

論文 / 著書情報
Article / Book Information

Citation	Hiroyoshi Tanabe, Shimpei Sato, Atsushi Takahashi "Fast 3D lithography simulation by convolutional neural network",Proc. SPIE 11614, Design-Process-Technology Co-optimization XV 2021, 116140M,,(2021, 2),doi: https://doi.org/10.1117/12.2583683
Copyright notice	Copyright 2021 Society of Photo-Optical Instrumentation Engineers (SPIE). One print or electronic copy may be made for personal use only. Systematic reproduction and distribution, duplication of any material in this paper for a fee or for commercial purposes, or modification of the content of the paper are prohibited.

PROCEEDINGS OF SPIE

SPIDigitalLibrary.org/conference-proceedings-of-spie

Fast 3D lithography simulation by convolutional neural network

Tanabe, Hiroyoshi, Sato, Shimpei, Takahashi, Atsushi

Hiroyoshi Tanabe, Shimpei Sato, Atsushi Takahashi, "Fast 3D lithography simulation by convolutional neural network," Proc. SPIE 11614, Design-Process-Technology Co-optimization XV, 116140M (22 February 2021); doi: 10.1117/12.2583683

SPIE.

Event: SPIE Advanced Lithography, 2021, Online Only

Fast 3D lithography simulation by convolutional neural network

Hiroyoshi Tanabe*, Shimpei Sato, Atsushi Takahashi

Tokyo Institute of Technology, 2-12-1 Ookayama, Meguro-ku, Tokyo 152-8550 Japan

ABSTRACT

Thin mask model has been conventionally used in optical lithography simulation. In this model the diffracted waves from the mask are assumed to be Fourier transform of the mask pattern. In EUV (Extreme UltraViolet) lithography thin mask model is not valid because the absorber thickness is comparable to the mask pattern size. Fourier transformation is not suitable for calculating the diffracted waves from thick masks. Rigorous electromagnetic simulations such as finite-difference time-domain method, rigorous coupled wave analysis and 3D waveguide model are used to calculate the diffracted waves from EUV masks. However, these simulations are highly time consuming. We reduce the calculation time by adapting a CNN (Convolutional Neural Network).

We calculate the far-field diffraction amplitudes from an EUV mask by using the 3D waveguide model. We divide the diffraction amplitudes into the thin mask amplitudes (Fourier transform of the mask pattern) and the residual mask 3D amplitudes. The incident angle dependence of the mask 3D amplitude for each diffraction order is fitted by using three parameters which represent the on-axis and the off-axis mask 3D effects. We train a CNN where the inputs are 2D mask patterns and the targets are the mask 3D parameters of all diffraction orders. After the training, the CNN successfully predict the mask 3D parameters. The CNN prediction is 5,000 times faster than the electromagnetic simulation.

We extend the transmission cross coefficient formula to include the off-axis mask 3D effects. Our formula is applicable to arbitrary source shapes and defocus. We can use the eigen value decomposition method to accelerate the calculation.

Keywords: lithography simulation, neural network, EUV mask

1. INTRODUCTION

The adoption of EUV lithography began in the mass production of logic and memory devices. Typical resolution of EUV lithography is 16 nm, assuming a moderate k_1 factor of 0.4. In this case the minimum mask pattern size is 64 nm. Standard thickness of Ta absorber is around 60 nm and the aspect ratio of the mask pattern is nearly 1. High aspect absorbers induce several mask 3D effects such as CD (Critical Dimension) errors and pattern shifts.^{1,2} They also change the effective phase-shift value near the absorber edge.³ It is necessary to include the mask 3D effects in EUV lithography simulation.

Rigorous electromagnetic simulation methods such as FDTD (Finite-Difference Time-Domain) method⁴, RCWA⁵ (Rigorous Coupled Wave Analysis) and 3D waveguide model⁶ accurately calculate the mask 3D effects. However, these simulations are highly time consuming. Semi empirical and fast methods have been developed to include the mask 3D effects approximately in the calculation.⁷⁻¹¹ Recently with the advent of the deep learning software and hardware platforms some attempts have been made to solve the mask 3D effect problems by using deep neural networks.¹²⁻¹⁴ The targets of the deep neural networks in these models are the near-field diffraction amplitudes of thick masks calculated by electromagnetic simulations. The near-field amplitudes are described in coordinate space.

Our model also uses a deep neural network to accelerate the calculation of mask 3D effects. Here we define the mask 3D amplitudes as the difference between the far-field diffraction amplitudes from a thick mask and those from a thin mask. Since the far-field diffraction amplitudes are described in momentum space, our model can be easily incorporated into TCC (Transmission Cross Coefficient) method¹⁵ and SOCS (Some Of Coherent Systems) model¹⁶ which is conventionally used in optical lithography simulations.

In Sec.2 we parametrize the mask 3D amplitudes as functions of diffraction orders and source positions. In Sec. 3 we construct a convolutional neural network which predicts the diffraction amplitudes from 2D EUV mask patterns. In Sec. 4 we extend the TCC method to include the off-axis mask 3D effects

[*tanabe.h.af@m.titech.ac.jp](mailto:tanabe.h.af@m.titech.ac.jp)

2. PARAMETRIZATION OF THE DIFFRACTION AMPLITUDE

One of the difficulties in the formulation of the 3D electromagnetic simulation is the number of the variables. The electric field has three components E_x , E_y and E_z while the light has only two polarizations. Therefore, one of the three electric field components is redundant to parametrize the diffraction amplitudes. 3D waveguide model⁶ uses the vector potential instead of the electromagnetic field. The model contains only two variables A_x and A_y which correspond to two polarizations. In the model A_z is fixed to be zero after gauge transformation.

Figure 1. shows a schematic view of light diffraction by an EUV mask. The light emitted from the secondary source at the position (s_x, s_y) illuminates the mask from the oblique direction (s_x, s_y) . The light is diffracted by the mask and it is scattered to the direction (k_x, k_y) . The diffracted light passes the pupil at the position (k_x, k_y) .

The diffraction momentum $(k_x - s_x, k_y - s_y)$ is the difference between the outgoing momentum (k_x, k_y) and the incident momentum (s_x, s_y) . Here we specify only x and y components of the momentum because z component obeys the energy conservation law.

As shown in Ref. 17 the diffraction amplitude $\mathbf{A}$ of a thick mask is decomposed into the thin mask amplitude $\mathbf{A}^{\text{FT}}$ and the mask 3D amplitude $\mathbf{A}^{\text{3D}}$. The thin mask amplitude $\mathbf{A}^{\text{FT}}$ is calculated from the Fourier transformation of the mask pattern and it depends on the diffraction momentum $(k_x - s_x, k_y - s_y)$.

$$\mathbf{A}(k_x, k_y; s_x, s_y) = \mathbf{A}^{\text{FT}}(k_x - s_x, k_y - s_y) + \mathbf{A}^{\text{3D}}(k_x, k_y; s_x, s_y). \quad (1)$$

The mask 3D amplitude $\mathbf{A}^{\text{3D}}$ is defined as the difference between the thick mask amplitude $\mathbf{A}$ and the thin mask amplitude $\mathbf{A}^{\text{FT}}$. The electromagnetic calculation of the thick mask amplitude is heavy. In Ref. 17 we fitted the mask 3D amplitudes of 1D mask patterns by using a CNN, which significantly reduced the computation time.

In this study we use 3D waveguide model to calculate the thick mask amplitudes of 2D mask patterns. We assume a periodic boundary condition with the mask pattern pitch L (Fig. 2). In this case the diffraction momentum has a discrete value,

$$k_x - s_x = l \frac{2\pi}{L}, \quad k_y - s_y = m \frac{2\pi}{L}, \quad (2)$$

where (l, m) represents the diffraction order. For convenience we also discretize the incident momentum as follows.

$$s_x = l_s \frac{2\pi}{L}, \quad s_y = m_s \frac{2\pi}{L}, \quad (3)$$

where (l_s, m_s) corresponds to the source position or incident angle. The pupil position or outgoing angle is specified by $(l + l_s, m + m_s)$. After the discretization of the momentum, Eq. 1 is written as follows.

$$\mathbf{A}(l, m; l_s, m_s) = \mathbf{A}^{\text{FT}}(l, m) + \mathbf{A}^{\text{3D}}(l, m; l_s, m_s). \quad (4)$$

All the elements of the thick mask amplitude are calculated simultaneously in 3D waveguide model. If we use FDTD method, we need to repeat the calculation for each incident angle.

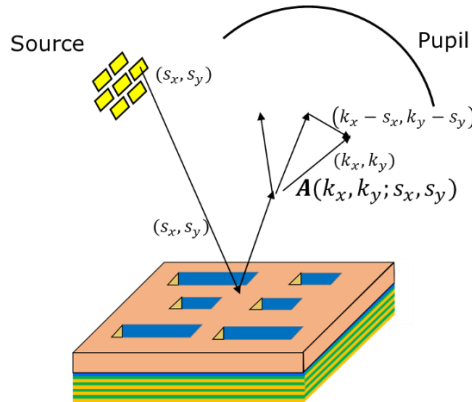


Fig. 1. Schematic view of light diffraction by an EUV mask.

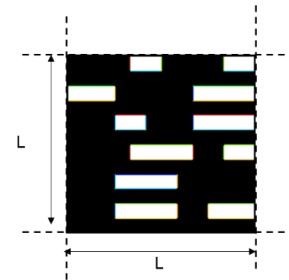


Fig. 2. Periodic mask pattern.

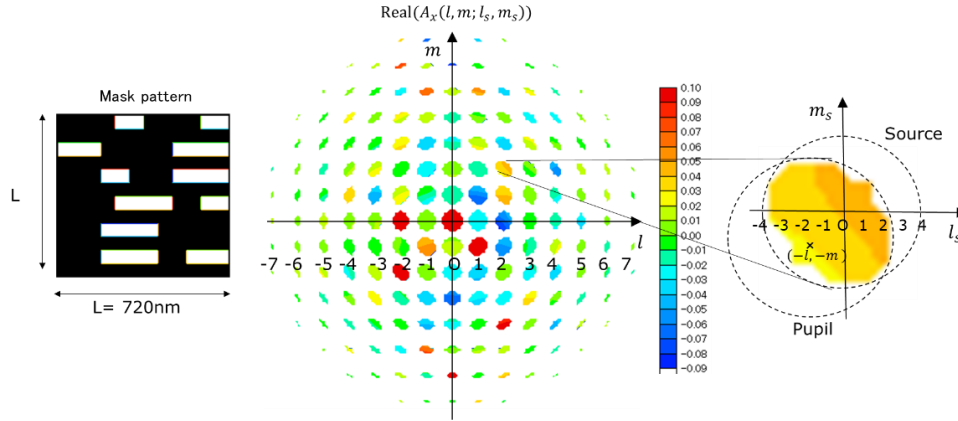


Fig. 3 Graphical representation of the diffraction amplitude.

The diffraction amplitude depends on four numbers, the diffraction order (l, m) and the source position (l_s, m_s) . These numbers are restricted by the source shape and the pupil shape. Assuming the maximum σ value of the source shape to be 1 the source position is limited by the following equation.

$$\sqrt{l_s^2 + m_s^2} \leq \frac{NA}{4} \frac{L}{\lambda} \quad (5)$$

where $NA=0.33$ is the numerical aperture of the projection optics and $\lambda=13.5$ nm is the wavelength. The magnification of the projection optics is 1/4. In the same way the pupil shape limits the pupil position $(l + l_s, m + m_s)$ as follows.

$$\sqrt{(l + l_s)^2 + (m + m_s)^2} \leq \frac{NA}{4} \frac{L}{\lambda} \quad (6)$$

Figure 3 shows the graphical representation of the diffraction amplitude. The center of the graph corresponds to the diffraction order $(0, 0)$ at all source positions. The peak value is 0.44 and it is the largest among all diffraction orders. The right-side figure enlarges the amplitude at the diffraction order $(2, 2)$. The amplitude depends on the source position (l_s, m_s) , which implies the incident angle dependence of the diffraction amplitude.

Figure 4 shows the polarization dependence of the diffraction amplitude. There are four possible combinations of incident and outgoing polarizations. The polarization change after diffraction is very small for EUV masks. For example, the maximum value of the diffraction amplitude from A_x to A_x polarization is 0.44 while the maximum amplitude from A_x to A_y polarization is 0.006. In the following discussion we will focus on the diffraction amplitude from A_x to A_x polarization.

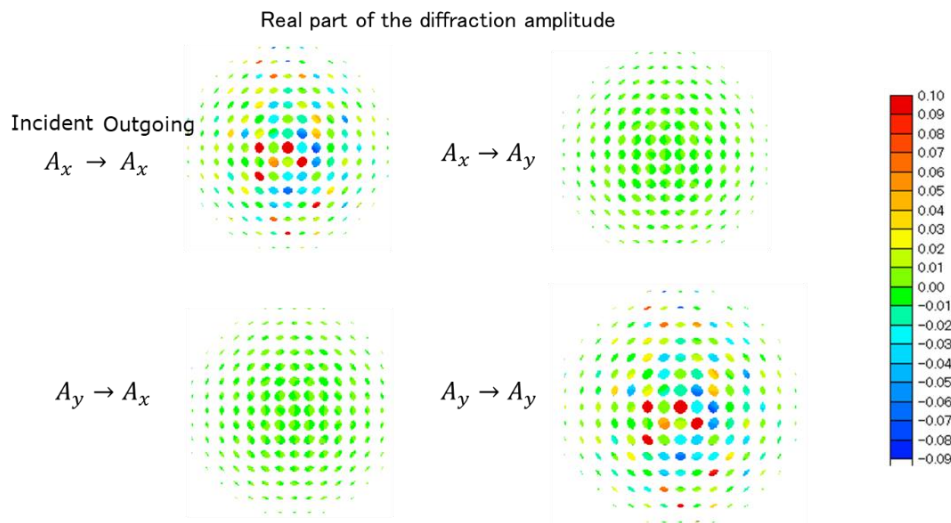


Fig.4 Polarization dependence of the diffraction amplitude.

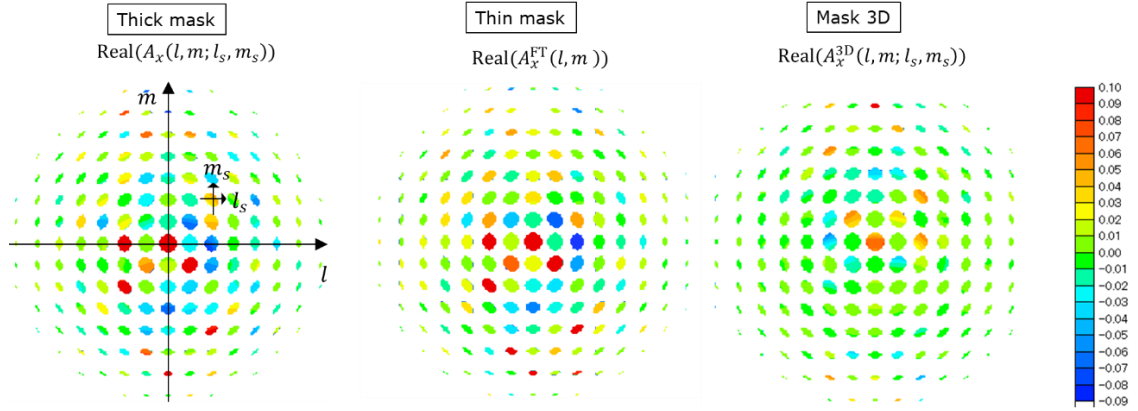


Fig. 5 Decomposition of the diffraction amplitude.

According to Eq. 4 the thick mask amplitude is divided into the thin mask amplitude and the mask 3D amplitude (Fig. 5). The contribution of the thin mask amplitude is dominant. The amplitude does not depend on the source position. It can be calculated from the Fourier transformation of the mask pattern.

The contribution of the mask 3D amplitude is small but not negligible. The amplitude depends on the source position. We parametrize the mask 3D amplitude at each diffraction order (l, m) as a linear function of the source position (l_s, m_s) .

$$A_x^{3D}(l, m; l_s, m_s) \cong a_0(l, m) + a_x(l, m) l_s + a_y(l, m) m_s, \quad (7)$$

where a_0 is the on-axis mask 3D amplitude at the center of the source position. Both a_x and a_y are the slopes of the amplitude in the x and y directions of the source plane. The second and third terms in Eq. 7 represent the off-axis mask 3D amplitude. We included here only the linear terms, but we can include higher orders if necessary.

3. CNN FOR THE MASK 3D AMPLITUDE

The thin mask amplitude $A_x^{FT}(l, m)$ is the Fourier transform of the mask pattern and it can be calculated very fast. The mask 3D amplitude $A_x^{3D}(l, m; l_s, m_s)$ is calculated by using the 3D waveguide model but the speed is slow. The mask 3D amplitude is characterized by three parameters a_0 , a_x and a_y . It is obvious that these three parameters are functions of the input mask pattern. We construct a CNN which reproduces the mask 3D amplitude. We expect that the CNN calculation is much faster than the electromagnetic simulation.

Figure 6 shows the architecture of our CNN. We use Keras on TensorFlow for the deep learning software. The input mask pattern has 720X720 binary data. We first convert them to 72X72 grayscale numbers by averaging the data. This is the

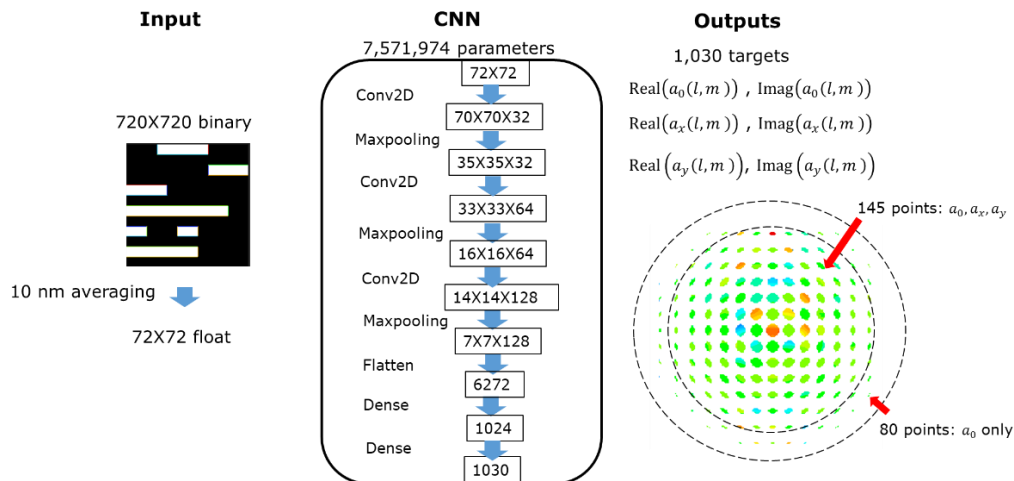


Fig. 6. Architecture of our convolutional neural network.

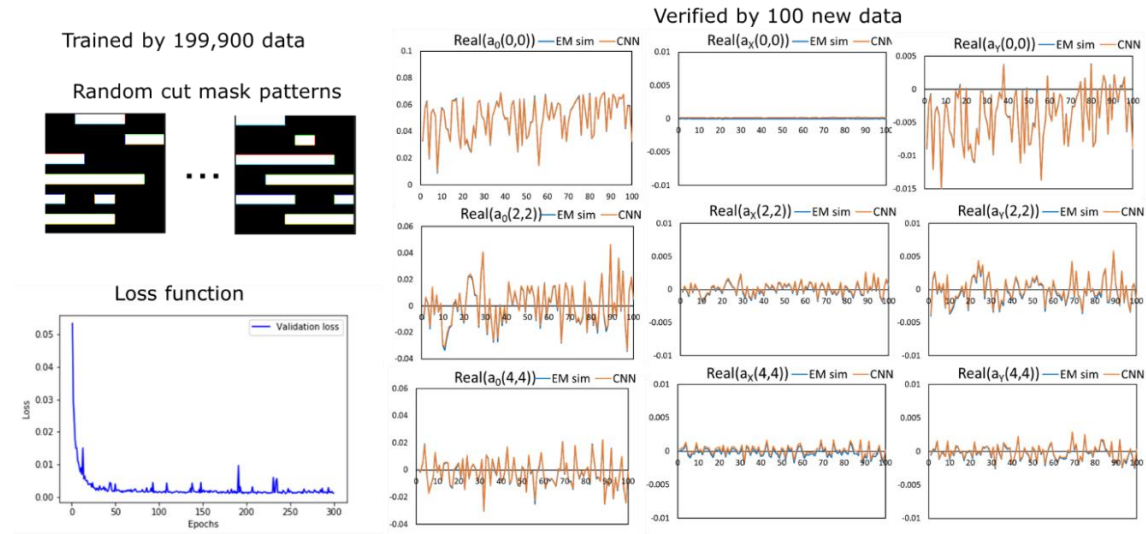


Fig. 7. Training and verification of our convolutional neural networks.

input to the CNN. Inside the CNN we repeat convolutions and maxpoolings three times. Outputs of the CNN are the real and imaginary parts of a_0 , a_x and a_y . As shown in Fig. 6, for large diffraction orders, we parametrize the amplitude with a_0 only because the area in the source plane, which represents the feasible combinations of diffraction orders and source positions is small. In total there are 1030 targets for CNN.

For training data, we generated 199,900 random cut mask patterns.¹⁸ The pattern pitch is 120 nm on the mask. The space width is 60 nm, and the space length has random number. As shown in Fig. 7, the fitting error decreased rapidly during the training. After the training we verified our CNN with 100 new data. The difference between the electromagnetic calculations and CNN is very small. We can successfully predict the mask 3D amplitude by using the CNN.

Figure 8 shows an example of the image intensities calculated by the electromagnetic simulation (EM), CNN and the Fourier Transformation (FT). It is well known that the mask 3D effect causes CD difference between two horizontal spaces at defocus positions, which is called as two-bar CD asymmetry.² CNN successfully reproduces the two-bar CD asymmetry. Thin mask model (FT) does not show CD asymmetry at defocus positions. The right-bottom figure shows the result of CNN which excludes a_x and a_y parameters. We cannot reproduce CD asymmetry without the off-axis mask 3D amplitude.

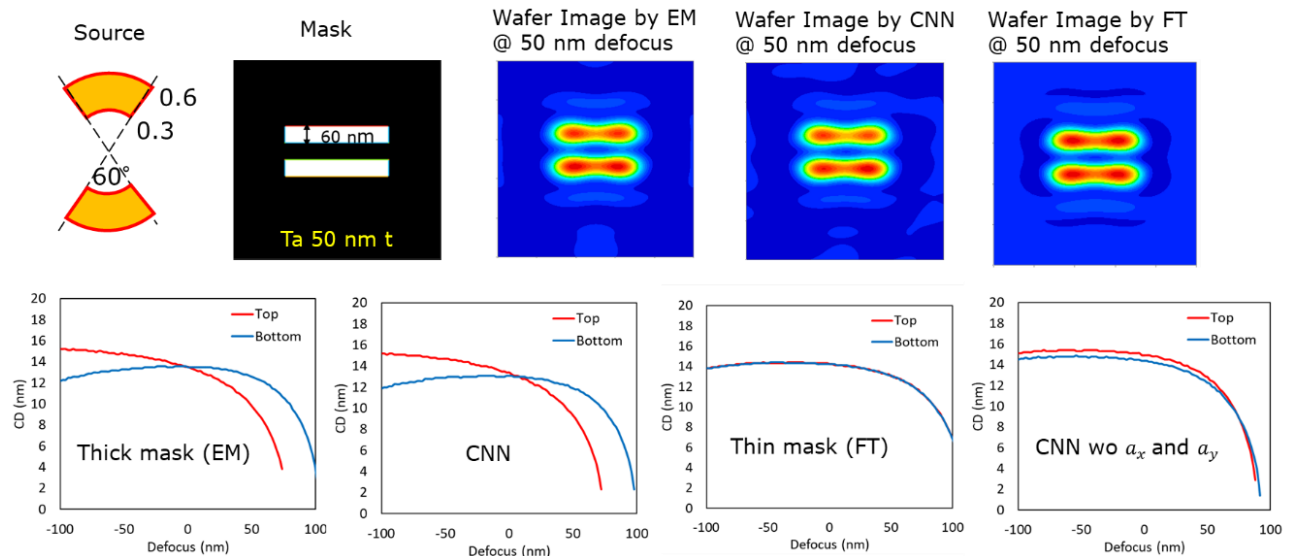


Fig. 8. Two-bar CD asymmetry simulations by EM, CNN and FT.

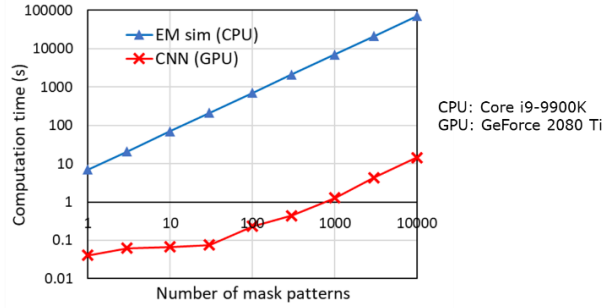


Fig. 9. Comparison of the computation time of the diffraction amplitudes.

Figure 9 compares the computation time of the EM simulation and the CNN prediction. Computation time is the time to calculate (or predict) the diffraction amplitudes. Other steps such as the intensity integration and the Fourier transformation are excluded.

The most time-consuming part of the EM simulation is the large matrix calculation. Although the EM simulation does not use GPU the matrix calculation is paralleled inside CPU by using Intel MKL. The computation time of the EM simulation linearly increased as a function of the number of the input data.

CNN prediction uses GPU. To get the full performance of GPU the number of the input data needs to be large. Computation time increases linearly beyond 300 input data. In this region the CNN prediction is 5,000 times faster than the EM simulation.

4. EXTENDED TCC FORMULA

SOCS model is often used in optical lithography simulation, especially in OPC (Optical Proximity Correction). The model decomposes TCC into a small number of eigen modes to accelerate the calculation. However, TCC formula is based on the thin mask model and the formula needs to be extended when the thick mask model is used. We derived the extended TCC formula for 1D mask patterns in Ref. 17. In this section we extend the TCC formula for 2D mask patterns.

We first rewrite the mask 3D amplitude in Equation 7 into the momentum form.

$$A^{3D}(\mathbf{k}; \mathbf{s}) \cong A^{3D}(\mathbf{k} - \mathbf{s}) + \partial_{sx} A^{3D}(\mathbf{k} - \mathbf{s}) s_x + \partial_{sy} A^{3D}(\mathbf{k} - \mathbf{s}) s_y. \quad (8)$$

As shown in Equation 1, the thick mask amplitude is the sum of the thin mask amplitude and the mask 3D amplitude.

$$A(\mathbf{k}; \mathbf{s}) = A^{FT}(\mathbf{k} - \mathbf{s}) + A^{3D}(\mathbf{k}; \mathbf{s}). \quad (9)$$

Next, we convert the vector potential to the electric field. In coordinate space the electric field is calculated from the vector potential as follows.⁶

$$\mathbf{E}(\mathbf{x}; \mathbf{s}) = ikA(\mathbf{x}; \mathbf{s}) + \frac{i}{k} \nabla_x (\nabla_x \cdot \mathbf{A}(\mathbf{x}; \mathbf{s})). \quad (10)$$

In momentum space the above equation is written as

$$\mathbf{E}(\mathbf{k}; \mathbf{s}) = ikA(\mathbf{k}; \mathbf{s}) - \frac{i}{k} (\mathbf{k} \cdot \mathbf{A}(\mathbf{k}; \mathbf{s})) \mathbf{k}. \quad (11)$$

Inserting Equations 8 and 9 into Equation 11 we obtain the following equation.

$$\mathbf{E}(\mathbf{k}; \mathbf{s}) \cong \mathbf{E}(\mathbf{k} - \mathbf{s}) + \partial_{sx} \mathbf{E}(\mathbf{k} - \mathbf{s}) s_x + \partial_{sy} \mathbf{E}(\mathbf{k} - \mathbf{s}) s_y. \quad (12)$$

Finally, the image intensity I is calculated from the electric field as follows (Abbe's method).

$$\begin{aligned} I(\mathbf{x}) &= \iint S(\mathbf{s}) \left| \iint \mathbf{E}(\mathbf{k}; \mathbf{s}) P(\mathbf{k}) e^{-ik \cdot \mathbf{x}} d\mathbf{k} \right|^2 d\mathbf{s} \\ &\cong \iint S(\mathbf{s}) \left| \iint (\mathbf{E}(\mathbf{k} - \mathbf{s}) + \partial_{sx} \mathbf{E}(\mathbf{k} - \mathbf{s}) s_x + \partial_{sy} \mathbf{E}(\mathbf{k} - \mathbf{s}) s_y) P(\mathbf{k}) e^{-ik \cdot \mathbf{x}} d\mathbf{k} \right|^2 d\mathbf{s} \\ &= \iint S(\mathbf{s}) \left| \iint (\mathbf{E}(\mathbf{k}) + \partial_{sx} \mathbf{E}(\mathbf{k}) s_x + \partial_{sy} \mathbf{E}(\mathbf{k}) s_y) P(\mathbf{k} + \mathbf{s}) e^{-ik \cdot \mathbf{x}} d\mathbf{k} \right|^2 d\mathbf{s}, \end{aligned} \quad (13)$$

where S is the source intensity and P is the pupil function. In the case of high NA optics, the pupil function is a matrix which changes the direction of the electric field from the entrance pupil to the exit pupil.¹⁹

By interchanging the order of the integrations in Equation 13 we get the following equation.

$$I(x) \cong \iint TCC(\mathbf{k}; \mathbf{k}') \mathbf{E}(\mathbf{k}) \cdot \mathbf{E}(\mathbf{k}')^* e^{-i(\mathbf{k}-\mathbf{k}') \cdot \mathbf{x}} d\mathbf{k} d\mathbf{k}' \\ + 2\text{Re}\{\iint TCC_x(\mathbf{k}; \mathbf{k}') \mathbf{E}(\mathbf{k}) \cdot \partial_{sx} \mathbf{E}(\mathbf{k}')^* e^{-i(\mathbf{k}-\mathbf{k}') \cdot \mathbf{x}} d\mathbf{k} d\mathbf{k}'\} \\ + 2\text{Re}\{\iint TCC_y(\mathbf{k}; \mathbf{k}') \mathbf{E}(\mathbf{k}) \cdot \partial_{sy} \mathbf{E}(\mathbf{k}')^* e^{-i(\mathbf{k}-\mathbf{k}') \cdot \mathbf{x}} d\mathbf{k} d\mathbf{k}'\}, \quad (14)$$

where TCC is the conventional transmission cross coefficient,

$$TCC(\mathbf{k}; \mathbf{k}') = \iint S(s) P(\mathbf{k} + s) P^*(\mathbf{k}' + s) ds \quad (15)$$

and TCC_x and TCC_y are extended transmission cross coefficients which represent the off-axis mask 3D effect.

$$TCC_x(\mathbf{k}; \mathbf{k}') = \iint s_x S(s) P(\mathbf{k} + s) P^*(\mathbf{k}' + s) ds, \quad (16)$$

$$TCC_y(\mathbf{k}; \mathbf{k}') = \iint s_y S(s) P(\mathbf{k} + s) P^*(\mathbf{k}' + s) ds. \quad (17)$$

TCC , TCC_x and TCC_y are Hermitian matrices. Therefore, we can use the eigen value decomposition method in the SOCS model. Our formula can be applied to arbitrary source shapes and defocus. We can accelerate the calculation by selecting small number of the eigen functions which have large eigen values.

5. SUMMARY

Diffraction amplitudes from thick EUV masks are calculated by using 3D waveguide model. The model contains two vector potentials A_x and A_y which represents two polarizations. There is no redundant variable in the model.

We define the mask 3D amplitude as the difference between the thick-mask and thin-mask diffraction amplitudes. We decompose the mask 3D amplitude into three parts. One is the on-axis mask 3D amplitude, and others are the off-axis mask 3D amplitudes in x and y directions which depend on the incident angle of the illumination.

We construct a CNN which can predict the diffraction amplitudes from 2D EUV mask patterns. After the training, the CNN successfully reproduce the mask 3D effects. CNN prediction is 5,000 times faster than the EM simulation.

We extend the TCC formula to include the off-axis mask 3D effect. Our formula can be applied to arbitrary source shapes and defocus. There are three TCCs in the formula. All of them are Hermitian matrices and we can use the eigen value decomposition method to accelerate the calculation.

In this study the mask area is small (720 nm X 720 nm) and the mask patterns are restricted to cut mask patterns. It is a big challenge to construct a CNN which covers large area and general mask patterns while keeping high accuracy.

REFERENCES

- [1] A. Erdmann, P. Evanschitzky, G. Bottiglieri, E. Setten and T. Fliervoet, "3D mask effects in high NA EUV imaging," Proc. SPIE 10957 (2019)109570Z.
- [2] T. Last, L. Winter, P. Adrichem and J. Finders, "Illumination pupil optimization in 0.33NA EUVL by intensity balancing for semi-iso dark field two-bar M1 building blocks," Proc. SPIE 10032 (2016)100320A.
- [3] M.C. Lare, F.J. Timmermans and J. Finders, "Alternative reticles for low-k1 EUV imaging," Proc. SPIE 11147 (2019)111470D.
- [4] A. Wong, "TEMPEST users' guide", UCB/ERL M94/64 (1994).
- [5] M.G. Moharam and T.K. Gaylord, "Rigorous coupled-wave analysis of planar-grating diffraction," J. Opt. Soc. Am. 71(1981) 811.
- [6] K.D. Lucas, H. Tanabe and A.J. Strojwas, "Efficient and rigorous three-dimensional model for optical lithography simulation," J. Opt. Soc. Am. A 13(1996)2187.

- [7] K. Adam and A. R. Neureuther, "Simplified modes for edge transitions in rigorous mask modeling," Proc. SPIE 4346 (2001)331.
- [8] A. Erdmann, C. Kalus, T. Schmoeller and A. Wolter, "Efficient simulation of light from 3-dimensional EUV-masks using field decomposition techniques," Proc. SPIE 5037 (2003)482.
- [9] P. Liu, X. Xie, W. Liu and K. Gronlund, "Fast 3D thick mask model for full-chip EUVL simulations," Proc. SPIE 8679 (2013)86790W.
- [10] H. Zhang, Q. Yan, D. Wei and E. Croffie, "A pattern- and optics-independent compact model of mask3D under off-axis illumination with significant efficiency and accuracy improvement," Proc. SPIE 9426 (2015)94260Q.
- [11] V. Domnenko, B. Kuechler, W. Hoppe, J. Preuninger, U. Klostermann, W. Demmerle, M. Bohn, D. Krueger, R. Ryoung, H. Kim, and L.E. Tan, "EUV computational lithography using accelerated topographic mask simulation," Proc. SPIE 10962 (2019)109620O.
- [12] S. Lan, J. Liu, Y. Wang, K. Zhao and J. Li, "Deep learning assisted fast mask optimization," Proc. SPIE 10587 (2018)105870H.
- [13] P. Liu, "Mask synthesis using machine learning software and hardware platforms," Proc. SPIE 11327 (2020)1132707.
- [14] R. Pearman, M. Meyer, J. Ungar, H. Yu, L. Pang and A. Fujimura, "Fast all-angle mask 3D ILT patterning," Proc. SPIE 11327 (2020)113270F.
- [15] M. Born and E. Wolf, "Principles of Optics," 7th Ed. 1999.
- [16] N.B. Cobb, "Fast optical and process proximity correction algorithms for integrated circuit manufacturing," Ph.D. dissertation (University of California, Berkeley, 1998).
- [17] H. Tanabe, S. Sato and A. Takahashi, "Fast 3D lithography simulation by convolutional neural network: POC study," Proc. SPIE 11518 (2020)115180L.
- [18] Y. Borodovsky, "EUV Lithography at insertion and beyond," International Workshop on EUV Lithography (2012).
- [19] M.S. Yeung, D. Lee, R. Lee and A.R. Neureuther. "Extension of the Hopkins theory of partially coherent imaging to include thin-film interference effects," Proc. SPIE 1927 (1993)452.