

WAVEFRONT SHAPING VIA A LEARNING ALGORITHM GUIDED,  
ADAPTIVE OPTICS DEVICE FOR FOCAL VOLUME CONTROL

by

HARRY NELSON BURTON III

A DISSERTATION

Submitted to the Faculty of Delaware State University  
in Partial Fulfillment of the Requirements for  
the Degree of Doctor of Science in the  
Optics Graduate Program of  
Delaware State University

Dover, Delaware  
August 2019

This dissertation is approved by the following members of the Final Oral Review Committee:

Dr. Thomas A. Planchon, Committee Chairperson, Division of Physical and Computational Sciences, Delaware State University

Dr. Gabriel Gwanmesia, Committee member, Division of Physical and Computational Sciences, Delaware State University

Dr. Gary Holness, Committee member, Division of Physical and Computational Sciences, Delaware State University

Dr. Renu Tripathi, Committee member, Division of Physical and Computational Sciences, Delaware State University

Dr. William Roach, External Committee member, Department of Biomedical Engineering, Vanderbilt University

## ACKNOWLEDGEMENTS

I would like to thank Dr's. Thomas Planchon and Wafa Amir for providing me this research opportunity, as well as giving me the guidance and mentorship needed to bring this research to completion. I would also like to thank my colleagues Christopher Debardeleben, Philip Chrostoski, Ashish Bachalava for their help during various points along this project. Finalizing this work would not have been possible without feedback from my external committee members. To that end, I would like to thank Dr's. Gabriel Gwanmesia, Renu Tripathi, Gary Holness, as well as external committee member Dr. William Roach (Vanderbilt University), for their role in evaluating the efficacy of my research; both at my defense and within this dissertation.

I would like to thank Delaware State University, the faculty and staff members of the Department of Physics and Engineering, and the Optical Science Center for Applied Research (OSCAR) center, for fostering my education. Funding for this research was provided by: The National Institute of Biomedical Imaging and Bioengineering (NIBIB) of the National Institute of Health (NIH) (*Award # R15EB026138*), the Delaware Space Grant Consortium (DSGC) Graduate Fellowship Program (*NASA Grant NNX15AI19H*), the National Science Foundation (*NSF-CREST grant #1242067*), and the National Institute of Health (*NIH – INBRE grant #P20GM1030446-14*). Finally, I would also like to thank my family, friends, and colleagues for supporting me in reaching this point in my academic career.

# **Wavefront Shaping via a Learning Algorithm Guided, Adaptive Optics Device for Focal Volume Control**

**Harry Nelson Burton III**

**Faculty Advisor: Dr. Thomas A. Planchon**

## **ABSTRACT**

The topic of this dissertation is the development of a learning algorithm for the shaping of a laser's three-dimensional focal intensity distribution, specifically for the interest of generating non-diffracting beams. These beams have garnered growing interest in fields such as microscopy and laser physics. The learning algorithm developed is based on a genetic algorithm (GA) approach. The goal of the algorithm is to find the necessary bi-dimensional phase (or wavefront) to apply to a Spatial Light Modulator (SLM) in order to optimize the laser spatial profile; which is measured by a Charged Coupling Device (CCD) camera. Two versions of the GA were developed, validated and used, the first one performing wavefront correction (or focal spot optimization), and the second one focal volume optimization.

The first version of GA is called single-plane GA, as it uses the signal of the CCD camera in one single transverse plane. This approach is similar to what is standard in wavefront correction schemes that are not using - for optimization - the direct measurement of the laser wavefront, but instead a signal generated by the laser that is dependent on the wavefront aberrations (or shape). An experimental setup is described, and wavefront optimization experiments validate the use of the single-plane GA as a way of correcting aberrations purposely induced on a Gaussian beam laser. The feedback signal, which can be maximized or minimized, is the spatial profile maximum intensity or its beam width.

As the single-plane GA uses the spatial profile, measured in one transverse plane, it is unable to optimize the lasers focal distribution in a predictable way. An extension of the method, called multi-plane GA, is later proposed. This novel approach to beam shaping uses, as feedback for algorithm optimization, the spatial profiles measured in multiple transverse planes. Recording multiple transverse planes at different longitudinal positions is achieved using the SLM, which applies a defocus term without physical translation of the CCD camera. This multiple recording capability requires a proper defocus calibration of the SLM, which was performed and validated via measurement of a non-aberrated Gaussian beam, as well as a SLM-generated astigmatic Gaussian beam, with comparisons taking the 3D recording of the focal distribution using manual translation of the CCD camera.

Different fitness methods using multiple transverse planes were developed and their implications and limits tested. The balance between the algorithm convergence speed, the degrees of freedom of the 2D phases induced by the SLM, the size of the algorithm search space, and ability of the GA to perform a given focal volume shaping goal, are discussed. Using the minimization of energy, size or similitude between spatial profiles in the five planes, convergence of the multi-plane GA towards a non-diffracting Bessel beam and Airy beam was demonstrated. Compared to the initial starting corrected Gaussian beam, the multi-plane GA was able to increase the laser depth of field up to 32 times for the Bessel beam, and up to 7 times for the Airy beam.

# TABLE OF CONTENTS

List of Tables .....	viii
List of Figures .....	ix
Abbreviations .....	xiii
<b>CHAPTER 1 INTRODUCTION</b> .....	<b>1</b>
1.1 Vision, Optics and Adaptive Optics.....	1
1.2 Dissertation outline .....	6
1.3 Beam wavefront and transverse spatial intensity profile .....	8
1.3.1 Maxwell’s equations and derivation of the wave equation.....	8
1.3.2 Diffraction and Huygens principle.....	11
1.3.3 Wavefront and Spatial Intensity Profile.....	13
1.3.4 The Gaussian beam .....	16
1.4 History of Adaptive Optics in Astronomy .....	20
1.5 AO Systems Design .....	25
1.5.1 Open-Loop AO systems.....	26
1.5.2 Closed-Loop AO systems .....	27
1.6 Active Optics .....	28
1.6.1 Deformable Mirrors (DM) .....	29
1.6.2 Spatial Light Modulator (SLM) .....	30
1.7 Adaptive Optics applications .....	32
1.7.1 Adaptive Optics in Astronomy .....	32
1.7.2 Adaptive Optics in Optical Manipulation .....	35
1.7.3 Adaptive Optics in Microscopy .....	36
1.7.4 Adaptive Optics in Medicine .....	37
1.7.5 Adaptive Optics for beam shaping.....	38
1.7.6 Adaptive Optics in high-energy laser Physics .....	39
1.7.7 Adaptive Optics in laser machining and fabrication.....	39
1.8 Wavefront Measurement.....	40
1.8.1 Interferometers .....	40
1.8.2 Shack-Hartmann Wavefront Sensor .....	43
1.8.3 Zernike Polynomials and the Classification of Aberrations .....	46
<b>CHAPTER 2 GENETIC ALGORITHM FOR WAVEFRONT CORRECTION</b> .....	<b>49</b>
2.1 Introduction.....	49
2.2 Experimental setup for wavefront correction .....	49
2.3 Phase Analysis .....	52
2.3.1 Derivation of the fringe pattern equation.....	53
2.3.2 Phase Extraction using Fourier Transform Analysis .....	55
2.4 Genetic Algorithm .....	59
2.4.1 Introduction.....	59

2.4.2	GA Initialization .....	61
2.4.3	Best Individual Selection and fitness score .....	63
2.4.4	Reproduction and Genetic Operators .....	64
2.4.5	GA Advantages and Disadvantages .....	66
2.5	Wavefront Correction using a learning algorithm .....	67
2.5.1	Compensation of the initial SLM distortion .....	67
2.5.2	Correction of Higher-Order Aberrations .....	77
2.6	GA Parameter Impacts on the Search Space.....	83
2.7	Discussion.....	87
<b>CHAPTER 3 CHARACTERIZATION OF 3D FOCAL INTENSITY DISTRIBUTIONS.</b>		<b>89</b>
3.1	Introduction.....	89
3.2	The 3D focal intensity distribution .....	90
3.3	Experimental setup.....	92
3.4	Calibration of the SLM defocus term .....	93
3.4.1	3D focal intensity distribution measured by CCD manual translation .....	93
3.4.2	Generation of 3D focal intensity distribution using SLM defocus term.....	96
3.5	3D focal intensity distribution scanning comparisons .....	101
3.5.1	Gaussian beam 3D focal intensity scan comparisons .....	101
3.5.2	Astigmatic beam 3D focal intensity scan comparisons .....	104
3.6	Discussion.....	106
<b>CHAPTER 4 FOCAL VOLUME OPTIMIZATION VIA MULTI-PLANE SELF-LEARNING ALGORITHM .....</b>		<b>107</b>
4.1	Introduction.....	107
4.2	Non-Diffracting Beams.....	108
4.3	Multi-plane Genetic Algorithm.....	109
4.3.1	Focal volume and depth of focus .....	110
4.3.2	Experimental setup for multi-plane GA.....	111
4.3.3	Fitness method construction .....	114
4.3.4	Illustrating the fitness score for idealized beams .....	117
4.3.5	Phase creation methods.....	119
4.4	Spatial profile similitude fitness experiment .....	121
4.4.1	Similitude fitness without normalization .....	122
4.4.2	Similitude fitness with normalization .....	123
4.4.3	Bessel beam vs Gaussian beam depth of focus comparison .....	125
4.5	Bessel beams of varying optimization ranges.....	126
4.6	Beam size difference fitness experiment .....	130
4.6.1	Beam size difference minimization using the Radial phase method .....	130
4.6.2	Beam size difference minimization using the Cartesian phase method.....	131
4.6.3	Airy beam vs Gaussian beam depth of focus comparison .....	133
4.7	Discussion.....	135

<b>CHAPTER 5 CONCLUSIONS AND FUTURE WORK .....</b>	<b>137</b>
5.1 Conclusions.....	137
5.2 Future Work.....	138
<b>REFERENCES.....</b>	<b>140</b>
<b>APPENDIX A. ZERNIKE POLYNOMIALS .....</b>	<b>150</b>
<b>APPENDIX B. SLM CALIBRATION.....</b>	<b>153</b>

## LIST OF TABLES

Table 1.1: Maxwell's Equations..	9
Table 2.1: GA wavefront correction parameters.....	70
Table 2.2: GA starting parameters for higher-order aberration correction.....	78
Table 2.3: GA convergence speeds and phase method effectiveness under various conditions ..	86
Table 4.1: List of Fitness methods used for axially extending the focal volume of the laser beam .....	114
Table 4.2: Starting parameters for the similitude method with no normalization of the intensity values .....	122
Table 4.3: Average beam size values and phase terms for Bessel beams of varying optimization ranges .....	129
Table 4.4: Starting parameters for the beam size difference fitness method using the Cartesian method.....	132

## LIST OF FIGURES

Figure 1.1: Illustration of image formation in the human eye. ....	3
Figure 1.2: Sunspot on the surface of the sun, with corrected (left) and uncorrected (right) images. ....	4
Figure 1.3: Huygens principle model of diffraction. ....	11
Figure 1.4: An optical point source and resulting spherical wavefront. ....	14
Figure 1.5: Illustration of a plane wave traveling along the propagation axis ( $z$ ) ....	15
Figure 1.6: Evolution with $z$ of the Gaussian beam width $w$ . $Z_R$ : Rayleigh Range. $\Theta$ : beam divergence. ....	17
Figure 1.7: Spatial Intensity Profile of a Gaussian beam (left) and its associated 1D intensity distribution (right).....	20
Figure 1.8: Aberrated image of Jupiter (left) and its corrected image (right).....	21
Figure 1.9: An early Eidophor projector. ....	23
Figure 1.10: “Seeing Compensator” proposed by Horace Babcock. ....	24
Figure 1.11: Example of an open-loop AO system.....	26
Figure 1.12: Illustration of the principle of a closed-loop AO system ....	28
Figure 1.13: Deformable Mirror ....	29
Figure 1.14: XY phase only SLM (Left) ....	32
Figure 1.15: AO system improvement of a imaging Uranus ....	33
Figure 1.16: Coronagraphic image of star AU Mic F606W ....	34
Figure 1.17: Optical tweezer system utilizing a spatial light modulator ....	35
Figure 1.18: AO correction of confocal fluorescence microscopy image ....	37
Figure 1.19: The LIGO sites in Livingston, Louisiana (left) and Hanford, Washington (right) ..	41
Figure 1.20: Example of a circular interference pattern ....	42
Figure 1.21: Detail of a lenslet array for wavefront sensing.....	45
Figure 1.22: 2D Illustration of the 36 <sup>th</sup> first Zernike Polynomials ....	48

Figure 2.1: Genetic algorithm for wavefront correction and phase measurement setup .....	50
Figure 2.2: Wavefront measurement arm of the original setup. ....	53
Figure 2.3: Fringe patterns subjected to well-known Zernike aberrations .....	56
Figure 2.4: Examples of known Fourier Transform $g(x) / G(f)$ pairs.....	57
Figure 2.5: Illustration of wavefront extraction using Fourier Transform method.....	58
Figure 2.6: Genetic Algorithm logical diagram .....	60
Figure 2.7: Illustration of search space in the case of two input genes .....	62
Figure 2.8: Example of the fitness scores evolution as a function of the GA generation.....	64
Figure 2.9: Crossover and Mutation methods.....	65
Figure 2.10: 2D phase of the SLM at 0V (a), Linecuts in the x and y-directions (b).....	68
Figure 2.11: Spatial intensity profile of uncorrected Gaussian beam at 750 mm lens focus.....	69
Figure 2.12: Correction results of initial SLM distortion .....	72
Figure 2.13: Fitness scores of the best individual as function of generations for each phase method.....	73
Figure 2.14: Spatial phase and linecuts before and after wavefront correction.....	74
Figure 2.15: Zernike coefficients before and after initial SLM distortion correction .....	75
Figure 2.16: SLM phase arrays after correction. ....	76
Figure 2.17: Gaussian beam spatial profile at focus with additional aberration from coverslip. .	78
Figure 2.18: Spatial profile before and after high-order aberration correction.....	79
Figure 2.19: Spatial phase and linecuts before and after aberration correction.....	80
Figure 2.20: Zernike coefficients before and after higher-order aberration correction .....	81
Figure 2.21: 512x512 SLM phase array and linecuts for both corrective methods.....	82
Figure 2.22: Fitness score evolution during correction of the aberrated Gaussian beam .....	83
Figure 2.23: 8x8 Checkerboard Array created using the checkerboard phase method .....	85
Figure 3.1: Displaying the 3D focal intensity distribution .....	91

Figure 3.2: Experimental setup for characterizing the laser 3D focal intensity distribution .....	93
Figure 3.3: Normalized MIP of the Gaussian beam over z.....	94
Figure 3.4: Gaussian beam size as function of z, extracted from the MIP shown in Fig. 3.3. ....	95
Figure 3.5: 2D defocus phase term sent to the SLM (left) and 1D linecut along x direction (right) .....	96
Figure 3.6: Beam size at $1/e^2$ as a function of the A2 defocus term applied by the SLM (a).....	98
Figure 3.7: Defocus term $A_2$ as a function of z, with linear fit depicted as dashed lines .....	99
Figure 3.8: Normalized MIP generated using the SLM calibrated for defocus.....	99
Figure 3.9: Beam size evolution comparison between the SLM and manual scans .....	100
Figure 3.10: Manual and SLM Gaussian beam scan comparisons .....	102
Figure 3.11: Manual and SLM spatial profile comparisons of the Gaussian beam .....	103
Figure 3.12: Beam size at $1/e^2$ comparison.....	103
Figure 3.13: Manual and SLM astigmatic beam scan comparisons .....	104
Figure 3.14: Manual and SLM spatial profile comparisons of the astigmatic beam .....	105
Figure 3.15: Astigmatic beam size comparison.....	106
Figure 4.1: Bessel beam wavefront and transverse spatial profile.....	108
Figure 4.2: Various non-diffracting beam types.....	109
Figure 4.3: Beam size of a Gaussian beam.....	110
Figure 4.4: Illustration of focal volume for an astigmatic beam.....	112
Figure 4.5: Experimental setup used by the multi-plane, learning algorithm.....	113
Figure 4.6: Manipulation of image for the similitude fitness method .....	115
Figure 4.7: The focal point mask fitness method.....	117
Figure 4.8: Illustration of the fitness score for different fitness methods implemented .....	118
Figure 4.9: 2D axicon phase term sent to the SLM (left), and linecut taken horizontally (right) .....	121
Figure 4.10: Results of the un-normalized similitude fitness method across all 5 planes .....	123

Figure 4.11: Iyz of the similitude method Bessel Beam.....	124
Figure 4.12: SLM phase results after normalized similitude method beam shaping.....	125
Figure 4.13: 1D longitudinal intensity for the Bessel beam (blue dashed line) and Gaussian beam (red solid line).....	126
Figure 4.14: Ixz of the Bessel beams w/ varying focal volumes. ....	128
Figure 4.15: SLM phase arrays (a) and their associated horizontal linecuts (b).....	129
Figure 4.16: Bessel-like beam generated using the beam size difference fitness method.....	131
Figure 4.17: 2D longitudinal intensity profile (Iyz) of the Airy Beam.....	132
Figure 4.18: SLM phase results after using the beam size difference method .....	133
Figure 4.19: Normalized 2D longitudinal intensity profile (Ixz) of the Airy Beam.....	134
Figure 4.20: 1D longitudinal intensity for the Airy beam (blue dashed line) and Gaussian beam (red solid line).....	135
Figure A.1: List of Zernike Polynomial equations. ....	152
Figure B.1: Experimental setup used for diffraction-based calibration of the SLM .....	154
Figure B.2: 4x4 Checker pattern applied by the SLM for diffraction calibration .....	155
Figure B.3: Beam intensity when a checkerboard pattern of $\pi$ -phase is applied to the SLM.....	156
Figure B.4: Intensity of the 0 <sup>th</sup> order diffraction as function of gray level.....	156
Figure B.5: LUT file that was generated after diffraction-method calibration.....	158

## LIST OF ABBREVIATIONS

AO	Adaptive Optics
ARPA	Advanced Research Projects Agency
CCD	Charged Coupling Device
cw	continuous wave
DM	Deformable Mirror
EM	Electromagnetic
eXAO	eXtreme Adaptive Optics
FFT	Fast Fourier Transform
FT	Fourier Transform
FWHM	Full Width, Half Maximum
GA	Genetic Algorithm
GUI	Graphical User Interface
IFFT	Inverse Fast Fourier Transform
IFT	Inverse Fourier Transform
LTAO	Laser Tomography Adaptive Optics
LIGO	Laser Interferometer Gravitational-Wave Observatory
LSFM	Light-Sheet Fluorescence Microscopy
MCAO	Multi-Conjugate Adaptive Optics
MIP	Maximum Intensity Projection
MMDM	Micromachined Membrane Deformable Mirror
MOAO	Multi-Object Adaptive Optics
PV	Peak-to-Valley
RGB	Red, Blue, Green
RMS	Root Mean Square

RTAC	Real Time Atmospheric Compensator
STEM	Science, Technology, Engineering, and Mathematics
SLM	Spatial Light Modulator

# CHAPTER 1 INTRODUCTION

## 1.1 Vision, Optics and Adaptive Optics

There are numerous fields of study within the world of STEM (Science, Technology, Engineering, and Mathematics), many of which tie back as far as early human civilization. As extensive and encompassing as these academic disciplines are, we wish to focus on one field in particular; Optics. Optics is a scientific field of study centered on the understanding of light (or electromagnetic waves), as well as the dynamics involved when light interacts with other matter or forces in nature. A burgeoning category of science, optics – as a field of study – has seen its popularity grow immensely in modern history, largely due to technical advancements and breakthroughs. Some of these advancements were within the field itself, such as our ability to detect, measure, and – ultimately – manipulate light for various purposes. This introductory chapter intends to not only provide a summarized breakdown of some of optics core fundamentals, but how they relate to a particular topic of interest that will be tackled in this dissertation: Adaptive Optics, and its use for shaping the focal volume of a laser beam. Let’s first introduce how the topic of vision links to adaptive optics.

Visual clarity (“Vision”) has been an indispensable tool to humanities ability to observe the world around it. Vision is how we are able to see the world through our day to day lives, how observers view the night sky and catalog the numerous stars and galaxies, or how a student is capable of seeing red blood cells through a microscope for the first time. The ability for humanity to learn through visual stimuli, and our desire to better perceive the world, has had marked technological, cultural, economic, and social impacts on human social behavior and living<sup>[1]</sup>. Not so dissimilar to Humans, optical-based devices - that we have designed - also rely on visual clarity to correctly process information.

An imaging camera can serve as an analog to the human eye, operating by using a lens to focus an incoming light source; similar to the human pupil. This light source then impinges upon the camera film (or sensor for digital cameras), where the intensity of the light source is captured on film through chemical emulsion, or by on-board collection within digital cameras. In the case of human eye, the retina collects the light and our brain processes various light stimuli. Cameras have become a centerpiece for capturing images of value. With inventor George Eastman spurring the proliferation of photographic cameras in the 19<sup>th</sup> century, they have had significant impact on both human culture and entertainment; with uses in photography, movies, journalism, and daily leisure. They have even made their way into our cell phones and, in the modern era, have been engineered with miniaturization in mind.

Human visual acuity is dependent upon many factors. These factors include: the sharpness of the retinal focus within the eye, the health and functioning of the retina, and the sensitivity of the interpretative faculties of the brain <sup>[2,3]</sup>. Often times, individuals with low visual acuity suffer from light not being focused properly when passing through their retina; called refractive error or “ametropia”. The top image in Fig. 1.1 showcases an eye with a cornea that is able to perfectly focus the light onto the retina. The bottom image shows a cornea with a surface that causes wavefront distortions. This produces a blurry, or out of focus, image. These distortions are subdivided into more commonly referred to visual conditions, such as Myopia, Hyperopia, Astigmatism, and Presbyopia, all of which are classified as either spherical or cylindrical aberrations. These types of aberrations are not just limited to human vision and are a common issue in optomechanical devices and instruments.

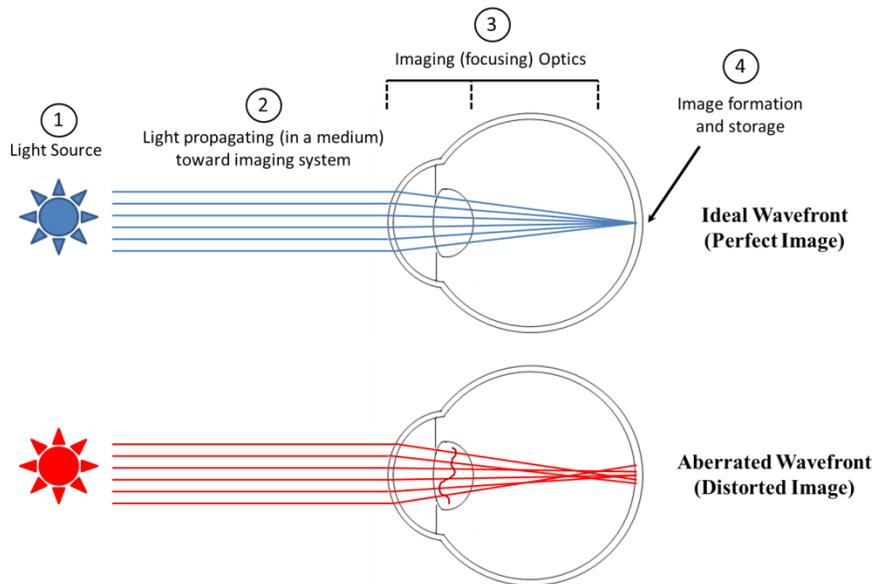


Figure 1.1: Illustration of image formation in the human eye. (Top) Ideal image formation (Bottom) Aberrated image formation.

To better describe the behavior of these aberrations, let us consider two instruments that lie along the extremes of optical imaging systems: the microscope – for discovering the structure of miniscule objects, and the telescope – for macroscopic observation of the universe. Microscopes have allowed for the imaging of miniscule objects of interest by focusing light through an objective; with other optics being used to allow for user visualization. This is juxtaposed by macroscopic observations performed through the science of Astronomy. One of mankind’s earliest forms of scientific discovery, our ability to view and study the stars has progressed significantly throughout the existence of our species. Using long-distance telescopes, both ground and space-based, humanity has acquired stunning images of our neighboring star systems, celestial bodies, and spatial anomalies. In fact, our technology has evolved so rapidly, we are at the point of even determining what may be the elemental composition of planets in distant galaxies; including their habitability. Within all these examples, the image formed does not always achieve the best quality possible; due to distortion. In both of these instruments, their spatial resolution (the amount of

detail) is a determining factor in revealing information about the object of interest. These distortions can be caused by the instrument itself (similar to aberrations in the human eye) or by the medium being looked through.

As the diagram in Fig. 1.1 showcases for the human eye, there are number of key components involved when forming an image of an object. There is the light emitted from the object (1), the medium in which light propagates before reaching the instrument (2), the imaging optics (3), and the recording/storage element (4). Each of these components can contribute to forming a poorly resolved image. In microscopy, much effort was initially made on improving the instrument itself (3) and (4), then on preventing distortion from the medium (2) <sup>[4-6]</sup>. Now, the field has moved towards trying to correct the aberrations originating from the source (sample) itself (1) <sup>[7-10]</sup>. For astronomy, the main contributor to image distortion is the medium (2), which is Earth's atmosphere. Specifically, this is an issue for ground-based telescopes, which is the main reason for sending smaller diameter telescopes into space, as the distortion brought on by this atmospheric turbulence is removed (see Fig. 1.2).

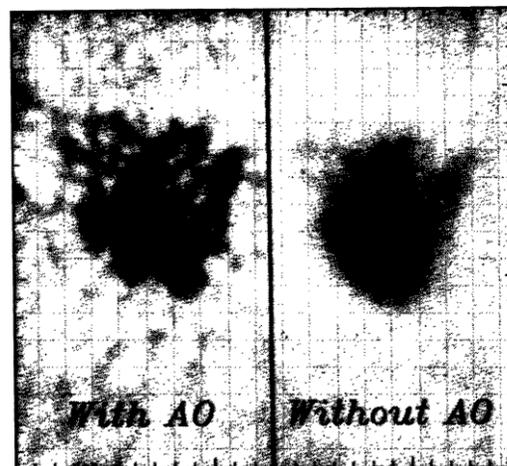


Figure 1.2: Sunspot on the surface of the sun, with corrected (left) and uncorrected (right) images. Images obtained with a Lockheed/SPO AO system <sup>[11]</sup>.

When astronomic ground-based telescopes are used to image and collect stellar information, by capturing the reflected light sent from a distant body or space, the returning light rays are altered from that of the original light source. This issue of turbulence can also be compounded if the surface of the mirror used within the ground-based telescope is not flat (smooth). Ultimately, the goal is to have the aberration free wavefront of the light source focus near-perfectly upon the light collecting optics in your system; or your eye. A methodology was developed for astronomy, called *Adaptive Optics (AO)*, in which an active optics is inserted in the system. This optical device can dynamically be deformed to induce the opposite effect caused by turbulence. AO methods have evolved drastically since their invention, and can be applied in a variety of fields, from military applications and laser science, to imaging in the biomedical field.

**This dissertation deals with the development of AO methods for wavefront correction and optical beam shaping, and more specifically, with the generation of shaped focal volumes.**

Active Optics are specially designed for transforming the wavefront of a light source. One of the key facets of an active optics is its ability to perform what is known as *wavefront correction*. At its core, correction is the act of providing a counteracting phase in order to “flatten” the wavefront back to its optimal shape; by way of the active optics. An aberrated light source is essentially a distorted (aberrated) wavefront. A wavefront can be envisioned as a ripple in a pond, comprised a several waves that flow along in space. Ripples that are close to the source that made them are more spherical, while those further away appear flat. A distorted wavefront can be thought of as deviations in the continuity of the ripple; i.e. a lack of smoothness. This deviation is the result of phase differences. The purpose of the active optics is to provide a phase that counteracts these phase differences. If an ideal “flat” wavefront is one with a value of 0, and a distortion added to

the wavefront is of value 1, then it's the function of the active optics to provide a -1 value that counteracts this distortion value.

*Optical beam shaping* shares similarities in some of the tools used for wavefront correction, but its goal is to shape the amplitude (or intensity distribution) of the light. Beam shaping can be done directly using amplitude shaping active optics, or phase shaping active optics that in turn influences the amplitude of the light source as the light propagates. Our study uses the phase shaping method specifically. In the same manner in which a phase is used to correct for a distorted wavefront, an additional phase can be imparted on the active optics in order to reshape incoming light; and its associated spatial profile. It is typically well suited to apply a beam shaping phase after wavefront correction, as aberrations in the original wavefront are likely to persist in the altered wavefront. This approach to beam shaping is quite beneficial for researchers and industries, given that active optics occupy a relatively small area (due to their small size), and the degree of flexibility offered by their ability to reshape by means of a controller or computer.

## **1.2 Dissertation outline**

This chapter is dedicated to discussing many of the core optics-based subjects associated with wave Optics and Adaptive Optics (AO). Section 1.3 discuss electromagnetic waves and their associated properties, such as phase, wavefront shapes, and the Gaussian beam; a common laser beam type which is used as a reference in this dissertation at multiple occasion. Section 1.4 is centered on wavefront distortion, while sections 1.5, 1.6, 1.7, and 1.8 discuss the history, system design, instruments, and applications comprising the field of AO's. Finally, section 1.9 discusses how a wavefront can be measured. This is an important detail to cover, as measuring the wavefront is essential to many of the experiments that are conducted within this dissertation.

Subsequent chapters touch on the various wavefront correction and shaping experiments that were performed. One approach to correcting a distorted wavefront, without measuring the aberrated wavefront, but instead its effect on a given signal (here - the laser beam profile), is laid out in **Chapter 2**. The experimental design described in chapter 2 is that of an interferometer, which also incorporates a Spatial Light Modulator (SLM) for correction. Working in tangent with the interferometer, and SLM, is a Genetic Algorithm (GA), which was developed as a feedback mechanism during the correction process. Large, static aberrations are applied to an incoming light source, and then corrected by the experimental setup. Diagnosis of the corrected wavefront, as well as the computational dynamics involved for the GA to find a solution, is also detailed.

**Chapter 3** describes the implementation of the SLM as a means of creating 3D axial intensity profiles. A novel calibration technique is employed as a means of imaging the beam source over the entirety of its propagation distance. Validation of this technique is made by obtaining the axial intensity profiles of both a Gaussian beam, one of the most common and fundamental laser beam types, as well as that of a beam suffering from induced aberrations.

Finally, in **Chapter 4**, a detailed experiment demonstrates the capacity for the GA, in conjunction with the active optics, to generate a beam with a distinct line focus; or non-diffracting qualities. This line focus generation is the result of imaging feedback from multiple transverse images (taken by a CCD camera) over several planes along the beam axis. This method will be referred to as multi-plane GA, in opposition to the single-plane GA method that is developed and presented in chapter 2, which only optimizes one transverse spatial profile. The line focus beams produced by the multi-plane GA have a number of real-world applications. As an example, in the field of microscopy, line focus beams can be used to generate light-sheets.

### 1.3 Beam wavefront and transverse spatial intensity profile

As the basis of optics is the study of light, this section is relegated to understanding the fundamentals of Electromagnetic (EM) waves, how they are capable of be described by a wavefront, and how the light distribution can be described by its transverse spatial profile. From a general standpoint, light is a byproduct of electromagnetic radiation. This radiation lies along an electromagnetic (EM) spectrum containing numerous frequencies, wavelengths, and energy. In our daily lives, we are most accustomed to light within the visible spectrum, with wavelengths ranging between the 400nm to 700nm range. There are numerous other types of radiation along this spectrum; including: Gamma, X-ray, Microwave, and Infrared, to name a few.

#### 1.3.1 Maxwell's equations and derivation of the wave equation

The generation of an electromagnetic wave lies within the theory and physics of electromagnetism, a field of scientific study that centers on the interactions between electrically charged particles; such as electrons. Electromagnetic waves are the byproduct of electromagnetic radiation, which is considered a long-range fundamental force in physics; long-range implying no direct physical interaction between two or more objects <sup>[12,13]</sup> Human curiosity of electromagnetism – in recorded history - stems as far back as 2750 BCE in ancient Egypt, with Egyptian texts describing those who handled electrically conducting animals (such as electric catfish or eels) receiving a shocking sensation <sup>[14]</sup>. While time has allowed for the breadth of knowledge regarding electromagnetism to flourish, such as Benjamins Franklins infamous kite flying experiment showing the connection between lightning and electricity, electromagnetic theory underwent a revolution during the 19<sup>th</sup> century. With experimental findings made by prominent scientists, such as Hans Christian Ørsted, Joseph Henry, Michael Henry, and Sir William Thomson, the physics of magnetism, electromagnetic induction (electric current), and

electric field oscillations, became observably proven and developed into the bedrock of current electromagnetic theory <sup>[15,16]</sup>. These independent discoveries would soon be modeled into a complete theory that would see itself as one of the principle fundamental laws of physics.

The nomenclature used today to describe electromagnetism is owed to James Clerk Maxwell. In the mid-19<sup>th</sup> century, after publishing his paper, “On Faradays lines of Force,” Maxwell would craft a set of 20 differential equations that would describe the interconnected behavior between electric and magnetic waves. These were detailed in his 1865 paper, “A Dynamical Theory of the Electromagnetic Field” <sup>[17,18]</sup>. This was later contracted into four partial differential equations, in part by Oliver Heaviside, using vector notation; as shown in Table 1.1 <sup>[19]</sup>. Maxwell was able to formulate the differential equations by adding a displacement current into André-Marie Ampère’s *Circuital Law*, which dealt with how a magnetic field projects itself around a closed loop with a current passing through it.

Gauss’ Law	$\vec{\nabla} \cdot \vec{D} = \rho$
Gauss’ Law of Magnetism	$\vec{\nabla} \cdot \vec{B} = 0$
Faraday’s Law of Induction	$\vec{\nabla} \times \vec{E} = \frac{-\partial \vec{B}}{\partial t}$
Ampère’s Law	$\vec{\nabla} \times \vec{H} = \vec{J} + \frac{\partial \vec{D}}{\partial t}$

Table 1.1: Maxwell’s Equations. D is the displacement field,  $\rho$  is the charge density, B is the magnetic field vector, E is the electric field vector, H is the magnetizing field, and J is the current density.

By taking the second derivatives and curls of the differential Maxwell Equations, one can find that the speed of light  $c$  is inversely proportional to the permittivity  $\epsilon_0$  and the vacuum permeability  $\mu_0$ , as shown in Eqn. (1.1).

$$c = \frac{1}{\sqrt{\epsilon_0 \mu_0}} \tag{1.1}$$

Maxwell's equations, and derivation for the speed of light, are pivotal discoveries within classical physics. It predicted that, not only are the electric and magnetic fields perpendicular to each other along to the propagation direction, but also that visible light itself is an electromagnetic wave. The formulae of the Maxwell Equations drew some level of criticism with regards to their mathematical simplicity and ability to describe – theoretically – the behavior of an electromagnetic wave. However, Maxwell's theories were experimentally proven by Heinrich Rudolf Hertz in 1887<sup>[20]</sup>, who proved that radio waves travel at the speed of light (Eqn. 1.1) and were also a form of light, solidifying Maxwell's work as one of the fundamental laws of physics.

Maxwell's Equations are capable of providing a versatile answer in describing many known optical phenomenon; from a mathematical standpoint. Specific to topic of adaptive optics, these equations are utilized in expressing the nature of a *wave*. Given this, it is appropriate to first define what is known as the *wave equation*. Much of the wave equation stems from the usage of Faradays Law. By taking the curl of Faradays Law of Induction, and then using Ampere's law (cf. Eqn. 1.2), the derivation of the wave equation (Eqn. 1.3) can be achieved.

$$\vec{\nabla} \times (\vec{\nabla} \times \vec{E}) = \vec{\nabla}(\vec{\nabla} \cdot \vec{E}) - \nabla^2 \vec{E} = -\frac{1}{c} \frac{\partial}{\partial t} (\vec{\nabla} \times \vec{B}) = -\frac{\mu\epsilon}{c^2} \frac{\partial^2 \vec{E}}{\partial t^2} \quad (1.2)$$

$$\left( \nabla^2 - \frac{1}{v^2} \frac{\partial^2}{\partial t^2} \right) \vec{E}(\mathbf{r}, t) = 0 \quad (1.3)$$

$$\text{where } v = \frac{1}{\sqrt{\mu\epsilon}} \quad \text{and } \nabla = \frac{\partial}{\partial x} \hat{i} + \frac{\partial}{\partial y} \hat{j} + \frac{\partial}{\partial z} \hat{k}$$

The symbol  $v$  denotes the speed of the wave, with the relative permittivity and permeability being  $\epsilon$  and  $\mu$ , and  $\mathbf{r}$  is the spatial coordinate that defines which dimension the wave equation travels (e.g.  $x$ ,  $y$ , or  $z$ ). One-dimensional wave equations would result in the use of only one of these variables and – generally – the variable  $z$  is used as a way of defining the waves direction of propagation. As such, the use of  $z$  will continue to be used in this dissertation when describing the

direction the wave is traveling; unless noted otherwise. As the wave propagates in  $z$ , interactions between the wave and matter can lead to diffraction, which is one of the more fundamental optical phenomena that is associated with wave behavior.

### 1.3.2 Diffraction and Huygens principle

*Diffraction* is the bending of light (or waves) around an edge or obstacle but can also occur naturally as a wave propagates in space. Originally discovered by scientist Ibn Sahl in c. 984 <sup>[21,22]</sup>, diffraction has been widely studied by many physicists; such as Francesco Maria Grimaldi, James Gregory, and Isaac Newton. However, the physicist Christian Huygens would set a new precedent on the understanding of diffraction. During his work in 1678, Huygens would expand upon the wave theory of light by proposing that each point source contained within a wavefront - the surface where all the wave value of disturbance is the same - would generate a secondary set of spherical point sources. These secondary sources would then go on to create an entirely new wavefront, comprising of spherical *wavelets*. (see Fig. 1.3).

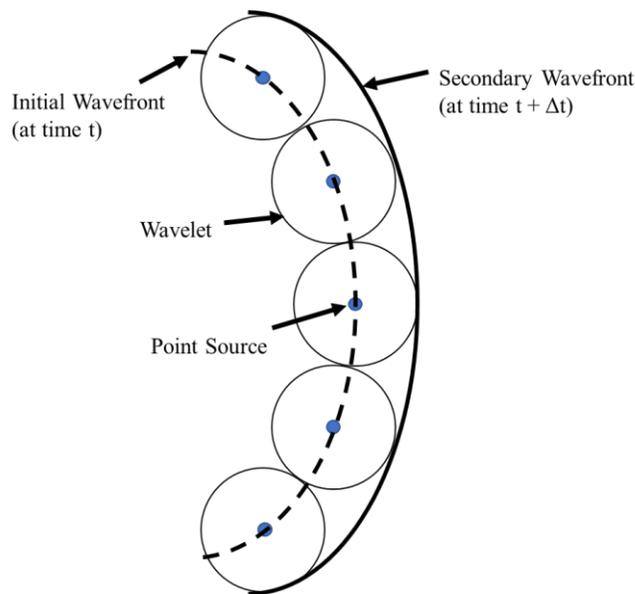


Figure 1.3: Huygens principle model of diffraction. An initial wavefront (at time  $t$ ) is seen as composed of spherical wavelets. After a time,  $t+\Delta t$ , when these wavelets add up, a secondary wavefront is formed.

This theoretical model would become a vital addition to the wave theory of light. This nature of diffraction would spark a number of breakthroughs within the understanding of light and wave theory. For example, by utilizing a two-slit aperture, and allowing a light source to propagate through the aperture, physician Thomas Young was able to showcase the optical phenomenon known as *interference*. This interference is grounded upon the theory of wave superposition, in which the phases and intensity of two waves superpose on one another to create regions of light and darkness. This demonstration of wave interference, along with the theoretical underpinnings Augustin-Jean Fresnel's theories on interference, would serve as the basis for the Huygens-Fresnel Principle. The Huygens-Fresnel Principle is a fundamental law of optics and serves as a classical model in evaluation wave propagation and diffraction. The principle is specific to a single point source in space,  $P_0$ , and is seen in Eqn. 1.4, where  $\theta$  is the angle between the normal  $\hat{n}$ , and the vector  $\vec{r}_{01}$  is pointing from  $P_0$  to  $P_1$ ; a secondary point. Finally, the rectangular form of this law can be seen in Eqn. 1.5, with  $x'$  and  $y'$  being the coordinates of the diffracting aperture,  $z$  being the propagation direction,

$$U(P_0) = \frac{1}{i\lambda} \iint_{-\infty}^{+\infty} U(P_1) \frac{e^{ikr_{01}}}{r_{01}} \cos(\theta) ds \quad (1.4)$$

$$U(x, y) = \frac{z}{i\lambda} \iint_{-\infty}^{\infty} U(x', y') \frac{e^{ikr_{01}}}{r_{01}^2} dx' dy' \quad (1.5)$$

$$r_{01} = \sqrt{z^2 + (x - x')^2 + (y - y')^2}$$

The Huygens-Fresnel Principle is useful in interpreting the diffraction nature of a wave both within the near field, and the far field. Mathematical representation of diffraction waves depends on whether someone intends to study the diffraction effect in the near field or far field. Near field diffraction occurs very close to the object causing the diffraction phenomenon, while

far field diffraction is the opposite. Complex wave amplitude approximations can be made for both the near and far field. By applying binomial expansion to  $r_{01}$ , we can obtain the near field approximation (Fresnel approximation) and far field approximations (Fraunhofer approximation). The Fresnel approximation is shown in Eqn. 1.6, while the Fraunhofer approximation is shown in Eqn. 1.7. Fresnel approximations assume a parabolic (spherical) wave, and Fraunhofer approximations assume that of the planar wave.

$$U(x, y) = \frac{e^{ikz}}{ikz} \iint_{-\infty}^{\infty} U(x', y') \exp \left\{ j \frac{k}{2z} [(x - x')^2 + (y - y')^2] \right\} dx' dy' \quad (1.6)$$

$$U(x, y) = \frac{e^{ikz} e^{i \frac{k}{2z}(x^2 + y^2)}}{ikz} \iint_{-\infty}^{\infty} U(x', y') \exp \left[ -j \frac{2\pi}{\lambda z} (xx' + yy') \right] dx' dy' \quad (1.7)$$

### 1.3.3 Wavefront and Spatial Intensity Profile

Electromagnetic waves are harmonic, meaning they oscillate – sinusoidally – as the wave propagates. A complex form of the wave equation can be derived that illustrates many key components of a traveling wave. By using separation of variables on the wave equation, and expressing it as a complex harmonic differential equation, we can obtain the complex electric field equation (Eqn. 1.8) for a wave propagating in  $z$ .

$$E(z, t) = E_0(z, t) [\cos(\omega t) + i \sin(\omega t)] \quad (1.8)$$

This equation can be further simplified using Euler's Formula (Eqn. 1.9) to provide the complex form of the electric field (as shown in Eqn. 1.10), with  $i$  is an imaginary unit number, and  $\omega$  is the angular frequency.

$$e^{iz} = \cos(z) + i \sin(z) \quad (1.9)$$

$$E(z, t) = E_0(z, t) e^{-i\omega t} \quad (1.10)$$

Given its importance in the shaping of the wave, let us discuss the concept of *phase*. When dealing with harmonic waves, each waveform travels cyclically, and a set point within this particular cycle is referred to as a phase. If we express a wave as a sinusoidal function, which has a specific periodicity, then the phase - and accompanying shifts - can be considered as  $\omega t + \delta$  <sup>[23]</sup>. In this equation,  $\omega t$  is the phase of the wave, and  $\delta$  is the phase shift. Due to the fact that a waveform is cyclical, this means that the phase operates from 0 to 360°; or 0 to  $2\pi$  (if expressed in radians).

A *wavefront* is generated as a consequence of radiated light rays. It can be considered as a series of points in space that – as a whole – travels along the same direction as the light ray; see Fig. 1.4. It should be noted, however, that the shape of the wavefront is perpendicular to the direction of light propagation at all points in space <sup>[24]</sup>. More importantly, a wavefront is best categorized by the similitude between the phases of each point, and how identical phases allow for the creation of a surface (wavefront).

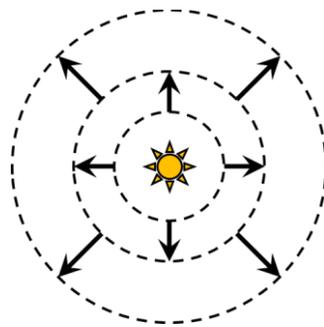


Figure 1.4: An optical point source and resulting spherical wavefront. The spherical surface in which the rays of light share a similar phase is called the wavefront; perpendicular to the direction the ray is traveling.

If a point source radiates a light wave, as shown in Fig. 1.4, the corresponding wave is called a *spherical wave*; as the shape of the wavefront would be spherical. A spherical wave has no  $\phi$  or  $\theta$  dependence; due to their angular symmetry <sup>[25]</sup>. In this particular case, one would simply

need to modify the wave equation into Eqn. 1.11, which can be rewritten in spherical form; see Eqn. 1.12.

$$\frac{\partial^2 \psi}{\partial r^2} + \frac{2}{r} \frac{\partial \psi}{\partial r} = \frac{1}{v^2} \frac{\partial^2 \psi}{\partial t^2} \quad (1.11)$$

$$\frac{\partial^2 (r\psi)}{\partial t^2} = v^2 \frac{\partial^2 (r\psi)}{\partial r^2} \quad (1.12)$$

If the light from the point source radiates symmetrically, then as the light travels an infinite distance away from the source, this spherical shape (from the point source origin) instead becomes that of a plane wave. It should be noted that, realistically, such a wave could not exist as it would need to fill the whole of space and would, therefore, require an infinite amount of energy. The plane wave instead serves a model for how a wavefront theoretically behaves as it propagates through space. A *plane wave* is composed of wavefronts that are parallel planes - or sheets - that are perpendicular to the direction the light is traveling, and have a constant phase along its surface (Fig. 1.5).

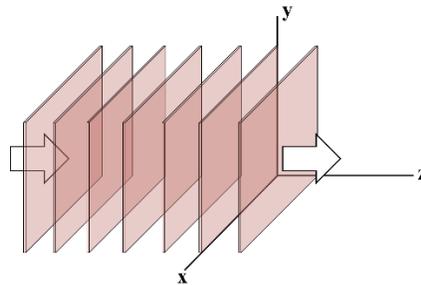


Figure 1.5: Illustration of a plane wave traveling along the propagation axis (z)

The wave equation can be used to illustrate a plane wave by expressing the traveling wave equation as an exponential. By doing so, both the complex and real components of the wave can be described, which is represented in Eqn. 1.13.

$$E(\mathbf{r}, t) = E_o e^{i(\mathbf{k}\cdot\mathbf{r} - \omega t + \phi)} \quad (1.13)$$

Here,  $\mathbf{k}$  is the wave vector equal to  $2\pi/\lambda$ , and  $\mathbf{r}$  is the position vector. These equations adequately express what is known as a *plane wave*, which is emphasized by the constant phase of the plane wave. Referring to Eqn. 1.13, the amplitude and the phase of the wave can be defined. From a general standpoint, *amplitude* describes the magnitude of a periodic function. The term is common place among many fields of physics and science and can be expressed in a multitude of ways. In Eqn. 1.13, the amplitude of the electric field would be denoted by ( $E_o$ ). The method by which someone evaluates the amplitude largely depends on the type of source the amplitude is coming from (e.g. electromagnetic, acoustic, etc.), and the preferred representation of said value. The term in the exponential in Eqn. 1.13 is the phase of the wave and contain factors that are due to the propagation in space and time, and the phase constant  $\phi$ . This phase constant if represented on a surface perpendicular to the light propagation would depend on transverse variables  $x$  and  $y$ , and would describe the wavefront shape.

For an optical signal in the visible spectrum, the oscillation of the amplitude is in the order of few femtoseconds, so most detector would not be fast enough to measure its time dependence, and an integrated value would be measured. A camera for example, would record the intensity profile,  $I \propto E_o^2$  ( $\propto$  being “proportional to”) and would also lose all information on the phase. Section 1.9 later in this chapter presents methods to circumvent this and measure the wavefront described by the phase  $\phi(x,y)$ .

#### 1.3.4 *The Gaussian beam*

The Gaussian beam is the most commonly used laser beam in laser systems and is well-known in both its mathematical form and optical properties; in addition to being easily produced.

Exploiting these properties has allowed its implementation to proliferate across a number of research, commercial, and manufacturing fields. Due to how well studied the Gaussian beam is, it also makes it a desirable reference to use when observing other types of beams and their behavior. It is for this reason that the experiments conducted in this dissertation utilize the Gaussian beam, as it will be reshaped into another beam and used for comparative analysis.

The Gaussian beam gets its name because its transverse spatial profile is described by a two-dimensional Gaussian function. Mathematically, the Gaussian beam stems from the paraxial wave equation, and its electric field magnitude can be seen in Eqn. 1.14, where  $r$  is the distance from the center of the beam, and  $w$  is the beam radius at  $1/e^2$  at propagation distance  $z$ .

$$E(r, z) = E_0 \frac{w_0}{w(z)} \exp \left[ -\frac{r^2}{w^2(z)} \right] \quad (1.14)$$

$E_0$  is the electric field amplitude at the beam focus, the point in which the size of the Gaussian beam is at its smallest. While the shape of the beam occupies the transverse plane ( $x$  and  $y$ ), it travels in a given direction ( $z$ ); see Fig. 1.6. The convergence and divergence of the Gaussian spherical waves creates a beam waist  $w_0$ , where the energy of the beam is most concentrated.

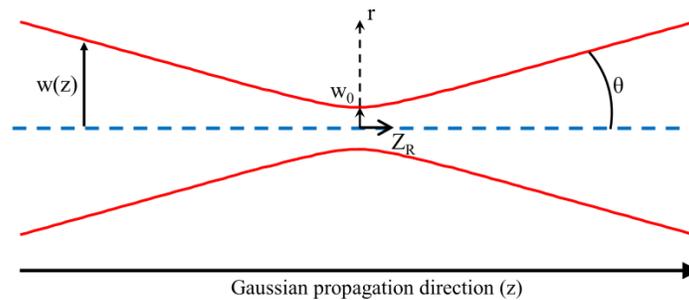


Figure 1.6: Evolution with  $z$  of the Gaussian beam width  $w$ .  $Z_R$ : Rayleigh Range.  $\Theta$ : beam divergence

As the Gaussian beam is a solution to the paraxial wave equation <sup>[26]</sup>, this means that it abides by the same diffraction and wavefront behavior outlined by the Huygens-Fresnel Principle

and Helmholtz Equation. This means that the Gaussian beam is the product of spherical wavefronts, which travel along the beam propagation axis ( $z$ ) and evolve over time. The beam radius ( $w$ ) is dependent upon the focal point of the Gaussian beam, and how far from this point (along the  $z$ -axis) the observer is. The beam waist,  $w_o$ , is marked by a convergence – or divergence – of spherical waves from a single focal spot. Eqn. 1.15 defines the beam radius as a function of both the beam waist, as well as the propagation distance  $z$ , and *Rayleigh Range*,  $Z_R$ . The Rayleigh Range is the point at which the cross-sectional area of the beam width is doubled and is expressed as Eqn. 1.16.

$$w(z) = w_o \sqrt{1 + \left(\frac{z}{Z_R}\right)^2} \quad (1.15)$$

$$Z_R = \frac{\pi w_o^2}{\lambda} \quad (1.16)$$

The Rayleigh range is an important tool in being able to model the way in which the Gaussian beam width evolves over the entirety of the propagation axis. When at a distance equal to  $Z_R$  from the beam waist, the width of the beam ( $w$ ) is  $\sqrt{2}$  larger than at the center of the beam waist. The way in which the “wings” of the beam profile converge or diverge, moving from the near field (waist) to the far field, is known as the *beam divergence angle* ( $\theta$ ); see Eqn. 1.17.

$$\theta = \frac{\lambda}{\pi w_o} \quad (1.17)$$

Depending on the divergence angle, this can have significant ramifications on the maximum intensity of the beam source along the propagation axis; before reaching  $1/e^2$  of the maximum beam intensity. If one were to minimize this divergence angle in the far field, then it would ultimately lead to a higher maximum intensity over a much large range of the beam. This would – in turn – produce a larger beam waist as result. This relationship is indicative of the

diffraction principle. When describing the *Irradiance* (Intensity) of the beam, it can be expressed as Eqn. 1.18, with  $I_0$  being the square of  $E_0$ .

$$I(r) = I_0 \exp\left(\frac{-2r^2}{\omega_0^2}\right) \quad (1.18)$$

What is unique about this equation is that it is almost similar to that of the electric field equation of a Gaussian beam (Eqn. 1.14), but placed at the transverse plane corresponding to the beam waist. Interestingly enough, when the Fourier Transform is applied to Eqn. 1.14, in order to go from the far-field to the near-field, the same Gaussian beam equation form is obtained. This allows for the Gaussian beam distribution to be modeled over any point within the optical system. The beam waist, Rayleigh Range, and beam divergence, are classified as the beam parameters for the Gaussian beam.

The **Spatial Profile** is another name for the transverse intensity distribution  $I(r)$  of a laser source. An optical measuring instrument, such as a Charged Coupling Device (CCD) camera, can be used to capture the intensity profile of the beam. From here, the measuring device ascribes a value for each pixel. This value is analogous to the intensity of the beam per pixel. As each pixel has its own individual intensity value, a plot of the intensity distribution can be generated. Fig.1.7 illustrates how a 2D profile of a Gaussian beam can be represented; with the 1D intensity distribution represented on the right panel.

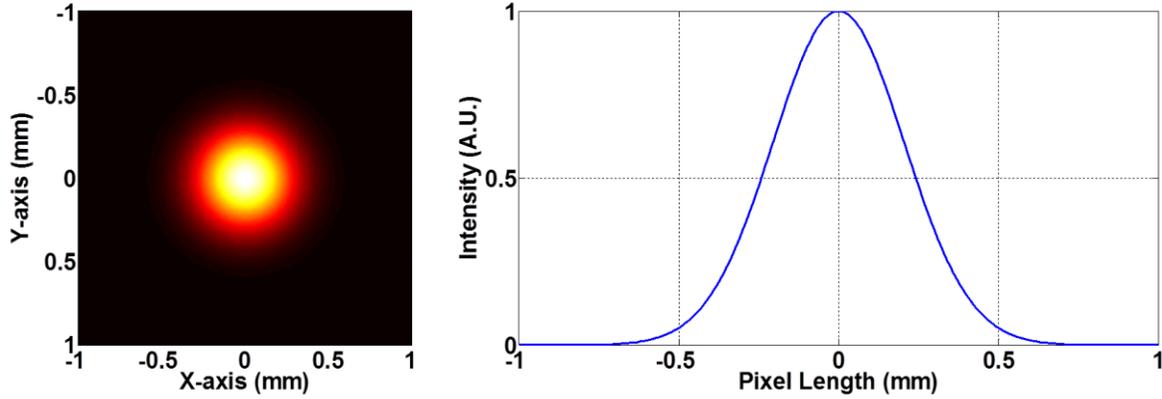


Figure 1.7: Spatial Intensity Profile of a Gaussian beam (left) and its associated 1D intensity distribution (right)

For a given intensity distribution, it is possible to define the width of the beam using the *Full Width Half Maximum* (FWHM) value. As its name would suggest, the FWHM is a measure of the intensity distributions width when it is at half of its maximum value. Each type of distribution has a corresponding FWHM. As an example, a Gaussian beam with the intensity distribution described in Eqn. 1.19 has an associated FWHM linked to its beam radius  $w$  at  $1/e^2$  is shown in Eqn. 1.20.

$$I(r) = I_0 e^{\frac{-2r^2}{w^2}} \quad (1.19)$$

$$FWHM = w\sqrt{2 \ln(2)} \quad (1.20)$$

#### 1.4 History of Adaptive Optics in Astronomy

While a theoretical spherical wave is considered to have a uniform shape, reality never tends to see such an outcome. Wavefront distortions are commonplace when dealing with optical waves and are characteristically viewed as a negative within many fields that require high visual quality or peak performance from an experiment; especially with the usage of lasers. Given the pervasiveness in having to deal with such a challenge, the most pertinent question to ask is: “What leads to such distortions in the first place, and what notable features are indicative of a distorted

wavefront?” To answer this, let us return to the field of astronomy; as an example. As was previously mentioned in the introduction, ground-based telescopes were limited in the quality of stellar imaging due to atmospheric turbulence. Solar radiation emitted by our Sun undergoes a variety of optical affects when entering the Earth’s atmosphere, altering its wave nature and intensity. Comprising of numerous gas molecules, particulates, and water vapor, our atmosphere is capable of producing several optical phenomena; the most prevalent being absorption. The atmosphere may be imagined as a material, with incoming solar radiation – and its accompanying wavelengths – meeting either a transparent or opaque surface; or somewhere in between [27,28]. When ground-based telescopes are used to image and collect stellar information, by capturing the reflected light sent from a distant body or space, the phase and amplitude of the returning wavefront is altered from that of the original wavefront. The distortion of a wavefront is typically due to the non uniformity of the medium optical index, diffraction, or scattering of the light source. Fig. 1.8 demonstrates the outcomes of aberrative effects on an image of Jupiter. These aberrations can be described by the mathematical formulae known as the Zernike Polynomials (described in section 1.8.3).

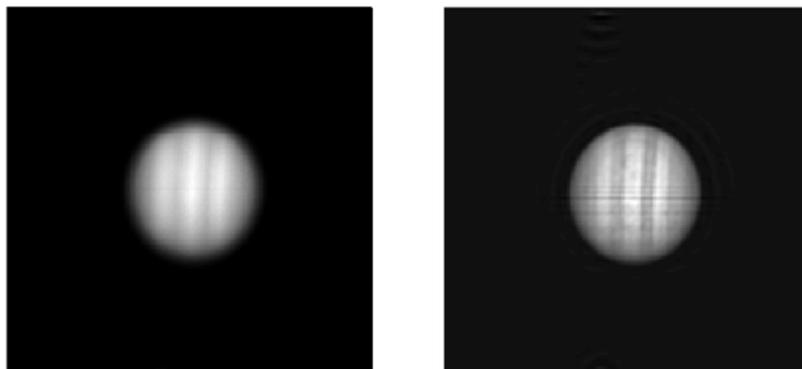


Figure 1.8: Aberrated image of Jupiter (left) and its corrected image (right) (from [29]).

*Seeing* was often complicated with astronomical telescopes prior to 1970, and still is today; in certain aspects. The angular resolutions of these telescopes were a determining factor in an observer's ability to accurately image stellar objects <sup>[30]</sup>. In the 1930's, early methods of correcting for turbulence were brought to bear with the development of autoguiders. These were astronomical devices which could help in tracking, and correcting, for such deviations. By incorporating a photodetector, such as a CCD camera or Complementary Metal–Oxide–Semiconductor (CMOS) camera, the autoguider takes in light captured by the telescope within a set exposure period. In our more modern time period, a computer measures any apparent changes in the position of neighboring stars or bodies around the object of interest in order to send position corrections to the telescope mount; which is computerized. The device has become prolific in astronomy, from amateur usage to professional research <sup>[31,32]</sup>. While autoguiders have led to a marked improvement in the capture and imaging of distant objects in space, there was still a continued need for mitigating turbulence effects.

During their initial conception, autoguiders were not designed to correct for defocusing or higher order aberrations <sup>[33]</sup>. It was not until decades after Horace Babcock originally theorized the deformable mirror that Adaptive Optics (AO) became more commonplace within the field of optics. Babcock, an American astronomer with a specialization in spectroscopy, wanted to address the issue of atmospheric turbulence impacting the image quality of ground-based telescopes. Two important advancements in wavefront correction would lead him to his theory of the first AO device.

The first occurred while working at the Mount Wilson and Palomar observatories at the Carnegie Institution of Washington in the late 1940's. Babcock developed his own autoguider, which was inspired after his frequent use the Foucault (Knife Edge) method. The Knife Edge

technique, which can be used to detect mirrors surface defects with relative high precision, is a feasible test in determining the properties of atmospheric turbulence. By incorporating the method into an autoguider, the stellar seeing disk (the point spread function used for viewing through the atmosphere) of the Mount Wilson telescope remained relatively motionless; a feat that was quite rare in astronomy at that time.

The second advancement occurred when he came across a device known as an Eidophor (see Fig.1.9), which was created by Dr. Fritz Fischer 1943 <sup>[34]</sup>. An Eidophor is used as a television projector and utilizes electrostatic charges from a scanned electron beam in order to deform a highly viscous oil that covers a mirrored disk; with the disk replacing the standard movie projector film at the time. Such devices were used in movie theaters and could even project television shows.

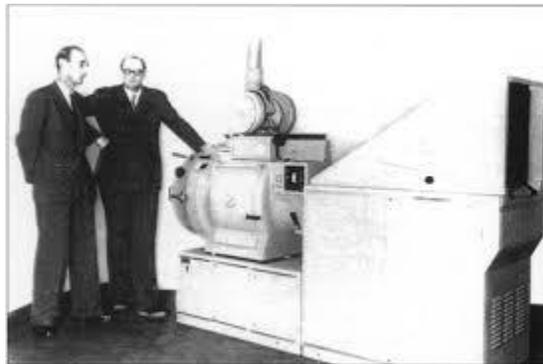


Figure 1.9: An early Eidophor projector. <sup>[35]</sup>.

Here is where Babcock's conceptualization of AO's bore fruit:

“If we had a means of continually measuring the deviation of rays from all parts of the mirror, and of amplifying and feeding back this information so as to correct locally the figure of the mirror in response to the schlieren pattern (optical inhomogeneity's in transparent materials – such as darkening/brightening), we could expect to compensate both for seeing and for any inherent imperfections of optical figure” <sup>[36]</sup>

By observing the designs of the Eidophor's ability to deform the oil through the use of an electric charge, Babcock was able to conceptualize an early prototype of a deformable mirror (called a "Seeing Compensator"); shown in Fig. 1.10. Notable in the figure is the use of both an autoguider, as well as an Eidophor. In his design, he uses a rotating knife-edge and photodetector to analyze the wavefront from a distant star. Based on feedback from the wavefront, correction signals are sent back to the compensator so that the optical thickness of the deformable oil film can be made on the Eidophor mirror. This feedback cycle is a mainstay of AO's, and when combined with a laser and optical control system that can be capably run by a computer, the benefits of AO's grew in popularity from the 1970's and onwards.

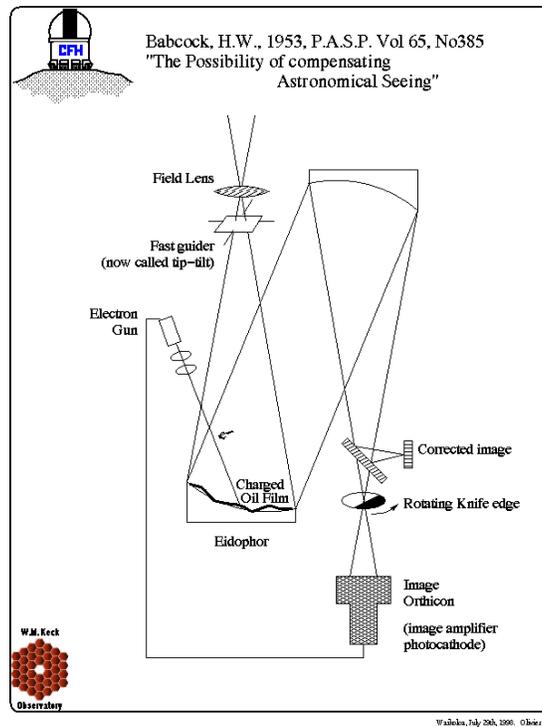


Figure 1.10: "Seeing Compensator" proposed by Horace Babcock. [35].

Experiments into wavefront correction of atmospheric turbulence initially began with the United States Advanced Research Projects Agency (ARPA) in 1972 (abbreviated DARPA today),

in conjunction with the contracting company Itek Optical Systems. The purpose of the department's experiments was for the tracking of Soviet satellites. Due to atmospheric turbulence, the necessity to compensate for such distortions in real time became a necessity for the United States military. During initial testing, the issue of finding the proper wavefront corrector, wavefront sensor, and a computational system able to analyze the wavefront data and corrections needed in real time, were demanding challenges. However, these challenges were ultimately surmounted in December of 1973, with the development of the Real-Time Atmospheric Compensator (RTAC) [37]. DARPA's research into AO's and wavefront correction would continue beyond this point, and advances in the science would later spur its usage in numerous other fields [38].

Research developments made after the construction of the RTAC saw numerous evolutions to the typical wavefront correction device. With wavefront correction becoming more precise, and with computers being more capable of processing information, AO's have seen widespread use across several STEM fields. Most notably, these fields include astronomy, optical-microscopy, and high-energy physics. Today's most common active optics are typically categorized as either a Spatial Light Modulator, or a Deformable Mirror, the latter of which is more commonly used in telescopes or in high intensity laser systems. Improvements in technology have made these devices significantly smaller in size, compared to their 20<sup>th</sup> century predecessors, and offer more in terms of wavefront correction and beam shaping. Numerous research groups and industries have sought the use of such devices in modifying existing techniques, or in fabricating entirely new ones.

## **1.5 AO Systems Design**

When constructing an AO system, it is important to decide upon how you intend to utilize the system. Depending on the type of environment the optical setup is operating in, the types of

aberrations, and how dynamic (responsive) the active optics needs to be, the components and layout of the AO setup will be constructed differently. Typically, there are two distinct AO system design setups: *Open-Loop* and *Closed-Loop*. The type of setup largely depends on whether the system has knowledge of the wavefront distortion before (Open-Loop), or after (Closed-Loop), the wavefront has been corrected. Regardless of which type of setup is used, all of them will – at a minimum – require an active optics, an imaging sensor (such as a CCD camera), a wavefront sensor, and a controller that sends a signal to the active optics for the purposes of altering the wavefront.

### 1.5.1 Open-Loop AO systems

An open-loop AO system often relies on pre-determined knowledge of the aberration that will impact the wavefront being received by the sensor. As this aberration is expected, it means that the system does not require iterative feedback for wavefront correction; such as that of a closed loop AO. A diagram of a typical open-loop system can be seen in Fig. 1.11.

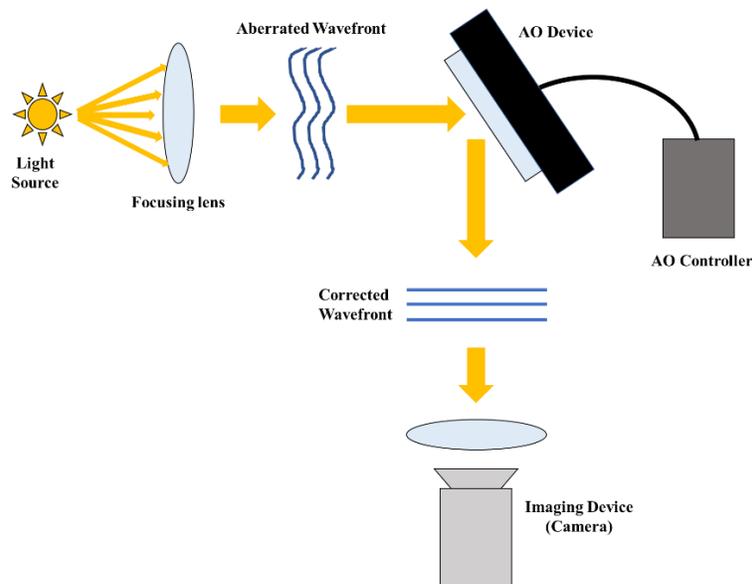


Figure 1.11: Example of an open-loop AO system. A controller sends a signal to the active optics for it to provide a new wavefront. This can then be used to correct for known aberrations or induce new ones.

This also means that a wavefront sensor is not necessarily needed for correction. Due to this, the wavefront that are generated for the active optics – in an open-loop – are already established and should apply a correction with a high degree of accuracy. In other instances, such as wanting to determine the effect a particular aberration has on the systems performance, an open-loop AO can be used to add an aberrative effect to an existing wavefront. Most often, this system will incorporate deformable mirrors into what is known as a Multi Object Adaptive Optics (MOAO) [39]. This technique uses multiple deformable mirrors in order to target and correct individual aberrations commonly associated with atmospheric turbulence.

### *1.5.2 Closed-Loop AO systems*

In a closed-loop AO system, the defining characteristic is that it relies on feedback of the wavefront before applying a correction. As such, it does not work on predefined knowledge of the aberrations within the system. Fig. 1.12 provides an illustration of an aberrated light source traveling within a closed-loop AO system. A light source with an aberrated wavefront impacts a beam splitter, and a wavefront sensor measures the wavefront shape, then provides feedback to the controller, which can then modify the surface of the AO device to correct the beam

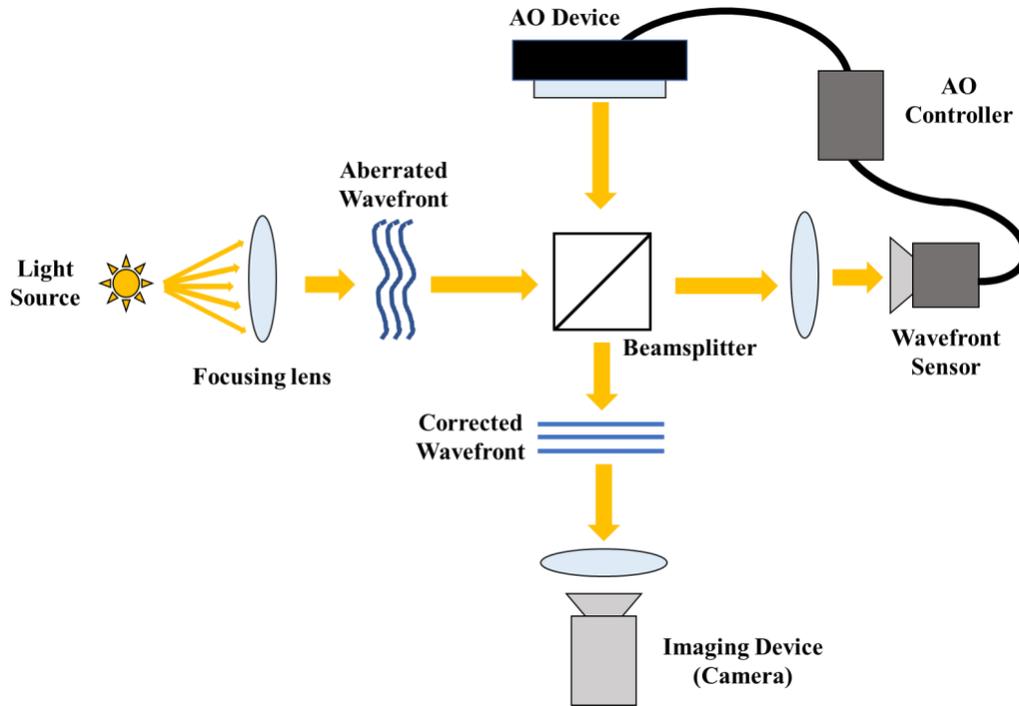


Figure 1.12: Illustration of the principle of a closed-loop AO system

This type of system is more relevant under an environment in which the aberrations present in the wavefront are unknown, in addition to scenarios in which the wavefront errors are smaller; relative to open-loop systems. Many closed-loop systems split the light source between a wavefront sensor and an imaging device. The sensor measures the level of aberration inherent in the wavefront, and the information from this can be used to determine the corrective wavefront needed to remove (or mitigate) aberrations in the wavefront. This process is continuous, and this feedback system can be used to iteratively correct the wavefront.

## 1.6 Active Optics

Shaping the wavefront for generating a line focus, the main goal of this dissertation, will require a device capability of altering the wavefront. Given the critical nature of wavefront correction devices within a detection system, it is of importance to categorize them. Two

commonly known active optics include Deformable Mirrors (DM) and Spatial Light Modulators (SLM). Both devices require the use of an electronic control in order to alter the structure of their shape; which can be handled via a PC. There are distinct differences in the way in which these devices need to be utilized within their respective fields. This section will elaborate on the various types of active optics that are available, as well as how they function.

### 1.6.1 Deformable Mirrors (DM)

Deformable Mirror (DM) were constructed as the byproduct of the United States military seeking to detect Soviet satellites by correcting for atmospheric turbulence. While the mechanical components of DM's have evolved over time, their core principles are still in place. DM's comprise of a collection of piezoelectric actuator stacks, forming what is known as a faceplate. By applying a voltage to the stack, its axial length can be altered <sup>[40]</sup>. This ability to change the axial length allows for the surface of the mirror to be deformed; illustrated in Fig. 1.13. Alternatively, segmented mirror arrays can be utilized to create the same effect. The required voltages needed for corrected can be found using the modal method, which measures the wavefront and shifts the actuators to apply a conjugate of the aberration <sup>[41]</sup>. This ability for deformation allows for the phase of an incident light source to be altered.

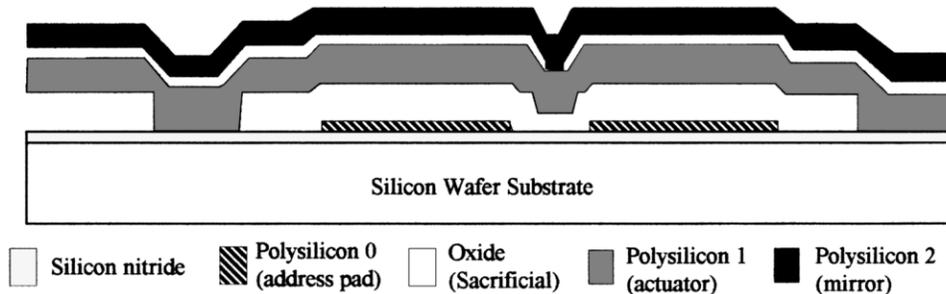


Figure 1.13: Deformable Mirror. Depending on if - and how much - voltage is applied through the actuator, the mirror will adjust themselves accordingly <sup>[42]</sup>

Deformable mirrors fall into three different types based on the deformable material/instrument used, and the method at which phase is altered. These are classified as: Elastomers, Membranes, and Cantilever/Torsion beams. Elastomers are elastic, synthetic polymers (i.e. rubber). A metalized elastomer is placed over an electronic circuit, and when the circuit is given electricity, the elastomer deforms due to electrostatic forces. Elastomer DM's are not as widely used; due to improvements in membrane-based deformable mirrors. They are also subjected to various structural issues, such as the elastomer bending on its own weight <sup>[43,44]</sup>. Both cantilever and torsional DM's utilize a thick reflective beam suspension; typically an air gap. Cantilevers are connected via a thin cantilever hinge, and when an electrical signal is received, the lever bends and the beam tip is deflected. For torsional DM's, the torsion beam is connected to two thin torsion hinges that, when an electrical signal is sent, twists the torsion hinges so that the beam rotates about the two hinges <sup>[45-47]</sup>.

Membrane DM's, more commonly referred to as Micromachined Membrane Deformable Mirror (MMDM), are among one of the more popular DM's due to the low-cost effectiveness in their construction and compartmentally small size; the latter of which is common among many DM's. MMDM's employ a polymer membrane that is metalized, and stretched over what is referred as a spacer; which forms an air gap between the membrane and MMDM circuit. When the circuit receives an electrical signal, the membrane deforms into the air gap <sup>[48-51]</sup>. DM's are popular for their laser printing capabilities, their benefits to optical communication and ophthalmology, and compatibility with spatial light modulators <sup>[52,53]</sup>.

### *1.6.2 Spatial Light Modulator (SLM)*

SLM's are compact devices that are capable of modulating the phase of an incoming wavefront, in reflection or transmission. This occurs on a fixed spatial pattern, where in a digital

driver receives a digital signal from a computer, and relays that information to the SLM as coherent optical information (or phase). Often employed in optical systems are liquid crystal layered SLM's with a silicon-based digital circuit underneath. Light first enters through a coverglass, which typically is coated with a transparent conductive film. It then impacts the liquid crystal layer, where it is finally reflected off of the silicon circuit; which contains a series of electrodes. The signal sent from the controller generates a 2D image containing a matrix of phase values that imprints on the silicon circuit. When this occurs, a voltage is produced for each corresponding electrode, which generates an electric field. This changes the liquid crystal structure and alters the phase of the incident beam.

An SLM is typically comprised of nematic liquid crystals. Nematic liquid crystals are a combination of organic molecules that resemble that of a rice grain. These molecules are aligned with the same orientation, typically, when no electrical response is given. However, when the nematic liquid crystals receive a voltage, an electro-optic response is given, whereby the molecules will rotate within their fixed position (disclination). This rotation creates a difference in the extraordinary ( $n_e$ ) and ordinary ( $n_o$ ) refractive indices, which creates a phase shift. Eqn. 1.21 showcases the voltage-dependent nature of the extraordinary and ordinary refractive indexes, and how they relate to the phase shift.

$$\varphi(V) = \frac{2\pi[n_e(V) - n_o]d}{\lambda} \quad (1.21)$$

In Eqn. 1.21,  $d$  is the thickness of the liquid crystal layer. Fig. 1.14 illustrates this rotational behavior, and how it affects the alignment of the crystals in relation to the silicon backplane and coverglass. Disclination of the nematic crystals is what ultimately allows for nematic SLM's to modulate the phase of an incoming wavefront and correct for aberrations.

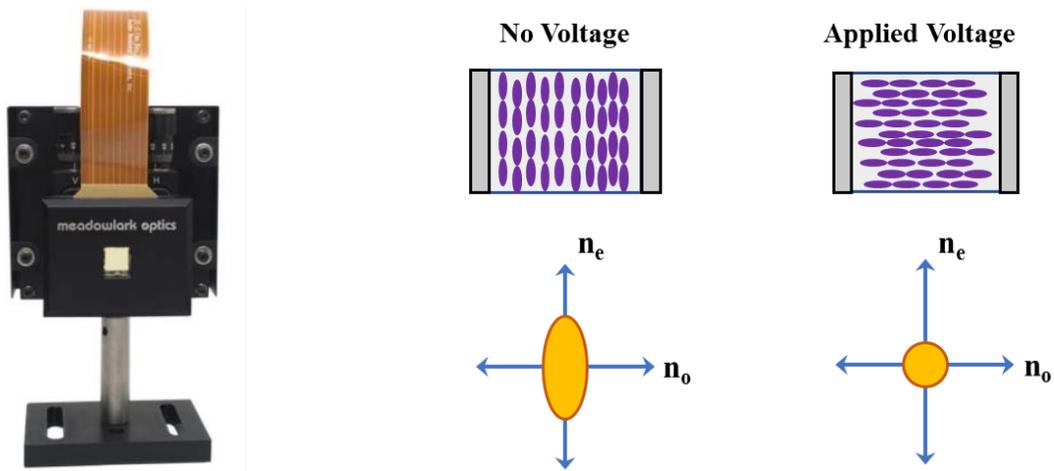


Figure 1.14: XY phase only SLM (Left) <sup>[54]</sup>. Polarization effect on the nematic crystals (Right)

## 1.7 Adaptive Optics applications

### 1.7.1 Adaptive Optics in Astronomy

With the digital era leading to ever more powerful and robust computers, the viability in harnessing active optics within ground and space-based telescopes has become more tangible. Large telescopes primarily operate with deformable mirrors, as they are capable of collecting a wider band of visible electromagnetic wavelengths and allow for continuous surfaces that produces higher quality images. The function of these deformable mirrors largely depends on whether the telescope is ground based, or in space. Ground based telescopes require that the deformable mirror remove any time-varying phase effects due to atmospheric turbulence. Space-based telescopes generally do not require the use of an AO, as they are not subjected to atmospheric turbulence; though there may be applications of DM's for space-based coronagraphy <sup>[55]</sup>. Space-based telescopes, in general, produce better quality images than their ground-based counterparts. This does not mean that ground-based telescopes have not seen significant improvements in their imaging quality; thanks in part to AO systems.

Numerous AO schemes have been developed in order to improve the smaller field of view provided by the AO wavefront sensor in ground-based telescopes. The schemes have proven their viability in enhancing the visual quality of stellar images for said telescopes; see Fig. 1.15 showcasing the planet of Uranus increased image quality with AO. Improvements could be seen - within the advent of the digital era - as early as 1988, with the development of a wavefront sensor technique known as Multi-Conjugate Adaptive Optics (MCAO). This technique addresses the cone effect that was common to many early telescopic AO systems, and extends the field of view such devices <sup>[56,57]</sup>. Modifications to the MCAO method have been performed in order to enhance field of viewing. As an example, using a double deformable mirror scheme that utilizes gradient information, in addition to scintillation, a corrected field of view up to 3 arcmin is possible <sup>[58]</sup>. Further developments have proven the effectiveness of such AO techniques <sup>[59-61]</sup>. There are other AO variations as well, including: eXAO (eXtreme-AO), GLAO (Ground-Layer AO), LTAO (Laser Tomography AO), and MOAO (Multiple Object AO).

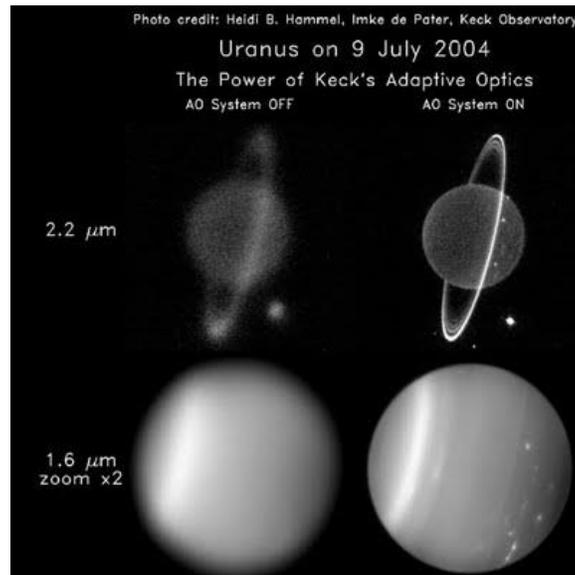


Figure 1.15: AO system improvement of a imaging Uranus <sup>[62]</sup>

Outside of such improvements, AO's allow for various methods of “seeing” through telescopes. Coronagraphy is the study of celestial bodies around a sun, where a telescope device known as a coronagraph is used to eclipse the sun; reducing solar noise and improving image quality. This method of astronomy benefits from the use of an AO system, allowing for diffraction-limited images that are accompanied with a high Strehl ratio <sup>[63,64]</sup>. This is merely only one type of viewing method, with direct imaging of star formations, our galactic center and solar system, and numerous other galaxies, likely continuing to be the most prominent type of viewing method when it comes to space observation. Many notable findings have been made by implementation of an AO to correct for atmospheric turbulence <sup>[65-67]</sup>. An example of a coronagraph can be observed below in Fig. 1.16. Images a) and b) is the star F606W and the reference PSF star HD 216149. Image d) is smoothed and shown using a stronger stretch to reveal the disk at greater heights and structure associated with the galaxy superposed on the northwest extension. A background star can be seen below the southeast extension.

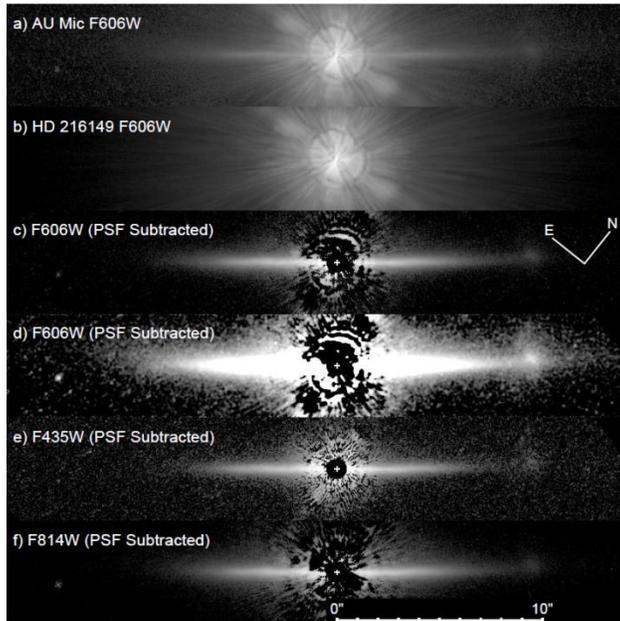


Figure 1.16: Coronagraphic image of star AU Mic F606W <sup>[68]</sup>

### 1.7.2 Adaptive Optics in Optical Manipulation

The use of AO's need not only be used to aid in the user's ability to detect or improve the visual quality of an object. They also offer a way forward in manipulating an optical source to suit a particular goal. One notable example is in the design of instruments known as Optical Tweezers. Optical tweezers utilize the gradient forces produced by a laser source (such as repulsive and attractive forces), in order to retain a particle inside of the beams waist; the particle is normally attached to a transparent bead (Fig. 1.17). The SLM is used, in conjunction with a computing device, in order to apply a series holographic patterns that creates optical. This allows for the microscopic beads (seen in the bottom right images) to be arranged in such a fashion. This technique has been of great use in biology, as it has allowed for the manipulation of live bacteria and viruses.

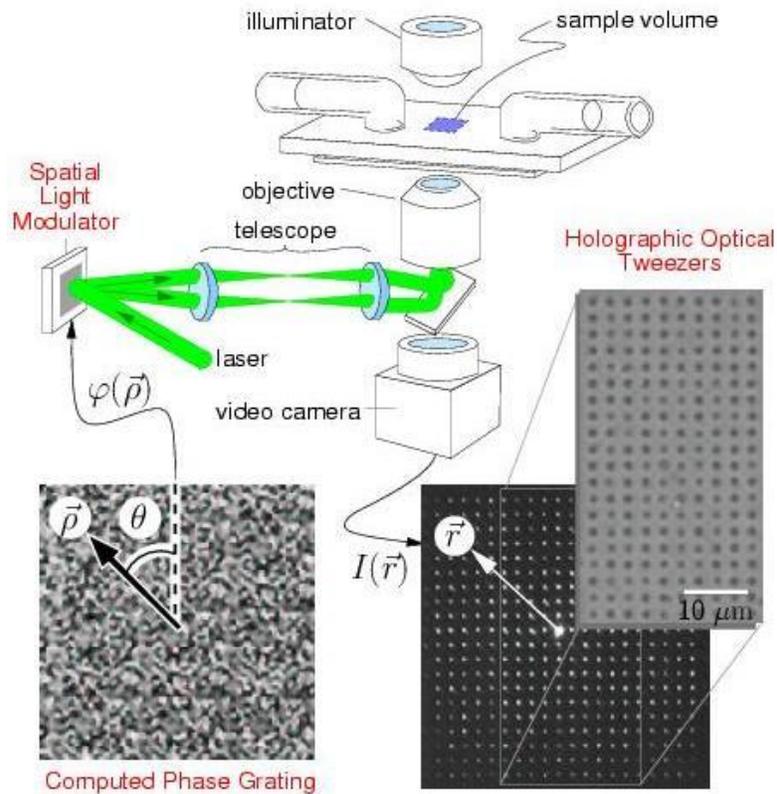


Figure 1.17: Optical tweezer system utilizing a spatial light modulator <sup>[69]</sup>

It is also applicable in the field of physics, as it allows for the measuring of forces between two particles; typically of micron size. AO's have been a key tool in altering the optical path and pitch of the light source used in such a setup, allowing for numerous optical traps to be generated [70]. Deformable mirrors, additionally, allow for aberration correction in such systems. Correcting these defects (such as spherical aberrations) improve the quality of the optical trap. Scientists at the University of Edinburgh were able to show that, by implementing a deformable mirror, the quality of the optical improved significantly [71]. Though their team did show that the trap quality diminished, when the trap is situated deeper inside the sample, the application of an AO still shows promising benefits for the science.

### *1.7.3 Adaptive Optics in Microscopy*

Within the life sciences, AO's have paved the way for improving the resolution and imaging potential of commonly utilized microscopy techniques. Researchers operating a scanning confocal microscope, a technique often utilized in the imaging of biological specimens, will likely attest to the difficulties of specimen-induced aberrations. These aberrations are the result of the differences in refractive indices among the various cell, membrane, and biological structures contained within the specimen [72,73]. This limitation hinders the imaging quality of the microscope and would only allow for a few number of cell layers to be imaged with a respectable level of detail. The same is also true for another imaging technique known as multiphoton microscopy. With the incorporation of AO's, such aberrations can be corrected (or outright removed), allowing for far deeper imaging into the sample, improved image quality, or optimum performance; see Fig. 1.18. The figure contains microscope images of labelled mouse intestine. Aberrations have been removed in order to improve visual quality

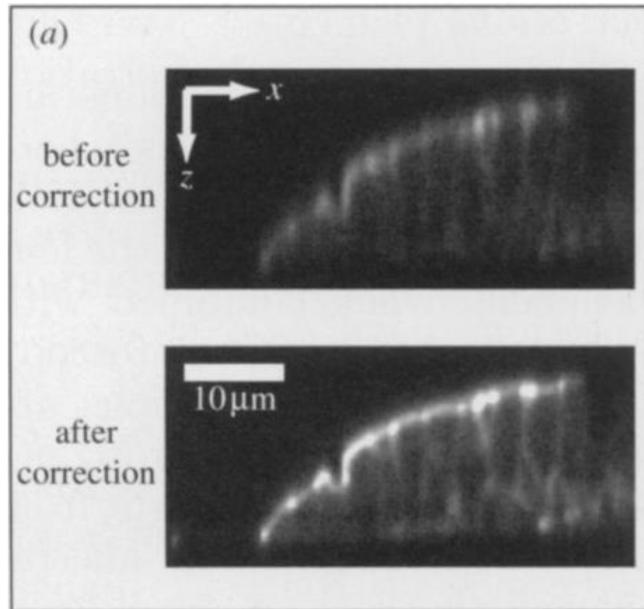


Figure 1.18: AO correction of confocal fluorescence microscopy image <sup>[72]</sup>

#### 1.7.4 Adaptive Optics in Medicine

As mentioned in the introduction, the human eye is prone to encountering several deleterious effects; in terms of imaging quality. This is caused due to several reasons, such as: Injury or damage to the eye or retina, hereditary disadvantage, or natural causes due to aging. In the medical branch of ophthalmology, improving upon one's eyesight would normally have the patient acquiring a pair of eyeglasses or contacts, which are fabricated with the corrected focal length and refractive index needed to improve viewing quality. While these visual aids can remove the effects of astigmatism and defocusing, they are still unable to correct for spherical, coma, and other irregular aberrations <sup>[74-76]</sup>. With advancements in AO, new techniques can be developed as a means of correcting for aberrations in the eye (to a greater degree than other techniques). As an example, an ophthalmoscope implementing a Shack-Hartmann wavefront sensor (detailed in section 1.8.2) and deformable mirror can be used to study the visual quality of a subject's. Their experiment resulted not only in the correction of the aforementioned uncorrected aberrations, but

also an improvement in the retinal Strehl ratio, and the generation of high-resolution retinal images. Though the initial costs of implementing a deformable mirror limited its availability during early investigations into its potential benefits, the practice has grown in recent years, with many differing schemes being proposed and developed for the correction – and discovery - of retinal defects <sup>[77-80]</sup>.

### *1.7.5 Adaptive Optics for beam shaping*

The ability for AO's to perform beam shaping has had a stark impact on numerous commercial industries and research areas that have required the use of a laser source; for various applications. Specifically, these advancements are seen in optical beam shaping, lithography, and machining. For an active optics to excel in areas such as these, not only must it provide a functionality that is equivalent – or better – than its optomechanical competitor, but it must also be of low cost; for it to be appealing. Given their fast response times, compact design, relatively low cost, and beam shaping qualities, AO's have developed into a powerful manufacturing tool. This section will briefly detail some of the major advantages that AO's devices have had in various industry fields, thanks in part to their beam shaping capabilities.

On a fundamental level, active optics allows for the reshaping of an incident beam into that of various other beam profiles. Though a specially crafted focusing lens, mask lens, or hologram, could be used to produce a specific profile, AO's provide a higher degree of flexibility. A spatial light modulator, for example, can be utilized as a means of altering a Gaussian beam into that of a Bessel Beam. A deformable mirror could be used in shaping both a Gaussian and super Gaussian beam. SLM's can also be used to generate a flat top beam <sup>[81,82]</sup>. Given that a particular beam profile can offer a host uses and benefits, the versatility that active optics provide is quite staggering. A Bessel-beam, for example, has applications within the realm of optical metrology

and light-sheet microscopy <sup>[83-86]</sup>. Beam shaping even allows for focusing through a scattered media <sup>[87]</sup>.

#### *1.7.6 Adaptive Optics in high-energy laser Physics*

In high-energy laser Physics, it is imperative for the optics system to have a highly efficient laser source. Often times, these systems employ various gain mediums and amplifiers to condense the beam, improving energy output, in order to generate peak power values ranging in the gigawatt range or higher. One of the most prominent aspects of the field is laser-induced thermonuclear fusion. Such a practice is designed to mimic the thermonuclear explosions that fuel our sun which, in terms of humanity, would mean an inexhaustible and pollution-free energy source; once contained <sup>[88]</sup>. This type of research is conducted primarily by government-based entities, such as the National Ignition Facility of the Lawrence Livermore National Laboratory in the U.S., LMJ (Laser Megajoule) facility in France, and the SG-III Laser Facility in China <sup>[89-91]</sup>. A common predicament in this field is high local fluence, which can cause damage to the optical instruments within the high-power laser system. Researchers have found that SLM's, in particular, are beneficial in this area because of their ability for improving the near-field beam quality; via beam shaping <sup>[92]</sup>. The SLM is capable of modulating the near field fluence in such a way that it allows for stable intensity propagation throughout the system. As a result of the SLM needing corrective information, this requires that it be housed within a feedback loop.

#### *1.7.7 Adaptive Optics in laser machining and fabrication*

High power lasers can also be put into place within the manufacturing industry. If enough energy is provided, a laser beam is more than capable of welding – or cutting – through metallic substances. This power is normally delivered via a laser pulse, typically on the nanosecond scale;

though it can reach into the femto- and picosecond scale. Delivery of an energy source within a small spot size, restricted to an even smaller time frame, allows a pulsed laser to ablate a target sample; removing its material. Deformable mirrors are already employed in this field, due to their beam shaping potential, as Gaussian beams are not always applicable to laser machining. SLM's have the capacity to compete with their bulkier cousin, but are generally hindered by high levels of peak power output on the ultrafast laser pulse scale. SLM's have seen developmental progress on this front, with demonstration of a water-cooled SLM to machine both a polyimide and metal film material <sup>[93]</sup>. Laser fabrication has also seen improvements from the usage of Active optics. The work of R. Simmonds showcases that using both a deformable mirror and SLM, in parallel, allows for improved three-dimensional fabrication results <sup>[94]</sup>. These techniques provide better beam quality, improved peak power, beam steering, and flexibility.

## **1.8 Wavefront Measurement**

With the fundamentals of wavefront theory presented, now the next critical part in most AO system is the wavefront sensor, the device that can measure the wavefront shape. This section will describe both the quantitative functions and optical instruments capable of allowing us to measure the wavefront of a light source.

### *1.8.1 Interferometers*

Interferometry is an optical technique that allows for the measurement of an optical light source wavefront. What is of importance to the research conducted in this dissertation, is its ability to use interferometry as a diagnostic tool when measuring the phase of the resulting interference pattern. It requires the use of two separate optical beams, which merge to a single propagation axis. This resulting superposition leads to an interference pattern; hence the name. Interferometers are a popular tool, as the nature of minute particle interactions within a light source makes it a

sensitive enough tool for making measurements of relatively negligible phenomenon. Today, this technique is employed in various research labs and industries, as well as highly sensitive projects. For example, interferometers are used to detect gravitational waves performed at LIGO (Laser Interferometer Gravitational-Wave Observatory) <sup>[95]</sup>; see Fig 1.19.



Figure 1.19: The LIGO sites in Livingston, Louisiana (left) and Hanford, Washington (right) <sup>[95]</sup>

Interferometers generally use a beam splitter in order to split a single beam source in two; through partial transmission. The two beams then impede upon reflective mirrors. Depending on the type of interferometer, one of the beams can be altered by adjusting the position of one of the mirrors. For example, the Mach-Zander Interferometer has one path (or “arm”) of the beam travel a slightly larger distance by moving the position of the mirror along the beams propagation axis. Regardless of the interferometer used, the two beams will then merge after reflecting off their respective mirrors. After recombination, a *fringe pattern* will form. The fringe pattern is caused by a shift in the phase of the wave coming from both beams. Eqn. 1.22 can be used to describe the amplitude of the fringes, with  $A(x,y)$  being the background illumination,  $B(x,y)$  being the amplitude modulation of the fringes,  $\varphi_c$  is the carrier frequency, and  $\varphi_s$  is the wanted phase signal <sup>[96, 97]</sup>.

$$I(x, y) = A(x, y) + B(x, y)\cos\{\varphi_c(x, y) + \varphi_s(x, y)\} \quad (1.22)$$

If a phase shift causes the waveform to be rotated 180 degrees, in comparison to the unaltered beam, this will create *deconstructive interference*; creating an area with little to no intensity. The smaller the degree in phase shift, the more intense the light source will be; referred to as *constructive interference*. An example can be seen in Fig. 1.20<sup>[98]</sup>. The black rings correspond to deconstructive interference, while the bright rings to constructive interference.

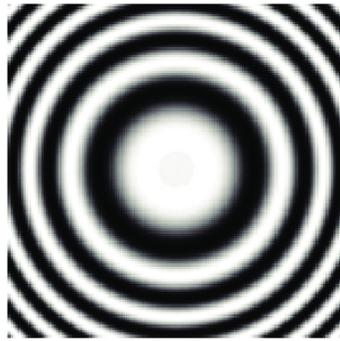


Figure 1.20: Example of a circular interference pattern<sup>[98]</sup>

Of all the interferometric methods available, the one that will be most suitable to the experiments detailed in this dissertation work is the Michelson Interferometer which uses a single beam splitter and measure in the merged beam the phase difference between the two arms. The method used to extract the 2D phase from the interferogram makes uses of two-dimensional Fourier Transform and is detailed in chapter 2, section 2.3. It should be noted that one core drawback of interferometry is that it is using itself as a reference, in that you cannot measure the incoming wavefront directly. Only the wavefront created between the two arms of the interferometer. Other wavefront sensors, like the Shack-Hartmann Wavefront Sensor, detailed in next section, are able to make absolute (not referenced) measurements of the wavefront.

### 1.8.2 Shack-Hartmann Wavefront Sensor

The Shack-Hartmann Wavefront Sensor was originally conceived as a means of improving the imaging quality of satellite images; a need made necessary by the United States Air Force in the 1960's<sup>[99]</sup>. A Hartmann sensor was originally constructed using an aperture with a circular hole. This aperture could be translated across an entrance pupil, and when the hole was moved a distance equal to its diameter, a measurement would be taken. This local tilt would be measured was dependent upon a value of the transverse ray aberration,  $T$ , which is shown in Eqn. 1.23.

$$\frac{dw}{dy} = -\left(\frac{nr}{R}\right)T \quad (1.23)$$

where  $dw/dy$  is the local tilt of the wavefront,  $n$  is the refractive index,  $r$  is the radius of the pupil, and  $R$  is the curvature of the reference sphere. By combining these localized tilts together, it is possible to rebuild the wavefront.

Reconstruction of the wavefront revolves around the careful analysis of the wavefront slope, as well as the manner in which the wavefront gradient is represented. Wavefront slope errors are the mathematical deviations of a wavefront across a surface from its mathematically perfect form<sup>[100]</sup>. This slope is measured by comparing the reference wavefront to that of the measured centroids; the optical profile produced along the optical axis exiting the lenslet. Assuming the reference centroids to be of the form  $(x_r, y_r)_k$ , and the measured centroids to be of the form  $(x_c, y_c)_k$ , the wavefront slope distribution can be illustrated in Eqn.1.24<sup>[101]</sup>.

$$\begin{pmatrix} \langle \partial w / \partial x \rangle \\ \langle \partial w / \partial y \rangle \end{pmatrix}_k \cong \frac{1}{L_H} \begin{pmatrix} x_c - x_r \\ y_c - y_r \end{pmatrix}_k \quad (1.24)$$

where  $L_H$  is the distance between the lenslet array and the detector. This slope formula ties into the wavefront gradient (Eqn. 1.25), which describes the amount of tilt deviation produced across the lenslet.

$$\nabla W = \frac{\partial w}{\partial x} \hat{i} + \frac{\partial w}{\partial y} \hat{j} \quad (1.25)$$

The wavefront gradient can be represented a few numerical notations that allows to reconstruct the wavefront. The two prevailing methods are *Zonal* and *Modal* reconstruction. In the case of Zonal reconstruction, integrating the wavefront gradients on a lenslet-by-lenslet basis can generate a wavefront slope of the form seen in Eqn. 1.26. Through least-square fitting, this leads to the reconstructed wavefront seen in Eqn. 1.27 <sup>[102]</sup>.

$$\beta_0^x = \frac{1}{2d_x} [w(x_0 + d_x) - w(x_0 - d_x)] \equiv \frac{1}{2d_x} (w_1 - w_{-1}) \quad (1.26)$$

$$w_{i+1,j} = w_{i-1,j} + 2d_x \beta_{i,j}^x \quad ; \quad w_{i,j+1} = w_{i,j-1} + 2d_y \beta_{i,j}^y \quad (1.27)$$

For Modal reconstruction, the wavefront is expressed as a polynomial expansion over the entire pupil. The measured wavefront slopes is fit against the wavefronts analytical derivatives, leading to the value of the wavefront modal coefficients. Using the mean-square error and the trace of  $(A^*A)^{-1}$ , where  $A^*A$  is an  $N \times N$  symmetrical matrix, the modal reconstruction can be representation as Eqn. 1.28.

$$C_{MR} = \frac{1}{h^2} \text{tr}(A^*A)^{-1} \quad (1.28)$$

The previous method of aperture wavefront sensing was eventually set aside for a more robust method. Instead, the practice of parallel processing was developed. This practice involved the usage of subapertures, each containing its own lens, with said lenses having their own position-

sensitive detector. This would be further refined by the use of a lenslet array. Each waveform tilt  $T$  is measured within a particular location of the pupil, where upon the value  $T$  is measured using a respective Point-Spread Function (PSF). A functional benefit of such a setup is the reduction in the PSF of each optical source. The point-spread function defines how much of the light, and to what radius (or distance), will travel from an optical point source. This setup is the most common of the Shack-Hartmann wavefront sensors used today. Fig. 1.21 showcases a lenslet array, where an incoming wavefront (suffering from wavefront distortion) produces circular optical patterns that do not align properly with the optical axis of each subaperture. When a distorted wavefront impacts the detector, the circular patterns that are produced are misaligned.

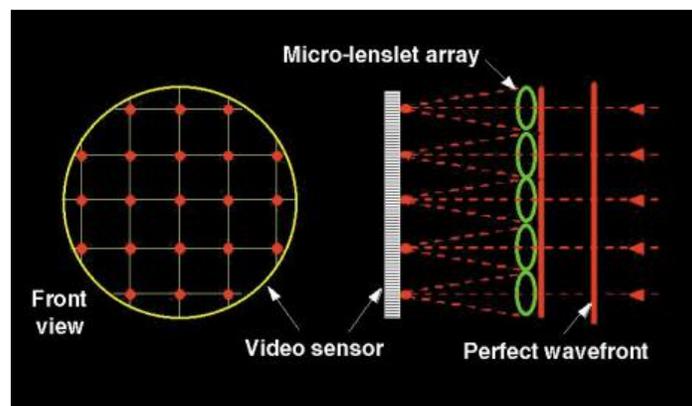


Figure 1.21: Detail of a lenslet array for wavefront sensing <sup>[103]</sup>

Shack-Hartmann wavefront sensors have enjoyed a great deal of popularity in modern history, due to their low-order aberration correction measurements and their ability to improve imaging quality. The device has become a staple in the astronomy community and has even been implemented within the Hubble Space Telescope using a proprietary sensor developed by Adaptive Optics Associates. Outside of astronomy, Shack-Hartmann sensors have also been utilized in ophthalmology as an aberrometer; which is utilized in the detection of higher-order aberrations in the pupil of a patient <sup>[104-106]</sup>. The Shack-Hartmann systems also see benefits when

implemented – in tandem – with an active optics. Often times, this implementation involves the removal of the microlens array in favor of a liquid crystal display or deformable mirror, which can generate microlenses on its own. Ultimately, this lends itself to the pre-correction of the initial wavefront by correcting for higher-order aberrations <sup>[107-109]</sup>.

### 1.8.3 Zernike Polynomials and the Classification of Aberrations

Each of the known types of aberrations are understood by their polynomial forms. Such polynomials fall under what are known as *Zernike Polynomials*. Derived by 1953 Nobel Prize winner Fritz Zernike, these circular polynomials were constructed to be orthogonal over the interior of a unit circle <sup>[110,111]</sup>. These polynomials were used to balance classical aberrations within a power series expansion and are often employed as a means of studying the behavior of small aberrations in diffraction images; within an optical system <sup>[112]</sup>. Typically, the case of a circular pupil is applied in structuring the polynomial of an aberration. When polar coordinates are considered for points across the pupil, the wave aberration function can be expressed as Eqn.1. 29.

$$W(\rho, \theta) = \sum_{n=0}^{\infty} \sum_{m=0}^n \left[ \frac{2(n+1)}{1+\delta_{m0}} \right]^{\frac{1}{2}} R_n^m(\rho) \cdot (c_{nm} \cos m\theta + s_{nm} \sin m\theta) \quad (1.29)$$

where  $c_{nm}$  and  $s_{nm}$  are the aberration coefficients,  $\rho = r/a$ , and  $a$  is the radius of the pupil <sup>[113]</sup>. The radial (circular) polynomials,  $R_n^m(\rho) \cos m\theta$ , and  $R_n^m(\rho) \sin m\theta$ , are expressed in polar coordinates as well, and are set within a radial notation; with  $l$  (angular momentum) and  $n$  (radial degree) being integers. In this particular case,  $n$  also represents the highest order of  $\rho$ . By setting the absolute value of the angular momentum equal to the azimuthal integer ( $m = |l|$ ), the radial polynomials can be expressed as Eqn. 1.30.

$$R_n^{\pm m}(\rho) = \sum_{s=0}^{\frac{n-m}{2}} (1)^s \frac{(n-s)!}{s! \left(\frac{n+m}{2}-s\right)! \left(\frac{n-m}{2}-s\right)!} \rho^{n-2s} \quad (1.30)$$

Aberrations are described by their shape, and fall under the following types: Piston, Tilt, Focus, Spherical, Coma, Astigmatism, Field Curvature, and Distortion. There are also higher-order aberrations as well. Each is given by their own ordered Zernike polynomial (Eqn. 1.31). Robert J. Noll would later tabulate these Zernike equations by ordering both their  $n$  and  $m$  components together and assigning each polynomial a Noll index ( $Z_j$ )<sup>[114]</sup>. A more detailed derivation of the Zernike Polynomials, as well as a list of their Noll indexes, can be found in Appendix A.

$$\begin{cases} Z_{\text{even}j}(\rho, \theta) = \sqrt{2(n+1)} R_n^m(\rho) \cos m\theta; & m \neq 0 \\ Z_{\text{odd}j}(\rho, \theta) = \sqrt{2(n+1)} R_n^m(\rho) \sin m\theta; & m \neq 0 \\ Z_j(\rho, \theta) = \sqrt{n+1} R_n^0(\rho); & m = 0 \end{cases} \quad (1.31)$$

As an example, a Zernike Mode number of 2 and 3 ( $Z_{2,3}$ ) would be indicative of a polynomial producing a “tilt”. An illustration of these polynomials can be seen in Fig.1.22. The Zernike polynomials are also used to categorize aberrations of the cornea and lens. This ability to categorize each known aberration will serve a role when seeking to correct the aberrated wavefront of a Gaussian beam used in chapter 2.

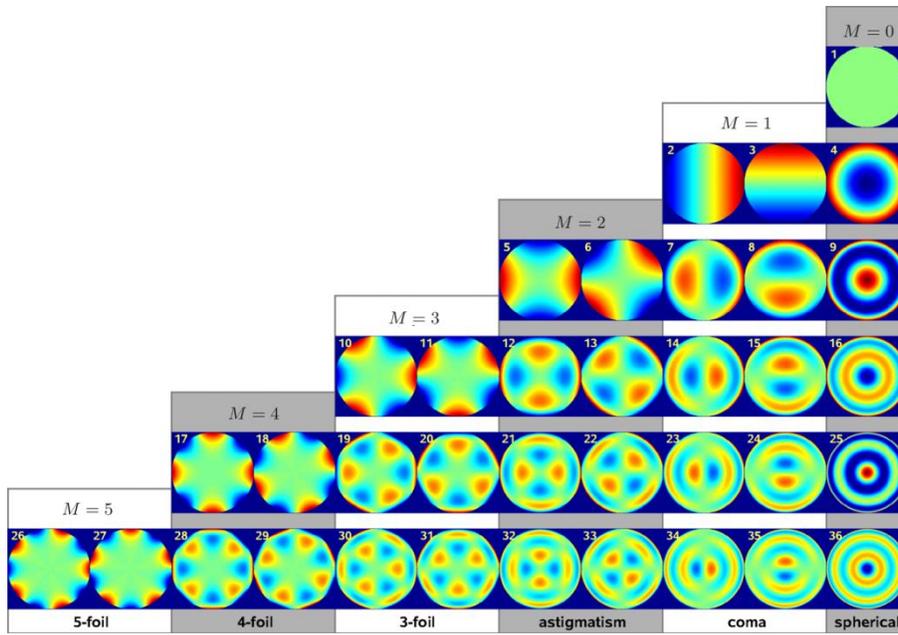


Figure 1.22: 2D Illustration of the 36<sup>th</sup> first Zernike Polynomials <sup>[115]</sup>

## **CHAPTER 2 GENETIC ALGORITHM FOR WAVEFRONT CORRECTION**

### **2.1 Introduction**

In the previous chapter, the basics of adaptive optics were introduced, with the main concept being the wavefront aberrations and the need to correct them. This chapter revolves around the use of a SLM (Spatial Light Modulator) to correct for wavefront aberrations that are purposely induced upon an incident laser beam, and more specifically, the description of the Genetic Algorithm (GA) that was developed to perform the wavefront correction. As there are several details that need to be outlined, such as the design of the wavefront correction system, an understanding of how phase is measured & extracted in the experimental setup, and the fundamentals of a GA, this chapter elaborates on each of these subjects. Afterwards, this chapter transitions to experiments involving wavefront correction, through a combination of GA optimization and SLM modulation, of both a Gaussian beam containing small aberrations induced by the non-flat initial SLM surface and an aberrated Gaussian beam, with large aberrations generated by a plastic coverslip. Finally, this chapter concludes with an analysis and discussion of these correction results, as well as assessing the algorithms performance. Ultimately, the goal of this chapter is to demonstrate how evolutionary algorithm techniques (such as the GA) can be extraordinarily powerful methods in shaping a wavefront and optimizing the focal plane distribution of a laser.

### **2.2 Experimental setup for wavefront correction**

Chapter 1 briefly presented that one of the approaches for wavefront correction is through direct wavefront measurement, with correction being performed in either a closed loop or open loop design (cf. section 1.5), and the wavefront information used as the feedback to know how to

control the active optics. Another approach is wavefront correction via the optimization of a signal created by the laser **other than the wavefront**, in conjunction with a learning algorithm using the signal to refine an optimization goal [116]. It is this second approach that has been explored in this dissertation, as the learning algorithm will instead replace the wavefront sensor commonly employed in a closed-loop AO system. The learning algorithm that was developed, a Genetic Algorithm (GA), is tested in this chapter in a classic wavefront correction manner, where a signal originating from the focal plane is used as the feedback for the correcting optics. The experimental setup is shown in Fig. 2.1. The signal used for feedback is the 2D spatial profile of a solid state 488 nm continuous wave (cw) blue laser measured by a camera, and the correcting optics is a XY, phase-only 512 pixels by 512 pixels SLM. The spatial profile CCD and SLM work in conjunction inside the GA to modulate the phase.

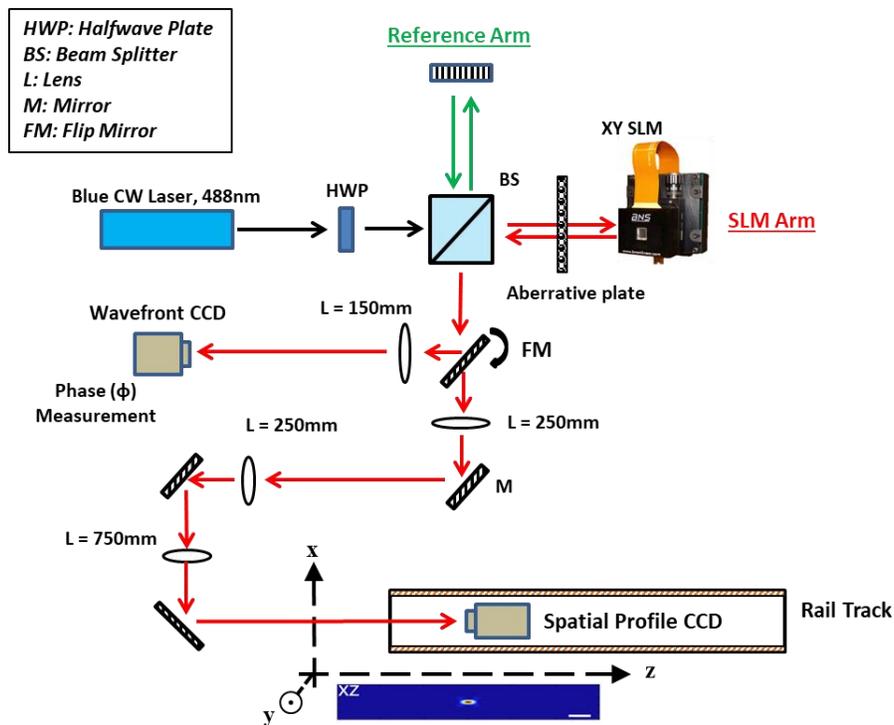


Figure 2.1: Genetic algorithm for wavefront correction and phase measurement setup

Given the phase modulation capabilities of the SLM, being able to acquire the optical wavefront - the bi-dimensional phase map – is critical to understanding what phase is created by the learning algorithm. For this experiment, a Michelson Interferometer has been utilized and its details are described in section 2.3 explaining the phase measurement. The SLM is stationary and serves to modulate the phase. In general, the surface of an active optics (SLM or DM) is not perfect at their neutral shape (0V applied to each pixel for our SLM), which creates small aberrations within the laser beam. However, since this method is intended to correct for wavefront aberrations, the learning algorithm would be able to correct for these small aberrative effects. The controller for the SLM sends the phase array in voltage, and a calibration procedure was developed to obtain a LUT (Lookup Table) (see procedures in Appendix B) to linearize the phase produced by the voltage input. The interferometer phase measurement is also used in order to check the validity of the LUT calibration procedure. The modulated beam can then travel down one of two arms (cf. Fig. 2.1), with the beam being directed by way of a flip mirror.

With the flip mirror in the upright position, the beam is sent to the **interference arm of the setup**, to measure the interference pattern, and hence deduce the wavefront (2D spatial phase). A 150mm focusing lens images the plane of the SLM by a 12-bit monochromatic CCD (Charged Coupling Device) camera (Sentech 1” USB3 camera), which will be referred to in this dissertation as the phase camera or the wavefront camera. More specifically, the lens allows the SLM image plane to be magnified (after focusing) in order to fill the wavefront camera.

With the flip mirror folded down, the beam can freely propagate along the **GA arm of the setup**, where the effect of the phase created by the SLM can be measured on the laser spatial profile. The beam first encounters a series of collimating lenses (both 250mm) that form an image of the SLM surface on a 750 mm lens, which then focuses the beam onto a 12-bit monochromatic

CCD camera (Allied Vision Technologies Manta G-033 CCD camera), which will be referred in this dissertation as the spatial profile camera. The spatial profile of the beam is recorded by this camera and serves as a feedback signal for GA experiments. The camera is set atop a one-meter track, allowing its position to be adjusted along the optical axis. As the intent is to image the modulated beam, when viewing from the GA arm, the reference mirror is blocked to only allow the modulated beam to be sent through the beam splitter.

Grabbing from the CCD cameras, applying a phase to the SLM, and operating the GA, requires the use of a computer interface. In this regard, a graphical user interface (GUI) known as National Instruments LabVIEW was used to develop the genetic algorithm, acquisition and control software. The advantage of using this type of application is that it allows for ready access in observing, diagnosing, and analyzing various elements within the setup, as LabVIEW is also capable of interfacing with a number of third-party elements; such as a CCD camera. Through this application, images collected by the CCD cameras are visualized in real-time, in addition to performing wavefront analysis when using the interferometer arm, or optimization with the learning algorithm developed for wavefront optimization when using the GA arm of the setup. Parameters associated with these experiments can also be changed in real-time.

### **2.3 Phase Analysis**

When using the interferometer arm, the wavefront CCD records an interference pattern that contains the phase difference between the arms, and consequently the 2D phase map (or wavefront) induced by the voltage pattern on the SLM. In chapter 1 (section 1.8.1), interferometry was introduced as a method of measuring the wavefront of an incoming light source. Let's now derive the equation for the fringe pattern corresponding to the Michelson interferometer used in the wavefront correction experiment.

### 2.3.1 Derivation of the fringe pattern equation

The interferometer-only portion of the experiment is shown in Figure 2.2. When the flip mount (FM) and the reference arm are used, a Michelson interferometer is formed which allows to measure the phase difference between the two arms, consequently the wavefront distortion from any object placed in the SLM arm. Typically, the standard design of a Michelson Interferometer consists of an optical source, two mirrors in each arm, a beam splitter, and a detector. A half waveplate is set before the beam splitter to polarize the incoming beam by  $180^\circ$  in order to obtain the correct polarization required by the SLM. The optical beam is first divided in two different directions using a cube beamsplitter. Specifically, half of the beam is reflected, while the other half is transmitted. Upon leaving the beam splitter, half of the beam encounters a dielectric mirror (reference mirror), while the other encounters the SLM (reflective surface).

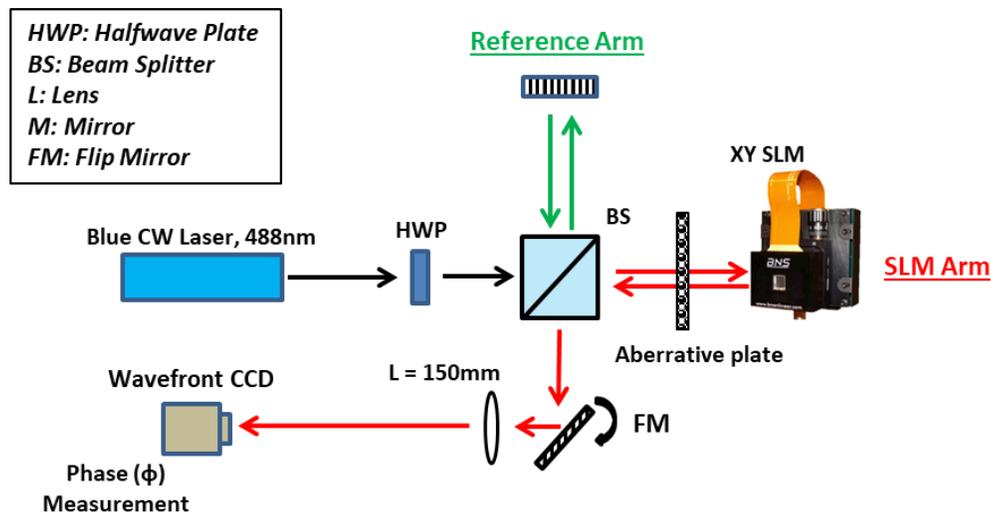


Figure 2.2: Wavefront measurement arm of the original setup.

The reference mirror can be tilted for the purpose of creating a linear phase difference between the beam reflected off its surface and the beam reflected off of the SLM. This difference is what creates the interference fringes. The beam produced by the 488nm diode laser has an

inherent phase term,  $\varphi_0(x,y)$ . Since the reference mirror only reflects the beam, this term is unchanged and can be represented as the following wave equation (Eqn. 2.1) for the reference arm.

$$E_{ref}(x, y) = E_0(x, y)e^{i[\varphi_0(x,y)]} \quad (2.1)$$

For the beam impeding upon the SLM, a phase term,  $\varphi_{SLM}(x,y)$  is added upon reflection, if the optical quality of the reflective surface is not perfect. This additional phase is often provided by the SLM manufacturer or can be measured when the SLM pixels are all at 0 volts. The voltage pattern written on the SLM adds a shaping phase,  $\varphi_{shaping}(x,y)$ . These phases add up to give the wave equation (Eqn. 2.2) for the light coming back from the SLM and recombining in the beamsplitter with the reference arm light.

$$E_{SLM}(x, y) = E_0(x, y)e^{i[\varphi_0(x,y)+\varphi_{SLM}(x,y)+\varphi_{shaping}(x,y)]} \quad (2.2)$$

Each beam contains the same amplitude ( $E_0$ ) as they originated from the same laser source. The superposition of both waves after recombination after the beamsplitter leads to a new wave with an altered amplitude and phase, notated as  $E_{sum}$  (see Eqn. 2.3).

$$E_{sum}(x, y) = E_0e^{i[\varphi_0(x,y)]}[1 + e^{i[\varphi_{SLM}(x,y)+\varphi_{shaping}(x,y)]}] \quad (2.3)$$

The camera does not record the electric field amplitude  $E_{sum}$ , but its intensity  $I_{sum} = |E_{sum}|^2 = E_{sum}E_{sum}^*$ . The complex amplitude multiplied by its complex conjugate leads to the following equation:

$$I_{sum}(x, y) = E_0^2(1 + e^{i\varphi})(1 + e^{-i\varphi})$$

$[\varphi = \varphi_{SLM}(x, y) + \varphi_{shaping}(x, y)]$  is the total phase difference between the two arms. Note that the phase term  $\varphi_0(x, y)$  disappeared, which means the initial wavefront from the laser cannot be measured using interferometry, which is the drawback of using interferometry which only measure a phase difference, not a light source absolute phase. Additional manipulation using Euler's formula  $e^{i\varphi} = \cos \varphi + i \sin \varphi$  leads to Eqn. 2.4 describing the 2D fringe pattern.

$$I_{Sum}(x, y) = E_0^2(1 + e^{i\varphi} e^{-i\varphi} + e^{i\varphi} + e^{-i\varphi}) = I_0(2 + 2 \cos \varphi)$$

$$I_{Sum}(x, y) = 2I_0[1 + \cos(\varphi_{SLM}(x, y) + \varphi_{shaping}(x, y))] \quad (2.4)$$

This equation is given further significance in section 2.5, where it is outlined as to how the setup detects both the correction and reshaping phase terms.

### 2.3.2 Phase Extraction using Fourier Transform Analysis

The interference pattern generated by two superimposing waves creates a two-dimensional interferogram described by (Eqn. 2.4). In the case of the Michelson interferometer with the reference mirror introduced a linear phase (tilt), much of the wavefront information collected by the wavefront CCD is received as a spatial frequency carrier, since the phase shift is created by tilting the reference mirror. Fig. 2.3 showcases the interferograms of multiple types of fringe patterns, with and without aberration <sup>[117,118]</sup>. In the absence of aberration, tilt, and focusing effects, the phase of the fringe pattern should be a similar across all points of its surface (i.e. “flat”); creating an ideal plane wave.

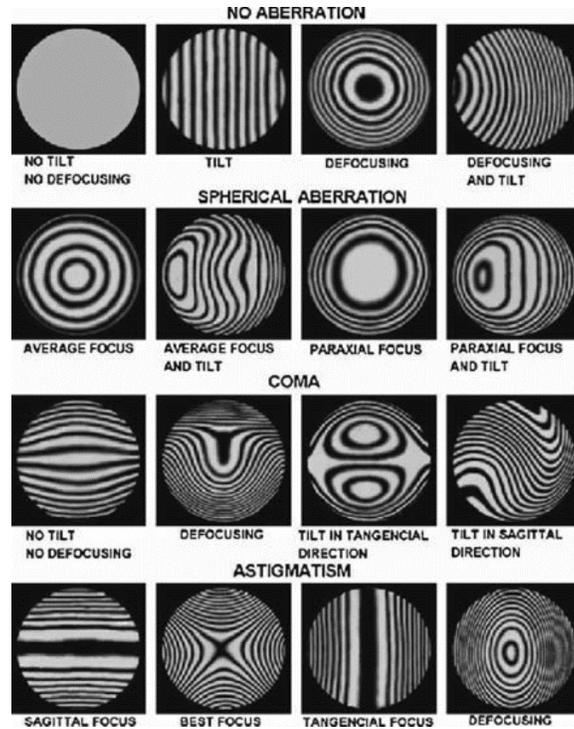


Figure 2.3: Fringe patterns subjected to well-known Zernike aberrations <sup>[118]</sup>

As phase values change across the interferogram, this leads to changes in the fringe pattern of the interferogram. When the phase has a sharp change, the fringes will look more closely spaced, as each change from bright fringe to the next bright fringe corresponds to a shift of  $2\pi$  in the phase. For example, going back to Fig.2.3, when looking at the interferogram with no aberration, but subjected to tilt (second panel labeled “tilt”), there a number of dark and bright lines interspersed throughout the interferogram, and the range of phase is roughly 10 waves of tilt.

To extract the raw phase and intensity values from the interferogram, it must first be demodulated. There are several demodulation techniques available, including: space-domain phase demodulation, sinusoidal window filter demodulation <sup>[119]</sup>, and phase-locked-loop demodulation <sup>[120-122]</sup>. Though many of these techniques have their own strengths and weaknesses, given that this experiment primarily deals with a linear carrier, a Fourier Transform phase demodulation technique is more suited for this experiment.

The *FT* (*Fourier Transform*) is a powerful mathematical and signal processing tool that transforms a function (or signal) that is described in the real domain (here spatial dimension  $x$  and  $y$ ) into a function that is described in frequency domain (here spatial frequencies  $f_x$  and  $f_y$ ). The equation for FT is represented in Eqn. 2.5, and the IFT (Inverse Fourier Transform) is represented in Eqn. 2.6. This is a 1D case, where  $x$  is the real-space domain variable,  $f$  is the frequency domain variable, and  $G(f)$  is the magnitude (or Fourier spectrum) of the real-space function  $g(x)$ .

$$G(f) = \int_{-\infty}^{\infty} g(x)e^{-i2\pi fx} dx \quad (2.5)$$

$$g(x) = \int_{-\infty}^{\infty} G(f)e^{i2\pi fx} df \quad (2.6)$$

By applying the FT to the interferogram, the spatial information is converted into the frequency domain; see Fig. 2.4 for some FT examples.

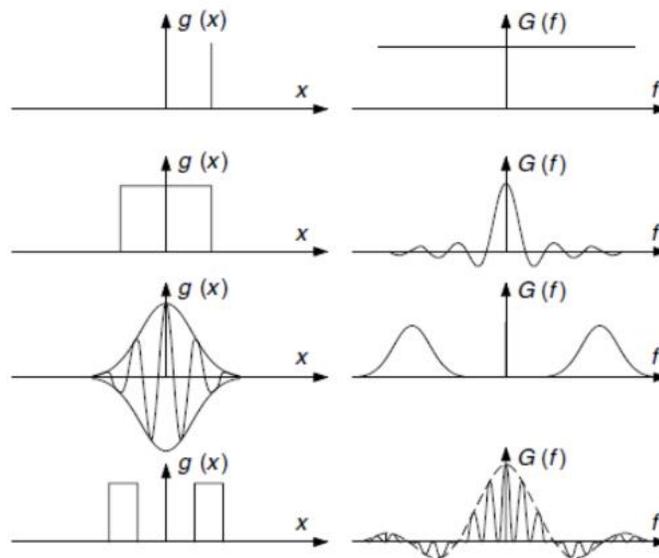


Figure 2.4: Examples of known Fourier Transform  $g(x) / G(f)$  pairs <sup>[118]</sup>

Figure 2.5 presents the different steps to go from the initial interferogram (Fig. 2.5a) to a 2D unwrapped phase (Fig. 2.5f), as was developed in the LabVIEW analysis software to allow

real-time phase extraction and display. Although a traditional discrete FT is mathematically sound, a *Fast Fourier Transform (FFT)* [96,123] method using matrix decomposition is preferred, as it requires less computational time. Subfigure 2.5b displays the magnitude of the interferogram FFT converted onto a logarithmic scale image to be able to visualize the central peak (zeroth-order frequency) and the sidelobes. Higher frequency values are situated further away from this zeroth-order frequency. The raw 2D phase and 2D intensity (Fig. 2.5d and 2.5e) are extracted from the 1<sup>st</sup> order sidelobe by inverse Fourier Transform.

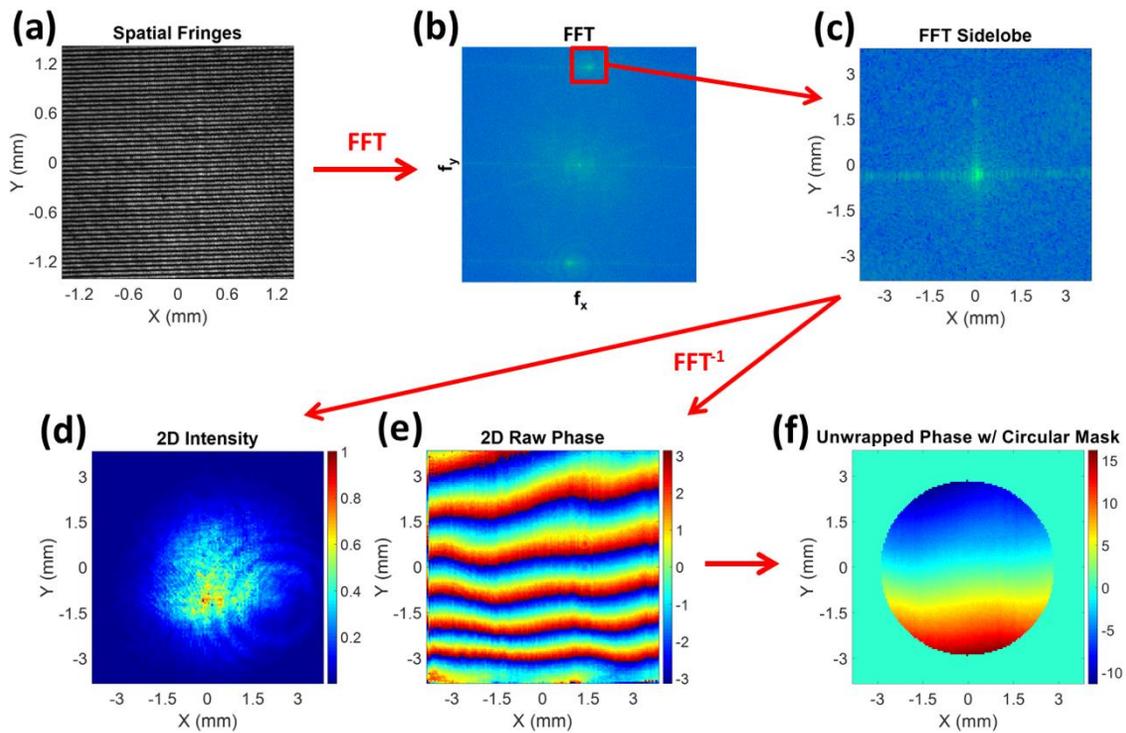


Figure 2.5: Illustration of wavefront extraction using Fourier Transform method.

The raw 2D phase contains  $2\pi$  phase jumps, which means the values representing the phase are restricted to a set boundary of  $-\pi < \phi < \pi$ ; or  $-180^\circ < \phi < 180^\circ$ . To unwrap the phase, whenever the phase values exceed this boundary, it can be increased or decreased by a multiple of  $2\pi$  (or

360°). Unwrapping allows the full extent of the phase to be displayed and helps in mapping the surface of the correction/reshaping wavefront being imparted unto the beam. During the unwrapping process, it is possible for noise fluctuations in the camera to create unwanted phase discontinuities as information is collected further away from the central beam spot, where the light intensity is close to the background noise of the camera. In order to alleviate this issue, a circular mask is overlaid on top of the wrapped phase before unwrapping. This masking technique mitigates these discontinuities, while also making it easier to measure the Peak to Valley (PV) and Root Mean Square (RMS) values of the 2D phase, as detailed in (Eqn. 2.7) and (Eqn. 2.8).

$$PV(\varphi) = Max(\varphi) - Min(\varphi) \quad (2.7)$$

$$RMS(\varphi) = \sqrt{\frac{1}{N} \sum_{all\ pixels} (\varphi - \langle \varphi \rangle)^2} \quad (2.8)$$

## 2.4 Genetic Algorithm

### 2.4.1 Introduction

Human history is rife with inventions that drew inspiration from natural and biological phenomenon. In the realm of information processing and computer science, the biological process of evolution serves as the basis for a computing technique known as a *Genetic Algorithm*. A Genetic Algorithm (GA) is an evolutionary algorithm that explores a number of randomized solutions in order to better optimize a target criterion of interest <sup>[124]</sup>. What is beneficial about this programming technique is its ability to find a solution that cannot be found analytically. Not only that, but due to the evolutionary and randomized nature of the algorithm, it is possible to generate a number of distinct solutions that could – potentially – achieve the same target.

The core reasoning behind the GA process stems from natural selection and evolution <sup>[125]</sup>. More precisely, the process is analogous to that of selective breeding. Selective breeding (or

artificial selection) is a process by which animals, or plants, breed with one another in order to express a desired phenotype (or trait). This phenotype is targeted in order to suit a particular goal. For example, a breeder may want to express a specific phenotype in plants that allows them to survive in colder climates. When viewed in the confines of this experiment, the flow of the GA can be best illustrated in a logical diagram as shown in Fig. 2.6.

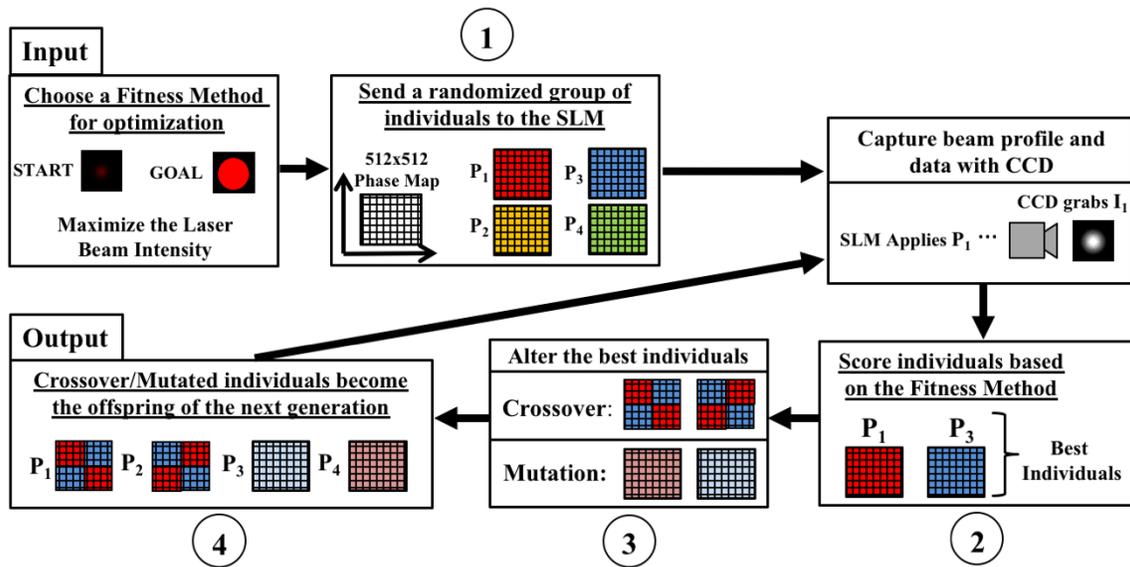


Figure 2.6: Genetic Algorithm logical diagram

A typical GA routine is defined as the following: (1) Create a set of initial, randomized solutions that serve as the “parents” of the first generation; (2) the GA selects a group of parents that have traits which are suitable to the targeted fitness method; (3) these parents reproduce with each other, creating new “offspring” through genetic operation techniques; (4) the GA then brings these offspring into the next cycle (generation). This process can be repeated up until a set termination point <sup>[126]</sup>. Termination of the GA process is dependent upon what the user determines to be an appropriate condition for ending it; such as: The target solution is found by meeting the minimum criteria, the maximum number of preset generations has been reached, the “best”

solution is observed by the viewer, etc. <sup>[127]</sup>. It is preferable to break down each of these steps to better understand the GA optimization process.

#### 2.4.2 *GA Initialization*

The ability for the GA to reach a desired solution first lies in the starting parameters. These variables alter the randomization, optimization, and reproduction of the GA, and ultimately if the algorithm can converge or not to a solution. These parameters include: the number of generations within each iteration of the GA, the population size, which fitness method (or operator) the GA uses, the required amount of “genetic” diversity within each generation, and the probability (or amount) of genetic crossover and mutation; to name a few. Both the parameters set and genes constructed – from a programming standpoint – are realistically given as bits; which are sent to a control system. Most of the time, these parameters do not act independently, and interact with one another at some point; known as *epistasis* <sup>[128,129]</sup>. Ultimately, the goal of setting these parameters is to optimize the functions being used so that they lead to a target solution. Depending on the complexity of the solution needed to be achieved, the number of parameters (or flexibility of said parameters) is likely to increase as well. This is in addition to alterations being made to the way genes interact with each other.

A useful term in defining how easily the GA can reach a solution, based on set parameters, is known as the **search space**. As the search space grows in size, based on the number of genes needed, or the information each parameter requires, the computational load increases. This translates as more time needed before the GA is able to terminate or find a solution. As an example, the SLM accepts values within an array of 512x512 pixels. The phase creation methods that are used only takes a select number of phase terms and creates a mesh grid that encompasses the entirety of the SLM array, with 512x512 phase arrays – when multiple terms are present – adding

upon each other. The maximum search space the GA could explore would be to have one phase term per SLM pixel, meaning that a search space containing 512x512 (or 262,144) terms (genes) could be chosen. A scenario like this is unreasonable, however, and would require too much computational effort and time. Fig. 2.7 illustrates the nature of how the search space operates in the case of only two input phase terms / genes. Two inputs (genes) are used as a means of finding a solution, which leads to a two-dimensional search space. The GA randomizes the values of these inputs in order to seek out the best solution.

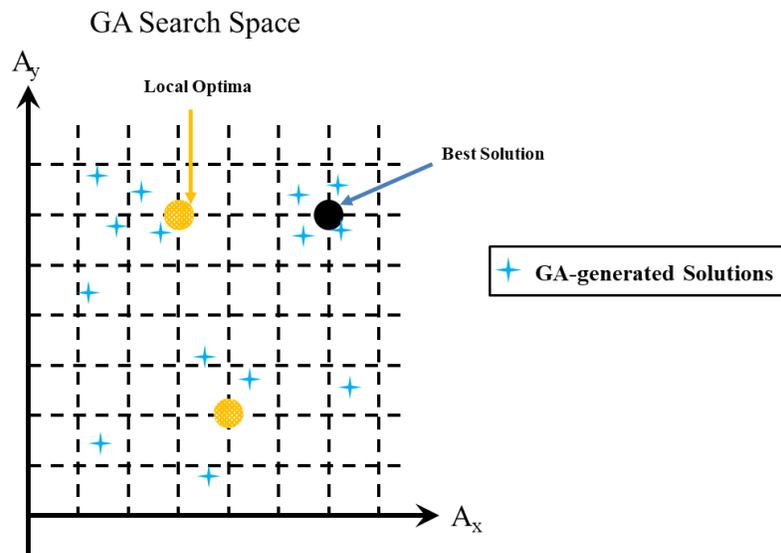


Figure 2.7: Illustration of search space in the case of two input genes

With each new input, the number of dimensions within the search space increases, and the amount of possible value combinations the GA could produce increases exponentially. The sheer size of the search space, when viewed in the context of the SLM, requires that there be restrictions on the amount of freedom the device can act upon. Another consequence of having a significantly large search space, is that the lack of specificity could negatively impact the GA's ability to reach the intended solution; or find an unnecessary one (called a *local-optimum*). A local optimum (displayed as yellow dots in Fig. 2.7) is a solution in a specific region of the search space that looks

like an optimal solution in respect to the defined goal; while – in actuality – there a better solution. Once parameters are set and defined, the GA sets about producing the population it needs to run its iterations. In the case of the wavefront correction experiment, the GA creates a randomized set of genes used to build the 512x512 phase array, which is sent to the SLM as a means of producing a phase modulation response.

### 2.4.3 *Best Individual Selection and fitness score*

Weighing the efficacy of the genes within each individual requires that they (the individuals) be given a value. This value is called a *fitness score*. The fitness score gives the GA an indication of which individuals (a set of genes) are best suited in meeting the fitness function parameter defined using the feedback signal, with the intent of directing the GA towards the intended solution. The fitness score can be outlined in several ways, such as  $f_i/\bar{f}$ , where  $f_i$  is the fitness score of a single Individual, while  $\bar{f}$  is the average fitness across all Individuals. Other methods of fitness evaluation include ranking the fitness of each individual <sup>[130,131]</sup>, or tournament selection <sup>[132]</sup>. The fitness function guides the GA towards the optimal solution. Fig. 2.8 illustrates an example of fitness scores evolution as function of the GA generation when the goal is to increase a laser maximum intensity measured by a camera. The maximum value of the fitness score could be 4096 counts here, which corresponds to the camera pixel depth. When the GA is close to the optimal solution, the fitness score reaches a plateau as there is no more improvement of its value.

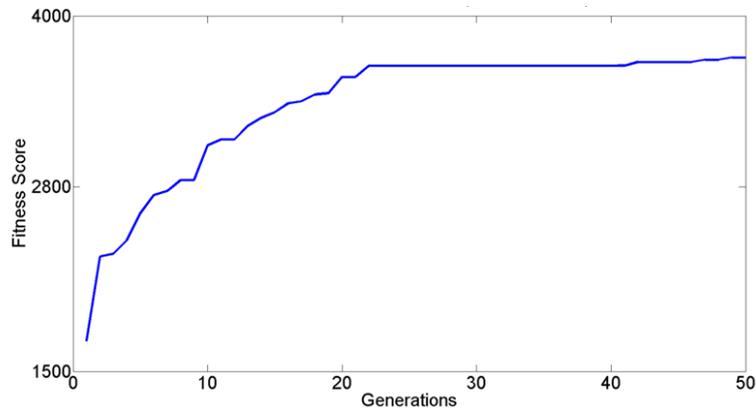


Figure 2.8: Example of the fitness scores evolution as a function of the GA generation

#### 2.4.4 Reproduction and Genetic Operators

Next comes the process of reproduction. The terms “parents” and “offspring” have been used up until this point for the sake of comprehension, but from here on, the terms “parent” and “offspring” will be defined as “individuals”. Each individual contains a 512x512 pixel phase array, and each generation incorporates a set number of individuals; the total being the population. After the best individuals are selected, they are then allowed to exchange values - within the phase map - with another individual, or provide some level of alteration to their existing values; making their way to the next generation. This process of making new individuals is facilitated by what are known as *genetic operators*. There are two types of genetic operators: *Crossover (Recombination)* and *Mutation* <sup>[133]</sup>; as depicted in Fig. 2.9.

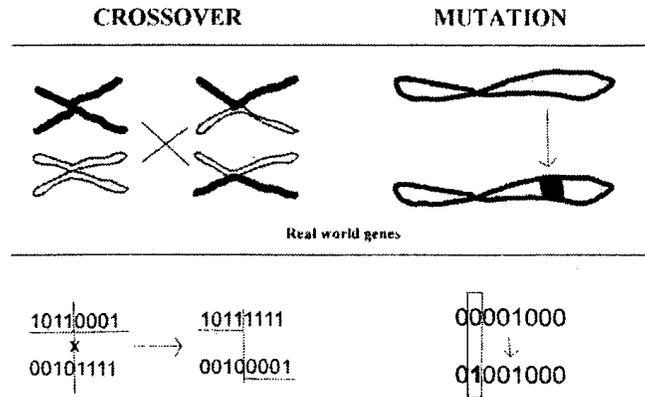


Figure 2.9: Crossover and Mutation methods <sup>[134]</sup>

**Crossover** takes two individuals and uses a combinational technique to interchange their phase values. Though what values are exchanged is random, there is a set order to which values are chosen. The premise behind the crossover technique is that a good solution can produce an even better solution <sup>[135]</sup>. In the case of the experiments performed in this chapter, random elements of each array are chosen interchanged with another individual, called uniform crossover. This ensures no bias between any two elements (bits) during reproduction.

**Mutation** takes one individual and alters the phase values within individual phase array, which then becomes a new individual for the next generation. Mutation operators can allow for uniform mutation, boundary mutation, or non-uniform mutation <sup>[136]</sup>. For the experiment, the amount of phase variation is set as an input parameter that can be adjusted for changing the GA convergence speed or success. A larger phase range for mutation translates into stronger phase modulation effects, allowing for greater exploratory opportunities. However, with increased exploration potential also comes the tradeoff of less precision. This lack of precision can potentially lock the GA into a suboptimal solution; known as exploitation. Balancing this ratio (Exploration vs Exploitation) is key in allowing for diversity and convergence.

#### 2.4.5 *GA Advantages and Disadvantages*

Advantages of a GA are quite numerous. It allows for the solving of a complex solution with no analytical form, it can optimize along continuous or discrete variables, it is capable of working in parallel computers, and it is generally well adapted to receive numerous input parameters <sup>[137]</sup>. As a tool, it is a reliable multi-conditional optimization technique. To serve as a juxtaposition, another optimization algorithm, known as simulated annealing <sup>[138]</sup>, works in a similar manner to that of the physical phenomena known as annealing. It takes one individual at a given “temperature” and adds a randomized state for the individual to explore; or not – if it chooses. The purpose of the simulated annealing technique is to optimize the individual to reach lowest temperature state possible; which would lead to the best solution. What the GA benefits intrinsically from is that its significantly larger population size enables it to search for a solution in fewer generations, as opposed to one individual making up the entirety of the population; as it is with simulated annealing. An example of this could be finding a needle in a very large haystack. Having multiple individuals involved shortens the time it takes to find the needle, as opposed to one. Essentially, simulated annealing is more time consuming than its GA counterpart, which is why it is often applied to more specific design problems <sup>[138]</sup>.

There are, however, limitations to the GA. Though capable of finding a solution, it may not be optimal to utilize a GA when needing to find an already established solution, as existing methods may already exist and are likely better optimized than the GA. Another issue of the GA, which was alluded to in the *GA initialization* section, is its ability to be trapped to a local optimum. Local optimums appear to be the desired solution the GA is searching for, and as a result these optimums can diminish a GA’s ability to search for other possible solutions; ones that may be even better. This is especially problematic if the GA is able to converge to its solution relatively quickly

<sup>[139]</sup>. There are methods of mitigating this type of convergence, and normally lies in a GA's ability to preserve diversity in the genes it produces. Mutation is typically what ensures diversity in the GA, and – preferably – what allows for there to be multiple solutions to a given function <sup>[140]</sup>.

## **2.5 Wavefront Correction using a learning algorithm**

The previous sections presented the GA and how it is used in the experimental arrangement shown in Fig. 2.1. To validate the GA, as far as its ability to correct the wavefront of an incoming aberrated beam, an aberrated medium (plastic see-through plate) was introduced on the beam to create additional higher order aberrations. The goal in doing so is to understand how well the GA can recover the initial wavefront by finding the inverse wavefront of the one given by the aberrated slide. Before the aberrated slide was introduced, an experiment was performed in which the SLM is turned on but has no phase applied (i.e. 0V); as a starting point of the correction. The GA would then compensate for these small initial aberrations (cf. Eqn. 2.2), serving as an indicator for how well the GA can correct an aberrated wavefront.

### *2.5.1 Compensation of the initial SLM distortion*

When the SLM is powered on, and zero volt sent to each pixel, the SLM lacks an ideal surface. As such, it is necessary to get a baseline of how much modulation takes place when the laser reflects off the SLM. With the beam directed along the interferometry arm, the phase modulation that is inherent in the SLM is measured and displayed in Fig. 2.10a. Linecuts (1D vector extractions of a 2D array) taken across the middle of the phase, in both the x and y-directions, also illustrates the phase not being constant across the surface of the SLM (Fig.2.10b).

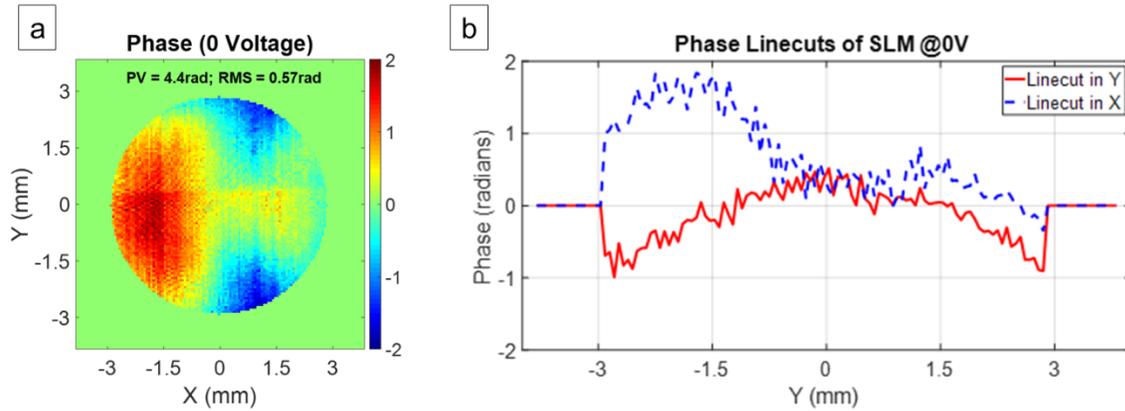


Figure 2.10: 2D phase of the SLM at 0V (a), Linecuts in the x and y-directions (b)

Taking the PV and RMS values of the phase, when the SLM is at 0V, gives a value of  $\sim 4.4$  and 0.57 radians; respectively. In order to determine the performance of the correction process used in this experiment, the Rayleigh Criteria defines the PV and RMS values needed to be achieved by the setup; see Eqn. 2.9. If the wavefront aberrations are reduced down to the level of (Eqn. 2.9), the system can be considered free of aberration. How close the system can come to these values would indicate its successful convergence.

$$PV(\varphi) < \frac{\pi}{2} \text{rad} = 1.6 \text{ rad} \quad \text{or} \quad RMS(\varphi) < \frac{\pi}{10} \text{rad} = 0.3 \text{ rad} \quad (2.9)$$

When focused by a lens (750 mm lens in Fig. 2.1), the wavefront aberrations transform the energy that is spread outside of the focal spot. When observing the spatial profile of the focused beam, also called the “focal spot” (see Fig. 2.11), the energy from the laser is dispersed unevenly around the central beam spot. This causes the peak intensity of the beam to be lower in the center.

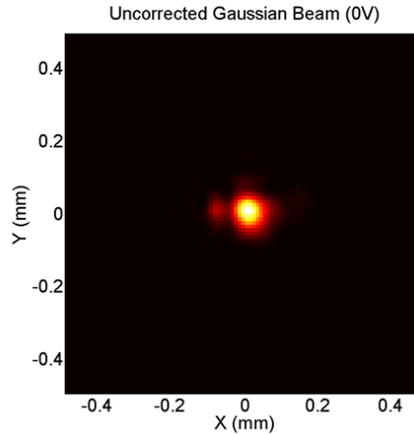


Figure 2.11: Spatial intensity profile of uncorrected Gaussian beam at 750 mm lens focus. Energy from the beam is dispersed, rather than concentrated at the central beam spot.

The profile of this beam is the result of the previously aforementioned optical phenomena, as well as other optical aberrations; such as: astigmatism or coma aberrations. In addition, the beam could not precisely be located in focus on the camera imaging plane. Addressing this issue requires the use of a phase method that can provide a modulation phase capable of minimizing aberrations; while also focusing the beam. When more energy is collected within the center of the beam spot, it produces a greater intensity. As such, in order to correct the wavefront of the beam the GA optimizes the wavefront by using a fitness method that takes feedback from the intensity measured by spatial profile camera; only to then maximize it.

Two correcting methods are provided to showcase how certain phase methods are more capable of wavefront correction than others. The first phase method is using what is known as the rectangular (Cartesian), polynomial phase method. The Cartesian phase method follows a typical polynomial mathematical form (Eqn. 2.10). The constant,  $A_n$ , is the phase value that is randomized and altered by the GA, and a mesh grid is used to map the phase method onto a 512x512 pixel array. As the array is 2D, two polynomial equations are utilized, with phase values assembled in both the  $x$  and  $y$  directions; such as  $A_{2x}$  and  $A_{2y}$ . With there being two directions of operation, this

method is also capable of some astigmatism correction, as one directional equation could provide a stronger focus term than the other.

$$\varphi_{Cartesian}(x) = A_{0x} + A_{1x}x + A_{2x}x^2 + \dots + A_{nx}x^n \quad (2.10)$$

The other phase method used is that of the Zernike polynomials, which encompass aberration formed on a circular pupil. This phase method should be better at correcting wavefront aberrations than the Cartesian phase method, as it can generate more complex phases. The different Zernike polynomials are described in Appendix A. By contrast, the strength of the Cartesian method lies in its ability to focus the beam; by way of the  $A_2$  term. This  $A_2$  term leads to a parabolic wavefront that mimics a focusing lens. Both methods can add a constant phase, as well as tilt, but these phase terms are disabled for the purposes of this experiment; as they do not allow for correction of the beam. For the Zernike method, these are the first three Zernike modes ( $Z_0$ ,  $Z_1$ ,  $Z_2$ ), and for the Cartesian method, this is the first two terms;  $A_0$  and  $A_1$ . Table 2.1 contains a list of GA parameters used during the correction process for both phase methods.

GA Parameter	Cartesian Phase Method	Zernike Phase Method
Best Individuals ( $N_{best}$ )	4	4
Mutated Individuals ( $N_{mutation}$ )	16	16
Crossover Individuals ( $N_{crossover}$ )	12	12
Number of terms (genes)	2	13 modes
Range of mutation (rad)	100	100

Table 2.1: GA wavefront correction parameters

When determining the amount of mutated individuals, the square of  $N_{best}$  is used ( $N_{mutated} = N_{best}^2 = 4^2 = 16$ ), while the number of crossover individuals is determined using the formula ( $N_{crossover} = N_{best} \times (N_{best} - 1)$ ). The best individuals are only chosen as a way for

their phase values to have both mutation and crossover operators applied to them for future individuals. This means that each generation only carries 28 individuals ( $N_{mutated} + N_{crossover}$ ). The best individuals from the previous generation do carry over, however, in the case that better individuals could not be found in the current generation. When mutation of the best individuals occurs, a technique known as *regularization* is employed. This technique takes the 4 mutated individuals, and then applies a logarithmic scaling factor to the phase value associated with those individuals; by orders of 10%, 1%, and 0.1%. This means each of the 4 mutated individuals has 3 additional individuals that are assigned to them, leading to 16 mutated individuals being produced. To give an example, if a mutated individual has a phase term value of 400 radians, its subsequent regularized individuals are 40, 4, and 0.4 radians. The benefit of this technique is that it allows the GA to essentially “roam” around a specific mutated individual, in order to see if better solutions can be found nearby.

As the GA iterates and creates new phase values to send to the SLM, the camera provides feedback to the LabVIEW system, enabling the GA to determine if a particular phase pattern (individual) is effective in maximizing the intensity of the laser beam. The results from correcting the Gaussian beam, with both the Cartesian and Zernike methods, can be seen in Fig. 2.12.

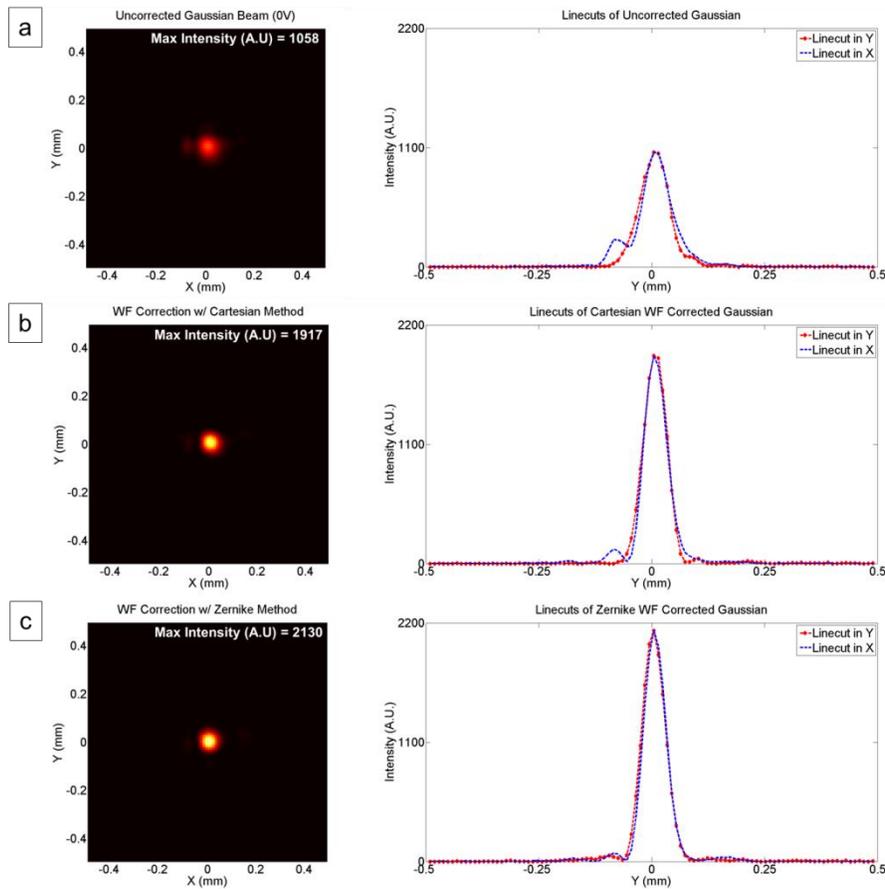


Figure 2.12: Correction results of initial SLM distortion. (a) Uncorrected Gaussian beam spatial profile and intensity distribution, (b) Spatial profile and intensity distribution of the beam after Cartesian phase method correction, (c) Spatial profile and intensity distribution of the beam after Zernike phase method correction

As seen in Fig.2.12, both phase methods demonstrate an ability to maximize the beam intensity, with up to a two-fold increase in the beam maximum intensity. Reductions in the artifacts present near the central focal spot is also noticeable. What is most noticeable is that the Zernike method demonstrated a higher degree of wavefront correction, as it results in the greatest increase in the beams intensity (2130 A.U. counts). Observation of the fitness scores, per generation (Fig. 2.13), further illustrates the Zernike methods advantage in aberration correction. While the Cartesian method is able to converge towards the target solution faster than the Zernike method (reaching a fitness score of 1.85x), ultimately the Zernike method is able to reach a more optimal solution (reaching a final fitness score of 1.89x). This is due to the restrictions placed on the

possible phase shapes for the Cartesian method, which allows the complex phase shapes of the Zernike method to find a more optimal solution. Section 2.6 goes into further details on the topic of exploration (search space), convergence, and phase efficacy. At certain times during the iteration process, the GA is unable to find a phase pattern that produces a fitness score that is better than the previous score from the last generation. When this occurs, the fitness score is held constant until a better score can be found.

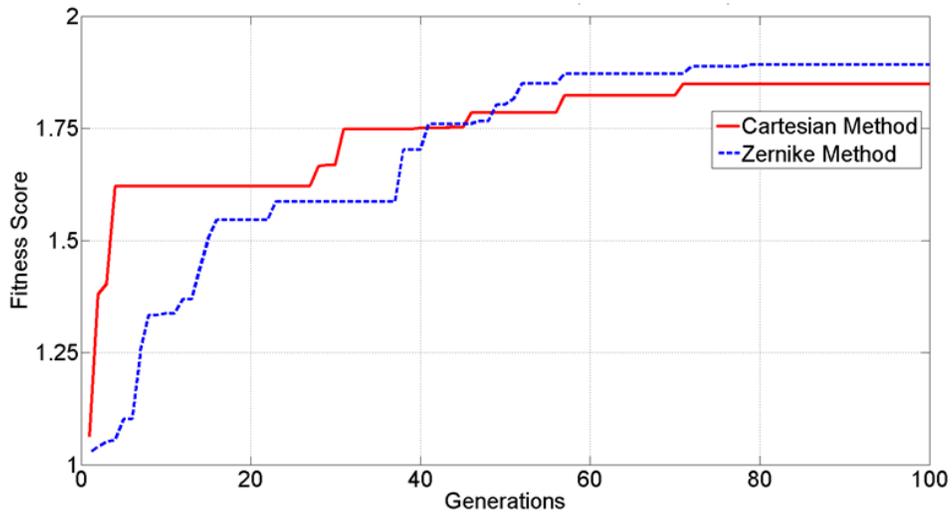


Figure 2.13: Fitness scores of the best individual as function of generations for each phase method. The fitness score used here are normalized to the reference fitness score for a beam with initial (0V) phase

Looking at the corrected phase (Fig. 2.14), both the Cartesian and Zernike methods produce wavefronts with a lower PV and RMS value, indicating that the wavefront aberrations are reduced. The Cartesian method is able to reduce the PV and RMS values down to 2.60 and 0.48 radians (Fig. 2.14b), while the Zernike method is able to achieve a reduction in the PV and RMS down to 2.23 and 0.42 radians (Fig. 2.14c). The shading of the color map associated with the unwrapped phase (for the corrected beam) reinforces how a lower PV & RMS value is akin to a near-constant phase distribution. As a PV of 1.6 radians and an RMS of 0.3 radians is desirable, the values that

are achieved through this correction technique show worthwhile signs as to the effectiveness of the learning algorithm.

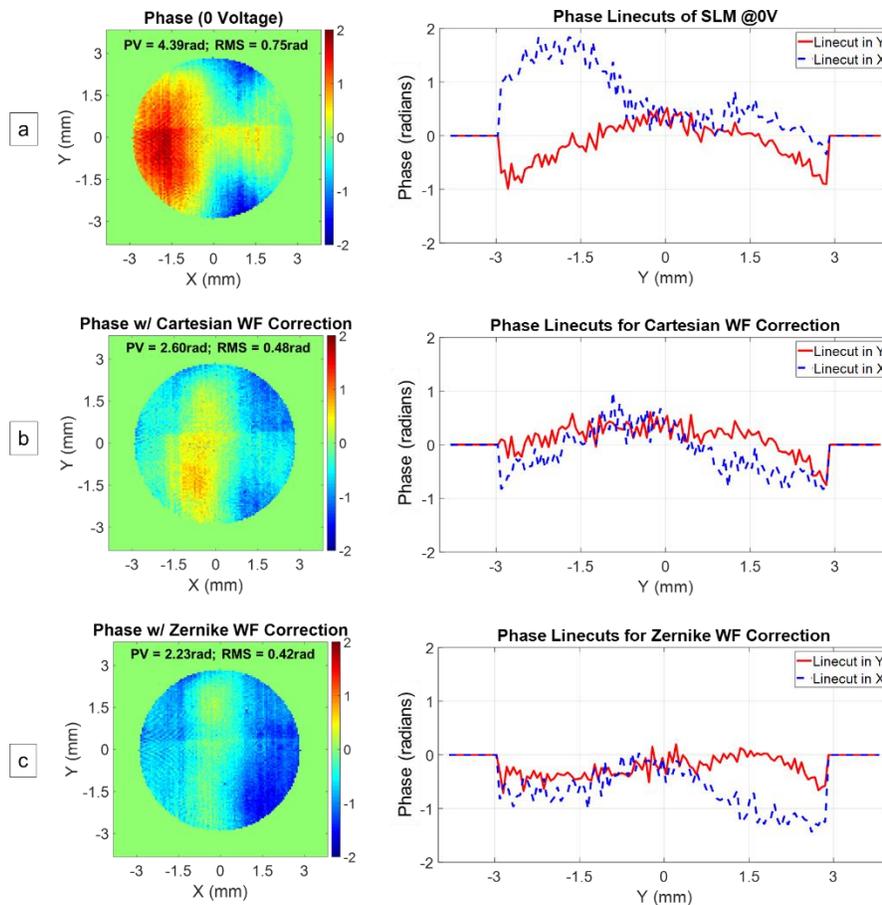


Figure 2.14: Spatial phase and linecuts before and after wavefront correction. (a) Uncorrected Gaussian beam phase and linecuts. (b) Cartesian phase and associated linecuts. (c) Zernike phase and associated linecuts

The phase linecuts of the Cartesian and Zernike phase methods showcase how both methods are capable in diminishing out-of-focus and astigmatic effects, optimizing the wavefront. In addition, as the Zernike corrected focus (Fig. 2.12.c) is clearly better than the Cartesian corrected focus (Fig. 2.12.b), it means that the phase measurement probably contains aberrative terms that are not really in the beam and are created by slight imperfection in the interferometer alignment. This further justifies the use of a learning algorithm for optimizing the signal of interest (camera signal), rather than measurement of wavefront that could contain additional aberrations.

While the GA can reduce the amount of aberration on the uncorrected beam wavefront, not meeting the Rayleigh criteria indicates that there may be more aberrations types that the GA could not correct. By analyzing the phase of the uncorrected beam and determining the Zernike mode (coefficients) amplitudes (over 36 modes –  $Z_{36}$ ), there are a number of higher-order aberrations still present within its wavefront; see Fig. 2.15. The modes related to  $Z_3$ ,  $Z_6$ ,  $Z_8$ ,  $Z_9$ ,  $Z_{10}$ ,  $Z_{19}$ ,  $Z_{21}$ , and  $Z_{35}$  display the greatest aberrative effect on the wavefront; see Fig. 2.15a. Observation of the correction results for Fig.2.15b (Cartesian correction) and 2.15c (Zernike correction) shows a reduction in the magnitudes of the  $Z_3$ ,  $Z_6$ , though with increases in the amplitude in other modes. One possibility for this behavior is the GA had not found the “best” solution which reduces all possible Zernike aberrations in which the Zernike phase method can operate upon. Additional Zernike phase terms, as well as GA iteration time, could see further improvements in the beam signal. It should be noted however, that with increased inputs from the Zernike phase method, the more time it would take for the GA to find the optimal solution.

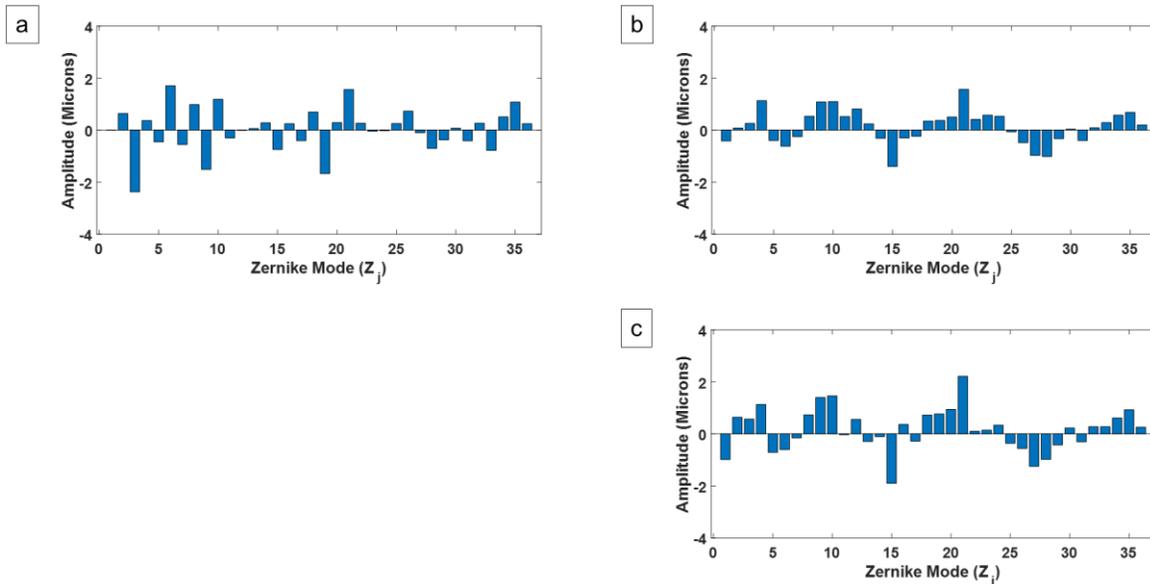


Figure 2.15: Zernike coefficients before and after initial SLM distortion correction. (a) Zernike coefficients of the aberrated beam. (b) Zernike coefficients after Cartesian phase method correction, and (c) the Zernike coefficients after Zernike phase method correction.

Transitioning to the 512x512 SLM phase array created when achieving these correction results, the pattern produced by the Cartesian method (Fig. 2.16a) shows some degree of astigmatic phase distribution, as well as a parabolic shape that is indicative of defocusing.

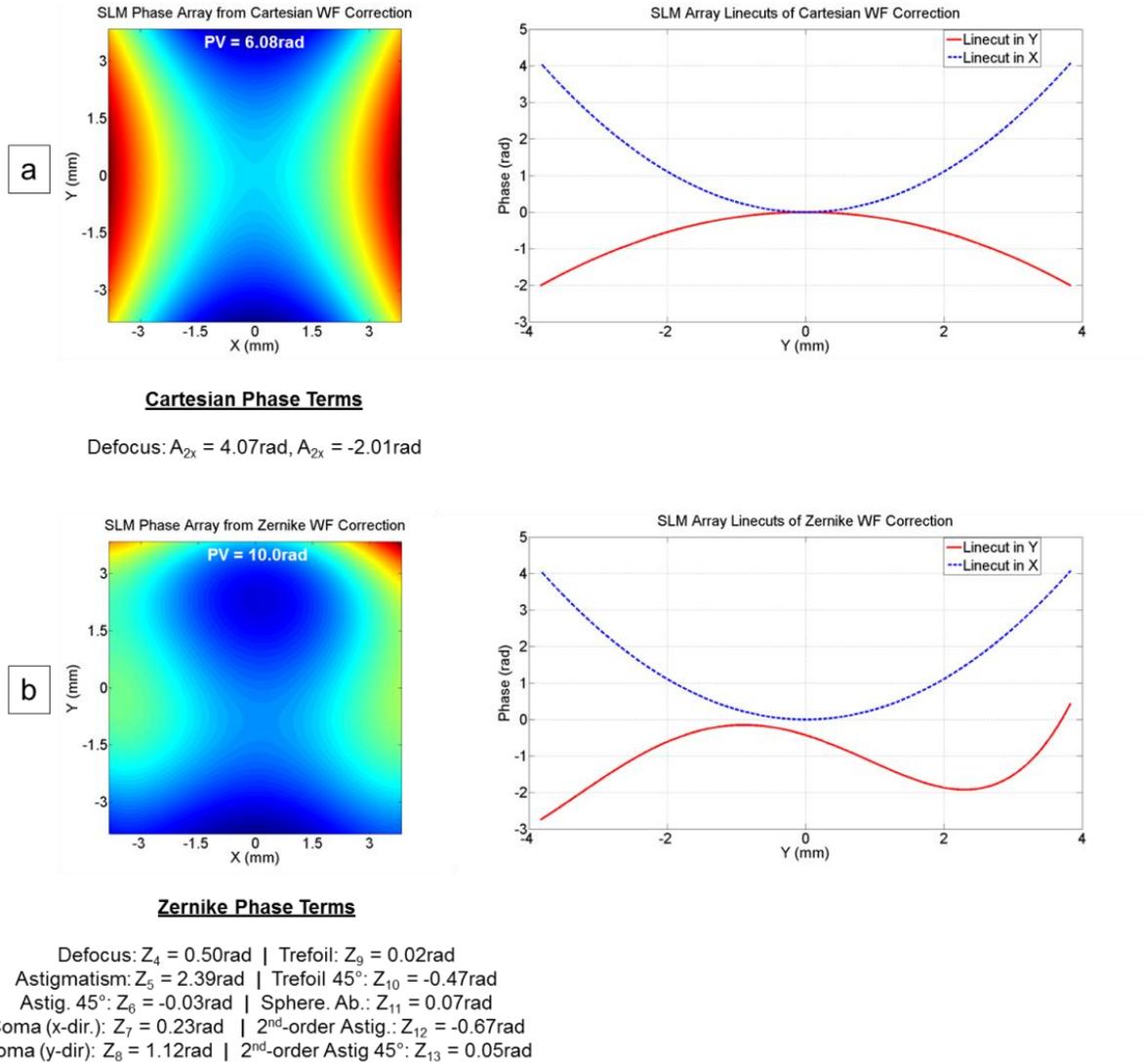


Figure 2.16: SLM phase arrays after correction. (a) SLM phase array and linecuts after GA optimization using the Cartesian method. (b) SLM phase array and linecuts after GA optimization using the Zernike method.

The Zernike method (Fig. 2.16b), however, displays a more unique phase array due to the various Zernike phase types that were utilized during correction. There is still a detectable defocusing phase pattern seen within the linecuts. What is more notable is the lack of similarity

between the two methods and their defocus terms. The Zernike method provides a stronger phase term with respect to 1<sup>st</sup>-order astigmatism ( $Z_5$ ), rather than pushing a strong defocus phase (unlike the Cartesian method), along with the correction of coma aberrations ( $Z_8$ ).

The phase (in radians) of the corrective SLM arrays is in good agreement with the uncorrected (0V) beam phase measured by the interferometer. The surface linecuts of the SLM array mirror that of the linecuts taken of uncorrected beam, showing that they are attempting to correct the beam with an inverse phase. Discrepancies between the PV and RMS of the SLM arrays used for correction, and the PV and RMS values of each correction methods respective phase, is primarily due to the circular mask used only in the phase measured by the interferometer. After demonstrating the ability for this learning algorithm setup to correct a standard Gaussian, attention is now turned to correcting for higher-order aberrations.

### 2.5.2 *Correction of Higher-Order Aberrations*

To better test the experimental setup and its control algorithm, and how well it can correct for aberrations, an aberrative object is introduced into the system. The object is an 8cm x 6cm see-through, plastic cover and is mounted in between the beam splitter and the SLM. Placing it in this position ensures that only the modulated wavefront is affected by the object, and ensures that both the fringe data and phase data is represented accurately along the interference arm of the setup. Fig.2.17. shows the spatial profile of the beam in the focal plane of the 750 mm lens, as a result of passing through the plastic cover. The maximum intensity of this aberrated beam is measured to be 86 counts, while the background noise is ~18 counts, illustrating the severity of the aberration when compared to the maximum count of 1058 for the uncorrected Gaussian beam. The diffuse shape of the beam also leads to a significantly large beam width, with the FWHM being 0.3 mm.

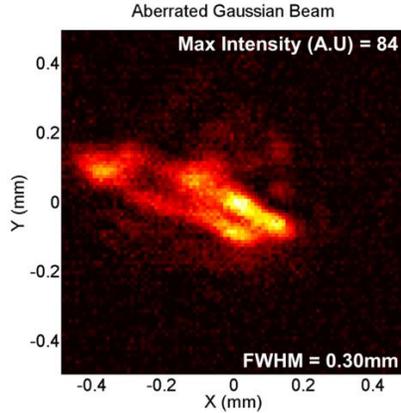


Figure 2.17: Gaussian beam spatial profile at focus with additional aberration from coverslip.

Both the Cartesian method and the Zernike method are tested for correction of these aberrations. As the goal is to produce a beam that is closer to an ideal Gaussian, reusing the same fitness method of maximizing the intensity seen in the CCD camera is still desirable. This concentrates the energy of the beam towards a central beam spot and allows for the mitigation of these higher-order aberrations. Table 2.2 shows the parameters used for the process of correcting the beam. Results of this correction can be seen in Fig. 2.18.

GA Parameter	Cartesian Phase Method	Zernike Phase Method
CCD Exposure	0.26ms	0.26ms
Number of Generations	50	50
Best Individuals ( $N_{best}$ )	4	4
Mutated Individuals ( $N_{mutation}$ )	16	16
Crossover Individuals ( $N_{crossover}$ )	12	12
Number of terms (modes)	2	13
Range of mutation (rad)	200	200

Table 2.2: GA starting parameters for higher-order aberration correction

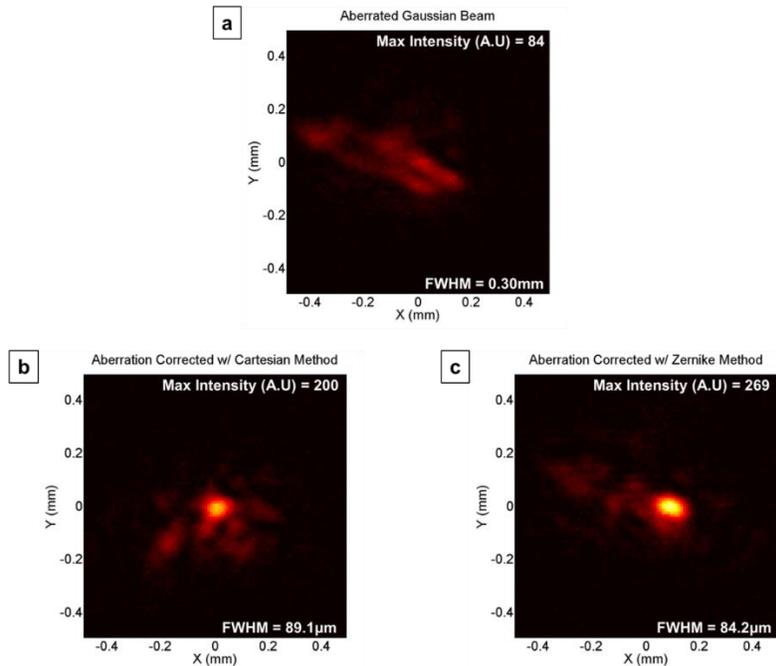


Figure 2.18: Spatial profile before and after high-order aberration correction. (a) The uncorrected beam. (b) Aberration correction using the cartesian method. (c) Aberration correction using the Zernike method.

Both the Cartesian (Fig. 2.18b) and Zernike (Fig. 2.18c) phase methods are able to correct for aberrations, producing a noticeably more intense beam. There is still a discernable level of aberration involved, which could be attributed to certain phase aberrations not being addressed by either method. The spatial profile for the Cartesian phase method shows a number of aberrative affects surrounding the beam spot. Given that is mostly capable of defocusing, there is a lack of corrective ability in removing these effects. The Zernike method, by contrast, is able to reduce the size of the existing aberration; while still improving the beams intensity. Using the Cartesian method for aberration correction, the original FWHM (Full Width Half Maximum) of the aberrated beam is reduced by a magnitude of 3.3x; registering as 89  $\mu\text{m}$ . Further reduction was demonstrated with the Zernike method, providing a reduction of 3.5x; the FWHM being 84.2  $\mu\text{m}$ . The impacts of correction are best seen when analyzing the phase before and after correction, as shown in Figure 2.19; along with its associated linecuts.

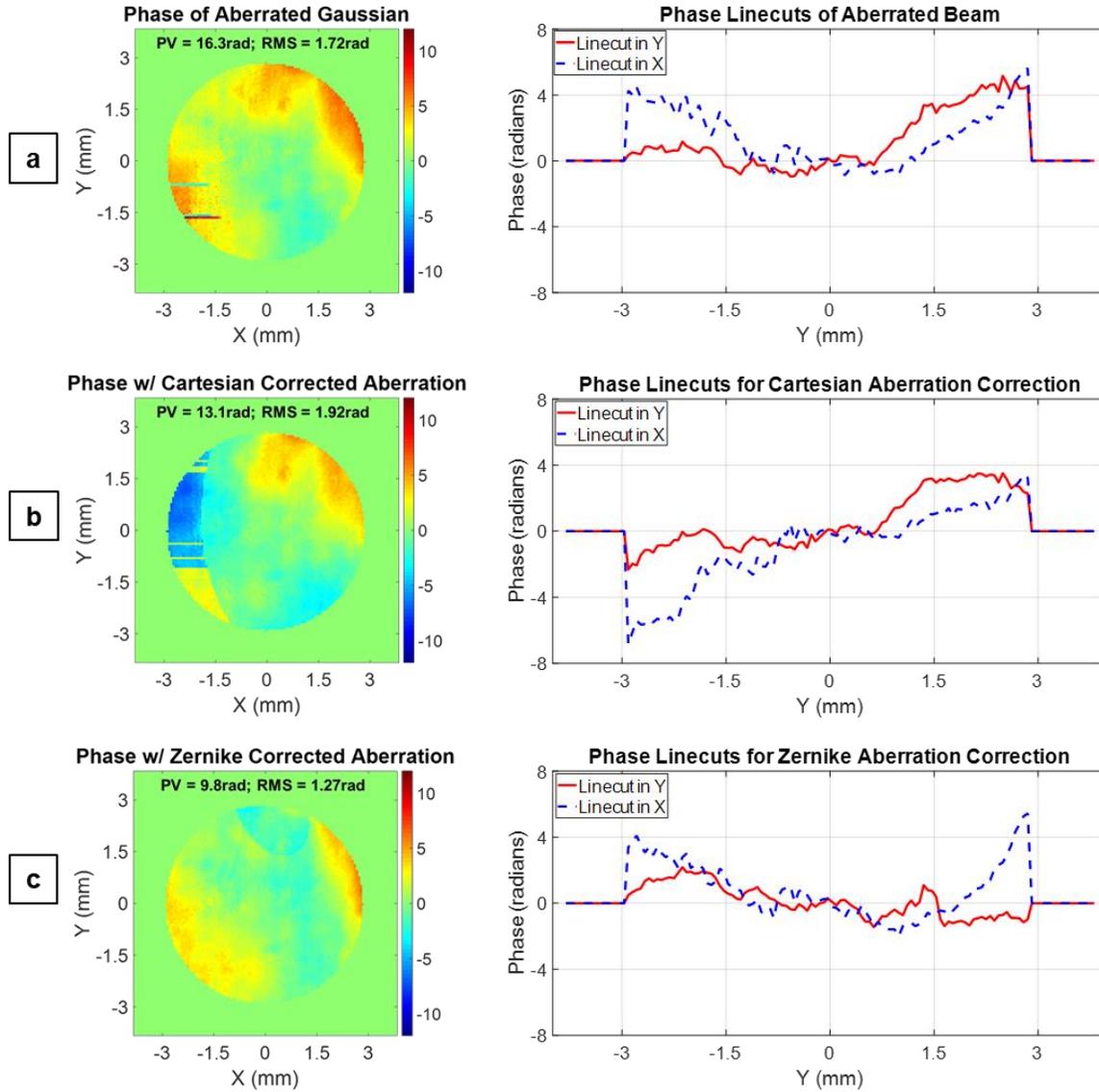


Figure 2.19: Spatial phase and linecuts before and after aberration correction. (a) Aberrated Gaussian beam (b) Cartesian correction (c) Zernike correction

The PV value of the aberrated Gaussian beam (Fig. 2.19a) is measured to be 16.3 radians, and an RMS of 1.72 radians. After performing GA optimization using both the Cartesian and Zernike methods, the PV value of the phase, after Cartesian correction (Fig. 2.19b), is reduced to 13.1 radians, with an RMS of 1.92 radians. This correction is rather small when viewed against the Zernike correction (Fig. 2.19c), which is able to reduce the PV value of the phase to 9.8 radians, and the RMS to 1.27 radians. Observation of the phase terms and their phase values showcases

how the Zernike method seeks to correct a significant number of higher-order aberrations that go unaccounted for by the Cartesian method. The Zernike coefficients, before and after corrections, can be seen in Fig. 2.20. Reductions in the amplitude of several Zernike modes, within the operating range of the Zernike phase method, can be seen between the aberrated beam (Fig. 2.20a) and the Zernike corrected beam (Fig. 2.20c). Higher-order aberrations, such as modes  $Z_{17}$ ,  $Z_{21}$ ,  $Z_{31}$ ,  $Z_{35}$ , are left uncorrected. An outlier in this case is mode  $Z_7$ , which corresponds to a coma (in x). No definitive conclusions can be ascertained regarding the inability for the Zernike phase method to correct this mode.

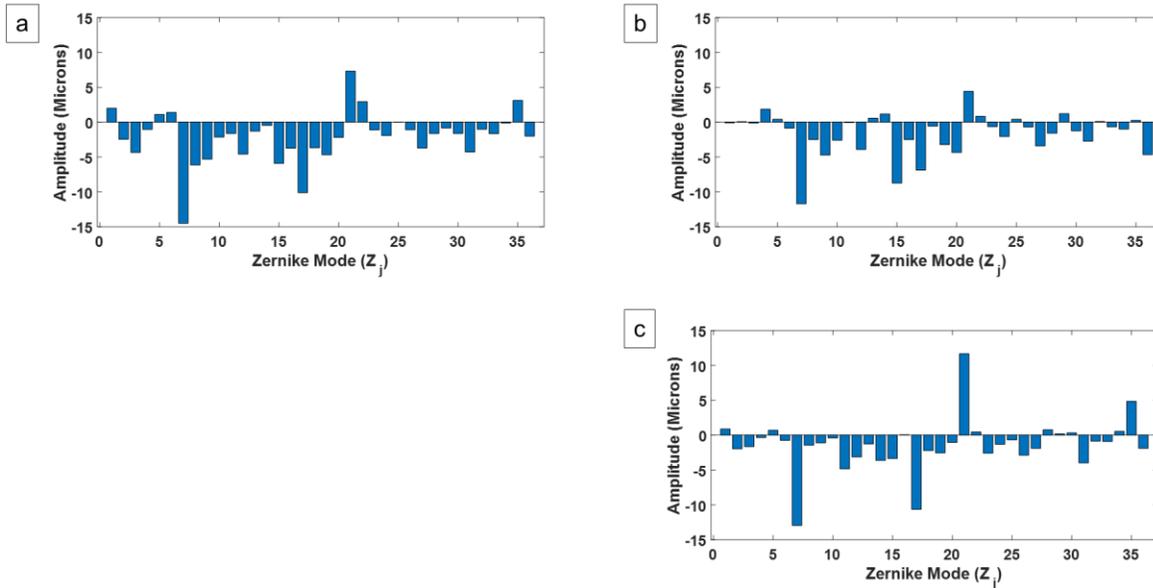


Figure 2.20: Zernike coefficients before and after higher-order aberration correction. (a) Aberrated beam. (b) After Cartesian phase method correction, and (c) after Zernike phase method correction

Fig. 2.21 includes images of the SLM phase arrays for both corrections, as well their respective linecuts. As the Cartesian method is only able to perform astigmatic and defocusing correction, the parabolic shape represents its limitation to correct higher-order aberrations. The Zernike phase array and linecut showcases a more complex structure, illustrating the ability for this method to correct for more aberrative effects; and to a stronger degree.

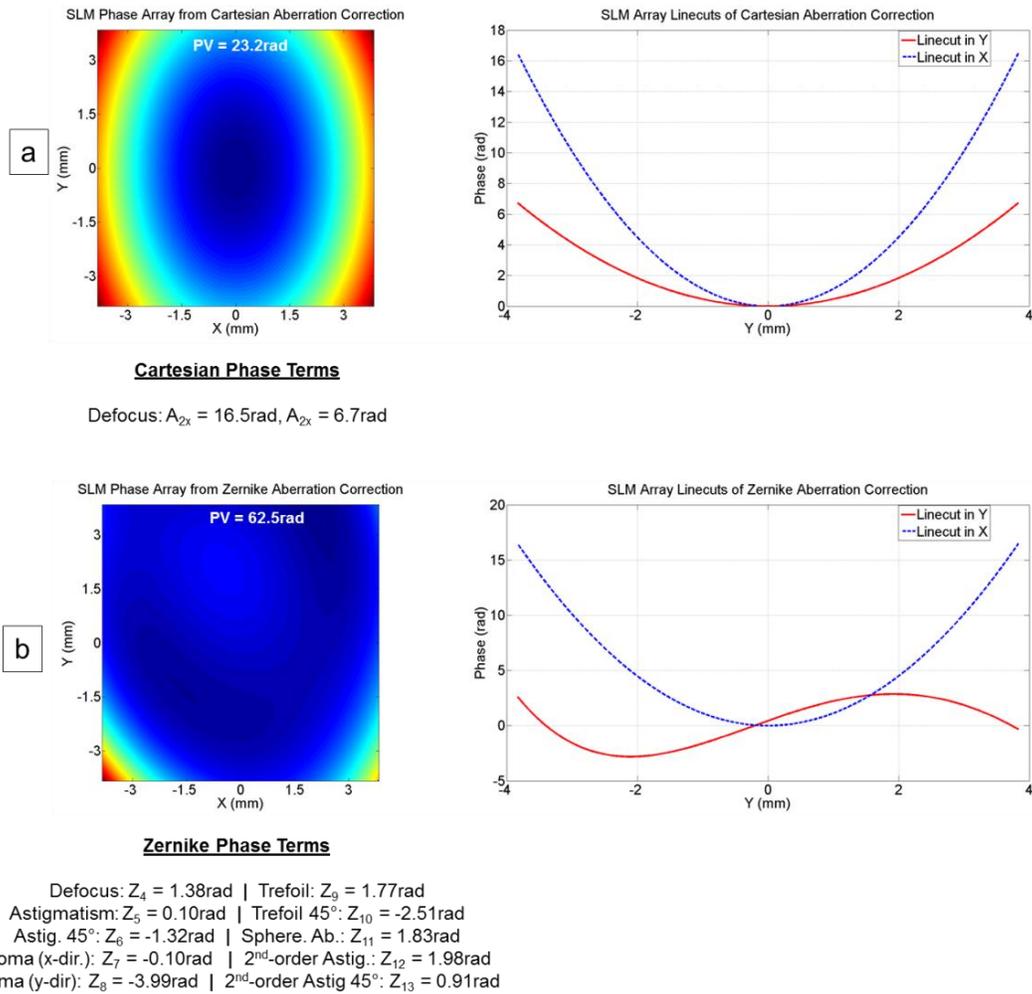


Figure 2.21: 512x512 SLM phase array and linecuts for both corrective methods. (a) Cartesian phase array and linecut after aberration correction. (b) The Zernike phase array and linecuts after aberration correction

Viewing the best fitness scores as a function of the GA generation (Fig. 2.22), both methods converge a near equal pace, but the Zernike can find the best solution for aberration correction. There are number of attributable factors involved in the stagnation of both methods when finding a better optimization solution. This includes: the number of aberrative orders, the amount of phase each aberration imparts, the phase method utilized, how the fitness & optimization methods are chosen, and the starting GA parameters. It is important to discuss the process of finding which parameters are best suited in helping the GA find the optimal solution, and what factors (such as the search space), play into how quickly a particular method converges towards a target solution.

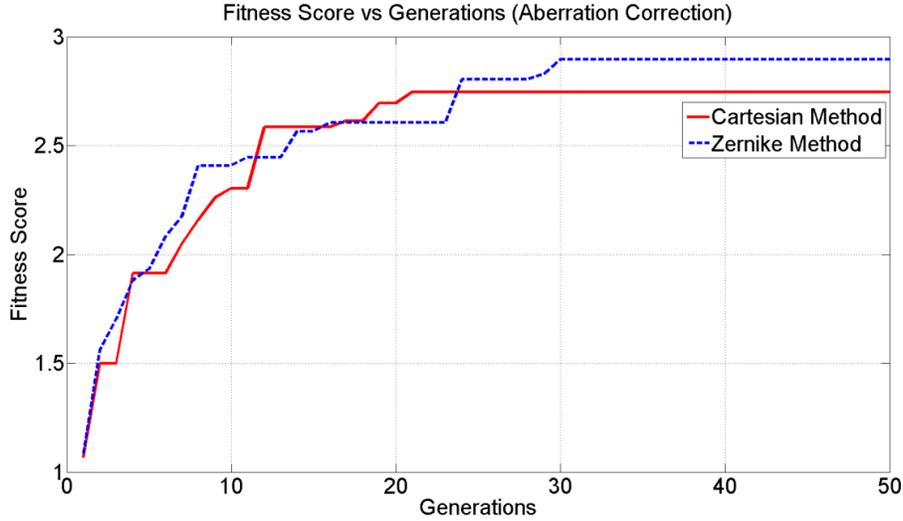


Figure 2.22: Fitness score evolution during correction of the aberrated Gaussian beam

## 2.6 GA Parameter Impacts on the Search Space

With experiments into wavefront correction concluded, it is necessary to analyze the efficacy of the parameters utilized in order to optimize the wavefront of the Gaussian beam. As mentioned before, the possible number of phase terms the GA could apply to the SLM is equivalent to the array (or LCD) screen size of said SLM; which is 512x512 pixels. For each pixel, the SLM only accepts unsigned 16-bit values, which ranges from 0 to 65535. Theoretically, an infinite number of phase terms can be sent to the SLM, but this is unrealistic. When viewed in terms of the Cartesian method, and the 2 defocus phase terms which were applied in section 2.5, the total number of solutions this method could formulate is  $65535^2$ ; or  $4.28 \times 10^9$ . If a phase term were to be applied to each individual pixel within the SLM, this would make the maximum number of solutions the GA could generate as  $65535^{(512 \times 512)}$ , which can be functionally seen as an infinite number of solutions. Due the vastness in how many solutions the GA can potentially produce, it is important to test the effectiveness of each method used in the previous section (Cartesian and

Zernike phase methods) to gauge how necessary additional phase terms are for a particular optimization task.

To provide a better understanding into how having too much freedom in the amount of phase terms can be detrimental, a third phase method, known as the Checkerboard method, has been used in conjunction with the Cartesian method. This checkerboard method is specifically designed to match the largest possible number of phase terms the SLM can take (512x512). As this is too computationally taxing, the full 512x512 array is broken up into user-defined “checkers”, with each checker being an  $M \times M$  pixel array. A phase term is associated with each checker, giving all the pixels contained within said checker the same value as the phase term. Using bi-linear interpolation, this  $M \times M$  array is expanded to a 128x128 pixel array, with this process repeating again in order to convert the – now – 128x128 array into 512x512 pixels. This interpolation is necessary to not only ensure the phase produced by the checkerboard method fills the entire SLM array, but also allows for smoother transitions between the values that were initially filled for each pixel; preventing large discontinuities. For simplicity and to reduce long computational times, an 8x8 checkerboard is used for analysis, meaning the 512x512 array is split into 64 checkers. Figure. 2.23 illustrates the design of the 8x8 checkerboard.

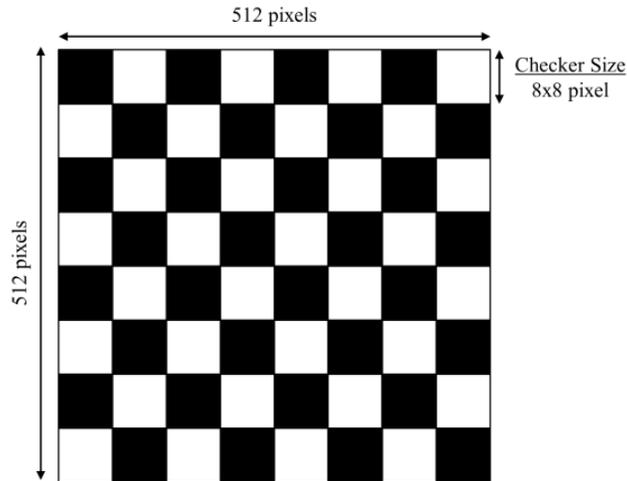


Figure 2.23: 8x8 Checkerboard Array created using the checkerboard phase method

As stated earlier, three distinct phase creation methods have been used for analysis. These phase methods were implemented in the following ways: The Cartesian method using only the defocus terms ( $A_{2x}$  and  $A_{2y}$ ), the Cartesian method using 4 additional phase terms ( $A_{3x/y} \dots A_{6x/y}$ ), the Zernike method using Zernike modes  $Z_3 \dots Z_8$ , and a combination of both the Cartesian (using  $A_{2x/y} \dots A_{6x/y}$ ) and the checkerboard method using an 8x8 checkerboard. The phase values for this 8x8 array are not preselected, but are instead randomized by the GA. Two fitness methods are chosen for optimization, with the first fitness method seeking to maximize the laser intensity, and the second method attempting to minimize the beam width of the laser. Scoring for beam width minimization uses the observed  $1/e^2$  beam width of the laser, though it is interpreted as pixels due to the step-wise nature of the SLM screen. Results of this testing can be seen in Table 2.3. Every method is tested across a variety of conditions, including:

- (1) the maximum intensity score it can reach,
- (2) the smallest beam width (in pixels) it can achieve, and
- (3) its convergence speed, which determines the number of generations it takes to reach 75% and 95% of the maximum intensity produced by each method.

Fitness Method	Phase Method	Corrected $I_{\max}$ (A.U)	Corrected Beam Width (pixels)	Convergence Speed (generations)	
				75%	90%
Max Int.	Cart. ( $A_{2x}, A_{2y}$ )	1876	X = 6, Y = 6	5	7
Max Int.	Cart. ( $A_{2x/y} \dots A_{6x/y}$ )	1538	X = 6, Y = 6	47	59
Max Int.	Cart. ( $A_{2x/y} \dots A_{6x/y}$ ), Checkerboard (8x8)	858	X = 6, Y = 6	12	59
Max Int.	Zernike ( $Z_3 \dots Z_8$ )	1939	X = 5, Y = 5	16	49
Min. Beam	Cart. ( $A_{2x}, A_{2y}$ )	1720	X = 6, Y = 6	5	10
Min. Beam	Cart. ( $A_{2x/y} \dots A_{6x/y}$ )	1815	X = 6, Y = 6	7	22
Min. Beam	Cart. ( $A_{2x/y} \dots A_{6x/y}$ ), Checkerboard (8x8)	110	X = 10, Y = 10	5	42
Min. Beam	Zernike ( $Z_3 \dots Z_8$ )	1120	X = 6, Y = 6	4	18

Table 2.3: GA convergence speeds and phase method effectiveness under various conditions

The Cartesian method using, only the defocus terms, showcases the fastest convergence speed of all the methods utilized in this analysis. However, this speed does not lend itself to achieving the best solution, which is demonstrated by the Zernike method achieving the highest intensity across both fitness methods; including a smaller beam width when maximizing the beam intensity. The number of terms implemented by the Zernike method requires a significantly greater of number of generations (at least 49) in order to reach 90% of the maximum beam intensity value. When viewed computationally, the seven Zernike modes being implemented ( $Z_3 \dots Z_8$ ) translated to a total combination of  $65535^7$ ; or  $5.13 \times 10^{33}$ .

This drawback of a significantly larger search space is taken to an extreme when the combination of the Cartesian and checkerboard method is observed. The significantly lower maximum intensity it can achieve (when attempting to maximize beam intensity) can be attributed to the significantly larger search space it needs to explore;  $65535^{6+(8 \times 8)}$ . The inability for this method to find a solution is further emphasized when viewing its performance when trying to minimize the beam size; where the corrected intensity is lower than the initial intensity of the beam. This combinational phase method is presumed to have encountered a local optimum where the beam size was perceived by the GA to be the “minimum.”

An outlier in this analysis is the Zernike method when minimizing the beam width. Theoretically, it should have been the most successful in not only minimizing the size of the beam, but in also achieving the highest maximum intensity. It is highly likely that the GA found a local optimum in which it believed that it had achieved the best solution, which is proven by the relatively small beam width of the Gaussian. While its width was small, this did not mean the beam was in focus, which may lie in the fact that the Zernike method did not apply a strong enough defocus term.

Another notable takeaway from this optimization test is the interplay between the number of phase terms used for a particular phase method, and how they behave under differing fitness methods. Focusing in on both Cartesian methods (of varying phase terms), implementing only the defocus term allows for both a quick convergence speed, as well as a greater beam intensity. Given that the Gaussian beam is already being focused, the corrective phase needed was largely restricted to the defocusing and astigmatism. The Cartesian method with additional phase terms ( $A_{3x/y} \dots A_{6x/y}$ ) introduces additional complexity into the GA, making it harder for it to find the best solution needed to maximize the beam intensity. This situation is reversed when the fitness method centers on minimizing the beam size, as those additional phase terms produce a beam with a greater intensity than its defocus phase-only counterpart, while also minimizing the beam width. As the intensity of the beam no longer factors into optimization, the GA benefits from having additional phase terms to work with, which allows for higher-order aberrations (which would previously have no importance in optimization) to be corrected for.

## **2.7 Discussion**

It has been demonstrated that 1) the experimental setup, which incorporates an SLM and the GA, is able to be used for both spatial profile observation and fringe analysis, 2) that the GA

is able to correct the wavefront of an uncorrected Gaussian beam, and 3) through GA optimization, it is possible to correct for higher-order aberrations. Results of wavefront correction has led to marked improvement of the intensity distribution of the corrected Gaussian beam, as well as a reduction in both the PV and RMS values of the spatial phase; an effect seen significantly with Zernike correction. An examination into the necessary number of phase methods, the efficacy of each phase method, as well as their convergence speeds, was also performed. From these observations, further insight was gained by understanding the dynamics between the amount of phase terms utilized (for each respective phase method), and how greater exploratory freedom also comes at the cost of finding the best solution needed to satisfy a particular fitness method.

It is important to recognize that multiple factors, such as the phase methods implemented and fitness method chosen, can have a large impact on the ability for the GA to converge to an optimal solution. For future studies, it is worthy to consider the use of a phase method that allows for the randomization of phase values within the SLM phase array. An optimization phase method that is wholly randomized would – theoretically – be able to find the best solution possible (compared to other phase methods). However, such a method would also require significantly increased computational time; given the exponential size of the search space. If such a method were to be applied, a novel technique would be necessary in order to minimize computation time, while also maximizing the possible search space for this particular phase method.

## CHAPTER 3 CHARACTERIZATION OF 3D FOCAL INTENSITY DISTRIBUTIONS

### 3.1 Introduction

The previous chapter presented wavefront correction experiments, with the Spatial Light Modulator (SLM) providing a phase modulation term to an incoming beam source. The required phase modulation was determined by a learning algorithm - more specifically a Genetic Algorithm (GA) - that used the signal from a Charged Coupling Device (CCD) camera as the feedback for optimization. This could be seen as a “classical” wavefront correction scheme, as it allows the optimization of the intensity distribution in the transverse plane in which the CCD camera is positioned; but does not constrain the light outside of the plane. One of the goals of this dissertation is to propose the optimization of a laser **focal volume**, instead of just the focal plane (focal spot), and explore if a learning algorithm is able to generate a focal volume that possesses a larger axial extend than that of a reference Gaussian beam. This means that the learning algorithm needs to have several input signals (CCD spatial profiles) at different points along the propagation axis.

As it is not feasible to manually move the CCD over a large range in a fast and reliable way, the SLM could provide a defocusing term that shifts the beam in the longitudinal z-direction, while also preserving the structure of the beam. A method of calibration needs to be devised in order to ensure the SLM sends the appropriate defocus phase term that adjusts the beam to the proper plane. This chapter is intended to discuss the fundamental concept known as the **3D focal intensity distribution**, develop a method to measure this distribution using a camera and the SLM, and obtain a proper calibration of the SLM defocus term for use in the multi-plane GA, an extension of the single-plane GA that was presented in chapter 2. The goal of this multi-plane GA is to shape the focal volume, and this work is presented in chapter 4. The effectiveness of the SLM

calibration is demonstrated in this chapter by running experimental comparisons between two longitudinal scanning methods, the manual translation of the CCD camera, and a longitudinal scan using the SLM to defocus the laser.

### 3.2 The 3D focal intensity distribution

The *3D focal intensity distribution* describes the characteristics and shape of the beam as it is traveling along the beam propagation direction ( $z$ ) while it is focused, and can be obtained by recording transverse spatial profiles  $I(x,y)$  at different longitudinal  $z$  positions. When visualizing the beam across its transverse (or spatial) profile, this requires observation of the beam in both the  $x$  and  $y$ -dimensions; or  $xy$ . This would be akin to viewing the point of a pencil. Alternatively, if the pencil were to be rotated, so that the entire length of the pencil was visible, this would be indicative of the longitudinal profile; which always requires the  $z$ -dimension. The viewer can then either choose to view the pencil from the top ( $xz$ ) or from the sides ( $yz$ ). The 3D focal intensity distribution would be akin to observing the full dimensions of the pencil.

Fig. 3.1 illustrates different ways of displaying the 3D focal intensity distribution in the case of a perfect Gaussian beam with similar characteristics than the one used in this dissertation ( $w_0 = 38$  microns,  $Z_R = 9.3$  mm). Fig. 3.1a shows nine 2D transverse intensity profiles  $I(x,y)$  as they would be recorded by a CCD camera while being translated in  $z$ . These 2D images constitutes a 3D stack, and only nine are displayed, but to obtain a finer axial resolution stack, images are saved at least every 5 mm (81 images for 400 mm range). Taking a 2D slice at  $y=0$  extracts from the stack the **2D longitudinal intensity profile  $I(x,z)$**  as displayed in Fig. 3.1b (normalized to 1) and Fig. 3.1c (not normalized, if the CCD exposure is kept constant while acquiring the stack).

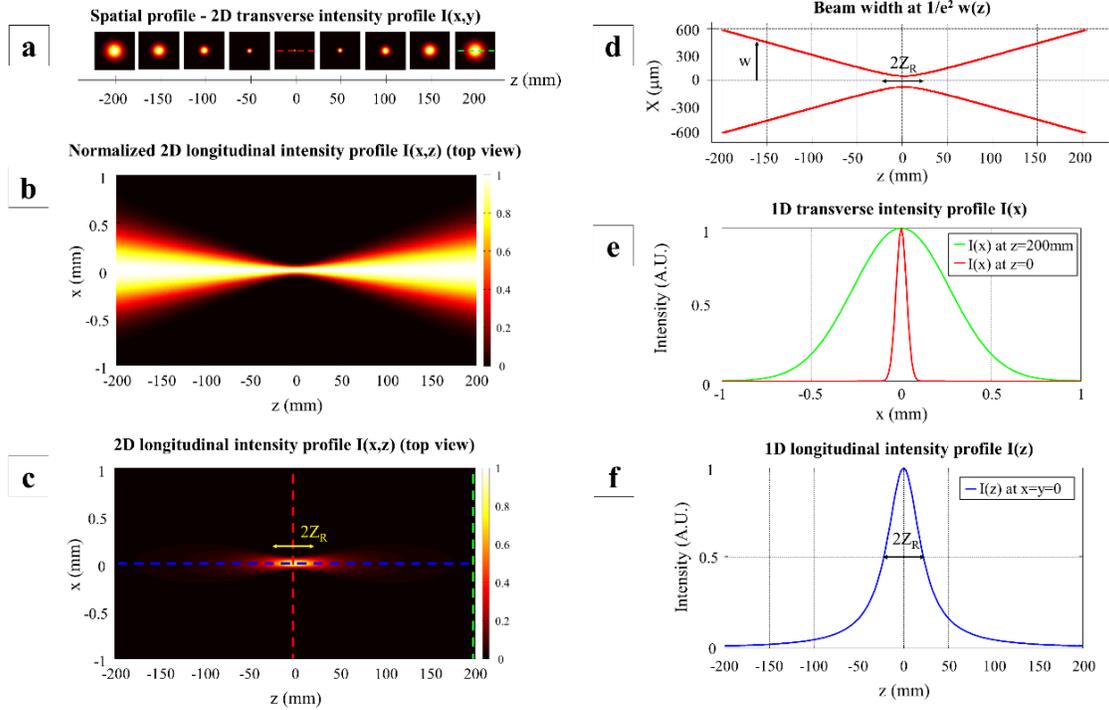


Figure 3.1: Displaying the 3D focal intensity distribution. (a) 2D spatial profiles are recorded to build a 3D stack. (b) Normalized 2D longitudinal intensity profile  $I(x,z)$  at  $y=0$ . (c) 2D longitudinal intensity profile  $I(x,z)$  at  $y=0$ . (d) Beam width  $w$  at  $1/e^2$  as function of  $z$ . (e) 1D transverse intensity profiles  $I(x)$  at focus ( $z=0$ ) and  $z=200$  mm. (f) 1D longitudinal intensity profile  $I(z)$  at  $x=y=0$ . Dashed lines in (a) and (c) correspond to the linecuts in (e) and (f).

To provide a visual support of the evolution of the beam size and divergence with  $z$ , the beam width radius at  $1/e^2$  in  $x$  and  $y$  can be plotted as shown in Fig. 3.1d. For a Gaussian beam, this beam width follows the dependence presented in Eqn. 1.15 seen in chapter 1. The **depth of focus** ( $2 Z_R$ ) is displayed in the figure and is a measure of the axial extend of the focal volume, region where the intensity is above 50% of the maximum peak intensity at the focal plane ( $z=0$ ). Fig. 3.1e displays the one-dimensional (1D) intensity distribution  $I(x)$  of the Gaussian beam, and can be found by taking a horizontal linecut across the center of the 2D spatial profile, or a linecut at given  $z$  positions in the 2D longitudinal profile. Dashed lines with the matching colors of the 1D profiles of Fig. 3.1e are used in Fig. 3.1a and Fig. 3.1c to indicate where the linecuts were

taken. Fig. 3.1f works in a similar manner, but is instead a linecut of the 2D longitudinal profile from Fig. 3.1c, a result of observing the beam as it travels along the longitudinal axis.

It is usual in optics to have only information about the spatial profile and phase in one transverse plane, and to have the 3D axial intensity distribution simulated or calculated. Unique in our case is the fact that the 3D intensity distributions of the beam profiles are obtained **experimentally**. The transverse (spatial) profiles are captured by the camera and they are overlaid to construct the 3D  $I(x,y,z)$  distribution. This allows each transverse profile to be linked to a respective propagation distance along the beam axis and creates a 3D stack of the beam which contains all the information to understand how the beam distribution evolves over the propagation axis  $z$ .

### **3.3 Experimental setup**

This experiment utilizes the same experimental setup arrangement as shown in chapter 2, but the interferometer and phase analysis arm are not used, so the beam is sent directly to the GA arm of the setup, and the reference arm is blocked (cf. Fig. 3.2). The spatial profile CCD camera is mounted on a translation stage that offers movement in the lateral direction, and the translation stage is affixed to a rail track that allows the camera to be moved along propagation direction  $z$ .

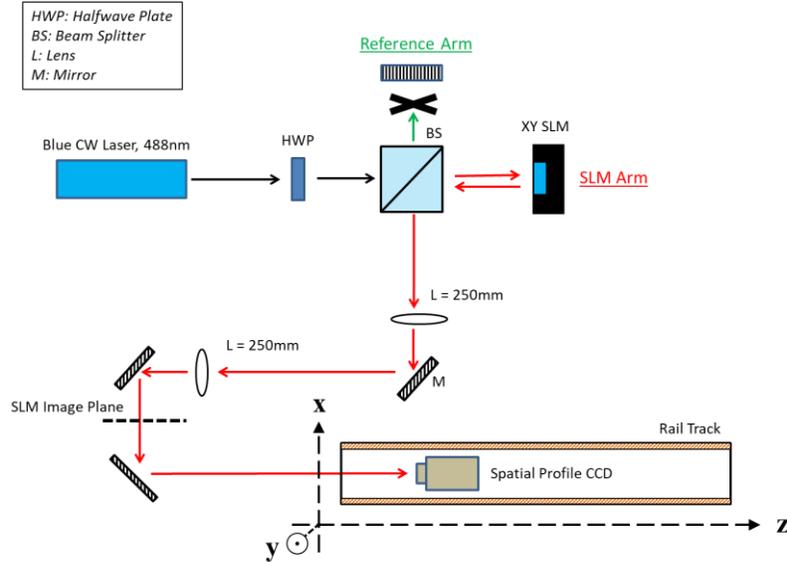


Figure 3.2: Experimental setup for characterizing the laser 3D focal intensity distribution

There is no focusing of the beam onto the spatial profile CCD camera, so the 750 mm lens from previous chapter experiment is removed. Instead, the SLM imparts a defocus term onto the beam, in addition to creating the distinct beam profiles seen in this experiment. For the purposes of this experiment, two distinct beam profiles are generated by the SLM, and their 3D focal intensity distributions are scanned using two different methods. This first method involves using the rail to manually reposition the spatial profile CCD camera at set intervals along a predetermined propagation range, while the second method is the SLM providing the defocus term necessary to simulate manual repositioning of the CCD camera at similar intervals.

### 3.4 Calibration of the SLM defocus term

#### 3.4.1 3D focal intensity distribution measured by CCD manual translation

The beam used for the calibration of the defocus term is a corrected Gaussian beam, with correction being handled by the SLM using the defocusing terms of the Cartesian method ( $A_{2x} = 140.5$  rad and  $A_{2y} = 135.5$  rad). The difference between  $A_{2x}$  and  $A_{2y}$  provides compensation of the

initial astigmatic shape on the SLM surface; which was measured and corrected in the experiment described in chapter 2. A distance of 1000 mm is chosen as our range of manual translation, with the beam focusing at 740 mm away from the SLM image plane (cf. Fig. 3.2). To use the same convention seen in Eqn. 1.15 for the beam size ( $w$ ), the focal plane at 740 mm is assigned a value of  $z=0$  mm, with the camera moving back and forth from this  $z$ -position by  $\pm 500$  mm.

Transverse profiles are recorded as the spatial camera is **translated manually** on a rail; by increments of 10 mm. The resulting 3D focal intensity distribution, shown as Maximum Intensity Projections (MIP) in the XZ and YZ planes, are shown in Fig. 3.3. Instead of displaying 2D slices at one given  $y$  value (similar to Fig. 3.1.b), a MIP<sub>xz</sub> takes the highest value pixel among all the  $y$  values, making it easier to visualize in a 2D representation any structure contained in the 3D stack.

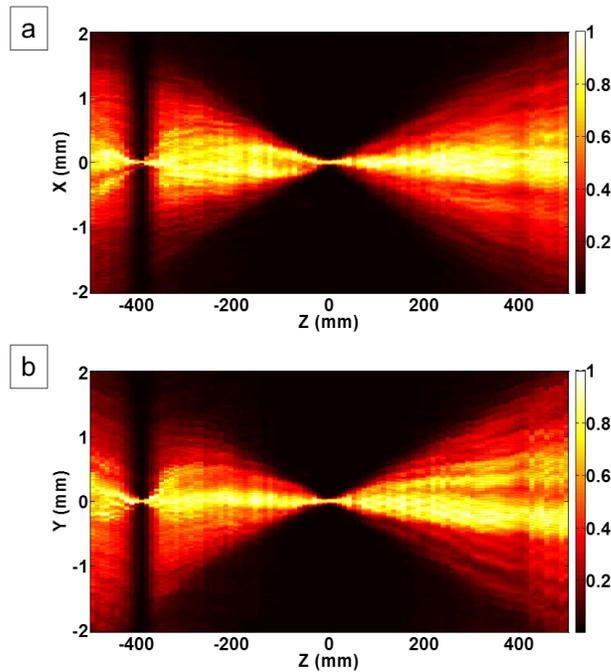


Figure 3.3: Normalized MIP of the Gaussian beam over  $z$ . (a) MIP<sub>xz</sub> manual scan. (b) MIP<sub>yz</sub> manual scan.

By recording the beam size values  $w_x$  and  $w_y$  obtained for each camera image, as a function of the propagation distance ( $z$ ), an experimental plot of  $w(z)$  is generated, shown as crosses in Fig. 3.4. These values can be fitted against the theoretical beam size of a perfect Gaussian beam (Eqn.

1.15). The blue dashed line in Fig.3.4 corresponds to the best fit to the theoretical beam size, which is obtained for a waist value of  $w_0 = 38\mu\text{m}$ .

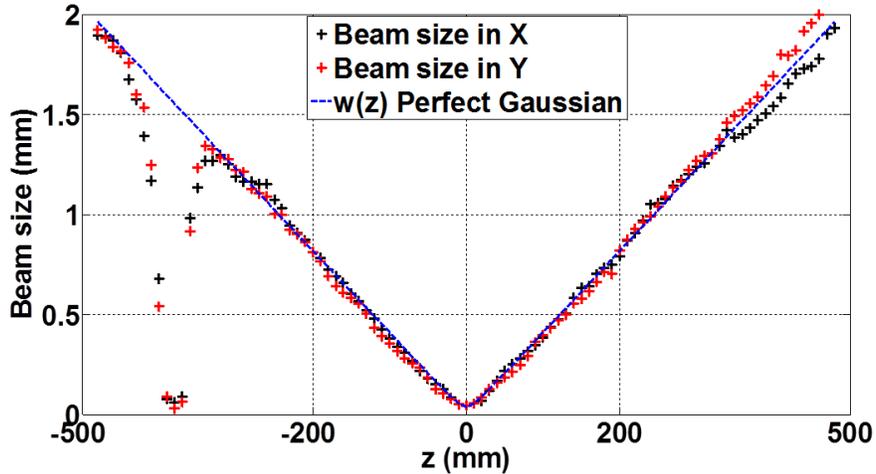


Figure 3.4: Gaussian beam size as function of  $z$ , extracted from the MIP shown in Fig. 3.3.

There is good agreement between the experimental and theoretical beam size evolution with  $z$ . Beam size values that are situated at the furthest  $z$ -values of the measurement range show noticeable deviations from the theoretical beam size evolution, but this is largely due to the beam profile being so dispersed that it makes calculating the beam size more prone to error. A secondary beam focus, at  $-380$  mm, is also apparent in Fig. 3.4. The intensity of this focus is 2.4% of the main focus (at  $z=0$ ). This secondary focus is likely the result of a front reflection of the laser beam on the SLM window whose antireflection coating is not optimized for 488 nm. As was mentioned in chapter 2, this could potentially be an issue for the GA to converge if, for example, a fitness method requiring the minimization of the beam size is the optimization criteria; as the secondary focus may have a smaller beam size.

### 3.4.2 Generation of 3D focal intensity distribution using SLM defocus term

For the purposes of the multiplane GA experiments presented in chapter 4, the goal of sending a phase with the defocus term is to allow the spatial profile CCD camera to see a beam transverse profile at different longitudinal values. In effect, this serves as a way of simulating the camera being moved through the beam path, allowing for observation of the beam structure and how it evolves over the propagation distance ( $z$ ). To successfully have the SLM assign the proper defocus term, it must first be calibrated so that it sends the correct value. A comparison with the previous section 3D axial intensity distribution, created by manually translating the spatial profile camera, provides validation of the calibration procedure. As the name suggests, the defocus phase term brings the image out of focus, causing blurring effects for an image, or beam size change for a laser beam. The defocus phase is highlighted by its quadratic dependence with beam size. Fig. 3.5 is an example of a 2D defocus term, as well as a plot of its shape along one direction.

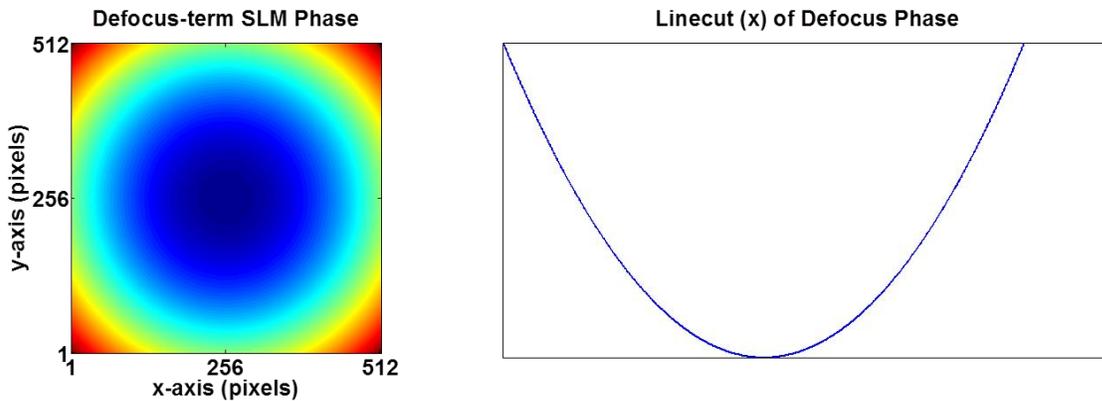


Figure 3.5: 2D defocus phase term sent to the SLM (left) and 1D linecut along x direction (right)

Creating the defocus term is fairly simple. Using some of the existing phase methods outlined before, such as the Cartesian and Zernike methods, the defocus term is created using certain Cartesian polynomial terms (such as the  $A_2$  terms) or Zernike modes (Zernike mode  $Z_4$ ).

For the SLM defocus calibration, the Cartesian method is used. For reference, the defocus terms for the Cartesian method can be seen in Eqn. 3.1.

$$\varphi_{Defocus}(x, y) = A_{2x}x^2 + A_{2y}y^2 \quad (3.1)$$

The question then becomes, “What defocus phase amount is necessary in order to move the beam focus to a distance  $z$ ? Finding this value requires that we measure how the beam size at  $(1/e^2)$  is impacted when moving the spatial camera at – and away – from the beam focus manually, and matching the change in beam size when the SLM defocus term is used to perform the beam size change. The next step is to use the SLM to defocus the beam using  $A_{2x}$  and  $A_{2y}$  terms, and record the corresponding spatial profiles and beam size values for each defocus  $A_2$  value. In this regard, the  $A_{2x}$  and  $A_{2y}$  defocus terms are set to increase in steps of 2 radians; from 0 radians to 400 radians. Both  $A_{2x}$  and  $A_{2y}$  work by applying a defocus term in one direction, with  $A_{2x}$  being in the x-direction, and  $A_{2y}$  being in the y-direction. As the SLM window surface shape is slightly astigmatic, it is necessary to calibrate for both terms separately by finding their respective beam sizes. By altering these terms, the beam width changes as a result of the beam focus being shifted. Fig. 3.6a. shows a plot of the beam width at  $1/e^2$  as a function of  $A_{2x}$  and  $A_{2y}$ , with a cropped version (Fig. 3.6b) excluding the secondary focus or values of  $A_2$  when the beam size is bigger than the CCD size, which result in the constant beam size value at 2 mm.

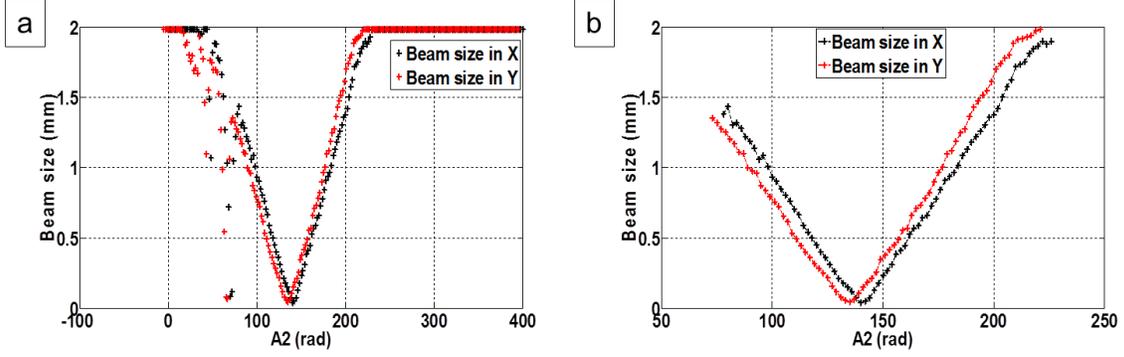


Figure 3.6: Beam size at  $1/e^2$  as a function of the  $A_2$  defocus term applied by the SLM (a). Beam size at  $1/e^2$  for the range of  $A_2$  values used for the defocus calibration

The next step is to associate one position  $z$  for each value of defocus  $A_2$ . Given that the beam waist at the focus ( $w_0$ ) is known, and the beam size is given by Fig. 3.6b, Eqn. 1.15 can be inverted in order to connect both  $A_2$  and  $z$ ; resulting in Eqn. 3.2.

$$z = z_R \sqrt{\left(\frac{w_0}{w}\right)^2 - 1} \quad (3.2)$$

From here, the defocus terms (as a function of  $z$ ) are plotted, see Fig. 3.7. The focal length (740 mm) is added in the  $z$ -axis so real values of distance are obtained for the calibration. Based on the apparent linearity of the  $A_2$  values, a linear fit is laid over them, from which the SLM defocus term calibration results is found; see Eqn. 3.3 and Eqn. 3.4.

$$A_{2x} = 175z + 10 \quad (3.3)$$

$$A_{2y} = 175z + 5 \quad (3.4)$$

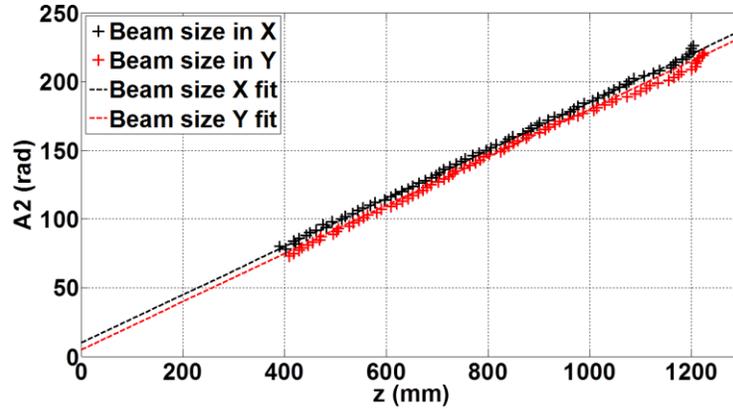


Figure 3.7: Defocus term  $A_2$  as a function of  $z$ , with linear fit depicted as dashed lines

The calibration terms (Eqn. 3.3 and 3.4) are then added in the control software for the SLM scanning process, which means the SLM should be able to create a 3D focal intensity distribution comparable to the one obtained when moving the CCD manually with this calibration. Fig. 3.8 shows the MIP results of the axial intensity scan when using the SLM, and not moving the CCD. These profiles are similar to that of the manual scanning method max projections; see Fig. 3.3.

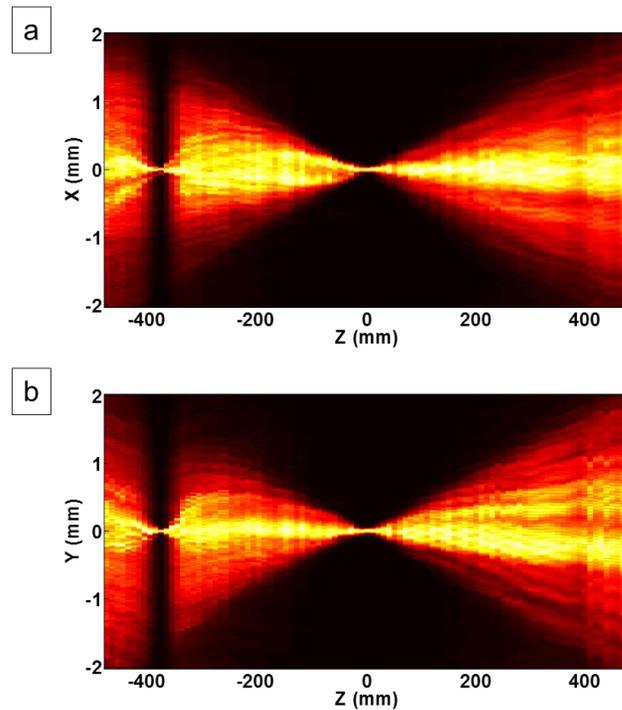


Figure 3.8: Normalized MIP generated using the SLM calibrated for defocus. (a) MIP<sub>xz</sub>, and (b) MIP<sub>yz</sub>.

Fig. 3.9 provides a comparison between the beam size values of both the SLM scanning method, and the manual scanning method. There is strong correlation between the beam size values for both methods. This is especially noticeable at both the main focus ( $z=0\text{mm}$ ) and the secondary focus ( $z=-380\text{mm}$ ). With confirmation that the calibration equation is working, by properly assigning the correct defocus phase value, experimental tests can now be conducted to see if this new calibration method is applicable for other types of beam distributions.

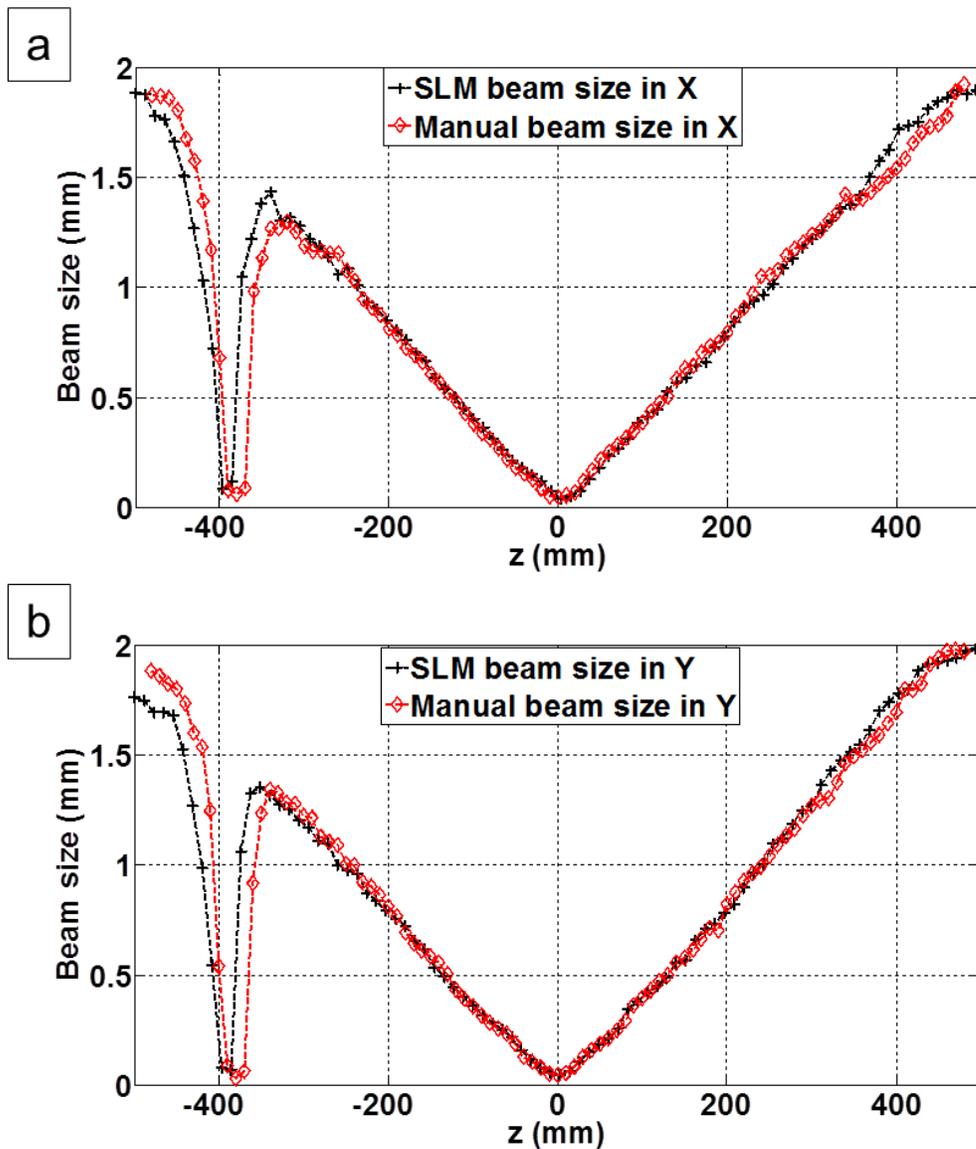


Figure 3.9: Beam size evolution comparison between the SLM and manual scans. (a) along x (b) along y

### 3.5 3D focal intensity distribution scanning comparisons

The SLM is programmed to generate the 3D focal distribution for two distinct beam profiles: a Gaussian beam and a beam with purposely introduced astigmatism (astigmatic beam). The Gaussian beam serves as a reference for calibration, as its behavior is well known. The astigmatic beam allows us to assess the efficacy of the calibrated defocusing term, as the aberrative effect becomes more pronounced the further the image scan is from the focal plane, and structures appearing in the spatial profiles should appear in both methods. All scans are performed over a distance of 400 mm. For the manual scans, 2D spatial profiles are captured at 10mm intervals which creates 41 images total. When performing the axial scans with the SLM, the calibration equations (3.3) and (3.4) are used in order to apply the appropriate defocusing term. This defocus term is created within a 512x512 pixel map similar to other phase creation methods. The additional phase maps created, in order to produce the necessary beam profiles, are then added onto this defocus map. Scanning in this manner is automated, allowing the camera to remain stationary while the SLM simulates the beams diffraction at the cameras observation plane. A benefit to this method being automated, is that it more images can be captured – in this case, 81 – in shorter time span and within the same scanning range.

#### 3.5.1 Gaussian beam 3D focal intensity scan comparisons

Fig. 3.10 displays the MIP taken along the xz-direction, and the yz-direction for both axial scanning methods. The MIP are all normalized to 1 to be able to compare more easily the fine structures in the beam. There are strong similarities between the max projections produced by both the manual and SLM scanning methods. As there is not enough visually distinction between the two scanning methods at the focal plane, the furthestmost wings of the Gaussian beam are used for comparison purpose.

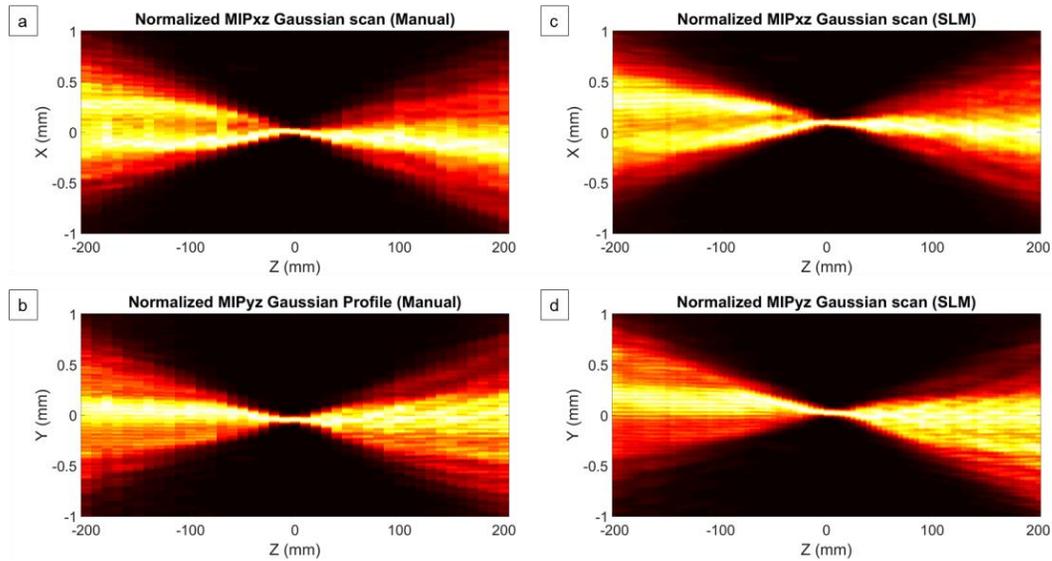


Figure 3.10: Manual and SLM Gaussian beam scan comparisons. Panels include MIP in the xz (left) and yz (right) directions for both the manual (a,b) and SLM (c,d) scans

The Gaussian beam transverse profile images captured at both -200 mm (first) and +200 mm (last) are showcased in Fig. 3.11, which includes both scanning methods. Visual observation of the images reveals relatively negligible differences between the two methods. Any pronounced differences are likely caused by small fluctuations in the background noise.

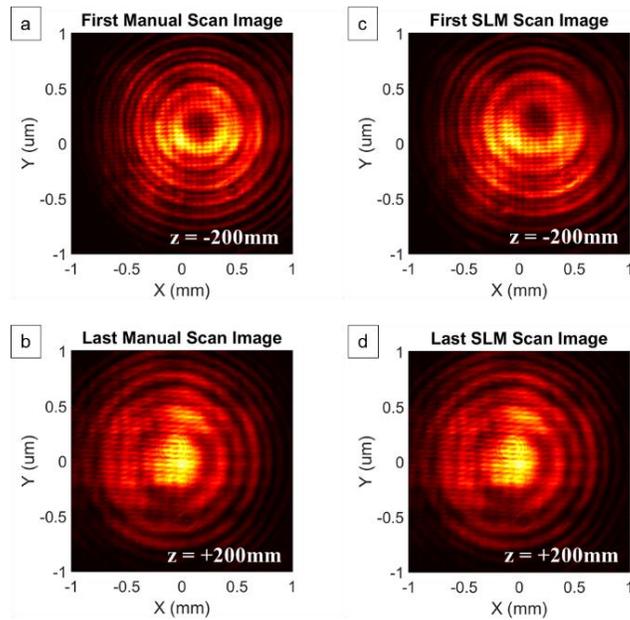


Figure 3.11: Manual and SLM spatial profile comparisons of the Gaussian beam. First ( $z = -200\text{ mm}$ ) and last ( $z = +200\text{ mm}$ ) transverse profile CCD images can be seen for both the manual (a,b) and SLM (c,d) scanning methods

While visual observation would suggest a strong similarity in both scanning methods, it is also important to test that the beam sizes are similar as well. By calculating the  $1/e^2$  beam size for each transverse plane, within the axial scan, the similarities between both methods become clearer; emphasized by Fig. 3.12. When the beam size is taken along the x and y directions of the spatial profile, there is a strong correlation between both scanning methods. Most notably is the beam waist – at focus – for both scanning methods are of nearly the same size, in addition to focusing at the same focal plane.

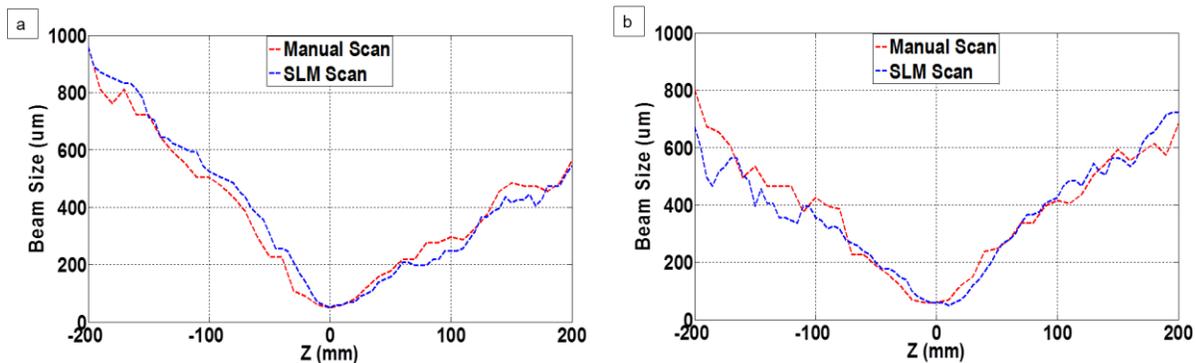


Figure 3.12: Beam size at  $1/e^2$  comparison. (a) x-direction (b). y-direction

### 3.5.2 Astigmatic beam 3D focal intensity scan comparisons

To generate an astigmatic beam, a positive defocus term was added to the x-direction and the same amount removed to the y-direction defocus term ( $A_{2x} = -10$  rad and  $A_{2y} = +10$  rad). Results displayed as MIP of the axial intensity scans for the astigmatic beam are presented in Fig. 3.13. Similar structures are visible in the intensity distribution. Also, visually, the focal planes - which should be different in X and Y for an astigmatic beam - for both methods are positioned at the same z-position.

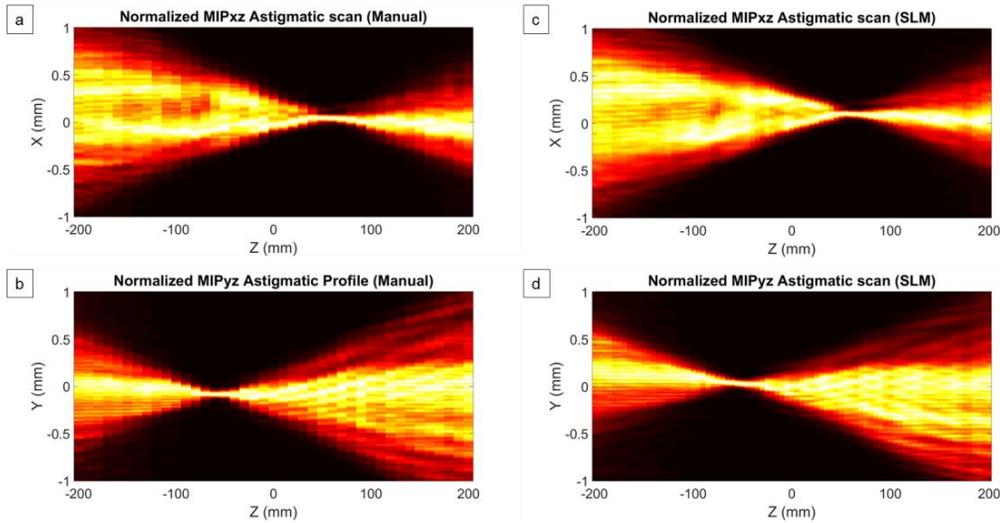


Figure 3.13: Manual and SLM astigmatic beam scan comparisons. Includes MIP along the xz and yz-directions for both the manual (a,b) and SLM (c,d) scan methods.

Supplemental to the MIP are the first and last transverse plane images for both scanning methods; see Fig. 3.14. The beam displays an elliptical shape by being elongated within the x-direction for the first transverse plane, while being elongated within the y-direction for the final transverse plane, which is indicative of astigmatic behavior. Once again, the spatial profiles are nearly indistinguishable between the manual scanning method and SLM scanning method.

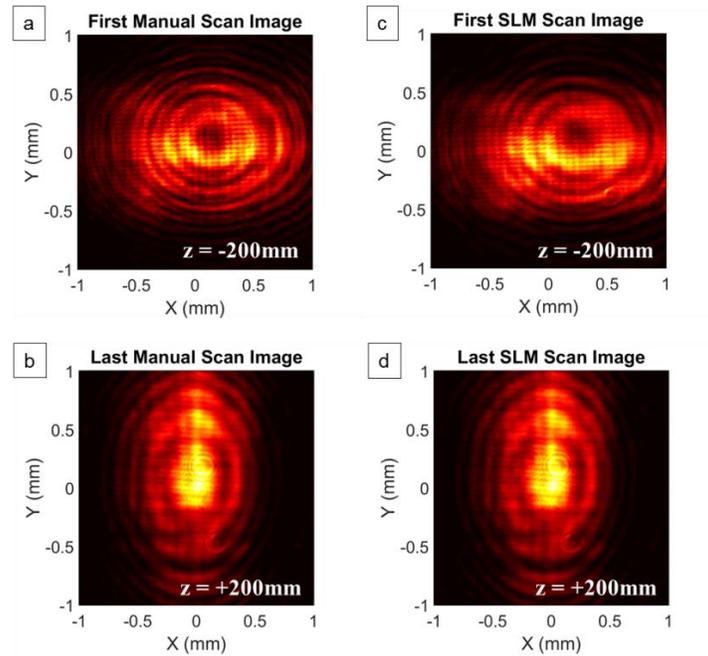


Figure 3.14: Manual and SLM spatial profile comparisons of the astigmatic beam

As before, plots of the beam sizes are shown in Fig. 3.15 for both scanning methods along the x and y-directions. The astigmatic effect has shifted the focus of the beam further towards one end of the scanning range; as opposed the Gaussian beam. In either case, there is matching beam size behavior between both scanning methods. The similarities between the two methods highlights the accuracy to which the defocusing term is calibrated. Observing similar beam waist values at each beam focal plane, as well as the same positioning in z of these planes (at  $f_x = +57$  mm and  $f_y = -55$  mm), is a strong indication that the SLM defocus calibration is correct.

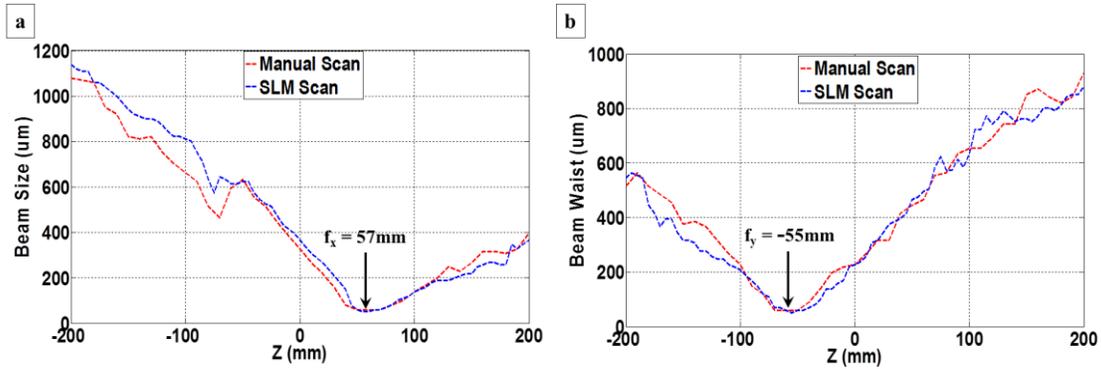


Figure 3.15: Astigmatic beam size comparison. (a) Beam size at  $1/e^2$  for the x-direction, displaying a focus at  $f_x = +57$  mm. (b) Beam size at  $1/e^2$  for the y-direction, displaying a focus at  $f_y = -55$  mm

### 3.6 Discussion

This chapter has demonstrated that it is not only possible to properly calibrate the SLM, so that it may apply the correct defocusing term, but that it is also possible to apply this term in order to create a 3D focal intensity distribution that is comparable to a manual scan; along the same propagation range. The effectiveness of the defocusing term is crucial to the multiplane GA experiments conducted within the next chapter, whose goal is to perform beam shaping and optimize the focal volume of the laser to extend the depth of focus and obtaining thin non diffracting beams.

## CHAPTER 4 FOCAL VOLUME OPTIMIZATION VIA MULTI-PLANE SELF-LEARNING ALGORITHM

### 4.1 Introduction

In chapter 2, it was shown that a Genetic Algorithm (GA) – in conjunction with a Spatial Light Modulator (SLM) as a correcting optics – is capable of wavefront correction of an aberrated Gaussian beam, using the spatial profile captured in a single focal plane as the feedback signal for the GA. One of the goals of this dissertation is to propose the optimization of the **focal volume** of a laser, instead of just the focal plane (focal spot), and explore if a learning algorithm is able to generate a focal volume that possesses a **larger axial extent** than that of a reference Gaussian beam. For that purpose, chapter 3 showcased how to calibrate the SLM for imaging a laser beam's spatial profile at different transverse planes. The generation of extended focal volume (larger extent axially in  $z$ ), has interesting application in several fields, for example in the shaping of the laser used in optical tweezers, or for the generation of thinner light sheets in light-sheet microscopy.

Chapter 4 seeks to provide the following:

- 1) provide a brief summary into non-diffracting beams (section 4.2);
- 2) present the multi-plane GA, a modification of the single plane GA (developed in chapter 2) that uses - as feedback - the signal from several transverse planes (section 4.3);
- 3) showcase results of the multi-plane learning algorithm generating beams with extended axial length (section 4.4 and 4.6); and
- 4) study the effects that the multi-plane optimization range has on the axial length based on images used for feedback (section 4.5).

## 4.2 Non-Diffracting Beams

Unique among laser beams are beams that are able to retain the same intensity distribution over an extended region of the optical axis; several times more than a typical Gaussian beam, which maintain its beam width over roughly two times the Rayleigh Range (cf. section 1.3.4). Beams such as these are classified as *Non-diffracting*. As its name would suggest, non-diffracting beams are able to limit the spread of the optical energy over the entirety of the optical axis. While retention of the beams shape is one of the more prominent aspects of a non-diffracting beam, there are number of other qualities associated with specific non-diffracting beam types. For example, a 1<sup>st</sup> order Bessel beam is capable of a process known as self-healing, where even if a portion of its wavefront is obstructed, it can reform its intensity profile further along the beam axis; see Fig. 4.1.

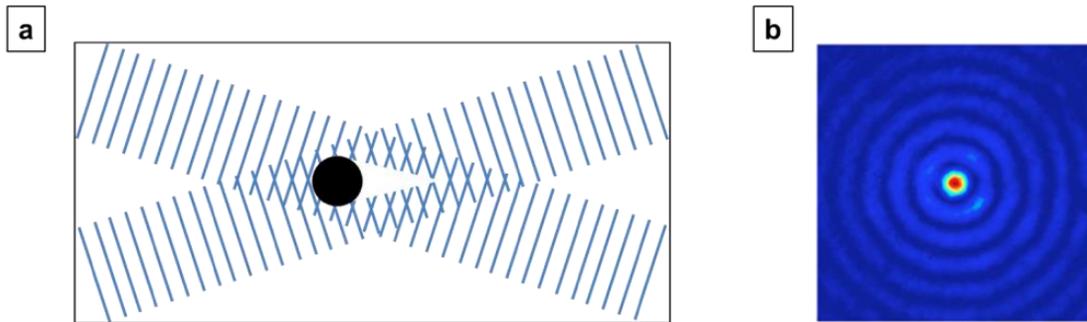


Figure 4.1: Bessel beam wavefront and transverse spatial profile. (a) When two planar waves intersect, they are able to form the Bessel Beam spatial pattern shown in (b) <sup>[141]</sup>. Even if a portion of the intersected wavefront is obstructed, the wave packet is still able to continue and reform shortly after the obstruction

The first notable non-diffracting beam, the Bessel beam, was experimentally demonstrated by Jim Durnin; along with its various modes <sup>[142,143]</sup>. The Bessel beam can be represented mathematically to determine its intensity distribution by Eqn. 4.1.

$$E(\mathbf{r}, t) = \exp[i(\beta z - \omega t)] \int_0^{2\pi} \exp[i\alpha(x \cos \varphi + y \sin \varphi)] \frac{d\varphi}{2\pi} = \exp[i(\beta z - \omega t)] J_0(\alpha \rho) \quad (4.1)$$

In this instance,  $\rho^2 = x^2 + y^2$  and  $J_0$  is the zero-order Bessel function of the 1<sup>st</sup> kind. Further discoveries into differing types of non-diffracting beams were made over several years, including Airy beams, Mathieu beams, and wave-based parabolic cylinder beams [83,144-147]; see Fig. 4.2.

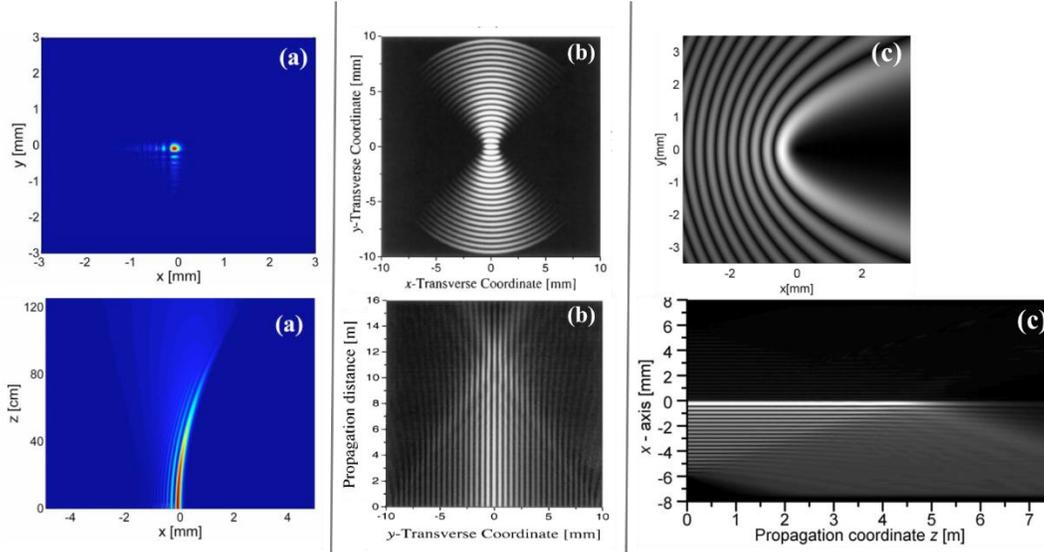


Figure 4.2: Various non-diffracting beam types. 2D spatial profile (top) and 2D longitudinal profiles (bottom) of the (a) Airy [146], (b) Mathieu [144], and (c) parabolic [147] non-diffracting beams

Non-diffracting laser beams are popular for their extended propagation range and reduced beam width [148,149]. The self-reconstructing behaviors they exhibit have a number of applications for fields requiring the use of optical beams. Today, several advancements have been made in generating these types of beams, such as the use of an axicon lens for producing Bessel beams [150], or using an active optics to generate non-diffracting beams [151].

### 4.3 Multi-plane Genetic Algorithm

After demonstrating the wavefront correction and shaping ability of the GA-guided SLM, as well as how calibrating the SLM for defocusing can simulate manual translation of the CCD (Charged Coupling Device), it is now possible to determine if this learning algorithm setup can generate extended focal volumes. Chapter 3 referred to the concept of a 3D focal intensity

distribution, containing several transverse images collected at various planes along a longitudinal axis which, when compiled, can generate of 3D shape (profile) of the beam. Important for this discussion is the axial length of this profile, as it is an important characteristic in defining what is known as the **Focal Volume**.

#### 4.3.1 Focal volume and depth of focus

The definition for focal volume used in this dissertation is the following: “The focal volume is a volumetric region in space where the light intensity is large than 50% of the maximum intensity at focus”. In the case of a Gaussian beam, it is equal to the depth of focus (also called “confocal parameter”) times the transverse (focal) area. The depth of focus is equivalent to  $Z = \pm Z_R$ , or twice the Rayleigh range and is shown in Eqn. 4.2 for a Gaussian beam.

$$\text{Depth of Focus (DOF)} = 2Z_R = \frac{2\pi w_0^2}{\lambda} \quad (4.2)$$

As a reminder,  $w_0$  is the beam size at  $1/e^2$  of the maximum intensity of the spatial profile at a given z-plane, and  $\lambda$  is the wavelength of the laser. Fig.4.3. illustrates how this depth of focus is tied to the Rayleigh range and waist of the Gaussian beam.

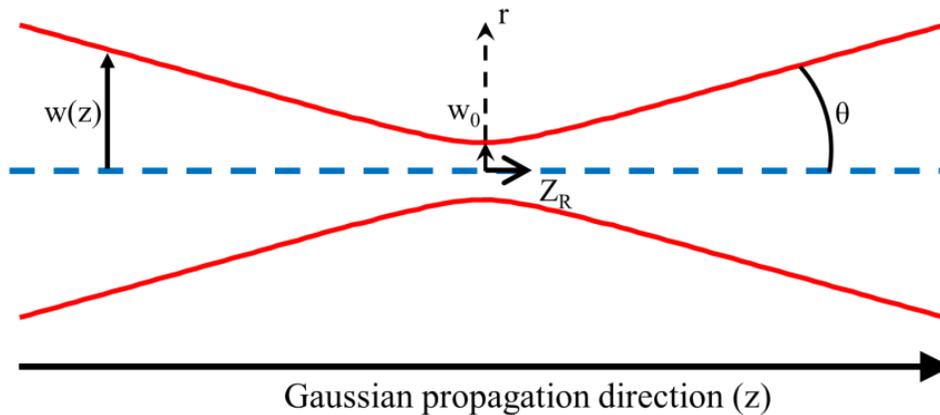


Figure 4.3: Beam size of a Gaussian beam.  $Z_R$  is indicative of the Rayleigh range.

The transverse area of a Gaussian beam is defined in Eqn. 4.3. The FWHM (Full Width, Half Maximum) of the Gaussian beam can then be related to the beam waist ( $w_0$ ) by  $FWHM = w_0\sqrt{2\ln(2)}$

$$\text{Transverse area } (A) = \pi \left( \frac{FWHM}{2} \right)^2 \quad (4.3)$$

Substituting the FWHM into Eqn. 4.3, and then multiplying both the depth of focus and transverse area, the solution for the focal volume can be found Eqn. 4.4.

$$\text{Focal Volume} = \text{DOF} \times A = \frac{(\pi \ln(2))^2}{\lambda} w_0^4 \quad (4.4)$$

Our goal is to extend the depth of focus of the focal volume. For the GA to optimize the distribution of the focal volume, instead of the distribution of a single transverse plane, the feedback measurements the GA receives should include **more than one transverse plane**. As the spatial CCD cannot be moved at different longitudinal positions accurately and quickly enough for the GA, changes in the longitudinal plane is performed by adding a focus phase on the SLM, in the same manner that was presented in chapter 3 to measure the full 3D focal intensity distribution. This approach is referred to as **multi plane GA**, in contrast with the **single plane GA** used in chapter 2, which only uses information from a single transverse plane.

#### 4.3.2 Experimental setup for multi-plane GA

Fig. 4.4 is an example of how an astigmatic beam would look like if only five transverse profiles were measured. It appears that the main characteristics of a 3D focal intensity distribution are visible even if only a few transverse planes are used to sample the beam structure, so it seems

reasonable to think that a GA using signals from these five planes should be able to optimize the focal volume; not just one single plane.

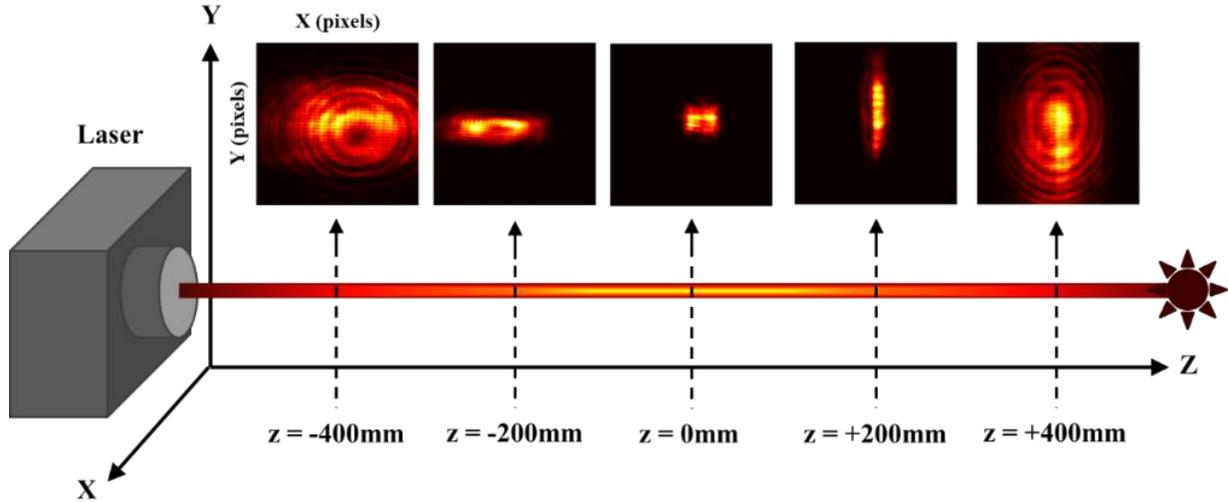


Figure 4.4: Illustration of focal volume for an astigmatic beam. The top five figures are spatial profiles taken at regular interval in  $z$

The experimental setup continues to use the Michelson interferometer layout shown in chapter 2 and 3; albeit with the spatial profile arm of the setup. The reference mirror is blocked, and the spatial profile camera is situated at a distance of 740 mm away from the SLM image plane. It is important to note that the spatial CCD must be situated at this range, as the defocus calibration results conducted in chapter 3 are only valid for this position. Attempting to provide a defocus term outside of this position would potentially lead to an inaccurate defocus term being produced. During the entirety of the GA optimization, the spatial CCD always stay fixed while still retaining its ability to capture transverse images at differing focal planes using the SLM defocus term. Fig. 4.5 displays the experimental arrangement, which includes the valid focal volume range. A transparent section shown atop of the camera and rail symbolizes the range in which an accurate defocus range can be produced.

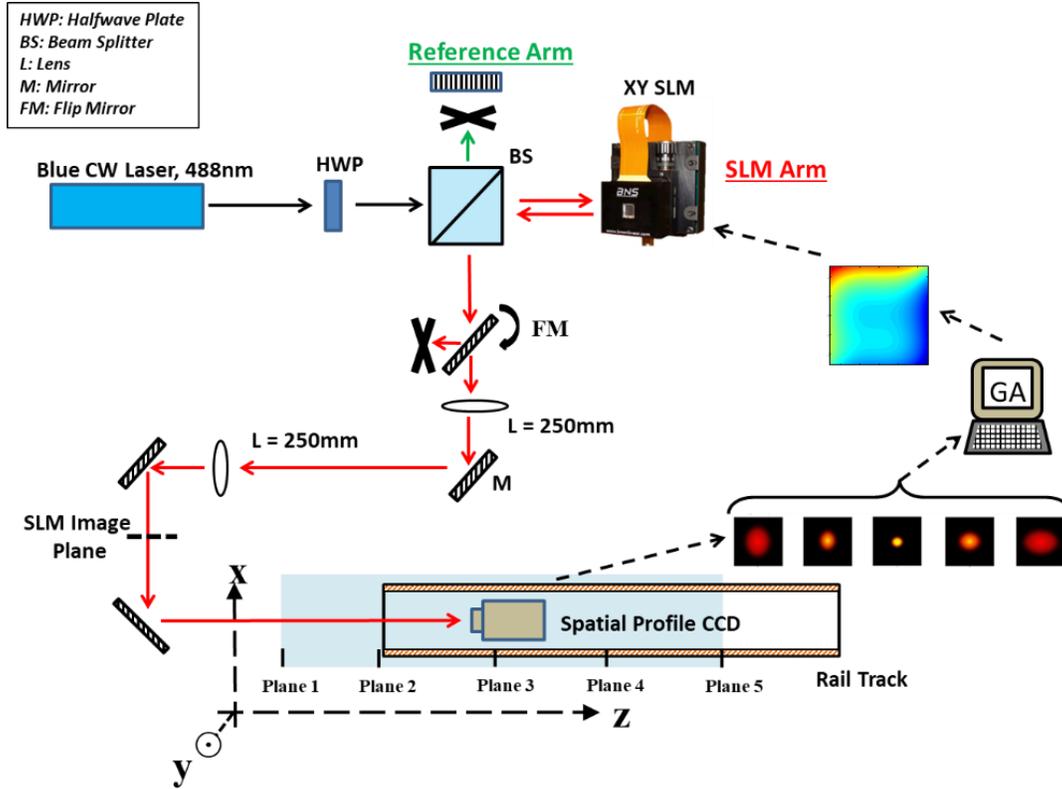


Figure 4.5: Experimental setup used by the multi-plane, learning algorithm.

Five transverse plane images are collected by the spatial profile camera, and a minimum ( $z_{\min}$ ) and maximum ( $z_{\max}$ ) longitudinal range is defined for optimization. There is no limit in how many transverse planes can be used to evaluate an individual in the multi-plane GA, but for all multi-plane optimizations performed in this dissertation at least five transverse planes were used. Each image taken lies at different distances within the focal volume. Two images are taken at  $z_{\min}$  and  $z_{\max}$ . Another is taken at the focal plane;  $z_0$ . The remaining two images lie halfway between  $z_{\min}$  and  $z_0$ , and  $z_0$  and  $z_{\max}$ . Capturing the spatial profiles from five transverse planes ensures we have enough information to generate an optimized focal volume, while also ensuring that we do not exacerbate the time needed for the GA to go through each iteration.

### 4.3.3 Fitness method construction

Section 2.4.3 detailed how the fitness method evaluates each individual (gene) in order to provide a fitness score for each one. The fitness methods used in this chapter has to evaluate multiple transverse planes, rather than a single plane. This means that each of the captured transverse images is used in the fitness evaluation process. In this section, details are provided into each of the fitness methods constructed for the purposes of axially extending the focal volume. Table 4.1 contains a list of each of the fitness methods that were developed to meet this challenge.

The first logical fitness methods to consider are similar to the ones applied for wavefront correction using the single plane GA presented in chapter 2. The first method, called “Maximum beam intensity average” records the maximum intensity in each plane, and averages the five values of maximum intensity. The second method, called “Beam size average”, calculates the beam size in each plane, and takes the average value of these five beam sizes. It seems logical that maximizing the beam intensity or minimizing the beam size across all planes should make a beam more constant in size and intensity; hence with larger depth of focus. While these methods work in the correct direction, they are not very successful because nothing prevents one plane from being considered better (more intense or smaller in size) than other planes, which limits potential for a good fitness score. When these methods were used, converging beams or diverging beams were generally obtained. Other methods had to be developed to improve GA efficiency, which **compare information between the planes** rather than adding values extracted from each of them.

Fitness method	Description / Goal of the method
Maximum beam intensity average	Maximize the beam intensity across all planes
Beam size average	Minimize the beam size across all planes
Spatial profile similitude	Have a similar intensity distribution across all planes
Beam size difference	Maintain a similar central beam size across all planes
Focal point mask	Restrict the beam into a circular mask

Table 4.1: List of Fitness methods used for axially extending the focal volume of the laser beam

## Spatial profile similitude method

By seeking similitude in 2D intensity distributions across each of the transverse images, the similitude fitness method should direct the multi-plane GA to find a phase array that would shape the beam to have similar spatial profiles across all 5 image planes. The similitude method is designed to minimize the squared difference (Eqn. 4.5) between each of the transverse images.

$$\text{Squared Difference} = \frac{1}{N} \sum (I - I')^2 \quad (4.5)$$

Here,  $I$  is the entire transverse image at one plane, while  $I'$  is the entire transverse image at another plane.  $N$  is total number of combinations that were made between different transverse images. An illustration of this procedure can be seen in Fig. 4.6.

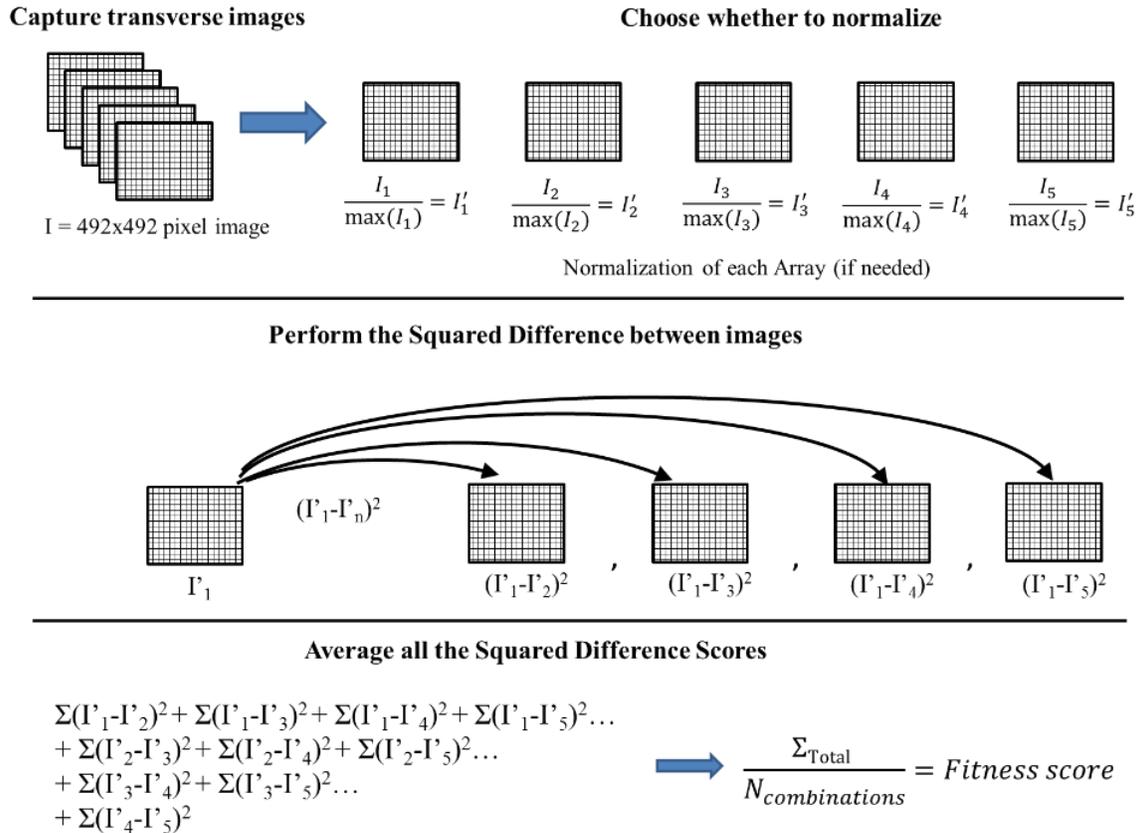


Figure 4.6: Manipulation of image for the similitude fitness method

This method can have the option to normalize - or not normalize - the intensity of each recorded image on the onset. Then, the squared difference between each of the images is found. Finally, upon obtaining the sum of all these values, they are divided by the total number of combinations in order to produce the fitness score for the GA. The squared difference value serves as the fitness score, and the multi-plane GA seeks to minimize this score. Lowering this score would mean the spatial profile for each plane is equally shared, which should result in a line-focus; or non-diffracting beam. In an ideal case, without noise, a beam with identical spatial profiles - across each image plane - would produce a fitness score of 0.

### **Beam size difference**

A common characteristic of a line focus or non-diffracting beam is that it maintains a well-defined, central beam spot over a relatively large propagation range. If the multi-plane GA could produce a beam shape with similar central beam sizes across each transverse plane, it could generate that same line focus. The fitness method proposed here is intended to minimize differences in the beam size over several transverse planes. This method evaluates the beam size (whether at  $1/e^2$  of the maximum intensity or FWHM) for one image plane, and then subtracts it from the beam size at a different plane. This is applied for each subsequent transverse image, with a vector containing each of the difference values being created. An average of this vector is then taken, and this value becomes the fitness score. As the first image has nothing to compare itself to, it is not counted when producing a fitness score.

### **Focal Point Mask**

The Focal Point Mask is a phase method that seeks to maximize the intensity of the beam within a circular region on the spatial profile camera. Within the dimensions of the 512x512 pixel

SLM array, a circular region with a radius  $R$  is created. The size of the mask can be changed internally to allow for a smaller region of interest, or much larger. Once this mask is created, the fitness method measures the intensity of each pixel within the mask, before averaging them. This average intensity value becomes the fitness score the multi-plane GA uses when determining the best individual. Overall, the purpose of this method is to restrict and shape the beam so that it can produce a line focus across multiple planes. Fig. 4.7 provides a diagram for how this method works.  $I_{pix}$  is the intensity measured at each pixel, and  $N_{pix}$  is the number of pixels measured.

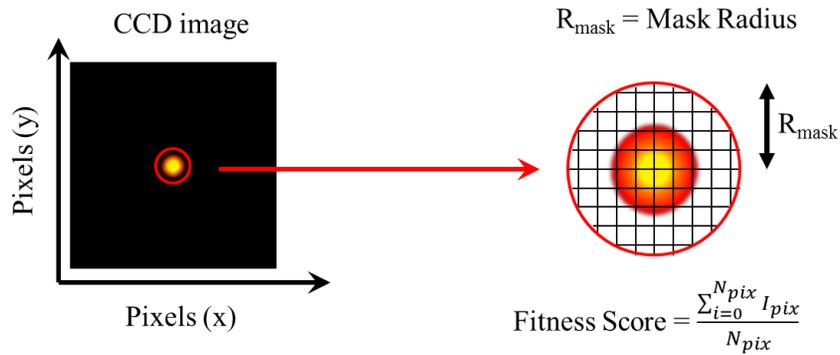


Figure 4.7: The focal point mask fitness method.

#### 4.3.4 Illustrating the fitness score for idealized beams

For the sake of illustrating some expected behavior of the GA when using each particular fitness method, this section takes a theoretical perfect Gaussian beam and several idealized beams (that may not be achievable at all) and displays their fitness scores. This provides a qualitative understanding of the usefulness of each of these different fitness methods, and how the multi plane GA using them would eventually converge if they are used. Fig. 4.8 displays the five spatial profiles corresponding to five  $z$ -planes, for four idealized beams. The first beam (Fig. 4.8a) is a theoretical Gaussian beam similar to the one generated in the geometry of our experimental setup. The second and third beams (Fig. 4.8b & Fig. 4.8c) are idealized beams that have constant profiles

corresponding to the Gaussian beam planes 2 and 3. The fourth beam (Fig. 4.8d) is simply noise (all planes have similar noise), which would happen if the SLM just scatter the light away. The maximum intensity and the beam size (FWHM) are indicated in the figure, along with the corresponding fitness scores for four fitness methods presented in the previous section. Beam (b) is the most desirable output, providing a large axial extend for a small beam size, then beam (c) is the next desirable output beam. Methods 1 and 2 (Maximum beam intensity average and beam size average) improve the fitness score for beam (b), but not for beam (c), as these methods give a better score for the initial reference Gaussian beam. Consequently, as indicated in the previous section, these methods converge very easily toward diverging or converging beams that contain one image plane with a tight focus. As a result, methods 3 and 4 (spatial profile similitude and beam size difference) for image comparison has been used more extensively.

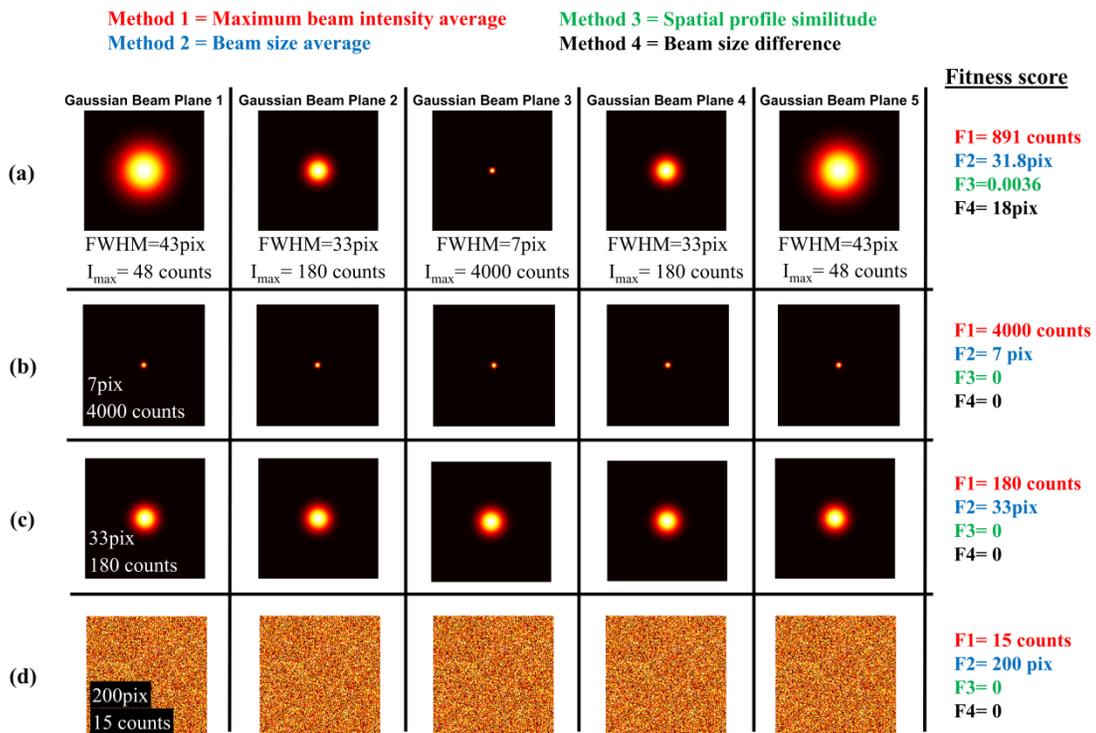


Figure 4.8: Illustration of the fitness score for different fitness methods implemented. The fitness scores are indicated to the right. (a) Theoretical Gaussian beam, (b and c) Idealized target non-diffracting beams that possess the profiles of plane 3 (b) and plane 2 (a) in all of their 5 planes. (d) noise filled camera

Nevertheless, the simple examples of Fig. 4.8 with their corresponding fitness scores  $F_3$  and  $F_4$ , all equal to 0 for beams (b) (c) and (d) show that these methods do not place any constraint on the beam size in any plane. This means that beams (b) and (c), while having different sizes, could both minimize the fitness and be a best solution for the GA to converge. It is also noteworthy that the beam (d) that has been scattered away from the camera, containing only background camera noise, also would be a best solution for the GA with methods 3 and 4.

Unique to these two methods is the necessity to include a thresholding mechanism, which is intended to reduce the possibility of the GA to converge to a solution that could lead to a very large beam size or a solution with no beam at all on the camera. With thresholding, at the beginning of every generation a corrected Gaussian beam is generated in order to serve as a reference from which the threshold is determined. Upon generation, the CCD measures the intensity of the Gaussian beam in the focal plane, and 5% of its maximum intensity is used as the threshold. From here, the GA measures the maximum intensity found across all five transverse planes to see if at least one plane meets the threshold. If the threshold is not exceeded, the best individuals from the previous generation are used for scoring of the current generation.

#### 4.3.5 *Phase creation methods*

In chapter 2, multiple phase methods were outlined and implemented with the intent of wavefront correction. These phase methods included the Cartesian method and Zernike method. Compared to the experiments in this chapter, the application of these phase methods was more targeted in chapter 2, since the goal of wavefront correction is to minimize the amount of aberration acting upon the Gaussian beam; within a single plane. A greater degree of exploratory freedom is needed on behalf of the multi-plane GA to find a beam with a line-focus (or non-diffracting), and one way of achieving this is by implementing new phase methods for the GA to use.

## **Annular Phase Method**

In seeking to create a non-diffracting beam, it is necessary to understand how these various beams can be generated. For a Bessel beam, one method is to use what is known as an annular lens (annulus). The annulus contains a ring with a predetermined thickness, which allows a Gaussian beam to pass through; known as masking. Depending on the minimum ( $R_{in}$ ) and maximum ( $R_{out}$ ) radius of the ring, the amount of energy throughput and profile produced by the annulus will change. While this phase method is well known in its ability to formulate a Bessel beam, the straightforwardness for it to do so limits its potential for generating any other type of beam with an extended focal volume.

## **Radial Phase Method**

Another alternative for generating a non-diffracting beam is the use of an axicon lens. The axicon is a specialized lens with a conical surface, and is able to generate with propagation a ring-like spatial profile that converts a Gaussian beam into a Bessel beam. Axicons work by creating the impinging wavefront seen in Fig. 4.1, which allows for a Bessel beam to be formed. In addition, the surface of the axicon can be curved, allowing for variable focusing of the Bessel beam; a trait the annular phase method does not provide. Rather than simply making a phase method solely designed for making an axicon, a new phase method is produced to provide this phase shape, as well as additional phase terms. This new method is called the radial phase method. The radial phase method is constructed within a polar coordinate style system. Both the x and y dimensions of the 2D SLM array are converted into the radial coordinate  $R$  using Eqn. 4.6. Each phase term corresponds to a higher-order of  $R$ , and the magnitude of the phase term is represented as  $A$ ; see Eqn. 4.7. An illustration of the  $A_1R$  phase term producing a phase equivalent to an axicon lens can be seen in Fig. 4.9.

$$R = \sqrt{x^2 + y^2} \quad (4.6)$$

$$\varphi_{Radial} = A_0 + A_1R + A_2R^2 \dots A_nR^n \quad (4.7)$$

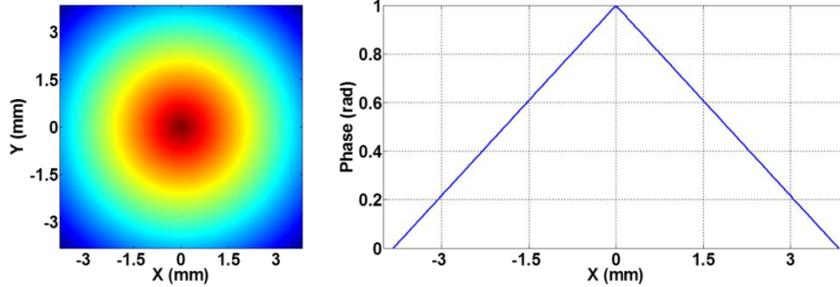


Figure 4.9: 2D axicon phase term sent to the SLM (left), and linecut taken horizontally (right)

As was previously discussed in section 2.6, greater exploratory freedom comes at a cost to computational agility and increased iteration times. Given that the multi-plane GA requires the analysis of multiple transverse images, there is already a reduction in the computational speed of the GA. Potentially compounding this issue could be the fitness method chosen for the GA, as the method of evaluation across multiple transverse images could also slow the GA. For there to be an even trade-off between exploratory freedom, and computational time, balancing the fitness method, phase method, and the number of phase terms, is the key.

#### 4.4 Spatial profile similitude fitness experiment

This section seeks to apply the “spatial profile similitude fitness method” as a mean of generating focal volumes with increased depth of focus. This fitness method is tested under a single phase method (Radial phase), but the efficacy of the fitness method is tested by having two scenarios where normalization is, and is not, used.

#### 4.4.1 Similitude fitness without normalization

In this experiment, optimization of the beam occurs over a range of 400 mm. This method begins by using the second and third phase terms ( $A_1$  and  $A_2$ ) of the radial phase. Table 4.2 contains the GA parameters used.

GA Parameter	Radial Phase Method
CCD Exposure	0.6ms
Number of Generations	100
Best Individuals ( $N_{\text{best}}$ )	4
Mutated Individuals ( $N_{\text{mutation}}$ )	16
Crossover Individuals ( $N_{\text{crossover}}$ )	12
Number of terms	2
Range of mutation (rad)	200

Table 4.2: Starting parameters for the similitude method with no normalization of the intensity values

Upon testing of the spatial profile similitude method (w/o normalization), the multi-plane GA was able to shape the reference Gaussian beam into different profiles. However, as the number of generations increased, the GA slowly began to make the laser more diffuse across each of the transverse planes. What is noteworthy about this change in behavior is that it is also tied to **a reduction in the fitness scores measured by the GA**. Further generations led to the GA spreading the laser intensity so greatly, that the intensity from the beam became indistinguishable from background noise measured by the spatial CCD. Fig. 4.10 displays the five transverse images captured after applying the un-normalized similitude method.

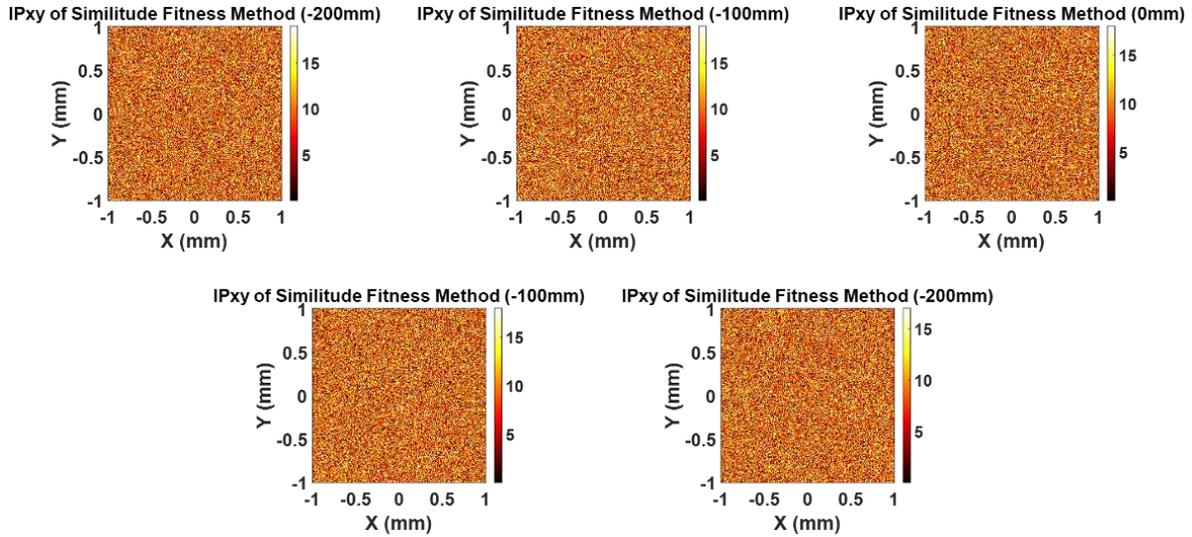


Figure 4.10: Results of the un-normalized similitude fitness method across all 5 planes

What can be inferred from this outcome is that the GA located an optimum; which was not the desired solution, but that was considered as the best solution by the GA based on the defined fitness. Essentially, through optimization, when it applied the squared difference it found less variance within the background noise than it did with any of the intensity profiles with beam seen in the camera. This outcome is not without justification, as the measured PV of the background noised is 15 counts, with an RMS of nearly 2 counts. This is likely seen by the GA as the better “solution” than that of making a beam with a line focus, but with greater fluctuations in the beams intensity. It is worth noting that the thresholding mentioned in section 4.3.4, to prevent the GA to converge toward noise, was not yet implemented when this experiment was performed.

#### 4.4.2 *Similitude fitness with normalization*

In an effort to mitigate the results of the previous section, this fitness now implements a normalization procedure for every transverse image. Normalizing each image means that the intensity values now only range from 0 to 1. This translates to a smaller RMS value than that of the background noise captured by the CCD, and also means that the squared difference value

between planes is much smaller as a result. An interesting notion to consider with this, is that normalization – in this manner – would mean that the GA seeks to optimize the shape of the beam, rather than its intensity value, across all planes.

This section uses the same radial phase method, as well as the same GA starting parameters and optimization range used in section 4.4.1. Initialization of the GA, using this new normalized similitude method, begins much in the same as the previous method. After several generations, the GA began to generate a Bessel beam that shifted focus across each of the different planes, in addition to an annulus forming at certain times. Continued iterations eventually saw the GA forming a Bessel-like beam across each of the transverse planes. Fig. 4.11 shows the 2D longitudinal profile ( $I_{yz}$ ) of the resulting Bessel beam, as well as the first and last transverse planes of the longitudinal profile. 81 transverse images encompass the longitudinal profile, with each separated by a distance of 5mm.

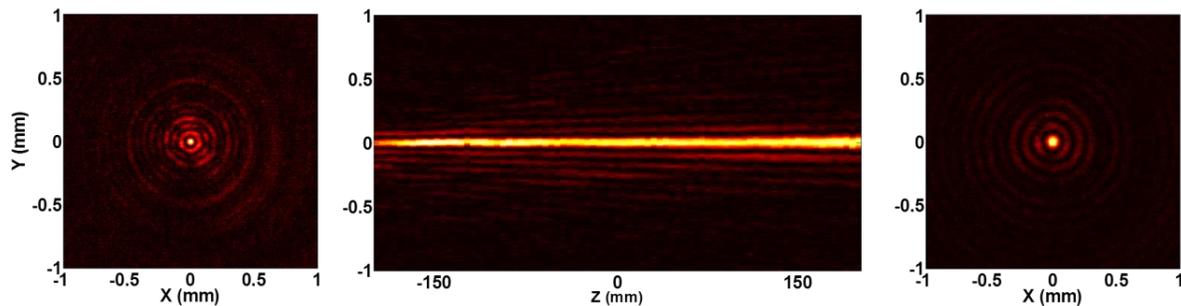


Figure 4.11:  $I_{yz}$  of the similitude method Bessel Beam. The transverse image on the left was captured at -200mm (540mm), and the one of the right was captured at +200mm (940mm)

The longitudinal profile of the Bessel beam showcases a strong, line focus which propagates over the entirety of the 400 mm optimization range. There is a discernable amount of divergence in the beam size as the beam travels from -200 mm to +200 mm, which is more noticeable in the higher-order fringes surrounding the central Bessel beam spot. This is also detected in the first and last transverse images, in which the central beam spot of the Bessel is

larger in the last plane than in first; along with fewer rings. Given that the radial phase is also imparting a focus term on the SLM phase array, this is a likely consequence. Calculating the beam width (as FWHM) of the central beam spot finds a value of  $58.5\ \mu\text{m}$ . Analyzing the 2D SLM phase array, by taking a horizontal linecuts across its center, showcases how the algorithm is able to achieve an axicon-like profile in order to generate the Bessel beam; see Fig. 4.12. The curvature of its surface is the result of the defocus term being applied to shift the focus of the Bessel beam.

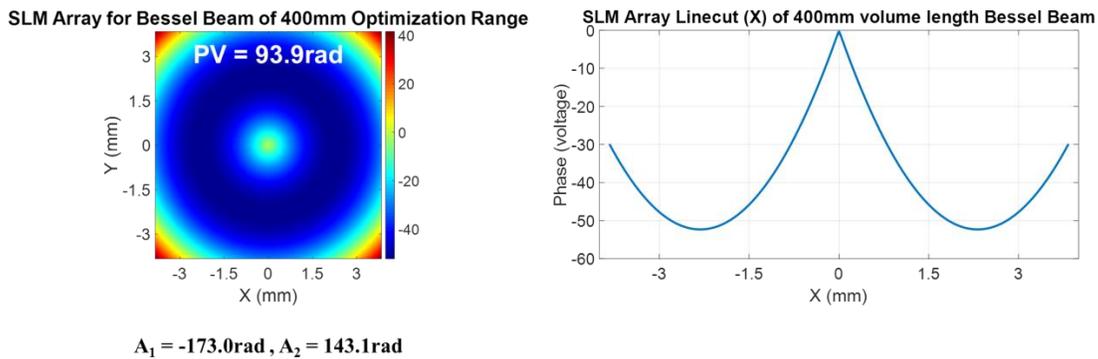


Figure 4.12: SLM phase results after normalized similitude method beam shaping. (left) 2D SLM phase array and horizontal linecut of the Bessel beam with an optimization range of 400mm (right)

The generation of the Bessel-like beam is a welcomed discovery, and highlights the importance that a fitness method can have on the multi-plane GA. The simple application of normalization was significant enough for the GA to go from finding a solution that only produced background noise, to a Bessel-like beam. These results also prove the effectiveness in having a phase method that is conducive to generating extended focal volume, while also being compatible with the fitness method.

#### 4.4.3 Bessel beam vs Gaussian beam depth of focus comparison

After generating the Bessel beam with the similitude method, comparisons can be made to determine how greatly the GA was able to extend the focal volume of the beam. As the focal

volume is dependent upon the depth of focus, this depth can be used to assess the degree in which the focal volume was impacted. A Gaussian beam will serve as a reference, and the Bessel beam that was optimized for 400 mm will be compared against it. Determining the depth of focus uses the 2D longitudinal profile of both beams, with a linecut taken across the central beam spot of both beam types. This produces an intensity plot in  $z$ , and taking the FWHM of this new plot will lead to the depth of focus value; seen in Fig. 4.13. Both beams are normalized to 1, the Bessel beam being normalized to the center of mass of the intensity region lying between -300 mm to 100 mm.

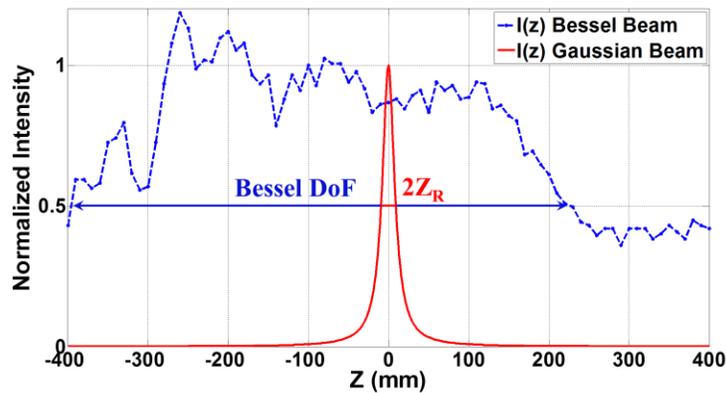


Figure 4.13: 1D longitudinal intensity for the Bessel beam (blue dashed line) and Gaussian beam (red solid line)

The depth of focus for the Gaussian beam is measured to be 18.6 mm, while the Bessel beam measures in at 609 mm; more than 32x the Gaussian beam. This extension of the depth of focus and – subsequently – the focal volume, demonstrates the potential for the multiplane GA to improve the axial dimension of the focal volume.

#### 4.5 Bessel beams of varying optimization ranges

While it has been demonstrated that the multi-plane GA is capable of producing a Bessel beam, it is necessary to prove that the algorithm is also able to optimize the beam for a definitive focal volume distribution. To illustrate this point, this section seeks to create three additional

Bessel-like beams using the phase and fitness method used in section 4.4.2. As there is a finite amount of energy in the laser beam, it is expected that a larger optimization range would drive the algorithm to expand the central beam size of the Bessel beam; to better distribute the beams energy. In turn, this would also lead to a reduction in the total intensity. Inversely, a smaller central beam size should be achievable by having a reduced optimization range, as energy does not have to be maintained over a larger propagation range.

Bessel beams with optimization ranges of 50mm (or 715mm to 765mm), 200mm (or 640mm to 840mm), and 600mm (440mm to 1040mm) were generated. Comparisons also include the Bessel beam made over the 400mm optimization range. In order to see how well these Bessel beams maintain their shape under these specific optimization ranges, a 2D longitudinal scan ( $I_{xz}$ ) was taken for each beam type over 800mm; or -400mm to 400mm. Fig. 4.14 displays each Bessel beam with varying optimization ranges indicated as dashed blue lines that correspond to the planes where the 5 images used during optimization were taken. Each of the Bessel beams produced by the multi-plane GA can carry themselves over a fairly large longitudinal range. The Bessel beam optimized for 50 mm displays a discernable level of diffusion, with the Bessel beam not forming until -280 mm. By contrast, the Bessel beams with 400 mm and 600 mm focal volume lengths are able to maintain central beam focus over the entirety of the scanning range. Each beam is subjected to a detectable level of divergence that is attributable to the focusing term. The diffusion and divergence seen here is most likely due to the fact that the GA is forced to optimize at a given distance from the SLM image plane, which restricts the maximum numerical aperture (NA) of the solution. Table 4.3 contains the average beam size values of the central beam spot, in addition to the  $A_1$  and  $A_2$  phase terms of the radial method. As theorized, a larger optimization range causes

an increase to the overall size of the central beam of the Bessel. What is also clear is that a significantly larger focus term is needed to extend across a greater optimization range

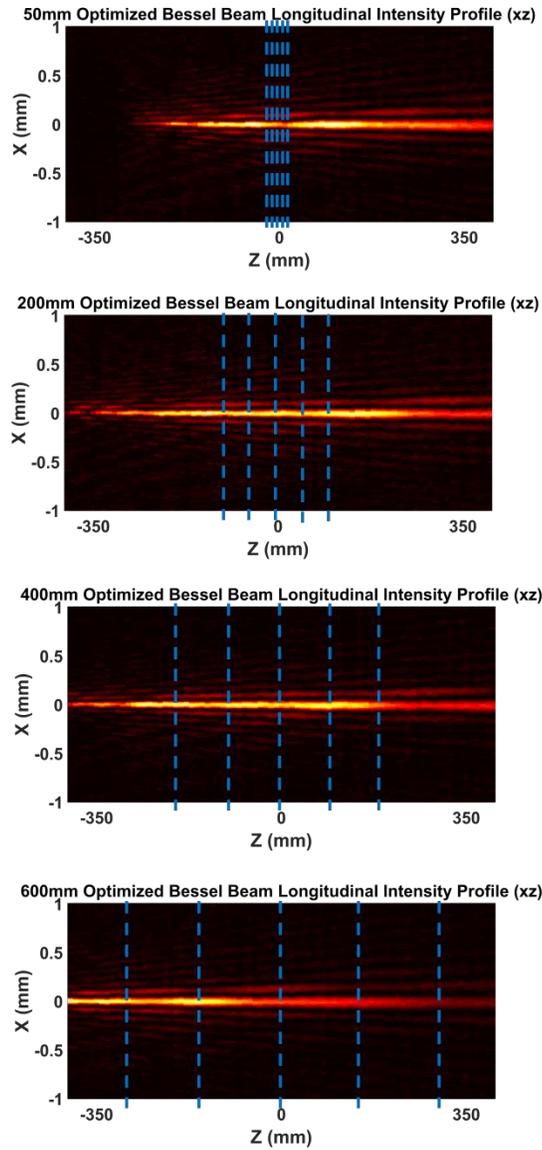


Figure 4.14:  $I_{xz}$  of the Bessel beams w/ varying focal volumes. The blue vertical lines indicate each z position a transverse image was taken during each beams respective multi-plane GA run.

Bessel Beam (Optimization Range)	Average Beam size ( $\mu\text{m}$ )	$A_1$ (radians)	$A_2$ (radians)
600mm	73.5	-145.8	154.7
400mm	58.5	-173.0	143.1
200mm	51.8	-174.8	135.2
50mm	48.2	-130.2	91.7

Table 4.3: Average beam size values and phase terms for Bessel beams of varying optimization ranges

Multiple 2D SLM phase arrays for each respective Bessel beam, in addition to a plot containing all their linecuts, are shown in Fig. 4.15. At the center of each array is a pointed tip (of value 0 radians). As the focal volume range increases, the amount of curvature added to the initial axicon phase pattern increases. This curvature provides a focusing effect that optimizes the Bessel beam for its respective optimization range.

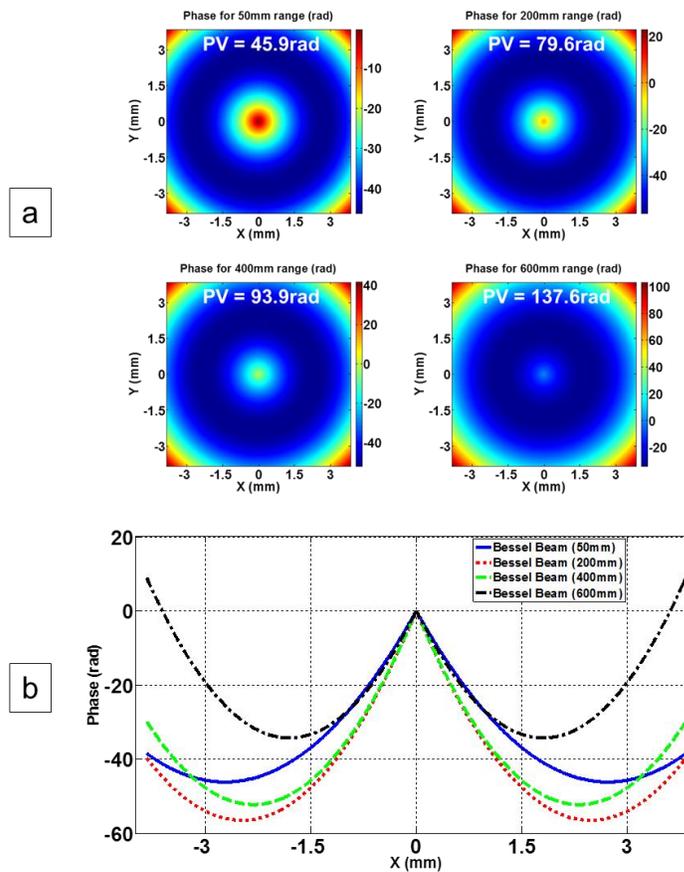


Figure 4.15: SLM phase arrays (a) and their associated horizontal linecuts (b)

When taking a horizontal linecut across this point, the shape of the axicon is more noticeable. At the wings of each axicon is a degree of curvature, which stems from the radial phase producing a parabolic phase shape for focusing. At greater optimization ranges, the parabolic shape becomes more pronounced, which is supported by the  $A_2$  values found in Table 4.3.

#### **4.6 Beam size difference fitness experiment**

This section seeks to test the effectiveness of the beam size difference fitness method and its ability to extend the focal volume. Differentiating this section from the section 4.4 will be the use of one fitness method (beam size difference) as a control, but with the application of two different phase creation methods. This will include the radial phase and the Cartesian phase. As the radial phase method was able to produce a Bessel-like beam, it has been adopted into this sections experiment to see how significant an impact a change in the fitness method can have on a phase method that was previously able to find a good solution. The Cartesian method will juxtapose the radial method, as its ability to operate in two separate dimensions could afford it greater shaping capabilities than that of the more uniform radial method.

##### *4.6.1 Beam size difference minimization using the Radial phase method*

Prior to starting the experiment, the GA parameters will be similar to the parameters in section 4.4.2. With a 0.6ms exposure rate for the camera, the reference Gaussian beam reaches an intensity value of nearly 3900 counts. At a factor of 5% of the Gaussians maximum intensity, this means that GA will only accept the best individuals of the current generation so long as one of the planes that is evaluated has a maximum intensity greater than 195 counts. Upon initializing this experiment, the multi-plane GA sought to generate a Bessel-like beam once again that spanned across all 5 planes. However, the Bessel-like beam produced under this fitness method experienced

a larger degree of divergence compared to those produced by the similitude fitness method. Figure 4.16 shows the first and last spatial profile image of the results reached by this fitness method, as well as the yz longitudinal intensity profile.

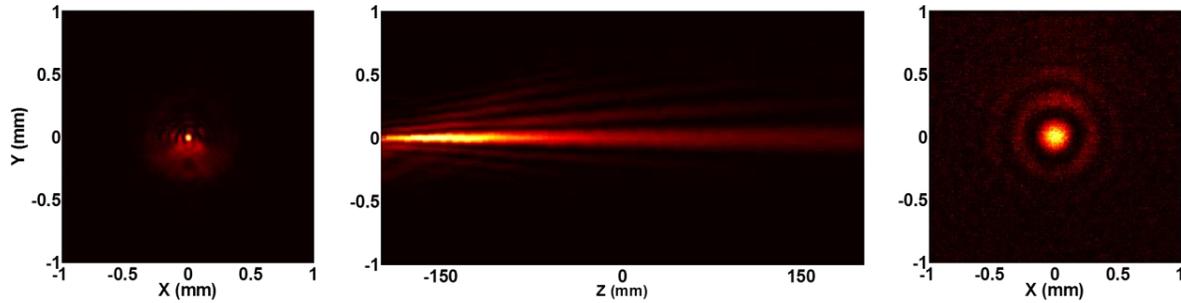


Figure 4.16: Bessel-like beam generated using the beam size difference fitness method

An intensity count of 450 at -200mm was measured, and an intensity count of 66 at +200mm. Such a low intensity value sends the beam closer to the background noise level of the CCD. When measuring the beam size (FWHM), the first plane (-200mm) has a beam size of 39.6  $\mu\text{m}$ , and enlarges to 178  $\mu\text{m}$ . Across all planes, the average beam width was measured to be 107  $\mu\text{m}$ . This larger beam width stands in stark contrast to the average beam size of the similitude method created Bessel beams, with the 600mm optimized Bessel beam maintaining a smaller beam size (73.5  $\mu\text{m}$ ) across a greater optimization range. This fitness method, in conjunction with the radial phase method, is unsuitable in generating a well-defined non-diffracting beam.

#### 4.6.2 *Beam size difference minimization using the Cartesian phase method*

This section will implement the Cartesian method to see if it can produce a non-diffracting beam. As the Cartesian method is able to apply a focus term, as well as modify itself in two separate dimensions (x and y), there is the possibility that it could alter the spatial profile of the beam so as to generate a non-diffracting beam. To add more complexity to this system, two additional phase

terms ( $A_{3xy}$  and  $A_{4xy}$ ) will be added in order to allow for greater freedom in the methods ability to shape the surface of the phase. As the use of only the  $A_2$  terms would merely create a focus at only a single plane, this action was needed. Table 4.4 contains the starting GA parameters.

GA Parameters	Cartesian Phase Method
CCD Exposure	0.6ms
Number of Generations	100
Best Individuals ( $N_{best}$ )	4
Mutated Individuals ( $N_{mutation}$ )	16
Crossover Individuals ( $N_{crossover}$ )	12
Number of terms	4
Range of mutation (rad)	200

Table 4.4.: Starting parameters for the beam size difference fitness method using the Cartesian method

After initializing this method, the GA was able to shape the beam into a new beam profile that we had not previously discovered. The spatial profile of the beam displays the notable asymmetrical, neighboring beam spots, which is indicative of an Airy beam; another type of non-diffracting beam. Fig.4.17 displays an image of the beams longitudinal profile, as well as the initial and final transverse images of the beam.

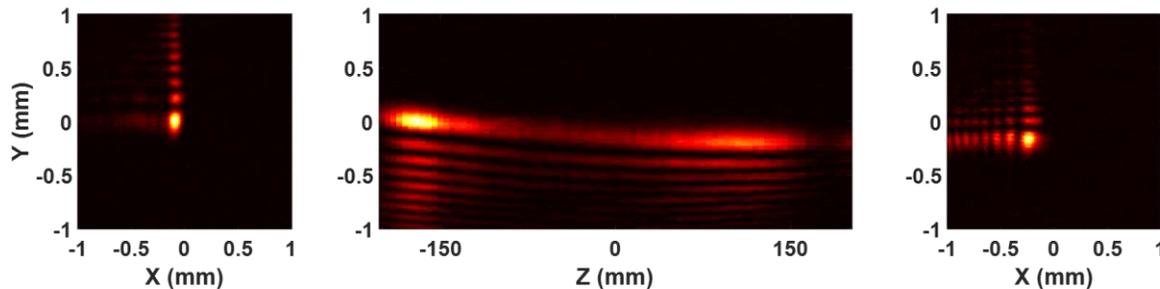


Figure 4.17: 2D longitudinal intensity profile ( $I_{yz}$ ) of the Airy Beam. The transverse image on the left was captured at -200mm (540mm), and the one on the right was captured at +200mm (940mm).

The Airy beam is distinct in that it contains a central beam spot, and translates itself along a curved path, within the transverse plane, as the beam propagates over the optical axis. In addition,

this type of beam has a number of adjacent – but less intense – beam spots surrounding it. A portion of the beam shows an apparent a loss in intensity compared to the outer edges of its longitudinal profile. This is likely the result of the Airy beams inherent translation, as its central beam spot shifts in position to form a parabolic arc. This shift in the central beam spot can draw certain parts of the Airy’s longitudinal profile out of focus when attempting to image. Across all planes, the central beam size (as FWHM) of the airy beam is observed to be 158  $\mu\text{m}$ .

Viewing the 2D SLM phase array shows a noticeable number of phase wrappings, with a distinct cubic function-like phase being detected when a linecut is passed horizontally across its surface; see Fig. 4.18. Supporting scientific literature into Airy beams showcase similar cubic surface phase shapes when generating an Airy beam [152,153]. A significant phase shape is also needed in order to create this beam profile, as a PV of 205.7 radians was generated by the SLM. This development is promising, as the multi-plane GA has now produced two non-diffracting beams; without a prior knowledge.

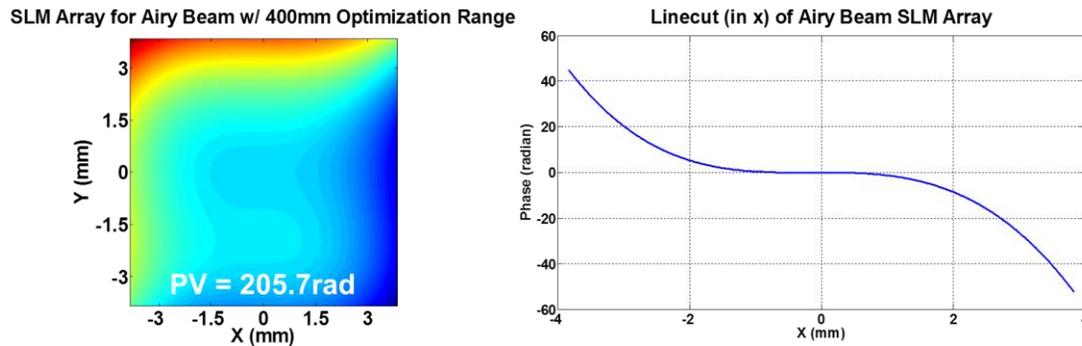


Figure 4.18: SLM phase results after using the beam size difference method. (left) SLM phase array and horizontal linecut of the 400mm volume length Airy beam (right)

#### 4.6.3 Airy beam vs Gaussian beam depth of focus comparison

After successfully generating an Airy beam, it is important to ascertain how significant an increase in the depth of focus was gained through the multi-plane GA. As the Airy beam translates

itself in both x and y directions, as the beam evolves over the longitudinal axis, this can lead to difficulties in determining an appropriate longitudinal range in which to evaluate its depth of focus. For this, it is more desirable to obtain information from a region of the beam that is kept in focus for an extended longitudinal distance. The 2D longitudinal intensity profile ( $I_{yz}$ ) shown in Fig. 4.17 is not sufficient enough in obtaining an accurate depth of focus measurement. As a result, the same Airy beam was measured from the XZ-axis, where the beam profile has minimal translation between -170mm to 60mm; see Fig. 4.19.

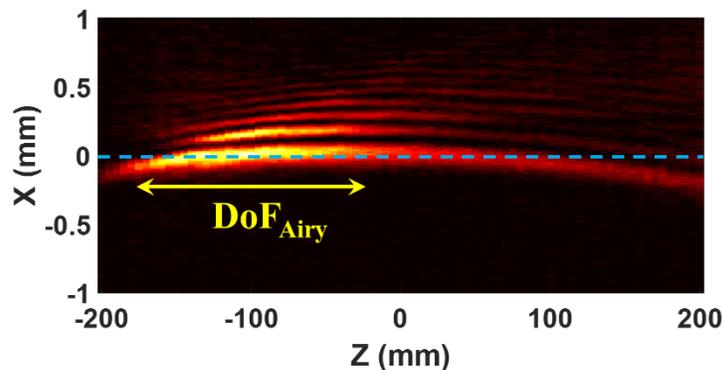


Figure 4.19: Normalized 2D longitudinal intensity profile ( $I_{xz}$ ) of the Airy Beam. The blue line represents the region at which the longitudinal intensity distribution was measured, and the yellow arrows indicate the depth of focus

The Gaussian beam used in section 4.4.3 will – once again – serve as a reference, with its depth of focus being 18.6mm. Fig.4.19 illustrates the depth of focus comparisons between the Airy and Gaussian beam, which are both normalized. For the Airy beam, when taking the region that has a light intensity greater than 50% of the maximum beam intensity (-152mm to -22mm), the depth of focus is determined to be 130mm. This amounts to a 7x increase in the depth of focus compared to that of the Gaussian.

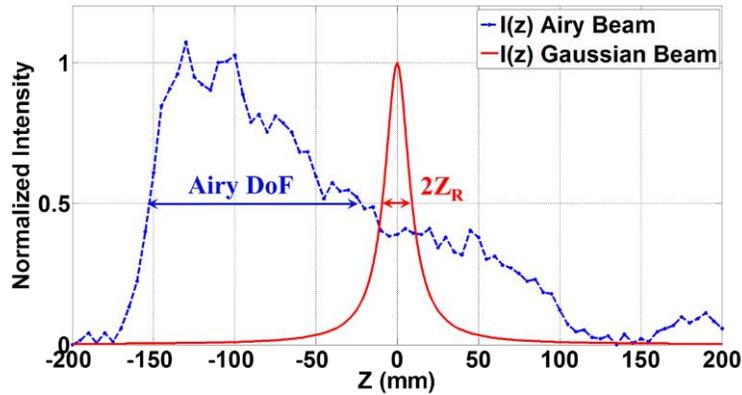


Figure 4.20: 1D longitudinal intensity for the Airy beam (blue dashed line) and Gaussian beam (red solid line)

## 4.7 Discussion

This chapter has demonstrated the ability of the self-learning, multiplane GA setup to generate multiple types of non-diffracting beams; including a Bessel Beam and an Airy beam. In addition, it was also demonstrated how this technique is capable of optimizing each type of beam for a specific focal volume. Altering the optimization range of these Bessel beam has proven to play a key role in its central beam size, as well as its ability to sustain its profile over relatively small, or large, focal volumes. Finally, by analyzing the 2D longitudinal intensity profile for both the Airy and Bessel beam, significant increases (by at least a factor of 7x) in the depth of focus can be seen over their Gaussian beam counterpart; which also translates to a greater focal volume.

Though non-diffracting beams were generated and optimized for various volume lengths, there are a number of hurdles still associated with using this technique in generating new types of non-diffracting beams; or beams with a line focus. As outlined in chapter 2, the GA requires initial parameters (such as the population size, fitness method, and phase methods used) that would allow it to travel around a search space that is sizable enough for exploration, in terms of finding the best solutions for a non-diffracting beam, but also mitigates the potential for finding a local optima; or no solution at all. With respect to the SLM, there are potential limitations on how “fine” a line

focus can be made within a restrictive focal volume. In addition, addressing the issue of divergence will also be necessary in order to further improve the intensity distribution and focal volume extension of the non-diffracting beams this method can produce.

## CHAPTER 5 CONCLUSIONS AND FUTURE WORK

### 5.1 Conclusions

An introduction (Chapter 1) was provided for the reader to understand the fundamental underpinnings behind wavefronts, adaptive optics, wavefront analysis, and the concept of aberrations; and subsequent correction. These topics served as a prelude to experiments conducted in chapters 2, 3, and 4, where the end goal was to produce non-diffracting beams without a prior knowledge of the wavefront needed to generate such a beam. To meet this goal, a learning algorithm, and specifically a Genetic Algorithm (GA), was designed and implemented for finding this wavefront, rather than relying on a wavefront sensor. Chapter 2 tested the validity of this learning algorithm by seeking to use the GA for the purpose of wavefront correction of aberrations in a Gaussian beam. By having the GA seek to maximize the intensity of the Gaussian beam, the intensity of the beam was significantly increased (by a factor of 2x) in the case of correcting the initial SLM distortion, as well as extracting a central beam spot after correcting for higher-order aberrations from the plastic coverslip.

Chapter 3 introduced the concept of the 3D focal intensity distribution of a beam and presented a method to characterize it, with the goal of optimizing it using the learning algorithm. As the focal volume occupies multiple transverse planes along the optical axis of the beam, it was necessary to have the GA receive feedback from multiple transverse images at these various planes. A defocus method was proposed and successfully demonstrated in chapter 3, which showcased the ability for the SLM to shift the focus of the beam so that the CCD camera can capture spatial profiles at different transverse planes; without needing to be translated. This served as the backbone for experiments conducted in chapter 4, whose goal were to generate non-diffracting beams (or beams with a line focus). By obtaining multiple transverse images, the GA

would then be able to act upon the 3D focal intensity distribution of the beam, in a method called **multi-plane GA**. Several fitness and phase methods were implemented and experimented upon (in chapter 4) to produce a beam with an extended focal volume. Using two new fitness methods, the spatial profile similitude and beam size difference fitness methods, both a Bessel beam and an Airy beam were generated; which are two well-known non-diffracting beams. These beams were able to achieve extended depths of focus (and subsequently - focal volumes) of up to 30x that of their Gaussian beam counterpart. It was also shown that by having the multi-plane GA operate over an extended optimization range, the central beam size and focal volume of the Bessel beam can be altered as a result.

## 5.2 Future Work

There are applications for the non-diffracting beams generated in chapter 4, which is the creation of thin light-sheets for LSFM (light-sheet fluorescent microscopy). When making light sheets, the ideal beam type would be a beam capable of the same non-diffracting qualities of the Bessel beam and Airy beam, but with minimal fringes or side lobes that accompany the central beam spot. For consideration, the theoretical “best method” for finding a new type of non-diffracting beam – to fit this need - would be to use a randomized phase method. Under this condition, the learning algorithm (and subsequently – the SLM) would be given the highest degree of freedom in shaping the beam to generate a non-diffracting beam with minimal side-lobes. Such a method would be able to find a new solution than any predefined phase method used within this dissertation.

The checkerboard phase method used section 2.6 is one such example. However, such a method also comes with notable drawbacks, as a purely random phase requires a significantly larger search space, leading to increased computational time. As was detailed in section 2.6, there

is a significantly large search space the GA could operate upon if every pixel in the SLM is manipulated independently ( $65535^{262,144}$ ). Other drawbacks to consider with this method is that, if the checker size is too low (4x4 for example), it could lead to less flexibility in the phase surface for the SLM to apply onto the beam; reducing the quality of the targeted GA solution. Too large of a checker size (ex. 64x64 pixel) would significantly increase the time needed to apply each random value to every checker, slowing down the GA.

Further work would be to compartmentalize the setup proposed in chapter 4, and have it implemented within a LSFM for in-situ beam shaping. As was demonstrated, the multi-plane GA has the potential for generating non-diffracting beams for the purposes of generating light-sheets. These beams have significantly extended focal volumes, allowing for greater imaging depth within the target sample. The extension of this volume – through the GA - should also be accompanied by a reduction in the transverse area of the beam, allowing for a central beam spot that is smaller than that of a Gaussian beam. This reduction in the beam spot would improve the resolution of the image, as the non-diffracting beam would be able to illuminate smaller elements of the sample. Finally, these non-diffracting beams require that their energy to be evenly distributed across the entirety of its depth of focus. As opposed to a Gaussian beam, which concentrates much of its intensity onto a finite focal spot, the larger intensity distribution of the non-diffracting beams enables the reduction of phototoxicity in live samples.

## REFERENCES

- [1] Lindsay, P. H., & Norman, D. A. (2013). *Human information processing: An introduction to psychology*. Academic press.
- [2] Riggs, L. A. (1965). Visual acuity. *Vision and visual perception*, 321-349.
- [3] Cline, D., Hofstetter, H. W., & Griffin, J. R. (Eds.). (1997). *Dictionary of visual science*. Butterworth-Heinemann.
- [4] Kam, Z., Agard, D. A., & Sedat, J. W. (1997). Three-dimensional microscopy in thick biological samples: A fresh approach for adjusting focus and correcting spherical aberration. *Bioimaging*, 5(1), 40-49.
- [5] Neil, M. A., Juškaitis, R., Booth, M. J., Wilson, T., Tanaka, T., & Kawata, S. (2000). Adaptive aberration correction in a two-photon microscope. *Journal of microscopy*, 200(2), 105-108.
- [6] Marsh, P. N., Burns, D., & Girkin, J. M. (2003). Practical implementation of adaptive optics in multiphoton microscopy. *Optics express*, 11(10), 1123-1130.
- [7] Rueckel, M., Mack-Bucher, J. A., & Denk, W. (2006). Adaptive wavefront correction in two-photon microscopy using coherence-gated wavefront sensing. *Proceedings of the National Academy of Sciences*, 103(46), 17137-17142.
- [8] Débarre, D., Botcherby, E. J., Watanabe, T., Srinivas, S., Booth, M. J., & Wilson, T. (2009). Image-based adaptive optics for two-photon microscopy. *Optics letters*, 34(16), 2495-2497.
- [9] Ji, N., Milkie, D. E., & Betzig, E. (2010). Adaptive optics via pupil segmentation for high-resolution imaging in biological tissues. *Nature methods*, 7(2), 141.
- [10] Wang, K., Milkie, D. E., Saxena, A., Engerer, P., Misgeld, T., Bronner, M. E., ... & Betzig, E. (2014). Rapid adaptive optical recovery of optimal resolution over large volumes. *nature methods*, 11(6), 625.
- [11] Beckers, J. M. (1993). Adaptive optics for astronomy: principles, performance, and applications. *Annual review of astronomy and astrophysics*, 31(1), 13-62.
- [12] Vainshtein, L. A. (1988). Electromagnetic waves. *Moscow Izdatel Radio Sviaz*.
- [13] Landau, L. D. (Ed.). (2013). *The classical theory of fields* (Vol. 2). Elsevier.
- [14] Brugsch, H. (1881). A History of Egypt Under the Pharaohs: Derived Entirely from the Monuments, to which is Added a Discourse on the Exodus of the Israelites (Vol. 2). J. Murray. H. D. Seymour, & P. Smith (Eds.). J. Murray.
- [15] Morrish, A. H. (2001). The physical principles of magnetism. The Physical Principles of Magnetism, by Allan H. Morrish, pp. 696. ISBN 0-7803-6029-X. Wiley-VCH, January 2001., 696.
- [16] Reitz, J. R., Milford, F. J., & Christy, R. W. (2008). *Foundations of electromagnetic theory*. Addison-Wesley Publishing Company.
- [17] Maxwell, J. C. (1861). Xxv. on physical lines of force: Part i.—the theory of molecular vortices applied to magnetic phenomena. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 21(139), 161-175.

- [18] Maxwell, J. C. (1865). A dynamical theory of the electromagnetic field. *Philosophical transactions of the Royal Society of London*, 155, 459-512.
- [19] Hunt, B. J. (1991). Michael Faraday, cable telegraphy and the rise of field theory. *History of Technology*, 13, 1-19.
- [20] Hertz, H., & Kelvin, W. T. B. (1893). *Electric waves* (p. 177). London: Macmillan.
- [21] Rashed, R. (1990). A pioneer in anaclastics: Ibn Sahl on burning mirrors and lenses. *Isis*, 81(3), 464-491.
- [22] Wolf, K. B., & Krotzsch, G. (1995). Geometry and dynamics in refracting systems. *European Journal of Physics*, 16(1), 14.
- [23] Born, M., & Wolf, E. (1980). Principles of optics: electromagnetic theory of propagation, interference and diffraction of light. Elsevier.
- [24] Porter, J., Queener, H., Lin, J., Thorn, K., & Awwal, A. A. (2006). *Adaptive optics for vision science: principles, practices, design and applications* (Vol. 171). John Wiley & Sons.
- [25] Bekefi, G., & Barrett, A. H. (1977). *Electromagnetic vibrations, waves, and radiation*. Mit Press.
- [26] Svelto, O., & Hanna, D. C. (1998). *Principles of lasers* (Vol. 4). New York: Plenum press.
- [27] Montgomery, A. J. (1969). Remote sensing of winds and atmospheric turbulence by cross-correlation of passive optical signals. *Space Science Reviews*, 10(2), 291-313.
- [28] National Aeronautics and Space Administration (NASA). *Remote Sensing: Absorption Bands and Atmospheric Windows*. Retrieved May 07, 2017, from [https://earthobservatory.nasa.gov/Features/RemoteSensing/remote\\_04.php](https://earthobservatory.nasa.gov/Features/RemoteSensing/remote_04.php)
- [29] Fadili, J. M., & Starck, J. L. (2006). Sparse representation-based image deconvolution by iterative thresholding. In *Astronomical Data Analysis ADA'06*.
- [30] Fried, D. L. (1966). Optical resolution through a randomly inhomogeneous medium for very long and very short exposures. *JOSA*, 56(10), 1372-1379.
- [31] Kibrick, R., & Robinson, L. (1987). The Lick observatory TV autoguider. *Publications of the Astronomical Society of the Pacific*, 99(619), 1014.
- [32] Bouchy, F., & Connes, P. (1999). Autoguider locked on a fiber input for precision stellar radial velocities. *Astronomy and Astrophysics Supplement Series*, 136(1), 193-204.
- [33] Hardy, J. W. (1998). *Adaptive optics for astronomical telescopes* (Vol. 16). Oxford University Press on Demand.
- [34] Sponable, E. I. (1953). Eidophor system of theater television. *Journal of the Society of Motion Picture and Television Engineers*, 60(4), 337-343.
- [35] Babcock, H. W. (1953). The possibility of compensating astronomical seeing. *Publications of the Astronomical Society of the Pacific*, 65(386), 229-236.
- [36] Kitsopanidou, K. (2003). The widescreen revolution and 20th Century-Fox's Eidophor in the 1950s. *Film History*, 15(1), 32-56.

- [37] Muller, R. A., & Buffington, A. (1974). Real-time correction of atmospherically degraded telescope images through image sharpening. *JOSA*, *64*(9), 1200-1210.
- [38] Fugate, R. Q., & Fried, D. L. (1991). Measurement of atmospheric wavefront distortion using scattered light from a laser guide-star. *Nature*, *353*(6340), 144.
- [39] Kellerer, A., Vidal, F., Gendron, E., Hubert, Z., Perret, D., & Rousset, G. (2012, September). Deformable mirrors for open-loop adaptive optics. In *Adaptive Optics Systems III* (Vol. 8447, p. 844765). International Society for Optics and Photonics.
- [40] Geary, J. M. (1995). *Introduction to wavefront sensors* (Vol. 18). Spie Press.
- [41] Liang, J., Williams, D. R., & Miller, D. T. (1997). Supernormal vision and high-resolution retinal imaging through adaptive optics. *JOSA A*, *14*(11), 2884-2892.
- [42] Bifano, T. G., Mali, R. K., Dorton, J. K., Perreault, J. A., Vandelli, N., Horenstein, M. N., & Castanon, D. A. (1997). Continuous-membrane surface-micromachined silicon deformable mirror. *Optical Engineering*, *36*(5), 1354-1361.
- [43] Hornbeck, L. J. (1990, May). Deformable-mirror spatial light modulators. In *33rd Annual Technical Symposium* (pp. 86-103). International Society for Optics and Photonics.
- [44] Wilbur, J. L., Jackman, R. J., Whitesides, G. M., Cheung, E. L., Lee, L. K., & Prentiss, M. G. (1996). Elastomeric optics. *Chemistry of materials*, *8*(7), 1380-1385.
- [45] Jaecklin, V. P., Linder, C., Brugger, J., De Rooij, N. F., Moret, J. M., & Vuilleumier, R. (1994). Mechanical and optical properties of surface micromachined torsional mirrors in silicon, polysilicon and aluminum. *Sensors and Actuators A: Physical*, *43*(1-3), 269-275.
- [46] Zhang, X. M., Chau, F. S., Quan, C., Lam, Y. L., & Liu, A. Q. (2001). A study of the static characteristics of a torsional micromirror. *Sensors and Actuators A: Physical*, *90*(1), 73-81.
- [47] Huang, J. M., Liu, A. Q., Deng, Z. L., Zhang, Q. X., Ahn, J., & Asundi, A. (2004). An approach to the coupling effect between torsion and bending for electrostatic torsional micromirrors. *Sensors and Actuators A: Physical*, *115*(1), 159-167.
- [48] Zhu, L., Sun, P. C., Bartsch, D. U., Freeman, W. R., & Fainman, Y. (1999). Adaptive control of a micromachined continuous-membrane deformable mirror for aberration compensation. *Applied Optics*, *38*(1), 168-176.
- [49] Bifano, T. G., Perreault, J., Mali, R. K., & Horenstein, M. N. (1999). Microelectromechanical deformable mirrors. *IEEE Journal of Selected Topics in Quantum Electronics*, *5*(1), 83-89.
- [50] Hishinuma, Y., & Yang, E. H. (2006). Piezoelectric unimorph microactuator arrays for single-crystal silicon continuous-membrane deformable mirror. *Journal of microelectromechanical systems*, *15*(2), 370-379.
- [51] Cornelissen, S. A., Bierden, P. A., Bifano, T. G., & Lam, C. V. (2009). 4096-element continuous face-sheet MEMS deformable mirror for high-contrast imaging. *Journal of Micro/Nanolithography, MEMS, and MOEMS*, *8*(3), 031308-031308.
- [52] Collins, D. R., Sampsel, J. B., Hornbeck, L. J., Florence, J. M., Penz, P. A., & Gately, M. T. (1989). Deformable mirror device spatial light modulators and their applicability to optical neural networks. *Applied optics*, *28*(22), 4900-4907.

- [53] Fernández, E. J., Vabre, L., Hermann, B., Unterhuber, A., Považay, B., & Drexler, W. (2006). Adaptive optics with a magnetic deformable mirror: applications in the human eye. *Optics Express*, 14(20), 8900-8917.
- [54] Meadowlark Optics. (2019). *Small 512x512 Spatial Light Modulator*. Retrieved from <https://www.meadowlark.com/product-family-details.php?pid=139>.
- [55] Malbet, F., Yu, J. W., & Shao, M. (1995). High-dynamic-range imaging using a deformable mirror for space coronagraphy. *Publications of the Astronomical Society of the Pacific*, 107(710), 386.
- [56] Beckers, J. M. (1988). Increasing the size of the isoplanatic patch with multiconjugate adaptive optics. In *European Southern Observatory Conference and Workshop Proceedings* (Vol. 30, p. 693).
- [57] Rigaut, F. J., Ellerbroek, B. L., & Flicker, R. (2000, July). Principles, limitations, and performance of multiconjugate adaptive optics. In *Adaptive Optical Systems Technology* (Vol. 4007, pp. 1022-1032). International Society for Optics and Photonics.
- [58] Berkefeld, T., Glindemann, A., & Hippler, S. (2001). Multi-Conjugate Adaptive Optics with Two Deformable Mirrors—Requirements and Performance. *Experimental Astronomy*, 11(1), 1-21.
- [59] Ragazzoni, R., Le Roux, B., & Arcidiacono, C. (2005). Multi-conjugate adaptive optics for elts: constraints and limitations. *Comptes Rendus Physique*, 6(10), 1081-1088.
- [60] Langlois, M., Moretto, G., Richards, K., Hegwer, S., & Rimmele, T. R. (2004, October). Solar multiconjugate adaptive optics at the Dunn Solar Telescope: preliminary results. In *Advancements in Adaptive Optics* (Vol. 5490, pp. 59-67). International Society for Optics and Photonics.
- [61] Marchetti, E., Brast, R., Delabre, B., Donaldson, R., Fedrigo, E., Frank, C., ... & Oberti, S. (2007). On-sky testing of the multi-conjugate adaptive optics demonstrator. *The Messenger*, 129(8).
- [62] Hammel H., Space Science Institute, University of California, Berkeley. (2004, Nov.). *Keck pictures of Uranus show best view from the ground*. Retrieved from [http://www.keckobservatory.org/keck\\_pictures\\_of\\_uranus\\_show\\_best\\_view\\_from\\_the\\_ground/](http://www.keckobservatory.org/keck_pictures_of_uranus_show_best_view_from_the_ground/).
- [63] Rouan, D., Riaud, P., Boccaletti, A., Clénet, Y., & Labeyrie, A. (2000). The four-quadrant phase-mask coronagraph. I. Principle. *Publications of the Astronomical Society of the Pacific*, 112(777), 1479.
- [64] Sivaramakrishnan, A., Koresko, C. D., Makidon, R. B., Berkefeld, T., & Kuchner, M. J. (2001). Ground-based coronagraphy with high-order adaptive optics. *The astrophysical journal*, 552(1), 397.
- [65] Marois, C., Macintosh, B., Barman, T., Zuckerman, B., Song, I., Patience, J., ... & Doyon, R. (2008). Direct imaging of multiple planets orbiting the star HR 8799. *Science*, 322(5906), 1348-1352.
- [66] Ghez, A. M., Salim, S., Weinberg, N. N., Lu, J. R., Do, T., Dunn, J. K., ... & Kremenek, T. (2008). Measuring distance and properties of the Milky Way's central supermassive black hole with stellar orbits. *The Astrophysical Journal*, 689(2), 1044.

- [67] Genzel, R., Burkert, A., Bouché, N., Cresci, G., Schreiber, N. F., Shapley, A., ... & Daddi, E. (2008). From rings to bulges: evidence for rapid secular galaxy evolution at  $z \sim 2$  from integral field spectroscopy in the SINS survey. *The Astrophysical Journal*, 687(1), 59.
- [68] Krist, J. E., Ardila, D. R., Golimowski, D. A., Clampin, M., Ford, H. C., Illingworth, G. D., ... & Bouwens, R. J. (2005). Hubble Space Telescope Advanced Camera for Surveys coronagraphic imaging of the AU Microscopii debris disk. *The Astronomical Journal*, 129(2), 1008.
- [69] Curtis, J. E., Koss, B. A., & Grier, D. G. (2002). Dynamic holographic optical tweezers. *Optics communications*, 207(1-6), 169-175.
- [70] Gibson, G. M., Leach, J., Keen, S., Wright, A. J., & Padgett, M. J. (2008). Measuring the accuracy of particle position and force in optical tweezers using high-speed video microscopy. *Optics Express*, 16(19), 14561-14570.
- [71] Theofanidou, E., Wilson, L., Hossack, W. J., & Arlt, J. (2004). Spherical aberration correction for optical tweezers. *Optics Communications*, 236(1-3), 145-150.
- [72] Booth, M. J. (2007). Adaptive optics in microscopy. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 365(1861), 2829-2843.
- [73] Nwaneshiudu, A., Kuschal, C., Sakamoto, F. H., Anderson, R. R., Schwarzenberger, K., & Young, R. C. (2012). Introduction to confocal microscopy. *The Journal of investigative dermatology*, 132(12), e3.
- [74] Smirnov, M. S. (1961). Measurement of the wave aberration of the human eye. *Biofizika*, 6, 687.
- [75] Howland, H. C., & Howland, B. (1977). A subjective method for the measurement of monochromatic\* aberrations of the eye. *JOSA*, 67(11), 1508-1518.
- [76] Walsh, G., Howland, H. C., & Charman, W. N. (1984). Objective technique for the determination of monochromatic aberrations of the human eye. *JOSA A*, 1(9), 987-992.
- [77] Fernández, E. J., Iglesias, I., & Artal, P. (2001). Closed-loop adaptive optics in the human eye. *Optics letters*, 26(10), 746-748.
- [78] Doble, N., Yoon, G., Chen, L., Bierden, P., Singer, B., Olivier, S., & Williams, D. R. (2002). Use of a microelectromechanical mirror for adaptive optics in the human eye. *Optics letters*, 27(17), 1537-1539.
- [79] Carroll, J., Neitz, M., Hofer, H., Neitz, J., & Williams, D. R. (2004). Functional photoreceptor loss revealed with adaptive optics: an alternate cause of color blindness. *Proceedings of the National Academy of Sciences of the United States of America*, 101(22), 8461-8466.
- [80] Prieto, P. M., Fernández, E. J., Manzanera, S., & Artal, P. (2004). Adaptive optics with a programmable phase modulator: applications in the human eye. *Optics Express*, 12(17), 4059-4071.
- [81] El-Agmy, R., Bulte, H., Greenaway, A. H., & Reid, D. T. (2005). Adaptive beam profile control using a simulated annealing algorithm. *Optics Express*, 13(16), 6085-6091.

- [82] Ma, H., Zhou, P., Wang, X., Ma, Y., Xi, F., Xu, X., & Liu, Z. (2010). Near-diffraction-limited annular flattop beam shaping with dual phase only liquid crystal spatial light modulators. *Optics express*, 18(8), 8251-8260.
- [83] Gutiérrez-Vega, J. C., Iturbe-Castillo, M. D., Ramirez, G. A., Tepichin, E., Rodríguez-Dagnino, R. M., Chávez-Cerda, S., & New, G. H. C. (2001). Experimental demonstration of optical Mathieu beams. *Optics Communications*, 195(1), 35-40.
- [84] Dartora, C. A., Zamboni-Rached, M., Nobrega, K. Z., Recami, E., & Hernández-Figueroa, H. E. (2003). General formulation for the analysis of scalar diffraction-free beams using angular modulation: Mathieu and Bessel beams. *Optics Communications*, 222(1), 75-80.
- [85] Fahrbach, F. O., Simon, P., & Rohrbach, A. (2010). Microscopy with self-reconstructing beams. *Nature Photonics*, 4(11), 780-785.
- [86] Planchon, T. A., Gao, L., Milkie, D. E., Davidson, M. W., Galbraith, J. A., Galbraith, C. G., & Betzig, E. (2011). Rapid three-dimensional isotropic imaging of living cells using Bessel beam plane illumination. *Nature methods*, 8(5), 417-423.
- [87] Vellekoop, I. M., & Mosk, A. P. (2007). Focusing coherent light through opaque strongly scattering media. *Optics letters*, 32(16), 2309-2311.
- [88] Thyagarajan, K., & Ghatak, A. (2011). Laser-Induced Fusion. In *Lasers* (pp. 403-415). Springer US.
- [89] Miller, G. H., Moses, E. I., & Wuest, C. R. (2004). The National Ignition Facility: enabling fusion ignition for the 21st century. *Nuclear fusion*, 44(12), S228.
- [90] Fleurot, N., Cavailler, C., & Bourgade, J. L. (2005). The Laser Megajoule (LMJ) Project dedicated to inertial confinement fusion: Development and construction status. *Fusion Engineering and Design*, 74(1), 147-154.
- [91] Zheng, W., Wei, X., Zhu, Q., Jing, F., Hu, D., Su, J., ... & Zhou, W. (2016). Laser performance of the SG-III laser facility. *High Power Laser Science and Engineering*, 4.
- [92] Bahk, S. W., Fess, E., Kruschwitz, B. E., & Zuegel, J. D. (2010). A high-resolution, adaptive beam-shaping system for high-power lasers. *Optics express*, 18(9), 9151-9163.
- [93] Beck, R. J., Parry, J. P., MacPherson, W. N., Waddie, A., Weston, N. J., Shephard, J. D., & Hand, D. P. (2010). Application of cooled spatial light modulator for high power nanosecond laser micromachining. *Optics express*, 18(16), 17059-17065.
- [94] Simmonds, R. D., Salter, P. S., Jesacher, A., & Booth, M. J. (2011). Three-dimensional laser microfabrication in diamond using a dual adaptive optics system. *Optics express*, 19(24), 24122-24128.
- [95] Harry, G. M., & LIGO Scientific Collaboration. (2010). Advanced LIGO: the next generation of gravitational wave detectors. *Classical and Quantum Gravity*, 27(8), 084006.
- [96] Takeda, M., Ina, H., & Kobayashi, S. (1982). Fourier-transform method of fringe-pattern analysis for computer-based topography and interferometry. *JosA*, 72(1), 156-160.
- [97] Pandey, N., Singh, M. P., Ghosh, A., & Khare, K. (2018). Optical surface measurement using accurate carrier estimation in Fourier transform fringe analysis and phase unwrapping based upon transport of intensity equation. *Journal of Optics*, 1-7.

- [98] Maciel, J. M., Rodríguez, F. J. C., González, M. M., Lecona, F. G. P., & Ramírez, V. M. D. (2017). Digital Processing Techniques for Fringe Analysis. In *Optical Interferometry*. InTech.
- [99] Platt, B. C., & Shack, R. (2001). History and principles of Shack-Hartmann wavefront sensing. *Journal of Refractive Surgery*, 17(5), S573-S577.
- [100] Kumler, J. J., & Caldwell, J. B. (2007, September). Measuring surface slope error on precision aspheres. In *Optical Manufacturing and Testing VII* (Vol. 6671, p. 66710U). International Society for Optics and Photonics.
- [101] Neal, D. R., Copland, J., & Neal, D. A. (2002, November). Shack-Hartmann wavefront sensor precision and accuracy. In *Advanced Characterization Techniques for Optical, Semiconductor, and Data Storage Components* (Vol. 4779, pp. 148-161). International Society for Optics and Photonics.
- [102] Southwell, W. H. (1980). Wave-front estimation from wave-front slope measurements. *JOSA*, 70(8), 998-1006.
- [103] Thibos, L. N. (2000, February). Principles of hartmann-shack aberrometry. In *Vision Science and its Applications* (p. NW6). Optical Society of America.
- [104] Thibos, L. N., & Hong, X. (1999). Clinical applications of the Shack-Hartmann aberrometer. *Optometry and vision science: official publication of the American Academy of Optometry*, 76(12), 817-825.
- [105] Maeda, N., Fujikado, T., Kuroda, T., Mihashi, T., Hirohara, Y., Nishida, K., ... & Tano, Y. (2002). Wavefront aberrations measured with Hartmann-Shack sensor in patients with keratoconus. *Ophthalmology*, 109(11), 1996-2003.
- [106] Cheng, X., Himebaugh, N. L., Kollbaum, P. S., Thibos, L. N., & Bradley, A. (2003). Validation of a clinical Shack-Hartmann aberrometer. *Optometry and Vision Science*, 80(8), 587-595.
- [107] Jiang, W., & Li, H. (1990, August). Hartmann-Shack wavefront sensing and wavefront control algorithm. In *Adaptive Optics and Optical Structures* (Vol. 1271, pp. 82-94). International Society for Optics and Photonics.
- [108] Seifert, L., Liesener, J., & Tiziani, H. J. (2003). The adaptive shack-hartmann sensor. *Optics Communications*, 216(4-6), 313-319.
- [109] Nicolle, M., Fusco, T., Rousset, G., & Michau, V. (2004). Improvement of Shack-Hartmann wave-front sensor measurement for extreme adaptive optics. *Optics letters*, 29(23), 2743-2745.
- [110] Zernike, F. (1935). Phase contrast. *Z. Tech. Physik.*, 16, 454.
- [111] Wang, J. Y., & Silva, D. E. (1980). Wave-front interpretation with Zernike polynomials. *Applied optics*, 19(9), 1510-1518.
- [112] Nijboer, B. R. A. (1942). *The diffraction theory of aberrations* (Doctoral dissertation, Wolters).
- [113] Mahajan, V. N. (1994). Zernike circle polynomials and optical aberrations of systems with circular pupils. *Applied optics*, 33(34), 8121-8124.
- [114] Noll, R. J. (1976). Zernike polynomials and atmospheric turbulence. *JOSA*, 66(3), 207-211.

- [115] Ruoff, J., & Totzeck, M. (2009). Orientation Zernike polynomials: a useful way to describe the polarization effects of optical imaging systems. *Journal of Micro/Nanolithography, MEMS, and MOEMS*, 8(3), 031404.
- [116] Planchon, T. A., Amir, W., Field, J. J., Durfee, C. G., Squier, J. A., Rousseau, P., ... & Mourou, G. (2006). Adaptive correction of a tightly focused, high-intensity laser beam by use of a third-harmonic signal generated at an interface. *Optics letters*, 31(14), 2214-2216.
- [117] Mantravadi, M. V., & Malacara, D. (1992). Newton, Fizeau, and Haidinger interferometers. *Optical shop testing*, 59, 1-45.
- [118] Malacara, Z., & Servin, M. (2016). *Interferogram analysis for optical testing*. CRC press.
- [119] Womack, K. (1984). Interferometric phase measurement using spatial synchronous detection. *Optical engineering*, 23(4), 234391.
- [120] Servin, M., & Rodriguez-Vera, R. (1993). Two-dimensional phase locked loop demodulation of interferograms. *Journal of Modern Optics*, 40(11), 2087-2094.
- [121] Hsieh, G. C., & Hung, J. C. (1996). Phase-locked loop techniques. A survey. *IEEE Transactions on industrial electronics*, 43(6), 609-615.
- [122] Kaura, V., & Blasko, V. (1997). Operation of a phase locked loop system under distorted utility conditions. *IEEE Transactions on Industry applications*, 33(1), 58-63.
- [123] Van Loan, C. (1992). Computational frameworks for the fast Fourier transform (Vol. 10). Siam.
- [124] Holland, J. H. (1975). Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence.
- [125] Grant, V. (1985). The evolutionary process: A critical review of evolutionary theory. Columbia University Press.
- [126] Ribeiro Filho, J. L., Treleaven, P. C., & Alippi, C. (1994). Genetic-algorithm programming environments. *Computer*, 27(6), 28-43.
- [127] Kumar, M., Husian, M., Upreti, N., & Gupta, D. (2010). Genetic algorithm: Review and application. *International Journal of Information Technology and Knowledge Management*, 2(2), 451-454.
- [128] Whitley, D. (1994). A genetic algorithm tutorial. *Statistics and computing*, 4(2), 65-85.
- [129] Deb, K., Pratap, A., Agarwal, S., & Meyarivan, T. A. M. T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE transactions on evolutionary computation*, 6(2), 182-197.
- [130] Baker, J. E. (1985, July). Adaptive selection methods for genetic algorithms. In *Proceedings of an International Conference on Genetic Algorithms and their applications* (pp. 101-111).
- [131] Whitley, L. D. (1989, June). The GENITOR Algorithm and Selection Pressure: Why Rank-Based Allocation of Reproductive Trials is Best. In *ICGA* (Vol. 89, pp. 116-123).
- [132] Goldberg, D. E. (1990). A note on Boltzmann tournament selection for genetic algorithms and population-oriented simulated annealing. *Complex Systems*, 4(4), 445-460.

- [133] Houck, C. R., Joines, J., & Kay, M. G. (1995). A genetic algorithm for function optimization: a Matlab implementation. *Ncsu-ie tr*, 95(09), 1-10.
- [134] Kim, H. S., & Cho, S. B. (2001). An efficient genetic algorithm with less fitness evaluation by clustering. In *Evolutionary Computation, 2001. Proceedings of the 2001 Congress on* (Vol. 2, pp. 887-894). IEEE.
- [135] Srinivas, M., & Patnaik, L. M. (1994). Adaptive probabilities of crossover and mutation in genetic algorithms. *IEEE Transactions on Systems, Man, and Cybernetics*, 24(4), 656-667.
- [136] Michalewicz, Z., & Janikow, C. Z. (1991, July). Handling constraints in genetic algorithms. In *ICGA* (pp. 151-157).
- [137] Anderson-Cook, C. M. (2005). Practical genetic algorithms.
- [138] Ingber, L. (1993). Simulated annealing: Practice versus theory. *Mathematical and computer modelling*, 18(11), 29-57.
- [139] Grefenstette, J. J. (1992, September). Genetic algorithms for changing environments. In *PPSN* (Vol. 2, pp. 137-144).
- [140] Kazarlis, S. A., Bakirtzis, A. G., & Petridis, V. (1996). A genetic algorithm solution to the unit commitment problem. *IEEE transactions on power systems*, 11(1), 83-92.
- [141] Milione, G., Dudley, A., Nguyen, T. A., Chakraborty, O., Karimi, E., Forbes, A., & Alfano, R. R. (2015). Measuring the self-healing of the spatially inhomogeneous states of polarization of vector Bessel beams. *Journal of Optics*, 17(3), 035617.
- [142] Durnin, J. (1987). Exact solutions for nondiffracting beams. I. The scalar theory. *JOSA A*, 4(4), 651-654.
- [143] Durnin, J., Miceli Jr, J. J., & Eberly, J. H. (1987). Diffraction-free beams. *Physical review letters*, 58(15), 1499.
- [144] Gutiérrez-Vega, J. C., Iturbe-Castillo, M. D., & Chávez-Cerda, S. (2000). Alternative formulation for invariant optical fields: Mathieu beams. *Optics letters*, 25(20), 1493-1495.
- [145] López-Mariscal, C., Bandres, M. A., Gutiérrez-Vega, J. C., & Chávez-Cerda, S. (2005). Observation of parabolic nondiffracting optical fields. *Optics express*, 13(7), 2364-2369.
- [146] Siviloglou, G. A., & Christodoulides, D. N. (2007). Accelerating finite energy Airy beams. *Optics Letters*, 32(8), 979-981.
- [147] Bandres, M. A., Gutiérrez-Vega, J. C., & Chávez-Cerda, S. (2004). Parabolic nondiffracting optical wave fields. *Optics letters*, 29(1), 44-46.
- [148] Voie, A. H., Burns, D. H., & Spelman, F. A. (1993). Orthogonal-plane fluorescence optical sectioning: Three-dimensional imaging of macroscopic biological specimens. *Journal of microscopy*, 170(3), 229-236.
- [149] Santi, P. A. (2011). Light sheet fluorescence microscopy: a review. *Journal of Histochemistry & Cytochemistry*, 59(2), 129-138.
- [150] Arlt, J., & Dholakia, K. (2000). Generation of high-order Bessel beams by use of an axicon. *Optics Communications*, 177(1-6), 297-301.

- [151] Chattapiban, N., Rogers, E. A., Cofield, D., Hill III, W. T., & Roy, R. (2003). Generation of nondiffracting Bessel beams by use of a spatial light modulator. *Optics letters*, 28(22), 2183-2185.
- [152] Broky, J., Siviloglou, G. A., Dogariu, A., & Christodoulides, D. N. (2008). Self-healing properties of optical Airy beams. *Optics express*, 16(17), 12880-12891.
- [153] Rose, P., Diebel, F., Boguslawski, M., & Denz, C. (2013). Airy beam induced optical routing. *Applied Physics Letters*, 102(10), 101101.
- [154] Neil, M. A., Booth, M. J., & Wilson, T. (2000). New modal wave-front sensor: a theoretical analysis. *JOSA A*, 17(6), 1098-1107.

## APPENDIX A. ZERNIKE POLYNOMIALS

The Zernike Polynomials<sup>[154]</sup> are appealing within the field of optics and ophthalmology thanks in part to the simplicity in their formulation, as well as their ability to categorize various optical aberrations in circular apertures (for example, an eye). Originally constructed for phase contrasting, this mathematical ensemble has been iterated upon for the benefit of optical design and testing. In the context of this dissertation, the Zernike polynomials serve as one of the phase methods for the purposes of wavefront correction. Formulation of the Zernike Polynomials is set within a polar coordinate regime;  $(r, \theta)$ . In this case,  $\rho = r/a$ , so that  $0 \leq \rho \leq 1$ , and  $0 \leq \theta \leq 2\pi$ . Using the Zernike circular polynomials,  $R_n^m(\rho)\cos(m\theta)$  and  $R_n^m(\rho)\sin(m\theta)$ , we can expand upon the wave aberration function  $W(\rho, \theta)$ :

$$W(\rho, \theta) = \sum_{n=0}^{\infty} \sum_{m=0}^n \left[ \frac{2(n+1)}{1+\delta_{m0}} \right]^{\frac{1}{2}} R_n^m(\rho) \times (c_{nm} \cos(m\theta) + s_{nm} \sin(m\theta))$$

where  $c_{nm}$  and  $s_{nm}$  are the aberration coefficients,  $n$  is the radial degree,  $m$  is the azimuthal frequency (with  $n - m \geq 0$ ), and  $\delta_{ij}$  is the Kronecker delta. The values for  $n$  and  $m$  are determined by the following:

$$n = \text{roundup} \left( \frac{-3\sqrt{9+8j}}{2} \right)$$

$$m = 2j - n(n+2)$$

Here,  $j$  is the Zernike Mode number, and defines each specific Zernike aberration. The Zernike mode number is defined as:

$$j = \frac{n(n+2) + m}{2}$$

For the circular polynomials:

$$R_n^m(\rho) = \sum_{s=0}^{(n-m)/2} \frac{(-1)^s (n-s)!}{s! \left(\frac{n+m}{2} - s\right)! \left(\frac{n-m}{2} - s\right)!} \rho^{n-2s}$$

in which  $n$  determines the polynomial degree. The radial polynomials are either even or odd:

$$R_n^m = \begin{cases} \delta_{m0}, & \frac{n}{2} \text{ even} \\ -\delta_{m0}, & \frac{n}{2} \text{ odd} \end{cases}; \quad R_n^m(1) = 1, \quad R_n^n(\rho) = \rho^n$$

It is important to recognize the orthogonal nature of the radial polynomials, in which their integral forms can be reduced into the following:

$$\int_0^1 R_n^m(\rho) R_n^{m'}(\rho) \rho d\rho = \frac{1}{2(n+1)} \delta_{nm'}$$

$$\int_0^{2\pi} \cos(m\theta) \sin(m'\theta) d\theta = 0$$

$$\int_0^{2\pi} \sin(m\theta) \sin(m'\theta) d\theta = \pi \delta_{mm'}$$

The aberration function, as orthonormal Zernike circular polynomials  $Z_j(\rho, \theta)$ , can be represented as:

$$W(\rho, \theta) = \sum_{j=1}^{\infty} a_j Z_j(\rho, \theta)$$

Incorporating the radial polynomials leads to the following Zernike mode expressions:

$$Z_{evenj}(\rho, \theta) = \sqrt{2(n+1)}R_n^m(\rho) \cos(m\theta), \quad m \neq 0,$$

$$Z_{oddj}(\rho, \theta) = \sqrt{2(n+1)}R_n^m(\rho) \sin(m\theta), \quad m \neq 0,$$

$$Z_j(\rho, \theta) = \sqrt{n+1}R_n^0(\rho), \quad m = 0,$$

Both Zernike and orthogonal aberrations consist of one or more classical aberrations; including spherical, astigmatism, coma, and tilt aberrations. Referencing these polynomials is typically performed by choosing a Zernike mode ( $Z_j$ ). As an example, a Zernike mode of  $Z_4$  corresponds to astigmatism along the x direction. In this case, the radial polynomial  $R_n^m(\rho)$  for astigmatism in the x direction can be represented as  $R_2^2(\rho)$ ; as well as  $\rho^2$ . Fig. A.1 lists the numerous Zernike polynomials, their aberration type, as well as the order of said aberration.

$i$	$n$	$m$	$Z_i(r, \theta)$	Aberration Term
1	0	0	1	Piston
2	1	1	$2r \cos(\theta)$	Tip
3	1	-1	$2r \sin(\theta)$	Tilt
4	2	0	$\sqrt{3}(2r^2 - 1)$	Defocus
5	2	2	$2\sqrt{3}r^2 \cos(2\theta)$	Astigmatism
6	2	-2	$2\sqrt{3}r^2 \sin(2\theta)$	Astigmatism
7	3	1	$2\sqrt{2}(3r^3 - 2r)\cos(\theta)$	Coma
8	3	-1	$2\sqrt{2}(3r^3 - 2r)\sin(\theta)$	Coma
9	3	3	$2\sqrt{2}r^3 \cos(3\theta)$	
10	3	-3	$2\sqrt{2}r^3 \sin(3\theta)$	
11	4	0	$\sqrt{5}(6r^4 - 6r^2 + 1)$	Spherical (1st)
12	4	2	$\sqrt{10}(4r^4 - 3r^2)\cos(2\theta)$	
13	4	-2	$\sqrt{10}(4r^4 - 3r^2)\sin(2\theta)$	
14	4	4	$\sqrt{10}r^4 \cos(4\theta)$	
15	4	-4	$\sqrt{10}r^4 \sin(4\theta)$	
16	5	1	$2\sqrt{3}(10r^5 - 12r^3 + 3r)\cos(\theta)$	
17	5	-1	$2\sqrt{3}(10r^5 - 12r^3 + 3r)\sin(\theta)$	
18	5	3	$2\sqrt{3}(5r^5 - 4r^3)\cos(3\theta)$	
19	5	-3	$2\sqrt{3}(5r^5 - 4r^3)\sin(3\theta)$	
20	5	5	$2\sqrt{3}r^5 \cos(5\theta)$	
21	5	-5	$2\sqrt{3}r^5 \sin(5\theta)$	
22	6	0	$\sqrt{7}(20r^6 - 30r^4 + 12r^2 - 1)$	Spherical (2nd)

Figure A.1: List of Zernike Polynomial equations.

## APPENDIX B. SLM CALIBRATION

Proper operation of the Spatial Light Modulator (SLM) assumes that the grey levels (a representation of the amount of voltage in one pixel) that are applied to the SLM pixels achieve a desired phase response. In addition, non-uniform or partial illumination of the SLM requires the location of the beam center and the beam size. By observing the behavior of the beams intensity, and modulating the phase in a periodic fashion, it is possible to derive the phase response of the SLM. From here, the Look-Up Table (LUT) for the SLM can be generated. In a general sense, LUT files are responsible for indexing and associating computational values to values found in a memory array. The ability for a computer to retrieve these associative values from memory significantly reduces the amount of processing time needed to find a particular value, and makes using such a method beneficial in run-time operations. A notable example of this is in the realm of image processing. Input values represented in unsigned, 16-bit format (0-65535) – for example – can be transformed into Red, Blue, Green (RGB) color format by associating the input values with those of the LUT file. In the case of adaptive optics, the LUT file is implemented in a manner so as to linearize the phase response of the liquid crystals from 0 to  $2\pi$ .

Specific to the SLM device used in this dissertation, the LUT file is represented as a text file with two columns. The SLM only accepts (and applies) unsigned, 16-bit values, which is represented from 0 to 65535. Column 1 is indicative of the input values sent to the SLM PCIe controller, while the second column is the associated mapped values. This column will increment, non-linearly, to a corresponding value that will result in a  $2\pi$  phase shift; up to 65535. SLM phase calibration is typically handled via interferometric methods. However, as using this methodology is fairly tedious, an alternate calibration method is utilized. Calibration of the SLM in this regard will utilize a diffraction method, in which a checkerboard phase pattern will be sent to – and

applied by – the SLM. When a coherent light source impacts an SLM located at the front focal plane of a lens, the Fourier Transform (FT) of the intensity pattern at the back focal plane is indicative of the phase pattern currently applied to the SLM. By changing the phase contrast of certain checkers, the amount of light diffracted from the zero<sup>th</sup> order power also changes. By mapping the gray levels as a function of the diffraction level, it is possible to determine the corresponding phase contrast values; which can be implemented into the LUT file.

Calibration is set within a Michelson interferometer. Fig. B.1 illustrates the design of the setup. A 488nm continuous wave (cw) laser passes through a series of focusing optics, followed by a half-wave plate. The beam then impinges upon a non-polarizing beam splitter, diverting the beam to both the SLM and the dielectric mirror. The SLM is of 512x512 pixel resolution (with 15 $\mu$ m pixel size), and is situated as the reference mirror within the interferometer.

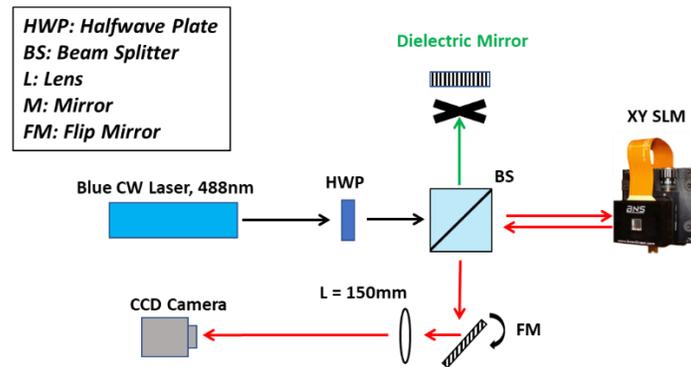


Figure B.1: Experimental setup used for diffraction-based calibration of the SLM

The SLM is implemented into a LabVIEW user interface for the purposes of driving the checker pattern. Light from the non-reference mirror is blocked off, as it is unnecessary. Once the beam from the SLM arm returns to the beam splitter, it then proceeds to impact a flip-mirror, which reflects the light towards a 150mm focusing lens. A 12-bit USB3.0 camera is situated at the focal point of the lens, and is used for image capture of the diffraction pattern. The checkerboard pattern

sent to the SLM (for diffraction) will consist of a 4x4 checkerboard, with each checker containing 128 pixels (see Fig. B.2).

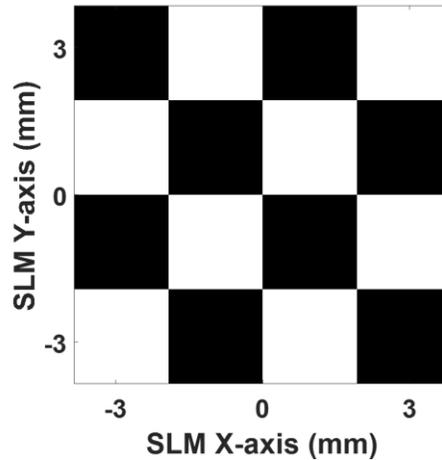


Figure B.2: 4x4 Checker pattern applied by the SLM for diffraction calibration

During testing, the white checkers were maintained at the maximum gray level (65535), while the black checkers start initially at 0. For the black checkers, the LabVIEW program used to apply the checkerboard onto the SLM would ramp the gray levels of the black checkers by increments of 128 gray levels; allowing for 512 measurable data points. Monitoring the intensity of the 0<sup>th</sup> order diffraction pattern with a cropped window (30x30 pixel), when altering the phase shift between the white and black checkers, allows for the applied grayscale to be matched with the output phase shift. Energy loss, due to diffraction, is at its highest when the phase contrast is equal to  $\pi$ , which occurs at an SLM gray level of 32,986. An example of the diffraction pattern produced when a phase of  $\pi$  is applied can be seen in Fig. B.3.

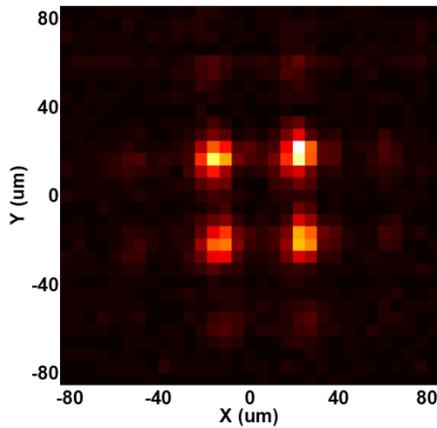


Figure B.3: Beam intensity when a checkerboard pattern of  $\pi$ -phase is applied to the SLM

A plot of the measured CCD intensity, as function of the applied SLM gray level, can be seen in Fig. B.4. A 10<sup>th</sup> degree polynomial fit is overlaid on top of the data points to generate a smoother intensity distribution; as function of gray level.

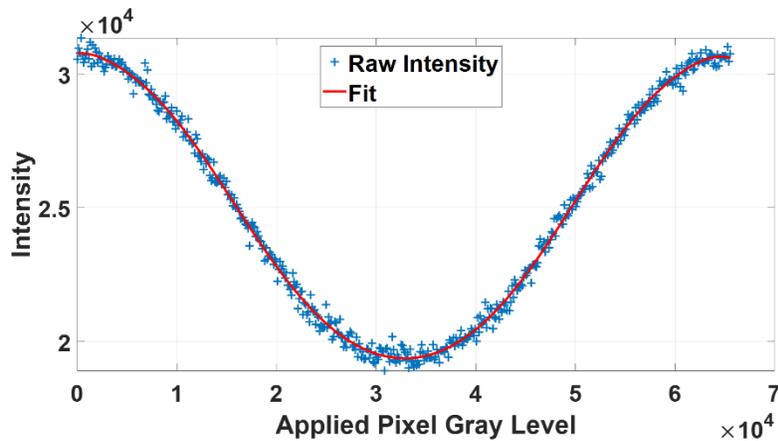


Figure B.4: Intensity of the 0<sup>th</sup> order diffraction as function of gray level. The intensity changes as the phase of the checkerboard pattern is increased from 0 to  $2\pi$ , with a  $\pi$  phase bringing the 0<sup>th</sup> order intensity to a minimum

Intensity measurements are normalized from 0 to 1, and normalization has to be applied separately for phase range of 0 to  $\pi$  and  $\pi$  to  $2\pi$ . This is necessary for the interpolation procedure that is needed to fill the LUT file towards the end of the calibration process; which needs single-valued data points. This also underpins the fact that there are not perfect boundaries between

neighboring pixels in the SLM, due to the liquid crystal molecules not occupying discrete positions above the pixels. After normalization, it is possible to match the intensity value to the phase (in radians) using Eqn. B.1 for the phase range of 0 to  $\pi$ , and Eqn. B.2 for the phase range of  $\pi$  to  $2\pi$ .

$$(\text{Phase} > \pi) \quad 2 \operatorname{acos}(\sqrt{\text{Intensity}}) \quad (\text{B.1})$$

$$(\text{Phase} < \pi) \quad 2[\pi - \operatorname{acos} \sqrt{\text{Intensity}}] \quad (\text{B.2})$$

The phase in radians can be subsequently converted to phase in waves, which merely requires dividing the values found in radians by  $2\pi$ . This allows for the phase values to track from 0 to 1; when decrementing from 65535 to 0. Once in waves, the phase can finally be converted to gray levels using Eqn. B.3. This same procedure is once again applied separately for the phase range 0 to  $\pi$ , and  $\pi$  to  $2\pi$ .

$$65535 - [65535 * (\text{Phase in Waves})] \quad (\text{B.3})$$

Once the gray level values have been determined, the gray levels found for both phase ranges can be placed within a 1D vector, and interpolated to 65535 points; originally from 512. When constructing the LUT, the first column linearly increments from 0 to 65535. For the second column, every gray level value that was found through conversion can be matched with each specific integer within the first column. If the gray level values that were obtained (via conversion) were used for this second column, several gaps would be present in the LUT file. The interpolation that was performed essentially fills in the holds that would have otherwise been left empty. This leads to the LUT file necessary for directing the proper phase response on the SLM. Fig. B.5 displays a small portion of the LUT file used for the SLM.

65535	65535	65485	65491	65435	65431
65534	65535	65484	65490	65434	65430
65533	65535	65483	65488	65433	65429
65532	65535	65482	65487	65432	65427
65531	65535	65481	65486	65431	65426
65530	65535	65480	65485	65430	65425
65529	65535	65479	65484	65429	65424
65528	65535	65478	65482	65428	65423
65527	65535	65477	65481	65427	65421
65526	65535	65476	65480	65426	65420
65525	65535	65475	65479	65425	65419
65524	65535	65474	65478	65424	65418
65523	65535	65473	65476	65423	65417
65522	65535	65472	65475	65422	65415
65521	65534	65471	65474	65421	65414
65520	65533	65470	65473	65420	65413
65519	65531	65469	65472	65419	65412
65518	65530	65468	65470	65418	65411
65517	65529	65467	65469	65417	65409
65516	65528	65466	65468	65416	65408
65515	65527	65465	65467	65415	65407
65514	65525	65464	65466	65414	65406
65513	65524	65463	65464	65413	65405
65512	65523	65462	65463	65412	65404
65511	65522	65461	65462	65411	65402
65510	65521	65460	65461	65410	65401
65509	65519	65459	65460	65409	65400
65508	65518	65458	65459	65408	65399
65507	65517	65457	65457	65407	65398
65506	65516	65456	65456	65406	65396
65505	65515	65455	65455	65405	65395
65504	65513	65454	65454	65404	65394
65503	65512	65453	65453	65403	65393
65502	65511	65452	65451	65402	65392
65501	65510	65451	65450	65401	65390
65500	65509	65450	65449	65400	65389
65499	65508	65449	65448	65399	65388
65498	65506	65448	65447	65398	65387
65497	65505	65447	65445	65397	65386
65496	65504	65446	65444	65396	65384
65495	65503	65445	65443	65395	65383
65494	65502	65444	65442	65394	65382
65493	65500	65443	65441	65393	65381
65492	65499	65442	65439	65392	65379
65491	65498	65441	65438	65391	65378
65490	65497	65440	65437	65390	65376

Figure B.5: LUT file that was generated after diffraction-method calibration. The left column increments by unit increments, and is mapped to the values in the second column