Line-Edge Roughness and the Impact of Stochastic Processes on Lithography Scaling for Moore's Law

Chris A. Mack
Lithoguru.com, 1605 Watchhill Rd, Austin, TX 78703

Abstract

Moore's Law, the idea that every two years or so chips double in complexity and the cost of a transistor is always in decline, has been the foundation of the semiconductor industry for nearly 50 years. The main technical force behind Moore's Law has been lithography scaling: shrinking of lithographic features at a rate faster than the increase in finished wafer costs. With smaller feature size comes the need for better control of those sizes during manufacturing. Critical dimension and overlay control must scale in proportion to feature size, and has done so for the last 50 years. But in the sub-50-nm feature size regime, a new problem has arisen: line-edge roughness due to the stochastic nature of the lithography process. Despite significant effort, this line-edge roughness has not scaled in proportion to feature size and is thus consuming an ever larger fraction of the feature size control budget. Projection of current trends predicts a collision course between lithography scaling needs and line-edge roughness reality. In the end, stochastic uncertainty in lithography and its manifestation as line-edge roughness will prove the ultimate limiter of resolution in semiconductor manufacturing.

Subject Terms: Moore's Law, lithography, line-edge roughness, linewidth roughness, LER, LWR

1. Introduction

Moore's Law, nearing its 50th anniversary, describes the rising number of components that can be economically integrated onto a single chip.[1,2] While generally described as a doubling of the number of components on a chip every 1 – 2 years, the true formulation of Moore's Law is the growing complexity of circuits made of minimum-cost components. It is the economic element of Moore's Law that is critical: the cost of a transistor has declined significantly each year for the past 50- years. The result is that each year the same chip can be made for a much lower cost, or the same price can be paid for a much more capable chip.

While many factors have contributed to the success of Moore's Law, the dominate force driving lower cost per transistor has been that transistor density rises faster than the cost of manufacturing a finished wafer. And by far, the biggest contributor to improved transistor density has been lithography-enabled shrinking of the feature sizes of the transistor. Over the 50-year period of Moore's Law feature sizes have shrank by about a factor of 1,000, from about 25 μm to about 25 nm (Figure 1). Lithography scaling results in features that are about 0.7X in size every technology generation, each generation lasting about 2 – 3 years.[3]

A continuation of Moore's Law requires continued lithography scaling. But such scaling is endangered by two unpleasant trends. First, improvements in lithography resolution today are enabled only by multiple patterning,[4,5,6] where higher resolution comes at a considerable increase in cost. But a further complication to continued lithography scaling is the impact of stochastic effects on the ability to produce and control features of sufficient quality. Stochastic effects in the exposure, baking, and development of photoresist features leads to roughness along the edges of those features. But while feature sizes are cut in half every 5 years so, the amount of roughness along the sides of those features

have not been scaling nearly as quickly (if at all). As a result, roughness is now a large percentage of the feature size, and growing larger with each generation.

To help understand the implications of these trends, this paper will explore the physics of line-edge roughness (LER) and linewidth roughness (LWR) formation, and what might be done to improve it. The principles involved can apply not just to semiconductor manufacturing, but to other areas of nanofabrication as well.

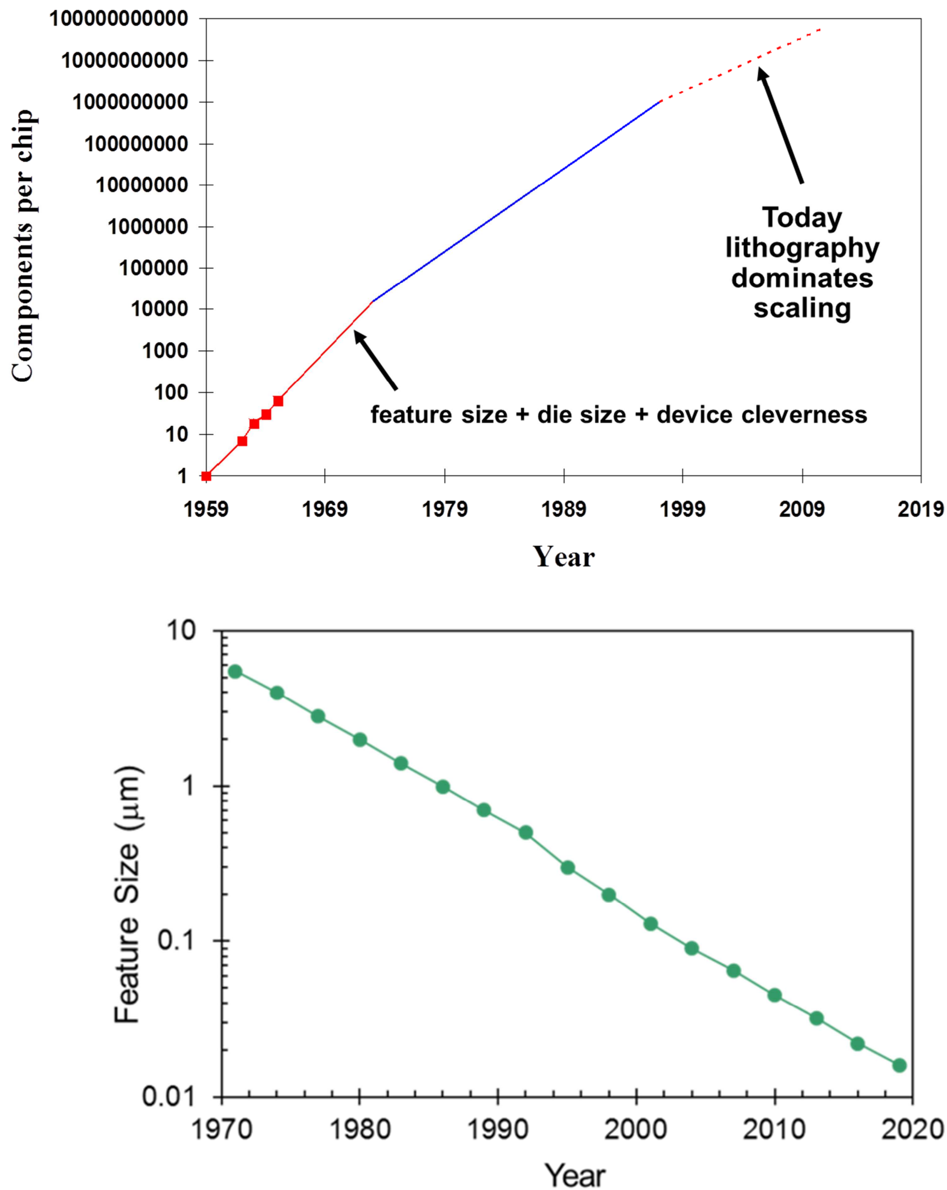


Figure 1. Moore's law as an increase in the number of components per chip, and the feature size reduction over time that has enabled it.

2. Stochastic Processes and Line-edge Roughness Formation

Randomness is everywhere present in nature. In most circumstances this randomness is hidden by the process of averaging: the visible response is the average response of many, many random events. If the number of events being averaged is sufficiently high, the stochastic nature of the process can be safely ignored and a mean field theory approach taken. In such an approach, quantities such as the number of molecules in a certain volume, the number of photons incident on a certain area, or the number of chemical species that react over a certain period of time are taken to be continuous and deterministic rather than discrete and random.

But if the number of events being averaged is small, looking at the average may fail to provide all the relevant information needed to understand the phenomenon. For example, if the volume of interest is sufficiently small, the number of photons absorbed in that volume may have an uncertainty that is too large to ignore. As lithography has scaled its feature sizes to the tens of nanometers, we have entered the regime where the randomness of the fundamental events in the lithography process cannot be ignored. The result is line-edge roughness, with a standard deviation of the edge position one the order of one or two nanometers (Figure 2). While such small amounts of roughness could easily be ignored in an era of 100 nm feature sizes and larger, today the consequences of even a single nanometer of uncertainty in the edge position of a feature is significant.

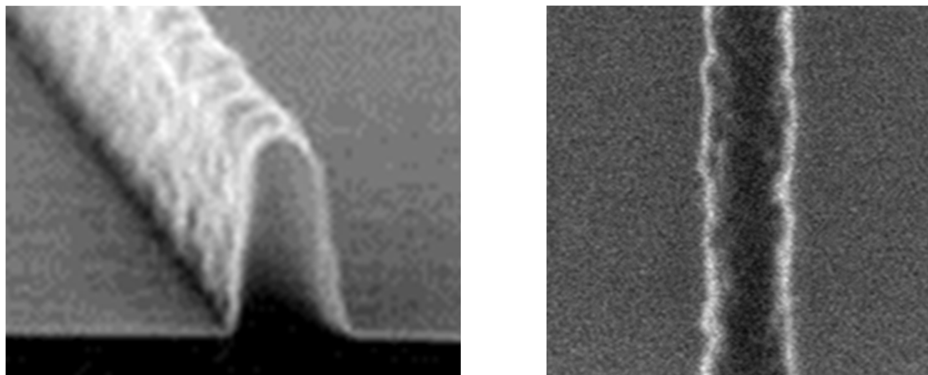


Figure 2. Examples of rough photoresist features exhibiting LER and LWR.

Some examples of processes that are fundamentally stochastic and give rise to roughness in lithographic features are:

- Photon count
- Photoacid generator positions
- Absorption
- Photoacid generation
- Polymer blocking group position
- Reaction-diffusion during post-exposure bake
- Polymer chain length
- Dissolution
- Etching

The stochastic nature of some of these steps is well understood. For others, very little is known. Some of these steps will now be described.

a) Photon and concentration shot noise, and photoacid generation

Standard photon counting statistics is Poisson and describes the uncertainty in the intensity of light. If $\langle n_{photons} \rangle$ is the mean number of photons incident on some area over some time, then the variance in the number of photons is also equal to $\langle n_{photons} \rangle$. As an example, consider the 193-nm exposure of a resist using a dose of 10 mJ/cm^2. At this wavelength, the energy of one photon is about 1.03×10^{-18} J. For an area of 1 nm X 1 nm, the mean number of photons for this dose is 97. The standard deviation of the number of photons is about 10 (the square root of the average), or about 10% of the average. For an area of 10 nm X 10 nm, the number of photons increases by a factor of 100, and the relative standard deviation decreases by a factor of 10, to about 1%. Since these are typical values for a 193-nm lithography process, we can see that shot noise contributes a noticeable amount of uncertainty as to the actual dose seen by the photoresist when looking at length scales less than about 10 nm.

Chemical concentration, the average number of molecules per unit volume, exhibits counting statistics identical to photon emission. If C is the average number of molecules per unit volume (the concentration), then the average number of molecules in a volume V will be CV, and the variance will also equal CV. The relative uncertainty in the number of molecules in a certain volume will be $1/\sqrt{CV}$. As an example, consider a typical 193-nm resist that has an initial photoacid generator (PAG) concentration of $\langle n_{0-PAG} \rangle$ = 0.042 molecules of PAG per cubic nanometer.[7] In a volume of $(10\text{ nm})^3$, the mean number of PAG molecules will be 42. The standard deviation will be 6.5 molecules, or about 15%.

Exposure of a chemically amplified resist involves absorption by the PAG followed by release of an acid. The number of absorption events that lead to the reaction of a PAG has greater uncertainty than the above photon and chemical concentration shot noises. Defining h as the concentration of acid relative to the initial concentration of unexposed PAG, the variance of this acid concentration in some volume V will be[8,9]

$$\sigma_h^2 = \frac{\langle h \rangle}{\langle n_{0-PAG} \rangle} + \frac{\left[(1-\langle h \rangle)\ln(1-\langle h \rangle)\right]^2}{\langle n_{0-PAG} \rangle \langle n_{photons} \rangle}, \qquad \langle h \rangle = e^{-C\langle E_{AVG} \rangle} \tag{1}$$

where $\langle h \rangle$ is the mean relative acid concentration resulting from exposure with an average dose of $\langle E_{AVG} \rangle$, and C is the exposure rate constant. For the case of the $(10\text{ nm})^3$ of 193-nm resist given above, the mean acid concentration will be about 0.4 and the standard deviation in acid concentration will be > 20% of the mean acid concentration. For this case, the impact of photon shot noise, absorption, and exposure [the right hand term of equation (1)] is minimal compared to variance in acid concentration caused by acid position uncertainty.

For extreme ultraviolet (EUV) resists, exposure involves a different mechanism. Photons are absorbed by the polymer rather than the PAG, leading to a photoionization event and a cascade of possibly several secondary electrons, each of which can potentially interact with a photoacid generator to create an acid. The resulting acid concentration variance will be[10]

$$\sigma_h^2 = \frac{\langle h \rangle}{\langle n_{0-PAG} \rangle} + \frac{\left[(1-\langle h \rangle)\ln(1-\langle h \rangle) \right]^2}{\langle n_{photoelectrons} \rangle} \tag{2}$$

where the mean number of generated photoelectrons is $\langle n_{photoelectrons} \rangle = \phi_e \langle n_{photons} \rangle (1-e^{-\alpha D})$, $\langle h \rangle = e^{-C\langle E_{AVG} \rangle}$, α = the absorption coefficient of a resist of thickness D, and ϕ_e is the electron generation efficiency (a number typical close to 1). For EUV resist exposure, the two terms in equation (2) are similar in magnitude, resulting in a much higher acid uncertainty than for a resist with a direct photon absorption and reaction mechanism (such as a 193-nm resist).

b) Reaction-diffusion resulting in polymer deblocking

The random processes described so far have no correlating mechanisms. The resulting noise is white for all length scales down to the molecular level. The reaction-diffusion deblocking of polymer in a chemically amplified resist does add a correlating mechanism that smooths high frequency roughness.

For a chemically amplified resist, the acid generated from exposure acts as a catalyst for a polymer deblocking reaction during a post-exposure bake in a process that is generically called a reaction-diffusion system. Acid diffuses until it comes within the reaction distance (a) of a blocked polymer site. Then, with a certain probability a reaction takes place deblocking that site (and increasing the probability that the polymer molecule will become soluble in developer). The acid is regenerated in the deblocking reaction and is free to diffuse to a new blocked site and to participate in a second deblocking reaction. Since these two deblocking reactions are caused by the same acid, they are not independent but rather are correlated.

One impact of the correlation caused by a diffusing catalyst is a smoothing of high-frequency uncertainty, that is, a smoothing of the uncertainty at length scales smaller than about the diffusion length of the acid. A rigorous treatment of this reaction-diffusion smoothing leads to an analytic expression for the power spectral density (PSD) of the time-averaged acid concentration (called the effective acid concentration, H_{eff}).[11]

$$PSD(f) = PSD(0)\left(\frac{1-e^{-(\pi \xi f)^2}}{(\pi \xi f)^2} \right)^2 \tag{3}$$

where f is the frequency, and the correlation length parameter ξ is determined by the acid diffusion length σ_D as $\xi = \sqrt{2}\sigma_D$. The zero frequency PSD in three dimensions is

$$PSD(0) = \frac{\pi^{3/2}\xi^3\sigma_{H_{eff}}^2}{4\left(2-\sqrt{2}\right)} \approx 2.37643\xi^3\sigma_{H_{eff}}^2 \tag{4}$$

Figure 3 shows a plot of this PSD.

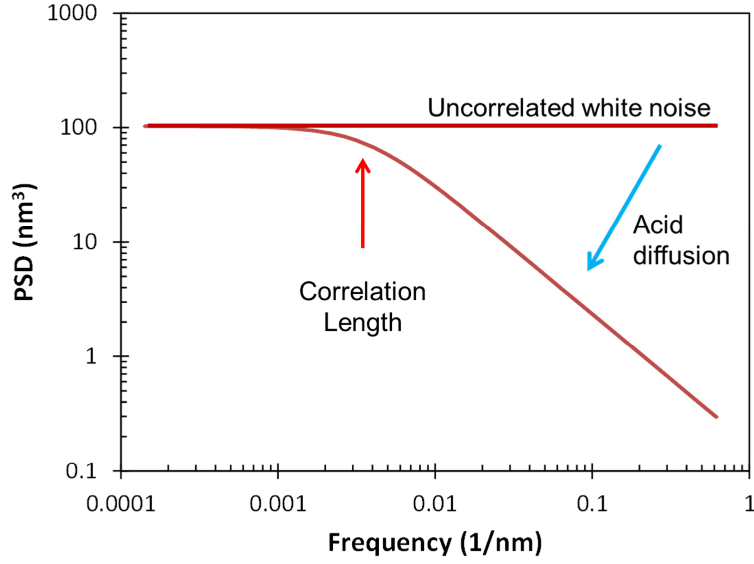


Figure 3. Illustration of how white noise, the uncertainty of the acid concentration after exposure, turns into the PSD of equation (3) through the correlating mechanism of acid catalyst diffusion.

This smoothing reduces the uncertainty in effective acid concentration compared to the acid concentration, though only at the high frequencies. The standard deviation of the effective acid concentration is approximated as[9]

$$\sigma_{h_{eff}} \approx \sqrt{2}\left(\frac{a}{\sigma_D}\right)\sigma_h \tag{5}$$

If the acid diffusion length (σ_D) is greater than the trapping distance a (the distance the acid must be to the blocked polymer site before a reaction is possible) then diffusion of the catalyst causes smoothing and correlation.

The effective acid concentration is the time average of the diffusing acid. This acid causes deblocking of blocked polymer sites. The blocked sites (with relative concentration m and a mean initial number in a given volume of $\langle n_{0-blocked}\rangle$) have a random, Poisson concentration distribution. Combining this uncertainty with the uncertainty of the deblocking reaction gives

$$\sigma_m^2 = \frac{\langle m\rangle}{\langle n_{0-blocked}\rangle} + \left(\langle m\rangle \ln\langle m\rangle\right)^2 \left(\frac{\sigma_{h_{eff}}}{\langle h_{eff}\rangle}\right)^2 \tag{6}$$

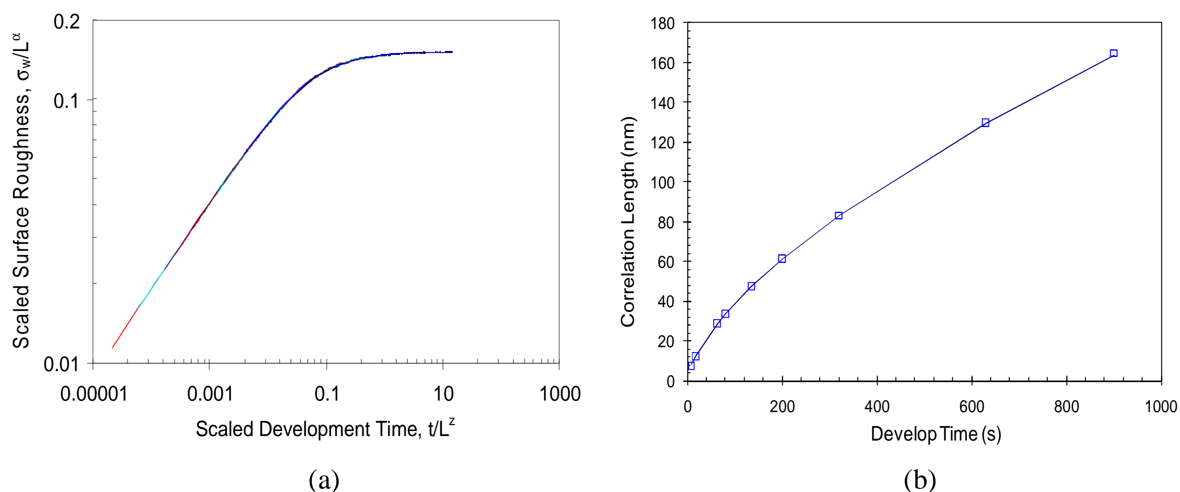
Using the example of a typical 193-nm resist, we will assume a typical density of blocked sites of 1.2/nm^3. Consider the case of $\langle h\rangle = \langle h_{eff}\rangle = 0.3$, and $\sigma_D/a = 5$. For a (10 nm)^3 volume, $\sigma_h/\langle h\rangle \approx 0.28$ and $\sigma_{h_{eff}}/\langle h_{eff}\rangle \approx 0.025$. The remaining blocked polymer will be assumed to take a typical value of $\langle m\rangle = 0.55$, giving $\sigma_m = 0.023$, or about 4.3%. For a (5 nm)^3 volume, $\sigma_m = 0.064$, or about 11%.

c) Photoresist development

Of all the process steps in a standard lithography process, development is the least understood in terms its impact on line-edge roughness.[12,13,14] The highly non-linear nature of development leads to extremely skewed distributions of probabilistic development rates. For example, if the underlying concentrations of blocked and deblocked polymer are normally distributed, the resulting development rate will approximately follow a generalized gamma distribution.[14] Additionally, the motion of a rough moving front at the resist-developer interface can be described using dynamical scaling.[12,13] Both the magnitude of the roughness and the development-induced correlation length grow with time (Figure 4).

One way in which the development process relates to the other mechanisms of roughness formation is through the volume over which the various physical and chemical processes are averaged. As the above examples have shown, averaging over a larger volume produces less uncertainty in every step, from photon absorption to polymer deblocking. In lithography the averaging volume of importance is the volume of one resist polymer molecule. Since the polymer molecule must dissolve as a unit, it is the average of its dissolution response over the entire molecule that controls its solubility.

More work is required to integrate our understanding of uncertainty and correlation in the steps leading to development with the uncertainty and correlation that derives from the development step itself.



(a) (b)

Figure 4. Simulations of open frame exposure and development of photoresist with surface roughness (σ_w) measured over an area of LXL: (a) the dynamical increase in surface roughness over time, scaled to show a universal scaling behavior, and (b) the increase in the correlation length of the surface roughness with development time. From Ref. 13.

d) Measuring roughness

Measuring line-edge or linewidth roughness involves the addition of both systematic and random errors to the measurement. The most valuable way to characterize roughness is with the power spectral density (PSD).[15] But the measurement of the PSD necessarily involves two limitations: the measurement extent is limited to a finite value L, and the sampling distance Δy is must be greater than zero. The non-infinite measurement length limits the low frequency information and results in a systematic bias called spectral leakage. The non-zero sampling distance limits the high-frequency information and results in a

systematic bias called aliasing. The natural length scale to compare to both L and Δy is the correlation length of the roughness, ξ.

Letting $PSD_d(f)$ be the discrete PSD as measured by sampling a continuous $PSD_c(f)$, we find that discrete PSD is equal to the continuous PSD modified by two error terms, ε_{alias} and $\varepsilon_{leakage}$.[16]

$$\langle PSD_d(f)\rangle = PSD_c(f)(1+\varepsilon_{alias})(1+\varepsilon_{leakage}) \qquad (7)$$

where $\varepsilon_{leakage} = \left(\frac{\xi}{L}\right)\left(\frac{(2\pi f\xi)^2-1}{(2\pi f\xi)^2+1}\right) + O\left(\frac{\xi}{L}e^{-L/\xi}\right)$, and $\varepsilon_{alias} \approx \left(\frac{\pi f \Delta y}{\sin(\pi f \Delta y)}\right)^2 - 1$.

These errors lead to very significant systematic biases in the measurement of PSD if not taken into account. Leakage can be significantly reduced through the use of data windowing, and aliasing can be at least partially compensated for by setting the sampling distance to be about twice the resolution of the measuring instrument.[16]

3. Impact of Line-edge Roughness on Linewidth Control

Line-edge and linewidth roughness are important in semiconductor manufacturing in ways that depend on the purpose of the feature being manufactured. High frequency roughness, for example, can lead to current leakage when present on the gate of a transistor, or electrical failure for a contact hole that is close to another electrical feature. But it is low-frequency roughness that is the biggest concern in semiconductor manufacturing. Low frequency roughness is indistinguishable from a linewidth error and thus contributes as one random component to the many components of linewidth uncertainty.

Consider a rectangular feature of long length L and small dimension CD. For roughness with a typical PSD having a correlation length ξ, roughness exponent H, and total roughness standard deviation (for an infinitely long line) of σ_{LWR}, the roughness-caused uncertainty in the CD (called the critical dimension uniformity, CDU) will be[17]

$$\sigma_{CDU} = \sigma_{LWR}\sqrt{\frac{(2H+1)\xi}{L}\left(1-\frac{\xi}{L}\right)} \qquad (8)$$

We want σ_{CDU} to be a small fraction of the nominal CD value. As this expression shows, the CDU is a function not only of the magnitude of the LWR (σ_{LWR}) but of the correlation length and the roughness exponent as well.

Equation (8) shows how linewidth roughness must scale to stay on track with Moore's Law. Typically, the linewidth control requirements for each new technology generation remains about constant as a percentage of the nominal feature size. A common rule of thumb is that the feature size must be controlled to within ±10% of the target feature size, and this control limit is taken to be a 3σ limit, *i.e.*, $\sigma < 0.033$*target CD. Thus, if the target CD is reduced by a factor of 2, the σ_{CDU} must be reduced by that same factor of 2, that is, σ_{CDU} should scale with the CD. In general, scaling dimensions in an integrated circuit is roughly uniform, so that L will also scale linearly with CD. Assuming that H remains about fixed for all processes, scaling the CD will require that σ_{LWR} scale linearly with CD, and that the correlation length ξ scale linearly with CD.

Scaling both the magnitude of the LWR and the correlation length of the roughness in proportion to the scaled CD has proven extremely difficult. Through significant effort LWR has been reduced somewhat over the last 10 years, but not nearly in proportion to the reductions in CD over that same time period. As a result, roughness-induced CD errors are a growing percentage of the CD and may soon come to dominate the sources of linewidth errors in semiconductor manufacturing.

4. Conclusions

For over 50 years Moore's Law has enabled continued growth in the semiconductor industry by providing ever more powerful chips at ever lower prices. Lithography scaling has been and continues to be the dominant driving force behind Moore's Law. Every five years or so the feature sizes of our devices are reduced by a factor of 2. With this reduction comes a similar reduction in the control requirements for the manufacture of those features: feature size uncertainty must scale with feature size. Throughout the 50+ years of resolution improvement in semiconductor lithography there has been a lock-step improvement in the ability to control those smaller features.

Stochastic variability is an inherent part of the lithography process. From the moment light shines on a film of photoresist, every part of lithography process has randomness that averages out to a consistent result only in the limit of large features. As feature sizes shrink, the scale over which we observe the average behavior of lithography shrinks as well. And with this shrinking scale comes the increasing importance of the randomness of the lithography process and its consequence, line-edge roughness.

Unfortunately, line-edge roughness does not scale easily with feature size. To date, line-edge roughness has been decreasing at a much slower pace than the feature sizes used to make our integrated circuits for leading edge devices. As consequence, we are fast approaching the day when roughness-induced linewidth uncertainty becomes the dominant factor in linewidth control. Ultimately, the physical limits of resolution in lithography may be caused by stochastic uncertainty and line-edge roughness. More work in formulating a comprehensive theory and model of LER formation is required before these limits can be fully understood.

References

[1] Gordon E. Moore, "Cramming More Components onto Integrated Circuits", *Electronics*, vol. 38, no. 8, pp. 114-117, Apr. 19, 1965.

[2] Gordon E. Moore, "Progress in Digital Integrated Electronics", *IEDM Technical Digest* (Washington, D.C.: 1975) pp. 11-13.

[3] Chris A. Mack, "Fifty Years of Moore's Law", *IEEE Transactions On Semiconductor Manufacturing*, **24**(2), pp. 202-207 (May, 2011).

[4] Andrew J. Hazelton, Shinji Wakamoto, Shigeru Hirukawa, et al., "Double-patterning requirements for optical lithography and prospects for optical extension without double patterning", *J. Micro/Nanolith. MEMS MOEMS*, **8**(1), 011003 (2009).

[5] Hubert Hody, Vasile Paraschiv, Emma Vecchio, et al., "Double patterning with dual hard mask for 28-nm node devices and below", *J. Micro/Nanolith. MEMS MOEMS*, **12**(4), 041306 (October 01, 2013).

[6] Frederick T. Chen, Wei-Su Chen, Ming-Jinn Tsai, and Tzu-Kun Ku, "Sidewall profile engineering for the reduction of cut exposures in self-aligned pitch division patterning", *J. Micro/Nanolith. MEMS MOEMS,* **13**(1), 011008 (March 25, 2014).

[7] Chris A. Mack, Fundamental Principles of Optical Lithography: The Science of Microfabrication, John Wiley & Sons (London: 2007), p. 240.

[8] *Ibid*., p. 245.

[9] Chris A. Mack, "Line-Edge Roughness and the Ultimate Limits of Lithography", *Advances in Resist Materials and Processing Technology XXVII*, SPIE Vol. 7639, p. 763931 (2010).
[10] Chris A. Mack, James W. Thackeray, John J. Biafore, and Mark D. Smith, "Stochastic Exposure Kinetics of EUV Photoresists: A Simulation Study", *Journal of Micro/Nanolithography, MEMS, and MOEMS*, **10**(3), 0330193 (Jul-Sep, 2011).
[11] Chris A. Mack, "Reaction-diffusion power spectral density" *Journal of Micro/Nanolithography, MEMS, and MOEMS*, **11**(4), 043007 (Oct-Dec, 2012).
[12] Chris A. Mack, "Stochastic Modeling in Lithography: The Use of Dynamical Scaling in Photoresist Development", *Journal of Micro/Nanolithography, MEMS, and MOEMS*, **8**(3), 033001 (Jul-Sep, 2009).
[13] Chris A. Mack, "Stochastic modeling of photoresist development in two and three dimensions", *Journal of Micro/Nanolithography, MEMS, and MOEMS*, **9**(4), 041202 (Oct-Dec, 2010).
[14] Chris A. Mack, "Defining and measuring development rates for a stochastic resist: a simulation study", *Journal of Micro/Nanolithography, MEMS, and MOEMS*, **12**(3), 033006 (Jul-Sep, 2013).
[15] Chris A. Mack, "Analytic form for the power spectral density in one, two, and three dimensions", *Journal of Micro/Nanolithography, MEMS, and MOEMS*, 10(4), 040501 (Oct-Dec, 2011).
[16] Chris A. Mack, "Systematic Errors in the Measurement of Power Spectral Density", *Journal of Micro/Nanolithography, MEMS, and MOEMS*, **12**(3), 033016 (Jul-Sep, 2013).
[17] Chris A. Mack, "Analytical Expression for Impact of Linewidth Roughness on Critical Dimension Uniformity", *Journal of Micro/Nanolithography, MEMS, and MOEMS*, **13**(2), 020501 (Apr-Jun, 2014).