

Estimating extremely low probability of stochastic defect in extreme ultraviolet lithography from critical dimension distribution measurement

Hiroshi Fukuda
Yoshinori Momonoi
Kei Sakai

Estimating extremely low probability of stochastic defect in extreme ultraviolet lithography from critical dimension distribution measurement

Hiroshi Fukuda,* Yoshinori Momonoi, and Kei Sakai

Hitachi High-Technologies Corporation, Tokyo, Japan

Abstract. Projection lithography using extreme ultraviolet (EUV) light at 13.5-nm wavelength will be applied to the production of integrated circuits below 7-nm design rules. In pursuit of further miniaturization, however, stochastic pattern defect problems have arisen, and monitoring such defect generation probabilities in extremely low range ($<10^{-10}$) is indispensable. We discuss a method for predicting stochastic defect probabilities from a histogram of feature sizes for patterns several orders of magnitude fewer than the number of features to inspect. Based on our previously introduced probabilistic model of stochastic pattern defect, the defect probability is expressed as the product sum of the probability for edge position and the probability that film defect covers the area between edges, and we describe the latter as a function of edge position. The defect probabilities in the order between 10^{-7} and 10^{-5} were predicted from 10^5 measurement data for real EUV-exposed wafers, suggesting the effectiveness of the model and its potential for defect inspection. © The Authors. Published by SPIE under a Creative Commons Attribution 4.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: [10.1117/1.JMM.18.2.024002](https://doi.org/10.1117/1.JMM.18.2.024002)]

Keywords: lithography; extreme ultraviolet; metrology; defect inspection; stochastic defect.

Paper 19006L received Jan. 29, 2019; accepted for publication Apr. 12, 2019; published online May 3, 2019.

1 Introduction

Projection lithography using extreme ultraviolet (EUV) light at the 13.5-nm wavelength is expected to achieve production of integrated circuits (ICs) below 7-nm design rules.¹ In pursuit of further miniaturization of semiconductor integrated circuit devices by EUV lithography, stochastic pattern defect problems have arisen.^{2–4} Stochastic pattern defects are fatal patterning failures such as bridging between neighboring pattern features or breakages of features, and its probability is extremely low (down to 10^{-12} or even below). Because cutting-edge integrated circuit devices today have more than 10^{12} critical features per a device layer on a 300-mm wafer, such a defect probability will result in an unacceptable level of defect density.

While suppressing the stochastic defect itself is indispensable for EUV lithography, monitoring and control of these defects is another crucial issue.^{4–7} When applying EUV lithography to IC manufacturing, design rules and nominal mask/process conditions should be set so that the stochastic defect probability is within a tolerable range (e.g., 10^{-12}). Since stochastic defect probability is very sensitive to resist feature size or the mask and process conditions, however, small deviations from the nominal condition can cause catastrophic wafer failure³ (e.g., a change in exposure dose of a few percent can in some cases change the defect probability by an order of magnitude). Detecting changes in stochastic defect probability in this extremely low range will be necessary but is a challenge. For directly inspecting a huge number (e.g., 10^{12}) features to detect below 10-nm size defects, present electron-beam-based inspection tools require unacceptably long inspection time,⁵ whereas the resolution

capability is marginal for optical inspection tools.⁶ In contrast, it has been reported that conventional indices, such as critical dimension (CD) and line edge roughness (LER), have correlations with defect probabilities though they are empirical without theoretical ground.⁷ Here, we propose an approach to predict an extremely low probability of stochastic defect from local CD uniformity (LCDU) data or CD histogram for a limited number of pattern features, typically several orders of magnitude lower than a number of features to be inspected. We previously introduced the probabilistic model for stochastic defects generation based on two mechanisms, cascading shot noises and long-range scattered photoelectrons.^{8,9} In this paper, we apply this model to predict an extremely low probability of stochastic defect generation on real wafers.

2 Probabilistic Model of Pattern Defects

Before discussing the defect prediction, here, we briefly review our model.^{8,9} We start from generating numbers of physical/chemical events in a resist film, such as photon absorption, secondary electron generation, chemical reaction, and solubility flipping of resist polymer/molecule using coupled Monte-Carlo simulation, which combines simulations for optical imaging, photoelectron scattering, and chemical amplification with acid diffusion [Fig. 1(a)]. We divide the resist film by three-dimensional grids and count the number of reactions in each voxel produced by the grids. We assume that the solubility of a particular voxel flips if the number of reactions in that voxel exceeds a certain threshold, and further, count the number, n_{SF} , of solubility-flipped voxels through thickness, which represents the degree of solubility change in a particular spot of resist film. From the histogram of this number n_{SF} under the same exposure dose, we obtain the probability density functions (PDFs) $pdf_{SF}(\vec{r}, n_{SF})$ for n_{SF} at location \vec{r} . Here, we

*Address all correspondence to Hiroshi Fukuda, E-mail: hiroshi.fukuda.zp@hitachi-hightech.com

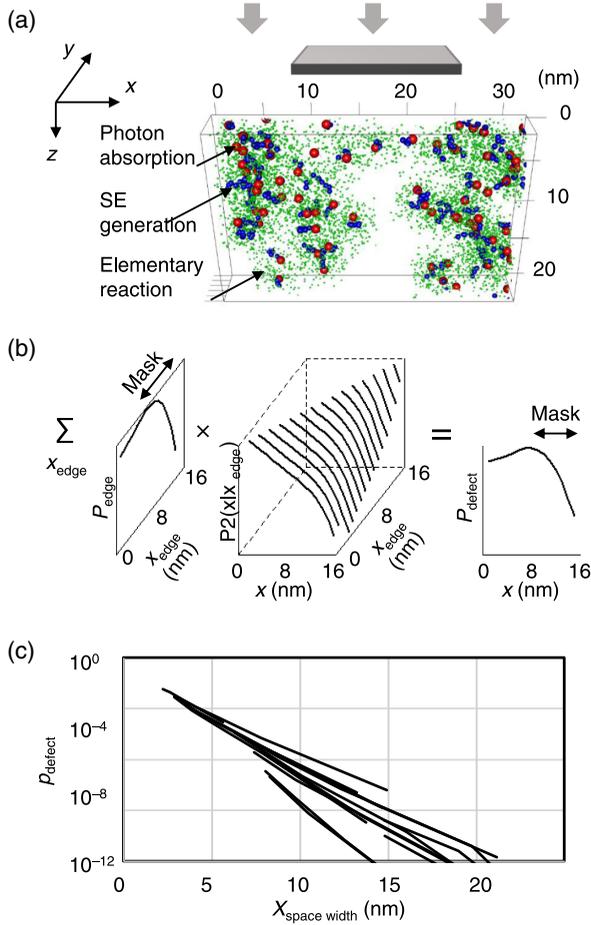


Fig. 1 Model of stochastic pattern defect. (a) Simulated distributions of photon absorptions (red spheres), SE generations (blue spheres), and acid-catalytic reactions (green spheres) for EUV-exposed chemically amplified resist. (b) Schematic procedure of defect probability calculation in Eq. (2). (c) Simulated relationships between defect probability and delineated pattern size for various exposure/material conditions. See Ref. 8 for details.

focus on bridge-type defects in negative-tone resist processes. We define a local spot pattern and a local spot defect so that they are generated when the number n_{SF} of solubility-flipped polymer/molecule through the film thickness exceeds a certain threshold Nc_{SF_X} (X = main pattern or film defect). Thus, the probabilities of local spot pattern/defect $P1_X$ per unit area (e.g., 1 nm^2) are expressed as

$$P1_X(\vec{r}_i, Nc_{\text{SF}_X}) = \int_{Nc_{\text{SF}_X}}^{\infty} pdf_{\text{SF}}(\vec{r}_i, n_{\text{SF}}) dn_{\text{SF}}, \quad (1)$$

where pdf_{SF} is the PDF for the number of solubility-flipped polymer/molecule through the film thickness. Main patterns are formed if the spot patterns cover over designated areas, whereas pattern defects are generated if the spot defects cover over critical areas of circuit features, such as residual film between main features. Assuming one-dimensional pattern for simplicity, the stochastic pattern defect probability (for mechanism A in Ref. 8) is obtained as the probability that the spot film defects cover the area between the main pattern edge at x_{edge} and the point x_d representing defect area as

$$P_{\text{defect A}}(x_d) = \int P_{\text{edge}}(x_{\text{edge}}) \cdot P2_{\text{defect}}(x_d|x_{\text{edge}}) dx_{\text{edge}}, \quad (2)$$

where

$$P2_{\text{defect}}(x_d|x_{\text{edge}}) = \prod_{x_{\text{edge}} < x < x_d} P1_{\text{defect}}(x, Nc_{\text{SF_defect}}), \quad (3)$$

where $P_{\text{edge}}(x_{\text{edge}})$ is the probability that the main pattern edge locates at x_{edge} and $P2_{\text{defect}}(x_d|x_{\text{edge}})$ is the probability that the spot film defects cover the area between x_{edge} and x_d . Figure 1(b) illustrates how we obtain $P_{\text{defect A}}(x_d)$ from $P_{\text{edge}}(x_{\text{edge}})$ and $P2_{\text{defect}}(x_d|x_{\text{edge}})$. A periodic structure with 32-nm pitch is assumed with the center of exposed and unexposed area located at $x = 0$ and 16 nm, respectively, and the mask edge at $x = 8$ nm.

Equation (2) shows that the probability of defect generation between x_d and x_{edge} depends on the horizontal location of edge x_{edge} . Although the actual edge location also varies in the depth direction along resist sidewall, the variations of edge location in the vertical direction are usually smaller than that in the horizontal direction (so-called LER), and we ignore the former in the present model. The above explanation assumed the defect generation mechanism A in Ref. 8 for simplicity, but the form of Eq. (2) holds also for mechanism B in the same reference. Optimization of exposure and material parameters to minimize defect probability showed clear trade-off relationship between defect probabilities and delineated pattern feature sizes as shown in Fig. 1(c), which is qualitatively consistent with experimental observations in Ref. 3. The exponential relationships between defect probabilities and exposure dosage required for obtaining designed size observed among varieties of resist materials⁴ are also explained by the model.⁹

3 Method of Defect Probability Estimation

Here, we apply the above-mentioned model for predicting defect probability on real wafers. In our method, the stochastic defect probability is expressed by the product sum of two probabilities $P_{\text{edge}}(x_{\text{edge}})$ and $P2(x|x_{\text{edge}})$ in Eq. (2). Our basic approach is to predict defect probability by evaluating P_{edge} and $P2$ in Eq. (2), not by directly inspecting full-pattern features. Evaluating probability in the order of P requires more than $1/P$ samples in general. Since both P_{edge} and $P2$ are larger than P_{defect} by orders of magnitude, we expect the same order of measurement time reduction. Here, P_{edge} is a histogram of local edge position and directly measurable using SEM, and thus, we focus on how we evaluate $P2$.

Let us suppose that defect probability increases due to some process variations, and we need to detect this change. According to the above model, these variations change the defect probability through P_{edge} and $P2$ in the following three pathways. First, process variations change the locations x_{edge} of pattern edges and their distribution P_{edge} . Second, the change in x_{edge} changes the value of $P2$ because $P2$ is a function of x_{edge} . Third, process variations change the function $P2$ itself because $P2$ is determined from chemical reaction density as explained from Eqs. (1) and (3).

We examined the changes in P_{edge} and $P2$ along each pathway using our above-described defect probability

model. Figure 2 shows the profiles of $P_{\text{edge}}(x_{\text{edge}})$, $P2(x_{\text{center}}|x_{\text{edge}})$, and $P_{\text{defect}}(x)$ for two exposure conditions, nominal and 20% overirradiation. Here, we assumed one of the exposure/material parameter sets optimized so as to minimize defect probability for 16-nm lines and spaces with 0.33 NA optics. Please see Ref. 8 for details. A 20% increase in irradiation dosage shifts the mean CD by 20% (corresponding to a 1.5-nm shift in edge position) with changing the histogram profiles [Fig. 2(a)]. While it also changes the profile of $P2$, this is small compared to its exponential dependence on x_{edge} [Fig. 2(b)]. In contrast, a 20% increase in dose changes P_{defect} by 2 orders of magnitude at the same location x [Fig. 2(c)]. This is because the linear change in x_{edge} is magnified by the exponential dependence of $P2$ on x_{edge} . Consequently, defect probability is exponentially dependent on the above amount of exposure dose variations through the first and second pathways. If we assume the shape of function $P2$ (dependences on x_{edge} and x) unchanged within the above ranges of exposure variations as an approximation, we can calculate the value of $P2$ from measured x_{edge} , and further P_{defect} as a product sum of P_{edge} and $P2$. Note, however, that $P2$ is a function of imaging and resist materials/processes conditions in general, and the above assumption needs to be examined when these conditions are changed.

Practically, two approaches can be taken for determining $P2$. In the first analytical approach, we directly calculate $P2$ using the probabilistic defect model as explained in Fig. 2. This requires model calibration as in every conventional lithography simulations. The other is an empirical approach, where we determine $P2$ so as to satisfy Eq. (1) with observed

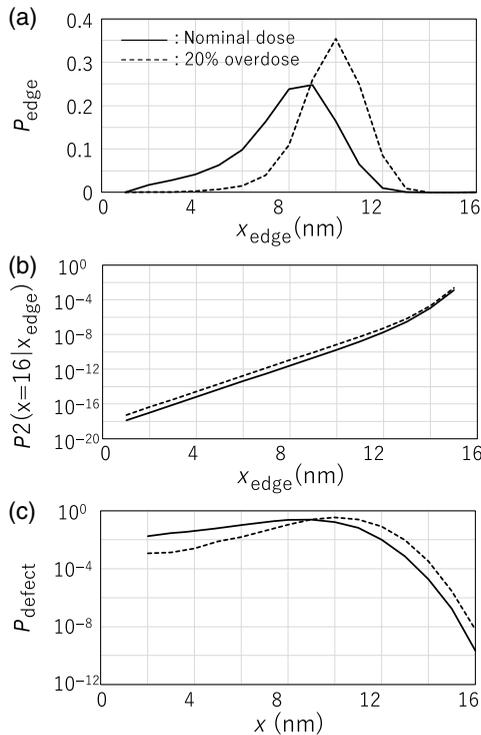


Fig. 2 Calculated probabilities: (a) P_{edge} that pattern edge locates at x_{edge} , (b) $P2$ that films defect generates between x_{edge} and clear space center ($x = 16$ nm), and (c) P_{defect} that pattern defect exists at the clear space center.

P_{edge} and P_{defect} . In Sec. 4, we examine the feasibility of our method using the latter approach.

4 Experimental Results and Discussions

We predict the defect probabilities in the order between 10^{-7} and 10^{-5} from 10^5 measurement data on real EUV-exposed wafers. Mask patterns containing two-dimensional array of more than 10^7 holes (24-nm diameter in 48-nm pitch) were exposed on a wafer ($\lambda = 13.5$ nm, NA = 0.33) with varying exposure dose to modulate defect probability. For each of the resist pattern groups exposed under 20 kinds of different exposure doses, each hole pattern size was measured by CD-SEM (Hitachi High-Technologies). The size of each feature was calculated from the area of ellipse best fitted to the shape defined by 50% threshold of signal intensity after applying Gauss filter to SEM images. With a 1-nm pixel size, about 50 pixels on edge contribute measurement, and the estimated error due to SEM noise is lower than 0.2 nm at the probe current (>100 pA) used in the experiment.¹⁰ We judge features below 9.5 nm as defects and calculate histograms of measured CD excluding these defects. CD histograms [1-nm bin, Fig. 3(a)] and defect probabilities [red

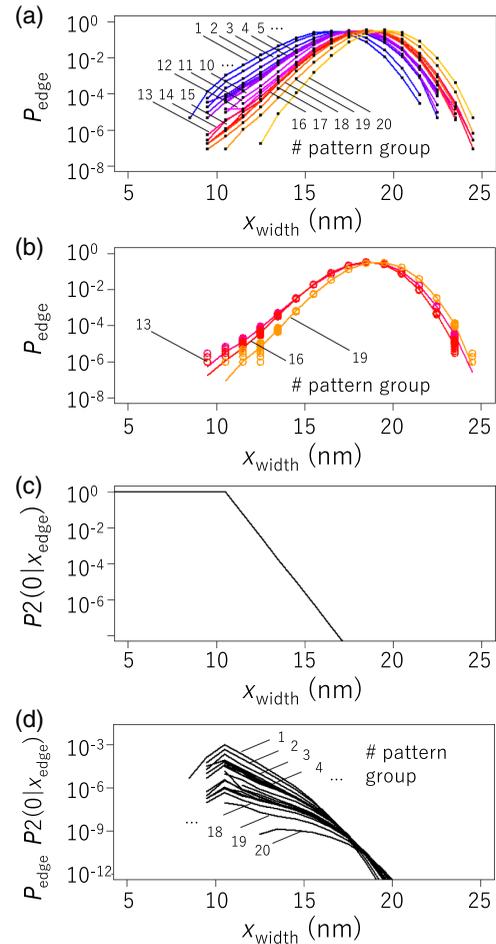


Fig. 3 (a) Distribution histograms of pattern size with varying exposure dosage (pattern groups #1 to #20). (b) Distributions of pattern size for groups #13, 16, and 19. Solid lines are for the results of full-pattern ($\sim 10^7$) measurement and small circles are for 10^5 pattern measurement. (c) Best fitted $P2$ to reproduce fully inspected results [red diamonds in Fig. 4]. (d) Integrand of Eq. (2) for pattern groups #1 to #20.

diamonds in Fig. 4] were obtained for 2×10^5 holes for the pattern groups #1 to #12 with relatively high ($>10^{-5}$) defect probability and for 10^7 holes for the pattern groups #13 to #20 with relatively low ($<10^{-5}$) defect probability. The defect probabilities exponentially decrease from 10^{-3} in group #1 to 10^{-7} in group #19 with increasing the average diameter of holes from 16.2 to 19.1 nm. Thus, a 3-nm decrease in feature size increases the defect probability by 4 orders of magnitude.

Here, we focus on the relationship between CD variations and pattern defect probabilities without discussing their root causes. In this experiment, we observed no definitive mask defect that prints on wafers regardless of exposure dosage. Although some defects observed in this experiment can be mask origin, their probabilities exponentially increase with decreasing exposure dose (or delineated hole size) similarly to as expected for other root causes, such as photon shot noises and stochastic variations in resist reactions discussed previously. We regard them equally as defects due to local variations in the amount of reactions, include them in the P_{edge} distribution, and apply the same $P2$ function in Eq. (2) no matter if their locations are fixed on the mask or random.

Our strategy is to determine the probability function $P2$ in Eq. (2) so that it best explains observed defect probabilities P_{defect} and CD histogram P_{edge} for every exposure conditions (pattern groups). In real application environments, it is desired to minimize the number of measurement points (time required for measurement) both in determining $P2$ and in predicting P_{defect} for unknown samples. Here, however, we utilized all the data in the group #1 to #20 for determining $P2$.

As a rough approximation of our simulated profiles for $P2$ [Fig. 2(b)], we assume that $P2$ exponentially decreases with the distance from the edge of main pattern and describes it in the form of $P2_0 \exp(-a \cdot x_{\text{width}})$. Here, we use the width of each feature ($x_{\text{width}} = x_{\text{right edge}} - x_{\text{left edge}}$) instead of x_{edge} to eliminate the influence of variation in pattern center positions. We calculate $P2$ ($P2_0$ and a) so that $\log(\int P_{\text{edge}} \cdot P2 dx_{\text{width}})$ best fits to $\log(P_{\text{defect}})$ for 19 groups (#1 to #19), and the obtained profile of $P2$ is shown in Fig. 3(c). Although $P2$ has no influence on calculated P_{defect} in $x_{\text{width}} < 9.5$ nm where we judge features as defects ($P_{\text{edge}} = 0$), $P2$ is set 1 for this region. From a statistical viewpoint, $P2$ can be regarded as the extreme-value cumulative distribution function that expresses the distribution of maximum distance for defects to continuously extend from the main pattern edge. Here, we leave the relationship between our assumption for $P2$ and varieties of functions used in this area open.

Next, we predicted the defect probabilities of groups #13 to #20 from 10^5 CD measurement data in each group with the above obtained $P2$. To examine the repeatability of the method, we repeated random sampling of 10^5 CDs from 10^7 CDs 100 times. Since the defect probabilities for the above groups range between 10^{-7} and 10^{-5} , each sampled CD data rarely contains defects (in average one defect in 10 samplings for $P_{\text{defect}} = 10^{-6}$). Predicted probabilities are shown by boxplots in Fig. 4, and they are in good agreement with the results of direct inspection of 10^7 features (red diamonds).

For the probabilities above 10^{-5} , the data used for prediction contain some defects, and the box plots are regarded as

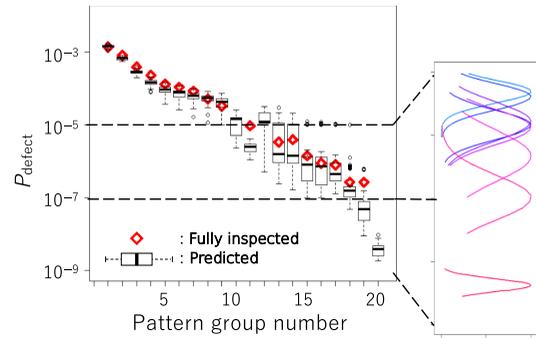


Fig. 4 Defect probabilities P_{defect} for pattern groups #13 to #20. Results of full inspection of 10^7 features (red) and those predicted from 10^5 pattern measurement (black). Box plots and normal distribution fits in the right are for 100 predictions using 10^5 CDs randomly sampled from 10^7 measurements.

the results of regression rather than of prediction. Between 10^{-7} and 10^{-5} , the data used for prediction usually contain no defect, and the predicted results (box plots) are verified by directly inspected results. Predicted results below 10^{-7} cannot be verified because it is beyond the limitation of direct measurement. These results show 2 orders of magnitude reduction in the time required for evaluating defect probability.

Predicted probabilities fitted into normal distribution are plotted for each of the seven groups in Fig. 4, and the prediction repeatability is in the range between 0.2 and 0.4 digit. Histograms of 10^5 measured CDs are shown for three groups (#13, 16, and 19) by circles in Fig. 3(b) with those for 10^7 measurement (solid lines). The frequencies of CDs in 10^5 histograms begin to scatter in the tail regions, and this limits the precision of the prediction.

To examine the range of edge position contributing to defect generation, the integrands in Eq. (2) [the product of Figs. 3(a) and 3(c)] are shown in Fig. 3(d) for the histograms of full-pattern measurement in every pattern groups. Peaks of the integrands spread to the range below 10 nm. Although histograms should cover this range, this often requires an unacceptably large number of measurement points (and thus long measurement time) for real manufacturing environment with low stochastic defect probability. Next, we extrapolate the tail of histogram to cover the desired range for such cases.

It was reported that CD histograms often deviate from the normal distribution and show exponential or multiple Gaussian distributions in their tails,^{3,5,7} and its relation to image profiles has also been pointed out.¹¹ This is observed also in our results [Fig. 3(a)]. Figure 5(a) shows histograms of 10^5 measured CDs randomly sampled from 10^7 CDs for 100 times (blue circles), histogram of for 10^7 measurement (red lines), and its normal distribution fit (black dotted line). The observed distribution start deviating from normal distribution for P_{defect} lower than 10^{-3} and approximately decreases exponentially with decreasing X_{width} . Thus, we extrapolate the tail of distribution for 10^5 measured CDs using the exponential function.

To suppress the influence of data scattering near the tail of distribution, here, we reject the data at the smallest CD bin of histogram, calculate the slope (decay coefficient) by averaging the slope between the second and the third smallest CD bins and that between the second and fourth smallest CD

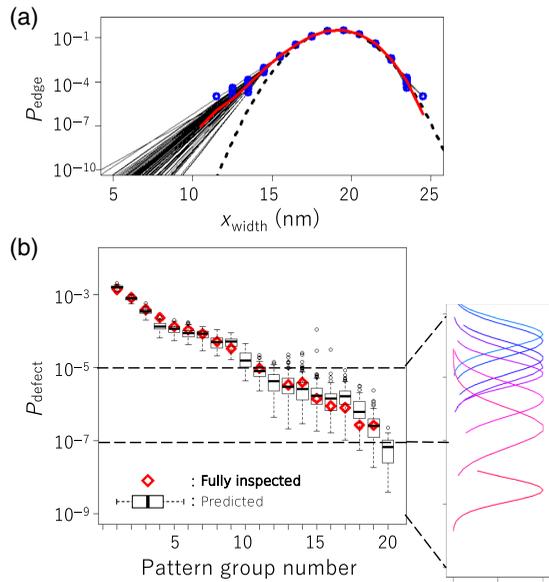


Fig. 5 (a) Distribution of pattern size for group #19. The 100 histograms of 10^5 CDs randomly sampled from 10^7 CD measurements (blue circles), distributions extrapolated from each of the 100 histograms (black solid lines), histogram of 10^7 CDs (red line), and normal distribution fit of 10^7 measurements (dotted line). (b) Defect probabilities predicted for pattern groups #1 to #20 using extrapolated distributions. Results of full inspection of 10^7 features (red) and those predicted from 10^5 pattern measurement (black). Box plots and normal distribution fits in the right are for 100 predictions using 10^5 CDs randomly sampled from 10^7 measurements.

bins, and connect the exponential function to the measured histogram at the second smallest CD bin. To examine the repeatability of the method, we repeated random sampling of 10^5 CDs from 10^7 CDs 100 times. Results of extrapolation are shown by black solid lines for the 100 samplings in Fig. 5(a). Predicted probabilities for groups #1 to #20 are shown by boxplots in Fig. 5(b), and they showed better agreement with the results of full-pattern inspection [red in Fig. 5(b)] than without using the extrapolation [Fig. 4]. The prediction repeatability is in the range between 0.2 and 0.3 digit.

Within the range of this study, it is reasonable to approximate P_{edge} , P_2 , and P_{defect} by exponential functions in the tail region of P_{edge} . However, the distributions below 10^{-7} need to be examined with various possibilities for statistical functions for modeling them. Finally, we comment on the relation of the present method to the reported defect probability dependence on tail CDs (e.g., defined as CD corresponding to 3σ limit).⁷ Assuming the exponential function $P_{\text{edge}} \propto \exp(b \cdot x_{\text{edge}})$ for x_{edge} in the tail region, suppose that the distribution of P_{edge} shifts by $-\delta x$ to $P'_{\text{edge}} \propto \exp([b(x_{\text{edge}} + \delta x)])$ due to change in exposure dosage for example. Then, P_{defect} changes to $P'_{\text{defect}} = \exp(b \cdot \delta x) P_{\text{defect}}(x_{\text{edge}})$ [$\log(P'_{\text{defect}}) = b\delta x + \log(P_{\text{defect}})$]

since the integrand of Eq. (2) is practically determined by the tail region. Thus, defect probability changes exponentially with the tail CD, and the present model explains the tail CD dependence of the defect probability.

In conclusion, applying the present method to plural spots on a chip or on a wafer visualizes the risk distribution of stochastic defects. Direct full inspection is needed only for the extracted risky area, and this is expected to reduce the required area of such a full inspection. Further, the verification results can be used for updating the model (function P_2). In this study, we predict stochastic defect probabilities from large-size LCDU data for a specific resist material/process. Note that any change in resist materials/processes can affect stochastic defect probability through the function P_2 as well as through the edge distributions (LCDU or LER).

Acknowledgments

The authors acknowledge P. De Bisschop and IMEC for the sample preparation and for their support of this work.

References

1. R. van Es et al., "EUV for HVM: towards and industrialized scanner for HVM NXE3400B performance update," *Proc. SPIE* **10583**, 105830H (2018).
2. R. Brainard et al., "Shot noise, LER and quantum efficiency of EUV photoresists," *Proc. SPIE* **5874**, 74–85 (2004).
3. P. De Bisschop, "Stochastic effects in EUV lithography: random, local CD variability, and printing failures," *J. Micro/Nanolithogr. MEMS MOEMS* **16**(4), 041013 (2017).
4. P. De Bisschop, "Stochastic printing failures in extreme ultraviolet lithography," *J. Micro/Nanolithogr. MEMS MOEMS* **17**(4), 041011 (2018).
5. F. Wang et al., "Assessment of variability and defectivity by high-throughput e-beam metrology for prediction of patterning defect probabilities," *Proc. SPIE* **10585**, 1058525 (2018).
6. K. Kaushik Sah et al., "EUV stochastic defect monitoring with advanced broadband optical wafer inspection and e-beam review systems," *Proc. SPIE* **10809**, 1080909 (2018).
7. M. J. Maslow et al., "Impact of local variability on defect-aware process window degradation," *Proc. SPIE* **10957**, 109570H (2019).
8. H. Fukuda, "Localized and cascading secondary electron generation as causes of stochastic defects in extreme ultraviolet projection lithography," *J. Micro/Nanolithogr. MEMS MOEMS* **18**(1), 013503 (2019).
9. H. Fukuda, "Impact of asymmetrically localized and cascading secondary electron generation on stochastic defects in EUV lithography," *Proc. SPIE* **10957**, 109570G (2019).
10. H. Fukuda et al., "Measurement of pattern roughness and local size variation using CDSEM," *J. Micro/Nanolithogr. MEMS MOEMS* **17**(4), 041004 (2018).
11. R. L. Bistol and M. E. Krysak, "Lithographic stochastics: beyond 3σ ," *J. Micro/Nanolithogr. MEMS MOEMS* **16**(2), 023505 (2017).

Hiroshi Fukuda joined Hitachi Central Research Laboratory in 1985, where he has engaged in various fields of lithography as well as nano-devices, MEMS, and hard disk drives, including research activities at Stanford University and Hitachi Europe Ltd. He has been with Hitachi High-Technologies since 2012. He received his BS, MS, and PhD degrees from Tokyo Institute of Technology in 1983, 1985, and 1994, respectively. He has published more than 30/80 journal/conference papers and holds over 20 patents.

Biographies of the other authors are not available.