Efficient proximity effect correction method based on multivariate adaptive regression splines for grayscale e-beam lithography

Wujun Mi, and Peter Nillius

Citation: Journal of Vacuum Science & Technology B, Nanotechnology and Microelectronics: Materials, Processing, Measurement, and Phenomena **32**, 031602 (2014); doi: 10.1116/1.4875955 View online: https://doi.org/10.1116/1.4875955 View Table of Contents: http://avs.scitation.org/toc/jvb/32/3 Published by the American Vacuum Society

Articles you may be interested in

Dose control for fabrication of grayscale structures using a single step electron-beam lithographic process Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures Processing, Measurement, and Phenomena **21**, 2672 (2003); 10.1116/1.1627808

Proximity effect correction for electron beam lithography by equalization of background dose Journal of Applied Physics **54**, 3573 (1983); 10.1063/1.332426

Corrections to proximity effects in electron beam lithography. I. Theory Journal of Applied Physics **50**, 4371 (1979); 10.1063/1.326423

Point exposure distribution measurements for proximity correction in electron beam lithography on a sub-100 nm scale Journal of Vacuum Science & Technology B: Microelectronics Processing and Phenomena **5**, 135 (1987);

Journal of Vacuum Science & Technology B: Microelectronics Processing and Phenomena 5, 135 (1987); 10.1116/1.583847

Proximity effect in electron-beam lithography Journal of Vacuum Science and Technology **12**, 1271 (1975); 10.1116/1.568515

Automated geometry assisted proximity effect correction for electron beam direct write nanolithography Journal of Vacuum Science & Technology B, Nanotechnology and Microelectronics: Materials, Processing, Measurement, and Phenomena **33**, 06FD02 (2015); 10.1116/1.4931691





Efficient proximity effect correction method based on multivariate adaptive regression splines for grayscale e-beam lithography

Wujun Mi^{a)} and Peter Nillius

Department of Physics, Royal Institute of Technology (KTH), SE-106 91 Stockholm, Sweden

(Received 15 February 2014; accepted 30 April 2014; published 13 May 2014)

Grayscale electron beam lithography is an important technique to manufacture three-dimensional (3D) micro- and nano-structures, such as diffractive optical devices and Fresnel lenses. However, the proximity effect due to the scattering of electrons may cause significant error to the desired 3D structure. Conventional proximity correction methods depend on the exposure energy distribution which sometimes is difficult to obtain. In this study, the authors develop a novel proximity effect correction method based on multivariate adaptive regression splines, which takes exposure energy and development into consideration simultaneously. To evaluate the method, a Fresnel lens was fabricated through simulation and experiment. The measurements demonstrate the feasibility and validity of the method. © 2014 American Vacuum Society. [http://dx.doi.org/10.1116/1.4875955]

I. INTRODUCTION

At present, the demands for the three-dimensional (3D) micro and nanostructures, such as 3D hologram, Fresnel lens, and blazed gratings, are on the increase. Various manufacturing methods have been proposed, such as laser micromachining,^{1,2} grayscale optical lithography,³ and nanoimprint lithography.⁴ Compared to these methods, grayscale electron beam lithography has two main advantages: no mask and higher resolution. The proximity effect, which refers to the forward scattering and backscattering when the incident electrons enter the resist and substrate, will however cause the exposure of the adjacent resist, which changes the pattern significantly and limits the resolution of the electron beam lithography.

So far, various methods have been proposed to realize 3D proximity effect correction to fabricate 3D structures. Lee et al. tried to eliminate the proximity effect based on the PYRAMID approach, which divided the exposed area into different regions and implements proximity correction by using a large lookup table.^{5–7} They also used the neural network⁸ and resist-profile-based methods⁹⁻¹² to achieve proximity correction. Unal et al. implemented proximity effect correction with a two-dimensional backscattering function through the self-developed software Layout BEAMER.¹³ Hirai et al. developed an iterative automatic dose optimization system, which calculates the energy distribution by Monte-Carlo simulation and acquires the resist profile through a development process model.^{14,15} Ogino et al. used the simplified electron energy flux (SEEF) model to calculate the backscattered energy distribution in a multilayered structure and achieved the correction through shape and dose modification.^{16,17}

In this study, we propose a novel proximity effect correction method, which is independent of exposure energy point spread function (PSF) in the resist. The method is based on multivariate adaptive regression splines (MARS) and also takes the development process into account. With this method, a MARS model is trained to show the relationship between the dose distribution and 3D structure profile. To evaluate the method, a Fresnel lens has been successfully manufactured through simulation and experiment.

II. GRAYSCALE E-BEAM LITHOGRAPHY AND PROXIMITY EFFECT

In conventional binary electron beam lithography, the resist is exposed with the dose of same value and its main concern is whether the features are developed or not, as shown in Fig. 1. While in grayscale electron beam lithography, varied dose is employed so that structures with different height of remaining resist can be manufactured.

In practice, the remaining height is determined by not only the dose at the corresponding position but also the dose distribution near the position. The reason is that the incident electrons will be scattered in the resist and substrate when the resist is exposed, which is known as the proximity effect. There are two types of scattering: forward scattering and backscattering. The exposure energy distribution E in the resist can be calculated by the convolution of the dose distribution D and the exposure PSF

$$E = D * \text{PSF.} \tag{1}$$

In terms of the exposure point spread function, conventional proximity effect correction methods can be classified into two groups: 2D and 3D proximity correction. 3D correction methods employ a three-dimensional exposure PSF(x, y, z) which can be acquired, for example, through Monte Carlo simulation. 2D proximity correction methods use an approximate two-dimensional double-Gaussian exposure point spread function¹⁸ as

$$PSF(r) = \frac{1}{1+\mu} \left(\frac{1}{\pi \alpha^2} e^{-\frac{r^2}{\alpha^2}} + \frac{\mu}{\pi \beta^2} e^{-\frac{r^2}{\beta^2}} \right),$$
 (2)

where *r* is the radial distance to the exposed point, and μ , α , and β are constants, which depend on the resist, electron energy, and substrate, α (a few nanometers to order of 100 nm) is the forward scattering range parameter, β (a few micrometers) is the backscattering range parameter, and μ is the ratio of the backscattering energy to the forward scattering energy.

^{a)}Electronic mail: wujun.mi@mi.physics.kth.se

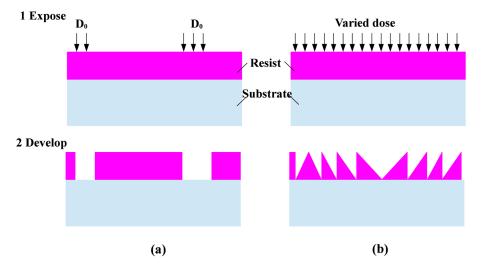


FIG. 1. (Color online) (a) Conventional binary electron beam lithography; (b) grayscale electron beam lithography.

In grayscale electron beam lithography, it is necessary to use a 3D correction method. One reason is that the relative exposure energy due to forward scattering and backscattering varies along the depth direction, especially for thicker resists. This means that the source of the proximity effect switches from being mainly from the forward scattering at the top of the resist, to be mainly due to backscattering at the bottom of the resist. Taking into account the developing process with a developing rate that is nonlinear to the exposed energy further complicates the matter. This phenomenon cannot be modeled using only 2D PSF. On the other hand, 3D methods rely on the 3D exposure point spread function, which is usually derived from Monte Carlo simulation. However, for newly released e-beam resist, like SML 1000, the chemical composition is often unknown, making it difficult to acquire the PSF through simulation. Our method provides true 3D proximity effect correction, while not requiring a full 3D exposure PSF. Instead, a dose correction function is obtained from test structures using machine learning.

III. PROPOSED METHOD

A. Overview of the method

As discussed above, due to proximity effect, the profile of the remaining resist at position (x, y) is determined by not only the dose at position (x, y) but also the nearby dose distribution, i.e., to derive the dose at position (x, y), the profile at the point and the contribution of nearby dose to the point have to be taken into account, which can be expressed as

$$D(x, y) = f(H(x, y), \{D(x_i, y_i); (x_i, y_i) \in N(x, y)\}), \quad (3)$$

where D(x, y) and H(x, y) are the dose and profile at position (x, y), N(x, y) is the neighborhood around the point, large enough to account for all proximity effects.

For a large structure with a uniform profile, the dose and neighboring dose are uniform, which means that the dose can be derived directly from the contrast function. For structures with a variable profile, the dose distribution is obviously nonuniform, which means that, relative to the contrast function,

J. Vac. Sci. Technol. B, Vol. 32, No. 3, May/Jun 2014

the dose needs to be adjusted depending on the contribution from the dose in the neighborhood of the point.

We assume that the adjustment factors can be written in terms of components that are the dose integrated over Gaussian neighborhoods of different size. Since the exposure point spread function in 3D can be written as linear combination of such components [as in Eq. (2)], it is reasonable to assume that the function f in Eq. (3) also can be written in terms of such components. We call this the proximity-corrected dose function (PCDF)

$$D(x, y) = f(H(x, y), C_1(x, y), C_2(x, y), \dots) , \qquad (4)$$

where C_i is the neighborhood dose, which can be expressed as

$$C_i(x,y) = \iint D(x - x', y - y') \times \frac{1}{\pi \alpha_i^2} e^{-\frac{x'^2 + y'^2}{\alpha_i^2}} dx' dy'.$$
(5)

Due to the proximity effect and development, the relationship between dose distribution and remaining height, neighborhood dose is so complicated that it is difficult to get an analytical solution. In this study, we use a machine learning method to fit the relationship.

B. MARS

MARS is a nonparametric regression approach that from training data finds the linear or nonlinear relationship between a set of predictor/input variables and response/output variables and is well suited for the high dimensional regression problems.^{19,20}

MARS uses one-sided truncated functions as the basic elements to complete the fitting process. These are defined as

$$(t-x)_{+} = \begin{cases} t-x & \text{if } x < t, \\ 0 & \text{otherwise,} \end{cases}$$
(6)

$$(x-t)_{+} = \begin{cases} x-t & \text{if } x \ge t, \\ 0 & \text{otherwise,} \end{cases}$$
(7)

where x is one of the variables and t is the so-called knot location. Each function is piecewise linear, and the two functions are called a reflected pair. Initially, a collection of reflected pairs \mathscr{C} is created, in which all values of predictor variables are treated as knots, such that

$$\mathscr{C} = \{ (X_j - t)_+, (t - X_j)_+ \}, \ _{t \in \{x_{1j}, x_{2j}, \dots, x_{Nj}\}, j = 1, 2, \dots, p}.$$
(8)

During the fitting procedure, all reflected pairs in the collection C are selected to fit the data step-wisely as the candidate functions, resulting in the MARS model of the form

$$\hat{f}(X) = a_0 + \sum_{m=1}^{M} a_m B_m(X) ,$$
 (9)

where a_0 is the constant coefficient, $X = (x_1, x_2, ..., x_p)$ is the predictor variables vector, $B_m(X)$ and a_m are the *m*th basic function and the *m*th coefficient, respectively. The basic

Start

function $B_m(X)$ has the form of function in \mathscr{C} or product of two or more such functions.

MARS builds the model in two steps, forward stepwise addition and backward stepwise deletion. In the forward stepwise addition procedure, all the reflected pairs are treated as candidate functions. Starting from a constant basic function, the reflected pair which minimizes the "lack of fit" (LOF) criterion is added to the model in each step and the coefficients are estimated by standard linear regression. The search continues until the number of basic functions reaches a preset value, resulting in an over-fitted model. In the backward stepwise deletion procedure, the basic functions are evaluated according to the LOF criterion and the one with the least contribution to the accuracy of the fit will be deleted in each step, which finally constructs a sequence of M - 1models, in which each has one less basic function than the previous one in the sequence. The optimal model can be selected by comparing the values of the LOF. The LOF is

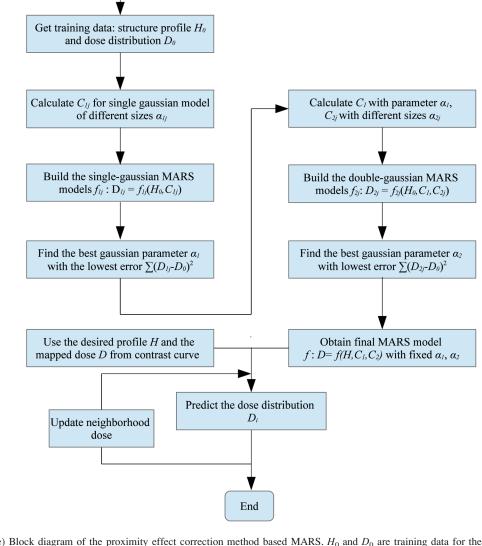


Fig. 2. (Color online) Block diagram of the proximity effect correction method based MARS. H_0 and D_0 are training data for the desired profile and dose distribution.

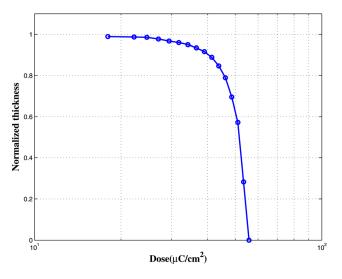


FIG. 3. (Color online) Contrast curve for ZEP 520 resist exposed with 25 keV on 1400 nm-thick resist. It is obtained by manufacturing a series of steps with 20 by 60 μ m, and separated by 10 μ m. The sample was developed in P-xylene for 40 s and resined in IPA for 30 s.

evaluated by a modified form of the generalized crossvalidation criterion and is defined as

$$\text{LOF} = \frac{1}{N} \sum_{i=1}^{N} \left[y_i - \hat{f}_M(x_i) \right]^2 / \left[\frac{1 - C(M)}{N} \right]^2, \quad (10)$$

where M is the number of basic functions (nonconstant), N is the size of the data set, and C(M) is the effective number

of parameters in the trained model. As we can see, the PCDF is a high dimensional function and the relationship between the input variables and the output variable is so complicated that it is difficult to find a predetermined model to fit it. MARS is suitable under this condition and the ARESLab toolbox will be used to build the MARS models in this study.²¹

C. Correction method

Learning the proximity-corrected dose function requires the manufacturing of test structures with known height and dose profiles. We have found sawtooth patterns with varying slopes and depths to be suitable in many cases. The structures should be wide enough to cover the influence of the backscatter lobe, which can be several micrometers wide. From this training data, we determine the proximitycorrected dose function and the sizes of the neighborhood dose functions.

The sizes of the neighborhood-dose Gaussians are determined through a greedy optimization procedure. First, a single Gaussian model is fitted using MARS for a range of variances. The α_1 , which predicts the dose with the lowest sum-of-squares error, is selected. Then, iteratively another Gaussian is added, and its variance is determined by fitting models for a range of values while keeping the previous neighborhood sizes fixed. This is repeated until all the neighborhood's sizes are determined, resulting in the learned proximity-corrected dose function. Figure 2 shows a block diagram of the algorithm.

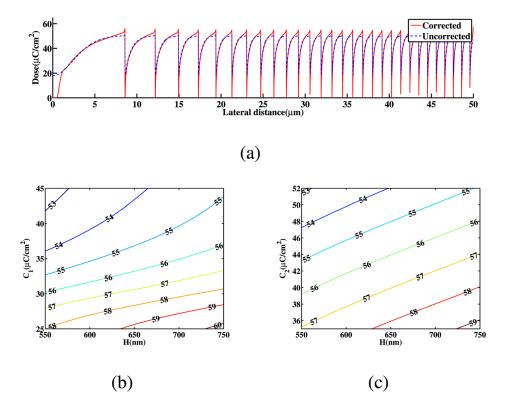


FIG. 4. (Color online) (a) Uncorrected and corrected dose distribution for the Fresnel lens. The uncorrected dose is directly mapped from the contrast curve. (b) and (c) Contour plot of the learned PCDF around the point corresponding to the leftmost peak of the corrected dose in (a) when (b) C_2 is fixed at 43.5 μ C/cm² and (c) C_1 is fixed at 32.8 μ C/cm².

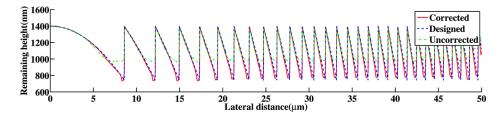


FIG. 5. (Color online) Simulated profile with the uncorrected and corrected dose as well as the desired lens profile.

To compute the dose distribution for a given desired structure, an iterative procedure is employed, as the neighborhood dose initially is unknown. First, the dose is estimated from the contrast function. From this, the neighborhood dose is computed, and a new dose distribution is predicted using the PCDF. This is repeated several times until we obtain the estimated proximity-corrected dose distribution.

IV. SIMULATION RESULT

In order to evaluate the proximity effect correction method, we have designed an e-beam lithography simulator. We used the resist ZEP 520 for which we could get the 3D point spread function through CASINO.^{22,23} The relationship between developing rate and exposure distribution was found by fitting a development rate model to data from manufactured steps. The simulator consists of an exposure calculation model and a development model. The exposure distribution in the resist is computed by the exposure calculation model; and with the energy distribution, the development model based on cell removal method²⁴ is employed to derive the remaining resist profile. The objective of the simulation was to fabricate a Fresnel lens for ultraviolet light, with 50 μ m radius and a focal length of 100 μ m.

As the training data, we chose sawtooth patterns, which are similar to the structures of the Fresnel lens. Here, we used sawtooth structures with depth of 0.2, 0.4, 0.6, 0.8, and 1.0 μ m and width of 0.5, 1, 2, 3, 4, 5, and 6 μ m. The dose for these structures was acquired by directly mapping the sawtooth profile to the contrast curve (Fig. 3). The simulator was used to get the remaining resist profile data after development.

The dose and profiles of the sawtooth structures were used to learn the PCDF using MARS as described in Sec. III. The best values for α_1 and α_2 were found to be 150 nm and 1700 nm, respectively. With the PCDF, the optimal dose for the Fresnel lens was computed in ten iterations, resulting in the proximity-corrected dose in Fig. 4(a). From the figure, we can see that the dose at the deep end of the sidewall has been increased to compensate for the lower contribution from neighboring dose. To illustrate the PCDF, we plot it around the point corresponding to the leftmost peak of the corrected dose, where D, H, C_1 , and C_2 are 56.2 μ C/cm², 655 nm, $32.8 \,\mu\text{C/cm}^2$, and $43.5 \,\mu\text{C/cm}^2$, respectively. Figure 4(b) shows the PCDF when C_2 is fixed at 43.5 μ C/cm², and Fig. 4(c) shows it when C_1 is fixed at 32.8 μ C/cm². We can see that dose increases with increment of depth and decreases with increment of neighborhood dose, as would be expected.

Finally, the development model was employed to get the profile of lens as shown in Fig. 5. The results show some artifacts in the deep ends near the vertical walls, but are otherwise in excellent agreement with the designed profile. The root-mean-square error (RMSE) is 64 nm.

V. EXPERIMENT RESULT

A Fresnel lens was manufactured to verify the correction method experimentally. The lens is the same as in the simulation with the difference that the e-beam resist is now SML 1000.

To get the contrast curve for SML 1000, a series of steps $(20 \times 60 \,\mu\text{m})$, and separated by $10 \,\mu\text{m})$, with gradually increasing dose were exposed using the Raith 150 electron beam lithography system, developed in methyl isobutyl ketone (MIBK) for 20 s and then resined in IPA for 30 s. After that, the steps were measured with a profilometer, resulting in the contrast curve shown in Fig. 6.

Similar with the simulation, the same sawtooth structures were fabricated using the Raith 150 and measured using the AFM to get the profile. Applying the correction method, the best values of α_1 and α_2 for SML 1000 resist are 250 nm and 3400 nm, respectively, and the optimum dose distribution for the Fresnel lens was predicted, as shown in Fig. 7(a). After exposure and development, the lens was measured

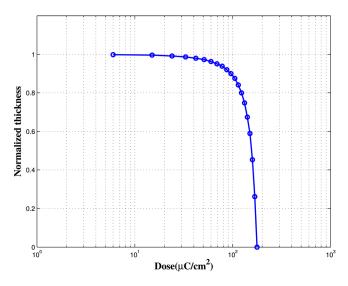


FIG. 6. (Color online) Contrast curve for SML 1000 resist exposed with 25 keV on 1350 nm-thick resist. The sample was developed in MIBK for 20 s and resined in IPA for 30 s.

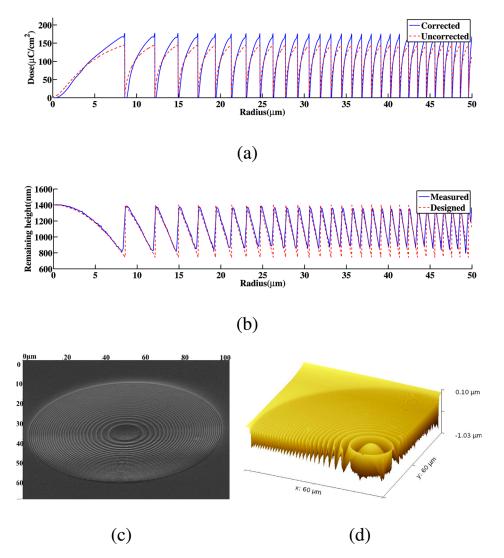


FIG. 7. (Color online) Experiment result with SML 1000 for the Fresnel lens. (a) the dose distribution along the radius; (b) the cross-section profile along the radius and the desired profile; (c) the SEM image of the manufactured Fresnel lens; (d) the AFM image of the lens.

using AFM. Figure 7(b) shows the measured cross-section profile of the Fresnel lens. Figures 7(c) and 7(d) are the SEM and AFM images of the fabricated lens.

From Fig. 7(b), we can see that the slope of the lens agrees well with the desired profile. The main deviation is at the sidewalls which are not perfectly vertical. This could partially be explained by the limitation of the AFM tip. The angles formed by the measured sidewalls and substrate plane are 63° , which is close to the angle of the AFM tip (MPP-11120-10 with an angle of $65^{\circ} \pm 2^{\circ}$). In this situation, it is hard for the tip to reach the bottom of the structures. To illustrate the error accurately, we present the RMSE for the profile with and without the sidewalls, which are 101.8 nm and 26.3 nm, respectively.

VI. SUMMARY

In this study, a three-dimensional proximity effect correction method based on multivariate adaptive regression has been presented. In contrast to conventional methods, this method is independent of the exposure point spread function, which means it can be used to complete the proximity effect correction even if we do not know the chemical composition of the e-beam resist. The method takes exposure distribution and development process into consideration simultaneously to eliminate the influence of the proximity effect. With sawtooth structures as the training data, a proximity-corrected dose function is built to obtain the dose distribution for 3D structures. To verify the method, a Fresnel lens was manufactured by simulation and experiment. From the result, it can be seen that the method is effective to realize proximity effect correction and to fabricate accurate 3D structures.

- ¹N. H. Rizvi and P. Apte, J. Mater. Process. Technol. **127**, 206 (2002).
- ²G. P. Behrmann and M. T. Duignan, Appl. Opt. **36**, 4666 (1997).
- ³C. Waits, B. Morgan, M. Kastantin, and R. Ghodssi, Sens. Actuators, A **119**, 245 (2005).
- ⁴M. Li, L. Chen, and S. Chou, Appl. Phys. Lett. **78**, 3322 (2001).
- ⁵K. Anbumony and S.-Y. Lee, J. Vac. Sci. Technol., B 24, 3115 (2006).
- ⁶F. Hu and S.-Y. Lee, J. Vac. Sci. Technol., B 21, 2672 (2003).
- ⁷S.-Y. Lee, S. Jeon, J. Kim, K. Kim, M. Hyun, J. Yoo, and J. Kim, J. Vac. Sci. Technol., B **27**, 2580 (2009).
- ⁸C. Guo, S.-Y. Lee, S. Lee, B.-G. Kim, and H.-K. Cho, J. Vac. Sci. Technol., B **27**, 2572 (2009).
- ⁹S.-Y. Lee and K. Anbumony, J. Vac. Sci. Technol., B 25, 2008 (2007).

¹⁰Q. Dai, S.-Y. Lee, S. Lee, B.-G. Kim, and H.-K. Cho, Microelectron. Eng. 88, 902 (2011).

- ¹¹Q. Dai, S.-Y. Lee, S.-H. Lee, B.-G. Kim, and H.-K. Cho, J. Vac. Sci. Technol., B **29**, 06F314 (2011).
- ¹²Q. Dai, S.-Y. Lee, S.-H. Lee, B.-G. Kim, and H.-K. Cho, J. Vac. Sci. Technol., B **30**, 06F307 (2012).
- ¹³N. Unal, D. Mahalu, O. Raslin, D. Ritter, C. Sambale, and U. Hofmann, Microelectron. Eng. 87, 940 (2010).
- ¹⁴Y. Hirai, H. Kikuta, M. Okano, T. Yotsuya, and K. Yamamoto, Proceedings of the 2000 International Microprocesses and Nanotechnology Conference (2000), pp. 160–161.
- ¹⁵M. Okano, Y. Hirai, and H. Kikuta, Jpn. J. Appl. Phys., Part 1 46, 627 (2007).
- ¹⁶M. Osawa, K. Ogino, H. Hoshino, Y. Machida, and H. Arimoto, J. Vac. Sci. Technol., B 22, 2923 (2004).

- ¹⁷K. Ogino, H. Hoshino, and Y. Machida, J. Vac. Sci. Technol., B 26, 2032 (2008).
- ¹⁸T. Chang, J. Vac. Sci. Technol. **12**, 1271 (1975).
- ¹⁹J. H. Friedman, Ann. Stat. **19**, 1 (1991).
- ²⁰T. Hastie, R. Tibshirani, and J. H. Friedman, *The Elements of Statistical Learning* (Springer, New York, 2001), Vol. 1.
- ²¹G. Jekabsons, Version 1.5.1, 2010.
- ²²D. Drouin, A. R. Couture, D. Joly, X. Tastet, V. Aimez, and R. Gauvin, Scanning **29**, 92 (2007).
- ²³H. Demers, N. Poirier-Demers, A. R. Couture, D. Joly, M. Guilmain, N. de Jonge, and D. Drouin, Scanning 33, 135 (2011).
- ²⁴Y. Hirai, S. Tomida, K. Ikeda, M. Sasago, M. Endo, S. Hayama, and N. Nomura, IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst. **10**, 802 (1991).