The Lithography Expert *(February 2008)*

Focus Averaging

Whether intentional or not, it is common for photoresist to be exposed by an image that is actually some average of images through focus. The term *focus averaging* (also called focus drilling, and sometimes known by the confusing term 'focus blur') can come from a variety of different sources. Chromatic aberrations in non-chromatic-corrected lenses (as is always the case for 248-nm and 193-nm lithography), coupled with a nonzero bandwidth source, cause focus averaging since the plane of best focus varies linearly with wavelength for such a lens. Wafer-stage vibrations in the focus direction produce such an averaging, as does the case where the wafer is tilted as it is being scanned. Occasionally, such focus averaging is intentionally induced in order to extend the tolerable range of focus (called the FLEX method by its inventors [1]).

What is the impact of this focus averaging on image quality? Is it always bad, or is there some circumstance where it can be beneficial? To investigate this phenomenon, we'll make a simplifying assumption about the response of the image to focus: at any point in space the aerial image will vary quadratically about best focus. In other words,

$$I(x,\delta) \approx I(x,0) - \delta^2 f(x) \tag{1}$$

where $I(x,0)$ is the image at best focus, δ is the defocus distance, and $f(x)$ is a property of that particular image. The approximation of equation (1) is a good one for small levels of defocus. Figure 1 shows an aerial image of a contact hole, as well as the function $f(x)$ for that image.

To describe the impact of different kinds of focus averaging in the most general terms, let $P(\delta)$ be the probability distribution of the defocus error δ. The focus-averaged image will be [2]

$$I_{focus-average}(x) \approx I(x,0) - f(x)\int_{-\infty}^{\infty} \delta^2 P(\delta)d\delta \tag{2}$$

The integral in equation (2) is called the raw second moment of the probability distribution, and can be calculated with the help of the definitions of the mean (μ_F) and variance (σ_F^2) of the distribution:

$$\mu_F \equiv \int_{-\infty}^{\infty} \delta P(\delta)d\delta$$

$$\sigma_F{}^2 \equiv \int_{-\infty}^{\infty} (\delta - \mu_F)^2 P(\delta)d\delta \tag{3}$$

so that the raw second moment becomes

$$\int_{-\infty}^{\infty} \delta^2 P(\delta) d\delta = \mu_F{}^2 + \sigma_F{}^2 \tag{4}$$

Thus, the focus-averaged image is simply

$$I_{focus-average}(x) \approx I(x,0) - f(x)\left[\mu_F{}^2 + \sigma_F{}^2\right] \tag{5}$$

As equation (5) shows, if the variance of the focus probability distribution goes to zero (i.e., there is no focus averaging), the mean focus value is just the defocus distance and equation (5) becomes equation (1). Equation (5) also shows the interesting result that focus averaging always makes the image worse – the focus variance is added to the square of the mean focus error so that the effective focus error is always larger. But this result contradicts many years of simulations and experience that says in some cases focus averaging can make the image better. What gives?

The source of the discrepancy comes from the approximation used in our quadratic focus model of equation (1). In fact, while for small focus errors the image does vary quadratically, for somewhat larger errors a quadratic change is too pessimistic. A better model includes a fourth-order term:

$$I(x,\delta) \approx I(x,0) - \delta^2 f(x)\left[1 - \beta\delta^2\right] \tag{6}$$

Since the parameter β is positive, this additional fourth order term works against the second order term to slow the degradation of the image. The resulting focus-averaged image can be calculated from the fourth moment of the probability distribution. For example, consider a normal (Gaussian) probability distribution. The raw fourth moment is

$$\int_{-\infty}^{\infty} \delta^4 P(\delta) d\delta = \mu_F{}^4 + 6\mu_F{}^2\sigma_F{}^2 + 3\sigma_F{}^4 \tag{7}$$

giving a focus-averaged image of

$$I_{focus-average}(x) \approx I(x,0) - f(x)\left[\left(\mu_F{}^2 + \sigma_F{}^2\right) - \beta\left(\mu_F{}^4 + 6\mu_F{}^2\sigma_F{}^2 + 3\sigma_F{}^4\right)\right] \tag{8}$$

Whether or not focus averaging makes the image better or worse depends on β and on the mean of the focus distribution. Using the normalized image log-slope (NILS) to judge image quality, consider for simplicity a desired dimension of the image at the isofocal point. This condition will greatly simplify the following mathematics without invalidating the conclusions under more general conditions. Calculating the NILS from equation (8),

$$NILS_{focus-average} \approx NILS_0 - \alpha\left[\left(\mu_F{}^2 + \sigma_F{}^2\right) - \beta\left(\mu_F{}^4 + 6\mu_F{}^2\sigma_F{}^2 + 3\sigma_F{}^4\right)\right] \tag{9}$$

where $NILS_0$ is the NILS at best focus and $\alpha = \frac{f'(x)CD}{I(x)}$ evaluated at the isofocal point, and thus is a constant for a given image. Without focus averaging (i.e., when $\sigma_F = 0$),

$$NILS_{no-averaging} \approx NILS_0 - \alpha\left[\mu_F{}^2 - \beta\mu_F{}^4\right] \tag{10}$$

Thus, equation (9) can also be written in terms of the equivalent NILS with defocus but no averaging:

$$NILS_{focus-average} \approx NILS_{no-averaging} - \alpha\sigma_F{}^2\left[1 - \beta\left(6\mu_F{}^2 + 3\sigma_F{}^2\right)\right] \tag{11}$$

The effects of focus averaging can now be made clear: averaging will make the image better whenever the above term in square brackets is negative. Thus,

$$\text{when } \mu_F > \sqrt{\frac{1}{6\beta} - \frac{\sigma_F{}^2}{2}}\text{, focus averaging improves the image}$$
$$\text{when } \mu_F < \sqrt{\frac{1}{6\beta} - \frac{\sigma_F{}^2}{2}}\text{, focus averaging worsens the image} \tag{12}$$

For small amounts of focus averaging, the variance in equation (12) can be neglected, giving a critical defocus (the value of defocus where focus averaging helps rather than hurts) of

$$\delta_{crit} \approx \sqrt{\frac{1}{6\beta}} \tag{13}$$

When the (mean) defocus exceeds this critical value, focus averaging improves image quality. For (mean) defocus less than this amount, focus averaging degrades the image. Comparing this analysis with equation (5), it is the fourth-order dependence of image intensity with defocus that enables the potential for benefit from focus averaging. [Note that although equations (12) and (13) were derived for normally distributed focus averaging, equation (13) applies generally to any reasonable focus probability distribution.]

Figure 2a shows a plot of NILS versus defocus for the contact-hole images of Figure 1. Also shown are fits of equation (10) to this data: a second order fit (α = 246 μm^{-2}, β = 0) over a small defocus range, and a quadratic fit (α = 246 μm^{-2}, β = 11.8 μm^{-2}) over a larger focus range. From this fit, the critical defocus value is, from equation (13), 0.12 μm. Figure 2b shows PROLITH simulations of the NILS under the same conditions, but with various amounts of normally-distributed focus averaging. From these curves, the critical defocus value is about 0.115 – 0.12 μm. Thus, the fourth-order approximation does a very good job of predicting the impact of focus averaging.

While normally distributed focus averaging was examined in detail, other probability distributions can be treated in the same way. For example, for a uniform focus distribution over a range Δ,

$$NILS_{focus-average} \approx NILS_{no-averaging} - \alpha\sigma_F{}^2\left[1-\beta\left(6\mu_F{}^2+\frac{9}{5}\sigma_F{}^2\right)\right] \qquad (14)$$

where $\sigma_F{}^2 = \frac{\Delta^2}{12}$.

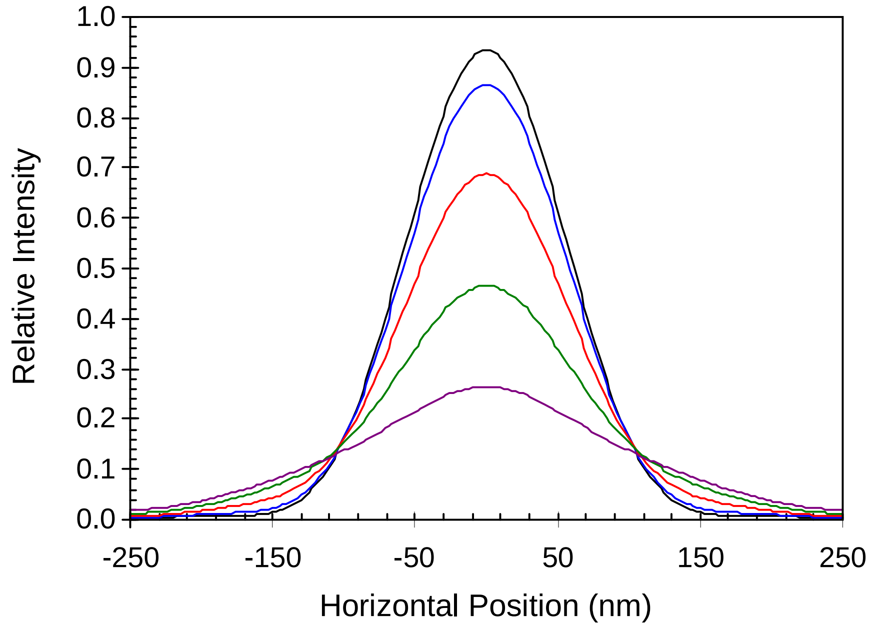
The approximate critical value of defocus remains the same as that given in equation (13).

How can this analysis be usefully employed? When considering whether intentional focus averaging may extend the depth of focus of a feature for a given process, the first step is to compare the critical focus value to the required DOF (that is, the range of focus errors expected in your process). If $2\delta_{crit}$ is about equal to or greater than the required DOF, focus averaging will only make things worse. If, however, the required DOF is appreciably greater than $2\delta_{crit}$, focus averaging will improve image quality at the extremes of focus, making focus averaging a viable DOF-enhancement approach.

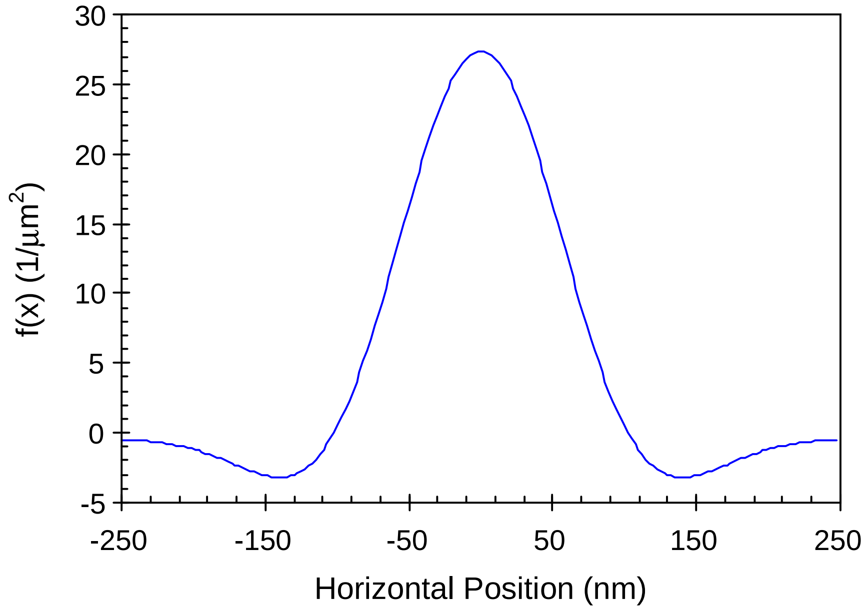
References

1. H. Fukuda, N. Hasegawa, T. Tanaka, T. Hayashida, "A New Method for Enhancing Focus Latitude in Optical Lithography: FLEX", *IEEE Electron Device Letters,* Vol. EDL-8, No. 4, (April 1987) pp. 179-180.

2. C. A. Mack, <u>Fundamental Principles of Optical Lithography: The Science of Microfabrication</u>, John Wiley & Sons (London: 2007).

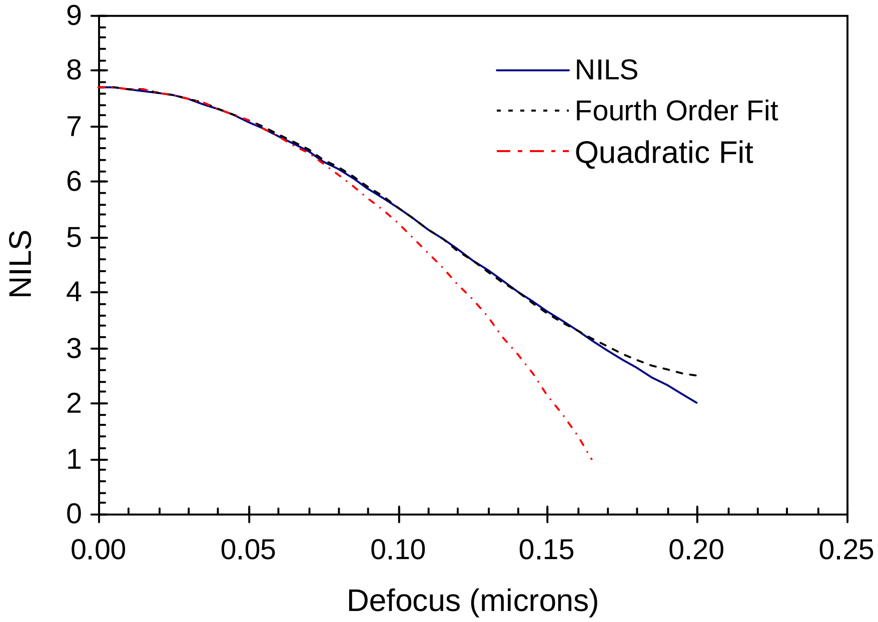


(a)

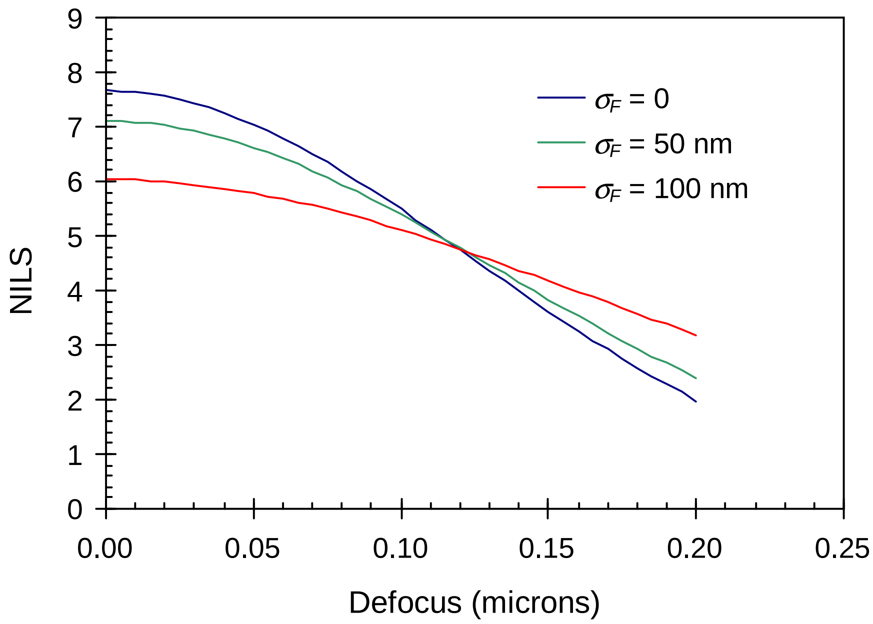


(b)

Figure 1. Aerial image of a contact hole (150 nm chrome on glass contact, λ = 193 nm, NA = 0.93, σ = 0.5) showing (a) images through focus in 0.05 μm steps, and (b) the resulting function $f(x)$.



(a)



(b)

Figure 2. (a) NILS versus defocus for no focus averaging and the images of Figure 1, along with second- and fourth-order fits, and (b) impact of Gaussian focus averaging.