NATIONAL UNIVERSITY OF IRELAND, GALWAY

Optimal post processing of AO corrected astronomical images: application to faint companion detection and characterisation

by

Daniel Burke

Supervisor:

Dr. Nicholas Devaney

A thesis submitted in partial fulfillment for the degree of Doctor of Philosophy

in the
Applied Optics Group
School of Physics
National University of Ireland, Galway

June 2011

"It's still magic even if you know how it's done."

Terry Pratchett

Abstract

High contrast adaptive optics correction of image degradation caused by imaging through the atmosphere of the Earth from the ground enables direct imaging of faint companions. Differential astrometry and photometry of faint companions is an important task in this branch of astronomy. However; the direct imaging of faint companions in high contrast adaptive optics corrected images is limited by instrumentally induced quasi-static speckles present in the PSF of the parent star of the faint companion.

In this thesis statistical decision theory is applied to the problem of detecting, locating and estimating the differential intensity of a faint companion in the presence of quasi-static speckle noise. The Hotelling observer, which is the optimal linear observer, is rigorously derived form the ideal observer to carry out this task.

This study is broken into two parts: the first is comparing the performance of the Hotelling observer, through using simulated and real adaptive optics corrected data, to other widely used algorithms for computing differential astrometry and photometry of faint companions, the second part addresses the difficulty in distinguishing the signal of a faint companion from that of a quasi-static speckle by combining the Hoteling observer with differential imaging techniques to suppress the quasi-static speckle noise and alternatively using multi-wavelength data to estimate the long-exposure point spread function of the imaging system.

Acknowledgements

Firstly I would like to thank my supervisor, Dr. Nicholas Devaney, for his constant guidance and encouragement throughout my Ph.D study. I have learnt so much from you in the last four years, thank you Nicholas. I am indebted to Professor Chris Dainty for giving me the opportunity to work in the Applied Optics Group. More than anything you have taught me how to look at difficult problems in a clear and simple way. I am proud to say I have been a student of yours. I am also very grateful to Dr. Szymon Gladysz for adding his insight to the project as well as key discussions on the development and application of my code.

I want to thank my friend, and mentor, Milo, we have come a long way together since our long philosophical discussions in Macs. Thanks to the *physics lads*: Kev, Kev, Sean, Jev and Elaine for all the laughs over coffee and help throughout my Ph.D.

Last but not least I want to thank my family. In particular my Mother and Father for their unconditional support throughout my life, I know at times my temper can get the better of me. I hope I have made you proud. Brigid, without you there is no way I could have finished my Ph.D. study. I would also like to thank my brother Pa, you have been a second father to me, I won't forget all the support and advice you have given me over the course of my *long* education. I also want to thank my aunt Mar for always being there when I needed her.

Dan Burke,
December 2010

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

ADI	**A**ngular **D**ifferential **I**maging
AEROC	**A**rea under the **EROC** curve
ALROC	**A**rea under the **LROC** curve
AO	**A**daptive **O**ptics
AUC	**A**rea under the **ROC** curve
BKE	**B**ackground **K**nown **E**xactly
EROC	**E**stimation **ROC** curve
FPF	**F**alse **P**ositive **F**raction
FWHM	**F**ull **W**idth **H**alf **M**aximum
IBD	**I**terative **B**lind **D**econvolution
IFU	**I**ntegral **F**ield **U**nit
IPA	**I**nterior **P**oint **A**lgorithm
IRAF	**I**mage **R**eduction **A**nalysis **F**acility
KKT	**K**arush **K**uhn **T**ucker conditions
LROC	**L**ocalisation **ROC** curve
MSE	**M**ean **S**quare **E**rror
OHE	**O**ptimal **H**otelling **E**stimator
OTF	**O**ptical **T**ransfer **F**unction
PAOLA	**P**erformance of **A**daptive **O**ptics for **L**arge and little **A**pertures
PDF	**P**robability **D**ensity **F**unction
PKE	**P**SF **K**nown **E**xactly
PSF	**P**oint **S**pread **F**unction

PSNR	**P**eak **S**ignal to **N**oise **R**atio
RMS	**R**oot **M**ean **S**quare
ROC	**R**eceiver **O**perating **C**haracteristic **C**urve
SAA	**S**hift **A**nd **A**dd
SDI	**S**pectral **D**ifferential **I**maging
SKE	**S**ignal **K**nown **E**xactly
SNR	**S**ignal to **N**oise **R**atio
SR	**S**trehl **R**atio
SSD	**S**tochastic **S**peckle **D**iscrimination
SSHE	**S**patial **S**canning **H**otelling **E**stimator
TPF	**T**rue **P**ositive **F**raction
2AFC	**2** **A**lternative **F**orced **C**hoice

To

Dad

Preface

Observing faint binary star companions from the ground using large astronomical telescopes is inherently difficult due to the image degradation caused by imaging through the turbulent atmosphere of the Earth. Adaptive optics (AO) correction of these turbulence effects enables very high angular resolution imaging. Given adequate resolution great difficulty then arises in distinguishing a faint companion from residual background noise.

After detection, differential astrometry and photometry of faint companions in adaptive optics observations is an important problem in astronomy (Roberts et al., 2005, 2007). Methods are currently being developed for application to Extremely Large Telescope (ELT) images of exoplanets (Kasper et al., 2008). For exoplanets, differential photometry can be used to derive the planetary mass assuming one of the theoretical models for either reflected or internal light (Baraffe et al., 2003). Hence, an estimate of the mass of the planet could in principle be derived from a single observation. The importance of accurate binary star photometry lies in determining one of the most basic properties of a star: it's mass. When the mass of a survey of stars is estimated accurately, this information can serve as a test of stellar formation and evolution models (Turner et al., 2008).

The acquisition of scientific data is carried out for a specific purpose. The quality of the data produced by a measuring system should be judged on how well the recorded data fulfilled that purpose. The primary motivation behind this thesis was to apply a task based objective assessment of image quality to adaptive optics corrected astronomical images, using the ideal linear observer to quantify the suitability of the data and the performance of the imaging system to the tasks of detecting, locating and deriving relative photometry of a faint binary companion. Although this task based assessment strategy is relatively new in astronomy (Barrett et al., 2006) it is now common practice in the field of medical radiology (Barrett & Myers, 2004).

The principal complication with using the ideal linear observer to perform faint companion detection is the presence of a diffraction limited point source like noise pattern referred to as quasi-static speckle noise. When applying the ideal linear observer to data this noise signal appears identical to faint point sources. Hence this thesis also sets out to investigate approaches to reducing the residual quasi-static noise source such that the ideal linear observer can be applied to the processed data.

Parts of the work presented in this thesis was carried out in collaboration with Professor Harrison Barrett and Luca Caucci of the College of Optical Sciences, University of Arizona, Dr. Szymon Gladysz of the European Organisation for Astronomical Research in the Southern Hemisphere, Dr. Lewis Roberts of the Jet Propulsion Laboratory, California Institute of Technology and Kevin Murphy of the Applied Optics Group. Details of our collaborative work is outlined in the synopsis below.

Synopsis

Chapter 1 presents the background to imaging through atmospheric turbulence. A general description of the propagation of light through the atmosphere is given, followed by a look at the specific details of imaging with adaptive optics that this thesis is concentrated on, namely point source detection and characterisation in the presence of quasi-static speckle noise. The source of this noise is described and two commonly used strategies for dealing with this type of noise are examined.

Chapter 2 introduces the mathematical framework for task based assessment of image quality, which is fundamental to the rest of this thesis. The chapter begins with a detailed overview of statistical decision theory, including a derivation of the receiver operating characteristic curve using a binary decision model. The remaining bulk of this chapter is given over to the derivation of the ideal linear observer, i.e. the Hotelling observer, from the ideal observer. A Gaussian model is used in the derivation to describe the long exposure adaptive optics corrected data upon which the Hotelling observer will operate. The optimal estimators for the extraction of differential faint companion astrometry and photometry are then derived. The mathematical similarities between the iterative blind deconvolution estimator for intensity and the Hotelling estimator for intensity is discussed.

Chapter 3 centres around the application of the Hotelling observer on simulated adaptive optics corrected data. The data simulation tool PAOLA is described. Work carried out to test the performance of the Hotelling observer in varying simulated observing conditions is reported upon. The robustness of the Hotelling observer to errors in estimate values of the data was also investigated. The following section compares three methods of applying the Hotelling observer to estimate the location of a faint companion in a simulated adaptive optics corrected image. The three algorithms are described in detail with their specific benefits and drawbacks discussed. This chapter concludes by extending the algorithm to also extract differential photometry from the data. The unbiased Hotelling intensity estimator is compared to several other photometric approaches.

The experiments described in section (3.5) were carried out in collaboration with Kevin Murphy of the Applied Optics Group, who carried out the experiment in the laboratory and used my proposed Shack-Hartmann data analysis as an input to his optical vortex detection algorithm.

Chapter 4 contrasts the differential astrometry and photometry calculated by the Hotelling observer to two state of the art algorithms widely used in astronomy. The PSF fitting routine StarFinder and the iterative blind deconvolution package Fitstars are reviewed in section (4.1). Real adaptive optics corrected data of point sources was available from the Lick observatory and was used to compare the methods by constructing artificial binary images. The results of this comparison are presented in section (4.2.2). As deconvolution and matched filtering are mathematically similar strategies the peak signal to noise ratio was used in section (4.3) to show the benefit of prewhitening the data. The chapter closes with the analysis of three observed real binary stars, with the Hotelling observer extracting comparable astrometry and photometry to the StarFinder algorithm. Sections (4.2) and (4.4) describes work carried out with Dr Szymon Gladysz, who provided and reduced the AO corrected Lick data and applied the StarFinder algorithm to the artificial binary data, and Dr. Lewis Roberts who processed the artificial binary data with the Fitstars code.

Chapter 5 addresses the primary limiting factor to faint point source detection, specifically the presence of residual quasi-static speckle noise. An alternative data simulation algorithm to PAOLA is presented, in which a model of the pupil phase is used to define the speckle noise. Section (5.2) reports upon the combination of spectral difference imaging with the Hotelling observer. The results of this merger illustrate the power of the Hotelling observer with dealing with spatially uncorrelated noise. The second part of this chapter derives and quantifies the performance of the angular differencing approach to suppressing quasi-static speckle noise. The primary goal of this method is to reconstruct a PSF which once subtracted from the data minimises the noise residuals. The residual image is processed by the Hotelling observer and the area under the estimation receiver operating characteristic curve is used to quantify the performance of the algorithm as a function of the angular separation of the star and faint companion. The angular differencing approach is extended in section (5.3.5) to also make use of multi-spectral data. The additional information provided by the multi-spectral data further attenuates the quasi-static speckle noise and leads to an increase in the area under the estimation receiver operating characteristic curve of a factor of 1.32.

Chapter 6 approaches the task of PSF reconstruction from multi-spectral data by a different route to that laid out in chapter (5). An alternate image formation model is

presented, where the intensity in the focal plane of a telescope is described by a non-linear relationship with the phase in the pupil plane, that is they are a Fourier transform pair. By perturbing the unknown phase in a known manner the unknown phase and hence the PSF can be recovered. A PSF recovery algorithm is presented which makes use of the change in phase with wavelength. The level of speckle noise attenuation achieved with this method is compared to the approaches employed in chapter (5).

Chapter 7 concludes the work presented in this thesis and discusses the implications for point source detection and characterisation in highly corrected adaptive optics images. Finally, some suggestions for future related topics of research are given.

Publications

- D. Burke, N. Devaney, S. Gladysz, H.H. Barrett, M. K. Whitaker and L. Caucci, "Optimal linear estimation of binary star parameters", Proc SPIE, Vol. 7015, 70152J, 2008
- D. Burke, S. Gladysz, L. C. Roberts, N. Devaney and C. Dainty, "An Improved Technique for the Photometry and Astrometry of Faint Companions", Publications of the Astronomical Society of the Pacific, Vol. 121, July 2009
- D. Burke, N. Devaney and S. Gladysz, "Determination of astrometry and photometry of faint companions in the presence of residual speckle noise", Proc 1st AO4ELT conference, 09001, 2010
- D. Burke, N. Devaney, S. Gladysz, and C. Dainty, "Enhanced Faint Companion Photometry and Astrometry Using Wavelength Diversity", in Adaptive Optics: Methods, Analysis and Applications, OSA Technical Digest (CD) (Optical Society of America, 2009), paper AOTuC5
- D. Burke, N. Devaney, J. Christou and M. Hartung, "Application of wavelength diversity for astronomical adaptive optics imaging", Proc SPIE, Vol. 7736, 221, 2010
- K. Murphy, D. Burke, N. Devaney and C. Dainty, "Experimental detection of optical vortices with a Shack-Hartmann wavefront sensor," Optics Express, Vol. 18, Issue 15, pp. 15448-15460, 2010
- D. Burke and N. Devaney, "Enhanced Faint Companion Photometry and Astrometry using Wavelength Diversity", JOSA A, Vol. 27, No. 11, November 2010

CHAPTER 1

ADAPTIVE OPTICS IMAGING

When light from an astronomical source passes through the atmosphere of the Earth the incoming wavefront is distorted. This distortion arises due to light passing through random fluctuations in the refractive index of the atmosphere caused by variations in temperature, humidity and pressure. These changes in refractive index will result in random optical path length differences both in space and time. The purpose of adaptive optics (AO) when imaging with astronomical telescopes is to correct, or compensate, for the distortion imposed upon the incoming wavefront.

1.1 The Kolmogorov Model of Turbulence

The distortion an incoming wavefront is subject to can be modelled using the Kolmogorov theory of turbulence (Kolmogorov, 1941). The Kolmogorov model describes how energy flow is distributed throughout a turbulent fluid medium. It is assumed that large-scale disturbances, on a scale called the outer scale L_0, transfer kinetic energy to other layers by breaking down into increasingly smaller disturbances. This process continues until the energy is transferred to the smallest disturbance scale called the inner scale l_0. The Kolmogorov model assumes no energy is lost until the inner scale is reached, at which point the energy is released as heat; see figure (1.1).

The mean refractive index of air is quite close to unity. However since the refractive index of air is proportional both to temperature and pressure, temperature fluctuations in the atmosphere result in refractive index fluctuations. Temperature changes are the

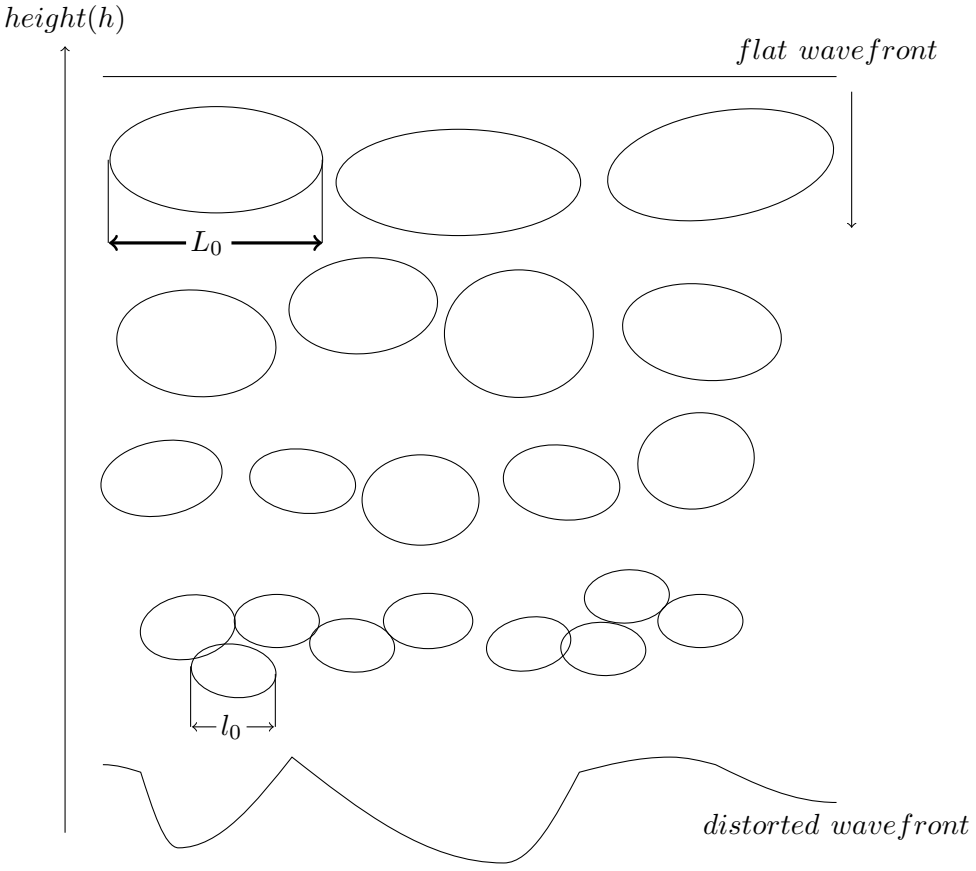


FIGURE 1.1: Kolmogorov energy flow in a turbulent fluid

dominant factor in this process as pressure changes are smoothed out at the speed of sound (Hardy, 1998).

Tatarskii (1961) introduced the use of the *structure function* to account for the slow change in the mean values of properties of the atmosphere, such as humidity, pressure and temperature. A structure function describes the mean square fluctuation, of the scalar quantity qk, between two points separated by Δr:

$$D_{qk}(r, \Delta r) =< |qk(r + \Delta r) - qk(r)|^2 >, \tag{1.1}$$

where <> denotes an ensemble average. Obukhov (1949) showed that the atmospheric temperature structure function can be expressed as:

$$D_t(\Delta r) = C_t^2 (\Delta r)^{\frac{2}{3}}, \tag{1.2}$$

where C_t^2 is the atmospheric temperature structure constant, a measure of the strength of the temperature fluctuations. Roddier (Roddier, 1981, 1988) shows that refractive

index fluctuations are caused by temperature changes in the atmosphere, hence the refractive index structure function, D_n, takes on a very similar form to equation (1.2):

$$D_n(\Delta r) = C_n^2(\Delta r)^{\frac{2}{3}}, \tag{1.3}$$

where C_n^2 is the refractive index structure constant, a measure of the strength of the refractive index variations. The power spectrum of the refractive index changes can be expressed as (Tatarskii, 1961):

$$\Phi_n(\kappa) = 0.033 C_n^2 \kappa^{-\frac{11}{3}}, \tag{1.4}$$

where κ is the spatial wave number, the modulus of a 3D vector. The quantity C_n^2 is a good measure of the contribution that the atmospheric turbulence makes to a wave propagating through it. It is important to note that the refractive index structure constant varies with the altitude h of the turbulent layer. The Kolmogorov power spectrum of the variation in the refractive index can therefore be re-written as:

$$\Phi_n(\kappa, h) = 0.033 C_n^2(h) \kappa^{-\frac{11}{3}}. \tag{1.5}$$

The spatial properties of the turbulent atmosphere can now be fully described by a single function, i.e. the C_n^2 function, under the assumption that κ is inside the inertial range, $2\pi/L_0 < \kappa < 2\pi/l_0$, where the outer scale is also dependent on altitude.

1.2 Propagation of light through the turbulent atmosphere

This section will address the phase changes to a plane wavefront due to passing through the turbulent atmosphere (Roddier, 1981). To approximate the continuous C_n^2 function the atmosphere can be modelled as a series of phase screens as shown by Lee & Harp (1969). Monochromatic plane waves, of wavelength λ, are considered propagating through an atmospheric layer of thickness δh towards a ground based observer, see figure (1.2).

The thickness of the atmospheric layer is large compared to the scale of the turbulent eddies and therefore Gaussian statistics can apply, but the layer is also thin enough such that diffraction effects can practically be ignored. Denoting the incident complex field as $\Psi_h = 1$ and the resulting complex field after emerging from the thin layer as:

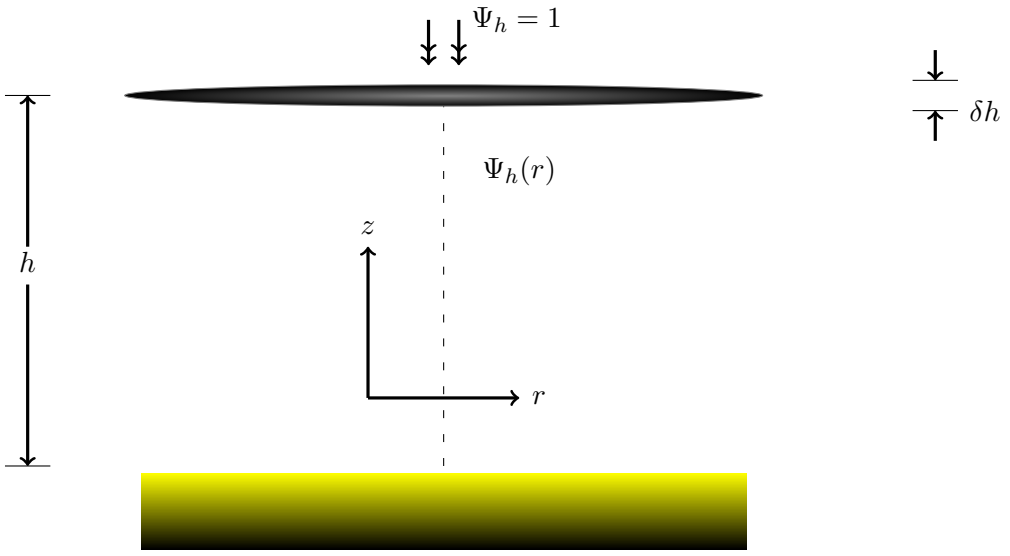


FIGURE 1.2: Model of thin turbulent layer of the atmosphere

$$\Psi_h(r) = e^{i\phi(r)}, \tag{1.6}$$

where the change in phase, $\phi(r)$ produced by the fluctuations in the refractive index $n(r, h)$ is given by:

$$\phi(r) = k \int_h^{h+\delta h} dz \; n(r, z), \tag{1.7}$$

where $k = 2\pi/\lambda$ is the wave number. Following on from here we now wish to define $D_\phi(r)$, the phase spatial structure function, in terms of the variations in refractive index. The covariance of the phase, $\phi(x)$, is given by:

$$B_\phi(r) =< \phi(x)\phi(x + r) >, \tag{1.8}$$

where <> denotes the spatial average and relates to the phase structure function via:

$$D_\phi(r) = 2[B_\phi(0) - B_\phi(r)], \tag{1.9}$$

Roddier (1981) shows that using equation (1.3) $D_\phi(r)$ reduces to:

$$D_\phi(r) = 2.914k^2 C_n^2 \delta h r^{\frac{5}{3}}, \tag{1.10}$$

and the coherence function $B_h(r)$ can be expressed as:

$$\begin{aligned} B_h(r) &= \exp\left[-\frac{1}{2}D_\phi\right], \\ &= \exp\left[-\frac{1}{2}(2.914k^2C_n^2\delta h r^{\frac{5}{3}})\right]. \end{aligned} \tag{1.11}$$

The atmosphere can be modelled as a stack of thin layers; from the perspective of the ground the effects of the atmospheric layers add together linearly. The contribution from each layer to the phase comes from the strength of the refractive index fluctuations over each layer, i.e. $C_n^2\delta h_j$. The coherence function for multiple layers is therefore found by summing over the layers:

$$B_0(r) = \exp[-\frac{1}{2}(2.914k^2r^{\frac{5}{3}}\sum_j C_n^2(h_j)\delta h_j)]. \tag{1.12}$$

This function is of great importance when determining the effect of the atmospheric turbulence on the performance of an imaging system.

An imaging system can be thought to consist of an astronomical telescope coupled with the atmosphere. The optical transfer function (OTF) of this system, for long exposure imaging, can be expressed as:

$$S(f) = B(f).T(f), \tag{1.13}$$

where $B(f)$ is the atmospheric transfer function, $T(f)$ is the telescope transfer function and f is the spatial frequency in cycles per radian. Fried (1966) introduced the concept that the resolving power of a telescope, R, could be defined as the integral of the optical transfer function of the system i.e.:

$$R = \int B(f).T(f)df. \tag{1.14}$$

Fried showed that for large apertures the resolving power only depends upon the atmospheric turbulence,

$$R_{\text{Large Apertures}} = \int B(f)df. \tag{1.15}$$

For small telescopes where the atmospheric effects are negligible the resolving power is a function of the imaging wavelength and the size of the telescope:

$$R_{\text{Small}} = \int T_d(f)df = \frac{\pi}{4}\left(\frac{d}{\lambda}\right)^2. \tag{1.16}$$

Fried also defined the diameter of a telescope, called r_0, which would have the same resolving power as the atmosphere:

$$\int T(f)df = \int B(f)df. \tag{1.17}$$

Comparing equations (1.15) and (1.16) we see that,

$$R = \int B(f)df = \frac{\pi}{4}\left(\frac{d}{\lambda}\right)^2. \tag{1.18}$$

From equation (1.12) Roddier (1981) shows that the atmospheric transfer function can be expressed as:

$$B(f) = \exp(-\frac{1}{2}Kf^{\frac{5}{3}}), \tag{1.19}$$

where K can be shown to be:

$$K = 3.44\left(\frac{r_0}{\lambda}\right)^{-\frac{5}{3}}, \tag{1.20}$$

these equations lead to an expression for the atmospheric transfer function in terms of the Fried coherence length (Fried, 1966) r_0:

$$\begin{aligned} B(f) &= \exp\left(-3.44\left(\frac{\lambda f}{r_0}\right)^{\frac{5}{3}}\right), \\ &= \exp\left(-3.44\left(\frac{r}{r_0}\right)^{\frac{5}{3}}\right). \end{aligned} \tag{1.21}$$

Equating this expression to that in equation (1.12), r_0 can be defined in terms of the integrated turbulence:

$$r_0 = \left[0.423k^2 \sum_j \delta h_j C_n^2(\delta h_j) \right]. \tag{1.22}$$

The phase spatial structure function can now be expressed in terms of the Fried coherence length:

$$D_\phi(r) = 6.88 \left(\frac{r}{r_0} \right)^{\frac{5}{3}}. \tag{1.23}$$

The Fried coherence length is a good measure of the turbulence strength of the atmosphere. Typically r_0 has a value on the order of 10 cm and so any telescope bigger than 10 cm is limited by turbulence, r_0 can therefore give an estimate of the performance of an atmospheric limited imaging system though the optical transfer function relationship (Hardy, 1998; Fried, 1965):

$$< S(f)_{\text{Long Exposure}} >= T(f) \exp \left(-3.44 \left(\frac{\lambda f}{r_0} \right)^{\frac{5}{3}} \right). \tag{1.24}$$

Furthermore Fried showed that the combined transfer function of an imaging system depends upon the size of the telescope D for short exposure imaging:

$$< S(f)_{\text{Short Exposure}} >= T(f) \exp \left(-3.44 \left(\frac{\lambda f}{r_0} \right)^{\frac{5}{3}} \left[1 - \left[\frac{\lambda f}{D} \right]^{\frac{1}{3}} \right] \right). \tag{1.25}$$

This equation can also apply to long exposure images if the short exposure images are re-centred, to remove tilt, and then co-added. When an AO system is used the performance of the imaging system is limited by the ability of the AO system to correct the incoming distorted wavefront. The following sections will outline the imaging of a faint companion with an AO system and how to overcome some of the problems encountered with this type of data acquisition.

1.3 Adaptive Optics

At a well selected astronomical site the coherence length of the atmosphere is typically on the order of $10-20$cm at visible wavelengths (Roddier, 1981). As early as the start of the twentieth century large telescopes, i.e. $D > 1\ m$, were being built to harness more light to image fainter sources. These instruments were limited in angular resolution

by the atmosphere having an effective diameter equal to the coherence length of the atmosphere.

An optical system based on feedback information which could compensate optical path length variations and hence flatten an incoming wavefront was proposed by Babcock (1953).

An astronomical adaptive optics system has three main tasks: measurement of the incoming distorted wavefront, computation of the control signals to be sent to the wavefront corrector and physically correcting the wavefront through the use of a wavefront corrector e.g. a deformable mirror.

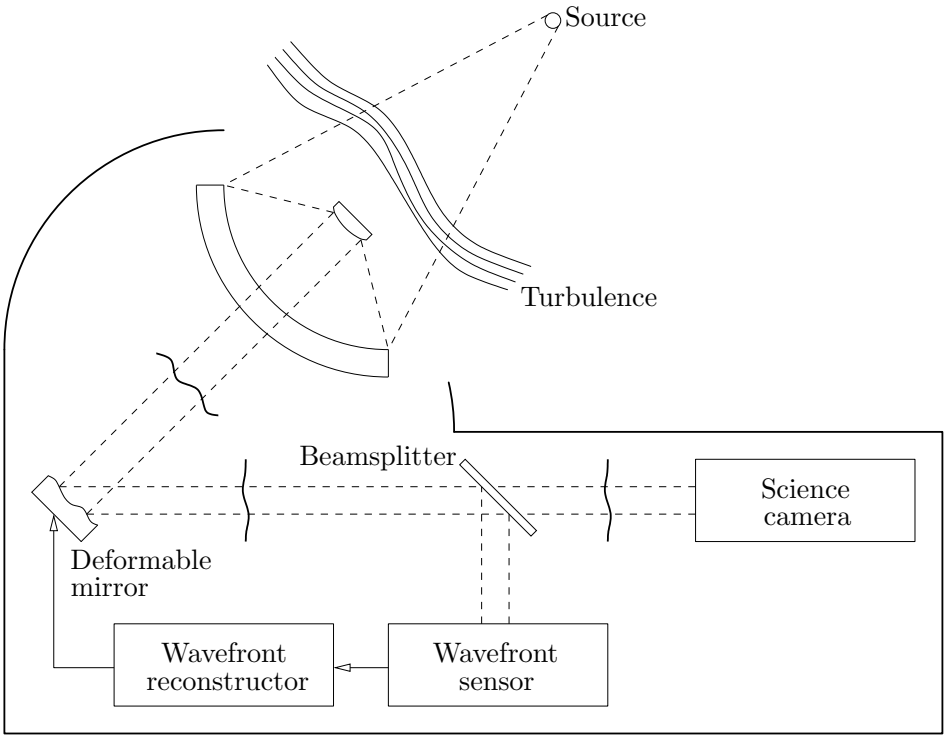


FIGURE 1.3: Setup of an Adaptive Optics Telescope (Caucci, 2007)

An incoming distorted wavefront is first deflected off the primary and secondary mirrors of the telescope, then onto the deformable mirror which attempts to correct, or flatten, the wavefront by applying an appropriate estimated phase correction. The light is then split between the science camera and the wavefront sensor. The wavefront sensor tries to estimate the instantaneous phase in the pupil plane of the telescope. This phase estimate is then used by the control system to determine the correct signals to send to the deformable mirror to correct the wavefront, see figure(1.3). It should be noted that in most systems, the telescope is not afocal as is shown in figure(1.3), the first element of the AO system will then typically be a collimator.

This arrangement of subsystems can be operated in two modes: open-loop and closed loop. In an open-loop setup the wavefront is measured in an uncorrected state. The necessary corrections are computed and applied to the wavefront. The time lag in sensing and correcting the wavefront must be less than the correlation time scale of the atmospheric turbulence. In general open-loop systems are only used when closed-looped system cannot be (Hardy, 1998), for example to pre-correct a laser beam for propagation up through the atmosphere. In closed-loop operation the wavefront sensor measures a corrected wavefront i.e. the wavefront corrector is now the first element in the system. Therefore the wavefront sensor only sees the residual, or compensated, error in the wavefront.

1.3.1 Wavefront Sensing

An astronomical wavefront sensor is required to measure the incoming wavefront and relay the compensated information to the wavefront reconstructor (Rousset, 1993). As natural reference sources emit broadband light it is not practical to make direct measurements of the optical phase, such as with an interferometer (Hardy, 1998). It is more reasonable to measure the direction of propagation of the local wavefront. The Shack-Hartmann setup (Roddier, 1981), see figure (1.4), is the most commonly used approach to astronomical wavefront sensing.

The Shack-Hartmann wavefront sensor spatially samples the wavefront with a two dimensional lenslet array, that is the lenslet array acts as a spatial low pass filter (Hardy, 1998). For an incoming planar wavefront the lenslet array will form a pattern of spots focused on a detector e.g. a CCD. The centroids of these spots can then be estimated. When the incoming wavefront is not a plane wave but has had some distortion imposed upon it the lenslet array will form spots in different locations from the plane wave case. This shift in the positions of each spot can be used to estimate the local slope of the wavefront in each lenslet sub-aperture. Using measurements from the entire array a two dimensional approximation of the wavefront can be constructed.

1.3.2 Wavefront Reconstruction and System Control

Taking the sampled wavefront slope measurements from the wavefront sensor, the reconstruction algorithm must stitch these together in two dimensions preserving the relative phase between the slopes. This is shown in one dimension in figure (1.5).

When the adaptive optics system is operating in a closed loop fashion, the wavefront sensor measures the open-loop phase distortion, ϕ, minus the applied phase correction,

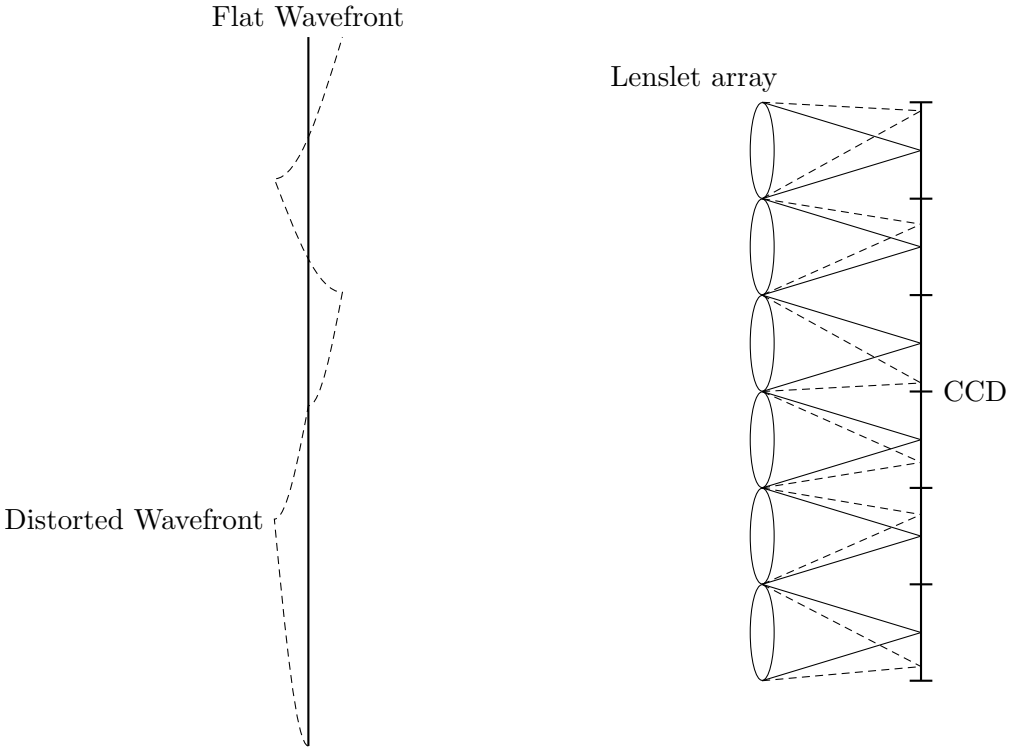


FIGURE 1.4: Shack-Hartmann Wavefront Sensor. The displacement of the image on the sensor is proportional to the wavefront distortion from the reference wavefront.

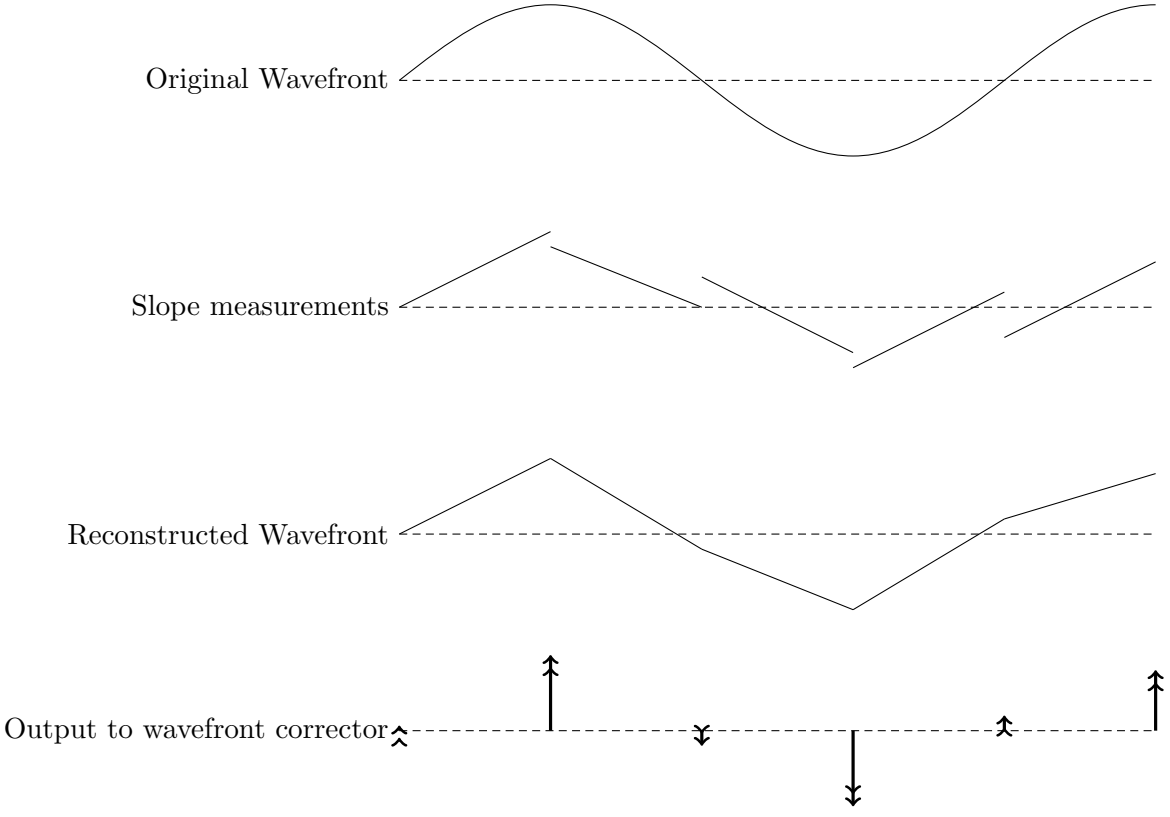


FIGURE 1.5: Wavefront Reconstruction using slope measurements.

φ. The goal of the control system is to compute a set of command signals which will minimise this residual error. There are two main approaches to wavefront reconstruction.

The first involves using a least-squares type reconstructor. This approach is based around minimising the measurement error, ϵ_s, of the wavefront sensor. If S denotes the wavefront sensor slope measurements, A the interaction matrix of the wavefront corrector and $\widetilde{\phi}$ the estimated phase values, the measurement error is given by:

$$\epsilon = ||S - A\widetilde{\phi}||^2, \tag{1.26}$$

where $||\ \ ||$ is the norm of a vector. The wavefront phase $\widetilde{\phi}$ is estimated to minimise ϵ_s. The least-squares solution verifies:

$$(A^T A)\widetilde{\phi} = A^T S, \tag{1.27}$$

where A^T is the transpose of A.

The second approach is formulated around minimising the wavefront residual variance. If $\widehat{\varphi}$ is the estimated wavefront correction, the residual wavefront variance, ϵ_φ, to be estimated is:

$$\epsilon_\varphi = \frac{1}{\text{Area}_{\text{aperture}}} \int_{\text{aperture}} \left\langle (\phi - \widehat{\varphi})^2 \right\rangle d\text{Area}_{\text{aperture}}. \tag{1.28}$$

The estimate of the phase can be given as a linear expansion:

$$\widehat{\varphi} = \sum_i \phi_i f_i, \tag{1.29}$$

where $\phi = \{\phi_i\}$ is the same as in the least squares reconstructor and the functions f_i are the influence functions of the wavefront corrector. The residual wavefront variance minimisation has a matrix solution B called the reconstruction matrix:

$$B \times \text{estimated wavefront slopes} = \text{correction on the wavefront corrector}, \tag{1.30}$$

or,

$$B = C_f^{-1} \left[C_\phi A^T (A C_\phi A^T + C_n)^{-1} \right], \tag{1.31}$$

where C_f is a $N \times N$ matrix of scalar products of the wavefront corrector influence functions f_i, N is the number of wavefront corrector commands or phase values, C_ϕ is the covariance matrix of ϕ and C_n is the covariance matrix of the noise in the wavefront sensor measurement.

1.3.3 Wavefront Correction

The task of the wavefront corrector is to take the information gathered and processed by the wavefront sensor and reconstructor and correct in real time the incoming distorted wavefront. In practice the phase change applied to the incoming wavefront is achieved by altering the optical path length through which the wavefront passes, this is akin to the phase change induced through passing through a varying refractive index medium. Deformable mirrors are mirrors whose surface shape can be altered in order to apply the required phase change to an incoming wavefront (Tyson, 1998), see figure (1.6).

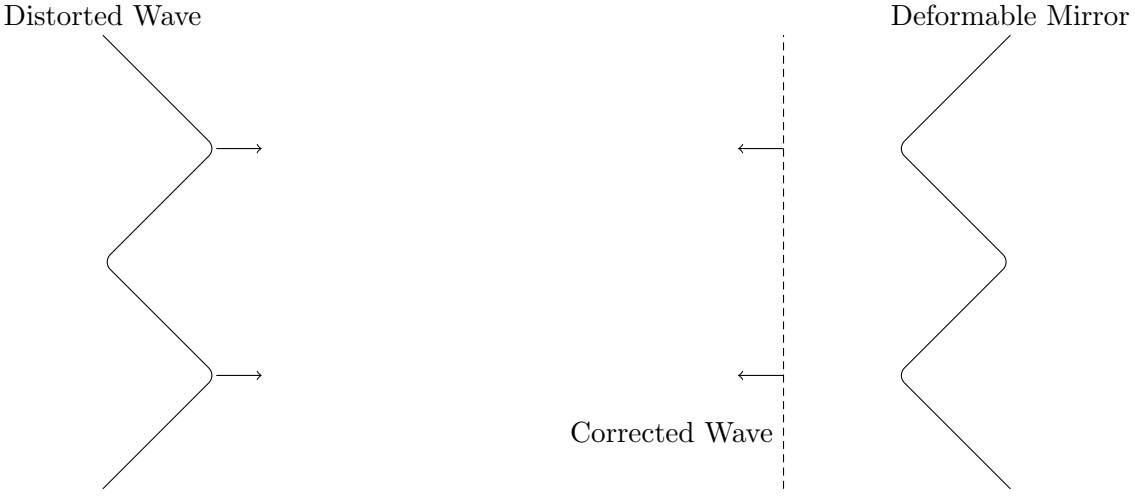


FIGURE 1.6: Wavefront Correction.

The shape assumed by the deformable mirror is controlled by an array of actuators beneath the surface of the mirror. Each actuator has a function which describes its influence upon the shape of the mirror when it is commanded to move. The mirror phase φ can be approximated as a linear superposition of the n actuator influence functions f_j by:

$$\varphi(x) = \sum_{j=0}^{n-1} c_j f_j(x), \tag{1.32}$$

where c_j is the j^{th} actuator command signal.

1.4 Speckle Noise in Adaptive Optics Images

Attempting to detect a faint companion above the halo produced by the diffraction wings of a star is very difficult. When imaging with adaptive optics from the ground, the uncorrected part of the incoming wavefront will produce a random intensity fluctuation in this halo. This pattern is referred to as residual speckle noise. These residual speckles arise primarily from two sources: short exposure 'atmospheric' speckle pinned to the diffraction pattern of the telescope and 'quasi-static' speckle from uncorrected aberrations in the system. The instrumentally induced quasi-static speckles (Marois et al., 2005) will not average out over time (Hinkley et al., 2007; Gladysz & Christou, 2008), but produce a time varying point spread function (PSF).

The relationship between the phase of the wavefront in the pupil plane of the telescope and the focal plane intensity distribution is very important in understanding the formation of residual speckle noise. It can be shown that the Fourier transform of the pupil autocorrelation function characterises the image of a point source i.e. an Airy disk in the case of a diffraction limited telescope with a circular aperture (Roddier, 1999). In the presence of atmospheric turbulence, the resolution of a telescope larger than r_0 is effectively turbulence limited, adaptive optics attempts to overcome this limitation.

A good description of the light intensity distribution and the wave amplitude in the focal plane is given by Aime & Soummer (2004). The complex amplitude of the wavefront in the pupil plane can be expressed as:

$$\Psi_1 = [A + a(x, y)]P(x, y), \tag{1.33}$$

where A corresponds to an incoming plane wave, $a(x, y)$ refers to the random term of the uncorrected part of the wavefront and $P(x, y)$ characterises the pupil transmission. The field in the focal plane is obtained by taking the Fourier transform of Ψ_1:

$$\Psi_2 = \underbrace{A\ F[P(x, y)]}_{C(x,y)} + \underbrace{F[a(x, y)P(x, y)]}_{S(x,y)}. \tag{1.34}$$

The $C(x, y)$ term above describes the wave amplitude when no turbulence is present i.e. an Airy pattern, whereas the random term $S(x, y)$ is associated with the speckle field produced by the 'frozen' atmosphere. How does this speckle field arise? It is the result

of filling the pupil of the telescope with r_0-sized sub-pupils which synthesises a filled-aperture interferometer affected by random phase errors. When the resulting fringes interfere constructively with each other a bright speckle of width $\approx \lambda/D$ appears in the focal plane, where D is the diameter of the pupil (Racine et al., 1999). Given enough realisations, the complex amplitude of the speckle field, $S(x, y)$, will follow a circular Gaussian distribution. This is equivalent to a time averaged view of the statistics of the turbulent atmosphere.

However on short time scales the picture is very different. Aime & Soummer (2004) describe the short exposure image plane intensity as:

$$|\Psi_2(x,y)|^2 = |C(x,y)|^2 + |S(x,y)|^2 + \underbrace{2Re[C^*(x,y)S(x,y)]}_{\text{speckle pinning}}. \tag{1.35}$$

Since the random field is modulated by the pupil diffraction pattern the last term in equation (1.35) leads to the phenomena of speckle pinning as described by (Bloemhof et al., 2001; Sivaramakrishnan et al., 2002; Bloemhof, 2003, 2004) see figure (1.7).

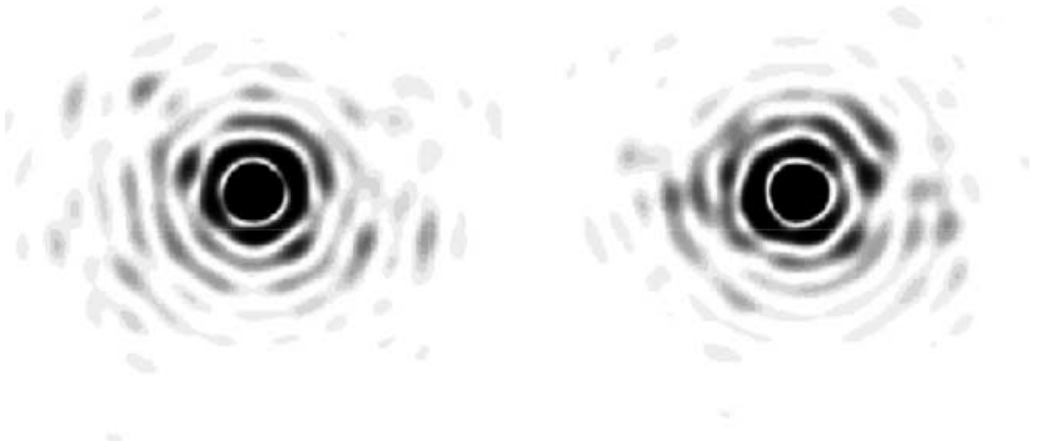


FIGURE 1.7: Two examples of pinned speckle Aime & Soummer (2004)

A better grasp of the speckle pinning effect can be achieved by looking at the probability density function (PDF) of the focal plane intensity distribution. This situation is equivalent to the case of laser speckle added to a coherent background (Goodman, 1975). Soummer et al. (2007) shows that the derivation of the PDF for the focal plane intensity leads to a "modified Rician distribution" derived by Goodman (1975) and used by Cagigal & Canales (2000). The modified Rician distribution is given as:

$$P_I(I) = \frac{1}{I_s} \exp\left[-\frac{I + I_c}{I_s}\right] I_0\left[\frac{2\sqrt{I}\sqrt{I_c}}{I_c}\right], \tag{1.36}$$

where I_0 is the zero order Bessel function of the first kind, I_s corresponds to the speckle halo produced by random intensity variations, I_c is proportional to the perfect PSF and therefore at successive Airy rings I_c is large and small; this then leads to an amplification in the variance of the speckles. When the PSF is at zero no amplification of the speckle occurs and the speckle statistics revert to the negative exponential PDF which characterises the statistics of classical laser speckle. Equation (1.36) holds for AO corrected data with a Strehl ratio greater than 0.4.

When attempting to detect a faint companion in this field a coronagraph could be used to reduce the speckle field. The variance of residual speckle is given as:

$$\sigma_I^2 = I_s^2 + 2I_s I_c. \tag{1.37}$$

Poisson statistics must also be accounted for when the number of photons is small:

$$\sigma_p^2 = I_c + I_s. \tag{1.38}$$

The variances add together and so the total variance becomes:

$$\sigma^2 = \underbrace{2I_c I_s + I_c}_{\sigma_c^2} + \underbrace{I_s^2 + I_s}_{\sigma_s^2}. \tag{1.39}$$

A perfect coronagraph can only remove the coherent part, $C(x, y)$, of the focal plane complex amplitude. Therefore only the variance σ_c^2 can be effected by cancelling I_c. The variance of the speckle term σ_s^2 will be unaffected. Aime & Soummer (2004) showed that a coronagraph will be efficient at reducing the speckle variance when $I_c > I_s$ in the original non-coronagraphic AO images.

The time averaged (long exposure) view of equation (1.35) is the sum of the telescope diffraction pattern with a halo of averaged speckles:

$$< |\Psi_2(x, y)|^2 >= \underbrace{|C(x, y)|^2}_{I_c} + \underbrace{< |S(x, y)|^2 >}_{I_s}. \tag{1.40}$$

1.4.1 Speckle Noise Discrimination

Christou et al. (2006) investigated the distribution of the on-axis intensity in AO short exposure images and found that the distribution showed a negative skewness. As stated above in short exposure AO images the off-axis intensity distribution can be modelled as a modified Rician distribution. Gladysz & Christou (2008) propose a speckle sensitive method which allows discrimination between scattered speckles and real faint sources in an image. They exploit the morphological differences in the on and off axes irradiance distributions to make the distinction between faint sources and speckles. Following on from this work Gladysz & Christou (2009) formulated a stochastic speckle discrimination (SSD) method. For a series of short exposures this technique computes I_c and I_s for each pixel in the focal plane from a time series for each pixel. The SSD algorithm then aims to build a map of local intensity statistics through a transform of spatial intensity variations. It should be noted that this SSD approach is typically used on recentred short exposure shift-and-add (SAA) images. The authors report that two spatial transforms gave good results:

$$I_{SAA} \mapsto \frac{I_s(x,y)}{I_c(x,y)}, \tag{1.41}$$

$$I_{SAA} \mapsto \frac{I_s(x,y)}{I_c(x,y)} m_3(x,y), \tag{1.42}$$

where m_3 is a sample skewness estimator defined as:

$$m_3 = \frac{1}{N} \sum_{j=0}^{N-1} \left(\frac{I_j - E(I)}{\sigma_I} \right)^3, \tag{1.43}$$

where N is the number of frames in the data cube. Gladysz & Christou (2009) processed data from observations of the binary star HD 8799 ($m_v = 4.8$) with this method, see figure(1.8). They noted that the statistical map smoothed out the anisotropies in the SAA image and that the faint companion was much more visible in the statistical map than it was in the original SAA image.

If one does not have access to short exposure data, alternative methods are available to reduce residual quasi-static speckle from long exposure data.

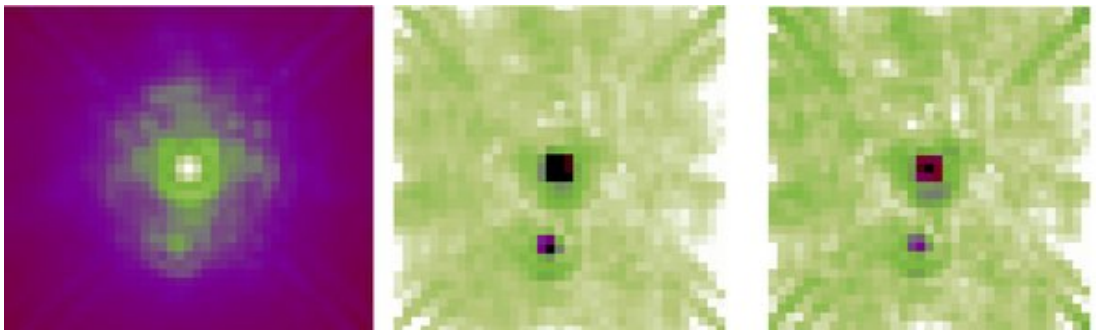


FIGURE 1.8: SAA image of HD 8799 in a square root scale (left), I_s/I_c (centre) and $m_3.I_s/I_c$ (right)(Gladysz & Christou, 2009)

1.5 Differential Imaging

Racine et al. (1999) proposed that quasi static speckle noise could be reduced by subtracting images obtained in two wavelength bands if the planet signal is much fainter in one of the bands. This would be the case when observing gas giant exoplanets in the methane band ($1.6\mu m$) (Rosenthal et al., 1996), see figure(1.9). This imaging scheme is referred to as Spectral Differential Imaging (SDI).

Before the two images could be subtracted from each other two main precautions are needed Marois et al. (2000):

1. One of the frames must be rescaled as the speckle pattern is proportional to wavelength (Marois et al., 2004).
2. The bandpass filters must be close together to preserve the similarities in the speckle structure.
3. Narrow bandwidths must be used to reduce chromatic elongation of the speckles.

Marois et al. (2000) showed that when images from three wavelengths are available a more efficient double differencing of the data can be carried out. Suppose three images are recorded at three wavelengths: $I_1(\lambda_1)$, $I_2(\lambda_2)$ and $I_3(\lambda_3)$ and $\lambda_1 < \lambda_2 < \lambda_3$, the double difference of these images is then given as:

$$dd = (I_1 - I_2) - k(I_1 - I_3), \tag{1.44}$$

where the constant k is given by the ratio of the Strehl ratio's (SR) of the data:

$$k = \frac{SR(I_1) - SR(I_2)}{SR(I_1) - SR(I_3)}. \tag{1.45}$$

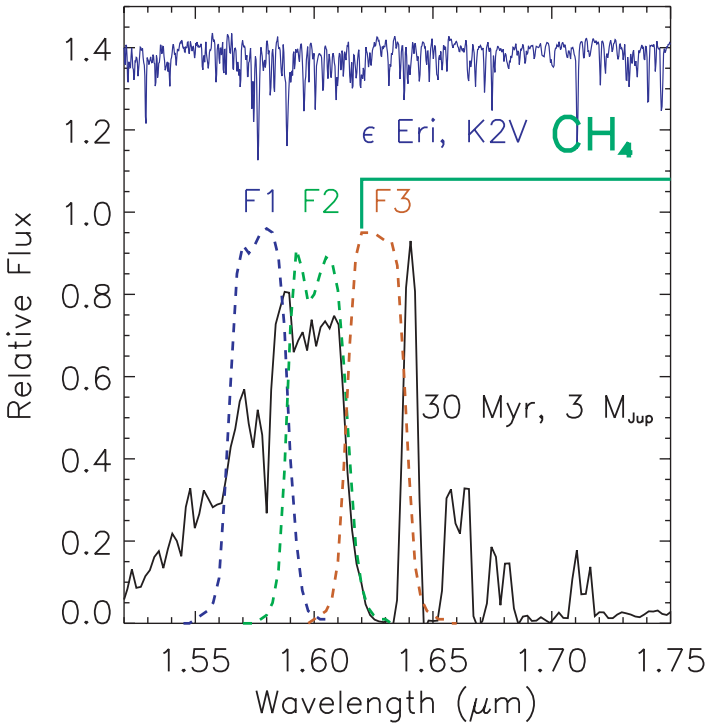


FIGURE 1.9: SDI filter transmission curves with an overlaid theoretical spectrum of a modelled 30 Million year old 3 Jupiter mass extrasolar planet (Biller et al., 2007). The F1 and F2 filters observer the 'off' 1.62 μm methane absorption feature, while the F3 filter samples directly 'on' the absorption feature. As a companion the spectrum of the K2 V star ϵ Eri (Meyer et al., 1998) is shown to be flat across this wavelength band. The subtraction of the 'on' and 'off' methane absorption images will reduce the star and quasi-static speckle noise without cancelling the signal of the companion.

Using this method speckle noise attenuation on the order of a factor of ~ 2 was reported Marois et al. (2005). Biller et al. (2007) conducted a survey of forty five nearby stars utilising this double difference SDI method. They achieved high H-band contrasts of greater than ten magnitudes at a separation of 1" for 45% of their data.

Another differential imaging approach to reduce quasi-static speckle noise was proposed by Marois et al. (2006), which is similar to the roll deconvolution technique (Schneider & Silverstone, 2003), developed to reduce the effect of static aberrations on the Hubble Space Telescope.

This method, referred to as Angular Differential Imaging (ADI), relies upon the fact that over the course of an observation the sky rotates about the celestial north pole relative to an observer. Hence if a sequence of images are acquired on a ground-based altitude / azimuth telescope with the instrument de-rotator of the telescope turned off a faint companion signal will also appear to rotate along with the field of view. Turning off the telescope de-rotator has the added benefit of improving the stability of the quasi-static PSF structure.

The ADI approach aims to reduce quasi-static speckle noise in two steps (Marois et al., 2006), see figure (1.10):

1. Reduction of correlated speckles by subtraction of a reference PSF,
2. Averaging of the residual uncorrelated noise by combination of the residual images after the field of view of each from have been aligned.

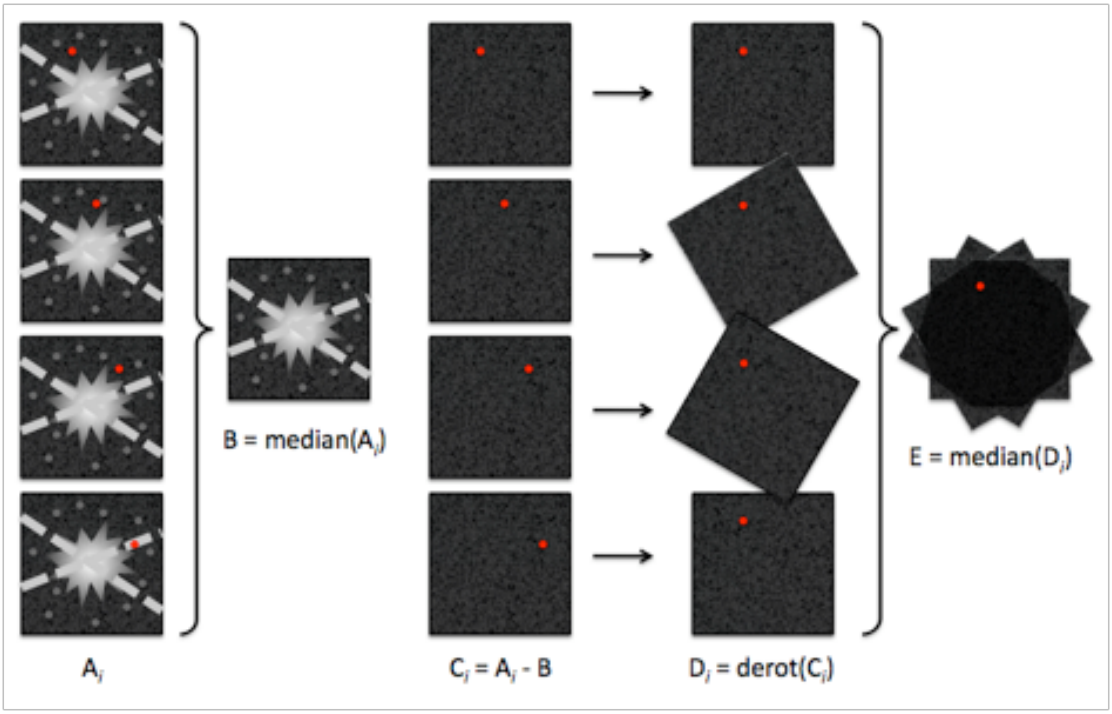


FIGURE 1.10: The ADI algorithm carries out two primary operations in the data: first a reference PSF is constructed by taking the median of the image sequence, this reference PSF is then subtracted from each individual frame in the sequence, the residual data is then derotated and medianly combined (Thalmann, 2009).

The reference PSF is created from the acquired sequence of images by taking the mean of a number of images for which the faint companion has moved significantly, but the PSF has not changed. The level of noise reduction is a function of the angular separation of the companion and the parent star, the number of reference images available to the algorithm and the image exposure time. Whilst imaging on Altair NIRI (Hodapp et al., 2003) at the 8m Gemini telescope a quasi-speckle noise attenuation of a factor of ~ 100 was reported while observing Vega.

CHAPTER 2

OPTIMAL DETECTION AND CHARACTERISATION OF FAINT COMPANIONS

This chapter introduces the subject of statistical signal decision theory. The decision of interest in this thesis is the detection of a faint companion buried in the intensity halo of its parent star. The receiver operating characteristic curve will be defined using a binary decision model as a method of quantifying the task performance of an imaging system and decision making algorithm. The area under the receiver operating characteristic curve will be used as a measure of the performance of decision making algorithms. The optimal linear observer, i.e. the Hotelling observer, will be rigorously derived from the ideal observer using the assumption that Gaussian statistics apply to the data. Furthermore the Hotelling estimator for differential astrometry and photometry will be presented.

2.1 Statistical Decision Theory

In general terms the objective, or task, of an imaging system is to gather information about the object which produced an observed image. This objective can be broken down into two types: classification and estimation. The goal of a classification task is to label the object, that is to assign the object to a particular class to which it belongs. The purpose of an estimation task is to extract useful numerical information from the image in question.

Science images are obtained for a distinct reason or purpose. The science task itself can define how well the task can be carried out with the imaging system. When a computer

algorithm and not human observers are used to carry out a task it is possible to define an ideal observer, which given the input data, e.g. an image of a binary star and information about the statistics of the imaging process, can achieve the 'best' possible performance for that task. Assessment of the task fulfilment will thus provide a measurement of the performance of the imaging system and a method of defining the quality of the images acquired for the task of interest. This type of objective assessment of image quality plays an important role in medical imaging (Barrett & Myers, 2004) and has recently been applied to the field of astronomy (Barrett et al., 2006; Burke et al., 2009).

2.1.1 Binary Decision Model

The goal of a classification task in faint companion detection is to determine whether or not a companion is present in an image. This is a binary classification problem (Barrett et al., 1998) or a two-alternative forced-choice (2AFC) test. The process of image formation is always corrupted by noise and due to the random nature of this noise the process is best described statistically. An example of a general classification task is shown in a flow chart in figure (2.1).

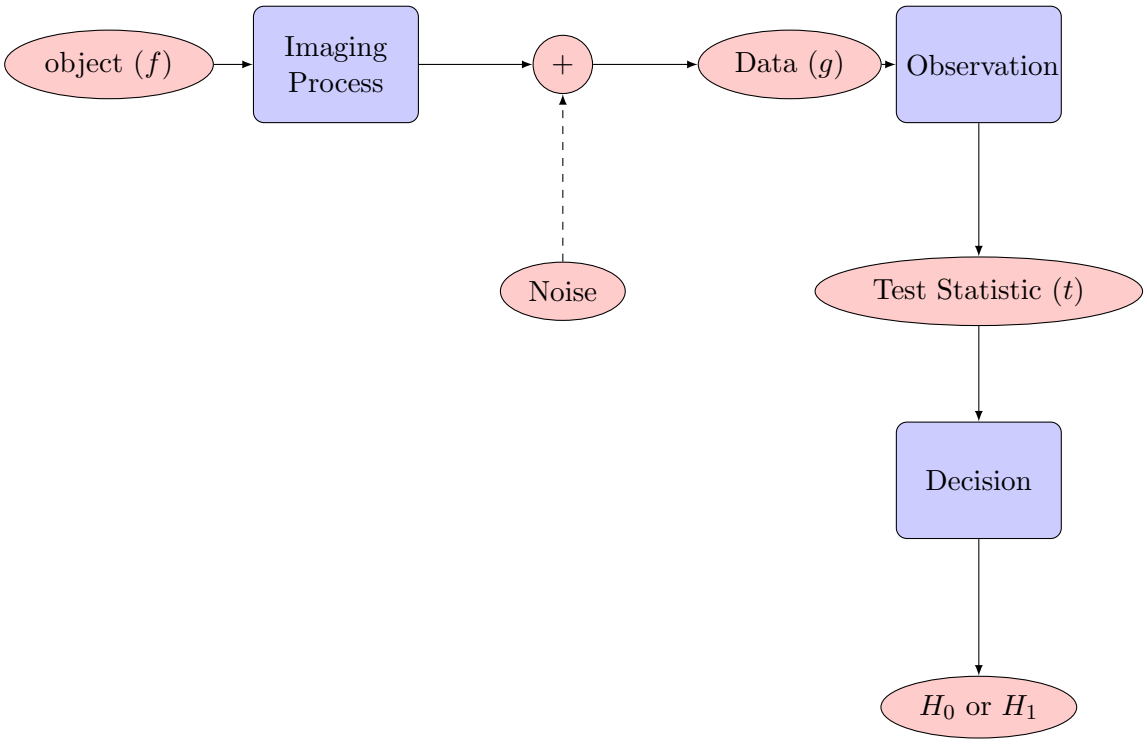


FIGURE 2.1: Flow chart representing a classification task

For each image, or data, under consideration we must decide between two hypotheses, H_0 (companion absent) or H_1 (companion present). Each image must be assigned into one

of two decision spaces Γ_0 or Γ_1. There cannot be any equivocation in the decision process which we will now describe. A scalar t, referred to as the test statistic, is used to make the classification decision. A discriminant function is applied to the data to compute the test statistic: this is the observation. The discriminant function or observer then compares the test statistic, t, to a decision threshold, t_c and hence assigns the data, g, to a particular decision space. When the test statistic is less than the decision threshold, $t < t_c$, the data is assigned to the H_0, companion absent, hypothesis, when the test statistic is greater than or equal to the decision threshold, $t \geq t_c$, the H_1 hypothesis is chosen. This decision process is never perfect because data is always corrupted by noise. For a 2AFC test there are four possible outcomes to this hit-miss detection task, as shown in table (2.1).

Observer Response	**Signal Present**	**Signal Absent**
YES	"hit" or **True Positive (TP)**	"false alarm" or **False Positive (FP)**
NO	"miss" or **False Negative (FN)**	"correct rejection" or **True Negative (TN)**

TABLE 2.1: Possible Answers for a hit-miss classification task

For a classification task the probability of a discriminant function choosing a false positive and the probability of choosing a true positive can be related to each other as points on the receiver-operating characteristic (ROC) curve.

2.1.2 The ROC, LROC and EROC curves

It has been shown (Barrett & Myers, 2004) that the performance of an observer in a binary decision task can be fully described by two fractions: the *true positive fraction* (TPF) and the *false positive fraction* (FPF). Let N be the total number of decisions made by an observer. Denote the number of true positive decisions by N_{TP} and so on for the other possible decisions such that $N = N_{TP} + N_{FP} + N_{TN} + N_{FN}$. If the number of decisions made by the observer approaches infinity then the TPF and FPF can be defined as:

$$\begin{aligned} TPF &= \left\{ \frac{N_{TP}}{N_{TP} + N_{FN}} \right\} = \lim_{N \to \infty} \left[\frac{\text{number of TP decisions}}{\text{actual number of positive cases}} \right], \\ FPF &= \left\{ \frac{N_{FP}}{N_{TN} + N_{FP}} \right\} = \lim_{N \to \infty} \left[\frac{\text{number of FP decisions}}{\text{actual number of negative cases}} \right]. \end{aligned} \tag{2.1}$$

It is useful to consider the probability density functions of the scalar test statistic, t, under the hypotheses H_0 and H_1, i.e. $pr(t|H_0)$ and $pr(t|H_1)$. From these probability density functions the TPF and FPF as a function of some threshold t_c can be formally defined:

$$\begin{aligned} TPF(t_c) &= \int_{t_c}^{+\infty} pr(t|H_1)dt, \\ FPF(t_c) &= \int_{t_c}^{+\infty} pr(t|H_0)dt. \end{aligned} \tag{2.2}$$

By varying the decision threshold, t_c, a plot showing the relationship between the TPF and the FPF can be mapped out. This plot known as the receiver operating characteristic curve is shown in figure (2.2).

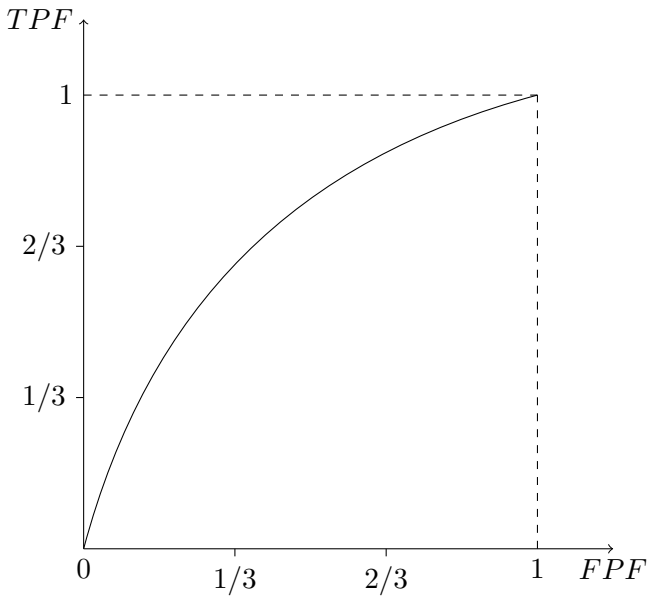


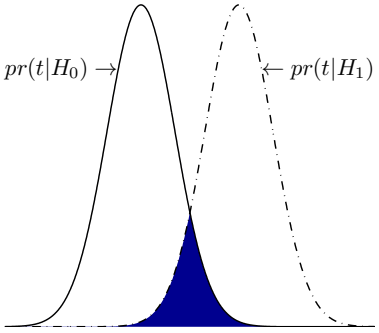
FIGURE 2.2: A sample ROC curve.

The area under the ROC curve (AUC) can be used as a figure of merit on the performance of the observer. The AUC is an indication of how well the densities $pr(t|H_0)$ and $pr(t|H_1)$ are separated from each other. The AUC can be defined in terms of the TPF and FPF as the decision threshold is varied:

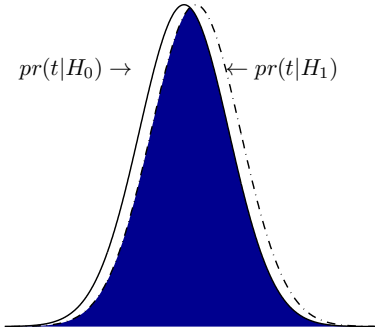
$$AUC = -\int_{-\infty}^{+\infty} TPF(t_c)\frac{dFPF(t_c)}{dt_c}dt_c. \tag{2.3}$$

When the densities $pr(t|H_0)$ and $pr(t|H_1)$ are well separated, see figure(2.3(a)), the AUC will be close to 1. However when the densities almost overlap the probability of making

a correct decision will almost equal the probability of making an incorrect decision and so the AUC will be close to 0.5, see figure(2.3(b)).



(a) Well Separated Test Statistic Density Functions



(b) Poorly Separated Test Statistic Density Functions

FIGURE 2.3: Test statistic density functions for a well performing decision making system (a) and a poorly performing system (b).

A classification task can be combined with an estimation task in the decision making process. An example would be faint companion detection where the location of the companion is unknown. In this case the observer will also return an extra parameter which represents an estimate on the location of the signal. The TPF is now the addition of the probability of correctly classifying an image and correctly localising the signal of the companion to within a given tolerance. A plot of the TPF vs. FPF for a detection-localisation task is referred to as a localisation ROC (LROC) curve (Barrett & Myers, 2004; Caucci et al., 2007; Khurd & Gindo, 2005). The LROC curve can be generalised into a detection-estimation task (Clarkson, 2007). For example an observer could be tasked with detecting a faint companion and estimating the position and intensity of the signal in an image. The plot of the TPF vs. FPF is now referred to as an estimation

ROC (EROC) curve. The binary correct detection-localisation function, which is used in LROC analysis, is replaced by a general utility function which measures the usefulness of a parameter estimate. The AUC is still used as the figure of merit for the LROC and EROC curves.

2.2 The Ideal Observer

The AUC has been defined as the figure of merit for classification-estimation tasks and provides a metric for the optimisation of the performance of the the observer used to carry out these tasks. For all discrimination tasks an ideal observer is that which maximises the AUC for a given task. It can be shown that the optimal discriminant function is the likelihood ratio (Barrett & Myers, 2004):

$$\Lambda(g) = \frac{pr(g|H_1)}{pr(g|H_0)}, \tag{2.4}$$

or its logarithm $\lambda(g) = ln[\Lambda(g)]$ (Barrett et al., 1998). To compute the likelihood ratio the probability density functions under both hypotheses need to be known and this requires descriptions of the objects to be classified-estimated and complete information regarding the data measurement processes and full knowledge of the statistics of the noise in the data.

2.3 The Hotelling Observer

In practice the likelihood ratio cannot always be calculated. It is then desirable to use a linear observer which also maximises the area under the ROC curve and takes the form:

$$t(g) = w^t g, \tag{2.5}$$

where g is the data and w is a template vector sought after such that it will maximise $t(g)$ and hence the AUC. Maximising the AUC is equivalent to maximising the class separation of $pr(t|H_0)$ and $pr(t|H_1)$ as shown in figure (2.3). It is convenient to introduce the nomenclature corresponding to faint companion detection and parameter estimation. It is assumed that the AO corrected astronomical data has already been pre-processed, so that it is flat-fielded and background-subtracted. Also, if the observation consists of

multiple images, we assume that these frames have been re-centred and co-added. The observer is now supplied with a pre-processed long exposure image.

Let $\overline{g}_0$ denote the mean image when a companion is absent where the over-bar refers to the mean of the data averaged over the measurement noise. When a companion is present, at the position r_{pl}, the mean image is written as $\overline{g}_{1,r_{pl}}$. The true signal of the companion is then given by:

$$s_{r_{pl}} = \overline{g}_{1,r_{pl}} - \overline{g}_0. \tag{2.6}$$

In real life situations astronomical images are contaminated by several types of noise, such as: cosmic ray hits, bad pixels on the detector, readout noise from the detector and Poisson noise due to the stochastic nature of photon detection. Bad pixels and cosmic ray hits can usually be compensated for by applying a median filter to the data (Artigau et al., 2008). The noisy images can be modelled under the two hypotheses as:

$$H_0 : g = \overline{g}_0 + n, \qquad H_{1,r_{pl}} : g = \underbrace{\overline{g}_0 + s_{r_{pl}}}_{\overline{g}_{1,r_{pl}}} + n. \tag{2.7}$$

The noise n is composed of Gaussian noise from the detector readout and Poisson noise arising from the detection of the incident radiation.

In the following we assume that the data with dimensions of $n \times n$ has been rearranged into column vectors of size $M \times 1$, where $M = n \times n$. This rearrangement makes it possible to use normal matrix operations on the data. The images averaged over noise can now be expressed as:

$$\overline{g}_0 = \underbrace{A_* h(r_*)}_{star\ image} + \underbrace{b}_{background}, \qquad \overline{g}_{1,r_{pl}} = \underbrace{A_* h(r_*)}_{star\ image} + \underbrace{A_{pl} h(r_{pl})}_{companion\ image} + \underbrace{b}_{background}, \tag{2.8}$$

where A_* is the intensity of the bright star located at position r_*, A_{pl} is the intensity of the companion located at r_{pl}, $h(r)$ is the PSF located at r and b is the background intensity.

While the intensity in AO-corrected short exposures can be approximated by a modified Rician distribution (Cagigal & Canales, 1998; Aime & Soummer, 2004), the sum of these intensities, over many exposures, can be approximated by a Gaussian distribution by the central limit theorem. Therefore we expect the intensity statistics in an ensemble of

long-exposure images to obey Gaussian statistics. I have also assumed that additional sources of noise such as quasi-static speckle noise have also been suppressed. We will see in Chapters (5, 6) how this may be attempted. Under the Gaussian assumption the densities $pr(g|H_0)$ and $pr(g|H_{1,r_{pl}})$ can then be expressed as (Helstrom, 1968; Caucci et al., 2007)

$$pr(g|H_0) = \left[\frac{1}{(2\pi)^M det(K_g)}\right]^{1/2} \times exp\left[-\frac{1}{2}(g-\overline{g}_0)^T K_g^{-1}(g-\overline{g}_0)\right], \tag{2.9}$$

$$pr(g|H_{1,r_{pl}}) = \left[\frac{1}{(2\pi)^M det(K_g)}\right]^{1/2} \times exp\left[-\frac{1}{2}(g-\overline{g}_0-s_{r_{pl}})^T K_g^{-1}(g-\overline{g}_0-s_{r_{pl}})\right]. \tag{2.10}$$

Where *det* denotes the determinant of a given matrix, ^{T} denotes transpose, $pr(g|H)$ is the conditional probability density function of the data under the hypothesis H, either H_0 or $H_{1,r_{pl}}$, and K_g is the covariance matrix of the data, of size $M \times M$.

The noise in the data is assumed to be uncorrelated i.e. the noise in one pixel is independent of the noise in other pixels. This is a reasonable assumption for the case of detector and photon noise. With the assumption of uncorrelated noise, the data covariance matrix, K_g, is diagonal, with elements given by (Caucci et al., 2007):

$$[K_g]_{m,m'} = [A_* h_m(r_*) + b_m + \sigma_m^2]\delta_{m,m'} \tag{2.11}$$

where $\delta_{m,m'}$ is the Kronecker delta function, the detector noise has been assumed to be Gaussian with variance σ_m^2, the background follows Poisson statistics and m represents the pixel index. This expression for K_g ignores the small addition to the covariance matrix due to the Poisson noise from the companion.

If the companion location, r_{pl}, is unknown, the ideal observer can be applied at a set of test locations R and is referred to as the ideal scanning observer (Caucci et al., 2007). This observer has a maximum test statistic at the true companion location, $r_{pl} \in R$. This means that the algorithm can be used to estimate differential astrometry of a companion. The ideal scanning observer takes on the form:

$$\Lambda(g) = \max_{r_{pl} \in R} \Lambda(g|r_{pl}) = \max_{r_{pl} \in R} \frac{pr(g|H_{1,r_{pl}})}{pr(g|H_0)}. \tag{2.12}$$

With the estimation of r_{pl} then computed as:

$$\widetilde{r}_{pl} = \arg\max_{r_{pl}\in R} \Lambda(g|r_{pl}). \tag{2.13}$$

Under the assumption of Gaussian noise (equations (2.9) and (2.10)) and taking the logarithm of $\Lambda(g)$ we obtain the log-likelihood ratio, $\lambda(g) = ln[\Lambda(g)]$, given by:

$$\begin{aligned}\lambda(g) &= \max_{r_{pl}\in R}\left[\frac{1}{2}(g-\overline{g}_0)^T K_g^{-1}(g-\overline{g}_0) - \frac{1}{2}(g-\overline{g}_0-s_{r_{pl}})^T K_g^{-1}(g-\overline{g}_0-s_{r_{pl}})\right], \\ &= \max_{r_{pl}\in R} -\frac{1}{2}\Big[-g^T K_g^{-1}(g-\overline{g}_0) + \overline{g}_0^T K_g^{-1}(g-\overline{g}_0) + g^T K_g^{-1}(g-\overline{g}_0-s_{r_{pl}}) \\ &\qquad - \overline{g}_0^T K_g^{-1}(g-\overline{g}_0-s_{r_{pl}}) - s_{r_{pl}} K_g^{-1}(g-\overline{g}_0-s_{r_{pl}})\Big].\end{aligned} \tag{2.14}$$

If K_g^{-1} is diagonal, as we assumed earlier (equation (2.11)) then the log-likelihood ratio becomes:

$$\begin{aligned}\lambda(g) &= \max_{r_{pl}\in R} -\frac{1}{2}\left[-g^T K_g^{-1} s_{r_{pl}} + \overline{g}_0^T K_g^{-1} s_{r_{pl}} - s_{r_{pl}}^T K_g^{-1} g + s_{r_{pl}}^T K_g^{-1}\overline{g}_0 + s_{r_{pl}}^T K_g^{-1} s_{r_{pl}}\right] \\ &= \max_{r_{pl}\in R} -\frac{1}{2}\left[2 s_{r_{pl}}^T K_g^{-1}(\overline{g}_0 - g) + s_{r_{pl}}^T K_g^{-1} s_{r_{pl}}\right], \\ &= \max_{r_{pl}\in R}\left[s_{r_{pl}}^T K_g^{-1}(g - \overline{g}_0 - \frac{1}{2}s_{r_{pl}})\right].\end{aligned} \tag{2.15}$$

The quantity that appears in square brackets in equation (2.15) is linear in g. The log-likelihood ratio, $\lambda(g)$, is therefore the ideal linear observer. Barrett et al. (2006) shows that the Hotelling observer, $t(g)$, is equal to the log-likelihood ratio if the data is normally distributed with equal covariances under both hypotheses i.e. $K_{g|H_1} = K_{g|H_0}$. In equation (2.15) the signal of the bright parent star and half of the predicted companion signal is removed from the data. The subtraction of half of the predicted companion signal does not affect the location of the maximum of $\lambda(g)$. The result is prewhitened, i.e. divided by the data covariance matrix K_g. The output of this operation is then processed with a matched filter using the expected companion signal located at the expected companion position, r_{pl}. Data prewhitening and matched filtering is not unique to the Hotelling observer; the generalised Wiener-Helstrom filter (Helstrom, 1968; Barrett

et al., 1995; Barrett et al., 2006) also uses second order statistics of the object, and forms its output by a linear operation on prewhitened data. The Wiener-Helstrom filter is biased towards a known a priori mean, when the number of unknown parameters is large this biasing can be very useful, i.e. the produced solution is a plausible one.

Following on from equation (2.15) and recalling from equation (2.11) the expression for $K_g = A_* h_m(r_*) + b_m + \sigma_m^2$ and $s_{r_{pl}} = \overline{g}_{1,r_{pl}} - \overline{g}_0 = A_{pl} h_m(r_{pl})$, an expression for the scanning Hotelling observer, $t_{Hot}(g)$, can be written in the form:

$$t_{Hot}(g) = \sum_{m=1}^{M} \frac{A_{pl} h_m(r_{pl})}{A_* h_m(r_*) + b_m + \sigma_m^2} \left[g_m - \overline{g}_{0,m} - \frac{1}{2} A_{pl} h_m(r_{pl}) \right] \tag{2.16}$$

The estimation of the companion location, r_{pl}, is then computed as:

$$\widetilde{r}_{pl} = \arg \max_{r_{pl} \in R} [t_{Hot}(g)] \tag{2.17}$$

In the following, equation (2.17) is referred to as the Spatial Scanning Hotelling Estimator (SSHE). The operator of the algorithm computes a set of scalars on the data g. These scalars are functions of the unknown position of the companion r_{pl}. The maximum of this set of scalars is taken as the value of the Hotelling observer, $t_{Hot}(g)$. In the above derivation of the SSHE it was assumed that the intensity of the companion A_{pl} was known. In general for faint companions this will not be the case. However the Hotelling method can be generalised to estimate both the position and intensity of a faint companion at the same time. The maximum of equation (2.16) is now dependent upon the Cartesian product of the two unknown parameters of the companion signal: r_{pl} and A_{pl}. Hence for every test companion location in R a set of test companion intensities is defined: A. Scanning over this three dimensional test space results in a maximum value of the Hotelling test statistic with a corresponding estimate of the companion position and intensity. However this is a computationally intensive approach.

The approach that follows relies on the fact that for every companion location estimate there exists a companion intensity estimate which maximises the Hotelling test statistic for that location, given that $t_{Hot}(g)$ is quadratic in A_{pl}. Recalling the log-likelihood ratio, equation (2.15), an estimate of the intensity of the companion, A_{pl}, for a defined set of test intensities A becomes:

$$\widetilde{A}_{pl} = \arg \max_{A_{pl} \in A} \left[s_{r_{pl}}^T K_g^{-1} (g - \overline{g}_0) - \frac{1}{2} s_{r_{pl}}^T K_g^{-1} s_{r_{pl}} \right]. \tag{2.18}$$

Equation (2.18) assumed r_{pl} is known. The log-likelihood ratio is maximised by taking the partial derivative of equation (2.18) with respect to A_{pl} and setting it equal to zero:

$$\frac{\partial t_{Hot}(g)}{\partial A_{pl}} = h(r_{pl})^T K_g^{-1}(g - \overline{g}_0) - A_{pl} h(r_{pl})^T K_g^{-1} h_m(r_{pl}) = 0, \tag{2.19}$$

where use was made of the fact that $s_{r_{pl}} = A_{pl} h(r_{pl})$. This leads to the following estimator for A_{pl}:

$$\widetilde{A}_{pl} = \frac{h(r_{pl})^T K_g^{-1}(g - \overline{g}_0)}{h(r_{pl})^T K_g^{-1} h(r_{pl})}. \tag{2.20}$$

Recalling the structure of the data covariance matrix K_g this estimator becomes:

$$\widetilde{A}_{pl} = \frac{\sum_{m=1}^{M} (h_m(r_{pl})/(A_* h_m(r_*) + b_m + \sigma_m^2))[g_m - \overline{g}_{0,m}]}{\sum_{m=1}^{M} h_m(r_{pl})^2/(A_* h_m(r_*) + b_m + \sigma_m^2)}, \tag{2.21}$$

where m is the pixel index of the image. We refer to equation (2.21) as the Optimal Hotelling Estimator (OHE). The estimation of the companion location is computed using the SSHE whilst substituting the estimation of A_{pl} from the OHE into the SSHE for each test location. The result is differential astrometry and photometry of an observed companion. Various methods of implementing this algorithm are explored in section (3.3).

The mathematical form of the OHE can be compared to that of the Iterative Blind Deconvolution estimator for intensity (Barnaby et al., 2000). This algorithm uses a conjugate gradient minimisation to minimise the error metric:

$$E_f = \sum_m \left[g_m - (\widetilde{o}_m * \widetilde{h}_m) \right]^2. \tag{2.22}$$

Where o represents the object (in the focal plane), h is an estimate of the PSF and g is the observed data. The object, o, can be divided into two signals (equation (4.1)) , the bright star and the companion signal:

$$E_f = \sum_m \left[g_m - A_* \delta(x - x_*, y - y_*) * \widetilde{h}_m - A_{pl} \delta(x - x_{pl}, y - y_{pl}) * \widetilde{h}_m \right]^2. \tag{2.23}$$

As above, instead of using a gradient minimisation, an expression for this minimum can be found by taking the partial derivative of E_f with respect to A_{pl} and setting it equal to zero.

$$\frac{\partial E_f}{\partial A_{pl}} = 2\sum_m \left[g_m - A_* \delta(x-x_*, y-y_*) * \widetilde{h}_m - A_{pl}\delta(x-x_{pl}, y-y_{pl}) * \widetilde{h}_m \right] \left[\delta(x-x_{pl}, y-y_{pl}) * \widetilde{h}_m \right] = 0. \tag{2.24}$$

This leads to the estimation of A_{pl}:

$$\widetilde{A}_{pl} = \sum_m \frac{[g_m - A_*(\delta(x-x_*, y-y_*)) * \widetilde{h}_m][\delta(x-x_{pl}, y-y_{pl}) * \widetilde{h}_m]}{[\delta(x-x_{pl}, y-y_{pl}) * \widetilde{h}_m][\delta(x-x_{pl}, y-y_{pl}) * \widetilde{h}_m]} \tag{2.25}$$

It is instructive to compare equations (2.21) and (2.25). The difference is that the OHE takes into account the variance of the noise in the data i.e. the OHE prewhitens the data. The importance of prewhitening will be shown in section (4.3).

CHAPTER 3

APPLICATION OF THE HOTELLING OBSERVER

In this chapter, we report upon the details of simulation work carried out to validate the Hotelling approach. The data simulation package PAOLA which was used to simulate long exposure AO corrected images is described. The second part of the chapter concentrates on faint companion detection at a known location, where the Hotelling observer is compared to other state of the art linear observers. The third section carries out experiments involving faint companion detection at an unknown location. For this task several techniques are introduced to calculate the Hotelling discriminant. Results are presented where the presence, location and intensity of a faint companion are required to be determined. Finally the Hotelling observer is applied to the task of detecting and locating Shack-Hartmann spot images in the presence of strong turbulence.

3.1 Data Simulation with PAOLA

PAOLA (Performance of Adaptive Optics for Large or Little Apertures) (Jolissaint et al., 2006) is an IDL-based analytic end-to-end AO simulation code which provides fast first order system performance estimates. This approach is based upon the derivation of Rigaut et al. (1998) which shows that a good approximation of the AO corrected long exposure optical transfer function can be found by modelling the AO corrected phase spatial power spectrum. The AO model is a Shack-Hartmann wavefront sensor with a least-square type reconstructor. The compensated phase is modelled by the addition of five error terms:

1. **Fitting Error** due to the deformable mirror being unable to compensate the incoming phase above its spatial cut-off frequency.

2. **Anisoplanatic Error** arising from the system attempting to compensate an incoming phase at a different angle from the analysed phase.

3. **Servo Lag Error** due to the time delay in applying the estimated wavefront correction to the incoming phase.

4. **Spatial Aliasing Error** in the wavefront sensor from phase spatial frequencies above the wavefront sensor cut-off frequency being sensed as low spatial frequencies.

5. **Additive Noise** in the wavefront sensor from detector readout and photon noise.

Analytical expressions are used to obtain an estimate of the residual phase power spectrum, S_{φ_c}, where φ_c denotes the compensated phase the phase structure function, D_ϕ can be expressed as:

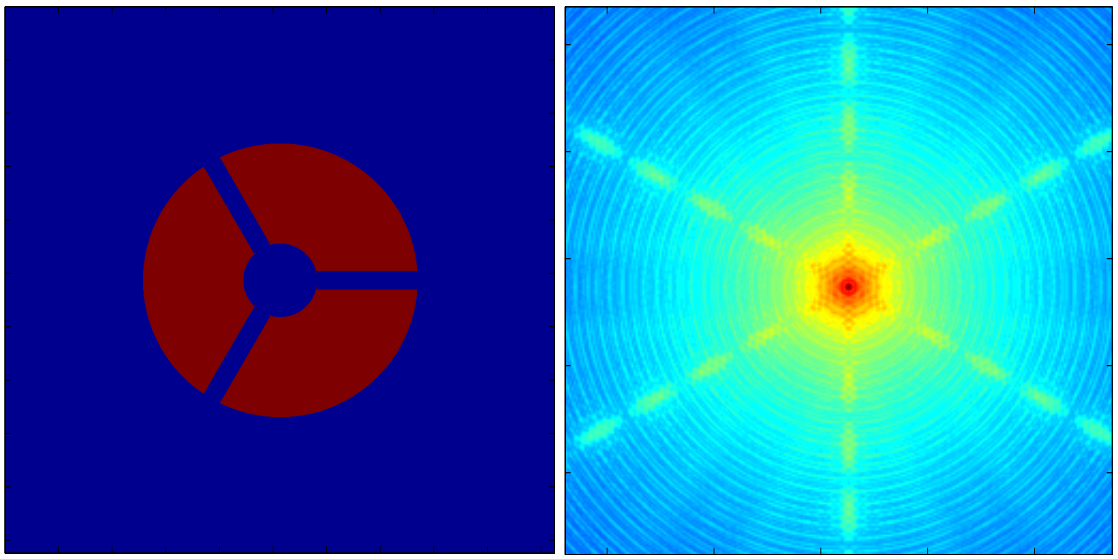
$$D_\varphi(\rho) = 2 \int \int_{-\infty}^{\infty} S_{\varphi_c}(k) \left[1 - cos\left(2\pi i k.\rho\right)\right] dk, \tag{3.1}$$

where ρ is the displacement vector, (x, y), between two points in the phase and k is the spatial frequency vector. The OTF of the system can hence be written as:

$$OTF(\rho/\lambda) = T(\rho/\lambda).exp\left[-\frac{1}{2} D_\varphi(\rho/\lambda)\right], \tag{3.2}$$

where λ is the imaging wavelength and T is the OTF of the telescope in the absence of turbulence. The long exposure PSF can be estimated by the inverse Fourier transform of the OTF.

The PAOLA package was chosen to simulate the data primarily due to its rapid execution time i.e. a few minutes for an eight metre class telescope compared to several hours for a Monte Carlo type algorithm (Jolissaint et al., 2006). The parameters of the AO system on the Lick Observatory $3m$ Shane telescope (Bauman et al., 1999; Fitzgerald & Graham, 2006; Lloyd et al., 2000) were used to simulate the data. In the simulations the deformable mirror of the AO system had 5 actuators across its diameter. The simulations were carried out in the K band ($2.2\mu m$) where the primary star was modelled as a $m_k = 5$ star and the natural guide star as a bolometric magnitude 10 star. The sky brightness was set at 19 magnitudes arcsec^{-2}. The atmosphere was modelled as consisting of three turbulent layers at altitudes of 1, 5 and 10 km, with wind speeds of 10, 15 and 25 ms^{-1}



(a) Simulated Lick Pupil (b) Estimated Lick PSF

FIGURE 3.1: Image of the simulated Lick pupil (a) and a log-scale image of the corresponding estimated PSF from PAOLA (b).

respectively. The seeing angle at $0.55\mu m$ was 0.8 arcsec. The optical transmission of the system in the K-band was set at 13%, the science filter bandpass at $0.32\mu m$, the quantum efficiency of the science camera at 67%, the gain of the science camera at 10 and an exposure time of $22ms$. These parameters correspond to the imaging campaign carried out at the Lick Observatory (Gladysz et al., 2006). Figure (3.1(a)) shows the modelled pupil of the Lick system and a corresponding simulated PSF, figure (3.1(b)).

3.2 Robustness of the Hotelling Observer

This section reports on applying the Hotelling Observer on a variety of different simulated data to perform a detection task. The area under the receiver operating characteristic curve has been used as the figure of merit for all observers in the determination of task performance. For the following experiments the observers were given complete knowledge of the signal to be detected, i.e. the position and intensity of the companion, the point spread function of the simulated AO system and the image background intensity and variance. This special case of signal detection is referred to as the "signal known exactly, point spread function known exactly and background known exactly" (SKE/PKE/BKE) case and will be used to assess the performance of the Hotelling observer.

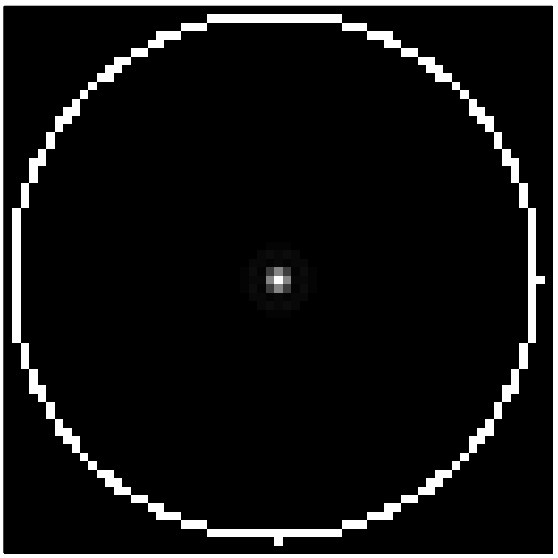


FIGURE 3.2: The white circle denotes the location of pixels which were used to estimate the background level and the variance of the detector readout noise.

Only the presence of the companion is left unknown to the observer. This case may not be as trivial as it first appears.

The intensity of the central star can be estimated using the known PSF in a least-squares algorithm. The variance of the detector readout (plus background) can be taken as the variance within an annulus of pixel values far away from the central star such that it is assumed the annulus only contains background information (Burke et al., 2009). For an example see figure(3.2) where variance of the detector readout, plus background, is estimated for a simulated Lick PSF. It is possible to estimate the long exposure PSF by several means: observing a reference star, reconstruction from wavefront sensor data (Véran et al., 1997), reconstruction from Angular Differential Imaging data (Marois et al., 2006), Spectral Differential Imaging data (Racine et al., 1999) or from focal plane data (Gonsalves, 1982), also see Chapter (6).

Recalling the notation from section (2.3) we denote: A_* as the intensity of the bright star located at position r_*, A_{pl} as the intensity of the companion located at r_{pl}, $h(r)$ as the simulated AO corrected PSF located at position r, b is the background intensity with a variance σ^2 and m is the pixel index. The long exposure images averaged over measurement noise are modelled as:

$$\overline{g}_0 = \underbrace{A_* h(r_*)}_{star\ image} + \underbrace{b}_{background}, \qquad \overline{g}_{1,r_{pl}} = \underbrace{A_* h(r_*)}_{star\ image} + \underbrace{A_{pl} h(r_{pl})}_{companion\ image} + \underbrace{b}_{background}. \quad (3.3)$$

It should be noted that when r_{pl} and $\overline{g}_0$ are fixed, the terms $s_{r_{pl}}^T K_g^{-1} \overline{g}_0$ and $s_{r_{pl}}^T K_g^{-1} s_{r_{pl}}$ in equation (2.15) do not vary with g and can therefore be treated as constants. Hence the Hotelling Observer for faint companion detection at a known location can be expressed as:

$$t_{Hot}(g|r_{pl}) = \sum_{m=1}^{M} \frac{A_{pl} h_m(r_{pl})}{A_* h_m(r_*) + b_m + \sigma_m^2} g_m. \tag{3.4}$$

It needs to be stressed that for this SKE/PKE/BKE example an exact estimate of A_{pl} is not needed. This is because $t_{Hot}(g|r_{pl})$ scales with A_{pl} and therefore this quantity can be factored into the threshold which the Hotelling Observer will be compared to.

As the Hotelling observer is a linear observer it was desirable to compare its performance to other linear observers. As the PSF and location of the companion signal are known, a straight matched-filter (Barrett & Myers, 2004; Poyneer, 2003) type observer was chosen as an initial comparison. The matched-filter employed in this study is simply the scalar product of the proposed companion signal i.e. a template vector and the data. This product can be thought of as a correlation (Barrett & Myers, 2004) and hence can be computed using Fourier methods. Given the PSF located at the companion position and the intensity of the companion the matched-filter observer is given by:

$$t_{MF}(g|r_{pl}) = \sum_{m}^{M} \underbrace{A_{pl} h_m(r_{pl})}_{\text{template vector}} g_m. \tag{3.5}$$

The formulation of this observer is very similar to the Hotelling observer. However in this observer the data vector has not been prewhitened. The performance of this observer is primarily influenced by how closely the PSF in the template vector matches the PSF of the data. This matched-filter is a convenient observer to compare the Hotelling observer to as no additional parameters beyond those required for the computation of the Hotelling observer are needed.

The second linear observer chosen is formulated around subtracting the signal of the primary star from the data and summing over a window. If k is the size of the window centred at r_{pl}, then the set of pixels inside the window are given as:

$$N(r_{pl}, k) \in \{1, \ldots, M\}, \tag{3.6}$$

where the window is $k \times k$ square. Having defined the window the star-subtraction observer can be defined as:

$$t_{sub}(g|r_{pl}) = \sum_{m \in N(r_{pl},k)} \left[g_m - b_m - A_* h_m(r_*) \right]. \tag{3.7}$$

Two sets of experiments will be reported upon below. The first set of experiments investigates the relationship between the AUC of the three observers for a detection task at a known location when the parameters of the data: A_{pl}, r_{pl}, b and σ are varied. How the AUC changed with the level of AO compensation was also investigated. The second set of experiments represent the first step away from the pure SKE/PKE/BKE scenario. In these experiments the observers are supplied with mismatched parameters for the data i.e. mismatched values for: A_{pl}, r_{pl}, b, σ and the level of AO correction. These tests will show how sensitive the Hotelling observer is to the estimates of these parameters.

3.2.1 Varying Companion Brightness

Adaptive Optics corrected images of a Lick-type system of size 64×64 pixels, with a pixel size of 76 mas, were simulated using PAOLA. The Hotelling, matched-filter and star subtraction observers were used to carry out a detection task where the data had the following parameters:

- $A_* = 9 \times 10^5 e^-$(total flux), $1 \times 10^5 e^-$(peak flux), $m_k = 5$, $\tau = 22\ ms$
- $b_m = 2 \times 10^{-2} e^-$, i.e. 19 magnitudes in the K band per arcsecond,
- $\sigma_m = 5e^-$,
- $r_{pl} = 5$ pixels or 380 mas from the primary star, diffraction limit $\lambda/D = 156mas$
- $A_{pl} = 1 \rightarrow 200e^-$ i.e. $15 \rightarrow 9$ magnitudes fainter than the primary,
- $k = 3$ pixels,

where e^- are photoelectrons and m is the pixel index. For each value of the companion brightness, 300 noisy images, i.e. Poisson and Gaussian noise was added, with and without a companion present were simulated. An ROC curve was then computed for each observer and the corresponding AUC was determined. As the companion brightness was varied a curve was mapped out for each observer on the A_{pl}-AUC plane. This experiment was repeated 10 times, the average A_{pl}-AUC curves are presented in figure (3.3) along with the standard deviation of the AUC for some values of A_{pl}.

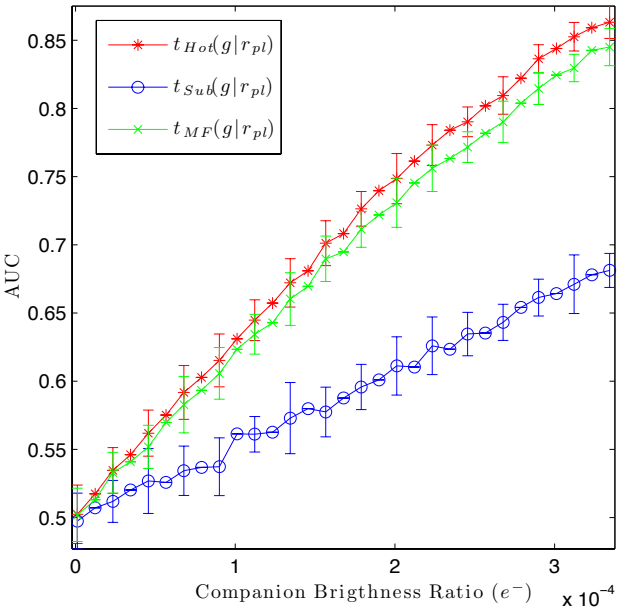


FIGURE 3.3: Plot of the A_{pl}-AUC plane for the Hotelling observer, simple matched filter and star subtraction observers.

The result of the $A_{pl}-$ AUC mapping, figure (3.3), shows that the Hotelling observer and the simple matched filter have very similar performances. However the difference in the performance of the observers may become more noticeable if the data covariance matrix is non-diagonal.

As can be seen from equation (3.4) and equation (3.5) the test statistics of the Hotelling observer and the simple matched filter scale with A_{pl}. This scaling effect appears to result in a linear relationship between the AUC and A_{pl} for these observers. The change in the AUC as a function of the companion brightness was modelled for the Hotelling observer using a linear interpolation, see figure (3.4). The correlation between the data and the result of the fit is very good with a correlation coefficient of 0.997. This result confirms the linearity of the scaling effect of the companion brightness on the AUC.

3.2.2 Increasing Companion Separation

After the primary star companion intensity ratio, the separation between the primary star and the companion is the next critical parameter to a faint companion detection task. In this experiment the data was simulated as in section (3.2.1). However the companion brightness was set at $A_{pl} = 220e^-$ and the companion separation increased from $r_{pl} = 0 \rightarrow 10$ pixels. For each companion position, 300 noisy images with and without a companion present were simulated. An ROC curve was then computed for each

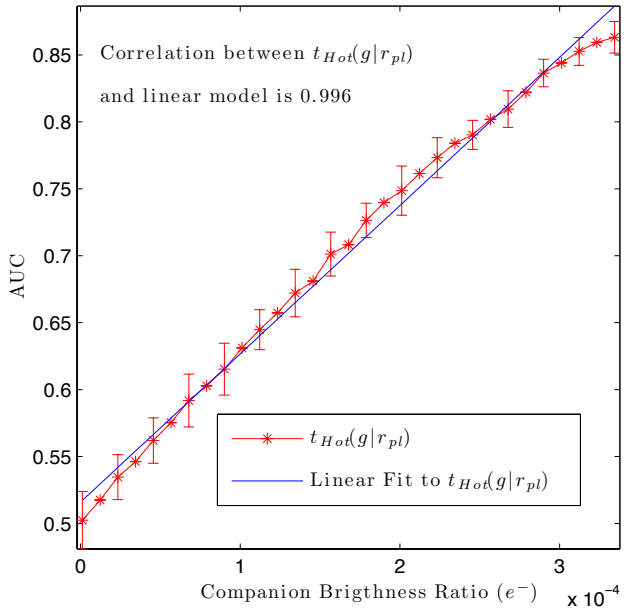


FIGURE 3.4: Investigation of the linear relationship between the companion brightness and the AUC for the Hotelling observer.

observer and the corresponding AUC was determined. As the companion separation was varied a curve was mapped out for each observer on the r_{pl}-AUC plane. This experiment was repeated 10 times, the average r_{pl}-AUC curves are presented in figure (3.5) along with the standard deviation of the AUC for some values of r_{pl}.

Figure (3.5) shows that the Hotelling observer outperforms the simple matched filter and star subtraction observers when the companion is close to the central star. However the Hotelling and simple matched filter observers perform almost identically far away from the primary star. This result stems from the fact that when the data covariance matrix is smooth i.e. far away from the central star, the Hotelling and simple matched filter observers will perform similarly.

3.2.3 Increased background intensity and variance of the detector readout noise

The importance of the background level and its variance on the Hotelling, matched-filter and star subtraction observers was investigated. In both of these experiments the data was simulated as in section (3.2.1), with the companion separation set to $r_{pl} = 5$ pixels and the companion intensity set to $A_{pl} = 220e^-$.

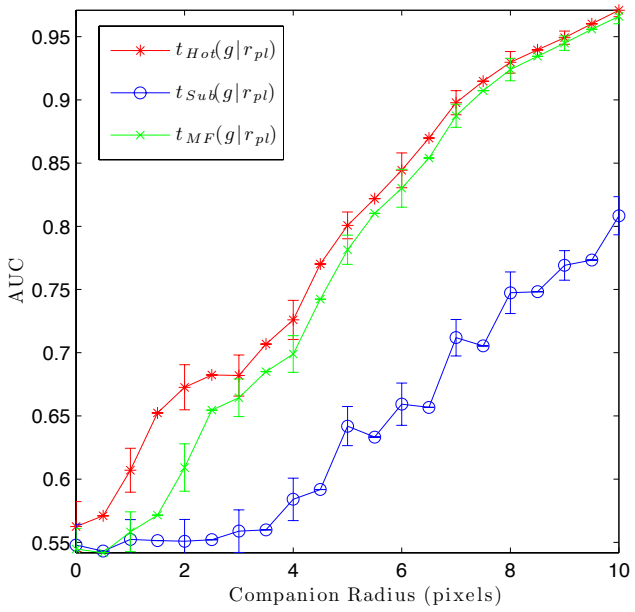
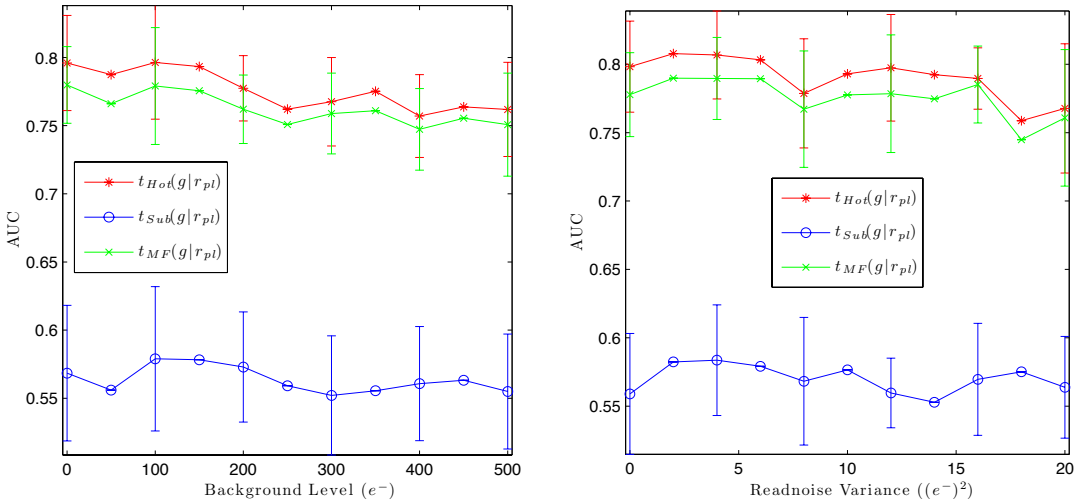


FIGURE 3.5: Plot of the r_{pl}-AUC plane for the Hotelling observer, simple matched filter and star subtraction observers.

Increasing the intensity of the uniform background, b_m, should in effect increase the level of Poisson noise. In the first test the background level was increased from $b_m = 0 \rightarrow 100e^-$ (peak value). For each value of b_m, 300 noisy images with and without a companion present were simulated. An ROC curve was then computed for each observer and the corresponding AUC was determined. As the background intensity was varied a curve was mapped out for each observer on the b_m-AUC plane. This experiment was repeated 10 times, the average b_m-AUC curves are presented in figure (3.6(a)) along with the standard deviation of the AUC for some values of b_m. The Hotelling observer outperforms the other two linear observers for all levels of the background tested here.

In a second experiment the variance of the detector readout noise was increased. The readout noise variance was increased from $\sigma_m^2 = 0 \rightarrow 20e^{-2}$. For each value of σ_m^2, 300 noisy images with and without a companion present were simulated. An ROC curve was then computed for each observer and the corresponding AUC was determined. As the readout noise variance was varied a curve was mapped out for each observer on the σ_m^2-AUC plane. This experiment was repeated 10 times, the average σ_m^2-AUC curves are presented in figure (3.6(b)) along with the standard deviation of the AUC for some values of σ_m^2. The Hotelling observer outperforms the other two linear observers for all levels of the readout noise variance tested here.

These experiments demonstrate the robustness of the Hotelling observer and simple matched filter to a wide range of background intensity levels and detector readout noise



(a) Plot of the b_m-AUC plane for the three linear observers

(b) Plot of the σ_m^2-AUC plane for the three linear observers

FIGURE 3.6: AUC planes for varying the background level and variance.

variances.

3.2.4 Varying Seeing Conditions

Over the course of an observation run the astronomical seeing angle will change and this leads to a varying performance in the adaptive optics system of an astronomical telescope. To study this effect PAOLA was utilised to simulate PSFs with Strehl ratios from $0.1 \rightarrow 0.97$, where Strehl ratio was used as a metric for the quality of the PSF. In this experiment the Strehl ratio of the data was varied by changing the value of the Fried parameter over the range $r_0 = 0.07m \rightarrow 1.01m$. The Hotelling, matched-filter and star subtraction observers were then used to carry out a detection task on this data cube to examine the effect of AO correction on the AUC. For each value of Strehl ratio 300 noisy images with and without a companion present were simulated. An ROC curve was then computed for each observer and the corresponding AUC was determined. As the Strehl ratio was varied a curve was mapped out for each observer on the Strehl ratio-AUC plane. This experiment was repeated 10 times, the average Strehl ratio-AUC curves are presented in figure (3.7) along with the standard deviation of the AUC for some values of Strehl ratio.

The Hotelling observer outperforms the matched-filter and star subtraction observers, in terms of the AUC, for all the seeing conditions considered in this test. As would be expected there appears an almost linear increase in the AUC with Strehl ratio for the three linear observers.

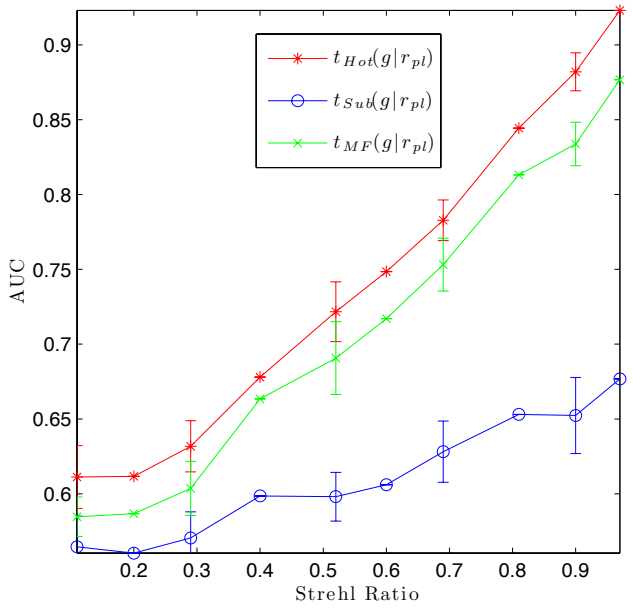


FIGURE 3.7: Plot of the Strehl ratio-AUC plane for the Hotelling observer, simple matched filter and star subtraction observers.

3.2.5 Mismatching the companion brightness in the template vector

The second set of experiments moves away from the pure SKE/PKE/BKE scenario. The sensitivity of the observers to mismatches, or errors, in their template parameters is investigated. In this experiment the dependence of the observers on the accuracy of the estimate of the companion brightness was investigated. AO corrected images of the Lick system, of size 64×64 pixels were again simulated using PAOLA as in section (3.2.1). The Hotelling observer and matched-filter observer were tasked with detecting a faint companion in data with the following parameters:

- $A_* = 9 \times 10^5 e^-$, $m_k = 5$, $\tau = 22\ ms$
- $b_m = 2 \times 10^{-2} e^-$ for all m, i.e. 19 magnitudes in the K band per arcsecond,
- $\sigma_m = 5e^-$ for all m,
- $r_{pl} = 5$ pixels,
- $A_{pl_{True}} = 220e^-$.

The star subtraction observer was not used in this test as this observer does not use the companion intensity as a search parameter. The observers were supplied with a range of test companion brightnesses $A_{pl_{Test}} = 0 \rightarrow 180e^-$. 300 noisy images with

and without a companion present were simulated. For each value of $A_{pl_{Test}}$ an ROC curve was computed for each observer and the corresponding AUC was determined. As the companion brightness was varied a curve was mapped out for each observer on the $A_{pl_{Test}}$-AUC plane. This experiment was repeated 10 times, the average $A_{pl_{Test}}$-AUC curves are presented in figure (3.8) along with the standard deviation of the AUC for some values of $A_{pl_{Test}}$.

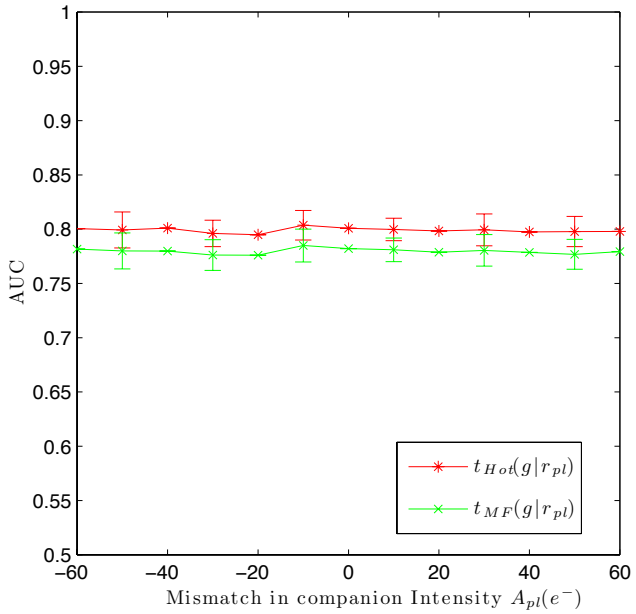


FIGURE 3.8: Plot of the $A_{pl_{Test}}$-AUC plane for the Hotelling and simple matched filter observers.

The results of this $A_{pl_{Test}}$-AUC mapping correlate well to the results shown in figures (3.3) and (3.4). That is the test statistics of the Hotelling and simple matched filter observers are linear in A_{pl}. Therefore an accurate estimate of A_{pl} will not effect the AUC in this case, the estimated value of A_{pl} simply changes the threshold to which the test statistics are compared to.

3.2.6 Mismatching the companion position in the template Vector

In the derivation of equation (3.4) it was assumed that the companion location, r_{pl}, was known and hence the value of the Hotelling observer for this formulation was not sensitive to variations of this parameter. It was therefore worthwhile to test this assumption. In this experiment the true value of the companion position was set at $r_{pl_{True}} = 5$ pixels. The other parameters of the simulation, i.e. the star brightness, companion brightness and background parameters, were kept the same as in the simulation in section (3.2.5).

The Hotelling, matched-filter and star subtraction observers were supplied with a range of test companion locations, $r_{pl_{Test}} = 4.5 \rightarrow 5.5$ pixels. 300 noisy images with and without a companion present were simulated. For each value of $r_{pl_{Test}}$ an ROC curve was computed for each observer and the corresponding AUC was determined. As the companion location was varied a curve was mapped out for each observer on the $r_{pl_{Test}}$-AUC plane. This experiment was repeated 10 times, the average $r_{pl_{Test}}$-AUC curves are presented in figure (3.9) along with the standard deviation of the AUC for some values of $r_{pl_{Test}}$.

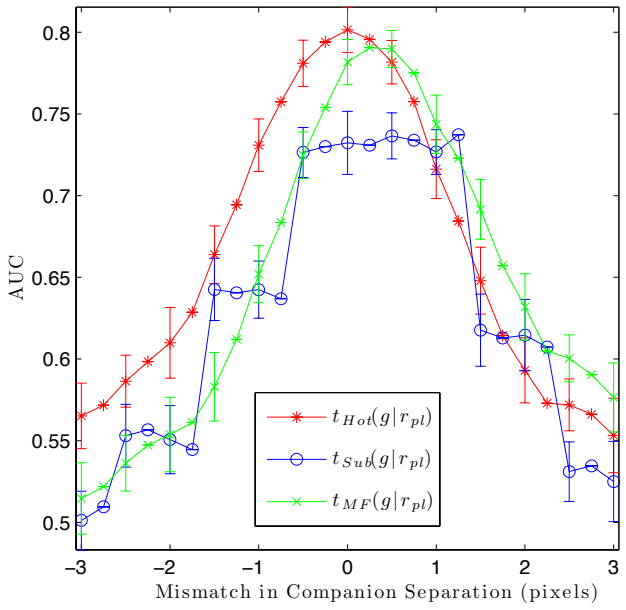


FIGURE 3.9: Plot of the $r_{pl_{Test}}$-AUC plane for the Hotelling observer, simple matched filter and star subtraction observers.

Figure (3.9) shows that the Hotelling and matched-filter observers appear very sensitive to an error in the companion position vector. This result confirms the assumption used to derive equation (3.4). However it should be noted that when the companion position is unknown the Hotelling observer reverts to the form expressed in equation (2.16) and becomes dependent on the companion location.

3.2.7 Mismatching the background intensity

As the background intensity is often estimated from the data itself it is important to understand how an error in this estimation would effect the detection of a faint companion. In this test the true value of the of the background intensity was set at $b_{m_{True}} = 100e^{-}$. The Hotelling, matched-filter and star subtraction observers were then

computed with a range of test intensities $b_{m_{Test}} = 0 \rightarrow 200e^-$. For each value of $b_{m_{Test}}$, 300 noisy images with and without a companion present were simulated. An ROC curve was then computed for each observer and the corresponding AUC was determined. As the background intensity was varied a curve was mapped out for each observer on the $b_{m_{Test}}$-AUC plane. This experiment was repeated 10 times, the average $b_{m_{Test}}$-AUC curves are presented in figure (3.10) along with the standard deviation of the AUC for some values of $b_{m_{Test}}$.

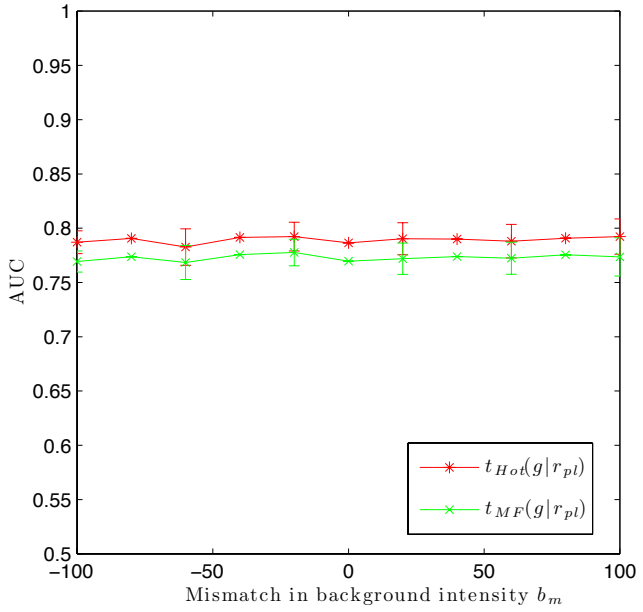


FIGURE 3.10: Plot of the $b_{m_{Test}}$-AUC plane for the Hotelling and simple matched filter observers.

The results of this experiment show that for the SKE/PKE test case the Hotelling and simple matched filter observers are insensitive to errors in the background estimation. Furthermore it may be possible to jointly estimate the background intensity as part of a multivariate estimation process as is done for the companion brightness.

3.2.8 Mismatching the variance of the detector readout

The variance of the detector readout, σ_m^2, is an important parameter in the analytic model of the data covariance matrix in the SKE/PKE/BKE test case. Hence an error in the estimation of this parameter should reduce the performance of the Hotelling observer for a detection task. To investigate this effect the variance of the detector readout was set at $\sigma^2_{m_{True}} = 10e^-$. The Hotelling and simple matched-filter observers were then supplied with a series of test variance estimates, $\sigma^2_{m_{Test}} = 0 \rightarrow 20e^-$. For each value of $\sigma^2_{m_{Test}}$,

300 noisy images with and without a companion present were simulated. An ROC curve was then computed for each observer and the corresponding AUC was determined. As the detector readout variance was varied a curve was mapped out for each observer on the $\sigma^2_{m_{Test}}$-AUC plane. This experiment was repeated 10 times, the average $\sigma^2_{m_{Test}}$-AUC curves are presented in figure (3.11) along with the standard deviation of the AUC for some values of $\sigma^2_{m_{Test}}$.

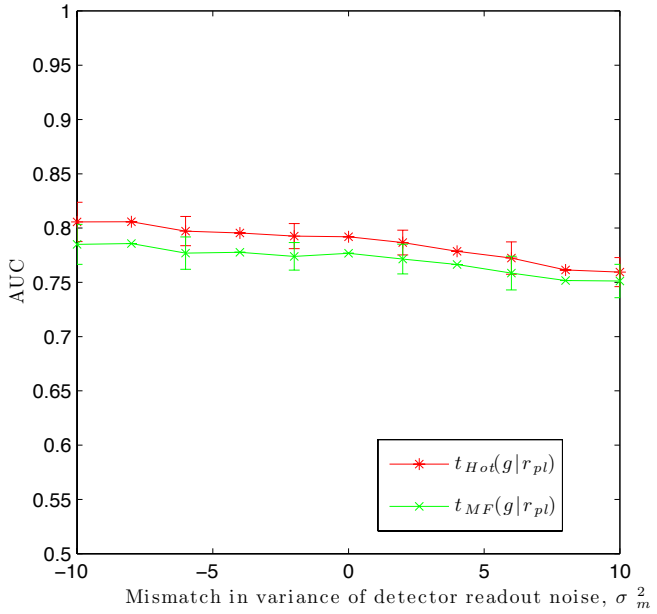


FIGURE 3.11: Plot of the $\sigma^2_{m_{Test}}$-AUC plane for the Hotelling observer, simple matched filter and star subtraction observers.

The results presented in figure (3.11) show that until the variance of the background intensity is over estimated by a factor of $\approx$ 1.5 the Hotelling observer outperforms the other tested observers. However after this cut-off the Hoteling and matched-filter observers perform on par with each other.

3.2.9 Mismatching seeing conditions

The Hotelling observer acts as a spatial matched filter, therefore it was crucial to study the dependence of the observer on the relation between the PSF of the data and the PSF used in the template vector. In this experiment PAOLA was again used to simulate PSFs. A set of PSFs were simulated with Strehl ratios of $0.1 \rightarrow 0.97$. In this experiment the Strehl ratio of the data was varied by changing the value of the Fried parameter over the range $r_0 = 0.07m \rightarrow 1.01m$. The Strehl ratio of the PSF used to simulate the data was selected as 0.52. The Hotelling, matched filter and star subtraction observers were

then calculated with a range of test PSFs with Strehl ratios ranging from $0.1 \rightarrow 0.97$. For each test PSF 300 noisy images with and without a companion present were simulated. An ROC curve was then computed for each observer and the corresponding AUC was determined. As the Strehl ratio of the test PSF was varied a curve was mapped out for each observer on the Strehl ratio mismatch - AUC plane. This experiment was repeated 10 times and the average Strehl ratio mismatch - AUC curve us presented in figure (3.12) along with the standard deviation of the AUC for some values of the Strehl ratio mismatch.

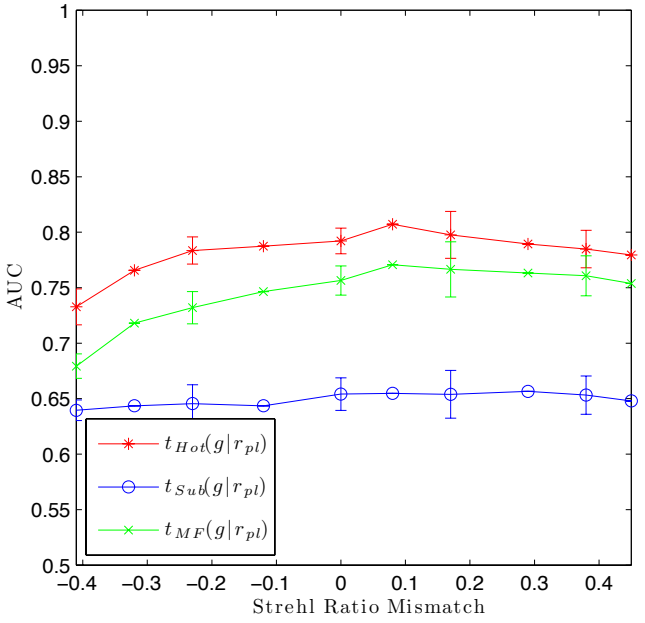


FIGURE 3.12: Plot of the Strehl ratio mismatch-AUC plane for the three linear observers

The results presented in figure (3.12) show that for the SKE/BKE/PKE test case knowledge of the exact profile of the PSF is not critical to the performance of the three observers.

3.3 Localisation Receiver Operating Characteristic Curves with simulated data

In real life situations the location of a faint companion in an image will in general not be known *a priori*. Hence it is necessary to include a localisation step along with the detection operation. In terms of the localisation ROC curve a detection task can

be coupled with a localisation task such that the True Positive Fraction is now the probability of correctly detecting a companion and correctly estimating the companion location to within a given tolerence, ϵ. A plot of the True Positive Fraction of correction decisions versus the False Positive Fraction of incorrect decisions is referred to as a localisation receiver operating characteristic curve.

In the following experiments the performance of the Hotelling observer is assessed for the tasks of companion detection, using equation (2.16) and companion localisation using the Spatial Scanning Hotelling Estimator, equation (2.17):

$$\begin{aligned} t_{Hot}(g) &= \max_{r_{pl}\in R} \sum_{m=1}^{M} \frac{A_{pl}h_m(r_{pl})}{A_* h_m(r_*) + b_m + \sigma_m^2} \left[g_m - \overline{g}_{0,m} - \frac{1}{2} A_{pl} h_m(r_{pl}) \right], \\ \widetilde{r}_{pl} &= \arg\max_{r_{pl}\in R} [t_{hot}(g)]. \end{aligned} \tag{3.8}$$

It should be noted that at this stage only the companion position is unknown, the other variables of the data, i.e. $A_*,\ A_{pl},\ b_m,$ PSF, and σ_m are assumed known.

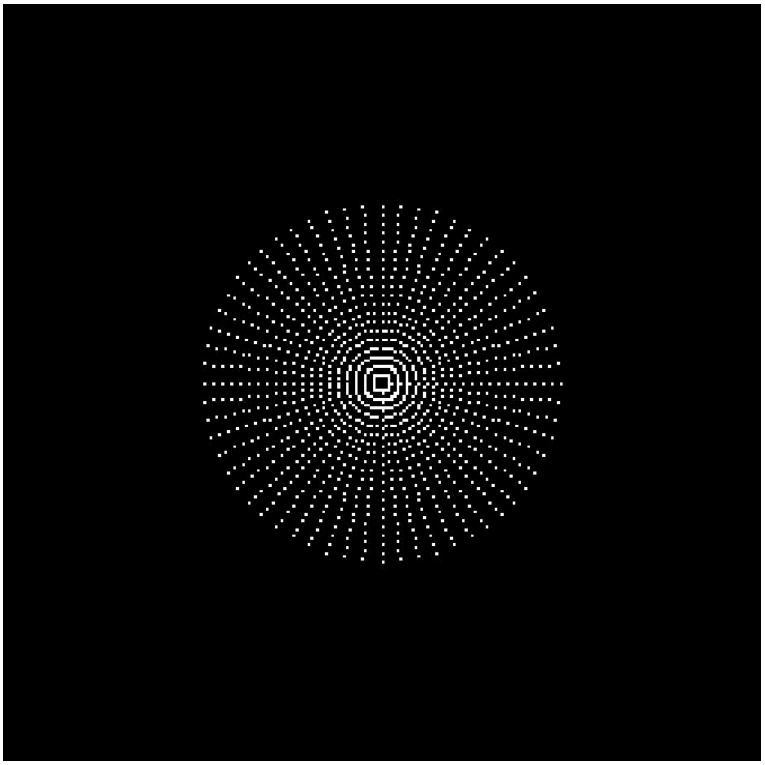
Three possible approaches to computing the SSHE were investigated. The first technique consists of defining a set of test locations, R, see figure (3.13(a)), such that the SSHE is at a maximum at the companion position, see figure (3.13(b)).

The set of test locations is defined by:

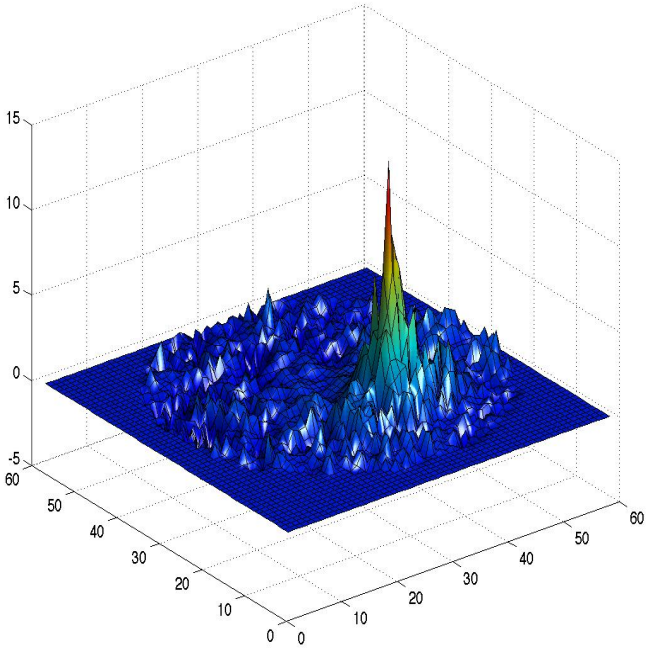
$$\begin{aligned} x_{i,j} &= \text{star position} + r_i * sin(\theta_j), \\ y_{i,j} &= \text{star position} + r_i * cos(\theta_j), \end{aligned} \tag{3.9}$$

where r_i is the radius of the search grid and the number of angular grid positions θ at each radius is given by $\theta(r_i) = 60 + i$. The drawbacks of this method are the heavy computation time required to calculate the observer at R locations as R needs to be large to cover an entire image and the spatial sampling of the test grid must be chosen by the user. However with this type of search no prior knowledge of the companion location is required. This approach to computing the values of the Hotelling test statistic will be referred to as $t_{Hot\ grid}$.

The second method is based around estimating the maximum value of the SSHE function in and around the location of the companion using a gradient index maximisation of the SSHE function. As the maximum value of the SSHE depends only upon the x & y



(a) Test locations, R, for the SSHE



(b) Values of the SSHE for the test locations in figure 3.13(a)

FIGURE 3.13: Set of test locations, (*top*), for the SSHE and the corresponding observer values, (*bottom*)

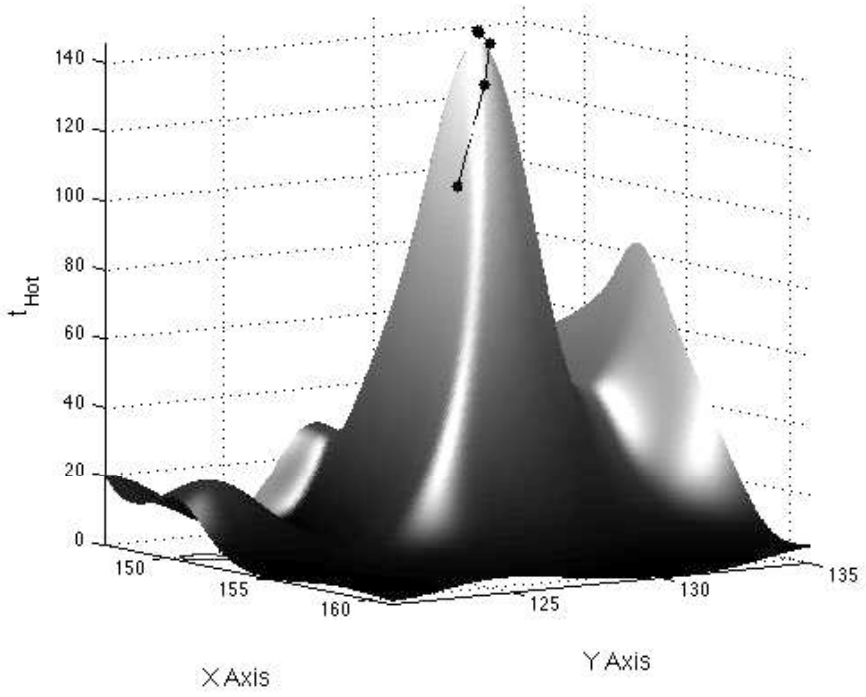


FIGURE 3.14: The value of the SSHE in the region around the location of the companion, the lines and dots show the maximisation test points.

position of the companion the Matlab function *fmincon* was used to estimate the maximum value of the function. The function *fmincon* carries out a bounded minimisation where the bounds were defined as the edges of the image. Bounding the minimisation in this manner prevents the wrap around of the template signal from one side of the image to the other. For some initial value of the companion position, r_0, the Matlab algorithm iteratively searches for the maximum of the function. The maximum is deemed to be reached when either the change in the position is less than 0.02 pixels or the change in the value of the function is less than 10^{-6}. A graphical representation of the values of the SSHE in the region around a faint companion are presented in figure (3.14). Finding the maximum of the SSHE using this technique reduces the number of test locations such that the computation time for an image of pseudo-Lick data takes less than a second compared to a minute using the search grid method. The drawback of this approach is the requirement of an initial guess for the companion position, it was observed that this estimated had to be within 3 pixels of the true location otherwise the observer would wander off and converge on a bright residual speckle in the image. This approach to computing the values of the Hotelling test statistic will be referred to as $t_{Hot\ ML}$.

The final method relies upon the fact that the scalar product of the template vector and the mean subtracted data vector can be viewed as a cross-correlation and hence can be estimated using Fourier transforms via the correlation theorem (Poyneer, 2003).

The PSF shift which produces the least-square difference between the signal template, $w = A_{pl}h_m r_{pl}$, and the prewhitened mean subtracted data, $K_g^{-1}(g - \overline{g}_{0,m})$ is the best answer. Formally the mean-square error (MSE) of the prewhitened mean subtracted data and the template at a displacement m, n is given by:

$$\begin{aligned} e[m,n] = &\Big[\sum\sum (w[i-m, j-n] - K_g^{-1}(g - \overline{g}_{0,m})[i,j])^2 \Big] \\ &\times [(N-m)(N-n)]^{-1}, \end{aligned} \tag{3.10}$$

where the images are of size $N \times N$ pixels, the summations are for $m \geq 0$, for $i = m$ to $i = N-1$, and for $m \geq 0$, from $i = 0$ to $i = N-1-m$. These summations hold likewise for n and j. Expanding (3.10) produces:

$$\begin{aligned} e[m,n] = &\Big(\sum\sum w[i-m, j-n]^2 + K_g^{-1}(g - \overline{g}_{0,m})[i,j]^2 - 2w[i-m, j-n] K_g^{-1}(g - \overline{g}_{0,m})[i,j] \Big) \\ &\times [(N-m)(N-n)]^{-1}. \end{aligned} \tag{3.11}$$

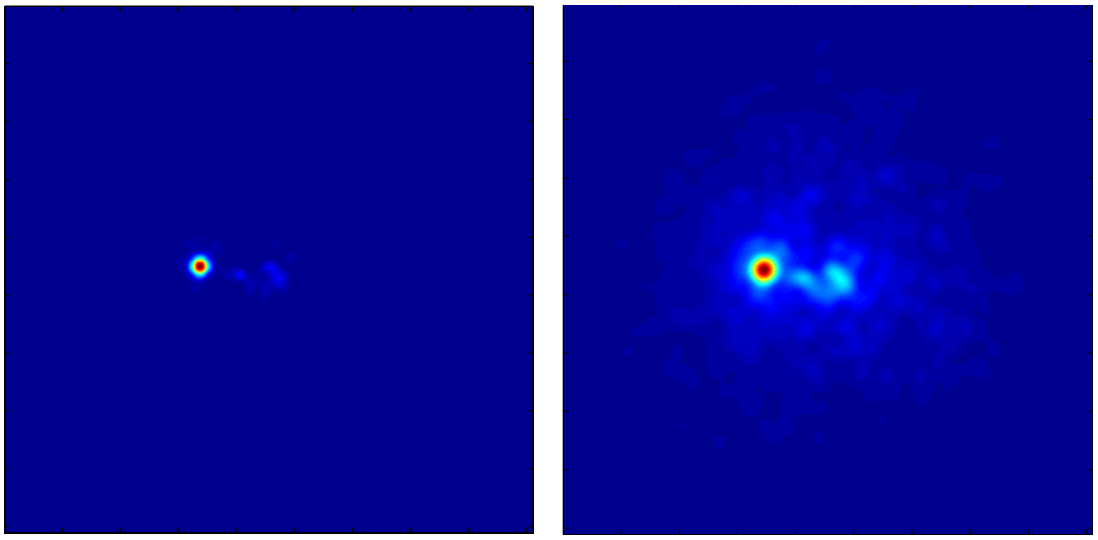
Poyneer (2003) notes that these two finite image signals can be thought of as single periods of an infinite periodic signal. Therefore the limits of the summations are now constant for all values of m and n:

$$e[m,n] = \left(\sum_{i=0}^{N-1} \sum_{j-0}^{N-1} w[i-m, j-n]^2 + K_g^{-1}(g - \overline{g}_{0,m})[i,j]^2 - 2w[i-m, j-n] K_g^{-1}(g - \overline{g}_{0,m})[i,j] \right) N^2. \tag{3.12}$$

Recalling the periodicity of the image signals, as one end of the signal moves away it will eventually wrap around from the other side. This results leads to a simplification of equation (3.12) as the two energy terms remain constant. The MSE equation can now be expressed as:

$$e[m,n] \propto -\sum_{i=0}^{N} \sum_{j=0}^{N} w[i-m, j-n] K_g^{-1}(g - \overline{g}_{0,m})[i,j], \tag{3.13}$$

which is the correlation of w and $g_{residual}$ calculated with periodic convolution. This correlation can be estimated efficiently in the Fourier domain by the use of the correlation theorem. To estimate the maximum value of the correlation function a parabolic interpolation is used to approximate the continuous correlation function and estimate its maximum value. This approach has the advantages of computational speed and no need of prior knowledge on the position of the companion. However this formulation does appear to be more sensitive to noise when compared to the gradient index maximisation or grid search approaches, see figure (3.15). This approach to computing the values of the Hotelling test statistic will be referred to as $t_{Hot\ MF}$.



(a) Image plane template fitting. (b) Fourier interpolated cross-correlation

FIGURE 3.15: The Hotelling observer computed using the grid search method *left* i.e. $t_{Hot\ grid}$ and the Hotelling observer interpolated using parabolic interpolation *right* i.e. $t_{Hot\ MF}$.

The performance of the Hotelling observer and matched filter observer in terms of area under the LROC curve (ALROC) was investigated. The three approaches outlined above for computing the Hotelling observer with an unknown companion location were used. The star subtraction observer was no longer used as it could only return integer shifts for the position of the companion which was too gross a shift when compared to the sub-pixel accuracy of the other methods. Artificial companions were simulated by scaling and shifting PAOLA simulated PSFs with a Strehl ratio of 0.52. The observers were tasked with detecting and locating a faint companion in data with the following parameters:

- $A_* = 9 \times 10^5 e^-$, $m_k = 5$, $\tau = 22\ ms$,
- $b_m = 2 \times 10^{-2} e^-$, i.e. 19 magnitudes in the K band per arcsecond,

- $\sigma_m = 5e^-$,
- $r_{pl} = 7$ pixels,
- $A_{pl} = 1800e^-$, i.e. $\Delta m_K = 6.75$.

For each experiment 10 positions for the companion were tested in order to minimise the bias from anisotropy in the PSF, see figure (3.16). In order to reduce computation time the grid search method, $t_{Hot\ grid}$, was constrained to search only within a region close to the location of the companion, with a spatial resolution of 0.1 pixels. The gradient index maximisation method, $t_{Hot\ ML}$, was given a starting position guess of plus or minus two pixels away from the maximum value of the mean subtracted data. For each companion position in figure (3.16) 300 noisy images with and without a companion present were simulated. An LROC curve was then computed for each observer and the corresponding ALROC was determined. In this simulation, the tolerance for a correct estimate of the companion position was set at $\varepsilon = 0.15$ pixels or 0.075 λ/D. The average ALROC estimates along with their corresponding standard deviation are reported upon in table (3.1).

TABLE 3.1: Comparison between methods of calculating the Hotelling observer

Observer	**ALROC**	**Std Dev ALROC**
$t_{Hot\ grid}(g\|r_{pl})$	084	0.1
$t_{Hot\ ML}(g\|r_{pl})$	0.85	0.04
$t_{Hot\ MF}(g\|r_{pl})$	0.8	0.03
$t_{MF}(g\|r_{pl})$	0.7	0.03

The results presented in table (3.1) demonstrate the limitations of the tested approaches. The $t_{Hot\ ML}(g|r_{pl})$ observer has the highest ALROC, at the cost of requiring an approximate companion location estimate. Therefore the rough companion location estimate, calculated using the Fourier based method, $t_{Hot\ MF}(g|r_{pl})$, could be input into $t_{Hot\ ML}(g|r_{pl})$ to yield a accurate companion location estimate without any use of prior companion location information.

3.3.1 Increasing Companion-Star Separation

Corresponding to the experiment reported upon in section (3.2.2) the effect of changing the companion radius upon the ALROC was investigated. Companions were placed at

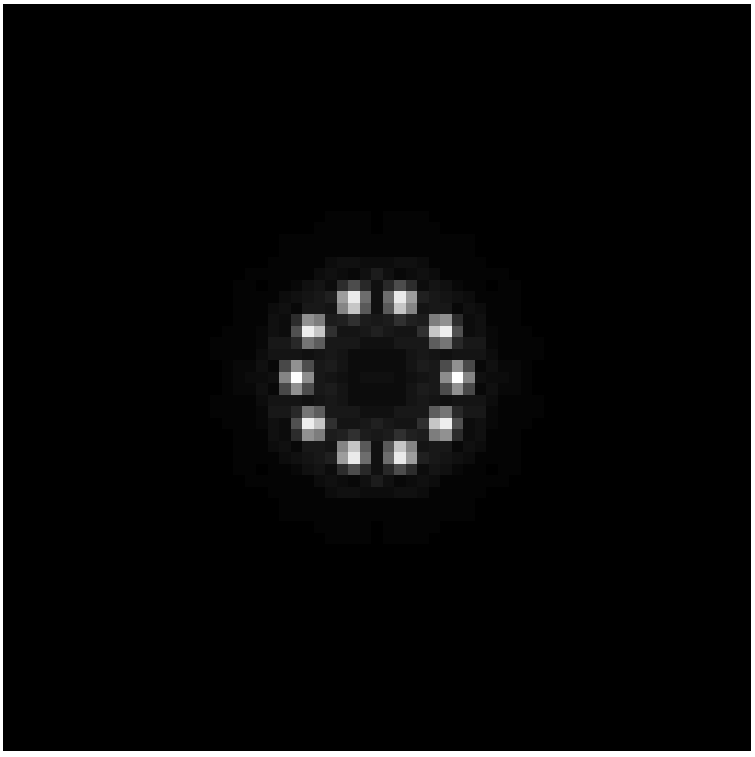
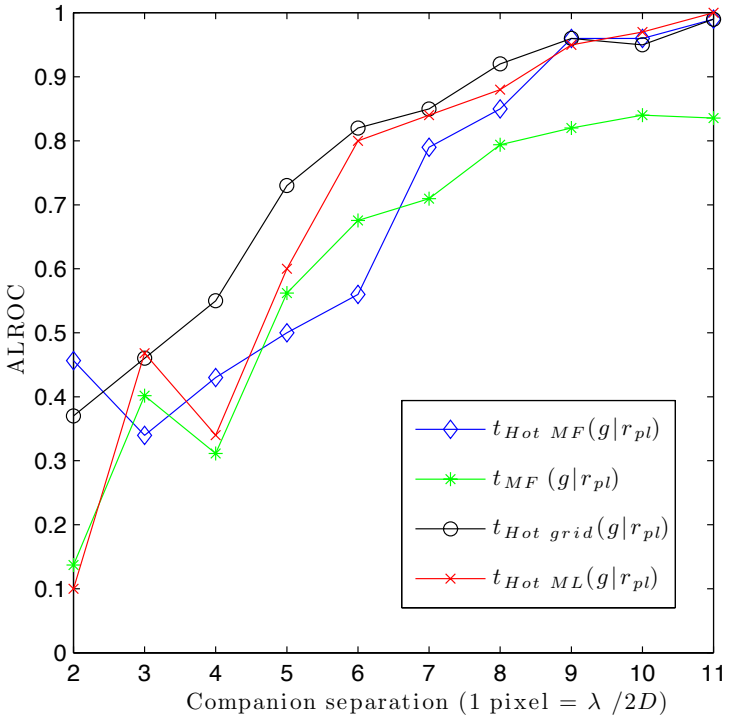


FIGURE 3.16: Ten artificial faint companion locations used to compute LROC curves for the various observers.

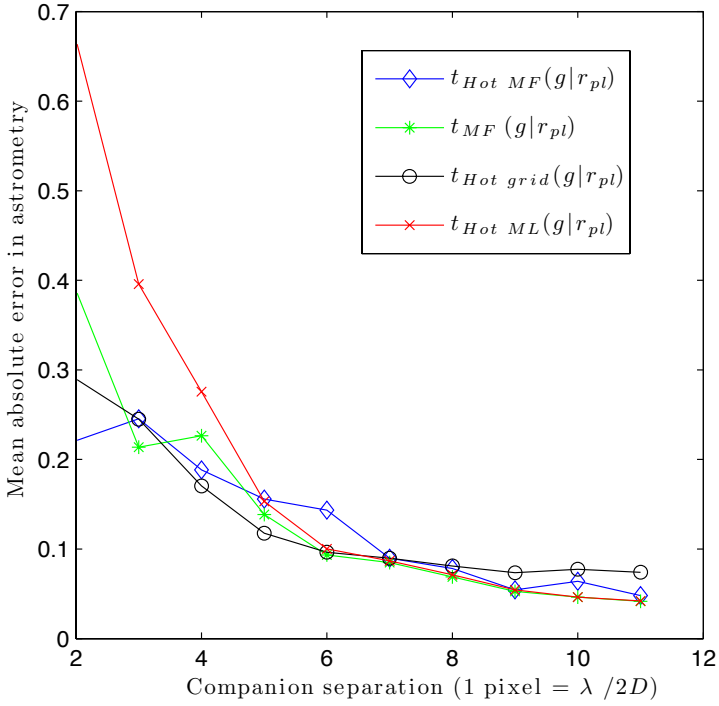
integer pixels from $r_{pl} = 0 \rightarrow 10$ pixels. For each value of r_{pl}, 300 noisy images with and without a companion were simulated. An LROC curve was then computed for each observer and the corresponding ALROC was determined. As the companion radius was varied a curve was mapped out on the r_{pl} - ALROC plane, see figure (3.17(a)).

The ALROC of all the observers increased with the star-companion separation distance. The minimisation and grid search approaches performed best. This is not a surprising result as this algorithm was always initiated within a few pixels of the true companion position. The two Fourier approaches performed similarly when the companion was close to the star. However the Hotelling observer outperformed the simple matched filter beyond 6 pixels from the star.

It was also possible to analyse how the astrometric accuracy of the observers varied with the star-companion separation distance, see figure (3.17(b)). As would be expected, the astrometric accuracy of all the observers increased with the star-companion separation distance. Beyond 6 pixels, or $2\lambda/D$, the the observers performed very similarly. Therefore the higher ALROC of the Hotelling observers, shown in figure (3.17(a)), must come from a better companion detection scheme compared to the straight matched filter, i.e. the prewhitening operation on the data must make a companion easier to detect above the background noise.



(a) Plot of the r_{pl}-ALROC plane for the four linear observers.



(b) Plot of the mean absolute error in astrometry for the four linear observers as a function of the companion-star separation distance.

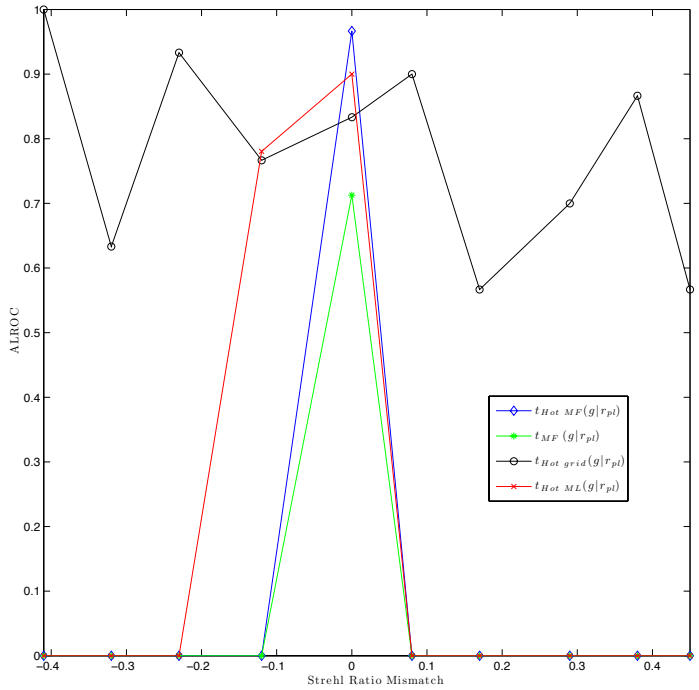
FIGURE 3.17: The area under the LROC curve as a function of the companion-star separation was investigated along with the corresponding mean absolute error in the estimation of the position of the companion.

3.3.2 Mismatch in Seeing Conditions

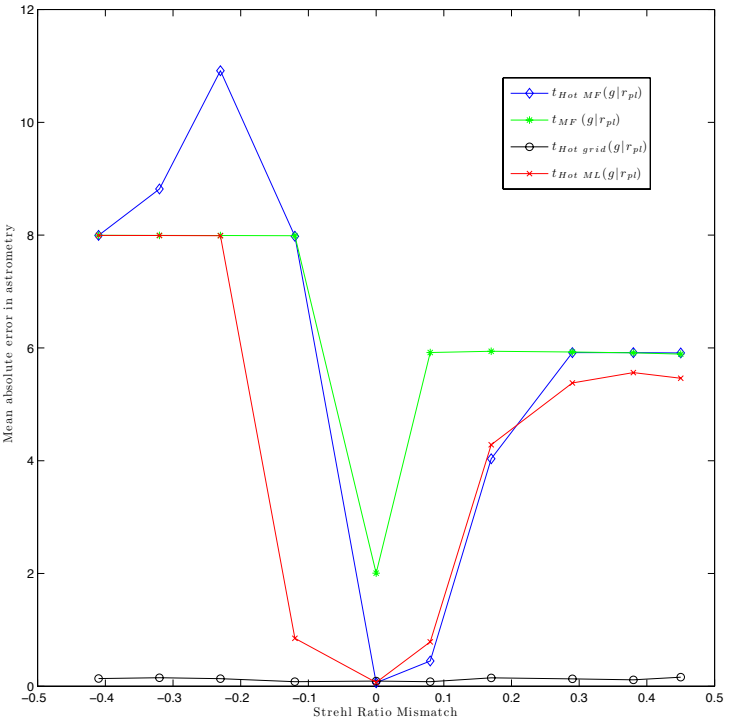
In previous experiments we have considered detection and localisation tasks where the PSF of the data was known exactly. In general this is an unrealistic assumption. We now consider an example where knowledge of the exact PSF of the data is unknown but the observer does have access to other reference PSFs acquired under different observational conditions. The goal of this experiment is to determine the sensitivity of the Hotelling observer to mismatches in the profile of the PSF used in the template vector. The simulated PSFs from section (3.2.4) were used in the following way: the data was simulated using a PSF with a SR of $\approx 52\%$ while the other PSFs with Strehl ratios ranging from $10\% \rightarrow 97\%$ were used by the Hotelling and matched filter observers to scan the data in an attempt to detect and localise a faint companion added to the data. In this experiment the Strehl ratio of the data was varied by changing the value of the Fried parameter over the range $r_0 = 0.07m \rightarrow 1.01m$. For each test PSF used, 300 noisy images with and without a companion were simulated. An LROC curve was then computed for each observer and the corresponding ALROC was determined. As the SR of the test PSF was varied a curve was mapped out for each observer on the SR mismatch - ALROC plane, see figure (3.18(a)).

Figure (3.18(a)) shows that all the scanning observers perform best with a properly matched PSF. The performance of the grid search method is relatively high as it is constrained to operate in the vicinity of the companion. In terms of astrometric accuracy, all the scanning observers compute a minimum astrometric error when using a properly matched PSF.

Figure (3.18(b)) shows that when a low Strehl ratio PSF is used to analyse the data the scanning observers nearly always converge upon the residual central star signal. This behaviour can be explained from the fact that the relatively sharp peak of the star signal will not be removed when using a broad low Strehl ratio PSF to perform the star subtraction operation. Accordingly when a high Strehl ratio PSF is used to subtract the central star signal a broad halo remains after the subtraction process. The scanning observers will then fixate upon this residual signal, this behaviour can be seen on the positive side of figure (3.18(b)).



(a) Plot of the SR mismatch - ALROC plane for the three linear observers



(b) Plot of the mean absolute error in astrometry for the four linear observers as a function of the Strehl ratio mismatch.

FIGURE 3.18: The area under the LROC curve as a function of the data and scanning PSF Strehl ratio mismatch was investigated along with the corresponding mean absolute error in the estimation of the position of the companion.

3.4 Estimation Receiver Operating Characteristic Curves with Simulated Data

Detection of a faint companion coupled with localisation and differential intensity estimation from simulated data is now reviewed. The True Positive Fraction is now the probability of correctly detecting a companion and correctly estimating the parameters of the companion to within a given tolerance, η. A plot of the True Positive Fraction of correction decisions versus the False Positive Fraction of incorrect decisions is referred to as an estimation receiver operating characteristic curve.

In section (2.3) it was shown that within the Hotelling framework there exists an unbiased estimator for the intensity of a companion at a given location i.e. equation (2.21) referred to as the Optimal Hotelling Estimator:

$$\widetilde{A}_{pl} = \frac{\sum_{m=1}^{M}(h_m(r_{pl})/(A_* h_m(r_*) + b_m + \sigma_m^2))[g_m - \overline{g}_{0,m}]}{\sum_{m=1}^{M} h_m(r_{pl})^2/(A_* h_m(r_*) + b_m + \sigma_m^2)}, \tag{3.14}$$

The performance of the OHE in terms of the area under the EROC curve (AEROC) was contrasted against other possible approaches. For the ten companion locations in figure (3.16) the observers listed in table (3.2) were tasked with detecting a faint companion with a differential magnitude of 7 in a simulated AO corrected image and estimating its position and intensity.

TABLE 3.2: Overview of the four observers used to detect and estimate the location and intensity of a faint companion.

Observer	**Search Method**
$t_{Hot\ grid}(g\|(a_{pl}, r_{pl}))$	contracting grid search in space and intensity
$t_{Hot\ ML\ Analytic}(g\|(a_{pl}, r_{pl}))$	gradient index maximisation for position, where the intensity is estimated at each position using the OHE
$t_{Hot\ MF}(g\|(a_{pl}, r_{pl}))$	Fourier domain estimation of cross-correlation on prewhitened data, using the OHE to extract the intensity of the companion
$t_{MF}(g\|(a_{pl}, r_{pl}))$	Fourier domain estimation of cross-correlation with a deconvolution estimator for intensity, i.e. equation (2.25)

For each position 300 noisy images with and without a companion were simulated. An EROC curve was then computed for each observer and the corresponding AEROC was determined. In this PSF known exactly experiment the tolerance for a correct estimate of the companion position was set at $\varepsilon = 0.15$ pixels and the tolerance on the estimate of the intensity was set at $\kappa = 0.1$ magnitudes. A table of the average area under the EROC curve is shown below, see table (3.3).

TABLE 3.3: Performance the four observers used to detect and estimate the location and intensity of a faint companion.

Observer	**AEROC**	**Std Dev AEROC**
$t_{Hot\ grid}(g\|(a_{pl}, r_{pl}))$	0.96	0.09
$t_{Hot\ ML}(g\|(a_{pl}, r_{pl}))$	0.95	0.08
$t_{Hot\ MF}(g\|(a_{pl}, r_{pl}))$	0.92	0.1
$t_{MF}(g\|(a_{pl}, r_{pl}))$	0.82	0.1

Similar to the results reported upon in table (3.1) when the PSF of the system is known exactly the performance of the Hotelling observers in terms of the AEROC curve are better then that of the straight matched-filter with a added deconvolution step. The differences in the AEROC of the three Hotelling observer implementations are due to the limitations in the various approaches. Even though the grid search method resulted in a high AEROC, this approach is very computationally intensive. The Fourier domain method fared worst of the Hotelling observers, this could be due to the unsuitability of the parabolic function used to model the continuous cross-correlation. The minimisation approached worked very well, however I observed that this method was very sensitive to the choice of the initial starting location. A more realistic case will follow where by the observers will be supplied with an inaccurate estimate of the PSF of the system.

3.4.1 Mismatch in Seeing Conditions

This experiment was designed to test the capability of the observers in table (3.2) at detecting a faint companion and estimating the position and intensity of a companion when the observers are supplied with a mismatched PSF. The simulated PSFs from section (3.2.4) were used in the following way: the data was simulated using a PSF with a SR of $\approx 52\%$ while the other PSFs with Strehl ratios ranging from $10\% \rightarrow 97\%$ were used by the Hotelling and matched filter observers to scan the data in an attempt to detect, localise and estimate the intensity of a faint companion added to the data. For each test PSF used, 300 noisy images with and without a companion were simulated. An EROC curve was then computed for each observer and the corresponding AEROC

was determined. As the SR of the test PSF was varied a curve was mapped out for each observer on the SR mismatch - AEROC plane, see figure (3.19).

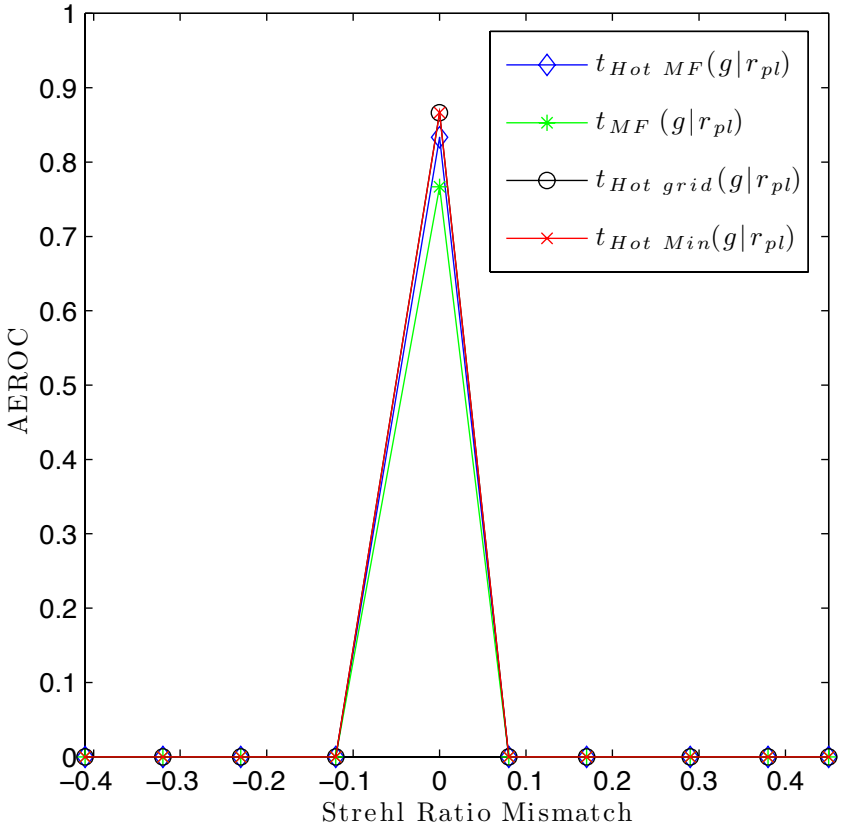


FIGURE 3.19: Plot of the SR mismatch - AEROC plane for the three linear observers

The results presented in figure (3.19) suggest that the intensity estimation of a faint companion is very sensitive to the correlation of PSF profile of the data and the scanning template. These results also demonstrate the advantage of the prewhitening operation of the Hotelling observer in flattening the residuals present in the mean subtracted data.

3.5 Detection and localisation of Shack-Hartmann spots for Wavefront Sensing in Strong Turbulence

Adaptive optics systems can be used in free space optical communication applications (Mackey, 2008) to reduce strong turbulence effects. Wavefront sensing is difficult in this application due to the presence of scintillation. Zeros of intensity begin to appear due to the presence of discontinuities in the phase, these discontinuous points in the phase, known as optical vortices, occur naturally when a laser beam propagates through the atmosphere (Roux, 1995).

Laboratory experiments were carried out to detect optical vortices in conditions typical to those experienced when a laser beam is propagated horizontally through the atmosphere. A spatial light modulator was used to simulate the effect of strong turbulence and a Shack-Hartmann wavefront sensor was utilised to measure the local tilts of the wavefront surface. The Hotelling observer was applied to the Shack-Hartmann spot images to detect and locate the spots in each sub-aperture. From the estimated spot positions an estimate of the wavefront could be recovered and input into the branch point potential method (Le Bigot & Wild, 1999; Wild & Le Bigot, 1999) to detect an optical vortex and characterise it if present.

This section describes work carried out with Kevin Murphy of the Applied Optics Group, who developed the experimental procedure for simulating optical vortices and measuring the resultant wavefront with a Shack-Hartmann wavefront sensor. He also analysed the outputs of the Hotelling observer with the branch point algorithm.

To test the application of the Hotelling observer to spot detection and localisation, ten thousand noisy spot images (size 19×19 pixels), with random spot locations, where the intensity of each spot was $A_{spot} = 10^3 e^-$, were simulated using a Gaussian PSF profile with a mean full-with-half-max of 2.5 pixels, varying normally with a standard deviation of 0.25 pixels from image to image. The noise in each image consisted of Poisson noise and normally distributed readout noise.

The Hotelling observer was applied to each spot image using the same Fourier domain method as $t_{MF}(g|(a_{pl}, r_{pl}))$ but using the OHE to derive the brightness of the spot, this method will be referred to as $t_{HotMF}(g|(a_{pl}, r_{pl}))$.

For each image the three matched-filter type algorithms as well as a classical centroiding algorithm, were tasked with locating a spot in the image. The PSF profile used in the template vectors was Gaussian in shape and had the same FWHM as the mean of the data. The t_{HotMF} showed the lowest mean error on the estimation of the position of the spot (in pixels) followed by t_{MF} and $t_{Hot\ ML\ Analytic}$, see table 3.4. The centroiding algorithm faired worst; however, it should be noted that this centroiding algorithm was not optimised as in Leroux & Dainty (2010). The centroiding algorithm used in the comparison is a simple, non-iterative centroiding algorithm. It thresholds the data and finds the centre of mass of a spot bigger than a certain size. The threshold is specific to each lenslet as it is intensity weighted for each subaperture. If two or more spots are detected above the threshold within one subaperture the brightest one is chosen as the focused Shack-Hartmann spot.

Observer	**Mean Error in Spot Position Estimation (Pixels)**	**AUC**	**SNR**
t_{HotMF}	3.6×10^{-4}	0.82	136
$t_{Hot\ ML\ Analytic}$	5×10^{-4}	0.76	206
t_{MF}	4×10^{-4}	0.80	81
Centroiding	1×10^{-3}		40

TABLE 3.4: Summary of testing the t_{HotMF}, $t_{Hot\ ML\ Analytic}$, t_{MF} and centroiding algorithms on ten thousand simulated spot images. The Hot_{MF} showed the lowest mean error in spot position estimation. This algorithm was also the most efficient spot detector having the highest AUC of the four tested observers.

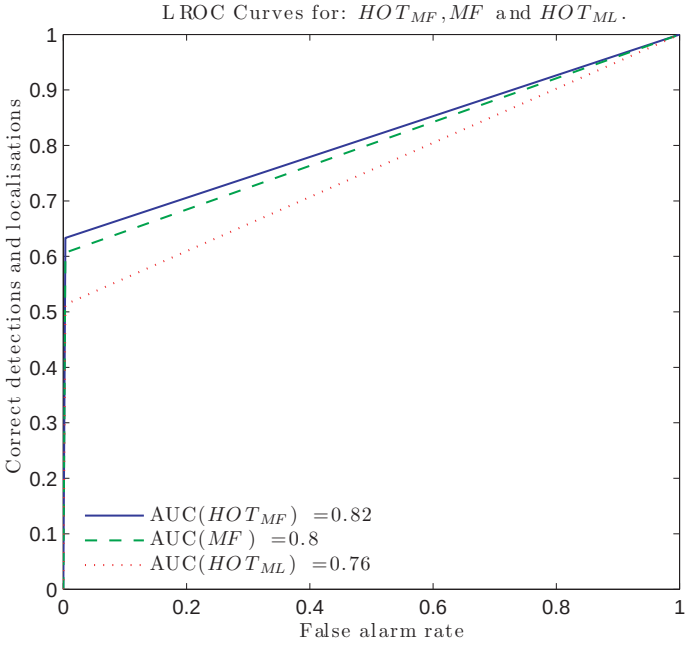


FIGURE 3.20: Plots of the LROC curves for: t_{HotMF}, $t_{Hot\ ML\ Analytic}$ and t_{MF}.

For the three matched-filter type observers it was possible to calculate a LROC curve where the tolerance on the accuracy of the location estimate was set at $\pm\epsilon = 4 \times 10^{-4}$ pixels.

Table (3.4) and figure (3.20) show that the cross-correlating Hotelling observer outperforms all the other tested algorithms in terms of mean centroiding error and in terms of the number of correct spot detections, i.e. t_{HotMF} has the highest area under the LROC curve, AUC= 0.82. Furthermore the t_{HotMF} has a much higher SNR compared to the straight matched filter, t_{MF}, and centroiding algorithm. Hence this increased SNR coupled with a high performance in spot detection and localisation shows that the t_{HotMF} is capable of detecting spots with very low intensities which is critical when analysing spot images in close proximity to a vortex. It should be noted that computing the Hotelling observer in the image plane could be improved by interpolation of the PSF

so that the resolution of the cross-correlation could be improved, as was done with the Fourier plane method.

In the laboratory experiment an initial plane wave, seeded with a single optical vortex, was simulated passing through 15 Kolmogorov phase screen realisations at a wavelength of $633nm$ with a propagation distance of $2.8km$. Figure (3.21(a)) shows an example of a Kolmogorov phase screen used in the simulation. Propagation of the optical field between phase screens was accomplished by multiplying the Fourier transform of the field by the Fresnel transfer function (Johnston & Lane, 2000; Martin & Flatte, 1990). The final degraded optical field was then encoded and applied to the spatial light modulator. The resultant wavefront was then sampled by a 46×46 lenslet Shack-Hartmann wavefront sensor.

Each Shack-Hartmann subaperture was defined by placing a 19×19 grid around each spot image, see figure (3.21(b)). The data covariance matrix was assumed to only compose noise terms and was expressed as:

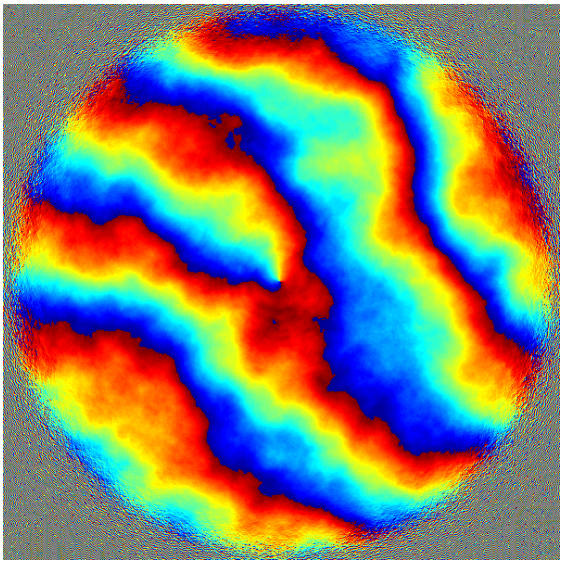
$$K_g = A_{spot} * h_m(r_{spot}) + \sigma_m^2 + b_m, \tag{3.15}$$

where A_{spot} is the intensity of the spot at position r_{spot}, σ_m^2 is the detector readout variance for the m^{th} pixel and b_m is the expected background intensity at the m^{th} pixel. These quantities were estimated as in section (4.3). A PSF was estimated as the mean of the spot images from the reference wavefront and the Hotelling observer was applied to each spot image using the Fourier approach to estimate each spot location. For each candidate spot location the Hotelling SNR_t was used as the spot detection metric:

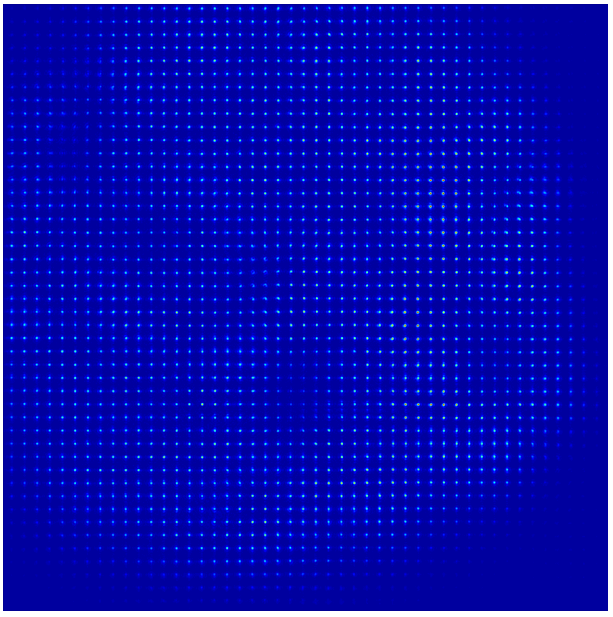
$$SNR_t = \frac{< t(g|H_1) > - < t(g|H_0) >}{\sqrt{\frac{1}{2}\sigma^2_{t(g|H_1)} - \frac{1}{2}\sigma^2_{t(g|H_0)}}}, \tag{3.16}$$

where $< \cdots >$ denotes ensemble average and $\sigma^2_{t(g|H_1)}$ and $\sigma^2_{t(g|H_0)}$ are the variances of $t(g|H_1)$ and $t(g|H_0)$. In this application $t(g|H_0)$ was estimated from the Hotelling test statistics of a ring surrounding the edge of the spot image, an estimate of the PDF of $t(g|H_0)$ was produced by taking the histogram of these test statistics, see figure (3.22). This histogram shows a Gaussian distribution, as was expected. The Hotelling SNR_t can be related directly to the AUC (Barrett & Myers, 2004) via:

$$SNR_t = 2\text{erf}^{-1}(2\text{AUC} - 1), \tag{3.17}$$



(a) Example of Kolmorogov phase screen used to simulate optical vortices



(b) Array of aberrated Shack-Hartmann spots

FIGURE 3.21: Example of phase screen applied to the spatial light modulator to produce an optical vortex and the acquired Shack-Hartmann spot image.

where erf^{-1} is the inverse of the error function,

$$\mathrm{erf}(x) = 2\pi^{-\frac{1}{2}} \int_0^x e^{-t^2} dt. \tag{3.18}$$

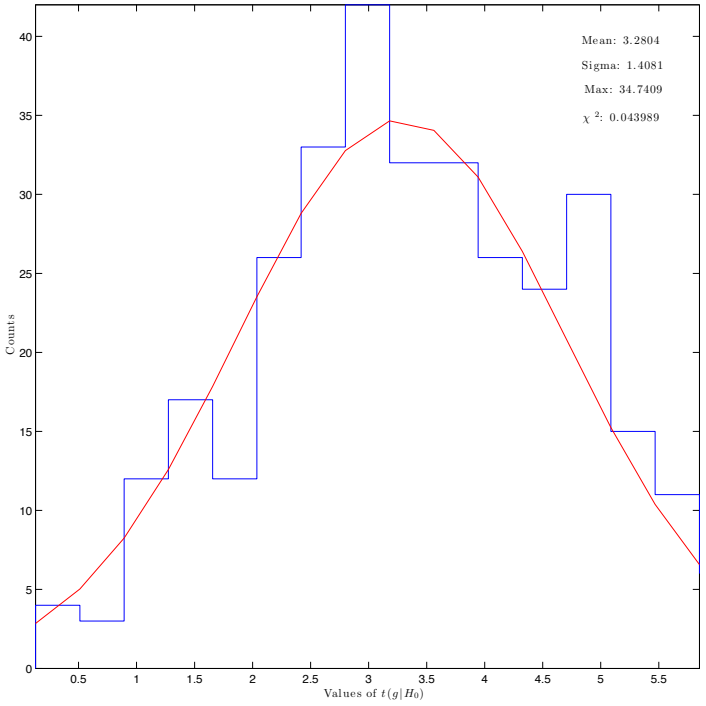


FIGURE 3.22: Histogram of the values of $t(g|H_0)$, sampled from a ring surrounding the edge of a spot image, and the Gaussian fit to this histogram.

A Hotelling SNR_t greater than 3, which equated to an AUC of 0.98, was selected as the threshold for a correct spot detection. When applied to the array of spot images in figure (3.21(b)) the Hotelling algorithm detected 2189 spots from 2200 analysed subapertures in contrast to 1974 spots detected by centroiding alone.

CHAPTER 4

BINARY STAR PARAMETER ESTIMATION

This chapter compares the accuracy of the Hotelling observer to other commonly used algorithms in astronomical photometry. The chosen algorithms are: StarFinder (Diolaiti et al., 2000) and Fitstars (ten Brummelaar et al., 2000). In section (4.1) a review of the current state of the art methods in determining differential astrometry and photometry of faint sources is presented. This review is followed by a set of experiments comparing the accuracy of the differential astrometry and photometry extracted using the Hotelling observer, StarFinder and Fitstars algorithms. In section (4.3) the advantages of prewhitening the image data is assessed using the peak-signal-to-noise ratio as a metric. This chapter concludes with the application of the Hotelling observer and StarFinder to a set of real observations of binary stars carried out at the Lick Observatory.

4.1 Current State of the Art Methods

There are three main approaches to deriving photometry from stellar images; aperture photometry, PSF fitting and deconvolution. The latter two techniques also extract relative astrometry. These algorithms were developed to analyse non AO-corrected images of crowded star fields. We will assume that they are applied to AO-corrected images, with a field of view smaller than the isoplanatic angle i.e. negligible variation in the PSF over the image.

Digital aperture photometry can be carried out using the APPHOT package (Davis, 1989) within the IRAF environment (Shames & Tody, 1986). The technique uses pixel

integration over a user-defined aperture in the image. Knowledge of the PSF is not required as PSF-fitting techniques are not used. This method works best on un-crowded star fields where there is no overlap of light from nearby stars.

It is often the case however that images of close binary stars have overlapping profiles, in which case aperture photometry is not suitable for extracting differential photometry. Stetson (1987) was one of the first to propose using PSF-fitting for this problem. In his PSF-fitting approach an analytic or empirical PSF is used together with a fitting algorithm to match scaled-and-shifted copies of the PSF to the data. The photometric accuracy of this method depends mostly on the accuracy of the PSF estimate. This estimate can be obtained by describing the PSF analytically or numerically. The most commonly used analytical PSFs are Gaussian, Lorentzian or Franz functions (Devaney, 1992). The analytical approach works best when the images are critically sampled (Nyquist), or under-sampled (Stetson et al., 1990), with the empirical approach the PSF is extracted directly from the observations. If there is a bright isolated star in an image, then a sub-array containing this isolated star will provide an empirical model of the PSF. This method is best implemented on over-sampled data (Stetson, 1992).

A combination of the above methods has been implemented in the DAOPHOT II software package (Stetson, 1992). An analytical function is first fitted to the data for several bright isolated stars, scaled copies of this profile are then subtracted from the original data. Sub-arrays containing the residuals are extracted and averaged together. The estimated PSF is then a combination of the analytical PSF and an interpolation on the residuals. This method works best when the analytic function accounts for over 90% of the profile shape within the "true" PSF.

Several approaches have also been developed to specifically extract relative astrometry and photometry from AO corrected astronomical fields. These methods include those presented by Mugnier et al. (1998); T. Fusco et al. (1999); Flicker & Rigaut (2005) and Mugnier et al. (2004). These algorithms are based on stochastic approaches which make use of prior knowledge about the science object and AO corrected PSF to perform a *myopic* deconvolution.

4.1.1 StarFinder

A recent implementation of the PSF-fitting approach is the StarFinder package (Diolaiti et al., 2000). The code, implemented in IDL, which has a graphical user interface, was developed for the specific purpose of measuring relative photometry and astrometry in AO-corrected stellar fields. The algorithm operates in two stages: initial PSF estimation and iterative PSF fitting. In the first stage, bright isolated stars in the image are

identified and background-subtracted. The resulting sub-images are then registered with sub-pixel accuracy, normalised, and median-combined to produce the first PSF estimate. A synthetic field, consisting of delta functions convolved with the PSF estimate is then subtracted from the data. In the second stage the residual image is searched via cross-correlation with the PSF template for additional sources. Images of the detected secondary sources are then fitted with the PSF estimate. The result of this procedure is relative astrometry and photometry, as well as an updated estimate of the PSF.

The StarFinder code terminates typically after $2-3$ iterations of its main fitting loop, see figure 2 in Diolaiti et al. (2000), when the number of detected sources approaches a constant value.

This analysis of stellar fields is similar to the Clean algorithm (Högbom, 1974) which also represents the analysed image as a collection of point sources.

4.1.2 Iterative Blind Deconvolution

In some cases, no a priori estimate of the PSF is available and one may have to attempt to estimate the PSF and object simultaneously. This situation is referred to as blind deconvolution in the literature, a thorough review of the subject can be found in Pantin et al. (2007) and Blanc-Féraud et al. (2010).

Lane (1992) proposed to derive relative astrometry and photometry by minimising an error metric function E which imposes constraints in both Fourier and image spaces. The minimisation technique he used was an unconstrained conjugate gradient approach (Fletcher, 1987). In practice the algorithm produces a series of estimates of the object, o, and the PSF, p, which have a decreasing value of the function E with each iteration. The algorithm converges when a local minimum of E has been found.

This iterative blind deconvolution (IBD) framework was further refined by Jefferies & Christou (1993). Their error metric function contained four contributions: as with Lane's approach an error metric in the image and Fourier spaces was included, as well as an error metric for a band-limit constraint on the PSF and an error metric for the Fourier modulus. The conjugate gradient routine from Press et al. (1986) was used to minimise the combined error metric. They applied their IBD algorithm to a wide range of astronomical images and detected a third component in the binary system 85 Pegasi.

Fitstars (ten Brummelaar et al., 2000) is an iterative blind deconvolution algorithm specifically optimised for binary stars. The algorithm extracts differential photometry and astrometry of two stellar components, as well as the PSF corresponding to the observations. The object is assumed to consist of two δ functions. Hence if there are

$N = 2$ number of stars in the field and the intensity of the ith star is A_i and its position is (x_i, y_i) the object can be written as:

$$o(x, y) = \sum_{i=1}^{N} A_i \delta(x - x_i, y - y_i), \tag{4.1}$$

the image can then be expressed as the convolution of the object and the PSF:

$$i(x, y) = o(x, y) * p(x, y). \tag{4.2}$$

Using an initial guess for the PSF, this equation can be solved in a least squares sense for the positions and magnitudes of the stars in the field. The sample PSF can be taken from an image of a single star, or the result from a previous operation of Fitstars. Even a relatively poor initial guess of the PSF will converge and produce similar results. The PSF changes substantially between observations of different targets due to changes in seeing and variations in AO performance on different targets. In order to compensate for this, a new model of the PSF is extracted from the data itself using the fitted estimates of the position and magnitude differences. An estimate of the PSF based on the star k is:

$$p_k(x, y) = i(x, y) - \sum_{j \neq k}^{N} \left(A_j \delta(x - x_j, y - y_j) * p_{old}(x, y) \right). \tag{4.3}$$

An estimate of the PSF can be extracted for each star in the field and a new PSF is formed by use of a weighted average over the PSF estimates (ten Brummelaar et al., 1996):

$$p_{new}(x, y) = \frac{1}{\sum_{j=1}^{N} A_j} \times \sum_{j=1}^{N} A_j p_j(x, y) \tag{4.4}$$

If the new values for the position and intensity of the binary star are not correct the new deconvolved PSF will contain energy from both sources, resulting in a PSF estimate with a large central peak and a smaller peak at the position of the second star. As initial estimates for the position and intensity of both sources are never exact, this leaking of energy can cause the PSF to become a binary itself (ten Brummelaar et al., 1996). To avoid this problem for both the objects a mask is made based on a single star image. This mask is set to one in the region of the jth star where the intensity of the star is

much greater than zero. Beyond this region the mask has a Gaussian roll-off to zero. Therefore each new PSF estimate is multiplied by this mask and then normalised:

$$\int p_{new}(x,y)dxdy = 1. \tag{4.5}$$

Equations (4.3), (4.4) and (4.5) are iterated until the root-mean-square (RMS) of the residual error of the fit of the model to the data changes less than 1%. Fitstars has been used with adaptive optics data sets from several telescopes with results which compared well to other methods (Horch et al., 2001; Pluzhnik, 2005).

4.2 Comparison of the Hotelling observer to Fitstars and Starfinder using AO corrected Lick data.

4.2.1 Observations

In order to compare the photometric and astrometric accuracy of the Hotelling observer with PSF-fitting and IBD we used data obtained with the Lick Observatory AO system on the 3-m Shane telescope (Bauman et al., 1999). Closed loop images of bright, single stars were obtained using the high-speed sub-array mode with a size of 64×64 pixels of the 256×256 pixel IRCAL camera (Fitzgerald & Graham, 2006). This corresponds to a field size of 4.864×4.864 arc-seconds. The sub-array measurements were captured with typical exposure times of $22ms$. Each data set comprised ten thousand images. All data were obtained in the K band ($2.2\mu m$) where the diffraction-limit is $151mas$ so that the data were effectively Nyquist sampled. The individual short exposures were registered with sub-pixel accuracy to produce shift-and-add images. The average Strehl ratio of these SAA images was around 40%. For the details of the observations and data reduction see Gladysz et al. (2006).

In our experiments we tested the accuracy of the photometry and astrometry on relatively-low (SR$\approx$ 30%), moderate (SR $\approx$ 40%) and high (SR $\approx$ 50%) Strehl data. All three algorithms required an estimate of the PSF. This estimate was provided in the form of a calibration PSF observed $10 \rightarrow 20$ minutes after the science observations. We tested the algorithms with properly-matched PSFs (ΔSR= 0.02%) and mismatched PSFs (ΔSR= 0.06%). We used observations of four different stars to create this test data: IRAS 21549+3929 ($m_V = 12.1,\ m_K = 5.96$), HD 18009 ($m_V = 8.23,\ m_K = 5.02$), HD153832 ($m_V = 7.25,\ m_K = 4.78$) and HD 143209 ($m_V = 6.3,\ m_K = 3.92$). Table (4.1) shows which PSFs were used to create the three data sets.

TABLE 4.1: PSF's used to simulate binary images

	Science PSF	**Reference PSF**
Low Strehl		
Matched	IRAS 21549+3929 (SR = 0.29)	IRAS 21549+3929 (SR = 0.32)
Mismatched	IRAS 21549+3929 (SR = 0.29)	HD18009 (SR = 0.35)
Medium Strehl		
Matched	HD15382 (SR = 0.43)	HD15382 (SR = 0.42)
Mismatched	HD15382 (SR = 0.49)	HD15382 (SR = 0.42)
High Strehl		
Matched	HD143209 (SR = 0.54)	HD143209 (SR = 0.52)
Mismatched	HD143209 (SR = 0.54)	HD15382 (SR = 0.49)

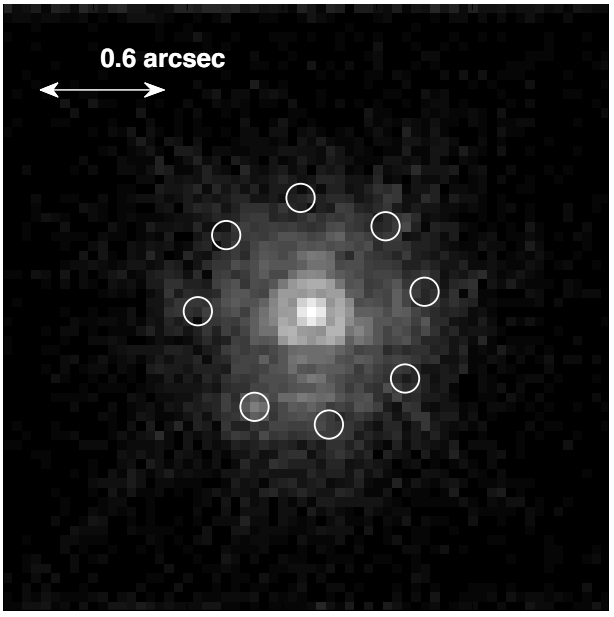


FIGURE 4.1: The eight artificial companion locations on a circle of radius 0.6 arcseconds.

Artificial binaries with differential magnitudes of $\Delta m_K = 3.5$ or 4.5 at a separation of $\theta = 0.6''$ were simulated by scaling and shifting the single-star SAA images. For each case 8 positions for the companion were tested in order to minimise the bias from anisotropy in the PSF, see figure (4.1). The mean absolute astrometric and photometric error was computed based on the results from these 8 positions.

While the Fitstars algorithm requires an initial estimate of the PSF, the estimate does not have to be very good e.g. a single star image from a different observation run. For these observations, Fitstars was tested with both properly matched and mismatched

calibration PSFs. In addition to the PSF estimate, Fitstars also requires an estimate of the initial locations of the stars in the image. These were obtained by visually examining the image. Fitstars was then able to iteratively derive a solution for each star.

As the StarFinder PSF-fitting algorithm was designed for the analysis of crowded fields imaged with AO, the fitting algorithm can take advantage of many estimates of the PSF within the field of view. In this experiment StarFinder was provided with the PSF estimate - an image of a single star. StarFinder had also to be given an estimate for the position of the companion and it was observed that these estimates had to be precise; otherwise the algorithm converged on the brightest static speckle. This situation was treated as non-convergence.

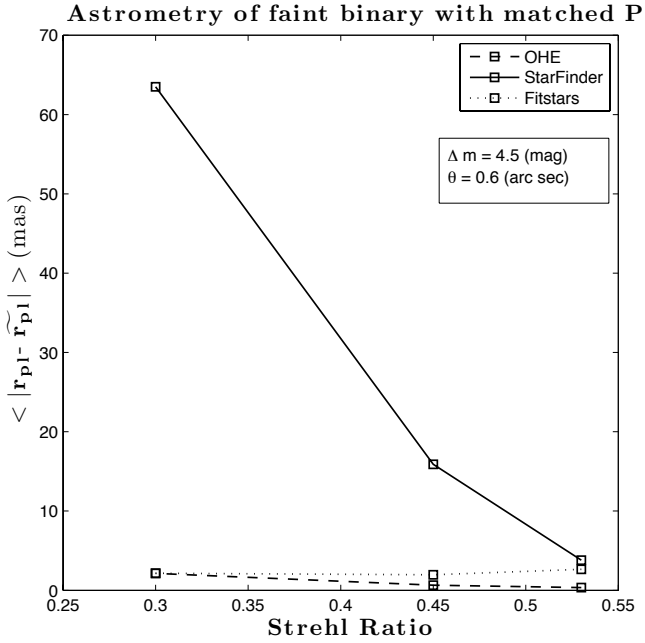
The Hotelling observer was supplied with the same PSF estimate as StarFinder and Fitstars. The star flux was estimated using a least-squares fit between the PSF and the data, were it was assumed that the companion flux was negligible in comparison to the flux of the star. The natural extension of this operation is to use the Hotelling observer to jointly estimate the star and companion locations and intensities as in T. Fusco et al. (1999). The observer was iterated using a gradient index maximisation for the position of the companion, where the intensity was estimated at each test position using the OHE. As with StarFinder and Fitstars the Hotelling observer required an initial estimate of the position of the companion, however this estimate did not need to be as precise as with the StarFinder code.

4.2.2 Results

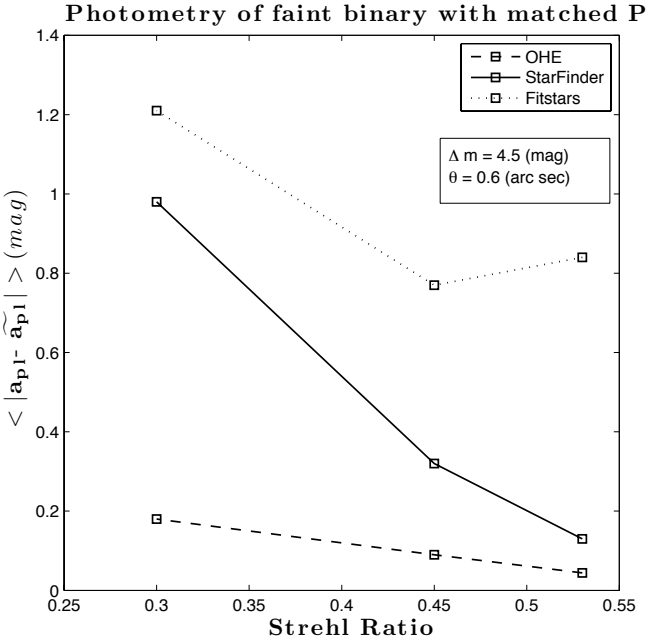
Tables (4.2) and (4.3) present the results for the mean absolute error in astrometry and photometry calculated using the three algorithms on the simulated data. The photometric and astrometric accuracy of the three observers, for the case of a faint companion with a matched PSF, is shown in figure (4.2).

The Hotelling algorithm always obtained the most accurate estimate for the position of the companion. In all tests, bar one, the Hotelling algorithm returned the most accurate estimate for the relative intensity of the companion. The Hotelling approach performed best on the high SR data set when using a properly matched calibration PSF, and the error in the results increase with decreasing SR.

As shown in figure (4.2(a)) the StarFinder error in astrometry is seen to depend very strongly on the Strehl ratio, with much weaker dependence in the case of Fitstars and the Hotelling observer. However, Fitstars fails to converge more often than the Hotelling



(a) Mean absolute error in the companion position estimates for the three data sets.



(b) Mean absolute error in the companion intensity estimates for the three data sets.

FIGURE 4.2: Error in astrometry for the three observers (*top*), error in photometry (*bottom*) of the faint companion with a matched PSF. One pixel = 0.76 mas.

TABLE 4.2: Comparative astrometry estimation task.

$\triangle$**m**	**PSF Match (SR)**	**Mean astrometric error (milli arc seconds) Hotelling Observer**	**StarFinder**	**Fitstars**
High Strehl ($\approx 50\%$)				
3.5	0.2%	0.28	3.0	0.01 (7 failures[1])
	0.6%	0.37	6.0	3.44 (2 failures)
4.5	0.2%	0.34	3.8	2.65
	0.6%	0.58	11.6	5.67 (5 failures)
Medium Strehl ($\approx 40\%$)				
3.5	0.1%	0.37	7.6	5.97 (4 failures)
	0.7%	0.37	4.5	3.45 (1 failure)
4.5	0.1%	0.65	15.9 (1 failure)	1.96 (3 failures)
	0.7%	0.63	12.8	6.11 (7 failures)
Low Strehl ($\approx 30\%$)				
3.5	0.3%	1.48	24.2	3.09 (5 failures)
	0.6%	1.1	9.8 (4 failures)	4.21 (4 failures)
4.5	0.3%	2.14 (2 failures)	63.5 (5 failures)	2.15 (3 failures)
	0.6%	1.76 (4 failures)	9.1 (7 failures)	3.95 (4 failures)

[1] failures refer to convergence of the algorithm on a bright speckle.

TABLE 4.3: Comparative photometry estimation task.

△m	PSF Match (SR)	Mean photometric error (magnitudes) Hotelling Observer	StarFinder	Fitstars
High Strehl ≈ 50%				
3.5	0.2%	0.04	0.05	3.42 (7 failures[1])
	0.6%	0.03	0.22	0.2 (2 failures)
4.5	0.2%	0.04	0.13	0.84 (1 failure)
	0.6%	0.08	0.43 (4 failures)	0.31 (5 failures)
Medium Strehl (≈ 40%)				
3.5	0.1%	0.05	0.14	1.17 (4 failures)
	0.7%	0.02	0.22	0.48 (1 failure)
4.5	0.1%	0.09	0.32	0.77 (3 failures)
	0.7%	0.1	0.52	0.36 (7 failures)
Low Strehl (≈ 30%)				
3.5	0.3%	0.09	0.46	1.27 (5 failures)
	0.6%	0.12 (2 failures)	0.32	0.125 (4 failures)
4.5	0.3%	0.18 (2 failures)	0.98 (5 failures)	1.21 (4 failures)
	0.6%	0.51 (4 failures)	0.37 (7 failures)	0.34 (4 failures)

[1] failures refer to convergence of the algorithm on a bright speckle.

algorithm or StarFinder. It is interesting that while StarFinder tends to provide less accurate astrometry than Fitstars, its photometric performance is better.

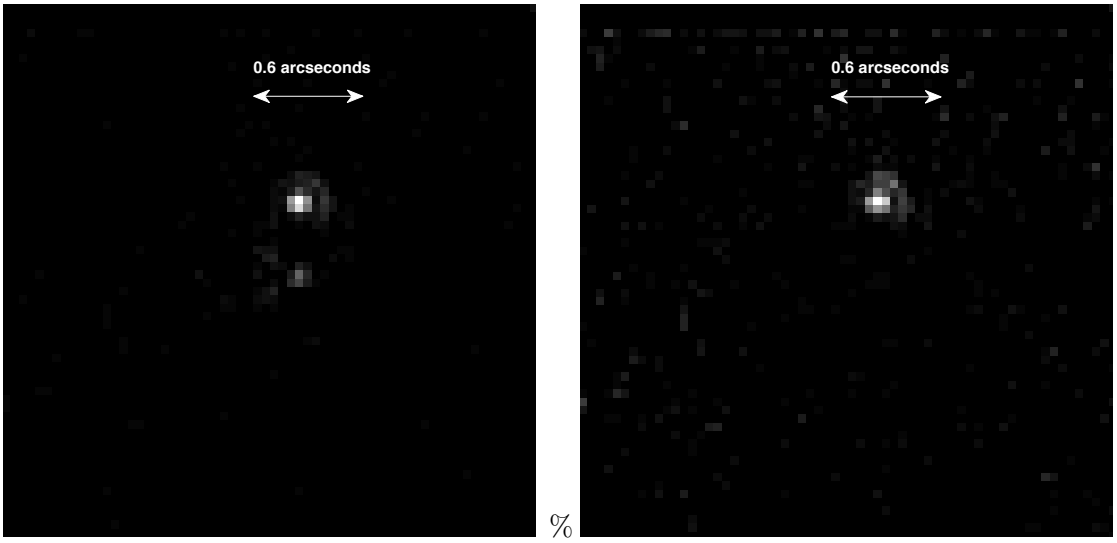
The ability of Fitstars to handle poor initial estimates of the PSF is shown in the results, where the accuracy using the mismatched PSF is practically the same as when using the matched PSF. This ability also causes the algorithm to perform poorly when the two stars are not well separated. When estimating the PSF, the estimate gets truncated at the midpoint between the two stars. In this data set the two stars are quite close together, $0.6\,arcsec \approx 4 \times FWHM$. Therefore a large amount of the PSF structure was not modelled completely and this leads to a decrease in the accuracy of the algorithm. This can be clearly seen in the photometry, figure (4.2(b)), because any mis-estimate of the PSF will cause large errors in the photometry. Astrometry only deals with the central point of the PSF, and is therefore more robust to this error. Fitstars has been shown to produce more accurate photometric measurements with more widely spaced binary stars (Roberts et al., 2005).

4.3 Effects of Prewhitening

In order to establish the effect of the prewhitening operation on the data, the peak-signal-to-noise ratio (PSNR) will be compared for the PSF-subtracted data and the prewhitened PSF-subtracted data, see figure (4.3). The signal-to-noise ratio (SNR) is defined as the ratio of the mean to the standard deviation of the measured data (Roggemann, 1996).

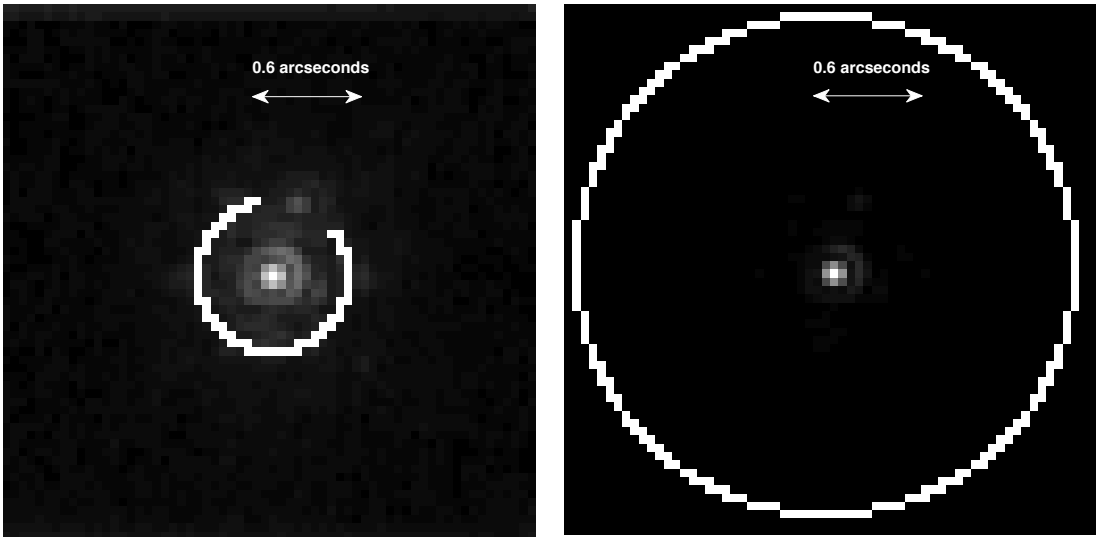
The binary system HD170648 ($m_k = 7.21,\ \Delta m_k = 3.07,\ separation\ \theta = 0.69$", spectral type A2) was observed along with a properly-matched calibration PSF (HD 173869: $m_V = 7.9,\ m_K = 7.53,$ spectral type A0), see Gladysz et al. (2008) for details. The covariance matrix of the data (equation (2.11)) was computed as follows: the intensity of the central star, A_*, is estimated using the calibration PSF in a simple least squares algorithm, the background, b, was estimated as the mean of an annulus centred at the bright star (see figure 4.4(b)) and the variance of the detector readout plus background was taken as the variance of the annulus pixel values. The PSNR is computed in the following way (Gladysz & Christou, 2008): the peak pixel value from the faint companion was compared to the standard deviation of the noisy pixels located the same distance away from the bright central star as the companion. This procedure is illustrated in figure (4.4(a)).

The PSNR is given by:



(a) PSF subtracted data. (b) Prewhitened PSF subtracted data.

FIGURE 4.3: The binary image minus the scaled PSF (a) and the subsequent prewhitened binary image (b).



(a) Annulus of pixels at the separation of the companion, used to estimate the PSNR.

(b) Annulus of pixels used to estimate the background level and the variance of the detector readout.

FIGURE 4.4: SAA image of the binary system HD170648 (a). The white circle denotes the location of pixels which were used to estimate the PSNR. The larger circle (b) is used to estimate the background.

$$\text{PSNR} = \frac{I_{Peak} - \overline{I}_{ring}}{\sigma(I_{ring})}. \tag{4.6}$$

The estimated PSNR was higher for the prewhitened data (PSNR = 277) than for the PSF subtracted data (PSNR = 164), see figure (4.5). Therefore the Hotelling observer

should be capable of operating on low SNR data there other classical observers would struggle.

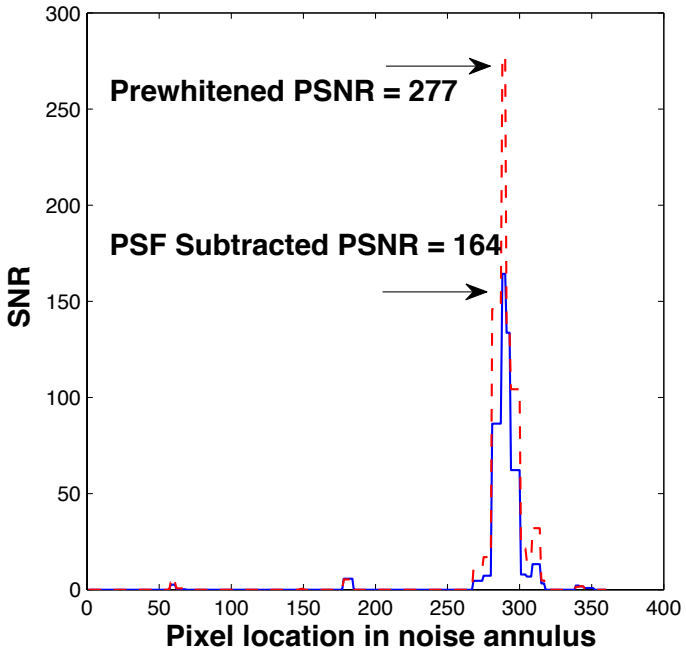


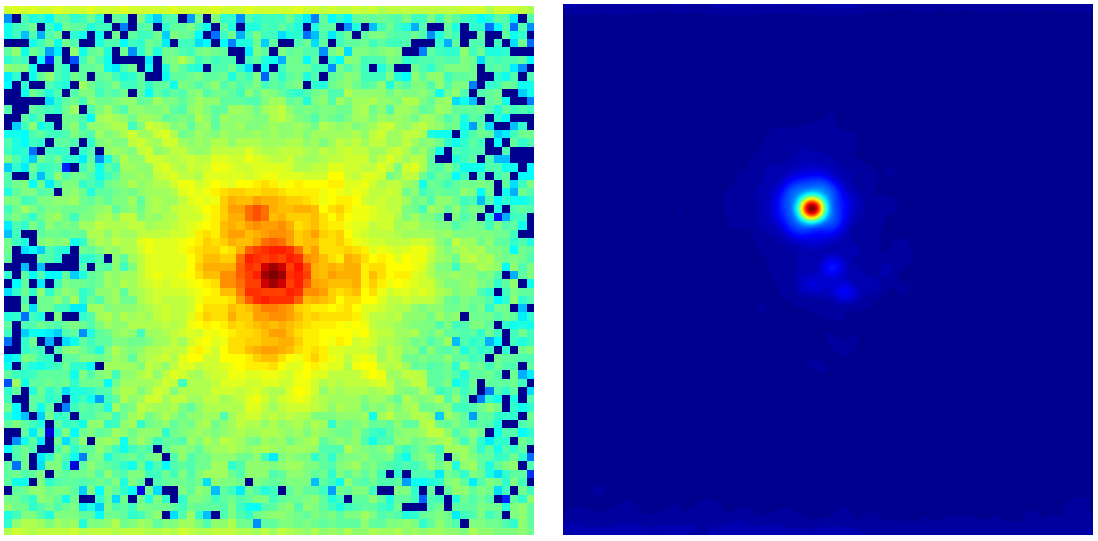
FIGURE 4.5: Comparison of pixel SNR for HD170648 in an annulus around the central star including the companion location.

4.4 Observations of Real Binary Stars

A small proportion of the observations carried out by Gladysz & Christou (2008) were aimed at binary stars. The analysis of two of the more interesting observations along with the analysis of observations of SAO 83636 are now presented. These observations were processed in the same manner as in section (4.2).

4.4.1 HD 235089

Figure (4.6(a)) shows the direct image of the binary star HD 235089 ($m_k = 6.41$). For this observation due to very stable atmospheric "seeing" a very good calibration PSF star was observed: HD 235160 ($m_k = 6.51$). Figure (4.6(b)) shows the application of the cross-correlating Hotelling observer, t_{HotMF}, to this data.



(a) Pre-processed image of the binary star HD 235089.

(b) Hotelling cross correlation of HD 235089 after subtraction of the reference PSF, shown on a linear scale.

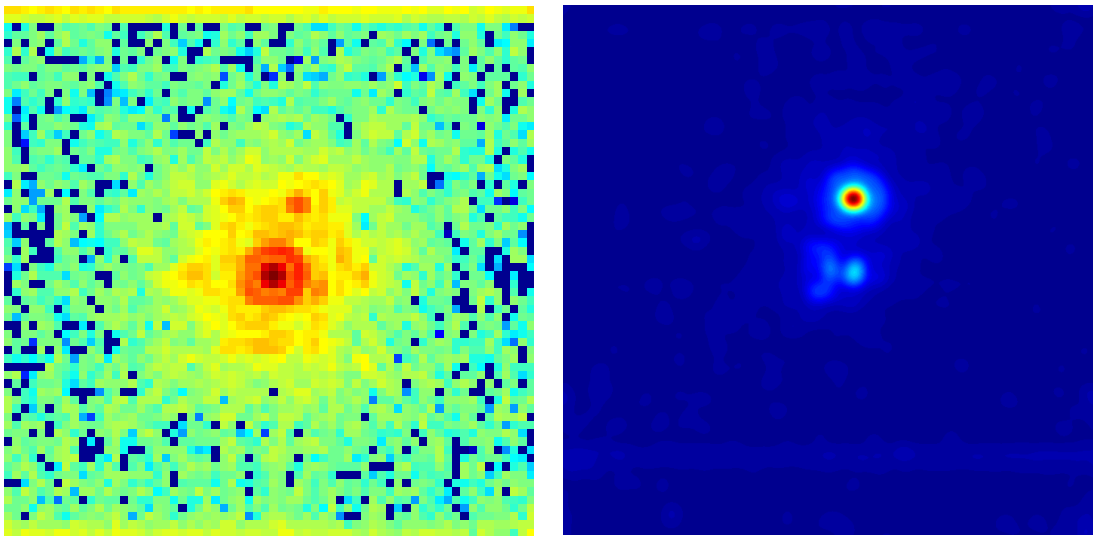
FIGURE 4.6: Observation of the binary star HD 235089 carried out at the Lick Observatory, shown on a log intensity scale and in false colour.

Figure (4.6(b)) shows that the prewhitening operation of the Hotelling observer smoothes out the residuals after the PSF subtraction step. The binary companion is easily detected above the smooth background, the PSF subtracted residual image has a PSNR = 67.8. Starfinder computed differential astrometry of 0.58 *arcseconds* and the Hotelling observer returned a very similar value of $0.579 \pm 2 \times 10^{-9}$ *arcseconds*. However there was a large difference in the differential photometry returned by the algorithms with Starfinder returning $\Delta m_k = 3.98$ and the Hotelling observer returning $\Delta m_k = 3.67 \pm 5 \times 10^{-10}$. A simple matched filter, equation (3.5) was applied to the data to produce a third differential intensity estimate. This observer returned differential photometry of $\Delta m_k = 3.67 \pm 1 \times 10^{-9}$, very close to the Hotelling estimate. Therefore I would be inclined to trust the results of the matched filter observers over that of StarFinder for this data.

4.4.2 HD 170648

As for the above data set, for HD 170648 ($m_k = 7.21$) a well matched calibration PSF star HD 173869 ($m_k = 7.53$) was observed within ten minutes of the science data. Figure (4.7(a)) shows the direct image of the binary star and figure (4.7(b)) presents the Hotelling analysis.

The Hotelling approach again produced a smooth cross correlation where the companion is clearly visible. Both the Hotelling observer and Starfinder produced very similar astrometry with values of $0.692 \pm 2 \times 10^{-9}$ *arcseconds* and 0.69 *arcseconds* respectively.



(a) Pre-processed image of the binary star HD 170648.

(b) Hotelling cross correlation of HD 170648 after subtraction of the reference PSF.

FIGURE 4.7: Observation of the binary star HD 170648 carried out at the Lick Observatory

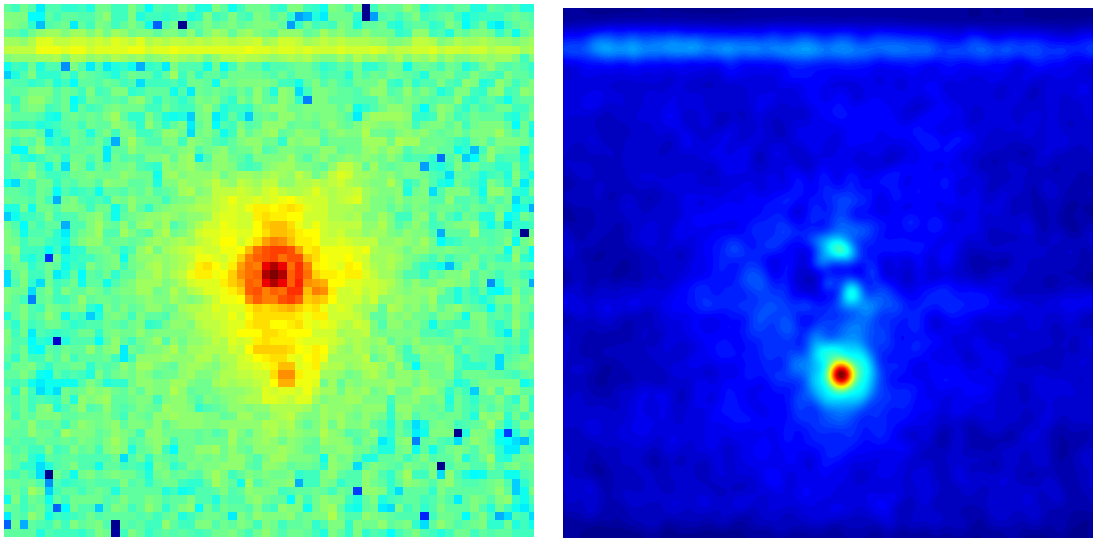
In contrast to the previous data set both methods also extracted similar values for the differential photometry of the companion, with the Hotelling observer returning a value of $\Delta m_k = 3.06 \pm 1 \times 10^{-8}$ and Starfinder extracting a value of $\Delta m_k = 3.07$.

4.4.3 WDS 15038+2006

This final data set consists of observations of the binary system WDS 15038+2006 ($m_k = 7.554$) and the calibration source TYC 149-123-1 ($m_k = 7.462$). Figure (4.8(a)) shows the direct image of the binary star and figure (4.8(b)) presents the Hotelling analysis. This data set had a significantly lower PSNR compared to the above data, with a PSNR = 20.9.

The Washington Double Star catalogue quoted the separation of this system as 0.9 arcseconds. The Hotelling approach again produced a smooth cross correlation where the companion is clearly detected. Both the Hotelling observer and Starfinder produced very similar astrometry with values of $0.924 \pm 1 \times 10^{-8}$ *arcseconds* and 0.932 *arcseconds* respectively. For this data the two algorithms also extracted almost identical values of photometry, $\Delta m_k = 3.18$.

The experiments reported upon in this chapter have shown that the Hotelling observer can extract differential astrometry and photometry from real AO corrected data on par,



(a) Pre-processed image of the binary star WDS 15038+2006.

(b) Hotelling cross correlation of WDS 15038+2006 after subtraction of the reference PSF.

FIGURE 4.8: Observation of the binary star WDS 15038+2006 carried out at the Lick Observatory.

and often better, compared to two widely used state of the art algorithms in astonomy. The benefits of prewhitening data has also been shown using the PSNR as a metric.

CHAPTER 5

THE QUASI-STATIC SPECKLE PROBLEM

As stated in section (1.4) instrumentally induced quasi-static speckles do not average out over time (Marois et al., 2005). This chapter reports upon simulation work carried out to implement the Simultaneous Difference Imaging (SDI) technique of Racine et al. (1999); Marois et al. (2000) and the Angular Differential Imaging (ADI) approach of Marois et al. (2006); Lafrenière et al. (2007). It should be stressed that simultaneous differential imaging and angular differential imaging are not competing strategies. A description of the multi-wavelength and sky rotated data will be outlined in section (5.1). The area under the estimation receiver operating characteristic curve and the peak signal to noise ratio will be used as the metrics to quantify the performance gain of the SDI and ADI algorithms. It will be shown in section (5.3.5) that these operations can be combined together to achieve greater speckle attenuation.

5.1 Data Simulation

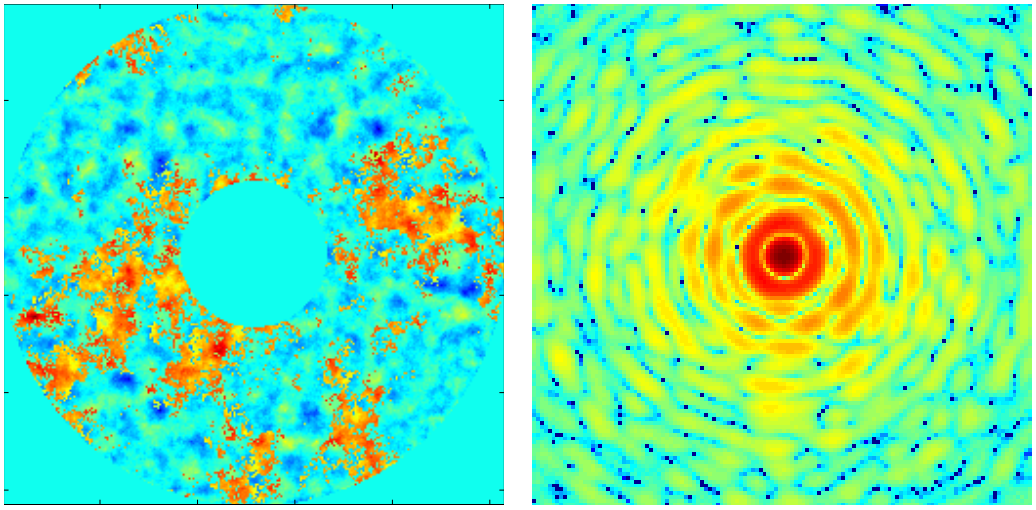
To test the differential imaging algorithms a Gemini-type telescope (Simons et al., 1995) was simulated in the Yorick development environment [1] with the following basic parameters:

- $D_{Primary} = 8.1m$,
- $D_{Secondary} = 1m$,
- $r_0 = 1.5m$ @ $1.25\mu m$,

[1] http://yorick.sourceforge.net/index.php

- Observing wavelength: $\lambda_1 = 1.578,\ \lambda_2 = 1.625$ & $\lambda_3 = 1.652\mu m$,

A fractal method was used to simulate independent Kolmogorov phase screens (Lane & Dainty, 1992) having $r_0 = 1.5m$ at $\lambda = 1.25$ microns, a $D_{Primary} = 8m$ and 10 sub-apertures across the pupil. Each screen is fitted with B-splines of third degree, the number of which correspond to the number of actuators in the system i.e. 100. The difference between the screen and the fit gives the residual phase, in this case $22nm$. A static phase screen is added to this; the static phase screens have an RMS error of $30nm$ and follow a powerlaw with a slope of -2.5, see figure (5.1(a)), an example of an AO corrected PSF simulated with this approach is presented in figure (5.1(b)). This simulation corresponds to a situation where the errors due to the telescope optics are constant during the observation, while the speckles due to residual atmospheric observations are independent. This is realistic since the exposure times required to give significant field of view rotation are much longer than the atmospheric coherence time. A long exposure phase screen was approximated as the sum of 100 independent Kolmogorov phase screens. A long exposure PSF was then simulated by taking the square of the absolute magnitude of the Fourier transform of this screen.



(a) Residual AO corrected phase screen. (b) Simulated AO corrected PSF.

FIGURE 5.1: Example of simulated phase screen used to simulate AO corrected PSFs.

5.2 Simultaneous spectral difference imaging and the Hotelling observer

The SDI approach of Racine et al. (1999) and Marois et al. (2000) was implemented using the double difference algorithm:

$$dd = (I_1 - I_2) - k(I_1 - I_3), \tag{5.1}$$

where three images are recorded at three wavelengths: $I_1(\lambda_1)$, $I_2(\lambda_2)$, $I_3(\lambda_3)$ and $\lambda_1 < \lambda_2 < \lambda_3$ and the constant k is given by the ratio of the Strehl ratios S of the data:

$$k = \frac{S_{\lambda_1} - S_{\lambda_2}}{S_{\lambda_1} - S_{\lambda_3}}. \tag{5.2}$$

The Strehl ratio at each wavelength, S_{λ_i}, was estimated as the ratio of the peak of the abberated PSF over the peak of a simulated diffraction limited PSF, where the intensity in both images have been normalised to unity. The factor k can also be defined as the mean ratio of the single difference images:

$$k = \left\langle \frac{I_1 - I_2}{I_1 - I_3} \right\rangle_{(x,y)}. \tag{5.3}$$

It should be noted that the planet signal will disappear as k approaches 1. This second method for computing k was adopted as it was found to produce smaller image residuals in the double difference image. The data simulated in the 'off' channels, i.e. the data at $\lambda = 1.625\ \mu m$ & $1.652\ \mu m$, needed to be rescaled before subtraction from the first wavelength channel so that the quasi-static speckle structure would line up. The data was scaled in the following manner:

1. the data were Fourier transformed,
2. the Fourier transformed data was scaled by a factor of λ_1/λ_i, where $i = 2, 3$, by either adding or removing zeros from the Fourier spectrum,
3. the scaled Fourier transformed data was then inverse Fourier transformed to give the image at the required wavelength.

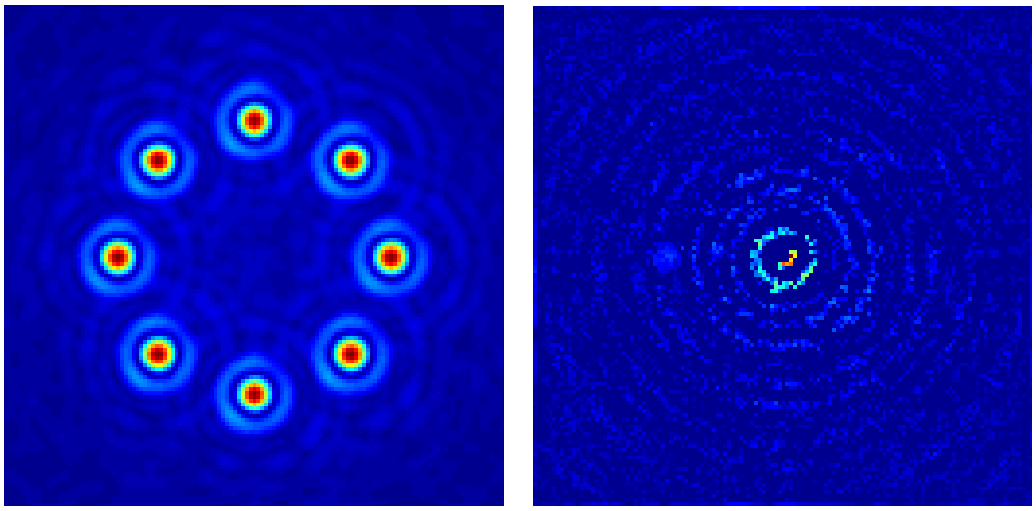
The major limitation of this scaling method is the requirement for integer padding of the Fourier transform. Cornia et al. (2010) have proposed a combination of zero padding the data in the Fourier and image domains to improve the scaling procedure. However this approach was not implemented here. For eight companion positions, see figure (5.2(a)), 300 noisy images, with and without a companion present, were simulated at the three selected wavelengths, where the noise in each channel consisted of static speckle, Poisson and detector readout noise. In this experiment the companion had a relative magnitude of $\Delta m_{\lambda_1} = 9$, $\Delta m_{\lambda_2} = \Delta m_{\lambda_3} = 14$. The double difference method was applied to

this data and the resulting residual image, see figure (5.2(b)), was processed with the Hotelling observer. The mean of the two 'off' channel scaled PSFs was used as the SDI double difference PSF estimate.

The Hotelling observer can be adapted to analyse SDI data (Barrett et al., 2006). In the SDI data the Poisson and detector readout noise is assumed to be independent in each wavelength channel. Therefore the noise terms will add together in the first and second SDI difference images and the data covariance matrix can be expressed as:

$$\begin{aligned}[K_g]_{m,m'} = &\left[\sigma^2_{m,\lambda_1} + \sigma^2_{m,\lambda_2} + b_{m,\lambda_1} + b_{m,\lambda_2} + A_{*,\lambda_1} + A_{*,\lambda_2}\right]\delta_{m,m'} \\ &+ k \times \left[\sigma^2_{m,\lambda_1} + \sigma^2_{m,\lambda_3} + b_{m,\lambda_1} + b_{m,\lambda_3} + A_{*,\lambda_1} + A_{*,\lambda_3}\right]\delta_{m,m'}.\end{aligned} \tag{5.4}$$

The detector readout variance plus background, $(\sigma_{m,\lambda_i} + b_{m,\lambda_i})$, can be estimated from each channel using an annulus centred around the parent star.



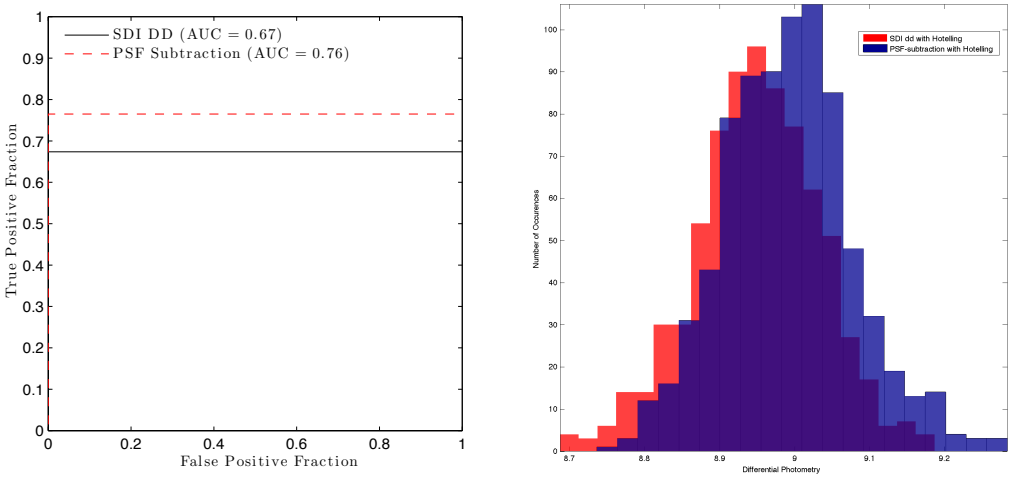
(a) Planet locations for SDI simulations, displayed on a square root scale.

(b) SDI double difference image for first companion location, displayed on a square root scale.

FIGURE 5.2: Artificial companion positions used to test the SDI double differencing method.

As a comparative technique another long exposure PSF was simulated at $\lambda = 1.578\mu m$, with the same static phase screen as the SDI data but with different realisations of Kolmogorov turbulence, with the same value of r_0. This PSF was subtracted from the first wavelength channel, λ_1 and the residual image was processed with the Hotelling observer using this PSF. This simulates having a perfect PSF calibrator at λ_1. For this data reduction method the Hotelling observer is applied in the same manner as in section (4.2).

An estimation receiver operating characteristic curve was computed for the SDI and PSF subtraction difference methods and the corresponding area under the EROC curve was determined. In this experiment, the tolerance for a correct estimate of the position of the companion was set at 0.2 pixels and the tolerance on the relative intensity estimate was 0.1 magnitudes. It should be noted that these two algorithms are not using the same information. The PSF subtraction method has exact knowledge about the nature of the uncorrected static aberrations in the system. Whereas the SDI DD approach only knows the ratio of the imaging wavelengths, λ_1/λ_2, such that the algorithm can spatially scale the data appropriately. An average EROC curve, presented in figure (5.3(a)), was produced by averaging the EROC curves over the set of eight companion positions. The area under the average EROC curves show that the PSF subtraction method performed on average slightly better that the SDI double difference subtraction method.



(a) Average EROC Curves for the SDI double differencing approach and PSF subtraction.

(b) Histograms for the photometry extracted from the SDI double differencing data and the PSF subtraction data using the Hotelling observer.

FIGURE 5.3: EROC curves and histogram of photometric accuracy for the SDI double differencing approach and PSF subtraction method.

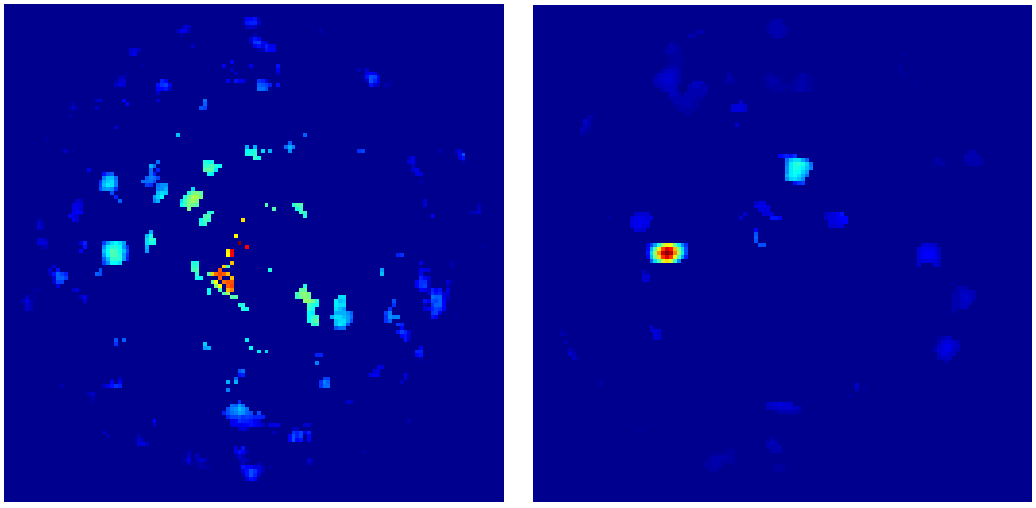
A downside of using a single metric, the AEROC, to analyse the combination of three tasks; faint companion detection and the computation of differential astrometry and photometry, is that one does not learn about how well an observer carried out one of the tasks independently of the other tasks. As the constraint on a correct estimate of the companion location was so tight in this case the performance of an observer could be investigated through the analysis of the computed differential photometry. For the eight companion positions in figure (5.2(a)), with 300 realisations of noise per position, a histogram was produced for the differential photometry calculated with the Hotelling observer using the SDI double differencing method and the PSF-subtraction method, given a correct companion location estimate has been made, see figure (5.3(b)). These

histograms show why the PSF-subtraction method had on average a higher AEROC compared to the SDI double difference subtraction method because the mean value of the PSF-subtraction differential photometry was 8.99 magnitudes whereas the mean value of the SDI double difference differential photometry was 8.95 magnitudes.

To illustrate the power of the Hotelling approach, the SDI double difference image was processed with an aperture photometry algorithm. For each position in the difference image the pixel intensities within a square aperture κ, of size 3×3, centred at r_{test} were summed and the mean of the aperture subtracted i.e.

$$t_{aperture}(g, r_{test}) = \sum \kappa - < \kappa > . \tag{5.5}$$

The resulting aperture photometry map was then thresholded to reject all values under 5 standard deviations of the noise. The noise and its standard deviation was estimated in increasing annuli centred on the star. In contrast the Hotelling likelihood map of the SDI double difference image was also thresholded to reject Hotelling test statistics under 5 standard deviations of the mean of the test statistics. Figure (5.4) shows these thresholded images.



(a) 5σ thresholded SDI double difference image processed with an aperture photometry algorithm (log scale).

(b) 5σ thresholded Hotelling Map of SDI double difference image.

FIGURE 5.4: 5 σ thresholded maps of SDI double differencing method.

The Hotelling approach shows a clear peak at the location of the companion whereas the peak in the aperture photometry map at the location of the companion (displayed on a logarithmic scale, see figure (5.4(a))) is difficult to distinguish from the various false positives surrounding it.

To show the effectiveness of the Hotelling observer in a more quantitative manner than a visual comparison the peak signal to noise ratio, equation (4.6), at the radius of the companion was estimated for the SDI double difference aperture photometry map and the SDI double difference Hotelling likelihhod map, see figure (5.5). The PSNR of the Hotelling likelihood map is approximately three times that of the pixel PSNR.

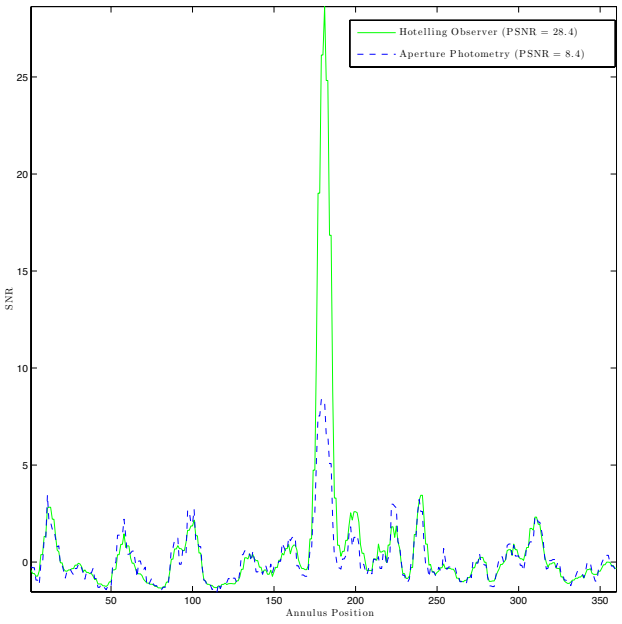


FIGURE 5.5: Peak SNR for the raw SDI double difference data and the Hotelling data map at the radius of the companion.

The results presented in figures (5.3, 5.4 and 5.5) show that the Hotelling observer can be combined with spectral difference imaging to improve the signal to noise ratio in the double difference residual image and detect, locate and extract photometry of faint companions.

5.3 Angular differential imaging (ADI) and the Hotelling observer

The PSF construction algorithm of Lafrenière et al. (2007); Marois et al. (2006) and the 'two-by-two' difference approach of Mugnier et al. (2009) were used to reduce simulated angular differential type data. Both differencing strategies assume that an image sequence is available to the algorithm whereby the limiting speckle structure is almost static over the course of the observation run. For both algorithms the Hotelling observer operates on the reduced difference image to compare the levels of quasi-static speckle attenuation.

5.3.1 Classical ADI

The ADI algorithm of Marois et al. (2006) is formulated around constructing a reference PSF in which a compromise is reached between the level of quasi-static speckle noise correlation in the reference PSF while making sure a sufficient companion displacement has occurred so that there is no self-subtraction of the companion signal. Marois states that a movement of at least twice the full width half maximum of the PSF is required. The time delay, τ_{min}, needed for the companion separation decreases as a function of the inverse of the angular separation between the companion and the host star. The initial step of this algorithm is to subtract the median of the image sequence from each image:

$$I_i^D = I_i - \text{median}(I_1, \ldots, I_n), \quad \text{for i} = 1 \ldots n, \tag{5.6}$$

where n is the number of images in the recorded sequence and the superscript D denotes difference image. This median subtraction minimises the noise in parts of the image which are dominated by pixel noise. The second step of this method takes the median of two images acquired before the target image and two images recored after the target image, in as close a time as possible, but which show enough field of view rotation to ensure no companion signal overlap. The reference PSF is broken into rings to accommodate the dependence of τ_{min} on the angular separation within the reference images. Therefore rings at different annuli will use different reference PSF images. The intensity inside each ring, i, is then scaled to minimise the noise after the subtraction of the reference PSF:

$$I_i^{ADI} = I_i^D - a \times \text{median}(I_{i-b}^D,\ I_{i-b-1}^D,\ I_{i+c}^D,\ I_{i+c+1}^D),\ \text{for } i \text{ rings}, \tag{5.7}$$

where a is the noise minimisation scaling factor, to be determined by the algorithm and the reference images recorded before and after the target image are I_b and I_c. The noise scaling factor a could be estimated for each ring using nonlinear optimisation algorithms to find the minimum of the square of the total residual inside each ring:

$$\widetilde{a} = \arg\min_a \left[\left(\sum_{ring} I_i^{ADI} \right)^2 \right]. \tag{5.8}$$

The set of difference images are then rotated to align their field of view rotation, θ, to that of the first image in the sequence via a bilinear interpolation. Finally a median is then taken over these aligned images:

$$I_{Final}^{ADI} = \text{median}(I_1^{ADI}, \text{rot}[I_2^{ADI}, \Delta\theta_{1-2}], \\ \text{rot}[I_3^{ADI}, \Delta\theta_{1-3}], \ldots, \text{rot}[I_n^{ADI}, \Delta\theta_{1-n}]). \tag{5.9}$$

5.3.2 ADI-LOCI

The PSF reconstruction approach developed by Lafrenière et al. (2007) builds upon the idea from classical ADI of splitting the reconstructed PSF into subsections in which the residuals in the difference image are minimised. In this algorithm the target image is cut up into small subsections and, independent for each new subsection, a linear combination of the set of reference images is sought such that the subtraction of the constructed PSF will minimise the noise. This algorithm is referred to as "locally optimised combination of images" or ADI-LOCI.

The coefficients used to weight the combination of reference images to create a subtraction subsection, S^T, are estimated by minimising the noise within a larger optimisation subsection, O^T. The optimal case is to make S^T as small as possible, however in practice the size of S^T is limited by computational resources. On the other hand the size of O^T is determined by the same requirements as classical ADI, i.e. the need to retain the signal of the companion. It should be noted that if O^T is too small the algorithm "sees" the signal of the companion as a residual to be minimised. The area of O^T is given by:

$$A = N_A \pi \left(\frac{W}{2}\right)^2, \tag{5.10}$$

where W is the full width half maximum of the PSF and N_A corresponds to the number of PSF cores within the subsection. As in classical ADI the set of reference images, K, must be selected such that the field of view rotation between frames is at least δ_{min} apart to negate self subtraction of the companion signal. The coefficients, c^k, are computed by the minimisation of the sum of the squared residuals of the optimisation subsection minus the reference subsections:

$$\sigma^2 = \sum_i m_i \left(O_i^T - \sum_k c^k O_i^k\right)^2, \tag{5.11}$$

where i is the pixel index, m is a mask used to reject bad pixels, k is the set of usable reference images and O_i^k are the reference subsections. This minimisation can be expressed as a linear problem of the form:

$$\begin{aligned} &\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}, \text{ where} \\ &\mathbf{A}_{jk} = \sum_i m_i \mathbf{O}_i^j \mathbf{O}_i^k, \\ &\mathbf{b}_j = \sum_i m_i \mathbf{O}_i^j \mathbf{O}_i^T, \text{ and} \\ &\mathbf{x}_k = \mathbf{c}^k, \end{aligned} \tag{5.12}$$

where bold font denotes a matrix or vector. Solving this system of equations results in an estimate of the coefficients c^k. Under the assumption that all the reference subsections, O^k are independent, A is always invertible, leading to a unique solution for c^k. The optimal reference subsection is then given by:

$$S^R = \sum_{k \in K} c^k S^k. \tag{5.13}$$

In the following experiments polar geometry has been used to define the optimisation subsections.

5.3.3 Initial Test of the Algorithm

Ten long exposure AO corrected PSFs were simulated at 1.578 μm. A field of view rotation of 18° from frame to frame was simulated in the image sequence such that the companion signal would not overlap in successive frames. Therefore the ADI LOCI algorithm should perform equally well at all angular separations for this data set. In practice such a field of view rotation can be achieved by either imaging very close to the zenith, i.e. $\approx 30°$ (Artigau et al., 2008) or by sampling the object at observing times that are spaced sufficiently. The latter method would have the drawback of possibly covering different seeing conditions.

The optimisation subsections were selected by dividing annuli azimuthally, see figure (5.6). The subtraction subsections were defined by their inner radius r, angular position, θ, radial width, Δr, and angular width, $\Delta\theta$. The optimisation subsections share their r, θ and $\Delta\theta$ with the subtraction subsections. The radial width of the optimisation

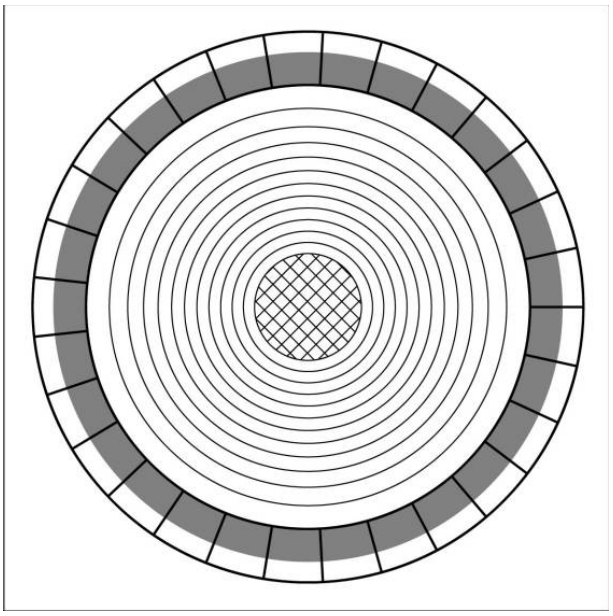


FIGURE 5.6: Example of subtraction subsections (grey) and optimisation subsections (thick lines) for the ADI-LOCI algorithm (Lafrenière et al., 2007).

subsections was selected as 3 pixels after testing several trial values. The angular width of the subsections is defined as:

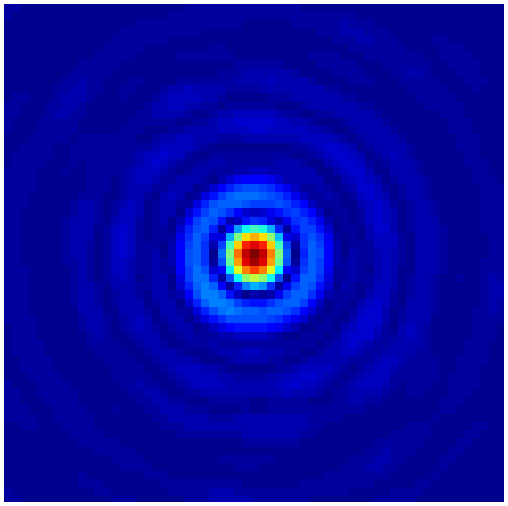
$$\Delta\theta = \left(\frac{2r}{W}\sqrt{\frac{1}{\pi N_A}}\right)^{-1}, \tag{5.14}$$

where $W = 2$ pixels and $N_A = 300$. A first test of the algorithm was carried out by inserting an artificial companion into the images with an angular separation of 16 pixels ($4\lambda/D$) and a relative magnitude of 10.

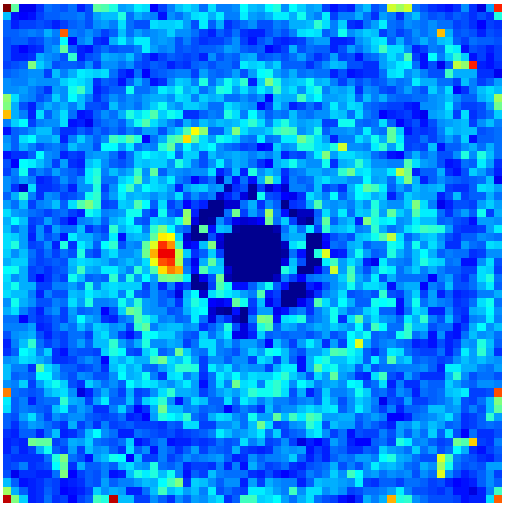
A radially symmetric average profile was subtracted from each image to reduce the seeing halo. The ADI-LOCI algorithm was then applied to the image sequence. The residuals were derotated, median combined and processed with the Hotelling observer, see figure (5.7). The Hotelling observer detected the companion and estimated an angular separation of 10.43 pixels, an error of 0.43 pixels or $\approx 0.1\lambda/D$ and extracted a value of 9.953 magnitudes for differential photometry.

5.3.4 EROC Curves in ADI LOCI reduced data

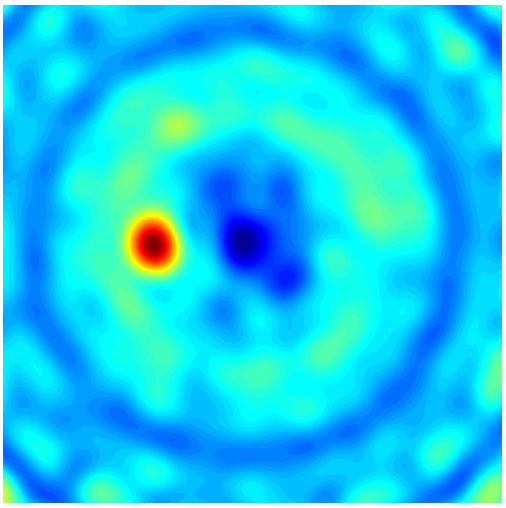
In this section I report upon investigating the performance of the ADI LOCI algorithm for faint companion signal detection, localisation and photometry. Artificial companions with a relative magnitude of $\Delta m_{\lambda=1.578\mu m} = 8$ were placed at λ/D intervals out to $25\lambda/D$



(a) Single noisy image of simulated ADI image sequence, displayed on a square root scale.



(b) Median combination of ADI LOCI reduced image sequence, displayed on a square root scale.



(c) Hotelling likelihood map of ADI LOCI reduced image with mean of reconstructed PSFs.

FIGURE 5.7: Initial run of ADI LOCI algorithm

away from the primary star. At each companion separation 300 noisy image sequences with and without a companion present were simulated. Both the companion absent and present sequences were reduced with the ADI LOCI algorithm.

In the ADI data it was assumed that the Poisson and detector readout noise in each image frame was independent. After the data was reduced using the ADI-LOCI algorithm the variance of the residual background, plus the mean background level, was estimated from an annulus of pixel values far away from the centre of the image.

The two-by-two angular subtraction method of Mugnier et al. (2009) was employed as a comparison to the ADI LOCI algorithm. This method works by subtracting angular

image pairs close enough in time such that there is little or no evolution of the static speckle field while far enough apart to allow enough angular movement of the companion to prevent self-subtraction. The difference images were derotated using bilinear interpolation, where a positivity constraint was applied to each frame. The derotated data was then median combined and processed with the Hotelling observer. As an estimate of the companion signal is needed by the Hotelling observer for the matched-filtering operation, the final combined difference image was subtracted from the first image in the ADI image sequence leaving only the star image. This star image was then used by the Hotelling observer.

An EROC curve was computed for both reduction algorithms with the tolerance on a correct companion location estimate of $0.1\lambda/D$ and the tolerance on a correct companion intensity estimate of 0.1 magnitudes. The corresponding area under the EROC curve was then determined. As the companion separation was increased a curve was mapped out on the angular separation - AEROC plane, see figure (5.8).

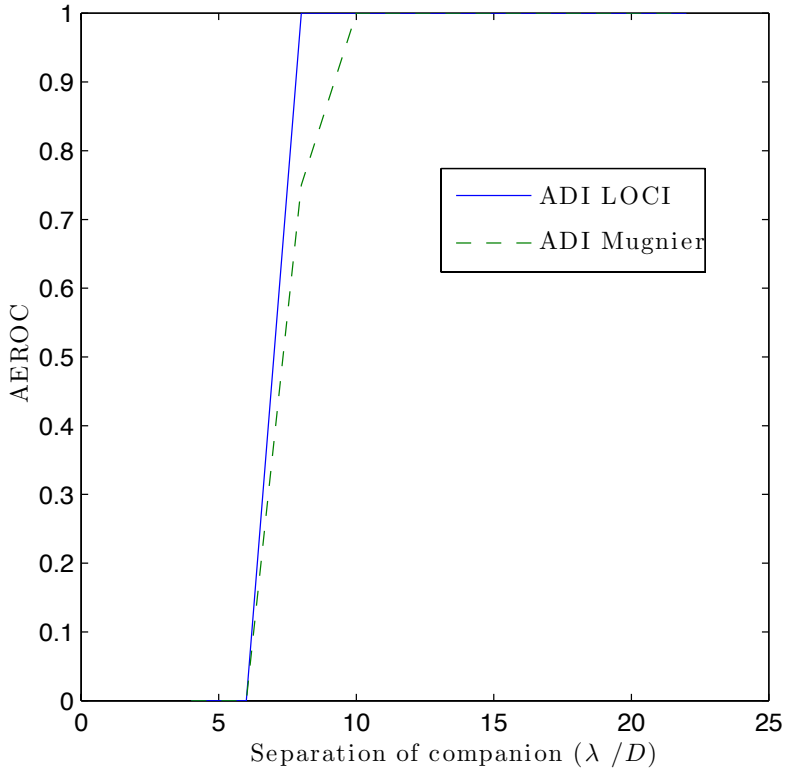


FIGURE 5.8: The area under the EROC curve as a function of the companion-star separation.

Figure (5.8) shows that beyond $6\lambda/D$ there does not appear to be a relationship between the AEROC and the angular separation of the companion. This is a logical result as the inner optimisation subsections have the same number of reference images as

the outer optimisation subsections and hence the level of speckle attenuation should be comparable. The number of reference images available to the algorithm does not depend upon angular separation in this simulation because the companion always moves 18° between frames. Therefore it was possible to reduce the size of the optimisation and reference subsections to single pixels. Figure (5.8) also shows that after $5\lambda/D$ the performance of the Mugnier two-by-two reduction algorithm is similar or better than the ADI LOCI approach.

5.3.5 Multi-wavelength ADI-LOCI

Differential imaging is not constrained to using angular differential data. If simultaneous multi-spectral data has been acquired in addition to angular differential data, see equation (5.15), the spectral data can be processed with the angular differential data to improve the level of quasi-static speckle attenuation.

$$\mathbf{I}(\lambda,\theta) = \begin{bmatrix} I(\lambda_1,\theta_1) & I(\lambda_2,\theta_1) & I(\lambda_3,\theta_1) \\ I(\lambda_1,\theta_2) & I(\lambda_2,\theta_2) & I(\lambda_3,\theta_2) \\ \vdots & \vdots & \vdots \\ I(\lambda_1,\theta_n) & I(\lambda_2,\theta_n) & I(\lambda_3,\theta_n) \end{bmatrix} \tag{5.15}$$

Where θ denotes the field of view orientation and n is the number of field of view angles. In this study the SDI and ADI methods were combined in series: first applying SDI, as in section (5.2), and then processing the residual SDI data with the ADI-LOCI algorithm. To investigate the effect of combing these differential imaging techniques a simulation was set up with the following parameters:

- SDI imaging wavelengths: $\lambda_1 = 1.578\mu m$, $\lambda_2 = 1.625\mu m$ and $\lambda_3 = 1.652\mu m$,
- Host star brightness $= 3 \times 10^7 e^-$
- Companion brightnesses: $\Delta m_{\lambda_1} = 10 = 3000e^-$, $\Delta m_{\lambda_2} = 16 = 12e^-$, $\Delta m_{\lambda_3} = 16 = 12e^-$,
- Companion separation $= 7 \times \lambda_1/D$.

For 300 realisations of noise (Poisson and detector readout) the simulated data were reduced using only the ADI-LOCI method and then the SDI double difference approach in combination with the ADI-LOCI algorithm. The variance of the residual noise plus background were then estimated directly from the two reduced images and inputed into

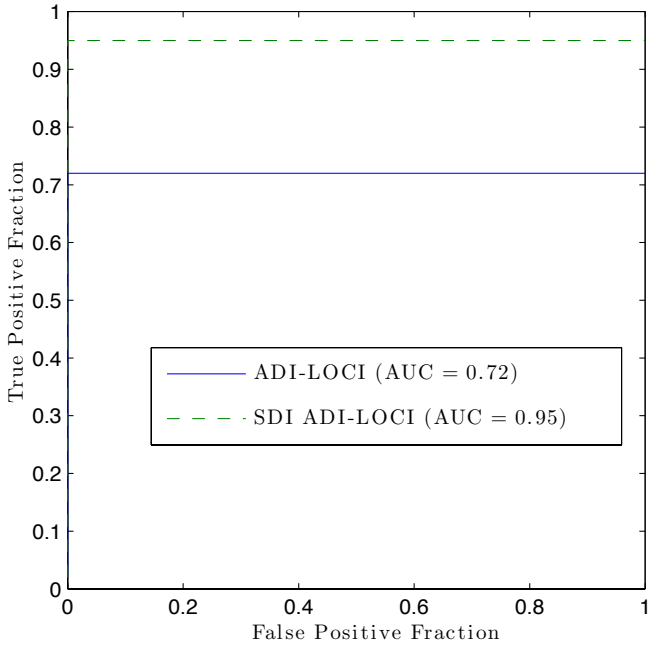


FIGURE 5.9: EROC curves for ADI-LOCI reduced data and SDI ADI-LOCI reduced data.

equation (2.11) for use in the Hotelling observer. The tolerance for a correct estimate of the companion location was set at $0.1\lambda_1/D$ and the tolerance on the estimate of the differential intensity was set at 0.1 magnitudes. An EROC curve was produced, see figure (5.9), and the AEROC computed for both reduction methods. The results from figure (5.9) show a significant increase in the AEROC when the data is processed with the SDI double differencing method in combination with the ADI-LOCI approach.

In a separate experiment four companions were placed into the same image with differential magnitudes of $\Delta m_{\lambda_1} = 10, 11, 12$ & 13 at a separation of $7\lambda_1/D$ from the central star with a 90° angular separation between each other, see figure (5.10). Each companion in turn had a differential magnitude of $\Delta m_{\lambda_2} = \Delta m_{\lambda_3} = 15, 16, 17$ & 18.

This data was reduced with SDI & ADI-LOCI and with ADI-LOCI alone. As before the required parameters for the data covariance matrix model were estimated directly from the residual. The Hotelling likelihood maps were calculated for both data reduction methods and are presented in figure (5.11) on a logarithmic scale.

The three brighter companions are apparent in the SDI-ADI LOCI Hotelling likelihood map however only the brightest source stands out above the residual noise in the ADI LOCI map. These Hotelling likelihood maps were then thresholded to reject Hotelling test statistic values under 3 standard deviations above the mean Hotelling noise level, see figure (5.12).

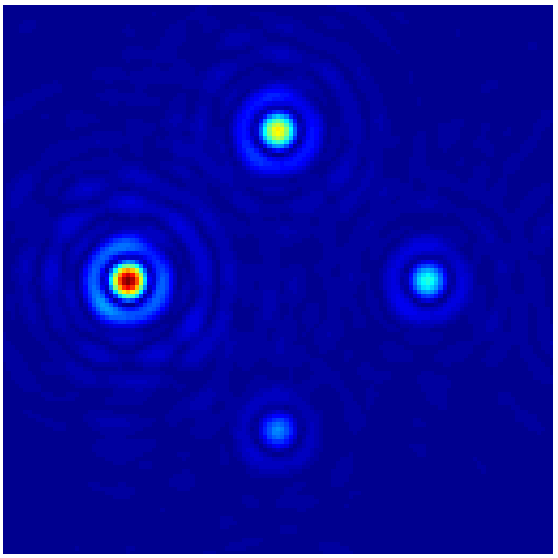
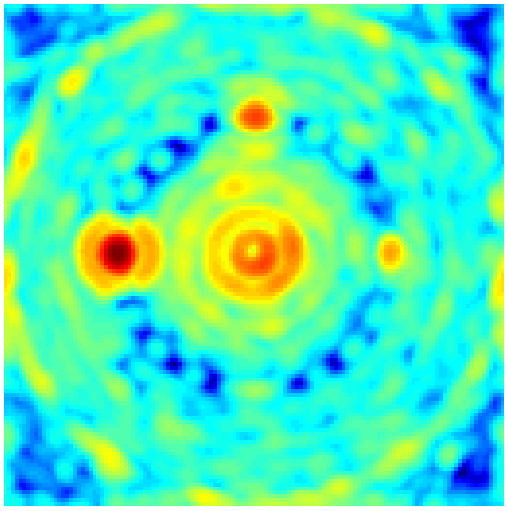
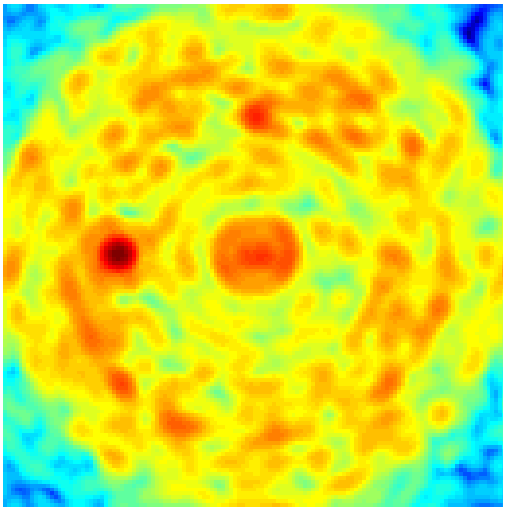


FIGURE 5.10: Four companions used to investigate using SDI in combination with ADI LOCI (square root scale)



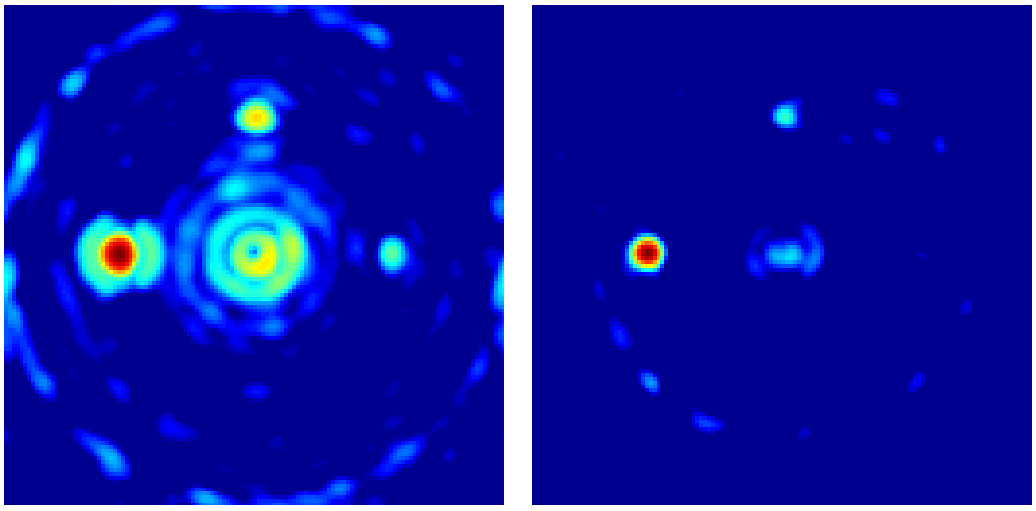
(a) Hotelling likelihood map for the data presented in figure (5.10) reduced using SDI and ADI-LOCI in series.



(b) Hotelling likelihood map for the data presented in figure (5.10) reduced using ADI-LOCI alone.

FIGURE 5.11: Hotelling likelihood maps for SDI & ADI-LOCI in combination and for ADI-LOCI alone, (logarithmic scale)

The thresholded SDI ADI-LOCI Hotelling likelihood map reveals the presence of the three brighter companions with a very weak signal at the location of the faintest companion. The thresholded ADI-LOCI Hotelling likelihood map also detects the two brighter companions, however the two fainter sources are not. It is therefore reasonable to conclude that the subtraction of properly scaled multi-spectral data can result in faint companion detections of an order of magnitude fainter that would otherwise be possible. Finally the SNR of the Hotelling likelihood maps at the radius of the companions was



(a) Thresholded Hotelling likelihood map for the map presented in figure (5.11(a)).

(b) Thresholded Hotelling likelihood map for the map presented in figure (5.11(b)).

FIGURE 5.12: Thesholded Hotelling likelihood maps for the data presented in figure(5.11), displayed on a logarithmic scale.

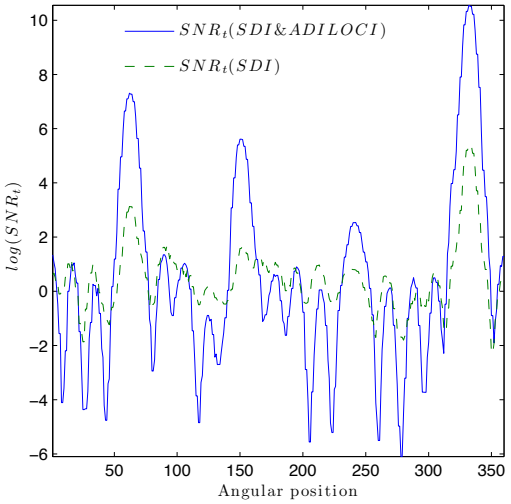


FIGURE 5.13: The logarithm of the Hotelling SNR at the radius of the companions for the two Hotelling likelihood maps above

estimated, see figure (5.13). This plot shows four clear peaks, one at the position of each companion, in the log-SNR of the ADI-LOCI Hotelling likelihood map whereas only two peaks are evident in the ADI-LOCI Hotelling likelihood log-SNR.

CHAPTER 6

PSF RECONSTRUCTION FROM MULTI-WAVELENGTH DATA

This chapter investigates an alternative approach to differential imaging to suppress quasi-static speckle noise and recover the PSF of the imaging system. This approach is based around using focal plane data to estimate the aberrated phase in the pupil plane of the telescope and hence recover the PSF via Fourier methods.

6.1 Image Formation

The image formation process can be expressed as the convolution of the system PSF and the observed object:

$$i(r) = o(r) * p(r) + n(r), \tag{6.1}$$

where r is the vector of focal-plane coordinates, i is the recorded image, o is an exact representation of the object in the focal plane, $p(r)$ is the system PSF, which is assumed to be isoplanatic i.e. it does not vary over the field, $*$ denotes convolution and n comprises all the additive noise terms.

The incoherent impulse response function of a system is the squared modulus of the inverse Fourier transform of the generalised pupil function:

$$p(r) = |FT^{-1}\left[H(\rho)\exp\{i\phi(\rho)\}\right]|^2, \tag{6.2}$$

where ϕ denotes the phase inside the pupil, $\rho = \lambda u$, λ is the imaging wavelength, u is the spatial frequency in radians and H is the binary pupil mask with a value of 1 inside the pupil and 0 everywhere else (Roddier, 1981).

With this image formation model the system PSF is fully described by the pupil phase ϕ. Furthermore equation (6.2) implies that the optical transfer function of the system can be expressed as the auto-correlation of the pupil function:

$$OTF(\rho) = H(u)\exp\{i\phi(\rho)\} \otimes H(\rho)\exp\{i\phi(\rho)\}, \tag{6.3}$$

where the correlation of two complex functions h_1 and h_2 is defined by:

$$[h_1 \otimes h_2](u) \triangleq \int h_1^*(t)h_2(t+x)dt. \tag{6.4}$$

6.2 Iterative PSF Estimation or the Phase Retrieval Problem

Fundamentally the phase retrieval problem aims to determine the phase of a complex function, i.e. the phase of an incoming wavefront, from information about its squared modulus, i.e. the system PSF. Using the image formation model above, equation (6.1) can be written in the Fourier domain as:

$$I = O \times P + N, \tag{6.5}$$

where capital letters denote the Fourier transformation of the function represented by the corresponding lower case letter i.e. I is the Fourier transform of i. This form suggests a simple deconvolution method - having a PSF, an estimate of the object could be obtained by inverting equation (6.5). Unfortunately this inversion leads to an amplification of noise, which is always present when making observations. This effect can be partly alleviated by using a Wiener filter in the inversion of equation (6.5) (Press et al., 1986).

If the pupil support of the system is known and the object being imaged is a point source there exists methods of estimating the pupil phase from a single focal plane intensity measurement (Gerchberg & Saxton, 1972; Gonsalves, 1976). This algorithm aims to

estimate a phase which best fits the known pupil and focal plane constraints via an iterative search method.

In AO-corrected astronomical observations the PSF is never known exactly. A further complication is that the AO-corrected PSF can have a complex asymmetric form. In this case both the object and the PSF have to be extracted from the data. We have already seen that this problem is referred to as blind deconvolution (Stockham et al., 1975). Ayers & Dainty (1988) proposed a Gerchberg-Saxton type iterative algorithm to solve for two positive functions, o and p which provide the best solution to equation (6.1).

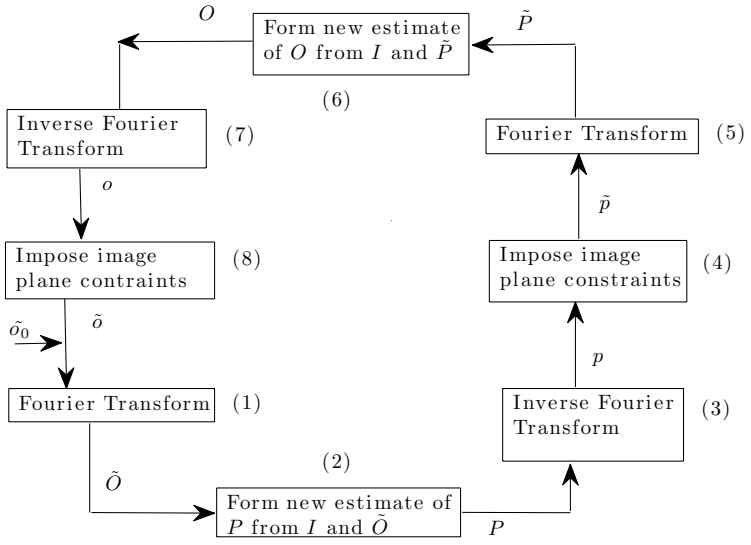


FIGURE 6.1: General deconvolution algorithm of Ayers & Dainty (1988).

The algorithm works in the following manner: firstly, a nonnegative-valued initial estimate $\tilde{o_0}$ is input to the iterative algorithm; this function is then Fourier transformed to yield $\widetilde{O}$; the tilde denotes an estimate of the function, $\widetilde{O}$ is then inverted to form an inverse filter and multiplied by I (equation (6.5)) to form P. This estimated Fourier spectrum is then inverse transformed to give p, the first estimate of the PSF. A non-negativity constraint is then applied to p and it is then Fourier transformed to give $\widetilde{P}$. This function is then inverted and multiplied by I to give the next estimate $\widetilde{O}_1$. The iterative loop is completed by inverse Fourier transforming $\widetilde{O}_1$ to give o_1 and applying a non-negativity constraint to yield $\widetilde{o}_1$. The loop is repeated until two positive estimates of o and p, which satisfy the required convolution i, are found. See figure (6.1).

Lane (1992) showed that this algorithm does not converge in a stable way, that is the solution is not unique. For example, an extra iteration of the Ayers-Dainty algorithm

may result in new estimates of o and p that are worse, both in terms of their error metric and visually, than the previous estimates. Therefore if the object being imaged or the pupil function are unknown additional information is required to recover the PSF.

6.3 PSF Reconstruction via Phase Diversity

Phase diversity refers to a method which can be used to infer phase aberrations in the pupil plane from focal plane image data. This methods allows one to estimate both an unknown observed object and the unknown phase aberrations of an imaging system. This technique was first proposed by Gonsalves & Childlaw (1979); Gonsalves (1982), later refined by Paxman et al. (1992); Paxman et al. (1996) and recently reviewed by Mugnier et al. (2006). The phase aberration parameters are estimated by minimising an objective function which requires at least two images of the object, but does not depend upon an object estimate. Typically one of the detected images is a conventional focal-plane image which is degraded by unknown aberrations such as those induced by atmospheric turbulence, misaligned optics or imperfections in the mirrors of the telescope. The second image of the same object is formed by perturbing these unknown aberrations in some known fashion e.g. adding a known amount of defocus or changing the observed wavelength (Ingleby & McGaughey, 2005; Lofdahl et al., 2001). The information provided by the second image leads to a unique solution for the unknown phase aberrations of the imaging system used to acquire the data.

Formally phase diversity is introduced into the second image by adding a known phase function into the generalised pupil function:

$$H_k = |H_k| \exp\{i[\phi + \theta_k]\}, \tag{6.6}$$

where ϕ is the unknown phase function to be estimated and θ_k is the known added phase function corresponding kth diversity image.

The unknown phase function ϕ can be expressed as the expansion on basis functions:

$$\phi(u) = \sum_{j=1}^{J} \alpha_j \theta_j(u), \tag{6.7}$$

where J is the number of coefficients used in the set $\{\alpha_j\}$ and $\{\theta_j\}$ is a set of basis functions such as complex exponentials which are orthogonal on a square support or the

set of discretised Zernike polynomials (Noll, 1976) for a circular aperture. As a large number of telescopes have circular pupils the Zernike polynomials are a convenient basis set as the polynomials form an orthogonal basis set on a unit disk. Each polynomial is produced by taking the product of a trigonometric function of polar angle θ with a function of radius r:

$$Z_i(r) = R_n^m(r)\Theta_n^m(\theta), \tag{6.8}$$

where:

$$\Theta_n^m(\theta) = \begin{cases} \sqrt{n+1} & \text{if } m = 0 \\ \sqrt{2(n+1)}\cos(m\theta) & \text{if } m \neq 0 \text{ and } i \text{ is even} \\ \sqrt{2(n+1)}\sin(m\theta) & \text{if } m \neq 0 \text{ and } i \text{ is odd} \end{cases} \tag{6.9}$$

and:

$$R_n^m(r) = \sum_{s=0}^{(n-m)/2} \frac{(-1)^s (n-s)!}{s!\,[(n+m)/2-s]!\,[(n-m)/2-s]!} r^{n-2s}. \tag{6.10}$$

The expansion of the phase using Zernike polynomials reads:

$$\phi(r) = \sum_{k=1}^{\infty} a_k Z_k(r). \tag{6.11}$$

However in practice the sum of the polynomials is limited to a finite number depending upon the problem. There is no loss in information in parameterising the phase in this manner as the basis functions could be composed of a point-by-point set of Kronecker delta functions to model the unknown phase function.

The problem is now as follows: given a set of diversity images $\{i_k\}$, a known phase diversity function $\{\theta_k\}$ and pupil functions $\{|H_k|\}$, estimate the object o and phase function basis coefficients α. If the dominant noise source in the data is modelled as an additive, independent and identically distributed random variable with a zero mean Gaussian PDF with a variance of σ_k^2 the normal probability density function is:

$$p(i_k; o, \alpha) = \frac{1}{(2\pi\sigma_k^2)^2} \exp - \left\{ \frac{(i_k - o * p_k)^2}{2\sigma_k^2} \right\}, \tag{6.12}$$

where $\tilde{p}_k$ is an estimate of the PSF in the kth diversity channel dependent upon the phase coefficients α. Maximising the logarithm of this likelihood function will lead to a maximum likelihood estimate for the object and phase parameters where the log-likelihood is given as:

$$L(o, \alpha) = -\sum_{k=1}^{K} [i_k - o * p_k]^2 \,. \tag{6.13}$$

Applying discrete versions of Parseval's theorem and the convolution theorem leads to the relation:

$$L(o, \alpha) = -\frac{1}{N^2} \sum_{k=1}^{K} |I_k - OP_k|^2, \tag{6.14}$$

where the diversity images are size $N \times N$ pixels and the capital letters denote discrete Fourier transforms of the corresponding lower case quanties.

At this point it should be noted that there may be an important difference in the signal to noise ratio in the diversity images. To compensate for this, in the case of $K = 2$, Löfdahl & Scharmer (1994) proposed to introduce a scaling parameter to compensate:

$$\gamma = \sigma_1^2 / \sigma_2^2. \tag{6.15}$$

where σ_1 and σ_2 are the RMS values of the noise in the two diversity images. The modified log-likelihood now reads as:

$$L(o, \alpha) = -\left\{ |I_1 - OP_1|^2 + \gamma |I_2 - OP_2|^2 \right\}. \tag{6.16}$$

Nonlinear optimisation algorithms could be directly applied to this modified log-likelihood function to estimate both the object and phase parameters. However this would require searching over a very large dimensional space of object and phase parameters. Paxmann (Paxman et al., 1992) noticed a significant reduction in the dimension of the required space for the case of two diversity images, $K = 2$. This reduction stems from deriving an objective function which only depends upon the phase parameters and not upon an object estimate. This is only possible due to the existence of a closed-form expression for the object which maximises the modified log-likelihood for a given set of phase functions:

$$O_M = \frac{I_1 P_1^* + I_2 P_2^*}{|P_1|^2 + \gamma |P_2|^2}, \tag{6.17}$$

where the asterisk as a superscript denotes complex conjugation. Substituting equation (6.17) into equation (6.16) results in an objective function which only depends upon the unknown phase parameters:

$$L_M(\alpha) = -\sum_u \frac{|I_1(u)P_2(u) - I_2(u)P_1(u)|^2}{|P_1(u)|^2 + \gamma |P_2(u)|^2}, \tag{6.18}$$

where it has been assumed that the estimates $\tilde{P}_1(u)$ and $\tilde{P}_2(u)$ do not both go to zero for the same u.

The Gonsalves objective function, equation (6.18) can be generalised to make use of an arbitrary number of wavelength diversity images. Paxman et al. (1992) shows that the general object which minimises the log-likelihood function, equation (6.12), for any number of diversity images is given by:

$$O_M = \frac{\sum_{k=1}^{K} I_k(u) P_k^*(u)}{\sum_{l=1}^{K} \gamma_l |P_l(u)|^2}, \tag{6.19}$$

where

$$\gamma_l = \sigma_1^2 / \sigma_l^2. \tag{6.20}$$

Furthermore they show what when equation (6.19) is substituted into equation (6.13) and dropping the $1/N^2$ scaling factor this results in a generalised objective function:

$$L_M(\alpha) = -\sum_u \frac{\sum_{j=1}^{K-1} \sum_{k=j+1}^{K} |I_j(u)P_k(u) - I_k(u)P_j(u)|^2}{\sum_{l=1}^{K} \gamma_l |P_l(u)|^2}. \tag{6.21}$$

In practice the objective function L_M is minimised numerically to obtain an estimate for the phase basis coefficients α. The minimisation is performed by a conjugate-gradient method which also employs a trust region (Byrd et al., 2000; Powell, 1978).

The bounds on the minimisation process are estimated empirically from the data. The Strehl ratio of the data is estimated, a random vector of Zernike coefficients is then generated which match this Strehl ration to $\pm 1\%$. The values of the basis coefficients are then bounded at twice the maximum value of this random vector of Zernike coefficients.

The minimisation process uses this random vector of Zernile coefficients as the initial starting point for the minimisation.

The minimisation terminates when the difference in the estimates of the phase coefficients from one iteration to the next approaches the precision of the machine.

6.4 Wavelength Diversity: An Application of Phase Diversity

We propose to use an approach referred to as Wavelength Diversity (Ingleby & McGaughey, 2005; Lofdahl et al., 2001) to estimate the residual AO corrected phase in the pupil of an astronomical telescope. Wavelength diversity is a variant of the classical phase diversity method presented above. Wavelength diversity introduces a known phase perturbation by changing the imaging wavelength by a known amount. Formally wavelength diversity alters the phase diversity equation (6.6) by scaling the unknown phase, ϕ, instead of adding a known phase function to it:

$$H_k = |H_k| \exp\{i[\phi\theta_k]\}. \tag{6.22}$$

Two diversity images were simulated at $\lambda_1 = 1.64\ \mu m$ and $\lambda_2 = 1.80\ \mu m$ with a known phase function, see figure (6.2(a)). The matlab function *fmincon* was then used to find minimum of equation (6.21) using the first 30 Zernike polynomials to model the pupil plane phase. A first guess at the vector of Zernike coefficients, $\alpha_{initial}$, was computed as a vector of zero mean Gaussian random numbers scaled to follow a f^{-2} power law. Bounds were placed upon the maximum and minimum values of the estimates of the Zernike coefficients and the phase in the pupil plane was bounded by a binary pupil mask. In this test the algorithm was given 27 random starts where the PSF with the minimum least square error between the estimated PSF and the data was taken as the best estimate. This result is shown in figure (6.2(a)). As can be seen in figure (6.2(a)) the error metric (equation (6.21)) appears to be insensitive to the sign on the even radial order Zernike coefficients. This error on the sign of the even radial order Zernike coefficients is to be expected since the PSF may be obtained by Fourier transforming the autocorrelation function of the complex amplitude in the pupil; if the pupil is symmetric, then the pupil autocorrelation of an even phase function does not depend on the sign of that function. This can be verified using the fact that the autocorrelation of the pupil complex amplitude, $a(u)$ and the autocorrelation of $a^*(-u)$ are equal. In this test we use a symmetric pupil function; therefore the algorithm does not retrieve a unique solution

to the phase. However the solution appears to be close enough to the global minimum to be a useful solution. In reality, it is probable that asymmetries in the pupil will remove this sign ambiguity.

The multi-spectral angular data, $I(\lambda, \theta)$, used in section (5.3.5), was processed angle-by-angle by the wavelength diversity algorithm in order to test the approach on a more realistic data set than above. The focus of this experiment is two fold: firstly; to investigate how well wavelength diversity algorithm estimated the pupil plane phase and secondly, to contast the SDI differencing method with the subtraction of the wavelength diversity PSF estimate.

For each field of view angle in the angular data, $\theta_{1,\ldots,n=11}$, the wavelength diversity algorithm estimated a pupil phase using the first three hundred Zernike polynomials and three corresponding PSF estimates, one per wavelength channel. For each wavelength channel the corresponding estimated PSF was used to subtract the signal of the central star. The three residual images were then medianly combined.

The set of n field of view residual images were then reduced using the ADI-LOCI algorithm to further attenuate any residual speckle noise. The final wavelength diversity ADI-LOCI residual image was operated upon by the Hotelling observer to extract differential astrometry and photometry of any faint companion present. The results of this operation are presented in table (6.1).

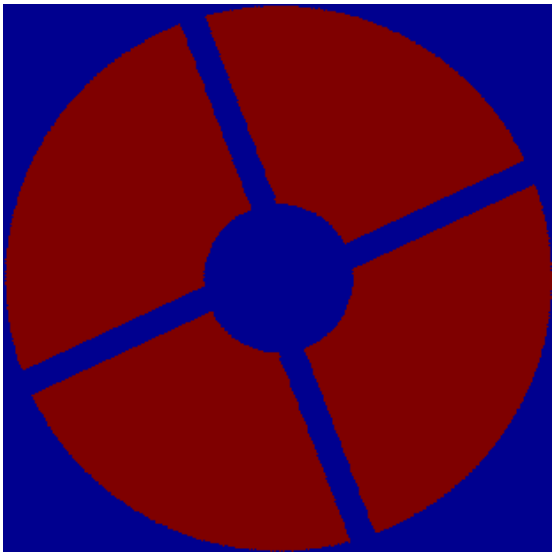
Table 6.1: SDI and ADI LOCI versus Wavelength diversity and ADI LOCI

	Error in Differential Astrometry (units of λ_1/D)	**Error in Differential Photometry (magnitudes)**
SDI ADI-LOCI	**0.057**	**0.29**
Wavelength diversity	**0.043**	**0.31**

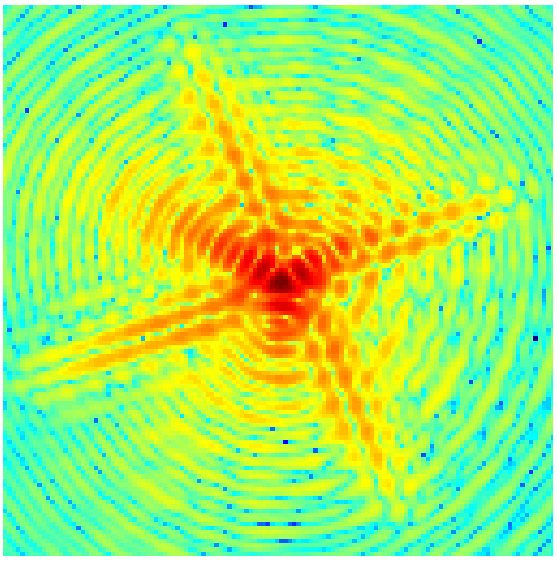
The AO corrected long exposure phase screen used to simulate the first field of view in the angular data, $I(\theta_1, \lambda_{1,2,3})$, is shown in figure (6.3(a)). The RMS of this residual phase screen was estimated at $37nm$. The wavelength diversity estimate of this phase screen is presented in figure (6.3(b)). The difference between the true phase screen and the estimated phase screen is presented in figure (6.3(c)). The estimated residual RMS difference between these phase screens was $6nm$. This residual phase error is comparable to the $5-10nm$ RMS residual static aberration which is required to detect warm Jupiters on a $8m$ class telescope (Mugnier et al., 2008).

To contrast the accuracy of the wavelength diversity phase estimation, the AO corrected phase screen was fitted directly with the first three hundred Zernike polynomials. The

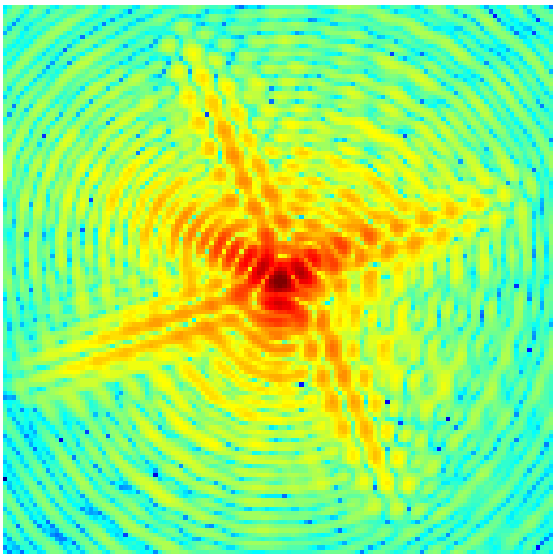
(a) Zernike Coefficient Estimates



(b) Gemini type pupil image



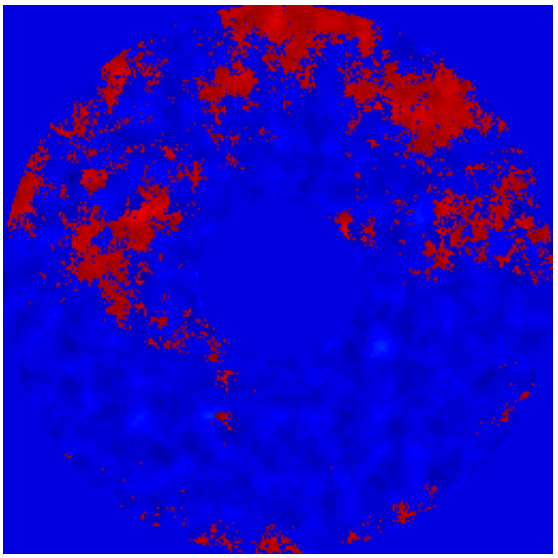
(c) PSF non-flipped Zernikes



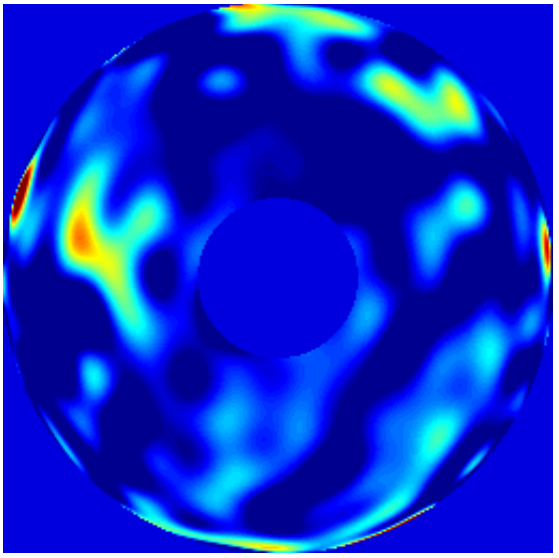
(d) PSF flipped Zernikes

FIGURE 6.2: Initial simulations of our wavelength diversity approach resulted in pupil phase estimates with a sign ambiguity on the even radial order Zernike coefficients. The corresponding PSF estimates, (c) & (d), were identical due to the fact that if the pupil is symmetric, then the pupil autocorrelation of an even phase does not depend on the sign of that function.

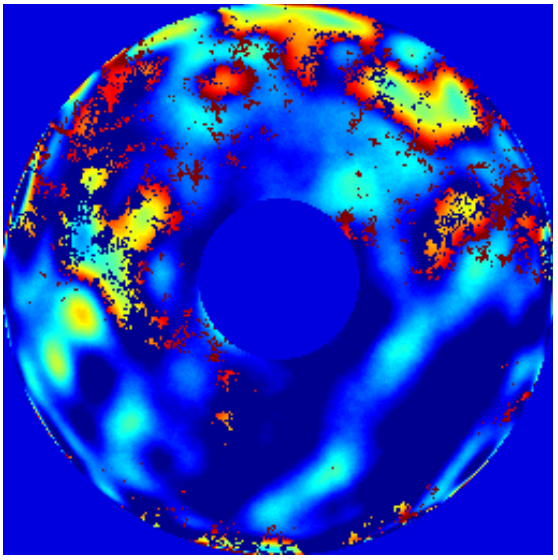
result of this fit is shown in figure (6.3(d)). When subtracted from the AO corrected phase screen the residual RMS error was estimated at $4nm$. Comparing this phase error to that of the wavelength diversity procedure shows that the wavelength diversity approach can recover the unkown AO corrected long exposure phase to a high degree of accuracy.



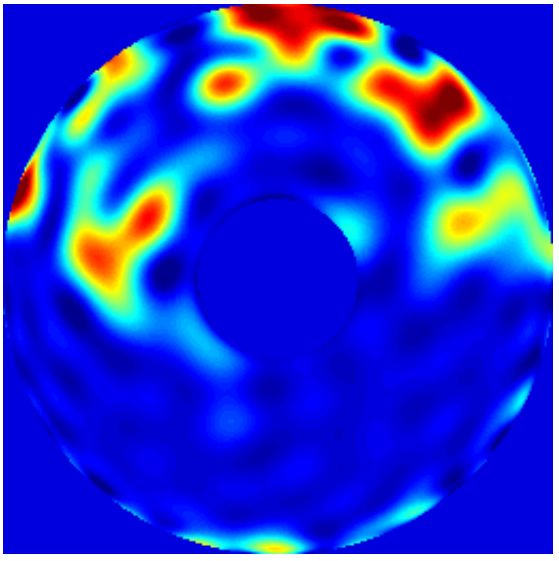
(a) AO corrected long exposure phase screen



(b) Wavelength diversity estimate of the phase in (a)



(c) Phase difference between the AO corrected phase screen and the Wavelength diversity phase estimate



(d) Zernike fit to the AO corrected phase screen in (a).

FIGURE 6.3: An AO corrected long exposure phase screen is presented in (a) along with the wavelength diversity estimate of this phase (b). The difference between these screens is shown in (c), the residual RMS difference of this phase screen was estimated at $6nm$. A direct Zernike fit to the AO corrected phase is shown in (d), the residual RMS difference between this screen and the AO corrected screen was estimated at $4nm$.

CHAPTER 7

CONCLUSION

7.1 Summary of Thesis Work

In this thesis the task of detecting, locating and estimating the intensity of a faint companion in adaptive optics corrected astronomical images has been reviewed. A general overview of the theory of light propagation through the turbulent atmosphere has been presented. The primary components of an adaptive optics system have been described, followed by an outline of the main factors which contribute to the difficulty in detecting a faint companion, i.e. differentiating the signal of a faint companion from residual quasi-static speckle noise. A framework for the application and assessment of the optimal linear observer, the Hotelling observer, to adaptive optics corrected astronomical data has been introduced. The main work during this thesis has been to investigate the performance of the Hotelling observer in detecting, locating and estimating the intensity of a faint companion in adaptive optics corrected astronomical images. This study has been carried out by numerical simulation of astronomical adaptive optics systems and also through the application of the observer to pseudo-real data. The residual quasi-static speckle noise problem has been addressed in two different approaches: differential imaging and PSF recovery from multi-wavelength data. A summary of the main investigation is presented below, followed by a discussion of topics which could benefit from future study.

Robustness of the Hotelling observer

The Hotelling observer was derived with the assumption that the noise in the data could be approximated by a Gaussian distribution, that the data covariance matrix was invertible and equal under both companion present and absent hypotheses and that the parameters of the data, e.g. star intensity or background level, could be correctly estimated from the data.

Using the adaptive optics simulation code PAOLA, the Hotelling observer was tested on varied data to try and establish the dependence of the observer upon the parameters of the data; when carrying out a detection task.

The results show that when the Hotelling observer is supplied with accurate estimates of the: star intensity, companion intensity, companion position, background level, detector readout variance and a precise PSF estimate a good model of the data covariance matrix is produced and along with a exact matched-filter template resulted in superior faint companion detection compared with a simple matched-filter or PSF subtraction approach. The results also show an almost linear increasing trend in the performance of the Hotelling observer with: companion brightness, companion separation from the central star and the level of adaptive optics correction. Apart from quasi-static speckle noise these parameters are the main contributors to the detection limits of an imaging system and hence as these constraints are relaxed it makes sense that the performance of the observer would increase accordingly. When the observer was feed incorrect estimates of the data parameters the only parameters which resulted in a noticeable drop in detection performance were mismatches in the companion position and mismatches in the PSF profile (Strehl ratio).

When the observer was tasked with also estimating the position and intensity of the companion the performance of the observer appeared to depend on the separation of the companion and the mismatch between the PSF profile of the data and the scanning template.

Binary star parameter estimation

A review of two state of the art algorithms for the computation of differential astrometry and photometry from adaptive optics corrected data is presented. The performance of the Hotelling observer was compared to these algorithms using observations of single star data obtained with the Lick observatory adaptive optics system on the 3m Shane telescope. The single star data was used to simulate artificial companions by scaling and shifting the single stars.

The experiments consisted of applying the Hotelling observer, the PSF-fitting algorithm StarFinder and the IBD algorithm Fitstars to data with properly matched and mismatched PSFs. The Hotelling observer always obtained the most accurate estimate for the position of the companion. In all cases, bar one, the Hotelling observer also returned the most accurate estimate for the intensity of the companion.

The superior performance of the Hotelling observer over the other two algorithms was shown via the PSNR to stem from prewhitening the PSF subtracted data to flatten the residuals.

The Quasi-static speckle problem

Instrumentally induced quasi-static speckle noise does not average out with increasing exposure time (Marois et al., 2005). Two complementary methods were investigated to deal with the two main drawbacks of applying the Hotelling observer: the inability of the observer to distinguish faint companions from point-like speckles and the necessity of an accurate PSF estimate.

A new data simulation algorithm, which simulated adaptive optics corrected Kolmogorov phase screen to generate long exposure PSFs was used to produce multi-wavelength and multi angular field of view data.

The first data reduction method consisted of using the spectral double differencing method of Marois et al. (2005), where a PSF was estimated by subtracting the double difference residual image from the original data, in combination with the Hotelling observer. The results of the data reduction show that the PSNR of the Hotelling data map is ≈ 2 times greater than the corresponding pixel residual image.

The second method involved using the angular differencing algorithm of Lafrenière et al. (2007). The ADI-LOCI algorithm was initially applied using a single wavelength channel. As part of the algorithm a PSF estimate was produced. Scanning the residual ADI-LOCI image with the Hotelling observer resulted in a faint companion ($\Delta m_k = 10$) detection, with an error in astrometry of $0.1\lambda/D$ and photometry error of 0.05 magnitudes.

When the spectral double differencing approach and the ADI-LOCI algorithm were used in series the PSNR in the residual Hotelling likelihood map increased by a factor of 10.

PSF recovery from multi-wavelength data

The final topic of this thesis investigated using multi-wavelength focal plane data to estimate the adaptive optics corrected phase in the pupil plane of an astronomical telescope and hence recover the system PSF at any wavelength via Fourier methods.

A general overview of the phase retrieval problem and phase diversity is presented. The framework for recovering the unknown corrected pupil phase from focal plane data is derived and the minimisation of multi-variate error metrics is reviewed. The approach used in this investigation denoted wavelength diversity is a generalisation of phase diversity which makes use of multi-wavelength data instead of defocus information. The phase was modelled using a Zernike decomposition.

Initial results using this approach revealed an insensitivity on the estimated sign of the even radial Zernike coefficients. This explained by the fact that the autocorrelation of the pupil complex amplitude, $a(u)$ and the autocorrelation of $a^*(-u)$ are equal.

Applying the wavelength diversity approach to realistic simulated data resulted in the accurate recovery of the system PSF and pupil phase. The computed PSF estimate, when used with the Hotelling observer, produced astrometry and photometry on par with the current state of the art differential imaging method. Furthermore the recovered pupil plane phase estimate closely matched a direct Zernike polynomial fit to the simulated phase.

Although the difficulties associated with applying the Hotelling observer to quasi-static speckle noise dominated AO-corrected images have not been fully resolved, the investigations in this thesis have set out a path for future study.

7.2 Future Work

If data is also recorded in time, the data set is a spatio-temporal random process. A spatio-temporal data covariance matrix can be estimated from the data and the corresponding spatio-temporal Hotelling observer applied to the data (Caucci et al., 2009). The data covariance matrix becomes three dimensional, two spatial and one time dimension, on the order of $N \times N \times J$, were each image is of size $N \times N$ and there are J recorded images. It would be interesting to investigate how this method would perform on data with dominant quasi-static speckle noise.

Concordantly simultaneous multi-wavelength data can be thought of as a spatio-spectral random process. It would be intriguing to estimate a three dimensional, two spatial and one spectral dimension, spatio-spectral data covariance matrix from multi-wavelength data to study the performance of the spatio-spectral Hotelling observer in the presence of quasi-static speckle noise.

The simulated adaptive optics corrected phase screens used in this thesis contained residual high spatial frequencies. In the investigations reported upon in section (6.4)

the unknown phase was modelled using Zernike basis functions as they are orthogonal on the unit circle, close to the optimal basis functions for describing a wavefront degraded by Kolmogorov turbulence, i.e. the Karhunen-Loeve functions (Fried, 1977; Lane & Tallon, 1992) and they are mathematically well defined and are tractable. However; a large number of polynomials, greater than 100, must be used to precisely model the phase. The phase model could be made simpler by trying to estimate the point-by-point phase values inside the telescope pupil. The basis functions could therefore be the Kronecker delta functions.

BIBLIOGRAPHY

Aime, C., & Soummer, R. 2004, The Astrophysical Journal, 612, L85

Artigau, E., Biller, B. A., Wahhaj, Z., Hartung, M., Close, L. M., Chun, M. R., Liu, M. C., G., T., Rigaut, F., Toomey, D. W., & Ftaclas, C. 2008, in Ground-based and Airborne Instrumentation for Astronomy II, ed. Ian S. McLean and Mark M. Casali, Vol. 7014 (Proc of SPIE), 70141Z

Ayers, G. R., & Dainty, J. C. 1988, Optics Letters, 13, 547

Babcock, H. W. 1953, Publications of the Astronomical Society of the Pacific, 65, 229

Baraffe, I., Chabrier, G., Barman, T. S., Allard, F., & Hauschildt, P. H. 2003, Astronomy and Astrophysics, 402, 701

Barnaby, D., Spillar, E., Christou, J. C., & Drummond, J. D. 2000, Astronomical Journal, 119, 378

Barrett, H. H., Abbey, C. K., & Clarkson, E. 1998, J. Opt. Soc. Am. A, 15

Barrett, H. H., Denny, J. L., Wagner, R. F., & Myers, K. J. 1995, J. Opt. Soc. Am. A, 12, 834

Barrett, H. H., & Myers, K. J. 2004, "Foundations of Image Science" (Weily Series in Pure and Applied Optics)

Barrett, H. H., Myers, K. J., Devaney, M. N., & Dainty, J. C. 2006, J. Opt. Soc. Am. A, 23, 3080

Barrett, H. H., Myers, K. J., Devaney, N., Dainty, J. C., & Caucci, L. 2006, in Proc of SPIE, ed. B. L. Ellerbroek & D. B. Calia, Vol. 6272

Bauman, B. J., Gavel, D. T., Waltjen, K. E., Freeze, G. J., Keahi, K. A., Kuklo, T. C., Lopes, S. K., Newman, M. J., & Olivier, S. S. 1999, Adaptive Optics Systems and Technology, 3762, 194

Bauman, B. J., Gavel, D. T., Waltjen, K. E., Freeze, G. J., Keahi, K. A., Kuklo, T. C., Lopes, S. K., Newman, M. J., & Olivier, S. S. 1999, Proc. SPIE, 3762, 194

Biller, B. A., Close, L. M., Masciadri, E., Nielsen, E., Lenzen, R., Brandner, W., McCarthy, D., Hartung, M., Kellner, S., Mamajek, E., Henning, T., Miller, D., Kenworthy, M., & Kulesa, C. 2007, The Astrophysical Journal Supplement Series, 173, 143

Blanc-Féraud, L., Mugnier, L. M., & Jalobeanu, A. 2010, Inverse Problems in Vision and 3D Tomography (ISTE), 97–121

Bloemhof, E. E. 2003, The Astrophysical Journal Letters, 582, L59

—. 2004, The Astrophysical Journal Letters, 610, L69

Bloemhof, E. E., Dekany, R. G., Troy, M., , & Oppenheimer, B. R. 2001, The Astrophysical Journal Letters, 558, L71

Burke, D., Gladysz, S., Roberts, L., Devaney, N., & Dainty, C. 2009, Publications of the Astronomical Society of the Pacific, 121, 767

Byrd, R., Gilbert, J., & Nocedal, J. 2000, Mathematical Programming, 89, 149

Cagigal, M. P., & Canales, V. F. 1998, Optics Letters, 23, 1072

Cagigal, M. P., & Canales, V. F. 2000, J. Opt. Soc. Am. A, 17, 903

Caucci, L. 2007, Point Detection and Hotelling Discriminant: An Application in Adaptive Optics, M.Sc. Thesis

Caucci, L., Barrett, H. H., Devaney, N., & Rodríguez, J. J. 2007, J. Opt. Soc. Am. A, 24, B13

Caucci, L., Barrett, H. H., & Rodriguez, J. J. 2009, Optics Express, 17, 10946

Christou, J. C., Gladysz, S., Redfern, M., W., B. L., & Roberts, Jr., L. C. 2006, Proc of The Advanced Maui Optical Space Surveillance Technologies Conference

Clarkson, E. 2007, J. Opt. Soc. Am. A, 24

Cornia, A., Mugnier, L. M., Mouillet, D., Vigan, A., Eggenberger, A., Rousset, G., Boccaletti, A., Carbillet, M., Dohlen, K., Fusco, T., Carson, J., & Montagnier, G. 2010, Proc. SPIE, 7736, 77361E

Davis, L. E. 1989, A User's Guide to the IRAF Apphot Package

Devaney, M. N. 1992, Astronomical Journal, 257

Diolaiti, E., Bendinelli, O., Bonaccini, D., Close, L., Currie, D., & Parmeggiani, G. 2000, Astronomy and Astrophysics, Supplimentary Series, 147, 335

Fitzgerald, M. P., & Graham, J. R. 2006, The Astrophysical Journal, 637, 541

Fletcher, R. 1987, "Practical Methods of Optimization" (Wiley, New York)

Flicker, R. C., & Rigaut, F. J. 2005, J. Opt. Soc. Am. A, 22, 504

Fried, D. L. 1965, J. Opt. Soc. Am., 55, 1427

—. 1966, J. Opt. Soc. Am., 56, 1380

—. 1977, J. Opt. Soc. Am., 67, 370

Gerchberg, R., & Saxton, W. 1972, Optik, 35, 237

Gladysz, S., & Christou, J. C. 2008, The Astrophysical Journal, 684

—. 2009, The Astrophysical Journal, 698, 28

Gladysz, S., Christou, J. C., Kenworthy, M., Law, N., & Dekany, R. G. 2008, Proc AMOS

Gladysz, S., Christou, J. C., & Redfern, M. 2006, Advances in Adaptive Optics II, 6272, 62720J

Gladysz, S., Christou, J. C., & Redfern, M. 2006, Proc. SPIE, 6272 I, 62720J

Gonsalves, R. A. 1976, J. Opt. Soc. Am., 66, 961

—. 1982, Optical Engineering

Gonsalves, R. A., & Childlaw, R. 1979, Proc. SPIE, 207, 32

Goodman, J. W. 1975, Laser Speckle and Related Phenomena (Berlin: Springer)

Hardy, J. W. 1998, Adaptive optics for astronomical telescopes (Oxford University Press)

Helstrom, C. W. 1968, Statistical Theory of Signal Detection (Pergamon)

Hinkley, S., Oppenheimer, B. R., Soummer, R., Sivaramakrishnan, A., Roberts, L. C., Jr., Kuhn, J., Makidon, R. B., Perrin, M. D., Lloyd, J. P., Kratter, K., , & Brenner, D. 2007, Astronomical Journal, 654, 633

Hodapp, K. W., Jensen, J. B., Irwin, E. M., H., Y., Chung, R., Flectcher, K., Robertson, L., Hora, J. L., Simons, D. A., Mays, W., Nolad, R., Bec, M., Merrill, M., & Fowler, A. M. 2003, Publications of the Astronomical Society of the Pacific, 115, 1388

Högbom, J. A. 1974, Astronomy and Astrophysics, Supplimentary Series, 15, 417

Horch, E., Ninkov, Z., & Franz, O. G. 2001, Astronomical Journal, 121, 1583

Ingleby, H. R., & McGaughey, D. R. 2005, Optics Letters, 30, 489

Jefferies, S. M., & Christou, J. C. 1993, The Astrophysical Journal, 415, 862+

Johnston, R. A., & Lane, R. G. 2000, Applied Optics, 39, 4761

Jolissaint, L., Véran, J.-P., & Conan, R. 2006, J. Opt. Soc. Am. A, 23, 382

Kasper, M. E., Beuzit, J.-L., Verinaud, C., Yaitskova, N., Baudoz, P., Boccaletti, A., Gratton, R. G., Hubin, N., Kerber, F., Roelfsema, R., Schmid, H. M., Thatte, N. A., Dohlen, K., Feldt, M., Venema, L., & Wolf, S. 2008, Proc. SPIE, 7015

Khurd, P., & Gindo, G. 2005, IEEE Trans. Med. Imaging, 24

Kolmogorov, A. 1941, Doklady Akademii Nauk SSSR, 30, 301

Lafrenière, D., Marois, C., Doyon, R., Nadeau, D., & Artigau, É. 2007, The Astrophysical Journal, 660, 770

Lane, R. G. abd Glindermann, A., & Dainty, J. C. 1992, Waves in Random and Complex Media (Taylor & Francis), 209–224

Lane, R. G. 1992, J. Opt. Soc. Am. A, 9, 1508

Lane, R. G., & Tallon, M. 1992, Applied Optics, 31, 6902

Le Bigot, E. O., & Wild, W. J. 1999, J. Opt. Soc. Am. A, 16, 1724

Lee, R. W., & Harp, J. C. 1969, Proc. IEEE, 57, 375

Leroux, C., & Dainty, C. 2010, Optics Express, 18, 1197

Lloyd, J. P., Liu, M. C., Macintosh, B. A., Severson, S. A., Deich, W. T. S., & Graham, J. R. 2000, Optical and IR Telescope Instrumentation and Detectors, 4008, 814

Lofdahl, M. G., Berger, T. E., & Seldin, H. 2001, Astronomy and Astrophysics, 377, 1128

Löfdahl, M. G., & Scharmer, G. B. 1994, Astronomy and Astrophysics, Supplimentary Series, 106, 243

Mackey, R. 2008, Adaptive Optics in Strong Turbulence for Application to Free Space Optical Communication, Ph.D. Thesis

Marois, C., Doyon, R., Nadeau, D., Racine, R., Riopel, M., Vallée, P., & Lafrenière, D. 2005, Publications of the Astronomical Society of the Pacific, 117, 745

Marois, C., Doyon, R., Racine, R., & Nadeau, D. 2000, Publications of the Astronomical Society of the Pacific, 112

Marois, C., Lafreniere, D., Doyon, R., Macintosh, B., , & Nadeau, D. 2006, The Astrophysical Journal, 641, 556

Marois, C., Racine, R., Doyon, R., Lafrenière, D., & Nadeau, D. 2004, The Astrophysical Journal Letters, 615, L61

Martin, J. M., & Flatte, S. M. 1990, J. Opt. Soc. Am. A, 7, 838

Meyer, M. R., Edwards, S., Hinkle, K., & Strom, S. 1998, The Astrophysical Journal, 508

Mugnier, L. M., Blanc, A., & Idier, J. 2006, Advances in Imaging & Electron Physics, Vol. 141 (Elsevier Inc.), 1–76

Mugnier, L. M., Conan, J.-M., Fusco, T., & Michau, V. 1998, Bayesian Inference for Inverse Problems, 3459, 50

Mugnier, L. M., Cornia, A., Sauvage, J.-F., Rousset, G., Fusco, T., & Védrenne, N. 2009, J. Opt. Soc. Am. A, 26, 1326

Mugnier, L. M., Fusco, T., & Conan, J.-M. 2004, J. Opt. Soc. Am. A, 21, 1841

Mugnier, L. M., Sauvage, J.-F., Fusco, T., Cornia, A., & Dandy, S. 2008, Optics Express, 16, 18406

Noll, R. J. 1976, J. Opt. Soc. Am., 66, 207

Obukhov, A. 1949, Izv. Akad. Nauk SSSR, Ser. Geogr. geofiz, 13:58

Pantin, E., Starck, J. L., & Murtagh, F. 2007, Blind image deconvolution: theory and applications (CRC Press), 277–317

Paxman, R. G., Schulz, T. J., & Fienup, J. R. 1992, J. Opt. Soc. Am. A, 9, 1072

Paxman, R. G., Seldin, J. H., Loefdahl, M. G., Scharmer, G. B., & Keller, C. U. 1996, The Astrophysical Journal, 446

Pluzhnik, E. A. 2005, Astronomy and Astrophysics, 431, 587

Powell, M. 1978, Nonlinear Programming 3, Academic Press

Poyneer, L. A. 2003, Applied Optics, 42, 5807

Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 1986, Numerical Recipes - The art of scientific computing (Cambridge University Press)

Racine, R., Walker, G. A. H., Nadeau, D., Doyon, R., & Marois, C. 1999, Publications of the Astronomical Society of the Pacific, 111, 587

Rigaut, F. J., Veran, J.-P., & Lai, O. 1998, Adaptive Optical System Technologies, 3353, 1038

Roberts, Jr., L. C., Turner, N. H., Bradford, L. W., ten Brummelaar, T. A., Oppenheimer, B. R., Kuhn, J. R., Whitman, K., Perrin, M. D., & Graham, J. R. 2005, The Astrophysical Journal, 130, 2262

Roberts, Jr., L. C., Turner, N. H., & ten Brummelaar, T. A. 2007, The Astrophysical Journal, 133, 545

Roddier, F. 1981, in Progress in optics. Volume 19. Amsterdam, North-Holland Publishing Co., 1981, p. 281-376., ed. E. Wolf, Vol. 19

Roddier, F. 1988, in Diffraction-Limited Imaging with Very Large Telescopes, Proceedings of the NATO Advanced Study Institute, held in Carge, Corsica, ed. D. M. Alloin & J. M. Mariotti

Roddier, F. 1999, Adaptive optics in astronomy (Cambridge University Press)

Roggemann, M. 1996, Imaging Through Turbulence (CRC Press)

Rosenthal, E. D., Gurwell, M. A., & P., H. P. T. 1996, Nature, 384

Rousset, G. 1993, Adaptive optics for astronomy (Springer), 115–137

Roux, F. S. 1995, J. Opt. Soc. Am. B, 12, 1215

Schneider, G., & Silverstone, M. D. 2003, Proc. SPIE, 4860, 1

Shames, P. M., & Tody, D. 1986, A User's Introduction to the IRAF Command Language

Simons, D. A., Robertson, D. J., & Mountain, C. M. 1995, Infrared Detectors and Instrumentation for Astronomy, 2475, 296

Sivaramakrishnan, A., Lloyd, J. P., Hodge, P. E., , & Macintosh, B. A. 2002, The Astrophysical Journal Letters, 581, L59

Soummer, R., Ferrari, A., Aime, C., & Jolissaint, L. 2007, The Astrophysical Journal, 669, 642

Stetson, P. B. 1987, Publications of the Astronomical Society of the Pacific, 99

—. 1992, ASPC, 25

Stetson, P. B., Davis, L. E., & Crabtree, D. R. 1990, in Astronomical Society of the Pacific Conference Series, Vol. 8, CCDs in astronomy, 289–304

Stockham, T. G., Cannon, T. M., & Ingebretsen, R. B. 1975, Proceedings of the IEEE, 63, 678

T. Fusco, J.-P. Véran, J.-M. Conan, & L.M. Mugnier. 1999, Astronomy and Astrophysics, Supplimentary Series, 134, 193

Tatarskii, V. 1961, Wave Propagation in a Turbulent Medium (Translated from Russian by R.A. Silverman, McGraw-Hill)

ten Brummelaar, T., Mason, B. D., McAlister, H. A., Roberts, Jr., L. C., Turner, N. H., Hartkopf, W. I., & Bagnuolo, Jr., W. G. 2000, Astronomical Journal, 119

ten Brummelaar, T. A., Mason, B. D., Bagnuolo, Jr., W. G., Hartkopf, W. I., McAlister, H. A., & Turner, N. H. 1996, Astronomical Journal, 112

Thalmann, C. 2009, A short introduction to Angular Differential Imaging, http://www.mpia.de/homes/thalmann/adi.htm

Turner, N. H., ten Brummelaar, T., Roberts, L. C., Mason, B. D., Hartkopf, W. I., & Gies, D. R. 2008, The Astrophysical Journal, 136, 554

Tyson, R. K. 1998, Principles of adaptive optics (Academic Press)

Véran, J., Rigaut, F., Maître, H., & Rouan, D. 1997, J. Opt. Soc. Am. A, 14, 3057

Wild, W. J., & Le Bigot, E. O. 1999, Optics Letters, 24, 190