Extreme ultraviolet lithography: A few more pieces of the puzzle

by

Christopher Neil Anderson

B.S. (University of California, Davis) 2004
M.S. (University of Arizona) 2005

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Engineering - Applied Science & Technology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor David Attwood, Chair
Dr. Patrick Naulleau
Professor Andrew Neureuther
Professor Avideh Zahkor

Spring 2009

Extreme ultraviolet lithography: A few more pieces of the puzzle

by

Christopher Neil Anderson

Abstract

Extreme ultraviolet lithography: A few more pieces of the puzzle

by

Christopher Neil Anderson

Doctor of Philosophy in Applied Science & Technology

University of California, Berkeley

Professor David Attwood, Chair

The work described in this dissertation has improved three essential components of extreme ultraviolet (EUV) lithography: exposure tools, photoresist, and metrology.

Exposure tools. A field-averaging illumination stage is presented that enables non-uniform, high-coherence sources to be used in applications where highly uniform illumination is required. In an EUV implementation, it is shown that the illuminator achieves a 6.5% peak-to-valley intensity variation across the entire design field of view. In addition, a design for a stand-alone EUV printing tool capable of delivering 15 nm half-pitch sinusoidal fringes with available sources, gratings and nano-positioning stages is presented. It is shown that the proposed design delivers a near zero line-edge-rougness (LER) aerial image, something extremely attractive for the application of resist testing.

Photoresist. Two new methods of quantifying the deprotection blur of EUV photoresists are described and experimentally demonstrated. The deprotection blur, LER, and sensitivity parameters of several EUV photoresists are quantified simultaneously as base weight percent, photoacid generator (PAG) weight percent, and post-exposure bake (PEB) temperature are varied. Two surprising results are found: 1) changing base weight percent does not significantly affect the deprotection blur of EUV photoresist, and 2) increasing PAG weight percent can simultaneously reduce LER and E-size in EUV photoresist. The latter result motivates the development of an EUV exposure statistics model that includes the effects of photon shot noise, the PAG spatial distribution, and the changing of the PAG distribution during the exposure. In addition, a shot noise + deprotection blur model is used to show that as deprotection

blur becomes large relative to the size of the printed feature, LER reduction from improved counting statistics becomes dominated by an increase in LER due to reduced deprotection contrast.

Metrology. Finally, this dissertation describes MOSAIC, a new wavefront metrology that enables complete wavefront recovery from print or aerial image based measurements. This new technique, based on measuring the local focal length of the optic at sampled positions in the pupil, recovers the curvature of the aberration and uses the curvature to recover the aberration itself. In a modeled EUV implementation, MOSAIC is shown to recover the SEMATECH Berkeley MET wavefront with a 4.2% RMS error: a $4\times$ improvement over the reported errors of the original lateral shearing interferometry wavefront measurement.

Dedicated to Mom, Dad, and Patrick Naulleau

Acknowledgements

I would like to thank all of the members of my dissertation committee for having been kind enough to read this thesis and serve on on my qualifying examination board. I am greatly indebted to Paul Denham, Brian Hoef, and Gideon Jones of the Center for X-Ray Optics (CXRO) at Lawrence Berkeley National Laboratory, who's nine-to-five is reflected in every piece of experimental data within this thesis. I am also indebted to the entire CXRO engineering team for building and maintaining the EUV exposure tool. I am eternally grateful to David Attwood for making all of this possible, and for so much more. I want to thank Rene Delano for giving me a reason to wear the "clydesdale's." And I thank Ken Goldberg for setting the bar high.

I am forever thankful to my mom for her endless support, infinite love, and for being one of my best friends in my adult years. I am forever thankful to my dad for teaching me to keep things simple, to be resourceful, and to be flexible. I am also grateful to the rest of my family for all of their encouragement and support. And to my friends: you make the juice worth the squeeze.

And finally, there is Patrick Naulleau. I could write pages, but I know I don't need to. To say it succinctly, you give us all something to aspire to. The research in this thesis was made possible by financial support from SEMATECH and the National Science Foundation EUV Engineering Research Center. This work was also supported by the Director, Office of Science, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

Contents

Chapter 1

Introduction

Contents

1.1 Photolithography

In early electronic devices, connections between components of an electric circuit were done by hand, with wire and solder (see Figure 1.1[i] top), a technique referred to as "point-to-point" wiring. Today's electronics utilize a far superior manufacturing process: photolithography. The basic difference between point-to-point and photolithographic wiring techniques can be

[i]These are pictures of the Moonlight 2W tube amplifier that I built in October 2005

summarized succinctly: instead of manually connecting wires to a network of pre-installed components (point-to-point, Figure 1.1 top), we now connect components to a network of pre-laid wires (photolithography, Figure 1.1 bottom).

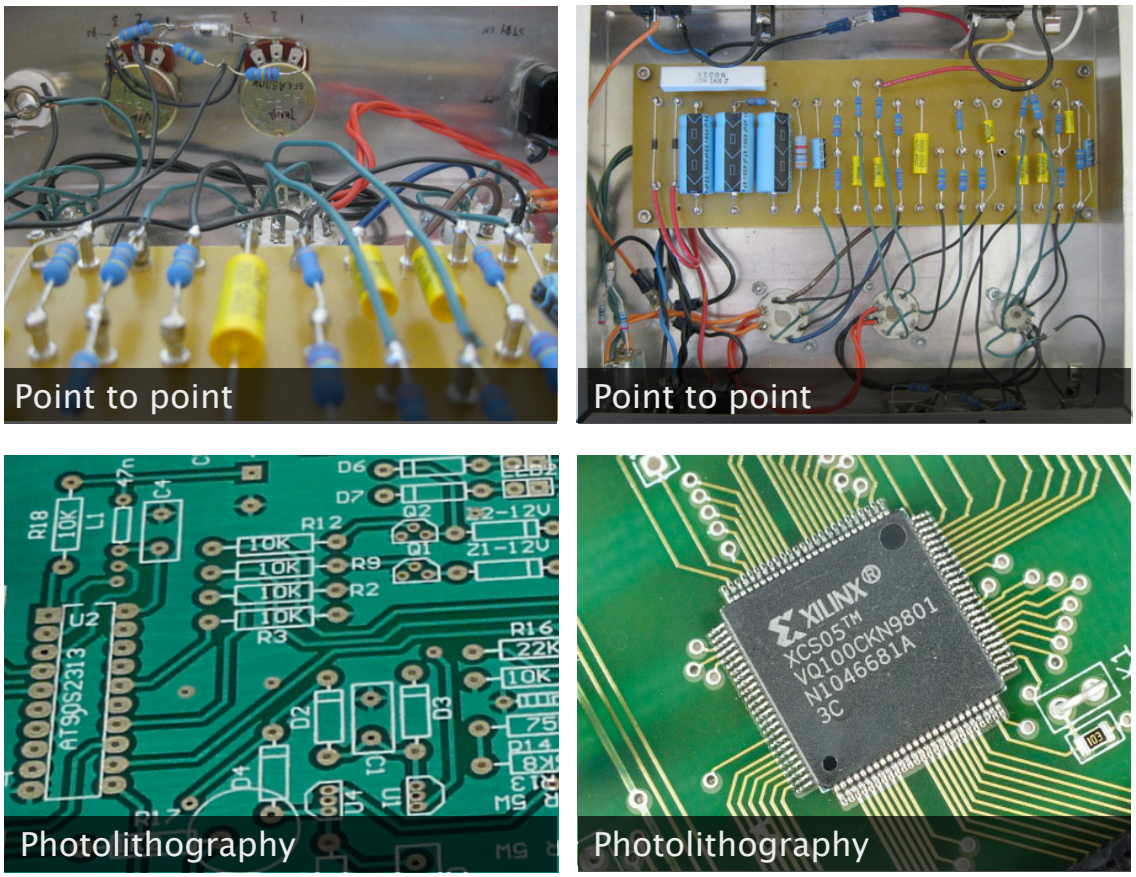

Figure 1.1: Top: point-to-point wiring - lots of wires, lots of time, lots of clutter. Bottom: photolithographic wiring. In the bottom left image, the white outlines show where electrical components will be "dropped in" to the pre-laid wiring network.

1.1.1 "Printing" is a misnomer

Photolithography is usually defined as a "printing" process. For example, circuit boards made with photolithography - ones with a pre-laid wire network - are called printed circuit boards (PCBs). The reality is that term "print" is fairly misleading in this context. When most people think about a printing process they think about a computer printer: a scanning "tiny

dot maker" that draws hundreds of thousands of little dots on a piece of paper - one at a time - to eventually create a picture. But this is not at all how a PCB is made.

1.1.2 Photolithography is an imaging technique

PCB's are made using an imaging system. This technique is illustrated in Figure 1.2 and is described below; think "camera." At the top of the Figure 1.2 is a light bulb, followed by an object called the mask. The mask is a scaled replica of the wiring pattern you wish to have on the circuit board. It's job is to block light where you want to have a wire and transmit light everywhere else. Below the mask is a lens that images (and de-magnifies) the mask pattern to the circuit board (Figure 1.3a), which has been coated with a layer of light sensitive material (film) called photoresist (Figure 1.3b).

When the light is turned on (Figure 1.3c), the lens "takes a picture" of the mask - the transparent regions of the mask show up as bright areas on the film (circuit board), exposing the film in these regions. The exposed areas of the film experience a chemical reaction rendering them dissolvable to a developer chemical. When the circuit board is developed the film is washed away in the regions that were exposed (Figure 1.3d). The underlying thin metal layer is now revealed in the places that were exposed. Next, the entire circuit board is rinsed in another chemical that dissolves all of the exposed regions of the metal layer (Figure 1.3e), this process is called "etching". Finally, another chemical is used to dissolve the thin layer of film off of the remaining geometrical patterns of metal + film (Figure 1.3f). The resulting structure is a thin layer of metal shaped like the geometrical patterns that were designed in to the mask at the beginning of the process.

1.1.3 A killer application

So far we have described photolithography as a process used to efficiently structure conductor material into "pre-laid wires on a board," or a PCB. The PCB is what the market calls a "better mouse trap" application of photolithography - it accomplishes something that could be done before (wiring a circuit) except better, faster, and cheaper.

The real value of photolithography, however, is in a different application: structuring *semi*conductor material. Instead of making tiny wires, you make tiny transistors (replace

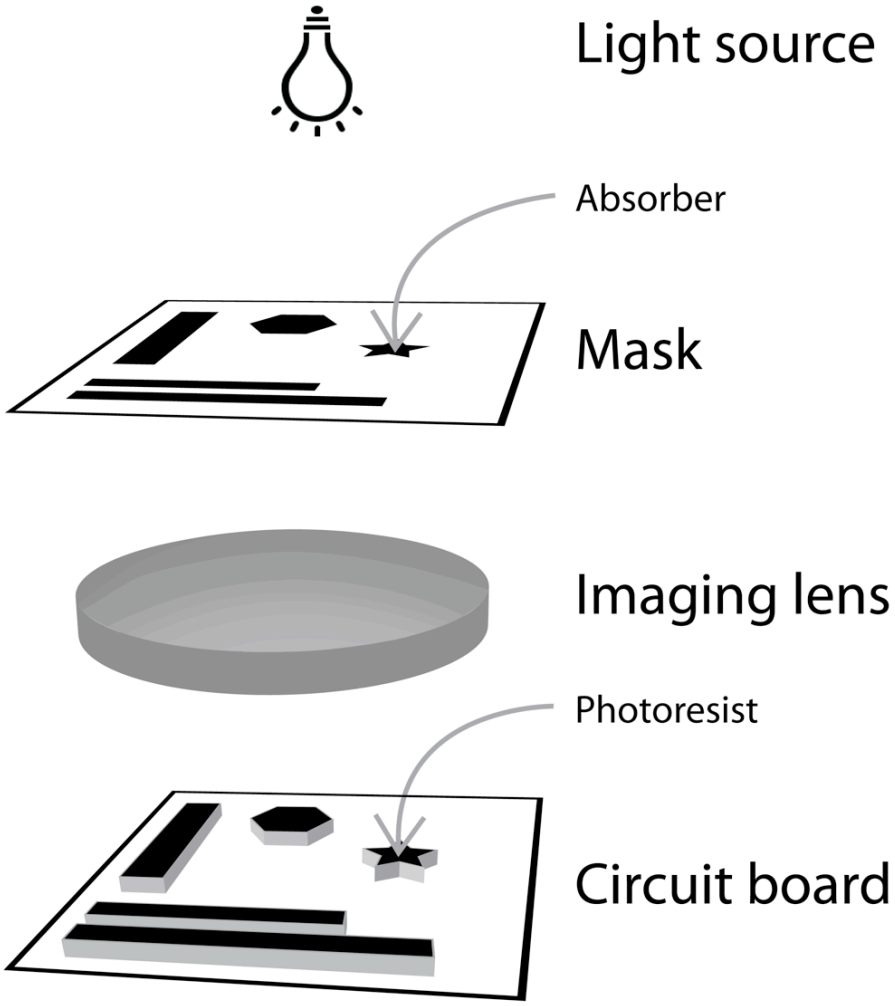

Figure 1.2: Schematic of a photolithographic imaging system.

the vacuum tube), tiny diodes, and digital memory. Instead of making a PCB, you make integrated circuits (ICs), central processing units (CPUs), and digital memory banks. In effect, this "killer" application of photolithography - one that enables something never before possible - created the digital electronics industry.

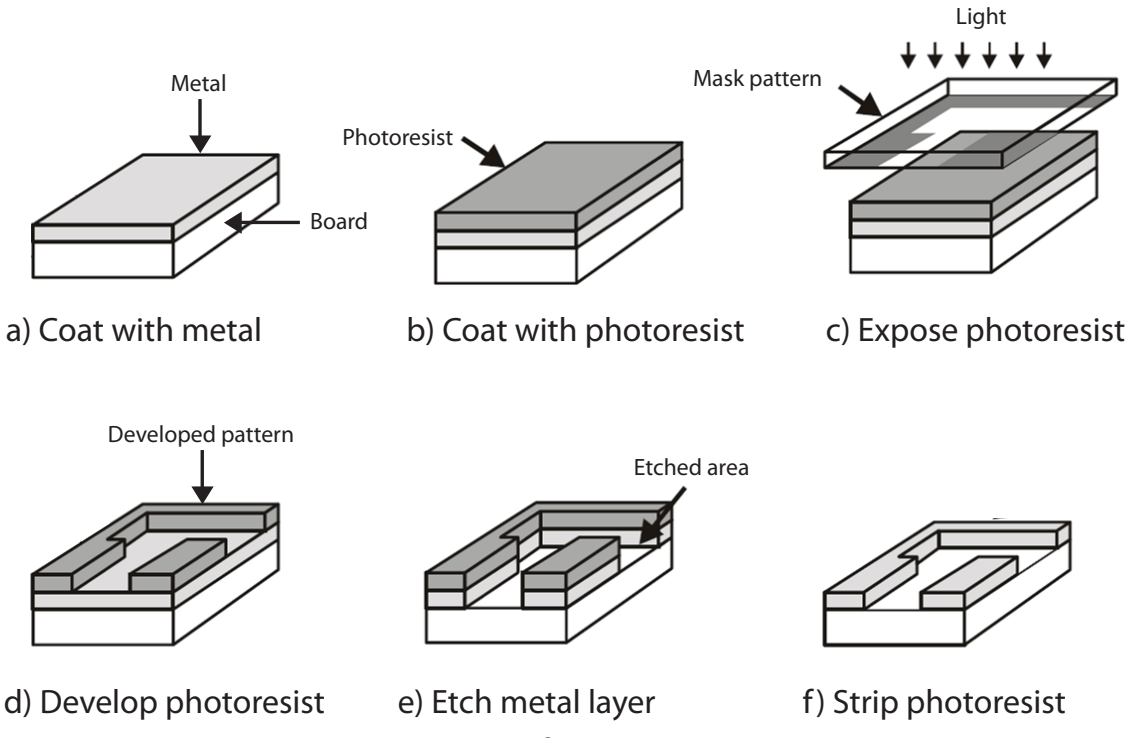

Figure 1.3: The photolithography process.

1.2 The digital electronics industry

1.2.1 Progress comes from incremental improvements

The continued success of the digital electronics industry is largely due to incremental improvements in photolithography. In 1990, Intel's 486 CPU was manufactured using a 1 micron lithography process. Last year, Intel's Core 2 Duo CPU was manufactured using a 45 nm process: a 500× improvement in the number of transistors on a chip, and a 500× boost in speed and capability.

1.2.2 Shorter wavelengths make better chips

When the wavelength of light used in an imaging system shrinks, the resolution of the system improves. As a result, the development of shorter wavelength sources has been intimately linked to the development of improved lithography systems that achieve better resolution

capabilities. For example, the wavelength of light used for lithography has evolved from 365 nm in 1987, to 248 nm in 1992, to 193 nm in 1999.

1.2.3 The next wavelength: 13.5 nm

The industry is currently evaluating the feasibility of its next proposed wavelength shift: from 193 nm, or deep ultraviolet (DUV), to 13.5 nm, or extreme ultraviolet (EUV), slated to improve the manufacturing node from a 45 nm double exposure process to a 22 nm single exposure process. Before its implementation, the industry must must overcome several major development hurdles including: EUV mask defects, EUV source power, and EUV photoresist, and several accessory challenges: optical contamination, and optical metrology, and reticle protection.

1.3 Dissertation overview

This dissertation research encompasses three critical aspects of the EUV lithography commercialization effort: exposure tools, photoresist, and metrology. In Chapter 3 an upgrade to the illuminator of the SEMATECH Berkeley EUV exposure tool is presented. A new uniformity stage utilizing field-averaging "fly's eye" optics is shown to achieve a 6.5% peak-to-valley intensity variation across the tool's entire field of view; a significant improvement over the previous illumination conditions.

In Chapter 4 a feasibility study of a stand-alone incoherent EUV interference lithography (IL) tool is presented. Analytical expressions for grating alignment tolerances are derived and grating quality issues are analyzed with a two-dimensional propagation model. It is shown that the IL tool can deliver 15 nm half-pitch sinusoidal fringes using available sources, gratings and alignment stages. In addition, the unique low line-edge-rougness (LER) aerial image of the incoherent IL tool described.

Chapter 5 begins the investigation of EUV photoresist. Two new methods of quantifying the deprotection blur[ii] of EUV photoresists are described and experimentally demonstrated.

[ii]The average width of the volume of resist polymer that is rendered dissolvable (and thus removed during development) by a single photo-generated acid

Experimental error sources are identified and a sensitivity analysis of all error sources is presented. Repeated blur extractions on three different resists are shown to produce error bars within those predicted by the sensitivity analysis.

In Chapter 6, the deprotection blur, LER, and sensitivity of several EUV resist platforms are quantified as resist constituents and process parameters are varied. Three independent attributes are studied: base weight percent, photoacid generator (PAG) weight percent, and post-exposure bake temperature. An in depth discussion of these experiments and their ties EUV LER is presented. An EUV exposure statistics model that includes the effects of photon shot noise, the PAG spatial distribution, and the changing of the PAG distribution during the exposure is described. In addition, a shot noise + deprotection blur model is used to show that as deprotection blur becomes large relative to the size of the printed feature, LER reduction from improved counting statistics becomes dominated by an increase in LER due to reduced deprotection contrast.

Finally, In Chapter 7 a new print-based wavefront metrology called MOSAIC is described. The method, based on mapping out the local focal length of a lens at sampled points in the pupil, recovers the the curvature of the aberration and uses the curvature to recover the aberration itself. Analytical expressions for obtaining the curvature from image data are presented. A least squares Gram-Schmidt process that recovers any function from samples of its curvature is described. Lastly, the subtleties of MOSAIC and its EUV implementation are described through a discussion that ties together the concepts of depth of focus, illumination spatial coherence, object structure, and optical aberrations.

Chapter 2

Background

Contents

2.1 The SEMATECH Berkeley MET

2.1.1 Overview

The SEMATECH Berkeley extreme ultraviolet (EUV) microfield exposure tool (MET) has been described in detail in the literature [1, 2, 3] and is only summarized here. The MET is a reflective, two-element, 5×-reduction optical system with a reflective reticle (mask). To support reflective masks with this on-axis system, the Scheimpflug principle is imposed tilting the mask by 4° and the wafer by 0.8°. The MET has an annular pupil with a central obscuration radius equal to 30% of the full pupil. The field of view is 1 mm × 3 mm at

the reticle plane (200 μm $\times$ 600 μm at the wafer plane). Figure 2.1 shows a CAD model depicting the major components of the Berkeley MET exposure station as well as the EUV beam path.

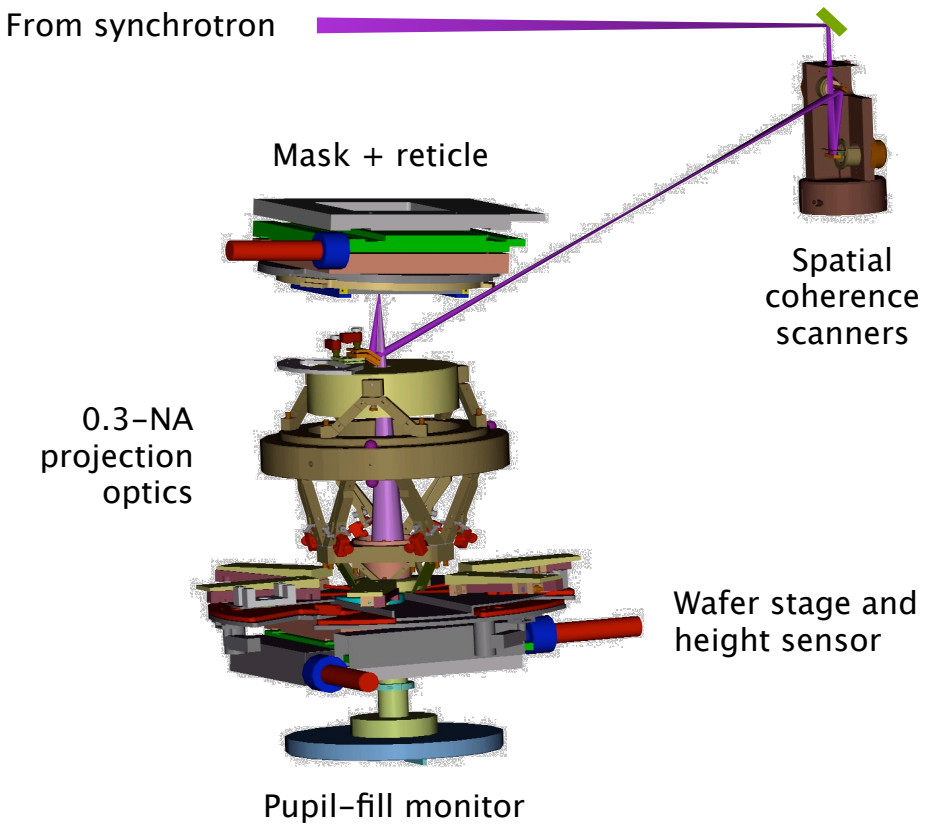

Figure 2.1: CAD model of major components and beam path for the SEMATECH Berkeley MET.

2.1.2 Programmable pupil fill

The Berkeley MET is situated on undulator beamline 12.0.3 of the Advanced Light Source at Lawrence Berkeley National Laboratory. Although the highly coherent undulator radiation [4] is ideal for interferometry it is not particularly well suited for direct use in lithography applications where such a source would result in σ values well below 0.1. Through the use of an active illumination system, however, this potential drawback can be readily turned into a significant benefit by enabling arbitrary pupil fill control in a virtually lossless manner.

Similar to the system depicted in Figure 2.2 (left), the MET illuminator uses a pair of scanning galvanometers to manipulate the angle that the impinging coherent undulator radiation strikes the mask [5]. By varying the illumination angle during the exposure, any desired pupil fill can be synthesized (Figure 2.2, right). The mutual incoherence of the various angles is ensured through the scanning process by virtue that none of the angles coexisting in time. Stationarity of the pupil fill across the field is ensured by re-imaging the scanning mirrors to the reticle. In addition, angular apodization is minimized by re-imaging the intrinsic source to pupil of the projection optic: each illumination angle on the mask corresponds to a single point in the pupil.

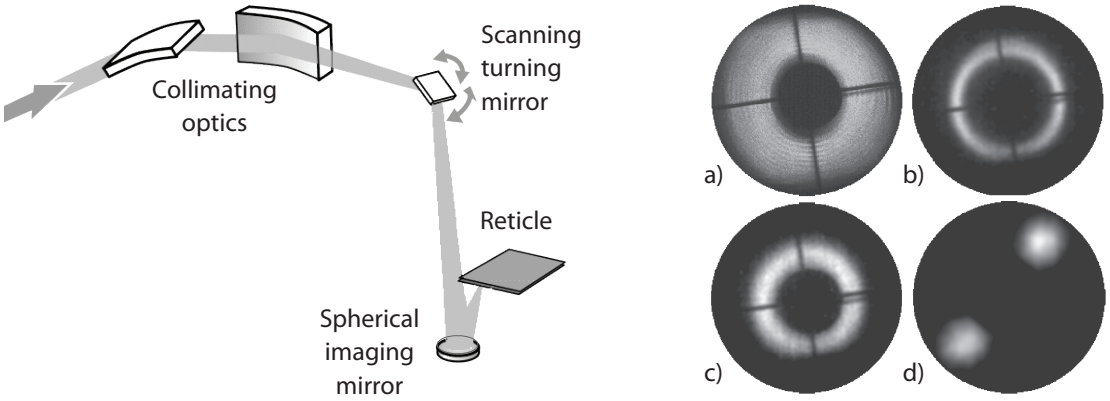

Figure 2.2: Left: schematic of the active illuminator used to synthesize arbitrary pupil fills. Right: a) 0.35 $< \sigma <$ 0.95 annular pupil fill. b) σ = 0.55 annular pupil fill. c) 0.35 $< \sigma <$ 0.55 annular pupil fill. d) 45° rotated dipole with σ = 0.1 poles. In the pupil-fill monitor images we also see the MET central obscuration as well as the arms used to support the direct-transmission-blocking baffles.

2.1.3 Resolution capabilities

With a numerical aperture of 0.3, the MET optic has a Rayleigh resolution (k_1 factor = 0.61) of 27 nm. Modified illuminations can push the k_1 factor as low as 0.25 yielding an ultimate resolution limit of 11.25 nm. The modeled contrast transfer function for three different illumination schemes is shown in Figure 2.3. With annular illumination (0.35 $< \sigma <$ 0.55) the resolution knee occurs at about 23 nm. With 45° dipole illumination, the resolution knee is pushed out to approximately 20 nm and the aerial-image contrast is generally enhanced.

Ultimate resolution on vertical lines can be achieved by using x-dipole illumination with an offset σ of 1, in which case the resolution knee is pushed down to 12.5 nm. The contrast dead band in the 20 to 35 nm range is caused by the central obscuration blocking the diffracted orders.

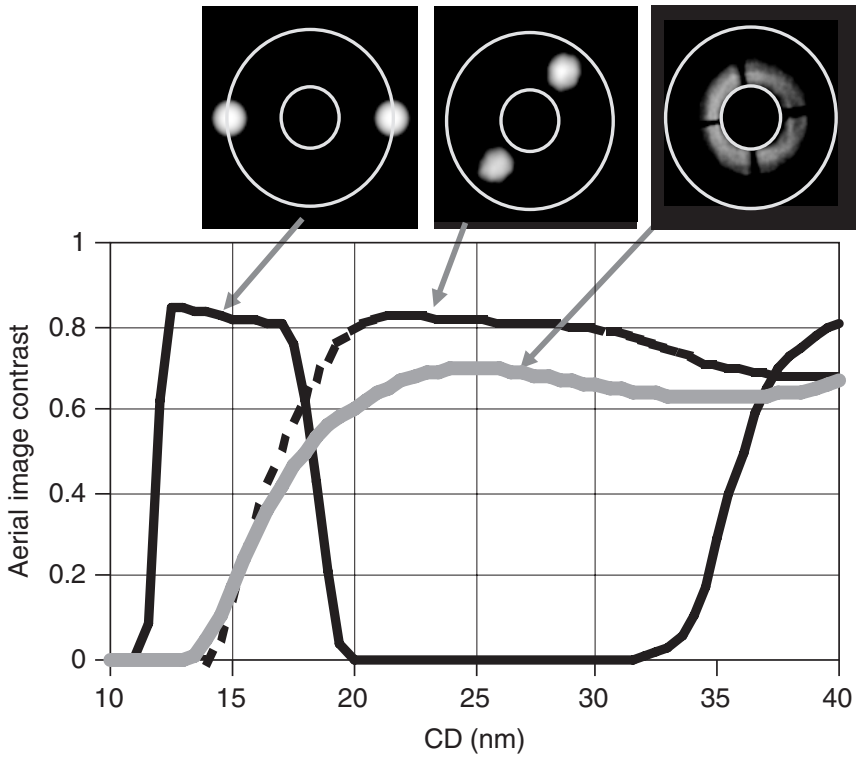

Figure 2.3: Modeled contrast transfer function of the SEMATECH Berkeley MET assuming three different illumination schemes: x-dipole with an offset of 1, 45° dipole, and $0.35 < \sigma < 0.55$ annular.

2.2 Photoresist models

2.2.1 The ideal binary model

During an exposure, the aerial image $I(x, y)$ illuminates a resist-coated wafer for an exposure time t delivering a spatially varying energy density $E(x, y) = I(x, y) \times t$. In the simple binary resist model, spatial locations that experience an energy density greater than the resist threshold T are rendered dissolvable by developer. As a result, lengthening or shortening the

exposure time changes the pattern that remains after development. Figure 2.4 (left) shows a mask and its resulting aerial image (1D). Figure 2.4 (right) shows the aerial energy densities for three different exposure times. As the exposure time increases, the exposed area that is above threshold increases, causing more resist to be removed during development.

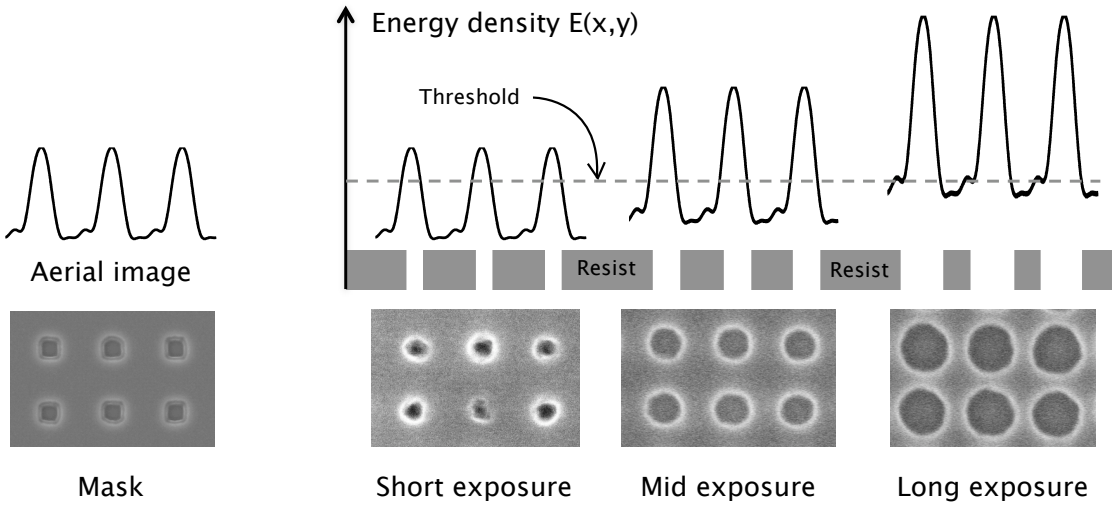

Figure 2.4: An overview of the lithographic patterning process highlighting the importance of exposure time (or dose) in making features print at a targeted size. In lithography, an optical system images the mask to the wafer. With the shutter open, photons are delivered to the resist. Resist that sees more photons than the threshold level are removed during development. Longer exposure times cause more areas to be above threshold causing features to print bigger.

2.2.2 A more complete model

While the above description captures high-level trends, it completely over simplifies the complex interactions that occur during the exposure, bake, and development processes. A more complete (but still high level) description is shown in Figure 2.5 (a) - (d). The first process step is the exposure: EUV photons ionize atoms in the resist and create cascades of secondary electrons at each photon absorption site (a). If photo acid generators (PAGs) are near a photon absorption site, the secondary electrons will most likely convert the PAGs to acids. The result is a discrete distribution of acids, one at each location where a PAG was activated (b). The second processing step is a post-exosure bake (PEB) that facilitates a catalyzed

reaction in which photoacids render the backbone polymer dissolvable to developer. This process, known as chemical amplification, is described very nicely in Harry J. Levenson's text *Principles of Lithography, 2nd Edition* [6]:

> The dissolution in developer of the backbone polymer of unexposed resist is prevented by protecting groups. Through a post-exposure bake, the acid can catalyze the cleaving of protecting groups from the backbone polymer. This cleaving reaction is catalytic, in the sense that the acid still remains after the reaction, and is therefore available to promote the cleaving of additional protecting groups. With prolonged heating, a single photoacid is able to diffuse and cleave multiple protecting groups. By these means a single exposure event is capable of producing multiple-deprotection events (gain), a process called chemical amplification.

The PEB processing step changes the discrete acid distribution to a continuous "deprotection profile" (c) . The third processing step is development, which removes resist in spatial locations with a deprotection higher than a specific threshold level (d).

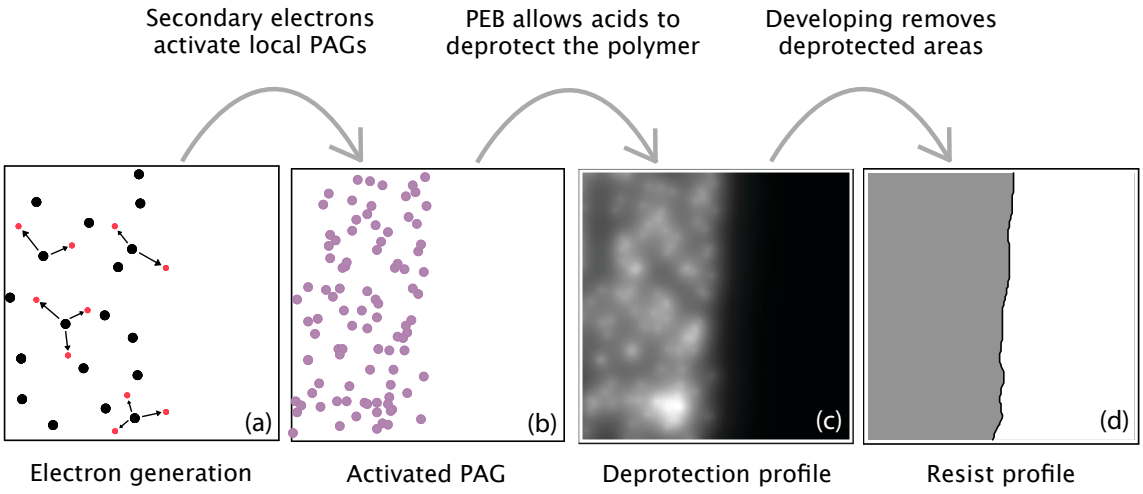

Figure 2.5: Illustration of the exposure, post-exposure bake, and development processes. Sub figures (b) - (d) are courtesy of G. Gallatin.

Chapter 3

MET illuminator upgrade

Contents

3.1 Introduction

Uniform illumination is critical to the performance of extreme ultraviolet microfield exposure tools. Prior to October 2006, daily changes in illumination uniformity often hampered the full-field printing capabilities of the SEMATECH Berkeley Microfield Exposure Tool (MET) at the Advanced Light Source at Lawrence Berkeley National Laboratory (LBNL). Figure 3.1 shows typical illumination conditions and the results of printing a clear field at a dose near the threshold of the photoresist [i]. With these illumination conditions, experiments such as full-field mask studies, aberrations measurements, and resist-profile analysis, which require a high degree of illumination uniformity over the entire field, were often difficult to

[i] To ensure we print close to the threshold dose of the photoresist we print a Focus-Exposure-Matrix (FEM) of clear fields and pick the dose which is just above the dose where nothing prints.

perform.

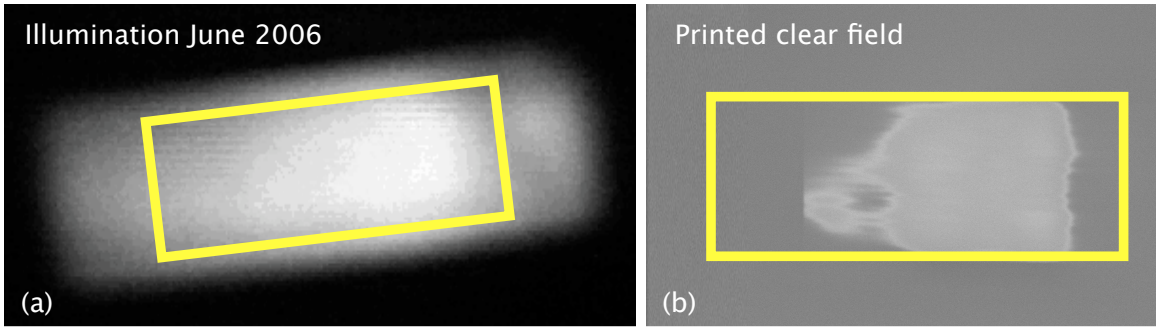

Figure 3.1: (a) Typical illumination field at reticle before the illuminator upgrade. Image is taken with a scintillator that converts EUV photons to visible photons. The outline shows the 1x3 mm object-side field. (b) Results of printing a clear field at a dose close to the threshold dose of the resist. Note that the printing results and scintillator image are not taken on the same day but do represent typical illumination conditions.

To alleviate these problems, a field-averaging uniformity stage was integrated into the existing MET illuminator [5]. In this chapter we present the new illuminator design, describe its implementation, and quantify its performance. The original design spec for the illumination uniformity of the SEMATECH Berkeley MET was to have less than 10% peak-to-valley intensity variation across the entire 200-μm x 600-μm wafer-side field of view. Before the upgrade, the uniformity spec was only met over a 100-μm x100-μm subset of the 200-μm x 600-μm wafer-side field. The upgraded illuminator is shown to satisfy the original design spec across the full wafer-side field, a dramatic improvement over the prior illumination conditions.

3.2 Field-averaging optics

At visible wavelengths, a particularly common method for increasing the effective uniformity of a light source is through the use of a field-averaging "flys eye" lens [7]. As depicted in Figure 3.2, in a reflection implementation, the field-averaging lens is an array of small lenslets that map neighboring sub-sections of an incoming illumination footprint to an overlap region wherein the expanded sub-fields are averaged to produce a highly uniform footprint.

In most implementations a fielding lens is also used, increasing efficiency by quickly redirecting the expanding sub-fields from different lenslets to a common overlap region. A typical arrangement is to place the field lens at the focus plane of the individual lenslets of the flys eye array. However if focus control is available before the field-averaging optic, the field lens can be eliminated by illuminating the field-averaging optic with a beam converging to the desired overlap plane.

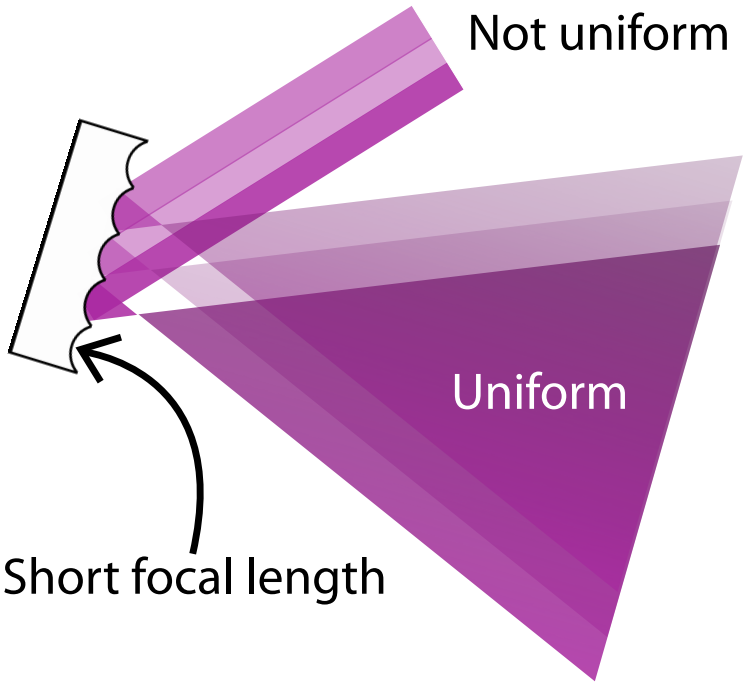

Figure 3.2: Schematic of a field-averaging optic operating in reflection mode. Non-uniformity of incoming beam is indicated by three different shades in the beam cross section. Individual lenslets of the field-averaging optic bring small subsections of the incoming beam quickly to focus and allow them to expand and overlap in the far-field.

3.3 Design and implementation at EUV

Integrating a field-averaging optic into the existing SEMATECH Berkeley MET illuminator required a unique implementation of the field-averaging concept. Operating in off-axis reflection mode, the astigmatic imaging requirements of integrating a single field averaging optic

into the existing illuminator would require the individual lenslet surfaces in the lenslet array to be toroidal. As toroidal surfaces are very difficult to fabricate, we chose to decouple the single field-averaging optic into a pair of cylindrical lenslet arrays, one for each orthogonal imaging direction. Adequate sampling of the incoming footprint is achieved by designing each cylinder-lens array to contain at least five cylindrical lenslets within the incoming footprint width in each imaging direction. The upstream Kirkpatric-Baez (KB) mirrors are used in a fielding configuration, illuminating each cylinder-lens array with a beam converging to the desired far-field overlap plane. This ensures that the footprints of the expanding sub-fields emerging from each individual lenslet spatially overlap completely in the far-field overlap plane and no light is wasted. Integration into the existing system required that far-field uniformity planes in x and y coincide with surfaces of the x and y Fourier-synthesis scanning mirrors [5].

Figure 3.3 shows a schematic of the upgraded illuminator. The two new x and y cylinder-lens arrays are called M7x and M7y, respectively and the existing x and y Fourier-synthesis scanning mirrors are called M8 and M9, respectively. We have designed the field-integrating uniformity stage to uniformly fill a 200-μm x 200-μm subset of the 200-μm x 600-μm wafer-side field. The cylinder-lens arrays are mounted to programmable servo scanners so that the remainder of the 200-μm x 600-μm field can be filled with scanning; this enables increased efficiency in applications that do not require the full field to be illuminated. We note that the small scan angles required to fill the full-field do not impede the ability to achieve maximal coherence when the Fourier-synthesis scanning mirrors are turned off.

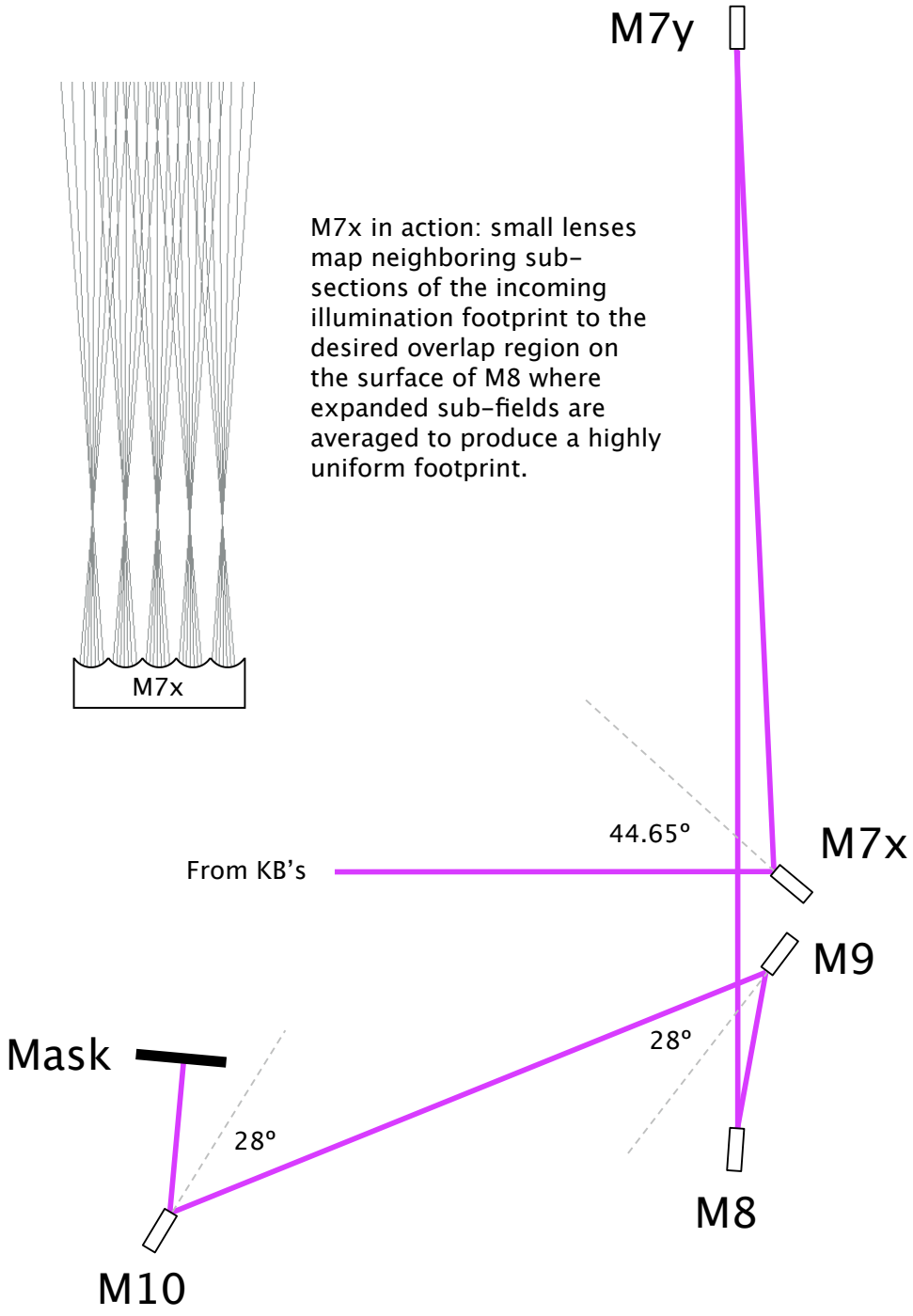

Figure 3.3: Schematic of the upgraded illuminator (M7x and M7y) and its integration into the existing Fourier-synthesis scanning mirrors (M8 and M9). M10 relays the surfaces of M8 and M9 to the mask. Distances and angles are not to scale.

The high spatial coherence of the illuminating synchrotron radiation [4] has the potential to cause unwanted interference between the beams emerging from different lenslet sites in each cylinder-lens array. To eliminate the possibility of observable fringes in the far-field overlap region, weve offset neighboring lenslets in each cylinder-lens array by the 500-nm illumination coherence length as shown in Figure 3.4; this ensures the fringes from different spectral elements (colors) are spatially shifted (dephased) enough to make the fringe contrast vanish completely. The illumination coherence length, determined by the relative spectral bandwidth arriving at the mask, is limited by the spectral pass-band associated with the concatenated series of multilayer reflections needed to get to the mask. The relative spectral bandwidth of of the multilayer set can approximated by $\Delta\lambda/\lambda\sqrt{n}$ where n is the total number of optics in the set and $\Delta\lambda/\lambda$ is the relative spectral bandwidth of each optic. Each 40-stack multilayer has $\Delta\lambda/\lambda$ of 0.06 so that the relative spectral bandwidth reaching the mask, $\Delta\lambda/\lambda_{\mathrm{MASK}}$, is 0.026 for n = 5; the coherence length at the mask, $\lambda^2/\Delta\lambda$, is 500 nm. The double pass in our reflection-based implementation ensures that the 500-nm lenslet offset will eliminate all fringe contrast in the illumination field.

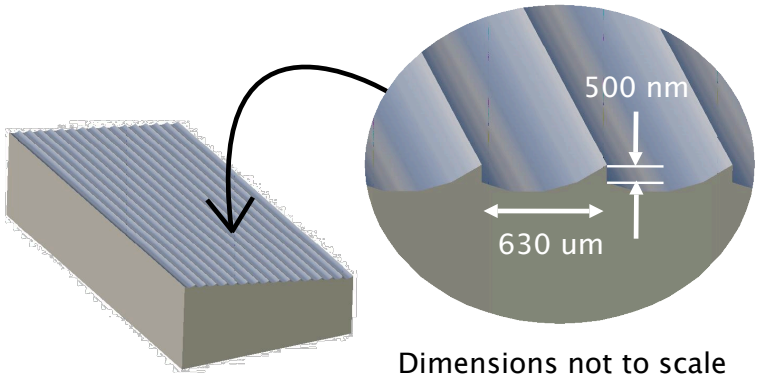

Figure 3.4: Schematic of the staircased field-averaging part. The illumination coherence length offset between lenslets dephases interference patterns from different colors, eliminating interference fringe contrast in the far-field overlap region.

Fabrication of the staircased cylinder-lens array surfaces was done by NU-TEK Precision Optical Corp [ii] using diamond turning techniques. Each cylinder-lens array has been designed to contain 20 neighboring cylindrical lenslets and have a width of 20 mm in the direction

[ii] NU-TEK Precision Optical Corp. 1202 Technology Dr., Aberdeen MD 21001

opposing focusing power. The additional lenslets (remember we only illuminate 5 of them) provide a buffer for alignment purposes and also bring the added benefit of providing extra copper for heat dissipation.

The diamond-turned lens surfaces were rendered suitable for multilayer deposition using an in-house smoothing technique [8]. Figure 3.5 shows profilometer surface height measurements from a small section of a non-staircased cylinder-lens array before and after smoothing. For comparison purposes, we show design data for the measured lens. Reflective multilayer parameters for M7x and M7y are chosen to optimize performance at the 13.5-nm lithographic wavelength while operating at the nominal angles of 44.65 degrees and 0.35 degrees, respectively. The small scan angles of the field scanners (M7x and M7y) and the Fourier-synthesis scanners (M8 and M9) ensure that all incidence angles striking these optics are within the acceptance cone of their respective multilayer coatings. After smoothing and coating, multilayer reflectivities of 61% and 64% were achieved for the M7x and M7y mirrors, respectively.

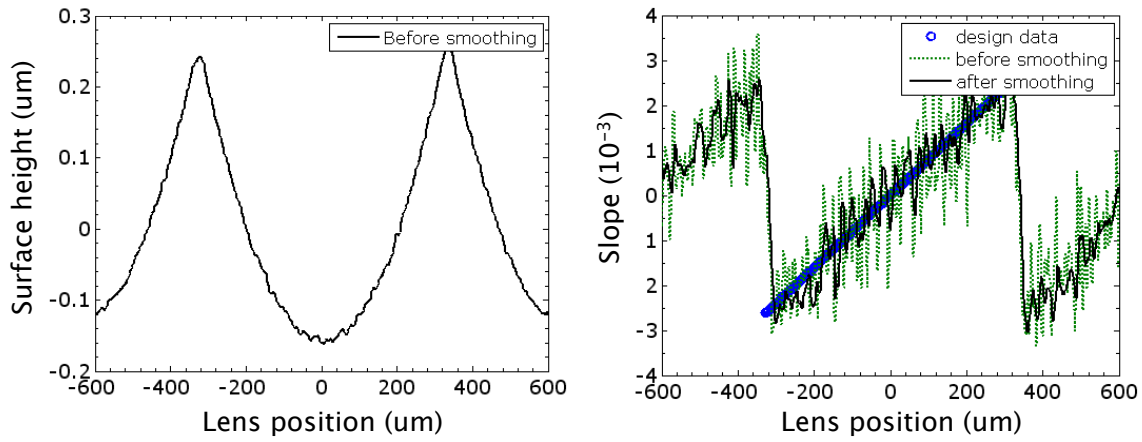

Figure 3.5: Left: cylinder-lens array surface height profile as measured with a profilometer before in-house surface smoothing. Right: cylinder-lens array surface slope before and after in-house surface smoothing; design data has been added for comparison.

The cylinder-lens arrays are mounted to Nutfield Technology Inc. [iii] Harmonic Scan HS-015 high-speed galvanometers controlled by in-house software and a Tektronix AWG7000 Arbitrary Waveform Generator. Although the field-averaging uniformity stage on its own is highly robust, its combined use with programmable scanners provides the added benefit of en-

[iii] Nutfield Technology, Inc. 49 Range Road, Windham NH, 03087

abling customizable scans to counteract any non-uniformities brought about by downstream misalignments, multilayer imperfections, etc. Integrating the field-averaging uniformity stage into the existing illuminator has required the x and y Fourier-synthesis optics (M8 and M9) to be modified. M8 and M9 have been modified to image the line-focus planes of the x and y cylinder-lens arrays to intermediate planes in x and y that are further relayed to the pupil of the main exposure tool optic.

3.4 System characterization

Figure 3.6 shows the illumination at the object (mask) following the upgrade and also shows the results of printing a clear field at a dose near the threshold of the photoresist. The uniformity improvement over the conditions before the upgrade (see Figure 3.1) is clearly evident. Note that in the before and after cases, the horizontal lines present in the scintillator images are a result of the scintillator crystal and do not represent real fluctuations in the illumination intensity at the object. The new field uniformity was quantified lithographically by measuring the fluctuations in printed feature size of an identical feature printed at nine different sites spanning the field. These measurements revealed a 6.5% peak-to-valley intensity variation across the nine points spanning the field. Finally, Figure 3.7 shows pattern printing results recorded one week after completion of the illuminator upgrade demonstrating that the tool is still in good working order. The line-space data was provided courtesy of Tom Wallow, Advanced Micro Devices.

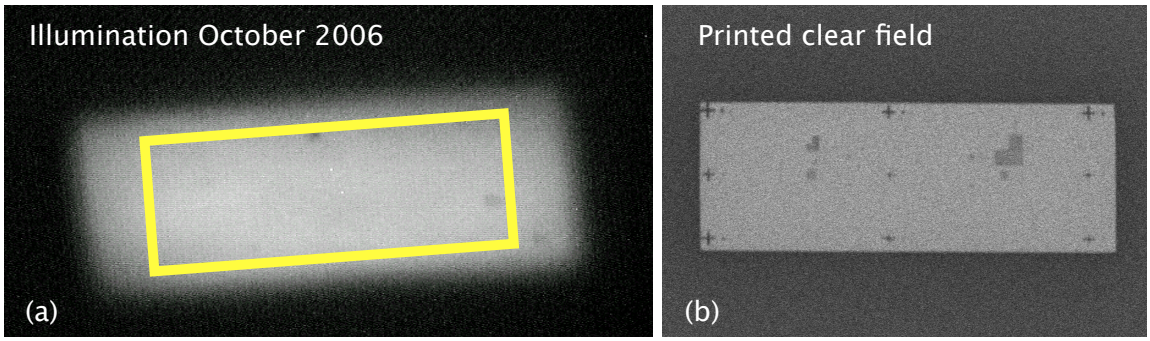

Figure 3.6: (a) Illumination field at reticle after the illuminator upgrade. Image is taken with a scintillator that converts EUV photons to visible photons. The outline shows the 1x3 mm object-side field. (b) Results of printing a clear field at a dose close to the threshold dose of the resist. Compare to the before upgrade case in Figure 3.1.

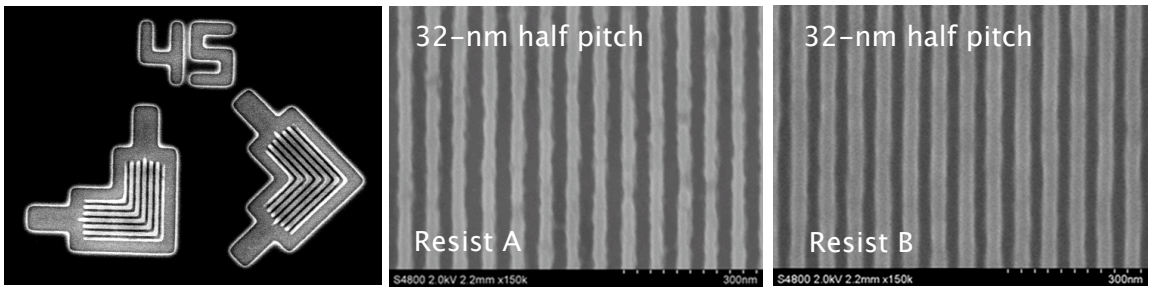

Figure 3.7: Scanning Electron Microscope (SEM) images of various printed features one week following completion of the illuminator upgrade. Line-space images courtesy of Tom Wallow, Advanced Micro Devices.

3.5 Summary

A scanning field-averaging uniformity stage has been developed that enables non-uniform, high-coherence sources to be used in applications where highly uniform illumination is required. An implementation of this system has been integrated into the existing Fourier-synthesis custom coherence illuminator of the SEMATECH Berkeley Microfield Exposure Tool at the Advanced Light Source at Lawrence Berkeley National Laboratory. Following the illuminator upgrade lithographic measurements have shown 6.5% peak-to-valley intensity variation across the entire 200-μm x 600-μm wafer-side field of view; this is a significant improvement over the previous illumination conditions.

Chapter 4

EUV Interference Lithography

Contents

4.1 Motivation

In order to address the crucial problem of high-resolution low line-edge roughness resist for extreme ultraviolet (EUV) lithography, researchers require significant levels of access to high-resolution EUV exposure tools. The prohibitively high cost of such tools, even microfield tools, has greatly limited this availability and arguably hindered progress in the area of EUV resists. To address this problem, the development of a new interference lithography (IL) tool capable of working with standalone incoherent EUV sources and based on a two-grating interferometer design has been proposed. Below we address several concerns about the feasibility of implementing incoherent EUV IL tools in practice including: grating alignment tolerances, grating fabrication tolerances, depth of focus, and source spatial/temporal coherence requirements. We also discuss the unique advantages of the two-grating configuration over conventional imaging optic implementations in terms of the spectral filtering properties of the second (imaging) grating, the homogenizing effects of incoherent illumination, and the resulting near zero-LER fringe pattern.

4.2 The two-grating interferometer

The geometry of the two-grating interferometer is shown in Figure 4.1. Incoming light strikes grating G_1 (sinusoidal with spatial frequency f_1) and splits into three diffracted orders (the zero-order transmitted beam is not shown since in practice it is blocked at the second grating). The ± 1 diffracted orders from G_1 propagate a distance z_1 to grating G_2 (sinusoidal with spatial frequency f_2) where they are redirected back towards the optic axis. The ∓ 1 diffracted orders from G_2 propagate a distance z_2 past the second grating to an interference plane where they may overlap and produce a modulated intensity pattern.

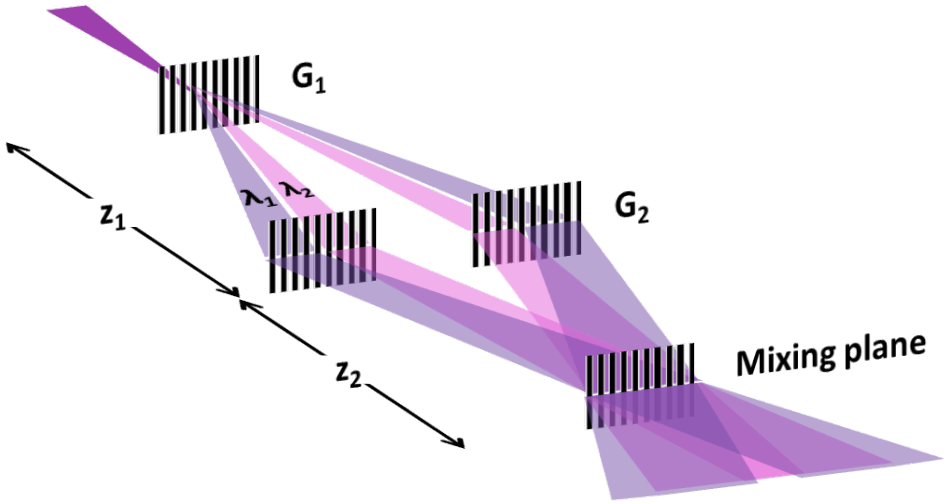

Figure 4.1: The two-grating interferometer. All diffraction gratings are assumed sinusodial. For instructive purposes, the illumination contains two distinct wavelengths, λ_1 and λ_2, in a small numerical aperture (NA) of incidence angles centered on-axis. It is assumed the zero-order transmitted beam is blocked at the second grating so it is not shown. Darker shades indicate locations where the two distinct colors spatially overlap.

4.3 Parallelism concerns are warranted

Many authors have studied the details of fringe formation in the two-grating interferometer [9, 10, 11, 12, 13, 14, 15]. To date, however, all of the theory describing fringe formation in the far-field two-grating interferometer has assumed parallel diffraction gratings. In near-field Talbot interferometry (which also uses two gratings) Patorski [16, 17] and Liu [18, 19] have shown that fringe formation is sensitive to small parallelism errors between the diffraction gratings. It is reasonable to assume that fringe formation in the far-field two-grating interferometer is also sensitive to small grating parallelism errors.

Cheng [14] has shown that the fringe depth of the parallel two-grating interferometer is inversely proportional to the spatial frequency of the gratings and the numerical aperture (NA) of the illumination. It is reasonable to believe that the effects of small grating parallelism errors may also scale inversely with grating spatial frequency and illumination NA. Experimental evidence suggests that when low-NA sources are used, the requirements on grating parallelism are within the capability of typical alignment stages [20] . On the other hand, for high spatial frequency implementations using higher-NA sources, it is reasonable to suspect that the requirements on grating parallelism might become a significant issue in practice. To

date, the best source options for stand-alone EUV IL tools are incoherent (broad) sources [21]. Due to the large collection NA of these sources and the high grating frequencies required to print features beyond the 32-nm technology node, there are serious concerns about the feasibility of implementing incoherent EUV IL tools in practice.

4.4 Previous work assuming parallel gratings

The parallel two-grating interferometer has been analyzed to several orders of accuracy in the literature. The analysis by Leith et al. [10, 11, 13] was based on a first-order approximation of the transfer function of free space. It was found that nonlocalized fringes form for polychromatic plane-wave illumination at any angle θ and localized fringes form for extended sources (multiple illumination angles), regardless of source spectral bandwidth (color content). This work concluded that in defocused interference planes, different illumination angles produce shifted versions of the same fringe pattern with the fringe patterns of all spectral elements (colors) coinciding. The net fringe dephasing between extreme illumination angles in the illumination cone was found to be proportional to the illumination NA and the longitudinal distance from the nominal interference plane.

In follow-up work by Cheng [14], a higher order analysis of the parallel two-grating interferometer was performed using a geometrical ray-tracing approach. This analysis showed that the $f_2 = 2f_1$ case, where f_1 and f_2 are the spatial frequencies of the first and second gratings, respectively, is a special configuration in which many of the second-order terms are mitigated. To second order with the $f_2 = 2f_1$ geometry, it was found that for on-axis illumination, the interference fringes produced by different spectral elements coincide with one another. It was also found that as the illumination angle goes off axis, the interference fringes produced by different spectral elements no longer coincide except in the nominal (zero-defocus) interference plane. The work by Cheng showed both the illumination NA and source spectral bandwidth play a role in fringe localization in the parallel two-grating interferometer.

To study the nonparallel two-grating interferometer we will use a phase tracking technique similar to the transfer function approach used by Leith. Instead of explicitly writing the illumination in terms of spatial frequency content we will leave the incidence angle and

illumination wavelength dependence separated so that we can easily study how angle content and spectral content independently affect the interferometer. This analysis will begin with a derivation of the propagation phase of free space between nonparallel planes as well as a derivation of the diffraction grating phase associated propagation through non-tilted and tilted gratings. Once this framework is developed we will track plane waves through the two-grating interferometer for arbitrary illumination angles, grating frequencies, axial distances and grating tilts. As has been done in the past will ignore diffraction effects from the edges of gratings and apertures [10, 11, 12, 13, 14, 15].

4.5 A framework that can accommodate tilt

4.5.1 Propagation phase

Consider a plane wave propagating at an angle θ with respect to the optical axis. To determine the phase acquired by the plane wave in propagating an axial distance z_1 we must compare the relative phase of the plane wave at the same transverse point x in the longitudinal planes $z = 0$ and $z = z_1$.

Parallel planes

Let's look at the transverse point $x = 0$ for simplicity. As shown in Figure 4.2 (left), the point $(x = 0, z = 0)$ lies on the surface of the original wavefront, on axis. The ray reaching the point $(x = 0, z = z_1)$ on G_2, however, comes from an off-axis location on the original wavefront. The ray reaching the point $(x = 0, z = z_1)$ must travel a distance $z_1 \cos\theta$ from the original wavefront surface that was in phase with the $(x = 0, z = 0)$ point. Using this geometry, we determine the propagation phase of free space propagation between two parallel planes [i]:

$$\phi(z, \theta, \lambda) = \frac{2\pi}{\lambda} z \cos\theta \tag{4.1}$$

[i] We note that we have written the propagation phase in terms of angle and wavelength rather than in terms of spatial frequency as done by Goodman [22]. It is for this reason we do not call call the propagation phase a transfer function. We also note that this phase is exact; the Fresnel approximation has not yet been made.

where z is the on-axis longitudinal separation between the two planes, θ is the propagation angle with respect to the optic axis and λ is the wavelength.

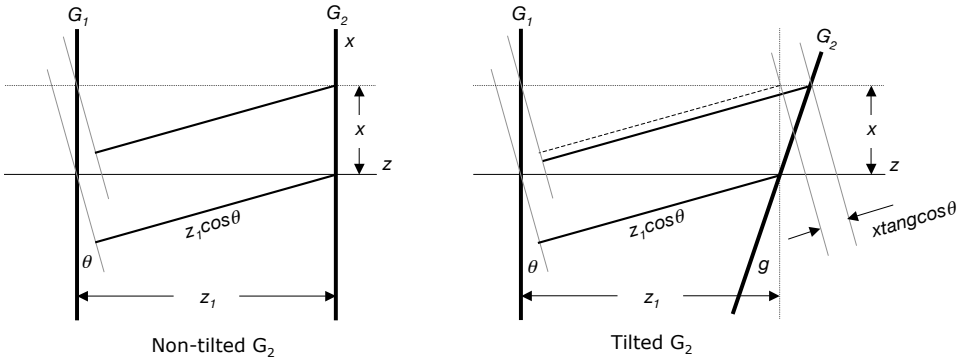

Figure 4.2: Computing the propagation phase of free space. This is a side-view schematic of the two-grating interferometer in non-tilted (left) and tilted (right) configurations. The dashed line in the tilted case shows the ray that is used for the non-tilted case. See section 2.A for an in-depth description.

Tilted planes

For the case of tilted planes we need to look at an off axis x point to observe the effect of the tilt. Let's again look at Figure 4.2 (right). The solid line shows the ray reaching G2 (now tilted) at a height x above the optic axis. The dashed line shows the ray that reaches the non-tilted plane at a height x above the optic axis. Note that the solid and dashed lines come from a different locations on the original wavefront. The grating tilt angle g causes the solid ray to travel an extra tilt-induced distance $x \tan g \cos\theta$ in addition to the distance $z_1 \cos\theta$ traveled in the parallel configuration. This extra path length lends itself to a nice physical picture, namely, the notion of an x dependent axial distance:

$$z_1'(x) = z_1 + x \tan g \tag{4.2}$$

By including the possibility of a tilt in the first plane, we develop the propagation phase of free space propagation between two tilted planes:

$$\phi(z, \theta, \lambda) = \frac{2\pi}{\lambda}(z - x\tan\theta_1 + x\tan\theta_2)\cos\theta \tag{4.3}$$

where $\theta_{1,2}$ are the tilt angles of planes 1 and 2 relative to the normal of the optic axis, θ is the propagation angle with respect to the optic axis, and z is the on-axis longitudinal separation

between planes 1 and 2. The two x dependent terms are understood as tilt corrections to the propagation phase.

4.5.2 Grating phase

It is well known [22] that diffraction gratings add a linear phase $\phi(x) = 2\pi m f x$ to the outgoing (diffracted) field where f and m are the grating spatial frequency and diffraction order, respectively . To correctly use this phase, however, the incoming field distribution must be written in the coordinate system whose optic axis is normal to the grating plane. Our geometry poses a problem for nonzero grating tilt g because in these configurations the field will be written in a coordinate system whose optic axis is tilted with respect to the grating normal. Nevertheless it is not very difficult to develop the framework to correctly describe the effect of grating phase with our geometry.

By definition, the grating phase is simply the phase difference between the field just before and just after the grating surface. Consider a plane wave propagating at an angle θ with respect to the optic axis (which we define as the normal to G_1) and impinging on G_2. For the non-tilted case (Figure 4.3 left), as you move away from the optical axis along the surface of the grating, you observe light that has travelled a distance $x \sin\theta$ further than the light that strikes the grating at the optical axis. If we tilt the grating (Figure 4.3 right) by an angle g and keep the same observation coordinate x, we observe that the light reaching the grating travels an extra tilt-induced distance $x \tan g \cos\theta$ in addition to the original distance $x \sin\theta$ traveled in the non-tilted configuration. Using these geometric distances we determine the phase (ϕ) of the field on the front surface of the grating as a function of the transverse distance x from the optical axis.

$$\phi_{front}(x) = \frac{2\pi}{\lambda} x(\sin\theta_{in} + \tan g \cos\theta_{in}) \tag{4.4}$$

where θ_{in} is the incoming propagation angle with respect to the optic axis. We can also compute the phase of the field on the rear surface of the grating using similar arguments:

$$\phi_{rear}(x) = \frac{2\pi}{\lambda} x(\sin\theta_{out} + \tan g \cos\theta_{out}) \tag{4.5}$$

where θ_{out} is the outgoing propagation angle with respect to the optic axis. The phase difference between the field at the front and rear surfaces of the grating, the grating phase, is given by:

$$\phi_g(x) = \phi_{rear} - \phi_{front} = \frac{2\pi}{\lambda} x \left[\sin\theta_{out} - \sin\theta_{in}\right] + \frac{2\pi}{\lambda} x \tan g \left[\cos\theta_{out} - \cos\theta_{in}\right] \quad (4.6)$$

To compute θ_{out} we use the grating equation:

$$\theta_{out} = \arcsin\left[\sin(\theta_{in} + g) + m\lambda f\right] - g \quad (4.7)$$

where f is the grating spatial frequency, m is the diffraction order and it is understood that we transfer into a coordinate system normal to the grating surface, use the standard grating equation [23], and then transfer back to the original coordinate system. With this propagation phase and grating phase framework we can now propagate a plane wave through the two-grating interferometer for arbitrary illumination angles, grating frequencies, axial distances and grating tilts.

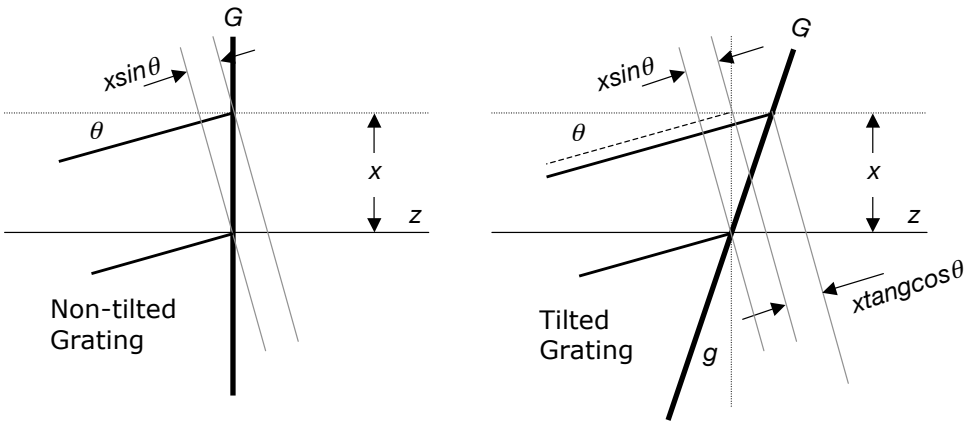

Figure 4.3: Computing the grating phase. This is a side-view schematic of a plane wave propagating at angle θ striking a diffraction grating in non-tilted (left) and tilted (right) configurations. For the tilted case to the right, the ray that would be used for the non-tilted case is shown with dashed lines.

4.6 Propagation through the interferometer

We now set out to determine the intensity distribution in the overlap region for arbitrary $\theta,\ \lambda,\ f_1,\ f_2,\ , z_1,\ z_2,\ g$ and interference plane tilt w. For bookkeeping purposes we will break the two-grating interferometer into two regions: 1 and 2; and two beams: top (T) (above the optic axis) and bottom (B) (below the optic axis) as shown in Figure 4.4. As an example of the nomenclature we write $\sin\theta_{T2}$ for the sin of the propagation angle of the top (T) beam in region 2. The optic axis will be defined as the normal to the surface of G_1 so that all grating tilts are absorbed into G_2. The grating phase assigned to a region will be the grating phase from the diffraction grating at the beginning of the region, i.e., region 1 is assigned the grating phase from G_1. The prescription for tracking the phase of a beam through the nonparallel two-grating interferometer is as follows:

1. Input a monochromatic (wavelength λ) unit amplitude plane wave propagating at angle θ with respect to the optic axis.
2. Apply grating phase for non-tilted G_1
3. Apply propagation phase for propagating the axial distance z_1 at angle θ_1 between non-tilted G_1 and tilted G_2
4. Apply grating phase for tilted G_2
5. Apply propagation phase for propagating the axial distance z_2 at angle θ_2 between tilted G_2 and tilted interference plane

The fields of the top and bottom beams in the tilted interference plane can be used to compute the intensity pattern created by the input plane wave:

$$\begin{aligned} I(x) = |E_T + E_B|^2 &= |E_T|^2 + |E_B|^2 + E_T E_B^* + E_T^* E_B \\ &= 2 + 2\cos\left(\phi(x, \theta,\ \lambda,\ f_1,\ f_2,\ z_1,\ z_2,\ g,\ w)\right) \end{aligned}$$

where $\phi = \phi_T - \phi_B$ is the phase difference between the top and bottom beams in the tilted interference plane; we will omit the explicit $x,\ \theta,\ \lambda,\ f_1,\ f_2,\ , z_1,\ z_2,\ g,\ w$ dependence from

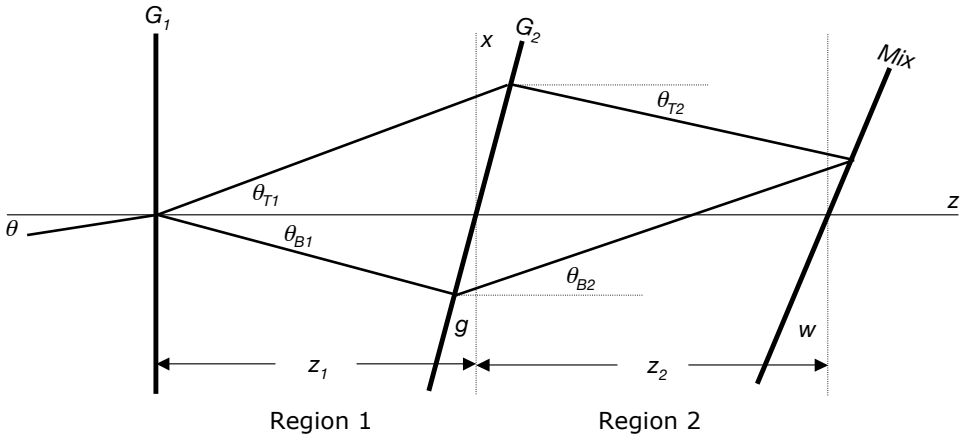

Figure 4.4: Two-grating interferometer nomenclature. This is a side-view schematic of the two-grating interferometer and describes the nomenclature used throughout this chapter for distances, angles, regions, etc.

here on out. Table 4.1 summarizes the propagation and grating phase acquired by each beam during propagation through the two-grating interferometer. To determine the propagation angles of the top and bottom beams in regions 1 and 2 we use Eq. (4.7):

$$\begin{aligned}
\theta_{T1} &= \arcsin\left[\sin\theta + \lambda f_1\right] \\
\theta_{B1} &= \arcsin\left[\sin\theta - \lambda f_1\right] \\
\theta_{T2} &= \arcsin\left[\sin(\theta_{T1} + g) - \lambda f_2\right] - g \\
\theta_{B2} &= \arcsin\left[\sin(\theta_{B1} + g) + \lambda f_2\right] - g
\end{aligned} \tag{4.8}$$

Table 4.1: Tracking the propagation and grating phase

Beam	Propagation Phase	Grating Phase
T1	$\frac{2\pi}{\lambda}(z_1 + x\tan g)\cos\theta_{T1}$	$\frac{2\pi}{\lambda}x(\sin\theta_{T1} - \sin\theta)$
B1	$\frac{2\pi}{\lambda}(z_1 + x\tan g)\cos\theta_{B1}$	$\frac{2\pi}{\lambda}x(\sin\theta_{B1} - \sin\theta)$
T2	$\frac{2\pi}{\lambda}(z_2 - x\tan g + x\tan w)\cos\theta_{T2}$	$\frac{2\pi}{\lambda}x(\sin\theta_{T2} - \sin\theta_{T1}) + \frac{2\pi}{\lambda}x\tan g(\cos\theta_{T2} - \cos\theta_{T1})$
B2	$\frac{2\pi}{\lambda}(z_2 - x\tan g + x\tan w)\cos\theta_{B2}$	$\frac{2\pi}{\lambda}x(\sin\theta_{B2} - \sin\theta_{B1}) + \frac{2\pi}{\lambda}x\tan g(\cos\theta_{B2} - \cos\theta_{B1})$

When computing the phase difference ϕ between the top and bottom beams in the tilted interference plane, all of the terms explicitly involving the G_2 tilt (g) cancel; the influence of the grating tilt remains buried inside the terms containing region 2 propagation angles.

$$\phi = \frac{2\pi}{\lambda} z_1 C_1 + \frac{2\pi}{\lambda}(z_2 + x\tan w)C_2 + \frac{2\pi}{\lambda} x S_2 \tag{4.9}$$

where

$$\begin{aligned} C_1(\lambda,\ \theta,\ f_1) &= \cos\theta_{T1} - \cos\theta_{B1} \\ C_2(\lambda,\ \theta,\ f_1,\ f_2) &= \cos\theta_{T2} - \cos\theta_{B2} \\ S_2(\lambda,\ \theta,\ f_1,\ f_2) &= \sin\theta_{T2} - \sin\theta_{B2} \end{aligned}$$

4.7 Revealing the impact of the grating tilt

Up to this point, there have been no approximations except that we have ignored diffraction effects from the edges of the gratings and apertures. In an attempt to unmask the implicit effect of the G_2 tilt buried in ϕ we now set out to simplify the troublesome terms involving the region 2 propagation angles governed by Eq. (4.7). Assuming a small grating tilt allows us to use the expansion:

$$\arcsin(\alpha + \delta) \approx \arcsin\alpha + \delta(1-\alpha^2)^{-1/2} \tag{4.10}$$

on Eq. (4.7) to pull the grating tilt *inside* the arcsin argument, i.e.,

$$\theta_{out} \approx \arcsin\left[\sin(\theta_{in} + g) + m\lambda f - g\left(1 - [\sin(\theta_{in} + g) + m\lambda f]^2\right)^{1/2}\right]$$

Furthermore, since small g is assumed, we can make another expansion, namely:

$$\sin(\theta_{in} + g) \approx \sin\theta_{in} + g\cos\theta_{in}$$

so that Eq. (4.7) is well approximated by:

$$\theta_{out} \approx \arcsin\left[\sin\theta_{in} + m\lambda f'\right] \tag{4.11}$$

where

$$f' = f + \frac{g}{m\lambda}\left[\cos\theta_{in} - \left(1 - [\sin\theta_{in} + g\cos\theta_{in} + m\lambda f]^2\right)^{1/2}\right]$$

can be considered the effective spatial frequency of a virtual non-tilted grating that generates the same propagation angles as the true tilted grating.

Continuing our efforts to unmask the effect of the G_2 tilt, we will now assume several restrictions on the interferometer geometry that make the problem tractable. As described by Cheng [14], the $f_2 = 2f_1$, $z_1 = z_2$ configuration of the parallel two-grating interferometer is a special case in which many of the second-order spectral bandwidth and angle bandwidth terms drop out. For the remainder of our analysis we will consider practical implementations of this configuration and assume $f_2 = 2f_1(1+\gamma)$ and $z_2 = z_1 + d$ where γ is understood as a small unitless G_2 pitch error and d is a small defocus parameter. With this configuration, the effective spatial frequency of G_2 for both beams can be easily derived.

Starting with Eq. (4.11), we wish to determine f'_{2T}, the effective spatial frequency of the virtual grating G'_2 for the top beam. We make the following substitutions in Eq. (4.11): $f = 2f_1(1+\gamma)$, $m = -1$, and $\sin\theta_{in} = \sin\theta_{T1} = \sin\theta + \lambda f_1$. We obtain:

$$f'_{2T} = 2f_1(1+\gamma) - \frac{g}{\lambda}\left[\cos\theta_{T1} - \left(1 - \left[\sin\theta_{B1} + g\cos\theta_{T1} - 2\lambda f_1\gamma\right]^2\right)^{1/2}\right]$$

where we've used $\sin\theta - \lambda f_1 = \sin\theta_{B1}$ inside the small-bracketed term. As we're only interested in keeping terms that are first order in the small parameters g and γ, we drop the $g\cos\theta_{T1} - 2\lambda f_1\gamma$ term before the squaring operation; we're able to do this because of the preceding g that multiplies everything inside of the large brackets. After simplification we obtain:

$$f'_{2T} \approx 2f_1(1+\gamma) - \frac{g}{\lambda}\underbrace{\left[\cos\theta_{T1} - \cos\theta_{B1}\right]}_{C_1}$$

The calculation for f'_{2B} (the effective spatial frequency of the virtual grating G'_2 for the bottom beam) is done in a similar fashion and yields the exact same result. We can now use Eq. (4.11) with the above f'_2 to write approximations for the region 2 propagation angles in Eq. (4.8):

$$\begin{aligned}
\theta_{T1} &= \arcsin\left[\sin\theta + \lambda f_1\right] \\
\theta_{B1} &= \arcsin\left[\sin\theta - \lambda f_1\right] \\
\theta_{T2} &\approx \arcsin\left[\sin\theta_{B1} - 2\lambda f_1\gamma + gC_1\right] \\
\theta_{B2} &\approx \arcsin\left[\sin\theta_{T1} + 2\lambda f_1\gamma - gC_1\right]
\end{aligned} \tag{4.12}$$

To simplify the C_2 term with the rewritten propagation angles we introduce the expansion:

$$\cos\left[\arcsin(\sin\alpha + \delta)\right] \approx \cos\alpha - \delta\tan\alpha - \frac{\delta^2}{2\cos^3\alpha} \tag{4.13}$$

where we have expanded the function $f(x) = \cos\left[\arcsin(x)\right] = (1-x^2)^{1/2}$ to second order about the nominal value $\sin\alpha$ and have assumed δ is a small correction term. Using Eq. (4.13) to expand C_2 to first order in the small parameters g and γ we obtain:

$$C_2 \approx -C_1 + \left(2\lambda f_1\gamma - gC_1\right)T_1$$

where:

$$T_1(\lambda,\ \theta,\ f_1) = \tan\theta_{T1} + \tan\theta_{B1}$$

Simplifying the S_2 term in Eq. (4.9) is relatively straightforward with the approximated region two angles in Eq. (4.12); the result is $S_2 \approx -2f_1(1+2\gamma)/\lambda$. Combining all of these approximations we achieve an expression for the phase that explicitly shows the impact of G_2 tilt and pitch error:

$$\begin{aligned}
\phi \approx &-\underbrace{2\pi x(2f_1)(1+2\gamma)}_{\text{desired modulaiton}} + \underbrace{2\pi x/\lambda\left[2gC_1 + \tan w\left(-C_1 + 2\lambda f_1\gamma T_1 - gC_1T_1\right)\right]}_{\text{unwanted modulation}} \\
&+ \underbrace{2\pi/\lambda\left[-dC_1 + z_1\left(2\lambda f_1\gamma - gC_1\right)T_1\right]}_{\theta,\ \lambda\text{ dependent fringe shift}}
\end{aligned}$$

By setting the interference plane tilt, w, to twice the grating tilt, g, the "unwanted modulation" term becomes second order in small parameters g and γ. For the remainder of the analysis we will assume that $w = 2g$ and drop the unwanted modulation term entirely. The expression for the phase becomes:

$$\phi \approx -\underbrace{d\,[2\pi C_1/\lambda]}_{\text{defocus } \Delta\phi} + \underbrace{\gamma\,[4\pi z_1 f_1 T_1]}_{\text{pitch mismatch } \Delta\phi} - \underbrace{g\,[2\pi z_1 C_1 T_1/\lambda]}_{\text{tilt } \Delta\phi} - \underbrace{2\pi x(2f_1)(1+2\gamma)}_{\text{desired modulation}} \tag{4.14}$$

and we remind the reader that C_1 and T_1 are both functions of $\theta,\ \lambda$, and f_1. The first and last terms in this expression represent the usual defocus term described in detail in the literature [14, 15] and the desired modulation at spatial frequency $2f_1(1 + 2\gamma)$, respectively. The second and third terms represent γ and g induced limitations, respectively, on the tolerable illumination $\Delta\theta$ and $\Delta\lambda$ that maintain fringe contrast in the interference plane – even with zero defocus.

4.8 Discuss defocus, pitch mismatch, tilt, and optimization

We begin by reminding the reader that the analysis leading to Eq. (4.14) extends only to the limited class of two-grating interferometers with small tilts between the gratings and with $z_2 \approx z_1$ and $f_2 \approx 2f_1$. In this special configuration we've obtained an expression for the phase ϕ that explicitly shows the effects of grating tilt, pitch mismatch and interference plane defocus.

The utility of the interferometer, however, doesn't depend on the phase itself, but how rapidly it changes as a function of illumination color and angle. For nonzero d, γ and g, the first three terms in Eq. (4.14) independently cause dephasing between the fringe patterns associated with different illumination colors and angles. To have an interferometer in working order (i.e., having reasonable fringe contrast) the fringe-pattern-shift from the two extreme illumination spatial frequencies should be less than the width of a fringe. Satisfying this specification requires the net dephasing from the first three terms in Eq. (4.14) to be less than π.

4.8.1 Illumination conditions

At this point it is useful to define two illumination classes that limit fringe contrast in clear-cut ways:

1. Temporally limiting illumination

- Polychromatic with bandwidth $\Delta\lambda$ and center wavelength $\bar{\lambda}$ (partially temporally coherent).
- One incidence angle at θ_0 (spatially coherent)

2. Spatially limiting illumination

- Monochromatic at wavelength $\bar{\lambda}$ (temporally coherent).
- Full-NA of incidence angles $\Delta\theta$ centered at θ_0 (partially spatially coherent).

For temporally limiting illumination in the ideal configuration (g = γ = 0) it has been shown [14, 15] that the temporal-limited depth of focus, $\text{DOF}_{\Delta\lambda}$, is given by:

$$\text{DOF}_{\Delta\lambda} = \frac{\cos^4\theta_0}{2\bar{\lambda}\Delta\lambda f_1^3 \tan\theta_0} \tag{4.15}$$

where the DOF is defined as the twice the largest d that limits the net dephasing from different colors in the illumination to a maximum value of π (term 1 in Eq. (4.14)) . For spatially limiting illumination in the ideal configuration, a spatially-limited depth of focus, $\text{DOF}_{\Delta\theta}$, can be derived. We begin by using Eq. (4.13) to expand C_1 to second order[ii] and then we differentiate the result with respect to θ.

$$C_1 \approx -2\lambda f_1 \tan\theta$$

$$\frac{\partial C_1}{\partial\theta} \approx -\frac{2\lambda f_1}{\cos^2\theta}$$

Using the above relationship, we can determine the value of d where the net dephasing (fringe-pattern-shift) of all the incidence angles within the illumination cone reaches a maximum value of π.

$$\Delta\phi = -d\frac{2\pi}{\lambda}\Delta\theta\left[\frac{\partial C_1}{\partial\theta}\right]_{\theta=\theta_0} = \pi$$

Defining the DOF as twice the maximum dephasing distance, we obtain:

[ii]Here the 'small correction term' does not contain a small parameter (g, γ, d) so we go to second order to maintain reasonable accuracy in the expansion.

$$DOF_{\Delta\theta} = 2d = \frac{\cos^2\theta_0}{2f_1\Delta\theta} \tag{4.16}$$

where $\Delta\theta$ is the full input angle range (2×NA) and θ_0 is the nominal illumination angle. When $\Delta\theta$ and $\Delta\lambda$ are both nonzero, the Eqs. (4.15) and (4.16) are useful for determining whether or not one form of dephasing dominates the other; this is often the case in practice. We define the illumination to be NA-limited when the dephasing from illumination spectral content ($\Delta\lambda$) is negligible in comparison to the dephasing from illumination angle content ($\Delta\theta$). The specification requires $\text{DOF}_{\Delta\theta} \ll \text{DOF}_{\Delta\lambda}$ or:

$$\frac{\Delta\lambda}{\bar{\lambda}} \ll \frac{\Delta\theta\cos^2\theta_0}{(\bar{\lambda}f_1)^2\tan\theta_0} \tag{4.17}$$

When spatially broad illumination is used, the illumination is often NA-limited. For example, when $\Delta\theta = 4^\circ$ (0.07 rad), $\theta_0 = 15^\circ$ and $\bar{\lambda}f_1 = 0.25$, NA-limitation requires $\Delta\lambda/\bar{\lambda} \ll 3.9$ which is satisfied by almost any existing source at any center wavelength. Although we've argued this case for the ideal configuration, is reasonable to assume that these characteristics also hold in the $g \neq 0,\ \gamma \neq 0$ configuration. For the remainder of the discussion we will consider the class of nonparallel interferometers that operate in the NA-limited regime: this enables us to treat the specific case of spatially limiting illumination and apply the results to the larger context of all NA-limited illumination schemes. We now focus our attention on Eq. (4.14), specifically, we wish to determine how θ_0, $\Delta\theta$, d, γ, and g affect fringe formation.

4.8.2 Grating parallelism tolerances

Before we examine the interplay between the various terms in Eq. (4.14) we will approximate C_1 and T_1 with Eq. (4.13) and with the following expansion:

$$\tan[\arcsin(\sin\alpha + \delta)] \approx \tan\alpha + \frac{\delta}{\cos^3\alpha} + \frac{3\delta^2\sin\alpha}{2\cos^5\alpha} \tag{4.18}$$

Here we have expanded the function $f(x) = \tan[\arcsin(x)] = x/(1-x)^{1/2}$ to second order about the nominal value $\sin\alpha$ and have assumed δ is a small correction term. Using Eqs. (4.13) and (4.18) we obtain:

$$\begin{aligned} C_1 &\approx -2\lambda f_1 \tan\theta \\ T_1 &\approx 2\tan\theta + \frac{3\lambda^2 f_1^2 \sin\theta}{\cos^5\theta} \\ C_1 T_1 &\approx -4\lambda f_1 \tan^2\theta \end{aligned} \tag{4.19}$$

where for small θ the first term in Eq. (4.19) for T_1 suffices for calculation purposes. With these approximations Eq. (4.14) can again be simplified:

$$\begin{aligned} \phi(\theta) \approx\ & d[4\pi f_1 \tan\theta] + \\ & \gamma[8\pi z_1 f_1 \tan\theta] + \\ & g[8\pi z_1 f_1 \tan^2\theta] + \\ & 2\pi x(2f_1)[1+2\gamma] \end{aligned} \tag{4.20}$$

To come up with a tolerance spec on g that enables a workable DOF (one where contrast reduction with defocus is dominated only by the defocus term), we require that the dephasing caused by term 3 is much less than the dephasing caused by term 1 when d is set to $\text{DOF}_{\Delta\theta}$.

For on-axis illumination with full input angle range = $\Delta\theta$ and $\gamma = 0$, the dephasing requirement is written as[iii]:

$$g 8\pi z_1 f_1 (\Delta\theta/2)^2 \ll d 4\pi f_1 \Delta\theta \tag{4.21}$$

setting d = $\text{DOF}_{\Delta\theta}$ and simplifying gives:

$$g \ll \frac{2\text{DOF}_{\Delta\theta}}{z_1 \Delta\theta} = \frac{1}{f_1 z_1 (\Delta\theta)^2} \tag{4.22}$$

As an example, we consider parameters suitable for an incoherent EUV IL tool: $1/f_1 = 30$ nm, $z_1 = 20$ mm and $\Delta\theta = 4^\circ$. These parameters require $g \ll 300$ μrad which is manageable in practice.

[iii]The tilt term is even in theta so only half of the input angle range is required. The defocus term is odd in theta so the full input angle range is required

For off-axis illumination with full-NA = $\Delta\theta$ and $\gamma = 0$, the dephasing requirement is written as[iv]:

$$g8\pi z_1 f_1 \Delta\theta \left[\frac{d(\tan^2\theta)}{d\theta}\right]_{\theta=\theta_0} \ll d4\pi f_1 \Delta\theta \left[\frac{d(\tan\theta)}{d\theta}\right]_{\theta=\theta_0} \tag{4.23}$$

The derivatives of $\tan\theta$ and $\tan^2\theta$ are given by:

$$\begin{aligned} \frac{d(\tan\theta)}{d\theta} &= \frac{1}{\cos^2\theta} \\ \frac{d(\tan^2\theta)}{d\theta} &= \frac{2\sin\theta}{\cos^3\theta} \end{aligned} \tag{4.24}$$

setting d = DOF$_{\Delta\theta}$ and simplifying gives:

$$g \ll \frac{\text{DOF}_{\Delta\theta}}{4z_1\tan\theta_0} = \frac{\cos^3\theta_0}{8f_1 z_1 \Delta\theta \sin\theta_0} \tag{4.25}$$

Eq. (4.25) is equivalent to Eq. (4.22) with $\Delta\theta$ replaced by $8\tan\theta_0$. With the EUV IL tool parameters form above we find that θ_0 = 5° requires $g \ll 30$ μrad; increasing θ_0 to 30° tightens the spec even more to $g \ll 5$ μrad. These findings suggest that interferometers operating in the NA-limited regime should avoid off-axis implementations (which would be required if reflective gratings are used), especially in applications where high spatial frequency gratings and larger numerical apertures are used.

4.8.3 Grating pitch mismatch and optimization

Up to this point we have assumed that γ, the grating pitch mismatch, is zero. We now include the possibility of a grating pitch mismatch, however, we treat it in the sense that the pitch mismatch is fixed at one value and cannot be altered. Because g and d can be manipulated in practice, there is the possibility of using a nominal defocus and/or grating tilt to partially mitigate the effects of the pitch mismatch.

One straightforward optimization scheme is to use the dephasing from a small amount of defocus to mitigate most of the dephasing due to the grating pitch mismatch. For this optimization we use term 1 to cancel term 2 in Eq. (4.14) at the center wavelength and center illumination angle. The optimized value for the defocus is $d_{ideal} = 2z_1\bar{\lambda}f_1\gamma T_1(\theta_0,\ \bar{\lambda})/C_1(\theta_0,\ \bar{\lambda})$.

[iv] Even and odd symmetry is irrelevant off axis so the full NA is used for both terms

In order to include the possibility of a focus error, now with respect to d_{ideal}, we redefine the defocus parameter: $d \equiv d_{ideal} + d'$ where d' is understood as the defocus with respect to to d_{ideal}. The expression for the phase in this optimization is written:

$$\phi \approx -\frac{2\pi}{\lambda} d' C_1 + 2\pi z_1 2 f_1 \gamma N - \frac{2\pi}{\lambda} z_1 g C_1 T_1 - 2\pi x (2 f_1)(1 + 2\gamma) \tag{4.26}$$

where

$$N(\theta,\ \lambda,\ \theta_0,\ \bar{\lambda}) = \left[1 - \frac{\bar{\lambda} T_1(\theta_0,\ \bar{\lambda}) C_1(\theta,\ \lambda)}{\lambda T_1(\theta,\ \lambda) C_1(\theta_0,\ \bar{\lambda})}\right] T_1(\theta,\ \lambda)$$

describes the γ-dephasing that could not be cancelled with defocus. For the remainder of the discussion we will drop the prime on d'. Assuming spatially limiting illumination we can simplify the expression for N with the substitution $\lambda \equiv \bar{\lambda}$. In section 4.8.2 we concluded that off-axis implementations should be avoided in applications with large NA's and high grating frequencies; we limit our discussion of pitch mismatch tolerances to illumination cones centered on-axis. When $\theta_0 = 0^\circ$ we can use Eq. (4.19) to write a very good approximation for N.

$$\underset{\theta_0=0}{N(\theta,\ \lambda)} \approx \frac{3\lambda f_1 C_1}{2} \left(1 - \cos^{-4}\theta\right) \approx -3\lambda f_1 C_1 \theta^2$$

where the final approximation is valid only for small θ. A workable DOF requires that the dephasing caused by the remaining pitch mismatch term in Eq. (4.26) (term 2) is much less than the dephasing caused by the DOF term (term 1). The dephasing requirement leads to the following condition:

$$\gamma \ll \frac{\text{DOF}_{\Delta\theta}}{6\lambda^2 f_1^2 z_1 (\Delta\theta)^2} = \frac{1}{24\lambda^2 z_1 (f_1 \Delta\theta)^3} \tag{4.27}$$

For the EUV IL tool example with $1/f_1 = 30$ nm, $z_1 = 20$ mm, $\Delta\theta = 4^\circ$ and λ = 13.5 nm, Eq. (4.27) requires $\gamma \ll 10^{-2}$ or a pitch mismatch much less than one part in a hundred. As the nominal operation wavelength increases and the illumination NA decreases this spec becomes more lenient. We remind the reader that the specification determined here is for the straightforward optimization scheme where defocus dephasing is used to compensate for pitch mismatch dephasing. The optimization parameter space, however, is quite large as it contains three interrelated parameters $\theta_0,\ g$, and d; determining the optimal combination of

g and d for each θ_0 would require a detailed numerical study that will not be pursued in this work.

4.9 A recap of what has been done

By using a two-dimensional phase tracking approach we have obtained the exact expression for the interference pattern in a two-grating interferometer when the angle between the gratings (g) and the angle between the final grating and the interference plane is arbitrary. When practical implementations of the $f_2 = 2f_1$, $z_2 = z_1$ configuration are considered and small g is assumed, several binomial expansions of the arguments inside the cosine function of the fringe pattern bring out the wavelength and incidence-angle dependence of grating tilt, grating pitch mismatch, interference plane tilt, and interference plane defocus.

For on-axis and off-axis nominal incidence angles we have derived specifications for tolerable g. It was found that for off-axis nominal incidence angles the specification for tolerable g can become 1 or 2 orders of magnitude more strict than the specification for $\theta_0 = 0°$. In general, the g-induced dephasing is proportional to g, f_1, z_1 and $\Delta\theta$ and increases substantially as θ_0 moves off-axis. For tilt angle g between the two gratings, we have found the optimal angle between the final grating and the interference plane is also g. In this configuration all unwanted modulation terms are of $\mathcal{O}(2)$ or higher in the small parameters g, γ, and d.

It was found that for a small grating pitch mismatch [we assume $f_2 = 2f_1(1+\gamma)$], nonzero nominal defocus and grating tilt may be used to partially mitigate the pitch mismatch dephasing in broad-source implementations. In the straightforward optimization scheme where dephasing from defocus is used to mitigate dephasing form pitch mismatch, a specification for tolerable pitch mismatch has been obtained.

We have found that experimental limitations in grating alignment, grating pitch matching and interference plane focus control have an increasingly significant impact on interferometer performance as the grating period decreases and the illumination NA increases. For incoherent EUV IL tools it is clear that off-axis reflection-based implementations should be avoided due to the impractical tolerance on grating alignment. The operation of on-axis transmission-based incoherent EUV IL tools, however, appears to be within the capabilities of existing

nano-positioning stages and nano-fabrication facilities.

4.10 Determination of grating quality requirements

While alignment, pitch-matching and focus control are important considerations for the two-grating interferometer, another major concern is the requirements on the gratings themselves. To address the grating quality issue we have developed 1D and 2D propagation codes (following the theory developed in Section 4.6) that include realistic representations of grating noise. Below we model the tool geometry described in Section 4.8.2 ($1/f_2 = 30$ nm, $z_1 = z_2 = 20$ mm, NA = 0.07) with the additional constraint of parallel gratings and nominal pitch-matching. We assume 4-mm-wide gratings and a 4-mm field size.

4.10.1 1-D modeling

Perfect gratings

We begin by verifying the DOF in the noise-free case assuming plane wave illumination components. With the DOF based on a fringe contrast limit of 50%, we found a DOF of 500 nm, very close to the value predicted by Eq. 4.16. Achieving plane wave illumination requires a near perfect illuminator and is not likely in practice. To study the potential impact of this, the modeling was expanded to include the possibility of non-planar constituent illumination components. Taking the collimator-induced divergence to have a 100-mm radius, we again find perfect fringes.

Imperfect gratings

Next we consider the grating quality issue explicitly. Grating noise can be viewed as spurious spatial frequency components in addition to the target spatial-frequency of the grating. Due to the tool geometry, we need only consider a limited range of grating noise frequencies. In the first grating, the highest spatial frequency noise of interest redirects the most extreme input spatial carrier (f = -NA/λ) so that its footprint just misses the second grating on the high side (see Figure 4.5). Simple geometry shows that largest relevant noise frequency in G1 is:

$$f_{max} = f_1 + \frac{\text{NA}}{\lambda} + \frac{w}{\lambda z} \tag{4.28}$$

where the first term is the nominal grating frequency, the second term is the spatial frequency of the most extreme input spatial carrier and the third term is the spatial frequency shift associated with laterally displacing the beams footprint one full width (w)at the second grating.

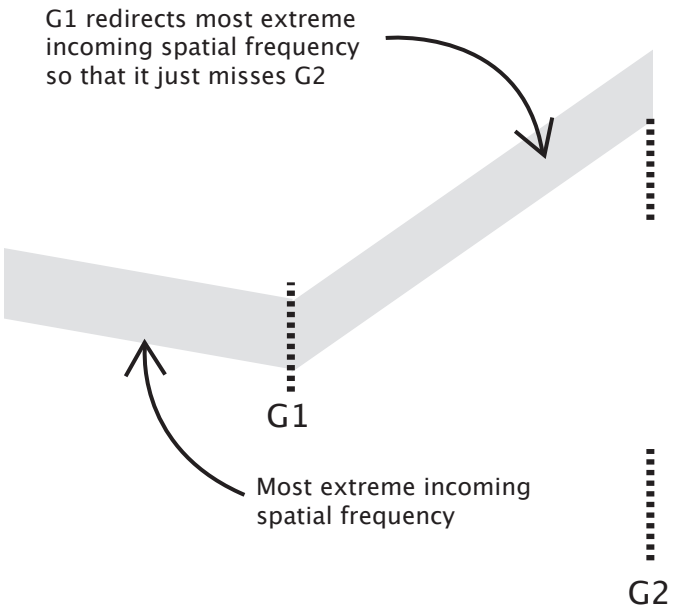

Figure 4.5: Plots of the C_1T_1 product (left) and it's derivative with respect to. θ (right) for for several values of $\eta \equiv \lambda f_1 \in [0.05, 0.25]$ with $-45 < \theta < 45$ degrees.

Realistic gratings are modeled in the frequency domain by a spectrum containing a unit amplitude delta function at $\pm f_1$ surrounded with Gaussian-weighted randomly phased and attenuated noise frequency samples within the geometrically defined noise bandwidth. We define the parameter σ as the relative bandwidth of the Gaussian weighting function, and define the noise percentage η as the intensity summation of all sampled noise frequencies relative to the intensity summation of all sampled frequencies. Figure 4.6 shows typical grating spectra with σ= .1 (left) and σ = .4 (right) ; η = .5 in both spectra.

The wafer-plane intensity distribution associated with a single input spatial frequency is determined as follows:

1. Pick a frequency from the G1 frequency spectrum.

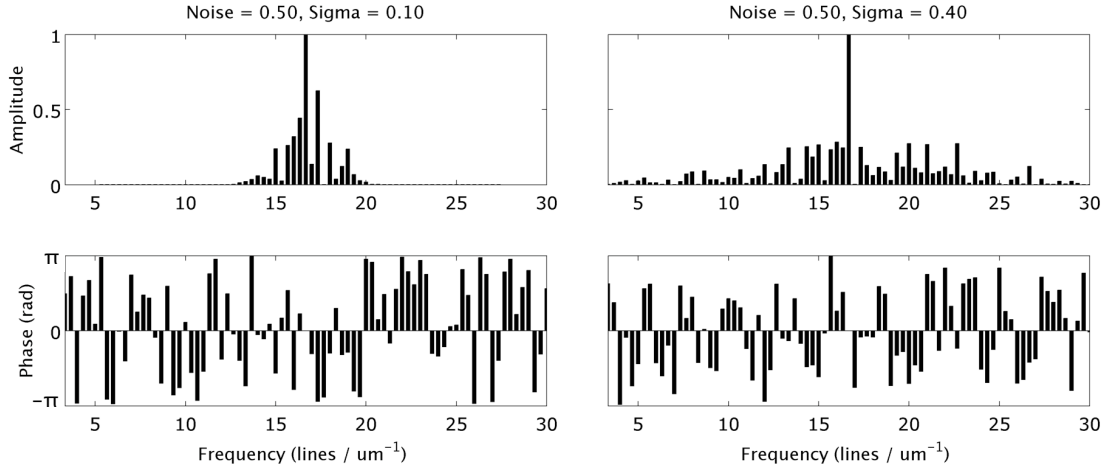

Figure 4.6: Simulated 1D grating frequency spectra for σ = 0.1 (left) and σ = 0.4 (right). η = 0.5. Field amplitude is shown at top with corresponding phase at the bottom.

2. Track the input spatial carrier through the tool as described in Section 4.6 assuming that G1 has the frequency that was picked in step 1 and that G2 is perfect.

3. Store the wafer-plane E-fields for the top and bottom branches for later use.

4. Go back to step 1, pick a different frequency and repeat until the top and bottom branch E-fields for each frequency in the G1 spectrum are stored, then proceed.

5. Linear superposition: weight each E-field pair (top + bottom branch) by its corresponding amplitude and phase in the G1 frequency spectrum and sum all E-fields together.

6. Square the result.

The final intensity distribution is the incoherent summation of the intensity patterns from each input spatial frequency.

This model has been used to test the sensitivity of the wafer plane intensity to changes in σ and η. Figure 4.7 (left) shows simulated 1D intensity cross sections over a 2 μm section of the wafer as σ is varied through [0:1] in the vertical dimension. Fringe contrast and fringe contrast uniformity are unaffected by the relative weighting of noise frequencies in the specified range. Figure 4.7 (right) shows simulated 1D intensity cross sections over a 2 μm section of the wafer as η is varied through [0:1] in the vertical dimension. We observe an almost linear relationship between fringe contrast and η: contrast = 1- η.

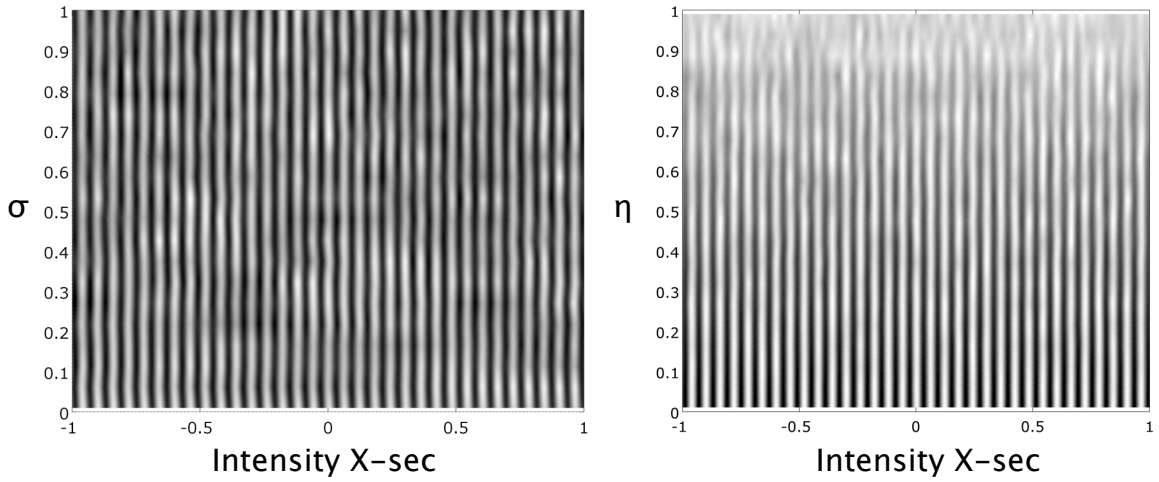

Figure 4.7: Fringe pattern vs. grating spectrum relative bandwidth, σ (left) and vs. grating noise percentage, η

4.10.2 2-D modeling

Ultimately, we are interested in the tool performance with realistic diffraction gratings that have noise components in both spatial dimensions. Figure 4.8 shows G1 frequency spectra with different relative bandwidths in x and y: $\sigma_x = \sigma_y = 1$ (top) and σ_x = 1, $\sigma_y = .03$ (bottom). Like the 1D model, interferometer performance is not sensitive to the widths (σ_x, σ_y) of the 2D grating noise spectrum. Figure 4.9 shows that the relationship between fringe contrast and η follows the same linear trend in 1D and 2D modeling. These results suggesting that 1D modeling is adequate to describe the impact of imperfect gratings on tool performance.

Figure 4.10 (middle row) shows the wafer plane intensity resulting from 2D gratings with η = 0.1, 0.3, and 0.5. The images of the modeled gratings are shown in the top row and the local fringe contrasts are shown in the bottom row. Grating images are obtained from the E-field leaving G1, which in turn is obtained by inverse transforming the under-sampled frequency spectrum leaving G1 to a 1/df-wide subset of the 4-mm field where df is the sample spacing in the frequency spectrum. The resulting E-field amplitude is squared and thresholded at 0.5 to produce the grating images. The robustness of the tool to poor grating quality will be discussed below.

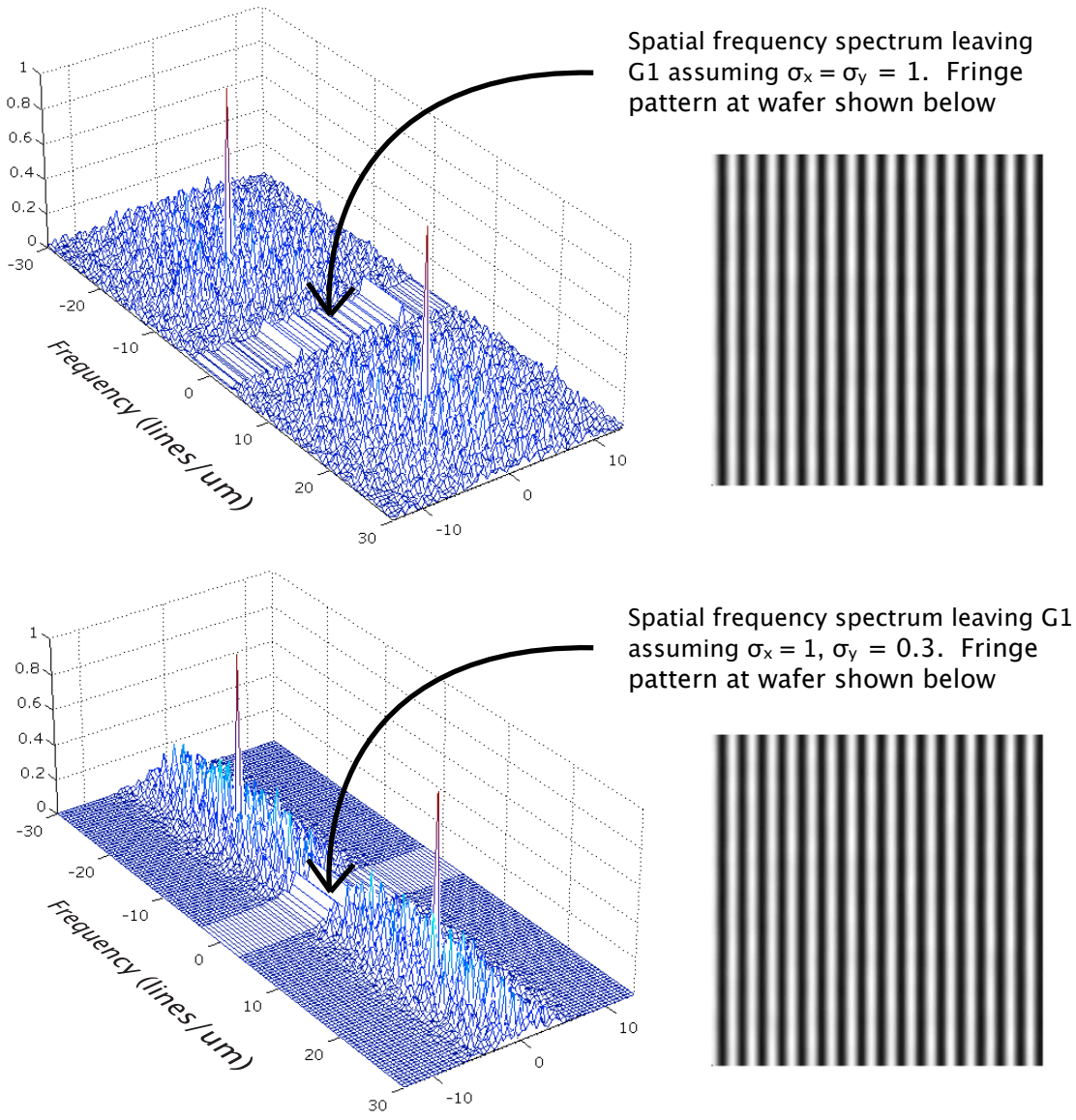

Figure 4.8: Top: G1 spectrum with $\sigma_x = \sigma_y = 1$ and resulting intensity at wafer. Bottom: G1 spectrum with $\sigma_x = 1$, $\sigma_y = 0.3$ and resulting intensity at wafer . There is no observable difference in fringe contrast.

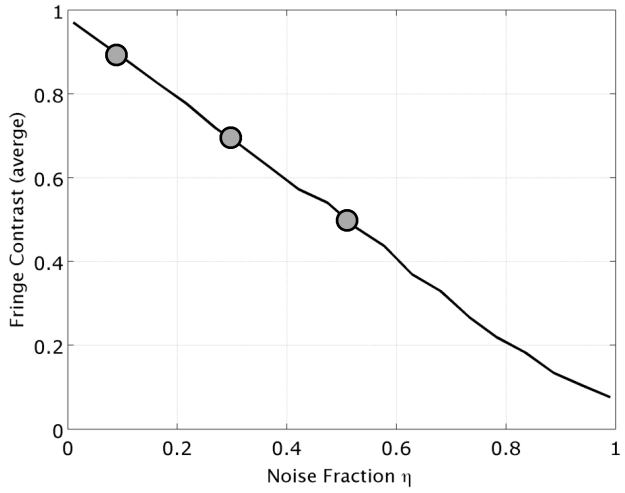

Figure 4.9: Fringe contrast vs. η. Three 2D data points are superimposed on the well-sampled 1D data from Figure4.7 (right). Results suggest that 1D modeling is adequate to model the effect of imperfect gratings.

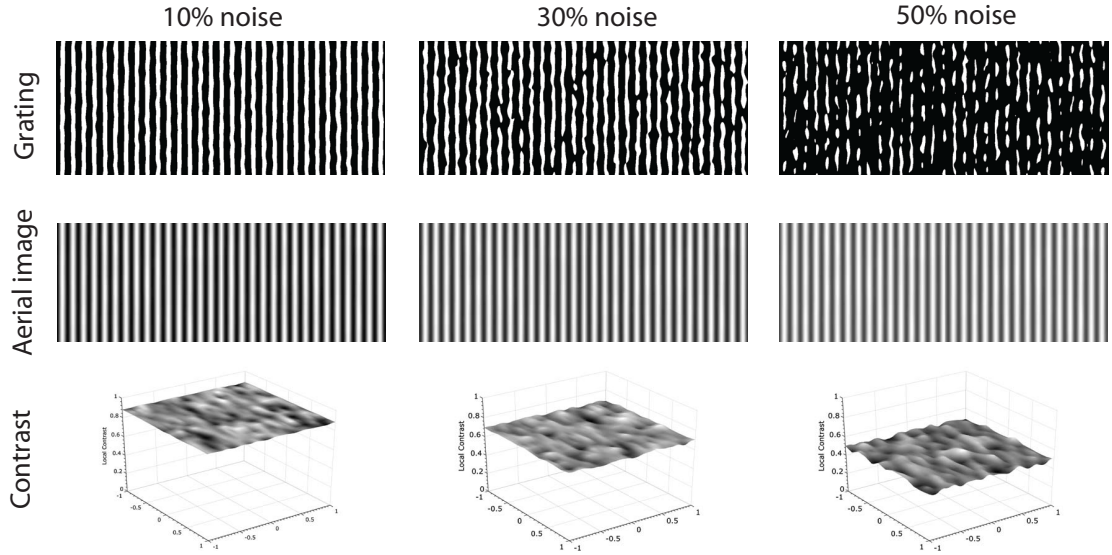

Figure 4.10: Top: typical 2D gratings with η = 0.1, 0.3, and 0.5. Middle: intensity patterns at the wafer assuming perfect G2. Bottom: local contrast of each intensity pattern.

4.11 The near zero-LER aerial image

One desirable property of the incoherent two-grating interferometer is that it has a very low LER aerial image. To describe how this is possible, it is useful to point out how the recombination grating acts differently than a lens would in its place.

Lenses *scale* incoming spatial frequencies, diffraction gratings *shift* incoming spatial frequencies, i.e., $f_{out} = f_{in}/M$ (lens) vs. $f_{out} = f_{in} + mf_g$ (grating) where M is the magnification, m is the diffraction order, and f_g is the grating frequency. The result is that with a recombining lens, diffracted rays from different object spatial frequencies converge in a single single z plane (Figure 4.11 top.) where as with a recombining grating, they converge in different planes (Figure 4.11 bottom). For the later, a consequence is that the focus plane for the object grating's nominal frequency is a defocused plane for all other grating noise frequencies. As a result, noise terms show up at the wafer as DC and contribute mainly to a loss of contrast in the fringe pattern.

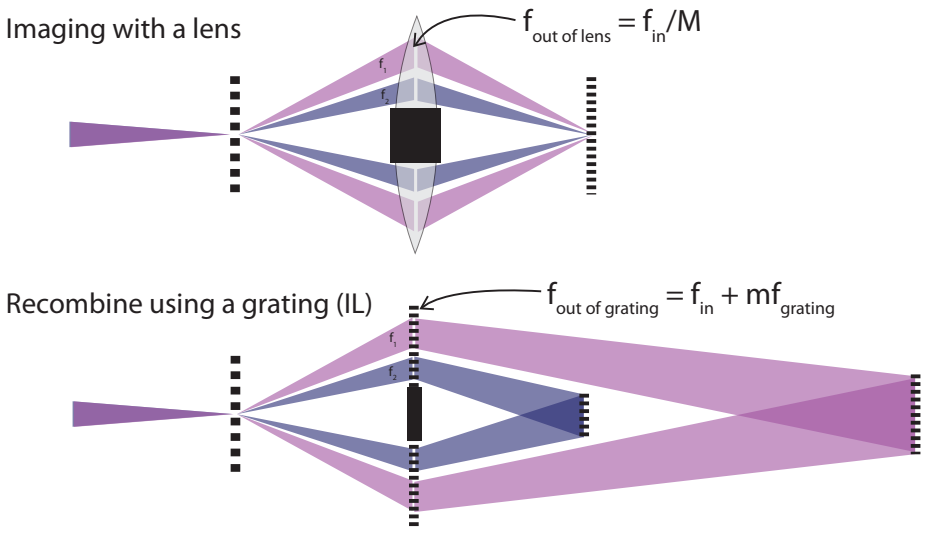

Figure 4.11: Structured patterning using two low-coherence approaches. This top-down view assumes spatially limited illumination (see Section 4.8.1) and that the object is a pair of vertically stacked sinusoidal gratings with spatial frequencies f_1 and f_2 ($f_1 > f_2$). The diffracted rays from grating f_1 (f_2) are labeled f_1 (f_2) and are colored pink (purple). Top: a lens scales incoming spatial frequencies so that all object frequencies are imaged to a single plane. Bottom: the grating shifts incoming spatial frequencies so that only one object frequency is imaged to a single plane.

In addition to the spatial filtering properties of the recombination grating, the illumination itself plays a role in the formation of the zero-LER aerial image. As described in Section 4.10.1, the fringe pattern in the two-grating interferometer can be viewed as a linear superposition of coherent interference patterns from each mutually incoherent spatial carrier in the illumination field. As each illumination angle interacts with a different area of the gratings, no two input angles produce the same coherent interference pattern. Through illumination incoherence, line imperfections in the gratings become completely homogenized in the final aerial image.

4.12 Utility of this work

In 2006, the feasibility of constructing a functioning interference lithography (IL) tool based on the incoherent two-grating platform was unknown. The work presented here has made it possible to write a value proposition for the potential customers of a commercialized incoherent IL tool: For EUV resist developers who want to test resists in house, the incoherent two-grating interferometer is a stand-alone exposure tool that can print line-space patterns down to 15 nm half-pitch using available sources, gratings and nano-positioning stages. Unlike interference tools that use coherent light, the incoherent tool produces a very low LER aerial image that is advantageous for resist characterization.

In this chapter we showed that a tool using a 13.5 nm Energetiq Z-pinch EUV source [21] with full NA = 0.07 and $1/f_1 = 30$ nm, $1/f_2 = 15$ nm, and $z_1 = z_2 = 20$ mm can operate with 15 μrad of tilt error, 1% grating pitch mismatch, and up to 20% grating noise energy while achieving a fringe contrast greater than 80%.

Chapter 5

Photoresist deprotection blur metrics

Contents

5.1 Motivation

Resists for extreme ultraviolet (EUV) lithography (λ =13.5 nm) are currently being optimized to meet the demanding specifications required at and beyond the 22 nm manufacturing node. At the present time the interplay between deprotection blur[i], line-edge-roughness (LER), and other factors contributing to the overall performance of EUV resists is not well understood. In practice, small perturbations in resist or process parameters almost always produce observable changes in printing performance yet the explanations for the observed changes are often speculative at best. Much of the difficulty is that we typically judge resist performance based on the final printed wafer, which of course is a product of the entire coating, exposure, and development process. Deconvolving the effects of the resist, moreover the effects from different constituents within the resist, is often a difficult task. In an attempt to better understand EUV resists, and how to improve them, there has been a large effort to develop resist metrics that can deconvolve the effects of deprotection blur, LER and other factors in producing observed performance changes.

5.2 Can we make deprotection blur quantifiable?

While resist sensitivity and LER are easy to quantify and compare based on direct observation of printing results, the determination of deprotection blur is less straightforward. Resist deprotection is the result of chemical, diffusive, and stochastic processes that occur during chemical amplification. Owing to the difficulty in generating deprotection models based on first principles, several authors [24, 25] have turned to a simplified point-spread function (PSF) linear systems approach to modeling resist exposure dynamics. The idea behind this approach is to lump all of the complicated chemical and diffusion processes during amplification into an effective blur function that converts the incident aerial image to a latent deprotection profile through a simple convolution process. Models based on the linear-systems PSF approach are convenient because they provide an intuitive link to the resist resolution limit.

[i]The average width of the volume of resist polymer that is rendered dissolvable (and thus removed during development) by a single photo-generated acid

5.3 Two new deprotection blur metrics

Figure 5.1 shows two observations that led to the development of the deprotection blur metrics that are described in this chapter. It was found that in general, improvements in line-space patterning (i.e., better line-edge-roughness (LER) and smaller supported pitches) are correlated with improved contact-hole exposure latitude and improved printing fidelity of sharp corners. Moreover, the PSF blur model predicts that reducing blur improves exposure latitude and the patterning fidelity of corners (see Figure 5.2). Given these consistent trends, we were compelled to investigate the development of blur metrics based on the contact-hole and corner features.

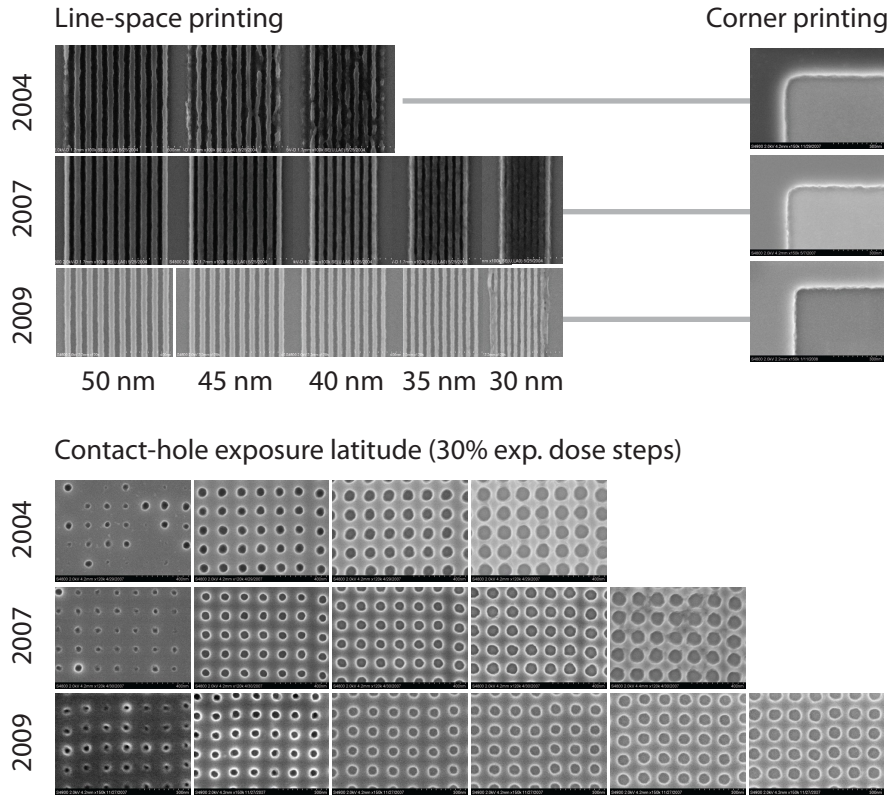

Figure 5.1: Through-pitch line-space printing (top left), corner printing (top right), and through-dose contact-hole printing (bottom) in the 2004, 2007, and 2009 SEMATECH Berkeley MET calibration resists. Top left: 50-30 nm half-pitch 1:1 line-space printing. Top right: the corner of a 700 nm elbow. Bottom: 50 nm 1:1.5 contacts, 30% dose steps. Observe that improved line-space printing (top left), improved corner printing (top right) and increased exposure latitude (bottom) occur simultaneously.

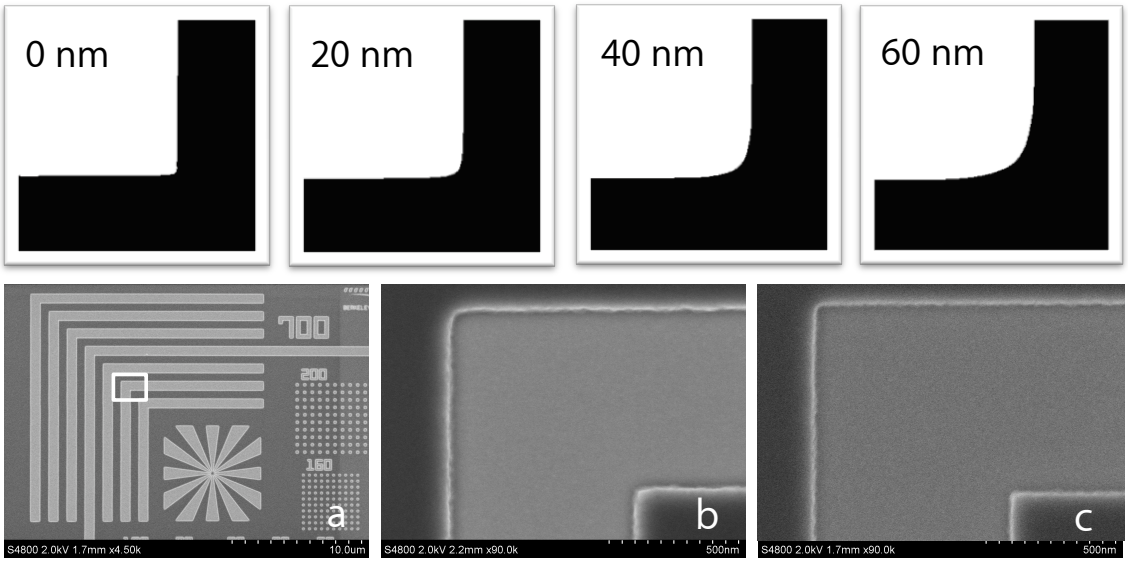

Figure 5.2: Top: Modeled corner rounding of a 700 nm elbow pattern assuming 0 nm, 20 nm, 40 nm, and 60 nm of resist blur. Bottom: (a) SEM image of a 700-nm elbow pattern. The white box shows the zoomed region in subfigures b and c; (b) 700 nm elbow printed in the 2004 SEMATECH Berkeley calibration resist; (c) 700 nm elbow printed in the 2009 SEMATECH Berkeley calibration resist.

5.3.1 The contact-hole blur metric

Determining resist blur with the contact-hole metric begins by shooting a focus-exposure matrix (FEM) of contact-hole arrays so that the FEM contains under-dosed contacts at the low doses and almost cleared contacts at the high doses. At each dose, SEM images of the best-focus contact-hole arrays are captured and offline software [26] is used to measure the average printed contact diameter. With these data, an experimental printed diameter (PD) vs. dose curve is generated (see Figure 5.3 left). Next, the aerial image is computed from a perfect (ideal) mask, using interferometrically measured aberrations [27] and assuming perfect focus. We then apply the HOST [28] PSF resist blur model to the aerial image, emulating a programmed level of resist blur. If desired, modeled PD vs. relative dose curves for a full range of resist blurs can be generated (see Figure 5.3 middle). To extract the blur of the tested resist, the programmed resist blur is varied until the mean-squared-error between the modeled and experimental PD vs. dose data is minimized (see Figure 5.3 right).

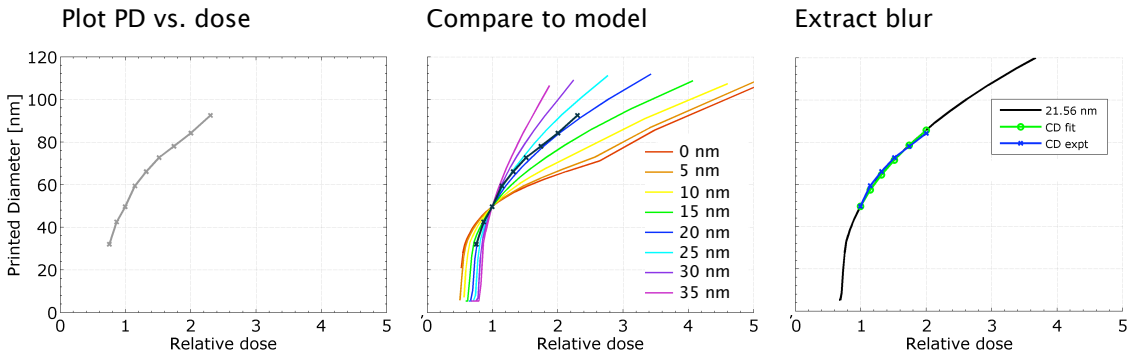

Figure 5.3: Left: experimental printed diameter (PD) vs. relative dose. The dose reference is the dose that coded 50 nm features print at 50 nm. Middle: experimental PD vs. dose curve compared to modeling curves assuming 0 - 35 nm of resist blur. Right: A resist blur of 21.5 nm minimizes the mean-squared-error between the modeling and experimental curves.

5.3.2 The corner rounding blur metric

The determination of resist blur with the corner metric begins by capturing a SEM image of a corner of a large printed feature. The SEM image is fed into commercial software [26] that determines the radius of curvature of the printed corner[ii]. Next, the aerial image is computed from a SEM image of the mask corner and a HOST [28] PSF resist blur is applied to the aerial image, emulating a programmed level of resist blur. To extract the blur of the tested resist, the modeled resist blur is varied until the modeled corner radius matches the experimental corner radius. Once the model data has been generated for a full range of deprotection blurs, the blur extraction step becomes a table look-up provided that mask and illumination conditions match those used for modeling.

5.4 Aerial image modeling limitations

The contact-hole and corner rounding blur metrics both require the ability to accurately model the aerial images that create the experimental patterning data. In 2004 the wavefront error of the SEMATECH Berkeley MET was measured using lateral shearing interferometry (LSI)

[ii]The measured radii of several identically coded features are generally averaged to mitigate error sources from mask imperfections and measurement uncertainty in the corner measurement software

[27]. The wavefront error of the MET was measured at 0.55 nm RMS (λ_{EUV}/24.5) in a 37-term annular Zernike series with an error of 0.07 nm RMS (λ_{EUV}/190), i.e., the error in the measurement of the wavefront error is 0.07 nm RMS. Consequently, there is a $\approx 13\%$ RMS error in the MET wavefront that is used in commercial aerial image packages, i.e., PROLITH [29]. In addition, FEM focus steps are typically on the order of 50 nm, meaning the random variable associated with defocus of the best-focus row in the FEM is uniform on the interval [-25, 25] nm with a standard deviation of 14 nm. As exposure conditions are always uncertain to some extent, it is of interest to investigate the impact this has on the ability of PSF-based resolution metrics to extract credible resolution numbers.

5.4.1 Sensitivity to aberration uncertainty

The measured wavefront error of the SEMATECH Berkeley MET can be written as a linear superposition of Zernike polynomials. In this representation, generating random aberration maps within the error-bars of measurement is as easy as making small, random changes to the Zernike coefficients. To investigate each metric's sensitivity to aberration uncertainty, we add 10%, 20% and 30% RMS random noise to the first 37 Zernike coefficients of the measured wavefront and generate aerial images of the test features using ten random aberration maps within each noise level. We then apply the HOST [28] PSF resist blur model to the aerial images, emulating a programmed level of resist blur.

Figure 5.4 (b) shows several PD vs. relative dose curves. Each gray level corresponds to a different resist blur, and within each resist blur are ten curves corresponding to each of the ten random aberration maps in the 10% RMS noise level. Figure 5.4 a, c, d show the equivalent plot for RMS noise levels of 0%, 20% and 30%. For all ten random aberration maps in a particular blur level, we find the blur whose nominal PD vs. dose curve (which assumes the nominal aberration map) is closest in terms of least-squared error (LSE). Table 5.1 summarizes the error-bars in extracted blur due to aberration uncertainty, which have been determined using this least-squares technique. Each blur has a slightly different error-bar, however they are close enough that we provide only the average value for each noise level. A similar procedure is followed for the corner rounding metric [30]; these results are also summarized in Table 5.1

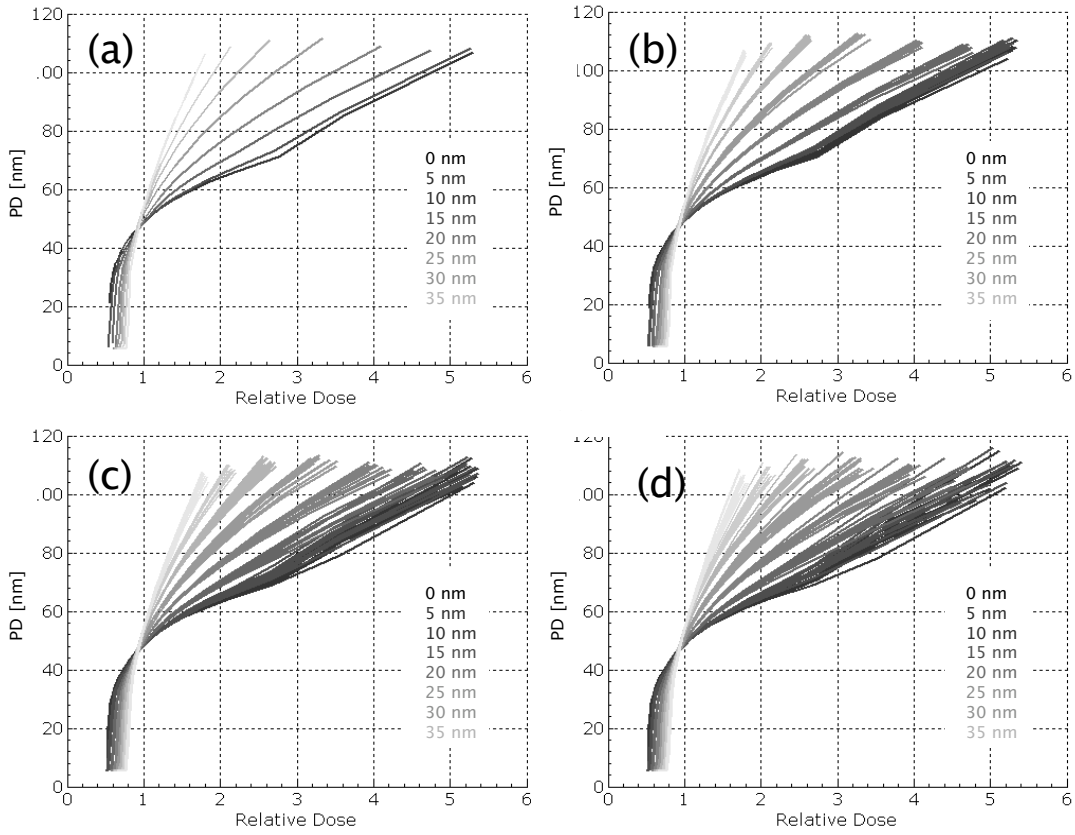

Figure 5.4: Modeled printed diameter (PD) vs. relative dose curves for the contact metric aberration study. Each graylevel corresponds to one blur. There are eight blurs spanning 0-35 nm in 5 nm steps. Each blur level contains 10 curves; one from each of the 10 randomly generated optical aberration maps within a given aberration noise level. Plots a, b, c, and d are associated with RMS aberration noise levels of: 0%, 10%, 20%, and 30%, respectively.

Table 5.1: Error-bars in extracted blur due to aberration uncertainty[a]

RMS noise	Contact P2V	Contact σ	Corner P2V	Corner σ
10%	1.4	0.5	1.2	0.4
20%	2.8	0.9	2.2	0.7
30%	4.9	1.5	3.4	1.1

[a] All error bars are in nm of blur

5.4.2 Sensitivity to focus uncertainty

To investigate each metric's sensitivity to focus uncertainty, we generate aerial images of the test features at several levels of defocus spanning the [-50 50] nm range. Treating each defocused aerial image as we treated each randomly aberrated aerial image we generate Figure 5.5 which is the equivalent to Figure 5.4 except that within each resist blur are five curves corresponding to five values of defocus: -50 nm, -25 nm, 0 nm, 25 nm, and 50 nm. Using the same least-squares technique described above, we find that the error-bars in extracted blur due to focus uncertainty are 1.7 nm peak to valley and 0.8 nm RMS for the contact metric and 1.6 nm peak to valley and 0.8 nm RMS for the corner rounding metric.

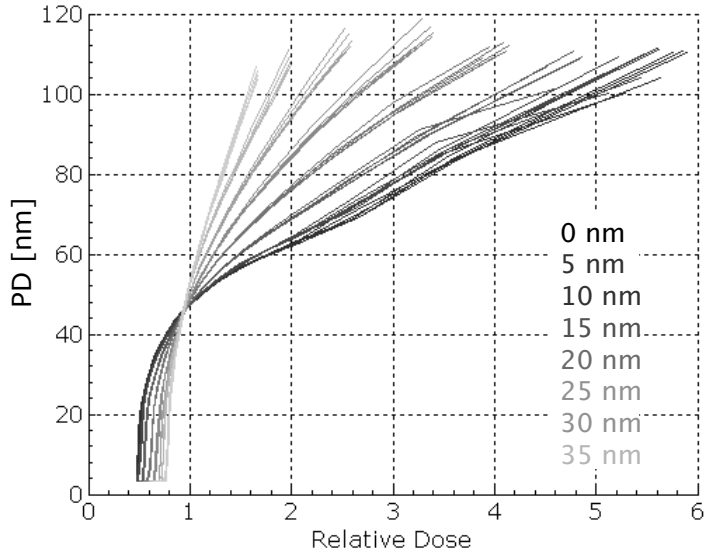

Figure 5.5: Modeled printed diameter (PD) vs. relative dose curves for the contact metric focus study. Each graylevel corresponds to one blur. There are eight blurs spanning 0-35 nm in 5 nm steps. Each blur level contains 5 focus curves associated with [-50, -25, 0, 25, 50] nm of defocus.

5.4.3 Discussion

The error bars from the 20% RMS noise level should provide an upper limit on error in extracted blur due to incomplete aberration knowledge at the Berkeley facility. Likewise, the error bars based on a [-50, 50] nm defocus range should provide an upper limit on error in extracted blur due to focus uncertainty if 50 nm focus steps are used during the exposure.

A the Berkeley facility the error the extracted blur is contributed to almost equally by wavefront error uncertainty and wafer focus uncertainty. Assuming independence, the error-bars associated with wavefront error uncertainty and focus uncertainty can be added in quadrature to estimate the total error in extracted blur for the contact and corner rounding metrics. The quadrature addition gives error-bars of 3.3 nm peak-to-valley and 1.2 nm RMS for the contact metric and 2.8 nm peak-to-valley and 1 nm RMS for the corner metric. We believe these error bars should be applicable with resist models other than the HOST model owing to the fact that they are associated with aerial image prediction limitations that are not unique to the resist model.

5.5 Sensitivity to experimental errors

The error-bars described above can be attributed to factors that limit the ability to accurately model the aerial image at the wafer surface in a given exposure. In addition to these error sources, there are several experimental factors that can potentially affect the result of a blur measurement. The most obvious of these are: picking the best-focused row from a FEM, SEM focus, and SEM image analysis. Below we will analyze each of these in detail and determine how they affect the blur determined with the contact-hole metric.

5.5.1 Picking the best-focus row in a FEM

With 50 nm focus steps, the random variable associated with aerial image defocus of the best-focus row in the FEM is uniform on the interval [-25, 25] nm. In practice, it is sometimes difficult to decide which row in the FEM is truly at best focus. Focus drifts through dose are periodically found to shift the best focus row by one step from the lowest to highest doses. Often times the true focus of the exposure tool is right between focus steps of the wafer, causing neighboring rows in the FEM to print almost identical through-dose data sets. Even for the most experienced persons, determining the best-focus row is routinely somewhat subjective.

A potential issue with the contact-hole metric is that it assumes the experimental data is from the best-focus row in the FEM. If the experimental data is taken from one row above

or below the best-focus row of the FEM, the random variable associated with aerial image defocus of the experimental set is uniform on the interval [75, 25] nm or [-25, -75] nm, respectively. One would expect that this level of defocus could alter the result of a resolution extraction when comparing the experimental data to modeled data generated with 0 nm of defocus.

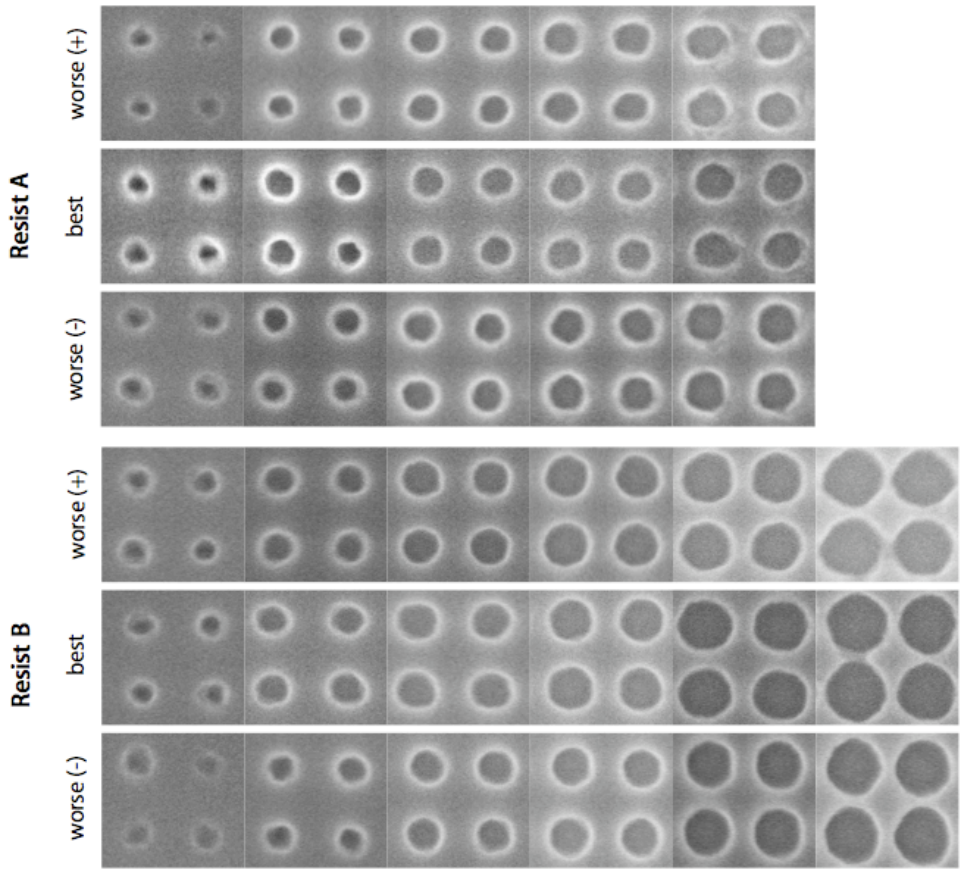

Figure 5.6: Through-dose SEM images of coded 1:1.5 50-nm dark-field contact-holes printed in resists A (top) and B (bottom) at three focus steps separated by 50 nm each. The relative dose step between SEM images is 1.32. The rows labeled 'best' are considered by the author to be in the best focus.

Figure 5.6 shows SEM images of printed contact-holes through-dose at neighboring focus steps separated by 50 nm in two resists; the relative dose step between images is 1.3 in both resists. The rows labeled 'best' in each figure are considered by the author to be in the best focus. This determination has been made by looking at printing characteristics at the lowest and highest doses. For doses that print well below the coded feature size, the best-focus row

generally has contacts with the cleanest edges and the least PD variation. At doses that print well above the coded feature size, the astigmatism of the SEMATECH Berkeley exposure tool [31] gives rise to shape changes in printed contacts at high doses, flattening them slightly in the horizontal and vertical directions on either side of best focus.

The through-dose contact-hole sets in Figure 5.6 have been analyzed and resist blurs have been extracted. The results are summarized in Table 5.2; we also show the results of another resist, resist H, although the corresponding SEM images are not included in the text. As expected, resists B and H show increases in reported blur on either side of best focus. For both of these resists the increase in extracted blur is in the range of the 1.2 nm RMS error-bar reported above. For resist A, however, we observe a 2 nm drop in extracted resolution in the defocused 'worse+' data set as compared to the 'best' data set. Perhaps the explanation is that the 'worse+' data is in fact the best focus data set.

Table 5.2: Extracted deprotection blur through focus.

Focus	Blur (nm)		
	Resist A	Resist B	Resist H
worse +	7.7	21.6	12.5
best	9.9	20.5	11.5
worse –	10.4	22.7	12.3

Overall the metric proves to be robust even with the ambiguity in experimental data selection. However, for resist formulations with small resolutions (i.e., resist A) these data show that cross-platform comparisons could benefit from repeated experimental trials to reduce the measurement uncertainty through statistical averaging.

5.5.2 SEM image edge-detection threshold

Once images are captured with the SEM they are analyzed using offline image analysis software to measure PD at a given intensity threshold. As one can imagine, changing the threshold level in the analysis software alters the result of the measurement. Table 5.3 shows the measured PD's for a through-dose set of SEM images at 5 different programmed threshold levels along with the extracted resolution for each threshold level. We observe no significant change in extracted resolution throughout the range of thresholds that are examined. These

data suggest that the contact-hole metric is relatively immune to changes in edge-detection schemes, however, cross-platform comparisons should make an effort to be as consistent as possible with image analysis methods.

Table 5.3: Blur dependence on edge-detection threshold.

Threshold (relative)	Measured PD through dose (nm)								Blur (nm)
0.40	41.8	48.4	53.1	57.5	63.1	67.2	72.3	81.7	13.9
0.45	43.0	49.8	54.3	58.6	64.4	68.5	73.6	83.1	13.9
0.50	44.2	51.0	55.5	59.9	65.6	69.6	74.9	84.4	13.9
0.55	45.4	52.1	56.6	61.0	66.7	70.7	76.0	85.7	14.0
0.60	46.4	53.2	57.5	62.1	67.8	71.8	77.1	86.1	14.2

5.5.3 SEM focus

There are many factors that affect the quality of an image captured in a SEM [6]. Assuming that the SEM is well-aligned and is properly corrected for stigma, perhaps the two biggest factors affecting image quality are signal-to-noise ratio and electron beam focus. Figure 5.7 shows through-SEM-focus images of coded 1:1.5 50 nm contacts printed at a dose slightly larger than dose-to-size. The images we show at each SEM focus are subsets of larger raw images containing > 28 contacts each; reported PD values are the average and variance (1σ) of 20 central contacts in each original image. To avoid SEM electron beam dosing between image captures a different (neighboring) part of the field is imaged at each SEM focus. For calibration purposes we also image five neighboring field sites at the middle SEM focus (see Figure 5.8) and perform a field uniformity test. The average and variance of the average printed diameter (PD_{avg}) at each cite are 60.9 and 0.4 nm, respectively, indicating that the 5 nm shift in measured PD through SEM focus is an artifact of focusing and not a variation in the contact sizes between the imaged subfields.

Ultimately we are interested in how SEM focus affects the extracted resist blur. We have imaged a through-dose set of contacts at three levels of deliberate SEM defocus; Figure 5.9 shows subsets of each through-dose set. Each set of data has been analyzed and curve-fit in the manner described in Sec 5.3.1 to extract a resist blur. The extracted resist blurs for best, worse, and worst SEM focus as defined in Figure 5.9 are 11.3 nm, 12.5 and 17.7 nm,

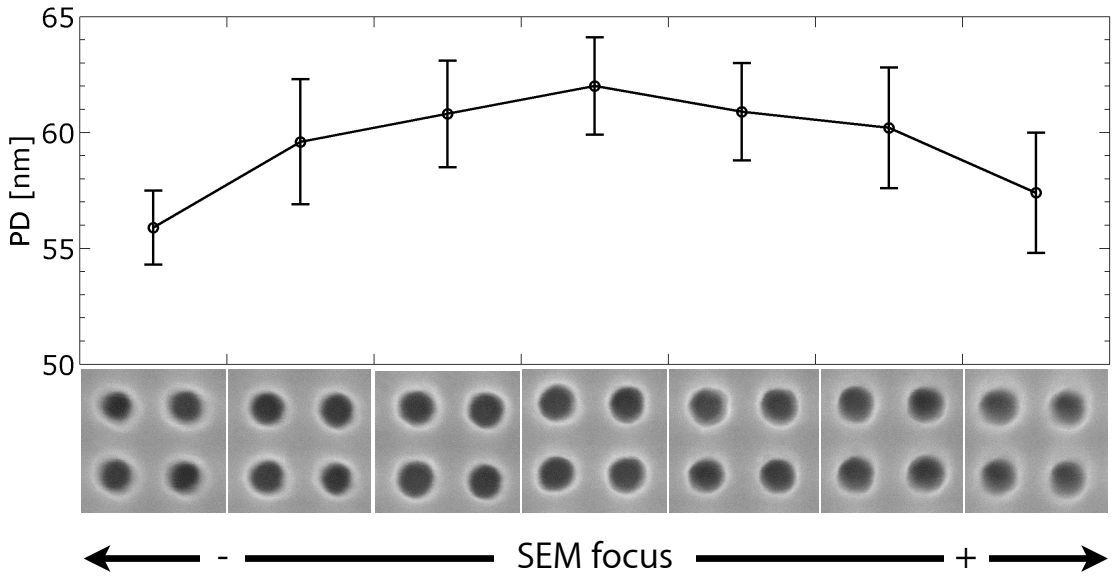

Figure 5.7: Measured printed diameter (PD) of coded 1:1.5 50 nm dark-field contact holes as a function of SEM focus. The plotted values are the average and standard deviation (1σ) of 25 neighboring contacts. To avoid SEM electron beam dosing between image captures a different part of the field is imaged at each SEM focus.

respectively.

The convolution process that generates the defocused SEM image does two things to affect the extracted blur: 1) it causes larger changes in measured PD for small contacts than it does for large ones (proximity effects), which steepens the slope of the PD vs. dose curve, and 2) it changes the dose-to-size value, which reshapes the PD vs. dose curves differently during dose normalization. The combination of these two effects gives rise to significant changes in the PD vs. relative dose curves through SEM focus.

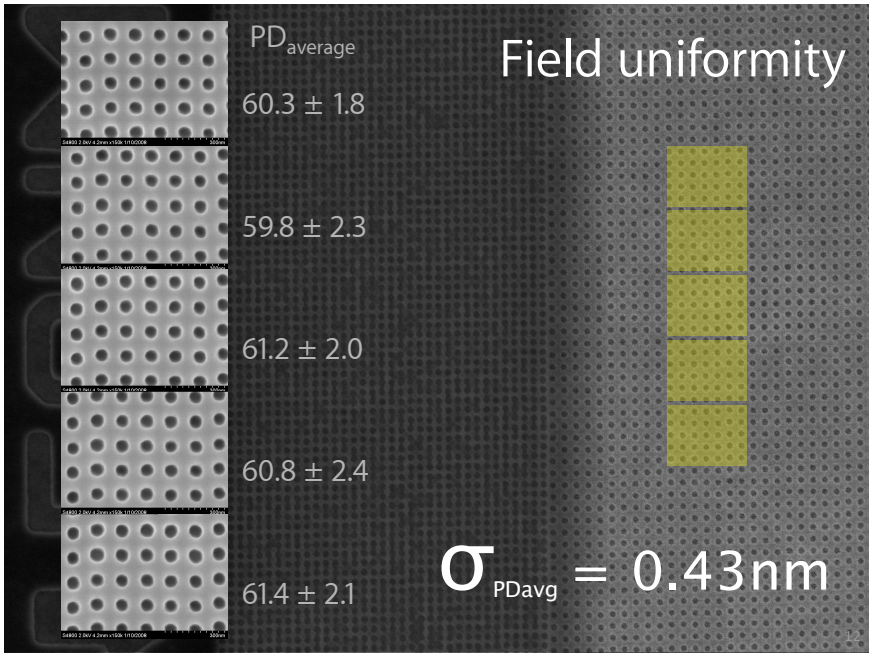

Figure 5.8: Contact field uniformity test. The variance in the average printed diameter (PD_{avg}) of contacts at neighboring sites in the field is 0.4 nm.

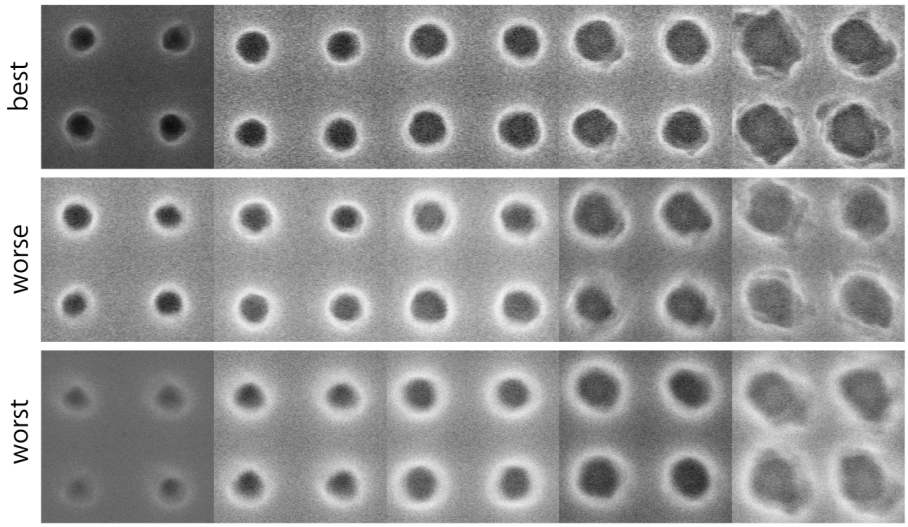

Figure 5.9: Through-dose images of coded 1:1.5 50 nm dark-field contact-holes at three levels of deliberate scanning electron microscope (SEM) defocus. The top row is well-focused and defocus increases in the downward direction. To avoid SEM electron beam dosing between image captures neighboring parts of the field are imaged at each SEM focus. Relative dose steps between images are 1.32.

5.5.4 Reproducibility

The SEMATECH Berkeley MET printing facility is known for it's long-term stability in terms of illumination conditions, optical aberrations, and focus control [1]. Here we examine the the reproducibility of the contact-hole metric by repeating identical experiments on different days for three resists. For a given resist formulation, the base coat, resist thickness, post-application bake, post-exposure bake, and illumination conditions remain fixed throughout the experiments. In each repeated trial we have made every effort to ensure SEM images are in focus, with emission current fixed throughout each through-dose set. All SEM images are captured by the same person (me). All PD measurements are made at the same threshold level (0.5) in the analysis software.

Table 5.4 shows the extracted blurs of Resist E, Resist F, and Resist G for each measurement trial. In each resist, the standard deviation of the blurs from repeated measurements is less than the 1.2 nm RMS error bar associated with aerial image modeling limitations. These results demonstrate that the modeled error bars adequately capture all of the experimental uncertainties associated with blur metrics.

Table 5.4: Reproducibility of the contact-hole blur metric.

Resist E	Resist F	Resist G
22.1	16.7	26.6
20.5	16.2	25.5
20.9		
Exposure dates		
12-04-2007	10-25-2007	09-30-2006
12-05-2007	10-26-2007	04-06-2007
09-07-2007		

5.6 Wrap up

Two new PSF-based resist blur metrics have been developed. Remarkably, both blur metrics are robust against aerial image modeling limitations and experimental error sources: the modeled 1.2 nm RMS error bar due to aerial image modeling limitations was larger than the standard deviation of repeated blur extractions on three different resists. The ability to quantify resist blur, and do so as resist constituents and process parameters are varied, will likely be an asset to resist developers as they aim to better understand resists and prepare them for commercialization at the 22 nm technology node.

Chapter 6

EUV RLS Experiments

Contents

6.1 Introduction

Current EUV resists do not simultaneously meet the resolution, line-edge roughness (LER), and sensitivity requirements (RLS requirements) for the 22 nm manufacturing node and beyond. Some of these shortcomings are reflected in the fact that the community still does not unanamously agree on several major resist-related issues; case and point, the pleathora of differing opinions about the factors that currently limit resist performanace. In all fairness, EUV resist development is not easy: the feedback loop can be delayed (the motivation for the interference lithography work) and, until recently, it lacked simultaneous quantification of the RLS parameters. Arguably, resist progress will be accelerated as resolution metrics are integrated into the feedback loop and a better understanding of the RLS interaction is achieved.

In this chapter we describe a series of experiments where all three RLS parameters are measured as resist constituents and process parameters are varied. Three independent attributes are studied: base weight percent, photoacid generator (PAG) weight percent, and post-exposure bake (PEB) temperature. Following a description of background information

and experimental methods, we present and analyze the results of these experiments and use them to drive a lengthy discussion about the relationships between deprotection blur, LER, and shot noise in EUV photoresists.

6.2 Background

6.2.1 The relationship between deprotection blur and LER

It is often claimed that the LER of an isolated edge can be improved by increasing resist deprotection blur. Fundamental to this assumption is that the "pixel size" relevant for counting statistics is determined by the average size of the diffusion or blur sphere. The resulting conclusion is that larger resist blurs lead to bigger counting bins, reduced shot noise, and reduced LER. However, Steenwinckel et al. [32] has shown that when patterning dense features with a half-pitch close to the width of the deprotection blur, increasing the deprotection blur can adversely impact LER. In this case, the advantage of having a larger detection bin may not overcome the disadvantage of the reduced deprotection contrast brought about by the large blur. Ultimately the aerial image properties (contrast, dose, and geometry) should dictate how the deprotection blur affects the LER of a line. In this respect, it is difficult to make generalizations that apply to all situations.

6.2.2 The relationship between sensitivity and LER

Two separate experiments have measured LER as base wt. % is varied in EUV photoresist [32, 33]. In both of these experiments, improvements in LER with increased base wt. % are attributed to reductions in shot noise that are correlated with reductions in photospeed (due to added base). The argument supporting this claim states: base inhibits chemical amplification → increasing base reduces chemical amplification → reducing chemical amplification requires an increase in photogenerated acid to maintain the same level of deprotection → increasing photogenerated acid requires more photons to be absorbed during the exposure → more photons absorbed means longer exposures, more dose, less shot noise.

6.3 Experimental methods

6.3.1 Imaging conditions

All resists were exposed to EUV (λ = 13.5 nm) radiation at the 0.3 numerical aperture SEMATECH Berkeley MET printing facility at the Advanced Light Source at Lawrence Berkeley National Laboratory (LBNL) using conventional $\sigma = 0.35 - 0.55$ annular illumination [1].

6.3.2 Features and metrology

All SEM analysis was performed at LBNL on a Hitachi S-4800 with a working distance of 2 mm and an acceleration voltage of 2.0 kV.

Corner rounding blur metric

For the corner rounding blur metric we print 700-nm dark field elbow patterns (see Figure 6.1) and use the printed corners with 270 degrees of resist remaining after development (see Figure 6.2). Radius of curvature values used for blur extraction are the average radius of seven identically coded corners (see Figure 6.1a). The radius of a single printed corner is determined using the "area fit" algorithm in an offline SEM image analysis package [26].

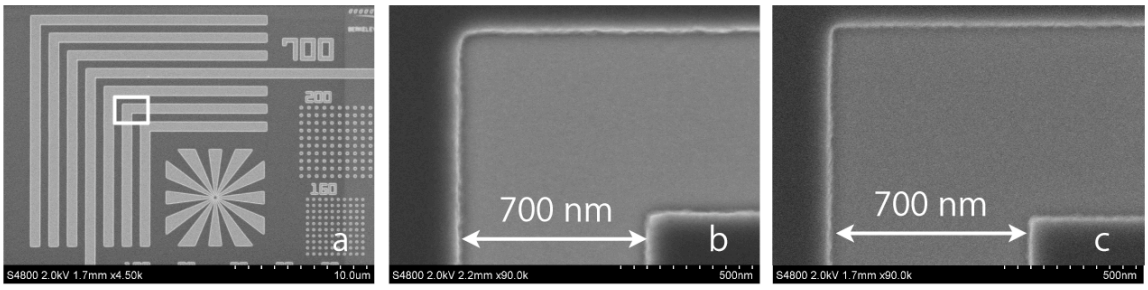

Figure 6.1: (a) SEM image of a 700-nm elbow pattern. Darker areas are the resist that remainins after development. The white box shows the zoomed region in subfigures b and c; (b) 700 nm elbow printed in Rohm and Haas XP 5496I photoresist; (c) 700 nm elbow printed in TOK EUVR P1123 photoresists.

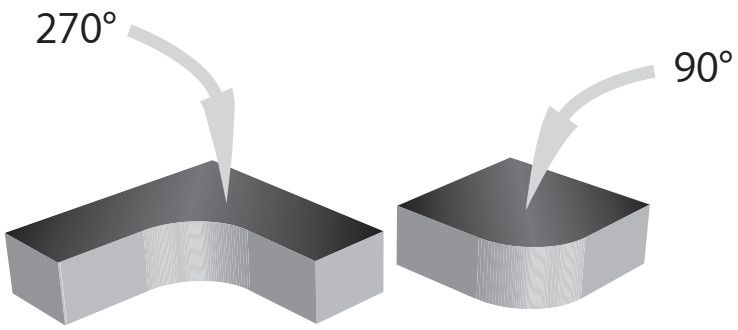

Figure 6.2: Dark material is resist. Left: a rounded corner with 270° of resist remaining after development. Right: a rounded corner with 90°of resist remaining after development.

Contact hole blur metric

For the contact hole blur metric we print 50 nm diameter, 125 nm pitch (1:1.5 duty cycle) dark-field contacts. Printed diameter (PD) values used for blur extraction are the average PD of 25 central contacts printed in an 10 x 10 array. All contact metric error sources identified in the previous chapter have been minimized by adhering to the suggested process guidelines: all SEM images are well focused, emission current is fixed throughout each through-dose set, all images are gathered by the same person, and all PD measurements are made at the same threshold level.

Patterned LER and intrinsic LER

To measure patterned LER[i] and intrinsic LER[ii] we print 50 nm half-pitch and 100 nm half-pitch 1:1 line space patterns, respectively. The reported LER values are obtained using a 3x3 process window (three dose steps and three focus steps) around the center-dose center-focus site in the focus-exposure matrix. For 100 nm 1:1 features the SEM magnification is set to 100k providing 6 patterned lines in each SEM image. The reported LER magnitude is the average of the 54 single-line LER values in the process window and the reported LER uncertainty is the 3σ standard deviation of the 54 single-line LER values divided by the square root of the number of lines in the process window (independence between single-line LER

[i]The LER of dense features whose pitch is close to the resolution limit of the resist

[ii]The LER of a feature much larger than the resolution limit of the resist

values is assumed). For 50 nm features the SEM magnification is set to 150k providing 8 patterned lines per SEM image and 72 lines in the process window. The spatial frequency spectrum of a single-line LER measurement is confined to a passband with a minimum period of 10 nm and a maximum period of 834 nm (the height of the SEM image).

6.3.3 Patterning limit and E-size

The patterning limit is defined as the smallest sized 1:1 lines that pattern in resist without excessive collapse, bridging, or top loss. E-size is defined as the dose required to print 50 nm 1:1 features at their coded size at best focus; we use the dose calibration adopted at SPIE 2008 [34].

6.4 Data: Base and PAG

6.4.1 Resists

We have tested three through-base resist series provided by Rohm and Haas based on commercially available XP 5435, XP 5271, and XP 5496 resist platforms as well as a through-base and through-PAG resist series provided by the University at Albany [35] based on the experimental EH27 platform. Table 6.1 summarizes the resist thickness, post-application bake (PAB), PEB, and development parameters for each resist platform; all process parameters were recommended by the resist suppliers. Four-inch hexamethyldisalizane-primed wafers were used for all experiments and all wafers were developed using a single puddle of Rohm and Haas MF26A developer.

Table 6.1: Resist process parameters

Supplier	Resist	Thickness (nm)	PAB (°C)	PAB (sec)	PEB (°C)	PEB (sec)	Dev. time (sec)
Rohm and Haas	XP 5435 E,F,D,G,H	120	130	60	130	90	45
Rohm and Haas	XP 5271 J,K,D	80	130	60	120	90	45
Rohm and Haas	XP 5496 H,I,F,J	80	130	60	90	90	45
University at Albany	EH27 A,C,D,E,F,G,H	125	130	60	130	90	45

6.4.2 Results

Figure 6.3 shows SEM images of bright field 1:1 lines at best focus printed in Rohm and Haas XP 5435 resist with different base weight percents shown at the left of each row. Figures 6.4 and 6.5 show the corresponding data for the XP 5496 and EH27 platforms, respectively. Figure 6.6 shows the corresponding data for the EH27 PAG loading series. Table 6.2 summarizes the RLS metrics for each resist.

In general, increasing base weight percent improves patterning ability and LER. For example, an 8× increase of base weight percent in XP 5435 reduces the patterning limit from from 50 nm to 36 nm, reduces 100 nm LER from 8.2 nm to 4.1 nm, and reduces 50 nm LER from 8.2 to 5.0 nm. Increasing PAG weight percent also influences printing performance, however the effects are both good and bad. On the upside, a 2× increase of PAG weight percent in EH27 improves 50 nm LER from 7.4 nm to 5.8 nm. On the downside, the same change in PAG weight percent leads to merging of the semi-isolated (outer) lines.

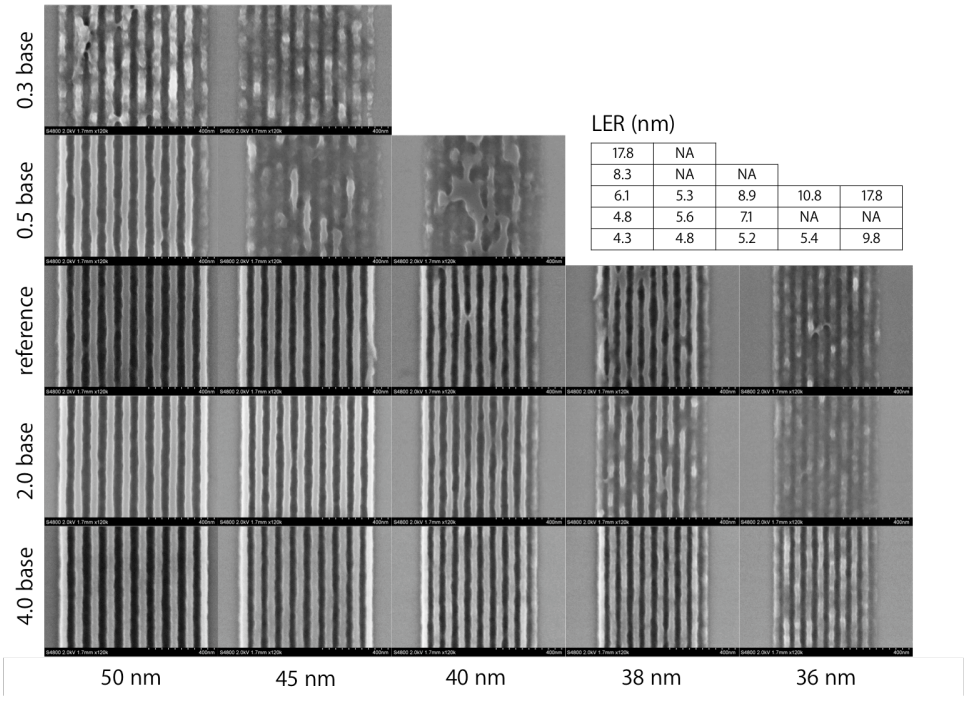

Figure 6.3: Base loading study. Through-pitch SEM images of bright field 1:1 lines printed in Rohm and Haas XP 5435 resist.

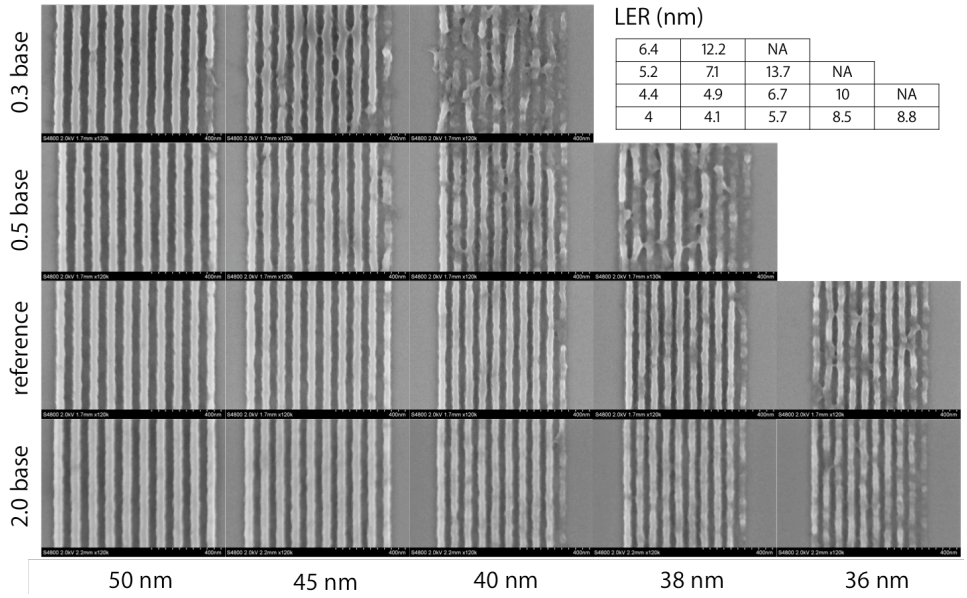

Figure 6.4: Base loading study. Through-pitch SEM images of bright field 1:1 lines printed in Rohm and Haas XP 5496 resist.

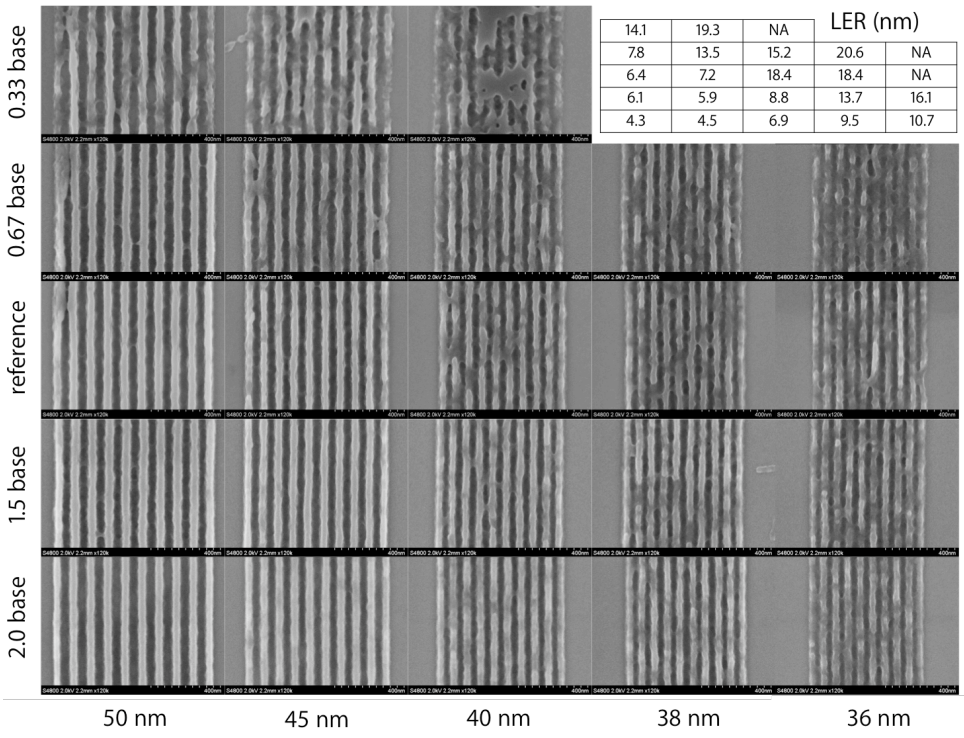

Figure 6.5: Base loading study. Through-pitch SEM images of bright field 1:1 lines printed in EH27 resist.

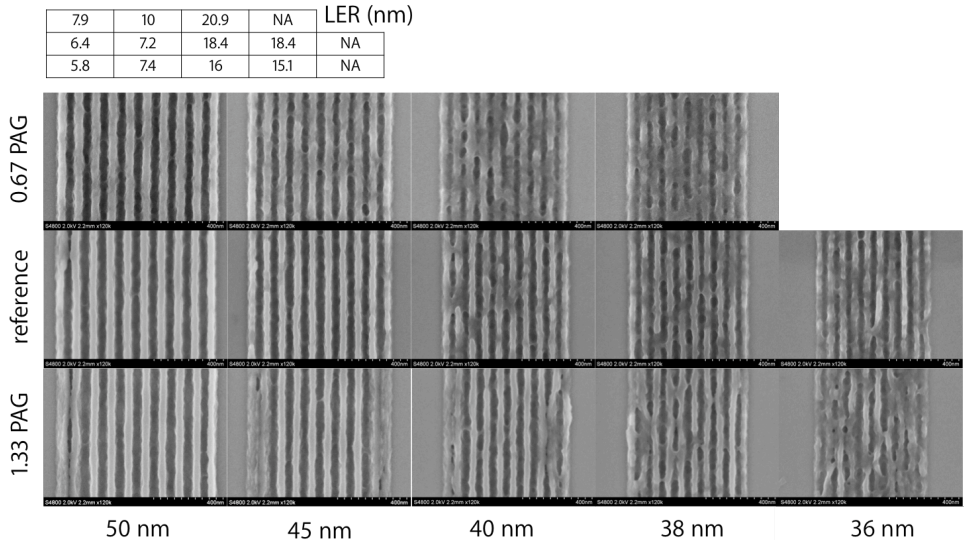

Figure 6.6: PAG loading study. Through-pitch SEM images of bright field 1:1 lines printed in EH27 resist.

Table 6.2: Base/PAG study RLS performance metrics

Resist	Label	E-size	Blur (nm)		Pattern limit	LER (nm)	
		(mJ/cm^2)	Contact	Corner	(nm 1:1)	50 nm 1:1	100 nm 1:1
XP-5435-E	0.3 base	1.6	32.1	36.8	52	13.7 ± 0.6	8.2 ± 0.1
XP-5435-F	0.5 base	2.3		38.4	50	8.2 ± 0.4	8.1 ± 0.1
XP-5435-D	Reference	3.2	31.3	35.0	42	6.1 ± 0.3	5.7 ± 0.2
XP-5435-G	2.0 base	6.4	26.2	33.8	40	5.5 ± 0.3	4.6 ± 0.1
XP-5435-H	4.0 base	14.0	25.1	30.0	36	5.0 ± 0.3	4.1 ± 0.1
XP-5271-J	0.5 base	4.0	27.9	31.2	47	13.4 ± 0.8	8.0 ± 0.2
XP-5271-K	Reference	6.5	25.4	32.3	43	6.7 ± 0.3	5.3 ± 0.1
XP-5271-D	2.0 base	12.5	23.8	34.8	39	6.7 ± 0.2	5.2 ± 0.1
XP-5496-H	0.3 base	3.0	26.5	27.6	48	8.1 ± 0.3	7.6 ± 0.1
XP-5496-I	0.5 base	4.7	26.4	28.1	44	7.9 ± 0.3	6.9 ± 0.1
XP-5496-F	Reference	7.6	24.6	27.0	38	6.5 ± 0.3	5.8 ± 0.1
XP-5496-J	2.0 base	15.2	25.0	29.2	38	5.3 ± 0.3	5.0 ± 0.1
EH-27-C	0.3 base	1.9	17.0	33.2	52	13.4 ± 0.7	6.9 ± 0.1
EH-27-D	0.7 base	3.2	17.3	35.0	47	8.8 ± 0.4	5.8 ± 0.1
EH-27-E	Reference	6.4	16.7	37.0	43	6.4 ± 0.2	5.1 ± 0.1
EH-27-F	1.5 base	7.8	15.0	32.3	42	5.3 ± 0.1	4.9 ± 0.1
EH-27-G	2.0 base	10.7	17.1	36.6	39	4.5 ± 0.1	4.1 ± 0.1
EH-27-A	0.7 PAG	8.5	17.0	33.2	47	7.4 ± 0.7	5.7 ± 0.1
EH-27-E	Reference	6.4	16.7	34.0	43	6.4 ± 0.2	5.1 ± 0.1
EH-27-H	1.3 PAG	4.9	16.1	33.6	40	5.8 ± 0.1	4.9 ± 0.1

6.5 Data: PEB temperature

6.5.1 Resists

We tested TOK EUVR P1123 EUV photoresist. As suggested by the supplier, the resist was spin-coated and sotfbaked at 120 °C for 60 seconds to yield a film thickness of 60 nm on hexamethyldisalizane-primed 4 in. wafers. Following the PEB, resists were developed using a single puddle of Rohm and Haas MF26A for 60 seconds. PEB temperatures of 80 °C, 90 °C, 100 °C, 120 °C, and 130 °C were tested.

6.5.2 Results

Figure 6.7 shows SEM images of bright field 1:1 lines printed through pitch in TOK EUVR P1123 resist for PEB temperatures of 80 °C, 90 °C, 100 °C, and 120 °C. LER improvements with decreasing PEB are measurable (see inset in Figure 6.7). At the 120 °C PEB temperature the semi-isolated lines completely clear at 45 nm and there is a significant LER increase over the 100 °C wafer in the smaller features. Figure 6.8 shows SEM images of a 130 °C PEB contact wafer, highlighting the onset of a lift-off mechanism that occurs at this PEB temperature. Due to this drastic change in printing behavior, the 130 °C PEB data is omitted.

Table 6.3 summarizes the RLS performance metrics for the TOK EUVR P1123 PEB study. Each PEB experiment was performed at least twice to test reproducibility and the results of each independent blur measurement are summarized in Figure 6.9. Within the 80-100 °C PEB temperature range, the extracted deprotection blur for repeated trials stays within the reported 1.75 nm RMS error bars of the contact-hole metric. The blurs at the 120 ° PEB temperature, however, differ by more than the reported error bar so it is likely that at 120 °C were observing the onset of the transition mechanism that shows up at the 130 ° PEB temperature.

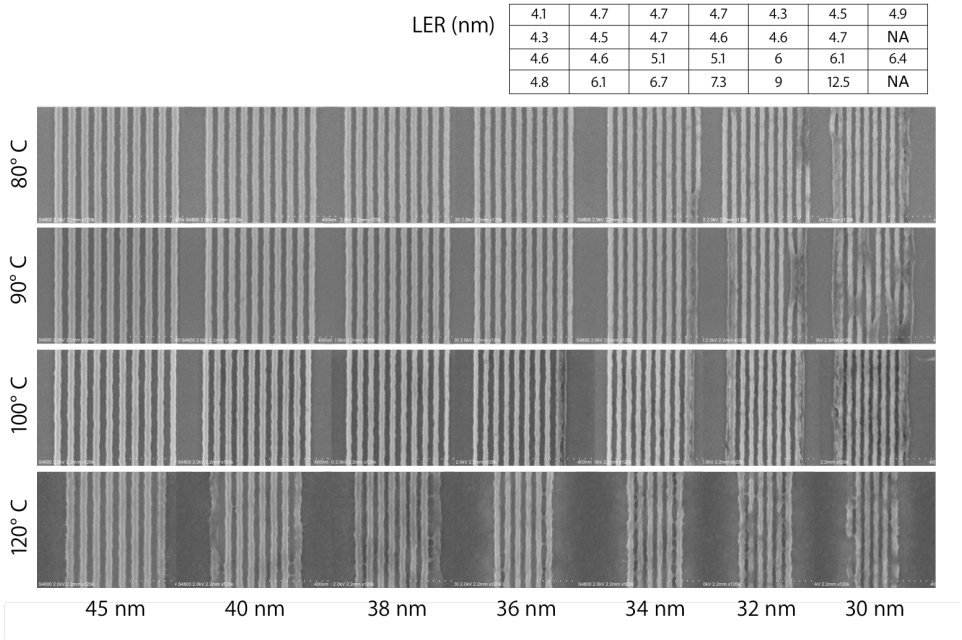

Figure 6.7: Through-pitch SEM images of bright field 1:1 lines printed in TOK EUVR P1123 resist with PEB temperatures of 80, 90, 100 and 120 °C.

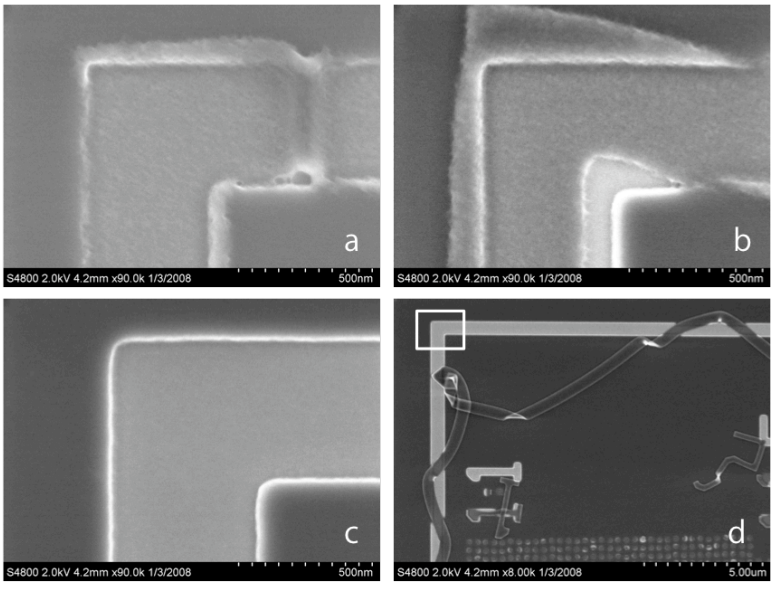

Figure 6.8: (a) - (c) are SEM images of a 500 nm elbow printed in TOK EUVR P1123 resist with a 130 °C / 90 second PEB with relative dose steps of 1.15 between each image; (d) is a zoomed out version of (c) where the field shown in (c) is outlined in white.

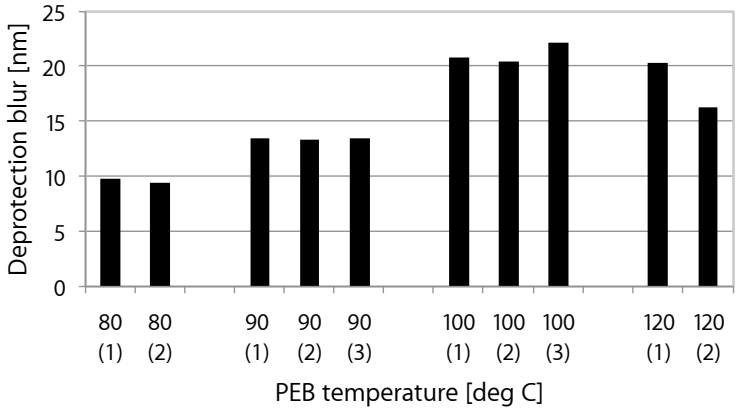

Figure 6.9: Measured deprotection blur of TOK P1123 resist with PEB temperatures of 80, 90, 100 and 120 °C. Each experiment was performed at least twice to test reproducibility.

Table 6.3: PEB study RLS performance metrics

Resist	PEB °C	E-size	Blur (nm)		Pattern limit	LER (nm)	
		(mJ/cm^2)	Contact	Corner	(nm 1:1)	50 nm 1:1	100 nm 1:1
P1123	80	11.9	9.7	11.0	28	4.3 ± 0.1	4.0 ± 0.1
P1123	90	8.8	13.5	15.1	28	4.5 ± 0.1	4.0 ± 0.1
P1123	100	8.2	21.1	22.3	19	4.8 ± 0.1	4.1 ± 0.1
P1123	120	6.8	18.3	18	30	4.8 ± 0.1	N/A

6.6 Discussion

6.6.1 Deprotection blur is largely insensitive to base, PAG wt. %

Of the three parameters that were tested (PAG weight percent, base weight percent, and PEB temperature) only the PEB temperature was found to significantly affect the deprotection blur. Prior to these experiments, it was expected that increasing base weight percent would reduce deprotection blur by consuming an increased fraction of acid by-products from the catalyzed deprotection reaction. However, in three out of four resists the deprotection blur stays within the error-bars of measurement as base is varied. Also, in the one resist where the deprotection blur does change (Rohm & Haas 5435), it only drops 19%[iii] as base weight percent is increased 13×.

6.6.2 PEB temperature: the RLS tradeoff

Reducing the PEB temperature of TOK EUVR P1123 photoresist simultaneously reduces its deprotection blur and sensitivity while improving 50 nm 1:1 LER and leaving 100 nm 1:1 LER unaffected. This interaction between the resolution (blur), LER and sensitivity (RLS interaction) warrants a thorough discussion of counting statistics. To make the problem tractable, two assumptions are made: 1) the photon "counting bin" is assumed a cube with width equal to the measured deprotection blur using the corner metric (i.e., $11 \times 11 \times 11$ nm^3 at 80 °C) , and 2) 50 % of the incoming photons are absorbed by the 60-nm-thick resist film. The amount of energy absorbed per counting bin in the 80 °C PEB wafer is:

$$\frac{11.9 \text{ mJ}}{\text{cm}^2} \cdot \frac{\text{cm}^2}{10^{14} \text{ nm}^2} \cdot \underbrace{\frac{11^2 \text{nm}^2}{1}}_{\text{bin area}} \cdot \underbrace{\frac{11 \text{ nm}}{60 \text{ nm}} \times 0.5}_{\text{frac. abs. in 11 nm}}$$

and in the 100 °C PEB wafer it is:

$$\frac{8.2 \text{ mJ}}{\text{cm}^2} \cdot \frac{\text{cm}^2}{10^{14} \text{ nm}^2} \cdot \underbrace{\frac{22^2 \text{nm}^2}{1}}_{\text{bin area}} \cdot \underbrace{\frac{22 \text{ nm}}{60 \text{ nm}} \times 0.5}_{\text{frac. abs. in 22 nm}}$$

[iii] This is the average of corner and contact data

Therefore, the 100 °C PEB wafer absorbs $\frac{8.2}{11.9} \cdot (\frac{22}{11})^3$ = 5.5 times more energy per counting bin than the 80 °C PEB wafer. If shot noise differences are directly mapped to LER differences, the 100 °C wafer should have $\sqrt{5.5}\times$ better LER than the 80 °C PEB wafer; however, these predictions oppose observed trends. If the height of the counting bin is stretched to the full 60 nm thickness of the resist, the 5.5× ratio is halved since the height ratio between counting bins (a factor of two) drops out of the math; however the prediction still opposes observed trends.

6.6.3 Deprotection blur can be a dominant LER contributor

Figure 6.10 (top) shows a cross section of a simulated aerial image of 100 nm and 50 nm 1:1 lines at the SEMATECH Berkeley MET with 0.35 - 0.55 annular illumination. The middle and bottom rows of Figure 6.10 show the latent image cross sections for 80 °C and 100 °C PEB temperatures, which are obtained by convolving the aerial image with 11 nm and 22 nm HOST resist blur kernels, respectively. The two-dimensional versions of the blurred aerial images are shown in Figure 6.11 (top), illustrating the degrading affect of the 22 nm blur on 50 nm 1:1 features. To include the affects of shot noise, each image is modified in the following manner.

1. The simulated latent image is copied and thresholded until the resulting binary features show up at their coded size. This level is called the "thresh-to-size" level.

2. The thresh-to-size level of the original image is scaled to the lowest experimental dose where the resist is rendered completely dissolvable to developer (E0)[iv]. This converts the latent image to an energy density profile that matches experimental conditions.

3. The energy density profile is down-sampled to a pixel size that matches the measured deprotection blur using the corner metric (i.e., 11×11 nm^2 at 80 °C), creating a map of incident energy density vs. counting bin.

4. The incident energy density at each bin is multiplied by the 50% absorption coefficient of the 60-nm-thick resist stack[v] and again by the cross sectional area of the counting

[iv] In general, E0 is a factor of two less than the E-size of 50 nm 1:1 features.

[v] Using the 60 nm thickness of the full stack (in place of 11-nm-thick and 22-nm-thick counting bins) assumes

bin. The result is the absorbed energy in each bin.

5. The absorbed energy in each bin is divided by the EUV photon energy, revealing the mean number of absorbed photons in each bin.

6. An *actual* number of absorbed photons in each bin is assigned by assuming each bin is a Poisson random variable with a rate equal to the mean photon number at that bin.

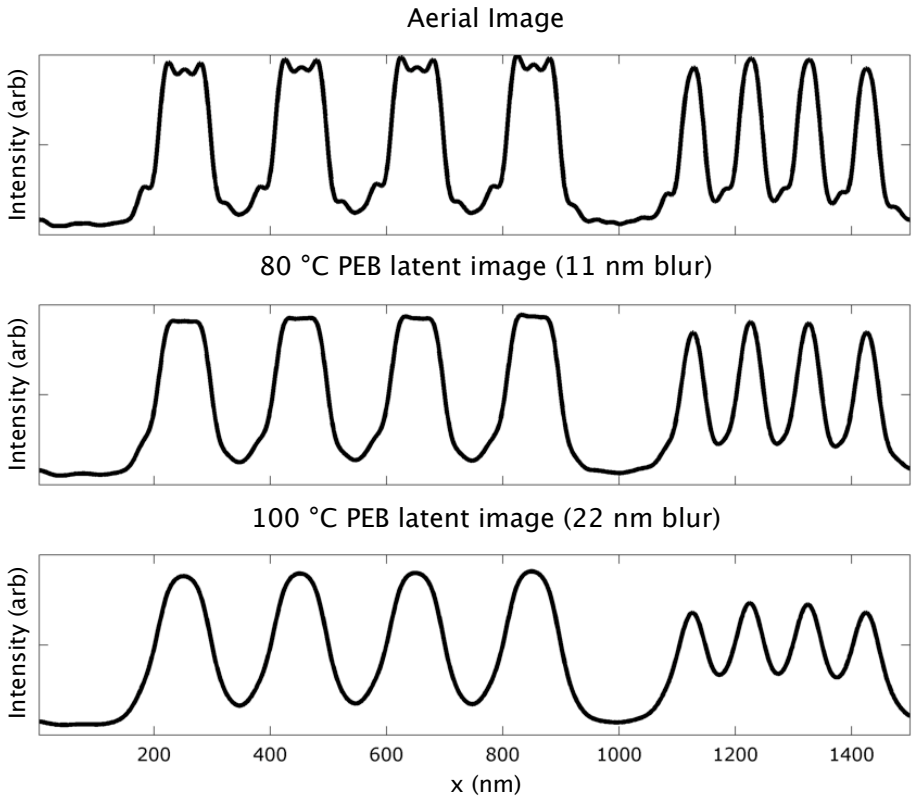

Figure 6.10: Top: simulated aerial image cross-section of 100 nm and 50 nm 1:1 lines at the SEMATECH Berkeley MET assuming 0.35 - 0.55 annular illumination. Middle and bottom: simulated latent images in TOK EUVR P1123 photoresist with an 80 °C PEB and 100 °C PEB, respectively.

Figure 6.11 (middle) shows one instance of the absorbed photon spatial distribution for the 80 °C PEB and 100 °C PEB wafers using this shot noise model. Figure 6.11 (bottom) shows the resulting resist images assuming a best-case binary development model. In practice,

that the top-down metrology process averages LER effects in the vertical dimension.

statistical fluctuations in deprotection and dissolution mechanisms most likely add additional noise.

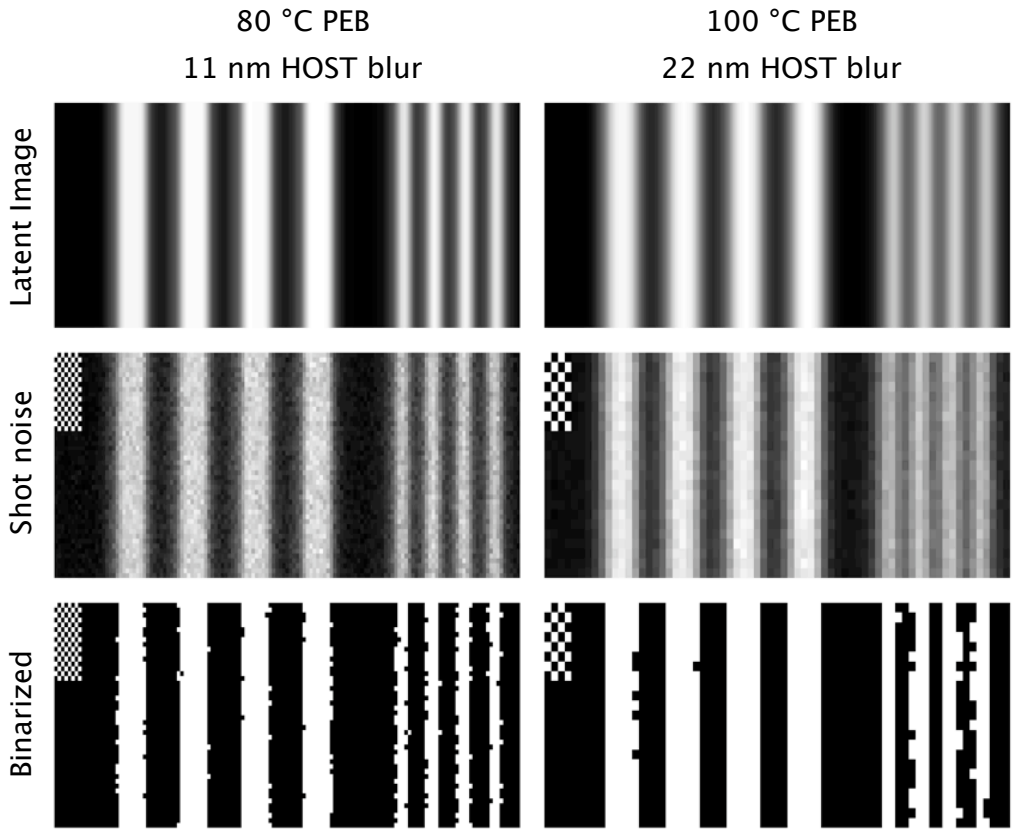

Figure 6.11: Top: simulated latent images of 100 nm and 50 nm 1:1 lines in TOK EUVR 1123 photoresist with an 80 °C PEB (left) and 100 °C PEB (right) at the SEMATECH Berkeley MET with 0.35 - 0.55 annular illumination. Middle: simulated latent images with shot noise added. Bottom: resulting resist profiles assuming a threshold development model. The pixel size is shown in the top left of the middle-row and bottom-row images.

This model illustrates that the 100 °C PEB wafer's $2.7\times$ higher number of absorbed photons per bin (compared to the 80°C PEB wafer) cannot compensate for the detrimental affect of the 22 nm blur on 50 nm 1:1 features. In this regime, the LER of 50 nm 1:1 lines is dominated by the deprotection contrast, not the absorbed photon signal-to-noise ratio. As the printed half-pitch grows relative to the deprotection blur, the blur's degradation of the aerial image becomes less severe. For example, in 100 nm 1:1 features the degrading effects of the 100 °C wafer's twice-as-large deprotection blur (compared to the 80 °C PEB wafer) are

more-or-less nulled-out by its 2.5× higher number of absorbed photons per pixel, as indicated by the equivalence of the measured LER on 100 nm 1:1 features in the 80 °C PEB and 100 °C PEB wafers.

These results are consistent with the Steenwinckel et al. [32] claim that as the acid diffusion length becomes large relative to the size of the feature being patterned, the LER reduction from improved counting statistics becomes dominated by an increase in LER due to reduced deprotection contrast. While the previous discussion was tied specifically to the PEB wafers, other examples from these experiments also defend this interpretation.

1. 50 nm 1:1 LER is larger than 100 nm 1:1 LER in every resist we tested where the deprotection blur is larger than 30 % of the 50 nm half pitch (five out of six resists)[vi].

2. At E-size ≈ 7 mJ/cm^2 (the only value common to every tested platform) the fractional difference between patterned and intrinsic LER is largest in the high blur platforms (5435 and 5271) and smallest in the low blur platform (1123).

6.6.4 Base loading: a simultaneous reduction of LER and sensitivity

Every tested base loading series exhibits a simultaneous reduction in LER and sensitivity as base weight percent is increased. These trends have been reported before [33, 32] and are in alignment with predictions based on shot noise arguments. Further discussion of this topic will be deferred to a later section.

6.6.5 PAG loading: a simultaneous reduction of LER and E-size

Increasing PAG weight percent in EH27 causes a simultaneous reduction of LER and E-size while leaving blur unaltered. This result opposes predictions based on photon shot noise arguments; it seems "too good to be true." As such, it warrants discussion. Throughout the PAG series, base weight percent, PEB temperature, and development parameters were fixed, making it likely that each PAG weight percent requires the same initial quantity of photo-generated acid to deprotect the same amount of resist. If this is true, each arriving photon in

[vi] The aerial image log slope of 50 nm and 100 nm line-space features is almost identical owing to their large size relative to the diffraction limit of the optic [27]

the higher PAG formulations must be comparatively more likely to activate a PAG since there are less incident photons for the same number of photo-generated acids.

One possibility is that increased PAG weight percent improves resist absorption and leaves the quantum yield[vii] unaffected. In this case, each absorbed photon would activate, on average, the same number of PAG molecules as in the reference formulation, however less arriving photons would be required to establish the number of absorbed photons needed for adequate deprotection. On the other hand, PAG weight percent may not affect resist absorption, meaning that a reduction in dose correlates to a reduction in the number of absorbed photons. In this case, each absorbed photon would have to generate more acids than in the reference formulation, i.e., the quantum yield would have to increase. Recent work suggests that this is in fact what happens, at least in the EH27 platform that is examined here [35]. In general, some mixture of these two situations is also possible.

In both of the described scenarios the higher PAG formulation absorbs at most the same number of photons as the reference formulation so reductions in photon shot noise cannot be responsible for improved LER. Of course this begs the question: what *is* improving LER with increased PAG? It is possible that the dissolution and / or mesoscale properties of the photoresist improve with PAG weight percent and advantageously affect LER [36, 37]. It is also possible that increasing the PAG weight percent improves the spatial distribution of PAG molecules, subsequently improving LER. It is also possible that it is something else all together.

While exploration of the first hypothesis is beyond the scope of this work, these experiments lend themselves nicely to a fruitful discussion of the PAG uniformity hypothesis. If the PAG spatial distribution obeys Poisson statistics, the plausibility of the hypothesis boils down to a numbers ratio: is the number of activated PAGs per bin much less than ($\ll$) or close to ($\approx$) the number of total PAGs per bin? If the PAG distribution does *not* follow a Poisson distribution, i.e., there are local aggregation or clumping sites, the treatment is more complicated. For the Poisson-distributed case, a relatively simple counting model can assess whether or not the PAG hypothesis is a reasonable one.

[vii] The average number of acids generated by an absorbed photon

6.7 A model for EUV exposure statistics

6.7.1 EUV photon absorption vs. acid generation

In deep ultraviolet (DUV) (λ = 248 nm and 193 nm) exposures, photons are absorbed by PAG molecules [35], creating acids at each absorption site. At EUV wavelengths, acid is generated with an entirely different mechanism. As described by Brainard et. al. [35], EUV photons are not absorbed by PAGs, they are absorbed by any atom. In EUV exposures, the absorption process ionizes an atom and subsequently creates a cascade of secondary electrons that can interact with PAG molecules to create acids. Consequently, an absorbed EUV photon is not guaranteed to create any acid. Moreover, the locations of the photon absorption sites do not in general coincide with the locations of the photo-generated acids.

6.7.2 EUV quantum yield is a random variable

The quantum yield of EUV photoresist is defined as the average number of acids created by an absorbed EUV photon. Brainard, et. al., has shown that the mean quantum yield of EH27 photoresist scales linearly with PAG density [35]. This result resonates well with physical arguments, namely that the quantum yield should be heavily influenced by number of PAGs that lie within the mean free path[viii] volume of a photoelectron cascade. Therefore a linear scaling of quantum yield with PAG density seems reasonable, at least up to the point that the number of PAGs per mean free path volume is so large that the entire photoelectron cascade is consumed.

Assuming that the PAG distribution obeys Poisson statistics, the number of PAG molecules in any particular counting bin is determined by a Poisson random variable. Since the PAG density in a counting bin is proportional to the number of PAGs in that bin, i.e., density = $\frac{\text{PAG}\#}{\text{volume}}$, the bin-to-bin statistics of the PAG density are identical to the bin-to-bin statistics of the PAG molecules. In other words the SNR of the PAG density is the same as the SNR of the PAG molecules. Since quantum yield is proportional to PAG density, it follows that quantum yield is a random variable and its SNR is the same as the SNR of the PAG molecules.

[viii] The average distance covered by a particle between subsequent impacts.

6.7.3 PAG statistics change throughout the exposure

The distribution of available PAG molecules changes during an exposure. For example, consider a time instant near the beginning of an exposure in a spatial location that is lit up. If a photon arrives in that instant, is absorbed, and subsequently converts PAGs to acids, then as far as the next arriving photons are concerned, it is as if those PAGs were never there in the first place. Said in other words, during the exposure, photons (electrons) remove available PAGs from the lineup so that the next arriving photons see a sparser set of available PAGs than the photons before them. Since the removal of available PAGs is done in a Poisson-distributed sense[ix], PAG consumption can be modeled as a reduction in the mean number of available PAGs per counting bin throughout the exposure. The reduction of available PAGs naturally affects the quantum yield, causing its statistics to follow those of the available PAG molecules.

If $\overline{P}$ is the mean number of PAGs per bin at the beginning of the exposure and $\overline{A}$ is the mean number of acids generated per bin during the exposure, the initial photons see a mean of $\overline{P}$ PAGs per bin, while the photons at the end of the exposure see a mean of $\overline{P} - \overline{A}$ PAGs per bin. The effective mean number of PAGs per bin can be obtained by finding the mean PAG number whose SNR matches the time-averaged PAG SNR during the exposure, including the effects of PAG consumption and quantum yield reduction. A reasonable approximation to this number is $\overline{P} - \frac{\overline{A}}{2}$, the average of the beginning and ending mean number of PAGs. Since the SNR of the quantum yield is identical to the SNR of the PAG molecules, the SNR of the effective quantum yield is determined by the SNR of a Poisson random variable with rate $\overline{P} - \frac{\overline{A}}{2}$.

6.7.4 Acid is the product of two random variables

The number of acids, A, generated in a counting bin is the product of the number of absorbed photons in that bin, N (a Poisson random variable), and the quantum efficiency of that bin, Q (a Poisson random variable). When multiplying two random variables X and Y, the fractional

[ix] PAGs are removed by the electron cascades, which more-or-less have the Poisson spatial distribution of the photon absorption sites

uncertainty[x] of the product is the quadrature addition of the fractional uncertainties of X and Y [38]. For example:

$$\frac{\sigma_{XY}}{\overline{XY}} = \frac{1}{SNR_{XY}} = \sqrt{\left(\frac{\sigma_X}{\overline{X}}\right)^2 + \left(\frac{\sigma_Y}{\overline{Y}}\right)^2} \tag{6.1}$$

To determine the fractional uncertainty of A, three quantities must be known: 1) $\overline{N}$, the mean number of absorbed photons per bin, 2) $\overline{Q}$, the mean quantum yield during the exposure, and 3) $\overline{P}$, the mean number of PAGs per bin at the beginning of the exposure. The product $\overline{NQ}$ is $\overline{A}$, the mean number of acids generated per bin. Since Q acquires the fractional uncertainty of the PAGs, its average fractional uncertainty throughout the exposure is that of a Poisson random variable with rate $\overline{P} - \frac{\overline{NQ}}{2}$:

$$\frac{\sigma_Q}{\overline{Q}} = \frac{1}{\sqrt{\overline{P} - \frac{\overline{NQ}}{2}}} \tag{6.2}$$

Likewise, N has the fractional uncertainty of a Poisson random variable with rate $\overline{N}$:

$$\frac{\sigma_N}{\overline{N}} = \frac{1}{\sqrt{\overline{N}}} \tag{6.3}$$

The fractional uncertainty of A is given by the quadrature addition of the fractional uncertainties of N and Q:

$$\frac{\sigma_A}{\overline{A}} = \sqrt{\frac{1}{\overline{P} - \frac{\overline{NQ}}{2}} + \frac{1}{\overline{N}}} \tag{6.4}$$

Eq. 6.4 can also be inverted to reveal the SNR of acids generated per bin, a quantity believed to be directly proportional to LER.

[x] The ratio of the standard deviation (σ_x) of the value/quantity and the mean ($\bar{x}$) of the value/quantity, or the inverse of the signal-to-noise ratio.

6.8 More discussion

6.8.1 PAG loading and base loading: the numbers breakdown

Table 6.4 summarizes all of the parameters that are relevant for comparing the photon, PAG, and acid statistics for the EH27 PAG/base series. Like the previous counting discussion, E0 doses are used since at line edges the resist is just barely rendered dissolvable to developer. Also, it is assumed that the top-down metrology process (imaging in a SEM) averages LER effects in the vertical dimension and the full 125 nm thickness of the resist stack is used for the vertical dimension of the counting bin. In addition, EH27's published absorption coefficient, E0 values and quantum yield values are used [33, 35].

The mean number of incident photons per counting bin is computed by multiplying E0 by a 33×33 nm^2 blur area[xi] and dividing by the EUV photon energy. Likewise, the mean number of PAGs per counting bin is obtained by multiplying the PAG density by the $33 \times 33 \times 125$ nm^3 bin volume. To obtain the mean number of absorbed photons per bin, the number of incident photons is multiplied by the 41% absorption coefficient of the 125-nm-thick film. Multiplying this number by the mean quantum yield reveals the mean number of acids generated per bin. Finally, the acid SNR of each formulation is computed with Eq. 6.4.

Table 6.4: EH27 photon and PAG statistics

Resist	Base	PAG		E0	Photons		QY	Acids	PAGs	Acid	LER
	(wt%)	(wt%)	(nm^{-3})	(mJ/cm^2)	Inc.	Abs.	(mean)	(mean)	(mean)	SNR	100 nm
EH-27-A	0.50	5.0	0.050	4.6	3340	1369	1.8	2465	6969	33.2	5.7
EH-27-E	0.50	7.5	0.075	3.6	2613	1071	2.3	2464	11100	31.1	5.1
EH-27-H	0.50	10	0.100	2.8	2032	833	3.0	2495	13940	28.0	4.9
EH-27-C	0.17	7.5	0.075	1.1	799	328	2.3	753	11100	17.8	6.9
EH-27-D	0.34	7.5	0.075	1.8	1307	536	2.3	1232	11100	22.6	5.8
EH-27-E	0.50	7.5	0.075	3.6	2613	1071	2.3	2464	11100	31.1	5.1
EH-27-F	0.75	7.5	0.075	4.4	3194	1310	2.3	3012	11100	34.0	4.9
EH-27-G	1.00	7.5	0.075	6.0	4356	1786	2.3	4108	11100	38.6	4.1

In EH27, changing the PAG concentration simultaneously scales the mean quantum yield

[xi] The corner metric blur is used since it agrees more with direct observation. Also, since the measured blurs of each PAG and base formulation are almost within the measurement uncertainty of the blur metric, a single blur is used for all of them.

and E0 in opposite directions such that their product remains fixed. Therefore, when the number of PAGs is larger than the number of absorbed photons, increasing PAG worsens the SNR of the absorbed photons more than it improves the SNR of the of the quantum yield. As a result it is not possible to create a PAG loading scenario where the PAG distribution gets "so much better" that it overcomes the reduced SNR of absorbed photons. In this respect, it is difficult to argue that EH27's LER improvements with PAG loading are due to a more uniform distribution of available PAG molecules; however, this statement assumes that the Poisson distribution approximation is a good one, which may not be the case. If clumping or aggregation is present and increasing the PAG concentration does improve PAG homogeneity, that would probably explain the simultaneous reduction of LER and E-size. It is also possible that the Poisson approximation is correct and the dissolution or mesoscale properties improved so much in the higher PAG formulations that the LER improves despite the worsened acid SNR; but there is no data to back that up, it is simply speculation.

Regarding base loading, the measured LER follows the acid SNR quite well. The highest base formulation just starts to generate enough acids that PAG consumption could noticeably affect the SNR of the quantum yield. If the PAG density and quantum yield were known for the Rohm & Haas base loading series' (5435, 5496, and 5271) it would be interesting to see if they generate enough acids to considerably consume the PAG distribution. In these three resists, the drop in LER in going from the second highest base weight percent to the highest base weight percent is at least a factor of two less than expected form the square root of the relative change in E-size; PAG consumption could be to blame.

6.9 Wrap up

Reducing the PEB temperature of TOK EUVR P1123 simultaneously reduces its deprotection blur and sensitivity while improving 50 nm 1:1 LER and leaving 100 nm 1:1 LER unaffected. Through a latent image + shot noise model, it was shown that in 50 nm 1:1 features an advantageous $1.7\times$ reduction in shot noise (achieved through an increase in the PEB temperature) is stifled by a dramatic reduction in deprotection contrast owing to the incurred deprotection blur's large size relative to the size of the printed feature. In general, these results support a previous claim by Steenwinckel, et. al., namely that as the acid diffusion length becomes

large relative to the size of the printed feature, LER reduction from improved counting statistics becomes dominated by an increase in LER due to reduced deprotection contrast.

Contrary to general expectations, it was found that PAG and base weight percent do not substantially affect the deprotection blur in the tested EUV photoresists. Not surprisingly, base loading was found to simultaneously reduce LER and sensitivity in every tested resist platform; however a puzzling result is that increasing PAG weight percent in EH27 causes a simultaneous reduction of LER and E-size. The exposure statistics model, which includes the effects of a changing PAG distribution throughout the exposure, showed that when PAG loading scales E0 and quantum yield such that their product remains fixed, it should not be able to improve LER, at least not from an improved acid SNR. At this point, the source of EH27's improved LER with increased PAG remains unknown. Repeating this study with other open-platform resists (ones where the constituent make-up is public domain) would be extremely valuable. It is plausible that other resists will respond differently than EH27 to a change in PAG weight percent, and maybe even follow the predictions of the exposure model. It is also possible that the Poisson distribution assumption is incorrect and there is more to the PAG homogeneity than the average number per detection bin. Nonetheless, a number of universities are now investigating the utility of high PAG platforms, acid amplifiers, and other agents designed to improve the quantity and uniformity of acids throughout the resist stack.

Acknowledgements: Special thanks to Jim Thackeray and Katherine Spear from Rohm and Haas, Koki Tamura, Chris Rosenthal, and Dave White from TOK, and Robert Brainard from the University at Albany for supplying resist and process support.

Chapter 7

MOSAIC: a new wavefront metrology

Contents

7.1 Motivation

The manufacturing of high performance semiconductors requires routine monitoring of aberrations in exposure tool optics to ensure that diffraction-limited imaging remains the status quo and optical proximity correction strategies can be maintained. There are currently several metrologies devoted to this task.

One metrology in widespread use today is an integrated lateral shearing interferometer (LSI) developed by ASML branded with the name ILIAS^{TM} [39]. There are also a handful of print-based tests in use that enable the quantification of specific aberrations, i.e., astigmatism, spherical error, or coma by imaging engineered mask features. Examples known to the authors are through-focus patterning of phase dots by Dirsken et. al [40], patterning of Zernike Fourier transform targets by Robins et. al. [41], and patterning of coma targets by Nomura et. al [42], all of which require the ability to pattern at the diffraction limit of the imaging optic. Finally, there is another class of print-based tests that use through focus patterning of vertical, horizontal, and oblique grating patterns to quantify astigmatism [43, 44] and spherical error [3].

As the industry progresses towards EUV lithography, integrated metrologies become increasingly difficult to implement. Moreover, print-based methods relying on diffraction-limited resist performance are hampered by the realities of resist limitations. Through-focus print-based techniques, however, have the ability to bypass resist limitations. Here we present a new wavefront metrology technique, MOSIAC (metrology of optical system aberrations by incoherent curvature sensing), which is a generalization of previously described through focus methods and enables complete[i] wavefront recovery from print or aerial image based measurements.

While the above discussion has been centralized to the topic of lithography (the original motivation for our work), we would would like to point out that MOSAIC is applicable to the

[i] Up to the two tilt Zernike terms, $Z = \rho\cos\theta$ and $Z = \rho\sin\theta$

broader context of all optical systems requiring routine wavefront characterization where it is impractical to remove the optics from the tool assembly for testing.

7.2 An introduction to MOSAIC

7.2.1 Goals

MOSAIC was designed with several goals in mind: (1) complete wavefront characterization, (2) simple integration into lithographic tools, (3) scalable to any numerical aperture (NA), and (4) at wavelength. MOSAIC, as you will see, satisfies all of the above criteria.

7.2.2 The basics of how it works

When an optical system is perfect, ray bundles emerging from different localized regions of the optic come to focus in the same plane (Figure 7.1 left). When aberrations are present, the same ray bundles focus in different longitudinal planes (Figure 7.1 right). The focus shift of a particular ray bundle is determined by the curvature of the aberration at the location the ray bundle emerges from the part.

The goal of MOSAIC is to image features that probe local regions of the optic and measure how focus varies with the probe point. In essence, MOSAIC aims to map out the local focal length of the entire pupil of an optic. From these data, the curvature (second derivative) of the aberration can be obtained and subsequently used to obtain the aberration itself.

7.2.3 A potential roadblock?

MOSAIC requires imaging in a regime where the object structure is much bigger than the diffraction limit and the source is reasonably coherent. Finding focus in this regime may be difficult owing to the large depth of focus (DOF) rendered by the imaging conditions. To address these concerns, we need to understand how the DOF is affected by object structure, illumination spatial coherence, and optical aberrations. The goal of the following section is to frame these concepts in a unified way, making it easier to intuit some of the details and subtleties of MOSAIC.

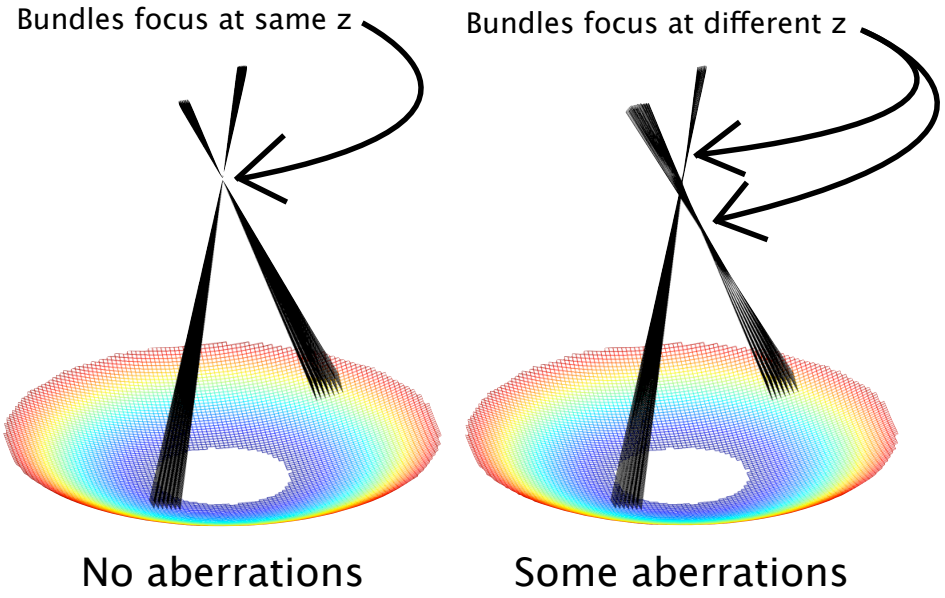

Figure 7.1: Left: ray bundles emerging from a perfect spherical wave come to focus in the same z plane. Right: ray bundles emerging from an aberrated spherical wave come to focus in different z planes.

7.3 The marriage between spatial coherence and DOF

7.3.1 The Köhler illuminator

There are three basic illumination classes: coherent, partially coherent, and incoherent. Using the Köhler illumination framework, it is easy to differentiate these illumination types and the way they influence imaging. Figure 7.2 shows a simplified Köhler illumination scheme illustrating the intimate link between illumination spatial coherence and illumination pupil fill[ii]. In a nutshell, the Köhler lens relays points in the source plane to points in the optic plane while mapping each source point into a unique illumination spatial carrier.

7.3.2 Coherent vs. partially coherent vs. incoherent

Spatially coherent illumination can be viewed as illumination containing one spatial carrier that probes the imaging optic in a single discrete location, i.e., the source is a single point.

[ii] "Pupil fill" is the term commonly used to describe the region of the imaging lens that is probed by the illumination field

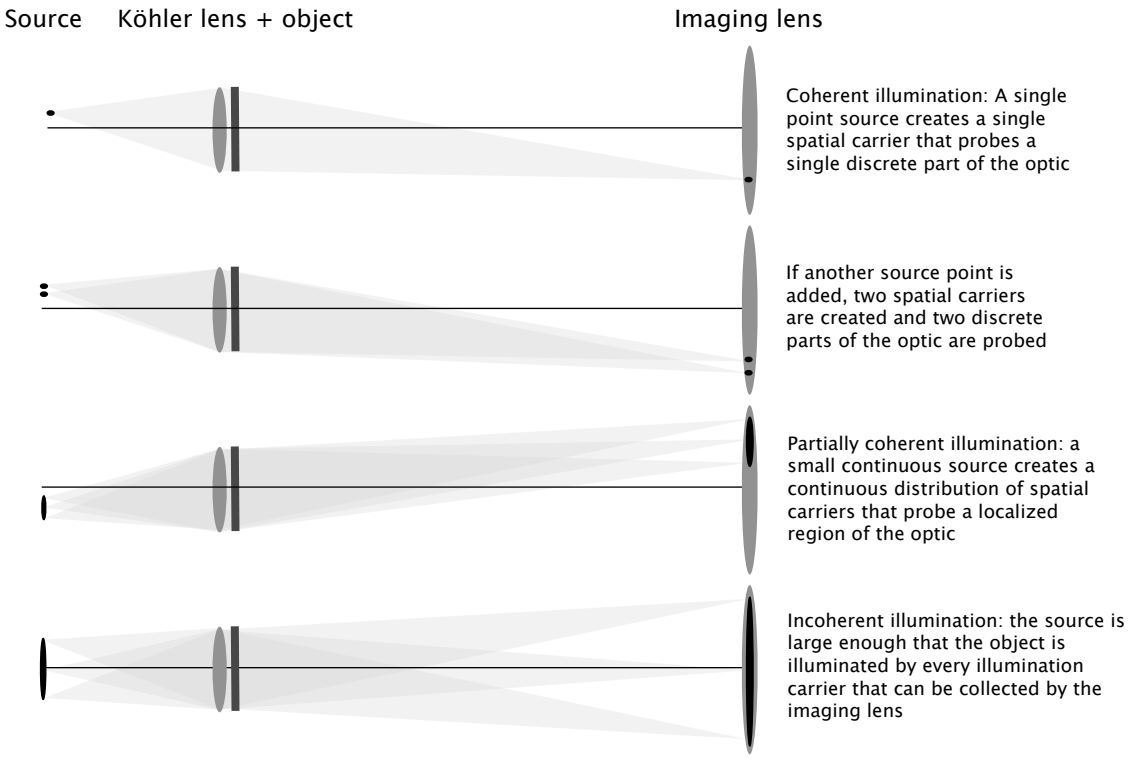

Figure 7.2: A Köhler illumination scheme. The Köhler lens relays points in the source plane to points in the lens plane while mapping each source point into a unique illumination spatial carrier. It is assumed that the point sources which make up the broad source are uncorrelated so that coherent aerial images from each source point are added together to produce partially coherent aerial images.

As the source broadens from a single point[iii], the illumination becomes partially coherent. Each source point creates a different spatial carrier that probes the optic in a different place. For example, the third row of Figure 7.2 shows that a small localized collection of point sources maps to a small localized pupil fill. As the source and pupil fill continue to broaden simultaneously, the system approaches the incoherent limit. As shown in the fourth row of Figure 7.2, incoherence is achieved when the object is illuminated by every spatial carrier that can possibly be collected by the lens including the effects of diffraction of the object.

The aerial image (AI) can be viewed as a summation of the coherent AIs of every spatial carrier in the illumination field, each of which is linked to unique source point and probe point. As a result, the Köhler illuminator provides an insightful framework to describe the

[iii] We assume all source points are uncorrelated as in a traditional incandescent filament

relationships between illumination spatial coherence and DOF. For the following discussion we will consider the object in Figure 7.3 which has structure large enough that most of the energy of its coherent diffraction pattern occupies a small localized region of the imaging lens. Since partially coherent AIs are built up as summation of coherent AIs, we will begin by analyzing the through focus behavior of a coherent AI.

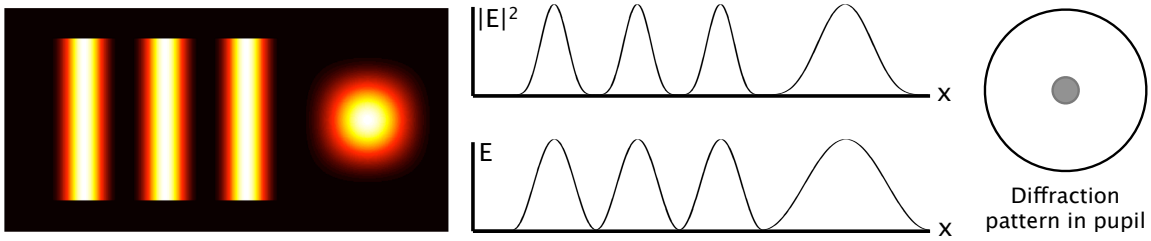

Figure 7.3: Modeled object. The object has large enough structure that its conerent diffraction pattern occupies a localized region of the imaging lens. The square on the right has been added for reference purposes (magnification inversion).

7.3.3 Defocus in non-aberrated coherent systems

Defocus in coherent optical systems adds a quadratic phase term to the coherent transfer function (CTF). If the probed region is small, a two-dimensional Taylor expansion of the quadratic phase about the centroid of the probed area is an insightful exercise since it breaks the problem into simple components while still remaining accurate. The constant term of the Taylor expansion adds a bulk phase to the frequency spectrum of the image, which does nothing in the spatial domain. The linear phase terms in the frequency domain (there will be x and y components in general) map to an image shift in the spatial domain. For example, as the image plane moves through the focal plane, the local slope of the quadratic phase of the CTF swings from positive to negative, giving rise to the image translation through the center of the field in the spatial domain.

The mapping of the linear phase term in the frequency domain to a shift in the spatial domain can be intuited by the convolution theorem, i.e., the inverse Fourier transform (IFT) of the product of the image spectrum and the linear phase from the CTF term is the convolution of the individual IFT's. Since the IFT of the linear phase term is a shifted delta function,

the convolution operation shifts the position of original image by an amount that depends on the probe location and the level of defocus. Likewise, the quadratic phase term of the Taylor expansion of the CTF phase maps to a convolution in the spatial domain with a Gaussian kernel whose width grows as defocus worsens. When the size of the probed region increases, the accuracy of the Taylor expansion model breaks down. Also, the finite extent of the lens and its low-pass filtering operation in the frequency domain become more apparent since high spatial frequencies from the object may be blocked by the CTF (in the spatial domain is equivalent to the feature diffracting light out of the lens).

Figure 7.4 shows through focus coherent AIs for three probe locations assuming zero aberrations. Note that magnification causes an inversion in the image. The first column of Figure 7.4 assumes a probe location on the x-axis close to the edge of the pupil. By direct observation, negative defocus (defined as shifting the image plane towards the imaging optic) shifts the image to the right of the field and and positive defocus shifts the image to the left of the field, a direct result of the linear component of the CTF quadratic phase. In fact, there is a general relationship between the probe location and the direction a coherent AI shifts through focus: as the image plane moves further from the imaging optic, a coherent AI always shifts in the direction of a vector pointing form the probe location to the center of the optic. The shift vectors at the bottom of Figure 7.4 illustrate this point. Also, Figure 7.5 shows that the quadratic component of the CTF causes increased degradation as the object size shrinks, i.e., the convolution operation with the gaussian kernel more significantly affects contrast, moreover the low-pass filtering operation removes some portions of the object frequency spectrum.

7.3.4 Defocus in non-aberrated partially coherent systems

Non-aberrated optical systems have a very unique property: there exists a plane (the focus plane) where the coherent AIs from every single possible probe location are spatially aligned. Figure 7.4 illustrates this point as all three probe locations have spatially aligned coherent AIs in the focus plane.

In a focused, non-aberrated system, one can continue adding additional illumination spatial carriers all the way up to the incoherent limit and the AI does not blur since the constituent

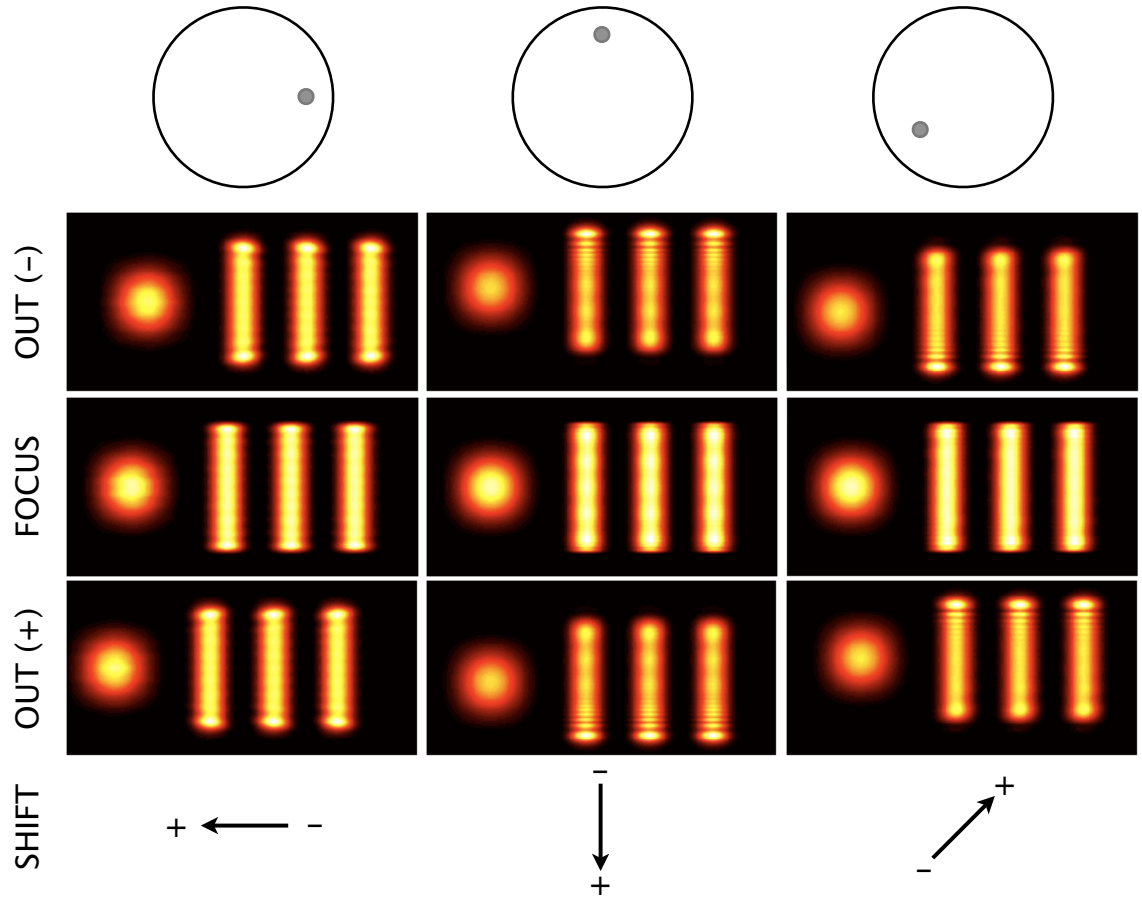

Figure 7.4: First row: pupil fill from a Köhler illuminator. Second row: aerial image in a defocused plane closer to the optic. Third row: aerial image in the plane of best focus. Fourth row: aerial image in a defocused plane further from the optic.

images do not shift. In defocused systems this is not the case. When a system is defocused, each illumination carrier has a unique coherent AI shift. As a result, the partially coherent AI becomes a summation of many uniquely shifted copies of roughly the same feature. Consequently, the AI becomes increasingly blurred as the illumination approaches the incoherent limit.

The DOF of an optical system can be defined as the distance over which the incurred image blur is smaller than the diffraction-limited resolution. When the pupil fill is on-axis monopole (a small disk centered on axis) there is a natural symmetry in the coherent AI shifts of the spatial carriers in the illumination: for each spatial carrier, there is one whose

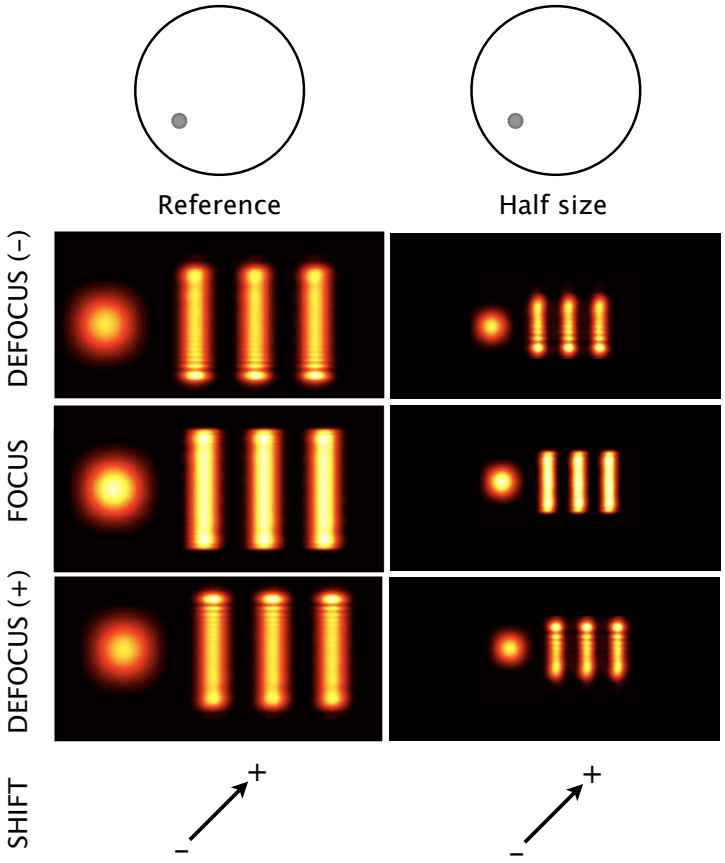

Figure 7.5: Through focus coherent aerial images of the object in Figure 7.3 (left) and a half-sized version of it (right). Defocus causes the same global image shift in both cases. In comparison to the focused AIs, defocus more severely degrades the AI of the half size feature because its coherent diffraction pattern occupies a larger region of the imaging lens and it acquires more spectral dephasing than the reference feature.

coherent AI translation in defocus is equal and opposite. Off-axis monopole pupil fill lacks the symmetry of the on-axis case. The result is that in defocused systems, off-axis monopole illumination blurs the AI *more* than the same monopole illumination centered on axis. In other words, the DOF worsens as the centroid of the probe location moves away from the optic axis.

7.3.5 Defocus in aberrated partially coherent systems: MOSAIC

In an aberrated optical system there is no image plane where the coherent AIs from every possible probe location are spatially aligned. Instead, small localized regions of the optic each

have their own plane of best focus, or a plane where the coherent AIs from all illumination carriers that probe that small region are best aligned. Experimentally, the goal of MOSAIC is to determine the plane of best focus (in two orthogonal imaging directions) at a sampling of locations throughout the pupil. Once the pupil is mapped out in terms of focus plane vs. probe location, the second derivatives of the wavefront (in two orthogonal directions) can be obtained, and subsequently be used to obtain the wavefront itself.

7.4 Mathematical foundation

Consider the electric field of a plane wave $E(\mathbf{r}) = \exp\left(i\frac{2\pi}{\lambda}\phi(\mathbf{r})\right)$ with:

$$\begin{aligned}\phi(\mathbf{r}) &= \mathbf{v}\cdot\mathbf{r}\\ \mathbf{v} &= \alpha\hat{\mathbf{x}} + \beta\hat{\mathbf{y}} + \gamma\hat{\mathbf{z}}\\ \mathbf{r} &= x\hat{\mathbf{x}} + y\hat{\mathbf{y}} + z\hat{\mathbf{z}}\end{aligned}$$

such that $\mathbf{v}$ is a unit vector ($\alpha^2+\beta^2+\gamma^2=1$) indicating the propagation direction. When a geometrical ray propagates a distance d in the direction of $\mathbf{v}$, its x, y, and z coordinates change by dα, dβ, and dγ, respectively. For example, changing the z coordinate by z' requires the ray to travel a distance $d' = z'/\gamma$, thus changing the x coordinate by $d'\alpha$ or $z'\alpha/\gamma$. Tracking the x and y coordinates of a ray as its z coordinate changes is trivial:

$$\begin{aligned}x(z) &= x_i - \frac{z\alpha}{\gamma} = x_i - \frac{z\alpha}{\sqrt{1-\alpha^2-\beta^2}}\\ y(z) &= y_i - \frac{z\beta}{\gamma} = y_i - \frac{z\beta}{\sqrt{1-\alpha^2-\beta^2}}\end{aligned}$$

where (x_i, y_i) is the (x, y) point of the ray in the initial z plane. For plane waves, it is obvious that the ray direction cosines α and β are the x and y derivatives of $\phi(\mathbf{r})$, respectively. In the general case, the x and y derivatives of $\phi(\mathbf{r})$ still give the ray direction cosines, however they are *local* direction cosines, and vary with (x_i, y_i). Below we show a general version of the x tracking equation using the derivative notation $\phi_x = \frac{\partial\phi}{\partial x}$.

$$x(z) = x_i - z \left. \frac{\phi_x}{\sqrt{1-\phi_x^2-\phi_y^2}} \right|_{(x_i,y_i)} \tag{7.1}$$

Now consider a second geometrical ray, this time leaving the initial z plane from the incrementally shifted x position: ($x_i + \delta$, y_i). The x position of this second ray is given by:

$$x(z) = x_i + \delta - z \left. \frac{\phi_x}{\sqrt{1-\phi_x^2-\phi_y^2}} \right|_{(x_i+\delta,y_i)} \tag{7.2}$$

When the wavefront is spherical, or close to it, these geometrical rays eventually cross (focus) in x. By equating Eq. 7.1 and Eq. 7.2 we can determine the z plane where this occurs.

$$\frac{1}{z} = \frac{\left. \frac{\phi_x}{\sqrt{1-\phi_x^2-\phi_y^2}} \right|_{(x_i+\delta,y_i)} - \left. \frac{\phi_x}{\sqrt{1-\phi_x^2-\phi_y^2}} \right|_{(x_i,y_i)}}{\delta} \tag{7.3}$$

Taking the limit as δ goes to zero we obtain a derivative:

$$\frac{1}{z} = \frac{\partial}{\partial x} \left(\frac{\phi_x}{\sqrt{1-\phi_x^2-\phi_y^2}} \right) \tag{7.4}$$

After the derivative is carried out, we obtain an equation for the second derivative (curvature) of ϕ, which, in essence, is the heart of MOSAIC:

$$\phi_{xx} = \frac{\frac{(1-\phi_x^2-\phi_y^2)^{3/2}}{z} - \phi_x\phi_y\phi_{xy}}{1-\phi_y^2} \tag{7.5}$$

To recap, we've derived Eq. 7.5 by: (1) tracing two rays that left an almost spherical wavefront from neighboring locations (x_i, y_i) and ($x_i + \delta$, y_i), (2) finding the z plane where the two rays cross in x, and (3) taking the limit of infinitesimally small δ. The outcome is a link between the x-crossing plane for a (x_i, y_i) point of origin and the second x derivative of the wavefront at that point.[iv]

7.4.1 Putting the math to work

Eq. 7.5 implies that even if you know z (the x-crossing plane of the two rays) a-priori knowledge of ϕ_x, ϕ_y, and ϕ_{xy} are required to extract ϕ_{xx}; however, ϕ_x, ϕ_y, and ϕ_{xy} of the base

[iv] It is trivial to write the analogue of Eq. 7.5 for focus in the y direction, or any general direction

sphere are sufficient. Curvature data are obtained through Eq. 7.5 by imaging features that probe local regions (x_i, y_i) of the optic and measuring how focus varies with (x_i, y_i). Once ϕ_{xx} and ϕ_{yy} are sampled through the pupil, the second x and y derivatives of the base sphere can be removed, revealing the second x and y derivatives of the aberration, which will be defined f_{xx} and f_{yy}, respectively.

7.5 Recovering the aberration from its curvature

To recover the underlying aberration, f, from f_{xx} and f_{yy} we use a least squares approach. Let $W(x_i, y_i)$ be a functional representation of the aberration at the sampled point (x_i, y_i) and assume that W can be represented by a linear combination of Zernike polynomials up to index L.

$$W(x_i, y_i) = \sum_{j=0}^{L} a_j Z_j(x_i, y_i) \tag{7.6}$$

Before writing the second derivatives of W we note that the piston, tilt x, tilt y, and 45° astigmatism Zernikes vanish when differentiated twice. In the Fringe indexing scheme [45], these correspond to Z_0, Z_1, Z_2, and Z_5, respectively. If we define α as the set of non-vanishing Zernike indices (3, 4, 6, 7, 8 ... L) we can write the second derivatives of W as follows:

$$\frac{d^2W(x_i, y_i)}{dx^2} = \sum_{j \in \alpha} a_j \frac{d^2Z_j(x_i, y_i)}{dx^2}$$
$$\frac{d^2W(x_i, y_i)}{dy^2} = \sum_{j \in \alpha} a_j \frac{d^2Z_j(x_i, y_i)}{dy^2}$$

In addition, we define the curvature variance, σ, as

$$\sigma = \sum_{i=1}^{N} \left[\frac{d^2W(x_i, y_i)}{dx^2} - f_{xx}(x_i, y_i)\right]^2 + \left[\frac{d^2W(x_i, y_i)}{dy^2} - f_{yy}(x_i, y_i)\right]^2$$
$$= \sum_{i=1}^{N} \left[\sum_{j \in \alpha} a_j \frac{d^2Z_j(x_i, y_i)}{dx^2} - f_{xx}(x_i, y_i)\right]^2 + \left[\sum_{j \in \alpha} a_j \frac{d^2Z_j(x_i, y_i)}{dx^2} - f_{yy}(x_i, y_i)\right]^2$$

where N is the number of probe points. If we wish to find the expansion coefficients a_j that make W most closely resemble[v] f, a solution is found when the partial derivative of σ with respect to each of the coefficients a_p is made equal to zero, i.e., $\partial\sigma/\partial a_p = 0$. After simplification, we obtain the following linear system:

$$\sum_{j \in \alpha} a_j G_{p,j} = F_p \quad \forall\, p \in \alpha \tag{7.7}$$

where

$$G_{p,j} = \sum_{i=1}^{N} \left[\frac{d^2 Z_p(x_i, y_i)}{dx^2} \frac{d^2 Z_j(x_i, y_i)}{dx^2} + \frac{d^2 Z_p(x_i, y_i)}{dy^2} \frac{d^2 Z_j(x_i, y_i)}{dy^2} \right]$$
$$F_p = \sum_{i=1}^{N} \left[f_{xx}(x_i, y_i) \frac{d^2 Z_p(x_i, y_i)}{dx^2} + f_{yy}(x_i, y_i) \frac{d^2 Z_p(x_i, y_i)}{dy^2} \right]$$

To better intuit what G and F are, it helps to introduce the idea of a "matrix vector:" a vector with matrices (rather than scalars) as elements. F_p can be viewed as the inner (dot) product between the curvature "matrix vectors" of f and Z_p, i.e.

$$F_p = \begin{bmatrix} f_{xx} & f_{yy} \end{bmatrix} \times \begin{bmatrix} Z_{p_{xx}} \\ Z_{p_{yy}} \end{bmatrix} \tag{7.8}$$

In other words, F_p can be intuited as the projection of the curvature of f onto the curvature of Z_p. Likewise, $G_{p,j}$ is the projection of the curvature of Z_p onto the curvature of Z_j.

The linear system in Eq. (7.7) can be solved using a Gram-Schmidt (GS) orthogonalization similar to the one outlined in Malacara's *Optical Shop Testing* [46]. The GS technique transforms the set of Zernike curvature matrix vectors ($j \in \alpha$) to a new set that are orthogonal over the sampled probe points. In this new basis, orthogonal in the "matrix vector inner product," all of the off-diagonal elements of the G matrix vanish, making the solution to the linear system trivial. Once the coefficients a_p are found in the orthonormal basis, the GS inverse transformation can be used to map the expansion coefficients back the original basis, revealing the aberration f up to piston, tilt x, tilt y and 45° astigmatism.

[v] up to pixton, tilt x, tilt y and 45° astigmatism

7.5.1 Acquiring the 45° astigmatism term

Astigmatism is one of a few low-order aberrations that can be minimized by optimizing system alignment. Moreover, astigmatism affects the through focus behavior of the aerial image. The fact that MOSAIC cannot obtain the 45° astigmatism term is problematic. Fortunately, this term can be acquired by repeating the measurement in a 45°-rotated system. Practically, this can be achieved several ways: rotating the mask 45°, rotating the part 45°, or encoding diagonal features on the mask that diffract at ±45° angles. The importance is that Z_4 in the rotated system (which is recoverable) maps to Z_5 in the non-rotated system (which is not recoverable). By completing the measurement in both systems, we capture both astigmatism Zernikes and two independent measurements of all other Zernike coefficients.

7.6 Additional details

By design, MOSAIC requires finding focus in a regime where only a small localized region of the optic is probed as the image is formed. From the above discussion, it is evident that the DOF of any given aerial image will be relatively large. Indeed, finding focus will be challenging, but not impossible. For example, employing sinusoidal phase shifting masks that eliminate the zeroth diffraction order make it possible to achieve frequency doubling in the aerial image, halving the DOF without increasing the probe area. Also, lithographic implementations can employ resists with a large intrinsic blur and low chemical contrast to increase patterning sensitivity to changes in aerial image contrast.

In addition, MOSAIC requires through-focus imagery form a multitude of probe locations throughout the pupil. This requires an illumination system that can probe only a local region of the optic, moreover, one that accommodates probing any desired point (x_i, y_i) within the pupil. One way this can be achieved is through a Köhler illuminator using a relatively small (σ = 0.1) source. Translating the source position between exposure sets allows through focus imagery to be gathered at each required probe location.

Alternatively, the use of holographic masks enables a single exposure to probe all of the required locations at once. For example, Figure 7.6 shows a holographic mask that contains a test feature encoded on two different spatial carriers A and B. If off-axis monopole illu-

mination is used to steers the zero order (undiffracted) light outside the pupil, the +1 order diffracted light from carriers A and B (containing the test feature) probe the optic at different locations. When the optic is aberrated, each probe region may have a different focal plane, which is illustrated by showing feature A in focus and feature B out of focus. If the mask is expanded to include spatial carriers that probe every desired location in the pupil, a single set of through-focus exposures will include the through-focus data for every probe location, eliminating the burden of translating the illumination between the through-focus set for each required probe point.

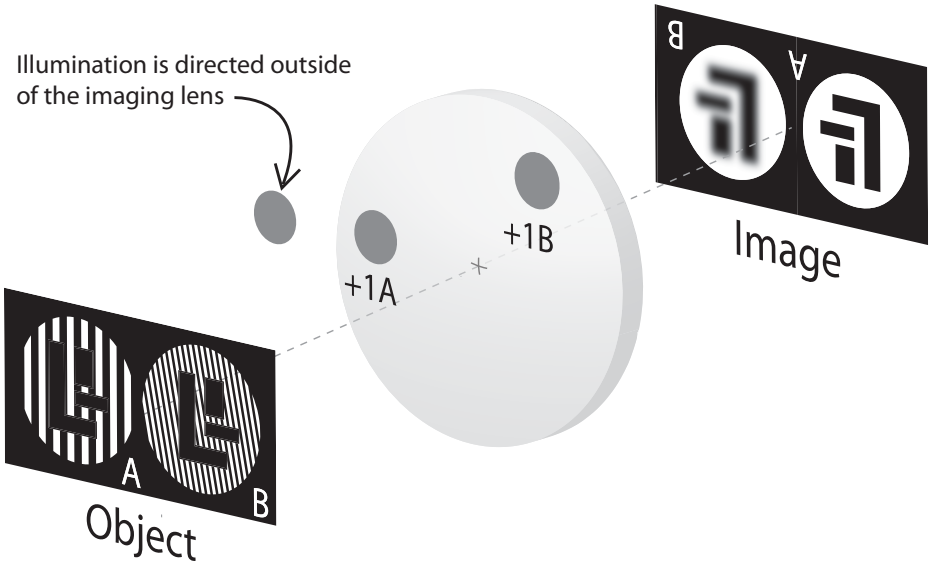

Figure 7.6: The holographic mask contains a test feature encoded on two different spatial carriers A and B. Off-axis coherent illumination steers the zero order (undiffracted) light outside the pupil. The +1 order diffracted light from carriers A and B (containing the test feature) probe the optic at different locations to form images of the test features in the image plane. When the optic is aberrated, each probe region may have a different focal plane. We illustrate this by showing feature A in focus and feature B out of focus.

7.7 Model-based proof of principle

We have modeled a MOSAIC experiment at the SEMATECH Berkeley EUV microfield exposure tool (MET) using commercial aerial image software. Using an object with transmittance

$t(x) = \sin(2\pi f x)$ and $f^{-1} = 400$ nm, we've suppressed the zeroth diffraction order to produce 100 nm features in the aerial image while probing only a $\sigma = 0.22$-wide (full width) area of the optic (see Figure 7.7). As shown in Figure 7.8 (left), we probe a 12 x 24, σ = 0.45 - 0.85 polar grid surrounding the MET's central obscuration[vi]. The maximum sampled radial position is set so that the furthest probing diffraction order of the object is just collected by the lens. We model monopole illumination with a coherence factor of σ = 0.08. The plane of best focus is defined at the plane with the highest aerial image contrast, which peaks at ≈ 1 (at focus) and drops to 0.8 at the extremes of the 3 μm focus range we test. The experiment is repeated twice, once in a reference coordinate system and once in 45-degree rotated system[vii]. Hence we capture both astigmatism Zernikes and two independent measurements of all other Zernike coefficients.

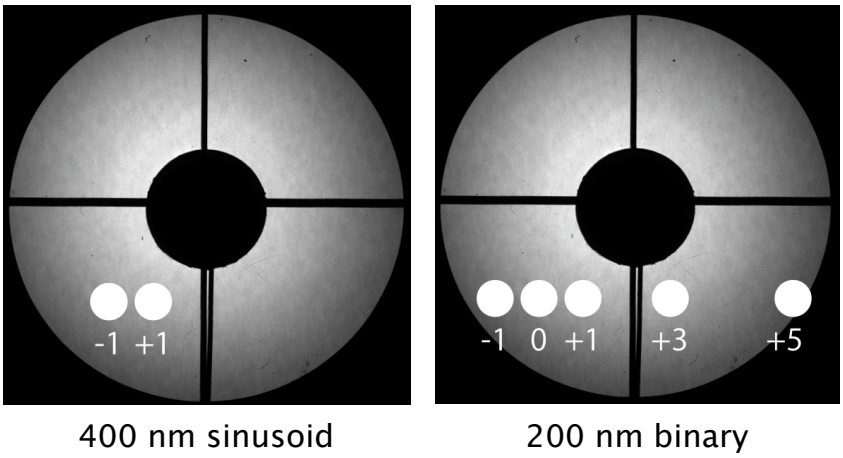

Figure 7.7: Left: diffraction pattern in MET pupil from a 400 nm sinusoidal grating (DC suppressed) with σ = 0.08 off-axis monopole illumination. Right: for reference, we show the diffraction pattern from a conventional 200 nm period binary grating with the same illumination.

Figure 7.8 (right) shows the "programmed" x-focus vs. probe map (solid lines) computed using an inverted version of Eq. 7.5 with known derivatives of the wavefront aberrations and the "recovered" x-focus vs. probe map (circles) computed from Eq. 7.5 and modeled aerial

[vi] This is overkill. A real experiment would only sample enough points to recover the spatial frequency content of the aberration, i.e., a 3 x 12 grid

[vii] We measure horizontal and vertical focus in the reference system and + 45 degree and -45 degree focus in the rotated system

images. Figure 7.9 shows the programmed and recovered wavefront aberrations. The results are in good agreement over the full pupil (σ = 0.3 - 1.0) despite the limited range (σ = 0.45 - 0.85) of the MOSAIC recovery. The RMS error in the recovered wavefront is 28 pm over a σ = 0.35 - 0.95 domain, corresponding to a 4.2% RMS error.

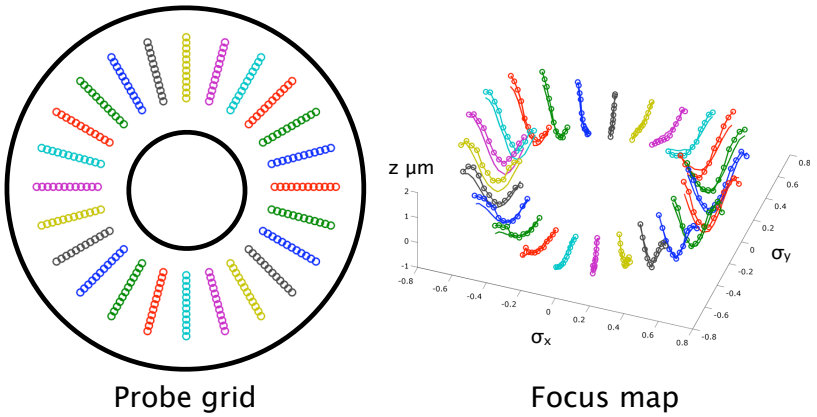

Figure 7.8: Left: 12 x 24, σ = 0.45 - 85 polar probe grid surrounding the σ = 0.3 central stop. Right: "programmed" x-focus vs. probe map (solid lines) vs. "measured" x-focus vs. probe map (circles).

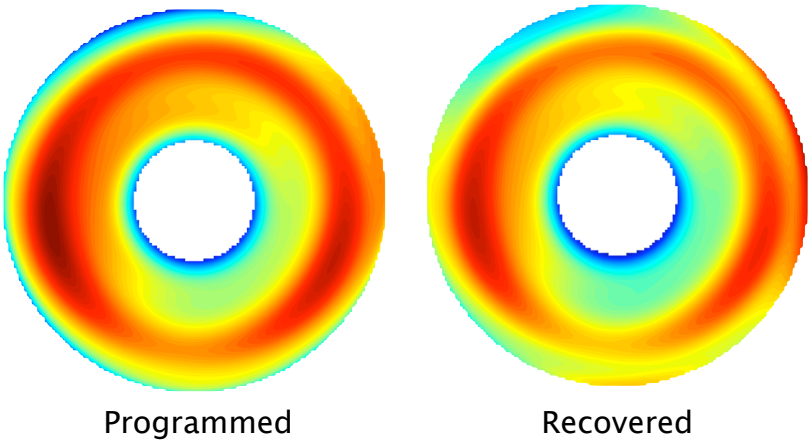

Figure 7.9: Left: programmed aberrations of the SEMATECH Berkeley MET pupil. Right: recovered aberrations of the full pupil ($\sigma = 0.3 - 1.0$) using MOSAIC with a σ = 0.45 - 0.85 probe domain.

7.8 Wrap-up

At wavelength metrology of 0.5 NA EUV exposure tools will not be easy; alignment tolerances for LSI are very demanding [47]; the pinholes and alignment requirements for PS/PDI are extremely demanding; and print-based techniques like MOSAIC have several limitations and technical challenges. For MOSAIC, the biggest implementation hurdle is capturing scanning electron microscope (SEM) images at each probe point: 36 probe points (3×12), 4 geometries, and 20 focus steps leads to 2880 SEM images. At one minute per image, that is two complete days of SEM time. If one SEM image can capture all four geometries, the time can be reduced to 12 SEM hours. While getting the data is laborious, manipulating it is fairly easy. Commercial SEM analysis packages like SuMMIT [26] provide automated routines to obtain LER, critical dimension, and other information from SEM images with very little effort. MOSAIC also has strict requirements on illumination coherence and as a result may not be practical unless production tools are designed with special illumination considerations in mind.

The model-based proof of principle described in this chapter is, however, very encouraging. The 4.2% RMS error in the MOSAIC recovery of the Berkeley MET wavefront is $4\times$ smaller than the error-bars of the original LSI measurement. If used for print-based applications, MOSAIC will most likely serve as a monitoring and characterization tool rather than an alignment tool since the analysis time is too long to provide adequate feedback for fine alignment.

Chapter 8

Conclusion

Contents

8.1 Summary

The work described in this dissertation has improved three essential components of extreme ultraviolet (EUV) lithography: exposure tools, photoresist, and metrology. In Chapter 3 a scanning field-averaging illumination stage was presented that enables non-uniform, high-coherence sources to be used in applications where highly uniform sources are conventionally required. An implementation at the SEMATECH Berkeley EUV exposure tool was shown to achieve a 6.5% peak-to-valley intensity variation across the tool's entire field of view.

In Chapter 4 a design for a stand-alone EUV printing tool capable of delivering 15 nm half-pitch sinusoidal fringes with available sources, gratings and nano-positioning stages was presented and analyzed. It was also shown that the proposed design delivers a near zero line-edge-rougness (LER) aerial image, something extremely attractive for the application of resist testing (the tool's intended use.)

In chapter 5 two robust methods of quantifying the deprotection blur of EUV photoresists were described and experimentally demonstrated. In Chapter 6, these metrics were used

to monitor the deprotection blur, line-edge-roughness (LER), and sensitivity parameters of several EUV photoresists as base weight percent, photoacid generator (PAG) weight percent, and post-exposure bake (PEB) temperature are varied. Two surprising results were found: 1) changing PAG and base weight percent does not significantly affect the deprotection blur of EUV photoresist, and 2) increasing PAG weight percent can simultaneously reduce LER and E-size in EUV photoresist. The puzzling nature of the latter result motivated the development of an EUV exposure statistics model that includes the effects photon shot noise, the PAG distribution, and the changing of the PAG distribution during the exposure. However, for Poisson-distributed PAG molecules, the model predicted that increasing PAG weight percent should not improve LER, at least not from an improved photoacid signal-to-noise ratio. The source of the LER improvement with increased PAG remains unknown.

It was also found that reducing the PEB temperature of EUV photoresist simultaneously reduces its deprotection blur and sensitivity. Through a shot noise + deprotection blur model, it was shown that in 50 nm 1:1 features an advantageous $1.7\times$ reduction in shot noise (achieved through an increase in the PEB temperature) is stifled by a dramatic reduction in deprotection contrast owing to the incurred deprotection blur's large size relative to the size of the printed feature. These results support a general conclusion that as deprotection blur becomes large relative to the size of the printed feature, LER reduction from improved counting statistics becomes dominated by an increase in LER due to reduced deprotection contrast.

Finally, Chapter 7 described MOSAIC, a new wavefront metrology that enables complete wavefront recovery from aerial image or print based measurements. While MOSAIC has several practical challenges, most notably the need to capture over 1000 SEM images for each experiment and the requirement of off-axis monopole illumination, the model-based proof of principle is very encouraging. Using MOSAIC, the Berkeley MET wavefront was recovered with a 4.2% RMS error: an error $4\times$ smaller than the reported errors of the original lateral shearing interferometry (LSI) measurement [27].

8.2 Future work

Continued development of MOSAIC may reveal that it can be used to recover only the low order aberrations - the ones that are most affected by, and can be corrected by, alignment. If

possible, this incarnation would require a much sparser sampling of the pupil, eliminating the burden of gathering upwards of 1000 SEM images for wavefront recovery; however, to avoid the aliasing of higher order aberrations, they must be known a priori. This would require a LSI measurement or a full MOSAIC recovery to have been done before hand.

In addition, further investigations into resist fundamentals are required to close the loop on the PAG - shot noise - LER issue. In conversations since SPIE Advanced Lithography 2009, Chris Mack (www.lithoguru.com, original creator of PROLITH [29]), John Biafore (KLA-Tencor), Mark Smith (KLA-Tencor), Patrick Naulleau (LBNL), and myself have agreed that it would be extremely valuable to continue research on the topic of PAG distributions in EUV photoresists. We have agreed that the next set of experiments should aim to measure: 1) the PAG distribution before an EUV exposure, 2) the acid distribution following the exposure, and 3) the LER of the developed wafer while PAG weight percent and PAG size are varied. Experiments such as these should once and for all provide conclusive evidence about the PAG - shot noise - LER relationships, eliminating the abundance speculation currently clouding the industry.

Bibliography

[1] P. Naulleau, "Status of EUV micro-exposure capabilities at the ALS using the 0.3-NA MET optic," *Proc. of SPIE*, vol. 5374, pp. 881–891, 2004.

[2] K. Goldberg, P. Naulleau, P. Denham, S. Rekawa, K. Jackson, J. Liddle, B. Harteneck, E. Gullikson, and E. Anderson, "Preparations for extreme ultraviolet interferometry of the 0.3 numerical aperture micro exposure tool optic," *JVST B*, vol. 21, no. 6, pp. 2706–2710, 2003.

[3] P. P. Naulleau, J. P. Cain, and K. A. Goldberg, "Lithographic characterization of the spherical error in an extreme-ultraviolet optic by use of a programmable pupil-fill illuminator," *Applied Optics*, vol. 45, no. 9, pp. 1957–1963, 2006.

[4] C. Chang, P. Naulleau, E. Anderson, and D. Attwood, "Spatial coherence characterization of undulator radiation," *Opt. Commun.*, vol. 182, no. 5, pp. 24–34, 2000.

[5] P. Naulleau, K. A. Goldberg, P. Batson, J. Bokor, P. Denham, and S. Rekawa, "Fourier-synthesis custom-coherence illuminator for extreme ultraviolet microfield lithography," *Applied Optics*, vol. 43, no. 5, pp. 820–825, 2003.

[6] H. J. Lenvinson, *Principles of Lithography, 2nd Edition*. SPIE Press, 2005.

[7] M. Foley and J. Munro, "Polymer fly's eye light integrator lens arrays for digital projectors," *Proceedings, Society for Information Display*, May 2000.

[8] F. Salmassi, P. Naulleau, and E. Gullikson, "Spin-on-glass coatings for the generation of super-polished substrates for use in the extreme ultraviolet regime," *Applied Optics*, vol. 45, no. 11, pp. 2404–2408, 2006.

[9] F. J. Weinberg and N. B. Wood, "Interferometer based on four diffraction gratings," *J. Sci. Instrum.*, vol. 36, pp. 227–330, 1959.

[10] E. N. Leith and B. J. Chang, "Space-invariant holography with quasicoherent light," *Applied Optics*, vol. 12, pp. 1957–1963, 1973.

[11] E. N. Leith and B. J. Chang, "Image formation with an achromatic interferometer," *Opt. Commun.*, vol. 23, pp. 217–222, 1977.

[12] B. J. Chang, *Grating-Based Interferometers*. PhD thesis, University of Michigan, 1975. University Microfilm 74-23-170.

[13] B. J. Chang, R. C. Alferness, and E. Leith, "Space-invariant achromatic grating interferometers: theory," *Applied Optics*, vol. 14, pp. 1592–1597, 1975.

[14] Y. S. Cheng, "Fringe formation in incoherent light with a two-grating interferometer," *Applied Optics*, vol. 23, pp. 3057–3059, 1984.

[15] Y. S. Cheng, "Temporal coherence requirement in a symmetric-path grating interferometer," *Applied Optics*, vol. 36, pp. 800–804, 1997.

[16] K. Patorski, "Talbot interferometry with increased shear," *Applied Optics*, vol. 24, pp. 4448–4453, 1985.

[17] K. Patorski, "Talbot interferometry with increased shear: further considerations," *Applied Optics*, vol. 25, pp. 1111–1116, 1986.

[18] Q. Liu and R. Ohba, "Effects of unparallel grating planes in Talbot interferometry," *Applied Optics*, vol. 38, pp. 4111–4116, 1999.

[19] Q. Liu and R. Ohba, "Effects of unparallel grating planes in Talbot interferometry II," *Applied Optics*, vol. 39, pp. 2084–2090, 2000.

[20] M. Wei, E. Gullikson, J. Underwood, T. Gustafson, and D. Attwood, "White-light spatial frequency multiplication using soft x-rays," *Proc. of SPIE*, vol. 2516, pp. 233–239, 1995.

[21] S. Horne, M. Besen, D. Smith, P. Blackborow, and R. DAgostino, "Application of a high-brightness electrodeless Z-pinch EUV source for metrology, inspection, and resist development," *Proc. SPIE*, vol. 6151, p. 61510P, 2006.

[22] J. W. Goodman, *Introduction to Fourier Optics, 2nd Edition*. McGraw-Hill Companies, 1996.

[23] E. Hecht, *Optics 3rd Edition*. Addison-Wesley Longman, 1998.

[24] P. Naulleau and C. Anderson, "Lithographic metrics for the determination of intrinsic resolution limits in EUV resists," *Proc. of SPIE*, vol. 6517, p. 65172N, 2007.

[25] C. Ahn, H. Kim, and K. Baik, "A novel approximate model for resist process," *Proc. of SPIE*, vol. 3334, pp. 752–763, 1998.

[26] SuMMIT software is distributed by EUV Technology, Martinez, CA 94553, http://www.euvl.com/summit.

[27] K. Goldberg, P. Naulleau, P. Denham, S. Rekawa, K. Jackson, E. Anderson, and J. A. Liddle., "At-wavelength alignment and testing of the 0.3 NA MET optic," *JVST B*, vol. 22, pp. 2956–2961, 2004.

[28] F. Houle, W. Hinsberg, M. Morrison, M. Sanchez, G. Wallraff, C. Larson, and J. Hoffnagle, "Determination of coupled acid catalysis-diffusion processes in a positive-tone chemically amplified photoresist," *JVST B*, vol. 18, pp. 1874–1885, 2000.

[29] "PROLITH is a registered trademark of KLA-Tencor, One Technology Drive, Milpitas, CA, 95035."

[30] C. Anderson and P. Naulleau, "Sensitivity study of two high-throughput resolution metrics for photoresists," *Applied Optics*, vol. 47, no. 1, pp. 56–63, 2008.

[31] P. Naulleau, C. Anderson, K. Dean, P. Denham, K. Goldberg, B. Hoef, B. L. Fontaine, and T. Wallow, "Recent results from the berkeley 0.3-NA EUV microfield exposure tool," *Proc. of SPIE*, vol. 6517, p. 65170V, 2007.

[32] D. Steenwinckel, J. Lammers, T. Koehler, R. Brainard, and P. Trefonas, "Resist effects at small pitches," *JVST B*, vol. 24, no. 1, pp. 316–320, 2006.

[33] R. Brainard, P. Trefonas, J. Lammers, C. Cutler, J. Mackevich, A. Trefonas, and S. Robertson, "Shot noise, LER and quantum efficiency of EUV photoresists," *Proc. of SPIE*, vol. 5374, pp. 74–85, 2004.

[34] P. Naulleau, E. Gullikson, A. Aquila, S. George, and D. Niakoula, "Absolute sensitivity calibration of extreme ultraviolet photoresists," *Optics Express*, vol. 16, pp. 11519–11524, 2008.

[35] R. Brainard, E. Hassanein, J. Li, P. Pathak, B. Thiel, F. Cerrina, R. Moore, M. Rodriguez, B. Yakshinskiy, E. Loginova, T. Madey, R. Matyi, M. Malloy, A. Rudack, P. Naulleau, A. Wuest, and K. Dean, "Photons, electrons and acid yields in EUV photoresists: a progress report," *Proc. of SPIE*, vol. 6923, p. 692325, 2008.

[36] V. Sarris, G. P. Patsis, V. Constantoudis, A. G. Boudouvis, and E. Gogolides, "A stochastic photoresist-polymer dissolution model combining the percolation and critical ionization models," *Japan Journal of Applied Physics*, vol. 44, pp. 7400–7403, 2005.

[37] J. Meiring, T. Michaelson, A. Jamieson, G. Schmid, and C. Willson, "Using mesoscale simulation to explore photoresist line edge roughness," *Proc. of SPIE*, vol. 5753, pp. 350–355, 2005.

[38] J. Taylor, *An introduction to error analysis, 2nd edition*. University Science Books, California, 1997.

[39] M. V. D. Kerkhof, W. D. Boeij, H. Kok, M. Silova, J. Baselmans, and M. Hemerik, "Full optical column characterization of DUV lithographic projection optics," *Proc. of SPIE*, vol. 5377, pp. 1960–1970, 2004.

[40] P. Dirksen, C. Juffermans, R. Pellens, M. Maenhoudt, and P. D. Bisschop, "Novel aberration monitor for optical lithography," *Proc. of SPIE*, vol. 3679, pp. 77–86, 2003.

[41] G. Robins, K. Adam, and A. Neureuther, "Measuring optical image aberrations with pattern and probe based targets," *JVST B*, vol. 20, no. 1, pp. 338–343, 2002.

[42] H. Nomura and T. Sato, "Techniques for measuring aberrations in lenses used in photolithography with printed patterns," *Applied Optics*, vol. 38, no. 13, pp. 2800–2807, 1999.

[43] J. Kirk, "Measurement of astigmatism in microlithography lenses," *Proc. of SPIE*, vol. 3334, pp. 848–854, 1998.

[44] P. Naulleau, J. Cain, K. Dean, and K. A. Goldberg, "Lithographic characterization of low-order aberrations in a 0.3-NA EUV microfield exposure tool," *Proc. of SPIE*, vol. 6151, p. 61512Z, 2006.

[45] J. Wyant and K. Creath, *Applied optics and optical engineering, Volume 11*. New York Academic Press, 1992.

[46] D. Malacara, *Optical Shop Testing, 2nd Edition*. Wiley, New York, 1992.

[47] R. Miyakawa, P. Naulleau, and K. Goldberg, "Analysis of systematic errors in lateral shearing interferometry for EUV optical testing," *Proc. of SPIE*, vol. 7272, p. 72721V, 2009.