Stochastic modeling in lithography: use of dynamical scaling in photoresist development

Chris A. Mack
Lithoguru.com
1605 Watchhill Road
Austin, Texas 78703
E-mail: chris@Lithoguru.com

Abstract. The concepts of dynamical scaling in the study of kinetic roughness are applied to the problem of photoresist development. Uniform, open-frame exposure and development of photoresist corresponds to the problem of quenched noise and the etching of random disordered media and is expected to fall in the Kadar-Parisi-Zhang (KPZ) universality class. To verify this expectation, simulations of photoresist development in 1+1 dimensions were carried out with random, uncorrelated noise added to an otherwise uniform development rate. The resulting roughness exponent α and the growth exponent β were found to match the 1+1 KPZ values exactly. © *2009 Society of Photo-Optical Instrumentation Engineers.* [DOI: 10.1117/1.3158612]

Subject terms: dynamical scaling; kinetic roughness; stochastic modeling; photoresist development; line-edge roughness (LER); line-width roughness (LWR).

Paper 09049PR received Mar. 11, 2009; revised manuscript received May 5, 2009; accepted for publication May 7, 2009; published online Jul. 2, 2009. This paper is revision of a paper presented at the SPIE conference on Advances in Resist Materials Processing Technology XXVI, February 2009, San Jose, California. The paper presented there appears (unrefereed) in SPIE Proceedings Vol. 7273.

1 Introduction

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. For example, the very idea of chemical concentration assumes that the volume one is interested in is large enough to contain many, many molecules so that an average number of molecules per unit volume can be used. While we can mathematically discuss the idea of the concentration of some chemical species at a point in space, in reality this concentration must be an average extended over a large-enough region. While in most cases the volumes of interest are large enough not to worry about this distinction, when trying to understand lithography down to the nanometer level, the continuum approximation begins to break down.

When describing lithographic behavior at the nanometer level, an alternative 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.

While stochastic modeling has been successfully applied to photoresist exposure and postexposure bake processes in recent years,[1,2] the stochastic behavior of resist dissolution is much less understood. Dissolution rate variance comes from both the variance in the polymer solubility itself and the resulting variation in the development path required to bypass randomly insoluble polymer molecules. Ultimately, the final result will be a roughness of the resist feature sidewalls that leads to line-edge roughness (LER) and line-width roughness (LWR) of the resist feature. One common approach to studying LER formation is through the use of Monte Carlo simulations.[3,4] While Monte Carlo methods are useful, there is also a need for the development of simple, analytical expressions that capture the essence of the LER formation mechanisms. One approach, which will be employed here, involves the development of scaling relationships as a means for elucidating fundamental mechanisms.[5]

Since the final LER will include all resist and aerial image contributions, studying LER to extract the contribution of resist development can be difficult. A simpler approach is to remove the aerial image from the experiment and study the resist surface roughness after a uniform open-frame exposure and development. Roberts, Fedynyshyn, and coworkers[6,7] used this technique to measure what they called the "innate material roughness" of a photoresist, although the description of the measurement as an "innate" property of the resist is speculative. The use of surface roughness after open-frame exposure and development as a probe for understanding the stochastic nature of resist development will be examined in detail in this paper. In particular, an analysis approach known as *dynamical scaling*

1932-5150/2009/$25.00 © 2009 SPIE

will be applied to photoresist development, and both experimental and simulated results will be analyzed in this way.

2 Background

Over the last 25 years, fractal concepts have been successfully applied to the appearance of roughness in surface growth phenomenon,[8] with many applications, including deposition and etching. Called *disorderly surface growth* or *kinetic roughness*, research has focused on determining how experimental roughness or the roughness predictions of specific models scale with time and distance. While the magnitude of the roughness is unique to the specific experiment or model and its circumstances/parameters, the scaling behavior of the roughness tends to be more universally applicable to a wide range of conditions so long as the basic mechanism remains consistent. This section will review current understanding of kinetic roughness as applied to a simple lithographic case: open-frame exposure and development of photoresist.

2.1 *Self-Affine Surfaces*

A *fractal* surface is by definition *self-similar* (also called *scale invariant*), meaning that it looks the same (is invariant) after a uniform (isotropic) scale transformation. For example, a single-valued function $h(x,y)$ is fractal if scaling by a constant b leads to

$$h(bx,by) \propto bh(x,y). \tag{1}$$

Thus, uniform scaling of the object in all dimensions leads to a self-similar object. For a *deterministic fractal*, Eq. (1) will be true exactly for certain values of b. For a *random* or *statistical fractal*, the scaling produces a function whose statistical properties (mean, standard deviation, etc.) scale as dictated by Eq. (1).

When a surface scales anisotropically, that surface is called *self-affine*. Such a surface may have, for example, the following scaling relationship:

$$h(bx,by) \propto b^{\alpha}h(x,y), \tag{2}$$

where α is called the self-affine exponent (although in our context it will be called the *roughness exponent*, but is also called the *Hurst exponent* and sometimes denoted by H). In other words, the h dimension scales separately from the x-y dimensions. Obviously, the special case of $\alpha=1$ leads to a self-similar fractal. If $\alpha=0$, the surface is completely smooth. An important consequence of the self-affine scaling relationship (2) is that the height-to-height difference $\Delta h=|h(x_1)-h(x_2)|$ for a given separation between points $\Delta x=|x_1-x_2|$ scales as

$$\Delta h \propto \Delta x^{\alpha}. \tag{3}$$

[Note that Eq. (3) applies to a deterministic self-affine surface. Many surfaces are found to be self-affine in the statistical sense, in which case the statistical equivalent of Δh, called the height-to-height correlation, discussed in the following, must be used.]

Consider the following example: open-frame exposure of photoresist coated on a planar wafer leading to, after development, a certain resist height remaining h. Including the effects of roughness, this surface height will be a function of wafer position, $h(x,y)$. If the possibility of overhangs is ignored (see Fig. 1), this surface height will be a single-valued function. Such surfaces are expected to be self-affine, since the nominally vertical direction of the isotropic development will cause the z dimension (the height dimension) of the resist surface to scale differently from the x-y dimensions. More will be said on the validity of this expectation in the following.

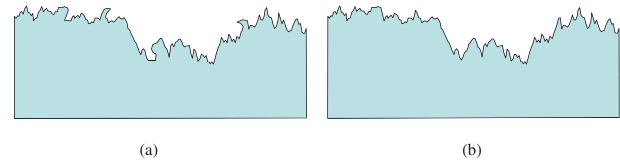

Fig. 1 Two rough surfaces: (a) with overhangs, and (b) with no overhangs so that the surface height $h(x,y)$ is single-valued.

For a statistical self-affine surface, the surface is characterized by statistical properties. For example, the mean height is

$$\langle h\rangle = \frac{1}{M^2}\sum_{i=1}^{M}\sum_{j=1}^{M} h(x_i,y_j), \tag{4}$$

where discrete measurements of height over a square area are assumed. If uniform spacing of $\Delta x=\Delta y$ between measurements is used, the height is averaged over an $L\times L$ area, where $L=M\Delta x$. The RMS surface height difference, often called the *interface width* or the *surface roughness*, is given by

$$\sigma_w = \left\{\frac{1}{M^2}\sum_{i=1}^{M}\sum_{j=1}^{M}[h(x_i,y_j)-\langle h\rangle]^2\right\}^{1/2}. \tag{5}$$

Note that the scaling relationship of Eq. (3), expressed in statistical terms, means that for a self-affine surface,

$$\sigma_w \propto L^{\alpha}. \tag{6}$$

As the size of the region being measured increases, the measured surface roughness (interface width) scales upward according to Eq. (6). As you will see in the following, however, resist surfaces are self-affine only over a region on the order of or less than the correlation length of the roughness. It is very important to note that when the statistics of an interface are measured using a sampling approach [as shown explicitly in Eqs. (5) and (6)], the accuracy of the resulting measures depends on an appropriate choice for Δx.

The dimensionality of the problem can be described in different ways. For the case of open-frame photoresist dissolution, the surface is two-dimensional (2-D), embedded in a three-dimensional (3-D) world. Since the resist height acts as a separate dimension, this problem is often described as a 2+1 dimensional case. In general, a $d+1$ dimensional problem is characterized by a d-dimensional interface. Thus, $d=2$ for the case of open-frame exposure and development of photoresist. (Note that some authors use d to represent the dimensionality of the space, rather than that of the interface surface.)

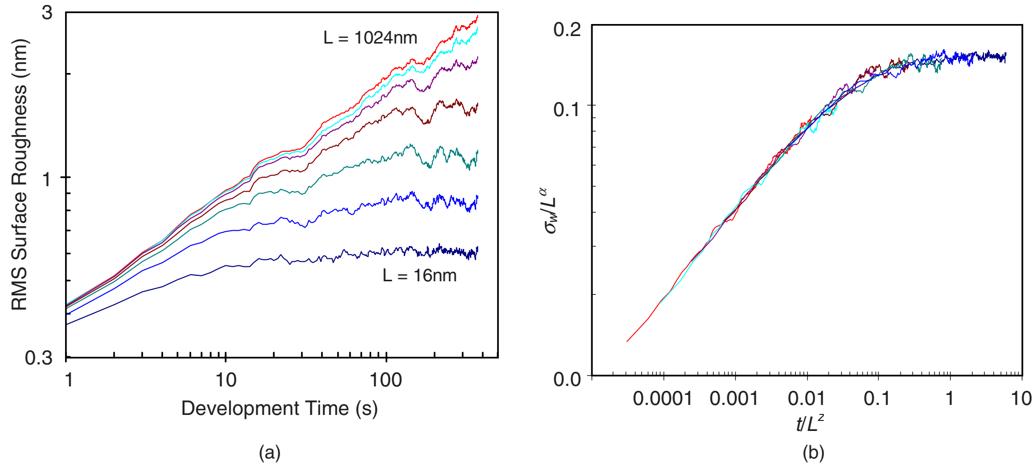

Fig. 2 An example of simulated surface roughness data as a function of development time and L (going from 16 nm to 1024 nm in powers of 2). (a) Raw data, and (b) the same data scaled according to Eq. (9) with $\alpha=0.5$ and $\beta=0.333$.

2.2 *Dynamical Scaling*

Consider again the case mentioned earlier of the development of a uniform, open-frame exposure of photoresist. Initially, the resist surface is perfectly smooth. But as development proceeds, stochastic effects lead to an increased roughening of the surface. Thus, the statistical properties of the interface (the mean height $\langle h\rangle$ and the amount of roughness σ_w) are a function of time. Ignoring absorption so that the resist receives a uniform exposure through its depth (and assuming a perfect antireflection coating so that no standing waves are present), the mean development rate will be constant. Thus, the mean surface height will scale linearly with time:

$$\langle h\rangle \propto t. \tag{7}$$

Empirically, many problems in etching and deposition show a roughness that, for moderately small times, grows as

$$\sigma_w \propto t^{\beta}, \tag{8}$$

where β is sometimes called the *growth exponent*. For extremely small times, as development is just getting started, the growth of roughness may behave somewhat differently (and so very small times are usually excluded from the kind of scaling analysis discussed here).

The growth in roughness as development proceeds does not continue indefinitely. For a given measurement size L, the interface roughness saturates after a long enough time. A typical example is shown in Fig. 2(a). However, since the roughness varies with L according to Eq. (6), the point of saturation with development time depends on the size of the measurement region. The overall scaling can be summarized as[9]

$$\sigma_w \propto L^{\alpha} f\left(\frac{t}{L^z}\right), \tag{9}$$

where

$$f(u)=\begin{cases} u^{\beta} & \text{for small } u \\ 1 & \text{for large } u \end{cases},$$

and $z=\alpha/\beta$ is called the *dynamic exponent*. Again, these scaling relationships are expected to hold for self-affine interfaces. As Fig. 2(b) shows, the proper choice of α and β allows the dynamic roughness data to collapse to a single universal curve for all L.

2.3 *Correlations of the Interface*

For our problem, the source of interface roughness will be a statistical uncertainty in the resist development rate r as a function of position. Thus, the development rate can be separated into a mean dissolution rate plus a random variable η:

$$r(x,y,h)=\langle r\rangle + \eta(x,y,h). \tag{10}$$

Obviously, η has been formulated to have zero mean. Assume for the moment that the underlying noise in development rate that gives rise to surface roughness is spatially uncorrelated. In other words,

$$\langle \eta(x,y,h)\eta(x',y',h')\rangle \propto \sigma_r^2 \delta(x-x')\delta(y-y')\delta(h-h'), \tag{11}$$

where the brackets represent averaging over different realizations of randomness, and σ_r^2 is the variance of the dissolution rate.

It is very important to note that even though the underlying noise is uncorrelated, the resulting rough resist surface will exhibit height-to-height correlations. The cause of these correlations is the isotropic nature of dissolution. If, due to random fluctuations, one point in the resist interface develops down more quickly than the rest, this dimple in the resist surface will begin to spread laterally. Thus, the neighboring points on the interface will have a resist height that is correlated with the original fast-developing point.

For a strictly self-affine surface, the height-to-height correlation function G for the interface will scale in the same way as σ_w:

$$G(\Delta x,\Delta t) = \{\langle[h(x,t) - h(x+\Delta x,t+\Delta t)]^2\rangle\}^{1/2} \propto \Delta x^{\alpha} f\left(\frac{\Delta t}{\Delta x^{z}}\right). \quad (12)$$

The typical distance over which heights interact (that is, over which G is sufficiently greater than zero) is called the *parallel correlation length*, $\xi_{\parallel}$. (There is also a *perpendicular correlation length*, which will not be discussed here.) Initially, this x-y plane correlation length is small, but it grows with time as[8]

$$\xi_{\parallel} \propto t^{1/z}. \quad (13)$$

The correlation length cannot grow to more than the measurement domain, so eventually it saturates at L.

Suppose instead that the underlying development rate noise is correlated. In fact, when the underlying noise is itself the result of a stochastic process, that noise will be correlated. For example, one could assume that the autocorrelation function for the noise exhibits a typical exponential decay with distance

$$\langle \eta(x,y,h)\,\eta(x',y',h')\rangle = \sigma_r^2 \exp[-(s/L_c)^{2\alpha_c}], \quad (14)$$

where $s=[(x-x')^2+(y-y')^2+(h-h')^2]^{1/2}$, L_c is the correlation length of the underlying development rate noise, and α_c is a roughness exponent related to the previous stochastic processes leading to the development rate noise. How will the resulting surface behave? There are two competing mechanisms correlating the final surface, and the final surface behavior will depend on which mechanism dominates. If the correlations of the underlying development rate noise are relatively weak, the correlations induced by the development process itself, as described by Eq. (13) and eventually saturating at a correlation length of L, will dominate. If the development rate noise correlations have a greater influence, the surface will have a correlation length that initially grows like Eq. (13), but eventually saturates at L_c rather than L. Thus, by measuring the dynamical scaling behavior of the roughness during open-frame development, the relative importance of development-induced roughness versus roughness induced by previous stochastic processes (exposure, reaction–diffusion, etc.) can be understood.

2.4 *Power Spectral Density*

Another approach to characterizing the roughness of a surface is to determine its power spectral density (PSD), the magnitude squared of the Fourier transform of the relative surface height:

$$PSD(f) = \lim_{L\to\infty} \frac{1}{L^2}\left|\int_{-L/2}^{L/2}\int_{-L/2}^{L/2} \tilde{h}(x,y)\exp[-i2\pi(f_x x + f_y y)]\mathrm{d}x\,\mathrm{d}y\right|^2, \quad (15)$$

where $f=(f_x^2+f_y^2)^{1/2}$ and $\tilde{h}=h-\langle h\rangle$. The self-affine scaling hypothesis of Eq. (9) can be translated into a scaling relationship for the surface roughness PSD. In the long-time limit and the long-wavelength limit (that is, $L\to\infty$), for a $d-$ dimensional interface, we expect that[10,11]

$$PSD(f) \propto \frac{1}{f^{d+2\alpha}}. \quad (16)$$

2.5 *Stochastic Differential Equations*

The evolution of a resist surface during dissolution can be described by a simple differential equation. Since r is the development rate normal to the resist surface, the rate at which the surface height at a specific (x,y) point changes is given by

$$-\frac{\partial h}{\partial t} = r(1+|\nabla h|^2)^{1/2}, \quad (17)$$

where ∇h is the gradient along the resist surface and represents the maximum slope of the interface at a given (x,y) point. If the surface is essentially horizontal with a small amount of roughness, $|\nabla h|^2 \ll 1$ and this equation can be approximated as

$$-\frac{\partial h}{\partial t} \approx r + \frac{r}{2}|\nabla h|^2. \quad (18)$$

Combining with Eq. (10) and assuming that the noise level is low ($\sigma_r \ll r$),

$$-\frac{\partial h}{\partial t} \approx \langle r\rangle + \frac{\langle r\rangle}{2}|\nabla h|^2 + \eta. \quad (19)$$

This result is a simplification of a common stochastic growth model called the Kadar-Parisi-Zhang (KPZ) equation.[12] In its full form, the KPZ equation is generally written as

$$\frac{\partial \tilde{h}}{\partial t} = \nu\nabla^2\tilde{h} + \frac{\lambda}{2}|\nabla\tilde{h}|^2 + \eta, \quad (20)$$

where $\lambda=\langle r\rangle$, $\tilde{h}=\langle r\rangle t-h$, and ν is a surface tension or diffusion term that relaxes the interface and contributes to smoothing. Lithography simulators assume $\nu=0$, but of course there could be some relaxation mechanism at work in actual development. Here, we will assume that the nonlinear terms dominate (both the $|\nabla h|^2$ and the η terms are nonlinear) and that ν is small.

For high-temperature deposition and etching processes, η in the KPZ equation is dominated by thermal noise, that is, it is a random variable in time. For the case of resist development, however, the noise is a spatial variation in development rate that, for a given resist instantiation, does not vary with time. Kessler, Levine, and Tu described this type of noise as *quenched*.[13] Materials that exhibit quenched noise are called *disordered media* or *random disordered media*. Studies of the KPZ equation with quenched random noise have shown that processes that can be described in this way all exhibit the same scaling exponents and are said to belong to the same *universality class*. In the $d=1$ case, the KPZ exponents can be determined exactly using renormalization group techniques,[8,14,15] to be $\alpha=1/2$ and $\beta=1/3$. For $d=2$, however, an exact determination of the exponents is not possible. For the case where ν is small, where $L \gg \sigma_w$, and where the underlying quenched

noise is uncorrelated, Hentschel and Family[16] have shown that a $d=2$ interface should have $\alpha=0.4$, and $\beta=0.25$, in agreement with simulation results by Kim and Kosterlitz.[17] Other simulation results, however, produced somewhat different exponents.[18] In general, scaling arguments (reinforced by renormalization group methods) show that $\alpha+z=2$ in all dimensions[8] (sometimes referred to as *Galilean invariance*). When this relationship holds,

$$\alpha = \frac{2\beta}{\beta+1} \quad \text{and} \quad \beta = \frac{\alpha}{2-\alpha}. \tag{21}$$

The effects of correlated noise on the scaling exponents is less well understood. The relationship $\alpha+z=2$ is expected to continue to hold, but β is thought to increase as the level of correlation increases.[19,20]

2.6 *Measuring α and β*

Measuring β is straightforward based on its definition but requires the measurement of surface roughness as a function of development time. While straightforward for simulations, it is time-consuming experimentally.

There are several methods for determining α, which ideally should all return the same value[8]:

- **Roughness versus** L: After ensuring that the surface has developed long enough to cross over into the long-time regime, the surface roughness should scale with the measurement domain size L as Eq. (6). A single sample can be used to provide the data, with different measurements using different L (that is, different portions of the developed sample).
- **Height-to-height correlation**: As Eq. (12) shows, the height-to-height correlation will scale with the distance between height measurements to the α power. Thus, from a single sample, a plot of height-to-height correlation versus correlation distance on a log-log scale should produce a straight line of slope α.
- **Full dynamical scaling**: Measure both α and β simultaneously using the full dynamic scaling relationship of Eq. (9). A plot of σ_w/L^α versus t/L^z for many different values of L (all other aspects of the experiment the same) should collapse to a single curve for the correct values of α and β.
- **Power spectrum**: Measure the power spectrum of the interface roughness. The slope of the PSD versus frequency on a log-log plot is $2+2\alpha$.

In all of the preceding methods, one single sample can be used to extract the roughness exponent. However, averaging over many different samples may be necessary to reduce the noise in the measurement. For large L, self-averaging can make the resulting exponent quite precise. As discussed earlier, the expected scaling relations for a self-affine surface apply only for domain sizes at or below the correlation length. Thus, comparing where the roughness saturates to the domain size L will show whether the final surface correlation is dominated by development effects or other correlations.

2.7 *Pinning, Depinning, and Directed Percolation*

In some circumstances, quenched noise can make the surface evolution of a deposition or etch process behave the same as a specific type of percolation problem. Consider a very high contrast resist such that a random distribution of deprotection levels in the resist leads to a certain probability that any given resist molecule will be essentially insoluble. As a simplified picture of the resist during development, each resist molecule is either soluble or insoluble with a certain probability. If the probability that a resist molecule is insoluble is low, the development front will move around any insoluble resist found in its path. As the probability of finding an insoluble resist molecule grows, these molecules will form clusters, the size of which increases with increasing probability of insolubility.

Eventually, a critical probability is reached such that the size of a typical insoluble cluster (the correlation length in the horizontal direction) will span the size of experimental domain (L). When this happens, it becomes highly probably that the developing front will encounter one of these large clusters and become *pinned*, or unable to continue developing. Such critical phenomenon are common in percolation problems. Since the moving interface is *directed*, meaning that it has a preferred direction of motion due to the nature of the dissolution problem, this phenomenon is called *directed percolation*.[21,22]

At the pinning (depinning) transition, the resulting pinned interface will be self-affine with a certain roughness exponent. While strictly speaking the rough surface is only self-affine at the pinning transition, conditions near the transition are often considered to be self-affine as well, so that the common roughness and growth exponents can be applied in this regime. It remains unclear whether resist development, especially development near a resist line edge, will have sufficiently large stochastic differences in dissolution rate to be considered percolative.

2.8 *Experimental Evidence*

There has been very little published data on the dynamic scaling of photoresist surface roughness. Reynolds and Taylor[23] measured the surface roughness of unexposed APEX-E as a function of development time for three different concentration of developers. The surface roughness was measured using an atomic force microscope (AFM) with $L=1.25$ μm (300×300 measurement points) with an uncertainty estimated at $\pm8\%$. Reading off the data from their published Fig. 3(a), the data was replotted here in log-log form in Fig. 3. The best-fit slopes for the 0.195N, 0.21N, and 0.302N developer concentrations are 0.093 ± 0.008, 0.097 ± 0.02, and 0.097 ± 0.02, respectively. (Standard errors based on the least-squares fits are used here for the uncertainties in the slopes, without taking into account the measurement uncertainty.) Thus, within experimental error, β is the same for each developer concentration and equal to about 0.1 ± 0.02.

It is clear from the data of Fig. 3 that the surface roughness has not saturated after the maximum 700-s development time. Thus, only β can be obtained from this data. Since the data are for unexposed resist, one can assume that the underlying development rate noise here was uncorrelated.

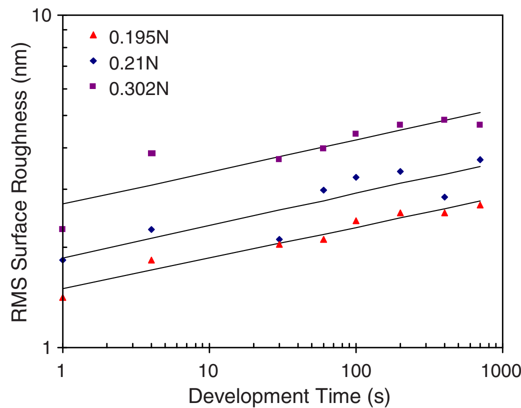

Fig. 3 Surface roughness measured on unexposed APEX-E as a function of development time for three different developers (data taken from Ref. 23).

Table 1 Determination of the growth exponent β based on the dissolution of pure polymers, with data found in Ref. 3.

Polymer	Growth exponent β	Error estimate for β
Poly-B	0.031	0.025
Poly-J	0.097	0.014
Poly-J #2	0.117	0.014
Poly-T	0.114	0.016

Roberts et al.[6] and Fedynyshyn et al.[7] measured the surface roughness after open-frame exposure and development as a function of development time in order to study what they have called "innate material roughness." In their approach, the development time was varied from 5 to 120 s, and RMS surface roughness was measured with an AFM over a 5 μm $\times$ 5 μm area, with a 19.5-nm distance between measurement points. While their plots of the data showed only RMS surface roughness versus mean resist thickness loss, some of their data from Ref. 6 have been replotted in Fig. 4 versus development time.

The authors of Ref. 6 provided the raw data used to generate their Fig. 3, which showed RMS roughness versus thickness loss for several polymers containing no PAG [and thus receiving no exposure or postexposure bake (PEB)]. Several of the polymers had very fast dissolution rates and thus could not be analyzed accurately here. The remaining three polymers were labeled Poly-B, Poly-J, and Poly-T (with chemical structures described in Ref. 6). In addition, Poly-J was measured twice, and both sets of data are shown here in Fig. 4. Fitting the log-log data to a straight line gave the growth exponent β as the best-fit slope. Least-squares fit results are shown in Table 1, with standard error estimates for the slope provided as well.

From the graphs of Fig. 4 and the data in Table 1, it is clear the Poly-B seems to be behaving differently from the other polymers. For Poly-J and Poly-T, the growth exponents are on the order of 0.1, matching the results found for unexposed APEX-E earlier. The second data set of Poly-J (called Poly-J #2) could possibly be leveling out at a saturation level of roughness, but the other data sets do not exhibit a clear saturation. Since the value of L used in this experiment (5 μm) is quite large, one might expect the saturation roughness to occur at a fairly long development time. It is unclear from this limited data whether the polymer Poly-B is anomalous or whether just this data set for Poly-B is anomalous. Since the data are for pure polymers, one can assume that the underlying development rate noise here was uncorrelated.

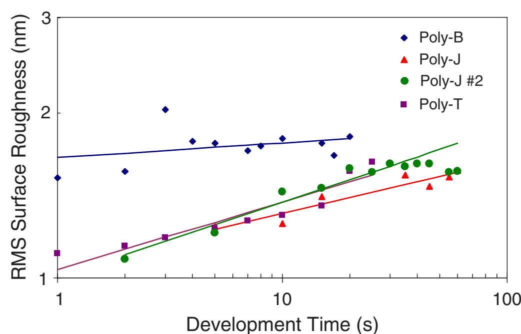

Fig. 4 Surface roughness measured on various polymers (formulated without PAG and receiving no exposure or PEB) as a function of development time (data taken from Ref. 6; Fig. 3).

Unfortunately, the data found in Refs. 6 and 23 were not analyzed to determine the roughness exponent, α. It seems reasonable, however, to claim that none of the data here have leveled off to a saturation level of roughness, so the resulting RMS surface roughness seen at the longest development time is not an innate property of the material.

2.9 *Open Questions*

There are several open questions concerning the application of this scaling approach to resist development and LER prediction. Some of these questions will be addressed in the following simulation study; others will be left to future work.

1. When applying the KPZ equation to open-frame exposure and development of photoresist, how valid is the assumption that $\nabla h \ll 1$? This is equivalent to asking whether overhangs on the resist surface are significant. Some work in this area[24,25] has shown that both α and β decline as the number of overhangs becomes significant.
2. Do the scaling exponents change with dose/average development rate? Do they change with the magnitude of the noise term?
3. How do correlations in the development rate noise affect the scaling exponents?
4. How does one apply the lessons learned from the constant $\langle r \rangle$ case (open-frame exposure and development) to the lithographically important case of surface roughness in the presence of large development rate gradients?
5. Does photoresist development in the low development regime exhibit directed percolation behavior?

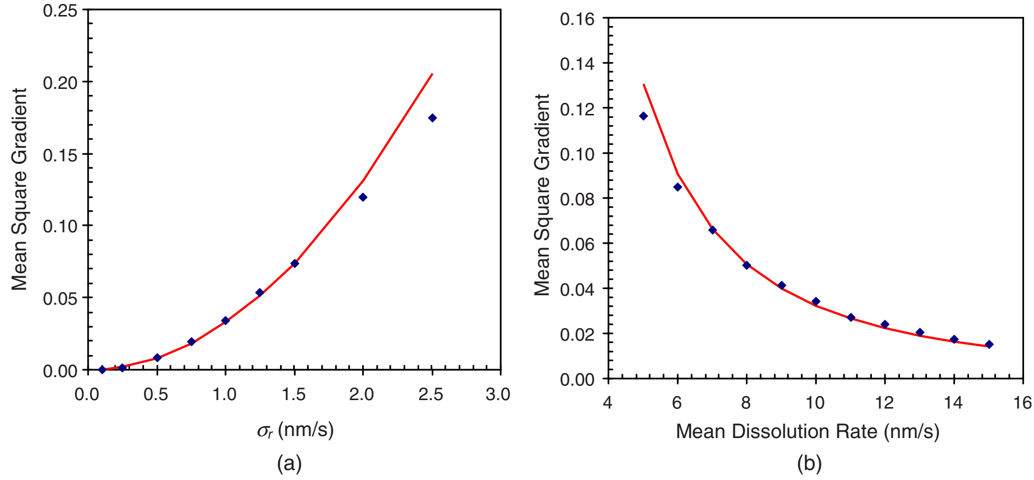

Fig. 5 Plots of the mean square gradient, $\langle |\nabla h|^2 \rangle$, as a function of (a) the dissolution rate standard deviation and (b) the mean dissolution rate of the simulations. Data points are simulation results, solid lines are from the empirical expression of Eq. (23).

3 Simulations of Photoresist Development

Simulation was used to predict the resist height as a function of development time for an open-frame exposure/development in the presence of stochastic dissolution rate noise. The dissolution rate of the photoresist followed Eq. (10), with a mean $\langle r \rangle$ and stochastic term η, set to be an uncorrelated, uniformly distributed random variable of standard deviation σ_r. To simplify the work, only the $d=1$ case was modeled, with a maximum $L=1024$ nm and a resist thickness of 500 nm or 4000 nm for the two following studies. The simulation grid was 1 nm in both x and z. The PROLITH v12.0 development algorithm was used to turn the $r(x,z)$ data into a resist surface as a function of development time. (PROLITH is a commercial lithography simulator available from KLA-Tencor.)

Development time was varied from 1 s to the time needed to clear the resist in increments of 1 s. At each development time, the mean and standard deviation of the resist height $h(x)$ was determined. L was varied by taking the 1024 nm width and breaking it up into two 512-nm regions, four 256-nm regions, etc., down to 16-nm-size regions. When multiple regions were cut from the one simulation, the analysis results were averaged.

In the first study, a resist thickness of 4000 nm was used, with $\langle r \rangle = 10$ nm/s and $\sigma_r = 2$ nm/s. Eight separate simulation trials were run, and the results averaged. These averaged results are shown in Fig. 2(a). By setting the roughness and growth exponents to their expected KPZ $d=1$ values ($\alpha=1/2$, $\beta=1/3$), the data collapse into one universal curve, as shown in Fig. 2(b). This is compelling evidence that the PROLITH model of the photoresist etching mechanism follows the KPZ universality class, as expected for the case of uncorrelated noise. By fitting the $L=1024$-nm data to a straight line on a log-log scale, the growth exponent was found. For the time interval from 1 to 20 s, $\beta=0.337\pm0.004$, and from 21 to 100 s, $\beta=0.336\pm0.004$. From 101 to 380 s, $\beta=0.266\pm0.004$, indicating that the curve is beginning to level off slightly. By fitting the saturation regions of the $L=16$ through $L=128$ nm cases to Eq. (6), the roughness exponent was found to be $\alpha=0.461\pm0.006$.

In a second study, the impact of $\langle r \rangle$ and σ_r on the mean resist height was examined. According to Eq. (19), the rate at which the mean resist height decreases with time is slightly greater than the mean dissolution rate due to the nonzero resist surface gradient:

$$-\frac{\partial \langle h \rangle}{\partial t} = \langle r \rangle + \frac{\langle r \rangle}{2}\langle |\nabla h|^2 \rangle \quad \text{so that} \quad \frac{1}{2}\langle |\nabla h|^2 \rangle = -\frac{1}{\langle r \rangle}\frac{\partial \langle h \rangle}{\partial t} - 1. \tag{22}$$

By measuring the slope of the mean resist height versus the development time curve from the simulations, the mean square resist surface gradient can be determined.

To study this effect, the resist thickness of the simulation was changed to 500 nm (to speed up the simulations), and the mean square gradient of the resist surface was calculated from the simulation results using Eq. (22). For $\langle r \rangle = 10$ nm/s, σ_r was varied from 0.25 nm/s to 2.5 nm/s. Also, setting $\sigma_r = 1$ nm/s, $\langle r \rangle$ was varied from 5 to 15 nm/s. The results are shown as the data points in Figs. 5(a) and 5(b). An interesting trend arises from this data. When the relative development rate variation, $\sigma_r/\langle r \rangle$, is small (less than about 0.15), the mean squared surface gradient scales as

$$\langle |\nabla h|^2 \rangle \approx 3.2\left(\frac{\sigma_r}{\langle r \rangle}\right)^2. \tag{23}$$

For larger $\sigma_r/\langle r \rangle$, the mean surface gradient grows less slowly than Eq. (23). This empirical model is shown in Fig. 5 as the solid curve.

It is clear from these results that the assumption that $|\nabla h|^2 \ll 1$, used in deriving the KPZ equation, is reasonable so long as the relative dissolution rate variation, $\sigma_r/\langle r \rangle$, remains small. In fact, the deviation of the simulation re-

sults seen in Fig. 5 from the empirical expression (23) roughly corresponds to the point where the simulated surface roughness begins to deviate from the assumptions of the KPZ derivation, at $|\nabla h|^2 \sim 0.1$. It should be noted that over the range of conditions simulated here, no overhangs in the resulting resist surface were observed.

4 Conclusions

From the work shown here, it is clear that the concepts of dynamical scaling can be (and should be) employed to study the growth of surface roughness during photoresist development. All evidence suggests that resist surfaces are self-affine and are governed by the scaling laws commonly used in the study of kinetic roughness. This scaling is characterized by two basic exponents, the roughness exponent α and the growth exponent β. The simple experiment of open-frame (uniform) exposure and development followed by measurement of the mean thickness and RMS surface roughness is a powerful tool for extracting these exponents and confirming the applicability of dynamical scaling.

What little experimental evidence has been published to date confirms the standard dynamical scaling, at least for the case of uncorrelated development rate noise. Simulation in $1+1$ dimensions also shows very nicely that the expected scaling behavior is followed almost perfectly (Fig. 2). Additionally, and not unexpectedly, these $1+1$ simulations match the KPZ universality class almost perfectly, giving $\alpha = 1/2$ and $\beta = 1/3$. Further, simulations have shown a linear dependence of the mean surface gradient on the relative development rate variation for small gradients and provided a numerical estimate of when the assumptions in the derivation of the KPZ equation might break down.

Much future work is needed on this topic. Simulations should be extended to $2+1$ dimensions, with the use of Gaussian rather than uniformly distributed noise as well. Further, the use of correlated development rate noise should show whether kinetic roughness during development dominates the final results or whether the underlying development rate noise, coming from earlier stochastic processes such as exposure and reaction–diffusion, controls the final surface characteristics. Last, a thorough understanding of the uniform, open-frame development case will lead the way to the more difficult and interesting case of roughness formation during development in the presence of a steep development rate gradient, as is found at the edge of a photoresist line.

Acknowledgments

I would like to thank Ted Fedynyshyn of MIT Lincoln Laboratories and Jeanette Roberts of Intel for sharing data from Ref. 6 with me and allowing me to analyze and display that data here. Also, thanks to Mark D. Smith and KLA-Tencor for providing a modified version of PROLITH that allowed the import of development rate data, thus enabling the simulations shown in this paper.

References

1. G. Gallatin, "Resist blur and line edge roughness," in *Optical Microlithography XVIII*, Bruce W. Smith, Ed., *Proc. SPIE* **5754**, 38–52 (2004).
2. Chris A. Mack, *Fundamental Principles of Optical Lithography: The Science of Microfabrication*, John Wiley & Sons, London (2007).
3. T. Mulders, W. Henke, K. Elian, C. Nolscher, and M. Sebald, "New stochastic post-exposure bake simulation method," *J. Microlithogr., Microfabr., Microsyst.* **4**, 043010 (2005): A. Philippou, T. Mülders, and 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**, 043005 (2007).
4. A. Saeki, T. Kozawa, S. Tagawa, H. Cao, H. Deng, and 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**, 043004 (2007).
5. V. Constantoudis, G. Patsis, and E. Gogolides, "Photoresist line-edge roughness analysis using scaling concepts," *J. Microlithogr., Microfabr., Microsyst.* **3** 429–435 (2004).
6. J. Roberts, R. Meagley, T. Fedynyshyn, R. Sinta, D. Astolfi, R. Goodman, and A. Cabral, "Contributions to innate material roughness in resist," in *Advances in Resist Technology and Processing XXIII*, Qinghuang Lin, Ed., *Proc. SPIE* **6153**, 61533U (2006).
7. T. Fedynyshyn, R. Sinta, D. Astolfi, A. Cabral, J. Roberts, and R. Meagley, "Deconstructing the resist to probe innate material roughness," in *Advances in Resist Technology and Processing XXIII*, Qinghuang Lin, Ed., *Proc. SPIE* **6153**, 615315 (2006).
8. A.-L. Barabasi and H. Stanley, *Fractal Concepts in Surface Growth*, Cambridge University Press, Cambridge (1995).
9. F. Family and T. Vicsek, "Scaling of the active zone in Eden process on percolation networks and the ballistic deposition model," *J. Phys. A* **18**, L75–81 (1985).
10. M. Plischke, Z. Racz, and D. Liu, "Time-reversal invariance and universality of two-dimensional growth models," *Phys. Rev. B* **38**(7), 3485–3495 (1987).
11. D. Liu and M. Plischke, "Universality in two- and three-dimensional growth and deposition models," *Phys. Rev. B* **35**(7), 4781–4787 (1988).
12. M. Kadar, G. Parisi, and Y.-C. Zhang, "Dynamic scaling of growing interfaces," *Phys. Rev. Lett.* **56**(9), 889–892 (1986).
13. D. Kessler, H. Levine, and Y. Tu, "Interface fluctuations in random media," *Phys. Rev. A* **43**(8), 4551–4554 (1991).
14. K. Wilson and J. Kogut, "The renormalization group and the ε expansion," *Phys. Rep.* 12(2) 75–199 (1974).
15. M. Fisher, "The renormalization group in the theory of critical behavior," *Rev. Mod. Phys.* **46**, 597–616 (1974).
16. H. Hentschel and F. Family, "Scaling in open dissipative systems," *Phys. Rev. Lett.* **66**(15), 1982–1985 (1991).
17. J. Kim and J. Kosterlitz, "Growth in a restricted solid-on-solid model," *Phys. Rev. Lett.* **62**(19), 2289–2292 (1989).
18. S. Buldyrev, S. Havlin, J. Kertesz, A. Shehter, and E. Stanley, "Surface roughening with quenched disorder in *d*-dimensions," *Fractals* **1**(4), 827–839 (1993).
19. M. Li, "Surface growth with spatially correlated noise," *Phys. Rev. E* **55**(1), 1178–1180 (1997).
20. E. Katzav and M. Schwartz, "Self-consistent expansion for the Kardar-Parisi-Zhang equation with correlated noise," *Phys. Rev. E* **60**(5), 5677–5680 (1999).
21. L.-H. Tang and H. Leschhorn, "Pinning by directed percolation," *Phys. Rev. A* **45**(12), R8309–R8312 (1992).
22. S. Havlin, L. Amaral, S. Buldyrev, S. Harrington, and H. Stanley, "Dynamics of surface roughening with quenched disorder," *Phys. Rev. Lett.* **74**(21), 4205–4208 (1995).
23. G. Reynolds and J. Taylor, "Factors contributing to sidewall roughness in a positive-tone, chemically amplified resist exposed by x-ray lithography," *J. Vac. Sci. Technol. B* **17**(2), 334–344 (1999).
24. A. Maritan, F. Toigo, J. Koplik, and J. Banavar, "Dynamics of growing interfaces," *Phys. Rev. Lett.* **69**(22), 3193–3195 (1992).
25. D. Ko and F. Seno, "Simulations of deposition growth models in various dimensions: the possible importance of overhangs," *Phys. Rev. E* **50**(3), R1741–R1744 (1994).

Chris A. Mack developed the lithography simulation software PROLITH and founded and ran the company FINLE Technologies for 10 years. He then served as vice president of lithography technology for KLA-Tencor for 5 years, until 2005. In 2003, he received the SEMI Award for North America for his efforts in lithography simulation and education, and in 2009, he received the SPIE Frits Zernike Award for Microlithography. He is also an adjunct faculty member at the University of Texas at Austin. Currently, he writes, teaches, and consults in the field of semiconductor microlithography in Austin.