Aberration retrieval for laser scanning microscopy

by

Pieter Henry Smid

submitted to The University of Nottingham

for the degree of Doctor of Philosophy

January 2019

University of Nottingham

UK | CHINA | MALAYSIA

To my beloved parents and in gratitude to my former supervisor

Prof. Nikolai Petersen

Abstract

This thesis explores different indirect wavefront sensing methods for detection of primary Zernike aberrations in laser scanning microscopes. All the presented aberration retrieval methods rely on analysing intensity distributions in the focal region and within the first dark ring of the Airy spot.

First a home-built reflection confocal microscope with a Zernike modal wavefront sensor is discussed (Chapter 3). Aberrations present in the sample or imaging system are measured indirectly by sequentially applying Zernike modes with a deformable membrane mirror (DMM) while maximising the detected intensity signal at the pinhole of the confocal microscope. When maximum intensity is reached at the pinhole, the Zernike mode(s) imposed by the DMM correct for the wavefront aberrations present in the sample and imaging system. The sensitivity of this modal method for measuring aberrations is discussed as a function of pinhole size for different Zernike modes and the difference between modal wavefront sensing in reflection and fluorescence is considered.

Large aberrations present a challenge for modal wavefront sensors since they can give rise to incorrect measurements due to cross-talk effects between the different Zernike terms. A way to solve this problem is to run through several correction iterations. This thesis proposes a new extension for modal wavefront sensing to tackle large sample induced aberrations more efficiently (Chapter 4). The new method uses an initial wavefront pre-correction which is based on a ray-tracing simulation of the sample. As a result, the number of iteration steps required is significantly reduced because the pre-correction removes the most relevant large aberrations present, thereby increasing the speed of the overall correction process.

Odd aberrations such as coma cannot be detected and corrected for in a reflection microscope because of a double-pass effect where the in-going light path and the

return light path pass of different sides of any element present in the system such as the DMM. As a result of this effect, odd aberrations are cancelled out after the second pass through the system and the measurement/correction system is not able to detect the presence of them. Nevertheless, the focal spot at the image plane will suffer from odd aberrations and these will affect the imaging performance of the microscope. A new method is presented in this thesis to break up the double-pass effect and allow odd aberrations to be detected and corrected for in a reflection confocal microscope (Chapter 5). To achieve this the beam is scanned across an edge and the edge response is used to determine the aberrations present as opposed to looking just at the intensity passing through the confocal pinhole. This method is illustrated by looking at coma, a common odd aberration found in optical microscopy. It is shown that the image of the edge (edge response) displays a characteristic distortion which is typical of coma and the amount of coma present in the imaging system can be estimated from the edge response curve.

Finally a novel aberration retrieval method is presented. This method is aimed at retrieving the amplitude of primary Zernike aberrations (astigmatism, coma, spherical aberration) in the pupil (Chapter 6). The primary Zernike aberrations are retrieved by fitting a set of orthogonal circle functions within the central region of the intensity distribution recorded at up to 3 different image planes, typically taken at focus and then either side of focus. Characteristic combinations of aberration sensitive fitting coefficients (so-called aberration indicators) are derived for each primary aberration (astigmatism, coma, spherical aberration) and it is shown that these indicators can be used for aberration retrieval. Importantly for aberration retrieval the indicators are selected so that there is a linear relationship between the aberration amplitudes and their respective indicators up to amplitude values of about 0.13λ. The issue of aberration cross-talk (when several aberrations are present) is also addressed and it is concluded that the new aberration retrieval method is successful as long as the rms wavefront deviation of all primary aberrations remains below 0.1λ. Benefits of this new approach as opposed to techniques such as the Gerchberg-Saxton algorithm are that it is fast, uses less intensity images and is non-iterative.

In summary, this PhD project makes new contributions to the field of aberration retrieval and adaptive optics in scanning microscopy by i) improving the modal aberration correction technique using an initial pre-corrected wavefront to significantly speed up the aberration correction procedure, ii) overcoming a double-pass cancellation issue in a reflection confocal microscope when looking at odd aberrations by using an edge scan to determine the odd aberrations present and iii) proposing a new phase retrieval technique that uses aberration indicators to retrieve the primary aberrations present in the pupil by looking at no more than three intensity images.

Acknowledgements

I would like to express my gratitude to my supervisors Dr. Amanda J. Wright and Dr. Chung W. See for their help and support throughout this thesis project. Many thanks to both of you to have make this possible and your encouragement to do my own independent research.

I am also grateful for comments and stimulating discussion on phase retrieval with Prof. Mike G. Somekh. A special thanks to Dr. Richard Smith and Dr. Baptiste Jayet for their help and advices on technical and safety questions. I am furthermore indebted to Dr. Solomon Ndiyang who taught me how to programme in LabVIEW.

I will never forget the good moments we shared together in the Optics & Photonics Research group over the last four years with my friends and colleagues. In particular: Dr. David Jung; Harriet Boyd; Dr. Mina Mossayebi; Dr. Sidahmed Abayazeed; and Victoria Ciampani. This experience would not have been the same without you.

I thank my parents for their love and support during this thesis. Without them I wouldn't have come that far. A very special thanks goes to my former supervisors Prof. Nikolai Petersen and Prof. Valery Shcherbakov who have taught me so much while studying Geophysics in Munich.

Conferences

- Poster "Aberration Correction in Confocal Microscopy using Adaptive Optics" at the DGAO (Deutsche Gesellschaft für Angewandte Optik) 2018 conference in Aalen
- Oral Presentation "Modal aberration correction in confocal microscope with CCD camera detection" at the conference OSA Imaging Systems and Applications 2018 Orlando Florida

Publications

- P. Smid, C. See, and A. J. Wright, "Modal aberration correction in confocal microscope with CCD camera detection," in Imaging and Applied Optics 2018 (3D, AO, AIO, COSI, DH, IS, LACSEA, LS&C, MATH, pcAOP), OSA Technical Digest (Optical Society of America, 2018), paper JTu5B.6.
- P. Smid, C. See, and A. J. Wright, “Adaptive optics modal correction of large aberrations using a ray-tracing pre-correction”, -in preparation-
- P. Smid, C. W. See, A. J. Wright, "Detection of odd aberrations in confocal reflection microscopy by means of an edge scan “, Optics Letters -in review-
- P. Smid, C. W. See, A. J. Wright, "Aberration retrieval based on the shape of intensity distributions in the vicinity of focus", -in preparation-

Contents

Glossary of abbreviations

Abbreviation	Term
AO	Adaptive optics
DMM	Deformable membrane mirror
ENZ	Extended Nijboer-Zernike theory
FFT	Fast Fourier transform
LSM	Laser scanning microscope
PSF	Point spread function
NA	Numerical aperture
SH	Shack-Hartmann sensor
SLM	Spatial light modulator

1. Introduction

1.1 Preface

Microscopes have been used for centuries to study and magnify images of microscopic samples. Zacharias Janssen is recognised as the inventor of the microscope in around 1590 in the Netherlands [1]. Thereafter, in 1676, Anthonie van Leeuwenhoek used his microscope for the observation of bacteria [2]. It was in the late 19th century, that Ernst Abbe developed a diffraction theory for describing the imaging process in a microscope, while working at the nowadays well known optics company Carl Zeiss Ag [3]. The 20th century saw the emergence of new types of microscopes such as the phase contrast microscope (which by means of a phase ring in the microscope's objective pupil, can image phase objects such as biological samples with increased contrast) and laser scanning microscopes (e.g. the confocal microscope). At the beginning of the 21st centuries, so-called super-resolution microscopes made their appearance. Examples of these types of microscopes are structured illumination, PALM (Photoactivated localization microscopy) and STED (Stimulated emission depletion microscopy) microscopes [4-6], whose resolution is below the classic theoretical diffraction limited resolution.

Although optical microscopes are designed to come close to their theoretical limits, residual system and sample induced aberrations will always eventually degrade the imaging performance. The theory of aberrations is closely related to the development of microscopy. Significant contributions in the field of optical aberrations were by Seidel who derived analytical expressions for primary aberrations (astigmatism, coma and spherical aberration) using geometrical optics, which were later named after him (Seidel aberrations) [7]. The diffraction theory of aberrations was pushed forward by Zernike and Nijboer [8-10]. Zernike introduced the famous Zernike aberration polynomials for his phase contrast method [10]. The Zernike aberration functions gained great popularity due to their mathematical properties (orthogonality over the unit circle, balanced aberration polynomials). A distorted wavefront in a circular aperture can be uniquely expressed in terms of a weighted series of orthogonal Zernike polynomials. Zernike polynomials are extensively used in different areas of optics, including optical design, adaptive optics, optical astronomy and wavefront sensing. In the beginning of the 21th century Braat and Janssen developed the extended Nijboer-Zernike theory, which provides semi-analytical expressions for describing the amplitude distribution around focus in the presence of aberrations [11-13]. The most commonly encountered aberration in microscopy is spherical aberration, which is often caused by a refractive index mismatch.

The use of adaptive optics in microscopy has helped to combat aberrations when imaging deep into samples by dynamically correcting for system and/or sample induced aberrations [14]. Here, an adaptive optics element corrects for aberration by shaping the wavefront with an equal but opposite distortions to produce an aberration-free image.

1.2 Motivation and objectives

The motivation behind this project was to explore novel aberration retrieval methods for microscopy with a particular emphasis on laser scanning microscopes. Although microscopes are designed to allow for imaging close to their theoretical limit, aberrations are often the limiting factor. Aberrations worsen the imaging performance and the results are blurry images with a loss of contrast and resolution. Measuring wavefront aberrations (i.e. the phase) is challenging because the phase information can not be directly derived from data obtained from intensity sensors. Interferometric devices or wavefront sensors are often used to quantify aberrations. To measure aberrations with a wavefront sensor or interferometer, often a modification or partial re-alignment of the optical components within the microscope is necessary. Wavefront sensing devices impose certain experimental requirements to ensure accurate wavefront measurements, such as, size of the light source, degree of coherence of the light source. In this thesis indirect wavefront sensing methods are studied which solely use the intensity signals measured by a camera. Some of these indirect wavefront sensing methods are not limited to laser scanning microscopes only but could also be used in a wider range of optical imaging systems, such as telescopes or optical lithography projection lenses.

1.3 Main contributions

My main contributions to the field of optics are in the development of novel methods to detect aberrations in laser scanning microscopes. The main contribution concerns a method which retrieves the primary Zernike aberrations from 3 intensity distribution images taken in the vicinity of focus. It is a non-iterative and relatively simple method. The method is not limited to certain types of microscopes but could be used in other optical imaging systems, such as telescopes and lithography projection lenses (although further work would be needed to test the capability of the method for retrieving aberrations in high numerical aperture (NA) systems). Furthermore, an optimisation strategy to tackle large aberrations more efficiently has been proposed. This method uses pre-corrections obtained from ray-tracing simulations of a sample. This optimisation strategy mitigates the effect of cross-talk between large aberrations and speeds up the correction process. Last but not least, a so-called double-pass effect which causes cancellation of odd aberrations in epi-illumination microscopes after reflection makes the detection of odd aberrations in reflection challenging. Nevertheless the focal spot in a scanning microscope would suffer from odd aberrations. An edge scan method is presented to break up the double-pass effect and allows detection of odd aberrations such as coma in reflection confocal microscopes. The edge scan method does not require additional hardware and can be used to detect, for example, coma in commercial confocal microscopes.

1.4 Synopsis

The main topic of this PhD thesis concerns aberration retrieval for scanning optical microscopy. The aberration retrieval methods considered are based on analysis of the intensity distribution in the vicinity of focus. Aberrations can be inferred from observation of the the intensity at the central part of a spot, or, alternatively, by determining the shape of the intensity distribution within a larger part of the Airy disk. The contents of the subsequent thesis chapters are briefly summarised below.

A review on laser scanning microscopes, optical aberrations, wavefront sensing methods and adaptive optics is given in Chapter 2.

In Chapter 3 a home-built laser scanning microscope with integrated adaptive optics for aberration retrieval is presented. The different components, microscope objective, lenses, detector, deformable and membrane mirror are discussed. A closed-loop calibration method of the deformable membrane mirror is explained and the quality of Zernike modes produced after calibration is shown. System aberrations were corrected with the calibrated deformable membrane mirror and the improvement on the axial and lateral point spread function of the microscope are shown.

In Chapter 4 a so-called confocal modal wavefront sensor is assessed with the home-built microscope described in Chapter 3. In modal wavefront sensing, aberrations are measured indirectly by optimisation of the confocal signal with an adaptive optics element (such as a deformable mirror) for each aberration separately. When the signal is at a maximum, all aberrations are corrected by the deformable mirror. The sensitivity of, and the differences between, modal sensing in reflection and in fluorescence are discussed. In the presence of large aberrations, modal sensing can become challenging due to cross-talk between aberrations, which often requires several optimisation iterations. A method is proposed to tackle large aberrations more efficiently by using ray-tracing pre-corrections.

Chapter 5 describes a method to detect coma in confocal reflection microscopy. The double-pass effect in reflection causes the cancellation of odd aberrations such as coma. By scanning the focal spot over an edge, the presence of coma can be detected.

The resulting edge response depends on coma and allows estimation of the amount of coma present.

Finally, in Chapter 6 a novel aberration retrieval method is presented. The method is aimed at retrieving primary Zernike aberrations by fitting a set of orthogonal functions to the intensity distribution within the first dark ring of the Airy spot in the vicinity of focus. Using characteristic combinations of fitting coefficients, the primary aberrations can be retrieved. The method works best in the presence of a single aberration, but can also retrieve multiple aberrations up to a certain extent.

In appendices details are given on: A) ray-tracing; and B) diffraction calculations using the extended Nijboer-Zernike theory (for the theoretical treatment of the aberration retrieval method described in Chapter 6); and C) camera images of intensity distributions through focus in the presence of primary Zernike aberrations are shown.

2. Literature review

This chapter aims to give an overview on wavefront aberrations and how they manifest themselves in optical imaging systems. The chapter starts with introducing scalar diffraction theory and how it can used to calculated the intensity distributions of aberrated beams in the vicinity of focus. Thereafter, the Zernike aberration polynomials, commonly used to describe wavefront aberrations in optics, are presented. Since this thesis is concerned with aberration retrieval for laser scanning microscopy, the imaging properties of such types of microscopes, with special emphasis on the popular confocal microscope, will be reviewed. At the end of the chapter it will be explained how wavefront aberrations can be dynamically corrected with the help of adaptive optics. Adaptive optics is a technology originating from optical astronomy where it is used in ground-based telescopes to correct atmosphere-induced wavefront aberrations. After briefly reviewing the use of adaptive optics in astronomy, the focus will be set on adaptive optics correction strategies in microscopy. The literature review chapter summarises what is known and constitutes the current state-of-the-art in the field of aberration retrieval and adaptive optics in optical microscopy.

2.1 Diffraction theory

In this section diffraction theory, coordinate systems definitions and the normalised optical coordinates used are explained. The Huygens-Fresnel principle describes electromagnetic field propagation through space. An electromagnetic wavefront is considered to be composed of an infinite number of point source emitters. Any later wavefront can be regarded as the envelope of the spherical wavefronts [7], emitted by all these individual point sources. The electromagnetic field at any point in space can then be determined using the superposition principle of waves. In this thesis, only scalar diffraction theory was applied because neither high numerical aperture (NA) objectives (NA $>$ 0.8) were put to use, nor polarisation effects (such as birefringence) had to be taken into account. At this point we introduce the coordinate systems used throughout this thesis which are shown in Fig. 2.1. It is noted that "It is, however, essential in the theory of image formation to identify the pupil surfaces with the pupil spheres, rather than the pupil planes, if Fourier transform theory is to be applied" [15].

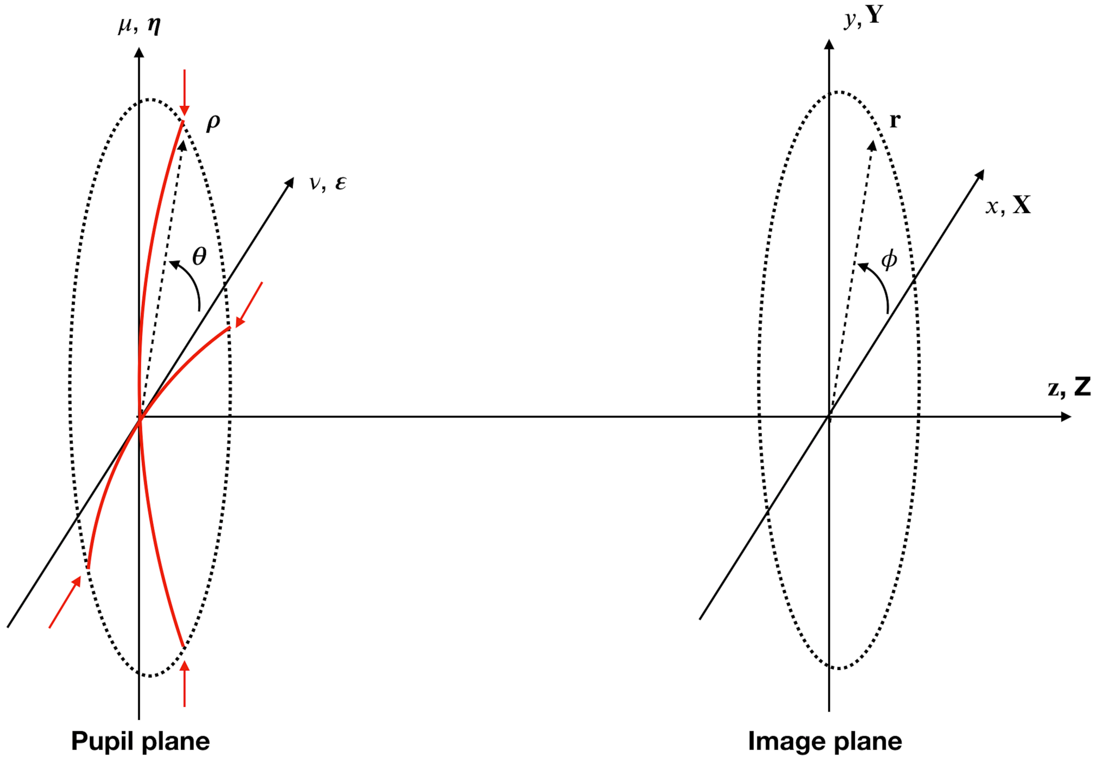


Figure 2.1: Coordinate system definition for the pupil and the corresponding image plane. Pupil plane: cartesian coordinates (ε, η) are normalised with respect to the pupil radius R to give (ν,μ). Image plane: cartesian coordinates (X, Y) are normalised with the diffraction unit λ/NA to give the normalised coordinates (x,y), where λ is the wavelength and NA is the numerical aperture. The red arrows indicate the outline of a spherical reference surface (converging wavefront) along two axes. The coordinate transformation for the normalised radial coordinates (r,ϕ) is: $x = r\cos(\phi)$; $y = r\sin(\phi)$ in the image plane and $\nu = \varepsilon/R$; $\mu = \eta/R$ in the pupil plane.

Where (ε, η) and (X, Y) are cartesian coordinates in the pupil and image plane, respectively before normalisation. The image plane coordinates (X, Y) are normalised with the aid of the diffraction unit [11] (NA/λ) and are of the form:

$$x = X\cdot(NA/\lambda);\quad y = Y\cdot(NA/\lambda) \tag{2.1}$$

where λ is the wavelength and NA is the numerical aperture. The pupil coordinates are normalised with respect to the pupil radius, R, such that:

$$\nu = \varepsilon / R; \quad \mu = \eta / R \tag{2.2}$$

For low to moderate NA values (NA < 0.6), the normalisation for the axial coordinate Z is given by [11]:

$$z = Z \cdot \left(NA^2 / 2\lambda \right) \tag{2.3}$$

One can use polar coordinates (ρ, θ) and (r, ϕ) in the pupil and image plane, respectively. The coordinate transformation from cartesian to polar coordinates is given by:

$$\begin{aligned} x &= r\cos(\phi); \ \mathrm{y} = r\sin(\phi) \\ \nu &= \rho\cos(\theta); \ \mu = \rho\sin(\theta) \end{aligned} \tag{2.4}$$

The following derivations can be found in the "Introduction to Fourier Optics" book [16]. Here the most relevant diffraction theory equations will be presented. The Huygen-Fresnel principle for scalar diffraction is expressed in terms of the diffraction integral [16]:

$$U(X,Y) = \frac{Z}{j\lambda} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U(\varepsilon,\eta) \frac{e^{jkr_{01}}}{r_{01}^2} d\varepsilon \, d\eta \tag{2.5}$$

where U is the electromagnetic field amplitude at the observation point $P_0(X,Y)$ originating from a point $P_1(\varepsilon,\eta)$ in the pupil plane, where j is the imaginary unit, and k the local wave vector. The distance r_{01} between P_0 and P_1 is given by [16]:

$$r_{01} = \sqrt{Z^2 + \left(X - \varepsilon\right)^2 + \left(Y - \eta\right)^2} \tag{2.6}$$

The rather complex integral (2.5) can be cast into a simpler form when looking at the diffracted amplitude at a distance far away. Equation (2.6) can be expanded using a Taylor series. Expanding equation (2.6) while assuming that Z is large in comparison to $X - \varepsilon$ and $Y - \eta$, it suffices to consider the lower order terms:

$$r_{01} = Z\sqrt{1 + \left(\frac{X - \varepsilon}{Z}\right)^2 + \left(\frac{Y - \eta}{Z}\right)^2} \approx Z\left[1 + \frac{1}{2}\left(\frac{X - \varepsilon}{Z}\right)^2 + \frac{1}{2}\left(\frac{Y - \eta}{Z}\right)^2\right] \tag{2.7}$$

Inserting (2.7) into (2.5), while dropping all higher order Z terms except the linear Z term in the denominator, one obtains the Fresnel diffraction integral [16]:

$$U\left(X,Y\right) = \frac{e^{jkZ}}{j\lambda Z}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} U\left(\varepsilon,\eta\right)\, e^{j\frac{k}{2z}\left[\left(X-\varepsilon\right)^2 + \left(Y-\eta\right)^2\right]}\, d\varepsilon\, d\eta = \frac{e^{jkZ}}{j\lambda Z} e^{j\frac{k}{2Z}\left(X^2+Y^2\right)} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left\{U\left(\varepsilon,\eta\right) e^{j\frac{k}{2Z}\left(\varepsilon^2+\eta^2\right)}\right\} e^{-j\frac{k}{Z}\left(X\varepsilon + Y\eta\right)}\, d\varepsilon\, d\eta \tag{2.8}$$

which can be seen as the Fourier transform of the product of the amplitude with a quadratic phase term in the pupil [16]. The limitations involved in using the Fresnel approximation is given by errors mainly introduced by the exponential in equation (2.5). The distance r_{01}, in the exponential, is multiplied by a large number k, which can introduce significant errors [16]. The error remains acceptable as long as the quadratic term of the Taylor expansion is smaller than a radian [16]. This corresponds to Z values of:

$$Z^3 >> \frac{\pi}{4\lambda}\left[\left(X - \varepsilon\right)^2 + \left(Y - \eta\right)^2\right]^2 \tag{2.9}$$

The Fresnel approximation, often referred to as the near-field approximation, is used when one is interested in the electromagnetic field distribution in the vicinity of a diffracting element (such as an aperture) [16, 17]. A detailed mathematical explanation about the limits of the Fresnel approximation is given in ref [16].

When observing the electromagnetic field at very large distances, or when the pupil is illuminated by a focussing (converging) spherical wavefront [16], the diffraction integral (2.8) can be further simplified. This is referred to as the Fraunhofer or far-field approximation. Under the Fraunhofer approximation $\left(Z >> k\left(\varepsilon^2+\eta^2\right)/2\right)$ the quadratic phase term in equation (2.8) is close to zero and we end up solely with the Fourier transform of the aperture:

$$U\left(X,Y,Z\right)=\frac{e^{jkZ}}{j\lambda Z}e^{j\frac{k}{2Z}\left(X^2+Y^2\right)}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}U\left(\varepsilon,\eta\right)\,e^{-j\frac{k}{Z}\left(X\varepsilon+Y\eta\right)}\,d\varepsilon\,d\eta \tag{2.10}$$

Expressing equation (2.10) in normalised coordinates:

$$U\left(x,y,z\right)=-j\left(\frac{R^2NA^2}{2\lambda^2 z}\right)e^{j\frac{4\pi}{NA^2}z}\,e^{j\frac{\pi}{2z}\left(x^2+y^2\right)}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}U\left(\nu,\mu\right)\,e^{-j\frac{\pi\left(NA\right)R}{z\lambda}\left(x\nu+y\mu\right)}\,d\nu\,d\mu \tag{2.11}$$

In this thesis, aberrated focal spots were simulated using the Fraunhofer approximation (equation 2.11). A phase aberration can be included by writing the complex amplitude at the pupil plane as:

$$U\left(\nu,\mu\right)=U\left(\rho,\theta\right)=A\left(\rho,\theta\right)\cdot e^{ik\Phi\left(\rho,\theta\right)} \tag{2.12}$$

Where $A\left(\rho,\theta\right)$ is the transmission function and $\Phi\left(\rho,\theta\right)$ is the aberration function. The intensity in the image plane is obtained by multiplying U with its complex conjugate U^*:

$$I=U\cdot U^* \tag{2.13}$$

In Chapter 6 (aberration retrieval based on focal spot shape) a different sign convention for the exponentials in equation (2.11) was chosen to use the same convention as used by the authors in the literature of the Extended-Nijboer Zernike theory [11-13]. The Fraunhofer diffraction integral (equation 2.11) was implemented in Matlab using the two dimensional FFT (fft2 function) and was computed on a NxN grid, where N was set to 2048 (2^11). The distance to the first dark ring of the calculated Airy spot was chosen to correspond to 18 elements in the Matlab matrix (with a total number of 2048x2048 elements). The pupil diameter corresponds to about 140 matrix elements. Two dimensional images of the pupil and a computed PSF (with homogeneous illumination and no phase aberration applied) are shown in Figure 2.2:

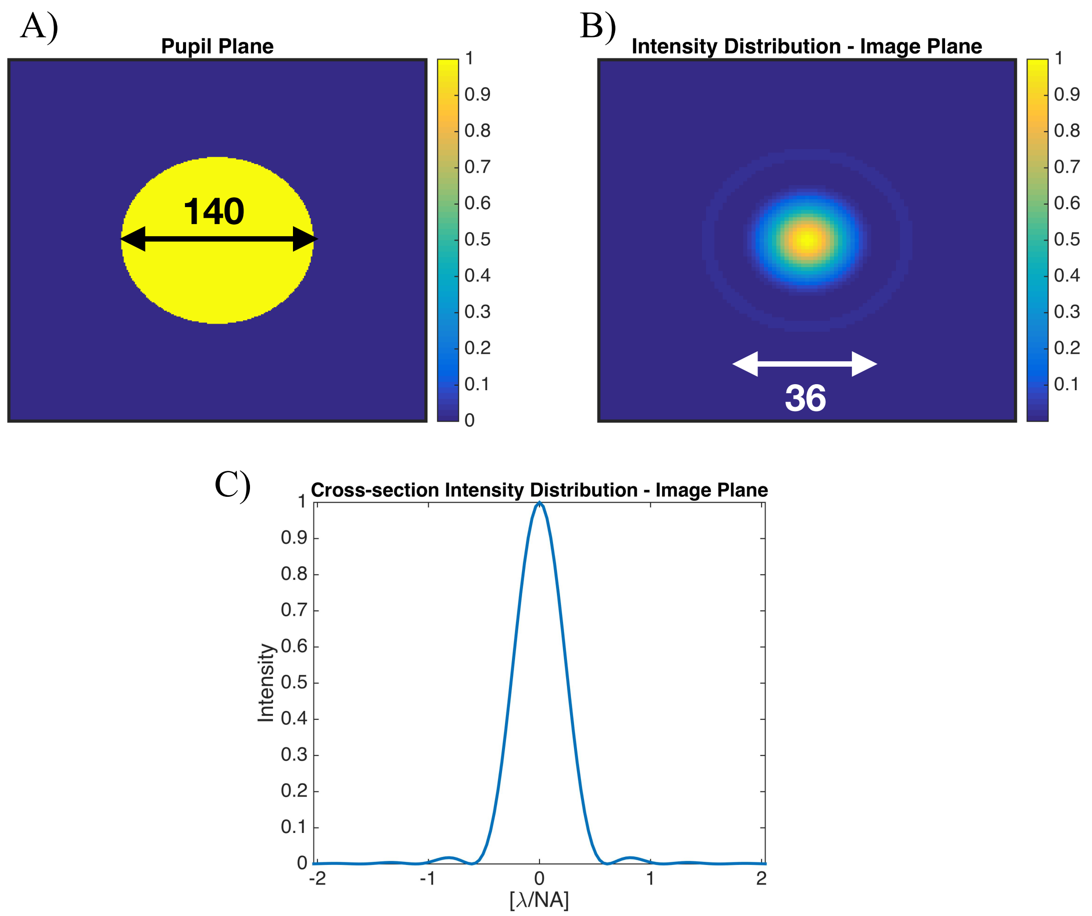


Figure 2.2: A) Circular pupil with homogeneous illumination, B) The computed intensity distribution using the two dimensional FFT (Matlab) on 2048x2048 grid and C) Cross-section of the intensity distribution through the centre of the PSF.

To assess the precision obtained with the computed intensity distribution, the intensity distribution function of the (ideal) Airy spot were subtracted from the intensity distribution computed with Matlab. The (ideal) Airy spot intensity distribution is given by:

$$I_{Airy}(r) = \left(\frac{2J_1(r)}{r} \right)^2 \tag{2.14}$$

where $J_1(r)$ is the first order Bessel function and $I_{Airy}(0) = 1$. The difference between the two distributions is shown in Fig. 2.3.

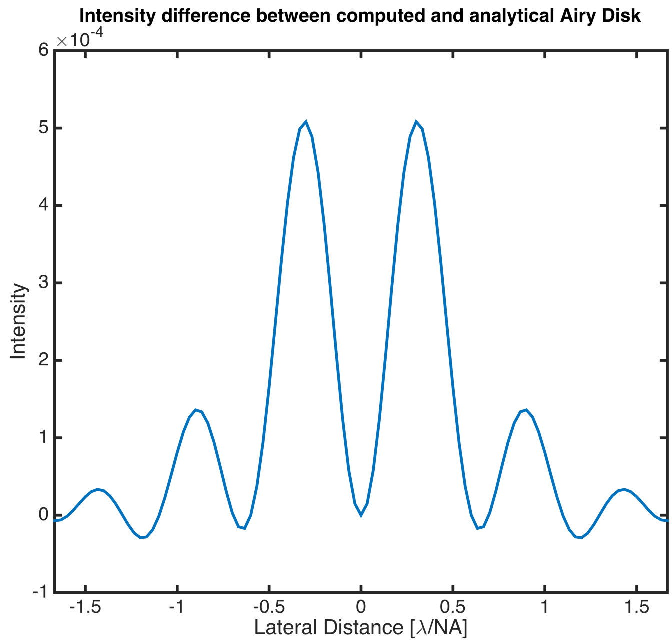


Figure 2.3: Intensity difference for an unaberrated PSF, comparing the two dimensional FFT (Matlab) on 2048x2048 grid to the intensity distribution function of the (ideal) Airy spot.

The error (i.e. the difference in intensity between the FFT result and the analytical solution shown in Fig. 2.3) is less than 10^{-3} in terms of intensity.

The scalar diffraction theory is held to be valid for NA up to 0.6 [11]. For NA > 0.6 or when polarisation effects have to be considered, such as birefringence [11], a vectorial diffraction theory should be used instead.

2.2 Zernike aberration polynomials

In the context of diffraction theory, as explained before, the phase aberration function $\Phi(\rho,\theta)$ in the pupil plane was introduced to account for the presence of wavefront aberrations. In optics, a popular set of functions for describing wavefront aberrations in the pupil are the Zernike polynomials. Zernike Polynomials were introduced when Fritz Zernike in 1934 studied the effects of aberrations on his phase contrast method [10]. Zernike polynomials are an orthogonal set of circle functions and are nowadays commonly used to describe aberrations in optical imaging systems. The lower order Zernike polynomials (nowadays often referred to as modes) are associated with commonly understood aberrations such as astigmatism, coma or primary spherical.

The Zernike Polynomials are defined over the unit circle and are given by:

$$\begin{aligned} Z_n^m(\rho,\theta) &= R_n^m(\rho)\sin m\theta \quad \text{for m negative} \\ &= R_n^m(\rho)\cos m\theta \quad \text{for m positive} \end{aligned} \tag{2.15}$$

where $R_n^m(\rho)$ is a finite hypergeometric series, ρ is the circle radius, n the radial order, m azimuthal order and θ the azimuthal angle.
The following rules apply for n and m: $n-m$ even, $m \le n$. $R_n^m(\rho)$ is given by:

$$R_n^m(\rho) = \frac{\rho^{-m}}{\left(\frac{n-m}{2}\right)!}\left(\frac{d}{d(\rho^2)}\right)^{\frac{n-m}{2}}\left\{\rho^{n+m}\left(\rho^2-1\right)^{\frac{n-m}{2}}\right\} \tag{2.16}$$

An arbitrary wavefront $W(\rho,\theta)$, when confined to a circular aperture can be decomposed into a weighted sum of orthogonal Zernike polynomials:

$$W(\rho,\theta) = \sum_{n,m} \alpha_n^m Z_n^m(\rho,\theta) \tag{2.17}$$

where α_n^m are the respective Zernike polynomial coefficients. Several Zernike polynomial notations have been introduced for the needs of different research fields, such as the widely used Noll and Malacara notations. The Noll notation is commonly used in Astronomy. Robert Noll introduced a modified set of orthonormal Zernike polynomials and discussed their mathematical properties and their application for correcting atmospheric turbulences as encountered in astronomical adaptive optics [18]. The Zernike coefficients output by the Thorlab's Shack-Hartmann Wavefront Sensor system are represented in the Malacara notation [19]. The Malacara notation involves a single indexing scheme as well as a normalisation of the Zernike polynomials. In this thesis, the Malacara normalisation is adopted, while using a (n,m) indexing scheme for the Zernike polynomials. The normalised Zernike polynomial are (in this thesis) written as:

$$\begin{aligned} Z_{n\ even}^{m}(\rho,\theta) &= \sqrt{2(n+1)}\ R_n^m(\rho)\cos(m\theta) \quad && m \neq 0 \\ Z_{n\ odd}^{m}(\rho,\theta) &= \sqrt{2(n+1)}\ R_n^m(\rho)\sin(m\theta) \quad && m \neq 0 \\ Z_n^m(\rho,\theta) &= \sqrt{n+1}\ R_n^0(\rho) && m = 0 \end{aligned} \tag{2.18}$$

The Zernike Polynomials, with their Malacara normalisations, up to radial order 4 and all the corresponding azimuthal orders are listed in table 2.1.

Table 2.1: Zernike Polynomials up to radial order 4 and the corresponding azimuthal orders using the Malacara normalisation

n	m	Normalisation Factor	Zernike Polynomial	Name
0	0	1	1	Piston
1	-1	2	$\rho\sin\theta$	Tip
1	1	2	$\rho\cos\theta$	Tilt
2	-2	$\sqrt{6}$	$\rho^2\sin 2\theta$	Oblique Astigmatism
2	0	$\sqrt{3}$	$2\rho^2-1$	Defocus
2	2	$\sqrt{6}$	$\rho^2\cos 2\theta$	Vertical Astigmatism
3	-3	$2\sqrt{2}$	$\rho^3\sin 3\theta$	Vertical Trefoil
3	-1	$2\sqrt{2}$	$(3\rho^3-2\rho)\sin\theta$	Horizontal Coma

n	m	Normalisation Factor	Zernike Polynomial	Name
3	1	$2\sqrt{2}$	$\left(3\rho^3-2\rho\right)\cos\theta$	Vertical Coma
3	3	$2\sqrt{2}$	$\rho^3\cos 3\theta$	Oblique Trefoil
4	-4	$\sqrt{10}$	$\rho^4\sin 4\theta$	Oblique Quadrafoil
4	-2	$\sqrt{10}$	$\left(4\rho^4-3\rho^2\right)\sin 2\theta$	Oblique Secondary Astigmatism
4	0	$\sqrt{5}$	$6\rho^4-6\rho^2+1$	Primary Spherical
4	2	$\sqrt{10}$	$\left(4\rho^4-3\rho^2\right)\cos 2\theta$	Vertical Secondary Astigmatism
4	4	$\sqrt{10}$	$\rho^4\cos 4\theta$	Vertical Quadrafoil

As can be seen from Table 2.1, in some Zernike polynomials the higher order radial term is "balanced" with lower order terms (by means of addition or subtraction), for example in case of spherical aberration $\left(n=4,\ m=0\right)$ the radial order term ρ^4 is "balanced" by a ρ^2 and a constant term. The addition of these lower order terms to the aberration polynomial is referred to as the "balancing" of Zernike polynomials. Zernike polynomial balancing ensures that for small aberrations (< 0.07λ), the normalised maximum intensity lies at the paraxial (or Gaussian) focus [7]. A geometrical interpretation was given by Nijboer [8]. He explained that for spherical aberration the ρ^2 term compensates for axial displacement caused by the ρ^4 term, while for coma the ρ term compensates for lateral displacement caused by the ρ^3 term. Thus the balancing ensures that the Zernike polynomials describe the wavefront with respect to the position of best focus in the presence of small aberrations. The term "best" focus will be used from here on to refer to the spot, in the vicinity of the geometrical focus, with the highest intensity in x,y,z.

Mathematical Properties of Zernike Polynomials

The Zernike polynomials are orthonormal on the unit circle which implies that:

$$\frac{1}{\pi}\int_0^1\int_0^{2\pi} Z_n^m(\rho,\theta) Z_{n'}^{m'}(\rho,\theta)\ \rho\ \mathrm{d}\rho d\theta = \delta_{nn'}\delta_{mm'} \tag{2.19}$$

The coefficients for the wavefront reconstruction can be computed as follows:

$$\alpha_n^m = \frac{1}{\pi}\int_0^1\int_0^{2\pi} W(\rho,\theta) Z_n^m(\rho,\theta)\rho\ \ \mathrm{d}\rho\mathrm{d}\theta \tag{2.20}$$

and the wavefront is then:

$$W(\rho,\theta) = \sum_{n,m} \alpha_n^m Z_n^m(\rho,\theta) \tag{2.21}$$

The mean value of the wavefront is:

$$\langle W(\rho,\theta)\rangle = \frac{1}{\pi}\int_0^1\int_0^{2\pi} W(\rho,\theta)\rho d\rho d\theta = \alpha_0^0 \tag{2.22}$$

So, all Zernike polynomials, except Piston (which has no physical relevance for describing aberrations), when integrated over the unit circle are equal to zero.

The root-mean square wavefront error (for small aberrations) is given by:

$$RMS = \sqrt{\langle W^2(\rho,\theta)\rangle - \langle W(\rho,\theta)\rangle^2} = \sqrt{\sum_{n,m>0}\left(\alpha_n^m\right)^2} \tag{2.23}$$

2.3 Image Formation - Microscopy

The three dimensional point spread function (PSF) h of a lens with circular aperture, under paraxial approximation, can be determined using an integral (using normalised cylindrical coordinates) of the form:

$$h(r,\phi,z) = Ce^{j\frac{4\pi}{NA^2}F} e^{j\frac{\pi}{2F}r^2} \int_0^{2\pi}\int_0^1 A(\rho,\theta) \cdot e^{ik\Phi(\rho,\theta)} e^{-jf\rho^2} e^{-j2\pi r\rho\cos(\phi-\theta)} \rho\, d\rho\, d\theta \qquad (2.24)$$

where $C = -j\left(\frac{R^2 s_0^{\ 2}}{2\lambda^2 z}\right)$ is a constant, z is the distance from focus, F is the distance from pupil to the image plane, $e^{-jkf\rho^2}$ is a defocusing term, representing a parabolic wavefront deviation in the pupil, Φ is a wavefront aberration expressed in wavelength, f is a defocus parameter [11] which is related to the axial coordinate z via $f = -2\pi z$. The intensity PSF is simply given by: $I = |h|^2$. It is noted that the intensity distribution at different axial planes can also be calculated by varying the value α_2^0 of Zernike defocus, instead of using the defocus parameter f. The wavefront Φ, as a function of an axial displacement Z, is given by [12]:

$$\Phi = Z\left[1 - \sqrt{1-\rho^2 (NA)^2}\,\right] \qquad (2.25)$$

In the low NA regime the Zernike defocus term can be used alternatively resulting in: $\Phi = \alpha_2^0 (2\rho^2 - 1)$. For higher NA systems a quadratic phase departure does not suffice to accurately describe an axial displacement [12]. Using the Taylor expansion for the square root in (2.25) and comparing it with the Zernike defocus polynomial as well as using the relationship [11] $f = -2\pi (NA)^2 Z/2\lambda = -2\pi z$:

$$Z = \frac{4\alpha_2^0}{NA^2} = -\frac{\lambda f}{\pi NA^2} \tag{2.26}$$

Formula 2.25 relates the axial displacement Z to the Zernike defocus amplitude and to the defocus parameter *f*. The intensity distribution at $f = 0$, i.e. at focus, is given by equation 2.14 and is shown in Fig. 2.4

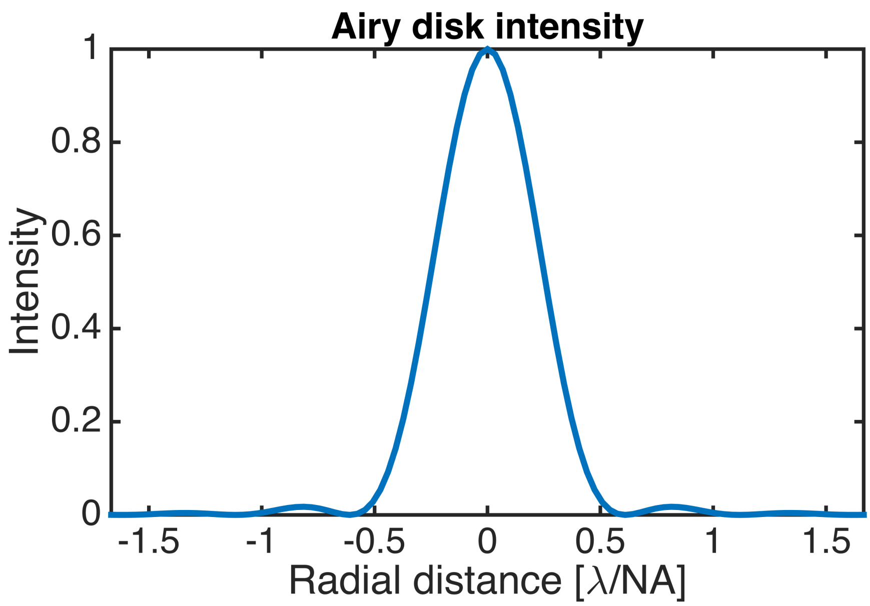


Figure 2.4: Airy disk intensity distribution.

Based on the Airy disk intensity distribution, the resolving power of a lens can be estimated. A commonly used criterium for resolution is the so-called Rayleigh criterium which is given by the lateral distance to the first zero point of the Airy disk function. In other words when two point objects, imaged through a microscope, come close to each other, Rayleigh considered that one can just resolve them when the maximum of one overlaps with the first minimum of the other which corresponds to a separation distance of:

$$d_{Rayleigh} = \frac{0.61\lambda}{NA} \tag{2.27}$$

Abbe used a different resolution criterium and stated that the maximum spatial frequency of an object that can be transmitted by a microscope objective is $2NA/\lambda$

(Abbe resolution limit). According to Abbe the size of the smallest detectable feature would be:

$$d_{Abbe} = \frac{0.5\lambda}{NA} \tag{2.28}$$

The PSF represents the impulse response of the lens. The impulse response is the response of a system, such as a lens, to a delta function. Consequently, the image of an object, produced by an optical imaging system, may be described as a convolution of the object with the system PSF. In mathematical terms this means that the image is the convolution of the object function t with the PSF, and for a coherent imaging system it would take the form:

$$I_{Object/Coherent}(r,\phi,z) = \left|h(r,\phi,z) \otimes t(r,\phi,z)\right|^2 \tag{2.29}$$

Whereas, in the case of incoherent imaging system, one obtains:

$$I_{Object/Incoherent}(r,\phi,z) = \left|h(r,\phi,z)\right|^2 \otimes \left|t(r,\phi,z)\right|^2 \tag{2.30}$$

2.4 Laser Scanning Microscopy

In laser scanning microscopes (LSM), an image is formed by scanning an object with a focussed beam. Scanning is achieved either by object or laser scanning. In the type 1 LSM microscope, when adopting the terminology of Wilson and Sheppard [20], the reflected (or transmitted) light from the sample is collected by a photodetector and used to create an image of the scanned sample. Probably the most popular type of LSM is the confocal microscope, which uses a pinhole in the detection plane, which is conjugate to the object plane, that functions as a spatial filter and blocks some of the out of focus light reaching the detector. Since its conception in the 1960s [21], the confocal microscope has gained great popularity in various scientific and industrial areas. Whether for imaging 3D biological samples or for inspection of industrial samples, the popularity mostly comes from its sectioning capabilities and contrast improvements. The removal of out focus light by the pinhole enables the confocal microscope to obtain contrast rich images of thin sections within a sample. The sectioning capability of the confocal microscope depends on the size of the pinhole. For small pinholes (radii smaller than half the distance to the first dark ring of the Airy disk) the sectioning performance hardly changes [22] and is comparable to that of a true confocal microscope (with an infinitely small pinhole). The sectioning properties start to deteriorate as the pinhole diameter increases. For large pinholes (~ ≥ 1.5AU) the microscope becomes effectively a conventional widefield microscope. As regards contrast, the true confocal microscope (with infinitely small pinhole) can image all spatial frequencies with higher contrast [23] than a widefield microscope. The sectioning capabilities of the confocal microscope are illustrated in Fig. 2.5.

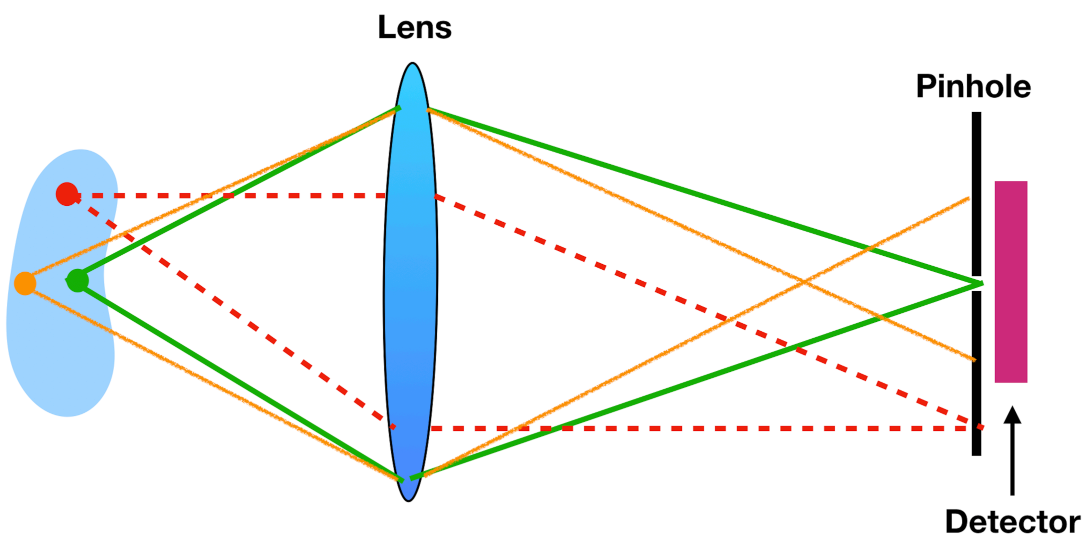


Figure 2.5: Confocal Microscope sectioning capability - light emitted from focus (green) reaches the detector whereas out of focus light (in orange and red) is blocked.

Both the illumination and the collection lens should ideally be diffraction limited. The distortion by aberrations of the PSF decreases the signal level at the pinhole, reducing depth discrimination and lowering the resolution [22, 24]. When imaging samples spherical aberration in particular can compromise the imaging performance of confocal microscopes because of refractive index mismatch conditions [25]. The confocal PSF is given by the product of the illumination PSF with the detection PSF. The difference between widefield and confocal microscopy is best notable when looking at the axial intensity distribution. The axial intensity distributions of a widefield and a confocal microscope, when looking at a point source emitter, are shown in fig 2.6

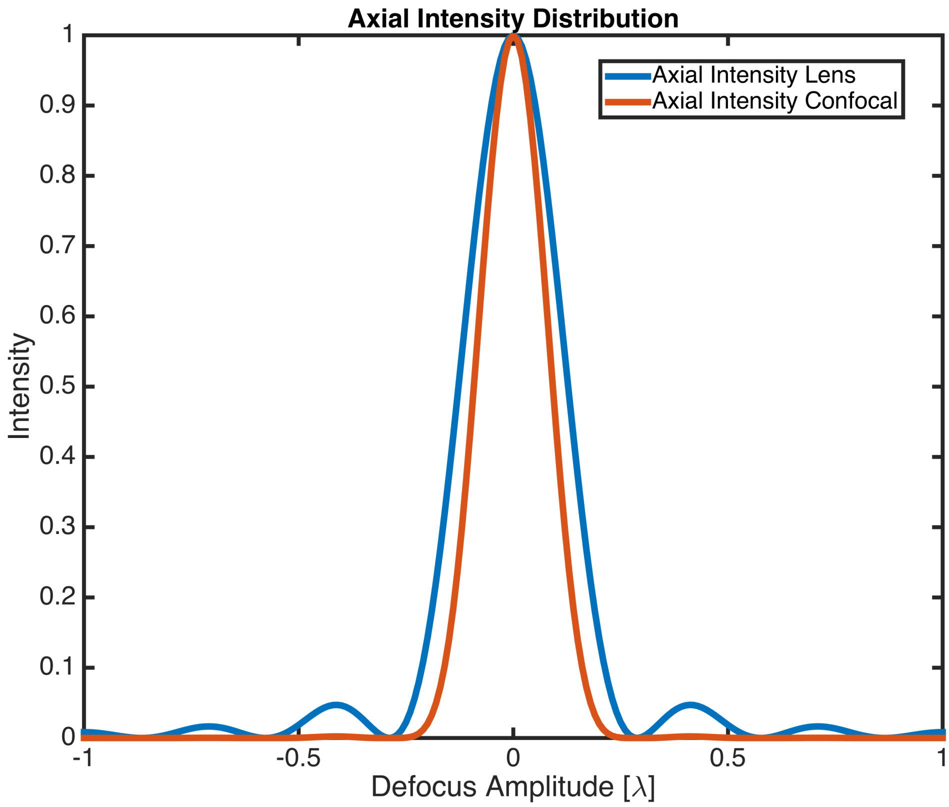


Figure 2.6: Axial intensity of a widefield and confocal microscope. The axial displacement was achieved by changing the value α_2^0 of the Zernike defocus term.

2.5 Adaptive optics

Adaptive Optics is an active wavefront shaping technology that locally imposes phase changes to a wavefront and can be used to correct for aberrations. Deformable mirrors are commonly used as the wavefront shaping device. The reflective surface of a deformable mirror is deformed by actuators.

2.5.1 Adaptive optics in astronomy

Adaptive optics (AO) is a technology originally developed in the field of optical astronomy to correct for optical aberrations caused by the turbulent atmosphere [26]. In the 1950s, the US air force drove scientific research aimed at finding ways to compensate for the wavefront distortions caused by the atmosphere. While the military was interested in developing new high powered laser systems for anti-rocket systems or satellite laser relay systems, astronomers were interested in achieving diffraction limited resolution for their telescopes to improve astronomical observation [27]. Babcock was the first to propose the idea of adaptive optics in 1953 [28]. His idea was to dynamically correct aberrations with a deformable mirror capable of changing its shape at frequencies up to about 30 Hz in order to keep up with changes in the turbulent atmosphere. It took many years before adaptive optics could establish itself in astronomy because the hardware was not available. The atmosphere induced wavefront distortions, after being measured, are corrected with a wavefront shaping device, which imposes a wavefront distortion of opposite sign but with the same magnitude. The problem of wavefront sensing in the atmosphere was solved by using the so-called (artificial) “laser guide star”. The mesosphere (atmospheric layer at a height of about 90km above sea level) contains enough sodium atoms to create a strong light beacon by laser excitation [29]. A high power laser is fired into the atmosphere to excite sodium atoms in the mesosphere and the re-emitted light is then used for measuring atmosphere induced wavefront aberrations [27]. In closed-loop with the wavefront sensor, the deformable mirror can then correct for atmosphere induced wavefront distortions. The advantage of laser guide stars is their bright intensity, and they can be created at a desired region in the sky. The region in the atmosphere where the wavefront distortion does not change significantly is referred to as the isoplanatic patch [30], thus laser guide stars are created close to the celestial bodies of interest to ensure that they are within the isoplanatic path. Alternatively, if an observed celestial body is bright enough, it can be used as natural guide star for wavefront sensing. The creation of an artificial laser guide by the MMT telescope in Southern Arizona is shown in Fig. 2.7A. Improvements realised by correcting the atmospheric wavefront aberrations with the adaptive optics system of the Canada-France-Hawaii Telescope are shown in Fig. 2.7B.

A)
B)

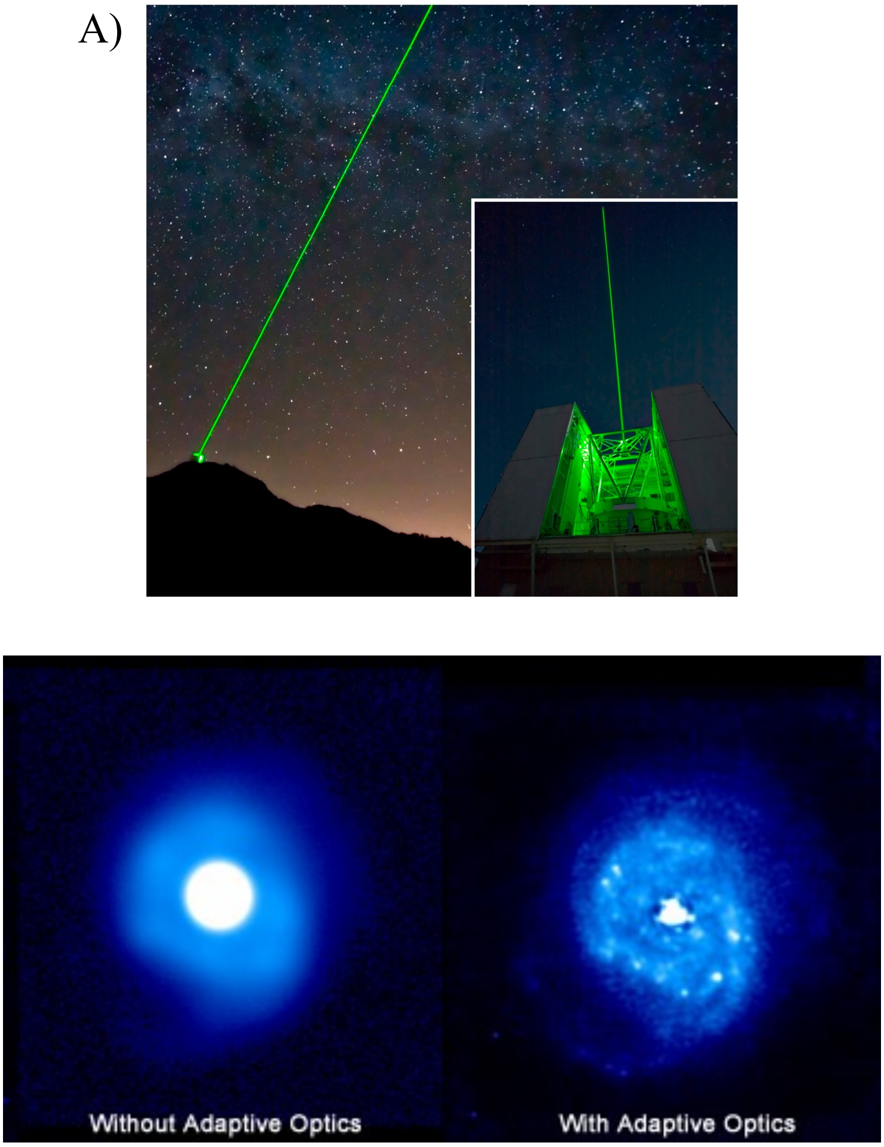


Fig 2.7: a) The laser-guide-star adaptive-optics (AO) system in operation at the 6.5m MMT (formerly the multiple-mirror telescope) in southern Arizona. (Images courtesy Thomas Stalcup, www.oldweb.lbto.org/index.htm). b) The nuclear region of the nearby galaxy NGC 7469 (Image Source: Canada-France-Hawaii Telescope www.cfht.hawaii.edu)) with and without adaptive optics

2.5.2 Adaptive Optics Systems

An AO correction system usually consists of a wavefront sensor and a wavefront corrector, typically a deformable mirror. After measuring a distorted wavefront, the deformable mirror imposes a wavefront distortion of opposite sign in order to correct the wavefront. AO correction systems are often operated in a closed-loop feedback control system wherein the wavefront corrections made by the deformable mirror are measured by a wavefront sensor and small corrections to the deformable mirror surface are iteratively applied until the desired wavefront shape is obtained, up to an acceptable level of quality. An illustration of a closed-loop system is shown in Fig. 2.8. In the following a short review on deformable mirrors, with special emphasis on the continuous deformable membrane mirror used in this thesis, will be given. Thereafter wavefront sensors and indirect wavefront sensing methods for microscopy will be discussed.

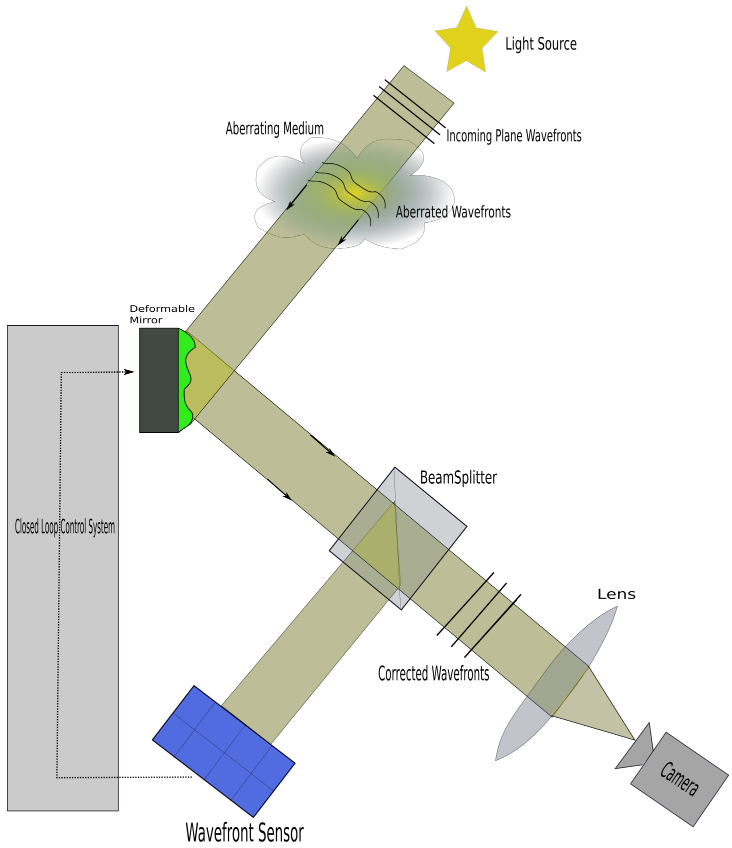


Figure 2.8: AO closed-loop correction system.

2.5.3 Deformable mirrors

Deformable mirrors are used for wavefront shaping applications in adaptive optics. With deformable mirrors local phase changes (or optical path differences) can be applied to the wavefront in order to compensate for wavefront distortions. Different types of deformable mirrors are known. Most common types of mirrors are deformable membrane mirrors (DMM) and segmented mirrors [31]. DMMs use a set of actuators to physically deform a reflective membrane. DMM actuators can be magnet-coil units, electrode pairs, electrostrictive elements or piezoelectric modules [32]. DMM actuators can be activated electromagnetically if the actuators are magnetic [33], make use of piezo-electric effects to deform piezo-electric crystals [31] or by applying a voltage to deform a conducting surface by means of an electric force. Segmented mirrors consist of an arrangement of small independent mirror segments which can be individually tilted to locally deform the wavefront.

As an alternative to deformable mirrors, spatial light modulators (SLM) can be used to shape the wavefront. SLMs consist of an array of pixels and a liquid crystal where the local refractive index can be changed by rotating the liquid crystals using electrostatic forces applied via the pixels [32]. Due to the large number of pixels available, wavefront shaping with SLMs can be done at a higher spatial resolution. Disadvantages of SLMs with regards to aberration correction are: their limited temporal bandwidth (~100Hz); the light source should be monochromatic; and the restricted amount of phase changes which can be applied (up to a wavelength) [32].
In this project the DMM, used for aberration correction and for applying controlled amounts of aberrations, is a Mirao 52e (Imagine Optics, France). The Mirao52e has 52 magnetic actuators which are fixed to the backside of a silver coated reflective surface [34]. The actuators are driven by a set of coils. The spacing between each actuator is about 2.5mm. Figure 2.9 shows a schematic drawing of an electromagnetic DMM.

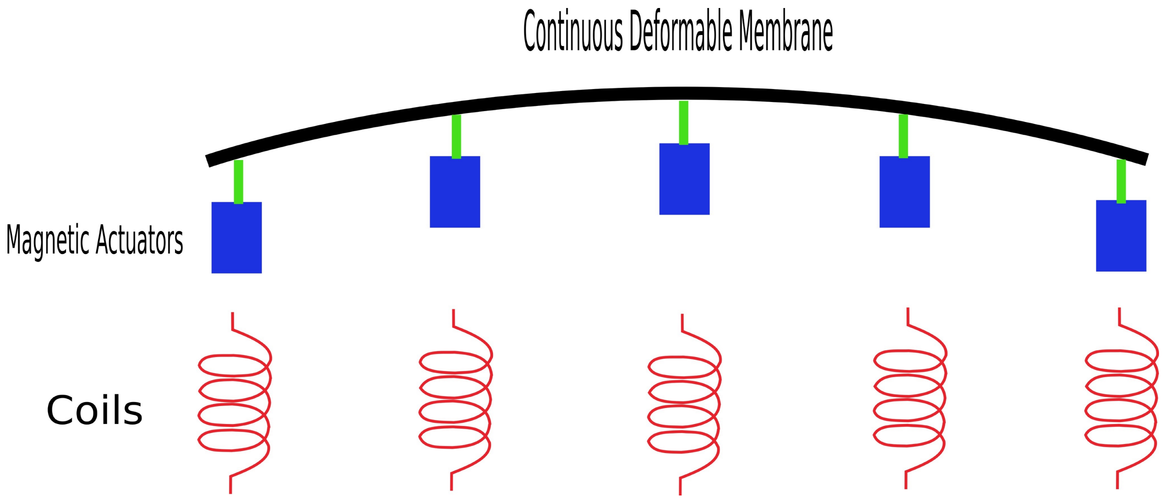


Figure 2.9: Deformable membrane mirror with magnetic actuators driven by a set of coils.

The mirror has an effective diameter of 15mm. It is recommended to use only a 12 mm area of the mirror in order to better reproduce high order spatial frequency aberrations [33]. Voltages between -1 Volt and +1 Volt can be applied to individual actuators. According to the manufacturer, the maximum peak-to-valley values of wavefront distortions are: 30 µm for primary astigmatism, 10 µm for primary coma mirror and 8 µm for primary spherical. The DMM operates with a bandwidth > 200Hz and has a good linear response [33]. The actuator arrangement, with 15mm and 12mm circular illumination areas, is shown in Figure 2.10. The membrane of the Mirao52 is bound outside the active region.

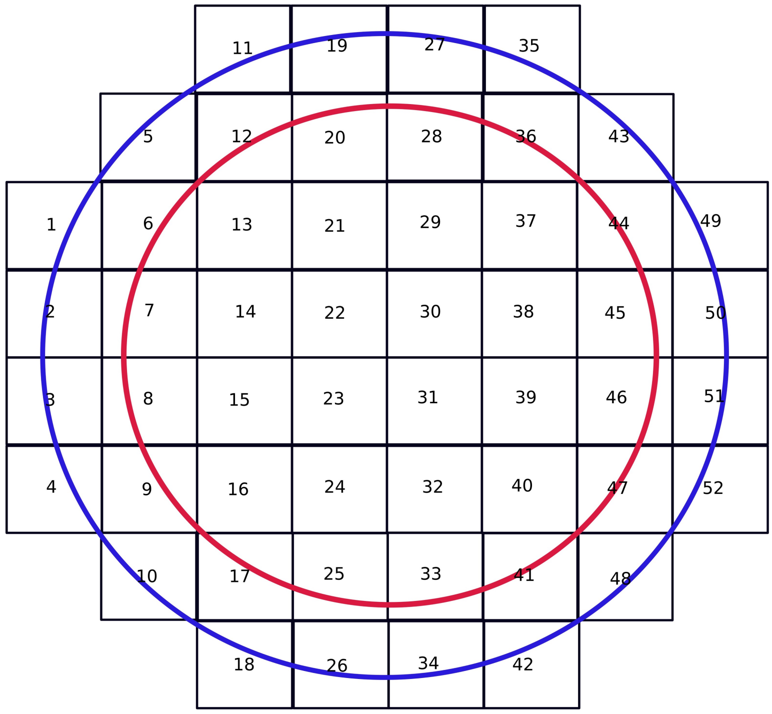


Figure 2.10: Mirao52e actuator configuration. The actuators are located under the centres of the numbered segments. The red circle has a diameter of 12mm and the blue circle has a diameter of 15mm.

A performance comparison between different off-the-shelf laboratory deformable mirrors is given in reference [35]. For aberration correction, the Mirao52 was one of the best performing mirrors in the study. The Mirao52 can also be used to correct for large amounts of spherical aberration in high NA imaging situations [34].

2.6 Wavefront Sensing

The phase of electromagnetic waves cannot be measured due to the high frequency oscillations of electromagnetic fields (hundreds of THz for electromagnetic radiation in the visible spectrum). Optical detectors for example record low frequency intensity signals and lack phase information. Wavefront sensing methods can be subdivided into direct and indirect wavefront sensing methods. Direct wavefront sensing methods directly measure a wavefront. Examples thereof are Shack-Hartmann sensors and interferometers. Indirect wavefront sensing methods derive wavefront information from a wavefront related parameter. In microscopy for example such a parameter could be the intensity measured by a pinhole detector. In the following an overview of these two types of sensing methods is given:

2.6.1 Direct Wavefront Sensing

Interferometry

In interferometry, phase information is obtained by interfering a probe beam with a reference beam and analysing the interference pattern. For interferometry to work, light sources with a certain degree of coherence are needed. Applications involving fluorescent light, for example, do not make use of interferometry due to the incoherent nature of fluorescence. Phase stepping interferometers record multiple interferograms with different amounts of phase shifts applied to one of the two paths of the interferometer. They have been used to measure wavefront aberrations produced by different biological samples in both low and high NA optical systems [36, 37]. Phase stepping interferometers are often arranged in a Mach-Zehnder or a Twyman-Green configuration. In shear interferometers the incoming beam of light is divided into two beams and the wavefront is measured by imposing a shear (or lateral displacement) to one of the beams with respect to the other [38]. Shear interferometers are often used in optical astronomy for wavefront sensing [27]. In scattering samples, interferometry is difficult because of backscattered light originating from outside the region of interest (e.g. focal plane). A way to mitigate this problem is to use low coherence light sources so that interference will only occur within a small region (within the coherence length). Feierabend et al. implemented a low coherence interferometry method called coherence gating wavefront sensing. Low coherence interferometry method is able to select in-focus light and to reject backscattered light [39] thereby allowing for wavefront sensing by means of interferometry within scattering samples.

Shack-Hartmann Sensor

A Shack-Hartmann Wavefront Sensor (SH) consists of an array of lenslets placed in front of a CCD array at a distance corresponding to the focal length of the lenslets [40, 41]. The incoming wavefront is split up by the lenslets into an array of focussed beams, which create focal spots on the CCD. In the case of an incoming plane wave, each lenslet would focus the light at a point on its optical axis. For an aberrated wavefront, the centre of some focal spots (often referred to as centroids) is shifted laterally, whereby the amount of shift of a centroid depends on the average tilt of the wavefront part that is incident on a lenslet. The displacement of a spot is proportional to the average wavefront gradient across the lenslet aperture [41]:

$$\begin{aligned} \delta x &= \frac{f}{A}\int_A \frac{\partial W(x,y)}{\partial x} dxdy \\ \delta y &= \frac{f}{A}\int_A \frac{\partial W(x,y)}{\partial y} dxdy \end{aligned} \tag{2.31}$$

where f is the focal length of a lenslet, A is the surface area of the lens let and W is the wavefront. The principle of a SH is depicted in Fig. 2.11. The wavefront is reconstructed using information of the local wavefront tilts, obtained from the centroid displacement data. Thereafter the distorted wavefront can be decomposed, for example, into Zernike aberrations [19] with appropriate software.

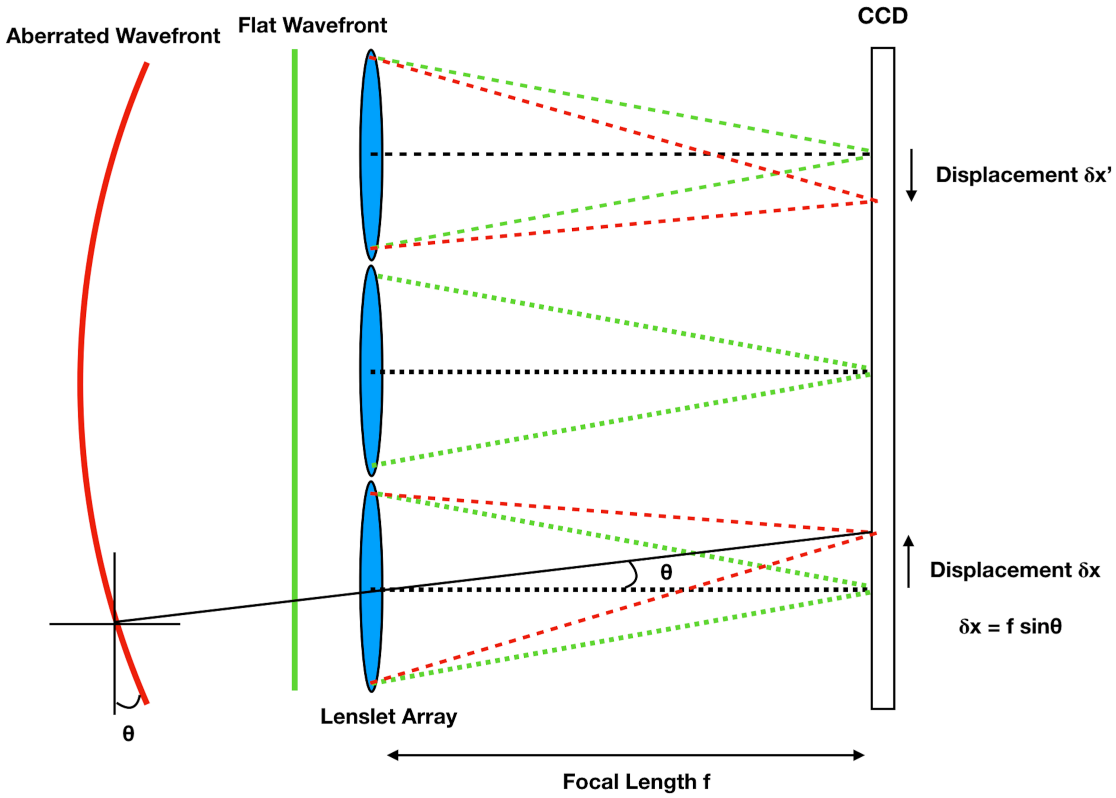


Figure 2.11 Principle of the Shack-Hartmann Wavefront sensor. A CCD array is placed at the focal position of the lenslet array. The focal points of the aberrated wavefront (red) are displaced relative to the centred focal spots of the flat wavefront (green)

SHs are commonly used in optical astronomy to measure atmosphere induced aberrations using the re-emitted light from laser guide stars. For reliable and accurate wavefront sensing, the SH requires light from a reference point source (or guide star). In microscopy, natural guide stars are rarely present in biological samples. Small fluorescent nanobeads have therefore been used as guide stars in conjunction with a SH sensor to correct for sample induced aberration in biological microscopy [42]. As an alternative to fluorescent beads, the fluorescent light emitted from large fluorescently labelled biological structures (such as neuron cells) has been used as guide stars, which provide slightly less accurate measurement results in comparison to diffraction limited fluorescent beads [43]. Direct wavefront sensing in biological samples can become challenging because of out-of-focus light. To reduce the amount of out-of-focus light reaching a SH sensor it has been proposed to use a large pinhole (placed in a plane conjugated to the image plane of the microscope) [44]. In analogy to the confocal microscope, this type of SH configuration is referred to as a confocal wavefront sensor. Although the pinhole removes out-of-focus light it also acts as a filter on the transmitted wavefront. As the pinhole gets smaller, the wavefront gets flattened due to high frequency components being filtered out, similar to a low-pass filter. A Pinhole with a 3 Airy Unit diameter has been shown to be a good compromise between out-of-focus light rejection and wavefront filtering [44]. Alternatively, SH sensors can measure the wavefront using the light from extended objects. The centroid shifts are determined from cross-correlation of each image produced by a lenslet with a reference image. The local wavefront gradients (see equation 2.3.1) are derived from the cross-correlation maxima positions [45]. This method has recently been used in light sheet microscopy for measuring aberrations using the fluorescent light emitted from the sample volume illuminated by the light sheet [46].

2.6.2 Indirect Wavefront Sensing

Modal Wavefront Sensing

In modal wavefront sensing, a wavefront related parameter, often referred to as the optimisation metric, is optimised by running through a set of aberration modes. An issue with modal wavefront sensing is that the underlying physics behind the imaging process has to be known in order to derive a set of aberration modes linked to the metric that requires optimisation. In order to be efficient, the aberration modes should have independent influence on the metric thereby mitigating cross-talk [32]. To apply aberration modes, wavefront shaping devices such as DMM or spatial light modulator have to be used. For a confocal microscope with a pinhole, Zernike modes are a convenient set of aberrations modes, because they are related to the measured intensity at the pinhole. Such an optimisation approach, often referred to as a modal Zernike wavefront sensor, has been used to determine wavefront aberrations in terms of its Zernike modes [47]. The principal of a modal Zernike wavefront sensor is illustrated in Fig.2.12. For the applied aberration mode in Fig. 2.12, the metric is evaluated for 3 different values of the amplitude of an aberration. The maximum of the metric can be found by curve fitting. At the position of the maximum, the aberration is compensated.

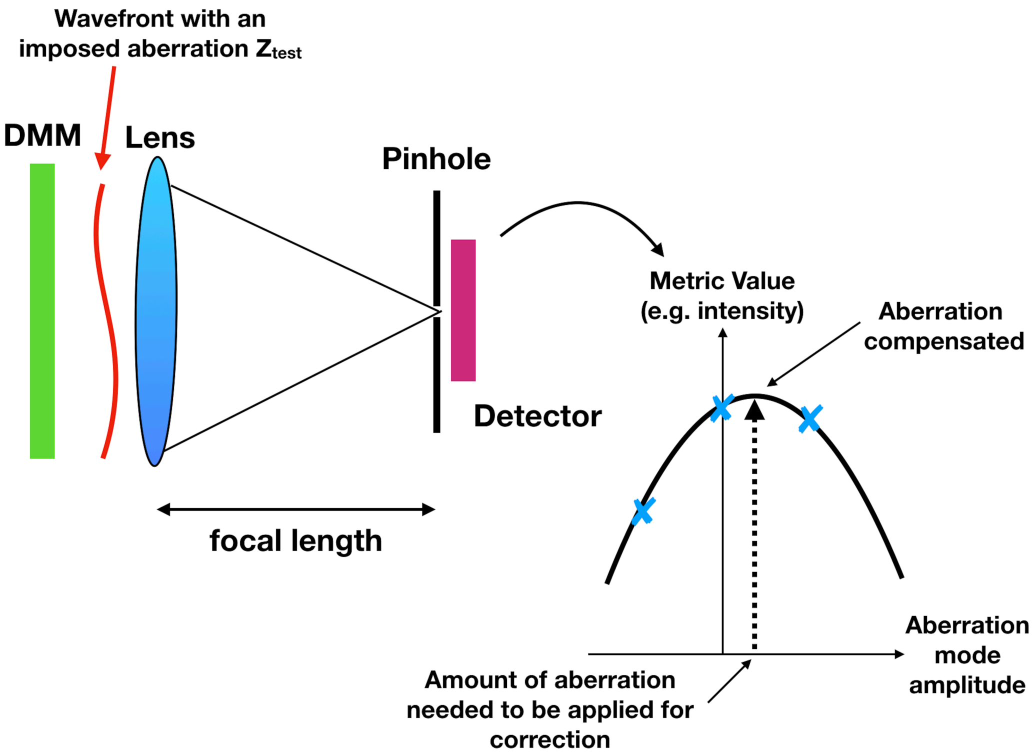


Figure 2.12 Working principle of a modal Zernike wavefront sensor shown for a single aberration mode

The method has also been used in a fluorescent confocal microscope [48]. A Zernike modal aberration correction method is limited to small aberration amplitudes (<0.071λ) due to the mutual influence of aberration modes on the metric (also known as cross-talk) for large aberration. The modal Zernike wavefront sensor can cope with larger aberrations but would need several optimisation iterations in order to achieve an acceptable correction [49]. A strongly aberrated focal spot can be efficiently corrected using the so called Lukosz modes by minimising the mean rms focal spot radius [50]. However, since the latter method requires measurements of the mean rms spot radius, it would be less suitable for confocal pinhole detection. For other types of microscopes (for example widefield microscopes, structured illumination microscope) different sets of aberration modes and metrics are used [51, 52].

Zonal Sensing

In zonal wavefront sensing, a single actuator or a group of actuators from the deformable mirror are selected, according to a chosen optimisation algorithm. The wavefront is locally deformed by actuators and a sensor is used to detect how these deformations affect the optimisation metric. Wright et al. [53] compared different iterative optimisation algorithms (hill climbing, random search and a genetic algorithm) to correct for sample induced aberrations (by focusing with an air microscope objective through various depths of water) in a confocal microscope setup. The random search algorithm selects an actuator randomly and applies a random perturbation to it (in terms of a voltage change) whereas the hill-climbing algorithm goes through the actuators sequentially while incrementally increasing the actuator voltage until a voltage cap is reached. In comparison, the genetic algorithm tests a population of random mirror shapes and generates an improved population for the next iteration based on the mirror shapes which improved the metric during the last iteration. The optimisation algorithms stop when the metric has not significantly improved after the last say 50 iterations. The hill climbing algorithm was the quickest (less than a minute) and achieved a good level of correction and with improvements (in terms of FWHM) slightly worse than the random search or genetic algorithms. The genetic algorithm gave the best results but at the cost of long optimisations. A random search optimisation showed to be the best compromise in terms of improvement and time to complete a optimisation [53]. The random search and genetic algorithms are more likely to find the global maximum as compared to a hill-climbing algorithm which can easily be stuck at a local maximum. Zonal methods have been used in non-linear microscopy [54] as well as in optical tweezers [55].

Phase Retrieval

It is reiterated here that phase retrieval methods aim to recover the phase information from image intensity data. Phase retrieval is used in X-ray crystallography [56], optical astronomy [57] and microscopy [58-60]. In X-ray crystallography, phase information, retrieved from the X-ray diffraction pattern of a crystal, gives information about the crystalline structure, while in astronomy and microscopy, one is usually interested in wavefront distortions in an aperture plane. This information can then be used for aberration correction [61] or deconvolution [59]. The following subsections give a brief description of the two prominent and commonly used phase retrieval methods, based on the Fourier transform and on the extended Nijboer-Zernike theorem.

Phase retrieval based on the Fourier transform

The famous Gerchberg-Saxton (GS) algorithm [62] was developed to retrieve the phase from intensity measurements in an aperture (or pupil plane) and in an image plane by iteration. The relationship between the electromagnetic field in the aperture and image plane is described by a Fourier transform, under the Fraunhofer approximation [16]. The iterative GS algorithm computes forward and inverse Fourier transforms of the aperture and the image plane until the calculated intensity patterns resemble the measured intensity data sufficiently. Constraints are applied in both Fourier planes by replacing the computed Fourier amplitudes with the square-root of the measured intensity. The GS algorithm minimises an error metric, calculated by taking the sum of the squared differences between the calculated Fourier modulus and the measured amplitude (i.e. the square root of the measured intensity) [62], until a pre-defined threshold is reached. A diagram of a GS algorithm for phase retrieval is shown in Fig. 2.13

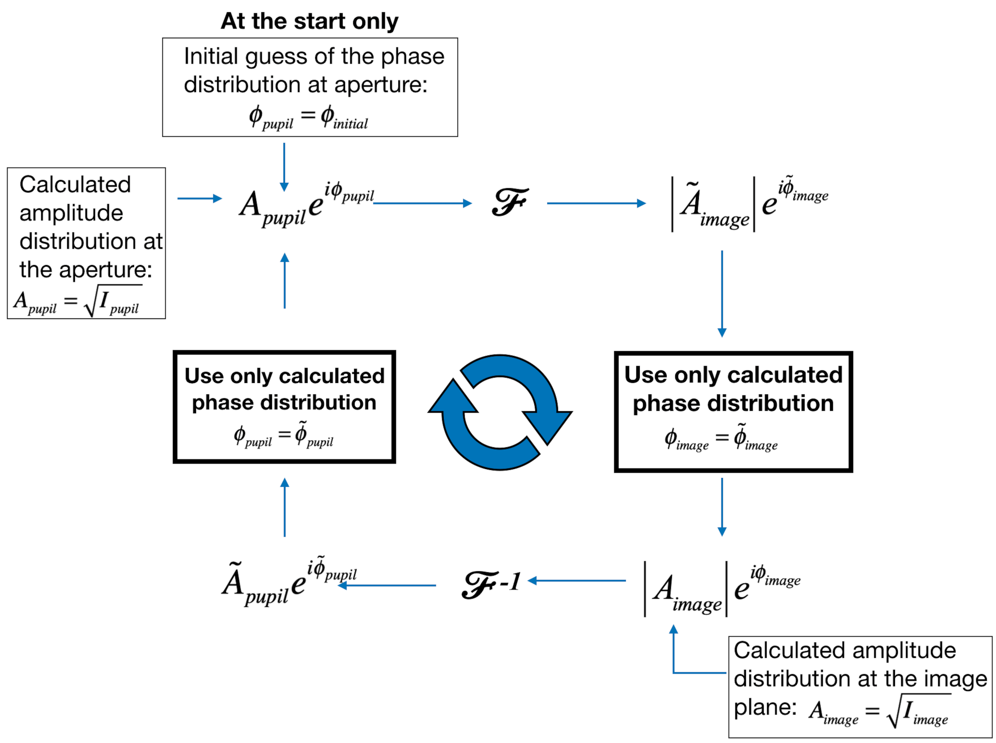


Figure 2.13: Iterative Gerchberg-Saxton algorithm diagram. $\mathscr{F}$ and $\mathscr{F}^{-1}$ represent the Fourier and inverse Fourier transform, respectively. I_{image} and I_{pupil} are the measured intensities in the image and aperture plane, respectively. $\boldsymbol{\phi}$ is a phase distribution function and A is an amplitude distribution function. The whole cycle is repeated until a predefined criterion is met.

Referring to Fig. 2.13, the GS algorithm starts with an initial guess for the phase distribution while the pupil amplitude distribution is calculated based on the measured intensity at the pupil plane. The Fourier transform of the complex amplitude function is then calculated. The obtained amplitude distribution $|\tilde{A}_{image}|$ is replaced by the square root of the measured intensity distribution $\left(\sqrt{I_{image}}\right)$ in the image plane. Next, the inverse Fourier transform of the modified complex amplitude function $|A_{image}|e^{i\phi_{image}}$ is calculated. The obtained amplitude distribution $\tilde{A}_{pupil}$ is replaced by the square root of the measured intensity distribution at the pupil $\left(\sqrt{I_{pupil}}\right)$. The whole cycle is repeated until a predefined criterion is met. When this criterium is met, the last found phase distribution is the final result. The original paper mentions that several iterations (> 60) were necessary for the GS algorithm to converge sufficiently

[62]. The downside of using a GS algorithm is that it can be slow and there can be ambiguities in the results produced [63, 64]. With phase diversity [65], phase retrieval algorithms are now more robust to noise, solve ambiguities [63] and, moreover, can also be applied to extended sources [65]. Phase diversity is implemented by taking a second image with an added phase function, usually a defocus. Phase diversity has been applied in astronomical telescopes [66], wavefront sensing [67] and also in microscopy [58-60]. In microscopy, one often uses high NA optics for which a phase retrieval algorithm extension is needed to take into account the vectorial character of light and polarisation effects. A scalar diffraction model as an approximation can still be used, when the curvature of the pupil function is taken into account by adding a quadratic phase term to the Fourier transform [58].

Extended Nijboer-Zernike Theory

Analytical expressions for the PSF exist for a limited number of cases. For an un-aberrated pupil, an analytical expression for the 3 dimensional PSF was found by Lommel [7]. Nijboer and Zernike derived expressions for weakly aberrated ($< 0.07\lambda$) focal spots in focus [9] and around focus (astigmatism and spherical). In most cases aberrated PSFs can nowadays be computed numerically with a FFT (a discrete Fast fourier transform). Braat and Janssen expanded upon the work of Nijboer and presented analytical solutions for calculating aberrated PSFs through focus [12] and allowing for high NA by taking polarisation into account. The method of Braat and Janssen is based on expanding the exponential function including the phase aberration terms using a Taylor series.

$$e^{i\Phi(\rho,\theta)} = 1 + i\Phi(\rho,\theta) - \frac{1}{2}\left[\Phi(\rho,\theta)\right]^2 + \ldots \tag{2.32}$$

where (ρ,θ) are cylindrical coordinates in the pupil plane and Φ is the phase distribution in the pupil. In the scalar extended Nijboer-Zernike theory (ENZ) [11-13] it was shown that the complex amplitude, at the image plane, can be described by:

$$U(r,\phi,f) = 2\sum_{n,m} \beta_n^m V_n^m(r,f) \begin{cases} \sin m\phi & \text{for m negative} \\ \cos m\phi & \text{for m positive} \end{cases} \tag{2.33}$$

where (r,ϕ) are cylindrical coordinates at the image plane and f is called the defocus parameter $f = -2\pi z$ with z being the axial coordinate, β_n^m are complex coefficients, which can be expressed in terms of the Zernike aberration coefficients, and the V_n^m functions are given by:

$$V_n^m(r,f) = \int_0^1 \rho R_n^m(\rho) e^{if\rho^2} J_m(2\pi\rho r)\mathrm{d}\rho = e^{if} \sum_{l=1}^{\infty} (-2if)^{l-1} \sum_{j=0}^{p} v_{lj} \frac{J_{m+l+2j}(2\pi r)}{l(2\pi r)^l} \tag{2.34}$$

where J are the first order Bessel functions and the coefficients v_{lj} are given by:

$$v_{lj} = \left[(-1)^p (m+l+2j) \begin{pmatrix} m+j+l-1 \\ l-1 \end{pmatrix} \begin{pmatrix} j+l-1 \\ l-1 \end{pmatrix} \begin{pmatrix} l-1 \\ p-j \end{pmatrix} \right] \bigg/ \begin{pmatrix} q+l+j \\ l \end{pmatrix} \tag{2.35}$$

The expressions in brackets represent binomial coefficients with $p = \frac{1}{2}(n-m)$ and $q = \frac{1}{2}(n+m)$.

In the presence of small aberrations (≤ 0.07λ), a simple linear approximation is possible which involves only the linear term in the Taylor expansion (Eq. 2.30). For large aberrations, the Taylor expansion of the phase function should also include higher order terms. Aberrated (intensity) PSFs through focus, calculated using the ENZ theory, are shown in Fig. 2.14 for a 0.5NA lens at a wavelength of 532nm in three focal planes (1um below focus, in focus, 1 um above focus).

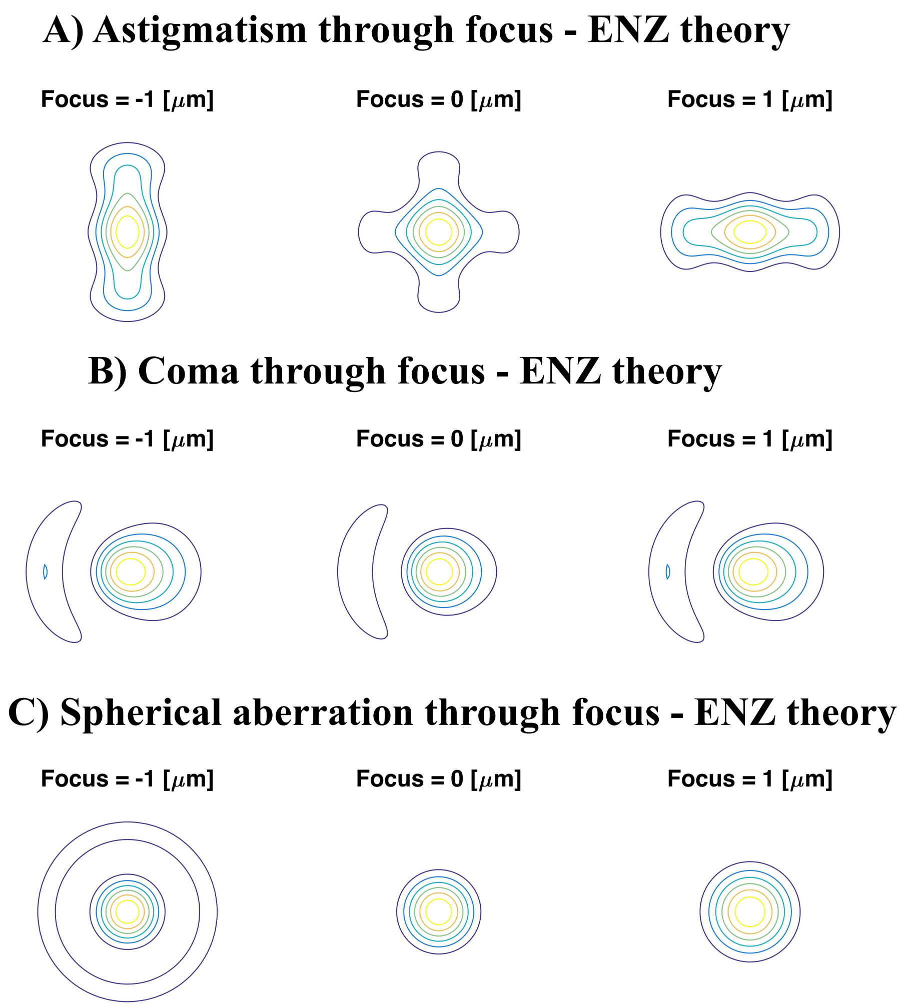


Figure 2.14: Aberrated PSFs through focus calculated using the extended Nijboer-Zernike theory. The calculations were done for a 0.5NA lens and a wavelength of 532nm and three focal planes (1um below focus, at focus, 1 um above focus).

For astigmatism (Fig. 2.14A), the focal spot undergoes a 90º rotation when going through focus. The coma focal spot (Fig. 2.14B) is symmetric around focus but the coma flare becomes more pronounced out of focus. Whereas for spherical aberration (Fig. 2.14C) the intensity distribution is asymmetric through focus. A scalar treatment of three dimensional focal spot intensity, in the presence of aberration, with the ENZ theory gives good results for NA up to 0.6. The scalar ENZ theory, described before, can be expanded to NA up to 0.8 by using a more complex expression for the defocus term [12]. For NA above 0.8, a vectorial ENZ theory may be used. The vectorial ENZ can take the vectorial character of the electromagnetic field, radiometric and polarisation effects into account [68]. The ENZ theory provides semi-analytical functions which describe the complex amplitude in the image plane based on the

Zernike polynomials which describe the wavefront in the pupil plane. Three dimensional aberrated intensity PSFs can be calculated using Equation (2.31) by multiplying the amplitudes with their complex conjugate. The inverse problem of retrieving the aberrations from intensity PSF data can be dealt with, in the case of small aberration amplitudes (less than about $< 0.07\lambda$), using either least square solvers or a method based on forming a set of equations using inner products [69]. For larger aberrations, an iterative retrieval method has been used [70].

2.7 Conclusion

This chapter covered the topics of scalar diffraction, optical aberrations, microscopy and wavefront sensing to give the reader an overview on the current state-of-the-art of aberration retrieval and adaptive optics in the field of optical microscopy. The information given should help to understand the upcoming Chapters which present novel aberration retrieval methods for laser scanning microscopy and proposes also a new strategy for indirect wavefront sensing.

3. Microscope with adaptive optics for confocal detection

The design and construction of a confocal microscope, incorporating Adaptive Optics (AO), is presented in this chapter. The purpose of the microscope was to study optical aberrations and their corrections [48], and to investigate new methods for aberration retrieval. Modal aberration correction and phase retrieval will be discussed in subsequent chapters. In 3.1 the general layout of the confocal microscope is presented. Distinct units of the confocal microscope: illumination, deformable membrane mirror, optical components, microscope objective and detection are discussed in this chapter. The topics covered in the following subsections are:

- 3.1.1: Microscope illumination
- 3.1.2: Deformable membrane mirror calibration and software control
- 3.1.3: Generation of unwanted aberrations in a double pass setup
- 3.1.4: Confocal Detection with a pixelated detector and a discussion of noise aspects

Finally in section 3.2 the microscope set-up which was built for aberration retrieval is described and characterised. In subsection 3.2.1, lateral and axial point spread functions obtained after correction of system aberrations are presented. In subsection 3.2.2, sectioning and imaging performance of the confocal microscope obtained by scanning a microscope resolution target in different focal planes are shown.

3.1 Experimental Setup

A first aspect to consider when building a scanning microscope is how to perform scanning. One uses either beam scanning or alternatively object scanning. Beam scanning requires moveable mirror systems in combination with a microscope objective that is corrected for aberrations over the whole scanning field. An advantage of object scanning is that there are, in principal, no requirements for field correction but it comes usually at the cost of lower scanning speed. The object is usually moved by a translation stage. This requires however that the object doesn't move with respect to the stage. For the setup used, object scanning was selected. The setup further includes Adaptive Optics (AO) for aberration correction, a schematic layout of the main units is given in Figure 3.1. We will first discuss the main building blocks of the microscope setup in detail in the following subsections. More technical details about the microscope setup as a whole will be given in sub-section 3.2.

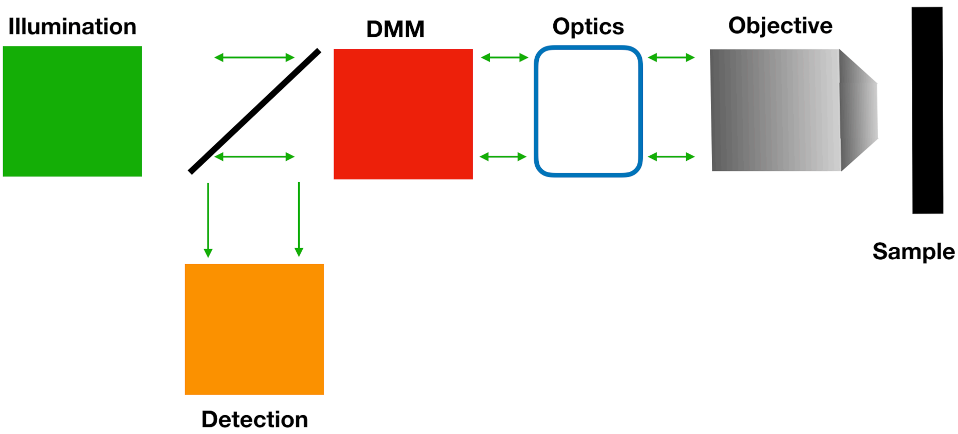


Figure 3.1: Scanning microscope setup with integrated Adaptive Optics.

3.1.1 Laser Fibre Illumination Unit

An illumination unit with a laser light source has the advantage of being monochromatic and is capable of delivering a high and constant intensity beam of coherent light. Laser light can either be used straight out of the laser unit or injected into a single mode fibre. The beam intensity profile coming out of the laser unit can be inhomogeneous. However, injecting laser light into a single mode fibre would produce, at the fibre output, an almost Gaussian beam intensity profile. For the setup, a green 532nm laser (B&W Tek) was injected into a single mode fibre with a lens (Qioptiq kineFlex single mode fibre system, cutoff wavelength 473nm), the fibre exit served as point source in the illumination unit of the scanning microscope. The laser beam exiting the fibre was collimated with a collimation lens L1 (focal length: 60mm) to a beam whose diameter was reduced to about 11mm with an iris (implies an NA of about 0.09). The illumination unit is shown in Figure 3.2

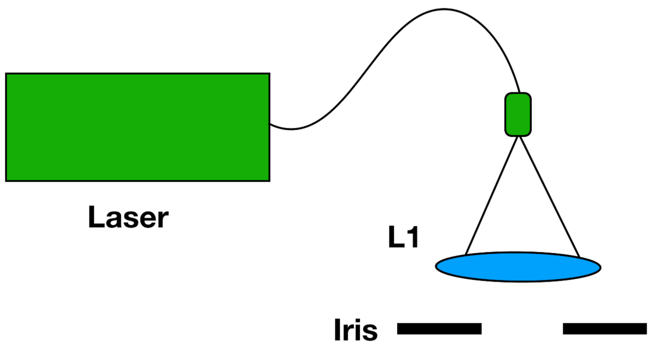


Figure 3.2: Laser-Fibre Illumination Unit.

3.1.2 Deformable mirror

The DMM used in the microscope was a Mirao52e, which was described in Chapter 2 subsection 5.2.3. The DMM was calibrated using a Shack-Hartmann Wavefront sensor (WFS150 Thorlabs, sensitivity $\lambda/50$ root mean square wavefront deviation). The dimensions of the CCD are 5.95 x 4.65 mm and the lenslet array comprises 1210 lenslets, about 550 of the lenslets were used for the calibration of the DMM, thereby assuring for reliable wavefront sensing, ease of alignment and avoiding overfilling. Behind each lenslet there are about 31x31 pixels to detect focal spot displacements. The focal spot size, produced by each lenslet, is about 45μm (or about 10 pixels) on the CCD. The WFS performed 10 frames averaging and the average signal was used for the Zernike analysis. A simple setup for calibrating a DMM with a WFS is shown in fig 3.3. The DMM surface is conjugated with the WFS lenslet array by means of the lens pair L2 and L3. The lens pair L2 and L3 also reduced the size of the beam to fit the WFS.

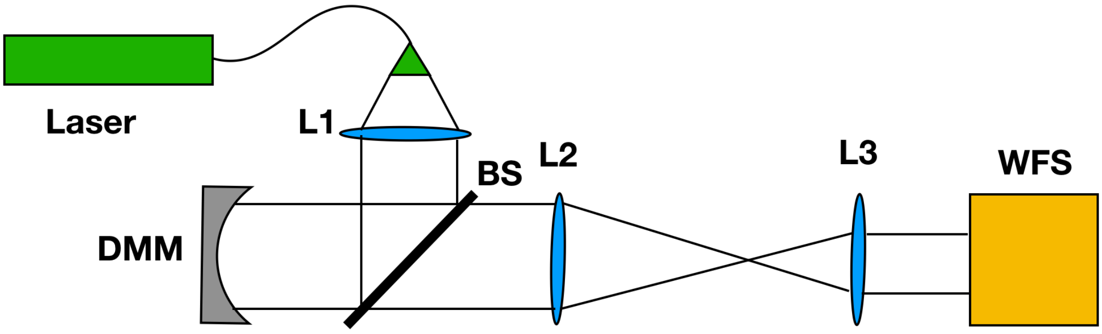


Figure 3.3: AO Calibration Setup. **L**: Lens; **BS**: Beamsplitter; **DMM**: Deformable Membrane Mirror; **WFS**: Wavefront Sensor;

The calibration procedure is explained below.

The wavefront detected by the WFS is decomposed by Thorlabs software into Zernike aberrations. We assume that the effects of the actuator control signals on the wavefront can be described by:

$$\vec{Z} = M \cdot \vec{V} + \vec{Z}_0 \tag{3.1}$$

where $\vec{Z}$ is the Zernike coefficients vector that describes the wavefront, V is the actuators voltage vector, and M is the influence matrix which describes how voltage signals translate into Zernike coefficients. $\overrightarrow{Z_0}$ is the Zernike coefficients vector that describes the initial shape of the mirror surface.

The previous equation can be re-arranged as:

$$\Delta\vec{Z} = \vec{Z} - \vec{Z}_0 = M \cdot \vec{V} \tag{3.2}$$

In order to determine a control Matrix C which translates a specific wavefront into voltage signals, the influence functions of every single actuator is measured first. The wavefront that is obtained when a specific actuator is activated is denoted the influence function of that actuator. The actuator influence functions are stored in the respective column of the influence matrix M.

An example of influence functions from actuators 3, 8, 15, 23, 31, 39, 46, 51 (which are arranged along a line as shown in Fig. 3.4) taken with respect to the flat mirror surface are shown in Fig. 3.4. The influence function of actuator 23 shows the most symmetrical and centred wavefront distortion pattern. This suggests that the wavefront centre is somewhat displaced in the direction of actuator 23.

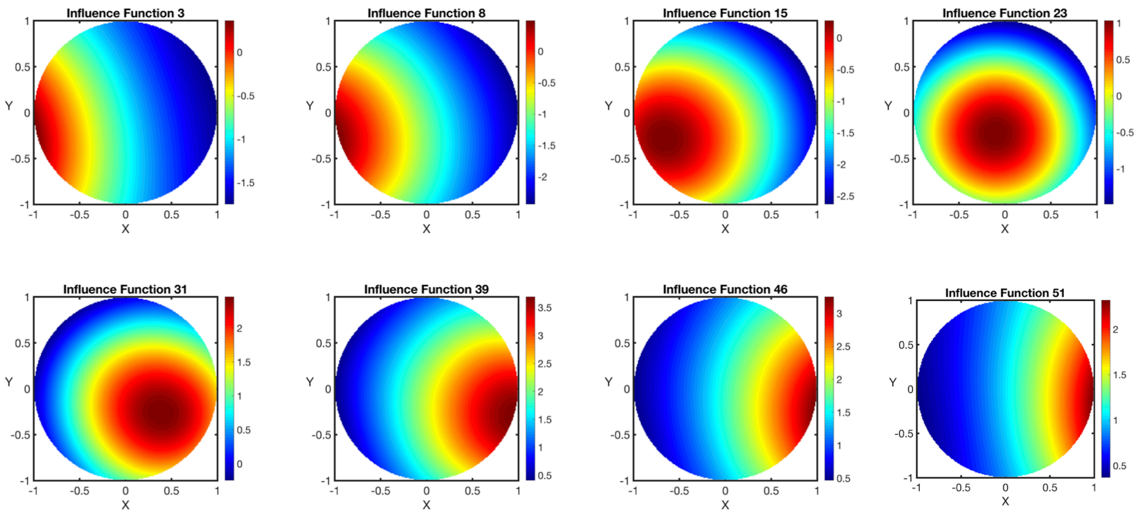


Figure 3.4: Influence functions of actuators 3, 8, 15, 23, 31, 39, 46, 51. Red colour represents highest amount of displacement while blue represents lowest amount of displacement. The scales are different for each influence function. Actuators 3 … 51 are arranged along a line below the DMM centre in Fig. 3.4. The Figure suggests further that the wavefront centre is somewhat displaced in the direction of actuator 23.

The control matrix is obtained by taking the inverse of M. Since M is not a square matrix a singular value decomposition is used to calculate the pseudo-inverse:

$$M = U\Lambda V^T \tag{3.3}$$

where U,Λ,V are three matrices. U is an orthogonal matrix containing the eigenmodes of the mirror whereas V is an orthogonal matrix containing the eigenmodes of the control signals. Λ is a diagonal matrix containing the eigenvalues [34, 71]. The inverse is then given by:

$$C = V\Lambda^{-1}U^T \tag{3.4}$$

To minimise noise contributions and to avoid mirror saturation, small singular values are set to zero [35, 71]. The required voltages to produce a specific wavefront can now be computed via the following equation:

$$\vec{V} = C \cdot \vec{Z} \tag{3.5}$$

The control algorithm, described before, is based on a linear model (see equation 3.1). To verify the linearity of the DMM, the deformation response of the mirror was studied by measuring the Peak-to-Valley (PV) values of actuators 5, 13 and 22. The following voltages were applied [0.05, 0.1, 0.2, 0.4, 0.8] Volts and the PV recorded by the SH sensor are plotted in fig 3.5. The data was fitted with straight lines having an offset and the respective R^2 values were calculated. For all three actuators the R^2 value was close to one, meaning that the data is well described by a linear relationship and that the wavefront deformation, in terms of PV, scales linearly with the voltage applied to an actuator. The mirror shows a good linear response and the data is in agreement with the results found in the literature [33].

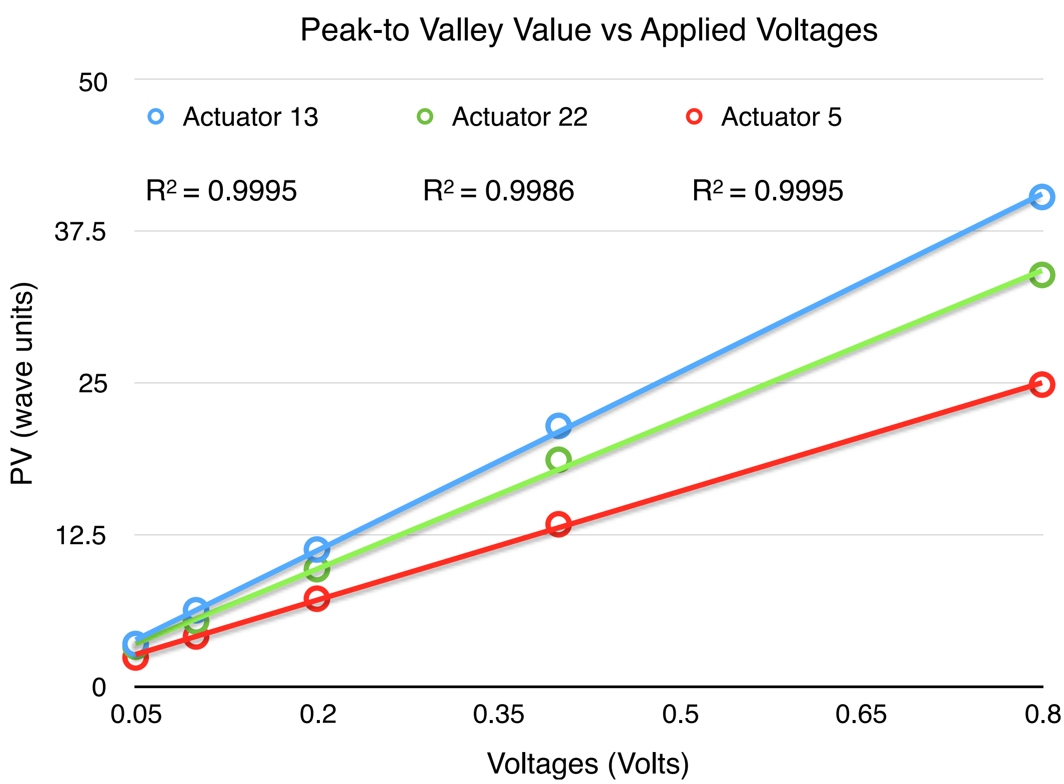


Figure 3.5: Peak-to-Valley Values of actuators 5, 13 and 22 plotted as a function of applied Voltages. A Linear fit was done and the R squared values are shown on the graph

Not all deformable mirror types show a linear deformation response which increases linearly with applied voltage. Electrostatic DMMs, for example, exhibit a non-linear response which becomes approximately linear when applying quadratic voltage increments [72].

Closed-loop AO performance

A closed-loop AO system is a feedback system, wherein the response to signals, which are applied to a DMM are measured by a WFS, and small corrections are iteratively applied to the signals until a predefined quality criterion is reached. The detected wavefront can be expressed in terms of Zernike polynomials as follows:

$$\Psi = \sum_{n=0}^{\infty} a_n Z_n \tag{3.6}$$

where Ψ is the wavefront, a_n are the Zernike aberration coefficients and Z_n^m the Zernike polynomials. To assess the wavefront quality of imposed aberrations we used the root-mean square (RMS) wavefront error $\Delta\Psi$:

$$\Delta\Psi = \sqrt{\frac{\int_0^1 \int_0^{2\pi} \left(\Psi - \overline{\Psi}\right)^2 \rho \, \mathrm{d}\rho \, \mathrm{d}\theta}{\int_0^1 \int_0^{2\pi} \rho \, \mathrm{d}\rho \, \mathrm{d}\theta}} = \sqrt{\overline{\Psi^2} - \left(\overline{\Psi}\right)^2} = \sqrt{\sum_{n=1}^{\infty} a_n^2} \tag{3.7}$$

where $\overline{\Psi}$ is the average wavefront (over the unit circle). The last equation is obtained when expressing the wavefront in terms of Zernike polynomials (using the Malacara notation [19]) by making use of their orthonormal properties. It is noted that the RMS wavefront error is also used to specify the quality of diffraction-limited optical imaging systems. An optical imaging system is said to be diffraction limited when it has a Strehl ratio of 0.8, which corresponds to a RMS wavefront error of < 0.071λ rms (often referred to as the Maréchal criterion [73]).

After each closed-loop iteration, the wavefront difference Ψ_{diff} between the desired $\Psi_{desired}$ and a detected wavefront $\Psi_{detected}$ is:

$$\Psi_{diff} = \Psi_{desired} - \Psi_{detected} \tag{3.8}$$

The RMS of the wavefront difference Ψ_{diff} (equation (3.8)) was used to assess the quality of the imposed aberration after each closed-loop iteration and is further denoted residual RMS wavefront error.

The correction applied at each close loop iteration is given by the following equation [74]:

$$\vec{V}_{n+1} = \vec{V}_n + \alpha \cdot C\left(\vec{Z}_{desired} - \vec{Z}_{detected}\right) \tag{3.9}$$

where: $\overrightarrow{V_n}$ is a vector containing the control voltages for each actuator; α is a control parameter varying from 0 to 1 which determines by how much the signal changes per iteration: C is the control matrix which transfers differences in Zernike terms into corresponding actuator voltage responses: and the $\vec{Z}$ terms represent vectors containing the Zernike coefficients of the desired and detected wavefront. The user interface of Labview software, which was programmed from scratch, for operating the DMM and WFS in open- or closed-loop is shown in Fig. 3.6.

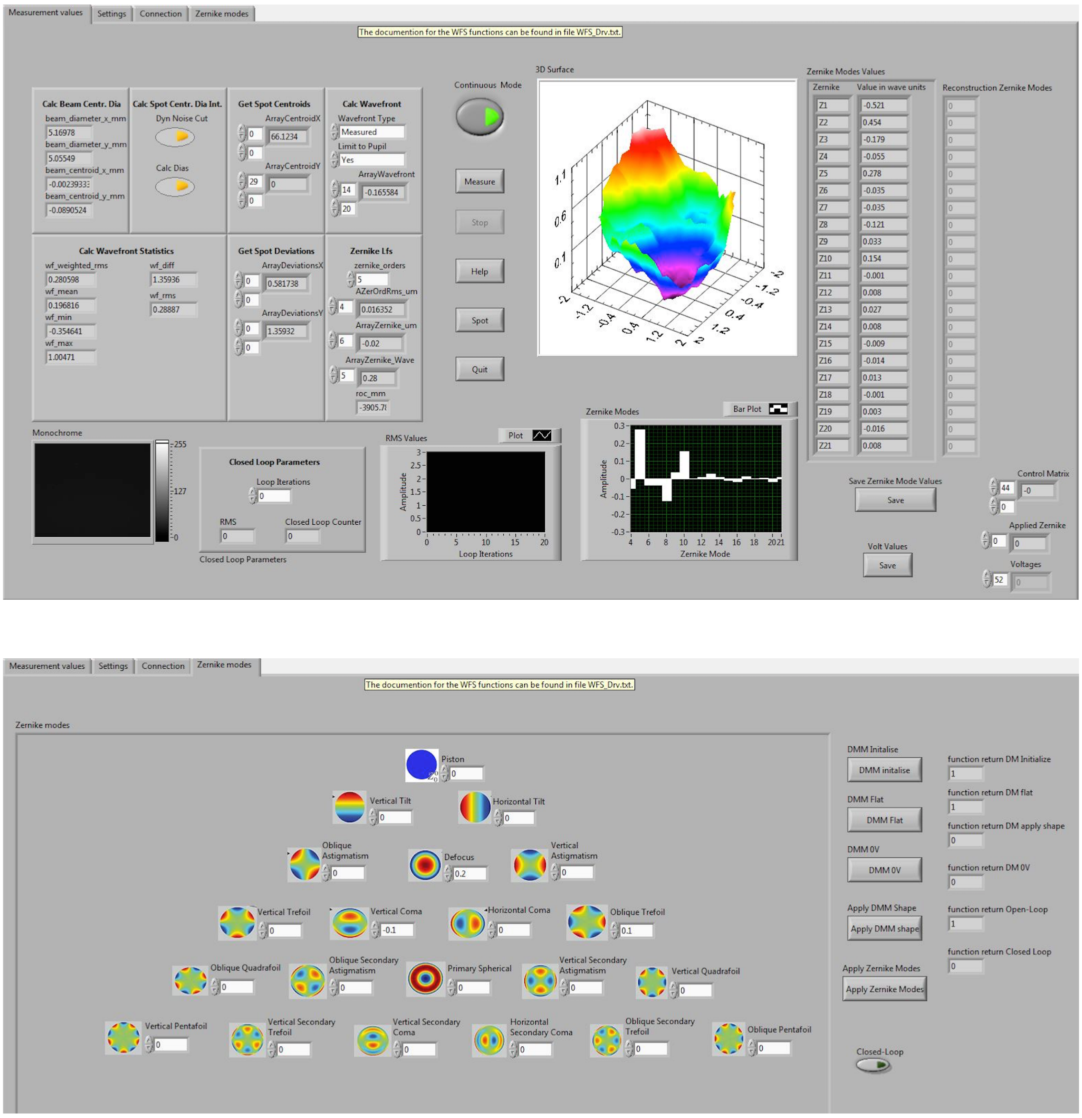


Figure 3.6: Labview Open/Closed-Loop Software user interface. Top interface is the wavefront sensor panel. Bottom interface is the deformable membrane mirror panel.

For closed-loop a value of 0.4 was used for the control parameter, as recommended previously [74]. A control parameter of 0.2 was also tested. It was found that the recommended value of 0.4 for the control parameter gives indeed give the best results in terms of residual RMS wavefront error Ψ_{diff} . The residual RMS wavefront errors Ψ_{diff} (see equation (3.10)), for primary vertical astigmatism, primary vertical coma, vertical trefoil, quadrafoil, secondary vertical astigmatism and spherical aberration, obtained for control parameter values of 0.2 and 0.4 in the microscope setup, which is described in section 3.2, are shown in Fig. 3.7. Similar results were obtained for other aberration orientations. For each aberration an amplitude of 0.2 λ rms was applied at the beginning. As expected, a low value of 0.2 for the control parameter needs more iterations to converge. The diffraction limit (0.071λ rms), for the different primary aberrations, quadrafoil and secondary astigmatism, was on average reached after about 8 iterations for a value of 0.4 of the control parameter. For a control parameter value of 0.2, the higher order aberrations (spherical, secondary astigmatism) didn't reach the value of 0.071λ rms within 20 iteration steps. The lower order (primary) aberrations (astigmatism, coma, trefoil) converged faster. The use of larger values for the control parameter leads to oscillations in the residual RMS wavefront error. Oscillations were observed for a control parameter of 0.7 (no data shown).

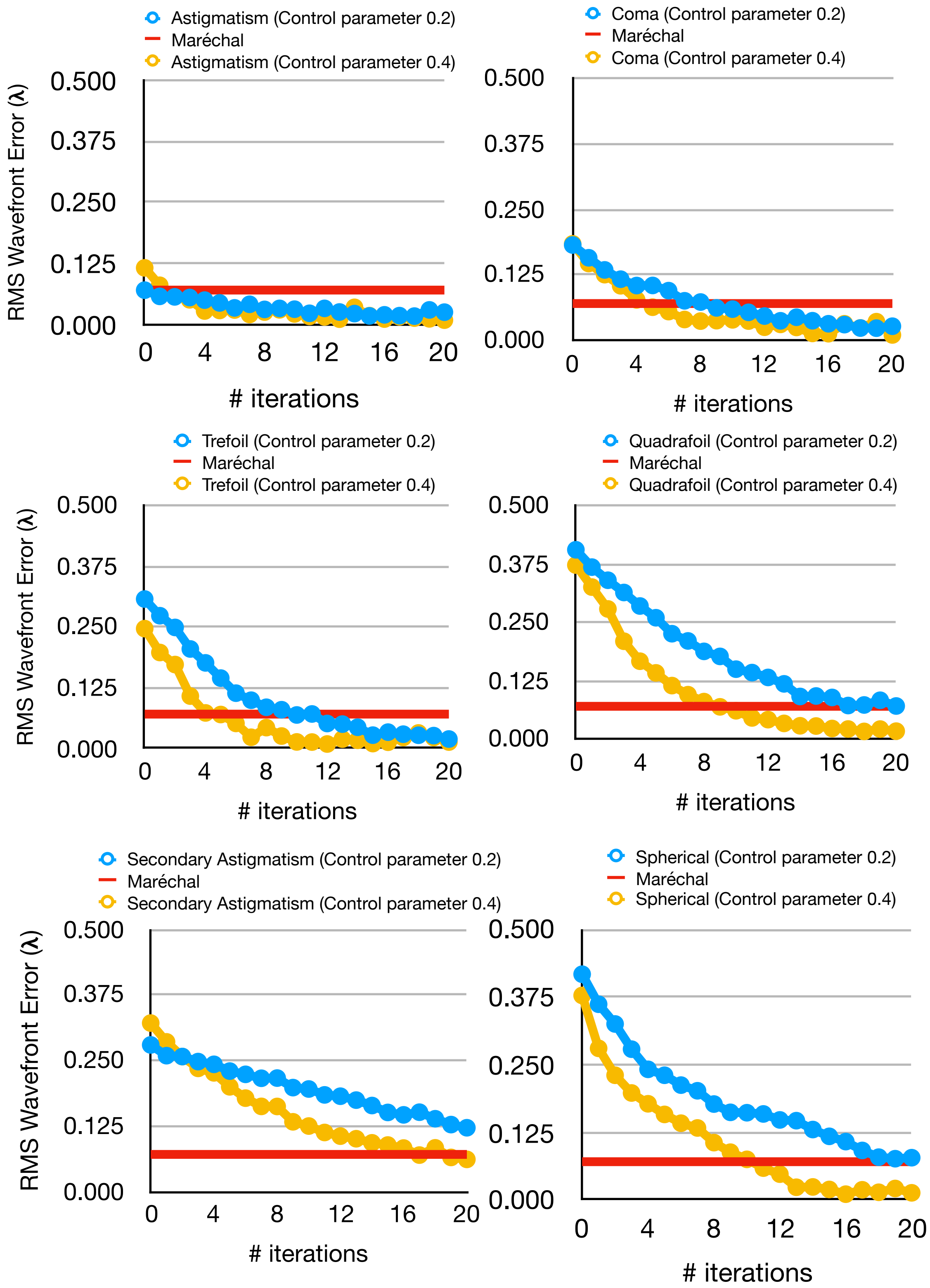


Figure 3.7: Closed-Loop Convergence curves. RMS values for astigmatism, coma, trefoil, quadrafoil, secondary astigmatism and spherical aberration are plotted versus number of iterations for control parameter 0.2 (blue) and 0.4 (yellow)

From the above it is clear that a value of the control parameter of 0.4 is well suited to our purpose and it was therefore used for further closed-loop control experiments. An iteration step was done every 2 seconds for better visualisation. With 10 frames averaging by the Shack Hartmann WFS (frame rate 15Hz max), the maximum correction speed obtained was 0.7 seconds. We tested the effect of omitting the tip/tilt value contributions in the closed-loop correction with a control parameter of 0.4. Figure 3.8 shows the RMS wavefront error curves of primary vertical astigmatism, primary vertical coma, trefoil and spherical aberration with/without taking tip and tilt into account.

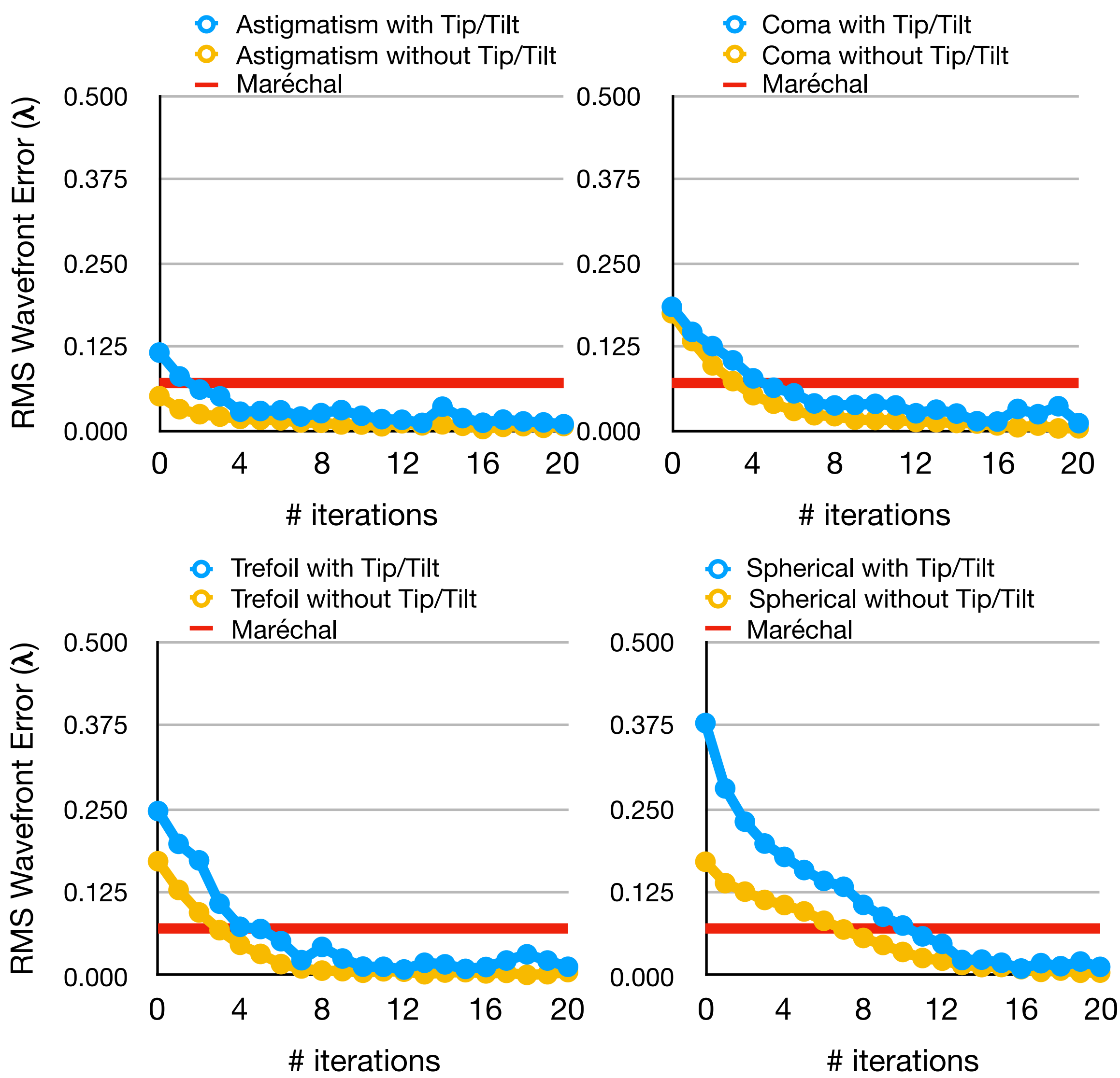


Figure 3.8: Closed-Loop Convergence curves. RMS values for astigmatism, coma, trefoil and spherical aberration are plotted versus number of iterations for control parameter 0.4 and in the presence (blue) of tip (Z2) and tilt (Z3) and without (yellow)

The figures show that tip/tilt affect the overall RMS wavefront error and more iterations are necessary to reach diffraction limited performance. It is noted that the Zernike modes tip, tilt, and defocus term Z_2^0 do not distort the focal spot per se, but shift it laterally or axially. Figure 3.10 and 3.11 show that the most common aberrations in microscopy (primary astigmatism, primary coma, primary spherical aberration) [36] can be sufficiently well corrected by the AO closed-loop system. Figures 3.10 and 3.11 show that coma and trefoil, which are both aberrations of radial order 3, converge faster than primary spherical which is an aberration of radial order 4. Compared to the results by Polo et al. the correction speed and the degree of improvement were higher. This can be explained by the properties of DMM (Mirao52e) used. The Mirao52e displays a good linear deformation response to applied voltages, whereas the mirror used by Polo et al. (Adaptica Srl.) has a non-linear (quadratic) response which requires additional corrections. It was shown that the control algorithm which does not take tip/tilt into account achieves diffraction limited performance for each aberration to be reached with less iterations and attains smaller values for the RMS wavefront error than a control algorithm algorithm which takes tip/tilt into account. Omitting tip/tilt during the closed-loop calibration would be detrimental for confocal microscopy because one doesn't want the focal spot to be shifted outside the pinhole. Zernike modes were calibrated in closed-loop (WFS sensitivity of λ/50 rms) and these calibrated mirror shapes were used in later experiments.

3.1.3 Microscope objective lens and relay optics

The microscope objective used was 0.75 numerical aperture (NA) multi-immersion Nikon objective lens (CFI Plan Fluor 20XMI (multi-immersion)). For all experiments, water was used as the immersion medium. In order to perform aberration correction, the DMM surface should be conjugated to the objective's back pupil. Furthermore, the beam size at the DMM should be adapted to the diameter of the objective pupil. Therefore, the Gaussian beam overfilled the microscope objective's pupil and was truncated to about 40% of its maximum intensity at the edge. These two conditions were met by placing a 4f lens system (the 4f optics unit in Fig. 1 which is, effectively, a Keplerian telescope) between the DMM and the microscope objective. The design of a 4f system is shown in Fig. 3.9. The magnification of the 4f system from DMM to objective was about -1.5 with a beam size of about 11mm at the DMM and roughly 17mm at the microscope objective.

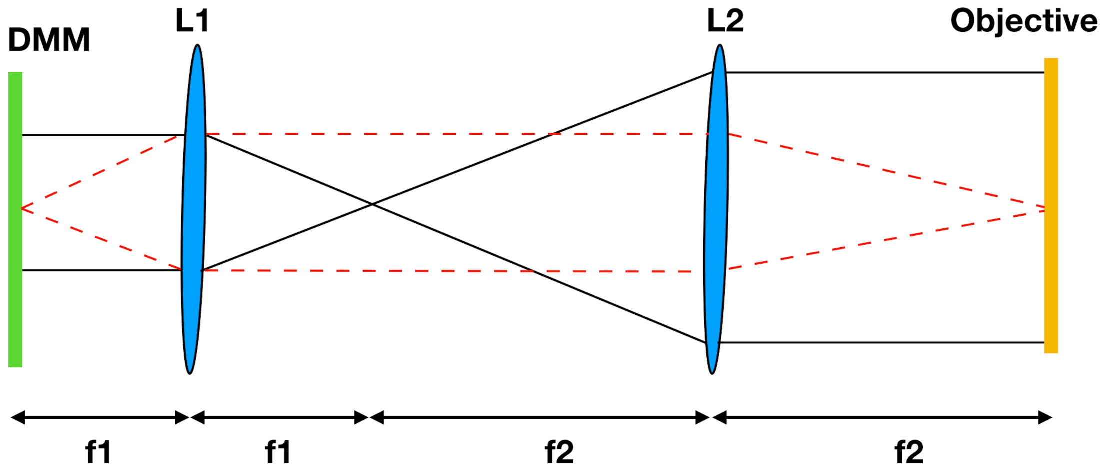


Figure 3.9: 4f system. L: Lens; f: focal length

When applying Zernike modes with a DMM, the presence of additional unwanted Zernike modes can compromise the performance of adaptive optics corrections. These contaminations can either be caused by residual Zernike modes generated by the DMM or misalignment of the optical system. In the following, we describe the aberration contamination caused by de-centring a lens pupil with respect to the optical axis.

The effect of de-centring the pupil of a lens in the system can be explained as follows. A wavefront can be described in terms of Zernike's polynomials. The Zernike polynomial Z_n^m are functions of the variables (x,y), which, when expressed in polar coordinates $\left(x = \rho\cos(\theta), y = \rho\sin(\theta)\right)$, can be of the form:

$$Z_n^m\left(\rho\cos\theta, \rho\sin\theta\right) = R_n^m\left(\rho\right)e^{il\theta} \tag{3.10}$$

where m (azimuthal order) and n (radial order) are non-negative integers, $n \geq |l|$ and $n-|l|$ is even (see [7]).

The radial polynomial $R_n^m(\rho)$ that describes the lowest order coma is:

$$R_3^1(\rho) = 3\rho^3 - 2\rho \tag{3.11}$$

and can also be written, in terms of (x,y), as

$$R_3^1(x,y) = 3\left(\sqrt{x^2+y^2}\right)^3 - 2\left(\sqrt{x^2+y^2}\right) \tag{3.12}$$

Now we consider the effect of de-centring the pupil. We assume that polynomial extends beyond the unit circle and introduce a small shift Δx in the x direction, i.e. for a displacement $\tilde{x} = x + \Delta x$. Developing $R_3^1(x,y)$ in a Taylor series (for the displacement in the x direction) then gives:

$$R_3^1(\tilde{x},y) = \left\{3\left(\sqrt{x^2+y^2}\right)^3 - 2\left(\sqrt{x^2+y^2}\right)\right\} + \left\{9\left(\sqrt{x^2+y^2}\right)x - \frac{2x}{\sqrt{x^2+y^2}}\right\}\Delta x + \ldots \tag{3.13}$$

or in polar coordinates:

$$R_3^1(x,y) = \left\{3\rho^3 - 2\rho\right\} + \left\{9\rho^2 - 2\right\}\cos\theta\ \Delta x + \ldots \tag{3.14}$$

Considering just the first two terms and inserting this expression in the real Zernike polynomial [7] describing coma, Z_3^{-1} or alternatively Z_3^1 (depending on the orientation), we obtain:

$$\begin{aligned} Z_3^{-1}(\rho,\theta) &= R_3^{-1}(\rho)\sin\theta = \left(\left\{3\rho^3 - 2\rho\right\} + \left\{9\rho^2 - 2\right\}\cos\theta\ \Delta x\right)\sin\theta \\ &= \left(\left\{3\rho^3 - 2\rho\right\}\sin\theta + \frac{1}{2}\left\{9\rho^2 - 2\right\}\sin 2\theta\ \Delta x\right) \end{aligned} \tag{3.15}$$

The first term represents coma. The second term, which is linear in Δx, describes an oblique astigmatism combined with an offset term. For Z_3^1, the second term varies with $\cos 2\theta$, representing a vertical astigmatism. Thus, astigmatism increases when increasing the displacement of the pupil. Analogously, we obtain for spherical aberration

$$Z_4^0 = R_4^0(\rho)\sin\theta = \left\{6\rho^4 - 6\rho^2 + 1\right\} + \left\{24\rho^3 - 12\rho\right\}\cos\theta\ \Delta x \tag{3.16}$$

The second term, which is linear in Δx, describes coma and tilt contributions. Thus by applying Zernike modes on the DMM, additional modes would be produced by an odd aberration in a system with a de-centred pupil, although the illumination and reflection path are overlapping. Figure 3.10 illustrates the effects of decentration on a coma aberration.

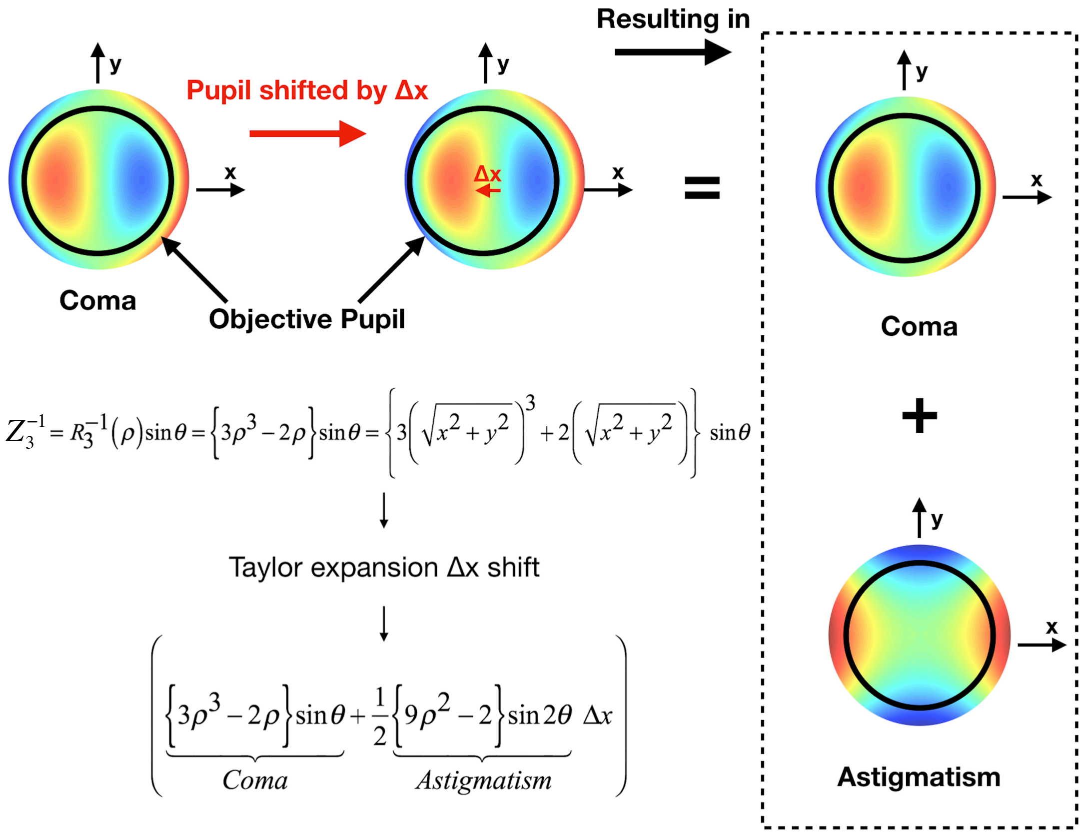


Figure 3.10: Effects of decentration on coma aberration

3.1.4 Pinhole Detection

The detection unit (see Fig. 3.1) of the confocal microscope consists of a light detector with a spatial filter. In most confocal systems, a photomultiplier tube (PMT) and a mechanical pinhole are used for confocal detection. This detection unit is shown in fig, 3.11. The circular pinhole is placed at the focal plane of the focussing lens (and is conjugated with the object plane). The pinhole diameter is about the size of the Airy disk, often referred to as an Airy unit (AU).

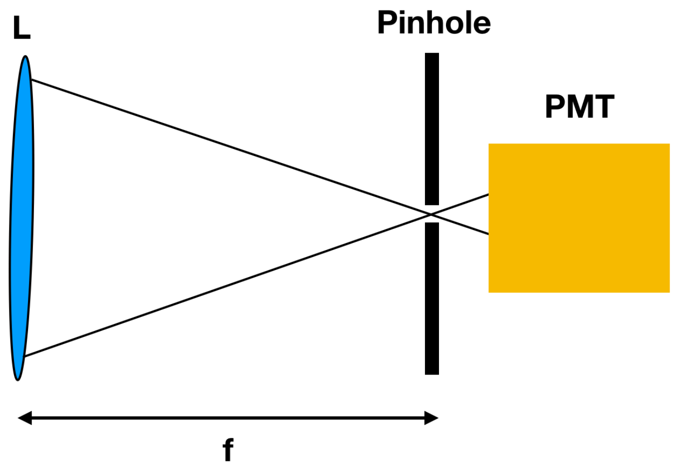


Figure 3.11: Confocal Detection unit consisting of a focussing lens L with a focal length f, a pinhole and a PMT: photomultiplier tube.

As an alternative to the pinhole and PMT, confocal detection can be achieved by placing a camera at the focal plane of the lens and summing the intensity signal of specific camera pixels. In a similar approach to that presented by See et al., confocal detection is achieved by summing the intensity of a number of pixels on the camera [75] . By adding only signals of pixels inside the pinhole area, the function of circular pinholes of different diameters can be approximated. An example of such a pinhole mask is shown in fig 3.12.

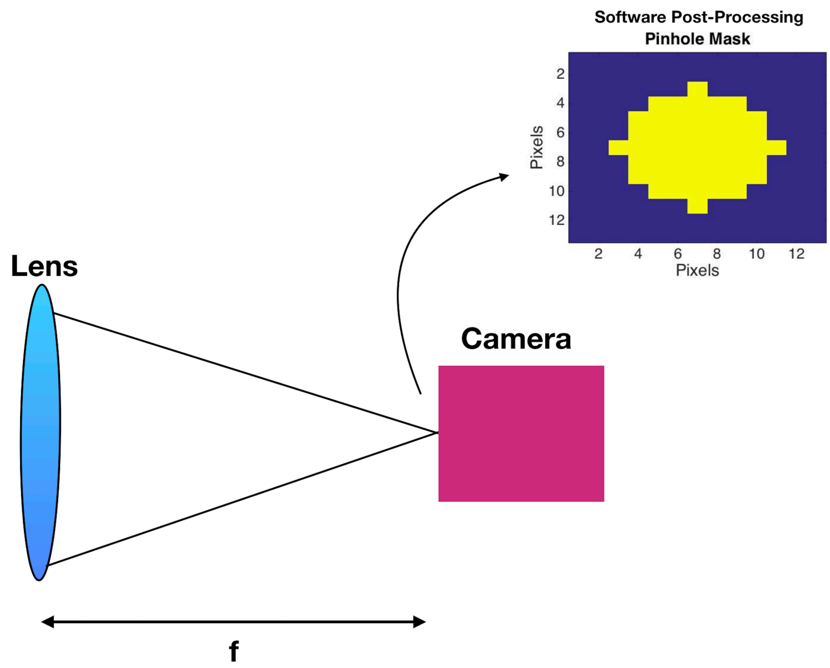


Figure 3.12: Confocal Detection unit using a camera and a pinhole mask; focussing lens L with a focal length f.

The central pixel of the pinhole mask is placed on the camera pixel with highest intensity. Pixels whose centre lie outside the pinhole mask do not contribute to the overall signal. The size of the Airy disk diameter on the camera depends on the magnification of the imaging system. In the experimental setup used here an electron multiplying charged couple device (EMCCD) camera was used with a lens having a focal length of 500mm. The Airy disk size was measured to be about 9 pixels in diameter. An Airy disk pinhole could be approximated by using a pinhole mask with a 9 pixel diameter, i.e. by adding the signals from 49 pixels of a 9x9 pixel sub-array (see Figure 3.12).

In a CCD detector, photons are converted into electrons by means of the photoelectric effect in the photo-sensitive semi-conductor structure of each pixel. Then the electrons are stored in the depletion zone of the semi-conductor until the full capacity is reached (full well reservoir capacity) [76]. The electrons are then shifted through a serial register by applying clock pulses to electrodes on the CCD chip [77]. The charges ultimately reach the output amplifier and an AD converter. Different noise sources in CCDs will be discussed in the following subsection. It is noted that an EMCCD, in comparison to a traditional CCD, has all the benefits of a CCD but it also reduces the readout noise, in low light conditions, because it uses on-chip gain circuitry provided at the output of the shift register, which by mean of impact ionisation generates additional electrons [77] , thereby amplifying the signal input for the AD converter. Furthermore, EMCCD sensors are usually cooled down to temperatures below -40ºC to reduce the amount of dark current.

Noise

The performance of a CCD detector is noise limited. The most commonly encountered noise sources in CCDs are shot noise, dark noise and readout noise. We will briefly review these different sources of noise and explain how they affect the EMCCD detector (Andor iXon 885) used for the experiments carried out in this thesis.

Shot noise:

Due to the quantum nature of light, the number of photons detected by a CCD varies from exposure to the next. The detected signal produced by the statistical arrival of photons can be described in terms of a Poisson distribution which for a large number of photons can be approximated by a Gaussian distribution. The noise associated to this phenomenon is often referred to as shot noise [78] and the standard deviation is given by:

$$N_{shot} = G \cdot F \cdot \sqrt{\eta \phi_p t} \tag{3.17}$$

where G is the CCD gain, F the noise factor, η the quantum efficiency, ϕ_p the mean incident photon flux in photons per pixels and t the exposure time. Typical values for F are 1 for a CCD, 1.3 for an EMCCD and 1.6-2 for ICCD (intensified CCD) [79]. The expression inside the square root represents the number of photons detected by the pixel during an exposure. For all non fluorescent experiments carried out in this thesis, no gain was applied to the EMCCD and thus G = 1. The full well capacity of the EMCCD is 30000 electrons. The shot noise contribution is therefore:

$$N_{shot} = 1.3 \cdot \sqrt{30000} \approx 225\ e^{-} \tag{3.18}$$

Dark noise:

Dark noise is generated by thermal fluctuations creating electron holes (absence of electrons) in the photosensitive semi-conductor of the pixel. An electric current will flow to compensate for the electron charge deficit at the electron holes. This unwanted dark current increases with temperature as more electron holes are created by thermal excitation. The dark current is given by [78]:

$$N = \left[2.55 \cdot 10^{15} N_{dc0} t \cdot d_{pixel}^{2} T^{\frac{3}{2}} e^{-\frac{E_g}{2kT}} \right]^{\frac{1}{2}} \tag{3.19}$$

where N_{dc0} is the dark current at 300K, d is the pixel size and T is the temperature and E_g is the energy band gap of the semi-conductor material. As can be seen from equation (3.19), the dark current decreases with decreasing temperature. The EMCCD camera was cooled down to -50°C and according to the manufacturers specification, the dark noise is about 0.06 e^{-} (mean value which has been extrapolated to -50°C).

Readout noise:

Readout noise is generated by the electronic circuitry of the camera [78] when electrons are transferred, shifted through the shift register, amplified and converted to a digital signal with an AD converter. These processes add noise to the signal by means of additional and unwanted electrons. The readout noise depends on the frame rate and will increase at higher frame rates [78]. The readout noise for a single pixel, according to the manufacturer, is about 28 e^{-}.

Total noise:

In the case of uncorrelated noise sources and Gaussian noise distributions, the different noises add up as the square root of the sum squared of each noise source:

$$N_{total} = \sqrt{N_s^2 + N_{dc}^2 + N_R^2} \qquad (3.20)$$

Because the dark current and readout are considerably less than the shot noise, the EMCCD is effectively limited by shot noise and, as a consequence, equation 4 simplifies to:

$$N_{total} \approx N_s \approx 225e^- \qquad (3.21)$$

The EMCCD is well known to be less affected by noise than traditional CCD detectors due to the cooling of the sensor but also because of the on-chip amplification of low intensity signals before the electric signals reaches the AD converter (when working in low light conditions). The signal to noise ratio (SNR) is given by:

$$SNR = \frac{\text{Number of detected photons}}{N_{total}} \approx \frac{\mathrm{N}_s}{\sqrt{N_s}} = \sqrt{N_s} \approx 225 \qquad (3.22)$$

3.2 Confocal Setup

The configuration of the confocal microscope, with integrated AO, used in our experiment is shown in Figure 3.13. Laser light (wavelength of 532 nm, continuous wave), injected in a single mode fibre, provides the illumination for the object scanning confocal microscope. Aberration correction is performed with a Mirao 52e (Imagine Optics, France) DMM. The DMM is conjugated to the pupil of the microscope objective (Ob) via a lens pair, L2 and L3 constituting a 4-f system (with a magnification -1.5). The microscope objective is a 0.75 numerical aperture, multi-immersion Nikon objective lens (20x magnification). The reflected light from the sample is focussed via lens L4 on an electron multiplying charge coupled device (EMCCD; iXon 885 Andor). Additionally, a LED light source was placed behind the sample stage for sample inspection in wide-field. The sample stage (PI Instruments P-733.3DD; a 3D stage with open-loop resolution of 0.1nm) is piezo driven in order to scan the sample and build up an image in x, y and z. The angle of incidence of the laser beam on the DMM is about 14^0 with respect to the normal. For fluorescent imaging, a fluorescent filter was placed between the the beamsplitter (BS) and the lens L4. Configuration b) and c) will be explained later in section 3.1.2.1 System aberration correction. The confocal microscope with the EMCCD in the optical path is represented in configuration a).

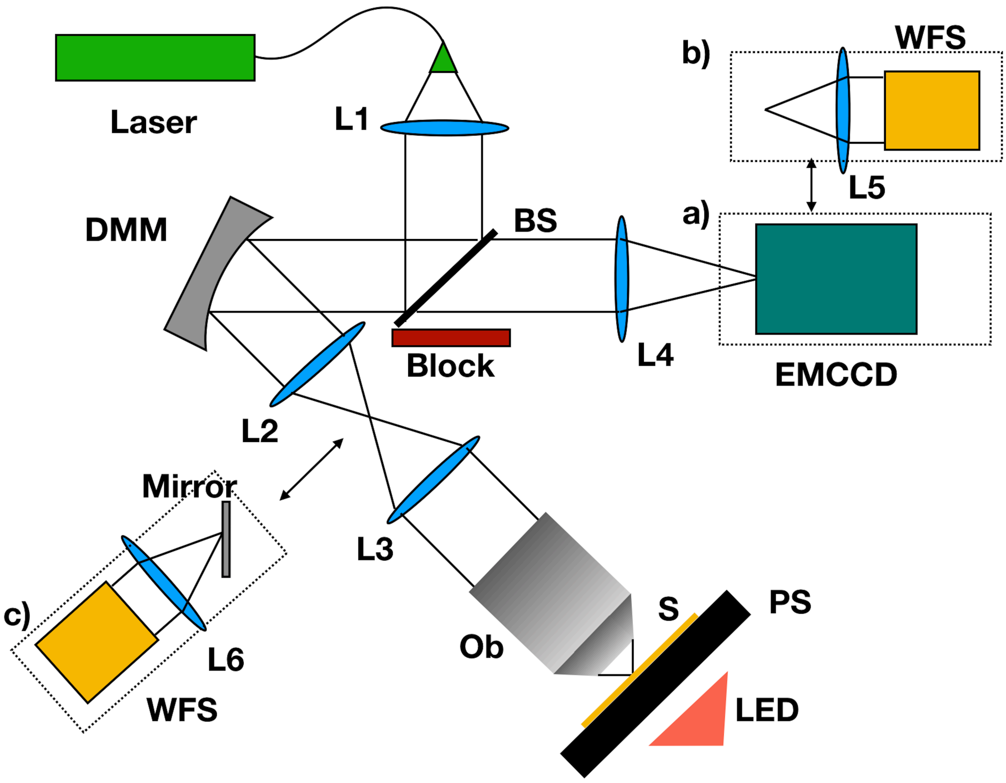


Figure 3.13: Optical confocal AO setup. L: Lens; BS: Beamsplitter; BS: Beamsplitter plate; DMM: Deformable Membrane Mirror; WFS: Wavefront Sensor; Ob: Objective; S: Sample; PS: Piezo-stage;

The DMM acts on both the incident as well as on the reflected wavefront and therefore forms part of a double-pass set-up. When imaging in reflection, odd aberrations are cancelled after the second reflection off the DMM, and even aberrations are doubled in terms of amplitude [14]. A screenshot of the user interface created with Labview software, which was written from scratch, to operate the confocal microscope is shown in Fig. 3.14.

Figure 3.14: Confocal Labview Software

3.2.1 System Aberration Correction

The DMM was calibrated with a WFS (WFS150, Thorlabs, sensitivity λ/50 rms wavefront deviation), using configuration c) in Fig. 3.16, in order to determine the required actuator settings for the first 15 Zernike modes, while taking into account the effects of unwanted aberrations due to oblique reflection off the DMM [80, 81]. The calibration routine used was closed-loop with the Malacara normalisation [19] for the Zernike modes. To determine the system aberrations, a mirror, placed on top of a coverslip, was brought in focus of the microscope objective. The reflected wavefront was measured by a wavefront sensor (configuration b) fig 3.13). The system aberrations could then be measured with the WFS and corrected for using the DMM. Subsequently, the WFS was replaced with the EMCCD camera (configuration a) fig 3.16) and the intensity of the pre-corrected focal spot at the EMCCD was further maximised by applying each Zernike mode in turn in order to optimise the central intensity of the focal spot. This additional optimisation further reduced the wavefront aberrations by about 0.08λ rms.

Images of spots with the related PSFs obtained before and after the system aberration correction can be seen in Fig. 3.18. This correction was used to remove the aberrations associated with the optical system in all further experiments. To measure the full width half maximum (FWHM) of the axial and lateral PSF a mirror was placed at the sample on top of a coverslip. For the axial PSF, the mirror was scanned through focus in 100nm steps and the intensity variation using a 3 pixel diameter pinhole mask was determined. The lateral PSF was taken from full images recorded with the EMCCD camera when the laser beam was focused on the mirror. The uncorrected/corrected measured lateral PSFs of the optical set-up (at best focus) are shown in Fig 3.15 and in Fig. 3.16 the lateral PSFs of Fig. 3.15c) were normalised to better show the effect of aberration correction on the FWHM.

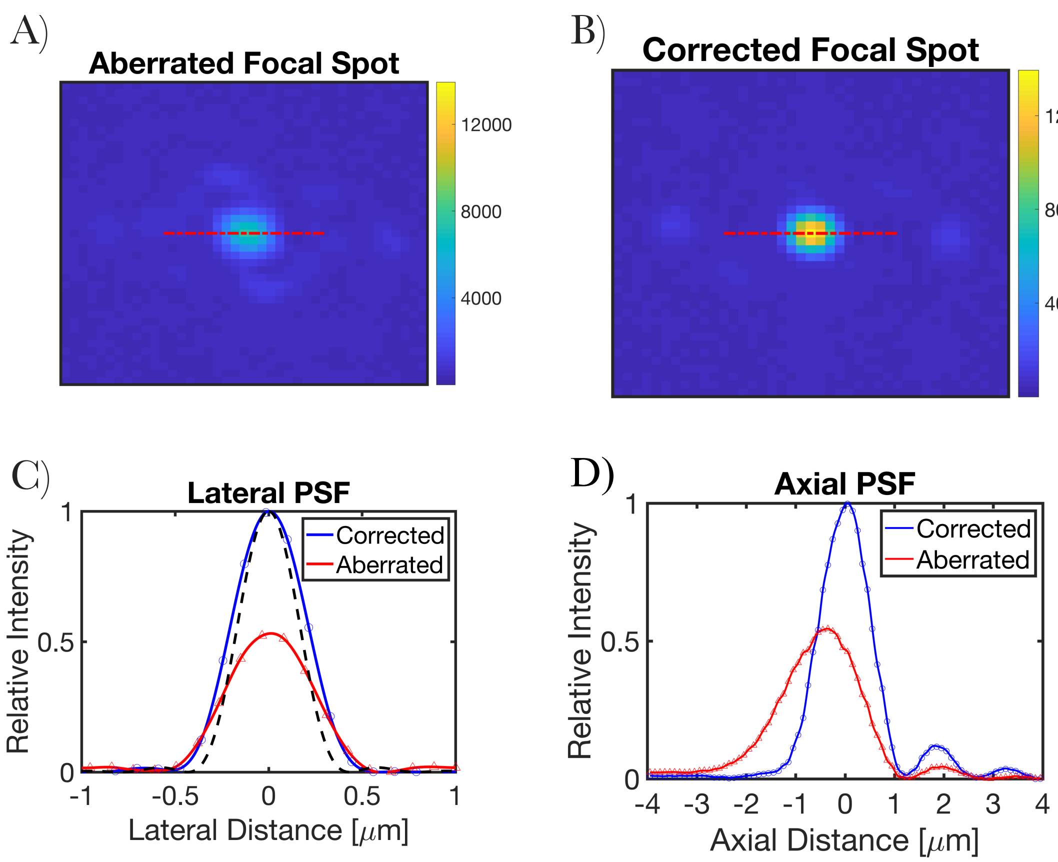


Figure 3.15: System aberration correction. A) an image of the laser beam focused on the mirror before aberration correction and B) an image of the laser beam focused on the mirror after system aberration correction . C) the measured lateral PSF with and without aberration correction, along the red dotted lines in A) and B) respectively, and D), the measured axial PSF with and without correction. For D), the axial PSF, a 3 pixel diameter pinhole mask was used, and the blue curve indicates the corrected PSF while the red curve indicates the uncorrected PSF. The black dashed curve in C), represents calculated Airy Disc cross section obtained with a 0.7NA objective at 0.532um wavelength. The length of the red dotted line is approximately 2.0μm.

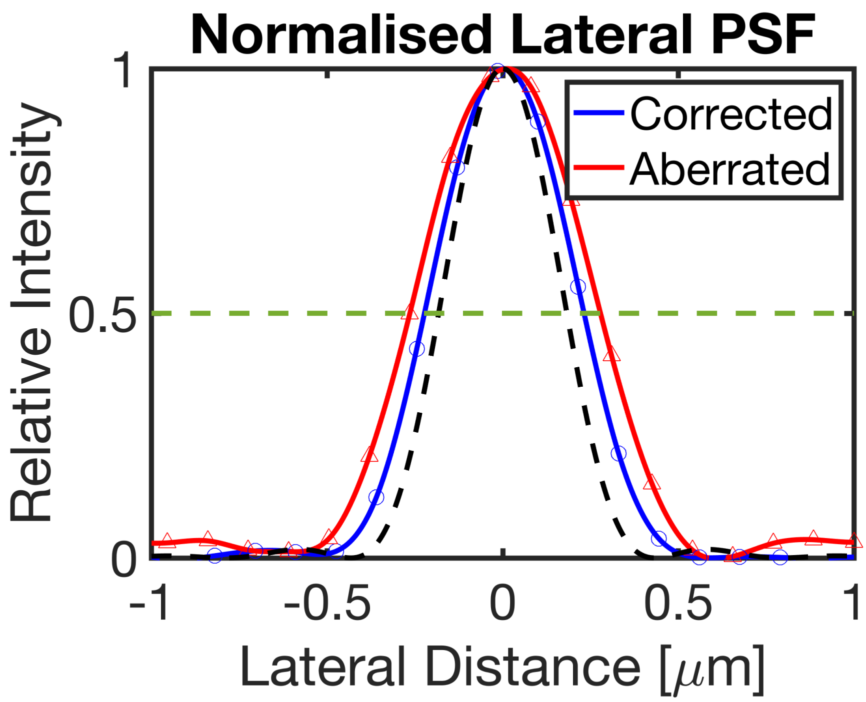


Figure 3.16: System aberration correction; the measured normalised lateral PSF with and without aberration correction, along the red dotted lines in fig 3.18 a) and b) respectively. The black dashed curve in c), represents calculated Airy Disc cross section obtained with a 0.7NA objective.

As can be seen in fig 3.15 and 3.16, applying the system aberration correction to the DMM has significantly improved the resolution of the optical system, increasing the maximum intensity by nearly a factor of 2 and reducing the width of the PSFs. The FWHM of the axial PSF has reduced from 1.81±0.03 µm without correction, to 1.21±0.03µm with correction, and the FWHM of the lateral PSF has reduced from 0.54±0.01 µm without correction, to 0.45±0.01 µm with correction, representing improvements of approximately 33% and 17% for the axial and lateral FWHMs respectively. The predicted axial and lateral FWHMs for a 0.75 NA objective lens would be 1.04 µm [22, Chapter 11, 82, page 2525-2544] and 0.37µm, respectively, in case of homogeneous pupil illumination. Therefore, the axial and lateral FWHMs obtained after aberration correction, are within 20% of the predicted theoretical values. The higher measured values can be explained, at least in part, by the fact that the pupil was illuminated with a Gaussian beam. The Gaussian beam was truncated to approximately 40% of its maximum intensity by the aperture of the objective lens. This would lower the effective NA of the objective lens from 0.75 to 0.7, in accordance with Marshall et al [83]. This fits better with what was measured. Satellite spots are visible around the focal spot in figs 3.18a and 3.18b, these diffraction spots can be explained by the periodicity in arrangement of the actuators behind the deformable membrane of the DMM [61]. The side lobes of the axial PSFs (fig 3.15d) are likely caused by second order spherical aberration. It has been shown that higher order spherical aberrations can give rise to asymmetrical sides lobes [84]. It is further noted that second order spherical aberration is less suitable for correction by the DMM because of the limited number of actuators. The intensity peaks of both axial PSFs in Fig. 3.15 do not coincide because the axial position of the mirror was re-adjusted after each scan. The FWHM of both axial and lateral PSFs of the microscope before and after aberration correction are listed in table 3.1. The relative intensity (with respect to the corrected PSF) is also given in table 3.1.

Table 3.1 - FWHM and relative intensity improvement of the microscope's

Sample	Lateral FWHM (µm)		Axial FWHM (µm)	
	Uncorrected	Corrected	Uncorrected	Corrected
PSF of the microscope system	0.54 ± 0.01	0.45 ± 0.01	1.81 ± 0.03	1.21 ± 0.03
Theoretical PSF for a 0.7 NA objective	-	0.39	-	1.15
Relative intensity (with respect to the corrected PSF)	~50%	100%	~50%	100%

3.2.2 Confocal images of microscope calibration target

To demonstrate the sectioning capability of the confocal microscope a small area of a reflective microscope calibration target (BCR photomask resolution standard from National Physical Laboratory, Row D calibration pattern: 50 μm long scale with opaque lines 2 μm apart) was scanned and imaged with the coverslip corrected water-immersion 0.75 NA objective of the confocal microscope. But, no coverslip was used. Aberrations were corrected using a modal correction (method described in detail in chapter 4) prior to scanning. The modal aberration correction was performed on a reflective area of the target. The target brought in focus and an area of 6 μm x 1μm was scanned. Thereafter the target was defocused by 0.5um and 1um (which correspond to about 40% and 80% of the FWHM of the axial PSF, respectively) and the same area was scanned again. Lateral scan steps are 50nm, which is below the Shannon criterion < 170 nm.

The maximum frequency transmitted by a lens is $2NA/\lambda$ (Abbe resolution limit). The Shannon sampling criterion, here defined for spatial frequencies, demands for the spacing Δx between adjacent scanning points that:

$$\frac{1}{\Delta x} \geq \frac{4NA}{\lambda} \tag{3. 23}$$

Thus:

$$\Delta x \leq \frac{\lambda}{4 \cdot NA} \tag{3.24}$$

For a NA of 0.75 and a wavelength of 532 nm we would have $\Delta x \leq 170nm$. On the Zeiss microscopy website [85]: “an interval 2.5 to 3 times the smallest resolvable feature is suggested”. To remove the baseline signal, the minimum measured intensity value was subtracted from each intensity value of the image. Typical exposure times per frames were 10ms. The confocal signal was obtained with a 5 pixel diameter pinhole mask (which roughly corresponds to a 0.6 AU pinhole). The in- and out of focus image are shown in Figure 3.17. Cross-sections perpendicular to the line pattern are shown at the bottom of the three images.

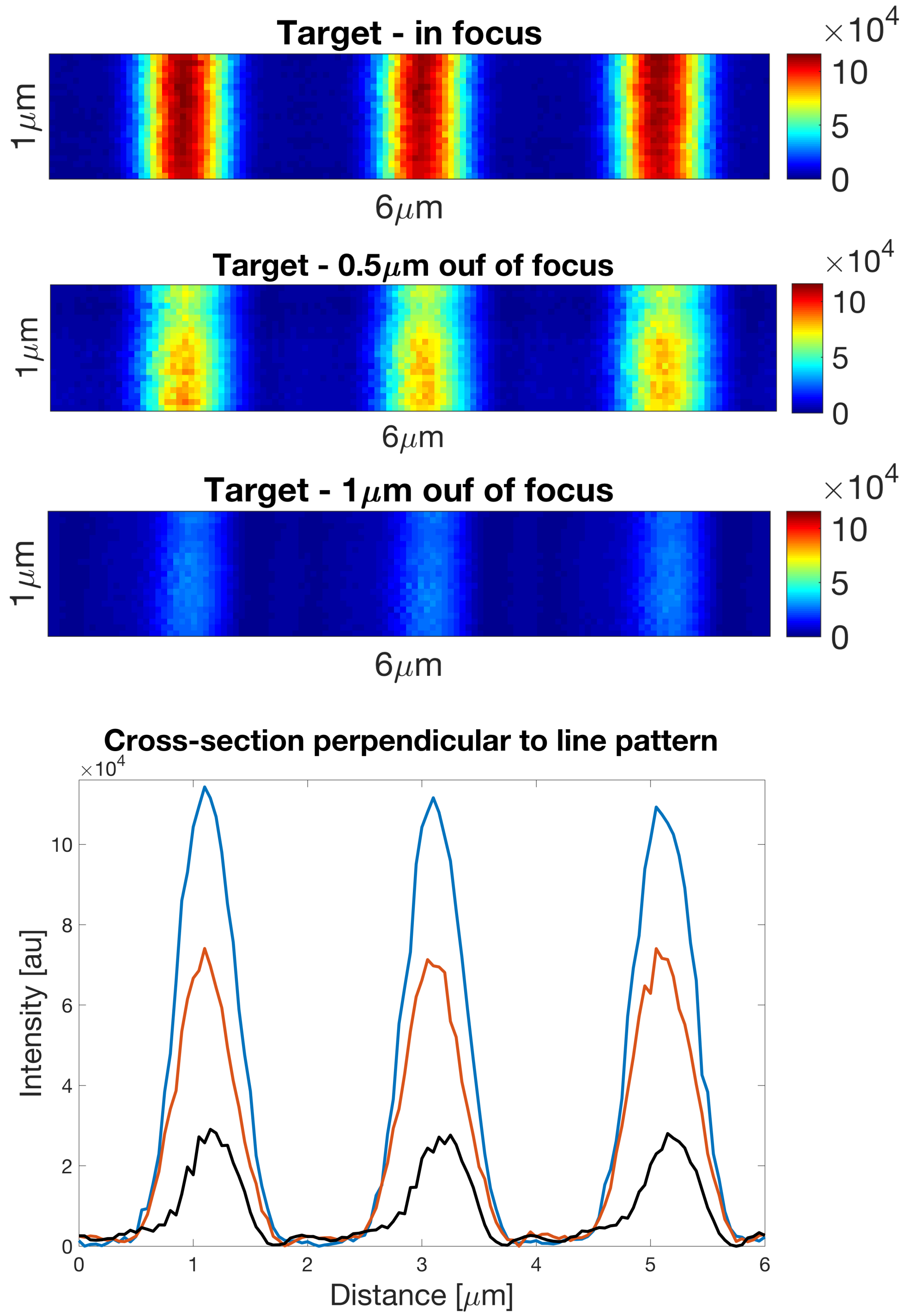


Figure 3.17: Confocal Image of a reflective line pattern taken in and out of focus (defocused by 0.5μm and 1μm, respectively). Images are scaled with respect to the in focus image. Below: Cross-section through the line pattern in focus (blue) and out of focus [0.5μm] (red) and [1μm] (black)

The confocal microscope shows a strong sectioning capability, as can be seen in Figure 3.20. The calibration target is poorly imaged, when defocused. A 1μm defocus reduces the maximum intensity, approximately, by a factor of 5.

3.3 Conclusion

In this chapter the design and construction of a confocal microscope with integrated AO were presented. The different microscope components were described and explained. Open- and closed-loop DMM calibration routines were compared and a closed-loop routine was chosen for accurately producing Zernike mode with the DMM. An unwanted source of aberration contamination caused by misalignment of the microscope objective, with respect to the DMM, was discussed and supported by a theoretical model. Confocal detection was implemented by using an EMCCD camera and selective pixel detection for enabling various pinhole size imaging. The spatial filtering, by the pinhole, was obtained by applying a pixel mask on camera images. Finally the system aberration correction procedure for the confocal microscope was explained and the lateral and axial PSF of the system, before and after aberration correction, were measured.

4. Modal Aberration Correction

In the previous chapter we discussed the design, construction and calibration of AO in an optical microscope with an EMCCD, that allows for software controlled (confocal) detection. Here we give a more detailed discussion on how to use a DMM for measuring and correcting wavefront distortions in a confocal microscope. In the so-called sensor-less modal wavefront sensing method [47] aberrations are measured indirectly by imposing sequentially different Zernike modes on the DMM, while monitoring the intensity variations in light passing through a pinhole. When the intensity signal at the pinhole is maximised, the wavefront distortion is corrected by the DMM. After introducing modal wavefront sensing, we'll take a closer look at the sensitivity of the method and explain the differences of confocal detection in reflection and in fluorescence. How cross-talk and local sub-maxima can affect modal wavefront sensing will be discussed thereafter. At the end, we present a new and effective modal optimisation strategy to correct large aberrations and demonstrate its use in confocal microscopy for diffraction-limited imaging deep into highly aberrating samples. This new approach uses ray tracing simulations to determine an initial pre-correction wavefront setting before using traditional modal correction routines. The result is a significant simplification of the iterative curve fitting procedure used previosuly to determine the magnitude of each Zernike mode present. As a result, the number of iteration steps required can be reduced by concentrating only on the relevant aberrations present, thereby reducing the speed of the overall correction process.

4.1 Modal aberration correction

For sufficiently small aberration amplitudes $\left(< 0.071\lambda\right)$ the optical system is considered diffraction limited (such a system is also referred to as satisfying the Maréchal criterium [73]). The normalised intensity I/I_0 at the diffraction limited focus is given by,

$$\frac{I}{I_0} \approx 1-\left(\frac{2\pi}{\lambda}\right)^2 * \left(\Delta\phi\right)^2 = 1-\frac{2\pi^2}{\lambda^2}\sum_{n=1}^{\infty}\sum_{m=0}^{n}\frac{\left(\alpha_n^m\right)^2}{n+1} \tag{4.1}$$

where I_0 is the peak intensity that would be obtained by an aberration free PSF, $\Delta\phi$ represents the variance of the wavefront at the aperture pupil with respect to a reference spherical wavefront, λ the wavelength, n the radial degree, m the azimuthal degree, and α_n^m the amplitudes of the Zernike coefficients [7] which describe the aberrated wavefront. I/I_0 is also referred to as the Strehl ratio. The above relations holds for Strehl ratios > 0.8. The corresponding Maréchal criterium, which sets a maximum to the wavefront aberration, is $\Delta\phi \le \lambda / 14$ (≤0.071λ rms). Optical systems that meet the Strehl criterium, and/or the Maréchal criterium are called diffraction limited. The Strehl ratio is proportional to the sum of the squared Zernike coefficients of each aberration. This highlights that the Zernike modes are an orthogonal set of functions that can be used to quantify the Strehl ratio for sufficiently small aberrations, and therefore, in theory, it is possible to correct for each aberration mode in turn on the basis of the Strehl ratio without being influenced by other aberrations. Such an optimisation approach with a pinhole, often referred to as a modal aberration correction, has been used to determine the wavefront aberrations present in terms of its Zernike modes [47]. Alternatively to modal aberration correction, a zonal optimisation approach can be chosen. Instead of applying Zernike modes, the DMM actuators are activated according to a specific optimisation algorithm [53] in order to maximise the intensity signal at the pinhole. Equation (4.1) holds for small aberrations. Imaging deep into samples often generates considerable amounts of aberrations which can be far beyond 0.071λ. For such large aberrations no simple equation describes the decrease of intensity and aberration cross-talk can become significant. In order to achieve diffraction limited imaging performance in highly aberrating samples, applying pre-corrections becomes essential to minimise aberration cross-talk and speed-up the optimisation process

When applying modal aberration correction, a small pinhole (< 1AU) in a confocal microscope can be used to measure the (maximum) spot intensity. Optimisation of the spot intensity (and the Strehl ratio) can be performed by applying defined amounts of each Zernike mode in turn using the DMM, so as to maximise the detected intensity. This technique has also been used in fluorescent confocal microscopy [48]. Facomprez et al. [49], showed that modal aberration correction can be performed on more strongly aberrated spots in a fluorescent two-photon microscope. The accuracy of the modal correction, as described by Facomprez, depends on the amount of wavefront aberration present, the signal to noise ratio and the number of intensity measurements per Zernike mode. For wavefront aberrations that satisfy the Maréchal criterion, at least 2N+1 intensity measurements (to do a simple parabolic fit through the data of each aberration) are needed per Zernike mode with N being the number of Zernike modes that are to be optimised. It follows from equation (4.1) that the influence of each Zernike mode on the Strehl ratio can be assessed separately. For strongly aberrated systems, the number of intensity measurements per Zernike mode needs to be increased. Facomprez et al. corrected 11 Zernike modes using up to 3 iterations, when the initial aberration was in the range of 0.09-0.3λ rms .

First the sensitivity of the confocal modal wavefront sensor to each Zernike mode was assessed with the confocal microscope described in chapter 3. The microscope's system aberrations were corrected before running the experiment (for more details see chapter 3, section 3.2.1). To measure the sensitivity of the confocal modal wavefront sensor, a mirror (with a drop of water between the cover slip and the mirror) was brought into focus and, then, the Zernike modes were sequentially applied, while measuring the changes in intensity at the pinhole. The results are shown in Fig. 4.1. Doing the measurement out of focus would reduce the sensitivity of the confocal modal wavefront sensor [48]. For each Zernike mode, the aberration amplitude was changed in steps of 0.019λ (which doubles for even aberrations after the second reflection on the DMM) and the intensity was measured with the EMCCD. Measurements were repeated five times for averaging. The CCD camera images were cropped so as to correspond to a 0.8 AU radius pinhole. In the set-up 1 AU corresponds to about 9 EMCCD camera pixels (see chapter 3, Figure 3.17). The intensity variation for a 0.8 AU pinhole are shown in Fig. 4.1 and will from now be referred to as Zernike sensitivity curves. Typical exposure times were 10-15 ms.

During a modal optimisation, the exposure time is kept constant. The EMCCD was cooled down to -50°C and no gain was applied. For each Zernike aberration sequence, a reference aberration-free image was taken. These aberration-free images can be used to estimate laser power fluctuations during a modal optimisation. For the aberration-free images, the intensity variations through a 1AU pinhole were less than 2%.

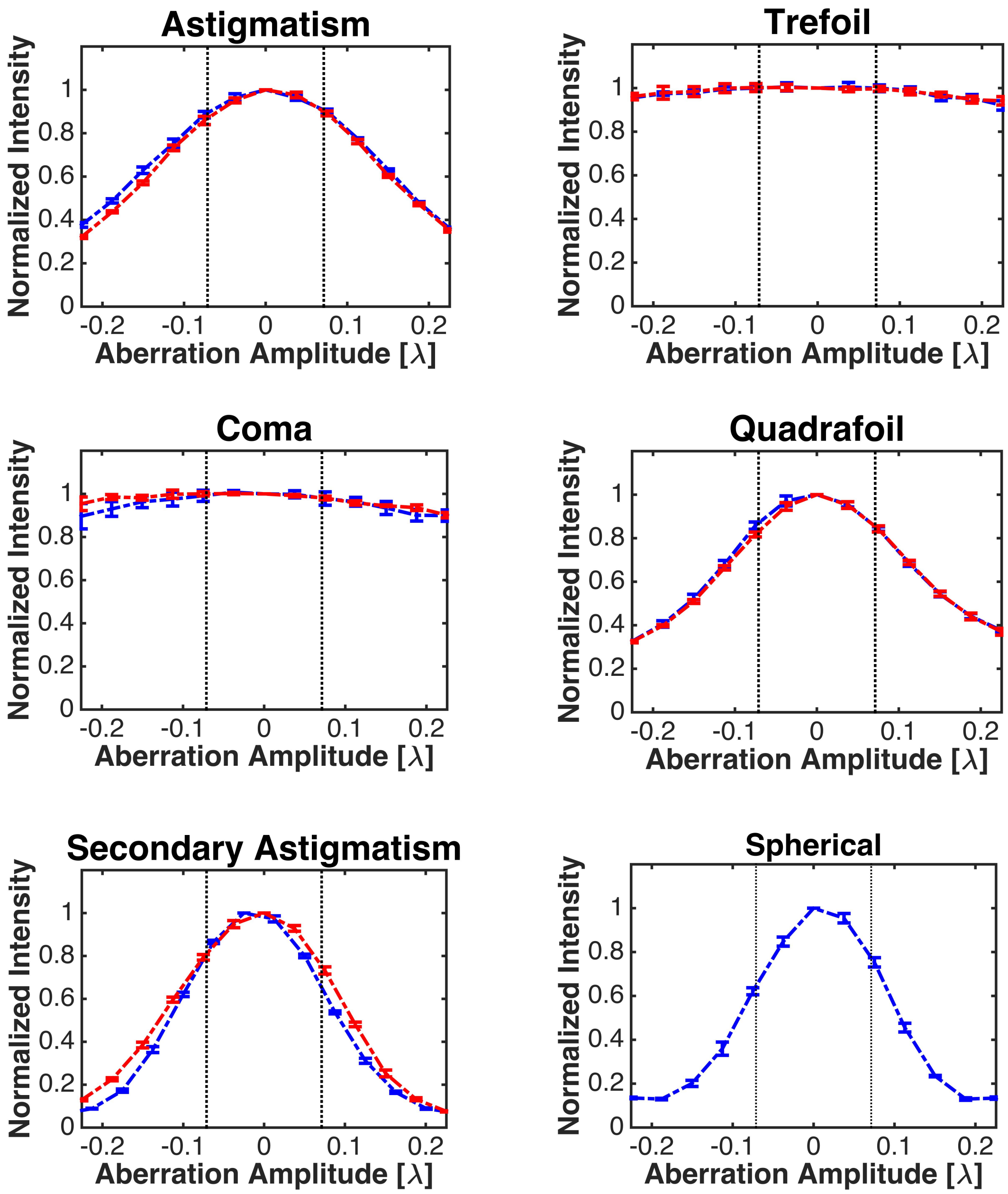


Figure 4.1: Measured Zernike sensitivity curves for a ~0.8 AU radius pinhole. The blue and red lines represent two orthogonal aberration orientations. The vertical dotted lines represent the Marechal criterion boundaries. Intensities are normalised with respect to the aberration amplitude at 0 λ.

As can be seen in Fig. 4.1, the intensity decrease is more dramatic for changes in spherical aberration and secondary astigmatism than for lower order astigmatism and quadrafoil. For spherical aberration and secondary astigmatism a 50% change in intensity is seen over the aberration amplitude range ±0.11 λ rms, compared to 40% intensity change for quadrafoil, 30% for astigmatism, and less than 10% for trefoil and coma. Odd aberrations, like coma and trefoil, show very little change in intensity over the full range of applied aberration amplitudes. When imaging in reflection, odd aberrations are cancelled after the second pass off the DMM, and even aberrations are doubled in terms of amplitude [14]. For even aberrations, and amplitudes up to ~0.071 λ rms, the curves presented in Figure 4.1 can be fitted around the maximum, for each aberration, assuming a parabola function of the form:

$$I = c_2\left(\alpha_n^m\right)^2 + c_1\alpha_n^m + c_0 \tag{4.2}$$

where α_n^m is a Zernike coefficient, c_i are fitting coefficients. In fitting equation 4.2 to data similar to that shown in fig 4.1, it is possible to determine the amplitude of each Zernike mode required to correct for the aberrations present in the sample.

In Fig. 4.2, different measured aberrated focal spots are shown, where for each Zernike aberration an amplitude of ~0.075λ (~0.15λ in double-pass) was applied by the DMM. The satellite spots around the aberration free focal spot can be explained as being caused by diffraction off the periodic actuator pattern of the DMM [61].

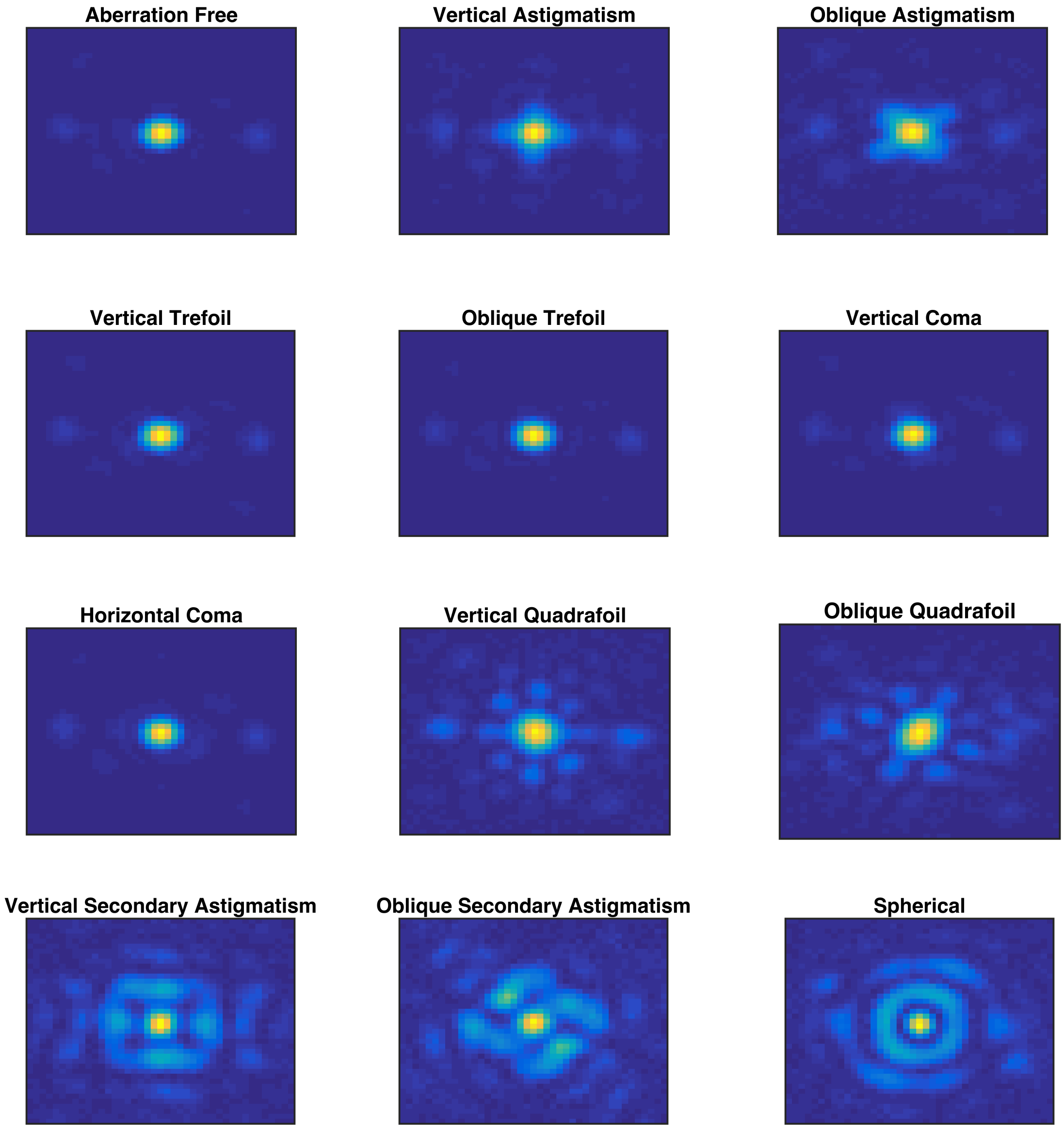


Figure 4.2: Different aberrated focal spots. For each aberration, an amplitude of 0.15λ was applied by the DMM

4.3 Modal Wavefront sensing sensitivity

Aberrations broaden the PSF and reduce the Strehl ratio. The Strehl equation (4.1) holds for infinitely small pinholes and for small aberration amplitudes. Larger pinholes improve the signal to noise ratio, are simpler to align and are also less affected by mechanical instabilities in the optical system. But for larger pinhole an increase of aberration amplitude leads to a smaller change in the detected intensity signal because some of the spread intensity is still passing through the larger pinhole. Thus, the pinhole size affects the sensitivity of modal wavefront sensing. Pinhole sizes up to 0.3 AU give responses comparable to those of an infinitely small pinhole [82] but for many applications pinhole sizes of 1 AU or larger are used. In our set-up by changing the diameter of the pinhole mask we can change the sensitivity of modal wavefront sensing. Various responses for different pinhole diameters are shown in Fig. 4.3 Only even aberration were considered. For primary astigmatism, quadrafoil and secondary astigmatism, solely the responses of the vertical component of each aberration were plotted.

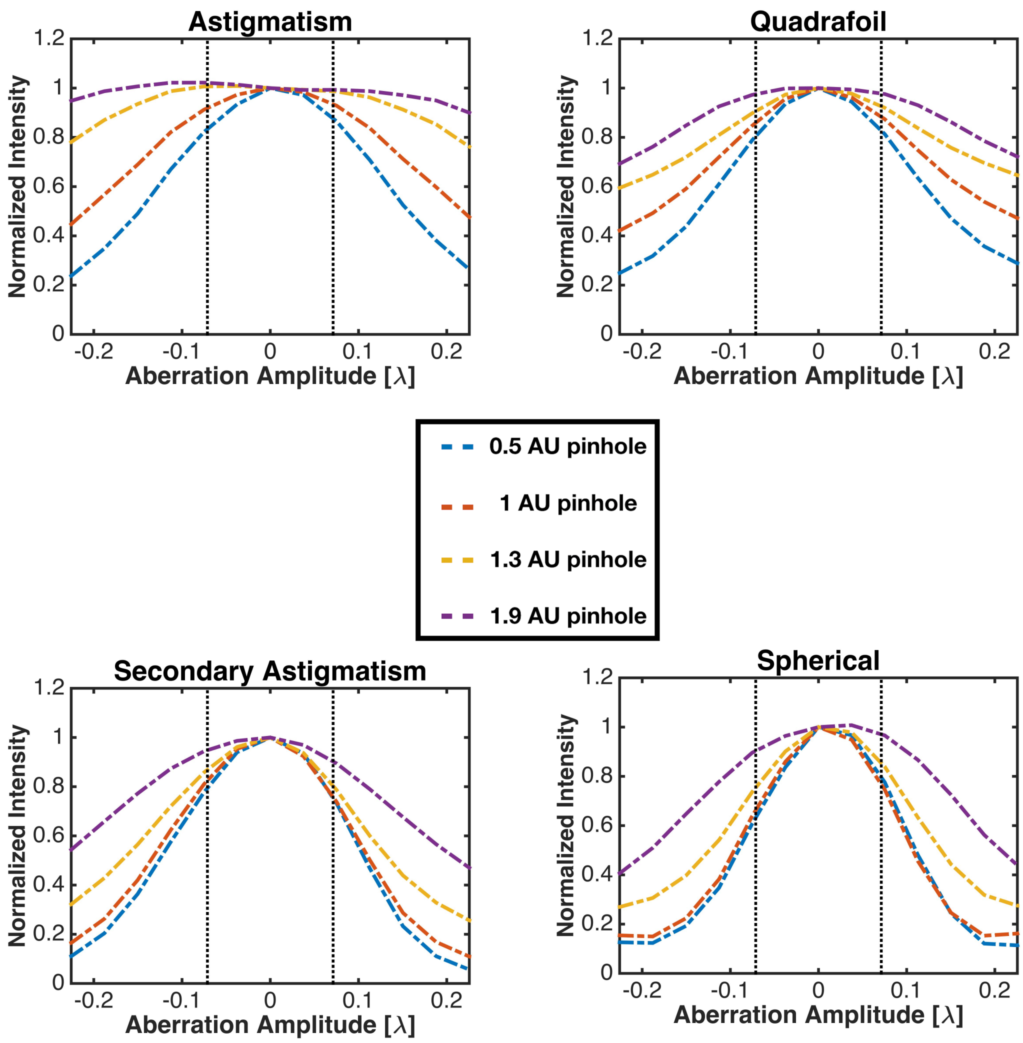


Figure 4.3: Measured Zernike sensitivity curves for astigmatism, spherical aberration, secondary astigmatism, quadrafoil with different pinhole mask sizes (0.5, 1, 1.3, 1.9) AU.

With increasing pinhole size, the sensitivity of modal wavefront sensing decreases. As can be seen from Fig. 4.3, for sensitive modal sensing pinhole sizes $\leq$ 1 AU (9 pixel pinhole) would be recommended for astigmatism and quadrafoil and ≤ 1.5 AU (13 pixel pinhole) for secondary astigmatism and spherical aberration. It is noted that higher order aberrations spread the intensity over larger areas than lower order aberrations. Consequently, modal wavefront sensing is more sensitive to higher order aberrations. For pinhole sizes of about 2 AU (17 pixel pinhole), astigmatism and quadrafoil become difficult to detect. The values for which a 50% intensity drop in intensity occurs for the different pinhole diameters and Zernike modes are listed in table 4.1.

Table 4.1 - 50% intensity drop at the pinhole for different pinhole diameters and Zernike modes

	Astigmatism	Quadrafoil	2nd Astigmatism	Spherical
0.5 AU pinhole	0.14λ	0.13λ	0.12λ	0.09λ
1 AU pinhole	0.2λ	0.2λ	0.13λ	0.09λ
1.4 AU pinhole	>0.22λ	>0.22λ	0.16λ	0.13λ
1.9 AU pinhole	>0.22λ	>0.22λ	>0.22λ	0.2λ

4.4 Comparison with theory

The performance of the modal wavefront sensor can be simulated using Fourier Optics as described in Chapter 2, section 2.1 (see also [16]). Zernike aberrations are applied to the pupil function. The pupil function is described in terms of Zernike aberrations and phase distribution over the pupil are calculated. Then a 2D Fourier transform is used to calculate the amplitude distribution over the image plane. The Fourier transform concerned 2048x2048 points and the pupil had a diameter of about 140 points while setting the remaining points to zero in order to effectively reduce the sampling interval at the image plane. Aberrated PSFs are computed by taking the squared modulus of the Fourier transform of the pupil function. A pinhole mask (see Fig. 3.15 in chapter 3) is placed on the PSF to evaluate the response of the modal wavefront sensor to various Zernike aberrations and aberration amplitudes. In Fig.

4.4, experimental and theoretical Zernike sensitivity curves for 0.8AU pinholes are compared.

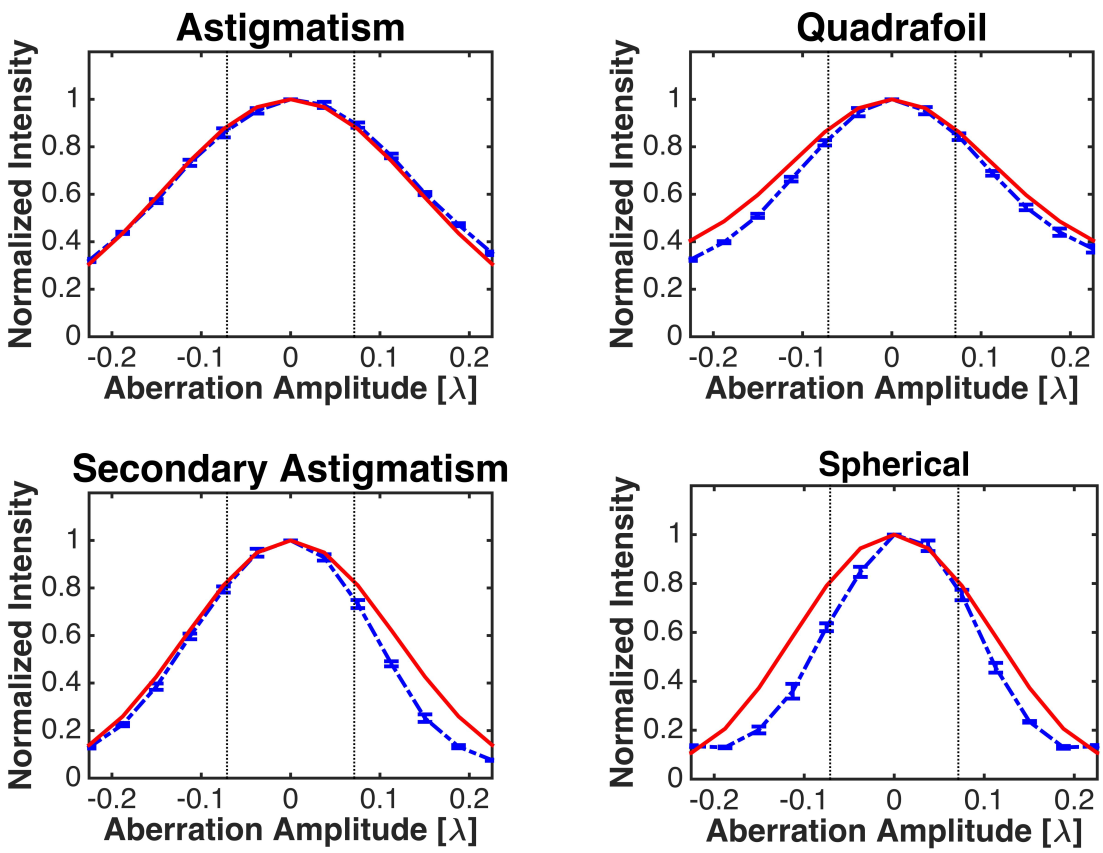


Figure 4.4: Measured (blue curve) and theoretical (red curve) Zernike sensitivity curves for astigmatism, spherical aberration, secondary astigmatism, quadrafoil for a 0.8AU pinole.

Fig. 4.4 shows that there is good agreement between experimental and theoretical data for lower order aberrations, especially for astigmatism. The experimental Zernike sensitivity curves for higher order aberrations lie below the theoretical curves. This discrepancy could be caused by the presence of unwanted aberrations. It is noted that a SH wavefront sensor, in double-pass, detected the presence of secondary spherical aberration when primary spherical aberration was applied with the DMM. Furthermore, generating higher order aberrations with a DMM having a limited number of actuators becomes more challenging.

4.5 Modal wavefront sensing in fluorescence

Fluorescent microscopy plays an important role in various biological and bio-medical research fields. Fluorophores can e.g. be used for visualising and studying processes happening at the sub-cellular level. Fluorescent imaging is used, amongst others, for protein detection, cell metabolism observation and for visualisation of ion or substance transport in microorganisms [86]. Fluorophores also allow for super-resolution microscopy. In STED microscopy [4] fluorophes are excited by using an excitation beam first and, then, a depletion beam is used to induce stimulated emission to force fluorescent molecules back in their ground state and solely detect fluorescent signal from within a volume smaller than the resolution limit given by diffraction. Other microscopy techniques such as PALM and STORM [6, 87] use both statistical photo-activation and localisation of fluorescent molecules to achieve super-resolution. Localisation can be done at higher precision than the diffraction limited resolution. Fluorescent emission occurs when ground state electrons of fluorescent substances are elevated to a higher energy level or excitation state. Relaxation of the excited molecule occurs partially through photon re-emission. This type of radiation, is referred to as fluorescent emission [88]. The fluorescence wavelength is larger than the excitation wavelength. The setup described in chapter 3 can easily be converted into a fluorescent confocal microscope by placing a fluorescent filter between the beamsplitter but before the lens focusing on the pinhole (Lens L4, Figure 3.16). In fluorescence, the microscope is not a double-pass system anymore because the fluorescent emitter acts as a new light source and the phase information of the laser beam is thus lost. With respect to modal sensing, the aberration, imposed by the DMM, will affect the intensity of both the excitation and fluorescent PSF. However, the fluorescent PSF on the EMCCD suffers only from the single-pass aberration. Fluorescent Beads (FluoSpheres™ Carboxylate-Modified Microspheres, 0.2 µm, orange fluorescent (540/560)) were diluted in water (~1/1e6). A drop was spread on a coverslip and left to dry. Some beads clustered. The laser was focussed on individual or smaller cluster of beads. The EMCCD gain (< 25) was turned on to ensure short exposure times (< 30ms). Modal optimisation was performed by sequentially applying the 11 Zernike modes with varying amplitudes. The intensity decrease at the pinhole (0.8 AU pinhole) for coma and trefoil, for five different beads are shown in Fig. 4.5.

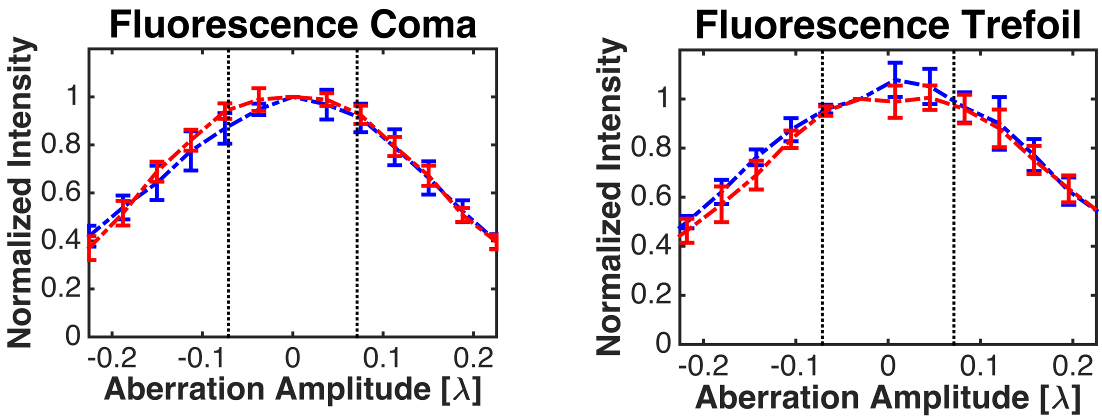


Figure 4.5: Measured Zernike sensitivity curves for a ~0.8 AU radius pinhole in fluorescence. The blue and red lines represent two different aberration orientations. The vertical dotted lines represent the Marechal criterion. Intensities are normalised with respect to the aberration amplitude at 0 λ rms.

The measured Zernike sensitivity curves are comparable to the curves obtained in reflection (see Fig. 4.1). Modal sensing is sensitive to odd aberrations in fluorescence. Other experiments on larger fluorescent beads and clusters (data not presented) have shown that modal sensitivity decreases with increasing size of the fluorescent emitter. The sensitivity depends on the size and geometry of the fluorescent emitter [89]. Peripheral regions of large fluorescent objects may emit light, when out of focus, thereby acting as stray light that may be detected by the pinhole, thereby decreasing the sensitivity of the confocal modal wavefront sensor.

4.6 Cross-talk in modal wavefront sensing

In modal wavefront sensing, crosstalk will be used here to refer to an undesired Zernike mode affecting the confocal signal from another Zernike mode. Aberration cross-talk affects modal wavefront sensing and can lead to inaccurate or erroneous measurements. While working on modal wavefront sensing, two different types of aberration cross-talk were identified. The first type type of cross-talk is caused by misalignment of the objective's pupil with respect to the DMM. This issue with misalignment was already described in chapter 3 section 3.1.3. Most important in the end is that the the confocal signal of an aberrated focal spot might not only be increased by correcting the relevant aberrations but also other Zernike modes can increase the intensity signal passing through the pinhole. In the presence of large aberrations (>> 0.071λ rms), cross-talk related to the diffracted intensity will occur and affects modal wavefront sensing. This has been investigated by Neil et al. [90].

4.7 Modal correction of large aberrations using a ray-tracing pre-correction

In the presence of large aberrations (> 0.07λ rms) modal wavefront sensing becomes challenging because of increased aberration cross talk and the possible presence of local sub-maxima. Previously, several modal iterations, involving at least five aberrations amplitudes per Zernike mode, were necessary to correct large aberrations [49]. Here, it is proposed to use a pre-correction as a starting point to efficiently correct large aberrations. The method proposed here uses a pre-correction determined from ray tracing simulations of a sample to simplify the modal correction process, with the aim of speeding up the correction process, reducing the number of interactions required, and mitigating problems associated with cross-talk. For more details about the way ray tracing was implement, reference is made to appendix A. The samples are simulated as geometrical objects (plates and cylinders, in this study) between a point source and the entrance pupil of an objective. For the cylinder case, it was assumed that the cylinder axis is perpendicular the optical axis of the objective. The point source was at a small distance (≈ 1µm) behind the cylinder, and along to the optical axis (< 20µm). Refractive indices of the materials simulated can be found in literature. The radius of curvature can be computed by estimating the object's typical dimensions and shape. Spatial variations of Zernike aberrations can be calculated by either changing the position of the sample or point source depending on sample or laser scanning mode. For predicting the sample induced aberrations, the numerical aperture was set to 0.7 (see chapter 3 section 3.2.1). For each simulation, it was ensured that the ray fan filled the objective pupil, by altering the angles of the ray cone accordingly, if necessary. The program gave a warning when total internal reflection would occur. In optical design, ray-tracing is used to determine aberrations such as those associated with an optical system. For sample induced aberrations, as long as the object is not too complex in terms of geometry and refractive index distribution, the same approach can be used, provided an approximated model of the object can be generated, the benefits of the approach, as stated above, can still be realised.

4.8 Results

In the following section, it will be shown that the results from ray tracing simulations can be effectively used as pre-corrections in the modal correction process. The method was tested for samples introducing spherical aberration and astigmatism. The three different test samples were: a glass coverslip (thickness: 0.13-0.16mm), a 125um diameter coreless termination fibre (FG125LA Thorlabs) and an air gap of ~34um between the coverslip and a mirror. These samples were chosen because they are either commonly used in microscopy (coverslip), or because they resemble objects which produce large amounts of spherical aberration (air gap) and astigmatism (fibre). The fibre could, for example, mimic the aberrations encountered when imaging deeper into samples such as a root or a worm. For ray tracing, samples were modelled as either layers (coverslip, air gap) or a cylinder (fibre). The three tested samples are shown in Fig. 4.5 top row. The first sample we consider is a coverslip (refractive index 1.52, thickness 0.145mm). By removing the coverslip and focussing directly on a mirror spherical aberration is generated because the microscope objective is corrected for imaging through a coverslip. Thus, the measured aberration will correspond to aberrations that would have been generated by a coverslip, but are of opposite sign. As a second sample, a fibre (refractive index 1.46) was immersed in a ≥99.5% percent glycerol solution (refractive index 1.47 at room temperature [91]). By immersing the fibre in glycerol, the astigmatism was below 0.15λ rms. The laser beam was focussed on the glycerol-mirror interface at the top of the fibre. Modal correction, involving ray-tracing pre-corrections, was performed for five positions behind the fibre and along the direction of the fibre axis (over a range of ±10μm), thereby avoiding total internal reflection. The third sample concerned an air gap. The fibre and glycerol were removed and two 34μm plastic shims were used as spacers between a coverslip and a mirror. The laser beam was focussed through the immersion water, the coverslip and the air gap, on to the mirror. For each sample, prior to running the modal optimisation, a pre-correction which was found by ray tracing simulations was applied to the DMM. The ray-tracing simulations were carried out with the refractive indices and dimension mentioned in the previous paragraph. Once the pre-correction had been applied, the focal spot was optimised using modal correction with a single optimisation. This process was repeated for five different positions for each sample (same z-plane but different (x,y) locations). Parabola functions were fitted through the Zernike sensitivity data (equation 4.1) because the residual aberrations, after pre-correction, were in all three cases $< 0.07\lambda$ rms. The parabolas were fitted through 5 intensity readings per Zernike mode. Comparable fitting results were obtained by using only 3 intensity data points. Figure 4.5 presents the corrected/uncorrected lateral and axial PSFs as well as images recorded on the EMCCD camera of the focal spots before and after correction.

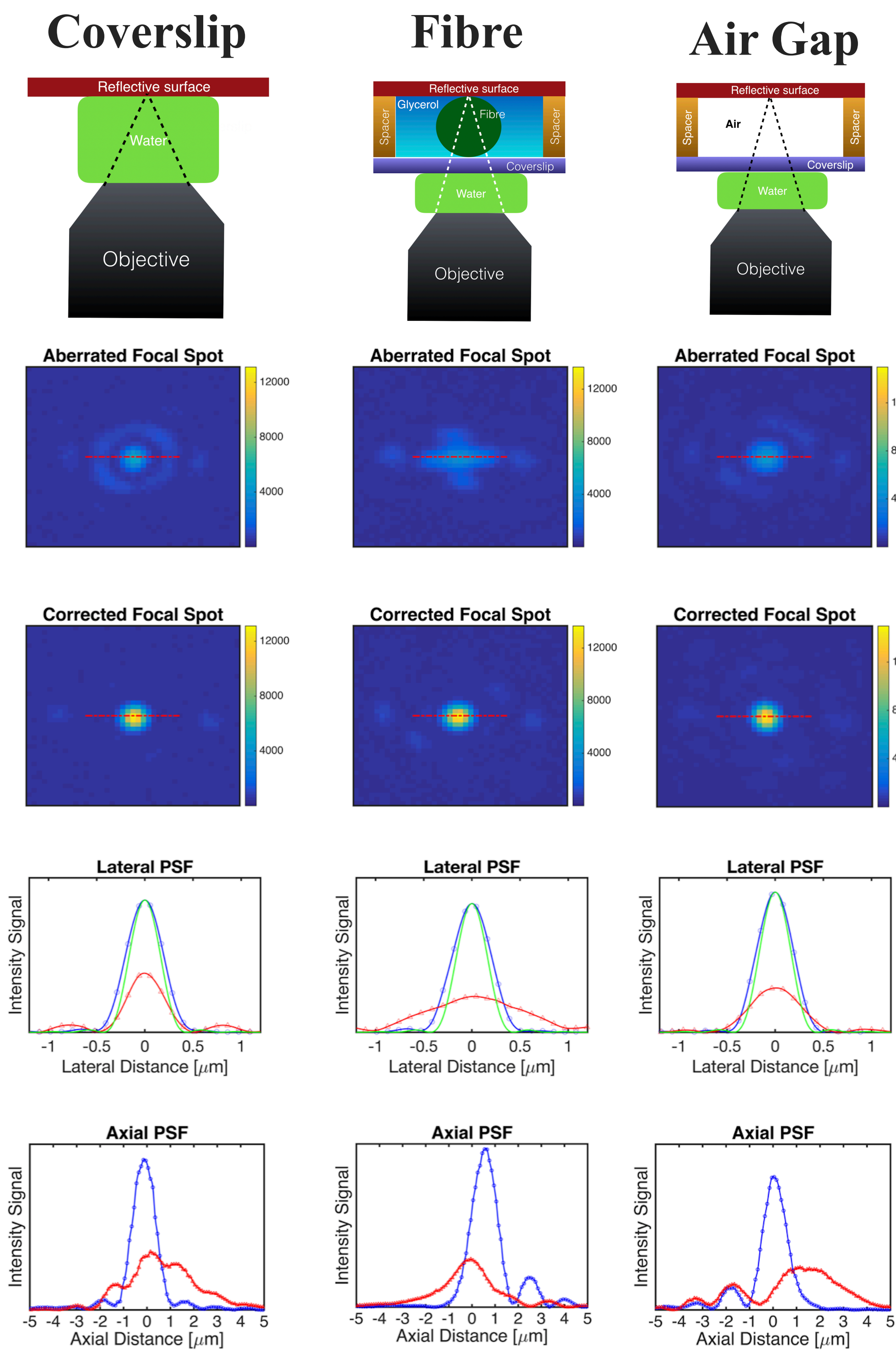


Figure 4.5: Modal aberration correction using a pre-determined starting point: First column: Direct reflection on the mirror; second column: coreless termination fibre; third column: air gap. First row: schematic of the sample. Second row: Uncorrected, aberrated, focal spots; Third row: corrected focal spot (after using a ray tracing pre-correction and 1 iteration of a modal aberration correction routine); Fourth row: Lateral PSFs (Corrected, Uncorrected, Airy Disc); Fifth row: Axial PSFs (Corrected/Uncorrected). For the axial PSF, a 3x3 pixel pinhole mask was used. The blue curves indicate the corrected spot and the red curves indicate the aberrated spot. Green curve represents calculated Airy Disc cross section for a 0.75 NA objective at 0.532µm. The red dotted line is ~ 2.0µm long.

To obtain the axial PSF, the reflective surface was moved through focus with the piezo-stage in steps of 100nm. The lateral PSFs were obtained from the camera image with highest intensity (see Fig.4.5).

For all three samples, the maximum intensity of the focal spot was increased by up to a factor of three after correction and the axial and lateral FWHM after aberration correction were comparable to the FWHM of the system corrected PSF (Figure 4.5 and table 4.2). For all samples, the Zernike coefficient values of the ray-tracing pre-correction were within 0.03λ rms (i.e. 0.06λ rms in double pass, which lies in the Marechal criterion range, see Fig. 3) of the values found after modal correction was implemented. Odd aberrations (coma and trefoil) did not significantly affect the measured intensity due to detecting reflected light in a double pass DMM configuration. After the pre-correction had been applied, the Zernike sensitivity curves were similar to those shown in Fig. 4.1. To avoid aberrations which do not have a meaningful maximum within the scan range, such as coma and trefoil (see Fig. 4.1), a minimum requirement was imposed on the coefficient (see equation 4.1), such that an intensity difference of at least 10%, over a range of 0.075λ rms was required.

Table 4.2. Lateral and Axial FWHM of the PSF for the 3 samples tested.

Sample	Lateral FWHM (μm)		Axial FWHM (μm)	
	Uncorrected	Corrected	Uncorrected	Corrected
PSF of the microscope system	0.54 ± 0.01	0.45 ± 0.01	1.81 ± 0.03	1.21 ± 0.03
Coverslip	0.42 ± 0.01	0.43 ± 0.01	2.58 ± 0.03	1.20 ± 0.03
Fibre	1.22 ± 0.01	0.48 ± 0.01	1.91 ± 0.03	1.27 ± 0.03
Air Gap	0.60 ± 0.01	0.44 ± 0.01	3.04 ± 0.03	1.22 ± 0.03
Theoretical PSF for a 0.75 NA objective	-	0.37	-	1.04

The main side lobes for the fibre and air gap sample of the corrected axial PSFs (see Fig. 4.5) is likely to be due to second order spherical aberration as predicted by ray tracing for the air gap sample. Second order spherical aberration is less suitable for correction by the DMM because of the limited number of actuators, making it hard to reproduce independently of other higher order Zernike modes.

Table 2 compares the ray tracing results used as a pre-correction with the final wavefront after modal correction. The difference in wavefronts was determined as follow:

$$\Psi_{diff} = \Psi_{raytracing} - \Psi_{\mathrm{modal}} \tag{4.3}$$

Subsequently the RMS wavefront error of Ψ_{diff} was calculated using:

$$\Delta\Psi_{diff} = \sqrt{\overline{\Psi^2_{diff}} - \overline{\Psi_{diff}}^2} = \sqrt{\sum_{n=1}^{\infty}\sum_{m=0}^{n}\left(\alpha_n^m\right)^2} \tag{4.4}$$

where α_n^m are the Zernike aberration coefficients, expressed using the Malacara notation. For the Zernike polynomials, reference is made to Appendix A. Only values above 0.01 λ rms are listed in table 4.3.

Table 4.3. Comparison of ray tracing pre-correction with the final correction after applying the pre-correction and a single iteration modal correction for the different samples

Sample	Aberration	Sample Aberrations present (λ rms)	Raytracing prediction (λ rms)	ΔΨ (λ rms)
Coverslip	Residual Aberrations	-	-	0.018
	Spherical	0.053	0.071	
Fibre	Residual Aberrations	-	-	0.018
	Astigmatism	0.120	0.132	
	Spherical	0.043	0.056	
Air Gap	Residual Aberrations	0.015	-	0.034
	Spherical	0.083	0.113	

The astigmatism coefficient was calculated using $Astigmatism = \sqrt{\left(\alpha_2^2\right)^2 + \left(\alpha_2^{-2}\right)^2}$. As can been seen from table 2, the ray tracing pre-correction alone corrects for >70% of the sample aberrations present and the residual aberrations lie well within the Maréchal range. The result of this is a greatly simplified Zernike sensitivity graph that means any remaining aberrations can be corrected for with a single iteration.

Figure 4.6 shows the effect of applying ray-tracing pre-corrections on the Zernike sensitivity curves. The Zernike sensitivity curves presented are those of the most relevant aberrations present, with pre-correction (blue curves) and without applying a pre-correction (red curves). The full modal aberration correction was repeated several times for each sample, the error bars represent the standard deviation of these measurements. Figure 4.6 presents data for the coverslip and the fibre sample. All curves are normalised with respect to the confocal signal measured when no aberrations are imposed. No data is presented for the air gap sample because this sample contains a large amount of spherical aberration. When large amounts of spherical aberrations are present (larger than 0.07 λ rms), changing the amount of spherical aberration alters the position of best focus (best focus is defined here as the axial plane where the focal spot has the highest maximum intensity). To account for this, the axial position of the reflective surface would have to be altered so that it is always located at best focus.

Figure 4.6 a) and d) show that after applying the pre-correction, the maxima of spherical and vertical astigmatism are shifted to within the Marechal criterion. Values within the Maréchal range allow for correction of that particular aberration using a simple parabolic fit through the data with a single iteration step. It is interesting to consider the impact of adding second order astigmatism on a wavefront already having large amounts of primary astigmatism. Looking at Fig. 4.6e (red curve), it is clear that, without pre-correction, adding negative second order astigmatism increases the confocal signal of the strongly (first order) astigmatic focal spot. Referring to the blue curve in Fig. 4.6e, recorded after pre-correction has been applied, one can see that the maximum is now close to zero aberration amplitude indicating that there is no-longer any second order astigmatism present. In this case it is clear that the second order astigmatism is an artefact of a strongly aberrated initial spot, and that the cross-talk between the first and second order astigmatism cannot be neglected when one of the aberrations is outside the Marechal criterion aberration range. It is further noted that the signal increases for odd aberrations when applying larger aberration amplitudes with the DMM (Fig.4.6 b) and f)). The Zernike sensitivity curve for coma is close to constant when a pre-correction is applied, as expected for a reflection double-pass AO system.

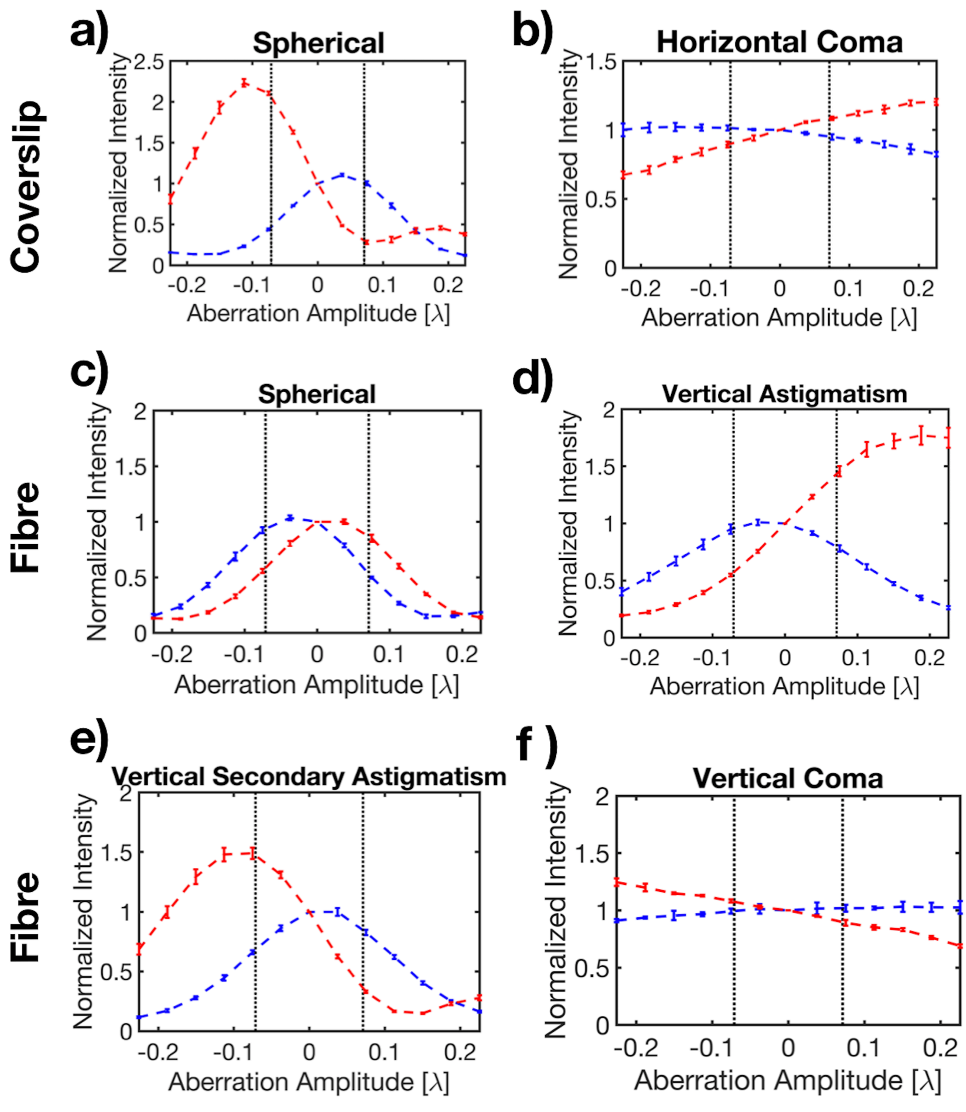


Figure 4.6: Zernike sensitivity curves with and without pre-corrections. The vertical dotted lines indicate the Marechal criterion limits. First row: coverslip aberrations; Second and third row: Fibre aberrations; blue curves: with pre-correction; red dashed curves: without pre-correction. Intensities are normalised with respect to the aberration amplitude at 0 λ rms.

4.9 Image degradation by aberrations

To show the effects of aberrations on an image taken with a reflection confocal microscope, the image formation process was simulated using coherent imaging theory [21]. A camera picture of the Robin Hood statue in Nottingham (1032 x 774 pixels taken with an iPhone 6 camera) was chosen as the object to be imaged by the reflection confocal microscope. To simulate how the Robin Hood picture would be imaged by a reflection confocal microscope, the amplitude of the Robin Hood picture (square root of iPhone camera intensity picture) was convolved with the microscope's PSF in accordance with [21]. Robin Hood's image was convolved once with an aberrated PSF and then with an aberration free PSF to show the improvement obtained by correcting the aberrations with an AO system. An aberrated PSF was generated using the measured fibre aberrations (see table 4.3, the measured aberration amplitudes were squared to take into account the double-pass effect; oblique astigmatism was chosen for better visualisation of distortions on the mostly horizontal and vertical features visible on the Robin Hood picture). The PSFs were calculated using Fourier optics [16], as described in section 2.1 and the Airy disk diameter was set to 6 pixels instead of 36 pixels (the 36 pixel setting was used for other simulations in this thesis). The PSF was changed from 36 pixels to 6 pixels so as to better fit with the picture size (1032 x 772 pixels). The scanned image at the detector was calculated using the following formula [21]:

$$I_{Object/Coherent}(r,\phi,z) = \left|h(r,\phi,z) \otimes t(r,\phi,z)\right|^2 \tag{4.5}$$

where the t represents the object to be imaged and which in this case given by the square root of the Robin Hood camera intensity picture. h is here the confocal microscope PSF (it is noted the confocal PSF is given by the product of illumination and detection PSF, which for a reflection confocal microscope equals the squared widefield PSF). The simulated images are shown in Fig. 4.7.

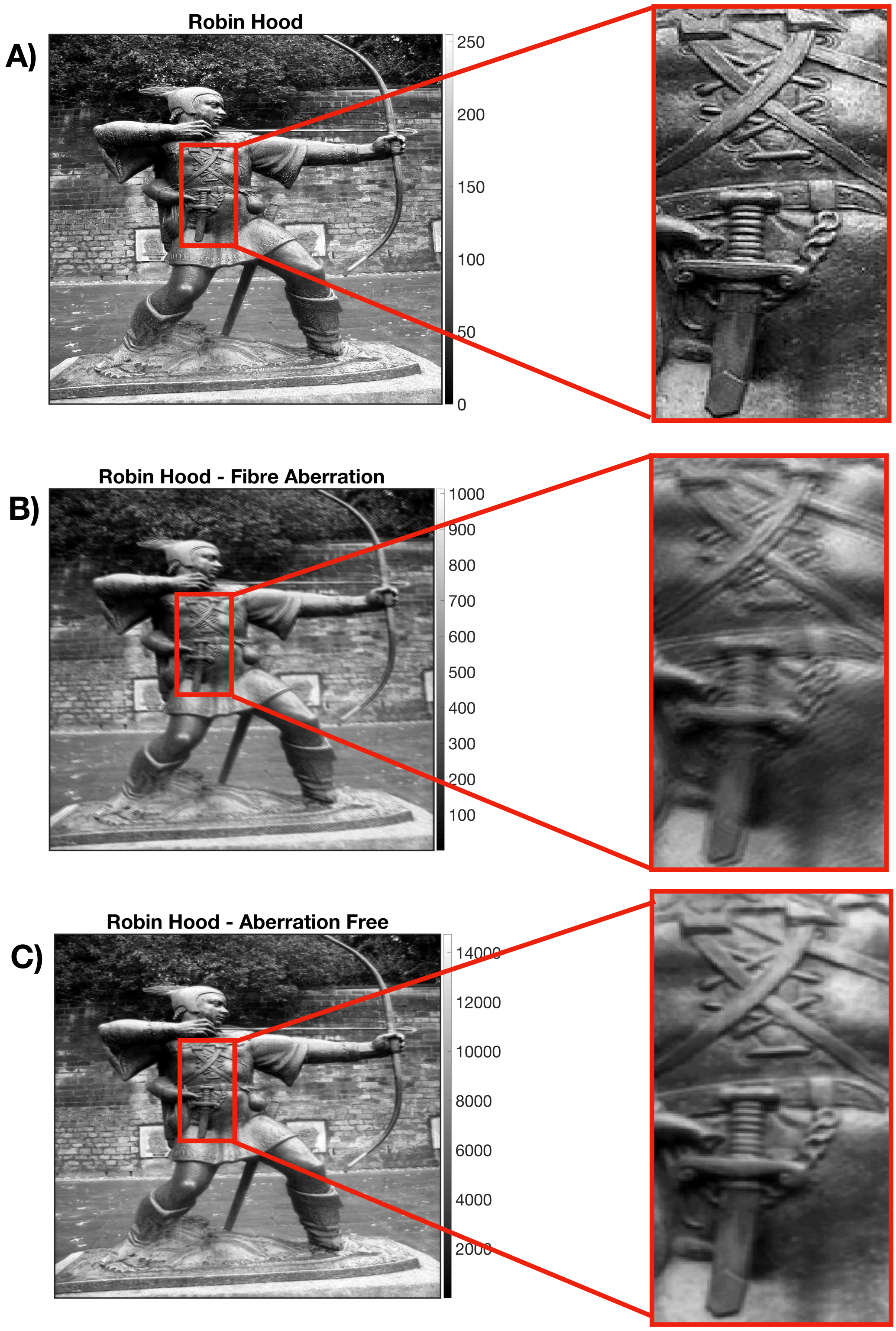


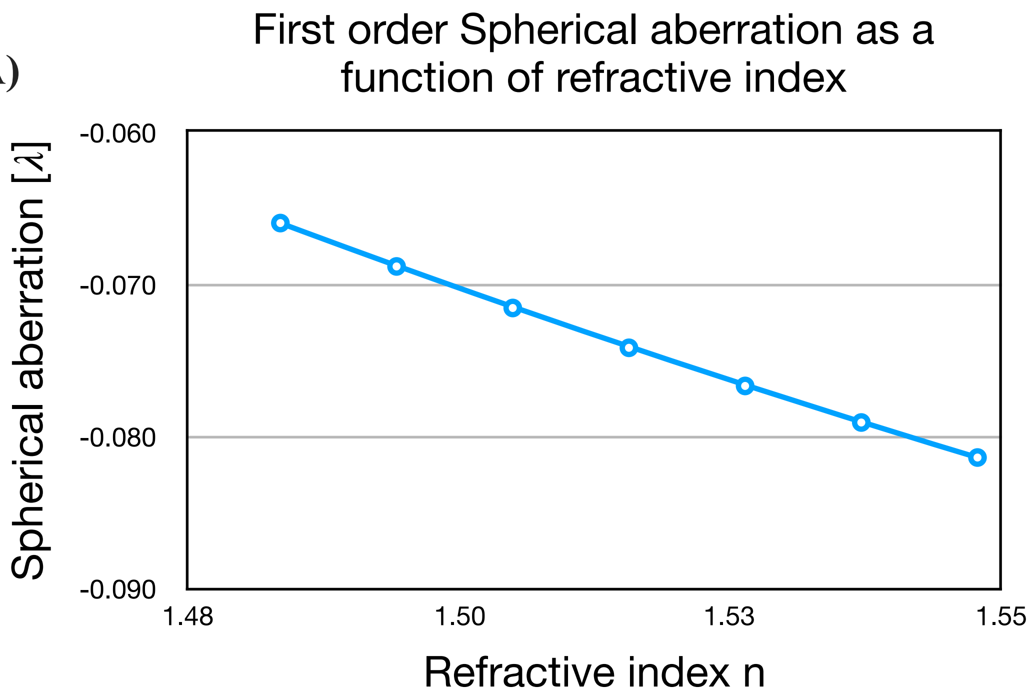
Figure 4.7: Image formation simulation of a confocal reflection microscope. A) iPhone 6 camera picture of the Robin statue in Nottingham used for the simulation. B) Aberrated image of Robin Hood as imaged by a reflection confocal microscope in the presence of the measured fibre aberrations. C) Image of Robin Hood that would be obtained with an aberration free confocal reflection microscope.

The Robin Hood camera picture taken with the iPhone 6 is shown in Fig. 4.7A. The aberrated image obtained from the Robin Hood picture (Fig. 4.7B) displays strong distortions and fine details are washed out as can be seen in the zoomed image of Robin Hood's shirt and belt. The metal chain, the dagger handle and the laces on Robin Hood's shirt are blurry and fine details can not be resolved. After correction of the fibre aberrations, the image obtained with an aberration free confocal PSF (Fig. 4.7C) is clearer and sharper. The aberration free image (Fig. 4.7C) is not as sharp as the original picture iPhone 6 picture (Fig. 4.7A), because the aberration free PSF used for the convolution slightly blurs the image, i.e. the diameter of the diffraction-limited Airy disk extends over 6 pixels.

4.10 Ray-tracing pre-correction accuracy

The success of the "Ray-tracing pre-correction" approach will depend how accurately one can estimate the typical aberrations of a sample using ray-tracing. For non-biological samples, the geometry and optical properties of samples are often roughly known. If the result of the pre-correction lies within $\approx 0.07\lambda$ rms of the best achievable correction, aberration cross-talk and local sub-maxima are negligible and sample aberrations can be determined by simple parabola fits through the data obtained after pre-correction. A pre-correction which is less accurate might require additional iterations and more than five sample points per aberration for finding the best correction. In any case, a pre-correction would still be beneficial, if it corrected for a significant amount of the initial aberrations present. The sensitivity of aberrations to ray-tracing parameters was estimated by simulating samples with slightly varying refractive indices and geometries (curvature, thickness, …). For the coverslip, I varied, independently from each other, the refractive index by 1.52 ±0.03 (~2%) and the thickness by up to ± 30% (~ 40μm). The variations in spherical aberrations for those varying parameters are shown in Fig. 4.8.

A)



B)

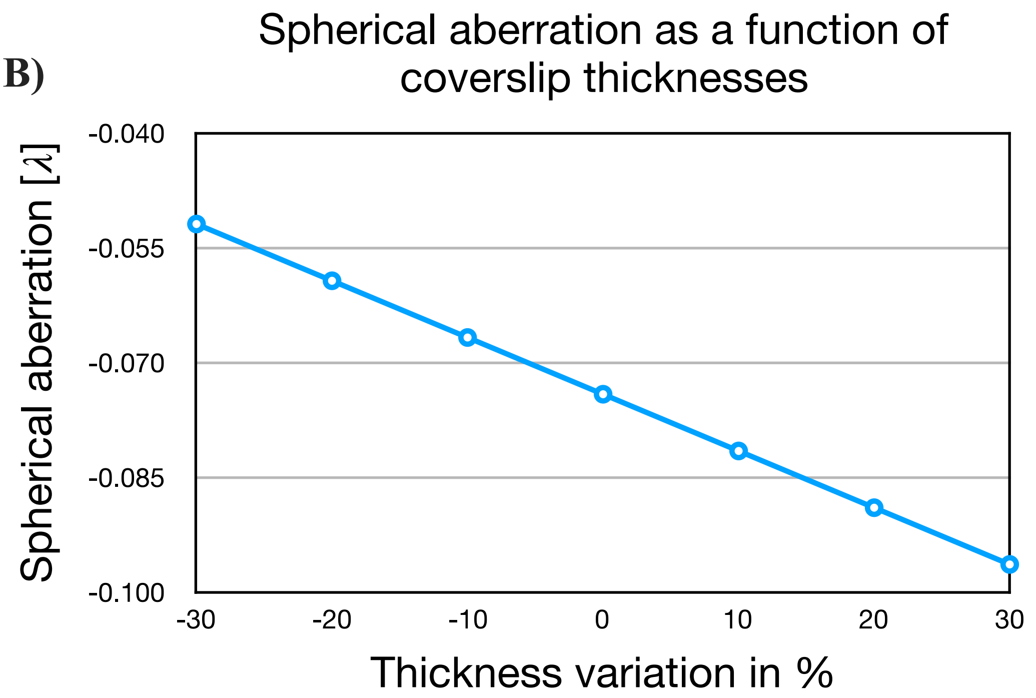
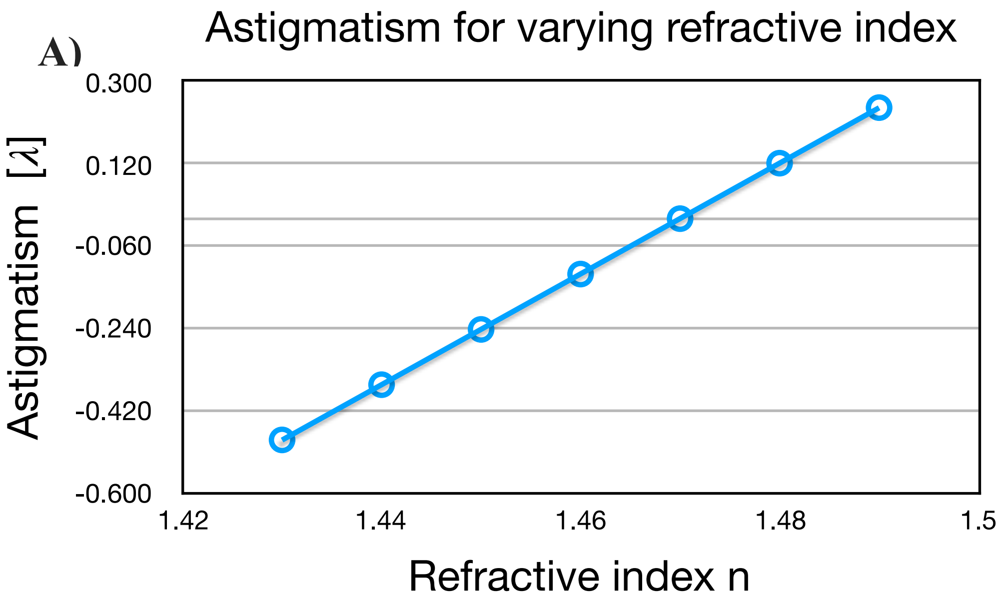


Figure 4.8: Variations of spherical aberration for A) varying refractive index and B) varying coverslip thicknesses

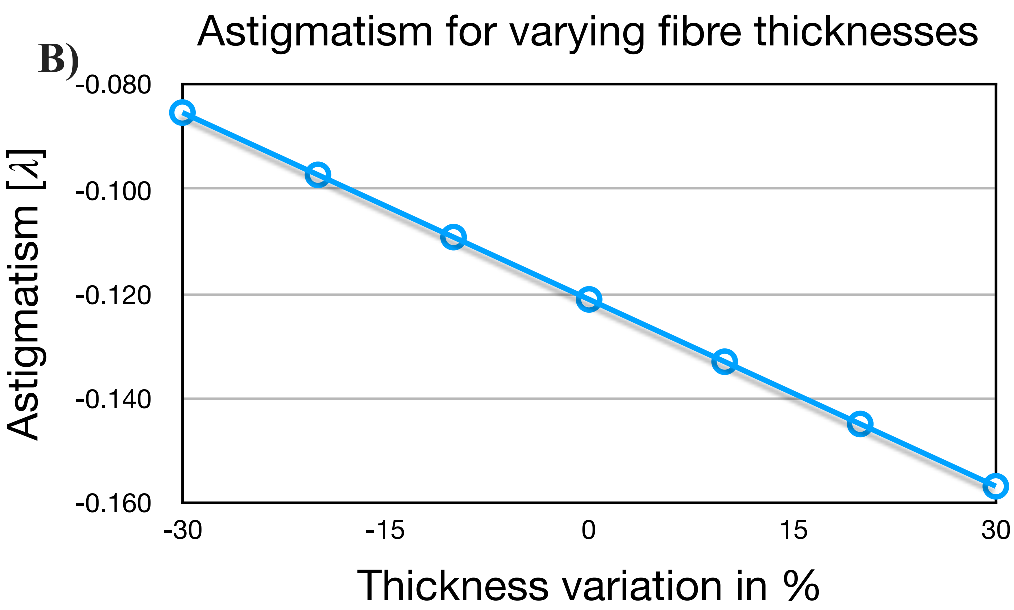
Spherical aberration varies approximately linearly with varying refractive index and as well as with coverslip thickness. An error of 0.02 in refractive index leads to a difference of about 0.007λ in spherical aberration. This lies well within the Maréchal range. For the coverslip, a 10% error in thickness leads to a change of about 10% in spherical aberration (~ 0.007λ rms), again in the Maréchal range. For residual aberrations within the Maréchal range, a single modal optimisation suffices for correcting residual aberrations after the pre-correction has been applied.

For the fibre, variations in astigmatism and spherical aberration were analysed for changes in refractive index and thickness. When changing the thickness of the fibre, the curvature of the fibre changed accordingly. The results for astigmatism and spherical aberration are shown in Fig. 4.9.

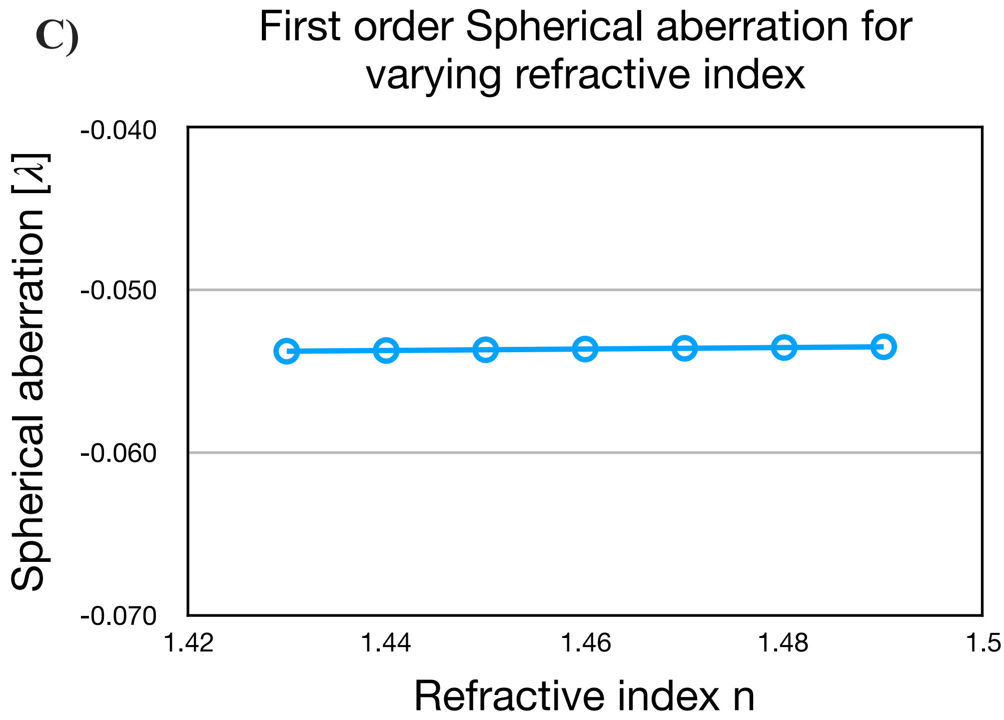
A)



B)



C)



D)

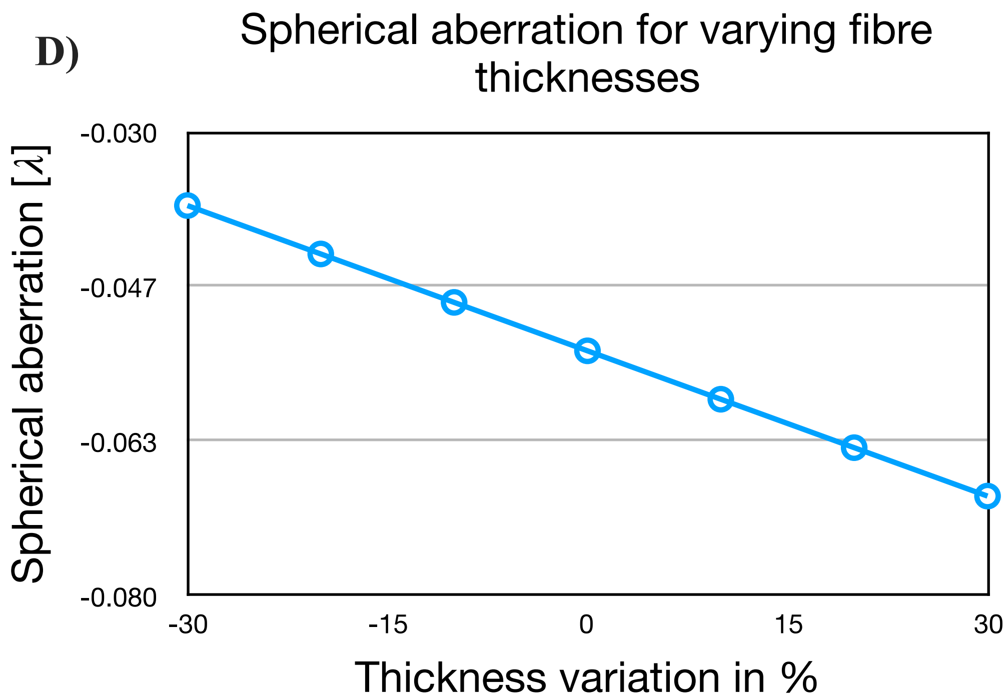


Figure 4.9: Variations of astigmatism and spherical aberration for varying fibre refractive index and varying fibre thicknesses

A change in thickness by about 20% leads to a change in spherical aberration of about 0.01λ rms. Astigmatism is more sensitive to changes in refractive index. A 0.006

change in refractive index (of either the fibre or the surrounding medium - glycerol) would result in an astigmatism amplitude change of about 0.07λ rms. According to *Hoyt* [91], that would correspond to a glycerol/water concentration error of about 7%. The glycerol solution used had a concentration of ≥ 99.5%. Variations in the fibre thickness of 10 % give rise to about 10% change in the aberration coefficients. As long as residual aberrations are within the Maréchal range (diffraction limited) a parabolic fit can be used for each Zernike mode to determine the final correction.

Relating these results to samples often used in microscopy. When it comes to biological samples, cell organelles are known to have different refractive indices, usually varying between 1.36-1.41 [92, 93] with a typical average refractive index of 1.38 for the whole cell [94] (the common refractive index variance within cells and tissues is usually < 0.03 [95]). These ranges are close to the ranges discussed before.

4.11 Efficiency of the ray-tracing pre-correction method

A comparison between the ray-tracing pre-correction method and a standard modal optimisation was made to demonstrate the efficiency of the proposed optimisation method by correcting the fibre aberrations (see Fig.4.5). The result is shown in Fig. 4.9. The top line of Figure 4.9 shows a standard modal correction where each Zernike mode was taken in turn and optimised by applying aberration amplitudes of $\pm 0.07\lambda$ to the DMM and measuring signal changes at the pinhole. At each iteration step, a wavefront correction was obtained by fitting parabolas through the obtained intensity data (see equation 4.2). The determined wavefront correction was used as the first step for the next iteration until no further improvement can be obtained. The bottom line in Figure 4.9 represents the newly proposed method, where a pre-correction wavefront is used as the starting point for the first modal optimisation. As can be seen in Figure 4.9, when a pre-correction is applied to the DMM, the modal correction speed has significantly increased and the number of iteration steps are reduced. For the example presented the number of iteration steps, to achieve a similar level of aberration correction, was decreased from three to one by using a pre-corrected wavefront as the starting point.

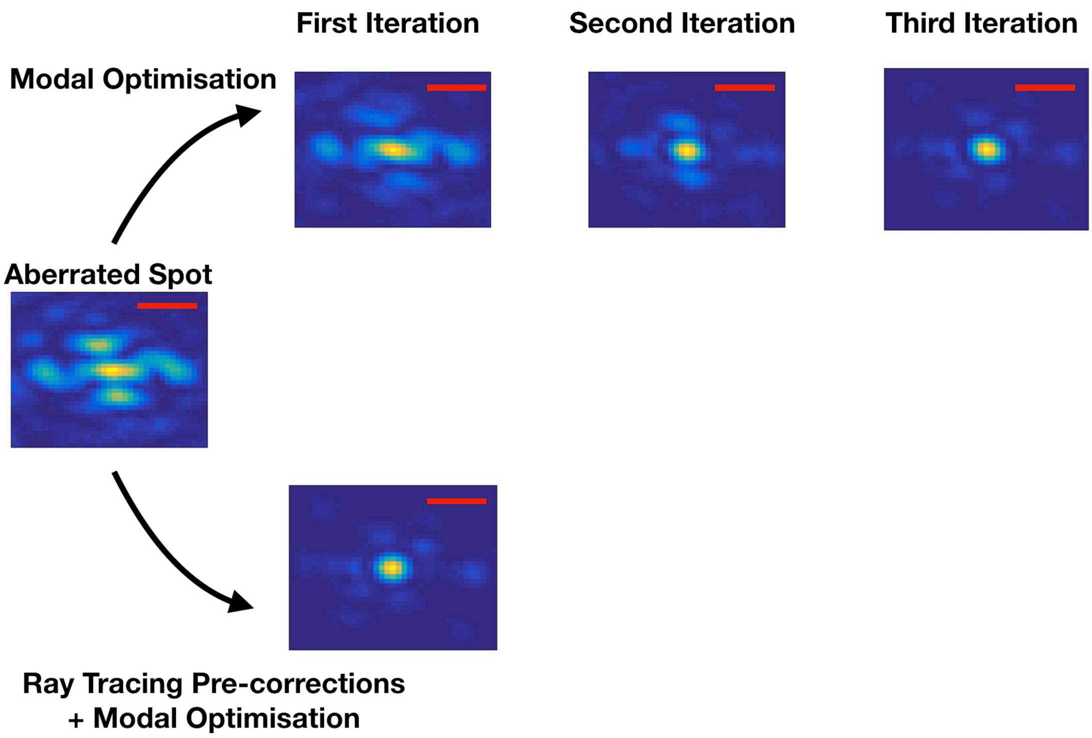


Figure 4.10: A comparison between a conventional iterative modal correction and the proposed new method using a ray tracing model to calculate a pre- correction. The sample used for this example was a 125um diameter coreless termination fibre (refractive index 1.4613), immersed in glycerol (refractive index 1. 4724). The length of the red line is approximately 2.0 um.

In the presence of large aberrations, the ray-tracing pre-correction method speeds up the aberration correction process. Best corrections for the tested samples were achieved after a single optimisation, while the traditional modal approach would have required up to 3 iterations to obtain comparable results. To put some numbers to these claims, let's consider the correction of the coverslip and fibre aberrations. Assuming a 20ms exposure time for each camera image and 5ms for applying an aberration with the DMM (the DMM has an operation bandwidth of 200Hz), the correction of the coverslip and fibre aberrations would require 23 camera images to be taken over a time period of 0.575s, when the pre-correction is utilised (based on 11 measured Zernike modes: 2 images per measured Zernike mode and a single reference image with the pre-correction applied which sums up to 23 images). Without the ray-tracing pre-correction, while using the Facompress et al. approach [50], the coverslip aberration correction would require a second optimisation which would result in a total of 50 camera images taken over a time interval of 1.15s. Correction of the coverslip aberration would thus require about twice the amount of data and would double the optimisation time when ray-tracing pre-correction is not used. As for tackling the large fibre aberrations, 5 images per measured Zernike mode with 3 iterations would be recommended [50] which translates to 165 taken camera images over a time period of 3.38s. Using a pre-correction for the fibre sample would make the optimisation about 6 times faster.

4.12 Conclusion

A modal wavefront correction method is used to determine aberrations by sequentially applying Zernike modes and monitoring the variation of a suited metric (e.g. intensity transmitted through a pinhole). For small pinhole sizes (< 0.3 AU), the transmitted intensity through the pinhole is a reasonable proxy of the Strehl ratio, equation (4.1), which predicts a parabolic dependence of the Strehl ratio on the Zernike coefficients, for aberration magnitudes below 0.07 λ rms. It is noted that our experiments were performed in a double pass setup which doubles the magnitude of even aberrations (odd aberrations are cancelled). We have shown that ray tracing simulation results provide reliable prior-information about the sample aberrations present. Applying a pre-correction, based on ray tracing, removes the majority of the aberrations allowing the final modal correction routine to be performed on a weakly aberrated focal spots. When using pre-correction the remaining aberrations can be determined by a simple parabola curve fitting of the data. Although, for the data presented, parabolic curve fitting was performed with five data points per Zernike mode, comparable results were obtained with only three data points.

This is compared with the method proposed by Facomprez et al. [49]. In the study of Facomprez et al, modal aberration correction was performed on the illumination path of a two-photon microscope. Facomprez et al. uses different strategies for modal aberration correction depending on the aberration amplitude, the number of measurements per Zernike mode (either 3, 5 or 9) and the amplitude of the aberration bias, the number of iterations performed changed from 1 to 3 [49]. If the Facomprez approach were used for our fibre and air gap samples, at least 3 iterations, using a minimum of 5 sample points per Zernike mode at each iteration, would be required to obtain diffraction limited imaging. For the coverslip experiment, at least one iteration, involving 5 measurements per Zernike mode, would suffice. Here we have demonstrated that by using a pre-correction as a starting point we can achieve diffraction limited resolution after one optimisation with only 3 measurements per aberration. The experimental results presented suggest that optimising only the Zernike modes predicted by ray tracing is sufficient to achieve a final PSF comparable to the one obtained when no sample is present and the system aberration corrected for, again speeding up the correction process. Alternatively, instead of using pre-corrections, Lukosz modes can be used, instead of Zernike modes, to correct aberrations larger than 0.07 λ rms [50]. However, such a method would require measurements of the mean rms spot radius and therefore is less suitable for confocal detection. Fast aberration corrections with a minimum number of iterations and measurements per Zernike modes are desired because of the need to reduce the amount of photo-damage caused by sample exposure to laser radiation.

Often the refractive index or the thickness of a sample are not precisely known and this can lead to an inaccurate ray tracing prediction. Aberration amplitudes tend to vary linearly with the dimensions of an object. As regards refractive index differences, a small change in refractive index gives also rise to an approximately linear variation in aberration. Even when there are some ambiguities in the true refractive index values and dimensions of the sample, the ray-tracing pre-correction is beneficial in speeding up the modal aberration correction process.

In conclusion, using a ray-tracing pre-correction approach simplifies the final correction procedure, reducing the time taken to complete a modal aberration correction, and therefore minimising radiation damage. With pre-corrections, large aberrations were corrected first, and, thereafter, the remaining aberrations were corrected in a single step with a modal optimisation routine. For scenarios, where sample parameters (refractive index, thickness, curvature, etc) are not well known, an estimated ray-tracing pre-correction would still reduce the sample aberration to a more manageable level first. Moreover, this technique would provide more flexibility when correcting for aberrations, e.g. one might go deeper into transparent samples, or one could make local corrections when scanning an image plane.

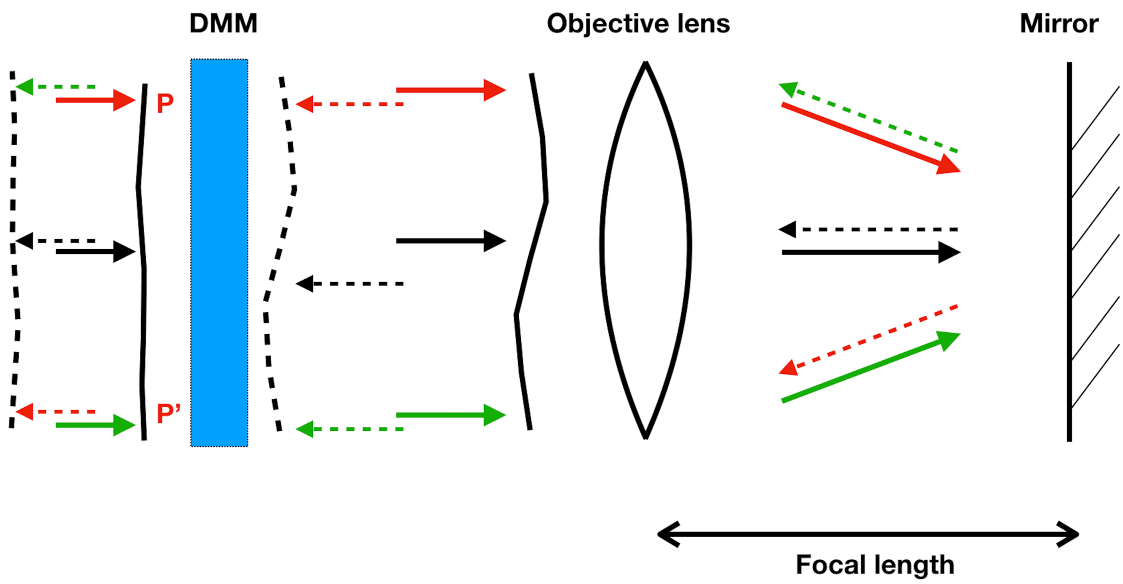
In order to achieve efficient and speedy correction when sequentially applying the modal correction routine, it is important to limit aberration cross-talk so that different Zernike modes have independent influence on the pinhole intensity and a correction made to one mode does not influence the magnitude of another mode. With the new method presented here, the effects of cross-talk are reduced because the modal correction is performed on a weakly aberrated spot. The main purpose of the proposed technique is to speed up the correction process, and under certain conditions, help the system to reach its optimum performance.

Regarding applications, non-contact surface profiling or imaging through layered structures or on immersed samples lies well within the capability of this technique. This technique could potentially speed up the correction of aberrations caused by biological samples which do not feature complex refractive index distributions, and can therefore be modelled using ray-tracing [96]. For more complex samples, this approach would be less suitable. However, a ray-tracing pre-correction could still provide a better starting point for modal correction. Ultimately aberration correction in biological samples is limited by scatter. Scatter increases with depth more strongly than aberrations. For highly scattering samples, aberration correction brings little improvement because significant amounts of signal are already lost by scatter [97]. Strongly scattering media require wavefront shaping techniques which impose local high frequency phase changes [98] involving a large number of actuators. Spatial light modulators have been used to focus light through a scattering media [99]

5. Coma detection in confocal reflection microscopy by means of an edge scan

In the previous chapter a confocal modal wavefront sensor, originally proposed by Neil et al. [47] was implemented and studied in a home-built reflection confocal microscope. Aberrations can be indirectly measured and corrected, by sequentially applying Zernike modes with a DMM and maximising the intensity passing through the pinhole of the confocal microscope. The confocal reflection microscope described in Chapter 3 is a double-pass system, where the incoming light path and reflected light path pass of different sides of any element present in the system such as the DMM and odd aberrations are cancelled after passing twice through the system. This is illustrated in Figure 5.1A). For illustration of the double-pass effect for a mirror sample, contour plots of aberrated spots at the object and detector plane are shown in Figure 5.1B).

A)



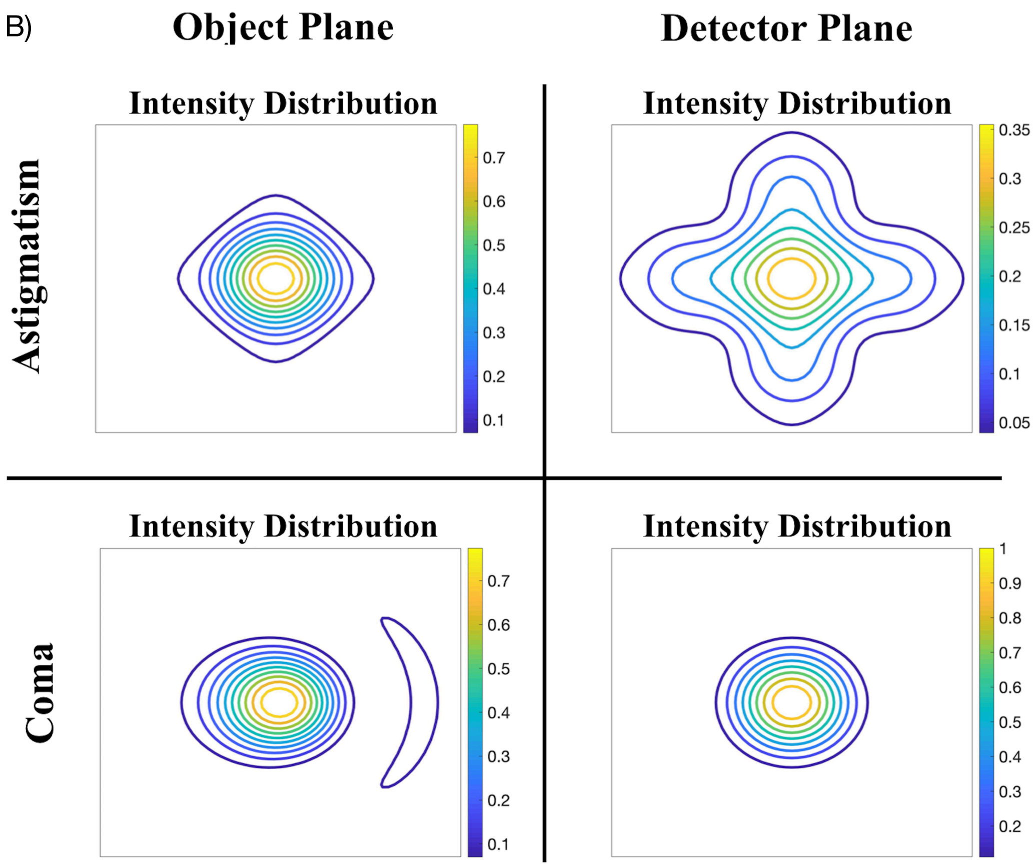


Figure 5.1: A) Illustration of the double-pass effect on a reflection microscope with an integrated DMM. The DMM imposes an odd (coma-like) aberration on the incoming wavefront (solid lines). The reflected wavefront and rays are represented by dashed arrows and lines. A ray passing through P will after reflection pass through the point P'. B) Illustration of the double-pass effect on focal spots in a reflection microscope. Aberrated spots at the object and detector plane are compared for an even aberration, astigmatism (top row), and odd aberration, coma (bottom row).

The DMM applies an odd (coma-like) aberration on the incoming wavefront (solid lines). The objective lens focusses the beam and will form an asymmetric coma-like spot on the mirror. Then, the light is reflected back from the mirror in reverse direction from right to left, and after a second pass through the objective it will be reflected off the DMM a second time, whereby an incident marginal ray which is reflected off the DMM at P in the first pass will be reflected at P' in the second pass. As a consequence, DMM imposed odd aberrations, such as coma, are cancelled, after the second reflection of the DMM. In a confocal microscope the light that is reflected off the DMM after the second pass will be used to form a spot on a pinhole, which will thus be devoid of information about asymmetric odd aberrations.

Thus, odd aberrations, such as coma and trefoil, can't be detected with a confocal reflective modal wavefront sensor. The same applies also to a Shack-Hartmann wavefront sensor or interferometric sensing device measuring the reflected wavefront. To detect odd aberrations, the wavefront should be measured in single pass, i.e. by measuring the wavefront in transmission, or by using fluorescent samples. Fluorescent emission would constitute a new source of light which lacks the phase information of the excitation beam. Odd aberrations in the detection path would thus affect the wavefront quality of the fluorescent signal.

It is important to mention that the scanning spot would be affected by odd aberrations and would result in a lower quality image of a sample. Normally the DMM would be used to correct for wavefront aberrations (system or sample induced aberrations), but for demonstration purposes the DMM was used in experiments to apply controlled amounts of coma to show the resulting degradation in image quality.

Edge scans can be used in confocal microscopy to assess lateral resolution. Gu et al. [100] studied the effects of defocus and spherical in terms of Seidel aberrations on the confocal image of a straight edge. They found that small amounts of defocus and spherical aberration can lead to a steeper edge response. However for large amount of spherical aberration, the steepness of the edge response decreases and inflection points appear at the top and bottom parts of the edge response. Interestingly for an ideal PSF, the intensity at the edge is around 0.25-0.33 of the maximum intensity, depending on the pinhole size (0.25 for an infinitely small pinhole and 0.33 for an infinitely large pinhole) [101]. In this chapter, a method is presented to detect coma in a double-pass reflection confocal microscope by means of an edge scan. The asymmetric edge breaks the double-pass effect. In the presence of coma, as will be shown, the resulting image of the edge will have a characteristic form and be less steep than if it would be when the edge is scanned with an unaberrated diffraction-limited focal spot. The orientation of the coma wavefront aberration (or the coma flare) can be determined by scanning the focal spot along two perpendicular edges. A comparison of experimental and theoretical curves makes estimation of the amount of coma possible.

5.1 Confocal image of an edge in the presence of aberrations

5.1.1 Coherent imaging

The image of an edge obtained with a confocal microscope, such as the one described in Chapter 3, can be simulated with the help of coherent imaging theory [20], using the Fourier transform and Fraunhofer approximation of the diffraction integral (see Chapter 2 section 2.1). However, for convenience we will use the normalised amplitude point spread function $h(x,y)$ at the focal plane which is given by [13]:

$$h(x,y,z)=\frac{1}{\pi}\iint_{\Omega}P(\nu,\mu)\; e^{-j2\pi(x\nu+y\mu)}\,d\nu\, d\mu \tag{5.1}$$

Where $P(\nu,\mu)$ is the pupil function, λ is the wavelength, j the imaginary unit, k the wave vector, F the focal length. The definition for the normalised optical coordinates is given in Chapter 2 section 2.1. The integral is evaluated over the unit circle Ω.
For the pupil function we use the following expression:

$$P(\nu,\mu)=A(\nu,\mu)\cdot e^{ik\Phi(\nu,\mu)} \tag{5.2}$$

Where $A(\nu,\mu)$ is the transmission function and $\Phi(\nu,\mu)$ is the aberration function which can be expressed in terms of Zernike polynomials:

$$\Phi(\nu,\mu)=\sum_{n,m}\alpha_n^m Z_n^m(\nu,\mu) \tag{5.3}$$

Here α_n^m represents a Zernike amplitude coefficient and Z a Zernike mode (see Chapter 2 section 2.2). $\Phi(\nu,\mu)$ describes the imposed aberrations (system aberrations, sample aberrations and aberrations imposed by a DMM).

The edge, in the image plane, can be described by a two dimensional binary function which has a value of 1 on one side of the edge and is nearly zero on the other side (we chose a value of 0.1 instead 0 for numerical stability reasons). The edge function, denoted by t, is:

$$t(y)=\begin{cases} 1 \ \text{ if } y \le y_{edge} \\ 0.1 \ \text{ if } y > y_{edge} \end{cases} \tag{5.4}$$

Where y_{edge} is the y position of the edge. The edge, programmed in Matlab, is shown in Fig. 5.2.

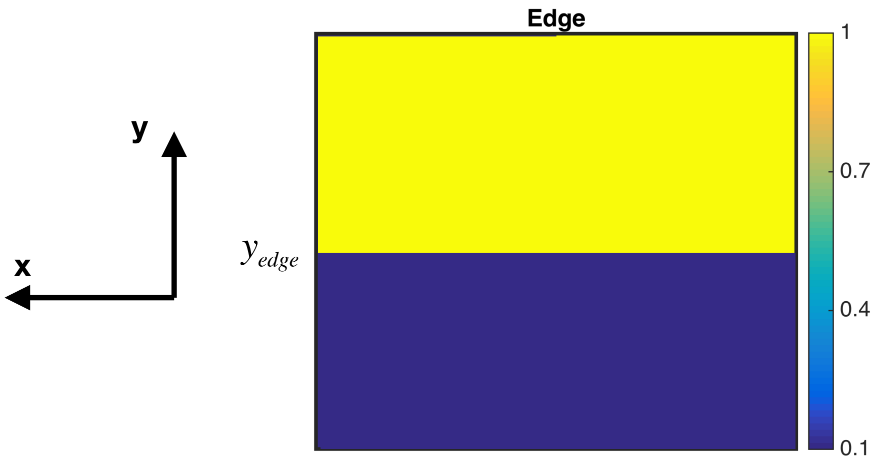


Figure 5.2: Edge in the image plane

The reflected amplitude at the pupil, $P'(\nu,\mu)$ is obtained by taking the inverse Fourier transform of the product of the amplitude PSF with the edge reflection function:

$$P'(\nu,\mu,z)=\pi\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x,y)\cdot t(x,y)\ e^{j2\pi(x\nu+y\mu)}\,dx\,dy \tag{5.5}$$

The pupil boundary conditions (Equation 5.5) then have to be applied on P'. The edge scan can be simulated by varying the value of y_{edge}. The final step is to compute the Fourier transform of P' with the wavefront aberration imposed after double-pass through the system. The imposed wavefront aberration on the return light path is given by $\Phi_{DP}(\nu,\mu)=\Phi(-\nu,-\mu)$. Φ_{DP} is obtained by mirroring Φ with respect to the pupil centre. The amplitude point spread function at the detector is then:

$$h_D(x_D,y_D,z_D)=\frac{1}{\pi}\iint_{\Omega} P'(\nu,\mu)\cdot e^{jk\Phi_{DP}(\nu,\mu)} \quad e^{-j2\pi(x_D\nu+y_D\mu)}\,d\nu\,d\mu \tag{5.6}$$

Where (x_D,y_D,z_D) are normalised cartesian coordinates in the detector plane. The intensity at the detector plane is obtained by multiplying the amplitude with its complex conjugate:

$$I_D=\left|h_D\right|^2 \tag{5.7}$$

Finally in order to obtain the confocal signal (intensity passing through the pinhole), one uses a circular mask, such as the one described in chapter 3, and integrates the intensity values within the mask. In fig 5.3 the edge response for an aberration free focal spot is shown for different pinhole sizes.

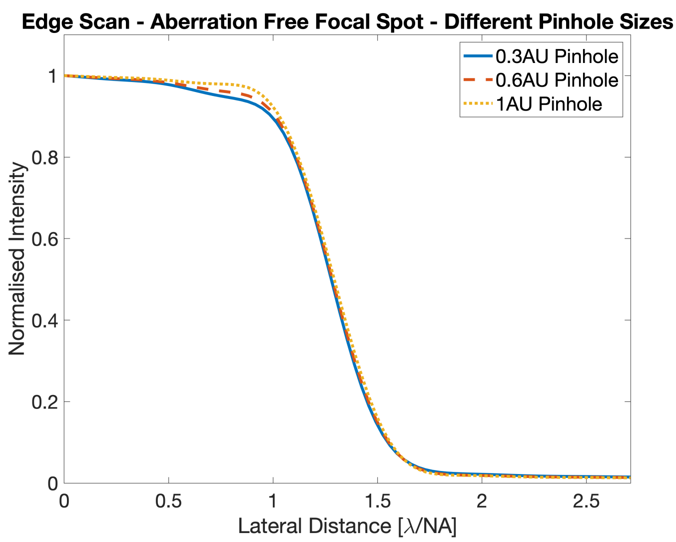


Figure 5.3: Confocal edge response for various pinhole sizes. Each curve is normalised with respect to its intensity value at $y=0$. AU stands for Airy Unit.

The edge response for the various pinhole are quite similar. It is noted that the 20-80% width of the edge responses is sometimes used as a criterion for the lateral resolution [102].

5.1.2 Simulations of confocal edge responses in the presence of aberrations

In the previous section the coherent imaging model used for calculating the edge response of a reflection confocal microscope with a DMM operating in a double-pass configuration was introduced. At focus the intensity distributions produced by even aberrations, such as astigmatism or spherical aberration, are not unique since one can not determine the sign of the wavefront aberration. The aberrated spots, in focus, of fig 5.4 A),B),E),F) show that the sign of even aberrations does not affect the diffraction pattern at focus. The intensity distribution is rotated by 180 degrees when the sign is changed, this result for astigmatism and spherical aberration in intensity distributions which remain unaffected by a sign change. The sign ambiguity of even aberrations can be overcome by observing the focal spot through focus, because the intensity distribution will be different on opposite sides of best focus (the term best focus will be used from here on to refer to the spot, in the vicinity of the geometrical focus, with the highest intensity in x,y,z).

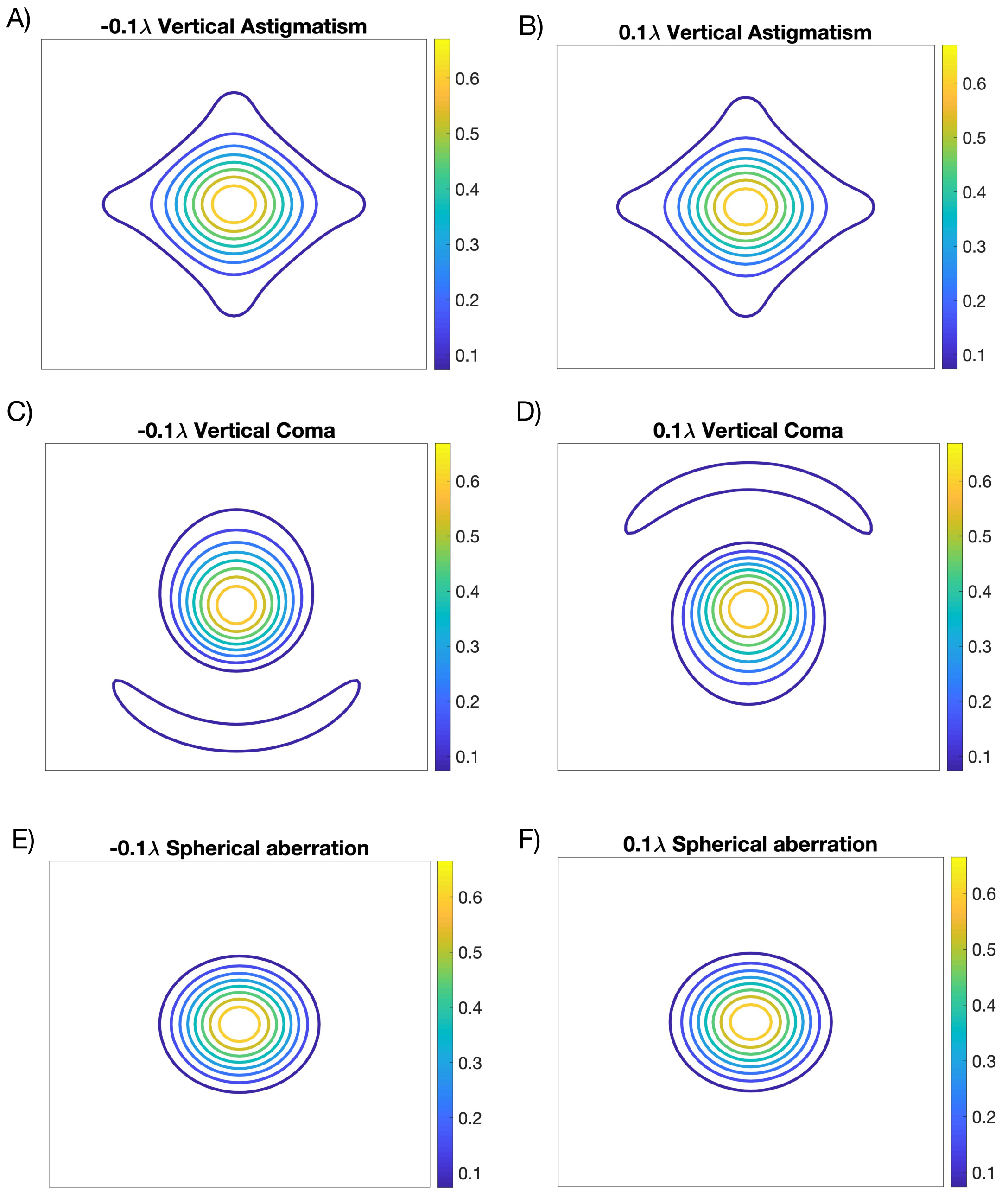


Figure 5.4: Intensity distributions at focus in the presence A-B) vertical astigmatism, C-D) vertical coma and E-F) spherical aberration. In the left column 0.1λ of wavefront aberration was applied whereas in the right column -0.1λ of wavefront was applied.

As regards odd aberrations, it is noted that depending on the sign of an odd aberration and the orientation of the edge (see Fig.5.5), the edge response can be different.

Two edge orientations were considered for the simulations, namely a horizontal or vertical edge, see Fig. 5.5 A) and B), respectively. In both figures, an aberrated spot with vertical coma or horizontal coma were added to show the orientation of the coma tail.

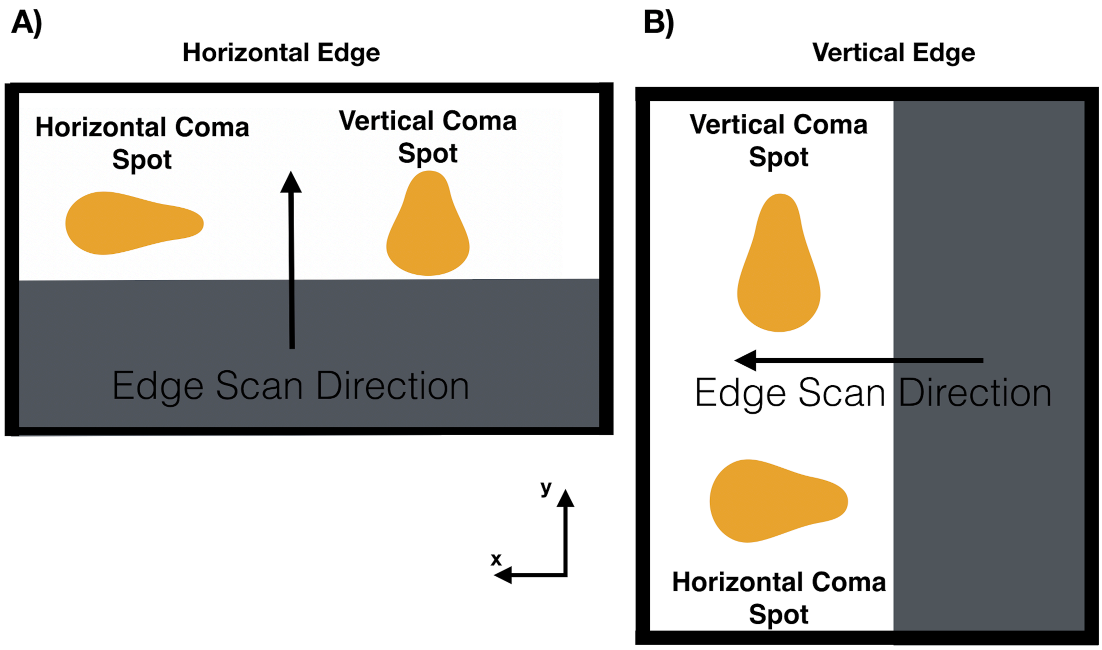


Figure 5.5: Horizontal and vertical edge. The orientation of the coma tail for a spot with vertical coma or horizontal coma is shown in both vertical and horizontal edge. The scan direction is depicted by the arrow. The white area represents the reflective surface whereas the grey area represents the non-reflective surface.

In Fig. 5.6 edge responses in the presence of primary spherical aberration, vertical astigmatism, oblique astigmatism, horizontal coma and vertical coma are plotted for different amounts of aberration amplitudes with an edge with extents in the horizontal direction (see Fig. 5.5A) in the best focus plane. The intensity at the detector was calculated using equations (5.7) and (5.8) and the phase aberration (equation 5.4) was expressed in terms of Zernike modes. The intensity, at the detector, was spatially filtered with a 0.6AU diameter pinhole. Each curve was normalised with its intensity value at the lateral distance $x = 0$. The aberration amplitude range for odd aberrations was chosen to be [-0.105λ, 0.105λ] which is larger than the range for even aberrations: [-0.07λ, 0.07λ] in order to avoid large aberrations at the detector because even aberrations are doubled in terms of amplitude (due to the double-pass effect).

At best focus a focal spot which is produced by even aberrations, such as astigmatism and spherical aberration, does not depend on the sign of the wavefront aberration as will the edge responses. Therefore, Fig. 5.6 A-C will look the same, if the sign of the wavefront aberrations becomes negative. Horizontal coma produces a “comet” shaped focal spot, with the coma tail orientation being parallel to the edge (see Fig. 5.5A). By changing the sign of the horizontal coma coefficient, one effectively flips the coma tail orientation by 180°. Therefore it becomes clear that the edge response, when scanning the spot perpendicular to the edge, will not depend on the sign of the horizontal coma wavefront aberration. Vertical coma, however, will produce two distinct types of edge responses depending on the sign of the wavefront aberration, as can be seen in Fig. 5.6 E) and F). For the Zernike aberration notation reference is made to Chapter 2 section 2.2.

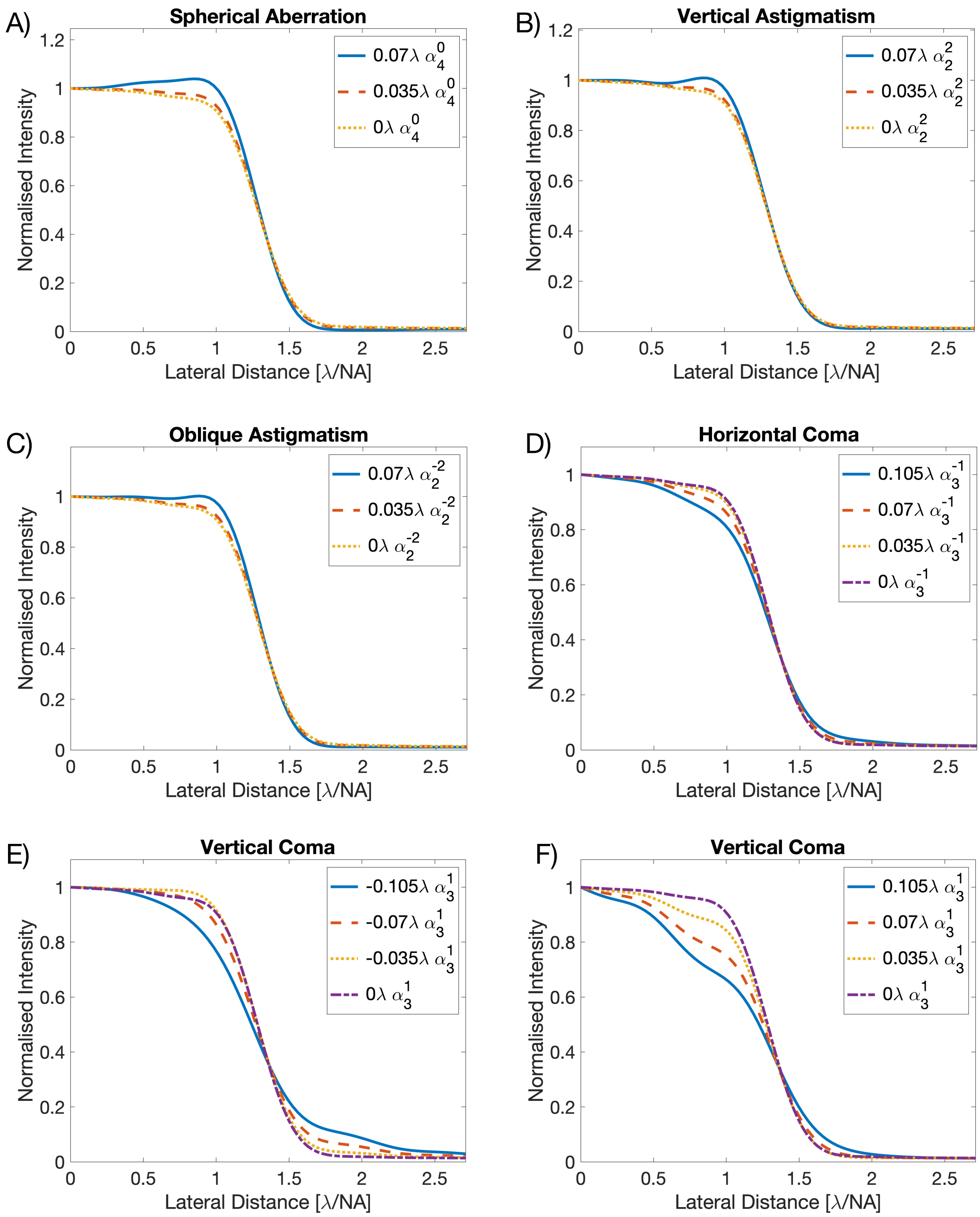


Figure 5.6: Confocal edge response of an edge extending in the horizontal direction. Each curve is normalised with respect to its intensity value at the lateral distance $y = 0$ .Primary aberrations: A) spherical aberration α_4^0 , B) vertical astigmatism α_2^2 , C) oblique astigmatism α_2^{-2} , D) horizontal coma α_3^{-1} , E) vertical coma α_3^1 (negative amplitudes) and F) vertical coma α_3^1 (positive amplitudes). Pinhole diameter: 0.6AU. The aberration amplitude was varied in the range of [-0.105λ, 0.105λ] for odd aberrations and [-0.07λ, 0.07λ] for even aberrations, in steps of 0.035λ .

The edge response one obtains will depend on whether the coma tail is scanned first or last over the edge. For negative amounts of vertical coma (i.e. when the coma tails points away from the edge, as shown in Fig. 5.5A vertical coma spot), the edge response at the bottom part decreases more slowly. Whereas for positive amounts of vertical coma (i.e. when the coma tails points towards the edge), the intensity decreases more slowly at the the top part of the edge. This slow decrease in intensity becomes more pronounced when the aberration amplitude increases. Especially the edge response in Fig. 5.6 F) is quite distinct and could thus be used to detect the presence of coma in a reflection confocal microscope. The slope of the edge response at 0.5 (normalised intensity) is visibly affected for aberration amplitudes > |0.07| λ. If the edge orientation is vertical as in Fig. 5.4B and the spot is scanned along the x-direction, the spot with horizontal coma would produce characteristic edge responses. The edge response would depend on the sign of α_3^{-1} whereas for vertical coma it wouldn't. The edge responses for a vertical edge for vertical and horizontal coma are shown in Fig. 5.7. The edge responses for vertical coma do not depend on the sign of α_3^1 and therefore only edge responses with positive of α_3^1 were plotted. Edge responses produced by a spot with astigmatism or spherical aberration are the same as in Fig. 5.6 A-C.

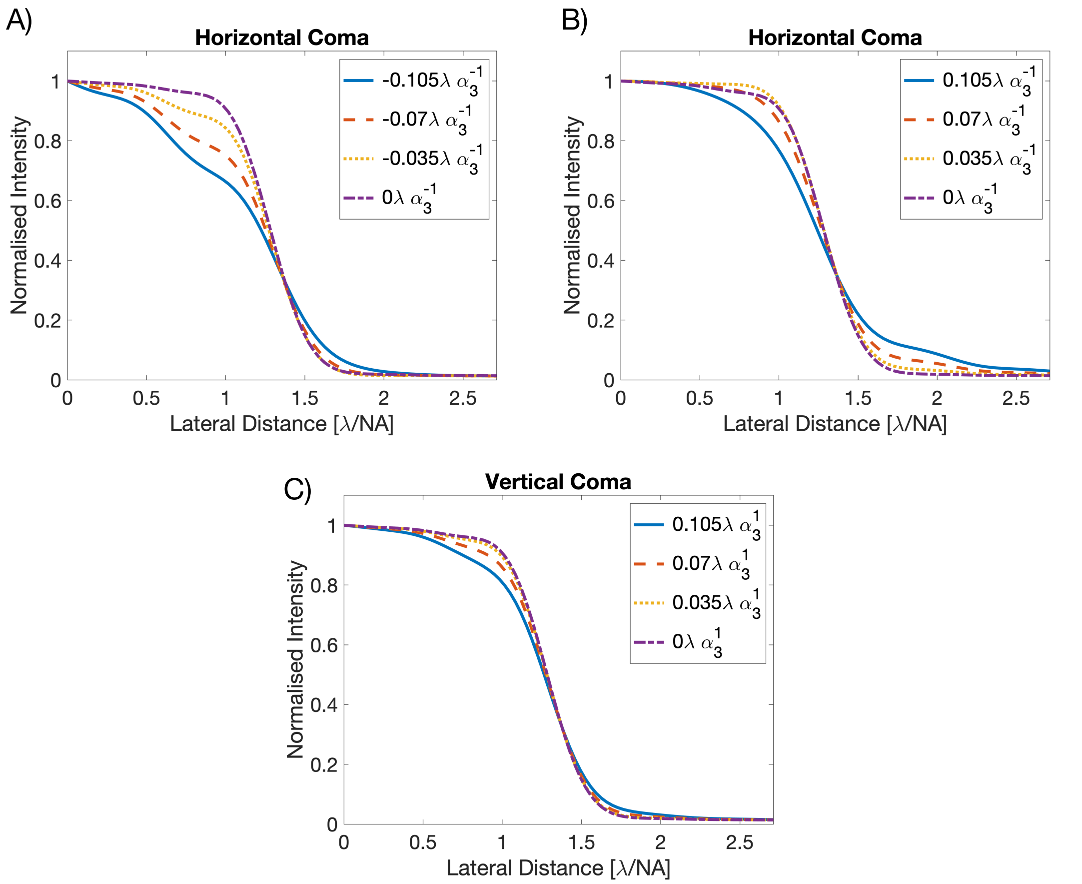


Figure 5.7: Confocal edge response of an edge extending in the vertical direction. Each curve is normalised with respect to its intensity value at the lateral distance $y = 0$.Primary aberrations: A) horizontal coma α_3^{-1} (negative amplitudes), B) horizontal coma α_3^{-1} (positive amplitudes) and C) A) vertical coma α_3^{1}. Pinhole diameter: 0.6AU. The aberration amplitude was varied in the range of [-0.105λ, 0.105λ] in steps of 0.035λ.

So, by scanning a focal spot across two edges which are perpendicular to each other, one can detect not only the sign and the orientation of the coma wavefront aberration, but can also give an estimation of the amplitude of the coma wavefront aberration. The influence of the pinhole size on the edge response of an edge extending in the horizontal direction (see Fig. 5.5 A) in the presence of vertical coma is shown in Fig. 5.8. The edge response for smaller pinholes is slightly more degraded.

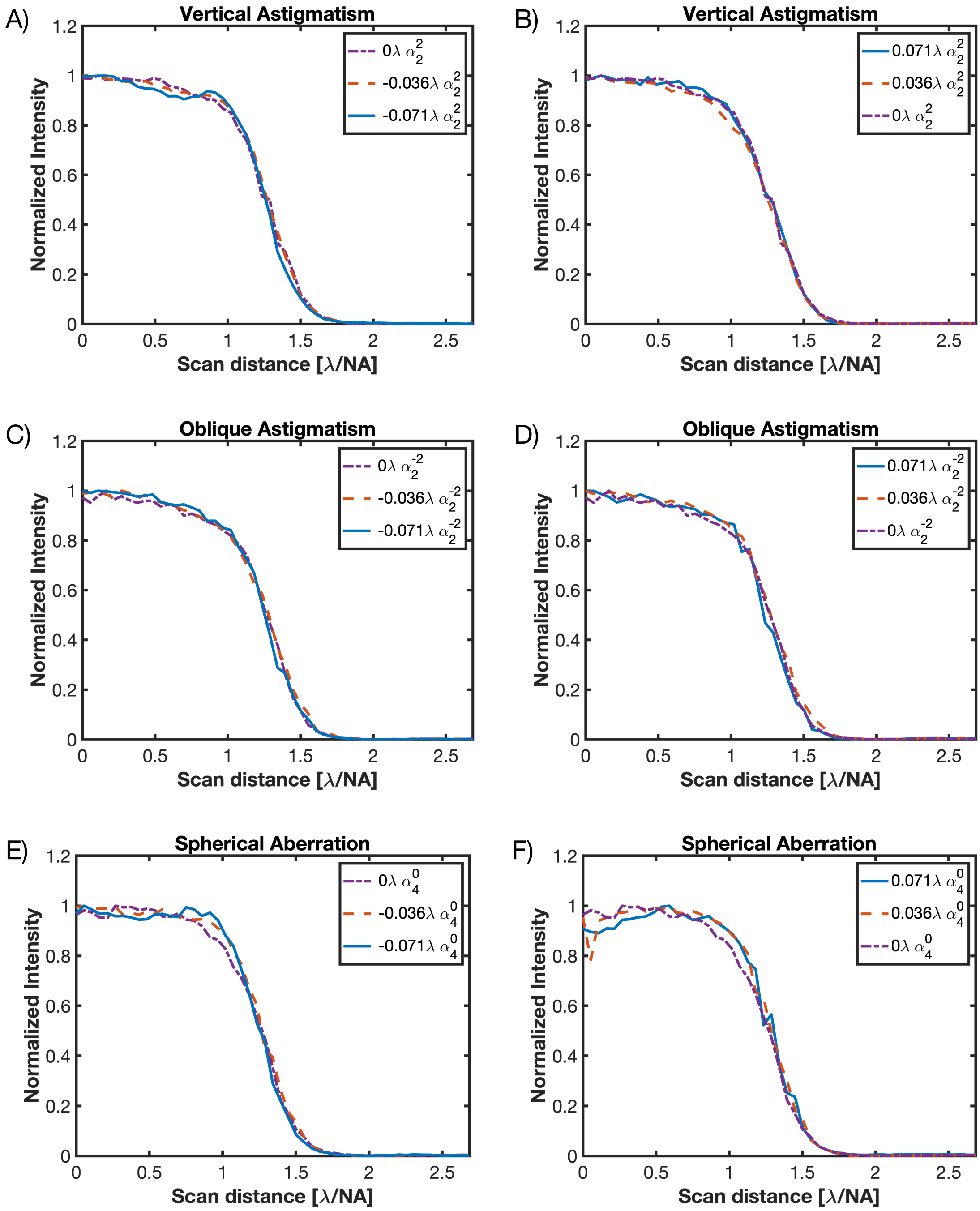


Figure 5.8: Confocal edge response of an edge extending in the horizontal direction in the presence of vertical coma for various pinhole sizes. The applied aberration amplitude is 0.07λ . Each curve is normalised with respect to its intensity value at the lateral distance $x = 0$. Chosen pinhole diameters: [0.3, 0.6, 1] AU.

5.1.3 Experimental confocal edge responses in the presence of aberrations

A USAF microscope resolution target - positive pattern (Edmund Optics, chrome pattern on glass) was scanned with the reflection confocal microscope described in Chapter 3. Primary astigmatism, coma and spherical aberration were applied independently from each other with the deformable membrane mirror (DMM). The resolution target was brought into focus by finding the axial position where the focal spot, reflected off the chrome pattern (far away from an edge), had its highest intensity (best focus). Each aberrated focal spot was then scanned over an edge with extents in the horizontal direction (see Fig. 5.5 A). The first and last scans were with a corrected focal spot in order to check that the edge responses were similar and to assure that the resolution target had not significantly drifted laterally during the scan series. A characteristic set of experimental edge responses for even aberrations is shown in Figure 5.9.

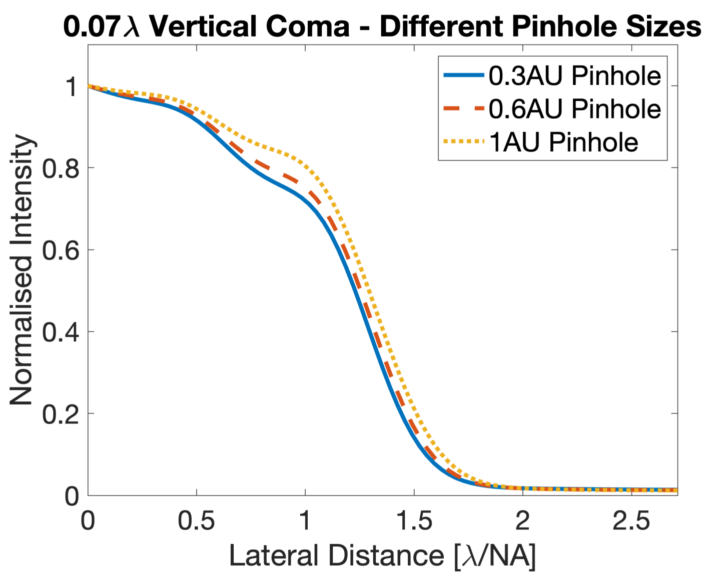


Figure 5.9: Experimental Confocal edge response of an edge extending in the horizontal direction. Each curve is normalised with respect to its intensity value at the lateral distance $x = 0$ .The edge responses for the different even primary aberrations are shown in A-B) vertical astigmatism, C-D) oblique astigmatism and E-F) spherical aberration. Pinhole diameter: 0.6AU. The aberration amplitude was varied in the range of ~[-0.07λ, 0λ, 0.07λ] in steps of $\sim 0.035\lambda$.

The intensity dip at the left side of Fig. 5.9F is somewhat similar to what has been calculated by Gu et al [100] in the presence of spherical aberration with some defocus. A characteristic set of experimental edge responses for coma (odd aberration) is shown in Figure 5.10.

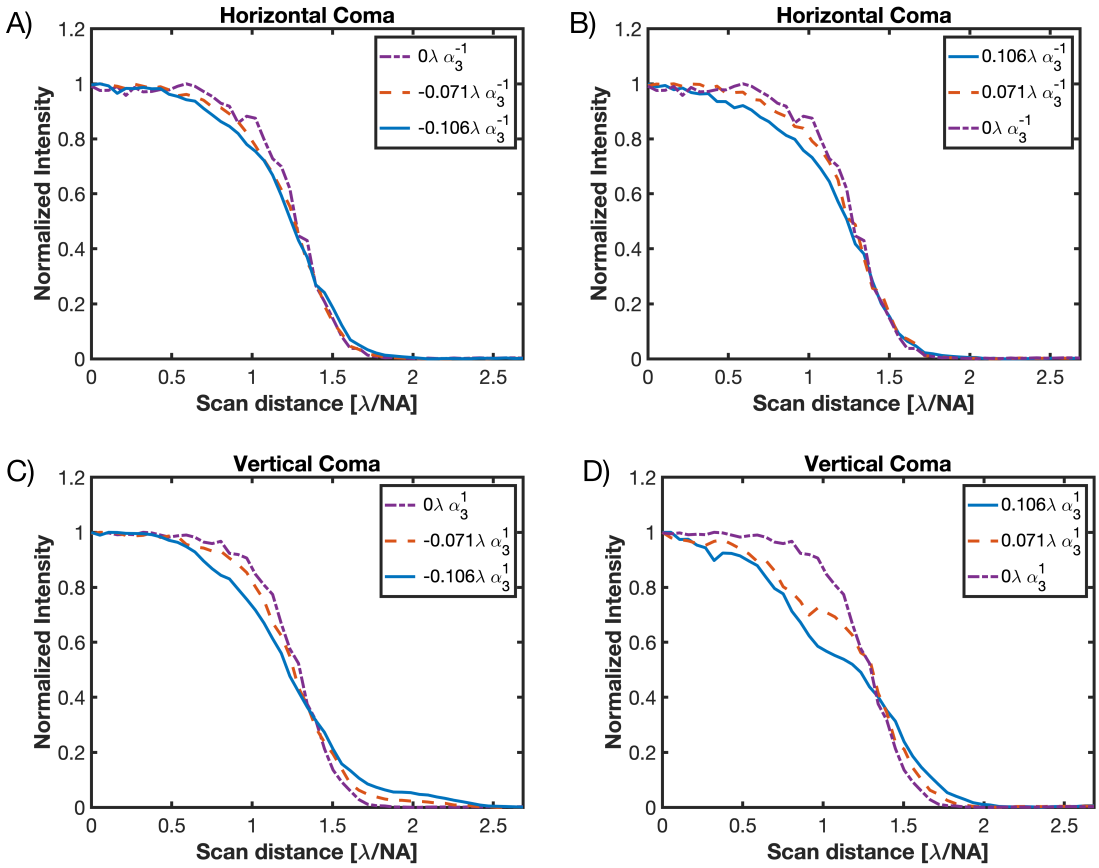


Figure 5.10: Experimental Confocal edge response for an edge extending in the horizontal direction. Each curve is normalised with respect to its intensity value at the lateral distance $x = 0$ .The edge responses in the presence of A-B) horizontal coma and C-D) vertical coma are shown. Pinhole diameter: 0.6AU. The aberration amplitude was varied in the range of ~[-0.106λ, 0λ, 0.106λ] in steps of $\sim 0.035\lambda$.

For all even aberrations (see Fig.5.9 A-D) as well as for horizontal coma (Fig.5.10 A-B), changing the sign of the wavefront aberration does not significantly change the shape of the edge response when scanning over an edge with extents in the horizontal direction (see Fig.5.5A). However, for vertical coma $\left(\alpha_3^1\right)$ one sees two distinct sets of edge responses depending on the sign of the aberration. The aberrated focal spots, at best focus, which were scanned over the edge resembled the focal spots shown in

Chapter 4 (Figure 4.2). For vertical coma edge response there is a slow decrease in intensity at the top / bottom of the edge depends on the sign of the coma wavefront aberration (see Fig.5.10 C-D).

To show the smearing, i.e. the slow decrease in intensity, caused by a coma aberrated spot on an image, a two dimensional area of 8x0.4µm of a USAF microscope resolution target was scanned in a x-y raster scan with different amounts of coma α_3^1 applied. The applied amplitudes for vertical coma were ~[-0.106λ, 0λ, 0.106λ]. Each image was normalised with respect to its highest intensity value. Pinhole diameter was chosen to be ~0.6AU. The USAF target was scanned with the piezo-stage. Scan steps were chosen to be 0.04µm. The scanned USAF microscope resolution target, for the different amounts of vertical coma, are shown in Fig. 5.11 .

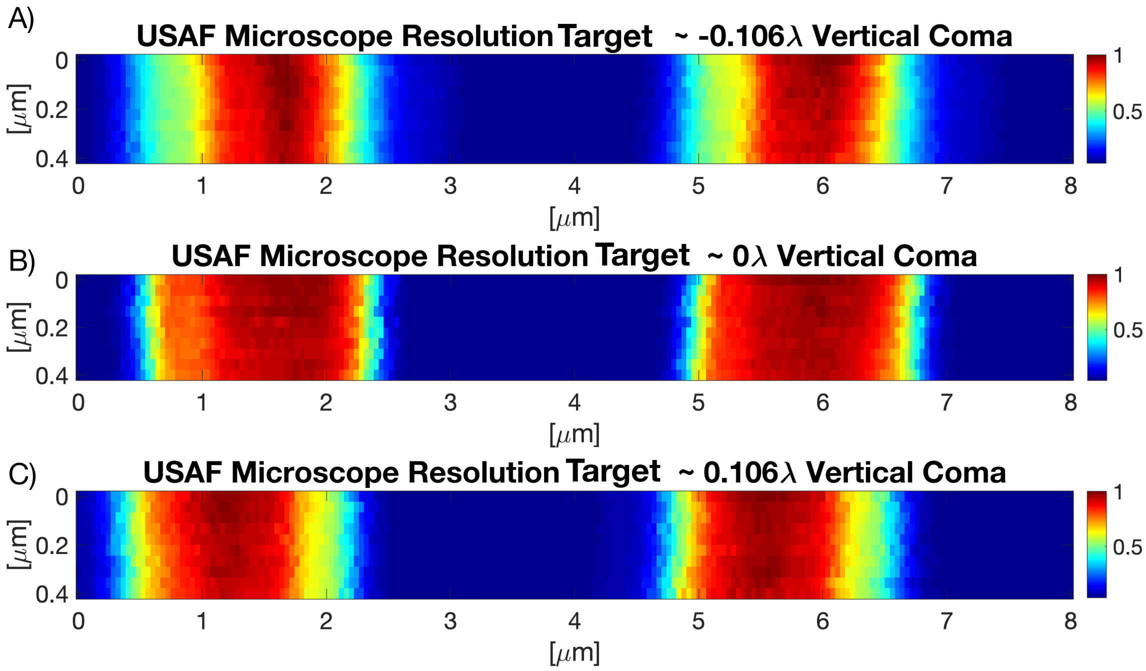


Figure 5.11: USAF microscope resolution target. A two dimensional area 8x0.4µm was scanned using 0.04µm scan steps. The applied wavefront aberration for vertical coma were ~[-0.106λ, 0λ, 0.106λ] for A),B) and C) respectively. The pinhole diameter was chosen to be ~0.6AU.

One can see that depending on the sign of the coma wavefront aberration (i.e. the orientation of the coma tail) one side of the stripes is more smeared than the other. Whereas for a corrected focal spot (Fig. 5.11B) both sides of the stripes look similar.

5.2 Comparison of simulation and experimental results for coma

In Figure 5.12, the simulated and experimental edge responses for coma are shown next to each other. The characteristic features of the vertical coma edge response (slow decrease in intensity at the top/bottom of the edge depends on the sign of the coma wavefront aberration) have been experimentally confirmed. The simulated and experimental edge responses show similar trends.

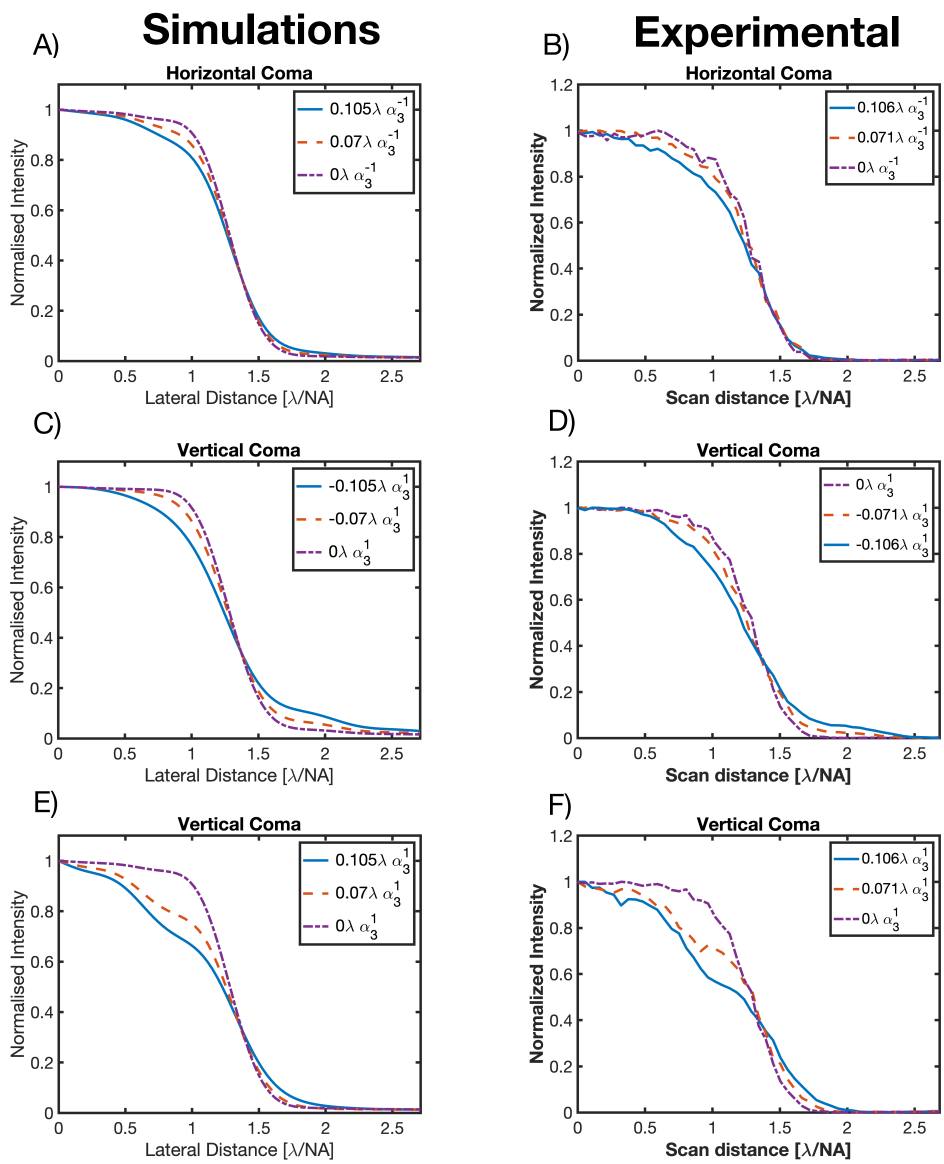


Figure 5.12: Comparison of Confocal edge response for an edge with extents in the horizontal direction in the presence of coma. Each curve is normalised with respect to its intensity value at the lateral distance $x = 0$. A/C/E are simulated edge responses and B/D/E are their experimental correspondent. The aberration amplitude was varied in the range of ~[-0.106λ, 0λ, 0.106λ] in steps of $\sim 0.035\lambda$.

5.3 Conclusions

The problem of the double-pass effect in reflection, which causes cancellation of odd aberrations after the second pass through the system, is that it makes detection of odd aberrations challenging. Wavefront sensors or indirect wavefront sensing, such as modal wavefront sensing, do not allow coma to be detected in a reflection setup. However, as presented here an edge scan allows coma to be detected in a confocal reflection microscope setup. The edge breaks up the double-pass effect and the edge response can be used to detect the presence of coma and allows for an estimation of the amount of coma. In this chapter, the effect of the primary aberrations on the image of an edge (or edge response) in a reflection confocal microscope were studied. Edge responses, in the presence of primary Zernike aberrations, were simulated and later measured with a home-built reflection confocal microscope (with a deformable membrane mirror to impose the aberrations, see Chapter 3). The edge responses were studied in the presence of a single aberration only, and not for aberration combinations. The results confirm that the presence of coma can be detected by observation of edge responses in a reflection confocal microscope, which coma aberrations would otherwise remain undetectable when using other sensing methods (e.g. wavefront sensor, modal sensing) because of the double-pass effect. The sign of the coma wavefront aberration as well as its orientation can be determined by scanning a focal spot across two edges (which should be perpendicular to each other in order to detect two orthogonal coma orientations, e.g. horizontal and vertical coma contributions). The smearing, i.e. the slow decrease in intensity, at the top of the edge response is a characteristic signature of a coma aberrated focal spot. The coma aberration amplitude can be estimated from the form of the edge response (see Figure 5.3 E) and F)). Small amounts of coma are, however, hard to detect because of noise, or could be confused with by local irregularities close to the edge. Nevertheless, coma amplitudes $> 0.035\lambda$ and up to about $\sim 0.14\lambda$, with an accuracy of $\pm 0.02\lambda$, are detectable with the proposed edge scan method. Furthermore the sensitivity to coma can be increased by reducing the pinhole size (see Fig. 5.5). Coma detection and correction in a confocal microscope could be achieved in a similar way to modal wavefront sensing by sequentially applying varying amounts of coma but by using another metric to optimise, i.e. “sharpest” edge response (see Fig. 5.6).

6. Aberration retrieval based on focal spot shape

This chapter presents a novel aberration retrieval method. This method is aimed at retrieving the amplitude of primary Zernike aberrations (astigmatism, coma, spherical aberration) in the pupil. These primary aberrations are retrieved by fitting a set of orthogonal functions to the intensity distribution of a beam as it propagates through focus. The resulting fitting coefficients are then combined to create characteristic aberration indicators that are sensitive to distortions in the intensity distributions caused by the primary aberrations. These indicators are ultimately used to retrieve the primary aberration amplitudes in the pupil.

Aberration retrieval is central to AO. Once the aberrations are measured, a DMM can be used for wavefront correction. In comparison to other direct wavefront sensing methods and modal/zonal optimisations, the aberration retrieval method proposed does not require any wavefront sensor or wavefront shaping device to measure aberrations but solely needs three intensity distribution images in the vicinity of focus to retrieve the primary aberrations. The method is quick in comparison to other commonly used phase retrieval methods. Intensity distributions are fitted within a circular region with a radius smaller than the distance to the first dark ring of the Airy spot and centred around the point of highest intensity. Due to the circular shape of the fitting region and the fact that orthogonal functions are well suited for mathematical fitting problems, Zernike polynomials were chosen as the fitting functions. Going forward the Zernike polynomials used for fitting will be referred to as orthogonal circle intensity polynomial to avoid confusion with the Zernike polynomials used to represent the aberrations present in the pupil. Thus the intensity distribution, within the circular region, is described as weighted sum of orthogonal circle intensity polynomials. Primary aberrations can be retrieved with a minimum of three intensity distributions, taken at focus and at two out of focus planes. The aberration indicators are computed from the fitting coefficients obtained at these three planes. There is an almost linear relationship between the aberration amplitudes and their respective indicators within an amplitude range of about [0.025λ, 0.13λ]. The issue of aberration cross-talk (when several aberrations are present) is addressed at the end of the chapter, and it is concluded that this aberration retrieval method is viable as long as aberrations do not become too strong (the rms wavefront deviation of all primary aberrations should remain below 0.1λ).

The outline of the aberration retrieval method, showing the essential steps, is shown in the block diagram (Fig. 6.1). The individual blocks are detailed in the different sections of this chapter.

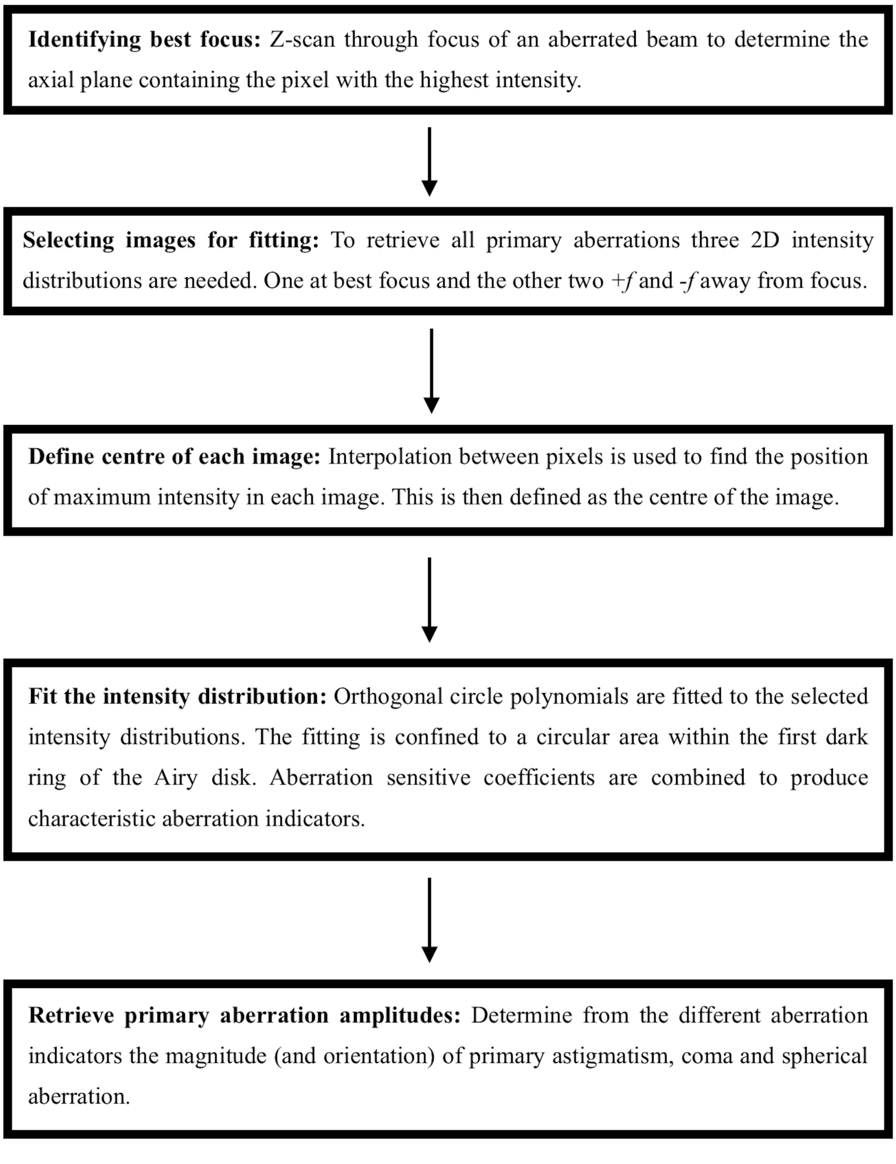


Figure 6.1: Outline of the aberration retrieval method based on fitting orthogonal circle polynomials to the intensity distributions of an aberrated beam in the vicinity of focus.

6.1 Intensity distribution distortions caused by aberrations

The aim of aberration retrieval in microscopy is to determine the wavefront in the pupil of the microscope objective (often expressed in terms of Zernike modes) by analysing the three dimensional intensity distribution of the focal spot. In chapter 2, two prominent aberration retrieval methods were presented, namely methods based on Gerchberg-Saxton algorithm or the ENZ theory. Here, a novel aberration retrieval method is proposed which detects distortions of the intensity distribution of an aberrated focal spot when going through focus. It is well known that wavefront aberrations manifest themselves in different characteristic intensity distributions at and at each side of focus. The intensity distributions of different primary aberrations at three axial planes are shown in Fig. 5.2. For the simulations, the intensity distributions were calculated using Fourier transforms and the Fraunhofer approximation [16]. Here, the pupil function, $P(\rho,\theta)$, is described as follows,

$$P(\rho,\theta) = P_0(\rho,\theta)e^{i2\pi\Phi(\rho,\theta)} \tag{6.1}$$

where P_0 is the amplitude of the electromagnetic field in the pupil plane, Φ is the phase distribution in the pupil and (ρ,θ) are cylindrical coordinates in the pupil plane. In the following simulations, it is assumed that the amplitude is constant over the entire pupil, thus $P_0 = 1$. The phase function Φ is expressed in terms of Zernike polynomials,

$$\Phi(\rho,\theta) = \sum_{n,m} \alpha_n^m Z_n^m(\rho,\theta) \tag{6.2}$$

where α_n^m represents a Zernike amplitude coefficient and Z_n^m a Zernike mode, n and m are the radial and azimuthal orders, respectively. The simulations were restricted to the three most commonly occurring Zernike aberrations in optical imaging systems, i.e. primary astigmatism, primary coma and primary spherical aberration. The intensity distributions at either side of focus were determined by adding the Zernike defocus term Z_2^0 to the pupil function. The amplitude in the image plane of the pupil, U, can be obtained via a 2D Fourier transform of the pupil function,

$$U = \mathscr{F}^{-1}\{P\} \tag{6.3}$$

where $\mathscr{F}^{-1}$ represents the inverse 2D Fourier transform. The intensity distribution is obtained by multiplying U with its complex conjugate.

$$I = U \cdot U^{*} \tag{6.4}$$

Intensity distributions of aberrated spots are shown below:

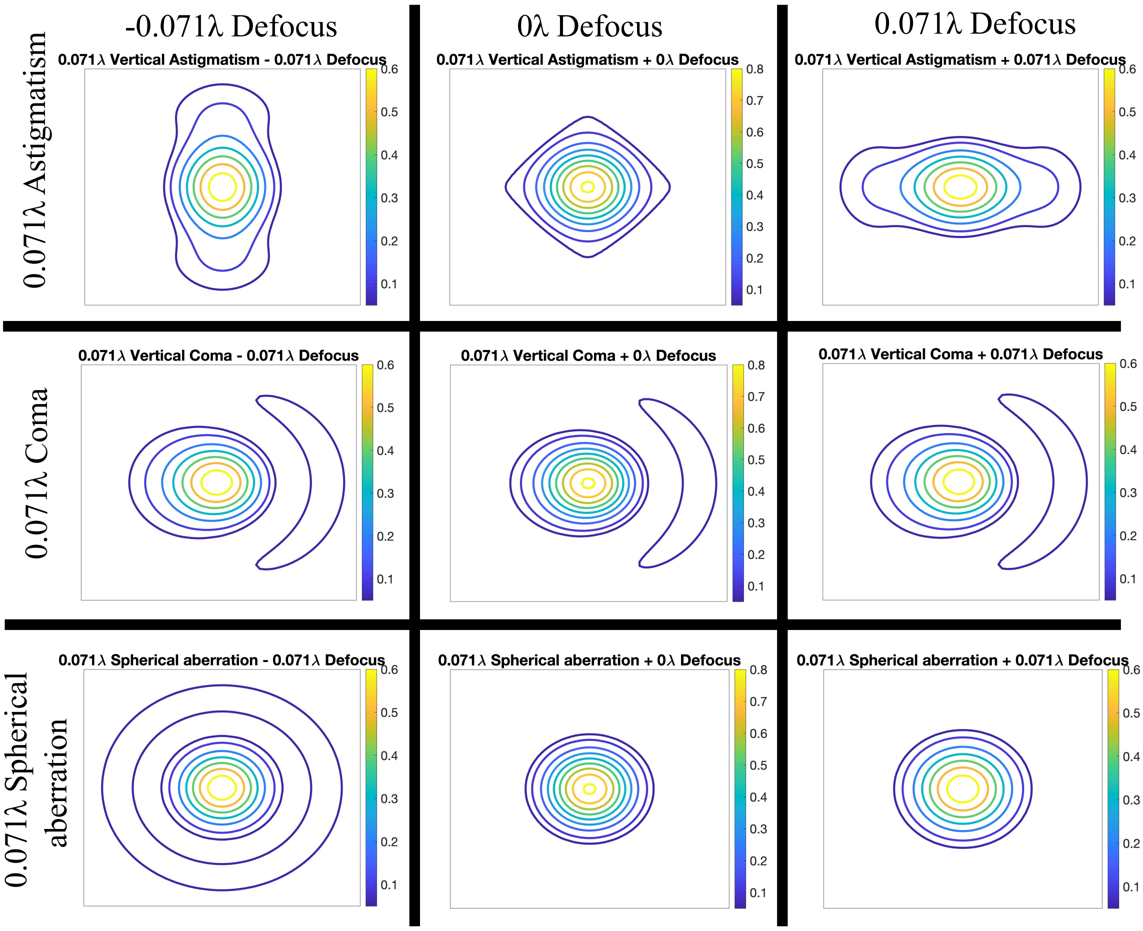


Figure 6.2: Aberrated spots at different axial planes. First row: Astigmatic spot through focus. Second row: Coma aberrated spot through focus. Third row: Spot with spherical aberration through focus

Nijboer derived mathematical expressions which describe the amplitude distribution at the image plane in the presence of aberrations and calculated the resulting intensity patterns in the vicinity of focus [9]. Based on Fig. 6.2 and Nijboer's work, characteristic features of the intensity distributions in the presence of spherical aberration, coma and astigmatism are listed below:

- Astigmatism: Out of focus, the intensity distributions is elongated along one direction. The direction of elongation rotates by 90 degrees when going through focus (first row, compare left and right images). At focus the intensity distribution has a 4 fold symmetry (see Fig. 6.2 top row central image).
- Coma: The intensity patterns are symmetric about focus as can be seen in Fig 6.2 (middle row, compare left and right image). However, the coma intensity distribution does not have rotational symmetry and has a tail which is oriented in the same direction in all axial planes (see fig 6.2 middle row).
- Spherical aberration: the intensity patterns before and after focus differ. On one side of focus, the central part of the intensity distribution is narrow and has a brighter outer intensity ring, whereas the spot broadens on the other side of focus (see Fig. 6.2 bottom row; compare left and right image).

6.2 Fitting of focal spot intensity distributions

A series of polynomials can be used to describe the central intensity distributions shown in Fig. 6.2. A set of orthogonal circle polynomials were fitted over a predefined circular area (in our case an area within the first dark ring of the ideal Airy spot):

$$I_{central}(r,\phi) \approx \sum_{n',m'} \gamma_{n'}^{m'} Z_{n'}^{m'}(r,\phi) \tag{6.5}$$

where γ_n^m are the coefficients of the orthogonal circle intensity polynomials $Z_{n'}^{m'}$ (which happen to be the Zernike polynomials) and r,ϕ are cylindrical coordinates in the image plane and the origin of the coordinate system lies at the point of maximum intensity. The intensity distributions were fitted in a least-square approach using singular value decomposition (SVD). Zernike polynomials are orthogonal over the unit circle and for our purpose they are better conditioned as a linear least square solver than standard polynomials. The method of fitting orthogonal circle polynomials to the central highest intensity part of the PSF in an axial plane is depicted in Fig. 6.3.

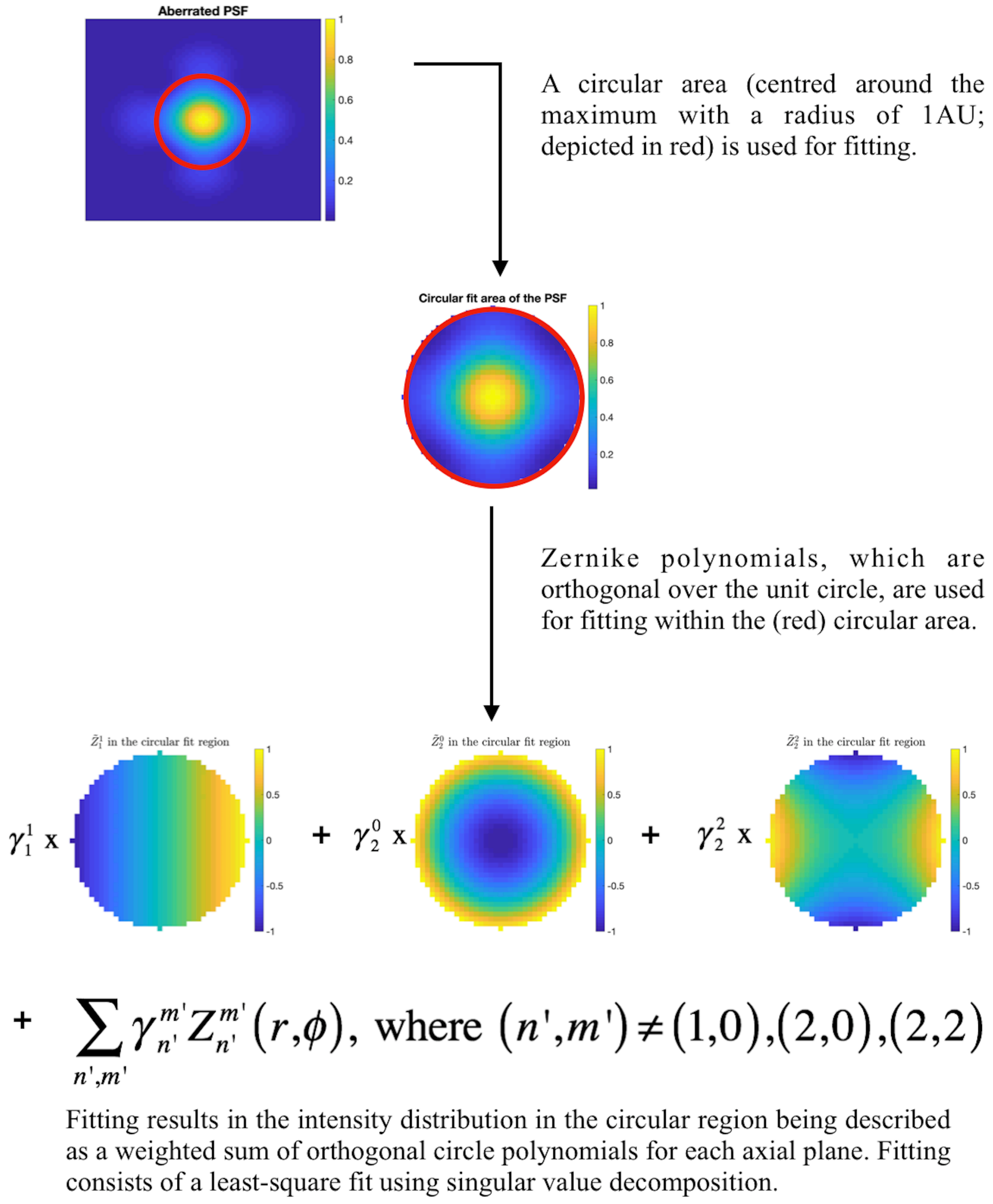


Figure 6.3: Flow chart - Fitting Zernike polynomials to the central part of the intensity PSF.

The relation between the γ_n^m coefficients in Eq. 6.5 and the α_n^m coefficients in Eq. 6.2 will be discussed later in Subsection 6.2.3 and Appendix B where analytical expressions for aberrated intensity distributions were derived using the ENZ theory.

6.2.1 Aberrations - intensity distributions through focus

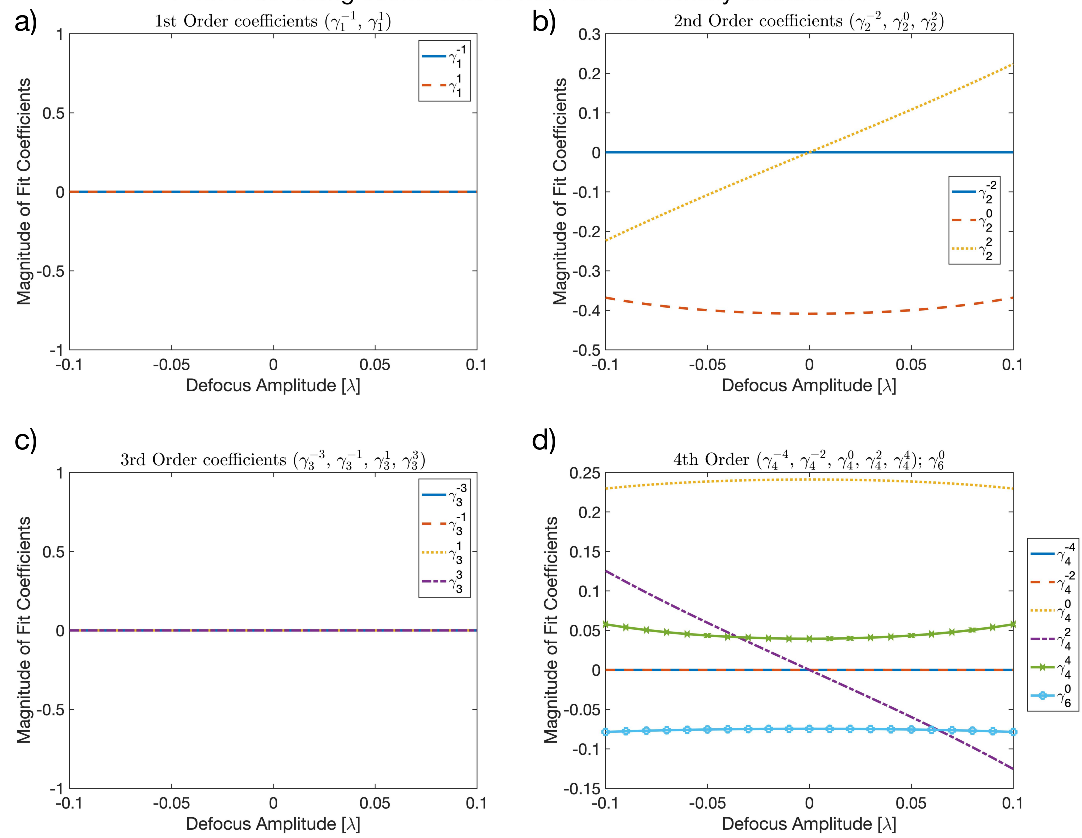
The intensity distribution at different axial planes was studied by varying the value of Zernike defocus $\left(\alpha_2^0\right)$ and setting a single primary aberration at a constant value. Defocus was varied in the range of [-0.1λ, 0.1λ]. Going back to chapter 2, the conversion for defocus parameter *f*, Zernike defocus amplitude α_2^0 and axial displacement Z is given by:

$$Z = \frac{4\alpha_2^0}{NA^2} = -\frac{\lambda f}{\pi NA^2} \qquad (6.6)$$

Formula 6.6 relates the axial displacement Z to the Zernike defocus amplitude α_2^0 and to the focal parameter *f*. Primary astigmatism, coma and spherical aberration were set to a constant value of 0.07λ. For each aberration, considered separately, the coefficients of the fitted circle intensity polynomials (in the image plane) are plotted as a function of defocus in Figure 6.4. In each axial plane, the intensity distributions were normalised with respect to the highest intensity value in that particular axial plane (if intensity measurements are performed with a CCD camera, then the camera can change the exposure time or amplification for each frame so as to avoid saturation). The intensity distributions were fitted using orthogonal circle polynomials up to radial order 4 (and also including the second order spherical orthogonal circle polynomial - radial order 6) using SVD. In each axial plane, the origin of the coordinate system (used for fitting) was placed at position of maximum intensity.

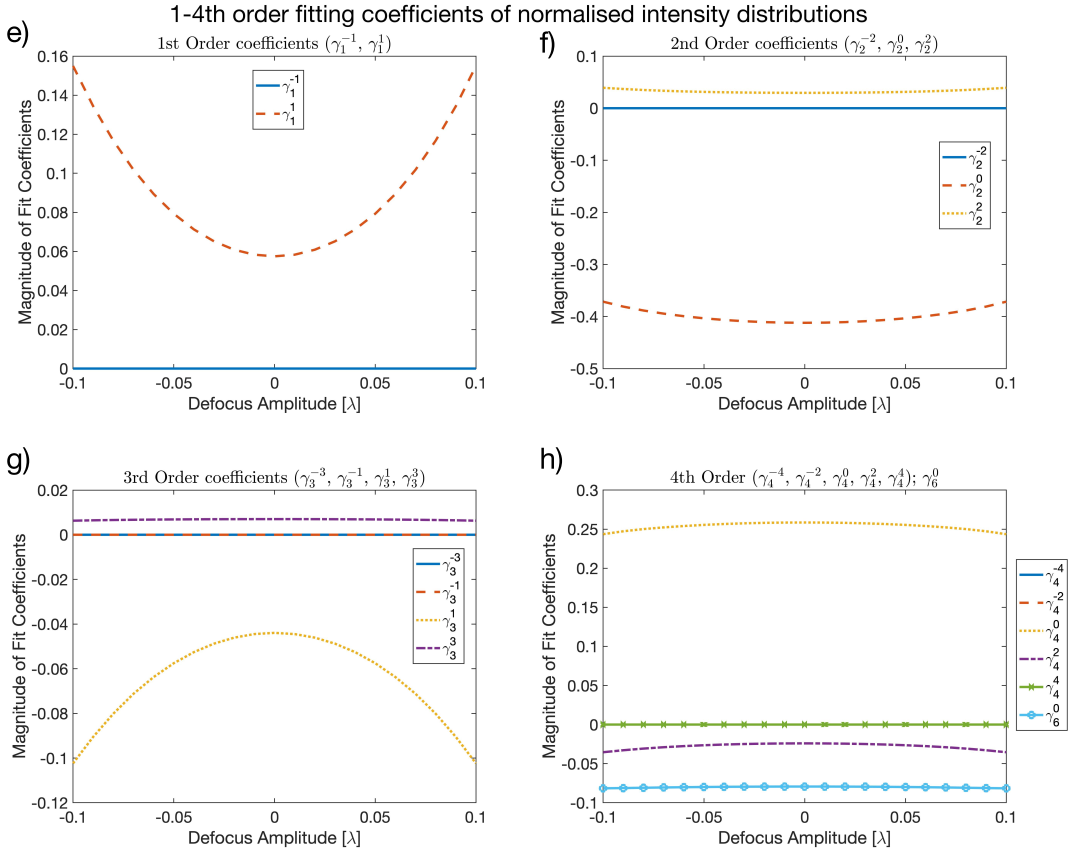
Constant Vertical Astigmatism

1-4th order fitting coefficients of normalised intensity distributions



Constant Vertical Coma

1-4th order fitting coefficients of normalised intensity distributions



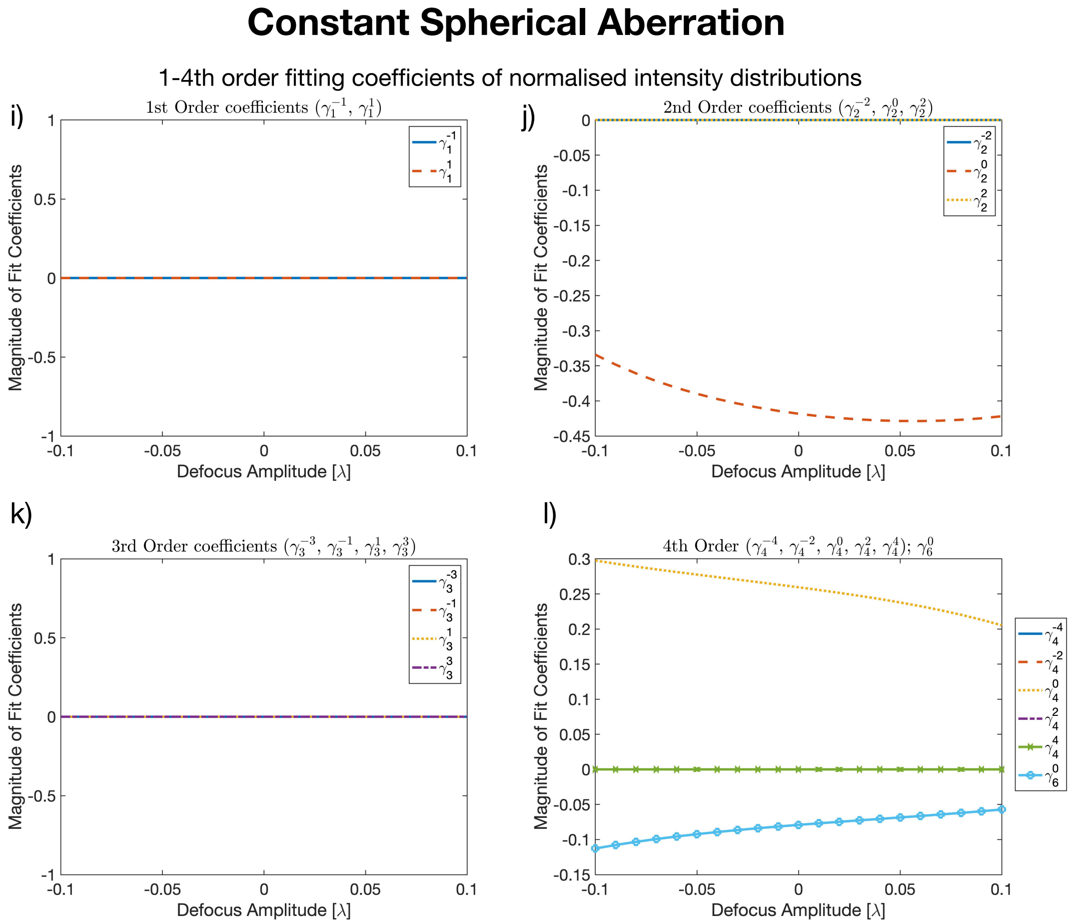


Figure 6.4: Fitting coefficients of Zernike polynomials. For a fixed amount (0.07λ) of vertical astigmatism, vertical coma and spherical aberration. The orthogonal circle polynomials are fitted around the central part of the intensity distribution using a circular mask with a diameter of 1 Airy units. The coefficients of the orthogonal circle polynomials are plotted for different amounts of defocus $\left(\alpha_2^0\right)$, the defocus range is $[-0.1\lambda, 0.1\lambda]$

The coefficients were plotted according to their radial order n. Only results for the vertical orientations of astigmatism and coma are plotted (oblique astigmatism and horizontal coma gave corresponding results).

When imposing vertical astigmatism the coefficients $\left(\gamma_1^{-1},\gamma_1^{1}\right)$ are zero (see Fig. 6.4a). The coefficient γ_2^2 is asymmetric with respect to best focus and varies linearly, while γ_2^{-2} is zero and γ_2^0 is symmetrical around best focus and changes slightly (see Fig. 6.4b). The third order coefficients, like the first order coefficients, are all zero (see Fig. 6.4c). The γ_4^0, γ_4^4, and γ_6^0 are symmetrical and more or less constant over the entire defocus range. γ_4^2 asymmetric with respect to best focus and varies linearly, while γ_4^{-2} is zero (see Fig. 6.4d). The change of sign of the coefficients $\left(\gamma_2^2,\gamma_4^2\right)$ through focus indicates a rotation of 90° of the intensity pattern.

For a coma wavefront aberration (see Fig. 6.4e, f), the $\left(\gamma_1^{-1},\gamma_1^1\right)$ and the $\left(\gamma_3^{-1},\gamma_3^1\right)$ coefficients, vary non-linearly through focus, as opposed to wavefront aberrations such as astigmatism or spherical aberration for which these terms are zero. The γ_4^0, γ_4^2, and γ_6^0 coefficients are symmetrical and more or less constant over the entire defocus range.

For a wavefront with spherical aberration, the γ_4^0 coefficient (see Fig. 6.4l) varies linearly with defocus and is not zero. In other words the intensity distribution is rotationally symmetric, but the diameter decreases at one side of best focus while it increases at the other side. The term γ_2^0 is not linear about focus (see Fig. 6.4j).

In the presence of coma, the fitting coefficients, through focus, are symmetrical but vary non-linearly (see Fig. 6.4 e-h). The resulting sign ambiguity (due to the symmetry about focus) and the non-linearity makes the coma aberration retrieval process more complicated than for astigmatism and spherical aberration. However, at best focus an approximately linear relationship exists between the coma amplitude and the fitting coefficients $\left(\gamma_1^1,\gamma_3^1\right)$ as shown in Fig, 6.5.

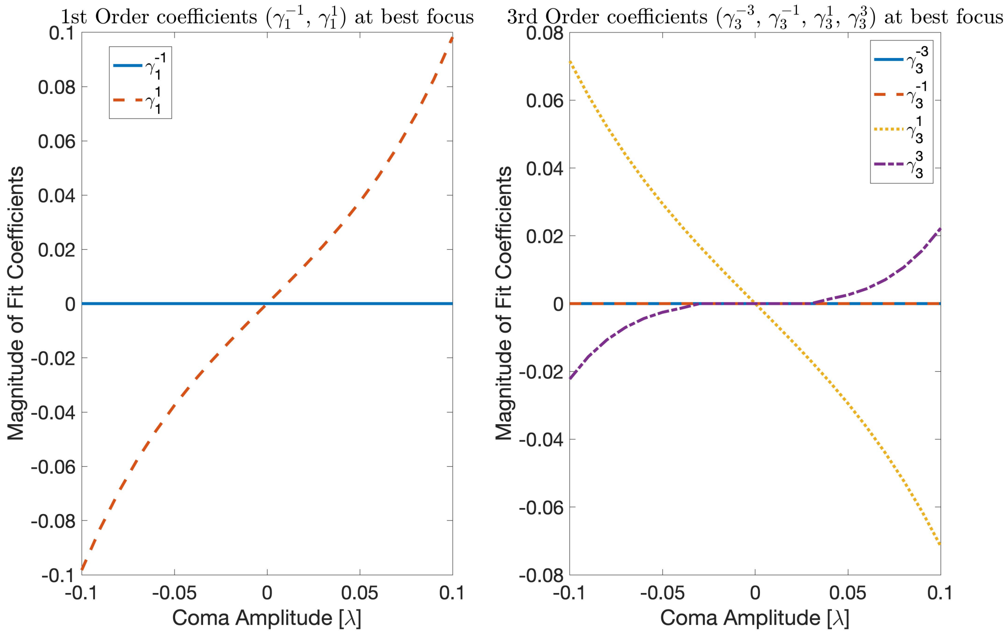


Figure 6.5: Fitting Coefficients of Zernike polynomials $\gamma_1^{-1},\gamma_1^{1},\gamma_3^{-3},\gamma_3^{-1},\gamma_3^{1},\gamma_3^{3}$. The orthogonal circle polynomials are fitted around the central part of the intensity distribution using a circular mask with a diameter of 1 Airy units. At the best focus plane the amplitude of vertical coma was varied in the range $[-0.1\lambda, 0.1\lambda]$

It should be noted that, because the fitting polynomials were fitted within the first dark ring of the Airy disk, features beyond the first dark ring are not detected. As can be seen from all figures (Fig. 6.4 and 6.5), the magnitudes of the lower order fitting coefficients are larger than the higher order coefficients by about an order of magnitude. Thus, the lower order fitting coefficients are more reliable and sensitive for aberration retrieval.

6.3 Aberration Indicators

In the previous section, it was shown by means of simulations that certain fitting coefficients are sensitive to primary aberrations and that they vary almost linearly with the aberration amplitude. A set of indicators, based on combinations of aberration sensitive fitting coefficients, suitable for aberration retrieval are presented in the following section. A reference axial plane f_0 is introduced in this subsection and represents an axial plane in the vicinity of best focus.

Astigmatism Indicator:

The intensity fitting coefficients $\left(\gamma_2^{\pm 2}, \gamma_4^{\pm 2}\right)$ are sensitive to astigmatism (see Fig. 6.3 vertical astigmatism). However, coma also affects to some extent $\left(\gamma_2^{\pm 2}, \gamma_4^{\pm 2}\right)$ (see Fig. 6.3 Vertical Coma) but the variations are symmetric about focus. So, if ones subtracts the $\gamma_2^{\pm 2}$ coefficients at two axial planes separated by the distance 2*f*, for example the two planes $(-f, f)$, any potential cross-talk caused by coma would be removed. Therefore an indicator for vertical astigmatism would be:

$$VA_{ind} = \left(\gamma_2^2\right)_{f_0+f} - \left(\gamma_2^2\right)_{f_0-f} \tag{6.7}$$

This indicator effectively detects the 90 degrees rotation of the intensity distribution when going through focus, which is characteristic for astigmatism. Analogously, an oblique astigmatism indicator would be:

$$OA_{ind} = \left(\gamma_2^{-2}\right)_{f_0+f} - \left(\gamma_2^{-2}\right)_{f_0-f} \tag{6.8}$$

Coma Indicator:

For a coma aberrated spot in the best focus plane both $\left(\gamma_1^{-1}, \gamma_1^1\right)$ and $\left(\gamma_3^{-1}, \gamma_3^1\right)$ vary linearly with $\alpha_3^{\pm 1}$. The coefficients $\left(\gamma_1^{-1}, \gamma_1^1\right)$ are better suited for coma retrieval because these coefficients are larger than the $\left(\gamma_3^{-1}, \gamma_3^1\right)$ coefficients. It should be noted that for wavefront aberrations such as spherical aberration and astigmatism the $\left(\gamma_1^{-1}, \gamma_1^1\right)$ coefficients are zero making these suitable indicator for coma. Thus to retrieve coma one can use the following indicator:

$$C_{ind} = \left(\gamma_1^{\pm 1}\right)_{f=0} \text{ or } \left(\gamma_3^{\pm 1}\right)_{f=0} \tag{6.9}$$

This indicator detects asymmetries in the intensity patterns

Spherical Aberration Indicator:

For a focal spot with spherical aberration, the coefficient γ_4^0 varies roughly linearly with α_4^0. The γ_2^0 coefficient is somewhat less linear (see Fig. 6.3 spherical aberration), but it can still be used to retrieve spherical aberration. Suitable spherical aberration retrieval indicators would be:

$$S_{ind} = \left(\gamma_2^0\right)_{f_0+f} - \left(\gamma_2^0\right)_{f_0-f} \text{ and } \left(\gamma_4^0\right)_{f_0+f} - \left(\gamma_4^0\right)_{f_0-f} \tag{6.10}$$

The spherical aberration indicator is sensitive to radial changes in the intensity distributions on opposite sides of focus.

The reference plane f_0 should lie in the vicinity of best focus in order to avoid fitting intensity distributions lying far from focus where the fitting coefficients vary non-linearly. Figure 6.6 illustrates how the vertical astigmatism, vertical coma and spherical aberration indicator would be evaluate based on intensity distributions at at $\left[f_0 - f, f_0, f_0 + f\right]$.

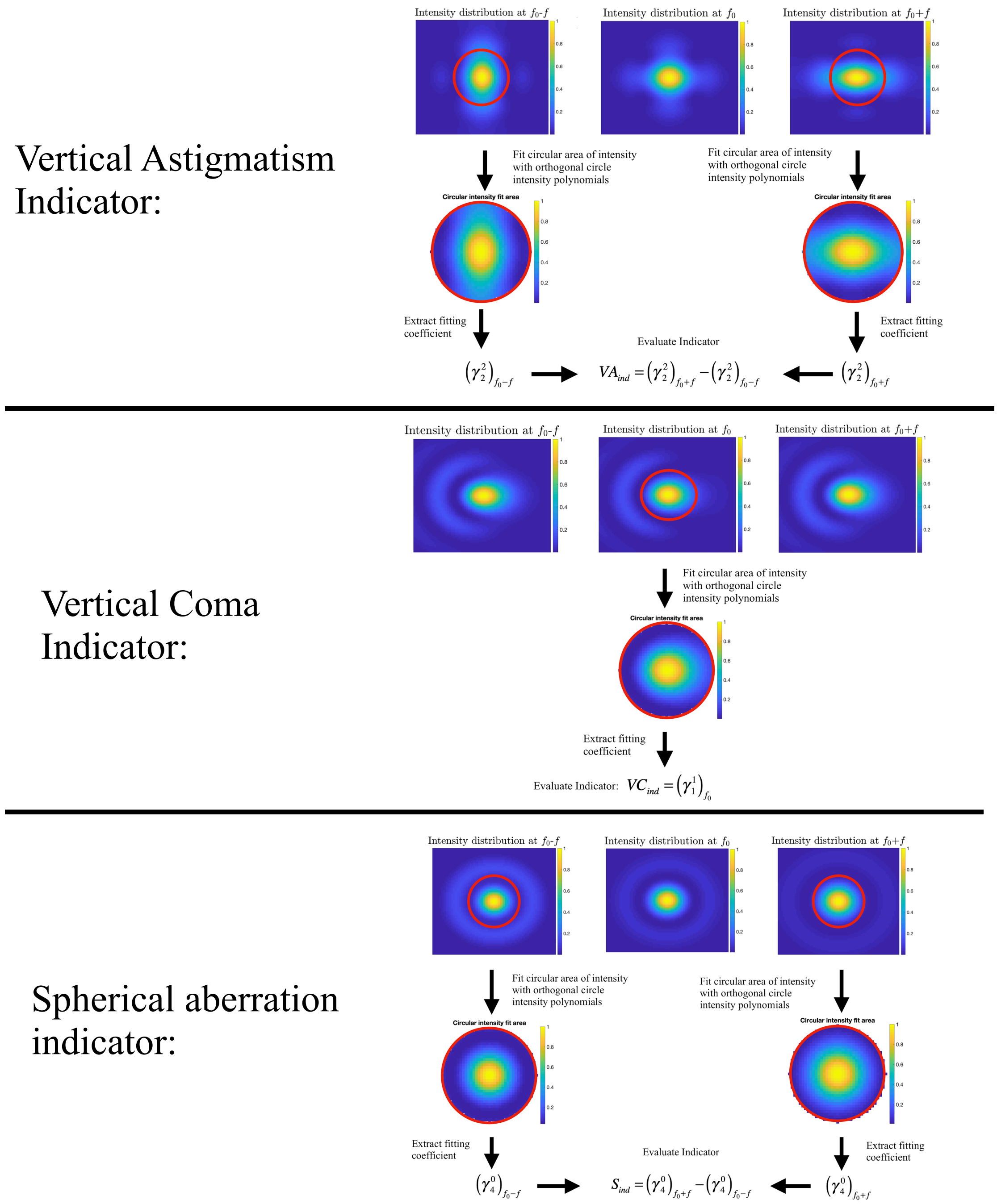


Figure 6.6: Illustration on how to evaluate the characteristic primary aberration indicators based on intensity distributions at different axial planes: $[f_0 - f, f_0, f_0 + f]$

The coma indicator, $(\gamma_1^{\pm 1})_{f=0}$ coefficients, varies almost linearly with the amplitude of coma $\alpha_3^{\pm 1}$ in the pupil, as was already shown in Fig. 6.5. The coefficients $(\gamma_2^{\pm 2})_f$ and $(\gamma_4^0)_f$ do also show an almost linear behaviour. In Fig. 6.7 the variations of the coefficients $\left(\gamma_2^2\right)_f$ and $\left(\gamma_4^0\right)_f$ at an axial plane $f = 1.07$ (or expressed in terms of Zernike $\alpha_2^0 = 0.05\lambda$ Malacara rms units) are shown when the respective primary aberration, namely astigmatism and spherical aberration, are varied in the range of $[-0.1\lambda,\ 0.1\lambda]$.

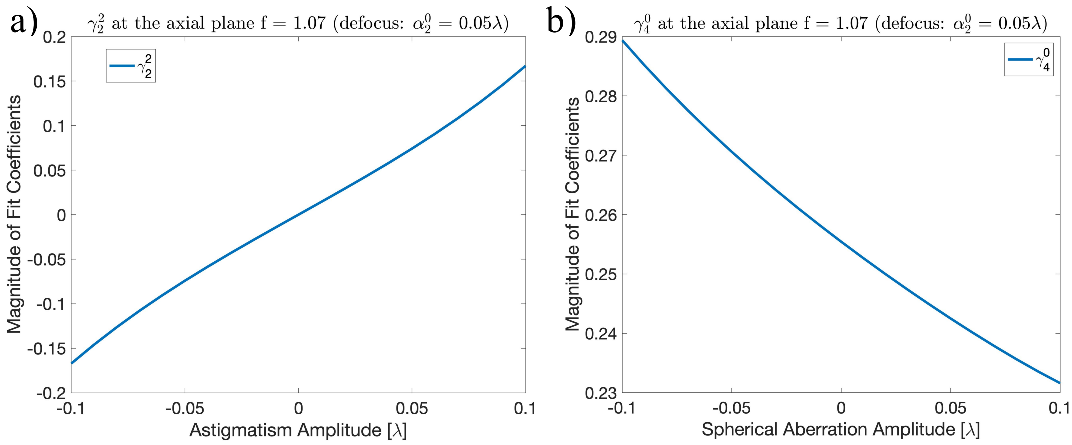


Figure 6.7: Fitting coefficients of Zernike polynomials a) γ_2^2 n the presence of vertical astigmatism and b) γ_4^0 n the presence of spherical aberration. The orthogonal circle polynomials are fitted around the central part of the intensity distribution using a circular mask with a diameter of 1AU. At the axial plane $f = 1.07$ (or expressed in terms of Zernike $\alpha_2^0 = 0.05\lambda$ Malacara rms units) the amplitude of a) vertical astigmatism, b) spherical aberration was varied in the range $[-0.1\lambda, 0.1\lambda]$

All the presented primary aberration indicators show an almost linear dependence with the their aberration amplitude in the pupil and are thus convenient for linear aberration retrieval.

The primary aberration trefoil is not very common in optical imaging systems and was therefore not studied further. However, this aberration could also be detected by looking at the $\gamma_3^{\pm 3}$ around best focus, similar to the coma indicator described (see Equation 6.9). Fig. 6.8 shows how the third order fitting coefficients vary at best focus in the presence of Z_3^3. The coefficient γ_3^3 increases almost linearly with increasing amount of Z_3^3. Thus, by using an indicator based on the $\gamma_3^{\pm 3}$ coefficients, one would be able to retrieve the amount of trefoil present in the pupil.

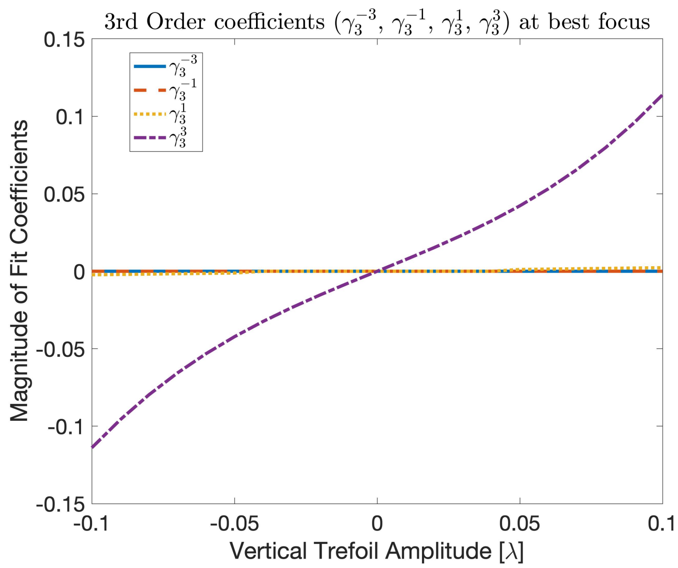


Figure 6.8: Fitting Coefficients $\gamma_3^{-3},\gamma_3^{-1},\gamma_3^{1},\gamma_3^{3}$ in the presence Z_3^3. The orthogonal circle polynomials are fitted around the central part of the intensity distribution using a circular mask with a diameter of 1 Airy units. At the best focus plane and the amplitude of vertical trefoil was varied in the range

6.3.1 Influence of defocus on the aberration indicators

Previously FFT simulations were used, for convenience, to simulate intensity distributions through focus in the presence of primary aberrations. The ENZ theory was used to obtain expressions which describe the intensity distribution for aberrated beams through focus. This was done to validate the approach proposed in this chapter, based on fitting orthogonal circle polynomials to the intensity distributions. Furthermore, from these expressions conclusions on the influence of defocus on the linearity of the aberrations indicators can be drawn as well as what happens to the coefficients when one changes the radius of the fit area.

For derivation of the intensity distributions expressions using the ENZ theory, please refer to Appendix B. Here the intensity expression with the most dominant terms including orthogonal circle polynomials polynomials (up to radial order 5) is given.

Astigmatism:

The intensity distribution of an astigmatic beam for small aberrations and close to best focus may be approximated by

$$I_{astig}\left(r,\phi;f,\alpha_2^2\right) \approx \begin{bmatrix} \left\{\left(0.228-0.013f^2+\ldots\right)-\left(0.012-\ldots\right)\left(\alpha_2^2\right)^2\right\}\mathrm{R}_0^0+ \\ \left\{\left(-0.419+0.041f^2+\ldots\right)+\left(0.078-0.008f^2-\ldots\right)\left(\alpha_2^2\right)^2\right\}\mathrm{R}_2^0(r)+ \\ \left\{\left(0.081-0.006f^2\right)\left(f\alpha_2^2\right)+\ldots\right\}\mathrm{R}_2^2(r)\cos 2\phi+ \\ \left\{\left(0.257-0.024f^2-\ldots\right)+\left(-0.057+0.005f^2+\ldots\right)\left(\alpha_2^2\right)^2\right\}\mathrm{R}_4^0(r)+ \\ \left\{\left(-0.048+0.003f^2\right)f\left(\alpha_2^2\right)+\ldots\right\}\mathrm{R}_4^2(r)\cos 2\phi+ \\ \left\{\left(0.033+\ldots\right)\left(\alpha_2^2\right)^2\right\}\mathrm{R}_4^4(r)\cos 4\phi \end{bmatrix} = \begin{bmatrix} \left(\gamma_0^0\right)\mathrm{R}_0^0+ \\ \left(\gamma_1^0\right)\mathrm{R}_2^0(r)+ \\ \left(\gamma_2^2\right)\mathrm{R}_2^2(r)\cos 2\phi+ \\ \left(\gamma_4^0\right)\mathrm{R}_4^0(r)+ \\ \left(\gamma_4^2\right)\mathrm{R}_4^2(r)\cos 2\phi+ \\ \left(\gamma_4^4\right)\mathrm{R}_4^4(r)\cos 4\phi \end{bmatrix}$$

(6.11)

Taking the coefficients γ_2^2 at two axial planes at a distance $2f$ apart, for example the two planes $(-f,f)$:

$$VA_{ind} = \left(\gamma_2^2\right)_f - \left(\gamma_2^2\right)_{-f} \approx \left\{0.081\left(1-0.074f^2\right)2f\left(\alpha_2^2\right)+\ldots\right\} = \left\{0.162\left(1-0.074f^2\right)f\left(\alpha_2^2\right)+\ldots\right\}$$

(6.12)

The $-0.074f^2$ term describes a deviation from a linear response. If one allows a 10% deviation from the linear response, one gets a close to linear range from about $+f$ to about $-f$ with $f \approx 1.16$ (or $f \approx 0.054\lambda$ when expressed in Malacara rms units):

$$f = \pm\sqrt{\frac{0.1}{0.074}} \approx 1.16 \Rightarrow \alpha_2^0 = \frac{1.16}{4\pi\sqrt{3}} \approx 0.054\lambda \qquad (6.13)$$

However it should be noted that the 10% deviation tolerance for *f* will be larger since the next higher order f term (which has been omitted) will reduce the deviation caused by the $-0.074f^2$ term because it has a different sign (positive).

<u>Coma:</u>

Using the ENZ theory, the intensity distribution in the presence of coma and close to best focus may be approximated (for small aberrations) by:

$$I_{coma}\left(r,\phi;\alpha_3^1,f\right)=\begin{bmatrix} \left\{\left(0.228-0.013f^2+\ldots\right)+\left(-0.025+0.003f^2-\ldots\right)\left(\alpha_3^1\right)^2\right\}\mathrm{R}_0^0+ \\ \left\{\left(0.040+0.010f^2+\ldots\right)\left(\alpha_3^1\right)+\ldots\right\}\mathrm{R}_1^1\left(r\right)\cos\phi+ \\ \left\{\left(-0.419+0.041f^2-\ldots\right)+\left(0.057-0.007f^2+\ldots\right)\left(\alpha_3^1\right)^2\right\}\mathrm{R}_2^0\left(r\right)+ \\ \left\{0.019-\left(6.59\cdot10^{-4}\right)f^2-\ldots\right\}\left(\alpha_3^1\right)^2\mathrm{R}_2^2\left(r\right)\cos2\phi+ \\ \left\{-\left(0.031+0.008f^2-\ldots\right)\left(\alpha_3^1\right)+\left(0.003+\left(3.72\cdot10^{-4}\right)f^2-\left(2.65\cdot10^{-5}\right)f^4\right)\left(\alpha_3^1\right)^3\right\}\mathrm{R}_3^1\left(r\right)\cos\phi+ \\ \left\{-0.001f^2+\left(6.1\cdot10^{-5}\right)f^4\right\}\left(\alpha_3^1\right)^3\mathrm{R}_3^3\left(r\right)\cos3\phi+ \\ \left\{-0.014-\left(2.21\cdot10^{-4}\right)f^2+\left(4.9\cdot10^{-5}\right)f^4\right\}\left(\alpha_3^1\right)^2\mathrm{R}_4^2\left(r\right)\cos2\phi+ \\ \left\{\left(0.257-0.024f^2-\left(8.07\cdot10^{-4}\right)f^4\right)+\left(-0.031+\ldots\right)\left(\alpha_3^1\right)^2\right\}\mathrm{R}_4^0\left(r\right)+ \\ \left\{\left(-0.015+0.001f^2-\ldots\right)\left(\alpha_3^1\right)+\ldots\right\}\mathrm{R}_5^1\left(r\right)\cos\phi \end{bmatrix}=\begin{bmatrix} \left(\gamma_0^0\right)\mathrm{R}_0^0+ \\ \left(\gamma_1^1\right)\mathrm{R}_1^1\left(r\right)\cos\phi+ \\ \left(\gamma_1^0\right)\mathrm{R}_2^0\left(r\right)+ \\ \left(\gamma_2^2\right)\mathrm{R}_2^2\left(r\right)\cos2\phi+ \\ \left(\gamma_3^1\right)\mathrm{R}_2^2\left(r\right)\cos\phi+ \\ \left(\gamma_3^3\right)\mathrm{R}_3^3\left(r\right)\cos3\phi+ \\ \left(\gamma_4^0\right)\mathrm{R}_4^0\left(r\right)+ \\ \left(\gamma_4^2\right)\mathrm{R}_4^2\left(r\right)\cos2\phi+ \\ \left(\gamma_3^1\right)\mathrm{R}_5^1\left(r\right)\cos\phi \end{bmatrix}$$

(6.14)

It should be noted that by looking at the derivation in the appendix B that equation B. 18-21, the intensity expression contains a third order term, i.e. $r^3\cos\theta$ but does not contain a linear radial term *r*. When using Zernike polynomials as the fitting functions a $R_3^1\left(r\right)$ polynomial has a 3rd order *r* term and also a 1st order *r* term. Therefore, the fitting will also generate a so called tilt term $R_1^1\left(r\right)$ to compensate for the *r* term contribution in the $R_3^1\left(r\right)$ polynomial. Hence, the magnitude of this tilt term is a measure of the amount of coma. Taking the coefficient γ_1^1 at $f=0$:

$$C_{ind}=\left(\gamma_1^1\right)_{f=0}\approx\left\{0.040\left(\alpha_3^1\right)+\ldots\right\} \tag{6.15}$$

Since the coma indicator is supposed to be applied in the best focal plane, one can estimate the deviation in *f* that is allowed by permitting a 10% deviation in the value of the coma indicator. In other words $0.010f^2<0.004$, resulting in: $f=\pm0.63$ (or $f\approx0.029\lambda$ when expressed in Malacara rms units).

$$f = \pm 0.63 \Rightarrow \alpha_2^0 = \frac{0.63}{4\pi\sqrt{3}} \approx 0.029\lambda \tag{6.16}$$

Spherical aberration:

The intensity distribution in the presence of spherical aberration for small aberrations and close to best focus may be approximated by:

$$I_{spherical}\left(r,\phi;f,\alpha_4^0\right) \approx \begin{bmatrix} \left\{\begin{matrix} \left(0.228-0.012f^2+\ldots\right)+ \\ \left(0.013-\left(8.55\cdot 10^{-4}\right)f^2+\ldots\right)f\left(\alpha_4^0\right)+ \\ \left(-0.0459+0.004f^2-\ldots\right)\left(\alpha_4^0\right)^2+ \\ \left(-0.001+\left(7.99\cdot 10^{-5}\right)f^2+\ldots\right)f\left(\alpha_4^0\right)^3 \end{matrix}\right\} R_0^0 + \\ \left\{\begin{matrix} \left(-0.419+0.0422f^2-\ldots\right)+ \\ \left(-0.017+\left(2.28\cdot 10^{-4}\right)f^2-\ldots\right)f\left(\alpha_4^0\right)+ \\ \left(0.085-0.010f^2+\ldots\right)\left(\alpha_4^0\right)^2+ \\ \left(+0.002-\left(3.18\cdot 10^{-5}\right)f^2-\ldots\right)f\left(\alpha_4^0\right)^3 \end{matrix}\right\} \mathrm{R}_2^0(r)+ \\ \left\{\begin{matrix} \left(0.258-0.023f^2+\ldots\right)+ \\ \left(-0.017+\left(7.28\cdot 10^{-4}\right)f^2-\ldots\right)f\left(\alpha_4^0\right)+ \\ \left(-0.049+0.006f^2-\ldots\right)\left(\alpha_4^0\right)^2+ \\ \left(0.002-\left(7.61\cdot 10^{-5}\right)f^2+\ldots\right)f\left(\alpha_4^0\right)^3 \end{matrix}\right\} \mathrm{R}_4^0(r) \end{bmatrix} = \begin{bmatrix} \left(\gamma_0^0\right)\mathrm{R}_0^0 + \\ \left(\gamma_1^0\right)\mathrm{R}_2^0(r)+ \\ \left(\gamma_4^0\right)\mathrm{R}_4^0(r)+ \end{bmatrix} \tag{6.17}$$

Taking the coefficient γ_4^0 at two axial planes at a distance 2*f* apart, for example the two planes $\left(-f,f\right)$:

$$S_{ind} = \left(\gamma_4^0\right)_f - \left(\gamma_4^0\right)_{-f} \approx -\left\{0.017\left(1-0.043f^2\right)2f\left(\alpha_4^0\right)+\ldots\right\} = \left\{0.034\left(1-0.043f^2\right)f\left(\alpha_4^0\right)+\ldots\right\} \tag{6.18}$$

The $-0.043f^2$ term describes the deviation from the linear response. If ones allows a 10% deviation from the linear response, one gets a close to linear range from about $+f$ to about $-f$ for $f \approx 1.53$ (or $f \approx 0.07\lambda$ when expressed in Malacara rms units):

$$f = \pm\sqrt{\frac{0.1}{0.043}} \approx 1.53 \Rightarrow \alpha_2^0 = \frac{1.53}{4\pi\sqrt{3}} \approx 0.07\lambda \tag{6.19}$$

Again it should be noted that the 10% deviation tolerance for f will be larger since the next higher order f term (which has been omitted) will reduce the deviation caused by the $-0.043f^2$ term because it has a different sign (positive).

All previously derived intensity equations (using the ENZ theory) were not normalised with respect to the maximum intensity at each focal plane (as was done for the FFT simulation results shown in Fig. 6.4 and Fig. 6.5). Normalising with respect to the maximum intensity is not trivial. The value of maximum intensity in a focal plane depends on the aberration coefficients and on the value of the focal parameter f. For small aberrations however one can determine the maximum intensity value by computing the on-axis intensity $I\left(0,0,f,\alpha_n^m\right)$ in each plane. It was found that especially for the astigmatism (equation 6.7) and spherical aberration (equation 6.10) indicators, normalising the intensity distributions is advantageous when the wavefront aberration increases because quadratic α_n^m terms (and also higher powers) would cause the indicators to behave non-linearly (see Appendix B). The aberration and focal parameter terms (in equation 6.12, 6.14 and 6.18) affect to some extent the linear indicator and tend to decrease its value when the wavefront aberration increases. This decrease in the indicator's value is partially compensated for by normalising the PSF because normalising increases the value of the fitted Zernike coefficient by $1/I\left(0,0,f,\alpha_n^m\right)$.

A comparison of the fitting coefficients found using the FFT simulations with the analytical values obtained with the ENZ theory (the chosen α coefficients values corresponds to 0.071λ in terms of Malacara rms wavefront deviation for the respective aberration) are shown in table 6.1. The derived equations (ENZ theory) were normalised by multiplying the obtained values with the on-axis intensity $1/I\left(0,0,f,\alpha_n^m\right)$ to compare them with the FFT results (the FFT's were normalised with respect to the maximum intensity at each focal plane, which is equivalent to an on-axis intensity normalisation when aberrations are small).

Table 6.1 - Fitting coefficient comparison between FFT fit and ENZ Theory

	FFT Fit Coma (α_3^1 = 1.26) (*f* = 0)	**ENZ Theory Coma (α_3^1 = 1.26) (*f* = 0)**	**FFT Fit Astigmatism (α_2^2 = 1.09) (*f* =1.55)**	**ENZ Theory Astigmatism (α_2^2 = 1.09) (*f* =1.55)**	**FFT Fit Spherical Aberration (α_4^0 = 1) (*f* =1.55)**	**ENZ Theory Spherical Aberration (α_4^0 = 1) (*f* =1.55)**
γ_0^0	0.235	0.237	0.282	0.286	0.269	0.261
γ_1^{-1}	0.000	0.000	0.000	0.000	0.000	0.000
γ_1^1	0.059	0.061	0.000	0.000	0.000	0.000
γ_2^{-2}	0.000	0.000	0.000	0.000	0.000	0.000
γ_2^0	-0.412	-0.410	-0.390	-0.402	-0.428	-0.432
γ_2^2	0.030	0.035	0.155	0.158	0.000	0.000
γ_3^{-3}	0.000	0.000	0.000	0.000	0.000	0.000
γ_3^{-1}	0.000	0.000	0.000	0.000	0.000	0.000
γ_3^1	-0.045	-0.041	0.000	0.000	0.000	0.000
γ_3^3	0.007	0.000	0.000	0.000	0.000	0.000
γ_4^{-4}	0.000	0.000	0.000	0.000	0.000	0.000
γ_4^{-2}	0.000	0.000	0.000	0.000	0.000	0.000
γ_4^0	0.259	0.260	0.236	0.233	0.226	0.229
γ_4^2	-0.025	-0.024	-0.086	-0.107	0.000	0.000
γ_4^4	0.000	0.001	0.049	0.051	0.000	0.000
γ_6^0	-0.079	-0.078	-0.076	-0.078	-0.064	-0.065

The values found with the ENZ theory correspond well with the FFT results (differences in general less than 10%), except the values of γ_4^2 in case of astigmatism and of γ_2^2 in case of 6.4.

Size of the fitting region

In experiments one might be restricted to fit the intensity distribution using circular masks smaller than 1AU because of the size and limited numbers of available pixels for measuring. To assess the influence of the fit area on the aberration indicators, so far the intensity distribution fitting was restricted to a circular area of 1AU in diameter. Reducing the size of the fit area affects the coefficients of the fitted orthogonal circle polynomials. These changes in the fitting coefficients are in fact the same as scaling down Zernike coefficients to smaller pupil sizes, which is described by Dai formula [103]. For demonstration, a comparison will be made between the coefficient values computed with Dai's formula and the ones obtained with the ENZ theory (equation 6.11, 6.14, 6.17) by substituting for a scaled radial coordinate $r' = \varepsilon(3.832r)$, where $0 \le \varepsilon \le 1$ is the pupil scaling factor. The value 3.832 represents the radial distance at which the first oder Bessel function has its first zero point (which corresponds to 1AU). The intensity equations, obtained with using the ENZ theory, were normalised by multiplying with the on-axis intensity $1\big/I\left(0,0,f,\alpha_n^m\right)$.

Dai's formula for non-normalised Zernike polynomials (it is reiterated here that orthogonal circle intensity polynomials are not normalised according to the Malacara normalisation) is given by [103] and the re-scaled coefficients $\tilde{\gamma}_n^m$ are:

$$\tilde{\gamma}_n^m = \varepsilon^n \left[\gamma_n^m + (n+1) \sum_{i=1}^{(N-n)/2} \gamma_{n+2i}^m \sum_{j=0}^{i} \frac{(-1)^{i+j}(n+i+j)!}{(n+j+1)!(i-j)!j!} \varepsilon^{2j} \right] \tag{6.20}$$

where N is the total number of radial orders used for the expansion. As can be seen from Equation 6.20, only higher order coefficients (with the azimuthal order m) affect

$\tilde{\gamma}_n^m$. In fig 6.9, results from the ENZ theory and Dai's formula for the spherical aberration indicators (see Equation 6.17) are shown. For Dai's formula the Zernike coefficients γ_n^m obtained with the ENZ theory for $\left(\alpha_4^0 = 1, f = \pm 1.55\right)$, $N = 8$ and $\varepsilon = 1$ were used to calculate the effects of re-scaling the coefficients to accommodate smaller pupil sizes:

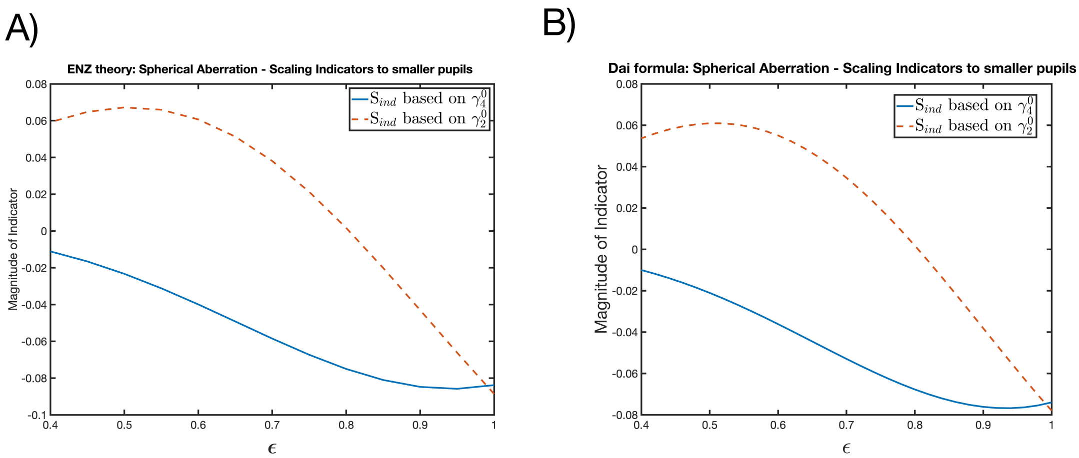


Figure 6.9: Scaling the two spherical aberration indicators (see equation 6.10) to a smaller pupil size using A) Extended Nijboer-Zernike (ENZ) theory, B) Dai's formula. The pupil were re-scaled by changing the scaling factor ε from 1 to 0.4. Pupil parameters are $\left(\alpha_4^0 = 1, f = \pm 1.55\right)$

Re-scaling can thus conveniently be computed using Dai's formula. Figure 6.9 (A or B) show that the indicator based on γ_2^0 for spherical aberration changes sign within the interval $\varepsilon \in [0.4, 1]$ and that it is zero at a value of about $\varepsilon = 0.8$. The indicator based on γ_2^0 can be used for retrieval of spherical for ε values in the range of either [0.4, 0.6] or for values somewhat less than 1. At ε values around 0.8, the indicator based on γ_2^0 is not suitable for indicating the presence of spherical aberration. In comparison the indicator based on γ_4^0 is most sensitive at ε values between [0.8, 1] but its magnitude continuously decreases with smaller values ε. The effects of decreasing the fit area on the indicators of coma $\left(\alpha_3^1 = 1.26, f = 0\right)$ and astigmatism $\left(\alpha_2^2 = 1.09, f = \pm 1.55\right)$ are shown in Fig. 6.10. It is noted that the two coma indicators γ_1^1 or γ_3^1 are measured at best focus (see Equation 5.13).

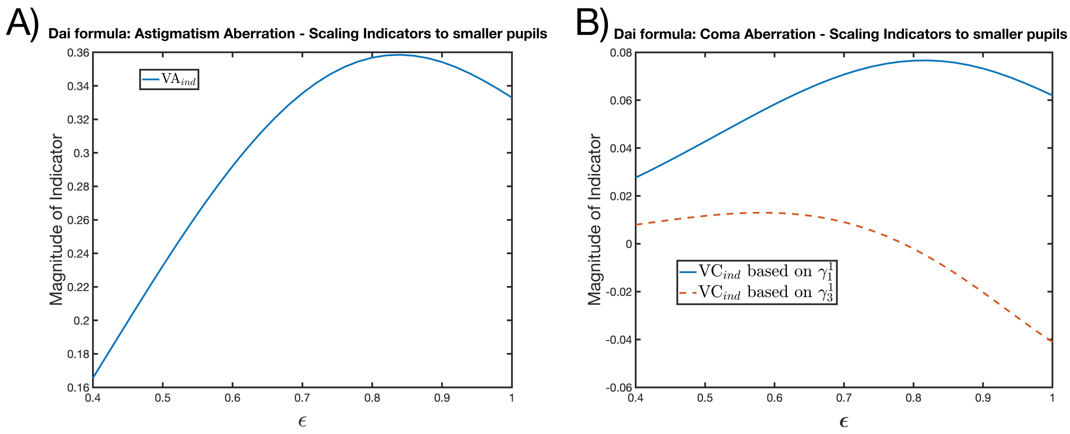


Figure 6.10: Scaling A) astigmatism and B) coma aberration indicators (see equation 6.7 and 6.9, respectively) to a smaller pupil size using Dai's formula. The pupil were re-scaled by varying the scaling factor ε from 1 to 0.4. Pupil parameters are $\left(\alpha_2^2 = 1.09, f = \pm 1.55\right)$ for astigmatism and $\left(\alpha_3^1 = 1.26, f = 0\right)$ for coma.

The indicator for astigmatism (see equation 6.7) is best evaluated within the range for ε = [0.7, 1]. For coma γ_1^1 terms are better suited for aberration retrieval than γ_3^1 terms, because γ_1^1 terms do not change sign over the entire range and the γ_1^1 indicator is never zero. The γ_1^1 tilt term is best measured in the range of ε = [0.7, 1]. From Figure 6.9 and 6.10, one can conclude that a circular fit region between 0.7-1 AU is best suited and most sensitive to retrieve spherical aberration, coma and astigmatism.

6.5 Experimental Validation

6.5.1 Aberration retrieval on experimental data

To validate the aberration retrieval method, an optical setup was built. The setup is illustrated in Fig. 6.11.

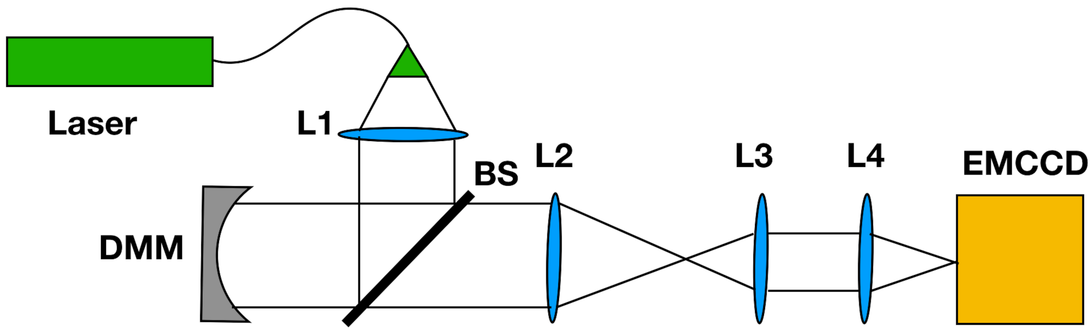


Figure 6.11: Experimental Setup for aberration retrieval. L: Lens; BS: Beamsplitter; BS: Beamsplitter plate; DMM: Deformable Membrane Mirror; EMCCD: Electron-Multiplying CCD.

Laser light (wavelength of 532 nm), injected in a single mode fibre, provides illumination for the experiments. Aberrations were applied with a Mirao 52e (Imagine Optics, France) DMM to light exitted by the fibre and collimated by lens L1. The DMM was calibrated in closed-loop mode as described in chapter 3 (section 3.1.2). Voltage combinations necessary to produce the first 15 Zernike modes were saved and used later on for open-loop operation of the DMM. The DMM is conjugated with the pupil of lens L4 (500mm) via a lens pair, L2 (100mm) and L3 (50mm) which reduced the beam diameter from about 10mm, at the DMM to about 5mm at L4. L4 focuses the laser beam onto the electron multiplying charged coupled camera (EMCCD; iXon 885 Andor).

The Zernike aberrations: primary astigmatism; primary coma; and spherical aberration were applied independently from each other with the DMM. Their amplitude was varied over a range of ~[-0.13λ 0.13λ] in steps of about ~0.027λ. Defocus was applied applied using the DMM by changing the value of the Zernike

Z_2^0 coefficient. Defocus was varied over a range of ~[-0.1λ 0.1λ] in steps of about ~0.025λ. The CCD sensor auto-exposure setting was used at each frame so that the maximum intensity for each frame was about 90% of the saturation value and the camera was cooled down to -50° to minimise thermal noise. The diameter of the first dark ring of the focal spot was measured to be about 32 pixels on the camera.

The intensity distribution in the different axial planes were fitted using orthogonal circle polynomials (Zernike polynomials) over a fixed circular area, centred around the pixel with maximum intensity. The fitting area of the focal spot was set by using a circular mask in the same way as described in chapter 3.

Under small aberration assumptions the position of maximum intensity can be determined by interpolation with a second order 2D polynomial fit around the pixel with the highest intensity in any axial plane. This will be further referred to as the "maximum intensity fit", which differs from the "intensity distribution fit" which involves fitting orthogonal circle polynomials. The position of maximum will be defined as the new centre of each image. When using FFT simulations, (m,n) represents the matrix element with the highest calculated intensity, while for a CCD, (m,n) is the pixel with highest intensity. Only second order polynomials were used, because contributions of higher order terms would be small around a maximum (the fit was in 2D area having a width smaller than 0.25AU). Hence, the following fit function, was used to estimate (x_0,y_0):

$$\mathrm{I(x,y) = fitresult(x,y) = p_{00} + p_{10}x + p_{01}y + p_{20}x^2 + p_{11}xy + p_{02}y^2} \tag{6.21}$$

That (x_0,y_0) is an extremum means that ∂I/∂x=0; and ∂I/∂y=0 when $(x,y)=(x_0,y_0)$. Thus

$$\begin{cases} \mathrm{p_{10} + 2p_{20}x_0 + p_{11}y_0 = 0} \\ \mathrm{p_{01} + 2p_{02}y_0 + p_{11}x_0 = 0} \end{cases} \tag{6.22}$$

Which may also be written in matrix form as:

$$\begin{pmatrix} p_{10} \\ p_{01} \end{pmatrix} + \begin{pmatrix} 2p_{20} & p_{11} \\ p_{11} & 2p_{02} \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = 0 \tag{6.23}$$

If the determinant of the matrix, which is $4p_{20} \cdot p_{02} - p_{11}^2$, is ≠ 0, then the solution, which involves matrix inversion, is

$$\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \frac{1}{p_{11}^2 - 4p_{20}p_{02}} \begin{pmatrix} 2p_{02} & -p_{11} \\ -p_{11} & 2p_{20} \end{pmatrix} \begin{pmatrix} p_{10} \\ p_{01} \end{pmatrix} = \frac{1}{p_{11}^2 - 4p_{20}p_{02}} \begin{pmatrix} 2p_{02}p_{10} - p_{11}p_{01} \\ 2p_{20}p_{01} - p_{11}p_{10} \end{pmatrix} \tag{6.24}$$

The fitting range is then shifted accordingly, provided that the maximum position would be shifted by more than 10% of the distance between two pixels (only significant changes should be taken into account):

$$\begin{cases} [-x_{\max}, x_{\max}] - x_0 \\ [-y_{\max}, y_{\max}] - y_0 \end{cases} \tag{6.25}$$

The parabola approximation used only holds in the vicinity of a maximum. For an aberration such as coma, the intensity distribution around a maximum may be better described by including higher order polynomials. However, finding the position of maxima in such circumstances is more complicated and a more viable method would require more complex algorithms which are beyond the scope of this thesis. To fit a parabola through data points, one needs a minimum of three data points. The parabola should be fitted in a region < ~0.3AU (range where the intensity can be well described with a parabola). Therefore, the diameter of the first dark ring of the Airy Disk should be at least 9 pixels.

For each individual aberration, the different indicators $VA_{ind}, OA_{ind}, C_{ind}, S_{ind}$ were plotted in Fig. 6.12-6.15 to show to what extent the characteristic indicator is sensitive to its particular aberration, while the others are hardly changing.

For coma only one orientation (vertical coma) is plotted (similar results were obtained for the horizontal orientation). For C_{ind} the average value of the coma indicators, measured at the three defocus values $\alpha_2^0 = (-0.025\lambda,\ 0,\ 0.025\lambda)$, was taken. Over such a small defocus range, the coma indicator will not change significantly (see Fig. 6.5, vertical coma: $\gamma_1^{-1}, \gamma_1^1, \gamma_3^{-1}, \gamma_3^1$ terms). Only values within the circular mask were used for the Zernike fit. It is reiterated here that the orthogonal circle polynomials are defined over a radius ranging from zero to one only. Thus, the radial orthogonal circle polynomials were normalised such that $r_{\max} = 1$. A circular mask with a diameter of 26 pixels was used for fitting (which corresponds to about 0.81AU). The position of maximum intensity was determined with a maximum fit using a square 0.21AU mask. Table 2.1 summarises the most important experimental parameters used for the aberration retrieval experiments.

Table 6.2 - Experimental parameters

Aberratio n	**Indicator**	**Number of axial planes necessary**	**Diameter of circular fit area (in AU)**	**Width of square maximum fit area (in AU)**	**Camera auto-exposure**
Z_2^2	VA_{ind}	2	0.81	0.21	Yes
Z_2^{-2}	OA_{ind}	2	0.81	0.21	Yes
$Z_3^{\pm 1}$	C_{ind}	at least 1	0.81	0.21	Yes
Z_4^0	S_{ind}	2	0.81	0.21	Yes

6.5.2 Experimental data on the coma indicator

The following plots show experimental data obtained with the optical setup shown in Fig. 6.11. The single aberration experiment whereby coma was applied is shown in Fig. 6.12. For each indicator, the y-axis limit was chosen to correspond to the highest measured indicator value for that particular aberration (thus ±0.2 for C_{ind} and S_{ind}; and ±1 for VA_{ind} and OA_{ind}; see Fig 6.12a-15a for the highest measured indicator values).

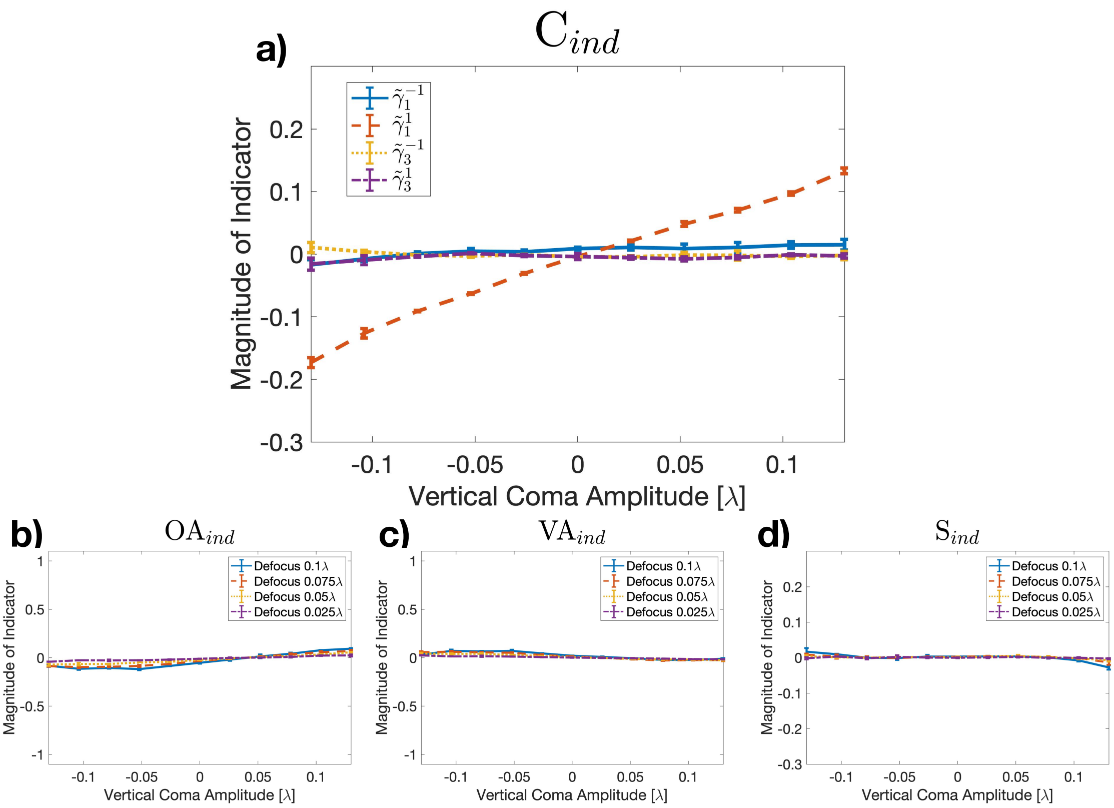


Figure 6.12: Experimental data. Primary aberration indicators for varying amounts of vertical coma. The coma indicator is measured at best focus. The orthogonal circle polynomials are fitted in the central part of the intensity distribution in a circular area with a diameter of about 0.81AU. The three figures at the bottom show the primary vertical astigmatism, primary oblique astigmatism and primary spherical aberration indicators. The γ_4^0 term was used for the S_{ind} and the γ_1^1 term was used for the C_{ind}. The coma wavefront aberration amplitude was varied over the range of $[-0.13\lambda, 0.13\lambda]$ whereas defocus varied over a range of $[-0.1\lambda, 0.1\lambda]$.

In Fig. 6.12a the coma indicator γ_1^1 shows an almost linear behaviour around zero but starts to become non-linear for larger aberration ($>0.11\lambda$, see in particular the left hand side of the graph). The other aberration indicators (see Fig. 6.12 b-d) are barely affected by the presence of a coma wavefront aberration.

6.5.3 Experimental data on the vertical astigmatism indicator

The single aberration experiment where vertical astigmatism was applied is shown in Fig. 6.13.

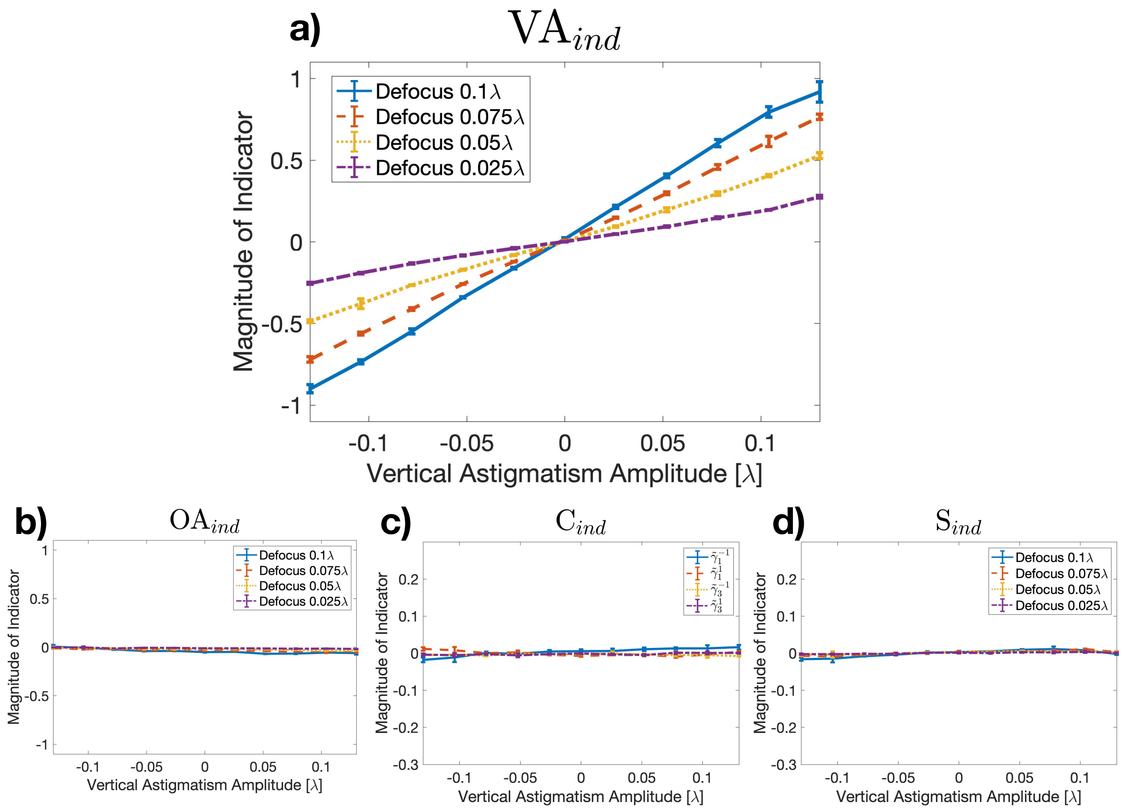


Figure 6.13: Experimental data. Primary aberration indicators for varying amounts of vertical astigmatism. The vertical astigmatism indicator is calculated at paired focal planes $(-f, f)$. The orthogonal circle polynomials are fitted in the central part of the intensity distribution in a circular area with a diameter of about 0.81AU. The three figures at the bottom show $OA_{ind}, C_{ind}, S_{ind}$. The γ_4^0 term was used for the S_{ind} and the γ_1^1 term was used for the C_{ind}. The vertical astigmatism wavefront aberration amplitude was varied in the range of $[-0.13\lambda, 0.13\lambda]$ whereas defocus varies in the range of $[-0.1\lambda, 0.1\lambda]$

The vertical astigmatism indicator shows an approximately linear behaviour around zero and becomes less linear for larger defocus ($>0.07\lambda$) and for larger values of vertical astigmatism $(>0.13\lambda)$ values (see Fig. 6.13a). To estimate the error of VA_{ind} at larger defocus, a linear curve was fitted through the data points in the aberration amplitude range of [0.075λ, 0.075λ] at 0.1λ defocus. The equation of the fitted curve $(R^2 > 0.98)$ was used to calculate by means of extrapolation the VA_{ind} values at

±0.125λ which were then compared with the measured values. The difference (or error) in the VA_{ind} values at ±0.125λ is ~9%. Doing the same estimation at a defocus value of 0.05λ (recommended value, see section 6.2.3) results in an error < 5%. The other aberration indicators are hardly affected by the presence of vertical astigmatism (see Fig. 6.13 b-d).

6.5.4 Experimental data on the oblique astigmatism indicator

The single aberration experiment where oblique astigmatism was applied is shown in Fig. 6.14.

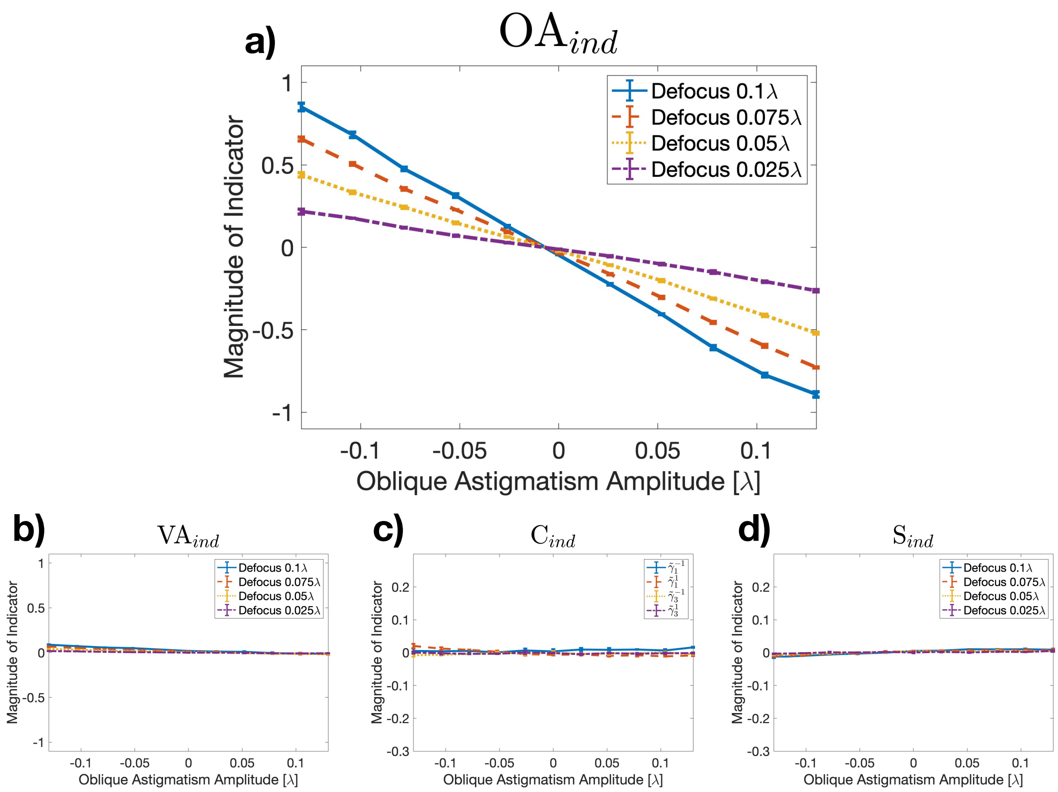


Figure 6.14: Experimental data. Primary aberration indicators for varying amounts of oblique astigmatism. The oblique astigmatism indicator is calculated at paired focal planes $(-f, f)$. The orthogonal circle polynomials are fitted in the central part of the intensity distribution in a circular area with a diameter of about 0.81AU. The three figures at the bottom show the primary vertical astigmatism, primary oblique astigmatism and primary spherical aberration indicators. The γ_4^0 term was used for the S_{ind} and the γ_1^1 term was used for the C_{ind}. The oblique astigmatism wavefront aberration amplitude was varied in the range of $[-0.13\lambda, 0.13\lambda]$ whereas defocus varies in the range of $[-0.1\lambda, 0.1\lambda]$

Similar to vertical astigmatism, the oblique astigmatism indicator shows an approximately linear behaviour (see Fig. 6.14a) around zero but which becomes less linear for larger defocus values $(>0.07\lambda)$ and for larger values of oblique astigmatism $(>0.13\lambda)$. To estimate the error of OA_{ind} at larger defocus, a linear curve was fitted through the data points in the aberration amplitude range of [0.075λ, 0.075λ] at 0.1λ defocus. The equation of the fitted curve $\left(R^2 > 0.98\right)$ was used to calculate by means of extrapolation the OA_{ind} values at ±0.125λ which were then compared with the measured values. The difference (or error) in the OA_{ind} values at ±0.125λ is ~12%. Doing the same estimation at a defocus value of 0.05λ (recommended value, see section 6.2.3) results in an error < 9%. The other aberration indicators are barely affected by the presence of oblique astigmatism (see Fig. 6.14 b-d).

6.5.5 Experimental data on the spherical aberration indicator

The single aberration experiment where spherical aberration was applied is shown in Fig. 6.15.

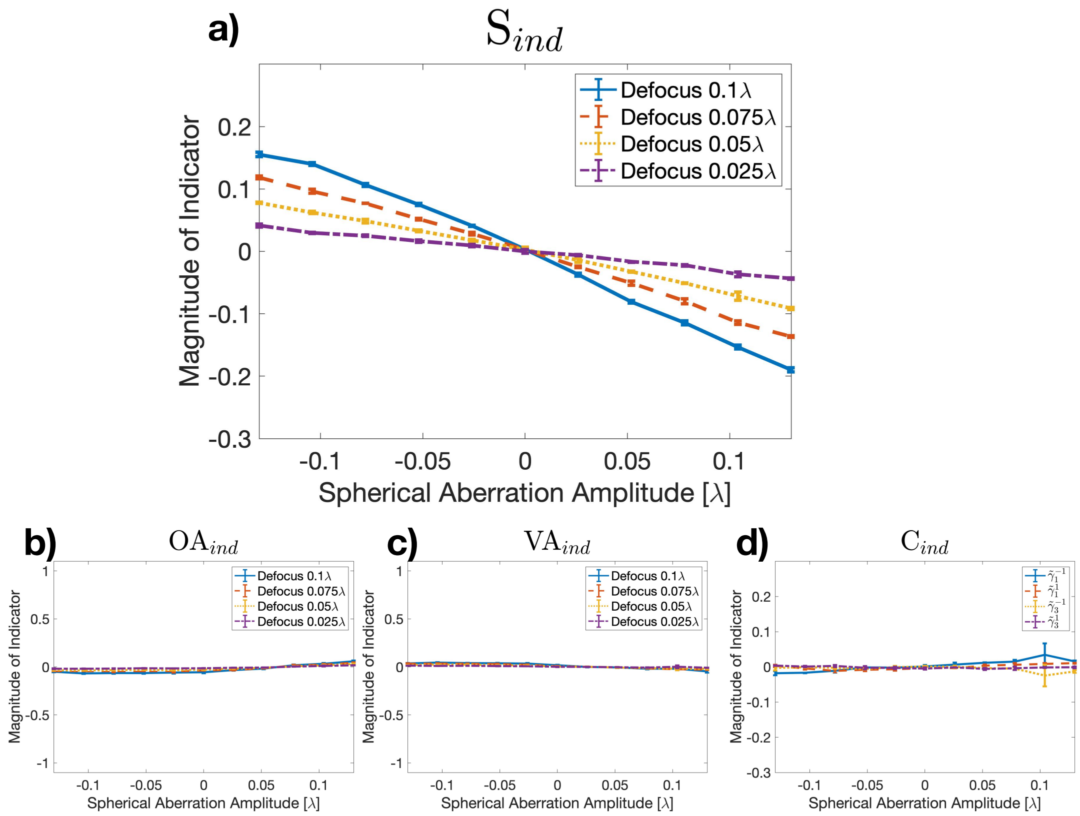


Figure 6.15: Experimental data. Primary aberration indicators for varying amounts of spherical aberration. The spherical aberration indicator is calculated at paired focal planes $(-f, f)$. The orthogonal circle polynomials are fitted in the central part of the intensity distribution in a circular area with a diameter of about 0.81AU. The three figures at the bottom show the primary vertical astigmatism, primary oblique astigmatism and primary spherical aberration indicators. The γ_4^0 term was used for the S_{ind} and the γ_1^1 term was used for the C_{ind}. The spherical wavefront aberration amplitude was varied in the range of $[-0.13\lambda, 0.13\lambda]$ whereas defocus varies in the range of $[-0.1\lambda, 0.1\lambda]$

As can be seen from Fig. 6.15a the spherical aberration indicator also shows an approximately linear behaviour when the amount of spherical wavefront aberration increases. For larger defocus ($\alpha_2^0 > 0.08\lambda$) and for larger values of spherical aberration $(> 0.13\lambda)$ the indicator is no longer linear. To estimate the error of S_{ind} at larger defocus values, a linear curve was fitted through the data points in the aberration amplitude range of [0.075λ, 0.075λ] at 0.1λ defocus. The equation of the fitted curve $\left(R^2 > 0.98\right)$ was used to calculate by means of extrapolation the S_{ind} values at ±0.125λ which were compared with the measured values. The difference (or error) in the S_{ind} values is ~20%. For a defocus value of 0.05λ (recommended value, see section 6.2.3) the error is < 9%. The other aberration indicators are hardly affected by the presence of spherical aberration (see Fig. 6.15 b-d).

6.5.6 Conclusions on experimental results

In the single aberration scenario, the amount of aberration which can be retrieved with the intensity distribution fit method, using orthogonal circle polynomials, lies in the range of about [0.03λ, 0.13λ] with a precision of roughly ±0.01λ. Smaller values than 0.03λ are difficult to retrieve because the intensity distributions are hardly distorted by aberrations, whereas for aberration amplitudes above 0.13λ the curves become non-linear and the primary aberration retrieval process is not trivial. It is noted that aberration smaller than 0.03λ barely affect the image quality of imaging system.

From the previous it follows from Fig. 6.12a-6.15a that the amplitude of primary aberration can be determined, when both the indicator value and the amount of defocus are known. We will later show (see section 6.5) that this can also be used when small amounts of additional primary aberrations are present.

6.5.7 Varying the fit region diameter

In Fig. 6.16, the effect of reducing the diameter of the fit region on experimental data is shown and compared with the values obtained with Dai's formula [103] (see equation 6.20). Experimental aberrated intensity distributions with a wavefront aberration of $\sim 0.08\lambda$ were fitted within a circular region with varying diameters (range 0.5-1 AU). The fitting procedure is explained in section 6.4. The obtained fitting coefficients were compared with the coefficients calculated with Dai's formula.

The input for Dai's formula (for a fit region of a 1AU) was calculated using the ENZ theory.

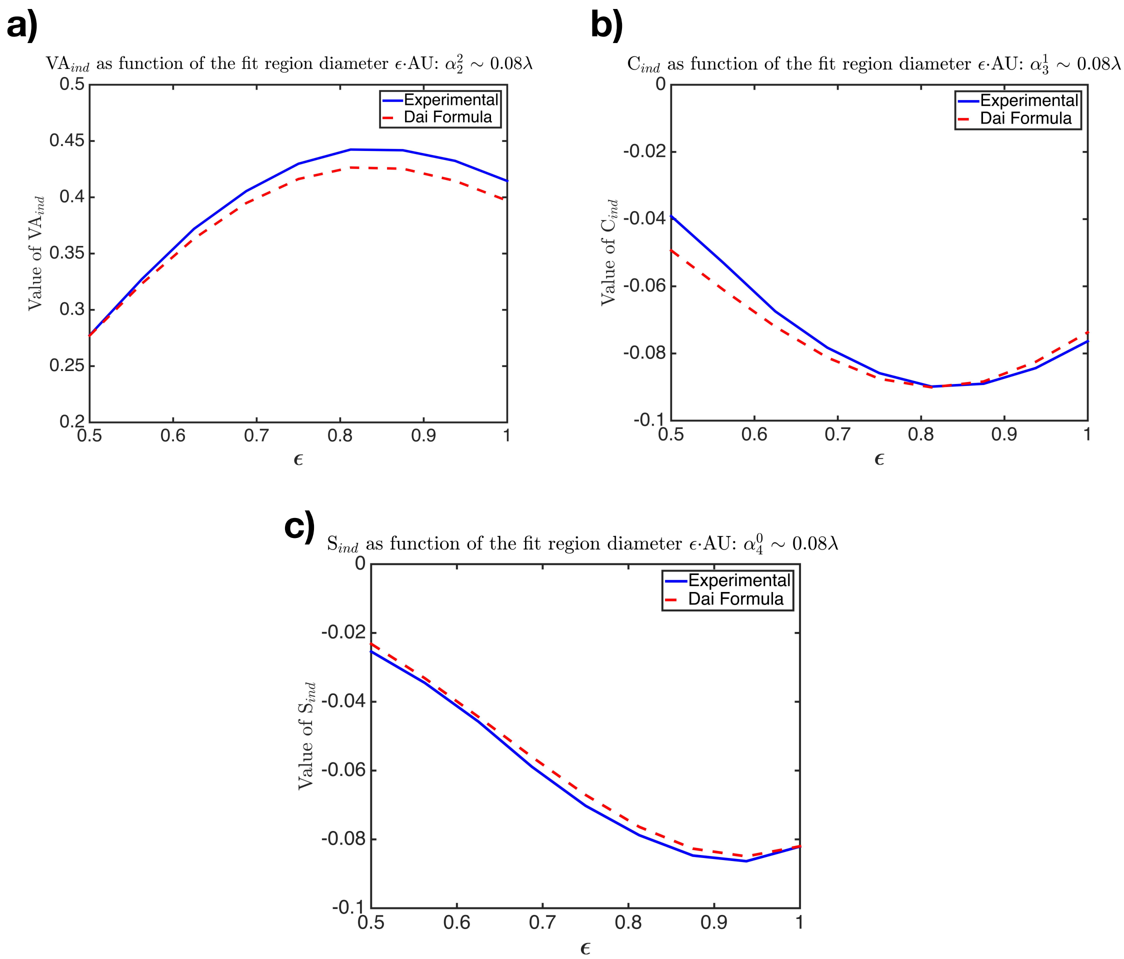


Figure 6.16: Comparison between scaling down the fitting coefficients to smaller fit regions on experimental data or by computation with the Dai formula for a) vertical astigmatism, b) vertical coma and c) spherical aberration.

Theoretical and experimental results seem to be in good agreement and thus it confirms that reducing the diameter of the fit region is in fact the same as scaling down the Zernike coefficients to smaller pupil sizes, as described by Dai's formula (see equation 6.20).

6.6 Multiple aberration scenarios

6.6.1 Simulation results

The performance of the aberration retrieval method in the presence of multiple aberrations will be discussed now. Often in optical imaging systems, such as microscopes, aberrations can appear in combinations rather than individually. The variations of VA_{ind} , C_{ind} or S_{ind} (see section 6.2.2) in the presence of other primary aberrations is studied using Fourier optics simulations. Ideally the indicators would not be affected by cross-talk of other aberrations and their values would stay constant. In the following the simulation results show the aberration amplitude range where the indicators stays sufficiently constant before they become affected by aberration cross-talk.

An analysis based on the ENZ theory became too complicated. It would require the extension of the Taylor series to include further terms, involving probably also higher order polynomials (>16). First try-outs resulted already in cumbersome expressions containing a large number of aberration product terms. Thus aberration cross-talk was studied using Fourier optics simulations only.

6.6.1.1 Cross-talk simulation - Astigmatism

In order to compare the FFT simulations with experimental results, the method for finding the position of the maximum intensity of the two dimensional PSF described previously (see equations 5.23-5.27) was used. First the two indicators of the two primary aberrations which have shown to be less influenced by crosstalk, i.e. astigmatism and spherical aberration, are discussed. Starting with the astigmatism indicator and coma cross-talk, the phase function in the pupil plane is:

$$\Phi = \alpha_2^2 Z_2^2\left(\rho,\theta\right) + \alpha_3^1 Z_3^1\left(\rho,\theta\right) \tag{6.26}$$

For astigmatism and spherical aberration, the phase function in the pupil plane is:

$$\Phi = \alpha_2^2 Z_2^2(\rho,\theta) + \alpha_4^0 Z_4^0(\rho,\theta) \tag{6.27}$$

The value of α_2^2, the coefficient for astigmatism, is kept constant while either α_3^1 or α_4^0 is varied. For α_2^2 three different values were chosen: [0.035λ, 0.07λ, 0.105λ] . The α_3^1 and α_4^0 coefficient were each varied within a range of [-0.1λ, 0.1λ]. VA_{ind} was evaluated at a defocus of $\alpha_2^0 = 0.05\lambda$. The results of these simulations are shown in Fig. 6.17.

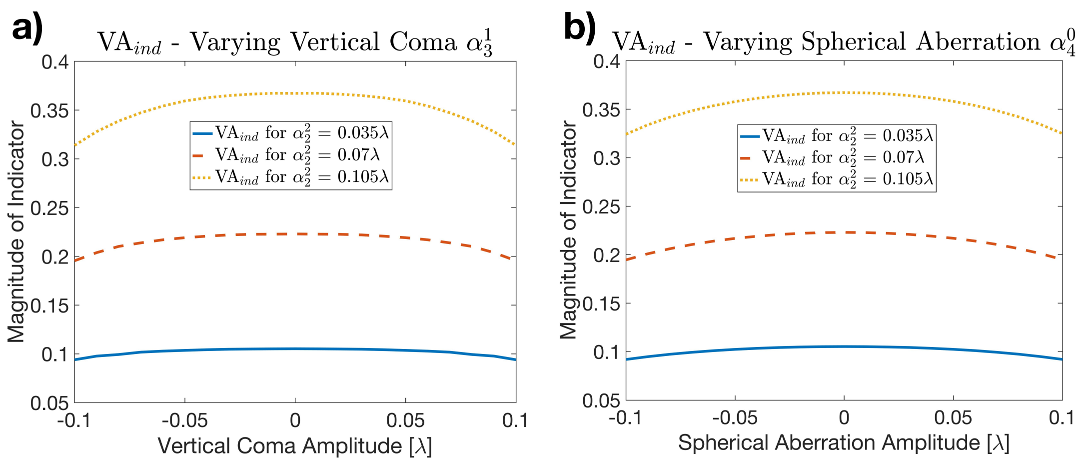


Figure 6.17: Influence of coma α_3^1 and spherical aberration α_4^0 on VA_{ind} for three different values of the astigmatism coefficient $\alpha_2^2 = [0.035\lambda, 0.07\lambda, 0.105\lambda]$ with varying amounts of: a) vertical coma α_3^1; b) spherical aberration α_4^0 .

Figure 6.17 shows that crosstalk becomes more significant with increasing amounts of astigmatism. Cross-talk effects for coma and spherical aberration are similar. For values of α_2^2 up to about 0.05λ, the astigmatism indicator is hardly affected by cross-talk. Thus, for $\alpha_2^2 > 0.05\lambda$, the astigmatism indicator may be considered sufficiently constant for aberration retrieval provided coma and spherical aberration are within a range of ~ ±0.07λ (VA_{ind} varies by less than 6%). Beyond the range ±0.07λ, the variation of VA_{ind} become larger than 10%. It is further noted that the astigmatism indicator is symmetrical around the origin both for coma as well as for spherical aberration.

6.6.1.2 Cross-talk simulation - Spherical aberration

Now the crosstalk effects on S_{ind} are considered. For spherical aberration and vertical astigmatism, the phase function in the pupil plane is:

$$\Phi = \alpha_2^2 Z_2^2(\rho,\theta) + \alpha_4^0 Z_4^0(\rho,\theta) \tag{6.28}$$

For spherical aberration and coma, the phase function in the pupil plane is:

$$\Phi = \alpha_3^1 Z_3^1(\rho,\theta) + \alpha_4^0 Z_4^0(\rho,\theta) \tag{6.29}$$

The α_4^0 is kept constant while either α_2^2 or α_3^1 is varied. For α_4^0 three different values were chosen: [0.035λ, 0.07λ, 0.105λ]. The values of the α_2^2 and α_3^1 coefficient were each varied within a range of [-0.1λ, 0.1λ]. S_{ind} was evaluated at a defocus of $\alpha_2^0 = 0.05\lambda$. The results of these simulations are shown in Fig. 6.18.

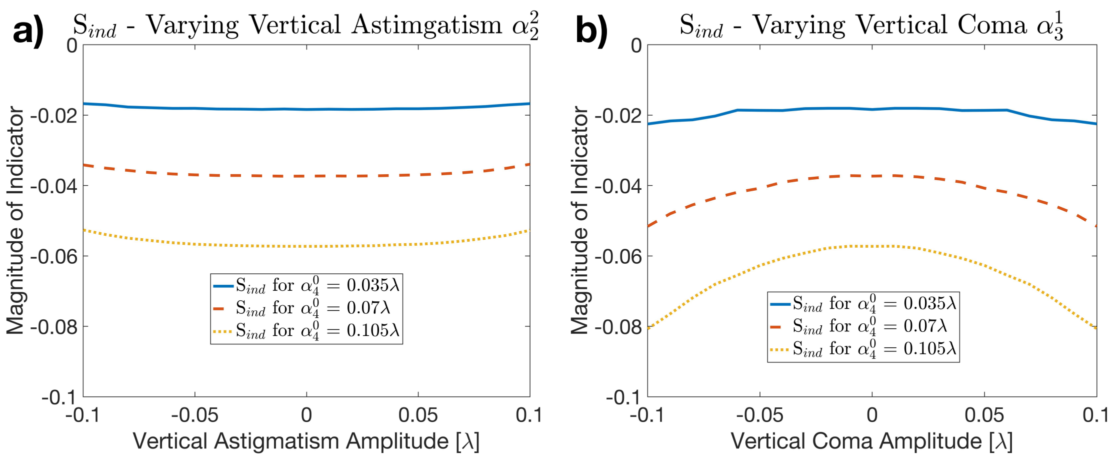


Figure 6.18: Influence of astigmatism α_2^2 and coma α_3^1 on S_{ind} for three different values of the spherical aberration coefficient $\alpha_4^0 = [0.035\lambda, 0.07\lambda, 0.105\lambda]$ with varying amounts of: a) astigmatism α_2^2; b) coma α_3^1.

S_{ind}, in each of the two different focal planes, is more affected by coma than by astigmatism. Figure 6.18 shows that crosstalk becomes more significant with increasing amounts of spherical aberration. Below 0.07λ, S_{ind} is hardly affected by cross-talk. For α_4^0 > 0.07λ, the indicator S_{ind} is considered sufficiently constant over the range ~ ±0.05λ (9% variation at 0.05λ of coma). The influence of astigmatism on the S_{ind} can be neglected (less than 8% variation at 0.1λ of astigmatism). It is further noted that S_{ind} is symmetrical around the origin both for astigmatism as well as for coma variations.

6.6.1.3 Cross-talk simulation - Coma

Now the crosstalk effects on the coma indicator are considered. For C_{ind}, the γ_1^1 was chosen (see equation 6.9) because it showed to be most sensitive to coma. The phase function in the pupil plane for coma and vertical astigmatism is:

$$\Phi = \alpha_2^2 Z_2^2(\rho,\theta) + \alpha_3^1 Z_3^1(\rho,\theta) \tag{6.30}$$

For coma and spherical aberration, the phase function in the pupil plane is:

$$\Phi = \alpha_3^1 Z_3^1(\rho,\theta) + \alpha_4^0 Z_4^0(\rho,\theta) \tag{6.31}$$

The value of α_3^1 is kept constant while either α_2^2 or α_4^0 is varied. For α_3^1 three different values were chosen: [0.035λ, 0.07λ, 0.105λ]. The values of the α_2^2 and α_4^0 coefficient were each varied within the range of [-0.1λ, 0.1λ]. The results of these simulations are shown in Fig. 6.19.

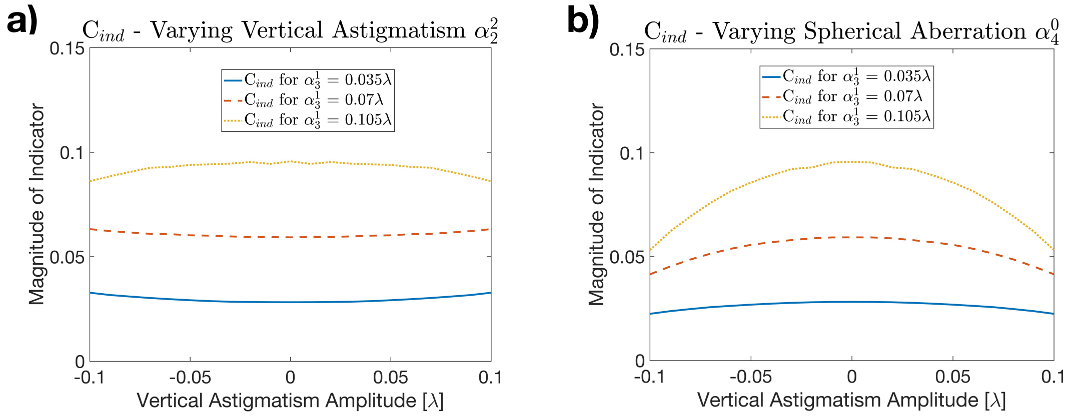


Figure 6.19: Impact of astigmatism α_2^2 and spherical aberration α_4^0 on the C_{ind} (using the γ_1^1 coefficient) for three different values of $\alpha_3^1 = [0.035\lambda, 0.07\lambda, 0.105\lambda]$ with varying amounts of: a) astigmatism α_2^2; b) spherical aberration α_4^0.

The coefficient γ_1^1 is more affected by a α_4^0 than by α_2^2. The amount of tolerable astigmatism is slightly higher than spherical aberration (for cross-talk resulting in a 10% change in C_{ind}, α_2^2 can be ~15% larger than α_4^0). Figure 6.19 shows that crosstalk becomes more significant with increasing amounts of coma. For small coma amplitudes α_3^1 (up to ~0.05λ), the value of C_{ind} changes by less than 10% when $\alpha_4^0 \leq 0.07\lambda$. For $\alpha_3^1 > 0.05\lambda$, C_{ind} is considered sufficiently constant (value varies by less than 10%) and reliable enough for aberration retrieval, provided the other primary aberrations are within a range of ~ ±0.05λ. Surprisingly, C_{ind} is less affected by astigmatism cross-talk at $\alpha_2^2 = 0.07\lambda$ than at $\alpha_2^2 = 0.035\lambda$ but then increases again for larger amounts of astigmatism. An almost flat curve is obtained for $\alpha_2^2 \approx 0.08\lambda$. It is noted that the C_{ind} is symmetrical around the origin both for astigmatism as well as for spherical aberrations variations.

Based on the previous results (Fig. 6.17-19) aberration retrieval, in the presence of multiple primary aberrations, is possible and reliable (error smaller than 10%) if, as rule of thumb, the rms wavefront aberrations is not considerably larger than 0.1λ.

6.6.2 Experimental results

In the following experiments, first a constant amount of ~0.06λ of a primary aberration (either vertical astigmatism, vertical coma or spherical aberration) was applied with the DMM. Then a second primary aberration was additionally applied and varied over the range of ~[-0.13λ, 0.13λ] in order to mimic the numerical simulation in section 6.6.1. Defocus α_2^0 (range ~[-0.1λ, 0.1λ]) was also applied with the DMM. The best focus plane was determined by selecting the axial plane wherein the camera image has the highest intensity pixel value of all (prior to using auto-exposure). Then, auto-exposure was used to assure that for each measured intensity distribution the pixel of highest intensity was at about 90% of the intensity saturation value of a pixel, this gave a high signal-to-noise ratio. The aberration indicators characteristic for the aberration combinations applied (astigmatism-coma, astigmatism-spherical aberration, coma-spherical aberration) were plotted (figs. 6.20-6.22). These figures show how the different aberration indicators are affected by cross-talk caused by the other primary aberrations. To evaluate VA_{ind} and S_{ind} two intensity distributions on opposite focus are needed. Aberration retrieval involved using five camera images: three camera images close to best focus (defocus values: ~[-0.025λ, 0, 0.025λ]) and two defocused images with the same amount of defocus but lying on opposite sides of best focus plane (defocus values: ~ ±[0.025λ, 0.05λ, 0.075λ, 0.1λ]). It is possible to retrieve coma with only one of these camera images but in this experiment three images were taken to average the C_{ind} value. To mitigate the effects of cross-talk by the DMM, the corresponding aberration indicator curves obtained for the single aberration scenario (Section 6.5) were subtracted from the plotted curves in Figure 6.20-22. For example, the measured values of when only varying spherical aberration and no astigmatism was present (see Fig. 6.15c) were subtracted from the measured curve that was obtained by applying a constant amount of astigmatism and varying only the amount of spherical aberration. To estimate the amount of the aberrations present, the measured indicators should be compared with their respective reference curve at the same defocus value (figs. 6.12a-6.15a, show single aberration indicator curves without aberration cross-talk).

6.6.2.1 Experimental cross-talk - Astigmatism

First, the crosstalk effects with regards to VA_{ind} (Equation 6.7) are discussed. A constant amount of vertical astigmatism $\alpha_2^2 \approx 0.06\lambda$ was applied. Vertical coma α_3^1 and spherical aberration α_4^0 were then varied separately from each other (range ~[-0.13λ, 0.13λ]). VA_{ind} was evaluated using axial plane pairs having different amounts of defocus $\alpha_2^0 = \pm 0.025\lambda; \pm 0.05\lambda; \pm 0.075\lambda; \pm 0.100\lambda$, respectively. The experimental results are shown in Fig. 6.20.

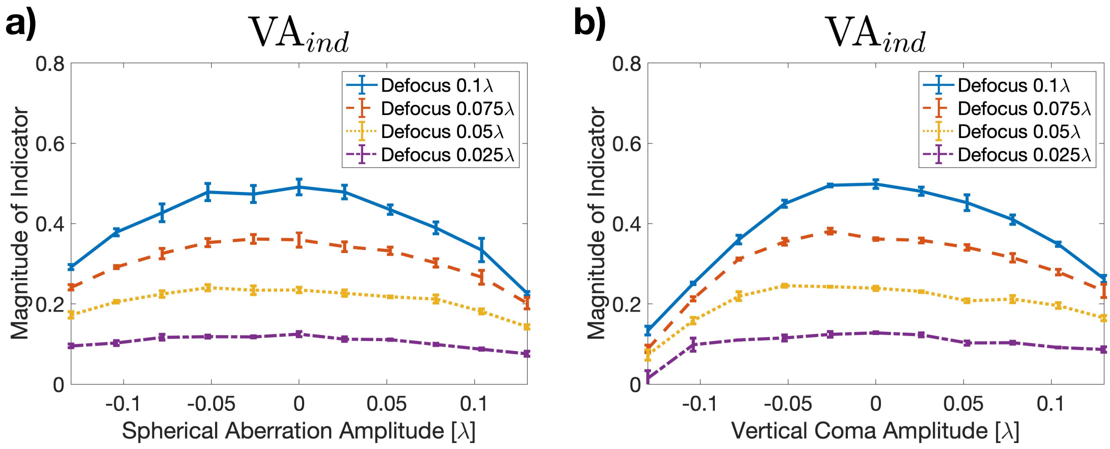


Figure 6.20: Experimental data on the influence of coma α_3^1 and spherical aberration α_4^0 on VA_{ind}: A constant amount of vertical astigmatism $\alpha_2^2 \approx 0.06\lambda$ was applied with the DMM while varying either a) α_4^0 or b) α_3^1. VA_{ind} was evaluated using axial plane pairs having different amounts of defocus $\alpha_2^0 = \pm 0.025\lambda; \pm 0.05\lambda; \pm 0.075\lambda; \pm 0.100\lambda$, respectively.

The astigmatism indicator VA_{ind} is more affected by coma (Fig.6.20b) than by spherical aberration (Fig. 6.20a). For increasing amounts of coma and spherical aberration, the value of VA_{ind} decreases and tends towards zero. The coma cross-talk is somewhat asymmetrical (decrease is stronger on the left hand side in Fig. 6.20b).

When using axial plane pairs with defocus values up to 0.05λ, VA_{ind} remains sufficiently constant within the aberration range of ±0.07λ resulting in an aberration estimation error smaller than 10%. When evaluating the indicator at larger defocus, the aberration range should not exceed ±0.05λ to satisfy an aberration estimation error smaller than 10%. To estimate the amount of astigmatism, the measured indicator value should be compared with its respective reference curve (see section 6.5.3 and Fig. 6.13, VA_{ind} in the sole presence of astigmatism without aberration cross-talk). For example, the value of VA_{ind} for $\alpha_4^0 = 0$ (centre of the graph Fig. 6.20a) is about -0.36 at a defocus value of 0.075λ which corresponds $\alpha_2^2 \approx 0.06\lambda$.

6.6.2.2 Experimental cross-talk - Spherical aberration

Now, the attention is drawn to the crosstalk affecting the spherical aberration indicator S_{ind} (Equation 6.10). A constant amount of spherical aberration $\alpha_4^0 \approx 0.06\lambda$ was applied. Then, either astigmatism α_2^2 or coma α_3^1 was varied (range ~[-0.13λ, 0.13λ]). S_{ind} was evaluated using axial plane pairs having different amounts of defocus $\alpha_2^0 = \pm 0.025\lambda; \pm 0.05\lambda; \pm 0.075\lambda; \pm 0.100\lambda$, respectively. The experimental results for S_{ind} are shown in Fig. 6.21.

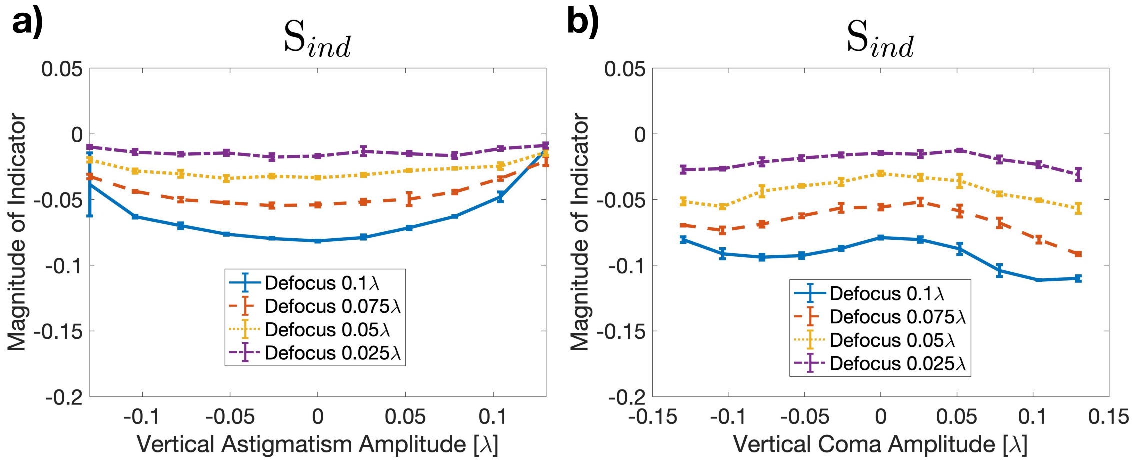


Figure 6.21: Experimental data on the influence of astigmatism α_2^2 and coma α_3^1 on S_{ind}: A constant amount of spherical aberration $\alpha_4^0 \approx 0.06\lambda$ was applied with the DMM while varying either a) α_2^2, or b) α_3^1. S_{ind} was evaluated using axial plane pairs having different amounts of defocus $\alpha_2^0 = \pm 0.025\lambda; \pm 0.05\lambda; \pm 0.075\lambda; \pm 0.100\lambda$, respectively.

When using axial plane pairs with defocus values up to ±0.05λ, S_{ind} remains almost constant and is hardly affected by the presence of either astigmatism within an amplitude range of ±0.07λ, or coma within an amplitude range of ±0.05λ. Outside these ranges the values of S_{ind} become dependent on the amount of astigmatism α_2^2 and coma α_3^1 present. The spherical aberration indicator S_{ind} is more affected by the presence of coma than by astigmatism. With increasing amounts of astigmatism α_2^2, the negative value of S_{ind} gets closer to zero, while, with increasing amounts of coma, the value of S_{ind} tends to increase to larger negative values. As a consequence, α_4^0 would be underestimated in the presence of large amounts of astigmatism and overestimated with increasing amounts of coma. Since the amount of S_{ind} and the amount of defocus are both known, the corresponding amount of spherical aberration can easily be determined (see section 6.5.5 and Fig. 6.15). For example, the value of S_{ind} for $\alpha_2^2 = 0$ (centre of the graph Fig. 6.21a) is about -0.05 at a defocus value of 0.075λ and which corresponds to $\alpha_4^0 \approx 0.06\lambda$.

6.6.2.3 Experimental cross-talk - Coma

Last, the crosstalk affecting C_{ind} (Equation 6.9) is considered. A constant amount of coma $\alpha_3^1 \approx 0.06\lambda$ was applied. Astigmatism α_2^2 and spherical aberration α_4^0 were then varied independently from each other (over a range of ~[-0.13λ, 0.13λ]). The experimental results are shown in Fig. 6.22.

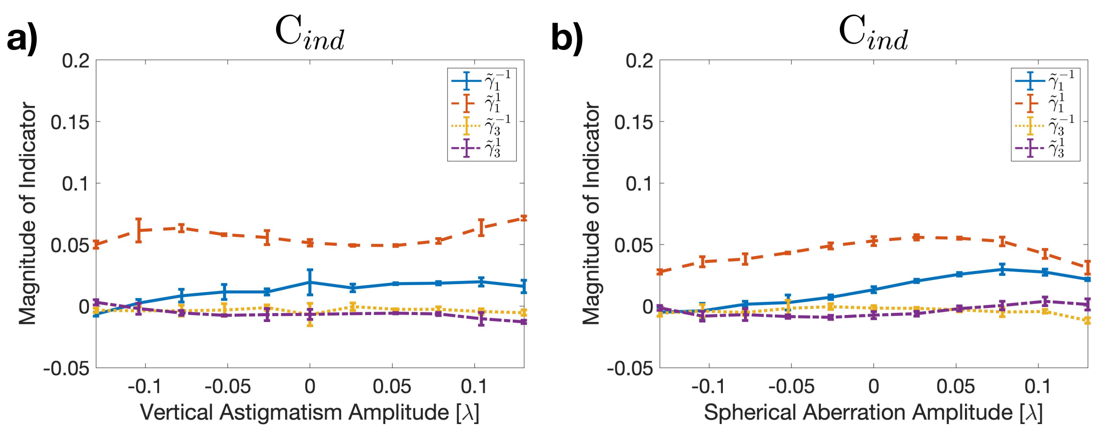


Figure 6.22: Experimental data on the influence of astigmatism α_2^2 and spherical aberration α_4^0 on C_{ind}: A constant amount of vertical coma $\alpha_3^1 \approx 0.06\lambda$ was applied with the DMM while varying a) α_2^2 or b) α_4^0.

The C_{ind} is less constant than VA_{ind} and S_{ind} (see Fig. 6.20 and 6.21, respectively). The value of C_{ind} decreases with increasing amounts of spherical aberration α_4^0, whereas the curve undulates with increasing amount of astigmatism α_2^2. In the aberration amplitude range ±0.05λ, the C_{ind} value is almost constant and the aberration estimation error is smaller than 10%. To estimate the amount of coma, the measured indicator value should be compared with its respective reference curve (see section 6.5.2 and Fig.6.12, C_{ind} in the sole presence of coma without aberration cross-talk). For example, the value of C_{ind} for $\alpha_2^2 = 0$ (centre of the graph Fig. 6.22a) is about 0.05 and which corresponds to $\alpha_2^2 \approx 0.06\lambda$. This cross-talk could be caused either by the DMM operated in open-loop, or by small misalignments of the pupil relative to the centre of the DMM (see chapter 3, section 3.1.3)

The experimental results (fig 6.20-6.22) resemble the FFT simulation results (fig 6.17-6.19). Especially when it comes to the onset of cross-talk (aberration amplitude at which cross-talk becomes significant), a comparable behaviour is observed. The indicator curves look similar, especially for VA_{ind} and S_{ind} . The C_{ind} differs most as can be seen in fig 6.22 where the curves seem to be more subject to noise issues. This might be caused by inaccuracies in finding the position of maximum intensity on the EMCCD detector. A second order polynomial does not describe well the intensity distribution of a coma aberrated spot around its maximum (especially in the direction of the coma flare). Including third order polynomials might be more accurate. Another possible explanation for the discrepancy between simulations and experimental results could be the open-loop control of the DMM. Applying combinations of different aberrations, in open-loop, can also create cross-talk, especially when the amount of an aberration increases. Nevertheless, as a proof of concept the open-loop experiment demonstrates the viability of the method but better results may be achieved by operating the DMM in closed-loop.

6.7 Conclusions

From what has been presented so far, it can be concluded that the aberration indicators proposed are suitable for linear aberration retrieval of primary aberrations (astigmatism, coma and spherical aberration). The astigmatism and spherical aberration indicator are best measured at paired defocused planes (±0.05λ defocus), whereas the coma indicator is best measured close to best focus. The smallest aberration amplitude that can be reliably detected is ~0.025λ. The indicator curves remain linear for aberration amplitudes up to at least 0.13λ (in the presence of a single primary aberration). In the presence of two different primary aberrations, in particular for astigmatism and spherical aberration, the indicators can, as a rule of thumb, be used for aberration retrieval for rms wavefront aberrations smaller than 0.1λ. The method can also potentially detect the three different primary aberrations. However, considering that the aberration amplitude detection range is ~ 0.03-0.13λ, its application range for measuring three different primary aberrations would be more limited.

The described aberration retrieval method, based on fitting orthogonal circle polynomials through intensity distributions, has the advantage of being robust in terms of least-square fitting (using orthogonal functions to fit experimental data) and mitigates adverse effects of noise by limiting the fit to a central relatively high intensity region. The proposed method is relatively simply to implement. It is a non-iterative method, and the aberration indicators vary linearly with their respective aberration and allow primary aberrations in optical imaging systems to be measured with a small number of measurements (< 5 camera images). There is no need to compute FFT's as with the Gerchberg-Saxton algorithms, nor is it necessary to use the complex analytical expressions of the ENZ theory for computing aberrated intensity distributions through focus. For the aberration range tested, our aberration retrieval method is non-iterative in contrast to both the Gerchberg-Saxton algorithm and the phase retrieval method based on the ENZ theory (which has been used for high NA optical lithography systems), which are both iterative. In comparison with the modal method, described in chapter 4, which requires ideally 5 measurements per aberration and a DMM to iteratively apply Zernike modes, the here proposed aberration retrieval method needs only 3 images of the aberrated intensity distribution through focus (a minimum of three images, but ideally five to average the coma indicator around best focus). It is recommended that there are at least 9 pixels within the Airy disk diameter for retrieving primary aberrations with this method. Other linear phase retrieval methods [104] are often limited to small aberration ranges ($< 0.07\lambda$) but with the new retrieval method aberrations with amplitudes up to 0.13λ can be retrieved.

6.8 Applications

A useful application for this aberration retrieval technique would be in testing the quality of manufactured lenses. The technique provides a simple, non-iterative, quick (3 or 5 images per lens) and also relatively cheap method (it only requires a scanning stage and a camera - no need for wavefront sensors, interferometers or DMMs). A simple setup for testing lenses is depicted in Fig. 6.23. Although this has only been demonstrated for a low NA lens, this aberration retrieval method could be further studied and possibly expanded to higher NA microscope objectives

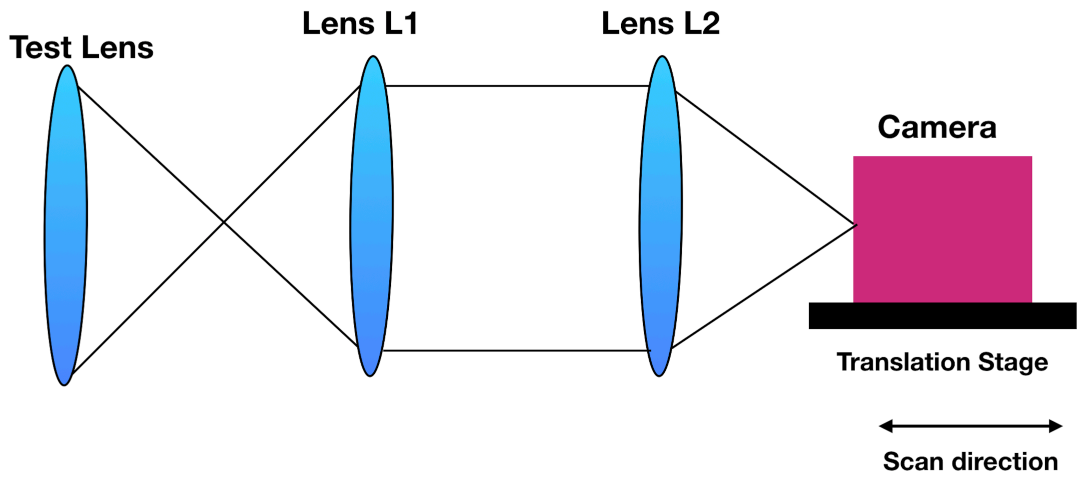


Figure 6.23: Optical setup With three lenses, i.e the test lens and an objective whose aberrations are known followed by a low NA lens, i.e. the NA should assure that there are enough pixels to span over the Airy Disk. The primary aberrations are retrieved using the aberration retrieval method involving Zernike polynomial fitting of an aberrated PSF through focus.

7. Conclusions and Future Work

7.1 Conclusions

The work and results presented in this thesis relate to aberration retrieval techniques based on intensity data in and around focus, as well as methods for subsequently correcting for the aberrations present. The methods were developed for laser scanning microscopes but the aberration retrieval method described in Chapter 6 could also be applied in a wider range of microscopes (e.g. widefield microscope) or other types of optical imaging systems, such as telescopes and optical lithography projection lenses. In the following concluding remarks on the different Chapters of this thesis are given.

In Chapter 3 a homebuilt confocal microscope, with integrated AO, was presented. The intensity detector consisted of an EMCCD camera, as opposed to the more conventional pinhole and photodetector typical used in confocal microscopy. Using the EMCCD, confocal detection was achieved by selecting pixels within a circular area (41 pixels were selected to mimic a 1AU pinhole in the microscope setup described in Chapter 3) and adding up their measured intensity values. It was shown that a closed-loop calibration of the DMM was necessary to obtain Zernike modes with a high degree of purity (to an accuracy of $\lambda/50$ rms). Furthermore, it was found that misalignment of the pupil of the objective lens with respect to the DMM pupil, could generate unwanted aberrations when applying Zernike modes with the DMM and thus compromise the quality of the aberration correction.

A confocal modal sensor-less wavefront sensor was studied in the home-built confocal microscope. The DMM was used to sequentially impose varying amounts of Zernike aberrations, with the aim of maximising the intensity of light passing through pinhole of the confocal microscope to correct for aberrations. The sensitivity of the method with respect to pinhole size and aberration type was investigated. The experimental results suggest that the pinhole diameter should be smaller than 1AU to have a good sensitivity for all tested Zernike aberrations (up to radial order 4). Modal sensitivity improves with decreasing pinhole diameter but no significant improvement is obtained below ~0.3AU. Large aberrations (> 0.071λ rms) make modal wavefront sensing challenging because of aberration cross-talk, and the potential presence of local sub-maxima which can lead to stagnation of the modal optimisation process. In this thesis, an optimisation strategy is proposed to tackle large aberrations more efficiently by using a wavefront pre-correction determined from ray-tracing simulations of the sample. Starting the optimisation with a ray-tracing pre-correction applied to the DMM simplifies the final correction procedure and reduces the time taken to complete a modal optimisation by minimising the number iteration steps necessary for correction. It was also shown that when sample parameters (refractive index, thickness, curvature, etc) are not well known, an estimated ray-tracing pre-correction would still reduce the sample aberration to a more manageable level and therefore speed up the correction process. Increasing the correction speed will reduce the risk of photon damage by exposing the sample to light for as little time as possible. It would also be beneficial for biological samples which do not feature complex refractive index distributions. The pre-correction optimisation approach gives more flexibility when correcting for aberrations, e.g. one might focus deeper into transparent samples, or one could make local pre-corrections when scanning an image plane.

The double-pass effect, which causes cancellation of odd aberration after second pass through the optical system, makes wavefront sensing in reflection setups challenging. A confocal edge scan, which involves scanning the focal spot over an edge, is shown to detect the presence of coma (an odd aberration) in a reflective confocal setup. The orientation of coma can be determined from scanning the spot over two edges which are perpendicular to each other. The amplitude of the coma aberration can be estimated by comparison of the measured edge responses with theoretical curves. The sensitivity of the method can be increased by reducing the size of the pinhole. Coma amplitudes > 0.035λ and up to about ~0.14λ, with an accuracy of ±0.02λ, were detectable with the proposed edge scan method. For larger amounts of coma, the edge response is strongly distorted which would make any further analysis difficult. The method offers a simple way to break the symmetry and detect coma in reflection confocal microscopes.

In the last chapter, a novel aberration retrieval method was presented to determine primary aberrations by making use of information about the intensity distribution at and around focus, within a volume that is bound by the first dark ring of the Airy spot (at a distance of 1 AU from the centre). The method requires 3 intensity distributions, preferably one intensity distribution at best focus and two intensity distributions at opposite side of best focus (defocus: $\left|\alpha_2^0\right| < 0.07\lambda$), to retrieve the primary aberrations. The intensity distribution within the first dark ring of the Airy spot are fitted with a set of orthogonal polynomials in this case Zernike polynomials are used for convenience. Using specific combinations of the coefficients from the fitted polynomials a set of aberration indicators are determined for the primary Zernike aberrations. These indicators are selected to vary linearly with the aberration of interest and to be insensitive to the presence of other aberrations. The method works best in the presence of a single aberration where amplitudes up to at least 0.13λ can be retrieved linearly with a precision of roughly ±0.01λ. The method can also retrieve aberration combinations as long as the total wavefront rms of the aberrations is within 0.1λ rms. The method is numerically robust because orthogonal circle polynomials are used for fitting the intensity distributions. In addition, the effects of shot-noise are mitigated, because the fit is performed only in a relatively high intensity area. For the aberration range tested, the aberration retrieval method is non-iterative in contrast to both the Gerchberg-Saxton algorithm and the phase retrieval method based on the ENZ theory (which has been used for high NA optical lithography systems), which are both iterative. In comparison with the modal method, described in Chapter 4, which requires ideally 5 measurements per aberration and a DMM to iteratively apply Zernike modes, the here proposed aberration retrieval method needs only 3 images of the aberrated intensity distribution through focus (a minimum of three images, but ideally five to average the coma indicator around best focus). To our knowledge the method described in Chapter 6 is the fastest phase retrieval method for determining primary aberrations.

7.2 Future Work

The work in this thesis has resulted in the development of novel aberration retrieval techniques and correction. In particular the method, described in Chapter 6, is a novel aberration retrieval technique which could be further explored. The method has been developed to retrieve primary aberrations in low NA systems with homogeneous illumination. Some aspects which could be addressed in the future are:

- How non-homogeneous pupil illuminations (e.g. Gaussian beam) affect the performance of the aberration retrieval method. Non-homogeneous illumination can easily be implemented in diffraction simulations by modifying the amplitude distribution in the pupil.

- Addressing the applicability of the method to high NA systems. This could be done by taking into account the vectorial character of light and polarisation effects. Simulations using vectorial diffraction theory, would shed light on the capability of the method to retrieve primary aberrations in high NA systems.

- Investigating cross-talk using the ENZ theory to obtain analytical expressions for intensity distributions in the presence of multiple aberrations. These expressions would help to get a better understanding of cross-talk effects and have the potential to allow an extension of the range over which aberration retrieval is possible (it would allow non-linear effects to be taken into consideration).

- Exploring the capability of the method to retrieve higher order aberrations. So far, the method has been limited to primary aberrations, which are, in general, the most important aberrations in optical systems. The effect of higher order aberrations on the intensity distribution is described by higher order Bessel functions (see Appendix C, Figure C.1). Higher order Bessel functions have their maxima outside the first dark ring of the Airy spot intensity pattern. It is thus conceivable that higher order aberrations can be retrieved by using intensity data outside the Airy disk. Enlarging the fit area or limiting the fitting to a ring region, outside the Airy disk, could allow retrieval of these higher order aberrations

As regards the edge scan method for detecting the presence of coma in reflection setups, the edge response produced by other odd aberrations, such as trefoil, could also be considered for further work. It should be mentioned however that trefoil is an aberration which, in general, does not deteriorate the quality of industrial optical imaging systems. Optical metrology is an area wherein the edge scan method would be useful. In metrology, microscopes are used for measuring the topography of samples. An odd aberration such as coma would limit the capability of the microscope to resolve surface features. Coma could be easily detected with the edge scan method and then be corrected for before measuring surface features.

Appendix A

Here the way ray-tracing was implemented in Matlab is discussed. In this thesis ray-tracing was used to determine sample induced wavefront aberrations in the pupil plane of the microscope's objective. The samples were simulated as simple geometrical objects (cylinders, plates). Ray-tracing is used in optical design to quantify and analyse the imaging performance of optical systems. Ray-tracing is a process of tracking the path of a ray (or rays) from surface to surface as a sequence of successive transfers between adjacent surfaces and entails changes in direction of a ray after refraction on a surface (or reflection). It involves the simple geometry of straight lines for the transfer, and uses Snell's law for finding the direction of the ray after refraction. Optical path differences of rays in the entrance pupil plane of an objective are used to determine the wavefront aberrations. Deviations from a spherical reference wavefront are caused by aberrations. A calculated wavefront may be decomposed in terms of Zernike aberrations (see Chapter 2). For ray-tracing we have chosen to follow the method described by Welford in [105].

Starting with Snell's law:

$$n'\sin I' = n\sin I \tag{A.1}$$

where n and n' are the refractive indices of the first and second medium, respectively. I and I' are the angle of incidence and refraction. For three dimensional ray tracing, one can use the vector form of Snell's law which is given by:

$$n'\left(\vec{r}'\times\vec{n}\right) = n\left(\vec{r}\times\vec{n}\right) \tag{A.2}$$

If this equations are multiplied by $\vec{n}$ we obtain $\vec{n}\times n'\left(\vec{r}'\times\vec{n}\right) = \vec{n}\times n\left(\vec{r}\times\vec{n}\right)$ which can also be written as $n'((\vec{n}\bullet\vec{n})\vec{r}' - (\vec{r}'\bullet\vec{n})\vec{n} = n((\vec{n}\bullet\vec{n})\vec{r} - (\vec{r}\bullet\vec{n})\vec{n})$ and thus as $n'(\vec{r}' - (\vec{r}'\bullet\vec{n})\vec{n} = n(\vec{r} - (\vec{r}\bullet\vec{n})\vec{n})$):

$$n'\vec{r'} - n\vec{r} = \left\{ n'(\vec{r'} \bullet \vec{n}) - n(\vec{r} \bullet \vec{n}) \right\} \vec{n}$$
$$n'\vec{r'} - n\vec{r} = (n'\cos I' - n\cos I)\vec{n} \tag{A.3}$$

where r and r' represent the incident and refracted ray vectors. Expanding the previous expression in scalar form we obtain the following set of equations:

$$\begin{aligned} n'L' - nL &= k\alpha \\ n'M' - nM &= k\beta \\ n'N' - nN &= k\gamma \end{aligned} \tag{A.4}$$

where

$$\begin{aligned} k &= n'\left(\vec{r'} \bullet \vec{n}\right) - n\left(\vec{r} \bullet \vec{n}\right) \\ &= n'\cos I' - n\cos I \end{aligned} \tag{A.5}$$

L, M, N are the direction cosines of the incident ray, L', M', N' are the direction cosines of the refracted ray, and α, β, γ are the direction cosines of the surface normal on the refractive surface. Following the ray tracing method described by Welford [105], rays propagate from one refracting surface to the next. Each refracting surface has a vertex plane, which is perpendicular to the z-axis and passes through the vertex of the refractive surface. Figure A.1 illustrates the Welford ray tracing method and the coordinates used.

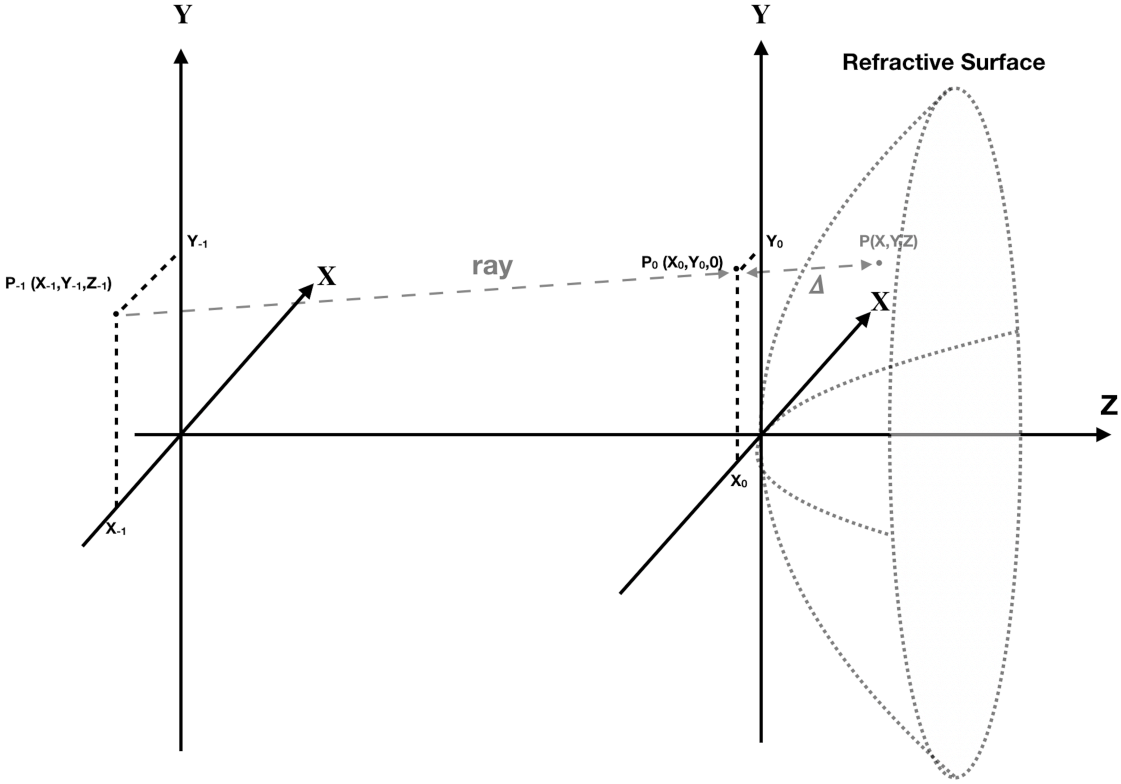


Figure A.1: Ray-tracing geometry used in accordance with the method described by Welford [105]

A ray leaves a "previous" surface at a point P_{-1} (X_{-1}, Y_{-1}, Z_{-1}) with direction cosines (L,M,N) and propagates towards a refractive surface. This is known as the ray transfer process. P_0 (X_0,Y_0,0) is where the ray meets the vertex plane of the refractive surface. The coordinates are given by:

$$\begin{aligned} X_0 &= X_{-1} + \frac{L}{N}\left(d - Z_{-1}\right) \\ Y_0 &= Y_{-1} + \frac{M}{N}\left(d - Z_{-1}\right) \end{aligned} \tag{A.6}$$

The distance from P_0 to a point P (x,y,z) which lies on a surface with curvature c is denoted Δ. The condition that P lies on the surface of curvature is given by (A.7) in case of a spherical surface, and by (A.8) in case of a cylindrical surface with the cylinder axis pointing in the x direction.

$$Z_{sphere} = \frac{1}{2} c\left(X^2 + Y^2 + Z^2\right) \tag{A.7}$$

$$Z_{cylinder} = \frac{1}{2} c\left(X^2 + Z^2\right) \tag{A.8}$$

Now the coordinates of P are:

$$\begin{aligned} X &= X_0 + L\Delta \\ Y &= Y_0 + M\Delta \\ Z &= N\Delta \end{aligned} \tag{A.9}$$

x, y, z can be eliminated from equation (A.7) or (A.8) by using equations (A.9) giving a quadratic equation with a solution of the form:

$$\Delta_{sphere} = \frac{F}{G + \sqrt{G^2 - cF}}, \quad \Delta_{cylindre} = \frac{F}{G + \sqrt{G^2 - c_y AF}} \tag{A.10}$$

where in the case of a sphere:

$$\begin{gathered} F = c\left(X_0^2 + Y_0^2\right), \\ G = N - c\left(LX_0 + MY_0\right) \end{gathered} \tag{A.11}$$

and in the case of a cylinder:

$$\begin{gathered} A = 1 - L^2, \\ F = c_Y Y_0^2, \\ G = N - c_Y MY_0 \end{gathered} \tag{A.12}$$

The optical path difference OPD is then: $OPD = n\left\{\left(d - Z_{-1}\right)/N + \Delta\right\}$.

The components of the normal n at the point of incidence at the cylinder are obtained by taking the partial derivatives of equation (A.8):

$$(\alpha, \beta, \gamma) = \left(0,\ -c_Y y,\ 1 - c_Y z\right) \tag{A.13}$$

and results in the following refraction equations:

$$
\begin{aligned}
n'L' &= nL \\
n'M' &= nM - K_Y Y \\
n'N' &= nN - K_Y Z + n'\cos I' - n\cos I
\end{aligned}
\tag{A.14}
$$

where $K_Y = c_Y\left(n'\cos I' - n\cos I\right)$.

cos$I$ and cosI' have to be determined first. We have

$$
\begin{aligned}
\cos I &= \vec{r} \bullet \vec{n} \\
&= N - c_Y\left(MY + NZ\right) \\
&= N - c_Y\left\{M\left(Y_0 + M\Delta\right) + N^2\Delta\right\} \\
&= N - c_Y\left(1 - L^2\right)\Delta - c_Y M Y_0
\end{aligned}
\tag{A.15}
$$

or

$$
\cos I = +\sqrt{G^2 - c_Y A F}
\tag{A.16}
$$

Thus cos$I$ has already been found in the transfer process (A.10) and we find cosI' by

$$
n'\cos I' = +\left\{n'^2 - n^2\left(1 - \cos^2 I\right)\right\}^{1/2} .
\tag{A.17}
$$

To determine aberrations in the back pupil plane of the objective, it is necessary to trace a fan of rays, originating from a point source, towards the objective. One has to find the coordinates where the rays, with equal optical path, will be. A sphere is fitted, in a least-square sense, through the data points. In reference [106], Braat describes different methods for determining the aberrated wavefront in the microscope objective's pupil based on the optical paths of rays. There is no general preferred method for all situations. We have chosen a different method, in which we fit a reference sphere through the set of coordinates obtained from rays with equal optical paths. As the optical path reference, the optical path of the marginal ray, in the XZ plane, to the pupil edge was chosen. The wavefront deviation is then calculated with respect to the fitted reference wavefront. This method has some similarities with method 4 or 5 in reference [106] in the sense that it depends on the radius of the reference sphere. It distinguishes itself from the other methods in regards to the centre of the spherical reference wavefront. The centre is obtained by least square fitting of a reference sphere (fitted through the set of coordinates obtained from rays with equal optical paths) and not using the marginal/aperture ray (as in method 4) or the paraxial focus (as in method 5). Figure A.2 illustrates how the reference wavefront is fitted.

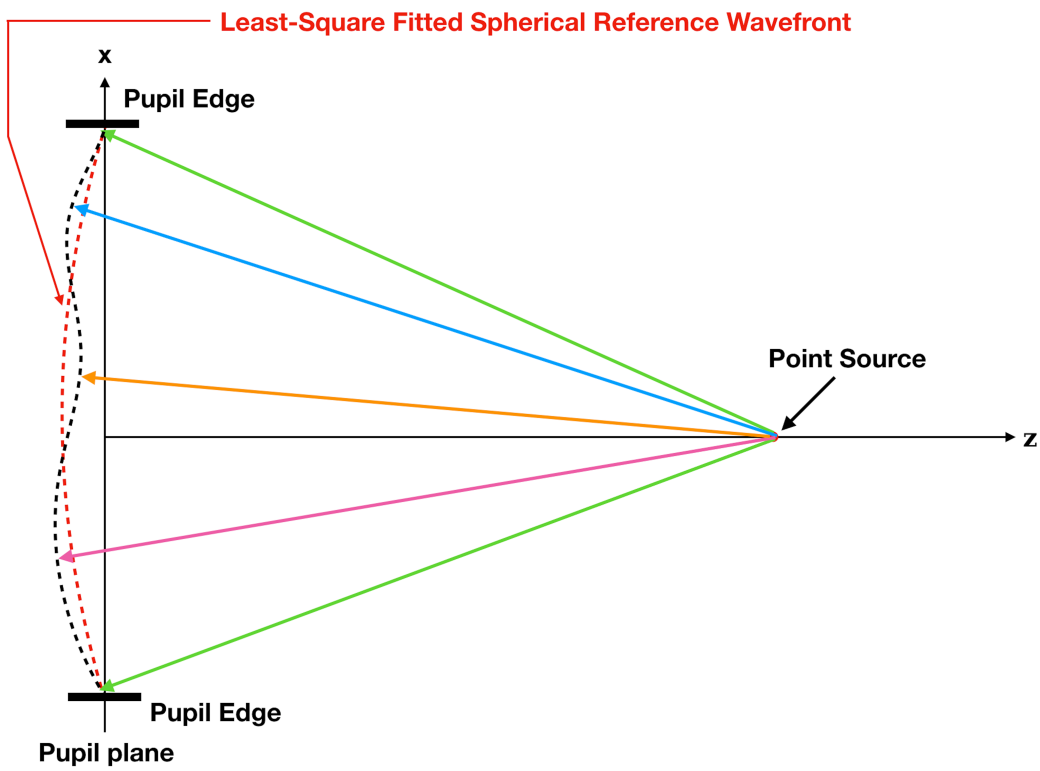


Figure A.2: Least-Square fit of a spherical reference sphere through the set of coordinates obtained from rays with equal optical paths. As the optical path reference, the optical path of the marginal ray, in the XZ plane, to the pupil edge was chosen.

The sphere shell depicts the un-aberrated spherical reference wavefront. The deviation from the reference sphere represents the aberrated wavefront in the objective's pupil. The Zernike polynomials are a set of orthogonal functions over the unit circle and are commonly used to describe wavefront aberrations (see Chapter 2 section 2.2). The aberrations in the objective's pupil are obtained by fitting Zernike polynomials to the aberrated wavefront. To check the results obtained by the ray tracing program, a comparison was made with literature [106]. Braat calculated the aberrations caused by a layer of glass with a refractive index of 1.5806 and a thickness of 600µm in the back pupil of a 0.6 NA objective with 650 nm wavelength. The results are shown in Table 1. Piston, Tip and Tilt were omitted since they won't reduce the intensity of the PSF and therefore would not degrade the imaging quality of the system. The Zernike polynomial notation from Born&Wolf [7] was adapted, as in reference [106]. In the Matlab programme, a ray-cone composed of 36 ray sections with an angular separation of $\pi/36$, each including 1021 rays, was traced towards the entrance pupil. The sine of the cone angle given by $\sin(\theta)$=NA/n, where n is the refractive index of the medium at the point source. For each simulation, it was ensured that the ray fan filled the objective's pupil, by altering the angles of the ray cone accordingly. The aberration were determined with respect to the fitted spherical reference wavefront.

Zernike Aberration	Braat	Ray Tracing Result	Ray Tracing with Shifted Reference Sphere
First order Spherical aberration	2.0454	-1.8735	-2.0454
Second order Spherical aberration	0.1793	-0.1678	-0.1793
Third order Spherical aberration	0.0154	-0.0141	-0.0149

Apart from the difference in sign, which is due to a different sign convention, the results are in good agreement. Differences in values can be explained by the different methods used for determining the aberrated wavefront. Aberration values close to the ones found by Braat were obtained by shifting the centre of the fitted reference sphere along the optical axis by -0.022mm. It is noted that our Ray tracing results, without a shift of the reference sphere, have lower values for spherical aberrations. This indicates that Braat's aberrations were not taken with respect to the best fit sphere.

Appendix B

Analytical expressions for the point spread function (PSF) exist for a limited number of cases. For an un-aberrated wavefront, an analytical solution for the 3 dimensional PSF was found by Lommel [7]. Nijboer and Zernike derived expressions for weakly aberrated (< 0.07 λ rms) focal spots around focus [9]. In most cases, however, aberrated PSFs can nowadays be computed numerically with a DFFT (discrete fast fourier transform). Braat and Janssen expanded the previous work of Nijboer and presented analytical solutions for calculating aberrated PSFs through focus [12] for high NA by taking polarisation into account. In the following a brief review of the scalar extended Nijboer-Zernike theory (ENZ), based on [12], will be presented. The scalar theory is at least sufficient for an NA up to 0.6. We recall that the pupil function, in cylindrical coordinates (ρ,θ), can be written as:

$$P(\rho,\theta) = A(\rho,\theta)e^{i\Phi(\rho,\theta)} \tag{B.1}$$

where A describes the amplitude distribution and Φ the phase distribution in the pupil.

Using the Huygens-Fresnel principle, the complex amplitude in the image plane is given by [11]:

$$U(r,\phi,f) = \frac{1}{\pi}\int_0^1 e^{if\rho^2}\rho\int_0^{2\pi} P(\rho,\theta)e^{i2\pi\rho r\cos(\theta-\phi)}\, d\theta \mathrm{d}\rho \tag{B.2}$$

where $f = -2\pi z$ is the defocus parameter, and (r,ϕ) are cylindrical coordinates in the image plane. The normalised axial optical coordinate is as follows: $z = \left[1-\sqrt{1-NA^2}\right]Z\big/\lambda$ which becomes $z = NA^2 Z/\lambda$ for low numerical apertures (NA) [11]. A wavefront, i.e. the phase distribution over a pupil, can be described in terms of Zernike polynomials (we consider here only the cosine terms):

$$\Phi(\rho,\theta)=\sum_{n,m}\alpha_n^m Z_n^m(\rho,\theta) \tag{B.3}$$

where α_n^m are Zernike coefficients and Z_n^m Zernike circle polynomials. The inner integral of equation (B.2) can be re-written, using a Taylor series expansion and assuming an uniform illumination $A(\rho,\theta)=1$, as:

$$\int_0^{2\pi} P(\rho,\theta)e^{i2\pi\rho r\cos(\theta-\phi)}d\theta=\int_0^{2\pi} e^{i\Phi(\rho,\theta)}e^{i2\pi\rho r\cos(\theta-\phi)}d\theta=\sum_{k=0}^{\infty}\frac{i^k}{k!}\int_0^{2\pi}\Phi^k(\rho,\theta)e^{i2\pi\rho r\cos(\theta-\phi)}d\theta \tag{B.4}$$

Using here the cosine-based expressions only was decided to improve the readability of the derivations (expressions related to the sine-based circle polynomials can be easily obtained). The last integral in (B.4) can be expanded using (B.3). By only considering cosine expressions of the form $\cos m\phi$ (i.e. replacing $\cos^m\phi$ terms by $\cos m\phi$ expressions) and by expanding the series up to the first order term $k=1$, which is allowed for small aberrations, equation (B.4) takes the form:

$$\int_0^{2\pi}\left[1+i\Phi(\rho,\theta)\right]e^{i2\pi\rho r\cos(\theta-\phi)}d\theta=\int_0^{2\pi} e^{i2\pi\rho r\cos(\theta-\phi)}\,d\theta+i\sum_{n,m}\alpha_n^m R_n^m(\rho)\int_0^{2\pi}\cos m\theta\cdot e^{i2\pi\rho r\cos(\theta-\phi)}d\theta \tag{B.5}$$

By using the following Bessel function identities [9]:

$$\begin{gathered}\int_0^{2\pi} e^{iqr\cos(\theta)}\cos(m\theta)d\theta=2\pi i^m J_m(qr)\\ \int_0^{2\pi} e^{iqr\cos(\theta)}\sin(m\theta)d\theta=0\end{gathered} \tag{B.6}$$

equation (B.5) can now be written as

$$\int_0^{2\pi}\left[1+i\Phi\left(\rho,\theta\right)\right]e^{i2\pi\rho r\cos(\theta-\phi)}d\theta=2\pi J_0\left(2\pi\rho r\right)+2\pi\sum_{n,m}\alpha_n^m i^{m+1}R_n^m\left(\rho\right)J_m\left(2\pi\rho r\right)\cos m\phi$$

(B.7)

and the complex amplitude in the image plane, equation (B.2), becomes [11]:

$$U\left(r,\phi,f\right)=2\int_0^1 e^{if\rho^2}\rho\, J_0\left(2\pi\rho r\right)\mathrm{d}\rho+2\sum_{n,m}\int_0^1 e^{if\rho^2}\rho\ \alpha_n^m i^{m+1}R_n^m\left(\rho\right)J_m\left(2\pi\rho r\right)\cos m\phi\ \mathrm{d}\rho$$

(B.8)

At focus $\left(f=0\right)$, $e^{if\rho^2}$ in (B.8) equals 1. This allowed Nijboer [9] to use the following formula:

$$\int_0^1 R_n^m(r)J_m(qr)rdr=(-1)^{\frac{n-m}{2}}\frac{J_{n+1}(q)}{q} \tag{B.9}$$

resulting for (B.8) in:

$$U\left(r,\phi,f\right)=2\sum_{n=0,m=0}^{\infty}\left(-1\right)^{\frac{n-m}{2}}\beta_n^m\frac{J_n\left(\upsilon\right)}{\upsilon}\cos m \quad \text{with } \upsilon=2\pi r \tag{B.10}$$

Where the β coefficients can be complex. Thus the amplitude in the focal plane can be represented as a weighted series of Bessel functions. This is referred to as the Nijboer-Zernike theory. In the extended Nijboer-Zernike theory compact series expressions were derived for describing the amplitude distribution in arbitrary focal planes. Janssen showed in reference [13] that the integral on the right hand side of equation (B.8) can be written in terms of a Bessel function series:

$$V_{nm}=\int_0^1\rho R_n^m\left(\rho\right)e^{if\rho^2}J_m\left(2\pi\rho r\right)\mathrm{d}\rho=e^{if}\sum_{l=1}^{\infty}\left(-2if\right)^{l-1}\sum_{j=0}^{p}v_{lj}\frac{J_{m+l+2j}\left(\upsilon\right)}{l\upsilon^l} \tag{B.11}$$

Where the v_{lj} coefficients are given by:

$$v_{lj} = (-1)^p (m+l+2j) \binom{m+j+l-1}{l-1} \binom{j+l-1}{l-1} \binom{l-1}{p-j} \Big/ \binom{q+l+j}{l}$$

(B.12)

The expressions in brackets represent binomial coefficients with $p = \frac{1}{2}(n-m)$ and $q = \frac{1}{2}(n+m)$. Consequently, the amplitude at an arbitrary focal plane f is given by:

$$U(r,\phi,f) = 2\sum_{n,m} \beta_n^m V_n^m \cos m\phi \qquad \text{(B.13)}$$

Where β are complex coefficients, which can be expressed in terms α_n^m. For large aberrations deriving the β coefficients from the α_n^m is not straightforward. It should thus be emphasised that equation B.5 is only valid for small aberrations. For large aberrations, the Taylor expansion of the phase function (equation B.3) should also include higher order terms.

The effect of aberrations on the amplitude (equation B.13) can be described in terms of Bessel functions. Amplitude variations, caused by primary aberrations, within the Airy disk are predominantly caused by Bessel functions up to order 6 [10]. The first 7 Bessel functions are illustrated in fig 1.B as well as the radial position at which the Airy Disk function has its first minimum (Airy radius).

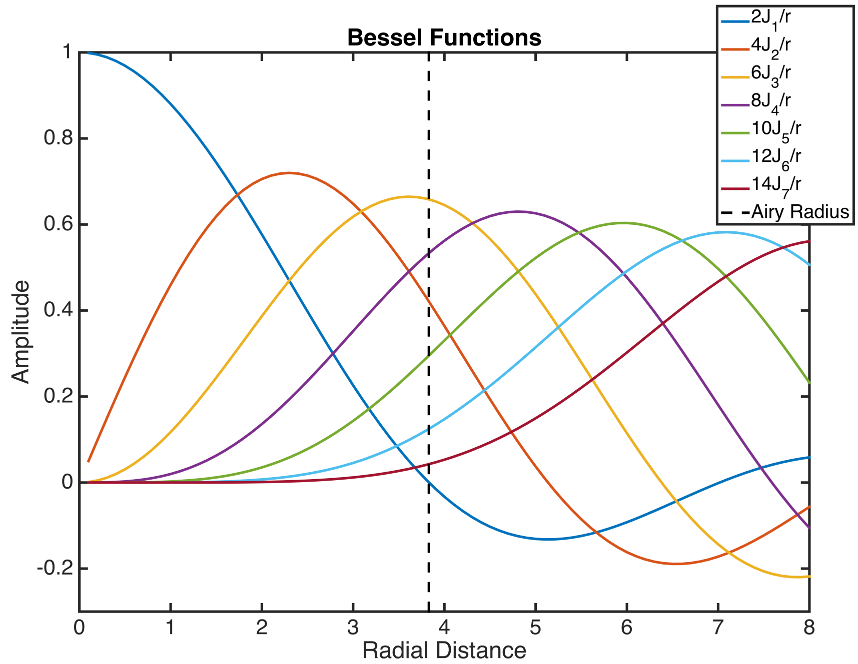


Figure 1.B: Bessel functions up to order 5 plotted as a function of the radial distance. The smallest radial distance (denoted as the Airy radius) at which the Airy Disk intensity becomes zero is marked by the dashed line.

$J_7(r)$ and higher order Bessel functions barely affect the amplitude within the Airy disk range For the analysis of aberrated PSFs, the Taylor series (equation B.3) was restricted to the second order terms $(k=2)$ as follows:

$$e^{i\Phi} \approx 1 + i\Phi - \frac{\Phi^2}{2!} \tag{B.14}$$

As shown in Chapter 6, such a restriction is allowable for aberration retrieval of aberrations in the range wherein we are interested. Expressions for PSFs with astigmatism, coma, spherical aberration will be derived below.

Focal spot with coma only

For coma $\Phi = \alpha_3^1 R_3^1(r)\cos(\phi) = \alpha_3^1(3r^3 - 2r)\cos(\phi)$ we have

$$e^{i\alpha_3^1 R_3^1(r)\cos\varphi} \approx 1 + i\alpha_3^1 R_3^1(r)\cos(\phi) - \frac{(\alpha_3^1)^2}{2!} A\left\{R_3^1(r)\right\}^2 \cos^2(\phi) \tag{B.15}$$

The quadratic term can be expanded in terms of the radial Zernike polynomials (see equations B.21) as follows:

$$\frac{\left(\alpha_3^1\right)^2}{2!}\left\{R_3^1\left(r\right)\right\}^2\cos^2\phi=\frac{\left(\alpha_3^1\right)^2}{2!}\left(9r^6-12r^4+4r^2\right)^2\frac{1+\cos 2\phi}{2}=$$
$$\frac{\left(\alpha_3^1\right)^2}{2\cdot 2!}\left\{\frac{9}{20}R_6^0\left(r\right)+\frac{1}{4}R_4^0\left(r\right)+\frac{1}{20}R_2^0\left(r\right)+\frac{1}{4}+\left(\frac{3}{5}R_6^2\left(r\right)+\frac{2}{5}R_2^2\left(r\right)\right)\cos\left(2\phi\right)\right\}$$

(B.16)

Plugging into equation (B.4):

$$\int_0^{2\pi}P\left(\rho,\theta\right)e^{i2\pi\rho r\cos(\phi)}d\theta=\sum_{k=0}^{2}\frac{i^k}{k!}\int_0^{2\pi}\Phi^k\left(\rho,\theta+\phi\right)e^{i2\pi\rho r\cos(\phi)}d\theta$$
$$=\int_0^{2\pi}\left[1+i\alpha_3^1R_3^1\left(r\right)\cos\left(\phi\right)-\frac{\left(\alpha_3^1\right)^2}{2\cdot 2!}\left\{\frac{1}{4}+\frac{1}{20}R_2^0\left(r\right)+\frac{1}{4}R_4^0\left(r\right)+\frac{9}{20}R_6^0\left(r\right)+\left(\frac{3}{5}R_6^2\left(r\right)+\frac{2}{5}R_2^2\left(r\right)\right)\cos\left(2\phi\right)\right\}\right]e^{i2\pi\rho r\cos(\phi)}d\theta$$

(B.17)

Thus for the amplitude, in terms of Bessel functions:

$$U_{coma}\approx e^{if}\begin{bmatrix}2\left(1-0.063\left(\alpha_3^1\right)^2\right)\left[\frac{J_1\left(r\right)}{r}+\ldots\right]+0.025\left(\alpha_3^1\right)^2\left[\frac{J_3\left(r\right)}{r}+\ldots\right]+0.2\left(\alpha_3^1\right)^2\cos\left(2\phi\right)\left[\frac{J_3\left(r\right)}{r}+\ldots\right]-\\2\left(\alpha_3^1\right)\cos\left(\phi\right)\left[\frac{J_4\left(r\right)}{r}+\ldots\right]-0.125\left(\alpha_3^1\right)^2\left[\frac{J_5\left(r\right)}{r}+\ldots\right]\end{bmatrix}$$

(B.18)

The complex amplitude, expressed in terms of Bessel functions, is obtained by using the formulae (B.6) and (B.11). The series were expanded using Matlab Symbolic routines. To assess the number of term necessary for the series expansion (B.11), a value of $f=4\pi\sqrt{3}\cdot\left(0.071\lambda\ \text{rms}\right)\approx 1.55$ for defocus (which corresponds to a Strehl ratio of 0.8) was used and the Strehl ratio was computed. A Strehl ratio of 0.8 was obtained when expanding the Bessel series (B.11) up to $l=5$. Bessel functions can in turn be expanded into polynomial series [107]:

$$J_{\upsilon}(z)=\left(\frac{z}{2}\right)^{\upsilon}\sum_{k=0}^{\infty}\frac{1}{k!\Gamma(k+\upsilon+1)}\left(\frac{-z^2}{2}\right)^{k} \tag{B.19}$$

Where Γ is the gamma function. Since we are interested in the Intensity distribution within the Airy disk region, we will replace the Bessel functions by their serial expansion up to the order r^8. The difference between the Airy spot intensity and the truncated series representation up to radial order 8 terms is shown in Figure 2.B. Thus, the residual error is less than 0.005.

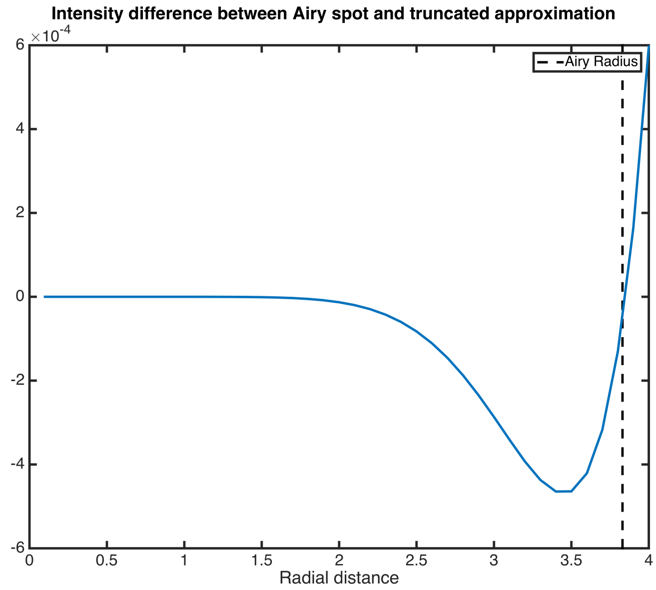


Figure 2.B: Intensity difference between Airy spot and truncated approximation. The smallest radial distance (denoted as the Airy radius) at which the Airy Disk intensity becomes zero is marked by the dashed line.

The series expansions of the first five Bessel functions of the first kind are:

$$
\begin{aligned}
\frac{J_1(r)}{r} &\approx \frac{1}{2}-\frac{r^2}{16}+\frac{r^4}{384}-\frac{r^6}{18432}+\frac{r^8}{1474560}+\ldots\\
\frac{J_2(r)}{r} &\approx \frac{r}{8}-\frac{r^3}{96}+\frac{r^5}{3072}-\frac{r^7}{184320}+\ldots\\
\frac{J_3(r)}{r} &\approx \frac{r^2}{48}-\frac{r^4}{768}+\frac{r^6}{30720}-\frac{r^8}{2211840}+\ldots\\
\frac{J_4(r)}{r} &\approx \frac{r^3}{384}-\frac{r^5}{7680}+\frac{r^7}{368640}+\ldots\\
\frac{J_5(r)}{r} &\approx \frac{r^4}{3840}-\frac{r^6}{92160}+\frac{r^8}{5160960}+\ldots
\end{aligned}
\tag{B.20}
$$

Expressing the Bessel series, in terms of the radial Zernike polynomials (see Appendix A) the following expressions were used:

$$
\begin{aligned}
1 &= R_0^0 \\
r &= R_1^0 \cos(\phi) \\
r^2 &= \frac{1}{2}\left(R_2^0 + R_0^0\right) \\
r^2 \cos(2\phi) &= R_2^2 \cos(2\phi) \\
r^3 \cos(\phi) &= \frac{1}{3}\left(R_3^1 + 2R_1^0\right)\cos(\phi) \\
r^3 \cos(3\phi) &= R_3^3 \cos(3\phi) \\
r^4 &= \frac{1}{6}\left(R_4^0 + 3R_2^0 + 2R_0^0\right) \\
r^4 \cos(2\phi) &= \frac{1}{4}\left(R_4^2 + 3R_2^2\right)\cos(2\phi) \\
r^4 \cos(4\phi) &= R_4^4 \cos(4\phi) \\
r^5 \cos(\phi) &= \frac{1}{10}\left(R_5^1 + 4R_3^1 + 5R_1^0\right)\cos(\phi) \\
r^6 &= \frac{1}{20}\left(R_6^0 + 5R_4^0 + 9R_2^0 + 5R_0^0\right) \\
r^6 \cos(2\phi) &= \frac{1}{15}\left(R_6^2 + 5R_4^2 + 9R_2^2\right)\cos(2\phi) \\
r^6 \cos(4\phi) &= \frac{1}{6}\left(R_6^4 + 5R_4^4\right)\cos(4\phi) \\
r^7 \cos(\phi) &= \frac{1}{35}\left(R_7^1 + 6R_5^1 + 14R_3^1 + 14R_1^0\right)\cos(\phi) \\
r^8 &= \frac{1}{70}\left(R_8^0 + 7R_6^0 + 20R_4^0 + 28R_2^0 + 14R_0^0\right) \\
r^8 \cos(2\phi) &= \frac{1}{56}\left(R_8^2 + 12R_6^2 + 20R_4^2 + 28R_2^2\right)\cos(2\phi) \\
r^8 \cos(4\phi) &= \frac{1}{28}\left(R_8^4 + 7R_6^4 + 20R_4^4\right)\cos(4\phi) \\
r^9 \cos(\phi) &= \frac{1}{126}\left(R_9^1 + 8R_7^1 + 27R_5^1 + 48R_3^1 + 18R_1^0\right)\cos(\phi) \\
r^{10} &= \frac{1}{252}\left(R_{10}^0 + 9R_8^0 + 35R_6^0 + 75R_4^0 + 90R_2^0 + 42R_0^0\right) \\
&\cdots
\end{aligned}
\tag{B.21}
$$

By further substituting $r' = 3.832r$ the radius is expressed in Airy units (AU, i.e. for the first dark ring r' equals 1) and by multiplying the amplitude with its complex conjugate, we obtain the intensity distribution (omitting f terms of power 5 and higher):

$$I_{coma}(r,\phi;\alpha_3^1,f)=\begin{bmatrix} \{(0.228-0.013f^2+(2.42\cdot10^{-4})f^4)+(-0.025+0.003f^2-(4.82\cdot10^{-5})f^4)(\alpha_3^1)^2\}\mathrm{R}_0^0+ \\ \{(0.040+0.010f^2-(3.29\cdot10^{-4})f^4)(\alpha_3^1)+((1.16\cdot10^{-4})+0.001f^2-(1.75\cdot10^{-5})f^4)(\alpha_3^1)^3\}\mathrm{R}_1^1(r)\cos\phi+ \\ \{(-0.419+0.041f^2-0.002f^4)+(0.057-0.007f^2+(3.23\cdot10^{-4})f^4)(\alpha_3^1)^2\}\mathrm{R}_2^0(r)+ \\ \{0.019-(6.59\cdot10^{-4})f^2-(7.7\cdot10^{-5})f^4\}(\alpha_3^1)^2\,\mathrm{R}_2^2(r)\cos2\phi+ \\ \{-(0.031+0.008f^2-(4.12\cdot10^{-4})f^4)(\alpha_3^1)+(0.003+(3.72\cdot10^{-4})f^2-(2.65\cdot10^{-5})f^4)(\alpha_3^1)^3\}\mathrm{R}_3^1(r)\cos\phi+ \\ \{-0.001f^2+(6.1\cdot10^{-5})f^4\}(\alpha_3^1)^3\,\mathrm{R}_3^3(r)\cos3\phi+ \\ \{-0.014-(2.21\cdot10^{-4})f^2+(4.9\cdot10^{-5})f^4\}(\alpha_3^1)^2\,\mathrm{R}_4^2(r)\cos2\phi+ \\ \{(0.257-0.024f^2-(8.07\cdot10^{-4})f^4)+(-0.031+0.003f^2+(9.93\cdot10^{-5})f^4)(\alpha_3^1)^2\}\mathrm{R}_4^0(r)+ \\ \{(-0.015+.001f^2-(1.72\cdot10^{-4})f^4)(\alpha_3^1)+((4.9\cdot10^{-4})+(1.41\cdot10^{-4})f^2+(3.13\cdot10^{-6})f^4)(\alpha_3^1)^3\}\mathrm{R}_5^1(r)\cos\phi \end{bmatrix}$$

(B.22)

Focal spot with astigmatism only

The amplitude distribution of an astigmatic focal spot is given by [13]:

$$U(r,\phi)=\sum_{l=0}^{\infty}D_l\left(\alpha_2^2\right)^l \tag{B.23}$$

where α_2^2 is the amplitude of astigmatism in the pupil plane and (r,ϕ) are cylindrical coordinates in the image plane. The function D is given by:

$$D_{2m}=2\left(-\frac{1}{4}\right)^m\sum_{k=0}^{m}\varepsilon_k\frac{T_{4m,4k}}{(m-k)!(m-k)!}\cos4k\phi$$

(B.24)

$$D_{2m+1}=2i\left(-\frac{1}{4}\right)^m\sum_{k=0}^{m}\varepsilon_k\frac{T_{4m+2,4k+2}}{(m-k)!(m+k+1)!}\cos2(2k+1)\phi$$

Where ε is the Neumann symbol $\left(\varepsilon_0 = 1,\ \varepsilon_n = 1 \text{ for n} \neq 0\right)$. The T function is given by the following Bessel series expansion:

$$T_{n,m} = e^{if} \sum_{l=1}^{\infty} \left(-2if\right)^{l-1} \sum_{j=0}^{p} t_{lj} \frac{J_{m+l+2j}\left(\rho\right)}{\rho^{l}} \tag{B.25}$$

where f represents the defocus parameter, $\rho = 2\pi\sqrt{x^2+y^2}$ the radial coordinate in the image plane $J_{m+l+2j}\left(\rho\right)$ are Bessel functions of the first kind and

$$t_{lj} = \left(-1\right)^{j} \frac{m+l+2j}{q+1} \begin{pmatrix} p \\ j \end{pmatrix} \begin{pmatrix} m+j+l-1 \\ l-1 \end{pmatrix} \Big/ \begin{pmatrix} q+l+j \\ q+1 \end{pmatrix} \tag{B.26}$$

with

$$\begin{pmatrix} p \\ j \end{pmatrix} = \frac{p!}{j!(p-j)!} = \frac{p\left(p-1\right)\cdots\left(p-j-1\right)}{j!} \tag{B.27}$$

and $p = \frac{1}{2}\left(n-m\right),\ q = \frac{1}{2}\left(n+m\right)$. Analogously to the derivation of the coma aberrated spot, we perform the substitutions (B.17 to B.21) and obtain for the intensity distribution of the astigmatic spot (omitting f terms of power 5 and higher):

$$I_{astig}\left(r,\phi;f,\alpha_2^2\right) \approx \begin{bmatrix} \left\{\left(0.228-0.013f^2+\left(2.42\cdot10^{-4}\right)f^4\right)-\left(0.012-0.001f^2-\left(1.00\cdot10^{-4}\right)f^4\right)\left(\alpha_2^2\right)^2\right\}\mathrm{R}_0^0+ \\ \left\{\left(-0.419+0.041f^2+0.006f^4\right)+\left(0.078-0.008f^2-\left(5.53\cdot10^{-4}\right)f^4\right)\left(\alpha_2^2\right)^2\right\}\mathrm{R}_2^0(r)+ \\ \left\{\left(0.081-0.006f^2\right)\left(f\alpha_2^2\right)+\left(-0.001+\left(8.45\cdot10^{-5}\right)f^2\right)f\left(\alpha_2^2\right)^3\right\}\mathrm{R}_2^2(r)\cos 2\phi+ \\ \left\{\left(0.257-0.024f^2-0.001f^4\right)+\left(-0.057+0.005f^2+\left(1.51\cdot10^{-4}\right)f^4\right)\left(\alpha_2^2\right)^2\right\}\mathrm{R}_4^0(r)+ \\ \left\{\left(-0.048+0.003f^2\right)f\left(\alpha_2^2\right)+\left(\left(4.01\cdot10^{-4}\right)-\left(6.26\cdot10^{-6}\right)f^2\right)f\left(\alpha_2^2\right)^3\right\}\mathrm{R}_4^2(r)\cos 2\phi+ \\ \left\{\left(0.033+\left(2.62\cdot10^{-4}\right)f^2-\left(1.22\cdot10^{-4}\right)f^4\right)\left(\alpha_2^2\right)^2\right\}\mathrm{R}_4^4(r)\cos 4\phi \end{bmatrix}$$

(B.28)

Focal spot with spherical aberration only

For spherical aberration $\Phi = \alpha_4^0 R_4^0(r) = \alpha_4^0(6r^4 - 6r^2 + 1)$ we have:

$$e^{i\alpha_4^0 R_4^0(r)\cos\varphi} \approx 1 + i\alpha_4^0 R_4^0(r) - \frac{(\alpha_4^0)^2}{2!} A\left\{R_4^0(r)\right\}^2 \tag{B.29}$$

The quadratic term can be expanded in terms of the radial Zernike polynomials (see equations B.21) as follows:

$$\frac{(\alpha_4^0)^2}{2!}\left\{R_4^0(r)\right\}^2 = \frac{(\alpha_4^0)^2}{2!}\left(\frac{1}{5}R_0^0 + \frac{2}{7}R_4^0 + \frac{18}{35}R_8^0\right) \tag{B.30}$$

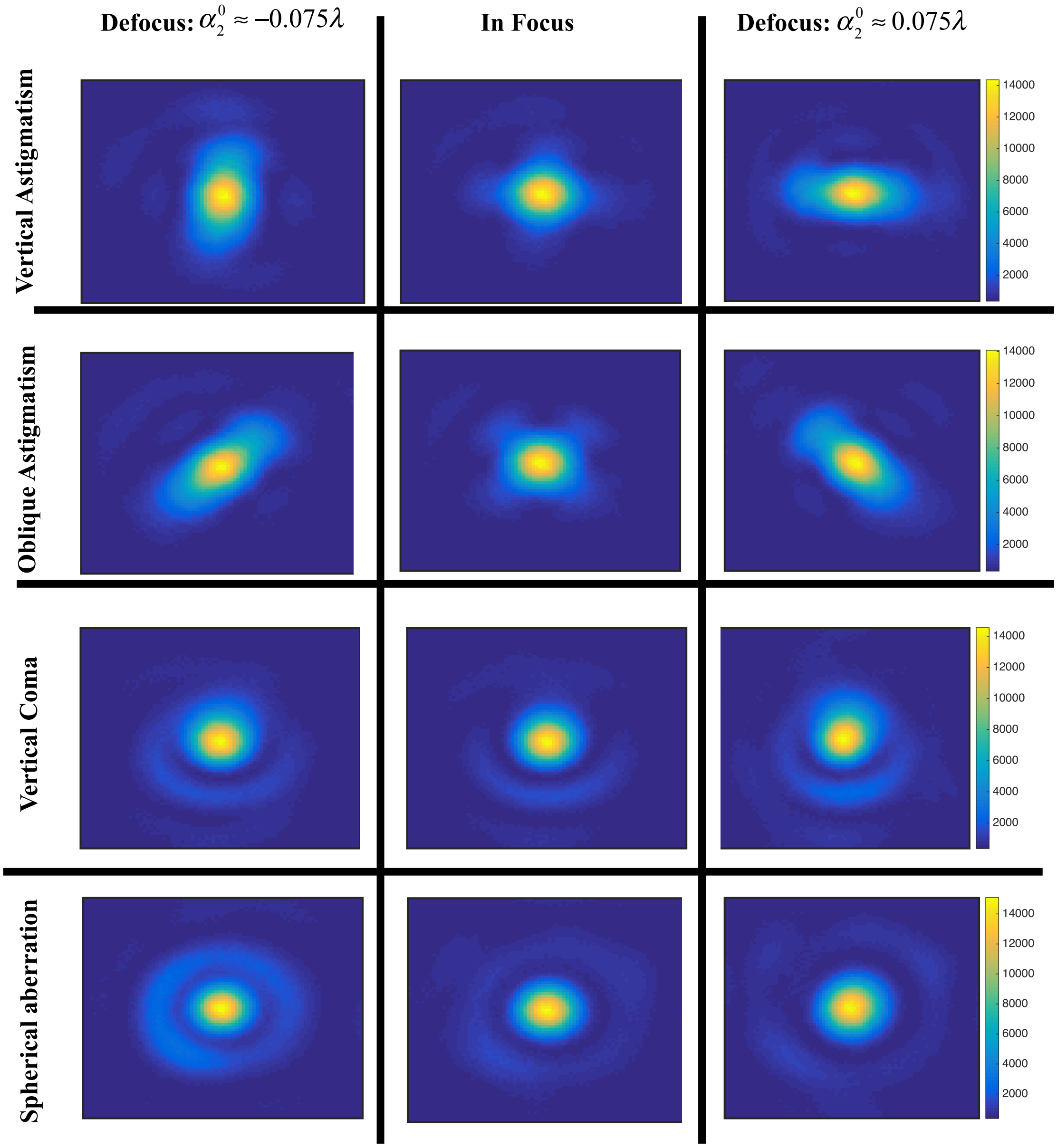
With α_4^0 being the spherical aberration amplitude. Like in the derivation of the coma aberrated spot, we perform the substitutions (B.17 to B.21) and obtain for the intensity distribution of the spherical aberrated spot (omitting f terms of power 5 and higher):

$$I_{spherical}(r,\phi;f,\alpha_4^0) \approx \left[\begin{array}{l}
\left\{\begin{array}{l}
\left(0.228 - 0.012 f^2 + (2.99\cdot 10^{-4}) f^4\right) + \\
\left(0.013 - (8.55\cdot 10^{-4}) f^2 + (1.65\cdot 10^{-5}) f^4\right) f(\alpha_4^0) + \\
\left(-0.0459 + 0.004 f^2 - 0.001 f^4\right)(\alpha_4^0)^2 + \\
\left(-0.001 + (7.99\cdot 10^{-5}) f^2 + (2.23\cdot 10^{-6}) f^4\right) f(\alpha_4^0)^3
\end{array}\right\} R_0^0 + \\
\left\{\begin{array}{l}
\left(-0.419 + 0.0422 f^2 - 0.002 f^4\right) + \\
\left(-0.017 + (2.28\cdot 10^{-4}) f^2 - (6.68\cdot 10^{-4}) f^4\right) f(\alpha_4^0) + \\
\left(0.085 - 0.010 f^2 + (1.44\cdot 10^{-4}) f^4\right)(\alpha_4^0)^2 + \\
\left(+0.002 - (3.18\cdot 10^{-5}) f^2 - (6.27\cdot 10^{-5}) f^4\right) f(\alpha_4^0)^3
\end{array}\right\} \mathrm{R}_2^0(r) + \\
\left\{\begin{array}{l}
\left(0.258 - 0.023 f^2 + 0.002 f^4 - (1.09\cdot 10^{-4}) f^6\right) + \\
\left(-0.017 + (7.28\cdot 10^{-4}) f^2 - (2.81\cdot 10^{-4}) f^4\right) f(\alpha_4^0) + \\
\left(-0.049 + 0.006 f^2 - (2.77\cdot 10^{-4}) f^4\right)(\alpha_4^0)^2 + \\
\left(0.002 - (7.61\cdot 10^{-5}) f^2 + (2.8\cdot 10^{-5}) f^4\right) f(\alpha_4^0)^3
\end{array}\right\} \mathrm{R}_4^0(r)
\end{array}\right]$$

(B.31)

Appendix C

Intensity distributions in and out of focus in the presence of primary aberrations



Bibliography

1. Rochow, T.G. and P.A. Tucker, *A Brief History of Microscopy*, in *Introduction to Microscopy by Means of Light, Electrons, X Rays, or Acoustics*, T.G. Rochow and P.A. Tucker, Editors. 1994, Springer US: Boston, MA. p. 1-21.

2. Lane, N., *The unseen world: reflections on Leeuwenhoek (1677) 'Concerning little animals'.* Philosophical transactions of the Royal Society of London. Series B, Biological sciences, 2015. **370**(1666): p. 20140344.

3. Abbe, E., Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung. Archiv für mikroskopische Anatomie, 1873. **9**(1): p. 413-418.

4. Hell, S.W. and J. Wichmann, Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. Opt Lett, 1994. **19**(11): p. 780-2.

5. Gustafsson, M.G., Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy. J Microsc, 2000. **198**(Pt 2): p. 82-7.

6. Betzig, E., G.H. Patterson, R. Sougrat, O.W. Lindwasser, S. Olenych, J.S. Bonifacino, M.W. Davidson, J. Lippincott-Schwartz, and H.F. Hess, *Imaging intracellular fluorescent proteins at nanometer resolution.* Science, 2006. **313**(5793): p. 1642-5.

7. Born, M., E. Wolf, and A.B. Bhatia, Principles of optics : electromagnetic theory of propagation, interference and diffraction of light. 7th (expanded) ed. 1999, Cambridge: Cambridge University Press. xxxiii, 952 p.

8. Nijboer, B.R.A., The diffraction theory of optical aberrations: Part I: General discussion of the geometrical aberrations. Physica, 1943. **10**(8): p. 679-692.

9. Nijboer, B.R.A., The diffraction theory of optical aberrations: Part II: Diffraction pattern in the presence of small aberrations. Physica, 1947. **13**(10): p. 605-620.

10. Zernike, A.F., Diffraction theory of the knife-edge test and its improved form, the phase-contrast method*. Journal of Micro/Nanolithography, MEMS, and MOEMS, 2002. **1**(2): p. 87-94.

11. Braat, J., Dirksen, P., van Haver, S., Janssen, A., Detailed description of the ENZ approach, http://www.nijboerzernike.nl.

12. Braat, J., P. Dirksen, and A.J. Janssen, Assessment of an extended Nijboer-Zernike approach for the computation of optical point-spread functions. J Opt Soc Am A Opt Image Sci Vis, 2002. **19**(5): p. 858-70.

13. Janssen, A.J., Extended Nijboer-Zernike approach for the computation of optical point-spread functions. J Opt Soc Am A Opt Image Sci Vis, 2002. **19**(5): p. 849-57.

14. Booth, M.J., *Adaptive optics in microscopy.* Philos Trans A Math Phys Eng Sci, 2007. **365**(1861): p. 2829-43.

15. Hopkins, H.H., Diffraction theory of laser read-out systems for optical video discs. Journal of the Optical Society of America, 1979. **69**(1): p. 4-24.

16. Goodman, J.W., *Introduction to Fourier optics*. 3rd ed. 2005, Englewood, Co.: Roberts & Company. xviii, 491 p.

17. Hecht, E., *Optics*. Fifth edition / ed. 2017, Harlow, England: Pearson Education. 722 pages.

18. Noll, R.J., *Zernike polynomials and atmospheric turbulence*.* Journal of the Optical Society of America, 1976. **66**(3): p. 207-211.

19. Malacara, D., *Optical shop testing*. 3rd ed. Wiley series in pure and applied optics. 2007, Hoboken, N.J.: Wiley-Interscience. xx, 862 p.

20. Wilson, T. and C. Sheppard, *Theory and practice of scanning optical microscopy*. 1984, London ; Orlando: Academic Press. ix, 213 p.

21. Minsky, M., *Microscopy apparatus*. 1961, Google Patents.

22. Pawley, J.B., *Handbook of biological confocal microscopy*. 2nd ed ed. Language of science. 1995, New York: Plenum Press. xxiii, 632 p., 4 p. of plates.

23. Wilson, T., Resolution and optical sectioning in the confocal microscope. Journal of Microscopy, 2011. **244**(2): p. 113-121.

24. M.J., B., N. M.A.A., and W. T., Aberration correction for confocal imaging in refractive−index−mismatched media. Journal of Microscopy, 1998. **192**(2): p. 90-98.

25. Egner, A. and S. Hell, Aberrations in Confocal and Multi-Photon Fluorescence Microscopy Induced by Refractive Index Mismatch, in Handbook Of Biological Confocal Microscopy, J.B. Pawley, Editor. 2006, Springer US. p. 404-413.

26. Glindemann, A., S. Hippler, T. Berkefeld, and W. Hackenberg, *Adaptive Optics on Large Telescopes.* Experimental Astronomy, 2000. **10**(1): p. 5-47.

27. Tyson, R.K., *Introduction to adaptive optics Robert K. Tyson*, in *Tutorial texts in optical engineering TT 41*. 2000, SPIE Press,: Bellingham, WA. p. 1 online resource (ix, 116 p.).

28. Babcock, H.W., *THE POSSIBILITY OF COMPENSATING ASTRONOMICAL SEEING.* Publications of the Astronomical Society of the Pacific, 1953. **65**(386): p. 229-236.

29. Happer, W., G.J. MacDonald, C.E. Max, and F.J. Dyson, Atmospheric-turbulence compensation by resonant optical backscattering from the sodium layer in the upper atmosphere. Journal of the Optical Society of America A, 1994. **11**(1): p. 263-276.

30. Beckers, J.M., *Increasing the Size of the Isoplanatic Patch with Multiconjugate Adaptive Optics.* Very Large Telescopes and their Instrumentation, ESO Conference and Workshop Proceedings, Proceedings of a ESO Conference on Very Large Telescopes and their Instrumentation, held in Garching, March 21-24, 1988, Garching: European Southern Observatory (ESO), 1988, edited by Marie-Helene Ulrich., p.693, 1988.

31. Manzanera, S., M.A. Helmbrecht, C.J. Kempf, and A. Roorda, MEMS segmented-based adaptive optics scanning laser ophthalmoscope. Biomed Opt Express, 2011. **2**(5): p. 1204-17.

32. Kubby, J.A., *Adaptive optics for biological imaging*. 2013, Taylor & Francis,: Boca Raton. p. xvi, 359 p.

33. Fernandez, E.J., L. Vabre, B. Hermann, A. Unterhuber, B. Povazay, and W. Drexler, *Adaptive optics with a magnetic deformable mirror: applications in the human eye.* Opt Express, 2006. **14**(20): p. 8900-17.

34. Shaw, M., S. Hall, S. Knox, R. Stevens, and C. Paterson, Characterization of deformable mirrors for spherical aberration correction in optical sectioning microscopy. Opt Express, 2010. **18**(7): p. 6900-13.

35. Devaney, N., D. Coburn, C. Coleman, J.C. Dainty, E. Dalimier, T. Farrell, D. Lara, D. Mackey, and R. Mackey. *Characterisation of MEMs mirrors for use in atmospheric and ocular wavefront correction*. in *MOEMS-MEMS 2008 Micro and Nanofabrication*. 2008. SPIE.

36. Schwertner, M., M.J. Booth, M.A.A. Neil, and T. Wilson, Measurement of specimen-induced aberrations of biological samples using phase stepping interferometry. Journal of Microscopy, 2004. **213**(1): p. 11-19.

37. Schwertner, M., M. Booth, and T. Wilson, Characterizing specimen induced aberrations for high NA adaptive optical microscopy. Opt Express, 2004. **12**(26): p. 6540-52.

38. Bates, W.J., *A wavefront shearing interferometer.* Proceedings of the Physical Society, 1947. **59**(6): p. 940.

39. Feierabend, M., M. Ruckel, and W. Denk, Coherence-gated wave-front sensing in strongly scattering samples. Opt Lett, 2004. **29**(19): p. 2255-7.

40. Platt, B.C. and R. Shack, History and principles of Shack-Hartmann wavefront sensing. J Refract Surg, 2001. **17**(5): p. S573-7.

41. Beverage, J.L., R.V. Shack, and M.R. Descour, Measurement of the three-dimensional microscope point spread function using a Shack-Hartmann wavefront sensor. J Microsc, 2002. **205**(Pt 1): p. 61-75.

42. Azucena, O., J. Crest, J. Cao, W. Sullivan, P. Kner, D. Gavel, D. Dillon, S. Olivier, and J. Kubby, Wavefront aberration measurements and corrections through thick tissue using fluorescent microsphere reference beacons. Optics Express, 2010. **18**(16): p. 17521-17532.

43. Tao, X., O. Azucena, M. Fu, Y. Zuo, D.C. Chen, and J. Kubby, Adaptive optics microscopy with direct wavefront sensing using fluorescent protein guide stars. Opt Lett, 2011. **36**(17): p. 3389-91.

44. Shaw, M., K. O'Holleran, and C. Paterson, Investigation of the confocal wavefront sensor and its application to biological microscopy. Opt Express, 2013. **21**(16): p. 19353-62.

45. Poyneer, L.A., Scene-based Shack-Hartmann wave-front sensing: analysis and simulation. Applied Optics, 2003. **42**(29): p. 5807-5815.

46. Lawrence, K., Y. Liu, R. Ball, A.J. VanLeuven, J.D. Lauderdale, and P. Kner. Scene-based Shack-Hartmann Wavefront Sensor for Light-Sheet Microscopy. in Imaging and Applied Optics 2018 (3D, AO, AIO, COSI, DH, IS, LACSEA, LS&C, MATH, pcAOP). 2018. Orlando, Florida: Optical Society of America.

47. Neil, M.A., M.J. Booth, and T. Wilson, Closed-loop aberration correction by use of a modal Zernike wave-front sensor. Opt Lett, 2000. **25**(15): p. 1083-5.

48. Booth, M.J., M.A. Neil, R. Juskaitis, and T. Wilson, *Adaptive aberration correction in a confocal microscope.* Proc Natl Acad Sci U S A, 2002. **99**(9): p. 5788-92.

49. Facomprez, A., E. Beaurepaire, and D. Debarre, Accuracy of correction in modal sensorless adaptive optics. Opt Express, 2012. **20**(3): p. 2598-612.

50. Booth, M.J., Wavefront sensorless adaptive optics for large aberrations. Opt Lett, 2007. **32**(1): p. 5-7.

51. Debarre, D., M.J. Booth, and T. Wilson, Image based adaptive optics through optimisation of low spatial frequencies. Opt Express, 2007. **15**(13): p. 8176-90.

52. Debarre, D., E.J. Botcherby, M.J. Booth, and T. Wilson, *Adaptive optics for structured illumination microscopy.* Opt Express, 2008. **16**(13): p. 9290-305.

53. Wright, A.J., D. Burns, B.A. Patterson, S.P. Poland, G.J. Valentine, and J.M. Girkin, Exploration of the optimisation algorithms used in the implementation of adaptive optics in confocal and multiphoton microscopy. Microsc Res Tech, 2005. **67**(1): p. 36-44.

54. Wright, A.J., S.P. Poland, J.M. Girkin, C.W. Freudiger, C.L. Evans, and X.S. Xie, *Adaptive optics for enhanced signal in CARS microscopy.* Opt Express, 2007. **15**(26): p. 18209-19.

55. Müllenbroich, M.C., N. McAlinden, and A.J. Wright, Adaptive optics in an optical trapping system for enhanced lateral trap stiffness at depth. Journal of Optics, 2013. **15**(7): p. 075305.

56. Millane, R.P., *Phase retrieval in crystallography and optics.* Journal of the Optical Society of America A, 1990. **7**(3): p. 394-411.

57. Fienup, J.R., J.C. Marron, T.J. Schulz, and J.H. Seldin, Hubble Space Telescope characterized by using phase-retrieval algorithms. Appl Opt, 1993. **32**(10): p. 1747-67.

58. Hanser, B.M., M.G. Gustafsson, D.A. Agard, and J.W. Sedat, *Phase retrieval for high-numerical-aperture optical systems.* Opt Lett, 2003. **28**(10): p. 801-3.

59. Hanser, B.M., M.G. Gustafsson, D.A. Agard, and J.W. Sedat, Phase-retrieved pupil functions in wide-field fluorescence microscopy. J Microsc, 2004. **216**(Pt 1): p. 32-48.

60. FOREMAN, M.R., C.L. GIUSCA, P. TÖRÖK, and R.K. LEACH, Phase-retrieved pupil function and coherent transfer function in confocal microscopy. Journal of Microscopy, 2013. **251**(1): p. 99-107.

61. Kner, P., L. Winoto, D.A. Agard, and J.W. Sedat, Closed loop adaptive optics for microscopy without a wavefront sensor. Proc SPIE Int Soc Opt Eng, 2010. **7570**.

62. Gerchberg, R.W., Saxton, W. O., A practical algorithm for the determination of phase from image and diffraction plane pictures. Optik, 1972. **35**(2): p. 237–250.

63. Fienup, J.R., Phase retrieval algorithms: a personal tour [Invited]. Appl Opt, 2013. **52**(1): p. 45-56.

64. Shechtman, Y., Eldar, Y.C., Cohen, O., Chapman, H.N., Miao, J., Segev, M., *Phase Retrieval with Application to Optical Imaging.* IEEE Signal Processing Magazine, 2015.

65. Gonsalves, R.A. Phase Retrieval And Diversity In Adaptive Optics. 1982. SPIE.

66. Dean, B.H., D.L. Aronstein, J.S. Smith, R. Shiri, and D.S. Acton. *Phase retrieval algorithm for JWST Flight and Testbed Telescope*. in *SPIE Astronomical Telescopes + Instrumentation*. 2006. SPIE.

67. Mugnier, L.M., A. Blanc, and J. Idier, Phase Diversity: A Technique for Wave-Front Sensing and for Diffraction-Limited Imaging, in Advances in Imaging and Electron Physics, P. Hawkes, Editor. 2006, Elsevier. p. 1-76.

68. Braat, J.J., P. Dirksen, A.J. Janssen, and A.S. van de Nes, Extended Nijboer-Zernike representation of the vector field in the focal region of an aberrated high-aperture optical system. J Opt Soc Am A Opt Image Sci Vis, 2003. **20**(12): p. 2281-92.

69. Braat, J.J.M., S. van Haver, A.J.E.M. Janssen, and P. Dirksen, Chapter 6 Assessment of optical systems by means of point-spread functions, in Progress in Optics, E. Wolf, Editor. 2008, Elsevier. p. 349-468.

70. Dirksen, P., J.J.M. Braat, A. Janssen, and C.A.H. Juffermans. *Aberration retrieval using the extended Nijboer-Zernike approach*. 2003. SPIE.

71. Larsson, U., Characterization of a micromachined membrane deformable mirror. Foi Defense Research Agency. http://www.foi.se/en/Search/Abstract/?rNo=FOI-R--1917--SE, 2006.

72. Fernandez, E. and P. Artal, Membrane deformable mirror for adaptive optics: performance limits in visual optics. Opt Express, 2003. **11**(9): p. 1056-69.

73. Maréchal, A., Étude des effets combinés de la diffraction et des aberrations géométriques sur l'image d'un point lumineux. Rev. Opt, 1947. **26**: p. 257-277.

74. Polo, A., A. Haber, S.F. Pereira, M. Verhaegen, and H.P. Urbach, An innovative and efficient method to control the shape of push-pull membrane deformable mirror. Optics Express, 2012. **20**(25): p. 27922-27932.

75. See, C.W., F. Hu, C.J. Chuang, and M.G. Somekh, Super-resolution imaging using proximity projection grating and structured light illumination. Opt Express, 2013. **21**(13): p. 15155-67.

76. Kenneth R. Spring, T.J.F.a.M.W.D. *Introduction to Charge-Coupled Devices (CCDs)*. Available from: https://www.microscopyu.com/digital-imaging/introduction-to-charge-coupled-devices-ccds.

77. Whirter, I.M. Electron Multiplying CCDs - New Technology for Low Light Level Imaging. in Proceedings of 33rd Annual European Meeting on Atmospheric Studies by Optical Methods. 2008.

78. Dussault, D. and P. Hoess. Noise performance comparison of ICCD with CCD and EMCCD cameras. in Optical Science and Technology, the SPIE 49th Annual Meeting. 2004. SPIE.

79. Denvir, D.J. and E. Conroy. Electron-multiplying CCD: the new ICCD. in International Symposium on Optical Science and Technology. 2003. SPIE.

80. Gomez-Vieyra, A. and D. Malacara-Hernández, Geometric theory of wavefront aberrations in an off-axis spherical mirror. Applied Optics, 2011. **50**(1): p. 66-73.

81. Raja Bayanna, A., R.E. Louis, S. Chatterjee, S.K. Mathew, and P. Venkatakrishnan, *Membrane-based deformable mirror: intrinsic aberrations and alignment issues.* Applied Optics, 2015. **54**(7): p. 1727-1736.

82. Driggers, R.G., Encyclopedia of Optical Engineering: Pho-Z, pages 2049-3050. 2003: Marcel Dekker.

83. Marshall, G.F., *Optical scanning*. Optical engineering. 1991, New York: Marcel Dekker, Inc. xx, 868 p.

84. Sheppard, C.J. and M. Gu, Aberration compensation in confocal microscopy. Appl Opt, 1991. **30**(25): p. 3563-8.

85. *Zeiss Digital Imaging*. Available from: http://zeiss-campus.magnet.fsu.edu/articles/basics/digitalimaging.html.

86. Combs, C.A., Fluorescence microscopy: a concise guide to current imaging methods. Curr Protoc Neurosci, 2010. **Chapter 2**: p. Unit2 1.

87. Dickson, R.M., A.B. Cubitt, R.Y. Tsien, and W.E. Moerner, On/off blinking and switching behaviour of single molecules of green fluorescent protein. Nature, 1997. **388**(6640): p. 355-8.

88. Williams, R.T. and J.W. Bridges, *Fluorescence of Solutions: A Review.* J Clin Pathol, 1964. **17**: p. 371-94.

89. Anisha, T., J. Alexander, W. Tony, and J.B. Martin, The influence of aberrations in third harmonic generation microscopy. Journal of Optics, 2010. **12**(8): p. 084009.

90. Neil, M.A., M.J. Booth, and T. Wilson, New modal wave-front sensor: a theoretical analysis. J Opt Soc Am A Opt Image Sci Vis, 2000. **17**(6): p. 1098-107.

91. Hoyt, L.F., New table of the refractive index of pure glycerol at 20 degrees C. Industrial and Engineering Chemistry, 1934. **26**: p. 329-332.

92. Liu, P.Y., L.K. Chin, W. Ser, H.F. Chen, C.M. Hsieh, C.H. Lee, K.B. Sung, T.C. Ayi, P.H. Yap, B. Liedberg, K. Wang, T. Bourouina, and Y. Leprince-Wang, *Cell refractive index for cell biology and disease diagnosis: past, present and future.* Lab on a Chip, 2016. **16**(4): p. 634-644.

93. Zhang, Q., L. Zhong, P. Tang, Y. Yuan, S. Liu, J. Tian, and X. Lu, Quantitative refractive index distribution of single cell by combining phase-shifting interferometry and AFM imaging. Scientific Reports, 2017. **7**(1): p. 2532.

94. Balaev, A.E., K.N. Dvoretski, and V.A. Doubrovski. *Refractive index of escherichia coli cells*. in *Saratov Fall Meeting 2001*. 2002. SPIE.

95. Shan, M., M.E. Kandel, and G. Popescu, Refractive index variance of cells and tissues measured by quantitative phase imaging. Opt Express, 2017. **25**(2): p. 1573-1581.

96. Schwertner, M., M.J. Booth, and T. Wilson, Simulation of specimen-induced aberrations for objects with spherical and cylindrical symmetry. J Microsc, 2004. **215**(Pt 3): p. 271-80.

97. Cha, J.W. and P.T. So. A Shack-Hartmann Wavefront Sensor Based Adaptive Optics System for Multiphoton Microscopy. in Biomedical Optics. 2008. St. Petersburg, Florida: Optical Society of America.

98. Mosk, A.P., A. Lagendijk, G. Lerosey, and M. Fink, Controlling waves in space and time for imaging and focusing in complex media. Nature Photonics, 2012. **6**: p. 283.

99. Vellekoop, I.M. and A.P. Mosk, Focusing coherent light through opaque strongly scattering media. Optics Letters, 2007. **32**(16): p. 2309-2311.

100. Gu, M. and C.J. Sheppard, Effects of defocus and primary spherical aberration on images of a straight edge in confocal microscopy. Appl Opt, 1994. **33**(4): p. 625-30.

101. Sheppard, C.J.R. and M. Gu, Edge-setting criterion in confocal microscopy. Applied Optics, 1992. **31**(22): p. 4575-4577.

102. Rea, N.P., T. Wilson, and R. Juškaitis, Semiconductor laser confocal and interference microscopy. Optics Communications, 1996. **125**(1): p. 158-167.

103. Dai, G.M., Scaling Zernike expansion coefficients to smaller pupil sizes: a simpler formula. J Opt Soc Am A Opt Image Sci Vis, 2006. **23**(3): p. 539-43.

104. Polo, A., A. Haber, S.F. Pereira, M. Verhaegen, and H.P. Urbach, *Linear phase retrieval for real-time adaptive optics*. 2013. Vol. 8. 2013.

105. Welford, W.T., *Aberrations of the symmetrical optical system*. 1974, London ; New York: Academic Press. ix, 240 p.

106. Braat, J., Analytical expressions for the wave-front aberration coefficients of a tilted plane-parallel plate. Appl Opt, 1997. **36**(32): p. 8459-67.

107. Abramowitz, M., Handbook of Mathematical Functions, With Formulas, Graphs, and Mathematical Tables. 1974: Dover Publications, Inc.