In-situ optical testing of exposure tools via localized wavefront curvature sensing

Ryan Miyakawa[1], Xibin Zhou[2], Michael Goldstein[2,3], Dominic Ashworth[2], Kevin Cummings[2], Yu-Jen Fan[2], Yashesh Shroff[3], Gregory Denbeaux[4], Yudhi Kandel[4], and Patrick Naulleau[1]

[1]Center for X-ray Optics, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA

[2]SEMATECH, Albany, New York 12203, USA

[3]Intel, 2200 Mission College Blvd. Santa Clara, CA 95054-1549, USA

[4]College of Nanoscale Science and Engineering, University at Albany, NY 12203, USA

ABSTRACT

We present a new form of optical testing for exposure tools based on measuring localized wavefront curvature. In this method, offset monopole illumination is used to probe localized regions of the test optic pupil. Variations in curvature manifest as focus shifts, which are measured using a photodiode-based grating-on-grating contrast monitor, and the wavefront aberrations are reconstructed using a least-squares approach. This technique is attractive as it is independent of the numerical aperture of the system and does not require a CCD or a separate interferometer branch.

Keywords: optical testing, wavefront, aberration, in-situ, focus sensor

1. INTRODUCTION

As exposure tools continue to push the boundaries of resolution, it becomes increasingly difficult to characterize the aberrations in the system. Interferometric techniques based on reference waves such as phase-shifting point diffraction interferometry (PS/PDI) become difficult to realize due to the strict requirements on coherence and reference wave quality. Self-referencing tests such as lateral shearing interferometry (LSI) suffer from tight tolerances on grating and detector tilt.[1]

2. OVERVIEW

In this paper, we propose a new method of optical testing that is based on measuring the localized curvature at various points across the pupil. Aberrations in the optical system cause a departure in the local curvature from that of an ideal optic which manifests as a small focus shift in the image plane. These focus shifts are recorded at each pupil probe location using a grating-on-grating contrast monitor and are converted into a series of wavefront curvature maps, depending on the orientation of the gratings. The wavefront is then reconstructed from its curvature maps using a least-squares approach.

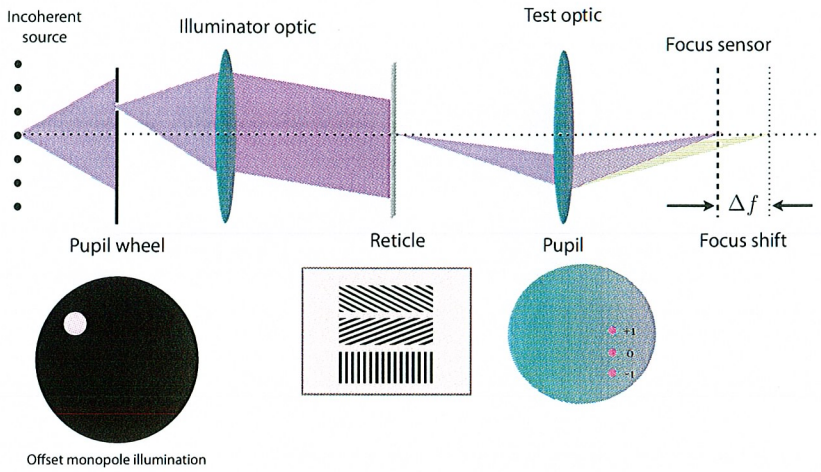

Figure 1. Schematic of measuring the local curvature of the test optic. Offset monopole illumination steers the reticle diffraction to specific points in the test optic pupil. A focus sensor in the image plane measures the plane of best focus

2.1 Setup

A schematic of the experimental setup is shown in Fig. 1.

Light from an extended incoherent source is incident on a pupil wheel containing several swappable illumination masks that control the range of spatial frequencies entering the system. Each illumination setting for this wavefront sensor uses a 0.15-σ offset monopole to steer the diffracted light from the reticle to a specific part of the test optic pupil. This light picks up some phase curvature caused by aberrations in the optic, which manifests as a small focus shift in the image plane.

The reticle contains three fields of 1:1 line-space grating patterns in three orientations: 0^o, 60^o, and 120^o. The pitch of these gratings determines the diffracted order separation and controls the spatial extent of the pupil probe. This probe size affects the sensitivity of the measurement, and is nominally chosen so that the ± 1 diffracted orders are $0.05 - 0.10$ times the entrance numerical aperture of the test optic. Figure 2.1 illustrates how illumination and grating orientation are used to control the pupil probe locations.

In the image plane, a focus sensor measures the plane of best focus for each grating orientation at each pupil probe location. Localized wavefront curvature is computed from the measured focus shifts and fed into an algorithm that reconstructs the aberrations in the system.

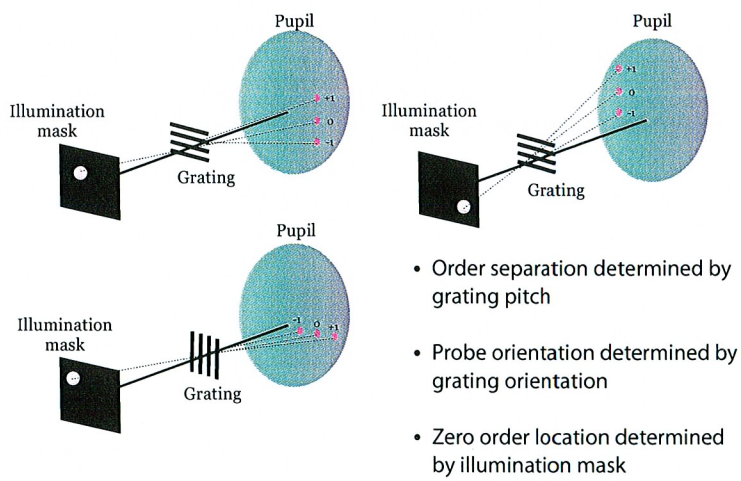

Figure 2. Pupil probes are controlled by illumination and grating orientation.

3. FOCUS SENSOR

The focus sensor implements a grating-on-grating contrast monitor that is described in the literature.[2] A schematic of the contrast monitor is shown in Figure 3.

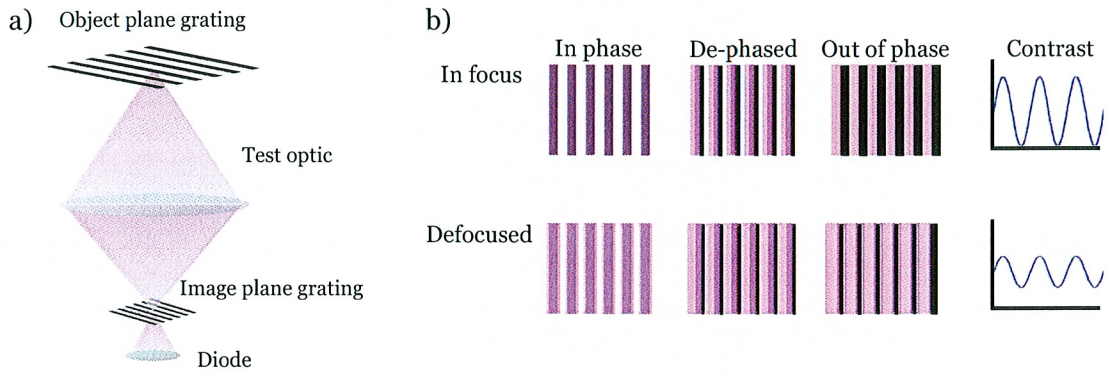

Figure 3. a) Focus sensor setup, b) Focus sensor working principle: the plane of best focus is given by plane which gives the highest contrast.

The reticle grating is imaged to a pitch-matched conjugate grating in the wafer plane which is scanned in the direction orthogonal to the grating lines. The integrated signal is captured by a photodiode and represents the convolution between the aerial image intensity and the image plane grating transmission function. After the signal is filtered to remove all but the 0 and ± 1 orders, the contrast computed as:

$$C_z = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} \tag{1}$$

This measurement is done at several locations through-focus, and the resulting curve is fit to a Gaussian model, where the plane of best focus is given by the value of z that maximizes the contrast.

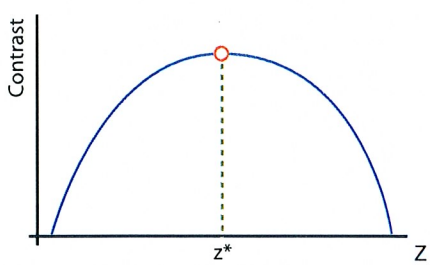

Figure 4. The plane of best focus is given by the z value that maximizes contrast

3.1 Accommodating multiple reticle grating orientations

Since the wavefront measurement requires the use of three separate grating orientations, the focus monitor must be designed accordingly. Ordinarily, the focus monitor would need to contain as many grating orientations as the reticle, however, we have replaced the image plane grating with a single hexagonal structure that is compatible with all three orientations shown in Figure 5.

Having a single structure on the wafer side means that there is no translational motion required on the focus sensor; all relative motions between the reticle and the focus sensor can be performed by the translating the reticle. This is important because stage requirements in a typical exposure tool will be much looser for the reticle than for the focus sensor due to the demagnification of the optical system.

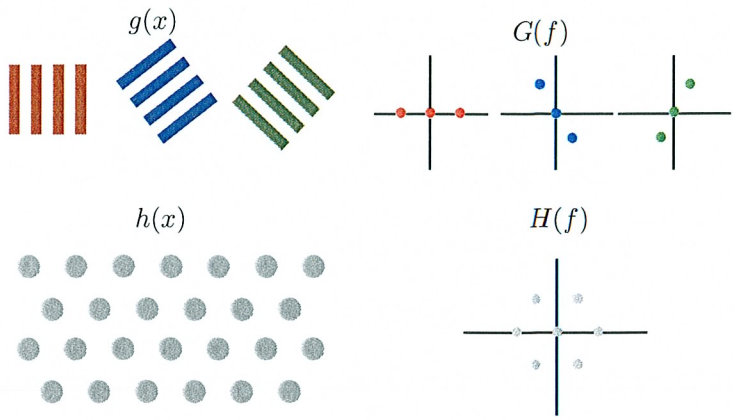

Figure 5. The hexagonal pinhole array $h(x)$ is compatible with all three orientations of gratings $g(x)$ on the reticle

4. WAVEFRONT RECONSTRUCTION

Here we outline the basic elements of the wavefront reconstruction by considering the coherent case (i.e. δ-function monopoles), although the extension to the partially coherent case is straightforward. We also consider an aberration signature that is represented by the summation of Zernike polynomials of maximum order N. Orders higher than N will cause errors in the wavefront reconstruction of comparable magnitude, and it is up to the user to specify an N that adequately represents the aberration signature in the optical system. Many exposure tool optics have higher order aberrations that are pre-characterized interferometrically. In these systems, where residual aberrations are generally low-order and caused by alignment and systematic effects, the higher order aberrations can be fed into the model to improve the accuracy of the reconstruction.

4.1 Overview

Because the algorithm uses curvature data to reconstruct the wavefront, Zernike terms without curvature: piston, x tilt, and y tilt, cannot be reconstructed. The recoverable system aberrations are represented by a vector $\vec{a} \in \mathbb{R}^{N-2}$ and the pupil function can be written:

$$W = \sum_{k=3}^{N} a_k Z_k \tag{2}$$

where Z_k represents the kth fringe Zernike polynomial.

Experimentally, the discretized curvature, defined as:

$$\frac{\Delta^2 W(\vec{u}_p)}{|\vec{\Delta u}_m|^2} = \frac{W(\vec{u}_p + \vec{\Delta u}_m) + W(\vec{u}_p - \vec{\Delta u},) - 2W(\vec{u}_p)}{|\vec{\Delta u}_m|^2} \tag{3}$$

will be measured at P locations across the pupil. Here the vector $\vec{u}_p$ specifies the location of the probe and $\vec{\Delta u}_m$ specifies the orientation of the curvature measurement ($\vec{\Delta u}_m$ also specifies the effective size of the probe which is related to the sensitivity of the measurement). If there are M curvature orientations, then there will be a total of $P \times M$ curvature measurements.

Applying the discretized curvature operator onto (2) gives:

$$\frac{\Delta^2 W(\vec{u}_p)}{|\vec{\Delta u}_m|^2} = \sum_{k=3}^{N} a_k \cdot \frac{\Delta^2 Z_k(\vec{u}_p)}{|\vec{\Delta u}_m|^2} \tag{4}$$

indicating that the measured curvatures will correspond to the sum discretized curvatures of the constituent Zernike polynomials evaluated at the same probe locations.

Since the process is linear, the relationship between the aberrations and the curvatures can be represented as a single linear operator C which maps a vector $\vec{a}$ of coefficients corresponding to the Zernike decomposition of the system aberrations to a set of curvature data $\vec{j}$. Let $V = \mathbb{R}^{N-2}$ be a vector space over $\mathbb{R}$, and $J \subseteq \mathbb{R}^{PM}$ be the image of C (PM is the dimension of the J, since every wavefront will give $P \times M$ curvature measurements).

$$C : V \to J \tag{5}$$

$$C(\vec{a}) = \begin{bmatrix} \Delta^2_{u_1}(\vec{p}) \\ \Delta^2_{u_2}(\vec{p}) \\ \vdots \\ \Delta^2_{u_M}(\vec{p}) \end{bmatrix} \cdot \vec{a} \equiv \vec{j} \tag{6}$$

where each $\Delta^2_{x_k}(\vec{p})$ represents a $P \times (N-2)$ block matrix of the form:

$$\Delta^2_{u_k}(\vec{p}) = \begin{bmatrix} \Delta^2 Z_3(p_1)/\Delta u_k^2 & \Delta^2 Z_4(p_1)/\Delta u_k^2 & \dots & \Delta^2 Z_N(p_1)/\Delta u_k^2 \\ \Delta^2 Z_3(p_2)/\Delta u_k^2 & \Delta^2 Z_4(p_2)/\Delta u_k^2 & \dots & \Delta^2 Z_N(p_2)/\Delta u_k^2 \\ \vdots & & & \vdots \\ \Delta Z_3(p_P)/\Delta x_k^2 & \Delta Z_4(p_P)/\Delta x_k^2 & \dots & \Delta Z_N(p_P)/\Delta x_k^2 \end{bmatrix} \tag{7}$$

Thus, in matrix form, $\mathbf{C}$ will be $P \cdot M$ rows by $(N-2)$ columns, and is constructed from the analytic forms of the Zernike polynomials. $\mathbf{C}$ represents a curvature library that relates pupil aberrations to a focus shift imparted on the probe. We can then write:

$$\mathbf{C} \cdot \vec{a} = \vec{j} \tag{8}$$

Equation (8) represents a over-specified system of linear equations, and a least-squares solution $\vec{a}^*$ can be solved by performing SVD on $\mathbf{C}$.

$$\vec{a}^* = \mathbf{C}^\dagger \cdot \vec{j} \tag{9}$$

4.2 Extension to finite monopoles

In the case that the offset monopoles have finite width, the effective curvature of each of the probes is given by the average curvature over the width of the probe. In this case, each element of $\Delta^2_{u_k}(\vec{p})$ integrated over the area defined by the illumination σ:

$$\text{coherent:} \quad \{\Delta^2_{u_k}(\vec{p})\}_{i,j} = \Delta^2 Z_{j+2}(p_i)/\Delta u_k^2$$

$$\text{partially coherent:} \quad \{\Delta^2_{u_k}(\vec{p})\}_{i,j} = \frac{\int_\sigma dS \cdot \Delta^2 Z_{j+2}(p_i)/\Delta u_k^2}{\int_\sigma dS} \tag{10}$$

The effect of this averaging is a reduction in sensitivity to curvature variations across the probe area. However, for Zernike terms of low radial order ($N < 24$), monopoles can be as large as $\sigma = 0.3$ without having a significant effect on the sensitivity of the reconstruction.

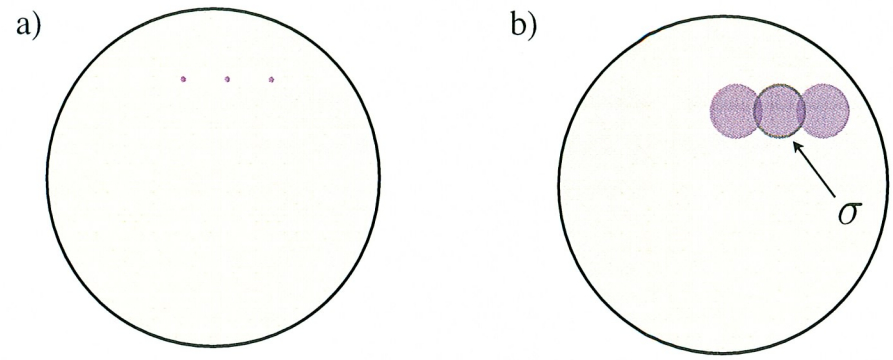

Figure 6. a) In the coherent case, curvature is computed from the Zernike values at the probe locations; b) For finite monopole width, curvature is integrated over the area defined by σ

4.3 Numerical stability and the condition number of C

The numerical stability of the operator C is paramount to the reliability of the reconstruction. Due to the effects of both noise and curvature contributions from higher order Zernike polynomials that are not being reconstructed, there will be a discrepancy $\vec{\Delta j}$ between the measured and ideal curvature maps. Therefore, C must be constructed in a manner such that this error does not translate into a disproportionately large error in the reconstructed aberrations $\vec{\Delta a}$. In other words, we must construct C so that equation (9) is well-conditioned.

A simple metric for the condition of C is the condition number, which can be represented as:

$$r = \frac{\sigma_{max}}{\sigma_{min}} \tag{11}$$

where σ_{max} and σ_{min} are the maximum and minimum singular values of $\mathbf{C}$. Broadly, a low condition number ($\log_{10} r < 2$) represents a well-conditioned problem, whereas a high condition number ($\log_{10} r > 6$) represents an ill-conditioned problem.[3]

As an example of ill-conditioned construction of C, consider the case where only 2 grating orientations (horizontal and vertical) are chosen. In this case, Z_5 (45^o astigmatism) is in the null space of C, because Z_5 has no curvature in the horizontal or vertical directions. I.e., from a curvature standpoint, Z_5 is geometrically equivalent to piston. Here the singular value of $\mathbf{C}$ corresponding to a_5 is 0, and the condition number $r = \infty$.

4.4 Optimizing probe locations using the condition number

When choosing the number, locations, and orientations of the probes, we strive to minimize the amount of data collection while still preserving a low condition number. For a fixed number of probes, we want to choose the probe configuration to minimize the condition number.

For example, if we are optimizing over the first 8 Zernike polynomials, and we are using 3 grating orientations and 4 probe points symmetrically positioned in a circle of radius R as shown in Figure 6a, we can ask which radius minimizes the condition number. Plotting the condition number vs radius gives the result in Figure 6b, which indicates that the optimal radius is $R = 0.47$.

In another example, we consider the first 15 Zernike polynomials, and we search for the minimum condition number in the case that we have two sets of 4 probe points each, located about circles of radii R_1 and R_2 respectively. In this case the optimal radii are found to be: $R_1 = 0.41$, $R_2 = 0.71$.

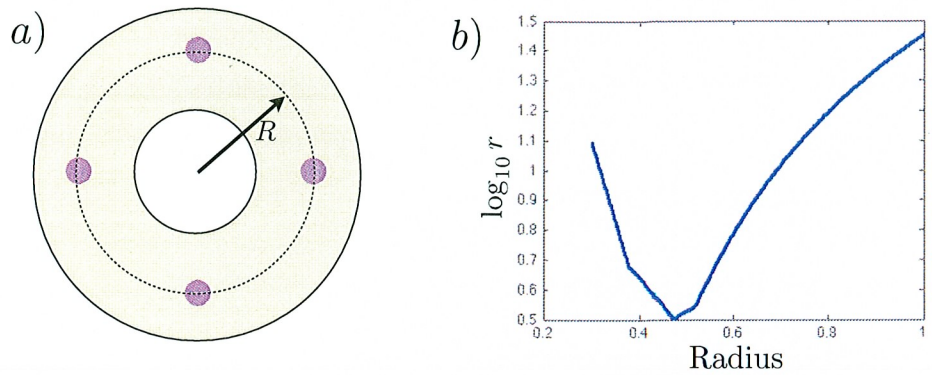

Figure 7. The influence of probe locations on the base 10 logarithm of the condition number of the reconstruction

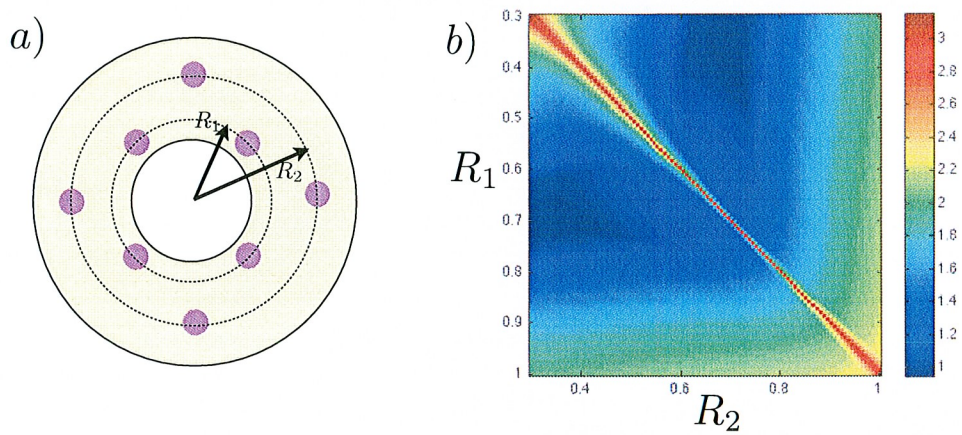

Figure 8. The influence of probe locations on the log of the condition number of the reconstruction for two rings of 4 probe points each. The color bar represents the base 10 logarithm of the condition number.

5. DISCUSSION

One of the key features of this reconstruction is its robustness to many external factors that directly affect the reconstructions in other techniques. Effects such as source fluctuation, photon and electronic noise, and stage uncertainty, all serve to reduce the contrast of the diode signal at each data point, but they do not bias it one way or another. This means that these effects do not generate errors directly; instead, they simply reduce the sensitivity of the experiment. This is in stark contrast to interferometry, for example, where noise in the interferogram can directly translate to aberrations.

The accuracy of the reconstruction is related to both the amount of data collected in the experiment and the number of Zernikes in the model. To reconstruct 15 Zernike polynomials, the reconstruction nominally requires 16 probe locations, 3 grating orientations, 11 focus steps, and 50 grating scan steps, which is expected to take around 1 hour to collect on the Sematch Albany MET. Simulations employing realistic experimental conditions in this configuration predict a wavefront accuracy of $\lambda_{EUV}/30$, which was computed as an average of 500 independent trials with randomized input wavefronts.

We have presented an in-situ optical test for measuring aberrations in exposure tools. Since the test uses a diode instead of a CCD, its integration is simpler than traditional interferometers. Additionally, it is independent of the numerical aperture of the system and therefore scales well to next generation tools. Moving forward, this technique provides a complementary method for wavefront metrology in systems with interferometric capabilities, as well as primary means for characterizing optics when interferometry is not available.

6. ACKNOWLEDGEMENTS

The authors would like to thank all of our collaborators at the Lawrence Berkeley National Lab, SEMATECH, Intel, and CNSE University at Albany.

This work was funded in part by SEMATECH and performed at Lawrence Berkeley National Laboratorys Advanced Light Source synchrotron facility which is supported by the DOE, Office of Science, Basic Energy Sciences.

REFERENCES

1. P. P. Naulleau, K. A. Goldberg, et al., "Extreme-ultraviolet phase-shifting point-diffraction interferometer: a wave-front metrology tool with subangstrom reference-wave accuracy," Applied Optics 38 (35), 7252-63 (1999),
2. Moest, Bearrach (Eindhoven, NL). Device and method for transmission image sensing. U.S. Patent 7675605, filed June 5, 2008, issued March 9, 2010.
3. A. Kline, C. Moler, G. Stewart, J Wilkinson, "An Estimate for the Condition Number of a Matrix," J. Numer. Anal., 16(2), 368375. (1979)

DISCLAIMER

This work was supported by the Director, Office of Science, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.