SIMULATION, DESIGN, AND TEST OF SQUARE, APODIZED PHOTON SIEVES FOR HIGH-CONTRAST, EXOPLANET IMAGING

THESIS

Thomas W. N. Dickinson, Second Lieutenant, USAF

AFIT-ENP-MS-16-M-065

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

DISTRIBUTION STATEMENT A
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

The views expressed in this document are those of the author and do not reflect the official policy or position of the United States Air Force, the United States Department of Defense or the United States Government. This material is declared a work of the U.S. Government and is not subject to copyright protection in the United States.

AFIT-ENP-MS-16-M-065

SIMULATION, DESIGN, AND TEST OF SQUARE, APODIZED PHOTON SIEVES FOR HIGH-CONTRAST, EXOPLANET IMAGING

THESIS

Presented to the Faculty

Department of Engineering Physics

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

in Partial Fulfillment of the Requirements for the

Degree of Master of Science in Applied Physics

Thomas W. N. Dickinson, B.S.

Second Lieutenant, USAF

March 24, 2016

DISTRIBUTION STATEMENT A
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

AFIT-ENP-MS-16-M-065

SIMULATION, DESIGN, AND TEST OF SQUARE, APODIZED PHOTON SIEVES FOR HIGH-CONTRAST, EXOPLANET IMAGING

THESIS

Thomas W. N. Dickinson, B.S.
Second Lieutenant, USAF

Committee Membership:

Lt. Col. Anthony L. Franz, PhD
Chair

Dr. Geoff P. Andersen
Member

Dr. Michael R. Hawks
Member

Dr. Michael A. Marciniak
Member

AFIT-ENP-MS-16-M-065

Abstract

A photon sieve is a lightweight, diffractive optic which is well-suited to be a deployable primary for a space telescope. Point spread functions (PSFs) can be altered by shaping and apodizing an aperture, and a PSF that drops off rapidly from the peak is desirable for high-contrast imaging. Due to diffraction, a square aperture concentrates much of the energy in the PSF along the vertical and horizontal axes in the image plane, but off-axis the PSF drops off more rapidly than the PSF of a circular aperture. Application of an appropriate apodization can further augment this effect. Because of this, square apodized photon sieves were simulated, designed, and tested for high-contrast performance and use in an exoplanet imaging telescope, with an algorithm being devised prior to fabrication for applying any apodization to a photon sieve.

Ten photon sieves were designed and experimentally tested. A resolution target was imaged with each sieve, with five of seven square apodized sieves resolving bar separations of 6.96 μm or smaller (compared to 6.20 μm for the unapodized sieves, and 7.41 μm for the Rayleigh criterion resolution). Though certain apodizations can significantly degrade resolution, images produced with these five sieves exceeded the Rayleigh resolution and nearly matched the imaging resolution of the unapodized square sieve. The sieves were characterized with a PSF analysis, for which a high dynamic range, scale-and-slice method was developed to drastically extend the detector's dynamic range and determine the PSF's shape down to below 10^{-4} of the peak irradiance. This is a general method that could be applied to any work where detailed experimental PSF characterization is important. The numerical Fresnel diffraction simulation agreed closely with this analysis, resulting in a maximum residual of 0.12

and an average residual of 0.008, relative to a peak of one. A novel semi-empirical analysis method was developed that uses the measured PSFs to assess high-contrast imaging performance more thoroughly than the best case scenarios typically presented in the literature. The square apodized sieves were shown to outperform an achromatic doublet lens and unapodized sieves for high-contrast imaging in the semi-empirical analysis and a two-beam illumination test. Cos^2 and pyramid ($1/3^{\mathrm{rd}}$) apodized square sieves achieved signal-to-noise ratios in the two-beam test more than three times greater than was achieved by an unapodized circular sieve. A laboratory-simulated exoplanet was shown to be possibly detectable using apodized square sieves with a peak value ratio below $10^{-3.69}$ and a peak-to-peak separation less than ten times the Rayleigh resolution.

The level of performance achieved (both simulated and experimentally observed) was not sufficient for directly imaging earth-like exoplanets, though it will be useful for other high-contrast applications. The ability to accurately and efficiently apply apodizations and conduct simulations for photon sieves, measure PSF shape across an extreme dynamic range, and conduct semi-empirical high-contrast imaging performance analyses will drive new PSF design and be useful for future high-contrast imaging work.

Dedicated to my wife, my parents, and my sister.

Acknowledgements

First I would like to thank my research advisor and committee chair, Lt. Col. Anthony Franz, for teaching me optics and providing excellent guidance and oversight for the full duration of the project. I would like to thank Dr. Michael Marciniak for his assistance in developing a working simulation and teaching me subjects necessary for this thesis. I would also like to thank Dr. Michael Hawks for providing advice on both experimental and theoretical work, and conducting a thorough Zemax analysis of the control lens. I would like to thank Greg Smith and Mike Ranft for their constant assistance with the laboratory equipment and setup.

Additionally, I would like to thank Dan Cook for the LaTeX formatting help and Steven Owens for helping brainstorm MATLAB implementation and optimization. Thank you to my wife for suggesting the best method for implementing the pyramid apodization and my sister for proofreading the document with painstaking attention to detail. Thank you to those who provided essential feedback on the document: Lt. Col. Franz, Dr. Marciniak, Dr. Hawks, Dr. Geoff Andersen, Dr. Olha Asmolova, and Dr. Geoff McHarg. Finally, thank you to Dr. Andersen and Dr. Asmolova for providing my first experience in an optics lab, suggesting this project, answering numerous questions and emails, hosting me for three days at the Laser and Optics Research Center to work on the project, and doing the fringe analysis of the interferograms. Thank you to all teachers, mentors, family, and friends who have inspired me throughout my life and supported me throughout this endeavor.

Thomas W. N. Dickinson

Table of Contents

Page

Abstract iv

Acknowledgements vii

List of Figures x

List of Tables xiv

List of Abbreviations xv

I. Introduction and Background 1

1.1 Problem Statement and Goal 1
1.2 Motivation 4
1.3 Exoplanet Detection 4
1.4 Fresnel Zone Plate and Photon Sieve Theory 7
1.5 Shaped, Apodized Optics 10

II. Simulation 13

2.1 Fresnel Diffraction Theory 13
2.2 Validity of the Numerical Fresnel Approximation 16
2.3 MATLAB® Implementation 19
2.4 Preliminary Simulation Results 23
2.4.1 Circular Sieve 24
2.4.2 Square Sieve 25
2.4.3 Pyramid Apodized Square Sieve 26
2.4.4 Sonine Apodized Square Sieve 29
2.4.5 Cosine Apodized Square Sieve 33
2.4.6 Gaussian Apodized Square Sieve 35
2.4.7 Apodized Square Sieve with Varying d/w 37
2.4.8 Varying Incident Wavelength 39
2.5 Noteworthy Results of the Simulation 40

III. Experimental Results 44

3.1 Equipment 44
3.1.1 Photon Sieve Design and Fabrication 45
3.1.2 Verdi Laser 53
Knife-Edge Measurement 54
Spatial Filtering 55
3.1.3 Detector Characterization 57
Detector Linearity Characterization 58

Page

Detector Blooming Characterization 60
3.1.4 HDR Imaging: The Scale-and-Splice Method 62
Steps of the Scale-and-Splice Method 62
Testing the Scale-and-Splice Method 65
3.1.5 Achromatic Doublet Lenses 68
3.2 Photon Sieve Characterization 71
3.2.1 Air Force Resolution Test Target Imaging 71
Resolution Target Imaging: No Diffuser 80
White Light Resolution Target Images 80
3.2.2 PSF Analysis and Simulation Comparison 82
Data Acquisition 83
Comparison to Simulation 90
Focusing Analysis 98
3.2.3 PSF Central Lobe Width 100
3.2.4 Interferometric Analysis 104
3.3 Laboratory Simulated Exoplanet Imaging 106
3.3.1 Semi-Empirical Method 107
3.3.2 Two-Beam Illumination Method 111
Two-Beam Illumination Test: 10^{-2} 112
Two-Beam Illumination Test: $10^{-3.5}$ 121

IV. Conclusion 127

Appendix A. Square Photon Sieve: Circularly Symmetric Apodization (or Controlled Unapodized) 132

Appendix B. Square Photon Sieve: Non-Circularly Symmetric Apodization 140

Appendix C. Photon Sieve Aperture Builder and Diffraction 145

Appendix D. PSF Image Calibration 153

Appendix E. HDR Scale-and-Splice, Peak Finding, and Semi-Empirical SNR 156

Bibliography 162

List of Figures

Figure		Page
1	Fresnel Zone Plate Geometry	8
2	Theoretical Airy Pattern Cross Section	11
3	Theoretical PSF Cross Section Comparison	12
4	Fresnel Diffraction Propagation Geometry	14
5	Numerical vs. Analytical Fresnel Integral, Low N_F	18
6	Numerical vs. Analytical Fresnel Integral, High N_F	19
7	Circular Photon Sieve	21
8	Apodized Circular Photon Sieve	22
9	Simulation Colorbar	24
10	Apodization Colorbar	24
11	Circular Sieve Exoplanet Simulation	25
12	Square Sieve Exoplanet Simulation	26
13	Square Pyramid Apodization Function	27
14	Pyramid Square Sieve Exoplanet Simulation (16% flat top)	27
15	Pyramid Square Sieve Exoplanet Simulation (no flat top)	28
16	Crossed Sonine Apodization Function	29
17	Sonine Square Sieve Exoplanet Simulation ($\nu = 6$)	30
18	Sonine Square Sieve Exoplanet Simulation ($\nu = 5$)	31
19	Sonine Square Sieve Exoplanet Simulation ($\nu = 4$)	31
20	Sonine Square Sieve Exoplanet Simulation ($\nu = 3$)	32
21	Cosine Apodization Functions	33
22	Cos^2 Square Sieve Exoplanet Simulation	34

Figure Page

23 Cos^4 Square Sieve Exoplanet Simulation 34

24 Gaussian Apodization Functions 35

25 Gaussian Square Sieve Exoplanet Simulation ($\sigma = 1.0$) 36

26 Gaussian Square Sieve Exoplanet Simulation ($\sigma = 0.5$) 37

27 Sieves with Differing d/w 38

28 Sieves with Differing d/w Exoplanet Simulation 38

29 Sonine Square Sieve PSF Simulation (varying λ) 40

30 PSF Images, Square vs. Circular Sieve 42

31 PSF Images, Square vs. Square Apodized Sieve 43

32 Fill Factor, Uncontrolled Unapodized 46

33 Fill Factor, Controlled Unapodized 47

34 Circular Apodizations: Function vs. Achieved 48

35 Apodizing Method 49

36 Apodization: Function vs. Achieved 50

37 Apodization: Function vs. Achieved, Residual 50

38 Rendering of Apodized Photon Sieves 52

39 Knife-Edge Measurement Diagram 55

40 CCD Linearity Graph 59

41 Imaged Airy Pattern with Bessel Function 60

42 CCD Blooming Displayed in Resolution Target Images 61

43 Testing the Scale-and-Splice Method 67

44 Fringes Produced with Shear Plate Collimation Tester 68

45 Spherical Aberration from the Control Lens,
Experimental Results 69

Figure Page

46 Spherical Aberration from the Control Lens, Zemax and Experimental Results 70

47 Resolution Target Imaging Setup Diagram 72

48 Resolution Target Images 76

49 Resolution Target Images Taken Without the Diffuser 80

50 White LED Source Spectrum 81

51 Resolution Target Images, Wide Bandwidth Illumination 82

52 PSF Imaging Diagram 84

53 PSFs: Log_{10} Images 86

54 PSFs: Log_{10} Surface Plots 87

55 PSFs: Log_{10} Diagonal Cross Sections 88

56 PSFs: Log_{10} Diag. Cross Sections (central portion, focused) 93

57 PSFs: Log_{10} Diag. Cross Sections (equal $f/\#$, focused) 94

58 PSFs: Log_{10} Diag. Cross Sections (central portion, unfocused) 95

59 PSFs: Log_{10} Diag. Cross Sections (equal $f/\#$, unfocused) 96

60 Residual, Central Portion Method (Focused and Unfocused) 97

61 Focusing Analysis 99

62 Focal Spot Size Test 102

63 Central Lobe Width as a Function of Loss Due to Apodization 103

64 Interferometer Diagram 104

65 Square Sieve Inteferogram 105

66 SNRs from the Semi-Empirical Analysis 109

Figure Page

67 Off-Axis SNRs from the Semi-Empirical Analysis 110

68 Two-Beam Illumination: Experimental Setup 112

69 Two PSFs: Surface Plots 115

70 Two PSFs: Log_{10} Images 116

71 Two PSFs: Log_{10} Surface Plots 117

72 Two PSFs: Log_{10} Diagonal Cross Sections 118

73 SNRs from the Two-Beam Illumination Test 121

74 Two PSFs: Log_{10} Images ($10^{-3.5}$ test) 123

75 Two PSFs: Log_{10} Diagonal Cross Sections ($10^{-3.5}$ test) 124

List of Tables

Table		Page
1	Photon Sieve Design Specifications	52
2	1951 USAF Resolution Target Information	73
3	Resolution Target Imaging Results	75
4	Two-Beam Illumination Test: Peak Separations and Peak Value Ratios	113
5	Two-Beam Illumination Test: Peak Separations and Peak Value Ratios ($10^{-3.5}$ test)	122

List of Abbreviations

Abbreviation		Page
FZP	Fresnel zone plate	1
PSF	point spread function	1
AU	astronomical units	2
ly	light years	2
PVR	peak value ratio	2
HST	Hubble Space Telescope	2
WFE	wavefront error	3
KST	Kepler Space Telescope	4
SNR	signal-to-noise ratio	7
resels	resolution elements	11
FFT	fast Fourier transform	15
CCD	charge-coupled device	44
ND	neutral density	44
LORC	Laser and Optics Research Center	44
EBL	electron beam lithography	45
TEC	thermoelectric cooling	57
ADU	analog-to-digital units	59
HDR	High dynamic range	62
AR	anti-reflection	68
OTF	optical transfer function	78

SIMULATION, DESIGN, AND TEST OF SQUARE, APODIZED PHOTON SIEVES FOR HIGH-CONTRAST, EXOPLANET IMAGING

I. Introduction and Background

A photon sieve is a new type of diffractive optic that can be made exceedingly lightweight and is thus very well-suited for space-based imaging. The typical photon sieve design is based on a Fresnel zone plate (FZP), with holes in place of the FZP's light zones. Dark-zone (antihole) photon sieves have also been successfully tested, but this project focused on only light-zone sieves in order to maintain a reasonable scope [5]. Replacing the zones with pinholes preserves structural integrity and means that photon sieves can be fabricated using flexible materials like Kapton, and potentially deployed in space as the primary optic of an imaging system. As an added benefit, eliminating the need for support struts across the FZP can reduce unwanted diffraction effects and improve the point spread function (PSF). FalconSAT-7 is an ongoing U.S. Air Force Academy research project to launch a 20-cm diameter photon sieve into low earth orbit in a 3U CubeSat, and deploy it as the primary optic of a solar telescope [6, 8, 9]. If successful, FalconSAT-7 could eventually pave the way for future missions with extremely large photon sieves (diameters greater than 20 m) that have been optimized for exoplanet imaging and detection.

1.1 Problem Statement and Goal

Detecting planets can be extremely difficult. In the cosmic sense, they are tiny and faint. Pluto, a dwarf planet within the solar system, went undiscovered until 1930 [36, p. 272]. It was too small and faint for over 300 years of telescope-using astronomers

to notice. Exoplanets add orders of magnitude of difficulty to the challenge. Pluto, at its most distant, lies 49.3 astronomical units (AU) from the Sun. An exoplanet that is 50 light years (ly) distant is roughly 64,000 times further from the Sun than Pluto. This is a relatively conservative distance for an exoplanet considering that fewer than 4,000 stars are within 50 ly of the sun [21].

Given the enormity of this challenge, the first exoplanet detection did not occur until 1989 [28]. Arguably the most interesting and challenging aspect of exoplanet research is detecting *earth-like* exoplanets. The two main problems in directly imaging exoplanets are the minuscule peak value ratio (PVR) between the exoplanet and its parent star and the tiny apparent angular separation between the exoplanet and its parent star. The PVR is defined as the ratio between the exoplanet's peak irradiance and the parent star's peak irradiance when imaged onto a detector. The smallest PVR capable of being detected by a telescope is referred to as the system's "inherent contrast" in much of the literature [26, p. 8]. Compared to most known exoplanets, earth-like exoplanets have smaller PVRs and smaller angular separations, making them even more difficult to detect and image. The PVR between an earth-like exoplanet and sun-like parent star is roughly 10^{-10}, meaning that a viable system must have an inherent contrast less than or equal to 10^{-10} [11,12,25]. For a habitable-zone exoplanet 50 ly from the sun, the angular separation between the exoplanet and its parent star would be (at most) 0.32 μrad or $0.07''$: $\frac{1}{25,000}$ the apparent angular diameter of the full moon and roughly equal to the limiting angular resolution of the Hubble Space Telescope (HST) [36, p. 167]. Combining the small apparent angular separation with the minuscule PVR means that even if a telescope is large enough that its diffraction limited resolution could theoretically resolve the exoplanet and its parent star as two separate objects, the exoplanet may still be effectively indistinguishable from the background diffracted light of the parent star.

The angular separation problem can be overcome with large aperture, high resolution telescopes. The PVR problem can be overcome by finding a way to shape the star's PSF, meaning to block or redistribute the diffracted light from the star such that the light from the exoplanet stands out from the background. Wave optics simulations suggest that this is possible with conventional optics such as lenses and mirrors [26, 34]. In order to shape the PSF with the consistency necessary for exoplanet imaging, the wavefront error (WFE) must be minimal [34]. The earth's atmosphere is a main contributer to WFE in astronomy, so being above the atmosphere is a major advantage of space telescopes [36, p. 167]. Space telescopes are also able to operate at wavelengths that are impractical for ground telescope imaging due to atmospheric absorption [19, p. 214].

Thanks to its ability to be fabricated on a flexible and lightweight membrane, a photon sieve is particularly well-suited to serve as the deployable primary optic of a high resolution space telescope with a diameter greater than 20 m [6, p. 1]. As an additional benefit, the surface flatness requirements for photon sieves are less than for conventional optics [6, 8]. If the PSF produced by a photon sieve could be appropriately shaped, a photon sieve could make an excellent optic for exoplanet imaging and detection with a high-contrast telescope.

The goal of this thesis was to evaluate promising photon sieve designs, which were chosen based on educated guesses through review of the relevant literature, first order theory, and Fourier optics simulations. Subsequently, a variety of promising photon sieves were fabricated and experimentally evaluated for use in a high-contrast, exoplanet imaging telescope.

1.2 Motivation

Currently prevalent methods of exoplanet detection are severely limited, with the ability to discover only a small fraction of exoplanets and biased towards those that are large and orbit closely to their parent star. These limitations are discussed in greater depth in Section 1.3. Direct imaging of exoplanets could complement and eventually outperform the existing methods. This would be hugely beneficial to the field of exoplanet research, especially if the direct imaging method were advanced to the point where it could be used to study earth-like exoplanets. Gathering information about large numbers of exoplanets would enable statistical studies that could provide insight into planetary system formation and how common habitable planets (and by extension life) might be in the Milky Way galaxy. Though a small number of exoplanets have been imaged successfully with conventional optics, new and innovative scientific and engineering solutions are needed to make the direct imaging method truly useful. Shaped, apodized photon sieves are one solution worth investigating. Additionally, exoplanet research is a driving force in developing new methods for high-contrast imaging which could drive further advances in many scientific fields.

1.3 Exoplanet Detection

Though telescopes such as NASA's Kepler Space Telescope (KST) are able to collect data that allows scientists to detect exoplanets and characterize their size and orbits, currently successful detection methods have severe limitations. The most successful detection method, the transit method, has been used by astronomers analyzing four years of KST mission data to detect over 1,000 independently confirmed exoplanets and over 4,000 exoplanet candidates [17]. KST, which was designed specifically for exoplanet detection, has even discovered eight near earth-size exoplanets orbiting in their stars' habitable zone [17]. Though the scientific and engineering achievements

are truly groundbreaking, a large fraction of exoplanets will go undetected using this method.

The transit method works by observing a star while its planet passes between the observer and the star. The flux collected from the star is recorded as a function of time, which results in a light curve. If there is no transit, the flux collected from the star will be constant as long as all other variables such as noise and atmospheric effects are accounted for. If there is a transit, there will be a dip in the light curve due to the fact that some of the light from the star is blocked by the exoplanet. Analyzing the light curve's shape provides insight into the relative sizes of the star and the exoplanet [36]. Determining the exoplanet's orbital period requires extended observation since at least two successive transits must be observed [36]. This is the first major limitation of the transit method. For earth-like exoplanets, this requires exceedingly long observations of a single patch of sky, especially since multiple consistent, successive dips in the light curve at regular time intervals are required (among other things) to confirm the presence of an exoplanet. The four years of data for a single patch of sky (0.25% of the full sky) acquired by KST are just enough to draw definitive conclusions about the orbital characteristics of earth-like exoplanet candidates.

The second major limitation of the transit method is that with all other factors being perfect the possible discovery space is still limited to only those exoplanets which have an orbit that carries them between earth and their parent star. In other terms, the earth must be very close to the exoplanet's orbital plane, otherwise no transit takes place and no dip in the light curve can be observed. Given that "the range of inclinations for which an exoplanet produces a transit is largest for planets close to their parent star," earth-like exoplanet detection using the transit method has unfavorable odds [36, p. 303]. The transit method is most suited for discovering "hot Jupiters" which are large exoplanets (since they cause a larger dip in the light

curve) that are close to their parent stars and have fast orbital periods [36, p. 304]. This is a likely alternative as to why such a large fraction of KST's discoveries are hot Jupiters compared to the explanation that hot Jupiters are more prevalent in the universe than earth-like planets [36, p. 304].

The second most prevalent method of exoplanet detection is the radial velocity method. Stars do not hold a fixed position in space, and stars and planets actually orbit each other's mutual center of mass. Since stars are much more massive, the center of mass of the system is close to the star's center of mass. This manifests as a "Doppler wobble," where the star's radial velocity can be observed to oscillate sinusoidally about the mean as it orbits the system's center of mass. Extended observations of a star must be made with a spectrometer to detect an exoplanet. The radial velocity is calculated from the Doppler shift present in the star's spectrum. The magnitude of the wobble will be more significant with a planet that has a large mass relative to its parent star, while the rate of the oscillation in the radial velocity is inversely related to the planet's orbital period. Therefore, like the transit method, the radial velocity method is most successful with massive exoplanets with short orbital periods: hot Jupiters.

Direct imaging of exoplanets, though challenging, has the potential to vastly increase the number of possible exoplanet discoveries. It does not require the observer lie in the same plane as the exoplanet; it does not necessarily require the exoplanet be massive (though a larger surface area to reflect light from its parent star is beneficial); and it does not necessarily require data be collected for an extended duration. A discovery can be made with a single image. Additionally, unlike the transit and radial velocity methods, direct imaging benefits from the exoplanet being far from its parent star (which means a long orbital period). This is because greater separation makes it easier to resolve the exoplanet and parent star as two individual

point sources. Finally, in contrast to the transit and radial velocity methods direct imaging also benefits from the observer lying outside the exoplanet's orbital plane, since this vantage point increases the apparent separation between the parent star and exoplanet.

Direct imaging of exoplanets allows spectroscopic measurements to be made and provides significant insight into specific planetary characteristics such as atmospheric content and mineral makeup. Direct imaging has already been successfully demonstrated, albeit with large exoplanets at a relatively large angular separation from their star, using angular differential and coronagraphic imaging [29]. Unfortunately, the bandwidth limitations of a conventional photon sieve telescope would render spectroscopic measurements impossible, though it is possible that the bandwidth limitations of the photon sieve could be exploited somewhat to increase the signal-to-background ratio for an imaged exoplanet (referred to for the remainder of the document as the signal-to-noise ratio (SNR)). Overall, the advantages of a lightweight, deployable, large aperture space telescope could easily outweigh the disadvantages. The two main challenges of directly imaging exoplanets are that the angular separation between the star and planet is extremely small and that the intensity of the starlight is much greater than the intensity of the light reflected from the exoplanet. This means that the signal intensity from the planet is effectively washed out in the intensity from its nearby parent star, which could be considered as noise.

1.4 Fresnel Zone Plate and Photon Sieve Theory

A brief introduction to the theory of Fresnel zone plates (FZPs) is presented. For thorough discussion including diagrams reference Chapter II of Major Christopher Tulip's work [42].

At its simplest, the photon sieve consists of a FZP where holes allow light to

pass through the light zones. Since photon sieves are based on FZPs, the same equations apply in many cases, though there are some subtle differences. FZPs work by passing alternating Fresnel zones, which interfere constructively given that there is 2π phase difference between the light passing through the center of subsequent alternative zones [24, p. 495]. A Fresnel zone, or half-period zone, can be constructed by dividing up a wavefront into rings such that when the light passing through those rings arrives at an on-axis point P a certain distance away, the light from the center of each successive zone is perfectly out of phase. This is because as the radial distance of the zone center from the optical axis increases, the path distance to P increases slightly, so the phase advances further and slightly leads the phase of the light from the inner ring [41].

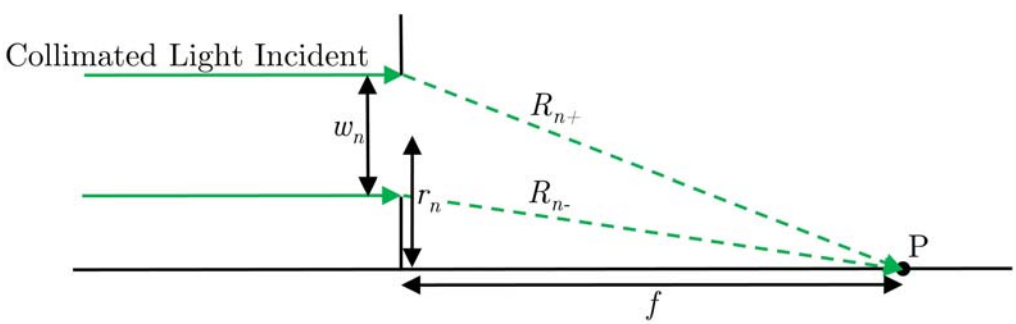


Figure 1. A cross section showing the geometry of a single FZP light zone is labeled using the same notation as is used in the equations presented in this section.

Choosing an optical path difference of $n\lambda$ between zones, where n is an integer indicating the zone number and λ is the wavelength will result in alternating zones that provide constructive interference. P is the point on the optic axis, and the distance from the center of the zone plate where $n = 0$ to P is chosen as the focal length, f [32]. As displayed in Fig. 1 the distance from P to the center of zone n is R_n, and the radial distance from the optic axis to the center of zone n is r_n. Since the desired optical path difference between zone centers is $n\lambda$, $R_n - f = n\lambda$, and

$R_n = n\lambda + f$ [32]. Since R_n is the hypotenuse of a right triangle and f and r_n are the legs, this can be substituted into the Pythagorean theorem to solve for the radial distance,

$$r_n = \sqrt{2nf\lambda + n^2\lambda^2}. \tag{1}$$

Given that λ is small, which it is in this case, λ^2 will be extremely small. Therefore, Eq. 1 can be truncated to

$$r_n \approx \sqrt{2nf\lambda}. \tag{2}$$

This can be solved for f to yield

$$f \approx \frac{r_n^2}{2n\lambda}. \tag{3}$$

Solving for the zone width, w, is slightly more involved. The radial distance to the inner zone edge is $r_n - \frac{w}{2}$, and the radial distance to the outer zone edge is $r_n + \frac{w}{2}$. Given that the distance from P to the outer zone edge is R_{n+} and the distance from P to the inner zone edge is R_{n-}, the path difference is

$$R_{n+} - R_{n-} = \frac{\lambda}{2}. \tag{4}$$

Using the Pythagorean theorem again, with $R_{n\pm}$ as the hypotenuses and $r_n \pm \frac{w}{2}$ and f as the legs, gives

$$R_{n\pm} = \sqrt{f^2 + \left(r_n \pm \frac{w}{2}\right)^2}. \tag{5}$$

Performing a binomial series expansion on each of the equations contained in Eq. 5, truncating to the first two terms, and substituting them both into Eq. 4 at the same time allows the zone width to be solved for as

$$w_n = \frac{f\lambda}{2r_n}. \tag{6}$$

It is important to note that $f > r_n + \frac{w}{2}$ is required in order for the binomial series to converge. More simply, the zone width can be solved for knowing that zones must have constant area $\pi\lambda f$ [2, p. 2976]. The combined area of all the light zones equals half the area of the full optic.

These are the basic equations required to design a photon sieve. It is possible to increase the photon sieve hole diameters, d, so that they are greater than the underlying Fresnel zones of width w. In fact, for a light-zone sieve, the optimum ratio is $d/w = 1.53$, as outlined in Section 2.4.7(p. 37) and covered in detail in the cited publications [2, 13, 27].

1.5 Shaped, Apodized Optics

As discussed in Section 1.1(p. 1), direct imaging of exoplanets is extremely challenging. Tailoring the PSF produced by a photon sieve is one possibility for overcoming these challenges. The PSF is the result of imaging a point source through an optical system, and can be thought of as the response of an optical system to a point source impulse. It tells how an optical system will spread and redistribute the energy emitted by a point source when it is imaged through an optical system and onto a detector. The point spread function of a diffraction limited, circular lens is the Airy pattern, shown in Fig. 2. As displayed, the irradiance does not stay below 10^{-4} of the peak until the radial distance from the peak is more than seven times the Rayleigh criterion resolution limit.

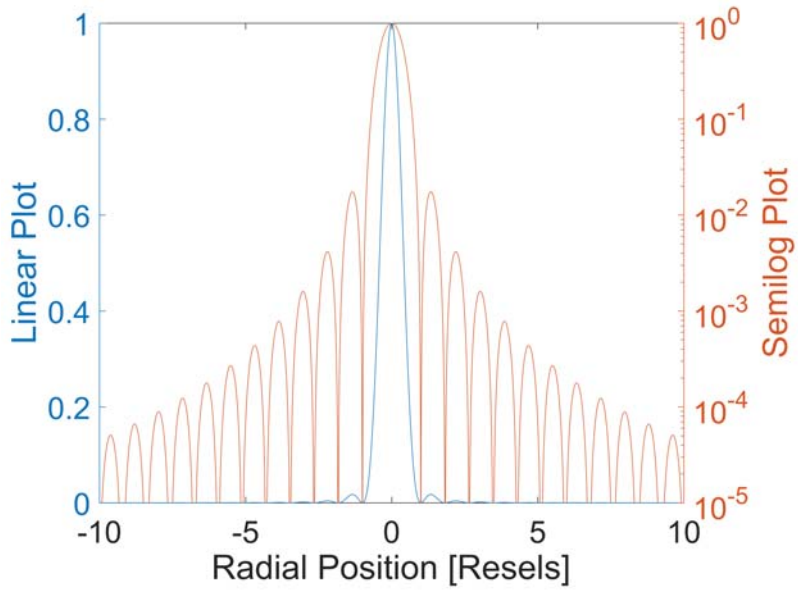


Figure 2. The Bessel function of the first kind of order one (the cross section through the peak of an Airy pattern) is shown plotted on both a linear scale and a semilog scale as a function of radial position in resels (Rayleigh criterion resolution elements).

Nisenson and Papaliolios have shown through simulation that square, apodized conventional optics are very promising for exoplanet imaging [34]. Since a square aperture distributes most of the diffracted energy along the axes, the irradiance of the PSF drops off rapidly along the diagonal, as shown in Fig. 3. Applying an appropriate apodization to the aperture further shapes the PSF, helping an exoplanet stand out. Nisenson and Papaliolios argued that (for conventional optics) with crossed Sonine and cosine apodized square apertures, exoplanet detection is possible for a 10^{-9} PVR with a 1/72 wave rms WFE and a separation of only six diffraction limits, or resolution elements (resels) [34].

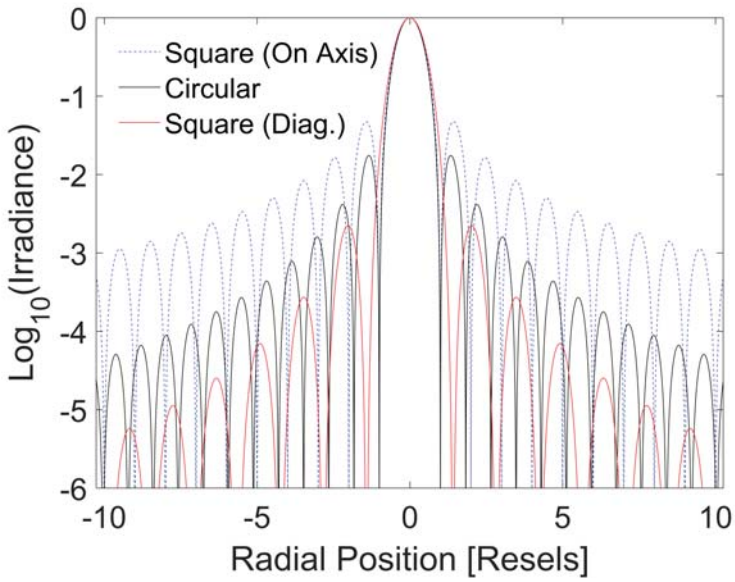


Figure 3. Theoretical $\log_{10}$ cross sections through PSFs produced by square and circular apertures. Along the diagonal the square aperture's PSF is shown to drop off more rapidly from the peak than the Airy pattern.

Much of what applies to conventional optics also applies to photon sieves, so it is plausible that results could be similar to Nisenson's. Cao and Jahns also presented a FZP design to produce a modified, Gaussian-like PSF by varying the zone width as a function of radial distance — similar to apodization [14]. The additional benefits of photon sieves make square, apodized photon sieves worth investigating.

II. Simulation

Wave optics simulations were conducted to help choose photon sieve designs that showed promise for exoplanet imaging. In order to produce timely results with the simulation, the Fresnel approximation to the Rayleigh-Sommerfeld integral and the numerical approximation to the Fresnel integral were both assumed to be reasonably valid. The simulation was built on these two key assumptions. The numerical Fresnel diffraction integral itself and a number of MATLAB programs provided by Schmidt ("`one_step_prop.m`", "`ft2.m`", "`circ.m`", "`rect.m`", and "`example_square_prop_one_step.m`") were used and adapted for the simulations [37].

2.1 Fresnel Diffraction Theory

Unlike in previous works, a simple Fourier transform was not sufficient to model the propagation of light through the photon sieve. With simulated photon sieve diameters on the order of 1 cm and individual hole diameters as large as 0.14 mm, Fraunhofer (far field) diffraction could not suitably model a 0.5-m propagation, as the focal plane of the photon sieve is firmly in the near-field. The Rayleigh-Sommerfeld diffraction integral provides the most accurate results, but at the expense of increased computational difficultly. The Rayleigh-Sommerfeld integral is written

$$U_2(x,y) = \frac{z}{j\lambda} \int \int_{\Sigma} U_1(\xi,\eta) \frac{\exp(jkr_{12})}{r_{12}^2} d\xi d\eta, \tag{7}$$

where the variables are as shown in Fig. 4. The field at a position in the observation plane is $U_2(x,y)$, z is the propagation distance to the observation plane, j is the imaginary unit, λ is the illuminating wavelength, Σ is the aperture, $U_1(\xi,\eta)$ is the field at a position in the source plane, k is the wavenumber $(2\pi/\lambda)$, and r_{12} is the

distance between a position in the source plane and a position in the observation plane [22, 37, 44]. It should be noted that the source plane as referenced in this investigation is the aperture plane (or the photon sieve plane) as shown in Fig. 4, and not the plane containing the true initial light source (the star). This naming convention conforms with Voelz and Schmidt, though Goodman consistently calls it the aperture plane instead of the source plane [22,37,44]. Simplifications can be made by assuming all light rays are relatively paraxial.

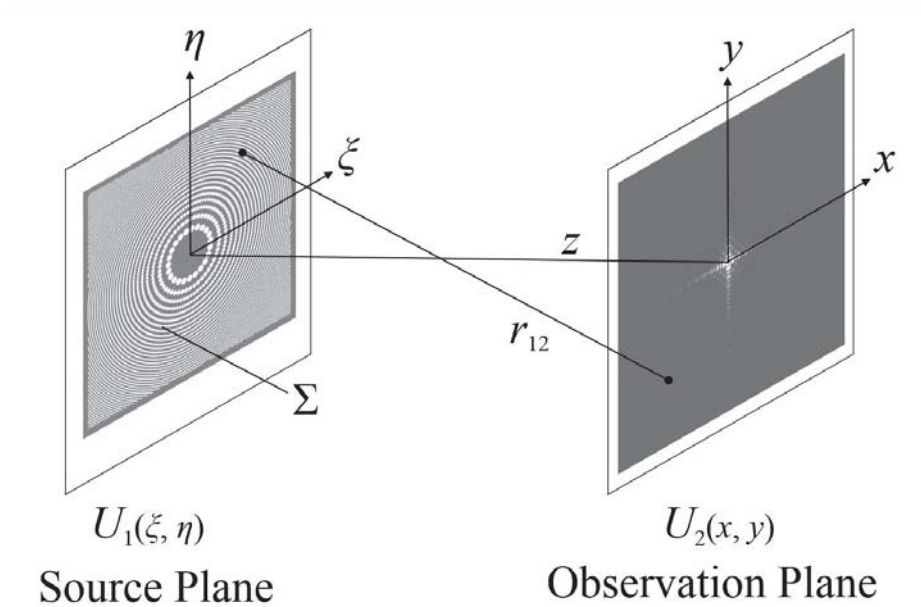


Figure 4. Fresnel diffraction propagation geometry assuming parallel source and observation planes, which is in accordance with the simulation and the experimental work. The source plane is the plane of the diffracting aperture (the photon sieve). This figure was inspired by Voelz's Figure 4.2 [44, p. 52].

The Fresnel integral is the first major step in approximating the Rayleigh-Sommerfeld integral. The distance between a given position in the source plane and a given position in the observation plane can be found in useful terms using the distance formula. This results in a square root inside an exponential in the integrand of the Rayleigh-

Sommerfeld integral. This distance term is labeled as r_{12} in Fig. 4, and is written

$$r_{12} = \sqrt{z^2 + (x-\xi)^2 + (y-\eta)^2}, \tag{8}$$

as described by [44, p. 52].

This makes computational simulations cumbersome but it can be closely approximated by applying a binomial expansion and truncating to the first two terms, which gives

$$r_{12} \approx z\left[1 + \frac{1}{2}\left(\frac{(x-\xi)}{z}\right)^2 + \frac{1}{2}\left(\frac{(y-\eta)}{z}\right)^2\right]. \tag{9}$$

This approximation is then applied to r_{12} in the phase of the exponential in Eq. 7, which is equivalent to making a parabolic approximation to a spherical wave from the point sources [44, p. 53]. Additionally, in the denominator inside the integrand of Eq. 7, the approximation $r_{12} \approx z$ is applied [44]. This results in the Fresnel diffraction expression, which is written

$$U_2(x,y) = \frac{e^{jkz}}{j\lambda z}\int\int_{\Sigma} U_1(\xi,\eta)\exp\left\{\frac{jk}{2z}\left[(x-\xi)^2 + (y-\eta)^2\right]\right\} d\xi d\eta. \tag{10}$$

The numerical approximation to this integral was used for the simulation portion of this research. The numerical approximation is accomplished by dividing the source plane into a grid and sampling based on that grid. For improved computational speed, the numerical approximation is built around the two-dimensional fast Fourier transform (FFT) algorithm, with the exponential phase term and scaling term being multiplied through after the FFT is completed. The numerical approximation to the Fresnel diffraction expression (Eq. 10), is written

$$U_2(x,y) = \frac{1}{j\lambda z}\exp\left\{\frac{jk}{2z}\left(x^2+y^2\right)\right\} \times FT2\left(U_1(\xi,\eta)\exp\left\{\frac{jk}{2z}\left(\xi^2+\eta^2\right)\right\}\right), \tag{11}$$

where FT2 is a simple function written by Schmidt that calls MATLAB®'s two dimensional FFT algorithm and then rearranges the output (an fftshift) so that low spatial frequencies are at the center of the matrix and high spatial frequencies are at the outside edges of the matrix [37].

2.2 Validity of the Numerical Fresnel Approximation

The Fresnel number can be used to represent the suitability of the Fresnel approximation, and is given by

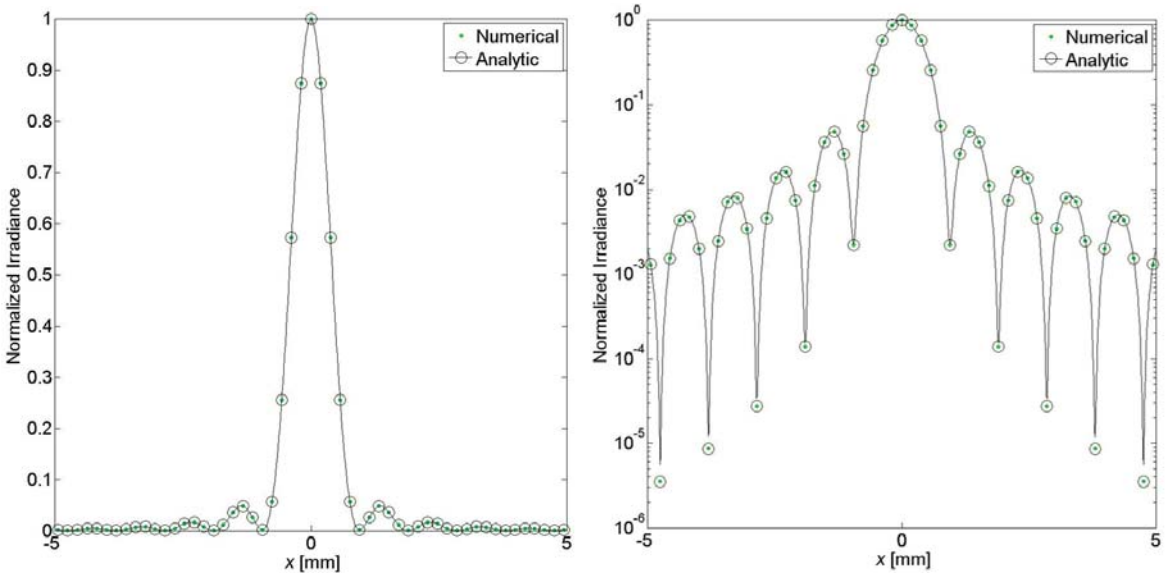
$$N_F = \frac{w^2}{\lambda z}, \tag{12}$$

where w is either the radius of a circular aperture in the source plane or the half-width of a square aperture [44, p. 55]. A Fresnel number less than one indicates that the approximation is appropriate, but the Fresnel approximation can still provide good results for Fresnel numbers up to 20 or 30 [44, p. 55]. Given the largest individual hole in a simulated sieve had a radius of roughly 0.14 mm, the propagation distance is 0.5 m and the incident wavelength is 532.1 nm, this gives a Fresnel number of 0.073, which is comfortably in the Fresnel region. The smallest holes in a simulated sieve are roughly 0.021 mm, which gives a Fresnel number of 0.0016 — firmly in the Fresnel region. However, for the sieves tested experimentally in Chapter III(p. 44), the half-width was roughly 1.8 cm. With the same propagation distance and incident wavelength, this resulted in a Fresnel number of 1212, placing the photon sieve as a whole firmly in the Rayleigh-Sommerfeld region. Due to computational constraints, the half-widths of simulated sieves were smaller — approximately 0.35 cm, with $N_F = 46$. This indicates that while the diffraction model was highly accurate for propagation through each of the individual holes, the combination of errors resulting from the interference of the wavelets produced at each hole could produce an appreciable error for the sieve as a whole. A hole at the corner of a square photon sieve with a half width of

1.8 cm is 2.9° off-axis given a 50-cm propagation distance. At optical wavelengths, this results in an undesirable Fresnel number, but since the simulation is being used primarily to compare different sieve designs, it is a relatively safe assumption that the Fresnel approximation can still be used to make rough estimates about the relative performance of different designs, as well as gain insight into the general shape of a certain design's PSF. It is possible that since the size of the holes decreases with increasing distance from the optical axis, the approximation is better than is indicated by the 1000+ Fresnel number for an entire photon sieve. The Fresnel integral is the first of the two main sources of uncertainty in the simulation.

The other main source of uncertainty in the simulation is the numerical approximation to the analytical solution of the Fresnel integral. The field in the observation plane (U_2) is computed from the field in the source plane (U_1) by a one-step propagation method. After multiplying a sampled version of the source field by a quadratic phase term, a Fourier transform is performed. The Fourier transform is a two-dimensional discrete Fourier transform (DFT), and is computed with MATLAB's fast Fourier transform algorithm. Next, a scaling factor is used to convert to spatial coordinates from spatial frequency coordinates and the whole thing is multiplied by a second quadratic phase factor [37, p. 90]. The accuracy of the result is dependent on how well-sampled the source plane is. Additional artifacts can be introduced in the numerical result by insufficiently padding the source plane matrix. To reduce the impact of numerical artifacts, all source matrices were padded with zeros such that the width of the full matrix was at least five times the width of the photon sieve. This is referred to as having a pad factor of five. Based on visual observations of the simulated point spread function, it was expected that numerical artifacts were insignificant. A comparison between numerical and analytical results is shown in Fig. 5 for a simple square aperture with a pad factor of five and a Fresnel number of

roughly 0.07. The numerical and analytic irradiance results agree closely, even when displayed on a $\log_{10}$ plot.

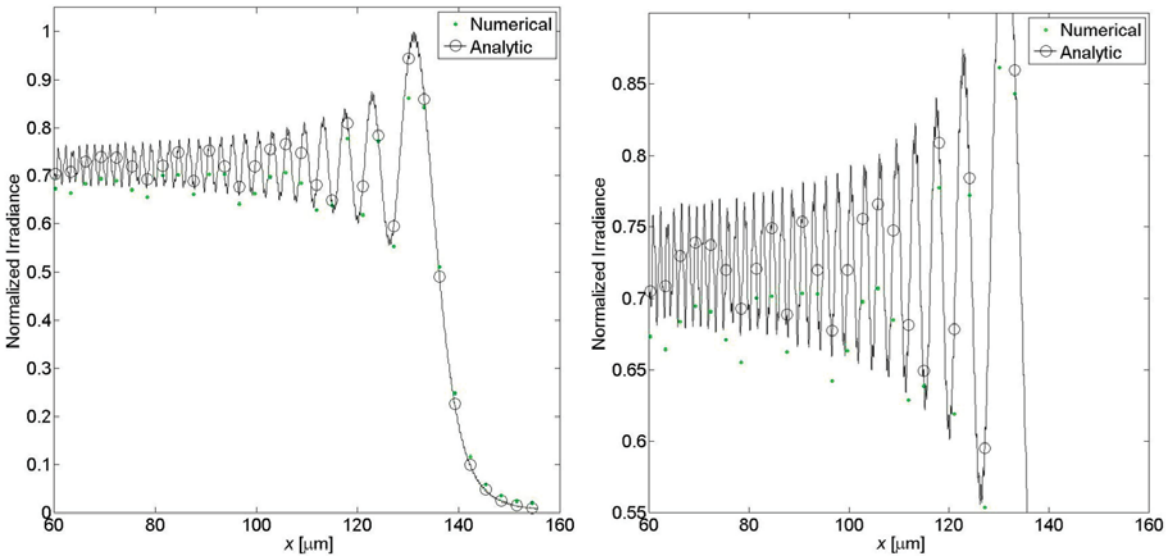


(a) (b)

Figure 5. Comparison of results from numerical and analytic Fresnel integrals for diffraction through a square aperture with $N_F \approx 0.07$. The numerical results are displayed as green data points, while the analytic results are displayed as black circles connected by a black line. The results are shown with irradiance plotted on a linear axis (a) and on a $\log_{10}$ axis (b). For this particular geometry the results are in close agreement. Irradiances were normalized separately for each method by dividing by the peak irradiance.

Differences between the numerical and analytic results start to become apparent above $N_F \approx 50$, with an example shown in Fig. 6. While it is understood that the Fresnel approximation to the Rayleigh-Sommerfeld diffraction integral breaks down at high Fresnel number, this shows that the discrepancy between simulation and reality will be exacerbated by the fact that the analytic and numerical results of the Fresnel approximation will begin to diverge. This suggests that using this method to model a photon sieve with $N_F > 1,000$ will give poor results, though this is not certain. The sieves tested in Chapter III(p. 44) had $N_F \approx 1,200$, and the sieve used in FalconSAT-7 has $N_F \approx 38,000$, given $w = 0.1$ m, $\lambda = 656.45$ nm, and

$z = 0.4$ m [6]. Taking this into consideration, it is suggested that future work focus on developing a more accurate simulation, perhaps based on work done by Cao and Jahns for high-numerical-aperture photon sieves [15].



(a) (b)

Figure 6. Comparison of results from numerical and analytic Fresnel integrals for diffraction through a square aperture with $N_F \approx 180$. The numerical results are displayed as green data points, while the analytic results are displayed as black circles connected by a black line. The results are shown with irradiance plotted with the y-axis starting at zero (a) to allow comparison with the low N_F results above in Fig. 5, and with a cropped y-axis (b) to better display the discrepancy between the two approaches. Note that the scale of the x-axis is now μm rather than mm as it is in Fig. 5 above. Irradiances were normalized separately for each method by dividing through by the peak irradiance.

2.3 MATLAB® Implementation

A series of wave optics simulations was conducted in order to aid in the design of the photon sieves. Both circular and square photon sieves were tested along with a number of different apodizations. Photon sieves were designed using MATLAB with code adapted from codes written by Andersen and Tulip [3, 31, 43]. Hole positions and diameters were saved for use in simulation. The code for the simulation was

adapted from code provided in *Numerical Simulation of Optical Wave Propagation* by Jason Schmidt and was run entirely using MATLAB [31, 37]. The simulations were conducted using a laptop with a CPU running at 2.40 GHz and 8 GB of RAM. The simulations were run with the highest sampling resolution possible in the source plane given the RAM limitation. After thoroughly optimizing the code for speed, simulation runtimes varied between 200 s and 3600 s. The runtime of the simulation was primarily dependent on the number of holes in the simulated photon sieve along with the size of the matrix used to represent the source plane.

Hole coordinates and diameters were calculated based on a desired focal length of 0.5 m and a desired operating wavelength of 532.1 nm. A binary matrix was calculated to represent the photon sieve, as shown in Fig. 7. Light incident on the sieve was approximated as a plane wave with zero tilt (and thus constant phase across the sieve), which is appropriate for an exceedingly distant source like a star outside the solar system. The desired simulation output was a time-averaged irradiance at the image plane, which is approximated as the focal plane for a distant source. Since the electric field is squared in calculating the time-averaged irradiance, the complex phase information is not preserved. Therefore, phase of the electric field in the image plane was not taken into consideration for the simulation.

For determining the PSF from a parent star, the binary matrix was multiplied by a constant representative of the PVR between the parent star and the exoplanet. The majority of the simulations were run with a PVR of 1/2,500 ($10^{-3.4}$), which is not representative of an actual extra solar system (where this ratio can be much smaller), but was sufficiently small to evaluate the photon sieve designs without filling the RAM of the test computer and crashing the program. A number of different apodizations were then applied by multiplying the functions across the mesh. Both circular and square photon sieves were tested. Square photon sieves were generated by using the

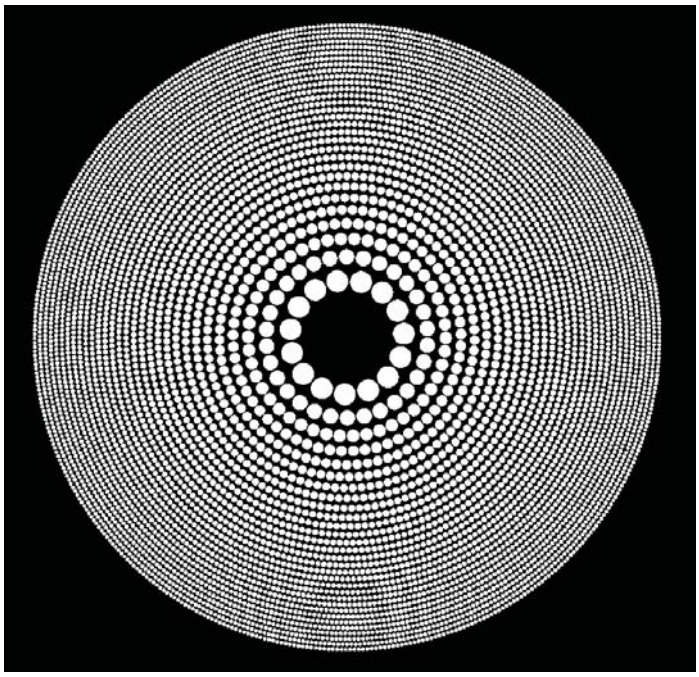


Figure 7. A circular photon sieve illuminated with a unit amplitude plane wave, modeled as a binary matrix of 1s and 0s in MATLAB. The 1s are displayed as white, and the 0s as black.

same code as the circular photon sieves and removing the holes on the edges. The aperture building and Fresnel diffraction code is provided in Appendix C(p. 145).

For this experiment, a number of different apodizations were simulated. The simulation results were used to select the most promising apodization functions and function parameters for photon sieve fabrication and experimental test. The most computationally intensive piece of code was the construction of the aperture matrix (the source plane) using the hole positions and diameters calculated using Eq. 1 and Eq. 6. Since the apodization matrix was applied to the aperture matrix through a simple multiplication, an aperture matrix could be saved and used many times with different apodizations, greatly increasing the efficiency of the code. A photon sieve with an apodization matrix applied to it is displayed in Fig. 8, in contrast to the unapodized sieve shown in Fig. 7. Note that the fabricated sieves differ from the simulated apertures because the apodization was applied by selectively removing

holes rather than by applying an amplitude mask. This is covered in Section 3.1.1(p. 45) where it is shown that the fabricated apodizations closely match the apodizing functions, so this should not be a significant source of uncertainty.

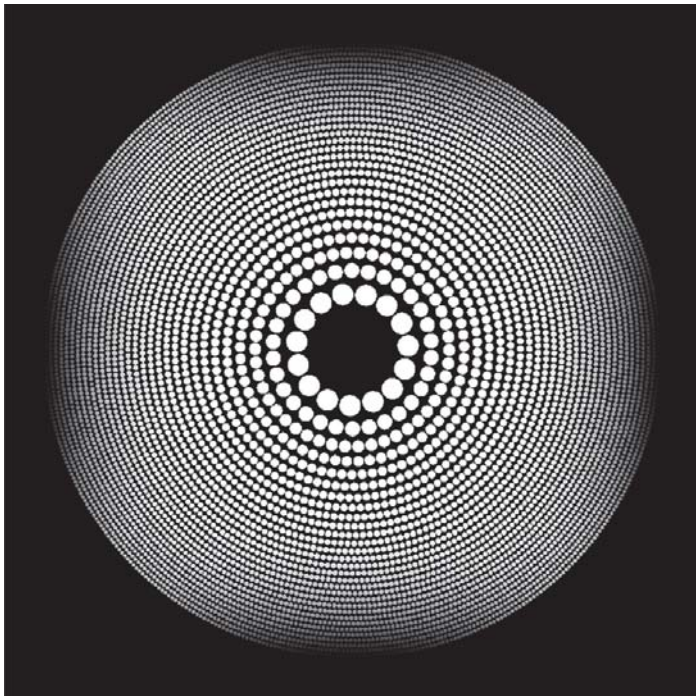


Figure 8. A circular photon sieve illuminated with a unit amplitude plane wave, modeled as a matrix of in MATLAB. An amplitude mask was then applied to apodize the sieve by multiplying the apodization function across the matrix. In this case the apodization function was a crossed Sonine function with $\nu = 4$. This apodization function can be seen to decrease the field amplitude at the edges of the sieve, where the white holes fade to black.

For each simulation, figures were automatically generated and saved and the relevant data were automatically saved to an Excel file. This was done for convenience, data synthesis and analysis, and to eventually enable the use of a supercomputer to run the simulation for future research. Considering that a supercomputer could complete hundreds of low-resolution simulations in rapid succession, implementation of an automated data recording system was necessary. It is suggested that future work involving photon sieve simulations use a supercomputer, such as the AFIT Linux cluster.

2.4 Numerical Fresnel Diffraction Simulations

This experiment focused on square photon sieves because shaping a PSF by using a square aperture rather than a circular one is an elegant and promising solution for exoplanet imaging [34]. The Fourier transform of the square aperture gives a 2-D Sinc, which is the field in the focal plane. Since most of the diffracted energy is concentrated along the axes perpendicular to the edges of the square, the irradiance in the PSF drops rapidly along the diagonals [22]. For a circular aperture, diffracted energy is distributed in the Airy pattern which is a series of concentric bright and dark rings. Since the diffracted energy is distributed with circular symmetry for a circular aperture, the PSF does not drop off as sharply as it does along the diagonals of the 2-D Sinc^2, as shown in Fig. 3(p. 12). Circular and square photon sieves appear to behave similarly to circular and square conventional optics, as shown in Fig. 30. To bring out the fainter details, simulated images of the PSF were scaled by taking the common logarithm ($\log_{10}$) of the data. Since initial numerical Fresnel diffraction simulations suggest that square photon sieves produce a PSF with diffracted energy concentrated along the axes, it is possible that the irradiance in the PSF decreases rapidly enough along the diagonal directions to be useful for exoplanet imaging. Additional beneficial changes to the PSF could possibly be accomplished by apodization.

In this simulation, a number of different transmission (apodization) functions were multiplied across the aperture, effectively acting as an amplitude mask. The purpose of apodization is to block and redistribute energy to shape a PSF to achieve a desired outcome. In this case, the desired outcome was suppression of secondary maxima (side lobes) to allow an exoplanet to stand out more than it would with a conventional, unapodized aperture. Unless otherwise stated, all simulations were run at the design wavelength of 532.1 nm. The focal length was 0.5 m and $d/w = 1.53$ [27, p. 185]. The circular sieve had 5,539 pinholes and the square sieve (with equal surface area to the

circular sieve) had 6,155 pinholes. The circular sieve had an $f/\#$ of 63.3. The same square sieve was used for all apodizations. The colorbars for the colormapped $\log_{10}$ images and apodization surface plots produced in this section are shown in Figures 9 and 10. The results of the simulation for the different apodizations investigated are presented and discussed below.

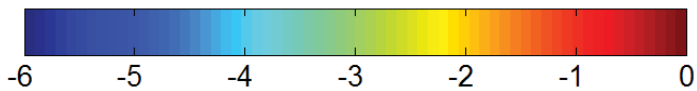


Figure 9. The color axis used for all $\log_{10}$, colormapped images. The labels of the colorbar denote the value of the exponent, e.g., "-6" on the colorbar corresponds to a data point with a value at 10^{-6} of the peak value.

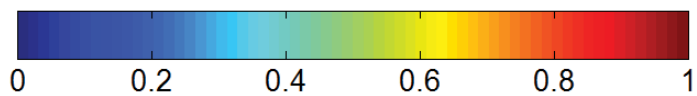
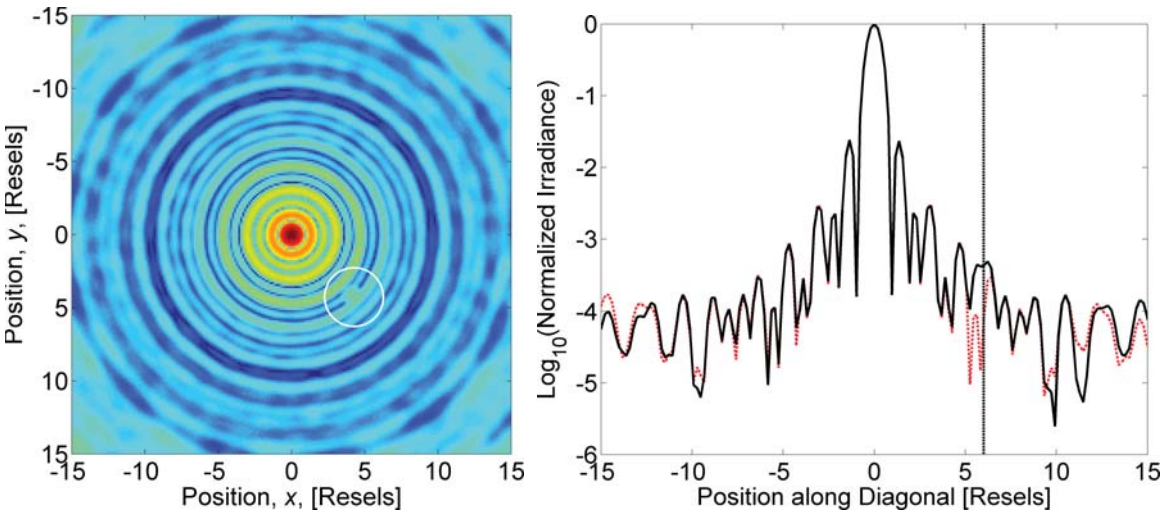


Figure 10. The color axis used for all colormapped apodization surface plots.

2.4.1 Circular Sieve.

The unapodized circular sieve is the most thoroughly studied type of photon sieve, but is likely the least promising for exoplanet imaging. However, for the sake of comparison a number of simulations were run using circular photon sieves. Fig. 11 shows that circular sieves produce rings similar to an Airy pattern, with significant energy outside the central lobe.

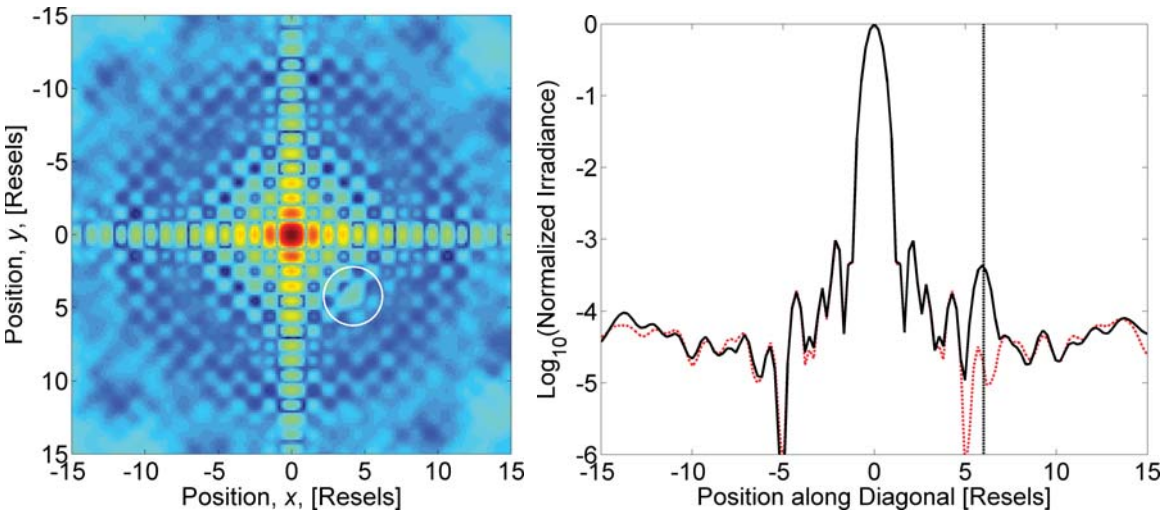


(a) (b)

Figure 11. Unapodized circular photon sieve normalized, $\log_{10}$ image (a) of two point sources with a 1/2,500 ($10^{-3.4}$) peak value ratio and six resels of separation. A diagonal cross section of the PSF is shown in (b), where the solid black line is with the presence of the exoplanet and the dotted red line is without. The position of the exoplanet is indicated by the white circle in (a) and the vertical, dotted black line in (b).

2.4.2 Square Sieve.

The unapodized square sieve was expected to show improvement compared to the unapodized circular sieve due to the fact that much of the diffracted energy in the PSF is concentrated along the axes. Fig. 12 supports these hypotheses, though the diagonal cross section did not improve as significantly as in simulations of conventional optics [34].

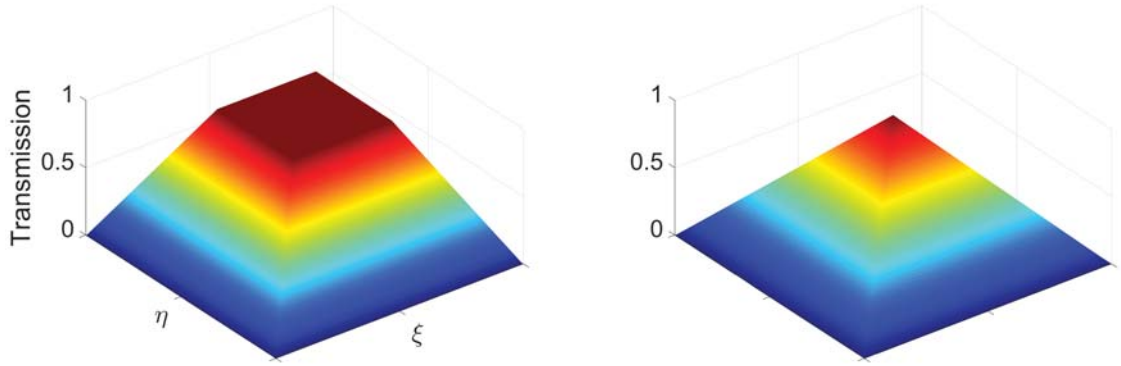


(a) (b)

Figure 12. Unapodized square photon sieve normalized, $\log_{10}$ image (a) of two point sources with a 1/2,500 ($10^{-3.4}$) peak value ratio and six resels of separation. A diagonal cross section of the PSF is shown in (b), where the solid black line is with the presence of the exoplanet and the dotted red line is without. The position of the exoplanet is indicated by the white circle in (a) and the vertical, dotted black line in (b).

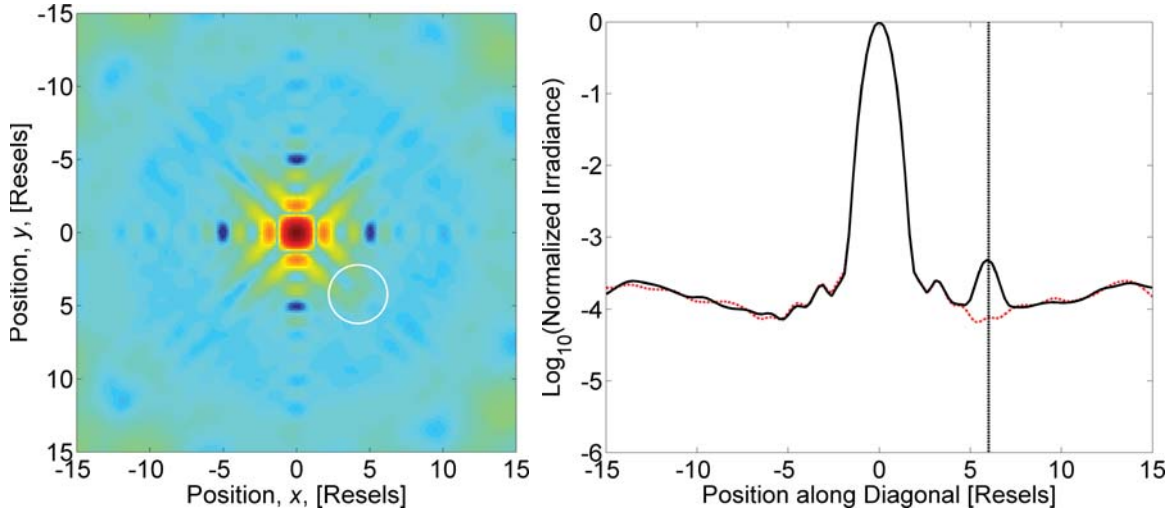
2.4.3 Pyramid Apodized Square Sieve.

This square pyramid apodization is an original apodization. Though this type of apodization does not appear in the literature, it is quite similar in nature to the crossed Sonine apodization. This function was selected for investigation because of its simplicity and its similar shape to the crossed Sonine apodization. The shape of the apodization suggests that it may block and redistribute diffracted energy favorably. The code was written such that a flat top could be applied to the square pyramid to increase throughput, as shown below in Fig. 13. With no flat top applied the throughput is reduced by 2/3 relative to an unapodized sieve. The overall transmission increases as the size of the flat top increases. Figures 14 and 15 show that apodization favorably smooths out the side lobes to produce a cleaner PSF, but does not reduce the PSF's overall floor. It is unclear whether or not this is more favorable for exoplanet detection.



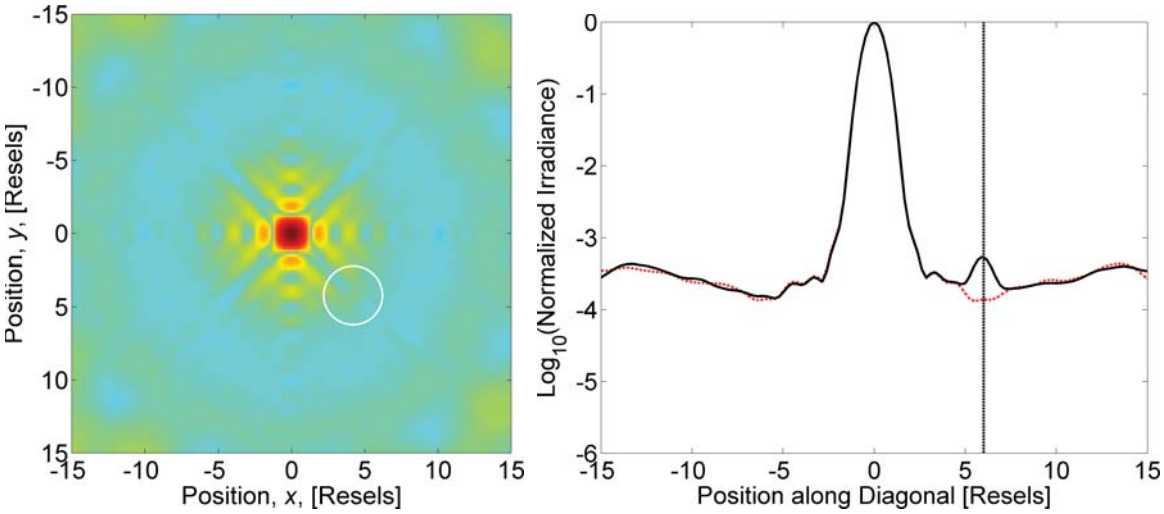
(a) 16% flat top by area (b) No flat top

Figure 13. The shape of the square pyramid apodization function with a 16% flat top by area (a) and no flat top (b). This apodization is a numerically generated square pyramid which slopes linearly from 100% transmission at the center to zero transmission at the edges and corners. Dark red corresponds to full transmission and dark blue to zero transmission.



(a) (b)

Figure 14. Pyramid apodized (16% flat top) square photon sieve normalized, $\log_{10}$ image (a) of two point sources with a 1/2,500 ($10^{-3.4}$) peak value ratio and six resels of separation. A diagonal cross section of the PSF is shown in (b), where the solid black line is with the presence of the exoplanet and the dotted red line is without. The position of the exoplanet is indicated by the white circle in (a) and the vertical, dotted black line in (b).



(a) (b)

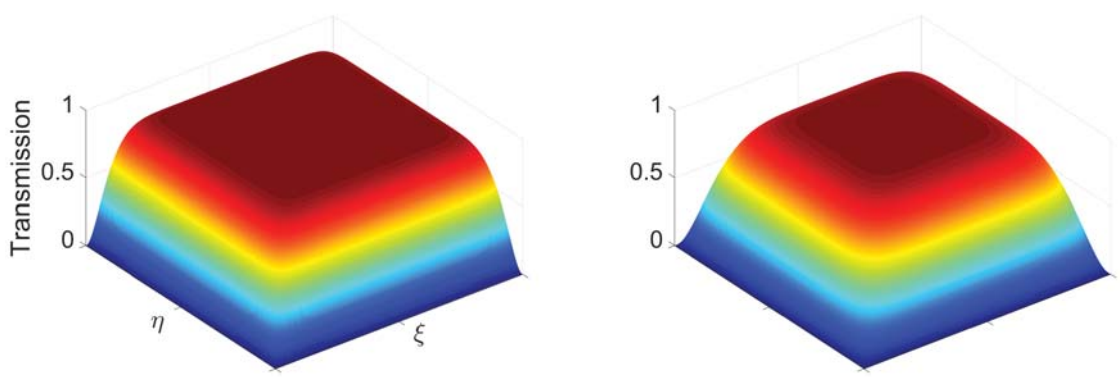
Figure 15. Pyramid apodized (no flat top) square photon sieve normalized, $\log_{10}$ image (a) of two point sources with a 1/2,500 ($10^{-3.4}$) peak value ratio and six resels of separation. A diagonal cross section of the PSF is shown in (b), where the solid black line is with the presence of the exoplanet and the dotted red line is without. The position of the exoplanet is indicated by the white circle in (a) and the vertical, dotted black line in (b).

2.4.4 Sonine Apodized Square Sieve.

Crossed Sonine functions provide transmission functions that take the form

$$T(\xi,\eta) = \left(1-\xi^2\right)^{\nu-1}\left(1-\eta^2\right)^{\nu-1}, \tag{13}$$

where T is the transmission between zero and one, ξ and η are the horizontal and vertical position coordinates in the source plane as shown in Fig. 4, and ν is a small integer such as 4, 5, or 6 [34]. The function is scaled such that the value of $T(\xi,\eta) = 0$ at the edge of the photon sieve. Beyond the edge of the photon sieve the value of the function is set to zero [34]. The shape of the apodization function is displayed in Fig. 16, along with how changes in ν affect the function's shape.

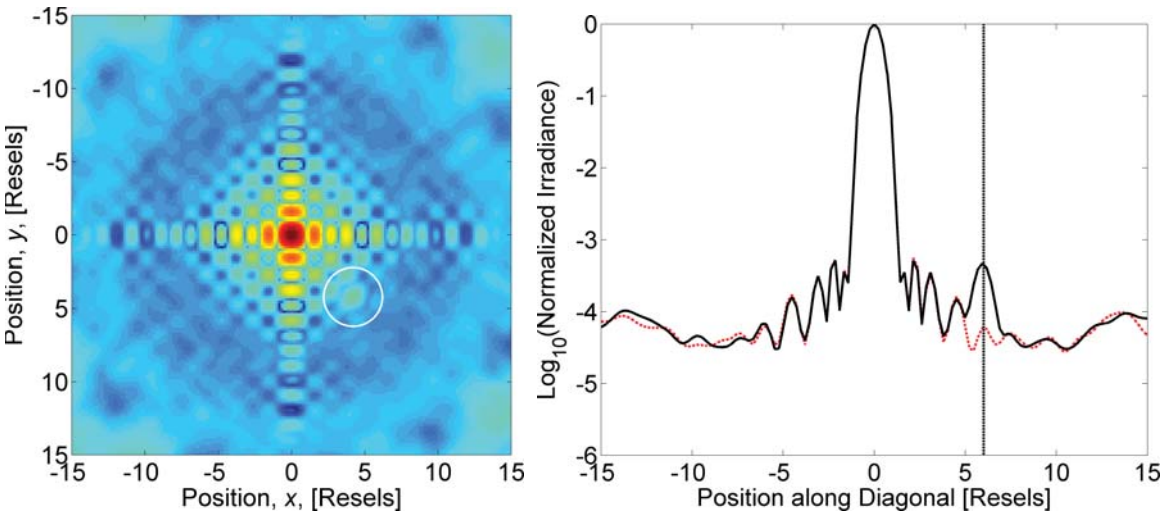


(a) $\nu = 7$ (b) $\nu = 4$

Figure 16. The crossed Sonine apodization function with $\nu = 7$ (a) and $\nu = 4$ (b) can be applied to a photon sieve so that the transmitted amplitude equals zero at the edges of the sieve. Dark red corresponds to full transmission and dark blue to zero transmission. Comparing (a) and (b) shows how increasing ν increases the overall transmission by causing the transmission to drop rapidly nearer to the edges. A similar apodization has been shown to improve the PSF for exoplanet detection with conventional optics [34].

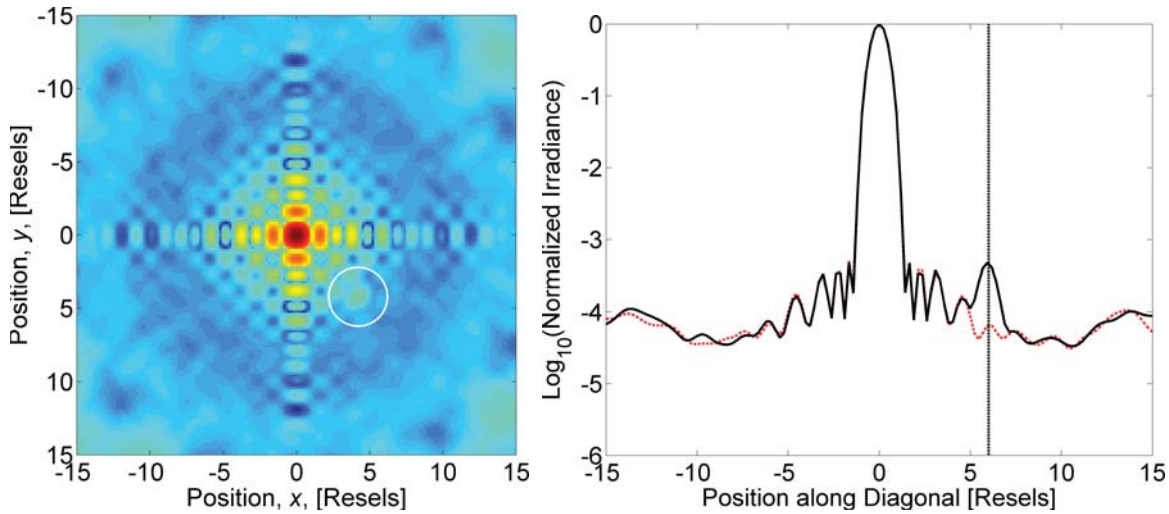
The crossed Sonine functions were the most successful apodization functions simulated with square conventional optics for exoplanet detection [34]. This made the

crossed Sonine apodization a logical first choice for simulation with a square photon sieve. As shown in Figures 17-20, the Sonine apodization works well for suppressing side lobes without raising the data's overall floor. Like the pyramid apodization it does not suppress the floor to the level achieved in simulations of conventional optics [34].



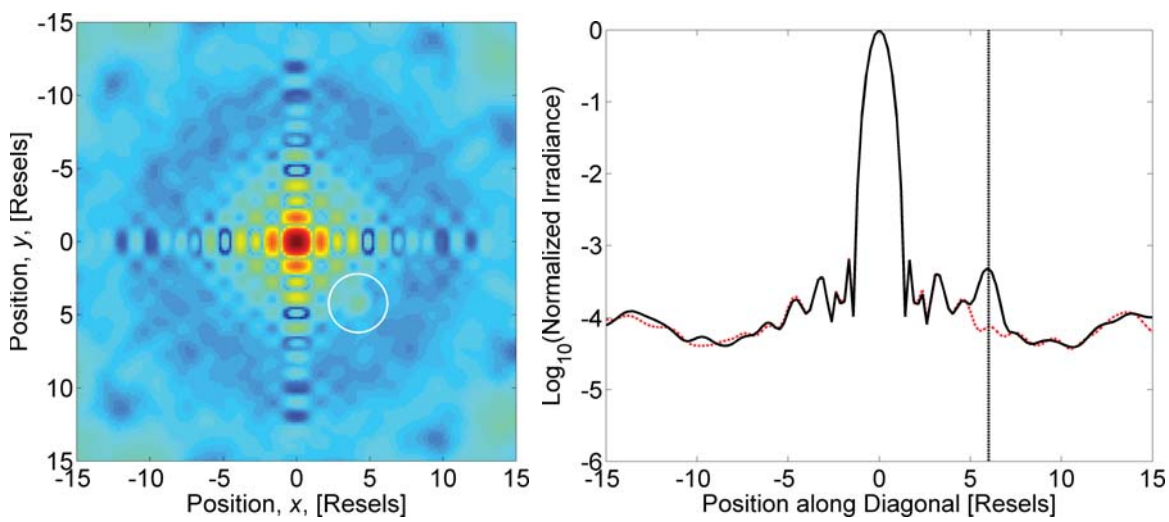
(a) (b)

Figure 17. Crossed Sonine apodized ($\nu = 6$) square photon sieve normalized, $\log_{10}$ image (a) of two point sources with a 1/2,500 ($10^{-3.4}$) peak value ratio and six resels of separation. A diagonal cross section of the PSF is shown in (b), where the solid black line is with the presence of the exoplanet and the dotted red line is without. The position of the exoplanet is indicated by the white circle in (a) and the vertical, dotted black line in (b).



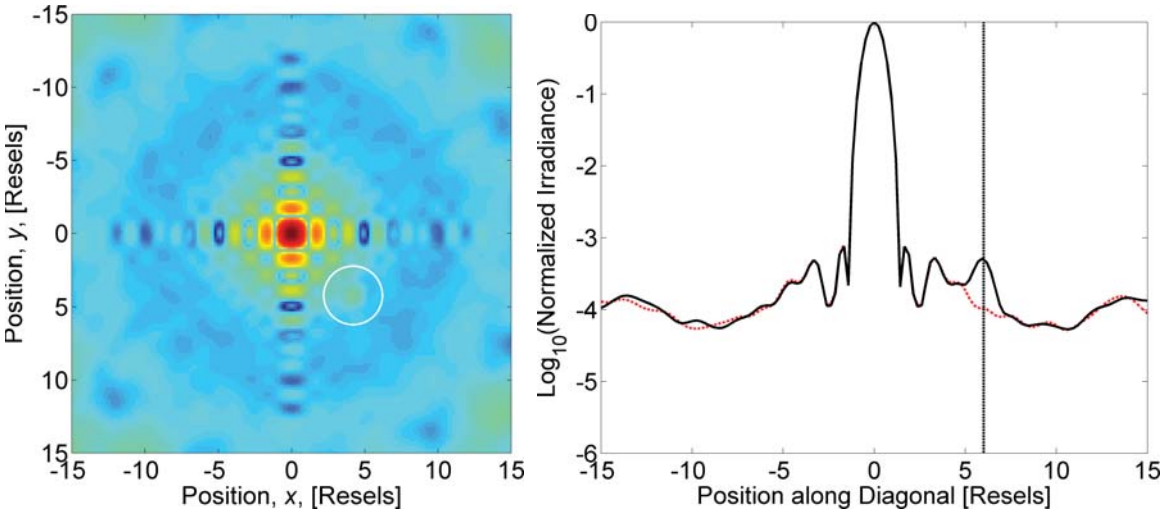
(a) (b)

Figure 18. Crossed Sonine apodized ($\nu = 5$) square photon sieve normalized, $\log_{10}$ image (a) of two point sources with a 1/2,500 ($10^{-3.4}$) peak value ratio and six resels of separation. A diagonal cross section of the PSF is shown in (b), where the solid black line is with the presence of the exoplanet and the dotted red line is without. The position of the exoplanet is indicated by the white circle in (a) and the vertical, dotted black line in (b).



(a) (b)

Figure 19. Crossed Sonine apodized ($\nu = 4$) square photon sieve normalized, $\log_{10}$ image (a) of two point sources with a 1/2,500 ($10^{-3.4}$) peak value ratio and six resels of separation. A diagonal cross section of the PSF is shown in (b), where the solid black line is with the presence of the exoplanet and the dotted red line is without. The position of the exoplanet is indicated by the white circle in (a) and the vertical, dotted black line in (b).

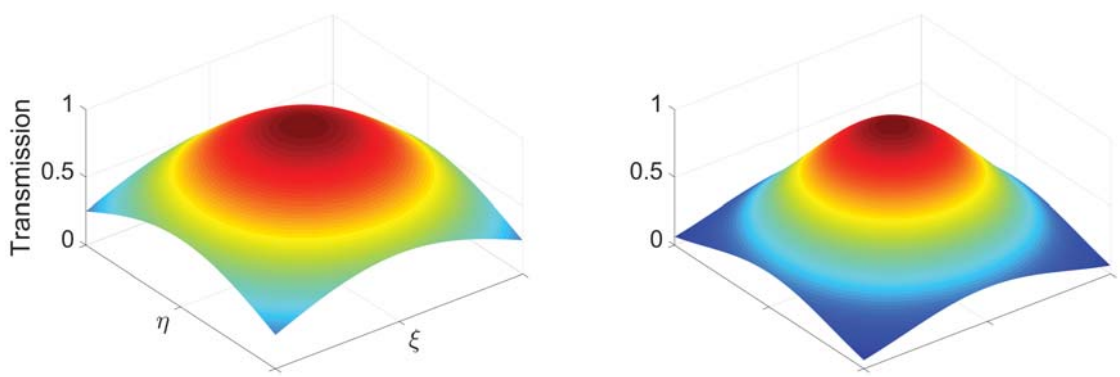


(a) (b)

Figure 20. Crossed Sonine apodized ($\nu = 3$) square photon sieve normalized, $\log_{10}$ image (a) of two point sources with a 1/2,500 ($10^{-3.4}$) peak value ratio and six resels of separation. A diagonal cross section of the PSF is shown in (b), where the solid black line is with the presence of the exoplanet and the dotted red line is without. The position of the exoplanet is indicated by the white circle in (a) and the vertical, dotted black line in (b).

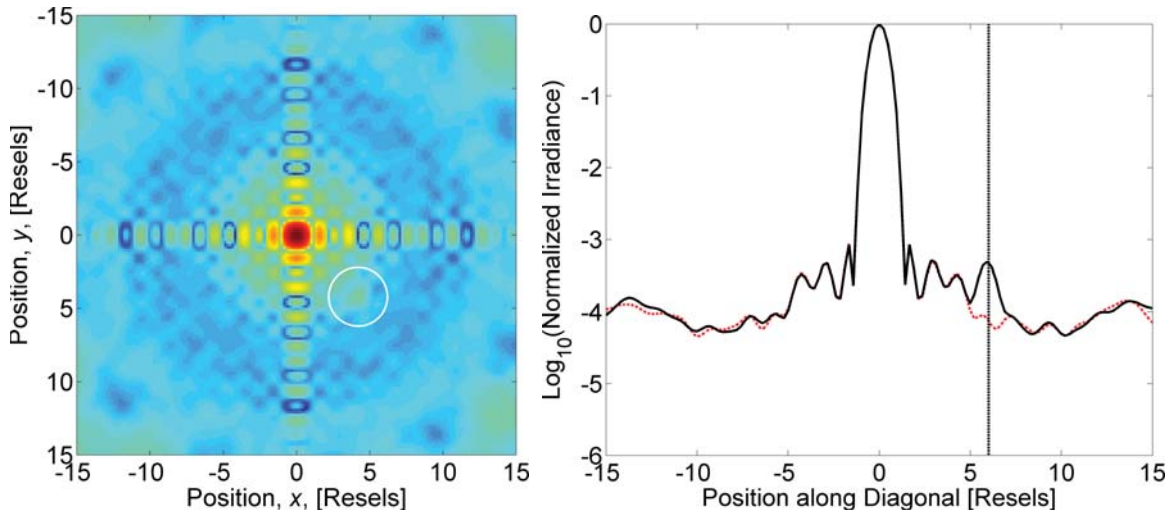
2.4.5 Cosine Apodized Square Sieve.

The cosine apodization either takes the form $\cos^2(\xi, \eta)$ or $\cos^4(\xi, \eta)$, which are both displayed in Fig. 21. This apodization was simulated successfully in the literature with square conventional optics, and therefore is promising for photon sieves as well. It has an overall transmission of roughly 20%, depending on how it is scaled [34]. Figures 22 and 23 show that the Cos^2 apodized square sieve produces results similar to the Sonine apodization, while the Cos^4 apodized sieves performs poorly and actually raises the side lobes, though they are smoothed.



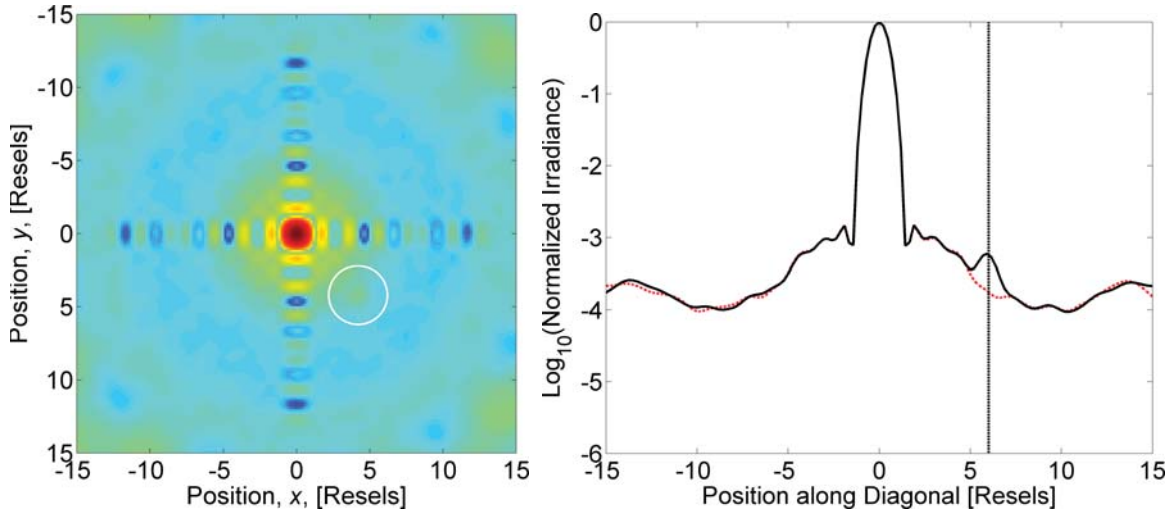
(a) $\cos^2(\xi, \eta)$ (b) $\cos^4(\xi, \eta)$

Figure 21. The cosine apodization functions can be applied to a photon sieve to shape the PSF. Dark red corresponds to full transmission and dark blue to zero transmission. Comparing (a) and (b) shows how $\cos^4$ decreases the overall transmission by falling off more rapidly. This apodization has been shown to improve the PSF for exoplanet detection with conventional optics [34].



(a) (b)

Figure 22. Cos^2 apodized square photon sieve normalized, $\log_{10}$ image (a) of two point sources with a 1/2,500 ($10^{-3.4}$) peak value ratio and six resels of separation. A diagonal cross section of the PSF is shown in (b), where the solid black line is with the presence of the exoplanet and the dotted red line is without. The position of the exoplanet is indicated by the white circle in (a) and the vertical, dotted black line in (b).



(a) (b)

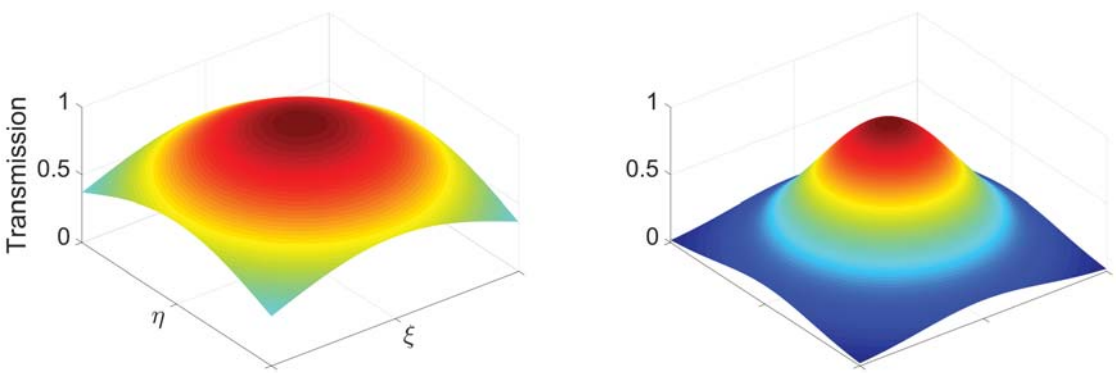
Figure 23. Cos^4 apodized square photon sieve normalized, $\log_{10}$ image (a) of two point sources with a 1/2,500 ($10^{-3.4}$) peak value ratio and six resels of separation. A diagonal cross section of the PSF is shown in (b), where the solid black line is with the presence of the exoplanet and the dotted red line is without. The position of the exoplanet is indicated by the white circle in (a) and the vertical, dotted black line in (b).

2.4.6 Gaussian Apodized Square Sieve.

The Gaussian apodization, which can be seen in Fig. 24, takes the form of the two-dimensional Gaussian where,

$$T(\xi, \eta) = \exp\left\{-\left(\frac{\xi^2}{2\sigma_\xi^2} + \frac{\eta^2}{2\sigma_\eta^2}\right)\right\}. \tag{14}$$

All Gaussian apodizations tested were symmetrical such that $\sigma_\xi = \sigma_\eta$. Varying σ varies the width of the Gaussian, or how quickly it falls off, as shown in Fig. 24.

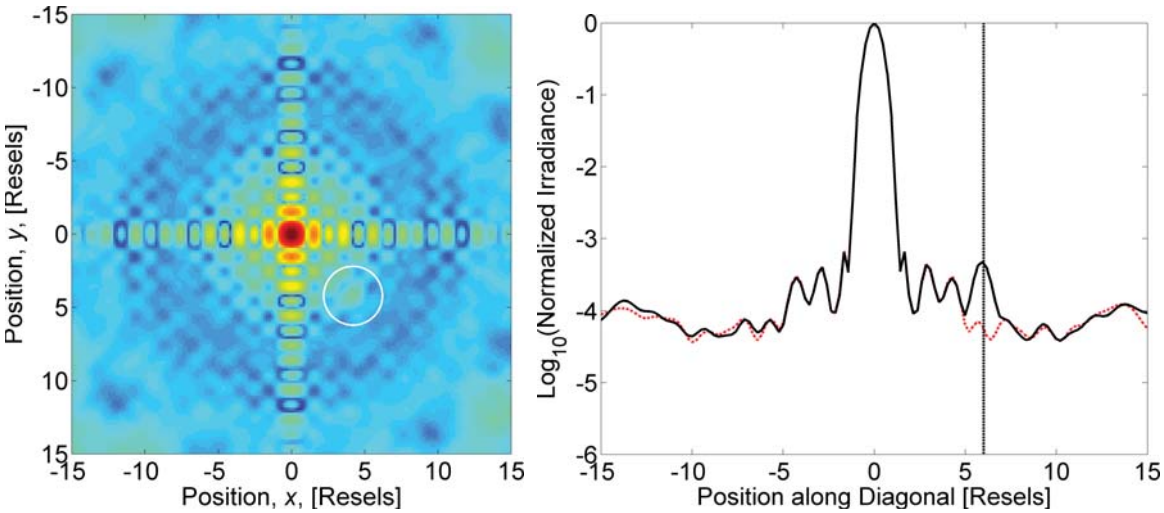


(a) $\sigma = 1$ (b) $\sigma = 0.5$

Figure 24. The Gaussian apodization function can be applied with varying widths to a photon sieve to shape the PSF. Dark red corresponds to full transmission and dark blue to zero transmission. Comparing (a) and (b) shows how decreasing the width decreases the overall transmission by falling off more rapidly.

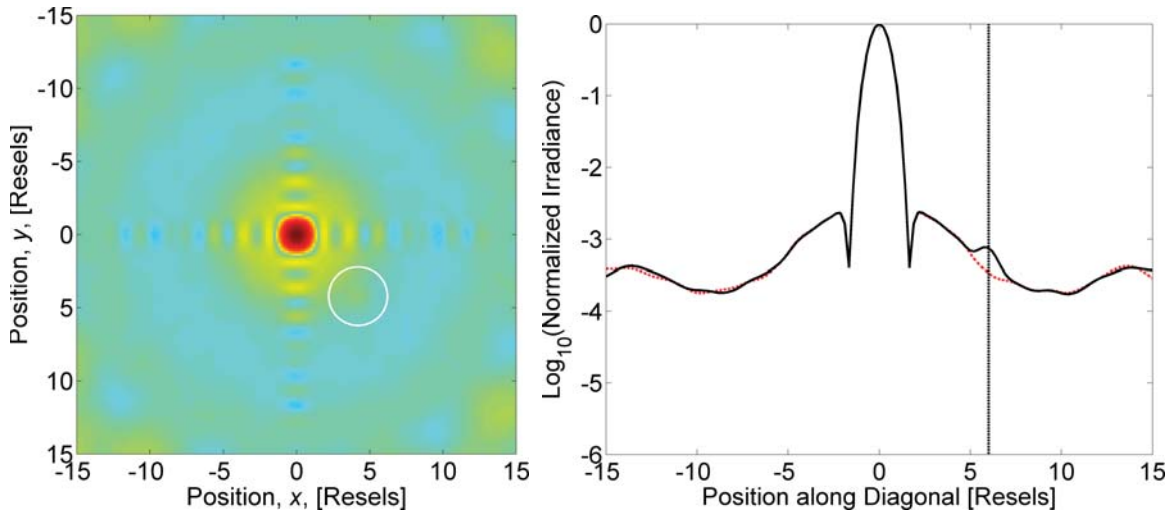
The Cosine and Gaussian apodizations have the benefit that they are easier than the Sonine and square pyramid apodizations to apply to a physical photon sieve, since they are circularly symmetric. The apodization can therefore be applied by simply thinning out the fill factor of the zones as the zone number, n, increases. For this research the fill factor was defined as the ratio between the total pinhole area in a zone to the area of the zone. Photon sieves can have fill factors greater than one, since the pinhole diameter can be greater than the zone width [2, p. 2976]. As displayed in

Figures 25 and 26, the $\sigma = 1.0$ Gaussian apodized square sieve performed similarly to the Sonine apodized square sieve and the $\sigma = 0.5$ sieve performed poorly.



(a) (b)

Figure 25. Gaussian apodized square photon sieve ($\sigma = 1.0$) normalized, $\log_{10}$ image (a) of two point sources with a 1/2,500 ($10^{-3.4}$) peak value ratio and six resels of separation. A diagonal cross section of the PSF is shown in (b), where the solid black line is with the presence of the exoplanet and the dotted red line is without. The position of the exoplanet is indicated by the white circle in (a) and the vertical, dotted black line in (b).



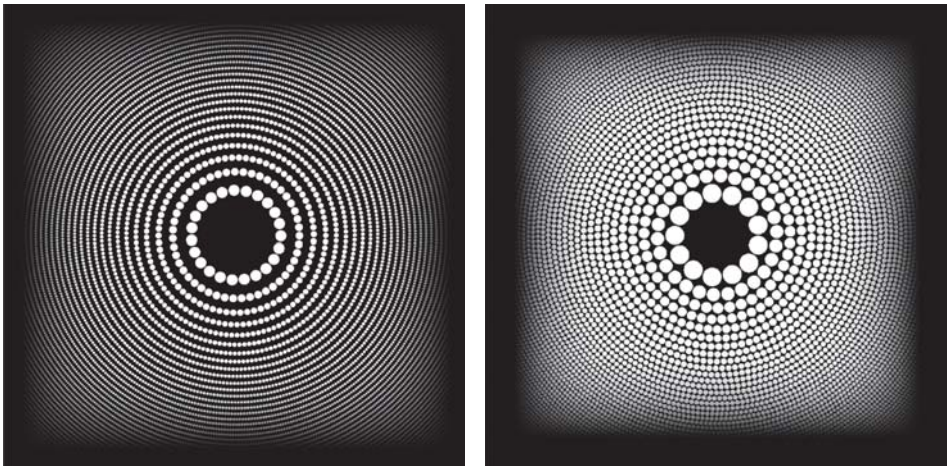
(a) (b)

Figure 26. Gaussian apodized square photon sieve ($\sigma = 0.5$) normalized, $\log_{10}$ image (a) of two point sources with a 1/2,500 ($10^{-3.4}$) peak value ratio and six resels of separation. A diagonal cross section of the PSF is shown in (b), where the solid black line is with the presence of the exoplanet and the dotted red line is without. The position of the exoplanet is indicated by the white circle in (a) and the vertical, dotted black line in (b).

2.4.7 Apodized Square Sieve with Varying d/w.

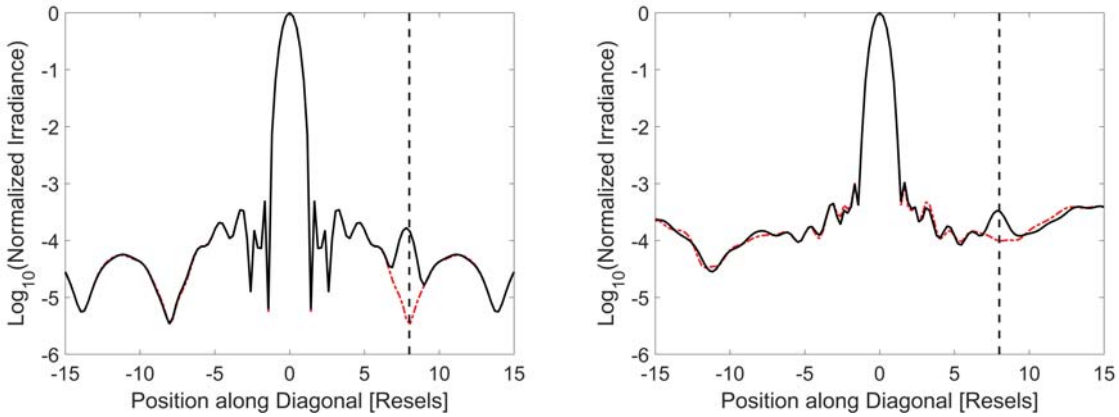
Square, Sonine apodized photon sieves with $\nu = 4$ were simulated with varying values of d/w, which is the ratio of pinhole diameter to zone width discussed briefly in Section 1.4. The ratios simulated were 0.8-1.8 in increments of 0.2. For comparison, the $d/w = 1.0$ and the $d/w = 1.8$ sieves are shown in Fig. 27.

The PSF became more favorable as d/w decreased, with the $d/w = 0.8$ sieve outperforming the $d/w = 1.8$ sieve considerably, as displayed in Fig. 28. However, due to manufacturing limitations and to maintain a reasonable scope for the thesis, varying d/w was not investigated further and may be a promising area for future research. The d/w ratio 1.53 was selected for the fabricated sieves since it is the smallest value of d/w resulting in a maximum for field strength at the focal point as a function of d/w [27, p. 185]; [13, p. 2390].



(a) $d/w = 1.0$ (b) $d/w = 1.8$

Figure 27. The widths, w, of the underlying Fresnel zones are the same for the sieves shown in (a) and (b). However, since the d/w ratio is selected to be different the hole sizes are different. The optimal ratio is 1.53 for normal sieves, but that may not be the case for apodized square sieves and may warrant future investigation.

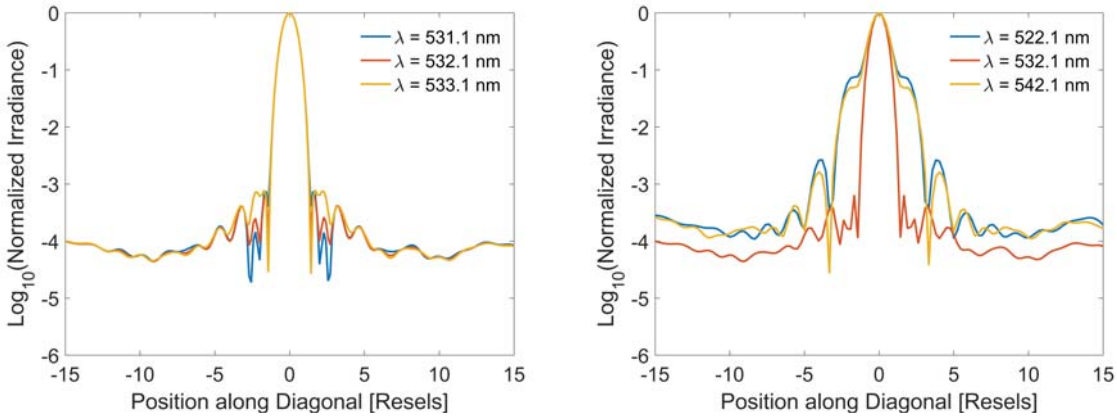


(a) $d/w = 0.8$ (b) $d/w = 1.8$

Figure 28. Sonine apodized ($\nu = 4$) square photon sieve normalized, diagonal cross sections through simulated $\log_{10}$ images of two PSFs with a 1/2,500 ($10^{-3.4}$) peak value ratio and six resels of separation with varying d/w. The solid black line is with the presence of the exoplanet and the dotted red line is without. The position of the exoplanet is indicated by the vertical, dotted black line. The PSF cross section for $d/w = 0.8$ (a) is more favorable with a central lobe that decreases more rapidly from the peak, more suppressed side lobes, and deeper nulls than with $d/w = 1.8$ (b).

2.4.8 Apodized Square Sieve with Varying Incident Wavelength.

Square, Sonine apodized photon sieves with $\nu = 4$ were simulated with varying incident wavelengths. The sieve was designed for a wavelength of 532.1 nm using the equations in Section 1.4(p. 7). The wavelengths tested were 522.1 nm, 531.1 nm, 532.1 nm (the design wavelength), 533.1 nm, and 542.1 nm. This is the design wavelength ±1 nm and ±10 nm. At $\lambda_{design} \pm 1$ nm the simulated sieve performed well, as shown in Fig. 29(a). Surprisingly, irradiance in the central lobe falls off most rapidly from the peak with $\lambda = 533.1$ nm, and least rapidly with $\lambda = 531.1$ nm, though the difference is minimal. The deepest null occurs with $\lambda = 531.1$ nm, though this is likely due to the low resolution of the simulation and not a true physical result. The simulation was repeated at the highest possible resolution (which is displayed in Fig. 29(a)) to see whether this prediction could be verified, but the changes in null depths were small. At $\lambda_{design} \pm 10$ nm the results were poor (unsurprising given the high dispersion produced by photon sieves) and significant excess energy can be observed in the side lobes in Fig. 29(b). At 10 nm from the design wavelength the innermost side lobes have so much excess energy that they have effectively merged with the central lobe, drastically widening it and decreasing imaging performance.



(a) $\lambda_{\text{design}} \pm 1$ nm (b) $\lambda_{\text{design}} \pm 10$ nm

Figure 29. Normalized, $\log_{10}$ diagonal cross sections through PSFs produced by a Sonine apodized ($\nu = 4$) square sieve with different illuminating wavelengths. The simulated sieve performed well with $\lambda_{\text{design}} \pm 1$ nm (a), but performed poorly with $\lambda_{\text{design}} \pm 10$ nm (b).

2.5 Noteworthy Results of the Simulation

As with the conventional optics simulations, it appears that square photon sieves improve the PSF along the diagonal relative to circular photon sieves. In Fig. 30 it can be seen that the PSF produced by the square sieve decreases rapidly along the diagonal direction, unlike the PSF produced by the circular sieve which appears similar to the Airy pattern.

The application of apodization to the square sieve also appears to improve the PSF, though not nearly as drastically as it did in Nisenson and Papaliolios's conventional optics simulations [34]. This could be because photon sieves are not well-suited for high-contrast imaging, or it may be because of the uncertainties with the numerical Fresnel approximation discussed in Section 2.2. Alternatively, it could also be attributed to some overlooked process or assumption used to generate Nisenson and Papaliolios's results that makes them inappropriate for comparison with these results. In Fig. 31, the PSF image produced by the apodized square sieve appears significantly cleaner outside the central region. The PSF has been smoothed, which would allow

a planet to stand out more easily, though unfortunately the data floor has not been suppressed. Additionally, less energy is present in the peak of the apodized PSF, which is logical considering that a portion of the incident light was blocked by the apodization. This is not displayed in the figure since each plot is normalized to its own peak, as the inherent contrast of the system is the relevant figure of merit.

Of the different apodizations investigated, the Sonine apodization and the square pyramid apodization (with high flat top percentage) seemed the most promising, and both of these were selected for experimental study. Due to the uncertainty of the simulation's accuracy and the relatively small differences in the simulated PSF cross sections the Gaussian and cosine apodizations were also selected for further investigation.

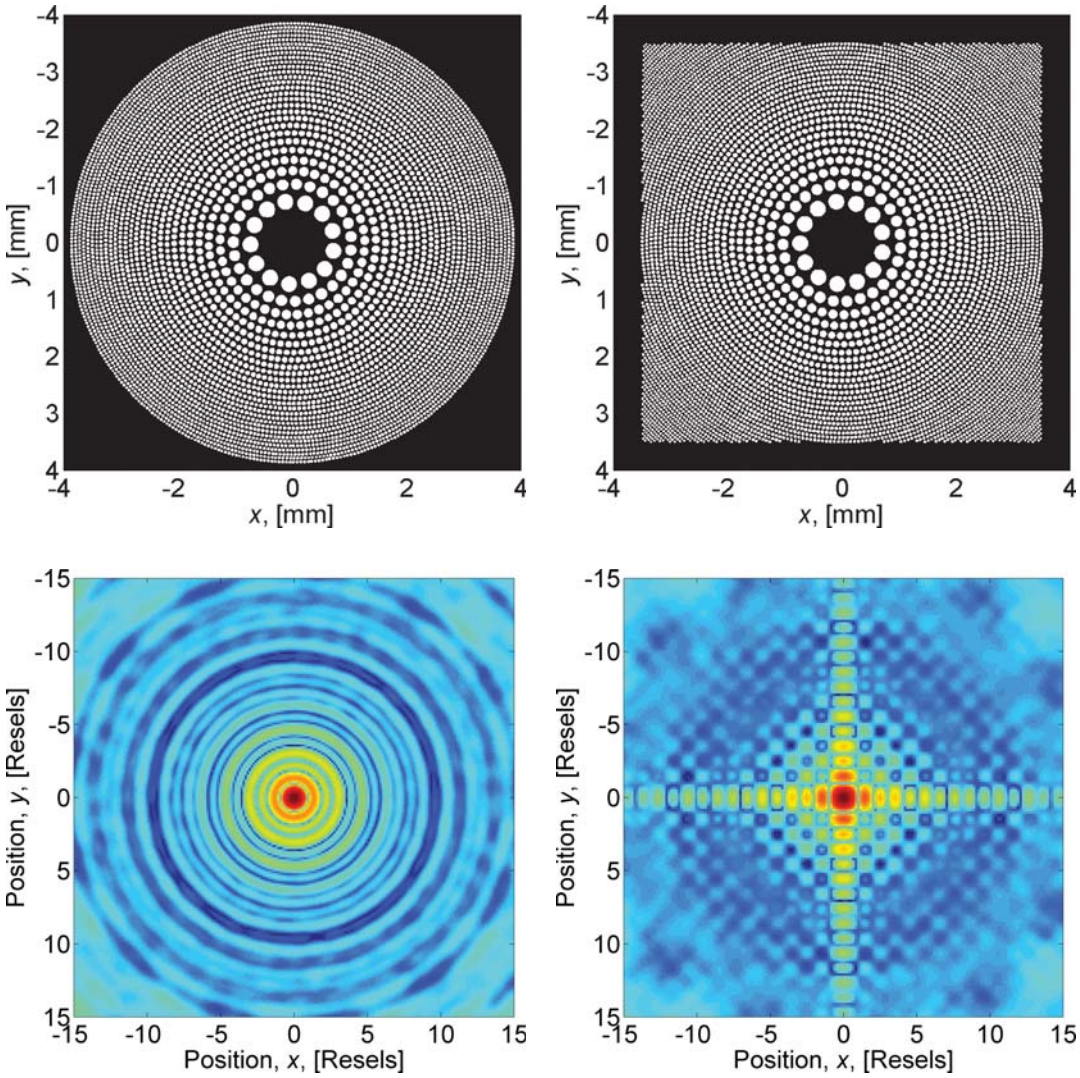


Figure 30. Comparison of the PSF images simulated using unapodized circular (lower left) and unapodized square (lower right) photon sieves. The sieves themselves are displayed above their respective PSFs. The surface area (and total hole area) of the two photon sieves is equal, and they are displayed above at the same scale.

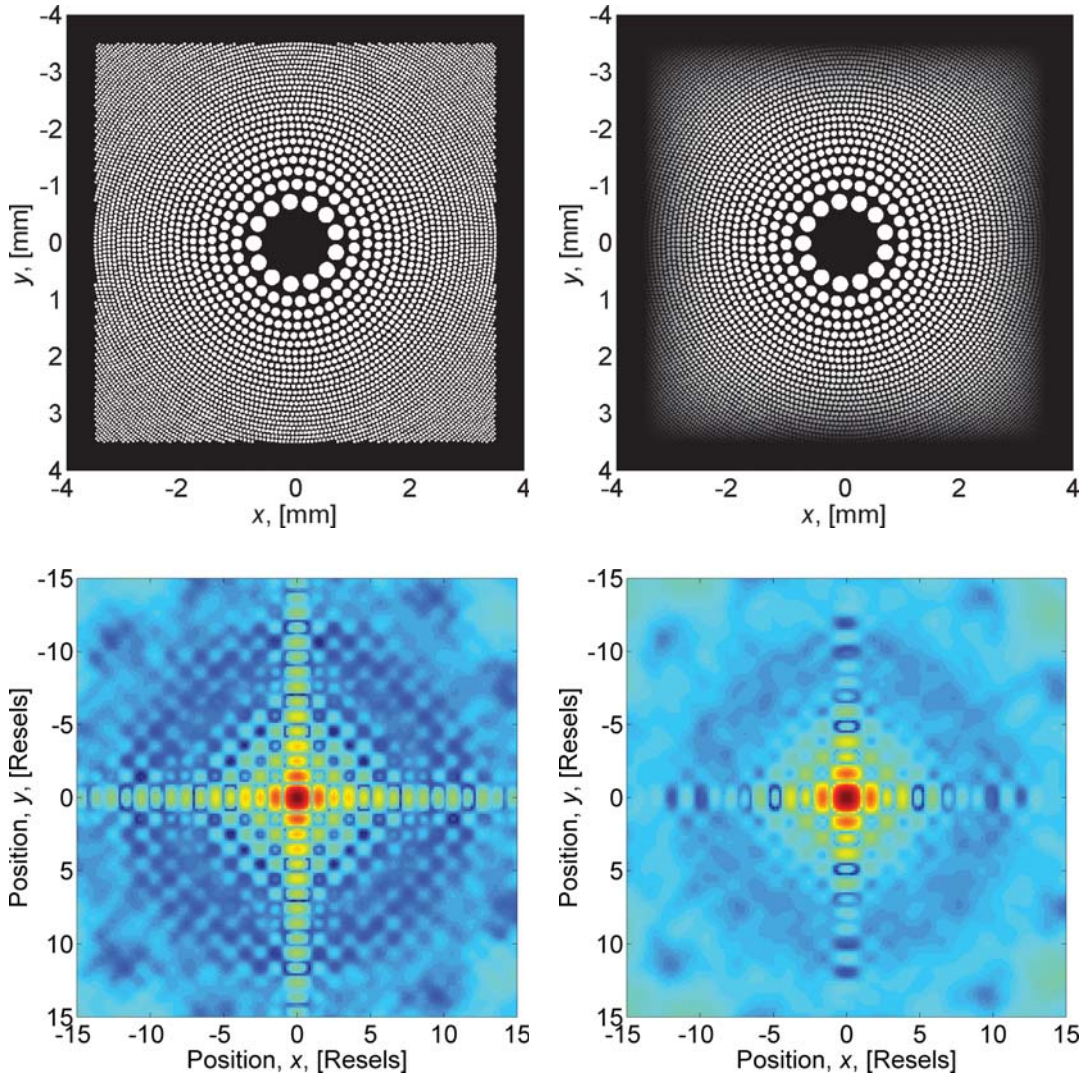


Figure 31. Comparison of the PSF images simulated using unapodized square (lower left) and crossed Sonine apodized square with $\nu = 4$ (lower right) photon sieves. The sieves themselves with apodization shown are displayed above their respective PSFs. The photon sieves are identical except for the applied apodization on the upper right.

III. Experimental Results

Photon sieves were designed and fabricated, and equipment was characterized to ensure it was appropriate for the experiment. A series of experimental tests with the photon sieves was conducted to provide concrete data allowing conclusions to be drawn regarding performance in a high-contrast telescope. Additionally, experimental results were compared to the simulated results to test the validity of the simulations for future work.

3.1 Equipment

The most important pieces of equipment in this experiment were the photon sieves themselves, the Verdi laser used as illumination, and the SBIG charge-coupled device (CCD) detector used to record PSFs and test the photon sieves' performance, and achromatic doublets used as collimating lenses and a control. These are discussed in greater detail below.

Other pieces of equipment required for this experiment (such as microscope objective lenses, mirrors, neutral density (ND) filters, and motion stages) were more standard and interchangeable, and therefore do not require extended discussion. AFIT faculty and lab technicians as well as the Air Force Academy's Laser and Optics Research Center (LORC) provided invaluable advice and oversight on the selection and purchase of dozens of pieces of equipment required for this experiment. All of the full page figures in this chapter were generated using a plotting routine adapted from code written by Martineau [30].

3.1.1 Photon Sieve Design and Fabrication.

Educated guesses about promising sieve designs were made using knowledge gained from the wave optics simulations. Ten photon sieves were designed in MATLAB after which the size and position data for the pinholes was converted to a Gerber (.gbr) file for use in fabrication [31]. The code used to design the sieves was adapted from codes written by Andersen and Tulip [3,43]. Two square chrome-coated quartz plates roughly 150 mm on a side and 2 mm thick were fabricated by Advance Reproductions using electron beam lithography (EBL), each containing four photon sieves. A third, smaller plate contained the final two photon sieves. It was required that holes be 1.5 μm or larger in diameter and hole edges be separated by at least 1.5 μm. According to the manufacturer, surface errors were less than 5 μm and manufacturing tolerance was ±0.25 μm. Having multiple sieves on each plate allowed different sieves to be placed into the optical path simply by rotating or switching plates. This provided the simplest and most consistent setup to thoroughly test each sieve.

The size of the individual photon sieves was limited by the size of the collimating lenses — 2″ or 5.08 cm in diameter. This means that for a square sieve to be fully illuminated by the expanded beam, the sieves had to be 5.08 cm or less on the diagonal, or 3.59 cm on a side. The radial distance to the outermost holes, r_n for $n = n_{max}$, was 2.54 cm. Using Eq. 6 with $f = 0.5$ m, $\lambda = 532.1$ nm, $r = 2.54$ cm, and $d/w = 1.53$ it was determined that the minimum hole width required for the square sieves was 8.02 μm, well within the manufacturer specifications.

The fill factor was the ratio of the total area of the pinholes in a given Fresnel zone to the area of that Fresnel zone. Even though the holes were separated by at least 1.5 μm, it was possible to achieve a fill factor greater than one since $d/w = 1.53$. As a result of the method used to meet the meet the manufacturer-required pinhole spacing the initial sieve-generating code did not result in a constant fill factor, but

rather one that decreased with increasing Fresnel zone radius, as shown in Fig. 32. This could be interpreted as an inherent apodization, and it was not ideal for the purpose of this investigation.

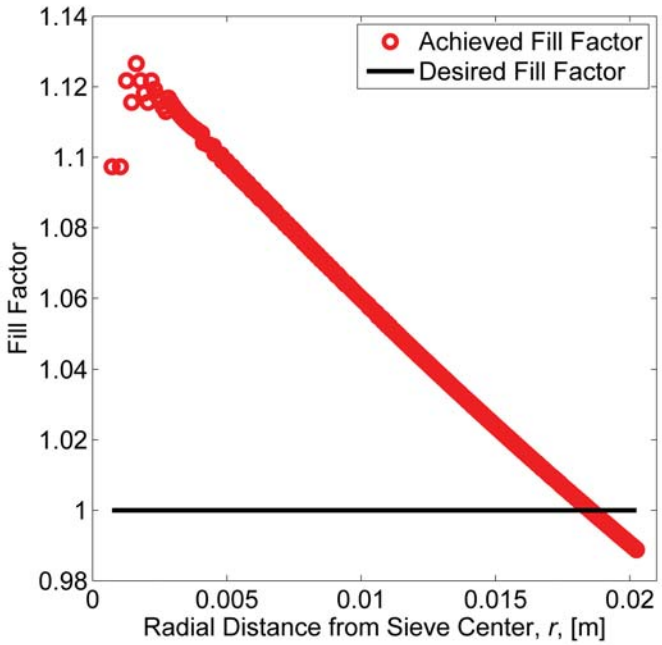


Figure 32. The fill factor is shown as a function of Fresnel zone radius for an uncontrolled unapodized photon sieve. The fill factor is shown to decrease roughly linearly with increasing Fresnel zone radius — a behavior that was undesired for this experiment.

It was decided that for the unapodized photon sieves, the fill factor would be forced as close to one as possible for each zone in order to provide better controls for the experiment. The area of each zone was calculated using Eq. 6 and Eq. 1. This was divided by the area of a pinhole in that zone to determine the number of pinholes giving a fill factor as close to one as possible (while meeting manufacturing spacing requirements). This quantity was then rounded down to the nearest whole number, since there was a discrete number of pinholes in each zone. Therefore the achieved fill factor was an approximation, and the approximation improved as the number of holes per zone increased (with increasing zone radius). The specific code showing the

application of the approximation is given in Appendix A (p. 132) lines 44-50, with the rounding taking place in line 50. This process gave results much nearer to the desired relationship between Fresnel zone radius and fill factor, as shown in Fig. 33.

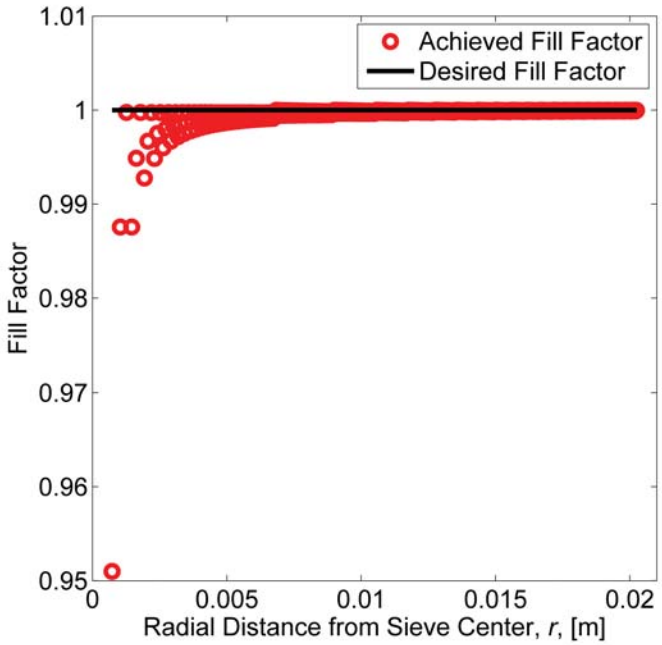


Figure 33. The fill factor is shown as a function of Fresnel zone radius for a controlled unapodized photon sieve. The fill factor remained within a percent of one, except for the innermost zones where there were so few holes that the approximation was poor. This method was used to generate the square and circular photon sieves, which also provided the baselines for the apodized sieves.

There are a number of different ways to apply an apodization to a photon sieve. For circularly symmetric apodizations such as the Gaussian apodization, the concentration of holes was forced to be higher in the inner Fresnel zones than in the outer Fresnel zones by varying the angular separation between the holes. This worked well and resulted in only small differences between the achieved and desired fill factors, as shown in Fig. 34. Holes were assigned a random angular position within their zones in order to prevent unwanted diffraction effects (since, for example, a line of holes leading radially outward would produce a diffraction spike), and manufacturer

spacing requirements were met. The fill factors for the circularly symmetric sieves are shown in Fig. 34.

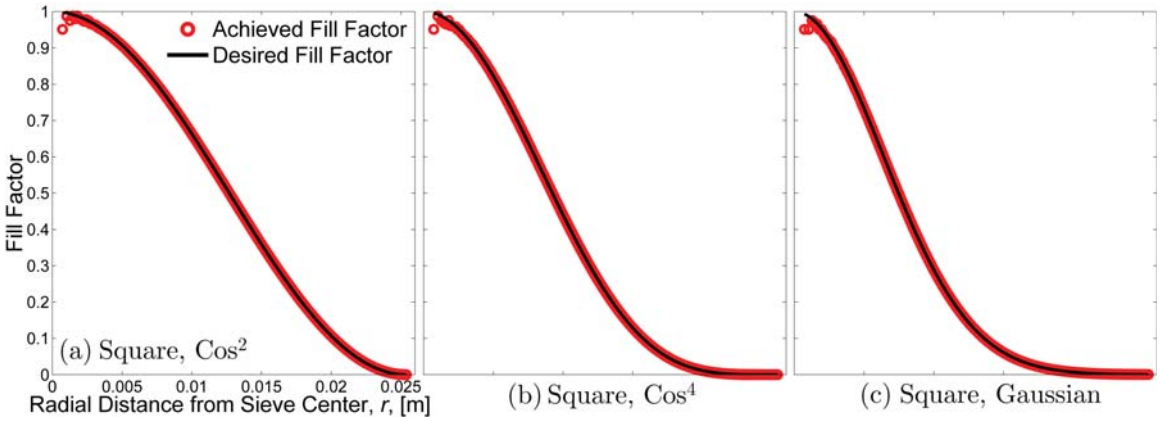


Figure 34. The fill factor is shown as a function of Fresnel zone radius for the sieves apodized with circularly symmetric functions. The fill factor remained within a percent of the desired function, except for the innermost zones where there were so few holes that the approximation (shown in Appendix A(p. 132) lines 44-50) was poor.

Non-circularly symmetric apodizations (such as the crossed Sonine) were more difficult to apply. The chosen method was to generate an unapodized photon sieve and then section it into a fine grid (250×250 elements) [20]. Each grid square contained roughly 100 pinholes to allow reasonable approximation. The value of the apodization function was computed at the center of each grid square, and the appropriate number of pinholes was removed from each grid square such that the fill factor would match the value of the apodization function. Removed holes within each grid square were selected randomly to prevent unwanted diffraction effects. A graphic of the method is shown in Fig. 35.

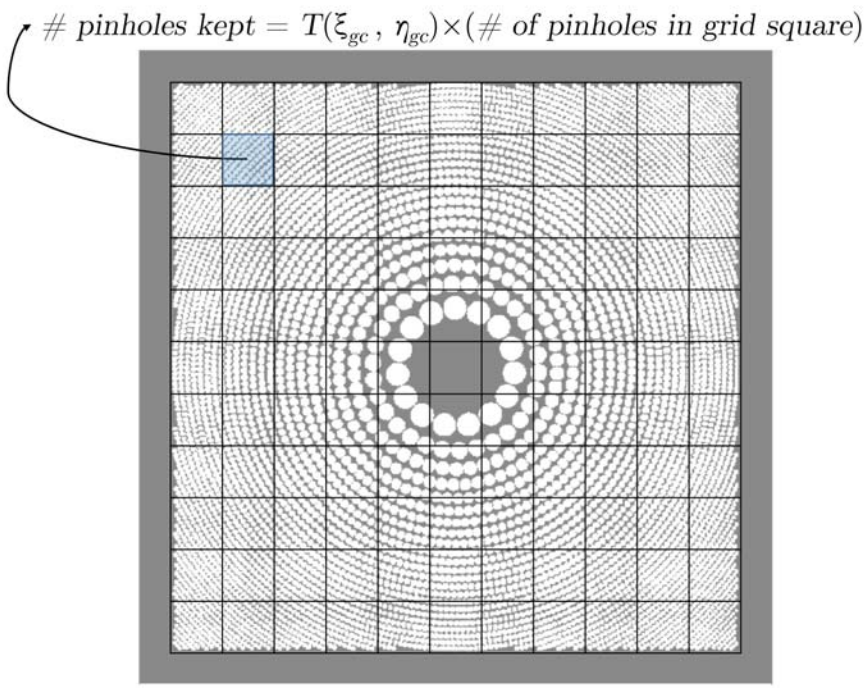


Figure 35. To apply the non-circularly symmetric apodizations each photon sieve was divided into a 250×250 element grid (shown simplified above as an 11×11 grid). The value of the apodizing function at the center of each grid square, $T(\xi_{gc}, \eta_{gc})$, was multiplied by the number of pinholes with centers within the grid square to calculate the number of pinholes in each grid square to keep. This number was reached by removing pinholes randomly from within each grid square.

The non-circularly symmetric apodizing method (given in Appendix B, lines 105-142) was implemented and optimized in MATLAB with assistance from Owens [35]. The selected grid resolution of 250×250 was adequate, as shown in Fig. 36, and the differences between the desired and achieved apodizations were minor. The residual between desired and achieved apodization for the square Sonine ($\nu = 5$) sieve is plotted in Fig. 37 to better display the small discrepancies.

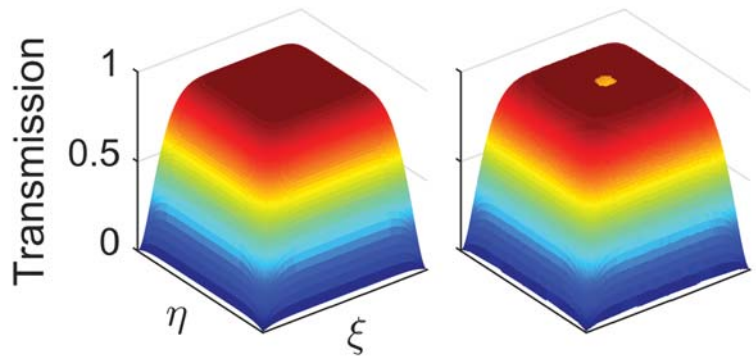


Figure 36. The apodizing function (left) and achieved apodization (right) for the square Sonine ($\nu = 5$) sieve. The difference is minor, which shows that this apodizing method was successful with a grid resolution of 250×250. The circular hole on top of the achieved apodization surface plots was expected, and was due to the fact that there are no pinholes in the center of a photon sieve. The differences between desired and achieved apodization were similarly minute for the other non-circularly symmetric apodizations.

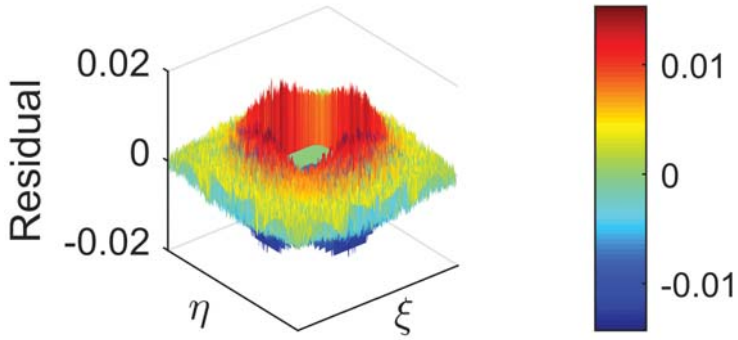


Figure 37. The residual between the apodizing function and the achieved apodization for the square Sonine ($\nu = 5$) sieve. The maximum residual was approximately 0.015 in magnitude.

An alternative method for applying non-circularly symmetric apodizations would have been to selectively decrease the hole diameters, which would have resulted in a complicated relationship between apodization and phase that was outside the scope of this thesis. However, with outermost holes sized at 8.01 μm and a minimum hole size of 1.5 μm, the area of the outermost holes could have potentially been reduced from 201.6 μm^2 to 7.1 μm^2, with roughly 3.5% of the original throughput.

After generating the Gerber files using MATLAB, they were reviewed for accuracy using GC-Prevue, which has the ability to generate a binary image from a Gerber file as shown in Fig. 38. The files were imported to the viewer using the RS-274X format. The RS-274X format only supports 1,000 different aperture sizes while the sieves contained 1,200+ zones. This made it necessary to divide each photon sieve into two separate Gerber files, with the first containing coordinates and diameters of pinholes in the inner zones and the second containing coordinates and diameters of pinholes in the outer zones. The outermost zone included in the first file was zone 989. The manufacturer was then informed to combine these two files into a single photon sieve during the manufacturing process.

Full specifications for the final photon sieve designs are listed in Table 1. The "% Throughput" column gives the ratio of the total pinhole area of each apodized sieve to its unapodized counterpart, meaning the Sonine apodized circular sieve is compared to the unapodized circular sieve, while the apodized square sieves are compared to the unapodized square sieve.

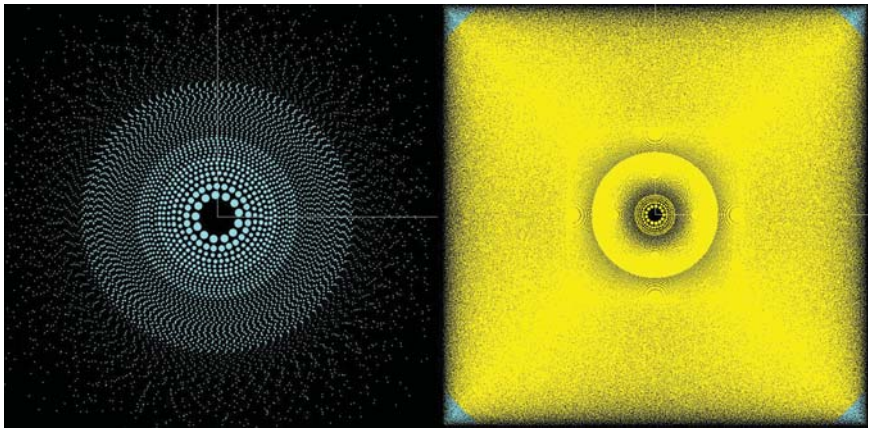


Figure 38. Example of a square, Gaussian apodized photon sieve (left) and square, pyramid (16% flat top) sieve (right) are shown as rendered by GC-Prevue. The pinholes are colored, while the black background is non-transmissive. The abrupt discontinuities in hole size are rendering artifacts and are not actually present. The pinholes in the corners of the pyramid sieve are a different color because they were loaded from a separate file due to RS-274X format limitations.

Table 1. Photon Sieve Design Specifications

Photon Sieve Type	Sieve Width [cm]	f/#	Total Zones	Total Pinholes	Minimum Hole Diameter, d_{min}, [μm]	Total Hole Area [mm^2]	% Area Covered	% Throughput (Apodization)	Code Runtime [min]
Circular	4.05	12.34	770	4,058,448	10.05	643	49.9	-	2.8
Circular, Sonine Apod. ($\nu = 5$)	4.05	12.34	770	3,320,053	10.05	573	44.5	89.1	164.3
Square	3.59	9.84	1,210	4,261,649	8.02	644	49.9	-	3.7
Square, Pyramid Apod. ($2/3^{rds}$ flattop)	3.59	9.84	1,170	2,852,928	8.15	524	40.6	81.3	161.0
Square, Pyramid Apod. ($1/3^{rd}$ flattop)	3.59	9.84	1,168	1,806,026	8.15	401	31.1	62.3	137.3
Square, Sonine Apod. ($\nu = 5$)	3.59	9.84	1,192	2,759,020	8.05	510	39.5	79.1	157.8
Square, Sonine Apod. ($\nu = 4$)	3.59	9.84	1,184	2,438,890	8.09	474	36.7	73.6	150.1
Square, Cos^2 Apod.	3.59	9.84	1,154	1,175,225	8.13	284	22	44.1	1.0
Square, Cos^4 Apod.	3.59	9.84	992	496,441	8.57	172	13.3	26.7	0.5
Square, Gaussian Apod. ($\sigma = r_{max}/4$)	3.59	9.84	1,089	296,924	8.04	125	9.6	19.3	0.3

3.1.2 Verdi Laser.

A frequency-doubled, diode-pumped Nd:YVO_4 laser ($\lambda = 532$ nm) was used as the light source for the majority of testing. The Verdi V-18, manufactured by Coherent Inc., operates from 10 mW to 18 W and was used between 10 mW and 1 W for this experiment. It was fitted with a closed-loop water cooling system and according to the manufacturer had a power stability of $\pm 1\%$. According to manufacturer specifications, it produced a diffraction-limited, vertically polarized beam with a beam diameter of 2.25$\pm$0.2 mm measured between the $1/e^2$ points at the exit port of the laser. The full-angle beam divergence was less than 0.5 mrad and the M^2 value was less than 1.1. The bandwidth of the beam was <5 MHz (<4.7 fm linewidth), which was ideal as it was far less than the focused image bandwidth of a photon sieve [2, p. 3708]. Not all of these specifications held true for the low power at which the laser was operated. The manufacturer recommended operating the laser at 14 W for the best performance, but this was deemed impractical due to the amount of power which would have to be absorbed or reflected to reduce the beam to the desired power. This power reduction might have been achieved using the back reflections from multiple uncoated, flat optics with graphite beam blocks in the path of the transmitted light, but the amount of stray light this setup would produce would have made much of the experimental work unmanageable. Because it was necessary to operate this laser far from its ideal power for this experiment, it is suggested that future work be conducted with a laser designed for low power operation (the V-18 is designed primarily for Ti:Sapphire pumping). In this case, no more suitable lasers were available. The laser was suitably spatially filtered and many of the experimental results aligned with predictions, so despite the manufacturer recommendation there is no reason to believe this invalidates the experiment.

Knife-Edge Measurement. A 90% - 10% knife-edge measurement was conducted to verify the beam width and determine the proper microscope objective and spatial filter pinhole combination for a beam expander with this laser. The setup for the measurement is shown in Fig. 39. The measurement was conducted with the laser power set to 10 mW since that was the setting used for the majority of the experiment, though the Silicon photodiode power meter read 13.3 mW for the unobstructed power measurement. A razor blade acted as the knife edge, and was translated through the beam using the micrometer until the remaining power incident on the detector dropped to 90% and then 10% of the unobstructed power. The micrometer reading was recorded at these two positions, and the distance between them was found to be 1.8 mm. The overwhelming source of uncertainty was the power meter reading, which fluctuated over a range of 0.7 mW (equating to ±23% uncertainty in the power for the 10% measurement). This was likely due to the limitations of the power meter, which provided unstable low power results for sources other than the laser. Assuming the beam had a Gaussian profile, the 90-10 result of 1.8 mm was multiplied by a factor of 1.56 to give a $1/e^2$ beam diameter of 2.8 mm [38, p. 1103]. Due to high uncertainty of the power meter readings and the fact that the manufacturer specification was within 20% of this value, the manufacturer specified $1/e^2$ beam diameter of 2.25 mm was used in subsequent calculations, though using the knife-edge result would not have changed the spatial filter pinholes selected.

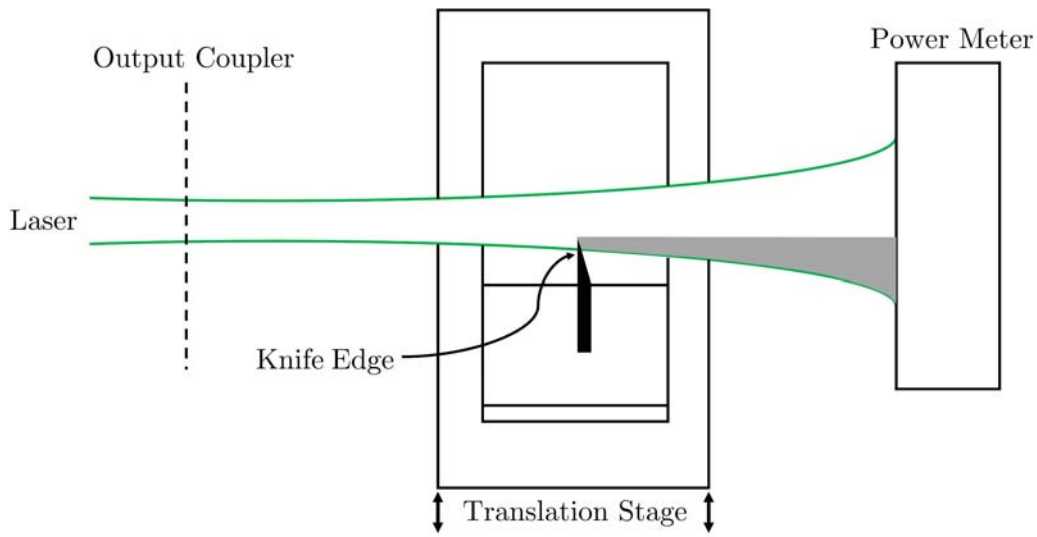


Figure 39. The experimental setup for the 90-10 knife-edge measurement. A razor blade was translated to block the beam using a micrometer, while the unblocked portion of the beam was incident on a power meter. This figure was adapted from course work done with Owens.

Spatial Filtering. In order to produce a clean, expanded beam, a microscope objective focused the laser through a spatial filter pinhole which was used to clip any higher spatial frequencies present in the beam. The simplest equation for determining the optimal spatial filter pinhole diameter is given by the Newport Corporation website as

$$D_{PINHOLE} = \frac{f\lambda}{w}, \tag{15}$$

where f is the effective focal length of the microscope objective lens and w is the $1/e^2$ beam radius (1.125 mm) [33]. According to the website, the equation was derived based on filtering out spatial wavelengths smaller than the diameter of the incident beam, and it results in a 99.3% power throughput. For 10× ($f = 15.4$ mm) and 20× ($f = 8.0$ mm) microscope objectives, Eq. 15 gave optimal pinhole diameters of 7.3 µm and 3.8 µm, respectively. A slightly different equation was given by the Edmund

Optics Inc. website with

$$D_{PINHOLE} = 0.9525 \times \frac{f\lambda}{w}, \tag{16}$$

which results in optimal pinhole diameters of 6.9 μm and 3.6 μm, respectively [18]. Pinholes are sold in 5 μm diameter increments, and the Newport Corporation website invariably recommended choosing the first available size greater than was calculated with Eq. 15. This is because though Eqs. 15 and 16 are optimal, they also provide somewhat of a minimum limit. With a pinhole any smaller, a significant portion of the low spatial frequency ideal spot will be clipped in the microscope objective's transform plane. The noise usually has much higher spatial frequencies than any contained in the ideal spot, so even a larger pinhole will properly filter it in the Fourier transform plane (for example, the Newport Corporation website recommends a 10 μm pinhole rather than a 5 μm even when the theoretical optical diameter is as small as 5.7 μm). As a result, a 10 μm pinhole was used with the 10× objective and a 5 μm pinhole was used with the 20× objective. Repeating the calculations with the knife-edge determined beam diameter of 2.8 mm resulted in the same selections. Spatial filter pinholes are easily damaged by a focused laser, and a damaged pinhole can lead to poor beam quality. Therefore, even though much of the experimental work was conducted with the laser power set to only 10 mW, high power pinholes were selected. Using these two microscope objective and pinhole combinations produced diverging beams with a clean central spot surrounded by a faint concentric ring. Ideally, the ring would not be present but it persisted no matter how much fine adjustment was made to the position of the spatial filter and the separation between it and the microscope objective. It is possible that due to the laser's low operating power, there was a slight presence of higher order modes with spatial frequencies low enough (such as the TEM_{01*} or TEM_{10} modes) that they were not fully clipped by the

spatial filter. This may have happened because the selected pinholes were larger than the theoretically ideal pinhole, as recommended [33]. A dozen different combinations of microscope objective and pinhole diameter were tested but none performed better than the pairs above. The long focal length (400 mm) of the collimating lenses meant that the filtered beam diverged considerably before reaching the collimating lens, and only the central third (by diameter) of the central spot was required to cleanly and evenly illuminate the collimating lens. Ultimately, the presence of the concentric ring had no apparent impact on the experimental work.

3.1.3 Detector Characterization.

The same imaging detector was used for the duration of the experiment, and is here referred to interchangeably as the detector, CCD, sensor, or camera. More precisely, the term "camera" refers to the full apparatus including CCD sensor, shutter, electronic cooling, and driving electronics, while the terms "sensor" and "CCD" refer to the doped silicon chip (comprised of millions of pixels) that converts photons to photoelectrons.

The detector was an SBIG STF-8300M camera fitted with a Kodak KAF-8300CE CCD. It possesses thermoelectric cooling (TEC) with a maximum temperature change of -40 ℃ to reduce dark current and thus noise. The dark current was very low, at 0.002 electrons per pixel per second at -10 ℃. The sensor has a resolution of 3352×2532 pixels and dimensions 17.96×13.52 mm^2. The manufacturer lists the pixel size as 5.4×5.4 μm^2, and did not respond to a request for an additional significant figure. The results presented for this experiment used a pixel size of 5.36 μm on a side, calculated by dividing the sensor width by the horizontal resolution. This figure was 0.7% smaller than the manufacturer specification, and this modest uncertainty had an insignificant impact on the results presented. The A/D converter was 16 bits

and the full-well capacity was roughly 25,500 photoelectrons. The camera possessed antiblooming technology (designed for up to 1000× the full-well-capacity), though some blooming was still observed as shown in Fig. 42(p. 61). Since it was designed for astrophotography the possible integration times ranged from 0.09 to 3600 s. Read noise was roughly 9.3 electrons. All of the specifications given for the camera were provided by the manufacturer.

A 3×3 binning mode allowed 0.1-s exposures to be downloaded in roughly 0.25 s, which meant the focus of an optical system could be adjusted and observed in near real time. For all experimental work the camera was operated with TEC set to 0 ℃ or -10 ℃, depending on the ambient temperature in the laboratory. With these settings the TEC was operating at 50-80% of its full capacity, which enabled it to keep the sensor's temperature stable at ±0.2 ℃ of the desired setting.

Detector Linearity Characterization. A linear detector has a linear relationship between the total energy incident on the detector and output pixel count, as shown in Fig. 40. Detector linearity was essential for reliably scaling data, as discussed in detail in Section 3.1.4. Because of this, it was important to verify detector linearity prior to gathering experimental data.

The TEC was set to 0 ℃ and all sources of stray light in the lab were removed or blocked as best as possible. After passing through ND filters to reduce the power, the laser was used with the 10× microscope objective and the 10-μm pinhole to produce a spatially filtered, diverging beam. The central portion of the beam was collimated with an achromat ($f = 400$ mm) and used to fully and evenly illuminate the detector. Exposures were taken with twenty-one different integration times, ranging from 0.1 to 55 s. For each integration time the shutter was closed and three dark frames with the same integration time were taken. Each set of three dark frames was averaged to produce a single dark frame image to subtract from each light image and

reduce the impact of noise on the measurement. After subtraction, the average pixel count, referred to from now on as analog-to-digital units (ADU), was found for each integration time. These ADU values are shown plotted as a function of integration time in Fig. 40.

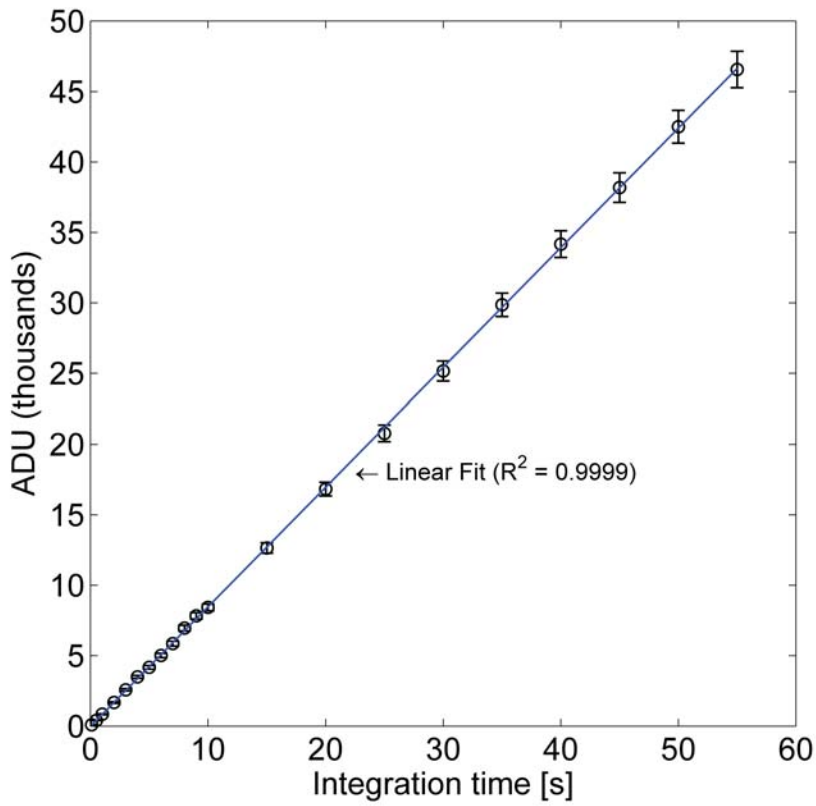


Figure 40. The results of a CCD linearity test showed that the detector is well-suited for this experiment. Mean (over the full image) pixel count (ADU) is plotted as a function of integration time with constant illumination as the black circles. The uncertainty for each data point (the standard deviation over the full image after dark frame subtraction) is also shown in black. The linear least squares fit is shown in blue.

A linear least squares fit was performed in MATLAB which resulted in $R^2 = 0.9999$ and the fitted line falling within the uncertainty bounds for all data points. This shows that the detector has excellent linearity for the integration time and ADU ranges investigated.

Detector Blooming Characterization. For this geometry,

$$D_{Airy} = 1.22 \times \frac{\lambda z}{2w} = 5.2 \text{ mm}, \tag{17}$$

gives the width of the central lobe of the Airy pattern, where z is the distance from the source plane to the observation plane and w is the half-width of the aperture [22, p. 78]. An image of an approximate Airy pattern was taken with the setup and procedures given in Section 3.1.4(p. 62). A horizontal cross section through HDR, scaled-and-spliced date is shown plotted with a Bessel function of the first kind of order one (the cross section through the central peak of the Airy pattern) overlaid in black in Fig. 41. This figure also displays the detector's excellent linearity, given that the linearly scaled data (in green) matches the unscaled data (in blue) almost perfectly. The Bessel function was defined such that the width of the central lobe of the Airy pattern matches the predicted width given by Eq. 17.

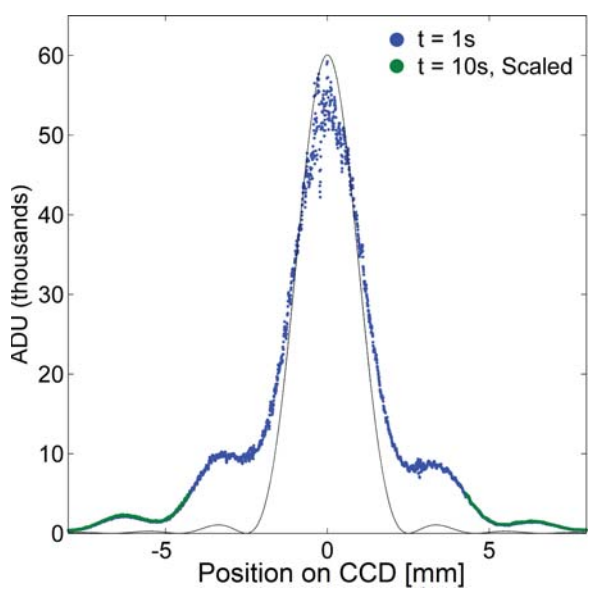


Figure 41. The results of a CCD linearity test with a Bessel function (black line) overlaid on top of the cross section through the peak of an Airy pattern image (green and blue points).

The mismatch between the data and the Airy pattern suggested that blooming

was an unavoidable issue for this detector. However, it was possible that the mismatch was due to an error in the experimental setup resulting in the diffraction pattern not being an Airy pattern, so an additional test was necessary.

A photon sieve was used to image a resolution target with the setup and procedure described in Section 3.2.1. Since the patterns of the resolution target have a known rectangular shape, blooming is easily observed. With the incident laser power kept constant, images (shown in Fig. 42) were taken with integration times doubling from 0.1 s to 3.2 s. This showed definitively that blooming was an unavoidable issue with this detector. It also showed that blooming was most severe along the CCD's horizontal axis. Much of the experimental work involved vertical and diagonal cross sections, which were apparently less susceptible to blooming. Since blooming redistributed ADU from the highest peaks to the deepest nulls, the experimental work conducted with this sensor gave a minimum performance threshold for these optics — the actual PSFs may be more favorable for exoplanet imaging than shown in the measurements. Though this was an unavoidable source of uncertainty, it was also useful in that it provided more practical data on the problem — any actual telescope will be imaging onto a detector that becomes susceptible to blooming at some point.

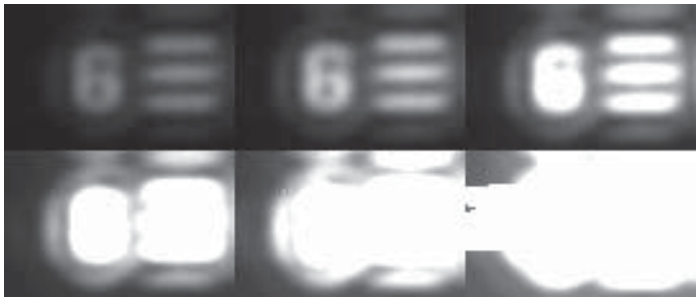


Figure 42. Images of a resolution target taken with a photon sieve with integration times doubling between 0.1 s and 3.2 s. Integration times increase from left to right then top to bottom. The 3.2 s image displays extreme blooming, especially along the CCD's horizontal axis.

Despite the fact that blooming was observed with this detector, the high linearity

and low noise meant it was an appropriate detector for this investigation.

3.1.4 HDR Imaging: The Scale-and-Splice Method.

For the purposes of this investigation, the dynamic range of a detector was the ratio between the largest and smallest possible signals — in this case 65,536 (this equals 2^{16}, where 16 is the detector's bit-depth). This was a simplification since the full-well-capacity of the CCD was roughly 25,000 photoelectrons (so data was never truly 65,536 discrete steps) and the smallest possible signal was increased significantly by noise (though noise was accounted for as best as possible). High dynamic range (HDR) data are data acquired using a process that increases the dynamic range beyond the typical value. In order to increase the dynamic range, use of a "scale-and-splice" method was necessary, and this method relied heavily on the detector's linearity. A linear CCD has a linear relationship between the energy incident on the detector and the pixel count that it outputs, as shown in Fig. 40.

A scale-and-splice method was developed and refined to extend the detector's dynamic range far beyond the typical value. The method was necessary to capture PSF detail at irradiances down to roughly 10^{-5} of the peak. The steps of the scale-and-splice method are listed, then described in more detail later in the section.

Steps of the Scale-and-Splice Method.

1. Acquire a set of short exposures of the PSF (short lights) with an integration time just short enough that the detector does not saturate. This is necessary to fully capture the peak of the PSF. Average these into a single image (short light master frame).

2. Acquire a set of long exposures of the PSF (long lights) with an integration time long enough to capture detail in the outer portions of the PSF. Average

these into a single image (long light master frame).

3. Block the light source and capture a set of calibration frames with the shorter integration time (short darks). Average these into a single image (short dark master frame).

4. With the light source still blocked capture a set of calibration frames with the longer integration time (long darks). Average these into a single image (long dark master frame).

5. Subtract the short dark master frame from the short light master frame to calibrate the data and create a calibrated short master frame.

6. Subtract the long dark master frame from the long light master frame to calibrate the data and create a calibrated long master frame.

7. *Scale* the calibrated long master frame linearly by dividing each pixel value by the ratio of the long integration time to the short integration time to create a scaled long frame.

8. Select the pixels of the scaled frame that have values less than 80% of the peak (saturated) value. These are the scaled, unsaturated pixels.

9. *Splice* the data by substituting the scaled, unsaturated pixels into their locations in the calibrated short master frame.

10. Divide all pixels by the peak pixel value to normalize the image and create normalized HDR ADU (scaled-and-spliced) data.

There were three main components of the scale-and-splice method: acquisition and calibration, scaling (dividing the long exposure data by the ratio of the long integration time to the short integration time), and splicing (replacing short exposure

pixels with scaled long exposure pixels). These three components are described in detail in the following paragraphs.

Steps 1-6: Acquisition and Calibration. The scale-and-splice method started by imaging a PSF with two different integration times (long and short). For each integration time, two different types of exposures were taken: light frames and dark frames (referred to as lights and darks). Lights were actual images of the PSF. Darks were calibration frames to be subtracted from the lights, and were captured after blocking the light source. This resulted in very little light being incident on the detector and produced a nearly black image, which is why they are known as darks. Dark subtraction accounted for the impact of detector noise and stray light. It is best to capture multiple light frames and dark frames at each integration time and average them to create a master light frame and master dark frame since this reduces the impact of vibration, detector temperature fluctuation, shot noise, and laser power fluctuation among other things. For each integration time the master dark frame was then subtracted from the master light frame to produce a calibrated image.

After the dark subtraction, some areas of the calibrated images had slightly negative values. To account for this discrepancy, the average minimum column pixel value was found, and the data was shifted up by the absolute value of this quantity. Had the data been simply shifted up by the absolute value of the minimum pixel value of the entire image, information would have been lost, since the minimum pixel values were due to the presence of a very small number of erratic pixels and were far from the average minimum column pixel value (approximately -2,000 versus -100). After the shift, any pixels with values ≤ 0 remaining were replaced with one, since that was the theoretical minimum possible ADU for the detector (in practice the minimum value is higher due to noise). With a ratio of 20 between long and short integration times, this process set the absolute floor of the data to roughly

10^{-6} of the peak, though as shown throughout the results (and best displayed in Fig. 54) the true floor was higher. The calibration code is provided in Appendix D(p. 153).

Step 7: Scaling. Because the integration time for the long exposure is much longer than it was for the short exposure, many more photons were incident on the detector, and the pixel counts were much higher. In order to account for this, the calibrated long master frame must be divided by a scaling factor: the ratio of the long exposure time to the short exposure time (e.g., dividing by ten when using a long integration time of 10 s and a short integration time of 1 s). This ratio was not always the best scaling constant, as displayed by some small discontinuities in the data in Fig. 55(p. 88). However, it minimized the discontinuities most consistently and was therefore selected for use.

Steps 8-10: Splicing. After producing the scaled frame, the pixels with values less than 80% of the peak (saturated) value were selected. This was done to preserve detector linearity (which is lost when the detector is saturated) and preserve the structure in the PSF's peak. These scaled, unsaturated pixels were then used to replace their corresponding pixels in the calibrated short master frame. The pixels were then divided by the peak pixel value to normalize the image. The result was normalized HDR ADU (scaled-and-spliced) data. The scale-and-splice code is provided in Appendix E(p. 156).

Testing the Scale-and-Splice Method. Section 3.1.3 shows that this detector was well-suited for this method due to its high linearity. To demonstrate its reliability, the scale-and-splice method was tested prior to its application to the experimental results discussed in Sections 3.2 and 3.3.

The same collimated beam as was used for the Section 3.1.3 linearity test was

incident on a 25-μm diameter pinhole. ND filters were used to reduce the power in the beam. Stray light in the laboratory was once again eliminated to the fullest extent possible. Diffracting the collimated light through a circular aperture produced an Airy pattern in the far field. Similar in nature to Eq. 12, but adapted for the far field, the Fraunhofer condition for this aperture is given by

$$z > \frac{2W^2}{\lambda} = 2.35 \text{ mm} \tag{18}$$

where z is the distance from the source plane to the observation plane and W is the width of the aperture [22, p. 75]. The distance from the source to the detector was selected as 100 mm to meet Eq. 18's condition and place the detector deep in the Fraunhofer region.

The scale-and-splice method was carried out following the steps of the list provided at the start of this section. Ten 1-s exposures and ten 10-s exposures were taken of the Airy pattern (light frames). The beam was blocked after the spatial filter pinhole using a graphite beam block (measuring roughly 10×10 cm^2) followed by a larger piece of black polyurethane foam (measuring roughly 30×20 cm^2). This, in combination with the measures taken to eliminate stray light in the laboratory, meant that almost no light was incident on the detector. Ten dark frames were then taken for each integration time. For each integration time, the light frames were averaged to produce two master frames of the Airy pattern, with the 10-s exposure being extremely overexposed in the center. The dark frames were also averaged for each integration time to produce two master dark frames. As described above, averaging multiple darks and lights served to reduce the impact of noise, vibration, and varying laser output on the experiment. The master dark frames were then subtracted from their corresponding master light frames to produce two calibrated images. This calibration method accounted for much of the remaining stray light as

well as the detector's read noise and Johnson noise. This calibration method was suggested by Dr. Geoff Andersen, and was used for the majority of the experimental work. After calibration, the data was scaled-and-spliced following the steps listed.

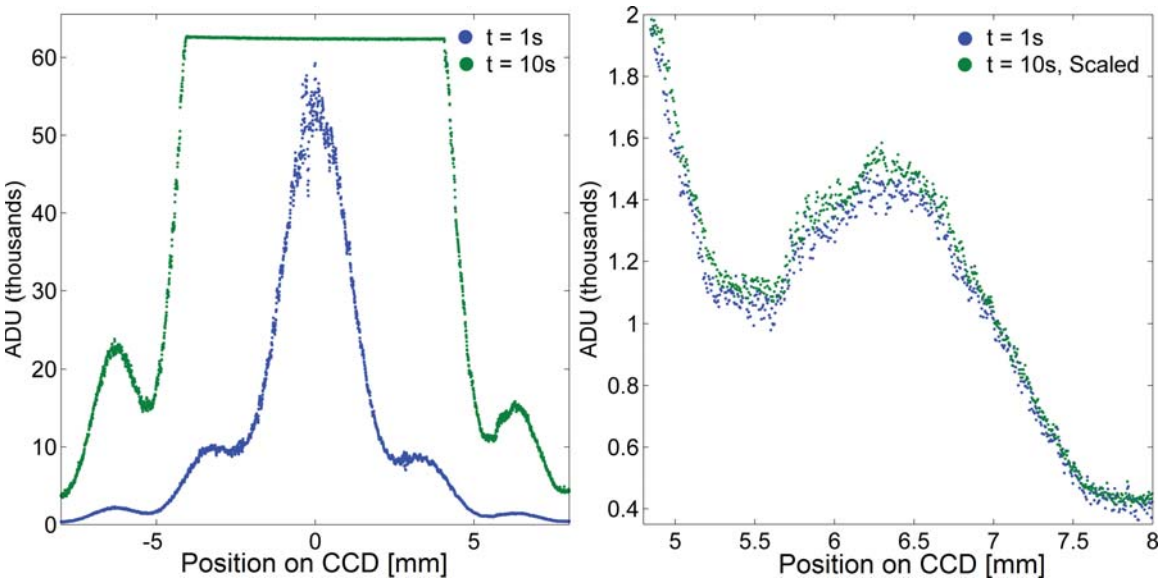


Figure 43. Horizontal cross sections though the peak of calibrated images of an Airy pattern with two different integration times (left) and a cropped view (right) showing the same data after scaling the 10-s data linearly by dividing the ADU by ten. Rather than splicing, the 1-s and 10-s data are shown overlaid for comparison. The green data points are the 10-s data, and the blue data points are the 1-s data.

Fig. 43 (left) shows the full horizontal cross section of the diffraction pattern (with long and short unscaled data overlaid), and Fig. 43 (right) shows a cropped view of the same data after dividing (scaling) the ADU for the 10-s data by ten. To preserve linearity, only 10-s data less than 80% saturated (ADU $<$ 52,429) was scaled. After calibration, the average standard deviation (between the values of corresponding pixels in ten frames) of the scaled-and-spliced 10-s data was 32 ADU, while the standard deviation of the corresponding data points (which would be replaced in the splice) was 69 ADU. The fact that the long exposure data had significantly less variation than the short exposure data illustrated the utility of the scale-and-splice method. The absolute value of the residual between the 1-s and the scaled 10-s data had an average of 68

ADU, only 0.1% of the full 65,536 ADU range of the detector. The standard deviation of the residual was 57 ADU. Therefore this HDR method was viable for PSF imaging.

3.1.5 Achromatic Doublet Lenses.

ThorLabs AC508-400-A model achromatic doublet lenses were used as collimating lenses for the duration of the experiment. The long focal length (f = 400 mm) of the collimating lenses helped to reduce spherical aberration since the high radius of curvature spherical surfaces of a longer focal length lens more closely approximated the ideal parabolic lens surface. As a result, the shear plate collimation tester showed relatively straight fringes as displayed in Fig. 44 (linear fringes indicate low aberration and horizontal fringes indicate a collimated beam).

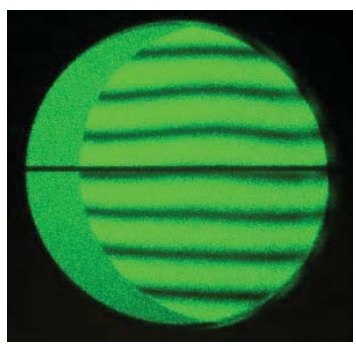


Figure 44. An image of a shear plate collimation tester showing relatively straight fringes produced by the setup used for much of the experimental work.

The large diameter allowed the photon sieves to be tested in the more appropriate infinite conjugate configuration (rather than $2f$ to $2f$), and the anti-reflection (AR) coating effective from 400-700 nm helped to reduce stray reflections. Unfortunately, due to the retaining rings that secured the lenses in their mounts, their effective diameter was reduced from 50.8 to 48.3 mm. As a result, a small portion of the square photon sieves (at the corners) was unilluminated during testing, as described in Section 3.2.1.

To reduce spherical aberration, it was important that the flatter surface (larger radius of curvature) of the lens faced nearby sources and the side with the more highly curved surface (smaller radius of curvature) faced distant sources [24, p. 255]. Since the spatial filter pinhole was only approximately 400 mm from the collimating lens, it was considered a nearby source and collimating lenses were always used with the flatter side of the lens facing the source.

A ThorLabs AC508-500-A achromatic doublet lens was also used as a control to compare with the photon sieves. Except for having $f = 500$ mm, it was identical to the collimating lenses. When focusing collimated light, this lens was oriented with the more highly curved surface facing the incident beam. Following the same exact procedure (described in Section 3.2.2 beginning on p. 82), images were taken of the PSF produced by the lens when it was oriented incorrectly, then correctly. The two images are shown as surface plots with a $\log_{10}$ scale in Fig. 45.

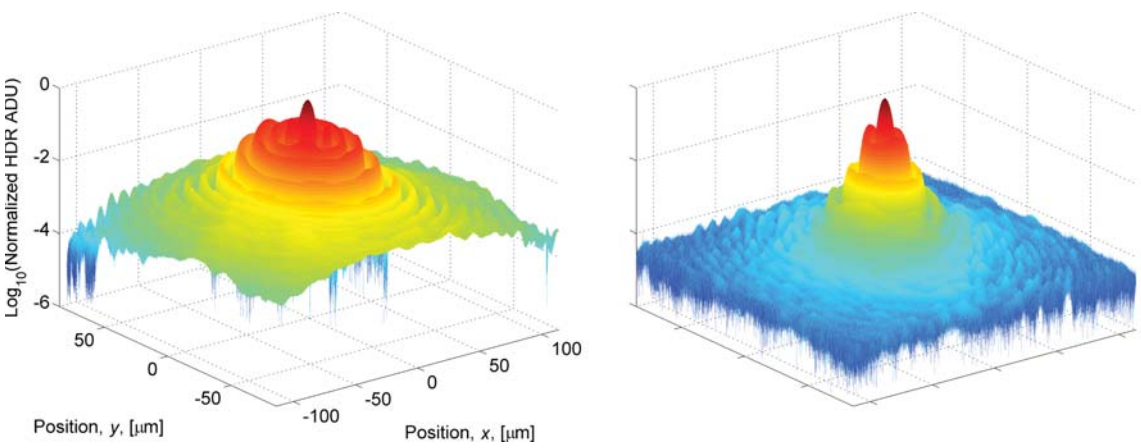


Figure 45. Surface plots of the PSFs produced by the control lens with the curved surface facing away from the collimated beam (left) and towards the collimated beam (right). The two surfaces are plotted with identical axes and colormaps. These PSFs were collected following the procedures given in Section 3.2.2(p. 82) and processed with the scale-and-splice method described in Section 3.1.4(p. 62).

The left plot shows a high degree of spherical aberration, "which essentially shifts light out of the central disk into the surrounding rings" [24, p. 254]. The right

plot, for which the lens was oriented correctly, shows minimal spherical aberration in comparison. To verify the experimental results, a PSF analysis of the lens was conducted with Zemax by Hawks [23]. The simulation assumed perfect alignment and monochromatic 532 nm, plane wave illumination. The results were a WFE of peak to valley = 1.27λ and rms = 0.37λ for the incorrect orientation, and a WFE of peak to valley = 0.07λ and rms = 0.02λ. The $\log_{10}$ scale images of the PSFs in Fig. 46 produced by the simulation display how much energy is shifted from the central spot to the outer rings, and can be seen to agree qualitatively with Fig. 45.

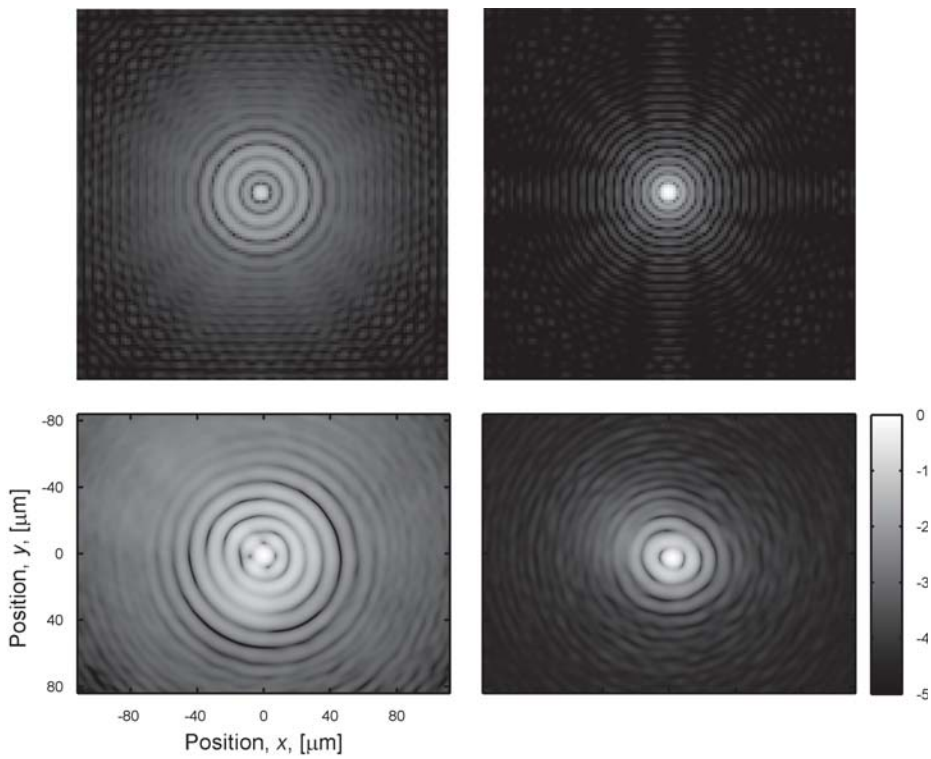


Figure 46. Log_{10} images of the PSFs produced by a Zemax analysis (top) and the HDR, scale-and-splice imaging method (bottom). The control lens was oriented with the curved surface facing away from the collimated beam (left) and towards the collimated beam (right). The images are all displayed with the same spatial and gray colormap scale. The Zemax analysis was conducted by Hawks [23].

For valid comparison with the photon sieves, it was essential that the lens be used in the proper orientation. The proper orientation was confirmed by both experimental

and Zemax analyses, which were shown to be in agreement by a qualitative comparison between Figures 45 and 46. The rough agreement between the experimental and theoretical results had the added benefit of providing some support for the sieves' experimental results presented in Sections 3.2.1-3.2.3 and 3.3.2(p. 111) which were produced following the same (or highly similar) procedure.

3.2 Photon Sieve Characterization

The performance of the photon sieves was thoroughly evaluated with a number of conventional optical tests. A resolution target imaging test was used to gauge the imaging performance of the photon sieves. Next a thorough PSF analysis was conducted including comparison against simulated results, a focusing analysis, and determination of focal spot size. Finally, an interferometric analysis was conducted to measure the WFE produced by the photon sieves.

3.2.1 Air Force Resolution Test Target Imaging.

In order to evaluate the photon sieves' resolutions, a negative 1951 U.S. Air Force Resolution Test Target was imaged after being placed at optical infinity. The Verdi laser ($\lambda = 532.1$ nm) was operated between 0.1 W and 0.4 W and was passed through absorptive ND filters to decrease the power. These filters were necessary because of the high sensitivity of the CCD and 0.09-s minimum integration time. The filters also reduced stray reflections and improved imaging contrast, which is why they were used in combination with a higher laser power instead of operating the laser at lower powers without the filters. Next, the beam passed through a rotating ground glass diffuser to eliminate the degrading effects of laser speckle on the test target image. The diffused beam was diffracted by the small elements of the resolution target, which caused the light to diverge. An achromatic doublet (collimating lens) was placed

with the resolution target at its focus to collimate the light, and after measuring the distance with a meter stick, a Shear Plate Collimation Tester was used to verify that the separation was exact. The collimating lens (shown in Fig. 47 as L1) was oriented with the flat surface facing the resolution target to minimize spherical aberration. The collimated light was then focused by the photon sieve (represented as "TO", the test optic, in Fig. 47) which was positioned with the chrome coating on the back surface. Irises were used to block stray light, including unwanted diffraction orders from the photon sieves. A 20× microscope objective with $f = 8.0$ mm (shown in Fig. 47 as L2) was used to magnify and image the focal plane onto the detector.

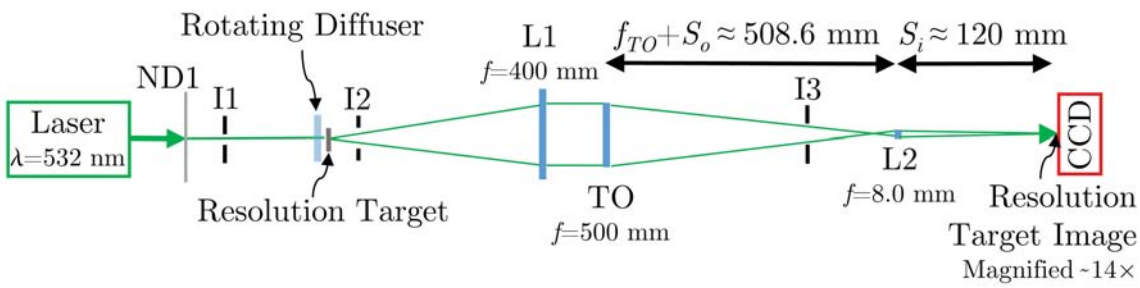


Figure 47. The experimental setup Resolution Test Target imaging. "ND" denotes neutral density filters, "I" denotes an iris, "TO" represents the test optic (each photon sieve), and "L" represents a lens. The number following the ND filter is just to differentiate ND filters, and does not indicate the optical density of the filter used.

The microscope objective was translated with a micrometer along the length of the optical system in order to adjust the object and image distances and reach perfect focus. Initial alignment was completed with the square photon sieve since its square diffraction orders and cross-shaped PSF were conducive to proper alignment. The 5.08 cm diameter collimating lens (L1) did illuminate the square photon sieves, which measured 50.8 mm along the diagonal, because of the 1.27-mm thick retaining rings needed to secure L1 in its mount. This meant that for all tests, 1.27 mm at each corner of the square photon sieves was left unilluminated, equating to 7 mm^2 (0.5%)

of the square sieve's surface. Though this almost certainly produced slightly degraded results, observation of the zero order and its absent corners aided in perfectly centering the collimated beam on the photon sieves.

Table 2. Resolution Target Data

Group	Element	Spatial Frequency, [Line Pairs/mm]	Line Width, [μm]	White Line Separation (Center-to-Center), [μm]
6	1	64.0	7.81	15.62
	2	71.8	6.96	13.92
	3	80.6	6.20	12.40
	4	90.5	5.52	11.04
	5	101.6	4.92	9.84
	6	114.0	4.38	8.76
7	1	128.0	3.91	7.82
	2	143.7	3.48	6.96
	3	161.3	3.10	6.20
	4	181.0	2.76	5.52
	5	203.2	2.46	4.92
	6	228.1	2.19	4.38

After aligning the system with the square sieve the CCD was adjusted so that the upper right corner of the rightmost vertical line in group 6, element 5 of the resolution target (denoted "6-5") was situated on the central pixel. Data for group 6, element 1 (6-1) through group 7, element 6 (7-6) are provided in Table 2. The CCD was kept in exactly the same position for imaging with the remaining photon sieves. The TEC was set to 0±0.2 ℃ and the CCD was operated in the 3×3 summation binning mode (combining groups of nine pixels into individual super-pixels by summing the ADU), which considerably reduced blooming. The laser power and neutral density filters were adjusted until 0.1-s exposures had a full-image mean ADU of 12,500±1,000. To minimize the impact of vibration, three 0.1-s exposures were taken, from which the sharpest was selected. This imaging process was repeated for each photon sieve, with consistent alignment being maintained by switching photon sieves without initially

moving any parts of the setup except for the photon sieves themselves. The new photon sieve was then translated vertically and horizontally parallel to the detector until the upper right corner of the rightmost vertical line in 6-5 was situated on the central pixel. Collimation was checked with the shear plate prior to imaging for each photon sieve and the focus was verified by once again shifting the microscope objective along the length of the optical system until the sharpest image was produced. Since all the photon sieves were designed with $f = 500$ mm, changes in the microscope objective's position between tests were minute (tens of micrometers) so magnification changes were insignificant.

The image distance from L2, S_i, was measured as 120±10 mm. The high uncertainty was primarily due to the sensor being recessed an unknown distance inside the camera body behind a shutter. Knowing the image distance and focal length, the Gaussian lens formula was used to calculate the object distance. Magnification was calculated as the ratio of the image distance to the object distance. However, because the resolution target was positioned at optical infinity, the exact magnification was not important for interpreting the results. The magnification could have been arbitrarily high without changing the results, because the intermediate image being produced by the photon sieve was in turn being imaged onto the detector. The quality of this intermediate image was fully dependent on the optics of the photon sieve. It was simply necessary that the magnification be high enough that the optics of the photon sieve were the limiting factor, rather than the pixel size of the detector.

The magnification used, 14±1, was ample given that for the circular sieve one resel (the half-width of the PSF's central lobe) equals 8.0 μm. With a 14× magnification that became 112 μm, which was seven times the binned super-pixel width. Using the size of a resel to predict resolution target results is an approximation (as discussed later in this section), but the approximation was certainly good enough

that the experimentally determined resolution was limited by the sieves rather than the pixel size given that the super-pixels were roughly $1/7^{th}$ the width of a resel. Initial results were significantly worse than expected, and it was observed that the test optics performed worse with the diffuser, which should not have been the case. It was discovered that the results could be improved considerably by placing the diffuser perfectly parallel with and immediately prior to the Resolution Target. Initially, the two were separated by approximately 50 mm, but the results presented were attained with the minimum possible separation of 3 mm. The full results of the resolution target test are presented in Table 3 and Fig. 48, with data on the resolution target itself being provided in Table 2.

Table 3. The Rayleigh prediction and the results (having met the Rayleigh criterion) for resolution target imaging with the photon sieves. A general trend of worsening resolution with decreasing throughput is observed, though for the high-throughput apodizations the resolution decrease is marginal. Note that this relation between throughput and resolution only holds for the general shape of the apodizing functions selected for this study (with high transmission near the center of the aperture and low transmission at the edges).

Photon Sieve	% Throughput (Apodization)	Rayleigh (Group-Element)	White Line Separation (Center-to-Center), [μm]	Resolved (Group-Element)	White Line Separation (Center-to-Center), [μm]
Circular	-	6-6	8.76	7-3	6.20
Circular, Sonine ($\nu = 5$)	89.2	-	-	7-3	6.20
Square	-	7-1	7.82	7-3	6.20
Square, Pyramid ($2/3^{rds}$)	81.4	-	-	7-2	6.96
Square, Pyramid ($1/3^{rd}$)	62.3	-	-	7-2	6.96
Square, Sonine ($\nu = 5$)	79.2	-	-	7-3	6.20
Square, Sonine ($\nu = 4$)	73.5	-	-	7-2	6.96
Square, Cos^2	44.1	-	-	7-2	6.96
Square, Cos^4	26.7	-	-	6-6	8.76
Square, Gaussian	19.2	-	-	6-6	8.76

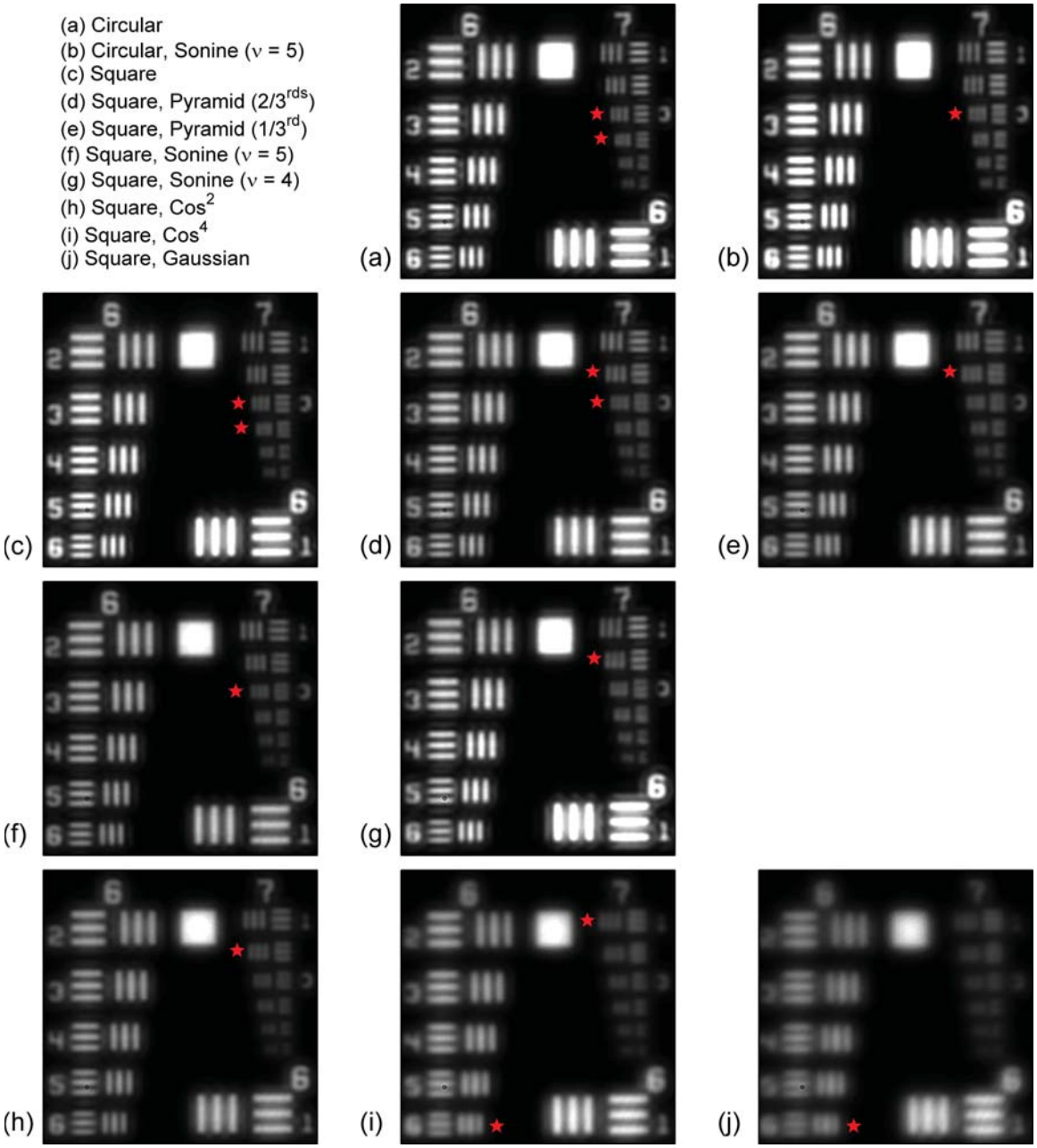


Figure 48. Images of the resolution target produced by the photon sieves. The smallest resolved elements for each photon sieve are labeled with an adjacent red star. Elements are distinguished by their row and column number. For example, in (a), the smallest resolvable element is somewhere between 7-3 and 7-4. The unapodized photon sieves exceeded the Rayleigh criterion for resolution. As expected, increasing apodization (meaning decreasing throughput) led to a change in shape of the optical transfer function and an effective decrease in the cut-off spatial frequency. Note that this relation between throughput and resolution only holds for the general shape of the apodizing functions selected for this study (with high transmission near the center of the aperture and low transmission at the edges).

For the unapodized circular sieve, the Rayleigh criterion is met when the central peak of the Airy pattern produced by a point source is directly on top of the first zero of the second source's Airy pattern. Therefore, according to the Rayleigh criterion, the minimum resolvable separation between two imaged point sources is the distance between the peak and the first zero (one resel), which is given by

$$(\Delta\ell)_{min} = \text{Resel}_{\text{circle}} = \text{ Min. Resolvable Separation } = 1.22 \times \frac{S_i \lambda}{2w}. \tag{19}$$

The resel is dependent on the image distance ($S_i = f$ in this case) and the half-width of the optic, w [22, p. 158]. Equations 19 and 17 are similar because both the far field and the field in the focal plane of an optic (for a source at infinity) are given by the Fourier transform of the source field, in other words — the focal plane of an optic also produces its Fraunhofer diffraction pattern [22]. With $f = 500$ mm, $2w = 40.5$ mm, and $\lambda = 532.1$ nm for the circular sieve, this gives $(\Delta\ell)_{min} = 8.01$ μm. Using Eq. 19 as an estimate, the center-to-center separation for the lines on the resolution target must be at least 8.0 μm for them to be resolved. Because the black and white lines have equal width, the center-to-center distance between adjacent white lines was twice the line width given in Table 2. This means that the smallest resolved element for the circular photon sieve should be 6-6, with center-to-center separations between white lines of 8.76 μm, slightly greater than one Resel. Element 7-1 had center-to-center separations of 7.82 μm, which were slightly less than one resel and should have been nearly resolved. As shown in Fig. 48(a), the circular photon sieve actually resolved 7-3, and possibly 7-4, which have center-to-center separations of 6.20 μm and 5.52 μm, respectively. As shown in Table 2, 7-3 had a spatial frequency of 161.3 line pairs/mm, where a line pair refers to a pair of adjacent black and white lines. According to the Rayleigh Criterion in Eq. 19, the optical transfer function

(OTF) cut-off spatial frequency for the circular photon sieve was significantly lower: 125 line pairs/mm.

Since the square sieve's PSF was much more similar to a 2-D Sinc^2 than an Airy pattern, the 1.22 in Eq. 19 is replaced with 1 to give

$$(\Delta\ell)_{min-square} = \text{Resel}_{\text{square}} = \text{Min. Resolvable Separation } = \frac{S_i\lambda}{2w}. \tag{20}$$

This equation gives $(\Delta\ell)_{min-square} = 7.41$ µm for the square photon sieve, for a spatial frequency of 135 line pairs/mm. This suggests the square photon sieve should have resolved 7-1 at 128.0 line pairs/mm and 7.82-µm center-to-center separation as shown in Table 2. As displayed in Fig. 48(c), the square photon sieve resolved 7-3 and possibly 7-4, like the circular photon sieve. This means the square photon sieve performed 20-34% better than predicted by Eq. 20.

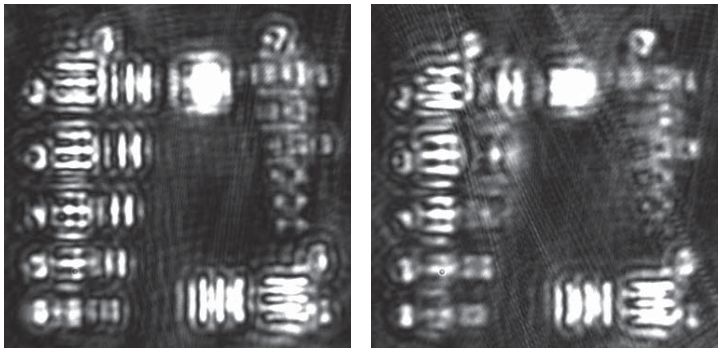
However, the above analysis does not take into account that the line patterns on the resolution target were different from point source illumination and that the Rayleigh criterion was "slightly less than the true cut-off frequency" [5]. The rectangular lines that made up the resolution target produced 2-D Sinc^2 functions rather than Airy patterns for the irradiance in the Fourier transform plane (in this case what was being imaged onto the detector) since they were not true point sources. The distance to the first null for a Sinc was 18% less than for an Airy pattern produced by a circular source of the same width placed at optical infinity, so the lines on the resolution target were more easily resolved than if they were point sources. Additionally, as stated by Andersen, the Rayleigh criterion does not give the true cutoff frequency [5].

The Sparrow condition, a second criterion for predicting resolution, gives estimates much closer to the true diffraction-limited cutoff frequency and has been shown to be more realistic for astronomical studies using equal brightness stars [24, p. 474].

The Sparrow condition is met when the overlapping irradiance distribution from two point sources has a flat top between the peaks and the second derivative of a cross section through the irradiance distribution equals zero at the center between the two peaks [24, p. 474]. For equal intensity linear sources — similar to the resolution target — the Sparrow condition gives a theoretical resolution 26% better than the Rayleigh criterion [39, p. 86]. In close agreement with this prediction, the circular sieve performed 29-45% better than the Rayleigh criterion and the square sieve performed 20-34% better than the Rayleigh criterion, depending on whether 7-4 was considered to be resolved. Closer inspection of the irradiance distribution suggests the Sparrow condition was met for 7-4 for both the square and circular photon sieves. Therefore, the results suggest that the performance of the photon sieves was diffraction limited as expected, and that any manufacturing or design defects were insignificant.

The apodized sieves also appeared to be producing diffraction limited images considering that all of the sieves except for the Gaussian apodized sieve met or exceeded the Rayleigh criterion for their unapodized counterpart. The Gaussian apodized square sieve resolved 6-6, compared to the Rayleigh criterion prediction of 7-1 for the unapodized square sieve. The Gaussian apodized square sieve had the most extreme apodization with a throughput 70.8% less than the square sieve, and was expected to perform the worst on resolution related tests. Additionally, these results showed that imaging resolution performance trends generally with the severity of the apodization — also in agreement with theoretical prediction. Note that this relation between throughput and resolution only holds for the general shape of the apodizing functions used in this experiment (with high transmission near the center of the aperture and low transmission at the edges). It is promising that high throughput apodizations could be applied with only a minor decrease in resolution: the pyramid and Sonine apodized sieves all resolved 7-2 or 7-3.

Resolution Target Imaging: No Diffuser. The rotating ground glass diffuser, shown in Fig. 47 immediately preceding the resolution target, was necessary for eliminating the degrading effect of laser speckle on the quality of the resolution target image. Images of the resolution target taken with the circular and square photon sieves without the rotating diffuser are shown in Fig. 49. These images show the same portion of the resolution target (groups 6 and 7) as displayed in Fig. 48. A comparison between the two figures illustrates how essential the diffuser was for attaining high resolution images.



(a) Circular Sieve (b) Square Sieve

Figure 49. Images of groups 6 and 7 of the resolution target taken without the diffuser with the circular sieve (a) and square sieve (b). These images illustrate the importance of the diffuser in achieving reliable results from the resolution target images.

White Light Resolution Target Images. Though wide bandwidth performance of photon sieves has been shown repeatedly to be poor, it was considered worthwhile to compare white-light imaging performance for the square and circular photon sieves [5]. The laser was blocked, the diffuser was removed, and a phosphor converted, white LED was put in its place immediately preceding the resolution target in the setup shown in Fig. 47. As displayed in Fig. 50, the white LED output peaked at 442 nm and 549 nm. Since the distance to the plane of best focus varies with illuminating wavelength for photon sieves, it is possible that better images could

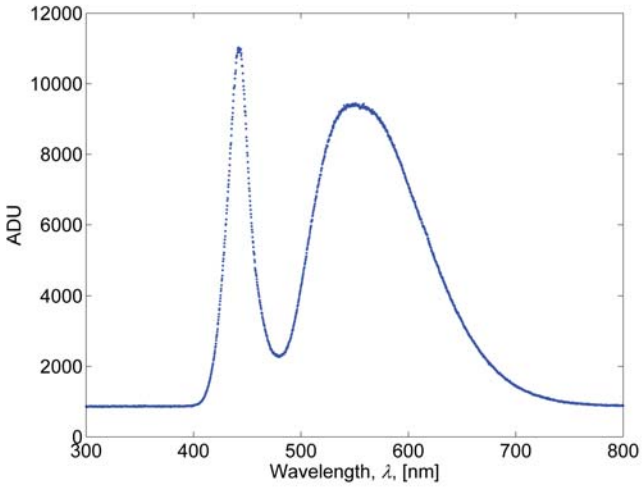
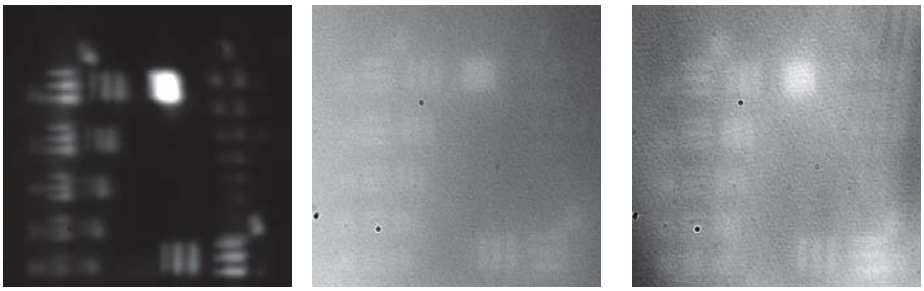


Figure 50. The spectrum of the white LED source used to illuminate the resolution target for wide bandwidth performance testing was measured with an Ocean Optics fiber spectrometer with a 3,648 pixel, 200-1100 nm range Silicon detector. According to the manufacturer specifications the corrected linearity of the detector is >99%. This graph shows that this source is appropriate for wide bandwidth testing since it produced photons over a roughly 300 nm range, with considerable power being output away from the 532.1 nm design wavelength of the photon sieves.

have been attained by adjusting the microscope objective to move its object plane closer to the focal plane of the 549-nm light. The unfocused light would still have been present, but since the source output more photons at 549 nm, there would have been a higher ratio of focused to unfocused photons, resulting in better contrast. This adjustment was attempted, but no noticeable improvement in the image contrast was observed, perhaps due to aberrations resulting from operating tens of nm away from the design wavelength. Since no improvement was observed, focus was achieved using the 532.1 nm laser as the source.

The resulting images are shown in Fig. 51, with an image produced by an achromatic doublet lens presented in (a) as a control. The histograms of all three images shown were clipped to maximize contrast. The histogram for (a) was only slightly clipped (inward shift of the black and white points by roughly 5%), while the histograms for (b) and (c) were appreciably clipped (inward shift of the black and white



(a) Doublet Lens (b) Circular Sieve (c) Square Sieve

Figure 51. Images of groups 6 and 7 of the resolution target taken with wide bandwidth (white LED) illumination with an achromatic doublet lens (a) the circular sieve (b) and square sieve (c). The histograms were clipped to maximize contrast for all three images. These images illustrate the difficulty of wide bandwidth imaging with the photon sieve, and show that there is no pronounced difference between the performance of the square and circular photon sieve.

points by roughly 35%). No adjustments whatsoever were made to the histograms of the images shown in Fig. 48; the only change necessary was increasing the laser power for the photon sieves with lower throughput. Without adjusting the histograms, there was little detail visible in Fig. 51 (b) and (c). Comparing (a) to (b) and (c) shows that, as expected, photon sieves perform much worse than conventional optics for wide bandwidth imaging. The differences between (b) and (c) are not pronounced, but it appears the square sieve produced a slightly higher contrast image. This suggests that a square photon sieve may have a marginally wider useful imaging bandwidth than a circular sieve, though no physical reason for why this would be the case is immediately apparent. Instead, it is likely that any difference between the contrast in (b) and (c) resulted from the histogram adjustment or unperceived variations in the experimental setup.

3.2.2 PSF Analysis and Simulation Comparison.

One of the most important factors in direct exoplanet imaging is the shape of the PSF. Considering the PSF produced by the parent star as noise, if the PSF drops off

rapidly enough from the peak, the exoplanet can produce a spike in irradiance on the detector sufficiently high above the noise floor that detection becomes possible. These photon sieves were designed following the investigation in Chapter II to produce PSFs with shapes suited to exoplanet imaging. To characterize these new photon sieve designs, their PSFs were imaged using the HDR scale-and-splice method (Section 3.1.4) and thoroughly analyzed. Additionally, this experimental analysis served as a useful test of the validity of the numerical Fresnel diffraction simulations carried out in Section 2.4. Since a goal of this thesis was to produce a photon sieve better suited than conventional optics for high-contrast imaging, a 50.8-mm diameter achromatic doublet (effective diameter = 48.3 mm due to the retaining rings) with the same focal length as the photon sieves (f = 500 mm) was used as a control for comparison with the sieves. This was the ThorLabs AC508-500-A lens described in Section 3.1.5.

Data Acquisition. Sources of stray light in the laboratory were removed or covered to the fullest possible extent using irises, graphite blocks, black polyurethane foam, black tape, black cloth, and black tubing. The TEC was set to -10±0.2 ℃. ND filters were used to reduce the power in the beam which was then spatially filtered and expanded with the 10× microscope objective, 10-μm pinhole, and f = 400 mm achromat, as shown in Fig. 52. The collimation was checked for each test using a shear plate collimation tester following L1. The collimated beam was then used to fully illuminate the test optic (except for the square photon sieves, which had their corners slightly clipped as discussed in Section 3.2.1). The alignment process was the same as described in Section 3.2.1, except with the center of the focal spot being positioned on the CCD's central pixel.

The focal spot was imaged and magnified onto the detector similarly to in Section 3.2.1, except with a different microscope objective for L2, with f = 15.36 mm±1% according to the manufacturer specifications. Image distance was measured as 1250±10

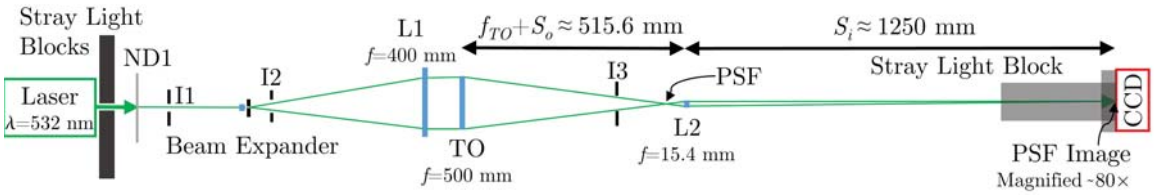


Figure 52. The experimental setup for PSF imaging. "ND" denotes neutral density filters, "I" denotes an iris, "TO" represents the test optic (each photon sieve), and "L" represents a lens.

mm, resulting in magnification = 80±1. The image acquisition software used, Nebulosity 4, was set to display the maximum single-pixel ADU for each image [40]. L2's position along the optic axis was adjusted finely with a micrometer while continuously taking 0.1-s exposures until the maximum single-pixel ADU reached a peak. This was determined to be imaging the focal spot, since the highest power was concentrated on a single pixel. For all test optics, this process resulted in the system imaging a spot 500 mm past the sieve (with a ±1 mm uncertainty).

After focusing, ND filters were adjusted until the maximum single-pixel ADU was roughly 60,000 for a 1-s exposure. This was done to ensure that as much dynamic range as possible was being used without saturating the detector. Following the HDR scale-and-splice method described in Section 3.1.4, ten 1-s light frames, ten 20-s light frames, ten 1-s dark frames, and ten 20-s dark frames were taken of the focal spot for each test optic. Averaging, subtraction, scaling, and splicing were done in MATLAB as described in Section 3.1.4(p. 62) with the pixels in the 20-s calibrated images with ADU below 80% of saturated being scaled linearly (dividing ADU by 20, the ratio of the exposure lengths) and replacing the same pixel in the 1-s calibrated image. The result was PSF images with dynamic ranges approximately 20× what could be captured with a single exposure (and relatively free of the impact of stray light because of the dark subtraction calibration).

The $\log_{10}$ of the data was taken to emphasize the subtle differences between the PSFs. Without doing this, it is much more difficult to compare the optics. After taking the $\log_{10}$, a colormap was applied. These results are presented in three different ways — as colormapped $\log_{10}$ images in Fig. 53, as colormapped $\log_{10}$ surface plots in Fig. 54, and as diagonal cross sections through the peak of the $\log_{10}$ data in Fig. 55.

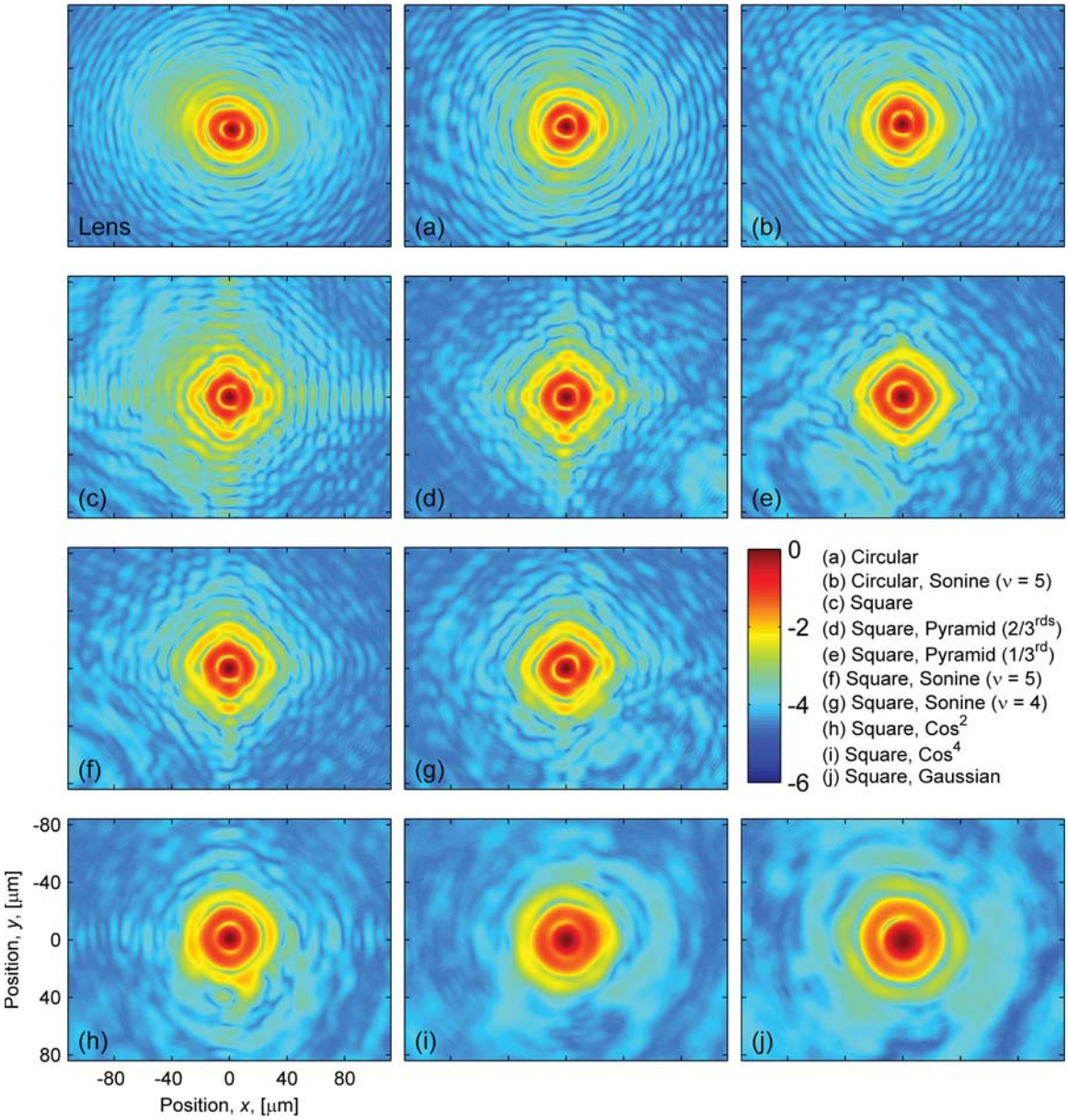


Figure 53. Log_{10} colormapped, HDR scaled-and-spliced images of the point spread functions produced by the control lens and the photon sieves. The positions given are in the plane of the PSF (the focal plane). The labels of the colorbar denote the value of the exponent, e.g., "-6" on the colorbar corresponds to a data point with a value at 10^{-6} of the peak value. All images are displayed at the same scale.

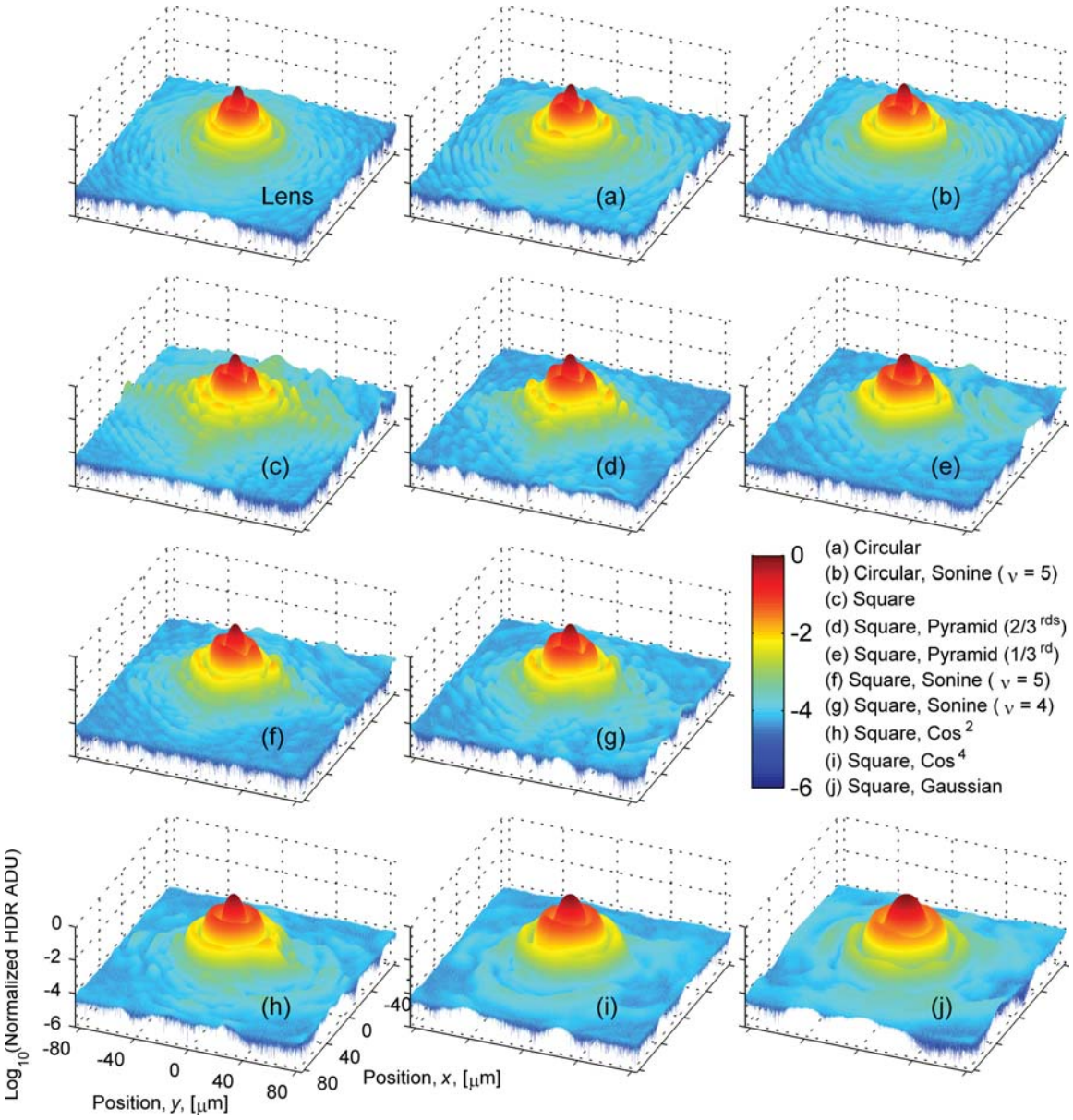


Figure 54. Log_{10} colormapped, HDR scaled-and-spliced surface plots of the point spread functions produced by the control lens and the photon sieves. The positions given are in the plane of the PSF (the focal plane). The labels of the colorbar denote the value of the exponent, e.g., "-6" on the colorbar corresponds to a data point with a value at 10^{-6} of the peak value. All surface plots are displayed at the same scale.

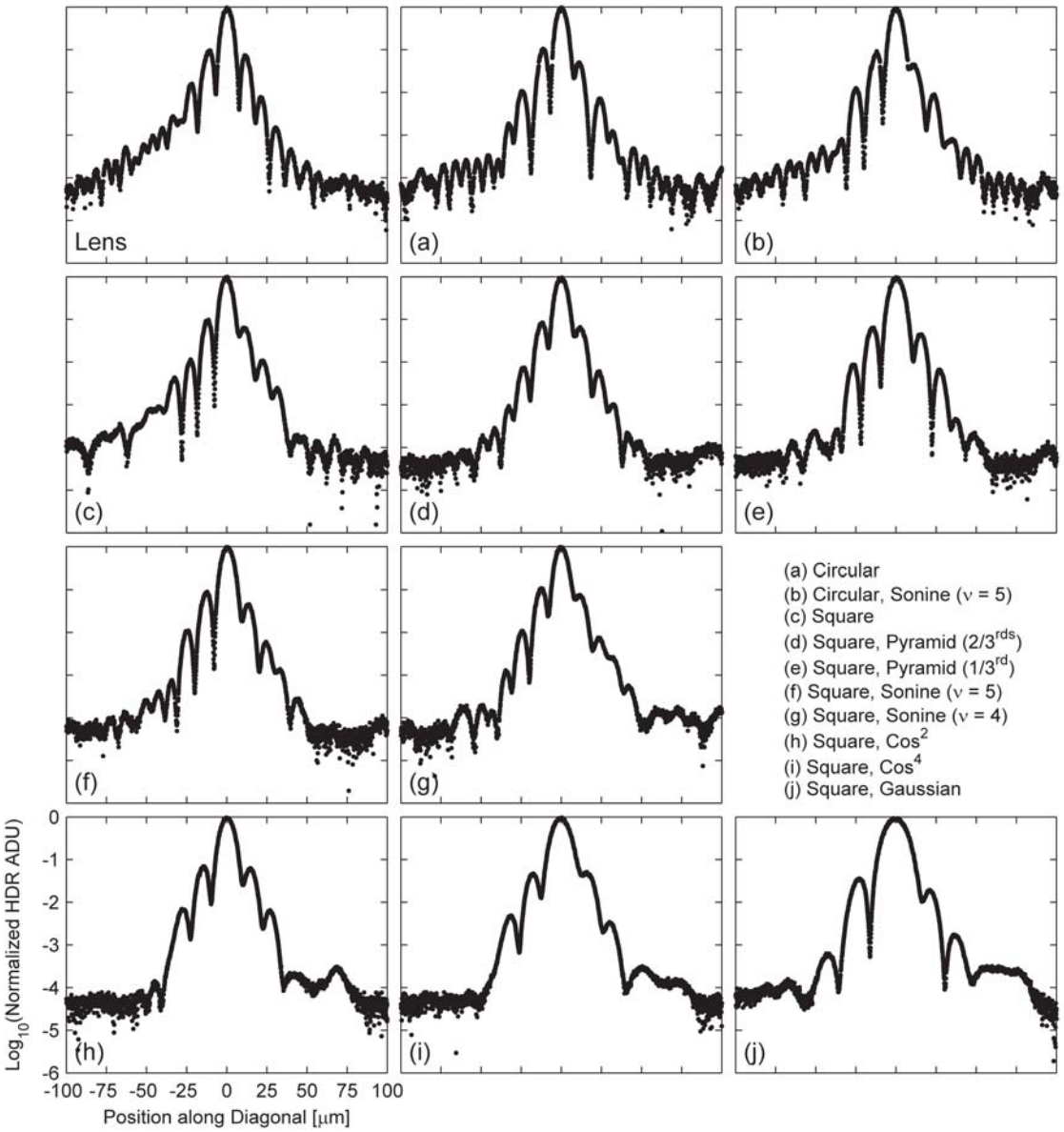


Figure 55. Log_{10} diagonal cross sections through HDR scaled-and-spliced images of the point spread functions produced by the control lens and the photon sieves. The positions given are in the plane of the focal spot. All plots are displayed at the same scale.

Fig. 53 best displays the impact of apodization on the shape of a PSF. Asymmetries present in the data are strongly exaggerated by the $\log_{10}$ scaling, and are likely due to back reflections and slight misalignment of the optical system. This is supported by the fact that asymmetries were observed (visually, with the laser power adjusted high enough to clearly see the magnified PSF) to shift while, for example, rotating L2. As expected, for the lens and the circular photon sieves (Lens, a, and b), the Airy pattern's rings are prominent. For the Sonine apodized circular sieve (b), the rings of the Airy pattern appear more square-like (rotated 45° relative to the orientation of square apodization applied to the optic), in accordance with the fact that a square-shaped apodization function was applied to a circular optic. For the square sieve (c), a significant amount of the diffracted energy is shown distributed along the vertical and horizontal axes, as expected. The square sieves with square shaped (pyramid and Sonine) apodizing functions (d, e, f, and g) have the horizontal and vertical diffraction spikes suppressed. The central lobes of the PSFs can be seen to widen with increasing loss due to apodization (decreasing throughput), and this is analyzed in depth in Section 3.2.3 and displayed in Fig. 63. This relation between throughput and resolution only holds for the general apodizing function shape used during this study (with high transmission near the center of the aperture and low transmission at the edges).

Fig. 54 provides the best display of the overall data floor, which appeared between 10^{-4} and 10^{-5} of the peak for all the test optics. The absolute data floor was set to approximately 10^{-6} by the processing method as described in Section 3.1.4, though it is possible that noise or stray light were the true limiting factors. In some of the plots, there is obvious structure (such as the rings for the circular optics) all the way to the edges of the plot. This suggests the HDR scale-and-splice method was capable of capturing some of the PSFs across this roughly 200×160-μm^2 field of view. Only

the peaks of the outer lobes were captured for the circular optics since the nulls were far too narrow and deep to be properly sampled. Capturing the peaks was enough for a proper characterization since a successful detection of an exoplanet would almost certainly require that its peak irradiance be greater than the peak irradiance of these outer lobes. It is unlikely that an exoplanet could be detected in one of these nulls, given how narrow they are and how easily they are disturbed by a small amount of WFE. For some of the other optics, it is more difficult to tell whether the PSF was properly captured over the full field of view. There is very little patterned structure apparent near the edges of the plots in (i) and (j), making it difficult to distinguish between data and noise.

Fig. 55 displays diagonal cross sections through the PSFs. Much of the data is asymmetric (likely due to slight errors in alignment of the optical system) though the underlying structure is apparent. Fig. 55(Lens) appears very similar to the predicted appearance shown in Fig. 2. Fig. 55(d) shows the most symmetric data. It is difficult to conclude whether the PSFs of the square sieves fall off more rapidly along the diagonal than the PSFs of the circular optics. Overlaying the diagonal cross sections for more direct comparison showed there to be no obvious difference in how rapidly the PSF drops off between the circular, square, unapodized, and apodized (with high throughput) optics.

Comparison to Simulation. Simulations were run using the method covered in Chapter II(p. 13) for comparison against the data. The square photon sieve used for the simulations had a 7.00 mm side length and the circular sieve had a 7.90 mm side length (resulting in total surface area equal to the square sieve). Two different versions of the simulation were run, one where the sieves were designed as the central portion of the fabricated sieves with $f = 500$ mm and a second where the sieves were designed with an $f/\#$ equal to that of the fabricated sieves, resulting in

$f = 97.5$ mm. For the central portion simulation, the square sieve had 6,155 pinholes with the smallest diameter = 42 μm and the circular sieve had 5,539 pinholes with the smallest diameter = 53 μm. For the equal $f/\#$ simulation, the square sieve had 161,544 pinholes with the smallest diameter = 8 μm and the circular sieve had 154,782 pinholes with the smallest diameter = 10 μm. Due to RAM limitations, the grid resolution for the simulated apertures was 9 μm, meaning that the smaller pinholes of the $f/\#$ simulated sieves were very poorly sampled for the numerical Fresnel simulation. Because of the increased number of pinholes, the apertures for the equal $f/\#$ sieves required roughly 36× the code runtime to generate (about 4300 s each versus 120 s each). The purpose of the different simulations was to determine the best manner for approximating photon sieves with millions of pinholes using limited computational resources.

The same apodizing functions as were applied to the fabricated sieves were used for these simulations. As in Chapter II, the apodization was applied as a simple amplitude mask rather than selectively removing pinholes. Random WFE with rms = $\lambda/20$ was multiplied across the aperture matrix as well to simulate imperfect manufacturing and alignment of the optics.

In all the comparisons shown here the lens' PSF was simulated using the Airy pattern, rather than the numerical Fresnel simulation. It is shown merely as a reference, and the data for the lens was left out in computing the mean residuals. It was observed that the inner side lobes had considerably more energy in the experimental data than in the simulations. This could be due to a number of factors including defocus, design and test wavelength mismatch, misalignment, or simply inaccuracy in the simulations. Simulations were conducted for imaging a point 5 mm past the focus ($1.01f$) for the central portion sieves and 0.195 mm past the focus for the equal $f/\#$ sieves ($1.002f$). The larger multiplicative factor for the central portion sieves

accounted for the fact that the larger $f/\#$ meant the light diverged at a slower rate after passing through the focus. The unfocused simulations were conducted after it was observed that improper focus could have accounted for much of the extra energy observed in the side lobes. A subsequent focusing analysis (p. 98) concluded that the method used to achieve focus (discussed on p. 84) was accurate, and that other sources of error such as spherical aberration and misalignment were much more likely the cause of the excess side lobe energy.

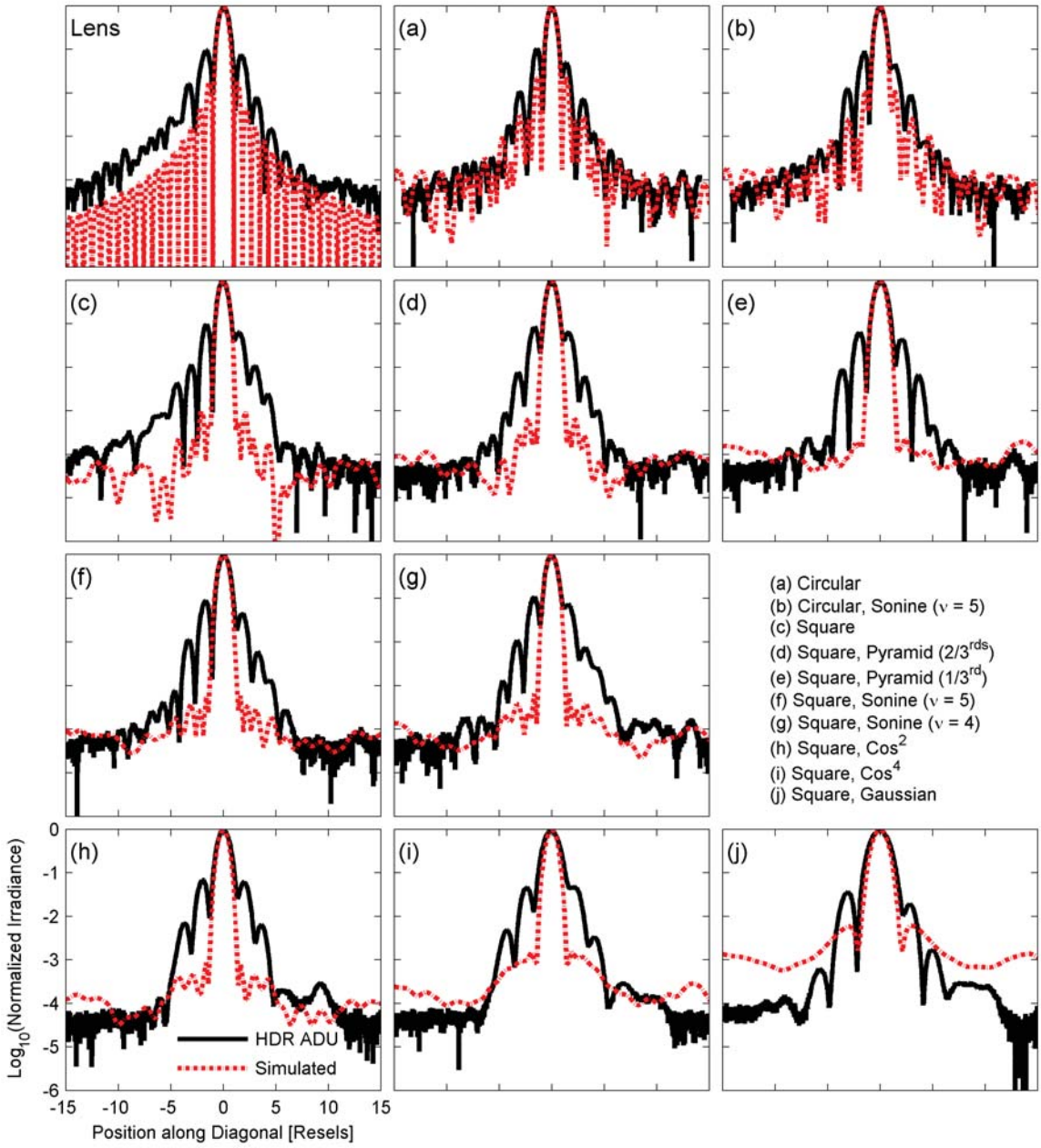


Figure 56. Central portion method. Log_{10} diagonal cross sections through HDR scaled-and-spliced images of the point spread functions produced by the control lens and the photon sieves plotted in black, with the numerical Fresnel simulation data plotted in red. Note that the simulated data for the control lens is from the Airy pattern and not the numerical Fresnel simulation.

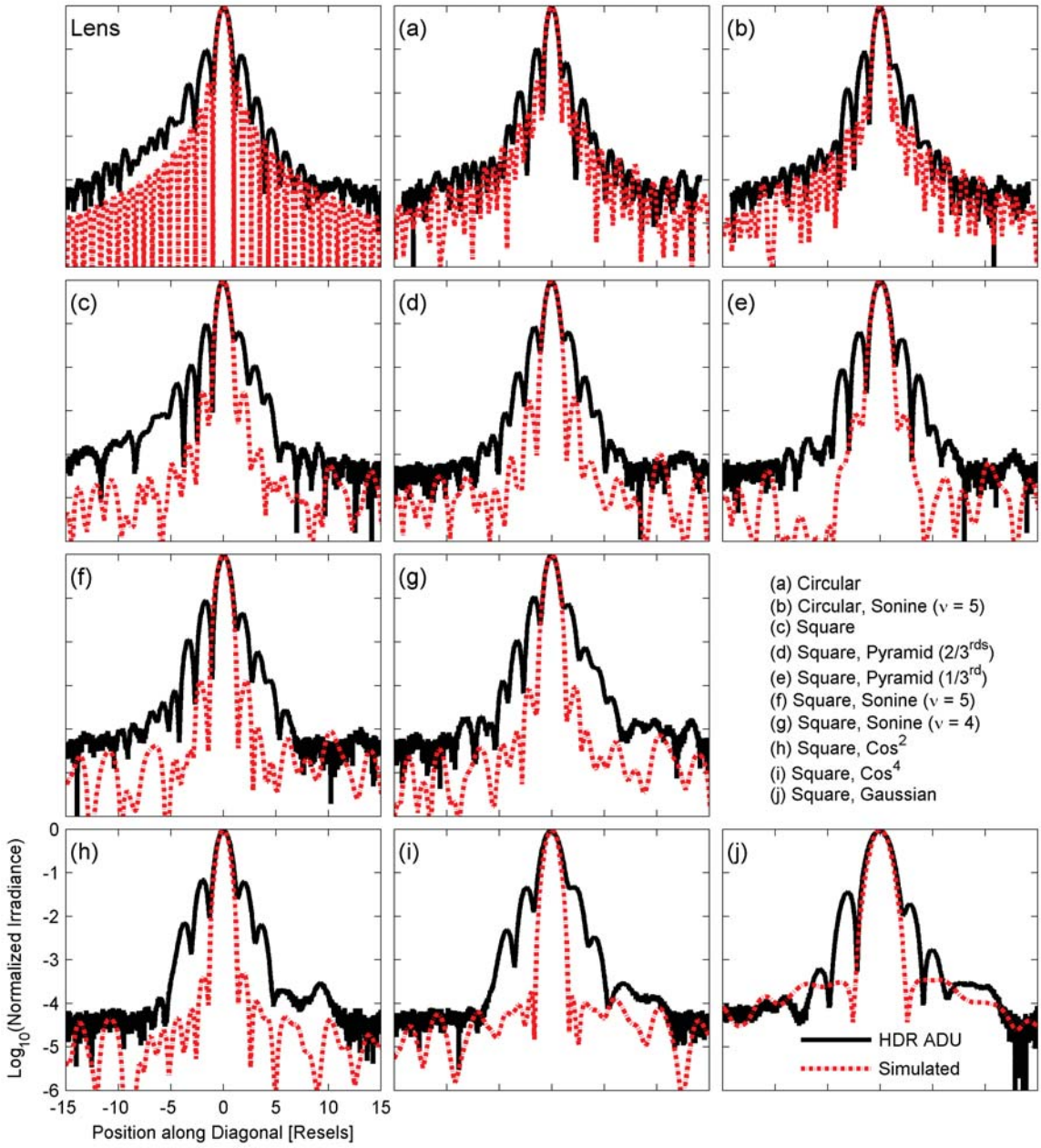


Figure 57. Equal $f/\#$ method. Log_{10} diagonal cross sections through HDR scaled-and-spliced images of the point spread functions produced by the control lens and the photon sieves plotted in black, with the numerical Fresnel simulation data plotted in red. The simulation was conducted with the equal $f/\#$ method. The simulated data for the control lens is from the Airy pattern and not the numerical Fresnel simulation.

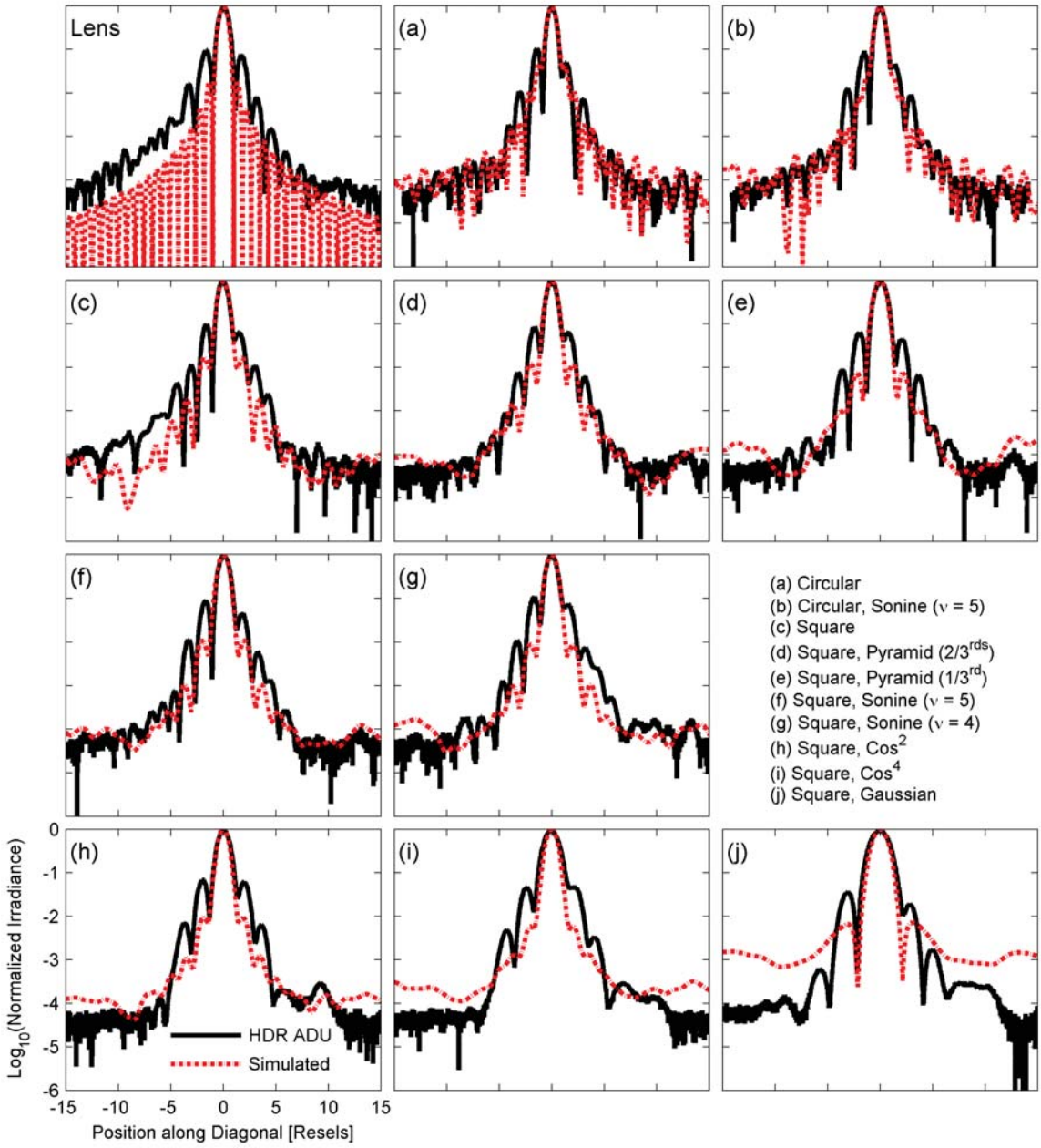


Figure 58. Central portion method, unfocused. Log_{10} diagonal cross sections through HDR scaled-and-spliced images of the point spread functions produced by the control lens and the photon sieves plotted in black, with the numerical Fresnel simulation data plotted in red. The simulation was conducted with the central portion method and was for 5 mm past the focal point. The simulated data for the control lens is from the Airy pattern and not the numerical Fresnel simulation.

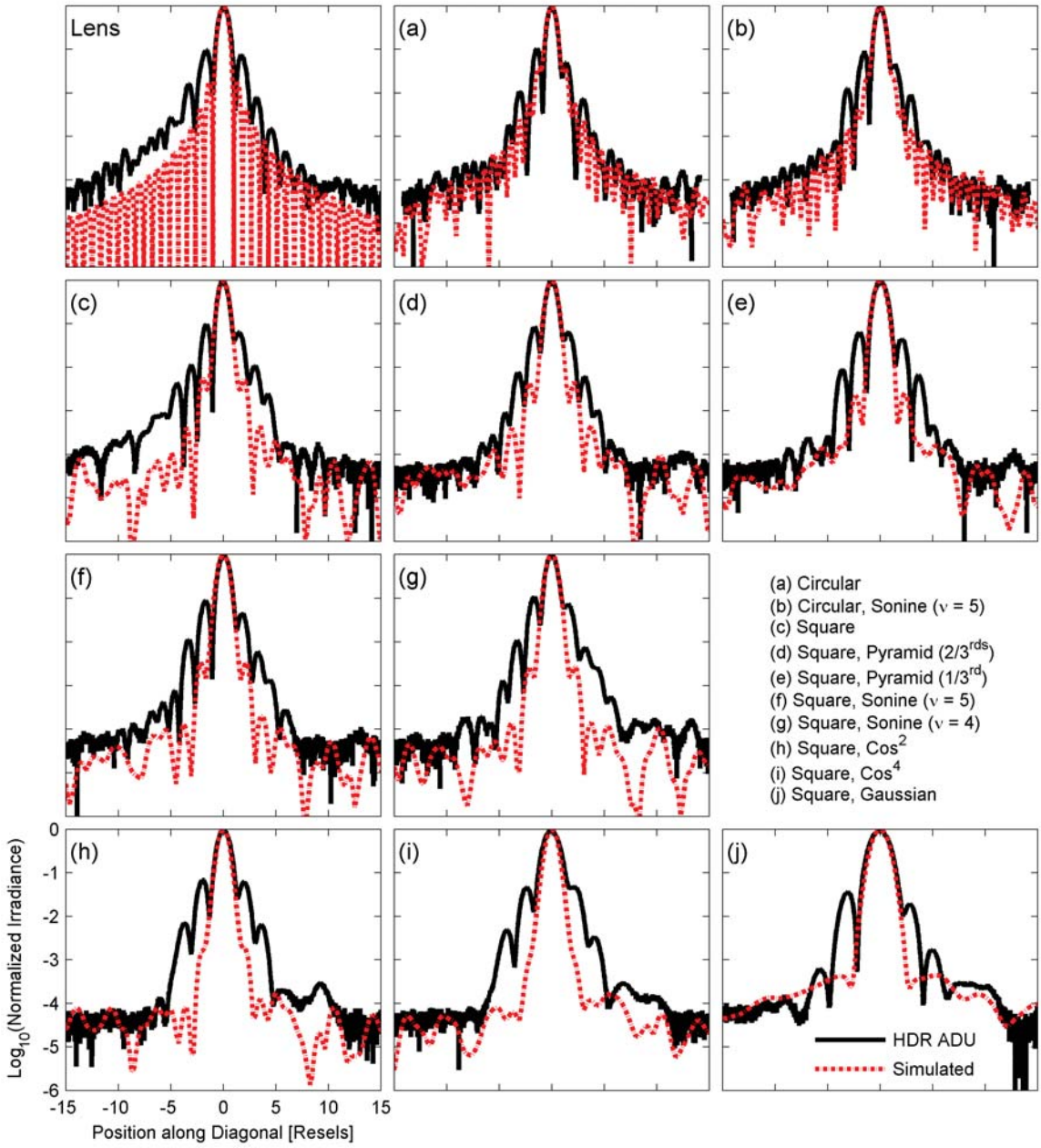
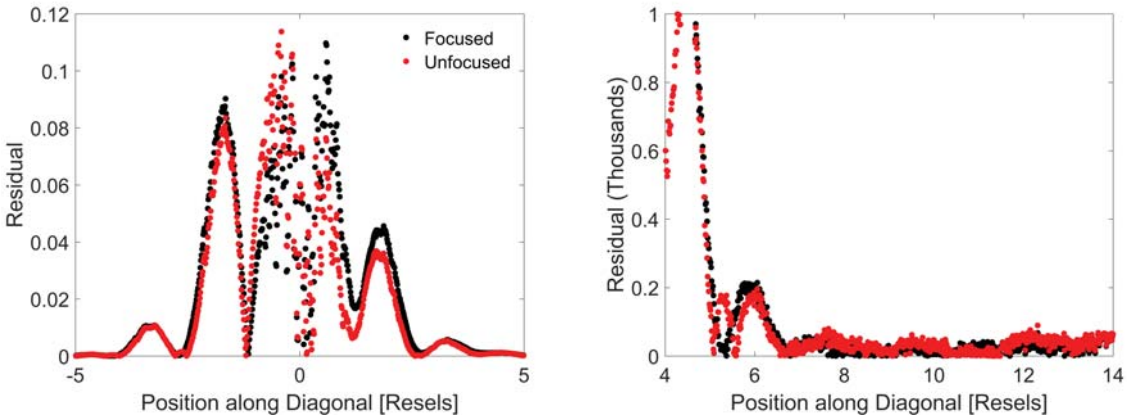


Figure 59. Equal $f/\#$ method, unfocused. Log_{10} diagonal cross sections through HDR scaled-and-spliced images of the point spread functions produced by the control lens and the photon sieves plotted in black, with the numerical Fresnel simulation data plotted in red. The simulation was conducted with the equal $f/\#$ method and was for 0.195 mm past the focal point. The simulated data for the control lens is from the Airy pattern and not the numerical Fresnel simulation.

The absolute value of the residual between the experimental and simulated data was calculated for each photon sieve. The mean and standard deviation of the residual were found for each photon sieve, then these were averaged to give a single measure of merit for each method. "Outer" residuals were also calculated, where only data points separated by more than four resels from PSF peak were included because the innermost side lobes were the largest contributors to the residual.



(a) Inner Residual (b) Outer Residual

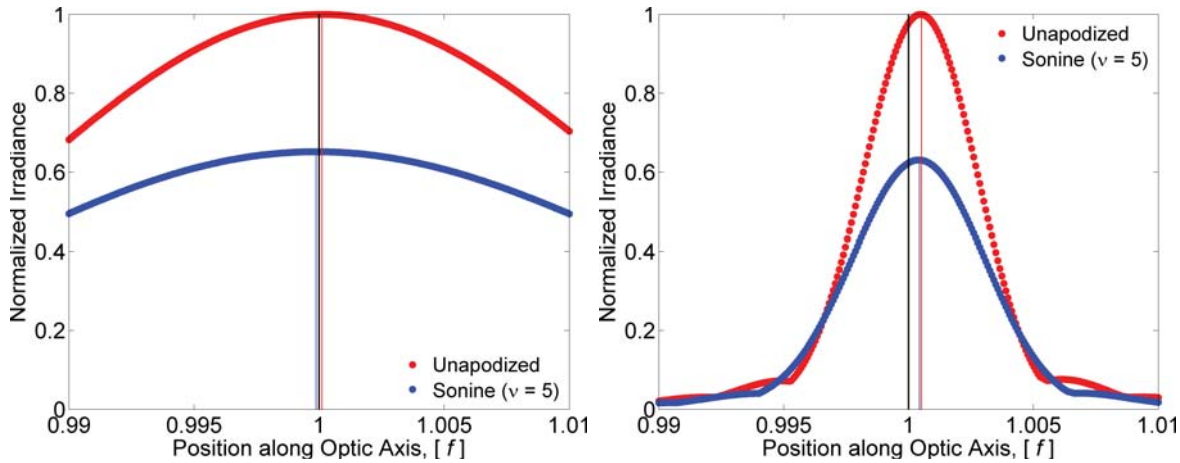
Figure 60. The inner (a) and outer (b) residuals for the focused and unfocused central portion method for the square Sonine ($\nu = 5$) sieve. In (a) it is shown that the unfocused simulation performed better overall, and the peak residual was less than 0.12. The outer residual plot (b) displays that more than four Resels from the peak the residual was roughly 0.001 or less.

Residuals are given on a linear scale with the $\pm$ uncertainty being the average standard deviation. The central portion method's (Fig. 56) overall average residual was 0.0084$\pm$0.022. The outer residual was 0.00025$\pm$0.00040. The equal $f/\#$ method's (Fig. 57) overall average residual was 0.0081$\pm$0.021. The outer residual was 0.00018$\pm$0.00043. The unfocused, central portion method's (Fig. 58) overall average residual was 0.0075$\pm$0.019. The outer residual was 0.00025$\pm$0.00037. The unfocused, equal $f/\#$ method's (Fig. 59) overall average residual was 0.0076$\pm$0.019. The outer residual was 0.00017$\pm$0.00042. The inner and outer residuals for the focused and

unfocused central portion method for the square Sonine ($\nu = 5$) sieve are shown in Fig. 60. This figure illustrates that most of the discrepancy was due to the inner side lobes.

Though the residuals appear very similar, these differences are exaggerated on a $\log_{10}$ scale, as can be seen by comparing Figures 56 and 57 to Figures 58 and 59. These very small differences are important for exoplanet imaging. For the overall residual, the unfocused, central portion method performed best. The unfocused, equal $f/\#$ method performed best for the outer residual. It is suggested that future work use the unfocused, central portion method since it was highly successful and the code runtime for generating the aperture was approximately 36× faster.

Focusing Analysis. The method for achieving focus described on p. 84 was suggested by Andersen [4]. The validity of this method was also supported by the Fresnel zone plate focusing analysis conducted by Cao and Jahns [16, p. 567]. A focusing analysis similar to the one presented in Fig. 5 on p. 567 of [16] was conducted for an unapodized square sieve and a square Sonine ($\nu = 5$) sieve using both the central portion (Fig. 61(a)) and equal $f/\#$ method (Fig. 61(b)) to determine if the focusing method resulted in extra energy in the side lobes.



(a) Central Portion Method (b) Equal $f/\#$ Method

Figure 61. A focusing analysis conducted for the unapodized square and square Sonine ($\nu = 5$) sieves using the two simulation methods. In both (a) and (b) the normalized peak irradiance is shown as a function of the position along the optic or propagation axis. The position is given in units of the sieve's focal length. In (a) and (b) both distributions were normalized to the unapodized distribution peak. The vertical black line shows the designed focus position and the vertical red and blue lines show the peak locations for the unapodized and apodized distributions, respectively.

In Fig. 61(a), the peak locations are shown to align almost exactly with the designed focus position. Fig. 61(b) shows how a lower $f/\#$ results in a narrower peak (as with conventional optics, photon sieves with low $f/\#$ have shallower depth of focus). Since the simulated sieves used to produce Fig. 61(b) were the same $f/\#$ as the fabricated sieves, it is expected that this is the more relevant distribution. The peak locations are farther along the optic axis than the designed focus position, with the peaks occurring at $1.0005f$. This would correspond to a 0.25-mm offset for the $f = 500$ mm fabricated sieves. The offset may not have been a physical effect and may have been due to a sieve design error or numerical artifact (such as the poor sampling of the pinholes in the outer zones, or the worsening of the Fresnel approximation with decreasing $f/\#$). This was less offset than the $1.002f$ propagation distance used to produce Fig. 59(p. 96), so even if this offset was physical, it could not have fully accounted for the extra energy observed in the inner side lobes. Further investigation

is required to determine whether the offset was a physical effect. Since spherical aberration shifts energy from the central lobe to the side lobes, it is likely that it was present in the experimental setup (and responsible for much of the residual between the experimental and simulated results), though its source is unknown [24, p. 254]. It is possible that CCD blooming, which was shown to be a problem in Section 3.1.3(p. 60), also contributed to the extra energy in the side lobes.

3.2.3 PSF Central Lobe Width.

The width of the central lobe of the focal spot is an excellent metric for characterizing and comparing optics. The data collected in Section 3.2.2 was further analyzed to determine the focal spot size for each optic.

The spot sizes were found using the HDR scaled-and-spliced images. Using MATLAB, the pixel with the maximum value in each image was found (P_1). This maximum pixel had a value of one, since it was the pixel to which the data were normalized. For every pixel in a 50×50 pixel region ($P_{1\text{-}2,500}$) centered on P_1, the average ADU of a 20×20 pixel region centered on that pixel ($P_{1\text{-}2,500}$) was found. The purpose of this was to smooth the data, and account for random fluctuation and dust on the optics. The size of the smaller region was chosen to give a large sample (400 pixels) without applying too much smoothing (<10% the width of the central lobe). From these 2,500 average values, the pixel with the maximum value was selected as the pixel (P) corresponding to the true peak of the PSF. The average of a 20×20 pixel region centered on each pixel in the column containing P was found, and the separation between the pixel locations of the first minima was taken to be the width of the central lobe. This process was done vertically to decrease the impact of blooming (shown in Fig. 42 (p. 61)). These widths ranged from 197 pixels for the control lens to 493 pixels for the Gaussian apodized sieve. The number of pixels was multiplied

by the pixel width (5.36 μm) then divided by the magnification (80) to convert the pixel width to a measurement with the same scale as the focal plane of the test optics. This gave the true central lobe widths in μm.

The central lobe widths are shown overlaid on normalized HDR images in Fig. 62, with the red circles defining the central lobe boundaries. The images are also labeled with the central lobe width in terms of resels. For the circular sieves, the size of one resel was computed with Eq. 19 (8.01 μm), for the square sieves, one resel was given in Eq. 20 (7.41 μm), and for the control lens, one resel was calculated with Eq. 19 using 48.3 mm as the optic's width to give 6.72 μm. As shown, the lens and unapodized photon sieves met or exceeded diffraction limited expectations (2 resels) for the central lobe width. As expected, the apodized sieves had wider central lobes than their unapodized counterparts. Overall, the apodized sieves performed well, with five of the eight apodized sieves producing central lobes less than one-third wider than their unapodized counterparts. Fig. 63 shows the percentage increase in the width of the central lobe as a function of the loss in throughput due to apodization for the square photon sieves, and shows that central lobe width increased with decreasing throughput as expected. The exact throughput values for the sieves are given in Table 1, and the % loss shown in Fig. 63 was computed by subtracting the % throughput from 100.

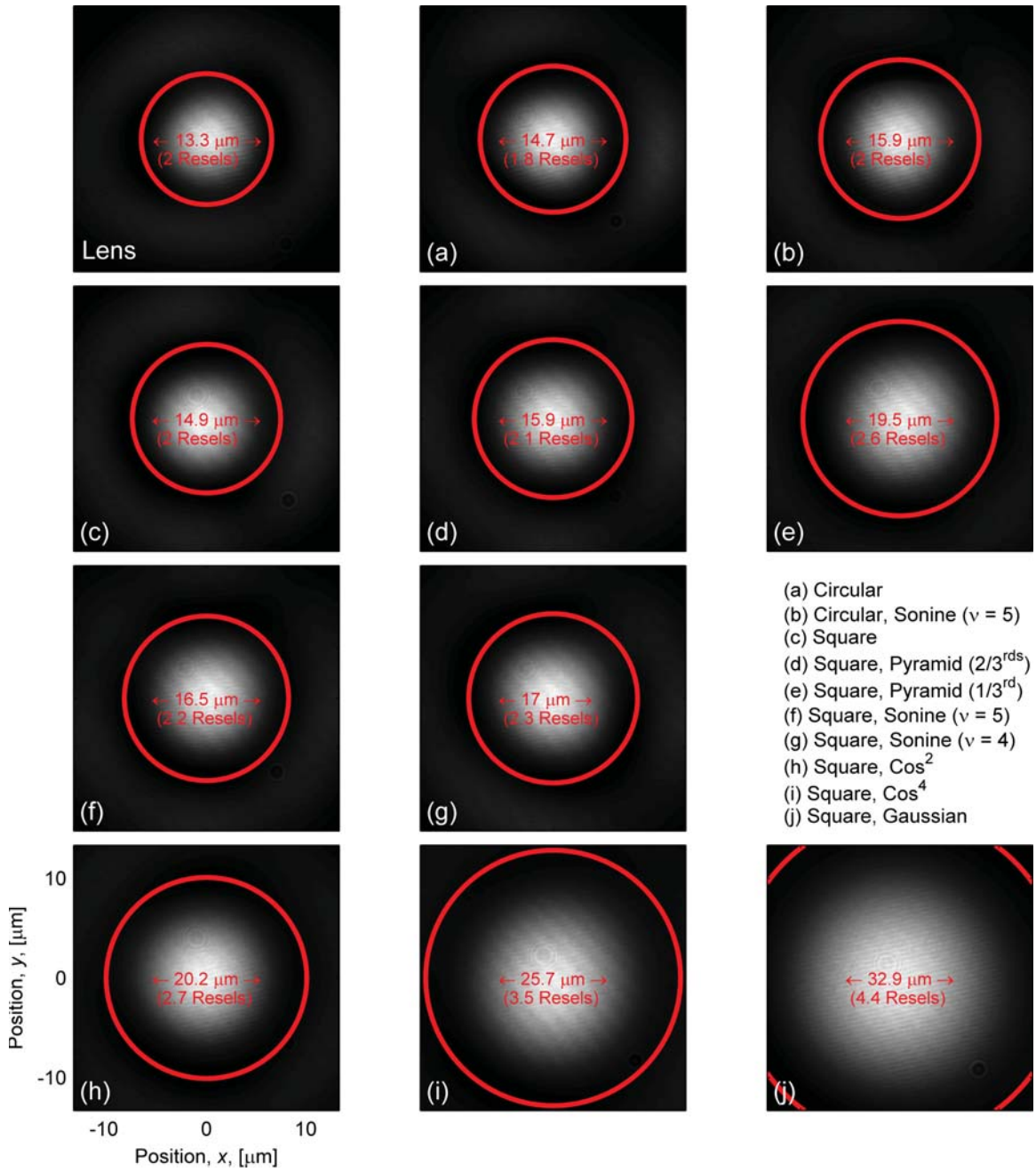


Figure 62. The results of the spot size measurement are shown overlaid on normalized HDR scaled-and-spliced images of the focal spots for each optic. The red circles define the boundaries of the central lobes. The images are displayed with a colormap of black = 0 and white = 1. All images are displayed at the same scale.

The results in Figures 62 and 63 further support the conclusion from Section 3.2.1 that photon sieves can be apodized considerably without significantly degrading image resolution. While the square, Sonine ($\nu = 5$) sieve performed slightly better on the resolution target test (resolving 7-3 compared to 7-2 for the square, pyramid ($2/3^{\text{rds}}$) sieve), it performed slightly worse on this test (with a central lobe width of 16.5 μm compared to 15.9 μm). As shown in Fig. 48, the square, pyramid ($2/3^{\text{rds}}$) sieve may have resolved 7-3, so this disagreement may be attributed to the somewhat subjective interpretation of the resolution target results. Aside from this single discrepancy, the resolution target and spot size results aligned exactly.

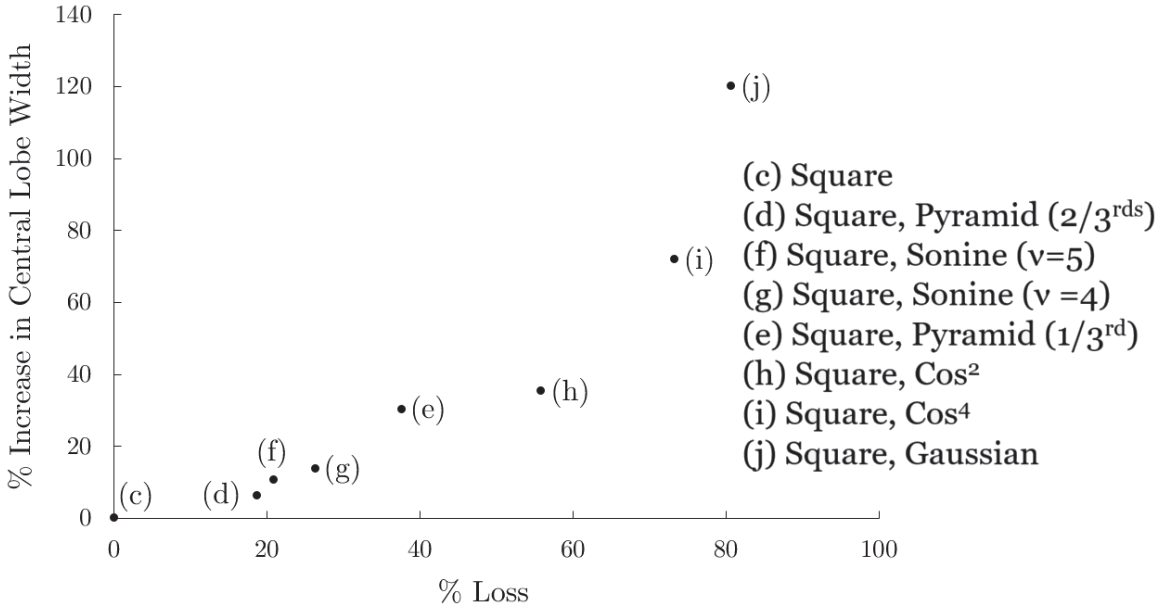


Figure 63. Central lobe width of the PSF plotted as a function of loss due to apodization for the unapodized (c) and apodized (d-j) square sieves. The central lobe width was observed to increase with increasing loss. This is in agreement with the resolution target results presented in Section 3.2.1 (p. 71). Note that this relationship between throughput and central lobe width only holds for the general shape of the apodizing functions used in this study (with high transmission near the center of the aperture and low throughput at the edges).

3.2.4 Interferometric Analysis.

An interferometer was used to measure the optical quality of the photon sieves. As shown in Fig. 64, the focused light from the test optic was recollimated with L3 and then combined and interfered with a reference beam using a pellicle beamsplitter. Collimation following L1, L2, and L3 was verified with a shear plate collimation tester, and the lenses were all oriented to minimize spherical aberration. The laser was blocked at the position of ND1 and a piece of paper with a barcode pattern was placed on the surface of the test optic. It was illuminated (from the right side of the TO, as diagrammed in Fig. 64), and the CCD was translated along the optic axis until the barcode's image on the CCD was in sharp focus as shown in Fig. 65(a).

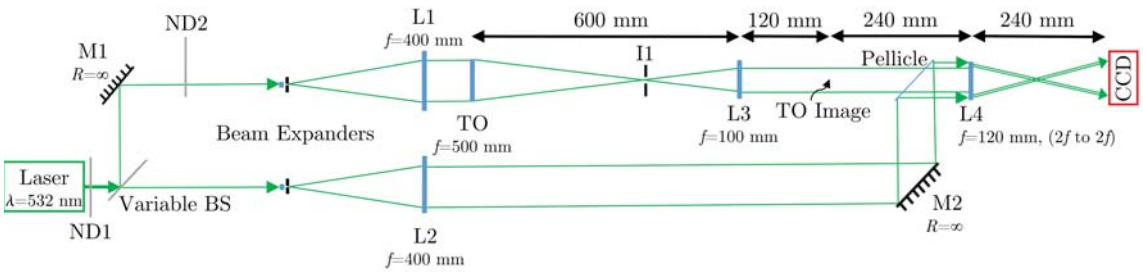
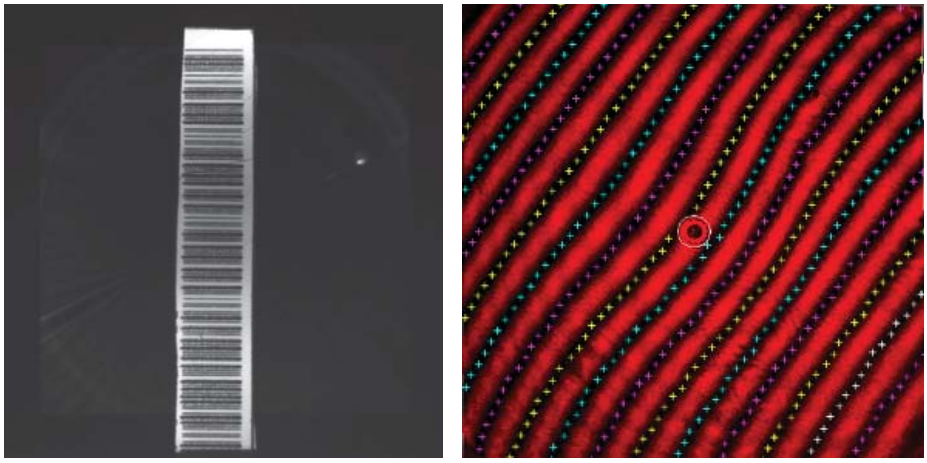


Figure 64. The experimental setup for the interferometer used to test the square and circular photon sieves as well as the control lens. "ND" denotes neutral density filters, "I" denotes an iris, "TO" represents the test optic, "L" represents a lens, and "M" represents a mirror.

The barcode was removed, the laser was unblocked, and ND1, ND2, and the variable BS were adjusted until interference fringes were observed (with good contrast between the dark and light fringes) on the CCD. Contrast was further improved by placing I1 at the TO's focus and blocking much of the zeroth order. A large number of 0.1-s exposures were taken of the interference fringes from which the best (having the highest contrast fringes) was selected for each test optic. This was necessary because

the optical table was not floating, so any small vibrations in the table significantly impacted the contrast of the fringes.



(a) Setup Verification (b) Square Sieve Interferogram

Figure 65. A barcode was placed on the TO's surface and illuminated (a). The CCD was translated along the optic axis until the barcode's image on the CCD (shown here) was in sharp focus. The test optic shown in this image is the square sieve, but the same process was repeated for the other test optics. The interferogram produced by the square sieve (b), after analysis using Quick Fringe by Asmolova [7].

Interferograms were analyzed by Asmolova using Quick Fringe, which traces the fringes to determine WFE as displayed in Fig. 65(b) [7]. The results for the control lens were a WFE of peak to valley = 0.66λ and rms = 0.14λ, which were significantly higher than the values predicted by the Zemax analysis in Section 3.1.5(p. 68) of peak to valley = 0.07λ and rms = 0.02λ. The circular sieve gave peak to valley = 0.48λ and rms = 0.10λ, and the square sieve gave peak to valley = 0.78λ and rms = 0.18λ. None of these values met Rayleigh's quarter wavelength rule (less than λ/4 peak to valley WFE) or Maréchal's condition of rms WFE below λ/14, both of which are indicators of a low-aberration system [10, p. 528].

The fringe analysis set a lower limit on the true performance of each test optic, as the interferogram incorporated the accumulated WFE from multiple other elements in the setup [2]. The excellent performance of the test optics in Sections 3.2.1(p. 71)

and 3.2.3(p. 100) suggests that the other elements (and misalignment) contributed the majority of the WFE observed in the interferograms. Assuming the control lens matched the schematic provided by the manufacturer, a best-case scenario that was used for the Zemax analysis (p. 69), more than 85% of the WFE observed in the fringe analysis resulted from the other optics in the experimental setup, and any misalignment of these optics. The recollimating lens, labeled L3 in Fig. 64, appeared to be contributing the majority of the WFE, since illuminating a shear plate collimation tester with L3 resulted in strongly curved fringes. After subtracting off this 85% outside contribution to WFE, which assumed this percentage held for the sieves, the circular sieve gave WFE of peak to valley = 0.07λ and rms = 0.02λ, and the square sieve gave WFE of peak to valley = 0.12λ and rms = 0.03λ. Given the assumptions that led to these results (the fact that the Zemax analysis presents a best-case scenario, giving the largest possible WFE fraction to subtract off, along with the assumption that this fraction holds reasonably well for the sieves) these values have very large uncertainties. However, these (best-case scenario) results, combined with the resolution target images and the central lobe width measurements, do suggest the sieves met Rayleigh's quarter wavelength rule and Maréchal's condition [10, p. 528].

3.3 Laboratory Simulated Exoplanet Imaging

A novel and important part of this investigation consisted of using the photon sieves to image sources that were be specifically designed to mimic extrasolar systems as closely as possible. These tests were most useful for characterizing the high-contrast imaging performance of the photon sieves. The semi-empirical and two-beam illumination methods presented in Sections 3.3.1 and 3.3.2 appeared the most promising, and were selected for use.

A third method, which involved passing the laser through a microscopic and closely

spaced pair of pinholes meant to mimic an extrasolar system, was decided to be physically implausible after further investigation. The method would have used the same manufacturing process as the photon sieves, and the manufacturing limitations imposed meant the desired damage threshold and feature sizes, and therefore resulting separation and PVRs, could not be achieved. Further investigation into this method, perhaps using a different manufacturing process, may be worthwhile. However, further investigation was deemed outside the scope of this thesis due to time constraints.

3.3.1 Semi-Empirical Method.

The semi-empirical analysis used the normalized, HDR scaled-and-spliced images captured and processed in Section 3.2.2. For each optic's PSF image, rings of pixels centered on the PSF peak with radii ranging from 20-80 μm (as measured in the PSF plane) in 0.05-μm increments were selected. This resulted in a total of 1,200 selected rings, each with more than 800 pixels. For each pixel in each of these rings, the SNR for an exoplanet with a PSF peak at 10^{-3} of the star's PSF peak was estimated by solving

$$\text{SNR} = \frac{\text{ADU} + \text{Dim Peak Value}}{\text{ADU}}, \tag{21}$$

where ADU was the value of the normalized, HDR scaled-and-spliced pixel. These SNRs were averaged together in three groups based on the radial and angular positions of the pixels to create plots of SNR as a function of angular position for three different radial distance ranges, as shown in Fig. 66. Each of the three solid lines shown for each test optic's plot was produced by averaging the SNRs from 400 rings, and smoothing with a robust lowess filter with a 5% span [31]. The MATLAB code is provided in Appendix E(p. 156). The radial distance ranges are 20-40 μm, 40-60 μm, and 60-80 μm, which equate to 2-5 resels, 5-7.5 resels, and 7.5-10 resels for the circular sieve. The lens had the advantage of having a greater width (and thus smaller resel)

than the circular and square sieves, so if not for the impact of apodization, the lens could be expected to outperform all the sieves. Even though the square sieve had a slightly smaller resel than the circular sieve (7.4 versus 8.0 μm), this was considered a fair comparison because the square and circular sieves had equal surface area and throughput. This means they would have nearly equal mass and be equally difficult to launch and deploy as the primary optic of a space telescope.

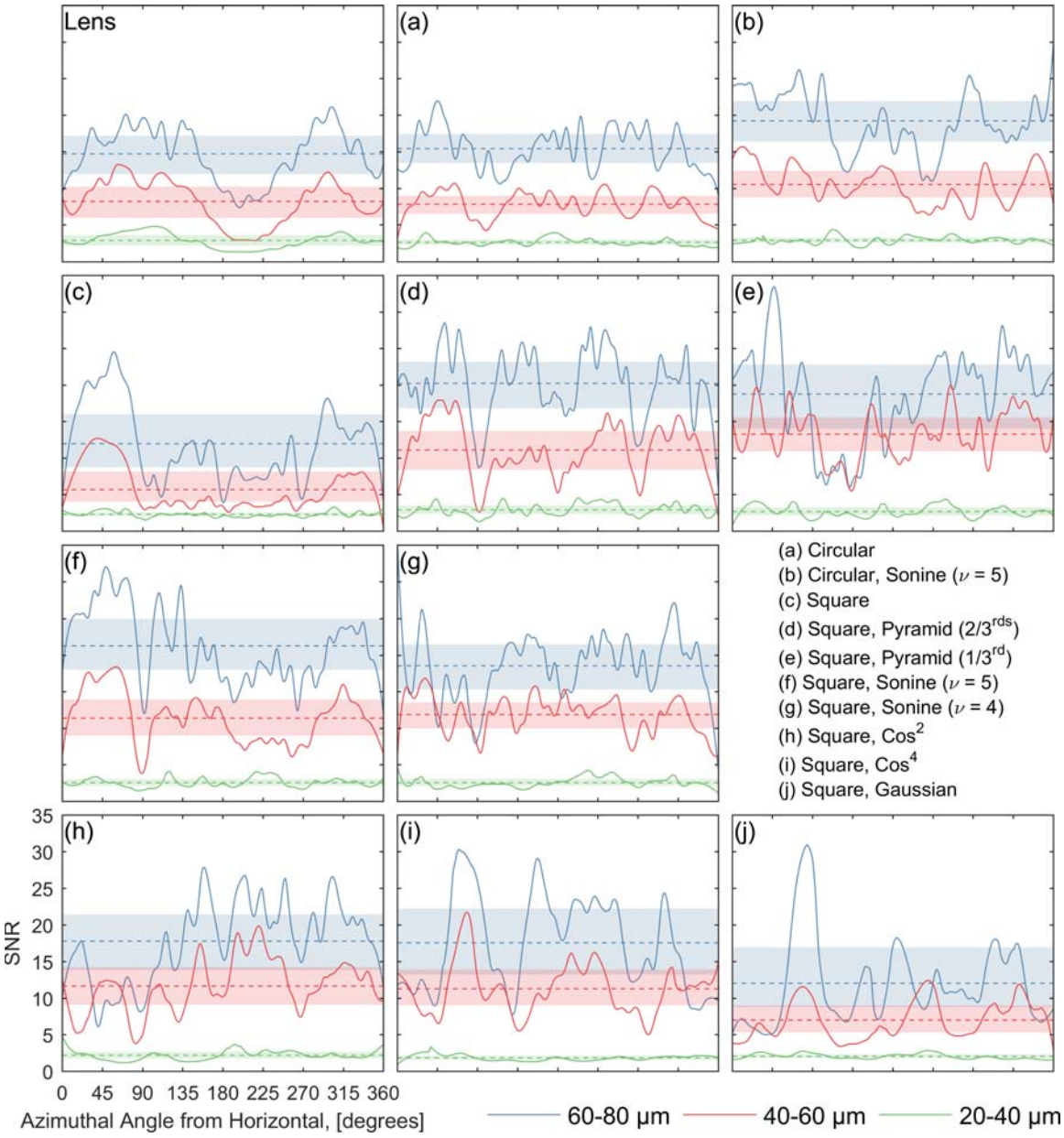


Figure 66. Estimated SNRs as a function of angular position (for three different distance ranges, shown as the solid lines) for an exoplanet with a PSF peak at 10^{-3} of the star's PSF peak. The dashed lines display the mean, averaged over the full 0-360° angular range, for each distance range. The distance ranges equate to 2-5 resels, 5-7.5 resels, and 7.5-10 resels for the circular sieve. The filled color regions display the bounds of the upper and lower semideviations, which were computed using code written by Alhouti [1]. The higher SNRs show that the apodized sieves largely outperform the unapodized optics.

As displayed in Fig. 66, the apodized square sieves (except for the Gaussian) outperformed the unapodized optics. As expected, dips in the SNR are visible on axis (at 0, 90, 180, and 270°) – especially for the square unapodized and pyramid ($2/3^{rds}$) sieves. SNR averages and upper and lower semideviations for positions more than 20° off-axis are shown for the three different distance ranges for each test optic in Fig. 67 [1]. Restricting the area of interest to positions more than 20° off-axis resulted in a 6% average improvement for the square sieves and a 9% average improvement for the square sieves with square shaped apodizations. The square, pyramid ($2/3^{rds}$) sieve improved the most (15%, leading to the best overall performance), while the performance of the circular Sonine sieve worsened by 7%. The square sieve improved by 13%, though it still performed poorer than expected.

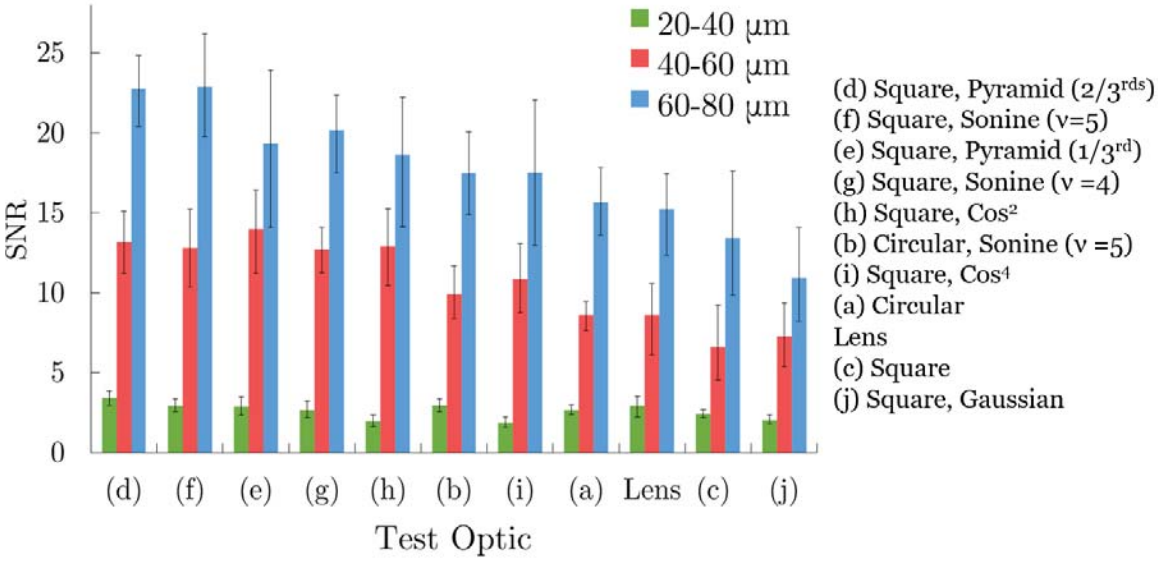


Figure 67. SNR averages from the semi-empirical analysis with the upper and lower semideviations for positions more than 20° off the axes of the image (near the diagonal) are shown for the three different distance ranges for each optic. The optics are sorted in order of decreasing performance (based on averaging together the SNR from the three distance ranges to get a single performance metric for each optic).

The results in Fig. 66 support the prediction that apodized sieves (except for the low-throughput Gaussian apodization) largely outperform unapodized optics for

high-contrast imaging. The results in Fig. 67 support the prediction that square sieves with square shaped apodizations have the best off-axis high-contrast imaging performance thanks to much of the diffracted energy being distributed on axis. The poor performance of the square sieve may have been due to unaccounted-for noise or stray reflection, since the square sieve's performance was roughly equal to the circular sieve's in the test presented in Section 3.3.2.

3.3.2 Two-Beam Illumination Method.

Two slightly offset collimated beams (with very different powers) were focused by each test optic to create slightly offset dim and bright PSFs, meant to simulate an exoplanet and its parent star. After passing the laser through ND1 and ND2 to reduce the power in the beam, as shown in Fig. 68, a beamsplitter was used to split the light and each beam was directed towards a spatial filter beam expander. The beam used to create the dim source was passed through ND3 (a variable ND wheel allowing fine adjustment) to decrease the power relative to other beam (the bright source's beam). While expanding, the two beams were recombined by reflecting the dim source's beam off a pellicle beamsplitter with a 45°angle of incidence. The bright source's beam passed directly through the pellicle, as shown in Fig. 68.

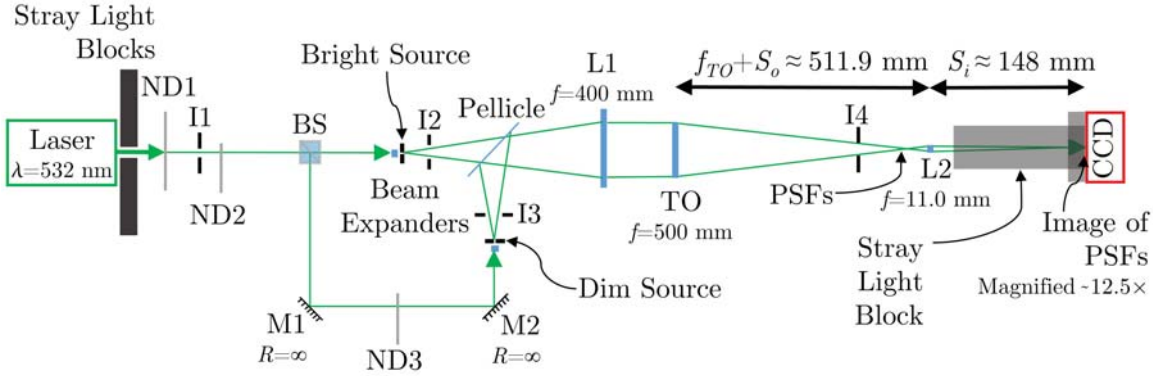


Figure 68. The experimental setup for simulated exoplanet imaging. "ND" denotes neutral density filters, "I" denotes an iris, "TO" represents the test optic (each photon sieve), "L" represents a lens, and "M" represents a mirror.

Two-Beam Illumination Test: 10^{-2}. The pellicle used was 92% transmissive and 8% reflective, which served to further increase the power ratio between the bright source's and dim source's beams. The test optic was used to focus the collimated light, and the focal plane was imaged and magnified with a microscope objective (L2) onto the CCD. This was repeated for each test optic, with alignment and focusing handled in the same manner as in Section 3.2.2 (p. 82). Very small adjustments to the positions of the beam expanders were used to offset the dim PSF along the bright PSF's diagonal with peak-to-peak separations ranging from 32-42 μm, as shown in Table 4. Dim and bright peak locations were found using the method described on p. 100, except using 10×10 pixel averages rather than 20×20 to account for the smaller magnification. The average separation was 36.8 μm (≈ 5 resels) with a standard deviation of 4 μm. Variations in separation and peak value ratio were largely due to slight mechanical drift in the three-axis spatial filters. Collimation was checked independently for each beam using a shear plate collimation tester.

Table 4. Two-Beam Illumination Test: Peak Separations and Peak Value Ratios

Photon Sieve	Peak-to-Peak Separation [μm]	Peak Value Ratio
Circular	31.9	$10^{-1.90}$
Circular, Sonine ($\nu = 5$)	41.8	$10^{-1.95}$
Square	35.5	$10^{-1.97}$
Square, Pyramid (2/3^{rds})	32.6	$10^{-1.97}$
Square, Pyramid (1/3rd)	42.3	$10^{-1.96}$
Square, Sonine ($\nu = 5$)	33.4	$10^{-1.91}$
Square, Sonine ($\nu = 4$)	32.8	$10^{-1.98}$
Square, Cos^2	38.6	$10^{-1.92}$
Square, Cos^4	39.2	$10^{-1.99}$
Square, Gaussian	39.8	$10^{-1.97}$

Prior to capturing data, stray light in the lab was blocked or eliminated to the fullest extent possible and the TEC was set to 0±0.2 ℃. Images were captured and calibrated following the method presented in Section 3.1.4(p. 62), with 20 s for the long exposures, 1 s for short exposures, and each exposure repeated three times with each test optic. Additionally, three long exposures were taken of only the dim PSF (with the bright beam blocked) with each test optic for use in determining the dim PSF's peak location and the PVR (the ratio of the dim peak's value to the bright peak's value). The ratio was set to roughly 10^{-2} as shown in Table 4 by rotating ND3. They ranged from $10^{-1.99}$-$10^{-1.90}$ with an average of $10^{-1.95}$ and a standard deviation of $10^{-3.08}$.

With $f = 11.0$ mm (with a ±1% tolerance according to the manufacturer) for L2 and the image distance (S_i) measured as 148±10 mm, the object distance (S_o) was calculated as 11.9±0.2 mm using the Gaussian lens formula. The ratio S_i/S_o was taken to find magnification = 12.5±1. The pixel width (5.36 μm) was divided by the magnification to match the scale to that of the test optic's focal plane (where a single pixel's width corresponded to 0.43±0.04 μm). This was used to properly scale the

results, and means that all distances presented with the results in this section have an uncertainty $\approx 10\%$. When comparing the results between test optics, this uncertainty becomes somewhat irrelevant. Since the final uncertainty resulted primarily from the uncertainty in S_i (which was constant for all test optics to within a fraction of a millimeter) and f (which was constant), any discrepancy in spatially scaling the data could be expected to impact all results equally.

The results of the two-beam illumination test are presented in four different ways: as colormapped surface plots in Fig. 69, as colormapped $\log_{10}$ images in Fig. 70, as colormapped $\log_{10}$ surface plots in Fig. 71, and as diagonal cross sections through the PSF peaks of the $\log_{10}$ data in Fig. 72.

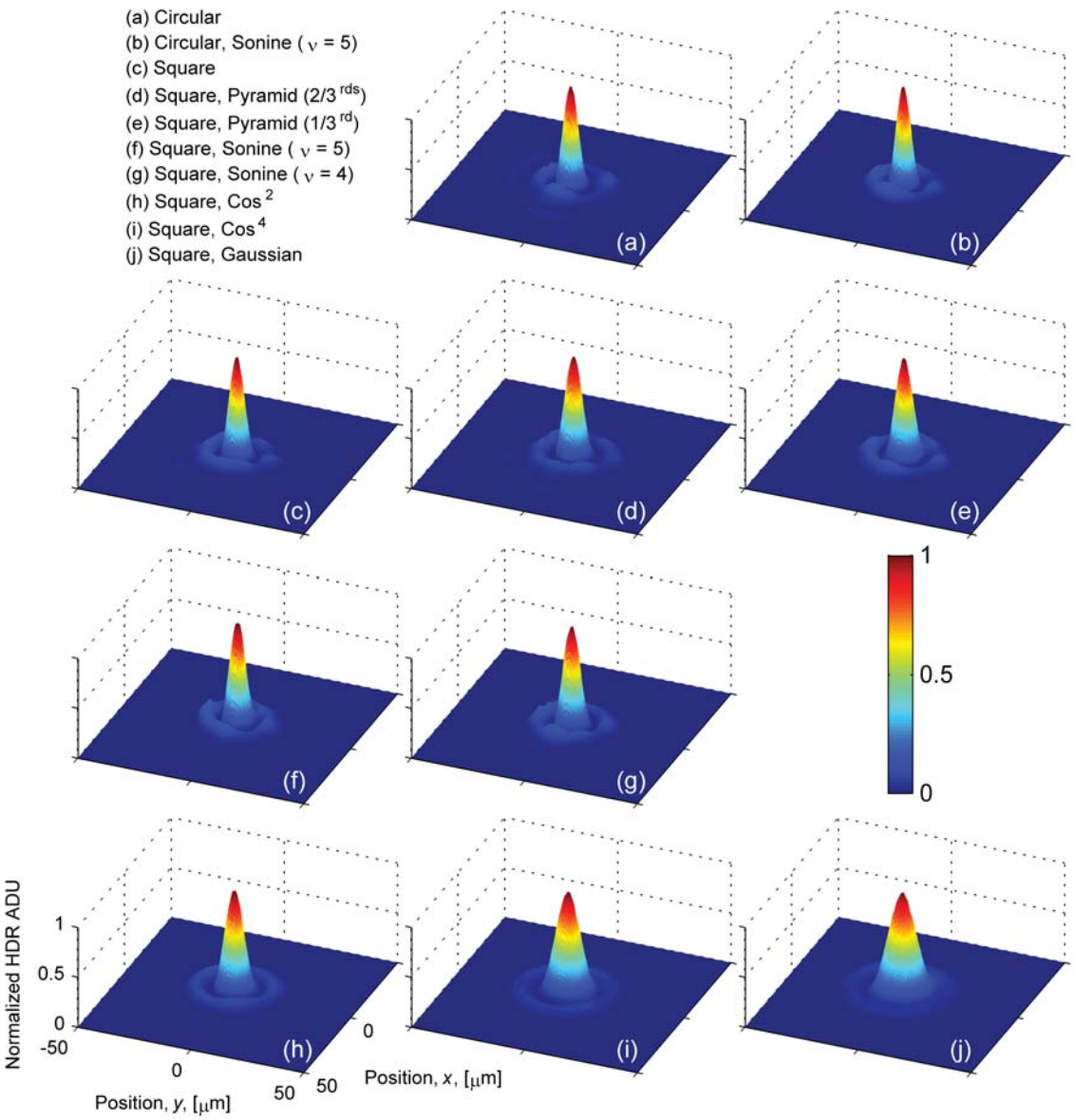


Figure 69. Surface plots displaying the colormapped, HDR scaled-and-spliced data from the two-beam illumination test. The power in the dim beam was decreased relative to the power in the bright beam using ND filters until the peak of its PSF was at roughly 10^{-2} of the bright source's peak. The bright PSF is centered in each surface plot, while the dim PSF (located near (25 μm, 25 μm)) is too faint to be visible. The positions given are in the plane of the PSFs (the focal plane) and all surface plots are displayed at the same scale.

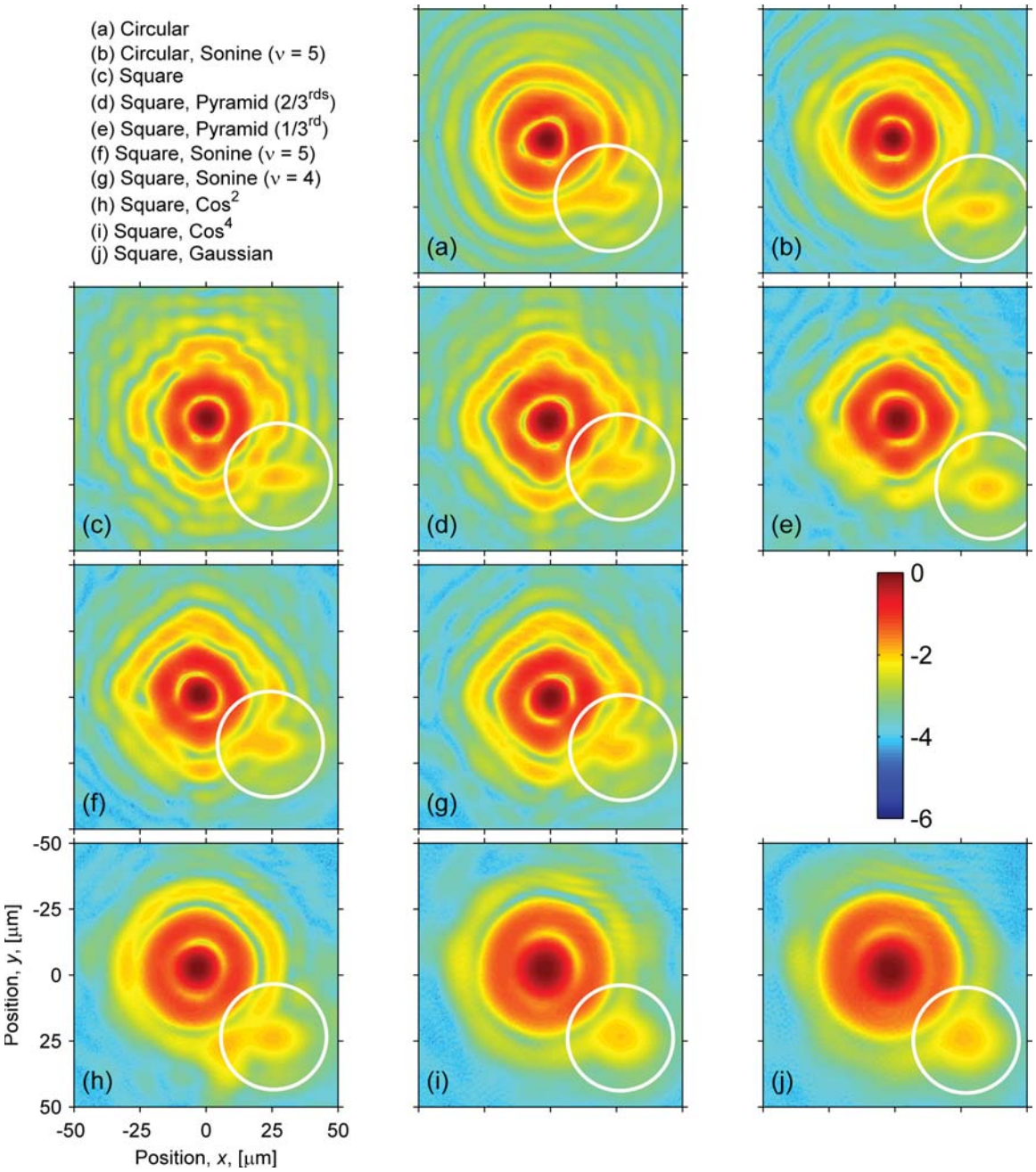


Figure 70. Log_{10} colormapped, HDR scaled-and-spliced images from the two-beam illumination test. The positions given are in the plane of the PSFs (the focal plane) and all images are displayed at the same scale. The labels of the colorbar denote the value of the exponent, e.g., "-6" on the colorbar corresponds to a data point with a value at 10^{-6} of the peak value. The power in the dim beam was decreased relative to the power in the bright beam using ND filters until the peak of its PSF was at roughly 10^{-2} of the bright PSF's peak. The bright PSF's peak is centered in each image, while the dim PSF's peak is shown circled in white.

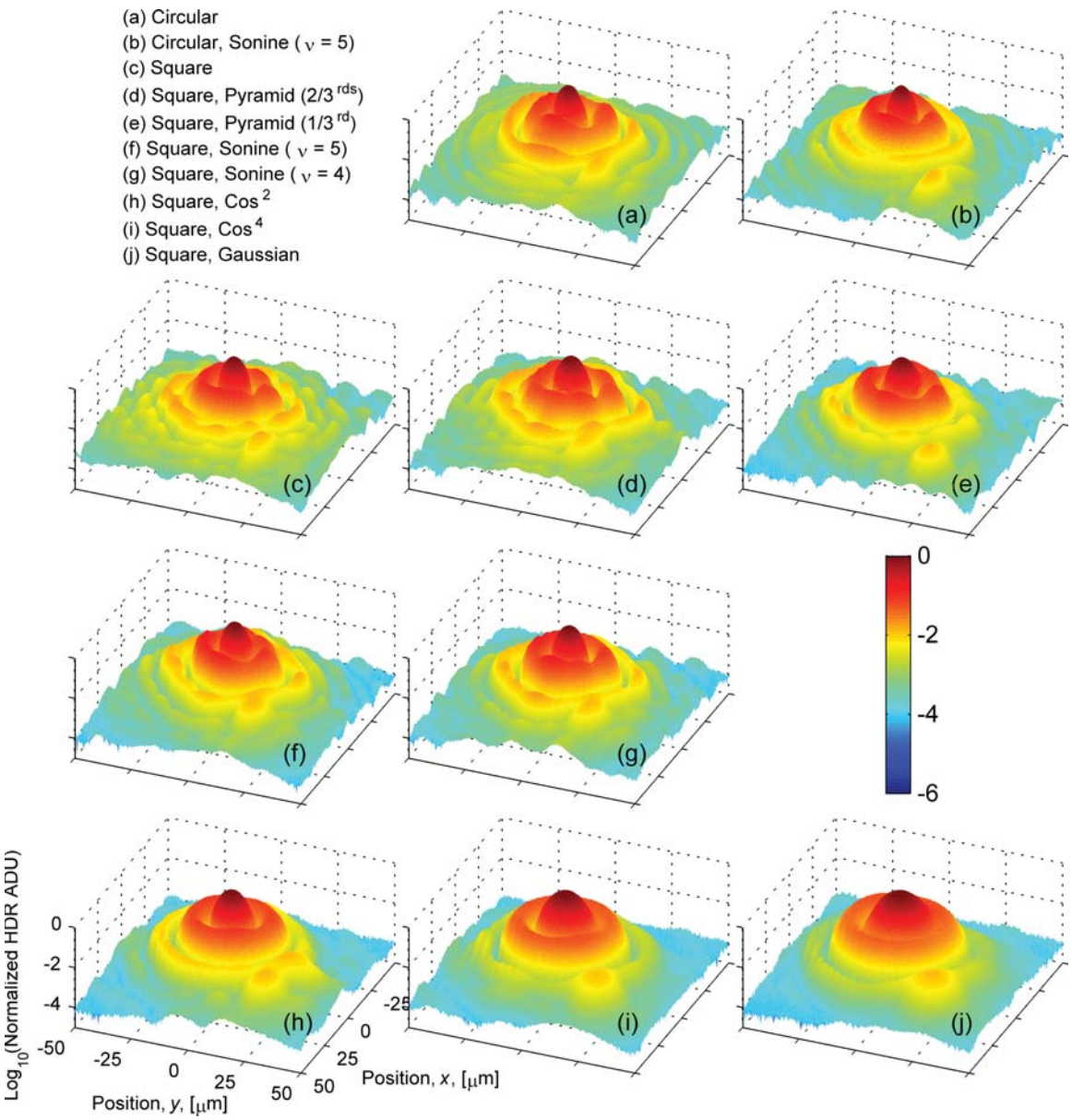


Figure 71. Log_{10} colormapped, HDR scaled-and-spliced surface plots from the two-beam illumination test. The positions given are in the plane of the PSFs (the focal plane) and all surface plots are displayed at the same scale. The labels of the colorbar denote the value of the exponent, e.g., "-6" on the colorbar corresponds to a data point with a value at 10^{-6} of the peak value. The bright PSF's peak is centered, while the dim PSF's peak is the protrusion visible slightly above and to the left of the letter labeling each plot.

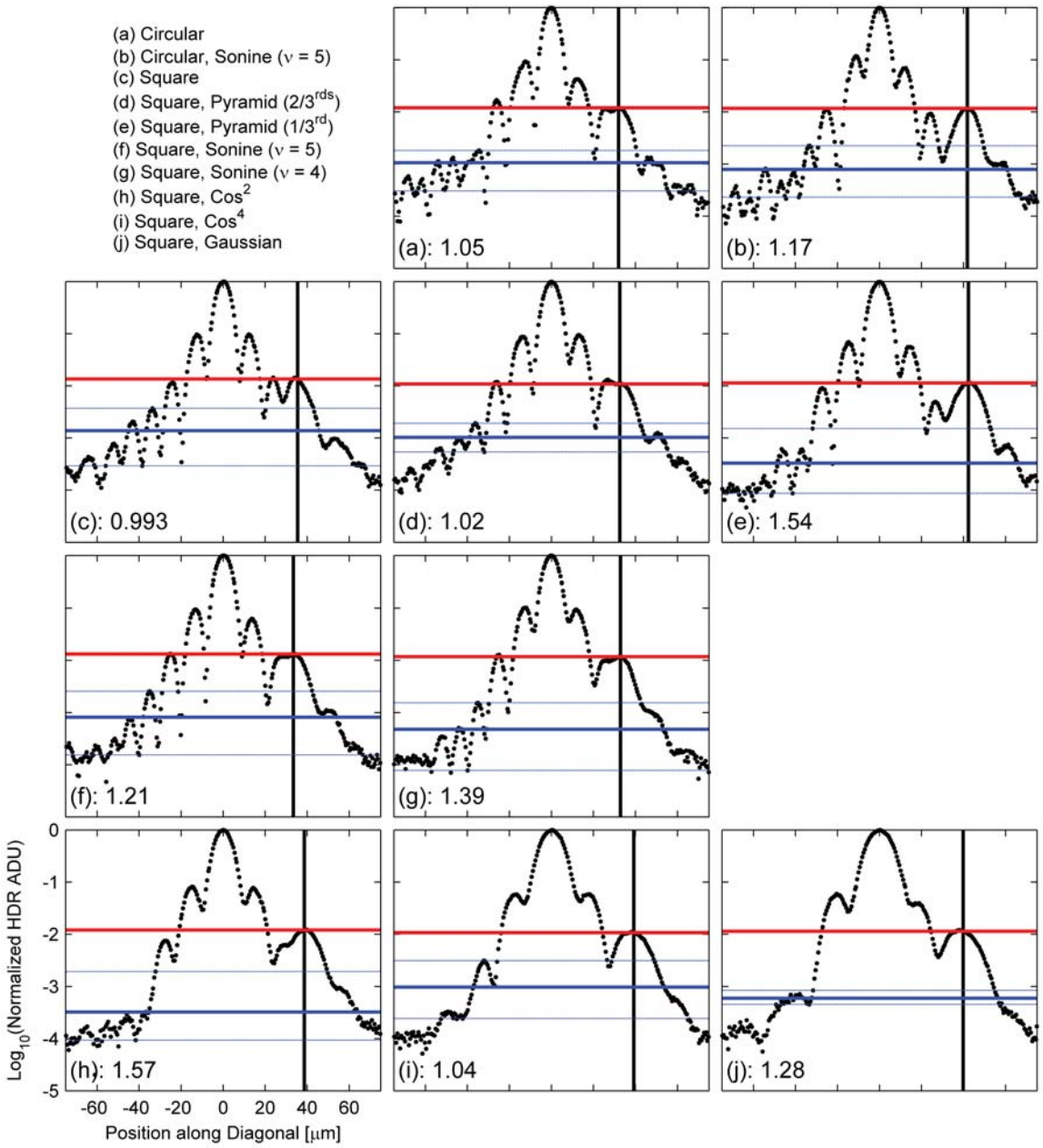


Figure 72. Log_{10} diagonal cross sections through HDR scaled-and-spliced images from the two-beam illumination test. The positions given are in the plane of the PSFs (the focal plane) and all plots are displayed at the same scale. The vertical black line indicates the position of the dim PSF's peak and the horizontal red line shows its value. The thick, horizontal blue line shows the average value of the displayed data on the bright peak's opposite side from 32-42 μm. The thin horizontal blue lines show the maximum and minimum values from that range. The number in the bottom corner of each plot is the separation between the red and thick blue lines in units of $\log_{10}$(normalized HDR ADU). Given the asymmetries in the data these numbers have large uncertainties.

Fig. 69 displays the utility of using a $\log_{10}$ scale to display these images, since at close to 10^{-2} of the bright peaks, the dim PSF peaks are not visible. It also displays the subtlety of the asymmetries in the data and the differences between the PSFs (both of which were exaggerated on the $\log_{10}$ scale). Fig. 69(a) is the only plot where part of the Airy pattern's second ring is observed (enlarging the plot may be necessary to see this). This supports the prediction that square, apodized photon sieves produce PSFs with suppressed outer lobes, at least near the diagonals. The central lobes are also observed to widen, in agreement with the results presented in Section 3.2.3 (p. 100) even though the HDR scaled-and-spliced images for the two-beam illumination test were collected separately.

Fig. 70 shows that the dim PSF peaks were placed approximately along the diagonals with the centers 32-42 μm from the bright PSF peak, with a mean of 36.8 μm and a standard deviation of 4 μm. The dim PSFs appear elongated horizontally due to vibrations in the pellicle beamsplitter, which were exacerbated by acoustic noise in the laboratory. It is suggested that any future projects making use of the two-beam illumination test strive to maintain an environment that is free of both stray light and acoustic noise. The vibration did not invalidate the results, however, since the ratio between the peak values of the dim and bright PSFs was held as constant as possible (though the dim PSF's shapes were impacted). In these images, the test optics appear to perform similarly.

Fig. 71 shows $\log_{10}$ colormapped, HDR scaled-and-spliced surface plots from the two-beam illumination test. The dim PSF's peak can be seen as the orange-yellow protrusion slightly above and to the left of the letters labeling each plot. The fact that the dim PSF peak is equal with or above the second "ring" of each bright PSF shows that the photon sieves performed well with such a small peak-to-peak separation (roughly 5 resels). However, no clear differences in performance between the sieves

are visible.

Fig. 72 shows $\log_{10}$ diagonal cross sections through HDR scaled-and-spliced images from the two-beam illumination test. Since the dim PSF peaks were not perfectly positioned along the diagonal, Bresenham's line algorithm was used to select the desired cross section (passing through both the bright and dim peaks) from each image [45]. The vertical black line in each plot indicates the position of the dim PSF's peak, with the bright peak centered. The horizontal red line shows the value of the dim PSF's peak, while the horizontal thick blue line shows the average value of the displayed data on the opposite side of the bright peak from 32-42 μm. The horizontal thin blue lines show the maximum and minimum values from that range, also on the side opposite the dim PSF peak. Data are presented in this manner (using the average over a range as the baseline comparison) because it was not possible to keep the separation between the dim and bright PSFs exactly constant. Comparing each dim PSF to its exact separation on the opposite side of the bright peak would not have been a fair comparison when the peak-to-peak separation varied by as much as 30% between the photon sieves. Along the diagonal, it was expected that the square sieve would noticeably outperform the circular sieve. Using the thick blue line as the baseline for comparison, the figure surprisingly shows that the performance of the square and circular photon sieves was roughly equal. The apodized sieves largely outperformed the unapodized sieves, with the square Cos^2, square pyramid ($1/3^{\text{rd}}$), and square Sonine ($\nu = 4$) sieves being the top performers. These results were in rough agreement with the results in Fig. 67(p. 110) where these three sieves placed fifth, third, and fourth out of eleven for overall high-contrast imaging performance near the diagonal. The values in the lower left corners of each plot give the separation between the red and thick blue lines (on the $\log_{10}$ scale) and show the square Cos^2 sieve outperforming the circular sieve by 50% on that scale.

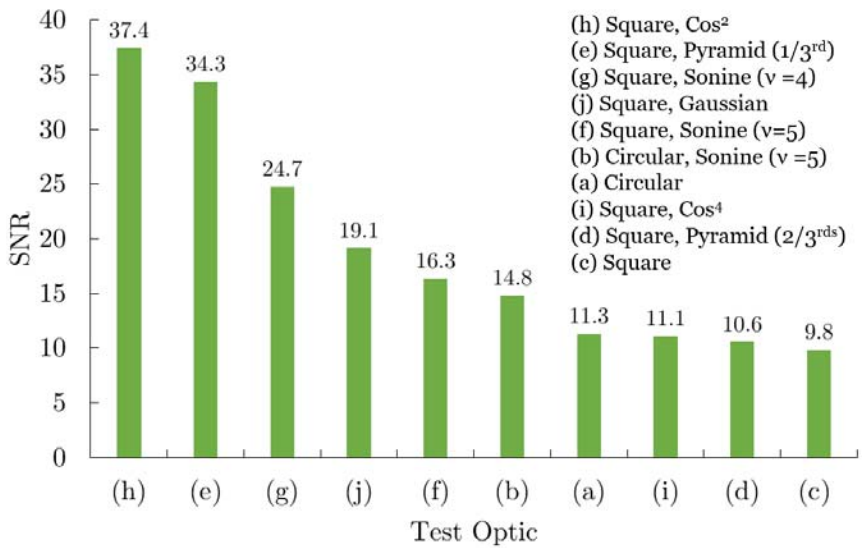


Figure 73. SNRs from the two-beam illumination test. These values were computed by taking the ratio between the dim PSF's peak value and the average over the 32-42 μm range on the bright peak's opposite side. The optics are sorted in order of decreasing performance.

The poor performance of the square pyramid ($2/3^{rds}$) sieve was unexpected after it displayed the best performance near the diagonal in the semi-empirical analysis (Section 3.3.1(p. 107)). The good results for this sieve in that section could be attributed to the fact that the PSF was much cleaner and better sampled, as can be seen by comparing Fig. 72(d) to Fig. 55(d). On a linear scale, performance was measured similar to the SNR shown in Eq. 21(p. 107) by taking the ratio between the dim PSF's peak value and the average over the indicated range on the bright peak's opposite side. By this measure, the square Cos^2 and square pyramid ($1/3^{rd}$) sieves outperformed the circular sieve by more than three times, as shown in Fig. 73.

Two-Beam Illumination Test: $10^{-3.5}$. A subset of the photon sieves was selected for a more rigorous test. ND3 (labeled in Fig. 68(p. 112)) was rotated to adjust the value of the dim PSF's peak until it was below $10^{-3.5}$ of the bright peak. The peak-to-peak separation was also adjusted to an average of 63.6 μm (8.6 resels

for the square sieve). The full list of peak-to-peak separations and PVRs is provided in Table 5. The long exposures were individual 100-s exposures rather than sets of three 20-s exposures. All other procedures were kept the same as for the 10^{-2} tests. Even with the increased peak-to-peak separation, this test placed the dim PSF at the very limit of detectability. The test results are presented as colormapped $\log_{10}$ images in Fig. 74 and as diagonal cross sections through the PSF peaks of the $\log_{10}$ data in Fig. 75.

Table 5. Two-Beam Illumination Test: Peak Separations and Peak Value Ratios ($10^{-3.5}$ test)

Photon Sieve	Peak-to-Peak Separation [μm]	Peak Value Ratio
Circular	69.8	$10^{-3.66}$
Square	59.9	$10^{-3.74}$
Square, Pyramid (1/3rd)	59.4	$10^{-3.79}$
Square, Sonine ($\nu = 4$)	68.4	$10^{-3.85}$
Square, Cos^2	60.5	$10^{-3.69}$

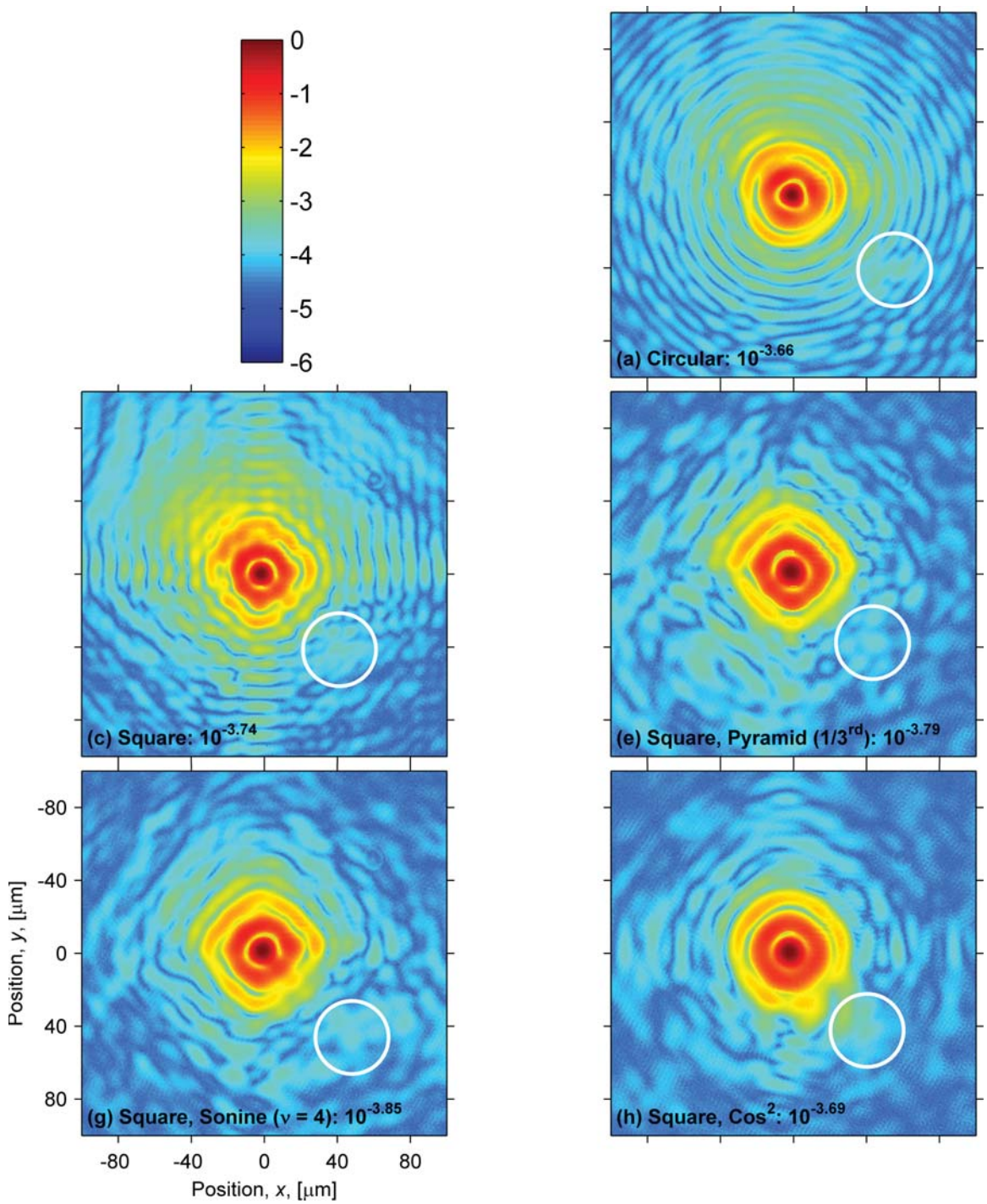


Figure 74. Log_{10} colormapped, HDR scaled-and-spliced images from the two-beam $10^{-3.5}$ illumination test. The positions given are in the plane of the PSFs (the focal plane) and all images are displayed at the same scale. The labels of the colorbar denote the value of the exponent, e.g., "-6" on the colorbar corresponds to a data point with a value at 10^{-6} of the peak value. The dim PSF's peak is shown circled in white in each image.

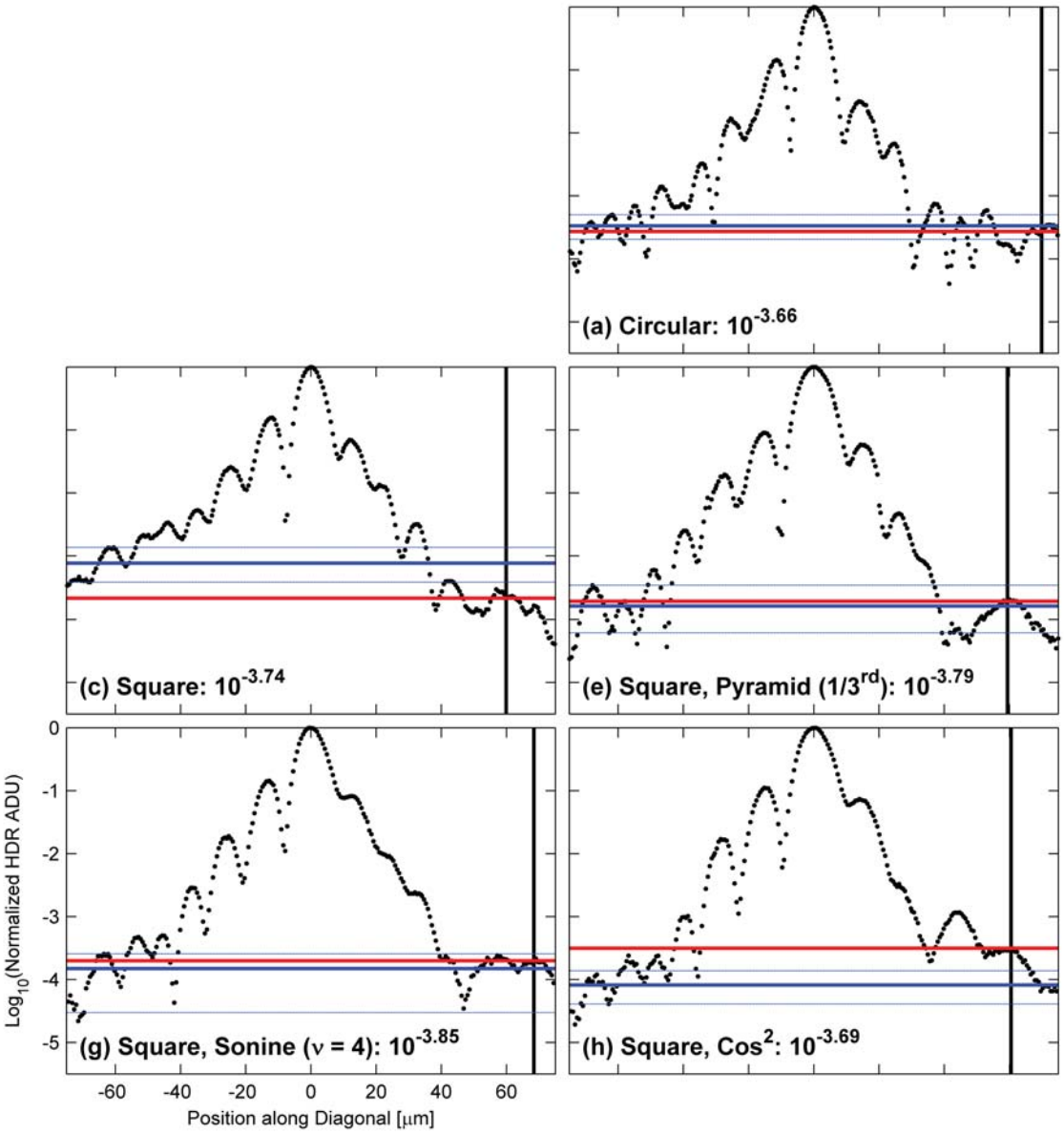


Figure 75. Log_{10} diagonal cross sections through HDR scaled-and-spliced images from the two-beam $10^{-3.5}$ illumination test. The positions given are in the plane of the PSFs (the focal plane) and all plots are displayed at the same scale. The vertical black line indicates the position of the dim PSF's peak and the horizontal red line shows its value. The horizontal thick blue line shows the average value of the displayed data on the bright peak's opposite side from 32-42 μm. The horizontal thin blue lines show the maximum and minimum values from that range. The number in the bottom corner of each plot is the separation between the red and thick blue lines.

Fig. 74 shows $\log_{10}$ colormapped, HDR scaled-and-spliced images produced by the subset of photon sieves. The dim PSF peaks can be seen placed approximately along the diagonals with the centers 59.4-69.8 μm from the bright PSF peak, with a mean of 63.6 μm and a standard deviation of 5 μm. Interpretation of these images is somewhat subjective, though it is clear that for the unapodized sieve, a detection would be extremely unlikely. For the square pyramid ($1/3^{\text{rd}}$) and square Sonine ($\nu = 4$) sieves, Fig. 74(e) and Fig. 74(g) show that a detection *may* be possible.

Fig. 75 shows $\log_{10}$ diagonal cross sections through HDR scaled-and-spliced images from the two-beam $10^{-3.5}$ illumination test. Since the dim PSF peaks were not perfectly positioned along the diagonal, Bresenham's line algorithm was once again used to select the desired cross section (passing through both the bright and dim peaks) from each image [45]. The plots are presented exactly the same way as in Fig. 72(p. 118) except for the fact that the numbers given with the label on each plot are the PVRs, rather than the separation between the red and thick blue lines. The vertical black line in each plot indicates the position of the dim PSF's peak, with the bright peak centered. The horizontal red line shows the value of the dim PSF's peak, while the horizontal thick blue line shows the average value of the displayed data on the opposite side of the bright peak from 59.4-69.8 μm. The horizontal thin blue lines show the maximum and minimum values from that range, also on the side opposite the dim PSF peak. For the unapodized sieves, the dim PSF's peak value was actually less than the opposite side average, as is shown by the thick blue lines being above the red lines. This can be partially attributed to asymmetries in the data, and the square sieve's PSF is noticeably lopsided. It should be noted that the circular sieve had the advantage of having the largest PVR (meaning the brightest dim PSF) and the square sieve had the third largest. Additionally, the circular sieve had the

advantage of having the greatest peak-to-peak separation (more than one standard deviation above the average). The square Cos^2, square pyramid ($1/3^{\text{rd}}$), and square Sonine ($\nu = 4$) sieves all produced images with the dim PSF's peak value greater than the opposite side average. The performance of the Cos^2 sieve was exaggerated by the fact that it had the second largest PVR at $10^{-3.69}$, compared to $10^{-3.79}$ and $10^{-3.85}$ for the square pyramid ($1/3^{\text{rd}}$) and square Sonine ($\nu = 4$) sieves, respectively. Even taking this into account, the Cos^2 performed excellently and was the only sieve with an opposite side average below 10^{-4}. When looking purely at the shape of the cross sections, the protrusion produced by the dim PSF appears to stand out most significantly for the square pyramid ($1/3^{\text{rd}}$) sieve, though that could be due to fortuitous placement of the dim PSF's peak directly on top of a lobe.

These results were in rough agreement with the semi-empirical analysis results shown in Fig. 67(p. 110) and the two-beam illumination 10^{-2} test results shown in Fig. 73(p. 121). The only clear conclusion that can be drawn is that the apodized sieves outperformed the unapodized sieves. Additionally, it is reasonable to rule out the square Gaussian, square Cos^4, and circular Sonine ($\nu = 5$) sieves as top performers among the apodized sieves considering that they were not among the top three performers in either Figure 67(p. 110) or 73(p. 121).

IV. Conclusion

Direct imaging of exoplanets is undeniably worthwhile due to how well it could complement the currently successful transit and radial velocity detection methods. However, it is a challenging endeavor, due to the minuscule angular separation and peak value irradiance ratio between an exoplanet and its parent star. Building high aperture space telescopes with PSFs optimized for high-contrast imaging is one potential way to overcome these challenges. Since they can be printed onto a lightweight and flexible membrane, photon sieves are well-suited to be a deployable primary optic for a large aperture space telescope. Additionally, photon sieves may be shaped and apodized to optimize for high contrast imaging. Because of this ten photon sieves (eight square, two circular) with varying apodizations were simulated, designed, and tested for high-contrast imaging performance and use in an exoplanet imaging telescope.

A numerical Fresnel simulation was performed to better understand how shaping and apodizing a photon sieve might impact the PSF. A method was devised for generating the simulated aperture and applying any apodization, while minimizing the computational cost. Both square and circular photon sieves were simulated, along with a number of different apodizations. Differing ratios between pinhole diameter and Fresnel zone width (d/w), as well as different incident wavelengths were also tested. The simulation produced promising results with small d/w ratios and with wavelengths 1 nm from the sieve design wavelength, though these results were deemed outside the scope of this thesis and were not investigated further. Two PSFs with a small peak value ratio were simulated, slightly offset, and combined to imitate exoplanet imaging and better judge the performance of the simulated sieves. A flat topped, square pyramid apodization was generated, simulated and tested for the first time, to the author's knowledge. The results of the simulation suggested

that photon sieves produce PSFs with floors between 10^{-4} and 10^{-5} of the peak irradiance, regardless of shape or apodization. This is significantly different from Nisenson and Papaliolios's work, where simulations predicted shaping and apodizing an aperture could suppress the floor below 10^{-15} for conventional optics [34]. As shape and apodization did not appear to have an overwhelmingly large affect on a photon sieve's PSF, a subset of apodizations was selected for further investigation.

Following the numerical Fresnel simulation, eight square and two circular photon sieves were selected for design and experimental test. Unapodized square and circular sieves were selected for design and test to serve as controls throughout the experiment. The square apodized sieves selected for design and test had cosine, Gaussian, crossed Sonine, and flat top pyramid shaped apodizations. A crossed Sonine apodized, circular sieve was also selected as an additional benchmark for comparison. In order to design the sieves it was necessary to write a new algorithm for applying any apodization to a photon sieve. This was necessary because only the cosine and Gaussian apodizations were circularly symmetric (meaning the apodization could be applied simply by selectively altering the inter-pinhole spacing in each Fresnel zone). For the non-circularly symmetric apodizations, the algorithm split the unapodized sieve into a 250×250 grid and randomly removed the appropriate number of pinholes from each grid square. As a result, the transmittance of each grid square matched the apodization with a maximum residual of only 0.015, or 1.5%. The photon sieves had between 300,000 and 4.3 million pinholes, with minimum pinhole diameter equal to 8 μm and maximum pinhole diameter equal to 279 μm. The square and circular sieves had equal surface area, with widths of 3.59 cm and 4.05 cm, respectively. Depending on the apodization, the throughput was between 19% and 81% relative to the unapodized square sieve, and the sieve-generating code required between 20 s and 10,000 s to run.

The sieves were characterized with a variety of tests. Resolution target imaging

showed that apodizations could be applied without significantly worsening resolution. Four of the square, apodized sieves resolved element 7-2 and the square, Sonine ($\nu = 5$) sieve resolved element 7-3. This means that five of seven square apodized sieves resolved bar separations of 6.96 μm or smaller (compared to 6.20 μm for the unapodized sieves and 7.41 μm for the Rayleigh criterion resolution). A detailed PSF analysis was carried out using a novel HDR scale-and-splice imaging method, which provided the PSF shape to below 10^{-4} of the peak irradiance, and simulated and experimental results were compared and shown to be in general agreement. Residuals between simulated and experimental PSFs were found to have an average of 0.008 and a standard deviation of 0.02, relative to a peak of one. The maximum residual was less than 0.12. This demonstrates that this method of simulation will be useful in future work. This experiment determined that if using a numerical Fresnel simulation, a photon sieve is best approximated using the central Fresnel zones with an equal focal length (this thesis used only about 50 light Fresnel zones, compared to roughly 1,200 in the experimental sieves). The central portion method performed similarly to the equal $f/\#$ method (where a sieve is approximated using a small diameter sieve with an $f/\#$ equal to the experimental sieve) and was more than thirty times faster to run. The PSF analysis confirmed that the PSF floor was between 10^{-4} and 10^{-5} of the peak irradiance, regardless of shape or apodization. A focal spot size measurement was conducted using the data from the PSF analysis and the results matched nearly perfectly with the resolution target test. The focal spot size test supported the conclusion from the resolution target test that photon sieves could be apodized without significantly degrading the resolution, with five of the eight apodized sieves producing central lobes less than one-third wider than their unapodized counterparts. An interferometric analysis was conducted which set a lower bound on the square sieve performance with WFE of peak to valley = 0.78λ and rms = 0.18λ. By comparing

the interferometric results of a control lens with a Zemax analysis, up to 85% of the WFE was determined to be a result of other optical elements in the setup, along with misalignment. Subtracting off this 85%, the square sieve gave a best-case scenario WFE of peak to valley = 0.12λ and rms = 0.03λ, which met Rayleigh's quarter wavelength rule and Maréchal's condition for low-aberration imaging. This demonstrated that photon sieves can be square shaped while maintaining wavefront quality.

Next, the sieves were tested for high-contrast imaging potential. A novel semi-empirical analysis was conducted using the PSF data, which provided a significantly more thorough characterization of sieve performance than a single best-case measurement. The semi-empirical analysis showed that the apodized sieves generally outperformed the control lens and unapodized sieves for high-contrast imaging, though asymmetries present in the PSF data prevented concrete conclusions from being drawn regarding exactly which sieve was the top performer. Surprisingly, the unapodized circular sieve outperformed the unapodized square sieve, though the difference between the two sets of results was not statistically significant. A two-beam illumination test was then conducted in which the Cos^2 and square pyramid ($1/3^{\mathrm{rd}}$) sieves produced simulated exoplanet SNRs more than three times higher than those of the unapodized circular sieve, though due to the asymmetries in the data this was not fully consistent with the results of the semi-empirical analysis. A laboratory-simulated exoplanet was shown to be possibly detectable using apodized square sieves with a peak value ratio below $10^{-3.69}$ and a peak-to-peak separation less than ten times the Rayleigh resolution, or ten resels.

Ultimately, square apodized sieves were shown to generally outperform the control lens and the unapodized sieves for high-contrast imaging, though not to the extreme degree the literature suggested was possible with conventional optics [34]. The observed improvement in the PSFs of the square, apodized sieves was not sufficient

for direct imaging of earth-like exoplanets, but will be useful for other high-contrast applications given that the PSF falls to 10^{-4} of the peak irradiance within ten resels. Since efficient methods were developed for designing and simulating apodized sieves, future work could improve upon and use these methods to optimize photon sieve apodization for any number of purposes. New shapes and apodizations could be investigated, for example a $1 - \cos(\xi, \eta)^2$ apodization that has high throughput at the edges of the sieve. Low d/w ratio photon sieves produced promising PSFs in the simulation, and should be further investigated. It is also recommended that future work investigate the possibility of improving upon the numerical Fresnel simulation by making fewer approximations to the Rayleigh-Sommerfeld diffraction integral. If using a numerical Fresnel simulation, the central portion method should be used. A supercomputer such as the AFIT Linux cluster could one day be used to help select the best photon sieve design for a future FalconSAT mission.

In summary, this work introduced new techniques for design of photon sieves and the characterization of high-contrast optical system performance. The method for applying non-circularly symmetric apodizations is applicable to all photon sieves. The HDR scale-and-splice method was successful in significantly extending the dynamic range of the detector. The semi-empirical analysis provides a more thorough and fairer assessment of high-contrast performance than the best-case scenarios typically presented in the literature.

Appendix A. Square Photon Sieve: Circularly Symmetric Apodization (or Controlled Unapodized)

```
% Code is adapted from code written by Andersen and Tulip
clc, clear all, close all

%% Select Photon sieve parameters
lambda=532.1E-9; %Sieve deisng wavelength [m]
f=0.5; %Focal length [m]
svdiam=0.0508; %Desired sieve diameter rounded up to capture all light [m]
square_sidelength=svdiam/sqrt(2); %desired square side length [m]
minholesep=1.5e-6; %Manufacturing limitation, min edge-edge hole separation [m]
reserror=0.25e-6; %error in feature size [m]
zone_split=989; %select the final zone to be in the first gerber file

%% Calculate additional information from user input
fnum=f/svdiam; %f/#
rmax=svdiam/2; %Approx. max radius [m] (from desired sieve diameter)
nmax=2*floor(((-f+sqrt(f^2+rmax^2))/lambda)/2);
%Outermost zone # (rounded down to nearest EVEN integer)
rmax=sqrt(2*nmax*lambda*f+(nmax^2)*(lambda^2)); %Max radius [m]
tic;

%% Position the pinholes
%initialize variables
holes_per_zone=zeros(1,nmax); %how many pinholes in each zone
zone_area=zeros(1,nmax); %area of each zone [m^2]
fill_factor=zeros(1,nmax); %ratio of total hole area to zone area for each zone
rn=zeros(1,nmax); %initialize zone radius
pf=zeros(1,nmax); %initialize the pad fraction for apodization
%initialize for vectors of all hole positions and diameters [m]
```

```
xcirc=0;
ycirc=0;
dcirc=0;
%These large vectors aren't ideal for efficiency, but are helpful for
%applying apodizations and are used as inputs to build the source plane for
%the numerical Fresnel simulation.

for j=1:nmax
    phi=2*pi*random('unif',1); %Random starting angle for each zone
    n=j; %set zone number to index number for loop
    rn(j)=sqrt(2*n*lambda*f+(n^2)*(lambda^2)); %center of zone "n"
    % use ff=1 for controlled unapodized
    ff=cos(.5*pi*rn(j)/rmax)^2; %desired fill factor from apodization
    w=lambda*f/(2*rn(j)); %width of that zone [m]
    d=1.53*w; %diameter of pinholes in zone [m]
    C=asin(d/2/rn(j)); %half angle of a hole in that zone [rad]
    zone_area(j)=pi*(rn(j)+(w/2))^2-pi*(rn(j)-(w/2))^2; %area of each zone [m^2]
    pf(j)=1-(minholesep/(d+reserror+minholesep))-...
        ((C*ff*zone_area(j))/(pi^2*(d/2)^2)); %pad fraction required to achieve desired
    scalefactor=1-(minholesep/(d+reserror+minholesep))-pf(j);
    %fraction of max # of holes per zone minus 5% buffer (for radial spacing)
    holes_per_zone(j)=floor(scalefactor*(2*pi/(2*C)));
    %max number of angular holes scaled
    holegap=(2*pi-holes_per_zone(j)*2*C)/holes_per_zone(j);
    %angular gap between holes [rad]
    fill_factor(j)=(holes_per_zone(j)*pi*(d/2)^2)/(zone_area(j));
    %achieved fill factor
    x=zeros(1,holes_per_zone(j)); %initialize for speed
    y=zeros(1,holes_per_zone(j)); %initialize for speed
        for i=1:holes_per_zone(j) %place each hole at an x,y coordinate
            x(i)=rn(j)*cos(phi); %x coordinate [m]
            y(i)=rn(j)*sin(phi); %y coordinate [m]
```

```
            phi=phi+2*C+holegap; %reset pinhole center
        end
    xcirc=cat(2,xcirc,x); %vector of all x coordinates [m]
    ycirc=cat(2,ycirc,y); %vector of all y coordinates [m]
    dcirc=cat(2,dcirc,d*ones(1,length(x))); %vector of all pinhole diams [m]
end

%get rid of 0 0 0 values (this was the easiest way to use "cat" command above)
xcirc=xcirc(2:end); %hole x coord [m]
ycirc=ycirc(2:end); %hole y coord [m]
dcirc=dcirc(2:end); %hole diameter [m]

%combine the data for further manipulation
circ_data=[xcirc;ycirc;dcirc];

%% Make it a Square
%cut out the elements that don't fit in the square
circ_data(:,(abs(circ_data(1,:)) > .5*square_sidelength)) = [];
circ_data(:,(abs(circ_data(2,:)) > .5*square_sidelength)) = [];

%% Simulation Input Variables
%save x positions, y positions, and diameters in single vectors with units
%all in meters in order to use as input for simulations
xcirc=circ_data(1,:);
ycirc=circ_data(2,:);
dcirc=circ_data(3,:);
%finds the halfwidth of the square
sieve_halfwidth=.5*(max(xcirc)-min(xcirc))+dcirc(xcirc==max(xcirc));

%save a file for use in the simulation
save('square_sieve_cos2_apod.mat','xcirc','ycirc','dcirc','sieve_halfwidth')
toc;
```

```
tot_hole_area=sum(pi*(dcirc/2).^2); %total hole area [m^2]
sieve_tot_area=(2*sieve_halfwidth)^2; %total sieve area [m^2]
sieve_trans=tot_hole_area/sieve_tot_area; %percent of sieve covered in holes

%% Convert data to form for Gerber file
% sort back into rings (needed for square sieve, not circle)
tot_holes=length(xcirc); %total number of holes
diameter_list=sort(unique(dcirc),'descend'); %list of diameters [m]
tot_zones=length(diameter_list); %total number of zones
num_holes=zeros(1,length(diameter_list)); %initialize num_holes
hc=1; %initialize
zonecounter=1; %initialize
for i=2:tot_holes; %counts up number of holes in each zone
    if dcirc(i)==dcirc(i-1);
        hc=hc+1;
    else num_holes(zonecounter)=hc;
        zonecounter=zonecounter+1;
        hc=1;
    end
end
num_holes(zonecounter)=hc;

%% Open the FIRST Gerber File for editing and write header/parameters applicable to the
% This file generates the inner zones
GerberFile1=fopen('C:\Directory\square_sieve_cos2_apod_1.gbr','w+');
fprintf(GerberFile1,'G75*\r\n'); %Set Multi Quadrant Mode; \r for Notepad viewing
fprintf(GerberFile1,'%%IPPOS*%%\r\n'); %Set image polarity to "positive"
fprintf(GerberFile1,'%%FSLAX35Y35*%%\r\n');
%Set coordinate Format Specification to omit Leading zeros, use Absolute notation,
%set number of integer positions in X,Y coordinates to 3 and decimal positions to 5
fprintf(GerberFile1,'%%LPD*%%\r\n'); %Set layer polarity to "dark"
```

```
fprintf(GerberFile1,'%%MOMM*%%\r\n'); %Set units to [mm]
fprintf(GerberFile1,'G75*\r\n'); %Set Multi Quadrant Mode; \r for Notepad viewing

%% Write Data to FIRST Gerber File
gbr_diam=1000*diameter_list; %hole diameters convert to [mm]
gbr_x=1e8*xcirc; %x position convert to [units of 10 nm]
gbr_y=1e8*ycirc; %y position convert to [units of 10 nm]

%make a list of "tools" from each aperture size
for i=1:zone_split;
    fprintf(GerberFile1,'%%ADD%dC,%.5f*%%\r\n',i+10,gbr_diam(i));
    %creates a numbered aperture for each circle
    %add 10 so there are no negative or 0 dimension tools
end

%select the tool and place the aperture
hc=0; %initialize the holecounter
for i=1:zone_split;
    fprintf(GerberFile1,'G54D%d*\r\n',i+10); %identifies which aperture to place
    %add 10 so there are no negative or 0 dimension tools
    for j=1:num_holes(i);
        fprintf(GerberFile1,'X%.0fY%.0fD03*\r\n',gbr_x(j+hc),gbr_y(j+hc));
        %places the circle at an coord (x,y)
    end
    hc=hc+num_holes(i);
end

%% Terminate and close FIRST Gerber File
fprintf(GerberFile1,'M02*'); %End Gerber File
fclose(GerberFile1);
toc;
```

```
%% Open the SECOND Gerber File for editing and write header/parameters applicable to the
% This file generates the outer zones
GerberFile2=fopen('C:\Directory\square_sieve_cos2_apod_2.gbr','w+');
fprintf(GerberFile2,'G75*\r\n'); %Set Multi Quadrant Mode; \r for Notepad viewing
fprintf(GerberFile2,'%%IPPOS*%%\r\n'); %Set image polarity to "positive"
fprintf(GerberFile2,'%%FSLAX35Y35*%%\r\n'); %Set coordinate Format Specification to omit
fprintf(GerberFile2,'%%LPD*%%\r\n'); %Set layer polarity to "dark"
fprintf(GerberFile2,'%%MOMM*%%\r\n'); %Set units to [mm]
fprintf(GerberFile2,'G75*\r\n'); %Set Multi Quadrant Mode; \r for Notepad viewing

%% Write Data to SECOND Gerber File

%make a list of "tools" from each aperture size
for i=zone_split+1:tot_zones;
    fprintf(GerberFile2,'%%ADD%dC,%.5f*%%\r\n',i+10-zone_split,gbr_diam(i));
    %creates a numbered aperture for each circle
    %add 10 so there are no negative or 0 dimension tools
end

%select the tool and place the aperture
for i=zone_split+1:tot_zones;
    fprintf(GerberFile2,'G54D%d*\r\n',i+10-zone_split);
    %identifies which aperture to place
    %add 10 so there are no negative or 0 dimension tools
    for j=1:num_holes(i);
        fprintf(GerberFile2,'X%.0fY%.0fD03*\r\n',gbr_x(j+hc),gbr_y(j+hc));
        %places the circle at an coord (x,y)
    end
    hc=hc+num_holes(i);
end

%% Terminate and close SECOND Gerber File
```

```
fprintf(GerberFile2,'M02*'); %End Gerber File
fclose(GerberFile2);
toc;

%% Manufacturing check
min_d=min(dcirc); %minimum hole size
max_d=max(dcirc); %max hole size
disp(['min. hole size = ' num2str(1e6*min_d) ' um'])
Del_cc_end=((xcirc(end)-xcirc(end-1))^2+(ycirc(end)-ycirc(end-1))^2)^.5;
%center-center sep.
Del_ee_end=Del_cc_end-.5*dcirc(end)-.5*dcirc(end-1); %edge-edge sep.
disp(['edge-edge separation at smallest holes = ' num2str(1e6*Del_ee_end) 'um'])
holechoice=1000; %choose a hole at which to find adjacent separation
Del_cc=((xcirc(holechoice)-xcirc(holechoice-1))^2+(ycirc(holechoice)-...
    ycirc(holechoice-1))^2)^.5; %center-center sep.
Del_ee=Del_cc-.5*dcirc(holechoice)-.5*dcirc(holechoice-1); %edge-edge sep.
disp(['edge-edge separation at chosen holes = ' num2str(1e6*Del_ee) 'um'])

%% Output useful information to a separate text file
InfoFile=fopen('C:\Directory\square_sieve_cos2_apod.txt','w+');
fprintf(InfoFile,'Square Sieve, cos^2 Apod. Specifications\r\n');
fprintf(InfoFile,'Incident Wavelength: %.1f [nm]\r\n',lambda*1e9);
fprintf(InfoFile,'Focal Length: %.2f [m]\r\n',f);
fprintf(InfoFile,'Target Width: %.4f [m]\r\n',svdiam);
fprintf(InfoFile,'Actual Width: %.4f [m]\r\n',2*sieve_halfwidth);
fprintf(InfoFile,'F Number: %.2f\r\n',fnum);
fprintf(InfoFile,'Total Rings: %.1f\r\n',tot_zones);
fprintf(InfoFile,'Total Holes: %.1f\r\n',tot_holes);
fprintf(InfoFile,'Max Hole Size: %.2f [um]\r\n',max_d*1e6);
fprintf(InfoFile,'Min Hole Size: %.2f [um]\r\n',min_d*1e6);
fprintf(InfoFile,'Min Zone Size: %.2f [um]\r\n',min_d/(1.53)*1e6);
fprintf(InfoFile,'nmax Hole Separation: %.2f [um]\r\n',Del_ee_end*1e6);
```

```
fprintf(InfoFile,'d/w: %.2f\r\n',1.53);
fprintf(InfoFile,'Code runtime: %.3f [min]\r\n',toc/60);
fprintf(InfoFile,'Total Hole Area: %.3f [mm^2]\r\n',1e6*tot_hole_area);
fprintf(InfoFile,'Fraction Area Covered: %.3f%\r\n',sieve_trans);
fclose(InfoFile);
```

Appendix B. Square Photon Sieve: Non-Circularly Symmetric Apodization

```
% Code is adapted from code written by Andersen and Tulip
clc, clear all, close all

%% Select Photon sieve parameters
% Same as circ. symmetric sieve code

%% Calculate additional information from user input
% Same as circ. symmetric sieve code

%% Position the pinholes
%initialize variables
holes_per_zone=zeros(1,nmax); %how many pinholes in each zone
zone_area=zeros(1,nmax); %area of each zone [m^2]
fill_factor=zeros(1,nmax); %ratio of total hole area to zone area for each zone
rn=zeros(1,nmax); %initialize zone radius
pf=zeros(1,nmax); %initialize the pad fraction for apodization
%initialize for vectors of all hole positions and diameters [m]
xcirc=0;
ycirc=0;
dcirc=0;
%These large vectors aren't ideal for efficiency, but are helpful for
%applying apodizations and are used as inputs to build the source plane for
%the numerical Fresnel simulation.

ff=1.00; %desired fill factor (removed from loop since constant)
for j=1:nmax
    phi=2*pi*random('unif',1); %Random starting angle for each zone
    n=j; %set zone number to index number for loop
```

```
    rn(j)=sqrt(2*n*lambda*f+(n^2)*(lambda^2)); %center of zone "n"
    w=lambda*f/(2*rn(j)); %width of that zone [m]
    d=1.53*w; %diameter of pinholes in zone [m]
    C=asin(d/2/rn(j)); %half angle of a hole in that zone [rad]
    zone_area(j)=pi*(rn(j)+(w/2))^2-pi*(rn(j)-(w/2))^2;
    %area of each zone [m^2]
    pf(j)=1-(minholesep/(d+reserror+minholesep))-...
        ((C*ff*zone_area(j))/(pi^2*(d/2)^2));
    %pad fraction required to achieve desired fill_factor
    scalefactor=1-(minholesep/(d+reserror+minholesep))-pf(j);
    %fraction of max # of holes per zone minus 5% buffer (for radial spacing)
    holes_per_zone(j)=floor(scalefactor*(2*pi/(2*C)));
    %max number of angular holes scaled
    holegap=(2*pi-holes_per_zone(j)*2*C)/holes_per_zone(j);
    %angular gap between holes [rad]
    fill_factor(j)=(holes_per_zone(j)*pi*(d/2)^2)/(zone_area(j));
    %achieved fill factor
    x=zeros(1,holes_per_zone(j)); %initialize for speed
    y=zeros(1,holes_per_zone(j)); %initialize for speed
        for i=1:holes_per_zone(j) %place each hole at an x,y coordinate
            x(i)=rn(j)*cos(phi); %x coordinate [m]
            y(i)=rn(j)*sin(phi); %y coordinate [m]
            phi=phi+2*C+holegap; %reset pinhole center
        end
    xcirc=cat(2,xcirc,x); %vector of all x coordinates [m]
    ycirc=cat(2,ycirc,y); %vector of all y coordinates [m]
    dcirc=cat(2,dcirc,d*ones(1,length(x))); %vector of all pinhole diams [m]
end

%get rid of 0 0 0 values (this was the easiest way to use "cat" command above)
xcirc=xcirc(2:end); %hole x coord [m]
ycirc=ycirc(2:end); %hole y coord [m]
```

```
dcirc=dcirc(2:end); %hole diameter [m]

%combine the data for further manipulation
circ_data=[xcirc;ycirc;dcirc];

%% Make it a Square
%cut out the elements that don't fit in the square
circ_data(:,(abs(circ_data(1,:)) > .5*square_sidelength)) = [];
circ_data(:,(abs(circ_data(2,:)) > .5*square_sidelength)) = [];

%% Apply the Apodization
%do everything using circ_data, since we have to keep things properly
%organized
xcirc=circ_data(1,:);
ycirc=circ_data(2,:);
dcirc=circ_data(3,:);

%find the bounds and center points of the grid and create grid vectors
grid_n=250+1; %how fine is the grid
x1=min(xcirc)-.5*dcirc(xcirc==min(xcirc));
x2=max(xcirc)+.5*dcirc(xcirc==max(xcirc));
y1=min(ycirc)-.5*dcirc(ycirc==min(ycirc));
y2=max(ycirc)+.5*dcirc(ycirc==max(ycirc));
xgv=linspace(x1,x2,grid_n); %x grid vector
ygv=linspace(y1,y2,grid_n); %y grid vector
xgvc=0.5*(xgv(1:end-1) + xgv(2:end)); %center points of x grid vector
ygvc=0.5*(ygv(1:end-1) + ygv(2:end)); %center points of y grid vector
[Xgc, Ygc]=meshgrid(xgvc,ygvc); %x and y grid center mesh

%break the sieve into a grid, count how many holes in each grid square
histdata=[xcirc', ycirc'];
histcnt=hist3(histdata,{xgvc' ygvc'}); %rows increase with x coord
```

```
                                        %columns increase with y coord
% replace the lines solving for T below with these to use any apodization
% function you want.
% nu=4;
% The apodization function. Put whatever function you want here.
% T=((1-((Xgc/max(abs(x1),x2)).^2).^(nu-1)).*...
%    ((1-((Ygc/max(abs(x1),x2)).^2).^(nu-1))));

%find the value of the apod function at center of each grid square
T=zeros(grid_n-1); %initialize
flattop_area=1/3;
flattop_fraction=sqrt(flattop_area);
flattop=round(flattop_fraction*((grid_n-1)/2));
rampvec=linspace(0,1,((grid_n-1)/2)-flattop);
for i=1:((grid_n-1)/2)-flattop;
    Ttemp=rampvec(i)*ones((grid_n-1)-2*i);
    T(i+1:(grid_n-1)-i,i+1:(grid_n-1)-i)=Ttemp;
end
for i=((grid_n-1)/2)-(flattop-1):(grid_n-1)/2
    Ttemp=ones((grid_n-1)-2*i);
    T(i+1:(grid_n-1)-i,i+1:(grid_n-1)-i)=Ttemp;
end
%rows increase with x coord and columns increase with y coord
%find closest kept/total hole ratio to the value of that function for each
%grid square
holeskept=round(T.*histcnt);
holesremoved=histcnt-holeskept;

%randomly remove appropriate number of holes from each grid square
for xcnt=1:grid_n-1;
    for ycnt=1:grid_n-1;
         circ_data_idx=find(circ_data(1,:) > xgv(xcnt) & circ_data(1,:) < ...
```

```
            xgv(xcnt+1) & circ_data(2,:) > ygv(ycnt) & circ_data(2,:)...
            < ygv(ycnt+1));
        circ_data_idx_s=datasample(circ_data_idx,holesremoved(xcnt,ycnt),...
            'Replace',false);
        circ_data(:,circ_data_idx_s)=[]; %remove it
    end
end

%% Simulation Input Variables
%save x positions, y positions, and diameters in single vectors with units
%all in meters in order to use as input for simulations
xcirc=circ_data(1,:);
ycirc=circ_data(2,:);
dcirc=circ_data(3,:);
%finds the halfwidth of the square
sieve_halfwidth=.5*(x2-x1);

%save a file for use in the simulation
save('square_sieve_pyramid33_apod.mat','xcirc','ycirc','dcirc','sieve_halfwidth')
toc;

tot_hole_area=sum(pi*(dcirc/2).^2); %total hole area [m^2]
sieve_tot_area=(2*sieve_halfwidth)^2; %total sieve area [m^2]
sieve_trans=tot_hole_area/sieve_tot_area; %percent of sieve covered in holes

%% Convert data to form for Gerber files, and write to Gerber files
% Same as circ. symmetric sieve code
```

Appendix C. Photon Sieve Aperture Builder and Diffraction

```
%% Numerical Fresnel Diffraction - Photon Sieve: 2 Sources
% Documentation - this code is adapted from Jason Schmidt's textbook:
% "Numerical Simulation of Optical Wave Propagation"

clc, clear all, close all

nameforplot='nu4_lam5321_square';

% parameters for different apodizations
nu=4; %nu for sonine apodization
sig=.5; %sigma for Gaussian apodization
cosval=4; %decide whether Cos^2 or Cos^4
flattop_area=.16; %ratio of pyramid flattop area to pyramid base area

%% Settings
Apodization =       'sonine'; %(sonine,cosine,gauss,squarepyra)
BuildAperture =     'square'; %(square,circle)
PlotSave =          'y'; %(y,n) do you want to save plots?
RandomWFE =         'y'; %(y,n) do you want to add 1/20th wave rms random WFE?
wvl = 532.1E-9; % optical wavelength [m]
k = 2*pi / wvl;
z = 0.500; % prop dist. [m]

%% Load the hole positions and diameters and display values
if 1==strcmp(BuildAperture,'square')
load('square_sieve_controlled_unapod');
rn=sqrt(2)*sieve_halfwidth;
resel=(wvl*z)/(2*sieve_halfwidth); %square resel [m]
elseif 1==strcmp(BuildAperture,'circle')
```

```
load('circ_sieve_controlled_unapod');
rn=sieve_halfwidth;
resel=(1.22*wvl*z)/(2*sieve_halfwidth); %circ resel [m]
end

disp('files loaded')
delta1goal = .2*min(dcirc); %grid resolution goal based on smallest hole size
disp(['Suggested Resolution: ' num2str(delta1goal) ' m'])

%% System Parameters
% to increase the resolution of the output, both N and L must be increased.
L_ap = 2*max(max(xcirc),max(ycirc))+min(dcirc) %fits aperture grid to sieve [m]
N_ap = L_ap/delta1goal; %estimates grid points for aperture
N_ap = 2.*ceil(N_ap/2) %round up to nearest even number for grid points for aperture
delta1 = L_ap/N_ap %calculate true grid resolution
padfactor = 6; %how much to pad aperture matrix PICK INTEGER!
N = padfactor*N_ap % number of total grid points per side (should be even)
L = N*delta1 % total size of the grid [m]
[x1, y1] = meshgrid((-N/2 : N/2-1) * delta1); %build the entire mesh

%% Build the Aperture
[x1ap, y1ap] = meshgrid((-N_ap/2 : N_ap/2-1) * delta1); %build aperture mesh
if 0==strcmp(BuildAperture,'n');
ap=zeros(N_ap); %initialize aperture matrix for speed
numcirc=length(xcirc); %initialize number of circles for speed
for counter=1:numcirc
% circ.m from Schmidt: Numerical Simulation of Optical Wave
% Propagation, 2010.
    hole=circ(x1ap-xcirc(counter),y1ap+ycirc(counter),dcirc(counter));
    ap = ap+hole; %adds each current hole to the aperture
end

```

```
ap_raw=ap; %saved without apodization or 0 padding
end

disp('Done building aperture')

%% Apodization, Intensity Ratio, and Matrix Padding
intensity_ratio=50; %set the intensity ratio

% Add 1/20th wave rms random WFE
% change the 0.05 in random_complex to whatever you want to set the rms WFE
if 1==strcmp(RandomWFE,'y')
ap_raw=complex(ap_raw);
random_complex=exp(1i*(.05*2*pi)*randn(N_ap,N_ap));
idx = ap_raw == 1;
ap_raw(idx) = random_complex(idx);
end

%use the squareroot of intensity ratio since we want electric field.
ap1=sqrt(intensity_ratio)*ap_raw; %make the star very bright
ap2=ap_raw; %and the exoplanet very dim

% do the apodization
if 1==strcmp(Apodization,'n')
    T=1;
elseif 1==strcmp(Apodization,'sonine')
    nu=nu;
    T=((1-((x1ap/sieve_halfwidth).^2).^(nu-1)).*...
        ((1-((y1ap/sieve_halfwidth).^2).^(nu-1))));
elseif 1==strcmp(Apodization,'cosine')
    %do cos4 math
    T=(cos(pi/4*(x1ap/sieve_halfwidth)).^cosval).*...
        (cos(pi/4*(y1ap/sieve_halfwidth)).^cosval);
```

```
elseif 1==strcmp(Apodization,'gauss')
    sigx=sig;
    sigy=sig;
    T=exp(-(((((x1ap/sieve_halfwidth)-0).^2)./(2*sigx^2))+...
        ((((y1ap/sieve_halfwidth)-0).^2)./(2*sigy^2))));
elseif 1==strcmp(Apodization,'squarepyra')
    T=zeros(N_ap); %initialize
    flattop_area=flattop_area;
    flattop_fraction=sqrt(flattop_area);
    flattop=floor(flattop_fraction*(N_ap/2));
    rampvec=linspace(0,1,(N_ap/2)-flattop);
    for counter=1:(N_ap/2)-flattop;
        Ttemp=rampvec(counter)*ones(N_ap-2*counter);
        T(counter+1:N_ap-counter,counter+1:N_ap-counter)=Ttemp;
    end
    for counter=(N_ap/2)-(flattop-1):N_ap/2
        Ttemp=ones(N_ap-2*counter);
        T(counter+1:N_ap-counter,counter+1:N_ap-counter)=Ttemp;
    end
else disp('Thats not an option for apodization')
end

%Apply the apodization
ap1=T.*ap1;
ap2=T.*ap2;

% pad the matrices with zeros for the Fresnel integral
ap1=padarray(ap1,[(N-N_ap)/2 (N-N_ap)/2]);
ap2=padarray(ap2,[(N-N_ap)/2 (N-N_ap)/2]);

if 0==strcmp(Apodization,'n')
% display the apodization so you can make sure it is correct
```

```
figure(1)
h=surf(T);
set(h,'LineStyle','none')
caxis([0 1])
zlim([0 1])
end

toc

%% Fresnel Integral
% one_step_prop.m from Schmidt: Numerical Simulation of Optical Wave
% Propagation, 2010.
[x2, y2, Uout1] = one_step_prop(ap1, wvl, delta1, z);
I1=Uout1.*conj(Uout1);
I2=I1./(intensity_ratio^2);

disp('Done with Fresnel Integral')
toc

%% Offset and Add the two focal spots
num_diff_lim=6; % how many times the diff. limit (resels) to make the offset
offset=num_diff_lim*resel; %[m]

% use the resolution to determine how many rows to pad
output_res=x2(1,2)-x2(1,1); % Fresnel integral output resolution
rows2pad=floor((offset/sqrt(2))/output_res); % calc # of rows to pad

% now pad "rows2pad" on the bottom and the right (for the 45deg offset)
I2pad=padarray(I2,[rows2pad rows2pad],0,'post');

% and now combine the focal spots
Itot=I2pad(rows2pad+1:end,rows2pad+1:end)+I1;
```

```
% normalize to the peak
normItot=Itot./max(max(Itot));

% create a grid in resels
Resel_xgv=x2(1,:)/resel;
Resel_ygv=(y2(:,1)/resel)';
Resel_mesh=meshgrid(Resel_xgv,Resel_ygv);

%% Log Image Plot
% Show the irradiance
croploc=15; %how many resels to crop to
Clim1=knnsearch(Resel_xgv',-croploc);
Clim2=knnsearch(Resel_xgv',croploc);
Rlim1=knnsearch(Resel_ygv',-croploc);
Rlim2=knnsearch(Resel_ygv',croploc);
dimcentx=knnsearch(Resel_xgv',num_diff_lim/sqrt(2))
dimcenty=knnsearch(Resel_ygv',num_diff_lim/sqrt(2))
normItot_crop=normItot(Rlim1:Rlim2,Clim1:Clim2);
normLog_crop=log10(normItot_crop);
normLog_crop=rot90(normLog_crop,2);

LogImage=figure(2);
h=imagesc(Resel_xgv(Clim1:Clim2),Resel_ygv(Rlim1:Rlim2),normLog_crop);
caxis([-6 0])
axis equal
axis square
xlim([-croploc croploc])
ylim([-croploc croploc])
set(gca,'XTick',-15:5:15,'YTick',-15:5:15)
caxis([-6 0])
hold on
```

```
th = 0:pi/200:2*pi;
xunit = (2)*cos(th) + Resel_xgv(dimcentx);
yunit = (2)*sin(th) + Resel_ygv(dimcenty);
hold on
plot(xunit, yunit,'-w','LineWidth',1.5);
set(gca,'FontSize',20);
xlabel('Position, {\it{x}}, [Resels]')
ylabel('Position, {\it{y}}, [Resels]')
if 1==strcmp(PlotSave,'y')
print(['LogImage_' nameforplot],'-dpng','-r800')
end
hold off

%% PSF Diag plot
p_psf=diag(normLog_crop);
%rotate by 90 degrees to get a cross section without the exoplanet present
s_psf=diag(rot90(normLog_crop,1));
diagResel=Resel_xgv*sqrt(2);

PSFDiag=figure(4);
axis square
h=plot(diagResel(Clim1:Clim2),s_psf,'-.r',diagResel(Clim1:Clim2),p_psf,'k-');
set(h,'LineWidth',2)
xlim([-15 15])
set(gca,'FontSize',20);
xlabel('Position along Diagonal [Resels]')
ylabel('Log_{10}(Normalized Irradiance)')
hold on
dashedlinehandle=line([num_diff_lim num_diff_lim],[-6 0],...
    'LineStyle','--','Color','black')
set(dashedlinehandle,'LineWidth',2)
ylim([-6 0])
```

```
set(gca,'XTick',-15:5:15,'YTick',-6:0)
hold off
if 1==strcmp(PlotSave,'y')
print(['PSFDiag_' nameforplot],'-dpng','-r800')
end
```

Appendix D. PSF Image Calibration

```
%% Thomas Dickinson - Thesis PSF Processing, Image Calibration
% This code performs the calibration on raw PSF image data
clc, close all %close everything for speed
optics={'500mmLens' 'circ_sieve_controlled_unapod' 'circ_sieve_sonine5_apod'...
    'square_sieve_controlled_unapod' 'square_sieve_pyramid66_apod'...
    'square_sieve_pyramid33_apod' 'square_sieve_sonine5_apod'...
    'square_sieve_sonine4_apod' 'square_sieve_cos2_apod'...
    'square_sieve_cos4_apod' 'square_sieve_gauss_apod'};

%% Loop through for each test optic to process all data at once
for opticnumber=1;
close all
clearvars -except optics opticnumbern %clear to free memory
optic=optics(opticnumber);
%set the input director - each test optic had its own folder
folder=['C:\CameraData\' char(optic) '\'];

%Initialize (data cube size is 2532 pixels X 3352 pixels X 10 images)
cube_S=zeros(2532,3352,10); %10 short lights
cube_CS=cube_S; %10 short calibrations (darks)
cube_L=cube_S; %10 long lights
cube_CL=cube_S; %10 long calibrations (darks)

%% Load all the raw, 16-bit .fit data
%Image filenames were generated with Nebulosity 4 during acquisition
for n=1:9
    cube_S(:,:,n)=fitsread([folder 'S_00' num2str(n) '.fit']);
    cube_CS(:,:,n)=fitsread([folder 'CS_00' num2str(n) '.fit']);
end
```

```
    cube_S(:,:,10)=fitsread([folder 'S_010' '.fit']);
    cube_CS(:,:,10)=fitsread([folder 'CS_010' '.fit']);

for n=1:9
    cube_L(:,:,n)=fitsread([folder 'L_00' num2str(n) '.fit']);
    cube_CL(:,:,n)=fitsread([folder 'CL_00' num2str(n) '.fit']);
end
    cube_L(:,:,10)=fitsread([folder 'L_010' '.fit']);
    cube_CL(:,:,10)=fitsread([folder 'CL_010' '.fit']);

%% Average the lights and calibration frames for a cleaner calibration
mean_CS=mean(cube_CS,3);
mean_CL=mean(cube_CL,3);
mean_S=mean(cube_S,3);
mean_L=mean(cube_L,3);

%Find the StDev of the Mean for the data sets
std_S=std(cube_S,0,3)/sqrt(10);
std_L=std(cube_L,0,3)/sqrt(10);

%clear the raw data to free memory
clear cube_S cube_CS cube_L cube_CL

%% Subtract the calibration frames from the data
S=mean_S-mean_CS;
L=mean_L-mean_CL;

%% Shift the data so smallest value is 1
%find the smallest element in each column, average those up, take the
%absolute value, and add it across the whole image
S=S+abs(mean(min(S))); %after subtraction some values are negative
L=L+abs(mean(min(L))); %because of this they must be shifted up
```

```
S(S<=0)=1; %after shifting replace any remaining negative values with 1
L(L<=0)=1;
save(['C:\ProcessorOutput\S_' char(optic) '.mat'],'S'); %Save short master
save(['C:\ProcessorOutput\L_' char(optic) '.mat'],'L'); %Save long master
end
```

Appendix E. HDR Scale-and-Splice, Peak Finding, and Semi-Empirical SNR

```
%% Thomas Dickinson - Thesis Subplot Generator Semi Empirical SNR
% This generates the plots for all the optics at once (after calibrating
% the data using the ImageCalib code). This code was hastily written and is
% probably difficult to follow. Please email thomas.dickinson.3@us.af.mil
% with any questions.

%% Initialize
clc, close all
optics={'500mmLens' 'circ_sieve_controlled_unapod' 'circ_sieve_sonine5_apod'...
    'square_sieve_controlled_unapod' 'square_sieve_pyramid66_apod'...
    'square_sieve_pyramid33_apod' 'square_sieve_sonine5_apod'...
    'square_sieve_sonine4_apod' 'square_sieve_cos2_apod'...
    'square_sieve_cos4_apod' 'square_sieve_gauss_apod'};
plotletters={'Lens' '(a)' '(b)' '(c)' '(d)' '(e)' '(f)' '(g)' '(h)' '(i)' '(j)'};
M=xlsread('C:\PSFData.xlsx','Sheet1','H2:H12'); %Calculated magnifications
int_S=1; %short exposure length [seconds]
int_L=20; %long exposure length [seconds]
angleint=linspace(0,360,500); %500 angles from 0 to 360 degrees
dimPeak=1E-3; %sets dim peak height (bright peak = 1)

%parameters for figure and panel size, using Martineau's "Perfect Subplot"
plotheight=18;
plotwidth=16;
subplotsx=3;
subplotsy=4;
leftedge=1.2;
rightedge=0.4;
topedge=1;
```

```
bottomedge=1.5;
spacex=0.2;
spacey=0.2;
fontsize=5;
sub_pos=subplot_pos(plotwidth,plotheight,leftedge,rightedge,bottomedge,...
    topedge,subplotsx,subplotsy,spacex,spacey);
counter=0;
opticnumberlist=[9 7 4 1 10 8 5 2 11 0 6 3]; %orders the subplots properly
f=figure('visible','on');
clf(f);
set(gcf, 'PaperUnits', 'centimeters');
set(gcf, 'PaperSize', [plotwidth plotheight]);
set(gcf, 'PaperPositionMode', 'manual');
set(gcf, 'PaperPosition', [0 0 plotwidth plotheight]);
colors=linspecer(3); %linspecer on Mathworks file exchange

for i=1:subplotsx
for ii=1:subplotsy
ax=axes('position',sub_pos{i,ii},'XGrid','off','XMinorGrid','off',...
    'FontSize',fontsize,'Box','on','Layer','top');

counter=counter+1;
if counter==10;
      ttt=text(0.25,0.6,{''; '(a) Circular';...
          '(b) Circular, Sonine (\nu = 5)'; '(c) Square';...
          '(d) Square, Pyramid (2/3^{rds})';...
    '(e) Square, Pyramid (1/3^{rd})'; '(f) Square, Sonine (\nu = 5)';...
    '(g) Square, Sonine (\nu = 4)'; '(h) Square, Cos^{2}';...
    '(i) Square, Cos^{4}'; '(j) Square, Gaussian'});
      axis off
      set(ttt,'FontSize',8);
end
```

```
if counter ~=10
opticnumber=opticnumberlist(counter);
optic=optics(opticnumber);
folder='C:\CameraData\'; %set the data directory
%Variables named "S" and "L" for calibrated short and long exposure
load(['S_' char(optic) '.mat']); %load the calibrated short exposure
load(['L_' char(optic) '.mat']); %load the calibrated long exposure

%% Scale-and-Splice
L_vect=reshape(L,[2532*3352,1]);
L_unsat=L_vect(L_vect < .80*max(L_vect)); %grab the unsaturated pixels
pix=1:2532*3352;
pix_unsat=pix(L_vect < .80*max(L_vect)); %grab the unsaturated pixels
%Scale the unsaturated long exposure data
L_unsat_scaled=L_unsat/(int_L/int_S);
HDR=S; %initialize the HDR data
%Splice
HDR(pix_unsat)=L_unsat_scaled; %replace with the unsaturated long exposure
HDRnorm=HDR./max(max(HDR)); %normalize to the peak
clear L_vect L_unsat pix pix_unsat L_unsat_scaled S L

%% Determine the Peak Location
[R, C]=find(HDR==max(max(HDR))); %starts with the peak value
SpotSizeBox=50; %search a 50x50 box centered on peak value
Vcs_avg=zeros(size(S)); %average 20x20 boxes around the 2500 pixels
for j=-ceil(.5*SpotSizeBox):ceil(.5*SpotSizeBox);
    for k=-ceil(.5*SpotSizeBox):ceil(.5*SpotSizeBox);
    Vcs_avg(R+j,C+k)=mean2(HDR((R+j)-10:(R+j)+10,(C+k)-10:(C+k)+10));
    end
end
[Rmax, Cmax]=find(Vcs_avg==max(max(Vcs_avg))); %Peak after smoothing

```

```
%% Set the coordinate system
xgv=linspace(-8.98,8.98,3352); %[mm] vectors using detector dimensions
ygv=linspace(-6.76,6.76,2532); %[mm]
scaled_xgv=1e3*(xgv/M(opticnumber)); %[um] scale w/ M to focal plane coords
scaled_ygv=1e3*(ygv/M(opticnumber)); %[um]
scaled_xgvcent=scaled_xgv-scaled_xgv(Cmax); %[um] center the data
scaled_ygvcent=scaled_ygv-scaled_ygv(Rmax);
%make a centered and scaled meshgrid
[scaled_Xgvcent, scaled_Ygvcent]=meshgrid(scaled_xgvcent,scaled_ygvcent);
clear xgv ygv scaled_xgv scaled_ygv scaled_xgvcent scaled_ygvcent

%% Perform the SemiEmp Math
cntr=0;
beep
stepsize=20;
for R=70:-stepsize:30; %radius of circle [um]... means at 20,40,60 um
    cntr=cntr+1;
    cntr2=0;
    SNRmatrix=zeros(stepsize*20+1,500); %initialize matrix to hold SNRs
    for Rstep=R-.5*stepsize:.05:R+.5*stepsize %step through 400 radii
    %Do the math
        cntr2=cntr2+1;
        disp(Rstep)
        Circle = abs((sqrt((scaled_Xgvcent).^2+(scaled_Ygvcent).^2))-...
            Rstep)<.01;
        [Rowcirc_ind, Colcirc_ind]=find(Circle==1);
        HDRnormCirc=zeros(1,length(Rowcirc_ind));
        angle=atan2d(Rowcirc_ind-Rmax,Colcirc_ind-Cmax); %careful w/ atan2d
        angle = angle+(angle < 0)*360; %convert to a 0 to 360 degree scale
        Combined=cat(2,angle,Rowcirc_ind,Colcirc_ind);
        Combined=sortrows(Combined,1);
        Rowcirc_ind=Combined(:,2);
```

```
        Colcirc_ind=Combined(:,3);
        angle=Combined(:,1);
        for n=1:length(Rowcirc_ind);
            HDRnormCirc(n)=HDRnorm(Rowcirc_ind(n),Colcirc_ind(n));
        end
        SNR=(HDRnormCirc+dimPeak)./HDRnormCirc;
        SNRmatrix(cntr2,:)=interp1(angle,SNR,angleint);
    end
    SNR=mean(SNRmatrix,1);
    SNR=smooth(angleint,SNR,0.05,'rloess');
    SNRmean=nanmean(SNR);
    SNRclean=SNR;
    SNRclean(isnan(SNRclean))=[];
    [LowerSTD, UpperSTD]=semistd(SNRclean); %semistd.m written by Alhouti
    SNRstd=nanstd(SNR);
    clear Circle Rowcirc_ind Colcirc_ind HDRnormCirc angle Combined SNRmatrix;

%% Make all plots at the same time in the loops
    hold on
    plot(angleint,SNR,'-','Color',colors(cntr,:))
    set(gca,'FontSize',8);
    xlim([0 360])
    ylim([0 35])
    hold on

    %***********DRAW FILLED IN BOUNDS************
    index=1:length(angleint);
    h1 = fill(angleint(index([1 1:end end])),...
          [(SNRmean-LowerSTD) (SNRmean+UpperSTD)*ones(1,length(angleint))...
          (SNRmean-LowerSTD)],...
         colors(cntr,:),'EdgeColor','none');
    set(h1,'FaceAlpha',0.2);
```

```
    set(get(get(h1,'Annotation'),'LegendInformation'),...
        'IconDisplayStyle','off');

    %average line
    dashedlinehandle1=line([0 360],[SNRmean SNRmean],'LineStyle','--',...
        'Color',colors(cntr,:),'LineWidth',.5);
    set(get(get(dashedlinehandle1,'Annotation'),'LegendInformation'),...
        'IconDisplayStyle','off');

    if opticnumber ==9
        xlabel('Azimuthal Angle from Horizontal, [degrees]')
        ylabel('SNR')
        set(ax,'XTick',0:45:360,'YTick',0:5:35)
    end
    if opticnumber ~= 9
        set(ax,'XTick',0:45:360,'YTick',0:5:35);
        set(ax,'XTickLabel','','YTickLabel','','ZTickLabel','')
    end
    tt3=text(2,33,['' char(plotletters(opticnumber)) '']);
    set(tt3,'FontSize',10)
    set(tt3,'Color','k')
    uistack(tt3,'top')
    hold off
end
end
end
end
```

Bibliography

1. Alhouti, A. Semideviation. `http://www.mathworks.com/matlabcentral/fileexchange/45251-semideviation`, 2014.

2. Andersen, G. Large optical photon sieve. *Optics Letters 30*, 22 (2005), 2976–2978.

3. Andersen, G. "Regular Photon Sieve", MATLAB Code, Private Communication. 2014.

4. Andersen, G. "PSF Imaging - Focus Method", Private Communication. 2015.

5. Andersen, G., and Tullson, D. Broadband antihole photon sieve telescope. *Applied Optics 46*, 18 (2007), 3706–3708.

6. Andersen, G. P., Asmolova, O., and Dickinson, T. FalconSAT-7: a membrane space telescope. *Proc. SPIE 9085* (2014), 908504X.

7. Asmolova, O. "Interference Fringe Analysis of Various Photon Sieves", Private Communication. 2016.

8. Asmolova, O., Andersen, G., Dearborn, M. E., McHarg, M. G., Quiller, T., and Dickinson, T. Optical testing of a membrane diffractive optic for space-based solar imaging. *Proc. SPIE 9006* (2014), 90060D.

9. Asmolova, O., Andersen, G., McHarg, M. G., Quiller, T., Maldonado, C., and Dickinson, T. Design and test of a novel solar imaging payload for small satellites. *Proc. SPIE 9602* (2015), 96020F.

10. Born, M., and Wolf, E. *Principles of Optics*, 7th (expanded) ed. Cambridge University Press, Cambridge, 2000.

11. Brown, R. A., and Burrows, C. J. On the feasibility of detecting extrasolar planets by reflected starlight using the Hubble Space Telescope. *International Journal of Solar System Studies 87* (1990), 484–497.

12. Brown, R. A., Burrows, C. J., Casertano, S., Clampin, M., Ebbets, D. C., Ford, E. B., Jucks, K. W., Kasdin, N. J., Kilston, S., Kuchner, M. J., Seager, S., Sozzetti, A., Spergel, D. N., Traub, W. A., Trauger, J. T., and Turner, E. L. The 4-m space telescope for investigating extrasolar Earth-like planets in starlight: TPF is HST2. *Proc. SPIE 4854* (2003), 95–107.

13. Cao, Q., and Jahns, J. Focusing analysis of the pinhole photon sieve: individual far-field model. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision 19*, 12 (2002), 2387–93.

14. Cao, Q., and Jahns, J. Modified Fresnel zone plates that produce sharp Gaussian focal spots. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision 20*, 8 (2003), 1576–1581.

15. Cao, Q., and Jahns, J. Nonparaxial model for the focusing of high-numerical-aperture photon sieves. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision 20*, 6 (2003), 1005–12.

16. Cao, Q., and Jahns, J. Comprehensive focusing analysis of various Fresnel zone plates. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision 21*, 4 (2004), 561–71.

17. Clavin, W., Chou, F., and Johnson, M. NASA's Kepler Marks 1,000th Exoplanet Discovery, Uncovers More Small Worlds in Habitable Zones., 2015.

18. EdmundOptics.com. Understanding Spatial Filters. `http://www.edmundoptics.com/resources/application-notes/lasers/understanding-spatial-filters/`. Accessed: 2015-10-15.

19. Eismann, M. T. *Hyperspectral Remote Sensing*. SPIE Press, Bellingham, Washington, 2012.

20. Franz, A. Mathematical Method for Applying Non Circularly Symmetric Apodizations to Photon Sieves, Private Communication. 2015.

21. Gliese, W., and Jahreiss, H. CNS3 - Gliese Catalog of Nearby Stars, 3rd Edition, 1991.

22. Goodman, J. W. *Introduction to Fourier Optics*, 3rd ed., vol. 35. Roberts & Company, Greenwood Village, CO, 1996.

23. Hawks, M. "Zemax Analysis of the ThorLabs AC508-500 Achromatic Doublet Lens", Private Communication. 2015.

24. Hecht, E. *Optics*, 4th ed. Addison-Wesley, San Francisco, 2002.

25. Kasdin, N. J., Brown, R. A., Burrows, C. J., Kilston, S., Kuchner, M., Littman, M. G., Noecker, M. C., Seager, S., Spergel, D. N., Turner, E. L., Traub, W. A., Vanderbei, R. J., and Woodruff, R. A. An optical/UV space coronagraph concept for the terrestrial planet finder. *Advances in Space Research 34* (2004), 625–630.

26. Kasdin, N. J., Vanderbei, R. J., Spergel, D. N., and Littman, M. G. Extrasolar Planet Finding via Optimal Apodized Pupil and Shaped Pupil Coronagraphs. *The Astrophysical Journal 582*, 2 (2003), 1147–1161.

27. Kipp, L., Skibowski, M., Johnson, R. L., Berndt, R., Adelung, R., Harm, S., and Seemann, R. Sharper images by focusing soft X-rays with photon sieves. *Nature 414*, 6860 (2001), 184–188.

28. Latham, D. W., Stefanik, R. P., Mazeh, T., Mayor, M., and Burki, G. The unseen companion of HD114762 - A probable brown dwarf. *Nature (ISSN 0028-0836) 339* (1989), 38.

29. Marois, C., Macintosh, B., Barman, T. S., Zuckerman, B., Song, I., Patience, J., Lafreniere, D., and Doyon, R. Direct Imaging of Multiple Planets Orbiting the Star HR 8799. *Science 322* (2008), 1348.

30. Martineau, P. Perfect subplot in Matlab. `http://p-martineau.com/perfect-subplot-in-matlab/`, 2014.

31. MathWorks. MATLAB®, 2014.

32. Myers, O. E. Studies of Transmission Zone Plates. *American Journal of Physics 19*, 6 (1951), 359.

33. Newport.com. Spatial Filters. `http://www.newport.com/Spatial-Filters/144910/1033/content.aspx`. Accessed: 2015-10-15.

34. Nisenson, P., and Papaliolios, C. Detection of Earth-like Planets Using Apodized Telescopes. *The Astrophysical Journal 548*, 2 (2001), L201–L205.

35. Owens, S. "An Efficient MATLAB Implementation for Applying Non Circularly Symmetric Apodizations to Photon Sieves", Private Communication. 2015.

36. Ryden, B., Peterson, B., and Demianski, M. *Foundations of Astrophysics*, vol. 78. Addison-Wesley, San Francisco, 2010.

37. Schmidt, J. D. *Numerical Simulation of Optical Wave Propagation*. Society of Photo-Optical Engineers (SPIE), Bellingham, WA, 2010.

38. Siegman, A. E., Sasnett, M. W., and Johnston, T. F. Choice of clip levels for beam width measurements using knife-edge techniques. *IEEE Journal of Quantum Electronics 27*, 4 (1991), 1098–1104.

39. Sparrow, C. M. On Spectroscopic Resolving Power. *The Astrophysical Journal 44* (1916), 76.

40. Stark, C. Nebulosity 4. `http://www.stark-labs.com/nebulosity.html`, 2015.

41. Sussman, M. Elementary Diffraction Theory of Zone Plates. *American Journal of Physics 28*, 4 (1960), 394.

42. Tulip, C. Photon Sieve Bandwidth Broadening by Reduction of Chromatic Aberration Effects Using Second-Stage Diffractive Optics, Master's Thesis. Air Force Institute of Technology. 2015.

43. Tulip, C. Primary Photon Sieve Design MATLAB Code, Master's Thesis, Appendix. Air Force Institute of Technology. 2015.

44. Voelz, D. *Computational Fourier Optics: A MATLAB Tutorial*. Society of Photo-Optical Engineers (SPIE), Bellingham, WA, 2011.

45. Wetzler, A. Bresenham optimized for Matlab. `mathworks.com/matlabcentral/fileexchange/28190-bresenham-optimized-for-matlab`, 2010.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704–0188

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704–0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202–4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE *(DD–MM–YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From — To)*
24–03–2016	Master's Thesis	June 2015 — Mar 2016

4. TITLE AND SUBTITLE

Simulation, Design, and Test of Square,
Apodized Photon Sieves for High-Contrast, Exoplanet Imaging

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Dickinson, Thomas W. N., 2d Lt, USAF

5d. PROJECT NUMBER

15P935

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Air Force Institute of Technology
Graduate School of Engineering and Management (AFIT/EN)
2950 Hobson Way
WPAFB OH 45433-7765

8. PERFORMING ORGANIZATION REPORT NUMBER

AFIT-ENP-MS-16-M-065

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

HQ USAFA/Department of Physics
Attn: Matthew G. McHarg
2354 Fairchild Dr, Suite 2A31
USAF Academy, CO 80840-6254
(719) 333-3510 (DSN: 333-3510) matthew.mcharg@usafa.edu

10. SPONSOR/MONITOR'S ACRONYM(S)

USAFA

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

DISTRIBUTION STATEMENT A:
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

13. SUPPLEMENTARY NOTES

This material is declared a work of the U.S. Government and is not subject to copyright protection in the United States.

14. ABSTRACT

A photon sieve is a lightweight, diffractive optic which is well-suited to be a deployable primary for a space telescope. Point spread functions (PSFs) can be altered by shaping and apodizing an aperture, and a PSF that drops rapidly from the peak is desirable for high-contrast imaging. For this reason, square apodized photon sieves were simulated, designed, and tested for high-contrast performance and use in an exoplanet imaging telescope. These sieves were shown to outperform conventional optics and unapodized sieves for high-contrast imaging in a number of tests. New methods were developed for apodizing sieves, measuring PSFs, and characterizing high-contrast performance. Tests indicated that square apodized sieves could detect exoplanets with irradiance below $10^{-3.69}$ of the star's PSF peak within ten diffraction limits of separation. This was not sufficient for directly imaging earth-like exoplanets, but will be useful for other high-contrast applications. The Fresnel diffraction simulation conducted for the sieves was shown to agree closely with the experimental results. The ability to accurately apply apodizations and conduct simulations for photon sieves, measure PSFs across an extreme dynamic range, and conduct high-contrast imaging performance analyses will drive new PSF design and be useful for future high-contrast imaging work.

15. SUBJECT TERMS

Photon Sieve, Apodized, Apodization, Diffractive Optic, Fresnel Zone Plate, Simulation, Exoplanet, High-Contrast Imaging

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Anthony L. Franz, Lt Col, PhD, AFIT/ENP
U	U	U	UU	182	**19b. TELEPHONE NUMBER** *(include area code)* (937) 255-3636, x4429; anthony.franz@afit.edu

Standard Form 298 (Rev. 8–98)
Prescribed by ANSI Std. Z39.18