Photolithography: Capabilities and Limitations

Gregg M. Gallatin* and J. Alexander Liddle
Center for Nanoscale Science and Technology
National Institute of Standards and Technology
Gaithersburg, MD 20899

INTRODUCTION

Photolithography is the workhorse of semiconductor computer chip manufacturing. It has evolved extraordinary capabilities over the past half-century. However, it does possess some fundamental limitations. Here we first detail its capabilities by way of debunking the three most common misconceptions about photolithography: (1) photolithography tools ("litho tools" for short) cannot print features significantly smaller than the wavelength of the light they use, (2) litho tools are the most expensive aspect of semiconductor fabrication and (3) the only thing a litho tool does is print small features somewhere on a silicon wafer. We then discuss photolithography's limitations and discuss the emerging techniques being developed to avoid them.

DEBUNKING THE MISCONCEPTIONS

1: Litho tools can't print features much smaller than the wavelength of light.

A typical litho tool operates by projecting a 4 x demagnified image of a mask pattern onto a thin (~ 100 nm thick) layer of photoactive material, called photoresist, which has been spin-cast onto the surface of a silicon wafer. For a positive (negative) tone photoresist the regions exposed to high intensity become soluble (insoluble) whereas those exposed to little or no light remain insoluble (soluble).[1,2,3] During development the soluble regions wash away, while the insoluble regions remain, capturing the desired pattern information on the wafer. After development, the wafer is processed, the

developed photoresist is stripped away, a new layer is spun on and the whole process is repeated.

Figure 1(a) provides a schematic that illustrates how a litho tool exposes the photoresist. For simplicity we consider the image of a square grating with a period small enough so that only the plus and minus first diffraction orders make it through the optical system. The transmitted diffraction orders are plane waves incident on the wafer at angles $\pm \theta$ and they interfere to produce a sinusoidal image intensity with a period of $\lambda/(2\sin(\theta))$ where λ is the wavelength. The sine of the maximum angle that can pass through the optical system is referred to as the "numerical aperture" or *NA* of the optics.[1,2,3] Although the image intensity is sinusoidal the developed photoresist pattern is a square wave due to the extreme nonlinearity of the photoresist development process, which effectively acts as a thresholding function. If the processing conditions are such that lines and spaces print with equal width – "dense lines and spaces" – then the width of the photoresist features will be $(1/2)\lambda/(2NA) = \lambda/(4NA)$. However, the nonlinearity of the response of the photoresist to the image intensity enables features much smaller than $\lambda/(4NA)$ to be produced simply by over-exposing the photoresist (Figure 1(b)). Over-etching the features when transferring them into the wafer can also shrink their width. This results in features with dimensions well below $\lambda/(4NA)$.

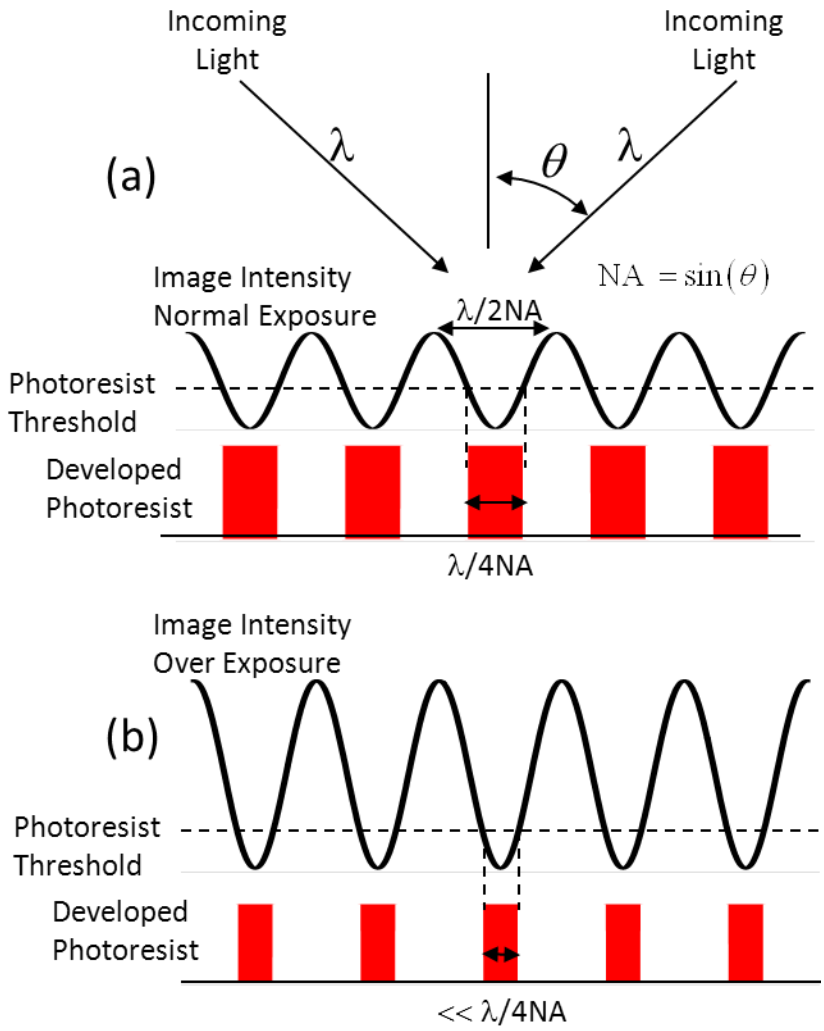

Figure 1: (a) Incoming plane waves of light at angles $+\theta$ and $-\theta$ produce a sinusoidal intensity pattern with period $\lambda/2NA$ where λ is the wavelength of the light and the numerical aperture, $NA = \sin(\theta)$. The nonlinear behavior of the photoresist means that for a normal exposure intensity the photoresist pattern produced comprises equal size lines and spaces with widths $\lambda/4NA$ and vertical side walls, i.e., a square wave. (b) Overexposing the photoresist leads to linewidths much smaller than $\lambda/4NA$ but the periodicity of the pattern remains fixed at $\lambda/2NA$. Feature sizes can be reduced as long as all 10^{12} to 10^{14} of them are in the right place and within ±10 % of the desired size

So, what does determine the minimum production-worthy printable feature size? It is set by the point at which it is no longer possible to maintain control over the size of the 10^{12} to 10^{14} features on every 300 mm (12 inch) diameter wafer to be within ± 10 % of the

desired dimension. Figure 2 shows a plot of the evolution of wavelength and minimum feature size (commonly referred to as the critical dimension or “CD") since 1980. Clearly, present day litho tools are printing features, to the ± 10 % specification, which are much smaller than the wavelength.

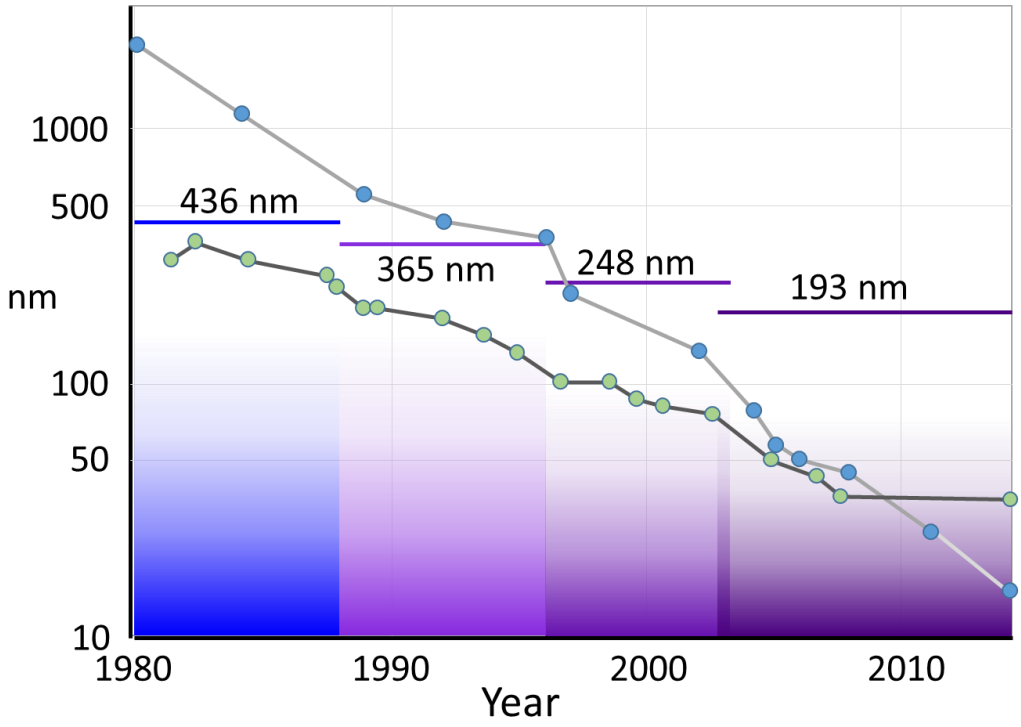

Figure 2: Evolution of minimum feature size or critical dimension (CD) (Connected curve with blue dots), exposure wavelength (horizontal bars), and minimum feature half-pitch ($\lambda/4NA$) (Connected curve with green dots) as a function of time since 1980. Before 1996-1997 CDs exceeded the wavelength. Since then, CDs have continued to shrink, passing the limit of $\lambda/4NA$. Immersion lithography at a wavelength of 193 nm has enabled printing features on the order of 20 nm and below which meet specification in production. The most recent Minimum Feature Size data point is 16 nm. Minimum feature size data is courtesy of Dr. Tim Brunner, IBM. $\lambda/4NA$ calculated from data in Matsuyama et al.[4]

2. Litho tool cost is the major contributor to chip cost

Present day litho tools cost ~$50-$80 million. But in a high volume manufacturing (HVM) environment, such as a chip making factory, they operate 24 hours a day, 7 days

a week, 365 days a year, at > 90 % up time, generally for at least 5 years. A state-of-the-art 193 nm wavelength litho tool requires only about 20 seconds to 30 seconds to completely expose a 300 mm diameter silicon wafer.[1,5,6] This corresponds to a "throughput" of ~120 to ~180 wafers per hour. Thus, during its 5 year lifetime, a tool will expose on the order of 5 million wafers which amortizes the ~ $50-80 million tool cost to only about $10-$16 per full-wafer exposure. A fully patterned 300 mm diameter wafer will yield a few hundred to a few thousand chips depending on the chip size. For a chip requiring 30 exposures to complete it, the tool cost contributes from only a few cents to a few dollars to the cost of each chip.

In fact, the major driver of chip cost is generally not the litho tool itself but the set of masks containing the different patterns needed to fabricate the chip.[5,6] As explained above, present day litho tools are printing features much smaller than the wavelength of the light they use and in this regime interference and diffraction effects[1,2,3] are extremely strong. In order to print such small features to meet specification, the patterns on the mask (coupled with the mask illumination) must be modified using so-called resolution enhancement techniques (RET) to compensate for these interference and diffraction effects. Mask patterns that can perform such compensation are extremely complex, and therefore very expensive, both to design and to fabricate. A single mask for printing features of the minimum size can cost from $100000 to $500000.[5,6] The contribution the mask makes to the cost of a single wafer exposure is simply the cost of the mask divided by the total number of times it is used to expose a wafer. This can be as low as about 1000 exposures for logic chips and as high as about 20000 exposures for memory chips.[6]

In the low mask usage case (~ one thousand exposures per mask) a $500000 mask contributes ~ $500 to the cost of printing one wafer level. The litho tool cost and the mask cost become roughly equal only in the high mask usage case (~ tens of thousands of exposures per mask). In fact, International Sematech notes that at the 32 nm node, if 5000 wafers are printed from a mask set "…the cost of an optical mask far outweighs litho tool costs because of the excessive data [*i.e. pattern complexity*] to write these masks".[5] The results of a complete cost of ownership (COO) model are discussed in Ref[5].

Here we simply illustrate the average cost per exposure over the lifetime of the tool using the simplified COO model[1,2,3]

Cost/Exposure = Tool Cost/(Exposures/Tool) + Mask Cost/(Exposures/Mask)

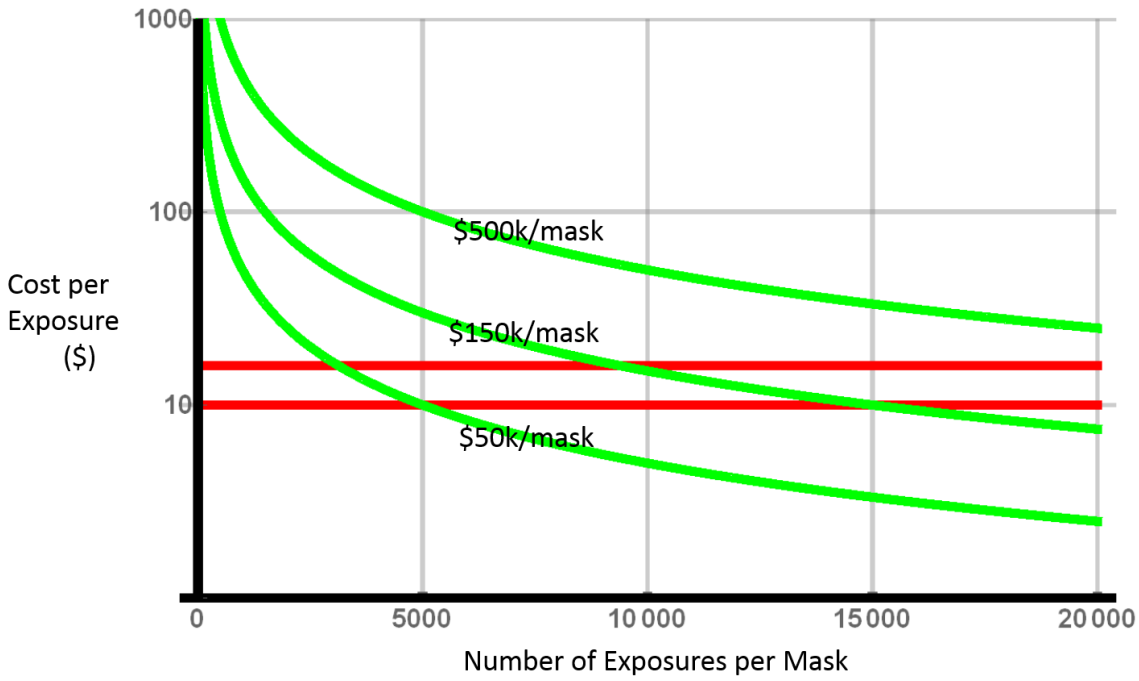

Figure 3: Result of the simplified COO model from Sematech for various technologies.[1,2,3] The two horizontal red lines indicate the range of amortized tool cost per full-wafer exposure, $10 to $16, as discussed in the text. For low mask usage, i.e, less than a few 1000 wafers/mask , the cost of the masks far outweighs the cost of the litho

tool. Depending on the mask cost the two become comparable in price only for high mask usage.

3. Litho tools just print small features somewhere on a wafer.

Litho tools do far more than just print small features somewhere on a wafer. Not only must all the 10^{12} to 10^{14} features be controlled to be within a few nanometers of their required size, they must also be placed within a few nanometers of their required position – not only relative to each other but also relative to the features and structures already on the wafer.[1,5] The accurate placement of successive patterns on top of one another on a wafer is termed "overlay".[1]

A computer chip is made of multiple patterned layers and for it to function all the patterns must properly overlay each other. Note that this does *not* require the patterns to be perfectly scaled or free of distortion in an absolute sense. The patterns can be distorted as long as they overlay each other with sufficient accuracy. Indeed computer chips expand and distort when operating due to differential heating and cooling and this distortion does not stop them from working. When a new wafer and/or a mask is loaded into a litho tool, the tool automatically senses and corrects the projected image of the pattern so that it properly overlays those already on the wafer. This "alignment" process not only translates and rotates the pattern as it is projected onto the wafer, it also distorts the pattern, accounting for all linear as well as many nonlinear distortion terms (see Figure 4) so as to achieve sufficient overlay. The alignment process is automatic and is performed at the production throughputs of 120 to 180 wafers per hour on each and every 300 mm diameter wafer the tool exposes with a given mask. The standard specification for how well patterns must overlay each other is that the pattern positions rotations and distortions

have to match to better than ~ 25 % of the feature size. At 22 nm the overlay has to be better than ≈ 5 nm. Summarizing the current state of the art, we see that a modern-day optical litho tool working in production will print 10^{12} to 10^{14} deep sub-wavelength (~ 30 nm scale) features per wafer at a rate of about one wafer every 20 to 30 seconds. All the features, with literally only a handful of defects per wafer, will be within a few nanometers of their desired size and within a few nanometers of their desired position, all for a cost of about $10 per exposure. This sets the bar for competing technologies.

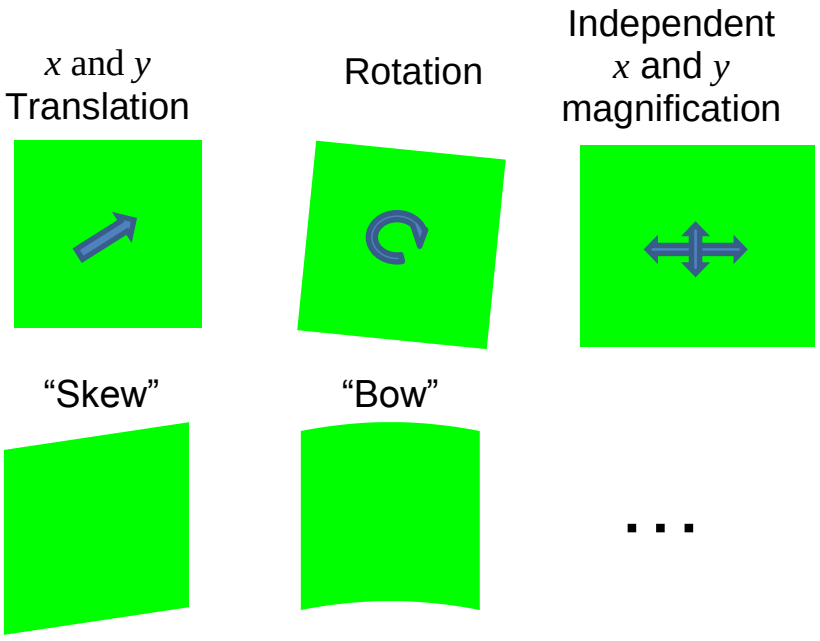

Figure 4: Some low order alignment/distortion terms that are automatically sensed and corrected for by the litho tool. Each green "square" represents a single ~ 26 mm X 32 mm exposure field on the 300 mm diameter wafer.

MAIN LIMITATION OF PHOTOLITHOGRAPHY

As impressive as the capabilities of photolithography tools are, they are suited physically and economically only to the production of integrated circuits. Further, as feature sizes become smaller, the cost of the lithography needed to maintain feature density is increased indirectly, through increasing mask costs, compounded by the introduction of

double, even quadruple,[7] patterning schemes and their associated additional process costs.[8] As we discuss below, this may open the door to alternative patterning techniques.[9] In addition, the drive to make circuits more capable by, for example, integrating on-chip optical communication systems, or sensors, is prompting the examination of numerous fabrication methods that can be used in conjunction with photolithography to add functionality to chips.

HISTORICAL CHALLENGES TO PHOTOLITHOGRAPHY

Over the years many technologies have attempted to displace photolithography from its dominant position in chip fabrication.[1,9] 1× systems, which have mask features the same size as the printed features, such as nanoimprint[10] and proximity x-ray lithography face difficulties achieving the required overlay accuracy because they lack the degrees of freedom available in projection-reduction systems to appropriately distort the image.[1] Electron- and ion-beam methods are capable of extremely high resolution, but space-charge effects, i.e. the interaction of charged particles with one another, mean that resolution is a function of beam current, severely constraining throughput at small feature sizes.[1,11,12,13] Although this can be mitigated by the use of multiple columns, this approach introduces formidable engineering challenges in terms of the simultaneous calibration and control of large numbers of independent electron/ion optical columns.[14,15,16] Optical techniques using reconfigurable masks based on micro-electromechanical systems (MEMS)[17,18] or massively parallel arrays of focused spots have been proposed.[19] While not suffering from the physical limitations to throughput that affect charged-particle systems, significant engineering challenges remain. The tradeoffs between pixel size, drive

voltage, demagnification and NA are difficult to manage for MEMS-based approaches, while data handling and spot uniformity are challenging for multi-beam systems.[20] In addition, the same diffraction effects that affect mask-based systems are present in maskless tools.

In Christensen's sense[21] the majority of these technologies are only "sustaining" since they offer no performance advantage orthogonal to the primary metrics of resolution (= minimum within-specification printable feature size), throughput, and overlay. Further, merely equaling the performance of photolithography is not enough, as there are substantial "insertion" costs associated with the introduction of any new technology into an existing manufacturing flow. Although there are reasons specific to each competing technology as to why it has not replaced photolithography, there is one generic reason why this is so: to displace photolithography, the new tool must, for the same cost, be able to exceed the performance of a standard litho tool in resolution and/or throughput and/or overlay at the current feature size, and extending for several generations of feature-size shrinks into the future. This is precisely the point made by Timothy Brunner in his 2003 paper "Why Optical Lithography Will Live Forever".[22] The evolution of photolithography is one of the principal factors underlying the continued vitality of Moore's Law[23,24] and has led to the cost per bit for both logic and memory to decrease by a factor of around 10^6 from 1980 to today.[23,24] The continuous and rapid progress of photolithography has so far made it impossible for any "disruptive" technology to displace it for chip manufacturing.[25] However, the "one size fits all" approach may be coming to an end.[9] The increasing importance of mask costs, particularly in low-volume manufacturing, has led to renewed

interest in maskless approaches – including ones based on electron-beams,[Error! Bookmark not defined.,14,17,19,20] even though their throughput is typically significantly lower than 120 to 180 wafers per hour. For high throughput, the heir apparent to 193 nm immersion lithography is Extreme Ultraviolet Lithography (EUV) which operates at a wavelength of ≈13.5 nm and, as a result of its significantly shorter wavelength, may avoid some of the mask-related problems discussed above. Although significant progress has been made over the past two decades in developing EUV for production it still has to meet throughput and mask defect targets[1,5,26] and in the end it may not prove to be an economically viable chip manufacturing technology.

The push to maintain feature density commensurate with feature size has provided an opportunity to use directed self-assembly in chip manufacturing[27,28,29,30] – a major departure from the top-down, deterministic methods employed so far. The approach with the greatest possibility of being used in chip manufacturing employs block copolymers (BCPs).[31] At the nanoscale these materials consist of two immiscible polymer chains covalently bonded together to form a single, long, linear molecule. BCPs phase separate on the nanoscale, exhibiting rich structural variations, with morphology and domain size determined by the relative volume fraction and molecular weight of the immiscible polymer chains. Highly ordered structures can be produced by templating the assembly of the BCP using topographic (graphoepitaxy) or chemical (chemoepitaxy) patterns. The crucial point is that the templating pattern can be of relatively low density with the BCP material "filling in the details". This makes it ideal for use in conjunction with photolithography processes that can produce small, precisely-placed features, but are

limited in the minimum possible pitch. Interestingly, the ability to generate large-area, nanoscale patterns with a minimum of top-down patterning is enabling other, more cost-sensitive applications, such as bit-patterned magnetic storage media.[27] Bottom-up, self-assembly methods rely on the transformation of a material, or collection of objects, from an initial disordered state to an ordered equilibrium configuration. This is not a deterministic process, and so the final configuration will contain an equilibrium distribution of defects or the process itself may be kinetically trapped and so not reach the desired ordered state. While it remains to be proven if templated BCP approaches can yield sufficiently low defect levels, early indications are promising.[32,33,34]

Semiconductor lithography is reaching a fascinating stage in its evolution. Optical lithography still reigns supreme, but disruptive technologies are starting to emerge.[35] On the one hand, maskless systems offer tremendous patterning flexibility, unconstrained by mask fabrication and its associated costs, but low throughputs are a major downside. On the other hand, directed self-assembly offers unprecedented access to low-cost, precise nanoscale patterning, but the limited set of pattern motifs is restrictive. Both of these approaches have their disadvantages, but both also provide benefits orthogonal to those of the incumbent technology and hence may find fertile ground in the existing manufacturing environment to grow into novel nanomanufacturing technologies.

*GMG is currently employed by Applied Math Solutions, LLC, Newtown, CT.

References

[1] Suzuki, K. & Smith, B. Editors, Microlithography: Science and Technology, Second Edition. CRC Press (2007).
[2] Mack C. Fundamental Principles of Optical Lithography: The Science of Microfabrication. John Wiley and Sons (2007).
[3] Wong, A. K-k. Optical Imaging in Projection Microlithography. SPIE Press (2005).

[4] Matsuyama, T., et al. The Lithographic Lens: its history and evolution. *Proc. SPIE* **6154**, 615403 (2006)
[5] International Roadmap for Semiconductors, 2009 Edition, Lithography, http://www.itrs.net/Links/2009ITRS/2009Chapters_2009Tables/2009_Litho.pdf . accessed 10/25/2011.
[6] Wuest, A., Estimation of Cost Comparison of Lithography Technologies at the 22 nm Half-pitch Node. *Proc SPIE* **7271**, 72710Y-1 (2009)
[7] Xu, P., et al. Sidewall spacer quadruple patterning for 15nm half-pitch. *Proc. SPIE* **7973**, 79731Q-1 (2011)
[8] Zimmerman, P. A. Extension Options for 193 nm Immersion Lithography. *J. Photopolym. Sci. Technol.* **22** 625-634 (2009)
[9] Preil, M. The Need for Multiple Alternatives for sub-20 nm Lithography. *Future Fab International*, Issue **38**, July 2011, www.future-fab.com,
[10] Malloy, M., Litt, L. C. Step and Flash Imprint Lithography for Semiconductor High Volume Manufacturing. J. Photopolym. Sci. Technol. **35** 749-756 (2010)
[11] Kruit, P. & Jansen, G. H. Chapter 7, Handbook of Charged Particle Optics. 2nd.Edition, J. Orloff, Ed., Taylor & Francis Group, LLC (2009).
[12]Petric, P. et. al. Reflective electron beam lithography: A maskless ebeam direct write lithography approach using the reflective electron beam lithography concept. *J. Vac. Sci. Technol. B* **28**, C6C6 (2010).
[13] McCord, M. A., et al. REBL: design progress toward 16 nm half-pitch maskless projection electron beam lithography. *Proc. SPIE* **8323** 832311-1 (2012)
[14] Yamada, A., et. al., Exposure Results with Four Column Cells in Multi Column EB Exposure System, Proc. SPIE **7488** 74881F (2009)
[15] van den Berg, C. et. al. Scanning exposures with a MAPPER multibeam system. *Proc. SPIE* **7970**, 79700D (2011).
[16] Takizawa, M., et al. Position Accuracy Evaluation of Multi Column E-beam Exposure System. *Proc. SPIE* **7970**, 79700B (2011)
[17] Lopez, D. et. al., Two-dimensional MEMS array for maskless lithography and wavefront modulation. *Proc. SPIE* **6589**, 65890S (2007)
[18] Zimmer, F., et. al., One-Megapixel Monocrystalline-Silicon Micromirror Array on CMOS Driving Electronics Manufactured With Very Large-Scale Heterogeneous Integration. *J. MEMS* **20** 564 (2011)
[19] Ljungblad, U., et. al. Phase shifted addressing using a spatial light modulator, *Microelectron. Eng.* **78-79** 398-403 (2005)
[20] Smith, H. I., et. al. Zone-plate-array lithography: A low-cost complement or competitor to scanning-electron-beam lithography. *Microelectron. Eng.* **83**, 956 (2006).
[21] Christensen, C. The Innovator's Dilemma, Harper Collins (2000).
[22] Brunner, T. A., Why optical lithography will live forever, *J. Vac. Sci. Tech.* **B21**, 2632 (2003).
[23] "Moore's Law", http://en.wikipedia.org/wiki/Moore's_law, accessed 10/25/2011.
[24] Moore, G. E. Cramming more components onto integrated circuits. *Electronics* **1965**, 38
[25] Appleyard, M.M., et. al. The innovator's non-dilemma: the case of next-generation lithography. *Manage. Decis. Econ.* **29**, 407 (2008).
[26] Bakshi., V Editor, EUV Lithography. Wiley (2008).
[27] Ruiz, R., et. al., Density Multiplication and Improved Lithography by Directed Block Copolymer Assembly. *Science* **321**, 936 (2008).
[28] Bita, I. et. al. Graphoepitaxy of Self-Assembled Block Copolymers on Two-Dimensional Periodic Patterned Templates. *Science* **321**, 939 (2008).
[29] Jeong, J. W., et. al. Highly Tunable Self-Assembled Nanostructures from a Poly(2-vinylpyridine-b-dimethylsiloxane) Block Copolymer. *Nano. Lett.* **10**, 1000 (2010).
[30] Dammel, R. R. Cost-effective Sub-20 nm Lithography: Smart Chemicals to the Rescue. *J. Photopolym. Sci. Technol.* **24** 33-42 (2011)
[31] Bates, F. S. & Glenn H. Fredrickson, G. H. Block Copolymers—Designer Soft Materials. **52**, *Phys. Today*. February, 32 (1999)
[32] Bencher, C., et. al. Directed Self-Assembly Defectivity Assessment (Part II). *Proc SPIE* **8323**, 83230N-1 (2012).
[33] Laachi,N., et. al. Self-consistent field theory of directed self-assembly in laterally confined lamellae-forming diblock copolymers. *Proc. SPIE* **8323**, 83230K-1 (2012).
[34] Gronheid, R., et al., Defect reduction and defect stability in IMEC's 14 nm half pitch chemo-epitaxy DSA flow. *Proc. SPIE* **9049**, 904905-1 (2014)

[35] Saavedra, H. M., et. al. Hybrid strategies in nanolithography, *Rep. Prog. Phys.* **73**, 036501 (2010).