

ACHROMATIC PHASE SHIFTING FOCAL PLANE MASKS

by

Kevin Newman

Copyright © Kevin Newman 2016

A Dissertation Submitted to the Faculty of the

COLLEGE OF OPTICAL SCIENCES

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2016

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Dissertation Committee, we certify that we have read the dissertation prepared by Kevin Newman, titled Achromatic Phase Shifting Focal Plane Masks and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

_____ Date: 3/30/2016
Olivier Guyon

_____ Date: 3/30/2016
Michael Hart

_____ Date: 3/30/2016
Tom Milster

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

_____ Date: 3/30/2016
Dissertation Director: Olivier Guyon

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of the requirements for an advanced degree at the University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that an accurate acknowledgement of the source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: Kevin Newman

ACKNOWLEDGEMENTS

This dissertation would not have been possible without the support of some very special people! I would first like to thank my advisor, Dr. Olivier Guyon, for his endless supply of innovative ideas and his contagious excitement for developing technology to study exoplanets. I would also like to thank my committee members, Dr. Michael Hart and Dr. Tom Milster, for their patience and guidance throughout the PhD process.

I would like to give a special thanks to the Ames Coronagraph Experiment group, where I have been integrated as part of the team for the past few years. Especially to Dr. Ruslan Belikov, who has been a fantastic mentor since the first day of my original summer internship with the group. Also thanks to Dr. Eduardo Bendek for the graduate school survival training, and to Dr. Dan Sirbu for the help with wavefront control simulations. I learned so much during my time with the ACE group, and I look forward to the success of their future projects!

I am forever grateful to Dr. Douglas O'Handley and the NASA Ames Academy for giving me a chance to become a part of the space exploration community. The Academy program has had a tremendously positive impact on my career and interest in space technology. The NASA Space Technology Research Fellowship program generously provided funding for my PhD, and the Arizona Space Grant Consortium helped to spark my interest in space research during my undergraduate studies. I cannot overstate the generous support and inspiration that I have received from NASA student programs throughout my academic life.

More generally, I would like to thank the entire College of Optical Sciences for providing a world-class education and for being accommodating during my off-campus studies. A special thanks to the Steward Observatory staff, especially Kim Chapman and Christina Siqueiros for

their patience and assistance. Also to James Conway for playing a critical role in "exposing" me to the world of nanofabrication.

Lastly, and perhaps most importantly, I would like to thank Elizabeth Qin for supporting and putting up with me. I will do my best to return the favor during your PhD! I would also like to thank my network of friends in the Bay Area for keeping me sane during long stretches of dissertation writing, and for providing constant motivation to finish.

DEDICATION

Dedicated to Larry and Darlene Newman, my most influential teachers.

Contents

1	Abstract	14
2	Introduction	16
2.1	Extrasolar planets	17
2.1.1	Indirect methods	17
2.1.2	Direct imaging	20
2.2	Adaptive Optics	27
2.3	Apodization	29
2.4	Focal plane masks	31
2.4.1	Amplitude masks	32
2.4.2	Phase masks	33
2.5	PIAACMC as a strong candidate for the future of direct imaging . . .	34
3	Science potential	37
3.1	Current capabilities	38
3.2	Potential science with new focal plane masks	39
3.3	Developing focal plane masks to advance exoplanet science	44
4	Achromatic size-scaling opaque focal plane mask	46
4.1	Achromatic size-scaling focal plane mask design	48
4.1.1	Achieving size-scaling with wavelength	48

Contents

4.1.2	Compatibility with wavefront control	49
4.2	Mask structure	50
4.3	Phase shift	54
4.4	Fabrication	56
4.5	Characterization of size-scaling with wavelength	57
4.6	Performance in a coronagraph	62
5	Phase masks for PIAACMC	68
5.1	Design	70
5.1.1	Smoothing mask profiles	76
5.2	Manufacturing	79
5.3	Results	80
5.4	Correcting for mask manufacturing errors	84
5.4.1	Surface roughness	89
5.5	Science potential with reflected light	91
6	Perspectives and future directions	94
6.1	Advancing the performance of coronagraphy	95
6.2	Future directions and considerations for space performance	97
7	References	99

List of Figures

2.1	The indirect methods for finding exoplanets, including the Transit method, Radial Velocity, Astrometry, and Microlensing.	21
2.2	Comparison of atmospheric infrared spectra of Venus, Earth, and Mars. Recording this type of spectra for exoplanets can help to determine the composition of the atmosphere, whether the planet could be habitable, and even if life currently exists on the planet. <i>Figure credit: The Pennsylvania State University</i>	23
2.3	The External Occulter method for directly imaging an exoplanet. The occulter casts a shadow over the telescope so that the starlight is blocked from the telescope and creates a dark region to observe nearby objects, such as an exoplanet.	25
2.4	A coronagraph is used to suppress starlight and enable direct imaging of exoplanets. These two panels have different gray scales to emphasize the planet light once the starlight has been suppressed in the right panel.	26
2.5	The contrast of a stellar image (log10). An Electric Field Conjugation algorithm is used to compute a Deformable Mirror shape that creates a dark region for exoplanet imaging on the right side of the stellar image. Additional speckles are created on the left side of the image.	28

List of Figures

2.6	A standard telescope pupil produces diffraction rings surrounding the PSF in the focal plane. Pupil plane apodization helps to suppress the diffraction rings and create a dark area surrounding the star.	31
2.7	PIAA uses aspheric optics to re-shape the light profile in the pupil plane. The resulting stellar image in the focal plane is free of diffraction rings.	32
2.8	A typical configuration for a coronagraph. The starlight is modified in the pupil and focal planes to suppress starlight and enable direct imaging of exoplanets.	32
2.9	Comparison of useful throughput for a PIAACMC design with the theoretical performance limit of coronagraphy on a stellar point source. <i>Figure credit: [37]</i>	36
3.1	List of planets that have been directly imaged as of December 2015 (source: Exoplanets.org).	38
3.2	Contrast and separation angle required to directly observe a hypothetical Earth-size exoplanet with reflected light at quadrature in the habitable zone around nearby stars.	40
3.3	Contrast and separation angle required to directly observe a hypothetical Jupiter-size exoplanet with reflected light at quadrature in the habitable zone around nearby stars. The giant planets that have been observed in thermal emission with previous coronagraphs are shown with red dots for comparison.	43

List of Figures

4.1	Schematic of the SCExAO instrument (<i>left</i>) and intensity profile in the pupil plane after PIAA (<i>right</i>). The achromatic size-scaling focal plane mask diffracts starlight onto a partially reflective Lyot stop which is then used in a low-order wavefront sensor. Planet light is passed by the focal plane mask and imaged onto the science camera.	48
4.2	A microscope image of a mask fabricated at NASA JPL Microdevices Laboratory, along with a diagram to define geometric parameters of the mask (<i>left</i>). Isometric view of mask profilometry (<i>right</i>).	51
4.3	Diffraction pattern produced by the binary structure of the achromatic size-scaling mask when it is illuminated with an apodized stellar PSF. When the binary pattern is made of 14 segments (<i>left</i>) the diffracted light is located closer to the optical axis compared with a binary pattern made of 30 segments (<i>right</i>).	53
4.4	Phase shift produced by adjacent cells in the achromatic mask binary pattern as a function of wavelength for $h = 1.81\mu\text{m}$. The phase is wrapped back to zero for shifts that are greater than one wavelength. Within the H-band (<i>dashed yellow box</i>), the phase shift is monotonic.	56
4.5	First simulation of the size-scaling property of the achromatic mask as a function of wavelength (<i>solid lines</i>). Results are compared to a conventional hard-edged mask (<i>dashed lines</i>). We measure throughput of an off-axis source and compare IWA for different wavelengths. . . .	58
4.6	Intensity as a function of distance from the center of the mask in measurement (<i>solid lines</i>) and simulation (<i>dashed lines</i>) for various wavelengths. The location of intensity increase, representing the effective edge of the mask, scales with wavelength.	61

List of Figures

4.7	Second simulation of the size-scaling property of the achromatic mask as a function of wavelength (<i>solid lines</i>). Results are compared to a conventional hard-edged mask (<i>dashed lines</i>).	63
4.8	Image of focal plane mask contrast recorded using the SCEXAO internal calibration source. The color scale is logarithmic and spans three orders of magnitude.	64
4.9	Simulation (<i>left</i>) and measured image (<i>right</i>) of starlight diffraction pattern in the Lyot plane. The logarithmic color scale spans three orders of magnitude and represents contrast relative to the maximum intensity of an unobstructed PSF. The scale for the measured image is offset by an order of magnitude relative to the simulated scale.	66
5.1	Schematic of PIAACMC system architecture.	69
5.2	Mask design process: compute the complex amplitude in the final image plane for each zone of the mask.	73
5.3	Apodization profile of the PIAA optics at the ACE lab (a). Zone geometry of a PIAACMC focal plane mask with ten concentric rings (b). Mask designs for PIAA optics at the ACE lab (c), including ideal designs (blue curve), ideally smoothed designs (red curve), and practically smoothed designs (green curve).	75
5.4	Optical microscope image of a manufactured mask.	81
5.5	AFM measurement of a manufactured mask (bottom), and profile of the zone depths compared with the designed profile (top).	82

List of Figures

5.6	Comparison of contrast over a 10% bandwidth for an ideal mask and a fabricated mask. The performance of the fabricated mask is more than one order of magnitude worse than the ideal mask. However, performance should improve with further fabrication development and implementation of wavefront control.	83
5.7	Radially symmetric apodization profile of the modified ACE PIAA optics (a) and structure of the focal plane mask design (b). A cross section profile of the Design 4 focal plane mask (c).	86
5.8	Contrast of an unobstructed pupil PIAACMC design with random manufacturing errors added to the focal plane mask (red dots). The contrast can be improved, up to a certain level, with the use of wavefront control (blue dots).	90
5.9	Contrast and separation angle required to directly observe a hypothetical Jupiter-size exoplanet in reflected light at quadrature in the habitable zone around nearby stars. The solid blue line shows the performance of the Design 4 PIAACMC system with a 10m diameter telescope aperture, and the dashed blue line shows the performance with a 2.4m telescope.	93

1 Abstract

The search for life on other worlds is an exciting scientific endeavor that could change the way we perceive our place in the universe. Thousands of extrasolar planets have been discovered using indirect detection techniques. One of the most promising methods for discovering new exoplanets and searching for life is direct imaging with a coronagraph. Exoplanet coronagraphy of Earth-like planets is a challenging task, but we have developed many of the tools necessary to make it feasible.

The Phase-Induced Amplitude Apodization (PIAA) Coronagraph is one of the highest-performing architectures for direct exoplanet imaging. With a complex phase-shifting focal plane mask, the PIAA Complex Mask Coronagraph (PIAACMC) can approach the theoretical performance limit for any direct detection technique. The architecture design is flexible enough to be applied to any arbitrary aperture shape, including segmented and obscured apertures. This is an important feature for compatibility with next-generation ground and space-based telescopes.

PIAA and PIAACMC focal plane masks have been demonstrated in monochromatic light. An important next step for high-performance coronagraphy is the development of broadband phase-shifting focal plane masks. In this dissertation, we present an algorithm for designing the PIAA and PIAACMC focal plane masks to operate in broadband. We also demonstrate manufacturing of the focal plane masks, and show laboratory results. We use simulations to show the potential performance of the coronagraph system, and the use of wavefront control to correct for mask manufacturing

1 Abstract

errors.

Given the laboratory results and simulations, we show new areas of exoplanet science that can potentially be explored using coronagraph technology. The main conclusion of this dissertation is that we now have the tools required to design and manufacture PIAA and PIAACMC achromatic focal plane masks. These tools can be applied to current and future telescope systems to enable new discoveries in exoplanet science.

2 Introduction

2.1 Extrasolar planets

For thousands of years, people have wondered if life exists on other worlds. Through scientific inquiry, we have discovered that our home, the Earth, is one of several planets orbiting around our host star, the Sun. The Sun is just one of approximately 100 billion stars in the galaxy, and our galaxy is just one of at least 100 billion galaxies in our universe. Given the sheer number of stars in our universe, one begins to wonder if we could possibly be alone. Perhaps it is only a matter of effort, ingenuity, and time to discover our galactic neighbors. My work focuses on the technology to help make it happen.

An extrasolar planet (exoplanet) is a planet that orbits around a star other than our own Sun. The first exoplanet discovery occurred in 1992 [1], and since then we have discovered thousands of exoplanets. The various methods for discovering and studying exoplanets can be grouped into two categories: indirect methods, and direct imaging.

2.1.1 Indirect methods

Most of the exoplanet discoveries to date have been made with indirect methods. In other words, these methods do not gather any information directly from the planet; they gather information from the host star, and look for the effects that a planet would have on the star. These methods are capable of discovering planets in certain situations, depending on the size of the planet and the orbital parameters, and can determine the orbital distance, mass, and radius of the planet.

With the transit method, we observe the intensity of light from a star, and look for a periodic decrease in the light that would occur as a planet passes in front of the star from our point of view, as shown in Figure 2.1a. The decrease in magnitude, given by the difference between the intensity during transit (I_T) and the intensity outside

2 Introduction

of transit (I_o), is proportional to the square of the ratio of the radius of the planet (R_B) to the radius of the star (R_A) as given in Equation 2.1.

$$\frac{I_o - I_T}{I_o} = \frac{\pi R_B^2}{\pi R_A^2} = \left(\frac{R_B}{R_A}\right)^2 \quad (2.1)$$

The transit signal is particularly strong for large planets and small stars. The transit occurrence rate is higher for planets located close to their host star because there is a wider range of orbital inclination angles for which a transit could occur. However, the probability that a planet's orbital orientation is compatible with the transit method is small, for example on the order of 1% for a Sun-Earth like system, and therefore most planets are invisible to the transit detection method. This method is particularly useful because a survey, such as the Kepler Space Telescope [2], can look for planets around hundreds of thousands of stars simultaneously. This is useful for determining planet population statistics in our galaxy [3].

Radial velocity is a method where one observes the spectrum of light from a star, and looks for periodic shifts in wavelength. The star-planet system can be roughly modeled as a two-body system, where both objects orbit around their common center of mass, as shown in Figure 2.1b. If a star is periodically moving towards or away from us throughout its orbit, we can observe that movement as a shift in the stellar spectrum, and infer the presence of a planet. The velocity of the star (V_{star}) can be calculated with the velocity of the planet (V_{pl}) and the ratio of the planet mass (M_{pl}) to the stellar mass (M_{star}) as shown in Equation 2.2. A larger planet mass will produce a larger change in the relative velocity of the star, and a larger shift in the stellar spectrum. The observed quantity is $V_{star} \cdot \sin(i)$, where i is the orbital inclination of the planet. An orbital inclination that is perpendicular relative to our line of observation will not produce a radial velocity signal. The semi-major axis (star to planet orbital distance, R) can be calculated using the observed orbital period

2 Introduction

(P) and Kepler's Third Law, which is shown in Equation 2.3 with the gravitational constant (G). The velocity of the planet can be calculated with Equation 2.4, and then the mass of the planet can be calculated as a function of orbital inclination, $M_{pl}\sin(i)$, using Equation 2.2. A planet located close to its host star will produce a stronger radial velocity signal than the same planet located further from the star. Radial velocity is best suited for large planets located close to their host star, and for planets whose orbital inclination is aligned such it pulls the star directly towards or away from our observation point.

$$V_{star} = V_{pl} * \frac{M_{pl}}{M_{star}} \quad (2.2)$$

$$R^3 = \frac{P^2 GM_{star}}{4\pi^2} \quad (2.3)$$

$$V_{pl} = 2\pi \frac{R}{P} \quad (2.4)$$

Astrometry is another method for observing the periodic movement of a star that is caused by the gravitational pull of an orbiting planet. In this case, shown in Figure 2.1c, we precisely measure the position of the star over time and look for periodic movement. The angular shift of the star (Θ) is proportional to the mass of the planet relative to the mass of the star, as shown in Equation 2.5. The signal is inversely proportional to the distance between the star and our observation point (d). The signal is stronger for a large orbital distance, which can be calculated from the orbital period using Equation 2.3. This technique requires precise calibration using other stars in the field of view for reference, and is best suited for nearby low-mass stars and large-mass planets.

$$\Theta = \frac{M_{pl}}{M_{star}} \frac{R}{d} \quad (2.5)$$

2 Introduction

Gravitational microlensing is an event where a foreground object, such as a star, acts as a lens by bending the light from a background object. When the two objects are aligned from our point of view, the microlensing effect can cause an increase in the observed flux from the background object. This principle can be used to detect exoplanets, as shown in the geometry of Figure 2.1d. When the microlensing event only involves a single foreground object, the flux amplification follows a smooth profile. When the foreground star has a planet, the planet contributes to the microlensing event, and modulates the flux amplification profile. Planets discovered with this method are difficult to observe again because the two-body alignment only lasts for a few days and does not occur again.

The indirect detection methods have begun to populate our map of the galaxy with exoplanets, and have produced a wealth of knowledge about planet occurrence rates and planet formation [4, 21]. The indirect methods can be complimentary to each other, and two or more indirect methods are often used to confirm the existence of a planet. Because the methods do not gather any light directly from the planet, the information about each planet is usually limited to its mass or size, and orbital distance. There have been some attempts at transit spectroscopy [6], but a more exact way to measure the spectra of an exoplanet is with direct imaging.

2.1.2 Direct imaging

Direct imaging is an exoplanet detection technique where we observe the thermal emission that is produced by the planet, or the light from the host star that is reflected by the planet. In this section we discuss the principles of direct imaging, and in Section 2.1 we will discuss the current work in the field. The technique can be sensitive to planets that are inaccessible to the indirect methods, such as planets that do not transit from our point of observation, or planets whose orbital inclination

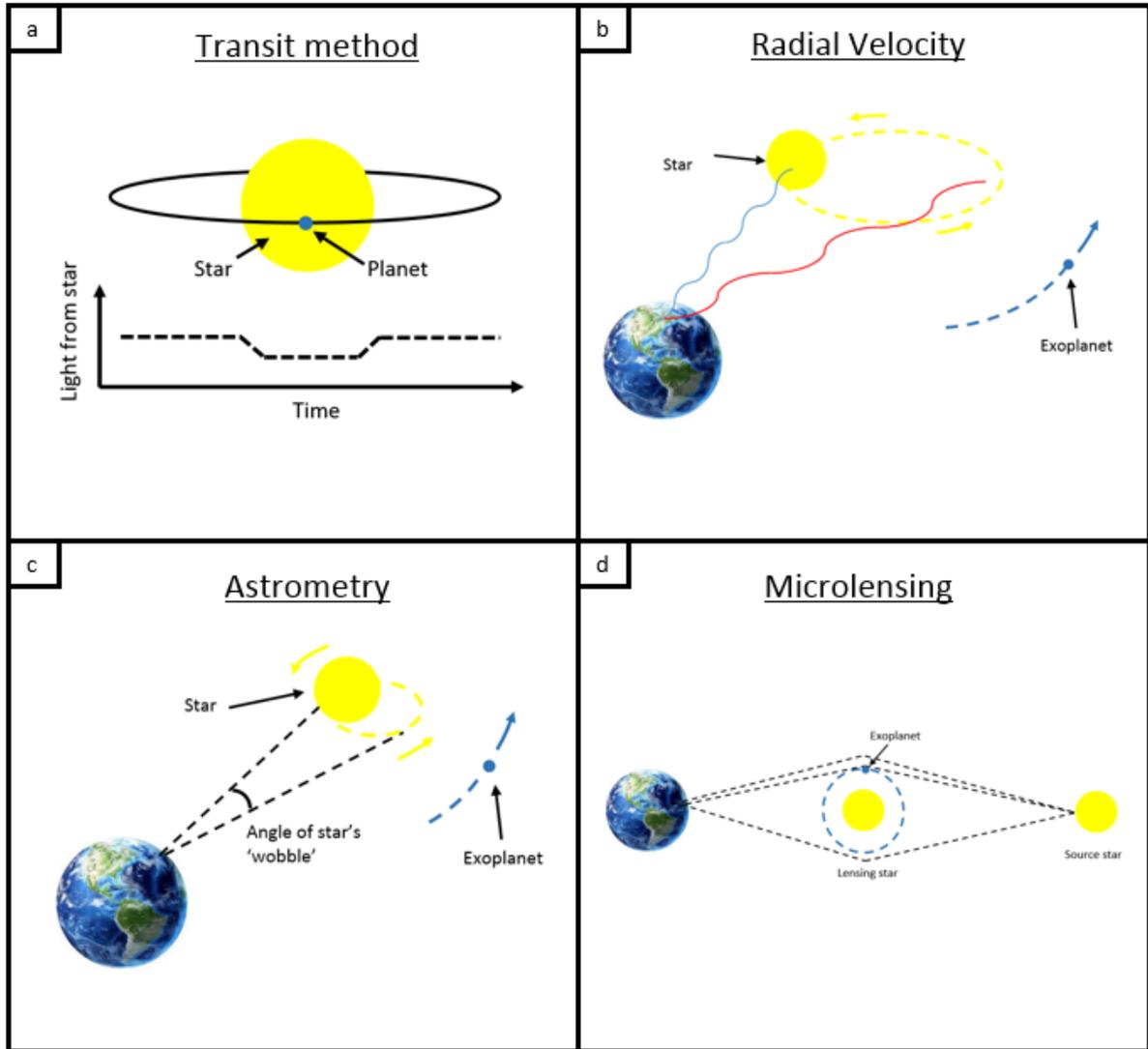


Figure 2.1: The indirect methods for finding exoplanets, including the Transit method, Radial Velocity, Astrometry, and Microlensing.

2 Introduction

is perpendicular to our line of sight, and therefore do not produce a radial velocity signal. The scientific advantage of direct imaging is that one can not only determine the mass and orbit of the planet, but can also potentially divide the planet light into its spectrum to determine the composition of the planet atmosphere [7, 8]. An example of this principle is shown in Figure 2.2. Planet spectroscopy can be used to look for evidence of life on the planet, such as an abundance of Oxygen.

Young giant planets produce their own thermal emission in the near-infrared, and gradually cool after formation. Older planets produce thermal emission at a wavelength of $10\mu m$ and longer. In both scenarios, we can design a direct imaging system to directly observe the thermal emission, but the demands on the telescope are different. Directly observing in the near-infrared is favorable to resolve the star-planet separation, but the contrast ratio is more difficult to achieve. Directly observing in the longer thermal infrared is favorable for contrast, but would require a large telescope or interferometric system to resolve the star-planet separation. Several systems, outlined in Section 2.1, have successfully observed the thermal emission of planets.

Planets can also be observed with reflected starlight from the host star in the visible spectrum. This method requires the planet to be located close to its host star because the reflected light becomes too dim when the planet is located further away. Resolving the star-planet separation becomes easier in the visible spectrum, but the contrast ratio is more challenging. Directly imaging planets in reflected light is most sensitive to large planets around nearby stars. However, direct imaging is potentially capable of observing all types of planets, including Earth-like planets located in the habitable zones of nearby low-mass stars [9]. The ideal scenario is to observe a planet at its quadrature, which is the greatest angular separation between the planet and the star from our point of view.

There are two fundamental challenges for direct imaging systems. The first chal-

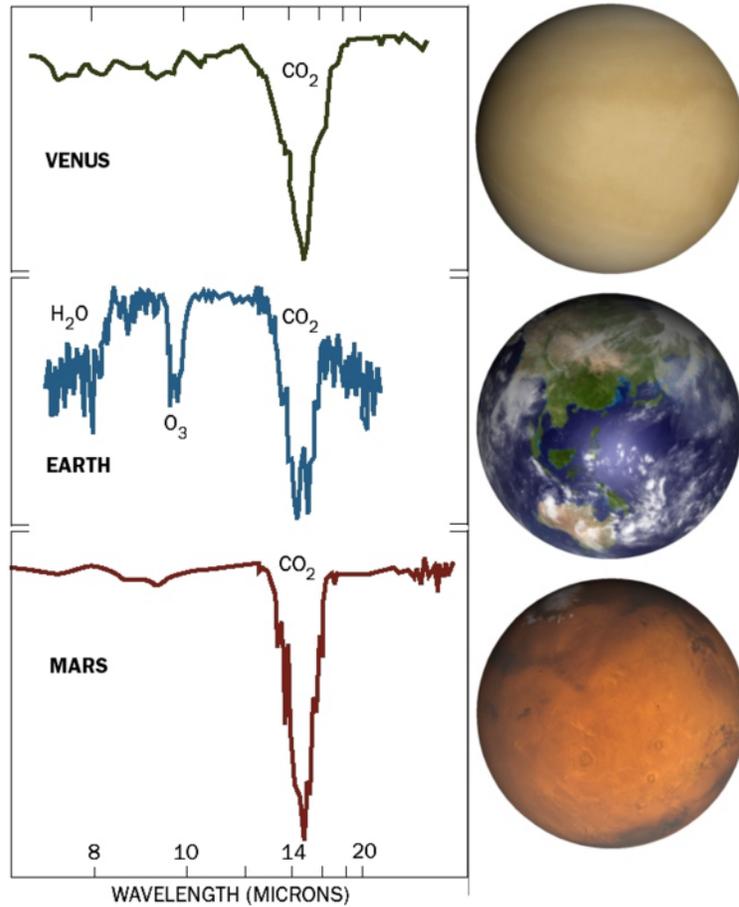


Figure 2.2: Comparison of atmospheric infrared spectra of Venus, Earth, and Mars. Recording this type of spectra for exoplanets can help to determine the composition of the atmosphere, whether the planet could be habitable, and even if life currently exists on the planet. *Figure credit: The Pennsylvania State University*

2 Introduction

challenge is to resolve the angular separation between the star and the planet. This is achieved by using a large diameter telescope, and observing at a wavelength sufficiently short for the telescope diffraction limit (λ/D) to be smaller than the planet-star separation. The scientific potential of a direct imaging system will depend on the size of the telescope that it is applied to and the range of wavelengths that are used for observation. It is useful to separate the system performance from these factors in order to directly compare performance between different types of direct imaging systems, so we typically use the λ/D unit to describe the angular separation between the planet and its host star.

The second challenge is to suppress the starlight that is produced by the host star. The starlight is so bright in comparison to the planet light that it overwhelms the image of the planet and makes it impossible to observe. The contrast ratio between the star and the planet varies based on the type of star, the size of the planet, the orbital distance, and the bandwidth that you are using to observe the planet. A young Jupiter-size planet orbiting around a G-type star can be observed in the near-infrared with a contrast ratio of 10^{-6} . An Earth-size planet located in the Habitable Zone of a Sun-like star has a contrast ratio of 10^{-10} in the visible spectrum. These contrast ratios are possible to achieve using modern advancements in starlight suppression techniques, such as external occulter, coronagraphy, and interferometric nulling.

An External Occulter, or 'Star Shade', is a large mask that is located between a telescope and the object star. As shown in Figure 2.3, the mask casts a shadow over the telescope, blocking the starlight before it reaches the telescope. This method can potentially reach very deep contrast, but has practical challenges such as precision manufacturing of a large occulter, and formation flying of the two objects in space [10, 11]. A typical distance between a telescope and an external occulter is on the order of tens of kilometers.

2 Introduction

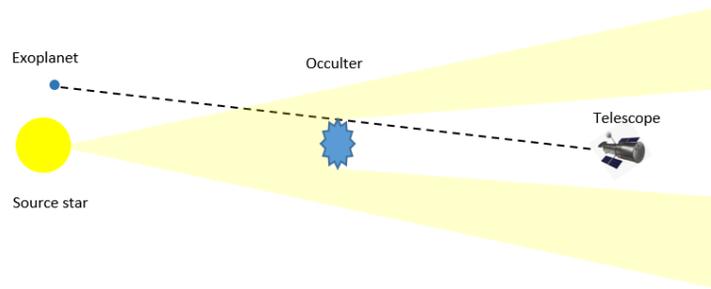


Figure 2.3: The External Occulter method for directly imaging an exoplanet. The occulter casts a shadow over the telescope so that the starlight is blocked from the telescope and creates a dark region to observe nearby objects, such as an exoplanet.

Nulling interferometry is a method that can combine the high resolution of an interferometric system with the starlight suppression that is necessary for direct imaging. The concept was first proposed for observing exoplanets by Dr. Bracewell [12], and achieves starlight suppression by combining two or more signals from the host star and forcing destructive interference. The Terrestrial Planet Finder Infrared Interferometer [13] was a NASA proposed project where the starlight collected by several space telescopes in a fixed structure or formation flying would be combined to null the star and observe exoplanets. A technology plan was developed to advance the mission [14], but it was eventually cancelled in 2011. Darwin was a similar mission concept proposed by the European Space Agency (ESA) for a constellation of space telescopes to observe terrestrial extrasolar planets in the mid-infrared wavelengths [15], but this mission was also not pursued past the concept phase. Nulling interferometry has the potential for high resolution imaging with a constellation of space telescopes, because long baselines between spacecraft may be more practical to achieve than a single large aperture telescope. However, the nulling concept requires high precision alignment and stable distances between the spacecraft. The beam combining concept is cur-

2 Introduction

rently being demonstrated in a more accessible ground-based environment with the Large Binocular Telescope Interferometer, which is using nulling interferometry for observing exozodiacal dust [16].

A coronagraph creates a focused image of the host star, and eliminates the starlight with a combination of pupil plane and focal plane modifications to the starlight. This concept was originally used to directly image the corona of our Sun [17]. The concept, shown in Figure 2.4, can be applied to both ground-based and space-based telescopes for directly observing exoplanets. Coronagraphs can achieve the deep contrast necessary to observe exoplanets using only moderately sized optical and mechanical components. Several different coronagraph architectures are discussed in Section 1.4, and we will spend the remainder of this thesis focusing on the development of technology to enable direct exoplanet imaging with a coronagraph.

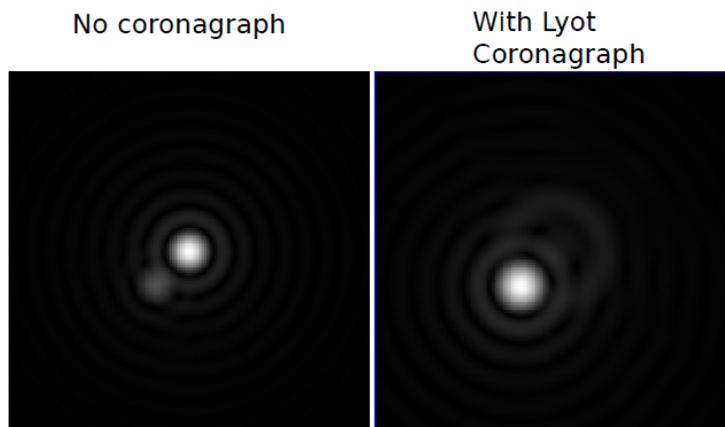


Figure 2.4: A coronagraph is used to suppress starlight and enable direct imaging of exoplanets. These two panels have different gray scales to emphasize the planet light once the starlight has been suppressed in the right panel.

The primary tool for a coronagraph to create a dark region for direct imaging is the starlight suppression system. Starlight suppression can be achieved with a combination of pupil plane modifications and focal plane masking or phase shifting.

2 Introduction

A coronagraph often includes apodization in the pupil plane, which is reshaping the light pattern to control diffraction in the focused image. The starlight suppression system is designed for a specific telescope aperture in order to compensate for the diffraction pattern of any obscurations or segments. Inevitably, small errors in the optical system, in addition to atmospheric turbulence if observing from the ground, cause wavefront errors which degrade the contrast in the dark region. For this purpose, an adaptive optics (AO) system is used to measure and correct for errors, and enhance the achievable contrast.

2.2 Adaptive Optics

Adaptive optics (AO) is a critical technology for enabling the direct imaging of exoplanets, as it is the primary tool for measuring and actively correcting optical errors in a telescope system. For ground-based observing, AO is most commonly used to correct for atmospheric turbulence and fast speckles [18], which distort the image of all celestial objects. For space-based observing, we use AO to correct for slowly-changing errors due to thermal variations and minor system misalignments. In both cases, AO helps to achieve a stable optical system. Extreme Adaptive Optics is used for high-contrast imaging [19, 20, 21].

A typical AO system has a wavefront sensor (WFS), at least one deformable mirror (DM), and a control loop. At least one guide star is also used to provide light to measure the wavefront aberrations. The guide star can be a natural star within the telescope field of view, or an artificial star generated with a high-powered laser [22]. The WFS is used to sample the wavefront so that a computer algorithm can determine the errors in the system, including atmospheric turbulence, telescope errors, and instrument errors. In some cases, the relatively simple Shack-Hartmann WFS [23] is sufficient for measurement, but recent advances have produced more sensitive

2 Introduction

WFS such as a pyramid, curvature, or non-linear WFS [24, 25].

The DM is a reflective optical device located in a pupil plane of a telescope which can be finely shaped to produce changes in the wavefront. The DM can be segmented or continuous, and the shape is usually controlled by a collection of actuators using an electrostrictive or magnetic force. Various technologies exist for shaping the DM, including a piezo stack, bimorph materials, a flexible membrane, and micro-mechanical systems (MEMS). The shape of the DM is determined with a control loop, which receives input from the WFS and calculates the proper shape of the DM to eliminate speckles and create a dark region. An example dark region is shown in the right side of Figure 2.5. There are several different wavefront control algorithms which can be used to determine the shape of the DM.

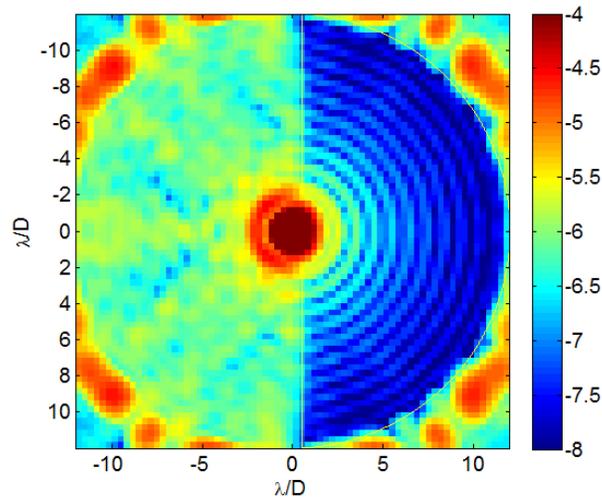


Figure 2.5: The contrast of a stellar image (\log_{10}). An Electric Field Conjugation algorithm is used to compute a Deformable Mirror shape that creates a dark region for exoplanet imaging on the right side of the stellar image. Additional speckles are created on the left side of the image.

Speckles in the image plane are caused by the scattering of starlight by optical defects in the direct detection system. They degrade the contrast of the dark region,

2 Introduction

and appear like planets when they are not corrected. The DM can be used to eliminate speckles, and there are a variety of algorithms to detect speckles and compute a shape for the DM to compensate. With a method called speckle nulling [26], the DM adds an additional speckle to the focal plane with a complimentary phase and amplitude to destructively interfere and eliminate the original speckle. This technique differs from conventional AO because the goal is not to maximize the Strehl ratio of an image, but instead to minimize the light that is scattered into the dark region to enhance contrast and enable faint object imaging. Energy minimization [27] is another correction algorithm based on matrix inversion; it is possible to obtain deeper dark holes with this algorithm, but the outer working angle is limited. Electric Field Conjugation (EFC) is a general speckle control method that uses a model of the system to calculate a phase map on the DM that will create a dark hole [28]. All of the speckle control techniques have the same fundamental goal of minimizing the energy in the dark region.

AO is certainly a critical dynamic tool to achieve a dark region for direct imaging, but it cannot achieve high-contrast imaging without some help from static optical components within the coronagraph. The starlight suppression system has the primary responsibility for removing starlight and creating a dark hole with a combination of pupil plane and focal plane modifications. One of the most commonly used tools for pupil plane modification is apodization.

2.3 Apodization

A standard telescope pupil has unity light intensity within the aperture and zero intensity outside of the aperture. In this scenario, a diffraction limited image of a star produces a diffraction pattern surrounding a central core of light, as shown in Figure 2.6. For an unobstructed circular aperture, the diffraction pattern is a series

2 Introduction

of concentric rings. In most cases of optical imaging, this is the ideal scenario because it indicates that the optical system is properly aligned. However, this is a problem for exoplanet imaging because the intensity of the diffraction rings is orders of magnitude brighter than the planet. We need to suppress the diffraction rings in order to observe the planet. This can be achieved with pupil apodization, which is a technique for changing the distribution of light in the pupil plane. There is a Fourier Transform relationship between the complex amplitude in the pupil plane and the focal plane, as given in Equation 2.6 for focal plane spatial variable x and pupil plane frequency variable ξ . We observe the resulting surface brightness in the focal plane, which is the square modulus of the complex amplitude. One strategy to remove diffraction rings using apodization is to reshape the light pattern in the pupil plane to a Gaussian-like profile, so that the focused image will also have a Gaussian-like profile which does not contain diffraction rings.

$$f(\xi) = \int_{-\infty}^{\infty} f(x)e^{-2\pi i x \xi} dx \quad (2.6)$$

The standard method for pupil apodization is to place a partially absorbing mask in the pupil plane. This method can achieve any arbitrary light intensity profile. The Apodized Pupil Lyot Coronagraph (APLC) [29] and Shaped Pupil Coronagraph [30] are examples of coronagraph systems which use absorption for apodization. However, some planet light is lost in the process of light absorbing for apodization, and the result is fewer photons for exoplanet detection and science. Guyon et al. [31] proposed the Phase-Induced Amplitude Apodization (PIAA) method for pupil plane apodization which achieves the light re-distribution using aspheric optics such as those shown in Figure 2.7. With PIAA, all of the planet light throughput is conserved. The main disadvantage of the PIAA technique is that it distorts the image of any off-axis objects that surround the central star. The coronagraph system uses inverse PIAA optics to

2 Introduction

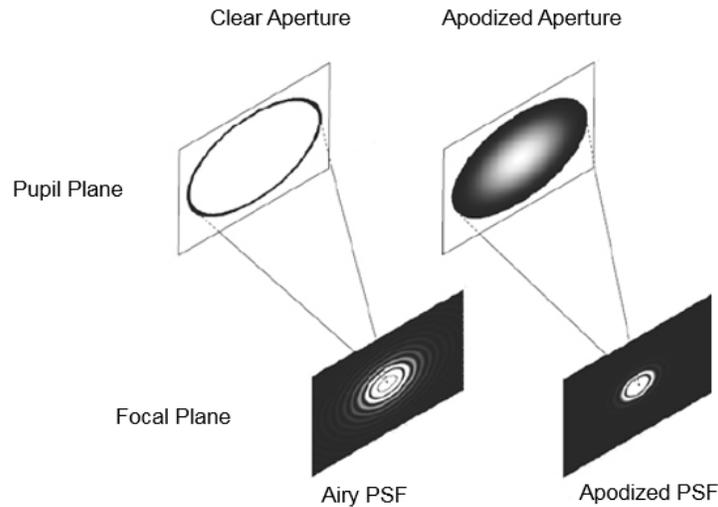


Figure 2.6: A standard telescope pupil produces diffraction rings surrounding the PSF in the focal plane. Pupil plane apodization helps to suppress the diffraction rings and create a dark area surrounding the star.

recover the original off-axis image shape.

2.4 Focal plane masks

In a typical coronagraph architecture, shown in Figure 2.8, the starlight is collected with telescope mirrors and focused to a small point. The location where the starlight is focused is known as the focal plane. A focal plane mask is placed at this location, overlapping with the focused image of the star. The focal plane mask is responsible for providing any focal plane modifications to the starlight that are desired for the coronagraph design. This is often the primary component of the coronagraph for starlight suppression, particularly for removing the central part of the stellar PSF.

There are two categories of focal plane masks: amplitude and phase masks. Both types of mask achieve the goal of directly interacting with the starlight so that it can be eliminated, but the suppression technique is different. An amplitude mask

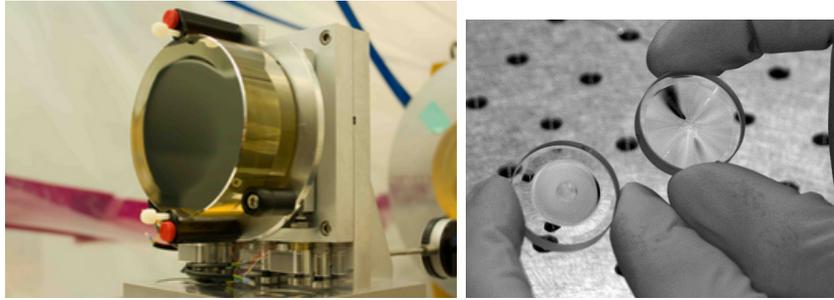


Figure 2.7: PIAA uses aspheric optics to re-shape the light profile in the pupil plane.

The resulting stellar image in the focal plane is free of diffraction rings.

generally blocks some or all of the starlight so that it does not continue through the rest of the optical system. A phase mask provides a phase shift to the starlight so that it can interfere destructively later in the system.

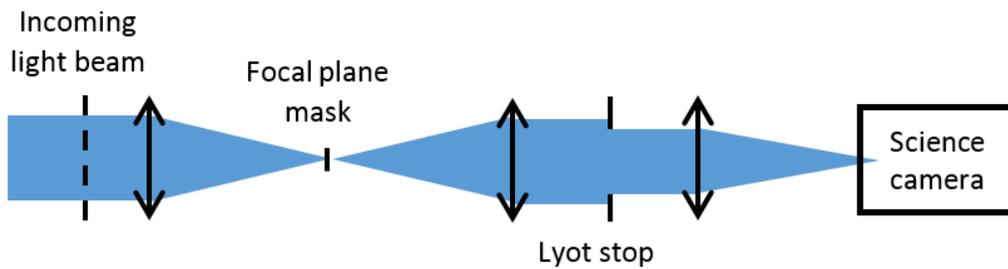


Figure 2.8: A typical configuration for a coronagraph. The starlight is modified in the pupil and focal planes to suppress starlight and enable direct imaging of exoplanets.

2.4.1 Amplitude masks

The simplest type of coronagraph focal plane mask is an opaque disk occulter. One advantage of using an opaque occulting mask is the simple manufacturing and alignment. The PIAA and APLC coronagraphs use an opaque disk occulter. The discarded

2 Introduction

starlight can be used for wavefront sensing [32]. One disadvantage of an opaque disk occulter is the 'size chromaticity' problem, namely, the fact that the size of a telescope point spread function (PSF) illuminating a coronagraphic mask will scale with wavelength, but coronagraphic masks remain a fixed size. In chapter 3, we discuss a new mask which operates conceptually as an opaque disk occulter, but uses a phase mask technique to improve performance and solve the 'size chromaticity' problem.

As an alternative to the opaque disk occulter, one could use another form of amplitude mask where amplitude apodization is used in the focal plane to match a simple Lyot stop in the pupil plane. This concept is used in the Band-limited Coronagraph [33], where the focal plane amplitude is modified to pass only band-limited frequencies, which are diffracted to the outer edge of the pupil plane and blocked with a Lyot stop. This concept can theoretically achieve total on-axis starlight extinction, but suffers from a relatively low planet light throughput.

2.4.2 Phase masks

Instead of modifying the amplitude of the stellar PSF, an alternative approach is to provide a phase shift to some or all of the starlight, and force the phase-shifted starlight to interfere destructively with the un-shifted starlight later in the optical path to achieve starlight suppression. There are several techniques for applying a phase shift in the focal plane. The simplest approach, introduced in the Phase Mask Coronagraph [34], uses a circular phase plate to produce a π phase shift in the core of the stellar image. This technique is simple to implement, but suffers from both size chromaticity and complex amplitude chromaticity. The Four-Quadrant Phase Mask [35] is a binary mask with alternating sections that produce 0 and π phase shift. By dividing the phase shift into quadrants, this approach solves the problem of size chromaticity for a circularly symmetric stellar PSF, but still does not provide a

solution for producing the π phase shift over a broad range of wavelengths.

The Dual-Zone Phase Mask [36] consists of two concentric phase masks where the thickness of each is optimized to null starlight. Compared with the other forms of binary phase masks, this technique can be effective over a larger bandwidth. The PIAA Complex Mask Coronagraph (PIAACMC) combines the circular disk phase mask approach with apodization from PIAA optics [37]. In this design, the phase mask is partially transmissive. In Chapter 4, we will discuss design methods to improve the chromaticity of the PIAACMC mask.

Instead of applying the phase shift in a binary pattern, one could apply the phase shift gradually. The Optical Vortex Coronagraph [38] uses a diffractive optical element with a purely azimuthal phase ramp to produce a smooth gradient phase profile in the azimuthal direction. One challenge with this approach is that the center of the vortex is an infinitely small point with an abrupt change in phase in all azimuthal directions; this is challenging to manufacture. The Annular Groove Phase Mask [39] induces an achromatic optical vortex by using a subwavelength zero-order grating.

Phase masks generally allow planet detection closer to the star and can achieve higher contrast when compared with the opaque occulting masks. However, phase masks can be particularly sensitive to manufacturing errors and instrument alignment errors.

2.5 PIAACMC as a strong candidate for the future of direct imaging

A standard Phase-Induced Amplitude Apodization (PIAA) coronagraph uses pupil apodization produced with PIAA optics to achieve a high contrast PSF while maintaining full throughput of planet light. An opaque disk focal plane mask is used to

2 Introduction

block the central core of the stellar PSF, and a Lyot stop is used to limit diffraction from the edges of the focal plane mask. The PIAACMC is an improvement on the PIAA coronagraph architecture, where the opaque disk focal plane mask is replaced with a partially transmissive phase-shifting focal plane mask. This technique combines the full throughput of PIAA with the low inner working angle (IWA) that is possible with a phase mask. The IWA is the angular separation from the host star at which the coronagraph can achieve 50% throughput of the planet light. This is an important metric for coronagraph performance because it describes the minimum angular separation at which the coronagraph can get sufficient light from the planet to identify the planet and perform scientific characterization.

The phase mask of a PIAACMC design only needs to provide a phase shift to a portion of the stellar PSF, and therefore the mask can be significantly smaller than an opaque mask in a conventional PIAA design, which needs to block the entire stellar PSF. The smaller mask allows for a smaller IWA. PIAA optics for a conventional PIAA coronagraph generally have strong apodization that approaches zero intensity towards the edge of the pupil. This technique limits diffraction ringing around the stellar PSF in the focal plane. With a phase mask, the system can accept more diffraction ringing in the focal plane because the mask can be designed to compensate with destructive interference. This trade-off allows the PIAA optics in a PIAACMC system to have a milder profile, which has the practical advantage of being easier to manufacture.

The phase mask in a PIAACMC system allows more design flexibility, which is important for compatibility with complex apertures, such as segmented and centrally obscured apertures. The primary challenge with complicated apertures is to control the diffraction, and the phase mask is an adjustable tool that can be designed for any aperture shape. PIAACMC performs favorably in comparison to other coronagraph

2 Introduction

architectures, especially considering the high throughput for planet light. In some cases, as shown in Figure 2.9, PIAACMC can approach the theoretical performance limit for any direct imaging instrument [9, 37].

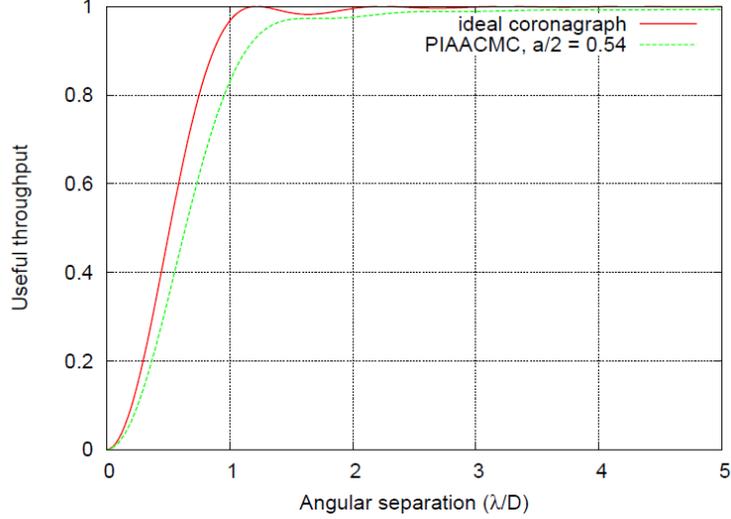


Figure 2.9: Comparison of useful throughput for a PIAACMC design with the theoretical performance limit of coronagraphy on a stellar point source. *Figure credit: [37]*

PIAA optics have been thoroughly developed and demonstrated [40, 41, 42], so the only remaining piece of technology that needs to be developed in order for PIAACMC to be successful is the phase-shifting focal plane mask. For observing in monochromatic light, the focal plane mask can be as simple as a circular phase plate. However, for the mask to work in broadband, the design and manufacturing of the mask becomes more challenging. In chapter 4, we will have a more detailed discussion of the challenges to develop the complex focal plane mask, and demonstrate strategies for design, manufacturing, and testing of the mask. Our goal is to develop the final piece of technology to enable PIAACMC to be applied to current and future telescopes and provide high-performance exoplanet science.

3 Science potential

3.1 Current capabilities

Direct exoplanet imaging is a relatively new field in astrophysics research, and there is a wealth of potential for new discoveries. The first planet to be directly imaged occurred in 2004 [43]. As of December 2015, there are a total of 8 confirmed directly imaged exoplanets [44], the majority of which are young planets, thermal emission imaged far from their host stars. These planets are the easiest targets for direct imaging, so they could be discovered with the early systems.

Planet Name	Semi-Major Axis (au)	Orbital Period (days)	Orbital Eccentricity	First Reference
HD 95086 b				Rameau 2013
HR 8799 b	67	1.64×10^5	0	Marois 2008
Fomalhaut b	113	3.2×10^5	0.120	Kalas 2008
GJ 504 b				Kuzuhara 2013
HR 8799 c	42	82000	0	Marois 2008
HR 8799 d	27	41000	0.1	Marois 2008
beta Pic b	8.8	7200	0.021	Lagrange 2009
HR 8799 e	14.5	16400	0	Marois 2010

Figure 3.1: List of planets that have been directly imaged as of December 2015 (source: Exoplanets.org).

There are several ground-based coronagraphs which are currently pushing the boundaries of exoplanet science. The Near Infrared Coronagraphic Imager for the Gemini South Telescope uses a Lyot coronagraph [45]. The High Contrast Instrument for the Subaru Next Generation Adaptive Optics (HiCIAO) [46] on the Subaru Telescope also uses a Lyot coronagraph. The Subaru Coronagraphic Extreme AO (SCEXAO) [47] system uses a PIAA coronagraph. The Spectro-Polarimetric High-contrast Exoplanet Research (SPHERE) [48] instrument on the European Southern Observatory Very Large Telescope uses a Lyot coronagraph and a four-quadrant phase mask. The Gemini Planet Imager (GPI) [49] on the Gemini Telescope uses a classical

Lyot coronagraph and an APLC.

Most of the current coronagraph systems are associated with ground-based telescopes. They are capable of observing the thermal emission of exoplanets, but not the reflected visible light because it is difficult to achieve the necessary contrast from the ground.

3.2 Potential science with new focal plane masks

There is a wealth of potential exoplanet discovery and science space that is yet to be explored. Many nearby stars are prime targets for potential exoplanet science with direct imaging. As we discussed in Section 1.1.2, there is potential science for observing exoplanets with both thermal emission and reflected light. These two observation cases have different requirements for the instruments necessary to observe. Young, hot planets can be observed with thermal emission in the near-infrared. As the planets cool, their thermal emission spectrum shifts towards the mid-infrared. Compared with reflected light, the contrast ratio with thermal emission is easier to achieve (typically 10^{-6}), but the star-planet separation is more difficult because the telescope size required scales with observation wavelength.

Observing exoplanets in reflected light (visible and near-infrared spectrum) has a more challenging contrast ratio, but the smaller requirement for telescope aperture size is more practical. For this reason, we will focus on directly observing in reflected light. The most scientifically interesting scenario is to study planets that could be potentially habitable, and are therefore located in the habitable zone of their host star. Figure 3.2 shows the star-planet separation angle and contrast that would be necessary to observe an Earth-like planet with reflected light in the habitable zone around our nearest stars, out to 20 parsecs distance. Beyond this distance, direct imaging becomes impractical with current methods. There is no guarantee that these

3 Science potential

theoretical planets exist, but we can use this analysis to determine what type of tools we need to develop to push the boundaries of exoplanet science.

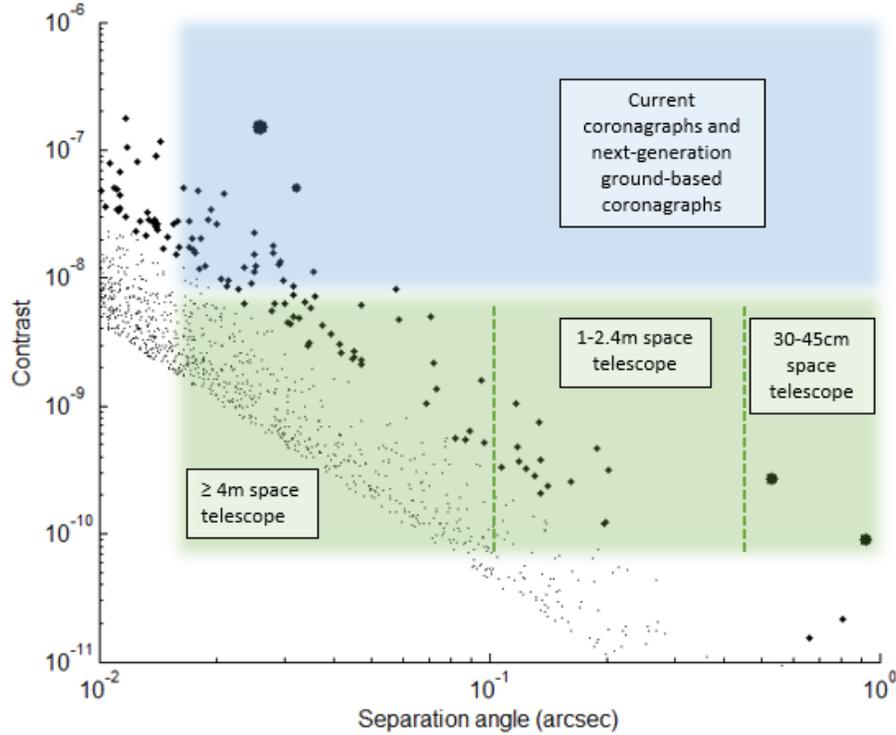


Figure 3.2: Contrast and separation angle required to directly observe a hypothetical Earth-size exoplanet with reflected light at quadrature in the habitable zone around nearby stars.

In general, the planets towards the lower left corner of Figure 3.2 are more challenging to observe, and the planets towards the upper right corner of the graph are less challenging. The size of each planet represented in the figure is inversely proportional to the distance of its host star relative to Earth, so a large dot represents a planet orbiting around a star that is located close to Earth, and a smaller dot represents a planet orbiting around a star that is located further from Earth. Pixel limits in the Figure only allow a few discrete dot sizes. A planet located around a given star that is located close to Earth can be observed with a relatively large angular separation. If

3 Science potential

the same star and planet are moved further away from Earth, the angular separation will be smaller and more challenging to resolve.

Another important characteristic is the host star's stellar type. For a faint star, such as an M-dwarf, the habitable zone is located relatively close to the star. Therefore a habitable planet would have a small angular separation, but it would be able to reflect a greater portion of the starlight, and could be observed with an easier contrast (10^{-7}). The habitable zone around a bright star is located relatively far from the star, so the angular separation of a habitable planet is greater, but the contrast is more challenging ($10^{-9} - 10^{-11}$). Future coronagraphs and systems that are currently in development will enable imaging for many of the potential planets. The potential exoplanet discoveries can be grouped into categories based on the location and size of the telescope that would be required to observe them.

Because of the effects of the atmosphere, coronagraphs are incapable of observing at deeper than 10^{-8} contrast from the ground. In space, contrast is only limited by the design of the system and the angular size of the star. Therefore all planets that require deeper than 10^{-8} contrast should be observed from space. However, it is more practical to construct a large telescope aperture on the ground. While space-based telescopes have been limited to relatively small aperture sizes, such as the Hubble Space Telescope (2.4m aperture) and the James Webb Space Telescope (6.5m aperture), the next generation of ground-based telescopes are designed to have primary mirrors that are tens of meters in diameter. The Thirty Meter Telescope (TMT) will have a 30m aperture [50], the Giant Magellan Telescope (GMT) will have a 24.5m aperture [51], and the European Extremely Large Telescope (EELT) will have a 39m aperture [52]. The large apertures can easily resolve the angular separation between many nearby stars and their habitable zones, however, there are relatively few potential Earth-size planets with a contrast that can be achieved from the ground.

3 Science potential

However, there is a unique science potential for these telescopes to observe Earth-size planets around the habitable zones of nearby M-dwarf stars. M-dwarfs provide a particularly favorable situation for observing in the habitable zone with reflected light because the stars are relatively dim, so the habitable zone is located closer to the star; this causes the contrast ratio in the habitable zone to be more favorable (10^{-7}) and within the limits of ground-based observing.

From space, there is great potential to observe Earth-like planets around many nearby stars. Some of the planets would require an aperture size that is unprecedented for space telescopes, but there is no theoretical barrier to the discovery of these planets. The space-based observations are separated into three categories in Figure 3.2 based on the size of the telescope aperture that would be required to observe them, assuming that we observe at 550nm wavelength with a coronagraph with a $2 \lambda/D$ IWA. The potential planets around our nearest star, Alpha Centauri, are particularly low-hanging fruit for exoplanet science, as they can be observed with only a 30-45 cm space telescope. The Alpha Centauri Exoplanet Satellite (ACESat) mission has been proposed to directly observe around these stars [53, 54].

For observing potential Jupiter-size planets in the habitable zone around nearby stars, the contrast is multiplied by a factor of 172. This multiplication factor is based on the square of the relative surface area for a Jupiter-size planet compared with an Earth-size planet, and the ratio of visual geometric albedo. The resulting planets, shown in Figure 3.3, are much easier to observe. In this case, there are many Jupiter-size planets that can be observed from the ground with current and next-generation telescopes. For comparison, the large planets that have already been observed with thermal emission are shown with red dots in Figure 3.3. The size of the red dots in Figure 3.3 was chosen for emphasis, and is not representative of the distance from Earth to the planet host star.

3 Science potential

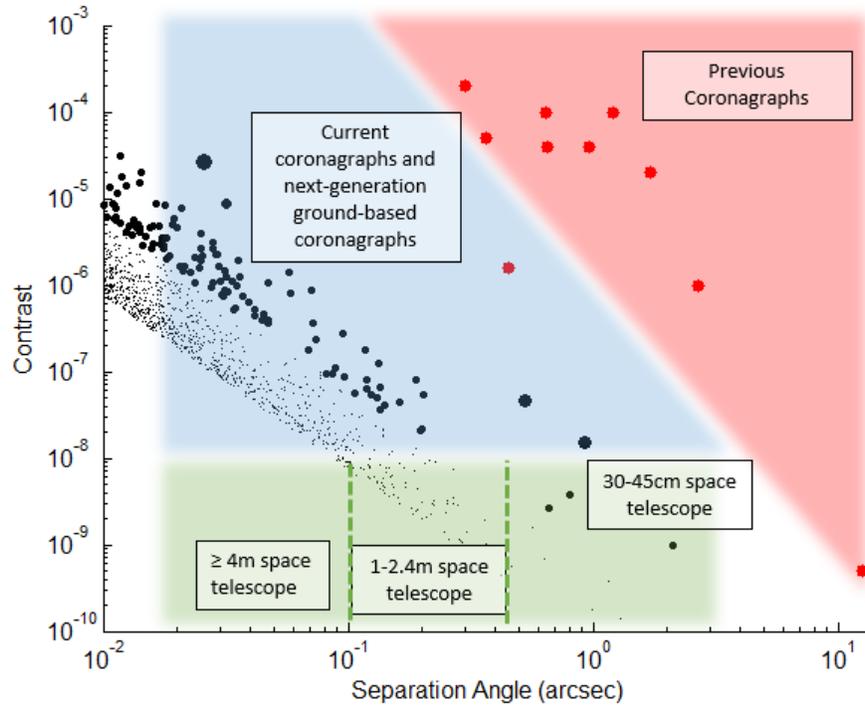


Figure 3.3: Contrast and separation angle required to directly observe a hypothetical Jupiter-size exoplanet with reflected light at quadrature in the habitable zone around nearby stars. The giant planets that have been observed in thermal emission with previous coronagraphs are shown with red dots for comparison.

In Chapters 3 and 4, we will discuss the development of technology that will enable the detection of these potential planets. We will re-visit this area for potential science and demonstrate that we have developed the tools necessary to expand the boundaries of exoplanet science, especially for coronagraphs that can be applied to the next-generation of ground-based telescopes. We can develop all of the components necessary to observe the planets using existing and proven technology.

3.3 Developing focal plane masks to advance exoplanet science

Large telescope apertures will enable new possibilities for ground-based coronagraphy, such as the direct imaging of Earth-like planets around nearby M-dwarf stars. These potential planets are shown in the lower left corner of the blue region of Figure 3.2. As telescope apertures become larger, it becomes more impractical to contain the entire primary mirror within a single piece of glass, and segmented pieces becomes more necessary. The next-generation of ground-based telescopes, such as the GMT, TMT, and E-ELT, will have large segmented apertures. As we discussed in Section 1.5, PIAACMC is a flexible coronagraph architecture that can be designed for any arbitrary aperture shape, including segmented and obscured apertures. PIAACMC is well suited to be designed for these complicated telescope apertures [55], so it is a good choice as a coronagraph architecture to be applied to these new telescopes.

The limited aperture size for space-based coronagraphs makes it especially important to have a system with a low IWA in order to maximize the available observation area close to a target star. As we discussed in Section 1.5, PIAACMC can approach the ideal IWA and useful throughput for any direct detection system. For a space-based coronagraph, the high throughput is desirable to reduce the observation time

3 Science potential

required for a given object. Mission longevity can be limited for space telescopes, so a higher throughput can result in more targets or more observations per target throughout the lifetime of the mission. PIAACMC focal plane mask development will provide a benefit for future space-based coronagraph missions, such as Exo-C [56].

There are two coronagraph systems which can immediately benefit from PIAACMC focal plane mask development. The SCExAO instrument [47] is currently in the process of implementing a set of PIAA optics that are specifically designed to be used with a PIAACMC focal plane mask. PIAACMC is also being considered for implementation with the space-based Wide-Field Infrared Survey Telescope (WFIRST) [66]. Advances in focal plane mask design, manufacturing techniques, and strategies for use with wavefront control will help these systems to reach their full scientific potential.

4 Achromatic size-scaling opaque focal plane mask

4 Achromatic size-scaling opaque focal plane mask

As we discussed in section 1.4.1, there are several coronagraph architectures which use a simple opaque disk occulter to block the on-axis starlight. This solution provides relatively simple manufacturing and alignment of the mask, but suffers from the 'size chromaticity' problem, namely, the fact that the size of a stellar PSF illuminating a coronagraphic mask will scale with wavelength, but coronagraphic masks remain fixed in size. In this chapter, we focus on a new focal plane mask concept which operates conceptually as an opaque disk occulter, but uses a phase mask technique to improve performance and solve the 'size chromaticity' problem.

The size of a typical occulting disk is designed to match the stellar PSF size at the longest wavelength in the coronagraph bandwidth, therefore blocking potential planet detection space from the shorter wavelengths. The alternative is to reduce the size of the disk; however, this allows starlight at longer wavelengths to pass around the mask and decreases the contrast. The size of an ideal mask would scale linearly with wavelength to match the size of the stellar PSF. This achromatic focal plane mask would maximize the potential planet detection space without allowing starlight leakage to degrade the system contrast. Compared with a conventional opaque disk mask, the achromatic mask allows coronagraph operation over a broader range of wavelengths and allows the detection of exoplanets closer to their host star [57, 58, 59].

We present a design for a size-scaling achromatic focal plane mask based on the diffractive optical filtering technique [60]. We manufacture and test a mask designed for the SCEXAO instrument [47, 61]. The same design principles can be applied to create an achromatic mask for other coronagraphs. We show results from laboratory testing which demonstrate the advantages of an achromatic mask compared with a conventional hard-edge opaque mask.

4.1 Achromatic size-scaling focal plane mask design

4.1.1 Achieving size-scaling with wavelength

Achromatic masks can be designed as substitutes for any coronagraph that normally uses a hard-edge opaque mask. We will focus on a mask that was specifically designed, manufactured, and tested for the SCEXAO instrument. The mask was designed to be used in the H-band ($\lambda = 1.5$ to $1.8 \mu\text{m}$) with a pupil apodized using the PIAA technique [31]. The main consequence of this setup is that when it reaches the focal plane, the beam has an equivalent focal ratio of $f/65$. Figure 4.1 shows the layout of SCEXAO and the beam intensity profile after PIAA apodization.

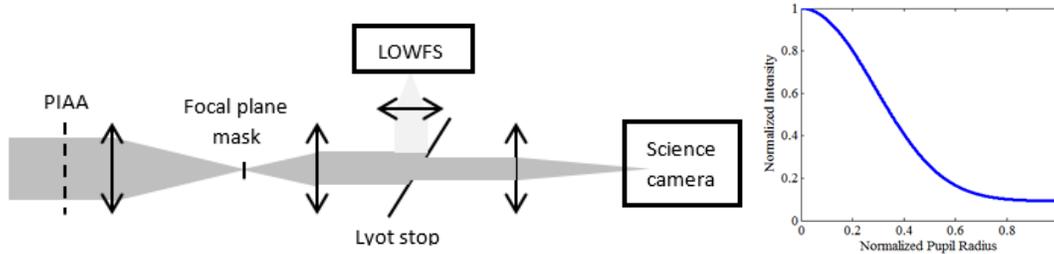


Figure 4.1: Schematic of the SCEXAO instrument (*left*) and intensity profile in the pupil plane after PIAA (*right*). The achromatic size-scaling focal plane mask diffracts starlight onto a partially reflective Lyot stop which is then used in a low-order wavefront sensor. Planet light is passed by the focal plane mask and imaged onto the science camera.

In Section 3.2, we describe how to design the opaque and diffractive structure of the focal plane mask to achieve a size-scaling effect with wavelength. The structure is based on optical filtering with a periodic phase shift as described in Section 3.3. The concept relies on an additional optical element in the coronagraph, a pupil plane stop. In any coronagraph that uses an opaque focal plane mask, starlight that passes around the edges of the mask is diffracted away from the center of the geometric pupil. This

light needs to be blocked by a separate mask, usually located in a re-imaged pupil plane after the mask, a component called the Lyot-stop. For the SCEXAO focal plane mask design, we also want the mask to be compatible with the low-order wavefront sensor, which is a measurement tool for wavefront control.

4.1.2 Compatibility with wavefront control

Any high performance coronagraph requires fine pointing control. On SCEXAO, this is achieved via a closed-loop system monitoring the light rejected by the coronagraph, forming a sub-system called the coronagraph low-order wavefront sensor (CLOWFS: [62, 63]). The starlight is rejected at the Lyot stop, so this component is made reflective so that the light can be separated from the science light, directed towards a lens, and re-imaged onto a camera. This sub-system is shown as a separated beam path in Figure 4.1.

CLOWFS monitors this image and uses it to drive a servo-loop that efficiently stabilizes the tip-tilt and other low-order modes such as focus and astigmatism [32]. The achromatic mask design offers the opportunity to further improve the sensitivity of this CLOWFS because the mask structure influences the spatial distribution of the light in the Lyot-plane. The reflective Lyot-plane stop within SCEXAO is a flat mirror with a circular hole cut through the center. Light that passes through the center hole of the Lyot stop is potential planet detection light, and it continues propagating toward the science camera. The starlight that is diffracted by the mask is incident upon the Lyot stop outside of the center hole, on the reflective area of the Lyot stop.

4.2 Mask structure

The effective size of the mask is designed to scale with wavelength throughout the system bandwidth. We will define the system bandwidth as spanning from the shortest wavelength (λ_{min}) to the longest wavelength (λ_{max}). The mask pattern, shown in Figure 4.2, is etched into a Silicon wafer, and consists of a central cone-shaped structure surrounded by an azimuthally periodic binary pattern. The cone is designed to reject all starlight incident on it as if it was an opaque disk, except instead of blocking the light directly, it diffracts it outside the pupil to be blocked by the Lyot-stop. Geometrically, the cone structure acts exactly like a conical lens. The radius of the cone (r_c) is set to the desired size of the occulter for a particular coronagraph design at λ_{min} . The desired size varies based on apodization profile, desired IWA, and desired contrast, but a typical occulter would be sized for $0.5 - 2\lambda/D$. The outer radius of the mask (r_m) is set to the desired size of the occulter at λ_{max} . The size of the stellar PSF scales with wavelength, so to keep a consistent size of the mask relative to wavelength, r_m scales with the system bandwidth. For example, a typical coronagraph bandwidth is 20%, so λ_{max} is 20% larger than λ_{min} ; in this case, r_m should be 20% larger than r_c .

There is a binary pattern around the outer edge of the mask which is similar in structure to the teeth of a circular gear. The relative height difference between adjacent teeth (h) produces a phase shift and causes the light to diffract. We describe in section 3.3 how to tune the parameter h to produce a phase shift that allows the size of the mask to scale with wavelength. The end result is that the region with binary structure is effectively opaque to long-wavelengths, yet transparent to short wavelengths, so that the size of the mask effectively scales with wavelength. The binary pattern region diffracts short wavelength light inside the Lyot stop opening, allowing it to pass to the science camera, but starlight at long wavelengths is diffracted

4 Achromatic size-scaling opaque focal plane mask

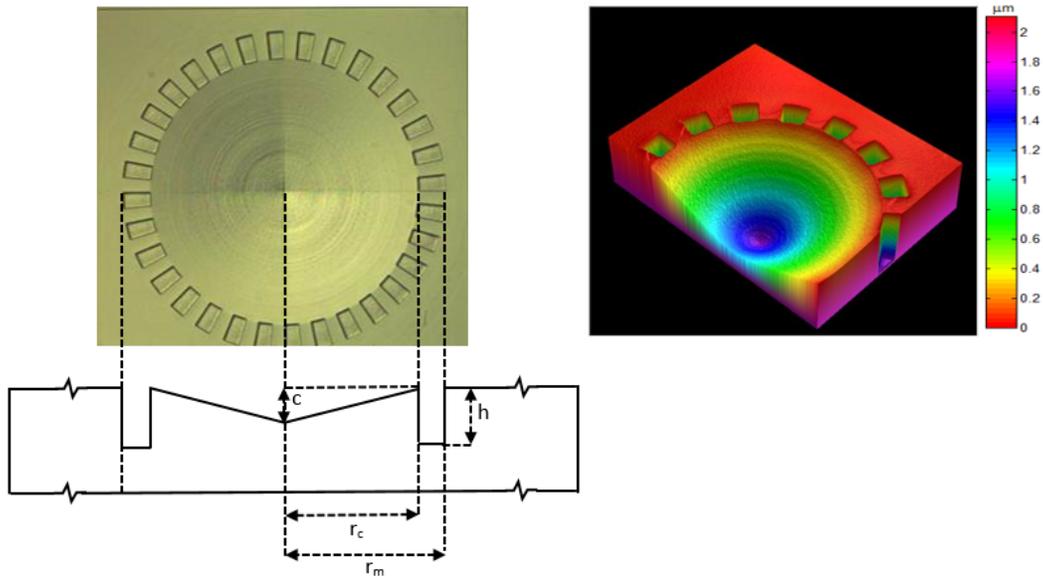


Figure 4.2: A microscope image of a mask fabricated at NASA JPL Microdevices Laboratory, along with a diagram to define geometric parameters of the mask (*left*). Isometric view of mask profilometry (*right*).

4 Achromatic size-scaling opaque focal plane mask

outside the Lyot stop, rejecting it from the system.

There are five parameters in the mask design: the radius of the cone (r_c), the depth of the cone (c), the outer radius of the mask (r_m), the depth of the teeth (h), and the number of segments in the binary pattern. Each of these design parameters has an impact on the system performance which can be described by the diffraction pattern of light in the Lyot plane. The design parameters are determined by system requirements, including the central wavelength, the bandwidth, and compatibility with the CLOWFS system. Light incident upon the central cone of the mask is always diffracted into a ring in the Lyot plane. The radius of the cone determines the amount of the stellar PSF at λ_{min} that is diffracted, and therefore blocked from the science camera. r_c should be set to match the desired size of the occulter at λ_{min} . The angle of the cone, computed by the ratio c/r_c , determines the radial position of diffracted light in the Lyot plane. A steeper angle moves the starlight further from the optical axis, and a shallow angle moves the starlight closer to the optical axis. If one is only concerned with removing the starlight from the system, then a steep angle within manufacturing limits would be appropriate. However, the rejected starlight can be useful for CLOWFS, so we adjust the angle of the cone to send the rejected starlight into collecting optics for the CLOWFS. In the case of SCEXAO, the starlight is collected in a tight circular area directly surrounding the outer extent of the Lyot mask.

Light at longer wavelengths that is incident upon the periodic binary structure is also diffracted into a ring in the Lyot plane, centered on the optical axis. The number of segments in the binary structure (and therefore the frequency between adjacent segments) determines the radial position of the diffracted light in the Lyot plane. An increase in the number of binary segments will move the diffracted light further from the optical axis. This concept is shown in Figure 4.3, where the images show the

4 Achromatic size-scaling opaque focal plane mask

diffraction pattern produced by the binary structure made with 14 segments (left) and 30 segments (right). For a greater number of segments, the diffraction pattern is located further from the Lyot stop boundary, which is shown with a white circle. To simply remove the diffracted starlight from the system, a high number of segments is appropriate within manufacturing limits.

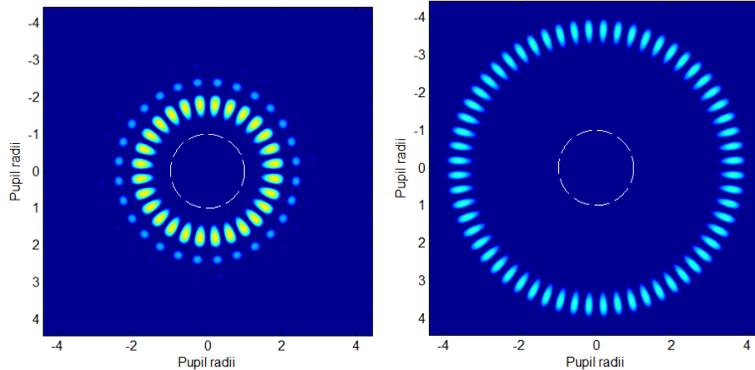


Figure 4.3: Diffraction pattern produced by the binary structure of the achromatic size-scaling mask when it is illuminated with an apodized stellar PSF. When the binary pattern is made of 14 segments (*left*) the diffracted light is located closer to the optical axis compared with a binary pattern made of 30 segments (*right*).

However, this diffracted starlight can also be useful for CLOWFS, so for SCEXAO, we adjust the number of segments to send the starlight into a tight ring surrounding the outer extent of the Lyot stop. A typical number of segments is between 14 and 30. The SCEXAO mask is optimized to position diffracted light between 4.5 and 12.5 mm radial distance from the optical axis in the Lyot plane, which is 1.0 and 2.8 pupil radii, respectively, and corresponds to the reflective region of the Lyot stop. The only remaining parameter of the mask design is the relative depth between adjacent segments of the binary structure, h . The segment depth, which creates a phase shift between adjacent segments, is a calculated parameter based on the system bandwidth.

4.3 Phase shift

The mask is modeled as a pure phase mask (i.e., amplitude transmission is 1). Any space not occupied by the profile of the silicon mask is occupied by air, so the effective phase shift of the mask is given relative to the phase shift that would have occurred if the mask were made of air. A generic calculation of phase shift can easily be achieved for any part of the mask, but here we focus on the periodic binary pattern at the edge of the mask because it determines the achromatic size-scaling property of the mask. The relative phase shift ($\Delta\phi$) between adjacent binary segments is given in Equation 4.1 as a function of the segment depth (h) and the refractive index of silicon (n_{Si}) for a particular wavelength (λ). The phase shift is computed in units of wavelengths. We make the approximation $n_{air} = 1.000$ because any further significant digits of the mask thickness usually amount to sub-nanometer levels that cannot be accurately manufactured.

$$\Delta\phi = (n_{Si}(\lambda) - 1) * h/\lambda \quad (4.1)$$

In the pupil plane, the Lyot stop acts as a low-pass filter, and blocks light that is diffracted far from the center of the optical axis, as shown in Figure 4.3. Equivalently, one can think of this process in the focal plane, where the Fourier Transform of the Lyot stop function (an Airy disk) is convolved with the electric field in the focal plane after it is modulated by the achromatic focal plane mask. Convolution with the Airy disk smooths the transition between adjacent segments in the mask binary structure, and selects only the low spatial frequencies of the mask where the electric field is mixed between adjacent segments. This low-pass filtering process operates in combination with a phase shift between adjacent mask segments to produce a size-scaling effect at the edge of the focal plane mask.

The size of an ideal focal plane mask would scale linearly with wavelength. However,

4 Achromatic size-scaling opaque focal plane mask

we are not aware of a phase profile as a function of wavelength that would produce a linear scaling of the mask size. Our best solution is to create constructive interference at the shortest wavelength in the observing band, and destructive interference at the longest wavelength in the observing band; this strategy creates an ideal size of the mask for λ_{min} and λ_{max} . In between the two extreme wavelengths, a monotonic scaling of phase shift produces a monotonic scaling of the effective mask size. The effective size is not necessarily the physical size of the mask at a given wavelength, but the size of the mask as seen through the Lyot stop, including the effects of constructive or destructive interference. For the SCEXAO mask, we compute h for a 20% bandwidth in the H-band, where $\lambda_{min} = 1.5\mu\text{m}$ and $\lambda_{max} = 1.8\mu\text{m}$. Equation 3.1 is plotted in Figure 4.4 for $h = 1.81\mu\text{m}$. At λ_{min} the phase shift is 2.995 wavelengths, which is close to a 6π phase shift, and at λ_{max} the phase shift is 2.475 wavelengths, which is close to a 5π phase shift. Therefore, adjacent segments in the binary structure produce constructive interference for λ_{min} , causing that part of the mask to behave as if it were transparent, and destructive interference for λ_{max} , causing that part of the mask to behave as if it were opaque. As shown in Figure 4.4, the relationship between wavelength and phase shift is monotonic within the H-band. This results in a mask with an effective size that scales monotonically with wavelength. We could create an achromatic mask for any of the slopes in Figure 4.4 that are highlighted in green, because there is constructive interference at λ_{min} and destructive interference at λ_{max} . The region in Figure 4.4 that is highlighted in red has slopes with the opposite effect compared with what is useful for an achromatic mask, destructive interference at λ_{min} , and constructive interference at λ_{max} . The phase curves keep a similar pattern, but shift for different values of h and for materials with different refractive indices. However, there is clearly a limit to the bandwidth over which this mask design can produce an effective phase shift that scales with wavelength; it would

be difficult to produce an achromatic design with greater than 20% bandwidth.

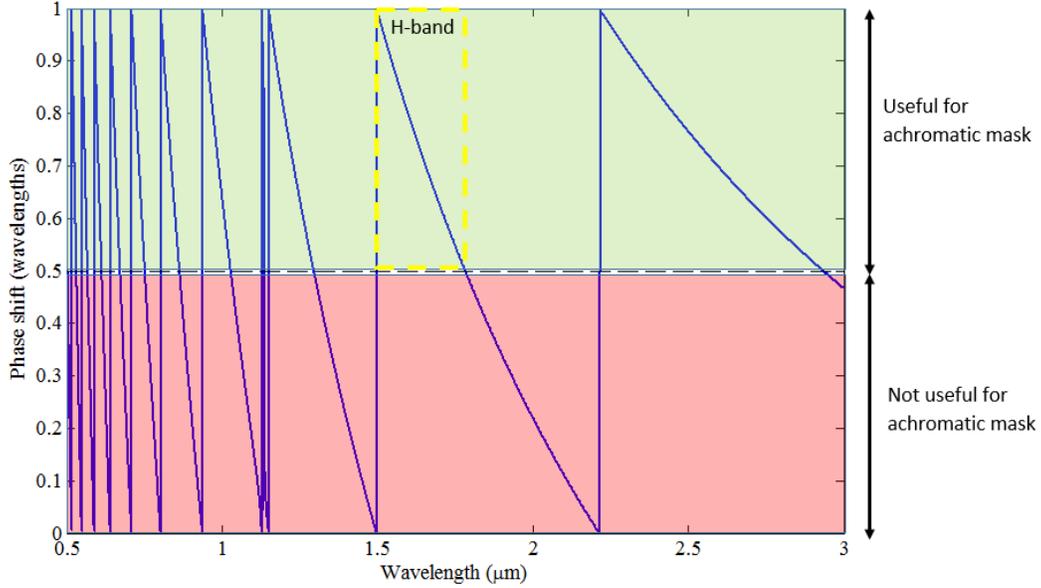


Figure 4.4: Phase shift produced by adjacent cells in the achromatic mask binary pattern as a function of wavelength for $h = 1.81\mu\text{m}$. The phase is wrapped back to zero for shifts that are greater than one wavelength. Within the H-band (*dashed yellow box*), the phase shift is monotonic.

4.4 Fabrication

The mask design contains small features that require precise nanofabrication. The NASA Jet Propulsion Laboratory (Microdevices Laboratory) has extensive experience making focal plane masks using various techniques [64]. They fabricated the achromatic masks using gray-scale electron beam lithography and plasma transfer etching. An array of masks were etched onto a silicon substrate with H-band antireflection coating on one side. It is beneficial to have masks with various sizes available on the same substrate, so we created four scaled versions of the same mask design

with $r_c = 130, 163, 203$ and $254 \mu\text{m}$, which represent approximately $1\lambda/D$, $1.25 \lambda/D$, $1.56 \lambda/D$, and $1.95 \lambda/D$ for the apodized beam described above, respectively. Balasubramanian et al. [64] provide a more detailed description of the mask fabrication process.

4.5 Characterization of size-scaling with wavelength

We can verify the size-scaling principle of the focal plane mask as a function of wavelength using simulations. For the first simulation, we create an apodized PSF which models the stellar PSF that is achieved in an actual coronagraph. At first, the PSF is centered on the mask, and the mask is blocking most of the starlight as it would in normal coronagraph operation. The PSF is then moved off axis in regular increments to simulate an off-axis source at various distances relative to the center of the mask. We measure the throughput of the PSF at each increment and calculate the point at which the throughput reaches 50%, known as the IWA. The IWA is considered to be the effective edge of the mask. The simulation is repeated for the shortest, middle, and longest wavelengths within the design bandwidth. Each throughput curve is normalized relative to the maximum throughput that is achieved when the PSF is completely unobstructed by the mask. Results are shown in Figure 4.5 for the $1\lambda/D$ achromatic mask designed for SCExAO, which has $r_c = 130\mu\text{m}$ and $r_m = 156\mu\text{m}$, and compared with a hard-edge mask with radius $156 \mu\text{m}$. This represents a typical situation where the hard-edge mask is sized for the longest wavelength in the observing bandwidth.

For both types of focal plane mask, the throughput is low when the mask is directly occulting the PSF. The throughput increases near the edge of the mask, and reaches a maximum when the PSF is no longer occulted by the mask. For the achromatic mask, the IWA is at nearly the same location for all wavelengths (in units of λ/D).

4 Achromatic size-scaling opaque focal plane mask

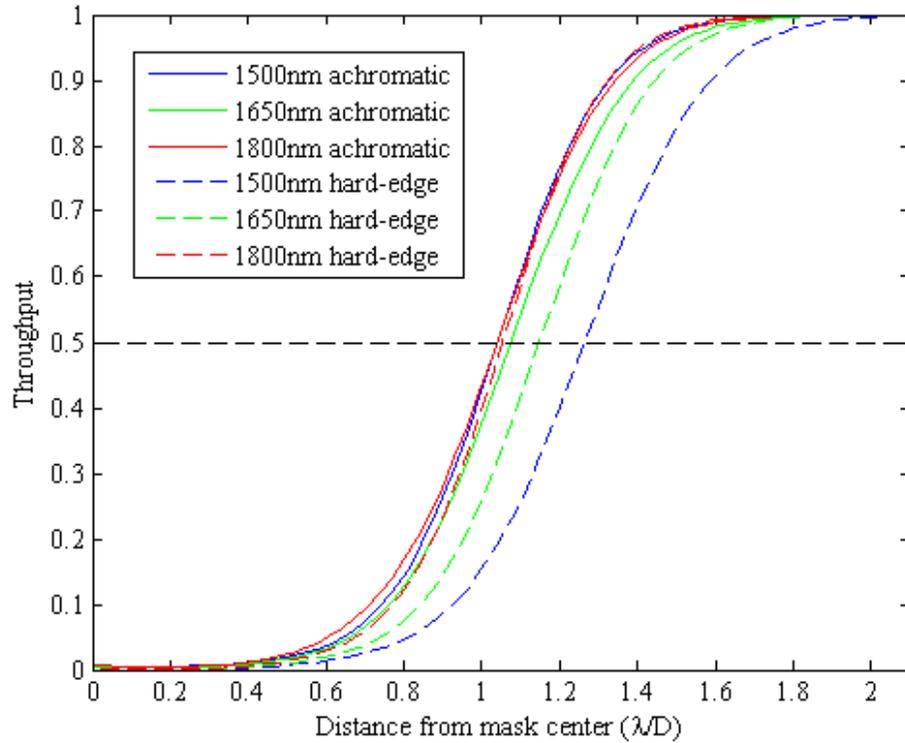


Figure 4.5: First simulation of the size-scaling property of the achromatic mask as a function of wavelength (*solid lines*). Results are compared to a conventional hard-edged mask (*dashed lines*). We measure throughput of an off-axis source and compare IWA for different wavelengths.

4 Achromatic size-scaling opaque focal plane mask

Over the entire 20% observational bandwidth, the effective size of the mask deviates by only 2.8% in units of λ/D . If the mask could provide a perfectly linear phase shift as a function of wavelength over the entire bandwidth, the chromatic dispersion would be 0%. However, as shown in Figure 4.4, the phase shift relationship is not exactly linear, and does not provide perfect destructive interference at $\lambda_{max} = 1.8\mu\text{m}$. In this simulation, the maximum deviation actually occurs for the middle wavelength. Despite the imperfect phase shifts, the achromatic mask provides a clear advantage over a hard-edge mask. For the conventional hard-edge opaque mask, the IWA scales by 20%. This is an expected result for a 20% bandwidth because the mask has a fixed size.

A similar experiment is easier to replicate within the SCExAO instrument. For this second experiment, the mask and the PSF remain on-axis as they would in normal coronagraph operation. In this case, we measure the intensity of the residual light as a function of distance relative to the center of the mask, which is also the center of the PSF. We intentionally use a PSF that is oversized with respect to the mask so that there is significant light leakage around the edge of the mask. Within SCExAO, we create an oversized PSF by stopping-down the light beam in a pupil plane. The PSF is apodized, so the intensity of the light leakage decreases with distance relative to the center of the mask. Therefore, the highest intensity of the light leakage occurs directly after the edge of the mask, and we can use this fact to identify where the edge of the mask is located. The experiment is repeated for different wavelengths, and we expect to see behavior that is similar to the effects shown in Figure 4.5. Each intensity curve is normalized relative to the greatest intensity within the curve. This simulation does not exactly measure throughput because the source remains on-axis. However, we can use the point of 50% maximum intensity of light leakage as a sort of modified inner working angle (mIWA) to compare the location of the edge of the mask for different

4 Achromatic size-scaling opaque focal plane mask

wavelengths. The mIWA is not a useful real-life performance metric of any scientific value, but it can be used as a metric to demonstrate the size-scaling property of the achromatic mask within SCExAO. We can then compare our measured mIWA from the achromatic mask with an mIWA that we would expect from a hard-edge mask.

The mIWA measurement was completed within SCExAO using a broadband supercontinuum source. To select only a narrow range of wavelengths at a time, we used bandpass filters centered at 1570, 1610, and 1650 nm wavelengths, each with FWHM of 12 nm. Figure 4.6 shows the resulting intensity measurements as the $r_c = 203\mu\text{m}$ achromatic mask was occulting the apodized PSF, along with a simulation for comparison. Unfortunately, the signal-to-noise ratio of the measurement was somewhat low due to the maximum output power of the supercontinuum source. However, we can identify features within the intensity curves that indicate the size-scaling effect of the achromatic mask. The mIWA dispersion is 2.8% in simulation and 3.2% in measurement, both of which are far less than what we would expect for a hard-edge mask. This is an important result because it shows that the size of the achromatic mask scales with wavelength.

To verify the difference in mIWA dispersion between an achromatic mask and a hard-edge mask, we completed the second experiment in simulation. Results are shown in Figure 4.7 for the $1\lambda/D$ achromatic mask compared with a hard-edge mask with radius $156\mu\text{m}$. Results of the simulation are somewhat different than those presented in Figure 4.6 because the experiment included parameters that were specific to the SCExAO optical layout.

The intensity curves in Figure 4.7 have a similar profile compared with the throughput curves of the first simulation shown in Figure 4.5. One noticeable difference is that the intensity at zero distance relative to the mask center is much higher. This occurs because we are measuring light leakage, which has far less intensity than the

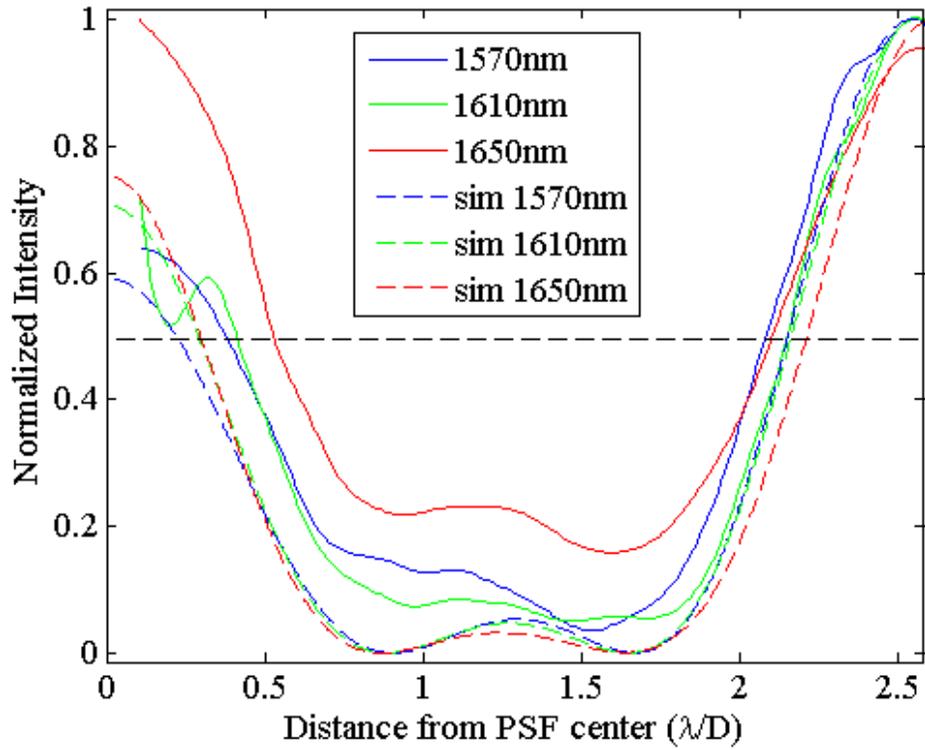


Figure 4.6: Intensity as a function of distance from the center of the mask in measurement (*solid lines*) and simulation (*dashed lines*) for various wavelengths. The location of intensity increase, representing the effective edge of the mask, scales with wavelength.

unobstructed PSF that is used for normalization of the throughput profiles. Another noticeable difference is that the mIWA occurs at a slightly greater distance than the IWA. Again, this is due to the fact that the light leakage reaches a maximum at a distance slightly beyond the edge of the mask. In this simulation, over the entire 20% observational bandwidth, the mIWA deviates by only 3.9% in units of λ/D . For the conventional hard-edge opaque mask, the mIWA scales by 17%. The mIWA dispersion numbers are close to the IWA dispersion, but differ slightly because the source remains on-axis. We can use this experiment to clearly demonstrate the size-scaling property of the achromatic mask. The achromatic mask has an advantage over the hard-edge mask for observing exoplanets because the achromatic mask is not oversized for short wavelengths in the observational band.

4.6 Performance in a coronagraph

The first step to verify the operation of a focal plane mask for coronagraphy is to test its response to monochromatic light. We simulated a stellar point source within SCEXAO by using a 1570 nm laser diode directed through an optical fiber. The simulated stellar PSF was apodized using PIAA lenses and focused onto the $r_c = 254\mu\text{m}$ mask. The science camera image shown in Figure 4.8 was recorded while the mask was blocking the source PSF. Images for the other three mask sizes show the same features, but have a different sized dark zone in the center of the image that corresponds to the size of the mask. The color scale represents contrast in the focal plane relative to the highest intensity pixel in an unobstructed PSF. The average contrast between 3.8 and 7.7 λ/D distance from the center of the coronagraph is 1.47×10^{-5} . From simulations, we expect to achieve an average contrast of 4.26×10^{-8} within the same boundaries. The brightness of the central peak is close to what we expect in simulation, but the diffraction spikes and system aberrations degrade the

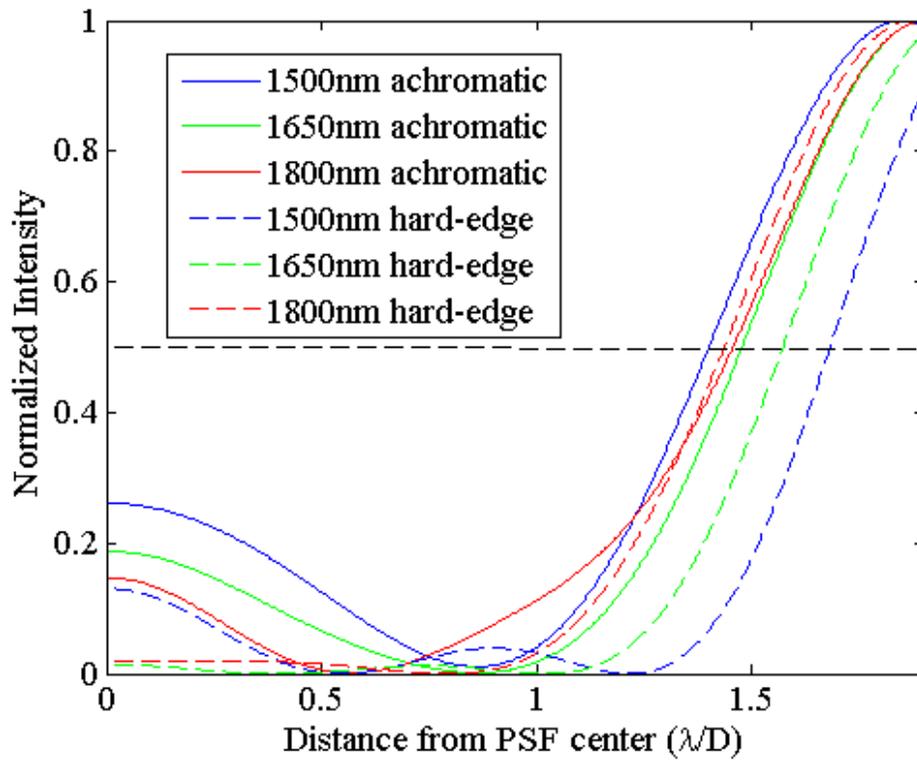


Figure 4.7: Second simulation of the size-scaling property of the achromatic mask as a function of wavelength (*solid lines*). Results are compared to a conventional hard-edged mask (*dashed lines*).

4 Achromatic size-scaling opaque focal plane mask

contrast further from the center of the coronagraph. The raw contrast presented is without the use of wavefront control, but we plan to utilize wavefront control in future iterations to achieve deeper contrast.

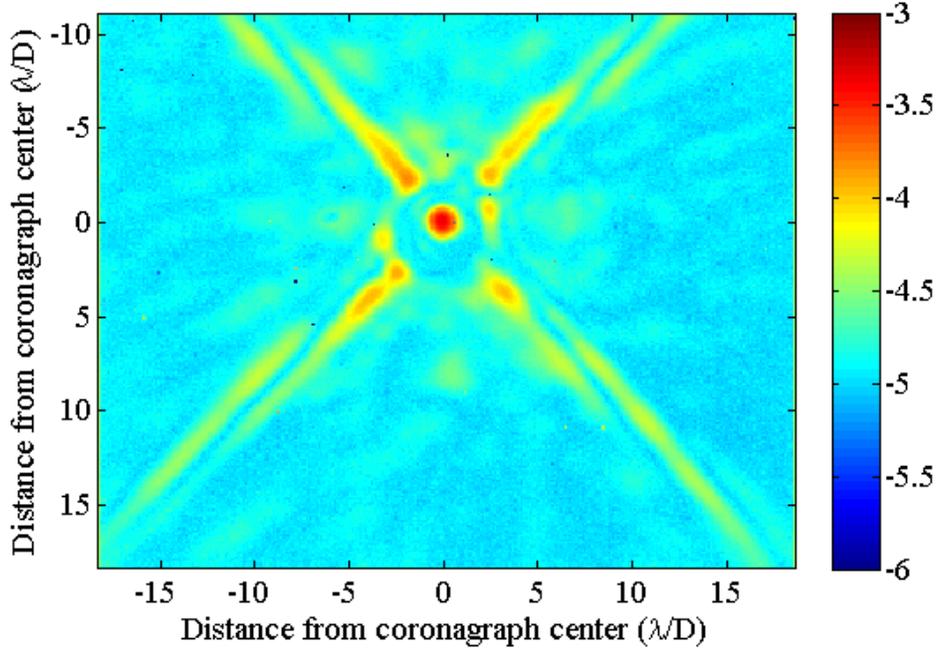


Figure 4.8: Image of focal plane mask contrast recorded using the SCExAO internal calibration source. The color scale is logarithmic and spans three orders of magnitude.

We observe two important features in Figure 4.8 that degrade the contrast: a central spot and four diffraction spikes. Manufacturing limits at the peak of the cone structure of the mask produce a flat surface that allows a small portion of the PSF to pass unobstructed. Unfortunately, the flat surface occurs at the point of greatest intensity of the PSF and produces a bright spot behind the center of the mask. We

4 Achromatic size-scaling opaque focal plane mask

plan to mitigate this effect in the future by making a small portion of the center of the mask opaque.

The four diffraction spikes are caused by the support beams of the secondary mirror of the Subaru Telescope. When testing with the SCEXAO internal calibration source, we included a pupil plane mask to simulate the effects of the Subaru Telescope central obscuration and support beams. However, we did not create a pupil plane mask to block the Subaru support beams after PIAA remapping. This is a major cause of light leakage into the science camera, and hence, the support beams are clearly visible in the images. Diffraction from the support beams degrades contrast over the entire image. Other blotches of light in the image are speckles caused by imperfections in the optical components, and these speckles further degrade the contrast. Speckles can be corrected using active wavefront control methods such as "speckle nulling" [26], or EFC [28]; however, these procedures were not utilized for testing the masks. We plan to achieve better contrast by installing a modified pupil plane stop to block diffracted light from the support beams and using speckle nulling to eliminate residual speckles.

As we discussed in Section 3.2, the focal plane mask was designed to diffract starlight into a ring structure in the Lyot plane so that it can be rejected by the coronagraph and used with CLOWFS. Ideally, the vast majority of the starlight would be diffracted into a tight ring on the Lyot stop mirror as shown in the simulation in Figure 4.9. The dashed white circle in Figure 4.9 marks the extent of the hole within the Lyot plane mirror, which is 0.56 pupil radii in size, and any light falling within this region is allowed to pass to the science camera. Our experimental data shows the same ring structure, but there is far more starlight remaining within the encircled region, which is caused by diffraction from the central obscuration and secondary support beams in the telescope pupil. Only 0.41% of the starlight is allowed to pass through the Lyot stop, but this is enough to cause degradation of the contrast at the

science camera. We can also see some vignetting in the measured Lyot plane image due to opto-mechanical mounts; however, this does not affect the operation of the mask or CLOWFS.

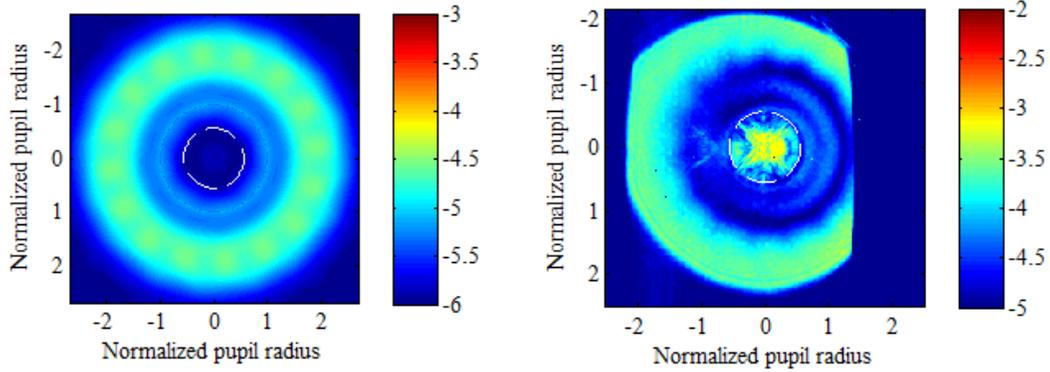


Figure 4.9: Simulation (*left*) and measured image (*right*) of starlight diffraction pattern in the Lyot plane. The logarithmic color scale spans three orders of magnitude and represents contrast relative to the maximum intensity of an unobstructed PSF. The scale for the measured image is offset by an order of magnitude relative to the simulated scale.

In this chapter, we have demonstrated an achromatic version of the opaque disk focal plane mask, which is a simple amplitude mask used in coronagraphs such as PIAA and APLC. When used in broadband science operation, the size-scaling achromatic mask provides a smaller IWA at the shorter wavelengths of the observational band, and opens up new observation space with the potential to discover and characterize additional exoplanets. This is a particularly attractive feature because the number of planets orbiting around a star is expected to increase close to the star. In addition, the star’s habitable zone is generally located relatively close to the star, and we are especially interested in observing planets within this region. The achromatic mask is an easy upgrade for coronagraphs which use an occulting disk because the mask can be implemented into these coronagraphs without any other design changes. Overall,

4 Achromatic size-scaling opaque focal plane mask

the size-scaling mask is a very useful solution for an achromatic amplitude mask. In the next chapter, we will apply similar design principles to develop an achromatic solution for phase masks. Specifically, we will demonstrate an achromatic phase mask for the PIAACMC architecture.

5 Phase masks for PIAACMC

5 Phase masks for PIAACMC

As we discussed in Section 1.5, the PIAACMC architecture is an ideal coronagraph solution for many current and future telescopes, and can achieve near the performance limit for any direct imaging system. The PIAACMC architecture is shown in Figure 5.1 with an example unobscured aperture. The telescope pupil is apodized by PIAA optics and focused onto the complex phase-shifting focal plane mask. Some of the starlight is phase-shifted by the mask, while some of the starlight passes around the mask and is not shifted. The light is re-collimated and filtered through a Lyot stop in the exit pupil plane, which can be as simple as an iris sized to match the size of the original input beam. The Lyot stop acts as a low-pass filter, and forces the phase shifted starlight to interfere destructively with the un-shifted starlight. At this point, the remaining light can be focused onto the science plane to directly image the area surrounding the nulled on-axis star. However, the apodization creates distortion in the off-axis image, so a set of inverse PIAA optics is often used to reverse the apodization and recover good off-axis imaging quality.

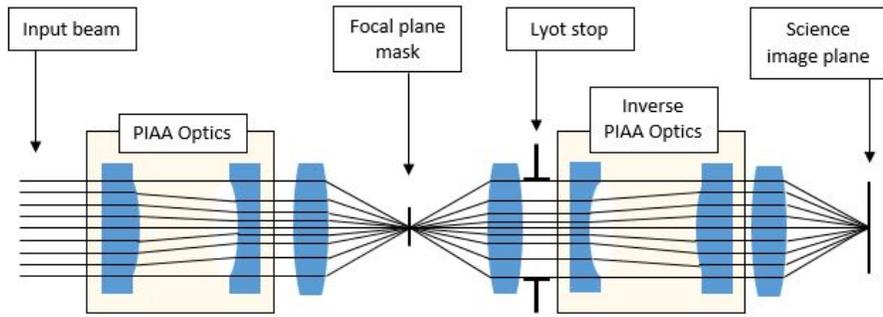


Figure 5.1: Schematic of PIAACMC system architecture.

PIAA optics have been well characterized and tested over the past few years, reaching contrast on the order of 1×10^{-8} between 2 and $4 \lambda/D$ in a 10% bandwidth with an opaque focal plane mask [65]. In this chapter, we show recent development of the phase shifting focal plane mask; specifically the design, manufacturing, and testing of

the mask to operate in a broad range of wavelengths. PIAACMC masks are also being developed as part of the WFIRST technology development [66]. The work shown here is complimentary to the WFIRST efforts because we focus on mask designs for the Ames Coronagraph Experiment (ACE) testbed [41] [42], and also generalize the design and manufacturing process enough that it can be applied to other systems. A PIAACMC mask has a complex transmission which ranges from 0 (opaque) to -1 (purely phase-shifting). For a monochromatic design with a purely transmissive mask, the focal plane mask could be as simple as a half wave plate that is sized to provide a π phase shift to half of the starlight. However, the broadband design is not so simple. The size of the mask should scale with wavelength. In Chapter 3, we discussed a solution to this challenge for a purely opaque mask. A phase mask has the added challenge of providing a π phase shift and partial complex transmission at every wavelength within the observing bandwidth.

Because both the size-scaling chromaticity challenge and the complex transmission chromaticity challenge need to be solved simultaneously, we are not aware of any analytical solution for the mask design. Instead, we use optimization to find a solution that is close to ideal. We will describe the mask design process in section 4.1, the manufacturing process in section 4.2, and show simulated mask performance based on manufacturing results in section 4.3. We will apply wavefront control to correct for mask manufacturing errors in section 4.4, and analyze the science potential of the resulting masks in section 4.5.

5.1 Design

The basic function of a PIAACMC mask is to provide a π phase shift and partial transmission to some of the starlight at all wavelengths in the observational bandwidth. We are not aware of any analytical solution for a mask structure profile that

5 Phase masks for PIAACMC

can achieve this. Instead, we get close to an ideal solution by optimizing the structure of the mask to provide the required phase shift. We start by choosing a mask size (typically 30-100 μm radius) which corresponds to the given apodization function and the desired IWA, as described in [37]. We then need to make the mask achromatic to operate within the entire range of wavelengths in the system bandwidth. To achieve an achromatic mask, we divide the mask into a collection of zones, for example concentric rings. A schematic of the process is shown in Figure 5.2, where each Fourier Transform operation between the focal and pupil planes is labeled as 'FT'.

Each individual zone is multiplied by the stellar focal plane field to calculate the location and portion of starlight that is affected by that zone. We compute the projection of the starlight within that zone in the Lyot plane, which is the exit pupil plane of the coronagraph system, and eliminate any light that falls beyond the boundary of the Lyot stop, which is displayed as a dashed yellow circle in Figure 5.2. The Lyot stop is sized to match the size of the entrance pupil, and any light falling beyond the boundary of the stop is rejected from the coronagraph and usually used for wavefront control, as described in [32]. The starlight that is passed by the Lyot stop is focused onto the final image plane. We compute the distribution of this starlight within a pre-defined optimization region; the optimization region is a sub-set of the entire image plane that is representative of the desired dark region for observing exoplanets (typically 2-10 λ/D), and the smaller size makes it easier to optimize.

We compute the electric field, E , as shown in Equation 4.1 for all values of x and y within the optimization region, and at a normally distributed selection of wavelengths λ within the design bandwidth. The height of each zone is designated as h_k , and n is the refractive index of the mask substrate. Equation 4.2 gives the total light intensity (I_{tot}) within the optimization region, so the optimization algorithm is designed to minimize this equation. We use a downhill simplex local optimization which is pro-

5 Phase masks for PIAACMC

grammed using the *fminsearch* function within Matlab. The local optimization begins at a starting point of zone heights which is randomly selected from a normal distribution between the acceptable zone height boundaries. The zone height boundaries are mainly constrained by manufacturing limits; a typical boundary would be from 0 to 2 microns depth. The optimization algorithm is repeated with more random starting points, typically up to 1000 iterations or until a target system contrast is achieved. The selection of the optimization region is dependent on the desired inner and outer working angles, typically from $x = 0$ to $10 \lambda/D$ and $y = 2$ to $4 \lambda/D$. One potential way to improve the optimization algorithm would be to implement a global optimizer to solve for the zone heights.

$$E(x, y; \lambda) = E_0(x, y; \lambda) + \sum_{k=1}^{N_{zones}} E_k(x, y; \lambda) [exp(i2\pi h_k(n(\lambda) - 1)/\lambda) - 1] \quad (5.1)$$

$$I_{tot} = \sum_{\lambda} \sum_x \sum_y |E(x, y; \lambda)|^2 \quad (5.2)$$

The optimized heights, h , of each zone in the mask are directly related to the complex field, ϕ , produced by the zone, as given in Equation 4.3 where n_{sub} is the refractive index of the mask substrate. The five parameters: mask size, mask material, number of zones, zone shapes, and zone heights; completely describe the mask structure. The zone shapes and zone heights are actually sets of several parameters that will be individually optimized for each zone. For manufacturing purposes, we describe the mask by its zone shapes and heights. However in simulation, the mask can be completely described by the complex field that it produces at a given wavelength.

$$\phi = exp(2\pi i(n_{sub} - n_{air})h/\lambda) \quad (5.3)$$

There is no theoretical restriction on the number of zones or shape of the zones used to create the mask. However, there are some practical considerations that affect

5 Phase masks for PIAACMC

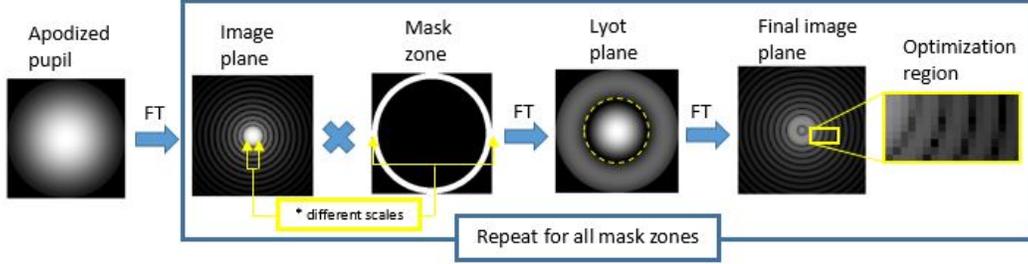


Figure 5.2: Mask design process: compute the complex amplitude in the final image plane for each zone of the mask.

the speed and performance of the optimization algorithm based on the coronagraph system configuration. We generally choose a zone shape that is easy to compute. For a circularly symmetric system with no obstruction, choosing a circularly symmetric mask structure can reduce the computation time. For a given mask size, a large number of small zones provides more degrees of freedom for optimization, and can lead to better performance. However, a large number of zones takes a longer time to optimize, and small zone features can be difficult to manufacture and measure. A good strategy for choosing the number of zones is to start with a large number, and repeat a time-constrained optimization with smaller numbers until the performance starts to degrade significantly. This, in effect gives us the smallest number of zones for which we can achieve a close to optimal solution under a reasonable time constraint.

We designed several PIAACMC masks for the existing PIAA optics at the ACE testbed. The PIAA optics are described in [41], and the apodization profile is shown in Figure 5.3a. The mirrors were originally designed for a traditional PIAA system, with a strong apodization profile to be used with an opaque occulter. As we discussed in Section 1.5, a PIAACMC system typically has a milder apodization profile [37], so for some mask designs we stop down the pupil and only use the inner portion of the PIAA optics that have a milder profile. We provide PIAACMC mask designs for

5 Phase masks for PIAACMC

both the full and the stopped-down apodization profile, although the stopped-down profile designs are more representative of a PIAACMC system.

Because the current PIAA system at ACE is unobstructed and circularly symmetric, we chose to divide the mask into a series of concentric ring zones. After trying different numbers of rings, we found that ten rings is adequate to provide the necessary phase shift over a 10% bandwidth centered at 580nm with adequate contrast. The structure of the mask zones is shown in Figure 5.3b. A one dimensional profile of the optimized zone heights is shown in Figure 5.3c for three different mask designs. All masks were designed to be used in reflection, but they could have alternatively been designed for transmission with different zone depths depending on the refractive index of the mask material. Design 1 uses the full PIAA apodization, has a maximum zone depth of $1\mu\text{m}$, and achieves 4.56×10^{-5} average contrast between 2 and 6 λ/D . Design 2 uses the stopped-down PIAA apodization, has a maximum zone depth of $1.2\mu\text{m}$, and achieves 5.64×10^{-7} contrast between 1.5 and 6 λ/D . Design 3 also uses the stopped-down PIAA apodization, has a maximum zone depth of $1.2\mu\text{m}$, and achieves 1.90×10^{-6} contrast between 1.1 and 6 λ/D .

Because Design 1 uses the full PIAA pupil with a strong apodization, the performance is limited to a relatively large IWA (2 λ/D) and relatively poor contrast as compared with the other designs. In this case, the phase mask of PIAACMC does not provide a significant advantage in achievable contrast when compared with an opaque occulter: this result was expected as described in [37], and the system is effectively similar to a traditional PIAA coronagraph. However, with the stopped-down pupil, which provides a milder apodization profile, the performance of Designs 2 and 3 are greatly improved. The primary difference between Designs 2 and 3 is the chosen IWA for the optimization region.

Our most advanced method for manufacturing the masks is electron beam lithog-

5 Phase masks for PIAACMC

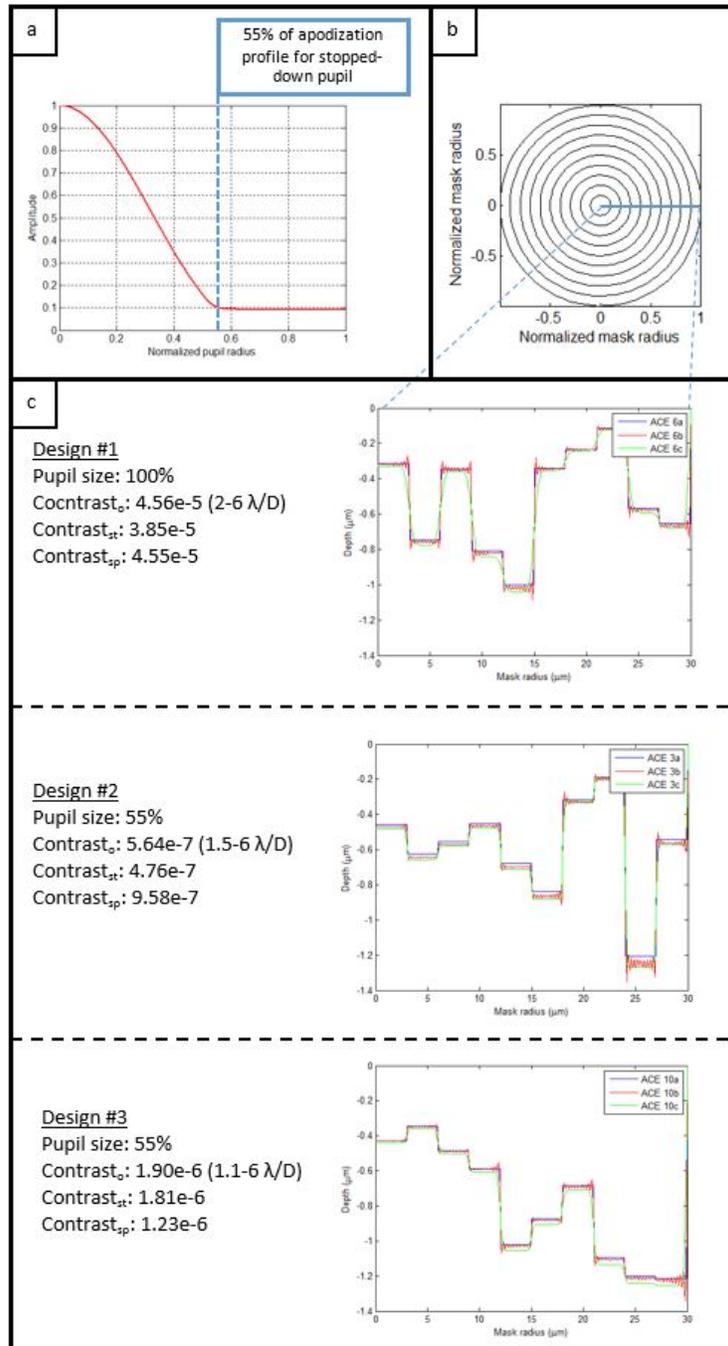


Figure 5.3: Apodization profile of the PIAA optics at the ACE lab (a). Zone geometry of a PIAACMC focal plane mask with ten concentric rings (b). Mask designs for PIAA optics at the ACE lab (c), including ideal designs (blue curve), ideally smoothed designs (red curve), and practically smoothed designs (green curve).

raphy and etching. There are a few guidelines for the mask design that make the manufacturing process easier and more accurate. The manufacturing process does not have infinite etch depth, so we set a limit on the maximum depth of a single zone, usually less than two micron. If the coronagraph system can accommodate a reflective mask, the zone depths are generally shallower because the phase shift is doubled in a double-pass mask system. The features of a reflective mask are generally a factor of 4 shallower than a transmissive mask for a mask material with a refractive index of 1.5. For a mask material with a high refractive index, such as Silicon ($n=3.4$ at $\lambda=1.5\mu\text{m}$), there is not a reduction in the feature depth for a reflective mask design. Also, it can be challenging to measure mask zones that have a larger depth than the zone width, so we set a minimum width-to-depth aspect ratio of at least one for all zones. These guidelines may be adjusted as we gain additional experience with mask manufacturing.

5.1.1 Smoothing mask profiles

The design process described in section 4.1 creates a mask profile with discrete zones. Adjacent zones almost always have different heights, and the transition between zones is only the size of a single pixel in simulation, which creates sharp edges at the zone boundaries. Kern et al. [66] has demonstrated that the sharp edges can be manufactured accurately using electron beam lithography, and that any errors in manufacturing the sharp features does not limit the performance of the mask. If sharper edges were needed, we could explore the use of a Bosch process for etching. However, smoothing out the sharp edges in a mask design could open the possibility for using alternative manufacturing techniques that cannot create sharp edges.

Guyon et al. [67] proposed a method for smoothing out the sharp edges without degrading the mask performance. The idea utilizes the fact that there is already a

5 Phase masks for PIAACMC

Lyot stop in the system after the focal plane mask. The Lyot stop acts as a low pass filter, so high frequencies in the focal plane are rejected from the system. We can smooth the sharp features by convolving the mask design with a kernel; as long as the Fourier Transform of the smoothing kernel is equal to unity everywhere within twice the radius of the Lyot stop boundary, the light that continues through the system is not affected, and the mask will have the same performance. The Fourier Transform of the smoothing kernel needs to be equal to unity for twice the boundary of the Lyot stop because of the autocorrelation of the Lyot stop. The nominal smoothing kernel is therefore the half-scale Airy pattern of the Lyot transmission function.

We convolved this *nominal kernel* with the three mask designs shown in Figure 5.3c, and indeed the system performance remained constant within computational errors. However, the nominally smooth mask profiles have *ringing* at the zone boundaries that could be equal or more difficult to manufacture compared with the original designs. The *ringing* is caused by the diffraction rings that surround the central core of the smoothing kernel Airy pattern. This problem was predicted in [67], and they proposed surrounding the unity pupil pattern with *gaussian wings* to reduce the ringing. This method was somewhat effective at reducing the ringing, but also reduced the smoothing effect of the kernel. The kernel with *gaussian wings* needs to be applied for several iterations to achieve the same level of smoothing as one iteration with the nominal smoothing kernels, and ultimately leads to the same result of ringing in the pattern.

To alleviate the *ringing* effect, we convolved the original mask designs with an alternative smoothing kernel that consists of only a small circularly symmetric core. This *practical kernel* does not theoretically achieve the original design performance, but it does effectively smooth the mask without creating *ringing* in the structure. After applying the *practical kernel* to the mask designs, we re-optimize the heights

of the entire structure to achieve close to the nominal performance.

The mask designs smoothed with the nominal kernel (red line) and the practical kernel (green line) are shown in Figure 5.3c along with the average contrast computed in the same optimization region as the respective original designs (blue line). Although the nominal smoothing kernel can retain close to the contrast of the original designs, the excessive ringing in the structure does not help the manufacturing process. The practical kernel, however, effectively smooths the sharp features of the mask design and can recover the original contrast performance within a factor of two. In the case of Design 3, the performance of the practically smoothed mask is actually better than the original design performance. This particular result is caused by computational errors, and we do not typically expect any smoothed mask designs to achieve better simulated performance than the original design.

Smoothing mask profiles with a practical kernel may be a good approach for easing manufacturing constraints, and may allow the use of alternative manufacturing processes that are less accurate for constructing sharp features. So far, we have only attempted to fabricate masks using electron beam lithography, but a smoothing design process could allow the possibility for other approaches.

Another approach to achieving a smooth mask profile could be to design the mask so that the phase shift is a continuous function over the radius of the mask. In our optimization algorithm, we need to discretize the smooth function in order to minimize Equation 4.2. For the masks designed in section 4.2, we settled on a discretization of 10 ring-shaped zones because that was a sufficient number of degrees of freedom to achieve our target contrast. One might suspect that if we increased the number of zones to 1000 or more, the solution would naturally form a smoothly varying approximation of a continuous function. Our mask design attempts with a larger number of rings (up to 40 rings) have not naturally produced a smoothly varying height profile.

We could impose a continuity constraint into the optimization algorithm, where the heights of adjacent zones could only be offset up to a maximum height difference. For a large number of zones and a tightly restricted maximum height difference, the continuity constraint could produce a smoothly varying height profile. However, our current method of local optimization with a random starting point is not efficient enough to reach a solution for 1000 or more degrees of freedom. With a more advanced optimization algorithm, we could explore the possibility of designing the mask to achieve a smoothly varying phase profile.

5.2 Manufacturing

Mask designs are typically on the order of 100 microns in diameter with a depth of up to 2 microns. Despite the small size and need for precise manufacturing, there are some fabrication techniques that can achieve high quality masks. We fabricate the masks at the Stanford Nanofabrication Facility using electron beam lithography with Polymethyl methacrylate (PMMA) resist. The resist is typically 420nm thick on a Silicon or Quartz wafer. Silicon is a good wafer material for transmissive masks in the near-infrared, or the wafer can be coated to act as a reflective mask. Quartz is a good wafer material for transmissive masks in the visible or near-infrared. The beam energy is 50keV with a 1.5nm beam radius. After resist development, the depth of the mask profile is limited to the depth of the resist, but we can increase the pattern depth by etching the pattern into the wafer. The pattern is transferred from the resist into the wafer by reactive ion etching; in this process the depth of the mask profile is multiplied by a factor of approximately 2 to 5 depending on the plasma composition.

Several different mask designs have been fabricated at SNF, and to characterize the manufacturing results, we will focus on a design that is useful for the ACE testbed. An optical image of the mask is shown in Figure 5.4, where the color variation for different

mask zones indicates that the zones are at different depths within the PMMA resist. Atomic Force Microscopy (AFM) is the best tool available for precisely measuring the depths of the mask zones. An AFM measurement of the mask is shown at the bottom of Figure 5.5, along with a one dimensional profile through the center of the mask, which can be used to determine the accuracy of the manufactured zone depths. The mask was measured just after the resist development and before reactive ion etching, so we compare the measured mask profile with the design profile that we expect to achieve in the resist. The zone edges appear smooth at the zone transitions, but this is mainly an artifact of the AFM measurement. We ignore the AFM measurements near the zone boundaries, and take an average of the profile depth within the center of each zone. The average error between the design and measured zone depths is 8 nanometers. We expect the profile depths to multiply by a factor of 1.7 in the etching process when the profile is transferred into the Silicon wafer, so we predict an average error of 14 nanometers in the Silicon wafer after etching. In future manufacturing iterations, these errors will decrease as we further refine the manufacturing process.

5.3 Results

We can use the AFM measurements to compute the expected performance of the manufactured mask. The errors in the mask zone depths translate to phase errors, as described in Equation 4.1. We include the measured zone depths in the ACE coronagraph simulation, and calculate the resulting contrast over a 10 % bandwidth centered at 580nm. Results are shown in Figure 5.6. The mask was designed to achieve an average contrast of 7.73×10^{-7} between 1.4 and 6 λ/D . The manufactured mask achieves an average contrast of 9.09×10^{-6} within the same region, which is a factor of 12 worse than the design.

There are three zones within the manufactured mask where the depth errors are

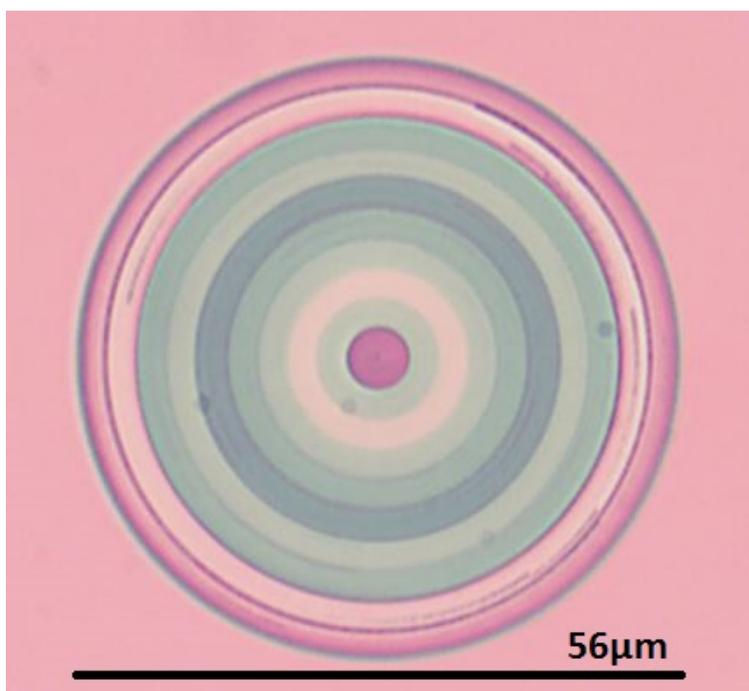


Figure 5.4: Optical microscope image of a manufactured mask.

greater than the other errors. These zones, shown in the 1D plot of Figure 5.5, are the center zone of the mask and the two zones closest to the outer edge of the mask. The errors in these zones dominate the average error. It is particularly important to correct the error at the center zone, which corresponds to the center of the stellar PSF where the starlight intensity is greatest. Therefore an error in this zone has the greatest effect on the phase shift of the starlight, and ultimately the mask performance. One efficient path to improvement is to refine the manufactured depth of this center zone. This is a typical process for mask manufacturing, where we refine the errors in the mask manufacturing over several iterations. As we gain more experience with mask manufacturing at SNF, we expect that the error refinement process will become more efficient. For example, as the process becomes more repeatable, we will have more accurate knowledge of the relationship between the electron beam dose and the resulting etch depth in the resist.

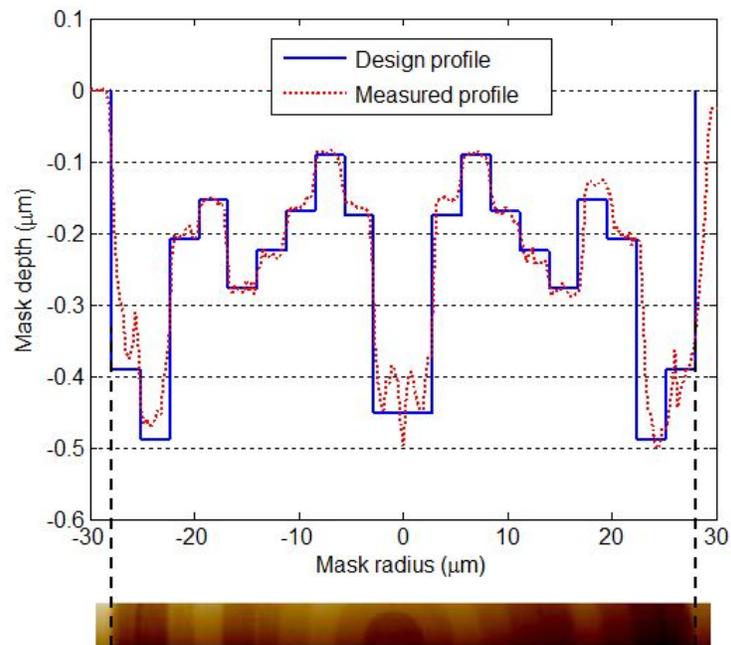


Figure 5.5: AFM measurement of a manufactured mask (bottom), and profile of the zone depths compared with the designed profile (top).

5 Phase masks for PIAACMC

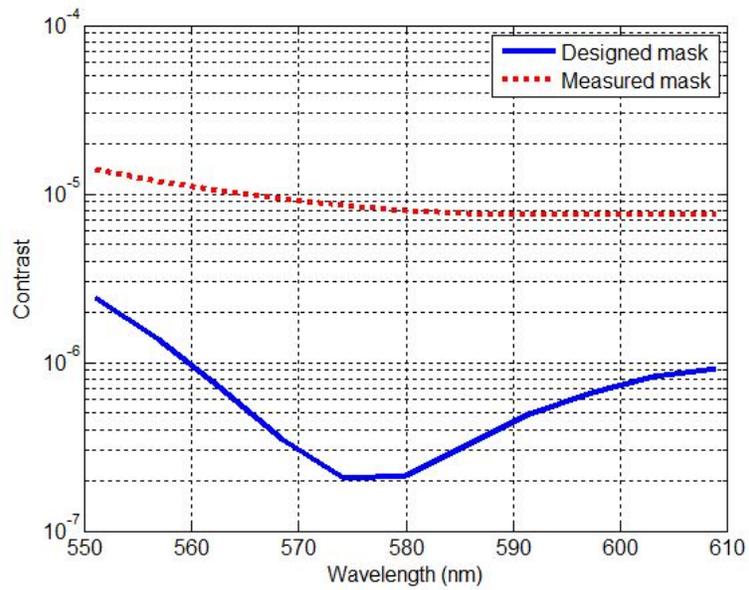


Figure 5.6: Comparison of contrast over a 10% bandwidth for an ideal mask and a fabricated mask. The performance of the fabricated mask is more than one order of magnitude worse than the ideal mask. However, performance should improve with further fabrication development and implementation of wavefront control.

5 Phase masks for PIAACMC

One important note is that these performance results are calculated as raw contrast without wavefront control. Wavefront control has been demonstrated with phase masks for PIAACMC in simulation [68], and it can provide a significant improvement to the system performance and compensate for some flaws in the mask manufacturing. The errors in the mask zone heights are comparable to typical errors of other optical components within the optical path, such as phase errors on mirrors, which can be corrected by wavefront control.

We can also expect some amount of performance increase from post-processing; some predict a factor of 30 [69], that will further enhance the contrast that is achievable with these designs. Some have achieved an improvement in contrast by a factor of 100 from PSF calibration [70]. Ultimately the important factor for planet detection is the signal-to-noise ratio, and post-processing cannot decrease the noise below the fundamental Poisson limit. However, the telescope integration time can be increased to provide a greater signal-to-noise ratio for planet detection, because the planet signal scales linearly with integration time, but the Poisson noise scales as a square root.

5.4 Correcting for mask manufacturing errors

In Section 1.2 we highlighted the importance of using adaptive optics with a high performance coronagraph system. AO is typically used to correct for small errors in the optical system, including correcting for atmospheric turbulence if the system is ground-based. Errors in the construction of the focal plane mask can be included in this category, as focal plane errors result in unexpected diffracted light. In this section we will explore the use of wavefront control with PIAACMC, specifically to correct for mask manufacturing errors. In section 4.1 we demonstrated that stopping-down the ACE PIAA optics provides a closer approximation to a typical PIAACMC

5 Phase masks for PIAACMC

apodization profile, so we will use this as a starting point for a new focal plane mask design.

The modified apodization profile is shown in Figure 5.7a. The focal plane mask is circular with a 60 micron outer diameter, which is divided into ten circularly symmetric ring-shaped zones, as shown in Figure 5.7b. The new focal plane mask, Design 4, is shown in Figure 5.7c. The mask is designed for a 10% bandwidth centered at 580nm. The depth of each zone is optimized to provide the phase shift necessary to achieve deep contrast over the entire bandwidth as described in Section 4.1. We aim for an average broadband raw contrast (PSF surface brightness) on the order of 10^{-7} , which is suitable for exoplanet science with ground-based telescopes, and is also suitable for space-based telescopes that are interested in observing circumstellar disks and Jupiter-size planets. Before wavefront control, Design 4 has 4.9×10^{-7} average contrast between 2 and 6 λ/D . The mask could be scaled laterally as a function of radius, and axially as a function of depth in order to accommodate another range of wavelengths. The PIAACMC architecture is certainly capable of achieving deeper contrast and smaller IWA than these designs [55, 66], but the goal of this section is to demonstrate wavefront control, so particularly high contrast is not necessary. Contrast levels on the order of 10^{-7} are already ambitious goals for ground-based observatories, and we will see in Section 4.5 that the design presented here is scientifically valuable.

A variety of PIAACMC focal plane masks have been manufactured at the Stanford Nanofabrication Facility and the NASA Jet Propulsion Laboratory, Microdevices Laboratory. The manufacturing process for such small and precise devices is not a trivial task, yet the labs have achieved good results from manufacturing [71, 66, 64]. However, there are still residual manufacturing errors, the most common of which is a

5 Phase masks for PIAACMC

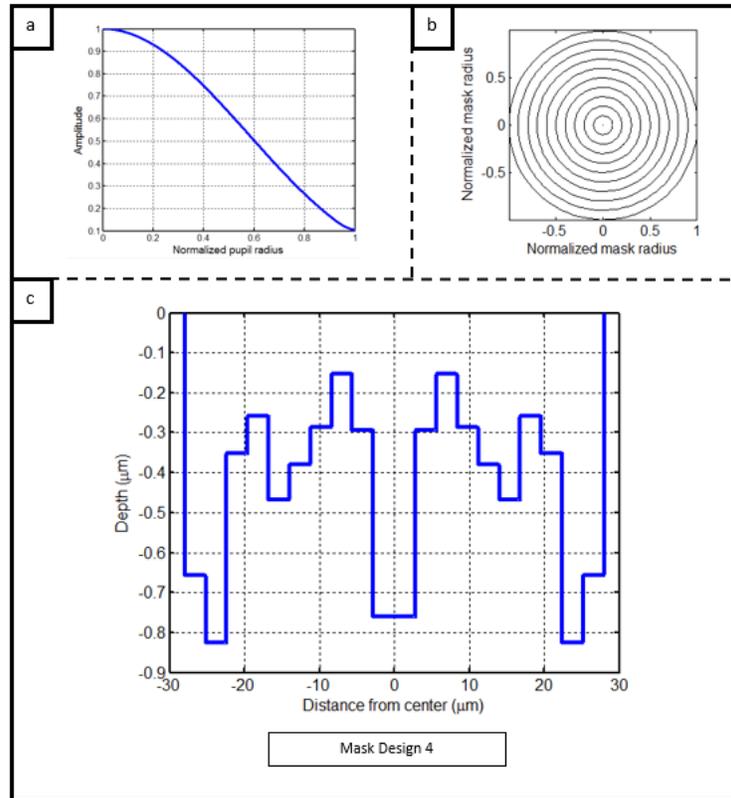


Figure 5.7: Radially symmetric apodization profile of the modified ACE PIAA optics (a) and structure of the focal plane mask design (b). A cross section profile of the Design 4 focal plane mask (c).

5 Phase masks for PIAACMC

difference between the designed and manufactured depth for a particular zone in the mask. Results from one particular manufacturing session were shown in Section 4.2, along with the resulting degradation of coronagraph performance shown in Section 4.3. In this section we will more broadly characterize how these typical mask manufacturing errors affect the performance of the coronagraph system, and how well we can compensate for these errors with the use of wavefront control. Starting with the mask Designs 4, we simulated adding random amounts of error to the depths of each zone in the mask, and calculated the resulting coronagraph performance. Results are shown with the red dots in Figure 5.8.

We simulated 150 trials of the Design 4 mask with independent randomly selected errors added to the depth of each zone. The errors were selected from a uniformly distributed range between $-N$ and $+N$ for different values of N between 0 and 200 nm. Errors assigned to different zones within a mask were uncorrelated. On the x-axis of Figure 5.8, we show the average of the absolute value of all depth errors in the mask. The y-axis in Figure 5.8 shows the resulting average contrast within a specified dark region. The dark region spans radially from 2 to 6 λ/D , and azimuthally from $+90^\circ$ to -90° . The region does not include the area within $0.5 \lambda/D$ in the x-direction to avoid a situation where the deformable mirror (DM) is attempting to correct regions that are separated by a 180° angle. The individual contrast values are fitted with a fourth-order polynomial that reaches a stable peak as the average zone error increases. For zero error, the system performance starts with the originally designed contrast, and as the error increases it follows a steep curve of worsening contrast before it peaks at 4.3×10^{-5} . There is a high degree of variability in the system performance, even for masks with similar amounts of error. However, we can identify a few important trends in the system performance. Without any strategy for performance correction, the average mask error would need to remain less than 5nm in order to achieve within

5 Phase masks for PIAACMC

an order of magnitude of the original design contrast. Once the average error in the mask has reached 40nm, the performance reaches a plateau, which suggests that any additional error in the system has little effect on the performance.

To correct for the effect of errors in the mask zone heights, we implemented an electric field conjugation (EFC) wavefront control algorithm [28]. In this investigation the mask errors in the PIAACMC focal plane create an aberrated field in the final focal plane. For these simulations, we assume that the rest of the coronagraph system does not contribute any aberrations, and that we have perfect knowledge of the electric field in the final focal plane. The EFC algorithm uses the spectrum of the aberrated electric field in the final focal plane and computes a corresponding DM setting that corrects these aberrations at a plane which is conjugate to the PIAA exit pupil (where the DM is located). EFC is a model-based wavefront correction technique and requires a linear model of the effect of each DM actuator poke (change in actuator height) in the correction region of the final image plane. For these simulations, we model the DM that is in use at the ACE laboratory, which is a Boston Micromachines DM with a grid of 32x32 actuators. A regularized least-squares solution computes the required DM actuator heights (DM settings) to conjugate these aberrations. DM settings are computed for the central wavelength (580 nm) and maintained as the wavelength band is extended to broadband. The WFC algorithm does not always reach an ideal solution, so we eliminated WFC results where the algorithm had diverged to a bad solution, which usually occurred because the regularization parameter search was not exhaustive enough. Many of these cases were clearly identified because the best contrast was at the outer edge of the design bandwidth; this should not occur for an ideal WFC solution because the system is corrected at the center wavelength. We also eliminated a few WFC solutions that resulted in a contrast worse than $1 \times 10^{-3.5}$, because these solutions were clearly outliers and not representative of the WFC

5 Phase masks for PIAACMC

algorithm performance. In total we eliminated approximately 25% of the WFC data.

The remaining WFC simulations, shown as blue dots in Figure 5.8, have a similar behavior compared with the uncorrected masks. However, the contrast follows a shallower curve as a function of mask zone error, and reaches a peak at 1.2×10^{-5} . For some masks with an average zone error of less than 15nm, WFC can achieve contrast that is better than the original design. With any greater amount of error, wavefront control cannot achieve the original design contrast. Therefore the average mask depth error should be kept to less than 15nm in order to achieve the full science potential of the coronagraph design. For errors greater than 50nm, there is little difference in the achievable contrast, but WFC can still improve the contrast by an order of magnitude in comparison to an uncorrected system. These results were obtained with a particular implementation of EFC that has been established at the ACE testbed for use with PIAA systems. It is possible that another implementation of EFC or a different WFC strategy altogether could produce better results for correcting mask performance.

5.4.1 Surface roughness

Surface roughness is another potential error in the focal plane mask manufacturing, as perfectly smooth surfaces are difficult if not impossible to achieve. We performed a simulation with surface roughness added to the Design 4 focal plane mask to determine if this type of error could have a significant effect on the coronagraph performance. Simulated mask pixels are 55nm squares, and we randomly added error to individual pixels with a uniform distribution between +/- 5nm. This is the maximum amount of surface roughness that we would expect from manufacturing. With surface roughness added to one hundred trial masks, all of the systems maintained performance at better than 4.94×10^{-7} contrast, which is almost no difference from the original

5 Phase masks for PIAACMC

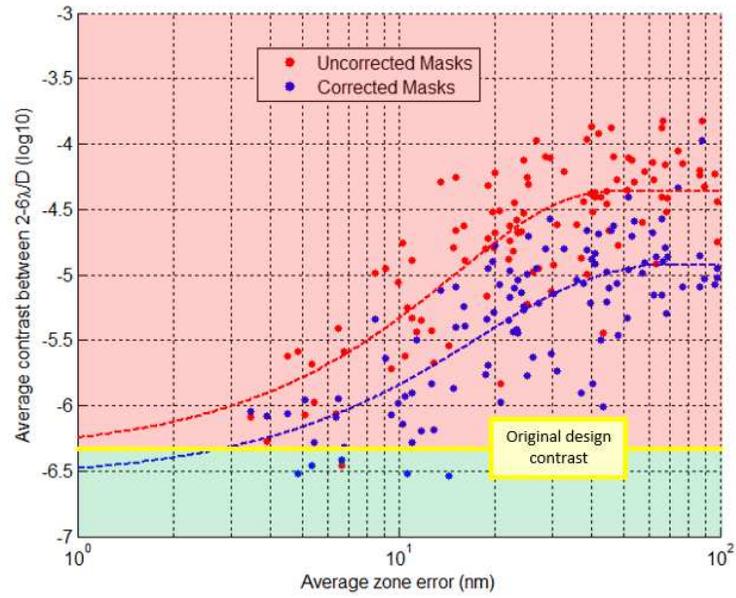


Figure 5.8: Contrast of an unobstructed pupil PIAACMC design with random manufacturing errors added to the focal plane mask (red dots). The contrast can be improved, up to a certain level, with the use of wavefront control (blue dots).

performance. The very minimal changes that we observed in contrast were likely caused by computational errors. This result was expected because surface roughness represents high frequency noise on the focal plane mask. These high frequencies are simply blocked from the system by the Lyot stop. We do not expect surface roughness on the focal plane mask to have any effect on the performance of any PIAACMC system.

5.5 Science potential with reflected light

In section 4.4 we studied how manufacturing errors and wavefront control contribute to PIAACMC system performance. In this section, we will investigate how the system performance relates to scientific capability of the instrument. Dr. Guyon created a model of potential extrasolar planets around nearby stars [72], and determined the contrast and separation angle that would be necessary to observe the planets with reflected starlight. We can use this model as a basis to determine, for a given PIAACMC instrument, how many stars can be observed to potentially discover extrasolar planets with reflected light. This is the same type of analysis that we looked at in Section 2.2, and now we can apply our mask manufacturing and wavefront control results to see what type of exoplanet science we can achieve given a realistic coronagraph performance. We will use the design 4 focal plane mask as a system baseline.

Results are shown in Figure 5.9 for potential Jupiter-size planets in the habitable zone of their host stars. The size of the dots represents the distance from Earth to the host star, where larger dots represent planets orbiting around the stars closest to Earth. If our design 4 PIAACMC system were applied to a telescope, the number of planets that can potentially be observed is greatly dependent on the size of the telescope aperture. For a ground-based telescope, we assume a primary mirror aperture of 10m diameter, although there are several next-generation ground-based telescopes

5 Phase masks for PIAACMC

that plan to have a greater aperture size. For a space-based telescope, we assume a 2.4m diameter aperture, although the James Webb Space Telescope and other future space telescopes could have larger diameter primary mirrors. We also assume a factor of 30 contrast improvement due to post-processing, as was the convention in [69], and we assume that a signal to noise ratio of 5 is necessary for a detection. Finally, we assume that the combination of wavefront control with the manufactured focal plane mask can achieve the original design 4 performance. In section 4.4 we determined that many focal plane masks can achieve the original design performance as long as they have less than 10nm average zone depth error. Recent manufacturing results from the Stanford Nanofabrication Facility have produced PIAACMC focal plane masks with an average zone depth error of 8nm. We correct the system with wavefront control over a region from 2 to 10 λ/D . The resulting system performance curves are shown in Figure 5.9.

For a ground-based telescope with a 10m primary mirror, the Design 4 PIAACMC system has the sensitivity to observe Jupiter-size planets in the habitable zones of 223 stars. The number of planets observed could be increased with a larger telescope aperture, but the contrast cannot be improved much due to the effects of the atmosphere. For a space-based telescope with a 2.4m aperture, the system has the sensitivity to observe Jupiter-size planets in the habitable zones of 3 stars. The number of planets observed could be greatly increased with a PIAACMC design that has a smaller inner working angle, or a larger telescope aperture. The number of planets observed from space could also be increased if the system were designed for deeper contrast, such as the AFTA PIAACMC system [66]. The Design 4 PIAACMC system is best suited for observing Jupiter-size planets with a ground-based telescope, but it could also provide some capabilities for a space-based telescope. The results presented here indicate that wavefront control is capable of correcting for the typical errors in

5 Phase masks for PIAACMC

most manufactured PIAACMC focal plane masks, and therefore manufacturing of the mask is not a limiting factor for coronagraph performance or scientific yield.

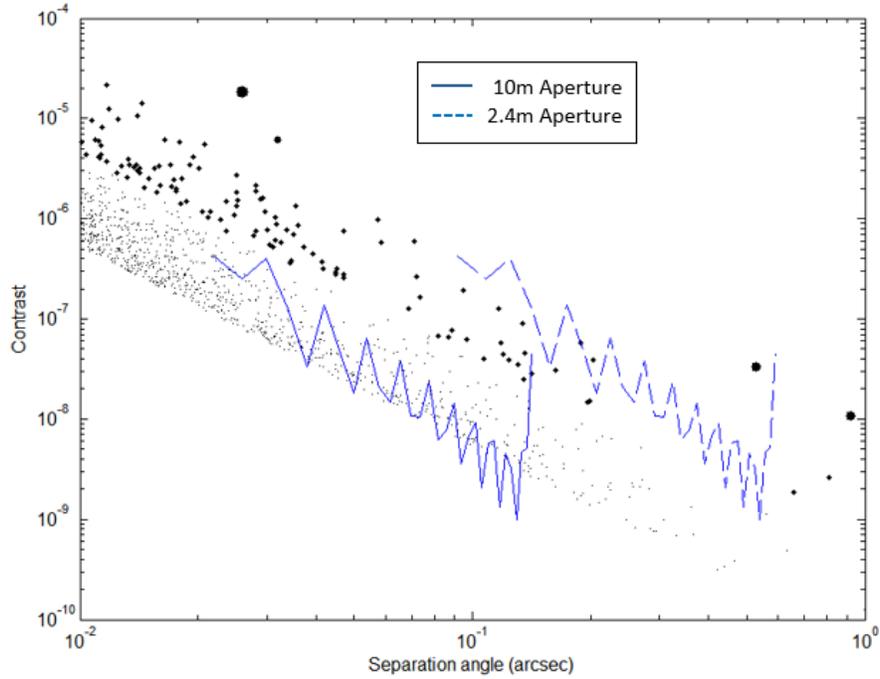


Figure 5.9: Contrast and separation angle required to directly observe a hypothetical Jupiter-size exoplanet in reflected light at quadrature in the habitable zone around nearby stars. The solid blue line shows the performance of the Design 4 PIAACMC system with a 10m diameter telescope aperture, and the dashed blue line shows the performance with a 2.4m telescope.

6 Perspectives and future directions

6.1 Advancing the performance of coronagraphy

The achromatic size-scaling focal plane mask that we discussed in Chapter 3 is a simple solution to improve the IWA for coronagraph architectures that use an opaque disk occulter, including the PIAA and APLC. The mask design enables a smaller IWA for short wavelengths in the observation bandwidth, which enables the potential for detection of additional planets around a given star. The mask is a direct replacement for the opaque disk occulter, and only requires a Lyot stop (or exit pupil) for its operation, which is a component that is usually already included in the coronagraph. With our design and successful laboratory demonstration of the size-scaling mask, the concept is ready to be applied to coronagraph systems to provide a performance boost.

Prior to the work described in Section 4, the status of PIAACMC architecture design was mainly an analysis of the concept in monochromatic light. However, direct imaging is more efficient in broadband. We outlined a process to design the partially phase-shifting focal plane mask for use in broadband, and demonstrated the manufacturing of the mask. A similar effort was performed for the WFIRST technology development [66], and our work is complementary because it is a more broad description of the mask design and manufacturing process that can be applied to other PIAACMC systems.

Prior to our work, all focal plane mask manufacturing was completed by the JPL Microdevices Laboratory [64]. We started a relationship with an additional source for mask manufacturing, the Stanford Nanofabrication Facility. This effort provided the practical advantages of having two sources for manufacturing, including the ability to complete more projects in a shorter amount of time, for a lower cost, while reducing the project load on JPL. Our relationship with SNF also allows the opportunity for trying new fabrication techniques, or working with new materials and designs. For

6 *Perspectives and future directions*

example, we have started an effort to fabricate PIAACMC focal plane masks on a Quartz substrate, which is a potential material for making transmissive masks in the visible spectrum.

As we discussed in Section 1.2, wavefront control is an extremely important technology for the success of coronagraph systems. It is important to demonstrate the use of wavefront control with any coronagraph architecture, and PIAACMC is no exception. In Section 4.4 we demonstrated the use of wavefront control with PIAACMC, specifically for the purpose of correcting for mask manufacturing errors. This analysis not only serves as a demonstration of WFC with PIAACMC, but also provides feedback to the mask manufacturing process to set tolerances on the accuracy needed to achieve high performance. Our analysis was specific to a PIAACMC design for the ACE laboratory PIAA optics, but the results can provide insights for other PIAA-based coronagraph systems.

The advanced state of affairs for coronagraphy now includes a solution to the size chromaticity challenge for an opaque disk occulter. It also includes a generalized solution to the size chromaticity and complex amplitude chromaticity challenges for the partially phase-shifting focal plane mask of PIAACMC. The advancement of these design challenges, along with demonstrations of manufacturing and wavefront control, provides the coronagraph community with a broadband solution for PIAACMC, which is one of the most promising architectures for current and future direct imaging science missions.

6.2 Future directions and considerations for space performance

PIAACMC has achieved high performance in simulation, and an important next step for advancing the Technology Readiness Level is to achieve high performance in the laboratory. Laboratory demonstration has been successful for PIAA [14, 42], and simulations have shown that fabricated phase-shifting focal plane masks should be adequate for high performance with PIAACMC [66]. The next important step is to demonstrate broadband high performance of a complete PIAACMC system in a laboratory environment.

For translating the laboratory coronagraph performance to space-based mission performance, stability of the telescope will be a critical factor. Instability leads to unwanted starlight leaking around the edges of the focal plane mask and degrading the contrast. The telescope designer has a few options for mitigating this effect: tightening the requirements for telescope stability, correcting for instability with WFC, or designing the coronagraph to tolerate more instability. Tightening the requirements for telescope stability is not always possible, but there are some advances to be made with WFC and coronagraph design.

Broadband WFC has been demonstrated in the laboratory with PIAA [41], and in simulation with PIAACMC (see Section 4.4). The next logical step is to demonstrate WFC in broadband with PIAACMC in the laboratory, as this is a crucial technology to achieve high performance. There is an effort underway to demonstrate 10% broadband WFC with PIAACMC as part of the WFIRST development. There is also an effort underway to demonstrate WFC with segmented apertures [73].

There is also an opportunity to adjust the coronagraph design to tolerate more instability. The flexibility of PIAACMC, and specifically the degrees of freedom of

6 Perspectives and future directions

the phase-shifting focal plane mask, provides a set of tools that the coronagraph designer can use to create a coronagraph which is more tolerant to instability. The same set of tools can be used to design the coronagraph to be more tolerant to stellar angular size, which becomes more important for large aperture telescopes. An important step in the development of PIAACMC is to learn how to take advantage of the design flexibility to alleviate stability requirements of the telescope.

7 References

References

- [1] A. Wolszczan, D. Frail, *Nature*, 355, p. 145-147 (1992)
- [2] G. Basri, W. Borucki, D. Koch, *New Astronomy Reviews*, 49, 7-9, p. 478-485 (2005)
- [3] J. Coughlin, F. Mullally, S. Thompson, et al., submitted to *ApJ* (2015)
- [4] N. Batalha, *PNAS*, Vol. 111-35, 12647-12654 (2014)
- [5] G. Marcy, L. Weiss, E. Petigura, et al., *PNAS*, Vol. 111-35, 12655-12660 (2014)
- [6] S. Seager, D. Sasselov, *ApJ*, 537, 916-921 (2000)
- [7] S. Seager, *PNAS*, Vol. 111-35, 12634-12640 (2014)
- [8] J. Trauger, K. Stapelfeldt, W. Traub, et al., *Proc. SPIE*, Vol. 7731 (2010)
- [9] O. Guyon, E. Pluzhnik, M. Kuchner, B. Collins, S. Ridgway, *ApJ*, 167-1 (2006)
- [10] W. Cash, C. Copi, S. Heap, et al., *External Occulters for the Direct Study of Exoplanets*, ExoPlanet Task Force (2007)
- [11] S. Seager, W. Cash, S. Domogal-Goldman, et al., *Exo-S: Starshade Probe-Class Exoplanet Direct Imaging Mission Concept* (2015)
- [12] R. Bracewell, *Nature*, 274, 780 (1978)
- [13] P. Lawson, O. Lay, and C. Beichman, *JPL Publication 07-1* (2007)
- [14] P. Lawson and J. Dooley, *JPL Publication 05-5* (2005)
- [15] C. Cockell, A. Leger, M. Fridlund, et al., *Astrobiology*, Vol. 9-1, pp. 1-22 (2009)
- [16] D. Defre, P. Hinz, A. Skemer, et al., *ApJ* 799, 42 (2015)
- [17] B. Lyot, *MNRAS*, Vol. 99, pp. 538 (1939)

References

- [18] F. Rigaut, *PASP*, Vol. 127-958, pp. 1197-1203 (2015)
- [19] T. Fusco, J. Sauvage, C. Petit, et al., Extreme adaptive optics for the detection of extrasolar planets, *SPIE Newsroom* (2015)
- [20] B. Macintosh, J. Graham, D. Palmer, et al., *Proc. SPIE*, 7015 (2008)
- [21] F. Martinache, O. Guyon, N. Jovanovic, et al., *PASP*, 126-565 (2014)
- [22] M. Hart, N. Milton, C. Baranec, et al, *Nature* 466, 727-729 (2010)
- [23] M. Nicolle, T. Fusco, G. Rousset, and V. Michau, *Opt Lett.* 29, 23, pp. 2743-2745 (2004)
- [24] F. Malbet, J. Yu, M. Shao, *PASP*, 107, 386 (1995)
- [25] O. Guyon, *ApJ*, 629, 1, pp. 592-614 (2005)
- [26] J. Trauger, C. Burrows, B. Gordon, et al., *Proc. SPIE*, 5487 (2004)
- [27] P. Borde, W. Traub, *Proc. IAU*, 1, C200, pp. 535-540 (2005)
- [28] A. Give'on, B. Kern, S. Shaklan, et al., *Proc. SPIE*, 6691 (2007)
- [29] R. Soummer, C. Aime, P. Falloon, *A&A*, 397, 1161 (2003)
- [30] J. Kasdin, R. Vanderbei, D. Spergel, et al., *ApJ*, 582, pp. 1147-1161 (2003)
- [31] O. Guyon, 2003, *A&A*, 404, 397
- [32] G. Singh, F. Martinache, P. Baudoz, et al., *PASP*, 126, 940, pp. 586-594 (2014)
- [33] M. Kuchner and W. Traub, *ApJ*, 570, 2, pp. 900-908 (2002)
- [34] F. Roddier, C. Roddier, *PASP*, 109, 815 (1997)
- [35] D. Rouan, P. Riaud, A. Boccaletti, et al., *PASP*, 112, 1479-1486 (2000)

References

- [36] R. Soummer, K. Dohlen, C. Aime, *A&A*, 403, 1, 369-381 (2003)
- [37] O. Guyon, F. Martinache, R. Belikov, et al., *ApJ*, 190-2, pp. 220-232 (2010)
- [38] D. Palacios, *Proc. SPIE*, 5905 (2005)
- [39] D. Mawet, P. Riaud, O. Absil, et al., *ApJ*, 633, 1191-1200 (2005)
- [40] P. Lawson, R. Belikov, W. Cash, *Proc. SPIE*, 8864-50 (2013)
- [41] D. Sirbu, S. Thomas, R. Belikov, et al., *Proc SPIE* 9605-18, (2015)
- [42] R. Belikov, E. Bendek, T. Greene, et al., *Proc. SPIE* 8864-31 (2013)
- [43] G. Chauvin, A. Lagrange, C. Dumas, et al., *A&A* 425, L29-L32 (2004)
- [44] <http://exoplanets.org/table/>
- [45] G. Toomey, C. Ftaclas, *Proc. SPIE*, 4841, pp. 889-900 (2003)
- [46] K. Hodapp, M. Tamura, R. Suzuki, et al., *Proc SPIE*, 6269 (2006)
- [47] N. Jovanovic, F. Martinache, O. Guyon, et al., *PASP*, in preparation
- [48] J. Beuzit, D. Mouillet, K. Dohlen, et al., *Proc. SPIE*, 9148 (2014)
- [49] B. Macintosh, J. Graham, D. Palmer, *Proc. SPIE*, 672 (2006)
- [50] TMT Website, <http://www.tmt.org/observatory> (2016)
- [51] GMT Website, <http://www.gmto.org/overview/> (2016)
- [52] E-ELT Telescope Design, <https://www.eso.org/sci/facilities/eelt/telescope/index.html>
(2016)
- [53] R. Belikov, E. Bendek, S. Thomas, et al., *Proc. SPIE*, 9605 (2015)

References

- [54] E. Bendek, R. Belikov, J. Lozi, et al., Proc. SPIE, 9605 (2015)
- [55] O. Guyon, P. Hinz, E. Cady, et al., ApJ 780-171 (2013)
- [56] <http://exep.jpl.nasa.gov/stdt/exoc/>
- [57] K. Newman, R. Belikov, O. Guyon, et al., Proc. SPIE, 8864 (2013)
- [58] K. Newman, R. Belikov, E. Pluzhnik, et al., Proc. SPIE, 9151, 216 (2014)
- [59] K. Newman, O. Guyon, K. Balasubramanian, et al., PASP (2015)
- [60] R. Belikov, O. Solgaard, Opt. Lett., 28, 6 (2003)
- [61] N. Jovanovic, O. Guyon, F. Martinache, et al., Proc. SPIE, 9147 (2014)
- [62] O. Guyon, T. Matsuo, R. Angel, ApJ, 693, pp. 77-84 (2009)
- [63] F. Vogt, F. Martinache, O. Guyon, et al., PASP, 123, 1434 (2011)
- [64] K. Balasubramanian, D. Wilson, V. White, et al., Proc. SPIE, 8864 (2013)
- [65] P. Lawson, R. Belikov, W. Cash, et al., Proc. SPIE, 8864-50 (2013)
- [66] B. Kern, O. Guyon, R. Belikov, et al., submitted to JATIS (2015)
- [67] O. Guyon, “Focal Plane Masks for Achromatic Coronagraphy,” (in preparation)
- [68] E. Sidick, B. Kern, R. Belikov, et al., Proc. SPIE, 9143 (2014)
- [69] D. Spergel, N. Gehrels, et al., “Wide-Field InfraRed Survey Telescope Astrophysics Focused Telescope Assets WFIRST-AFTA,” 2015 Report by Science Definition Team (SDT) and WFIRST Study Office (2015)
- [70] A. Skemer, P. Hinz, S. Esposito, et al., ApJ, 753-1 (2012)

References

- [71] K. Newman, J. Conway, R. Belikov, et al., “Focal plane phase masks for PIAA: design and manufacturing,” *PASP* (to be published)
- [72] <http://www.naoj.org/staff/guyon/04research.web/14hzplanetsELTs.web/catalog.web/content.htm>
- [73] J. Codona, O. Guyon, NASA Early Stage Technology Workshop, ‘Wavefront Control for High Performance Coronagraphy on Segmented and Centrally Obscured Telescopes’ (2015)