Journal of
Micro/Nanolithography,
MEMS, and MOEMS

Nanolithography.SPIEDigitalLibrary.org

Understanding the efficacy of linewidth roughness postprocessing

Chris A. Mack

SPIE.

Understanding the efficacy of linewidth roughness postprocessing

Chris A. Mack*
Lithoguru.com, 1605 Watchhill Road, Austin, Texas 78703, United States

Abstract. Lack of progress in reducing linewidth roughness of lithographic features has led to investigations of the use of postlithography process smoothing techniques. However, it remains unclear whether such postprocessing will sufficiently reduce the detrimental effects of feature roughness. Thus, there is a need to understand the efficacy of postprocessing on not just roughness reduction, but on the negative device impacts of roughness. We derive model equations of how roughness impacts lithographic performance and incorporates smoothing using postprocessing. These models clearly show that postprocess smoothing works best by increasing the correlation length. Increasing the correlation length is very effective at reducing high-frequency roughness that impacts within-feature variations but is less effective at reducing low-frequency roughness that impacts feature-to-feature variations. It seems that postprocess smoothing is not a substitute for reducing the initial roughness of resist features. [DOI: 10.1117/1.JMM.14.3.033503]

Keywords: line-edge roughness; linewidth roughness; power spectral density; postprocessing; smoothing.

Paper 15066P received Apr. 22, 2015; accepted for publication Jun. 17, 2015; published online Jul. 14, 2015.

1 Introduction

Line-edge roughness (LER) and linewidth roughness (LWR) are becoming increasingly important sources of error in lithographic processing because feature sizes have been shrinking faster than the magnitude of the LWR. For extreme ultraviolet lithography in particular, reducing LWR has been a vexing problem, with resist LWR on the order of 4 to 5 nm (3σ),[1] but with requirements of less than 2 nm (12% of the physical gate length, according to the International Technology Roadmap for Semiconductors[2]). One potential solution is the use of postlithography processing, such as electron or ion beam or ultraviolet light exposure, annealing in a hydrogen environment, HBr plasma treatment, or the etch process itself.[3–10] Such processes have been shown to reduce the high-frequency roughness and the overall 3σ of the roughness. There is some debate, however, about whether such postprocess smoothing is effective at reducing roughness in a meaningful way—a way that will reduce the detrimental impact of roughness on device performance.

The impact of LER and LWR on semiconductor devices is a function of the nature of the roughness, which includes both the amount of roughness and its frequency content. Low-frequency roughness, occurring over long length scales, behaves like an error in mean critical dimension (CD) or edge position, resulting in feature-to-feature variation.[11] High-frequency roughness gives within-feature variation that we classically recognize as a "rough" feature, and can impact the performance of interconnect lines in particular. The frequency behavior of the roughness is usually characterized by its power spectral density (PSD), which describes how much variance in the feature can be found in each increment of frequency. As will be explained, knowing the PSD allows one to translate characteristics of the roughness into the effects of the roughness.

These two regimes of high-frequency and low-frequency roughnesses can be demarcated by considering the following application of lithography. Suppose our goal is to print many identical lines (nominally rectangular in shape) of (narrow) width w and (longer) length L. Ignoring all other aspects of process variations that might impact the printing of these lines, we ask "How will roughness affect these lines?" There are two categories of impacts. The "high" frequency impact will be the LWR of each feature about its mean CD, $\bar{w}$. For a line of length L, we shall denote this LWR as $\sigma_{\mathrm{LWR}}(L)$. The "low" frequency impact will be the feature-to-feature variation of $\bar{w}$, the roughness-induced critical dimension uniformity (CDU). For a line of length L, we shall denote this mean linewidth variation as $\sigma_{\mathrm{CDU}}(L)$. Thus, the goal of this paper will be to understand the efficacy of postprocess smoothing on these two aspects of lithographic roughness. As will be shown, the impact of postprocess smoothing will depend on whether a complementary lithography scheme is used (where smoothing takes place after long lines have been formed, but before they have been cut up into smaller segments) or standard lithography is used to make the lines (where smoothing is performed on the final line segments).

2 Characterizing High and Low Frequency Roughnesses

Power spectral densities for measured lithographic features exhibit common general tendencies. For low frequencies, the PSD is flat, indicating uncorrelated, white noise. For high frequencies, the PSD falls as a power of 1/frequency, indicating correlated, fractal behavior. The transition occurs at a frequency corresponding to the correlation length. The

*Address all correspondence to: Chris A. Mack, E-mail: chris@lithoguru.com

1932-5150/2015/$25.00 © 2015 SPIE

Palasantzas PSD function[12] has been found to describe well this basic shape of LER and LWR spectral densities for wide range of conditions both pre- and postprocessing, in the resist and after etch. The function is

$$\mathrm{PSD}(f) = \frac{\mathrm{PSD}(0)}{[1+(2\pi f\xi)^2]^{H+1/2}},$$
$$\mathrm{PSD}(0) = 2\sigma_{\mathrm{LWR}}^2\xi\left(\frac{\sqrt{\pi}\Gamma\left(H+\frac{1}{2}\right)}{\Gamma(H)}\right), \quad (1)$$

where f is the frequency, ξ is the correlation length, H is the roughness exponent, Γ is the gamma function, and σ_{LWR} is the standard deviation of the linewidth for an infinitely long line, which we shall also denote as $\sigma_{\mathrm{LWR}}(\infty)$. A graph of a typical PSD is shown in Fig. 1, indicating the role of each of the three parameters in defining the shape of the PSD.

Given this analytical function for the PSD, we can now provide expressions for the two quantities of interest: $\sigma_{\mathrm{LWR}}(L)$, the LWR for a feature of length L; and $\sigma_{\mathrm{CDU}}(L)$, the feature-to-feature variation of the mean linewidth of a line of length L as caused by roughness. The general expression for CDU has been previously derived[11]

$$\sigma_{\mathrm{CDU}}^2(L) = \frac{(2\,H+1)\xi\sigma_{\mathrm{LWR}}^2(\infty)}{L}\left[1-\frac{\xi}{L}(1-\mathrm{e}^{-L/\xi})\right]. \quad (2)$$

Additionally, a single, dimensionless metric of LWR has been proposed[11]

$$M_{\mathrm{LWR}} \equiv \frac{\sigma_{\mathrm{LWR}}}{\mathrm{CD}}\sqrt{\frac{(2H+1)\xi}{\mathrm{CD}}}. \quad (3)$$

For reasonably long lines ($L \gg \xi$), this metric is directly proportional to the relative CDU ($\sigma_{\mathrm{CDU}}/\mathrm{CD}$) and is thus preferred over simply quoting the $3\sigma_{\mathrm{LWR}}$ value for most applications.

For $\sigma_{\mathrm{LWR}}(L)$, analytical expressions can be derived for the cases of $H = 0.5$ and $H = 1.0$, the extremes of the expected range of values for the roughness exponent. These derivations, however, involve some important subtleties and can be deduced in different ways. In the first approach, we will derive $\sigma_{\mathrm{LWR}}(L)$ from the frequency domain by integrating the PSD for high frequencies, $|f| \geq f_L$, and thus, subtracting out the low-frequency roughness.

$$\sigma_{\mathrm{LWR}}^2(L) = \int_{-\infty}^{-f_L}\mathrm{PSD}(f)\mathrm{d}f + \int_{f_L}^{\infty}\mathrm{PSD}(f)\mathrm{d}f$$
$$= \sigma_{\mathrm{LWR}}^2(\infty) - \int_{-f_L}^{f_L}\mathrm{PSD}(f)\mathrm{d}f. \quad (4)$$

For the Palasantzas PSD, this results in

$$H = 0.5, \quad \frac{\sigma_{\mathrm{LWR}}^2(L)}{\sigma_{\mathrm{LWR}}^2(\infty)} = 1-\frac{2}{\pi}\tan^{-1}(2\pi\xi f_L),$$

$$H = 1.0, \quad \frac{\sigma_{\mathrm{LWR}}^2(L)}{\sigma_{\mathrm{LWR}}^2(\infty)} = 1-\left[1+\left(\frac{1}{2\pi\xi f_L}\right)^2\right]^{-1/2}. \quad (5)$$

Previously, in Ref. 13, a value of $f_L = 1/L$ was chosen as the low-frequency cut-off. However, upon careful consideration the correct choice is $f_L = 1/(2\,L)$. While the lowest frequency found in a PSD taken from a line of length L will be at $f_L = 1/L$, the PSD at this frequency is best thought of as a bin of PSD data centered at this frequency. The missing bin, corresponding to PSD(0), runs from $f = -1/(2\,L)$ to $f = +1/(2\,L)$. Thus, it is this bin that must be subtracted out in Eq. (4). This results in the final form of our $\sigma_{\mathrm{LWR}}(L)$ derivation using the PSD method.

$$H = 0.5, \quad \frac{\sigma_{\mathrm{LWR}}^2(L)}{\sigma_{\mathrm{LWR}}^2(\infty)} = 1-\frac{2}{\pi}\tan^{-1}\left(\frac{\pi\xi}{L}\right),$$

$$H = 1.0, \quad \frac{\sigma_{\mathrm{LWR}}^2(L)}{\sigma_{\mathrm{LWR}}^2(\infty)} = 1-\left[1+\left(\frac{L}{\pi\xi}\right)^2\right]^{-1/2}. \quad (6)$$

A plot of the square-root of these two equations is given in Fig. 2. While an analytical expression does not exist for

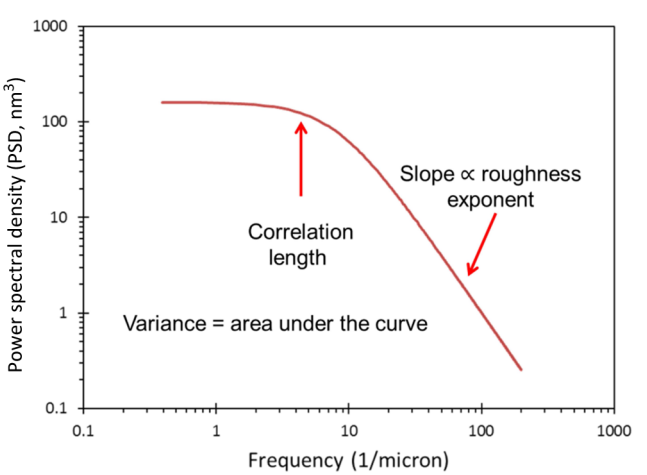

Fig. 1 Example of a typical power spectral density, using $\sigma_{\mathrm{LWR}} = 2$ nm, $\xi = 20$ nm, and $H = 0.5$.

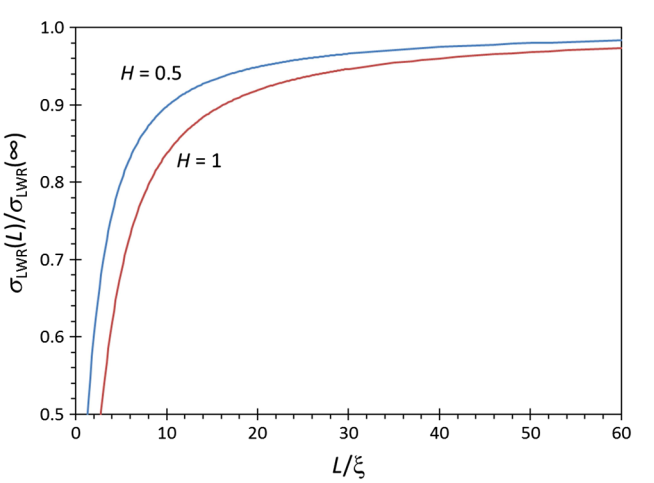

Fig. 2 A plot of the square-root of Eq. (6) as a function of the line length L, in multiples of the correlation length ξ.

intermediate values of H, the resulting $\sigma_{\mathrm{LWR}}(L)$ behavior will be intermediate between these two curves. For $L/\xi > 5$, a simple third-order Taylor series will approximate these expressions quite well.

$$H = 0.5, \quad \frac{\sigma^2_{\mathrm{LWR}}(L)}{\sigma^2_{\mathrm{LWR}}(\infty)} \approx 1 - \frac{2\xi}{L}\left(1 - \frac{1}{3}\left[\frac{\pi\xi}{L}\right]^2\right),$$

$$H = 1.0, \quad \frac{\sigma^2_{\mathrm{LWR}}(L)}{\sigma^2_{\mathrm{LWR}}(\infty)} \approx 1 - \frac{\pi\xi}{L}\left(1 - \frac{1}{2}\left[\frac{\pi\xi}{L}\right]^2\right). \tag{7}$$

The PSD method of deriving $\sigma_{\mathrm{LWR}}(L)$ suffers from the problem of spectral leakage,[14] and so the results from Eq. (6) are slightly in error (as will be shown). Avoiding spectral leakage, an expression for $\sigma_{\mathrm{LWR}}(L)$ can also be derived from the autocovariance function, which is the Fourier transform of the PSD. Here, I will follow the approach of Zhao et al.[15] Unfortunately, their published result contains an error in their equation (2.63), so the correct result will be given here. The definition of $\sigma_{\mathrm{LWR}}(L)$ as an RMS roughness of a line of length L can be obtained from

$$\sigma^2_{\mathrm{LWR}}(L) = \left\langle \frac{1}{L}\int_{-L/2}^{L/2} [w(x) - \bar{w}]^2 \mathrm{d}x \right\rangle$$
$$\text{where } \bar{w} = \frac{1}{L}\int_{-L/2}^{L/2} w(x)\mathrm{d}x, \tag{8}$$

and $\langle \ldots \rangle$ denotes an average over many rough lines. Since the mean linewidth of a given feature $\bar{w}$ is calculated from the $w(x)$ data from that feature, $\bar{w}$ will be correlated with $w(x)$ in Eq. (8). Bringing the expectation into the integral and carrying out the square gives

$$\sigma^2_{\mathrm{LWR}}(L) = \sigma^2_{\mathrm{LWR}}(\infty) - \frac{1}{L^2}\int_{-L/2}^{L/2}\mathrm{d}x \int_{-L/2}^{L/2} \mathrm{ACF}(x - r)\mathrm{d}r, \tag{9}$$

where ACF is the autocovariance function of the roughness. For the case of $H = 0.5$, the ACF is an exponential.

$$H = 0.5, \quad \mathrm{ACF}(r) = \sigma^2_{\mathrm{LWR}}(\infty)\mathrm{e}^{-|r|/\xi}. \tag{10}$$

Plugging this ACF into Eq. (9) and carrying out the integrations gives the final result, previously derived by Leunissen et al.[16]

$$\frac{\sigma^2_{\mathrm{LWR}}(L)}{\sigma^2_{\mathrm{LWR}}(\infty)} = 1 - \frac{2\xi}{L}\left[1 - \frac{\xi}{L}\left(1 - \mathrm{e}^{-L/\xi}\right)\right]. \tag{11}$$

The difference between Eq. (11) derived from the ACF and the $H = 0.5$ Eq. (6) derived from the PSD is small, as seen in Fig. 3 [the relative difference in predicted $\sigma_{\mathrm{LWR}}(L)$ is approximately equal to $(\xi/L)^2$ when $\xi \ll L$]. Simulations of randomly rough lines, using the approach presented previously,[14] match the ACF result. Furthermore, measurements based on discrete sampling will add other biases to the determination of $\sigma_{\mathrm{LWR}}(L)$,[17] although such biases can be kept small and will not impact the analysis presented.

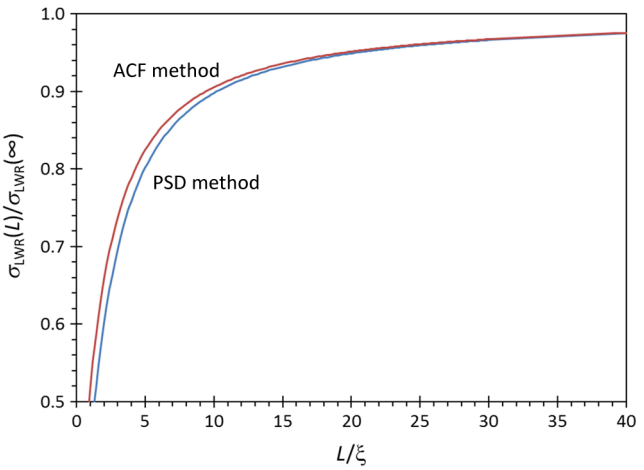

Fig. 3 Comparing the models for $\sigma_{\mathrm{LWR}}(L)$ based on the power spectral density (PSD) method [Eq. (6) for the case of $H = 0.5$] and the autocovariance function method [Eq. (11)].

For $H = 1.0$, the resulting ACF is r multiplied by a modified Bessel function of the second kind of order 1

$$H = 1.0, \quad \mathrm{ACF}(r) = \sigma^2_{\mathrm{LWR}}(\infty)(|r|/\xi)K_1(|r|/\xi). \tag{12}$$

Although an analytical expression for the ACF is possible in this case, it does not produce a convenient result when attempting to integrate with Eq. (9). However, another approach can be used to find an approximate expression for $\sigma_{\mathrm{LWR}}(L)$ for this case. The total linewidth variance of an infinitely long line can be divided into the constituent parts of CDU and LWR for a shorter line of length L.[16] In other words,

$$\sigma^2_{\mathrm{LWR}}(\infty) = \sigma^2_{\mathrm{LWR}}(L) + \sigma^2_{\mathrm{CDU}}(L). \tag{13}$$

For $H = 0.5$, this division of variance components can be confirmed using Eqs. (2) and (11). Since Eq. (2) is approximately correct for all values of H (it is exact for $H = 0.5$), one can use Eq. (13) to derive a general, approximate expression for $\sigma_{\mathrm{LWR}}(L)$ for all roughness exponents.

$$\frac{\sigma^2_{\mathrm{LWR}}(L)}{\sigma^2_{\mathrm{LWR}}(\infty)} = 1 - \frac{(2\,H + 1)\xi}{L}\left[1 - \frac{\xi}{L}(1 - \mathrm{e}^{-L/\xi})\right]. \tag{14}$$

3 Impact of Postprocessing on Complementary Lithography

To begin, we will examine the impact of postlithography smoothing processes when applied to complementary lithography. Complementary lithography begins by printing long lines that, for practical purposes, can be assumed to be infinitely long. Smoothing will be performed on these long lines. Then a second patterning step will cut these long lines up into smaller line segments of length L.

In applying the results given in the Sec. 2 to the topic of postprocessing for complementary lithography, two important assumptions will be made. First, we will assume that

the true PSD of the features, both pre- and postprocessing, can be well-modeled by the Palasantzas PSD of Eq. (1). In that way, we can characterize any postprocessing as a change in one or more of the three PSD parameters, σ_{LWR}, ξ, and H, as applied to very long lines. Second, we will assume that postprocessing can never reduce the PSD at frequencies below one over the length of the line being smoothed. In particular, we will assume that the PSD(0) of the printed features is constant and unaffected by postprocessing. In fact, it is possible to increase PSD(0), since noise can always be added to a process, but we will assume that any postprocessing that does this will be rejected out of hand. We can now apply these assumptions to the equations of Sec. 2 for $\sigma_{LWR}(L)$ and $\sigma_{CDU}(L)$, under the constraint that PSD(0) is constant.

Writing the CDU in terms of PSD(0)

$$\sigma_{CDU}^2(L) = K^2(H)\frac{PSD(0)}{L}\left[1-\frac{\xi}{L}(1-e^{-L/\xi})\right],$$
$$K^2(H) = \frac{(H+\frac{1}{2})\Gamma(H)}{\sqrt{\pi}\Gamma\left(H+\frac{1}{2}\right)}. \quad (15)$$

First, consider a postprocessing method that increases the roughness exponent H. For constant PSD(0), all of the CDU dependence on roughness exponent H comes from the function $K(H)$. Examining this function over the range of H from 0.5 to 1.0 reveals that $K(H)$ goes from a maximum of 1.0 to a minimum of 0.975. Thus, $K(H)$ is effectively constant, and since CDU is proportional to $K(H)$, changing the roughness exponent has essentially no impact on CDU when PSD(0) is constant. The CDU expression thus becomes

$$\sigma_{CDU}^2(L) \approx \frac{PSD(0)}{L}\left[1-\frac{\xi}{L}(1-e^{-L/\xi})\right]. \quad (16)$$

From Eq. (16), we can see that the only way to improve roughness-induced loss of CDU during complementary lithography is to increase the correlation length. For a given line length, the worst-case CDU comes when $\xi \ll L$, where

$$\xi \to 0, \sigma_{CDU}(L) \to \sqrt{\frac{PSD(0)}{L}} = \sigma_{CDU}(\max). \quad (17)$$

A plot of Eq. (16) is shown in Fig. 4, normalized by $\sigma_{CDU}(\max)$. Increasing the correlation length reduces $\sigma_{CDU}(L)$, but only slowly. If $\xi = L/10$ (a fairly large correlation length in most circumstances), the $\sigma_{CDU}(L)$ is only 5% below its maximum. If a postprocessing scheme doubles ξ to $L/5$, the $\sigma_{CDU}(L)$ is now 10% below its maximum. Growing ξ further to $L/2$ (a very large correlation length for most features), the $\sigma_{CDU}(L)$ is still only 25% below its maximum. A correlation length of $\xi = L$ results in a 40% reduction in $\sigma_{CDU}(L)$. Growing the correlation length does improve CDU, but the correlation length must grow to a large fraction of the final (after cutting) line length to make a noticeable improvement.

Likewise, we can rearrange Eq. (6) in terms of PSD(0).

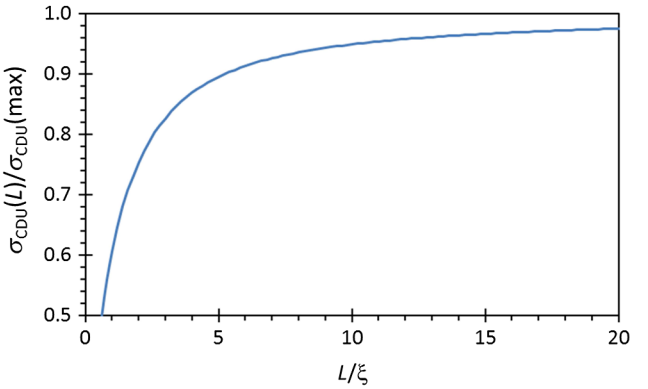

Fig. 4 A plot of Eq. (16) for the CDU at constant PSD(0) as a function of the line length L, in multiples of the correlation length ξ.

$$H = 0.5, \quad PSD(0) = 2\xi\sigma_{LWR}^2(\infty),$$
$$H = 1.0, \quad PSD(0) = \pi\xi\sigma_{LWR}^2(\infty), \quad (18)$$
$$H = 0.5, \quad \sigma_{LWR}^2(L) = \frac{PSD(0)}{2\xi}\left(1-\frac{2}{\pi}\tan^{-1}\left(\frac{\pi\xi}{L}\right)\right)$$
$$H = 1.0, \quad \sigma_{LWR}^2(L) = \frac{PSD(0)}{\pi\xi}\left(1-\left[1+\left(\frac{L}{\pi\xi}\right)^2\right]^{-1/2}\right). \quad (19)$$

We can plot Eqs. (18) and (19), creating a dimensionless $\sigma_{LWR}(L)$ by multiplying it by $\sqrt{L/PSD(0)}$ [that is, dividing by $\sigma_{CDU}(\max)$]. The results are shown in Fig. 5. As can be seen, increasing either the roughness exponent or the correlation length can have a large impact on the within-feature

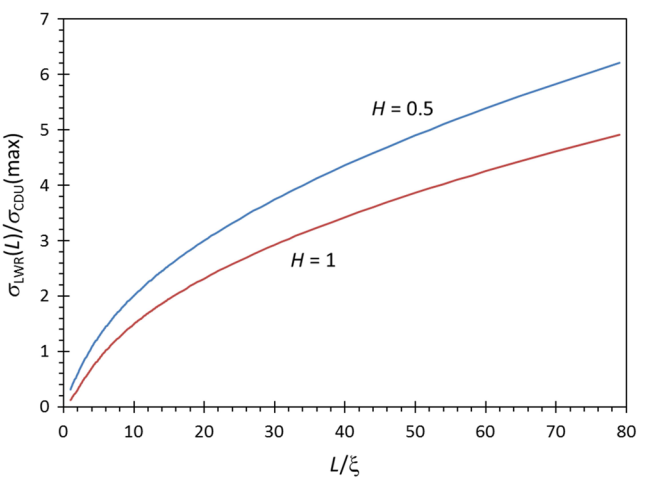

Fig. 5 A plot of the square-root of Eq. (19) for the short-length linewidth roughness (LWR) at constant PSD(0) as a function of the line length L, in multiples of the correlation length ξ.

roughness. If the correlation length is doubled such that L/ξ drops from 40 to 20, the within-feature LWR is reduced by 1/3. Doubling the correlation length again produces a 55% drop from the original LWR value. These increases in correlation length are quite achievable for many different smoothing processes (given the 10- to 30-nm correlation lengths typical of high-resolution photoresist processes). Thus, it is reasonable to expect that postprocess smoothing could be effective at reducing the high-frequency, within-feature LWR.

4 Impact of Postprocessing on Standard Lithography

Consider now the impact of postlithography process smoothing on standard lithography rather than the complementary lithography discussed in the Sec. 3. In standard (not complementary) lithography, a line of length L is formed in photoresist and then the smoothing process is performed. For such a case, the postprocess smoothing can have no impact on the CDU of any feature.

To see this, consider a hypothetical smoothing process that does have the ability to reduce the $\sigma_{\mathrm{CDU}}(L)$ for a collection of lines of length L. This means that the smoothing process must be able to change the mean CD of these lines in the proper direction. If the mean width of a certain feature is too small, the smoothing technique must make it larger. If the mean width of a different instance of that feature on the same wafer is too large, the smoothing technique must make that one smaller. In other words, the smoothing process must be smart enough to know whether a given feature's mean CD is too small or too large and make the appropriate correction to move that CD closer to its target value. It must be able to take material away from a line that is too big and add it to a different line that is too small. There are some systems, such as directed self-assembly (DSA), which have a preferred feature size and a thermodynamic tendency to approach that preferred size. However, no currently proposed postprocess smoothing technique works in this way. Thus, it seems unlikely that any smoothing technique currently conceived has the possibility of lowering the $\sigma_{\mathrm{CDU}}(L)$ for any value of L when standard (not complementary) lithography is employed.

As for $\sigma_{\mathrm{LWR}}(L)$, the equations derived in the Sec. 3 should be accurate whenever the correlation length is small compared to the line length, $\xi \ll L$. Thus, both increasing the roughness exponent and increasing the correlation should improve within-feature variation for standard (not complementary) lithography.

As an extreme example of smoothing, one could imagine a "liquefaction" process that melts the resist line and turns the actual CD all along the line into the mean CD.[18] Such a process would set the within-feature LWR to zero, but all variation in the mean CD between features that existed before the liquefaction process would remain. CDU would not be improved.

5 Discussion and Conclusions

The aforementioned results provide some very clear guidance on how to think about the impact of postprocess smoothing. For either standard or complementary lithography, it is possible to significantly reduce $\sigma_{\mathrm{LWR}}(L)$, the within-feature variation, using postprocessing. This is valuable, especially for interconnect layers that are susceptible to this high-frequency roughness. However, the problem of reducing the feature-to-feature variation is much harder. For standard (noncomplementary) lithography, postprocess smoothing will have no impact on CDU. Thus, only for complementary lithography is it even possible to consider using such smoothing techniques to reduce $\sigma_{\mathrm{CDU}}(L)$.

Postprocessing can either increase the roughness exponent, or increase the correlation length, or both. For example, Azarnouche has shown that a combination of a VUV cure followed by etching steps increased the roughness exponent of a very long 80-nm wide feature from 0.57 to 0.9, and increased the correlation length of that feature from 16 to 41 nm.[19] As the aforementioned derivations show, within-feature variation, $\sigma_{\mathrm{LWR}}(L)$, is affected by both the increase in correlation length and the increase in roughness exponent. For the Azarnouche data, a 200-nm long line would be expected to experience a 3× decrease in $\sigma_{\mathrm{LWR}}(L)$. But for standard lithography, this smoothing would not impact the CDU of any feature. For complementary lithography, only the increase in correlation length will improve CDU. For example, if these 80-nm features were cut into 200-nm long segments, the aforementioned expressions predict that the CDU would be improved by only about 6% for Azarnouche's smoothing process, an essentially unnoticeable amount (for example, decreasing from 2.0 nm to 1.9 nm). If the features were cut into short 100 nm long segments, the improvement in CDU would be a somewhat more noticeable; 14%. However, if we needed our smoothing process to lower the roughness-induced CD nonuniformity by a factor of 2, these small improvements in CDU will not be enough to save the day. Instead, the only way to obtain a significant reduction in $\sigma_{\mathrm{CDU}}(L)$ for complementary lithography is to have a correlation length that approaches (or even exceeds) L.

Whereas the models developed in this paper are for line or space features, the thought processes apply equally well to contact holes. Postprocess smoothing may be capable of reducing the high-frequency roughness found in contact holes, but would be incapable of improving the roughness-induced CDU (sometimes called local CDU in the case of contact holes). No conventional postlithography smoothing process (excluding a DSA process) can simultaneously make holes that are too large become smaller and holes that are too small become bigger. However, more work is required to better apply the concepts of this paper to contact hole roughness.

The models presented in this paper make various points about postprocessing clear, but are based on two very important assumptions: that the true PSD of the features can be well-modeled by the Palasantzas PSD expression, and that the zero-frequency PSD of the printed feature is unaffected by postprocessing. Both of these assumptions are deserving of scrutiny.

Experience has shown that most PSDs follow the general shape of the Palasantzas PSD. Whereas some deviations in the shape are certainly possible—there is no theory stating that the roughness must follow exactly the Palasantzas PSD expression—there is little chance that these small deviations in shape will affect the conclusions of this paper. Any PSD that follows the general shape of being flat at low frequencies, followed by decreasing PSD above a certain critical

frequency, will produce the same basic conclusions as derived here for the Palasantzas PSD.

The second assumption is more interesting, and potentially controversial. To reject this assumption and say that it is possible to lower PSD(0) is to say that it is possible to change the mean CD of a very long line in the proper direction. Using the same logic as described for the CDU in standard lithography, lowering the PSD(0) means that variation of the mean CD of a very long line must be reduced by the smoothing technique, requiring a smoothing process that knows which direction the mean CD of every line must be changed.

On the other hand, some experimental data have exhibited the behavior of lower PSD(0) after smoothing. However, this can be explained as simply an artifact of the metrology: the apparent PSD(0) was greater than the true PSD(0) preprocessing. An example of such an effect was described by Wallow et al.[20] As another example, if the SEM metrology was set to measure the PSD of the top of a resist feature, but the etch process responded to the bottom of the resist feature, it is certainly possible that the after-etch PSD could have a lower PSD(0) than the before-etch PSD. This does not mean that the etch process "smoothed" the PSD(0) down. It is simply that the original before-etch metrology measured incorrectly the PSD(0) of the resist feature. In such a circumstance, one might find a smoothing technique that lowers the before-etch resist PSD(0), but not the after-etch PSD(0). Since the after-etch PSD is the only PSD that really matters, the smoothing was not actually effective in lowering PSD(0).

Another example is SEM image noise. Noise in the SEM image has the effect of uniformly raising the PSD.[14,21] Often, resist SEM images must be taken at very low-electron doses in order to avoid damaging the resist. The result of these low doses is high SEM electron shot noise. However, when imaging after etch, a much higher electron dose can be used since there is no resist to be damaged. This gives better metrology, with lower SEM image noise, and a lower PSD. However, this renders comparing the before and after etch PSDs problematic. A lower after-etch PSD(0) could be due solely to metrology noise differences, not an actual reduction in low-frequency PSD.

Thus, despite some measurements that show postprocessing lowering the PSD at low frequencies, there has been no demonstration to date that these reduced low-frequency PSD values survive through the etch process. Given that there is no logical mechanism for reducing PSD(0) with postprocessing, it seem likely that the assumption of a constant PSD(0) is a good one.

For complementary lithography, increasing correlation length is the only viable approach for postprocess smoothing, and the models developed in this paper help to quantify the benefits of a given increase in correlation length. Correlation lengths must approach, or even exceed, the final feature length to provide meaningful improvement in CDU. How far can correlation length be increased? So far, it seems that none of the proposed smoothing approaches can achieve the needed increase in correlation length.

References

1. R. Gronheid et al., "Extreme-ultraviolet secondary electron blur at the 22-nm half pitch node," *J. Micro/Nanolith. MEMS MOEMS* **10**(3), 033004 (2011).
2. The International Technology Roadmap for Semiconductors, Semiconductor Industry Association, San Jose, 2013, http://www.itrs.net (1July2015).
3. A. V. Pret et al., "Resist roughness evaluation and frequency analysis: metrological challenges and potential solutions for extreme ultraviolet lithography," *J. Micro/Nanolith. MEMS MOEMS* **9**(4), 041308 (2010).
4. P. Foubert et al., "Impact of post-litho linewidth roughness smoothing processes on the post-etch patterning result," *J. Micro/Nanolith. MEMS MOEMS* **10**(3), 033001 (2011).
5. A. Yamaguchi et al., "Analysis of line-edge roughness in resist patterns and its transferability as origins of device performance degradation and variation," *J. Photopolym. Sci. Technol.* **16**(3), 387–394 (2003).
6. A. Pawloski et al., "The transfer of photoresist LER through etch," *Proc. SPIE* **6153**, 615318 (2006).
7. A. V. Pret, R. Gronheid, and P. Foubert, "Roughness characterization in the frequency domain and linewidth roughness mitigation with post-lithography processing," *J. Micro/Nanolith. MEMS MOEMS* **9**(4), 041203 (2010).
8. P. De Schepper et al., "Line edge and width roughness smoothing by plasma treatment," *J. Micro/Nanolith. MEMS MOEMS* **13**(2), 023006 (2014).
9. P. De Schepper et al., "Hydrogen plasma treatment: the evolution of roughness in the frequency domain," *Proc. SPIE* **9054**, 90540C (2014).
10. D. L. Goldfarb et al., "Effect of thin-film imaging on line edge roughness transfer to underlayers during etch processes," *J. Vac. Sci. Technol.* **B22**(2), 647–653 (2004).
11. C. A. Mack, "Analytical expression for impact of linewidth roughness on critical dimension uniformity," *J. Micro/Nanolith. MEMS MOEMS* **13**(2), 020501 (2014).
12. G. Palasantzas, "Roughness spectrum and surface width of self-affine fractal surfaces via the K-correlation model," *Phys. Rev. B* **48**(19), 14472–14478 (1993).
13. C. A. Mack, "Understanding the efficacy of linewidth roughness post-processing," *Proc. SPIE* **9425**, 94250J (2015).
14. C. A. Mack, "Systematic errors in the measurement of power spectral density," *J. Micro/Nanolith. MEMS MOEMS* **12**(3), 033016 (2013).
15. Y. Zhao, G.-C. Wang, and T.-M. Lu, *Characterization of Amorphous and Crystalline Rough Surface: Principles and Applications*, p. 29, Academic Press, San Diego (2001).
16. L. H. A. Leunissen et al., "Full spectral analysis of line width roughness," *Proc. SPIE*, **5752**, 578–590 (2005).
17. C. A. Mack, "Generating random rough edges, surfaces, and volumes," *Appl. Opt.* **52**(7), 1472–1480 (2013).
18. S. Y. Chou and Q. Xia, "Improved nanofabrication through guided transient liquefaction," *Nat. Nanotechnol.* **3**, 295–300 (2008).
19. L. Azarnouche et al., "Plasma treatments to improve line-width roughness during gate patterning," *J. Micro/Nanolith. MEMS MOEMS* **12**(4), 041304 (2013).
20. T. Wallow et al., "Line edge roughness in 193 nm resists: lithographic aspects of etch transfer," *Proc. SPIE* **6519**, 651919 (2007).
21. J. S. Villarrubia and B. D. Bunday, "Unbiased estimation of linewidth roughness," *Proc. SPIE* **5752**, 480–488 (2005).

Chris A. Mack developed the lithography simulation software PROLITH and founded the company FINLE Technologies in 1990. He served as a vice president of lithography technology for KLA-Tencor for five years, until 2005. He received the SEMI award for North America in 2003 and the SPIE Frits Zernike award for microlithography in 2009. He is a fellow of SPIE and IEEE, and is an adjunct faculty member at the University of Texas at Austin.