The Lithography Expert *(May 2006)*

Diffusion and Resolution

Chris A. Mack, *Austin, Texas*

The diffusion of chemical species during post-exposure bake is a necessary evil. For a chemically amplified resist, acid must diffuse in order to find reactive sites on the polymer to deblock. It is the deblocking that changes the resist solubility, and the greater the amount of diffusion, the greater the amount of deblocking. But diffusion will also reduce resolution. Consider a diffusion length (the average distance that a particle will diffuse) that is greater than the feature size – all of the information of the feature will disappear as it smears itself out due to this diffusion. Of course, a good resist used properly will have a diffusion length much smaller than the desired feature size. But how much diffusion is too much? Since reducing diffusion reduces the amount of deblocking (and thus the sensitivity of the resist), we don't want to drive diffusion too low. In this column we'll derive some simple rules to quantitatively explain the impact of diffusion on resolution.

While the full reaction-diffusion kinetics of a chemically amplified resist are manageably complex, here we'll consider the simpler case of a conventional resist so that diffusion can be treated by itself. Fickean diffusion, where the diffusivity of the diffusing material remains constant, can be treated as a simple convolution problem. If *m(x)* is the original latent image (the spatial distribution of chemical species in the resist), the new latent image after diffusion *m*(x)* can be found by convolving the original chemical distribution with a Gaussian:

$$m^*(x) = m(x) \otimes Gaussian = \frac{1}{\sqrt{2\pi}\sigma_D} \int_{-\infty}^{\infty} m(\tau) e^{-(x-\tau)^2/\sigma_D^2} d\tau$$

where σ_D is the diffusion length. While a one-dimensional case is shown here, it is easy to extend the convolution to two or three dimensions (though things do get more complex when boundary conditions are applied).

Consider first the case of an isolated line (dense features will be considered next). To make things very simple, let's assume the isolated line latent image before diffusion is a Gaussian of width *w*. That makes the convolution calculation very simple: the convolution of two Gaussians is itself a Gaussian that is wider than the original two.

$$w^* = \sqrt{w^2 + w_D^2} \qquad (1)$$

where w^* is the width of the latent image after diffusion and w_D is the width of the diffusion Gaussian. If we measure the width of the feature at the 30% threshold, then $w_D \approx 3.1\sigma_D$. As an example, if the diffusion length is 10% of the original feature size (a not uncommon case), the post-diffusion feature size will be 5% larger. If the diffusion length is increased to 20% of the feature size, the final feature will grow by 18% compared to the original.

But tracking the width of the final isolated feature is not necessarily the most appropriate way to explain the impact of diffusion on resolution. When trying to print a feature at a given target size, increased diffusion can be compensated by a change in exposure dose to get the correct final dimension. The effect of diffusion then is a loss in process latitude due to a reduction in the slope of the latent image. For a Gaussian image shape, the log-slope at any given position is inversely proportional to the square of the width of the Gaussian. Thus, while a diffusion length that is 10% of the pre-diffusion feature size will grow the feature size by about 5%, it will also reduce the log-slope of the latent image at the nominal feature edge by about 10% (see Figure 1). The impact of reducing the latent image slope will be a loss in exposure latitude and thus process window. For an isolated feature, resolution is limited by your ability to control smaller features, so a reduction in process latitude translates directly into worse resolution.

For dense features, the impact of diffusion is a bit different. The pitch resolution, the smallest pitch that can be imaged, is determined by the frequency cut-off of the imaging lens. Thus, strictly speaking, diffusion has no role to play in the ultimate pitch resolution. But like the isolated feature discussed above, diffusion reduces process latitude by lowering the slope of the latent image. Consider a generic latent image for a repeating line/space pattern of pitch p described as a Fourier series:

$$m(x) = \sum_{n=0}^{N} a_n \cos(2\pi n x / p) \tag{2}$$

where a pattern symmetrical about $x = 0$ is assumed so that there are no sine terms in the series. Larger values of n represent higher frequency terms (harmonics) in the image, though a typical high resolution dense pattern will have an upper limit of $N = 2$ or 3. The effect of diffusion has been previously described [1] as simply a reduction in the amplitude of each harmonic.

$$m^*(x) = \sum_{n=0}^{N} a_n^* \cos(2\pi n x / p) \tag{3}$$

where $$a_n^* = a_n e^{-2(\pi n \sigma_D / p)^2}$$

Several interesting observations can be made from this expression. Obviously, a greater diffusion length leads to a greater degradation of the latent image (a_n is reduced). But it is the diffusion length relative to the pitch that matters. Thus, for the same diffusion length, smaller pitch patterns are degraded more than larger pitch patterns. Also, the higher frequency terms degrade faster than the lower frequency terms. In fact, each frequency term can be said to have an effective diffusion length equal to $n\sigma_D$ and it is the ratio of this effective diffusion length to the pitch that determines the amount of damping for that frequency component (Figure 2).

Given the discussion above, which pattern is more sensitive to diffusion, a dense feature or an isolated feature? Let's assume that both features have the same nominal feature size, that the dense features are equal lines and spaces, and that the dense features are high resolution so that only the zero and first harmonics are found in the latent image. For the dense pattern, the latent image log-slope will be proportional to a_1/a_0. For manageably small levels of diffusion, the impact of diffusion on the latent image log-slope for each feature type becomes

$$\text{Isolated Line:} \qquad log-slope_{final} \approx log-slope_{initial}\left[1-10\left(\frac{\sigma_D}{w}\right)^2\right]$$

$$\text{Dense Line:} \qquad log-slope_{final} \approx log-slope_{initial}\left[1-5\left(\frac{\sigma_D}{w}\right)^2\right] \tag{4}$$

Thus, a small isolated line is about twice as sensitive to diffusion as a small dense pattern of lines and spaces. In reality, even the smallest dense features often have some higher frequency components (n = 2 and possibly 3) in their latent images, so that the difference between dense and isolated feature sensitivity to diffusion is not as great as that given in equation (4).

As a final note, it is interesting to compare the effects of diffusion to the effects of defocus. In general, the impact of defocus on an image is much more complicated than diffusion, since defocus adds a phase error to electric field diffraction orders that can then recombine in non-linear (and sometimes non-obvious) ways. However, for some simple cases a direct comparison can be made. In a previous edition of this column (*MLW*, November 2004), the impact of defocus on an alternating phase shifting mask was described. For a pattern of small lines and spaces such that only and all of the two first orders are captured, the effect of defocus is to reduce the magnitude of the primary harmonic of the intensity image:

$$\frac{a_{1(defocus)}}{a_{1(in-focus)}} = \frac{J_1(2\pi\,\delta\,\sigma\,NA/\,p)}{\pi\,\delta\,\sigma\,NA/\,p} \tag{5}$$

where δ is the defocus distance, *NA* is the numerical aperture, and σ is the partial coherence (not to be confused with the diffusion length!). For small amounts of defocus and diffusion, the Bessel function defocus effect of equation (5) and the exponential diffusion effect of equation (3) can both be expanded as Taylor series.

$$\frac{J_1(2\pi\,\delta\,\sigma\,NA/\,p)}{\pi\,\delta\,\sigma\,NA/\,p} \approx 1-0.5(\pi\delta\sigma NA/\,p)^2\,, \qquad e^{-2(\pi n\sigma_D\,/\,p)^2} \approx 1-2(\pi n\sigma_D\,/\,p)^2 \tag{6}$$

Comparing these two approximate expressions, a small amount of defocus is equivalent to a small amount of diffusion.

$$\sigma_D \approx \frac{\delta\sigma NA}{2} \tag{7}$$

Again, there is not always perfect correlation between diffusion and defocus, depending on the specifics of the imaging case. But sometimes, the idea that a focus blur behaves just like a diffusion blur is as accurate as it is intuitive.

References

1. M. D. Smith and C. A. Mack, "Examination of a simplified reaction-diffusion model for post exposure bake of chemically amplified resists," *Advances in Resist Technology and Processing XVIII*, *Proc.*, SPIE Vol. 4345 (2001) pp.1022-1036.

List of Figures

Figure 1. Effect of diffusion on the latent image log-slope for an isolated line (Gaussian latent image shape assumed).

Figure 2. Effect of diffusion on the latent image frequency components for a dense line.

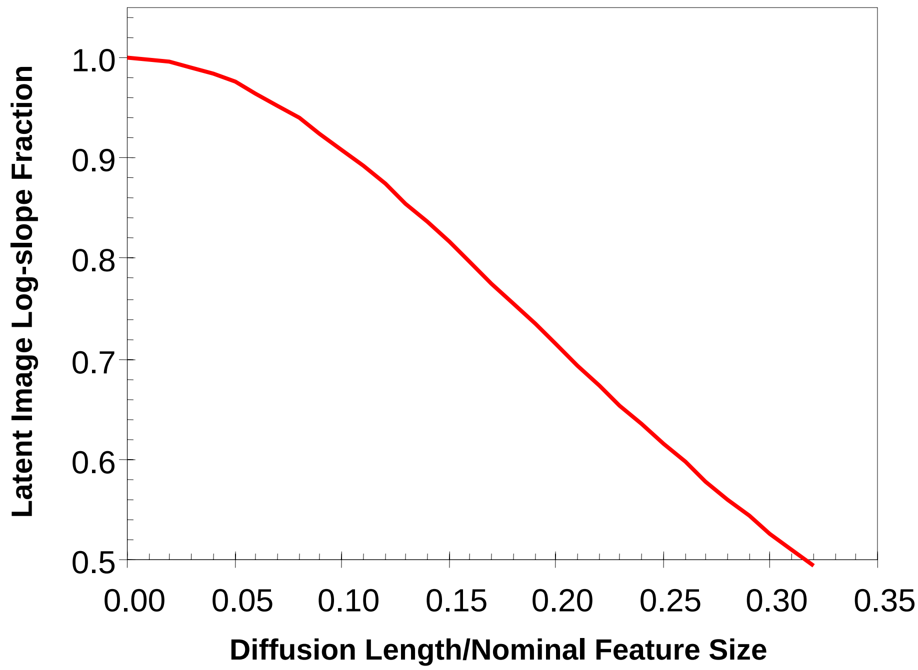

Figure 1. Effect of diffusion on the latent image log-slope for an isolated line (Gaussian latent image shape assumed).

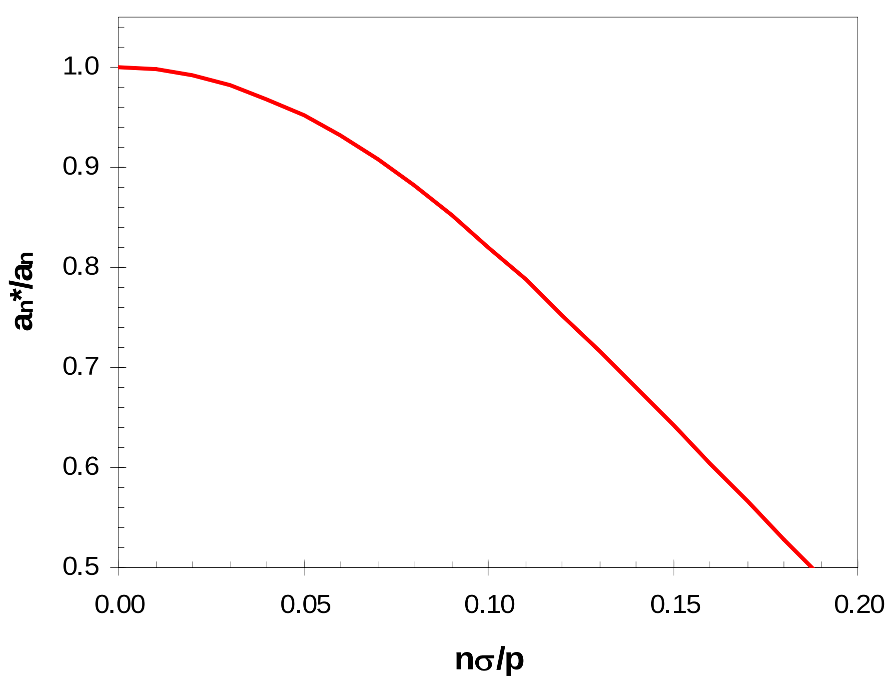

Figure 2. Effect of diffusion on the latent image frequency components for a dense line.