

Rochester Institute of Technology

RIT Digital Institutional Repository

Theses

10-28-2018

Measurement of Amplitude and Phase Pupil Variation for EUV Lithography Systems

Zachary A. Levinson
zal2186@rit.edu

Follow this and additional works at: <https://repository.rit.edu/theses>

Recommended Citation

Levinson, Zachary A., "Measurement of Amplitude and Phase Pupil Variation for EUV Lithography Systems" (2018). Thesis. Rochester Institute of Technology. Accessed from

This Dissertation is brought to you for free and open access by the RIT Libraries. For more information, please contact repository@rit.edu.

R·I·T

MEASUREMENT OF AMPLITUDE AND PHASE PUPIL VARIATION FOR EUV LITHOGRAPHY SYSTEMS

by

Zachary A. Levinson

A dissertation submitted in partial fulfillment of the requirements for the
degree of Doctorate of Philosophy in Microsystems Engineering

Microsystems Engineering Program
Kate Gleason College of Engineering

Rochester Institute of Technology
Rochester, New York
October 28, 2018

MEASUREMENT OF AMPLITUDE AND PHASE PUPIL VARIATION FOR EUV LITHOGRAPHY SYSTEMS

by

Zachary A. Levinson

Committee Approval:

We, the undersigned committee members, certify that we have advised and/or supervised the candidate on the work described in this dissertation. We further certify that we have reviewed the dissertation manuscript and approve it in partial fulfillment of the requirements of the degree of Doctor of Philosophy in Microsystems Engineering.

Dr. Bruce W. Smith
Committee Chair and Dissertation Advisor

Date

Dr. Zoran Ninkov
Committee Member

Date

Dr. Stefan Preble
Committee Member

Date

Dr. Timothy Brunner
Committee Member

Date

Dr. Obert Wood
Committee Member

Date

Dr. Robert Pearson
Committee Member

Date

Certified by:

Dr. Bruce W. Smith
Director, Microsystems Engineering Program

Date

Dr. Doreen Edwards
Dean, Kate Gleason College of Engineering

Date

ABSTRACT

Kate Gleason College of Engineering
Rochester Institute of Technology

Degree: Doctor of Philosophy

Program: Microsystems Engineering

Author's Name: Zachary A. Levinson

Advisor's Name: Dr. Bruce W. Smith

Dissertation Title: MEASUREMENT OF AMPLITUDE AND PHASE PUPIL VARIATION FOR EUV LITHOGRAPHY SYSTEMS

Aberration control and characterization in a state of the art photolithographic lens have the tightest tolerances of any optical system. This is especially true in next generation extreme ultraviolet lithography systems with estimates for the wavefront tolerance below 500 pm RMS. These systems use radiation at a wavelength of 13.5 nm. No materials sufficiently refract this radiation, so reflective lens designs must be used. The mirrors are constructed as Bragg reflectors and much of the intense power of the source is ultimately distributed through the system as heat with each reflection. Moreover, the angle dependent reflection of these mirrors can also lead to amplitude asymmetries across the pupil. While interferometric techniques are the de-facto standard of wavefront analysis, they require the use of additional optics and are therefore difficult to implement during system use. Moreover, interferometric techniques cannot measure amplitude pupil variation.

In this work both the pupil amplitude and phase variation of several EUV lithography systems will be measured using images of binary targets formed by each system. Using the systems' own images to monitor its wavefront has the benefit of providing an aberration monitor during system use. Models will be constructed between wavefront variation and a space-domain basis in which the effects of aberrations are separable. This allows both the amplitude and pupil variation to be rapidly extracted from these systems. Finally, the theory of anamorphic primary aberrations will be developed and the image-based method will be extended to these types of systems.

ACKNOWLEDGMENTS

Many people helped to guide and support me on my endeavor to become a researcher of photolithography. The following people made significant contributions to my efforts and mental well-being, and for that I am enormously grateful.

Specifically, I would like to thank Dr. Bruce Smith for encouraging me to pursue a doctoral degree, for supporting me along the way whenever and however necessary, and for helping to hone my independent research and communication skills. I would not be writing this without him. I would also like to thank my group members: Andrew Burbine, Germain Fenger, Burak Baylav, and Andrew Estroff. Your friendship and conversations helped shape my research in ways for which I cannot begin to thank you enough. I would like to thank Lisa Zimmerman for everything she does to keep the Microsystems Engineering department running and her continual patience.

The process of obtaining this degree would have been much more difficult without Valerie Altounian sneaking research-themed illustrations into my bags when travelling to conferences, and of course without her unconditional emotional support. Likewise, my parents, Diane and Michael Levinson, and sister, Samantha Levinson, provided necessary and unconditional support.

I would like to thank my committee members, Dr. Stefan Preble, Dr. Zoran Ninkov, Dr. Obert Wood II, and Dr. Timothy Brunner for their support and guiding me through the Ph.D. process.

My work would also not be possible without my many collaborators, many of whom have been already mentioned, but I would also like to thank: Erik Verduijn, Antoine Wojdyla, Markus Benk, and Kenneth Goldberg. My experimental work would have not been possible without them, nor without the financial support of GLOBALFOUNDRIES and SEMATECH. Actinic imaging and analysis was performed by the University of California, Lawrence Berkeley National Laboratory under the auspices of the U.S. Department of Energy, Contract No. DE-AC02-05CH11231. Finally, I would like to acknowledge and thank KLA Tencor for their use of the PROLITH™ simulator.

Table of Contents

Abstract.....	iv
Acknowledgments.....	v
List of Figures.....	ix
List of Tables.....	xxi
List of Acronyms.....	xxiii
1. Introduction.....	1
1.1 Photolithography Systems.....	2
1.2 EUV Lithography.....	8
1.2.1 Anamorphic systems.....	12
1.3 Lens Quality Metrics.....	14
1.3.1 Point Spread Function.....	14
1.3.2 Apodization.....	16
1.3.3 Aberrations.....	17
1.3.4 Flare.....	19
2. Background.....	22
2.1 Theory of Image Formation.....	22
2.1.1 Geometrical Image Formation.....	22
2.1.2 Coherent Image Formation.....	32
2.1.3 Partially Coherent Image Formation.....	39
2.2 Linear Algebra Techniques.....	45
2.2.1 Eigenvector Decomposition.....	45
2.2.2 Singular Value Decomposition.....	46
2.2.3 Principal Component Analysis.....	47
2.2.4 SOCS decompositions.....	48
3. Measuring Pupil Variation.....	50
3.1 A naïve example.....	50
3.2 Interferometric techniques.....	54
3.3 Image-based techniques.....	55
3.4 Optimization of image-based metrology targets.....	59
4. Extraction of amplitude variation.....	63

4.1	Choosing a basis to represent amplitude pupil variation	63
4.2	Comparison of the properties of several orthonormal bases	71
4.2.1	Zernike Polynomials.....	71
4.2.2	Legendre Polynomials	73
4.2.3	Hermite Polynomials.....	74
4.3	Extraction of amplitude variation	75
5.	The QUIP Algorithm	81
5.1	The QUIP Algorithm	81
5.2	Δ CD Eigenfunctions	87
5.3	Obtaining inverse pupil solutions	93
5.4	Measuring high-order aberrations.....	97
5.5	The effects of noise on aberration retrieval.....	100
5.6	QUIP Software.....	103
5.6.1	Image viewer.....	104
5.6.2	Model builder.....	106
5.6.3	Wavefront analyzer	108
5.6.4	Program Setup.....	109
6.	Experimental Results	112
6.1	SHARP Pupil Extraction.....	112
6.2	AIMS EUV™ Pupil Extraction	116
6.2.1	Third-Order Only	116
6.2.2	Third and Fifth Order	123
6.3	NXE:3300 Pupil Extraction.....	126
6.3.1	First Extraction.....	127
6.3.2	Second Extraction	133
6.3.3	Third-Order Only	133
6.3.4	Third and Fifth Order	135
7.	Aberrations in Anamorphic lithography systems	138
7.1	Primary aberrations of anamorphic optics.....	139
7.2	A new aberration basis	144
7.3	Simulating anamorphic aberrations	154

7.4	Using Zernike polynomials to describe anamorphic aberrations	156
7.5	Image-based metrology of anamorphic aberrations.....	157
8.	Conclusions	159
9.	References	161
10.	Appendix A: Annular anamorphic polynomials.....	166
11.	Appendix B: Annular anamorphic polynomial tables	169

LIST OF FIGURES

Figure 1. An example of a projection lithography system used for relief pattern transfer to a wafer.	3
Figure 2. Typical spectrum of a mercury arc lamp. ⁶	4
Figure 3. Comparison of on-axis illumination and off-axis illumination for a single coherent point source.....	5
Figure 4. Common lithography source shapes, a) conventional centered monopole, b) offset monopole, c) dipole, d) annulus	6
Figure 5. ITRS evaluation of next-generation patterning methods. ¹²	6
Figure 6. ITRS patterning requirements for the next ten years. ¹²	7
Figure 7. Laser produced plasma EUV source. ¹⁵	8
Figure 8. Schematic view of a typical EUV lithography system. The dashed black line shows the chief ray.....	9
Figure 9. A typical EUV mask stack.....	10
Figure 10. Reflectivity as a function of incident angle for a typical EUV multilayer mirror. The specific mirror plotted consists of 40 alternating pairs of 2.8 nm Mo and 4.2 nm Si films with no interdiffusion. ²⁰	10
Figure 11. Shadowing in an EUVL system. Horizontal features (left) appear to increase in size while vertical features (right) do not have this problem. The right panel is in the plane of a space feature.....	11
Figure 12. Geometrical considerations of high-NA EUVL. a) Reflective system close to the maximum NA, b) Larger NA with the same chief ray angle, c) the chief ray angle must increase to allow a higher-NA	11

Figure 13. Comparison of the entrance pupils of an isomorphic projection lens (a) and an anamorphic projection lens (b), and the (c) exit pupil of the same anamorphic lens.....	12
Figure 14. An example of an anamorphic EUVL system.....	14
Figure 15. Transfer of light through a linear imaging system via the point spread function.	15
Figure 16. The optical transfer function and point spread function of an ideal pupil.....	16
Figure 17. The optical transfer function and point spread function of a Gaussian apodized pupil.....	17
Figure 18. An example of the type of imaging system being studied.	22
Figure 19. Coordinate systems in the object and image planes.	23
Figure 20. Geometry of the exit pupil and image planes.....	24
Figure 21. The geometry of an ideal spherical interface	30
Figure 22. The geometry of the coherent diffraction problem.....	33
Figure 23. Contour plots of the 3 rd order Zernike polynomials. Indices are given for the single index classification, along with their classical names.....	39
Figure 24. Geometry of a partially coherent source and its receiving plane	40
Figure 25. Data representing a line in R^3 with Gaussian noise (blue circles) with the orthogonal directions of highest variation identified (black arrows) via PCA.	48
Figure 26. Setup for a lateral shearing interferometer in a lithography system, as in Integrated Lens Interferometer at Scanner (ILIAS).....	55
Figure 27. An SEM micrograph of the cylindrical phase hole on a reticle with a depth of $\lambda/2$ (left) and an image of the phase hole formed in resist (right) ⁶³	56
Figure 28. Transmission electron micrographs of (a) 50 nm and (b) 100 nm phase features post etch.....	57

Figure 29. (a) Examples of metrology targets for each primary aberration. The red lines denote measurement locations. The aberration is interrogated by the CD difference at measurement sites. (b) A flowchart for the iterative algorithm of image-based pupil characterization.	58
Figure 30. Variation in aerial image CD for vertical lines and spaces through focus and different levels of Z5 (a), and the variation in the difference in vertical and horizontal aerial image CD through focus and Z5 (b).	59
Figure 31. Diffraction information and resulting aerial images for 20nm (a) vs. 33nm (b) line/ space array overlaid with 3rd and 5th order 90° astigmatism respectively; 0.33 NA, 0.2σ.	60
Figure 32. NILS contours for astigmatism 90° (top left), coma X (top right), and trefoil X (bottom).	62
Figure 33. a) The image which was rotated then expanded in orthonormal bases. The units are normalized distance so that the function is defined over the unit circle. b) Zernike expansion coefficients for 3rd order astigmatism, coma, spherical aberration, and trefoil for the image in part (a) through rotation.	72
Figure 34. The image in Figure 35a was again rotated and expanded in Legendre polynomials and Hermite polynomials. The moments for a) Legendre polynomials for P_{52} , P_{72} , P_{92} , P_{102} and b) Hermite polynomials for H_{52} , H_{72} , H_{92} , H_{102} are plotted as a function of rotation angle.	74
Figure 35. Flowchart of pupil amplitude reconstruction. First the pupil function is sampled via the diffracted spectrum of the targets used for phase interrogation. The value of these samples is determined via iterative fitting to aerial image simulations. Next	

Barnes analysis is used to interpolate between the samples to construct the original function. Finally, the interpolated function is expanded in a Fourier-Zernike series. 78

Figure 36. NRMSE distribution of the reconstructed amplitude functions with and without interpolation via Barnes objective analysis. The distribution for functions reconstructed without interpolation is semi-transparent to show the entirety of both distributions..... 79

Figure 37. Comparison of an orthogonal function expansion of an interpolated amplitude function in a) Zernike polynomials, and b) the first 36 combinations of Cartesian Legendre polynomials. The RMSE after adding each additional polynomial for both series expansions is given in part c). 80

Figure 38. A 1D aerial image through defocus (a) represents an intensity volume which is part of a higher dimensional space (b) when aberrations are present in an optical system.82

Figure 39. The first two principal components, or eigenfunctions, of the dataset shown in Figure 38. 82

Figure 40. Eigenfunction for a 60 nm pitch line/space array with astigmatism. Because the action of astigmatism is orientation dependent the eigenfunction in this case consists of both plots together. Coordinates are given at the mask level with the dashed lines representing the mask edge. 83

Figure 41. Eigenfunction for a 30 nm line at 60 nm pitch and 90 nm pitch with spherical aberration. Because the action of spherical aberration is pitch dependent the eigenfunction in this case consists of both plots together..... 84

Figure 42. Eigenfunction for a 50 nm CD five-bar pattern with coma (left) and eigenfunction for a 35 nm brick wall structure with trefoil (right). The left most feature in this trefoil eigenfunction represents the right side of a bar..... 84

Figure 43. Eigenfunction for a 60 nm pitch line/space array with Z_{A5} and Z_{A6} . Because the action of these aberrations are orientation dependent the eigenfunction in this case consists of both plots together..... 84

Figure 44. Eigenfunction for a 30 nm line at 60 nm pitch and 90 nm pitch with Z_{A9} . Because the action of Z_{A9} is pitch dependent the eigenfunction in this case consists of both plots together..... 85

Figure 45. Eigenfunction for a 50 nm CD five-bar pattern with Z_{A7} and Z_{A8} (left) and the eigenfunction for a 35 nm brick wall structure with Z_{A10} and Z_{A11} (right). The left most feature of the Z_{A10} and Z_{A11} eigenfunction represents the right side of a bar. 85

Figure 46. Comparison of $\epsilon\mathbf{1}(\mathbf{x})$ and the PCA eigenfunction for an imaging system with coma (Z_7) with zero defocus. 86

Figure 47. Coefficient for aerial images aberrated with coma projected onto the coma eigenfunction. The slope of the line may change with changes in the imaging system, but this function is representative of the projection function for coma..... 87

Figure 48. An example of a Δ CD eigenfunction (left) and the projection curve (right) for astigmatism. Astigmatism is interrogated by the CD difference between orthogonal lines and spaces. 89

Figure 49. An example of a Δ CD eigenfunction (left) and the projection curve (right) for coma. Coma is interrogated by the CD difference between the outer two bars of a five bar structure. 89

Figure 50. An example of a Δ CD eigenfunction (left) and the projection curve (right) for spherical aberration. Spherical is interrogated by a heuristic computed from a constant CD line at pitches 1:1, 1:2, and 1:4.	90
Figure 51. An example of a Δ CD eigenfunction (left) and the projection curve (right) for trefoil. Trefoil is interrogated by the CD difference between orthogonal lines and spaces.	90
Figure 52. An example of a Δ CD eigenfunction (left) and the projection curve (right) for spherical aberration over an extended range of $\pm 250m\lambda$	91
Figure 53. Comatic eigenfunction with varying amounts of trefoil. There are five lines plotted, despite only three being visible. This is because two are on top of each other because the eigenfunction is even in the Z10 axis.	92
Figure 54. Projection coefficient for coma Δ CD functions onto the appropriate eigenfunction with different amounts of trefoil.	92
Figure 55. A flowchart for the aberration separable basis algorithm of image-based pupil characterization.	94
Figure 56. An example of model verification using 25 Gaussian distributed points. Training the model on only seven aberration treatment conditions gives an RMSE on the order of single picometers.	96
Figure 57. An example of model verification using 25 Gaussian distributed points for coma in the presence of trefoil.	96
Figure 58. An iterative solution for both coma and trefoil obtained via the QUIP eigenfunction algorithm.	97
Figure 59. An example of measuring a synthetic wavefront composed from only third-order astigmatism using the QUIP algorithm.	99

Figure 60. An example of measuring a synthetic wavefront composed from only fifth-order astigmatism using the QUIP algorithm.	99
Figure 61. An example of measuring a synthetic wavefront composed from both third- and fifth-order astigmatism using the QUIP algorithm.	100
Figure 62. Histogram of residuals for the left CD of a five-bar structure. The residuals are computed between individual CD measurements and the mean at that dose. The residuals closely follow a normal distribution with a mean of 0 nm and a 0.42 nm standard deviation.	102
Figure 63. Synthetic CD data computed from the line $CD=0.5642 \cdot E^{-2.1448}$ with Gaussian noise with a 0.42 nm standard deviation. The dashed black line is the true line from which the data was computed. A 95% confidence interval for a linear fit to the data is shown in the shaded region between the thick black lines. The true line is within the confidence interval, but is not the line of best fit, given by the solid thin black line.	102
Figure 64. Flowchart of aberration extraction using the image-based method adapted for confidence intervals.	103
Figure 65. The Image Viewer interface showing a sample image from an EUV actinic mask microscope.	105
Figure 66. Example of SEM micrograph image processing for NXE:3100 case study. a) A comparison of the original (left) and median filtered (right) images, b) plot of a intensity line scan across the line in part (a) for both the original and filtered images. ...	105
Figure 67. (a) An example image at the original resolution and after Fourier interpolation (inset). (b) Column-wise average of the interpolated region in the inset of panel (a) to estimate the aerial image.	105

Figure 68. The Model Builder interface showing the setup for a sample model to interrogate coma pupil variation. The top screen shows the simulation setup while the bottom screen shows choosing the response variable and the option to add interacting terms. 107

Figure 69. The Model Builder interface showing the result of building a model for coma.108

Figure 70. (a) The Wavefront Analyzer interface showing a sample setup to interrogate the primary phase aberrations. (b) The Wavefront Analyzer showing the results running the setup shown in panel (a) using the iterative algorithm. 109

Figure 71. An example of the CSV header structure for data files..... 110

Figure 72. A schematic view of the SHARP microscope at the Advanced Light Source. EUV light at a wavelength of 13.5 nm from a bending magnet is focused on the photomask. The photomask is imaged on a CCD sensor using Fresnel zone plate lens. 113

Figure 73. Pupil amplitude and phase extracted from SHARP with only third-order terms. (Z_5 – Z_{11})..... 116

Figure 74. A schematic view of the projection optics in the AIMS EUV system. ⁷⁴..... 117

Figure 75. Comparison of the targets used for aberration interrogation and their images as formed on the AIMS EUV tool..... 118

Figure 76. CD data for vertical spaces (left), horizontal spaces (middle), and their difference (right). These are the targets used to interrogate astigmatism 90° (Z_5)..... 119

Figure 77. CD data for 45° spaces (left), 135° spaces (middle), and their difference (right). These are the targets used to interrogate astigmatism 45° (Z_6)..... 120

Figure 78. CD data for the left and right bars of a vertical five-bar structure (left and middle plots, respectively), and their difference (right). This is the target used to

interrogate coma X (Z7). The grey region denotes the region used for aberration extraction.	120
Figure 79. CD data for the top and bottom bars of a horizontal five-bar structure (left and middle plots, respectively), and their difference (right). This is the target used to interrogate coma Y (Z8). The grey region denotes the region used for aberration extraction.	121
Figure 80. CD data for the left and right ends of a horizontal T-bar structure (left and middle plots, respectively), and their difference (right). This is the target used to interrogate trefoil X (Z10). The grey region denotes the region used for aberration extraction.	121
Figure 81. CD data for the top and bottom ends of a vertical T-bar structure (left and middle plots, respectively), and their difference (right). This is the target used to interrogate trefoil Y (Z11). The grey region denotes the region used for aberration extraction.	122
Figure 82. Extracted wavefront from AIMS EUV plotted with all third-order terms. (Z ₅ –Z ₁₁).	123
Figure 83. Wavefront extracted from AIMS EUV with all third-order terms and their fifth-order counterparts. (Z ₅ –Z ₁₆ , Z ₁₉ ,Z ₂₀)	125
Figure 84. Comparison of the targets used for aberration interrogation and their images in resist via NXE:3300 exposure.....	128
Figure 85. 95% confidence intervals for a second order polynomial fit to CD data for vertical and horizontal lines. This is the target used to interrogate astigmatism 90° (Z ₅).	

The markers are the raw CD-SEM data, while the thin black line is the line of best fit. The thick black lines represent the edges of the confidence interval..... 129

Figure 86. 95% confidence intervals for a second order polynomial fit to CD data for 45° and 135° lines. This is the target used to interrogate astigmatism 45° (Z_6). 130

Figure 87. 95% confidence intervals for a linear regression to CD data through dose for a vertical five-bar structure. This is the target used to interrogate coma X (Z_7). 130

Figure 88. 95% confidence intervals for a linear regression to CD data through dose for a horizontal five-bar structure. This is the target used to interrogate coma Y (Z_8). 131

Figure 89. 95% confidence intervals for a linear regression to CD data through dose for horizontal T-bars. This is the target used to interrogate trefoil X (Z_{10}). 131

Figure 90. 95% confidence intervals for a linear regression to CD data through dose for vertical T-bars. This is the target used to interrogate trefoil Y (Z_{11}). 132

Figure 91. Wavefront extracted from NXE3300 with all third-order (Z_5 - Z_{11}) 133

Figure 92. Wavefront extracted from IMEX NXE:3300 with all third-order terms. (Z_5 - Z_{11}). 135

Figure 93. Wavefront extracted from IMEC NXE:3300 with all third-order terms and their fifth-order counterparts. (Z_5 - Z_{16} , Z_{19} , Z_{20}) 137

Figure 94. Bausch & Lomb anamorphic lens. (top) The lens in its purpose-built stand, and (bottom) imaging the entrance pupil of the lens. 140

Figure 95. A comparison of the images of the same object as an anamorphic projection lens is rotated. The top shows the effective demagnification in the horizontal and vertical pupil axes. 141

Figure 96. Two crossed cylindrical lens forming a stigmatic image. This is the simplest anamorphic optical system..... 142

Figure 97. Schematic view of the anamorphic objective patent lens used for simulation of residual wavefronts shown in the XZ plane..... 143

Figure 98. Orthogonal function expansions of residual wavefronts for (left) an isomorphic singlet and (right) the anamorphic objective patent lens. The residual wavefront was computed in ZEMAX OpticStudio in both the ideal and detuned states..... 144

Figure 99. The 3rd and 5th order aberrations of an anamorphic optical system. The names provided are the closest isomorphic aberrations. 148

Figure 100. Interferograms for the third-order anamorphic circle polynomials. 152

Figure 101. Interferograms for the third-order annular anamorphic polynomials with a 20% obscuration..... 153

Figure 102. (a) Orthogonal function expansions of residual wavefronts the anamorphic objective patent lens in both a Zernike basis and the proposed anamorphic basis. The residual wavefront was computed in ZEMAX OpticStudio in both the ideal and detuned states. Note that defocus Y and spherical Y do not exist in the Zernike polynomials, Zernike defocus and spherical is shown as being in the X direction in this plot. Similarly, astigmatism 90° does not exist in the anamorphic basis. (b) Transfer matrix between Zernike polynomials and the anamorphic primary aberrations. The stars are placed in the terms which have the largest magnitude for each particular term. 154

Figure 103. Screenshot of software written to generate anamorphic pupil filters. 155

Figure 104. Screenshot of the same software showing the simulation set capabilities of the software..... 156

Figure 105. The binary targets used to interrogate aberrations both anamorphic and isomorphic aberrations. 158

LIST OF TABLES

Table 1. The primary Zernike polynomials in both polar and Cartesian coordinates.....	38
Table 2. Phase and amplitude RMSE between $\epsilon_1(\mathbf{x})$ and eigenfunctions computed via PCA.....	86
Table 3. Standard deviation of CD measurements for each structure type	101
Table 4. Aberrations extracted from confidence intervals based around the ideal CD (no aberration) with noise given by the standard deviations in Table 3.....	103
Table 5. Recommended settings for wavefront extraction via QUIP	111
Table 6. CDs of the metrology targets used for pupil function extraction on SHARP...	114
Table 7. Required polynomial orders, and their associated RMSE for phase and amplitude variation, for each Zernike polynomial.	115
Table 8. A comparison of the extracted phase and amplitude variation using both the iterative algorithm and QUIP algorithm.....	115
Table 9. CDs of the metrology targets used for pupil function extraction on AIMS EUV.	119
Table 10. Extracted Zernike aberration values for each aberration type for the AIMS EUV.....	122
Table 11. Additional pitches used to measure third- and fifth-order aberrations from AIMS EUV.	124
Table 12. Extracted Zernike aberration values for each aberration type for the AIMS EUV.....	126
Table 13. CDs of the metrology targets used for pupil function extraction on NXE3300.	128
Table 14. Summary of the exposed wafers.....	129

Table 15. Extracted Zernike aberration values for each aberration type for the IMEC NXE:3300. The margin of error was obtained by extracting aberrations from the confidence intervals on the CD difference of metrology structures.....	132
Table 16. Extracted Zernike aberration values for each aberration type for the IMEC NXE:3300.....	134
Table 17. Additional pitches used to measure third- and fifth-order aberrations from the IMEC NXE:3300 exposure system.....	136
Table 18. Extracted Zernike aberration values for each aberration type for the IMEC NXE:3300.....	137
Table 19. Analytic solutions for the first 40 anamorphic circle polynomials.....	149
Table 20. Analytic solutions for the first 18 annular anamorphic polynomials.....	150
Table 21. Results of ANOVA for DOE to determine the sensitivity of T-Bar targets to A ₁₁ –A ₁₄ . Fields marked with an X were found to be statistically significant.....	158
Table 22. Annular anamorphic polynomials for $\epsilon = 0.0$	169
Table 23. Annular anamorphic polynomials for $\epsilon = 0.05$	170
Table 24. Annular anamorphic polynomials for $\epsilon = 0.10$	171
Table 25. Annular anamorphic polynomials for $\epsilon = 0.15$	172
Table 26. Annular anamorphic polynomials for $\epsilon = 0.20$	173
Table 27. Annular anamorphic polynomials for $\epsilon = 0.25$	174
Table 28. Annular anamorphic polynomials for $\epsilon = 0.30$	175
Table 29. Annular anamorphic polynomials for $\epsilon = 0.35$	176

LIST OF ACRONYMS

ADT	Alpha Demo Tool
AIC	Akaike Information Criterion
ART	Aberration ring test
ARTEMIS	Aberration ring test exposed with multiple illumination settings
ASB	Aberration separable basis
CCD	Charge coupled device
CD	Critical dimension
DOF	Depth of focus
DUV	Deep ultraviolet
ENIAC	Electronic Numerical Integrator and Computer
EUV	Extreme ultraviolet
EUVL	Extreme ultraviolet lithography
FEM	Focus exposure matrix
HVM	High volume manufacturing
IC	Integrated circuit
ILIAS	Integrated Lens Interferometer at Scanner
ITRS	International Technology Roadmap for Semiconductors
LPP	Laser produced plasma
NA	Numerical aperture
NILS	Normalized image log slope
NRMSE	Normalized root mean square error
OTF	Optical transfer function
PCA	Principal component analysis
PSF	Point spread function
RMS	Root mean square

RMSE	Root mean square error
SEM	Scanning electron microscope
SHARP	SHARP High-NA Actinic Reticle Review Project
SOCS	Sum of coherent systems
SVD	Singular value decomposition
TAMIS	TIS at multiple illumination settings
TCC	Transmission cross coefficient
TIS	Total integrated scatter; Transmission image sensor

1. INTRODUCTION

Just one year before the invention of the bipolar junction transistor the ENIAC computer was presented to the public. It was a state of the art computing machine, and a tremendous engineering endeavor—containing about 18,000 vacuum tubes and weighing over 30 tons.¹ Over its nine-year life ENIAC computed more than humans had prior to its creation. Despite its achievements though, the invention of the transistor, and subsequently the integrated circuit, quickly made ENIAC and its vacuum tubes obsolete.¹

The integrated circuit (IC) allowed the same circuits to be produced in a smaller, more efficient, and more reliable package.² ICs enable transistors, resistors, and capacitors to be fabricated in close proximity on the same semiconducting substrate. In 1965 Gordon Moore observed that the number of transistors in ICs had doubled roughly every two years and predicted that it would continue to do so for at least another decade.³ This observation—which has now become known as Moore’s law—became an almost prophetic driving force for the microelectronics industry.

To continually increase the number of transistors in a given area the transistor size must be scaled smaller. This has the additional benefit of decreasing the operating voltage, and therefore the power consumption.^{4,5} The features on the wafer are formed via a microelectronic manufacturing process called photolithography.¹ Thus, to continue increasing the number of transistors the minimum resolution of photolithographic processes must continue to decrease.⁶

1.1 Photolithography Systems

The goal of a photolithography process is to transfer the image of a mask into a relief pattern on a wafer. A photosensitive film called photoresist is cast across a substrate and then selectively exposed to radiation. The radiation initiates a photochemical reaction, which allows the relief structure to be formed in a developing solution.⁷ For example, the radiation may cause a change in dissolution properties, thus allowing the exposed areas to be removed with a developing solution.

Photolithography systems rely on an optical setup similar to a microscope, as shown in Figure 1. A Köhler illumination system is used so that light from a source point evenly illuminates the mask.⁶ The mask contains the pattern which is going to be transferred to the wafer, and it must block light in some areas and not in others. The mask diffracts the light from the source, which is subsequently collected by the objective lens system. Finally, the objective lens system focuses the light into the photoresist film.^{6,8}

It is common in microelectronic manufacturing to use periodic patterns because this makes the photolithography process simpler. The space over which the pattern repeats is called the pitch. The minimum resolvable pitch by a photolithographic process is given by,

$$P = 2k_1 \frac{\lambda}{n \sin \theta} = 2k_1 \frac{\lambda}{NA}, \quad \mathbf{1.1}$$

where k_1 is a process dependent factor, λ is the wavelength, and NA is the numerical aperture, which characterizes the angle of the most extreme ray entering the optical

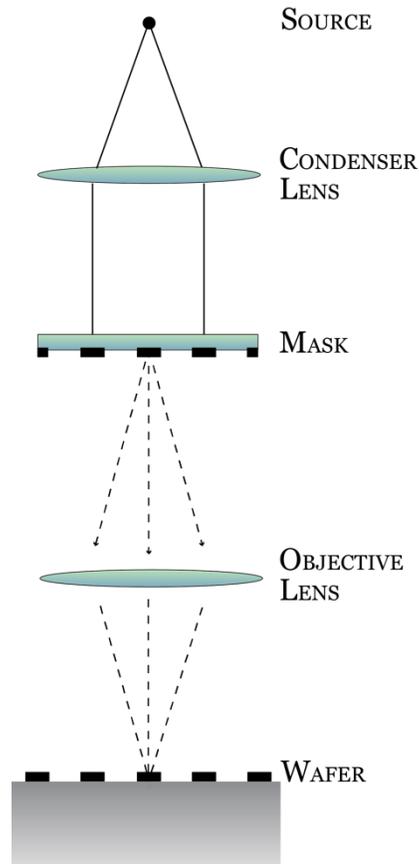


Figure 1. An example of a projection lithography system used for relief pattern transfer to a wafer.

system.^{6,9} This relationship is called the Rayleigh criterion and shows that there are two ways to make smaller features on the wafer: 1) decrease wavelength, 2) increase NA.

Photolithography processes used a mercury arc lamp source for over thirty years.¹⁰ The spectrum of this type of source is shown in Figure 2. Broadband radiation was used until the 1980s, when the spectral peak at 436 nm was isolated. Changing the wavelength requires new materials, new lenses, and new processes, so there is an economic incentive to not change wavelengths often. The NA of these 436 nm systems was increased until around 1990 when the industry moved to the peak at 365 nm.¹⁰ In the mid-1990s excimer lasers had matured enough to be used in high volume manufacturing.

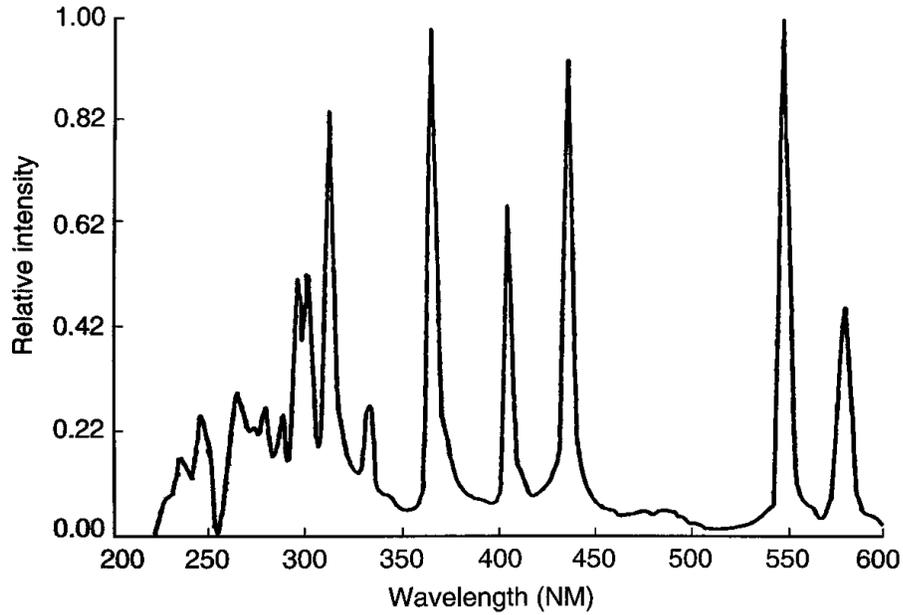


Figure 2. Typical spectrum of a mercury arc lamp.⁶

KrF excimer lasers at a 248 nm wavelength were used through the early 2000s when they were surmounted by ArF excimer lasers at a 193 nm wavelength.¹⁰ This is the historical trend of the photolithography industry: decrease the wavelength, increase NA.

Increasing NA is not necessarily a trivial matter though. For example, to offer superior imaging performance to a lower NA lens, aberrations must be sufficiently controlled. Additionally, the depth of focus for a lithography system is given by,

$$DOF = \pm k_2 \frac{\lambda}{NA^2}, \quad 1.2$$

where k_2 is another process dependent factor.^{6,8} While increasing NA allows for smaller features to be resolved it also decreases the depth of focus. This limits the thickness of photoresist and allowable wafer topography.

These trade-offs are the reason why resolution enhancement techniques, such as off-axis illumination, were developed. As seen in Figure 3, using an off-axis source tilts the diffraction orders with respect to the optical axis.¹¹ This allows the objective lens to collect diffracted energy that may not be collected if an on-axis source were used. This type of illumination also has the benefit that the depth of focus will be infinite if the 0th and 1st diffraction orders are symmetric about the lens, assuming the lens is ideal.⁶

Typically, a collection of mutually incoherent off-axis points will be used to illuminate the reticle in a lithography system. Some common configurations of off-axis illumination are given in Figure 4. The choice of source shape is largely pattern dependent, as each shape has its own unique advantages and disadvantages.

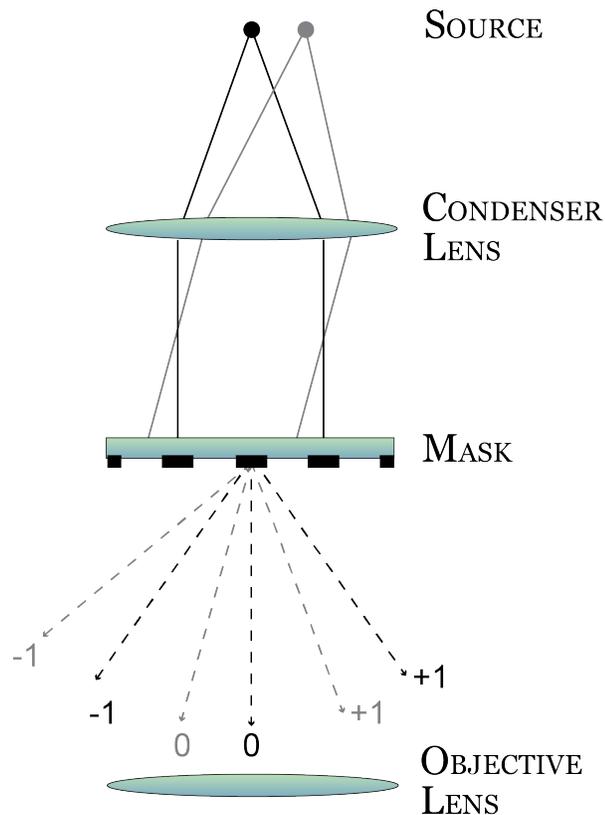


Figure 3. Comparison of on-axis illumination and off-axis illumination for a single coherent point source.

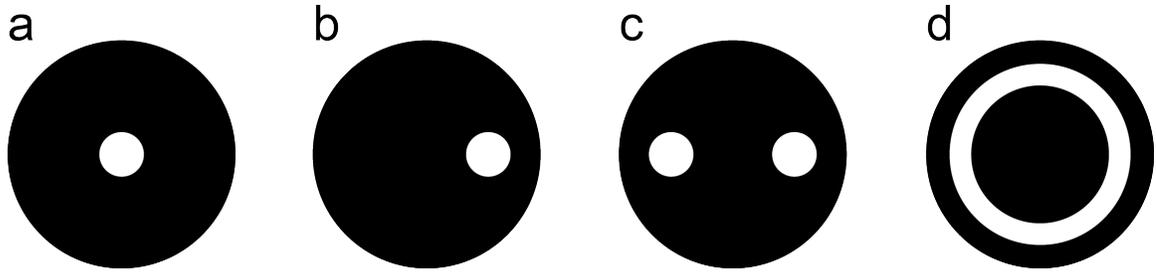


Figure 4. Common lithography source shapes, a) conventional centered monopole, b) offset monopole, c) dipole, d) annulus

The International Technology Roadmap for Semiconductors (ITRS)—a consortium which creates goals to push the semiconductor industry forward—predicts that to progress with continued scaling a major change needs to take place.¹² Figure 5 shows a comparison of four different next-generation patterning technologies. All of these technologies, with the exception of extreme ultraviolet (EUV) lithography, represent a move away from traditional wavelength scaling.

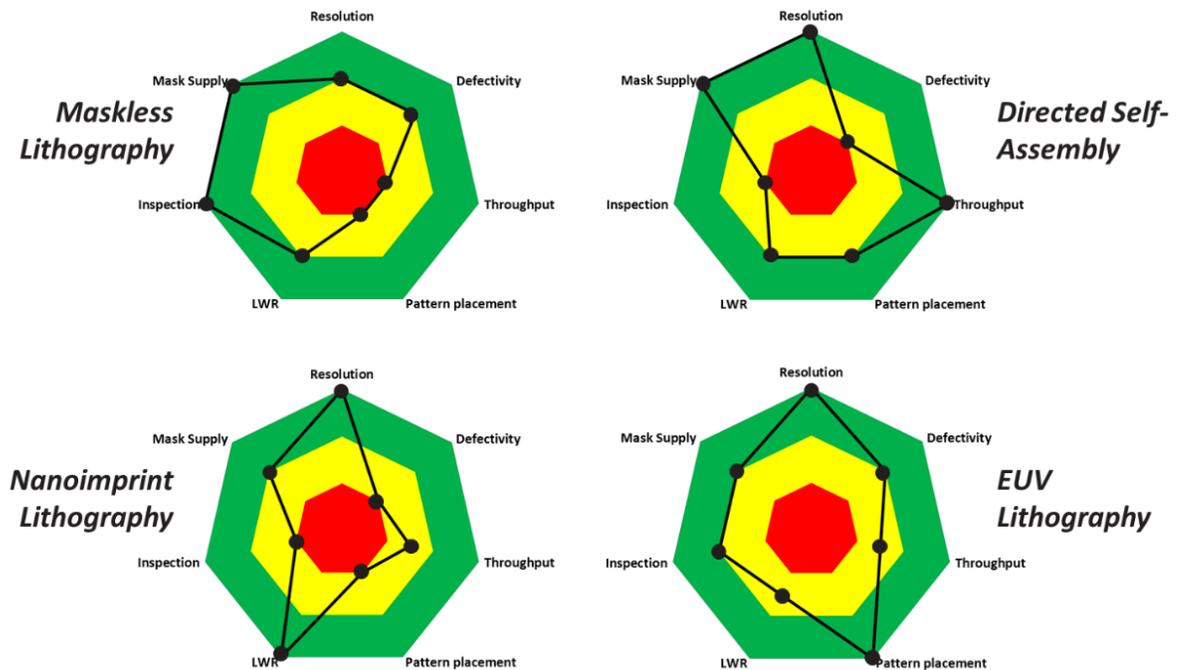


Figure 5. ITRS evaluation of next-generation patterning methods.¹²

At the present moment, none of these technologies are mature enough to be inserted into high-volume manufacturing (HVM). There is however a high motivation for EUV lithography (EUVL) to play a role in future patterning, as the industry has made a significant financial investment in this technology. State of the art HVM patterning requires multiple photolithographic exposures to obtain smaller pitches. There are various forms of multiple patterning, but they all significantly increase the cost per layer. The ITRS patterning requirements for the next ten years are given in Figure 6 and show that there is a point where the costs of multiple patterning exceed the costs of even a low throughput EUVL process.^{12,13} Therefore, it is reasonable to expect that EUVL will play a key role in future HVM patterning despite the challenges it faces.

	2014	2016	2019	2021	2024	2026
Leading IDM/foundry	N14	N10	N7	N5	N3	N1.5
Ground rules CPP=x0.78, MP=x0.65						
Contacted poly pitch – [nm]	70	52	42	32	25	25
	LELE	LELE,SADP	SADP	SAQP	SAQP	SAQP
Metal pitch – [nm]	52	36	24	16	10	10
Metal patterning	LELELE, SADP	SAQP, EUV	SAQP, EUV	SAOP, DSA, EUV	SAOP, DSA, EUV	SAOP, DSA, EUV
Fin pitch – [nm]	42	27	18	12		
Fin patterning	SADP	SAQP	SAOP, DSA	SAOP, DSA	SAOP, DSA	SAOP, DSA
Vertical pitch – [nm]			18	15		
Fin width – [nm]	8	6	-	-		
Nanowire diameter – [nm]			6	5	5	5
Channel length – [nm]	22	16	12	10	10	10
Spacer width – [nm]	15	10	7	5	5	5
Contact CD - [nm]	18	16	14	12	-	-
Device structure	finFET FDSOI	finFET FDSOI	finFET Lateral NW	Lateral NW Lateral NS	Stacked Vertical NW	Stacked Vertical NW

Figure 6. ITRS patterning requirements for the next ten years. ¹²

1.2 EUV Lithography

EUVL should be considered an extension of optical lithography, though it provides an entirely unique set of challenges. The first challenge is creating sufficient radiation for photoresist exposure. Although several types of sources are being considered, the strongest contender is the laser produced plasma (LPP) source, as shown in Figure 7.^{14,15} A CO₂ laser beam creates a high density Sn plasma where excited electrons emit EUV radiation when relaxing to their ground state. The radiation from the source is then formed into a beam by the collector mirror.

An example of an EUVL optical system is shown in Figure 8. There are no materials which refract EUV wavelengths ($\lambda=13.5$ nm), so EUVL systems must use reflective optics.¹⁴ The non-zero chief ray angle, caused by the requirement to use reflective optics, is the source of many challenges which distinguish EUVL from its longer wavelength optical counterparts.¹⁶

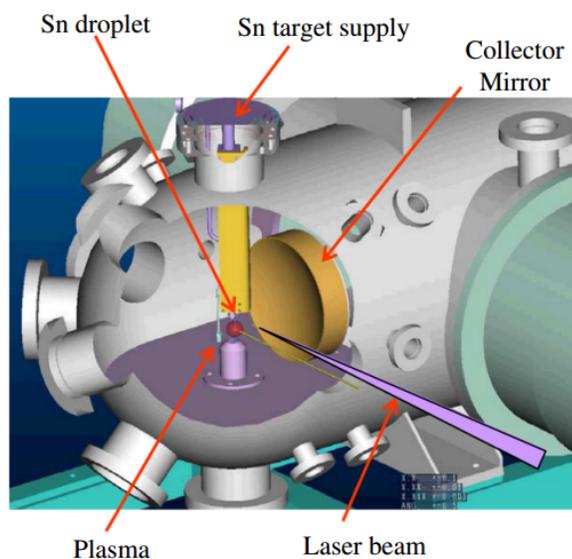


Figure 7. Laser produced plasma EUV source.¹⁵

The mirrors in an EUVL system are Bragg reflectors fabricated using alternating layers of molybdenum and silicon. Some of the incident radiation is reflected at each interface, so significant reflectivity can be obtained by using many pairs of layers engineered to maximize constructive interference.¹⁷⁻¹⁹ The reticle is constructed in a similar way and tantalum nitride (TaN) is used as an absorbing material, as shown in Figure 9. This mirror configuration may only have a peak reflectivity around 70%, and is highly dependent on the incident angle, as shown in Figure 10. In a system with such mirrors and 13 reflections only 0.97% of the source power makes it to the wafer plane. The rest of the source power is distributed throughout the system as heat and scattering. This heat is expected to contribute to system drift and thermal dynamic aberrations.

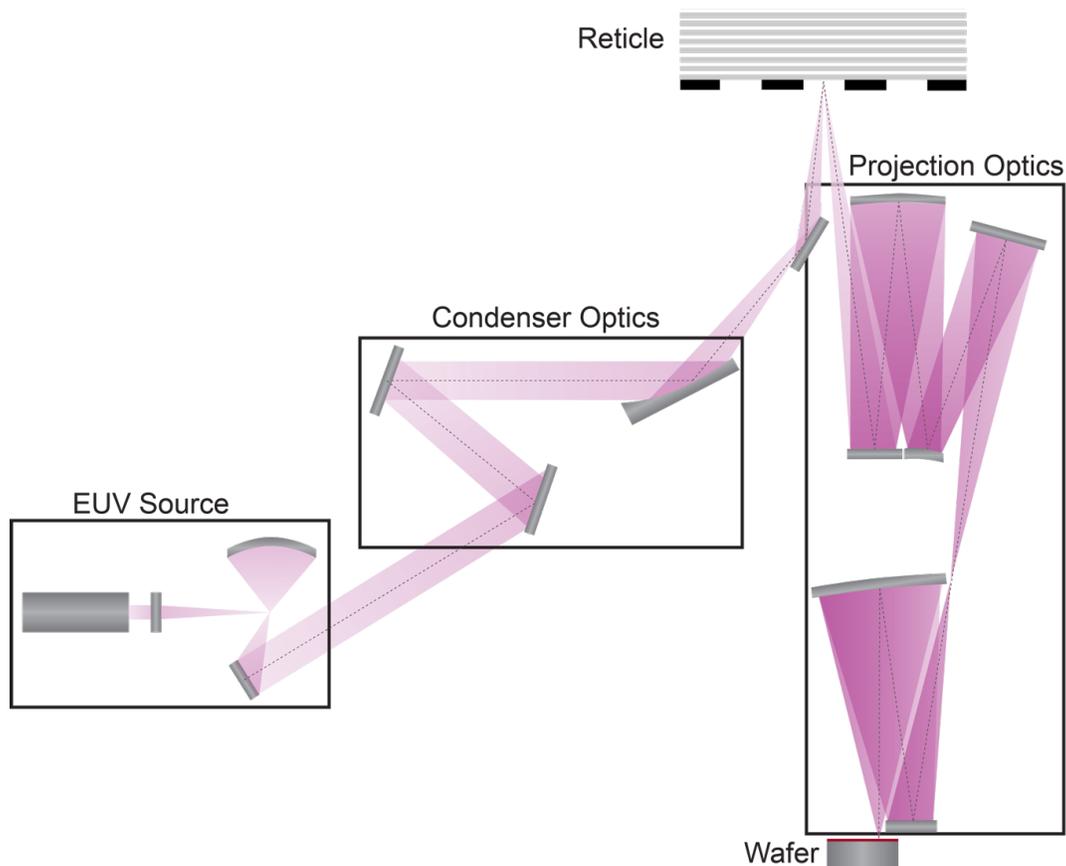


Figure 8. Schematic view of a typical EUV lithography system. The dashed black line shows the chief ray.

Further complications arise when considering high-NA EUV lithography. Due to the non-orthogonal beamline there is a maximum NA for a given chief ray angle such that the incident and diffracted cones of light do not overlap, as shown in Figure 12a and 12b.^{16,23} In order to increase NA the chief ray angle must be increased as well, as in Figure 12c, but this is non-ideal due to the decrease in reflectivity with incident angle. In order for EUVL to be a viable technology there must be a solution to high-NA imaging.

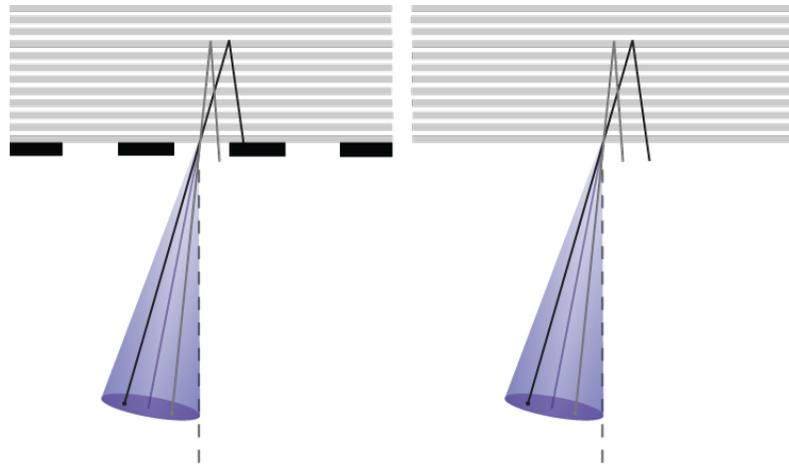


Figure 11. Shadowing in an EUVL system. Horizontal features (left) appear to increase in size while vertical features (right) do not have this problem. The right panel is in the plane of a space feature.

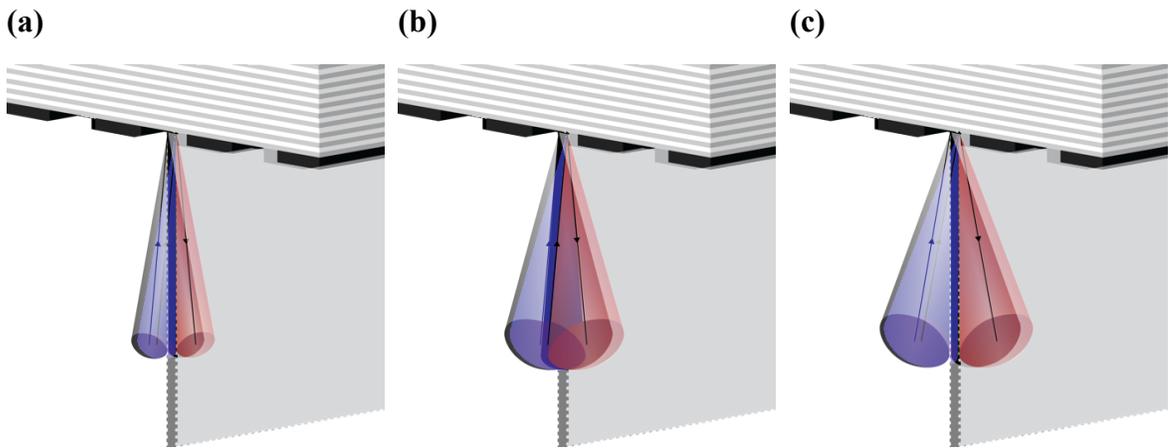


Figure 12. Geometrical considerations of high-NA EUVL. a) Reflective system close to the maximum NA, b) Larger NA with the same chief ray angle, c) the chief ray angle must increase to allow a higher-NA

1.2.1 Anamorphic systems

The solution to the problem of high-NA lithography is to use an anamorphic optical system. Anamorphic systems have different magnifications in different axes of the lens.^{23,24} Historically these systems have primarily been used to fit a wide aspect ratio motion picture as a nearly square frame on a film strip.^{24,25} A comparison of looking directly into an isomorphic projection lens and an anamorphic projection lens is given in Figure 13a and b. The entrance pupil is defined as the image of the aperture stop seen from the front of the optic, while the exit pupil is the image as seen from the back of the optic.^{26,27} Asymmetric magnification causes the entrance pupil of the anamorphic lens to appear as an ellipse. Looking from the back of the lens, the pupil appears to be circular, as in Figure 13c. The test pattern shown in all three panels of Figure 13 is elliptical—as it appears through the isomorphic lens—but appears to be circular through the anamorphic lens.²⁸

This is useful in an EUVL system because the NA can be geometrically increased in one direction and effectively increased in the other direction via a higher magnification.^{23,29} An example of anamorphosis in an EUVL system is shown in Figure 14. To maintain the

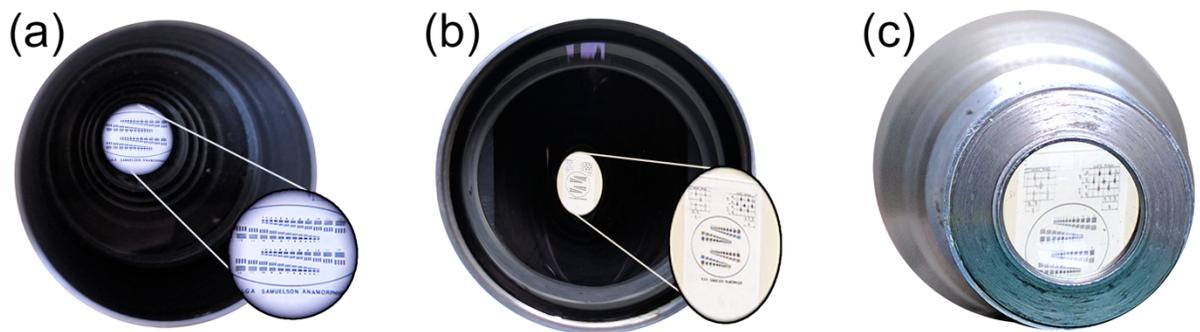


Figure 13. Comparison of the entrance pupils of an isomorphic projection lens (a) and an anamorphic projection lens (b), and the (c) exit pupil of the same anamorphic lens

same feature size as in Figure 12 the reticle pattern must be stretched. A full field now occupies twice the space on the reticle in one direction as compared to an isomorphic system. Therefore, anamorphic lithography systems would require two half-field exposures for one equivalent isomorphic exposure field.²⁹

The NA is identical in both axes of the lens and the incident and diffracted beams do not overlap. Because the entrance and exit pupils are different shapes, anamorphic lenses have different properties than traditional lenses and must be analyzed separately from isomorphic lenses. Anamorphic imaging is expected to be implemented in next generation EUVL systems.

In both isomorphic and anamorphic systems, increasing NA, by definition, means increasing the angular spread of light through the optic. The last mirror in the optical train of current generation EUVL systems is tilted so that the second to last mirror does not block the optical path. However, this is not possible with the increased angles at higher numerical apertures. At the most extreme angles required for $NA > 0.5$, the multilayer reflectivity is near zero. The solution to reduce the angular spread, yet still maintain high-NA, is to create a hole in the second to last lens element—a central obscuration. As of the date of writing the size of the central obscuration has not been published. However, ASML has determined that it should be smaller than 35% of the pupil radius. There are certain angular frequencies beyond this point which the optic can sufficiently resolve.³⁰

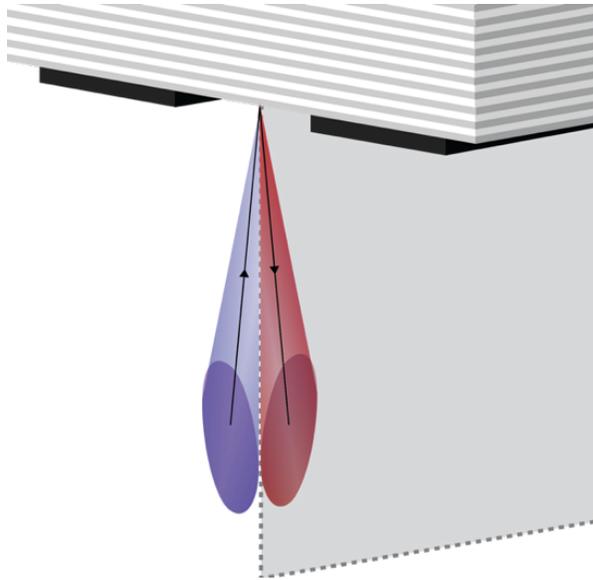


Figure 14. An example of an anamorphic EUVL system

1.3 Lens Quality Metrics

So far we have presumed that all lenses are ideal. Even ideal lenses have aberrations, but there are also many possible non-idealities in real lenses. It is therefore necessary that we can establish and quantify the quality of an optic. This section will introduce non-idealities in optics and several metrics for judging optical quality.

1.3.1 *Point Spread Function*

The transfer of light through a linear optical system can be fully characterized by the point spread function. If the point spread function (PSF) of a system is known, the image can be determined by convolution of the PSF with the object, as in Figure 15.³¹ The PSF is defined as the image of a coherent point object. The PSF could therefore be measured via the image of a contact hole—a common lithography pattern. Photoresist forms a relief structure though and will not record the full PSF. It is therefore impossible to directly measure the PSF of a lithography system through photoresist exposure. Further, most lithography systems are not even linear due to the use of partial coherence.⁶

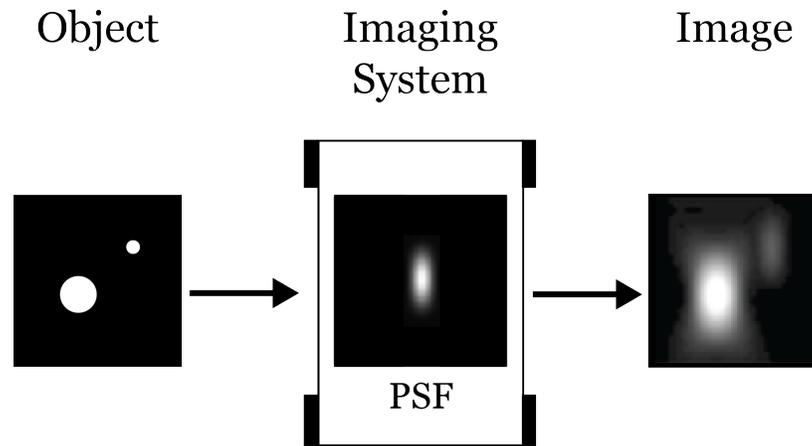


Figure 15. Transfer of light through a linear imaging system via the point spread function.

The Fourier transform of the PSF is called the optical transfer function (OTF) and is often a more convenient form for system analysis.³¹ The OTF and PSF of an ideal optical system are shown in Figure 16. The ideal PSF is called the Airy disk, and its contrast has been stretched in the two-dimensional representation of Figure 16. Stretching the contrast shows the characteristic sharp peak surrounded by concentric rings, also seen in the $y = 0$ cutline of the Airy disk in Figure 16.

From a more physical perspective, the transfer of light through an optical system can be characterized by amplitude and phase errors in the system's exit pupil.²⁶ In the case of a linear optical system, the transfer is characterized by the pupil function. The PSF is the impulse response of a linear optical system and can therefore be computed from the system's transfer function. The pupil function is useful even in non-linear systems because the OTF can be computed from it. A simple way to quantify the quality of an optic is the Strehl ratio given by,

$$S = \frac{\left| \int_{r=0}^1 \int_{\theta=0}^{2\pi} P(r, \theta) r dr d\theta \right|^2}{\left| \int_{r=0}^1 \int_{\theta=0}^{2\pi} P_0(r, \theta) r dr d\theta \right|^2}, \quad 1.3$$

where $P(r, \theta)$ is the actual pupil function, and $P_0(r, \theta)$ is the ideal pupil of the same system.³² The Strehl ratio should be as close to unity as possible for ideal imaging.

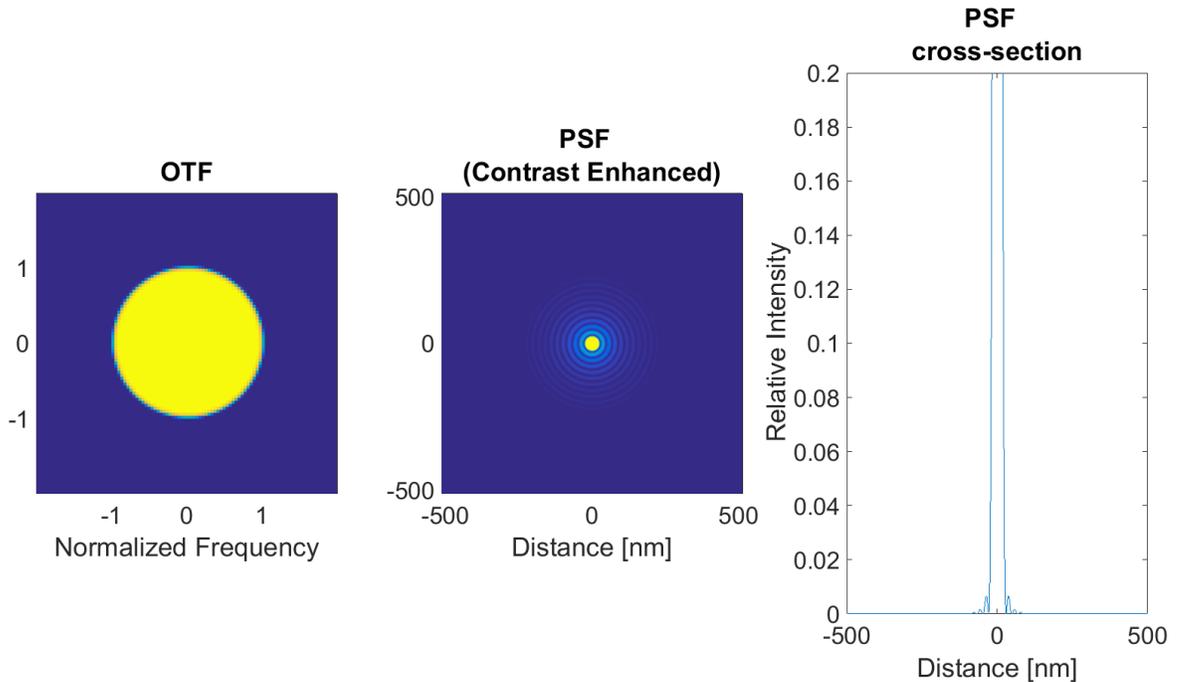


Figure 16. The optical transfer function and point spread function of an ideal pupil

1.3.2 Apodization

Apodization is a technique where the transmission of a lens' pupil is varied from the ideal top-hat profile.³² We will use the term apodization to refer only to intentional amplitude variation. The OTF and PSF of an apodized optical system is shown in Figure 17. The pupil is defined by the function,

$$P(r) = e^{-3r^2} \quad 0 \leq r \leq 1 \quad , \quad 1.4$$

so that the transmission is just 5% at the edge of the pupil. Rings are not visible in the two dimensional PSF, despite being processed identically as that in Figure 16. In the cut-line through the $y = 0$ plane it is clear that the rings have been removed through the use of apodization—hence the literal translation of “removing the foot”. Apodization improves through-focus performance of a lens and has uses in numerous imaging applications but is not used in lithography.^{33,34} Therefore, any apodization of a lithography system is non-ideal and will instead be referred to as amplitude variation.

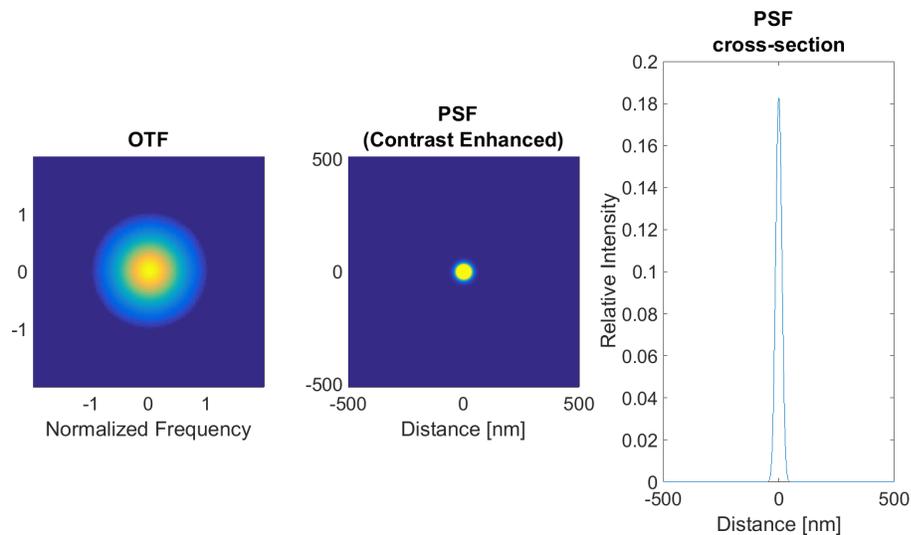


Figure 17. The optical transfer function and point spread function of a Gaussian apodized pupil

1.3.3 Aberrations

Traditionally, aberrations are defined to be any error which causes a point object to not be imaged as a point. This type of error is caused by phase variation in the pupil.^{26,32} Marechal noted the following inequality for pure phase variation,^{35–37}

$$S \geq \frac{1}{\pi^2} \left| \int_{r=0}^1 \int_{\theta=0}^{2\pi} \cos \frac{2\pi}{\lambda} \Phi(r, \theta) r dr d\theta \right|^2, \quad 1.5$$

where $\Phi(r, \theta)$ is the phase error caused by aberrations. For small aberrations this can be simplified by a Taylor expansion as,

$$S \geq \left(1 - \frac{2\pi^2}{\lambda^2} \sigma \right)^2, \quad 1.6$$

where σ is the mean of the squared phase function, $\Phi^2(r, \theta)$.³² This inequality therefore relates the Strehl ratio to RMS wavefront error.

Historically several authors have attempted to quantify the maximum wavefront error under which an optical system can be considered diffraction limited. Lord Rayleigh proposed that a system can be considered diffraction limited if the peak-valley wavefront error does not exceed $\lambda/4$.^{32,38} This results in different wavefront RMS for different types of aberrations, but corresponds to a range of Strehl ratios between 0.8 and 0.92 for the primary aberrations. Later Maréchal proposed that the Strehl ratio should be greater than 0.82, corresponding to a wavefront RMS of approximately $\lambda/14.5$.³⁵⁻³⁷

When aberrations are present in a lithography system they primarily affect CD and image placement error.⁶ It is therefore prudent to evaluate these criteria in the context of lithography. We will use the Maréchal criterion to evaluate the performance of a lithography system with primary aberrations. In a purely astigmatic system the Maréchal criterion sets a limit of $169 m\lambda$. This leads to a system where with just 30 nm of defocus 25 nm horizontal lines resolve while vertical lines do not. (e.g. an EUVL system with 0.33 NA and a conventional circular illuminator with 0.25σ) This difference in useable

depth of focus is likely incompatible with state of the art lithography processes. In the same system but with primary coma, fully dense lines have a placement error of 1.0 nm. Further, the placement error increases to over 7 nm for fully dense 38 nm lines and spaces. The ITRS specifies an overlay error of less than 3.4 nm for the 7 nm node.³⁹ This placement error is nearly a third of the overlay budget in the best-case scenario, and is therefore not tolerable in a lithography system.

Aberration control and characterization in a state of the art photolithographic lens have the tightest tolerances of any optical system. State of the art DUV lithography systems have wavefront RMS below 2 nm ($\lambda/100$). Because wavefront error scales with wavelength, EUV lithography has an even tighter tolerance. Some authors have estimated the required wavefront tolerance to below 500 pm RMS ($\lambda/27$) for EUVL optics.⁴⁰

1.3.4 Flare

An ideal mirror will reflect an incident point of light at the same angle, but a non-ideal mirror only reflects *most* of the light at the same angle. The rest of the light is scattered at some other angles.⁴¹ In lithography any scattered light which makes it to the image plane is called flare. The total amount of light scattered by a surface can be quantified by the total integrated scatter (TIS) which is given by,⁴²

$$TIS = R \left(1 - e^{-(4\pi\sigma \cos \theta/\lambda)^2} \right), \quad 1.7$$

where R is the reflectance of the optic, σ is the RMS roughness of the surface, and θ is the incident angle. Flare was a concern in DUV lithography but will be even worse in

EUVL systems because TIS varies as λ^{-2} . The total scatter will be 137 times worse in an EUVL system than a ArF DUV system given identical roughness.

To determine the effect of flare on imaging, surface roughness must be considered as a random process. The roughness can be measured experimentally and its autocorrelation function can be computed.⁴³ The power spectral density of the roughness characterizes the frequency content of the roughness and is the Fourier transform pair of the roughness autocorrelation function. Assuming perfect condenser optics, the image intensity with rough projection optics can be computed as,

$$\langle I(x) \rangle = PSF_{scatter} \star I_0, \quad 1.8$$

where $\langle I(x) \rangle$ is the ensemble average of the image intensity, I_0 is the ideal intensity without scattering, and $PSF_{scatter}$ is the point spread function of the scattering.⁴⁴ This depends on the characteristics of the roughness. Assuming the roughness is characterized by a zero-mean, stationary, ergodic, and Gaussian distributed variable, the PSF can be given by,

$$PSF_{scatter} = \mathcal{F}^{-1} \left\{ e^{-\sigma_p^2} e^{\sigma_p^2 C_\phi} \right\}, \quad 1.9$$

where σ_p is the RMS roughness, C_ϕ is the roughness autocorrelation function, and \mathcal{F} is the Fourier transform operator.⁴⁴ Gaussian roughness therefore blurs the image intensity, resulting in a loss of contrast.

Roughness in EUVL optics can be the millimeter regime, so modeling, characterization, and compensation is critical to image process optimization.⁴³ Typically,

the flare PSF is convolved with the ideal aerial image to obtain the effect of flare on the aerial image, which is called a flare map. This flare map can then be used to compensate on the mask for flare.

2. BACKGROUND

2.1 Theory of Image Formation

To define the inverse imaging problem, we will first consider the forward imaging problem. The geometrical theory of aberrations will be formalized. Next the wave theory of aberrations will be established.

2.1.1 Geometrical Image Formation

We consider a general imaging system as shown in Figure 18. Light travelling from a point on the potentially extended source is imaged onto the mask. This light diffracts at the mask, creating a new wavefront that serves as the input of the imaging system. The system transforms that light in a unique way and outputs a new wavefront, which can vary in amplitude and phase from the ideal wavefront. In the ideal case the new wavefront converges to a point in the image plane.^{26,27}

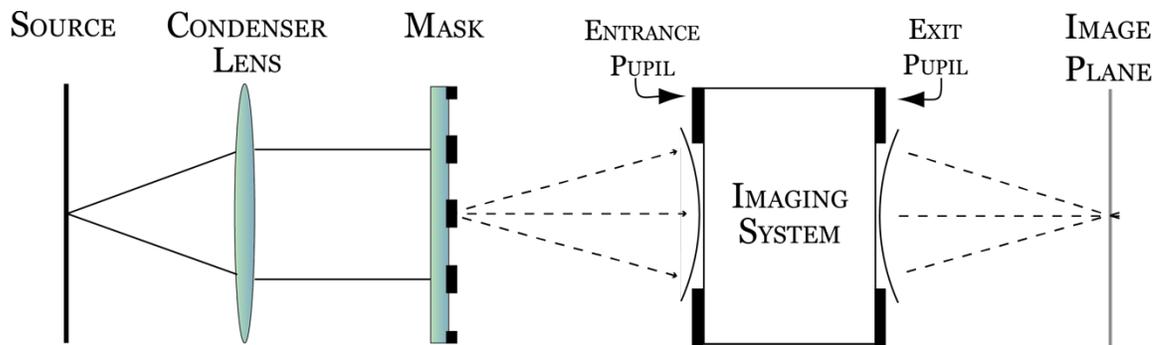


Figure 18. An example of the type of imaging system being studied.

The coordinate systems which will be used for analysis of this imaging system are shown in Figure 19. Let P_1^* be the Gaussian image of an object located at P_0 . The image

at P_1^* is formed by a spherical wavefront centered at that point and going through the center of the exit pupil plane, O_1' . In the presence of aberrations the image of an object at P_0 may form at some different point P_1 .

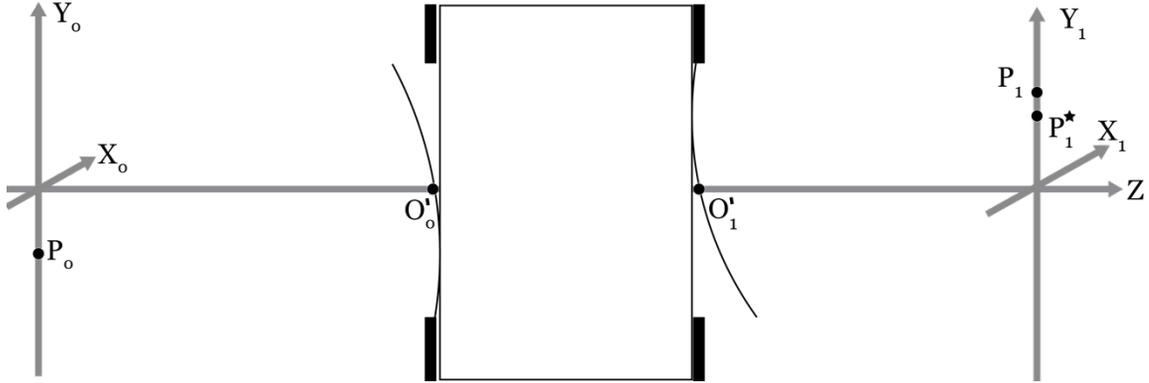


Figure 19. Coordinate systems in the object and image planes.

The vector $\overline{P_1^*P_1}$ characterizes the effect of an aberration from the perspective of the rays, and is therefore called the ray aberration.⁴⁵ Likewise, the difference between the actual wavefront W and the Gaussian reference wavefront S is given by the optical path difference,

$$\Phi = [Q\bar{Q}] , \quad 2.1$$

where \bar{Q} and Q are the intersection of the ray $P_1^*P_1$ with S and W , respectively, as shown in Figure 20. The wave aberration can be rewritten as,

$$\Phi = [P_0Q] - [P_0\bar{Q}] . \quad 2.2$$

Because the center of the exit pupil is on the same wavefront as \bar{Q} ,²⁶

$$\begin{aligned} X_1^* &= MX_0 \\ Y_1^* &= MY_0 \end{aligned} \quad . \quad \mathbf{2.6}$$

Therefore, the wave aberration is only a function of the coordinates (X_0, Y_0, X_1, Y_1) .

For now, our optical system is defined to be rotationally invariant. Rotational invariance is useful in optical systems because a conventional lens' aberrations should be independent of its rotation. This requirement means that the wavefront aberration must also be rotationally invariant. If the system is rotated through an angle θ we can relate the new rotated coordinates to the unrotated coordinates by the relations,

$$\begin{aligned} X' &= X \cos \theta - Y \sin \theta \\ Y' &= X \sin \theta + Y \cos \theta \end{aligned} \quad . \quad \mathbf{2.7}$$

We now consider the quantity $(X^2 + Y^2)$, which is given in the rotated system by,

$$X'^2 + Y'^2 = X^2 \cos^2 \theta + Y^2 \sin^2 \theta + Y^2 \cos^2 \theta + X^2 \sin^2 \theta = X^2 + Y^2 \quad . \quad \mathbf{2.8}$$

The quantity $(X^2 + Y^2)$ has the same functional form regardless of the rotation, and therefore it is rotationally invariant. It can also be shown that the quantity $(X_0X + Y_0Y)$ is rotationally invariant. We can therefore expect that a power series expansion of the wave aberration will include powers of $(X_0^2 + Y_0^2)$ and $(X^2 + Y^2)$ and $(X_0X + Y_0Y)$. The wave aberration can be expanded as,^{32,45}

$$\begin{aligned}
\Phi(X_0, Y_0, X, Y) = & a_0 + b_0(X_0^2 + Y_0^2) + b_1(X^2 + Y^2) + b_2(X_0X + Y_0Y) \\
& + c_0(X_0^2 + Y_0^2)^2 + c_1(X^2 + Y^2)^2 + c_2(X_0X + Y_0Y)^2 \\
& + c_3(X_0^2 + Y_0^2)(X^2 + Y^2) + c_4(X_0^2 + Y_0^2)(X_0X + Y_0Y) \\
& + c_5(X^2 + Y^2)(X_0X + Y_0Y) + \dots
\end{aligned} \tag{2.9}$$

The rays are normal to the wavefront, therefore for ray QP_1 , which passes through the points (X, Y, Z) and $(X_1, Y_1, 0)$, the direction cosines are given by,³²

$$n_1 \cos \alpha = \frac{\partial V}{\partial X} = n_1 \frac{X_1 - X}{R'}, \tag{2.10}$$

$$n_1 \cos \beta = \frac{\partial V}{\partial Y} = n_1 \frac{Y_1 - Y}{R'},$$

$$n_1 \cos \gamma = -n_1 \frac{Z}{R'}.$$

Using implicit differentiation of the Gaussian reference sphere,

$$\frac{\partial}{\partial X} [(X - X_1^*)^2 + (Y - Y_1^*)^2 + Z^2] = \frac{\partial R^2}{\partial X}. \tag{2.11}$$

Simplifying,

$$\frac{\partial Z}{\partial X} = -\frac{X - X_1^*}{Z}. \tag{2.12}$$

Through differentiation of the wave aberration, as given in Equation 2.4, we also have,

$$\frac{\partial \Phi}{\partial X} = \frac{\partial V}{\partial X} + \frac{\partial V}{\partial Z} \frac{\partial Z}{\partial X}. \tag{2.13}$$

Combining,

$$\frac{\partial \Phi}{\partial X} = n_1 \frac{X_1 - X}{R'} + \left(-n_1 \frac{Z}{R'} \right) \left(-\frac{X - X_1^*}{Z} \right) . \quad 2.14$$

We can simplify this expression and repeat the same process in Y to obtain,^{26,32,45}

$$X_1 - X_1^* = \frac{R'}{n_1} \frac{\partial \Phi}{\partial X} , \quad 2.15$$

$$Y_1 - Y_1^* = \frac{R'}{n_1} \frac{\partial \Phi}{\partial Y} .$$

These expressions connect the wave aberration to the ray aberration exactly. However, they require *a priori* knowledge of the aberrations to know R' . Nevertheless, these expressions show that aberrations must involve X and Y . Therefore, in our series expansion of the wave aberration the terms corresponding to a_0 , b_0 , c_0 , and so on do not represent aberrations.⁴⁵

The quantity $(X^2 + Y^2)$ corresponds to b_1 and represents an optical path difference which is null at the center of the pupil but varies quadratically moving towards the edges. This represents a wavefront which has a different radius than the Gaussian reference, and will therefore converge either faster or slower than the reference sphere. No matter which, the image point will be focused in a different plane than the reference.⁴⁵

The quantity $(X_0X + Y_0Y)$ corresponds to b_2 and represents a tilted wavefront. Using Equation 2.15 we can find,

$$X_1 - X_1^* = \frac{R'}{n_1} X_0 \quad 2.16$$

$$Y_1 - Y_1^* = \frac{R'}{n_1} Y_0 .$$

This corresponds to a transverse shift of the image point in the Gaussian image plane.

The quantity $(X^2 + Y^2)^2$ corresponds to c_1 and is the only term which does not vanish for $(X_0, Y_0) = (0,0)$. In expanded form this term is given as,

$$X^4 + 2X^2Y^2 + Y^4 . \tag{2.17}$$

We can now make analogies between this form and the 1st order focus and tilt errors. There are no odd-parity terms, so the image point will not be shifted within its focal plane. This term is also independent of the object coordinates and contributes quadratically increasing phase shifts, or in other words: quadratically increasing focal shifts. Thus, rays entering the lens at different heights will be focused in different planes.^{26,32,45} These effects describe the term commonly called spherical aberration.

The quantities corresponding to c_2 and c_3 are generally considered together. If we take $c_2 = c_3$, this term can be expanded to,

$$X_0^2(2X^2 + Y^2) + Y_0^2(X^2 + 2Y^2) + 2XYX_0Y_0 . \tag{2.18}$$

Again, we can now make analogies between this form and the 1st order focus and tilt errors. In this case, points in different pupil axes will be focused in different planes. This effect of orientation dependent focus errors is the common effect of astigmatism.^{26,32,45}

The term corresponding to c_4 has an odd-parity component, so we can expect it to contribute transverse image shifts. Expanding this quantity, we obtain,

$$X(X_0Y_0^2 + X_0^3) + Y(X_0^2Y_0 + Y_0^3) \quad 2.19$$

This term corresponds to an image point shift which varies cubically with the object point coordinates. This effect is also known as distortion.⁴⁵

Finally, expanding the quantity corresponding to c_5 we obtain,

$$X_0X^3 + X^2YY_0 + Y^2XX_0 + Y_0Y^3 \quad 2.20$$

As the object is moved further off-axis, the image point is shifted further from its Gaussian point in both directions. It is however shifted more in one direction, which creates the comet tail smear associated with coma.⁴⁵

2.1.1.1 Refraction at an ideal spherical interface

We now briefly consider an ideal spherical interface, as shown in Figure 21. The object and image points are given by black circles, while the vertex and center of curvature are given by gray circles. The optical path length from the object point to the image point is,

$$OPL = n_1l_o + n_2l_i \quad 2.21$$

We can analyze the triangle defined by the object point, the incident point, and the center of curvature using the law of cosines to obtain,

$$l_o^2 = R^2 + (s_o + R)^2 - 2R(s_o + R) \cos \phi \quad 2.22$$

Similarly, the length of the ray on the image side can be written as,

$$l_i^2 = R^2 + (s_i - R)^2 + 2R(s_o - R) \cos \phi \quad 2.23$$

Therefore, the optical path length can be written as,

$$OPL = n_1 \sqrt{R^2 + (s_o + R)^2 - 2R(s_o + R) \cos \phi} + n_2 \sqrt{R^2 + (s_i - R)^2 + 2R(s_o - R) \cos \phi} \quad 2.24$$

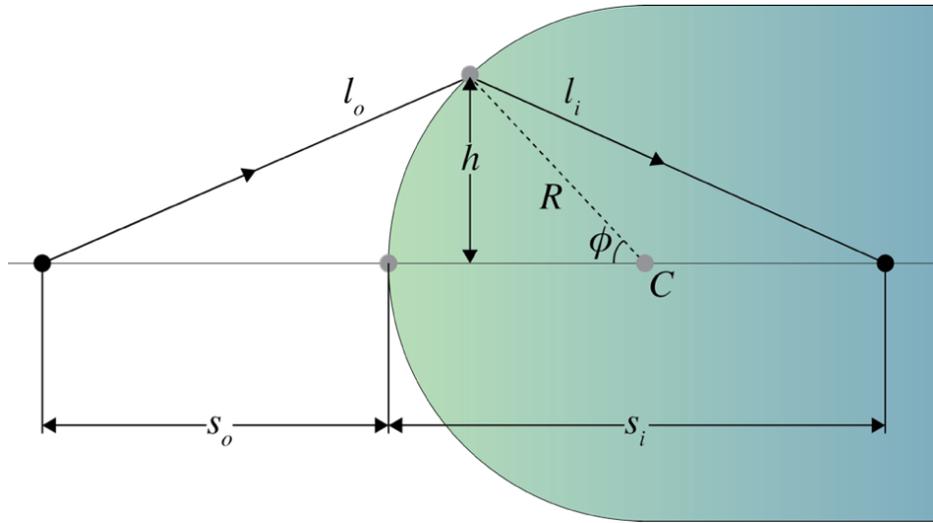


Figure 21. The geometry of an ideal spherical interface

Fermat's principle states that light travels along the path that takes the least time. In more mathematical terms, the derivative of an optical path length with respect to position is zero. In this case, the angle ϕ can be taken as the position variable so that,

$$\frac{dOPL}{d\phi} = \frac{n_1 R (s_o + R) \sin \phi}{l_o} - \frac{n_2 R (s_i - R) \sin \phi}{l_i} = 0 \quad 2.25$$

In all cases $R > 0$ and $\sin \phi > 0$ so this expression can be simplified to,

$$\frac{n_1}{l_o} + \frac{n_2}{l_i} = \frac{1}{R} \left(\frac{n_2 s_i}{l_i} - \frac{n_1 s_o}{l_o} \right) \quad 2.26$$

This expression is exact, but requires knowledge of the lengths of the rays. We can avoid that by using Equations 2.22 and 2.23. Using a Taylor expansion of $\cos \phi$ and $(1 + x)^{-1/2}$ keeping only the first two terms in both expansions, we obtain,

$$l_o^{-1} \approx s_o^{-1} - \frac{R(s_o + R)}{2s_o^3} \phi^2, \quad 2.27$$

$$l_i^{-1} \approx s_i^{-1} + \frac{R(s_i - R)}{2s_i^3} \phi^2. \quad 2.28$$

For small angles, we observe,

$$\phi \approx \sin \phi = \frac{h}{R}. \quad 2.29$$

Combining with Equations 2.27 and 2.28,

$$l_o^{-1} \approx s_o^{-1} - \frac{(s_o + R)h^2}{2s_o^3 R}, \quad 2.30$$

$$l_i^{-1} \approx s_i^{-1} + \frac{(s_i - R)h^2}{2s_i^3 R}. \quad 2.31$$

Inserting these expressions into Equation 2.26 we can find,²⁷

$$\frac{n_1}{s_o} + \frac{n_2}{s_i} = \frac{n_2 - n_1}{R} + h^2 \left[\frac{n_1}{2s_o} \left(\frac{1}{s_o} + \frac{1}{R} \right)^2 + \frac{n_2}{2s_i} \left(\frac{1}{R} - \frac{1}{s_i} \right)^2 \right]. \quad 2.32$$

This equation allows computation of the image location formed by an ideal spherical interface. In practice, it is most convenient to solve this equation numerically. The focal length of a lens is defined as the image location of an infinitely distant object, which can be approximated to the third-order as,

$$\frac{1}{f} = \frac{n_2 - n_1}{n_2 R} + \frac{h^2}{2f} \left(\frac{1}{R} - \frac{1}{f} \right)^2 \quad 2.33$$

In the paraxial, or first-order, approximation only the first term occurs, and the object is focused to a point. Including the higher order term, we see that the focal length depends quadratically on the ray height at the interface. This is exactly spherical aberration as derived in the previous section. Further, in two dimensions an off-axis object point will not be focused to a point due to oblique astigmatism. Thus, aberrations are not only due to manufacturing defect, and can be considered on a theoretical basis.^{8,27}

Aberration balancing is considered a crucial step in modern lens design. Typically, several rays are traced through the system and the optical path length from surface to surface is computed. The wavefront is defined by surfaces of constant path length, and the wavefront aberration can be determined through comparison to a Gaussian reference.⁴⁷ Aberrations can be balanced with optical components that contribute nearly equivalent aberrations in the opposite direction. For example, spherical aberration can be corrected by adding a compensation plate which contains an aspheric surface. There are several competing objective lens system designs for EUVL scanners. The composite RMS wavefront error is 0.12λ – 0.23λ in these designs at 0.25 NA.¹⁴

2.1.2 Coherent Image Formation

We now consider again a linear shift invariant optical system that transforms an input object to an output image, as in Figure 18. A monochromatic point source of light far away from an aperture creates even illumination with parallel wavefronts.

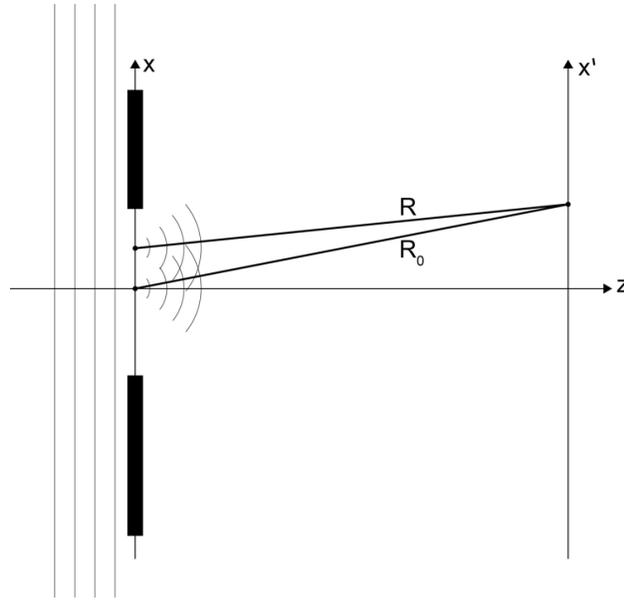


Figure 22. The geometry of the coherent diffraction problem

Via Huygen's principle we assume the electric field can be represented as a sum of point sources emitting spherical waves as,^{48,49}

$$dE(x') = E_0(x) \frac{e^{ikR}}{R} dx \quad . \quad 2.34$$

The distance from a point in the aperture to the screen is given by

$$R = \sqrt{(x - x')^2 + z^2} \quad . \quad 2.35$$

and the distance from the center of the aperture to the screen is given by

$$R_0 = \sqrt{x'^2 + z^2} \quad . \quad 2.36$$

The distance between the screen and a point can then be simplified and Taylor expanded to obtain,

$$R = R_0 \left[1 + \frac{x^2 - 2xx'}{2R_0^2} \right] \quad 2.37$$

We can conclude from Equation 2.37 that $\frac{1}{R} \approx \frac{1}{R_0}$ for large distances. The electric field in the receiving plane is then given by,

$$E(x') = \int_{x=-\infty}^{\infty} dx \frac{E_0(x)}{R_0} e^{ikR_0 \left(1 + \frac{x^2 - 2xx'}{2R_0^2} \right)} \quad 2.38$$

After some algebraic manipulation, this can be expanded to,

$$E(x') = \frac{1}{R_0} e^{ikR_0} \int_{x=-\infty}^{\infty} dx E_0(x) e^{ik \frac{x^2}{2R_0}} e^{-ik \frac{xx'}{R_0}} \quad 2.39$$

We would prefer to ignore the quadratic phase term in Equation 2.39, so we must determine where it will be negligible.⁴⁸ This occurs when that term is near unity, or,

$$\frac{kx^2}{2R_0} \ll 1 \quad 2.40$$

Equation 2.40 implies,

$$R_0 \gg \frac{\pi a^2}{\lambda}, \quad 2.41$$

where a is the half-width of the aperture. Therefore, the screen must be very far away as compared to the size of the aperture. We have already assumed that R_0 is large, so we can therefore ignore the quadratic phase term in the integrand. We also note that,

$$\frac{x'}{R_0} = \sin \theta \therefore k' \frac{x'}{R_0} = k'_x \quad 2.42$$

The x-component of the wave vector can be interpreted as a spatial frequency, which we will call u . Finally, the integral can be simplified to,

$$E(u) = \frac{e^{ikR_0}}{R_0} \int_{x=-\infty}^{\infty} dx E_0(x) e^{-i2\pi ux} \quad 2.43$$

This is the Fraunhofer diffraction integral and is a form of Fourier integral.^{48,49} Therefore, the electric field in the receiving plane of a uniformly illuminated aperture is given by the Fourier transform of the aperture. The receiving plane will therefore be called the frequency domain.

Equation 2.43 represents the electric field in the entrance pupil of the imaging system. We know from Section 2.1 that aberrations are caused by phase in the exit pupil which is characterized by the pupil function. For an ideal pupil, the pupil function simply characterizes the shape of the aperture. The electric field in the exit pupil is expressed through multiplication with the pupil function as an inverse Fourier transform, or,

$$E(u') = P(u') \int_{x=-\infty}^{\infty} dx E_0(x) e^{-i2\pi ux} \quad 2.44$$

We also know that the imaging system transforms the incident radiation into a wavefront converging on the image plane. Therefore, the electric field in the image plane can be expressed as,^{48,49}

$$E(x^{(i)}) = \int du' P(u') E(u') e^{i2\pi u' x^{(i)}} \quad 2.45$$

Using the shift theorem of the Fourier transform,⁵⁰

$$E(x^{(i)}) = \int du' E(u') \int dx P(u' - x^{(i)}) e^{i2\pi u' x} \quad 2.46$$

After carrying out the Fourier transform with respect to u' and rearranging,

$$E(x^{(i)}) = \int dx m(x) H(u' - x^{(i)}) \quad 2.47$$

where $m(x)$ is the mask function, and $H(x)$ is the point spread function and is the Fourier transform of the pupil function. Equation 2.47 is easily recognized as a convolution integral. The image plane intensity can be easily obtained through the squared modulus of this integral.^{6,26,41}

We therefore see that the pupil function is the transfer function of the imaging system in the coherent limit. For an ideal grating object, the diffracted spectrum samples the pupil function. In general, the pupil function is complex valued and can be given as,

$$P(u, v) = \alpha(u, v) e^{i2\pi W(u, v)} \quad 2.48$$

where $\alpha(u, v)$ represents the pupil amplitude function, and $W(u, v)$ represents the pupil phase function.²⁸ In an unaberrated system, the amplitude function is uniformly transmitting and the phase is zero at best focus. It is customary to expand the phase function in a Fourier-Zernike series as,

$$W(u, v) = \sum_{n=0}^{\infty} a_n Z_n(\rho, \phi) \quad 2.49$$

where (ρ, ϕ) represents polar coordinates defined in terms of (u, v) , and Z_n represents the n -th Zernike polynomial.^{6,28}

The Zernike polynomials form a complete set of orthogonal polynomials over the unit circle and fulfill,⁴⁵

$$r^2(1 - r^2) \frac{\partial^2 U}{\partial r^2} + [r - 3r^3] \frac{\partial U}{\partial r} + \frac{\partial^2 U}{\partial \phi^2} + \gamma r^2 U = 0 \quad 2.50$$

This differential equation can be solved via separation of variables by assuming a solution of the form,

$$Z_n^m(r, \phi) = R_n^m(r) \Phi_m(\phi), \quad 2.51$$

where n and m are integers with $-n < m < n$ and even $n - m$. The azimuthal component can be shown to be given by,

$$\Phi_m(\phi) = \begin{cases} \sin m\phi \\ \cos m\phi \end{cases}, \quad 2.52$$

while the radial component can be shown to be given by,

$$R_n^m(r) = \sum_{k=0}^{(n-m)/2} (-1)^k \frac{(n-m)!}{k! \left(\frac{n+m}{2} - k\right)! \left(\frac{n-m}{2} - k\right)!} r^{n-2k} \quad 2.53$$

The radial solution can be obtained by the appropriate change of variables to transform Equation 2.50 into a hypergeometric equation. The complete Zernike polynomial is then given by the product of Equations 2.52 and 2.53, as in Equation 2.51.⁴⁵

The Zernike radial polynomials are commonly normalized so that $R_n^m(r = 1) = 1$ and $R_n^m(r = 0) = 0$ for $m \neq 0$. In lithography it is customary to order the polynomials by a single index, rather than by n and m .⁶ The primary Zernike polynomials are given in Table 1 in both polar and Cartesian forms. It is extremely uncommon to work with the Zernike polynomials in Cartesian coordinates. This form is provided in Table 1 for comparison to Equations 2.6 – 2.9. The Zernike polynomials exactly reproduce the form of the geometric aberrations derived in Section 2.1. Contour plots of the first few terms are given in Figure 23.

Table 1. The primary Zernike polynomials in both polar and Cartesian coordinates

N	Name	Polar Equation	Cartesian Equation
1	Piston	1	1
2	Tilt X	$r \cos \theta$	u
3	Tilt Y	$r \sin \theta$	v
4	Defocus	$2r^2 - 1$	$2(u^2 + v^2) - 1$
5	Astigmatism 90°	$r^2 \cos 2\theta$	$u^2 - v^2$
6	Astigmatism 45°	$r^2 \sin 2\theta$	$2uv$
7	Coma X	$(3r^3 - 2r) \cos \theta$	$3u(u^2 + v^2) - 2u$
8	Coma Y	$(3r^3 - 2r) \sin \theta$	$3v(u^2 + v^2) - 2v$
9	Spherical	$6r^4 - 6r^2$	$6(u^2 + v^2)^2 - 6(u^2 + v^2)$
10	Trefoil X	$r^3 \cos 3\theta$	$u^3 - 2uv^2$
11	Trefoil Y	$r^3 \sin 3\theta$	$2u^2v - v^3$

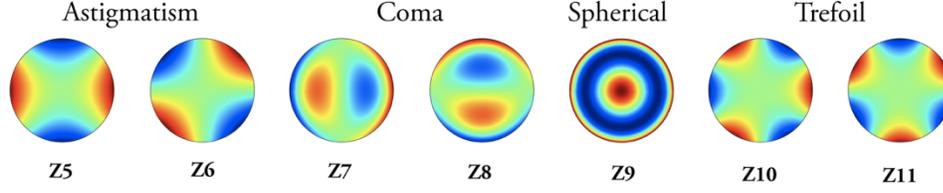


Figure 23. Contour plots of the 3rd order Zernike polynomials. Indices are given for the single index classification, along with their classical names.

2.1.3 Partially Coherent Image Formation

Informally, the coherence of light refers to its ability to create an interference pattern. An ideal monochromatic wave is sinusoidal and can interfere with any other ideal monochromatic wave. However, no wave is strictly monochromatic and interference requires a statistical similarity. In general, there are two types of coherence: 1) temporal and 2) spatial.²⁶ These refer to the ability of a light field to interfere with a temporal or spatially shifted version of itself. Light which is perfectly correlated is said to be coherent, while light which is perfectly uncorrelated is said to be incoherent. Anything in between is considered partially coherent.²⁶

To develop an imaging theory for a partially coherent source we will consider a single point of a spatially extended source at two different points in the receiving plane, as in Figure 24. We begin with the mutual coherence function, as defined by Wolf,^{26,51}

$$\Gamma_{12}(\tau) = \Gamma(P_1, P_2, \tau) = \langle E_1(t + \tau) E_2^*(t) \rangle, \quad 2.54$$

where E_n is an electric field at P_n , $\tau = t_2 - t_1$ is the time difference between the two field points, and the angle brackets denote a long time average, i.e.,⁵¹

$$\langle f \rangle = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T f dt \quad . \quad 2.55$$

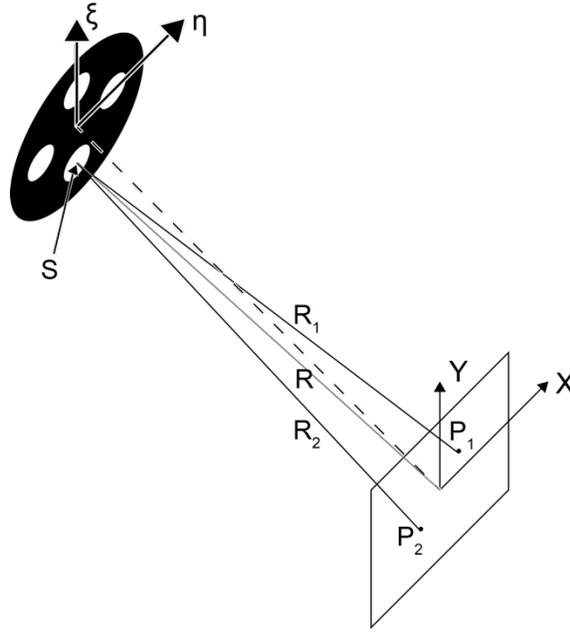


Figure 24. Geometry of a partially coherent source and its receiving plane

We must now determine how the mutual coherence of the spatially extended source of a lithography system propagates to the image plane. The intensity in the image plane can be determined by letting $P_2 = P_1$ and $\tau = 0$ so that,²⁶

$$\Gamma_{11} = \langle E_1(t) E_1^*(t) \rangle = I(P_1) \quad . \quad 2.56$$

The spectral width of lithography sources is quite small so we will make a quasi-monochromatic approximation and assume that our source is temporally coherent.^{6,51,52}

Under this assumption the mutual coherence function becomes,

$$\Gamma_{12}(0) = \langle E_1(t) E_2^*(t) \rangle \quad , \quad 2.57$$

which is the cross-correlation of the two electric fields. If the receiving plane is sufficiently far away from the source, the electric field can be given by,²⁶

$$E_n(t) = \mathcal{E}_n(t - R_n/c) \frac{e^{-i\omega(t-R_n/c)}}{R_n}, \quad 2.58$$

where R_n is the distance between the source point and P_n , \mathcal{E} is the complex amplitude of the field. The electric fields in this form, and therefore their mutual coherence, are solutions to the wave equation. The mutual coherence function is then,

$$\Gamma_{12} = \int_S dS \langle \mathcal{E}_1(t - R_1/c) \mathcal{E}_2^*(t - R_2/c) \rangle \frac{e^{i\omega R_1/c} e^{-i\omega R_2/c}}{R_1 R_2}. \quad 2.59$$

The time origin in a temporally coherent field is arbitrary, so we can offset the time axis by R_1/c to obtain,

$$\Gamma_{12} = \int_S dS \left\langle \mathcal{E}_1(t) \mathcal{E}_2^* \left(t + \frac{R_1 - R_2}{c} \right) \right\rangle \frac{e^{i\omega R_1/c} e^{-i\omega R_2/c}}{R_1 R_2}. \quad 2.60$$

If the source is far away then $R_1 \approx R_2$ so,

$$\Gamma_{12} = \int_S dSI(P_S) \frac{e^{i\bar{k}(R_1 - R_2)}}{R_1 R_2}, \quad 2.61$$

where $I(P_S)$ is the intensity distribution of the source and \bar{k} is the mean wavenumber.

The distance between the source point and the field points is given by,

$$R_n = \sqrt{R^2 + (x_n - \xi)^2 + (y_n - \eta)^2}, \quad 2.62$$

where R is the distance between the source point and the mid-point of the two field points. So, the path difference between the two field points is,

$$R_2 - R_1 = R \left[\sqrt{1 + \frac{(x_2 - \xi)^2}{R^2} + \frac{(y_2 - \eta)^2}{R^2}} - \sqrt{1 + \frac{(x_1 - \xi)^2}{R^2} + \frac{(y_1 - \eta)^2}{R^2}} \right] \quad 2.63$$

Using a Taylor series and some algebra,

$$R_2 - R_1 \approx \frac{x_2^2 - x_1^2 + y_2^2 - y_1^2}{2R} - \frac{(x_1 - x_2)\xi + (y_1 - y_2)\eta}{R} \quad 2.64$$

We can express the spatial frequency in terms of this geometry as,²⁶

$$u = \frac{\bar{k}}{2\pi R}(x_2 - x_1) \quad 2.65$$

$$v = \frac{\bar{k}}{2\pi R}(y_2 - y_1)$$

Combining, the mutual intensity can be written as,

$$\Gamma_{12} = \frac{e^{i\bar{k}(x_2^2 - x_1^2 + y_2^2 - y_1^2)/2R}}{R_1 R_2} \int_S dSI(P_S) e^{-i2\pi(u\xi + v\eta)} \quad 2.66$$

This again is a Fourier integral and is called the van Cittert-Zernike theorem.²⁶ This theorem implies that a quasi-monochromatic incoherent source creates mutual coherence in a far field plane given by the Fourier transform of its intensity function. The phase term at the front can be safely ignored if the distance between P_1 and P_2 is much less than the mean wavelength.²⁶

We found in Equation 2.47 that the electric field in the image plane due to a coherent source is given by the convolution of the transmitted electric field with the point spread function,

$$E(x^{(i)}) = \int dx m(x) \mathcal{E}(x) H(x - x^{(i)}) \quad 2.67$$

Because the mutual coherence follows its own wave equation, we can find the mutual coherence in the image plane via,

$$\Gamma_{12}^{(i)} = \int \int dx_1 dx_2 m(x_1) m^*(x_2) \langle \mathcal{E}(x_1) \mathcal{E}^*(x_2) \rangle H(x_1 - x_1^{(i)}) H^*(x_2 - x_2^{(i)}) \quad 2.68$$

Taking the Fourier transform of the mutual coherence function,

$$\Gamma_{12}^{(i)} = \mathcal{F}^{-1} \left\{ \int \int du_1 du_2 M(u_1) M^*(u_2) S(u) P(u_1 - u_1^{(i)}) P^*(u_2 - u_2^{(i)}) \right\} \quad 2.69$$

The intensity in the image plane can be obtained by letting $P_1 = P_2$ which implies that $u_1^{(i)} = u_2^{(i)} = u$. The intensity is then given by,

$$I(x) = \int \int \int du du_1 du_2 M(u_1) M^*(u_2) S(u) P(u_1 - u) P^*(u_2 - u) e^{i2\pi(u_1 - u_2)x} \quad 2.70$$

The source intensity and pupil function can be grouped together so that the image plane intensity can be written as,

$$I(x) = \int \int du_1 du_2 TCCM(u_1) M^*(u_2) e^{i2\pi(u_1 - u_2)x} \quad 2.71$$

where TCC is the transmission cross coefficient and is given by,

$$TCC = \int du S(u) P(u_1 - u) P^*(u_2 - u) \quad 2.72$$

This can be extended to two dimensions as,

$$I(x, y) = \int \int TCC \cdot M(u', v') M^*(u'', v'') e^{-i2\pi[(u' - u'')x + (v' - v'')y]} du' dv' du'' dv'' \quad 2.73$$

$$TCC(u', v', u'', v'') = \int \int S(u, v) P(u + u', v + v') P^*(u + u'', v + v'') dudv \quad 2.74$$

Equation 2.74 is often referred to as the Hopkins imaging equation and provides a general solution to computing partially coherent imagery in the presence of aberrations.^{26,53} The TCC allows the action of a partially coherent imaging system to be separated from the object. These solutions though are considerably more difficult to interpret physically than in the coherent limit. The TCC can be interpreted as an overlap integral between the source, the pupil function, and a reflection of the pupil function. The pupil function is averaged over the partially coherent source, so an ideal grating object would sample the source-average pupil function.

2.2 Linear Algebra Techniques

2.2.1 Eigenvector Decomposition

Eigenvector decomposition is a special form of matrix factorization which applies only to square diagonalizable matrices. For a matrix \mathbf{A} which has these properties, the eigenvector decomposition is stated as,^{54,55}

$$\mathbf{A}\mathbf{X} = \mathbf{X}\mathbf{\Lambda} , \tag{2.75}$$

where \mathbf{X} is a matrix of *eigenvectors* and $\mathbf{\Lambda}$ is a diagonal matrix of *eigenvalues*. The eigenvectors fulfill the property,

$$\mathbf{A}\hat{x}_i = \lambda_i\hat{x}_i , \tag{2.76}$$

where \hat{x}_i is the i -th column of \mathbf{X} and λ_i is the i -th element of the diagonal of $\mathbf{\Lambda}$. The eigenvectors are orthogonal and span the entire space of \mathbf{A} . Therefore, they form a basis in which \mathbf{A} can be perfectly represented as,

$$A = \sum_{i=1}^k \lambda_i \hat{x}_i \hat{x}_i^\dagger , \tag{2.77}$$

where the dagger symbol represents the Hermitian conjugate of a matrix.⁵⁵ To find the eigenvectors of a matrix one must solve the system of equations generated by,

$$\det (\mathbf{A} - \lambda\mathbf{I}) = 0 , \tag{2.78}$$

where \mathbf{I} is the appropriate identity matrix. This system is not linearly independent, so any matrix may have multiple forms of equivalent eigenvectors.⁵⁶ It is therefore possible to

find one set of eigenvectors which point in the opposite direction of another set. Both sets are factorizations valid however and should be considered equivalent.

2.2.2 Singular Value Decomposition

Eigenvector decomposition is a very useful technique and—despite its restriction to only square matrices—applies to many problems. Singular value decomposition (SVD) is a generalization of eigenvector decomposition to rectangular matrices. Any matrix multiplied by its transpose will generate a square symmetric matrix.⁵⁴ Symmetric matrices are diagonalizable and can therefore be analyzed with eigenvector decomposition. Let \mathbf{B} represent an arbitrary $n \times m$ matrix. The eigenvector decomposition of $\mathbf{B}^\dagger\mathbf{B}$ states,

$$(\mathbf{B}^\dagger\mathbf{B}) \mathbf{V} = \mathbf{V}\mathbf{\Lambda} . \quad 2.79$$

The eigenvectors of $\mathbf{B}^\dagger\mathbf{B}$ will not fulfill the eigenvalue equation for \mathbf{B} , but we can make the following definition,⁵⁷

$$\mathbf{B}\hat{v}_i = \sqrt{\lambda_i}\hat{u}_i . \quad 2.80$$

\mathbf{V} is an orthogonal matrix, so this definition implies that,

$$\mathbf{B} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^\dagger , \quad 2.81$$

where $\mathbf{\Sigma}$ is a diagonal matrix containing the square roots of the the diagonal elements of $\mathbf{\Lambda}$. This is the singular value decomposition of the matrix \mathbf{B} .

In the case that \mathbf{B} is a square matrix, \mathbf{U} and \mathbf{V} can be interpreted as rotation matrices because they are both orthogonal matrices. The matrix $\mathbf{\Sigma}$ is diagonal and

therefore represents a scaling. SVD can be interpreted geometrically as a rotation and stretching and rotation of the unit disc.⁵⁷

2.2.3 *Principal Component Analysis*

Principal component analysis (PCA) is a statistical technique to identify the directions of highest variation in a high dimensional dataset. To begin, the covariance matrix of a mean-centered dataset is computed and decomposed into its eigenvectors.^{57,58} The eigenvectors are then sorted by decreasing eigenvalue. The direction of highest variation is given by the eigenvector with the highest eigenvalue and is termed PC1. SVD can be used equivalently to compute PCA. Therefore, PCA can be interpreted geometrically as determining the basis in which the original data is rotated to reduce redundancy and highlight variation.

This process is depicted in Figure 25 where PCA was used to analyze data representing a line in \mathbf{R}^3 with Gaussian noise. In Figure 25 this is the direction which appears to be a least-squares line of best fit to the data. In actuality, this direction minimizes the error orthogonal to the data—not orthogonal to the axis.

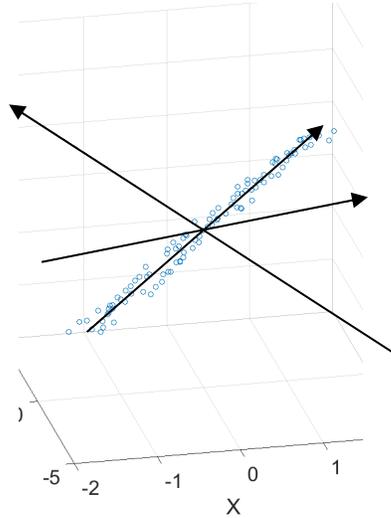


Figure 25. Data representing a line in \mathbb{R}^3 with Gaussian noise (blue circles) with the orthogonal directions of highest variation identified (black arrows) via PCA.

2.2.4 SOCS decompositions

Sum of coherent systems (SOCS) decomposition is a technique to compute partially coherent imagery fast via eigenvector decomposition of the TCC. We can see from Equation 2.74 that the TCC function can be described by a Hermitian matrix \mathbf{T} because,

$$TCC(f', g', f'', g'') = TCC^*(f'', g'', f', g'), \quad 2.82$$

This matrix can become quite large, but it can be decomposed into a product of a smaller matrix,

$$\mathbf{T} = \mathbf{P}^\dagger \mathbf{P}, \quad 2.83$$

where \mathbf{P} is called the pupil shift operator and \mathbf{P}^\dagger is its Hermitian adjoint.^{59,60} Further decomposition of the pupil shift operator via singular value decomposition gives the relation,

$$\mathbf{P} = \mathbf{U} \mathbf{L} \mathbf{V}^\dagger, \quad 2.84$$

where \mathbf{U} is a matrix whose rows represent the eigenvectors of $\mathbf{P} \mathbf{P}^\dagger$, \mathbf{V}^\dagger is a matrix whose columns represent the eigenvectors of $\mathbf{P}^\dagger \mathbf{P}$, and \mathbf{L} is a diagonal matrix with the singular values of \mathbf{P} .

Noting Equation 2.83, it is apparent that \mathbf{V} contains the eigenvectors of the TCC matrix, which leads to the eigenvalue decomposition of \mathbf{T} ,

$$\mathbf{T} = \mathbf{V}\mathbf{L}^2\mathbf{V}^\dagger . \quad \mathbf{2.85}$$

The TCC eigenfunctions, Φ_k , are obtained by reshaping the columns of \mathbf{V}^\dagger . The image plane intensity can in turn be decomposed using the eigenfunctions of the TCC function as,⁶¹

$$I(x, y) = \sum_{k=1}^N L_{kk}^2 |\Psi_k(x, y)|^2 , \quad \mathbf{2.86}$$

where Ψ_k is the k-th image intensity eigenfunction given as,

$$\Psi_k(x, y) = \mathcal{F}^{-1} \{ \Phi_k(x, y) M(x, y) \} . \quad \mathbf{2.87}$$

3. MEASURING PUPIL VARIATION

3.1 A naïve example

Consider a 1D slit pupil with a pupil function given by,

$$P(u) = \alpha(u)e^{iW(u)}. \quad 3.1$$

Diffraction of a line-space array with pitch P from a coherent source has the spectrum,

$$\tilde{M}(u) = \frac{1}{2}\delta(u) + \frac{1}{\pi}\delta\left(u - \frac{1}{P}\right) + \frac{1}{\pi}\delta\left(u + \frac{1}{P}\right), \quad 3.2$$

assuming only the 0th and $\pm 1^{\text{st}}$ diffraction orders are collected by the lens. This spectrum samples the pupil function as,

$$\tilde{E}(u) = \tilde{M}(u)P(u) = \frac{P(0)}{2}\delta(u) + \frac{P(\rho_1)}{\pi}\delta(u - \rho_1) + \frac{P(-\rho_1)}{\pi}\delta(u + \rho_1), \quad 3.3$$

where ρ_1 is defined to be the spatial frequency, or $1/P$. The electric field in the image-domain is given by the Inverse Fourier Transform, or,

$$E(x) = \frac{P(0)}{2} + \frac{P(\rho_1)}{\pi}e^{i2\pi\rho_1x} + \frac{P(-\rho_1)}{\pi}e^{-i2\pi\rho_1x}. \quad 3.4$$

The image is therefore given by,

$$I(x) = E(x)E^*(x). \quad 3.5$$

After simplification,

$$\begin{aligned}
I(x) = & \frac{\alpha(0)^2}{4} + \frac{\alpha(\rho_1)^2}{\pi^2} + \frac{\alpha(-\rho_1)^2}{\pi^2} - \frac{2\alpha(\rho_1)\alpha(-\rho_1)}{\pi^2} \\
& + \frac{\alpha(0)\alpha(\rho_1)}{\pi} \cos[2\pi\rho_1x + W(\rho_1) - W(0)] \\
& + \frac{\alpha(0)\alpha(-\rho_1)}{\pi} \cos[2\pi\rho_1x + W(0) - W(-\rho_1)] \\
& + \frac{4\alpha(\rho_1)\alpha(-\rho_1)}{\pi^2} \cos^2 \left[2\pi\rho_1x + \frac{W(\rho_1) - W(-\rho_1)}{2} \right].
\end{aligned} \tag{3.6}$$

Suppose the pupil function is given by an ideal top-hat transmission with some combination of, for example, third-order and fifth-order coma aberration. This can be given as,

$$P(u) = e^{i[A(3u^3-2u)+B(10u^5-12r^3+3r)]}, \tag{3.7}$$

where A is the amount of third-order coma, and B is the amount of fifth-order coma.

Coma is an odd-parity aberration, therefore the pupil function has the property,

$$W(-\rho_1) = -W(\rho_1), \tag{3.8}$$

and in this specific case,

$$W(0) = 0. \tag{3.9}$$

The image can then be given by,

$$I(x) = \frac{1}{4} + \frac{2}{\pi} \cos[2\pi\rho_1x + W(\rho_1)] + \frac{4}{\pi^2} \cos^2[2\pi\rho_1x + W(\rho_1)]. \tag{3.10}$$

Both of the sinusoidal terms have the same phase shift given by $W(\rho_1)$. The spatial frequency ρ_1 is known and we can measure the aerial image $I(x)$. The phase shift can be fit to an experimental aerial image with, for example, non-linear least squares. Therefore, the value of the pupil function at ρ_1 can be determined experimentally from the aerial image. However, the amount of third-order coma can be written as,

$$A = -B \frac{10\rho_1^5 - 12\rho_1^3 + 3\rho_1}{3\rho_1^3 - 2\rho_1} + \frac{W(\rho_1)}{3\rho_1^3 - 2\rho_1}. \quad 3.11$$

Therefore, there are multiple combinations of A and B that can lead to the same aberrated aerial image. Measuring the image of one pitch is not sufficient to determine both low and high order aberrations. A single binary target that can measure both low and high order aberrations from the same image does not exist. The value of the pupil function at ρ_1 is given by,

$$W(\rho_1) = 10B\rho_1^5 + (3A - 12B)\rho_1^3 + (3B - 2A)\rho_1, \quad 3.12$$

Therefore, low- and high-order effects get lumped into one effective low-order term if the high-order term is ignored. If we also measure the image of another pitch with spatial frequency ρ_2 then we have,

$$A = -B \frac{10\rho_2^5 - 12\rho_2^3 + 3\rho_2}{3\rho_2^3 - 2\rho_2} + \frac{W(\rho_2)}{3\rho_2^3 - 2\rho_2}. \quad 3.13$$

Combining,

$$\frac{W(\rho_2)}{3\rho_2^3 - 2\rho_2} - \frac{W(\rho_1)}{3\rho_1^3 - 2\rho_1} = B \left[\frac{10\rho_2^5 - 12\rho_2^3 + 3\rho_2}{3\rho_2^3 - 2\rho_2} - \frac{10\rho_1^5 - 12\rho_1^3 + 3\rho_1}{3\rho_1^3 - 2\rho_1} \right]. \quad 3.14$$

Simplifying,

$$B = \frac{W(\rho_2)(3\rho_1^3 - 2\rho_1) - W(\rho_1)(3\rho_2^3 - 2\rho_2)}{(10\rho_2^5 - 12\rho_2^3 + 3\rho_2)(3\rho_1^2 - 2\rho_1) - (10\rho_1^5 - 12\rho_1^3 + 3\rho_1)(3\rho_2^2 - 2\rho_2)}. \quad 3.15$$

If instead the pupil function were even-parity it would have the property,

$$W(\rho_1) = W(-\rho_1). \quad 3.16$$

Equation **3.10** therefore becomes,

$$I(x) = \frac{1}{4} + \frac{2}{\pi} \cos[W(\rho_1) - W(0)] \cos[2\pi\rho_1 x] + \frac{4}{\pi^2} \cos^2[2\pi\rho_1 x]. \quad 3.17$$

The sinusoidal term involving the pupil function acts as a phase shift and can also be experimentally determined from an aerial image. Therefore, for either even- or odd-parity aberrations, if we have images from two different pitches we can determine both the low- and high-order aberrations from experimentally measured quantities.

This example serves two purposes: 1) to illustrate the most basic form of the inverse problem we are solving, and 2) to show that there is not enough information in a single image to measure high-order aberrations. In practice aberrations are not correlated with sinusoidal phase shifts to measure aberrations in lithography systems. This is because lithographers generally cannot directly measure the aerial image and because this method could be extremely sensitive to noise. All aberration measurements do correlate an image-domain quantity with aberrations though.

3.2 Interferometric techniques

Interferometric techniques use the interference of a reference beam and a test beam, which has been modified by the test optic, to infer the pupil phase variation. There are two general categories of interferometric aberration measurements: 1) common path and 2) separate path interferometry. The reference beam takes a separate path in the latter, while the beams are split after the test optics in the former. Separate path interferometry is highly sensitive to atmospheric variation and vibration and is therefore considered incompatible with EUVL systems.

A form of common path interferometry is already available onboard modern lithography scanners. Integrated Lens Interferometer at Scanner (ILIAS) uses a point source at the reticle plane to image the pupil plane.⁶² The point source is transformed by the test optic and is imaged in the wafer plane approximately as a point. The wafer plane contains a set of gratings, instead of a wafer, which causes a far-field interference pattern proportional to the gradient of the wavefront in the shearing direction. This setup is shown in Figure 26. The aberrated wavefront of the test optic can then be inferred from measuring the wavefront gradients in multiple directions.

Interferometric techniques, such as ILIAS, are the de facto standard for aberration measurement, including DUV lithography systems. These techniques can be very accurate, but require additional optics and do not accurately reflect the system while it is in use. Moreover, interferometric techniques cannot measure pupil amplitude variation. Therefore, an alternative measurement technique is required for EUVL systems.

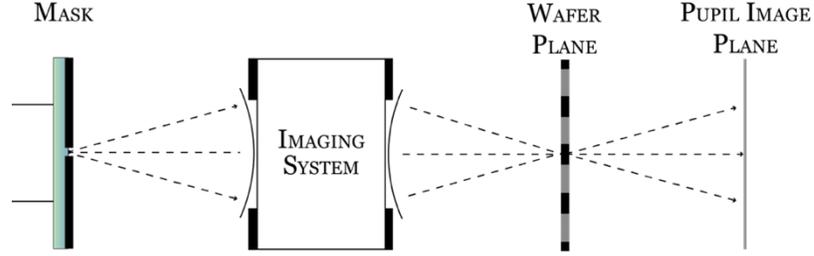


Figure 26. Setup for a lateral shearing interferometer in a lithography system, as in Integrated Lens Interferometer at Scanner (ILIAS)

3.3 Image-based techniques

In 1999 Dirksen et al. introduced the aberration ring test (ART) to measure the lens aberrations of a scanner from images formed by the scanner itself.⁶³ A cylindrical hole is etched into a reticle to form a $\lambda/2$ phase feature, as in Figure 27. An unaberrated system should image the hole as a circular ring. Any deviation from a circle can therefore be attributed to aberrations. The edges of the ring are expanded in a Fourier series as,

$$r(z, \Psi) = a_0(z) + \sum_m a_m(z) \cos m\Psi + \sum_m b_m(z) \sin m\Psi \quad 3.18$$

where (r, Ψ) are polar coordinates in the wafer plane and z is the defocus. Each Fourier component is then correlated to an aberration type. For example for coma X,

$$C_x = a_1 = \beta_7 Z_7, \quad 3.19$$

where β_7 is a model parameter and Z_7 is the amount of coma X. Higher-order terms are lumped into an equivalent low-order term and can be determined by exposing an aberration ring at multiple illumination settings (ARTEMIS).⁶⁴ This requires solving a system of equations with the Fourier components at each illumination setting, i.e.,

$$\begin{aligned}
C_x^{(1)} &= a_1^{(1)} = w_7^{(1)} Z_7 + w_{14}^{(1)} Z_{14} + w_{23}^{(1)} Z_{23} + \dots \\
C_x^{(2)} &= a_1^{(2)} = w_7^{(2)} Z_7 + w_{14}^{(2)} Z_{14} + w_{23}^{(2)} Z_{23} + \dots \\
C_x^{(3)} &= a_1^{(3)} = w_7^{(3)} Z_7 + w_{14}^{(3)} Z_{14} + w_{23}^{(3)} Z_{23} + \dots
\end{aligned}
\tag{3.20}$$

ARTEMIS can measure up to Z25 with a 3σ reproducibility between 8 and 23 m λ (1.5–4.5 nm on a DUV system and 0.11–0.31 nm in EUV) depending on the aberration type.⁶⁴

An alternative to ARTEMIS, which requires a special test structure on the reticle, is to have test structures directly built into the reticle and wafer stages. The transmission

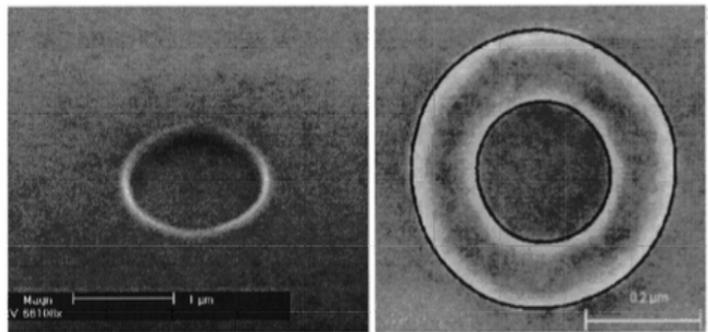


Figure 27. An SEM micrograph of the cylindrical phase hole on a reticle with a depth of $\lambda/2$ (left) and an image of the phase hole formed in resist (right)⁶³

image sensor (TIS) method uses a grating in the reticle stage and correlates measurements from a photodetector in the wafer stage with the effects of aberrations. Even parity aberrations are correlated with best focus shifts, while odd parity aberrations are correlated with image shifts. The TIS method can measure astigmatism, coma, and spherical aberration, but cannot measure three-foil because three-foil does not affect the imaging of horizontal and vertical gratings. Higher-order terms can be measured by using multiple illumination settings, in which case the test is referred to as TIS at multiple illumination settings (TAMIS) and has 3σ reproducibility of 10-15 m λ (2–3 nm on a DUV system or 0.14–0.20 nm in EUV).

Phase features have proven to be difficult to create on EUV reticles, so ART/ARTEMIS are not an option for EUVL. An example of 50 nm and 100 nm phase features on an EUVL reticle is shown in Figure 28. There is extreme etch undercutting, which destroys the 50 nm features, and significantly blurs the 100 nm features. Similarly, TIS/TAMIS have been difficult to implement in EUVL because image sensors can be unreliable. This, combined with its inability to measure three-foil, makes TIS/TAMIS not a viable option for EUVL.

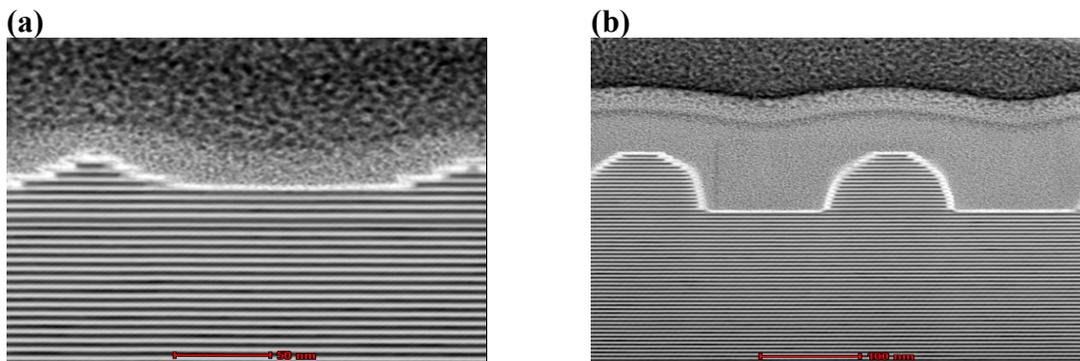


Figure 28. Transmission electron micrographs of (a) 50 nm and (b) 100 nm phase features post etch.

The ideal aberration measurement technique for EUVL uses binary targets and images formed in resist. To this end, we have identified a number of pattern types, already present on most masks, which are sensitive to specific types of Zernike polynomials, as shown in Figure 29a. If a certain aberration is present, then there will be a phase or amplitude difference between the measurement sites of these targets.

The pupil variation of an optical system can then be extracted via the images of these targets as in the algorithm described by Figure 29b. To begin, we assume that the system is aberration free. Several aerial image simulations are computed and a model is

fit to these simulations via non-linear least squares. This model can then be inverted to fit to experimental data (either aerial image or CD data) and obtain an initial guess for that aberration. This process is repeated for each aberration to obtain an initial guess for the aberrated wavefront. Finally, the process repeats until the algorithm converges on a solution for the pupil variation.

As an example, if astigmatism 90° is present in the system the diffraction information for a vertical grating and that of a horizontal grating of the same pitch will experience opposite phase shifts, causing an orientation dependent shift in CD. Moreover, this difference in CD will be focus dependent, as seen in Figure 30. These targets have been designed to interrogate third order aberrations only. It is difficult to directly interrogate high order terms due to the partially coherent source averaging the pupil.

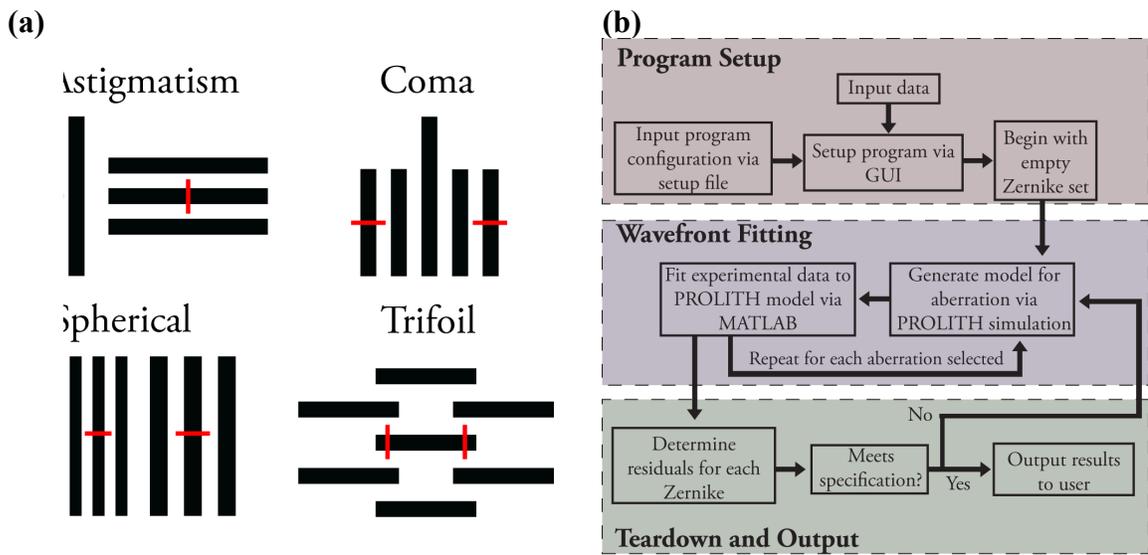


Figure 29. (a) Examples of metrology targets for each primary aberration. The red lines denote measurement locations. The aberration is interrogated by the CD difference at measurement sites. (b) A flowchart for the iterative algorithm of image-based pupil characterization.

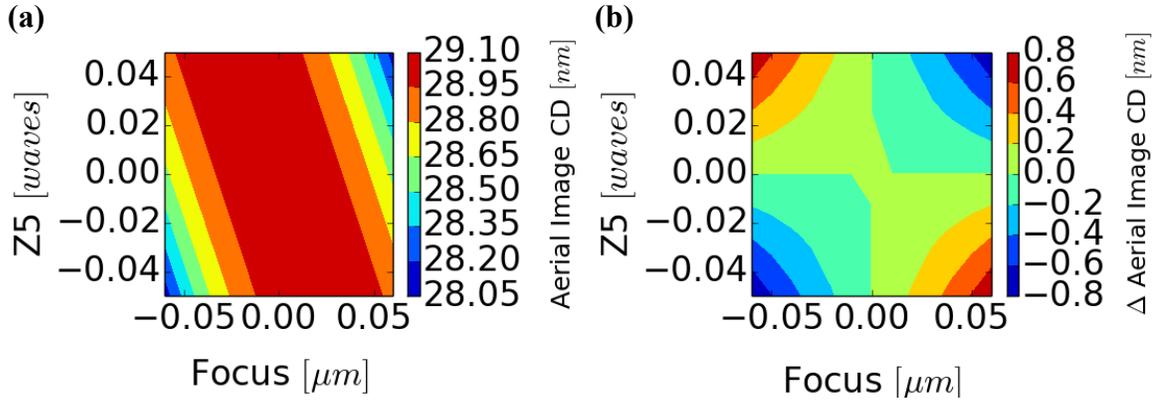


Figure 30. Variation in aerial image CD for vertical lines and spaces through focus and different levels of Z5 (a), and the variation in the difference in vertical and horizontal aerial image CD through focus and Z5 (b).

3.4 Optimization of image-based metrology targets

The dimensions of the targets need to be optimized for each system and illumination setting. A target will be most sensitive when its diffracted spectrum is centered at the extrema of the Zernike polynomial it is meant to interrogate. Using this technique to optimize targets to interrogate third and fifth order astigmatism, there is only a minor difference in the aberration phase RMS between the two aberrations, as seen in Figure 31. There is, however, a large difference in aerial image contrast and NILS. Therefore, we observe a trade-off between target printability and aberration sensitivity.

Smaller structures are more sensitive to aberrations, but the most sensitive structures may not be printable. Instead, we balance printability and aberration sensitivity to determine the optimum target by considering the worst-case scenario. After choosing a source shape, the required depth of focus (DOF) is determined. These are used as inputs in aerial simulations that are constrained by determining the minimum NILS to be considered printable and the desired aberration tolerances.

For example, we will consider target optimization for an NXE:3100 system. We chose an annular source with $0.5\text{--}0.8\ \sigma$ at 0.25NA and a maximum of $\pm 75\text{nm}$ defocus. A NILS threshold of 2 was used to define printability, while aberration tolerances were determined via the mean aberration levels extracted using the image-based technique in the Alpha Demo Tool (ADT), a prototype EUVL system. It was assumed that the NXE:3100 system would have less severe aberrations than the ADT.

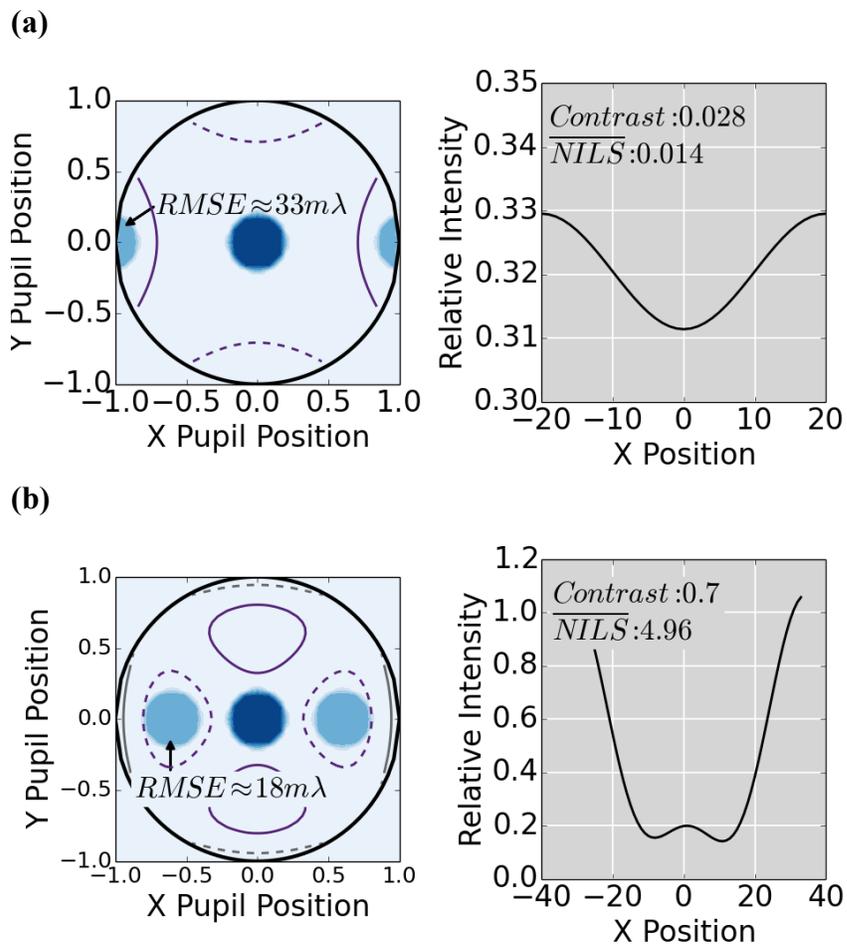


Figure 31. Diffraction information and resulting aerial images for 20nm (a) vs. 33nm (b) line/space array overlaid with 3rd and 5th order 90° astigmatism respectively; 0.33 NA, 0.2σ .

Using this approach NILS contours were constructed for each metrology target, as shown in Figure 32. The target to interrogate spherical aberration is not present in this figure because spherical aberration is measured by shifts in best focus, rather than by CD. Instead, this target was optimized by centering the $\pm 1^{\text{st}}$, $\pm 2^{\text{nd}}$, and $\pm 3^{\text{rd}}$ diffraction orders at the pupil edge. Next, a point was chosen in between the pitch determined for the $\pm 1^{\text{st}}$ and $\pm 2^{\text{nd}}$ diffraction orders, and likewise for the $\pm 2^{\text{nd}}$ and $\pm 3^{\text{rd}}$ diffraction orders. This gives a total of five pitches to interrogate spherical aberration. For this source we determined pitches of 42 nm, 64 nm, 85 nm, 105 nm, and 125 nm to be optimal.

The remaining aberrations were interrogated via a CD difference and can be optimized using this balance of printability and aberration sensitivity. In Figure 32 the aberration levels extracted from the ADT are given by the dashed lines. One can simply find the intersection of the dashed lines and the NILS contour to find the optimal target. For the five-bar structure used to interrogate coma, things are a bit more complicated. Here, we must maintain printability for both edges of the exterior bars. Therefore, the most sensitive edges are shown with the average NILS for both edges. With this consideration in mind the optimization procedure for the coma target is identical to that of the astigmatism and trefoil targets. Using this methodology it was determined that the optimal pitches for the chosen annular source were 55 nm, 51 nm, and 51 nm, for astigmatism, coma, and trefoil, respectively.

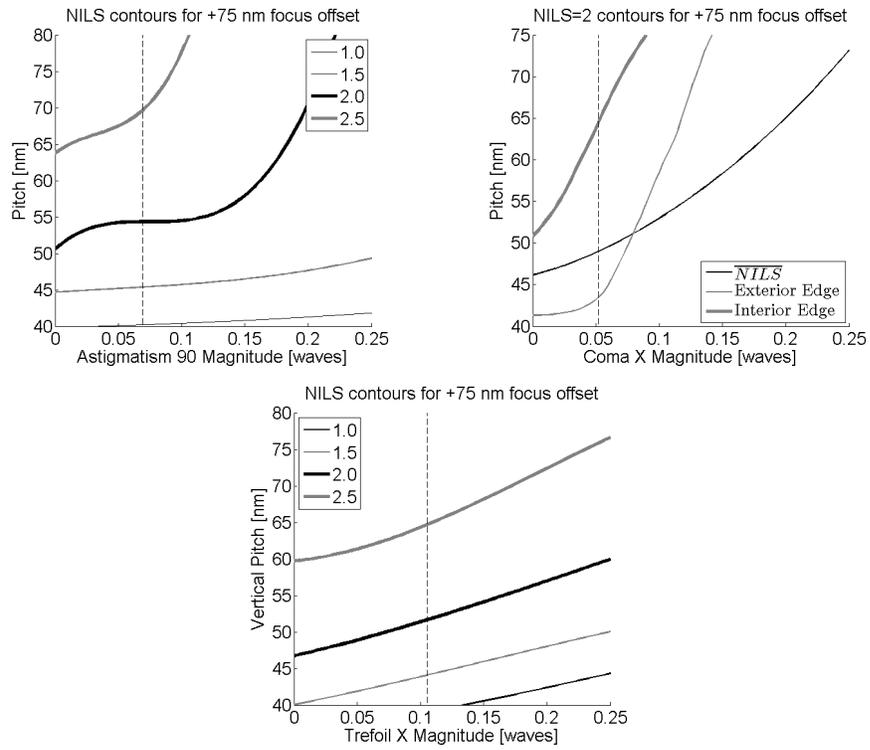


Figure 32. Nils contours for astigmatism 90° (top left), coma X (top right), and trefoil X (bottom).

4. EXTRACTION OF AMPLITUDE VARIATION

4.1 Choosing a basis to represent amplitude pupil variation

We have seen that Zernike polynomials are used to express pupil phase variation because they reproduce the terms derived from a geometrical treatment of aberrations. This makes the Zernike polynomials the most efficient basis for expressing pupil phase variation, though this is not guaranteed to be the case for pupil amplitude variation. The Zernike polynomials also have the property that they are orthonormal over the unit circle and rotationally invariant. As mentioned previously, these are very useful qualities for a function describing an isomorphic optic with a circular pupil.

We therefore seek some set of functions, other than the Zernike polynomials, which are orthonormal across the unit circle and which are rotationally invariant. The function we are looking for, $V(x,y)$, must be a superposition of rotational invariants, and therefore satisfies,

$$\nabla^2 V + \alpha \left(x \frac{\partial}{\partial x} + y \frac{\partial}{\partial y} \right)^2 V + \beta \left(x \frac{\partial}{\partial x} + y \frac{\partial}{\partial y} \right) V + \gamma V = 0. \quad 4.1$$

In polar coordinates the differential operators can be expressed as,

$$\rho \frac{\partial}{\partial \rho} = x \frac{\partial}{\partial x} + y \frac{\partial}{\partial y}, \quad 4.2$$

and

$$\frac{\partial}{\partial \phi} = y \frac{\partial}{\partial x} - x \frac{\partial}{\partial y}. \quad 4.3$$

where ρ and ϕ are polar coordinates defined for x and y in the usual way. The differential equation then becomes,

$$(1 + \alpha\rho^2) \frac{\partial^2}{\partial \rho^2} V + \left[\frac{1}{\rho} + (\alpha + \beta)\rho \right] \frac{\partial}{\partial \rho} V + \frac{1}{\rho^2} \frac{\partial^2}{\partial \phi^2} V + \gamma V = 0. \quad 4.4$$

Assuming that the function V is separable,

$$V = R(\rho)G(\phi). \quad 4.5$$

Combining these two equations it is easy to see that,

$$G'' = -m^2 G, \quad 4.6$$

and therefore,

$$G = \begin{cases} \cos m\phi \\ \sin m\phi \end{cases}, \quad 4.7$$

The differential equation for the radial component becomes,

$$\rho(1 + \alpha\rho^2)R'' + [1 + (\alpha + \beta)\rho^2]R' + \left[\gamma\rho - \frac{m^2}{\rho} \right] R = 0. \quad 4.8$$

Now let $x = r^2$. We can show that,

$$\frac{dR}{dr} = 2\sqrt{x} \frac{dR}{dx} \quad 4.9$$

and

$$\frac{d^2 R}{dx^2} = 4x \frac{d^2 R}{dx^2} + 2 \frac{dR}{dx} \quad 4.10$$

Using these to change the variable of the differential equation,

$$x(1 + \alpha x) \frac{d^2 R}{dx^2} + \left[1 + \left(\alpha + \frac{1}{2} \beta \right) x \right] \frac{dR}{dx} + \frac{1}{4} \left[\gamma - \frac{m^2}{x} \right] R = 0. \quad 4.11$$

The coefficients corresponding to the first derivative and the function itself are not analytic for $x = 0$ and $x = -1/\alpha$, therefore these are singular points of the differential equation. This corresponds to $r = 0$ and $r^2 = -1/\alpha$. This means that the singularity at $x = -1/\alpha$ will occur in the real plane for negative α . We can further examine the singularities by multiplying by $(x - x_0)^2/x(1 + \alpha x)$ to obtain,

$$(x - x_0)^2 R'' + (x - x_0)^2 \frac{1 + (\alpha + \beta/2)x}{x(1 + \alpha x)} R' + \frac{(x - x_0)^2}{4x(1 + \alpha x)} \left[\gamma - \frac{m^2}{x} \right] R = 0. \quad 4.12$$

Using $x_0 = 0$, we obtain,

$$x^2 R'' + \frac{1 + \left(\alpha + \frac{\beta}{2} \right) x}{1 + \alpha x} x R' + \frac{1}{4(1 + \alpha x)} [\gamma x - m^2] R = 0. \quad 4.13$$

It is clear that in this case the terms associated with R' and R are analytic for $x = 0$.

Similarly using $x_0 = -1/\alpha$,

$$(x + 1/\alpha)^2 R'' + \frac{1 + \left(\alpha + \frac{\beta}{2} \right) x}{\alpha x} \left(x + \frac{1}{\alpha} \right) R' + \frac{\left(x + \frac{1}{\alpha} \right)}{4\alpha x} \left[\gamma - \frac{m^2}{x} \right] R = 0 \quad 4.14$$

which is also analytic for $x = -1/\alpha$. We can therefore conclude that the singularities at $x = 0$ and $x = -1/\alpha$ are both regular singular points.

We are interested in solutions around $x = 0$, but we also know that the singularity at $x = -1/\alpha$ is a weak singularity because it is a regular singular point. The differential

equation can be reduced to an associated Euler equation by using the Taylor series of the terms corresponding with R' and with R ,

$$x^2 R'' + xR' - \left(\frac{m}{2}\right)^2 R = 0 \quad 4.15$$

Now assuming a solution,

$$R(x) = x^{r_0} \quad 4.16$$

We can find,

$$r_0(r_0 - 1) + r_0 - \left(\frac{m}{2}\right)^2 = 0 \quad 4.17$$

Therefore we find that the exponent of the singularity is,

$$r_0 = \pm m/2 \quad 4.18$$

The solution around $r = 0$ is therefore of the form,

$$R(x) = x^{m/2} F(x) \quad 4.19$$

We can find,

$$\frac{dR}{dx} = x^{m/2} \left[\frac{m}{2x} F(x) + F'(x) \right] \quad 4.20$$

And

$$\frac{d^2 R}{dx^2} = x^{m/2} \left[F'' + \frac{m}{x} F' + \frac{m^2 - 2m}{4x^2} F \right] \quad 4.21$$

Therefore,

$$\begin{aligned} \alpha x(1 + \alpha x)F'' + \left[(m + 1)\alpha + \alpha(m + 1)\alpha x + \frac{1}{2}\beta\alpha x \right] F' \\ + \alpha \left[\frac{1}{4}\gamma + \frac{m}{4}(\alpha m + \beta) \right] F = 0 \end{aligned} \quad 4.22$$

Letting $z = -\alpha x$ we obtain,

$$z(1 - z)F'' + \left[\alpha(m + 1)(z - 1) + \frac{1}{2}\beta z \right] F' - \left[\frac{1}{4}\gamma + \frac{m}{4}(\alpha m + \beta) \right] \alpha F = 0 \quad 4.23$$

This is in the form of a hypergeometric equation,

$$z(1 - z)F'' + [c - (a + b + 1)z]F' - abF = 0 \quad 4.24$$

Clearly,

$$c = -(m + 1)\alpha \quad 4.25$$

We can also find,

$$a^2 + \left[\alpha(m + 1) + \frac{1}{2}\beta + 1 \right] a + \alpha \left[\frac{1}{4}\gamma + \frac{m}{4}(\alpha m + \beta) \right] = 0 \quad 4.26$$

And

$$b = \frac{\alpha}{a} \left[\frac{1}{4}\gamma + \frac{m}{4}(\alpha m + \beta) \right] \quad 4.27$$

Solving these simultaneous equations,

$$a = -\frac{1}{2} \left[\alpha(m + 1) + \frac{1}{2}\beta + 1 \right] + \frac{1}{2} \sqrt{1 - \alpha\gamma + D} \quad 4.28$$

And

$$b = -\frac{1}{2}\left[\alpha(m+1) + \frac{1}{2}\beta + 1\right] - \frac{1}{2}\sqrt{1 - \alpha\gamma + D} \quad 4.29$$

Where

$$D = \alpha^2(2m+1) + 2\alpha(m+1) + \alpha\beta + \beta\left(\frac{\beta}{4} + 1\right) \quad 4.30$$

Therefore,

$$R(x) = x^{m/2}F(a, b, -\alpha(m+1), -\alpha x) \quad 4.31$$

If we choose $\alpha = -1$, the coefficient of R'' is zero for both $x = 1$ and $x = 0$, which allows us to normalize the functions to the edge of the pupil. The solution to the differential equation is then,

$$R(r) = r^m F\left(\frac{1}{2}\left[m - \frac{\beta}{2}\right] + \frac{1}{2}\sqrt{\frac{\beta^2}{4} + \gamma}, \frac{1}{2}\left[m - \frac{\beta}{2}\right] - \frac{1}{2}\sqrt{\frac{\beta^2}{4} + \gamma}, m+1, r^2\right) \quad 4.32$$

This will be singular at $r^2 = 1$ if $c - a - b = 0$ or,

$$(m+1) = \left[(m+1) - \frac{1}{2}\beta - 1\right] \quad 4.33$$

Therefore if $\beta = -2$, the function will only be finite within the pupil if F is a terminating polynomial.

$$\text{Let } \frac{1}{2}\left[m - \frac{\beta}{2}\right] - \frac{1}{2}\sqrt{\frac{\beta^2}{4} + \gamma} = -\frac{n-m}{2}, \text{ therefore,}$$

$$\gamma = n(n - \beta) \quad 4.34$$

And

$$R(r) = r^m F\left(\frac{m+n-\beta}{2}, -\frac{n-m}{2}, m+1, r^2\right) \quad 4.35$$

For even $n-m$ this results in a terminating sum of polynomials. We will normalize the functions to the edge of the pupil so that,

$$R_n^m(r=1) = 1 \quad 4.36$$

Therefore,

$$R_n^m(r) = (-1)^{\frac{n-m}{2}} F\left(\frac{n+m-\beta}{2}, -\frac{n-m}{2}, m+1, r^2\right) r^m \quad 4.37$$

The hypergeometric function can be rewritten in terms of Jacobi polynomials using the identity,

$$F(p+n, -n, q, x) = G_n(p, q, x) \quad 4.38$$

Where G_n is a Jacobi polynomial. Thus,

$$R_n^m(r) = (-1)^{\frac{n-m}{2}} \binom{n+m/2}{m} r^m G_{\frac{n-m}{2}}\left(m - \frac{\beta}{2}, m+1, r^2\right) \quad 4.39$$

Jacobi polynomials are defined as,

$$G_n(p, q, x) = \frac{x^{1-q}(1-x)^{q-p}}{q^n} \left\{ \frac{d}{dx} \right\}^n \{x^{q+n-1}(1-x)^{p+n-q}\} \quad 4.40$$

Using this definition with the solution for the radial function, it can be shown that,

$$R_n^m(r) = \sum_{k=0}^{\frac{n-m}{2}-\gamma} \frac{(-1)^k \left(\frac{n-m}{2} - \gamma\right)! (n-k-\gamma)!}{k! \left(\frac{n-m}{2}\right)! \left(\frac{n-m}{2} - k - \gamma\right)! \left(\frac{n+m}{2} - k - \gamma\right)!} (r^2 - 1)^\gamma r^{n-2k-\gamma} \quad 4.41$$

Where

$$\gamma = \frac{\beta + 2}{2} \quad 4.42$$

The series requires both that $n > m$ and $n - m - \beta$ be even parity so that $\frac{n-m}{2} - \gamma$ is a whole number. Therefore, $n - m$ must have the same parity as β . If β is odd, then there will be some terms in $R_n^m(r)$ that go as,

$$\left(\sqrt{r^2 - 1}\right)^{\beta+2} \quad 4.43$$

This term is only real valued for $r^2 > 1$, which is outside of the pupil. Therefore, β should be even parity. We can also observe that the lowest order term in the basis will occur for $n = m + \beta + 2$.

Suppose, for example, $\beta = 2$. The lowest order (n, m) combination is (5,1), in which case the lowest order term is $r^3(r^2 - 1)^2$. This is inadequate to account for low-order pupil variation. Low order terms occur last in the series, so we can require the lowest order term to be linear in r by requiring,

$$n - 2\left(\frac{n - m}{2} - \gamma\right) - \gamma = 0 \quad 4.44$$

This implies,

$$\beta = 2(m - 1) - 2 \quad 4.45$$

If we require the lowest order term to coincide with $m = 1$, then $\beta = -2$ and the solution then becomes,

$$R_n^m(r) = \sum_{k=0}^{\frac{n-m}{2}} \frac{(-1)^k (n-k)!}{k! \left(\frac{n-m}{2} - k\right)! \left(\frac{n+m}{2} - k\right)!} r^{n-2k} \quad 4.46$$

This is the definition of the radial component of the Zernike polynomials. Therefore, the only set of functions with the properties that we desire over the unit circle are the Zernike polynomials. Therefore, if we want to express amplitude pupil variation in a basis that is: 1) orthogonal over the unit circle, 2) rotationally invariant, the only choice is to use Zernike polynomials.

This does not mean that the Zernike polynomials are the only choice of basis for amplitude pupil variation. It only means that the only basis over the unit circle which has the properties that we desire to express pupil variation is the Zernike polynomials. If we loosen our requirements to, for example, not require rotational invariance, other bases may efficiently express amplitude pupil variation.

4.2 Comparison of the properties of several orthonormal bases

4.2.1 Zernike Polynomials

Any function defined over the unit circle, $f(r, \phi)$, can be expanded in terms of Zernike polynomials by solving the integral given by Equation 4.52.⁶⁵ The Zernike polynomials are designed to be rotationally invariant, which means that the expansion coefficients of a version of f rotated by θ , given by $f(r, \phi - \theta)$, are only modified by a linear azimuthal phase term.⁴⁵ The Zernike moments of a rotated function are therefore given by Equation 4.48. This can be seen by rotating the example image of Figure 35a and then expanding it in Zernike polynomials. The choice of this image is arbitrary--any function defined over the unit circle is an equally valid choice. Some of the resulting Zernike moments are

shown in Figure 35b, where it is seen that the moments vary only by an azimuthal phase factor.⁶⁵

$$a_{|n|m} = \frac{n+1}{\pi} \int_{\phi=0}^{2\pi} \int_{r=0}^1 f(r, \phi) R_n^m(r) e^{-im\phi} r dr d\phi. \quad 4.47$$

$$a_{|n|m}^{(\theta)} = a_{|n|m} \cdot e^{-im\theta}. \quad 4.48$$

Rotational invariance is useful in optical systems because a lens' aberrations should be independent of its rotation. For example, the total amount of 3rd order

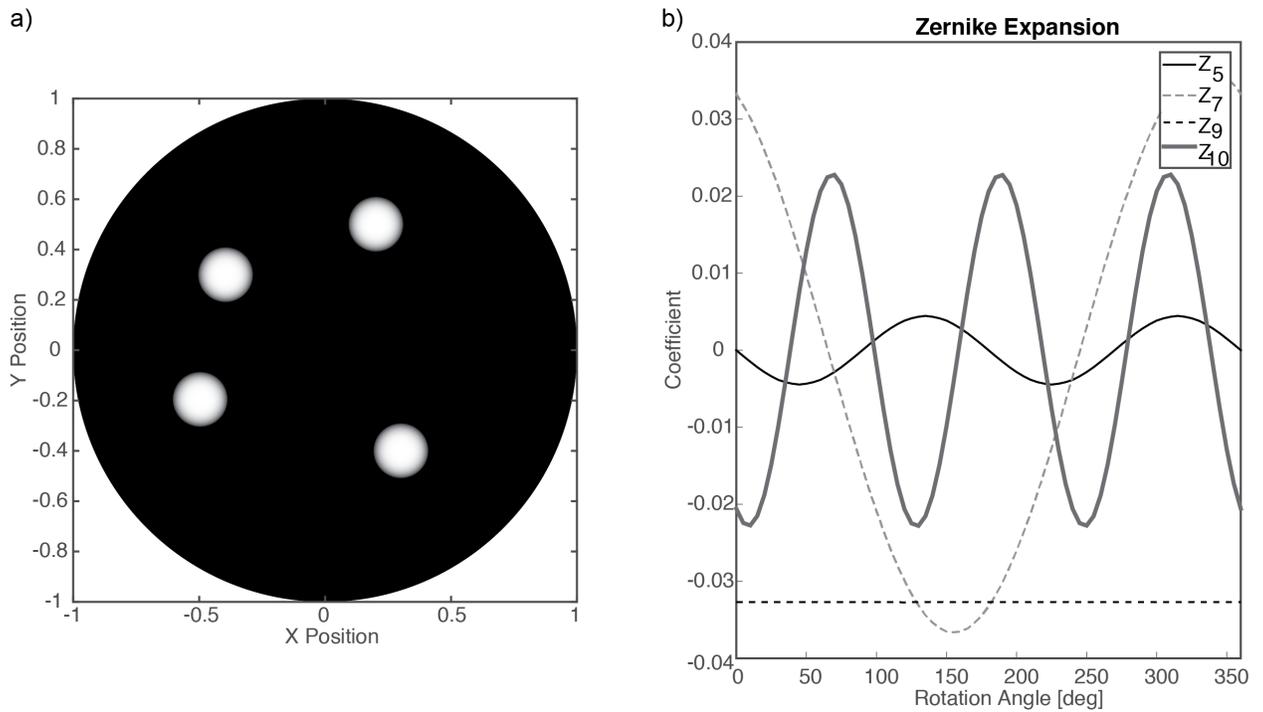


Figure 33. a) The image which was rotated then expanded in orthonormal bases. The units are normalized distance so that the function is defined over the unit circle. b) Zernike expansion coefficients for 3rd order astigmatism, coma, spherical aberration, and trefoil for the image in part (a) through rotation.

astigmatism in a lens will remain constant through the lens' rotation, though the amount of 3rd order astigmatism 90° (Z_5) and 3rd order astigmatism 45° (Z_6) contributing to this total 3rd order astigmatism may change. Put another way, the quantity $a_5^2 + a_6^2$ remains constant as a function of rotation angle, while a_5 and a_6 themselves change through rotation. This gives rise to the sinusoidal behavior of the moments in Figure 35b.

4.2.2 Legendre Polynomials

The Legendre polynomials form a complete orthogonal set over the interval $-1 \leq x \leq 1$. The polynomials can be generated via the Rodrigues' formula, Equation 4.49, where n is a positive integer.⁶⁶

$$P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} (x^2 - 1)^n . \quad 4.49$$

These polynomials are normalized so that the values at the edge of the interval are $P_n(x = 1) = 1$ and $P_n(x = -1) = (-1)^n$. Here we extend the Legendre polynomials over the 2D Cartesian plane by using all of the possible products of the Legendre polynomials in x and y , denoted by $P_N^{(2)}$. This combination of polynomials is orthogonal over a unit square and fulfills the orthogonality condition,

$$C_{MN} \int_{x=-1}^1 \int_{y=-1}^1 P_M^{(2)}(x, y) P_N^{(2)}(x, y) dx dy = \delta_{MN}, \quad 4.50$$

where δ_{MN} is the Kronecker delta and C_{MN} is an orthogonality constant.²⁸ These polynomials are not orthogonal over the unit circle, but that does not prevent them from

being used as an expansion basis. One must simply recompute all of the 2D-Legendre moments if another term is added to the series expansion.

Figure 34a shows the rotation of Figure 35a and its subsequent expansion in Legendre polynomials. This time the variation in expansion coefficients is not a simple sinusoid, and therefore the Legendre polynomials are not a rotationally invariant basis.

4.2.3 Hermite Polynomials

The Hermite polynomials form a complete orthogonal set over the interval $-\infty \leq x \leq \infty$.

The Rodrigues' formula for Hermite polynomials is given by Equation 4.51.⁶⁶ There is no analytic formula for the value $H_n(x = \pm 1)$.

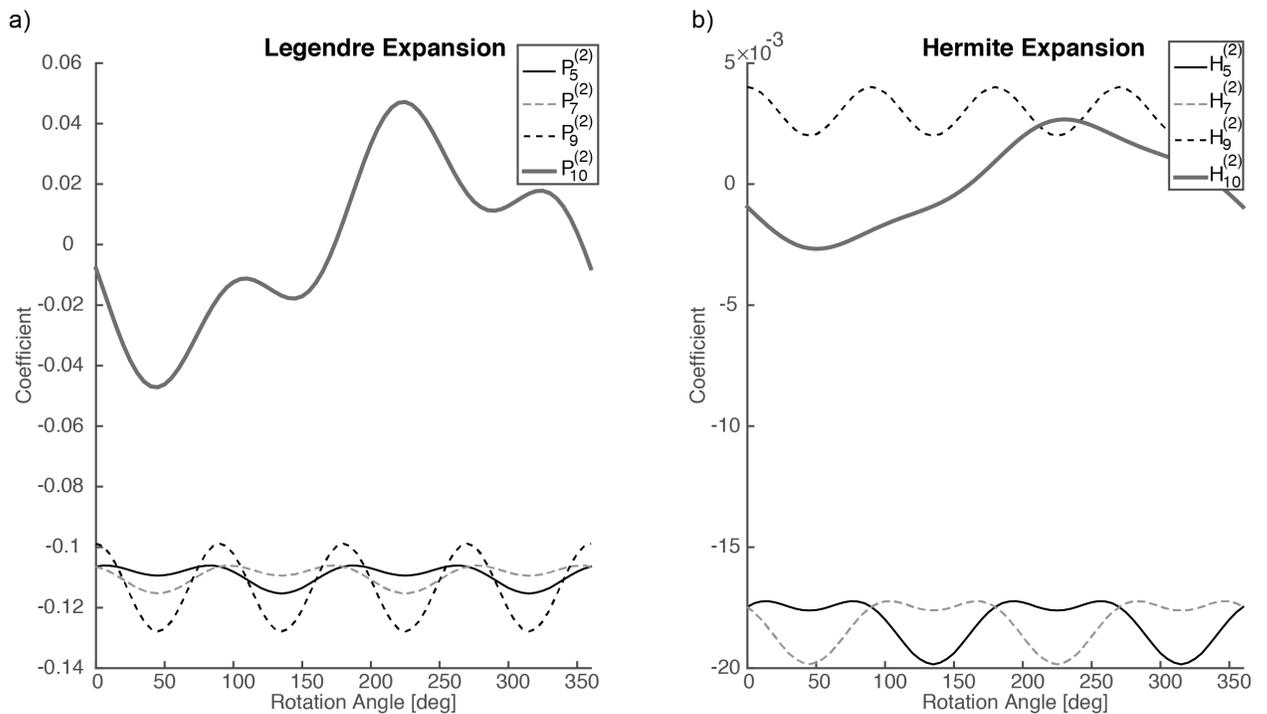


Figure 34. The image in Figure 35a was again rotated and expanded in Legendre polynomials and Hermite polynomials. The moments for a) Legendre polynomials for $P_5^{(2)}$, $P_7^{(2)}$, $P_9^{(2)}$, $P_{10}^{(2)}$ and b) Hermite polynomials for $H_5^{(2)}$, $H_7^{(2)}$, $H_9^{(2)}$, $H_{10}^{(2)}$ are plotted as a function of rotation angle.

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2}. \quad 4.51$$

Again we extend the basis over the 2D Cartesian plane by using all of the possible products of the Hermite polynomials in x and y . The notation $H_N^{(2)}$ is used to refer to the N -th product of two Hermite polynomials. The same orthogonality argument made for Legendre polynomials holds for Hermite polynomials. Similarly, like Legendre polynomials, the Hermite polynomials are not rotationally invariant. This can again be seen in Figure 34b where Figure 35a was again rotated and expanded this time in Hermite polynomials.

4.3 Extraction of amplitude variation

The multilayer reflectivity of mirrors in an EUVL system can vary dramatically as a function of incidence angle. This introduces a diffraction intensity imbalance in partially coherent imagery. This intensity imbalance can cause image variation through focus. Any optical system can be affected by this type of variation, but these effects become non-negligible with the tighter tolerances of EUVL systems. For this reason, we have adopted a more general definition of system aberration: any variation during imaging which introduces error to the pupil.

We redefine the amplitude function in terms of a separate Fourier-Zernike series,

$$\alpha(u, v) = 1 + \sum_{n=0}^{\infty} b_n Z_n(\rho, \phi) \quad 4.52$$

Extraction of amplitude variation follows a similar procedure to that for phase variation, but is considered separately with a yet unknown basis. To reconstruct the pupil amplitude

function $A(u, v)$, image data is iteratively fit to simulations to determine the values of the average value of the function over each diffraction order. This value corresponds to the source-averaged sample of the pupil amplitude function. This approach is identical to that used to extract pupil phase with one complication: the best choice of basis is unknown, so the source-averaged value of the amplitude function needs to be determined for each diffraction order instead of an expansion coefficient.

This pupil sampling process is illustrated in Figure 35. The amplitude function can be reconstructed through interpolation after iteratively fitting samples. Barnes objective analysis is used to interpolate across the pupil between the samples. This procedure—commonly used in meteorological modeling—uses an initial guess for each grid point, then iteratively refines it from the error computed from the known values. The weight of each error is proportional to the inverse of its distance from other points.^{67,68}

To begin the objective analysis we assume that the amplitude function $A(\rho, \theta)$ is Fourier decomposable. Then a corresponding filtered function $\tilde{A}(\rho, \theta)$ is constructed as,

$$\tilde{A}(\rho, \phi) = \int_{\theta=0}^{2\pi} \int_{r=0}^{\infty} A(\rho + r, \phi + \theta) W(r) r dr d\theta, \quad 4.53$$

where r and θ are polar coordinates defined with respect to a point in the pupil (ρ, ϕ) , and the filter $W(r)$ is given as,

$$W(r) = \frac{1}{4\pi k} e^{-r^2/4k}, \quad 4.54$$

where k is an arbitrary shape parameter. Equation 4.53 can be written in terms of the known discrete values of A as,

$$\tilde{A}_0(\rho, \phi) = \frac{\sum_{i=0}^N A_i e^{-r_i^2/4k}}{\sum_{i=0}^N e^{-r_i^2/4k}}, \quad 4.55$$

where N is the number of known data points. A corrected amplitude function can then be computed from the smoothed error field as,

$$\tilde{A}_1(\rho, \phi) = \tilde{A}_0 + \frac{\sum_{i=0}^N (A_i - \tilde{A}_{0,i}) e^{-r_i^2/4\gamma k}}{\sum_{i=0}^N e^{-r_i^2/4\gamma k}} \quad 4.56$$

where γ is a number chosen to be between zero and one. This corrected field can be iteratively computed until the error reaches some prescribed limit. This interpolation has high accuracy even when the samples are disordered and unevenly spaced. At this step, the amplitude function has been determined but further expansion of the interpolated function in an orthogonal basis is useful. Zernike polynomials were chosen in Figure 35 because of the *a priori* knowledge that the original function was composed in this basis.

This method can reproduce the original function with little error from a small number of samples. To illustrate this, 500 random amplitude functions composed of third-order Zernike amplitude polynomials ($Z_{A5} - Z_{A11}$) were sampled in the $\rho = 0.5$ and $\rho = 0.9$ pupil zones with a partially coherent 0.1σ source. The amplitude functions were randomly generated with a mean range of 19.09% of the pupil transmission deviation. These functions were then reconstructed using the scheme proposed in Figure 35. The RMS of each initial function varied, so it would be inappropriate to directly compare the RMS error (RMSE) from these cases. Instead, the error was normalized to the range of the original function to facilitate comparison. The NRMSE (as a percentage of the range) closely follows a normal distribution with a mean of 7.08% and a standard

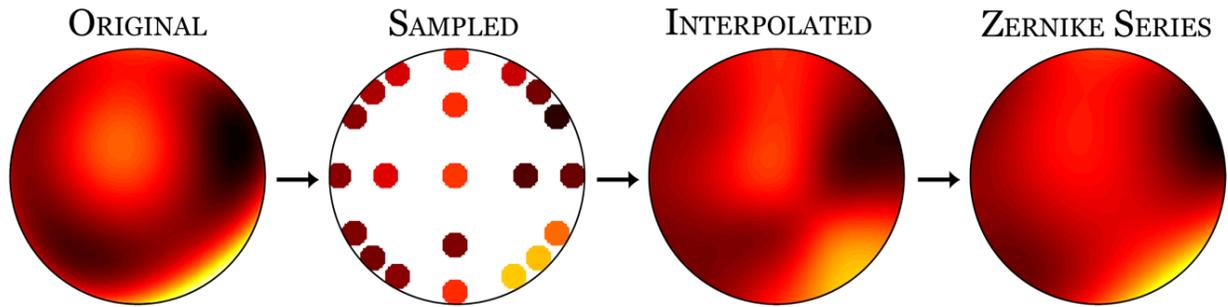


Figure 35. Flowchart of pupil amplitude reconstruction. First the pupil function is sampled via the diffracted spectrum of the targets used for phase interrogation. The value of these samples is determined via iterative fitting to aerial image simulations. Next Barnes analysis is used to interpolate between the samples to construct the original function. Finally, the interpolated function is expanded in a Fourier-Zernike series.

deviation of 1.96%. The NRMSE distribution of the 500 randomly generated functions is shown in Figure 36 with and without the use of interpolation via Barnes objective analysis, this shows that the use of interpolation decreases the mean NRMSE.

Closer inspection of Figure 35 would reveal that the error manifests itself as a smoothing of the original function. This can be seen by comparing the extracted series coefficients for one of the reconstructed functions to the exact coefficient, Figure 37a. The original functions were composed of random arrangements of third-order Zernike polynomials, but higher order terms appear in the reconstructed function. This blurring effect is from the pupil-averaging of the partially coherent source, which can be reduced by increasing both the source coherence and the number of pupil samples. Still, the function can be reconstructed with little error with only the 3rd order Zernike polynomials used to compose the original function. This is compared to an expansion in the first 36 combinations of Cartesian Legendre polynomials (Figure 37b) where the

expansion would require more terms to reach a similar RMSE. This is verified by plotting the RMSE after adding each term for both series expansions, as in Figure 37c. In this plot, the RMSE decreases at a faster rate when a Zernike basis is used.

We have used this method successfully to extract pupil variation from an EUV imaging system. The algorithm converges in around ten iterations, which takes 3.2 hours to complete on a machine with a 3.6 GHz quad-core Intel i7 processor and 16GB of RAM. This must be completed every time a wavefront is extracted.

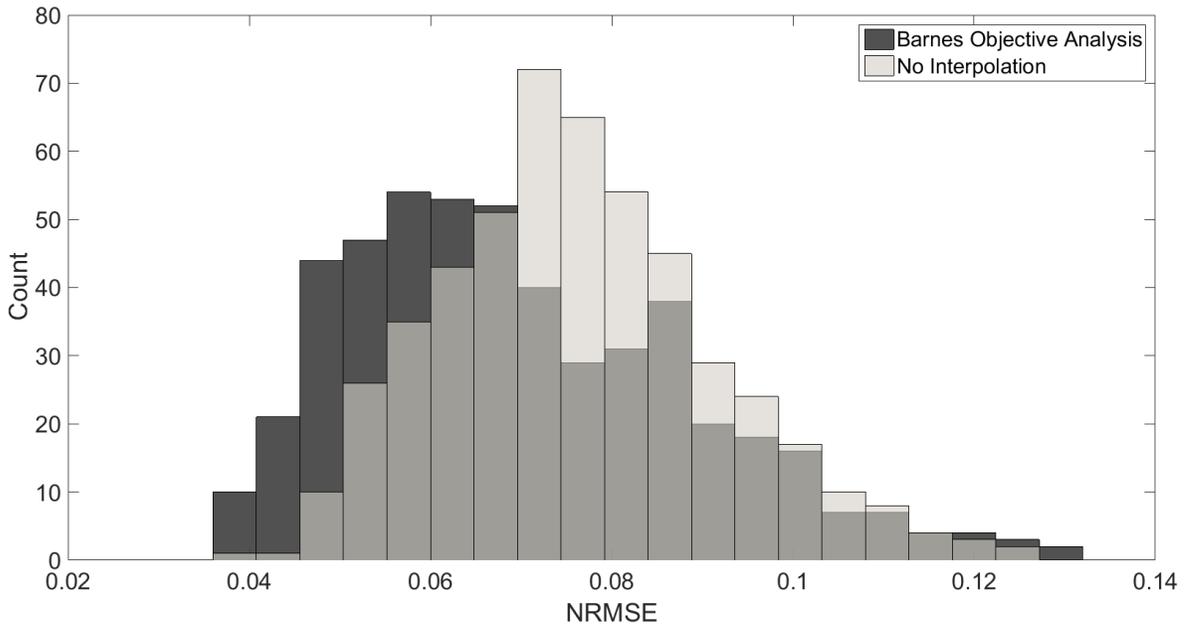


Figure 36. NRMSE distribution of the reconstructed amplitude functions with and without interpolation via Barnes objective analysis. The distribution for functions reconstructed without interpolation is semi-transparent to show the entirety of both distributions.

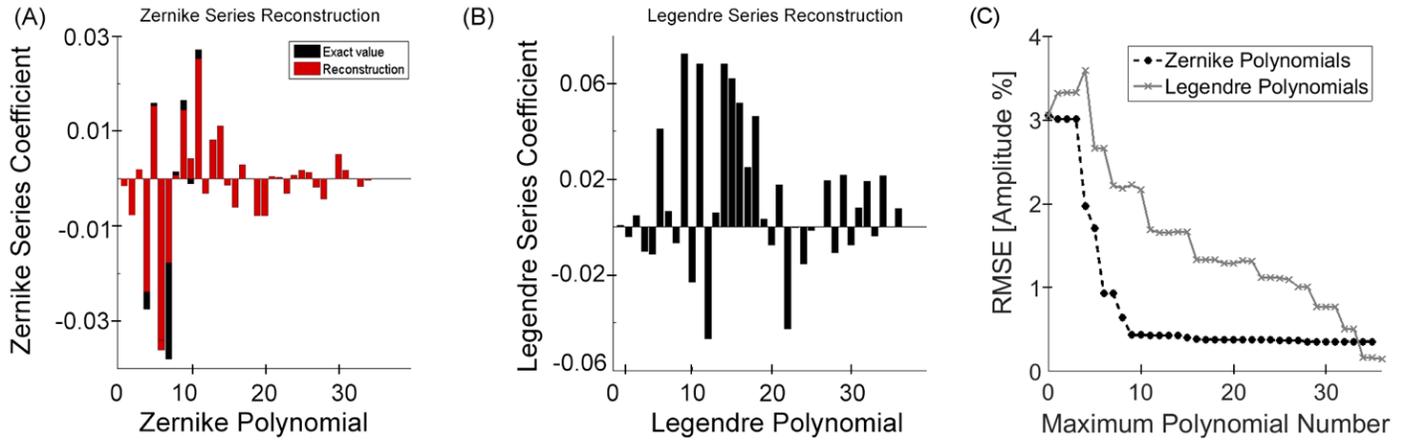


Figure 37. Comparison of an orthogonal function expansion of an interpolated amplitude function in a) Zernike polynomials, and b) the first 36 combinations of Cartesian Legendre polynomials. The RMSE after adding each additional polynomial for both series expansions is given in part c).

5. THE QUIP ALGORITHM

5.1 The QUIP Algorithm

To speed up the process of wavefront extraction, we must separate the model building and wavefront fitting steps. This can be achieved by building models linking wavefront variation to image plane behavior. This may result in a lengthy model building step, but this only needs to be completed once for each imaging system and illuminator. Afterwards, wavefront variation can be extracted rapidly.

First we will consider a 1D aerial image through focus, which we call a through-focus intensity volume, as in Figure 38a. This is part of a higher dimensional dataset when aberrations are present in the system. For example, a system with an unknown amount of spherical aberration can be described by a four dimensional dataset, as shown in Figure 38b. This dataset can be decomposed using PCA and each principal component can be interpreted as its own intensity volume, as shown in Figure 39. We will refer to these principal components as *eigenfunctions*.

The eigenfunction shows the aerial image variation caused by a certain aberration. This can be seen in Figure 39 where the first eigenfunction shows the maximum intensity of lines increasing with positive defocus. This has the effect of shifting the best focus of that pattern. The eigenfunctions of the primary phase aberrations are given in Figures 40–42. It can be seen from the eigenfunctions that astigmatism causes an orientation dependent change in best focus, spherical aberration causes a pitch dependent change in best focus, coma causes placement error and asymmetric smear across the field, and finally trefoil causes a CD difference in patterns with

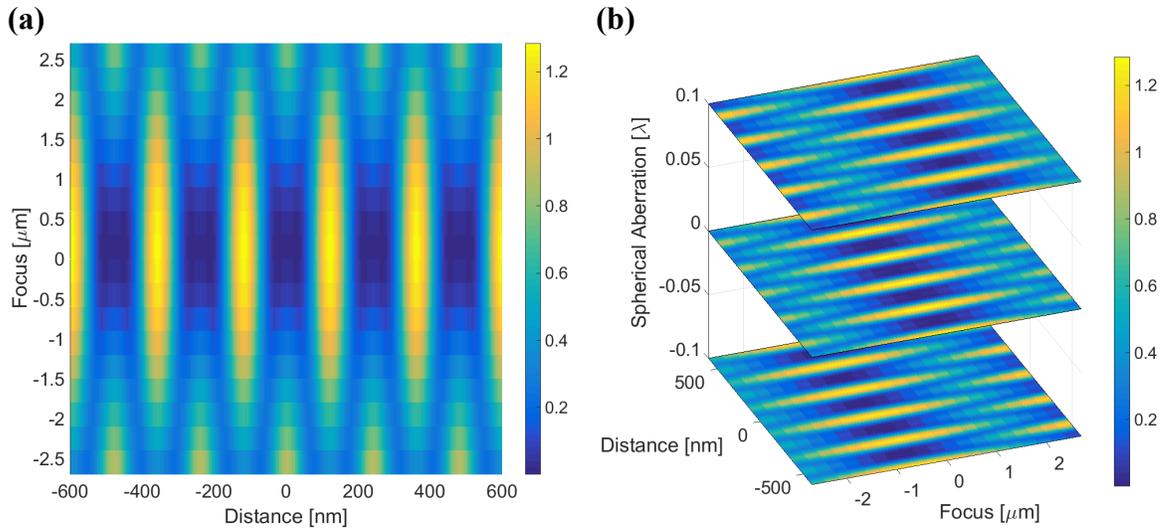


Figure 38. A 1D aerial image through defocus (a) represents an intensity volume which is part of a higher dimensional space (b) when aberrations are present in an optical system.

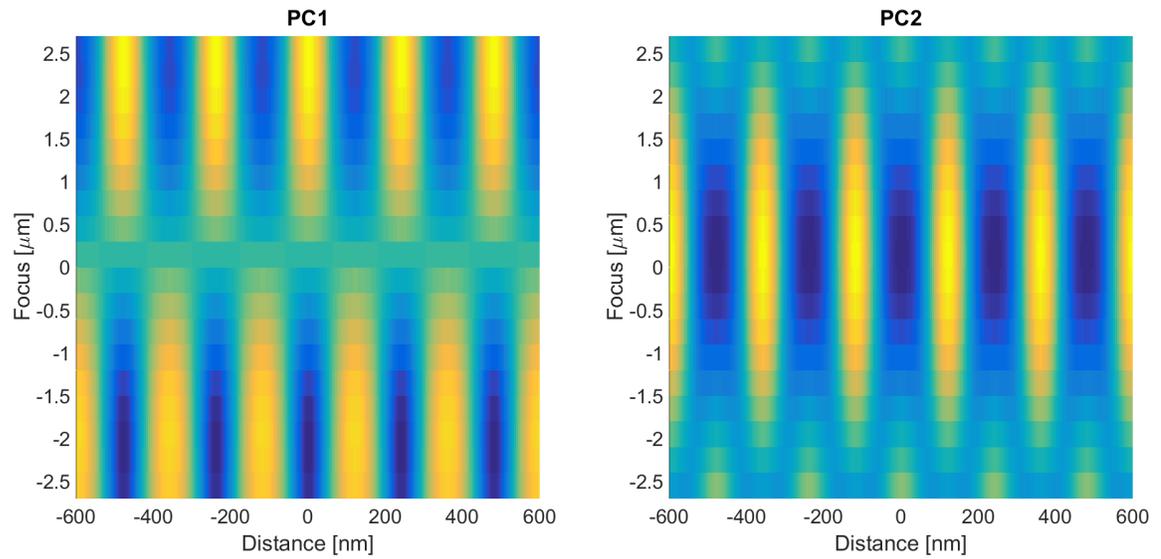


Figure 39. The first two principal components, or eigenfunctions, of the dataset shown in Figure 38.

60° symmetry. It is also apparent that even-parity phase polynomials (as given in Figures 40 and 41) cause through-focus variation unlike odd-parity phase polynomials (as given in Figure 42).

We can now determine the action of amplitude aberrations using this same type of analysis. The eigenfunctions of the primary Zernike polynomials as applied to amplitude pupil variation (given as $Z_{A,n}$) are given in Figures 43–45. It is seen that, unlike their phase counterparts, the even-parity Zernike amplitude polynomials do not cause through-focus variation, while the odd-parity components do. $Z_{A,5}$ and $Z_{A,6}$ can be seen to cause orientation dependent changes in contrast, $Z_{A,9}$ causes pitch dependent changes in contrast, $Z_{A,7}$ and $Z_{A,8}$ cause a field and focus dependent pattern smear, and finally $Z_{A,10}$ and $Z_{A,11}$ cause focus dependent changes in contrast and placement error.

We now define the function,

$$\epsilon_k(x, y) = |\Psi_k^{(0)}(x, y)|^2 - |\Psi_k(x, y)|^2 \quad (5.1)$$

where $\Psi_k^{(0)}$ is the k-th image intensity eigenfunction, computed via SOCS decomposition, of the unaberrated system and Ψ_k is the k-th image intensity eigenfunction of the aberrated system.

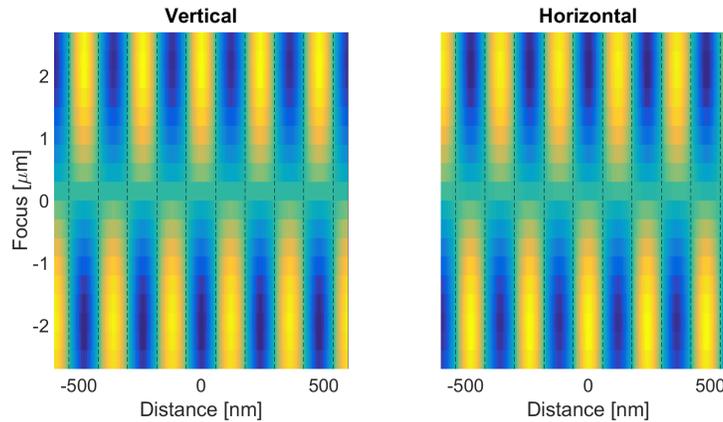


Figure 40. Eigenfunction for a 60 nm pitch line/space array with astigmatism. Because the action of astigmatism is orientation dependent the eigenfunction in this case consists of both plots together. Coordinates are given at the mask level with the dashed lines representing the mask edge.

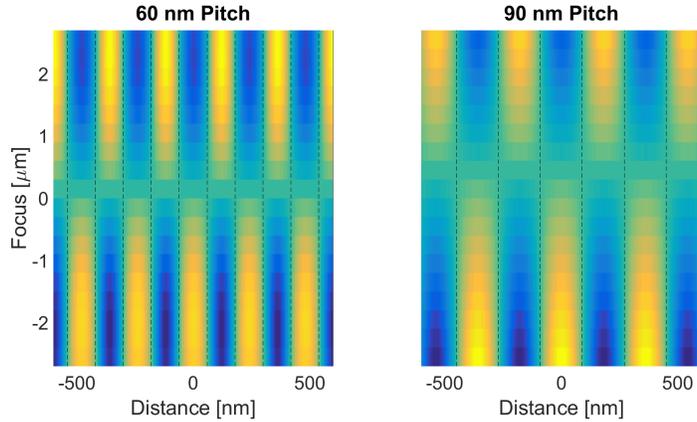


Figure 41. Eigenfunction for a 30 nm line at 60 nm pitch and 90 nm pitch with spherical aberration. Because the action of spherical aberration is pitch dependent the eigenfunction in this case consists of both plots together.

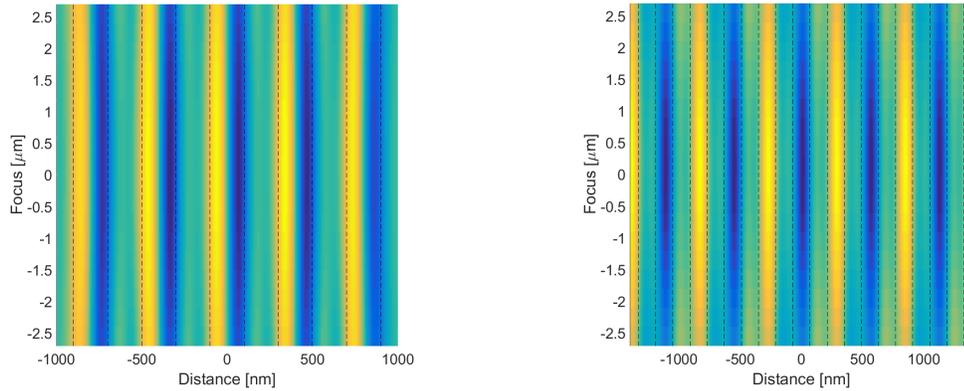


Figure 42. Eigenfunction for a 50 nm CD five-bar pattern with coma (left) and eigenfunction for a 35 nm brick wall structure with trefoil (right). The left most feature in this trefoil eigenfunction represents the right side of a bar.

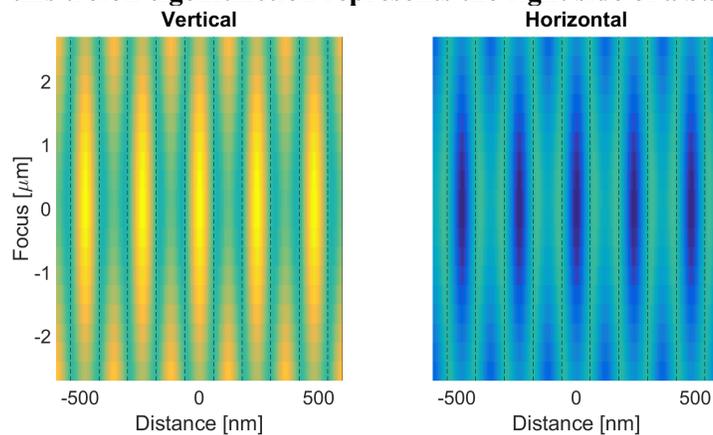


Figure 43. Eigenfunction for a 60 nm pitch line/space array with Z_{A5} and Z_{A6} . Because the action of these aberrations are orientation dependent the eigenfunction in this case consists of both plots together.

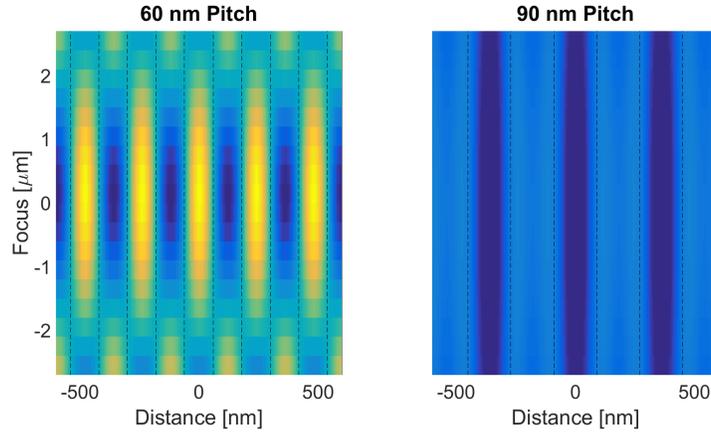


Figure 44. Eigenfunction for a 30 nm line at 60 nm pitch and 90 nm pitch with Z_{A9} . Because the action of Z_{A9} is pitch dependent the eigenfunction in this case consists of both plots together.

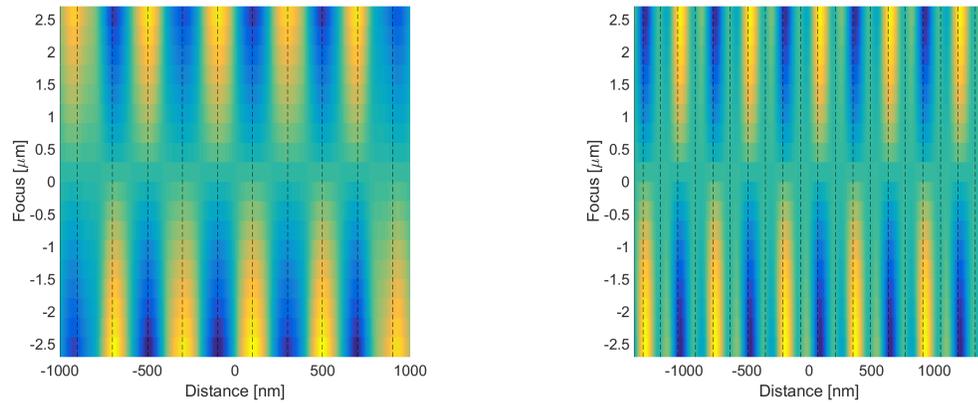


Figure 45. Eigenfunction for a 50 nm CD five-bar pattern with Z_{A7} and Z_{A8} (left) and the eigenfunction for a 35 nm brick wall structure with Z_{A10} and Z_{A11} (right). The left most feature of the Z_{A10} and Z_{A11} eigenfunction represents the right side of a bar.

The coma eigenfunction (Figure 42) is compared in Figure 46 to $\varepsilon_1(x)$ computed via Equation (5.1 for this system with coma, as shown in. There is a RMS error (RMSE) of 0.67×10^{-3} between these two curves. Similar error, as given in Table 2, is obtained by comparing the other eigenfunctions shown in Figures 40–45 to the appropriate $\varepsilon_k(x)$ through defocus. It is clear now that $\varepsilon_k(x)$ represents the directions of highest variation in the image intensity when aberrations are present.

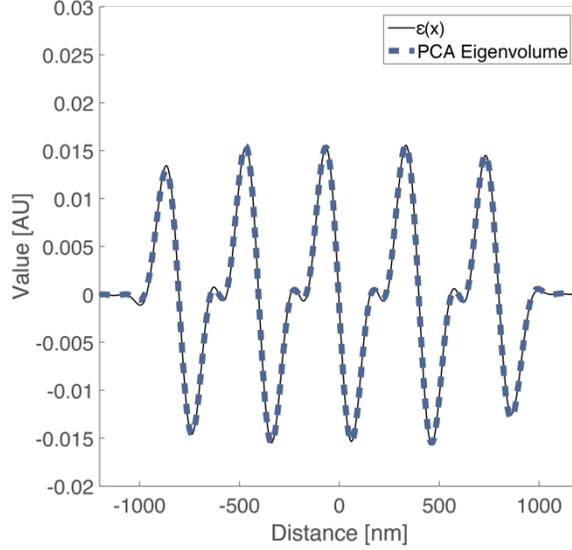


Figure 46. Comparison of $\varepsilon_1(x)$ and the PCA eigenfunction for an imaging system with coma (Z_7) with zero defocus.

Table 2. Phase and amplitude RMSE between $\varepsilon_1(x)$ and eigenfunctions computed via PCA.

Name	Phase RMSE	Amplitude RMSE
Z_5/Z_6	0.50×10^{-3}	0.19×10^{-3}
Z_7/Z_8	0.20×10^{-3}	0.20×10^{-3}
Z_9	0.67×10^{-3}	0.10×10^{-3}
Z_{10}/Z_{11}	0.89×10^{-3}	0.82×10^{-3}

The projection, or inner product, of an aberrated aerial image onto the eigenfunction yields a single number. The curve of projection coefficients therefore forms a function that relates the amount of aberration in the pupil to an image space measurement. An example of this function for coma is given in Figure 47. The amount of aberration in an aerial image can be determined if the projection coefficient has odd parity and can be inverted into a function. More details on obtaining inverse pupil solutions will be given in Section 5.3.

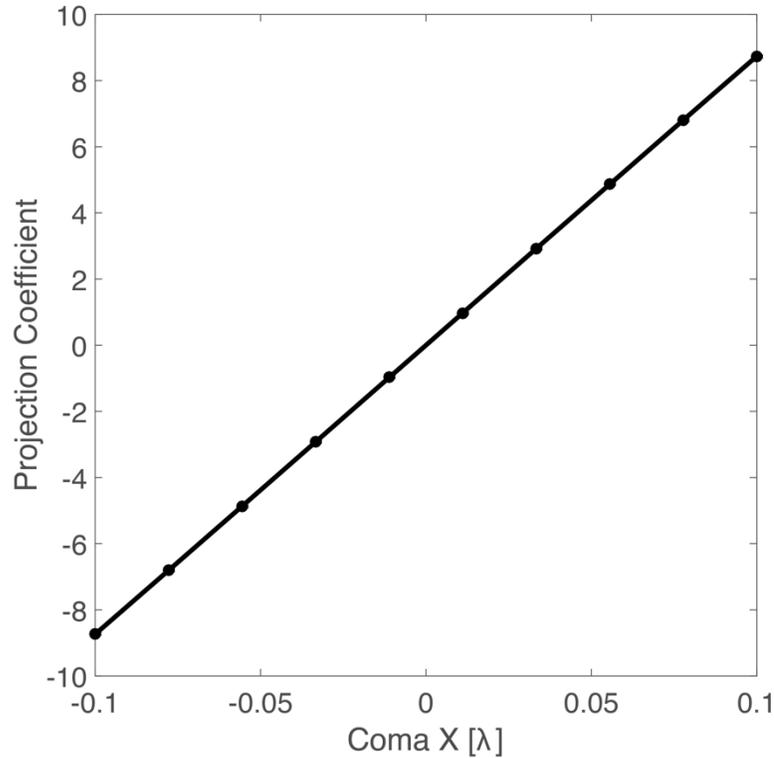


Figure 47. Coefficient for aerial images aberrated with coma projected onto the coma eigenfunction. The slope of the line may change with changes in the imaging system, but this function is representative of the projection function for coma.

5.2 Δ CD Eigenfunctions

As mentioned previously, lithographers do not typically directly measure the aerial image. It is therefore advantageous to consider the CD analogue to aerial image aberration eigenfunctions. We have seen that the eigenfunctions of an aberrated intensity volume are related to the aberrated components of the TCC matrix. Therefore, one could calculate the eigenfunctions using either method and obtain identical eigenfunctions. While it is always possible to compute aerial images from a TCC, it is not possible in general to find CDs from an aerial image. For the simplest of cases an analytic CD function may be possible, but in practice root-finding algorithms are the safest approach. This property means that while the aerial image aberration eigenfunctions could be

computed directly, there is no such CD eigenfunction that can be computed directly from either the aerial image or aerial image eigenfunction.

There are two ways to calculate the aerial image eigenfunctions, so we should consider using PCA to find eigenfunctions of the ΔCD function. Examples of this are given in Figures 48–51 for astigmatism, coma, spherical aberration, and trefoil. To compute these the ΔCD function is computed as in Figure 30, then PCA is used to find the eigenfunction representing the variation due to aberrations. The projection of ΔCD functions onto their respective eigenfunctions is also shown. Each of the curves is invertible into a single-valued function and can therefore be used to measure aberrations from experimental CD measurements. It appears in Figure 50 that the model for spherical aberration is given by a quadratic curve, and therefore would exhibit degeneracy over a wider aberration range. However in Figure 52 it can be seen that over a range of $\pm\lambda/4$ the projection function is actually a cubic function rather than quadratic. Therefore, there is no risk of degeneracy in modeling spherical aberration in this way.

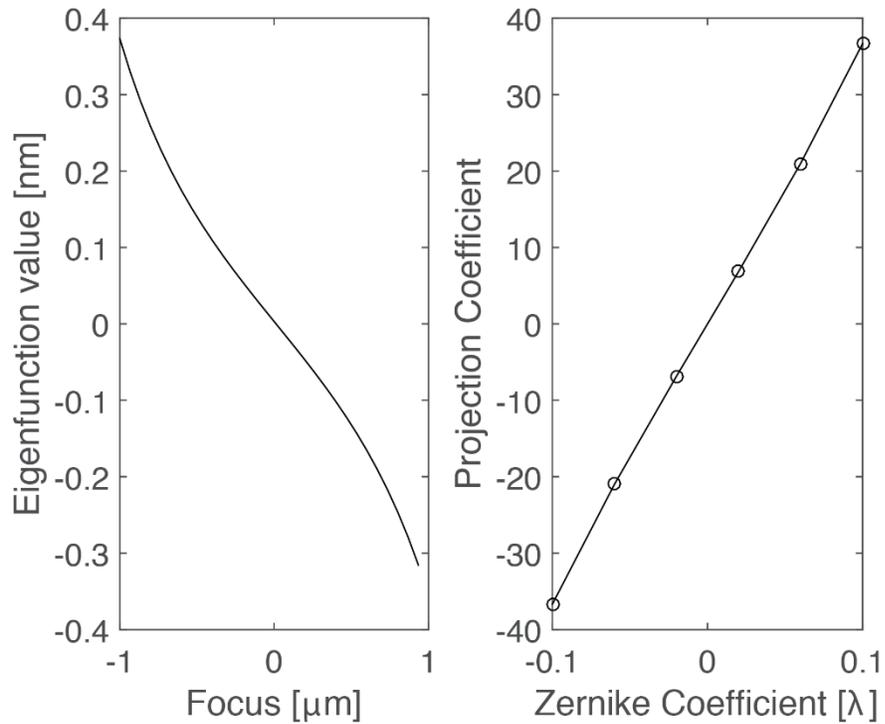


Figure 48. An example of a Δ CD eigenfunction (left) and the projection curve (right) for astigmatism. Astigmatism is interrogated by the CD difference between orthogonal lines and spaces.

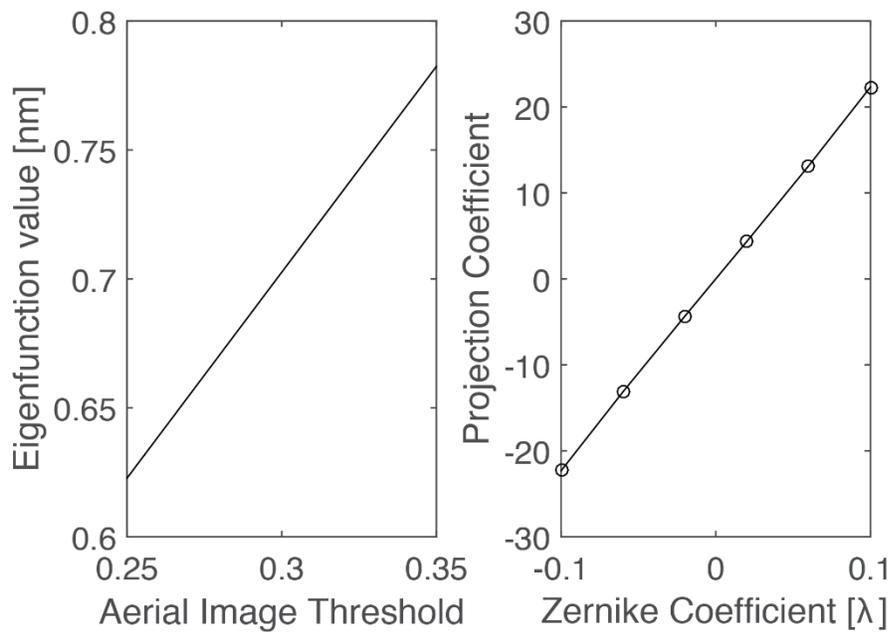


Figure 49. An example of a Δ CD eigenfunction (left) and the projection curve (right) for coma. Coma is interrogated by the CD difference between the outer two bars of a five bar structure.

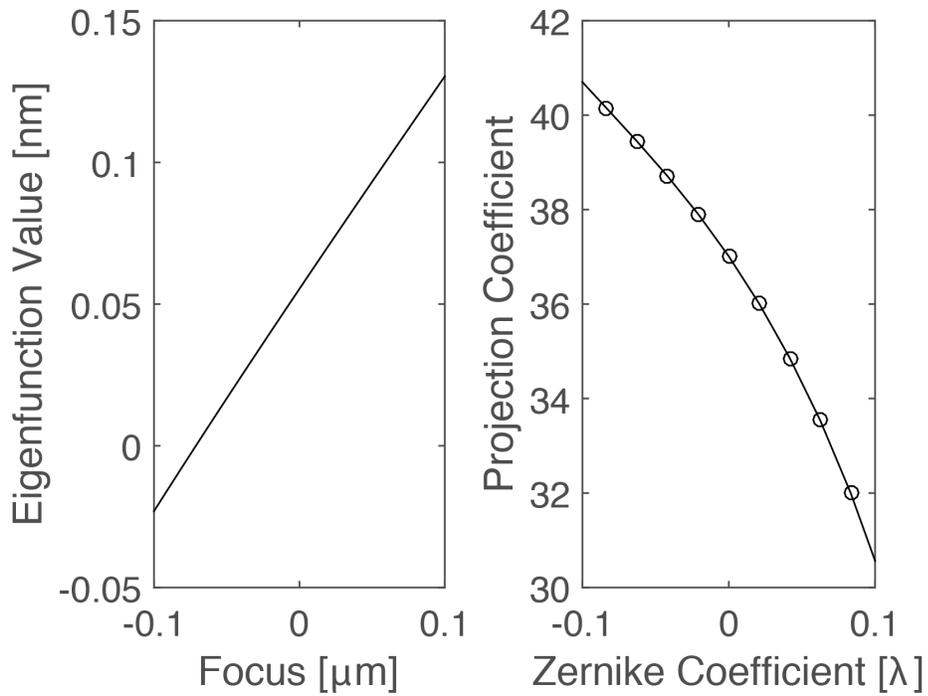


Figure 50. An example of a Δ CD eigenfunction (left) and the projection curve (right) for spherical aberration. Spherical is interrogated by a heuristic computed from a constant CD line at pitches 1:1, 1:2, and 1:4.

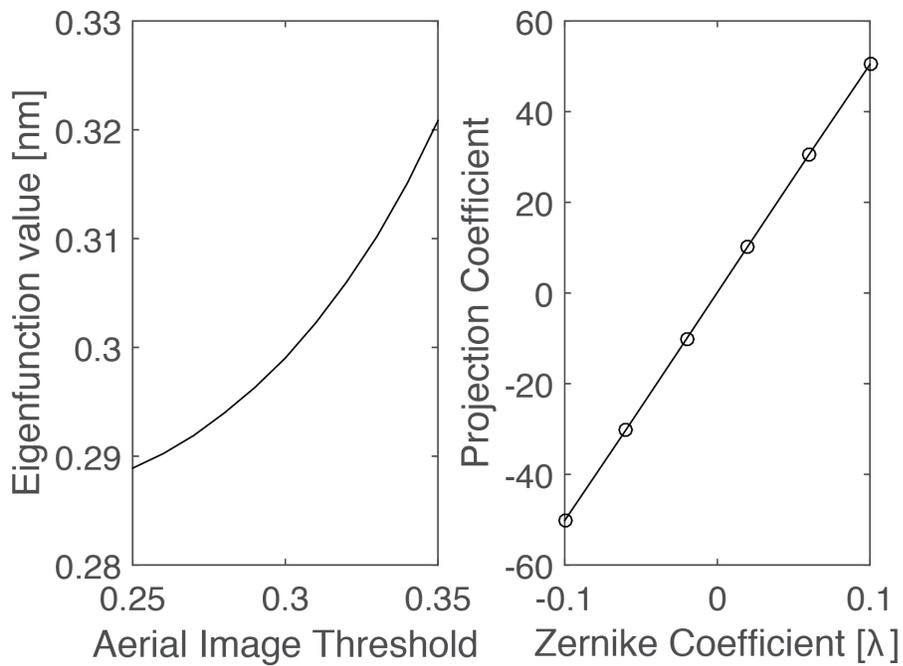


Figure 51. An example of a Δ CD eigenfunction (left) and the projection curve (right) for trefoil. Trefoil is interrogated by the CD difference between orthogonal lines and spaces.

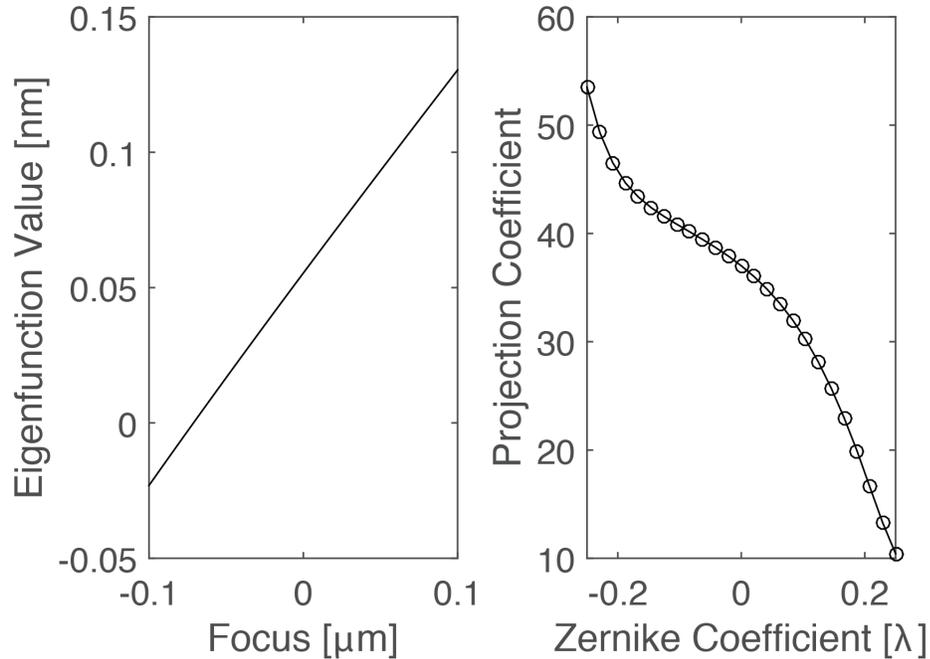


Figure 52. An example of a ΔCD eigenfunction (left) and the projection curve (right) for spherical aberration over an extended range of $\pm 250\text{m}\lambda$.

As we saw in Section 3.1, the effects of certain aberration terms are confounded in the image domain. The eigenfunction can therefore be interpreted as being a function of multiple aberration terms. For example, the eigenfunction corresponding to coma is shown in Figure 53 for varying amounts of trefoil. There are five lines plotted in Figure 53, despite it appearing that there are only three, because the eigenfunction variation is even-parity in the Z_{10} axis. However, the projection coefficient onto the correct eigenfunction is unique for coma given varying amounts of trefoil, as seen in Figure 54. The amount of coma in the presence of trefoil can therefore be determined because each (Z_7, Z_{10}) point maps to a unique projection coefficient.

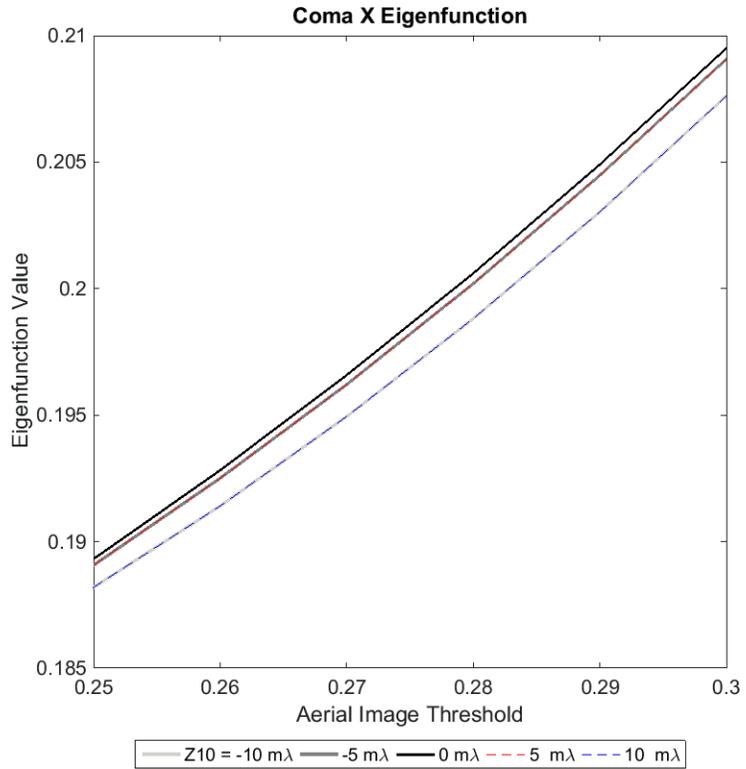


Figure 53. Comatic eigenfunction with varying amounts of trefoil. There are five lines plotted, despite only three being visible. This is because two are on top of each other because the eigenfunction is even in the Z10 axis.

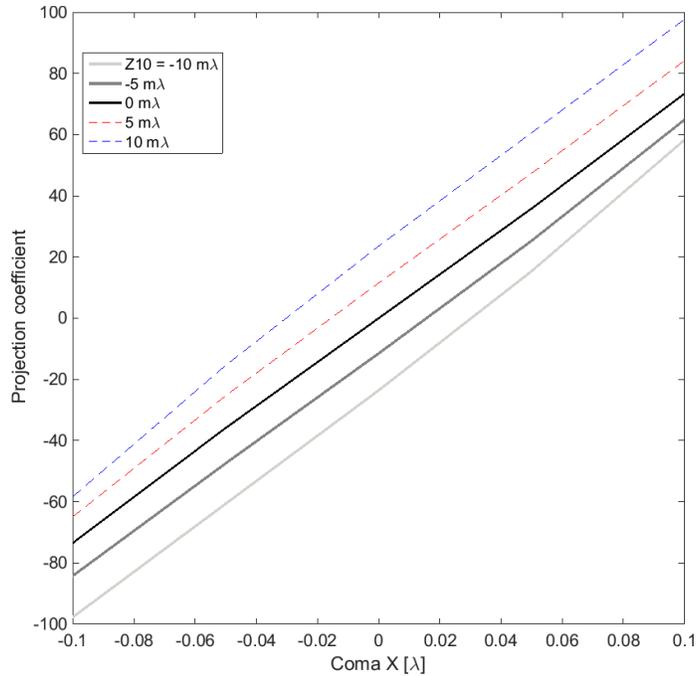


Figure 54. Projection coefficient for coma Δ CD functions onto the appropriate eigenfunction with different amounts of trefoil.

5.3 Obtaining inverse pupil solutions

Aberrations can be extracted using the QUIP algorithm as follows: 1) the appropriate eigenfunctions are determined, 2) the projection functions are determined for each eigenfunction, 3) experimental data is projected onto the eigenfunctions with a guess for interacting terms, 4) the model is inverted, 5) the process is repeated iteratively. This process is shown as a flowchart in Figure 55 and is explained in more detail in the rest of this section.

First the targets and source shape are chosen. The aberration separable basis (ASB) can be pre-computed via computation of the TCC if aerial image data is used. Otherwise, the basis will need to be computed via PCA. Next, a simulated full factorial experiment is run. Using aerial image data this experiment results in a set of intensity volumes, and results in a set of ΔCD functions if using CD data. Each simulated function (intensity volume or ΔCD) is projected onto the ASB eigenfunctions and a model is constructed between the treatment combinations and projection coefficients. Each model consists of a system of polynomial equations fit to the projection coefficients via non-linear least squares. If the relationship is linear the model can be formalized as

$$C^{(n)} = \mathbf{tc} \cdot \beta^{(\mathbf{n})} + \bar{C}^{(n)}, \quad (5.2)$$

where $C^{(n)}$ is the projection of each intensity volume onto the n^{th} eigenvolume, $\beta^{(n)}$ represents the model coefficients, and $\bar{C}^{(n)}$ represents the mean projection coefficient onto the n^{th} eigenvolume. In the linear case this could be formed into a matrix equation as

$$C^{(n)} = \begin{bmatrix} tc_{11} & tc_{12} & tc_{13} & \dots & tc_{1n} \\ tc_{21} & tc_{22} & tc_{23} & \dots & tc_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ tc_{m1} & tc_{m2} & tc_{m3} & \dots & tc_{mn} \end{bmatrix} \cdot \begin{bmatrix} \beta_1^{(n)} \\ \beta_2^{(n)} \\ \vdots \\ \beta_n^{(n)} \end{bmatrix} + \bar{C}^{(n)} \quad (5.3)$$

Finally, an experimental intensity volume can be projected onto the eigenvolumes and the model can be inverted as,

$$\mathbf{X} = \beta^{-1}(\hat{\mathbf{C}} - \bar{\mathbf{C}}), \quad (5.4)$$

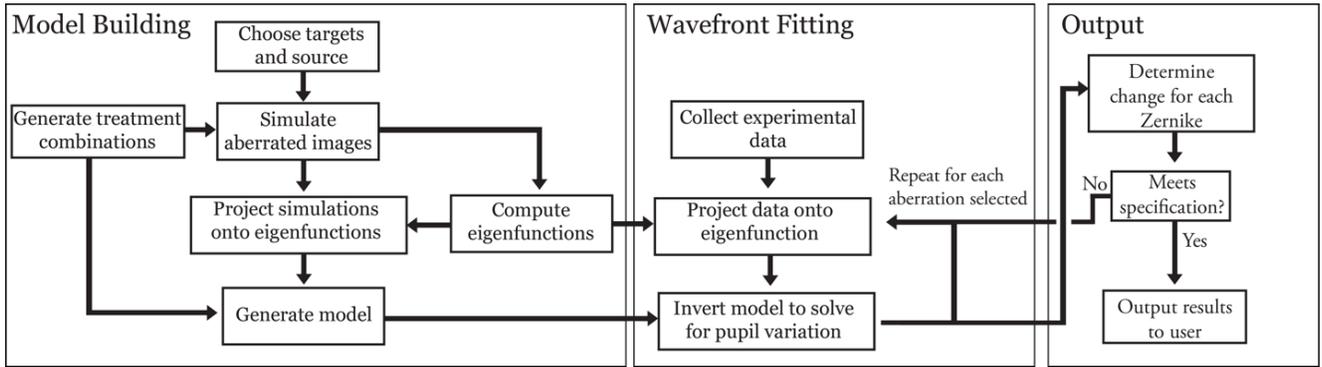


Figure 55. A flowchart for the aberration separable basis algorithm of image-based pupil characterization.

where $\hat{\mathbf{C}}$ is the projection of an experimental intensity volume onto the model eigenfunctions. In reality, the relationship between treatment combinations and projection coefficients is rarely linear, as seen in Figure 54. The techniques outlined in Equations 13-15 can be easily adapted to higher-order polynomials, or the projection function can be interpolated through known points. For example, in a four-dimensional space with only spherical aberration the forward model is generalized with an n-degree polynomial as

$$C^{(1)} = \bar{C}^{(1)} + \beta_1^{(1)} a_9 + \beta_2^{(1)} a_9^2 + \dots + \beta_n^{(1)} a_n \quad (5.5)$$

In the non-linear case the model cannot be written in a matrix form and must be inverted by solving the system of equations numerically.

If the eigenfunctions include interacting terms, then an initial guess must be given to project the function into a single dimension in aberration space. For example, when fitting coma in the presence of trefoil a guess must be given for trefoil. Similarly, when fitting trefoil in the presence of coma a guess must be given for coma. This process can be repeated to refine both terms.

Fitting aberrations to Δ CD eigenfunctions can have errors below $0.25 m\lambda$ ($3.4 pm$ for $\lambda = 13.5 nm$) with only a small number of simulations. A set of points randomly distributed within the parameter space can be used to verify a model. In Figure 56, 25 random points were used to verify a model trained with only seven aberration treatment combinations to obtain an RMSE of $0.12 m\lambda$. Similarly, coma in the presence of trefoil achieves an RMSE of $0.22 m\lambda$ using a model generated from a 7^2 full factorial experiment.

These models only need to be built once, as compared to the purely iterative algorithm where many of the same simulations would be repeated with each iteration. An example of an iterative solution is shown in Figure 58 for both third-order coma and trefoil. This requires two models be built: 1) coma in the presence of trefoil and 2) trefoil in the presence of coma. Because the models have already been built, each iteration is very fast—100 iterations can run in under one second.

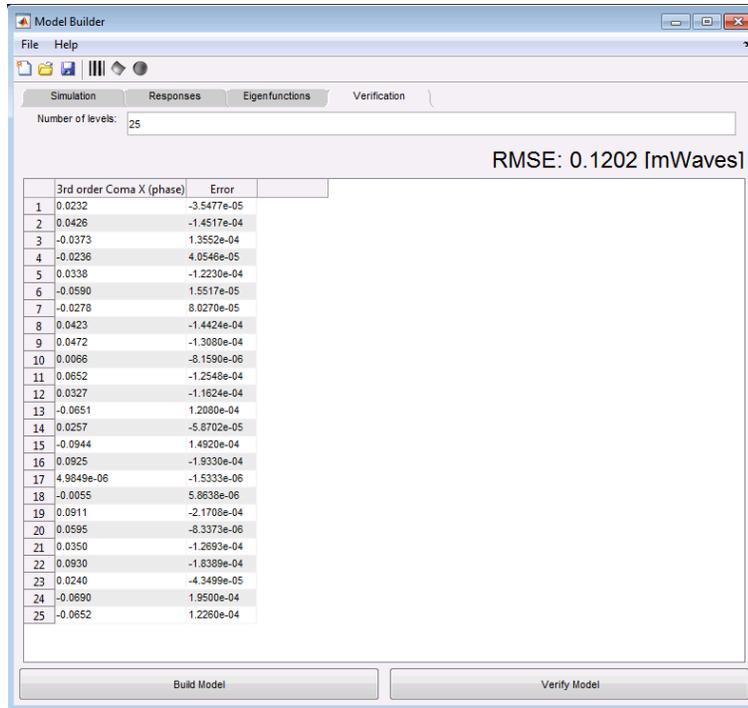


Figure 56. An example of model verification using 25 Gaussian distributed points. Training the model on only seven aberration treatment conditions gives an RMSE on the order of single picometers.

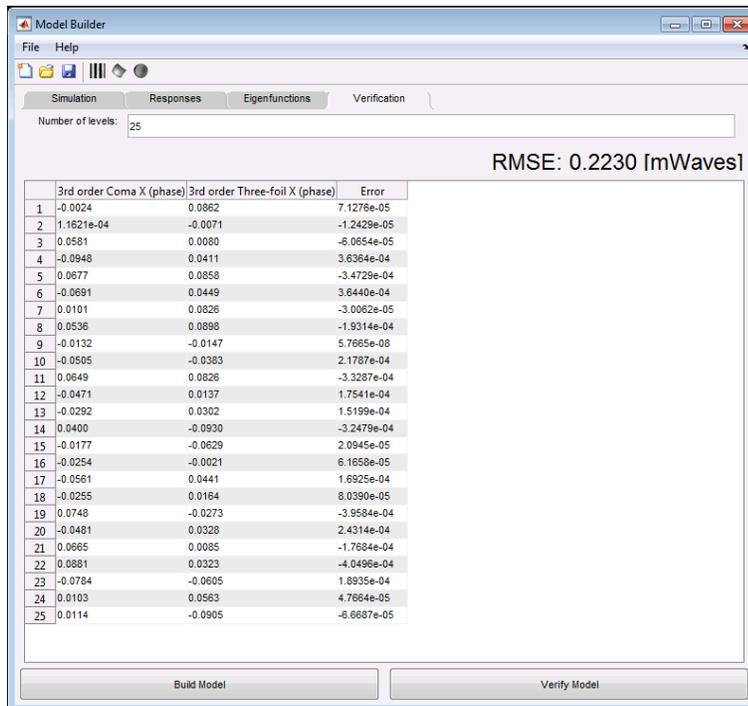


Figure 57. An example of model verification using 25 Gaussian distributed points for coma in the presence of trefoil.

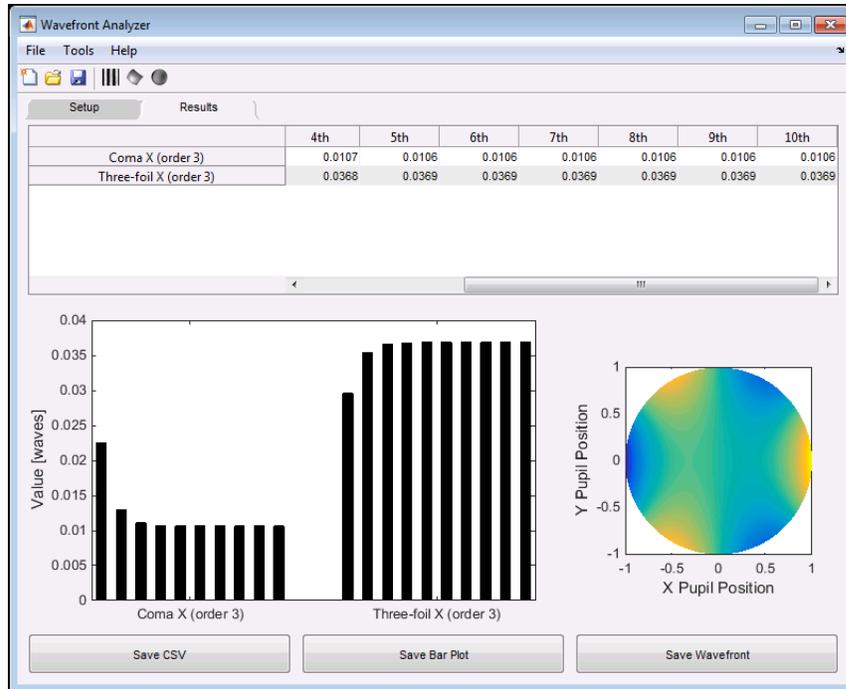


Figure 58. An iterative solution for both coma and trefoil obtained via the QUIP eigenfunction algorithm.

5.4 Measuring high-order aberrations

Pure low order Zernike aberrations do not exist in real lithography systems. The aberrated wavefront is most accurately described by a combination of low- and high- order terms. Trefoil is not a high-order comatic term, but patterns that are sensitive to coma will be sensitive to trefoil and vice-versa. This is equally true for low- and high-order terms of the same kind of aberration. Therefore, measuring high-order terms poses the same computational challenges as measuring coma and trefoil together.

As outlined in Section 3.1, high-order terms can be measured by using multiple pitches of the same type of target. This gives enough information to distinguish between low- and high-order terms in the image domain. Using the QUIP algorithm, high-order terms are fit by creating a model for the low-order term in the presence of the high-order term, and vice versa. Then the algorithm proceeds as above for coma and trefoil.

Any algorithm that measures high-order aberrations needs to perform accurately under three types of situations: a wavefront composed from only third-order, one composed from only fifth-order aberrations, and one composed from a mixture of both third- and fifth-order aberrations. These three test cases assure that the algorithm is not changing the low-order terms and is correctly fitting the high-order terms.

Examples of these three test cases are shown in Figures 59 –61, where the solved value of each aberrations is shown at the end of each iteration. In the case of Figures 59 and 60, it can be seen that the value of fifth- and third- order astigmatism (respectively) is zero. Similarly, it can be seen in Figure 61 that the iterative solution seems to oscillate between two values while converging. This is expected because the two terms are confounded. We begin by assuming no aberrations, and so the algorithm tried to correct phase errors with both terms. These sum into one term that is effectively double what it should be. In response, both terms will decrease due to the initial overcompensation. Both terms will in turn increase due to an undercompensation, *ad infinitum*.

An additional 24 trials of each case were run (25 total for each case). In each case the mean error was below 0.001λ . This error is slightly higher than can be achieved with synthetic data composed of purely third-order aberrations. However, it is still a very acceptable level of error.

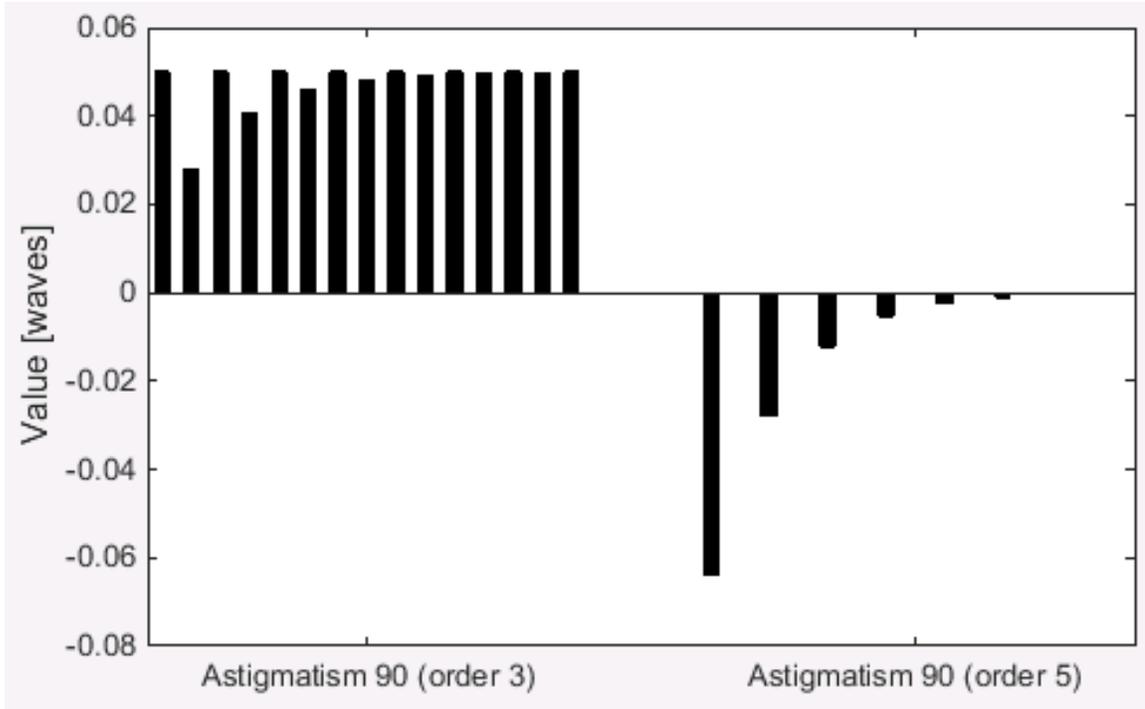


Figure 59. An example of measuring a synthetic wavefront composed from only third-order astigmatism using the QUIP algorithm.

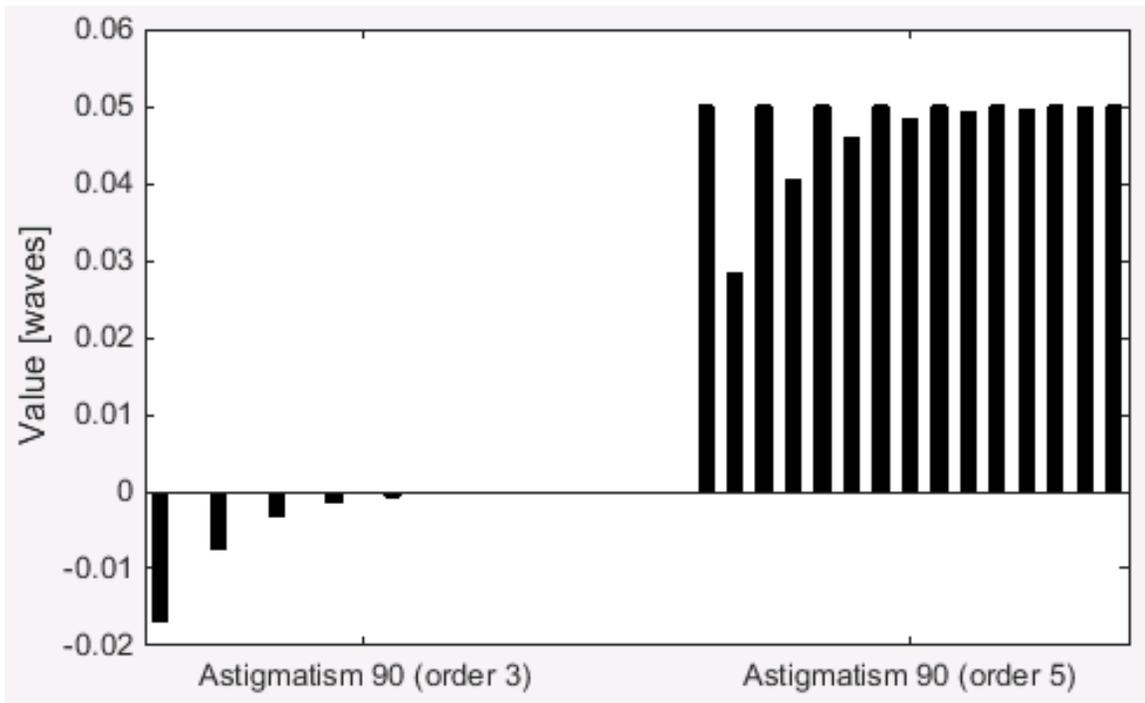


Figure 60. An example of measuring a synthetic wavefront composed from only fifth-order astigmatism using the QUIP algorithm.

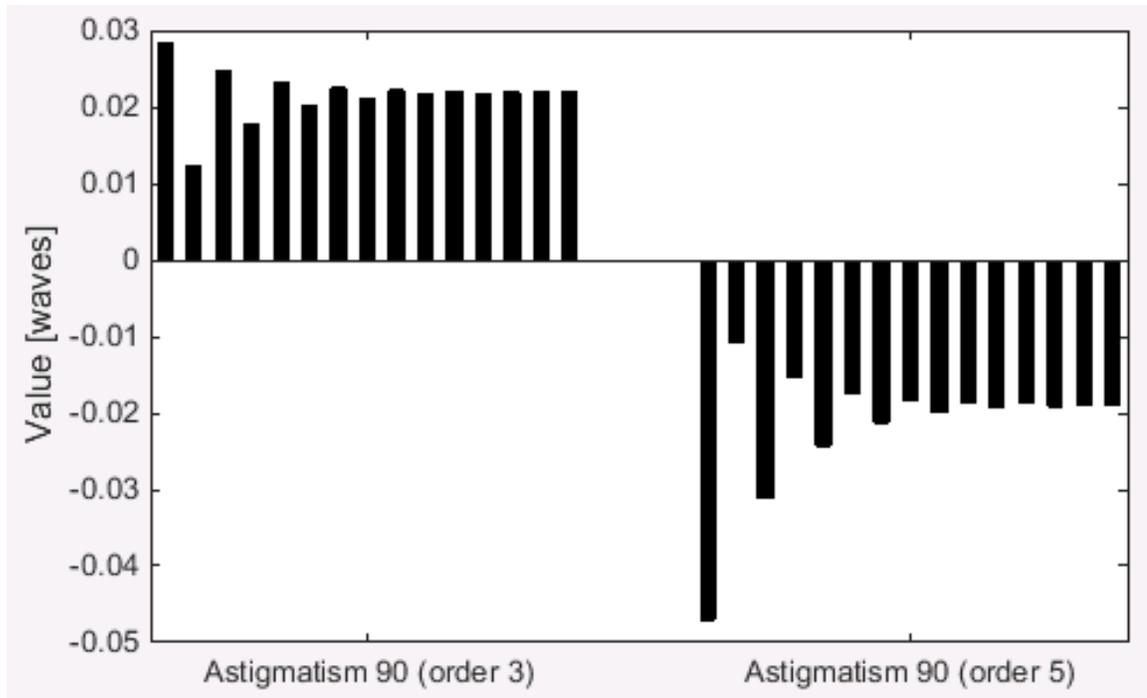


Figure 61. An example of measuring a synthetic wavefront composed from both third- and fifth-order astigmatism using the QUIP algorithm.

5.5 The effects of noise on aberration retrieval

There can be a considerable amount of variation in CD-SEM measurements and the aberration retrieval method involves non-linear least squares regression. It is therefore worthwhile to consider the effect of this noise on regression. The CD variation was found to closely follow a normal distribution. An example histogram of the difference between individual measurements and the mean CD at each dose is given in Figure 62 for the left CD of a five-bar structure. In this specific example, the standard deviation is 0.42 nm and the standard deviations of the other structures are given in Table 3. Synthetic data was then computed by adding Gaussian noise to a linear regression to the experimental data. This synthetic data can stand in as an analogue of the real data for further analysis.

Figure 63 shows the linear regression of the synthetic data with a 95% confidence interval for the slope. The actual line that the data was computed from is contained within the confidence interval but does not match the linear regression on the noisy data. The aberration levels are extracted from a difference in CD and will not be accurate if only a simple linear regression is used. However, the confidence intervals provide bounds on the CD, and therefore the aberration levels. Aberration extraction should proceed as follows: 1) compute confidence intervals for both CDs used for extraction, 2) determine the confidence interval for the difference in CDs from the edges of the CD confidence intervals, 3) fit aberrations to the edges of the CD difference confidence interval to determine bounds on the aberration levels, as in the flowchart in Figure 64.

To better understand the error limits of this method CDs were computed for the unaberrated case and Gaussian noise was added to the data. The mean was zero for all cases and the standard deviation was as given in Table 3. Confidence intervals were computed and aberrations were extracted as in Figure 64. This process was repeated 10 times and the mean extracted value for each aberration type is given in Table 4. These values provide an error measurement for the algorithm which depends on the standard deviation of the noise for each structure. In all cases the uncertainty is below 0.001λ .

Table 3. Standard deviation of CD measurements for each structure type

Aberration Name	Structure Type	Standard Deviation [nm]
Astigmatism	Line/Space Array	0.36
Coma	5-bar Structure	0.42
Trefoil	T-Bar	0.45

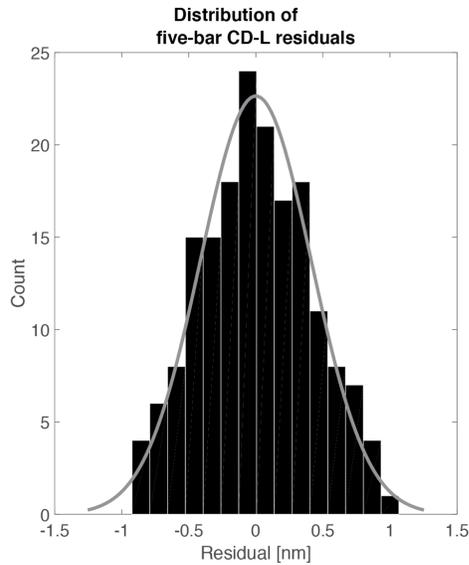


Figure 62. Histogram of residuals for the left CD of a five-bar structure. The residuals are computed between individual CD measurements and the mean at that dose. The residuals closely follow a normal distribution with a mean of 0 nm and a 0.42 nm standard deviation.

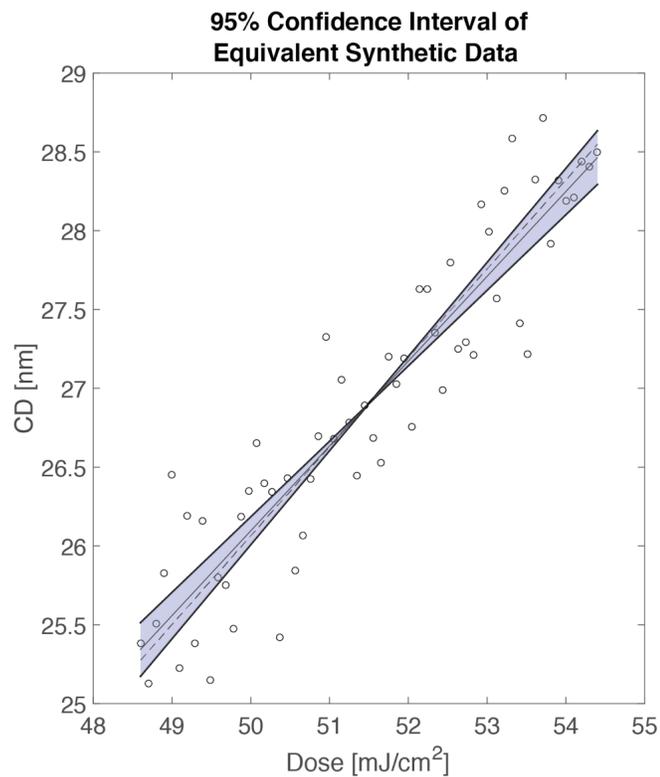


Figure 63. Synthetic CD data computed from the line $CD=0.5642 \cdot E-2.1448$ with Gaussian noise with a 0.42 nm standard deviation. The dashed black line is the true line from which the data was computed. A 95% confidence interval for a linear fit to the data is shown in the shaded region between the thick black lines. The true line is within the confidence interval, but is not the line of best fit, given by the solid thin black line.

Table 4. Aberrations extracted from confidence intervals based around the ideal CD (no aberration) with noise given by the standard deviations in Table 3.

Aberration Name	X aberration [$m\lambda$]	Y aberration [$m\lambda$]
Astigmatism	0.005	-0.548
Coma	0.336	-0.466
Trefoil	-0.752	-0.396

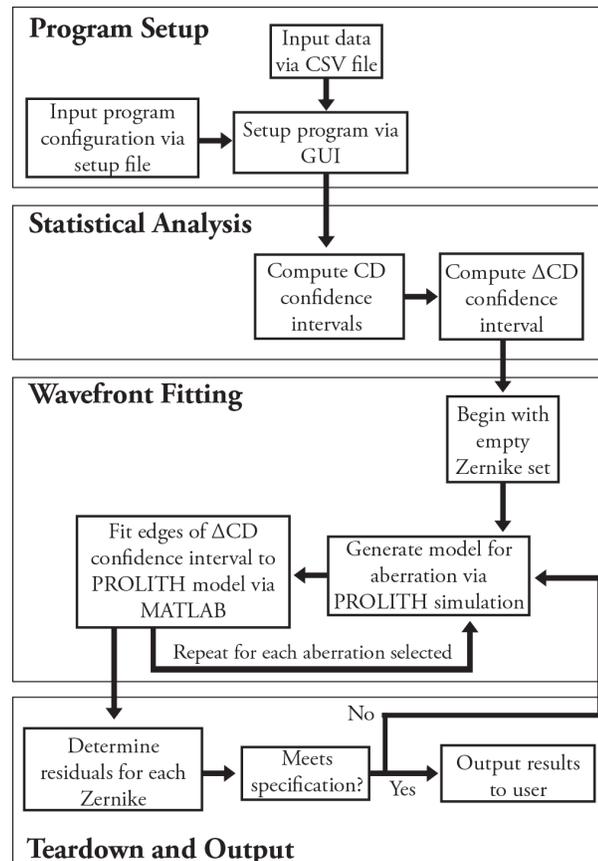


Figure 64. Flowchart of aberration extraction using the image-based method adapted for confidence intervals.

5.6 QUIP Software

QUIP provides a methodology to measure the aberrations of a lithography system from images formed by that system. The software package contains three major components:

1) image viewer, 2) model builder, 3) wavefront analyzer. Additionally, there are two

different wavefront fitting algorithms that can be used. The necessary software components change depending on which algorithm is in use.

5.6.1 *Image viewer*

The first step to wavefront analysis in QUIP is to use the image viewer to extract the necessary image data. The software can read CD-SEM and scanner lot reports to automatically link this data to image measurements. The image viewer interface is shown in Figure 65. To begin, either CD-SEM micrographs or through focus image stacks are loaded into the software. Stage errors can be corrected through subpixel image rotation and registration algorithms. If necessary, noise can be filtered from the image and, finally, a region of interest is selected and optionally interpolated. The CD or through-focus intensity volume can then be extracted from this region.

In CD-SEM micrographs there is often a large amount of high frequency noise originating from secondary electron noise. This type of noise affects CD measurements but can be filtered by using either a median filter or a Savitzky-Golay filter. An example of unfiltered vs. median filtered image is given in Figure 66a. The effect of the filter is more easily seen when comparing intensity line scans between the two images, Figure 66b.

Rotation correction is computed by selecting an edge which should be straight. The software can then determine the appropriate offset to correct the rotation of this edge. Other stage errors can be corrected through the registration of the image stack. Image translation is determined with subpixel accuracy via the frequency domain phase of the cross-correlation of adjacent images.⁶⁹ Next, the image can be interpolated to a higher pixel grid, which also deconvolves the response of the CCD sensor. This process is

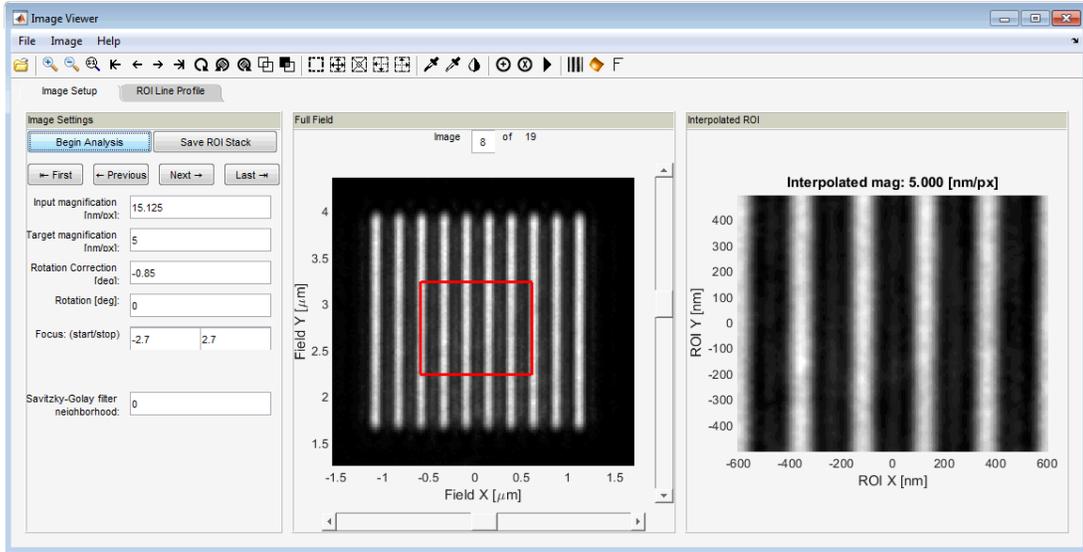


Figure 65. The Image Viewer interface showing a sample image from an EUV actinic mask microscope.

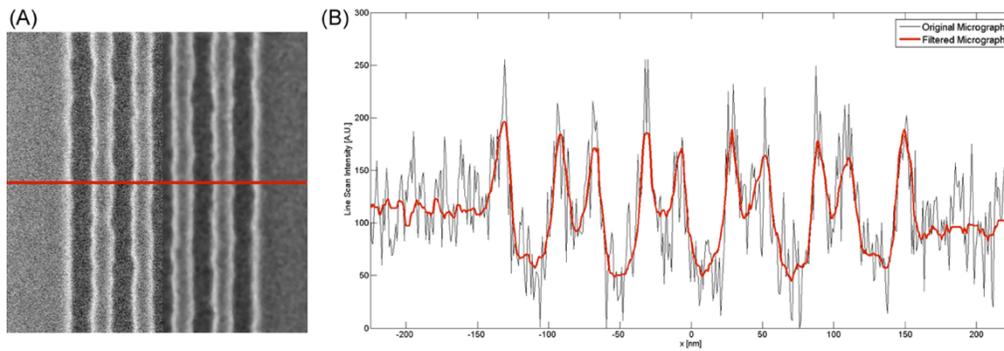


Figure 66. Example of SEM micrograph image processing for NXE:3100 case study. a) A comparison of the original (left) and median filtered (right) images, b) plot of a intensity line scan across the line in part (a) for both the original and filtered images.

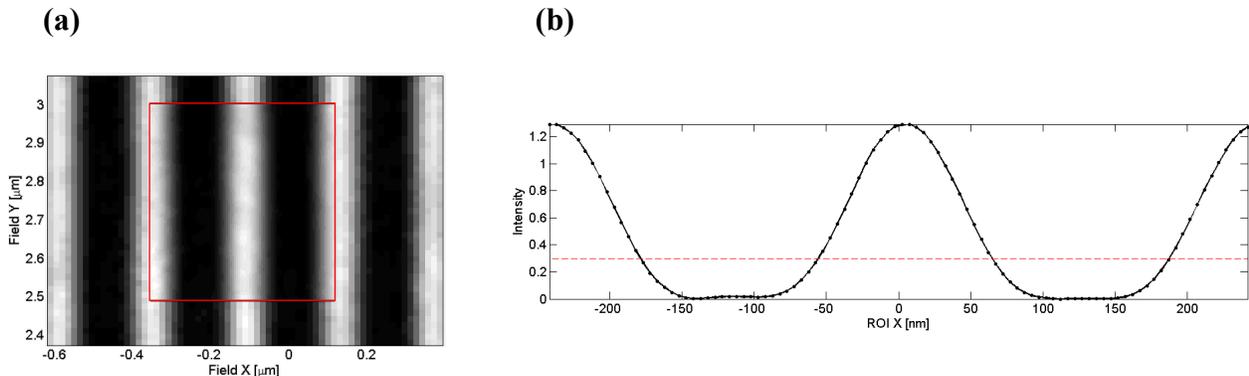


Figure 67. (a) An example image at the original resolution and after Fourier interpolation (inset). (b) Column-wise average of the interpolated region in the inset of panel (a) to estimate the aerial image.

shown in Figure 67a, where an image collected at approximately 15 nm/px was interpolated to approximately 5 nm/px. Finally, the aerial image or SEM line scan can be approximated via the normalized column-wise median of the interpolated region. The CD can then be measured by thresholding the profile, or the through-focus intensity volume can be captured by saving the profile from a through-focus image stack.

5.6.2 *Model builder*

The next step in wavefront analysis in QUIP is to pre-build a model if the PCA-based algorithm is to be used. The Model Builder interface is shown in Figure 68. The software assists in setting up and running the aerial image simulations necessary to build the desired models. Response variables can be amplitude and/or phase pupil variation specified by Zernike polynomials. An optional setup file can be used to specify the default values of the interface.

After running the appropriate simulations, the eigenvolumes obtained via PCA can be viewed directly in the software. The eigenvolume shows the aerial image variation caused by a certain aberration. A polynomial model is constructed between the simulated treatment combinations and projections of intensity volumes onto the selected eigenvolumes. The appropriate order of the polynomial can be determined via Akaike information criterion. This is an information theoretic approach to quantifying overfitting risk, and is built-in to the software. The RMSE of the models is also computed to quantify the robustness of the models.

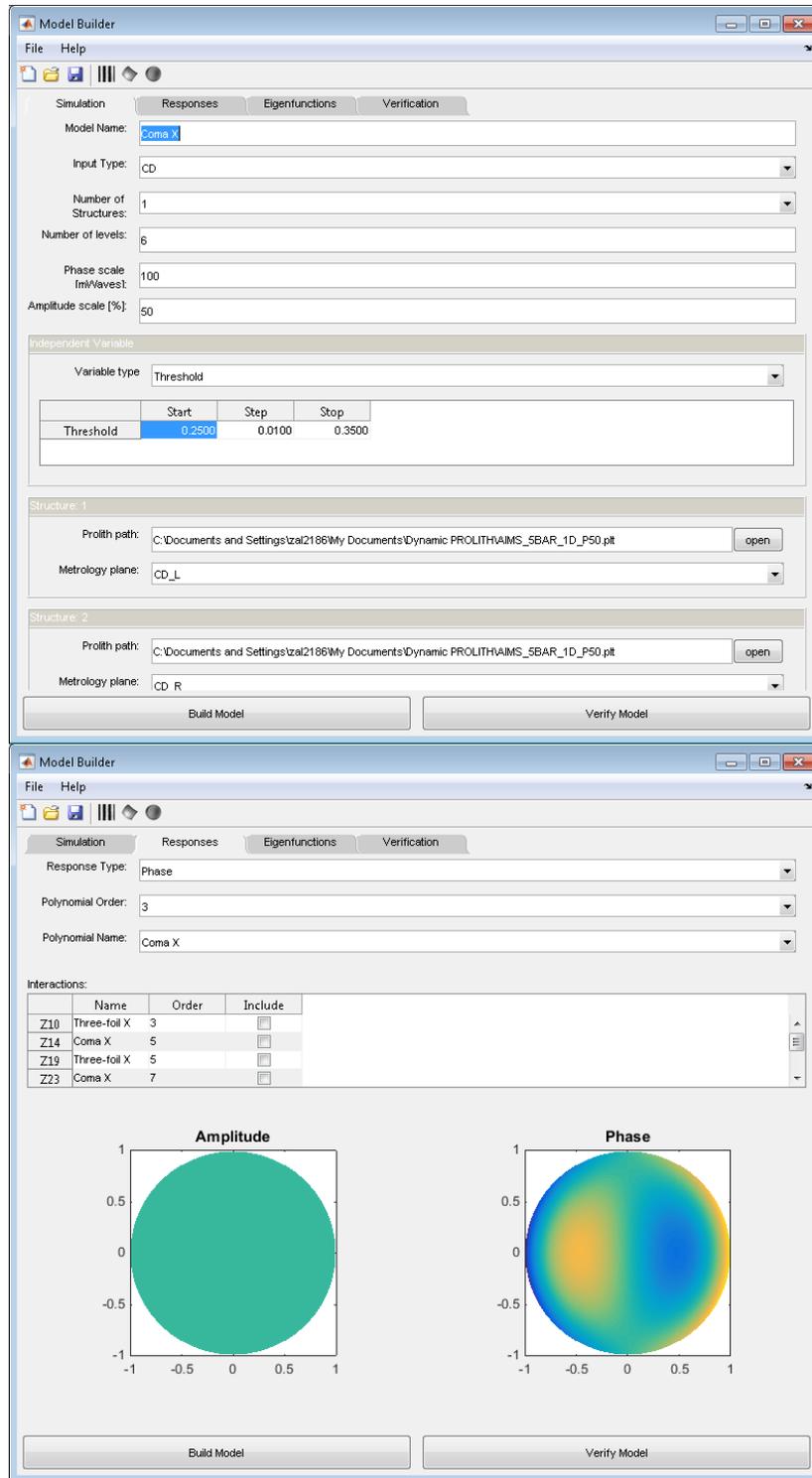


Figure 68. The Model Builder interface showing the setup for a sample model to interrogate coma pupil variation. The top screen shows the simulation setup while the bottom screen shows choosing the response variable and the option to add interacting terms.

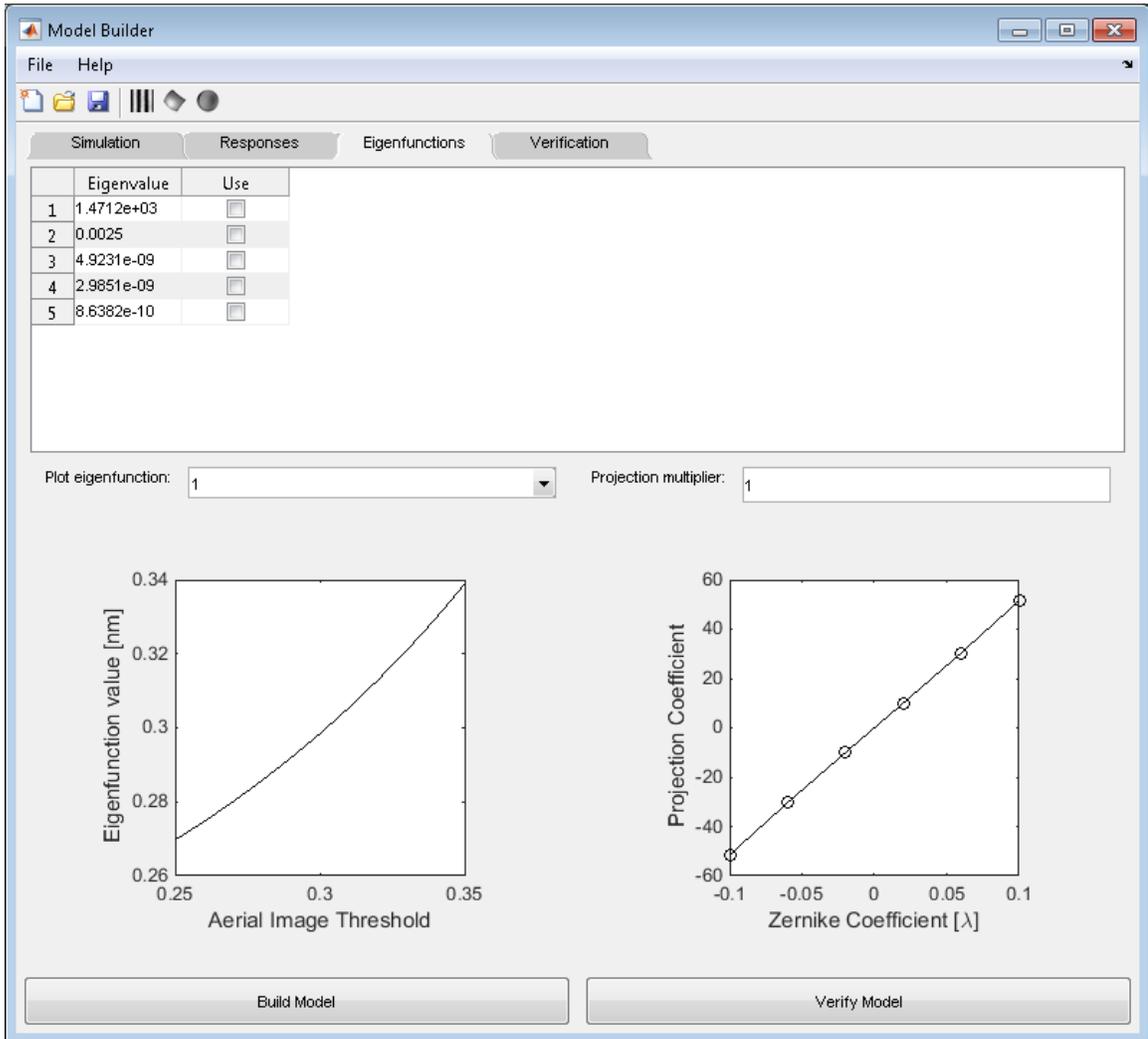


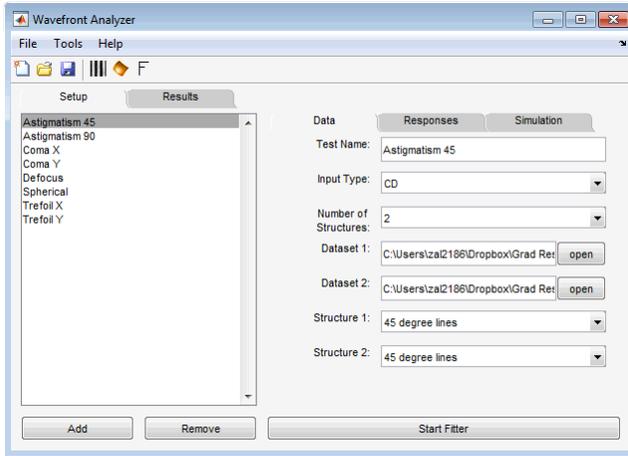
Figure 69. The Model Builder interface showing the result of building a model for coma.

5.6.3 Wavefront analyzer

The final step in wavefront analysis in QUIP is to use the wavefront analyzer. The interface of this component is shown in Figure 70. Like the Model Builder, an optional setup file can be used to specify default values. If a model has been pre-built in the Model Builder, then this can be loaded into the software and an inverse pupil solution can be obtained rapidly by loading the corresponding experimental data. If the iterative

algorithm is to be used, then the appropriate simulations must be set up at this point. Regardless of algorithm, the results of the analysis are output both numerically and graphically. In addition, the results may be saved to a CSV file.

(a)



(b)

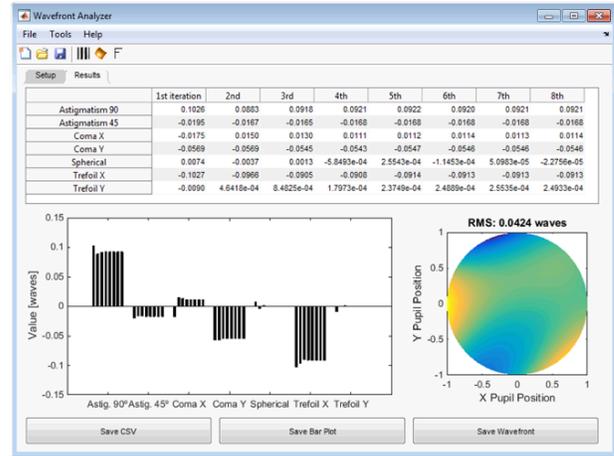


Figure 70. (a) The Wavefront Analyzer interface showing a sample setup to interrogate the primary phase aberrations. (b) The Wavefront Analyzer showing the results running the setup shown in panel (a) using the iterative algorithm.

5.6.4 Program Setup

No matter which algorithm is used, image data must be obtained first. This can either be micrographs from a CD-SEM or focus stacks via SHARP or AIMS EUV. If the wavefront is to be fit to micrographs then CDs need to be measured. This can be done either in the offline SEM software or in the QUIP Image Viewer. In either case care should be taken to ensure that the pattern is being recognized and measured correctly. Mismeasurements do not always stand out as outliers and contribute to error in the extracted wavefront.

After image data has been prepared the data must be formatted into an appropriate CSV file that the QUIP Wavefront Analyzer can read. The image data for all of the used structures can be placed into a single CSV file, or into their own files. Let M be the number of CDs measured for a given target and let N be the number of test cases for each CD. The data for any given structure will be given by a $(M+1) \times (N+2)$ array of cells. The first cell of the first row should contain the name of the structure and the remaining cells in that row should be empty. The second row should contain the headings for the independent variable and the CDs. These cells act as labels for the data and appropriate names should be chosen to aid in setting up the software. The remaining cells should contain the measured data. This format is shown graphically in Figure 71. Additional data can be added in the same file in adjacent columns.

	A	B	C
1	AST30_ORTHO		
2	Focus	CD-V	CD-H

Figure 71. An example of the CSV header structure for data files

The Wavefront Analyzer provides a very general framework for fitting aberrations to image data. The software can be setup to fit any response (aberration) to any image data, but only certain combinations will give an accurate measurement. Therefore, care must be taken to setup the software correctly. After initial setup a plain text file can be saved to recall settings at a later date. A summary of the recommended settings is given in Table 5. Finally, before fitting can begin the appropriate PROLITH files must be provided. These files can be created manually or can be generated automatically by the QUIP software. (Chose Tools > PROLITH File Generator) Data measured through dose must be calibrated to an aerial image threshold, which can also be

done automatically by the QUIP software. (Choose Tools > Dose Calibration) After fitting the user will be presented with the final result of the wavefront extraction, as well as tabulated results and a bar plot of the wavefront after each iteration.

Table 5. Recommended settings for wavefront extraction via QUIP

Structure Name	Independent Variable	Number of Structures	CD1	CD2	CD3	Response
Vertical/Horizontal Lines	Focus	2	Vertical	Horizontal		Astigmatism 90°
45°/135° Lines	Focus	2	45°	135°		Astigmatism 45°
5-Bar Vertical	Dose	1	Left	Right		Coma X
5-Bar Horizontal	Dose	1	Top	Bottom		Coma Y
Line Through Pitch	Focus	3	1:1	1:2	1:4	Spherical
T-Bar Horizontal	Dose	1	Left	Right		Three-foil X
T-Bar Vertical	Dose	1	Top	Bottom		Three-foil Y

6. EXPERIMENTAL RESULTS

6.1 SHARP Pupil Extraction

The SHARP High-NA Actinic Reticle Review Project (SHARP) is an EUV lithography mask microscope at Lawrence Berkeley National Laboratory. A schematic view of the system is shown in Figure 72. SHARP forms a magnified image of the mask at 13.5 nm wavelength onto a CCD sensor via interchangeable Fresnel zone plate lenses. Because the lenses are interchangeable a wide range of magnified numerical apertures are available, from 0.0625 to 0.156 NA (equivalent to 0.25 to 0.625 NA on a 4x lithography system). Synchrotron radiation is focused onto the MEMS mirror in Figure 72 and is subsequently imaged onto the photomask. The MEMS mirror scans the pencil-like synchrotron beam across the aperture, synthesizing arbitrary partially coherent source shapes.⁷⁰

The target dimensions of the structures used for this experiment are given in Table 6. These values were determined by calculating the pitches required to sample the 50% and 90% pupil zones. All data was collected using a 0.0625 NA lens with a conventional 0.1σ source. This source was chosen in order to resemble a fully coherent source and yet introduce a small amount of pupil averaging. Each target was imaged through its full depth of focus with the smallest step size available— $\pm 2.7 \mu\text{m}$ in $0.3 \mu\text{m}$ steps.

Images were subsequently analyzed in the QUIP Image Analyzer.⁷¹ First the images were registered to eliminate stage errors. Next, a small region of interest was inspected using interpolation. The images were collected at approximately 15 nm/px and were interpolated to 5 nm/px. The image intensity was approximated by finding the column-

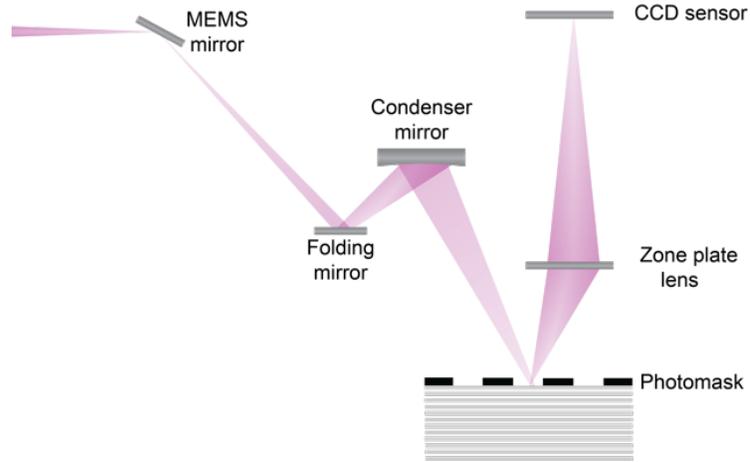


Figure 72. A schematic view of the SHARP microscope at the Advanced Light Source. EUV light at a wavelength of 13.5 nm from a bending magnet is focused on the photomask. The photomask is imaged on a CCD sensor using Fresnel zone plate lens.

wise median of the interpolated region. This is subsequently normalized such that the intensity volume peaks at the ideal value at best focus.

A 10^2 full factorial experiment was run for each Zernike polynomial. The phase design unit was $100 \text{ m}\lambda$ and the amplitude design unit was 50%. The aberration separable basis was computed via custom code written in MATLAB. Vector aerial image simulations were then computed in PROLITH⁷², a commercial lithography simulation engine. Models were created between treatment combinations and projections of intensity volumes onto the aberration basis. The polynomial order was chosen via Akaike information criterion (AIC), an information theoretic approach to quantifying over-fitting risk.⁷³ 100 points normally distributed through the parameter space were used for model verification via AIC. Finally, the pupil variation was extracted using these models.

The required polynomial order (along with its associated AIC weight) is presented in Table 7 with the RMSE for both phase and amplitude models. The RMSE corresponds

Table 6. CDs of the metrology targets used for pupil function extraction on SHARP.

Aberration Name	Structure Type	Target CD [nm]
Astigmatism 90° (Z_5)	Vertical/Horizontal Lines	120
Astigmatism 45° (Z_6)	45°/135° Lines	120
Coma X (Z_7)	Vertical 5-bar	200
Coma Y (Z_8)	Horizontal 5-bar	200
Spherical (Z_9)	Line through pitch	120
Trefoil X (Z_{10})	Horizontal T-Bar	140
Trefoil Y (Z_{11})	Vertical T-Bar	140

to the error of the 100 normally distributed points and is on the order of 0.2 mλ and 0.2% for phase and amplitude, respectively. This amount of error is comparable to the error of the iterative algorithm from previous studies. All models were built on a machine with 16 GB RAM using a 3.6 GHz quad-core Intel i7 processor. The QUIP eigenfunctions were computed directly from the TCC in 32.5 seconds and total runtime to build all models was 20 hours. The bulk of this time is spent on full vector image simulations.

A comparison of the pupil variation extracted with both the iterative and QUIP algorithms is given in Table 3. The iterative algorithm took 3.2 hours to converge on a solution on the same machine described above, while the QUIP algorithm took 0.2 seconds to obtain an inverse solution. The amplitude and phase pupil variation are plotted in Figure 73, as extracted by the QUIP algorithm. The difference between the two algorithms is insignificant and is likely caused by numerical complications when fitting to aerial image data. When CD data is used there is no difference between the two algorithms.

Table 7. Required polynomial orders, and their associated RMSE for phase and amplitude variation, for each Zernike polynomial.

Name	Polynomial Order	Phase RMSE [mλ]	Amplitude RMSE [%]
Z ₅	6 (0.632 AIC weight)	0.18	0.35
Z ₆	5 (0.500 AIC weight)	0.19	0.25
Z ₇ /Z ₈	4 (0.496 AIC weight)	0.10	0.20
Z ₉	6 (0.995 AIC weight)	0.10	0.15
Z ₁₀ /Z ₁₁	6 (0.742 AIC weight)	0.12	0.18

Table 8. A comparison of the extracted phase and amplitude variation using both the iterative algorithm and QUIP algorithm.

Name	Iterative Phase [mλ]	Iterative Amplitude [%]	QUIP Phase [mλ]	QUIP Amplitude [%]
Z ₅	-1.0	+14.9	-3.5	+5.1
Z ₆	-3.5	+0.3	-2.4	+0.4
Z ₇	0.0	-10.0	-0.8	-10.1
Z ₈	0.0	-15.4	+0.1	-15.7
Z ₉	+2.7	+2.8	+1.8	+4.9
Z ₁₀	0.0	-10.3	+0.2	-10.2
Z ₁₁	-23.2	+3.9	-23.7	+4.3

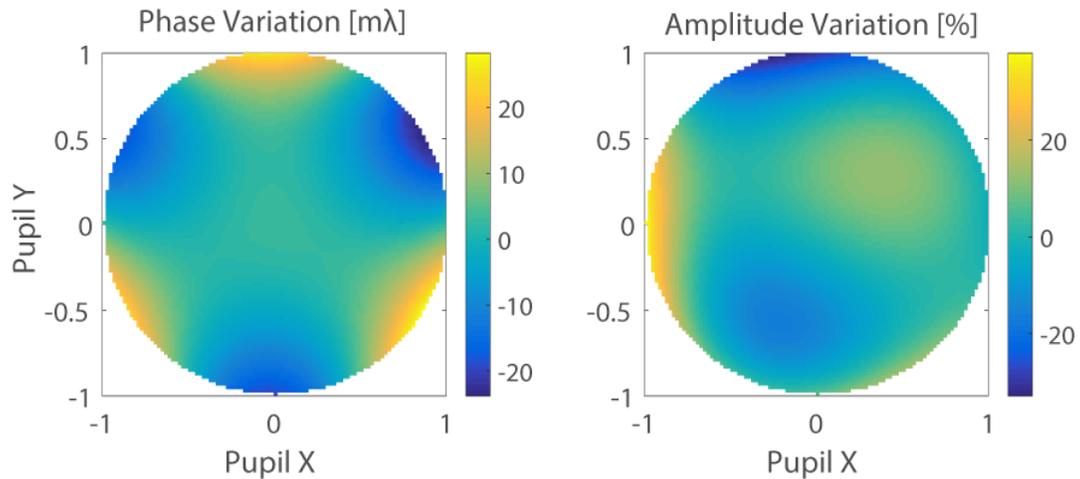


Figure 73. Pupil amplitude and phase extracted from SHARP with only third-order terms. (Z_5-Z_{11})

6.2 AIMS EUV™ Pupil Extraction

AIMS, like the SHARP microscope, is a system built for actinic reticle inspection. However, unlike SHARP, AIMS is designed to use optics more similar to those used in production scanners. Instead of a zone plate lens, AIMS uses reflective optics to form an image on a CCD sensor with 750x magnification. An aspheric mirror is used for the primary mirror, while the remaining three mirrors are spherical, as in Figure 74.⁷⁴

6.2.1 Third-Order Only

A preliminary wavefront was extracted from the AIMS EUV exposure system using the QUIP algorithm. Examples of these structures compared to images as formed on the AIMS EUV tool are given in Figure 84. The target CDs, at the mask level, of the structures used for this experiment are given in Table 13. These values were chosen to match previous experiments on NXE:3300 systems. In those experiments the target CDs were determined using an optimization procedure that balances aberration sensitivity and target printability. A total of 38 image sequences were captured from the AIMS EUV



Figure 74. A schematic view of the projection optics in the AIMS EUV system. ⁷⁴

system through $\pm 1 \mu\text{m}$ of defocus with a defocus step of 66.7 nm. All exposures were completed using an annular source with $0.702\sigma_i - 0.894\sigma_o$ at 0.0825 NA (0.33 4xNA). This report concerns only the 10 sequences for the features described in Table 13.

The red lines in Figure 84 denote the measurement locations used for wavefront extraction. CDs were measured using custom offline metrology software developed in MATLAB. Images were provided by ZEISS as MATLAB arrays with floating point values ranging from zero to unity. In this normalization unity corresponds to the intensity of a large open area and zero corresponds to no light. Each focal plane in the array was written to a 16-bit PNG image using the same normalization. This converts the values to integers ranging from zero to 65535. This process can then be undone in the metrology software in order to restore the original tool normalization. This software was calibrated to the AIMS EUV system by measuring the CD in pixels of both the lines and spaces of three arrays with different pitches. This information was used to compute the spatial resolution of the images as 4.420 nm/px in mask level coordinates.

This resolution was used for to obtain CD data in nanometers. CDs are measured from the images by averaging a region either column-wise or row-wise depending on the feature orientation. The images are normalized to the ideal aerial image extrema, as obtained via a PROLITH simulation. The width of the averaging regions was 500 nm with the length varying depending on the target type. Where it was possible to do so the data was averaged over several identical features.

The measurements obtained from the images are plotted in Figure 76–81. In the case of coma and trefoil, not all of the data was used for aberration extraction. The gray region in the plot denotes the regions that were used to extract the wavefront. The estimated Zernike coefficients for a third-order wavefront is given in Table 10 and the wavefront is plotted in Figure 82. Due to a confidentiality agreement the extracted values are normalized to the highest term. Both the purely iterative and QUIP algorithms were used with the CD data extracted from the AIMS images. The models converged after eight iterations with no difference between the two algorithms.

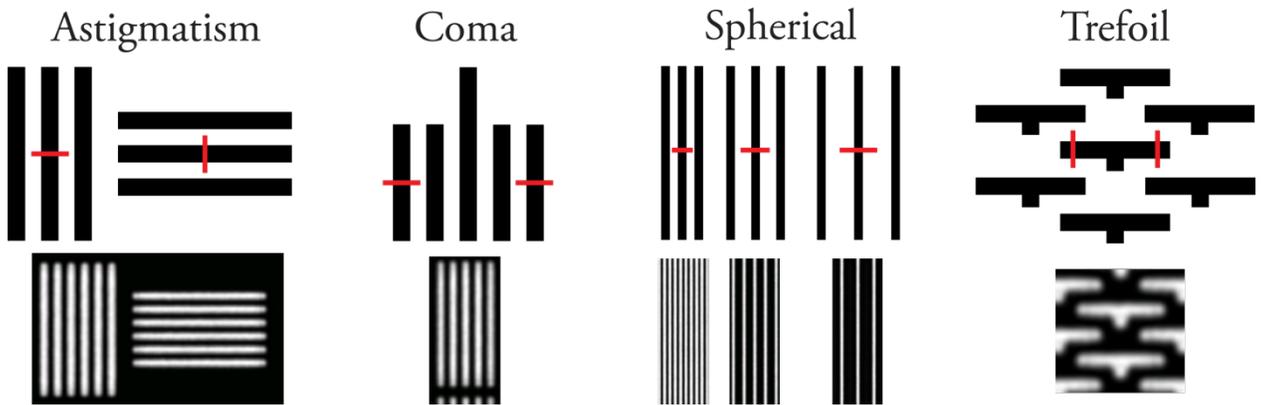


Figure 75. Comparison of the targets used for aberration interrogation and their images as formed on the AIMS EUV tool.

Table 9. CDs of the metrology targets used for pupil function extraction on AIMS EUV.

Aberration Name	Structure Type	Target CD [nm]
Astigmatism 90° (Z_5)	Vertical/Horizontal Lines	120
Astigmatism 45° (Z_6)	45°/135° Lines	120
Coma X (Z_7)	Vertical 5-bar	100
Coma Y (Z_8)	Horizontal 5-bar	100
Spherical (Z_9)	Line through pitch	104
Trefoil X (Z_{10})	Horizontal T-Bar	120
Trefoil Y (Z_{11})	Vertical T-Bar	120

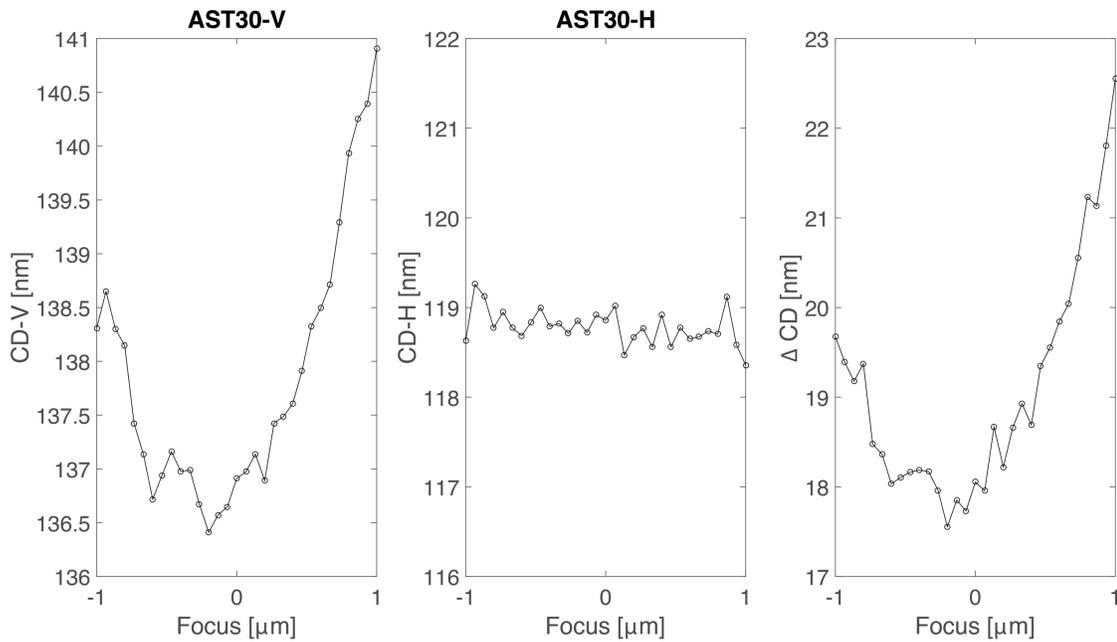


Figure 76. CD data for vertical spaces (left), horizontal spaces (middle), and their difference (right). These are the targets used to interrogate astigmatism 90° (Z_5).

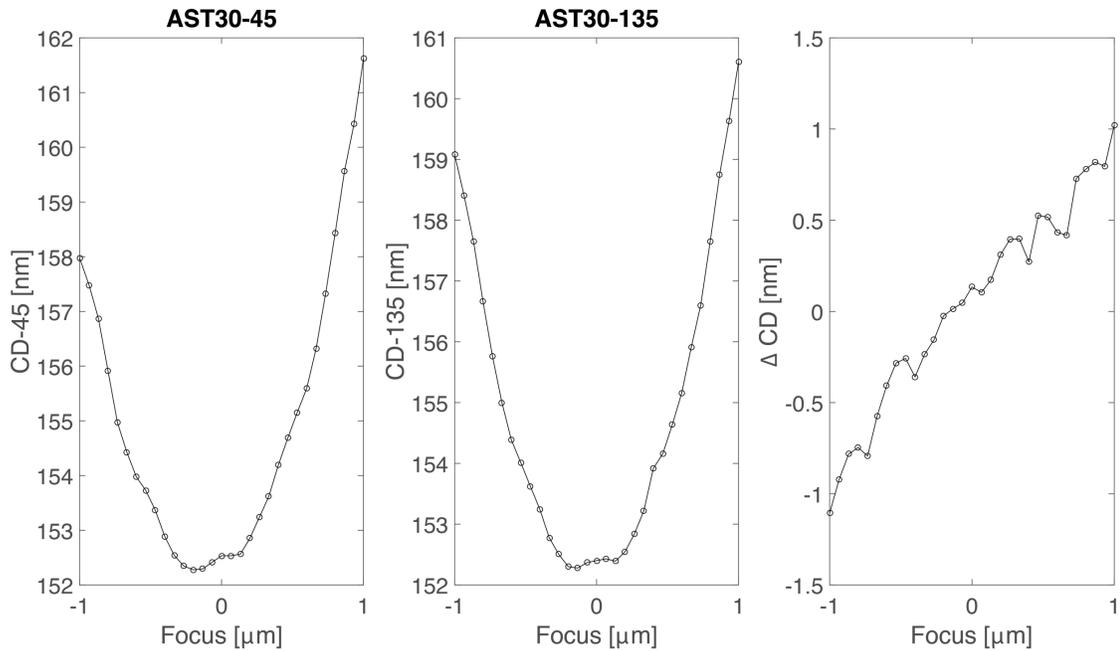


Figure 77. CD data for 45° spaces (left), 135° spaces (middle), and their difference (right). These are the targets used to interrogate astigmatism 45° (Z6).

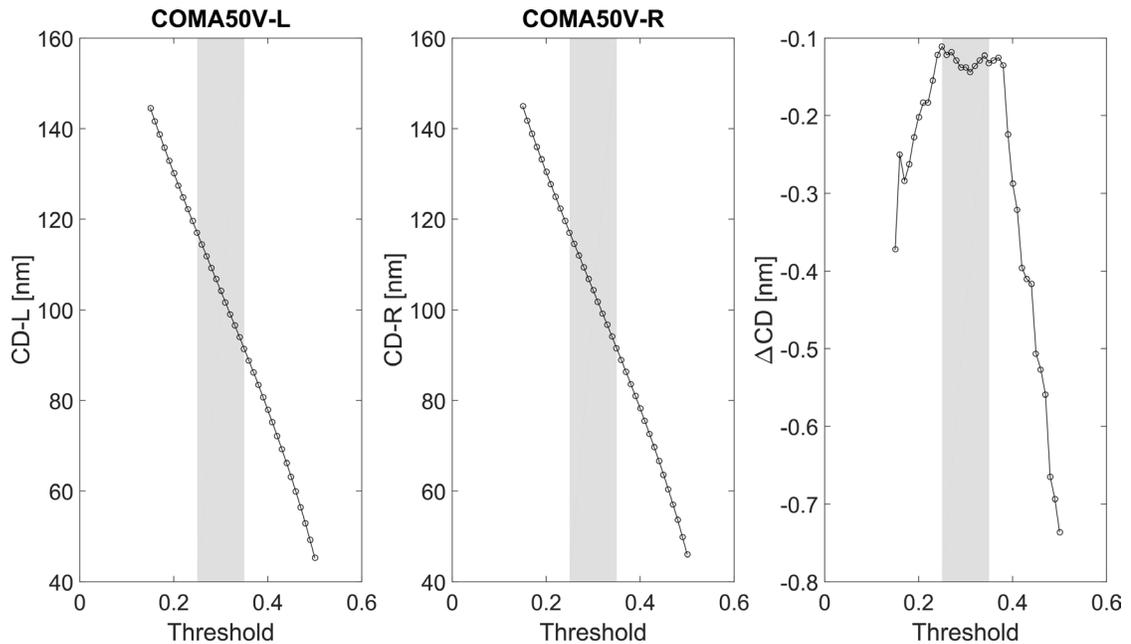


Figure 78. CD data for the left and right bars of a vertical five-bar structure (left and middle plots, respectively), and their difference (right). This is the target used to interrogate coma X (Z7). The grey region denotes the region used for aberration extraction.

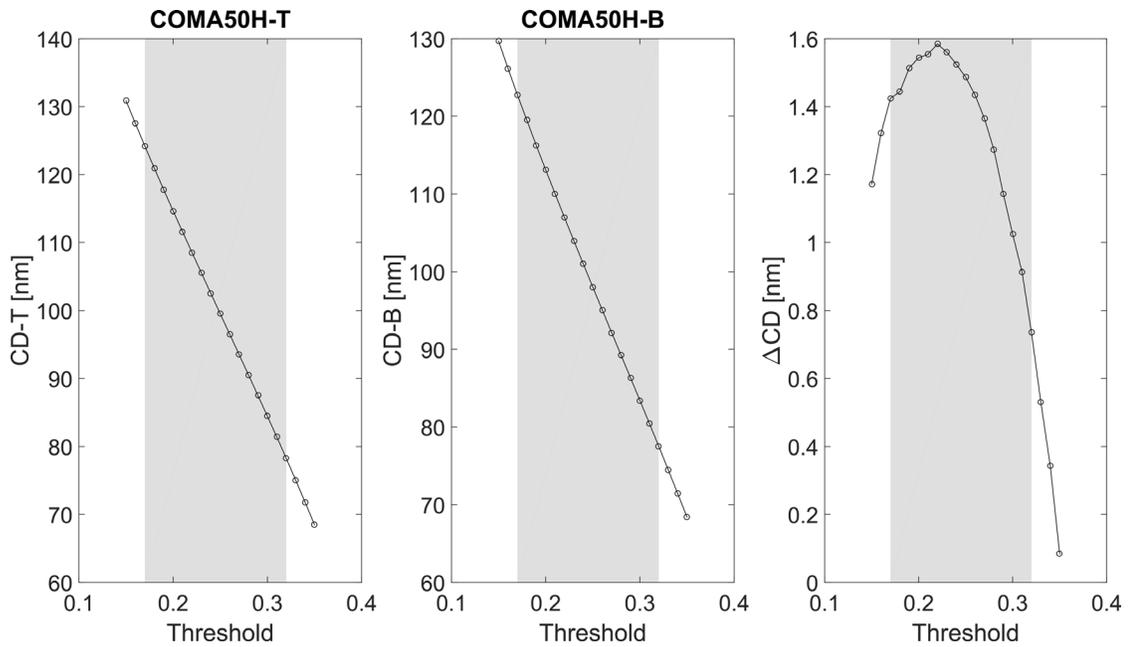


Figure 79. CD data for the top and bottom bars of a horizontal five-bar structure (left and middle plots, respectively), and their difference (right). This is the target used to interrogate coma Y (Z8). The grey region denotes the region used for aberration extraction.

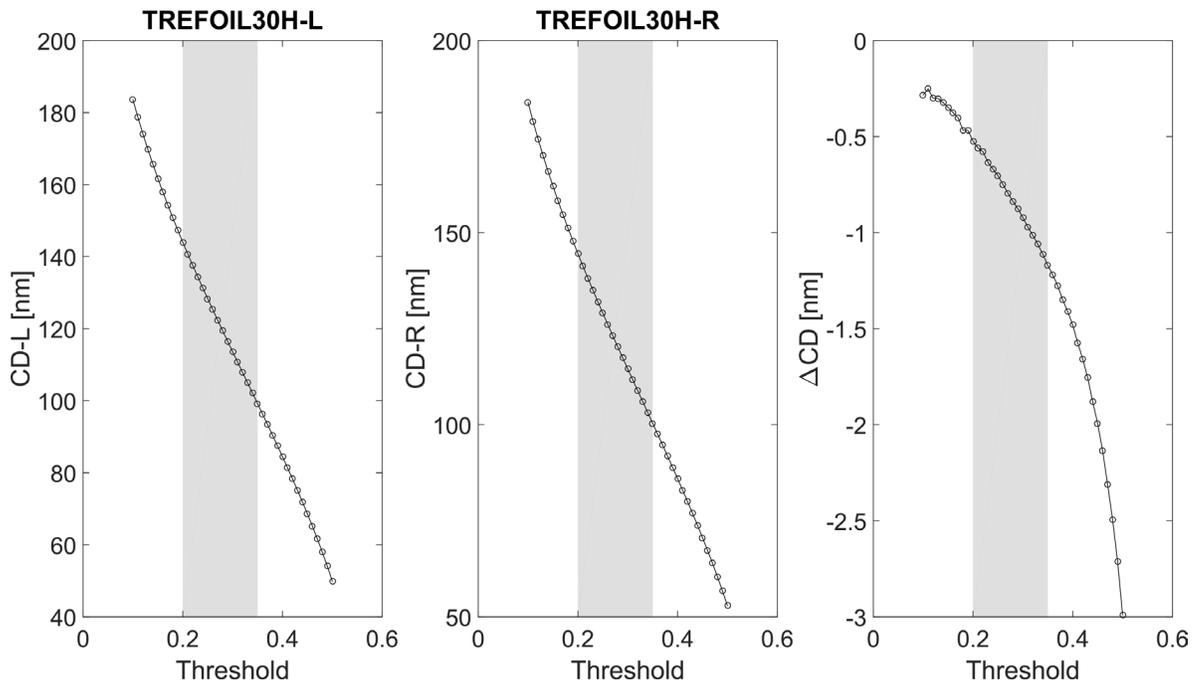


Figure 80. CD data for the left and right ends of a horizontal T-bar structure (left and middle plots, respectively), and their difference (right). This is the target used to interrogate trefoil X (Z10). The grey region denotes the region used for aberration extraction.

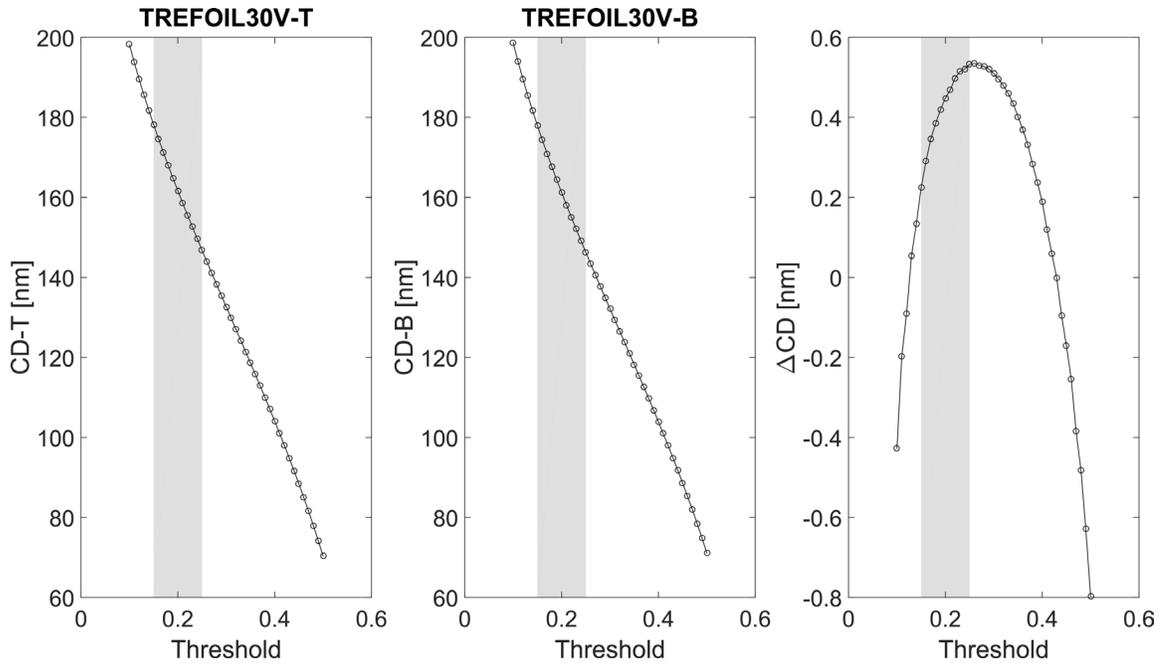


Figure 81. CD data for the top and bottom ends of a vertical T-bar structure (left and middle plots, respectively), and their difference (right). This is the target used to interrogate trefoil Y (Z11). The grey region denotes the region used for aberration extraction.

Table 10. Extracted Zernike aberration values for each aberration type for the AIMS EUV.

Zernike Term	Aberration Name	Direction	Iterative Phase [A.U]	QUIP Phase [A.U]
Z ₅	Astigmatism	90°	0.859	0.859
Z ₆		45°	0.658	0.658
Z ₇	Coma	X	-0.268	-0.268
Z ₈		Y	-0.745	-0.745
Z ₉	Spherical		-0.215	-0.215
Z ₁₀	Trefoil	X	1.000	1.000
Z ₁₁		Y	-0.255	-0.255

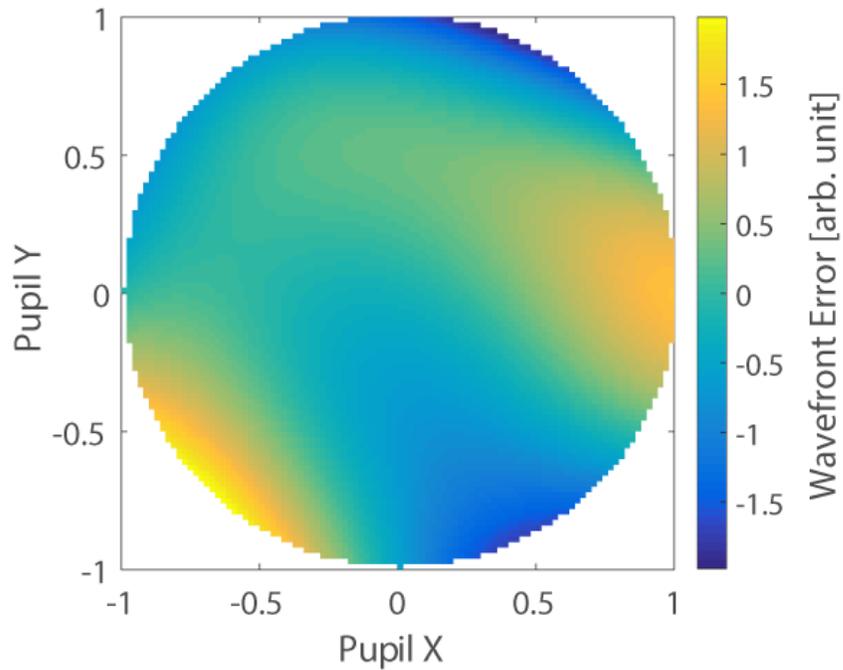


Figure 82. Extracted wavefront from AIMS EUV plotted with all third-order terms. (Z_5 – Z_{11})

6.2.2 *Third and Fifth Order*

Third- and fifth- order terms were measured from the AIMS EUV using the additional pitches given in Table 11. The procedure for data analysis was identical to that of the previous section. This time, only the QUIP algorithm was used. The extracted pupil phase variation is plotted in Figure 83 and is tabulated in

Table 12. The results here are normalized to the same value as in in order to facilitate comparison between the two techniques. Comparing Figure 83 to Figure 82, the wavefront obtained by fitting only low-order terms is very similar to the wavefront obtained using third- and fifth-order terms. This is expected because the extracted third-order terms contain the effects of all high-order terms lumped into the single low-order term. It should be noted that four-foil was not measured in this analysis because it would require a third pitch of the targets used to interrogate astigmatism.

Table 11. Additional pitches used to measure third- and fifth-order aberrations from AIMS EUV.

Aberration Name	Structure Type	Target CD [nm]
Astigmatism 90° (Z_5)	Vertical/Horizontal Lines	160
Astigmatism 45° (Z_6)	45°/135° Lines	160
Coma X (Z_7)	Vertical 5-bar	140
Coma Y (Z_8)	Horizontal 5-bar	140
Spherical (Z_9)	Line through pitch	80
Trefoil X (Z_{10})	Horizontal T-Bar	100
Trefoil Y (Z_{11})	Vertical T-Bar	100

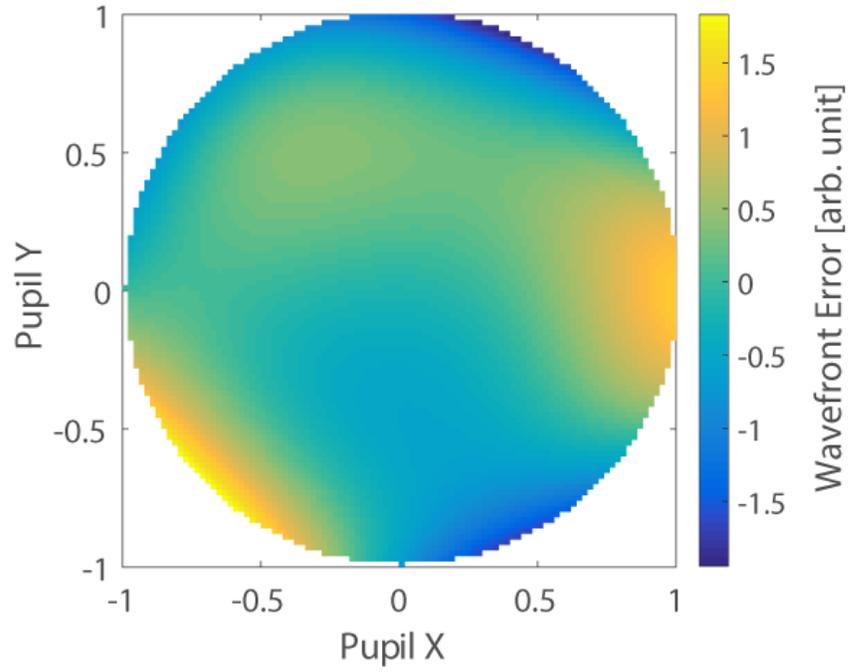


Figure 83. Wavefront extracted from AIMS EUV with all third-order terms and their fifth-order counterparts. (Z_5 – Z_{16} , Z_{19} , Z_{20})

Table 12. Extracted Zernike aberration values for each aberration type for the AIMS EUV.

Aberration Name	Aberration Order	Value [A.U.]
Astigmatism 90°	3	0.752
	5	0.134
Astigmatism 45°	3	0.383
	5	0.228
Coma X	3	-0.161
	5	-0.081
Coma Y	3	-0.691
	5	-0.013
Spherical	3	-0.201
	5	-0.027
Three-Foil X	3	0.879
	5	0.074
Three-Foil Y	3	-0.188
	5	-0.087

6.3 NXE:3300 Pupil Extraction

The NXE:3300 is a state-of-the-art production EUV lithography tool produced by ASML. It uses a full-field catoptric lens with 0.33 NA. An approximate wavefront was extracted from the IMEC NXE:3300 exposure system using the RIT image-based method and the QUIP algorithm. The tool is capable of free-form sources through the use of the FlexPupil system, but all exposures presented here were completed with an annular

source. The sources in modern lithography processes are co-optimized with the mask layout to optimize the contrast, MEEF, process window, and other imaging effects. Another way of looking at free-form sources is that they increase the performance of certain patterns by decreasing the performance of other patterns. For example, a horizontal dipole source will print vertical lines very well and will not print most pitches of horizontal lines at all. In our case, we do not care about the process window of our structures, and are using them as a way to sample the pupil. Because we do not want to bias any particular pupil direction, an azimuthally symmetric source is the best choice for aberration extraction. Partially coherent “conventional” sources and annular sources are both good choices. CD measurements from micrographs of resist images is fit to pupil wavefront error via aerial image simulations. Examples of these structures compared to images formed in resist on the IMEC NXE:3300 system are given in Figure 84.

6.3.1 *First Extraction*

The target CDs of the structures used for this experiment are given in Table 13. These values were determined using an optimization procedure that balances aberration sensitivity and target printability. Six exposures were completed on the IMEC NXE:3300 system, as detailed in Table 14. All exposures were completed using an annular source with $0.702\sigma_i - 0.894\sigma_o$. The first two wafers were used to center the process window for the remaining four wafers used for wavefront extraction. The red lines in Figure 84 denote the measurement locations used for wavefront extraction.

Confidence intervals are shown in Figures 5–10 for all the metrology structures used in this work. In each case the confidence intervals are shown for both CDs as well as

the computed confidence intervals for the CD difference. Through-focus data is known to vary quadratically so a second-order polynomial is used for these structures.

A total of 10 iterations were needed for the wavefront extraction models to converge on a solution. The extracted wavefront is given in Table 15 with margin of error and a plot of the final extracted wavefront is given in Figure 91. The wavefront has a total RMS of 49.00 mλ or 0.66 nm. Only one pitch of each target type was measured in this dataset, so only third-order aberrations can be extracted.

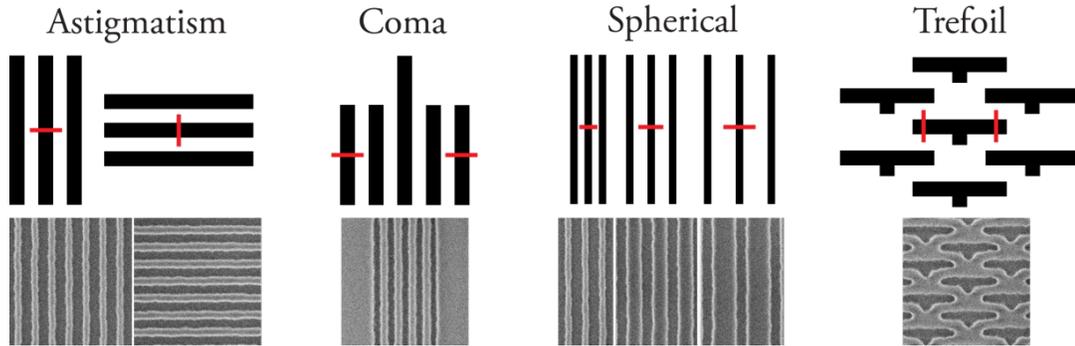


Figure 84. Comparison of the targets used for aberration interrogation and their images in resist via NXE:3300 exposure

Table 13. CDs of the metrology targets used for pupil function extraction on NXE3300.

Aberration Name	Structure Type	Target CD [nm]
Astigmatism 90° (Z_5)	Vertical/Horizontal Lines	30
Astigmatism 45° (Z_6)	45°/135° Lines	30
Coma X (Z_7)	Vertical 5-bar	25
Coma Y (Z_8)	Horizontal 5-bar	25
Spherical (Z_9)	Line through pitch	30
Trefoil X (Z_{10})	Horizontal T-Bar	30
Trefoil Y (Z_{11})	Vertical T-Bar	30

Table 14. Summary of the exposed wafers.

Wafer Number	Type
1	FEM
2	CDU
3	Energy meander
4	Fine focus meander
5	Coarse focus meander
6	FEM

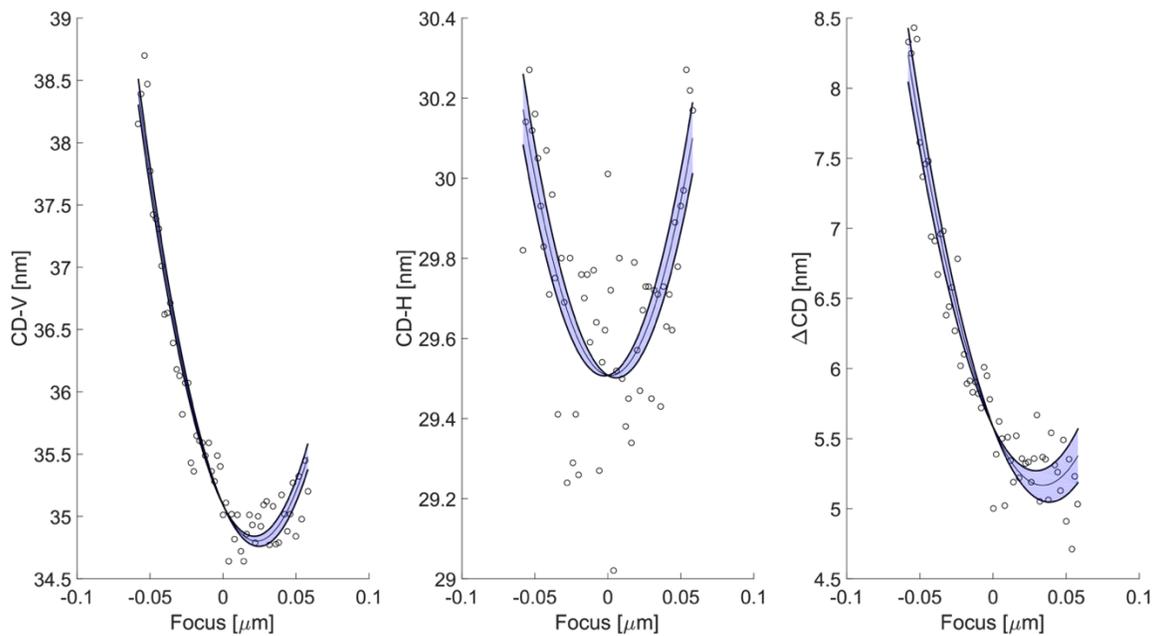


Figure 85. 95% confidence intervals for a second order polynomial fit to CD data for vertical and horizontal lines. This is the target used to interrogate astigmatism 90° (Z_5). The markers are the raw CD-SEM data, while the thin black line is the line of best fit. The thick black lines represent the edges of the confidence interval.

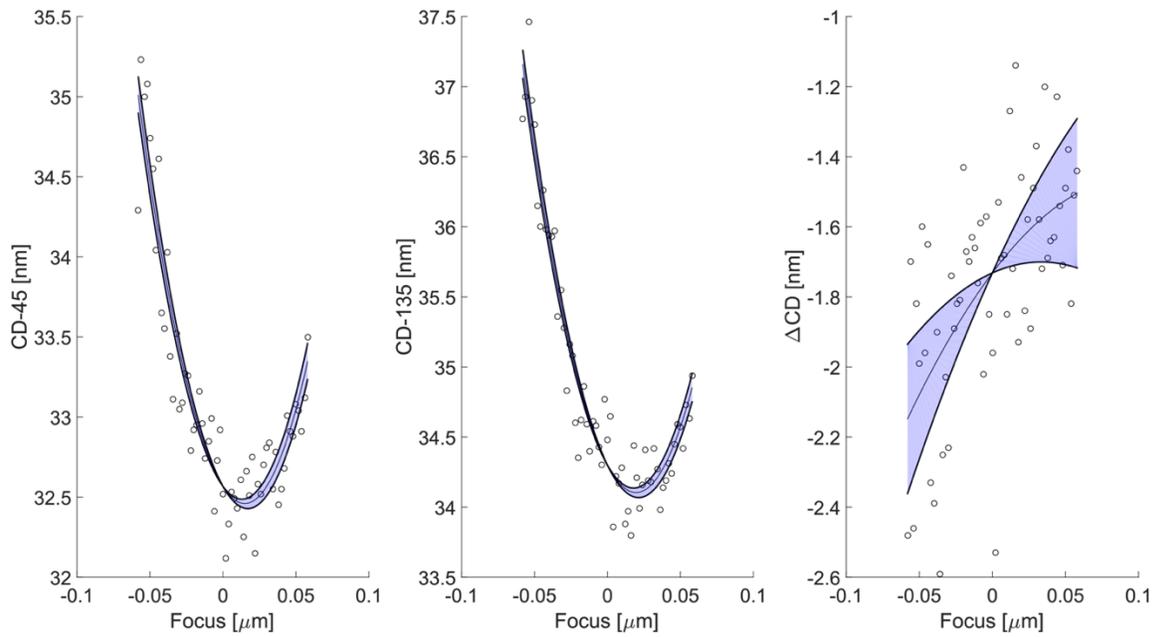


Figure 86. 95% confidence intervals for a second order polynomial fit to CD data for 45° and 135° lines. This is the target used to interrogate astigmatism 45° (Z₆).

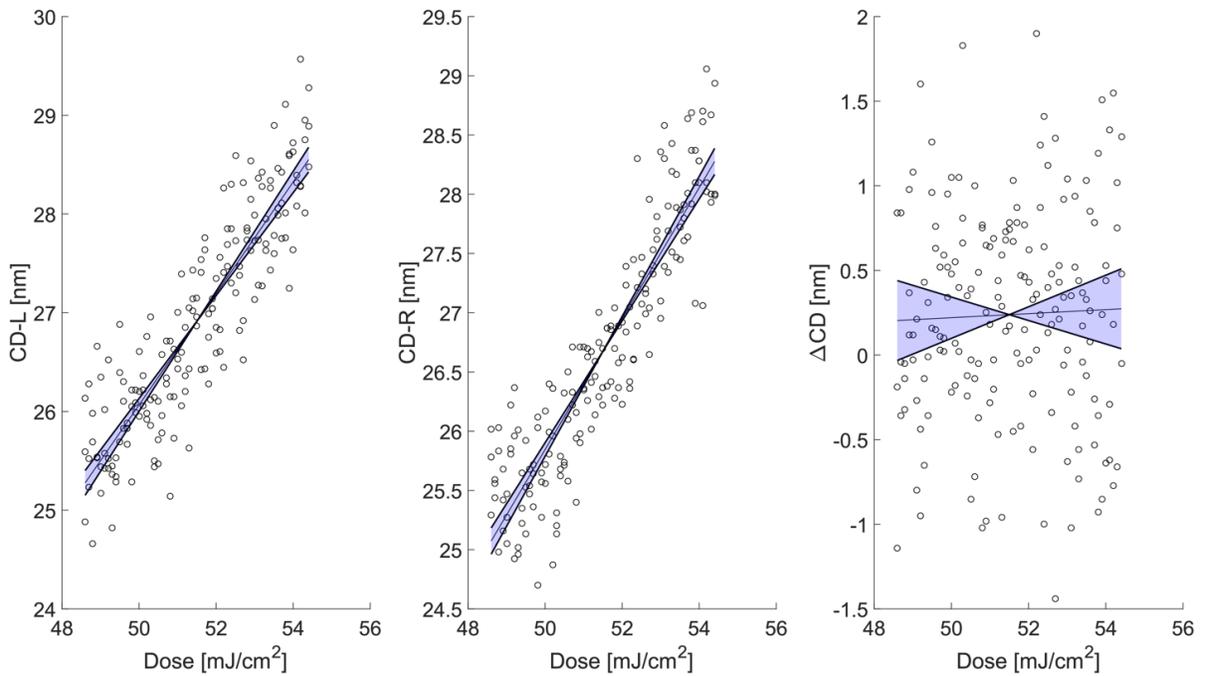


Figure 87. 95% confidence intervals for a linear regression to CD data through dose for a vertical five-bar structure. This is the target used to interrogate coma X (Z₇).

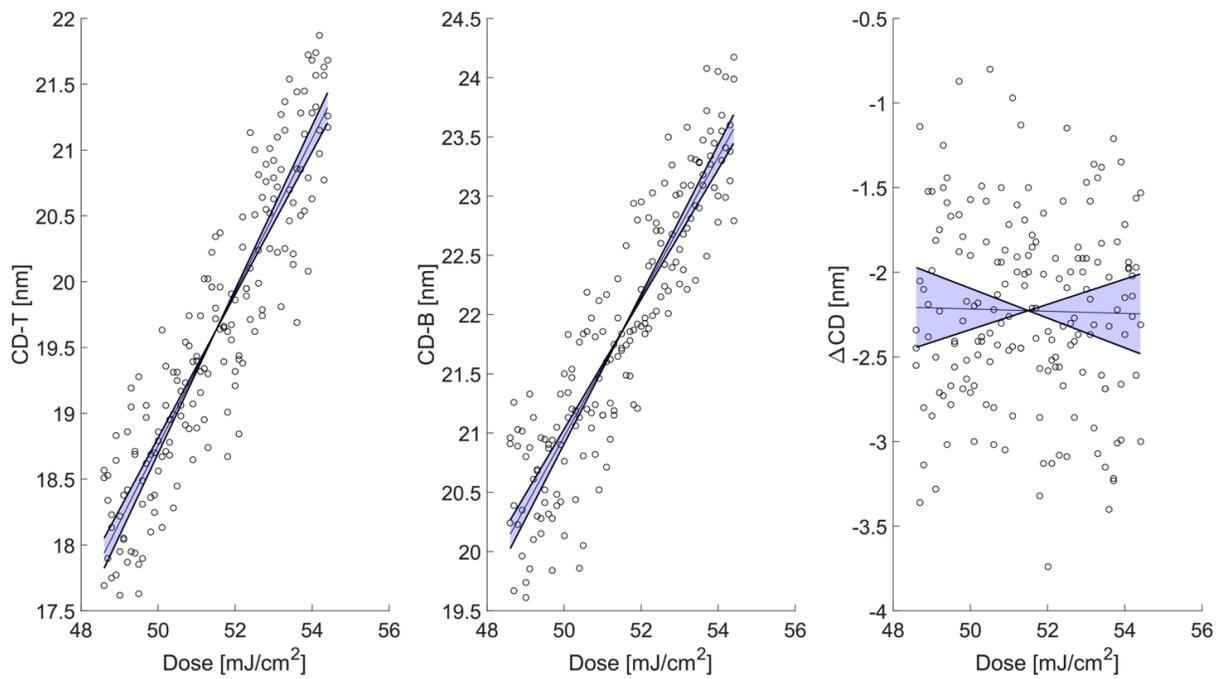


Figure 88. 95% confidence intervals for a linear regression to CD data through dose for a horizontal five-bar structure. This is the target used to interrogate coma Y (Z_8).

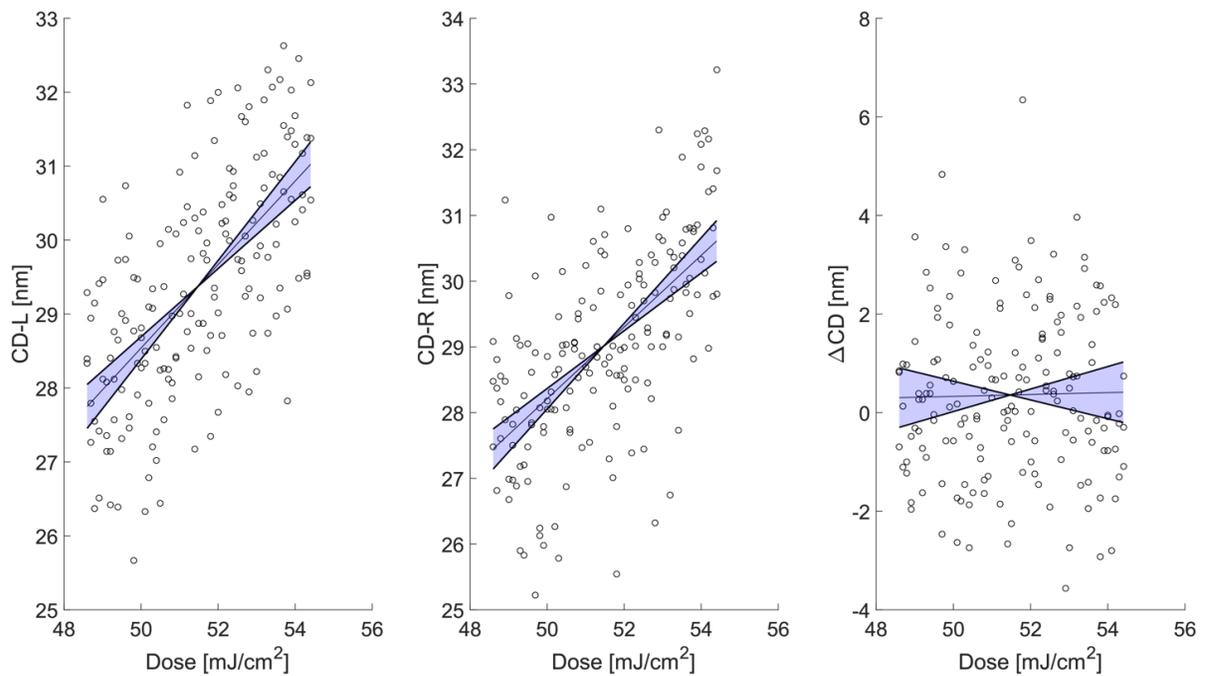


Figure 89. 95% confidence intervals for a linear regression to CD data through dose for horizontal T-bars. This is the target used to interrogate trefoil X (Z_{10}).

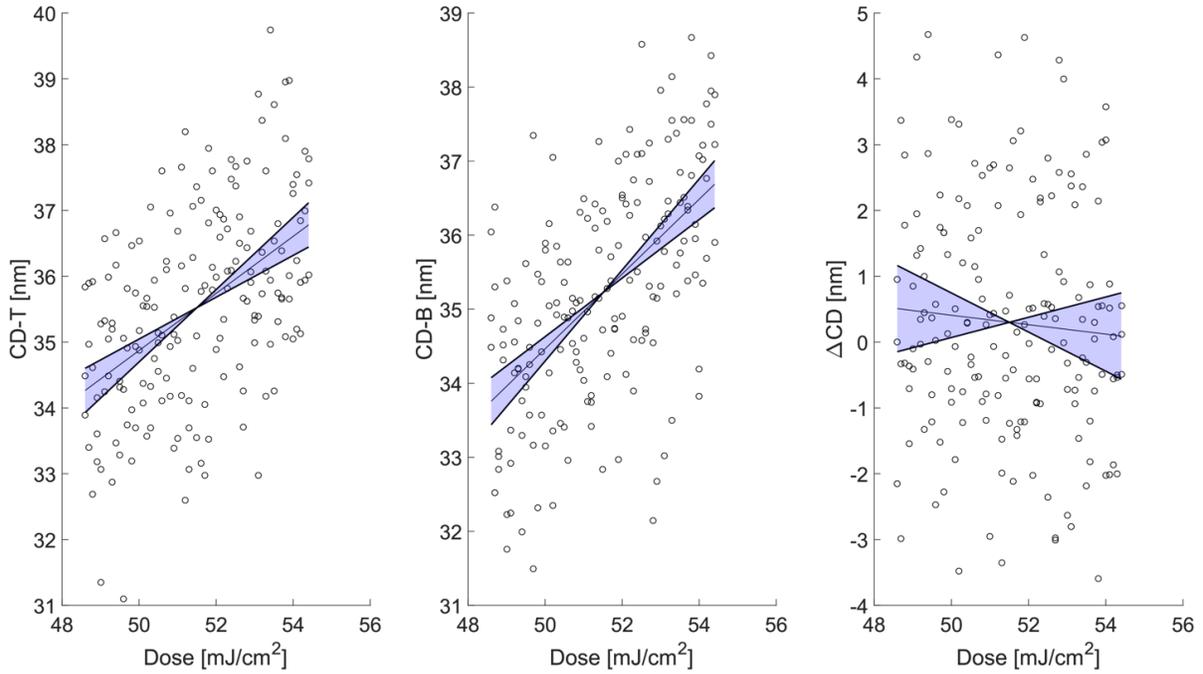


Figure 90. 95% confidence intervals for a linear regression to CD data through dose for vertical T-bars. This is the target used to interrogate trefoil Y (Z_{11}).

Table 15. Extracted Zernike aberration values for each aberration type for the IMEC NXE:3300. The margin of error was obtained by extracting aberrations from the confidence intervals on the CD difference of metrology structures.

Aberration Name	Structure Type	Zernike estimate [$m\lambda$]
Astigmatism 90° (Z_5)	Vertical/Horizontal Lines	86.56 ± 0.55
Astigmatism 45° (Z_6)	$45^\circ/135^\circ$ Lines	27.49 ± 0.70
Coma X (Z_7)	Vertical 5-bar	0.14 ± 0.18
Coma Y (Z_8)	Horizontal 5-bar	86.17 ± 0.19
Spherical (Z_9)	Line through pitch	-10.42 ± 0.53
Trefoil X (Z_{10})	Horizontal T-Bar	18.39 ± 0.20
Trefoil Y (Z_{11})	Vertical T-Bar	95.87 ± 0.60

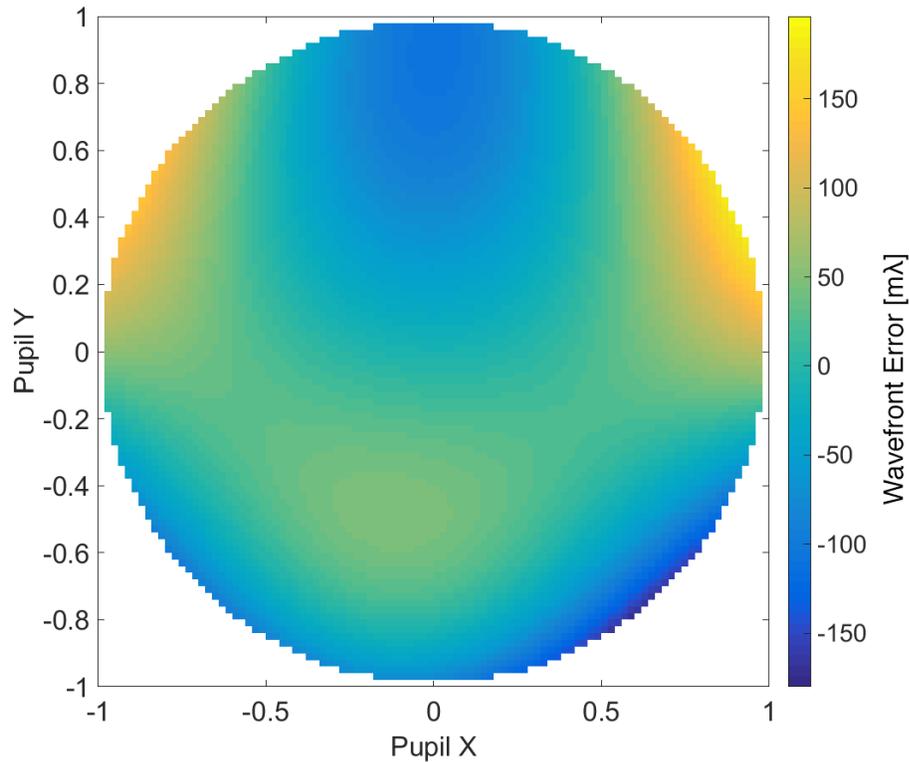


Figure 91. Wavefront extracted from NXE3300 with all third-order (Z_5 – Z_{11})

6.3.2 *Second Extraction*

In the time between exposure and when analysis was completed, the ILIAS sensor was cleaned on the IMEC NXE:3300. This sensor is used for on-board aberration measurements which are used for feed forward controls to attempt to reduce aberrations. In response to this cleaning, a second dataset was exposed. This dataset also contained additional pitches to measure high order terms.

6.3.3 *Third-Order Only*

The procedure to measure third-order aberrations remained unchanged between the first and second runs. However, the dose range on the energy meander wafer was expanded to make it easier to extract the odd-parity aberrations (coma and trefoil) over a wider change in aerial image

threshold. The wavefront extraction converged in 15 iterations, and the results are given in Table 16. The RMS of the wavefront is $32.27 \text{ m}\lambda$ (0.44 nm). Comparing these results to Table 15, there is a clear resemblance between the two wavefronts. This is perhaps clearer when considering Figure 92, where the wavefront from this run was extracted. Coma and three-foil Y have remained the largest terms. The X-direction counterparts of these terms have remained small. The total amount of astigmatism has been reduced. This is a good indication of the repeatability of this method and shows how this technique could be used to track system variation over time.

Table 16. Extracted Zernike aberration values for each aberration type for the IMEC NXE:3300.

Aberration Name	Value [$\text{m}\lambda$]
Astigmatism 90°	-20.3
Astigmatism 45°	+22.5
Coma X	-05.3
Coma Y	-59.9
Spherical	+08.9
Three-Foil X	-03.3
Three-Foil Y	-57.7

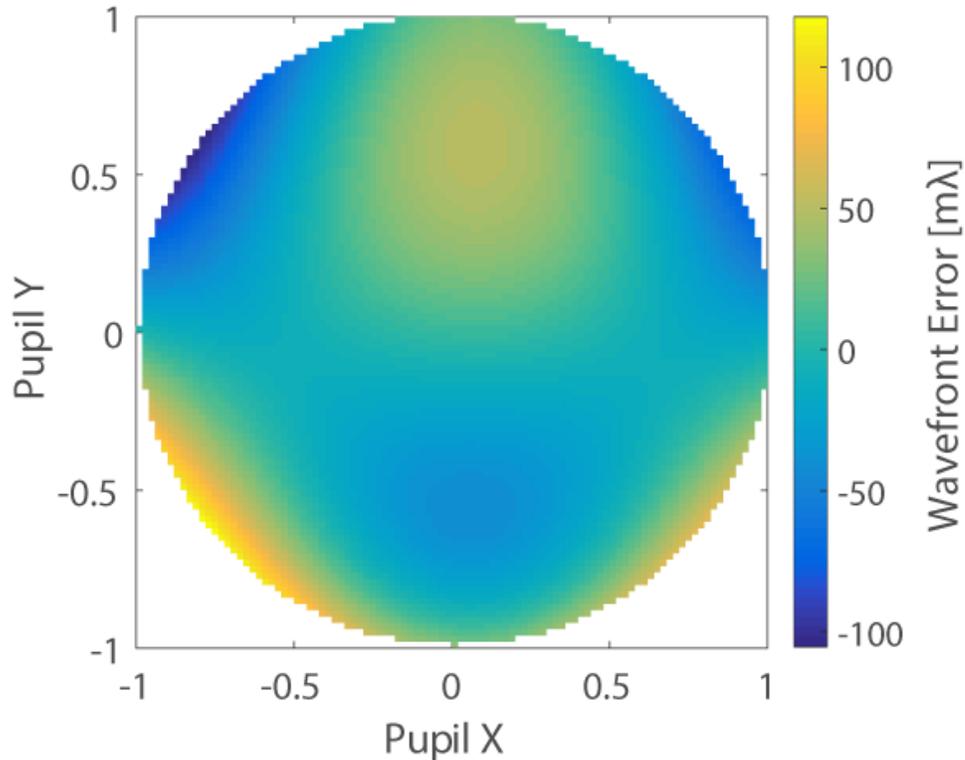


Figure 92. Wavefront extracted from IMEX NXE:3300 with all third-order terms. (Z_5 – Z_{11})

6.3.4 *Third and Fifth Order*

Third- and fifth-order terms were extracted from CD data from images formed by the IMEC NXE:3300 system in resist. The same structures used to extract third order terms were used, in addition to the ones listed in Table 17. The analysis is identical to the previously listed experiments. The QUIP models converged in 25 iterations, and the results are given in Table 18 and plotted in Figure 93. The extracted wavefront had an RMS of $25.30\ m\lambda$ (0.34 nm). Again, this wavefront bears resemblance to the purely third-order wavefront. This is because the third-order wavefront represents a lumped contribution of high-order terms into a single low-order term. When third- and fifth- order terms are extracted the high-order contribution is lumped into the fifth-order terms. Again, four-foil was not extracted in this analysis due to the requirement of a third pitch of the astigmatism target.

Table 17. Additional pitches used to measure third- and fifth-order aberrations from the IMEC NXE:3300 exposure system.

Aberration Name	Structure Type	Target CD [nm]
Astigmatism 90° (Z_5)	Vertical/Horizontal Lines	40
Astigmatism 45° (Z_6)	45°/135° Lines	40
Coma X (Z_7)	Vertical 5-bar	35
Coma Y (Z_8)	Horizontal 5-bar	35
Spherical (Z_9)	Line through pitch	40
Trefoil X (Z_{10})	Horizontal T-Bar	30
Trefoil Y (Z_{11})	Vertical T-Bar	30

Table 18. Extracted Zernike aberration values for each aberration type for the IMEC NXE:3300.

Aberration Name	Aberration Order	Value [$m\lambda$]
Astigmatism 90°	3	-16.7
	5	-5.7
Astigmatism 45°	3	+17.4
	5	+2.3
Coma X	3	-3.6
	5	-5.6
Coma Y	3	-40.8
	5	-9.3
Spherical	3	+05.5
	5	0.0
Three-Foil X	3	-6.4
	5	-1.2
Three-Foil Y	3	-48.5
	5	-12.2

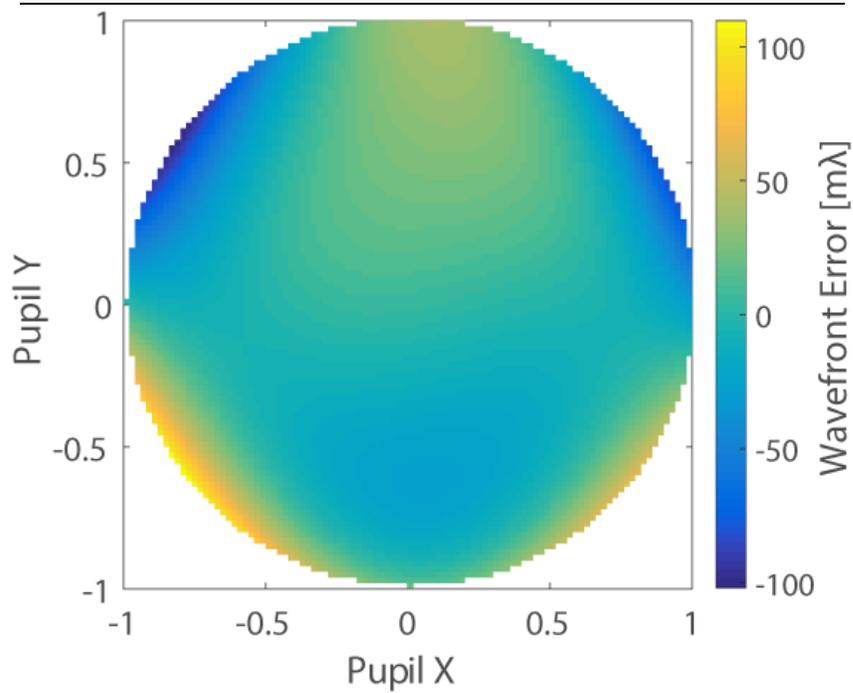


Figure 93. Wavefront extracted from IMEC NXE:3300 with all third-order terms and their fifth-order counterparts. (Z_5 - Z_{16} , Z_{19} , Z_{20})

7. ABERRATIONS IN ANAMORPHIC LITHOGRAPHY SYSTEMS

As we have seen, Zernike aberrations are used as the aberration basis in isomorphic optics with rotational symmetry because they reproduce geometrical aberrations. That physical interpretation is the reason why Zernike polynomials specifically are used over any other basis. This is important because it allows lens designers to minimize the residual wavefront and allows for wavefront compensation and correction.

It is well established that Zernike polynomials do not describe the primary aberrations of optical systems with non-circular pupils. For example, ISO standard 14999-4:2015 describes how to interpret interferometric measurements and states that it is not possible to fit Zernike polynomials to non-circular pupils.⁷⁵ Cassegrain telescopes are catoptric and have a central obscuration. The most famous is undoubtedly the Hubble Space telescope, which has an $\epsilon = 0.33$ central obscuration. The Hubble telescope is also famous for being launched with a large figure error in its primary mirror, causing a large amount of residual spherical aberration. In a comparison of the aberrations pre- and post-correction NASA scientists state, “As is customary in HST optical studies, they [Zernike polynomials] are orthonormalized for a 33% central obscuration, which accounts for the OTA’s secondary mirror.”⁷⁶ Other authors, besides Mahajan who has written extensively on the subject, have derived aberration bases for non-circular pupils.⁷⁷⁻⁷⁹ Therefore, it is reasonable to believe that anamorphic optics may have different aberrations than isomorphic optics.

7.1 Primary aberrations of anamorphic optics

Anamorphic optics are defined as having different magnification factors in different directions of the pupil.^{24,29,80} This leads to differently shaped entrance and exit pupils. In a lithographic anamorphic projection lens the entrance pupil will be elliptical while the exit pupil will be circular. However, as one rotates an anamorphic optic the magnification changes, which in turn changes the action of the optic.

Of course, the optics of a lithography system would never be rotated, which can make this point difficult to visualize, but no less important. It is helpful to use a real anamorphic objective, such as the one by Bausch & Lomb in Figure 94. That lens was held in a purpose-built stand and the entrance pupil was imaged as a function of rotation in Figure 95. As the optic is rotated the image is skewed from perfect circles to an ellipse with a vertical major axis. It is only after a rotation of 180° that the optic again images the object as circles, therefore anamorphic optics can only be considered *reflection invariant*.²⁸

Counterintuitively, because the action of the optic changes through rotation, phase errors in the pupil can lead to different kinds of effects in the image through rotation. Therefore, the phase error should be described by different kinds of aberrations as the optic is rotated. Because anamorphic optics are only reflection invariant we can expect that their wavefront aberration will have the same quality.^{28,80-82} The wavefront aberration can be expanded in a power series using the reflection invariant terms: X_0^2 , Y_0^2 , X^2 , Y^2 , XX_0 , YY_0 , as,

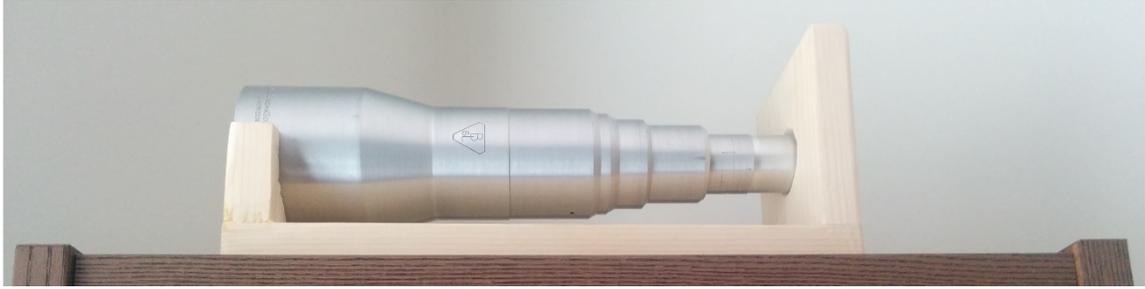


Figure 94. Bausch & Lomb anamorphic lens. (top) The lens in its purpose-built stand, and (bottom) imaging the entrance pupil of the lens.

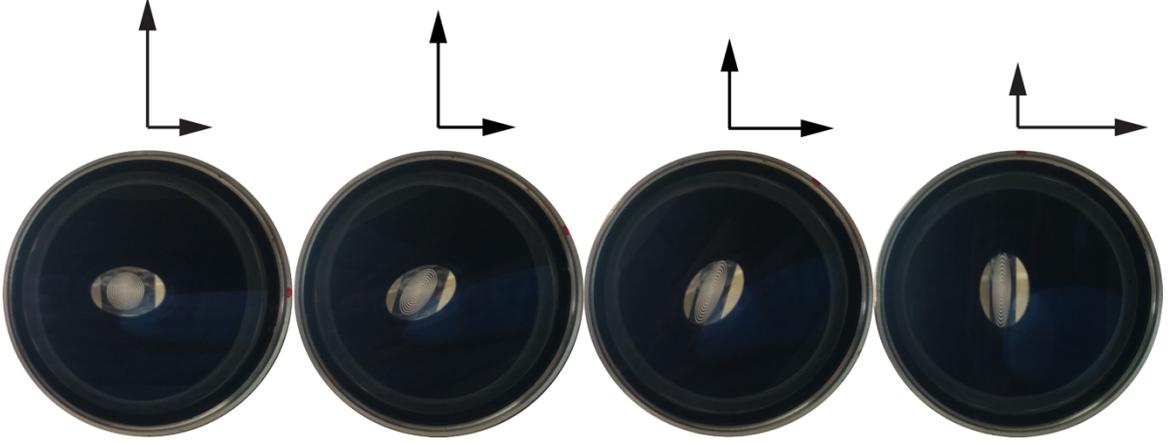


Figure 95. A comparison of the images of the same object as an anamorphic projection lens is rotated. The top shows the effective demagnification in the horizontal and vertical pupil axes.

$$\begin{aligned}
 \Phi(X_0, Y_0, X, Y) = & a_0 + b_0 X_0^2 + b_1 Y_0^2 + b_2 X^2 + b_3 Y^2 + b_4 X_0 X + b_5 Y_0 Y + c_0 X_0^4 + \\
 & c_1 Y_0^4 + c_2 (X_0 Y_0)^2 + c_3 X^4 + c_4 (X_0 X)^2 + c_5 X_0^3 X + c_6 X_0 X^3 + c_7 Y^4 + c_8 (Y_0 Y)^2 + \\
 & c_9 Y_0^3 Y + c_{10} Y_0 Y^3 + c_{11} (XY)^2 + c_{12} X^2 Y_0 Y + c_{13} X_0^2 Y_0 Y + c_{14} Y_0^2 X_0 X + \\
 & c_{15} Y^2 X_0 X + c_{16} (X_0 Y)^2 + c_{17} X_0 X Y_0 Y + c_{18} (Y_0 X)^2 + \dots,
 \end{aligned} \tag{7.1}$$

where X and Y are pupil plane coordinates and X_0 and Y_0 are object space coordinates. If the system were rotationally invariant this expression would simplify to the Seidel representation of aberrations. Instead, we find that the primary aberrations of anamorphic optics are only separable in Cartesian coordinates, instead of polar coordinates, and therefore can be written as a product of 1D aberrations. Ignoring terms that do not contain X or Y , we can see that there are sixteen primary aberrations of anamorphic systems. This

description is equivalent to having separate Seidel aberrations in each plane of symmetry of the optic, as described by Wynne.^{80,82}

We now consider the anamorphic analogue to an isomorphic singlet: two crossed cylindrical lenses, as in Figure 96. In general, this system is not stigmatic and has at most two Gaussian conjugates. Moving either lens in this system would cause defocus only in one direction. This is the definition of astigmatism in an isomorphic optic, but it is more appropriate for this system to interpret this kind of phase error as defocus. In this case, the method for correction of the phase error is the same, but this is not guaranteed to be the case.

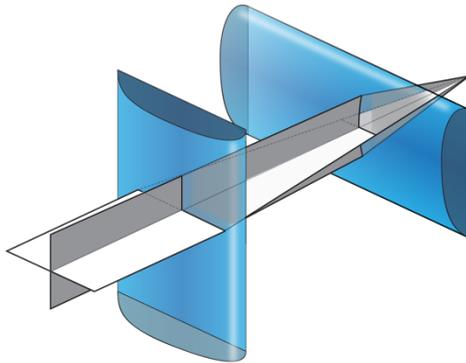


Figure 96. Two crossed cylindrical lens forming a stigmatic image. This is the simplest anamorphic optical system.

To this point, we simulated the residual wavefronts of an anamorphic objective patent lens, Figure 97.⁸³ This lens is made of four distinct sections: 1) an isomorphic section which forms an intermediate image of the object, 2) an isomorphic doublet used to focus the optic, 3) and 4) two separate anamorphic objective sections. Residual wavefronts for this optic were computed using ZEMAX OpticStudio⁸⁴. The lens was detuned by tilting the first cylindrical surface in the X direction by 2 mm. Residual

wavefronts in the ideal and detuned cases are shown in Figure 98. Tilting an isomorphic singlet is known to cause an apparent increase in both tilt and coma. The tilting action does not cause coma in the lens, but coma is balanced by tilt. Therefore, tilting the lens unbalances it and causes an apparent increase in coma. Detuning the anamorphic objective in the same way should create the same kinds of aberrations, but instead we see a very large amount of trefoil. Attempting to correct this phase error as trefoil instead of coma could introduce new phase errors, so it would be incorrect in this case to interpret this as trefoil.

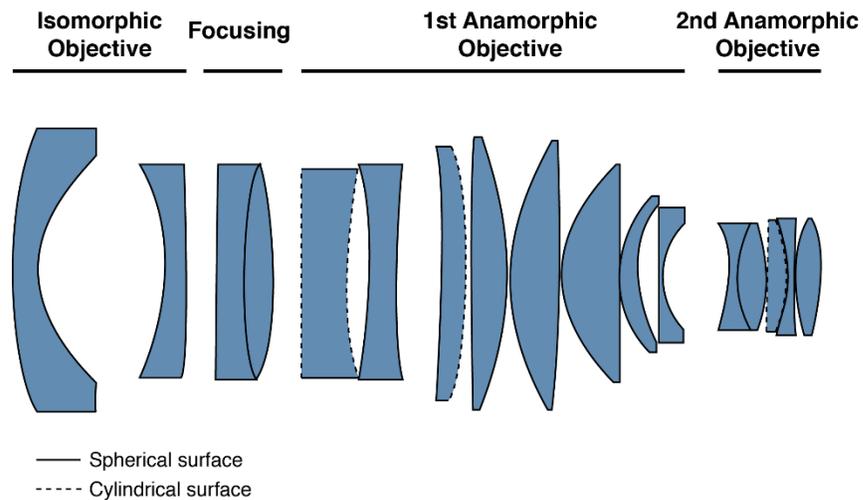


Figure 97. Schematic view of the anamorphic objective patent lens used for simulation of residual wavefronts shown in the XZ plane.

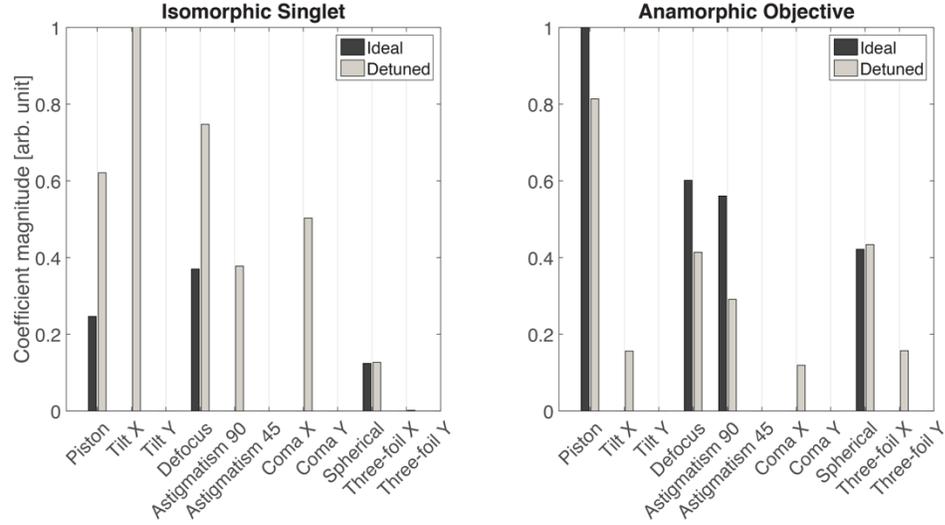


Figure 98. Orthogonal function expansions of residual wavefronts for (left) an isomorphic singlet and (right) the anamorphic objective patent lens. The residual wavefront was computed in ZEMAX OpticStudio in both the ideal and detuned states.

7.2 A new aberration basis

We have seen that the Zernike polynomials are defined as the solutions to,

$$(1 - r^2) \frac{\partial^2 V}{\partial r^2} + (1 - 3r^2) \frac{\partial V}{\partial r} + \frac{1}{r^2} \frac{\partial^2 V}{\partial \phi^2} + n(n+1)rV = 0, \quad (7.2)$$

where r and ϕ are polar coordinates. Transforming Equation (7.2) to Cartesian coordinates from polar coordinates,

$$\nabla^2 V - \left(x \frac{\partial}{\partial x} + y \frac{\partial}{\partial y} \right)^2 V - 2 \left(x \frac{\partial}{\partial x} + y \frac{\partial}{\partial y} \right) V + n(n+1)V = 0. \quad (7.3)$$

Alternatively, if only one dimension is considered the differential equation becomes,

$$(1 - x^2) \frac{d^2 V}{dx^2} - 2x \frac{dV}{dx} + n(n+1)V = 0. \quad (7.4)$$

This is easily recognized as Legendre's differential equation. Therefore, Legendre polynomials represent balanced aberrations in a 1D slit pupil.²⁸ Because the primary

anamorphic aberrations are separable in Cartesian coordinates, products of Legendre polynomials should be used as an orthonormal basis to represent balanced anamorphic aberrations. Products of Legendre polynomials are orthogonal over the unit square, so they must be reorthogonalized over the unit circle. This is achieved via the Gram-Schmidt process as,^{28,85-87}

$$A_{j+1}(u, v) = P_{j+1}(u, v) - \sum_{k=1}^j \frac{A_k(u, v)}{2\pi} Q_{j+1}^k, \quad (7.5)$$

where A_j is an anamorphic aberration function, P_j is a 2D Legendre polynomial, and Q_{j+1}^k is given by,

$$Q_{j+1}^k = \int_{u=-1}^{+1} \int_{v=-\sqrt{1-u^2}}^{+\sqrt{1-u^2}} P_{j+1}(u, v) A_k(u, v) dudv. \quad (7.6)$$

These polynomials can be re-orthogonalized again over an annulus to account for a central obscuration,

$$A_{j+1}(u, v; \epsilon) = A_{j+1}(u, v; \epsilon = 0) - \sum_{k=1}^j \frac{A_k(u, v; \epsilon)}{2\pi} Q_{j+1}^k, \quad (7.7)$$

where ϵ is the obscuration as a percentage of the pupil radius, $A_j(u, v; \epsilon)$ is an annular anamorphic aberration function, $A_j(u, v; \epsilon = 0)$ is one of the anamorphic circle polynomials derived in Equation (7.7), and Q_{j+1}^k is given by,

$$Q_{j+1}^k = \int_{u=-1}^{+1} \int_{v=-\sqrt{1-u^2}}^{+\sqrt{1-u^2}} A_{j+1}(u, v; \epsilon = 0) A_k(u, v; \epsilon) dudv - \int_{u=-\epsilon}^{+\epsilon} \int_{v=-\sqrt{\epsilon^2-u^2}}^{+\sqrt{\epsilon^2-u^2}} A_{j+1}(u, v; \epsilon = 0) A_k(u, v; \epsilon) dudv. \quad (7.8)$$

It is also possible to orthogonalize the Legendre polynomials with a gap in their domain centered at the origin with a width of 2ϵ , then orthogonalize the resulting terms

over an annulus. Two Gram-Schmidt processes are required either way, and the computation complexity is equivalent.

The Gram-Schmidt process generates the functions in an arbitrary order, so we will need to determine the best way to order the terms in this basis. The Zernike polynomials each have a radial order n and an azimuthal order m , but are often given ordered by a single index. One of the most common ordering schemes is the one proposed by Wyant.⁸⁸ First a new quantity is calculated for each combination of (n, m)

$$n' = \frac{n+m}{2}. \quad (7.9)$$

The polynomials are then ordered by increasing n' and decreasing m .

Similarly, each anamorphic aberration function is proportional to a 2D Legendre polynomial, which in turn is a product of a different Legendre polynomial in each pupil direction. Therefore, the 2D Legendre polynomial can be referred to by combinations (q, r) , and

$$A_{j+1} = A_q^r \propto P_q(u)P_r(v). \quad (7.10)$$

From the orders of the Legendre polynomials we define an intermediate quantity,

$$\begin{aligned} n &= q + r \\ m &= q - r \end{aligned} \quad (7.11)$$

The terms n and m are analogous to the Zernike radial and azimuthal orders. It is important to remember however that the anamorphic basis is not separable in polar coordinates, so the terms in Equation (7.11) can only provide an analogy to the Zernike basis. This generates a sequence that is similar to that of the Zernike polynomials, but aberration types appear out of order. To fix this we define two new ordering coefficients,

$$\begin{aligned} n' &= \frac{n+m'}{2} \\ m' &= -|m| + n + \text{mod}(n, 2) \end{aligned} \quad (7.12)$$

Finally, the anamorphic aberration functions can be ordered by increasing n' and decreasing m' . This puts the functions in a familiar order for those already familiar with the Wyant Zernike ordering scheme. 3rd and 5th order anamorphic aberrations ordered in this proposed scheme are shown in Figure 99. Analytic solutions are given in Tables 19 and 20 for the anamorphic circle polynomials and annular polynomials. More details on computation can be found in Appendix A: Annular anamorphic polynomials, and the annular polynomials are computed for several obscurations in Appendix B: Annular anamorphic polynomial tables. Finally, interferograms for third-order anamorphic polynomials are given with 0% and 20% obscurations in Figures 100 and 101.

Notably, spherical aberration is no longer described by a single azimuthally symmetric term but rather by two separate terms, as expected from the previous descriptions of anamorphic aberrations. Additionally, while four-foil is a higher order aberration in isomorphic optics, it is a primary aberration in anamorphic optics.

This residual wavefront of the detuned patent lens was expanded in the proposed anamorphic basis in Figure 102a. There are several terms which only appear in either the Zernike basis or anamorphic basis. The terms in Figure 102a have been grouped together to facilitate comparison. Zernike defocus and spherical aberration are azimuthally symmetric but are plotted in comparison to the corresponding anamorphic aberration in the X direction. Similarly, astigmatism 90° does not exist in the anamorphic basis because it is indistinguishable from separate defocus in the X and Y directions. Noting these differences, we can see that trefoil does not appear in the anamorphic expansion

and instead has been correctly attributed to coma. The expansion in the proposed anamorphic basis makes the necessary correction for the phase error clear, while the Zernike expansion does not.

This is made clearer when examining the transfer matrix between the proposed anamorphic basis and the Zernike polynomials, Figure 102. Many terms in the anamorphic basis are represented in part by the same Zernike terms. In the example patent lens, correcting the apparent trefoil error from the Zernike expansion would help to decrease the anamorphic coma, but it would introduce anamorphic trefoil. Amongst the anamorphic primary aberrations, this same relationship between the two interpretations also applies to the defocus, spherical, and four-foil terms.

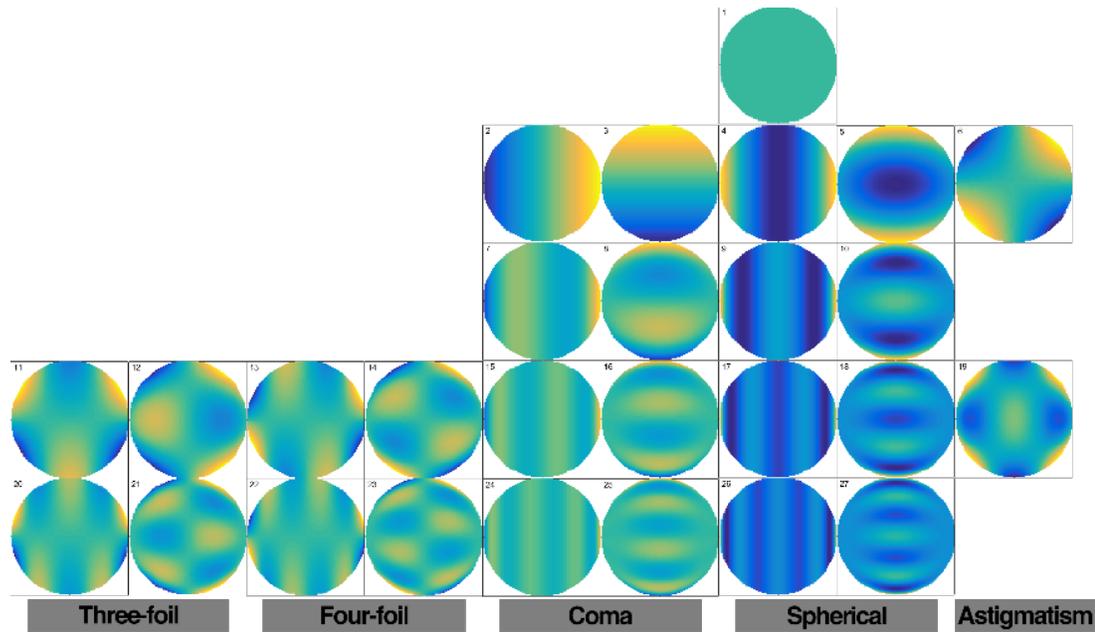


Figure 99. The 3rd and 5th order aberrations of an anamorphic optical system. The names provided are the closest isomorphic aberrations.

Table 19. Analytic solutions for the first 40 anamorphic circle polynomials.

N	n'	m'	Anamorphic Circle Polynomial ($A_N(u, v)$)
1	0	1	1
2	1	1	$2u$
3	1	1	$2v$
4	1	0	$4u^2 - 1$
5	1	0	$\sqrt{2}(u^2 + 3v^2 - 1)$
6	2	2	$2\sqrt{6}uv$
7	2	1	$8u^3 - 4u$
8	2	1	$4v(3u^2 + 5v^2 - 3) / \sqrt{5}$
9	2	0	$16u^4 - 12u^2 + 1$
10	2	0	$\sqrt{2/7}(3u^4 + 6(5v^2 - 1)u^2 + 35v^4 - 30v^2 + 3)$
11	3	3	$4v(6u^2 - 1) / \sqrt{5}$
12	3	3	$4u(u^2 + 3v^2 - 1)$
13	3	2	$2\sqrt{2}u(8u^2 - 3)v$
14	3	2	$4\sqrt{2}uv(3u^2 + 5v^2 - 3)$
15	3	1	$32u^5 - 32u^3 + 6u$
16	3	1	$2\sqrt{2/21}v(15u^4 + 10(7v^2 - 3)u^2 + 63v^4 - 70v^2 + 15)$
17	3	0	$64u^6 - 80u^4 + 24u^2 - 1$
18	3	0	$2(5u^6 + 15(7v^2 - 1)u^4 + 15(21v^4 - 14v^2 + 1)u^2 + 231v^6 - 315v^4 + 105v^2 - 5) / \sqrt{33}$
19	4	4	$\sqrt{10/7}(8u^2 - 1)(u^2 + 3v^2 - 1)$
20	4	3	$2\sqrt{3/35}(80u^4 - 48u^2 + 3)v$
21	4	3	$2\sqrt{6/7}u(3u^4 + 6(5v^2 - 1)u^2 + 35v^4 - 30v^2 + 3)$
22	4	2	$4u(24u^4 - 20u^2 + 3)v$
23	4	2	$4uv(15u^4 + 10(7v^2 - 3)u^2 + 63v^4 - 70v^2 + 15) / \sqrt{3}$
24	4	1	$8u(16u^6 - 24u^4 + 10u^2 - 1)$
25	4	1	$8v(35u^6 + 105(3v^2 - 1)u^4 + 21(33v^4 - 30v^2 + 5)u^2 + 429v^6 - 693v^4 + 315v^2 - 35) / \sqrt{429}$
26	4	0	$256u^8 - 448u^6 + 240u^4 - 40u^2 + 1$
27	4	0	$\sqrt{2/715}(35u^8 + 140(9v^2 - 1)u^6 + 210(33v^4 - 18v^2 + 1)u^4 + 28(429v^6 - 495v^4 + 135v^2 - 5)u^2 + 6435v^8 - 12012v^6 + 6930v^4 - 1260v^2 + 35)$
28	5	5	$4\sqrt{2/7}u(10u^2 - 3)(u^2 + 3v^2 - 1)$
29	5	5	$4\sqrt{2/15}(10u^2 - 1)v(3u^2 + 5v^2 - 3)$
30	5	4	$2(40u^4 - 20u^2 + 1)(u^2 + 3v^2 - 1) / \sqrt{3}$
31	5	4	$\sqrt{2/11}(12u^2 - 1)(3u^4 + 6(5v^2 - 1)u^2 + 35v^4 - 30v^2 + 3)$
32	5	3	$8v(112u^6 - 120u^4 + 30u^2 - 1) / \sqrt{21}$
34	5	3	$8u(5u^6 + 15(7v^2 - 1)u^4 + 15(21v^4 - 14v^2 + 1)u^2 + 231v^6 - 315v^4 + 105v^2 - 5) / \sqrt{33}$
35	5	2	$4\sqrt{3/5}u(128u^6 - 168u^4 + 60u^2 - 5)v$
36	5	2	$8\sqrt{6/143} \cdot uv(35u^6 + 105(3v^2 - 1)u^4 + 21(33v^4 - 30v^2 + 5)u^2 + 429v^6 - 693v^4 + 315v^2 - 35)$
37	5	1	$2u(256u^8 - 512u^6 + 336u^4 - 80u^2 + 5)$
38	5	1	$2\sqrt{2/2431}v(315u^8 + 420(11v^2 - 3)u^6 + 126(143v^4 - 110v^2 + 15)u^4 + 36(715v^6 - 1001v^4 + 385v^2 - 35)u^2 + 12155v^8 - 25740v^6 + 18018v^4 - 4620v^2 + 315)$
39	5	0	$1024u^{10} - 2304u^8 + 1792u^6 - 560u^4 + 60u^2 - 1$
40	5	0	$2 \cdot [63u^{10} + 315(11v^2 - 1)u^8 + 210(143v^4 - 66v^2 + 3)u^6 + 630(143v^6 - 143v^4 + 33v^2 - 1)u^4 + 45(2431v^8 - 4004v^6 + 2002v^4 - 308v^2 + 7)u^2 + 46189v^{10} - 109395v^8 + 90090v^6 - 30030v^4 + 3465v^2 - 63] / \sqrt{4199}$

Table 20. Analytic solutions for the first 18 annular anamorphic polynomials.

N	n'	m'	Anamorphic Circle Polynomial ($A_n(u, v; \epsilon) = N_n/D_n$)
1	0	0	$N_1 = 1$
2	1	1	$N_2 = 2u$ $D_2 = \sqrt{\epsilon^2 + 1}$
3	1	1	$N_3 = 2v$ $D_3 = \sqrt{\epsilon^2 + 1}$
4	1	0	$N_4 = -4u^2 + \epsilon^2 + 1$ $D_4 = \sqrt{\epsilon^4 + 1}$
5	1	0	$N_5 = \sqrt{2}(\epsilon^6 + 2\epsilon^4 + 2\epsilon^2 - 3v^2(\epsilon^4 + 1) - u^2(\epsilon^4 + 4\epsilon^2 + 1) + 1)$ $D_5 = (\epsilon^2 - 1)\sqrt{\epsilon^8 + \epsilon^6 + 2\epsilon^4 + \epsilon^2 + 1}$
6	2	2	$N_6 = 2\sqrt{6}uv$ $D_6 = \sqrt{\epsilon^4 + \epsilon^2 + 1}$
7	2	1	$N_7 = 4u(\epsilon^4 + \epsilon^2 - 2u^2(\epsilon^2 + 1) + 1)$ $D_7 = \sqrt{\epsilon^{10} + 3\epsilon^8 + 3\epsilon^2 + 1}$
8	2	1	$N_8 = 4v(3(\epsilon^8 + 2\epsilon^6 + 6\epsilon^4 + 2\epsilon^2 + 1)u^2 + v^2(5\epsilon^8 + 10\epsilon^6 + 6\epsilon^4 + 10\epsilon^2 + 5) - 3(\epsilon^{10} + 2\epsilon^8 + 3\epsilon^6 + 3\epsilon^4 + 2\epsilon^2 + 1))$ $D_8 = (\epsilon^2 - 1)\sqrt{(5\epsilon^8 + 10\epsilon^6 + 6\epsilon^4 + 10\epsilon^2 + 5)(\epsilon^{10} + 5\epsilon^8 + 6\epsilon^6 + 6\epsilon^4 + 5\epsilon^2 + 1)}$
9	2	0	$N_9 = \epsilon^8 + 5\epsilon^6 - 6(v^2 - 1)\epsilon^4 + (5 - 6v^2)\epsilon^2 + 16uv^4(\epsilon^4 + \epsilon^2 + 1) - 6u^2(2\epsilon^6 + 3\epsilon^4 + 3\epsilon^2 + 2) + 1$ $D_9 = \sqrt{\epsilon^{16} + 2\epsilon^{14} + 5\epsilon^{12} + 2\epsilon^8 + 5\epsilon^4 + 2\epsilon^2 + 1}$
10	2	0	$N_{10} = \sqrt{2} \cdot [3(\epsilon^{16} + 30\epsilon^{14} + 90\epsilon^{12} + 114\epsilon^{10} + 130\epsilon^8 + 114\epsilon^6 + 90\epsilon^4 + 30\epsilon^2 + 1)u^4 + 6(-\epsilon^{18} - 18\epsilon^{16} - 60\epsilon^{14} - 102\epsilon^{12} - 119\epsilon^{10} - 119\epsilon^8 - 102\epsilon^6 - 60\epsilon^4 - 18\epsilon^2 + 5v^2(\epsilon^{16} + 6\epsilon^{14} + 18\epsilon^{12} + 26\epsilon^{10} + 18\epsilon^8 + 26\epsilon^6 + 18\epsilon^4 + 6\epsilon^2 + 1) - 1)u^2 + 5v^4(7\epsilon^{16} + 18\epsilon^{14} + 54\epsilon^{12} + 62\epsilon^{10} + 78\epsilon^8 + 62\epsilon^6 + 54\epsilon^4 + 18\epsilon^2 + 7) - 6v^2(5\epsilon^{18} + 18\epsilon^{16} + 60\epsilon^{14} + 102\epsilon^{12} + 115\epsilon^{10} + 115\epsilon^8 + 102\epsilon^6 + 60\epsilon^4 + 18\epsilon^2 + 5) + 3(\epsilon^{20} + 9\epsilon^{18} + 36\epsilon^{16} + 84\epsilon^{14} + 111\epsilon^{12} + 118\epsilon^{10} + 111\epsilon^8 + 84\epsilon^6 + 36\epsilon^4 + 9\epsilon^2 + 1)]$ $D_{10} = (\epsilon^2 - 1)^2 \sqrt{(\epsilon^{16} + 5\epsilon^{14} + 15\epsilon^{12} + 19\epsilon^{10} + 20\epsilon^8 + 19\epsilon^6 + 15\epsilon^4 + 5\epsilon^2 + 1)(7\epsilon^{16} + 18\epsilon^{14} + 54\epsilon^{12} + 62\epsilon^{10} + 78\epsilon^8 + 62\epsilon^6 + 54\epsilon^4 + 18\epsilon^2 + 7)}$
11	3	3	$N_{11} = 4v(\epsilon^4 + \epsilon^2 - 6u^2(\epsilon^2 + 1) + 1)$ $D_{11} = \sqrt{(\epsilon^2 + 1)(5\epsilon^8 + 10\epsilon^6 + 6\epsilon^4 + 10\epsilon^2 + 5)}$
12	3	3	$N_{12} = -4u(-\epsilon^{10} - 2\epsilon^8 - 3\epsilon^6 - 3\epsilon^4 - 2\epsilon^2 + u^2(\epsilon^8 + 2\epsilon^6 + 6\epsilon^4 + 2\epsilon^2 + 1) + 3v^2(\epsilon^8 + 2\epsilon^6 - 2\epsilon^4 + 2\epsilon^2 + 1) - 1)$ $D_{12} = (\epsilon^2 - 1)\sqrt{(\epsilon^8 + 2\epsilon^6 - 2\epsilon^4 + 2\epsilon^2 + 1)(\epsilon^{10} + 5\epsilon^8 + 6\epsilon^6 + 6\epsilon^4 + 5\epsilon^2 + 1)}$
13	3	2	$N_{13} = 2\sqrt{2}uv(8u^2(\epsilon^4 + \epsilon^2 + 1) - 3(\epsilon^6 + \epsilon^4 + \epsilon^2 + 1))$ $D_{13} = \sqrt{(\epsilon^4 + \epsilon^2 + 1)(\epsilon^{12} + 2\epsilon^{10} + 3\epsilon^8 + 3\epsilon^6 + 3\epsilon^4 + 2\epsilon^2 + 1)}$
14	3	2	$N_{14} = -4\sqrt{2}uv(3(\epsilon^{12} + 2\epsilon^{10} + 3\epsilon^8 + 8\epsilon^6 + 3\epsilon^4 + 2\epsilon^2 + 1)u^2 + 5v^2(\epsilon^{12} + 2\epsilon^{10} + 3\epsilon^8 + 3\epsilon^6 + 3\epsilon^4 + 2\epsilon^2 + 1) - 3(\epsilon^{14} + 2\epsilon^{12} + 3\epsilon^{10} + 4\epsilon^8 + 4\epsilon^6 + 3\epsilon^4 + 2\epsilon^2 + 1))$ $D_{14} = (\epsilon^2 - 1)\sqrt{(\epsilon^{12} + 2\epsilon^{10} + 3\epsilon^8 + 3\epsilon^6 + 3\epsilon^4 + 2\epsilon^2 + 1)(\epsilon^{16} + 5\epsilon^{14} + 15\epsilon^{12} + 19\epsilon^{10} + 20\epsilon^8 + 19\epsilon^6 + 15\epsilon^4 + 5\epsilon^2 + 1)}$
15	3	1	$N_{15} = 32(\epsilon^{10} + 5\epsilon^8 + 6\epsilon^6 + 6\epsilon^4 + 5\epsilon^2 + 1)u^5 - 8(4\epsilon^{12} + 20\epsilon^{10} + 33\epsilon^8 + 36\epsilon^6 + 33\epsilon^4 + 20\epsilon^2 + 4)u^3 + 6(\epsilon^{14} + 5\epsilon^{12} + 15\epsilon^{10} + (19 - 12v^2)\epsilon^8 + (19 - 16v^2)\epsilon^6 - 3(4v^2 - 5)\epsilon^4 + 5\epsilon^2 + 1)u$ $D_{15} = \sqrt{(\epsilon^{10} + 5\epsilon^8 + 6\epsilon^6 + 6\epsilon^4 + 5\epsilon^2 + 1)(\epsilon^{20} + 6\epsilon^{18} + 3\epsilon^{16} + 12\epsilon^{14} + 12\epsilon^{12} - 20\epsilon^{10} + 12\epsilon^8 + 12\epsilon^6 + 3\epsilon^4 + 6\epsilon^2 + 1)}$

Table 20. (Continued from previous page)

N	n'	m'	Anamorphic Circle Polynomial ($A_n(u, v; \epsilon) = N_n/D_n$)
16	3	1	$N_{16} = 2\sqrt{2}v(15(\epsilon^{24} + 8\epsilon^{22} + 91\epsilon^{20} + 240\epsilon^{18} + 425\epsilon^{16} + 448\epsilon^{14} + 574\epsilon^{12} + 448\epsilon^{10} + 425\epsilon^8 + 240\epsilon^6 + 91\epsilon^4 + 8\epsilon^2 + 1)u^4 + 10(v^2(7\epsilon^{24} + 56\epsilon^{22} + 277\epsilon^{20} + 720\epsilon^{18} + 1295\epsilon^{16} + 1456\epsilon^{14} + 1378\epsilon^{12} + 1456\epsilon^{10} + 1295\epsilon^8 + 720\epsilon^6 + 277\epsilon^4 + 56\epsilon^2 + 7) - 3(\epsilon^{26} + 8\epsilon^{24} + 61\epsilon^{22} + 170\epsilon^{20} + 330\epsilon^{18} + 439\epsilon^{16} + 491\epsilon^{14} + 491\epsilon^{12} + 439\epsilon^{10} + 330\epsilon^8 + 170\epsilon^6 + 61\epsilon^4 + 8\epsilon^2 + 1)u^2 + 3v^4(21\epsilon^{24} + 168\epsilon^{22} + 471\epsilon^{20} + 1200\epsilon^{18} + 2125\epsilon^{16} + 2208\epsilon^{14} + 2614\epsilon^{12} + 2208\epsilon^{10} + 2125\epsilon^8 + 1200\epsilon^6 + 471\epsilon^4 + 168\epsilon^2 + 21) - 10v^2(7\epsilon^{26} + 56\epsilon^{24} + 187\epsilon^{22} + 510\epsilon^{20} + 990\epsilon^{18} + 1313\epsilon^{16} + 1437\epsilon^{14} + 1437\epsilon^{12} + 1313\epsilon^{10} + 990\epsilon^8 + 510\epsilon^6 + 187\epsilon^4 + 56\epsilon^2 + 7) + 15(\epsilon^{28} + 8\epsilon^{26} + 37\epsilon^{24} + 112\epsilon^{22} + 244\epsilon^{20} + 400\epsilon^{18} + 450\epsilon^{16} + 496\epsilon^{14} + 450\epsilon^{12} + 400\epsilon^{10} + 244\epsilon^8 + 112\epsilon^6 + 37\epsilon^4 + 8\epsilon^2 + 1))$ $D_{16} = (\epsilon^2 - 1)^2(21\epsilon^{24} + 168\epsilon^{22} + 471\epsilon^{20} + 1200\epsilon^{18} + 2125\epsilon^{16} + 2208\epsilon^{14} + 2614\epsilon^{12} + 2208\epsilon^{10} + 2125\epsilon^8 + 1200\epsilon^6 + 471\epsilon^4 + 168\epsilon^2 + 21) (\epsilon^{26} + 13\epsilon^{24} + 56\epsilon^{22} + 160\epsilon^{20} + 340\epsilon^{18} + 444\epsilon^{16} + 486\epsilon^{14} + 486\epsilon^{12} + 444\epsilon^{10} + 340\epsilon^8 + 160\epsilon^6 + 56\epsilon^4 + 13\epsilon^2 + 1)^{1/2}$
17	3	0	$N_{17} = -\epsilon^{22} - 14\epsilon^{20} + (36v^2 - 69)\epsilon^{18} - 2(15v^4 - 96v^2 + 100)\epsilon^{16} + (-150v^4 + 408v^2 - 331)\epsilon^{14} + (-210v^4 + 564v^2 - 385)\epsilon^{12} - 5(42v^4 - 120v^2 + 77)\epsilon^{10} + (-210v^4 + 564v^2 - 331)\epsilon^8 + (-210v^4 + 408v^2 - 200)\epsilon^6 - 3(50v^4 - 64v^2 + 23)\epsilon^4 - 2(15v^4 - 18v^2 + 7)\epsilon^2 + 64v^6(\epsilon^{16} + 5\epsilon^{14} + 15\epsilon^{12} + 19\epsilon^{10} + 20\epsilon^8 + 19\epsilon^6 + 15\epsilon^4 + 5\epsilon^2 + 1) - 10u^4(8\epsilon^{18} + 43\epsilon^{16} + 135\epsilon^{14} + 221\epsilon^{12} + 253\epsilon^{10} + 253\epsilon^8 + 221\epsilon^6 + 135\epsilon^4 + 43\epsilon^2 + 8) + 12u^2(2\epsilon^{20} + 13\epsilon^{18} + (46 - 5v^2)\epsilon^{16} + (104 - 25v^2)\epsilon^{14} + (137 - 75v^2)\epsilon^{12} + (146 - 95v^2)\epsilon^{10} + (137 - 95v^2)\epsilon^8 + (104 - 75v^2)\epsilon^6 + (46 - 25v^2)\epsilon^4 + (13 - 5v^2)\epsilon^2 + 2) - 1$ $D_{17} = [(\epsilon^{16} + 5\epsilon^{14} + 15\epsilon^{12} + 19\epsilon^{10} + 20\epsilon^8 + 19\epsilon^6 + 15\epsilon^4 + 5\epsilon^2 + 1)(\epsilon^{28} + 5\epsilon^{26} + 18\epsilon^{24} + 8\epsilon^{22} + 34\epsilon^{20} + 35\epsilon^{18} + 27\epsilon^{16} - 56\epsilon^{14} + 27\epsilon^{12} + 35\epsilon^{10} + 34\epsilon^8 + 8\epsilon^6 + 18\epsilon^4 + 5\epsilon^2 + 1)]^{1/2}$
18	3	0	$N_{18} = 2(5(\epsilon^{40} + 147\epsilon^{38} + 1568\epsilon^{36} + 7614\epsilon^{34} + 25580\epsilon^{32} + 66154\epsilon^{30} + 138376\epsilon^{28} + 229193\epsilon^{26} + 300975\epsilon^{24} + 340422\epsilon^{22} + 352440\epsilon^{20} + 340422\epsilon^{18} + 300975\epsilon^{16} + 229193\epsilon^{14} + 138376\epsilon^{12} + 66154\epsilon^{10} + 25580\epsilon^8 + 7614\epsilon^6 + 1568\epsilon^4 + 147\epsilon^2 + 1)u^6 + 15(-\epsilon^{42} - 87\epsilon^{40} - 938\epsilon^{38} - 4894\epsilon^{36} - 17370\epsilon^{34} - 47144\epsilon^{32} - 103557\epsilon^{30} - 183749\epsilon^{28} - 263856\epsilon^{26} - 319418\epsilon^{24} - 345236\epsilon^{22} - 345236\epsilon^{20} - 319418\epsilon^{18} - 263856\epsilon^{16} - 183749\epsilon^{14} - 103557\epsilon^{12} - 47144\epsilon^{10} - 17370\epsilon^8 - 4894\epsilon^6 - 938\epsilon^4 - 87\epsilon^2 + 7v^2(\epsilon^{40} + 27\epsilon^{38} + 248\epsilon^{36} + 1254\epsilon^{34} + 3980\epsilon^{32} + 9838\epsilon^{30} + 19996\epsilon^{28} + 32453\epsilon^{26} + 42795\epsilon^{24} + 48222\epsilon^{22} + 49872\epsilon^{20} + 48222\epsilon^{18} + 42795\epsilon^{16} + 32453\epsilon^{14} + 19996\epsilon^{12} + 9838\epsilon^{10} + 3980\epsilon^8 + 1254\epsilon^6 + 248\epsilon^4 + 27\epsilon^2 + 1) - 1)u^4 + 47\epsilon^{42} + 508\epsilon^{40} + 2904\epsilon^{38} + 11105\epsilon^{36} + 32119\epsilon^{34} + 74933\epsilon^{32} + 143270\epsilon^{30} + 224182\epsilon^{28} + 292059\epsilon^{26} + 332477\epsilon^{24} + 345290\epsilon^{22} + 332477\epsilon^{20} + 292059\epsilon^{18} + 224182\epsilon^{16} + 143270\epsilon^{14} + 74933\epsilon^{12} + 32119\epsilon^{10} + 11105\epsilon^8 + 2904\epsilon^6 + 508\epsilon^4 + 47\epsilon^2 + 7v^4(3\epsilon^{40} + 41\epsilon^{38} + 304\epsilon^{36} + 1402\epsilon^{34} + 4260\epsilon^{32} + 10214\epsilon^{30} + 20288\epsilon^{28} + 32259\epsilon^{26} + 42085\epsilon^{24} + 47986\epsilon^{22} + 49816\epsilon^{20} + 47986\epsilon^{18} + 42085\epsilon^{16} + 32259\epsilon^{14} + 20288\epsilon^{12} + 10214\epsilon^{10} + 4260\epsilon^8 + 1402\epsilon^6 + 304\epsilon^4 + 41\epsilon^2 + 3) - 2v^2(7\epsilon^{42} + 129\epsilon^{40} + 1105\epsilon^{38} + 19020\epsilon^{36} + 5698\epsilon^{34} + 49388\epsilon^{32} + 105519\epsilon^{30} + 183743\epsilon^{28} + 261852\epsilon^{26} + 183743\epsilon^{24} + 105519\epsilon^{22} + 49388\epsilon^{20} + 19020\epsilon^{18} + 5698\epsilon^{16} + 1106\epsilon^{14} + 129\epsilon^{12} + 261852\epsilon^{10} + 317006\epsilon^8 + 342782\epsilon^6 + 342782\epsilon^4 + 317006\epsilon^2 + 22380\epsilon^{32} + 52774\epsilon^{30} + 102868\epsilon^{28} + 159549\epsilon^{26} + 209555\epsilon^{24} + 238046\epsilon^{22} + 247136\epsilon^{20} + 238046\epsilon^{18} + 7) + 1)u^2 + 7v^6(33\epsilon^{40} + 331\epsilon^{38} + 2024\epsilon^{36} + 7622\epsilon^{34} + 22380\epsilon^{32} + 52774\epsilon^{30} + 102868\epsilon^{28} + 159549\epsilon^{26} + 209555\epsilon^{24} + 238046\epsilon^{22} + 247136\epsilon^{20} + 238046\epsilon^{18} + 7) + 15v^4(21\epsilon^{42} + 227\epsilon^{40} + 1498\epsilon^{38} + 6374\epsilon^{36} + 20350\epsilon^{34} + 51504\epsilon^{32} + 107697\epsilon^{30} + 183729\epsilon^{28} + 259576\epsilon^{26} + 314498\epsilon^{24} + 183729\epsilon^{22} + 51504\epsilon^{20} + 107697\epsilon^{18} + 1498\epsilon^{16} + 227\epsilon^{14} + 15v^2(7\epsilon^{44} + 89\epsilon^{42} + 676\epsilon^{40} + 3408\epsilon^{38} + 12215\epsilon^{36} + 33913\epsilon^{34} + 77051\epsilon^{32} + 144650\epsilon^{30} + 222754\epsilon^{28} + 289773\epsilon^{26} + 330179\epsilon^{24} + 343070\epsilon^{22} + 330179\epsilon^{20} + 289773\epsilon^{18} + 222754\epsilon^{16} + 144650\epsilon^{14} + 77051\epsilon^{12} + 33913\epsilon^{10} + 12215\epsilon^8 + 3408\epsilon^6 + 676\epsilon^4 + 89\epsilon^2 + 7) - 5(\epsilon^{46} + 23\epsilon^{44} + 240\epsilon^{42} + 1532\epsilon^{40} + 6529\epsilon^{38} + 20619\epsilon^{36} + 51922\epsilon^{34} + 107788\epsilon^{32} + 183727\epsilon^{30} + 259471\epsilon^{28} + 314046\epsilon^{26} + 340352\epsilon^{24} + 340352\epsilon^{22} + 314046\epsilon^{20} + 259471\epsilon^{18} + 183727\epsilon^{16} + 107788\epsilon^{14} + 51922\epsilon^{12} + 20619\epsilon^{10} + 6529\epsilon^8 + 1532\epsilon^6 + 240\epsilon^4 + 23\epsilon^2 + 1))$ $D_{18} = (\epsilon^2 - 1)^3[(\epsilon^{40} + 14\epsilon^{38} + 105\epsilon^{36} + 460\epsilon^{34} + 1430\epsilon^{32} + 3504\epsilon^{30} + 7056\epsilon^{28} + 11318\epsilon^{26} + 14857\epsilon^{24} + 16840\epsilon^{22} + 17455\epsilon^{20} + 16840\epsilon^{18} + 14857\epsilon^{16} + 11318\epsilon^{14} + 7056\epsilon^{12} + 3504\epsilon^{10} + 1430\epsilon^8 + 460\epsilon^6 + 105\epsilon^4 + 14\epsilon^2 + 1)(33\epsilon^{40} + 331\epsilon^{38} + 2024\epsilon^{36} + 7622\epsilon^{34} + 22380\epsilon^{32} + 52774\epsilon^{30} + 102868\epsilon^{28} + 159549\epsilon^{26} + 209555\epsilon^{24} + 238046\epsilon^{22} + 247136\epsilon^{20} + 238046\epsilon^{18} + 159555\epsilon^{16} + 159549\epsilon^{14} + 102868\epsilon^{12} + 52774\epsilon^{10} + 22380\epsilon^8 + 7622\epsilon^6 + 2024\epsilon^4 + 331\epsilon^2 + 33)]^{1/2}$

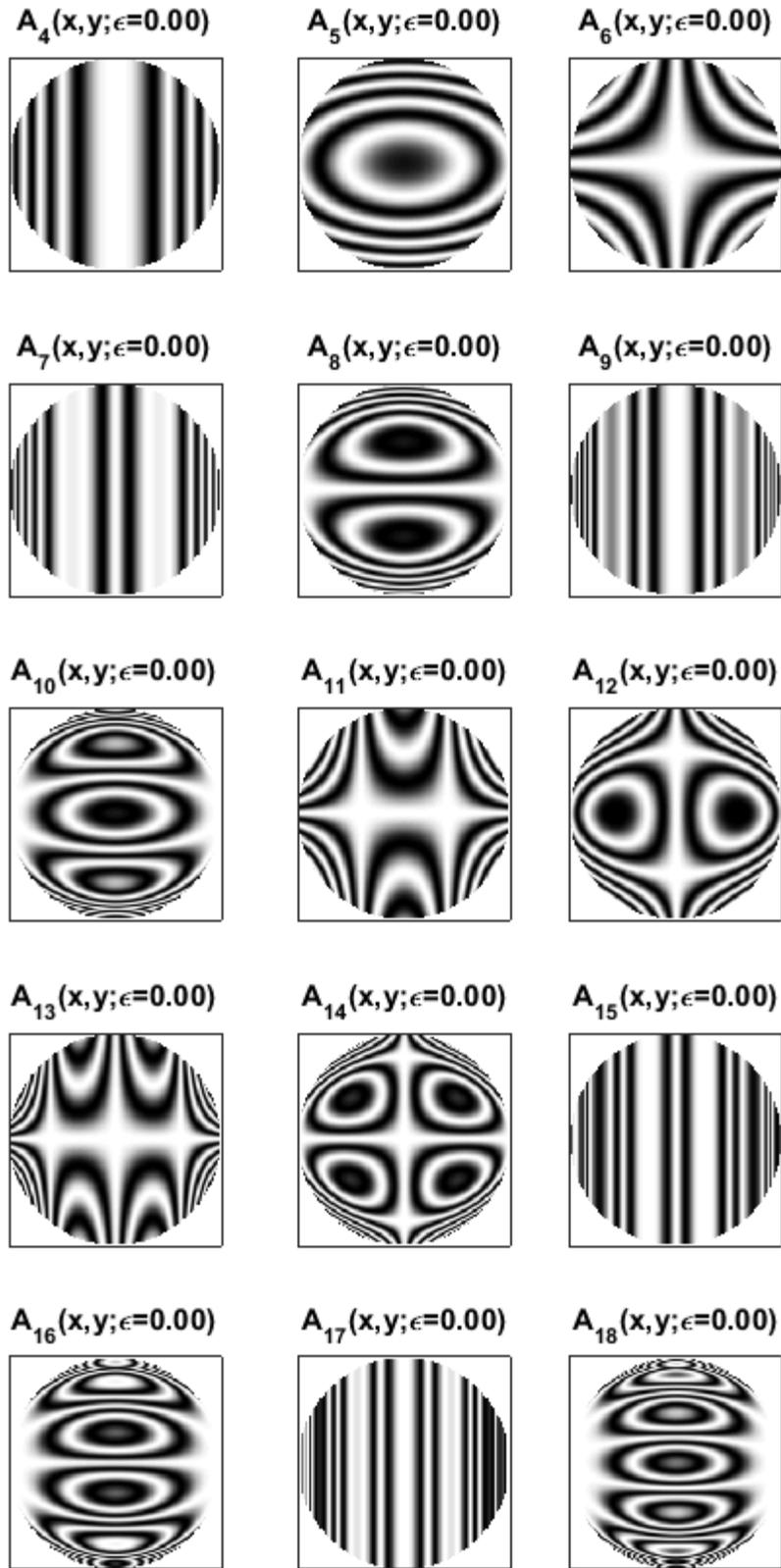


Figure 100. Interferograms for the third-order anamorphic circle polynomials.

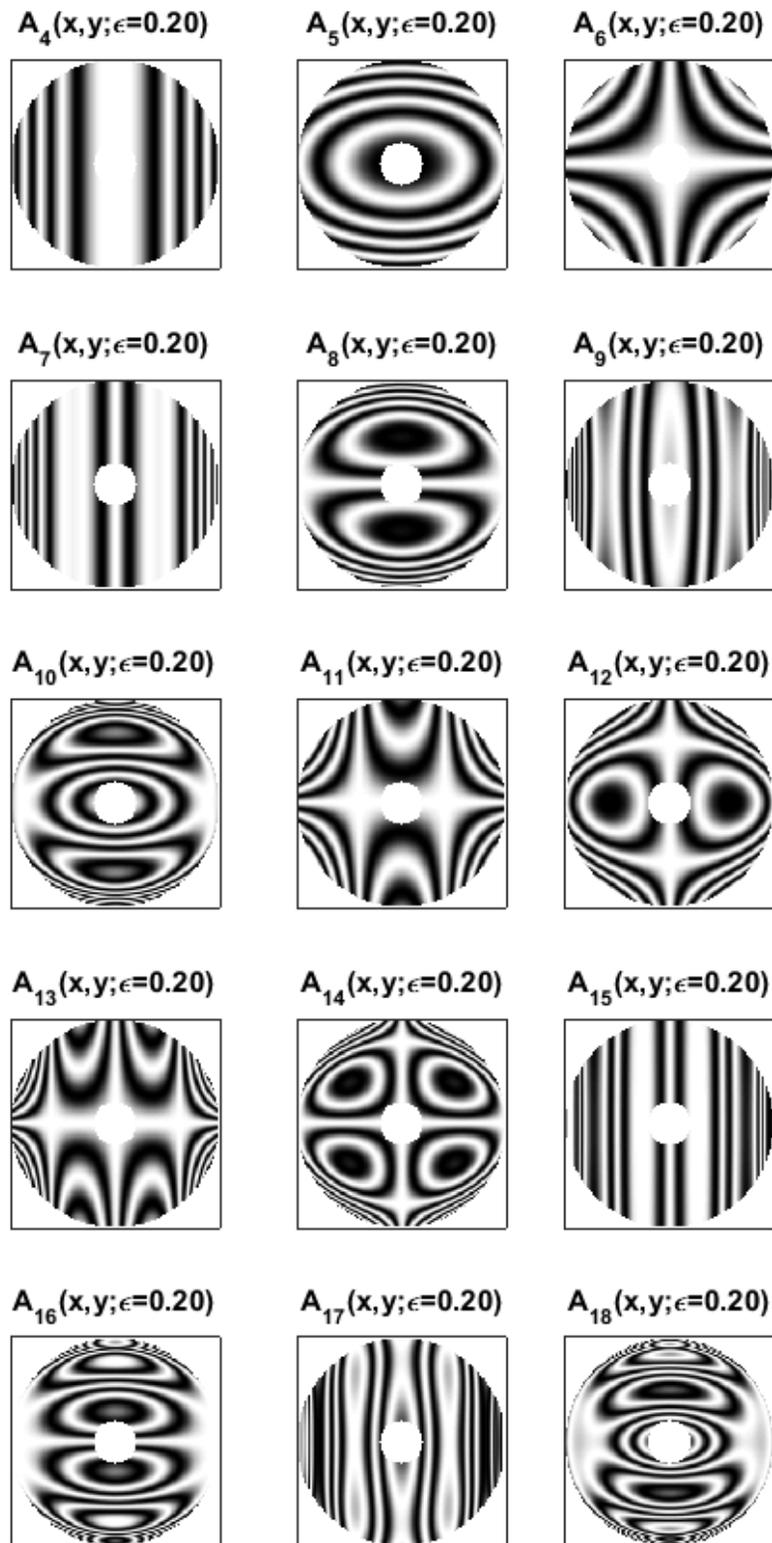


Figure 101. Interferograms for the third-order annular anamorphic polynomials with a 20% obscuration.

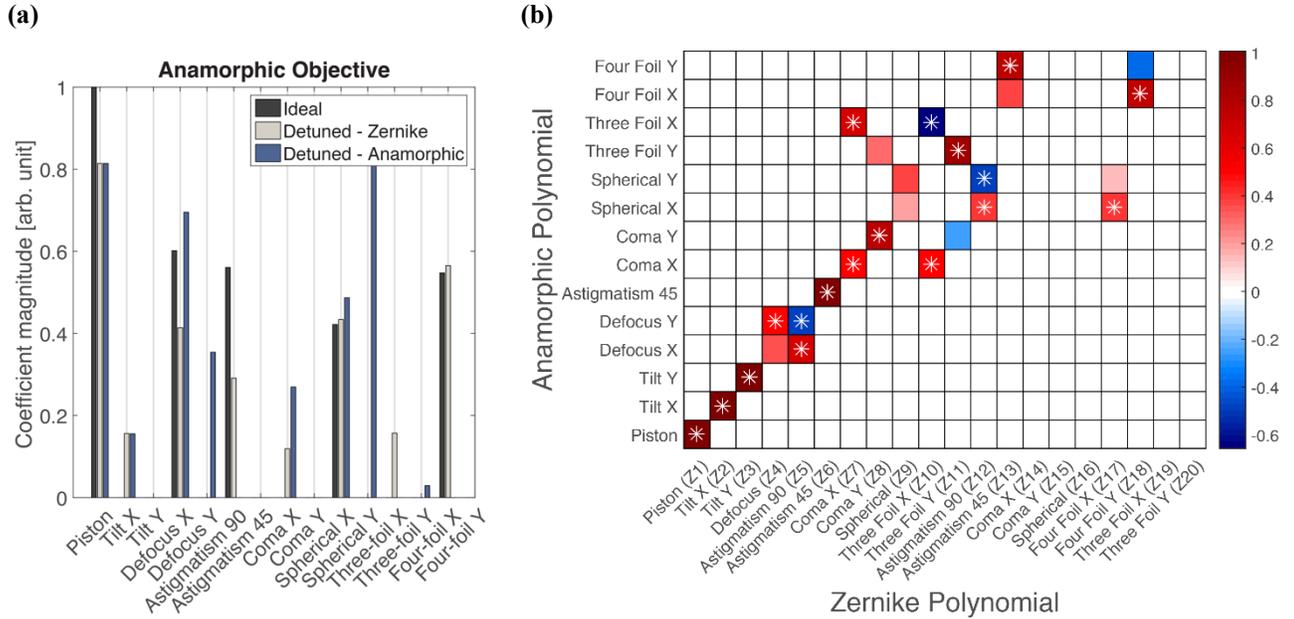


Figure 102. (a) Orthogonal function expansions of residual wavefronts the anamorphic objective patent lens in both a Zernike basis and the proposed anamorphic basis. The residual wavefront was computed in ZEMAX OpticStudio in both the ideal and detuned states. Note that defocus Y and spherical Y do not exist in the Zernike polynomials, Zernike defocus and spherical is shown as being in the X direction in this plot. Similarly, astigmatism 90° does not exist in the anamorphic basis. (b) Transfer matrix between Zernike polynomials and the anamorphic primary aberrations. The stars are placed in the terms which have the largest magnitude for each particular term.

7.3 Simulating anamorphic aberrations

We have written a custom application in MATLAB to compute aberrations using the proposed basis for anamorphic aberrations, as seen in Figure 103. The Gram-Schmidt process described by Equation (7.5) is a computationally expensive procedure. Computing high order terms can be quite slow, so the anamorphic basis is pre-computed up to A_{68} . This ensures that computation of pupil filters is fast on any machine. The application supports creating a wavefront with both phase and amplitude errors in the proposed anamorphic basis. Once an anamorphic wavefront has been generated the user has several options: the wavefront can be exported directly to PROLITH⁷², it can be written to a

PROLITH pupil filter file, or it can be expanded in Zernike polynomials for simulation in any existing software.

It is often necessary to simulate the effects of aberrations at multiple levels, a so-called simulation set. The application will compute the equivalent sets of Zernike coefficients to vary in groups to achieve the desired simulation. Varying levels of anamorphic aberrations cannot be run directly through PROLITH as simulation sets, and instead must be run as a series of single simulations with pupil filter files. This is because some simulations may require Zernike coefficients to be simulated in decreasing order to achieve the desired simulation, which PROLITH does not allow. Instead, the application can write the necessary PROLITH pupil filter files for a simulation set along with a MATLAB script to run the simulation. The interface for simulation sets is shown in Figure 104. This software is freely available.⁸⁹

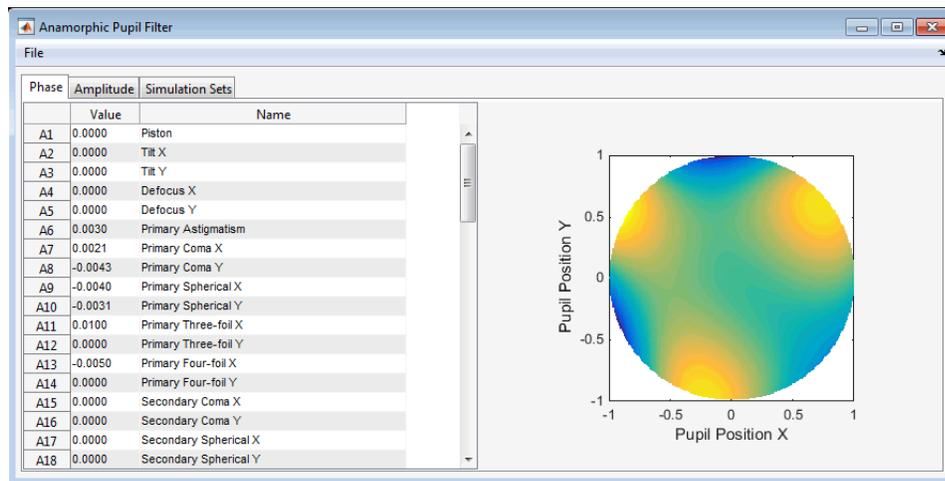


Figure 103. Screenshot of software written to generate anamorphic pupil filters.

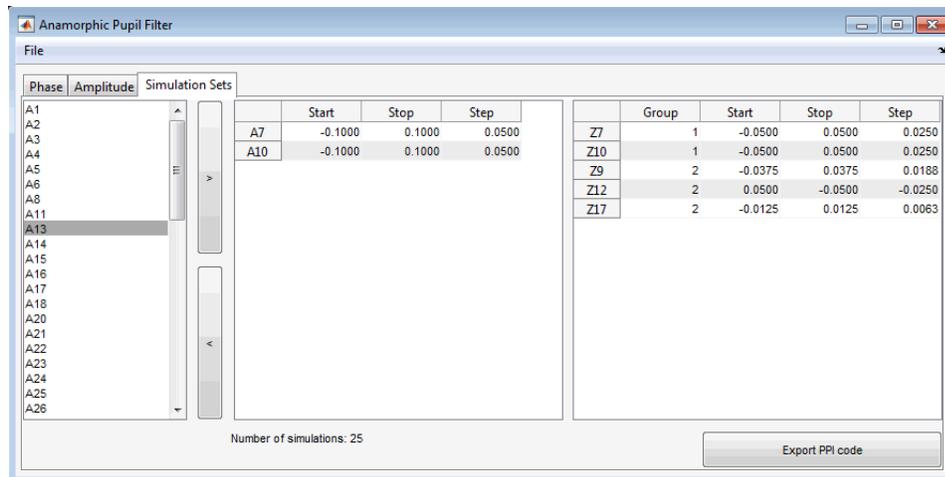


Figure 104. Screenshot of the same software showing the simulation set capabilities of the software.

7.4 Using Zernike polynomials to describe anamorphic aberrations

As noted in the previous section, it is sometimes acceptable to simulate anamorphic aberrations by way of Zernike polynomials. The wavefront should be composed in the anamorphic basis, then transformed into the Zernike basis. This is simply a mathematical technique to enable existing simulation software to carry out simulations with the new anamorphic basis. Carrying out simulations with Zernike polynomials directly would incorrectly describe the aberrations.

If the phase error in the pupil is perfectly known, using either the proposed basis or a Zernike basis will yield equivalent image simulations. However, the Zernike basis will provide the incorrect physical interpretation for correction. Therefore, using a Zernike basis is inadequate for designing, optimizing, and correcting an anamorphic optic. Moreover, if the phase errors are not known and are instead being measured experimentally the correct basis must be used. Interferometric techniques, which correlate Zernike pupil error with experimental measurements of the wavefront gradient, can be

easily adapted to use the proposed basis. Image-based techniques, which correlate Zernike pupil error with resist CD variation, can be similarly adapted.

7.5 Image-based metrology of anamorphic aberrations

The targets used for isomorphic systems are shown in Figure 105. The targets for astigmatism and spherical aberration are imaged through-focus, while the targets for coma and trefoil are imaged through exposure dose.

The effects of the terms in the proposed anamorphic aberration basis are similar to their isomorphic counterparts. Therefore, many of the metrology targets can be reused to interrogate the anamorphic aberration basis. Astigmatism 90° does not occur in anamorphic basis, so only 45° and 135° lines and spaces are necessary. The targets for coma remain unchanged. Spherical aberration is interrogated by a line of constant CD at various pitches, but in the anamorphic basis there are two kinds of spherical aberration. Therefore, we will need a set of spherical targets in both a vertical and horizontal orientation. Referring to Figure 97, the extrema of anamorphic three-foil and four-foil are very close together. This presents a difficulty when using a partially coherent source to image the targets.

To investigate this further, a full factorial experiment was run with anamorphic three-foil (A_{11} and A_{12}) and four-foil (A_{13} and A_{14}). The existing targets for three-foil were simulated through focus and through aerial image threshold. Each aberration was run at three levels with a design unit of $100 m\lambda$, focus and threshold were run at five levels each with design units of $25 nm$ and 0.05 , respectively. The results of ANOVA for this DOE are given in Table 21. A_{11} - A_{14} can be interrogated by measuring the existing T-bar targets both through focus and through exposure.

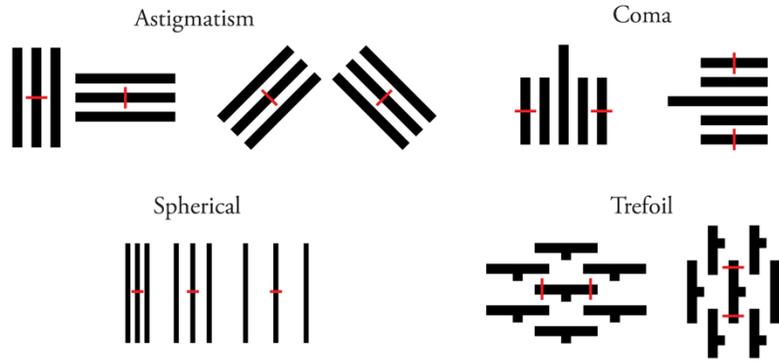


Figure 105. The binary targets used to interrogate aberrations both anamorphic and isomorphic aberrations.

Table 21. Results of ANOVA for DOE to determine the sensitivity of T-Bar targets to A₁₁–A₁₄. Fields marked with an X were found to be statistically significant.

Term	T-Bar CD_L-	T-Bar CD_T-
A11	X	X
A12	X	X
A13		X
A14	X	X
A11*Focus	X	
A12*Focus	X	
A13*Focus		X
A14*Focus	X	
A11*Threshold	X	X
A12*Threshold	X	X
A13*Threshold		X
A14*Threshold	X	X

8. CONCLUSIONS

State-of-the-art diffraction limited lithography processes have the tightest aberration tolerances of any optic. However, the aberrations of in-use tools cannot be easily monitored using the standard aberration retrieval methods. The proposed method allows one to measure both third- and fifth-order pupil variation from CD-SEM micrographs of images formed in resist. The pupil variation can either be in amplitude, phase, or a mixture of the two.

EUV lithography systems have aberration tolerances measured in picometers. However, in terms of wavelength this is a looser restriction than their DUV counterparts. This makes EUV lithography systems a better candidate for image-based techniques than state-of-the-art DUV systems because the image process variation caused by aberrations can be easily measured as CD variation. In addition, roughly 30% of the light at each reflective multilayer interface is lost to heat. This leads to thermal aberrations that vary throughout tool use. This aspect of EUV lithography systems means that it is imperative that aberrations can be measured during tool use.

It has been shown both theoretically and experimentally that the best basis to use to express amplitude pupil variation is the Zernike polynomials. Amplitude pupil variation was measured experimentally by interrogating individual pupil samples.

Aberration retrieval is achieved by using one-dimensional and two-dimensional binary targets which are sensitive to a specific type of aberration at multiple pitches, in combination with an algorithm based in-part on principal component analysis. This

approach removes the computationally expensive simulations from outside of the iterative loop, thus allowing rapid pupil characterization. Once models are built, the pupil can be characterized in under one second. Additionally, a new method of measuring spherical aberration was presented. This removed the necessity of subjective analysis from the aberration retrieval scheme.

This approach is very flexible and has been demonstrated on three different kinds of tools: 1) an actinic EUV mask microscope, 2) a production lithography scanner, and 3) a catoptric mask inspection tool.

Finally, a new basis was proposed to express the primary aberrations of anamorphic lithography systems. This basis is derived from first reducing the Zernike polynomials to one dimension, then reorthogonalizing a product of one-dimensional aberrations over the unit circle. Zernike polynomials are used for isomorphic optics because both the optics and the Zernike basis are rotationally invariant. Additionally, they accurately reproduce the primary aberrations achieved by expanding the wavefront error in a power series. We found the analogous power series for the reflection invariant anamorphic systems, and then constructed a new basis which reproduces these aberrations. This new basis can be ordered using a scheme analogous to the Wyant Zernike scheme. This enables direct comparison between the two bases. Anamorphic EUVL systems will have a central obscuration in the objective lens, so results were also computed for several different annular pupils.

9. REFERENCES

- [1] Wolf, S., [Microchip Manufacturing], Lattice Press, Sunset Beach, Calif. (2003).
- [2] Lécuyer, C., Brock, D. C. and Last, J., [Makers of the Microchip: A Documentary History of Fairchild Semiconductor], The MIT Press, Cambridge, Mass (2010).
- [3] Moore, G. E., “Cramming more components onto integrated circuits,” *Electronics* **38**(8), 114 (1965).
- [4] Streetman, B. and Banerjee, S., [Solid State Electronic Devices, 6th Edition, 6th edition], Prentice Hall of India (2005).
- [5] Neamen, D., [An Introduction to Semiconductor Devices], McGraw-Hill Education, Boston (2005).
- [6] Smith, B. W., “Optics for Photolithography,” [Microlithography: Science and Technology], B. W. Smith and K. Suzuki, Eds., CRC Press (2007).
- [7] Smith, B. W., “Resist Processing,” [Microlithography: Science and Technology], B. W. Smith and K. Suzuki, Eds., CRC Press (2007).
- [8] Mack, C. A., [Inside Prolith: A Comprehensive Guide to Optical Lithography Simulation for the PROLITH Family of Lithography Simulation Tools, v5.0, 2nd Rep edition], Finle Technologies Inc, Austin, Tex. (1997).
- [9] Rayleigh, Lord., “On the Theory of Optical Images, with special reference to the Microscope,” *Journal of the Royal Microscopical Society* **23**(4), 474–482 (1903).
- [10] Bruning, J. H., “Optical lithography: 40 years and holding,” *Proc. SPIE* **6520**, 652004–652004–652013 (2007).
- [11] Kamon, K., Miyamoto, T., Myoi, Y., Nagata, H., Kotani, N. and Tanaka, M., “Photolithography System Using a Combination of Modified Illumination and Phase Shift Mask,” *Jpn. J. Appl. Phys.* **31**(12S), 4131 (1992).
- [12] ITRS 2.0., “More Moore Roadmap–ITRS 2.0 White Paper” (2015).
- [13] Raley, A., Thibaut, S., Mohanty, N., Subhadeep, K., Nakamura, S., Ko, A., O’Meara, D., Tapily, K., Consiglio, S. and Biolsi, P., “Self-aligned quadruple patterning integration using spacer on spacer pitch splitting at the resist level for sub-32nm pitch applications,” *Proc. SPIE* **9782**, 97820F–97820F – 14 (2016).
- [14] Bakshi, V., ed., [EUV Lithography], SPIE, 1000 20th Street, Bellingham, WA 98227-0010 USA (2008).
- [15] Fujimoto, J., Abe, T., Tanaka, S., Ohta, T., Hori, T., Yanagida, T., Nakarai, H. and Mizoguchi, H., “Laser-produced plasma-based extreme-ultraviolet light source technology for high-volume manufacturing extreme-ultraviolet lithography,” *J. Micro/Nanolith. MEMS MOEMS* **11**(2), 021111–1 (2012).
- [16] Burbine, A., Levinson, Z., Schepis, A. and Smith, B. W., “Study of angular effects for optical systems into the EUV,” presented at SPIE, 17 April 2014, 90482N.
- [17] Scholze, F., Tu“mmmler, J., Gullikson, E. and Aquila, A., “Comparison of extreme ultraviolet reflectance measurements,” *J. Micro/Nanolith. MEMS MOEMS* **2**(3), 233–235 (2003).

- [18] Kritsun, O., La Fontaine, B., Hao, Y., Li, J., Wood, O., Raghunathan, S., Brunner, T., Koay, C.-S. and Mizuno, H., “Characterization of a 0.25NA full-field EUV exposure tool,” Proc. SPIE **7271**, 727121-727121-11 (2009).
- [19] Tennant, D. M., Bjorkholm, J. E., Eichner, L., Freeman, R. R., Jewell, T. E., MacDowell, A. A., Pastalan, J. Z., Szeto, L. H., Waskiewicz, W. K., White, D. L., Windt, D. L. and Wood II, O. R., “Comparison of reflective mask technologies for soft x-ray projection lithography,” Proc. SPIE **1604**, 91–104 (1992).
- [20] “Multilayer Reflectivity.”, <http://henke.lbl.gov/optical_constants/multi2.html> (18 June 2016).
- [21] Song, H., Zavyalova, L., Su, I., Shiely, J. and Schmoeller, T., “Shadowing effect modeling and compensation for EUV lithography,” Proc. SPIE **7969**, 796910-796910 – 9 (2011).
- [22] Raghunathan, S., McIntyre, G., Fenger, G. and Wood, O., “Mask 3D effects and compensation for high NA EUV lithography,” Proc. SPIE **8679**, 867918-867918–13 (2013).
- [23] Migura, S., Kneer, B., Neumann, J. T., Kaiser, W. and van Schoot, J., “Anamorphic high-NA EUV lithography optics,” presented at SPIE, 2015, 96610T-96610T – 9.
- [24] Kingslake, R., “Anamorphic Lenses,” Journal of Photography and Motion Pictures of the George Eastman House **5**(9), 208–213 (1956).
- [25] Chretien, H., “Anamorphic lens system and method of making the same,” US1962892 A (1934).
- [26] Born, M. and Wolf, E., [Principles of Optics, 7th ed.], Press Syndicate of the University of Cambridge (1999).
- [27] Hecht, E., [Optics, 4 edition], Addison-Wesley, Reading, Mass (2001).
- [28] Mahajan, V. N., [Optical Imaging and Aberrations, Part III: Wavefront Analysis], SPIE Press, Bellingham, Wash (2013).
- [29] van Ingen Schenau, K., Bottiglieri, G., van Schoot, J., Neumann, J.-T. and Roesch, M., “Imaging performance of the EUV high NA anamorphic system,” presented at SPIE, 2015, 96610S-96610S – 10.
- [30] van Schoot, J., van Setten, E., Rispen, G., Troost, K. Z., Kneer, B., Migura, S., Neumann, J. T. and Kaiser, W., “High-numerical aperture extreme ultraviolet scanner for 8-nm lithography and beyond,” Journal of Micro Nanolithography, MEMS, and MOEMS **16**(4), 041010 (2017).
- [31] Gaskill, J., [Linear Systems, Fourier Transforms, and Optics], John Wiley & Sons (1978).
- [32] O’Neill, E. L. and Physics., [Introduction to Statistical Optics], Dover Publications, Mineola, N.Y (2004).
- [33] “FUJINON LENS XF56mmF1.2 R APD.”, Fujifilm Global, <http://www.fujifilm.com/products/digital_cameras/x/fujinon_lens_xf56mmf12_r_apd/> (30 November 2016).
- [34] Quanz, S. P., Meyer, M. R., Kenworthy, M. A., Girard, J. H. V., Kasper, M., Anne-Marie Lagrange, Apai, D., Boccaletti, A., Bonnefoy, M., Chauvin, G., Hinz, P. M. and Lenzen, R., “First Results from Very Large Telescope NACO Apodizing Phase Plate: 4 μm Images of The Exoplanet β Pictoris b,” ApJL **722**(1), L49 (2010).
- [35] Brown, J., [Microwave Lenses], Methuen & Co (1953).

- [36] Maréchal, A., Thesis **University of Paris** (1948).
- [37] Maréchal, A. and Françon, M., [Diffraction, Structure des Images, Influence des Aberrations] (1960).
- [38] Rayleigh, J. W. S. B., [Scientific Papers: 1869-1881], Cambridge University Press (1899).
- [39] “ITRS 2013 Edition - Lithography Tables.”, (2013).
- [40] Wurm, S. and Gwyn, C., “EUV Lithography,” [Microlithography: Science and Technology], B. W. Smith and K. Suzuki, Eds., CRC Press (2007).
- [41] Jackson, J. D., [Classical Electrodynamics Third Edition, 3 edition], Wiley, New York (1998).
- [42] Harvey, J. E., Schröder, S., Choi, N. and Duparré, A., “Total integrated scatter from surfaces with arbitrary roughness, correlation widths, and incident angles,” *Opt. Eng* **51**(1), 013402–1 (2012).
- [43] Milster, T. D. and Beaudry, N. A., “Scattering and coherence in EUVL,” presented at SPIE, 1998, 537–543.
- [44] Lai, K., Wu, C. J. and Proglor, C. J., “Scattered light: the increasing problem for 193-nm exposure tools and beyond,” *Proc. SPIE* **4346**, 1424–1435 (2001).
- [45] Nijboer, B., “The Diffraction Theory of Aberrations,” Thesis (1942).
- [46] Hamilton, W. R., “On Some Results Of The View Of A Characteristic Function In Optics,” *British Association Report*, 360–370 (1833).
- [47] Smith, W. J., [Modern Optical Engineering, 4th Ed., 4 edition], McGraw-Hill Education, New York (2007).
- [48] Pedrotti, F. L., Pedrotti, L. M. and Pedrotti, L. S., [Introduction to Optics, 3 edition], Pearson, Upper Saddle River, N.J (2006).
- [49] Voelz, D. G., [Computational Fourier Optics: A MATLAB Tutorial], SPIE Press, Bellingham, Wash (2011).
- [50] Easton Jr., R. L., [Fourier Methods in Imaging, 1 edition], Wiley, Chichester, West Sussex (2010).
- [51] George B. Parrent, Jr., “Basic Theory of Partial Coherence,” presented at AFIPS ’66 (Spring), 1966, Proceedings of the April 26-28, 1966, Spring joint computer conference.
- [52] Underwood, J. H., Gullikson, E. M., Koike, M., Batson, P. J., Denham, P. E., Franck, K. D., Tackaberry, R. E. and Steele, W. F., “Calibration and standards beamline 6.3.2 at the Advanced Light Source,” *Review of Scientific Instruments* **67**(9), 3372–3372 (1996).
- [53] Hopkins, H. h., “Calculation of the Aberrations and Image Assessment for a General Optical System,” *Optica Acta: International Journal of Optics* **28**(5), 667–714 (1981).
- [54] Emmett J. Ientilucci., “Using the Singular Value Decomposition” (2003).
- [55] Jackson, J. D. and Mathematics., [Mathematics for Quantum Mechanics: An Introductory Survey of Operators, Eigenvalues, and Linear Vector Spaces, Dover Ed edition], Dover Publications, Mineola, N.Y (2006).
- [56] Kiusalaas, J., [Numerical Methods in Engineering with Python, 2 edition], Cambridge University Press (2014).
- [57] Shlens, J., “A Tutorial on Principal Component Analysis,” <http://arxiv.org/abs/1404.1100> (2014).

- [58] Smith, Lindsay I., “A tutorial on Principal Components Analysis” (2002).
- [59] Lian, Y. and Zhou, X., “Fast and accurate computation of partially coherent imaging by stacked pupil shift operator,” *Proc. SPIE* **7488**, 74883G-74883G – 10 (2009).
- [60] Yamazoe, K., “Computation theory of partially coherent imaging by stacked pupil shift matrix,” *Journal of the Optical Society of America A* **25**(12), 3111 (2008).
- [61] Cobb, N., “Sum of Coherent Systems Decomposition by SVD” (1995).
- [62] van de Kerkhof, M. A., de Boeij, W., Kok, H., Silova, M., Baselmans, J. and Hemerik, M., “Full optical column characterization of DUV lithographic projection tools,” *Proc. SPIE* **5377**, 1960–1970 (2004).
- [63] Dirksen, P., Juffermans, C. A., Pellens, R. J., Maenhoudt, M. and De Bisschop, P., “Novel aberration monitor for optical lithography,” *Microlithography’99*, 77–86 (1999).
- [64] Dirksen, P., Juffermans, C. A., Engelen, A., De Bisschop, P. and Muellerke, H., “Impact of high-order aberrations on the performance of the aberration monitor,” *Microlithography 2000*, 9–17 (2000).
- [65] Khotanzad, A. and Huga Hong, Y., “Invariant Image Recognition by Zernike Moments,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*. **12**(5), 489 (1990).
- [66] Lipschutz, S., Spiegel, M. and Liu, J., [Schaum’s Outline of Mathematical Handbook of Formulas and Tables, 4th Edition: 2, 400 Formulas + Tables, 4th ed.], McGraw-Hill Education (2012).
- [67] Barnes, S. L., “A Technique for Maximizing Details in Numerical Weather Map Analysis,” *J. Appl. Meteor.* **3**(4), 396–409 (1964).
- [68] Barnes, S., “Mesoscale objective map analysis using weighted time-series observations,” NOAA Technical Memorandum **ERL NSSL-62**.
- [69] Manuel Guizar., [Efficient subpixel image registration by cross-correlation] (2008).
- [70] Naulleau, P. P., Goldberg, K. A., Batson, P., Bokor, J., Denham, P. and Rekawa, S., “Fourier-synthesis custom-coherence illuminator for extreme ultraviolet microfield lithography,” *Applied Optics* **42**(5), 820 (2003).
- [71] Levinson, Z., Smith, J., Fenger, G. and Smith, B. W., “An automated image-based tool for pupil plane characterization of EUVL tools,” presented at SPIE, 2016, 97762M.
- [72] KLA Tencor., “PROLITH,” Version X5\1, 2015, <<http://www.kla-tencor.com/lithography-modeling/chip-prolith.html>>.
- [73] Burbine, A., Fryer, D. and Sturtevant, J., “Akaike information criterion to select well-fit resist models,” presented at SPIE, 2015, 94270J-94270J – 7.
- [74] Feldmann, H., Ruoff, J., Harnisch, W. and Kaiser, W., “Actinic review of EUV masks,” *Proc. SPIE* **7636**, 76361C-76361C – 11 (2010).
- [75] “ISO 14999-4:2015 - Optics and photonics - Interferometric measurement of optical elements and optical systems - Part 4: Interpretation and evaluation of tolerances specified in ISO 10110.”, <<https://webstore.ansi.org/Standards/ISO/ISO149992015>> (28 October 2018).
- [76] Krist, J. E. and Burrows, C. J., “Phase-retrieval analysis of pre- and post-repair Hubble Space Telescope images,” *Appl. Opt.*, **AO** **34**(22), 4951–4964 (1995).
- [77] Barakat, R. and Riseberg, L., “Diffraction Theory of the Aberrations of a Slit Aperture,” *J. Opt. Soc. Am.*, **JOSA** **55**(7), 878–881 (1965).

- [78] Chanan, G. and Troy, M., "Strehl ratio and modulation transfer function for segmented mirror telescopes as functions of segment phase error," *Appl. Opt.*, AO **38**(31), 6642–6647 (1999).
- [79] Rayces, J. L., "Least-squares fitting of orthogonal polynomials to the wave-aberration function," *Appl. Opt.*, AO **31**(13), 2223–2228 (1992).
- [80] Wynne, C. G., "The Primary Aberrations of Anamorphic Lens Systems," *Proc. Phys. Soc. B* **67**(7), 529 (1954).
- [81] Mahajan, V. N., "Orthonormal aberration polynomials for anamorphic optical imaging systems with circular pupils," *Applied Optics* **51**(18), 4087 (2012).
- [82] Welford, W. T., [Aberrations of Optical Systems, 1 edition], CRC Press, Bristol, Eng. ; Philadelphia (1986).
- [83] Dodoc, A., Bannert, C., Blahnik, V. and Sehr, H., "Anamorphic objective," US20130022345 A1 (2013).
- [84] ZEMAX., "OpticStudio," Version 16.5, 2017, <<http://www.zemax.com/os/opticstudio>>.
- [85] Schmidt, E., "Zur Theorie der linearen und nichtlinearen Integralgleichungen.," *Mathematische Annalen* **65**, 370–399 (1908).
- [86] Dai, G. and Mahajan, V. N., "Nonrecursive determination of orthonormal polynomials with matrix formulation," *Optics Letters* **32**(1), 74 (2007).
- [87] Gram, J. P., "Ueber die Entwicklung reeller Functionen in Reihen mittelst der Methode der kleinsten Quadrate.," *Journal für die reine und angewandte Mathematik (Crelle's Journal)* **1883**(94), 41–73 (2009).
- [88] Wyant, J. C. and Creath, K., "Basic Wavefront Aberration Theory for Optical Metrology," presented at *Applied Optics and Optical Engineering*, Volume XI, 1992, 2.
- [89] Levinson, Z., "Anamorphic Pupil Filters," Version 1.0, 2017, <<http://www.zaclevinson.com/SPIE2017>>.

10. APPENDIX A: ANNULAR ANAMORPHIC POLYNOMIALS

This section provides more details on computing the proposed anamorphic aberration basis. The aberration basis can be computed using the Gram-Schmidt process, or by a matrix formulation proposed by Mahajan.⁸⁶ Both methods are presented here in the Mathematica programming language. The same thing can be done in MATLAB, but Mathematica provides two specific benefits here: 1) it is better for obtaining analytic solutions, and 2) the Mathematica code is easier to understand than MATLAB code.

To obtain the anamorphic circle polynomials from the Legendre polynomials using the matrix method:

```
MaxPoly = 70;
LOrder := {{0, 0}, {1, 0}, {0, 1}, {2, 0}, {1, 1}, {0, 2},
{3, 0}, {2, 1}, {1, 2}, {0, 3}, {4, 0}, {3, 1}, {2, 2}, {1, 3},
{0, 4}, {5, 0}, {4, 1}, {3, 2}, {2, 3}, {1, 4}, {0, 5}, {6, 0},
{5, 1}, {4, 2}, {3, 3}, {2, 4}, {1, 5}, {0, 6}, {7, 0}, {6, 1},
{5, 2}, {4, 3}, {3, 4}, {2, 5}, {1, 6}, {0, 7}, {8, 0}, {7, 1},
{6, 2}, {5, 3}, {4, 4}, {3, 5}, {2, 6}, {1, 7}, {0, 8}, {9, 0},
{8, 1}, {7, 2}, {6, 3}, {5, 4}, {4, 5}, {3, 6}, {2, 7}, {1, 8},
{0, 9}, {10, 0}, {9, 1}, {8, 2}, {7, 3}, {6, 4}, {5, 5}, {4, 6},
{3, 7}, {2, 8}, {1, 9}, {0, 10}, {11, 0}, {10, 1}, {9, 2}, {8,
3}, {7, 4}, {6, 5}, {5, 6}, {4, 7}, {3, 8}, {2, 9}, {1, 10},
{0, 11}, {12, 0}, {11, 1}, {10, 2}, {9, 3}, {8, 4}, {7, 5},
{6, 6}, {5, 7}, {4, 8}, {3, 9}, {2, 10}, {1, 11}, {0, 12},
{13, 0}, {12, 1}, {11, 2}, {10, 3}, {9, 4}, {8, 5}, {7, 6},
{6, 7}, {5, 8}}

L2D[j_] :=
LegendreP[LOrder[[j]][[1]], u]*LegendreP[LOrder[[j]][[2]], v]

AreaFactor = Pi;
Cpp[m_, n_] := Integrate[Integrate[ L2D[m]*L2D[n], {v, -Sqrt[1 -
u^2], Sqrt[1 - u^2]}], {u, -1, 1}, Assumptions -> u \[Element]
Reals && v \[Element] Reals]/AreaFactor
```

```

Q = Assuming[u \[Element] Reals && v \[Element] Reals,
CholeskyDecomposition[Array[Cop, {MaxPoly, MaxPoly}]]];
M = Assuming[u \[Element] Reals && v \[Element] Reals,
Inverse[Transpose[Q]]];
A[J_] := Simplify[Sum[M[[J]][[j]]*L2D[j], {j, 1, J}], u
\[Element] Reals && v \[Element] Reals]

n[i_] := (LOrder[[i]][[1]] + LOrder[[i]][[2]])
m[i_] := (LOrder[[i]][[1]] - LOrder[[i]][[2]])
mprime[i_] := -Abs[m[i]] + n[i] + Mod[n[i], 2]
nprime[i_] := (n[i] + mprime[i])/2
Acircle = Array[A, {MaxPoly}];
TableForm[Table[{i, nprime[i], mprime[i], Acircle[[i]]}, {i, 1,
MaxPoly}]]

```

To obtain the annular anamorphic polynomials from circle polynomials using the Gram-Schmidt method:

```

MaxPoly = 28;
F = {{1}};

PupilIntegrate [f_] := PupilIntegrate[f] =
Integrate[
Integrate[f, {v, -Sqrt[1 - u^2], Sqrt[1 - u^2]}], {u, -1, 1},
Assumptions -> u \[Element] Reals && v \[Element] Reals] -
Integrate[ Integrate[
f, {v, -Sqrt[\[Epsilon]^2 - u^2],
Sqrt[\[Epsilon]^2 - u^2]}], {u, -\[Epsilon], \[Epsilon]},
Assumptions -> \[Epsilon] >= 0 && \[Epsilon] <
1 && \[Epsilon] \[Element] Reals && u \[Element] Reals &&
v \[Element] Reals]
AreaFactor = Pi*(1 - \[Epsilon]^2);
c[j_, k_] := c[j, k] = -1/AreaFactor *
PupilIntegrate[Acircle[[j]]*F[[k]]]

Gj[j_] := Gj[j] = Sum[c[j, k]*F[[k]], {k, 1, j - 1}] +
Acircle[[j]]

Fj[j_] := AppendTo[F, Simplify[Gj[j]/
Sqrt[PupilIntegrate[Gj[j]^2]/AreaFactor], \[Epsilon] >=
0 && \[Epsilon] < 1 && \[Epsilon] \[Element] Reals,
TimeConstraint -> 10000]]

For[i=2,i<MaxPoly+1,i++,Fj[i]; Print[F[[i]]]]

```

The anamorphic circle polynomials could also be computed in Mathematica using the Gram-Schmidt method. Many of the computations in the matrix method can be

parallelized, but can become very resource intensive. This is because it involves a matrix inversion, so computing higher order terms can become impossible using this technique. The matrix inversion could be replaced with a least squares regression for the appropriate column vector from the identity matrix, but this is also computationally limiting.

The Gram-Schmidt process offers an alternative to the resource intensity of the matrix method, but this comes at the sake of computation time. The 28th annular anamorphic polynomial, for example, took around 12 hours to compute on a iMac 3.1 GHz Core i7 with 16 GB RAM. Of course all of this can be avoided by just computing the terms numerically. In this case the Gram-Schmidt process should not be used due to its numerical instability.

11. APPENDIX B: ANNULAR ANAMORPHIC POLYNOMIAL TABLES

This section provides the annular anamorphic polynomials computed over several annular pupils. All solutions are computed from the results given in Table 20. This includes the results for $\epsilon = 0$, which reproduce the results given in Table 19.

Table 22. Annular anamorphic polynomials for $\epsilon = 0.0$.

N	n'	m'	Anamorphic Circle Polynomial ($A_n(u, v; \epsilon = 0)$)
1	0	0	1
2	1	1	$2u$
3	1	1	$2v$
4	1	0	$4u^2 - 1$
5	1	0	$-\sqrt{2}(-u^2 - 3v^2 + 1)$
6	2	2	$2\sqrt{6}uv$
7	2	1	$-4u(1 - 2u^2)$
8	2	1	$4v(3u^2 + 5v^2 - 3) / \sqrt{5}$
9	2	0	$16u^4 - 12u^2 + 1$
10	2	0	$\sqrt{2/7}(3u^4 + 6(5v^2 - 1)u^2 + 35v^4 - 30v^2 + 3)$
11	3	3	$-4v(1 - 6u^2) / \sqrt{5}$
12	3	3	$4u(u^2 + 3v^2 - 1)$
13	3	2	$2\sqrt{2}u(8u^2 - 3)v$
14	3	2	$4\sqrt{2}uv(3u^2 + 5v^2 - 3)$
15	3	1	$32u^5 - 32u^3 + 6u$
16	3	1	$2\sqrt{2/21}v(15u^4 + 10(7v^2 - 3)u^2 + 63v^4 - 70v^2 + 15)$
17	3	0	$64u^6 - 80u^4 + 24u^2 - 1$
18	3	0	$2(5u^6 + 15(7v^2 - 1)u^4 + 15(21v^4 - 14v^2 + 1)u^2 + 231v^6 - 315v^4 + 105v^2 - 5) / \sqrt{33}$

Table 23. Annular anamorphic polynomials for $\epsilon = 0.05$.

N	n'	m'	Anamorphic Circle Polynomial ($A_n(u, v; \epsilon = 0.05)$)
1	0	0	1
2	1	1	$1.9975u$
3	1	1	$1.9975v$
4	1	0	$-0.999997(1.0025 - 4v^2)$
5	1	0	$-1.41598(-1.01001u^2 - 3.00002v^2 + 1.00501)$
6	2	2	$4.89285uv$
7	2	1	$-3.98508u(1.00251 - 2.005u^2)$
8	2	1	$1.77775v(3.01511u^2 + 5.02504v^2 - 3.01506)$
9	2	0	$0.997494(16.0401u^4 - 12.0451u^2 + 0.0025(5 - 6v^2) + 1.)$
10	2	0	$0.532133(3.22669u^4 + 6(5.07556v^2 - 1.04538)u^2 + 35.2267v^4 - 30.2723v^2 + 3.06818)$
11	3	3	$-1.78217(1.00251 - 6.015u^2)v$
12	3	3	$3.97522u(1.00504u^2 + 3.01496v^2 - 1.00502)$
13	3	2	$2.81783u(8.02005u^2 - 3.00752)v$
14	3	2	$5.62157uv(3.01506u^2 + 5.02509v^2 - 3.01506)$
15	3	1	$0.98641(32.4012u^5 - 32.4017u^3 + 6.075u)$
16	3	1	$0.604311v(15.3086u^4 + 10(7.14174v^2 - 3.06115)u^2 + 64.2689v^4 - 71.4118v^2 + 15.3035)$
17	3	0	$0.987554(64.806u^6 - 81.0835u^4 + 12(0.0025(13 - 5v^2) + 2.)u^2 - 0.005(15v^4 - 18v^2 + 7) - 1)$
18	3	0	$0.340381(6.8871u^6 + 15(7.48349v^2 - 1.22344)u^4 + 15(21.731v^4 - 14.659v^2 + 1.12072)u^2 + 236.882v^6 - 323.654v^4 + 108.402v^2 - 5.29512)$

Table 24. Annular anamorphic polynomials for $\epsilon = 0.10$.

N	n'	m'	Anamorphic Circle Polynomial ($A_n(u, v; \epsilon = 0.1)$)
1	0	0	1
2	1	1	$1.99007u$
3	1	1	$1.99007v$
4	1	0	$-0.99995(1.01 - 4u^2)$
5	1	0	$-1.42127(-1.0401u^2 - 3.0003v^2 + 1.0202)$
6	2	2	$4.87443uv$
7	2	1	$-3.94132u(1.0101 - 2.02u^2)$
8	2	1	$1.7454v(3.06181u^2 + 5.10061v^2 - 3.06091)$
9	2	0	$0.989905(16.1616u^4 - 12.1818u^2 + 0.01(5 - 6v^2) + 1.00001)$
10	2	0	$0.524939(3.92735u^4 + 6(5.30913v^2 - 1.1861)u^2 + 35.9273v^4 - 31.1166v^2 + 3.28106)$
11	3	3	$-1.76233(1.0101 - 6.06u^2)v$
12	3	3	$3.90343u(1.0206u^2 + 3.05941v^2 - 1.0203)$
13	3	2	$2.78611u(8.0808u^2 - 3.0303)v$
14	3	2	$5.51655uv(3.06092u^2 + 5.1015v^2 - 3.06091)$
15	3	1	$0.947465(33.6194u^5 - 33.6267u^3 + 6.3u)$
16	3	1	$0.567998v(16.3402u^4 + 10(7.58843v^2 - 3.25882)u^2 + 68.185v^4 - 75.7922v^2 + 16.2572)$
17	3	0	$0.950874(67.2972u^6 - 84.4372u^4 + 12(0.01(13 - 5v^2) + 2.)u^2 - 0.02(15v^4 - 18v^2 + 7) - 1)$
18	3	0	$0.317925(13.1734u^6 + 15(9.07266v^2 - 1.96887)u^4 + 15(24.0929v^4 - 16.813v^2 + 1.52382)u^2 + 255.642v^6 - 351.396v^4 + 119.417v^2 - 6.278)$

Table 25. Annular anamorphic polynomials for $\epsilon = 0.15$.

N	n'	m'	Anamorphic Circle Polynomial ($A_n(u, v; \epsilon = 0.15)$)
1	0	0	1
2	1	1	$1.977787u$
3	1	1	$1.977787v$
4	1	0	$-0.999747(1.0225 - 4u^2)$
5	1	0	$-1.43004(-1.09051u^2 - 3.00152v^2 + 1.04602)$
6	2	2	$4.84358uv$
7	2	1	$-3.87147u(1.02301 - 2.045u^2)$
8	2	1	$1.69439v(3.14418u^2 + 5.22815v^2 - 3.13966)$
9	2	0	$0.977049(16.3681u^4 - 12.4142u^2 + 0.0225(5 - 6v^2) - 0.0030375(v^2 - 1) + 1.00006)$
10	2	0	$0.512894(5.16569u^4 + 6(5.72207v^2 - 1.43657)u^2 + 37.1653v^4 - 32.6194v^2 + 3.66513)$
11	3	3	$-1.73003(1.02301 - 6.135u^2)v$
12	3	3	$3.79171u(1.04806u^2 + 3.13203v^2 - 1.04655)$
13	3	2	$2.73358u(8.18405u^2 - 3.06905)v$
14	3	2	$5.34459uv(3.13983u^2 + 5.2326v^2 - 3.1397)$
15	3	1	$0.888034(35.6994u^5 - 35.737u^3 + 6(1.1125 - 0.00151875(4v^2 - 5))u)$
16	3	1	$0.514419v(18.4337u^4 + 10(8.40877v^2 - 3.63871)u^2 + 75.098v^4 - 83.6074v^2 + 18.0011)$
17	3	0	$0.89205(71.7002u^6 - 90.3843u^4 + 12(0.0225(13 - 5v^2) + 2)u^2 - 0.045(15v^4 - 18v^2 + 7) - 0.00151875(50v^4 - 64v^2 + 23) - 1)$
18	3	0	$0.283372(25.9749u^6 + 15(12.2389v^2 - 3.49285)u^4 + 15(28.6547v^4 - 21.065v^2 + 2.35079)u^2 + 290.955v^6 - 404.16v^4 + 140.803v^2 - 8.29125)$

Table 26. Annular anamorphic polynomials for $\epsilon = 0.20$.

N	n'	m'	Anamorphic Circle Polynomial ($A_n(u, v; \epsilon = 0.20)$)
1	0	0	1
2	1	1	$1.96116u$
3	1	1	$1.96116v$
4	1	0	$-0.999201(1.04 - 4u^2)$
5	1	0	$-1.44227(-1.1616u^2 - 3.0048v^2 + 1.08326)$
6	2	2	$4.80015uv$
7	2	1	$-3.77963u(1.0416 - 2.08u^2)$
8	2	1	$1.6285v(3.26919u^2 + 5.41025v^2 - 3.25499)$
9	2	0	$0.958704(16.6656u^4 - 12.7496u^2 + 0.04(5 - 6v^2) - 0.0096(v^2 - 1) + 1.00032)$
10	2	0	$0.496038(7.05492u^4 + 6(6.35256v^2 - 1.82285)u^2 + 39.0529v^4 - 34.937v^2 + 4.26982)$
11	3	3	$-1.6863(1.0416 - 6.24u^2)v$
12	3	3	$3.65008u(1.08973u^2 + 3.23079v^2 - 1.085)$
13	3	2	$2.66084u(8.3328u^2 - 3.12499)v$
14	3	2	$5.11107uv(3.25596u^2 + 5.42404v^2 - 3.2552)$
15	3	1	$0.814553(38.7199u^5 - 38.8415u^3 + 6(1.2 - 0.0048(4v^2 - 5))u)$
16	3	1	$0.451122v(22.2314u^4 + 10(9.73275v^2 - 4.28812)u^2 + 85.6682v^4 - 95.7451v^2 + 20.8055)$
17	3	0	$0.814774(78.4172u^6 - 99.5082u^4 + 12(0.0016(46 - 25v^2) + 0.04(13 - 5v^2) + 2.)u^2 - 0.08(15v^4 - 18v^2 + 7) - 0.0048(50v^4 - 64v^2 + 23) - 1)$
18	3	0	$0.240804(49.7448u^6 + 15(17.9784v^2 - 6.34377)u^4 + 15(36.5972v^4 - 28.697v^2 + 3.91071)u^2 + 350.205v^6 - 494.139v^4 + 178.422v^2 - 12.1055)$

Table 27. Annular anamorphic polynomials for $\epsilon = 0.25$.

N	n'	m'	Anamorphic Circle Polynomial ($A_n(u, v; \epsilon = 0.25)$)
1	0	0	1
2	1	1	$1.94029u$
3	1	1	$1.94029v$
4	1	0	$-0.998053(1.0625 - 4u^2)$
5	1	0	$-1.45793(-1.25391u^2 - 3.01172v^2 + 1.13306)$
6	2	2	$4.744uv$
7	2	1	$-3.67058u(1.06641 - 2.125u^2)$
8	2	1	$1.55197v(3.44682u^2 + 5.65096v^2 - 3.41245)$
9	2	0	$0.934717(17.0625u^4 - 13.1982u^2 + 0.0625(5 - 6v^2) - 0.0234375(v^2 - 1) + 1.00124)$
10	2	0	$0.474625(9.76948u^4 + 6(7.2598v^2 - 2.38621)u^2 + 41.7616v^4 - 38.3169v^2 + 5.17634)$
11	3	3	$-1.63243(1.06641 - 6.375u^2)v$
12	3	3	$3.48967u(1.14894u^2 + 3.35307v^2 - 1.13748)$
13	3	2	$2.5689u(8.53125u^2 - 3.19995)v$
14	3	2	$4.82447uv(3.41616u^2 + 5.68383v^2 - 3.41328)$
15	3	1	$0.733463(42.7993u^5 - 43.1057u^3 + 6(1.31251 - 0.0117188(4v^2 - 5))u)$
16	3	1	$0.384936v(28.8152u^4 + 10(11.779v^2 - 5.35581)u^2 + 101.003v^4 - 113.714v^2 + 25.1401)$
17	3	0	$0.724281(88.0676u^6 - 112.729u^4 + 12(0.00390625(46 - 25v^2) + 0.0625(13 - 5v^2) + 2.)u^2 - 0.125(15v^4 - 18v^2 + 7) - 0.0117188(50v^4 - 64v^2 + 23) - 1)$
18	3	0	$0.195137(93.1699u^6 + 15(28.2369v^2 - 11.6133)u^4 + 15(50.1786v^4 - 42.2366v^2 + 6.836)u^2 + 446.972v^6 - 644.43v^4 + 243.886v^2 - 19.3593)$

Table 28. Annular anamorphic polynomials for $\epsilon = 0.30$.

N	n'	m'	Anamorphic Circle Polynomial ($A_n(u, v; \epsilon = 0.30)$)
1	0	1	
2	1	1	$1.91565u$
3	1	1	$1.91565v$
4	1	0	$-0.995974(1.09 - 4u^2)$
5	1	0	$-1.47707(-1.3681u^2 - 3.0243v^2 + 1.19693)$
6	2	2	$4.67503uv$
7	2	1	$-3.54914u(1.0981 - 2.18u^2)$
8	2	1	$1.46896v(3.69037u^2 + 5.95622v^2 - 3.61987)$
9	2	0	$0.905122(17.5696u^4 - 13.7745u^2 + 0.09(5 - 6v^2) - 0.0486(v^2 - 1) + 1.00371)$
10	2	0	$0.44922(13.5641u^4 + 6(8.53049v^2 - 3.18893)u^2 + 45.5406v^4 - 43.1318v^2 + 6.51264)$
11	3	3	$-1.56986(1.0981 - 6.54u^2)v$
12	3	3	$3.32102u(1.23012u^2 + 3.49597v^2 - 1.20662)$
13	3	2	$2.45934u(8.7848u^2 - 3.29649)v$
14	3	2	$4.49645uv(3.63102u^2 + 6.02255v^2 - 3.62249)$
15	3	1	$0.650147(48.1059u^5 - 48.7666u^3 + 6(1.45009 - 0.0243(4v^2 - 5))u)$
16	3	1	$0.320907v(39.9438u^4 + 10(14.903v^2 - 7.08768)u^2 + 122.892v^4 - 140.v^2 + 31.7998)$
17	3	0	$0.626799(101.554u^6 - 131.428u^4 + 12(0.0081(46 - 25v^2) + 0.09(13 - 5v^2) + 2.)u^2 - 0.18(15v^4 - 18v^2 + 7) - 0.0243(50v^4 - 64v^2 + 23) - 1)$
18	3	0	$0.151069(173.181u^6 + 15(46.792v^2 - 21.4786)u^4 + 15(73.6876v^4 - 66.6562v^2 + 12.4278)u^2 + 606.089v^6 - 898.755v^4 + 360.31v^2 - 33.573)$

Table 29. Annular anamorphic polynomials for $\epsilon = 0.35$.

N	n'	m'	Anamorphic Circle Polynomial ($A_n(u, v; \epsilon = 0.35)$)
1	0	0	1
2	1	1	$1.88772u$
3	1	1	$1.88772v$
4	1	0	$-0.99258(1.1225 - 4u^2)$
5	1	0	$-1.49988(-1.50501u^2 - 3.04502v^2 + 1.27685)$
6	2	2	$4.59334uv$
7	2	1	$-3.41967u(1.13751 - 2.245u^2)$
8	2	1	$1.38312v(4.01682u^2 + 6.33455v^2 - 3.88803)$
9	2	0	$0.870224(18.2001u^4 - 14.4972u^2 + 0.1225(5 - 6v^2) - 0.0900375(v^2 - 1) + 1.00942)$
10	2	0	$0.420738(18.8036u^4 + 6(10.2887v^2 - 4.32333)u^2 + 50.7439v^4 - 49.9339v^2 + 8.47741)$
11	3	3	$-1.50006(1.13751 - 6.735u^2)v$
12	3	3	$3.15307u(1.33894u^2 + 3.65667v^2 - 1.29601)$
13	3	2	$2.33425u(9.10005u^2 - 3.41803)v$
14	3	2	$4.14099uv(3.91638u^2 + 6.45376v^2 - 3.89509)$
15	3	1	$0.568528(54.8711u^5 - 56.155u^3 + 6(0.00183827(19 - 16v^2) - 0.0450187(4u^2 - 5) + 1.61293)u)$
16	3	1	$0.262215v(58.4545u^4 + 10(19.6774v^2 - 9.88859)u^2 + 154.209v^4 - 178.684v^2 + 42.1327)$
17	3	0	$0.52861(120.166u^6 - 157.644u^4 + 12(0.00183827(104 - 75v^2) + 0.0150063(46 - 25v^2) + 0.1225(13 - 5v^2) + 2.)u^2 + 0.00183827(-210v^4 + 408v^2 - 200) - 0.245(15v^4 - 18v^2 + 7) - 0.0450187(50v^4 - 64v^2 + 23) - 1)$
18	3	0	$0.112099(323.496u^6 + 15(81.0977v^2 - 40.3835)u^4 + 15(115.41v^4 - 111.935v^2 + 23.4317)u^2 + 873.97v^6 - 1341.94v^4 + 574.97v^2 - 62.5277)$