Adaptive Optics For Stellar Interferometry

Nazim Ali Bharmal

Cavendish Astrophysics
and Selwyn College, Cambridge

16th June 2004

A dissertation
submitted for the degree of Doctor of Philosophy
in the University of Cambridge

Declaration of Originality

This dissertation is the result of work carried out in the Astrophysics Group of the Cavendish Laboratory, Cambridge, between October 2000 and June 2004. The work contained in this thesis is my own except where stated otherwise. No part of this dissertation has been submitted for a degree, diploma or other qualification at this, or any other university. The total length of this dissertation does not exceed sixty thousand words.

Nazim Bharmal
16th June 2004

Acknowledgements

I acknowledge financial support from PPARC in the form of a research studentship, and from the Cavendish Laboratory during my final year. I also thank Selwyn College for their support, both financial and administrative, during the past eight years.

The COAST group has made my time at the Cavendish a unique experience, and taught me many important, and fun, lessons about professional astronomy (and astronomers.) In particular, I want to thank my supervisors Dave Buscher & Peter Warner, for helping me to think more carefully about physics and making my project less intimidating, Chris Haniff, for helping my writing to be far more legible, and John Young, for creating the COAST data reduction software which let me avoid a task worthy of Ulysses. I am also grateful to John B, Donald, Roger, Nathalie, Bodie, Debbie, Bob, Amanda.

Without LaTeX and BibTeX this thesis would have been less organised, and without Vim, Python, NumPy, Linux (the kernel and the distributions), PGPLOT, and David Schmenk's MX-5 CCD drivers, a lot more money would have been spent for no extra gain in quality. Not only is it amazing that such software is simply given away, but it is literally the foundation upon which this thesis was constructed.

Of course, Cambridge isn't just about work (or maybe I mean rowing, punting, and May balls.) Having been here with me for nearly eight years, Michael and Ai-Lan have proved themselves to be great friends and a benchmark against which I shall judge all other chemists/couples. TonyJohn proved that leaving Cambridge didn't mean leaving Cambridge, and Alex RW almost proved the opposite. Helen helped me decide that physics didn't necessarily mean a life amongst the dreaming spires, and the other fellows that made this PhD more than academic include Alejandro, Rachel P, Ross W, James, Ed, Gayoung, Elsa, Bridget, Rosie, and Harriet.

My family have sometimes wondered at the length of time that I've spent at university, but they have supported me throughout—as well as always letting me escape back to London. And to Julie-Ann; the phone calls and visits to Rome were much needed sanity checks, and without you all the past year would have been infinitely harder.

A man's errors are his portals of discovery.
- James Joyce

There is nothing so ridiculous but some philosopher has said it.
- Cicero

Glossary

AO (Adaptive Optics) AO corrects for phase aberrations in real time (at a rate greater than 10 Hz.)

Autoguider AO system in the COAST interferometer which corrects for linear phase aberrations.

Curvature Sensing (CS) Curvature Sensing uses the intensities in two defocused planes, equally distanced but in opposite directions from the focus, to measure a wavefront.

CWFS (Curvature Wavefront Sensor) A CWFS uses Curvature Sensing as its principle of operation.

Difference Signals The Difference Signals, inter-plane and intra-plane, are the calculations made from the intensities in a DPS-based WFS.

DM (Deformable Mirror) A DM is an optical element which can change the phase across the aperture of a light beam incident upon it (not necessarily a reflective mirror).

DPS (Diffractive Phase Sensing) A new method of wavefront measurement, introduced in this dissertation, that uses the intensities measured in an intra-focal and extra-focal plane.

Extra-focal A plane perpendicular to the optical axis that is after the focal plane.

Fast mirror A flat mirror in each COAST telescope which has PZT actuators that can rotate the mirror at high speeds (at least 100 Hz).

FOV (Field of View) The angular region over the sky from which useful measurements can be made given a fixed pointing direction.

G-tilt The image motion from the average phase gradient across an aperture.

IIR (Infinite Impulse Response) A type of digital filter used in the autoguider control software.

Intra-focal A plane perpendicular to the optical axis that is before the focus.

Intra-plane A difference signal which is computed from the difference of diametrically opposite intensity measurements, where the intensity is an average of that in the intra-focal and extra-focal planes.

Inter-plane A difference signal which is computed from the difference of intensity measurements from the same points in the intra-focal and extra-focal planes.

Isophote A contour of equal intensity.

$m-$even A phase aberration that has even symmetry when rotated by π.

$m-$odd A phase aberration that has odd symmetry when rotated by π.

MTF (Modulation Transfer Function) The modulus of the OTF, it contains all information about the image resolution.

MX-5 (Starlight Xpress MX-5) A model of CCD camera used for the NES Prototype.

NES (Nine Element Sensor) Numerical and prototype WFS design for evaluating the DPS technique.

OPD (Optical Path Difference) The difference in path distance of two differing light rays between two points.

OTF (Optical Transfer Function) The Fourier transform of the PSF, which provides information on the sampling by, and aberrations from, the telescope aperture.

Piston The mean phase aberration introduced at a telescope aperture, irrelevant for imaging but important for interferometry.

POB (Parallel Output Beamsplitter) A form of beamsplitter, with a polarisation-selecting interface, that has output beams which propagate in parallel.

PSF (Point Spread Function) The intensity at the focal plane of a telescope, from a point source at infinity.

PZT (Piezoelectric Transducer) Substance, often ceramic, which deforms upon application of a voltage across its surface.

Measurement matrix Matrix containing the difference signals for a linear wavefront sensing scheme (the pseudo-inverse is used for wavefront sensing).

QC (Quadrant Cell or Quadcell) A 4 pixel WFS which is used to measure the amplitude of tip/tilt phase aberrations.

Quadratic Aberrations The three Zernike polynomial aberrations containing only quadratic terms: defocus, astigmatism 1 and astigmatism 2.

Shuffled Charge Imaging Used with the MX-5 camera for the NES Prototype to read out the charge within an entire CCD sequentially.

Strehl Ratio Measurement of the variance of a wavefront across an aperture.

SVD (Singular Value Decomposition) Matrix decomposition algorithm, used to form the pseudo-inverse of the measurement matrix.

TC (Telescope Controller) A system within the COAST telescope responsible for tracking the sky rotation via the telescope siderostats.

Tip/tilt Linear phase aberration which causes the gross offset of the image at the focal plane.

USB (Universal Serial Bus) The I/O interface used by the MX-5 camera.

Visibility Corresponds to the amplitude squared of the complex visibility, unless explicitly stated otherwise.

WFS (Wavefront Sensor) AO component which measures the wavefront across an aperture.

Z-tilt The image motion from the least squares fit of tip/tilt to the phase aberrations across an aperture.

Contents

1. Introduction

1.1. Astronomy as an Imaging Science

The era of modern astronomy began with the first observations of the moons orbiting Jupiter by Galileo Galilei in 1611. The technological advance which made the observation possible was the telescope, which Galileo constructed to magnify images of distant objects in the sky. With the naked eye, points that have an angular separation of less than 1 arc minute cannot be individually resolved, but his telescope reduced this limit to ~ 3 arc seconds.

Instruments offering higher resolutions than that previously available have invariably resulted in new, and often unexpected, astrophysics. The minimum angular separation which a diffraction-limited telescope, of diameter D, can resolve is $\sim \lambda/D$. For example, with a 10 m aperture this limiting resolution at visible wavelengths is ~ 0.01 arc seconds which is approximately the size of the largest stars in the sky. However, when observing through the atmosphere the limit is ~ 1 arc seconds—equivalent to a diffraction-limited 10 cm aperture.

The collective effect by the atmosphere on visible and near-IR light passing through it is called seeing. Its origin is turbulent mixing of air parcels which have differing temperatures, and this variation causes corresponding variations in the refractive index. Above the atmosphere, wavefronts

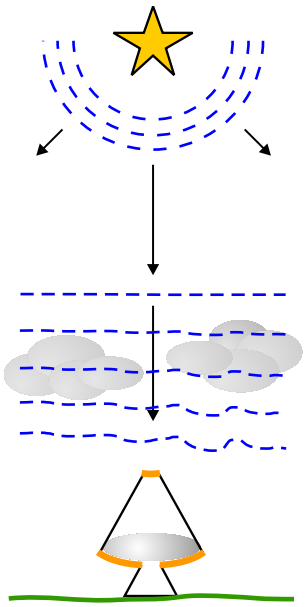

Figure 1.1.: A wavefront from a star arrives flat at the Earth due to the great distance from the object. As it passes through the atmosphere, the wavefront is distorted and accumulates phase aberrations.

from stars and other distant objects arrive flat at the Earth. These wavefronts then travel through the atmosphere and become distorted. When they reach the telescope aperture, the resultant phase aberrations reduce the coherence of the light at different points in the pupil. The coherence can be described by a scale length, which governs the image resolution, and a typical value is $\sim 10\,\mathrm{cm}$.

The seeing is reduced at high altitudes, and observatories on mountain tops have obtained seeing-limited resolutions as small as ~ 0.2 arc seconds. Space based telescopes avoid these problems entirely, but with consequent disadvantages such as (astronomical!) cost. The seeing limit is increasingly problematic both as optical telescopes become larger and as non-optical facilities, such as ALMA and VLBI, offer better resolutions. In order to have comparable fidelity in optical images, differing strategies have been developed to combat seeing at ground level.

The two methods that will be studied here are Adaptive Optics and Stellar Interferometry. Adaptive Optics (AO) corrects for the phase aberrations at the telescope aperture, and increases the image resolution up to the diffraction limit. Stellar Interferometry instead coherently interferes light from two or more separated apertures. The van Cittert-Zernike theorem states that the resulting interference fringes measure the Fourier transform of sky intensity at a frequency inverse to the aperture separation. The apertures sizes can be relatively small, because it is their spacing that determines the final resolution, and so seeing is less problematic for an interferometer.

However, the use of small apertures limits the sensitivity, and to use larger apertures it becomes necessary to correct the phase aberrations. Using Adaptive Optics with Stellar Interferometry achieves this, and the aim here is to examine this union of methods.

1.1.1. High Resolution Imaging with Adaptive Optics

Adaptive Optics was first conceptualised by Horace Babcock in 1953 (3). Closed loop AO consists of a feedback mechanism between a corrector and wavefront measurement sensor, and is the form used exclusively by astronomy. Measurements made of the wavefront are corrected and remaining phase errors are iteratively corrected. Open loop AO, conversely, does not use direct feedback and will not be further considered.

The typical layout of a closed loop AO system is shown in figure 1.2. The telescope aperture is imaged onto a deformable mirror (DM), which actively changes shape to match the inverse wavefront of the incoming light. The reflected light will ideally have a flat wavefront, but this correction will not be exact and, via a beamsplitter (BS), the remaining phase is detected by a Wavefront Sensor (WFS). The WFS measurements are sent to a Control System which updates the inferred wavefront and reshapes the DM to converge on the actual wavefront. As the phase aberrations evolve, the DM correction is only accurate for a short period of time and so its shape is repeatedly updated from new WFS measurements. This process requires short integrations of wavefront samples from a source in the sky, called the guide star, which is typically a star or other bright and compact object in the field of view.

The technologies needed to compensate for phase aberrations (~ 10000 independent correction areas for a 10 m telescope) and make rapid wavefront measurements (approximately at 1 kHz) are challenging. The first telescope AO system (26) was constructed for the the US military in the late 1970's, and the first use of AO in scientific astronomy came a decade later, in 1989, with

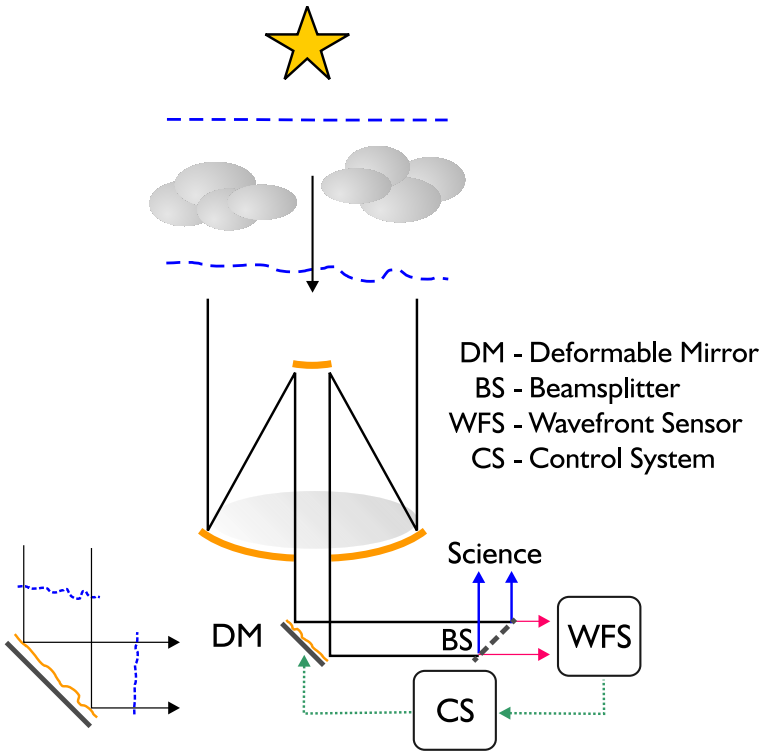

Figure 1.2.: The layout of a closed loop Adaptive Optics system.

the COME-ON instrument (68). Since that time, the design of almost every new telescope with apertures larger than 2 m has incorporated some form of AO system (88) and many older telescope have been retrofitted with AO (73). Current AO research now focuses on using artificial guide stars from lasers focused in the atmosphere (31), extending the corrected field of view (32), and developing correction for extremely large telescopes (40) (ELTs) which have diameters greater than 10 m.

1.1.2. High Resolution Imaging with Stellar Interferometry

In 1869 Fizeau suggested the use of slits cut in an otherwise opaque mask to create interference fringes from starlight. A quantitative theory of interferometry in astronomy was subsequently developed by Michelson in 1890, and he used it to publish the first measurements of stellar diameters in 1921. Fizeau and Michelson developed astronomical interferometry, independently it is thought, as a technology to overcome the seeing-resolution limit i.e., the object of AO, but achieved in an indirect fashion.

A diagram of Michelson's 20 ft interferometer is drawn in figure 1.3. The outer mirrors are displaced, along the baseline, $B \leq 20\,\text{ft}$, and the inner mirrors allow the light to be coherently combined at the focus of a telescope—analogous to a Young's double slit experiment. As the diameter of the telescope, $D = 2.5\,\text{m}$, is smaller than the baseline, the fringes measure higher resolution information than the telescope mirror alone could. The need for this high resolution was shown when Michelson used a 3 m baseline to measure the diameter of Betelgeuse (Alpha Orionis),

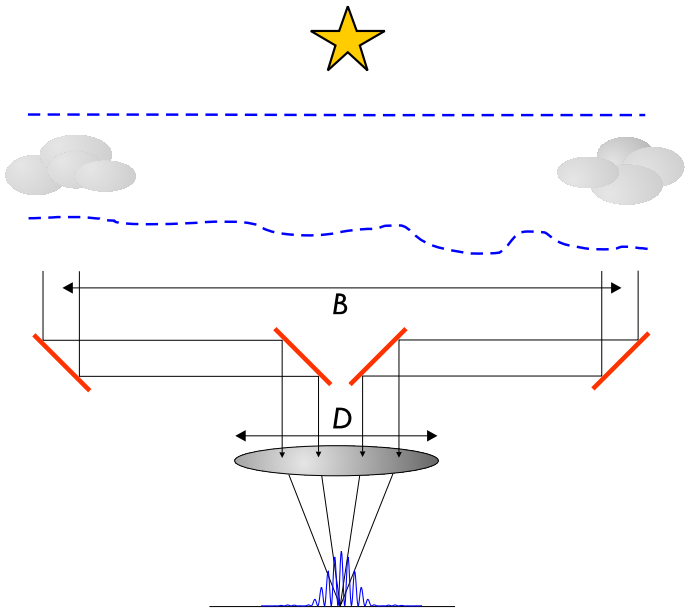

Figure 1.3.: A schematic diagram of Michelson's 20 ft interferometer. The distance between the two outer (movable) mirrors defines the baseline, B, and the interference fringes are formed the telescope focus.

the largest star in the sky. Even today, the largest telescopes have diameters of 10 m and, assuming diffraction limited performance, can resolve only a handful of the nearest large stars.

Michelson's 20 ft interferometer was, however, difficult to use because seeing causes both a reduction in the fringe contrast—known as the visibility—and introduces motions into the fringe pattern. As his instrument had small apertures (~ 25 cm) and he visually observed the fringes at the focus, these effects could be overcome. The stellar diameters were therefore inferred by changing the baseline until the intrinsic visibility was zero, and could only be determined for the brightest stars. The successor to the 20 ft, the 50 ft interferometer, never produced significant scientific results because of worse seeing effects with longer baselines and the difficulty in maintaining equal path lengths between each aperture and the telescope focus.

Later developments in stellar interferometry concentrated on using two (or more) physically separated apertures. In these schemes, the light is propagated to a central location (laboratory) via a path compensator which maintains coherence between the beams. In the laboratory, the light is interfered in a beam combiner and the fringes are recorded electronically. This approach is called Long Baseline Interferometry and the basic design is shown in figure 1.4. The technology to control path differences only became available 50 years after Michelson's experiments, and in 1975 the I2T (37) became the first long-baseline interferometer to produce fringes. The advantage of the design is that the baselines can be extended over many metres and the current record is ~ 300 m, at the CHARA array (47), which is equivalent to a resolution of ~ 0.3 milliarc seconds.

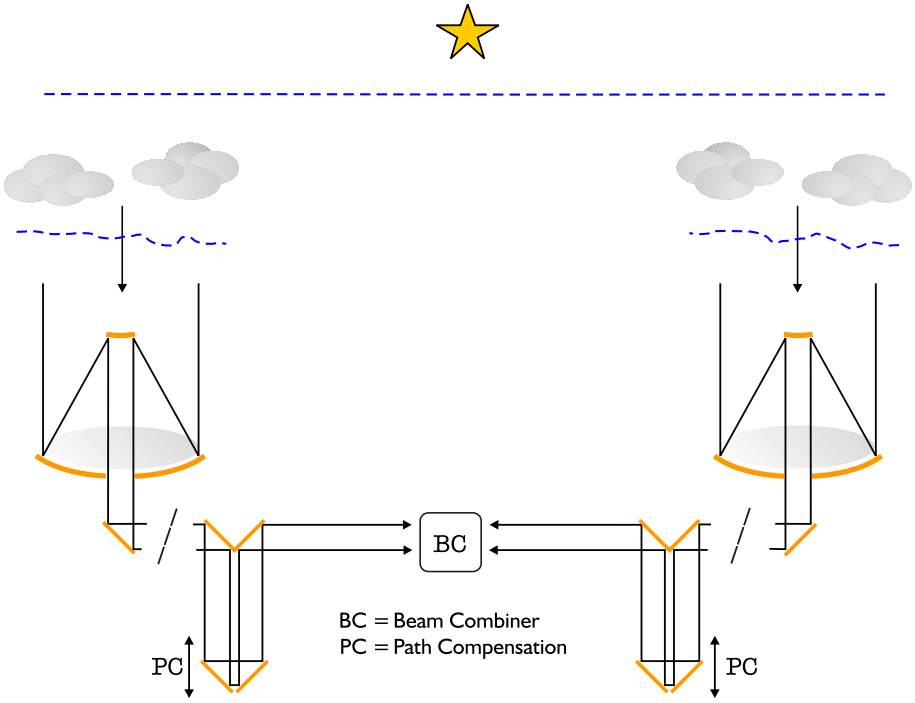

Figure 1.4.: Long Baseline Interferometry; two apertures are shown and the light from each is propagated, via a path compensation system, to the beam combiner where interference fringes are formed.

Figure 1.5.: COAST in profile; visible in this photograph are the 5 array telescopes (3 are located together at the centre of the picture), the laboratory bunker, and the light propagation pipes.

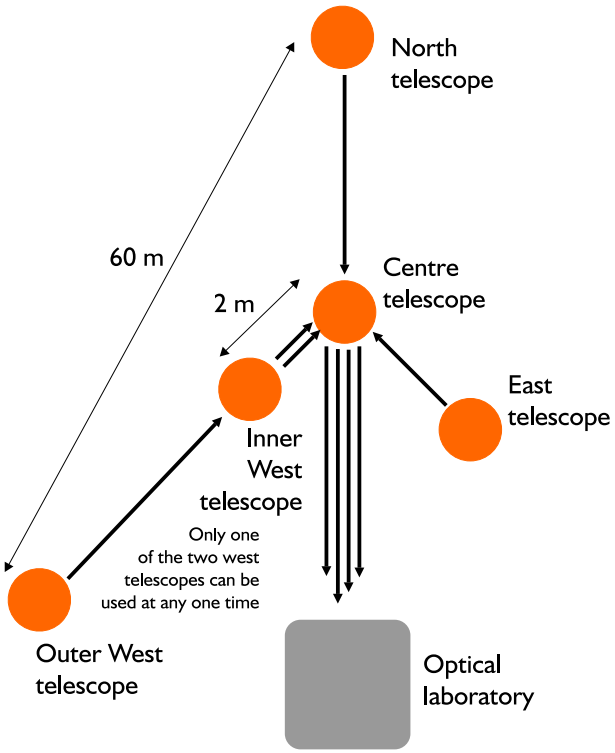

Figure 1.6.: Layout of the 5 apertures of the COAST telescope (spring 2004). The longest and shortest baselines, 2 m and 60 m, are indicated and the arrows indicate the light propagation from each telescope to the laboratory.

1.2. The Cambridge Optical Aperture Synthesis Telescope

An example of a long-baseline stellar interferometer, used for the research in this dissertation, is the Cambridge Optical Aperture Synthesis Telescope (COAST) (23) (an external photograph is shown in figure 1.5). COAST operates in the visible and near-IR ($0.65 - 2.3\,\mu$m) at the Lord's Bridge Observatory, located at sea-level near the city of Cambridge, UK. The interferometer has 5 apertures, referred to as "telescopes" in the COAST parlance, and these are laid out in a Y-shaped array as in figure 1.6, with a baseline range of ~ 2 m to ~ 60 m.

Each telescope is an afocal horizontal-pointing design with an aperture of 0.4 m. The telescope is illuminated by starlight using a flat mirror (siderostat) which rotates to track the science target. This scheme makes each telescope very stiff mechanically since the only moving part is the siderostat—so it is good for path length control. After being compressed by the primary and secondary mirrors, the light beam reflects from a mirror rotated by piezo-electric actuators. This mirror, called a "fast mirror", is a DM in the AO system, the autoguider, which corrects the linear phase aberrations across each aperture (commonly known as tip/tilt). The primary aim of the AO system is to reduce losses in the fringe visibilities, but other benefits include stopping beam "wander" and preventing vignetting.

The light from four (of the five) telescopes travels within aluminium pipes to the laboratory. The laboratory is in a bunker covered with soil which thermally insulates the interior, and it has a diurnal temperature range of less than 1°C. The layout of the optics and instruments within

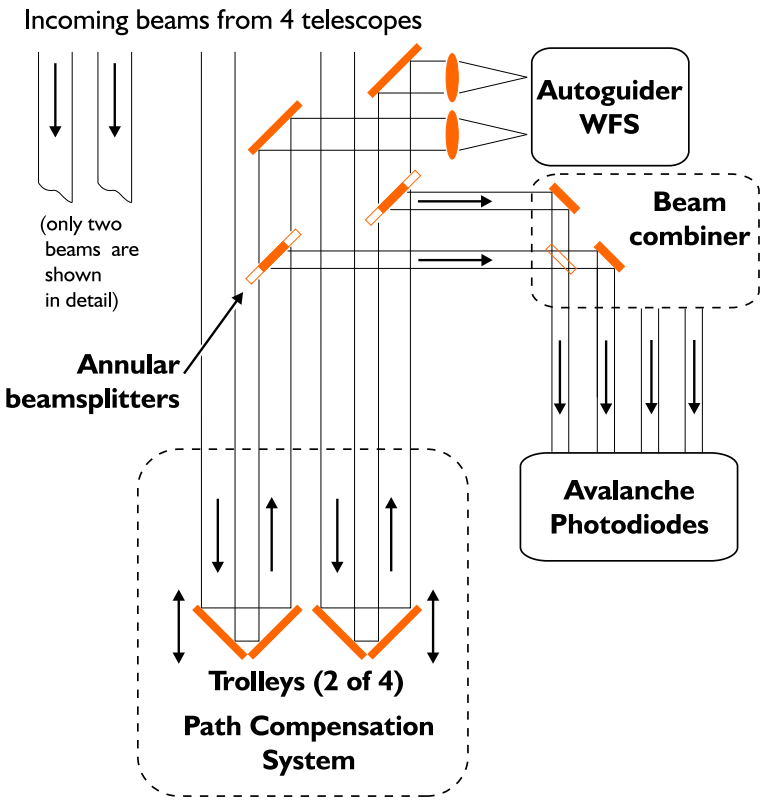

Figure 1.7.: Overview of the optical laboratory in COAST, the labels describe systems which are discussed in the text.

the laboratory are shown in figure 1.7. The path compensation system is the first optical system encountered by the incoming light beams and consists of four trolleys on rails. A single trolley consists of a "roof mirror" arrangement which retro-reflects the light, and by moving to different positions can introduce arbitrary path delay, in order to maintain equal path lengths between beams as the target moves across the sky. The trolleys also incorporate a voice-coil actuator on the cat's eye to allow small changes in path delay without needing to physically move the entire trolley.

The beams leave the path compensation system and are intercepted by spot mirror beamsplitters. The outer annulus of each beam is sent to the autoguider WFS, whose measurements control the fast mirrors in each telescope. The inner part of the beam is instead reflected into the beam combiner which has four inputs and four outputs. Each input corresponds to one beam entering the laboratory and each output contains 1/4 of of each input beam. The beam combiner consists of a set of 50% intensity beam splitters and these interfere every input beam together. The fringes from the interfered (output) beams are then measured using four APDs (Avalanche Photo Diodes), each of which uses a narrow bandwidth filter. Alternatively, one APD can be replaced with a spectrometer (39) which measures dispersed fringes over a large bandwidth.

1.3. Combining Adaptive Optics and Stellar Interferometry

During the late 1980's and early 1990's a number of prototype stellar interferometers were constructed (23)(48)(11)(1)(72)(15). They all used small apertures ($\sim 25\,\mathrm{cm}$, a few times the coherence

scale) and were designed to develop the technologies of long-baseline interferometry. Successes from these instruments included the first model-independent images from an optical interferometer (4) and measurements of many stellar diameters and binary separations.

Tango and Twiss (77) showed that even for apertures sizes equivalent to the coherence length, a limited form (low order) of AO called tip/tilt correction is required to prevent excessive losses in the visibility. In a single aperture (monolithic) telescope the tip/tilt aberrations cause gross image movement, and their correction is often incorporated into the guiding system which tracks the sky rotation. All modern long-baseline interferometers therefore incorporate low order AO in their guider design, as does COAST, although it is not often explicitly referred to as AO.

The limitations of current stellar interferometers is their low sensitivity, and the next generation will account for this by using larger apertures (13)(35). The phase aberrations will then become poorly approximated by linear functions and higher spatial frequency (higher order) AO systems will need to be considered. Accordingly, this dissertation will first examine the problem that seeing causes in stellar interferometers. The application of Adaptive Optics in Stellar Interferometry will then consider these results to achieve the final goal: reduced losses in fringe visibility and increased sensitivity.

- **Chapter 2:** This chapter begins with a discussion of phase aberrations; their origin and effect on image resolution and fringe visibility. These are used to quantify and compare performance metrics in AO and interferometry, and the specific benefits of AO for interferometry and its method of implementation are used to highlight areas of research that are discussed in other chapters.

- **Chapter 3:** A solution to the problem of making high sensitivity wavefront measurements is presented in this chapter. Starting with existing WFSs used in interferometer AO systems, the methods of measuring high order aberrations are considered. A new WFS method, Diffractive Phase Sensing, is presented and an implementation is described in the context of a specific WFS design: the Nine Element Sensor (NES).

- **Chapter 4:** This chapter concerns numerical simulations of the NES to evaluate its performance in an AO system. Comparisons are made with two existing WFS designs, one commonly used in astronomical AO and the other in use within current interferometer AO. The conclusions drawn specify the observation regimes for which each of the three WFS designs is most appropriate.

- **Chapter 5:** The design and construction of a NES prototype is discussed in this chapter. The prototype WFS is first tested in the laboratory, and its novel optic and CCD detector operation were analysed prior to use. The prototype was then used to make measurements of defocus phase aberrations at COAST, and results from these observations are presented and discussed to understand their implication.

- **Chapter 6:** The final chapter considers the existing AO system at COAST—the autoguider—and its measurements of tip/tilt aberrations. The aim and method used to parameterise the atmospheric turbulence is detailed, and the results are verified with measurements from a

DIMM and with fringe visibilities. Using the autoguider, the statistics of the seeing at the COAST site is presented from a year long dataset .

- **Chapter 7:** The conclusion of the thesis considers the success and failures of implementing a WFS designed specifically for interferometer AO systems. Suggestions for improvements and directions of future work in the area are given. The shortcomings found in the current AO system at COAST are reviewed, and future extensions to correcting visibilities for seeing effects, using the autoguider, are given.

2. Adaptive Optics in Interferometry

2.1. Introduction

To design an AO system for a telescope, the effect that the seeing has must be first understood. Generically, the seeing causes a loss in SNR, either of the image in a monolithic (single aperture) telescope or the fringe visibility in an interferometer. This reduces both the signal and the measurement SNR, and correcting the phase aberrations will always reduce these losses. An AO system will, however, have its own measurement SNR and can only operate with a suitably bright guide star, so if this limiting magnitude does not allow for observations of fainter targets then using AO will lead to no increase in the telescope's limiting sensitivity. As will be explained, an interferometer does not require an AO system to gain higher resolution. The question that will therefore be studied in this chapter is how to build an AO system for an interferometer that will allow the observation of fainter targets with larger apertures. Understanding the noise of the WFS measurements and what level of correction maximises the limiting magnitude of the entire interferometer array will provide the answer.

The chapter begins with a review of turbulence-induced phase aberrations and describes their impact on both the image creation process, using one aperture, and interferometer fringes, using a pair of apertures. The conclusions drawn from the comparison will explain why an AO system in an interferometer should differ from conventional AO in order to offer useful correction capabilities. In the context of these conclusions the analysis of the currently available AO technologies will highlight the particular research needs for interferometer AO.

2.2. Origin of atmospheric phase aberrations

The cause of phase aberrations from the atmosphere is the variation in refractive index between different vertical paths. A description of how the wavefront is distorted upon reaching the ground will provide the parameters used to describe the effects of seeing. Using the wave description of electromagnetic radiation, the scalar field of light, Ψ, at the top of the atmosphere can be written as,

$$\Psi(\mathbf{r}, t) = A \exp\left[i\mathbf{k}.\mathbf{r} + i2\pi\nu t\right], \tag{2.1}$$

where $\mathbf{r}$ represents any point in space. The amplitude, A, is a scalar function and the argument of Ψ is referred to as the phase. The variables $\mathbf{k}$ and ν have their usual meaning, wavevector and frequency respectively. The phase is constant in a plane perpendicular to $\mathbf{r}$ so the wavefront is flat. The seeing causes variations in the field at ground level, and two consequences are observable. The

phase is no longer constant in a plane, which leads to the distorted wavefront, and diffraction can introduce amplitude changes. Writing the phase, in the plane perpendicular to $\mathbf{k}$, as a complex quantity,

$$\Delta(\mathbf{r}, t) = \Phi(\mathbf{r}, t) + i\Lambda(\mathbf{r}, t), \tag{2.2}$$

the field at the ground is

$$\Psi(\mathbf{r}, t) = A \exp\left[i\mathbf{k}.\mathbf{r} + i2\pi\nu t + i\Delta(\mathbf{r}, t)\right]. \tag{2.3}$$

The quantities in Δ are the (real) phase, Φ, and the log amplitude, Λ. Variations in the latter causes scintillation, which is the mechanism behind the twinkling of stars.

2.2.1. Theory of Phase from Atmospheric Turbulence

The changes in refractive index of the atmosphere are a random and homogeneous process in space and time. The proper description is therefore statistical and uses ensemble averages and variances.

The dynamics of the turbulence which causes seeing was first described by Kolmogorov (36). The "Kolmogorov turbulence model" proposed that energy is input into the atmosphere at large scales, the outer scale $\mathcal{L}_0$, by solar heating or other forcing mechanisms. This energy is then transferred without loss to smaller scale turbulent cells, until the viscous friction between molecules causes the energy to dissipate. The scale at which the dissipation occurs is the inner scale, l_0 , and has a length of the order of millimetres (63). The value of $\mathcal{L}_0$ is more contentious and estimates range from 2 m (52) to over 1 km (10). For the purpose of this dissertation, it is assumed that $\mathcal{L}_0$ is effectively infinity as the relevant length scales for AO under consideration here are ~ 1 m.

The effects of turbulence with an infinite outer scale on optical propagation were developed by Tatarski (78). The complex phase variations in light received at a plane are zero mean Gaussian distributed and isotropic. As a result, the description of phase aberrations uses the concept of structure functions; the second order structure function, $D(\chi; \mathbf{r})$, is defined as,

$$D(\chi; \mathbf{r}) = \left\langle \left[\chi(\mathbf{r}') - \chi(\mathbf{r}' + \mathbf{r})\right]^2 \right\rangle. \tag{2.4}$$

This quantity implicitly ignores the mean of χ and instead emphasises its fluctuations. The structure function of the complex phase, Δ, can be separated into the structure functions of the phase and log amplitude, as these are independent,

$$D(\Delta; \mathbf{r}) = D(\Phi; \mathbf{r}) + iD(\Lambda; \mathbf{r}), \tag{2.5}$$

noting that $\mathbf{r}$ is now considered to be a vector perpendicular to $\mathbf{k}$. The amplitude fluctuations at the surface are often small (63) and can be neglected, $D(\Lambda; \mathbf{r}) \simeq 0$. The Kolmogorov velocity structure function, defined in the inertial subrange $l_0 \leq r \leq \mathcal{L}_0$, is of the form $D(v; \mathbf{r}) = C_v^2 r^{2/3}$. The work of Tatarski showed that the refractive index changes have the same structure function,

$$D(n; \mathbf{r}) = r^{2/3} \int dh\, C_n^2(h). \tag{2.6}$$

The integral of the variable C_n, which varies with altitude h, represents the total refractive index change for a vertically propagating light ray. The integral can usually be represented by an average value, obtained from site measurements using generalised SCIDAR (86) for example.

2.2.2. Spatial Statistics of the Phase

The structure function of the phase, Φ, was calculated by Fried (20),

$$D\left(\Phi;\mathbf{r}\right)=\left\langle\left[\Phi(\mathbf{r}')-\Phi(\mathbf{r}'+\mathbf{r})\right]^2\right\rangle=6.88\left(r/r_0\right)^{5/3}, \tag{2.7}$$

where

$$r_0=\left[0.423k^2(\sec\theta_z)\int dh\; C_n(h)\right]^{-3/5}.$$

The variable r_0, known as the coherence length, is the spatial scale of the seeing effects. Pointing a telescope at an angle away from the zenith, θ_z, involves observing through more of the atmosphere and r_0 becomes proportional to $\cos^{3/5}\theta_z$. The refractive index changes cause chromatic phase shifts so r_0 is also proportional to $\lambda^{6/5}$—its values are given for a canonical wavelength of 500 nm unless explicitly specified otherwise. There are several physical interpretations of r_0, including the diameter of a circular aperture for which the phase of light has a variance, σ_Φ^2, of approximately 1 rad rms (53). In general, for a telescope of diameter D, $\sigma_\Phi^2=1.03\left(D/r_0\right)^{5/3}$. Typical values of r_0 range from 10 cm (poor seeing, 1 arcsec) to 20 cm (good seeing, 0.5 arcsec), and also varies in time at any one location.

Another statistical description of the turbulent-origin phase is the power spectrum of the phase. Using the structure function, the equivalent Weiner spectrum can be written as

$$W_\Phi\left(\mathbf{f}\right)=0.033C_n^2\,\mathbf{f}^{-11/3}, \tag{2.8}$$

where $\mathbf{f}$ is the spatial frequency. Low spatial frequencies therefore have more power than high frequencies, so the bulk of phase aberrations are also at low spatial frequencies. The pole at the origin implies that the mean phase has infinite variance, but this it occurs at a scale larger than $\mathcal{L}_0$ and can therefore be ignored.

2.2.3. Temporal Statistics of the Phase

The coherence length, r_0, gives information on the spatial scale of the phase aberrations, and it is useful to define a temporal analogue, t_0. The temporal structure function of the phase is,

$$D\left(\Phi;t\right)=\left(t/t_0\right)^{5/3}. \tag{2.9}$$

The physical interpretation of t_0 includes the effects of turbulence evolution at any point and the bulk motion of the same turbulence relative to the ground. A simple model of the time evolution of phase aberrations, Taylor's Hypothesis (79), assumes that there is no intrinsic evolution of the turbulence and the only effect is that the air moves at a velocity $\mathbf{v}$. The seeing effects can then be described by using a (thin and static) phase screen, which imparts phase aberrations for light

passing through it, that moves relative to the ground. A screen is fully specified by r_0 and the "wind" velocity $\mathbf{v}$. In this model, t_0 is the time during which the phase, for light arriving at a fixed point on the ground, changes by one rad rms (78),

$$t_0 = 0.314\, r_0/v.$$

Measured values for t_0, at 500 nm, are of the order of a few milliseconds. Neglecting turbulence evolution means that this model becomes a poorer approximation as the aperture of a telescope increases. Although there is evidence that the turbulence occurs in well defined thin layers, in reality there may be several phase screens moving at different relative velocities. Finally, values derived for $\mathbf{v}$, when one phase screen models the aberrations, often do not relate to physical wind speeds (74).

2.2.4. Modal Representation of Phase

The phase aberrations at ground level can be represented as the sum of functions—the modes of a basis—which are orthogonal when integrated over a circular aperture. A commonly used basis is the set of complete Zernike polynomials. The indexing of Zernike polynomials used here is after Noll (53) and a summary of the properties is given in appendix A. As this basis is complete, the aperture phase aberrations can be written as a decomposition,

$$\Phi(r,\theta) = \sum_{i=1}^{\infty} a_i\, Z_i(r,\theta),$$

where r is a normalised radius, so that at the aperture diameter $r = 1$. The mode amplitudes, a_i, can be found from the orthogonality relationship,

$$a_i = \int_{\theta=0}^{2\pi} \int_{r=0}^{1} dr\, d\theta\; \Phi(r,\theta) Z_i(r,\theta).$$

A convenient property of Zernike polynomials is that the first few also represent common optical (Seidel) aberrations and are referred to by their Seidel-equivalent names. The Zernike polynomials are also indexed by radial and angular indices and the relationship of the indices and the Seidel-equivalent naming is summarised in table 2.1.

2.3. Phase Aberrations in Imaging and Interferometry

To understand the benefits of AO correction, the effects of phase aberrations on the image creation process in a single aperture (monolithic) telescope and the analogous fringe creation in an interferometer are described. The Strehl Ratio, $\mathcal{S}$, measures the degradation in the image and the analogue for the loss of fringe contrast is the Visibility Coherence, η. The analytic expressions of these quantities are related (77) and numerical simulations have been confirmed that Strehl ratio can be used as a proxy for visibility coherence. The improvement in Strehl ratio for one aperture is the performance metric for AO, but can also be used to infer the increase in coherence without explicitly needing to model an interferometer.

Radial index	**Azimuthal index**, m			
	0	1	2	3
0	Z_1 piston			
1		Z_2, tip Z_3, tilt		
2	Z_4 defocus		Z_5, astig. x Z_6, astig. y	
3		Z_7, coma y Z_8, coma x		Z_9, trefoil x Z_{10}, trefoil y
4	Z_{11} spherical aberration			

Table 2.1.: The first 11 Zernike polynomials with radial index, n, and azimuthal index, m.

The Strehl Ratio quantifies the coherence of light across the telescope aperture. With no phase errors, there is complete constructive interference at the centre of a focused image but phase aberrations reduce the coherence and therefore the on-axis intensity. The intensity when aberrations are present, I_Φ, divided by the intensity without aberrations, I_0, is one definition of $\mathcal{S}$,

$$\mathcal{S} = \frac{I_\Phi}{I_0} \leq 1. \tag{2.10}$$

In the case of a stellar interferometer, phase aberrations affect the fringe contrast. Michelson's original definition of visibility, $\mathcal{V}^2$, is the normalised difference in the maximum and minimum intensity of the fringes,

$$\mathcal{V}^2 = \frac{I_{max} - I_{min}}{I_{max} + I_{min}}. \tag{2.11}$$

The phase aberrations across each aperture of the interferometer reduces the coherence of light between the apertures.

This causes the fringe visibility to be reduced and η can be defined similarly to $\mathcal{S}$ as the ratio of visibilities with and without phase aberrations,

$$\eta^2 = \frac{\mathcal{V}_\Phi^2}{\mathcal{V}_0^2} \leq 1. \tag{2.12}$$

2.3.1. Theory of Imaging using One Aperture

Consider the monolithic telescope in figure 2.1, which has an aperture diameter of D and a focal length f. The complex amplitude of light at the aperture is described by $A(x, y) = P(x, y) \exp[i\Phi(x, y)]$, where Φ is the phase. The pupil transmission function, $P(x, y)$, is equal to one within the aperture and zero outside.

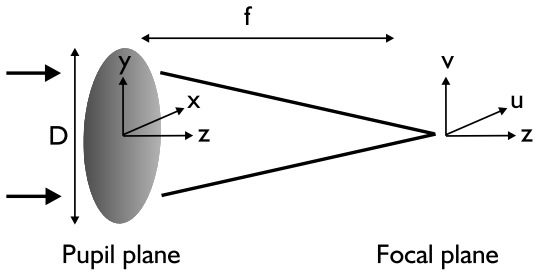

Figure 2.1.: Schematic diagram of a single aperture (monolithic) telescope. The aperture has a diameter of D and the optics a focal length of f. The Cartesian coordinates are (x, y) in the pupil plane and (u, v) in the focal plane.

The image at the focal plane depends on the amplitude at the pupil, which in turn is a function of the sky intensity distribution. Disregarding phase aberrations, for the case of an on-axis monochromatic point source, the image intensity is a result of Fraunhofer diffraction,

$$I_\delta(u,v) = |A(u,v)|^2 = \frac{i}{f\lambda}\left|\int\int dx\,dy\; P(x,y)\exp\left[-i\frac{k}{f}(xu+yv)\right]\right|^2. \tag{2.13}$$

The image, I_δ, is effectively the modulus squared of a scaled Fourier transform of P. This is called the Point Spread Function (PSF), and it is an Airy disc for a diffraction-limited circular aperture. For a more general sky intensity distribution F, the focal plane image is,

$$I(u,v;\lambda) = F(f\alpha, f\beta;\lambda) * I_\delta(u,v), \tag{2.14}$$

where (α, β) are the angular coordinates on the sky. When phase aberrations are present, the PSF varies with the pointing angle on the sky and in time,

$$I_\delta(\alpha,\beta;t) \propto \left|\int\int dx\,dy\; P(x,y)\exp\left[-i\frac{k}{f}(xu+yv)+i\Phi(x,y;\alpha,\beta;t)\right]\right|^2. \tag{2.15}$$

For the remainder of this section, these variations are ignored but will be important later. The image resolution can be defined using the Fourier transform of the PSF; this latter quantity is called the Optical Transfer Function (OTF),

$$\begin{aligned}
\mathcal{O}(\omega,\eta;\lambda) &= \int\int du\,dv\; I(u,v;\lambda)\exp\left[i2\pi(u\omega+v\eta)\right] \Big/ \int\int du\,dv\; I(u,v;\lambda) \\
&= \int\int dx\,dy\; A(x,y)A^*\left(x+\frac{\omega f}{k}, y+\frac{\eta f}{k}\right) \Big/ \int\int dx\,dy\;\; |A(x,y)|^2 \\
&= \mathcal{M}(\omega,\eta)\exp\left[i\mathcal{P}(\omega,\eta)\right].
\end{aligned} \tag{2.16}$$

The second line shows that the OTF is equal to a normalised auto-correlation of the pupil amplitude, and is a consequence of the Wiener-Khinchin theorem. The amplitude of the OTF, $\mathcal{M}$, is the Modulation Transfer Function (MTF) and, using Schwarz's theorem, the integral of the

MTF is maximised when there are no phase aberrations. The integral of the MTF is therefore one measure of image resolution, and is called the Resolving Power,

$$\mathcal{R} = \int d\mathbf{f}\ \mathcal{M}(\mathbf{f}). \tag{2.17}$$

For an aberration-free circular aperture the resolving power is $\mathcal{R} = (\pi/4)(D/\lambda)^2$ c.f., the equivalent Raleigh Criterion resolution is $1.22\,(\lambda/D)$. The expression for imaging through the atmosphere requires that terms for phase aberrations are included in the MTF,

$$\begin{aligned}\mathcal{M}'(\omega,\eta) = \int\int dx\,dy\ \Big\{ P(x,y)P\left(x+\frac{\omega f}{k}, y+\frac{\eta f}{k}\right) \\ \times \left\langle \exp\left[i\left(\Phi(x,y) - \Phi\left(x+\tfrac{\omega f}{k}, y+\tfrac{\eta f}{k}\right)\right)\right]\right\rangle \Big\} \\ \Big/ \int\int dx\,dy\ |P(x,y)|^2\ .\end{aligned} \tag{2.18}$$

The average incorporating the phase can be written as the Phase Transfer Function (PTF) $\mathcal{B}$ and the total MTF is a product of the unaberrated MTF and the average PTF: $\mathcal{M}' = \mathcal{M}\langle\mathcal{B}\rangle$. Over timescales much larger than t_0, $\langle\mathcal{B}\rangle$ can be written in terms of the phase structure function (2.7),

$$\begin{aligned}\langle\mathcal{B}(\omega,\eta)\rangle &= \left\langle \exp\left[i\left(\Phi(x,y) - \Phi\left(x+\frac{\omega f}{k}, y+\frac{\eta f}{k}\right)\right)\right]\right\rangle \\ &= \exp\left[-\frac{1}{2} D\left(\Phi; \sqrt{\left(\frac{\omega f}{k}\right)^2 + \left(\frac{\eta f}{k}\right)^2}\right)\right].\end{aligned} \tag{2.19}$$

The PTF allows for an alternative definition of r_0,

$$\langle\mathcal{R}\rangle = \frac{\pi}{4}\left(\frac{r_0}{\lambda}\right)^2 = \int d\nu\ \langle\mathcal{B}(\nu)\rangle. \tag{2.20}$$

Returning to the resolving power of a telescope imaging through the atmosphere, the aberrated MTF is suppressed at high frequencies by the exponential form of $\langle\mathcal{B}\rangle$ and at low frequencies $\mathcal{M} \simeq 1$. The joint MTF can therefore be approximated as $\langle\mathcal{M}'\rangle \simeq \langle\mathcal{B}\rangle$. The resolving power will therefore be equal to using a telescope of diameter r_0 which results in the seeing resolution-limit for any diameter telescope with $D > r_0$.

The time averaging above is equivalent to considering long exposures. For exposures shorter than t_0, the turbulence evolves slowly compared to the integration time and $\mathcal{R}$ is $(\pi/4)(3.5\,r_0/\lambda)^2$. This result will be considered when discussing the phase aberration effects in an interferometer.

The resolving power is relatively insensitive to small phase aberrations, but this is where the Strehl ratio is most sensitive. For example, a r_0-sized aperture has a phase variance of 1.03 rad rms and suffers no loss in resolving power but the Strehl ratio is reduced to ~ 0.37. An alternative definition of $\mathcal{S}$ uses the MTFs defined above,

$$\mathcal{S} = \int d\mathbf{f}\ \mathcal{M}(\mathbf{f})\,\mathcal{B}(\mathbf{f}) \Big/ \int d\mathbf{f}\ \mathcal{M}(\mathbf{f})\ . \tag{2.21}$$

Neither the resolving power nor the Strehl ratio depends on the mean phase across the aperture, $\bar{\Phi}$, because it is irrelevant for imaging.

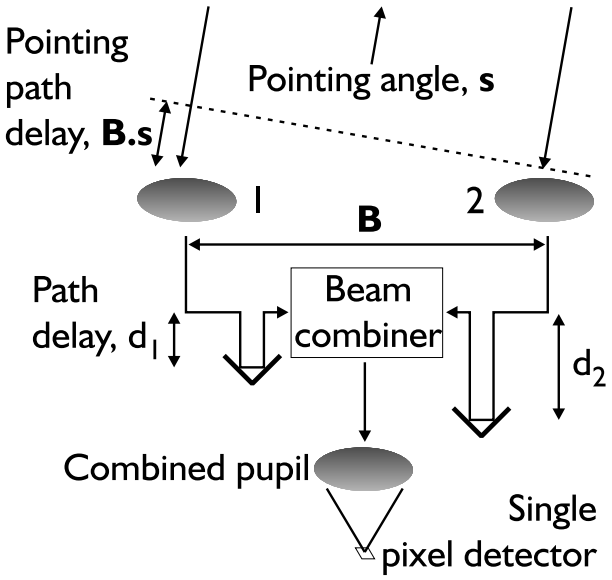

Figure 2.2.: Schematic diagram of a two aperture Michelson interferometer. The source is at an angle **s** and the *added* path delays for each beam are d_1 and d_2.

2.3.2. Theory of Stellar Interferometry with Two Apertures

The interferometer shown in figure 2.2, with two apertures separated by a baseline **B**, will be used to describe the effects of phase aberrations on interference fringes. The variables associated with each aperture are identified by their subscript, either 1 or 2. The complex amplitude at the two apertures is $A_1(x_1, y_1)$ and $A_2(x_2, y_2)$ and the only difference is in the phase aberrations,

$$\begin{aligned} A_1(x_1, y_1) &= P(x_1, y_1; \lambda) \exp\left[\Phi_1(x_1, y_1)\right] \\ A_2(x_2, y_2) &= P(x_2, y_2; \lambda) \exp\left[\Phi_2(x_2, y_2)\right]. \end{aligned} \tag{2.22}$$

The pupils of each aperture are interfered at the beam combiner, and the interfered pupil is $\Psi(\bar{x}, \bar{y})$, where the coordinates with bars indicate a point relative to each aperture's centre. The path compensation system changes each beam's propagation distance by adding or removing path delay, d_1 and d_2. The interfered pupil is then imaged onto a single pixel detector, as it is the image's on-axis intensity that is of interest. For the case of a monochromatic point source, with brightness F_0 and no phase aberrations present, the intensity is,

$$I_\Psi(\Delta d) = F_0 \left(2 + 2\cos\left[k\left(B_\alpha \alpha_0 + B_\beta \beta_0 + \Delta d\right)\right]\right) \left| \int\int d\bar{x}\, d\bar{y}\; P(\bar{x}, \bar{y}) \right|^2. \tag{2.23}$$

This is a sinusoidal function of the path delay difference, $\Delta d = d_1 - d_2$, and phase from the projected baseline, **B**.**s**, if the pointing angle from the zenith is $\mathsf{s} = (\alpha_0, \beta_0)$. To obtain zero optical path difference (OPD) between the beams, the path delay added to *one* beam is $-$**B**.**s**. To record the fringe pattern ("fringe packet") a small modulation of the OPD is made to "scan" the fringes. At COAST, this is done by modulating the roof mirrors with the voice coil actuator on each trolley.

Note that altering the aperture diameter affects only the value of the integral term, but this changes only the intensity of the fringes and not their contrast. Hence the theoretical resolving power of an interferometer is unaffected by aperture shape or size.

It was mentioned above that the variation with sky angle of the phase aberrations does not affect an interferometer. This can be be explained by considering the measurement of a chromatic point source with a finite bandpass. Assume that this bandpass is rectangular, $\lambda_0 - \Delta\lambda/2 \leq \lambda \leq \lambda_0 + \Delta\lambda/2$ and the intensity of the source is equal at all wavelengths. The interfered intensity is now

$$\begin{aligned} I_\Psi(\Delta d, \Delta\lambda, \lambda_0) &= \int_{\lambda_0-\Delta\lambda/2}^{\lambda_0+\Delta\lambda/2} d\lambda \; I_\Psi(\Delta d) \\ &= F_{\lambda_0}\Delta\lambda \left(2 + 2\,\mathrm{sinc}\left[2\pi\Delta\lambda/\lambda_0^2 \; (B_\alpha\alpha_0 + B_\beta\beta_0 + \Delta d)\right]\right. \\ &\quad \left. \times \cos\left[2\pi/\lambda_0 \; (B_\alpha\alpha_0 + B_\beta\beta_0 + \Delta d)\right]\right) \left| \int\int d\bar{x}\, d\bar{y} \; P(\bar{x}, \bar{y}) \right|^2 . \end{aligned} \tag{2.24}$$

The fringes are modulated by a sinc function and examples for 10% and 25% bandpasses are shown in figure 2.3. The result of using chromatic light is a finite coherence length, $\lambda_0^2/\Delta\lambda$, and it reduces the visibility contrast except at zero OPD. If the path delays are not continuously adjusted to accommodate the rotation of the sky, the fringe contrast is rapidly lost. Equivalently, if the apertures point at angle $\mathbf{s}$ and the target is at $\mathbf{s} + \delta\mathbf{s}$ then the fringe packet will no longer be measurable when

$$\delta\mathbf{s}.\mathbf{B} \simeq \lambda_0^2/\Delta\lambda.$$

This defines the field of view (FOV) for an interferometer, and is smaller with larger bandpasses and larger baselines. For the case of $\Delta\lambda/\lambda_0 = 0.1$ and $\lambda_0 = 0.5\,\mu$m, the FOV is $\theta_B = \lambda_0^2/(B\,\Delta\lambda) \simeq 1/B$ arcsec. The isoplanatic patch is the angular diameter within which the phase aberrations for different angular directions are correlated to ~ 1 rad rms and a typical diameter is ~ 1 arcsec. As this is much larger than the FOV, the angular variation in Φ_1 and Φ_2 can be neglected.

The fringe packet from an extended source will now be considered, assuming it is monochromatic and the brightness is $F(\mathbf{s};\lambda)$. The phase aberrations are also included and, analogously to path delay, they are written as the difference in phase between the apertures, $\Delta\Phi = \Phi_1 - \Phi_2$. The complex visibility, $\mathbf{V}$, of the source is defined as,

$$\mathbf{V}(\mathbf{B};\lambda) = \int d\mathbf{s} \; \sqrt{F(\mathbf{s};\lambda)} \exp\left[i\frac{2\pi}{\lambda}\mathbf{B}.\mathbf{s}\right]. \tag{2.25}$$

This is the Fourier transform of the sky intensity at a spatial frequency equal to the baseline projection, $\mathbf{B}.\mathbf{s}$, and is an example of the van Cittert-Zernike theorem. It represents the information that a stellar interferometer measures, as is now shown. Using the expressions for complex visibility and the difference in phase aberrations, the interfered intensity of the fringe packet is,

$$I_\Psi(\Delta d) \;=\; 2I_\Sigma + 2I_\Sigma \,\mathrm{Re}\left\{\mathbf{V}(\mathbf{B}) \exp\left[ik\Delta d\right] \int\int d\bar{x}\, d\bar{y} \; P \exp\left[i\,\Delta\Phi\,\right]\right\}. \tag{2.26}$$

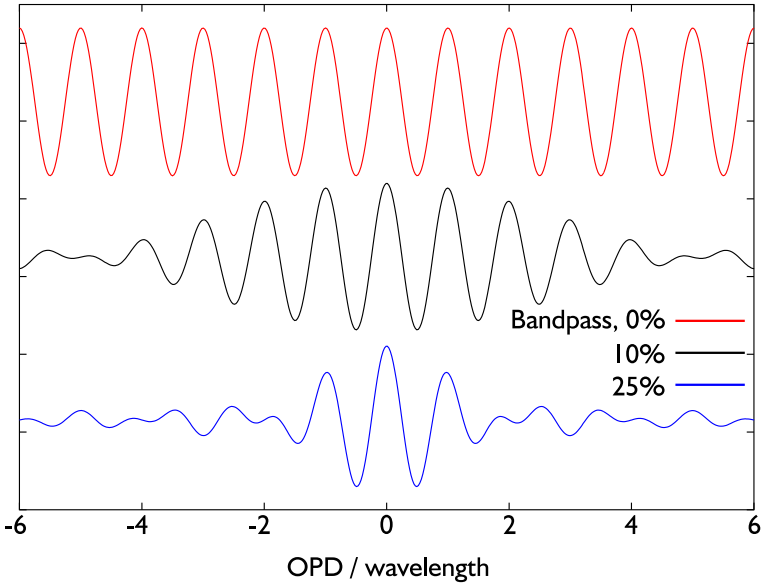

Figure 2.3.: Fringes with 0% (monochromatic), 10%, and 25% bandpasses.

The first term, I_Σ, is the mean intensity in the fringe packet and is a function of the aperture sizes, as before. The terms in the bracket form the oscillatory part of the fringe packet: the real part of the complex visibility multiplied by a function of $\Delta\Phi$. The aberrations can therefore be seen to have two effects. First, the amplitude of $\mathbf{V}$ is lowered unless $\Delta\Phi$ is a constant (i.e., the fringe contrast is reduced), and secondly the phase of $\mathbf{V}$ is changed by the difference of mean phase. Both these effects corrupt the measured value of $\mathbf{V}$, and often the squared modulus of the complex visibility is measured (7),

$$\left\langle |\mathbf{V}|^2 \right\rangle = \left\langle \left| \int\int d\bar{x}\, d\bar{y}\; A_1\, A_2^* \right|^2 \right\rangle \Big/ \left\langle \int\int d\bar{x}\, d\bar{y}\; |A_1|^2 + \int\int d\bar{x}\, d\bar{y}\; |A_2|^2 \right\rangle . \tag{2.27}$$

The visibility, unless otherwise stated, can now be taken to mean the amplitude squared of the complex visibility i.e., $V \equiv |\mathbf{V}|^2$.

Measuring the Phase of the Complex Visibility

The measured intensity in an interferometer is the real part of the (corrupted) complex visibility for a given path delay, and the visibility phase can be inferred from the fringes by modulating the OPD. The measured visibility phase will, however, include the effects of the mean phase aberrations. An alternative phase quantity, the closure phase, can be constructed from the sum of three (measured) visibility phases from different baselines and is immune to aberrations.

Let the intrinsic visibility phase, from the baseline of apertures i and j, be $\Phi_{V;ij}$ and the mean phase aberration at an aperture be $\Phi_{A;i}$. The measured complex visibility, $\mathbf{V}'$, will have a phase,

$$\Phi_{V';\Sigma;ij} = \Phi_{V;ij} + (\Phi_{A;i} - \Phi_{A;j})\,.$$

The measured visibility phase for the three baselines, and their sum, is,

$$\begin{aligned}
\Phi_{V';12} &= \Phi_{V;12} + (\Phi_{A;1} - \Phi_{A;2}) \\
\Phi_{V';23} &= \Phi_{V;23} + (\Phi_{A;2} - \Phi_{A;3}) \\
\Phi_{V';31} &= \Phi_{V;31} + (\Phi_{A;3} - \Phi_{A;1}) \\
\Sigma\Phi_{V';ij} &= \Phi_{V;12} + \Phi_{V;23} + \Phi_{V;31} \equiv \Sigma\Phi_{V;ij},
\end{aligned} \tag{2.28}$$

where the final quantity is the closure phase which is independent of $\Phi_{A;i}$. For N apertures, where $N \geq 3$, the number of independent closure phases, ${}^{N-1}C_2$, is smaller than the number of visibilities, ${}^{N}C_2$. While with $N = 3$, only 33% of the phase information can be recovered, with $N = 6$ it is 66%, and $N = 12$ results in 83% of the visibility phase being measurable.

Measuring the Amplitude of the Complex Visibility

The loss in visibility amplitude will now be discussed. It is understood to mean the decrease in the squared modulus of the visibility from its intrinsic value. The visibility coherence factor, using equation (2.27), is the ratio of the numerator when phase aberrations are present to numerator without aberrations,

$$\begin{aligned}
\left\langle |\eta|^2 \right\rangle &= \langle V' \rangle / \langle V \rangle \\
&= \left\langle \left| \int\int d\bar{x}\, d\bar{y}\; P(\bar{x},\bar{y})^2 \exp\left[\Delta\Phi(\bar{x},\bar{y})\right] \right|^2 \right\rangle \Big/ \left\langle \left| \int\int d\bar{x}\, d\bar{y}\; P(\bar{x},\bar{y})^2 \right|^2 \right\rangle .
\end{aligned} \tag{2.29}$$

The relationship to the Strehl ratio for one aperture is manifest if η is rewritten in terms of the MTFs of the apertures. If the baseline is longer than the outer scale, $B > L_0$, then the (uncorrected) phase aberrations at each aperture are uncorrelated and can be treated as independent. From this assumption, Tango & Twiss (77) derive the appropriate expression,

$$\left\langle |\eta|^2 \right\rangle = \int d\mathbf{f}\; \mathcal{M}(\mathbf{f})\, \langle \mathcal{B} \rangle^2(\mathbf{f}) \Big/ \int d\mathbf{f}\; \mathcal{M}(\mathbf{f}) \, . \tag{2.30}$$

Compared to equation (2.21), the difference is the square of $\langle \mathcal{B} \rangle$. For the case of two apertures with uncorrected phase aberrations, the ratio of $\mathcal{S}$ to η is shown in figure 2.4. The important result is that Strehl ratio is a monotonic function of visibility coherence (82) and therefore $\mathcal{S}$ will be used as a proxy for η.

Time variation in Mean Phase Aberration Difference

The atmosphere is weakly dispersive and consequently the phase aberrations it causes can be described as wavelength-independent path differences. The difference in the mean of the phase aberrations from the two apertures, $\Delta\overline{\Phi}(t,\lambda)$, can then be equivalently described as a change in the OPD. This component of the OPD, the atmospheric OPD, will change randomly in time and cause the fringe packet to move unpredictably about its nominal zero position. For short integrations of the intensity from the beam combiner, this means the fringes are shifted away from the expected position.

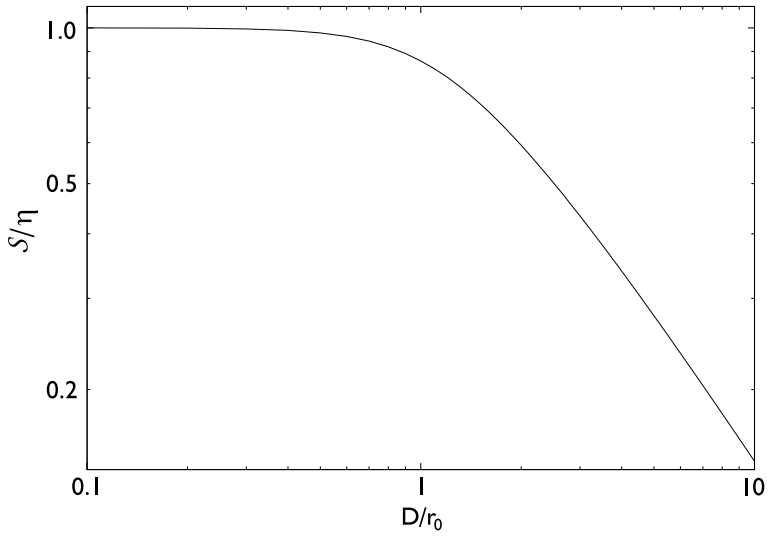

Figure 2.4.: The ratio of the Strehl ratio, $\mathcal{S}$, to the square root of the visibility coherence, η, with no AO.

The result for longer integrations, greater than a few t_0, is that the fringe pattern is "washed out" and the visibility cannot be measured. The baselines which are unaffected by the variation in atmospheric OPD have a mean phase shift of ≤ 1 rad rms and, from the structure function, have $B \leq r_0$. Therefore, coherent integration of the fringes is prevented for long baselines.

This effect is analogous to the linear phase aberrations (tip/tilt) causing movement of an image in a monolithic telescope; the presence of these aberrations does not change the MTF, and so does not cause a reduction in resolving power as long as they are constant. Short exposures for images therefore have higher resolution than long exposures because the the image motion is fixed during the integration time.

2.4. Application of AO to a Stellar Interferometer

At this stage, the statistics and effects on the visibility of phase aberrations have been described. The design of an AO system can now be explained in general terms, and in particular, the appropriate "order" to use can be justified. Order of AO can be defined using Zernike polynomials: an AO system which has order N senses and corrects the first N Zernike polynomials after piston.

2.4.1. Order of AO Correction for Different Aperture Sizes

The residual variance of phase after removing the first N Zernikes from atmospheric phase aberrations on a circular aperture has been calculated by Noll (53). The results are summarised in table 2.2 and the values should be scaled by $(D/r_0)^{5/3}$ to obtain the value in rad rms.

The residual variance decreases with increasing order, and the rate of decrease increases except for the removal of Zernike polynomials with equal radial index n. Therefore, an useful increment in AO order represents removing all Zernike polynomials up to an increased radial index. For $N > 10$, the residual also scales approximately as $N^{-\sqrt{3}/2}$, whereas it descends more steeply for the terms

up to $N = 9$. Low order AO is therefore defined as $N \leq 10$ and high order is $N > 10$—the division is for correcting up to radial index $n = 4$.

Z_1	Z_2	Z_3	Z_4	Z_5	Z_6	Z_7	Z_8	Z_9	Z_{10}
1.031	0.582	0.134	0.111	0.088	0.065	0.059	0.053	0.046	0.040

Table 2.2.: Residual phase variance after correcting up to and including the stated Zernike polynomial. The units are $(D/r_0)^{5/3}$ rad rms. The vertical bars separate Zernike polynomials with differing radial index.

The order which leaves ~ 1 rad rms phase variance (equivalent to an uncorrected r_0 sized aperture) is shown in table 2.3 for increasing D/r_0. The scaling of r_0 with wavelength means that $(D/r_0)^{5/3} \propto \lambda^{-2}$ so the order also decreases for longer wavelengths. The next generation of interferometers will have aperture diameters of 1–2 m, for example CHARA (45), VLTI auxiliary array (35), and MROI (13). These apertures are a factor of 3–10 smaller than current telescopes with AO systems, so the order for interferometer AO will automatically be smaller than that for current astronomical AO.

For the following calculations it will be assumed that $D = 1$ m and four estimates of N will be made: in the visible and near-IR and for the (observed) median and lower seeing limits at high altitudes. In 1 arc sec (poor) seeing, at 500 nm (V-band) $D/r_0 = 10$ and $N \geq 20$, whilst at 800 nm (I band), $D/r_0 = 6$ and $N \geq 8$. In (average) seeing of 0.5 arc sec, at 500 nm $D/r_0 = 5$ and $N \geq 5$ whilst at 800 nm $D/r_0 = 3$ and $N \geq 2$. It can therefore be concluded that low order AO is appropriate for 1 m apertures except in the visible with poor seeing.

2.4.2. Comparing the Single-Telescope PSF and Interferometer Fringes

The PSF and the modulated fringes are the equivalent quantities, respectively, in a monolithic telescope and an interferometer which are corrupted systematically by phase aberrations. In the presence of aberrations, the PSF has its resolution systematically lowered and similarly the visibility of the fringes is systematically lowered. In principle, a dataset which contains such systematic errors can be corrected by characterising the error and then accounting for it.

Aperture size / D/r_0	**Required order** / N	**Effective correction**
1	0	No AO
3	2	Tip/tilt correction
5	5	Correction to astigmatism
7	10	
9	18	
11	25	

Table 2.3.: The order of AO correction which leaves a residual phase variance of ~ 1 rad rms.

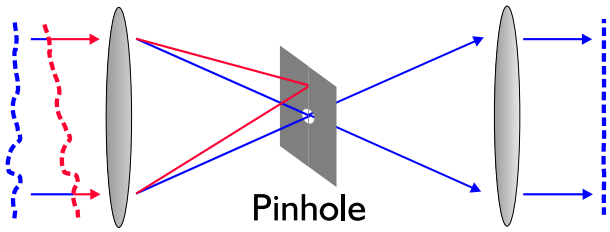

Figure 2.5.: A pinhole spatial filter. The blue wavefront, with high order aberrations, couples well into the pinhole. The red wavefront, with a low order tilt aberration, has a large loss of throughput.

The difficulty of attempting to characterise the PSF is that, at any one time, it varies over the position in the image and consists of spatial frequencies from zero (the total intensity) up to the diffraction-limit of the telescope. It is therefore difficult to make an average measurement unless it is from a long exposure, and in this case the loss in image resolution makes the calibration extremely challenging because of low SNR. In contrast, the fringes consist of one frequency (which is baseline dependent and therefore known) and are measured within the isoplanatic patch thus are not dependent on sky angle either.

The visibility coherence, η, then becomes only a function of the seeing in the direction of the target. The scheme currently used to characterise η is to make visibility measurements which are intrinsically known e.g., from a point source which is unresolved and therefore has a visibility of one. Such targets are known as calibrators and bracketing the observation of a general science target with those of a calibrator allows η to be estimated for the target.

This calibration approach is correction applied after the data has been collected and only reduces systematic errors. To improve the measurement SNR itself requires correction before the measurements are taken, and this can be either active, such as AO, or passive. An example of a passive method to remove phase aberrations in light is spatial filtering. The angular width of the PSF from an aperture in the interferometer, $\sim \lambda/D$, is larger than the interferometer FOV, $\lesssim \lambda/B$. One effect of phase aberrations is to cause the PSF to spread, and only light which is in phase interferes at the centre of the PSF. A spatial filter allows this "unaberrated" light to pass through it and rejects "aberrated" light. The filtered light has a smaller phase variance and therefore larger η when interfered. Spatial filtering cannot be used for a monolithic telescope because its FOV is larger than the PSF, and spatial filtering would therefore reduce the resolving power to the size of the filter.

Two types of spatial filter have been successfully used in interferometry: single mode optical fibres (70)(57) and pinholes (33)(74). The difficulty in using these devices is that the coupling of light into them is low because uncorrected phase aberrations shift the focused image around an area $\approx (D/r_0)^2$ larger than the filter input aperture, see figure 2.5. The aberrations which cause this reduction in throughput are the low order aberrations which pinholes do not effectively filter (34). Simulations have shown that a tip/tilt correction system is mandatory when using a pinhole to obtain an increase in the visibility SNR.

A numerical model of a two aperture interferometer (8), using a single mode fibre, has been used to investigate the improvement in visibility SNR when using AO. A single mode fibre is a perfect filter as the phase variance of the filtered (output) light is zero. The AO system modelled was also "perfect" in that it corrected N Zernike polynomials exactly, therefore the results quantify the increase in coupling efficiency assuming ideal correction. When $N = 5$, the peak SNR (without the fibre) occurs when D/r_0 is ~ 4 c.f., $D/r_0 = 5$ for the same N in the simple argument above. With the fibre in place and $N = 5$, the peak SNR is ~ 2 times bigger and $D/r_0 \simeq 7$ i.e. the diameter increase gives a flux increase of about two. Since an AO system would use less than 50% of the input flux, the realistic flux available for interferometric measurements increases.

2.4.3. Correction of Piston Aberrations

The time variation of the atmospheric OPD prevents long exposures of fringes, whilst for short exposures the path delay modulation (to scan the fringe packet) may not encompass the changing zero OPD position.

There are two approaches to compensate for the fringe "movement": the fringe packet can be tracked and the path delays adjusted to remove the atmospheric OPD, or the scan modulation amplitude can be increased to encompass the expected variance in atmospheric OPD. The latter, being a passive method, can operate at any light level but the amplitude needed may be much larger than the coherence length: for a baseline of 100 m at 800 nm and a bandpass of 10%, the atmospheric OPD has a standard deviation of $\sim 60\,\mu$m whereas the coherence length is $\sim 4\,\mu$m.

Actively tracking the fringe packet (essentially AO) can be divided into two forms: group delay estimation and phase delay estimation. The former maintains the zero OPD to within the coherence envelope length so the scanning is more efficient. Phase delay estimation tracks the position of the white light fringe, which is at the zero OPD position and so enables the fringes to be coherently integrated for longer than several t_0 (71)(38).

An increase in integration time by using fringe tracking within an interferometer is also the benefit offered by AO on monolithic telescopes: short exposures become highly correlated. As the SNR of phase delay measurements is proportional to V, using AO in an interferometer can potentially allow phase delay estimation with fainter targets. This last argument, however, is only true if the AO WFS also has sufficient sensitivity to operate with the science target.

2.4.4. Optimum Diameter of Interferometer Apertures

A simple method to reduce the decreases in visibility coherence is to stop down (reduce) the aperture diameters. The aberrations across smaller apertures have a smaller phase variance, but collect less light; at low light levels there is an optimum diameter which maximises the visibility SNR when using tip/tilt correction. At high light levels, the SNR is maximised when the phase variance is zero i.e., for point-like apertures.

The visibility SNR is maximised at low light levels with tip/tilt correction when $D/r_0 \sim 3$ (7). With higher order correction, the aperture diameter at which the SNR is maximised increases. Using AO correction of order $N = 5$, the optimum diameter is $D/r_0 \sim 4$. Higher order correction, $N = 9$, has a maximum SNR at $D/r_0 \sim 7$. These results also assume perfect AO correction without

spatial filtering, but larger apertures collect more light and increase the WFS SNR. Assuming seeing of 0.5 arc seconds, the optimal AO-corrected diameters (in the visible) are ~ 0.8 m and ~ 1.4 m for $N = 5$ and $N = 9$.

2.4.5. Benefits of AO in an Interferometer

Correcting the phase aberrations across the apertures of an interferometer produces several benefits, and these are outlined here,

- **Reduction in visibility loss**: This has been covered above, and for aperture sizes greater than r_0, the improvement in the visibility coherence factor can be significant. This is a primary motivation, and if the correction was very high order i.e. residual phase was $\ll 1$ rad rms then the remaining (slowly varying) errors could be characterised by observing calibrator targets once or twice per night. Very high order correction will be most realistically implemented using a spatial filter.

- **Longer piston coherence time**: The timescale for the piston aberration to change by 1 rad rms, analogous to t_0, increases with larger aperture diameters (74). If D is increased from 0.2 m to 1 m then the timescale, proportional to $D^{1/6}$, increases by $\sim 1/3$. This increase will reduce the maximum bandwidth needed for fringe tracking system.

- **Bootstrapping**: The SNR for fringe tracking is proportional to the visibility amplitude, V. Using a long baseline, the intrinsic visibility may be sufficiently small that its measurement is prevented using short integrations. However, it is precisely these fringes which benefit from the coherent integration available with fringe tracking. The OPD can, however, be inferred by measurements on (shorter) intermediate baselines. This is called Bootstrapping and AO allows for the bootstrapping measurements to made on longer baselines without a reduction in their SNR (65).

- **Reducing diffraction losses**: The propagation of light from each aperture to the beam combiner will result in diffraction losses. Phase aberrations increase these losses although the diffraction it can paradoxically increase the visibility SNR (30) by an effect analogous to spatial filtering. AO therefore increases the visibility SNR and reduces propagation losses in flux. Correction of aberrations becomes more important as baseline lengths increase, especially when fringe tracking is implemented.

The advantages offered by using AO on a monolithic telescope are higher sensitivity and improved image resolution. Interferometer AO instead only enhances the fringe measurement sensitivity.

2.4.6. Conclusion

A stellar interferometer uses several small apertures which are intrinsically suited to low order correction, whereas the larger apertures of monolithic telescopes require high order correction. The instruments in an interferometer use incoherent (short exposure) integration, in either the beam combiner (with passive OPD scanning) or in the fringe tracker (when using phase delay tracking.)

An interferometer AO system will therefore operate with instruments that require some minimum flux, so the AO WFS cannot therefore take increasingly large fractions of the guide star flux without rendering the other instruments inoperable. This is very different from monolithic telescope AO where all instruments can integrate. There is therefore a need to emphasise the wavefront measurement sensitivity over the requirement to correct high order aberrations.

2.5. Interferometer AO Design

In practice the interferometer AO system will often operate in the low light regime and its performance is then primarily limited by the wavefront measurement sensitivity. The optimum AO design will therefore aim to maximise the WFS sensitivity and minimise other sources of noise.

The sources of errors in an AO system can be divided into several areas; the measurement noise in the Wavefront Sensor (WFS), noise in the reconstructed wavefront, and errors in wavefront correction. The choice of detector for the WFS affects the measurement noise, but the noise propagation into the measured wavefront depends on the WFS method. The Control System (CS) is some form of signal processing device, typically a digital computer, and processes the inferred wavefront before transferring it to the Deformable Mirror (DM). This last step introduces negligible noise but can introduce time delays. The DM will not reconstruct the exact wavefront shape demanded, and all correction devices suffer from mechanical errors. The overall error is a contribution from the individual sub-systems, $\sigma^2_\Sigma = \sigma^2_{DM} + \sigma^2_{CS} + \sigma^2_{WFS}$.

2.5.1. Wavefront Sensor Analysis

Existing WFSs used in astronomy (27) fall into several method categories which naturally sort themselves with respect to the sensitivity. A measure for sensitivity is the SNR, and more precisely the limiting magnitude of a guide star for which the SNR is one. The AO system in which the following WFSs are analysed has an order of $N = 5$ for an aperture of 1 m. The recent availability of sub-electron read-noise CCD arrays (43) means that only photon noise need be considered.

- **Zonal sensors:** These wavefront sensors subdivide the pupil into independent sub-apertures i.e., zones. Each sub-aperture measures intensity which is only related to the phase aberrations for that sub-aperture. The Shack-Hartmann sensor consists of an array of lenslets which focus the sub-apertures onto an array detector, and the tilt from each lenslet image is measured using a Quadrant Cell (QC). The Curvature WFS uses the intensity difference between two defocused images to measure the wavefront Laplacian and gradient at the pupil edge. Other types exist but these two are the designs used most commonly in astronomical AO. Figure 2.6 shows example sub-aperture configurations for measuring order $N = 5$; the SH sensor uses 4 pixels per sub-aperture (36 in total), and the CWFS uses 2 pixels per sub-aperture (18 in total). The sizes of sub-apertures in both cases are $\sim 2r_0$. The SNR for a QC is proportional to $\sqrt{n_p}$, where n_p is the flux. For an AO system operating at 200 Hz ($t_0 \simeq 5$ ms), the flux limit is $n_p \simeq 10$. Using all the light at I-band, the limiting magnitude is $m \approx 13$. Using a curvature sensor, the photon noise is reduced but the tip/tilt sensitivity is worse so a similar limit applies.

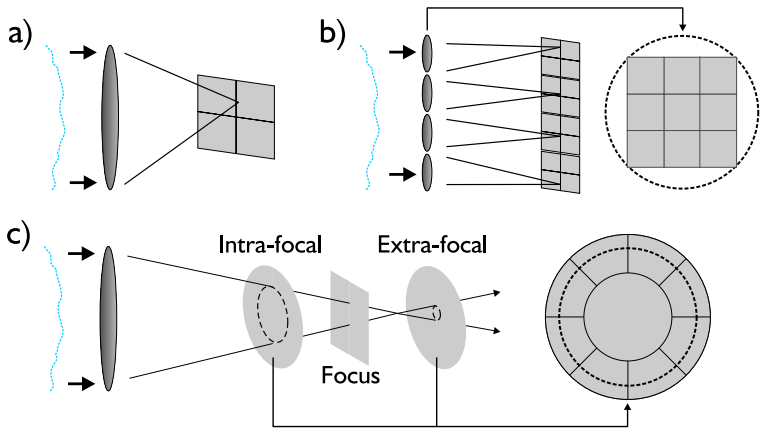

Figure 2.6.: Wavefront sensors used in astronomy; a) Quadrant Cell, b) Shack-Hartmann, and c) Curvature WFS. The dotted outline represents the pupil in the plane indicated by the arrow, and solid lines are the sub-aperture outlines.

- **Modal sensors:** An alternative is to use all the light from the pupil (one sub-aperture) and measure the focused light at a plane where diffraction is important. In doing so, the phase across the aperture modifies the intensity at all points in the image c.f., the Strehl ratio. The image distortions can then be related to a modal decomposition of phase. An example of a modal WFS is the Quadrant Cell, see figure 2.6. It can be thought of as a single sub-aperture SH sensor and therefore has a limiting magnitude of $m \approx 16$. It provides order $N = 2$ measurement and therefore only at long wavelengths is this order of correction optimum. Alternatively, if the aperture were stopped down to $D/r_0 = 3$ then the limiting magnitude would be $m \approx 15$. A QC-style sensor capable of sensing defocus ($N = 3$) has been developed but its defocus measurements are only applied at $\sim 1\,\mathrm{Hz}$ (85). A QC has also been combined with a spatial filter by placing a pinhole at the centre of the detector (24).

- **Hybrid sensors:** The Pyramid sensor (59) and the astigmatic Shack-Hartmann sensor (56) are hybrid modal-zonal WFSs which have increased sensitivity compared to zonal WFSs. For the low order sensors being considered here, their advantages are small; both sensors still require 9 sub-apertures so have a similar magnitude limit to the SH sensor.

- **Non-linear methods:** Examples of non-linear wavefront sensing includes Point Diffraction Interferometry and Phase Diversity (16), the latter being potentially the most sensitive wavefront sensing method. These techniques have, however, only been only been studied for high order correction in astronomy and the needed real-time non-linear algorithms, to reconstruct the wavefronts, are currently an unsolved problem.

In imaging applications, the guide star is often different from the science target. As the interferometric FOV is small, an off-axis guide star would be incoherent as concerns the beam combiner and so all its light may be used for the WFS. This has an obvious potential for increasing in (effective) WFS sensitivity. However, the decorrelation of Zernike polynomial amplitudes between two pointing angles—equivalent to the changes in PSF with sky angle—limits the usefulness of this

option. It has been shown (67) that the rate of decorrelation with angle is faster for higher orders; compared to Z_2/Z_3, Z_4 decorrelates more rapidly by a factor of ~ 7, and Z_7/Z_8 by ~ 17.

An alternative to using stars are Laser Guide Stars which are artificial light sources generated in the atmosphere. A LGS projects a laser up and the light either backscatters (Rayleigh beacons) or excites sodium atoms which re-emit light (Sodium beacons). Their usefulness is marginal for low order AO because they do not produce a signal for tip/tilt aberrations.

Discussion

The wavefront sensors currently developed for astronomy are too insensitive for use in an interferometer; for a full aperture a SH sensor can have a limiting magnitude smaller than that from a QC if the aperture was stopped down. Because of the decorrelation of high order Zernike aberrations, it is unlikely that a separate guide star will be used. The WFS therefore does not need a large FOV and need not constrain the DM placement. If the WFS is placed at a position where the corrected beams are close together, then the potential exists to use one physical detector such as a CCD to simplify the design.

2.5.2. Control System Analysis

The WFS requires finite integration times and so introduces a temporal lag, which is controlled by the AO Control System. The integration time, t_{int}, should be maximised while still being able to sample the evolution of the highest measured order. This limit on t_{int} has been calculated by Roddier (67) under the condition that N aberrations are measured and the uncorrected error of these terms is less than twice that of the terms not being corrected. The maximum t_{int} for an order $N = 2$ is ~ 2 times bigger than that for $N = 5$ and ~ 3 times bigger than for $N = 10$. A $N = 5$ AO system with a modal sensor therefore has a decrease in limiting magnitude limit of $\Delta m \simeq 0.6$ and $N = 10$ has a decrease of $\Delta m \simeq 1.2$.

The other temporal lag come from the processing of the WFS measurement into a wavefront and applying control algorithms. The requirements for low order AO are no different from, for example, a phase delay fringe tracker and this aspect of the control system is a "solved problem".

2.5.3. Deformable Mirror Analysis

The phase aberrations from the atmosphere can be described as wavelength-independent path delays so the AO correction should be similarly achromatic. A realistic wavelength range to correct is 0.6–2.2 μm, covering the visible and near-IR. In AO for a monolithic telescope, a DM can introduce piston phase aberrations—known as the piston response—without consequence for the image resolution. In an interferometer this additional piston changes the OPD and reduces the fringe visibility, so the piston response should either be zero or calibrated and removed dynamically via the path delay modulation. These considerations, being the most important, are compared with the major types of DM available (27),

- **Segmented Mirrors**: These mirrors use discrete mirror segments which are controlled individually. Using three actuators per sub-aperture, these mirrors have excellent temporal

response and can have a large stroke range. However, they are complicated to manufacture because they require a large number of degrees of freedom to replicate a specific shape. The non-continuity of the surface also leads to diffraction losses. They have no piston response problem and, like any mirror, offer achromatic correction.

- **Membrane Mirrors**: Using a single surface, the mirror faceplate is made very thin and its shape directly dependent on the actuators supporting it. This requires the mirror to be supported at fixed points, and controlling the piston response becomes non-trivial. The actuator spacing controls the maximum stroke because a greater actuators density reduces the permissible strain on the mirror surface. A low order mirror (fewer actuators) allows larger spacings and therefore bigger stroke.

- **Bimorph Mirrors**: This mirror is again a thin layer supported by a baseplate. Instead of using discrete actuators, bimorphs rely on the contractions/expansion of the baseplate. This is constructed of a piezoelectric ceramic, for example Lead Zirconate. The mirror is physically supported at fixed points (usually the edge), so has the same piston control issue as membrane mirrors. Bimorphs also suffer from actuator spacing restricting the stroke. They are more reliable, however, by virtue of their solid state design.

- **Liquid Crystal Modulators**: These are transmissive devices whose refractive index changes with applied voltage. The physical mechanism of phase correction is wavelength independent using, for example, the Pancharatnam phase (42). The temporal response ($> 20\,\mathrm{ms}$) is slow though and dispersion effects restrict the bandwidth over which they are useful, $\sim 400 - 900\,\mathrm{nm}$.

- **Dynamic Optics**: Also called Dirigible Optics, these devices correct aberrations by moving fixed optics to produce low order Zernike aberrations: tip/tilt, defocus and astigmatism (28). The bandwidth of these devices are limited to either the visible to near-IR or the near-IR to mid-IR, and no such system has been demonstrated in a working AO system.

One further consideration for a DM is its conjugation to the height of the dominant (most turbulent) phase screen, figure 2.7. Using conjugation with a monolithic telescope improves the FOV for which there is effective AO correction, but this benefit is not relevant for an interferometer. However, the diffraction from a phase screen causes amplitude variations across the telescope aperture and, although these are small, reduces the effectiveness of the AO correction. The use of AO conjugation in an interferometer is to account for this last effect and so improve the accuracy of the correction.

The conjugation altitude is site dependent, for example the Altair (29) AO system on Mauna Kea is conjugate to $\sim 7\,\mathrm{km}$ but there is evidence that the ground layer can be a strong phase screen (67) so the final choice must depend on site measurements. There have been few at (candidate or existing) locations for interferometers, which are considerably flatter than the sites of monolithic telescopes.

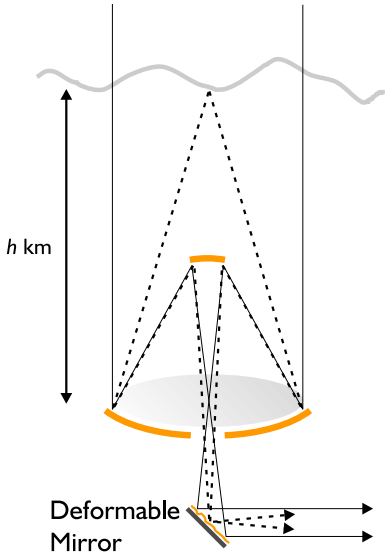

Figure 2.7.: Conjugation of the DM to a dominant phase screen at an altitude of h km.

Discussion

The membrane and bimorph deformable mirror types are available commercially and have been demonstrated as suitable for low order AO (60). The bimorph piston response can be compensated for (84) and an AO system with such a DM is planned for the GI2T interferometer. Dynamic optics are an unproven solution and will have lower throughput than a reflective mirror solution but may prove cost effective. As with the control system, the only issues for wavefront correction are implementation specific. There is also a requirement for specific interferometer-site measurements of conjugation altitude.

2.6. Conclusion

The effects of phase aberrations on the fringe creation process, and the resulting degradation of the complex visibility, have been reviewed. Using Strehl ratio as a proxy to evaluate the visibility coherence is shown to be valid. The availability of passive phase aberration correction, using spatial filters, and the generally smaller apertures in an interferometer suggest that low order correction is the correct approach. Several AO components, namely the deformable mirror and control system, are available in a suitable form commercially. The requirement of high sensitivity wavefront sensors, because of the necessity for short exposure measurements within some system of the interferometer, currently holds back the implementation of a working AO system.

3. Wavefront Sensing Methods

3.1. Introduction

The interferometer AO system analysis identified that a high sensitivity low order WFS is needed. Such designs have generally not been considered for astronomical AO, where instead high order WFSs have been preferred. In order to maximise the sensitivity of a WFS, the intensity must be measured at a plane where it is most sensitive to the phase aberrations. This plane is the focal plane, but then diffraction causes the image to becomes a non-linear function of the phase. Several approaches to using the focal—or near focal plane—intensity have been used to infer the phase but these methods do not have the associated real-time algorithms needed for use in an AO system.

The low order phase aberrations, however, cause focused images to distort in a simple and, crucially, independent fashion so intensity changes can be described with respect to aberration modes. When these modes have a small amplitude, the intensity changes for each can be approximated as linear functions of the mode's amplitude and the problem of inferring the modal amplitudes from the intensity changes becomes more tractable for real-time operation.

In this chapter, the model of intensity changes outlined above is exploited to formulate a high sensitivity modal WFS. Beginning with the principles of measuring tip/tilt aberrations with a focal plane image, the method used to measure low order Zernike polynomial aberrations will be discussed. As sensitivity of the WFS is important, the maximisation of this quantity is investigated. The main results from the analysis demonstrate the principles of measuring the amplitudes of aberration modes using two near-focal plane images. Based on this principle, a new wavefront sensing method called Diffractive Phase Sensing (DPS) is presented here and it is the basis of the proposed WFS design.

This DPS WFS is discussed qualitatively to understand which aberrations can be measured most efficiently. The concept of integrating spatial filtering into the operation of a WFS was briefly mentioned in the previous chapter and incorporation of a pinhole spatial filter with a DPS WFS is described here in more detail.

3.2. Low Order Wavefront Sensing

The WFS currently used in (most) interferometers is the Quadrant Cell (QC), and it measures the image motion at the focus to calculate the amplitude of linear phase aberrations. These aberrations, referred to as tip/tilt, can be represented by the first two Zernike polynomials, Z_2 and Z_3. The QC is a low order ($N = 2$) and modal WFS; the latter because the intensity in the image results from the diffraction of the amplitude across the pupil. The use of only four detector pixels minimises detector noise to maximise the measurement sensitivity.

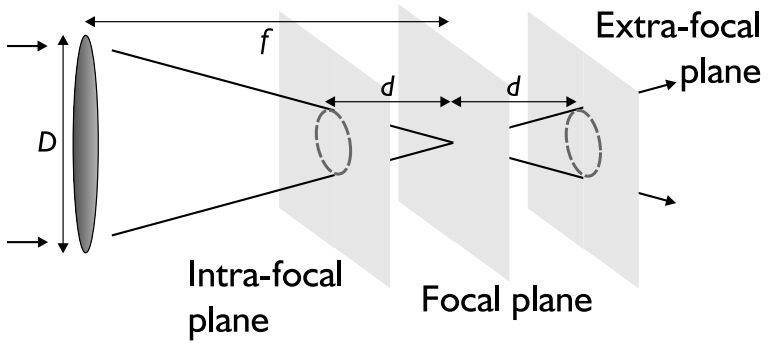

Figure 3.1.: Optical layout showing the relationship between the focal plane, intra-focal plane and extra-focal plane.

Remaining at the focal plane, the effects on an image of the next 3 Zernike aberrations are considered. These aberrations from the next radial order above Z_2/Z_3 are also referred to as the quadratic aberrations: defocus (Z_4), astigmatism 1 (Z_5), and astigmatism 2 (Z_6). The image is blurred in the presence of these aberrations, and importantly this effect is independent of the aberration sign; the effect of a positive defocus is indistinguishable from that of a negative defocus.

To make measurements that break this sign degeneracy, two images are required from two planes that are offset along the optical axis (equivalently, the planes are defocused). This approach is used for two current wavefront sensing methods: Curvature Sensing (64) and Phase Diversity (21). The former uses two equally but oppositely defocused planes that are sufficiently far from the focus that geometrical optics applies. The resulting differences in intensity between equivalent points in the planes is then proportional to the curvature (Laplacian) of the wavefront. The Phase Diversity method typically infers the wavefront from intensities, recorded at the focus and a defocused plane, and uses an iterative algorithm. The defocused plane in this case uses a small defocus such that the intensity cannot be described by geometrical optics.

The wavefront sensing method that will be discussed in this chapter is a hybrid of these two approaches; the intensities at two equally defocused planes (c.f., Curvature Sensing) are used when Fourier optics (diffraction) applies (c.f., Phase Diversity). The two planes are called the Intra-Focal Plane and the Extra-Focal Plane, and their layout relative to the focus is shown in figure 3.1. Examples of the effects that the first 5 Zernike aberrations have on the defocused images at these planes are shown in figure 3.2, under the assumption that the aberrations have small amplitudes. Using near-focus images increases the wavefront measurement sensitivity (c.f. Phase Diversity) (19) but using equally defocused planes allows the intensity changes to be linearly related to the aberration amplitudes in a modal fashion (c.f. Curvature Sensing). The modal behaviour occurs because each aberration changes the intensity across the entire image but in an individual fashion: for example, tilts move the image, defocus causes a radial expansion of the image, and astigmatisms contract and expand the image along perpendicular axes.

Similar near-focus methods have been used for high sensitivity measurements of one Zernike aberration, for the phasing of the primary mirror segments in the Keck Telescopes (9) and in confocal microscopy AO (51). The following sections will examine how to more generally infer the

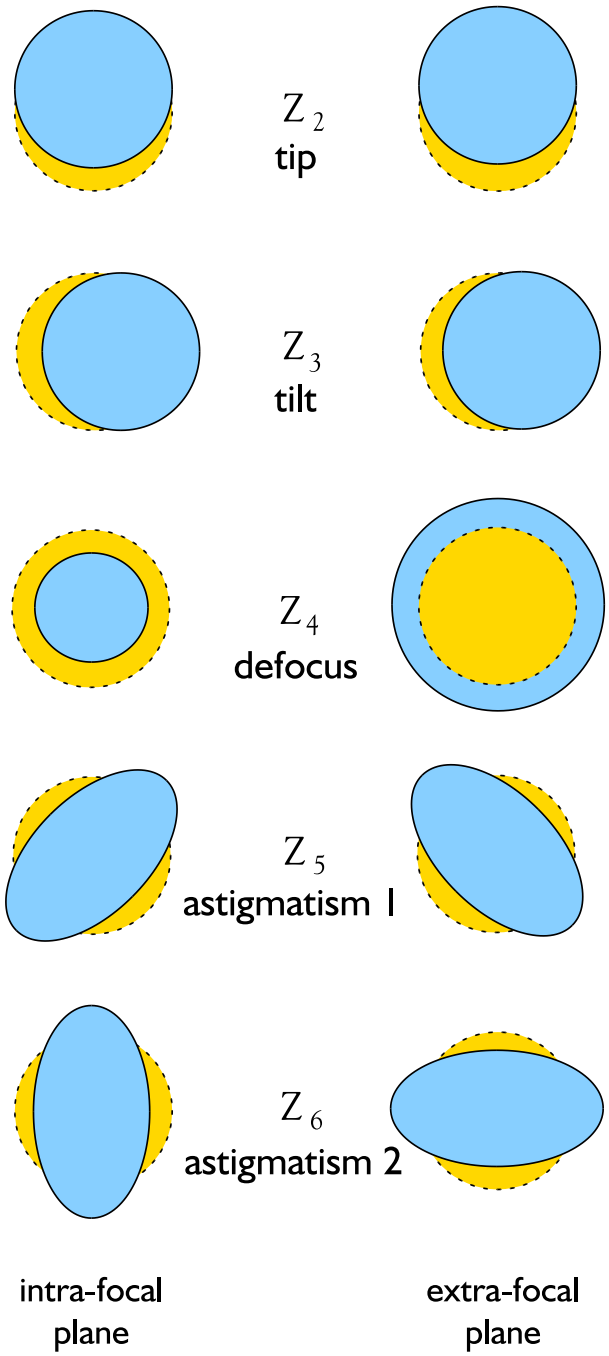

Figure 3.2.: Isophotes (outlines) of the images on the two defocused planes, for the effects of the first 5 Zernike polynomial aberrations. The blue images represent the effect of each aberration and the unaberrated equivalent is in yellow.

aberration amplitudes from the two defocused images and how the measurement sensitivity can be maximised.

3.2.1. Aberration Effects at the Focus and Near-Focus

The isophotes shown in figure 3.2 show an important difference between the tip/tilt and quadratic aberrations: the former cause identical isophote distortions on each plane whereas the latter cause opposite distortions. This is closely related to why only the amplitude, and not the sign, of quadratic aberrations can be measured from a focal plane image (62). In this section, the use of intensity measurements from two defocused planes will be quantitatively analysed and, in doing so, the breaking of the sign degeneracy formally shown.

For the analysis here, it is more straightforward to use the OTF of each defocused image than the the image intensities themselves. A pupil, of diameter D, is defined by the pupil transfer function P,

$$P(x,y) = \begin{cases} 1 & x^2+y^2 \leq D^2/4 \\ 0 & x^2+y^2 > D^2/4 \end{cases}.$$

The phase across the pupil, $\Phi(x,y)$, can be written as a sum of even and odd terms, $\Phi = \Phi_E + \Phi_O$. The even terms are defined as being symmetrical, $\Phi_E(-x,-y) = \Phi_E(x,y)$, and the odd terms are anti-symmetrical, $\Phi_O(-x,-y) = \Phi_O(x,y)$. The Zernike polynomials also have these symmetry properties: Z_2 and Z_3 are symmetric whereas Z_4, Z_5, and Z_6 are anti-symmetric. As "even" and "odd" in the context of Zernike polynomials have conventional definitions (appendix A) that are unrelated to this symmetry, two new definitions are introduced here: the symmetric Zernike polynomials are described as $m-$even and the anti-symmetric polynomials as $m-$odd.

With these definitions in hand, the effect of a general phase aberration can be considered. The unnormalised OTF of a telescope with phase aberrations can be written as,

$$\begin{aligned}\mathcal{O}(\boldsymbol{r}) = &\int d\boldsymbol{x}\, P(\boldsymbol{x}-\boldsymbol{r}/2)\, P(\boldsymbol{x}+\boldsymbol{r}/2) \\ &\times \exp\left[i\Phi_O(\boldsymbol{x}-\boldsymbol{r}/2) - i\Phi_O(\boldsymbol{x}+\boldsymbol{r}/2) + i\Phi_E(\boldsymbol{x}-\boldsymbol{r}/2) - i\Phi_E(\boldsymbol{x}+\boldsymbol{r}/2)\right]. \end{aligned} \tag{3.1}$$

The phase terms can be written more compactly as,

$$\Delta\Phi(\boldsymbol{x};\boldsymbol{r}) = \Phi(\boldsymbol{x}-\boldsymbol{r}/2) - \Phi(\boldsymbol{x}+\boldsymbol{r}/2). \tag{3.2}$$

Rewriting (3.1), and expanding only the complex exponential containing the even phase terms into cosines and sines,

$$\begin{aligned}\mathcal{O}(\boldsymbol{r}) = &\int d\boldsymbol{x}\, P(\boldsymbol{x}-\boldsymbol{r}/2)\, P(\boldsymbol{x}+\boldsymbol{r}/2) \\ &\times \left\{\cos\left[\Delta\Phi_E(\boldsymbol{x};\boldsymbol{r})\right] + i\sin\left[\Delta\Phi_E(\boldsymbol{x};\boldsymbol{r})\right]\right\} \exp\left[i\Delta\Phi_O(\boldsymbol{x};\boldsymbol{r})\right]. \end{aligned} \tag{3.3}$$

Without loss of generality, the offset $\boldsymbol{r}$ can be written as parallel to one Cartesian coordinate i.e., $\boldsymbol{r} = (0,r)$ when $\boldsymbol{x} = (x,y)$. The 2D integral in (3.3) can then be written as a sum of 1D integrals,

$$\int d\boldsymbol{x}\, f(\boldsymbol{x}) \rightarrow \int_{-\infty}^{\infty} dx \, \left(\int_{0}^{\infty} dy \; f(x,y) + \int_{-\infty}^{0} dy \; f(x,y) \right).$$

Let the limits of the second integral within the brackets be reversed. From (3.2),

$$\begin{aligned} \Delta\Phi_O\,(x,y;r,0) &= \Delta\Phi_O\,(x,-y;r,0) \\ \Delta\Phi_E\,(x,y;r,0) &= -\Delta\Phi_E\,(x,-y;r,0)\,. \end{aligned}$$

All terms containing $\sin\left[\Delta\Phi_E\,(u,v;r,0)\right]$ in (3.3) therefore cancel and the remaining terms are

$$\begin{aligned} \mathcal{O}\,(r,0) = 2 \int_{-\infty}^{\infty} \int_{0}^{\infty} dx\, dy \\ \times P\,(\boldsymbol{x}-\boldsymbol{r}/2\,)\, P\,(\boldsymbol{x}+\boldsymbol{r}/2\,) \cos\left[\Delta\Phi_E\,(\boldsymbol{x};\boldsymbol{r})\right] \exp\left[i\Delta\Phi_O\,(\boldsymbol{x};\boldsymbol{r})\right]. \end{aligned} \tag{3.4}$$

As the even phase term $\Delta\Phi_E$ is only present in a cosine, if the sign of Φ_E is changed then the OTF does not change. This is the sign degeneracy and can be overcome by adding a known even phase, Φ_D, to the pupil. This deliberately added phase aberration is known as the Diversity, and for defocused planes the diversity aberration is defocus. The diversity for the intra-focal image is then equal but has an opposite sign to that for the extra-focal image.

The difference in the OTFs of images from the intra-focal and extra-focal plane is equivalent to the Fourier transform of the difference of the images from these planes. Let the diversities for the two planes be written as $\pm\Phi_D$ and the corresponding OTFs as $\mathcal{O}_{\pm}$. The difference of their OTFs is then,

$$\begin{aligned} \mathcal{O}_{-}\,(\boldsymbol{r}) - \mathcal{O}_{+}\,(\boldsymbol{r}) = 4 \int_{-\infty}^{\infty} \int_{0}^{\infty} dx\, dy \; P\,(\boldsymbol{x}-\boldsymbol{r}/2\,)\, P\,(\boldsymbol{x}+\boldsymbol{r}/2\,) \\ \times \sin\left[\Delta\Phi_E\,(\boldsymbol{x};\boldsymbol{r})\right] \sin\left[\Delta\Phi_D\,(\boldsymbol{x};\boldsymbol{r})\right] \exp\left[i\Delta\Phi_O\,(\boldsymbol{x};\boldsymbol{r})\right]. \end{aligned} \tag{3.5}$$

There are two important results here: the even phase is now present in a sine so the OTF difference depends on the sign of Φ_E, and the magnitude of the difference also depends on the amplitude of the diversity.

3.2.2. Sensitivity of measurement to diversity amplitude

The previous section showed that the intensity differences between equally defocused planes breaks the sign degeneracy for measurements of $m-$even Zernike aberrations. These differences form the WFS signal, and the signal sensitivity as a function of the diversity amplitude will now be examined. The sine term in (3.5) containing the diversity, in the difference of the OTFs means that when the diversity is zero (the planes coincide) the signal is (obviously) zero. If the diversity is increased but remains small, then the sine term can be approximated as a linear function and the sensitivity increases as the two planes move away from—but remain near—the focus.

Curvature Sensing also use the subtraction of intensity between intra-focal and extra-focal planes to measure the wavefront. If the intra-focal and extra-focal planes are a distance d from the focus, and the focal length is f, the Curvature Sensing signal (64) is,

$$\frac{I_+(\boldsymbol{r}) - I_-(-\boldsymbol{r})}{I_+(\boldsymbol{r}) + I_-(-\boldsymbol{r})} = f\frac{(1-d/f)}{d/f}\left[\nabla^2\Phi\left(\frac{d}{f}\boldsymbol{r}\right) + \delta_n\frac{\partial\Phi}{\partial\boldsymbol{n}}\left(\frac{d}{f}\boldsymbol{r}\right)\right]. \tag{3.6}$$

The sensitivity of this signal is a function of the relative defocus from the focal plane, d/f, and increases without bound as the planes approach the focus. This apparent contradiction with the previous result is because of the requirement of geometrical optics for the Curvature Sensing method i.e., it only applies when the defocus diversity is large.

Combining the sensitivity results, it is argued that the sensitivity of wavefront sensing using two defocused images is maximised when d is small and diffraction is important. This will be quantified by approximating the focal plane image with a Gaussian function,

$$I_0(\rho,\theta;k) \simeq C\exp\left[-\rho^2/2\omega^2\right], \tag{3.7}$$

and the diffraction-limited FWHM of the Gaussian is $0.44\,f/kD$. The intensity at the defocused planes can now be written analytically for arbitrary defocus diversity. In the intra-focal and extra-focal planes, the images are also Gaussian,

$$I_\pm(\rho,\beta;k) = C\,\frac{\omega^2}{\sqrt{\omega^4+F^2}}\exp\left[-\frac{\rho^2}{2}\left(\frac{f}{f-d}\right)^2\left\{\frac{\omega^2}{\sqrt{\omega^4+F^2}}\right\}\right]. \tag{3.8}$$

The (intra-focal) defocus diversity is

$$\begin{aligned}\Phi_D &= -\left(\frac{k}{f}\right)^2 F\,r^2,\\ F &= \frac{f}{2k}\frac{d}{f-d}.\end{aligned} \tag{3.9}$$

The magnitude of the diversity describes if diffraction is important. For large diversity ($F \gg \omega$), the image width is a linear function of the geometrical variables d, f, and D. With small diversity ($F \leq \omega$), the contribution from the diffraction-limited FWHM, ω, causes the image size to tend to the diffraction-limited value.

In the pupil, defocus and astigmatism aberrations can be added by a multiplication of the (real) Gaussian amplitude by a (complex) Gaussian. Adding a defocus aberration, $\Phi = -\gamma\,r^2$, the defocused images remain Gaussian and symmetric,

$$I_\pm(\rho,\beta;k) = C\,\frac{\omega^2}{\sqrt{\omega^4+(F\pm f\gamma/k)^2}}\exp\left[-\frac{\rho^2}{2}\left(\frac{f}{f-d}\right)^2\left\{\frac{\omega^2}{\sqrt{\omega^4+(F\pm f\gamma/k)^2}}\right\}\right], \tag{3.10}$$

where I_+/I_- is the intensity in the intra-focal/extra-focal plane. A plot of the difference between I_+ and I_-, integrated over the image planes, is shown in figure 3.3 and is normalised by the largest difference. The sensitivity to $m-$even aberrations is maximised when the defocus distance is small and diffraction is important—as argued above.

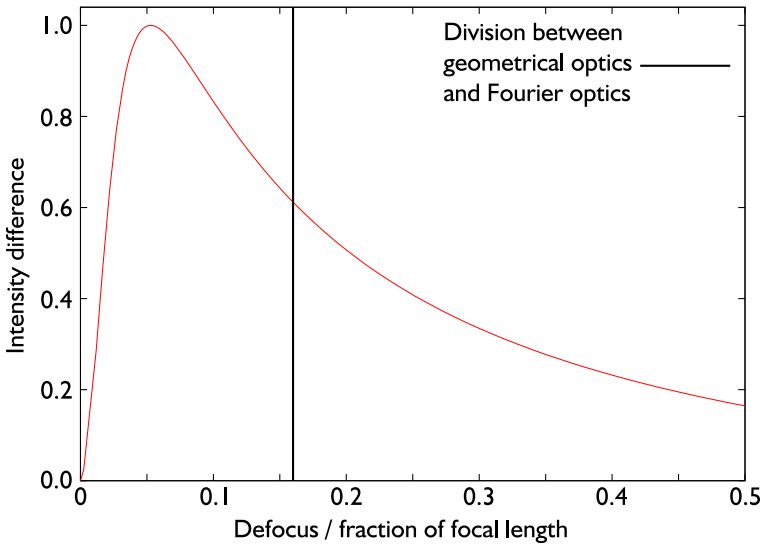

Figure 3.3.: Normalised integrated difference between the intensity of two defocused Gaussian images with small defocus aberration.

3.3. Diffractive Phase Sensing

In summary, the amplitude of $m-$even and $m-$odd aberrations can be inferred with maximum sensitivity from the intensities in two defocused planes which are near the focus. The Diffractive Phase Sensing method of wavefront measurement uses these principles and its details are discussed in this section. It is noted that so far only defocus has been considered as the diversity, because it is physically easy to implement, although in principle any $m-$even aberration can be used.

The defocus diversity amplitude depends on the distance of the defocused planes from the focus d, the focal length f, and the pupil diameter D. The defocus diversity, $\pm\Phi_D$, can then be written as

$$\Phi_D\,(x,y) = -\frac{kd}{2f(f-d)}\left(x^2+y^2\right) \equiv a_d\,Z_4\,(x,y)\,.$$

The equivalence shows that the diversity can be written as a Zernike polynomial. It is noted that small diversity implies $a_d \lesssim 1\,\text{rad rms}$, so the upper bound on d is,

$$d \lesssim 16\sqrt{3}f^2\,/kD^2\,. \tag{3.11}$$

This limit is a function of the wavelength and the focal ratio D/f and, assuming $\lambda = 500\,\text{nm}$ and $D/f = 1/50$, $d \lesssim 5\,\text{mm}$.

The assumption that the amplitude of the diversity is small is now extended to the amplitudes of aberrations being measured. Under this condition, the difference of intensities can be shown to be proportional to the amplitude of Zernike polynomial aberrations. This proportionality occurs because the equal defocus of the two planes means that with no aberrations the two images are identical and their difference becomes trivially zero. The proof of linearity is derived in appendix B and the main results are given here.

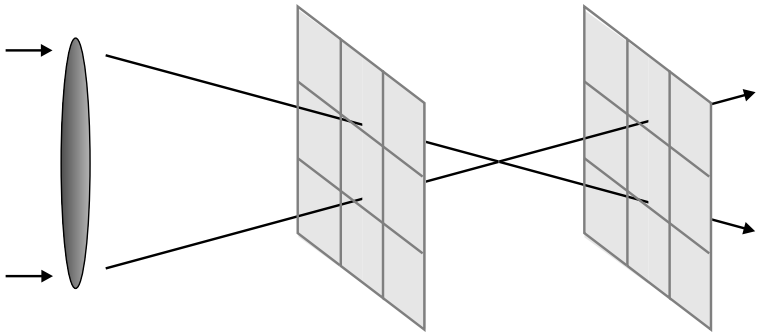

Figure 3.4.: The Nine Element Sensor design, using 9 pixels per intra-focal and extra-focal plane (note that the diagram is not to scale).

The even phase is decomposed using the subset of Zernike polynomials which are $m-$even, $Z_{j;e}$. The odd phase is similarly decomposed by using the $m-$odd Zernike polynomials, $Z_{j;o}$. The reason for this division is that two possible treatments of the intensities are possible: the inter-plane difference and the intra-plane difference. To first order, the former only measures the amplitude of $m-$even terms and the latter only measures the amplitude of $m-$odd terms.

The inter-plane difference, $\Delta I_\delta(u,v)$, is the subtraction of the intensities in each plane between equivalent points,

$$\begin{aligned}&\Delta I_\delta(u,v) = [I_+ - I_-](u,v) \propto \\ &\sum_{n=0}^{\infty} a_n \left(\mathcal{F}^s\{P\}\,\mathcal{F}^s\{P\Phi_D\,Z_{n;e}\} - \mathcal{F}^s\{P\Phi_D\}\,\mathcal{F}^s\{P\,Z_{n;e}\}\right)(u,v),\end{aligned} \tag{3.12}$$

where the last line is equation (B.6) from appendix B. The intra-plane difference, $\Delta I_\Sigma(u,v)$, is the subtraction of the summed intensity, from the two planes, between diametrically opposed points,

$$\begin{aligned}\Delta I_\Sigma(u,v) &= [I_+ + I_-](u,v) - [I_+ + I_-](-u,-v) \\ &\propto i\sum_{m=1}^{\infty} a_m\,\mathcal{F}^s\{P\}\,\mathcal{F}^s\{P\,Z_{m;o}\}(u,v),\end{aligned} \tag{3.13}$$

where the last line is equation (B.9) from appendix B. The operator $\mathcal{F}^s\{\}$ represents a Fourier transform with scaled variables on the Fourier plane c.f., the Fraunhofer Diffraction integral. The two types of intensity differences are collectively called the Difference Signals, and form the measurements used to infer the wavefront.

3.4. Properties of Diffractive Phase Sensing

An example WFS design is shown in figure 3.4 which can make intensity measurements that can be converted to difference signals. It is called the Nine Element Sensor (NES) and uses a 3×3 array of square light-detector elements (pixels) in each defocused plane. The choice of square pixels is a pragmatic one and allows the intensities to be recorded directly onto a CCD, without needing

either specially shaped pixels or the use of a lenslet array. The properties of DPS, using the NES as the WFS, are now described qualitatively.

3.4.1. Qualitative Description of NES

The method by which the NES measures the first 5 Zernike aberrations is now explained to explain the WFS design. The tip aberration causes a horizontal movement of the image. The NES measures this via a subtraction of the flux in the lefthand column of pixels from the righthand column, ignoring the central column. The tilt, causing vertical movement, is measured similarly by subtracting the flux in the top row of pixels from the bottom row of pixels. Because the movement is identical on both planes, the subtraction gives the same result in either plane. The sensitivity can be increased by adding the fluxes between the two planes before subtracting across the summed intensity—this is the intra-plane difference.

The defocus aberration increases the image size in one plane and decreases it in the other. The effect of astigmatism aberrations is represented by a contraction and expansion of the images in perpendicular directions. The directions of contraction and expansion in one plane are reversed in the other plane. These quadratic aberrations can be measured by the subtraction of intensity measured by equivalent pixels in each planes—this is the inter-plane difference.

By using two types of difference signals, the measurements of $m-$odd aberrations are independent from the measurements of $m-$even aberrations i.e., the signals from tip and tilt aberrations do not affect the signals from defocus and the astigmatism aberrations, as long as the amplitudes are small.

3.4.2. Order of Measurement

A modal wavefront sensor design should be matched to the modes it is intended to measure and be insensitive to other terms. It has been indicated in section 2.4.1 that a suitable order of an interferometer AO system is $N = 5$, hence the NES should measure the tip/tilt and quadratic aberrations.

The intra-plane and inter-plane differences for the first 12 polynomials are plotted in figure 3.5, using amplitude of 0.1 rad rms for both the diversity and the aberration. As both $m-$even and $m-$odd aberrations produce a signal which is symmetrical about the image centre, the number of independent aberrations that can be measured equals the number of pixels in one plane c.f., the same result for Curvature WFSs. This can also be considered in terms of the number of intra-plane and inter-plane differences. The NES can therefore measure 9 aberration modes, from 4 intra-plane and 5 inter-plane differences.

The contrast between the simple signals for Z_2 and Z_3 and the more complex signals for the higher order $m-$odd terms is the reason for only measuring the first two $m-$odd aberrations. The detector pixel sampling that would be required to measure these high order terms is much larger than that needed for tip/tilt. Comparing the signal from Z_4, Z_5, and Z_6, to the the signals from higher order $m-$even terms also shows the same increase in complexity. This explains why the only $m-$even aberrations measured are the quadratic terms.

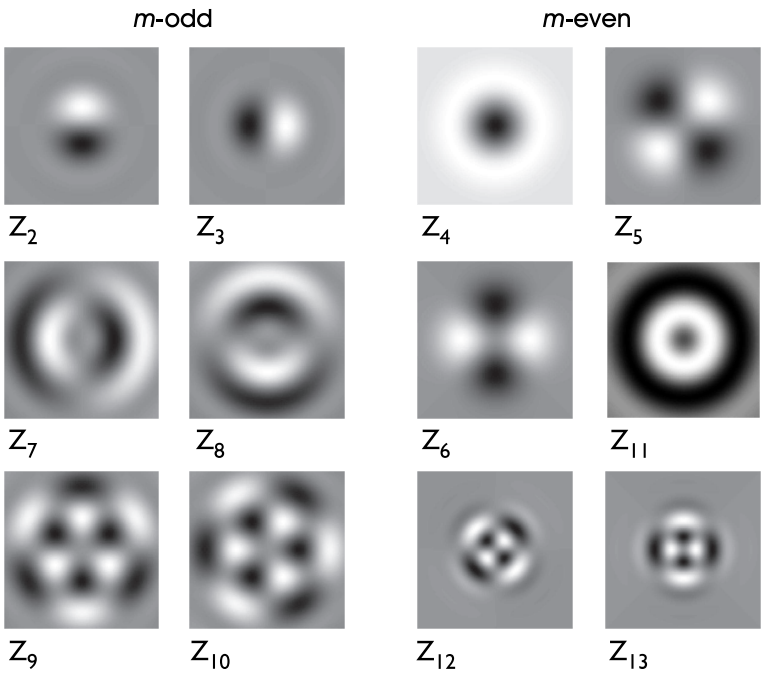

Figure 3.5.: Intra-plane differences for the first 6 $m-$odd Zernike polynomials, and the inter-plane differences for the first 6 $m-$even Zernike polynomials.

3.4.3. Orthogonality of Difference Signals

The difference signal is orthogonal between $m-$odd and $m-$even Zernike aberrations only when their amplitudes are small, ≤ 1 rad rms. An aperture of diameter $D/r_0 \sim 5$ has an uncorrected tip/tilt variance of ~ 7 rad rms and—even when the AO loop is closed—tip/tilt errors greater than 1 rad rms can occur. This amplitude is no longer small, and causes coupling between the signals from $m-$odd and $m-$even modes. Alternatively, the coupling is described by considering the resulting image motion: the motion is identical in two images and this effectively moves the inter-plane differences relative to the detector pixels. The measured intensities are therefore integrated over different parts of the image from that assumed when the images are centred, and this in turn affects the calculated difference signals. The coupling is shown in the top line of figure 3.6.

A method of correcting for this coupling *after* the intensities are measured is now presented. The OTF of a defocused (or focused) image with only $m-$even aberrations is real, see equation (3.4), so the MTF is identical to the OTF. When Z_2 and Z_3 aberrations are present, the OTF changes but the MTF is unchanged. This can be understood by recalling that the MTF is used to determine the image resolution, and the instantaneous image resolution is unaltered by tip/tilt aberrations. By replacing the OTF with the MTF, the effect of image motion for the inter-plane differences can be removed without affecting the inter-plane difference signals.

The second largest $m-$odd terms after Z_2/Z_3 are coma and trefoil. Their variance is ~ 75 times smaller than for tip/tilt and uncorrected coma has a standard deviation of ~ 0.6, over a $D/r_0 \sim 5$ sized aperture. It will therefore be assumed that the odd phase over such an aperture is represented only by the linear terms, $\Phi_O \equiv \Phi_L$ and the even phase, Φ_E, is unchanged,

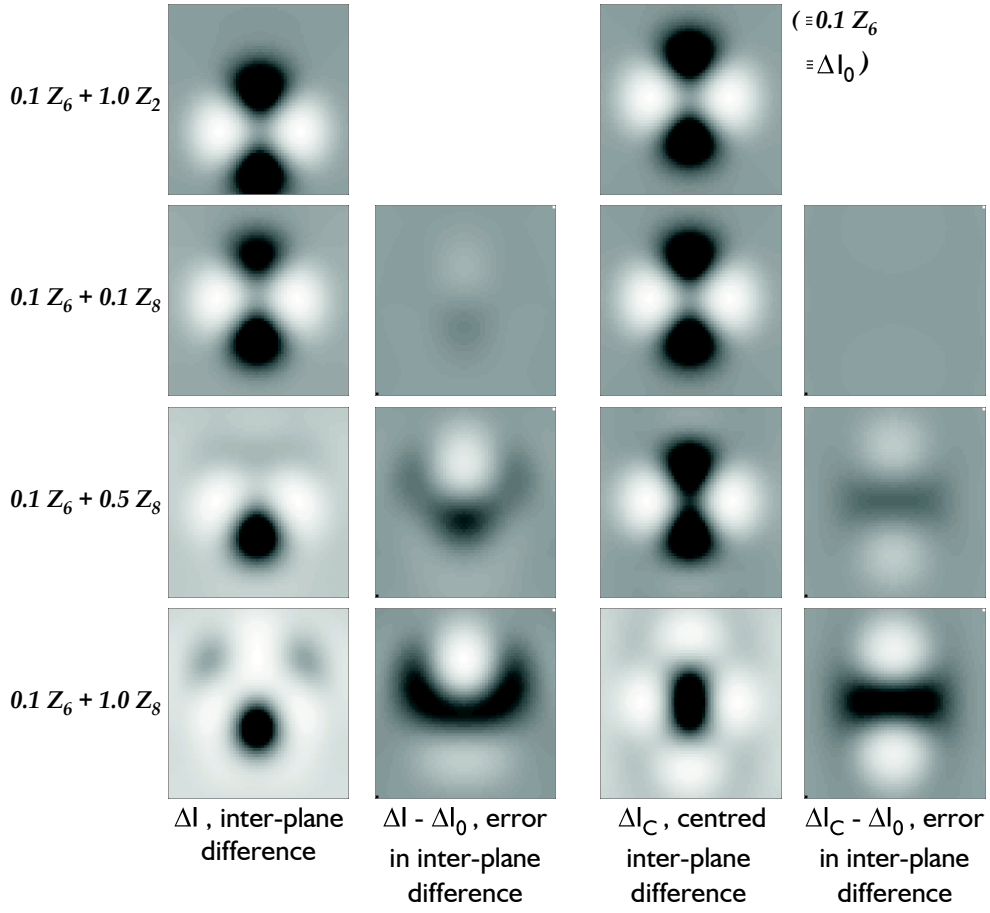

Figure 3.6.: Inter-plane differences before and after centring. The difference signal where there are no $m-$odd aberrations, $\Delta I_{\delta 0}$, is the goal of the centring algorithm.

$$\begin{aligned} \Phi(x,y) &\simeq \Phi_E(x,y) + \Phi_L(x,y), \\ \Phi_L(x,y) &= \alpha_x x + \alpha_y y. \end{aligned} \tag{3.14}$$

The OTF of the intensities in the two defocused planes, equation (3.3), is written with the phase terms as differences, $\Delta\Phi$. The difference of the linear term is only a function of the offset, $\boldsymbol{r} = (r_x, r_y)$, and not location in the pupil, (x, y),

$$\Delta\Phi_L(\boldsymbol{x};\boldsymbol{r}) \equiv \Delta\Phi_L(\boldsymbol{r}) = \alpha_x r_x + \alpha_y r_y. \tag{3.15}$$

The linear phase terms are taken outside of the integral in the OTF, and the difference of the OTFs, equation (3.4), is

$$\begin{aligned} \mathcal{O}_-(\boldsymbol{r}) - \mathcal{O}_+(\boldsymbol{r}) &= 4\exp[i\,\alpha.\boldsymbol{r}] \\ &\times \int_{-\infty}^{\infty}\int_{0}^{\infty} dx\,dy\; P(\boldsymbol{x}-\boldsymbol{r}/2)\,P(\boldsymbol{x}+\boldsymbol{r}/2)\sin[\Delta\Phi_E(\boldsymbol{x};\boldsymbol{r})]\sin[\Delta\Phi_D(\boldsymbol{x};\boldsymbol{r})]. \end{aligned} \tag{3.16}$$

Taking the modulus of this difference, and therefore discarding the complex exponential outside the integral,

$$
\begin{aligned}
&|\mathcal{O}_{-}(\boldsymbol{r}) - \mathcal{O}_{+}(\boldsymbol{r})| \\
&\quad = \quad 4\int_{-\infty}^{\infty}\int_{0}^{\infty} dx\, dy\; P(\boldsymbol{x} - \boldsymbol{r}/2)\, P(\boldsymbol{x} + \boldsymbol{r}/2) \sin\left[\Delta\Phi_E(\boldsymbol{x};\boldsymbol{r})\right] \sin\left[\Delta\Phi_D(\boldsymbol{x};\boldsymbol{r})\right] \\
&\quad \equiv \quad \mathcal{M}_{-}(\boldsymbol{r}) - \mathcal{M}_{+}(\boldsymbol{r}) \,. \qquad (3.17)
\end{aligned}
$$

Formally, this shows that the modulus of the difference of the OTFs is equivalent to the difference of the MTFs when all linear odd phase terms are removed, leaving only the effect of even phase terms. The centring algorithm is therefore,

- Apply a Fourier transform to the inter-plane differences; which gives the equivalent of the difference of the OTFs
- Discard the phase; convert the transform to the equivalent of the the MTF differences
- Apply an inverse Fourier transform.

Algebraically, writing the Fourier transform operator as $\mathcal{F}\{\}$,

$$\Delta I_{\delta c}(u,v) = \mathcal{F}^{-1}\left\{\left|\mathcal{F}\left\{\Delta I_{\delta}(u,v)\right\}\right|\right\}, \qquad (3.18)$$

where the centred inter-plane difference is $\Delta I_{\delta c}$ and the uncentred equivalent is ΔI_{δ}. Examples of the inter-plane difference with and without centring are shown in figure 3.6. The amplitude of the defocus diversity is 0.1 rad rms and the $m-$even aberration is Z_6 which also has an amplitude of 0.1 rad rms. To this aberration is added $m-$odd Z_2 and Z_8 (tilt and coma). The difference between the difference signal from only Z_6 ($\Delta I_{\delta 0}$) and those before (ΔI_{δ}) and after centring ($\Delta I_{\delta c}$) are also plotted.

The coupling from the tilt aberration, which only shifts the difference signal, is correctly removed when only Z_6 is also present. The asymmetry in the difference signal with coma and astigmatism becomes significant when the coma amplitude is greater than 0.5 rad rms. The centring algorithm then forces the difference signals to become centred in an incorrect fashion. The error in centring is not large for small coma, and indicates that the assumption of ignoring $m-$odd phase terms other than tip/tilt may be acceptable.

3.4.4. Linearity

The wavefront sensor concept here is based on the assumption that the signals measured are linearly related to the wavefront aberrations. Whilst this is true for very small amplitudes, there will be some limit where non-linear effects will become important. The gradient of the sensor measurements, with respect to the input aberration amplitude, is one when linear but non-linearities may cause the gradient to go to zero, or diverge in a negative or positive direction. The first two will still result in (useful) negative feedback with a closed AO loop, but the last would cause positive feedback and a divergence in the estimated wavefront from that present. This potential issue will be considered more carefully in the next chapter.

3.5. DPS and Spatial Filtering

As previously mentioned, a WFS and spatial filter can be integrated and this possibility is outlined here. The defocused images in a DPS WFS can be thought of as focused images convolved with a "blurring" function. The intensity at the focal plane is described by the Fraunhofer Diffraction integral, which takes the form of a scaled Fourier transform. Therefore, low and high spatial frequency phase aberrations respectively affect the intensity nearer and further from the optical axis. By placing a pinhole at the centre of the intra-focal/extra-focal plane—at the centre of a DPS WFS—the light that passes through has the high frequency aberrations removed. The proposed WFS-filter combination could simply replace a spatial filter within the interferometer and the filter would operate in the same way as before. The advantage of combining a pinhole filter and WFS is that the throughput of the filter-component is implicitly measured by the WFS-component, because if the throughput goes down then the WFS signal increases and acts to restore the throughput. Thus, the alignment of each component is "automatic" when the AO feedback loop is closed.

The effect the integration would have on the WFS-component measurements can be explained by the principles by which a pinhole spatially filters light. A pinhole can be thought of only allowing through light which is in phase across the aperture. The rejected light falls on an opaque surface and is scattered away. In terms of a DPS WFS, the light that is in phase across the aperture does not contribute to the difference signals because it produces identical (centred) intensity patterns on both defocused planes. The light that is not in phase, and is rejected by the pinhole, is that which gives rise to the intensity differences and results in the difference signals (around the centre of the defocused images). The exception to this is with defocus aberrations; the diversity is a defocus aberration so any defocus phase aberrations causes the intensity at *all* points in the images to vary. The intensity at the centre of the defocused images is therefore sensitive only to defocus aberrations and only its measurements will be made less sensitive.

This argument can easily be made quantitative, as all the on-axis intra-plane differences are, by definition, zero and the on-axis inter-plane differences (3.12) are,

$$\begin{aligned}\Delta I_{\delta}(u,v) \simeq & \\ & 2C \sum_{n=1}^{\infty} a_n \left(\int_{-\infty}^{\infty} d\boldsymbol{x}\, P(\boldsymbol{x}) \int_{-\infty}^{\infty} d\boldsymbol{x'}\, P(\boldsymbol{x'})\, \Phi_D(\boldsymbol{x'})\, Z_{n;e}(\boldsymbol{x'}) \right. \\ & \left. - \int_{-\infty}^{\infty} d\boldsymbol{x}\, P(\boldsymbol{x})\, \Phi_D(\boldsymbol{x}) \int_{-\infty}^{\infty} d\boldsymbol{x'}\, P(\boldsymbol{x'})\, Z_{n;e}(\boldsymbol{x'}) \right). \end{aligned} \tag{3.19}$$

These integrals are identical to the inner product defining the orthogonality of the Zernike polynomials (appendix A). The second set of integrals is therefore always zero, as the mean of all $Z_{n;e}$ are zero, except for the piston term. For the case of piston, the first set of integrals evaluates to the same value as the second set and all terms cancel. For the other Zernike polynomials, the first set of integrals is also zero, except for $Z_{n;e} \equiv \Phi_D = Z_4$. As the diversity is equivalent to Z_4, the on-axis difference signal is only non-zero for the defocus aberration.

In the context of the NES design, this means that not using the flux in the central pixels only causes a loss in the sensitivity of defocus measurements. An example spatial filter-WFS design is

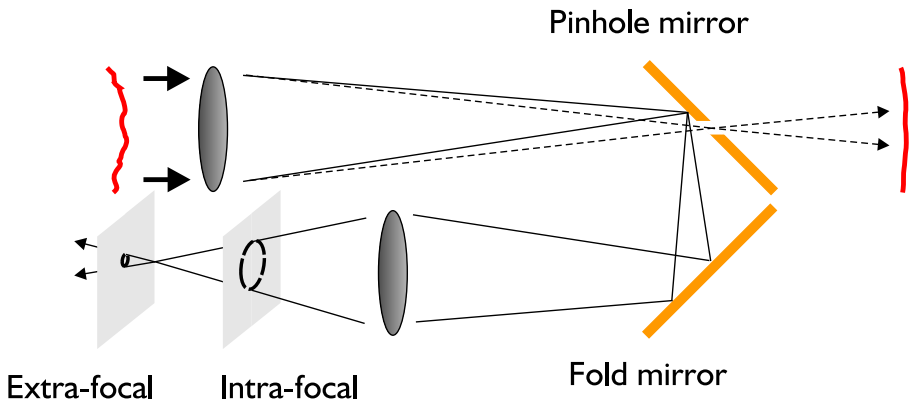

Figure 3.7.: Combined pinhole spatial filtering and wavefront sensing.

shown in figure 3.7. The opaque parts of a pinhole filter are made reflective, and instead of being scattered the rejected light is sent to the WFS optics. This design also allows the use of a single mode fibre in place of a pinhole.

3.6. Conclusion

The DPS wavefront sensing method presented here, with the NES as an example WFS design, is meant specifically for measuring low order Zernike aberrations with high sensitivity. For the NES, only a few detector pixels, ~ 18, are needed which is a considerably smaller number than employed in other WFS designs that measure higher order terms than tip/tilt. The general properties of the NES were explained and a centring algorithm was introduced to remove the coupling between the signals from even phase terms and tip/tilt aberrations. What has not been shown is how to optimise such a sensor and its performance in a closed feedback loop with turbulent phase aberrations. Furthermore, the arguments and descriptions were quantitative. These deficiencies are addressed in the next chapter with numerical simulations.

4. Numerical Simulations

4.1. Introduction

The NES design, which was qualitatively introduced in the previous chapter, will be analysed quantitatively here. Numerical simulations allow the intensities at the intra-focal and extra-focal planes to be calculated for arbitrary aberration amplitudes, and so the NES measurement characteristics can be described. An important component of this is the breakdown of the small amplitude approximation which was used to derive the DPS result of linear wavefront sensing. Another description of the NES is its sensitivity, and this can be quantified by propagating the WFS measurement noise to the noise of the inferred wavefront. Knowing the wavefront noise then also allows the NES design to be optimised for particular measurement characteristics.

These results allow the flux and seeing limits above which a NES WFS provides "useful" measurements to be evaluated. The criterion defining useful is the Strehl ratio in a closed loop AO simulation. Results using the NES are presented and compared with existing WFS designs. It then becomes straightforward to find the improvement in limiting magnitude when using the NES as a WFS in AO. Using the NES in COAST is evaluated as part of this work, and appropriate simulation parameters are included. The extension of the sensor design to incorporate spatial filtering is discussed at the end of this chapter.

4.2. Numerical Methods of WFS Simulation

A numerical simulation of the NES first involves propagating the complex amplitudes from a discretely sampled pupil to the intra-focal and extra-focal planes. The WFS detector pixels are then modelled by binning the intensity values across the images. The difference signals are calculated from the pixel fluxes, and linear wavefront reconstruction is used to find the aberrations in terms of the Zernike polynomial amplitudes.

4.2.1. Pupil Propagation to Defocused Image Planes

The intensities in the intra-focal, I_+, and extra-focal plane, I_-, are calculated from the complex amplitudes propagated from the pupil. The pupil amplitude is defined by the multiplication of the pupil transmission function P with the complex amplitude at the aperture A. The propagation is then calculated using a Fast Fourier Transform, and details of the method are given in appendix C and summarised here.

The discrete intensities, in a numerical array of size $N \times N$, from the defocused planes are written as a Discrete Fourier Transform, $\mathcal{D}\{\}$, of the complex amplitude in the pupil,

$$I_{\pm;\mu\nu} \propto \left|\mathcal{D}\left\{A_{\alpha\beta} P_{\alpha\beta} \exp\left[\pm i \Phi_{D;\alpha\beta}\right]\right\}\right|^2, \tag{4.1}$$

where Φ_D is the defocus diversity. The indices μ, ν & α, β for the imaged and pupil planes are defined to be zero at the centre of the numerical array. The pupil transmission function is one inside a radius of R array elements and zero outside, corresponding to a physical circular aperture of radius r. The results can then be presented in scale-free units: lengths on the defocused planes are in multiples of $1.22\,N/R$ and the diversity amplitude is in rad rms. The Airy disc diameter at the focal plane, if the diversity was zero, is then 2 in the scale-free units. The defocus diversity can be written as a discrete Zernike polynomial, $\Phi_{D;\alpha\beta} = a_d\, Z_{4;\alpha\beta}$.

Discrete Zernike polynomials are samples of the continuous functions, and it was found that they maintain their orthonormality relationship to within 0.5% when their radius (equal to R) was at least 32 array elements.

The radius R affects both the sampling of the defocused images (in terms of the image width in numerical array elements) and the maximum defocus diversity. The former is proportional to N/R whereas the latter is proportional to R, and it was found that the range of defocus required to maximise sensitivity was small. Having therefore chosen a suitable minimum R, N/R was maximised under the condition that the time to compute the propagation didn't become excessive. The sampling of phase across the aperture also depends on R, but the restriction on this from the Nyquist sampling criterion can be ignored (because it requires that $R \gtrsim 0.5\,(D/r_0)$ and $R \sim 32$ and $D/r_0 < 10$ in this work.)

4.2.2. Linear Wavefront Reconstruction

The method of wavefront reconstruction used for the NES (and other WFS models) is the matrix method used commonly for zonal WFSs (14). In this scheme, the WFS measurements are converted to the amplitudes of the modes in a basis set, here the Zernike polynomials, using matrix algebra. Two calculation stages are required: the difference signals are first calculated for each aberration alone and used to create the Measurement Matrix, and then this matrix is inverted and used for calculating the mode amplitudes given general difference signals. An advantage of the matrix approach is that the noise propagation to the amplitudes of the Zernike aberrations is linear.

The measured NES WFS signals are first normalised and written as a vector, s, which consists of both types of (normalised) difference signal,

$$\mathsf{s} = \begin{pmatrix} \uparrow \\ \Delta I_\delta \,/\Sigma I_\delta \\ \downarrow \\ \uparrow \\ \Delta I_\Sigma \,/\Sigma I_\Sigma \\ \downarrow \end{pmatrix}, \tag{4.2}$$

where the difference signals are defined as:

inter-plane:

$$\begin{aligned} \Delta I_{\delta i} &= I_{+;i} - I_{-;i} \\ \Sigma I_{\delta i} &= I_{+;i} + I_{-;i} \end{aligned}$$

intra-plane:

$$\begin{aligned} \Delta I_{\Sigma i} &= \{I_{+;i} + I_{-;i}\} - \left\{I_{+;i'} + I_{-;i'}\right\} \\ \Sigma I_{\Sigma i} &= \{I_{+;i} + I_{-;i}\} + \left\{I_{+;i'} + I_{-;i'}\right\} \end{aligned} \tag{4.3}$$

The intensities are the detector pixel fluxes, and the first part of the subscript, $+/-$, refers to which plane the pixels are in and the second part to which particular pixel. Note that i' is the pixel which is diametrically opposite from pixel i.

A subset of the Zernike polynomials, Z_s, correspond to the measured modes of the phase aberrations. The amplitude of the modes in Z_s are written as a vector, a, and the relationship between the two vectors and the measurement matrix, M, is

$$\mathsf{s} = \mathsf{M}\mathsf{a}. \tag{4.4}$$

To construct M, the vector s is calculated when $a_i = 0.1\,\delta_{ij}$. Column j in M is then equal to this s,

$$M_{ij} = s_i \;\text{ for }\; \Phi = 0.1\,Z_j \text{ for } Z_j \in Z_s. \tag{4.5}$$

The amplitude of the polynomial is small (0.1 rad rms) because this allows the assumption to be made that the difference signals are a linear function of the amplitude. The WFS measures the vector s and it is the vector a which is needed for inferring the aberrated wavefront. This requires an inversion of M but as this matrix is not necessarily square the pseudo-inverse must instead be used. The pseudo-inverse, M^+, is defined as $M^+_{ij}M_{jk} = \delta_{ik}$ i.e., a generalised inverse, noting that $M^+_{ij}M_{jk} \neq M_{ij}M^+_{jk}$. The amplitude vector, whose elements are the amplitudes of Z_s in units of 0.1 rad rms, is then

$$\mathsf{a}' = \mathsf{M}^+\mathsf{s}, \tag{4.6}$$

where the amplitude vector is written with a dash to denote it as being inferred rather than known.

4.2.3. Singular Value Decomposition

The calculation of the pseudo-inverse uses Singular Value Decomposition (SVD). SVD (58) is a general algorithm which takes a matrix, A, and returns its decomposition in terms of 3 matrices: $\mathsf{A} = \mathsf{U}\mathsf{W}\mathsf{V}^\mathsf{T}$. Since U and V are column orthonormal and W is diagonal, the pseudo-inverse of A can be written straightforwardly as,

$$A^+_{ij} = V_{ik}\,(1/W)_{kl}\,U^T_{lj}. \tag{4.7}$$

The SVD algorithm is named because the diagonal elements of W are the eigenvalues of the space of A; if any are zero then it implies that A is singular.

Optimising the Subset of Aberrations

The decomposition of the measurement matrix is not only useful for calculating the pseudo-inverse, but provides information on the difference signals from each aberration in Z_s and their suitability for being measured with the NES.

The columns of U are normalised vectors, and when U is multiplied by W the columns of the resultant matrix can be linearly combined to form the difference signals in M. Therefore W provides information on the sensitivity of the NES to each aberration; if W contains a zero eigenvalue then there are non-independent difference signals.

The combination of the columns of UW is controlled by the values within each column of V^T (or equivalently, each row of V.) If the aberrations in Z_s result in measured difference signals which are orthogonal (independent) then the columns of M are also orthogonal. The SVD is then straightforward and each row of V will have only one non-zero value i.e. each column of UW corresponds to a column of M. A NES then measures the aberration amplitudes independently and without confusion.

If instead two or more *measured* differences signals are not orthogonal, then the SVD will result in V with more than one non-zero value in two or more rows. This situation will arise when the NES is poorly suited to measuring the relevant aberrations; examples might be defocus and spherical aberration whose inter-plane difference signals (see figure 3.5) cannot be distinguished by using only 3 pixels across the images. Given two non-orthogonal difference signals, the SVD of M will have one column in U which contains the common (correlated vector) part of these difference signals and another column with the difference (anti-correlated vector) of the signals. The anti-correlated vector is then associated with an element in W that has a small magnitude, and the difference of the signals will have a non-linear component which is neglected but should not be.

The anti-correlated signal cannot therefore be used to distinguish between the two aberrations, and the effect of the correlated signal is to couple the measured amplitudes of the aberrations: the presence of one aberration will affect the measured amplitude of the other and vice-versa. Although this coupling effect occurs between all aberrations when the assumption of linearity becomes invalid (non-small amplitudes), in this case it affects all amplitudes larger than those used to form M. Furthermore, the anti-correlated vector causes the measured amplitudes to become significantly non-linear and thus (effectively) useless.

In summary, comparing the eigenvalues in W indicates which aberrations are "poorly" measured by the NES and should be eliminated from Z_s.

Noise Propagation using SVD

The propagation of noise to the amplitudes of Z_s is calculated from the SVD decomposition, and proceeds from the (implicit) assumption that the variance on each element of s is independent and equal to 1. The independence occurs as a result of the inter-plane and intra-plane differences measuring mutually different types of aberrations, $m-$even and $m-$odd respectively. The two

types of difference signal can then be considered separately: neither the intra-plane nor inter-plane signals in s make use of any one pixel's intensity more than once, and the result is no correlation between the elements of s. Under this condition, the measurement noise is incorporated into W by scaling s before constructing M,

$$\overline{s_i} = \frac{s_i}{\sigma_{s_i}}. \tag{4.8}$$

The scaled measurement matrix, and its decomposed components, are designated by a bar over the relevant variables—noting that the matrices $\overline{\mathsf{V}}$ and $\overline{\mathsf{U}}$ are unchanged since these are unitary by definition. The amplitude variances are then,

$$\langle \sigma_{a_i}^2 \rangle = \sum_j \frac{\overline{V}_{ij}^2}{\overline{W}_{jj}^2}. \tag{4.9}$$

The total variance of the amplitudes, σ_{a}^2, can include the cross-correlation terms $\langle \sigma_{a_i} \sigma_{a_j} \rangle$. In the WFS design where V is sparse (as described above) these terms are zero ($\sigma_{\mathsf{a}}^2 \equiv \sum_i \sigma_{a_i}^2$). The total variance is written analogously to that for the individual amplitudes,

$$\langle \sigma_{\mathsf{a}}^2 \rangle = \sum_j \frac{1}{\overline{W}_{jj}^2}. \tag{4.10}$$

Several Monte Carlo simulations, using a NES model with simulated detector and photon noise, were used to find the variance of an ensemble of measured aberration amplitudes. The results were compared with the SVD predictions and found to give equivalent values.

4.2.4. Parameter Optimisation

The physical geometry of the NES, using square detector pixels, is specified by two parameters: the length of a pixel and the defocus diversity amplitude. To optimise a NES, the minimum total amplitude variance, as a function of these parameters, was found for detector readout noise and for photon noise. The method is via a search of parameter space, and is only practical because the SVD noise calculation is fast.

4.2.5. Comparison with Existing WFS Designs

The NES is compared with two WFS designs currently used in astronomical AO: the Curvature WFS and Quadrant Cell. The CWFS is representative of zonal sensors and, like the NES, measures the wavefront through intensity differences in two defocused planes. The Quadrant Cell is the most sensitive of the 3 sensors to tip/tilt errors and the WFS most commonly found in current interferometer AO.

Curvature WFS Model

The CWFS model also uses the propagation of light from the pupil to an intra-focal and an extra-focal plane. As the CWFS uses a geometric optics description of the defocused images, the defocus length must allow diffraction effects to be neglected. The details on the numerical methods of

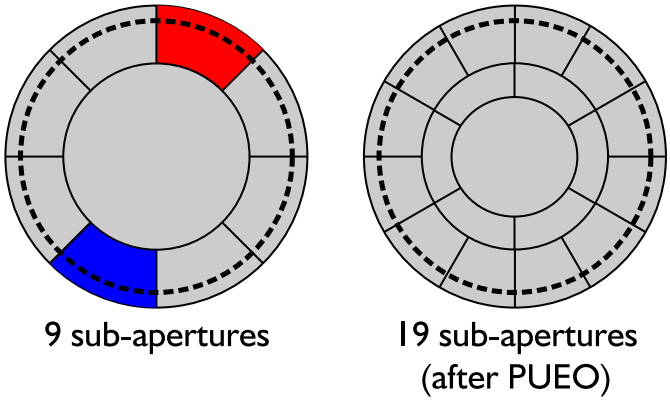

Figure 4.1.: The sub-aperture configuration for the simulated 19 sub-aperture and 9 sub-aperture Curvature WFSs. The dotted line represents the defocused pupil, and the red and blue shading one pair of diametrically opposed sub-apertures (see text for details).

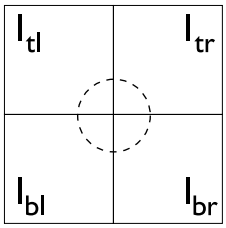

Figure 4.2.: The four detector pixels (quadrants) used in a Quadrant Cell. The dotted line represents the focused image.

propagation (in Appendix C) show that the propagation method first propagates the light to the focus and then from the focus to the defocused planes.

The defocus distance used is $d = 0.005\,f$, and was chosen to be near the focus whilst satisfying the geometric optics limit. Two designs of the CWFS sub-aperture geometry were used: one with 19 sub-apertures and the other with 9 sub-apertures. These correspond to high-order and low-order designs, and in particular the latter is chosen to measure only the first 5 Zernike polynomials after piston (Z_2 to Z_6). These sub-aperture configurations are shown in figure 4.1.

The WFS signal for a CWFS is the normalised intensity difference between planes,

$$s_i = \frac{I_{+;i} - I_{-;i'}}{I_{+;i} + I_{-;i'}}. \tag{4.11}$$

The extra-focal image is rotated by π, when compared with the inter-focal image, because it is beyond the focus. As a result, the terms referring to each plane in (4.11) correspond to intensity measurements in diametrically opposed sub-apertures; an example of such a pair are those shaded red and blue in figure 4.1. Otherwise the numerical method is as for a NES.

Quadrant Cell Model

The Quadrant Cell (QC) is modelled using intensities at the focal plane. The focused image is binned into four square quadrants, each representing a detector pixel as in figure 4.2. Let the intensity in the quadrants be written as: I_{tl} (top left), I_{tr} (top right), I_{bl} (bottom left), and I_{br} (bottom right). The signals for measuring the tip/tilt aberrations are then,

$$\begin{aligned} s_1 &= \frac{I_{tl} + I_{bl} - I_{tr} - I_{br}}{I_{tl} + I_{bl} + I_{tr} + I_{br}}, \\ s_2 &= \frac{I_{tl} + I_{tr} - I_{bl} - I_{br}}{I_{tl} + I_{bl} + I_{tr} + I_{br}}. \end{aligned} \tag{4.12}$$

To allow integration into a generalised AO simulation code, these signals were not used to directly infer the amplitudes of Z_2 & Z_3. Instead a measurement matrix (of dimension 2×2) was calculated and the tip/tilt amplitudes found via M^+, as for the CWFS and NES.

4.3. NES Results with Zernike Aberrations

Analysis of the SVD decomposition of the NES measurement matrix suggested that the optimum order to measure is $N = 5$ (Z_2 to Z_6), because higher order Zernike resulted in non-orthogonal difference signals within M. The SVD of the 9-element CWFS measurement matrix also showed the same lack of orthogonality for N greater than 5. The measurements using the CWFS and NES are of signals (curvature and intra/inter-plane respectively) which have an intrinsic circular symmetry. Note that because the limitation on the number of terms in Z_s is identical for the CWFS and NES, this does not occur as a consequence of a square pixel geometry. The remainder of this work uses a NES measuring these first 5 Zernike aberrations; Z_2 and Z_3 ($m-$odd), and Z_4, Z_5, and Z_6 ($m-$even, the quadratic aberrations).

The performance of a WFS design can be quantified by several metrics. The distinction between open loop and closed loop AO is important here; for the latter it is the remaining phase aberrations after correction that are important. If an aberration changes little whilst the closed loop AO system converges on its amplitude then the correction is still useful. Hence, there are two distinct categories of performance metric: those that give information on the WFS response to individual Zernike polynomials, roughly equivalent to operating with an open loop, and those concerned with phase aberrations in a closed loop. Both these sets will be compared with the equivalent results from a CWFS and a QC.

4.3.1. Wavefront Measurement Noise

The NES design is specified by two parameters, the detector pixels length and the defocus diversity amplitude. The total rms amplitude error was calculated for all combinations of pixel widths between 0.1 and 2.0 and diversity between 0.025 and 1.5 rad rms. This range of diversity is sufficiently small to allow propagation directly from the pupil.

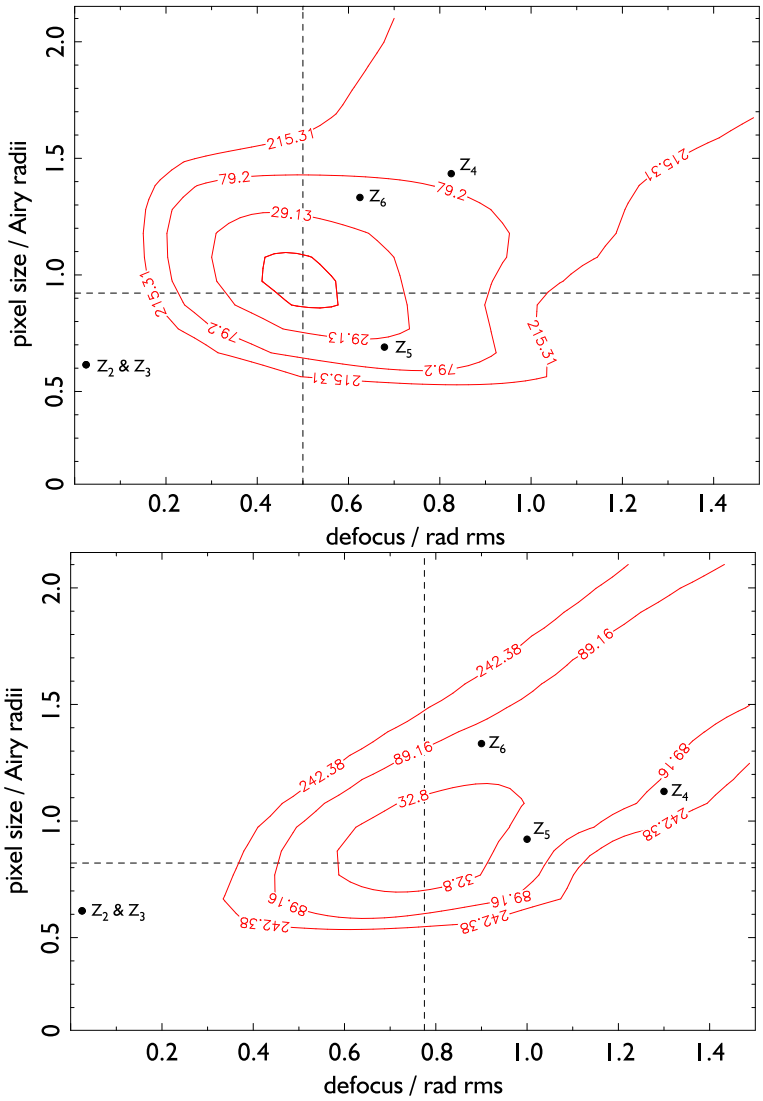

Figure 4.3.: NES total rms error for varying pixel width and diversity. The top plot is for photon noise and the bottom plot is for detector noise.

NES Optimisation and Measurement Noise

Two noise scenarios were considered; for photon noise, the input flux per plane is 1 photon, and for detector noise, the input flux per plane is 9 photons and the read noise per pixel is 1. The variances from using these values can be easily scaled to arbitrary fluxes and detector noise; for N photons, the 1 photon noise variances should be divided by N and the 1 detector read noise variances should be divided by N^2.

The rms error plots—a search of the 2D parameter space—are shown in figure 4.3 and the maximum SNR is marked by the intersection of the dashed lines. The minimum rms errors for each individual Zernike aberration are shown by filled points. The contours show a similar trend of error in both plots; as the defocus gets larger, the SNR is maintained only if the pixel size increases. There are also minimum limits on both parameters below which the error increases rapidly. The

Noise scenario	**Optimised parameters**	
	Defocus diversity / rad rms	Pixel length / Airy radii
photon noise-dominated	0.50	0.92
read noise-dominated	0.78	0.82

Table 4.1.: Optimised NES parameters corresponding to the two noise scenarios.

Parameters	**Amplitude rms error** / rad rms				
	Z_2	Z_3	Z_4	Z_5	Z_6
photon noise optimised	0.27	0.27	0.27	0.24	0.31
read noise optimised	0.41	0.41	0.23	0.26	0.27

Table 4.2.: RMS errors on the Zernike aberrations for the two parameter optimisations, for an input flux of 10 photons and no read noise.

result for diversity corresponds well with the sensitivity arguments for optimum defocus diversity in the previous chapter.

As expected from the dependence of the difference signals on the diversity, the error on the amplitude of Z_2/Z_3 is minimised for the smallest defocus amplitude whilst errors for the amplitudes of Z_4–Z_6 are minimised with finite defocus. The minimum total error is the compromise between the optimum for each aberration.

The optimised parameters for the two noise scenarios are shown in table 4.1. Qualitatively, the smaller pixel length and larger defocus is favoured for Gaussian noise statistics because the flux is spread out more evenly over the pixels. Conversely, when the noise has Poisson statistics the effect of noise is reduced by maximising the difference signals.

The central pixel only contributes to the difference signal for the defocus aberration, and this makes the minimum error on the amplitude of Z_4 more sensitive to both the pixel length and defocus. The astigmatisms, Z_5 and Z_6, show a minimum variance for distinctly different pixel lengths. This difference is explained by the square pixel geometry; Z_5 is measured by the corner pixels and Z_6 by the middle pixels along each edge. The two aberrations have similar errors when the distance between these pixels and the image centre is the same, so the pixel length favoured for Z_5 is smaller than for Z_6. A suitable circular pixel arrangement would return equal pixel dimensions. The rms errors on the amplitudes of each Zernike aberration are shown in table 4.2, when the input flux is 10 photons and only photon noise is present.

QC and CWFS Wavefront Measurement Noise

The corresponding errors for the QC and 19 sub-aperture CWFS are shown in table 4.3. The errors for the 9 sub-aperture CWFS were not significantly different. The QC has the lowest error for tip/tilt, as there is no defocus diversity, whereas the CWFS has the highest error and the largest diversity. The errors on the $m-$even aberration amplitudes measured by the CWFS are all

Wavefront	**Amplitude rms error** / rad rms				
Sensor	Z_2	Z_3	Z_4	Z_5	Z_6
Quadrant Cell	0.16	0.16	-	-	-
Curvature WFS	30.77	30.77	9.16	13.26	12.78

Table 4.3.: Quadrant Cell and Curvature WFS amplitude measurement errors, for an input flux of 10 photons and photon noise, c.f., table 4.2.

higher than the equivalent results for the NES, as expected with a smaller defocus diversity for the latter. The small difference between the errors of Z_5 and Z_6 from the CWFS is due to the specific sub-aperture geometry used here.

4.3.2. Linearity

The linearity of the measured amplitudes is the maximum amplitude for which the measurement is unaffected by non-linearities in the difference signals. This is the limit with which a NES can accurately measure aberrations in an open loop AO system. A definition of linearity is the amplitude at which the measurement deviates by more than 5% from the actual value,

$$\text{linear if } 95\% \leq \frac{|a_{measured}|}{|a_{input}|} \leq 105\%. \tag{4.13}$$

NES Linearity

The response of the NES to individual aberrations is plotted in figure 4.4, one plot for each set of optimised values. As the response is symmetrical for positive and negative amplitudes, the input amplitude range plotted here is from 0 to 1.0 rad rms.

For small amplitudes, ≤ 0.1 rad rms, the response is linear as predicted. The non-linearity of the NES in both plots occurs for amplitudes greater than ~ 0.2 rad rms, and the response for each aberration is significantly different. To investigate the sensitivity of the linearity to the choice of NES parameters, the linearity to 5% was calculated for each aberration by varying one parameter whilst keeping the other fixed i.e., along the dotted lines in figure 4.3. The results are shown in tables 4.4 and 4.5.

As the defocus diversity is increased, the linearity for all aberrations increases and in particular for Z_2/Z_3. This change for tip/tilt is because the defocus increases the image size so reduces the image movement as a fraction of the image width. Hence, as non-linearities occur at roughly a constant image offset relative to its width, the linearity limit increases. The increases seen for Z_5 and Z_6 are similarly explained. However, measurements of Z_4 use the central pixel intensity and the decrease of linearity when the diversity is 1 rad rms shows this aberration's increased sensitivity to non-linear effects in that pixel.

The change in pixel length shows that for small sizes the linearity of the $m-$even aberrations are increased. At these scales, the majority of the difference signals come from the Airy disc itself but are of a small amplitude. For widths above ~ 0.7 Airy radii, there is little change in linearity

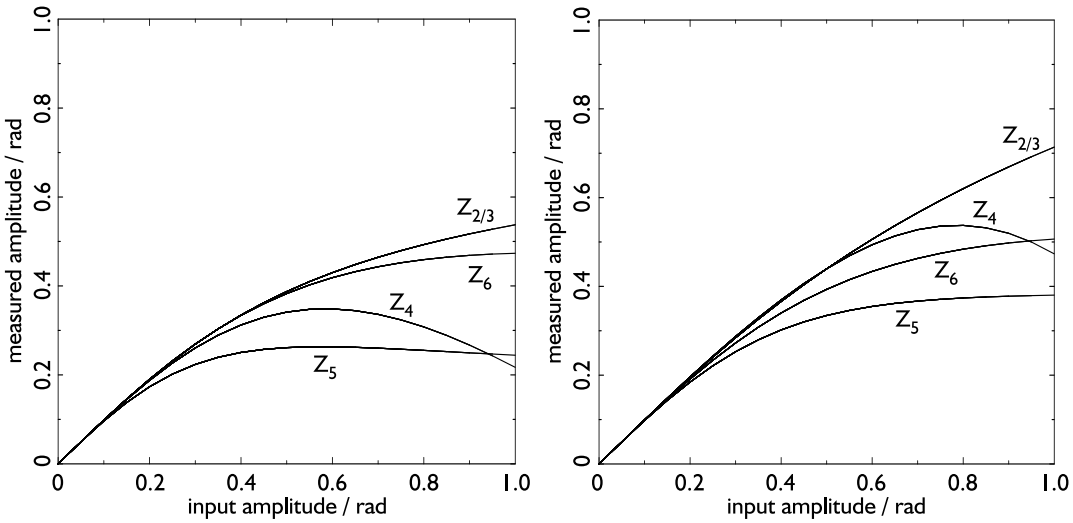

Figure 4.4.: Response of NES for a given input Zernike aberrations with varying amplitudes. The left-hand plot corresponds to a photon-noise optimisation and the right-hand plot to detector-noise optimisation.

Defocus	**Linearity limit** / rad rms				
/ rad rms	Z_2	Z_3	Z_4	Z_5	Z_6
0.20	0.17	0.17	0.11	0.10	0.17
0.30	0.18	0.18	0.13	0.10	0.17
0.40	0.19	0.19	0.15	0.11	0.18
0.50	0.21	0.21	0.18	0.12	0.18
0.60	0.24	0.24	0.22	0.14	0.19
0.70	0.30	0.30	0.28	0.15	0.20
0.80	0.38	0.38	0.32	0.17	0.21
0.90	0.56	0.56	0.30	0.19	0.22
1.00	0.89	0.89	0.23	0.21	0.23

Table 4.4.: The limit of linearity to 5% for a central pixel width of 0.92 Airy radii and varying defocus diversity.

until the central pixel approaches a Airy disc radius. Then the linearity of only the tip/tilt terms increases significantly because the image motion relative to the pixel size is now reduced.

QC and CWFS Linearity

The CWFS is linear, as expected for a zonal WFS, for aberration amplitudes (tested to a limit of 2 rad rms) smaller than the amplitude of its diversity defocus. The QC is more comparable with the NES, being a sensor design with a limited linearity range. Its linearities for measuring Z_2 and Z_3 are equal at 0.4 rad rms.

Pixel width	**Linearity limit** / rad rms				
/ Airy radii	Z_2	Z_3	Z_4	Z_5	Z_6
0.41	0.27	0.27	0.61	0.37	0.61
0.51	0.21	0.21	0.61	0.21	0.61
0.61	0.20	0.20	0.36	0.15	0.36
0.71	0.20	0.20	0.21	0.13	0.26
0.82	0.20	0.20	0.19	0.12	0.21
0.92	0.21	0.21	0.18	0.12	0.18
1.02	0.22	0.22	0.18	0.12	0.17
1.12	0.22	0.22	0.17	0.12	0.16
1.23	0.24	0.24	0.17	0.12	0.15
1.33	0.31	0.31	0.17	0.13	0.15
1.43	0.61	0.61	0.17	0.13	0.15

Table 4.5.: The limit of linearity to 5% for a defocus diversity of 0.5 rad rms and varying pixel width.

Discussion

The linearity limit of the NES is sensitive to the choice of parameters, and only general comments about the change of linearity with parameter variation can be made. When the aberrations have amplitudes of $\lesssim 0.1$ rad rms then the NES measures each aberration linearly. For larger errors, the NES is still expected to provide stable measurement in closed loop AO since the measured amplitudes do not suffer a sign change. The results from the QC are an example of this; the measurements from this sensor saturate at ± 0.5 rad rms so the AO correction converges on the actual tip/tilt aberrations when they are larger than this limit.

4.3.3. Non-linearities

There are two quantities which will describe the effects of non-linearities within the difference signals: cross-coupling and aliasing. The cross-coupling error is caused by the amplitudes of aberrations in Z_s causing difference signals which are significantly non-linear and so corrupt the measured amplitudes of terms in Z_s. Aliasing occurs when difference signals from higher order aberrations, not in Z_s, produce difference signals which result in spurious measured amplitudes.

Both of these errors can be quantified from the amplitude orthogonality matrix,

$$E_{ij}(\alpha, \beta) = a'_i - \alpha \text{ for } \Phi = \alpha\, Z_i + \beta\, Z_j, \tag{4.14}$$

where $Z_i \in Z_s$. Each element is the amplitude measured for Z_i when the input aberration is Z_i and Z_j, for arbitrary input amplitudes α and β. The matrix from an ideal WFS is diagonal, $E_{ij} = \delta_{ij}\beta$, and the deviation from this measures the non-linearity. The cross-coupling error, X_j, and aliasing error, L_j, can be defined as,

$$\begin{aligned} X_{ij}(\alpha,\beta) &= E_{ij}(\alpha,\beta) \text{ for } Z_j \in Z_s, \\ L_j(\beta) &= \sum_i E_{ij}^2(0,\beta) \text{ for } Z_j \notin Z_s. \end{aligned} \tag{4.15}$$

The definitions are such that each quantity is zero when the error it measures is not present. With an aperture diameter of $D/r_0 = 5$, the 2-σ amplitude limit of the 3 quadratic $m-$even aberrations is ~ 1 rad rms. Hence, the amplitudes of aberrations considered will be 0.1 rad rms (representing small aberrations) 0.5 rad rms, and 1.0 rad rms. The effectiveness of the centring algorithm (see chapter 3) to break cross-coupling between tip/tilt and the $m-$even aberrations can also evaluated from the cross-coupling error. For the NES, the cross-coupling and aliasing were calculated for both sets of optimised parameters.

Cross-coupling

The results for NES cross-coupling were similar for either parameterisation and for varying α. To clarify the results, only $X_{ij}(0.1,\beta)$ where β is 0.5 and 1.0 rad rms will be shown. The cross-coupling when β was 0.1 was less than 3×10^{-3} rad rms and verifies that using finite-sized detector pixels allows the calculation of orthogonal difference signals for small amplitude aberrations.

The cross-coupling shown in table 4.6, for the photon noise-optimised parameters, is largest between the astigmatisms and defocus but smaller between the two astigmatisms. This $m-$even behaviour occurs because the defocus aberration is the only term measured using all 9 pixels and so is most sensitive to cross-coupling. It is problematic for these $m-$even terms when $\beta = 1.0$ rad rms as the coupling amplitude is larger than the actual $m-$even amplitude of 0.1 rad rms. The cross-coupling between tip/tilt and the $m-$even aberrations is relatively small, and with the centring algorithm it was removed entirely.

In comparison, the 9 and 19 sub-aperture CWFS designs show little cross-coupling, $X_{ij}(0.1, 1.0)$ $\leq 5 \times 10^{-3}$ rad, and the QC shows no cross-coupling because the signals in (4.12) are defined to be orthogonal.

Aliasing

The NES aliasing was determined by using one Zernike aberration Z_j (which is not in Z_s) and examining all the measured amplitudes (which are in Z_s). The aliasing was found to be onto one aberration i.e., there was a one-to-one mapping from Z_j to some aberration in Z_s, and this justifies the use of the summation in the expression for L. The reason for this unique mapping is that the difference signals from the aberration Z_j are aliased by the finite NES detector pixels, and the symmetry of the difference signals becomes imposed onto the symmetry of the NES pixel grid. There can then only be an amplitude confusion with the aberration in Z_s whose difference signals also have the same measured symmetry.

As with the cross-coupling, the aliasing for either parameterisation of the NES was found to have similar characteristics and the results are given in table 4.7 for the photon noise-optimised parameters. Overall, the aliasing from all aberrations is small and does not increase significantly

$X_{ij}(0.1, 0.5)$			**j**		
i	2	3	4	5	6
2	-	0.02	0.02	0.01	0.01
3	0.02	-	0.02	0.01	0.01
4	0.02	0.02	-	0.05	0.05
5	<0.01	<0.01	0.05	-	0.02
6	<0.01	<0.01	0.05	<0.01	-
$X_{ij}(0.1, 1.0)$			**j**		
i	2	3	4	5	6
2	-	0.04	0.04	0.04	0.04
3	0.04	-	0.04	0.04	0.04
4	0.04	0.04	-	0.12	0.12
5	0.02	0.02	0.12	-	0.03
6	0.02	0.02	0.12	0.03	-

Table 4.6.: Cross-coupling for the photon noise-optimised NES; $X_{ij}(0.1, 0.5)$ and $X_{ij}(0.1, 1.0)$.

β	$\mathbf{L_j}$								
	7	8	9	10	11	12	13	14	15
0.25	0.05	0.05	<0.01	<0.01	0.04	0.05	0.03	0.01	<0.01
0.50	0.07	0.07	0.02	0.02	0.06	0.09	0.08	0.01	<0.01
0.75	0.10	0.10	0.02	0.02	0.07	0.14	0.12	0.01	<0.01
1.00	0.11	0.11	0.01	0.01	0.06	0.17	0.17	0.01	<0.01

Table 4.7.: Aliasing, $L_j(\beta)$, for the photon noise-dominated optimised NES design (j corresponds to Z_7 to Z_{15} inclusive).

with increasing amplitude. The significant aliasing is from Z_7/Z_8 onto Z_2/Z_3, and Z_{12}/Z_{13} onto Z_6/Z_5. From a phase screen with Kolmogorov statistics, the variance of Z_{12}/Z_{13} is ~ 0.1 times smaller than the (uncorrected) variance of Z_6/Z_5, so these effects can be neglected. However, this last point is invalid if Z_6/Z_5 are well corrected, so the aliasing will affect the accuracy of the wavefront correction of these terms in a closed loop.

The aliasing on the QC measurement occurs only from Z_7/Z_8 and has approximately the same effect as for the NES: $L_{7/8}(1.0) \sim 0.15$. The 19 sub-aperture CWFS has no aliasing because it already measures Z_7 to Z_{15}.

The 9 sub-aperture CWFS measures the same 5 aberrations as the NES, and its aliasing is shown in table 4.8. The aliasing, when it is not small, is significant and particularly so for Z_7/Z_8. The NES detector pixels effectively "blur" the difference signals from high order aberrations so aliasing effects are relatively small. This does not happen for the sub-apertures of either CWFS

$\mathbf{L_j(1.0)}$								
7	8	9	10	11	12	13	14	15
2.59	2.59	0.24	0.24	< 0.01	0.62	0.56	< 0.01	0.22

Table 4.8.: Aliasing, $L_j(1.0)$, for the 9 sub-aperture CWFS.

design. For comparison, the 19 sub-aperture CWFS was modelled with a measurement matrix containing only the first 5 Zernike aberrations and the aliasing results were similar to those from the 9 sub-aperture version. The conclusion is that a low order CWFS requires a minimum number of sub-apertures (greater than 9 but less than 19) to avoid deleterious aliasing effects with Zernike polynomial aberrations.

4.3.4. Summary

The non-linearities of the NES have been shown to be mainly from cross-coupling of the $m-$even aberrations, that is between defocus, astigmatism 1 and astigmatism 2. The amplitudes of these cross-couplings is relatively small, however, and the actual amplitudes of these aberrations will be small in a closed loop. These effects, together with the linearity limit, suggest that the NES is only suitable for small diameter telescopes, $D/r_0 \simeq 5$. The equivalent results from the QC are comparable and in some cases are worse than the NES results. For example, the NES response to tip/tilt—which are the largest variance aberrations from the atmosphere and so the most important to correct—do not saturate for ≤ 1 rad rms whereas the QC response saturates at 0.5 rad rms. The expected improvement over the QC, in a closed loop, will be quantified in the next section. In comparison, the 9 sub-aperture CWFS was shown to suffer from severe aliasing because of its poor sampling. Further results using the CWFS are only from use of the 19 sub-aperture design.

4.4. Atmospheric Aberration Results

The discussion so far has considered the effect on measurements where the phase aberrations consist of one or two Zernike polynomials, which is an idealised situation. The results from using closed loop simulations with atmospheric aberrations will provide the evaluation of how the NES will perform in an interferometer AO system. Using the Strehl ratio to quantify the level of correction, comparisons will be made between using the NES, CWFS, and QC within the AO model. The parameter space of measurement noise and seeing conditions explored will provide results relevant to the aperture sizes of AO-equipped interferometers. In particular, it can be assumed that COAST-sized apertures are approximated with $D/r_0 \simeq 3$ at near-IR wavelengths. In chapter 6, it is shown that poor seeing corresponds to $r_0(500\,\text{nm}) \simeq 5\,\text{cm}$ at the COAST site and, as the telescope apertures have diameters $D = 24\,\text{cm}$, this leads to the seeing value given above. The results from this section also allow the effects of cross-coupling and aliasing errors on closed loop correction to be understood.

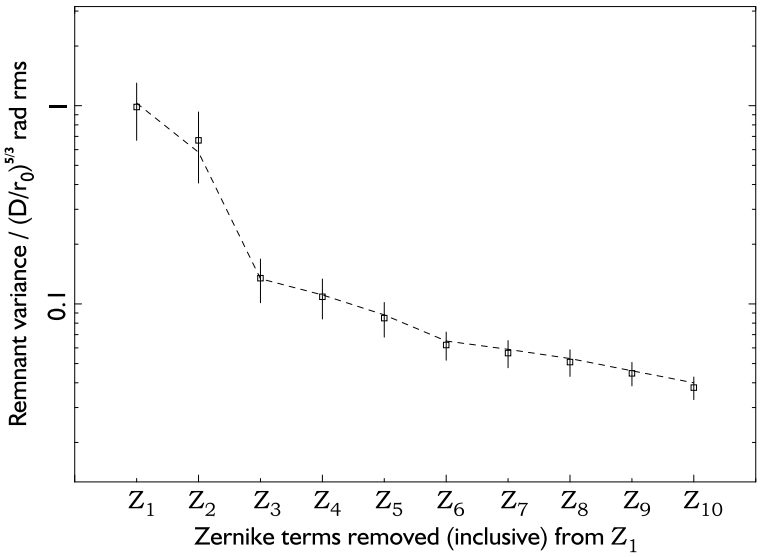

Figure 4.5.: The remnant rms phase error after the inclusive removal of Zernike polynomials from Z_1 to that indicated, with the theoretical value shown by the dotted line ($D/r_0 = 1$). Note this particular example, used for the simulations, does not show the oft-seen systematic reduction in variance for Z_2 and Z_3.

4.4.1. Method of Closed Loop AO Simulations

The simulation of an AO closed loop involves calculating the atmospheric phase aberrations across the aperture, applying a correction to this phase, and measuring the remnant (corrected) phase with a WFS model. This allows an evaluation of how well a WFS can measure and correct a temporally evolving wavefront. Increasing the "strength" of the seeing, by decreasing r_0 relative to D, allows the limiting aperture size to be determined by this method. The other limit is the faintest guide star for which the WFS correction will not be dominated by noise and can be established by specifying the input flux and modelling photon and detector read noise appropriately.

The time evolution of the aberrations is simulated by using a large phase screen and Taylor's hypothesis. For each iteration in the simulation, a phase screen (size $L \gg R$) has an offset segment extracted. This segment is then multiplied by the pupil transmission function and this defines the complex amplitude in the pupil plane. At the end of the iteration the offset is incremented, and this causes the screen to "flow" past the aperture. The wavefront correction is implemented by a "DM" which subtracts from the wavefront a sum of Zernike aberrations, as reconstructed from previous WFS measurements. The corrected phase then forms the wavefront at the input aperture of the WFS.

The phase screen is obtained from a discrete Fourier transform of white noise modulated by a power spectrum (46) which generates Kolmogorov spatial statistics. A common criticism of this approach is that spatial frequencies below the inverse of the array dimensions are not incorporated into the final phase screen. These frequencies also have the largest amplitudes, so their neglection causes a reduction in the variance of low order Zernike aberrations. However, because $L > R$, the absent spatial frequencies are at a scale larger than the aperture diameter so power in the Zernike polynomial aberrations (excepting Z_2 and Z_3) are not adversely affected. Figure 4.5 shows

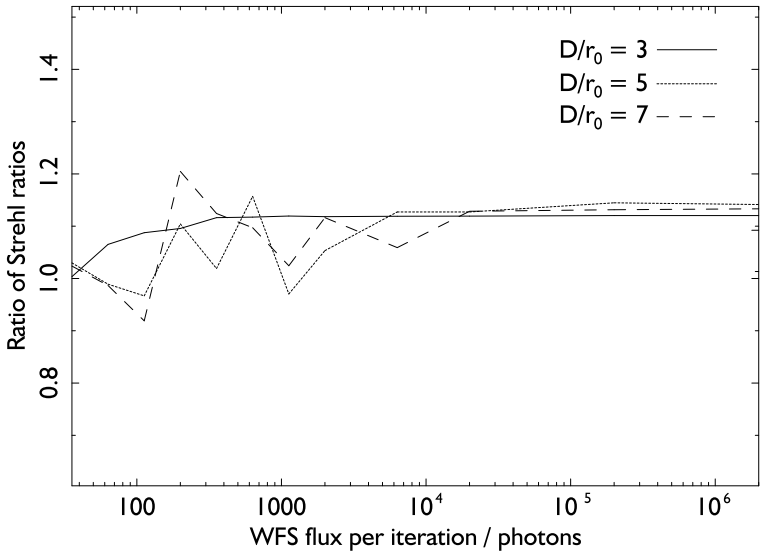

Figure 4.6.: The mean Strehl ratio without centring divided by the mean Strehl ratio with centring, when using the NES in a closed loop for three seeing conditions.

an example of the rms phase error after removing increasing numbers of Zernike polynomials from a phase screen (size of $L = 2048$) that is circularly-truncated by an aperture (radius $R = 48$). Although in general there is a systematic reduction in the variance of Z_2 and Z_3, the rms errors for this particular phase screen are consistent with the theoretical predictions. The following results also use this phase screen, and it is considered representative of a single-phase screen atmosphere.

4.4.2. Evaluation of the NES Centring Algorithm

The centring algorithm was shown to be effective in breaking the cross-coupling between the tip/tilt and $m-$even Zernike aberrations for the NES with single measurements. The change in the mean Strehl ratio when using the centring algorithm in a closed loop AO simulation allows its effect to be evaluated when higher order $m-$odd Zernike aberrations are also present.

The NES parameters used in the following simulations are a defocus diversity of 0.6 rad rms and a pixel width of 1 Airy radius. Three values of the seeing were used: in terms of D/r_0, 3, 5, and 7. For each value of D/r_0, the closed loop simulations were run with simulated photon noise and repeated for various fluxes. After each iteration, the Strehl ratio was calculated and at the end of a simulation the ensemble mean was returned. The results presented are the average of these ensemble means for 5 independent starting positions on the phase screen. This means each run, for given input flux and seeing, is further averaged (effectively) over the phase screen. The final results at a flux of 10^6 differed by less than 10% when calculated using a different phase screen.

A plot of the mean Strehl ratio without using the centring algorithm divided by Strehl ratio with the centring algorithm applied (the ratio of the Strehl ratios) is shown in figure 4.6. The surprising result is that the ratio is always greater than one, indicating that the NES measurements are better without using the centring algorithm.

To investigate whether the centring algorithm was affected by higher order $m-$odd aberrations, the simulations were run again but Zernike aberrations higher than Z_6 were removed before being

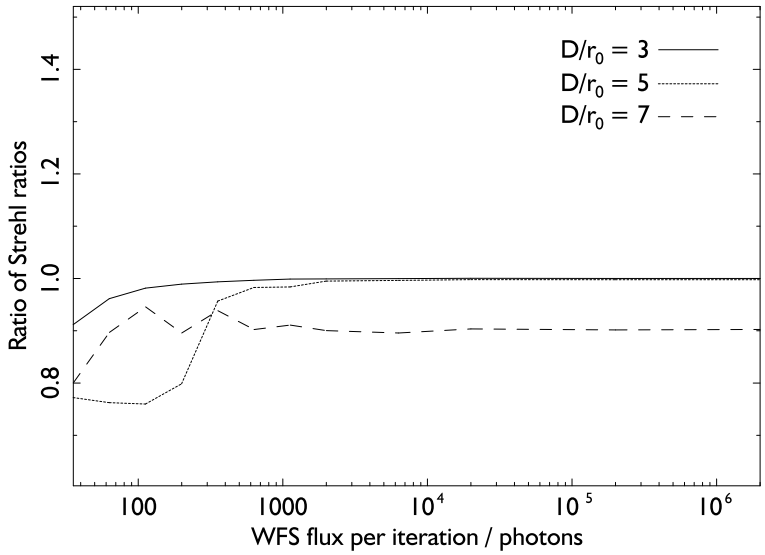

Figure 4.7.: The mean Strehl ratio without centring divided by the mean Strehl ratio with centring c.f., figure 4.6 but higher order phase aberrations than Z_6 were removed prior to being measured.

	RMS of Zernike aberration amplitudes / rad rms				
	$m-$odd			$m-$even	
	Z_2	Z_3	Z_4	Z_5	Z_6
actual	0.587	0.327	0.414	0.155	0.177
remnant with centring	0.198	0.095	0.244	0.049	0.054
remnant without centring	0.198	0.095	0.262	0.068	0.078

Table 4.9.: The variance of individual aberration amplitudes for 500 open loop measurements with the NES.

measured by the WFS. The remaining aberrations, comprising of Z_2 to Z_6 inclusive, were then used as before. The resulting ratio of Strehl ratios is shown in figure 4.7, and can now be seen to be less than or equal to one. It can therefore be assumed that it is the first higher order $m-$odd aberrations after tip/tilt i.e., Z_7/Z_8 (coma), which causes the centring algorithm to fail in a closed loop. This aberration is also the highest variance term after tip/tilt.

Finally, the use of a NES for measurements (with no correction) of aberrations on one moving phase screen was analysed. The phase amplitudes were small, corresponding to $D/r_0 = 1$, and noise was not simulated. The resulting amplitudes are equivalent to those measured if a NES was used in an open loop AO system. A comparison can be made between the measured and actual aberrations by examining the actual variance of the aberrations and the variance of these aberrations when corrected by the NES measurements (again, effectively open loop AO). These remnant variances with and without use of the centring correction are displayed in table 4.9, and demonstrate that the centring algorithm does improve the NES measurements by a small amount, even when higher order $m-$odd Zernike aberrations are present.

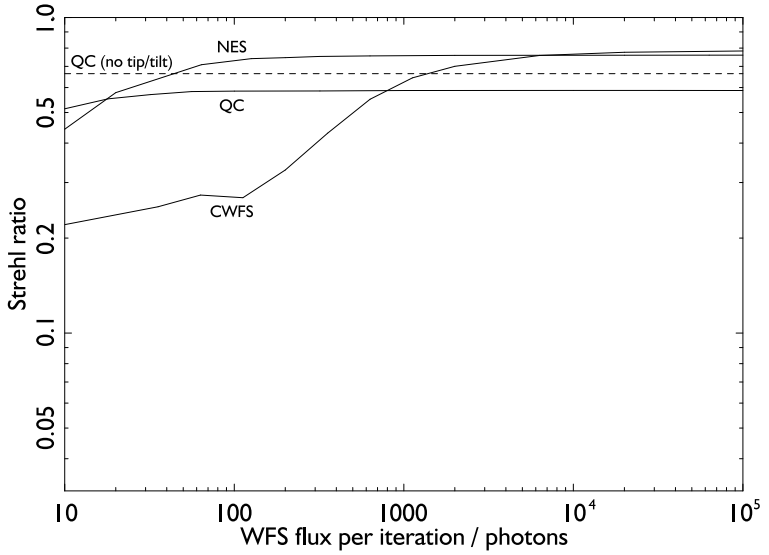

Figure 4.8.: Mean Strehl ratios from AO simulations with a CWFS, NES, and QC for $D/r_0 = 3$. The simulations include photon noise only. The dashed line is the equivalent Strehl ratio when remnant tip/tilt aberrations are removed (for clarity, only shown for the QC in this figure.)

It can be concluded that the use of the centring algorithm is only beneficial when making measurements of uncorrected aberrations or when the image motion is particularly large. In a closed loop the assumption that the $m-$odd phase is dominated by the linear tip/tilt terms can become invalid when these terms are well corrected and then small. The naïve centring of intensities then results in the centring error causing larger measurement errors than the cross-coupling. Accordingly, the remainder of the closed loop simulations in this chapter do not use the centring algorithm.

4.4.3. Comparison of the NES with a CWFS and QC

Having established how to optimally control a NES, the closed loop simulations comparing the use of the NES with the 19 sub-aperture CS and QC were undertaken. These use identical numerical models except for the differing WFS model. The physical parameters simulated were an aperture of 1 m, a wind speed of $10\,\mathrm{ms}^{-1}$ and an AO loop iteration time of 1 ms. The simulations were calculated for three values of seeing, D/r_0 equal to 3, 5, and 7. These correspond to good, intermediate, and poor seeing in the near-IR as might be expected at a high-altitude site. The input flux is quoted in photons per iteration and can be scaled to photons per second by multiplying by 10^3.

The results are plotted in figures 4.8, 4.9, and 4.10 for each seeing value and WFS separately. The solid lines represents the Strehl ratio with the correction measured by each WFS and the dashed lines are the Strehl ratio after any remaining tip/tilt aberrations are removed. As tip/tilt are the most important aberrations to correct, a comparison between the dashed and solid lines reveals how accurate the tip/tilt measurement of a WFS is.

In all the figures, the sensitivity of each sensor is apparent. In terms of decreasing sensitivity, the sensors can be ordered as the QC, NES, and CWFS, as expected from the SNR results above. Comparing the limiting fluxes below which the correction with each WFS degrades significantly:

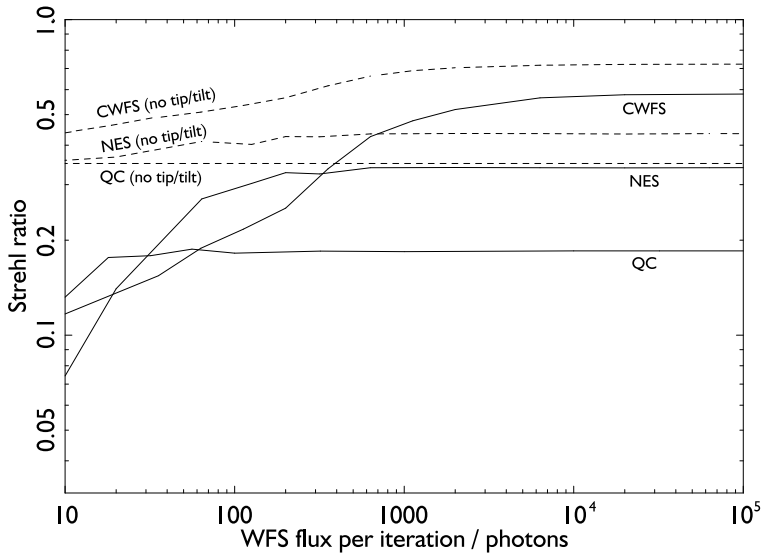

Figure 4.9.: Mean Strehl ratios from AO simulations with $D/r_0 = 5$, otherwise as figure 4.9.

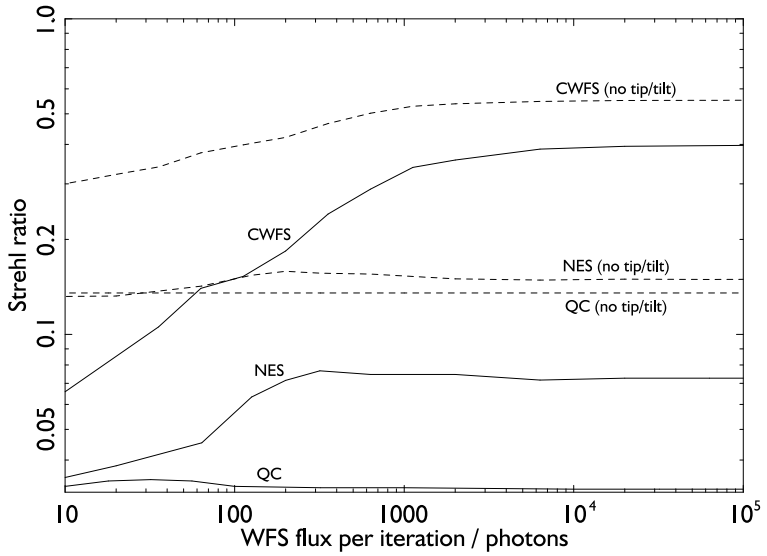

Figure 4.10.: Mean Strehl ratios from AO simulations with $D/r_0 = 7$, otherwise as figure 4.10.

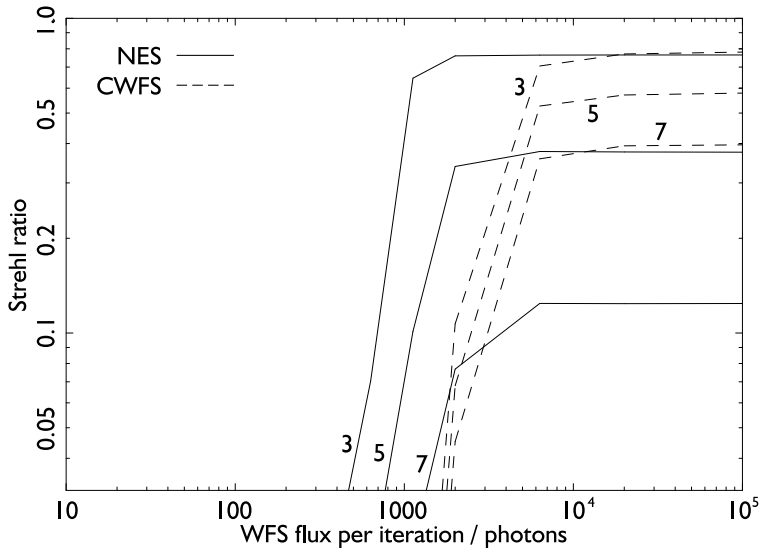

Figure 4.11.: Mean Strehl ratios from AO simulations with a CWFS (dashed) and NES (solid). The read noise is equivalent to 15 photons per pixel. The numbers refer to the value of D/r_0.

for the CWFS it is ~ 1000 photons per iteration, the NES limit is ~ 100 photons per iteration, and with the QC it is ~ 20 photons per iteration. The consequence on the Strehl ratios can be seen most clearly for $D/r_0 = 3$. The correction with the CWFS is hardly better than the NES at high light levels but is lower for fluxes below ~ 800 photons per iteration. Only at low light levels, ~ 20 photons, does the QC do better than the NES.

The results with average seeing conditions, $D/r_0 = 5$, also shows the usefulness of the NES. Although at high light levels the CWFS AO simulation gives higher Strehl ratios, the reason is simply its higher order correction. Examining the difference between the dashed and solid lines for the NES and CWFS, the high light level differences are smaller for the NES (0.09) than the CWFS (0.14). Therefore the NES-based tip/tilt correction is better, and improves relatively as the flux is reduced. The QC results compared to the other two sensors show the advantage of using low order AO. At high light levels, the NES is better than the QC and almost equivalent to perfect tip/tilt correction (the dashed line for the QC). Only when the flux is smaller than ~ 20 photons is the QC the best performing sensor.

In poor seeing conditions, $D/r_0 = 7$, it becomes clear that a high order WFS is necessary. The correction with the CWFS is the best regardless of the flux. The NES and QC, whose results are significantly worse than the CWFS, are more easily compared with each other. Comparing their dashed lines, there is little difference and it can be surmised that NES measurements of $m-$even aberrations contributes little to wavefront correction. The NES is better, however, than the QC at measuring tip/tilt at all fluxes, although not significantly so below ~ 100 photons per iteration. The saturation of the QC measurement for tip/tilt aberrations larger than 0.5 rad rms is probably the most important explanation for this difference. Furthermore, the aliasing of coma in the NES simulates tip/tilt errors and, although it results in "wrong" measurements, it leads to a rise in the Strehl ratio.

The simulations were ran again for the NES and CWFS, with read noise equivalent to 15 photons per detector pixel. The mean Strehl ratios are plotted in figure 4.11 for the three seeing conditions. At the highest light levels, the results are equivalent to those previously discussed. The relative sensitivity of the two sensors is clearly demonstrated by the drastic reduction in Strehl ratio below a certain flux; $\lesssim 4000$ photons for the CWFS and $\lesssim 2000$ photons for the NES. Interestingly, for better seeing the limiting flux for the NES is reduced, not because the noise changes but instead because the wavefront convergence time decreases and the difference signals tend to saturate less often. Conversely, as the seeing worsens the gain of the WFS is (effectively) reduced and so the effects of noise become worse.

4.4.4. Discussion

The conclusion to be drawn from these results is that each sensor is optimal in different conditions. In poor seeing, high order wavefront correction is needed with linear wavefront measurements so a CWFS is the best choice. In better seeing, only when the interferometer target is faint does using the QC become the sensible choice. The NES is the best sensor otherwise; in good and moderate seeing and for relatively faint targets it is the best choice. As the good seeing case corresponds to average or poor seeing as measured at the COAST site i.e., it is frequently better, it can be concluded that the NES is a suitable WFS for use in AO with this interferometer. The comparative limiting magnitudes (with zero read noise detectors) can be estimated for the three WFSs from these results: if all flux in I-band is used from a 1 m aperture with 0.5 arcsec seeing, the CWFS, NES, and QC limits are $m_I \simeq 12$, 15, and 17 respectively.

4.5. Integrated Spatial Filtering Results

The concept of integrating a spatial filter into a WFS will now be discussed. The integrated filter and WFS is called by the WFS name prefixed with 'SF' e.g. SF-NES. The proposed scheme in the previous chapter, which was to place a pinhole spatial filter at the focus and reflect the rejected light to the WFS, was evaluated in a numerical simulation. These simulations modelled the complex amplitude around a pinhole in the focal plane and propagated them back to a pupil plane where the defocus diversity could be added. The poor sampling of amplitude around the central Airy disc in the focal plane (see Appendix C) meant that this SF-NES model could not be made to work.

An alternative approach was used to model the SF-NES. The light from the telescope aperture is propagated to an intra-focal and extra-focal plane, and the central portion of the complex amplitude—corresponding to a transparent hole—in the intra-focal plane is further propagated to a pupil plane. This pupil then contains the filtered light, and overall is equivalent to using a defocused pinhole spatial filter. The intensity in the intra-focal and extra-focal planes is calculated with that in the central "hole" removed.

The optimum diameter for a round pinhole spatial filter at the focus, when considering the visibility SNR, is 0.82 Airy disc diameters (34). This is larger than the size of the NES pixels used above (table 4.1) so only the outer 8 pixels of the SF-NES need be modelled. These 8 pixels are then extended to the edge of the numerical array to maximise the detected flux, pictorially

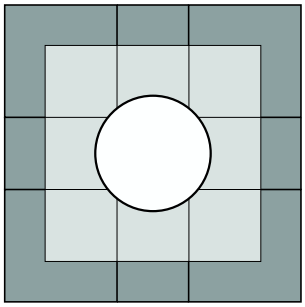

Figure 4.12.: Outline of the pixels of the NES (light grey), the SF-NES (dark grey) and the pinhole.

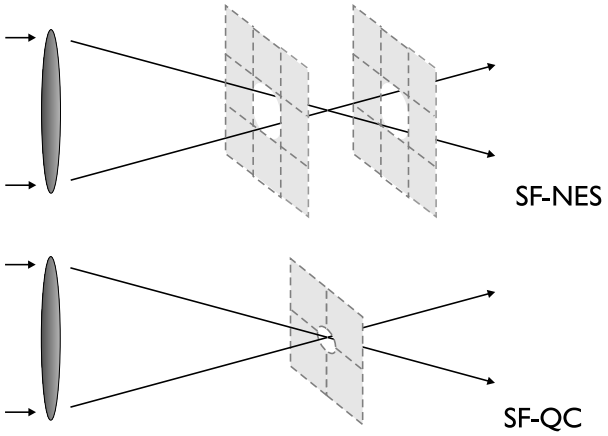

Figure 4.13.: Optical outline of the SF-NES and SF-QC operate. This diagram does not correspond to a physical design of the SF-NES.

shown in figure 4.12. The size of the "central pixel" (which is encompassed by the pinhole) is 1 Airy disc radius. It is noted that the size of the pinhole was not specifically optimised because the small defocus diversity of the SF-NES (0.5 rad rms) means that the intra/extra-focal images were assumed to be blurred focal plane images. As a comparison with the SF-NES, a QC integrated with the same sized pinhole—a SF-QC—was used. A graphical comparison of the two SF-WFSs is shown in figure 4.13.

Numerical simulations of a closed loop AO system using each sensor, without simulated noise, were run for varying seeing conditions. As before, the final results for each seeing value are an average of runs with identical seeing but different starting points in the phase screen. The Strehl ratio was calculated for both the corrected input beam and the spatially filtered light. These ratios are closely correlated, since if the input beam is poorly corrected then the throughput of the filtered beam is low. The effect of throughput can be seen if the on-axis intensity of the filtered light is plotted, and this quantity is simply related to visibility measurements; the on-axis intensity is effectively the Strehl ratio ($\mathcal{S}$) multiplied by the total flux ($\mathcal{N}$) whereas the SNR of a visibility measurement is $\sqrt{\mathcal{N}V^2}$, assuming photon noise only. Using the Strehl ratio as a proxy for visibility then allows the on-axis intensity to be used as a proxy for the visibility SNR. The Strehl ratios and

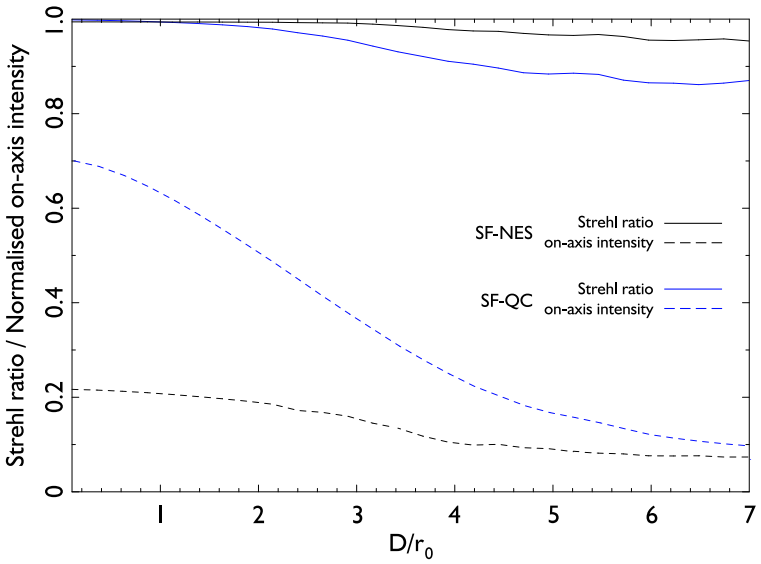

Figure 4.14.: Strehl ratios and normalised on-axis intensities for the filtered light from a SF-QC and a SF-NES.

on-axis intensity, the latter normalised by the intensity obtained from an unfiltered aberration-free beam, is shown in 4.14.

The filtered light Strehl ratio shows that the SF-NES is a better spatial filter, in terms of reducing the phase aberration variance, than the SF-QC. As the Strehl ratios for both WFSs are always very high, the on-axis intensity can be approximated as the throughput and this shows that the SF-NES has a lower—though more consistent—throughput than the SF-QC. The first result is because the more accurate wavefront sensing by the NES improves the image-pinhole alignment more than the defocus reduces the spatial filtering effect. The second result is because the defocused images couple less well to the pinhole than a focused image. At the expense of less effective spatial filtering, this last point may be alleviated if the pinhole size is increased.

The effect of the pinhole on the wavefront measurements of the SF-QC and SF-NES was also evaluated from closed loop simulations. The correction applied to the input beam of the WFS was calculated for the SF-QC, SF-NES, and the equivalent WFS-only designs. The resulting Strehl ratios are shown in figure 4.15. As expected, correction using the NES is always better than correction with the QC. The SF-QC makes marginally worse measurements than the QC, and the Strehl ratios for the QC WFSs essentially overlap for all seeing values considered here. The SF-NES measurements are noticeably worse than those from the NES, and also the SF-QC/QC for $D/r_0 \gtrsim 3.5$. Also, the Strehl ratio from using the SF-NES shows greater variation than the other WFSs. The central pixel in the NES is clearly important for measuring defocus, and without it the correction with the SF-NES becomes unstable. The cross-coupling of defocus to the astigmatisms, as shown in section 4.3.3, makes correction of the defocus aberration particularly important.

4.5.1. Discussion

The results of integrating a pinhole spatial filter into a WFS are encouraging: the resultant filtered light always has a higher Strehl ratio than the input beam. The throughput of the filter is markedly

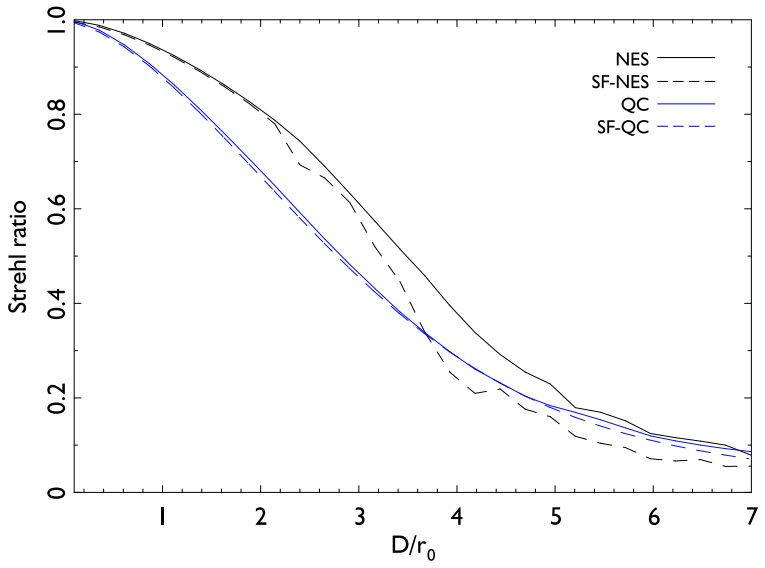

Figure 4.15.: Strehl ratios for the input (unfiltered) beam, using the integrated spatial filter WFS and the equivalent WFS-only designs.

different between the SF-QC and SF-NES, and is higher for the former WFS but more consistent for the latter. The impact of removing the central part of the image on the wavefront sensing is negligible for the QC but significant for the NES. Together these results suggest that using a pinhole integrated into a Quadrant Cell gives the benefits first suggested: automatic alignment and more efficient use of the light by not requiring a separate beamsplitter. However, two factors that have not been considered here are effects of detector/photon noise and the fact that the design parameters of the SF-NES were not optimised—although it is not clear what the criterion for optimisation should be.

4.6. Conclusions

The numerical simulations of the NES shows the (expected) linear measurement characteristics, except when the input phase aberrations are not small. The aliasing and cross-coupling, defined specifically to measure the effects of these non-linearities, show that these errors are not large for the NES under the typical conditions for an interferometer aperture. The conclusion is that the NES is only suitable for open loop AO if D/r_0 is small (~ 1) but it is possible to use it for closed loop AO with worse seeing. This is confirmed by the closed loop simulations, which show that the NES is the best WFS, compared to a Curvature WFS and Quadrant Cell, in terms of sensitivity and wavefront correction when the seeing corresponds to $D/r_0 = 3$ or $D/r_0 = 5$—the former corresponding to the case of using the NES with COAST-sized apertures. For seeing of $D/r_0 = 7$, the CWFS is the best performing sensor regardless of light level so high-order AO is required for large apertures: greater than 1.2 m in I-band for 1 arc sec seeing, or 2.4 m for 0.5 arc sec seeing. The centring algorithm for the NES was found to be effective, but only when the odd phase aberrations consists mostly of tip/tilt as in open loop AO. The results of modelling a WFS with integrated spatial filter suggests that using a SF-QC is a better option than the SF-NES. For

the original proposal of using the NES with the apertures in COAST, the NES design offers better wavefront correction, when photon noise is included, than a more conventional WFS, except for extremely faint sources. A realistic alternative may be the SF-QC since introducing the pinhole does not affect the SF-QC wavefront measurements and QCs are already used at COAST (chapter 6). A more detailed investigation on the topic of pinholes integrated with wavefront sensors is required, however, before a definitive choice can be made of whether combination is superior to individual components.

5. Nine Element Sensor Prototype

5.1. Introduction

The numerical simulations of the NES confirmed that the design is suitable for low order AO with small apertures, for example those in the COAST telescope array. As with any new WFS method, a prototype which can measure aberrations using a telescope is its best verification. To that end this chapter concerns the design of and experiments with the NES prototype. The lack of DM or other correcting optics limited these experiments to passive measurements with the COAST interferometer.

The design stage was split in two; the design of the optical components and the corresponding design of the detector's mode of operation. The intended location in the COAST laboratory bunker was the beam-combiner table, and this is identical to one in the optics laboratory at the Cavendish Laboratory. This permitted a design to be tested in the laboratory and then directly installed within COAST and, given the limited space available, a compact WFS with no moving parts was envisaged. As the autoguider AO system and the tip/tilt aberrations at COAST are investigated in the next chapter and the data from the NES Prototype was used in this chapter to investigate higher order aberrations.

The measurements at COAST were divided into two parts: assessing the effects of internal turbulence on beam propagation within the laboratory bunker, and measurements of seeing on the sky. The results from both of these experiments were unexpected and improvements and further work using the NES prototype are discussed.

5.2. Optical Design

The optical requirement for the NES prototype is relatively unusual in that two defocused images from the same pupil are needed simultaneously. In principle, this involves splitting the input beam into two components and letting each have a different defocus diversity.

For practical reasons, further restrictions are applied and summarised as invariability and efficiency. The invariability of a design refers to the degrees of freedom for altering the optical paths and their variation over time. Generally, as the number of optical components increases, their alignment, relative and absolute, becomes more sensitive to external disturbance. Also the efficiency of the optics, which includes the throughput and the prevention of introducing phase aberrations, decreases as the number of optics increases. Together, these two criteria imply that the design should use as few parts and optics as possible. As a result, the NES prototype was constructed using one CCD camera, one lens and a novel beamsplitter.

5.2.1. Defocused Plane Imaging Methods

The Curvature WFS has prompted two solutions to the problem of imaging defocused images onto a fixed (detector) plane. The astronomical implementation uses a membrane mirror with variable curvature (66). As this mirror's focal length is modulated, the imaged plane changes periodically from the intra-focal to extra-focal plane and vice-versa. Disadvantages of this approach are that the integration time for each image is less than half of the mirror oscillation period and that the mirror control loop requires synchronisation with the detector integration, which adds complexity.

The second method also combines the defocusing and "beam splitting" by using a (static) diffractive optic (5). A common diffractive optic is the Fresnel zone plate (which acts as a lens) and the diffractive optic used for the CWFS is effectively a Fresnel zone plate with each zone centre offset from the centre of the optic's aperture. It can therefore be thought of as a combination of a diffraction grating and lens. As with a diffraction grating, there are varying orders of images—zero order is on-axis—and the +1 and -1 orders are the equally and oppositely defocused images required. Disadvantages of using this optic include its low throughput, because of the alternating opaque and transparent areas of the offset Fresnel zones, and its commercial unavailability.

5.2.2. Parallel Output Beamsplitter

The two solutions above are either too complex or too inefficient, so the NES prototype design will use a separate beamsplitter and defocus diversity optic. Using one detector implies that the beamsplitter needs parallel outputs, using optics to fold the output beams if necessary. Several NES designs were constructed, involving conventional beamsplitters and two lenses, to add the diversity separately to each beam. The alignment difficulties in these designs could not be overcome: the problem of invariability. Instead, the concept finally used was to add the diversity via one lens. This lens is placed before the beamsplitter and the different path lengths of the output beams then determine their relative diversity difference. As the NES has a small diversity, the fold optics can be incorporated into the beamsplitter: if the input beam has a diameter of 15 mm (as at COAST) and the lens has a focal length of 500 mm at 1.0 μm, to have a diversity difference of ~ 0.5 rad rms then the required path length difference between the beams is ~ 5 mm. This distance is small compared to the physical size of the beam and, inspired by the "ghost" reflection in a plate beamsplitter, a new beamsplitter design was conceived.

In a standard cube beamsplitter the output beams are the reflected beam and the transmitted beam. The former is at a right angle, and the latter parallel to, the input beam. The beamsplitter can be considered to be two right-angled glass prisms which are placed with their largest faces touching. If the prism from which the transmitted beam exits is replaced by a glass plate, then the angle of incidence of that beam with the rear surface of the glass plate exceeds the critical angle. The beam transmitted across the interface then undergoes total internal reflection to become parallel to the reflected beam. It then passes through the beamsplitter interface a second time and exits from the same surface as the reflected beam. This new design is called, for obvious reasons, the Parallel Output Beam-splitter (POB), and a comparison of the beams paths with those in a normal cube are shown in figure 5.1.

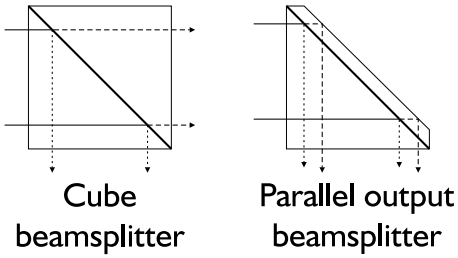

Figure 5.1.: A cube polarisation beamsplitter and a Parallel Output Beamsplitter.

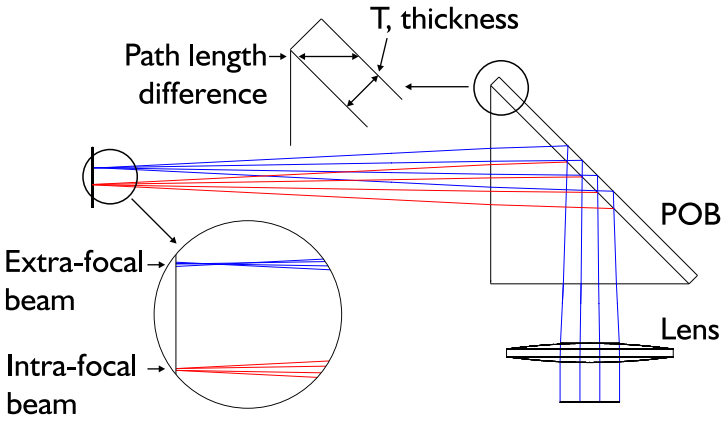

Figure 5.2.: Lens and POB arrangement to image two equally defocused images, on opposite sides of the focus, at the same plane.

To prevent the transmitted beam from being split again the POB interface must be of a polarisation type. The two output beams are therefore polarised but, as light sources in the lab and stars on the sky are (mostly) unpolarised, their intensities are equal.

The path length difference between the two output beams is defined by the POB, and is measured from the beamsplitter interface to the rear surface of the plate. To add the defocus diversity, a lens—called the diversity lens—is placed before the POB. The reflected beam will then focus further away from the POB exit face than the transmitted beam. The detector is placed midway between the focus positions and the images from the reflected and transmitted beams become equivalent to the intra-focal and extra-focal images. The reflected and transmitted beams can therefore be referred to as the intra-focal and extra-focal beams. The lens and POB and formation of the two images is all shown in a ray tracing diagram, figure 5.2.

The thickness of the glass plate, measured from the interface to the rear surface along a direction perpendicular to these planes, is a free variable, T. The path length difference can then be written as $\sqrt{2}T$ and, when the focal length of the lens is chosen, determines the defocus diversity. If the input aperture has radius r, the lens has a focal length of $f \gg T$, and a monochromatic wavelength λ is assumed, then the positive/negative defocus diversity for the intra-/extra-focal beam is,

$$\Phi_D(x,y) = \pm\sqrt{\frac{2}{3}}.\frac{kr^2T}{8f^2}\left(x^2+y^2\right) \tag{5.1}$$

Format: 500×290 pixels	Pixel size: $9.8\mu\text{m} \times 12.6\mu\text{m}$
CCD size: $4.9\,\text{mm} \times 3.6\,\text{mm}$	Readout noise: $\sim 15\text{e}^-$ rms
DAC: 16 bit (photoelectrons per pixel, $\lesssim 54000$)	Quantum Efficiency: ~ 0.5
Cooling: thermoelectric to $\sim -30°\text{C}$	Peak sensitivity: 550 nm

Table 5.1.: Specifications of the Starlight Xpress MX-5 CCD camera

5.3. CCD Detector Design

The CCD detector chosen for the NES prototype is a Starlight Xpress MX-5 CCD camera. The camera consists of a cooled CCD connected to a computer via a USB I/O interface and the overall specification is shown in table 5.1. The availability of open source software drivers (69) allowed customisation of the CCD readout method for the specific requirement of reading out a few detector pixels at high speed.

The minimum detector requirements for the NES prototype are a readout of 18 square pixels at a rate of $\sim 100\,\text{Hz}$. By binning 4 pixels horizontally ($39.2\,\mu\text{m}$ width) and 3 vertically ($37.8\,\mu\text{m}$ height) it was possible to achieve a readout rate of 333 Hz with the MX-5.

The time for reading out the digitised charge from a CCD array is a function of the size of the exposed readout region within the full frame—the sub-frame. The conventional readout of the MX-5 occurs as a set of subsequent processes, known as a readout cycle, and, from the perspective of the software in the computer, the readout cycle is a number of requests sent to, and acknowledged by, the I/O interface. The I/O interface uses an internal microprocessor with customisable software and communicates with the camera head, which contains the CCD and associated electronics. The initial analysis indicated that the communication between the I/O interface and the computer limited the maximum readout rate (of one pixel) to be $\sim 30\,\text{Hz}$. A different readout cycle was programmed into the I/O interface microprocessor, designed originally to readout a 64 pixel row from a MX-5 camera at 500 Hz (54).

In this readout cycle, after the computer sends the first exposure request, the I/O interface enters a loop of sequential exposures until explicitly instructed otherwise. An additional increase in readout rate was obtained by preventing the normal charge clearing process. Using these two enhancements, the resulting readout cycle is called Shuffled Charge Imaging (SCI). It can only be used because the MX-5 CCD employs an interline design; each column of light-sensitive pixels is physically interleaved with equivalent readout pixels which are covered by an opaque mask. Above the light-sensitive pixels is placed a contiguous lenslet array which restores the "fill factor" of the CCD array and recovers the "lost" light. The readout pixels act simply as a storage for the charge transfered from the light-sensitive array and, during one readout cycle, the photoelectrons are clocked (transfered periodically) from this array to the readout register. There, the number of electrons is digitised and the resulting values sent to the I/O interface and then finally to the computer.

A conventional exposure begins with transferring the photoelectrons at the end of the exposure to the readout array from the light-sensitive array. Only the charge within the subframe on the readout array is clocked from the readout register to a ADU (Analogue to Digital Unit), for digitisation,

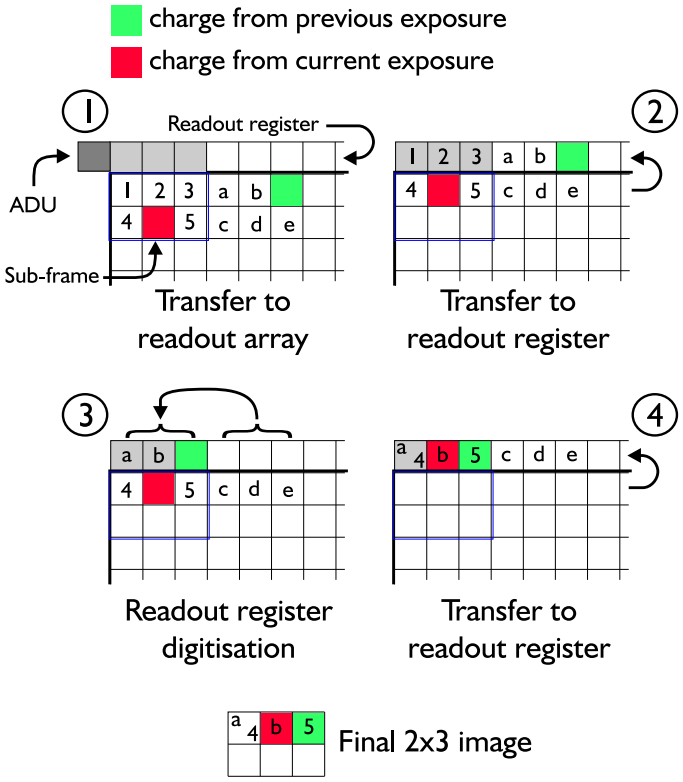

Figure 5.3.: The scheme of Shuffled Charge Imaging. The numbered stages are referred to in the text.

and the remainder is discarded from the register. The clocking steps clears the readout array so that the next transfer to it from the light-sensitive array will only contain photoelectrons from a single exposure. The readout process in the CCD can thus be divided into several steps;

- Transfer the charge from the light-sensitive array to the readout array.
- Transfer each row in the readout array to the readout register.
- Transfer each pixel in the readout register towards the ADU.
- If the pixel next to the ADU corresponds to one in the sub-frame then digitise it, else discard the charge.
- Repeat the charge transfer to the readout register until all the rows of the readout array have been clocked out.

Shuffled Charge Imaging (SCI) eliminates all clocking steps bar those for sub-frame digitisation, and allows the readout rate to be increased to 500 Hz. As a readout is also continuous, the exposure time becomes is a function of the time to digitise the sub-frame and is not independently definable. The lack of charge clearing means that readouts essentially occur in parallel to exposures, with the current readout equating to charge from previous exposures. An essential feature of SCI is that, regardless of the sub-frame size, the charge in the entire readout array is eventually digitised. As a consequence, the sub-frame size is chosen to provide the required readout rate and need not

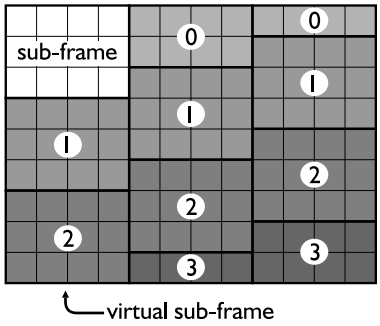

Figure 5.4.: The grid of virtual sub-frames for Shuffled Charge Imaging. Shaded sub-frames are virtual and the numeral represents the delay in exposures for their readout.

to encompass the locations of the two defocused images on the light-sensitive array. A graphical schematic of the readout method and charge movement is shown in figure 5.3, and the circled numbers will be referred to in the text.

By not clearing the accumulated charge *after* an exposure, the transfer at the end of an exposure adds the charge in the light-sensitive array to that already in the readout array. The readout array can now contain photoelectrons from two or more separate exposures (stage 1). The sub-frame (marked in blue) is defined as starting at the readout register and each row from it is clocked into the readout register. This also transfers charge from outside the sub-frame. Each time this vertical clocking occurs, each row in the readout array moves one row closer to the readout register (stage 2). The number of pixels transfered to the ADU for digitisation is now the width of the sub-frame, but pixels after this do not have their charge discarded (stage 3). When the next row's charge is transfered into the readout register, it adds to the the (offset) charge that was originally outside the sub-frame (stage 4). This means that each digitised pixel can contain values from charge originating in more than one pixel in the light-sensitive array. Consecutive readouts will therefore eventually readout the entire CCD.

The sub-frame is therefore augmented with virtual sub-frames which are of the same size and arranged in a warped grid across the entire readout array, an example layout is shown in figure 5.4. The intensity in a pixel in the final image comes from the charge in corresponding pixels of every virtual sub-frame. The virtual sub-frames next to the readout register have their charge readout with that from the real sub-frame during each readout cycle. The charge in other virtual sub-frames are then readout during subsequent exposures as their charge is shuffled up by one sub-frame per readout.

In general, this would lead to confusion as the pixels in the digitised image no longer have a one-to-one correspondence to pixels in the readout array. However, from the NES prototype perspective, the full frame has two compact images so each exposure only transfers significant charge into two virtual sub-frames. As long as each defocused image, relative to the virtual/real sub-frame it lies in, is offset from the other, the final image from each exposures consists of two well-separated images.

The frame rates for various sub-frame sizes and binnings of CCD pixels were measured (appendix D) and it was found that a sub-frame size of 60×24 unbinned pixels (binned to 15×8 pixels) gave

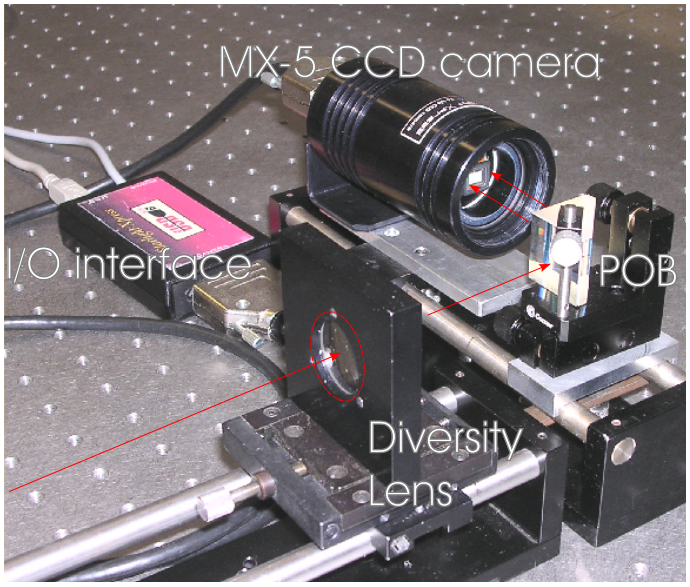

Figure 5.5.: Photograph of the NES Prototype with major parts labelled. The red line(s) represents the light beam(s), which enters the WFS at the lower left of the picture.

a stable rate of 333 Hz (3 ms integration per frame). The stability is important because the time between readouts determines the exposure time of each frame and this is required to be constant for temporal analysis.

5.4. NES Prototype Laboratory Experiments

The NES prototype design consists of two optical benches, one containing the MX-5 camera and the POB and the other holding the diversity lens, as shown in figure 5.5. The lens is removable, to allow flexibility in the focal length, whilst the POB and MX-5 are initially aligned with a laser and then not touched. Once the WFS is aligned, the two benches were baffled with a box to reduce stray light and the camera I/O interface connected to a computer; in this state the WFS is ready to be operated. Two diversity lenses were used, one to configure the WFS to operate as a CWFS and the other for operation as a NES.

5.4.1. POB Characterisation Experiment

A POB was fabricated from a near-IR cube beamsplitter, available from the standard stock list of a local company[1]. The cube had the glass sliced away from one corner and the resulting optic is shown in figure 5.6.

Specification

The specification of the cube beamsplitter is listed in table 5.2. The choice of the near-IR bandpass beamsplitter was dictated by the COAST bunker windows which cut off wavelengths below ~

[1]Comar Instruments, 70 Hartington Grove, Cambridge CB1 7UH, UK

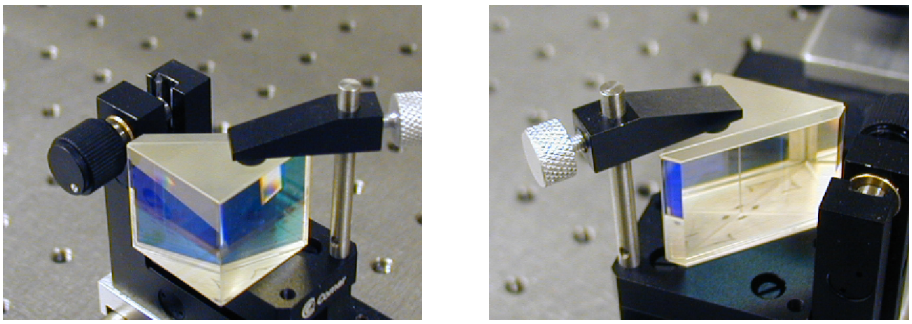

Figure 5.6.: Photograph of the POB fabricated from a cube beamsplitter.

Bandwidth	700-1064 nm
Transmitted beam polarisations (S/P)	0.2%/97%
Reflected beam polarisations (S/P)	3%/99%
Material	SF-15 glass
Dimensions	25 mm
Anti-reflection coatings	All external surfaces except rear

Table 5.2.: Manufacturers specifications of the cube beamsplitter from which the POB was cut.

600 nm. The thickness past the interface was specified as $T = 3$ mm (a path difference of 4.24 mm) for two reasons: the manufacturing process was limited to thicknesses greater than 2 mm and the upper limit of 3.46 mm is because the thickness controls the spacing between the output beam. In fact the spacing is identical to the path difference so the width of the light-sensitive array in the MX-5 CCD requires the thickness be less than 3.46 mm.

Polarisation Characterisation

One important characteristic of the POB is the "quality" of the polarisation for the output beams. For a polarisation beamsplitter, the variables T_i and R_i represent the fractions of the polarisation states for the output beams. The transmitted beam is ideally of one polarisation, $T_1 = 1$ and $T_2 = 0$, and similarly for the reflected beam $R_1 = 0$ and $R_2 = 1$. In the POB design, if $T_2 \neq 0$ then there will be more than two output beams when the extra-focal beam is split again at the interface.

Measurements of the polarisation quality for specific wavelengths in the range 600-800 nm were determined by using a prism to disperse the light output from the POB. The intensities were calibrated by measuring the dispersion without the POB in place to measure the spectral transmittance of the other optics. Calibrated intensities of the first three output beams from the POB are shown in figure 5.7. Although there is a third output beam because the interface is not perfect, its intensity is very small over the wavelength range measured. For broadband light, the combined intensity of the intra-focal and extra-focal beams is $\sim 97\%$ of the input.

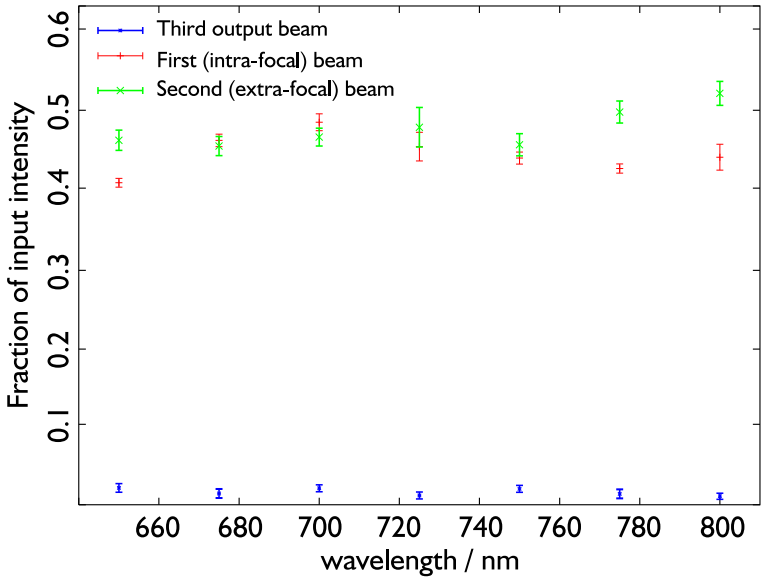

Figure 5.7.: Normalised intensities of the first three output beams from the POB for several wavelengths.

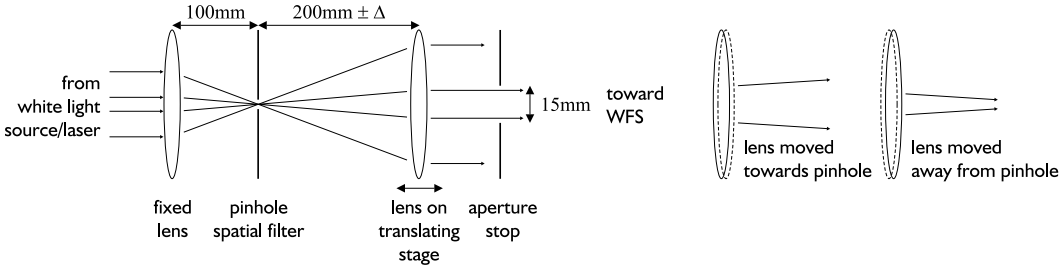

Figure 5.8.: The beam expander used to introduce defocus aberrations. The variable Δ is the offset of the lens from the zero defocus position.

5.4.2. Defocus Aberration Control Experiment

Introducing controllable phase aberrations to test the response of the WFS was achieved by translating bulk optics. Hardy (28) noted that the the controlled misalignments of two lenses can be used to correct low order Zernikes, Z_2 to Z_8 inclusive. Inverting this concept to introduce tip/tilt, defocus, or astigmatism, only misalignments (translations and rotations) in one lens are needed.

Introducing defocus aberrations requires only moving the second (defocus) lens in a beam expander along its optical axis. To control tip/tilt and astigmatism, the lens also needs to be rotated. The goal of these experiments was to confirm the NES prototype could make measurements of defocus, so no rotation of the defocus lens was allowed. This lens was attached to a stage with a micrometer screw gauge, allowing 10 μm accuracy in translation, Δ (defined to be zero when no defocus aberrations is introduced). To reduce other unwanted aberrations, the beam expander also incorporated a pinhole spatial filter and an outline of the optics is shown in figure 5.8. The output beam has a diameter of 15 mm from use of an aperture stop.

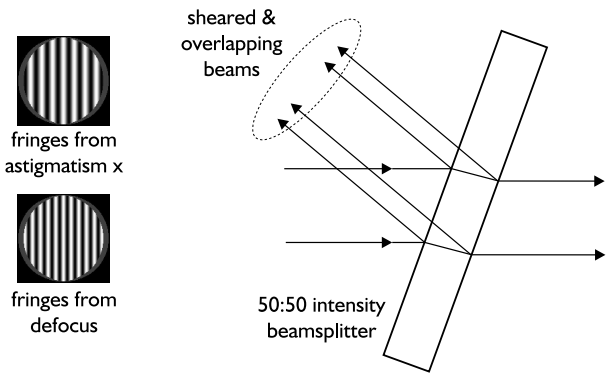

Figure 5.9.: Using a plate beamsplitter to form a shearing interferometer, looking down onto horizontally propagating beams.

To ensure that the defocus aberration could be correctly inferred from the defocus lens offset, read from the stage's micrometer, a shearing interferometer was intially used. It was implemented by simply placing a plate 50:50 intensity beamsplitter after the defocus lens at a large angle of incidence to the beam, $\sim 40°$ (50). The reflections from its front and rear surfaces result in two reflected beams which are offset (sheared) and overlap. This causes interference between the sheared beams and, using a laser for high coherence, vertical fringes could be seen across the overlap region, see figure 5.9. The period of these fringes is a function of the *horizontal* amplitude component of defocus i.e., if the second order terms of a phase aberration are written as $\Phi = a\,x^2 + b\,y^2 + c\,xy$, then the fringe period is proportional to $|a^{-1}|$ only. The defocus aberration introduced by a translation of Δ with a lens of focal length $f \gg \Delta$ is

$$\Phi(x, y) = \frac{k\Delta}{2f^2}\left(x^2 + y^2\right).$$

The fringe number (the inverse of the period) is then proportional to Δ. Initially, the defocus lens was moved away from zero defocus until there were a large number of fringes which could not be individually discerned. As the lens was moved back towards the zero defocus position, the fringe pattern remained vertical but the frequency decreased. The number of fringes across the interferogram was recorded at several distance intervals until the number became too small to measure. The lens was moved past zero defocus and, as the fringe frequency again increased, the number of fringes seen was again recorded.

The expected behaviour of the beam expander's defocus aberrations with defocus lens offset is seen in a plot of the number of fringes (negative after the zero defocus position to allow a linear fit) versus defocus lens offset, figure 5.10. However, at the position indicated to introduce zero defocus, fringes could still be seen (with a period of $\sim 1/4$ across the overlap region). As these fringes were vertical, it is assumed that they were from astigmatism aberrations (25). Using the defocus error from the graph which gives the same period of fringes, the amplitude of the assumed astigmatism can be estimated to be ~ 1 rad rms at the laser wavelength. An estimate of the rotation angle in the defocus lens, that would result in this astigmatism, is $\sim 1°$. The optics were realigned several times to remove this effect but it could not be reduced below the level presented here.

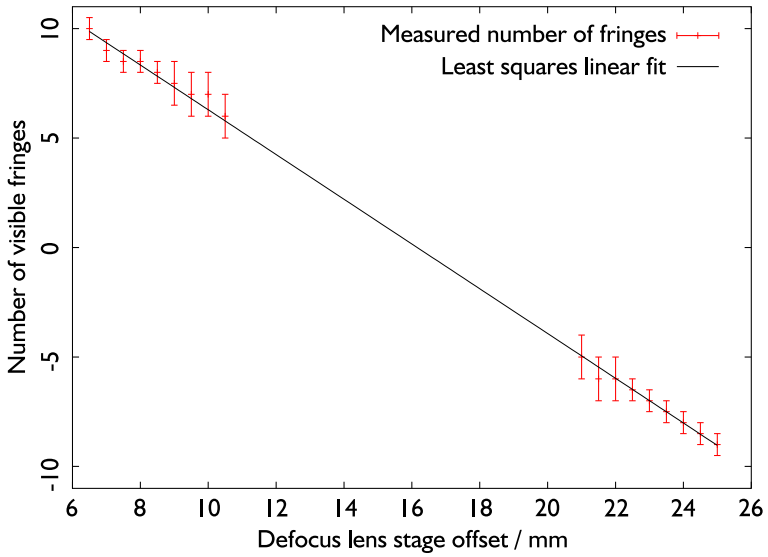

Figure 5.10.: Number of interference fringes in the overlap region from the shearing interferometer plotted against the defocus lens offset.

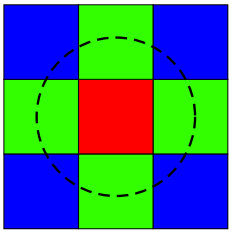

Figure 5.11.: Sub-apertures used in the CWFS experiment with the POB. The dotted line represents the defocused image outline and the colour coding corresponds to sub-apertures for measuring specific aberrations

5.4.3. CWFS Experiment

Using a short focal length diversity lens with the POB allows a Curvature WFS to be created, and this was implemented in the laboratory. The defocus distance from the focus is half the path difference, $d = T/\sqrt{2} \ll f$, and the requirement for CWFS is,

$$d \gg \frac{\lambda f^2}{D^2}, \tag{5.2}$$

where f is the focal length of the lens. The inequality can be satisfied using a 150 mm focal length achromat and a 700 ± 10 nm bandwidth filter with a white light source. The purpose of this experiment was to ensure that defocus aberrations could be measured using the two defocused images from a POB. Additionally, confirmation was sought that the additional aberrations inferred from the shearing interferometry experiment were actually present.

The experiment involved taking high-resolution frames with the CCD camera for several defocus lens offsets. A frame consists of the two defocused images, each of which was extracted via software into two equally sized sub-frames. The centring algorithm was applied to these sub-frames, and

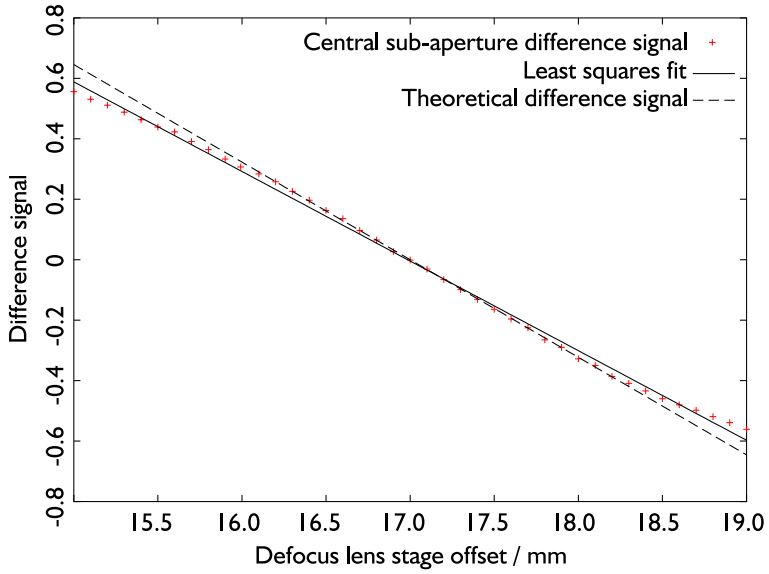

Figure 5.12.: Defocus aberration difference signals measured with the POB and a 150 mm focal length achromat, in the Curvature WFS configuration.

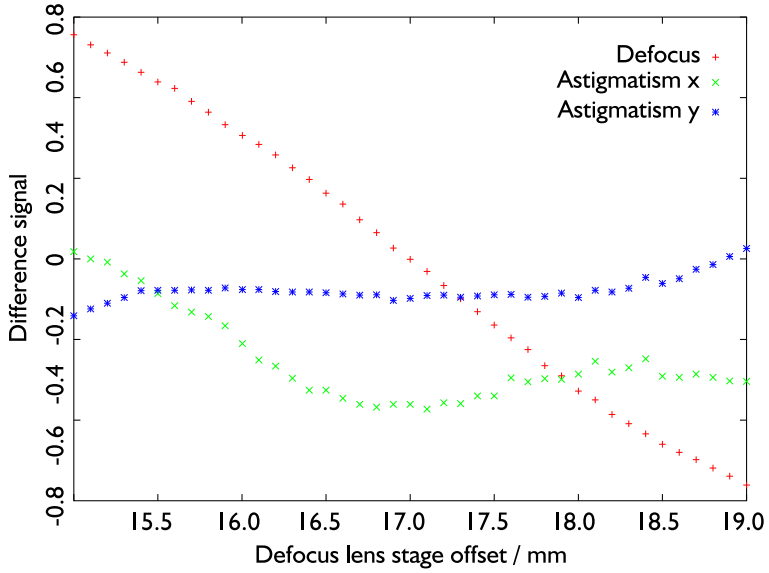

Figure 5.13.: Difference signals corresponding to defocus, astigmatism x, and astigmatism y aberrations, c.f. figure 5.12.

finally a 3×3 grid of binned pixels was formed, see figure 5.11. The centring algorithm is the Fourier transform method previously described; apply a FT to the intensity, set the phase at all points to be zero, and then apply an inverse FT. The grid defined 9 sub-apertures per plane and the central sub-aperture was always smaller than the image size: the CWFS measurement is then the subtraction between equivalent sub-apertures in the intra-focal and extra-focal planes.

The curvature of the defocus aberration is directly measured by the difference signal of the central sub-aperture, marked red in figure 5.11. The two astigmatism aberrations can be qualitatively

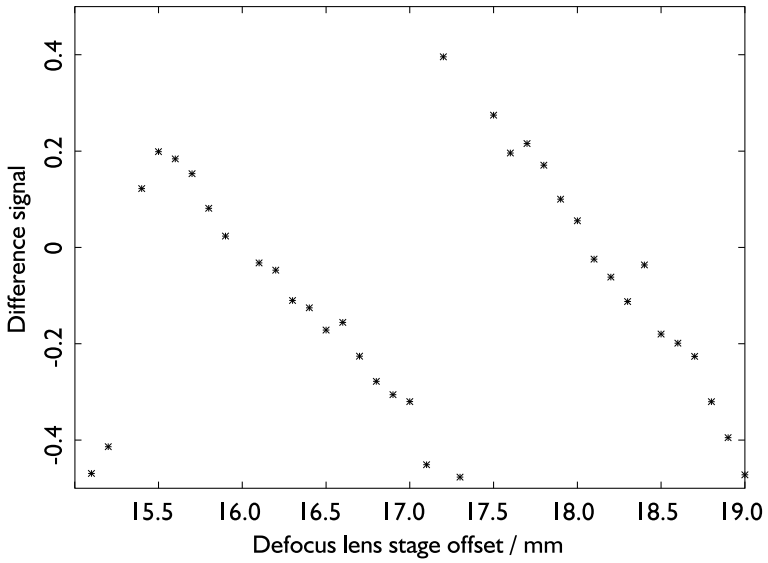

Figure 5.14.: Horizontal position of one image—the discontinuity occurs whenever the image moves across an entire unbinned CCD pixel.

measured by difference signals from independent sub-apertures, marked green and blue. A plot of the defocus difference signal with the best linear fit and expected theoretical fit is shown in figure 5.12, and the difference signals for the three quadratic aberrations together are plotted in figure 5.13.

The measured defocus amplitudes show a general agreement with the predicted line, and the mismatch is attributed to uncertainties in the focal length of the diversity lens used. This result shows that the principle of using the POB for defocus aberration measurements is valid. The difference signals for the astigmatisms indicate that there does appear to be additional aberrations. A further measurement was the horizontal position of the images, obtained from a "centre of mass" measurement of the sub-framed images before their centring. The positions are plotted for one image in figure 5.14, and the discontinuity corresponds to a physical image movement of one CCD pixel.

The movement of the images is linearly related to the defocus lens offset and is consistent with the defocus lens's optical axis being fixed but not parallel to the beam direction. As the defocus lens moves relative to the CCD, the introduced tilt aberration changes as the distance between the lens's optical axis and the beam's optical axis changes.

5.4.4. NES Experiment

The CWFS configuration proved that the diversity lens and POB combination allows the defocus aberration to be measured from the two defocused images. That experiment also showed that the image movement across the CCD with the beam expander optics introduced significant tilt and astigmatism. The former can be corrected with the centring algorithm and the latter did not significantly affect the defocus measurements. The results using of the POB with a longer focal length lens to form the NES prototype will be discussed in this section.

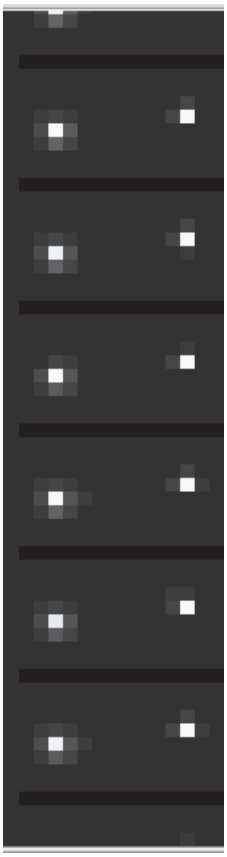

Figure 5.15.: Example strip of consecutive frames from the shuffled charge imaging mode.

The diversity lens used for the NES has a focal length of 300 mm, which gives a diversity amplitude of ± 0.82 rad rms at 800 nm. Analysis of the POB and this lens was carried out with the ZEMAX ray tracing software: the introduced aberrations in the two images consisted mainly of defocus (as required) and the next significant term was spherical aberration ($\sim 3\%$ of the diversity amplitude). Finally, it is noted that the binning of the CCD pixels was not chosen to optimise the resulting pixel width (~ 1.8 Airy radii) for sensitivity of the NES prototype.

A difference from the previous chapter is that broadband operation is considered here, using the same white light source as the CWFS experiment. The general concept for wavefront measurement involves a numerical simulation of the optics to calculate the pseudo-inverse of the measurement matrix, and for this a single wavelength is needed to be specified. The data was therefore first analysed to establish an effective wavelength for broadband light.

The experimental procedure was similar to the CWFS experiment above; the defocus lens was moved to several positions, and for each location the two defocused images on the CCD were recorded. Here the SCI mode of the CCD was used, and for each lens offset a 15 s strip of frames was taken. Each frame in a strip file is a sub-frame on the CCD, and a strip file consists of 5000 frames, so the readout rate is 333 Hz (an example section of a strip is shown in figure 5.15).

From each frame, a 7×7 pixel sub-frame is extracted for the two images and a 3×3 grid is used to calculate the intra-plane differences. The intensities in the 7×7 sub-frames then have the centring

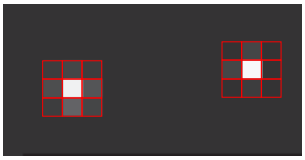

Figure 5.16.: One frame from the shuffled charge imaging mode. The grids represent the 3×3 grid used for calculating difference signals.

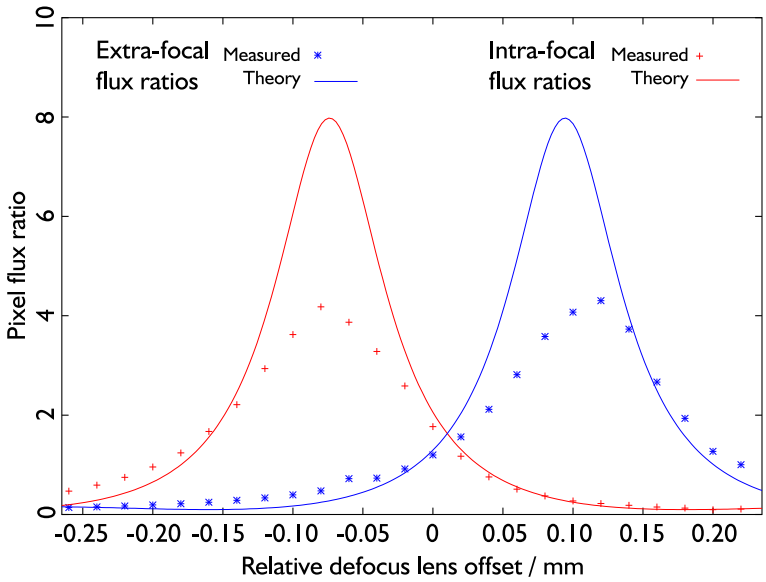

Figure 5.17.: The ratio between the central pixel flux and the sum of the outer 8 pixel fluxes. The theoretical measurement assumes a wavelength of 800 nm.

algorithm applied and the inter-plane difference are calculated from the 3×3 grid, and only results from the inter-plane difference are presented here. One NES detector pixel therefore corresponds to a sub-frame pixel, as shown in figure 5.16, and notice that the extra-focal (righthand) image in each frame is one pixel higher than the intra-focal (lefthand) image. This occurs because the images are located in adjacent virtual sub-frames on the readout array.

A simplistic measure of the defocus of an image is its size, and can be quantified for the NES measurements by the ratio of the flux in the central pixel to the flux in the outer 8 pixels. A numerical simulation of the prototype NES was run for various defocus aberration amplitudes and this ratio calculated. Similarly, the ratio was calculated for the experimental data and a plot showing a comparison between the data is shown in figure 5.17. To scale the defocus aberration in the simulation from rad rms to an equivalent defocus lens offset, several wavelengths were empirically used and the value finally chosen was 800 nm. The data is only an approximate match to the simulated values but note that both the peaks and the crossover point, at zero defocus lens offset, correspond well. The peaks are the point where the defocus diversity is cancelled by the defocus aberration i.e., each image is focused. Compared to the theoretical curves, the ratios for the data are smaller at the peak, and the explanation is simply that the (white light) focused images are not diffraction-limited and therefore larger than the simulated images.

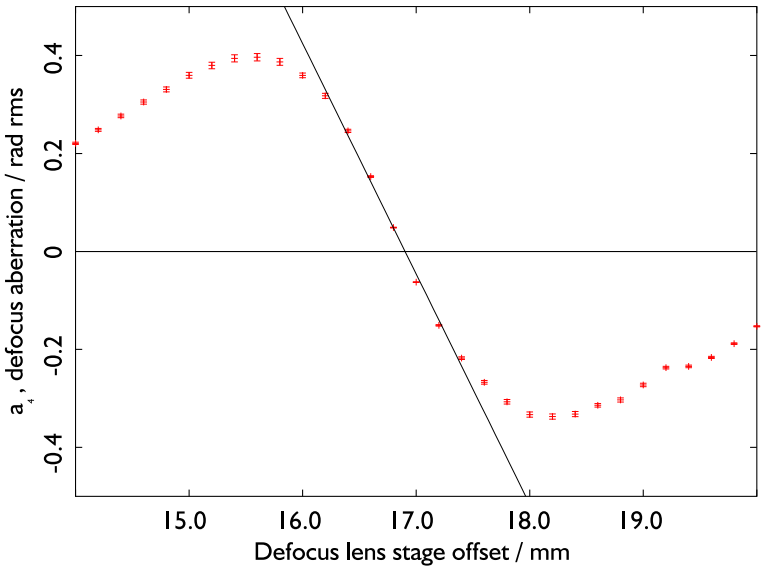

Figure 5.18.: The defocus aberration amplitude as measured by the NES prototype with SCI.

Although only qualitative, the correspondence provides the basic confirmation that the reduction of data from the NES prototype can use a measurement matrix from the simulations. Accordingly, the pseudo-inverse M^+ was calculated for an effective wavelength of 800 nm and used to reconstruct the wavefront. A plot of the amplitude of Z_4 (defocus), from the data used in figure 5.17, is plotted in figure 5.18. The amplitudes and error bars are the mean and standard deviation from the measurements in each frame.

As expected, there is a symmetric linear region around zero defocus and for larger defocus aberrations the measured amplitude is non-linear and falls away towards zero. The line represents a linear fit to the points about the zero defocus position. The predicted position of zero defocus, ~ 17 mm, is consistent with the CWFS experiment, figure 5.12, and the results from the shearing interferometry experiment. However, the measured defocus amplitudes do not correspond to the actual amplitudes expected from the lens movement. This is a result of assuming an effective wavelength and neglecting broadband effects. To measure the actual defocus aberration amplitude, an additional (empirical) multiplicative constant—the amplitude calibration factor—is required, and an estimate of its value is 1.95 ± 0.01.

A plot showing the amplitudes of Z_4, Z_5, and Z_6 for the defocus lens movement is shown in figure 5.19. Although the measured amplitude of Z_5 is small and consistent with zero, the NES measures a significant amplitude of Z_6. In particular, it shows a behaviour that is offset from the Z_4 amplitudes. The measured amplitudes of Z_6 are not thought to correspond to the actual aberration but instead are an artifact of the centring algorithm.

The effects of the image movement from the defocus lens misalignment can be seen if the measured amplitudes of Z_4 with and without centring are plotted, figure 5.20. As every frame in a strip file has unchanging aberrations from the beam expander, a significant variance in defocus results from image intensity being divided between the centre pixel and an adjacent pixel. The amplitude then become extremely sensitive to the noise, as can be seen for the lens offset of 18 mm. At the lens

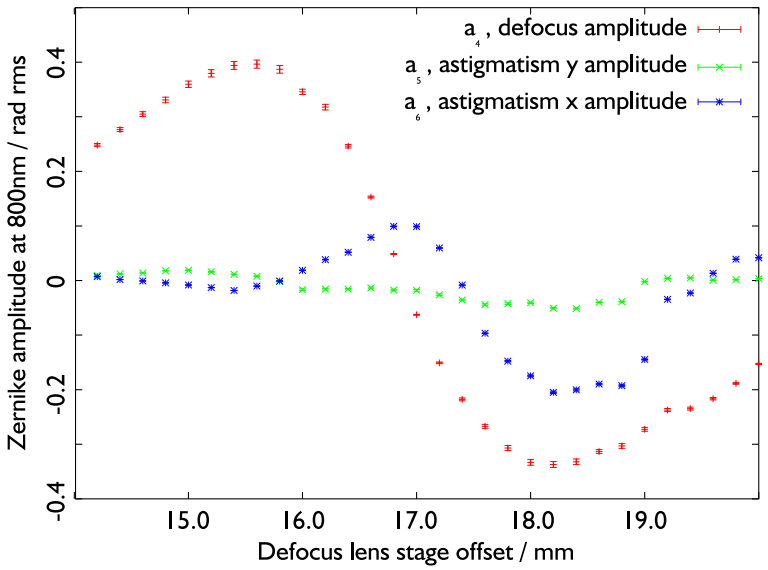

Figure 5.19.: The amplitudes of Z_4, Z_5, and Z_6 as measured by the NES prototype.

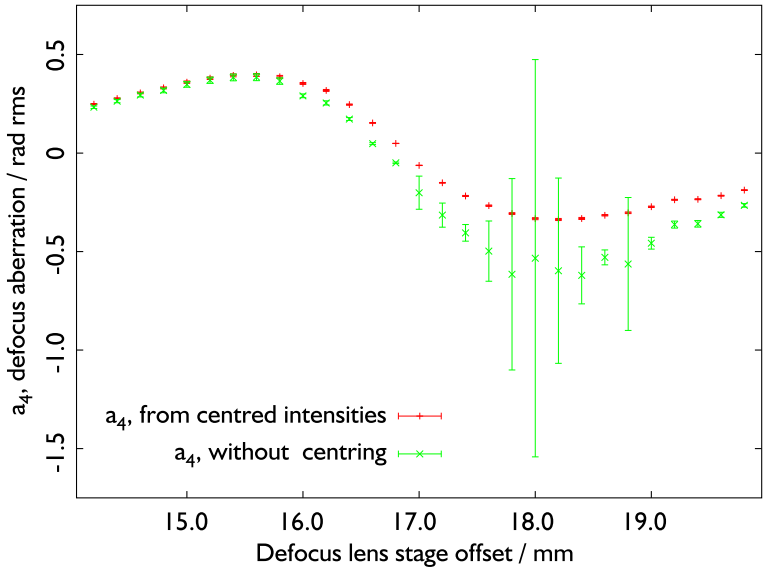

Figure 5.20.: The amplitudes of Z_4 with and without using the centring algorithm on the intensities.

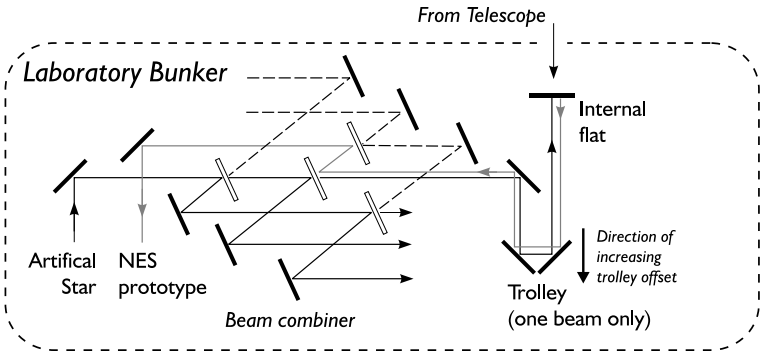

Figure 5.21.: The NES prototype experiment location in the COAST laboratory. The diagram shows the internal flat inserted for the artificial star light source, and removing the mirror allows starlight to be used instead.

location of 18 mm, the variance is highest and there is a marked bias in the Z_4 aberration amplitude. Also, this lens offset is where the measured amplitude of Z_6 reaches a minimum. As the centring is applied to intensities which are sampled over finite sized pixels, it introduces increasingly erroneous intensities as the image movement approaches half a pixel width. These errors affect the corner pixels in the 3×3 grid most and it is these which measure Z_6, hence the difficulty of measuring this aberration.

5.4.5. Conclusion

The laboratory experiments with the NES prototype have shown that the basic design concept, using a single lens and the POB, can be used to measure the amplitude of defocus aberrations. Two specific results are that the POB can be used with a short focal length lens to form a Curvature WFS, and that the NES prototype can be used to make linear measurements of small defocus aberrations at a rate of 333 Hz. The evidence of additional aberrations introduced by the beam expander optics was only qualitative. It did, however, provide an opportunity to prove the usefulness of the centring algorithm with the NES data, albeit for a slightly different purpose than originally envisaged.

5.5. NES Prototype Measurements at COAST

The remainder of this chapter will concern measurements taken at the COAST interferometer with the NES prototype. The experiment was placed within the COAST laboratory bunker on the beam-combiner table, as shown in figure 5.21. As mentioned previously, the table is of the same type as in the Cavendish Laboratory so a direct transplantation of the prototype was possible. The data was recorded for three distinct situations: to measure the seeing between the internal light source and WFS; the seeing from beam propagation within the bunker; and seeing from the atmosphere. Correspondingly, the results are subdivided into internal and external seeing measurements. As the amplitude calibration factor is uncertain, the values presented here are in unscaled units from the software and not calibrated values. In common with the laboratory experiment, several strip files

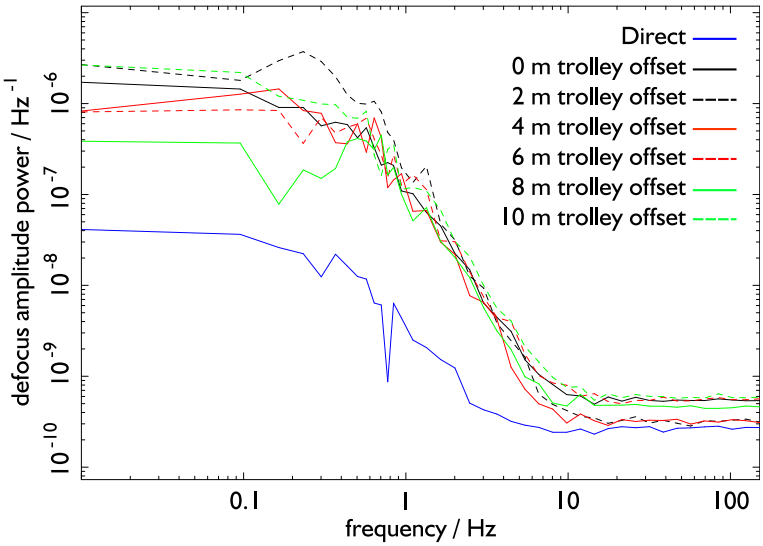

Figure 5.22.: Power spectra for NES-measured defocus aberration amplitudes, for the internal seeing measurements.

were taken for each measurement and the images were extracted and had the centring algorithm applied before using the pseudo-inverse measurement matrix.

5.5.1. Internal seeing

To measure the seeing within the bunker, the internal light source, referred to as the artificial star, was used as a reference beam. It inputs a light beam into the beam-combiner, which is split into four beams at its output although only one output was used here (and the others were physically blocked.) The beam is then reflected from a spot mirror beamsplitter and travels, via one path compensation trolley, toward an internal flat mirror. The reflected light retraces its path back to the spot beamsplitter, which also acts as an aperture stop, and re-enters the beam-combiner. Upon exiting the beam-combiner, there are again four light beams and one output is directed toward the NES prototype. The light then travels the same path twice except for a small distance from the light source to the beam-combiner and from the beam-combiner to the NES.

To measure the seeing intrinsic in the artificial star output, a mirror was used to redirect the light, that would otherwise enter the beam-combiner, toward the NES. An aperture stop, to match that from the spot mirror beamsplitters, was placed before the NES defocus lens to ensure the correct diversity for the images. This optical arrangement is described as 'direct'. The remainder of the experiments used the light path described above for several positions of the path compensation trolley. Each light path is then described in the text by reference to the trolley offset in metres. Particular care was taken to minimise the impact of turbulence since the bunker is not an isolated environment e.g., only necessary electrical equipment was powered on and the experimenter operated the WFS from a distance.

The results from only the measured defocus aberrations (Z_4) are presented here, to simplify the presentation and because it was this aberration that was fully tested in the laboratory experi-

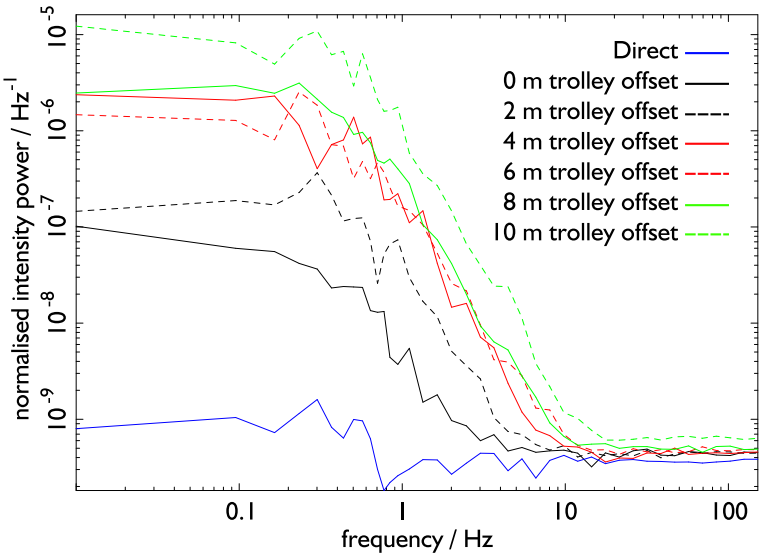

Figure 5.23.: Power spectra of the scintillation (normalised intensity) for the internal seeing measurements.

ments. Two other measurements can be made from the recorded images: the scintillation and the polarisation. The scintillation index is the variance of the normalised total intensity,

$$S = \frac{\langle (I_+ + I_-)^2 \rangle}{\langle (I_+ + I_-) \rangle^2} - 1.$$

where I_+ and I_- is the total flux within each 7×7 subframe around the defocused images. The polarisation is calculated from the flux ratio between the two images, I_+ / I_-. The use of this variable is to test the assumption made that the two images have equal intensity, or equivalently that the light entering the NES is unpolarised.

The power spectra of the defocus measurements are shown in figure 5.22 for the direct path and 6 trolley offsets. Each spectrum is an average of 5 separate spectra, and then is binned logarithmically (each bin width is equal in log frequency). The variances corresponding to each spectrum are shown in table 5.3.

The measured power for the direct path is smaller than for all paths involving the trolley and the aberrations in the light output from the artificial star can therefore be ignored when considering the paths involving the trolley. When examining the spectra from trolley paths, for all the positions the high frequency behaviour is similar to the noise limit (at 10 Hz). The variances themselves vary considerably and show no particular trend with trolley offset.

The power spectrum of scintillation can be computed from the normalised image intensity,

$$\frac{(I_+ + I_-)}{\langle (I_+ + I_-) \rangle},$$

and the averaged spectra are shown in figure 5.23, while the scintillation index is listed in table 5.4. The scintillation for the trolley paths show a markedly different character from the defocus amplitudes, because as the trolley offsets increase, the scintillation index also increases in general.

Light path	Defocus variance
Direct	$0.44 \pm 0.06 \times 10^{-4}$
0 m	$5.99 \pm 3.18 \times 10^{-4}$
2 m	$10.81 \pm 3.65 \times 10^{-4}$
4 m	$5.30 \pm 2.52 \times 10^{-4}$
6 m	$3.97 \pm 0.94 \times 10^{-4}$
8 m	$2.67 \pm 0.85 \times 10^{-4}$
10 m	$7.66 \pm 3.67 \times 10^{-4}$

Table 5.3.: Measured NES amplitude variance for internal defocus aberrations. The light path represents either the direct path or the offset of the path compensation trolley.

Light path	Polarisation	Scintillation index
Direct	1.026 ± 0.001	$0.42 \pm 0.02 \times 10^{-4}$
0 m	1.097 ± 0.007	$0.77 \pm 0.07 \times 10^{-4}$
2 m	1.094 ± 0.020	$1.41 \pm 0.16 \times 10^{-4}$
4 m	1.109 ± 0.005	$7.30 \pm 3.21 \times 10^{-4}$
6 m	1.167 ± 0.003	$6.26 \pm 2.16 \times 10^{-4}$
8 m	1.144 ± 0.006	$12.33 \pm 1.94 \times 10^{-4}$
10 m	1.185 ± 0.011	$43.19 \pm 19.04 \times 10^{-4}$

Table 5.4.: Polarisation—the ratio of the two defocused image intensities—and scintillation indices for the internal seeing measurements. The light path represents either the direct path or the offset of the path compensation trolley.

This increase is also evident when comparing the relevant power spectra. The cause is simply the larger diffraction losses, around the beamsplitter aperture stop, for longer light paths.

The polarisation is also shown in table 5.4, and for the direct path the image intensities show a small polarisation effect, $\sim 3\%$, which is similar to the value measured in the laboratory. When the light passes through the beam-combiner, the polarisation increases and there is evidence that it becomes larger with longer path differences. Although these results invalidate the assumption of unpolarised light required for calculating scintillation-free inter-plane difference signals, it is noted that there is no evidence of correlation between the polarisation measurements and either scintillation index or the defocus amplitudes.

The recorded images can be summarised as being affected by polarisation of the artificial star light but this does not appear to affect the measurement of defocus aberrations. There is an increase of the defocus variance for paths including the trolley compared to the direct path, but there is little evidence that the increase of trolley offset is correlated with the total power in this aberration. Upon comparison of the power spectra for defocus, it is more apparent that there is a common source of seeing for the trolley paths which implies a source located between the artificial

star and path compensation system. It is postulated that this source may be the electronics for the laser interferometer, which is part of the trolley metrology system and are always powered on. This is located below the light path at the entrance to the path compensation system, and its heat output could cause the aberrations that are seen. The effect of increased trolley paths are instead seen in the scintillation power spectra. There is also evidence that the polarisation increases with trolley offset, but it is dominated by the (fixed) polarisation from the beam-combiner.

5.5.2. Atmospheric seeing

Using the same experimental technique as for the internal measurements, the seeing from the atmosphere was measured by removing the internal flat mirror and positioning a trolley at 0 m. Observations presented here were made using two COAST telescopes (Centre and Inner West) on one night (10^{th} December 2003) using α Tau as the reference star. This star has a magnitude of $m_I = -1.8$ in I-band and a zenith angle at transit of $\sim 40°$. It is therefore bright and has a low declination which allows the neglection of polarisation effects from the telescope siderostat pointing angle. To eliminate the effects of seeing introduced at the telescope itself (dome seeing), the telescopes were allowed to reach the ambient air temperature by opening their covers 1 hour before the observations (4 hours after sunset). Comparison of seeing measurements using artificial star light reflected from a siderostat with the internal seeing measurements showed no evidence of dome seeing.

The data was again reduced to find the defocus amplitudes, scintillation index, and polarisation. Beginning with the polarisation measurements, the results from the two telescopes were significantly different: the image intensity ratio is 1.14 ± 0.01 and 1.39 ± 0.05, for the Centre and Inner West telescopes. The trolley used for the internal seeing measurements also corresponds to that used with the light beam from the Centre telescope. Since the internal polarisation measurement has a similar intensity ratio to that measured using the Centre telescope, the source of the internal polarisation can more precisely determined to be the light path through the beam-combiner from the spot mirror beamsplitter to the NES.

The scintillation indices are $S = 0.024 \pm 0.004$ and $S = 0.028 \pm 0.003$ from the Centre and Inner West telescopes, and are greater than the internal results. The atmospheric scintillation is expected to be much larger because of the relatively small telescope aperture size (17) (24 cm on the sky). Diffraction of light passing through the atmosphere results in wavefront fluctuations being partially converted to amplitude fluctuations at the ground. As telescope apertures become larger, there is an effective averaging of these fluctuations and the character of scintillation changes, one effect being that the variation is reduced. Conversely, small diameter telescopes suffer from a larger variability in the intensity as is seen here.

The defocus amplitude spectra are plotted for both beams in figure 5.24, and the corresponding plots of the scintillation spectra is in figure 5.25. As before, the power spectra are averaged over 11 individual estimates, each of 60 s length, and then is logarithmically binned. To ensure that the same time period from each telescope was used for NES measurements, the strip file measurements were interleaved between the two telescope. The asymmetry of image intensities was found to be larger for the Inner West than from the Centre telescope, and this effect is also apparent in the

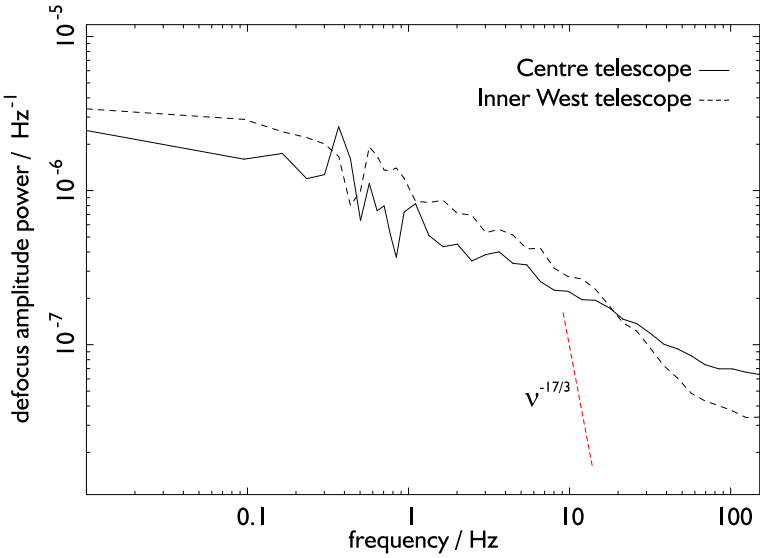

Figure 5.24.: Power spectra for NES-measured defocus aberration amplitudes from Alp Tau on 10/12/2003.

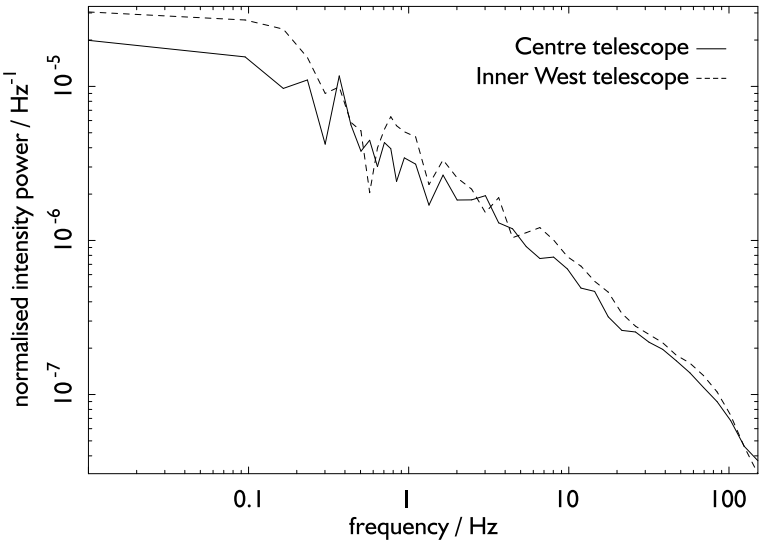

Figure 5.25.: Power spectra of the scintillation measured from Alp Tau on 10/12/2003.

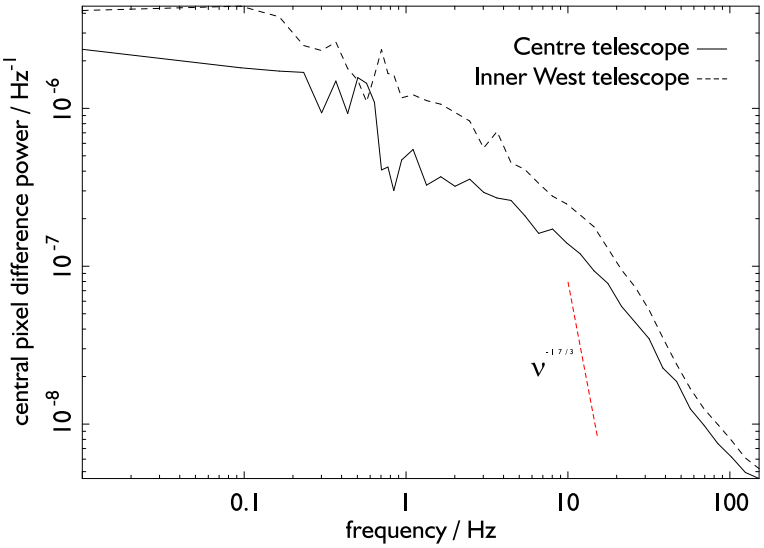

Figure 5.26.: Power spectra of the central pixel differences, Δ, from Alp Tau on 10/12/2003.

defocus spectra. In particular, the defocus spectrum measured from the Inner West telescope shows a knee at $\sim 10\,\mathrm{Hz}$ which is not apparent in the spectrum from the Centre telescope. However, the theoretical prediction for the power spectrum (67) is a $2/3$ power law at low frequencies and a $-17/3$ power law at high frequencies. Although the mismatch when compared with the experimental spectra at low frequencies can be explained by the effects of internal seeing, c.f. figure 5.22, the high frequencies for the Centre telescope spectra follow a too-shallow power law of $\sim \nu^{-1}$. Power spectrum slopes matching theory have been found in other experimental studies (87)(67)(44) using various types of WFS, so the lack of correspondence to theory here is assumed to be a flaw in the NES prototype.

The effects of scintillation on the measured defocus from using the two telescopes can also be shown by fitting a straight line to the instantaneous total flux versus the instantaneous defocus amplitude: the gradients are 0.00 ± 0.06 and -0.12 ± 0.08 for the Centre and Inner West telescopes. The correlation between flux and defocus aberrations and the power spectra shapes allows the conclusion to be made that measurements using the Inner West telescope are significantly affected by scintillation, and those from the Centre telescope are less so.

An attempt to understand the effects of scintillation on defocus measurements was made by using only the intensities in the central pixel of the 3×3 grid. This approach is simpler as it does not involve the pre-computing of the pseudo-inverse measurement matrix, hence any uncertainties in this component can be neglected. As was shown in chapter 3, the on-axis intensity difference between the intra-focal and extra-focal planes is proportional to only the diversity aberration—hence for the NES this is a measure of the defocus aberrations. Before computing this difference signal, Δ, the scintillation can be accounted for by normalising the central pixel flux, $I_{+/-;c}$, with the total flux from each image,

$$\Delta = \frac{I_{+;c}}{I_{+}} - \frac{I_{-;c}}{I_{-}}. \tag{5.3}$$

Telescope	Defocus amplitude variance	Defocus amplitude mean
Centre	$1.1 \pm 0.2 \times 10^{-2}$	$4.3 \pm 1.0 \times 10^{-2}$
Inner West	$1.0 \pm 0.3 \times 10^{-2}$	$1.4 \pm 0.2 \times 10^{-1}$

Table 5.5.: Measured NES defocus amplitude mean and variance from Alp Tau on 10/12/2003.

The averaged power spectra of Δ are shown in figure 5.26, and they are the analogue of those in figure 5.24. As well as the improved similarity of each telescope's spectra, these results differ in their form from the defocus amplitude spectra. The high frequency slopes still do not correspond to the theoretical expectation and it is thought that this is because the dark frame and flat fielding of the CCD were not sufficiently accurate. Dark frame files and flat fielding files were taken by inserting beam blockers into the beam combiner but a closer comparison of these files with the files containing the defocused images suggested there was an extra background flux in the latter. This may possibly be from "leakage" of the trolley metrology laser into the NES field of view. As the dark frame values are then unusable, they were inferred directly from the strip files by using the unilluminated areas of each frame. This approach cannot take into account any variation in background intensity i.e., it assumes no flat fielding is necessary. Nevertheless, the comparison of central pixel difference signals with the defocus amplitudes derived above does confirm that scintillation is also a significant problem for the NES with the current approach to data collection and analysis

The defocus amplitudes calculated from the matrix method were used to form a distribution and a Gaussian was then fitted to find the variance. The variance and means are shown in table 5.5, and the summary is that variance from each telescope is comparable but the mean defocus differs. As mean defocus amplitude from the internal seeing measurements was also not consistent with these results—the average from light paths involving the trolley is $1.9 \pm 2.3 \times 10^{-2}$—no specific conclusions can be drawn, as again the results may be influenced by scintillation. It is, however, consistent with there being an offset in the amplitudes which is caused by unequal image intensities; the "extra" flux in the brighter image does not contribute to the defocus measurement (equation 4.2) and becomes a constant error term.

Nonetheless, the defocus amplitude variance measured from the Centre telescope is approximately the same as that from the Inner West telescope (the former is 10% larger) and both are considerably larger than the internally measured variance. They can be converted to D/r_0 with two assumptions: first, the seeing is fully described by by Kolmogorov statistics, so the expected variance of defocus aberrations is $0.022\ (D/r_0)^{5/3}$ and, secondly, the amplitude calibration factor is equal to that derived in the laboratory (1.95). The results then correspond to D/r_0 (at 500 nm) of 2.53 and 2.60 for the Centre and Inner West telescopes, or r_0 of 9.5 cm and 9.2 cm. These values have the correct order of magnitude for the COAST site but there is obvious uncertainty as to their accuracy.

To investigate the effect scintillation has on the NES measurements, a numerical simulation was used, incorporating propagation from the compressed telescope aperture (diameter 25 mm) to the spot mirror beamsplitters (diameter 15 mm). Propagation was modelled using Fresnel diffraction

(22) over a distance of 15 m, the seeing was set to be $D/r_0 = 3$, and the wind speed of the single phase screen was 15 ms^{-1}. The simulation used three image intensity ratios to model the polarisation effects: 1.0 (no polarisation effects), 1.1 and 1.5.

The scintillation index from the simulations is 0.006, notably smaller than the experimental values because the scintillation at the telescope aperture is neglected. Simulating longer propagation paths to increase the value of the index would become unphysical for two reasons. The internal propagation introduces a spatial filtering effect as aberrated parts of the wavefront "diffract" out of the beam, and the diffraction from the sky differs in character because there is no aperture boundary so light can diffract out of and into the telescope aperture.

The defocus amplitude variances when using equal image intensities resulted in a loss of less than 1%, and this decrease is from propagation alone. When compared with the results for intensity ratios of 1.1 and 1.5, the decrease in defocus amplitude variance was 2% and 10% respectively. Also, as expected, the mean defocus amplitude increases with larger image intensity ratios. The polarisation therefore clearly affects the defocus amplitudes and these results are consistent with the experimental measurements. The power spectra from simulations, however, did not show the experimentally predicted high-frequency slopes. The theoretically predicted high frequency behaviour—a slope of $-17/3$—was seen in all the simulation spectra and, when these compared with each other, the effects of both propagation and polarisation on the measured gradients was negligible. Modelling more than one moving phase screen, with non-parallel velocities, also did not alter the lack of correspondence. The experimental power spectra, therefore, remain unexplained. It is finally noted that the non-linear measurement behaviour of the NES, as seen experimentally within the laboratory and in simulations (chapter 4), was observed in these simulation as a flattening (saturation) of power at low frequencies. As there is no evidence of this effect in the internal or atmospheric defocus measurements in this chapter, it is suggested that non-linearities are not a significant factor here.

5.6. Conclusion

The design of the NES prototype in hardware and software has been described and the performance verified in laboratory experiments. Using two focal length lenses, either a CWFS or a NES can be created. The results of measuring defocus with either configuration are shown to be linear, within a small region near zero amplitude for the latter as expected. The use of broadband light with the NES demonstrated that there are chromatic effects that were not previously described, but these can be quantified by using an amplitude calibration factor. The prototype WFS design presented is nonetheless useful for making qualitative measurements. The subsequent measurements, made within the COAST bunker and of atmospheric seeing, used this prototype and were affected by scintillation. An unpredicted polarisation effect, which made the intensity of the POB images differ, means that an alternative image normalisation is required. The alternative defocus estimate proposed, using only the difference of central pixel fluxes, showed that scintillation effects were reduced but not eliminated. The variance of defocus aberrations measured on the sky suggested that these measurements are of the correct order of magnitude, and comparisons with a numerical

simulation incorporating propagation effects explained some of the discrepancies between theory and the experimental results. However, the power spectra of the experimental defocus amplitudes remained problematic, as the high-frequency behaviour was not described by either theory or the simulation results.

6. COAST Autoguider Analysis

6.1. Introduction

When observing with a stellar interferometer, being able to parameterise the seeing, in terms of the temporal (t_0) and spatial (r_0) behaviour, can be valuable. Knowing the spatial seeing allows for quantitative planning of observations e.g., poor seeing can prevent high accuracy measurements of small intrinsic visibilities and time may be better spent on less demanding observations. Temporal measurements are already available at COAST, from the power spectra of recorded fringes, but spatial seeing has so far only been qualitatively estimated.

The COAST autoguider AO system has operated since the telescope's inception but it has not previously been studied in any detail. The autoguider is responsible for correcting the tip/tilt aberrations in the four beams entering the laboratory bunker, and aids the telescope control system to track the target star via the telescope siderostats. The system can thus be split into two: slow guiding (at a rate of $\sim 1\,\mathrm{Hz}$) which corrects tracking errors, and fast guiding (operating at $\sim 100\,\mathrm{Hz}$) to remove atmospheric tip/tilt aberrations. These aberrations are corrected for with fast mirrors, which are flat mirrors with piezo-electric actuators which allow the mirror to rotate about the vertical and horizontal axes. Similar AO systems exist at other interferometers and these have also been used for spatial seeing studies (83)(55)(41).

The COAST autoguider was initially modified to automatically record the data from its WFS onto computer disc during regular observations. In the context of the theory of tip/tilt aberrations and their correction, the data from the autoguider system was analysed to understand its appropriate theoretical description. Both the temporal power spectra and distributions of the fast mirror angles were examined. The former revealed various problems with the AO system feedback loop and required a more detailed study of the slow guiding implementation. As a result, a numerical simulation of the entire autoguider was used to investigate the slow guiding performance. The remainder of the chapter uses distributions of fast mirror angles to measure the spatial seeing. From a year long dataset, the statistics of the spatial seeing at the Lord's Bridge Observatory, the location of COAST, are presented.

6.2. Theory of tip/tilt correction

To measure the tip/tilt correction, the autoguider uses a Quadrant Cell (QC) for up to four beams from the telescope array. When fast guiding, the aim is to to reduce both the vignetting of the beams and visibility coherence losses. It is correct to state that *most* image motion results from the tip/tilt Zernike polynomials (Z_2/Z_3), and their amplitudes are called the Z-tilt. However, higher order $m-$odd phase aberrations also cause image motion.

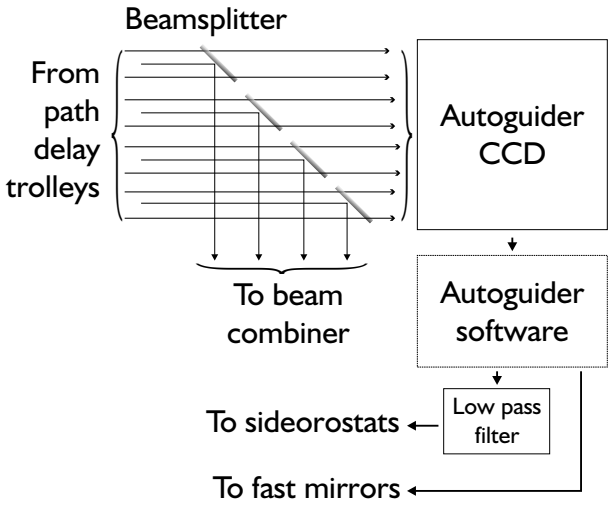

Figure 6.1.: Schematic diagram of the COAST autoguider. The beamsplitters are spot mirror which allow the outer annulus of the incoming beams to be used for the autoguider, while the central part is reflected towards the beam combiner.

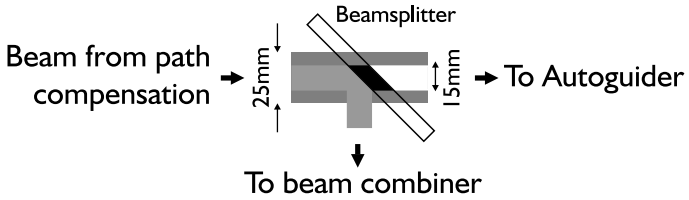

Figure 6.2.: Spot mirrors used as beamsplitters for the COAST autoguider. The diameter of the outgoing beam is 15 mm and the outer annulus has a diameter of 25 mm with a thickness of 5 mm.

The total effect of phase aberrations on image motion can be shown by first writing the phase across a beam pupil, $\Phi(x, y)$, as a Taylor expansion,

$$\Phi(x, y) = \Phi_0 + \Phi_x\, x + \Phi_y\, y + O\left(xy, x^2, y^2\right), \tag{6.1}$$

where Φ_0 is the mean phase, and Φ_x and Φ_y are the amplitude of the linear terms. The mean phase and high order terms can be ignored, and the PSF c.f., equation (2.15), of the aberrated beam is

$$I^*(u, v) = \frac{i}{f\lambda}\left|\int\int dx\, dy\; P(x, y) \exp\left[-i\frac{k}{f}(xu + yv) + i\Phi_x\, x + i\Phi_y\, y\right]\right|^2. \tag{6.2}$$

The linear phase terms can be incorporated into the focal plane coordinates: $(u', v') = (u, v) - f/k\,(\Phi_x, \Phi_y)$. In comparison to an unaberrated PSF, this PSF has a translation in the focal plane—the image motion—which is a function of Φ_x and Φ_y. These are the average phase gradients in two directions across the pupil; $\Phi_x = \langle \partial\Phi/\partial x \rangle$ and $\Phi_y = \langle \partial\Phi/\partial y \rangle$. The QC signals therefore measure these gradients which are called the G-tilts. It has been shown (81) that the amplitude of G-tilt is

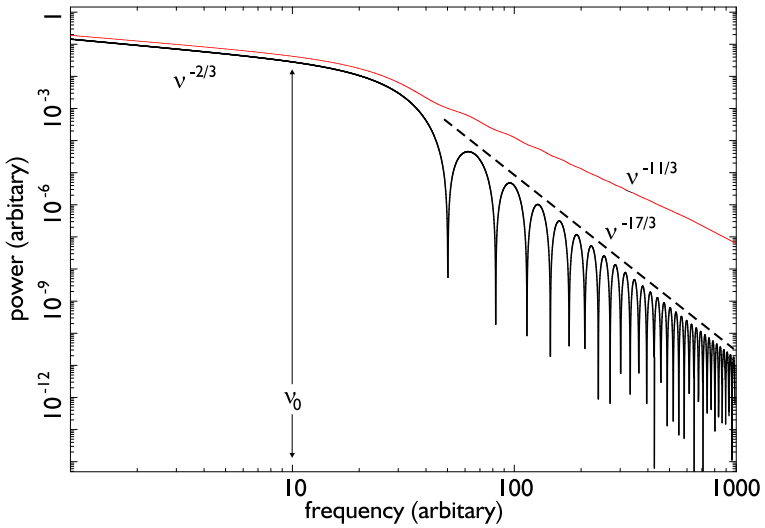

Figure 6.3.: Theoretical power spectra for Z-tilt (black) and G-tilt (red). The oscillations for the Z-tilt, which are not of interest here, have unequal minima for numerical reasons.

equivalent to the amplitudes of all Zernike polynomials with angular index $m = 1$, which include terms proportional to x or y,

$$\begin{aligned}\left\langle \frac{\partial \Phi}{\partial x_i} \right\rangle &= \left\langle \frac{\partial \Phi}{\partial \rho} \right\rangle \frac{\partial \rho}{\partial x_j} + \left\langle \frac{\partial \Phi}{\partial \theta} \right\rangle \frac{\partial \theta}{\partial x_j} \\ &\equiv \sum_{n=0} a_{2n+1,1}\, Z_{2n+1,1}\,(1,\theta)\,. \end{aligned} \tag{6.3}$$

The variance, in rad rms, of the G-tilt along one axis is then

$$\begin{aligned}\sigma_G^2 &= \sum_{n=0}\sum_{n'=1} \frac{1}{4} Z_{2n+1,1}\,(1,\theta)\; Z_{2n'+1,1}\,(1,\theta)\,\langle a_{2n+1,1} a_{2n'+1,1}\rangle \\ &\simeq \sum_{n=0}^{2} \frac{1}{4} Z_{2n+1,1}^2\,(1,\theta)\,\langle a_{2n+1,1}^2\rangle = 0.598\left(\frac{D}{r_0}\right)^{5/3}. \end{aligned} \tag{6.4}$$

The final line is a result of two approximations; the cross-correlation between terms are neglected, as they are small, and only the first three terms are used, as the remainder increase the value by less than 1%. The equivalent Z-tilt variance is $0.456(D/r_0)^{5/3}$ and is therefore significantly less.

The theoretical temporal 1D power spectra of Z-tilt and G-tilt, using Taylor's hypothesis to convert from spatial to temporal coordinates, has been calculated by several workers (81)(77)(12). The spectra share a common low frequency component, proportional to $\nu^{-2/3}$ where ν is the frequency, and then a knee followed by a steeper power law. The Z-tilt spectrum beyond the knee is bound by $\nu^{-17/3}$ whereas the G-tilt spectrum is described by $\nu^{-11/3}$. The knee corresponds to the time taken for a (assumed) single phase screen to move a distance equal to the telescope aperture diameter. A comparison of the theoretical power spectra are plotted in figure 6.3.

To find r_0 from the recorded image motions, there are two possible approaches. Either the one axis variance can be computed directly from the fast mirror angles or the power spectrum of the

angles can be used. The latter has the advantage of also being able to measure the temporal behaviour of the seeing. Existing experimental results, however, show some disagreement in the measured spectral power laws: estimates of the high-frequency slopes range from $-8/3$ (75)(61) to $-11/3$ (80). The question of which image motion to use—Z-tilt or G-tilt—will first be determined from the high-frequency behaviour of the autoguider data.

6.3. COAST autoguider overview

The autoguider WFS consists of a CCD upon which four light beams are imaged. Each image has a (virtual) QC associated with it, consisting of 4 binned CCD pixels, and 16 pixels are readout every ~ 6 ms to calculate the instantaneous image motion. These errors are added to the tip/tilt aberration estimate for each beam, to form the current image motion estimate per beam. These estimates are range-checked—the values are truncated within a pre-defined range—and then sent in arbitrary units, known as pseudo-volts, to amplifiers controlling the fast mirror actuators. The angle estimates are also low-pass filtered and sent to the Telescope Controller (TC). This filtered signal is used for the slow guiding, to correct tracking errors and prevent the fast mirrors saturating at their range limits.

The autoguider software was modified to record the QC error signals rather than the actual angles of the fast mirrors. These angles are unavailable, unlike in previous studies on other AO systems, and must instead be reconstructed offline. The only time that the mirror angles are known is when fast guiding begins because, at the beginning of the night, the mirrors are aligned to a zero angle perpendicular to a reference beam (from the artificial star).

6.3.1. Annular tip/tilt

The autoguider WFS uses beamsplitters to remove light from the incoming beams, but their design is unusual; the light is divided spatially by picking off an outer annulus as shown in figure 6.2. The phase at the edge of an aperture (an infinitely thin annulus) is sufficient to determine either the G-tilt or Z-tilt within it. However, a finite-thickness annulus will measure a different image motion from that using the central aperture. The effect of this arrangement was first analysed using a numerical simulation.

The simulations used a QC in a closed-loop, first using the entire aperture for the WFS and then an annulus outside of the aperture (as at COAST). The resulting power spectra of the measured image motions and the actual Z-tilt are plotted in figure 6.4. The QC signals from using the central aperture are a good approximation to Z-tilt, whereas the annulus results measure the low frequency Z-tilts but underestimate the power beyond the knee. As the power law slopes for the annular case are consistent with Z-tilt, the difference between the WFS apertures is only characterised by a shift of the spectrum knee frequency. Since most of the image motion in either the G-tilt or Z-tilt case is at low frequencies, these results confirm that the annulus measurement can measure the image motion power.

The main result is that the power spectrum of image motions measured using a finite thickness annulus cannot be compared directly with the theoretical power spectrum from using a filled

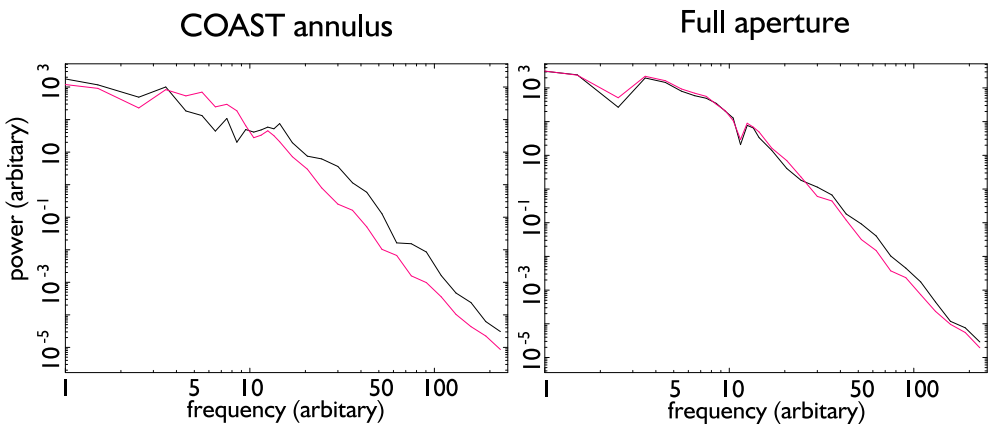

Figure 6.4.: Power spectra for numerical simulations using an outer annulus (COAST dimensions) and the full aperture with a Quadrant Cell. The Z-tilt power spectrum is in black, the tip/tilt errors in red, and the knee at $\sim 10\,\mathrm{Hz}$.

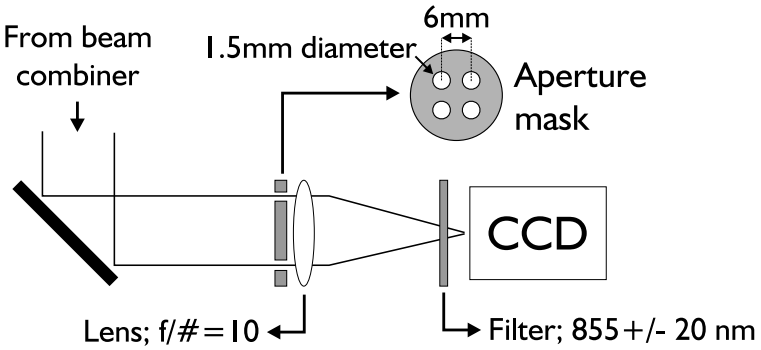

Figure 6.5.: The outline of the Fizeau Interferometer used for calibrating the fast mirrors.

aperture. There is no equivalent theoretical spectrum for using an annulus, and therefore only qualitative analysis of the autoguider power spectra can be made. It is also noted that the Z-tilt high frequency power law in these simulations is $\sim \nu^{-15/3}$ rather than the expected $\nu^{-17/3}$, so these results do not resolve whether Z-tilt or G-tilt is the image motion measured by the autoguider.

6.3.2. Fast mirror calibration

The autoguider AO loop is a closed feedback loop and so operates by an (essentially) iterative process. The feedback does not require the control software to know the exact angle of each fast mirror nor the rotation of the fast mirrors per pseudo-volt. However, this calibration factor is needed for the reconstruction of the fast mirror angles in software, otherwise the recorded data can only be scaled to pseudo-volts.

The angle of rotation for each axis of a fast mirror, in terms of pseudo-volts, was measured using a Fizeau Interferometer (FI). The experiment is shown in figure 6.5, and its position on the beam combiner table within the COAST laboratory bunker is shown in figure 6.6. The FI consists of a lens and an aperture mask with four 1.5 mm holes spaced 6 mm apart and arranged at the corners

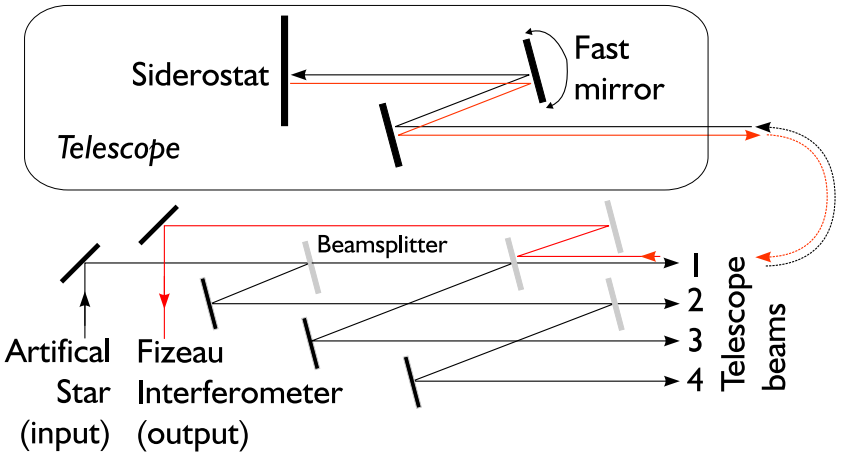

Figure 6.6.: A sketch of the calibration experiment configuration in COAST. The light from the artificial star (black) travels out to the telescope siderostat via the fast mirror. The returned beam (red), reflects again from the fast mirror and is then diverted to the Fizeau Interferometer.

of a square. This results in two pairs of baselines, each perpendicular to the other, which enhances the flux reaching the CCD detector. The bandwidth filter (855 ± 20 nm) was used to both define the wavelength and increase the fringe contrast. A fast mirror rotation causes a phase gradient in the incoming light, and this causes the fringes to move on the CCD. A Fourier transform of the image reveals this movement by a change in the phase at the frequency of the relevant fringes. By aligning the mask with the mirror rotation axes, each pair of baselines measures rotation of the mirror around one axis only.

The light from the artificial star is first sent through the beam combiner, splitting it four ways, and then reflected from the spot mirror beamsplitters. These mirrors restrict the outgoing beam diameters to 15 mm so the tilts in the beam returning from a telescope cannot be measured by the autoguider QCs. Each telescope can be tested independently by inserting beam blockers into the beam combiner so that there is only one output. The output light then passes through the path compensation system, with the trolleys static, and enters the telescope. It reflects from the fast mirror and then the siderostat, which is positioned so that the beam is reflected back toward the laboratory. The returning beam reflects again from the fast mirror and retraces its path to the beam combiner. One exit of the beam combiner output is then diverted into the FI.

An example of a typical fringe pattern and the associated Fourier transform is shown in figure 6.7, and the phases at the points marked '1' and '2' are those used in this experiment. Using interference fringes allows their movement to be directly related to the wavefront gradient via two variables: the baseline length and the wavelength. These two quantities are well defined so precise focusing of the fringes is not important. It was found that measurements during the daytime were not possible because the turbulence generated within the beam transport pipes blurred the fringes, and the experiment was carried out approximately 1 hr after sunset on a cloudy day.

A similar measurement of mirror rotation was carried out at the SUSI telescope (80) but instead used a Michelson interferometer arrangement with a laser. An advantage of the FI is that a coherent

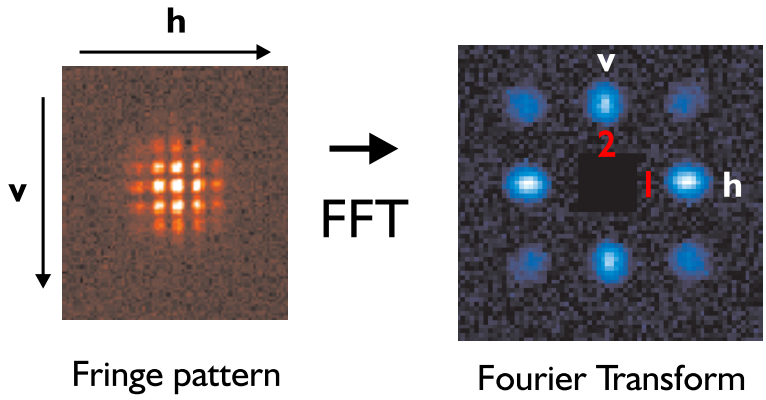

Figure 6.7.: An example fringe pattern and the resulting Fourier transform (to enhance the contrast, the zero frequencies have been removed). The letters refer to the horizontal and vertical baselines.

source is not needed and the optics are compact; this is important as COAST is an operational telescope and the experiment had to fit into the existing space without significantly disrupting any future observations.

The experiment was carried out on two telescopes: the Centre and East. The mirror actuators use PZT materials and are therefore susceptible to hysteresis, so the experiment was begun at the zero position of the fast mirrors. A 15 s exposure was taken per 10 pseudo-volt increase of an actuator channel, and after reaching the highest value (200 pseudo-volts) the mirror was rotated to the opposite of its range (10 pseudo-volts). Significant hysteresis was introduced during this movement and is seen as a discontinuity in the phase of the FI fringes. The experiment was then resumed until the mirror reached its zero position; its position was confirmed by visual examination of images on a TV monitor. The experiment was then repeated on the other mirror actuator channel. Fringes recorded with mirror rotations that caused significant vignetting were not used in the analysis.

The phases determined from the fringes are modulo-2π and were unwrapped by hand. After correcting for the hysteresis, by offsetting the phases from negative rotation angles by a fixed amount, a straight line fit to measure the gradient was made. The stages in the data processing and a sample fit are shown in figure 6.8.

The fringes from both baselines were analysed for each actuator channel to confirm that the mirror rotation was independent about each axis. The calibration factors, converted to micro-radians per pseudo-volt, are shown in table 6.1 and show that the actuator channels are independent. Interestingly, the rotation per pseudo-volt is not equal between the mirrors, nor between the two axes of the East telescope fast mirror. The calibration factor about the vertical axis for both mirrors was also determined from an independent method (using calibrated image displacements on a TV monitor) and this gave consistent results within the experimental error.

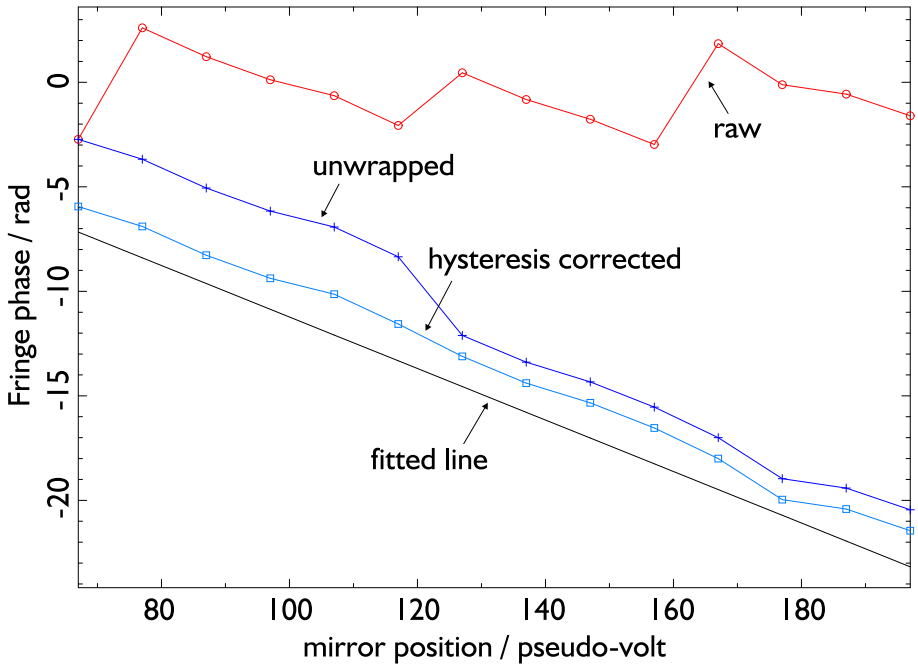

Figure 6.8.: Example fringe phases from the Fizeau Interferometer experiment. The data and fit have been purposely offset from each other since only the gradient is of interest.

Telescope	**Axis of rotation**	**Rotation/ 10^{-6} rad channel 1**	**channel 2**
Centre	Horizontal	-0.03 ± 0.13	-2.79 ± 0.05
	Vertical	2.78 ± 0.06	-0.19 ± 0.07
East	Horizontal	0.30 ± 0.14	-3.32 ± 0.13
	Vertical	2.27 ± 0.13	-0.21 ± 0.10

Table 6.1.: Measured fast mirror rotation in micro-radians per pseudo-volt. Note that these values are half as big for reflection from starlight.

6.4. Fast Mirror Rotation Angle Reconstruction

The fast mirror rotation angles are reconstructed by integrating the QC signals, using a replica of the autoguider software algorithms. If the image position measured by the QC, and scaled to pseudo-volts, at a given time is x_n then, using a gain of β (nominally 1), the mirror angle is y_n,

$$y_n = y_{n-1} + \mathrm{Int}\left[\beta\, x_n\right], \tag{6.5}$$

where $y_0 = 0$ and the 'Int' operator represents rounding down to the nearest integer. Using the calibration factors determined above, the angles can be converted to radians for the Centre and East telescopes only. Also, the angle on the sky is 16× smaller because of the magnification from telescope primary and secondary mirrors (which weren't used in the calibration experiment). Both of these factors are a scaling of angles and can be ignored until r_0 is calculated, hence the angular units used for analysis are in pseudo-volts unless otherwise specified. The reconstructed mirror angles are then either plotted as a distribution or a temporal power spectrum.

6.4.1. Mirror Angle Distributions

The image motion from atmospheric phase aberrations has a Gaussian distribution with a zero mean. The distributions of angles from four mirrors (equally telescopes) from 240 s of fast guiding is shown in figure 6.9. All the distributions show a non-zero mean and several are asymmetric. The sharp peaks at the highest and lowest angles represent the fast mirrors saturating at their range limits. This angular range is $\approx \pm 2.5$ arc seconds, which is the 2-σ limit of poor seeing ($r_0(500\,\mathrm{nm}) \simeq 4\,\mathrm{cm}$). An example time series of fast mirror angles is shown in figure 6.10, and it can be seen that the mirror moves from one end of its range (at time $\sim 35\,\mathrm{s}$) to the other (at $\sim 45\,\mathrm{s}$). The explanation of the saturation is therefore unlikely to be from seeing.

6.4.2. Mirror Angle Power Spectra

The cause behind the saturation and distribution asymmetry can be found in the power spectra of the fast mirror angles. An example spectrum from the same data as figure 6.9 is plotted in figure 6.11. It was first logarithmically binned, so the bin widths are constant in log frequency, and the red line is a fit using simple power laws,

$$P(\nu) = P_0 \left(\left(\frac{\nu}{\nu_0} \right)^{\alpha\gamma} + \left(\frac{\nu}{\nu_0} \right)^{\beta\gamma} \right)^{1/\gamma}.$$

The parameters α and β are the underlying low and high frequency power laws, γ is the knee "sharpness", and ν_0 and P_0 are scaling factors. As the fit shows, neither the low nor high frequencies match the theoretical spectra for either Z-tilt or G-tilt. However, if the rounding component of equation (6.5) is not present then the fit for this data returns $\alpha = -0.86$ and $\beta = -3.68$ ($-2.58/3$ and $-11.80/3$ respectively). The high frequency image motion is of increasingly smaller amplitude and so is more affected by the rounding. Therefore, the correct measure of image motion by the autoguider is G-tilt. A distribution of fits to 449 power spectra of mirror angles with and without

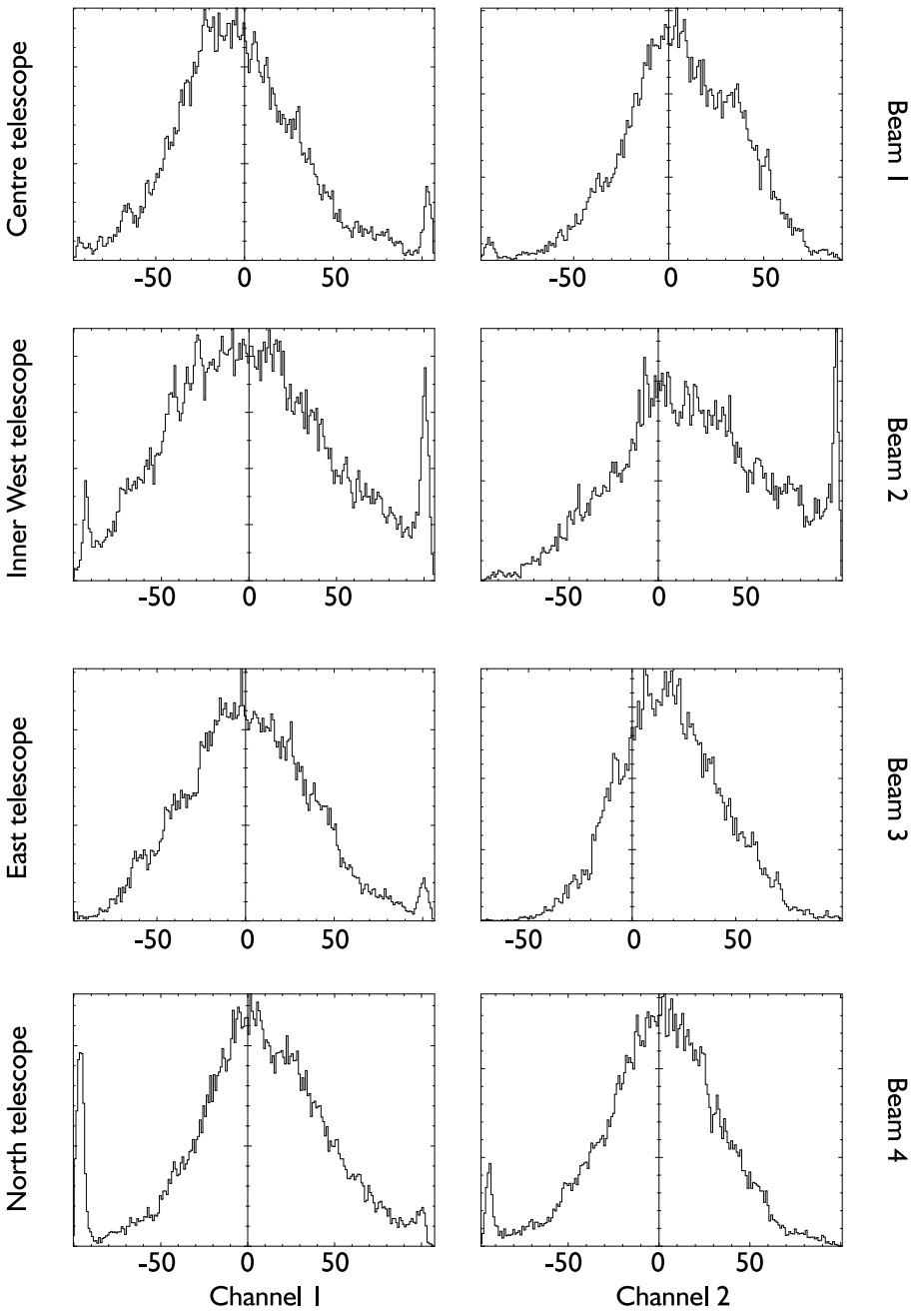

Figure 6.9.: Fast mirror angle distributions, in pseudo-volts, from $\sim 240\,\mathrm{s}$ of fast guiding.

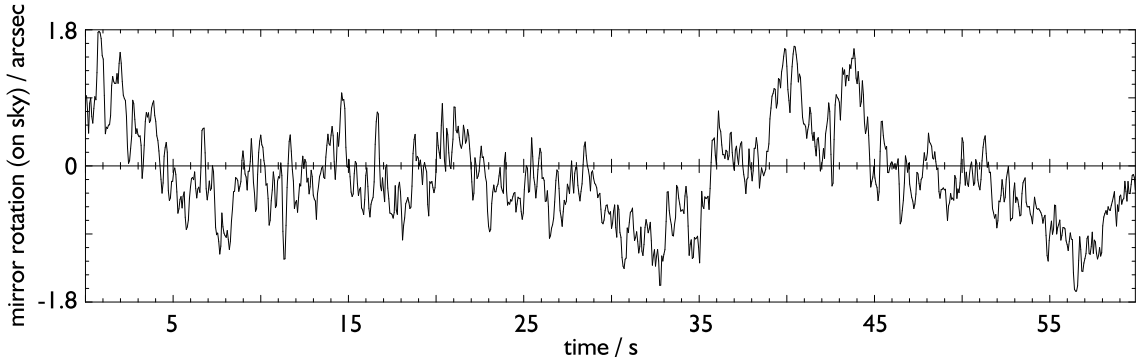

Figure 6.10.: Fast mirror angle for 60 s of fast guiding, from one actuator channel of the Centre telescope.

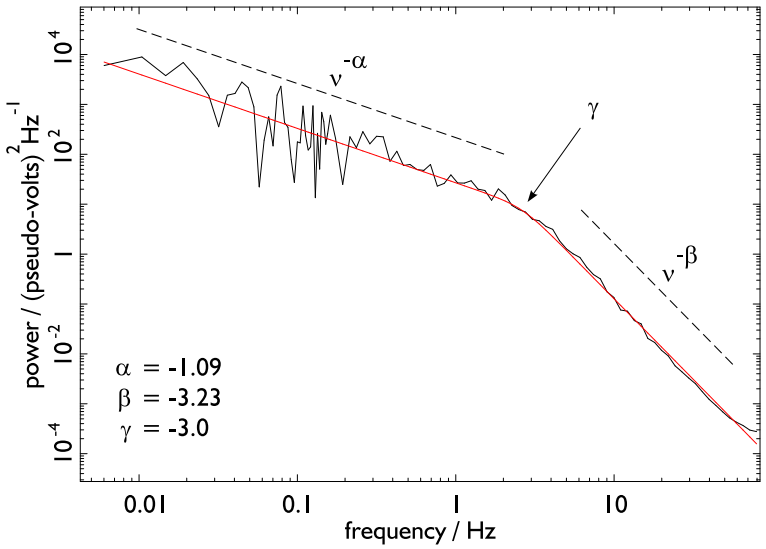

Figure 6.11.: Power spectrum from one actuator channel of the Centre telescope, the same data as plotted in figure 6.9, and the associated fit.

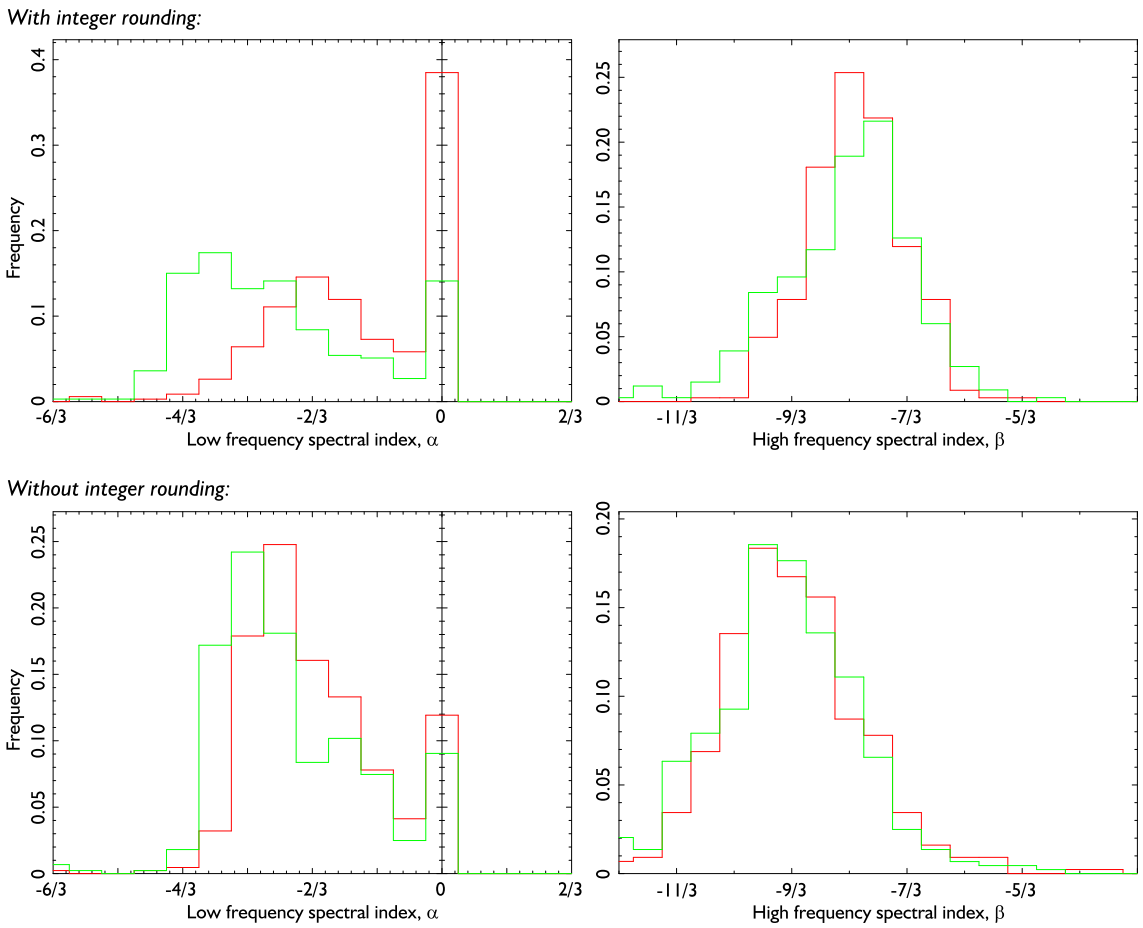

Figure 6.12.: Distribution of parameters α and β from fits to the spectra of the reconstructed mirror angles. Only data from the Centre telescope, vertical (red) and horizontal (green) axes, are used.

the integer rounding operator is shown in figure 6.12, and shows that not using rounding gives more theoretically consistent results.

The effect of integer rounding can also be shown by examining the power spectrum of the quadcell signals themselves. The Discrete Fourier Transform is written as $\mathcal{D}\{\}$, and the power spectrum of the reconstructed mirror angles as $\mathcal{P}_M = |\mathcal{D}\{y_n\}|^2$. Let the spectrum of the quadcell signals then be written as $\mathcal{P}_Q = C|\mathcal{D}\{y_{n+1} - y_n\}|^2$. The transform in the latter can be written as,

$$\begin{aligned}
\mathcal{D}\{y_{n+1} - y_n\}_m &= \mathcal{D}\{y_{n+1}\}_m - \mathcal{D}\{y_n\}_m \\
&= \mathcal{D}\{y_n * \delta_{n+1}\}_m - \mathcal{D}\{y_n\}_m \\
&= \mathcal{D}\{y_n\}_m \ (\exp[i\,2\pi m/N] - 1) \\
&= \mathcal{D}\{y_n\}_m \ \exp[i\,\pi m/N] \ \sin(2\pi m/N).
\end{aligned}$$

The spectrum of the quadcell signals is then,

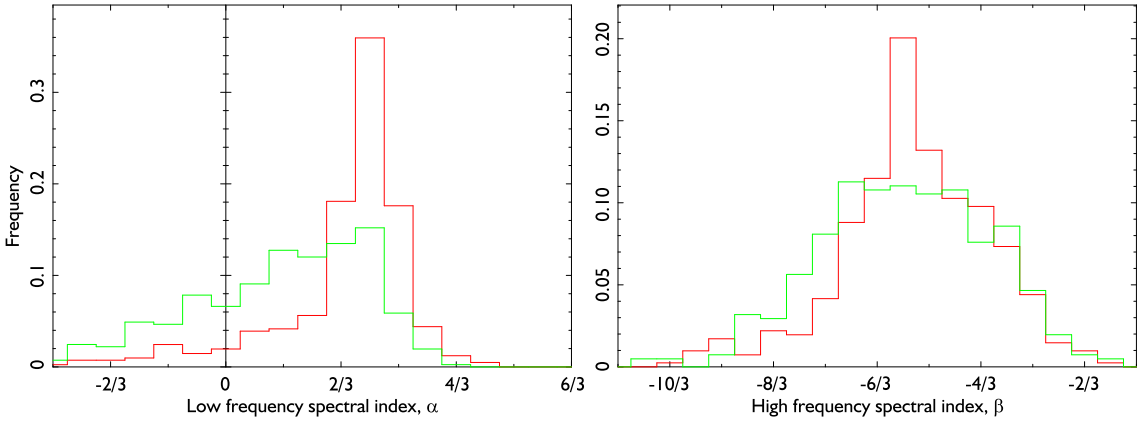

Figure 6.13.: Distribution of parameters α and β from the fit to the power spectrum of the quadcell signals. The fits use the data from vertical (red) and horizontal (green) axes of the Centre telescope only.

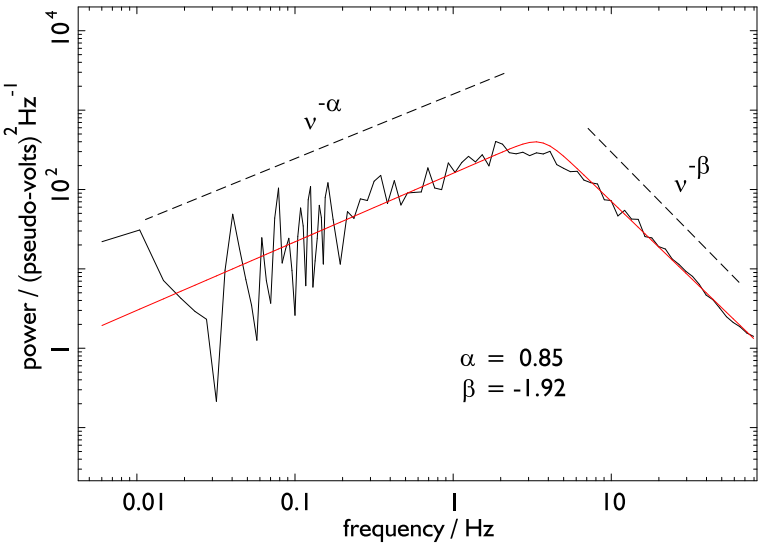

Figure 6.14.: Power spectrum for the quadcell signals (horizontal axis) for the Centre telescope, from the same data as plotted in figure 6.9.

$$\begin{aligned} \mathcal{P}_Q(m) &= C\,\mathcal{P}_M(m)\,\sin^2(2\pi m/N) \\ &\simeq \frac{4\pi^2 C}{N^2}\,\mathcal{P}_M(m)\,m^2 \end{aligned} \tag{6.6}$$

where the final line shows an approximation for low frequencies (because the argument of the sine is equal to or less than π so quartic and higher terms may be neglected except near the Nyquist frequency.) The power laws are therefore altered from the mirror angle spectrum by multiplication of ν^2: low frequencies should follow a 4/3 power law and high frequencies a $-5/3$ power law. To account for the approximation made above, power at frequencies near the Nyquist frequency are not used in the fits, and one example is shown in figure 6.14. A distribution of fits to quadcell

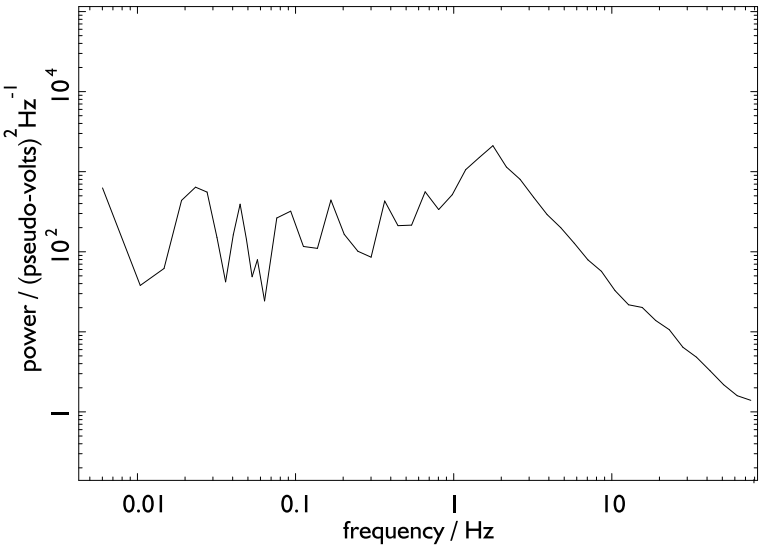

Figure 6.15.: Power spectrum for the quadcell signals (horizontal) from the Inner West telescope, note the extra power around the knee at 2 Hz.

signal spectra for both axes of the Centre telescope fast mirror is shown in figure 6.13: the mean high frequency fit for the two axes is for slopes of -1.83 and -1.76 (equivalent to $-5.27/3$ and $-5.46/3$). Note that these averages have a better correspondence to the G-tilt theory than the fits to the spectra of the mirror angles. One reason for this may be that the reconstruction needed for mirror angles is non-linear, because it involves range checking.

Neither the high nor low frequencies fits to the spectra of the reconstructed (with rounding) mirror angles correspond to the theoretical predictions for Z/G-tilt. There cannot therefore be a certainty as to which description is correct. As the rounded data is *better* described by G-tilts, the remainder of this work will use the *assumption* that the variance of the reconstructed mirror angles, using rounding, is described by the G-tilt one-axis variance (6.4).

The low frequency image motion does not show the expected behaviour nor a decrement in the spectrum from slow guiding. These imply that the fast mirrors are being used, instead of the siderostats, to correct for a large fraction of the tracking errors. This will then cause the fast mirrors to reach their range limits. These errors are then seen in the distributions as the saturation and the asymmetries.

Examining the log-log spectra of the quadcell signals from the Inner West telescope revealed anomalous power at the knee transition, an example spectrum is shown in figure 6.15. The distribution of mirror angles for this telescope also reached the range limits more often than those from the other telescopes. Plotting the angle spectrum using a frequency times power versus log frequency has the advantage that the area under the curve is then proportional to power,

$$\int_{\log \nu_1}^{\log \nu_2} d \log \nu \; \nu \, \Phi(\nu) \equiv \int_{\nu_1}^{\nu_2} d\nu \; \Phi(\nu). \tag{6.7}$$

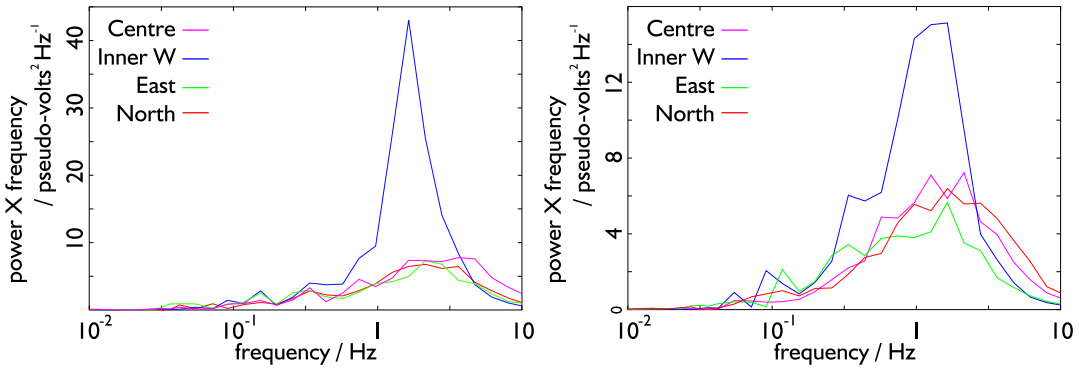

Figure 6.16.: Mirror angle power times frequency against log frequency for the four beams, from the vertical (left) and horizontal (right) axes.

Such a spectrum for the four beams is plotted in figure 6.16 which shows that most power is concentrated around the knee frequency. It also shows an excess of power for the Inner West telescope when compared with the other telescopes. To explain these results, the response of the fast mirror in the Inner West telescope was compared to those in the Centre and East telescopes: an image was visually observed on a TV monitor and a step movement was sent to the fast mirror actuator channels. The motion of the Centre and East telescope's fast mirrors was faster than the eye could follow, as expected, but the Inner West telescope's mirror took $\sim 1\,\mathrm{s}$ to travel 50% of its range.

The poor response of this fast mirror means that only low frequency image motions are correctly compensated, and at around $\sim 1\,\mathrm{Hz}$ the mirror response would suffer from a significant phase lag. Positive feedback would be the result and this manifests itself in the mirror angle reconstruction. At the time of writing, this fast mirror had been identified as installed at a different time to the other mirrors and further investigations were taking place.

It is interesting to note that this mirror had been operating in the Inner West telescope for some years. A considerable dataset of fringes has been obtained using it and no particular bias has been found. The problem identified here, however, does prevent the inference of r_0 from this telescope's fast mirror angles. The North and Outer West telescopes were also not used in this study as they were not operated as frequently, for technical reasons, as the Centre and East telescopes. The reduced autoguider data is therefore only from the latter two telescopes.

6.5. Slow Guiding Analysis

The sky rotation is calculated by the telescope controller (TC) during fast guiding and the errors in the resulting siderostat rotation feeds back via the filtered signals from the autoguider. This process can be divided into two stages: first, the WFS signals are filtered and sent to the TC and, secondly these signals are used by the TC to control the siderostat.

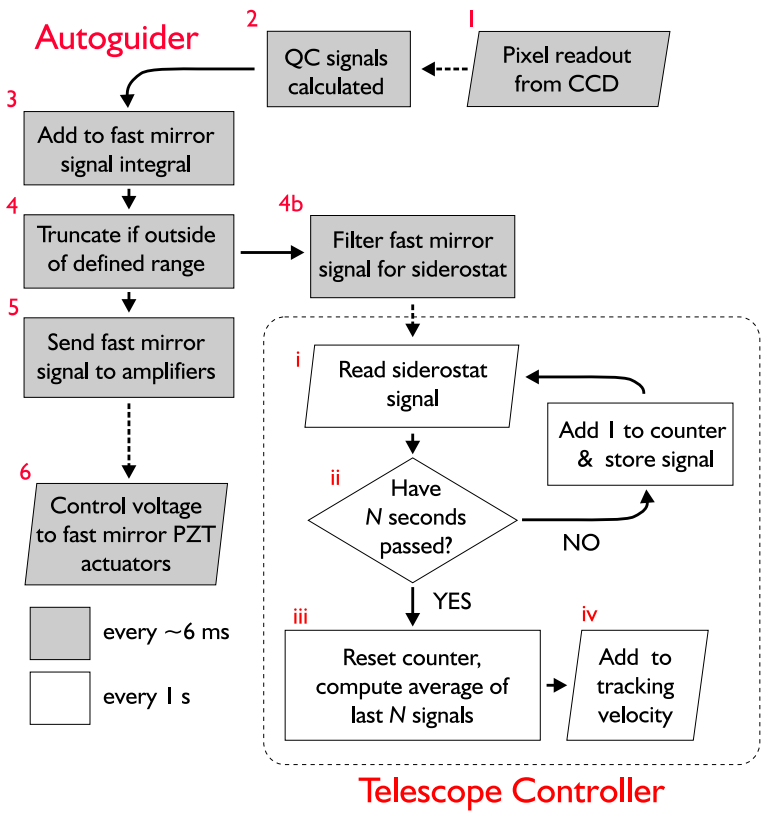

Figure 6.17.: Flow diagram for the Autoguider and Telescope Controller. Only the components relevant to fast guiding are displayed.

6.5.1. Autoguider Filter Response

The following description assumes a basic knowledge of system control theory and the Z-transform. The response of the autoguider digital filter will be studied here to understand at which frequencies the correction is dominated by the fast mirror and where the siderostat correction is expected to take over.

The description here will use one beam as an example; it is identical and occurs in parallel for 4 beams (or subset thereof.) During each iteration of the autoguider feedback loop, taking Δ seconds, calculations are made of the fast mirror angle, y_i, (6.5) and its filtered version, l_i. The filter used is of an exponential Infinite Impulse Response (IIR) type,

$$l_i = e^{-\alpha}\, l_{i-1} + (1 - e^{-\alpha})\, y_i, \tag{6.8}$$

where α, by analogy with a R-C filter, can be converted to a time constant: $\tau = \Delta / \alpha$. This filtered signal is called the autoguider siderostat signal and is the one sent to the TC, as shown on the flowchart in figure 6.17.

Analysis of equation (6.8) in frequency space is more straightforward. Its Z-transform is

$$L_j = H(z)\, Y_j = \frac{(1 - e^{-\alpha})}{1 - e^{-\alpha}\, z}\, Y_j. \tag{6.9}$$

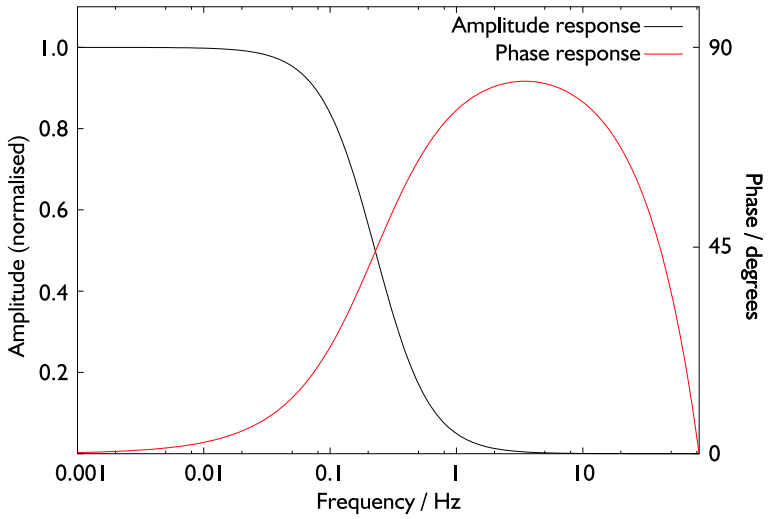

Figure 6.18.: Amplitude and phase response of the autoguider filter.

The feedback transfer function, H, is of a standard form and the equivalent response in the time domain (sampled every Δ seconds) is a convolution with an exponential, $\exp[-\alpha\Delta\, i]$. Converting H to continuous form,

$$H(s) = \frac{(1 - e^{-\alpha})}{1 - e^{-s\Delta}e^{-\alpha}}. \tag{6.10}$$

The frequency response is derivable by replacing s with $i\omega$ and computing the (square) amplitude of $H(i\omega)/H(0)$,

$$\frac{|H(i\omega)|^2}{|H(0)|^2} = \frac{1 + e^{-2\alpha} - 2e^{-\alpha}}{1 + e^{-2\alpha} - 2e^{-\alpha}\cos[\omega\Delta]} \text{ where } \omega \leq \frac{\pi}{\Delta} \tag{6.11}$$

The amplitude and phase response of this filter is plotted in figure 6.18 and of interest here is the -3 dB response, where the amplitude is half. The time constant depends on the CCD integration time (usually 1 ms) and the sampling time ($\Delta \sim 6 \times 10^{-3}$ s): normally $\tau = 0.7$ s. The -3 dB response frequency is then 0.40 Hz which, compared to a knee frequency of $\sim$ 1 Hz in figure 6.11, mainly allows through power at frequencies lower than that of the knee. This filter is therefore a good choice if the siderostat response at these low frequencies is non-zero and has small phase lag.

6.5.2. Siderostat Response

To understand the response of the siderostat and compare it with the filter response, a numerical simulation is preferable because the TC processes the autoguider siderostat signal in a complicated fashion before using it. The TC updates the position of the siderostat, for sky rotation, every second by sending a velocity. During the second between updates, motors move the siderostat linearly such that its instantaneous position approximately corresponds to the target elevation and azimuth. The errors in tracking are then meant to be corrected via the autoguider siderostat signal. Although the siderostat signal is sent during each iteration of the autoguider loop (every 6 ms), the TC only reads the value every 1 s. The TC then calculates a mean every 3 s of the last 3 values

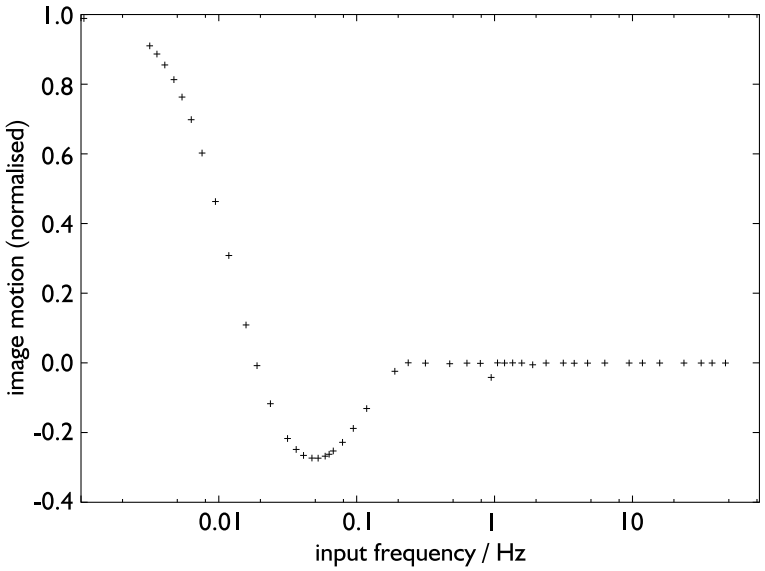

Figure 6.19.: Siderostat response (complementary sensitivity transfer function) with 3 s sampling and the running mean, as currently implemented at COAST.

i.e, a running mean. This mean is then used to compute a correction velocity that is constant for the next three seconds.

The simulation models image motions as sine waves with amplitudes smaller than the fast mirror range limits. The image motion is corrected by the siderostat and then the fast mirror; the response of the siderostat alone can be shown by plotting the complementary sensitivity transfer function of the TC. This is the ratio of the feedback signal (which drives the siderostat) to the command signal (which is the tip/tilt aberration including the tracking errors.) The transfer function is plotted in figure 6.19, and shows that the response is only above 50% for frequencies $\lesssim 0.01$ Hz and has a component with a phase lag around 0.05 Hz. In this latter region, the response is negative and the autoguider siderostat signal causes an image motion amplification and therefore increases the tip/tilt errors that the fast mirrors must compensate. The phase lag is increased because the phase response of the autoguider filter is not zero at these frequencies: at 0.25 Hz, the filter output amplitude is 0.68 and the phase is 0.83 rad (48°). The artifacts at 1 Hz and 2 Hz can be ignored, because these result from the TC sampling rate equalling the image motion frequency which in turn causes the mean siderostat position to be non-zero (compensated for by the fast mirror.)

A set of further simulations were ran with different TC sampling parameters to improve the siderostat response. These algorithms are,

- Not using a windowed mean and instead using the sampled signal at 1 Hz i.e., an effective increase of the sampling rate.
- As above, but with the autoguider filter time constant reduced to $\tau = 0.07$ s.
- As above, but with the sampling increased to 2 Hz and the gain halved.

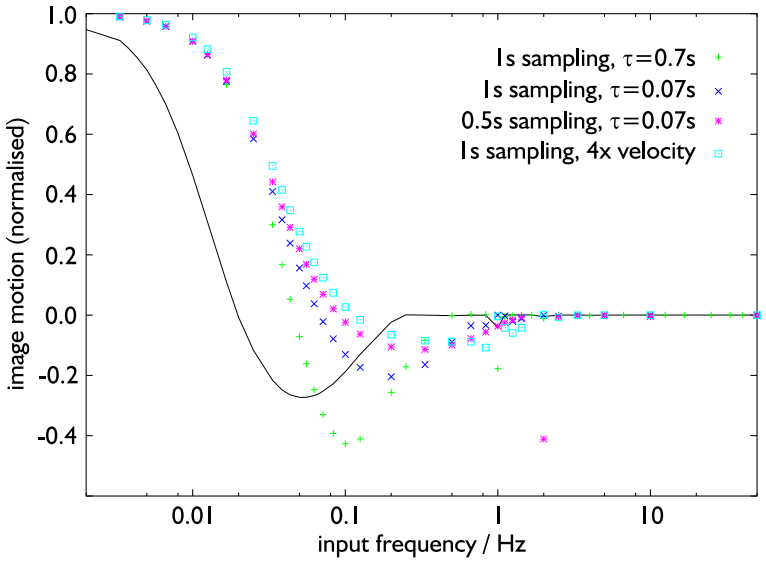

Figure 6.20.: Siderostat response, for the four algorithms described in the text (points) and the current algorithm at COAST (line).

- As in the second case (1 Hz sampling), but using a quadrupled velocity for the first quarter of the second and set to zero for the remaining three quarters i.e., a reduction in the siderostat response latency.

The results of the siderostat response are plotted in 6.20, with the line representing the existing response. The first algorithm (green points) removes the 3 s averaging and, although the response is better at low frequencies, the response around the phase lag becomes worse. Introducing a shorter time constant (dark blue points), to shift the filter phase response to higher frequencies, results in a reduction in the magnitude of the negative amplitude region. Increasing the sampling to 2 Hz (magenta[1] points) improves the response around the siderostat phase lag a small amount but is otherwise similar. Finally, increasing the siderostat velocity (cyan points) instead of the sampling rate, gives the best response, except around 1 Hz. At these frequencies the step behaviour of the siderostat, when the velocity becomes zero after 0.25 s of movement, causes an effect similar to aliasing of the autoguider siderostat signal.

These alternative algorithms were designed to be compatible with the existing COAST hardware. They show that simply removing the averaging is not sufficient and it must be combined with a modification of the autoguider siderostat signal filter. The further modifications, that of a sampling rate greater than 1 Hz or using a velocity increase for part of a second, did not improve the response as greatly, and in the case of the latter introduce further effects which would require further investigation before implementation.

Summary

The slow guiding diagnosis above can be summarised into several points. First, the assumption that the siderostat can be used to correct errors from tracking the sky rotation is false, and results in the

[1]For those reading in greyscale, I of course mean light black.

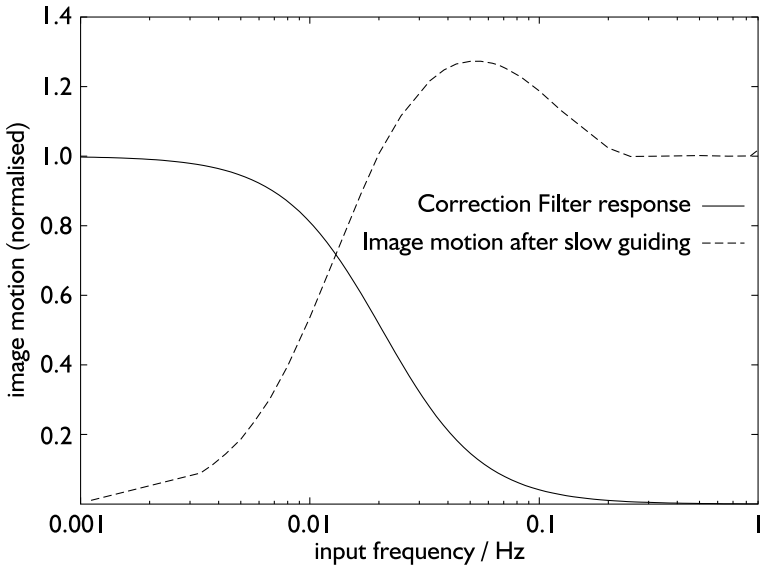

Figure 6.21.: Response of the Correction Filter, with the remnant image motion from slow guiding (sensitivity transfer function).

fast mirrors having to be used for this purpose. Second, the siderostat response is relatively poor and not well matched to the autoguider siderostat signal filter. Finally, the siderostat transfer function shows that using the current algorithm results in a significant amplification of image motion when the siderostat suffers from a phase lag. The overall consequence is that the fast mirrors are used to compensate for low frequency tip/tilt errors, and results in these mirrors reaching their range limits and saturating. The simulations with the alternative algorithms are proposals to overcome these deficiencies, via changes in the TC control algorithm. Removing the running mean and changing the autoguider filter time constant provides the largest immediate benefit to the siderostat response.

6.6. The Seeing at COAST

One method originally proposed for calculating r_0 was to use the distribution of the reconstructed fast mirror angles to compute the variance. As the recorded image motions also include tracking errors, the question of whether seeing estimates would be adversely affected must be answered. The first step to eliminate the errors is to only use the mirror angles which are within the range limits, to eliminate the "spikes" in the distributions (figure 6.9). An attempt was then made to remove, or at least reduce, the effect of the remaining (low frequency) tracking errors by high-pass filtering the fast mirror angles.

Beam	**Mean**/pseudo-volts		**Variance**/pseudo-volts^2		$\mathbf{r_0}$**(500 nm)/cm**	
	vert	**horiz**	**vert**	**horiz**	**vert**	**horiz**
Uncorrected						
1	-3.8 ± 0.4	8.9 ± 0.3	991 ± 17	812 ± 14	5.9 ± 0.3	6.7 ± 0.2
3	2.3 ± 0.3	19.6 ± 0.3	1196 ± 16	607 ± 10	6.8 ± 0.2	6.4 ± 0.2
Corrected						
1	-0.7 ± 0.2	1.6 ± 0.3	877 ± 12	695 ± 10	6.4 ± 0.2	7.3 ± 0.2
3	-0.2 ± 0.4	0.9 ± 0.2	939 ± 16	556 ± 8	7.8 ± 0.2	6.8 ± 0.2

Table 6.2.: Fits of normal distributions to fast mirror angle distributions from 240 s of fast guiding.

6.6.1. Data Correction Method

To correct the angles, they are first filtered with a (low-pass) IIR filter, equation (6.8), using an arbitrary time constant: $\tau = 0.07\,\mathrm{s}$. The corrected angles, $y_{corr;i}$, are then the original angles, y_i, with the filtered angles, y_i^*, subtracted,

$$y_{corr;i} = y_i - y_i^* \tag{6.12}$$

The image motion after correction by the slow guiding and the response of the filter is shown in figure 6.21. The filter only has a non-zero response for frequencies where the siderostat response is also non-zero and it removes all power at very low frequencies, but causes only a small reduction of the overall variance. As will be shown, this drawback is negligible compared to the improvement provided. The filter's main effect, therefore, is to reduce image motion at frequencies (around 0.05 Hz) where the siderostat response has a phase lag.

An example of a mirror angle distribution for one actuator channel, before and after correction, is shown in figure 6.22. The parameters for fits of a normal distribution, using a Levenberg-Marquardt least-squares minimisation (58), are shown in table 6.2. The errors on the distribution parameters were inferred from the assumption that the errors on each bin of the distribution were Poisson.

The fit remnants differ mainly in that the non-Gaussian features are less apparent after using the correction filter. The argument that the subtraction in equation (6.12) will always tend to form a more Gaussian distribution, because of the Central Limit Theorem, is incorrect since the two data sets are not statistically independent. In fact the correction method described is equivalent to using a high-pass filter. It must be emphasised that this correction method is "ad-hoc", and to validate the applicability its results were compared with two independent measures of spatial seeing.

Comparison with a DIMM

The variances from fits to the fast mirror angle distributions can be scaled to radians rms and the value of D/r_0 calculated from the theoretical one axis variance. The value of D depends on the aperture used for the autoguider QC; in this case the annuli from the spot mirror beamsplitters so the average of the inner and outer diameters was used, and D is assumed to be 32 cm on the sky.

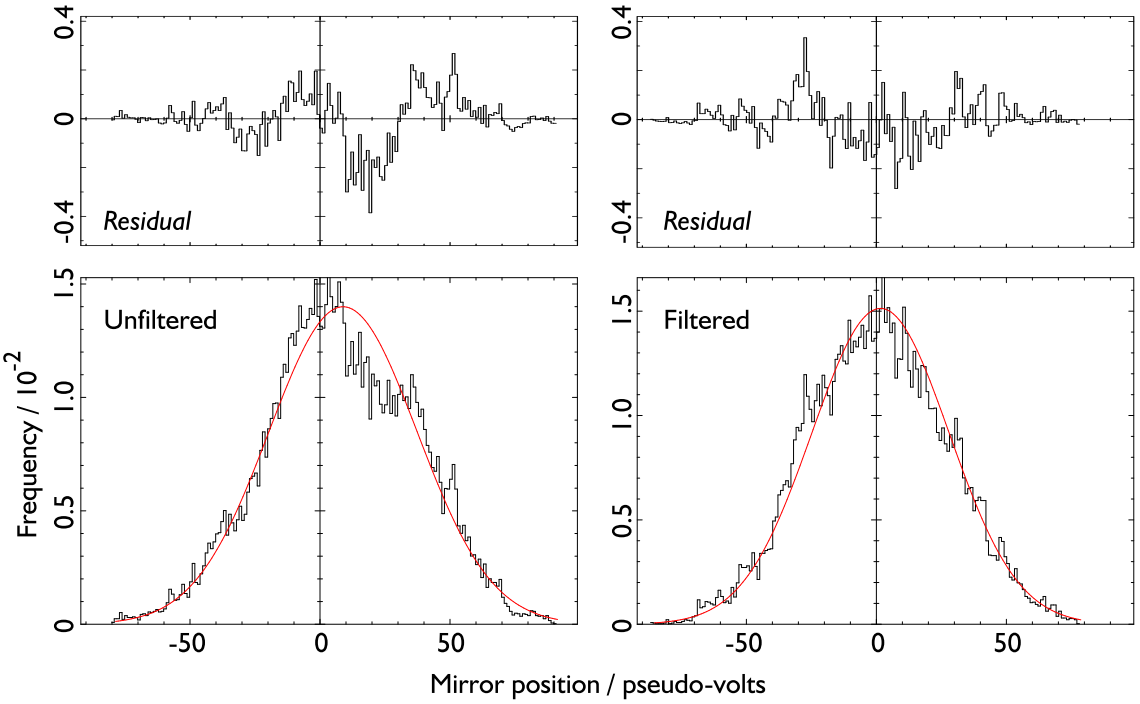

Figure 6.22.: Mirror angle distribution before and after using the correction filter described in the text.

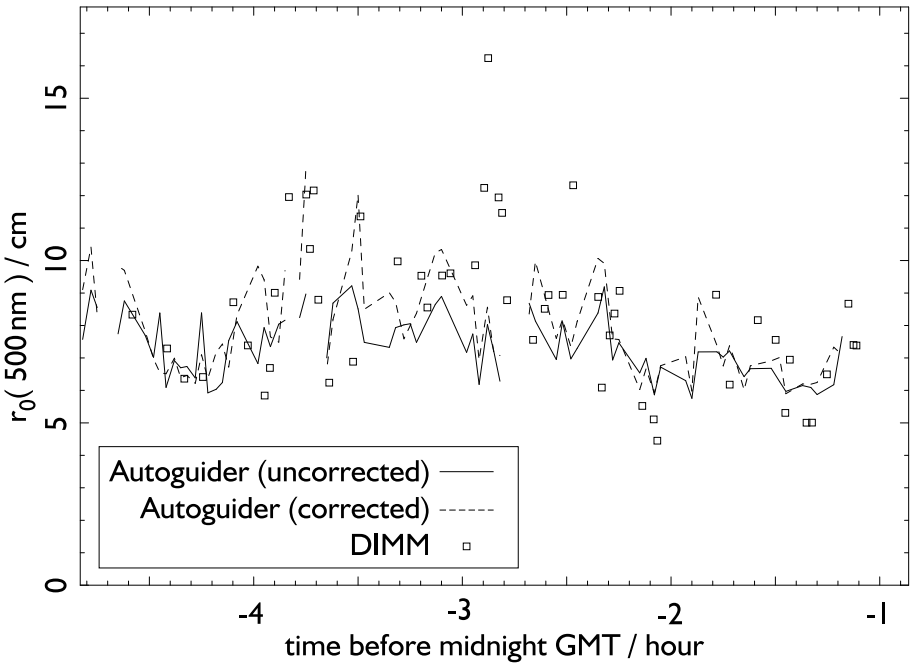

Figure 6.23.: Seeing measured on 01/03/2004 from the autoguider data (with and without the correction filter) and the DIMM.

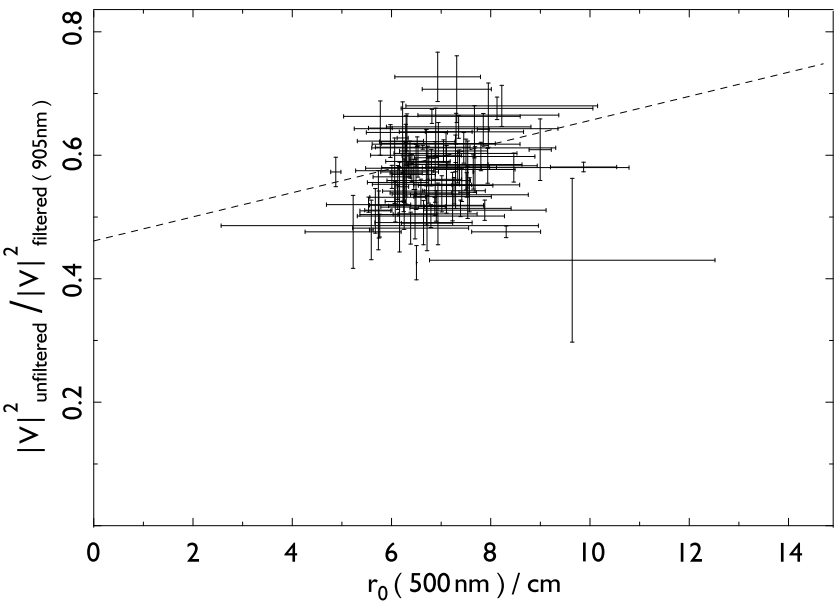

Figure 6.24.: Ratio of fringe visibilities plotted against r_0 (from autoguider data without correction).

A plot of $r_0(500\,\text{nm})$ is shown in figure 6.23 for one night (1st March 2004). The dashed and solid lines represents $\langle r_0 \rangle$—the average r_0 from both axes of beams 1 and 3. In comparison, the points are seeing measurements from a Differential Image Motion Monitor (DIMM) installed in a COAST telescope (54). The autoguider seeing measurements are consistent with those from the DIMM, confirming that the autoguider can be used to characterise the spatial seeing. It also shows that using the correction filter gives a better correspondence of seeing measurements with the DIMM.

Comparison with Fringe Visibilities

The source of phase aberrations in an ideal telescope is solely from the atmosphere, but in a realistic instrument the misalignments of internal components cannot be neglected. Compared to these two terms, the magnitude of internal seeing is small and can therefore be neglected. The total sum of aberrations, causing the loss in visibility coherence, is then a poor measure of the atmospheric seeing alone. A pinhole spatial filter is installed in COAST (33) at one output of the beam combiner, and was used to measure spatially filtered visibilities simultaneously with unfiltered visibilities. The filter removes phase aberrations from the seeing, and therefore provides a measure of the internal misalignments. The ratio of the visibility from unfiltered light, V_{uf}, to that from spatially filtered light, V_{sf}, is then expected to be dependent on the atmospheric seeing alone and can be compared with the estimates of r_0.

Figure 6.24 and 6.25 show the plots of $V_{uf}\,/V_{sf}$ against autoguider-derived r_0, with straight line fits. The data was taken over 10 nights and both the autoguider results and the visibility ratios were averaged in 300 s bins to remove the factor of data acquisition time. The linear fits resulted in a gradient of $2.0 \pm 0.7 \times 10^{-2}\text{cm}^{-1}$ for the unfiltered autoguider data and $4.4 \pm 0.3 \times 10^{-2}\text{cm}^{-1}$ for the filtered autoguider data. Although the relationship between the visibility ratio and r_0 is not

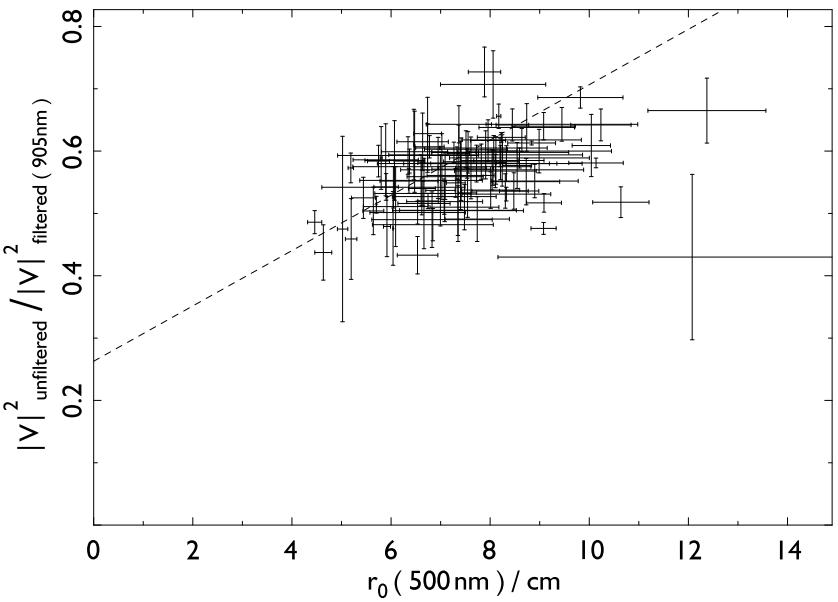

Figure 6.25.: Ratio of fringe visibilities plotted against r_0 (using the data correction method).

linear, the increased correlation (smaller error in the gradient) using the correction filter suggests again that it gives a more accurate measure of the seeing.

In principle, the visibility ratios can be converted to an equivalent r_0 and then directly compared to the autoguider results. However, this requires modelling the spatial filter and the associated effects of phase aberrations from sources other than the atmosphere, and attempts to create a simulation of the spatial filter did not result in a good match of visibility ratio-derived r_0 to autoguider-derived r_0. The simple description given in the introduction to this section neglects factors such as different outputs from the beam combiner having different static aberrations and different misalignments in the APDs measuring the fringes. In other words, the visibility ratio may include effects other than from the atmosphere. Finally, the simulation naïvely assumed the autoguider has no temporal lag so applied an immediate correction to the aperture phase aberrations. Given the difficulties in accurately simulating the required autoguider and beam combiner properties, no further time was spent refining the model.

Conclusion

Fitting a normal distribution to the reconstructed mirror angles, when the mirror operates within its range limits, gave a good agreement with the seeing estimates from the DIMM. The data correction method proposed, which can be summarised as applying a high-pass filter, results in comparable values of r_0 to not using the filter, but, when compared with the fringe visibilities, it provides a more accurate measure of the seeing. It will now be assumed that the correction filter can be used with the data to give the valid r_0 results, but it is acknowledged that the correction, as implemented, is only qualitative.

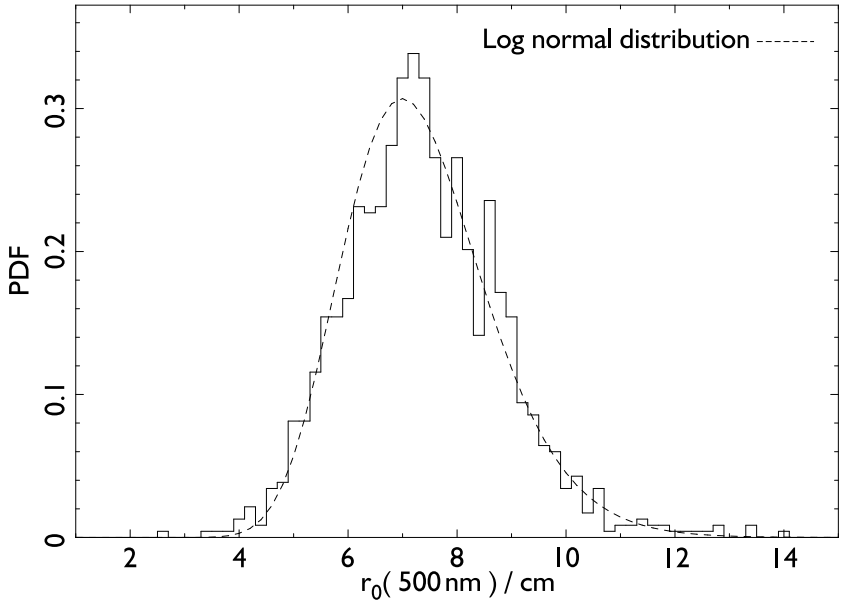

Figure 6.26.: Distribution of $\langle r_0 \rangle$ from 16/01/2003 to 17/12/2003.

6.6.2. Results of a Seeing Measurement Campaign

Autoguider data from January to December 2003, representing 74 nights, was analysed to derive the distribution of seeing at the COAST site. For a given autoguider data file, the mirror angles, after using the correction filter, are split into 120 s segments: these numbered 1167 in total. After fitting to a normal distribution of each segment, the value of $\langle r_0 \rangle$is calculated and the resulting distribution is shown in figure 6.26 with a log-normal distribution fit.

The mean value at 500 nm is $7.3^{+1.4}_{-1.2}$ cm or, equivalently, the mean of the logarithm is 1.43 ± 0.18. The median and the mode are 8.3 ± 0.1 cm and 7.2 cm. These values are consistent with sea-level measurements made using the SUSI interferometer autoguider-equivalent (80) system, which resulted in a mean value for r_0 of $7.6^{+1.5}_{-1.3}$ cm for the Narrabri site. If a normal distribution is fitted to the COAST $\langle r_0 \rangle$ data, then a smaller χ^2 value is obtained but, as r_0 is always a positive variable, this distribution would be unphysical and is therefore not considered here. Seeing measurement campaigns using DIMMs at Cerro Paranal (Chile), Roque de Los Muchachos Observatory (49) (Canary Islands), and Maidanak Observatory (18) (Uzbekistan), have found the distribution of r_0 to be consistently log-normal.

An important omission in the calculation of $\langle r_0 \rangle$ is ignoring the angle of the target from the zenith, θ. This is because the correlation of the position of the target with the autoguider data would have had to be inferred by hand from written observing logs, and as these do not contain sufficient data for direct automation, it would have been extremely time consuming. The inferred values of r_0 can be corrected by being divided by the zenith angle correction factor, $\cos^{-5/3}\theta$. The result of not doing this, is that the distribution presented is biased towards smaller values. The quoted mean r_0 value is therefore representative of the seeing during observations at "typical" declinations from the zenith. Finally, not using the correction filter resulted in only 786 samples from the data set that could have their angle distributions fitted—the remainder were rejected

as having too large a residual—and the resulting mean r_0 is $6.6^{+1.6}_{-1.3}$ cm (the median value was 6.7 ± 0.1 cm, and the mode was 6.2 cm.)

6.7. Conclusion

The aim of this chapter was to use the autoguider data to estimate the instantaneous seeing when regular observations were underway. The analysis of the full system, and in particular the telescope controller and siderostat, suggests problems with the design which limited this study. Improvements for controlling the telescope siderostats, given the current hardware, have been suggested, and the obvious step to optimise the AO loop is to remove the telescope controller running mean and simultaneously reduce the autoguider filter time constant. A more detailed study would be required, however, to produce a fully optimised algorithm. The result of the tracking errors introduced by siderostat response meant that the power spectra of the fast mirror angles could not be directly used to calculate the atmospheric image motions. This is compounded by the spectra of reconstructed mirror angles not corresponding to the theoretical spectrum at high frequencies. Therefore, to derive the seeing, the distributions of the mirror angles were utilised and comparison with measurements from a DIMM confirmed that r_0 can be estimated in this way. A heuristic correction method, involving a high-pass filter, has been implemented for the mirror angle distributions and it improves the correlation of seeing estimates when compared to results from the DIMM and to visibility ratios. Using the correction filter, the distribution of r_0 at the COAST site was found to have a mean of $7.3^{+1.4}_{-1.2}$ cm.

7. Conclusions and Future Work

7.1. Conclusions

The use of Adaptive Optics for Stellar Interferometry is an interesting juxtaposition of two technologies whose original aim is the same: to make high resolution astrophysical measurements that are not limited by seeing. This dissertation has examined this problem by starting with the causes of seeing, its resultant effects on the fringe visibility, and what AO can do to alleviate these problems for an interferometer.

7.1.1. Summary of chapters

As explained in chapter 2, the goal of AO for interferometry should be to reduce the visibility coherence loss. As AO does not allow an interferometer to make coherent integrations longer than a few t_0, the WFS used must use relatively little flux and have high SNR i.e., be particularly sensitive. The small aperture sizes and availability of spatial filtering for an interferometer imply that low order high-sensitivity AO is the optimum requirement. It was shown that an AO system which corrects the first 5 Zernike polynomial aberrations, $N = 5$, is a suitable definition of low order for moderate seeing conditions in the near-IR wavelength region. A review of available AO technologies concludes that the Control System and Deformable Mirrors are currently available or implementable using commercial components, but no suitable Wavefront Sensor exists.

Chapter 3 examines the unavailable AO component: a low order high-sensitivity Wavefront Sensor. A new wavefront sensing concept, Diffractive Phase Sensing (DPS), is introduced and shown to be more sensitive than the related Curvature Sensing method. Using the DPS method, the design and basic properties of the Nine Element Sensor are described. The NES is a WFS for the purpose of measuring small amplitude phase aberrations, rather than general astronomical AO. Its characteristics for wavefront sensing, using the DPS method, are different from more commonly used WFS. They are therefore examined (qualitatively) to explain why the design of the NES is optimum for this low order AO application. In particular, a centring algorithm is developed to solve the coupling of measured signals between tip/tilt aberrations and $m-$even abberations (such as defocus).

The numerical simulations of the NES in chapter 4 are intended to evaluate its wavefront sensing properties, such as sensitivity and the measurement non-linearities. A further motivation was to analyse the suitability of such a WFS for use in a COAST AO system. In comparison with two WFSs currently used for AO, the Quadrant Cell and the Curvature WFS, it is demonstrated that the NES is the best performing WFS when using small telescope apertures, $D/r_0 \leq 5$, with low fluxes, less than 1000 photons per iteration of the AO loop. The smallest aperture size considered

here, $D/r_0 = 3$, was chosen to corresponded to a COAST aperture. In two other cases, the other WFSs are better choices: with very low fluxes, less than 30 photons per iteration, the QC offers the best correction, and when the seeing is relatively poor, $D/r_0 \geq 7$, a high order WFS such as the CWFS is needed. It can be pointed out that an alternative interpretation is that the NES is better suited to wavefront sensing in the infra-red, where D/r_0 is larger than in optical wavelengths. A disadvantage of this region of the spectrum is that detectors have finite read noise, so the overall gain may be less than intially supposed. The concept of creating a hybrid WFS with an integrated pinhole spatial filter, to improve throughput and efficiency of both spatial filtering and wavefront sensing, was also investigated via simulations. When using a QC and NES as the WFS component of the hybrid sensor, the wavefront sensing measurement is less affected for the QC. In the case of no measurement errors, the QC is essentially unaffected by the central pinhole. Also, the pinhole at the focus, such as with a QC hybrid, has a significantly better throughput than the NES hybrid, which essentially uses a pinhole at a defocused plane.

As the NES was shown in the simulations to be suitable for an interferometer with COAST-sized apertures, a NES prototype was built and installed into COAST. The description of the NES prototype—its design in terms of the optics and the CCD detector—are detailed in chapter 5. The laboratory experiment used to introduce controllable defocus aberrations was evaluated and shown to introduce additional, unwanted. aberrations, namely tilt and astigmatism. Despite these errors, the experiment using the NES prototype, in its Curvature WFS and NES configurations, was shown to measure the expected defocus aberrations. The NES configuration was used with an assumed wavelength for operation with broadband light, and it was found that an amplitude calibration factor was required to convert the measured amplitudes to the actual aberration amplitude. The rapid readout mode of the CCD was tested in the above experiment and shown to allow defocus aberration measurements to be made at a rate of 333 Hz. Using this readout mode, the NES prototype was installed at COAST and used to make qualitative internal and external seeing measurements. Both of these produced unexpected results. The internal measurements suggested that there is a significant source of defocus aberrations located near the beam combiner table. Additionally, the assumption that the light entering the NES would be unpolarised was shown to be false, although there was no evidence this affected the internal seeing results. The measurements from the atmosphere showed severe effects of scintillation and, in particular, the shape of the experimental defocus aberration power spectra could not be explained. Under the assumption that the measured defocus aberration variance was not affected by scintillation, the value of r_0 from one night of data with the NES was ~ 9 cm at 500 nm.

The final chapter concerned the measurement of the seeing at COAST using the autoguider instrument. An analysis of the measured image motions suggested two problems with the current tip/tilt correction: first, the siderostat control algorithm is poorly matched to the autoguider filter and secondly, the fast mirror on the Inner West telescope has an unacceptably slow response. The former problem was analysed to elicit a data correction method which acted as a high-pass filter. Although "ad-hoc", it was shown to improve the accuracy seeing estimates when compared with contemporaneously measured fringe visibilities, and it was used to correct data from a year-long dataset of autoguider measurements. The resulting values of r_0 from this seeing campaign yielded a

fit to a log-normal distribution, as found in other seeing measurement campaigns. At a wavelength of 500 nm, the mean value of r_0 was 7.3 cm using the data correction method or 6.6 cm using uncorrected data.

7.2. Future work

The uncertainties in the results from using the NES prototype at COAST are the most significant area which require further investigation. The implication is that the interferometer environment within the laboratory bunker at COAST—which is open to sources of additional turbulence such as heat from computers and the observers themselves—creates a source of non-Kolmogorov phase aberrations. Future interferometers will require more stringently controlled environments and evaluating the internal seeing in COAST will aid their design. Internal measurements using the current NES prototype within COAST can be repeated with the Curvature Sensor configuration to verify these results. Furthermore, a non-linear algorithm, which did not make the assumptions of small aberration amplitudes, to determine phase aberrations from unequal intensity images could be developed and tested with the internal data. This would then have two benefits: the existing data can be re-reduced without the current uncertainties due to scintillation effects, and this algorithm would allow the neglection of cross-coupling of Zernike aberration amplitudes.

As computing processor speeds increase, the numerical time required to execute such an algorithm, which may be of a Phase Diversity type, decreases and a real-time implementation becomes more feasible. Another area of future investigation would be to utilise the linearity results found in this work to improve convergence speeds of iterative methods.

Another use of the NES prototype, in its Curvature Sensor configuration, is in telescope alignment. Given the complexity of the optical arrangement within COAST, moving a telescope within the array to form a different baseline is a time consuming task. In particular, refocusing the telescope is important, and more so for longer baselines, to preventing vignetting and subsequent loss in flux. Having a WFS which gave quantitative feedback as to the focus would save considerable time as the current method involves a theodolite which is a qualitative instrument. The current experiment is lacking in usability, as the data acquisition and analysis are separate command line software. A more usable instrument would be able to be operated over the internal COAST network and, for example, would allow a laptop to display real-time feedback on the adjustment of a telescope's optics.

The analysis of the COAST autoguider AO system concluded that the telescope controller does not adequately act on the error signals sent from the autoguider. As was shown, within the context of the current control loops, changing the siderostat control algorithm allows a better response to be obtained. However, even the best performance suggested by the simulations is not necessarily a good match to the autoguider siderostat signal filter. The analysis in this dissertation was brief, and to correctly re-implement the slow guiding component at COAST will require a more detailed understanding of how the telescope controller can be practically altered. An example is the assumption made that the siderostats are ideal with perfect temporal response, which clearly is not the case. An alternative approach is to alleviate the tracking errors by improving the telescope

pointing models. This would reduce or prevent these errors and make the slow guiding a less important component of tip/tilt correction.

Finally, a potential use for the autoguider seeing analysis is outlined. Since there is only one spatial filter currently at COAST, the seeing causes significant visibility coherence loss on 3 of 4 beam combiner outputs. The correlation between the autoguider-derived r_0 and the visibility ratios suggests that the effects of seeing on visibility coherence can be modelled and corrected for. In doing so, the static phase aberrations and misalignments of the beam combiner and APDs will remain. The errors from using calibrator target observations, which are not synchronous with science target observations, can be compensated for from the autoguider measurements. The systematic errors are then correctly removed by the calibration process. A less ambitious scheme is to use the r_0 measurements to quantitatively ignore visibility measurements that are likely to be mis-calibrated due to short timescale changes in the seeing.

A. Zernike Polynomials

A.1. Summary

The definitions of the basis of complete Zernike polynomials is given here. A graphical depiction of the 2nd to 6th polynomials is shown. An important symmetry property used in this work, that of being $m-$even or $m-$odd, is defined.

A.2. Definitions

The complete basis of Zernike polynomials (6) are an orthogonal set of modes defined in polar coordinates, (r, θ), within a circle. This circle is nominally the unit circle, $r \leq 1$, but a circle with diameter D can be used by scaling the radial coordinate, $r \to r/D$.

Each polynomial, $Z_{mn}(r, \theta)$, is defined by two indices: the angular index, m, and radial index, n. When $m \neq 0$ each polynomial has two forms, even and odd, and when $m = 0$ there is one form. Even Zernike polynomials, Z_{mn}^{even}, have a cosine term and odd polynomials, Z_{mn}^{odd}, have a sine term. There are two further constraints: $n - m$ is even and $m \leq n$.

The formal definition of $Z_{mn}(r, \theta)$, is therefore written as

$$Z_{mn} = \begin{cases} \sqrt{2n+2} R_n^m(r) \begin{cases} \cos(m\theta) & \text{even} \\ \sin(m\theta) & \text{odd} \end{cases} & m \neq 0 \\ \sqrt{n+1} R_n^0(r) & m = 0, \end{cases} \tag{A.1}$$

where

$$R_n^m(r) = \sum_{j=0}^{(n-m)/2} \frac{(-1)^j (n-j)! r^{n-2j}}{j! \left[(n+m)/2 - j\right]! \left[(n-m)/2 - j\right]!}.$$

The function R_n^m is normalised such that the integral of the multiplication of two Zernike polynomials is orthonormal,

$$\int_0^1 dr \int_0^{2\pi} d\theta \; Z_{mn}(r, \theta) Z_{m'n'}(r, \theta) = \begin{cases} \delta_{mm'} \delta_{nn'} & \text{if } Z_{mn} \;\&\; Z_{m'n'} \text{ are both even or odd} \\ 0 & \text{otherwise.} \end{cases} \tag{A.2}$$

Any arbitrary function, defined at all points in the unit circle, can be decomposed into the amplitudes of Z_i, a_i,

$$f(r, \theta) = \sum_{i=0}^{\infty} a_i Z_i(r, \theta),$$

Radial index, n	**Azimuthal index,** m 0	1	2	3
0	Z_1 piston			
1		Z_2, tip Z_3, tilt		
2	Z_4 defocus		Z_5, astig. 1 Z_6, astig. 2	
3		Z_7, coma 2 Z_8, coma 1		Z_9, trefoil x Z_{10}, trefoil y
4	Z_{11} spherical aberration			

Table A.1.: The first 11 Zernike polynomials with radial index, n, and azimuthal index, m. The Seidel-equivalent name is also shown.

and the amplitudes can be calculated from an integral,

$$a_i = \int_0^1 dr \int_0^{2\pi} d\theta \; Z_{mn}(r,\theta) \, f(r,\theta) \, .$$

The polynomials have zero-mean, $\int\int Z_i = 0$, and, due to being normalised, the decomposition of phase, $\Phi(2r/D, \theta)$, across a circular aperture of diameter D returns amplitudes whose units are rad rms,

$$\sigma_\Phi^2 = \sum_{i=0}^{\infty} a_i^2$$

A.2.1. Single Index Mapping

There also exists a one-to-one mapping of each polynomial to one index (53), i,

$$Z_i(r,\theta) \rightarrow \begin{cases} Z_{mn}^{even} & i \text{ is even} \\ Z_{mn}^{odd} & i \text{ is odd.} \end{cases}$$

Table A.1 shows the relationship of the first 11 Zernike polynomials to these indices, and common Seidel-equivalent names for the phase aberrations they represent.

The explicit expressions for the first 11 Zernike polynomials are,

$$
\begin{aligned}
Z_1 &= 1 \\
Z_2 &= 2r\,\cos(\theta) \\
Z_3 &= 2r\,\sin(\theta) \\
Z_4 &= \sqrt{3}(2r^2-1) \\
Z_5 &= \sqrt{6}r^2\,\sin(2\theta) \\
Z_6 &= \sqrt{6}r^2\,\cos(2\theta) \\
Z_7 &= \sqrt{8}(3r^3-2r)\,\sin(\theta) \\
Z_8 &= \sqrt{8}(3r^3-2r)\,\cos(\theta) \\
Z_9 &= \sqrt{8}r^3\,\sin(3\theta) \\
Z_{10} &= \sqrt{8}r^3\,\cos(3\theta) \\
Z_{11} &= \sqrt{5}(6r^4-6r^2+1)
\end{aligned}
$$

A.2.2. Symmetry: m-even & m-odd

An important property of Zernike polynomials, used in this work but not elsewhere, is their symmetry. As the terms even and odd are already defined, two new definitions are provided. A Zernike polynomial is described as $m-$even if it is symmetric when rotated by π. Conversely, is it $m-$odd if it is anti-symmetric under the same rotation.

$$
\begin{aligned}
Z_{m-\text{even}}(r,\theta) &= Z_{m-\text{even}}(-r,\theta+\pi) \\
Z_{m-\text{odd}}(r,\theta) &= -Z_{m-\text{odd}}(-r,\theta+\pi)\,.
\end{aligned}
$$

The naming for this property comes from the angular index m: when m is even/odd, the angular component of the polynomial is symmetric/anti-symmetric.

A.2.3. Graphical Summary

The first 5 Zernike polynomial after Z_1 are drawn in figure A.1[1]. These correspond to the lowest order phase aberrations, order $N = 5$, and are the first two $m-$odd and first three $m-$even modes.

[1] From `http://wyant.optics.arizona.edu/zernikes/zernikes.htm`, 24/March/2004

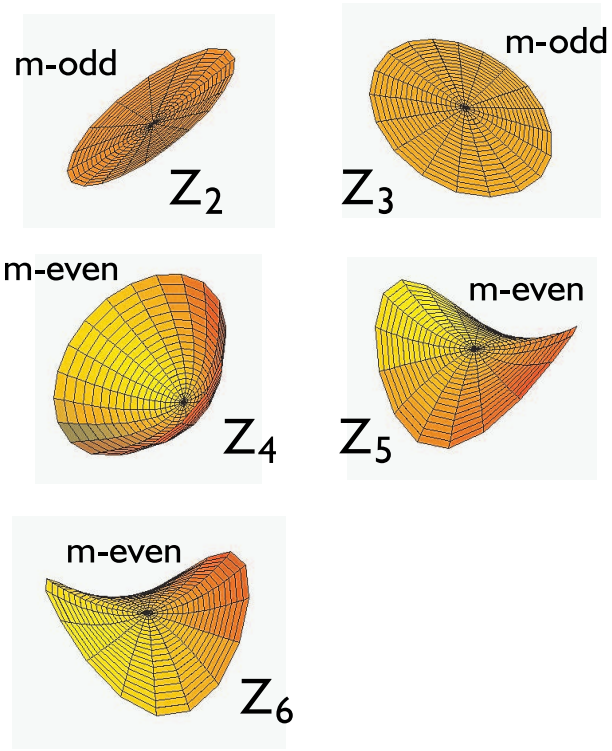

Figure A.1.: Graphical representation of 5 Zernike polynomials, Z_2 to Z_6, with the symmetry property also labelled. Note that each graphic has been rotated and their axes are not necessarily parallel.

B. Linear Measurements of Aberration Amplitudes

B.1. Summary

In this appendix, the proof is given of the linearity, with respect to the amplitude of phase aberration amplitudes, of the differences of intensities measured in two defocused planes. The assumptions made are that the planes have equal and opposite even diversity phase, applied to their respective pupils, and that the amplitude of all phase terms are small ($\ll 1$ rad rms).

The intensity from the two planes can be subtracted in two ways to obtain the amplitudes of $m-$even and $m-$odd Zernike polynomials. The inter-plane difference, ΔI, is sensitive, to 1^{st} order, to the even aberrations and the gain is proportional to the diversity amplitude. The intra-plane difference, $\delta\Sigma I$, is sensitive, to 1^{st} order, to the odd aberrations and has fixed gain.

B.2. Background to Method

Let the complex amplitude at the input aperture be written as,

$$A(x,y) = P(x,y)\exp\left[i\Phi(x,y)\right], \tag{B.1}$$

where P is the pupil transmission function and Φ is the phase aberration. The diversity will be written as $\pm\Phi_D\,(x,y)$, with positive sign for the intra-focal plane image and negative sign for the extra-focal plane image. It is noted here that the terms "intra-focal" and "extra-focal" need not specifically refer to defocused planes i.e., the diversity can be any even function of phase, and this includes defocus.

The complex amplitude is propagated to the two defocused planes using Fraunhofer diffraction, and the modulus squared of the amplitude at these planes is the intensity. The intensity at the intra-focal plane is denoted by I_+ and at the extra-focal plane by I_-. The Fraunhofer diffraction integral may be written as a scaled Fourier transform, $\mathcal{F}^s\{\}$, so the intensities are

$$I_{+/-}\,(u,v) = C\left|\,\mathcal{F}^s\left\{P(x,y)\exp\left[\pm i\,\Phi_D(x,y) + i\,\Phi(x,y)\right]\right\}\right|^2, \tag{B.2}$$

where C is an arbitrary constant to scale for the flux received at the aperture. For a complex exponential $\exp\left[iX\right]$ $(X \ll 1)$ then a Taylor series expansion to 2^{nd} order in X can be used as an approximation,

$$\exp\left[iX\right] = 1 + iX + \frac{X^2}{2!} + \ldots \tag{B.3}$$

Using this result, and writing the phase aberrations as a sum of even and odd terms, $\Phi = \Phi_E + \Phi_O$, the complex exponential in (B.2) can be written as

$$\begin{aligned} I_{+/-} &= C\left|\mathcal{F}^s\left\{P\left[1 + i\left(\pm\Phi_D + \Phi_O + \Phi_E\right) + \frac{\left(\pm\Phi_D + \Phi_O + \Phi_E\right)^2}{2} + \ldots\right]\right\}\right|^2 \\ &\simeq C\left|\left[\mathcal{F}^s\left\{P\right\} + i\mathcal{F}^s\left\{P\Phi_O\right\} + \frac{\mathcal{F}^s\left\{P\left(\Phi_D^2 + \Phi_O^2 + \Phi_E^2\right)\right\}}{2} \pm \mathcal{F}^s\left\{P\Phi_D\Phi_E\right\}\right]\right. \\ &\quad \left. + \left[-i\mathcal{F}^s\left\{P\left(\pm\Phi_D + \Phi_E\right)\right\} \pm \mathcal{F}^s\left\{P\Phi_D\Phi_O\right\} + \mathcal{F}^s\left\{P\Phi_E\Phi_O\right\}\right]\right|^2. \end{aligned}$$

The second line uses the linearity property of Fourier transforms, and has been separated into two parts delineated by the square brackets. The terms in the first bracket are real i.e., the transforms are either real, or imaginary and multiplied by i. The second group of terms are imaginary i.e., the transforms are imaginary, or real and multiplied by i. Therefore the modulus squared is equivalent to the sum of the brackets squared,

$$\begin{aligned} I_{+/-} &\simeq C\left[\mathcal{F}^s\left\{P\right\} + i\mathcal{F}^s\left\{P\Phi_O\right\} + \frac{\mathcal{F}^s\left\{P\left(\Phi_D^2 + \Phi_O^2 + \Phi_E^2\right)\right\}}{2} \pm \mathcal{F}^s\left\{P\Phi_D\Phi_E\right\}\right]^2 \\ &\quad - \left[-i\mathcal{F}^s\left\{P\Phi_D\right\} \mp i\mathcal{F}^s\left\{P\Phi_E\left(1 - i\Phi_O\right)\right\} + \mathcal{F}^s\left\{P\Phi_D\Phi_O\right\}\right]^2. \end{aligned} \tag{B.4}$$

B.3. Inter-Plane Differences

To recover the amplitudes of the even phase polynomials, first expand the brackets from (B.4). The intensity in the two planes are then subtracted at equal points, $\Delta I_\delta(u,v) = I_+(u,v) - I_-(u,v)$. The remaining terms to 3$^{\text{rd}}$ order in phase, are

$$\begin{aligned} \Delta I_\delta &\simeq 2C\left(\mathcal{F}^s\left\{P\right\} + \mathcal{F}^s\left\{P\Phi_D\Phi_O\right\}\right)\mathcal{F}^s\left\{P\Phi_D\Phi_E\right\} \\ &\quad -2C\left(\mathcal{F}^s\left\{P\Phi_D\right\} + i\mathcal{F}^s\left\{P\Phi_D\Phi_O\right\}\right)\mathcal{F}^s\left\{P\Phi_E\left(1 - i\Phi_O\right)\right\}. \end{aligned} \tag{B.5}$$

Retaining only the lowest order terms, to emphasise the result for small amplitudes, the even phase is decomposed into $m-$even Zernike polynomials, $Z_{j;e}$, the odd phase is decomposed into $m-$odd Zernike polynomials, $Z_{j;o}$ (the definitions of $m-$even and $m-$odd are in appendix A).

$$\Delta I_\delta\left(u,v\right) \simeq 2C\sum_{n=1}^{\infty} a_n\left(\mathcal{F}^s\left\{P\right\}\mathcal{F}^s\left\{P\Phi_D\, Z_{n;e}\right\} - \mathcal{F}^s\left\{P\Phi_D\right\}\mathcal{F}^s\left\{P\, Z_{n;e}\right\}\right)(u,v). \tag{B.6}$$

This subtracted intensity is called the Inter-Plane Difference, ΔI_δ, because it is the subtraction between each plane. It is the signal measured by a WFS and is proportional to only the $m-$even Zernike polynomial amplitudes, being zero otherwise. Therefore, to 1$^{\text{st}}$ order in phase, the continuous inter-plane difference signals from different $m-$even Zernike polynomials are orthogonal. The gain of the signal is also proportional to the amplitude of the diversity, Φ_D.

B.4. Intra-Plane Differences

The corresponding signal to measure $m-$odd Zernike polynomial amplitudes, the intra-plane difference, is developed in a similar way. The terms in (B.4) which are linear with respect to the odd phase were removed by subtracting the intensity between planes. Instead, summing the intensity at equal points between the planes, $\Sigma I\left(u,v\right)=I_{+}\left(u,v\right)+I_{-}\left(u,v\right)$, gives the following expression,

$$\begin{aligned}\Sigma I &\simeq C\left(\mathcal{F}^{s}\left\{P\right\}+i\mathcal{F}^{s}\left\{P\Phi_{O}\right\}+\frac{\mathcal{F}^{s}\left\{P\left(\Phi_{D}^{2}+\Phi_{O}^{2}+\Phi_{E}^{2}\right)\right\}}{2}\right)^{2}+\left(\mathcal{F}^{s}\left\{P\Phi_{D}\Phi_{E}\right\}\right)^{2}\\&\quad-\left(-i\mathcal{F}^{s}\left\{P\Phi_{D}\right\}+\mathcal{F}^{s}\left\{P\Phi_{D}\Phi_{O}\right\}\right)^{2}+\left(-\mathcal{F}^{s}\left\{P\Phi_{E}\left(1-i\Phi_{O}\right)\right\}\right)^{2}.\end{aligned}\tag{B.7}$$

The terms which are transforms of even functions, which are either terms with the even phase or the odd phase squared, are real and even. Conversely, the terms which are transforms of odd functions are imaginary and odd. Rewriting (B.7) to separate these symmetric and anti-symmetric terms with brackets, and retaining terms to 3^{rd} order in phase,

$$\begin{aligned}\Sigma I &\simeq C\left[\mathcal{F}^{s}\left\{P\right\}^{2}+\mathcal{F}^{s}\left\{P\Phi_{D}\right\}^{2}+\mathcal{F}^{s}\left\{P\Phi_{E}\right\}^{2}+\mathcal{F}^{s}\left\{P\right\}\mathcal{F}^{s}\left\{\Phi_{D}^{2}+\Phi_{O}^{2}+\Phi_{E}^{2}\right\}\right]+\\&\quad Ci\left[2\mathcal{F}^{s}\left\{P\Phi_{O}\right\}\left(\mathcal{F}^{s}\left\{P\right\}+\mathcal{F}^{s}\left\{\Phi_{D}^{2}+\Phi_{O}^{2}+\Phi_{E}^{2}\right\}^{2}/2\right)+i2\mathcal{F}^{s}\left\{P\Phi_{O}\right\}^{2}\right.\\&\quad\left.-2\mathcal{F}^{s}\left\{P\Phi_{D}\right\}\mathcal{F}^{s}\left\{P\Phi_{D}\Phi_{O}\right\}+2\mathcal{F}^{s}\left\{P\Phi_{E}\right\}\mathcal{F}^{s}\left\{P\Phi_{O}\right\}\right].\end{aligned}$$

The terms which are linear with respect to the odd phase are within the second, anti-symmetric group. The symmetric terms can be removed by subtracting the summed intensity at a given point from the summed intensity at the point rotated by π,

$$\Delta I_{\Sigma}\left(u,v\right)=\Sigma I\left(u,v\right)-\Sigma I\left(-u,-v\right),$$

and the intensity difference is

$$\begin{aligned}\Delta I_{\Sigma}=4Ci&\left[\mathcal{F}^{s}\left\{P\Phi_{O}\right\}\left(\mathcal{F}^{s}\left\{P\right\}+\mathcal{F}^{s}\left\{\Phi_{D}^{2}+\Phi_{O}^{2}+\Phi_{E}^{2}\right\}^{2}/2\right)\right.\\&\left.+i\mathcal{F}^{s}\left\{P\Phi_{O}\right\}^{2}-\mathcal{F}^{s}\left\{P\Phi_{D}\right\}\mathcal{F}^{s}\left\{P\Phi_{D}\Phi_{O}\right\}+\mathcal{F}^{s}\left\{P\Phi_{E}\right\}\mathcal{F}^{s}\left\{P\Phi_{O}\right\}.\right].\end{aligned}\tag{B.8}$$

This subtraction is called the Intra-Plane Difference, ΔI_{Σ}, because the intensities are subtracted between points on a (summed) plane. Retaining the lowest order phase terms and decomposing these into Zernike polynomials gives

$$\Delta I_{\Sigma}\left(u,v\right)=4Ci\sum_{m=1}^{\infty}a_{m}\,\mathcal{F}^{s}\left\{P\right\}\mathcal{F}^{s}\left\{P\,Z_{m;o}\right\}\left(u,v\right).\tag{B.9}$$

To 1^{st} order, this difference signal is proportional to only the $m-$odd Zernike polynomial amplitudes. Unlike the subtraction between planes, this result does not contain the diversity term so the signal gain is constant.

C. Methods of Defocused Plane Imaging

C.1. Summary

The method used to propagate the complex amplitude from the pupil plane to a defocused plane, at an arbitrary distance before the focus, is described here. It is initially developed using continuous variables, and then extended to the case of discrete variables. The limitations of the discrete case are carefully considered, and limiting conditions in terms of defocus length are given. An example of the on-axis intensity of an image as the defocus increases is presented.

C.2. Continuous Formulation for Defocused Plane Imaging

The propagation of the monochromatic complex amplitude from the pupil plane to an arbitrary image plane can be described by a Fresnel Diffraction Integral (22). Let the pupil transmission function be P, the complex amplitude A, the propagation distance l, and the imaging optics (a lens) have a focal length f. The integral is then

$$A(u,v;l) = \exp\left[i\frac{k}{2l}(u^2+v^2)\right]\frac{i}{\lambda l}\int dx\,dy\; A(x,y)\; P(x,y) \\ \times \exp\left[-i\frac{k}{l}(xu+yv)\right]\exp\left[i\frac{k}{2l}(x^2+y^2)\right]\exp\left[-i\frac{k}{2f}(x^2+y^2)\right]. \tag{C.1}$$

The coordinates and distances are shown diagrammatically in figure C.1, and the two assumptions made are that a paraxial lens is used and l is not small compared to the aperture dimension. When $f = l$, the integral reduces to the Fraunhofer Diffraction Integral. The intensity at the image plane does not depend on the quadratic phase outside the integral and this term will be further

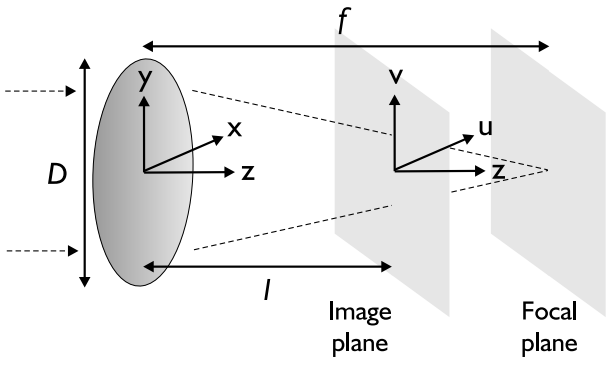

Figure C.1.: Propagation from the pupil to an image plane which is not the focal plane.

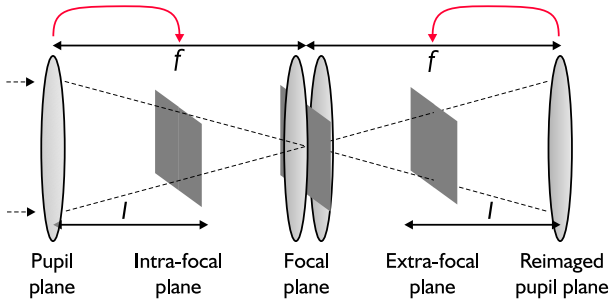

Figure C.2.: The intra-focal and extra-focal planes. Their propagation from the relevant pupil is denoted by the red arrows.

neglected—formally, its removal is equivalent to placing a paraxial lens of focal length l just before the image plane. The integral can be rewritten as a Fourier transform, $\mathcal{F}\{\}$,

$$\begin{aligned} A\left(u',v';l\right) &= \frac{i}{\lambda l}\int dx\,dy\; A\left(x,y\right)P\left(x,y\right)\exp\left[i2\pi(xu'+yv')\right]\exp\left[i\Phi_D\left(x,y\right)\right] \\ &\quad \text{where } \Phi_D\left(x,y\right) = -\frac{k\left(f-l\right)}{2lf}\left(x^2+y^2\right) \\ &= \frac{i}{\lambda l}\,\mathcal{F}\left\{A\left(x,y\right)P\left(x,y\right)\exp\left[i\Phi_D\left(x,y\right)\right]\right\}\left(u',v'\right). \end{aligned} \tag{C.2}$$

The coordinates (u',v') are scaled according to wavelength and distance from the pupil: $(u',v') = (u/\lambda l, v/\lambda l)$. It is well known that the amplitude at the focus is the Fourier transform of the complex pupil, AP. The amplitude at a plane *before* the focus—an intra-focal plane—is also a Fourier transform of a complex pupil, but the phase of AP has added to it a diversity, Φ_D. In the context of Fraunhofer diffraction, the focal plane corresponds to the plane at infinity and the intra-focal plane to a plane at some finite distance.

There is no meaning to the intensity on a plane after infinity, which is the naïve interpretation of an extra-focal plane. Instead, the amplitude after the focus must be considered separately from the amplitude before the focus. A Fourier transform of the amplitude at the focus returns the amplitude at the pupil (the re-imaged pupil) but rotated by π. In terms of light propagation, this second transform is equivalent to placing a paraxial lens after the focal plane and propagating the light a further distance f. The extra-focal plane can now be thought of as being between this re-imaged pupil and the focus.

Therefore, to calculate the amplitude at the extra-focal plane, the amplitude from the re-imaged pupil is propagated *back* towards the focus. This is shown diagrammatically in figure C.2, and the reverse propagation can be written analogously to (C.1),

$$\begin{aligned} A\left(u,v;-l\right) = \exp\left[i\frac{k}{2l}(u^2+v^2)\right]\frac{i}{\lambda l}\int dx\,dy\; A\left(-x,-y\right)\; P\left(-x,-y\right) \\ \times\exp\left[i\frac{k}{l}(xu+yv)\right]\exp\left[-i\frac{k}{2l}(x^2+y^2)\right]\exp\left[i\frac{k}{2f}(x^2+y^2)\right]. \end{aligned} \tag{C.3}$$

Note that the change in this expression is in the sign of the propagation kernel and the rotation of the pupil coordinates. If the signs of the latter are reversed—de-rotating the re-imaged pupil—then the expression becomes,

$$A\left(u,v;-l\right) = \exp\left[-i\frac{k}{2l}(u^2+v^2)\right]\frac{i}{\lambda l}\int dx\,dy\; A\left(x,y\right)\; P\left(x,y\right) \\ \times \exp\left[-i\frac{k}{l}(xu+yv)\right]\exp\left[-i\frac{k}{2l}(x^2+y^2)\right]\exp\left[i\frac{k}{2f}(x^2+y^2)\right]. \tag{C.4}$$

This is identical to (C.1) except the quadratic phase terms have reversed sign. Letting the amplitude on the intra/extra-focal planes be denoted by A_+/A_-, the propagation to either defocused plane can be written as one Fourier transform with the difference being a sign change of the diversity,

$$A_\pm\left(u',v';l\right) = \frac{i}{\lambda l}\,\mathcal{F}\left\{A\left(x,y\right)P\left(x,y\right)\exp\left[\pm i\Phi_D\left(x,y\right)\right]\right\}\left(u',v'\right). \tag{C.5}$$

This demonstrates that the diversities have the same amplitude and opposite signs when describing intra-focal and extra-focal planes which are equal distances from the focus.

C.3. Discrete Formulation for Defocused Plane Imaging

The Discrete Fourier Transform (DFT) is used in the numerical simulations for modelling propagation. The DFT of a $N \times N$ array of discrete values (henceforth a numerical array) is defined as,

$$\begin{aligned} A_{\mu\nu} &= \sum_{\alpha=0}^{N-1}\sum_{\beta=0}^{N-1} A_{\alpha\beta}\exp\left[i\frac{2\pi}{N}\alpha\mu+\beta\nu\right] \\ &= \mathcal{D}\left\{A_{\alpha\beta}\right\}_{\mu\nu}. \end{aligned} \tag{C.6}$$

The values in the numerical array can specified by indices, α,β and μ,ν for the pupil and imaged planes, and are defined as zero at the array centre. The scaling between the continuous variables and the indices is

$$\begin{aligned} (\alpha,\beta) &\rightarrow \left(\frac{x}{\Delta_p},\frac{y}{\Delta_p}\right) \\ (\mu,\nu) &\rightarrow \left(\frac{u}{\Delta_i},\frac{v}{\Delta_i}\right). \end{aligned} \tag{C.7}$$

Comparing the Fourier kernel in (C.6) and (C.1), the scaling factors can be written as,

$$\begin{aligned} \frac{2\pi}{N\Delta_p\Delta_i} &= \frac{k}{l} \\ \Delta_i &= \frac{\lambda l}{N\Delta_p} \end{aligned} \tag{C.8}$$

The pupil scaling factor, Δ_p, is the ratio of discretely sampled pupil diameter $2R$ to the physical aperture diameter D. It is therefore a free parameter along with the other physical variables, k and l. These define the image scaling Δ_i, which is linearly proportional to the propagation distance and so the physical distance between adjacent array elements increases with increasing defocus. This scaling is equal in the intra-focal and extra-focal planes, which makes the manipulation of intensities between these planes straightforward. The discrete equivalent of the propagation, equation (C.5), is,

$$\begin{aligned} A_{\pm;\mu\nu} &= \frac{i}{N\Delta_p\Delta_i}\,\mathcal{D}\left\{A_{\alpha\beta}P_{\alpha\beta}\exp\left[\pm i\Phi_{D;\alpha\beta}\right]\right\}, \\ &\quad \text{where } \Phi_{D;\alpha\beta} = -\frac{\Delta_p^2 k\,(f-l)}{2lf}\left(\alpha^2+\beta^2\right). \end{aligned} \tag{C.9}$$

C.3.1. Propagation from the Pupil Plane

The use of discrete transforms requires that the Nyquist sampling criterion be considered for the diversity phase, assuming that the complex amplitude $A_{\alpha\beta}$ already satisfies this requirement. The criterion must be evaluated at the pupil boundary,

$$\left(x^2+y^2\right)^2 \leq \frac{D^2}{4} \rightarrow \left(\alpha^2+\beta^2\right) \leq R^2, \tag{C.10}$$

where R $(D\Delta_p/2)$ is the aperture radius in array elements. The Nyquist criterion is

$$\Phi_{D;R\,0} - \Phi_{D;(R-1)0} \leq \pi, \tag{C.11}$$

and can be expanded to

$$\frac{k\,(f-l)}{2lf} \leq \frac{\pi}{(2R-1)\Delta_p^2}. \tag{C.12}$$

It is hard to immediately interpret the meaning of this expression, and two additional variables are introduced. The defocus distance is the distance *from* the focal plane, $d = f - l$, and the Fresnel number F is

$$F = \frac{D^2}{4f\lambda} = \frac{R^2\Delta_p^2}{f\lambda}. \tag{C.13}$$

The inequality in (C.12) can now be written as

$$d \lesssim \frac{f^2\lambda}{R\Delta_p^2} = \frac{fR}{F}, \tag{C.14}$$

where it is assumed that $d \ll f$ or alternatively that $F/R \gg 1$ ($R \simeq 50$ and $F \gtrsim 1000$ for realistic optics at visible wavelengths.) This is the largest physical distance, from the focus, for which a discrete numerical simulation can accurately propagate light from a circular pupil.

C.3.2. Propagation from the Focal Plane

The distance limitation of light propagation from the pupil matters when the Curvature WFS is modelled. This wavefront sensing method uses propagation to the intra/extra-focal planes whose defocus lengths are sufficiently large that diffraction effects can be neglected. The limit on the defocus length (64) in this case can be written as

$$d \gg \frac{f^2\lambda}{R^2\Delta_p^2}. \tag{C.15}$$

This limit is generally incompatible with the inequality from propagating light directly from the pupil, equation (C.14). The plane being propagated from, so far, has implicitly been assumed to be the pupil plane but the only explicit use of this fact was in deriving the Nyquist criterion limit above.

The case of propagating the light first to the focus, and then from there using the same DFT technique to the defocused plane will now be considered. The formalism developed in the previous section can be immediately applied with a suitable change of variables. The Nyquist inequality is now considered on the defocus diversity added to the focal plane and, since there is no explicit boundary, the criterion is evaluated at the corners of the numerical array,

$$\Phi_{D;(N/\sqrt{2})(N/\sqrt{2})} - \Phi_{D;(N/\sqrt{2}-1)(N/\sqrt{2}-1)} \leq \pi. \tag{C.16}$$

The inequality can be written in terms of the physical variables, as for equation (C.12), by substituting Δ_i for Δ_p, $N/\sqrt{2}$ for R, and d for l,

$$\frac{k\,(f-d)}{2df} \leq \frac{\pi}{(\sqrt{2}N-1)\Delta_i^2}. \tag{C.17}$$

Rewriting the inequality terms of Δ_p and using l, the distance from the pupil plane to the defocused image plane,

$$l \lesssim \frac{f}{(2\sqrt{2}N-1)f\lambda)/(N^2\Delta_p^2)+1}. \tag{C.18}$$

The second term in the denominator is large compared to the first and allows the use of a Taylor expansion to derive

$$l \lesssim f - \frac{\sqrt{8}\lambda f^2}{N\Delta_p^2}. \tag{C.19}$$

if it assumed that $N \gg 1$.

C.3.3. Notes on Derived Inequalities

The inequalities (C.14) and (C.19) depend on both the discrete variables and the optics being modelled, via λ, d, f and F. The size of the discrete array, N, only appears in the second inequality but it is implicit in defining Δ_p once R is chosen. Therefore, the variables which may be chosen freely are N and R.

To increase the limit on d, and maintain the same physical variables, R must be increased and this decreases Δ_p. This approach also increases Δ_i and implies poorer sampling of the image plane.

If the limit on d is still restrictive, as for modelling Curvature Sensing, then propagation via the focus is needed. To image at any plane, the overlap between the two methods can be evaluated. Let the inequality for imaging via the focus be written using the defocus length d. Comparing the inequalities,

$$\begin{aligned} &\text{propagation from focus:} \quad f \geq d \gtrsim \frac{\sqrt{8} f^2 \lambda}{N \Delta_p^2} \\ &\text{propagation from pupil:} \quad 0 \leq d \lesssim \frac{f^2 \lambda}{2 R \Delta_p^2}. \end{aligned} \tag{C.20}$$

To have an overlap, the requirement is that

$$R \leq \frac{N}{\sqrt{32}}, \tag{C.21}$$

so the radius of the aperture must be ~ 6 array elements times smaller than the size of the numerical array. The overlap does not imply, however, that obtaining an image at a given plane by using the two methods in turn will result in each image having the same scaling factor. Writing the scaling on the imaged plane as Δ_f and the distance from the pupil plane as l, the scaling factors are

$$\begin{aligned} &\text{propagation from focus:} \quad \Delta_f = \frac{\Delta_p \, (f - l)}{f} \\ &\text{propagation from pupil:} \quad \Delta_f = \frac{\lambda l}{N \Delta_p}. \end{aligned} \tag{C.22}$$

The first expression is a geometric scaling of the image, as for geometrical optics. The second involves the wavelength and array size, as for Fourier optics.

C.4. Example of Discretely Imaging a Defocused Image

As an example, the intensity of a defocused image along the optical axis will be studied. For increasing defocus, as the diversity amplitude becomes periodically equal to an even number of Fresnel zones, the on-axis intensity goes to zero (6). This can be described by writing the propagation explicitly

$$I(0,0) = 0 = C \left| \int_0^{D/2} dr \; r \exp \left[i \frac{k(f-l)}{2fl} \, r^2 \right] \right|^2, \tag{C.23}$$

where C is a scaling constant. The defocus distance for which this occurs is

$$d = f \frac{2n}{F + 2n} \quad \text{where } n \in \mathbb{I}^+. \tag{C.24}$$

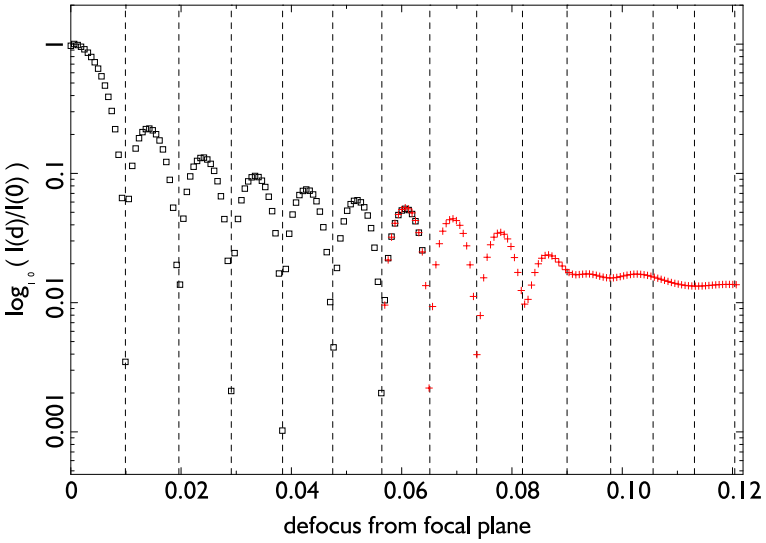

Figure C.3.: On-axis intensity for a defocused image, normalised to the zero defocus value. The vertical lines represent where the intensity is expected to equal zero.

The discretely sampled intensity for increasing defocus was calculated using values of R and N which resulted in an overlap. Figure C.3 shows the resulting on-axis intensity, with the intensities propagated from the pupil shown with squares and those via the focus shown with pluses. The defocus distances marked by vertical lines are those for which the intensity is expected to be zero. The predicted intensity nulling can be seen until $d/f \simeq 0.08$. The oscillation is then seen to "die away", but only when the propagation method is via the focus.

This lack of zero intensity results from the discrete sampling of the pupil. At the focal plane, the sampling leads to inaccuracies in the Airy rings surrounding the central Airy disc, especially as the distance from the centre of the image increases. As the defocus distance increases, the defocus diversity added to the focal plane amplitude becomes smaller. The complex amplitude from high order rings then become more important for propagation to suppress the on-axis intensity, but the nulling fails when high order rings are needed to be accurately sampled. The light propagated from the focus is therefore better described by geometrical optics, but this limitation is unimportant as this propagation method is only used for modelling the Curvature WFSs.

D. Readout Rates for Shuffled Charge Imaging

D.1. Summary

Using a Starlight Xpress MX-5 CCD camera and customised code for its USB I/O interface (54), the Shuffled Charge Imaging (described in chapter 5) readout rate was tested for various sub-frame sizes and pixel binning. The results are presented here.

D.2. Timing results

The start time of each frame transfer to the readout array of the MX-5 CCD was measured in two ways. The first used a signal transmitted between the USB I/O interface and the camera head, and an oscilloscope was used to monitor the signal, via a pin on a connecting cable, to infer the readout rate. This rate was compared with the rate determined by the CCD control software, running on Linux (kernel version 2.4.18) and using a computer-based hardware timer. The host computer had an Intel motherboard with a UHCI controller. Similar tests on a computer with a non-Intel motherboard and a OHCI controller could not produce stable frame rates without first ensuring that the interrupt used for the USB controller was not shared with any other devices. The timings with this second computer were otherwise similar to those given here.

The internal timer results matched the rate from the oscilloscope to within $\sim 100\,\mu$s, and the table below presents the software timings of the mean readout for one sub-frame. A sub-frame size or pixel binning of $X \times Y$ means X pixels horizontally and Y vertically.

Each frame transfer to the CCD readout array must be spaced in time by a multiple of 1 ms, as the USB clock tick operates at 1 kHz. Non-integral values for the frame readout time have variable readout rates and therefore larger variance. In general, the readout rate is stable until the number of pixels being readout is $\sim 32^2$. This prevented the continuous read out of a sub-frame which is the width of the MX-5 CCD. There also appears to be no simple rule governing the frame rate, for example see the results for 2×2 binning. The sub-frame chosen for the NES tests, 15×8 binned pixels with a binning of 4×3, gave a frame readout time of 2.99 ± 0.01 ms (equivalently, a readout rate of 333 Hz.)

sub-frame size /unbinned pixels	**binning /CCD pixels**	**frame integration time,** $\langle t_{int} \rangle$/ms	**variance** $\sigma_{t_{int}}$/ms^2
1×1	1×1	1.99	0.03
2×2	1×1	1.99	0.02
4×4	1×1	1.99	0.03
8×8	1×1	1.99	0.03
16×16	1×1	3.98	0.03
32×32	1×1	11.95	0.14
64×64	1×1	42.81	1.84
128×128	1×1	167.87	32.42
16×32	1×1	6.97	0.02
32×16	1×1	6.97	0.02
16×32	2×1	4.98	0.01
32×16	2×1	3.98	0.01
16×32	1×2	3.98	0.01
32×16	1×2	3.98	0.01
16×32	2×2	2.98	0.01
30×16	2×2	2.99	0.01
32×16	2×2	2.50	0.26
34×16	2×2	2.98	0.01
16×32	2×3	2.49	0.26
32×16	2×3	2.49	0.26
16×32	3×2	2.49	0.27
32×16	3×2	1.99	0.03

Bibliography

[1] Armstrong J.T. et al., 1998. The Navy Prototype Optical Interferometer. *Astrophysical Journal*, 496:550.

[2] Avila R., Vernin J., & Masciadri E., 1997. Whole atmospheric-turbulence profiling with Generalized SCIDAR. *Applied Optics*, 36:7898–7905.

[3] Babcock H.W., 1953. The possibility of compensating atmospheric seeing. *Publications of the Astronomical Society of the Pacific*, 65:229–236.

[4] Baldwin J.E. et al., 1996. The first images from an optical aperture synthesis array: mapping of Capella with COAST at two epochs. *Astronomy and Astrophysics*, 306:L13–L16.

[5] Blanchard P.M., & Greenaway A.H., 1999. Simultaneous Multiplane Imaging with a Distorted Diffraction Grating. *Applied Optics*, 38:6692–6699.

[6] Born M., & Wolf E., 1980. *Principles of Optics*. Pergamon Press.

[7] Buscher D.F., 1988. Optimizing a ground-based optical interferometer for sensitivity at low light levels. *Monthly Notices of the Royal Astronomical Society*, 235:1203–1226.

[8] Buscher D.F., & Shaklan S.B., 1994. Low-order adaptive optics and single-mode fibers in stellar interferometry. In *Proceedings SPIE 2201, Adaptive Optics in Astronomy*, pages 980–988.

[9] Chanan G., Troy M., & Sirko E., 1999. Phase Discontinuity Sensing: A Method for Phasing Segmented Mirrors in the Infrared. *Applied Optics*, 38:704–713.

[10] Colavita M.M., Shao M., & Staelin D.H., 1987. Atmospheric phase measurements with the Mark III stellar interferometer. *Applied Optics*, 26:4106–4112.

[11] Colavita M.M. et al., 1999. The Palomar Testbed Interferometer. *Astrophysical Journal*, 510:505–521.

[12] Conan J., Rousset G., & Madec P., 1995. Wave-front temporal spectra in high-resolution imaging through turbulence. *Optical Society of America Journal*, 12:1559–1570.

[13] Creech-Eakman M.J. et al., 2003. The Magdalena Ridge Optical Interferometer and its Science Drivers. *Meeting of the American Astronomical Society*, 203.

[14] Cubalchini R., 1979. Modal wave-front estimation from phase derivative measurements. *Journal of the Optical Society of America*, 69:972–977.

[15] Davis J. et al., 1999. The Sydney University Stellar Interferometer - I. The instrument. *Monthly Notices of the Royal Astronomical Society*, 303:773–782.

[16] Dean B.H., & Bowers C.W., 2003. Diversity selection for phase-diverse phase retrieval. *Journal of the Optical Society of America A*, 20:1490–1504.

[17] Dravins D., Lindegren L., Mezey E., & Young A.T., 1998. Atmospheric Intensity Scintillation of Stars. III. Effects for Different Telescope Apertures. *Publications of the Astronomical Society of the Pacific*, 110:610–633.

[18] Ehgamberdiev S.A. et al., 2000. The astroclimate of Maidanak Observatory in Uzbekistan. *Astronomy and Astrophysics Supplement*, 145:293–304.

[19] Fienup J.R., Thelen B.J., Paxman R.G., & Carrara D.A., 1998. Comparison of phase diversity and curvature wavefront sensing. In *Proceedings SPIE 3353, Adaptive Optical System Technologies*, pages 930–940.

[20] Fried D.L., 1965. Statistics of a Geometric Representation of Wavefront Distortion. *Optical Society of America Journal*, 55:1427–1435.

[21] Gonsalves R.A., 1982. Phase retrieval and diversity in adaptive optics. *Optical Engineering*, 21:829–832.

[22] Goodman J.W., 1996. *Introduction to Fourier Optics*. McGraw-Hill.

[23] Haniff C.A. et al., 2000. COAST: the current status. In *Proceedings SPIE 4006, Interferometry in Optical Astronomy*, pages 627–633.

[24] Haniff T.M., Hu A.K., & Green E.D., 1994. Self-aligning spatial filter. In *Proceedings SPIE 2220, Sensing, Imaging, and Vision for Control and Guidance of Aerospace Vehicles*, pages 6–14.

[25] Harbers G., Kunst P.J., & Leibbrandt G.W.R., 1996. Analysis of lateral shearing interferograms by use of Zernike polynomials. *Applied Optics*, 35:6162–6172.

[26] Hardy J.W., 1993. Compensating Imaging System. In *ESO Conference Proceeedings 48, Active and Adaptive Optics*, pages 107–114.

[27] Hardy J.W., 1998. *Adaptive optics for astronomical telescopes*. Oxford University Press.

[28] Hardy J.W., & Wallner E.P., 1994. Wavefront compensation using active lenses. In *Proceedings SPIE 2201, Adaptive Optics in Astronomy*, pages 77–87.

[29] Herriot G. et al., 1998. Innovations in Gemini adaptive optics system design. In *Proceedings SPIE 3353, Adaptive Optical System Technologies*, pages 488–499.

[30] Horton A.J., Buscher D.F., & Haniff C.A., 2001. Diffraction losses in ground-based optical interferometers. *Monthly Notices of the Royal Astronomical Society*, 327:217–226.

[31] Hubin N., & Noethe L., 1993. Active Optics, Adaptive Optics, and Laser Guide Stars. *Science*, 262:1390–1394.

[32] Johnston D.C., & Welsh B.M., 1994. Analysis of multi-conjugate adaptive optics. *Journal of the Optical Society of America A*, 11:394–408.

[33] Keen J.W., 2003. Design and Implementation of Spatial Filtering at COAST. Ph.D. thesis, University of Cambridge.

[34] Keen J.W., Buscher D.F., & Warner P.J., 2001. Numerical simulations of pinhole and single-mode fibre spatial filters for optical interferometers. *Monthly Notices of the Royal Astronomical Society*, 326:1381–1386.

[35] Koehler B., & Flebus C., 2000. VLTI auxiliary telescopes. In *Proceedings SPIE 4006, Interferometry in Optical Astronomy*, pages 13–24.

[36] Kolmogorov A.N., 1941. Dissipation of energy in the locally isotropic turbulence. *Comptes rendus de l'Académie des Sciences de l'U.R.S.S.*, 32:16–18.

[37] Labeyrie A., 1975. Interference fringes obtained on Vega with two optical telescopes. *Astrophysical Journal Letters*, 196:L71–L75.

[38] Lane B.F., & Colavita M.M., 2003. Phase-referenced Stellar Interferometry at the Palomar Testbed Interferometer. *Astronomical Journal*, 125:1623–1628.

[39] Lawson P.R. et al., 1998. Multi-wavelength fringe measurement with a CCD spectrometer at COAST. In *Proceedings SPIE 3350, Astronomical Interferometry*, pages 753–761.

[40] Le Louarn M., Hubin N., Sarazin M., & Tokovinin A., 2000. New challenges for adaptive optics: extremely large telescopes. *Monthly Notices of the Royal Astronomical Society*, 317:535–544.

[41] Linfield R.P., Colavita M.M., & Lane B.F., 2001. Atmospheric Turbulence Measurements with the Palomar Testbed Interferometer. *Astrophysical Journal*, 554:505–513.

[42] Love G.D. and Bhandari R., 1994. Optical properties of a QHQ ferroelectric liquid crystal phase modulator. *Optics Communications*, 11:475–478.

[43] Mackay C.D., Tubbs R.N., Bell R., Burt D.J., Jerram P., & Moody I., 2001. Subelectron read noise at MHz pixel rates. In *Proceedings SPIE 4306, Sensors and Camera Systems for Scientific, Industrial, and Digital Photography Applications II*, pages 289–298.

[44] Madec P.Y, Conan J.M., & Rousset G., 1993. Temporal characterisation for wave-front in Adaptive Optics. In *ESO Conference Proceedings 42, Telescope and Instrumentation Technologies*, pages 471–474.

[45] McAlister H.A. et al., 2000. The CHARA array on Mt. Wilson, California. In *Proceedings SPIE 4006, Interferometry in Optical Astronomy*, pages 465–471.

[46] McGlamery B.L., 1976. Computer Simulations of Compensation of Turbulence Degraded Images. In *Proceedings SPIE 74*, pages 225–233.

[47] Monnier, J. D., 2003. Optical interferometry in astronomy. *Reports of Progress in Physics*, 66:789–857.

[48] Mourard D. et al., 2003. Status report on the GI2T interferometer. In *Proceedings SPIE 4838, Interferometry for Optical Astronomy II*, pages 9–18.

[49] Muñoz-Tunón C., Vernin J., & Varela A.M., 1997. Night-time image quality at Roque de Los Muchachos Observatory. *Astronomy and Astrophysics Supplement*, 125:183–193.

[50] Murty M., 1964. The use of a single plane parallel plate as a lateral shearing interferometer with a visible gas laser source. *Applied Optics*, 3:531–534.

[51] Neil M.A.A., Booth M.J., & Wilson T., 2000. New modal wave-front sensor: a theoretical analysis. *Journal of the Optical Society of America A*, 17:1098–1107.

[52] Nightingale N.S., & Buscher D.F., 1991. Interferometric seeing measurements at the La Palma Observatory. *Monthly Notices of the Royal Astronomical Society*, 251:155–166.

[53] Noll R.J., 1976. Zernike polynomials and atmospheric turbulence. *Optical Society of America Journal*, 66:207–211.

[54] O'Donovan B. et al., 2003. DIMMWIT Comparing atmospheric seeing values measured by a Differential Image Motion Monitor, Which Is Transportable and COAST. In *Proceedings SPIE 4838, Interferometry for Optical Astronomy II*, pages 794–802.

[55] Ohishi N. et al., 2003. Wavefront tilt correction system for MIRA-I.2. In *Proceedings SPIE 4838, Interferometry for Optical Astronomy II*, pages 736–747.

[56] Paterson C., & Dainty J.C., 2000. Hybrid curvature and gradient wavefront sensor. *Optics Letters*, 25:1687–1689.

[57] Perrin G., Coudé du Foresto V., Ridgway S.T., Mariotti J., & Benson J.A., 1995. Fibered recombination unit for the Infrared-Optical Telescope Array. In *Proceedings SPIE 2476, Fiber Optics in Astronomical Applications*, pages 120–128.

[58] Press W.H. et al., 1993. *Numerical Recipes in C*. Cambridge University Press.

[59] Ragazzoni R., 1996. Pupil plane wavefront sensing with an oscillating prism. *Journal of Modern Optics*, 43:289–293.

[60] Rigaut F. et al., 1998. Performance of the Canada-France-Hawaii Telescope Adaptive Optics Bonnette. *Publications of the Astronomical Society of the Pacific*, 110:152–164.

[61] Robbe S., Sorrente B., Cassaing F., Rabbia Y., & Rousset G., 1997. Performance of the angle of arrival correction system of the I2T + ASSI stellar interferometer. *Astronomy and Astrophysics Supplement*, 125:367–380.

[62] Robinson S.R., 1978. On the problem of phase from intensity. *Journal of the Optical Society of America*, 68:87–93.

[63] Roddier F., 1981. The effects of atmospheric turbulence in optical astronomy. In *Progress in Optics XIX*, pages 281–376.

[64] Roddier F., 1988. Curvature sensing and compensation: a new concept in adaptive optics. *Applied Optics*, 27:1223–1225.

[65] Roddier F., 1988. Passive Versus Active Methods in Optical Interferometry. In *ESO Conference Proceedings 29, High-Resolution Imaging by Interferometry*, pages 565–586.

[66] Roddier F., Northcott M., & Graves J.E., 1991. A simple low-order adaptive optics system for near-infrared applications. *Publications of the Astronomical Society of the Pacific*, 103:131–149.

[67] Roddier F., Northcott M.J., Graves J.E., McKenna D.L., & Roddier D., 1993. One-dimensional spectra of turbulence-induced Zernike aberrations: time-delay and isoplanicity error in partial adaptive compensation. *Journal of the Optical Society of America*, 10:957–965.

[68] Rousset G., Fontanella J.C., Kern P., Gigan P., & Rigaut F., 1990. First diffraction-limited astronomical images with adaptive optics. *Astronomy and Astrophysics*, 230:L29–L32.

[69] Schmenk D., 2004. `http://home.earthlink.net/~dschmenk/`.

[70] Shaklan S., & Roddier F., 1988. Coupling starlight into single-mode fiber optics. *Applied Optics*, 27:2334–2338.

[71] Shao M., & Colavita M.M., 1992. Potential of long-baseline infrared interferometry for narrow-angle astrometry. *Astronomy and Astrophysics*, 262:353–358.

[72] Shao M. et al., 1988. The Mark III stellar interferometer. *Astronomy and Astrophysics*, 193:357–371.

[73] Shelton J.C., Schneider T., McKenna D., & Baliunas S.L., 1995. First tests of the Cassegrain adaptive optics system of the Mount Wilson 100 inch telescope. In *Proceedings SPIE 2534, Adaptive Optical Systems and Applications*, pages 72–77.

[74] St-Jacques D., 1998. Astronomical Seeing in Space and Time — a study of atmospheric turbulence in Spain and England, 1994–98. Ph.D. thesis, University of Cambridge.

[75] St-Jacques D. et al., 1997. The JOSE atmospheric seeing monitor at the William Herschel Telescope. *Monthly Notices of the Royal Astronomical Society*, 290:66–74.

[76] Starlight Xpress Ltd., 2004. Starlight Xpress Ltd., `http://www.starlight-xpress.co.uk/`.

[77] Tango W.J., & Twiss R.Q., 1980. Michelson stellar interferometry. In *Progress in Optics XVII*, pages 239–277.

[78] Tatarski V.I., 1961. *Wavefront Propagation in a Turbulent Medium*. Dover.

[79] Taylor G.I., 1931. The Spectrum of Turbulence. *Proceedings of the Royal Society of London A*, 164:476–490.

[80] ten Brummelaar T.A., 1993. Taking the twinkle out of the stars: An adaptive wavefront tilt correction servo and preliminary seeing study for SUSI. Ph.D. thesis, University of Sydney.

[81] ten Brummelaar T.A., 1995. The contribution of high order Zernike modes to wavefront tilt. *Optics Communications*, 115:417–424.

[82] ten Brummelaar T.A., Bagnuolo W.G., & Ridgway S.T., 1995. Strehl ratio and visibility in long-baseline stellar interferometry. *Optics Letters*, 20:521–523.

[83] ten Brummelaar T.A., & Tango W.J., 1994. A wavefront tilt correction servo for the Sydney University Stellar Interferometer. *Experimental Astronomy*, 4:297–315.

[84] Vérinaud C., & Cassaing F., 2001. Piston control with adaptive optics in stellar interferometry. Application to the GI2T interferometer and bimorph mirrors. *Astronomy and Astrophysics*, 365:314–323.

[85] Vaughnn D., Claver C.F., & Richardson E.H., 2000. Four-quadrant error sensor that yields position and focus utilizing an internal mirrorlette array. In *Proceedings SPIE 4008, Optical and IR Telescope Instrumentation and Detectors*, pages 1413–1422.

[86] Vernin J., & Muñoz-Tunón C., 1994. Optical seeing at La Palma Observatory. 2: Intensive site testing campaign at the Nordic Optical Telescope. *Astronomy and Astrophysics*, 284:311–318.

[87] Wilson R., 1998. Results of the JOSE site evaluation project for adaptive optics at the William Herschel Telescope. *New Astronomy Review*, 42:465–469.

[88] Wizinowich P. et al., 2000. First Light Adaptive Optics Images from the Keck II Telescope: A New Era of High Angular Resolution Imagery. *Publications of the Astronomical Society of the Pacific*, 112:315–319.