

A Simple Model of Line-Edge Roughness

Chris A. Mack
www.lithoguru.com

Abstract

A stochastic modeling approach is used to predict the results of the exposure and post-exposure bake of a chemically amplified photoresist. The statistics of photon shot noise, chemical concentration, exposure, reaction diffusion and amplification are derived. The result, though preliminary, is a prediction of the standard deviation of the final deprotection level of polymer molecules in the resist using simple analytical expressions. Combining this result with ongoing work to characterize the stochasticity of resist development will eventually lead to a full model of the line-edge roughness of a resist feature. The current model is used to elucidate the impact of resist properties and processing on line-edge roughness. The optimization of acid diffusion length and the exposure dose/thermal dose trade-off leads to a minimum LER. Optimization of resist molecular size should be possible once the development step is fully modeled.

Line-edge roughness (LER), and the associated line-width roughness (LWR), are likely the ultimate limiters of resolution in optical lithography for semiconductor

manufacturing. As feature sizes have scaled smaller to the cadence of Moore's Law, the accompanying roughness of the edges of these features has stubbornly refused to scale. Thus, this roughness has become a more prominent feature of lithography patterning with each new technology generation. Current devices are being manufactured with resist LWR (measured as the 3σ line-width variation along a single line) that approaches or exceeds 10 percent of the feature size. Further scaling of feature size will only make a bad problem worse. Significant effort is now being expended by the industry to address LER. One essential component of such efforts is the development of predictive models for line-edge roughness.

Most theoretical descriptions of lithography make an extremely fundamental and mostly unstated assumption about the physical world being described: the so-called *continuum approximation*. Even though light energy is quantized into photons, and chemical concentrations are quantized into spatially distributed molecules, the descriptions of aerial images and latent images in standard lithography simulators ignore the discrete nature of

these fundamental units and instead use continuous mathematical functions. When describing lithographic behavior at the nanometer level, an alternate approach, and in a very real sense a more fundamental approach, is to build the quantization of light as photons and matter as atoms and molecules directly into the models used. Such an approach is called *stochastic modeling*, and involves the use of random variables and probability density functions to describe the statistical fluctuations that are expected. Of course, such a probabilistic description will not make deterministic predictions – instead, quantities of interest will be described by their probability distributions, which in turn are characterized by their moments, such as the mean and variance. Thus, feature size, for example, is described by a mean (called the critical dimension, CD) and a standard deviation (the LWR).

One common approach to studying LER formation is through the use of Monte Carlo simulations[1-3] and mesoscale modeling.[4] These approaches can be extremely valuable since they can be made rigorous at the length scale of interest and can be used to test the impact of various fundamental stochastic mechanisms that may be at work. The drawback to Monte Carlo approaches, however, is their lengthy execution times resulting from the need to run each stochastic step a large number of times to provide proper statistical results. Often important physical insights can remain undiscovered beneath the mountains of statistical data that a Monte Carlo simulator can generate.

While Monte Carlo methods can be extremely useful, there is also a need for the development of simple, analytical expressions that capture the essence of the

LER formation mechanisms. By formulating the equations describing the fundamental processes and kinetics of exposure, baking and development as stochastic equations, one might hope for a solution to these stochastic equations that mimic the mean-field solutions that are used in physical lithography simulators today. Alas, attempts at such a formulation are certain to be disappointing as the fundamental stochastic equations remain immensely complicated.[5] One approach, then, is to look for solutions that provide, rather than the full stochastic nature of each intermediate variable, an approximation to the variance of each term. Thus, while the mean-field theory of the continuum models gives the mean of the distribution for each variable in a tractable mathematical form, the goal here is to find similar tractable expressions for the variance of each term. This paper provides a progress report on this author's as-yet incomplete effort.

Much of the treatment given below follows that provided in Ref. 6, with more recent advances included.[7] A model is presented that includes the statistics of photon shot noise and chemical concentrations combined with probabilities of absorption and exposure to give the variance of the acid concentration after exposure. During post-exposure bake, acid diffusion and reaction is first formulated to give the effective acid concentration and its variance, followed by the level of polymer deprotection and its variance. The stochastics of photoresist development are touched upon next, but are not included in the model used here due to their lack of maturity. Finally, pulling the results together, an attempt at a comprehensive line-edge roughness model is provided, though many deficiencies remain.

1. Photon and Chemical Concentration Shot Noise

Consider a light source that randomly emits photons at some average rate into some area A . Assume further that each emission event is independent. Over some time T , the probability that exactly n photons will be emitted in this time period is given by a Poisson distribution. From this distribution, the mean and standard deviation of n , as well as the resulting intensity of light I , can be calculated:

$$\langle I \rangle = \frac{\langle n_{\text{photons}} \rangle}{TA} \left(\frac{hc}{\lambda} \right), \quad \sigma_I = \frac{1}{TA} \left(\frac{hc}{\lambda} \right) \sigma_n = \frac{\langle I \rangle}{\sqrt{\langle n_{\text{photons}} \rangle}} \quad (1)$$

where h is Planck's constant, c is the vacuum speed of light, and λ is the vacuum wavelength.

As this equation shows, the uncertainty of getting the mean or expected intensity grows as the number of photons is reduced, a phenomenon known as *shot noise*. As an example, consider a 193nm exposure of a resist with a dose-to-clear of 10 mJ/cm². At the resist edge, the mean exposure energy ($=\langle I \rangle T$) will be on the order of the dose-to-clear. At this wavelength, the energy of one photon, hc/λ , is about 1.03×10^{-18} J. For an area of 1nm X 1nm, the mean number of photons during the exposure, from equation (1), is about 97. The standard deviation is about 10, or about 10 percent of the average. For an area of 10nm X 10nm, the number of photons increases by a factor of 100, and the relative standard deviation decreases by a factor of 10, to about 1 percent. Since these are typical values for a 193nm 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 10nm.

For extreme ultraviolet lithography (EUVL), the situation is considerably worse. At a wavelength of 13.5nm, the energy of one photon will be 1.47×10^{-17} J, about 15 times greater than at 193nm. Also, the goal is to have EUV resist sensitivity that is 2-4 times better than 193nm resists (though it is unclear whether this goal will be achieved). Thus, the number of photons will be 30-60 times less for EUV than 193nm lithography. A 1nm X 1nm area will see only two to three photons, and a 100nm² area will see on the order of 200 photons, with a standard deviation of 7 percent.

Chemical concentration, the average number of molecules per unit volume, exhibits counting statistics identical to photon emission (for reasonably low concentrations). Let C be the average number of molecules per unit volume. For some volume V , the probability of finding exactly n molecules in that volume will be given by a Poisson distribution with an average number of molecules in the volume equal to CV , and the variance also equal to CV . The relative uncertainty in the number of molecules in a certain volume will be

$$\frac{\sigma_n}{\langle n \rangle} = \frac{1}{\sqrt{\langle n \rangle}} = \frac{1}{\sqrt{CV}} \quad (2)$$

As an example, consider a 193nm resist that has an initial PAG concentration of 3 percent by weight, or a concentration of about 0.07 mole/liter (corresponding to a resist density of 1.2 g/ml and a PAG molecular weight of 500 g/mole). Converting from moles to molecules with Avogadro's number, this corresponds to 0.042 molecules of PAG per cubic nanometer. In a volume of (10nm)³, the mean number of PAG molecules will be 42. The standard deviation will be 6.5 molecules, or about 15 percent. For 248nm resists, the PAG loading is

typically 3 times higher or more, so that closer to 150 PAG molecules might be found in a $(10\text{nm})^3$ volume, for a standard deviation of 8 percent. Note that when the mean number of molecules in a given volume exceeds about 20, the Poisson distribution can be well approximated with a Gaussian distribution.

2. Photon Absorption and Exposure

What is the probability that a photon will be absorbed by a molecule of light-sensitive material in the resist? Further, what is the probability that a molecule of sensitizer will react to form an acid? As discussed above, there will be a statistical uncertainty in the number of photons in a given region of resist, a statistical uncertainty in the number of PAG molecules, and additionally a new statistical uncertainty in the absorption and exposure event itself.

Defining h as the concentration of acid relative to the initial concentration of unexposed PAG, and leaving the details of the derivation to Ref. 6, the standard deviation of this acid concentration in some volume V will be

$$\sigma_h^2 = \frac{\langle h \rangle}{\langle n_{0-PAG} \rangle} + \frac{[(1-\langle h \rangle)\ln(1-\langle h \rangle)]^2}{\langle n_{photons} \rangle} \quad (3)$$

where $\langle n_{0-PAG} \rangle$ is the mean number of PAGs initially found in that volume and $\langle h \rangle$ is the mean acid concentration resulting from exposure. This result is reasonably intuitive. The first term on the right-hand side of equation (3) is the expected Poisson result based on exposure kinetics – the relative uncertainty in the resulting acid concentration after exposure goes as one over the square root of the mean number of acid molecules generated with-

in the volume of interest. For large volumes and reasonably large exposure doses, the number of acid molecules generated is large and the statistical uncertainty in the acid concentration becomes small. For small volumes or low doses, a small number of photogenerated acid molecules results in a large uncertainty in the actual number within that volume. The second term accounts for photon shot noise and adds to the variance due to chemical concentration shot noise. For the case of the $(10\text{nm})^3$ of 193nm resist given above, the standard deviation in initial acid concentration near the resist edge (where the mean acid concentration will be about 0.4) will be > 20 percent of the acid concentration. For 193nm resists, the impact of photon shot noise is minimal compared to variance in acid concentration caused by simple molecular position uncertainty.

For EUV resists, exposure entails an extra mechanism. Absorption of a photon leads to ionization and the release of possibly several secondary electrons, each of which can potentially be captured by a photoacid generator to create an acid. This mechanism will not be treated here but has been investigated by others.[8]

3. Acid-Catalyzed Reaction Diffusion

In this section we'll consider reaction diffusion and the polymer deblocking reaction. While the details of the derivation of the results presented below will be left out (see Refs. 6 and 7), the approach makes use of the concept of the von Smoluchowski trap. When an acid approaches a blocked (protected) site on the polymer within a distance given by the capture radius a , there is some probability that a reaction will take place. Thus, the stochastics of diffusion (a random walk

through the three-dimensional resist matrix, with a diffusion length σ_D) can be coupled with the probability of capture and reaction through the von Smoluchowski trap radius. Since the probability that any given blocked site will react is proportional to the time average of the acid concentration seen by that site (that is, passing within a distance a of that site) during the bake cycle, we can define this time-average acid concentration as h_{eff} and derive the uncertainty of this quantity.

$$\sigma_{h_{eff}}^2 = \left(\frac{\sqrt{2}a}{\sigma_D} \right)^2 \sigma_h^2 \quad (4)$$

This result is extremely interesting. Because the reaction depends on a diffusing catalyst, the uncertainty in the effective (time-average) acid concentration is reduced whenever the diffusion length is greater than the capture distance of the deblocking reaction.

The effective acid concentration can be used to calculate the amount of deblocking that occurs during PEB. Letting m be the relative concentration of blocked polymer sites,

$$m = e^{-\alpha_f h_{eff}}, \quad \langle h_{eff} \rangle \approx \langle h \rangle$$

$$\sigma_m^2 = \frac{\langle m \rangle}{\langle n_{0-blocked} \rangle} + \left(\frac{\langle m \rangle \ln \langle m \rangle}{\langle h \rangle} \right)^2 \left(\frac{\sqrt{2}a}{\sigma_D} \right)^2 \left(\frac{\langle h \rangle}{\langle n_{0-PAG} \rangle} + \frac{[(1-\langle h \rangle) \ln(1-\langle h \rangle)]^2}{\langle n_{photom} \rangle} \right) \quad (5)$$

where $n_{0-blocked}$ is the initial number of blocked polymer sites found in the given volume V . Or, in a slightly different form,

$$\left(\frac{\sigma_m}{\langle m \rangle} \right)^2 = \frac{1}{\langle n_{0-blocked} \rangle \langle m \rangle} + (K_{amp} t_{PEB})^2 \left(\frac{\sqrt{2}a}{\sigma_D} \right)^2 \left(\frac{\langle h \rangle}{\langle n_{0-PAG} \rangle} + \frac{[(1-\langle h \rangle) \ln(1-\langle h \rangle)]^2}{\langle n_{photom} \rangle} \right) \quad (6)$$

where K_{amp} is the amplification rate constant and t_{PEB} is the post-exposure bake time (the amplification factor is $\alpha_f = K_{amp} t_{PEB}$).

While the above equations show how

fundamental parameters affect the resulting variance in the final blocked polymer concentration, interpretation is somewhat complicated by the fact that these parameters are not always independent. In particular, the Byers-Petersen model shows a relationship between $K_{amp} t_{PEB}$ and $\sigma_D a$, a topic that will be further discussed below.

Consider the example of a typical 193nm resist, with initially 1.2 blocked groups/nm³, 0.042 PAGs/nm³, $K_{amp} t_{PEB} = 2$, $\langle h \rangle = \langle h_{eff} \rangle = 0.3$, and $\sigma_D/a = 5$. For a (10nm)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 have $\langle m \rangle = 0.55$ and $\sigma_m = 0.023$, or about 4.3 percent. For a (5nm)³ volume, $\sigma_m = 0.064$, or about 11 percent.

The exposure dose dependence of σ_m can be deduced using equation (5). As the exposure dose goes to zero, $\langle h \rangle$ goes to 0 and $\langle m \rangle$ goes to 1. For an infinite dose, $\langle h \rangle$ goes to 1 and $\langle m \rangle$ goes to 0. The resulting values for σ_m are

$$\sigma_m^2(dose=0) = \frac{1}{\langle n_{0-blocked} \rangle}$$

$$\sigma_m^2(dose=\infty) = \frac{e^{-\alpha_f}}{\langle n_{0-blocked} \rangle} + (\alpha_f e^{-\alpha_f})^2 \left(\frac{\sqrt{2}a}{\sigma_D} \right)^2 \frac{1}{\langle n_{0-PAG} \rangle} \quad (7)$$

The variation of σ_m with dose is shown in Figure 1.

Because a single acid molecule diffuses and potentially causes many reactions, these reactions will be stochastically correlated.[9] If the diffusion of the acid catalyst is the only mechanism by which the concentration m becomes spatially correlated, the autocorrelation of the reaction-diffusion point spread function (RDPSF) will define this spatial correlation. Consider first the (non-normalized) autocorrelation of the effective acid concentration. Assuming that the initial distribution of

the catalyst is stochastically uncorrelated,

$$R_{H_{\text{eff}}} = \sigma_H^2 (RDPSF \otimes RDPSF) \quad (8)$$

It will be useful to normalize the auto-correlation function to be one at the origin. For the 1D case,

$$\tilde{R}_{H_{\text{eff}}}(\tau) = \frac{\int_{-\infty}^{\infty} RDPSF(x)RDPSF(x+\tau)dx}{\int_{-\infty}^{\infty} [RDPSF(x)]^2 dx} \quad (9)$$

Analytical evaluation of equation (9) for the 1D, 2D and 3D cases does not seem possible, so numerical integrations were performed and the results fit to a standard exponential correlation function:

$$\tilde{R}_{H_{\text{eff}}}(\tau) = e^{-(|\tau|/\zeta)^{2\alpha}} \quad (10)$$

where ζ is the correlation length and α is the Hurst (roughness) exponent. Fitting the numerical evaluation of equation (9) to the empirical function (10) produces extremely good fits. For the important 3D case, $\alpha \approx 0.9$ and the correlation length is equal to $1.52\sigma_D$.

4. Acid-Base Quenching

The acid-base neutralization reaction due to the presence of quencher may pose the greatest challenge to stochastic modeling of the sort being derived here. While acid concentrations in chemically amplified resists are low, base quencher concentrations are even lower, leading to greater statistical uncertainty in concentration for small volumes. Further, since the reaction is

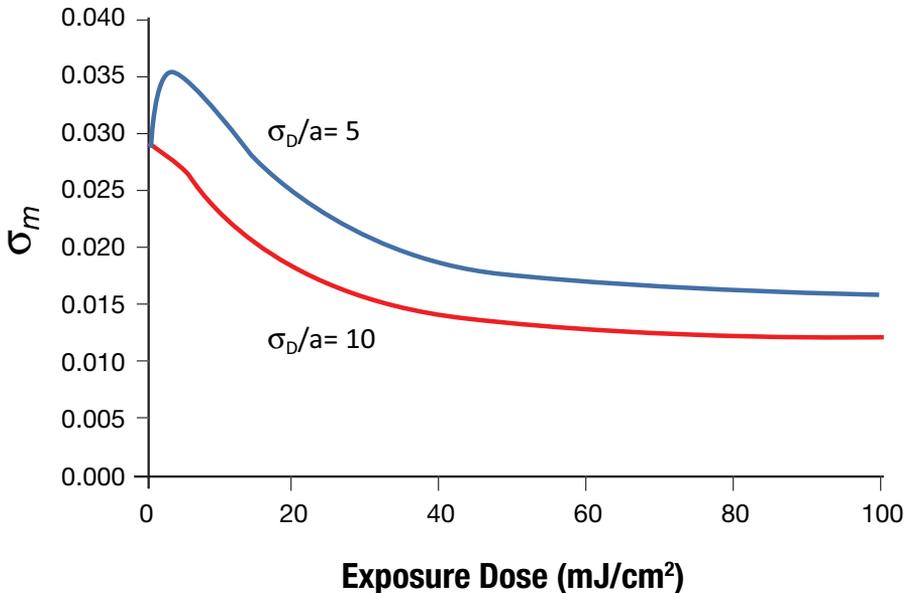


Figure 1. The Dose Dependence of σ_m Using Equation (5) and the Parameters Given in the Text, an Exposure Rate Constant of $0.05 \text{ cm}^2/\text{mJ}$, and a $(10\text{nm})^3$ Volume

one of annihilation, statistical variations in acid and base concentrations can lead effectively to acid-base segregation, with clumps of all acid or all base.[10,11] Such clumping is likely to lead to low-frequency line-edge roughness. The presence of quencher, however, also leads to dramatic improvements in the gradient of acid which, as will become clear below, leads to improvement in the final line-edge roughness.[12] Much further work is needed to study and model this phenomenon. Thus, while acid-base quenching is extremely important in its impact on LER, it will not be considered in the model presented here.

5. Development

The surface-limited reaction of a partially deprotected polymer with developer can be treated in a stochastic nature.[13,14] However, dissolution rate couples with the path of dissolution to produce the final photoresist edge, so that the stochastic nature of this dissolution path must also be taken into account. One approach to studying the stochastic nature of photoresist dissolution involves the characterization of scaling relationships as a means for elucidating fundamental mechanisms.[15,16] In this paper, however, the development step will be essentially ignored, except for one aspect. Since a polymer molecule must either dissolve or remain behind, the size of the polymer will determine the volume V over which the uncertainty in deblocking concentration plays out. If a typical 193nm resist polymer has a radius of gyration of 5nm, the volume is equivalent to a cube of about 8nm on a side. This then becomes the volume needed to determine $\langle n_{photons} \rangle$, $\langle n_{0-PAG} \rangle$, and $\langle n_{0-blocked} \rangle$.

Note that larger resist polymer molecules actually reduce line-edge roughness by increasing the volume of resist over

which the stochastic reactions are averaged. Since, however, the full development step is not yet accounted for in this model, the impact of molecular size on development itself is not included. It seems obvious that if larger molecules are removed or left behind on the resist sidewall, the result will be greater roughness. Thus, resist molecular size will provide two competing mechanisms for sidewall roughness: Larger molecules provide greater stochastic averaging of the events leading up to that molecule's solubility (and thus reducing the uncertainty in its solubility), but also produce larger "pixels" of roughness as those molecules dissolve or remain behind. Thus, it is clear that there will be an optimum molecular size that minimizes the overall sidewall roughness of the resist feature. Determining what this optimum molecular size is will have to wait for more detailed modeling of the stochastic development process.

6. Line-Edge Roughness: An Overall Model

In the sections above, a stochastic model for exposure and reaction-diffusion of chemically amplified resists was developed. This stochastic model will now prove useful for the prediction of certain line-edge roughness trends. While development should also be included, for the sake of simplicity we will assume an infinite contrast development process so that the line edge will be determined by the blocked polymer latent image. Thus, a simple threshold model for the latent image will determine the resist critical dimension. A Taylor series expansion of the blocked polymer concentration as a function of position, cut off after the linear term, allows us to predict how a small change in blocked polymer concentration (Δm) will result in a change in edge position (Δx):

$$\Delta x = \frac{\Delta m}{dm/dx} \quad (11)$$

From this, we can devise a simple quantitative model for line-edge roughness. This equation, however, predicts that an image that produces an infinite gradient of blocked polymer sites will have zero LER. In reality, the non-zero molecular size being dissolved will produce non-zero roughness, even if an infinite gradient of blocked sites could be obtained. Thus, the standard measure of line-edge roughness, from a top-down SEM, will be proportional to the standard deviation of blocked polymer concentration divided by its gradient perpendicular to the line edge plus a term to account for polymer size (σ_0):

$$\sigma_{LER} = \frac{\sigma_m}{dm/dx} + \sigma_0 \quad (12)$$

Michaelson[12] plotted measured LER versus calculated values of dm/dx and found that many different resists followed an almost universal curve, which I have fit to this model:

$$LER = \frac{33}{dm/dx} + 5 \quad (13)$$

where LER is the 3σ value, in nanometers, and dm/dx is in units of $1/\mu\text{m}$. This fit corresponds to $\sigma_m = 0.011$, which is on the order of the values shown in Figure 1, and a value of σ_0 consistent with typical resist polymer sizes.

To achieve a low LER it will be necessary to make the standard deviation of the deprotection small and make the gradient of deprotection large. A main topic of Chapter 9 of Ref. 6 is how process parameters can be used to maximize the latent image gradient given an aerial image log-slope:

$$\frac{\partial m}{\partial x} \approx \frac{1}{\eta} e^{-\alpha_f h_{eff}} (1 - e^{-\eta \alpha_f}) (1 - h_{eff}) \ln(1 - h_{eff}) \frac{\partial \ln I}{\partial x} \quad (14)$$

where

$$\eta = \frac{\pi^2 \sigma_D^2}{2L^2 K_{amp} t_{PEB}} = \frac{\pi^2 D}{L^2 K_{amp}}$$

The term η represents the ratio of the rate of diffusion for a feature of size L to the rate of deblocking reaction. Comparing equation (5) to equation (14), there is one interesting variable in common to both: acid diffusion. Increasing acid diffusion will reduce σ_m , but will also reduce the latent image gradient. One would expect, then, an optimum level of diffusion to minimize the LER.

To investigate the impact of diffusion on LER, we can combine equations (5) and (14) into (12). Thus, for the no-quencher case, and ignoring photon shot noise to make the equations a bit simpler,

$$\sigma_m = \sqrt{\frac{\langle m \rangle}{\langle n_0 - \text{blocked} \rangle} + \left(\langle m \rangle \ln \langle m \rangle \right)^2 \left(\frac{\sqrt{2}a}{\sigma_D} \right)^2 \left(\frac{1}{\langle n_0 - \text{PAG} \rangle \langle h \rangle} \right)} \quad (15)$$

$$\frac{\partial m}{\partial x} \propto \frac{1}{\sigma_D^2} \left(1 - e^{-\pi^2 \sigma_D^2 / 2L^2} \right)$$

so that

$$LER \propto \frac{\sigma_D^2}{1 - e^{-\pi^2 \sigma_D^2 / 2L^2}} \sqrt{1 - (K_{amp} t_{PEB}) \langle m \rangle \ln \langle m \rangle \left(\frac{\sqrt{2}a}{\sigma_D} \right)^2 \frac{\langle n_0 - \text{block} \rangle}{\langle n_0 - \text{PAG} \rangle}} \quad (16)$$

Figure 2 shows the trend of LER versus acid diffusion for a 45nm feature for four different values of the deprotection capture range a , 0.5, 1.0, 2.0 and 3.0nm. In each case, there is a diffusion length that minimizes the LER. Below the optimum diffusion length, LER is limited by σ_m so that increasing the diffusion will improve LER. Above the optimum diffusion length

the LER is gradient limited, so that increases in diffusion further degrade the gradient and worsen the LER. This optimum diffusion length is given approximately by

$$\sigma_D^2 \approx \sqrt{2K} \left(\frac{aL}{\pi} \right) - \frac{Ka^2}{2} \quad (17)$$

where

$$K = -2(K_{amp} t_{PEB}) \langle m \rangle \ln \langle m \rangle \frac{\langle n_{0-block} \rangle}{\langle n_{0-PAG} \rangle}$$

Note that the optimum diffusion length is constrained by the feature size at one end and the deblocking reaction capture range at the other:

$$a \ll \sigma_D \ll L \quad (18)$$

As L decreases, there becomes less room for the diffusion length to fit within these constraints.

Unless, of course, a is allowed to decrease as well. This capture range for the deblocking reaction is not an easy parameter for the resist chemist to manipulate, but it can be adjusted. There is a consequence, however. The rate of the deblocking reaction is a strong function of this capture range. In fact, assuming that the amplification reaction is in the diffusion-limited regime, the amount of amplifi-

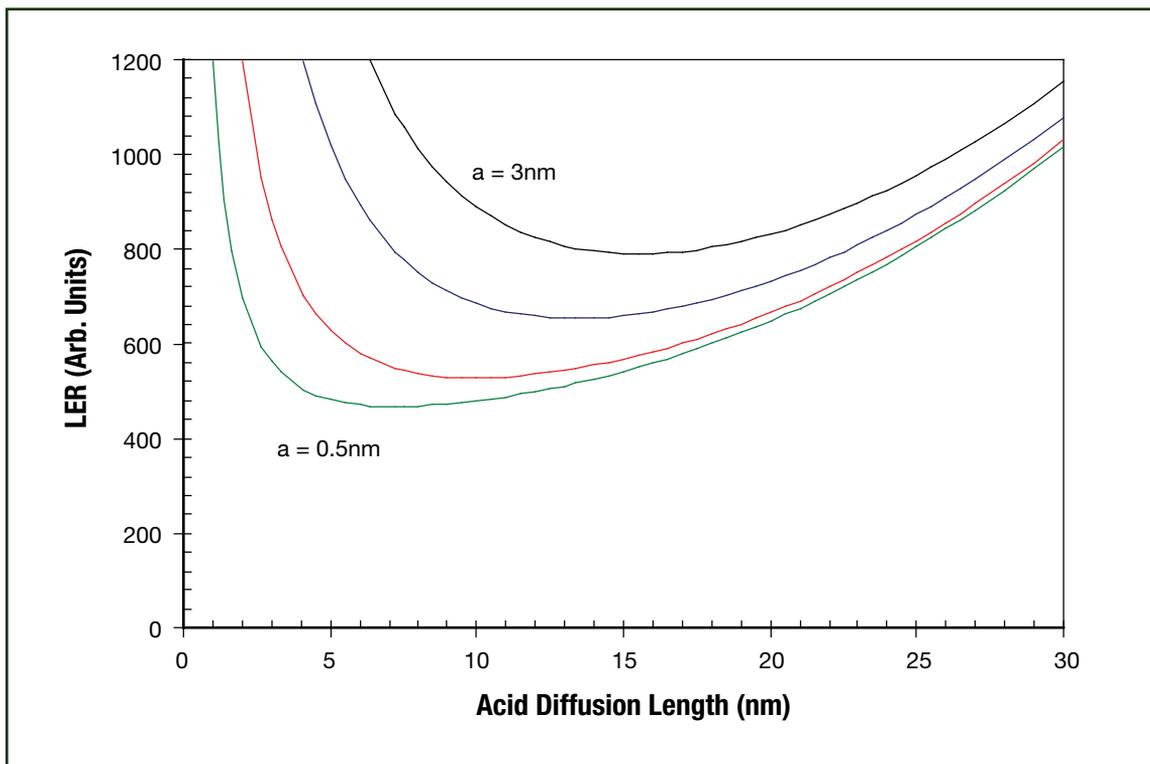


Figure 2 Prediction of LER Trends for a 45nm Feature for Four Values of the Deblocking Reaction Capture Range a (0.5, 1, 2, and 3nm)

cation will be controlled by the amplification factor α_f which is proportional to both a and the diffusion length squared:

$$\alpha_f = K_{amp} t_{PEB} = 2\pi\sigma_D^2 a G_0 N_A \quad (19)$$

where G_0 is the initial PAG concentration and N_A is Avogadro's number. To keep line-edge roughness small for smaller features, both the diffusion length and the reaction capture range should be lowered in proportion to L . But this means that the amplification factor will decrease as L^3 . Lower amplification factor will require increased exposure dose to cause the same amount of amplification, meaning

that dose would have to rise dramatically to keep LER low in the presence of shrinking feature sizes. There is one other term, however, that can slow this unfortunate scaling relationship. By increasing the PAG loading G_0 , the amplification factor can be kept higher while diffusion and capture range are decreased. There are very real, practical limits to PAG loading, however, and it is doubtful that this lever will provide sufficient long-term relief. It seems that the fundamental stochastic nature of resist chemistry creates a need for much higher exposure dose to keep small features from being dominated by LER.

The LER model presented here can also be used to investigate the impact of expo-

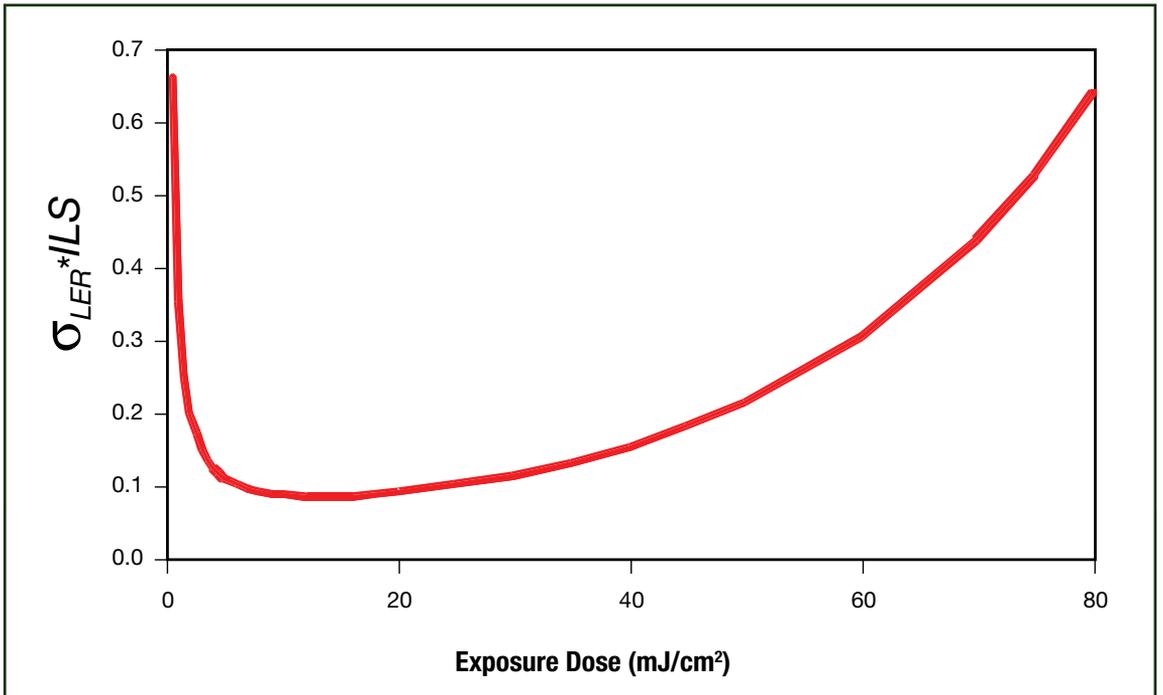


Figure 3. Prediction of LER (1σ , relative to the image log-slope, ILS), as a function of dose (all other parameters held constant at the values used for Figure 1, and with $\sigma_D = 5\text{nm}$, $a = 0.5\text{nm}$, and $\sigma_0 = 0$) for a 30nm feature. It is important to note that no attempt was made to adjust the process to keep the resultant lithographic feature size constant.

sure dose and the exposure dose/thermal dose trade-off on LER. Figure 3 shows the impact of dose on LER, keeping all other parameters of the process constant and letting $\sigma_0 = 0$. Since the entire process is held constant except exposure dose, this changing LER will be accompanied by a large change in the resulting feature dimension as well. Figure 4 shows a more interesting case where both the exposure dose and the thermal dose (PEB conditions) were adjusted to keep the level of deprotection constant ($\langle m \rangle = 0.5$), and thus the resultant feature width constant. The results show that there is an optimum trade-off between exposure dose and thermal dose to keep the LER at a minimum. Higher doses will not always result in lower LER. Note, however, that the acid diffusion length was kept constant in these calculations. In reality, changing the thermal dose will most certainly result in a change in acid diffusion length as well.

7. Conclusions

In this paper, an attempt has been made to develop a comprehensive stochastic model for LER based on deriving approximate expressions for the variance and correlations that occur at each step in the lithography process. While some progress has been made, the resulting model is far from complete.

The work begins with photon shot noise. Speckle has not been discussed here, though recent studies have made very good progress in understanding this phenomenon for 193nm lithography.[17,18] Along with chemical concentration shot noise, the result is a Poisson distribution. Combining these distributions with the probability of absorption and exposure gives a nearly Poisson acid shot-noise distribution. Reaction diffusion provides an incredibly interesting and important result: Diffusion of the reaction catalyst means that the uncertainty in the effective acid

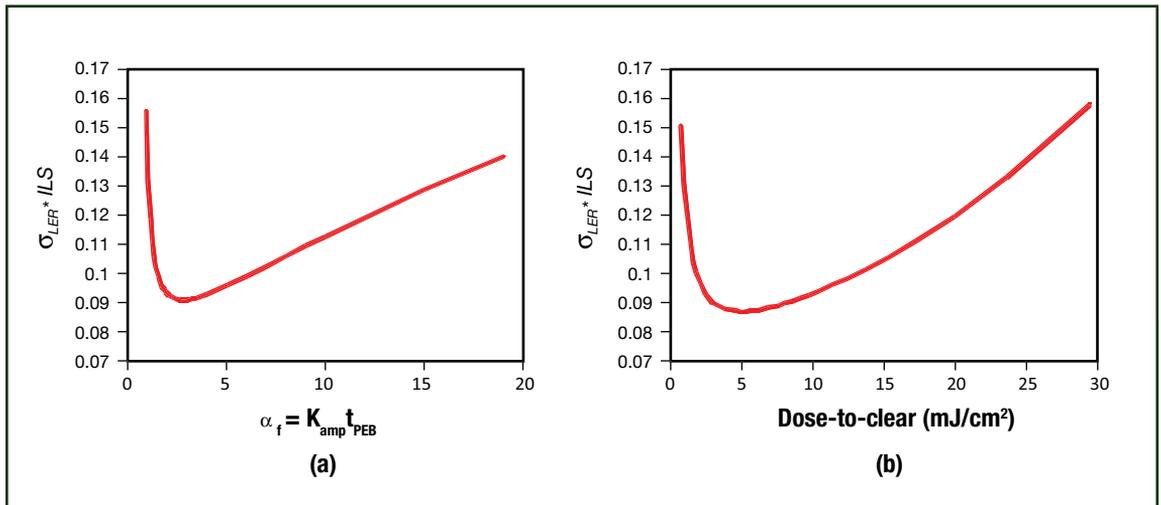


Figure 4. Prediction of LER (1σ , relative to the image log-slope, ILS), as a function of exposure and thermal dose where the level of deprotection was kept constant at 0.5 (all other parameters held constant at the values used for Figure 1, with $\sigma_D = 5\text{nm}$, $a = 0.5\text{nm}$ and $\sigma_0 = 0$) for a 30nm feature. This is equivalent to adjusting the process to keep the resulting lithographic feature size constant.

concentration is reduced whenever the acid diffusion length is greater than the von Smoluchowski trap radius. Thus, acid (catalyst) diffusion reduces stochastic uncertainty in the effective acid concentration. Since, however, increased acid diffusion also degrades the acid gradient, there is an optimum diffusion length for minimizing LER. The trade-off between exposure dose and thermal dose affects the optimum latent image gradient, and thus contributes to LER as well. For any given resist and feature size, there will be an optimum balance of exposure dose and thermal dose (PEB conditions) to minimize LER.

Development is likely to be a very significant generator of roughness. Unfortunately, our current understanding of how development dynamically roughens a surface is insufficient to include these effects in the present model. It seems likely that the polymer molecule size will bring with it the volume scale required to turn the variance expressions derived in this paper into quantitative predictors of LER.

Since the very early days of semiconductor manufacturing, researchers have attempted to predict the limits of optical lithography. As barriers to improvements in resolution were discovered, novel means of defying the limits were inevitably found. Stochastic limits to resolution, in the form of line-edge roughness, may be the most fundamental limit to lithographic resolution. It is unclear how low line-edge roughness can be pushed, but progress in reducing LER has been painfully slow over the last decade. A comprehensive and physically accurate stochastic model of lithography is needed before the ultimate limits of optical lithography will be known, and eventually reached.

Endnotes

1. T. Mulders, W. Henke, K. Elian, C. Nolscher, and M. Sebald, "New stochastic post-exposure bake simulation method," *J. Microlith. Microfab. Microsyst.*, 4, p. 043010 (2005).
2. A. Philippou, T. Mülders, E. Schöll, "Impact of photoresist composition and polymer chain length on line-edge roughness probed with a stochastic simulator," *J. Micro/Nanolith. MEMS MOEMS*, 6, p. 043005 (2007).
3. A. Saeki, T. Kozawa, S. Tagawa, H. Cao, H. Deng, M. Leeson, "Exposure dose dependence on line-edge roughness of a latent image in electron beam/extreme ultraviolet lithographies studied by Monte Carlo technique," *J. Micro/Nanolith. MEMS MOEMS*, 6, p. 043004 (2007).
4. Gerard M. Schmid, Michael D. Stewart, Sean D. Burns, and C. Grant Willson, "Mesoscale Monte Carlo Simulation of Photoresist Processing," *J. Electrochem. Soc.*, Vol. 151, pp. G155-G161 (2004).
5. Mark D. Smith, "Mechanistic model of line-edge roughness," *Advances in Resist Technology and Processing XXIII, Proc.*, SPIE Vol. 6153, p. 61530X (2006).
6. Chris A. Mack, "Fundamental Principles of Optical Lithography: The Science of Microfabrication," J. Wiley and Sons (London: 2007).
7. Chris A. Mack, "Line-Edge Roughness and the Ultimate Limits of Lithography," SPIE Vol. 7639 (2010).
8. John J. Biafore, Mark D. Smith, Tom Wallow, Patrick Nalleau, David Blankenship, and Yunfei Deng, "Pattern prediction in EUV resists," *Proc. SPIE* 7520, 75201P (2009).
9. C. A. Mack, "Stochastic Modeling in Lithography: Autocorrelation Behavior of Catalytic Reaction-Diffusion

- Systems,” *Journal of Micro/Nanolithography, MEMS, and MOEMS*, Vol. 8, No. 2, p. 029701 (2009).
10. D. ben-Avraham and S. Havlin, “Diffusion and Reactions in Fractals and Disordered Systems,” Cambridge University Press (Cambridge: 2000), Chapter 13.
 11. John J. Biafore, Mark D. Smith, Stewart A. Robertson, and Trey Graves, “Mechanistic simulation of line-edge roughness,” Proc. SPIE 6519, 65190Y (2007).
 12. T. B. Michaelson, *et al.*, “The Effects of Chemical Gradients and Photoresist Composition on Lithographically Generated Line-edge roughness,” *Advances in Resist Technology and Processing XXII*, SPIE Vol. 5753, pp. 368-379 (2005).
 13. P. Tsiartas, L. Flanagan, C. Henderson, W. Hinsberg, I. Sanchez, R. Bonnacaze, and C.G. Willson, “The Mechanism of Phenolic Polymer Dissolution: A New Perspective,” *Macromolecules*, Vol. 30, pp. 4656-4664 (1997).
 14. C. Mack, “Stochastic approach to modeling photoresist development,” *Journal of Vacuum Science & Technology*, Vol. B27, No. 3, pp. 1122-1128 (2009).
 15. V. Constantoudis, G. Patsis, and E. Gogolides, “Photoresist line-edge roughness analysis using scaling concepts,” *J. Microlithogr. Microfabrication, Microsyst.*, Vol. 3, p. 429 (2004).
 16. C. A. Mack, “Stochastic Modeling in Lithography: The Use of Dynamical Scaling in Photoresist Development,” *J. Micro/Nanolith. MEMS MOEMS*, 8, (2009).
 17. G. Gallatin, N. Kita, T. Ujike, and B. Partlo, “Residual speckle in a lithographic illumination system,” *J. Micro/Nanolith. MEMS MOEMS*, 8, 043003 (2009).
 18. O. Noordman, A. Tychkov, J. Baselmans, J. Tsacoyeanes, G. Politi, M. Patra, V. Blahnik, and M. Maul, “Speckle in optical lithography and its influence on line-width roughness,” *J. Micro/Nanolith. MEMS MOEMS*, 8, 043002 (2009).

About the Author

Chris A. Mack

Dr. Chris A. Mack developed the lithography simulation software PROLITH, and founded and ran the company FINLE Technologies for 10 years. He then served as vice president of Lithography Technology for KLA-Tencor for five years, until 2005. In 2003, Dr. Mack received the SEMI Award for North America for his efforts in lithography simulation and education, and in 2009, he received the SPIE Frits Zernike Award for Microlithography. Dr. Mack is also an adjunct faculty member at the University of Texas at Austin. Currently, he writes, teaches and consults on the field of semiconductor microlithography in Austin, Texas. ■