Journal of
Micro/Nanolithography,
MEMS, and MOEMS

SPIEDigitalLibrary.org/jm3

Defining and measuring development rates for a stochastic resist: a simulation study

Chris A. Mack

SPIE

Defining and measuring development rates for a stochastic resist: a simulation study

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

Abstract. Photoresist development rate can be defined microscopically (the development rate at a point) or macroscopically (the propagation rate of an average resist height). In the presence of stochastic noise, these two rates will be different. In order to properly calibrate lithography simulators, the difference between these two definitions of development rate should be quantified. Using theoretical derivations and a stochastic (Monte Carlo) development simulator, the propagation rate of a resist surface is characterized in the presence of stochastic variation in the resist deprotection concentration and a nonlinear development rate response. The resulting propagation rate (macroscopic development rate) can be more than an order of magnitude higher than for the case of no stochastic noise. Correlation in the development rate creates an effective surface inhibition over a depth into the resist proportional to the correlation length, with results that are qualitatively different for two-dimensional versus three-dimensional simulations. The differences between microscopic and macroscopic dissolution rate can have an important effect on how development rate models should be calibrated, depending on their use in continuum or stochastic lithography simulators. [DOI: 10.1117/1.JMM.12.3.033006]

Subject terms: development rate; stochastic resist; lithography simulation; line-edge roughness; line width roughness; correlation length.

Paper 13006P received Feb. 11, 2013; revised manuscript received Jun. 5, 2013; accepted for publication Jun. 17, 2013; published online Jul. 17, 2013.

1 Introduction

Photoresist development rates are commonly measured to characterize photoresist dissolution behavior. Such data are frequently used to calibrate development models for simulators. An important though frequently unstated assumption of this use is the equivalence of microscopic and macroscopic development rates. Microscopic development rate is the development rate at a point in the resist and is the rate used by simulators [Fig. 1(a)]. Macroscopic development rate is the mean rate at which a large open area of resist develops down [Fig. 1(b)] and is the quantity measured when development rates are measured, e.g., when using a dissolution rate monitor. In the absence of stochastic effects that result in surface roughness, these two rates are identical.

In the real, stochastic case, the mean propagation rate of a large open area is a strong function of the stochastic uncertainty of the development rate, especially in regions of moderately low dissolution rates. This paper will explore the impact of stochastic uncertainty in microscopic development rate, including the effects of correlations, on the macroscopic development rate through the use of both analytical derivations and Monte Carlo simulations. Note that this general problem, the difference between the local and global front propagation rate in the presence of noise, is encountered in many fields.[1–3]

First, the probability distribution for development rates will be derived. Then, the various cases of macroscopic development [one-dimensional (1-D), two-dimensional (2-D), and three-dimensional (3-D), with and without correlations] will be studied both analytically and with simulations. Finally, simulations will show that the disconnect between microscopic and macroscopic development rates can be significant, affecting the accuracy of development model calibration.

2 Theory

Dissolution rate uncertainty will inevitably result from uncertainty in the underlying inhibitor concentration (e.g., the concentration of protecting groups in a chemically amplified resist). Consider a simple development rate function,[4]

$$r = r_{\max}\frac{(a+1)(1-m)^n}{a+(1-m)^n} + r_{\min},$$
$$a = \frac{(n+1)}{(n-1)}(1-m_{\text{th}})^n, \tag{1}$$

where r is the development rate, m is the relative inhibitor (protecting group) concentration, and $r_{\max}$, $r_{\min}$, n, and m_{th} are model parameters. Here, we will neglect $r_{\min}$ as small compared to the development rate in the region of interest. The edge of a photoresist feature will necessarily have a protection level that is near the knee of the development rate curve, so that $m > m_{\text{th}}$. Thus, if $n \gg 1$ (which is the case for a modern, high-contrast photoresist), the development rate in this region will be well approximated by

$$r \approx r'_{\max}(1-m)^n, \quad r'_{\max} = r_{\max}\frac{a+1}{a}. \tag{2}$$

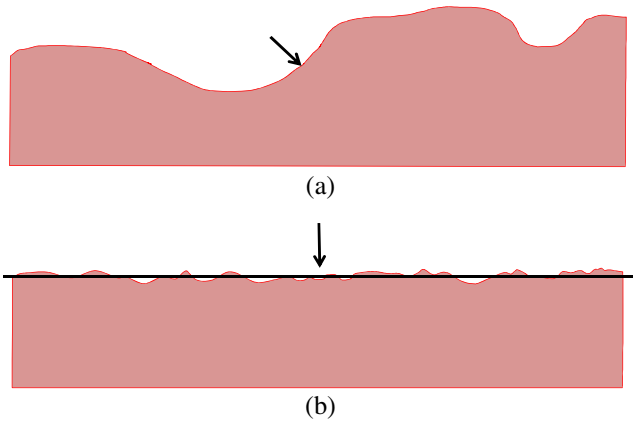


Fig. 1 Photoresist development rate is defined in two ways: (a) microscopic development—the rate at which a point on the resist surface moves perpendicular to that surface, and (b) macroscopic development—the rate at which the mean surface position (the mean development front) propagates for a large open-frame exposure.

This development rate expression will be used throughout this paper. While simple, it accurately reflects the nonlinear development rate response for the exposure and deprotection levels expected near a photoresist feature edge.

Suppose that m is a random variable with a normal distribution, so that $m \sim \mathrm{N}(\mu, \sigma_m)$. If the randomness of m is the only source of uncertainty in the resulting development rate, a probability distribution function (pdf) for r can be calculated

$$\mathrm{pdf}_r = \left|\frac{\mathrm{d}m}{\mathrm{d}r}\right| \mathrm{pdf}_m, \quad \left|\frac{\mathrm{d}m}{\mathrm{d}r}\right| = \frac{1}{nr}\left(\frac{r}{r'_{\max}}\right)^{\frac{1}{n}}, \tag{3}$$

$$\mathrm{pdf}_r = \frac{1}{\sqrt{2\pi}\sigma_m}\frac{1}{nr}\left(\frac{r}{r'_{\max}}\right)^{\frac{1}{n}} e^{-[1-\mu-(r/r'_{\max})^{1/n}]^2/2\sigma_m^2}. \tag{4}$$

Unfortunately, the moments of this distribution (in particular, the mean and the variance) cannot be analytically calculated, making the utility of this pdf expression questionable. The mode, however, does have an analytical form:

$$r_{\mathrm{mode}} = r'_{\max}(1 - m_{\mathrm{mode}})^n,$$
$$1 - m_{\mathrm{mode}} = \frac{1-\mu}{2}\left[1 + \sqrt{1 - \frac{4(n-1)\sigma_m^2}{(1-\mu)^2}}\right]. \tag{5}$$

This expression gives real roots so long as $\sigma_m < (1-\mu)/(2\sqrt{n-1})$.

While not always applicable, the case of a small amount of noise in the underlying inhibitor concentration will allow some useful approximations. For small σ_m (i.e., for $\sigma_m \ll 1-\mu$),

$$1 - m_{\mathrm{mode}} \approx (1-\mu)\left[1 - (n-1)\left(\frac{\sigma_m}{1-\mu}\right)^2\right] \tag{6}$$

so that

$$r_{\mathrm{mode}} \approx r(\mu)\left[1 - (n-1)\left(\frac{\sigma_m}{1-\mu}\right)^2\right]^n. \tag{7}$$

Again for the case of small σ_m, the actual pdf of Eq. (4) is well approximated by a generalized gamma distribution (GGD):[5]

$$\mathrm{pdf}_{\mathrm{GGD}} = \frac{\beta}{\lambda\Gamma(\alpha)}\left(\frac{r}{\lambda}\right)^{\alpha\beta-1} e^{-(r/\lambda)^\beta}, \tag{8}$$

which has the following properties:

$$\mathrm{mode} = \lambda\left(\alpha - \frac{1}{\beta}\right)^{\frac{1}{\beta}}, \quad \langle r\rangle = \lambda\frac{\Gamma\left(\alpha + \frac{1}{\beta}\right)}{\Gamma(\alpha)},$$
$$\langle r^2\rangle = \lambda^2\frac{\Gamma\left(\alpha + \frac{2}{\beta}\right)}{\Gamma(\alpha)}, \tag{9}$$

where $\langle \ldots \rangle$ is the average over many instances. Matching the mode of the GGD to the mode of r, we find that

$$\lambda = r_{\mathrm{mode}}\left(\alpha - \frac{1}{\beta}\right)^{-\frac{1}{\beta}}. \tag{10}$$

The GGD can then be conveniently calculated by defining

$$\gamma \equiv \left(\frac{r}{r_{\mathrm{mode}}}\right)^\beta\left(1 - \frac{1}{\alpha\beta}\right) \tag{11}$$

so that

$$\mathrm{pdf}_{\mathrm{GGD}}(r) = \left(\frac{\beta}{r}\right)\left(\frac{\alpha^\alpha e^{-\alpha}}{\Gamma(\alpha)}\right) e^{\alpha(1-\gamma+\ln\gamma)}. \tag{12}$$

As we shall see, the assumption of small σ_m is equivalent to saying $\alpha \gg 1$, so that Sterling's approximation to the gamma function can be used

$$\Gamma(\alpha) = \alpha^\alpha e^{-\alpha}\sqrt{\frac{2\pi}{\alpha}}\left(1 + \frac{1}{12\alpha} + \frac{1}{288\alpha^2} + \ldots\right)$$
$$\approx \alpha^\alpha e^{-\alpha}\sqrt{\frac{2\pi}{\alpha}} \tag{13}$$

giving, for $r > 0$,

$$\mathrm{pdf}_{\mathrm{GGD}}(r) \approx \left(\frac{1}{r}\right)\left(\frac{\beta\sqrt{\alpha}}{\sqrt{2\pi}}\right) e^{\alpha(1-\gamma+\ln\gamma)}. \tag{14}$$

The values of α and β can now be determined by empirically fitting the GGD to numerical evaluations of Eq. (4). Excellent fits are obtained when

$$\alpha = \left(\frac{1-\mu}{3\sigma_m}\right)^2, \quad \beta = \frac{3}{n}. \tag{15}$$

This leads to an approximate pdf for r of

$$\mathrm{pdf}(r) \approx \frac{1}{\sqrt{2\pi}}\left(\frac{1}{r}\right)\left(\frac{1-\mu}{n\sigma_m}\right) e^{\alpha(1-\gamma+\ln\gamma)}, \tag{16}$$

where

$$\gamma = \left(\frac{1-m}{1-m_{\text{mode}}}\right)^3 \left[1 - 3n\left(\frac{\sigma_m}{1-\mu}\right)^2\right] \approx \left(\frac{1-m}{1-\mu}\right)^3 \left[1 - 3\left(\frac{\sigma_m}{1-\mu}\right)^2\right]. \quad (17)$$

Using the small σ_m approximation for γ [the right-hand side of Eq. (17)] requires a small adjustment to the scaling for the pdf, giving

$$\text{pdf}_{\text{approx}}(r) = \frac{1}{2.54}\left(\frac{1}{r}\right)\left(\frac{1-\mu}{n\sigma_m}\right) e^{\alpha(1-\gamma+\ln\gamma)}. \quad (18)$$

Note that γ, and thus the exponential term in this pdf, is not dependent on the development rate function parameters for the case of small σ_m. A plot of Eq. (18) compared to Eq. (4) is shown in Fig. 2 for the case of $n = 10$ (a common value for state-of-the-art photoresists today). Similar fits are obtained over a range of n from 2 to 15, though the match is not as good for larger n.

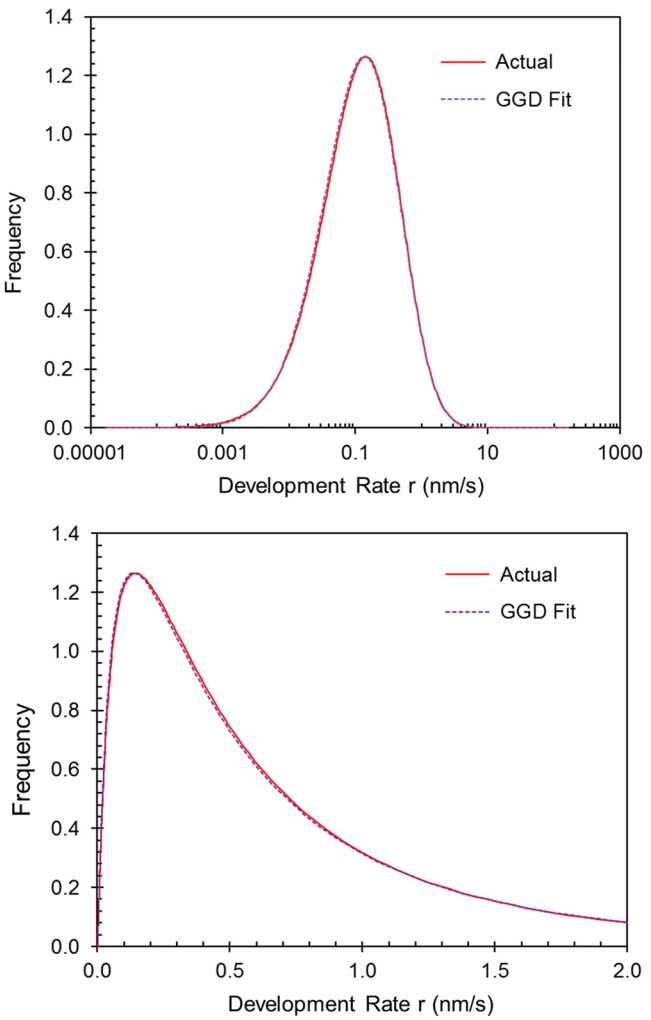


Fig. 2 Comparison of the numerical calculation for the pdf_r (labeled as "Actual") to the approximate generalized gamma distribution (GGD) form of Eq. (18), for $r_{\text{max}} = 200$ nm/s, $m_{\text{th}} = 0.5$, $n = 10$, $\mu = 0.73$, and $\sigma_m = 0.03$. The two graphs differ only in the use of linear and log scales for the x-axis.

The advantage of using the GGD is the ability to calculate the mean and variance of the development rate analytically. Applying Sterling's approximation to the gamma functions in Eq. (9)

$$1 + \left(\frac{\sigma_r}{\langle r\rangle}\right)^2 = \frac{\Gamma(\alpha)\Gamma\left(\alpha+\frac{2}{\beta}\right)}{\left[\Gamma\left(\alpha+\frac{1}{\beta}\right)\right]^2} \approx \frac{\left(1+\frac{2}{\alpha\beta}\right)^{\alpha+\frac{2}{\beta}-\frac{1}{2}}}{\left(1+\frac{1}{\alpha\beta}\right)^{2\alpha+\frac{2}{\beta}-1}}. \quad (19)$$

For $\alpha\beta \gg 1$,

$$\left(\frac{\sigma_r}{\langle r\rangle}\right)^2 \approx \left(\frac{n\sigma_m}{1-\mu}\right)^2 + \frac{1}{2}\left(1-\frac{3}{n}\right)^2\left(\frac{n\sigma_m}{1-\mu}\right)^4 + O\left(\frac{n\sigma_m}{1-\mu}\right)^6. \quad (20)$$

The mean value of the development rate becomes

$$\langle r\rangle = \lambda\frac{\Gamma\left(\alpha+\frac{1}{\beta}\right)}{\Gamma(\alpha)} \approx r_{\text{mode}} e^{-1/\beta}\left(1-\frac{1}{\alpha\beta}\right)^{-\frac{1}{\beta}}\left(1+\frac{1}{\alpha\beta}\right)^{\alpha+\frac{1}{\beta}-\frac{1}{2}}. \quad (21)$$

For $\alpha\beta \gg 1$,

$$\frac{\langle r\rangle}{r(\mu)} \approx 1 + \frac{1}{2}\left(1-\frac{1}{n}\right)\left(\frac{n\sigma_m}{1-\mu}\right)^2 + O\left(\frac{n\sigma_m}{1-\mu}\right)^4. \quad (22)$$

The accuracy of these approximations will be tested below with simulations, along with simulations of the development front propagation rate (i.e., the macroscopic development rate).

3 1-D Development Front Propagation (Macroscopic Development) Rate

The propagation of the mean resist surface height for an open-frame exposure is called the average development front propagation rate (r_{prop}) or the macroscopic development rate. For the 1-D case, this quantity can be derived analytically since the path of dissolution is obviously known. For a development rate as a function of depth into the resist $r(z)$, development to a depth D occurs over a development time t_{dev} according to

$$t_{\text{dev}} = \int_0^D \frac{dz}{r}. \quad (23)$$

For a development rate that only exhibits random variation (no systematic variations), and for a sufficiently large D,

$$\left\langle\frac{1}{r}\right\rangle \approx \frac{1}{D}\int_0^D \left(\frac{1}{r}\right) dz. \quad (24)$$

Since r_{prop} can be approximated as D/t_{dev}, we see by comparing Eqs. (23) and (24) that

$$r_{\text{prop}} = \frac{1}{\left\langle \frac{1}{r} \right\rangle}. \tag{25}$$

The mean of the reciprocal development rate can be derived from the approximate pdf of Eq. (18), giving

$$\left\langle \frac{1}{r} \right\rangle = \frac{\Gamma\left(\alpha - \frac{1}{\beta}\right)}{\lambda \Gamma(\alpha)}. \tag{26}$$

Thus,

$$\frac{r_{\text{prop}}}{\langle r \rangle} = \frac{1}{\langle r \rangle \left\langle \frac{1}{r} \right\rangle} = \frac{[\Gamma(\alpha)]^2}{\Gamma\left(\alpha + \frac{1}{\beta}\right)\Gamma\left(\alpha - \frac{1}{\beta}\right)}. \tag{27}$$

Applying, as before, Sterling's approximation to the gamma functions

$$\frac{r_{\text{prop}}}{\langle r \rangle} \approx \left(1 + \frac{1}{\alpha\beta}\right)^{\frac{1}{2}-\alpha-\frac{1}{\beta}} \left(1 - \frac{1}{\alpha\beta}\right)^{\frac{1}{2}-\alpha+\frac{1}{\beta}}. \tag{28}$$

For $\alpha\beta \gg 1$,

$$-\ln\left(\frac{r_{\text{prop}}}{\langle r \rangle}\right) \approx \left(\frac{n\sigma_m}{1-\mu}\right)^2 + \frac{9}{2n^2}\left(\frac{n\sigma_m}{1-\mu}\right)^4 + O\left(\frac{n\sigma_m}{1-\mu}\right)^6 \tag{29}$$

and

$$-\ln\left(\frac{r_{\text{prop}}}{r(\mu)}\right) \approx \frac{1}{2}\left(1 + \frac{1}{n}\right)\left(\frac{n\sigma_m}{1-\mu}\right)^2 + O\left(\frac{n\sigma_m}{1-\mu}\right)^4. \tag{30}$$

The results above show that for 1-D development the impact of underlying development randomness is to decrease the macroscopic development rate compared to the case of development without noise. Since the development path is constrained to a straight line, the time required to develop down is dominated by the slowest developing bits of resist. For the cases of 2-D and 3-D development, the results are qualitatively different. Since the path of development is not dimensionally constrained, the path can go around insoluble bits of resist.[6] To explore the impact of development rate noise on the macroscopic development rate in 2-D or 3-D, Monte Carlo simulations will be employed.

4 2-D and 3-D Monte Carlo Simulations: Uncorrelated

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.[7,8] A standard fast marching level set method[9,10] was used to convert a regular grid of development rates into a resist surface as a function of development time, using a linear interpolation of development rates between grid points. The development model discussed above in Eq. (2) was used, with $m \sim \mathrm{N}(\mu, \sigma_m)$ selected using a random number generator for each grid point. The grid size was set to 1 nm. For 2-D (1 + 1) simulations, the simulation width was 4096 grid elements and the resist thickness was 4096 grids. The development time was adjusted in each case so that ~2000 time steps would allow the front to reach the bottom of the resist. For 3-D simulations (2 + 1), the widths in x and y were 512 grids, the resist thickness was 512 or 1024 grids, and ~500 or 1000 times steps were used. The development model parameters were $r_{\max} = 200$ nm/s, $m_{\text{th}} = 0.5$, and n was varied between 5 and 15 (typical parameters for common resist materials). At each time step, a resist surface was extracted and the average resist height was calculated.

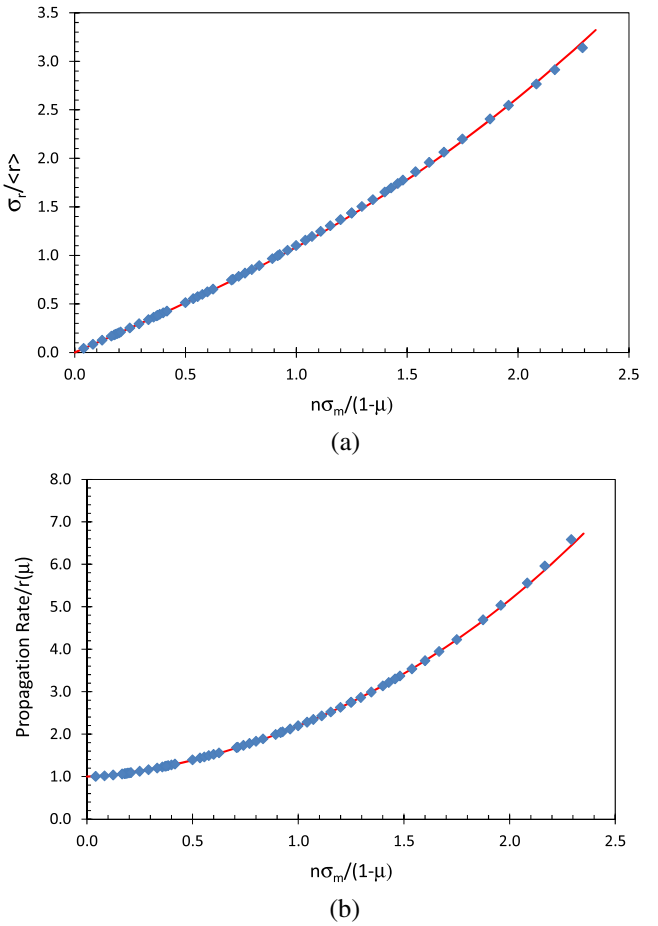


Fig. 3 Two-dimensional (2-D) simulation results for $n = 10$. Each data point is the average of 100 trials. μ was varied between 0.70 and 0.79 and σ_m was varied between 0.005 and 0.050. (a) Development rate noise and (b) relative macroscopic development rate.

For each combination of μ and σ_m evaluated, the mean and standard deviation of the microscopic development rate were calculated, and the development front propagation rate was determined by fitting the average resist surface height versus time to a straight line. An example of the 2-D simulation results, for $n = 10$, is shown in Fig. 3. An example of 3-D simulation results is shown in Fig. 4. Interestingly, all results follow a single curve as a function of one parameter, $n\sigma_m/(1-\mu)$.

The simulation results were fit to semi-empirical expressions (guided by the results of the theory section above). Equation (20) predicts an approximately linear variation of the relative development rate uncertainty with relative deprotection uncertainty, but for higher amounts of uncertainty a higher-order term is required. For 2-D and 3-D simulations and for all values of n, simulation results follow

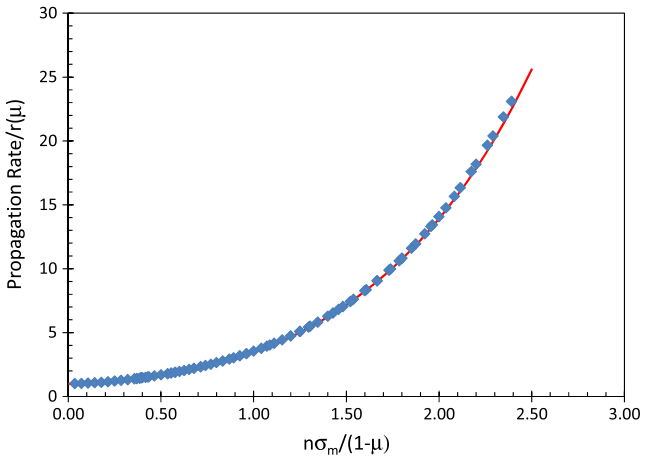


Fig. 4 Three-dimensional (3-D) simulation results for $n = 10$. Each data point is the average between 4 and 50 trials. μ was varied between 0.72 and 0.77 and σ_m was varied between 0.001 and 0.055. The resulting $r(\mu)$ went from 0.08 to 1.92 nm/s and the propagation rate varied from 0.1 to 6.7 nm/s.

$$\left(\frac{\sigma_r}{\langle r\rangle}\right)^2 \approx \left(\frac{n\sigma_m}{1-\mu}\right)^2 + \left(0.50 - \frac{2.7}{n}\right)\left(\frac{n\sigma_m}{1-\mu}\right)^4. \tag{31}$$

The front propagation rate, r_{prop}, was also fit to empirical equations.

$$\text{2-D: } r_{\text{prop}} \asymp r(\mu)\left[1 + \left(\frac{n\sigma_m}{1-\mu}\right)^{1.5} + \left(0.27 - \frac{1}{n}\right)\left(\frac{n\sigma_m}{1-\mu}\right)^3\right], \tag{32}$$

$$\text{3-D: } r_{\text{prop}} \asymp r(\mu)\left[1 + 2.2\left(\frac{n\sigma_m}{1-\mu}\right)^2 + (0.04n - 0.188)\left(\frac{n\sigma_m}{1-\mu}\right)^4 + (0.0016n - 0.008)\left(\frac{n\sigma_m}{1-\mu}\right)^6\right] \tag{33}$$

In the above expressions, the propagation rate was expressed relative to the development rate at the mean deprotection concentration, $r(\mu)$. It is also interesting to see how the front propagation rate compares to the mean development rate, $\langle r\rangle$. In a previous study that assumed an uncorrelated $r \sim N(\langle r\rangle, \sigma_r)$, the empirical results were[5]

$$\text{2-D: } r_{\text{prop}} \approx \langle r\rangle\left[1 + 0.93\left(\frac{\sigma_r}{\langle r\rangle}\right)^{1.67}\right], \tag{34}$$

$$\text{3-D: } r_{\text{prop}} \approx \langle r\rangle\left[1 + 1.47\left(\frac{\sigma_r}{\langle r\rangle}\right)^{1.67}\right]. \tag{35}$$

Thus, for Gaussian development rate noise, increasing the development rate uncertainty always resulted in an increase in the macroscopic development rate (over the range of noise studied, which kept $\sigma_r/\langle r\rangle$<0.3).

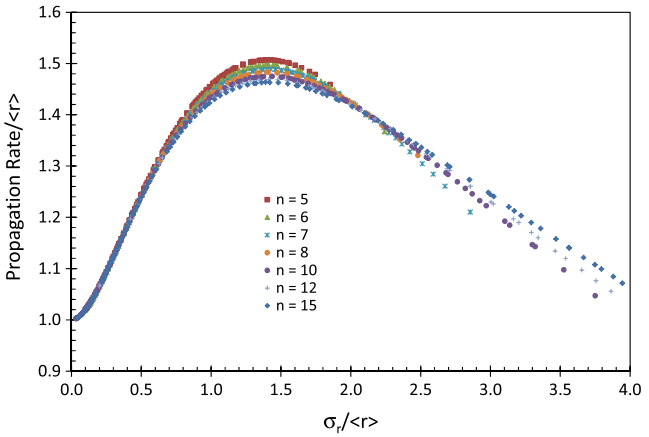


Fig. 5 2-D simulation results for $n = 5$ to 15 for GGD noise. Each data point is the average of 100 trials.

For development noise that follows the highly skewed GGD, the results were extended to higher noise levels to reveal a different regime. Figure 5 shows that a plot of $r_{\text{prop}}/\langle r\rangle$ as a function of $\sigma_r/\langle r\rangle$ produces a curve almost independent of n. Further, there are two apparent regimes. For low $\sigma_r/\langle r\rangle$ (less than ~1.5), the macroscopic development rate grows with increasing development rate noise, just as for the case of Gaussian noise. But for higher levels of noise the macroscopic development rate actually slows down (as it always does for the 1-D case).

Empirically, the 2-D data of Fig. 5 follows this expression:

$$r_{\text{prop}} \simeq \langle r\rangle\left[1 + a\left(\frac{\sigma_r}{\langle r\rangle}\right)^{1.55}\right]\left[\frac{s_0^b}{\left(\frac{\sigma_r}{\langle r\rangle}\right)^b + s_0^b}\right] \tag{36}$$

with $a = 1$, $s_0 = 1.42 + 0.23/n$ and $b = 2.13 + 0.48/n$. This model, along with the data for $n = 10$, is shown in Fig. 6.

For 3-D simulations, we see the same basic trend as for 2-D, but the peak propagation rate occurs at a much higher noise level (greater than 4). The same Eq. (36) matches the 3-D data using $a = 1.88$–$1.1/n$, $s_0 = 1.83 + 2.4/n$ and $b = 1.77 + 0.67/n$. The fact that the macroscopic

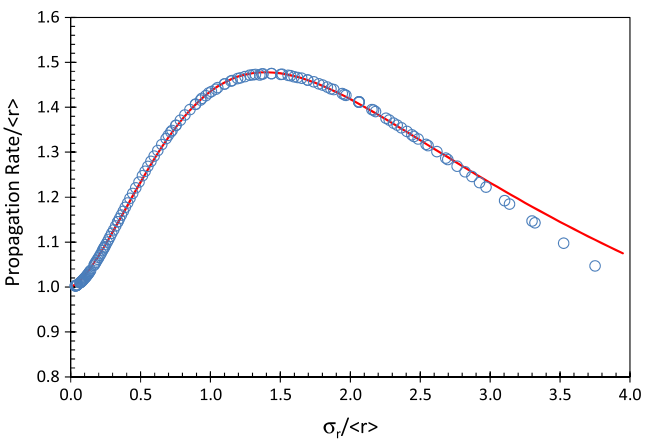


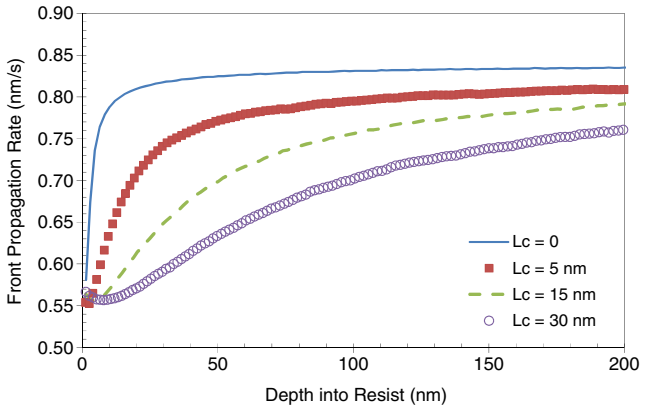
Fig. 6 2-D simulation results for $n = 10$ (open circles) and the model (solid line) of Eq. (36).

development rate slows down relative to the mean at a much higher level of noise reflects the increased opportunities for the development paths to avoid slow-dissolving bits of resist in 3-D compared to 2-D.

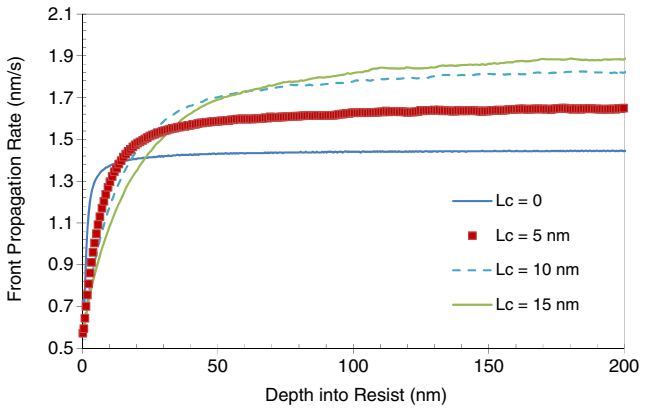
5 Monte Carlo Simulations: Correlated

Simulations using correlated noise employed the same parameters as described above for uncorrelated simulations. To generate the correlated rate noise, the previously derived reaction–diffusion PSD[11] and the numerical method proposed by Thorsos[12] were used. For each combination of μ and σ_m evaluated, the development front propagation rate was determined by fitting the average resist surface height versus time with a straight line, excluding the top half of the photoresist. The reason for excluding the top portion of the resist can be seen by looking at the development rate as a function of depth into the resist (Fig. 7).

In all cases, the dissolution front begins propagating at $\langle r \rangle$ (equal to 0.572 nm/s in this case), then grows to a steady-state propagation rate (about 0.835 nm/s in the 2-D case). For uncorrelated m, the front propagation rate grows quickly, reaching halfway between $\langle r \rangle$ and the steady-state rate in about 3 nm (for these 2-D simulations). When the underlying inhibitor concentration is correlated, the front propagation rate holds about steady until a resist thickness equal to



(a)



(b)

Fig. 7 Simulation results of front propagation rate as a function of depth into the resist for $n = 10$, $\mu = 0.73$, $\sigma_m = 0.03$ (giving $\sigma_r/\langle r \rangle = 1.25$ in 2-D) and varying deprotection correlation lengths (labeled as Lc) for (a) 2-D (average of 8000 trials) and (b) 3-D (average of 100 trials).

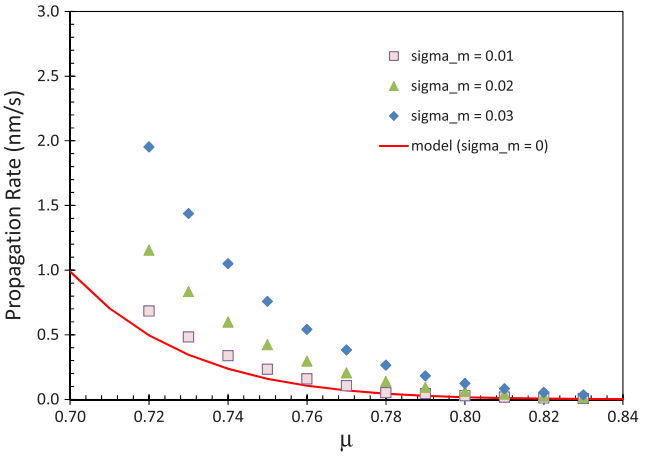


Fig. 8 Comparison of 3-D simulation results of the development front propagation rate (for $n = 10$) for different σ_m (uncorrelated noise). The "model" curve corresponds to $\sigma_m = 0$.

the correlation length is consumed. Then, the propagation rate rises more slowly, reaching halfway between $\langle r \rangle$ and the steady-state rate at a depth equal to about 3.5 times the correlation length. Note that the final propagation rate is about independent of the correlation length for the 2-D case.

Interestingly, the 3-D case shows far less surface inhibition due to correlated noise than the 2-D case (reaching halfway between $\langle r \rangle$ and the steady-state rate at a depth about equal to one correlation length, rather than 3.5). Further, the final, steady-state propagation rate increases significantly with increasing correlation length. It is clear that 2-D simulations of stochastic development may not be a good proxy for a 3-D reality, even qualitatively. It is also clear that for sufficiently thin resists (with thickness on the order of the correlation length) a "bulk" behavior of the development rate will not be reached.

6 Discussion and Conclusions

Lithography simulators, whether making the continuum approximation or performing stochastically, require a model relating the dissolution rate at a point in the resist to the level of deprotection at that point. The data to calibrate such a model universally comes from measurements of film thickness versus development time for large open-frame exposures. As the simulations performed in this work show, the average development front propagation rate, as might be measured by a dissolution rate monitor, can be many times larger than the development rate at the mean deprotection concentration when stochastic variations in deprotection rate are present. Figure 4 shows propagation rates >20 times higher than the expected rate. A summary of this effect is shown in Fig. 8, where the impact of stochastic noise is indistinguishable from an increase in both m_{th} and r_{min} (Fig. 9).

The reason for this behavior is the dramatically skewed probability distribution for development rate that arises from a normally distributed deprotection level when the dissolution rate dependence is highly nonlinear (i.e., for a high value of n). The results presented above give rise to two concerns. When using dissolution rate data to calibrate a continuum model, differences in the variance of the deprotection level will affect the data and thus, the model fit (Fig. 9). If the

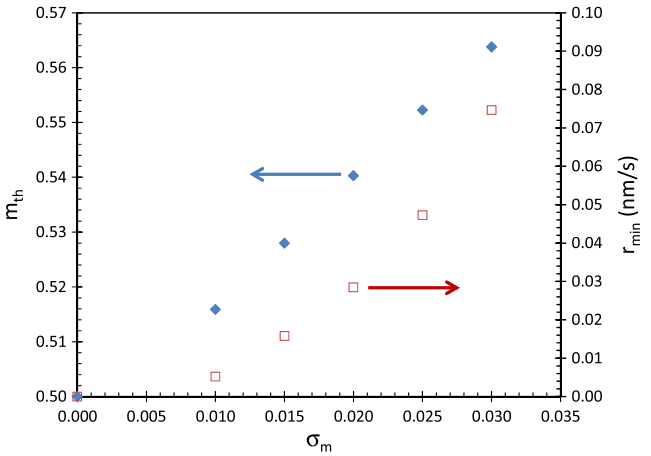


Fig. 9 The data from Fig. 8 matches the development rate model of Eq. (1) if m_{th} (diamonds) and r_{min} (squares) are allowed to increase as a function of stochastic noise.

continuum simulator is to be used for cases where the variance in deprotection level matches that of the dissolution rate measurement experiment, all is well. But if not, there will be some concern as to the faithfulness of the simulation results. When using dissolution rate data to calibrate a stochastic model, the calibration process should include stochastic effects as well. Otherwise, a continuum model calibration process could greatly overestimate the dissolution rate values to be used in a stochastic simulator near the knee of the development curve.

Further, both the qualitative and quantitative differences between the results of 2-D and 3-D simulations of stochastic development indicate that the common practice of using simpler and faster 2-D simulations as a proxy for a 3-D reality may be suspect. Full 3-D simulations will likely be required to understand even qualitatively the impact of stochastic development on lithographic results.

Finally, a grid size of 1 nm was used throughout this paper. An alternate interpretation of the grid, however, would be to consider each 3-D grid element as a resist polymer molecule. Typical resists employ polymers that occupy volumes on the order of 10 nm^3. Thus, in this interpretation a correlation length of 5 would not be 5 nm, but would be a correlation that reaches across five polymer molecules. All of the results presented in this paper could also be interpreted in this way.

References

1. J. Armero et al., "External fluctuations in front propagation," *Phys. Rev. Lett.* **76**(17), 3045–3048 (1996).
2. J. Armero et al., "Ballistic and diffusive corrections to front propagation in the presence of multiplicative noise," *Phys. Rev. E* **58**(5), 5494–5500 (1998).
3. E. Brunet and B. Derrida, "Effect of microscopic noise on front propagation," *J. Stat. Phys.* **103**(1–2) 269–282 (2001).
4. C. A. Mack, *Fundamental Principles of Optical Lithography*, p. 260, John Wiley & Sons, London (2007).
5. E. W. Stacy, "A generalization of the gamma distribution," *Ann. Math. Stat.* **33**(3), 1187–1192 (1962).
6. C. A. Mack, "Stochastic approach to modeling photoresist development," *J. Vac. Sci. Technol.* **B27**(3), 1122–1128 (2009).
7. C. A. Mack, "Stochastic modeling in lithography: the use of dynamical scaling in photoresist development," *J. Micro/Nanolithogr. MEMS MOEMS* **8**(3), 033001 (2009).
8. C. A. Mack, "Stochastic modeling of photoresist development in two and three dimensions," *J. Micro/Nanolithogr. MEMS MOEMS* **9**(4), 041202 (2010).
9. J. A. Sethian, *Level Set Methods and Fast Marching Methods*, 2nd ed., Chapter 8, Cambridge University Press, Cambridge (1999).
10. J. A. Sethian, "Fast marching level set methods for three-dimensional photolithographic development," *Proc. SPIE* **2726**, 262–272 (1996).
11. C. A. Mack, "Reaction–diffusion power spectral density," *J. Micro/Nanolithogr. MEMS MOEMS* **11**(4), 043007 (2012).
12. E. I. Thorsos, "The validity of the Kirchhoff approximation for rough surface scattering using a Gaussian roughness spectrum," *J. Accoust. Soc. Am.* **83**(1), 78–92 (1988).

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 a fellow of SPIE and IEEE and is also an adjunct faculty member at the University of Texas at Austin. In 2012, he became editor-in-chief of the *Journal of Micro/Nanolithography, MEMS, and MOEMS* (JM3). Currently, he writes, teaches, and consults on the field of semiconductor microlithography in Austin, Texas.