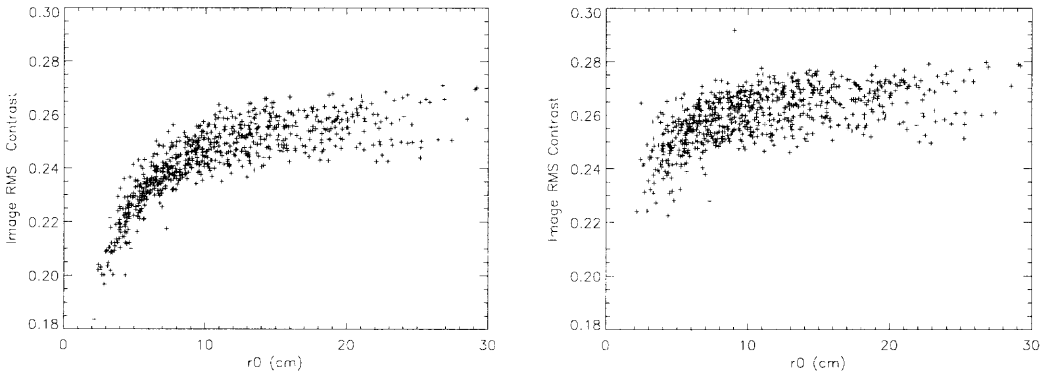

Figure 6.29 Image contrast versus the Fried parameter r_0 from the wing images captured on May 2004. From the original images (left) and deconvolved images (right).

larger width of the pinhole indicates the presence of a small amount of non-common path aberrations that reduce the maximum spatial resolution of the optical set up.

The accuracy of the estimated PSFs also depends critically on the values of the AO system calibration factors, as discussed in Chapter 5. The correct estimation of the residual wavefront and the applied correction depends on the accurate determination of these calibration factors. Using an incorrect DM calibration factor would skew the estimation of r_0 obtained from the DM commands. Similarly, the residuals measured by the WFS require an accurate WFS calibration factor to account for the exact residuals that reach the image.

The calibration factors obtained theoretically from the technical data of the system (shown in Table 5.1) were used to computed the estimated PSFs from the AO loop data. A PSF was estimated for each image in the time series of data captured on May 6th 2004. Each solar image was deconvolved with its estimated PSF. The RMS contrast of the image was computed before and after deconvolution and stored together with the estimated value of r_0 and the Strehl ratio of the PSF. Figure 6.8 shows the Strehl ratio obtained from this calculations plotted against the value of r_0. The RMS contrasts of the AO-corrected images before and after deconvolution are displayed in Figure 6.29 as a function of r_0.

As illustrated by Figure 6.8, the correction that the AO system provides depends strongly on the seeing conditions. The same dependency applies to the RMS contrast of

an AO-corrected image. It is higher for good seeing conditions and lower for bad seeing conditions. There is a maximum RMS contrast that a particular image can have that is determined by the RMS contrast of the object and the diffraction-limited PSF of the optical system. A diffraction-limited image will have this maximum possible RMS contrast.

Ideally, after deconvolution with their corresponding estimated PSF the RMS contrast for every wing image should be restored to the maximum diffraction-limited RMS contrast of the image and be independent of seeing conditions. This should produce a flat distribution of the RMS contrast after deconvolution when plotted against the value of r_0.

The left panel in Figure 6.29 illustrates the dependence of the RMS contrast of the AO-corrected wing images on the seeing conditions. The right panel in Figure 6.29 displays the RMS contrast of those same wing images after deconvolution with their corresponding estimated PSFs. The RMS contrast is clearly increased for all the wing images and its dependency on the seeing conditions is reduced, as indicated by a flatter distribution. However, the seeing dependency is not completely eliminated from the deconvolved RMS contrast. This is to be expected, since for extremely bad seeing conditions, the performance of the AO system is severely reduced and some of the high spatial frequency information in the images may be lost beyond recovery in the noise. This is illustrated in the power spectra shown in Figure 6.20.

CHAPTER 7

THE EVERSHED FLOW: FLOW GEOMETRY AND ITS TEMPORAL EVOLUTION

7.1 Introduction

Sunspot penumbra still present many mysteries to solar astronomers. In particular, the complex magnetic field structure and the dynamic evolution of penumbral fine-structure are not yet well understood. See the recent review articles by Thomas & Weiss [93] and Weiss [103] for a summary of outstanding problems concerning the physics of the penumbra.

The penumbral magnetic field is highly inhomogeneous. Several observations performed over the last decade have shown that the inclination of the field vector is different between dark and bright filaments [95, 43, 82]. The magnetic field is generally more vertical in bright filaments. In particular, this is true near the outer penumbra where dark filaments are found to be nearly horizontal and the difference in inclination angle between the more vertical bright filaments and nearly horizontal dark filaments can be $30-40$ deg [42].

Based on these observations, the picture of an interlocking-comb magnetic field structure, sometimes also referred to as "uncombed" or "fluted" penumbra, has now been generally accepted as an accurate description of the penumbral magnetic field geometry. For a comprehensive discussion of this subject, see Thomas & Weiss [92, 93], Solanki [81], or Weiss [103]. This intuitively rather unexpected field geometry has been referred to as "the most remarkable feature of sunspot magnetic fields" [93].

The extensively studied Evershed flow [22] is an important aspect of penumbral physics that has to be explained by any successful penumbral model (see e.g., Thomas & Weiss [93]). It is also considered a firmly established fact that the Evershed flow occurs along magnetic field lines [7, 11], as fundamental physics would dictate, if the Evershed

effect is indeed caused by a plasma flow. The correlation between Evershed flow and dark and bright filaments has been somewhat controversial amongst observers. The early observations of Beckers & Schröter [6] found the Evershed effect co-located with dark filaments. Other observers did not find a clear correlation between Evershed flow and dark filaments [105, 44]. Although, a number of more recent high-resolution observations seem to provide strong evidence that the Evershed flow is predominantly co-located with the horizontal dark filaments (e.g., Title et a. [95], Langhans et al. [42]). This correlation is generally stronger when velocity and intensity signals are compared that form at approximately the same height in the atmosphere [59, 60, 84]. Schlichenmaier et al. [74] find that the correlation between dark filaments and Evershed flow varies from the center-side to the limb-side penumbra and high correlation coefficients are found only locally. They conclude that the Evershed flow occurs in bright and dark filaments.

The Evershed flow is time dependent. Flow speeds along flow channels vary on time scales of 10-20 minutes [58, 79]. It has been argued that since these velocity variations are observed to be coherent across several individual flow channels, it could mean that they are produced by larger scale wave motions and are not an inherent feature of the actual Evershed flow along individual flux tubes [93].

Theoretical models of the Evershed effect include the siphon flow model [91, 49, 50] and the moving tube model [76, 75].

A siphon flow develops along a magnetic loop, the foot-points of which have different magnetic field strengths (at the same geometrical height). The resulting pressure difference between the two foot-points of a penumbral flux tube is what drives the Evershed flow.

Schlichenmaier et al. [76, 75] model the dynamic evolution of a thin flux tube inside the penumbra. A flux tube initially located at the magneto-pause becomes buoyant due to radiative heating and rises. Radiative cooling at the photosphere produces pressure differences along the loop, driving an outward directed flow along the flux tube as it rises through the penumbra. Thomas & Weiss [93] point out that the outflow produced by the moving

tube model is also the result of a gradient in gas pressure and requires, or leads, to higher magnetic field strength at the outer foot-point as well. The moving tube model assumes that, at the outer edge of the penumbra, flow continues along the magnetic canopy, where an open boundary condition is applied.

The moving tube model predicts (see Figure 8 of Schlichenmaier et al. [76]) a filament structure that is characterized by a nearly vertical upflow of hot plasma at the foot-point of the filament, which within a few hundreds km turns into a horizontal filament that is elevated by about 100 km above the surface. Schlichenmaier et al. discuss observational consequences of their model. At the foot-point, the tube's temperature is higher than the corresponding background values and the foot-point appears bright. They point out that the foot-point of the flux tube inside the penumbra could therefore be identified with a bright penumbral grain. It has also been suggested that these hot upflows may significantly contribute to the heating of the penumbra [77]. As the plasma flows along the tube, it loses energy by radiation, cools off and the filament becomes as dark as its surroundings. Rimmele [62] observed flows along penumbral filaments that seem to confirm at least some aspects of the Schlichenmaier et al. [76, 75] model.

Inspired by the discovery of dark-cored penumbral filaments [71], Spruit & Scharmer [83] have taken a new look at penumbral structure. The authors propose that penumbral filaments are due to convection in field-free, radially aligned gaps just below the visible surface of the penumbra. The model produces a horizontal magnetic field at some height above the gap, producing an environment for where the Evershed effect could occur. Although, the authors do not claim that the model provides a satisfactory explanation for the Evershed effect.

Although direct measurements of magnetic field were not performed, the highly resolved observations of penumbral flows, and in particular, their temporal evolution over an extended period of time presented here do provide important clues that will aid in the inter-

pretation and development of realistic penumbral models. For example, the observations provide strong evidence in favor of some aspects of the moving tube model.

7.2 Observations

The observations and experimental setup were described in Chapter 6. Due to the broad filter passband of the UBF, the measured wing intensities are averages over large portions of the wing profile. Hence, the dopplergrams give a measure of the mean position of the spectral line. Obviously, any information about vertical gradients in the velocity field is lost due to the coarse spectral resolution. However, this velocity is relatively insensitive to changes of the spectral line profile and line parameters, such as line strength and line width, that have been observed to occur on filamentary scales in the penumbra [59, 60, 96].

The exposure time was set to 1.2 seconds. The data was collected over a period of about 118 min, during which the seeing conditions were good but variable. The high-order adaptive optics system [63] was deployed to correct for the majority of atmospheric seeing. However, variable seeing conditions result in a variable degree of correction of the adaptive optics system and hence, in variations of the PSF observed at the detector plane. In order to alleviate the negative impact of these residual seeing effects, the PSF estimation method was used to obtain a PSF from the AO loop data for each individual UBF exposure.

7.3 Data Reduction

After flat and dark correction, the UBF filtergrams were corrected for residual differential image motion visible across the extended field-of-view by using a destretch algorithm. A total of 286 dopplergrams (line-of-sight (LOS) velocity maps) were computed in the usual manner, indicated by Equation (6.1).

The dopplergrams were calibrated using a calibration factor (c_v) determined by modelling the spectral profile of the observed lines as observed through the UBF. The Liege atlas profile was convolved with the UBF passband, from which the calibration factor can

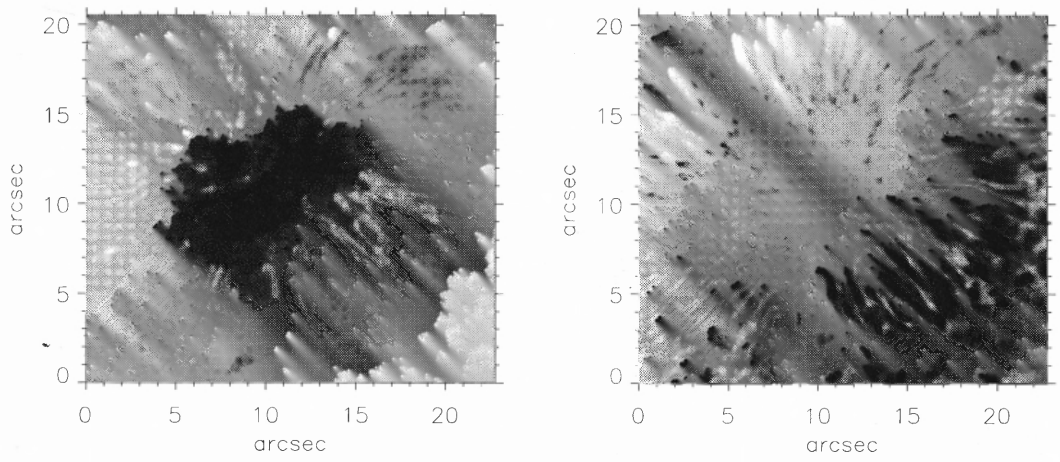

Figure 7.1 Narrow band filtergram and dopplergram of a sunspot observed at approximately 30 deg away from disk center. The effective exposure time was 2.4 sec. The dopplergram shows the signature of the Evershed flow. The solar limb is toward the upper left corner of these images. The velocities are encoded in a grey-scale: bright: red-shift; dark: blue-shift.

be derived directly. Blue- and red-wing filtergrams were added to produce an intensity map corresponding to approximately the same height as the velocity map. The effective exposure time for both velocity and intensity maps is therefore 2.4 sec. The average umbral velocity was chosen as the velocity reference point (see Rimmele [60]).

The PSF estimated from the stored AO telemetry data is used to de-convolve the corresponding UBF filtergrams using a simple Wiener filter. Through this process, the amplitudes (contrast) of the images can be recovered within the stated accuracy. This is of particular importance for the dopplergram observations performed for this study. Since the filtergrams that are combined to generate a dopplergram are not taken simultaneously, variations in the PSF between the subtracted filtergrams can cause spurious velocity signals.

Using the PSF estimation technique, a consistent high-resolution time sequence of filtergrams and dopplergrams was produced in spite of variable seeing conditions. Figure 6.29 plots the image contrast of the filtergrams as a function of the Fried parameter as observed (left) and after deconvolution with the estimated PSFs. The contrast of the "raw" sunspot filtergrams shows a strong dependence on r_0, in particular for Fried parameters

$r_0 \leq 7$ cm. This deterioration of AO performance (Strehl ratio) for small Fried parameters is expected since the WFS subaperture size is 7.5 cm. In the deconvolved sequence, the contrast curve is flatter, indicating that the PSF estimation accurately recovers the amplitudes over a wide range of seeing conditions. However, the plot also indicates that for small Fried parameters the reconstruction becomes more noisy.

A temporal low pass filter was applied to the time sequences of filtergrams and dopplergrams in order to remove fast (e.g. 3 min, 5 min) oscillatory components that are definitely visible in the time sequences. The goal was to study the long term evolution of the Evershed flow. The filter therefore passes frequencies ≤ 2 mHz. It has been shown that the Evershed effect evolves within 10-20 minutes [58, 79], time scales, which are well within the applied filter.

Due to the large number of photons collected (long-exposure), the statistical errors in the doppler measurements can be estimated from the photon noise in the filtergrams and Equation (6.1), and are of order 20 m/s only. Although the adaptive optics greatly reduces seeing effects, the dominant error source in the uncorrected and unfiltered dopplergrams remains residual seeing effects. These random errors can be estimated from the difference of consecutive dopplergrams. Because of the fast cadence, a significant change of the penumbral structure is not expected, i.e., the difference signal is mostly due to residual seeing effects. The random errors are of order 65-80 m/s. These relatively small random errors are reduced further to about 15 m/s by applying the PSF deconvolution and the temporal filter.

7.4 Results

Figure 7.1 displays a sample narrowband filtergram and the corresponding dopplergram. In spite of the long exposure time, the resolution of both intensity and velocity maps is excellent and approaches the diffraction limit. Bright points as well as umbral dots are

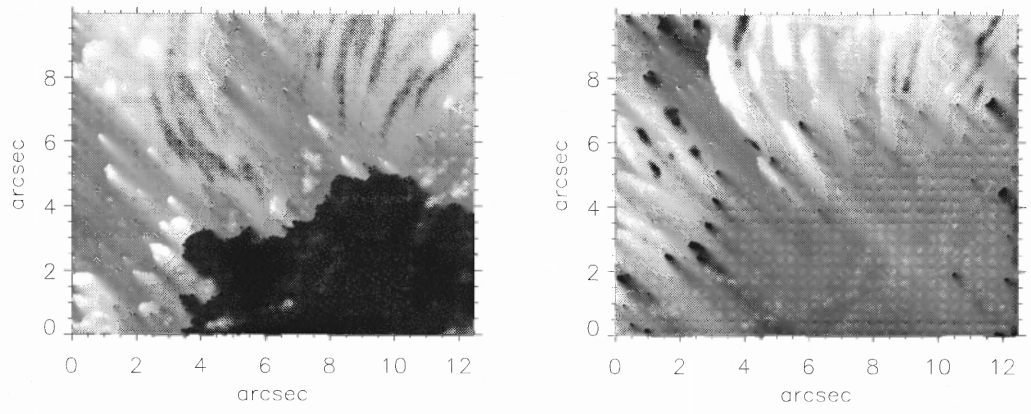

Figure 7.2 Dopplergram and filtergram of limb side penumbra. In the dopplergram, dark (bright) areas mark blue-shifted (red-shifted) line profiles.

visible in the filtergrams. This demonstrates the effectiveness of the real-time adaptive optics correction combined with the post-facto PSF correction.

It has been well established that the Evershed flow is mostly horizontal, i.e., parallel to the solar surface (see Thomas & Weiss [93] and references therein). Since the spot was observed away from disk center, the observations present the typical Evershed flow pattern that is characterized by red-shifted spectral profiles in the limb-side penumbra (upper-left) and blue-shifted profiles in the center-side penumbra (lower right). The measured maximum velocities range between ± 2 km/s. Assuming a perfectly horizontal flow and taking into account projection effects ($\sin\Theta = 0.5$) this amounts to a flow amplitude of about 4 km/s. Previous publications quote Evershed velocities of up to 6 km/s. However, the broad filter band path and the resulting averaging over large portions of the spectral line wings have to be considered, which will lower inferred velocity amplitudes compared to bisector velocities from high spectral resolution data.

The limb-side penumbra allows to distinguish between upflow (blue-shift), which is expected to appear at the inner foot-points of Evershed loops, and outflow (red-shift) along the nearly horizontal magnetic flux tube. In the center-side penumbra, both flow components would appear blue-shifted and are not easily distinguished in simple LOS dopplergrams. Therefore, all further analysis will be focused on the limb side penumbra.

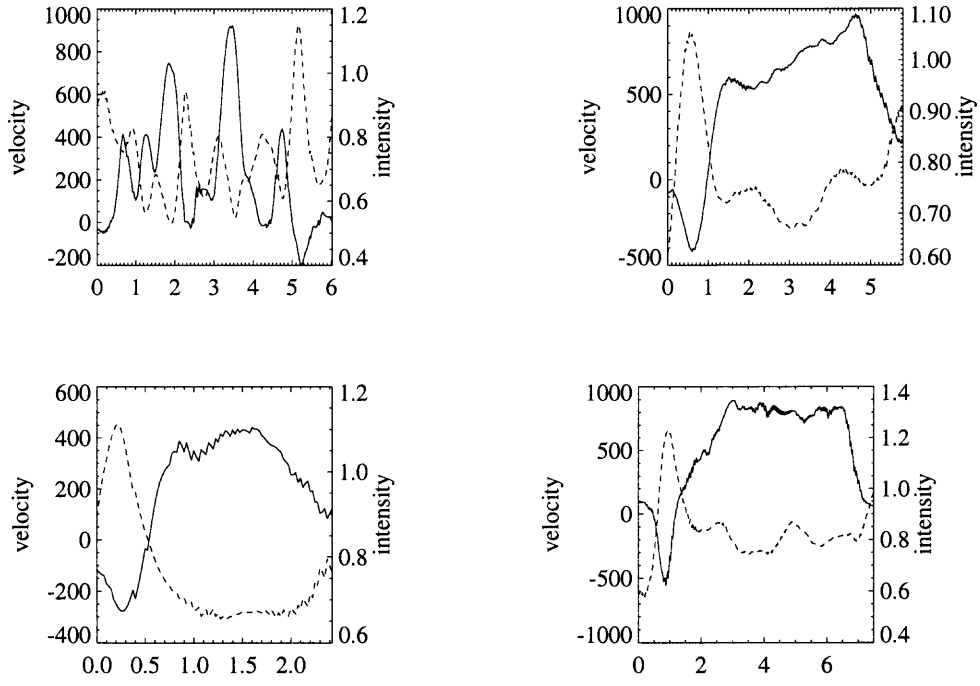

Figure 7.3 Velocity (solid line) and intensity (dashed line) traced across several filaments in the azimuthal direction (top left) and along several filaments in the radial direction (top right and bottom). Positive velocity corresponds to red-shifted line profiles. Note the strong correlation between upflow and bright point as well as the steep decrease of the velocity near the outer penumbra.

7.4.1 Evershed Flow Geometry Along Individual Filaments

Figure 7.2 shows velocity and intensity maps of a portion of the limb side penumbra. The most striking features seen in the velocity map are the small-scale (0.″2) upflows located at the inner foot-points of a flow-channel. Close inspection reveals that upflows can be identified at the inner foot-point for most, if not all, of the observed flow channels. Comparing intensity and velocity maps, it can be seen that upflows are co-located with bright features, often referred to as penumbral grains, while the horizontal red-shifted outflow is correlated with a dark filament. This correlation persists throughout the entire time sequence. Complex, twisted and entangled, flow channels make a clear association more difficult (but not impossible). Hence, a few flow channels that can be easily identified as individual filaments

are selected and their velocity and intensity signals in the radial direction are traced, i.e., along the axis of a filament, and across several filaments in the azimuthal direction.

Figure 7.3 shows representative examples of such traces from different time steps in the sequence. The azimuthal trace clearly shows that red-shifted flows are correlated with dark filaments. The radial traces start at the inner penumbral foot-point and reveal the close correlation between upflow and "bright grain" as well as the correlation between horizontal flow and dark penumbral filaments. The intensity of the bright points usually exceeds the average brightness of the surrounding "quiet" photosphere. Values in the range of I = 1-1.2I_{quiet} are observed. The spatial extent of the foot-points along the filament axis can be inferred from Figure 7.3, and is of order $\leq$ 0.″4 FWHM. Another striking feature is the very steep rise of the velocity signal from blue-shift to red-shift. The intensity signal changes from bright to dark over a similar distance. The distance over which the flow turns from vertical to horizontal (bright to dark) is measured as the distance between the location of the maximum upflow and the point from where the red-shift only increases gradually. This distance is between 0.″5 and 0.″7 or 380 - 500 km.

After the initial steep rise, the velocity increases only very gradually along the radial direction until at the outer boundary of the flow channel, which is usually located near the outer edge of the penumbra, an even steeper decline of the velocity is observed. The length of the flow channels in the radial direction for different filaments ranges from 2-6″(1500 - 4500 km), as measured from the maximum upflow to the half-point of the velocity drop-off at the outer edge. In a very few cases, unusually short and short-lived filaments were observed (see Figure 7.6).

7.4.2 Flows in Dark-Cored Penumbral Filaments

The resolution in the observations is good enough to identify dark-cored penumbral filaments [71] in the filtergrams and study the corresponding flow pattern. The observational signature of dark-cored filaments is a narrow dark line in the center of bright filaments,

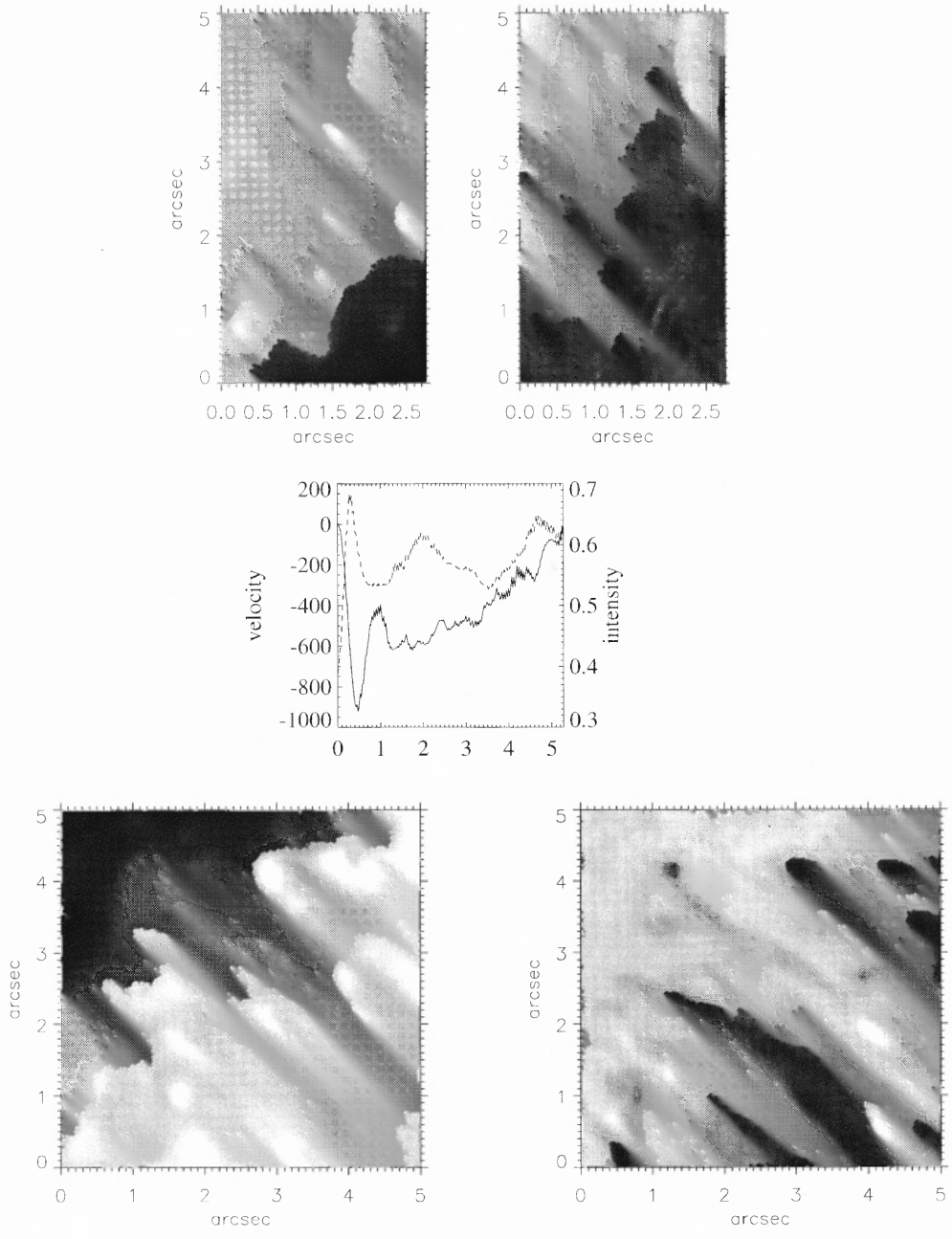

Figure 7.4 Close-up view of dark-cored penumbral filaments. Left - intensity; right - velocity. Examples from the limb-side (top) and center-side penumbra (bottom) are shown. The artifacts seen in the velocity map of the upper right panel are from an imperfect flat field correction. The line plot shows a trace along a dark core in the center-side penumbra. Solid - velocity; dashed - intensity. The plot demonstrates that the bright, inner foot-point of the flow channel is seen also for the center-side penumbra.

extending along the axis of the filament. Dark-cored filaments are often seen in the inner penumbra but are more easily identified in the center-side penumbra [88], suggesting that they are shallow features [103].

Figure 7.4 shows scaled up images of a small section of the inner penumbra that contains what are believed to be dark-cored penumbral filaments. Examples are shown from both the limb-side and center-side penumbra. The corresponding velocity map is shown as well.

These images provide strong evidence that dark-cored penumbral filaments originate at bright foot-points near the penumbra-umbra boundary. The foot-points are co-located with upflows, while the dark-cored filaments are co-located with red-shift. The dark cores appear to be the signature of well defined individual flow channels. It should be noted again that the dark portion of the filament is associated with the more horizontal outflow, which is usually referred to as the Evershed flow. Figure 7.4, and in particular, the temporal evolution of these flows and their corresponding intensities, which are discussed in the following section, provide strong evidence that bright grains and dark cores belong to the same flow system along a magnetic loop, i.e., they are both signatures of the Evershed effect.

It is not entirely clear whether the examples of dark-cored filaments shown from the limb side penumbra fit the definition given above as well as the examples from the center-side do. However, the flow pattern measured along the dark cores is very similar in the center-side and limb-side penumbra, respectively. The line plot in Figure 7.4 shows a trace along a dark-cored filament from the center-side penumbra. Also here, a bright blue-shifted feature at the inner foot-point of the flow channel is clearly seen. Instead of the quick turn to red-shift observed in the limb-side penumbra (Figure 7.3), the velocity remains blue-shifted as would be expected for a horizontal outflow simply because of LOS effects.

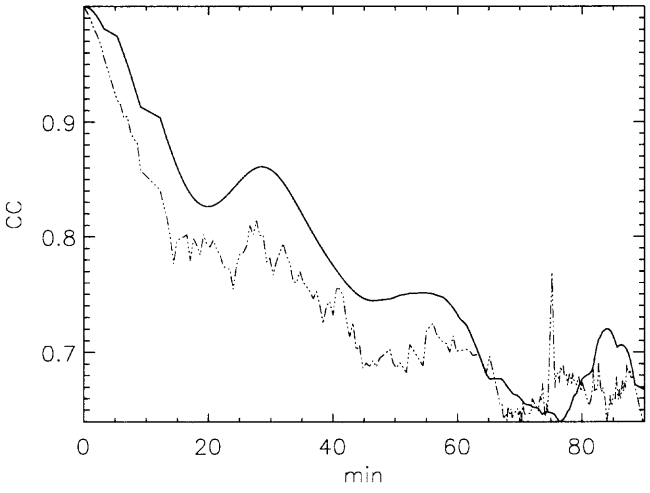

Figure 7.5 Maximum of 2-D cross-correlation as a function of time of velocity (solid) and intensity (dashed) maps seen in Figure 7.2. Note the secondary and tertiary maxima at about 30 min and 55 min.

One notes that there are a number of red-shifted areas visible in center-side penumbra. It might be tempting to interpret these downflows as Evershed return flows that have been reported to exist within the outer penumbra [104]. However, the downflows seen in Figure 7.4 are not readily associated with individual flow channels, i.e., as end points (outer foot-points) of the Evershed flows.

7.4.3 Temporal Evolution of Evershed Flows

Before discussing the temporal evolution of Evershed flow channels in detail, the temporal cross-correlation of the penumbral velocity field is discussed. The time sequence of 2-D cross-correlation for the field-of-view is computed and shown in Figure 7.2. All consecutive frames were tracked and correlated to this reference. Figure 7.5 plots the maximum value of the 2-D cross-correlation as a function of time for both intensity and velocity. The plots indicate how the scene de-correlates over time and thus, gives an indication of the lifetime of penumbral structures. Low spatial frequencies (general shape of sunspot) de-correlate slowly, which explains the high degree of correlation that persists even after

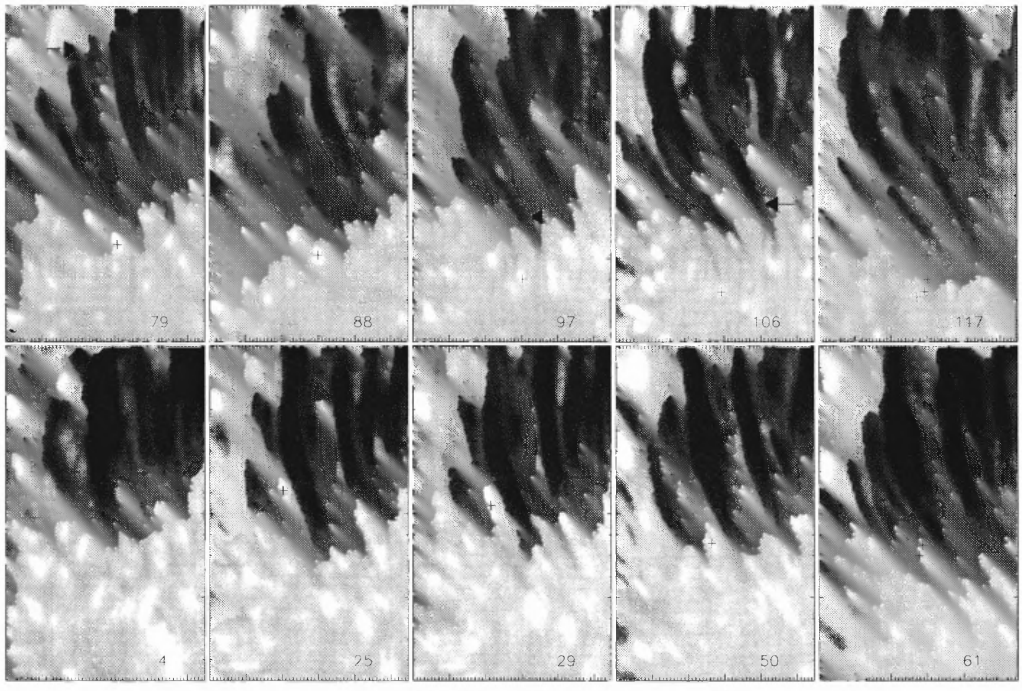

Figure 7.6 Time sequence of velocity maps. The field-of-view is $5.5 \times 8\text{arcsec}^2$. The time steps are 4, 25, 29, 50, 61, 79, 88, 97, 106, and 117 minutes. The temporal evolution of a filament from the time when is first becomes visible to the time when it disappears is depicted with these images. The crosses mark the position of the inner foot-point (upflow), which steadily moves inward toward the penumbra-umbra boundary with a proper motion velocity of about 0.5 km/s. The average proper motion speed can be infered from the distance the inner foot-point traveled (trace marked by the crosses) divided by the traveling time. The gray-scale has been reversed for better showing, i.e., dark: red-shift; bright: blue-shift. The arrows point to examples of "crossing filaments". Evidence for unusually short flow channels is seen near the outer penumbra at time step 79 min (arrow).

90 min. The rapid de-correlation observed in the first 15-20 min, followed by a more gradual drop in the correlation, can be explained by the evolution of penumbral fine-structure.

An interesting feature of these cross-correlations is the presence of the secondary and tertiary maxima at about 30 min and 55 min. These maxima are clearly visible in the velocity correlation and less prominent in the intensity correlation. These secondary and tertiary maxima might be the signature of recurring Evershed flow packages observed by Rimmele [58] and Shine et al. [79].

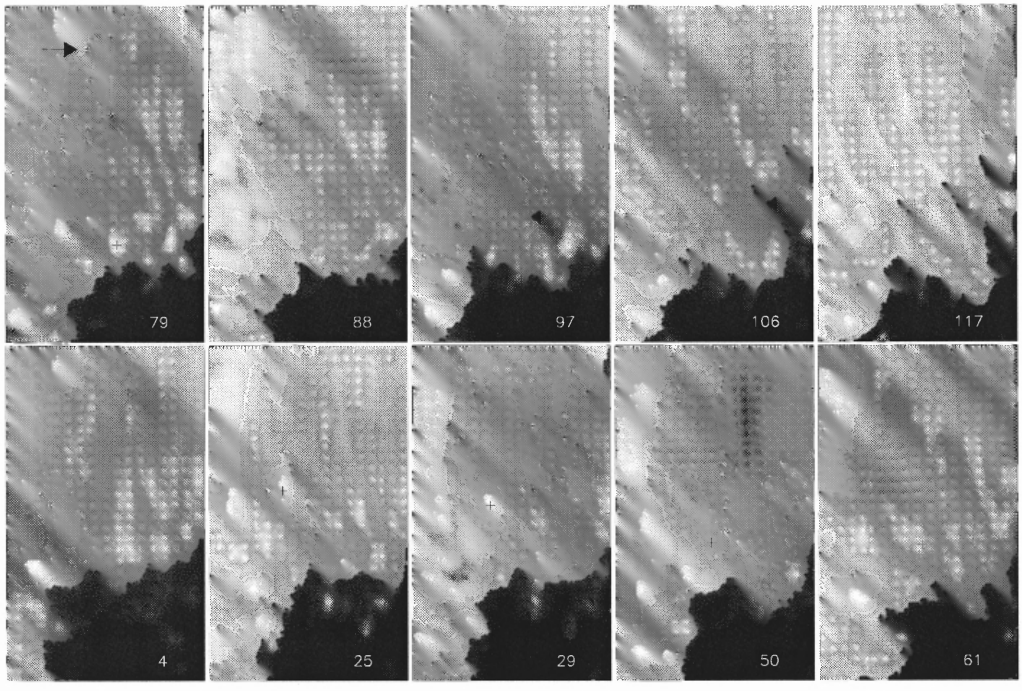

Figure 7.7 Time sequence of intensity maps corresponding to velocity maps shown in 7.6.

Figures 7.6 and 7.7 show a time sequence of doppler maps and corresponding intensity maps. A small $5.5 \times 8''$ field-of-view was selected in order to highlight, at high spatial resolution, the evolution of a few representative penumbral filaments. The flow channels were followed throughout their entire life cycle. A particularly nice example is marked in this sequence of velocity and intensity images. A flow channel first appears in an area that shows what appears to be a small scale convective pattern penetrating into the outer penumbra. The crosses mark the position of the inner foot-point of the flow channel. Immediately after the flow channel appears, the inner foot-point begins to move inward, while during these early stages, the outer boundary of the flow remains more or less in the same position, i.e., the filament becomes more and more elongated. At time step 61 min, two foot-points become visible, indicating that either previously unresolved flow channels have split sufficiently to become resolved or a second flow channel has emerged at this location.

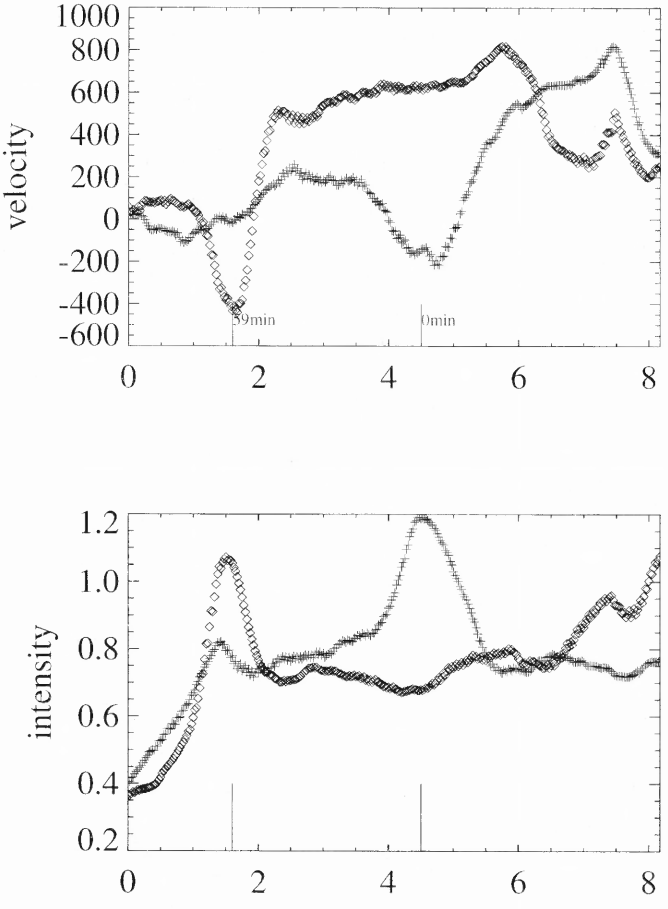

Figure 7.8 Evolution of flow channel. Traces are in the radial direction along the filament axis. Velocity - top; intensity - bottom. Two different time steps are shown. 0 minutes (crosses) refers to shortly after the flow first becomes visible. At 59 minutes (diamonds), the inner foot-point has moved close to the penumbra - umbra boundary. The sharp outer edge of the flow channel is seen to move inward as well.

Continuing with time step 79 min, a single flow channel is observed again, the clearly visible foot-point of which continues to move toward the penumbra-umbra boundary. During these later stages of the evolutionary path of the filament, the outer edge of the flow channel is moving inward as well (see also Figure 7.8). Adjacent flows channels behave in a similar fashion. Toward the end of the sequence, the flow channel fades away. The crosses in the last image trace the motion of the inner foot-point from the outer parts of the penumbra to

It is difficult to determine whether, near the beginning of the time sequence, an already existing flow emerges into the photosphere from below or whether the flow forms in a magnetic loop that is already located in the photosphere. Figure 7.9 holds a possible clue. It shows the evolution of the velocity along the radial axis of the flow as the flow channel appears. The fact that blue-shift and red-shift appear simultaneously argues in favor of an emerging pre-existing flow. One would expect a flow that develops along an existing magnetic loop to first exhibit an upflow at the inner foot-point that then propagates outward. Also, this figure shows that the location of the maximum blue-shift begins moving inward immediately, which is consistent with an emerging loop as well.

7.4.4 Stacked Flow Channels

Numerous examples where penumbral flow channels appear twisted and tangled, apparently crossing each other in the azimuthal direction, were identified in the data. This is more readily seen in the velocity maps where individual filaments (flow channels) can be more easily identified and traced. The arrows in Figure 7.6 point to two examples where such crossing of filaments (flows) occurs giving the impression that, at the crossing points, filaments are stacked on top of each other in the vertical direction. Areas ("filaments") with velocities close to zero were observed in between flow channels (or clusters of flow channels).

An increased velocity signal is observed at the location where flow channels cross. Figure 7.10 shows an example with traces in the radial and azimuthal direction. Near the inner foot-points, two individual filaments are clearly resolved. Moving toward the outer penumbra to the location where the flow channels cross, the flows are no longer distinguishable as separate features, whereas near the foot-points, the structures are clearly distinguishable. At the crossing point, the flow amplitude suddenly increases. This is not surprising since two flow channels, stacked on top of each other, fall within the response function of the line, i.e., the "velocity fill factor" increases in both the vertical and the

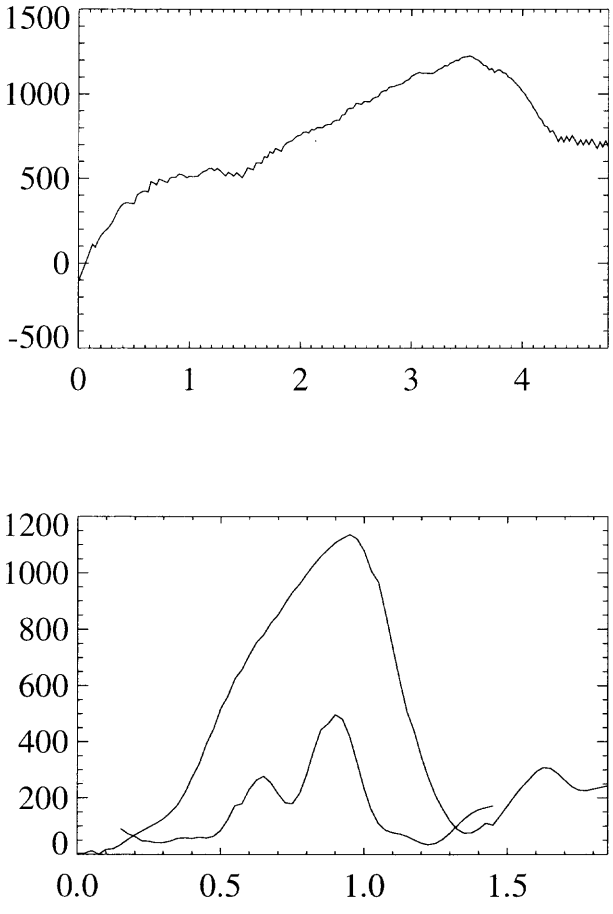

Figure 7.10 Evidence for crossing flows. Bottom: Trace across two individual flow channels near their foot-points, where the channels are resolved, and at the point where the two filaments appear to cross in azimuthal direction and at different heights. Top: Trace in along one of the flow channels. At the location where the flow channels cross, a sudden increase in flow velocity is observed.

horizontal direction. It would be interesting to obtain observations of crossing filaments with significantly higher spectral resolution in order to study the line profiles and bisectors. One would expect to see more complex bisector shapes at locations where two (or more) flow channels, located at different heights in the atmosphere, fall within the line response function. Unfortunately, due to the broad filter bandpass of the UBF, the data is not suited to provide bisector information.

7.5 Summary and Discussion

The main observational results can be summarized as follows:

- Individual Evershed flow channels can be identified. The lateral extent of the flow channels is very near the diffraction limit of the telescope ($\leq 0.''2$) and it is expected that the actual size of the flow is even smaller.

- Flow channels often seem to cluster together, forming a conglomerate of what appears to be twisted, tangled flow channels that in many cases cross each other at different heights in the atmosphere. One might also call this "uncombed" penumbral structure. Penumbral fine-structure appears to be even more complex than the already surprisingly complex, widely accepted interlocked-comb structure would suggest.

- Areas ("filaments") with velocities close to zero are observed in between flow channels (or clusters of flow-channels).

- For the vast majority of Evershed flow channels, an inner foot-point can be identified. This inner foot-point is a small ($\leq 0.''2$), point-like, bright upflow that has all the characteristics of what has been dubbed "bright penumbral grain" in the literature, i.e., penumbral grains are the inner foot-points of Evershed flow channels. The plasma flow vector - expected to be field aligned [7] - is more vertical in bright up-flowing regions that mark the inner foot-points of the Evershed flow.

- The intensity of the bright points generally exceeds the average photospheric brightness.

- The upflows observed in the bright foot-point turn over to nearly horizontal flows within a very short distance ($0.''5$ - $0.''7$ or 380 - 500 km). This horizontal part of the flows is correlated with a dark filament or dark-cored filament.

- In the limb-side penumbra, the LOS component of the upflow is of order 300 - 500 m/s. In the center-side penumbra, LOS upflow velocities of order 1000 m/s are observed.

- An extremely steep decline of the velocity at the outer end of the individual flow channels is observed. It is interesting to note that this sharp drop-off of the Evershed flow is not only found at the outer edge of the penumbra, but also flow channels that end well within the penumbra (even the inner penumbra) exhibit this rapid decay over, in some cases, less than 0.″2.

- The observations allowed the study of the temporal evolution of individual flow channels, tracing them from first appearance to their disappearance. While undergoing significant changes, individual flow channels could be traced as they moved from the outer parts of the penumbra to the penumbra-umbra boundary, where the flow disappears. The proper motion speed was determined by tracking the upflow at the foot-point (see caption of Figure 7.6) and is of order 0.5 - 1 km/s, consistent with previous measurements of penumbral grain motion (e.g., Sobtka & Sütterlin [80]).

- The bright inner foot-point and the dark filament carrying the horizontal outflow are observed to evolve as a unit. This indicates that dark filaments and bright grains are part of the same magnetic flux tube that carries the Evershed flow.

- In the early stages of the flow channel's evolution, the sharp outer edge of the flow channel appears stationary, while during the later stages, the outer edge is observed to also move inward toward the penumbra-umbra boundary. Often, this outer edge is located well within the penumbra.

- A flow channel was observed to emerge from a convective pattern near the outer penumbra. The emerging flow is "arch like" (inverse U). As the flow moves inward

and in particular, near the penumbra-umbra boundary, the flow geometry is characterized by the quick turn from upflow to horizontal flow over a very short distance.

- The length of the majority of flow channels varies between 1500 km and 4500 km. A few examples of much shorter loops were observed in the outer penumbra.
- The life-time of flow channels ranges from 30 - 115 min.
- Several red-shifted (downflow) areas are found in parts of the center-side penumbra (where the Evershed effect produces blue-shift). However, an interpretation of these downflows as Evershed return flows is not obvious from these data.

Many aspects of the observations provide strong support for the the moving flux tube model [76, 75]. The observed flow geometry, i.e., upflow in a bright inner foot-point that quickly turns over into a nearly horizontal flow along a dark filament, is consistent with predictions of this model. Evidence was also found of Evershed flow channels emerging (rising) from below. However, the most important observation is considered to be the fact that the bright inner foot-point and the dark filament carrying the more horizontal outflow are moving and evolving as a unit, which provides convincing evidence that dark filaments and bright grains are part of the same moving magnetic flux tube that carries the Evershed flow. The Evershed flow is associated with both bright and dark penumbral features. The observed movement of the inner foot-points of those flow channels toward the penumbra-umbra boundary with a velocity of about 0.5-1.0 km/s is also expected in the moving tube model. The time scales predicted by the moving tube model for the rise of a penumbral flux tube (120 min) are consistent with the ones observed here.

In order to estimate the actual upflow velocity at the inner foot-points of the flow channels, it is assumed that the inclination of the flow vector at the inner foot-point is between $35 - 40$ deg, which is the inclination measured for the strong magnetic field component by, e.g., Borrero et al. [11] and Langhans et al. [42]. Taking into account the sunspot's

position on the disk (LOS angle), an upflow amplitude of 0.5 - 1.5 km/s in the bright foot-points seen in the limb-side penumbra is obtained. For the particular observing position of the sunspot on the solar disk, no LOS correction is necessary in the center-side penumbra, i.e., the LOS is such that the actual upflow amplitude is observed in the center-side penumbral grains. In the example of Figure 7.4, the flow amplitude is 950 m/s. These upflows are significantly slower than predicted by the moving tube model, which predicts velocities around 3 km/s at the foot-point [76]. The horizontal outflow amplitudes measured are lower, as are the ones predicted by these models. The discrepancy might be explained by still insufficient spatial resolution and the definitely insufficient spectral resolution.

Heating of the penumbra by hot upflows along magnetic flux tubes has been suggested as a mechanism sufficient to explain the penumbral brightness [77]. However, Spruit & Scharmer [83] argue that these upflows are unlikely able to provide the uniform heating of the penumbra along its length.

The occurrence of downflows associated with Evershed flow channels is not obvious at all in the data. This is somewhat inconsistent with the findings of bisector analysis of the same spectral line (Fe I 5576 Å) by Schlichenmaier et al. [73], who concluded that downflows in deep layers can explain line asymmetries. Again, the poor spectral resolution of the observations combined with the formation height of the line (mid-upper photosphere), may prevent the detection of flows deep in the atmosphere.

The interpretation of the observations in the context of alternative models of the Evershed effect is less straightforward. Siphon flow models are stationary solutions and therefore, are inherently unable to explain the temporal evolution and proper motion of the flow channel, which is found to be common to the vast majority of Evershed flow channels.

It is difficult to reconcile the observations with the penumbral model recently put forward by Spruit & Scharmer [83]. As mentioned in the introduction in this model, the Evershed flow is located along horizontal magnetic field structures above the field-free gap. However, the flow can only be carried over a finite distance, i.e., the model is allowing for

flows that are rather local and transient. The authors, therefore, suggest that a local version of the moving tube model might be at work. They further suggest that even though, on average, the Evershed flow may appear smooth and steady, it is in fact a locally transient phenomena. However, the observations suggest a more global version of the moving tube model to be responsible for the Evershed effect. Flows that are consistent with the predicted flow geometry, extending over and propagating across large parts of the penumbra, are found. There is a single well defined inner foot-point and the lifetime of the flows can be as much as 2 hours. It seems unlikely that these observational signatures could be produced by a local phenomenon. The observed hyper-structure of dark-cored filaments can be explained by un-combed penumbra models, i.e., it can be produced by a nearly horizontal flux tube embedded in a more vertical background field [10].

Although the definition of a dark-cored filament is rather subjective, it is found that dark-cored filaments are more readily identified near the penumbra-umbra boundary in the center-side penumbra, which agrees with previous observations [11, 88, 42]. The dark-cored filaments are identified as such in both the limb-side and center-side penumbra and can be associated with regular Evershed flow channels originating in a hot upflow. This is consistent with Bellot Rubio et al. [8], who find flows in dark-cored filaments to be mostly horizontal with a small upflow component. Interestingly, it is found that for this data set, a clear identification of bright and dark filaments is difficult for the limb-side penumbra and much more easily performed for the center-side penumbra.

It is still considered a major challenge to explain the sharp ($\leq 0.''2$) outer boundary of the Evershed flow. The fact that this sharp decline in velocity is seen not only at the penumbra-quiet sun boundary, but throughout the penumbra, argues that this effect is not linked to the transition between magnetic and non-magnetic environments. If the sharp boundary is due to the flow returning back down to the surface, the flux tube would have to bend over a short distance and at least as quickly as it appears to be the case for the inner foot-point. Flux pumping [94] may provide the physical mechanism. However, Spruit &

Scharmer [83] argue that the flux pumping mechanism is inconsistent with observations since it only operates at the outer penumbral boundary and not within the penumbra.

It is further noted that as the inner foot-point moves toward the inner penumbra, the outer edge of the flow (sharp boundary) initially stays in place (Figures 7.6 and 7.9), but then also begins to move inward (see Figure 7.8). If the sharp boundary marks the onset of the return flow, this would indicate that the outer foot-point, where the downflow occurs, moves inward as well. If the the sharp edge marks the point where the flow enters into the canopy instead, as was suggested by Schlichenmaier et al. [76, 75], it does so in many cases well within the penumbra. More recent simulation results of the moving tube model [72] produce flux tubes that, during the course of their evolution, develop sea serpent like waves. This structure produces downflow arches well within the penumbra. However, any direct evidence for such sea serpent like flow structures is not seen in the data.

Infrared polarimetric observations of penumbral fine-structure using two-component inversion techniques have provided strong evidence for the picture of an un-combed penumbra, where the penumbra is composed of a penumbral flux tube embedded in a magnetic background field [11]. The authors study a portion of the limb-side penumbra, as in this paper, and find that the Evershed effect is confined to the nearly horizontal flux tube component, while the background is essentially at rest. The lower right panel of Figure 4 of Borrero et al. [11] plots the inferred velocity as a function of penumbral radius for the flux tube and the background component, respectively. The flux tube component shows positive velocity throughout the penumbra, consistent with a horizontal outflow that is only slightly inclined in the inner penumbra. The background field is at rest only in the outer penumbra, while in the inner penumbra the inversion returns negative velocities (upflow). However, the observations clearly indicate that upflow occurs in the flux tube component at their inner foot-points. It should be noted that the upflow derived for the background component from the inversion is very similar to the upflows observed at the filament foot-points (order 1 km/s).

In order to understand this apparent discrepancy, one must keep in mind that vector magnetic field measurements often suffer from a lack of spatial resolution [46]. This is particularly true for infrared observations, such as the ones by Borrero et al. [11]. In fact, this is the primary reason for deploying multi-component inversion techniques. However, uniqueness of the solution is often a question and the interpretation is often difficult compared to observations where the structures are resolved. Borrero et al. [11] estimate the achieved spatial resolution to be about 1″. At this resolution, upflows with spatial extent 0.″2 - 0.″5 and horizontal outflows are definitely mixed together in one pixel. It appears that the two-component inversion associates the upflows with the background component and the horizontal flows with the flux tube component, even though according to the data, they are part of the same flow along a flux tube.

The observations have provided important new information that will help to distinguish between various existing penumbral models and hopefully will lead to more refined penumbra models. In the future, these observations must be confirmed and extended by obtaining simultaneous high-resolution vector magnetic field measurements, which are now possible with instruments such as the Diffraction-Limited Spectro-Polarimeter (DLSP).

The DLSP is fed by a high-order adaptive optics corrected beam and can deliver 0.″2 resolution vector polarimetric data. The DLSP can be combined with the UBF and high-resolution g-band speckle imaging, which might provide additional important information.

CHAPTER 8

CONCLUSIONS

AO systems have revolutionized ground-based astronomy by providing real-time correction for atmospheric aberrations and producing diffraction-limited observations. However, due to limited temporal and spatial bandwidth, wavefront sensor noise and other limitations, the correction provided by the AO system is only partial. With knowledge of the PSF, the image quality can be further improved.

The lack of point sources in the field-of-view of solar observations makes a direct measurement of the PSF impossible. A method to estimate the long-exposure AO-corrected PSF from AO loop data was implemented specifically for the solar AO systems at the DST and at the Big Bear Solar Observatory. This method was originally proposed by Jean-Pierre Véran [101, 100] and implemented for the PUEO AO system at the Canada-France-Hawaii Telescope. However, PUEO is a curvature wavefront sensor based AO system designed for night-time observations. The method was adapted for a correlating Shack-Hartmann based AO system designed for solar observations.

The method presented estimates the long-exposure PSF from the telemetry data generated by the AO system during its operation. The data measured by the solar cross-correlating Shack-Hartmann wavefront sensor provide information about residual wavefront aberrations, while the commands sent to the deformable mirror provide information about the original seeing conditions. The AO telemetry data produced during the capture of a solar image completely characterize the AO-corrected image quality and contain enough information to produce an estimation of the long-exposure PSF. Using this method, each AO-corrected image can be further corrected using its own estimated long-exposure PSF through post-processing techniques.

The PSF estimation method was tested on real solar observations and an attempt was made to validate it by applying the method to AO-corrected observations of the star Sirius. Although for some cases, the observed PSF for Sirius could be well matched with the corresponding long term PSF estimate from the AO telemetry, a consistent match was not found for all star images. In that sense, the observations of the star Sirius cannot be considered a sufficient validation of the method at this time. Several issues related to the performance of the AO system under low light levels, for which it was not designed, resulted in low correcting bandwidth, high wavefront sensor noise and consequently produced low Strehl ratios in the AO-corrected Sirius images. These problems led to the violation of crucial assumptions required by the PSF estimation method. The stellar observations should be repeated under more favorable conditions, closer to the solar operating conditions. This may require the observations to take place during the night when the seeing conditions are better and there is significantly less background light. Also, the use of different AO hardware, such as a more sensitive and less noisy wavefront sensor camera. Alternatively, a different approach that can be used during solar observations, such as phase diversity or speckle reconstruction, may be attempted. Further validation of the results produced by the PSF estimation method is highly recommended to ensure the accuracy of the PSFs produced.

The utility of the PSF estimation method was demonstrated with AO-corrected solar observations, where an estimation of the AO-corrected PSF is normally difficult. The quantitative measurements and scientific data extracted from the solar observations were significantly improved by deconvolution with the estimated AO-corrected PSFs. This is particularly true for the case of the velocity maps, or dopplergrams, created by combining images captured at different times

The PSFs estimated from the solar observations provide a measurement of the performance of a solar AO system during different seeing conditions. This is the first time such a measurement has been obtained for a solar AO system. The Strehl ratio of the AO-corrected solar images is computed from the estimated PSFs and plotted against the

Fried parameter r_0, which characterizes the seeing conditions. At this time, measuring the Strehl ratio from the estimated PSFs is the only method available to directly measure the performance of the solar AO system. The measured dependency of the AO performance against seeing conditions, which is consistent with the predictions of performance models, has already been used by other researchers [107] to modify speckle reconstruction code to produce more accurate AO-corrected speckle transfer functions.

The PSF estimation method was applied to the study of the photospheric Evershed effect on the penumbra of a sunspot. The observations, published by Rimmele & Marino [57], consist of a 120 min time sequence of diffraction-limited and post-processed velocity measurements. The study of the temporal evolution and structure of the Evershed flow produced new ground-breaking results that demonstrate the connection between the bright penumbral grains as the foot-points of the dark penumbral filaments. Some aspects of the observations provide strong support for the moving tube model of the Evershed flow [76, 75]. These observations have already guided the development of new penumbral models [35].

The dopplergrams computed from the solar observations were significantly improved by deconvolution of the wing images with their corresponding estimated PSFs. In particular, dopplergrams computed from wing images that were captured during very different seeing conditions presented very high intensity-velocity cross-talk that would, in some cases, destroy the velocity signal. Deconvolution of the wing images significantly reduced the amount of intensity-velocity cross-talk in the dopplergrams and greatly improved the quantitative velocity measurements obtained.

A very interesting result obtained from the observations relates to the RMS contrast values of the quiet sun granulation obtained after deconvolution. The obtained values (around 9%) are significantly lower than the contrast values predicted by current MHD models, which predict a quiet sun granulation contrast of >20%. This result is consistent with the latest results obtained by other image reconstruction techniques, all of which are unable to reproduce the high RMS contrasts of MHD models [98].

The long-exposure PSF estimation method from AO loop data can be applied to a wider range of solar observations and techniques. It is especially useful for applications that require accurate high-resolution quantitative measurements produced by combining data captured at slightly different times, such as magnetograms and dopplergrams. The method can benefit the observations produced by advanced instruments, such as the interferometric bi-dimensional spectrometer (IBIS), the diffraction-limited spectro-polarimeter (DLSP), the spectro-polarimeter for infrared and optical regions (SPINOR) and other post-focus instrumentation at the DST, Big Bear Solar Telescope and other solar telescopes.

The method can be expanded to produce long exposure PSFs for solar multi-conjugate adaptive optics (MCAO) systems, which provide correction for a much wider field-of-view. The expanded method could provide long exposure PSF estimates at different points in the field-of-view from the loop data produced by MCAO systems. Different parts of the field can be deconvolved with their corresponding PSF, producing improved quantitative measurements over a larger field.

APPENDIX A

THE KARHUNEN-LOÈVE MODES

Zernike polynomials are usually the functions of choice to describe wavefronts or surfaces in optical applications. They are an orthonormal basis of functions for the space defined inside the unit circle. This is very adequate for optical systems with circular pupils, such as telescopes with annular pupils. Also, the first Zernike polynomials represent the most common aberrations found in optical systems, such as focus, astigmatism and coma. Another advantage of Zernike polynomials is that they can be calculated using relatively simple analytical expressions. In the case of AO systems, Zernike polynomials have been commonly used to describe the wavefront aberrations induced by atmospheric turbulence.

The statistical properties of atmospheric turbulence are well known and are described by the Kolmogorov model [41]. Noll [52] provides an expression for the temporal covariance of the distorted phase when expressed in terms of Zernike polynomials. This temporal covariance is not a diagonal matrix, which implies that different Zernike terms are not statistically independent. Therefore, Zernike polynomials are not the optimal basis of functions to decompose atmospheric phase fluctuations.

Zernike polynomials constitute an orthonormal set of functions, i.e., they are not spatially correlated. Hence, the Zernike polynomials are eigenmodes of the spatial covariance matrix of atmospheric turbulence. However, since the Zernike modes are temporarily correlated for Kolmogorov turbulence, they are not eigenmodes of the temporal covariance matrix. The temporal variations of the Kolmogorov atmospheric turbulence cannot be optimally represented by Zernike modes.

The optimal basis to describe a wavefront distorted by Kolmogorov turbulence is another set of orthonormal functions with a diagonal temporal covariance matrix. Such a set of functions can be calculated by diagonalizing the Zernike temporal covariance matrix

calculated by Noll. These new functions are called Karhunen-Loève functions [15, 16, 70, 102], more commonly known as KL modes.

There is no analytical representation for these KL modes. They are calculated by diagonalizing the covariance matrix expressed in terms of Zernike polynomials. Thus, the KL modes are linear combinations of Zernike polynomials.

The Zernike covariance matrix $\mathbf{C}_Z$ is a real symmetric matrix and hence a Hermitian matrix. A complex matrix $\mathbf{C}$ is said to be Hermitian or self-adjoint [87, 1] if $\mathbf{C} = \mathbf{C}^\dagger$, where $\mathbf{C}^\dagger$ indicates the adjoint matrix of $\mathbf{C}$, i.e., the transpose of the complex conjugate: $\mathbf{C}^\dagger = (\mathbf{C}^*)^T$. Since $\mathbf{C}_Z$ is a real matrix, the adjoint operation is simply the transpose, and the transpose of a symmetric matrix is the same matrix: $\mathbf{C}_Z^\dagger = \mathbf{C}_Z^T = \mathbf{C}_Z$. Therefore $\mathbf{C}_Z$ is a Hermitian matrix.

A Hermitian matrix can be decomposed [87] as: $\mathbf{C} = \mathbf{U}\,\mathbf{\Lambda}\,\mathbf{U}^{-1}$ where $\mathbf{\Lambda}$ is a diagonal matrix, and $\mathbf{U}$ is a unitary matrix. A matrix is unitary if: $\mathbf{U}^\dagger = \mathbf{U}^{-1}$. The elements of the diagonal matrix $\mathbf{\Lambda}$ are the eigenvalues of $\mathbf{C}$ and the columns of the unitary matrix $\mathbf{U}$ are the corresponding eigenvectors of $\mathbf{C}$ that form an orthonormal basis. The elements of matrix $\mathbf{\Lambda}$ are real since the eigenvalues of a Hermitian matrix are always real. In the particular case of a real $\mathbf{C}$ matrix, as is the case of $\mathbf{C}_Z$, the unitary matrix $\mathbf{U}$ is also a real matrix. A real unitary matrix $\mathbf{U}$ is also orthogonal since $\mathbf{U}^T = \mathbf{U}^{-1}$.

The Zernike covariance matrix $\mathbf{C}_Z$ is decomposed in this manner with the help of a diagonal matrix $\mathbf{C}_{KL}$ and a unitary matrix $\mathbf{U}$. The columns of the unitary matrix $\mathbf{U}$ are the eigenvectors of $\mathbf{C}_Z$, i.e., the KL modes. The matrix $\mathbf{C}_{KL}$ is the temporal covariance matrix expressed in the new KL modes, hence a diagonal matrix.

$$\mathbf{C}_Z = \mathbf{U}\,\mathbf{C}_{KL}\,\mathbf{U}^{-1} \tag{A.1}$$

$$\Rightarrow \mathbf{C}_Z = \mathbf{U}\,\mathbf{C}_{KL}\,\mathbf{U}^{T} \tag{A.2}$$

There are several methods available to diagonalize the Zernike covariance matrix and obtain the matrices $\mathbf{C}_{KL}$ and $\mathbf{U}$. IDL provides a routine to diagonalize real symmetric matrices called *EIGENQL*. This routine uses an algorithm based on Householder reductions and the QL method with implicit shifts [53] to obtain the eigenvalues and eigenvectors of the matrix. However, the workload of the QL algorithm [53] used to compute the diagonalizations is $O(n^3)$. Thus, the diagonalization of a large matrix can become a very time consuming computation.

The fact that the Zernike covariance matrix is a sparse matrix, i.e., most of its elements are zero, can be exploited to increase the performance of the diagonalization computation. Only Zernike terms with the same azimuthal order and the same index parity [68] are temporarily correlated and produce non-zero values in the covariance matrix. However, by default Zernike polynomials are ordered following the definition suggested by Noll [52], which produces a covariance matrix with non-zero values distributed over all rows and columns. The order of the Zernike modes can be adjusted so Zernike modes with same azimuthal order are grouped together. Using this new modal order, the Zernike covariance matrix is block diagonal, i.e., the non-zero elements are grouped in clusters around the diagonal. Each individual block is not correlated with the rest of the matrix and can be diagonalized independently.

Due to the $O(n^3)$ workload of the diagonalization algorithm, the computation of several small diagonalizations, one for each small block, is preferable to the computation of just one diagonalization of the whole matrix. Therefore, reordering the Zernike polynomials to produce a block diagonal covariance matrix significantly speeds up the diagonalization process. Figure A.1 schematically demonstrates how the Zernikes polynomials are reordered to produce a block diagonal covariance matrix.

The sets of eigenvalues and eigenvectors for each block are grouped together to produce the eigenvalues and eigenvectors of the whole covariance matrix. Thus, the obtained eigenvalues correspond to the diagonal elements of the matrix $\mathbf{C}_{KL}$ from Equation (A.2),

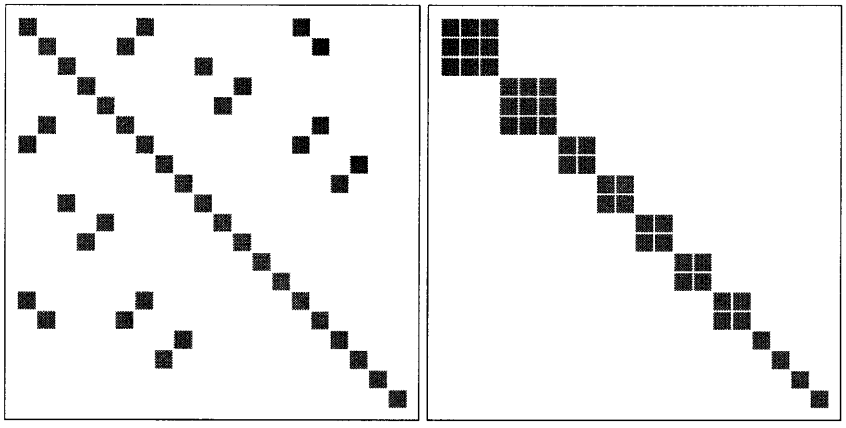

Figure A.1 Nonzero elements of 20×20 Zernike covariance matrix before (left) and after (right) reordering.

Order on left: (2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21)

Order on right: (2 8 16 3 7 17 4 11 5 13 6 12 9 19 10 18 14 15 20 21)

while the eigenvectors constitute the columns of matrix **U**. The eigenvectors are expressed as a linear combination of Zernike polynomials, i.e., the columns of matrix **U** contain the Zernike coefficients that define each eigenvector. The eigenvectors of the Zernike covariance matrix are the KL modes, thus the matrix **U** contains the decomposition of the KL modes in terms of Zernike polynomials.

The particular order in which the KL modes are computed, i.e., the order of the columns of **U**, depends on the reordering applied to the Zernikes to produce the block diagonal matrix. It should also be noted that the rows of matrix **U** are affected by this same reordering process and thus contain Zernike coordinates expressed in the new order. The inverse ordering transformation can be applied to the rows of matrix **U** to restore the Zernike coefficients to the original Noll ordering.

The order in which the computed KL modes are indexed is, in principle, arbitrary. A valid indexing scheme is to order the KLs by decreasing eigenvalue. This produces a set of KL modes with a variance that monotonically decreases with mode order, such as the Noll ordering of Zernike polynomials does. However, some of the eigenmodes have identical eigenvalues, in which case, some other sorting criteria must be applied between them to

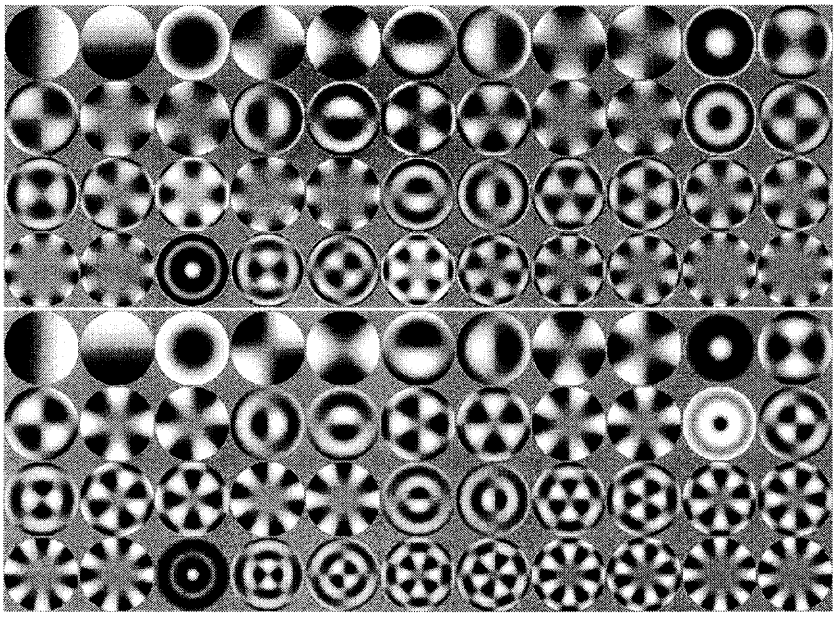

Figure A.2 First 44 Zernike polynomials (top) and KL modes (bottom). The KL modes are ordered as to maintain maximum similarity with Zernike polynomials of the same index. All modes are computed for the unit circle. Mode intensity is scaled independently for each mode to maximize contrast.

break the ambiguity. Another ordering scheme consists of applying the same inverse ordering operation to both the columns and rows of matrix U. This KL ordering does not produce a smoothly decreasing variance with decreasing mode index, as illustrated by the Kolmogorov variance fit shown in Figure 4.9. However, it maintains maximum similarity between KL modes and Zernike polynomials of the same index. Figure A.2 shows the first 44 Zernike and KL modes and illustrates how the KL ordering maintains Zernike similarity, specially for the lower order modes. This second ordering scheme can be used locally to discriminate between KL modes with the same eigenvalue when sorting by decreasing eigenvalue.

In addition to the diagonal temporal covariance matrix th KL modes produce, KL modes present another advantage over Zernike polynomials for AO systems with Shack-Hartmann wavefront sensors. Zernike polynomials of increasing order show a tendency to distribute most of their structure on the outside rim of the unity radius circle, whereas the

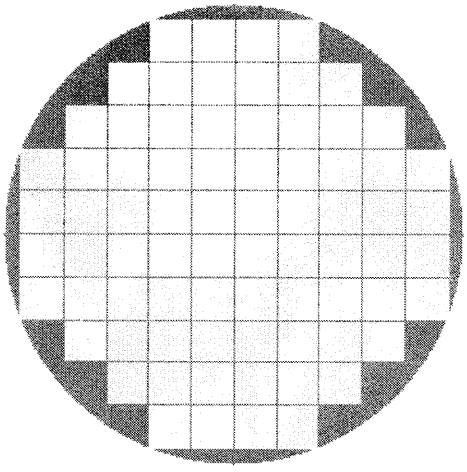

Figure A.3 Shack-Hartmann wavefront sensor 10x10 pupil geometry.

KL modes are distributed closer to the pupil center. This can be clearly observed on the higher order Zernike polynomials shown in Figure A.2. A Shack-Hartmann wavefront sensor divides the pupil space into a regularly spaced grid of square subapertures (Figure A.3). The mean $x-$ and $y-$ slopes of the wavefront are then measured at each subaperture. Therefore, having most of the mode structure lying on just a few of the subapertures close to the edges of the pupil tends to stress the measuring capabilities of the Shack-Hartmann wavefront sensor. Furthermore, the edges of the pupil is the location where the distribution of wavefront sensor subapertures is more sparse.

Figures A.4 and A.5 illustrate how the KL and Zernike modes fit on to the Shack-Hartmann array. Both Zernike and KL modes are of equivalent order (mode number 45) but the structure distribution on the pupil is quite different. The Zernike mode is mostly flat everywhere but close to the edge of the pupil, where most of its structure is concentrated. The KL mode presents a more extended structure that reaches farther into the center of the pupil. When each of these modes is sensed by the Shack-Hartmann wavefront sensor, the subapertures capture more information for KL modes.

This can be further explored by calculating the reconstruction matrix given by this geometry using Zernike polynomials and KL functions. The reconstruction matrix is calcu-

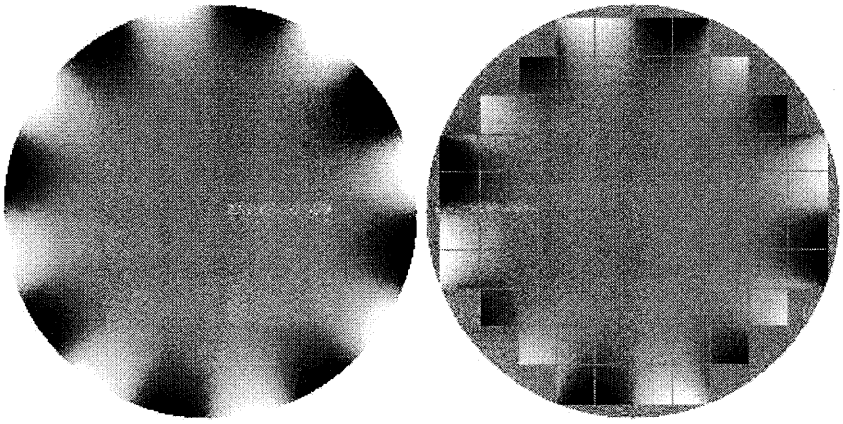

Figure A.4 Zernike mode 45 in the unit circle and how it is seen by WFS geometry. Mode intensity is scaled independently for each mode to maximize contrast.

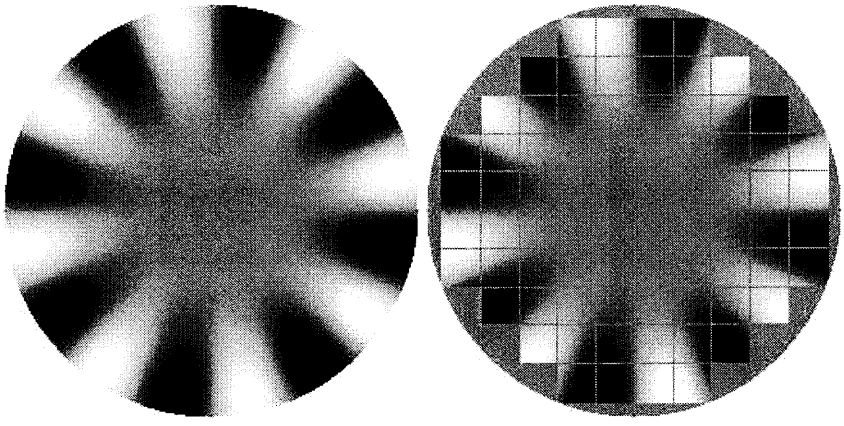

Figure A.5 Karhunen Loève mode 45 in the unit circle and how it is seen by WFS geometry. Mode intensity is scaled independently for each mode to maximize contrast.

lated by computing the pseudo-inverse of the interaction matrix, as shown in Appendix C. The eigenvalues of the interaction matrix contain the sensitivity of the WFS to a particular eigenmode of the interaction matrix. The inversion of a very small eigenvalue, which corresponds to an eigenmode with low sensitivity, produces a very large eigenvalue of the reconstruction matrix and may dominate its response. If the response of the reconstruction matrix is dominated by a mode that cannot be properly detected by the WFS, the reconstruction matrix will be mostly amplifying noise [67] and result in an unstable servo loop.

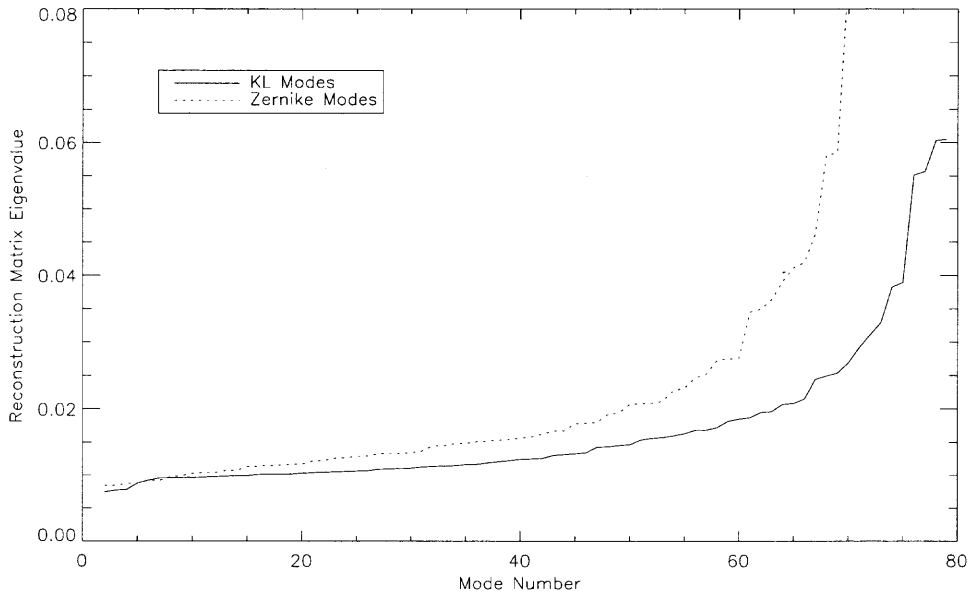

Figure A.6 Reconstruction matrix eigenvalues using Zernike polynomials and Karhunen-Loève functions.

Thus, very small eigenvalues in the interaction can reduce the signal to noise ratio of the reconstruction matrix.

Figure A.6 shows a plot of the reconstruction matrix eigenvalues obtained using KL modes and Zernike polynomials. The eigenvalues obtained using Zernike polynomials diverge much quicker than the ones obtained with KL modes. This implies that more KL modes than Zernike modes can be reliably reconstructed. Also, the eigenvalues of the Zernike reconstruction matrix are consistently higher than the KL ones, suggesting that the KL reconstruction matrix will produce a wavefront reconstruction with a higher signal to noise ratio. In conclusion, a Shack-Hartmann wavefront sensor is more sensitive to KL modes than to Zernike polynomials.

APPENDIX B

CALIBRATION OF ADAPTIVE OPTICS SYSTEM

This section discuses the different strategies attempted to obtain an accurate and complete characterization of the different components of the AO system. The PSF estimation method uses the AO loop data produced by the AO system. This includes WFS measurements that give an estimation of the residual errors after correction, and DM commands that provide information about the seeing conditions. It is critical for the accuracy of the PSF method that these measurements are correctly quantified.

These measurements are recorded by the AO system as raw data. The recorded WFS measurements are expressed as pixel shifts measured from the WFS camera. The DM commands are expressed as counts that can be directly sent to the DM controller. These raw system measurements must be converted to real wavefront measurements in real physical units, such as radians of wave. The transformation is performed with calibration factors. In the case of the WFS, this calibration factor relates wavefront tilt in each subaperture of the WFS to pixel shift in the WFS camera. The DM factor relates the counts sent to the DM controller to actual push of each DM actuator. These calibration factors are vital for the correct interpretation of the AO loop data and must be accurately estimated.

B.1 Theoretical Calibration

The calibration factors can be estimated from knowledge of the physical characteristics of the AO system components, such as the plate scale of the WFS, which are in principle known and should provide good estimates of the required factors.

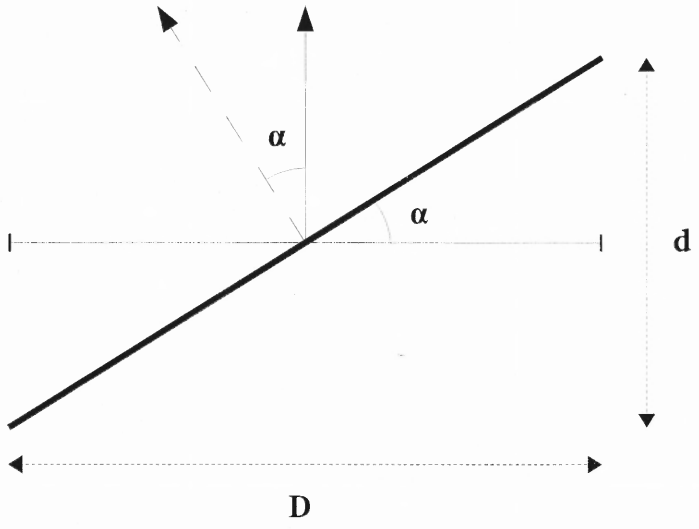

Figure B.1 Schematic representation of a flat wavefront tilted by an angle α in the pupil of a telescope of aperture D.

B.1.1 Wavefront Sensor Calibration

The WFS produces an array of subaperture images on its camera. The angular resolution of the WFS relates pixel distance in the produced images to angular field separation in the real object being imaged. This angular resolution of the WFS, also referred to as its pixel scale, provides enough information to relate measured pixel shifts to wavefront tilt. The pixel scale of the WFS in the AO system has been previously measured: $P_s = 1$ arcsec/pixel.

Consider a flat tilted wavefront that enters the pupil of the telescope. The wavefront is tilted by an angle α with respect to the optical axis in the pupil of diameter D, as shown in Figure B.1. The distance d in the figure is expressed as:

$$\alpha \simeq \tan\alpha = \frac{d}{D} \Rightarrow d = D\,\alpha \tag{B.1}$$

where the following approximation for small angles was applied: $\tan\alpha \simeq \alpha$ since common wavefront tilts are on the order of seconds of arc, i.e., $\alpha \sim 10^{-6}$ rad. The light propagates from the pupil of the telescope down to the AO system. Inside the AO system, a small

fraction of the light is separated by a beam splitter and used for wavefront sensing, while the rest continues to the science path. A wide filter is used on the WFS light to only allow light of wavelength around $\lambda_{\text{wfs}} = 550$ nm into the WFS. Hence, the wavelength of the light reaching the WFS can be considered λ_{wfs}. This allows the conversion of d from distance units to wave units (radians of wave).

$$d[rad] = \frac{2\pi}{\lambda_{\text{wfs}}} D\,\alpha \tag{B.2}$$

The tilted wavefront is imaged by the lenslet array and rescaled by the imaging optics of the WFS, producing the subaperture images on the WFS camera. The optical layout of the WFS is designed to produce a known pixel scale on the camera P_s. Thus, features recorded one pixel apart in the WFS camera are P_s radians apart on the sky, i.e., a pixel shift S_{px} measured by a subaperture of the WFS corresponds to a wavefront tilt of $\alpha = P_s\,S_{px}$ radians on that subaperture. A global tilt produces the same wavefront tilt in all the subapertures, and all measure a pixel shift of S_{ps}. Substituting this in Equation B.2 produces a relation between measured pixel shift in the WFS and actual wavefront tilt on the pupil:

$$d[rad] = \frac{2\pi}{\lambda_{\text{wfs}}} D\,P_s\,S_{px} \tag{B.3}$$

The PSF estimation algorithm uses modal decomposition to express wavefront. This measured tilt must be scaled into modal coefficients that reproduce the correct measured tilt. The simplest example is to use Zernike polynomials as the modal basis. The first Zernike polynomial (of order n=2) corresponds directly to tilt, shown in Figure B.3. However, as illustrated by the figure, the Zernike polynomial of order n=2 ranges from a minimum value of -2 to a maximum value of 2. The distance d produced by a wavefront described by a

Zernike coefficient $z_2 = 1$ would be $d = 4$. There is a factor 4 between d in radians and the Zernike coefficient z_2

$$z_2 = \frac{1}{4}\frac{2\pi}{\lambda_{\text{wfs}}} D\, P_s\, S_{px} \tag{B.4}$$

The previous equation constitutes a relationship between the pixel measurements from the WFS and the Zernike coefficients that describe the wavefront on the pupil. This represents the actual response of the WFS system. However, the wavefront is reconstructed from the WFS measurements with a reconstruction matrix $\mathbf{D}^+$, i.e., the pseudo-inverse of the interaction matrix. The interaction matrix is computed by simulating the WFS response to synthetic wavefront inputs, as described in Appendix C. Thus, a synthetic wavefront described by a Zernike coefficient $\hat{z}_2 = 1$, which has a slope of 2 on the x axis, produces a WFS simulated x shift response of 2 pixels in all the subapertures. When the computed $\mathbf{D}^+$ encounters such a WFS measurement, it reconstructs the wavefront as described by $\hat{z}_2 = 1$ instead of what is indicated by Equation (B.4):

$$z_2(2\,px) = \frac{1}{4}\frac{2\pi}{\lambda_{\text{wfs}}} D\, P_s\, 2 \tag{B.5}$$

Therefore, the Zernike coefficients reconstructed by $\mathbf{D}^+$ must be scaled by a calibration factor k_{wfs} that would make $\hat{z}_2 = z_2$. This factor is the WFS calibration factor and it relates Zernike coefficients restored by $\mathbf{D}^+$ to Zernike coefficients that describe the actual wavefront:

$$k_{\text{wfs}} = \frac{1}{2}\frac{2\pi}{\lambda_{\text{wfs}}} D\, P_s \tag{B.6}$$

B.1.2 Deformable Mirror Calibration

In the case of the DM, the DM controller applies a voltage to each actuator according to the commands received from the AO system. Each actuator then pushes the thin membrane of the DM proportionally to the voltage applied. The values that the AO system sends to the DM controller are expressed in units called *counts*, which are proportional to the voltage applied. The DM calibration factor relates the commands in counts sent to the DM to actual push of the actuators.

The technical specifications of the DM state that the maximum stroke of each actuator is 4.8 microns. It also indicates that a value of 500 counts sent to an actuator produces a push 1/8 of the total range. Thus, a value of 1000 counts produces a linear push of 1.2 micrometers. The ratio between counts and distance pushed by the actuator is: $1000/1.2 = 833.333\text{counts/micron}$. The distance an actuator pushes when instructed to push C counts is written as:

$$d[\mu\text{m}] = \frac{1.2}{1000} C \tag{B.7}$$

The PSF estimation algorithm requires an estimation of the shape of the DM to obtain a fit of the Fried parameter r_0. Since the Fried parameter is defined at a wavelength of 550 nm, the distance d an actuator pushes can be written as radians of wave.

$$d[\text{rad}] = \frac{2\pi}{0.550\mu m} \frac{1.2}{1000} C \tag{B.8}$$

The DM applies its correction to the wavefront by reflection, therefore an actuator pushing a certain distance will introduce twice that change into the corrected wavefront. A certain value of C counts sent to a particular actuator will introduce a change of d' in the wavefront:

$$d'[\mathrm{rad}] = 2\frac{2\pi}{550\mu m}\frac{1.2}{1000}C \tag{B.9}$$

The previous equation relates counts sent to the DM controller to actual wavefront change produced. The factor that relates these two quantities is the DM calibration factor:

$$k_{\mathrm{dm}} = 2\frac{2\pi}{550\mu m}\frac{1.2}{1000} \tag{B.10}$$

The wavefront correction introduced by the DM is then reconstructed by using the DM calibration factor k_{dm} and the influence functions. This wavefront is then decomposed in terms of a modal basis and fitted to the Kolmogorov model to obtain a value of the Fried parameter r_0.

B.2 Empirical Calibration

These calibration factors can also be measured from the AO system itself by measuring the response of the WFS and the DM to a known wavefront input. The simplest wavefront input to consider is a flat tilted wavefront. The WFS and DM response to this wavefront relate internal raw AO system measurements (in pixel shift and count units) to external actual wavefront measurements.

A pinhole was placed at the main focus of the telescope to provide a high contrast image with no aberrations for the AO system to lock on. The tilted wavefront was introduced by shifting this pinhole in the focal plane. The amount of tilt introduced can be easily measured by taking images of the pinhole with a camera of known pixel scale in prime focus.

Initially, the pinhole is centered on the field-of-view. An image is captured of this "zero" position with the camera as a reference. Also, a reference subaperture is captured on the AO system, so the WFS sensor is measuring zero shifts with the pinhole in this position.

So far the AO system is working in "open-loop" mode, i.e., the AO system is working on a passive mode where the wavefront is being continuously measured by the WFS but the DM is not providing any correction. The DM in open-loop behaves as a flat mirror.

The pinhole is shifted in the focus plane. An image of the shifted pinhole is taken with the camera. The difference between the images captured with centered and shifted pinhole will provide a measurement of the wavefront tilt introduced. The WFS measures the wavefront coming from the shifted pinhole as a tilted wavefront with respect to the centered pinhole. The WFS measurements produced are stored. These measurements, which relate a known wavefront tilt to WFS pixel shifts, will be used to obtain the WFS calibration factor.

B.2.1 Empirical Calibration: Wavefront Sensor

The DM calibration factor is obtained by letting the DM correct the wavefront tilt introduced by the shifted pinhole. The AO loop is closed and the shape of the DM is adapted to correct for the tilt that the WFS was measuring. When correction is complete, the shape of the DM emulates the measured tilt and corrects for it, which produces zero tilt measurements in the WFS. The DM commands that produce the tilt shape are stored and will be used to estimate the DM calibration factor.

The amount of tilt introduced by shifting the pinhole is extracted from the images taken with the science camera. Knowing the pixel scale in prime focus that the optical setup produces on the science camera, i.e., how many arcseconds on the sky correspond to each pixel on the camera, the tilt introduced can be easily measured.

The pixel shift between the centered and shifted pinhole images can be estimated by two different methods. The first method consists of computing the cross-correlation of the centered and shifted pinhole images. The position of the cross-correlation peak provides a measurement of the pixel shift between pinhole images. An alternative approach consists

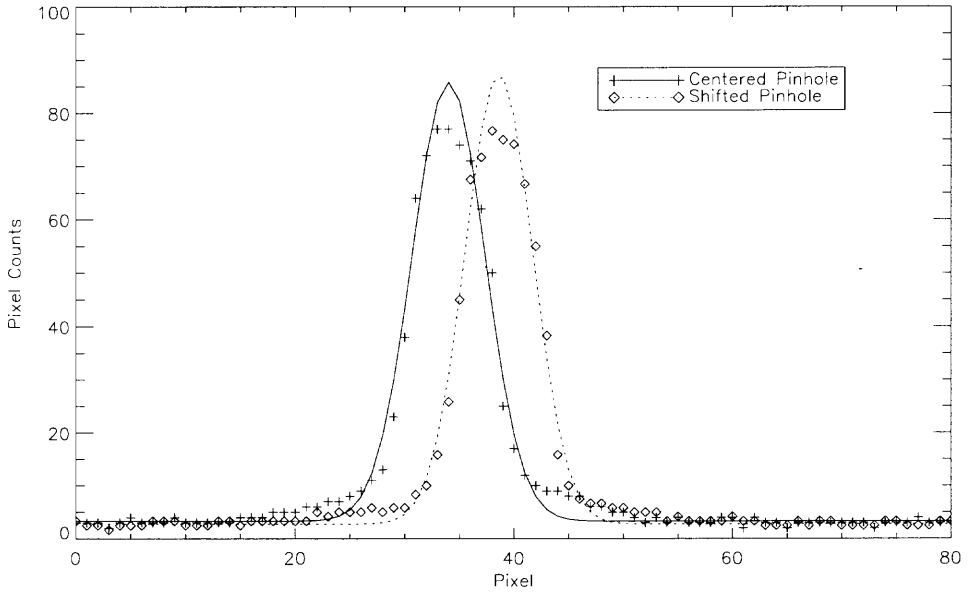

Figure B.2 Cross-section of pinhole images. Data points are plotted as symbols. Lines represent the Gaussian fits.

of fitting a gaussian function to each pinhole image and obtaining their positions from the fit.

The width of the cross-correlation peaks was considerably large, introducing some uncertainty in the shift estimates. This is believed to be caused by small shape differences between both pinhole images. The gaussian fits seemed to provide better more consistent position measurements.

Figure B.2 shows a cross-section of both pinhole images, showing the actual data points and the gaussian fits. The pinhole image pixel shift S_{px} is calculated by subtracting the positions of the gaussian fits. Since the camera pixel scale of the setup is known $P_s = 0.1875$ arcsec/pixel, the actual tilt introduced to the AO system is obtained by:

$$\alpha = P_s \, S_{px} \qquad [\text{arcsec}] \tag{B.11}$$

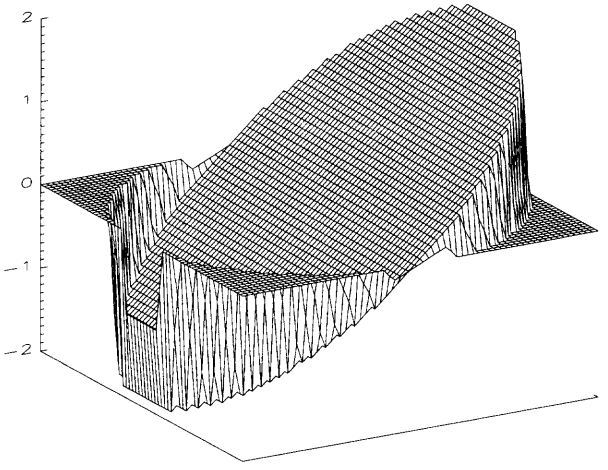

Figure B.3 Zernike polynomial Z_2, i.e., tilt in the x direction.

Knowing the angle α of the wavefront tilt and the wavelength of the light used for wavefront sensing, the distance d from Figure B.1 can be computed in radians of wave with Equation (B.2).

The PSF estimation algorithm follows a modal decomposition of the measured wavefront. The distance d that describes the wavefront tilt must be expressed as a coefficient to the tilt Zernike, i.e., Zernike polynomial of order n=2 (shown in Figure B.3). The wavefront can be expressed as: $\phi = z_2 Z_2$. As previously, a Zernike coefficient $z_2 = 1$ would produce a d of 4 radians.

For a known d, the correct value of z_2 is $z_2 = \frac{d}{4}$. The value of d was measured from the centered and shifted pinhole images, hence the value of the Zernike coefficient z_2 is written as:

$$z_2 = \frac{d}{4} = \frac{\pi}{2\lambda_{\text{wfs}}} D \tan(\alpha) \tag{B.12}$$

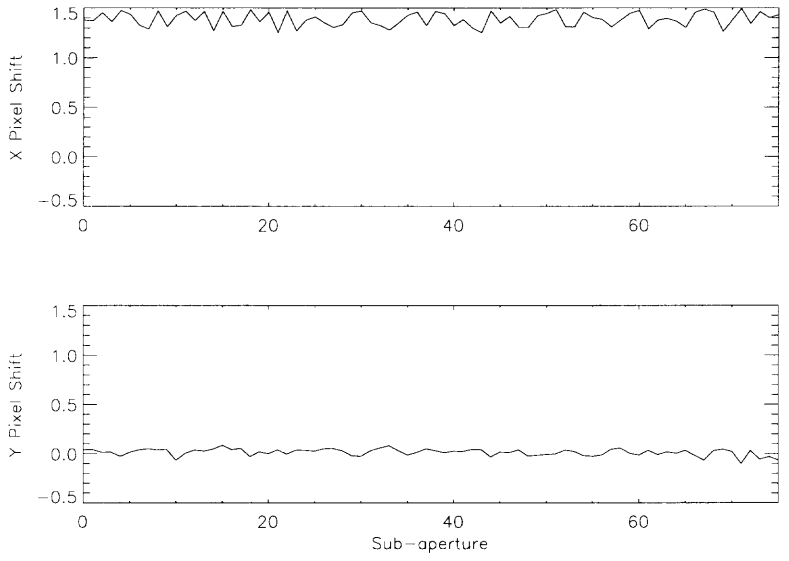

Figure B.4 WFS pixel shifts produced by tilt in the x direction.

The modal coefficient decomposition of the wavefront tilt has been estimated from the science camera images. Now it must be compared to the coefficient decomposition obtained from the WFS measurements. The WFS pixel shift measurements, shown in Figure B.4, are expressed as modal coefficients with the help of the reconstruction matrix D^+ (see Appendix C).

Thus, the Zernike coefficient decomposition $\vec{z}$ that describes the wavefront tilt is obtained from the WFS measurements $\vec{w}$ and the reconstruction matrix: $\vec{z}' = \mathrm{D}^+\vec{w}$. This new z_2' coefficient is compared to the one obtained from the shifted pinhole images z_2 to calculate the WFS calibration factor. The coefficient z_2 is already correctly scaled to describe the wavefront tilt, thus the WFS calibration factor must be applied to the z_2' coefficient obtained from the WFS measurements:

$$z_2 = k_{\mathrm{wfs}} z_2' \implies k_{\mathrm{wfs}} = \frac{z_2}{z_2'} \tag{B.13}$$

The WFS calibration can be incorporated into the reconstruction matrix $\mathbf{D}^+$ that was used to calculate z_2'. This reduces the number of multiplications when the calibration factor must be applied to a large number of WFS measurements sets.

B.2.2 Empirical Calibration: Deformable Mirror

The DM calibration factor is computed following a similar procedure. The AO loop was closed and the introduced wavefront tilt was corrected by the DM. Normally, any tip or tilt measured in the wavefront would be corrected by the tip-tilt mirror before it reaches the DM. During the calibration procedure, the tip-tilt correction was disabled to allow the DM to correct for any tilt.

The stored actuator commands are in units called *counts*, which are proportional to the voltage applied to the actuators. The actuator commands in tilt will be related to the previously measured wavefront tilt, see Equation (B.11).

The influence function of each actuator describes the shape that the actuator produces when pushed. The shape of the DM is decomposed as a linear combination of influence functions with the actuators commands as the coefficients of the decomposition. This assumes that the influence functions constitute a set of orthonormal functions that describe the space of DM shapes. In reality, there is some cross-talk between adjacent actuators, i.e., the shape resulting from pushing two adjacent actuators is not exactly the same as the shape obtained by adding the contribution of each actuator independently. However, while a linear decomposition is not exactly accurate, it provides a good approximation [97] to the actual shape of the DM.

The influence functions were directly measured from the AO system by poking each actuator individually and measuring the resulting DM shape with an interferometer. Results from this measurement are shown in Figure 4.7. To eliminate the noise from the influence functions, a synthetic set was fitted to the measured influence functions. Two different types of synthetic influence functions were tested: Gaussian functions and a combination

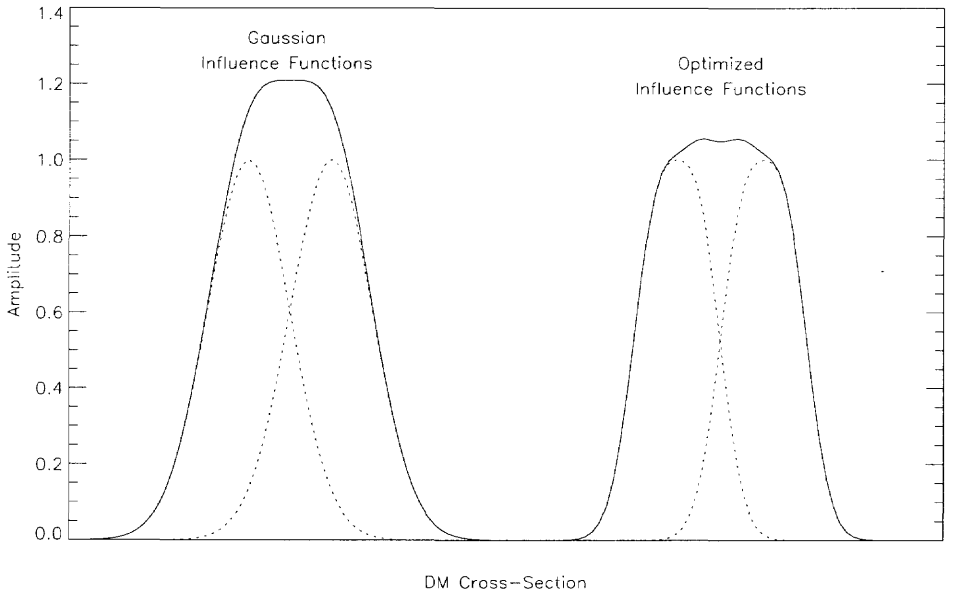

Figure B.5 Cross-section of shape produced by two adjacent actuators pushing. Using Gaussian and combination of 4 Gaussians influence functions.

of 4 smaller Gaussians. The second type of influence function is constructed by adding 4 small Gaussians located on the corners of a square. The resulting shape is less pointy than a single Gaussian and presents a slight square shape that is a better match to the actual shape of the measured influence functions.

Another advantage of the second type of influence functions is the lower cross-talk produced by simulating the poking of two adjacent actuators, as illustrated by Figure B.5. Hence, the influence functions obtained by combining four small Gaussians were chosen to generate the synthetic set of influence functions. Their width and position were fitted from the measured influence functions.

The shape of the DM ϕ_{dm} is reconstructed from the actuator commands and the synthetic influence functions. Figure B.6 shows the reconstructed shape of the DM while correcting for the wavefront tilt, computed from the actuator commands stored. This shape is decomposed as a linear combination of Zernike polynomials:

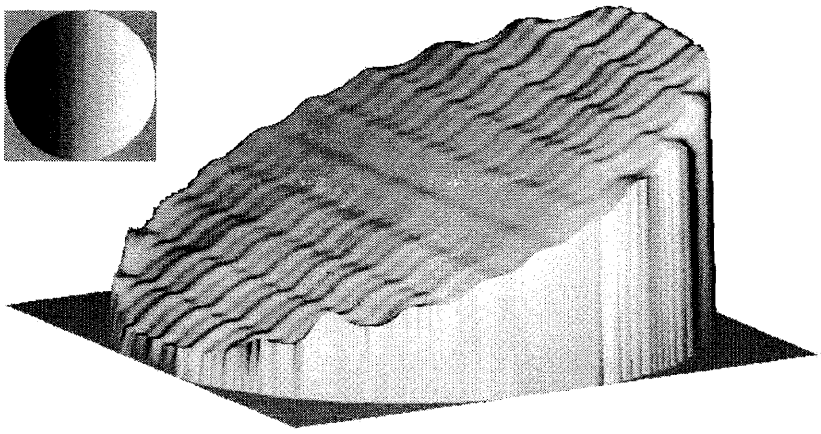

Figure B.6 Reconstruction of the shape of the DM from stored mirror commands.

$$\phi_{\mathrm{dm}}(\vec{x}) = \sum_j z'_j Z_j(\vec{x}) \tag{B.14}$$

The process of transforming the mirror commands into Zernike coefficients is just a change of basis from influence functions to Zernike polynomials. The change can be directly performed if the projection matrix of the change of basis is known. The calculation of the projection matrix B_Z is described in Appendix C. The Zernike coefficients $\vec{z}'$ are computed from the actuator commands m_i with the help of the projection matrix: $\vec{z}' = \mathrm{B}_Z \vec{m}$. Figure B.7 shows an example of Zernike coefficients obtained in this way. The Zernike coefficients from the figure were computed from the actuator commands stored during the calibration process. All of the coefficients are zero or close to zero except for z'_2, clearly indicating that the wavefront they describe is mostly composed of tilt.

The Zernike coefficient z'_2 obtained from the decomposition is compared to the one obtained from the pinhole measurements z_2 to obtain the DM calibration factor. As before, z_2 is already scaled correctly so the DM calibration factor k_{dm} must be applied to the coefficient z'_2 obtained from the actuator commands.

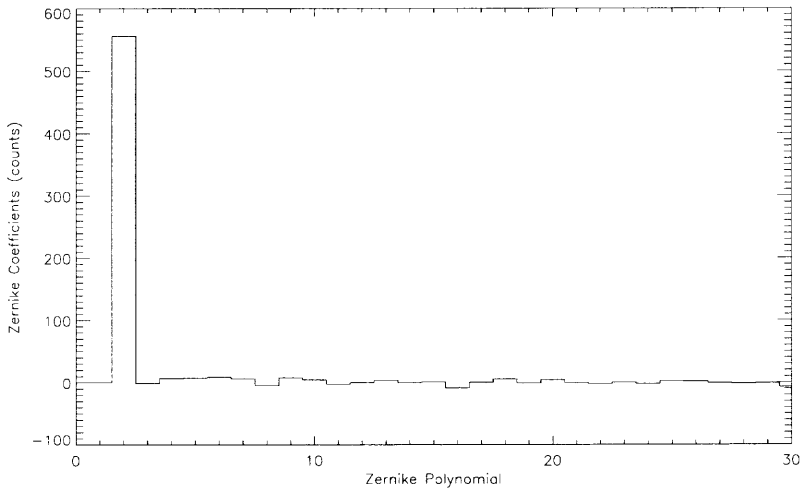

Figure B.7 Zernike coefficients that describe the shape of the DM correcting the wavefront tilt.

$$z_2 = k_{\mathrm{dm}}\, z_2' \tag{B.15}$$

$$k_{\mathrm{dm}} = \frac{z_2}{z_2'} \tag{B.16}$$

As in the case of the WFS calibration factor, the DM calibration factor can be incorporated inside the projection matrix to reduce the number of multiplications required when processing a large number of data sets.

APPENDIX C

ADAPTIVE OPTICS CONTROL

This section summarizes the basic concepts and requirements of a solar AO control system [33, 67, 61, 97] and how they relate to the PSF estimation method.

In an AO system, the residual wavefront after correction by the DM is measured by the Shack-Hartmann WFS. At each AO iteration, the shape of the DM is updated in such a way as to minimize the measured wavefront residuals. In this way, the AO system is a closed-loop control system. The system computing the updated DM shape from the measured wavefront measurements at each AO iteration is the *AO Control System*, as shown in Figure 2.1.

The AO control system processes the measurements from the WFS and computes the new shape of the DM. A Shack-Hartmann WFS measures the average x and y slope of the wavefront at different points in the pupil. The slopes are measured by the WFS as a series of x and y shifts between subaperture images formed by lenslets onto the WFS camera. The DM consists of a thin mirror plate in front of a series of actuators that can push or pull on it. With the measured x and y shifts, the AO control system computes the optimum shape of the DM in the least-square sense as a series of actuator commands, effectively closing the loop between wavefront sensing and correction.

The control process is formally a matrix multiplication [97, 33] between the *Control Matrix* C^{+} of the system and the WFS measurements. The WFS measurements, i.e., the x and y shifts from each subaperture, are arranged in vector form: $\vec{w}$. The DM commands are also arranged in vector form: $\vec{m}$. The vector $\vec{m}$ is then obtained by multiplying the control matrix by the WFS measurements vector $\vec{w}$: $\vec{m} = \mathrm{C}^{+}\,\vec{w}$. Equation (C.1) shows this in matrix form.

$$\begin{pmatrix} m_1 \\ m_2 \\ \vdots \\ m_A \end{pmatrix} = \underbrace{\begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1S} \\ c_{21} & c_{22} & \cdots & c_{2S} \\ \vdots & \vdots & \ddots & \vdots \\ c_{A1} & c_{A2} & \cdots & c_{AS} \end{pmatrix}}_{C^{+}(A\times S)} \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_s \end{pmatrix} \tag{C.1}$$

The WFS has $S/2$ subapertures that measure wavefront tilt in the x and y directions, hence the WFS vector $\vec{w}$ contains S elements. The DM has a number A of actuators, hence the DM commands vector $\vec{m}$ contains A elements. The control matrix C^{+} is a $A \times S$ matrix, generally $S > A$ for AO systems. It should be noted that the WFS shifts are arranged in the vector $\vec{w}$ as follows: $\vec{w} = \{S_{1x}, S_{1y}, S_{2x}, S_{2y}, \ldots\}$, where S_{1x} refers to the x shift from subaperture 1. In addition, the control matrix is constrained to set the piston mode to zero since it cannot be measured by the Shack-Hartmann WFS.

C.1 Theoretical Calculation

The control matrix of the AO system can be theoretically computed by simulating how the AO system works. For this, the characteristics of both the WFS and the DM must be known.

The Shack-Hartmann WFS consists of an array of lenslets placed in a pupil image. Each lenslet forms an image from one section of the wavefront on a section of the WFS camera. The relative shifts of each one of these images with respect to a reference one are proportional to the mean wavefront slope at the lenslet position on the pupil [33, 97]. This process can be numerically simulated since the position and size of each lenslet in the pupil is known. Given a wavefront, the $x-$ and $y-$slope of the wavefront at each point in the pupil can be calculated. These slopes are then averaged inside each lenslet producing the final $x-$ and $y-$ WFS measurement for each subaperture. Since the pixel scale of the

WFS camera is known, i.e., how much tilt corresponds to a shift of one pixel, the calculated mean slopes can be related to pixel shifts in the camera.

The Xinetics DM consists of a thin mirror plate placed in front of an array of actuators. Each actuator can push or pull on the mirror to produce the desired shapes. The shape that an actuator will produce when pushed is called its influence function, which can be measured with an interferometer. Therefore, a mirror shape can be decomposed in terms of these individual influence functions. In other words, the amount of pushing for each actuator necessary to generate a certain mirror shape can be numerically computed.

The response of a WFS to a given wavefront can be simulated and used to compute the required actuator commands that reproduce that wavefront on the DM. However, these simulation calculations rely on some assumptions about the actual AO system: it is assumed that the WFS and DM configurations with respect to the pupil are exactly known. Any small misalignment between the actual system and the model will affect the obtained results.

As discussed in Appendix A, the wavefront can be decomposed as a linear combination of some modal basis. Karhunen-Loève (KL) functions are chosen as the modal basis in this case. Thus, a wavefront φ is described as a linear combination of KL functions K_i:

$$\varphi = \sum_{i}^{K} k_i K_i \tag{C.2}$$

where K_i are the KL modes and k_i are the KL coefficients that describe wavefront φ, i.e., the coordinates of the wavefront when expressed in terms of KL modes.

Since the wavefront can be expressed as a linear combination of KL modes, it is enough to know the response of the WFS to each KL mode in order to obtain its response to any arbitrary wavefront. Each individual KL mode K_i is applied to the WFS model and the produced pixel shifts are recorded. The shifts produced by KL mode K_i are placed in a vector named $\vec{w}_i$. Thus, assuming a number K of KL modes, the following list of measurements vectors is obtained: $\{\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_K\}$.

The properties of matrix multiplication can be used to simplify the computations. Each vector $\vec{w}_i$ is placed as the column of a matrix, as shown in Equation (C.3).

$$\mathbf{D} = \begin{pmatrix} w_{11} & w_{12} & \dots & w_{1K} \\ w_{21} & w_{22} & \dots & w_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ w_{S1} & w_{S2} & \dots & w_{SK} \end{pmatrix} \tag{C.3}$$

This matrix is called the *Interaction Matrix* $\mathbf{D}$ and its dimensions are $S \times K$, where S is twice the number of subapertures in the WFS and K is the number of KL modes used in the wavefront decomposition. This matrix $\mathbf{D}$ relates KL coefficients to WFS shifts. The shifts measured by the WFS $\vec{w}$ are computed by the matrix multiplication of $\mathbf{D}$ with the KL coefficients $\vec{k}$, described in Equation (C.2) that describe the wavefront:

$$\vec{w} = \mathbf{D}\,\vec{k} \tag{C.4}$$

A simple example can help to clarify the previous statement. Assume a wavefront consisting only of the 2nd KL mode. In this case, the KL coefficient decomposition of the wavefront is: $\vec{k}_{ex} = (0, 1, \dots, 0)$. The interaction matrix $\mathbf{D}$ is then multiplied by the KL coefficient vector $\vec{k}_{ex}$. The result is just the 2nd column of the interaction matrix: $\vec{w}_2$, i.e., the shifts vector that the 2nd KL mode produced on the WFS: $\vec{w}_2 = \mathbf{D}\,\vec{k}_{ex}$. This is more clearly shown in matrix form in Equation (C.5).

$$\begin{pmatrix} w_{12} \\ w_{22} \\ \vdots \\ w_{S2} \end{pmatrix} = \begin{pmatrix} w_{11} & w_{12} & \dots & w_{1K} \\ w_{21} & w_{22} & \dots & w_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ w_{S1} & w_{S2} & \dots & w_{SK} \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} \tag{C.5}$$

A matrix $\mathbf{D}$ relating KL coefficients to WFS measured shifts has been obtained. However, what is needed is a matrix that relates measured WFS shifts to KL coefficients so that the wavefront can be estimated from the WFS measurements. Such a matrix can be obtained by multiplying Equation (C.4) on the left by the inverse matrix of the interaction matrix. However, the interaction matrix $\mathbf{D}$ normally constitutes an over-determined system of linear equations, and thus, not a square matrix [67, 87]. The dimensions of $\mathbf{D}$ are $S \times K$, where the number of shift measurements is larger than the number of KL modes considered ($S > K$). Since $\mathbf{D}$ in general is not a square matrix, it cannot be inverted. A pseudo-inverse matrix $\mathbf{D}^+$ that follows $\mathbf{D}^+ \mathbf{D} = \mathcal{I}$, where $\mathcal{I}$ is the identity matrix, must be computed instead:

$$\mathbf{D}^+ \vec{w} = \mathbf{D}^+ \mathbf{D} \vec{k} \tag{C.6}$$

$$\vec{k} = \mathbf{D}^+ \vec{w} \tag{C.7}$$

This pseudo-inverse matrix $\mathbf{D}^+$ is called *Reconstruction Matrix*, and its dimensions are $K \times S$. It relates measured WFS shifts to KL coefficients that describe the wavefront. The KL coefficients are calculated from the measured WFS shifts through a matrix multiplication: $\vec{k} = \mathbf{D}^+ \vec{w}$.

The reconstruction matrix $\mathbf{D}^+$, i.e., the pseudo-inverse of $\mathbf{D}$, can be computed using the method of *singular value decomposition* (SVD). The SVD method [53, 87] decomposes a matrix as a product of orthonormal and diagonal matrices:

$$\mathbf{D} = \mathbf{U}\,\mathbf{W}\,\mathbf{V}^T \tag{C.8}$$

Matrix $\mathbf{U}$ is a column-orthonormal $S \times K$ matrix, $\mathbf{V}^T$ is the transpose of an orthonormal $K \times K$ matrix and $\mathbf{W}$ is a diagonal $K \times K$ matrix. The diagonal elements of $\mathbf{W}$ are called the *singular values* of matrix $\mathbf{D}$. An orthonormal matrix [87] is a real matrix that

satisfies: $\mathbf{U}\mathbf{U}^T = \mathcal{I}$, where $\mathcal{I}$ is the identity matrix. Thus, the transpose of an orthogonal matrix is equal to its inverse: $\mathbf{U}^T = \mathbf{U}^{-1}$.

Once $\mathbf{D}$ is decomposed as indicated by Equation (C.8), its inversion is straightforward:

$$\mathbf{D}^+ = \mathbf{V}\,\mathbf{W}^{-1}\,\mathbf{U}^T \tag{C.9}$$

The inverse matrix of a diagonal matrix $\mathbf{W}$ is another diagonal matrix $\mathbf{W}^{-1}$, such that: $w_{ii}^{-1} = 1/w_{ii}$, where w_{ii} and w_{ii}^{-1} are the diagonal elements of $\mathbf{W}$ and of $\mathbf{W}^{-1}$, respectively. In addition, the inverse of the orthonormal matrices are their transpose matrices.

A very small or null diagonal element of $\mathbf{W}$ indicates that the system represented by $\mathbf{D}$ is not sensitive to a particular configuration. When this very small singular value is inverted, it produces a very large singular value in the pseudo-inverse matrix. That is, the inversion process is giving a very large weight to some configuration to which the system is not sensitive, hence mostly amplifying noise [67]. This can make the system represented by the pseudo-inverse matrix unstable. It is preferable to set the very large inverted singular value to zero.

As shown in Equation (C.7), the KL decomposition of the wavefront $\vec{k}$ is obtained by a matrix multiplication of the measured WFS shifts with the reconstruction matrix $\mathbf{D}^+$. The measured wavefront is described in terms of KL modes as shown in Equation (C.10).

The DM attempts to correct the measured wavefront φ by adapting its shape to replicate it. The shape of the DM is described by the DM commands and the DM influence functions, as shown in Equation (C.11). Thus, the measured wavefront φ can be expressed equally in terms of KL modes K_i and in terms of influence functions I_i:

$$\varphi = \sum_{i}^{K} k_i K_i \tag{C.10}$$

$$\varphi = \sum_{i}^{A} m_i I_i \tag{C.11}$$

The DM influence functions I_i are the characteristic shape that each actuator produces on the DM when pushed. The coefficients m_i describe the wavefront expressed in terms of the influence functions. However, these m_i coefficients are not the actual DM commands stored by the AO system. The actuator commands produced by the system are expressed in some particular units called *counts*, which are proportional to the voltage applied to the DM actuators. The transformation between the m_i coefficients and the DM actuator commands in counts is performed with a proportionality constant, as described in Section 4.4.2.

The dependence on the KL modes K_i must be eliminated from Equation (C.10) in order to translate the measured wavefront into influence function coefficients m_i. For this, each KL mode is projected into the new basis of influence functions with the help of a projection matrix B. Since both basis K_i and I_i are known, the elements of the projection matrix b_{ij} can be easily derived, as will be discussed later in this section. Thus, the projection of the KL modes on the influence functions is written as:

$$K_i = \sum_{j}^{A} b_{ij} I_j \tag{C.12}$$

Equation (C.12) is substituted in Equation (C.10) to produce:

$$\varphi = \sum_{i}^{K} k_i \sum_{j}^{A} b_{ij} I_j \tag{C.13}$$

Equations (C.11) and (C.13), which provide two different decompositions of the measured wavefront in terms of influence functions, can be equated producing the following relations:

$$\sum_{j}^{A} m_j I_j = \sum_{i}^{K} k_i \sum_{j}^{A} b_{ij} I_j \tag{C.14}$$

$$\sum_{j}^{A} m_j I_j = \sum_{j}^{A} \left(\sum_{i}^{K} k_i b_{ij} \right) I_j \tag{C.15}$$

$$\Rightarrow \; m_j = \sum_{i}^{K} k_i \, b_{ij} \tag{C.16}$$

Equation (C.16) relates the influence function coefficients m_j that describe the shape of the mirror to the KL coefficients k_i that describe the wavefront measured by the WFS. This relation can be rewritten in matrix form as:

$$\begin{pmatrix} m_1 \\ m_2 \\ \vdots \\ m_A \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1K} \\ b_{21} & b_{22} & \dots & b_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ b_{A1} & b_{A2} & \dots & b_{AK} \end{pmatrix} \begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ k_K \end{pmatrix} \tag{C.17}$$

The elements b_{ij} of the projection matrix are the only unknowns in the previous equation. They can be computed by taking the scalar product of the influence function I_n on both sides of Equation (C.12):

$$K_i \cdot I_n = \sum_{j}^{A} b_{ij} \, I_j \cdot I_n \tag{C.18}$$

To simplify matters, both basis are kept of the same dimension, which is fixed by the number A of actuators in the DM. Thus, all the indexes i, j, n are iterated from 1 to A and Equation (C.18) represents A different linear equations. All the possible scalar products

defined by $I_j \cdot I_n$ can be written as the elements of a $A \times A$ matrix denoted by S_{II}. The same can be said for the scalar products $K_i \cdot I_n$, denoted by S_{KI}. The projection matrix $\mathbf{B}$ is then another $A \times A$ matrix. Thus, Equation (C.18) can be rewritten in matrix form as:

$$\mathbf{S}_{KI} = \mathbf{B}\,\mathbf{S}_{II} \tag{C.19}$$

Equation (C.19) is multiplied on the right by the inverse matrix of $\mathbf{S}_{II}$, denoted by $\mathbf{S}_{II}^{-1}$, to produce an expression of the projection matrix $\mathbf{B}$:

$$\mathbf{B} = \mathbf{S}_{KI}\,\mathbf{S}_{II}^{-1} \tag{C.20}$$

To summarize, the KL coefficients $\vec{k}$ that describe the measured wavefront φ are obtained from the reconstruction matrix $\mathbf{D}^+$ and the WFS measurements $\vec{w}$. They are transformed into influence function coefficients $\vec{m}$ by the projection matrix $\mathbf{B}$, and converted to counts units to be sent to the DM controller to update the shape of the DM. In this way, the loop from WFS measurements to DM commands has been closed.

The reconstruction described above is performed by two individual operations, as described by Equation (C.4) and Equation (C.16). These two operations can be combined into one requiring just one matrix multiplication:

$$\left.\begin{aligned} \vec{k} &= \mathbf{D}^+\,\vec{w} \\ \vec{m} &= \mathbf{B}\,\vec{k} \end{aligned}\right\} \quad \Rightarrow \quad \vec{m} = \underbrace{\left[\mathbf{B}\,\mathbf{D}^+\right]}_{\mathbf{C}^+}\,\vec{w} = \mathbf{C}^+\,\vec{w} \tag{C.21}$$

$$\mathbf{C}^+ \equiv \mathbf{B}\,\mathbf{D}^+ \tag{C.22}$$

This matrix $\mathbf{C}^+$ is the control matrix of the system, defined as the product of the projection matrix $\mathbf{B}$ and the reconstruction matrix $\mathbf{D}^+$. The control matrix directly relates WFS measurements to DM actuator commands.

C.2 Empirical Calculation

The previous theoretical derivation is based on assumptions about the characteristics of the actual AO system. For example, it is assumed that the exact position of the subapertures of the WFS with respect to the pupil is known, and that the exact position of the influence functions and their shapes are known. It is also assumed that all actuators in the DM have a perfectly linear response and that they all have the same response. Any deviation of the system characteristics from these assumptions will affect the performance of the computed control matrix. A more reliable and realistic control matrix must be measured from the system itself [67].

The control matrix can be obtained by measuring the response of the WFS to each one of the actuators being pushed. The WFS measurements produced by individually pushing each of the A actuators in the DM, is obtained: $\vec{w}_0, \vec{w}_2, \ldots, \vec{w}_A$. The WFS response to each actuator is placed as a column in a matrix to create the *poke matrix* of the AO system.

$$\mathbf{C} = \begin{pmatrix} w_{11} & w_{12} & \ldots & w_{1A} \\ w_{21} & w_{22} & \ldots & w_{2A} \\ \vdots & \vdots & \ddots & \vdots \\ w_{S1} & w_{S2} & \ldots & w_{SA} \end{pmatrix} \tag{C.23}$$

The poke matrix $\mathbf{C}$ is of dimensions $S \times A$ and relates DM actuator commands to WFS measurements: $\vec{w} = \mathbf{C}\,\vec{m}$. Similarly to the previous theoretical derivation, this matrix is multiplied by a vector of DM actuator commands to obtain the WFS measurements that they would produce. Again, a simple example can be used to show this. The DM actuator vector describing actuator number 2 being pushed is written as: $\vec{m}_{ex} = (0, 1, 0, \cdot, 0)$. This vector is multiplied by the poke matrix C, producing a result that is equal to the 2nd column of $\mathbf{C}$, i.e., the WFS measurements obtained by poking actuator number 2: $\vec{w}_2 = \mathbf{C}\,\vec{m}_{ex}$.

The control matrix $\mathbf{C}^+$ is computed as the pseudo-inverse of the poke matrix $\mathbf{C}$, since $\mathbf{C}$ is in general not a square matrix. Thus, the control matrix $\mathbf{C}^+$ relates WFS measurements to DM actuator commands.

$$\mathbf{C}^+ \vec{w} = \mathbf{C}^+ \mathbf{C} \vec{m} \tag{C.24}$$

$$\vec{m} = \mathbf{C}^+ \vec{w} \tag{C.25}$$

However, one must be careful when computing the pseudo-inverse matrix of the poke matrix. In a sense, the poke matrix represents the internal characteristics of the AO system working in closed-loop. It only contains configurations that the AO system can produce and detect. Anything that is invisible to the AO system or that the AO system cannot reproduce is not reliably represented by the poke matrix. When computing the pseudo-inverse of $\mathbf{C}$, many of these problematic configurations may appear in the result, making the control matrix $\mathbf{C}^+$ unstable.

The matrices that the SVD method produces contain information about configurations that are inherent to the AO system, in a sense the eigenmodes and eigenvalues of the system. The noise propagation on a certain eigenmode of the poke matrix is proportional to the inverse of the corresponding eigenvalue [67]. Thus, an eigenmode with a very small eigenvalue must be filtered out from the final control matrix. The eigenmodes can be extracted from the $\mathbf{U}$ and $\mathbf{V}$ matrices produced by the SVD algorithm and then filtered to ensure no unstable configurations are present in the pseudo-inverse.

C.3 Application to PSF Estimation Method

Knowledge about the residual wavefront that was measured by the WFS and about the shape of the DM is required by the long-exposure PSF estimation method. For the method, the main concern is not the calculation of the control matrix of the AO system, but the individual matrices that relate WFS measurements and DM commands to wavefront, derived in

Section C.1. The PSF estimation method requires a matrix to transform the WFS measurements of the residual wavefront to KL coefficients; and another matrix to transform DM commands describing the shape of the mirror to KL coefficients. As shown previously, they can be theoretically calculated when the characteristics of the WFS and DM are known.

The reconstruction matrix $\mathbf{D}^+$ is computed following the method described in Section C.1. This matrix produces an estimation of the residual wavefront measured by the WFS in terms of KL coefficients. A projection matrix $\mathbf{B}'$ that transforms DM commands into KL coefficients is also computed as described. It should be noted that this new projection matrix $\mathbf{B}'$ is different from the one computed previously $\mathbf{B}$. This new projection matrix is calculated by expressing the influence functions in terms of KL modes:

$$I_i = \sum_j^K b'_{ji} K_j \tag{C.26}$$

The projection matrix elements b'_{ji} are computed following the same method as before, by taking the scalar product of both sides of Equation (C.26) with K_n. The same approach described in Eqs. (C.19) and (C.20) is followed.

$$I_i \cdot K_n = \sum_j^K b'_{ji} K_j \cdot K_n \tag{C.27}$$

$$\mathbf{S}_{IK} = \mathbf{B}' \mathbf{S}_{KK} \tag{C.28}$$

$$\mathbf{B}' = \mathbf{S}_{IK} \mathbf{S}_{KK}^{-1} \tag{C.29}$$

where $\mathbf{S}_{IK}$ is the matrix defined by the $I_i \cdot K_n$ scalar products and the $\mathbf{S}_{KK}^{-1}$ is the inverse of the matrix containing the $K_j \cdot K_n$ scalar products.

The PSF estimation method requires knowledge about the wavefront residuals after correction and about the original atmospheric seeing distortions, which are estimated from the DM commands. Thus, the PSF estimation method requires these two matrices: $\mathbf{D}^+$

and $\mathbf{B}'$. With the reconstruction matrix $\mathbf{D}^{+}$, the WFS measurements are interpreted as KL coefficients, i.e., the residual wavefront after correction. The projection matrix $\mathbf{B}'$ provides the shape of the DM described in terms of KL modes from the DM actuator commands, which is used to estimate the atmospheric seeing conditions characterized by the value of the Fried parameter r_0.

APPENDIX D

SHACK-HARTMANN WAVEFRONT SENSOR NOISE

This section provides an in-depth derivation of the WFS measurement noise in a cross-correlating Shack-Hartmann WFS working on extended sources, as is the case in solar AO. The work presented here is a compilation of non-published work by V. Michau [48, 47] from ONERA.

A cross-correlating Shack-Hartmann WFS measures the slope of the wavefront by performing cross-correlations between the subaperture images and a reference. For any particular subaperture, the position of the cross-correlation peak is proportional to its shift with respect to the reference subaperture and, hence, to the relative slope of the wavefront at that position. The subaperture images are produced by a lenslet array and captured by a the WFS camera. The image captured by the camera is affected by several noise sources, such as: photon noise (also called shot noise), readout noise, thermal noise (also called dark current) and inhomogeneous pixel response. The last two can be calibrated out by the AO system by capturing dark current and flat field images.

Readout noise is the noise produced by the camera amplifier, which converts the electron charge captured by a pixel into a voltage that can be measured. Photon noise arises from the inherent statistical nature of the photon detection process and is known to follow a Poison distribution. Thus, the standard deviation of the photon noise can be estimated as the square root of the measured signal. For solar observations, the main noise contributor is photon noise, due to the high light levels present, and readout noise can be safely neglected.

The presence of noise in the subaperture images affects the results of the cross-correlation functions and introduces noise in the wavefront sensor measurements. A study of the cross-correlation noise properties is required to obtain better estimations of the mea-

sured wavefront. An expression of the variance of the cross-correlation noise in a Shack-Hartmann WFS was derived by Michau [48] and later corrected [47]. Its derivation is given in this Appendix.

Consider the case of the cross-correlation between a subaperture image I and a subaperture reference I_r. Assume the relative shift between them is described by the vector (α, β). The cross-correlation function of I and I_r is defined by:

$$C(x,y) = \int_{S_r} I_r(u,v) I(u+x, v+y)\, du dv \tag{D.1}$$

All functions considered here are discrete functions due to the matrix nature of the WFS detector. The calculations using continuous functions and their integrals can be translated to discrete functions and discrete sums, neglecting the effects of this discretization. The above integral is an integral over S_r, which is the region where I_r is defined. This assumes that S, the region where I is defined, contains S_r for all the shifts considered.

The relative shift between I and I_r is determined by calculating the cross-correlation function C and computing the position of the center of gravity of the cross-correlation peak after thresholding. In general, subaperture images are affected by noise and introduce an uncertainty on the determination of the cross-correlation maximum.

Consider different realizations of a subaperture image, all of them images of the same object but affected differently by noise. The noise free subaperture images are defined as the ensemble average of all these realizations: $\langle I \rangle$ and $\langle I_r \rangle$. The autocorrelations of $\langle I \rangle$ and $\langle I_r \rangle$ are defined:

$$C_{\langle I_r \rangle}(x,y) = \int_{S_r} \langle I_r(u,v) \rangle \langle I_r(u+x, v+y) \rangle du\, dv \tag{D.2}$$

$$C_{\langle I \rangle}(x,y) = \int_{S_r} \langle I(u,v) \rangle \langle I(u+x, v+y) \rangle du\, dv \tag{D.3}$$

Consider D as the domain for which $C(x,y) \geq s$ where s is a threshold value. The center of mass of the cross-correlation function inside D determines the position of the cross-correlation peak. Thus, the abscissa of the center of mass of the thresholded cross-correlation function is given by:

$$x_g = \frac{\int_D x(C(x,y) - s)\, dxdy}{\int_D (C(x,y) - s)\, dxdy} \tag{D.4}$$

Some other useful quantities are defined below:

$$x_g = \frac{N_g}{D_g} \tag{D.5}$$

$$N_g = \int_D x(C(x,y) - s)\, dxdy \tag{D.6}$$

$$D_g = \int_D (C(x,y) - s)\, dxdy \tag{D.7}$$

$$\sigma_C^2(x,y,x',y') = \langle C(x,y)C(x',y')\rangle - \langle C(x,y)\rangle\langle C(x',y')\rangle \tag{D.8}$$

In the following derivations, the subaperture images I and I_r will be considered as random variables and the relative shift of I as a parameter. In the case where I and I_r are affected by noise, the measurement x_g will also be affected by noise. An expression for its variance $\sigma_{x_g}^2$ will be derived below.

D.1 Case Where I is Noisy and I_r is Noise Free

A simplified case where only I is affected by an additive noise n_i will be studied first. As indicated before, I_r is assumed fixed while I is shifted with respect to I_r by a vector (α, β).

$$I(x,y) = \langle I(x,y)\rangle + n_i(x,y) \tag{D.9}$$

$$I_r(x,y) = \langle I(x-\alpha, y-\beta)\rangle \tag{D.10}$$

The variance of x_g can be written from Equation (D.4) in terms of N_g and D_g. The fluctuations of the denominator can be neglected with respect to the fluctuations of the numerator producing:

$$\sigma^2_{x_g} = \frac{\langle N_g^2 \rangle - \langle N_g \rangle^2}{\langle D_g^2 \rangle} \tag{D.11}$$

The numerator of the previous expression can be expanded according to the previous definition of N_g:

$$\begin{aligned}
&\langle N_g^2 \rangle - \langle N_g \rangle^2 = \\
&= \langle \left(\int_D x(C(x,y) - s)\, dxdy \right)^2 \rangle - \langle \int_D x(C(x,y) - s)\, dxdy \rangle^2 \\
&= \int_D \int_D xx' \left[\langle C(x,y)C(x',y') \rangle - \langle C(x,y) \rangle \langle C(x',y') \rangle \right] dxdydx'dy' \\
&= \int_D \int_D xx' \sigma_C^2(x,y,x',y')\, dxdydx'dy'
\end{aligned} \tag{D.12}$$

The cross-correlation terms in Equation (D.12) can be also expanded as:

$$\begin{aligned}
&\langle C(x,y)C(x',y') \rangle = \\
&= \langle \int_{S_r} I_r(u,v) I(u+x, v+y)\, dudv \int_{S_r} I_r(u',v') I(u'+x', v'+y')\, du'dv' \rangle \\
&= \int_{S_r} \int_{S_r} I_r(u,v) I_r(u',v') \langle I(u+x, v+y) I(u'+x', v'+y') \rangle dudvdu'dv'
\end{aligned}$$

$$\begin{aligned}
&\langle C(x,y) \rangle \langle C(x',y') \rangle = \\
&= \langle \int_{S_r} I_r(u,v) I(u+x, v+y)\, dudv \rangle \langle \int_{S_r} I_r(u',v') I(u'+x', v'+y')\, du'dv' \rangle \\
&= \int_{S_r} \int_{S_r} I_r(u,v) I_r(u',v') \langle I(u+x, v+y) \rangle \langle I(u'+x', v'+y') \rangle dudvdu'dv'
\end{aligned}$$

Assuming the noise on I is white noise, i.e., noise affecting different pixels is decorrelated:

$$\langle I(x,y)I(x',y')\rangle - \langle I(x,y)\rangle\langle I(x',y')\rangle = \begin{cases} \sigma^2_{n_i}(x,y) & : (x,y) = (x',y') \\ 0 & : (x,y) \neq (x',y') \end{cases}$$

Thus, from Equation (D.12) it follows:

$$\sigma^2_C(x,y,x',y') = \int_{S_r} I_r(u,v) I_r(u+x-x', v+y-y') \sigma^2_{n_i}(u+x, v+y)\, du dv \quad \text{(D.13)}$$

The previous expression is greatly simplified assuming that the variance of the noise is constant across the subaperture image I: $\sigma^2_{n_i(u+x,v+y)} = \sigma^2_{n_i}$. From the fact that I_r is assumed to be noise free $I_r = \langle I_r \rangle$ and Equation (D.2):

$$\sigma^2_C(x,y,x',y') = \sigma^2_{n_i} C_{\langle I_r\rangle}(x-x', y-y') \quad \text{(D.14)}$$

Thus, the numerator of Equation (D.11) can be written as:

$$\langle N_g^2\rangle - \langle N_g\rangle^2 = \sigma^2_{n_i} \int_D \int_D x x' C_{\langle I_r\rangle}(x-x', y-y')\, dx dy dx' dy' \quad \text{(D.15)}$$

D.2 Case Where I_r is Noisy and I is Noise Free

In this case, subaperture image I is assumed to be noise free and subaperture reference I_r is affected by an additive noise n_{i_r}. As previously, I is shifted by a vector (α, β) with respect to I_r.

$$I_r(x,y) = \langle I_r(x,y)\rangle + n_{i_r}(x,y) \quad \text{(D.16)}$$

$$\langle I_r(x,y)\rangle = I(x-\alpha, y-\beta) \quad \text{(D.17)}$$

The cross-correlation of two functions is a symmetric function with respect to the exchange of I and I_r, as can be seen from its definition in Equation (D.1). Therefore, the derivation presented in the previous section can be directly applied to this case with I and I_r exchanged.

$$\langle N_g^2 \rangle - \langle N_g \rangle^2 = \sigma_{n_{i_r}}^2 \int_D \int_D xx' C_{\langle I \rangle}(x - x', y - y')\, dxdydx'dy' \tag{D.18}$$

where $\sigma_{n_{i_r}}^2$ represents the variance of the noise affecting the reference subaperture image I_r.

D.3 Case Where I_r and I are Both Noisy

Normally, both subaperture images are affected by noise. Again the images are assumed to be affected by a white additive noise n_i and n_{i_r}, respectively. Subaperture image I is shifted by a vector (α, β) with respect to I_r.

$$I_r(x, y) = \langle I_r(x, y) \rangle + n_{i_r}(x, y) \tag{D.19}$$

$$I(x, y) = \langle I(x, y) \rangle + n_i(x, y) \tag{D.20}$$

$$\langle I_r(x, y) \rangle = \langle I(x - \alpha, y - \beta) \rangle \tag{D.21}$$

The noise terms n_{i_r} and n_i are not correlated. In the calculation of σ_C^2, it is possible to decompose I (and I_r), as a function of $\langle I \rangle$ and n_i ($\langle I_r \rangle$ and n_{i_r}). Also, all the terms of the type $\langle n_i n_{i_r} \rangle$ are zero. The results from sections D.1 and D.2 can be combined to obtain a general result for noisy I and I_r:

$$\langle N_g^2 \rangle - \langle N_g \rangle^2 =$$
$$= \int_D \int_D xx' \left[\sigma_{n_{i_r}}^2 C_{\langle I \rangle}(x - x', y - y') + \sigma_{n_i}^2 C_{\langle I_r \rangle}(x - x', y - y') \right]$$
$$dxdydx'dy' \tag{D.22}$$

D.4 Simplified Expressions

Several approximations can applied to obtain a simplified expression for the noise variance σ_{x_g}. The following approximation relates the autocorrelations $C_{\langle I \rangle}$ and $C_{\langle I_r \rangle}$ around the maximum of $C_{\langle I_r \rangle}$.

$$C_{\langle I_r \rangle}(x, y) \approx C_{\langle I \rangle}(x, y)(S_r \otimes S_r)(x, y) \tag{D.23}$$

where the function $(S_r \otimes S_r)$ represents the convolution of S_r with itself. The autocorrelation function $C_{\langle I \rangle}$ has a maximum on $(0, 0)$ and its value is given by:

$$|I|_r^2 = \int_{S_r} \langle I(u, v) \rangle \langle I(u, v) \rangle dudv \tag{D.24}$$

Furthermore, the shape of the autocorrelation function $C_{\langle I \rangle}$ around the central peak can be approximated by the following parabolic function:

$$C_{\langle I \rangle}(x, y) = |I|_r^2 \left(1 - \frac{x^2 + y^2}{2\delta^2} \right) \tag{D.25}$$

where the parameter δ represents the half width at half maximum of the cross-correlation peak. The convolution function $(S_r \otimes S_r)$ can be expressed analytically in the case where S_r is a rectangle of dimensions $a_r \times b_r$. Thus, inside the rectangle defined by $[-a_r, a_r] \times [-b_r, b_r]$ this function is defined by:

$$(S_r \otimes S_r)(x,y) = \left(1 - \frac{|x|}{a_r}\right)\left(1 - \frac{|y|}{b_r}\right) \tag{D.26}$$

Using these relations, Equation (D.22) can be re-written in a much more explicit form:

$$\langle N_g^2 \rangle - \langle N_g \rangle^2 = |I|_r^2 \int_D \int_D xx' \left[1 - \frac{(x-x')^2 + (y-y')^2}{2\delta^2}\right] \times \left[\sigma_{n_{ir}}^2 + \sigma_{n_i}^2 \left(1 - \frac{|x-x'|}{a_r}\right)\left(1 - \frac{|y-y'|}{b_r}\right)\right] dxdydx'dy' \tag{D.27}$$

The same approach can be applied to calculate the denominator of Equation (D.11). From the definition in Equation (D.7) and the above relations, the following expression for $\langle D_g \rangle^2$ is found:

$$\langle D_g \rangle^2 = |I|_r^4 \left[\int_D \left(1 - \frac{x^2+y^2}{2\delta^2} - s\right) dxdy\right]^2 \tag{D.28}$$

The calculation of these integrals requires an explicit definition of the domain D that constitutes the domain for which $C(x,y) \geq s$, where s is the threshold applied in the center of mass calculations. Using the approximation from Equation (D.25) as the cross-correlation function, an estimation of the domain D can be computed as follows:

$$D \approx \begin{cases} 1 - \frac{x^2}{2\delta^2} \geq s & : \forall x \in [-a_s, a_s] \\ 1 - \frac{y^2}{2\delta^2} \geq s & : \forall y \in [-a_s, a_s] \end{cases} \tag{D.29}$$

$$\Rightarrow \quad a_s = \delta\sqrt{2(1-s)} \tag{D.30}$$

Once the domain D is defined, the integrals can be calculated and a final result for the variance of x_g, the measured abscissa of center of mass of the cross-correlation peak, is obtained:

$$\sigma_{x_g}^2 = \frac{4\delta^2 \left(\sigma_{n_i}^2 + \sigma_{n_{ir}}^2\right)}{|I|_r^2} \tag{D.31}$$

Moreover, if the noise variances of the image and the reference are assumed to be the same $\sigma_{n_i}^2 = \sigma_{n_{ir}}^2$, the previous expression can be further simplified as:

$$\sigma_{x_g}^2 = \frac{8\,\delta^2\,\sigma_n^2}{|I|_r^2} \tag{D.32}$$

With:

- σ_n^2: noise variance of subaperture image and reference
- δ: half width at half maximum of $C_{\langle I \rangle}$ the autocorrelation function of the noise free subaperture image
- $|I|_r^2$: maximum value of $C_{\langle I \rangle}$

D.5 WFS Noise from Real Data

The noise affecting the WFS shift measurements can be estimated from the power spectrum of real open-loop WFS measurements. Parseval's theorem states the relationship between the power of the signal computed in the time space and in the frequency space. It specifies that the power of the signal is the same in both spaces, i.e., it is not changed by a Fourier Transform. Parseval's theorem [12] for a real continuous signal is written as:

$$\int_{-\infty}^{\infty} |f(x)|^2\, dx = \int_{-\infty}^{\infty} |F(\nu)|^2\, d\nu \tag{D.33}$$

The time series of WFS measurements constitutes a discrete signal, which requires discrete Fourier transforms. The direct and inverse discrete Fourier transforms [12] are defined as:

$$F_u = \sum_{k=0}^{N-1} f_k \exp\left[-j2\pi uk/N\right] \tag{D.34}$$

$$f_k = \frac{1}{N} \sum_{u=0}^{N-1} F_u \exp\left[j2\pi uk/N\right] \tag{D.35}$$

Parseval's theorem for discrete signals [12] transforms into Equation (D.36).

$$\sum_{k=0}^{N-1} |f_k|^2 = \frac{1}{N} \sum_{u=0}^{N-1} |F_u|^2 \tag{D.36}$$

Parseval's theorem can be used to relate the statistical variance of a signal to its power spectral density (PSD). The variance of a signal is calculated as illustrated by Equation (D.37) and Equation (D.38) in the special case where the mean of the signal is zero. The angled parenthesis $\langle\rangle$ indicate temporal mean.

$$\sigma_f^2 = \langle f_k^2 \rangle - \langle f_k \rangle^2 \tag{D.37}$$

$$\langle f_k \rangle = 0 \;\rightarrow\; \sigma_f^2 = \langle f_k^2 \rangle \tag{D.38}$$

Both sides of Parseval's theorem are divided by the number of points in the signal minus one $N - 1$. The signal is assumed to have zero mean. If that is not the case, it can always be achieved by subtracting the mean from the signal.

$$\frac{1}{N-1} \sum_{k=0}^{N-1} |f_k|^2 = \frac{1}{N-1} \frac{1}{N} \sum_{u=0}^{N-1} |F_u|^2 \tag{D.39}$$

$$\sigma_f^2 = \frac{1}{N-1} \frac{1}{N} \sum_{u=0}^{N-1} |F_u|^2 \tag{D.40}$$

The left side of Equation (D.39) is directly the definition of the variance of the signal. The variance of the signal can be obtained from the modulus square of the Fourier transform

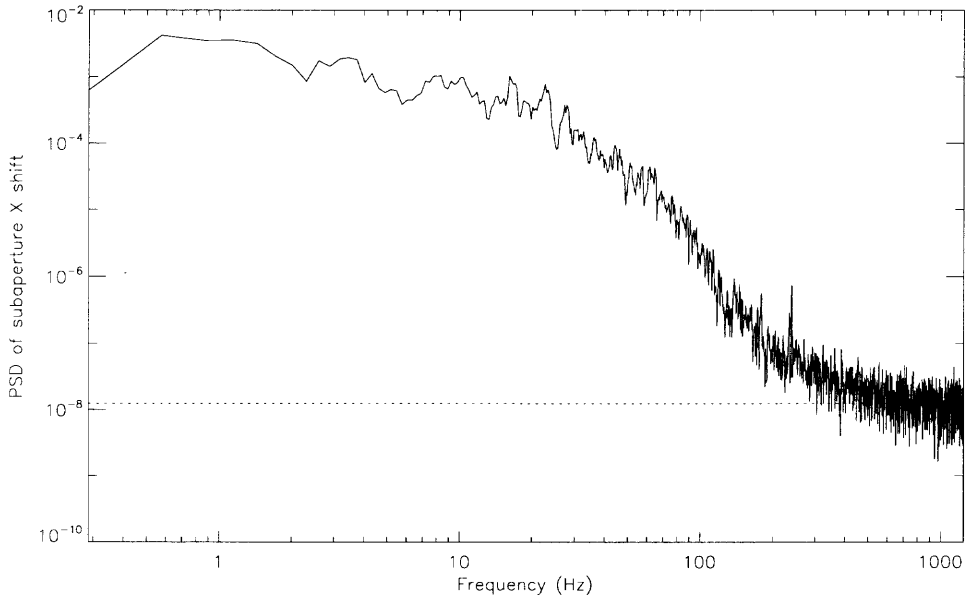

Figure D.1 Power spectrum of solar WFS x-shift open-loop data. The dotted line represents a fit to the noise tail of the power spectrum, which provides information about the noise variance of the signal.

amplitudes of the signal, i.e., the PSD of the signal. Hence, the variance of the signal is computed as the integral of its PSF.

$$\mathrm{PSD}\,(f)_u = |F_u|^2 \tag{D.41}$$

$$\sigma_f^2 = \frac{1}{N-1}\frac{1}{N}\sum_{u=0}^{N-1}\mathrm{PSD}\,(f)_u \tag{D.42}$$

The WFS x- and y-shift measurements are considered to be affected by additive white noise, i.e., noise that presents constant amplitude for all frequencies. Figure D.1 shows the PSD of the x-shifts measured by a certain subaperture of the WFS, i.e., the power that the signal has at each frequency. At very high frequencies, the power of the signal disappears and only the power of the noise remains. The tail of the plot in Figure D.1 corresponds to the power of the white noise that affects the signal.

The level of the tail of the PSD can be used to extrapolate the power of the white noise that affects all frequencies of the signal. The white noise power is represented in the figure by a dotted line. The integral of this white noise PSD over all frequencies can be used in Equation (D.42) to obtain the variance of the noise affecting the signal. The noise variance of the open-loop WFS x-shifts measurements obtained from the PSD shown in Figure D.1 was computed as: $\sigma^2_{x_g} \simeq 9 \times 10^{-5}\,\text{pixels}^2$ (or $\sim$ 4 nm of wavefront standard deviation). The variance of the signal is computed by integrating the PSD, which produces: $\sigma^2_{\text{signal}} \simeq 0.4\,\text{pixels}^2$. Hence, the signal to noise ratio of the WFS measurements is calculated to be around 45.

APPENDIX E

ALIASING ERROR

A description of the effects of aliasing [16] in the WFS measurements is given in this section. Insufficient sampling of a signal can make high frequency components show up as lower frequency components. This effect is called aliasing. Nyquist sampling theorem [12] states that a signal should be sampled at a rate at least twice its maximum frequency ν_M, the Nyquist frequency $\nu_N = 2\nu_M$. A signal sampled at a lower rate than its Nyquist frequency will be affected by aliasing.

Figure E.1 shows a very simple example demonstrating the effects of aliasing. The original signal presents just one frequency component at 9 Hz. This signal is then sampled at a rate of 10 Hz, which is lower than the required Nyquist sampling for this signal, i.e., 18 Hz. Sampling this signal at the given rate produces a signal of 1 Hz, i.e., the real signal of 9 Hz is *aliased* as a signal of 1 Hz.

The wavefront distortions produced by atmospheric turbulence can in general be considered as formed by an infinite number of spatial frequency components. A Shack-Hartmann WFS samples the wavefront with a limited number of subapertures of finite size placed in a particular geometry, both of which define its sampling spatial frequency. Since this sampling frequency is finite, the measurements produced by a Shack-Hartmann WFS are affected by aliasing.

A Shack-Hartmann WFS measures the slope of the wavefront and not the wavefront directly. The slopes are measured as pixel shifts between the subaperture images on the WFS camera. The wavefront is then reconstructed from the pixel shifts as a linear combination of basis functions, such as Zernike polynomials or KL functions. KL functions were chosen in this case because they have been proven to reduce the reconstruction error [16]. Even though both the KL functions and Zernike polynomials form an orthogonal basis of

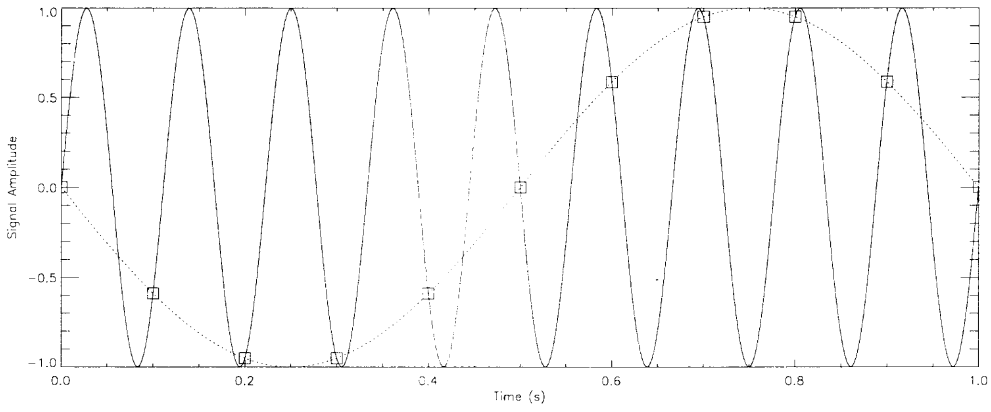

Figure E.1 Signal of frequency 9 Hz sampled at rate of 10 Hz produces a signal of 1 Hz.

the circular aperture space, their first derivatives do not. Thus, the reconstructed wavefront will be affected by cross-coupling errors arising from KL modes that were not included in the reconstruction, i.e., high-order KL modes.

The coefficients of the KL decomposition of the wavefront are obtained from the measured pixel shifts by means of the reconstruction matrix $\mathbf{D}^{+}$. The reconstruction matrix is the pseudo-inverse matrix of the interaction matrix $\mathbf{D}$, as described in Appendix C. The interaction matrix provides the WFS pixel shifts that a wavefront described by a series of KL coefficients would produce, as shown in Equation (C.4). The list of pixel shifts measured by the WFS are indicated by $\vec{w}$, and the KL coefficients that describe the wavefront are indicated by $\vec{k}$.

The series of infinite KL coefficients that describe a particular wavefront can be split into two sub-series: the low order KL coefficients that the AO system can correct $\vec{k}_{\parallel}$, and the high-order KL coefficients that the AO system cannot correct $\vec{k}_{\perp}$. An infinite interaction matrix $\mathbf{D}_{\infty}$ would consider all possible KL modes and could be similarly split in two components: $\mathbf{D}_{\parallel}$ and $\mathbf{D}_{\perp}$. The $\mathbf{D}_{\parallel}$ component refers to the low order KL modes component, which is the same interaction matrix previously defined $\mathbf{D}_{\parallel} \equiv \mathbf{D}$. Using this infinite interaction matrix, Equation (C.4) can be rewritten as:

$$\vec{w} = \mathbf{D}_{\infty}\,\vec{k} = [\mathbf{D} \quad \mathbf{D}_{\perp}] \begin{bmatrix} \vec{k}_{\parallel} \\ \vec{k}_{\perp} \end{bmatrix} \tag{E.1}$$

Due to the finite sampling frequency of the WFS, the higher order KL components $\vec{k}_{\perp}$ cannot be directly measured by the WFS. Their presence translates into aliasing and cross-coupling errors that affect the measurable KL components $\vec{k}_{\parallel}$. During normal AO operation, the reconstructed KL coefficients $\vec{k}'_{\parallel}$ that describe the low order components of the wavefront are obtained from the WFS measurements. This is done with the reconstruction matrix:

$$\vec{k}'_{\parallel} = \mathbf{D}^{+}\,\vec{w} \tag{E.2}$$

The prime in $k'_{\parallel}$ indicates that the KL coefficients are reconstructed from the WFS measurements. Equation (E.1) is substituted into Equation (E.2) producing:

$$\begin{aligned} \vec{k}'_{\parallel} &= \mathbf{D}^{+}\,\mathbf{D}_{\infty}\,\vec{k} \\ &= \mathbf{D}^{+}\,[\mathbf{D} \quad \mathbf{D}_{\perp}] \begin{bmatrix} \vec{k}_{\parallel} \\ \vec{k}_{\perp} \end{bmatrix} \\ &= \mathbf{D}^{+}\left(\mathbf{D}\,\vec{k}_{\parallel} + \mathbf{D}_{\perp}\,\vec{k}_{\perp}\right) \\ &= \vec{k}_{\parallel} + \mathbf{C}\,\vec{k}_{\perp} \end{aligned} \tag{E.3}$$

where the fact that the reconstruction matrix is the pseudo-inverse matrix of the interaction matrix has been taken into account: $\mathbf{D}^{+}\,\mathbf{D} = \mathcal{I}$. The newly defined matrix $\mathbf{C}$ is called the cross-talk matrix [16] and is written as:

$$\mathbf{C} \equiv \mathbf{D}^{+}\,\mathbf{D}_{\perp} \tag{E.4}$$

The $\mathbf{D}_{\perp}$ component is computed in the same manner as the regular interaction matrix $\mathbf{D}$ using the high-order KL modes. It describes the response of the WFS to these high-order KL modes. Practical purposes forbid the calculation of $\mathbf{D}_{\perp}$ using an infinite number of modes. An arbitrarily large KL mode order N_{∞} is chosen as the maximum KL mode.

The cross-talk matrix $\mathbf{C}$ quantifies the influence that high-order KL modes have on the lower order reconstructed KL coefficients, i.e., the effects of aliasing and cross-coupling. However, the required high order KL components cannot be directly measured, thus Equation (E.3) cannot be used in its present form. This difficulty is solved by the fact that the PSF estimation method only employs statistical quantities and only the covariance of $\vec{k}_{\parallel}$ is required. Taking the covariance of Equation (E.3) produces:

$$\operatorname{cov}\left(k'_{\parallel}, k'_{\parallel}\right) = \operatorname{cov}\left(k_{\parallel}, k_{\parallel}\right) + \mathbf{C}\operatorname{cov}\left(k_{\perp}, k_{\perp}\right)\mathbf{C}^{T} \tag{E.5}$$

Since the AO system cannot correct for the modes described by $\vec{k}_{\perp}$, they are not affected by correction and will follow the statistical properties described by the Kolmogorov model. Their covariance matrix is then calculated from the Kolmogorov model as given by Noll [52]. In Section 4.4.1, an expression for the covariance of the measured KL coefficients that describe the residual wavefront was introduced. Ignoring detection noise, Equation (4.42) can be compared with the previous relation to produce the following expression for $langler_i r_j\rangle$, the aliasing and cross-coupling covariance term:

$$\langle r_i r_j \rangle = \mathbf{C}\operatorname{cov}\left(k_{\perp}, k_{\perp}\right)\mathbf{C}^{T} \tag{E.6}$$

REFERENCES

[1] G. B. Arfken and H. J. Weber, *Mathematical methods for physicists* (Harcourt/Academic Press, 2001), 5th edn.

[2] H. W. Babcock, "The Possibility of Compensating Astronomical Seeing," Pub. of the Astron. Soc. of the Pac. **65**, 229-236 (1953).

[3] J. M. Beckers, "The Sacramento Peak Observatory Universal Birefringent Filter," in *Bulletin of the American Astronomical Society* pp. 269-+ (1973).

[4] J. M. Beckers, "Material motions in sunspot umbrae," Ap. J. **213**, 900-905 (1977).

[5] J. M. Beckers, "Increasing the Size of the Isoplanatic Patch with Multiconjugate Adaptive Optics," in *ESO Conference and Workshop Proceedings* pp. 693-+ (1988).

[6] J. M. Beckers and E. H. Schröter, "The Intensity, Velocity and Magnetic Structure of a Sunspot Region. IV: Properties of a Unipolar Sunspot," Solar Physics **10**, 384-+ (1969).

[7] L. R. Bellot Rubio, H. Balthasar, M. Collados, and R. Schlichenmaier, "Field-aligned Evershed flows in the photosphere of a sunspot penumbra," A&A **403**, L47-L50 (2003).

[8] L. R. Bellot Rubio, K. Langhans, and R. Schlichenmaier, "Multi-line spectroscopy of dark-cored penumbral filaments," A&A **443**, L7-L10 (2005).

[9] M. Born and E. Wolf, *Principles of Optics* (Pergamon Press, London, 1970), 4th edn.

[10] J. M. Borrero, "-," Priv. commun.

[11] J. M. Borrero, S. K. Solanki, L. R. Bellot Rubio, A. Lagg, and S. K. Mathew, "On the fine structure of sunspot penumbrae. I. A quantitative comparison of two semiempirical models with implications for the Evershed effect," A&A **422**, 1093-1104 (2004).

[12] E. O. Brigham, *The Fast Fourier Transform* (Prentice-Hall, Inc., New Jersey, 1974).

[13] J.-M. Conan, *Étude de la correction partielle en optique adaptative*, Ph.D. thesis, Université de Paris 11, Orsay, France (1994).

[14] J.-M. Conan, T. Fusco, L. M. Mugnier, et al., "Deconvolution of adaptive optics images: from theory to practice," in *Proceedings of the SPIE* **4007**, 913-924 (2000).

[15] G.-M. Dai, "Modal compensation of atmospheric turbulence with the use of Zernike polynomials and Karhunen-Loève functions," J. Opt. Soc. Am. A **12**, 2182-2193 (1995).

[16] G.-M. Dai, "Modal wave-front reconstruction with Zernike polynomials and Karhunen-Loève functions," J. Opt. Soc. Am. A **13**, 1218-1225 (1996).

[17] L. Delbouille, G. Roland, and L. Neven, *Atlas photometrique du spectre solaire de [lambda] 3000 a [lambda] 10000* (Liege: Universite de Liege, Institut d'Astrophysique, 1990, 1990).

[18] C. Denker, D. Mascarinas, Y. Xu, W. Cao, G. Yang, H. Wang, P. R. Goode, and T. Rimmele, "High-Spatial-Resolution Imaging Combining High-Order Adaptive Optics, Frame Selection, and Speckle Masking Reconstruction," Solar Physics **227**, 217-230 (2005).

[19] D. Dravins, L. Lindegren, E. Mezey, and A. T. Young, "Atmospheric Intensity Scintillation of Stars I. Statistical Distributions and Temporal Properties," Pub. of the Astron. Soc. of the Pac. **109**, 173-207 (1997).

[20] J. D. Drummond, "The Adaptive Optics Lorentzian Point Spread Function," in *Proceedings of the SPIE* **3353**, 1030-1037 (1998).

[21] B. L. Ellerbroek, "Linear systems modeling of adaptive optics in the spatial-frequency domain," Journal of the Optical Society of America A **22**, 310-322 (2005).

[22] J. Evershed, "Radial movement in sun-spots," The Observatory **32**, 291-292 (1909).

[23] D. L. Fried, "Optical Resolution Through a Randomly Inhomogeneous Medium for Very Long and Very Short Exposures," J. Opt. Soc. Am. **56**, 1372-1379 (1966).

[24] U. Frisch, *Turbulence. The Legacy of A.N. Kolmogorov* (Cambridge University Press, 1995), 1st edn.

[25] T. Fusco, C. Petit, G. Rousset, J.-F. Sauvage, K. Dohlen, D. Mouillet, J. Charton, P. Baudoz, M. Kasper, E. Fedrigo, P. Rabou, P. Feautrier, M. Downing, P. Gigan, J.-M. Conan, J.-L. Beuzit, N. Hubin, F. Wildi, and P. Puget, "Design of the extreme AO system for SPHERE, the planet finder instrument of the VLT," in *Proceedings of the SPIE* Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference **6272** (2006).

[26] G. D. Gatewood and C. V. Gatewood, "A study of Sirius," Ap. J. **225**, 191-197 (1978).

[27] D. T. Gavel, "Suppressing anomalous localized waffle behavior in least-squares wavefront reconstructors," in *Proceedings of the SPIE* **4839**, 972-980 (2003).

[28] J. W. Goodman, *Statistical Optics* (John Wiley & Sons, 1985), 1st edn.

[29] D. Gratadour, "NAOS PSF Reconstruction," Presentation given at Workshop on PSF Reconstruction 2004, May 10-12, Victoria, BC Canada.

[30] D. P. Greenwood, "Bandwidth specification for adaptive optics systems," J. Opt. Soc. Am. **67**, 390-393 (1977).

[31] D. P. Greenwood, "Mutual coherence function of a wave front corrected by zonal adaptive optics," J. Opt. Soc. Am. **69**, 549-554 (1979).

[32] U. Grossmann-Doerth, W. Schmidt, and E. H. Schroeter, "Size and temperature of umbral dots," A&A **156**, 347-353 (1986).

[33] J. W. Hardy, *Adaptive Optics for Astronomical Telescopes* (Oxford University Press, New York, 1998), 1st edn.

[34] K. Hartkorn and T. Rimmele, "Velocity Measurements of Umbral Dots," in *ASP Conf. Ser. 286* (2003), pp. 281-+.

[35] T. Heinemann, A. Nordlund, G. B. Scharmer, and H. C. Spruit, "MHD simulations of penumbra fine structure," ArXiv Astrophysics e-prints (2006).

[36] J. Herrmann, "Cross coupling and aliasing in modal wave-front estimation," J. Opt. Soc. Am. **71**, 989-992 (1981).

[37] J. O. Hinze, *Turbulence. An Introduction to Its Mechanism and Theory* (McGraw-Hill, 1959), 1st edn.

[38] L. Jolissaint, *Optique Adaptive au Foyer d'un Télescope de la Classe 1 mètre*, Ph.D. thesis, Université de Genève, Geneva, Switzerland (2001).

[39] L. Jolissaint, J.-P. Veran, and J. Marino, "OPERA, an automatic PSF reconstruction software for Shack-Hartmann AO systems: application to Altair," in *Proceedings of the SPIE* **5490**, 151-163 (2004).

[40] C. U. Keller, R. F. Stein, and A. Nordlund, "Observational Signatures of a Solar Small-Scale Global Dynamo," Proc. 200th AAS Meeting **34**, 792-+ (2002).

[41] A. N. Kolmogorov, "Local structure of turbulence in incompressible fluids with very high Reynolds number," Dokl. Akad. Nauk SSSR **30**, 301-305 (1941).

[42] K. Langhans, G. B. Scharmer, D. Kiselman, M. G. Löfdahl, and T. E. Berger, "Inclination of magnetic fields and flows in sunspot penumbrae," A&A **436**, 1087-1101 (2005).

[43] B. W. Lites, D. F. Elmore, P. Seagraves, and A. P. Skumanich, "Stokes Profile Analysis and Vector Magnetic Fields. VI. Fine Scale Structure of a Sunspot," Ap. J. **418**, 928-+ (1993).

[44] B. W. Lites, A. Skumanich, and G. B. Scharmer, "High-resolution spectra of solar magnetic features. I - Analysis of penumbral fine structure," Ap. J. **355**, 329-341 (1990).

[45] R. B. Makidon, A. Sivaramakrishnan, L. C. Roberts, Jr., B. R. Oppenheimer, and J. R. Graham, "Waffle mode error in the AEOS adaptive optics point-spread function," in *Proceedings of the SPIE* **4860**, 315-323 (2003).

[46] V. Martínez Pillet, "Spectral signature of uncombed penumbral magnetic fields," A&A **361**, 734-742 (2000).

[47] V. Michau, "Calcul de la précision de mesure d'un analyseur de Shack-Hartmann fonctionnant avec une source étendue," Internal ONERA Report (2002).

[48] V. Michau, G. Rousset, and J. C. Fontanella, "Wavefront Sensing from Extended Sources," Proceedings of 13th NSO/SP Summer Workshop **13**, 124-128 (1992).

[49] B. Montesinos and J. H. Thomas, "The Evershed effect in sunspots as a siphon flow along a magnetic flux tube," Nature **390**, 485-+ (1997).

[50] B. Montesinos and J. H. Thomas, "The Evershed Effect in Sunspots: A Theoretical Explanation," Astrophysics and Space Science **263**, 323-326 (1998).

[51] I. Newton, *Opticks* (Dover Publications, 1952).

[52] R. J. Noll, "Zernike polynomials and atmospheric turbulence," J. Opt. Soc. Am. **66**, 207-211 (1976).

[53] W. H. Press, S. A. Teukolsky, et al., *Numerical Recipes in C* (Cambridge University Press, Cambridge, 1994), 2nd edn.

[54] A. Quirrenbach, "Atmospheric turbulence," (2002), lecture at the CfAO summer school.

[55] S. H. Reiger, "Starlight scintillation and atmospheric turbulence," A. J. **68**, 395-406 (1963).

[56] D. Ren, S. L. Hegwer, T. Rimmele, L. V. Didkovsky, and P. R. Goode, "Optical design of high-order adaptive optics for the NSO Dunn Solar Telescope and the Big Bear Solar Observatory," in *Proceedings of the SPIE* **4853**, 593-599 (2003).

[57] T. Rimmele and J. Marino, "The Evershed Flow: Flow Geometry and Its Temporal Evolution," Ap. J. **646**, 593-604 (2006).

[58] T. R. Rimmele, "On the temporal behaviour of the Evershed effect." A&A **290**, 972-982 (1994).

[59] T. R. Rimmele, "Evidence for thin elevated evershed channels." A&A **298**, 260-+ (1995).

[60] T. R. Rimmele, "Sun center observations of the Evershed effect," Ap. J. **445**, 511-516 (1995).

[61] T. R. Rimmele, "Solar Adaptive Optics," in *Proceedings of the SPIE* **4007**, 218-231 (2000).

[62] T. R. Rimmele, "Plasma Flows Observed in Magnetic Flux Concentrations and Sunspot Fine Structure Using Adaptive Optics," Ap. J. **604**, 906-923 (2004).

[63] T. R. Rimmele, "Recent advances in solar adaptive optics," in *Proceedings of the SPIE* Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference **5490**, 34-46 (2004).

[64] T. R. Rimmele and R. R. Radick, "Solar adaptive optics at the National Solar Observatory," in *Proceedings of the SPIE*, D. Bonaccini and R. K. Tyson, eds., Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference **3353**, 72-81 (1998).

[65] L. C. Roberts, Jr., M. D. Perrin, F. Marchis, A. Sivaramakrishnan, R. B. Makidon, J. C. Christou, B. A. Macintosh, L. A. Poyneer, M. A. van Dam, and M. Troy, "Is that really your Strehl ratio?" in *Proceedings of the SPIE* **5490**, 504-515 (2004).

[66] F. Roddier, "The Effects of Atmospheric Turbulence in Optical Astronomy," Prog. Optics **19**, 281-376 (1981).

[67] F. Roddier, *Adaptive Optics in Astronomy* (Cambridge University Press, 1999), 1st edn.

[68] N. Roddier, "Atmospheric wavefront simulation using Zernike polynomials," Opt. Eng. **29**, 1174-1180 (1990).

[69] M. C. Roggemann and B. Welsh, *Imaging Through Turbulence* (CRC Press Inc., 1996).

[70] G. B. Scharmer, P. M. Dettori, M. G. Lofdahl, and M. Shand, "Adaptive optics system for the new Swedish solar telescope," in *Proceedings of the SPIE* **4853**, 370-380 (2003).

[71] G. B. Scharmer, B. V. Gudiksen, D. Kiselman, M. G. Löfdahl, and L. H. M. Rouppe van der Voort, "Dark cores in sunspot penumbral filaments," Nature **420**, 151-153 (2002).

[72] R. Schlichenmaier, "Penumbral fine structure: Theoretical understanding," Astronomische Nachrichten **323**, 303-308 (2002).

[73] R. Schlichenmaier, L. R. Bellot Rubio, and A. Tritschler, "Two-dimensional spectroscopy of a sunspot. II. Penumbral line asymmetries," A&A **415**, 731-737 (2004).

[74] R. Schlichenmaier, L. R. Bellot Rubio, and A. Tritschler, "On the relation between penumbral intensity and flow filaments," Astronomische Nachrichten **326**, 301-304 (2005).

[75] R. Schlichenmaier, K. Jahn, and H. U. Schmidt, "A Dynamical Model for the Penumbral Fine Structure and the Evershed Effect in Sunspots," Astrophysical Journal Letters **493**, L121+ (1998).

[76] R. Schlichenmaier, K. Jahn, and H. U. Schmidt, "Magnetic flux tubes evolving in sunspots. A model for the penumbral fine structure and the Evershed flow," A&A **337**, 897-910 (1998).

[77] R. Schlichenmaier and S. K. Solanki, "On the heat transport in a sunspot penumbra," A&A **411**, 257-262 (2003).

[78] E. J. Seykora, "Solar scintillation and the monitoring of solar seeing," Solar Physics **145**, 389-397 (1993).

[79] R. A. Shine, A. M. Title, T. D. Tarbell, K. Smith, Z. A. Frank, and G. Scharmer, "High-resolution observations of the Evershed effect in sunspots," Ap. J. **430**, 413-424 (1994).

[80] M. Sobotka and P. Sütterlin, "Fine structure in sunspots. IV. Penumbral grains in speckle reconstructed images," A&A **380**, 714-718 (2001).

[81] S. K. Solanki, "Smallscale Solar Magnetic Fields - an Overview," Space Science Reviews **63**, 1-+ (1993).

[82] S. K. Solanki and C. A. P. Montavon, "Uncombed fields as the source of the broad-band circular polarization of sunspots," A&A **275**, 283-+ (1993).

[83] H. C. Spruit and G. B. Scharmer, "Fine structure, magnetic field and heating of sunspot penumbrae," A&A **447**, 343-354 (2006).

[84] D. C. H. Stanchfield, II, J. H. Thomas, and B. W. Lites, "The Vector Magnetic Field, Evershed Flow, and Intensity in a Sunspot," Ap. J. **477**, 485-+ (1997).

[85] J. L. Starck, E. Pantin, and F. Murtagh, "Deconvolution in Astronomy: A Review," Pub. of the Astron. Soc. of the Pac. **114**, 1051-1069 (2002).

[86] M. Stix, *The Sun, An Introduction* (Springer-Verlag, Berlin Heidelberg, 1989), 1st edn.

[87] G. Strang, *Linear Algebra and Its Applications* (Saunders College Publishing, 1988), 3rd edn.

[88] P. Sütterlin, L. R. Bellot Rubio, and R. Schlichenmaier, "Asymmetrical appearance of dark-cored filaments in sunspot penumbrae," A&A **424**, 1049-1053 (2004).

[89] V. I. Tatarski, *Wave Propagation in a Turbulent Medium* (McGraw-Hill, New York, 1961).

[90] J.-C. Thelen and F. Cattaneo, "A Dynamo Driven by Turbulent Thermal Convection," in *Bulletin of the American Astronomical Society* **31**, 858-+ (1999).

[91] J. H. Thomas and B. Montesinos, "A siphon-flow model of the photospheric Evershed flow in a sunspot," Ap. J. **407**, 398-401 (1993).

[92] J. H. Thomas and N. O. Weiss, "The theory of sunspots," in *NATO ASIC Proc. 375: Sunspots. Theory and Observations*, J. H. Thomas and N. O. Weiss, eds., pp. 3-59 (1992).

[93] J. H. Thomas and N. O. Weiss, "Fine Structure in Sunspots," Annual Review of Astronomy & Astrophysics **42**, 517-548 (2004).

[94] J. H. Thomas, N. O. Weiss, S. M. Tobias, and N. H. Brummell, "Downward pumping of magnetic flux as the cause of filamentary structures in sunspot penumbrae," Nature **420**, 390-393 (2002).

[95] A. M. Title, Z. A. Frank, R. A. Shine, T. D. Tarbell, K. P. Topka, G. Scharmer, and W. Schmidt, "On the magnetic and velocity field geometry of simple sunspots," Ap. J. **403**, 780-796 (1993).

[96] A. Tritschler, R. Schlichenmaier, L. R. Bellot Rubio, the KAOS Team, T. Berkefeld, and T. Schelenz, "Two-dimensional spectroscopy of a sunspot. I. Properties of the penumbral fine structure," A&A **415**, 717-729 (2004).

[97] R. K. Tyson, *Principles of Adaptive Optics* (Academic Press, 1991).

[98] H. Uitenbroek, A. Tritschler, and T. Rimmele, "The contrast in the G-Band in high-resolution filtergrams," Submitted to Ap. J. (2007).

[99] N. Vedrenne, V. Michau, C. Robert, J.-M. Conan, and T. Fusco, "Wavefront sensing on extended sources: anisoplanatism effects on phase estimation," in *Proceedings of the SPIE* **5894**, 259-268 (2005).

[100] J.-P. Véran, *Estimation de la Reponse Impulsionnelle et Restauration d'Image en Optique Adaptative: Application au Systeme d'Optique Adaptative du Telescope Canada-France-Hawaii*, Ph.D. thesis, Ecole nationale supérieure des télécommunications, Paris, France (1997).

[101] J.-P. Véran, F. Rigaut, H. Maître, and D. Rouan, "Estimation of the Adaptive Optics Long-Exposure Point-Spread Function Using Control Loop Data," Opt. Soc. of Am. **14**, 3057-3069 (1997).

[102] J. Y. Wang and J. K. Markey, "Modal compensation of atmospheric turbulence phase distortion," Journal of the Optical Society of America (1917-1983) **68**, 78-87 (1978).

[103] N. O. Weiss, "The Puzzling Structure of a Sunspot," in *Solar MHD Theory and Observations: A High Spatial Resolution Perspective*, J. Leibacher, R. F. Stein, and H. Uitenbroek, eds. (2006), pp. 213-+.

[104] C. Westendorp Plaza, J. C. del Toro Iniesta, B. Ruiz Cobo, V. Martinez Pillet, B. W. Lites, and A. Skumanich, "Evidence for a downward mass flux in the penumbral region of a sunspot," Nature **389**, 47-+ (1997).

[105] E. Wiehr and G. Stellmacher, "Velocity and magnetic field fluctuations in penumbral fine-structures," A&A **225**, 528-532 (1989).

[106] V. Wilken, C. R. de Boer, C. Denker, and F. Kneer, "Speckle measurements of the centre-to-limb variation of the solar granulation." A&A **325**, 819-824 (1997).

[107] F. Wöger and O. von der Lühe, "Amplitude Calibration of Adaptive Optics Supported Solar Speckle Imaging," in *OSA Topical Meeting: Adaptive Optics Analysis and Methods* (2007).