Line-Edge Roughness and the Ultimate Limits of Lithography

Chris A. Mack

www.lithoguru.com, Austin, Texas

Abstract

In this paper, 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 stochastics 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 acid diffusion on 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 ignore the discrete nature of these fundamental units and use instead 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.

One common approach to studying LER formation is through the use of Monte Carlo simulations[1,2,3] 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. The goal of this paper is provide a progress report on this as-yet incomplete effort.

Much of the treatment given below follows that provided in Ref. 6, with more recent advances included. First, the statistics of photon shot noise are reviewed, providing the standard Poisson statistics of photon counting. Chemical concentrations also result in counting statistics that are Poisson. The probabilities of absorption and exposure are combined with photon and chemical concentration shot noise 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 is touched upon next, but the details are left to another publication. Finally, pulling the final results together, a fist attempt at a comprehensive line-edge roughness model is attempted, though many deficiencies remain.

1.   Photon Shot Noise

Consider a light source that randomly emits photons at an average rate of L photons per unit time into some area A. Assume further that each emission event is independent. Over some small time interval dt (smaller than $1/L$ and small enough so that it is essentially impossible for two photons to be emitted during that interval), either a photon is emitted or it is not (a binary proposition). The probability that a photon will be emitted during this interval will be Ldt. Consider now some long time T ($\gg dt$). What can we expect for the number of photons emitted during the period T? This basic problem is called a Bernoulli trial and the resulting probability distribution is the well-known *binomial distribution*. If $N = T/dt$, the number of time intervals in the total time, then the probability that exactly n photons will be emitted in this time period is given by a binomial distribution $P(n)$. The binomial distribution is extremely cumbersome to work with as N gets large. If, however, $NLdt = TL$ remains finite as N goes to infinity, the binomial distribution converges to another, more manageable equation called the Poisson distribution:

$$P(n) = \frac{(TL)^n}{n!} e^{-TL} \tag{1}$$

Since there is no limit to how small dt can be made, letting dt go to zero will by default make N go to infinity for any nonzero time interval T and nonzero photon emission rate L.

The Poisson distribution can be used to derive the statistical properties of photon emission. The expectation value of n [that is, the mean number of photons that will be emitted in a time interval T, denoted by the notation $E(n)$ or $\langle n \rangle$] is TL (a very reasonable result since L was defined as the average rate of photon emission). The variance (the standard deviation squared) is also TL. To use these statistical properties, we must convert from number of photons to a more useful measure, intensity. If $n_{photons}$ is the number of photons that cross an area A over a time interval T, the mean intensity of light will be

$$\langle I \rangle = \frac{\langle n_{photons} \rangle}{TA} \left(\frac{hc}{\lambda} \right) = \frac{L}{A} \left(\frac{hc}{\lambda} \right) \tag{2}$$

where h is Planck's constant, c is the vacuum speed of light, and λ is the vacuum wavelength. The standard deviation of the intensity can also be computed from the properties of the Poisson distribution.

$$\sigma_I = \frac{1}{TA}\left(\frac{hc}{\lambda}\right)\sigma_n = \frac{\langle I \rangle}{\sqrt{TL}} = \frac{\langle I \rangle}{\sqrt{\langle n_{photons} \rangle}} \tag{3}$$

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 193-nm exposure of a resist with a dose-to-clear of 10 mJ/cm^2. 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 1 nm X 1 nm, the mean number of photons during the exposure, from equation (2), is about 97. The standard deviation is about 10, 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.

For *Extreme Ultraviolet Lithography* (EUVL), the situation will be considerably worse. At a wavelength of 13.5 nm, the energy of one photon will be 1.47×10^{-17} J, about fifteen times greater than at 193 nm. Also, the goal for resist sensitivity will be to have EUV resists that are 2–4 times more sensitive than 193-nm 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 193-nm lithography. A 1 nm X 1 nm area will see only two to three photons, and a 100-nm^2 area will see on the order of 200 photons, with a standard deviation of 7%.

2. Chemical Concentration

Concentration, the average number of molecules per unit volume, exhibits counting statistics identical to photon emission. Let C be the average number of molecules per unit volume, and dV a volume small enough so that at most one molecule may be found in it (thus requiring that the concentration be fairly dilute, so that the position of one molecule is independent of the position of other molecules). The probability of finding a molecule in that volume is just CdV. For some larger volume V, the probability of finding exactly n molecules in that volume will be given by a binomial distribution exactly equivalent to that for photon counting. And, as before, this binomial distribution will also become a Poisson distribution by letting dV go to zero.

$$P(n) = \frac{(CV)^n}{n!}e^{-CV} \tag{4}$$

The average number of molecules in the volume will be CV, and the variance will also be CV. The relative uncertainty in the number of molecules in a certain volume will be

$$\frac{\sigma_n}{\langle n \rangle} = \frac{1}{\sqrt{\langle n \rangle}} = \frac{1}{\sqrt{CV}} \tag{5}$$

[The requirement that the concentration be 'dilute' can be expressed as an upper limit to the Poisson distribution – for a given molecule size, saturation occurs at some n_{max} molecules in the volume V. So long as $P(n_{max})$ is small, the mixture can be said to be dilute[7].]

As an example, consider a 193-nm resist that has an initial PAG concentration of 3% by weight, or a concentration of about 0.07 mole/liter (corresponding to a 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 $(10 \text{ nm})^3$, the mean number of PAG molecules will be 42. The standard deviation will be 6.5 molecules, or about 15%. For 248-nm 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%. 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.

As mentioned briefly above, Poisson statistics apply only for reasonably low concentrations. The random distribution of molecules assumes that the position of each molecule is independent of all the others. As concentrations get higher, the molecules begin to 'crowd' each other, reducing their randomness. In the extreme limit, molecules become densely packed and the uncertainty in concentration goes to zero. This saturation condition is a function of not only the concentration, but the size of the molecule as well. To avoid saturation, the volume fraction occupied by the molecules under consideration must be small.

3. 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.

Consider a single molecule of PAG. First-order kinetics of exposure can be used to derive equation the concentration of PAG remaining after exposure (and, as well, the concentration of acid generated) in the continuum approximation (this is also called the *mean-field* solution to the kinetics of exposure). From a stochastic modeling perspective, this kinetic result represents a probability density function for reaction: G/G_0 is the fraction of PAG that is unreacted in some large volume, and by the Law of Large Numbers this must be the probability that any given PAG will remain unexposed. Let y be a random variable that represents whether a given single PAG molecule remains unexposed or was converted to acid by the end of the exposure process. Thus $y = 0$ means an acid has been generated (PAG has reacted), and $y = 1$ means the PAG has not been exposed (no acid generated). A kinetic analysis of exposure gives us the probability for each of these states, given a certain intensity-in-resist I:

$$P(y=0\,|\,I)=1-e^{-CIt}, \qquad P(y=1\,|\,I)=e^{-CIt} \tag{6}$$

The probability of exposing *one* acid molecule after exposure can now be translated into a mean and uncertainty of the overall acid concentration after exposure. Consider a volume V that initially contains some number $n_{0\text{-}PAG}$ PAG molecules. After exposure, the number of remaining (unexposed) PAG molecules Y will be

$$Y = \sum_{i=1}^{n_{0\text{-}PAG}} y_i \tag{7}$$

Assuming that each exposure event is independent, the mean of Y becomes

$$\langle Y \rangle = \langle n_{0-PAG} \rangle \langle y \rangle \tag{8}$$

The variance will be[6]

$$\sigma_Y^2 = \langle Y \rangle + \langle Y \rangle^2 \left(e^{(Ct\sigma_I)^2} - 1 \right) \tag{9}$$

Since σ_I, as given by equation (3), will in general be small,

$$\sigma_Y^2 \approx \langle Y \rangle + \frac{\langle Y \rangle^2 (C \langle I \rangle t)^2}{\langle n \rangle} \tag{10}$$

The variance of Y has two components. The Poisson chemical distribution gives the first term, $\langle Y \rangle$. Photon shot noise adds a second term, inversely proportional to the mean number of photons.

At this point it is useful to relate the number of remaining PAG molecules per unit volume Y to the concentration of acid H, and the initial number of PAGs n_{0-PAG} to the initial PAG concentration G_0.

$$H = G_0 - \frac{Y}{N_A V}, \quad G_0 = \frac{n_{0-PAG}}{N_A V} \tag{11}$$

where N_A is Avogadro's number. We can also define a relative acid concentration h to be

$$h = \frac{H}{\langle G_0 \rangle} = \frac{G_0}{\langle G_0 \rangle} - \frac{Y}{\langle n_{0-PAG} \rangle} = \frac{n_{0-PAG}}{\langle n_{0-PAG} \rangle} - \frac{Y}{\langle n_{0-PAG} \rangle} \tag{12}$$

The means of these quantities can be related by

$$\langle h \rangle = \frac{\langle H \rangle}{\langle G_0 \rangle} = 1 - \frac{\langle Y \rangle}{\langle n_{0-PAG} \rangle} = 1 - \langle y \rangle \tag{13}$$

Using equation (12), the variance of the acid concentration can be calculated as

$$\sigma_h^2 = \frac{\langle h \rangle}{\langle n_{0-PAG} \rangle} + \frac{\sigma_Y^2 - \langle Y \rangle}{\langle n_{0-PAG} \rangle^2} \tag{14}$$

Finally, using equation (10), the variance in acid concentration will be

$$\sigma_h^2 = \frac{\langle h \rangle}{\langle n_{0-PAG} \rangle} + \frac{[(1 - \langle h \rangle) \ln(1 - \langle h \rangle)]^2}{\langle n_{photons} \rangle} \tag{15}$$

This final result, which accounts for photon fluctuations, uncertainty in the initial concentration of photoacid generator, and the probabilistic variations in the exposure reaction itself, is reasonably intuitive. The first term on the right-hand side of equation (15) 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 within 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. For the case of the $(10 \text{ nm})^3$ of 193-nm 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%. For 193-nm 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 and acid. This mechanism will not be treated here but has been investigated by others.[8]

4.   Acid Catalyzed Reaction–Diffusion

In this section and the next, we'll consider the polymer deblocking reaction. In the continuum limit, the amount of blocked polymer left after the PEB is given by

$$M = M_0 e^{-K_{amp} t_{PEB} h_{eff}}$$

$$h_{eff}(x, y, x) = \frac{1}{t_{PEB}} \int_0^{t_{PEB}} \left(h(x, y, z, t = 0) \otimes DPSF \right) dt = h(x, y, z, t = 0) \otimes RDPSF \qquad (16)$$

As before, the latent image of acid after exposure, $h(x, y, z, t = 0)$ used in the continuum approximation is actually the mean acid concentration $\langle h \rangle$, with a standard deviation given above. The effective acid concentration, however, has a very specific interpretation: it is the time average of the acid concentration at a given point. The interesting question to be answered, then, is whether this time-averaging effect of diffusion coupled with the acid-catalyzed reaction affects the uncertainty in the effective acid concentration compared to the original acid concentration uncertainty.

To determine the statistical properties of the effective acid concentration, we'll begin by looking at the diffusion of a single molecule of acid. Let the binary random variable $y_i(t)$ represent whether that molecule is found in some small volume dV located a distance r_i from its original location, during the interval of time between t and $t + dt$. It will be given by the standard Gaussian diffusion kernel:

$$P(y_i(t) = 1) = \left(2\pi\sigma_D^2 \right)^{-3/2} e^{-r_i^2 / 2\sigma_D^2} dV , \quad \sigma_D^2 = 2Dt \qquad (17)$$

where D is the acid diffusivity and σ_D is the acid diffusion length. For n_i acid molecules at this location that then diffuse, the total number of acid molecules in that volume dV and over the same time interval will be $Y_i(t)$:

$$Y_i(t) = \sum_{j=1}^{n_i} y_{ij}(t), \quad \langle Y_i(t) \rangle = \langle n_i \rangle \langle y_i(t) \rangle, \quad \sigma_{Y_i}^2 = \langle Y_i(t) \rangle \tag{18}$$

Adding up the contributions from all of the locations that could possibly contribute acid molecules into the volume dV during the interval of time between t and $t + dt$ produces the standard convolution result:

$$Y(t) = \sum_i Y_i(t)$$

$$\langle Y(t) \rangle = \sum_i \langle n_i \rangle \langle y_i(t) \rangle = \langle n \rangle \otimes DPSF(t) \tag{19}$$

We now wish to integrate over time, from 0 to t_{PEB}.

$$Y = \frac{1}{t_{PEB}} \int_0^{t_{PEB}} Y(t) dt$$

$$\langle Y \rangle = \frac{1}{t_{PEB}} \int_0^{t_{PEB}} \langle Y(t) \rangle dt = \frac{1}{t_{PEB}} \int_0^{t_{PEB}} \langle n \rangle \otimes DPSF(t) dt = \langle n \rangle \otimes RDPSF \tag{20}$$

Thus, as expected, the effective acid concentration used in the continuum approximation is in fact the mean value of a stochastic random variable. The uncertainty of Y, however, involves some extra complications. Leaving the details of the derivation to Ref. 6, we obtain

$$\sigma_Y^2 = \langle n \rangle \otimes CovPSF$$

$$CovPSF = \frac{1}{dV\, t_{PEB}^2} \int_0^{t_{PEB}} \int_0^{t_{PEB}} E(y(t)\, y(t')) dt\, dt' \tag{21}$$

where $CovPSF$ is a new function that I call the 'covariance point spread function'.

Carrying out the integration above numerically (and again leaving the details of the derivation aside), the result (in 3D) becomes

$$CovPSF(r) \approx 2 \left(\frac{a}{\sigma_D} \right)^2 RDPSF(r) \tag{22}$$

where a is the capture radius of the deblocking reaction (called the von Smoluchowski trap radius).

The impact of the $CovPSF$ can now be determined. The shape of the $CovPSF$ is very similar to that of the $RDPSF$. Since the effective acid concentration near the line edge does not differ appreciably from the acid concentration when convolved with the $RDPSF$, the same will be true when convolved with the $CovPSF$. Thus, the effective acid concentration can be approximated as

$$\langle h_{eff} \rangle = \langle h \rangle \otimes RDPSF \tag{23}$$

The standard deviation of the effective acid concentration is approximated as

$$\sigma_{h_{eff}} \approx \sqrt{2}\left(\frac{a}{\sigma_D}\right)\sigma_h \tag{24}$$

As equation (24) indicates, if the acid diffuses a distance less than the reaction capture range, the catalytic nature of the amplification reaction actually increases the stochastic variation in the effective acid concentration compared to the original acid concentration. If, however, the diffusion length is greater than this capture range, the time-averaging effect of the catalytic reaction will smooth out stochastic roughness. It is not diffusion, *per se*, that reduces stochastic uncertainty, but rather the diffusion of a reaction catalyst that does so. Since in real resist systems the diffusion length will invariably be greater than the reaction capture distance, the net affect will always be a reduction in the effective acid concentration standard deviation.

5. Reaction–Diffusion and Polymer Deblocking

The stochastics of the deblocking of a single blocked site will follow along the same lines as the single PAG exposure analysis of section 3. Let y be a random variable that represents whether a given single blocked site remains blocked by the end of the PEB. Thus $y = 1$ means the site remains blocked, and $y = 0$ means the site has been deblocked. As before, the continuum kinetic analysis gives us the probability that a single site is deblocked for a given effective acid concentration.

$$P(y=1\,|\,h_{eff}) = e^{-K_{amp}t_{PEB}h_{eff}} \;,\qquad P(y=0\,|\,h_{eff}) = 1 - e^{-K_{amp}t_{PEB}h_{eff}} \tag{25}$$

The probability distribution of h_{eff}, however, is not obvious. While the relative acid concentration has a Poisson distribution, the time-averaging effect on the acid diffusion turns the discrete acid random variable into a continuous effective acid random variable.

It will be reasonable to assume that h_{eff} is normally distributed with mean and standard deviations as given in the previous section. Thus, the mean value of y becomes

$$\langle y \rangle = \frac{1}{\sqrt{2\pi}\sigma_{h_{eff}}} \int_{-\infty}^{\infty} \left(e^{-K_{amp}t_{PEB}h_{eff}}\right) e^{-\left(h_{eff}-\langle h_{eff}\rangle\right)^2 / 2\sigma_{h_{eff}}^2}\, dh_{eff}$$

$$\langle y \rangle = e^{-K_{amp}t_{PEB}\langle h_{eff}\rangle} e^{\frac{1}{2}\left(K_{amp}t_{PEB}\sigma_{h_{eff}}\right)^2} \tag{26}$$

The random variable y has a log-normal probability distribution and equation (26) can be recognized as the standard result for a log-normal distribution.

The total number of blocked groups remaining in a certain small volume will be given by Y.

$$Y = \sum_{i=1}^{n_{0-block}} y_i \tag{27}$$

The mean of Y can be easily computed, as before.

$$\langle Y \rangle = \langle n_{0-block} \rangle \langle y \rangle \tag{28}$$

The variance of Y can be found with a result similar to that for photon shot noise during exposure:

$$\sigma_Y^2 = \langle Y \rangle + \langle Y \rangle^2 \left(e^{\left(K_{amp} t_{PEB} \sigma_{h_{eff}} \right)^2} - 1 \right) \tag{29}$$

From the definitions of M and Y,

$$\langle M_0 \rangle = \frac{\langle n_{0-blocked} \rangle}{N_A V}, \quad \langle M \rangle = \frac{\langle Y \rangle}{N_A V}, \quad m \equiv \frac{M}{\langle M_0 \rangle} \tag{30}$$

Thus,

$$\langle m \rangle = \langle y \rangle \tag{31}$$

and

$$\sigma_m^2 = \frac{\sigma_M^2}{\langle M_0 \rangle^2} = \frac{\sigma_Y^2}{\langle n_{0-blocked} \rangle^2} \tag{32}$$

giving

$$\sigma_m^2 = \frac{\langle m \rangle}{\langle n_{0-blocked} \rangle} + \langle m \rangle^2 \left(e^{\left(K_{amp} t_{PEB} \sigma_{h_{eff}} \right)^2} - 1 \right) \tag{33}$$

For small levels of effective acid uncertainty,

$$\sigma_m^2 \approx \frac{\langle m \rangle}{\langle n_{0-blocked} \rangle} + \langle m \rangle^2 \left(K_{amp} t_{PEB} \sigma_{h_{eff}} \right)^2 = \frac{\langle m \rangle}{\langle n_{0-blocked} \rangle} + \left(\langle m \rangle \ln \langle m \rangle \right)^2 \left(\frac{\sigma_{h_{eff}}}{\langle h_{eff} \rangle} \right)^2 \tag{34}$$

As before, the first term captures the Poisson uncertainty due to the initial distribution of blocked polymer. The second term captures the influence of the effective acid concentration uncertainty. Combining this expression with the variance of the effective acid concentration,

$$\sigma_m^2 = \frac{\langle m \rangle}{\langle n_{0-blocked} \rangle} + \left(\langle m \rangle \ln \langle m \rangle \right)^2 \left(\frac{\sigma_h}{\langle h \rangle} \right)^2 \left(\frac{\sqrt{2} a}{\sigma_D} \right)^2 \tag{35}$$

Finally, using equation (15) for the variance of the acid concentration,

$$\sigma_m^2 = \frac{\langle m \rangle}{\langle n_{0-blocked} \rangle} + \left(\frac{\langle m \rangle \ln \langle m \rangle}{\langle h \rangle} \right)^2 \left(\frac{\sqrt{2}a}{\sigma_D} \right)^2 \left(\frac{\langle h \rangle}{\langle n_{0-PAG} \rangle} + \frac{[(1 - \langle h \rangle) \ln(1 - \langle h \rangle)]^2}{\langle n \rangle} \right) \qquad (36)$$

Or, in a slightly different form,

$$\left(\frac{\sigma_m}{\langle m \rangle} \right)^2 = \frac{1}{\langle n_{0-blocked} \rangle \langle m \rangle} + \left(K_{amp} t_{PEB} \right)^2 \left(\frac{\sqrt{2}a}{\sigma_D} \right)^2 \left(\frac{\langle h \rangle}{\langle n_{0-PAG} \rangle} + \frac{[(1 - \langle h \rangle) \ln(1 - \langle h \rangle)]^2}{\langle n \rangle} \right) \qquad (37)$$

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 $\sigma_D a$.

Using the example of a typical 193-nm resist, $M_0 N_A = 1.2$ /nm^3, $G_0 N_A = 0.042$ /nm^3 and $K_{amp} t_{PEB} = 2$. Consider the case of $\langle h \rangle = \langle h_{eff} \rangle = 0.3$, and $\sigma_D/a = 5$. For a $(10\ \text{nm})^3$ 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 have $\langle m \rangle = 0.55$ and $\sigma_m = 0.023$, or about 4.3%. For a $(5\ \text{nm})^3$ volume, $\sigma_m = 0.064$, or about 11%.

6. Autocorrelation Behavior of Reaction-Diffusion

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 *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,

$$R_{H_{eff}} = \sigma_H^2 \left(RDPSF \otimes RDPSF \right) \qquad (38)$$

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

$$\tilde{R}_{H_{eff}}(\tau) = \frac{\displaystyle\int_{-\infty}^{\infty} RDPSF(x) RDPSF(x + \tau) dx}{\displaystyle\int_{-\infty}^{\infty} [RDPSF(x)]^2 dx} \qquad (39)$$

For the 2D and 3D cases, integrations are best done in polar and spherical coordinates, respectively. This allows the double and triple integrals, respectively, to become single integrals over distance r [of the form of equation (39)] by multiplying the 2D *RDPSF* by $\sqrt{|r|}$ and the 3D *RDPSF* by $|r|$. Analytical evaluation of equation (39) for the 1D, 2D, and 3D cases does not seem possible, so numerical integrations were

performed. Figure 1 shows the results. Each of these results can be extremely well approximated by a standard exponential correlation function:

$$\tilde{R}_{H_{eff}}(\tau) = e^{-\left(|\tau|/\zeta\right)^{2\alpha}} \tag{40}$$

where ζ is the correlation length and α is the Hurst (roughness) exponent. Fitting the numerical evaluation of equation (39) to the empirical function (40) produces the results shown in Table I, where both a linear fit to the autocorrelation function and to the logarithm of the autocorrelation function were performed. The resulting fits are extremely good – plotting the linear fits on Figure 1 would produce lines indistinguishable from the calculated results from equation (39). Obviously, the linear fit does a better job of matching the small-τ behavior while logarithmic fitting results in better matching to the large-τ region.

Table I. Results of the best fit of equation (40) to the numerically evaluated equation (39), using a least-squares fit to $\tilde{R}_{H_{eff}}$ (linear fit) or to its logarithm (logarithmic fit).

	Linear Fit		Logarithmic Fit	
	ζ/σ_D	α	ζ/σ_D	α
1D	1.266	0.848	1.252	0.817
2D	1.532	0.936	1.515	0.901
3D	1.528	0.900	1.519	0.879
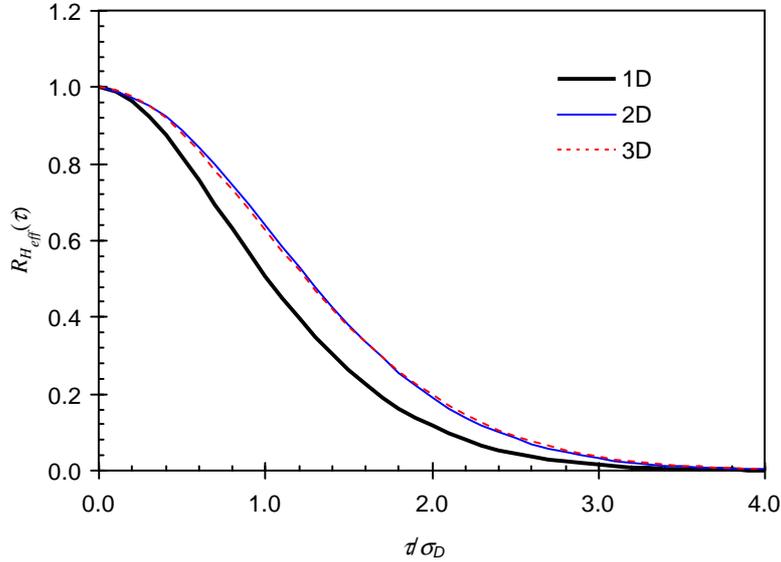
Figure 1. Numerical evaluation of the *RDPSF* autocorrelation for the 1D (thick solid black line), 2D (thin solid blue line), and 3D (dashed line) cases.

These results show that diffusion of the catalyst in a first-order reaction-diffusion system produces persistent correlation ($\alpha > 0.5$), with a correlation length that is a multiple of the diffusion length (as expected). For the important 3D case, $\alpha \approx 0.9$ and the correlation length is just over 50% greater than the catalyst diffusion length.

7. 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 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 being presented here.

8. 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] An accompanying publication in this volume[17] addresses recent work on this subject, but the current state of that research does not yet allow its integration with the model presented here.

9. 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, 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):

$$\Delta x = \frac{\Delta m^*}{dm^*/dx} \qquad (41)$$

From this, we can devise a simple qualitative model for line-edge roughness. 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:

$$LER \propto \frac{\sigma_{m^*}}{dm^*/dx} \tag{42}$$

Michaelson[12] plotted measured LER versus calculated values of dm^*/dx and found that many different resists followed an almost universal curve. The curve, however, was slightly different than that given by equation (42). In fact, it is well fit by the following empirical expression:

$$LER = \frac{33}{dm^*/dx} + 5 \tag{43}$$

where LER is the 3σ value, in nanometers, and dm^*/dx is in units of $1/\mu m$. The constant term at the end has been the subject of much speculation, and could be related to the influence of development on LER.

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 by

$$\frac{\partial m^*}{\partial x} \approx \frac{1}{\eta} e^{-\alpha_f h_{eff}} \left(1 - e^{-\eta \alpha_f}\right)\left(1 - h_{eff}\right)\ln\left(1 - h_{eff}\right)\frac{\partial \ln I}{\partial x} \tag{44}$$

where $\alpha_f = K_{amp}t_{PEB}$, and 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 reaction. From equation (36) we can see how to minimize the statistical uncertainty in deprotection. There is one interesting variable in common to both: acid diffusion. Increasing acid diffusion will reduce σ_{m^*}, but will 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 (36) and (44) into (42). Thus, for the no-quencher case, and ignoring photon shot noise,

$$\sigma_{m^*} = \sqrt{\frac{\langle m^* \rangle}{\langle n_{0-blocked} \rangle} + \left(\langle m^* \rangle \ln\langle m^* \rangle\right)^2 \left(\frac{\sqrt{2}a}{\sigma_D}\right)^2 \left(\frac{1}{\langle n_{0-PAG} \rangle \langle h \rangle}\right)}$$

$$\frac{\partial m^*}{\partial x} \propto \frac{1}{\sigma_D{}^2}\left(1 - e^{-\pi^2 \sigma_D{}^2/2L^2}\right) \tag{45}$$

so that

$$LER \propto \frac{\sigma_D{}^2}{1 - e^{-\pi^2 \sigma_D{}^2/2L^2}} \sqrt{\frac{\langle m^* \rangle}{\langle n_{0-blocked} \rangle} + \left(\langle m^* \rangle \ln\langle m^* \rangle\right)^2 \left(\frac{\sqrt{2}a}{\sigma_D}\right)^2 \left(\frac{1}{\langle n_{0-PAG} \rangle \langle h \rangle}\right)} \tag{46}$$

and

$$LER \propto \frac{\sigma_D{}^2}{1-e^{-\pi^2\sigma_D{}^2/2L^2}}\sqrt{1-(K_{amp}t_{PEB})\langle m*\rangle \ln\langle m*\rangle \left(\frac{\sqrt{2}a}{\sigma_D}\right)^2 \frac{\langle n_{0-block}\rangle}{\langle n_{0-PAG}\rangle}} \qquad (47)$$

Figure 2 shows the trend of LER versus acid diffusion for a 45-nm feature for three different values of the deprotection capture range a, 0.5, 1.0, 2.0 and 3.0 nm. 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

$$\sigma_D{}^2 \approx \sqrt{2K}\left(\frac{aL}{\pi}\right) - \frac{Ka^2}{4} \qquad (48)$$

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

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

$$a \ll \sigma_D \ll L \qquad (49)$$

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 amplification will be controlled by the amplification factor α_f:

$$\alpha_f = K_{amp}t_{PEB} = 2\pi\sigma_D{}^2 aG_0 N_A \qquad (50)$$

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 much 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.
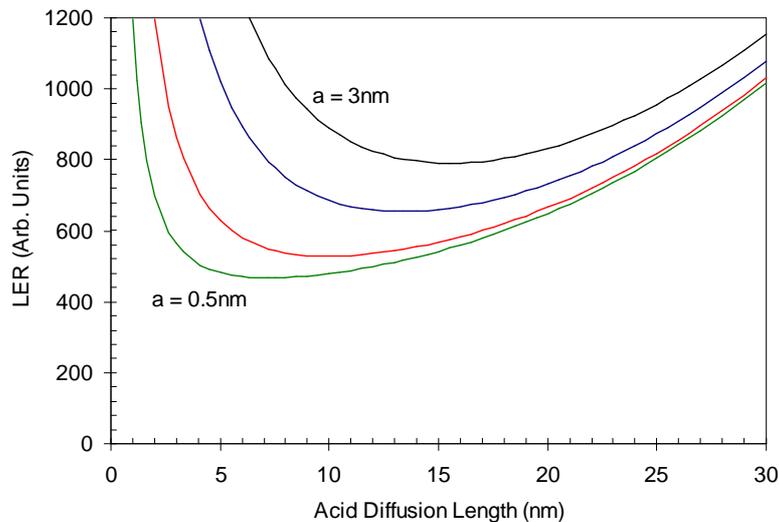
Figure 2   Prediction of LER trends for a 45 nm feature for three values of the deblocking reaction capture range a (0.5, 1, 2, and 3 nm).

10. 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 here begins with photon shot noise. Speckle has not been discussed here, though recent studies have made very good progress in understanding this phenomenon for 193-nm lithography.[18,19] 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 concentration is reduced whenever the acid diffusion length is greater than the von Smoluchoski trap radius. Thus, acid 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.

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.

[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] Jeffrey Byers, private communication.

[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. Flanagin, C. Henderson, W. Hinsberg, I. Sanchez, R. Bonnecaze, 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] Chris A. Mack, "Stochastic modeling of photoresist development in 2D and 3D", *these proceedings*.

[18] G. Gallatin, N. Kita, T. Ujike, and B. Partlo, "Residual speckle in a lithographic illumination system" *J. Micro/Nanolith. MEMS MOEMS*, 8, 043003 (2009).

[19] 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 linewidth roughness", *J. Micro/Nanolith. MEMS MOEMS*, 8, 043002 (2009).