Resist Blur and Line Edge Roughness

Gregg M. Gallatin
IBM T. J. Watson Research Center, Yorktown Heights, NY 10598, gallatin@us.ibm.com

ABSTRACT

A straightforward analytic model of resist line edge roughness is presented which predicts all the known scaling laws as well as the shape of the experimentally seen frequency content or power spectrum of the roughness. The model implies there are strong basic limitations to achieving, simultaneously, low roughness, low dose and high resolution in any standard chemically amplified resist process. A simple model of how roughness maps to device performance is also presented.

1. INTRODUCTION

As shown in Cobb, et. al.[1] and elsewhere[2−6] resist line edge roughness (LER) is typically dominated by spatial frequencies lower than roughly 20 to 30 cycles/micron. The rms roughness, σ_{LER}, on the other hand, is typically on the order of only a few nm. Thus the roughness is long range compared to its amplitude which indicates it should have little to do with molecular weight or chemical statistics. In fact, as has recently been shown by Brainard, et. al.[6] and Cutler, et. al.[7] LER is relatively independent of molecular weight, at least for one class of chemically amplified resists. This is not to say that such effects do not contribute at all to LER. They no doubt do contribute. But, given the fact that almost the entire value of σ_{LER} comes from low frequencies it follows that high frequency effects cannot be the dominant contributors and to understand LER we must understand the origin of the low frequency content. Here, a straightforward analytical model of LER is presented which predicts the low frequency dominated shape of the roughness power spectrum and also explains the standard scaling laws that have been seen experimentally. The model is fully three dimensional with resist line edges being treated as two dimensional surfaces.

The model consists of two parts. First, the aerial image, properly normalized, is taken as the probability distribution for where a photon will be absorbed and release an acid. Up to saturation effects, the joint probability to release N acids is the product of N single acid release probabilities as is consistent with the rules of quantum mechanics. Second, the shape of the post exposure bake (PEB) deprotection "blur" which occurs around each released acid is derived from the reaction diffusion equations[8] and combined with the exposure statistics to yield the deprotection statistics. Resist development, implemented using what can be viewed as a large scale equivalent to the critical ionization model[9−12] , yields the complete functional form for the parametric dependence and frequency content of the LER. Similar approaches to modeling and understanding LER have been presented by Fukuda[13,14] and others.[6,15−17] Here, the ideas incorporated in those approaches are extended and combined to yield an analytical representation of LER.

The net perspective of the model is that the dominant contributor to LER is the randomness in the positions of acids released during exposure combined with the superposition of the resist deprotection "blurs" generated around each released acid during the PEB. One consequence of the model is that it appears that it will be extremely difficult, if not essentially impossible, in any standard chemically amplified resist process, to achieve, simultaneously, low dose, low LER and high resolution. Such a conclusion, if true, has important implications for the future of lithography.

Given the conceptual simplicity of the model it is interesting that it automatically incorporates some "extra" effects. First, the intensity distribution in exposure tools commonly covers an area which is large compared to individual image features, e.g., a millimeter or larger exposure area as compared to fractions of a micron for individual features. It follows directly from the probability distribution used by the model that, for individual features, the number of acids released obeys Poisson statistics even if the overall total number of released acids is fixed. Hence the model automatically incorporates shot noise effects and they do not have to be put in by hand. Note that although real exposure tools don't count photons they are designed to operate so that the total dose, i.e., the total number of photons, N, absorbed across the entire exposure field is controlled to a given

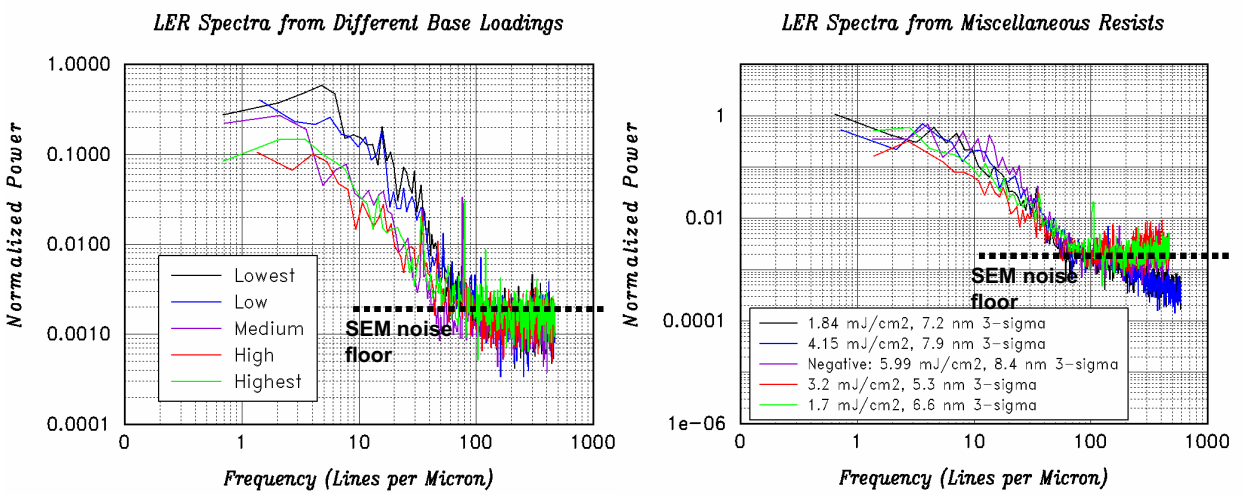

Figure 1. Data from Cobb, et. al.[1] showing that LER is dominated by low frequencies for various resists and processing conditions. The flattening of the spectrum a high frequencies corresponds to pixel-to-pixel white noise in the SEM images. Given that σ^2_{LER} is the area under the spectrum it follows that almost all the contribution to σ_{LER} comes from length scales longer than roughly 30nm. This specific data is for EUV, but the same frequency content is seen in LER for 248, 193 and 157nm resists.[2–6]

value. Of course, over the nominal exposure area N is extremely large and so its fractional fluctuation, $1/\sqrt{N}$ is much smaller than the dose control requirement. Second, in most, but certainly not all current exposure tools, the resist blur is small compared to the image resolution. In the model this condition leads automatically to the standard reciprocal of the image slope dependence of LER seen in many LER models, i.e., the lower the resolution the worse the LER. But, the model automatically switches behavior in the opposite case where the blur is larger and shows that LER will still be present even for infinite edge slope but will follow a different scaling law. Third, the model yields the standard $1/\sqrt{dose}$ dependence seen in essentially all LER models, but again, without explicitly imposing Poisson statistics. This scaling law implies that LER should go to zero in the limit of a large, effectively infinite, dose. But, this ignores saturation effects which imply that the *dose* must asymptotically approach a maximum value dependent on the quantum efficiency and acid loading of the resist and thus LER must approach a fixed nonzero value for fixed acid loading even for infinite dose. Finally, the model explicitly shows that the PEB blur range sets the frequency content of the roughness and that σ_{LER} should decrease slowly with increasing blur range, all else being held fixed. Each of these results is described below.

2. EXPOSURE AND ACID RELEASE STATISTICS

Consistent with the rules of quantum mechanics the probability for a photon to be absorbed, and release an acid, at position $\vec{r} = (x, y, z)$ inside the volume of the resist is proportional to the local dose $E(\vec{r}) = \varepsilon t_{\text{exp}} I(\vec{r})$ where ε is the quantum efficiency for acid release, t_{exp} is the exposure time and $I(\vec{x})$ the image intensity. The exact formula is somewhat complicated and depends on various parameters[18] but this form is sufficient for our purposes. Of course for an acid to be released at $\vec{x}$, a photoacid generator (PAG) must be present initially at that position. The probability distribution of the PAG is nominally uniform throughout the resist volume and if the dose is below saturation then we can ignore the PAG statistics and we have that the probability for an acid to be released at $\vec{x}$ during exposure is given by

$$p(\vec{r}) = \frac{E(\vec{r})}{\int_V E(\vec{r})\, d^3r} = \frac{I(\vec{r})}{\int_V I(\vec{r})\, d^3r} \tag{1}$$

where $d^3r = dxdydz$ The integral in the denominator normalizes the probability distribution so that. $\int_V p(\vec{r})\, d^3r = 1$, with the integral taken over the volume $V = AT$ of the resist being exposed. Here A is the total exposed area (again on the order of square mm) and T is the thickness of the resist.

Up to saturation effects which imply that a given acid cannot be released twice, the probability distribution for N acids to be released at positions $\vec{r}_1, \vec{r}_2, \cdots, \vec{r}_N$ is given by

$$P(\vec{r}_1, \vec{r}_2, \cdots, \vec{r}_N) = p(\vec{r}_1)\, p(\vec{r}_2) \cdots p(\vec{r}_N) \tag{2}$$

which is automatically normalized as $\int_V P(\vec{r}_1, \cdots, \vec{r}_N)\, d^3r_1 \cdots d^3r_N = 1$. This probability distribution is consistent the old quantum experiment of sending photons through, say, a double slit one at a time and watching the interference intensity distribution get filled in discretely, step-by-step, with each absorption.[19,20]

Importantly, even with the net number of released acids fixed at N on a large scale, the probability distribution defined above automatically incorporates shot noise into the model on the scale of individual features. The total number of absorptions in a volume $v = aT$ where a is the area of a small feature is given by $\sum_{m=1}^{N} \theta_v(\vec{r}_m)$ where $\theta_v(\vec{r}) = 1$ for $\vec{r}$ in v and is zero otherwise. The probability Q_n for n absorptions in v is then given by the integral of $P(\vec{r}_1, \cdots, \vec{r}_N)$ over V times a Kronecker delta-function $\delta\left(n, \sum_{m=1}^{N} \theta_v(\vec{r}_m)\right)$ which equals 1 for $n = \sum_{m=1}^{N} \theta_v(\vec{r}_m)$ and is zero otherwise and thus counts only those cases where $\sum_{m=1}^{N} \theta_v(\vec{r}_m) = n$, i.e.,

$$Q_n = \int_V P(\vec{r}_1, \cdots, \vec{r}_N)\ \delta\left(n, \sum_{m=1}^{N} \theta_v(\vec{r}_m)\right) d^3r_1 \cdots d^3r_N$$

Substituting the Fourier representation of $\delta(n, m)$, which is $1/2\pi \int_0^{2\pi} \exp(i\omega(n-m))\, d\omega$, using the binomial theorem and defining q as the probability for getting 1 acid release in volume v, i.e., $q = \int_v p(\vec{r})\, d^3r$ gives

$$\begin{aligned}
Q_n &= \frac{1}{2\pi} \int_0^{2\pi} \left[e^{i\omega n} \int_V P(\vec{r}_1, \cdots, \vec{r}_N) \exp\left(-i\omega \sum_{m=1}^{N} \theta_v(\vec{r}_m)\right) d^3r_1 \cdots d^3r_N \right] d\omega \\
&= \frac{1}{2\pi} \int_0^{2\pi} e^{i\omega n} \left(q e^{-i\omega} + (1-q)\right)^N d\omega \\
&= \sum_{m=0}^{N} \frac{N!}{m!\,(N-m)!} q^m (1-q)^{N-m} \frac{1}{2\pi} \int_0^{2\pi} e^{i\omega(n-m)} d\omega \\
&= \frac{N!}{n!\,(N-n)!} q^n (1-q)^{N-n}
\end{aligned}$$

For a nominal dose, N is extremely large since A is large. For a small feature with area $a \ll A$ we have, on average, that $q \ll 1$ and $n \ll N$. In this limit $N!/(N-n)! \to N^n$ and $(1-q)^{N-n} \to (1-q)^N \simeq \exp[-Nq]$ and thus

$$Q_n \simeq \frac{(Nq)^n}{n!} e^{-Nq} \tag{3}$$

which is the Poisson distribution. Hence even with a fixed total number of released acids on the scale of the exposure area A, the Poisson distribution, i.e., shot noise, still controls the statistics of the number of acids released in any given small feature with the average number being Nq as one would expect.

3. DEPROTECTION STATISTICS

The deprotection density, after PEB, surrounding an acid released at position $\vec{r}$ is derived in the Appendix. Neglecting saturation effects the net deprotection density for N acids released at positions $\vec{r}_1, \vec{r}_2, \cdots \vec{r}_N$ is then given by the sum of the corresponding deprotection blurs as

$$\rho_D^{net}(\vec{r}, \vec{r}_1, \cdots \vec{r}_N) = \sum_{n=1}^{N} \rho_D(\vec{r} - \vec{r}_n) \tag{4}$$

Note that ρ_D^{net} depends explicitly on the acid release positions $\vec{r}_1, \vec{r}_2, \cdots \vec{r}_N$. We assume that all its statistical behavior follows from statistics of the $\vec{r}_1, \cdots \vec{r}_N$ as given by $P(\vec{r}_1, \cdots \vec{r}_N)$. In other words, it is the randomness in the positions of the acid release that generates the dominant contribution to LER.

This approach does ignore the nanoscale chemical statistics or atomistic roughness which, although it must contribute as discussed above and elsewhere.[21,22,23] is apparently not the dominant generator of LER. In effect, here we are taking the deprotection distribution $\rho_D(\vec{r})$ as a continuous smooth function whereas, in fact, it should be treated as a probability distribution and sampled appropriately. Treating it as a smooth non-statistical function can be justified by the fact that the model developed below does a good job predicting the experimentally seen behavior of LER including the functional form of its frequency content.

Finally, it should be noted that treating resist development as the threshold of the convolution of the image, in the resist, with a resist deprotection blur function has been shown to do a good job describing real resist behavior.[24−30]

4. LER STATISTICS

For our purposes, the critical ionization model[9−12] can be interpreted to state that the edge of a developed resist line is determined implicitly by

$$\tau = \rho_D^{net}(\vec{r}_{edge}, \vec{r}_1, \cdots, \vec{r}_N) \tag{5}$$

where τ is the threshold deprotection density for resist development and $\vec{r}_{edge}$ maps out the entire shape and position of the two dimensional resist surface for the given values of $\vec{r}_1, \vec{r}_2, \cdots, \vec{r}_N$. Position on the two dimensional surface represented by $\vec{r}_{edge}$ requires two independent coordinates. Given the orientation of the edge considered below (see Figure 2) we will be able to take these coordinates to be x and z with x the horizontal position along a resist edge and z the vertical position. The appropriate value of τ to use in any given case depends on the net combination of the resist process conditions. For our purposes we don't need an explicit value of τ and require only the implicit relation given above.

The average resist surface is defined by integrating ρ_D^{net} times $P(\vec{r}_1, \cdots, \vec{r}_N)$ over V, setting the result equal to τ and solving for $\vec{r}_{edge}$. Doing the integration yields

$$\begin{aligned}
\tau &= \left\langle \rho_D^{net}(\vec{r}_{edge}, \vec{r}_1, \cdots, \vec{r}_N) \right\rangle \\
&= \int_V \rho_D^{net}(\vec{r}_{edge}, \vec{r}_1, \cdots, \vec{r}_N) P(\vec{r}_1, \cdots, \vec{r}_N) d^3r_1 \cdots d^3r_N \\
&= N \int_V \rho_D(\vec{r}_{edge} - \vec{r}) p(\vec{r}) d^3r
\end{aligned}$$

Thus, on average, the net resist surface is determined implicitly by the convolution of the resist blur with the properly normalized image intensity inside the volume of the resist, just as one would expect.[24−30]

To solve for $\vec{r}_{edge}$ consider the specific case of a line edge oriented along the x axis in the region of $y = 0$ so that $I(\vec{r}) = I(y)$ as shown in Figure 1. Thus $p(\vec{r}) = p(y)$ and so only y is determined from the above equation and we have

$$\tau = \left\langle \rho_D^{net}(\vec{r}_{edge}, \vec{r}_1, \cdots, \vec{r}_N) \right\rangle \Longrightarrow \tau = \left\langle \rho_D^{net}((x, y_{edge}, z), \vec{r}_1, \cdots, \vec{r}_N) \right\rangle \tag{6}$$

with x and z being independent and used to label position on the edge itself.

The above equation ignores line end effects as well as top and bottom (of the resist) effects. This is consistent with the experimental definition of LER which works with only a finite portion of a longer line and hence explicitly stays away from line ends. The neglect of top and bottom effects is justifiable when the resist is much thicker than the range of ρ_D as is often the case as resist thicknesses are generally greater than 100nm and the width of ρ_D is on the order of 30 to 50nm. Experimentally, some change in the character of LER near the top and bottom of the resist has been seen[31] but this change is certainly not large. Also, topcoats and bottomcoats are often acid loaded to help reduce "t-topping" and footing which also reduces top and bottom boundary effects. Resist absorption can be included by allowing I to depend on z as well as y. In any case details such as these are not relevant for understanding the general cause and properties of LER, which is our present purpose.

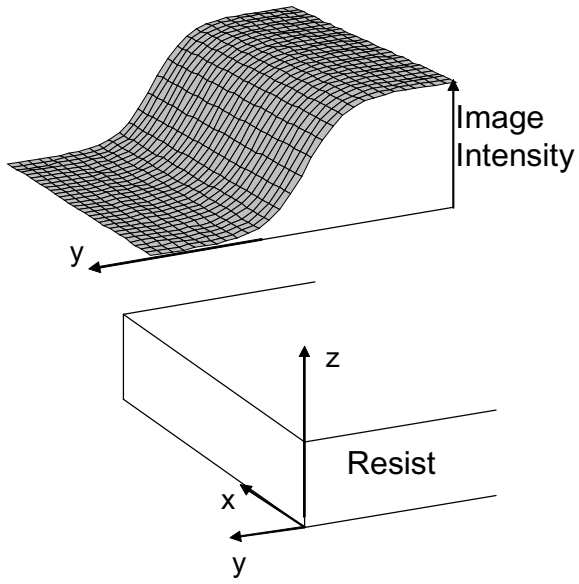

Figure 2. Geometry used for the derivation of LER given in the text. LER is reprsented as fluctuations in the resist surface δy about the nominal resist surface at $y = 0$.

Without loss of generality we can take the position of the average edge to be at $y = y_{edge} = 0$ with x covering a finite range along the line and z ranging from 0 to T over the resist thickness. Geometrically, as shown in the bottom of Figure 2 this is a flat vertical plane located at $y = 0$.

Fluctuations in $\rho_D^{net}\left(\left(x, y_{edge}, z\right), \vec{r}_1, \cdots, \vec{r}_N\right)$ about its average value due to the statistical variation in the acid release positions $\vec{r}_1, \cdots, \vec{r}_N$ cause the resist edge position to fluctuate in order to maintain the critical ionization criteria. Since LER is on the order of a few nm 1σ we can determine these fluctuations as follows. The fluctuation of ρ_D about the nominal average edge value is given by

$$\begin{aligned}\delta\rho_D\left(x, z, \vec{r}_1, \cdots, \vec{r}_N\right) &\equiv \rho_D\left(\left(x, y_{edge}, z\right), \vec{r}_1, \cdots, \vec{r}_N\right) - \left\langle \rho_D^{net}\left(\left(x, y_{edge}, z\right), \vec{r}_1, \cdots, \vec{r}_N\right)\right\rangle \\ &= \rho_D\left(\left(x, y_{edge}, z\right), \vec{r}_1, \cdots, \vec{r}_N\right) - \tau\end{aligned}$$

In order to compensate for this fluctuation and stay at the value τ the surface must shift by $\delta y\left(x, z, , \vec{r}_1, \cdots, \vec{r}_N\right)$ where the the position dependence of δy follows from the position dependence of $\delta\rho_D$. The critical ionization criteria then becomes

$$\begin{aligned}\tau &= \left\langle \rho_D^{net}\left(\left(x, y_{edge} + \delta y, z\right), \vec{r}_1, \cdots, \vec{r}_N\right)\right\rangle + \delta\rho_D\left(x, z, \vec{r}_1, \cdots, \vec{r}_N\right) \\ &= \left\langle \rho_D^{net}\left(\left(x, y_{edge}, z\right), \vec{r}_1, \cdots, \vec{r}_N\right)\right\rangle \\ &\quad + \left[\frac{\partial}{\partial y}\left\langle \rho_D^{net}\left(\left(x, y, z\right), \vec{r}_1, \cdots, \vec{r}_N\right)\right\rangle\right]_{y=y_{edge}} \delta y\left(x, z, \vec{r}_1, \cdots, \vec{r}_N\right) \\ &\quad + \delta\rho_D\left(x, z, \vec{r}_1, \cdots, \vec{r}_N\right)\end{aligned}$$

Note that the first term on the right hand side is just τ. Solving for δy, substituting for $\delta\rho_D$ and setting $y_{edge} = 0$ yields

$$\delta y\left(x, z, \vec{r}_1, \cdots, \vec{r}_N\right) = -\frac{\rho_D\left(\left(x, 0, z\right), \vec{r}_1, \cdots, \vec{r}_N\right) - \tau}{\partial\left\langle \rho_D^{net}\left(\left(x, 0, z\right), \vec{r}_1, \cdots, \vec{r}_N\right)\right\rangle / \partial y_{edge}} \tag{7}$$

Again, the dependence of δy on $\vec{r}_1, \cdots, \vec{r}_N$ is included specifically since the exact position and shape of the surface as defined by δy depends on the exact positions of the released acids before PEB. The term in the denominator has already been averaged and hence it does not depend on $\vec{r}_1, \cdots, \vec{r}_N$. In the denominator we must remember to set $y_{edge} = 0$ only after taking the derivative.

Taking the expectation value of Eq. (7) yields

$$\left\langle \delta y\left(x,z,\vec{r}_1,\cdots,\vec{r}_N\right)\right\rangle = -\frac{\left\langle \rho_D\left(\left(x,0,z\right),\vec{r}_1,\cdots,\vec{r}_N\right)-\tau\right\rangle}{\partial\left\langle \rho_D^{net}\left(\left(x,0,z\right),\vec{r}_1,\cdots,\vec{r}_N\right)\right\rangle/\partial y_{edge}} = 0 \tag{8}$$

as expected.

Letting $\vec{r}_{edge} = (x,0,z)$ and $\vec{r}\,'_{edge} = (x',0,z')$ be two positions on the average edge, the autocorrelation function of the roughness is given by

$$\begin{aligned}
\left\langle \delta y\left(x,z\right)\delta y\left(x',z'\right)\right\rangle &= \frac{\left\langle \left(\rho_D\left(\left(x,0,z\right),\vec{r}_1,\cdots,\vec{r}_N\right)-\tau\right)\left(\rho_D\left(\left(x',0,z'\right),\vec{r}_1,\cdots,\vec{r}_N\right)-\tau\right)\right\rangle}{\left(\partial\left\langle \rho_D^{net}\left(\left(x,0,z\right),\vec{r}_1,\cdots,\vec{r}_N\right)\right\rangle/\partial y_{edge}\right)^2} \\
&= \frac{-\tau^2+\sum_{n,m}\int\rho_D\left(\vec{r}_{edge}-\vec{r}_n\right)\rho_D\left(\vec{r}\,'_{edge}-\vec{r}_m\right)p\left(\vec{r}_1\right)\cdots p\left(\vec{r}_N\right)d^3r_1\cdots d^3r_N}{\left(N\int\rho_D\left(\vec{r}_{edge}-\vec{R}\right)\partial p\left(\vec{R}\right)/\partial Y\,d^3R\right)^2} \\
&= \frac{-\tau^2+\left(1-1/N\right)\tau^2+N\int\rho_D\left(\vec{r}_{edge}-\vec{R}\right)\rho_D\left(\vec{r}\,'_{edge}-\vec{R}\right)p\left(\vec{R}\right)d^3R}{\left(N\int\rho_D\left(\vec{r}_{edge}-\vec{R}\right)\partial p\left(\vec{R}\right)/\partial Y\,d^3R\right)^2} \\
&= \frac{1}{N}\frac{\int\rho_D\left(\vec{r}_{edge}-\vec{R}\right)\rho_D\left(\vec{r}\,'_{edge}-\vec{R}\right)p\left(\vec{R}\right)d^3R}{\left(\int\rho_D\left(\vec{r}_{edge}-\vec{R}\right)\partial p\left(\vec{R}\right)/\partial Y\,d^3R\right)^2}
\end{aligned} \tag{9}$$

in the limit of large N. Here, to avoid confusion, $\vec{R} = (X,Y,Z)$ is used as the integration variable. The above form follows from using $\int\rho_D\left(\vec{r}_{edge}-\vec{R}\right)p\left(\vec{R}\right)d^3R = \int\rho_D\left(\vec{r}\,'_{edge}-\vec{R}\right)p\left(\vec{R}\right)d^3\vec{R} = \tau/N$ and the denominator has been rewritten using

$$\begin{aligned}
\left[\int\frac{\partial\rho_D\left(\vec{r}_{edge}-\vec{R}\right)}{\partial y_{edge}}p\left(\vec{R}\right)d^3R\right]_{y_{edge}=0} &= -\int\frac{\partial\rho_D\left(\vec{r}_{edge}-\vec{R}\right)}{\partial Y}p\left(\vec{R}\right)d^3R \\
&= \int\rho_D\left(\vec{r}_{edge}-\vec{R}\right)\frac{\partial p\left(\vec{R}\right)}{\partial Y}d^3R
\end{aligned}$$

where the last step follows by integration by parts and $y_{edge} = 0$ on the right-hand side.

To make further progress we must evaluate the above integrals. This can be done numerically for any given case and it can be done analytically for particular functional dependencies. Here we will first approximate the integrals assuming just the generic properties of a blur shape to obtain the overall scaling law dependence of the mean square roughness on various parameters. Then, using the functional form of the deprotection distribution derived in the Appendix we will derive the expected shape of autocorrelation function of the LER which, when Fourier transformed, yields the frequency content or equivalently the power spectrum of the LER.

The mean square roughness σ_{LER}^2 is given by $\left\langle \delta y\left(x,z\right)^2\right\rangle$. To simplify the derivation let $I\left(\vec{r}\right) = I_o i\left(\vec{r}\right)$ with $i\left(\vec{r}\right)$ normalized to 1 in the middle of large features, e.g. open frame so that $i\left(\vec{r}\right)$ takes values in the range 0 to a little greater than 1, to allow for diffraction ringing and I_o is the open frame image intensity in units of incident photons/(area $\times$ time). The factors of I_o cancel in the definition of $p\left(\vec{r}\right)$ and so $p\left(\vec{r}\right) = i\left(\vec{r}\right)/\int_V i\left(\vec{r}\right)d^3r$. All finite regions where $i\left(\vec{r}\right)$ is strictly zero don't contribute and so, without loss of generality, we can restrict V to just those regions where $i\left(\vec{r}\right)\neq 0$. Assuming a reasonably transparent resist it follows that $\int_V i\left(\vec{r}\right)d^3r \simeq fV = fAT$ where f corresponds roughly to the pattern density. To simplify the derivation of how σ_{LER} scales with various parameters we assume here an on-off deprotection blur shape, i.e., $\rho_D\left(\vec{r}\right) = \bar{\rho}_D =$ non-zero constant for $r \leq R_{PEB}$ and is 0 outside R_{PEB}. (The derivation of the LER power spectrum given below will use the full form for ρ_D

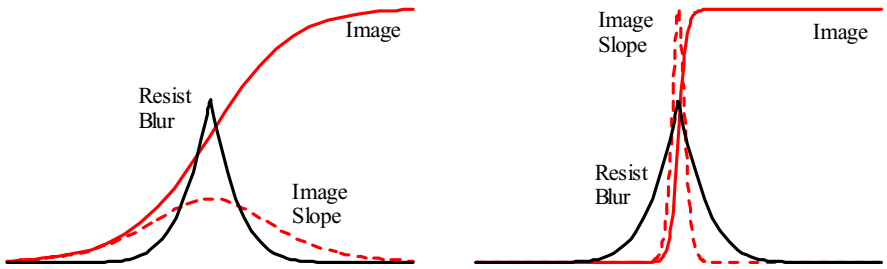

Figure 3. There are two different resolution regimes relevant for discussing LER. On the left the image has lower resolution than the resist. On the right this is reversed and the resist has lower resolution than the image. The effect of both cases on LER is discussed in the text.

derived in the Appendix.) Using this we have

$$\begin{aligned}\int_V \rho_D \left(\vec{r}_{edge}-\vec{R}\right)^2 p\left(\vec{R}\right) d^3R &\approx \frac{1}{fV}\int_V \rho_D \left(\vec{r}_{edge}-\vec{R}\right)^2 i\left(\vec{R}\right) d^3R \\ &\approx \left(\frac{i_{edge}\bar{\rho}_D^2}{f}\right)\left(\frac{V_D}{V}\right) \end{aligned} \tag{10}$$

where $V_D = 4\pi R_{PEB}^3/3 \simeq 4R_{PEB}^3$ is the volume of the deprotection blur. The resist diffusion blur or simple resist blur R_{PEB} is given by $R_{PEB} = (Dt_{PEB})^{1/2}$ where D is the diffusion constant and t_{PEB} is the PEB time.

The denominator in Eq. (10) has two generic forms depending on the relative scales of R_{PEB} and the image resolution as shown in the left and right graphs in Figure 3. If R_{PEB} is much smaller than the image resolution then the image edge slope is roughly constant over the range $2R_{PEB}$ and we have

$$\begin{aligned}\int_V \rho_D \left(\vec{r}_{edge}-\vec{R}\right) \partial p\left(\vec{R}\right)/\partial Y\, d^3R &\approx \frac{1}{fV}\int_V \rho_D \left(\vec{r}_{edge}-\vec{R}\right) \partial i\left(\vec{R}\right)/\partial Y\, d^3R \\ &\approx \left(\partial_\perp i_{edge}\right)\left(\frac{\bar{\rho}_D}{f}\right)\left(\frac{V_D}{V}\right) \end{aligned} \tag{11}$$

where $\partial_\perp i_{edge}$ indicates the image edge slope in the direction perpendicular the edge evaluated at the nominal edge position.

On the other hand, if the image has higher resolution than the blur then $\partial i\left(\vec{R}\right)/\partial Y \simeq \delta\left(Y - y_{edge}^{image}\right) \simeq \delta(Y)$ with the image edge position taken as $y_{edge}^{image} \simeq y_{edge} = 0$ in which case

$$\begin{aligned}\int_V \rho_D \left(\vec{r}_{edge}-\vec{R}\right) \partial p\left(\vec{R}\right)/\partial Y\, d^3R &\simeq \frac{1}{fV}\int_V \rho_D \left(\vec{r}_{edge}-\vec{R}\right) \partial i\left(\vec{R}\right)/\partial Y\, d^3R \\ &\simeq \left(\frac{\bar{\rho}_D}{f}\right)\left(\frac{A_D}{V}\right) \end{aligned} \tag{12}$$

where $A_D \simeq \pi R_{PEB}^2 =$ the blur area. This result shows that even in the limit of infinite edge slope, LER will not vanish but instead will follow a scaling law independent of the image edge slope.

Finally, combining the above results with the total number of acids released, which is given by

$$N = \varepsilon \int_A E(x,y,T)\, dxdy = \varepsilon I_o t_{\exp} \int_A i(x,y,T)\, dxdy \approx \varepsilon I_o t_{\exp} f A = \varepsilon E_o f V/T \tag{13}$$

where $E_o = I_o t_{\exp}$ is the open frame dose and using $E_o = E_{edge}/i_{edge}$, we have

$$\begin{aligned}
\sigma_{LER} &\approx \sqrt{\frac{T\left(\frac{i_{edge}\bar{\rho}_D^2}{f}\right)\left(\frac{V_D}{V}\right)}{\varepsilon\left(\frac{E_{edge}}{i_{edge}}\right) fV\left(\partial_\perp i_{edge}\right)^2\left(\frac{\bar{\rho}_D}{f}\right)^2\left(\frac{V_D}{V}\right)^2}} \\
&\approx \left(\frac{i_{edge}}{\partial_\perp i_{edge}}\right)\sqrt{\frac{T}{\varepsilon E_{edge} V_D}} \\
&\approx \left(\frac{i_{edge}}{\partial_\perp i_{edge}}\right)\sqrt{\frac{T}{\varepsilon E_{edge} 4R_{PEB}^3}} \qquad (14)
\end{aligned}$$

Although this formula is meant to be just a scaling law, the approximations used to obtain it are not too extreme and so we should expect it to be approximately numerically correct. Accurate numerical values of course require a detailed numerical evaluation of Eq. (9) with $\vec{r}_{edge} = \vec{r}\,'_{edge}$. In the high resolution case Eq. (12) must be used instead of Eq. (11) in which case the same result is obtained as above but with $1/\partial_\perp i_{edge}$ replaced by R_{PEB}. This result will be discussed in detail in the next section.

To evaluate the functional form of the LER frequency content, i.e., the shape of the power spectrum, we can ignore any overall constants or normalization and so only need to consider the numerator in Eq. (9) In the region of interest $p(\vec{r}) = p(y)$ and using the form for ρ_D derived in the Appendix for three dimensions yields

$$\begin{aligned}
&\int \rho_D\left(\vec{r}_{edge} - \vec{R}\right)\rho_D\left(\vec{r}\,'_{edge} - \vec{R}\right) p\left(\vec{R}\right) d^3R \\
&= \int \rho_D\left(x - X, -Y, z - Z\right)\rho_D\left(x' - X, -Y, z' - Z\right) p\left(Y\right) d^3R \\
&= \frac{1}{(2\pi)^4}\int d\beta_x d\beta_z d\beta_y d\beta'_y \left\{\left(\frac{1 - \exp\left[-t_{PEB} D\left(\beta_x^2 + \beta_z^2 + \beta_y^2\right)\right]}{D\left(\beta_x^2 + \beta_z^2 + \beta_y^2\right)}\right)\left(\frac{1 - \exp\left[-t_{PEB} D\left(\beta_x^2 + \beta_z^2 + \beta_y'^2\right)\right]}{D\left(\beta_x^2 + \beta_z^2 + \beta_y'^2\right)}\right)\right. \\
&\qquad \left.\times \exp\left[i\beta_x\left(x - x'\right) + i\beta_z\left(z - z'\right)\right]\int p\left(Y\right)\exp\left[-i\left(\beta_y + \beta'_y\right)Y\right] dY\right\}
\end{aligned}$$

In the region of the edge $p(Y)$ is, to a good approximation a linear function, i.e., $p(Y) = p_{edge} + sY$ where $s = \partial p_{edge}/\partial Y$ is a constant and therefore

$$\int p\left(Y\right)\exp\left[-i\left(\beta_y + \beta'_y\right)Y\right] dY = \left(p_{edge} + si\frac{\partial}{\partial\beta'_y}\right) 2\pi\delta\left(\beta_y + \beta'_y\right)$$

Noting that the factor in parentheses $\left(\frac{1-\exp\left[-t_{PEB} D\vec{\beta}^2\right]}{D\vec{\beta}^2}\right)$ is a function only of β_y^2 and, calling it $\mathrm{f}\left(\beta_y^2\right)$, we have

$$\int d\beta_y d\beta'_y \mathrm{f}\left(\beta_y^2\right)\mathrm{f}\left(\beta_y^{2\prime}\right)\frac{\partial}{\partial\beta'_y}\delta\left(\beta_y + \beta'_y\right) = -\int d\beta_y \mathrm{f}\left(\beta_y^2\right)\frac{\partial}{\partial\beta_y}\mathrm{f}\left(\beta_y'^2\right) = 0$$

and so the term proportional to s makes no contribution. The autocorrelation function then reduces to

$$\begin{aligned}
\left\langle \delta y\left(x, z\right)\delta y\left(x', z'\right)\right\rangle &\equiv \sigma_{LER}^2 C\left(x - x', z - z'\right) \\
&= \frac{\sigma_{LER}^2}{K}\int d^3\beta e^{i\beta_x\left(x - x'\right) + i\beta_z\left(z - z'\right)}\left(\frac{1 - \exp\left[-R_{PEB}^2\left(\beta_x^2 + \beta_z^2 + \beta_y^2\right)\right]}{R_{PEB}^2\left(\beta_x^2 + \beta_z^2 + \beta_y^2\right)}\right)^2
\end{aligned}$$

where $K = 4\pi\left(2\sqrt{\pi} - \sqrt{2\pi}\right)/R_{PEB}^3 \simeq 4\pi/R_{PEB}^3$ is a normalization constant chosen so that $C\left(0, 0\right) = 1$. Note that C depends only on the coordinate differences $x - x'$ and $z - z'$ as expected. The frequency content or power

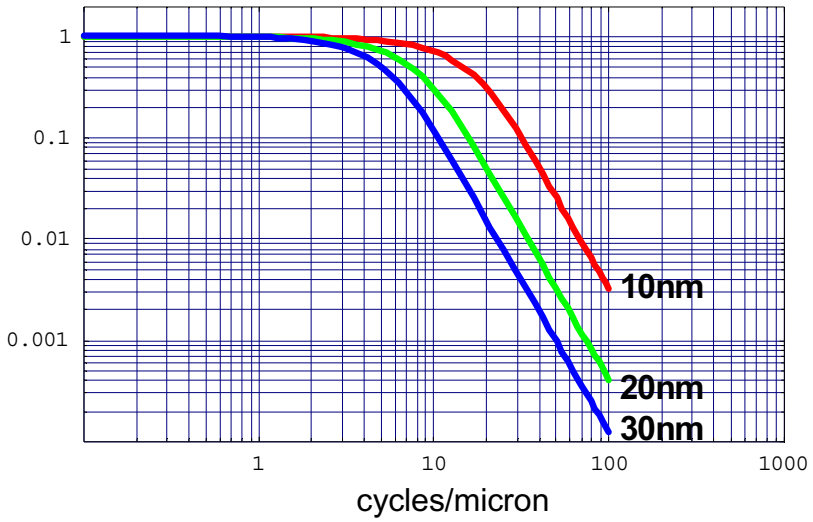

Figure 4. LER frequency spectrum predicted by the model, Eq. (15), normalized to unity for small β for three values of R_{PEB}. The shape compares well to measured LER frequency distributions such as those shown in Figure 1.

spectrum $F\left(\beta_x, \beta_y\right)$ of the LER is the Fourier transform of the autocorrelation function. Letting $\beta_x^2 + \beta_z^2 = \beta_\perp^2$, $\beta_y = \alpha$ we have

$$\begin{aligned} F\left(\beta_{x,}\beta_z\right) &= \sigma_{LER}^2 \int \exp\left[-i\left(\beta_x x + \beta_z z\right)\right] C\left(x,z\right) \frac{dxdz}{\left(2\pi\right)^2} \\ &= \frac{\sigma_{LER}^2}{4\pi/R_{PEB}^3} \int d\alpha \left(\frac{1 - \exp\left[-R_{PEB}^2\left(\beta_\perp^2 + \alpha^2\right)\right]}{R_{PEB}^2\left(\beta_\perp^2 + \alpha^2\right)} \right)^2 \\ &= \frac{\sigma_{LER}^2}{8\, R_{PEB}^3\, \beta^3} \left[\begin{array}{c} \frac{2}{\sqrt{\pi}} \left(\sqrt{2} e^{-2\, R_{PEB}^2\, \beta^2} - 2\, e^{-R_{PEB}^2\, \beta^2} \right) R_{PEB}\, \beta \\ + \left(2 - 4\, R_{PEB}^2\, \beta^2\right) \mathrm{erf}(R_{PEB}\, \beta) \\ + \left(-1 + 4\, R_{PEB}^2\, \beta^2\right) \mathrm{erf}(\sqrt{2}\, R_{PEB}\, \beta) \end{array} \right] \end{aligned} \tag{15}$$

This function, normalized to unity for small β, is plotted in Figure 4 for $R_{PEB} = 10, 20$ and 30 nm. Clearly this matches the shape of the frequency content of LER seen experimentally implying that the model presented here is capturing the dominant contribution to LER. Note that the resist sidewall roughness as described by the above formula is isotropic as seen experimentally.[31]

Fourier transforming the power spectrum yields the shape of the autocorrelation function. This is shown for one dimension in Figure 5. The horizontal axis is in units of R_{PEB} indicating the autocorrelation length defined as the value of x where C drops to $1/e$ of its peak value is roughly $R_{PEB}/\sqrt{2}$.

5. DISCUSSION

The scaling-law result for σ_{LER} given in Eq. (14) contains the "known" LER dependencies: proportionality to the reciprocal of the image log slope and proportionality to $1/\sqrt{dose}$ as seen in many studies.[32,33] The *dose* here is the dose at the line edge. Thus, the $1/\sqrt{dose}$ dependence will not be seen if the dose is varied simply by over exposing or under exposing a given image with a given resist since the resist edge will move so as to remain at nominally same dose value. The variation in LER which occurs with such over exposing or under exposing is due rather to the edge moving to positions with different image slope values and can be understood graphically as shifting the resist blur function left and right relative to the image slope in Figure 3. This shift reduces the contribution to the integrals in both Eqs. (11) and (12) and hence increases the LER. The effect this has on

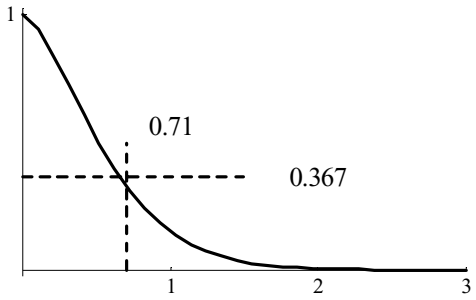

Figure 5. The autocorrelation function derived from the analytic form of the power spectrum. The horizontal axis is in units of R_{PEB}. The impact of the finite correlation length on the device performance is discussed in the text.

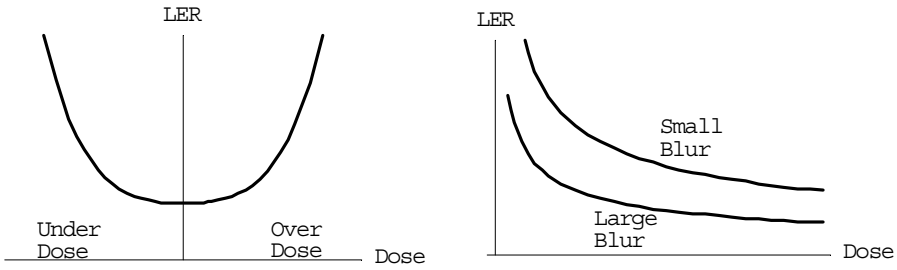

Figure 6. The graph on the left shows, schematically, the LER behavior predicted by the model when the resist edge is shifted with respect to the image edge as would happen when over dosing or under dosing a given resist. Depending on the image shape and flare level the curve can be non-symmetric and/or shifted horizontally. The graph on the right shows, schematically, the LER behavior predicted by the model if the resist is altered so that different doses are required to keep the resist edge in the same position with respect to the image edge.

LER is shown schematically in the left graph in Figure 6. To see the $1/\sqrt{dose}$ dependence requires varying the resist and/or the resist process, such as, for example, by changing base loading, so that the edge stays fixed at the same position relative to the image for the different dose values.[6] Finally, as mentioned in the introduction, for very large doses we should expect the $1/\sqrt{dose}$ behavior to saturate which means LER will cease to decrease with increasing dose.

As discussed above the image log slope dependence is present only when the blur is short range compared to the image edge slope. A different scaling law, still with finite LER, holds for the opposite case, even for perfect image resolution, i.e., infinite image log slope.

The derivation above ignored boundary effects at the top and bottom of the resist. Given that, the dependence of LER on T can be understood as follows. Increasing T with N fixed increases the average spacing between released acids and the greater this average distance, the greater the roughness, all else being fixed. A different dependence on T should be expected when the boundary conditions at the top and bottom of the resist come into play as is necessary for a thin resist process.

Finally, LER decreases with increasing R_{PEB} since, for a given density of released acids, increasing the blur increases the smearing effect of the resist which decreases the roughness but this also amounts to a decrease in resist resolution precisely because of the smearing effect it has on the image as captured by the resist.

The model is fully three dimensional with the resist edge treated as a two dimensional surface with isotropic roughness. This is consistent with expermental data.[31] In the two dimensional approximation where the resist edge is treated as a one dimensional line, repeating the above derivation yields the same basic dependencies but without a T in the numerator in Eq.(14) and R_{PEB} occurs in the denominator only to the square rather than the cubic power. In both the two and three dimensional cases the power spectrum as given in Eq. (15) will have the same form.

Although Eq. (14) is just a scaling law, and not an exact result, it is worth showing that it predicts roughly correct values for $3\sigma_{LER}$. The image resolution factor $i_{edge}/\partial_{\perp} i_{edge}$ depends on the imaging system, the illumination conditions, the dose level, and the pattern, and can vary over a rather wide range even for a given NA and wavelength. Here we use the approximation $i_{edge}/\partial_{\perp} i_{edge} \sim \lambda/2\pi NA$. Letting $\lambda = 193nm$, $NA = 0.75$, $T = 200nm$, $\varepsilon = 0.1$ so that the resist is 90% transmitting, $E_{edge} \sim E_0/4$ with $E_0 = 20mJ/cm^2$, and $R_{PEB} = 20nm$ yields $3\sigma_{LER} \sim 6nm$ which is certainly the right order of magnitude. For EUV with $\lambda \simeq 13nm$ and $NA = 0.25$ the image resolution factor $i_{edge}/\partial_{\perp} i_{edge}$ is no longer valid, as discussed above. Using Eq. (12), instead of Eq. (11) it is replaced in Eq. (14) by $i_{edge} R_{PEB}$ with the square root factor remaining the same. Ignoring flare and taking $\varepsilon = 0.5$, $E_{edge} \sim E_0/4$ with $E_0 = 4mJ/cm^2$, $R_{PEB} = 10nm$ to allow for the required increase in resolution at EUV feature sizes, and $T = 150nm$ gives $3\sigma_{LER} \sim 3nm$. Again, these numerical estimates should not be taken to be accurate predictions or limitations of any technology. They are only meant to indicate that the scaling law Eq. (14), and it's high resolution version as applied to EUV, are predicting at least the correct order of magnitude of the LER. When this is combined with the prediction of the frequency content from Eq. (15) the implication is that the model is properly capturing the dominant cause of LER.

We end the discussion with a derivation of the impact of LER on device performance following that given by Croon.[34] First, it is gate length variation or roughness, and not resist roughness itself, which is directly relevant and it must be remembered that the pattern transfer process itself may significantly alter the character of the roughness.[31] [NOTE: terminology switches between device considerations and lithography considerations, device gate length = lithography line width and device gate width = lithography line length, and so LER maps to gate length variation or gate length roughness when considering device performance.] Here we assume that all the intervening processes between resist development and the final physical device can be represented as factors multiplying the LER and its coherence length. That is, we assume that the gate length roughness $\sigma_{GLR} \propto \sigma_{LER}$ and the coherence length of the gate length roughness, $\xi \propto R_{PEB}$. This is obviously an oversimplification since factors other that LER contribute to σ_{GLR}, such as etch parameters as discussed by Goldfarb, et. al.[31] In any case, the gate length roughness can be characterized by particular values of σ_{GLR} and ξ and we can evaluate their consequences, independent of their origin. Let x label position along the width (= lithographers length) of a gate of total width W and let $L(x)$ be the gate length (= lithographers width) at position x. Now consider some relevant device parameter such as a current I which, neglecting short channel effects can be written as

$$I = \int_0^W J(L(x))\,dx = \int_0^W J\left(\bar{L} + \delta L(x)\right)dx = J\left(\bar{L}\right)W + \frac{\partial J\left(\bar{L}\right)}{\partial \bar{L}} \int_0^W \delta L(x)\,dx + \cdots$$

We have expanded to first order in the fluctuation $\delta L(x)$ of L about its average value $\bar{L}$ and $J(L)$ is the current density for a channel of length L. We take $\delta L(x)$ to have a statistical average of zero, i.e., $\langle \delta L(x) \rangle = 0$ so that $\langle I \rangle = J\left(\bar{L}\right)W$. The fluctuation in I, σ_I, follows from

$$\sigma_I^2 = \left\langle \left(I - \langle I \rangle\right)^2 \right\rangle = \left(\frac{\partial J\left(\bar{L}\right)}{\partial \bar{L}}\right)^2 \int_0^W \int_0^W \langle \delta L(x)\,\delta L(x') \rangle\,dx dx' \cong \left(\frac{\partial J\left(\bar{L}\right)}{\partial \bar{L}}\right)^2 \sigma_{GLR}^2 2\xi W$$

where we have used $\langle \delta L(x)\,\delta L(x') \rangle = \sigma_{GLR}^2 C(x - x')$ with C having an autocorrelation length of ξ satisfying $W >> \xi$. This leads to the relative fluctuation in I of the form

$$\frac{\sigma_I}{\langle I \rangle} \cong \frac{\partial\left(\ln\left(J\left(\bar{L}\right)\right)\right)}{\partial \bar{L}} \left(\frac{\sigma_{GLR}}{\sqrt{W/2\xi}}\right)$$

This result implies that gate length roughness including any contribution coming from resist LER is deweighted by the squareroot of the ratio of the gate width (= lithographers length) to the autocorrelation length of the

roughness. In particular it implies that when $W >> \xi$ the impact of roughness on device performance should be minimal. The above derivation stopped at first order in δL. It is straightforward to carry this to higher order and derive the quadratic, cubic, etc., dependence of $\sigma_I / \langle I \rangle$ on σ_{GLR} if desired. For a more detailed discussion and evaluation of the impact of LER on device performance see Oldiges, et. al.[35] and also[36,37] .

6. SUMMARY

The fact that the functional form of the power spectrum given in Eq (15) has the same shape as the measured frequency content for essentially all experimental cases and that the area under this shape accounts for almost all the contribution to σ^2_{LER} implies that the model presented here describes the dominant contribution to LER. Given that and rewriting Eq (14) heuristically as

$$\sigma^2_{LER} \times dose \times blur \simeq \text{constant} \tag{16}$$

we see that it appears to be extremely difficult, if not impossible, for a resist to have, simultaneously, low LER, low dose and high resolution, i.e., low blur, such as is required by EUV. This balance between the different factors is shown schematically in the right graph in Figure 6. It might be argued that the solution is to lower the value of the constant and certainly it can be varied[6] but there appears to be an almost purely geometric limitation to how low its value can be, at least in any standard chemically amplified resist. To understand this, consider the following. At low dose the acids are far apart and the volume in between them needs to be filled in to a sufficient deprotection density by the PEB in order for the resist to develop and this implies a large blur. On the other hand, a small blur requires the acids to be sufficiently close together for them to combine and add up to the required deprotection level. It should be noted that the term "low" as used as used here is relative and specific values have to be compared to detailed specifications for any particular case. Finally the model shows that even in the limit of high resolution, LER remains finite while ceasing to have a strong dependence on image slope.

7. APPENDIX

The Hinsberg-Houle chemical kinetics model and associated code[8] accounts for four basic chemical processes plus a coarse grained acid diffusion. Let $[P]$ represent protected polymer, $[D]$ represent deprotected polymer and $[A]$ represent acid. The processes are: (1) Acid catalyzed deprotection: $[P] + [A] \rightarrow [D] + [A]$. (2) Local hopping of the acid from a protected site to a deprotected site and vice versa: $([D] + [A]) + [P] \rightleftharpoons [D] + ([P] + [A])$. The rates for the forward and reverse processes are different and represent an environment dependent random walk on the atomic scale which is equivalent to an environment dependent local diffusion of the acid. Properly coarse grained this process should merge into the standard diffusion equation and I will treat it as such. (3) Thermally activated deprotection: $[P] \rightarrow [D]$. (4) "Autocatalysis": $[P] + 2[D] \rightarrow 3[D]$. In this case the deprotected polymer acts to deprotect more polymer. (5) Acid diffusion on the nm and greater length scale with diffusion dependent on the local environment, i.e., the diffusion constant depends on the local density of protected and deprotected polymer.

These processes can be represented as

$$\begin{aligned} \partial_t \rho_P(\vec{r},t) &= -k_A \rho_A(\vec{r},t)\,\rho_P(\vec{r},t) - k_T \rho_P(\vec{r},t) - k_D \rho_D(\vec{r},t)^2\,\rho_P(\vec{r},t) \\ \partial_t \rho_A(\vec{r},t) &= \vec{\partial} \cdot \Big(D(\rho_P,\rho_D)\,\vec{\partial}\rho_A(\vec{r},t) \Big) \end{aligned}$$

Here ρ_A, ρ_P, ρ_D are the densities of acid, protected and deprotected polymer respectively and $\partial_t = \partial/\partial t$ with $\vec{\partial} = (\partial_x, \partial_y, \partial_z)$ = gradient.

The densities are assumed to be defined on a coarse grained scale on the order of a cubic nm or so. The various k coefficients are the rate constants for each step. The three terms in the first equation represent steps 1,3,4 respectively. The second equation is the diffusion equation with $D(\rho_P, \rho_D)$ the acid diffusion constant. We can normalize things (neglecting resist shrinkage) by writing $\rho_P + \rho_D = 1$ in which case $D(\rho_P, \rho_D) \rightarrow D(\rho_P)$. (Actually we should write $\rho_P + \rho_D = \bar{\rho}$ where $\bar{\rho}$ is the real density of crosslinks, but if we redefine ρ_P and ρ_D by

dividing by $\bar{\rho}$ we get the normalization used above. Note that this rescaling does not affect the linear terms in the first equation but it does alter the scaling of the nonlinear term.) Under the assumption that for the cases of interest the density of deprotected polymer stays relatively small we can approximate $D(\rho_P) \simeq D = \text{constant}$. The diffusion equation is linear in ρ_A. With D depending on ρ_P or equivalently ρ_D, substituting the solution of the diffusion equation into the first equation will yield a nonlinear first term. But, if D is taken to be constant then both the first and second term in the first equation are explicitly linear. In this case the equation with just these two terms is solvable analytically and this is the assumption that we will make here. The last term in the first equation is explicitly nonlinear. If only this term was present on the right hand side then again this nonlinear equation would be analytically solvable. But with all three terms present a complete analytic solution is apparently not possible. Base loading effects are not included here as they make the equations explicitly nonlinear as well.

NOTE: The first equation is kinetic only, i.e., it has no space derivatives and so, with the initial condition that $\rho_P(\vec{r}, t=0) = \text{constant}$, all the position dependence of ρ_D and ρ_P enters via the position dependence of ρ_A. This is an important point. It indicates that it is the solution to the diffusion equation which contains the image pattern information and that ρ_P and ρ_D, although they may "evolve" or propagate this dependence via their kinetics, they inherit the position information from ρ_A and cannot generate any on their own.

With $D = \text{constant}$ the solution to the diffusion equation can be written as

$$\rho_A(x,t) = e^{tD\vec{\partial}^2}\rho_A(x,0)$$

This rather arcane way of writing the solution is particularly convenient since if $\rho_A(\vec{r},0)$ is expressed as a superposition of basis functions satisfying particular boundary conditions then $\rho_A(\vec{r},t)$ will also satisfy the same boundary conditions. Hence boundary conditions can easily be incorporated into the solution.

Consider the Greens function form of the solution in one dimension, i.e., $\vec{r} \to x$. Take $\rho_A(x,0) = C\delta(x) = C\int \frac{d\beta}{2\pi} e^{i\beta x}$ which is the appropriate superposition for the infinite space $-\infty \le x \le +\infty$ and C is a constant which normalizes the initial acid density to the desired value. Substituting this above yields

$$\rho_A^{1D}(x,t) = C\int e^{tD\partial_x^2} e^{i\beta x}\frac{d\beta}{2\pi} = C\int e^{-tD\beta^2 + i\beta x}\frac{d\beta}{2\pi} = \frac{C}{\sqrt{4\pi Dt}} e^{-x^2/4Dt}$$

which is the standard result.

Substituting the $1D$ solution for the diffusion equation into the first equation and ignoring the third (nonlinear) term yields

$$\rho_P^{1D}(x,t) = Exp\left[-k_A C\int_0^t \frac{e^{-x^2/4Dt}}{\sqrt{4\pi Dt}}dt - k_T t\right]$$

The integral in the exponent is

$$\int_0^t \frac{e^{-x^2/4Dt}}{\sqrt{4\pi Dt}}dt = \sqrt{\frac{t}{\pi D}}e^{-x^2/4Dt} + \frac{x}{2D}\left(\text{erf}\left(x/\sqrt{4Dt}\right) - 1\right)$$

and so

$$\rho_D^{1D}(x,t) = 1 - \rho_P^{1D}(x,t) = 1 - Exp\left[-k_A C\left(\sqrt{\frac{t}{\pi D}}e^{-x^2/4Dt} + \frac{x}{2D}\left(\text{erf}\left(x/\sqrt{4Dt}\right) - 1\right)\right) - k_T t\right]$$

The thermal term is small and we shall ignore it from here on. Using the following values for the parameters: $D = 7.8nm^2/\sec$, $t = 120\sec$ and $k_A C = 0.2$ yields the result for ρ_D shown in Figure 5. This matches closely the result presented in Hinsberg et. al.[24] Essentially the same formula for ρ_D was presented recently in[38] .

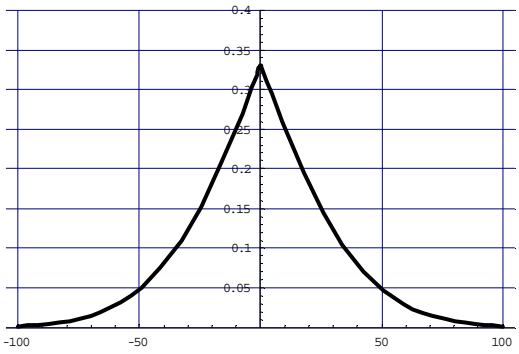

Figure 7. One dimensional PEB deprotection blur using the formula and values given in the text. The result matches closely the data presented in Hinsberg et. al.[24] The horizontal scale is in nm and the vertical scale indicated deprotection density.

In three dimensions we have

$$\begin{aligned}
\rho_D(\vec{x},t) &= 1-\exp\left[-k_A C\int_0^t\int e^{-tD\vec{\beta}^2}e^{i\vec{\beta}\cdot\vec{x}}\frac{d^3\beta}{(2\pi)^3}dt\right] \\
&= 1-\exp\left[-k_A C\int\frac{1-\exp\left(-tD\vec{\beta}^2\right)}{D\vec{\beta}^2}e^{i\vec{\beta}\cdot\vec{x}}\frac{d^3\beta}{(2\pi)^3}\right] \\
&\simeq k_A C\int\frac{1-\exp\left(-tD\vec{\beta}^2\right)}{D\vec{\beta}^2}e^{i\vec{\beta}\cdot\vec{x}}\frac{d^3\beta}{(2\pi)^3}
\end{aligned}$$

where the last line is valid for $k_A C << 1$.

8. ACKNOWLEDGEMENT

I would like to acknowledge many fundamental and useful discussions with Frances Houle, Jonathan Cobb, Dario Goldfarb, Alex Liddle, Patrick Naulleau and Robert Brainard. This work was supported, in part, under DARPA Contract #N66001-00-C8803.

REFERENCES

1. Jonathan L. Cobb, et. al. Proc. SPIE **4688**, 412 (2002)
2. V. Constantoudis, et. al., J. Vac. Sci. Technol. **B22**, 1974 (2004)
3. V. Constantoudis, et. al., J. Vac. Sci. Technol. **B21**, 1019 (2003)
4. Atsuko Yamaguchi and Osamu Komuro, Jpn. J. Appl. Phys. **42**, 3763 (2003)
5. J. A. Croon, et. al., IEDM 2002
6. Robert L. Brainard, et. al., Proc. SPIE **5374**, 74 (2004)
7. Charlotte A. Cutler, Proc. SPIE **5037**, 406 (2003)
8. F. A. Houle, et. al., J. Vac. Sci. Technol. **B18**, 1874 (2000)
9. P. C. Tsiartas, et. al., Macromolecules **30**, 4656 (1997)
10. L. W. Flanagin, et. al., Macromolecules **32**, 5337 (1999)
11. S. D. Burns, et. al., J. Vac. Sci. Technol. **B20**, 537 (2002)
12. William Hinsberg, et. al., Proc. SPIE **5376**, 352 (2004)
13. Hiroshi Fukuda, J. Photopolymer Sci. and Tech. **15**, 389 (2002)
14. Hiroshi Fukuda, Jpn. J. Appl. Phys. **42**, 3748 (2003)

15. Lei Yuan and Andrew R. Neureuther, Proc. SPIE **5376**, 312 (2004)
16. Gerard M. Schmid, Proc. SPIE **4690**, 381 (2002)
17. G. M. Gallatin, Proc. SPIE **4404**, 123 (2001)
18. Rodney Loudon, *The Quantum Theory of Light*, Oxford University Press (2000).
19. R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals*, Chapter 1, McGraw-Hill (1965).
20. Richard P. Feynman, et. al., *The Feynman Lectures on Physics*, Vol III, Chapter 1, Addison-Wesley (1965).
21. Jonathan L. Cobb, et. al., Proc. SPIE **5037**, 397 (2003)
22. G. M. Gallatin, et. al., J. Vac. Sci. Technol. **B21**, 3172 (2003)
23. F. A. Houle, et. al., J. Vac. Sci. Technol. **B20**, 924 (2002)
24. William Hinsberg, et. al., Proc. SPIE **5039**, 1 (2003)
25. Gerard M. Schmid, et. al., Proc. SPIE **5376**, 333 (2004)
26. J. A. Hoffnagle, et. al., Opt. Letts. **27**, 1776 (2002)
27. William D. Hinsberg, et. al., Proc. SPIE **3999**, 148 (2000)
28. D. Fuard, et. al., Proc. SPIE **5040**, 1536 (2003)
29. Michael D. Shumway, et. al., Proc SPIE **5374**, 454 (2004)
30. Patrick P. Naulleau, Appl. Opt. **43**, 788 (2004)
31. Dario Goldfarb, et. al., J Vac. Sci. Technol. **B22**, 647 (2004)
32. Martha I. Sanchez, et. al., Proc. SPIE **3678**, 160 (1999)
33. Adam R. Pawloski, et. al., Proc. SPIE **5376**, 414 (2004)
34. J. A. Croon, et. al., ESSDERC 2003
35. P. Oldiges, et. al., SISPAD 2000
36. T. Linton, et. al., IEDM 2002
37. Asen Asenov, et. al., IEEE Trans. Electron Dev. **50**, 1254 (2003)
38. Uzodinma Okoroanyanwu and J. H. Lammers, "Resist Road to the 22nm Technology Node", Future Fab. Intl. **17**, Chapter 5, Section 5 (2004)