

Iterative Deconvolution for Exposure Planning in Scanning Laser Lithography

Omid T. Ghalehbeygi¹, John O'Connor², Ben S. Routley¹ and Andrew J. Fleming¹

Abstract—Laser scanning lithography is a maskless method for exposing photoresist during semiconductor manufacturing. In this method, the power of a focused beam is modulated while scanning the photoresist. This article describes an iterative deconvolution method for determining the exposure pattern. This approach is computationally efficient as there is no gradient calculation. Simulations demonstrate the accurate fabrication of a feature with sub-wavelength geometry.

I. INTRODUCTION

Lithography is a key process in semiconductor manufacturing and accounts for approximately half of the production costs [1]. To address the high cost of mask fabrication and exposure infrastructure, methods for maskless lithography are under development. A number of maskless lithography processes such as scanning electron beam lithography and ion beam lithography are already commercially available [2].

Scanning laser lithography is a cost-effective maskless method which has the advantage of using standard photoresist and process chemistry. In this method, a focused laser is scanned over the photoresist while modulating the beam power. Improvements in scanning methods [3], [4] and positioning bandwidth [5]–[8] have significantly improved the feasibility of this approach.

To increase the resolution of scanning laser lithography, probe-based enhancement has been demonstrated to overcome the diffraction limit [9], [10]. Feature sizes of 30-nm have been reported using wavelengths in the visible spectrum of laser light [11]. The foremost difficulties associated with this technique are the low throughput and complicated exposure planning.

The problem of exposure planning is shared by all maskless lithography methods that involve a scanning beam. The exposure plan defines the position and intensity of the laser, electron, or ion beam. The objective is to find an exposure plan that minimizes differences between the desired and predicted feature. This class of problem is widely referred to as inverse lithography [12]–[18]. Nonlinear programming methods [13], [19], and gradient-based algorithms [12], [18] have both applied to find exposure plans. However, these methods are computationally intensive as they require the first and second-derivative calculation for function with millions or billions of variables.

*This work was supported by the Australian Research Council Grants DP150103521, FT130100543, and DP120100487.

¹Omid T. Ghalehbeygi, Andrew J. Fleming, and Ben S. Routley are with the Faculty of Engineering and Built Environment, University of Newcastle, Callaghan, NSW 2308, Australia.

²John O'Connor is with the The School of Mathematical and Physical Sciences, University of Newcastle, Australia

This article investigates the use of iterative deconvolution for solving the inverse lithography problem. The iterative deconvolution method was proposed in 1961 [20] for inverting the response of physical sensors. This method has also been applied to the inversion of other scattering and convolution problems [20]–[23]. Iterative deconvolution is derivative-free and therefore significantly less computationally intensive than previously proposed methods [12], [13], [18], [19].

II. MODELING PROCESSES

In this section, the scanning laser lithography process is modeled as a convolution operation followed by a non-linear operator. The photoresist layer is assumed to be sufficiently thin so that the beam profile remains constant throughout its depth. The optical properties of the film, which are a function of the exposure state, are also assumed to be constant. Other optical effects such as scattering and cavity formation are ignored.

A. Beam Profile

The light intensity (in W/m^2) at the focal point of the objective lens can be analytically expressed as [24]

$$B(x, y) = \frac{2P}{\pi w_0^2} e^{-\frac{2(x^2+y^2)}{w_0^2}}, \quad (1)$$

where x and y represent the transverse axes of the beam at focal point w_0 , and P is the power.

This normalized intensity is illustrated in Fig. 1, where the focal point is located in the center of the workspace (B_ϕ).

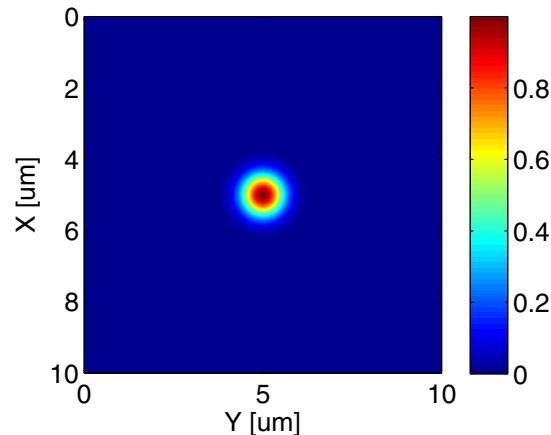


Fig. 1. The normalized intensity of a Gaussian beam, where the focal point is located in the middle of the workspace. The beam width is $w_0 = 450$ nm.

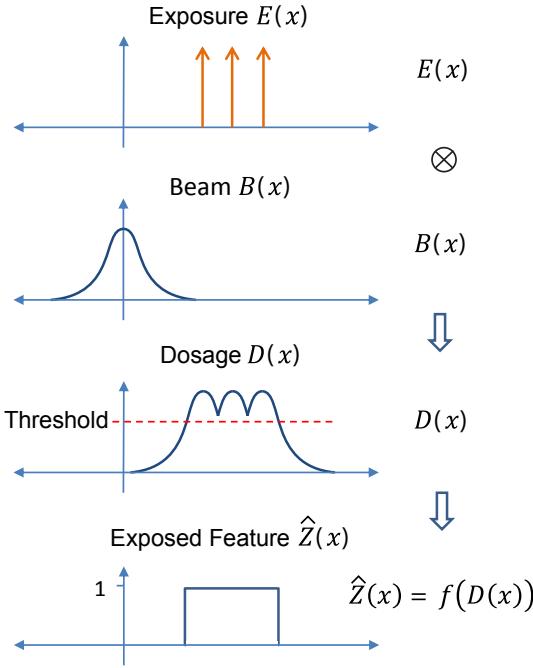


Fig. 2. A simplified one-dimensional model of scanning laser lithography. In this example, the exposure pattern $E(x)$ is three discrete exposures of equal energy. The resulting dosage $D(x)$ is the sum of each exposure point convolved with the beam profile $B(x)$. Finally, the photoresist function $f(\sigma)$ maps the cumulative dosage $D(x)$ to the predicted feature $\hat{Z}(x)$.

B. Continuous Exposure Modeling

A one-dimensional model of the exposure process along the x -axis (i.e. $y = 0$) is illustrated in Fig. 2. The exposure profile $E(x)$ represents the energy delivered at a position x . In this work, the exposure energy is modulated by controlling the time interval for which the laser shutter is open. Since the beam power is constant, the time interval is proportional to the resulting dosage. Other possibilities include modulating the beam power or the scanning speed.

The light intensity (in W/m^2) is a Gaussian function described in 1. To calculate the dosage $D(x)$ (in J/m^2) at a single point, the intensity is multiplied by the exposure time, that is $D(x) = t_{on}B(x, 0)$. Where multiple exposures t_i are involved at arbitrary locations x_i , the total dosage is

$$D(x) = \sum_{i=1}^N t_i B(x - x_i, 0). \quad (2)$$

The above equation is a convolution operation which can be generalized to discrete or continuous exposures in one or more dimensions. That is, in general

$$D(x, y) = E(x, y) \otimes B(x, y). \quad (3)$$

where \otimes is the convolution operator. When the exposure function is discrete, the dosage can be expressed as

$$D(x, y) = \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} E_{i,j} B(x - x_i, y - y_j), \quad (4)$$

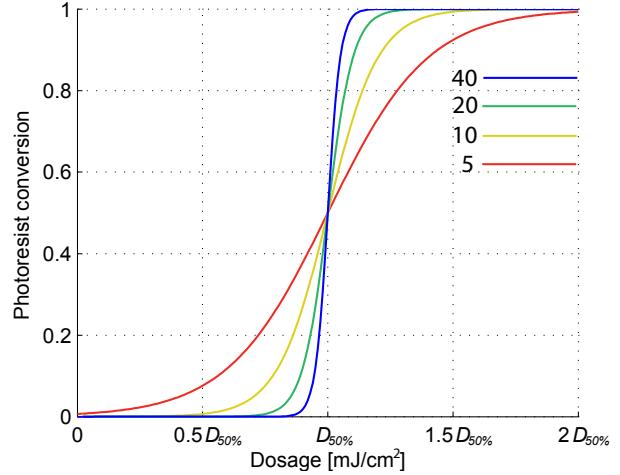


Fig. 3. The sigmoid function representing the percentage of converted photoresist versus cumulative dosage. The steepness parameter is varied from $\alpha = 5$ to 40

where the i, j element of the matrix $E \in \mathbb{R}^{N_x \times N_y}$ represents the exposure energy at a location $x = \mathbf{x}_i$ and $y = \mathbf{y}_j$, where $\mathbf{x} \in \mathbb{R}^{N_x}$ and $\mathbf{y} \in \mathbb{R}^{N_y}$.

C. Photoresist Modeling

The photoresist model describes the fraction of chemical conversion as a function of dosage. The simplest model is a threshold function which indicates 100% conversion when the pixel's dose is above a threshold. For example,

$$\hat{Z}(x, y) = \begin{cases} 1 & D(x, y) \geq T \\ 0 & D(x, y) < T \end{cases}, \quad (5)$$

where \hat{Z} is the predicted exposure result, and T is the threshold.

In practice, the photoresist conversion is a continuous function of dosage. The sigmoid function is an approximation of this process

$$\hat{Z}(x, y) = f(D(x, y)) = \frac{1}{1 + e^{-\alpha(D(x, y) - D_{50\%})}} \quad (6)$$

where the parameter α dictates the steepness of the sigmoid, and $D_{50\%}$ is the dosage where half of the photoresist is converted. When α is large, the sigmoid approaches the threshold model. Fig. 3 illustrates the behavior of the sigmoid function with different values of α .

D. Discrete Exposure Modeling

To implement deconvolution, the functions for exposure, beam profile and dosage are replaced by matrices that represent these functions at discrete locations. The workspace is an $N \times N$ square grid of points along the x and y axes with a spatial resolution of Δ , that is

$$\mathbf{x} = \mathbf{y} = [0, \Delta, 2\Delta, \dots, (N-1)\Delta], \quad (7)$$

The exposure and dosage matrices are defined by $E \in \mathbb{R}^{N \times N}$ and $D \in \mathbb{R}^{N \times N}$ respectively. The peak light intensity is normalized to one for simplicity. That is, the element

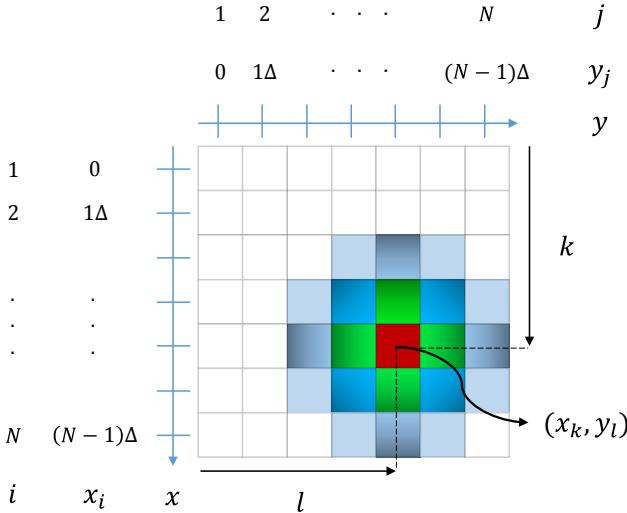


Fig. 4. The dosage matrix resulting from a single exposure at (x_k, y_l) .

$E_{k,l}$ represents the exposure energy at location (x_k, y_l) , where $E_{k,l}$ refers to the k^{th} row and l^{th} column. The dosage D and predicted feature \hat{Z} and the desired Z have the same structure.

The beam profile matrix $B^{k,l} \in \mathbb{R}^{N \times N}$ is the beam power over the workspace for a focal point located at x_k, y_l and $B_{i,j}^{k,l}$ indicates the intensity at the location (x_i, y_j) . Thus, the array of beam profile matrices is

$$B_{i,j}^{k,l} = \frac{2P}{\pi w_0^2} \exp\left(-\frac{2(x_i - x_k)^2 + (y_j - y_l)^2}{w_0^2}\right) \quad (8)$$

for $i, j = 1, \dots, N$, and, $k, l = 1, \dots, N$.

The dosage D represents the cumulative energy per unit area, and is calculated by multiplying the exposure energy at a point by the corresponding beam profile matrix, $B^{k,l}$. The dosage D matrix is

$$D = \sum_{k=1}^N \sum_{l=1}^N E_{k,l} B^{k,l}, \quad (9)$$

where an individual element is

$$D_{i,j} = \sum_{k=1}^N \sum_{l=1}^N E_{k,l} B_{i,j}^{k,l}. \quad (10)$$

Fig. 4 illustrates the matrix D , when E contains a single non-zero entry at (x_k, y_l) . The D matrix is obtained by multiplying the matrix $B^{k,l}$ by the scalar $E_{k,l}$.

For convenience, the matrices are vectorized by stacking the rows. That is, the “vec” operator is

$$\text{vec}\{\mathbf{E}\} \triangleq \begin{bmatrix} \mathbf{E}_{:,1} \\ \mathbf{E}_{:,2} \\ \vdots \\ \mathbf{E}_{:,N} \end{bmatrix}, \quad (11)$$

where Matlab notation is used and $\mathbf{E}_{:,k}$ refers to column k of the matrix \mathbf{E} . As a vector, the exposure matrix becomes

$$\mathbf{e} \triangleq \text{vec}\{\mathbf{E}\}. \quad (12)$$

The dosage matrix can also be vectorized $\mathbf{d} \triangleq \text{vec}\{\mathbf{D}\}$ so that convolution operation can be replaced with the multiplication. To do this, it is required to define an $N \times N$ array of $N \times N$ matrices Ω , where the columns of Ω are the vectorized versions of $B^{k,l}$, that is

$$\Omega = [\text{vec}\{\mathbf{B}^{1,1}\}, \dots, \text{vec}\{\mathbf{B}^{N,1}\}, \text{vec}\{\mathbf{B}^{1,2}\}, \dots, \text{vec}\{\mathbf{B}^{N,N}\}]. \quad (13)$$

Equation (9) can be rewritten By Using this definition as

$$\mathbf{d} = \Omega \mathbf{e}, \quad (14)$$

where $\Omega \in \mathbb{R}^{N^2 \times N^2}$ and $\mathbf{d}, \mathbf{e} \in \mathbb{R}^{N^2}$.

The vectorized predicted feature \hat{Z} can be estimated by applying the thresholding function (5) element wise to \mathbf{d} .

Finally, the original form of the matrices \mathbf{E} , \mathbf{D} , and \hat{Z} can be reconstructed by reshaping the vectors \mathbf{e} , \mathbf{d} , and \hat{Z} respectively.

III. ITERATIVE DECONVOLUTION

The aim of iterative deconvolution is to find a vector e with non-negative entries that satisfies (14). This approach is described in the following subsection with a one-dimensional example.

A. One Dimensional Example

Consider the convolution of two vectors, $\mathbf{y} = \mathbf{h} \otimes \mathbf{x}$, where the entries of \mathbf{x} cannot be negative. This can be expressed in matrix form,

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_M \end{bmatrix} = \begin{bmatrix} h_1 & h_M & \dots & h_2 \\ h_2 & h_1 & \dots & h_3 \\ \vdots & \vdots & \ddots & \vdots \\ h_M & h_{M-1} & \dots & h_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_M \end{bmatrix}, \quad (15)$$

$$\mathbf{y} = \mathbf{H} \mathbf{x}, \quad (16)$$

where \mathbf{H} is a circulant matrix derived from the vector \mathbf{h} . A direct solution to (16) is not possible due to singularities in \mathbf{H} [25], [26] and the non-negative constraint on \mathbf{x} . Iterative deconvolution is a method for solving ill conditioned inverse problems. This method has been used in applications such as spectroscopy [27] and image processing [28].

The algorithm proposed by Van Cittert in 1930 [29] is

$$\hat{\mathbf{x}}^{(n+1)} = \hat{\mathbf{x}}^{(n)} + \mu(\mathbf{y} - \mathbf{H} \hat{\mathbf{x}}^{(n)}), \quad (17)$$

where $\hat{\mathbf{x}}^{(n)}$ is the n^{th} iteration of the solution and μ is a relaxation factor of convergence. To guarantee convergence, it is required that $0 \leq \mu_i \leq 2/\lambda_i$, where λ_i is the eigenvalue of \mathbf{H} at the i^{th} row. This inequality implies that the \mathbf{H} matrix should be positive definite [27]. As \mathbf{H} is a symmetric matrix, its eigenvalues are real; in turn, $\mathbf{H}^T \mathbf{H}$ becomes a positive definite matrix because its eigenvalues are the

eigenvalues of \mathbf{H} , squared. Thus, Van Cittert linear iterative algorithm expressed as

$$\hat{\mathbf{x}}^{(n+1)} = \hat{\mathbf{x}}^{(n)} + \mu(\mathbf{H}^T \mathbf{y} - \mathbf{H}^T \mathbf{H} \hat{\mathbf{x}}^{(n)}), \quad (18)$$

A common method for choosing μ is to guaranteeing convergence by using the maximum eigenvalue λ_{max} ; however, this significantly reduces convergence speed.

In reference [26], the convergence speed was increased by using a local variable relaxation factor, described by

$$\mu_i = \frac{\hat{x}_i^{(n)}}{(\mathbf{H}^T \mathbf{H})_i \hat{x}^{(n)}} \quad (19)$$

By inserting (19) into (18), an equation known as the Gold deconvolution algorithm is obtained, that is

$$\hat{\mathbf{x}}^{(n+1)} = \hat{\mathbf{x}}^{(n)} \odot \frac{\mathbf{H}^T \mathbf{y}}{\mathbf{H}^T \mathbf{H} \hat{\mathbf{x}}^{(n)}}, \quad (20)$$

$$\hat{\mathbf{x}}^{(n+1)} = \hat{\mathbf{x}}^{(n)} \odot \frac{\mathbf{y} \otimes \mathbf{h}}{\mathbf{h} \otimes (\hat{\mathbf{x}}^{(n)} \otimes \mathbf{h})}, \quad (21)$$

where \odot indicates the element wise multiplication.

B. Iterative Exposure Deconvolution

The Gold iterative deconvolution method in (21) can be applied directly to find a non-negative exposure vector e which satisfies (14). That is,

$$\hat{\mathbf{e}}^{(n+1)} = \hat{\mathbf{e}}^{(n)} \odot \frac{\Omega^T \mathbf{d}}{\Omega^T \Omega \hat{\mathbf{e}}^{(n)}}, \quad (22)$$

The above algorithm can also be rewritten when the arguments are two dimensional matrices,

$$\hat{\mathbf{E}}^{(n+1)} = \hat{\mathbf{E}}^{(n)} \odot \frac{\mathbf{D} \otimes \mathbf{B}_\phi}{\mathbf{B}_\phi \otimes \underbrace{(\hat{\mathbf{E}}^{(n)} \otimes \mathbf{B}_\phi)}_{\hat{\mathbf{D}}}}. \quad (23)$$

where $\hat{\mathbf{D}}$ is the predicted dosage at each iteration and \mathbf{D} is the desired dosage resulting from the desired feature \mathbf{Z} . \mathbf{D} can be found by solving (6) for \mathbf{D} , that is

$$\mathbf{D} = D_{50\%} - \left[\frac{1}{\alpha} \odot \ln \left(\frac{1}{\mathbf{Z}} - 1 \right) \right]. \quad (24)$$

IV. SIMULATION RESULTS

In the simulation, the threshold energy is $D_{50\%} = 0.5$, and the Gaussian beam has a width of 450 nm, as illustrated in Fig. 1. The workspace is $10 \mu\text{m} \times 10 \mu\text{m}$ with a resolution of 100 nm, which results in a 100×100 matrix.

The desired feature is illustrated in the right column of Fig. 5, where the values are either zero or 0.9, which represents a zero or 90% photoresist conversion. The initial exposure condition $\hat{\mathbf{E}}^0$ is a scaled version of the desired feature. As expected, the dosage produced by the initial exposure pattern in Fig. 5(b) results in a gross over-exposure in Fig. 5(c). However, after 50 and 10,000 iterations, the difference between the desired and predicted feature is reduced. The evolution of the error versus iteration is plotted in Fig. 6.

For 10,000 iterations, the total calculation time is 30 seconds, which is significantly faster than the gradient descent method [12], which was 90 seconds for the same problem size.

V. CONCLUSION

This article describes the application of iterative deconvolution to inverse lithography. Due to the computational simplicity, this method is significantly faster and requires less memory than competing methods. Therefore, the proposed method is well suited to large inverse lithography problems.

Current research involves experimental application, and the derivation of an optimal iteration step size and termination condition.

REFERENCES

- [1] L. R. Harriott, "Limits of lithography," *Proceedings of the IEEE*, vol. 89, no. 3, pp. 366–374, Mar. 2001.
- [2] M. Altissimo, "E-beam lithography for micro-nanofabrication," *Biomicrofluidics*, vol. 4, no. 2, p. 026503, 2010.
- [3] Y. R. Teo, Y. K. Yong, and A. J. Fleming, "A comparison of scanning methods and the vertical control implications for scanning probe microscopy," *Asian Journal of Control*, vol. 19, no. 2, pp. 1–15, Mar. 2017.
- [4] A. J. Fleming and A. G. Wills, "Optimal periodic trajectories for band-limited systems," *IEEE Transactions on Control Systems Technology*, vol. 13, no. 3, pp. 552–562, May 2009.
- [5] A. J. Fleming and Y. K. Yong, "An ultra-thin monolithic xy nanopositioning stage constructed from a single sheet of piezoelectric material," *IEEE/ASME Transactions on Mechatronics*, vol. 22, no. 6, pp. 2611–2618, Dec. 2017.
- [6] M. Ratnam, B. Bhikkaji, A. J. Fleming, and S. O. R. Moheimani, "PPF control of a piezoelectric tube scanner," in *IEEE Conference on Decision and Control and European Control Conference*, Seville, Spain, Dec. 2005.
- [7] Y. R. Teo, A. A. Eielsen, J. T. Gravdahl, and A. J. Fleming, "A simplified method for discrete-time repetitive control using model-less fir filter inversion," *Journal of Dynamic Systems, Measurement and Control*, vol. 138, no. 8, pp. 1–13, Aug. 2016.
- [8] A. A. Eielsen, Y. R. Teo, and A. J. Fleming, "Improving robustness filter bandwidth in repetitive control by considering model mismatch," *Asian Journal of Control*, vol. 19, no. 4, pp. 1–11, Jul. 2017.
- [9] M. Born, E. Wolf, and A. Bhatia, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*. Cambridge University Press, 1999.
- [10] B. S. Routley, J. L. Holdsworth, and A. J. Fleming, "Optimization of near-field scanning optical lithography," in *SPIE Advanced Lithography*, San Jose, CA, 2015.
- [11] Y. Lin, M. Hong, W. Wang, Y. Law, and T. Chong, "Sub-30nm lithography with near-field scanning optical microscope combined with femtosecond laser," *Applied Physics A*, vol. 80, no. 3, pp. 461–465, 2005.
- [12] O. T. Ghalehbeygi, G. Berriman, A. J. Fleming, and J. L. Holdsworth, "Optimization and simulation of exposure pattern for scanning laser lithography," in *IEEE Multiconference on Systems and Control*, Sydney, 2015.
- [13] A. J. Fleming, A. G. Wills, O. T. Ghalehbeygi, B. S. Routley, and B. Ninness, "A nonlinear programming approach to exposure optimization in scanning laser lithography," in *American Control Conference*, Boston, MA, 2016.
- [14] A. Poonawala and P. Milanfar, "Mask design for optical microlithography – an inverse imaging problem," *Trans. Img. Proc.*, vol. 16, no. 3, pp. 774–788, Mar. 2007.
- [15] F. Liu and X. Shi, "An efficient mask optimization method based on homotopy continuation technique," in *2011 Design, Automation Test in Europe*, March 2011, pp. 1–6.
- [16] L. Pang, G. Dai, T. Cecil, T. Dam, Y. Cui, P. Hu, D. Chen, K.-H. Baik, and D. Peng, "Validation of inverse lithography technology (ilt) and its adaptive sraf at advanced technology nodes," pp. 69 240T–69 240T–12, 2008.
- [17] Y. Granik, "Fast pixel-based mask optimization for inverse lithography," *Journal of Micro/Nanolithography, MEMS, and MOEMS*, vol. 5, no. 4, pp. 043 002–043 002–13, 2006.
- [18] O. T. Ghalehbeygi, A. G. Wills, B. S. Routley, and A. J. Fleming, "Gradient-based optimization for efficient exposure planning in maskless lithography," *Journal of Micro/Nanolithography, MEMS, and MOEMS*, vol. 16, pp. 16 – 16 – 9, 2017.

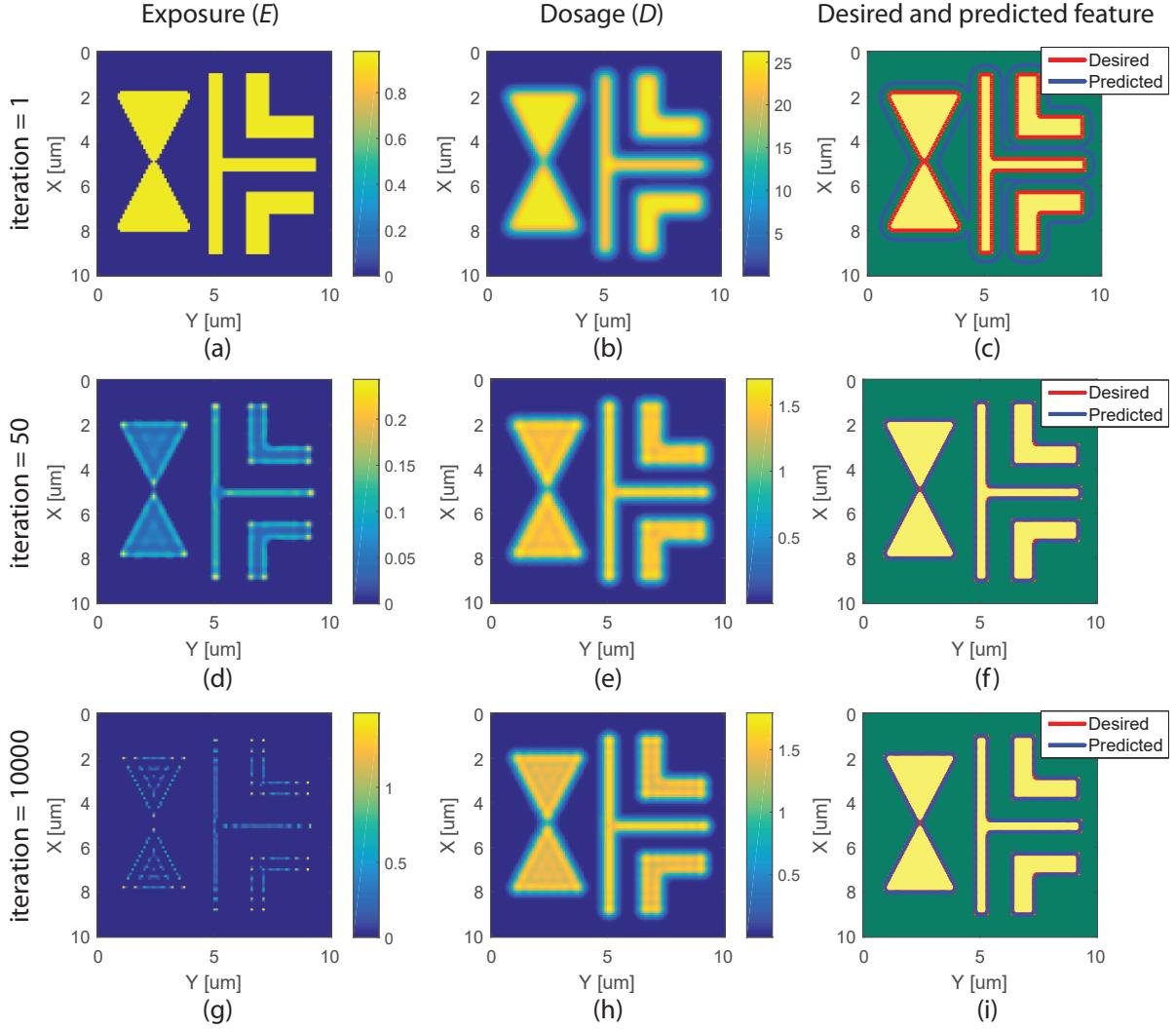


Fig. 5. The result of iterative deconvolution with 1, 50, and 10000 iterations. The exposure, resulting dosage, and predicted feature are plotted in the left, middle, and right column. The color intensity represents the exposure energy in the left column and energy density in the middle column.

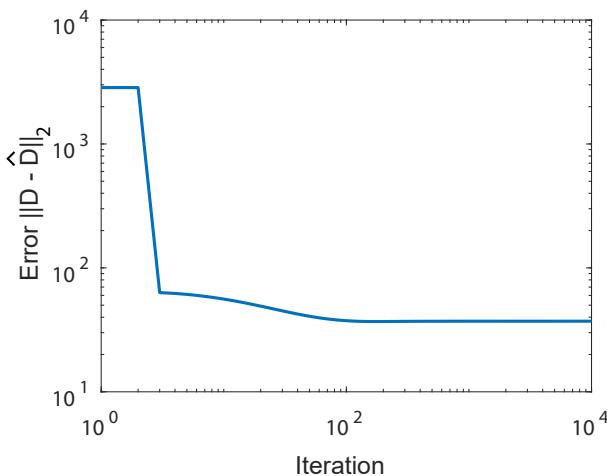


Fig. 6. The error between the desired dosage D and predicted \hat{D} after 10000 iterations. The error was plotted in logarithmic scale.

- [19] A. J. Fleming, O. T. Ghalehbeygi, B. S. Routley, and A. G. Wills, "Experimental scanning laser lithography with exposure optimization," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 8662 – 8667, 2017.
- [20] L. D. Skarsgard, H. E. Johns, and L. E. S. Green, "Iterative Response Correction for a Scintillation Spectrometer," *Radiation Research*, vol. 14, no. 3, pp. 261–280, 1961.
- [21] P. E. Siska, "Iterative unfolding of intensity data, with application to molecular beam scattering," *The Journal of Chemical Physics, Volume 59, Issue 11, p.6052-6060 (JChPh Homepage)*, December 1973.
- [22] L. B. Lucy, "An iterative technique for the rectification of observed distributions," *The astronomical journal*, vol. 79, p. 745, 1974.
- [23] M. K. Khan, S. Morigi, L. Reichel, and F. Sgallari, "Iterative methods of richardson-lucy-type for image deblurring," *Numerical Mathematics: Theory, Methods and Applications*, vol. 6, no. 1, p. 262275, 2013.
- [24] NEWPORT. (2017, March) Gaussian beam optics.
- [25] R. Gold, *An Iterative Unfolding Method For Response Matrices*. Argonne National Laboratory, Dec 1964.
- [26] C. Xu, I. Aissaoui, and S. Jacquey, "Algebraic analysis of the van cittert iterative method of deconvolution with a general relaxation factor," *J. Opt. Soc. Am. A*, vol. 11, no. 11, pp. 2804–2808, Nov 1994.
- [27] P. Banduch, M. Morh, and J. Kritiak, "Study of the van cittert and gold iterative methods of deconvolution and their application in the deconvolution of experimental spectra of positron annihilation," *Nuclear Instruments and Methods in Physics Research Section A*:

- Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 384, no. 2, pp. 506 – 515, 1997.
- [28] M. B. Cannell, A. McMorland, and C. Soeller, “Image enhancement by deconvolution,” *Handbook of biological confocal microscopy*, pp. 488–500, 2006.
 - [29] P. H. van Cittert, “Zum Einfluß der Spaltbreite auf die Intensitätsverteilung in Spektrallinien. II,” *Zeitschrift für Physik*, vol. 69, pp. 298–308, May 1931.