Reaction-diffusion power spectral density

Chris A. Mack
Reaction-diffusion power spectral density

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

1 Introduction

Stochastic variations in photoresist exposure, reaction-diffusion, and development ultimately lead to sidewall surface roughness of the final photoresist features, described as line-edge roughness (LER) or linewidth roughness (LWR). Full characterization of the LER requires its description over the spectrum of roughness frequency components, most commonly using the power spectral density (PSD). While a comprehensive model for LER formation in lithography does not yet exist, several pieces of the puzzle have been explored. For example, models for the prediction of the uncertainty in acid concentration after exposure of a chemically amplified resist have been developed for 193-nm lithography.

Full characterization of the LER requires its description over the spectrum of roughness frequency components, most commonly using the power spectral density (PSD). While a comprehensive model for LER formation in lithography does not yet exist, several pieces of the puzzle have been explored. For example, models for the prediction of the uncertainty in acid concentration after exposure of a chemically amplified resist have been developed for 193-nm lithography.

2 Reaction-Diffusion PSD—the Effective Acid Concentration

The random distribution of acid after exposure of a chemically amplified resist is essentially uncorrelated (and thus its PSD will be flat—white noise) when ignoring such a correlating mechanism as speckle during 193-nm exposure, or secondary electron generation for 13.5-nm exposure. Reaction-diffusion, where the acid diffuses and acts as a catalyst for a polymer deprotection reaction, adds correlation to the resulting uncertainty. One acid can deprotect multiple polymer sites; as a result, those deprotection reactions are not independent. In a previous study, the autocorrelation function of the effective acid concentration during PEB was derived.

\[
\tilde{R}_{\text{eff}}(\tau) = \frac{\int_{-\infty}^{\infty} R_{\text{PSF}}(s)R_{\text{PSF}}(s+\tau)ds}{\int_{-\infty}^{\infty} |R_{\text{PSF}}(s)|^2 ds}. 
\]

where \( \tau \) is the radial distance from the origin, and the reaction-diffusion point spread function \( R_{\text{PSF}} \) has analytical forms in one, two, and three dimensions.

\[
1-D: R_{\text{PSF}}(x) = \frac{1}{\sqrt{2\pi \sigma_x^2}} e^{-\frac{x^2}{2\sigma_x^2}} \text{erfc} \left( \frac{|x|}{\sqrt{2\sigma_x}} \right), \\
2-D: R_{\text{PSF}}(r) = -\frac{1}{2\pi \sigma_y^2} \text{Ei} \left( -\frac{r^2}{2\sigma_y^2} \right), r = \sqrt{x^2 + y^2}, \\
3-D: R_{\text{PSF}}(r) = \frac{1}{2\pi \sigma_x^2} \left[ \frac{\sigma_y^2 e^{-\frac{r^2}{2\sigma_y^2}}}{\sqrt{2\sigma_y^2}} \right], r = \sqrt{x^2 + y^2 + z^2} 
\]

The effective acid concentration \( H_{\text{eff}} \) is the time average of the acid concentration during the bake and is the convolution of the initial acid concentration after exposure with the dissolution/etching algorithm that in turn will predict the final surface roughness (the subject of future work).
reaction-diffusion point spread function.\(^7\) Note that the PSD is simply the Fourier transform of the autocorrelation function, and in general, the resulting PSD will have different forms in different dimensions.\(^8\)

While the integrals in Eq. (1) cannot be carried out analytically in any dimension, the Fourier transform of Eq. (1) can, yielding the PSD.

\[
\text{PSD}(f) = \text{PSD}(0) \left(1 - \frac{e^{-\left(\frac{2\pi f}{\xi}\right)^2}}{\left(\frac{2\pi f}{\xi}\right)^2}\right)^2, \tag{3}
\]

where the correlation length-like parameter \(\xi\) is related to the acid diffusion length \(\sigma_a\) by \(\xi = \sqrt{2}\sigma_a\). The zero frequency PSD is calculated from Parseval’s theorem.

\[
1 - D: \quad \text{PSD}(0) = \frac{3\sqrt{2\pi\sigma_a^3}}{8\sqrt{2 - 1}} \approx 1.60466\sigma_a^2H_{\text{diff}},
\]

\[
2 - D: \quad \text{PSD}(0) = \frac{\pi\sigma_a^2\sigma_{\text{diff}}^2}{2\ln 2} \approx 2.26618\xi^2\sigma_{\text{diff}}^2, \tag{4}
\]

\[
3 - D: \quad \text{PSD}(0) = \frac{\pi^{3/2}\sigma_a^3}{4(2 - \sqrt{2})} \approx 2.37643\xi^3\sigma_{\text{diff}}^2.
\]

In the high frequency limit, where \(\pi\xi f \gg 1\), we see that PSD \(\approx 1/f^4\), corresponding to a Hurst exponent of \(\alpha = 0.5\) for the 3-D case.

Interestingly, the reaction-diffusion PSD is the same, to within a scale factor, for one, two, and three dimensions. This is not generally assumed to be the case. For example, it has been common to assume that a stretched exponential autocorrelation function can apply.\(^9,10\)

\[
\hat{R}(s) = \sigma^2 e^{-\left(s/\xi\right)^\alpha}. \tag{5}
\]

For \(\alpha = 0.5\), the resulting PSD can be analytically derived.\(^8\) Letting \(d\) be the dimensionality of the problem,

\[
\text{PSD}_d(f) = \frac{a_1\sigma^2 e^{2\pi f\xi_1}}{[1 + (2\pi f\xi_1)^{2\alpha + 2}]}, \tag{6}
\]

where \(a_1 = 2, a_2 = 2\pi,\) and \(a_3 = 8\pi.\) For the 3-D case, this expression produces the high frequency \(1/f^4\) behavior found for the reaction-diffusion PSD. For 1-D and 2-D problems, however, Eq. (6) predicts \(1/f^2\) and \(1/f^3\) behavior, respectively, while the reaction-diffusion PSD of Eq. (3) predicts a high frequency \(1/f^4\) behavior for all dimensions.

Comparing the reaction-diffusion PSD of Eq. (3) to the 3-D version of Eq. (6), the frequency behavior can be made to match in the low-frequency regime, or in the high-frequency regime, but not both. Matching at low frequencies gives

\[
\text{PSD}_d(0) = 8\pi\sigma^2\xi_3 = \text{PSD}_{\text{RD}}(0) = \frac{\pi^{3/2}\xi^3\sigma_{\text{RD}}^2}{4(2 - \sqrt{2})},
\]

\[
\xi_3 = \xi \left(\frac{\sqrt{\pi}}{32(2 - \sqrt{2})}\right) \approx 0.4556\xi. \tag{7}
\]

Matching at high frequencies gives

\[
\xi = \sqrt{\frac{2(2 - \sqrt{2})}{\sqrt{\pi}}} \approx 0.6610\xi. \tag{8}
\]

In either case, the mid-frequency regime will not match, as seen in Fig. 1. The differences, while not dramatic, are noticeable.

3 Correlation Length

There is no universally accepted definition of correlation length (\(L_{\text{corr}}\)) that can pin down its value except to within a multiplicative factor. One definition, for example, is based on the frequency that reduces the PSD by a factor of 2

\[
\text{PSD} \left( f = \frac{1}{2\pi L_{\text{corr}}} \right) = \frac{\text{PSD}(0)}{2}. \tag{9}
\]

By this definition, the correlation length of the reaction-diffusion system is related to \(\xi\) by \(L_{\text{corr}} = 0.5819\xi = 0.8229\sigma_a\) (for 1-D, 2-D, and 3-D). Another common definition, proposed by Stratonovich,\(^11\) is the area under the autocorrelation function curve:

\[
\int_0^{L_{\text{corr}}} R(s) \, ds = \frac{1}{2\pi L_{\text{corr}}}. \tag{10}
\]

\[
\int_0^{L_{\text{corr}}} \sigma^2 e^{-\left(s/\xi\right)^\alpha} \, ds = \frac{\sigma^2}{2\pi L_{\text{corr}}}. \tag{11}
\]

\[
\xi = \sqrt{\frac{2(2 - \sqrt{2})}{\sqrt{\pi}}} \approx 0.6610\xi. \tag{12}
\]

In either case, the mid-frequency regime will not match, as seen in Fig. 1. The differences, while not dramatic, are noticeable.

where the autocorrelation function is assumed to be an even function of the distance s. The second integral can be simplified by changing the order of integration, giving a result that depends on the dimensionality of the problem

\[ L_{\text{corr}} = \int_0^\infty R ds = \frac{1}{\sigma^2} \int_0^\infty F^{-1}\{\text{PSD}\} ds, \quad (10) \]

Carrying out these calculations for the reaction-diffusion PSD,

\[ 1 - D: \quad L_{\text{corr}} = \frac{3\sqrt{\pi}}{16(\sqrt{2} - 1)} \xi \approx 0.80233 \xi, \]
\[ 2 - D: \quad L_{\text{corr}} = \frac{4\sqrt{\pi}(\sqrt{2} - 1)}{3 \ln 2} \xi \approx 1.41225 \xi, \]
\[ 3 - D: \quad L_{\text{corr}} = \frac{\sqrt{\pi} \ln 2}{2(2 - \sqrt{2})} \xi \approx 1.04865 \xi. \quad (12) \]

Thus, depending on the definition used and the dimensionality of the problem, the correlation length is between 0.58\( \xi \) and 1.41\( \xi \). It seems reasonable, therefore, to define the correlation length for the reaction-diffusion system as being synonymous with the parameter \( \xi \).

### 4 Reaction-diffusion PSD—the Protecting Group Concentration

The relationship between the effective acid concentration and the relative concentration of protecting groups in the chemically amplified resist (\( m \)) is exponential, making calculation of the PSD for \( m \) difficult.

\[ m = e^{-\alpha_t h_{\text{eff}}}, \quad (13) \]

where \( \alpha_t \) is the amplification factor (with typical values between 1 and 2), equal to the amplification rate constant times the PEB time, and \( h_{\text{eff}} \) is the effective acid concentration relative to the initial photoacid generator concentration.\(^7\) When the effective relative acid concentration near the resist line-edge is sufficiently small that a linear approximation to this exponential relationship is reasonable, the relative protecting group concentration will have a PSD of the same form as Eqs. (3) and (4), where the variance of the effective acid concentration is replaced by the variance of the relative protecting group concentration, \( \sigma_{m}^2 \).

For the more general case, numerical simulations can be used to explore the shape of the PSD of the relative protecting group concentration. Given a normally distributed random variable \( h_{\text{eff}} \sim N(\mu_h, \sigma_h) \) and a desired PSD, how does one generate a random volume of \( h_{\text{eff}}(x, y, z) \) on a regular grid? While there are several methods available, I prefer the approach proposed by Thorsos.\(^1^2\) The goal is to create a grid of random numbers with a Gaussian distribution and with spatial correlations that would produce, on average, a given PSD. Thorsos described the algorithm in one dimension, which will be reproduced briefly here. Given \( N_x \) grid points with spacing \( \Delta x \) covering a distance \( L_x = N_x \Delta x \), the relative effective acid concentration at the point \( x_n = n \Delta x \) is given by

\[ h_{\text{eff}}(x_n) = \mu_h + \frac{1}{L_x} \sum_{j=-N_x/2}^{N_x/2-1} F(f_j)e^{i2\pi f_j x_n}, \quad (14) \]

where this calculation is performed as the Fast Fourier Transform (FFT) of \( F \) on a grid of frequencies \( f_j = j/L_x \). The function \( F \), in turn, is calculated from the amplitude of the PSD (for \( j \geq 0 \))

\[ F(f_j) = \sqrt{L_x \text{PSD}(f_j)} \left\{ \frac{(\eta_1 + i\eta_2)/\sqrt{2}, \quad j \neq 0, N_x/2}{\eta_1, \quad j = 0, N_x/2} \right\}, \quad (15) \]

where \( \eta_1 \) and \( \eta_2 \) are two independent \( N(0,1) \) random numbers. Since \( h_{\text{eff}}(x_n) \) must be real, the negative frequencies of \( F \) are obtained from a symmetry relationship: \( F(f_j) = F^\ast(f_{-j}) \).

The Thorsos algorithm can easily be extended to two and three dimensions, so long as care is taken to properly produce the boundary conditions (a purely real random number is used at the origin and at the outer edges of the volume) and the symmetry to result in a purely real \( h_{\text{eff}} \). In two dimensions, this requires

\[ \text{Re}\{F(f_x, f_y)\} = \text{Re}\{F(-f_x, -f_y)\}, \]
\[ \text{Re}\{F(-f_x, f_y)\} = \text{Re}\{F(f_x, -f_y)\}, \]
\[ \text{Im}\{F(f_x, f_y)\} = -\text{Im}\{F(-f_x, -f_y)\}, \]
\[ \text{Im}\{F(-f_x, f_y)\} = -\text{Im}\{F(f_x, -f_y)\}. \quad (16) \]

In three dimensions,

\[ \text{Re}\{F(f_x, f_y, f_z)\} = \text{Re}\{F(-f_x, -f_y, -f_z)\}, \]
\[ \text{Re}\{F(-f_x, f_y, f_z)\} = \text{Re}\{F(f_x, -f_y, -f_z)\}, \]
\[ \text{Re}\{F(f_x, -f_y, f_z)\} = \text{Re}\{F(-f_x, f_y, -f_z)\}, \]
\[ \text{Im}\{F(f_x, f_y, f_z)\} = -\text{Im}\{F(-f_x, -f_y, -f_z)\}, \]
\[ \text{Im}\{F(-f_x, f_y, f_z)\} = -\text{Im}\{F(f_x, -f_y, -f_z)\}, \]
\[ \text{Im}\{F(f_x, -f_y, f_z)\} = -\text{Im}\{F(-f_x, f_y, -f_z)\}. \quad (17) \]

The random \( m \) is produced by putting the random value of \( h_{\text{eff}} \) for each grid point into Eq. (13). Then, the resulting volume of random \( m \) was analyzed by extracting its PSD. Since the result is inherently spherically symmetric [dictated by the symmetry of Eq. (3)], the PSD as a function of the three spatial frequency dimensions was interpolated onto one radial-direction grid. This provides an added benefit of significant averaging for high spatial frequencies (though none for the lowest frequency). Further averaging is obtained...
by repeating the numerical simulation numerous times and averaging the resulting PSDs.

For a Gaussian distribution of \( h_{\text{eff}} \), the distribution of \( m \) will be log-normal, with mean and standard deviation given by

\[
\langle m \rangle = e^{\alpha f \mu m} e^{\sigma_m^2/2}, \quad \sigma_m^2 = \langle m \rangle^2 (e^{\sigma_m^2} - 1).
\]

(18)

The shape of the PSD of \( m \) was explored in the 2-D case for \( \sigma_m = 0.05 \) and varying \( \langle m \rangle \), using \( \alpha_f = 1.5 \). The resulting numerical PSDs (the average of 2000 trials for each set of parameters) matched Eqs. (3) and (4) extremely well so long as \( \mu_h \) was not too close to 0 or 1. In particular, a good match to the PSD expression occurred when

\[
\mu_h - 3\sigma_h > 0 \quad \text{and} \quad \mu_h + 4\sigma_h < 1.
\]

(19)

Note that these are also the approximate constraints on describing the distribution of \( h_{\text{eff}} \) as Gaussian. Also note that under these conditions and for \( \sigma_m = 0.05 \) or less, the log-Normal distribution for \( m \) is approximately Normal in shape.

## 5 Development Rate PSD

Dissolution rate uncertainty will inevitably result from uncertainty in the underlying inhibitor concentration. Consider a simple development rate function

\[
r = r_{\text{max}} \frac{(a + 1)(1 - m)^n}{a + (1 - m)^n} + r_{\text{min}},
\]

(20)

where \( r \) is the development rate, and \( r_{\text{max}}, r_{\text{min}}, n, \) and \( m_{\text{th}} \) are model parameters. Here, we will neglect \( r_{\text{min}} \) 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 \( n \gg 1 \), the development rate in this region will be well approximated by

\[
r \approx r_{\text{max}}^* (1 - m)^n, \quad r_{\text{max}}^* = r_{\text{max}} \frac{a + 1}{a}.
\]

(21)

This development rate expression will be used below. While simple, it accurately reflects the non-linear development rate response for the exposure and deprotection levels expected near a photoresist feature edge.

Using the Thorsos method described above, a Gaussian distribution of \( m \) with a PSD given by Eqs. (3) and (4) was used in the development rate Eq. (21) to create a random volume of development rate. An example development rate PSD is shown in Fig. 2 (\( r_{\text{max}} = 200 \text{ nm/s}, \ m_{\text{th}} = 0.5, \) and \( n = 10 \)).

The shapes of the resultant development rate PSDs are similar to, but not exactly the same as, the PSD function of Eq. (3). By fitting Eq. (3) to the numerical results, the impact of the highly nonlinear development rate function is seen to be a small decrease in the correlation length as a function of the development nonlinearity, \( n \), multiplied by the relative noise in the protecting group concentration, \( \sigma_m/(1 - \langle m \rangle) \). Figure 3 shows these results, along with a quadratic fit to the data.

## 6 Conclusions

The result of the derivations presented in this paper is a simple, new expression for the PSD of the effective acid concentration for a chemically amplified resist. Under conditions of moderately small variations in the protecting group concentration, the PSD of the protecting group concentration also follows this same PSD expression. Finally, near the photoresist line-edge the resulting development rate will also follow this PSD expression (approximately), though with a slightly smaller correlation length. The decrease in the development rate correlation length with increasing development rate noise is a new and unexpected result.

These results can usefully be employed in a comprehensive stochastic model of lithography to predict LER. Previous work to describe the time evolution of rough surfaces
during development has assumed uncorrelated development rate noise. 

By applying the analytical PSD expressions derived in this paper, future work to simulate the time evolution of developed rough surfaces should show whether kinetic surface roughness caused by dissolution itself dominates the final lithographic roughness or whether the underlying development rate noise, coming from the earlier stochastic reaction-diffusion process, controls the final surface characteristics.

References


