

# Nanometer-Level Semiconductor Imaging for Micrometer-Level MEMS

Burn J. Lin\*

Taiwan Semiconductor Manufacturing company, Ltd.  
168 Park Ave. 2, Hsinchu Science Park, Hsinchu County, Taiwan 308-44 R.O.C.

## ABSTRACT

Lithography spearheaded the semiconductor industry to nanometer-level feature sizes. The MEMS industry, having started later and being less developed in economy of scale, can take advantage of the experience of semiconductor patterning to make MEMS patterning cheaper and faster. Even though most MEMS devices are still in the micrometer regime, there are still many semiconductor lithography techniques to benefit from. Four types of lithography are used for MEMS fabrication: proximity printing, nanoimprint, projection printing, and maskless direct write. Projection printing stands out as the best candidate for MEMS high volume manufacturing.

The MEMS technology places more emphasis on DOF than resolution. It often requires fabrication of obliquely oriented devices. It also needs to test many innovative ideas before committing to mass production. Several semiconductor methods and some MEMS-specific methods to extend the depth of focus are covered. Multiple-e-beam direct-write systems are discussed with a focus on the suitability to high volume manufacturing in cost and lithographic performance for MEMS.

**Keywords:** optical lithography, e-beam lithography, maskless lithography, MEMS, projection printing, proximity printing, nanoimprint.

## 1. INTRODUCTION

There is a lot of synergy between imaging in semiconductor and that in MEMS. They both use lithography imaging equipment, even though the equipment for MEMS are generally less expensive and only have to handle micrometer resolution instead of nanometer. It is clear that both industries take advantage of miniaturization. They both are cost sensitive, thus need the economy of scale to grow. However, MEMS has an advantage. It can take advantage of the imaging experience in semiconductor lithography and shorten the development cycle.

MEMS fabrication takes advantage of many available imaging techniques. Proximity printing, due to its simplicity and low-cost equipment, is a popular imaging scheme. In addition to simplicity, oblique illumination to create a tilted resist image is quite easy. A full range of wavelength from visible<sup>1</sup>, uv<sup>2</sup>, to x-ray<sup>2</sup> can be used. However optical diffraction and the requirement of a contact-preventing gap between the mask and the wafer, limits the resolution of proximity printing to 0.5  $\mu\text{m}$  with deep-uv light<sup>3</sup> and no better than 0.25  $\mu\text{m}$  with 1-nm x-ray<sup>4</sup>. Proximity printing is a full-wafer exposure technique, facilitating high wafer per hour (wph) but for the same reason, the wafer size is limited by the size of the mask. If the popular 6-inch mask size is to be taken advantage, proximity printing cannot support wafers larger than 6 inch in diameter. The most severe handicap of proximity printing is probably the small proximity that the wafer has to be separated from the mask, making it defect prone. Frequent mask cleaning is necessary but cannot completely make it competitive in defect levels against projection printing. Alignment through a mask-to-wafer gap and across the entire wafer is inherently not very accurate. Hence, proximity printing is not a high volume manufacturing (HVM) technique for MEMS.

Another popular imaging technique for MEMS is nanoimprint<sup>5</sup>. Using a template as a mold to produce the resist image makes nanoimprint an extremely high resolution technique down to teen nanometers with high aspect ratio. However, it is an intimate contacting scheme. The defect level has to be similarly unacceptable as proximity printing. Moreover,

---

\* burnlin@tsmc.com; phone +886-3-563-6688 ext 722-5858; fax +886-3-578-7737.

the molding process can generate bubble defects. Mold lifting to separate the template from the wafer can break the delicate wafer image.

The mold has a much shorter lifespan than the optical mask. For HVM, child and grandchild molds have to be replicated using the same molding technique. They need large storage and complicated handling logistics. They have to be frequently inspected for damages. Molding is a slow pace process. Pressing the template into the fluid takes time to overcome the resistance from the viscosity of the fluid and to avoid creating bubble-producing turbulence. Waiting for bubbles to dissolve before molding takes time. Waiting for the fluid to solidify also takes time. Pulling out the mold is also a slow process. Over all, nanoimprint is a slow process.

Producing a uniform residual film under the bottom of the template not only takes time but also requires stringent spec of the flatness and roughness of the mold and the wafer as well as the forbiddance of nano-particles under the template. In addition to the above intra-template unevenness of the residual film, the thickness of the film between moldings is also difficult to control. Making the template accurately aligned to the previous layers on the wafer is also quite difficult.

Hence, just as proximity printing, nanoimprint is not a HVM technique for MEMS. This leaves projection printing using a stepper or scanner and direct write lithography as the candidates for MEMS HVM for low cost, high resolution, and low defect. However, the specific needs of MEMS for large depth of focus(DOF) and oblique incidence must be met.

We will discuss methods to increase DOF in a stepper or scanner environment, provision of oblique illumination, and enlargement of field size in this paper. Some of the techniques have been presented in the IEEE MEMS 2010 conference<sup>6</sup> and are presented here for completeness and with new insights wherever possible.

## 2. EXTENDING THE DOF

The resolution and DOF in optical projection printing are governed by the resolution and DOF scaling equations.

$$W = k_1 \cdot \frac{\lambda}{\sin \theta} = k_1 \cdot \frac{\lambda}{NA} \tag{Eq. (1)}$$

$$DOF = k_3 \cdot \frac{\lambda}{\sin^2(\theta/2)} = k_3 \cdot \frac{\lambda}{NHA^2} \tag{Eq. (2)}$$

Where  $W$  is the feature width;  $\lambda$ , imaging wavelength in the imaging media;  $\theta$ , aperture half angle as shown in Fig. 1.  $NA$  is the numerical aperture;  $NHA$ , the numerical half aperture. Explanation of these equations can be found in most lithography literature<sup>7</sup>. The scaling coefficients  $k_1$  and  $k_3$  are for resolution and DOF, respectively. When  $\theta$  is increase to gain resolution, DOF is traded off.

With the help of these two fundamental equations, we now discuss four methods to extend DOF, namely 2-beam imaging, optimizing  $k_1$  and  $NA$ , focus drilling, and stacked mask.

### 2.1 2-beam imaging

If the illumination is coherent, infinite DOF can be achieved for a fixed pitch. Even though we can only use partially coherent illumination in actual patterning, 2-beam imaging provides more DOF than with 3-beam imaging<sup>8</sup>.

The 3-beam and 2-beam imaging techniques are depicted in Fig. 1. When a grating of pitch  $p$  is illuminated, it diffracts the light into many orders. For semiconductor imaging, a high-NA lens to capture many orders of diffraction for the minimum feature is not economical. Only

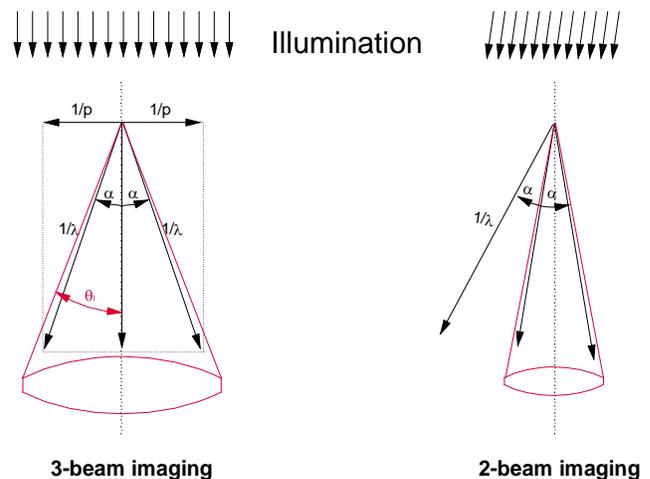


Fig. 1 3-beam and 2 beam imaging

the 0<sup>th</sup> and the 1<sup>st</sup> orders need be included, as shown in the figure. The angle  $\alpha$  that the 1<sup>st</sup> order beam made with the 0<sup>th</sup>-order beam is  $\sin^{-1}(p/\lambda)$ . In the 3-beam case, when  $\alpha$  is smaller than the aperture angle  $\theta$ , the grating can be resolved by the imaging lens. Two-beam illumination is achieved by tilting the illumination off axis. This way, only the 0<sup>th</sup> and one of the 1<sup>st</sup> order beams are captured by the imaging lens.

Analytically, the electric field of the -1<sup>st</sup> order beam C is,

$$E_c = A_c \exp[i\kappa_1 z_1 \cos\alpha_1 + i\kappa_1 x_1 \sin\alpha_1] \quad \text{Eq. (3)}$$

where  $A$  is the amplitude of the light wave and  $\kappa_1 \equiv 2\pi/\lambda_1$  in media 1. Similarly, the electric field of the 0<sup>th</sup> and 1<sup>st</sup> order beams, D and E, in media 1 are,

$$E_d = A_d \exp[i\kappa_1 z_1] \quad \text{Eq. (4)}$$

and

$$E_e = A_e \exp[i\kappa_1 z_1 \cos\alpha_1 - i\kappa_1 x_1 \sin\alpha_1]. \quad \text{Eq. (5)}$$

The total intensity passing through the lens is the square of the absolute amplitude of the summation of the three fields.

$$\begin{aligned} I &= [E_d + E_c + E_e]^* [E_d + E_c + E_e] \\ &= a_d^2 + 2a_c^2 \\ &\quad + 4a_d a_c \cos\kappa_1(x_1 \sin\alpha_1) \cos[\kappa_1 z_1(1 - \cos\alpha_1) - \delta_c] \\ &\quad + 2a_c^2 \cos\kappa_1(2x_1 \sin\alpha_1) \end{aligned} \quad \text{Eq. (6)}$$

where  $\delta_c$  is the phase difference between beams E and C. the x-dependent terms in Eq. (6) determines the resolution of the 3-beams and the z-dependent terms, the DOF. Let  $\kappa_1 z_1(1 - \cos\alpha_1) - \delta_c = 0$  be the focal plane. We can define the DOF as between  $\kappa_1 z_1(1 - \cos\alpha_1) = \delta_c - \eta$  and  $\delta_c + \eta$ .  $\eta$  is a phase term which is defines the acceptable defocus.

$$\kappa_1 \Delta z (1 - \cos\alpha_1) = 2 \eta. \quad \text{Eq. (7)}$$

$$DOF = \frac{\eta \lambda_1}{\pi(1 - \cos\alpha_1)} = k_3 \frac{\lambda_1}{\sin^2 \frac{\alpha_1}{2}} \quad \text{Eq. (8)}$$

Hence, the DOF of 3-beam imaging is just as the DOF predicted by Eq. (2).

In 2-beam imaging, the two interfering beams are,

$$E_c = A_c \exp i\kappa_1 [z_1 \cos\alpha_c - x_1 \sin\alpha_c] \quad \text{Eq. (9)}$$

and

$$E_d = A_d \exp i\kappa_1 [z_1 \cos\alpha_d - x_1 \sin\alpha_d] \quad \text{Eq. (10)}$$

The combined intensity is

$$I = [E_c + E_d]^* [E_c + E_d] \quad \text{Eq. (11)}$$

$$= a_c^2 + a_d^2 + 2a_c a_d \cos[\kappa_1 x_1 (\sin \alpha_c - \sin \alpha_d) - \kappa_1 z_1 (\cos \alpha_c - \cos \alpha_d) + \delta]$$

When  $\alpha_c = -\alpha_d$ , Eq. (11) becomes

$$I = a_c^2 + a_d^2 + 2a_c a_d \cos[2\kappa_1 x_1 \sin \alpha_c + \delta] \quad \text{Eq. (12)}$$

There is no  $z$  degradation in  $I$ . The DOF is infinite. This is the case for 2-beam interference when the illumination is coherent and the tilt angle is optimized for the particular grating pitch. At off-ideal conditions but not as extreme as isolated patterns, the DOF of 2-beam interference is expected to be better than that of 3 beams. The best approach is to use simulation to evaluate the expected improvement.

### 2.2 Optimize NA and $k_1$

Equations (1) and (2) can be used to optimize DOF with  $k_1$  and  $\lambda$ , at a given requirement of resolution. For example, for 0.5- $\mu\text{m}$  resolution at  $\lambda=365$  nm, using  $k_1=0.6$  needs  $\text{NA}=0.438$  to reach  $\text{DOF}=1.66$   $\mu\text{m}$  based on  $k_3=0.23$ . With  $k_1=0.4$ , DOF is now 3.85  $\mu\text{m}$ , based on  $\text{NA}=0.292$  and also  $k_3=0.23$ . For the same resolution, DOF is 3.27 and 7.40  $\mu\text{m}$  for  $k_1=0.6$  and 0.4 respectively, if  $\lambda=193$  nm is used instead of 365 nm. The situation is plotted in Fig. 2, where the wavelength tradeoff to the DOF of 0.2- $\mu\text{m}$  resolution is also shown.

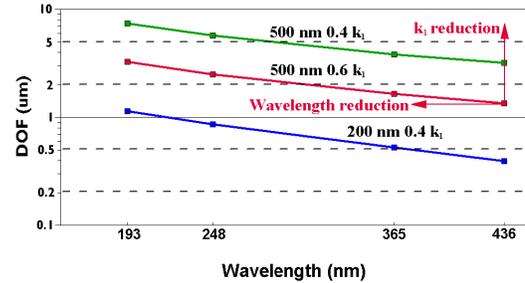


Fig. 2 Optimize  $k_1$  and NA for DOF.

### 2.3 FLEX – Focus Drilling

A series of exposures at different focal planes can be superimposed to gain DOF as proposed by Hayashita et al<sup>9</sup>. It takes advantage of extra image contrast in the focal plane to compensate for lower contrast at the out-of-focus planes. The technique is depicted in Fig. 3 showing the superposition of two exposures with a shift of focus and that of three exposures to extend the DOF even further. FLEX is a focus drilling technique. The acronym stands for Focus Latitude Enhancement eXposure. The practice of FLEX on a stepper and a scanner is shown in Fig. 4. In the former case, the wafer is exposed twice or three times with a shift of its longitudinal position at each exposure. For the latter, multiple exposures are avoided by just tilting the wafer during scan. This way, the defocal images are continuously superimposed.

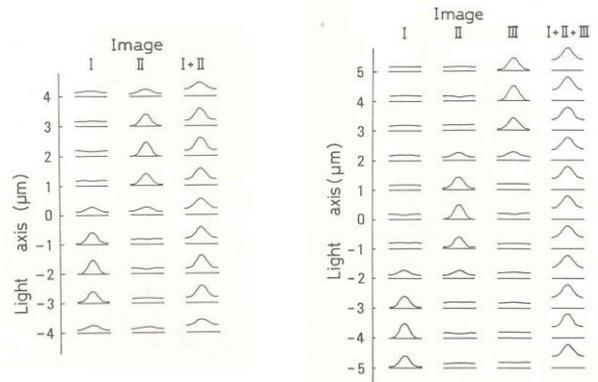


Fig. 4 Focus drilling - FLEX

Focus drilling does not have to be limited by mechanical means. The imaging wavelength can be slightly shifted during laser pulsing to achieve superposition of defocused images<sup>10</sup>.

### 2.4 Stacked Mask

FLEX can extend DOF by a factor of two or three. However, MEMS imaging over extremely large topography is often required. If only two discrete planes have to be patterned, the usual practice is to expose twice with reloading of mask and refocusing of the substrate in the longitudinal direction. Double exposures and re-

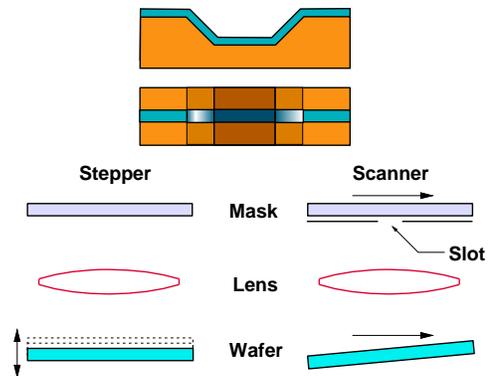


Fig. 3 Practicing focus drilling on a stepper and a scanner.

loading masks reduce the throughput of the imaging tool, resulting in less productivity and higher cost. There are also possibilities of misalignment associated with multiple exposures.

The mask stacking technique proposed here combines the two exposures with a composite mask as shown in Fig. 5. One mask is turned upside down and stacked with a normally oriented mask with a spacer at the periphery. The height of the spacer is  $m^2h$  where  $m$  is the reduction ratio of the imaging system and  $h$  is the vertical distance of the two largely separated wafer planes to image. The two masks can be pre-aligned with care to ensure consistent alignment at the wafer level. This alignment at the masks is desirable not because it is a one-time effort but the alignment accuracy is also repeatable. The only foreseeable challenge of this technique is in the upside-down imaging of the lower mask. Steppers and scanners are designed for imaging with the mask absorber facing the imaging lens, with its substrate away from the critical imaging path, so that the flatness and homogeneity of the mask substrate are not critical. With those MEMS patterns that have redundant resolution capability, the flatness and homogeneity of commercially available masks may suffice. If not, the substrate homogeneity and flatness will have to be specified tighter with a higher cost. The cost has to be weighted against the productivity loss of the imaging tool.

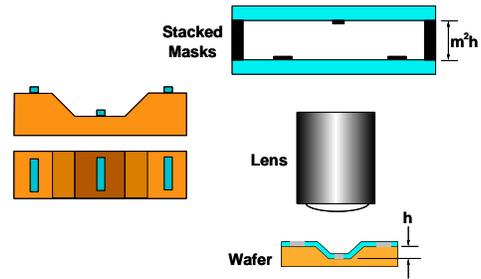


Fig. 5 Imaging on two largely separated planes with a composite mask.

### 3. OBLIQUE ILLUMINATION

Semiconductor manufacturing seldom calls for an oblique beam such as for Inclined Lithography to make fiber holders<sup>11</sup>. It is relatively easy to modify a proximity printer for this but achieving oblique illumination on a stepper or scanner for HVM is not trivial. The off-axis illumination setup on these projection printing tools may help to realize an obliquely incident beam. Such set up is shown in Fig. 6. With normal 2-beam imaging, the aperture stop of the lens is centered, with a sufficient opening to accept the diffraction angle of the two centered beams. With oblique imaging, the aperture stop is off center to select the two off-center beams.

Let the tilting angle be  $\phi$ . Then, the oblique illumination angle is  $\alpha/2 + \phi$ . The total extent of the diffracted beams cannot exceed the lens aperture angle  $\theta$ . Therefore,  $\alpha + \phi \leq \theta$  sets the limit of  $\phi$ . Table 1 lists some sample settings. To achieve  $45^\circ$  inclination, an I-line NA=0.73 stepper suffices. However, such high NA is not available. One needs to use KrF tools for that NA at a higher cost.

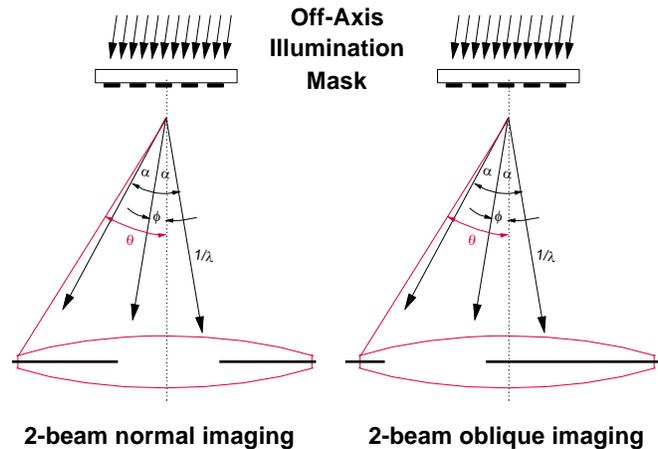


Fig. 6 Off-axis illumination for oblique imaging.

Table 1 Sample settings for oblique imaging.

Pitch ( $\mu\text{m}$ )	$\alpha$ (deg)	$\alpha/2+\phi$ (deg)	$\lambda$ (nm)	NA
5	4.19	45	365	0.73
5	2.84	45	248	0.72
5	4.19	30	365	0.53

### 4. MEB DW

Multiple e-beam direct write (MEB DW) systems offer high resolution, large DOF, flexibility, and freedom from masks. The high resolution aspect may not be important for MEMS applications, except for NEMS. However, there are many other useful aspects for MEMS. One MEB DW system particularly useful for MEMS, is a reflective e-beam lithography (REBL) system developed by KLA-Tencor<sup>12</sup>, as shown in Fig. 7. It uses a beam bender to redirect the incident

beam from the illumination optics, so that the illumination beam perpendicularly incidents on the programmable reflective mask called dynamic pattern generator (DPG), which is basically a CMOS circuit with its last metal layer facing the incoming beam. Each metal pad is programmed to switch between 0 and 2 V. The illuminating beam is decelerated substantially before reaching the DPG. The pads at zero voltage absorb the incoming electrons while those at 2 V reflect them. The reflected beam is accelerated back to either 50 or 100 keV and demagnified to the desired spot size. As many as 2000x2000 pixels can be fabricated on the DPG. REBL is virtually an e-beam projection printing system with an e-beam reflection mask that can be programmed at electronic speed. The projected image is scanned across a group of wafers on a rotary or linear stage. Here, the rotary stage is shown.

The impact of REBL on MEMS is multifold. First, because of scanning and the maskless feature, the field size need not be confined to the typical 26 x 33 mm<sup>2</sup> area provided by all scanners for at least eight technology nodes. With REBL, the field size can be as large as the entire wafer as long as sufficient electronics is installed in the datapath of the system.

Second, with a 50~100 keV system, the DOF can be extremely large. A recent simulation on a 100-keV system shows a capability of many micrometers. Exposure-Defocus (E-D) trees<sup>13</sup> of isolated and dense features at the center and the edge of the electro optics field are constructed from simulation results. The simulation is based on the requirement of ion implanting wells requiring a resist thickness in the order of 700 nm and feature size of 150 nm in 300-nm pitch. The beam current on wafer is 4 μA at 16 mrad, with 20-nm acid diffusion length and the resist blur as a function of the beam blur at the best focus. The E-D trees and the corresponding E-D window are shown in Fig. 8. The E-D window measures 26% in exposure latitude and 5.9 μm in DOF. It is more than sufficient for ion implantation and many MEMS applications.

Third, there is no need to make mask, thus saving mask cost, cycle time, and the subsequent problems of mask contamination, damage, and inspection. The mask CD control budget and image placement budget may all be reallocated to the wafer.

Fourth, the throughput and cost of the system can be scaled up and down for the application<sup>14</sup>. For HVM, the throughput can be as high as 196 wph for 130-nm half pitch at 20% pattern density using 9 e-beam columns on the multi-wafer stage, with cost less than scanners with equivalent throughput. For research, the number of columns and the size of the datapath can be reduced for lower cost. Because of the leverage of HVM, the scaled down system can enjoy the debugging efforts on the HVM systems to make it reliable and economical.

## 5. CONCLUSION

For MEMS to go into HVM, it has to adopt HVM tools in semiconductor manufacturing. These tools can be developed for unique MEMS applications such as large DOF, oblique illumination, flexible field size, and low cost.

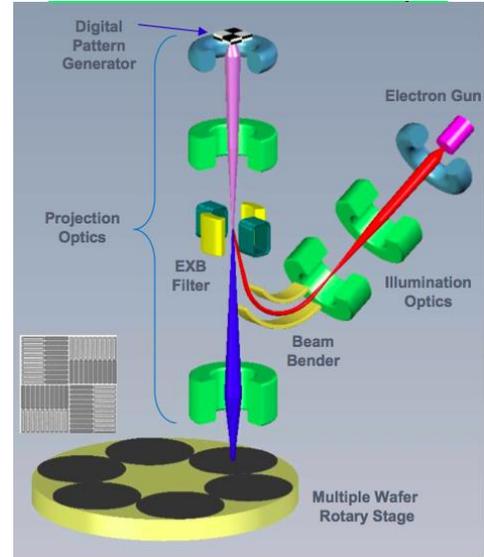


Fig. 7 REBL MEB DW system.

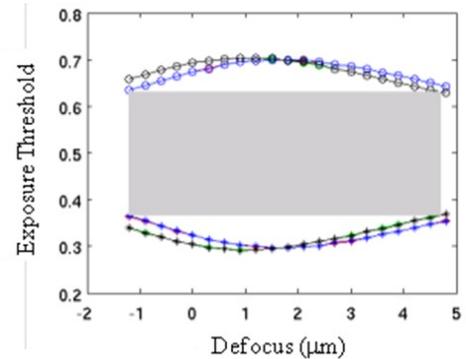


Fig. 8 E-D trees and E-D window for well implant.

## REFERENCES

- 
- <sup>1</sup> J.S. Kilby, "Miniaturized Electronic Circuits", US Patent 3,138,743, 1964.
  - <sup>2</sup> B.J. Lin, "A Comparison of Projection and Proximity Printings -- From UV to X-Ray", *Microelectronic Engineering* vol. 11, p. 137, Elsevier Science Publishers B.V., 1990.
  - <sup>3</sup> B.J. Lin, "Optical methods for fine line lithography" in "Fine line lithography", edited by R. Newman, North-Holland Publishing Co. 1980.
  - <sup>4</sup> M. Hasegawa, Y. Nakayama, K. Yamaguchi, T. Terasawa and Y. Matsui, "Printing characteristics of proximity x-ray lithography and comparison with optical lithography for 100-nm node and below", *SPIE Proceedings*, vol. 3997, p. 95, 2000.
  - <sup>5</sup> M. Coburn et al, "Step and flash imprint lithography: A new approach to high-resolution patterning", *SPIE Proceedings* vol. 3676, 379, 1999.
  - <sup>6</sup> B.J. Lin, "Can MEMS take advantage of advances in semiconductor lithography?", *Proceedings of The 23<sup>rd</sup> IEEE International Conference on MEMS 2010*.
  - <sup>7</sup> B.J. Lin, "Optical Lithography: Here Is Why" SPIE Press 2009.
  - <sup>8</sup> Ch. 3 of Reference 7.
  - <sup>9</sup> T. Hayashida, H. Fukuda, T. Tanaka, and N. Hasegawa, "A novel method for improving the defocus tolerance in step and repeat photolithography", *SPIE Proceedings*, vol. 772, p. 66, 1987.
  - <sup>10</sup> I. Lalovic et al, "Focus drilling for increased process latitude in high-NA immersion lithography", *SPIE Proceedings* vol. 7973, 797328, 2011.
  - <sup>11</sup> Z. Ling, and K. Lian, "New fabrication techniques of SU-8 fiber holder with cantilever-type elastic microclips by inclined UV lithography using single Mylar mask", ISSN 0946-7076 (print) 1432-1858 (online), *Microsyst. Technol.*, 2009.
  - <sup>12</sup> P. Petric, "Reflective electron beam lithography(REBL), a novel approach to high speed maskless e-beam direct write lithography", *52<sup>nd</sup> International Conference on Electron, Ion, and Photon Beam Technology and Nanofabrication (2008)*
  - <sup>13</sup> Chapter 4 or Reference 8.
  - <sup>14</sup> B.J. Lin, "Future of MEB DW systems", *SPIE 2012 Advanced Lithography Symposium [8323-1]*.