Analytic Derivation and Minimization of Line Edge Roughness in Electron-beam Lithography

by

Rui Guo

A dissertation submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Auburn, Alabama
December 16, 2017

Keywords: Electron-beam Lithography, Line Edge Roughness, Analytic Method, Stochastic Developing Rate Distribution, Development Path, LER Minimization

Approved by

Soo-Young Lee, Professor of Electrical and Computer Engineering
Gopikrishna Deshpande, Associate Professor of Electrical and Computer Engineering
Hulya Kirkici, Professor of Electrical and Computer Engineering
Guofu Niu, Professor of Electrical and Computer Engineering

Abstract

Electron-beam (e-beam) lithography is widely employed in a variety of areas such as fabrication of photomasks, imprint lithography molds, and experimental circuit patterns, etc., because of its ability to transfer ultra-fine features onto the resist and eventually to the substrate. Its main limitations are the low throughput due to the pixel-by-pixel or feature-by-feature writing and the proximity effect caused by electron scattering. The importance of developing effective and efficient schemes for correcting the proximity effect has been well recognized for a long time, and various methods were proposed and implemented by many researchers. As the feature size decreases well below microns into nanoscale, line edge roughness (LER) becomes an increasingly important factor that cannot be ignored since it does not scale with the feature size. It can significantly affect the minimum feature size and maximum feature density realizable in practice, and also the functionality of a device. Therefore, it is unavoidable to minimize the LER in order to maximize the feature density and enhance the yield of fabricated devices.

One important step required in developing an effective method to minimize the LER is to understand the characteristics of LER. That is, it is necessary to be able to estimate the LER accurately. A possible approach is to rely on a simulation. Given a circuit pattern, the stochastic exposure (energy deposited in the resist) distribution is either obtained directly through the Monte Carlo simulation or computed through the convolution with stochastic point spread functions generated via the Monte Carlo simulation. Then, the remaining resist profile (from which the LER is quantified) is derived through a resist development simulation. While the simulation approach is flexible, the main drawback is that it is computationally intensive, since it requires an excessive computation to analyze the quantitative relationship between the LER and e-beam lithographic parameters, e.g., the dose (the amount of electron charge given to a unit area). Once the set-up of e-beam lithographic process is changed, e.g., for a different type of resist, such a process needs to be repeated. An analytic approach to LER estimation is an alternative, which can avoid the repetitive procedure, since the explicit relationship between the LER and e-beam lithographic parameters can be derived.

In this dissertation, an analytic method of deriving and minimizing the LER for a single-line pattern caused by the stochastic variation of developing rate in the resist in the e-beam lithography is described with a realistic 3-D model. The specific objective is to derive an accurate analytic expression of the LER from the moments of developing rate distribution, i.e., the mean and variance. With the developing rate distribution, development paths in the resist development process are derived and used to estimate the variation of edge locations, i.e., LER. The analytic expression of LER is used to minimize the critical dimension (CD) error and LER by optimizing the dose. Then, the analytic method is extended to large-scale uniform patterns exposed with a uniform dose by modeling the global exposure distribution. In addition, to verify the analytic method by using the experimental results, a method of extracting the stochastic information needed by the analytic method from experimental results, i.e., SEM images, is developed. A specific goal of this study is to obtain an explicit analytic expression of LER which can be used in the minimization of LER. In order to achieve the goal, in some steps of derivations, certain assumptions and approximations are made and numerical computations are employed, i.e., the derivations of LER are not completely analytic.

Acknowledgments

First, I would like to express my sincere gratitude to my advisor, Dr. Soo-Young Lee, for giving me the opportunity to join in the research of modeling and minimizing the line edge roughness in electron-beam lithography. In the doctorate program of past six years, Dr. Lee continuously guided and motivated me to get through the difficulties in this research with his vast knowledge and skills. Without his valuable suggestions, this work would not be possible.

I would like to thank Dr. Gopikrishna Deshpande, Dr. Guofu Niu, and Dr. Hulya Kirkici, for serving on my committee, and Dr. Nedret Billor, for being my university reader. They provided me with valuable comments and kindly help.

Also, I would like to extend my gratitude to the former and current members of our research group, including Qing Dai, Xinyu Zhao, Jian Bao, and Dehua Li for the collaborations. Thanks to Samsung Electronics Co., Ltd. for funding this research.

I would like to thank all the friends I have met in Auburn, especially, Yuan Bian, Yang Chao, Song Gao, Daoqi Hou, and Haoran Zhang, for giving me a wonderful experience in Auburn. Finally, I am grateful for the support of my family, including my parents, Yaohong Guo and Feng Zhao, and my wife, Selena Liu, for their love, patience, and understanding.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

Electron-beam lithography (e-beam) is a lithographic process used to transfer circuit patterns onto the resist and eventually to the substrate such as silicon. It employs a focused beam of electrons to expose a circuit pattern on the resist coated on the surface of substrate. In nanofabrication, the e-beam lithography plays an important role because of its ability to transfer circuit patterns with a much higher patterning resolution than the conventional optical lithography which is limited by the diffraction of light. Therefore, it is suitable for the fabrication of nanoscale features, e.g., photomasks, imprint lithography molds, and experimental circuit patterns, etc. [1, 2, 3, 4, 5, 15].

A well-known issue in the e-beam lithography is the proximity effect due to electron scattering, which causes a deviation from the target dimensions of a feature in the written pattern. The importance of developing effective and efficient schemes for correcting the proximity effect has been well recognized for a long time, and various methods were proposed and implemented by many researchers [6, 7, 9, 10]. However, as the feature size decreases well below micron into nanoscale, line edge roughness (LER) eventually puts a fundamental limit on the minimum feature size and maximum pattern density that can be realized, since the LER does not scale with the feature size [12, 15]. Therefore, it is essential to minimize the LER in order to achieve the highest resolution possible by the e-beam lithographic process.

1.1 Line Edge Roughness

The LER refers to the random deviation of a feature boundary from the designed boundary (see Fig. 1.1). It can be caused by a number of statistically fluctuating effects in the e-beam lithographic process, e.g., the unevenness of chemical species in the resist such as photoacid generators, the random walk nature of acid diffusion during chemical amplification, and the nonzero size of resist polymers being dissolved during development, etc. The LER analyzed in this dissertation is the random variation of edge location due to the stochastic fluctuation of exposure (energy deposited into the resist) caused by the shot noise (variation of electron flux) and random scattering of electrons.

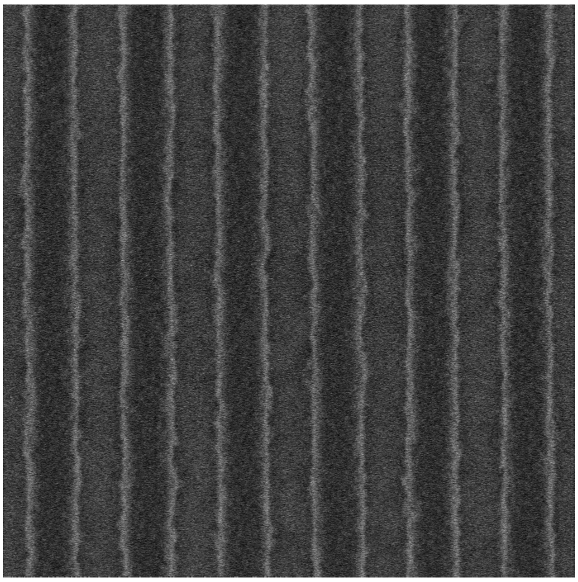


Figure 1.1: An example of SEM image in a top-down view of a line-space pattern.

1.2 Review of Previous Work

Many of the previous efforts made to estimate and reduce the LER are based on an empirical or trial-and-error approach via experiment or simulation [13, 14, 15]. However, such a method can be very time-consuming and expensive since repetitive simulations or experiments may be required.

1.2.1 Simulation Approach

In the previous study [15], a computational approach to minimizing the LER was taken. Before minimizing the LER, the dependency of LER on various lithographic parameters needs to be analyzed through an extensive simulation. In the simulation results, the following behaviors of LER were observed. (1) The LER is substantially larger at a lower layer of resist. (2) The LER decreases as the dose (the amount of electron charge given to a unit area) increases (within a reasonable range of dose). (3) As the edge location is moved from the inside of a feature to the outside, the LER decreases. (4) The LER is smaller for a smaller exposing interval. (5) For different types of resist, though the LER level may be different, the similar behaviors of LER are exhibited.

Based on the behaviors of LER, two LER minimization methods have been developed. Unlike the conventional methods [7, 8] where a two dimensional (2-D) model is mostly used, i.e., ignoring any variation along the resist depth dimension, a 3-D model of the substrate system is employed in these two methods to consider the different behaviors of LER and CD error at different resist layers. In one method, only the shape of a feature is adjusted with a uniform dose to minimize the LER, taking the critical dimension (CD) error also into account. In the other method, both the shape of a feature and the spatial distribution of dose within the feature are controlled.

Even though the methods are flexible and accurate, they require an excessive computation in analyzing the relationship between various lithographic parameters and the LER. And such a process needs to be repeated once the set-up of e-beam lithographic process, e.g., resist type, changes.

1.2.2 Analytic Approach

In order to avoid the simulation and experiment, one may take an analytic approach to estimating and minimizing the LER. Most of the efforts of using an analytic approach are in the direction of estimating the effects of LER on the device behaviors [16, 17]. In the work of Croon in 2002 [16], the simple analytical expressions were presented, which calculate the impact of LER on MOSFET parameter fluctuations. It is experimentally demonstrated that the LER has no impact on 80 nm gate length transistors. Simulations show that the LER becomes significant for 32 nm channel length devices. In another study of Diaz in 2001 [17], an analytical model was introduced to represent the effects of LER on both off-state leakage and drive currents for sub-100-nm devices. The model partitions a given device into small unit cells along its width, and each unit cell assumes a constant gate length (i.e., the cell width is small compared to the LER spatial frequency). An analytical model is used to represent the saturated-threshold-voltage dependency on the gate length of unit cell. Even though these theoretical models can be simple and accurate, an analytic expression of LER was not derived. Therefore, the minimization of LER by optimizing the e-beam lithographic parameters analytically is impossible.

Mack has made a lot of efforts in studying and developing analytically the model of LER. In one of his work in 2009 [18], a stochastic approach to modeling resist development was presented. A continuous approximation to the critical ionization (CI) model was derived and shown to closely match the CI model, i.e., the mechanistic model of phenolic polymer-based resist dissolution [34]. It has the potential for providing the theoretical framework required to determine the variance of developing rate in the dissolution path. Since one of the main sources of LER is from the randomness in the resist development process, this model may provide a way of modeling the LER due to such randomness. Later in his work in 2010 [19], a stochastic modeling of resist development was extended in two and three dimensions. The concepts of dynamic scaling in the study of kinetic roughness were applied to the problem of resist development. To verify the results, simulations of resist development in $1+1$ and $2+1$ dimensions were carried out with various amounts of random, uncorrelated noise added to an otherwise uniform developing rate. Even though these studies provide good understandings in the stochastic factors that affect the LER, an analytic expression of LER was not derived.

In another study [20], a comprehensive stochastic model of LER in the optical lithography was formulated with the expression of the standard deviation of the final deprotection level of polymer molecules in the resist. In the optical lithographic process, the variations of photon shot-noise, chemical concentration (spatially distributed molecules in the resist), exposure, and acid-catalyzed reaction diffusion are first analyzed individually. Then, all these variations are combined to get the standard deviation of the final deprotection level of polymer molecules in the resist, which is used to estimate the LER by setting a threshold. However, the physical and chemical mechanisms in the e-beam lithography may not be the same as those in the optical lithography. More importantly, during the derivation, an infinite contrast of development process was assumed to obtain the LER directly from the blocked-polymer latent image without modeling the development process, which is unrealistic.

Even though the LER estimated by our previous analytic method [21] is close to that generated by the simulation, the final expression of LER is not derived as a function of the e-beam lithographic parameters such as dose and, therefore, it is not suitable for the analytic procedure of minimizing the LER with respect to those parameters.

1.3 Objectives

In the previous efforts, either an analytic expression of LER is not derived, or the LER expression is not a function of e-beam lithographic parameters. Another problem is that it may not be accurate enough to be employed in the minimization procedure.

This dissertation addresses a specific issue of estimating and minimizing the LER in the e-beam lithography through an analytic method, i.e., deriving the explicit relationship between the LER and an e-beam lithographic parameter, i.e., dose, and minimizing the LER with respect to the dose. More specifically, optimize the dose with all the other parameters, i.e., beam energy, developing time, etc., fixed. Without having to obtain and refer to the remaining resist profile, the mathematical expression of LER is derived explicitly from the stochastic fluctuation of developing-rate distribution on which the remaining resist profile mainly depends. One of the main advantages is that the proposed method should be able to analyze and minimize the LER caused by the stochastic variation of developing rate in the resist without excessive experiments and simulations. It should be pointed out that even though this study focuses on the stochastic variation of the developing rate, it is expected that other factors can be considered in this method, by adjusting the stochastic developing rate to reflect their direct or equivalent effects on the developing rate. For example, to consider the randomness in the developing process, an additional fluctuation may be added to the developing rate.

The main objectives of this study are

- the analytic estimation and minimization of LER for a single-line pattern,

- the analytic estimation and minimization of LER for a large-scale uniform pattern, and
- the verification of the analytic results through the simulation and experiments.

A specific goal of this study is to obtain an explicit analytic expression of LER which can be used for minimization of LER. In order to achieve the goal, in some steps of derivations, certain assumptions and approximations are made. Therefore, it needs to be pointed out that the derivations of the LER in this dissertation are not completely analytic.

1.4 Organization of Dissertation

The rest of the dissertation is organized as follows:

- Chapter 2 introduces the essential analytic model of e-beam lithography. The concepts are defined and the simulation procedure used to verify the results of the analytic method is described.
- Chapter 3 describes the analytic method of estimating and minimizing the LER for a single-line pattern.
- Chapter 4 extends the analytic method to large-scale uniform patterns, i.e., patterns of lines/spaces.
- Chapter 5 describes a method of extracting the essential information required by the analytic method from the experimental results by modeling the e-beam lithographic process from SEM images.
- Chapter 6 verifies the results of the analytic method using the experimental results first. Then, a thorough comparison between the results of the analytic method and those from the simulation is made by individually controlling the parameters that affect the LER.
- Chapter 7 presents the conclusions of this dissertation and the possible future work.

Chapter 2

Analytic Model and Simulation

In this chapter, the model of e-beam lithographic process is described first. The main concepts in the analytic estimation and minimization of LER are explicitly defined. Then, the simulation, which is used to verify the results of the analytic method, is depicted.

2.1 Analytic Model

2.1.1 Substrate Model

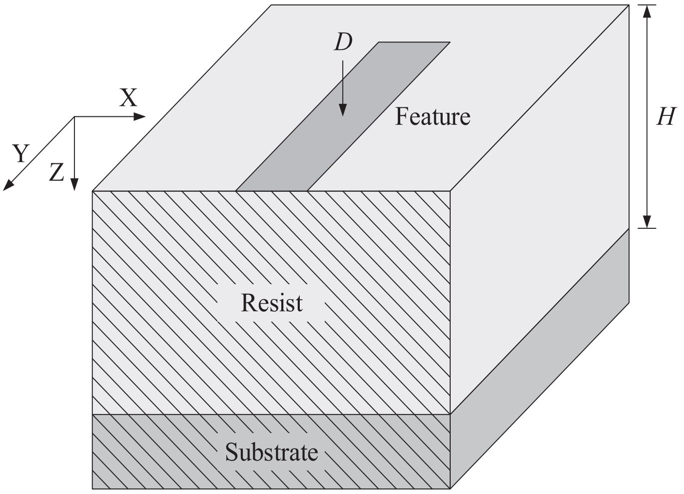


Figure 2.1: A typical 3-D model of substrate system with a resit layer on top of the substrate where a line feature is exposed with a certain dose level D.

A 3-D (three-dimensional) model of substrate system is employed in this study. As shown in Fig. 2.1, a resist layer is on the top of the substrate where the X-Y plane corresponds to the top surface of resist, and the resist depth is along Z dimension. A line feature is exposed with a certain dose level D. In the e-beam lithography, the dose depicts the amount of electron charge given to a unit area, and its unit is μC/cm^2.

2.1.2 Exposure Model

The exposure in the e-beam lithography depicts the energy deposited in the resist. Its distribution can be computed from the dose distribution of a circuit pattern through the convolution with a proximity function or point spread function (PSF), which shows how the energy is distributed throughout the resist when a single point is exposed [24, 25].

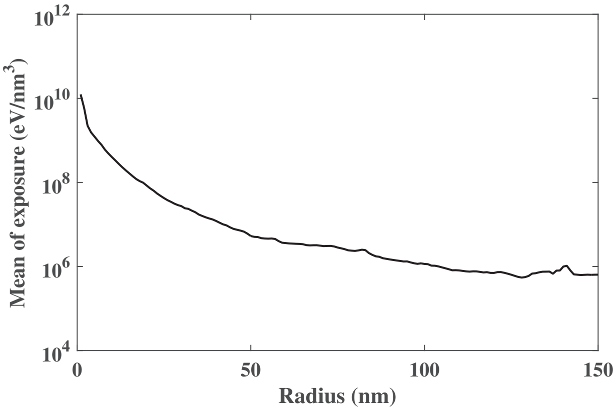
In this study, the 3-D PSF, denoted by $psf(x, y, z)$, is generated by the CASINO Monte Carlo simulation software [33]. By generating a set of stochastic PSF's through the Monte Carlo simulation for the same set-up of e-beam lithographic parameters, i.e., beam energy, resist thickness, etc., the moments of the stochastic PSF's can be extracted. For example, the mean and standard deviation of PSF for the resist thickness, dose, beam diameter, and beam energy of 300 nm PMMA on Si, 640 μC/cm^2, 3 nm, and 50 keV, respectively, are illustrated in Fig. 2.2 for the 1-D case.

Let the dose distribution on the surface of resist be denoted by $D(x, y)$ and the exposure at the point (x, y, z) in the resist by $e(x, y, z)$. Then, the 3-D spatial distribution of exposure can be expressed by the following convolution as illustrated in Fig. 2.3 for the 1-D case:

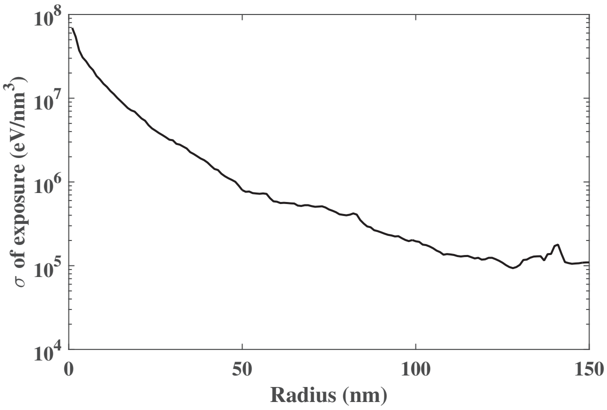
$$e(x, y, z) = \int\int D(x - x', y - y')psf(x', y', z)dx'dy'. \tag{2.1}$$

The exposure $e(x, y, z)$ is stochastic when the PSF $psf(x, y, z)$ is stochastic. For a uniform substrate and a stable e-beam tool, the stochastic properties of PSF must be space-invariant. That is, the mean and variance of PSF (See Fig. 2.2), denoted as $m_{psf}(x, y, z)$ and $\sigma^2_{psf}(x, y, z)$, respectively, are space-invariant. Then, the mean exposure, $m_e(x, y, z)$, can be expressed as

$$m_e(x, y, z) = \int\int D(x - x', y - y')m_{psf}(x', y', z)dx'dy'. \tag{2.2}$$



(a)



(b)

Figure 2.2: The moments of stochastic PSF generated by the Monte Carlo simulation for 300 nm PMMA on Si with a dose of 640 μC/cm^2, beam diameter of 3 nm, and beam energy of 50 keV: (a) mean and (b) standard deviation.

Also, assuming that $psf(x, y, z)$ and $psf(x', y', z)$ are uncorrelated, the variance of exposure, $\sigma_e^2(x, y, z)$, can be derived as

$$\sigma_e^2(x, y, z) \quad = \quad \int\int D(x - x', y - y')\sigma_{psf}^2(x', y', z)dx'dy'. \tag{2.3}$$

2.1.3 Exposure-to-Rate Conversion Formula

The relationship between the exposure and developing rate is generally not linear. During the writing process on a resist, electrons come across the resist and deposit their energy (exposure) which ionizes the phenol groups in the polymer. Each initially blocked (protected) phenol group in the polymer has a chance to be ionized. If the number of ionized phenol groups

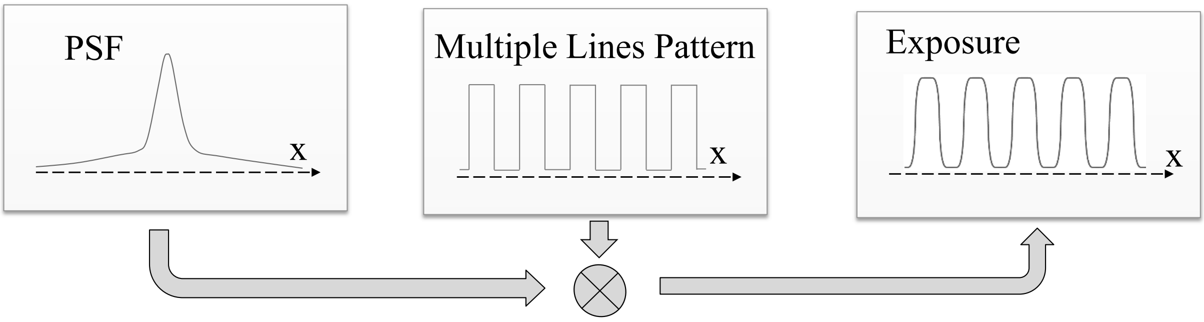


Figure 2.3: From the PSF to the exposure through the convolution.

exceeds the critical fraction, the polymer becomes soluble [18]. The probability of a given phenol group being ionized is directly related to the energy (exposure) deposited in that group, and thus the developing rate in the resist can be derived from the exposure. The developing rate is also related to other factors such as molecular weight and pH of the developer.

In this study, the developing rate $R(x, y, z)$ at each point in the resist is calculated from the exposure $e(x, y, z)$ through a nonlinear exposure-to-rate conversion formula $F[\cdot]$. The conversion formula may be derived experimentally as follows. A single line is exposed with a spatially-uniform dose, and after the resist development process, the depth in the cross-section of remaining resist profile is measured at the center of line. This process is repeated with different dose levels. Note that the resist is developed only vertically at the center of line when the dose is spatially uniform. The resist profile can also be obtained through the resist development simulation by computing the exposure distribution in the resist and converting the exposure into the developing rate. The relationship between the exposure and developing rate (i.e., conversion formula), which minimizes the difference between the measured and simulated depths, is obtained. The following conversion formula $F[\cdot]$ was derived (also refer to Fig. 2.4) using the experimental results where the substrate system is composed of 300 nm PMMA on Si:

$$R(x, y, z) = F[e(x, y, z)] = 3700 \cdot \exp\left(-\left(\frac{e(x, y, z) - 1.0e11}{5.8e10}\right)^2\right) - 152.5 \tag{2.4}$$

where $R(x, y, z)$ is in nm/minute and $e(x, y, z)$ in eV/nm^3.

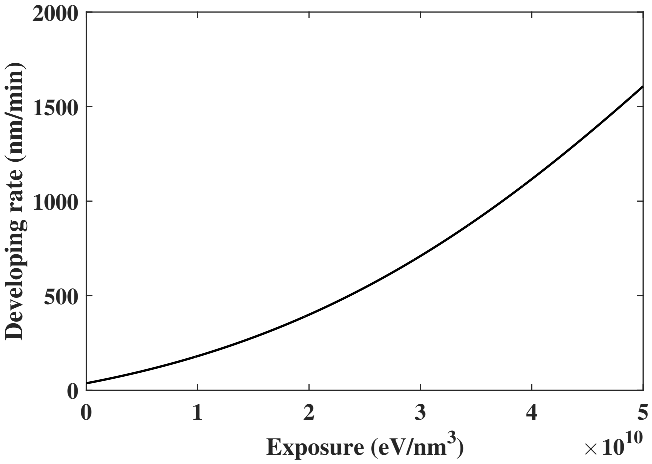


Figure 2.4: The experimentally derived nonlinear exposure-to-rate conversion formula $F[\cdot]$. For the experiments, the Si wafer was pin-coated with 300 nm PMMA and soft-baked at 160 °C for 1 min. The line pattern was exposed with different doses with beam energy of 50 keV, and the sample was developed in MIBK:IPA = 1:2 for 40 s.

The validity of this conversion formula was verified in the previous work [11]. The mean and variance of $R(x, y, z)$ are denoted by $m_R(x, y, z)$ and $\sigma_R^2(x, y, z)$, respectively.

2.1.4 Resist-Development Model

The resist layer is partitioned into cubic cells, and the stochastic exposure is calculated at each cell. The stochastic developing rate $R(x, y, z)$ of each cell is calculated from the exposure $e(x, y, z)$ through the nonlinear exposure-to-rate conversion formula. Given a spatial distribution of the developing rate, the remaining resist profile is determined through the resist development process. Assume that the resist development process can be modeled by "development paths" where a development path is defined as a path along which the resist is developed, as illustrated in Fig. 2.5. Each development path starts from the resist surface toward the boundary of resist profile. Given a possible edge location (x, y, z), there exists only one development path that depicts the development process from the top surface of the resist layer to that edge location, and its length is denoted by s. Note that the LER is proportional to the length fluctuation of the development path of an edge location. The developing rate along the development path is expressed by $R(p)$ where $0 \leq p \leq s$ (see Fig. 2.5). The analytic expression of LER is derived utilizing the concept of development path later.

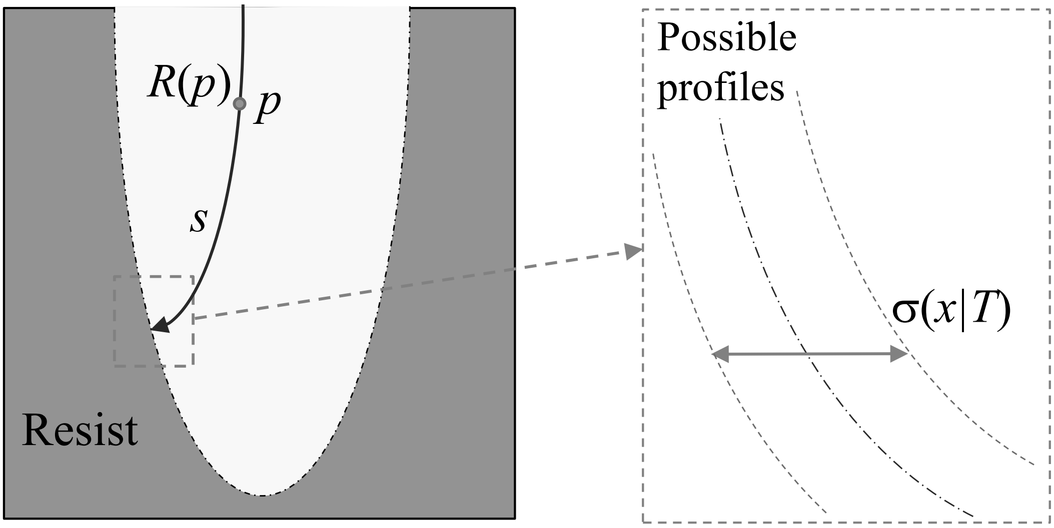


Figure 2.5: The development path in the resist and the variation ($\sigma(x|T)$) of edge location with a fixed developing time. Note that $p = 0$ and $p = s$ correspond to the starting and ending points of a development path, respectively. $R(p)$ is the developing rate at p.

2.1.5 Definition of LER

The LER is defined as the variation of the edge location at a layer of the remaining resist profile shown in Fig. 2.6, and is quantified as the standard deviation of the edge location x measured in the lateral dimension given a developing time T, i.e.,

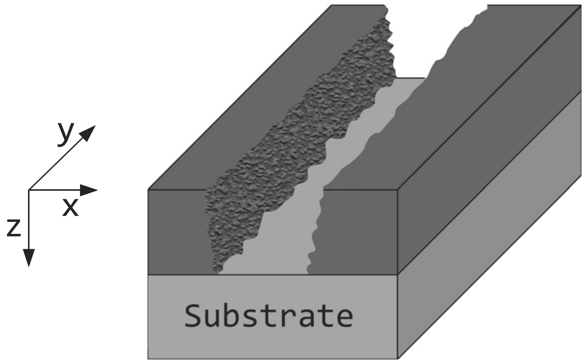
$$LER = \sigma(x|T). \tag{2.5}$$

2.1.6 List of Notations

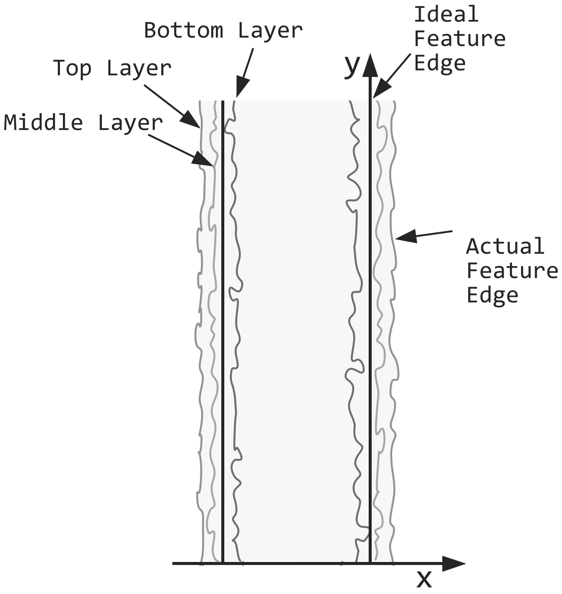
A summary of the notations is shown in Table 2.1.

2.2 Resist-Development Simulation

Given the stochastic developing rate distribution generated by the Monte Carlo simulation, a resist-development simulation is used to derive the remaining resist profile where the CD



(a)



(b)

Figure 2.6: (a)The LER in a 3-D model of substrate system. (b) The LER at each layer is quantified individually as the standard deviation of edge location along the length dimension of a feature.

and LER can be calculated. There are several commonly used simulation methods, e.g., cell removal method, fast marching method.

The Cell removal method is widely used [26, 27, 28] because it has been proved to be stable and robust, and easy to implement. In the 3-D version of the cell removal method, the resist layer is partitioned into cubic cells. The developing process is started from the top surface of the resist layer. In each iteration of the simulation, the surface (sides) of a cell that is in contact with the developer is determined and each side is adjusted (developed) based on the developing rate. If a cell is fully developed, the states of its neighbors are updated, i.e.,

Notation	Defination
$psf(x, y, z)$	Point spread function which depicts the energy distribution when a single point is exposed
$m_{psf}(x, y, z)$	Mean of a set of PSF's at the point (x, y, z)
$\sigma_{psf}(x, y, z)$	Standard deviation of a set of PSF's at the point (x, y, z)
$D(x, y)$	Dose distribution in a pattern
$e(x, y, z)$	Exposure at the point (x, y, z)
$R(x, y, z)$	Developing rate at the point (x, y, z)
$R(p)$	Resist developing rate at p on the development path
$m_R(p)$	Mean of developing rate at p on the development path
$\sigma_R(p)$	Standard deviation of developing rate at p on the development path
T	Developing time
D_0	Uniform-dose level

Table 2.1: Notations and definations

mores sides are considered to be exposed to the developer or the undeveloped cells start to be developed. Given a specified developing time, the remaining resist profile is derived from the boundary between the developed and undeveloped cells. A fast marching level set method [26] is based on level set methods which are numerical techniques for computing the position of propagating fronts. It relies on an initial value of the partial differential equation for propagating a level set function and uses the techniques borrowed from the hyperbolic conservation laws [29].

In this study, a novel method for 3-D resist development simulation [22] is utilized in order to reduce the simulation time of an iterative method such as the cell removal method. It employs the concept of "L-shape development paths," (see Fig. 2.7) which start from the top surface of resist toward the boundaries of the final resist profile, to model the development process.

The resist development process is isotropic as the resist is developed in all possible directions. However, in the path-based method, the development process is modeled by two orthogonal processes, i.e., vertical development and lateral development, in order to simplify the simulation procedure. Accordingly, each L-shape development path (just path hereafter) consists of two orthogonal path segments, i.e., vertical (to depict vertical development) and lateral (to depict lateral development) path segments (refer to Fig. 2.7). More specifically, each

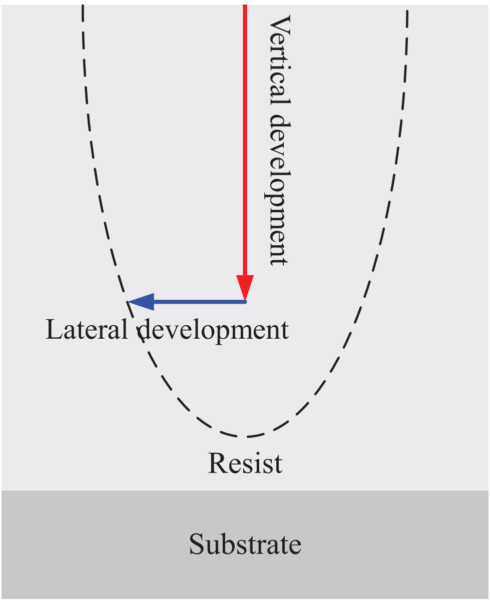
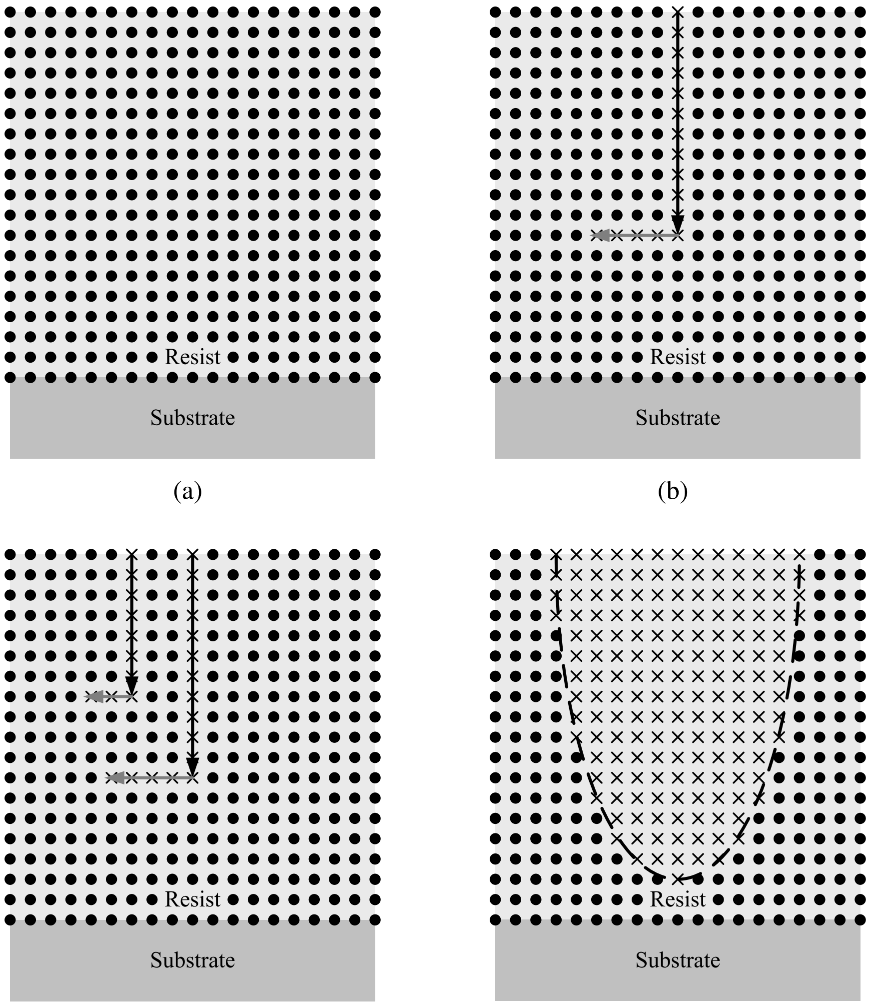


Figure 2.7: Illustration of a development path with the vertical and lateral development where the dashed curve represents a cross-section of resist profile.

path has one vertical path segment (along the Z dimension) and may have one or more lateral path segments (along the X dimension or Y dimension). Each path is “computed” (followed) individually, i.e., finding the farthest point in the resist the path can reach given a developing time T. The computation of each path terminates when the sum of the times spent on its path segments is equal to the given developing time T. The final resist profile is determined by tracing the boundaries between the developed and undeveloped points after all possible paths are computed (refer to Fig. 2.8). The method has been shown to be as accurate as the widely-used methods, e.g., cell-removal and fast-marching methods, but with a much less requirement of computation [22].



(a) (b) (c) (d)

Figure 2.8: (a) All the points in the resist are marked as “undeveloped” (‘·’) before the resist development simulation. (b) After a path is computed, the points which it passes through are marked as “developed” (‘x’). (c) After another path is computed, more points are marked as “developed”. (d) The final resist profile is determined by tracing the boundaries between those developed points and those which are not.

Chapter 3

A Single Line Feature with a Uniform Dose

In this chapter, an analytic method of estimating the LER at any layer of resist is described for the case where a long single line is exposed with a uniform dose. The line is sufficiently long along the Y dimension that any variation, i.e., the mean and variance of developing rate, along the Y dimension can be neglected. The analytic (mathematical) expression of LER is derived explicitly from the stochastic fluctuation of developing rate distribution on which the remaining resist profile mainly depends. Also, the LER is minimized analytically with respect to the dose with all the other e-beam lithographic parameters, i.e., beam energy, developing time, etc., fixed. Note that the goal is to obtain an explicit analytic expression of LER for a single line feature with a uniform dose, which can be used for minimization of LER, and to achieve the goal, the derivation of LER is not completely analytic, i.e., in some steps certain approximations are made or an iterative method is employed.

The rest of the chapter is organized as follows. The analytic estimation of LER is described in Sec. 3.1. The procedure for minimizing the LER is presented in Sec. 3.2. The results of estimating and minimizing the LER are discussed in Sec. 3.3, followed by a summary in Sec. 3.4.

3.1 Estimation of LER

In this section, the LER is analytically estimated as a function of the edge location with the dose fixed at a certain level D and, therefore, D is not explicitly considered in the estimation procedure. However, in Sec. 3.2, to minimize the LER and CD error with respect to the dose level, the relationship of edge location versus D is included where D is a variable and will be explicitly considered.

The derivation of the analytic expression of LER is carried out in three steps: (1) given an edge location x at a layer (z_0) at which the LER is to be estimated, find the development path reaching the location and derive the fluctuation of developing time T taken by the development path, (2) estimate the direction of development path at the edge location x, and (3) derive the expression of LER given a developing time T.

3.1.1 Development Path and Fluctuation of Developing Time

An edge location is determined by the development path reaching the location. To estimate the LER at an edge location (x, z_0), its development path needs to be found first. However, since the developing process is isotropic, it is not easy to derive the development path analytically in one step. From the distribution of developing rate, the time taken by a path (from the top surface of the resit layer to the edge location (x, z_0)) can be calculated. Given the idea that the development path of an edge location is the path that consumes the minimum time to reach (x, z_0) among all possible paths, i.e., any path from the top of the resist surface to (x, z_0) (see Fig. 3.1), an iterative method is employed to derive the development path shown in Fig. 3.1).

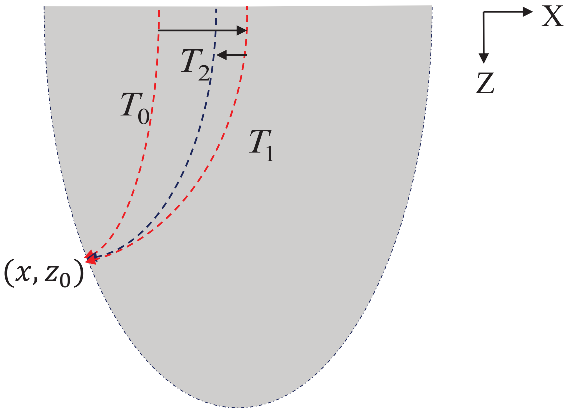


Figure 3.1: The illustration of iterative method to derive the development path by tracing back and forth from (x, z_0). The time consumed by each path is calculated along the path, i.e., T_0, T_1, and T_2, where $T_0 > T_1 > T_2$.

The iterative method may be described as follows. In the first step, the iterative method traces back from the edge location to the top surface of resist. For example, given the edge location (x, z_0), the point (x_1, z_1) that takes the minimum developing time from (x, z_0) to (x_1, z_1) among all possible x_1, where $z_1 < z_0$, is found. From (x_1, z_1), the next point (x_2, z_2) can be derived by the same procedure. Repeating this procedure until $z_i = 0$ which corresponds to the top surface of resist. Then, $\{(x, z_0), (x_1, z_1), (x_2, z_2), ...\}$ represents a possible development path to (x, z_0). Note that even though the developing time between two adjacent points on the path is the minimum, the time consumed by this path may not be the minimum among all possible paths. Therefore, in the second step, with (x, z_0) fixed, additional possible paths are found by adjusting each of the remaining points and tracing back from the point as in the first step. Of all the derived paths, the path that takes the minimum time is taken as the development path for the edge location (x, z_0).

The fluctuation of developing time can be directly derived from the fluctuation of developing rate on the development path. Given $R(p)$ on the development path (see Fig. 2.5), the developing time T is expressed as the integral of $\frac{1}{R(p)}$ along the path,

$$T = \int_0^s \frac{dp}{R(p)}. \tag{3.1}$$

Assuming that the correlation of $\frac{1}{R(p)}$ at any two points (p_1, p_2) on the development path denoted as $Cov_{\frac{1}{R}}(p_1, p_2)$ is known, according to Eq. 3.1, the mean and variance of developing time are computed as

$$\begin{aligned} m(T|s) &= \int_0^s m_{\frac{1}{R}}(p)dp, \\ \sigma^2(T|s) &= \int_0^s \sigma^2_{\frac{1}{R}}(p)dp + \int_0^s \int_0^s Cov_{\frac{1}{R}}(p_1, p_2)dp_2 dp_1. \end{aligned} \tag{3.2}$$

3.1.2 The Direction of the Development Path

The direction of resist development is normal to the surface of remaining resist profile, and is not always lateral (X dimension). Therefore, the variation of the development-path length may not be quantitatively the same as the LER to be measured in the lateral dimension shown in

Fig. 3.2. The direction of the development path is expressed as the angular deviation θ from the lateral direction and needs to be derived to get an accurate estimate of LER.

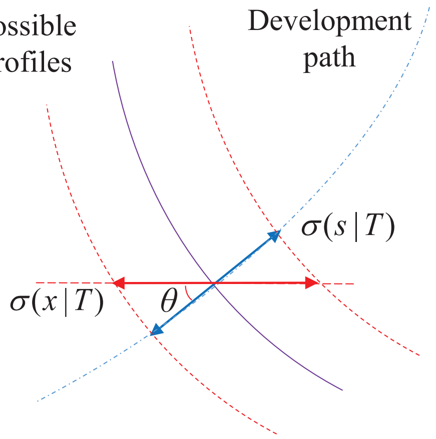


Figure 3.2: The fluctuation of the development-path length is "projected" onto the X (lateral) dimension.

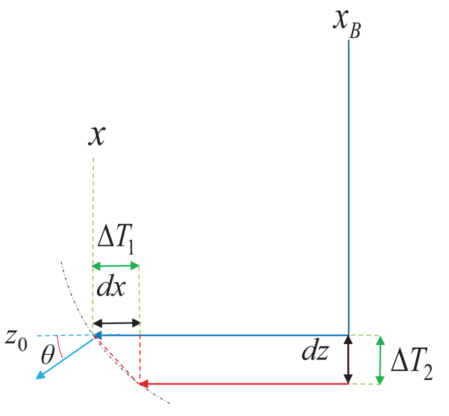
To facilitate the derivation of angle θ, the resist development process is modeled by "L-shape" paths (see Sec. 2.2). Note that the remaining resist profiles estimated using L-shape paths are as accurate as those by the cell removal and fast marching methods [22].

In the region close to the feature edge, the resist development usually progresses both laterally and vertically (see Fig. 3.3(a)). On the other hand, the resist development progresses mostly vertically in the region sufficiently away from the feature edge (see Fig. 3.3(b)). Let x_B represent the boundary between the two regions.

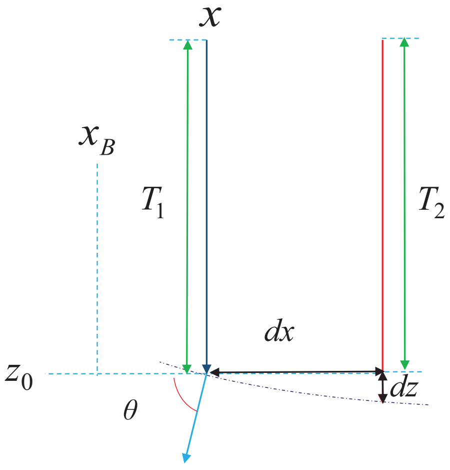
For a possible edge location (x, z_0) close to the feature (left) edge ($x \leq x_B$), the L-shape path starting from $(x_B, 0)$ with its lateral segment on the X-Y plane at $z = z_0$ is the development path reaching (x, z_0) shown in Fig. 3.3(a). Consider another L-shape path from $(x_B, 0)$, which reaches an adjacent X-Y plane at $z = z_0 + dz$, ending at $(x_B - dx, z_0 + dz)$. When dz is sufficiently small, the angle θ is derived as

$$\theta(x, z_0) = \tan^{-1}\left(\frac{dx}{dz}\right) = \tan^{-1}\left(\frac{m_R(x, z_0) \cdot \Delta T_1}{m_R(x_B, z_0 + dz) \cdot \Delta T_2}\right) \qquad (3.3)$$

where ΔT_1 is the difference of developing time between the two development paths in the lateral dimension, and ΔT_2 is that in the vertical dimension as illustrated in Fig. 3.3(a).



(a)



(b)

Figure 3.3: The direction of development path in a region (a) close to the left edge of a feature and (b) close to the feature center.

Since the two paths take the same developing time and share the segments which also take the same amount of developing time except ΔT_1 and ΔT_2, ΔT_1 must be equal to ΔT_2 when dz is close to 0. The angle θ can be simplified to

$$\theta(x, z_0) = \tan^{-1}\left(\frac{m_R(x, z_0)}{m_R(x_B, z_0)}\right). \tag{3.4}$$

For an edge location close to the feature center ($x > x_B$), the L-shape path can be approximated to consist of only a vertical segment as shown in Fig. 3.3(b). Consider two adjacent paths starting from $(x, 0)$ and $(x+dx, 0)$ reaching the edge location (x, z_0) and $(x+dx, z_0+dz)$, respectively. Since the two paths take the same developing time, the following relationships are derived:

$$\begin{aligned}
T_1 &= \int_0^{z_0} m_{\frac{1}{R}}(x,z)dz, \\
T_2 &= \int_0^{z_0} m_{\frac{1}{R}}(x+dx,z)dz, \\
T_1 &= T_2 + dz \cdot m_{\frac{1}{R}}(x+dx,z_0).
\end{aligned} \tag{3.5}$$

From Eq. 3.5,

$$\int_0^{z_0} m_{\frac{1}{R}}(x,z)dz = \int_0^{z_0} m_{\frac{1}{R}}(x+dx,z)dz + dz \cdot m_{\frac{1}{R}}(x+dx,z_0) \tag{3.6}$$

and

$$\begin{aligned}
\frac{dz}{dx} &= \frac{\int_0^{z_0} m_{\frac{1}{R}}(x,z)dz - \int_0^{z_0} m_{\frac{1}{R}}(x+dx,z)dz}{dx} \cdot \frac{1}{m_{\frac{1}{R}}(x+dx,z_0)} \\
&= -\int_0^{z_0} \frac{\partial m_{\frac{1}{R}}(x,z)}{\partial x} dz \cdot \frac{1}{m_{\frac{1}{R}}(x+dx,z_0)}.
\end{aligned} \tag{3.7}$$

When $dz \to 0$ and $dx \to 0$, the angle θ is derived as

$$\theta(x,z_0) = \tan^{-1}\left(\frac{dx}{dz}\right) = \tan^{-1}\left(-\frac{m_{\frac{1}{R}}(x,z_0)}{\int_0^{z_0} \frac{\partial m_{\frac{1}{R}}(x,z)}{\partial x} dz}\right). \tag{3.8}$$

In the region close to the feature edge, the angle θ evaluated by Eq. 3.4 is larger than that by Eq. 3.8 since the developing rate varies fast in the lateral dimension. On the other hand, in the region close to the feature center, the angle θ by Eq. 3.8 is larger than that by Eq. 3.4. Therefore, in this study, instead of deriving x_B analytically which is not straightforward, the angle θ is evaluated with the approximation of $m_R(x_B, z_0) \approx m_R(\frac{W}{2}, z_0)$ as follows:

$$\theta(x,z_0) = \max\left(\tan^{-1}\left(\frac{m_R(x,z_0)}{m_R(\frac{W}{2},z_0)}\right), \tan^{-1}\left(-\frac{m_{\frac{1}{R}}(x,z_0)}{\int_0^{z_0} \frac{\partial m_{\frac{1}{R}}(x,z)}{\partial x} dz}\right)\right). \tag{3.9}$$

As can been seen from the equation, the angle θ is determined by the average developing-rate distribution in the resist. When it is close to the feature edge, the angle θ is only determined by the mean of developing rate R at the layer ($z = z_0$). However, when it is close to the feature center, the angle θ is affected accumulatively by the mean of $\frac{1}{R}$ from resist surface ($z = 0$) to the edge location ($z = z_0$).

3.1.3 LER at a Layer

The fluctuation of the development-path length (s given T) may be related to the fluctuation of developing time taken by this path. If the developing rate is constant (R) around s, the fluctuation of developing time can be directly converted into the fluctuation of s given T as

$$\sigma(s|T) = \sigma(T|s) \cdot R. \tag{3.10}$$

In reality, the developing rate varies spatially and is stochastic at each point. However, it would be a reasonable approximation that the developing rate around s is represented by the mean of developing rate at s, i.e., $m_R(s)$ (refer to Fig. 3.4). Then, a more realistic relationship between $\sigma(s|T)$ and $\sigma(T|s)$ can be expressed as

$$\sigma(s|T) \approx \sigma(T|s) \cdot m_R(s). \tag{3.11}$$

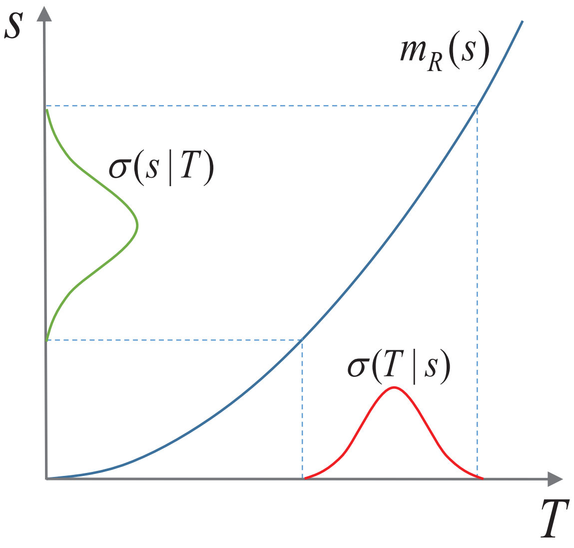


Figure 3.4: The $\sigma(s|T)$ is derived from $\sigma(T|s)$ through the mean of developing rate ($m_R(s)$).

Since the LER is measured in the lateral dimension (X dimension), $\sigma(s|T)$, which is derived along the direction of development path, needs to be projected onto the X dimension (see Fig. 3.2). According to the definition of LER,

$$LER(x, z_0) = \sigma(x|T) = \frac{\sigma(T|s) \cdot m_R(x, z_0)}{\cos(\theta(x, z_0))}. \tag{3.12}$$

As can be seen from the equation, the LER at (x, z_0) is determined by three parts, the standard deviation of developing time T given s, average developing rate, and angle θ, i.e., $\sigma(T|s)$, m_R, and $\cos(\theta)$. The m_R and $\cos(\theta)$ are both functions of the location (x, z_0). Given an edge location (x, z_0), its development path and the length s of this path are determined and, therefore, $\sigma(T|s)$ calculated along this path at s is uniquely determined by (x, z_0), i.e., $\sigma(T|s)$ is a function of (x, z_0).

3.1.4 Adjustment

It needs to be pointed out that the interaction between adjacent development paths is not taken into account in the above derivation. The interaction tends to decrease the variation (of path length) and reduce the LER. Therefore, an adjustment factor needs to be employed to compensate the interaction between the paths.

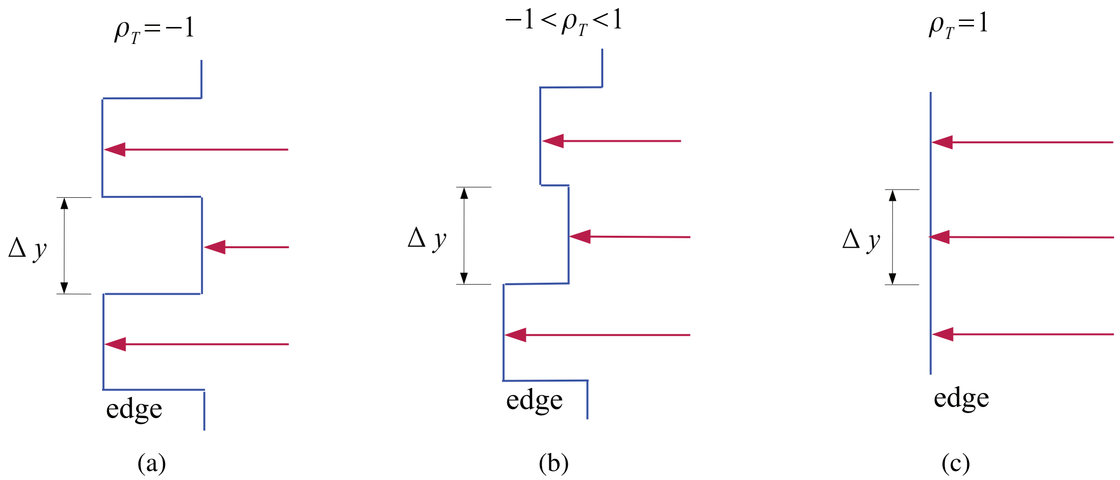


Figure 3.5: The interaction between paths when the correlation of developing time between adjacent paths is (a) $\rho_t = -1$, (b) $-1 < \rho_t < 1$, (c) $\rho_t = 1$.

The correlation coefficient of developing time ($\rho_T(s)$) between adjacent development paths is used to derive the adjustment factor, which is derived by the integral of the covariance of $\frac{1}{R}(s)$ along the path. When $\rho_T(s)$ is closer to -1 (illustrated in Fig. 3.5 (a)), the developing time for the same edge location shows a larger difference between adjacent development paths. This leads to a higher level of interaction between the adjacent paths and, therefore, a larger adjustment (reduction) is to be made. On the other hand, when $\rho_T(s)$ is closer to 1 (illustrated in Fig. 3.5 (c)), the adjacent development paths share the same edge location, i.e., a lower level of interaction, and thus the adjustment (reduction) needs to be smaller. A reasonable multiplicative adjustment factor may be formulated as $1 - \alpha\frac{1-\rho_T}{2}$ where α determines the maximum adjustment and $0 < \alpha < 1$. Note that the adjustment factor is 1 (i.e., no adjustment) when ρ_T is 1 and $1 - \alpha$ when ρ_T is -1. And a reasonable adjustment factor may be expressed as

$$\left(1 - \alpha\frac{1 - \rho_T(s)}{2}\right). \tag{3.13}$$

The correlation coefficient ρ_T is computed between adjacent development paths and, therefore, is affected by the distance between the adjacent paths, i.e., sampling interval, denoted by Δy. For example, a smaller Δy makes $\rho_T(s)$ larger (closer to 1). Hence, this effect needs to be considered. Also, the larger $\sigma(s|T)$ is, a larger adjustment is to be made in general. If $\sigma(s|T)$ is much smaller than Δy, e.g., $\frac{\sigma(s|T)}{\Delta y}$ close to zero, there is almost no interaction between adjacent development paths even when $\rho_T(s) = -1$ shown in Fig. 3.5 (a). On the other hand, when $\sigma(s|T)$ is much larger than Δy and $\rho_T(s) = -1$, most of the difference (LER) between adjacent development paths can be compensated due to the interaction. Considering these two facts and noting that α needs to be limited to 1, α must be quantified to be $\min(1, \frac{\sigma(s|T)}{\Delta y})$. In this study, the Δy is fixed to 1 nm and $\sigma(s|T)$ is normally larger than 1 nm. Therefore, to simplify the equation, the adjustment factor is expressed as

$$\left(1 - \frac{1 - \rho_T(s)}{2}\right) = \frac{\rho_T(s) + 1}{2}. \tag{3.14}$$

Finally, the estimate of LER is given as

$$LER(x, z_0) = \frac{\rho_T(s) + 1}{2} \frac{\sqrt{\left(\int_0^s \sigma^2_{\frac{1}{R}}(p)dp + \int_0^s \int_0^s Cov_{\frac{1}{R}}(p_1, p_2)dp_2dp_1\right) \cdot m_R(x, z_0)}}{\cos(\theta(x, z_0))}. \quad (3.15)$$

As can be seen from Eq. 3.15, given a location (x, z_0), the fluctuation of $\frac{1}{R}$ and the average developing rate at (x, z_0) together determine the fluctuation of developing-path length and have a positive relationship (correlation) with the LER. When the direction of development process is more lateral (angle θ is closer to 0), this fluctuation leads to a larger LER. And the interaction between paths tends to reduce the LER.

3.2 Minimization of LER

The analytic method of minimizing the LER based on Eq. 3.15 is described in this section. In this study, the LER is minimized with respect to the dose which is uniformly applied to the entire area of a feature. Since the LER and CD error are both functions of edge location, the optimal edge location is first derived without explicitly involving the dose, and then the optimal dose is related to the optimal edge location, without introducing an extra complexity. The LER minimization is specific to a certain layer, e.g., the bottom layer as it has the largest LER among all layers and, therefore, the notation of LER, i.e., $LER(x, z_0)$, is simplified to $LER(x)$.

3.2.1 Optimal Path Length

As the dose increases, the edge location moves from the inside of a feature to the outside. Inside the exposed area of a feature, the LER decreases rapidly toward the boundary of the exposed area [15]. It continues to decrease over the boundary and right outside the exposed area, and then almost levels off or slightly increases in some cases. Therefore, the LER is not minimal at the feature boundary (where the CD error is minimal, i.e, zero) in general. That is, the LER and CD error cannot be minimized at the same time (the same edge location or equivalently the same dose). The cost function $C(x)$ employed in this study involves both CD error and LER as follows:

$$C(x) = CD\ error + 3 \cdot LER = |x - x_t| + 3 \cdot LER(x) \tag{3.16}$$

where x_t is the target edge location and its corresponding path length is denoted by s_t.

Since the LER is the standard deviation of edge location, the cost function of $CD\ error + 3 \cdot LER$ may be considered practically as the worst-case deviation of edge location from the target location (e.g., 99.7% in the case of Gaussian distribution). Also, $3 \cdot LER$ corresponds to the 3σ-LER typically used in the research and industry communities.

When the edge location is inside the feature, e.g., $x < x_t$ close to the right edge, both LER and CD error decrease as x increases. Therefore, the optimal x cannot be inside the feature. When $x \geq x_t$,

$$C(x) = x - x_t + 3 \cdot LER(x). \tag{3.17}$$

Around the target edge location x_t, the remaining resist profile is almost vertical and, therefore, $\theta(x) \approx 0$ and $\cos\theta(x) \approx 1$.

The s_t denotes the corresponding path length for the target location x_t. Since $\sigma(T|s)$ changes very slowly around s_t,

$$\sigma(T|s) \approx \sigma(T|s_t) = \sqrt{\left(\int_0^{s_t} \sigma^2_{\frac{1}{R}}(p)dp + \int_0^{s_t}\int_0^{s_t} Cov_{\frac{1}{R}}(p_1, p_2)dp_2 dp_1\right)}. \tag{3.18}$$

A similar approximation can be made to the $\rho_T(s)$, i.e., $\rho_T(s(x)) \approx \rho_T(s_t)$.

From Eq. 3.18, the LER (Eq. 3.15) around x_t may be expressed as

$$\begin{aligned} LER(x) &= \frac{\rho_T(s_t)+1}{2} \cdot \sigma(T|s_t) \cdot m_R(x) \\ &= \frac{\rho_T(s_t)+1}{2} \sqrt{\left(\int_0^{s_t} \sigma^2_{\frac{1}{R}}(p)dp + \int_0^{s_t}\int_0^{s_t} Cov_{\frac{1}{R}}(p_1, p_2)dp_2 dp_1\right)} \cdot m_R(x). \end{aligned} \tag{3.19}$$

Assuming that the cost function is a simple convex function, the optimal edge location s_{opt} is obtained from

$$\frac{dC(x)}{dx} = 1 + \frac{3(\rho_T(s_t)+1)}{2} \cdot \sigma(T|s_t) \cdot m'_R(x) = 0 \tag{3.20}$$

where $m'_R(x)$ is the first order derivative of the mean developing rate $m_R(x)$ with respect to x.

The optimal edge location x_{opt} is derived as

$$\begin{aligned} x_{opt} &= {m'_R}^{-1}\left(\frac{2}{3(\rho_T(s_t)+1)\cdot\sigma(T|s_t)}\right) \\ &= {m'_R}^{-1}\left(\frac{2}{3(\rho_T(s_t)+1)\cdot\sqrt{\left(\int_0^{s_t}\sigma^2_{\frac{1}{R}}(p)dp + \int_0^{s_t}\int_0^{s_t} Cov_{\frac{1}{R}}(p_1,p_2)dp_2dp_1\right)}}\right) \end{aligned} \tag{3.21}$$

where ${m'_R}^{-1}(\,)$ is the inverse function of $m'_R(\,)$.

3.2.2 Optimal Dose

From this point on, the dose D is considered explicitly, which affects the developing rate. Let $m_{\frac{1}{R}}(x, D)$ denote the mean of $\frac{1}{R}$ at x when the dose is D. D_t and D_{opt} are the doses required to reach x_t and x_{opt} in T, respectively. D_t is assumed to be known, e.g., derived by the simulation or experiment. Note that $D_{opt} > D_t$ since $x_{opt} > x_t$.

In order to relate the optimal dose to the optimal edge location, the developing time, ΔT, taken by the development path to go from x_t to x_{opt} when $D = D_{opt}$ is computed. Since T is the same for both D_t and D_{opt}, ΔT is approximately equal to the change (decrease) in the developing time, required to reach x_t when the dose is changed (increased) from D_t to D_{opt}.

$$\Delta T = \int_{x_t}^{x_{opt}} m_{\frac{1}{R}}(x, D_{opt})dx \approx s_t \cdot \left(\bar{m}_{\frac{1}{R}}(D_{opt}) - \bar{m}_{\frac{1}{R}}(D_t)\right) \tag{3.22}$$

where $\bar{m}_{\frac{1}{R}}(D)$ is defined as $\frac{1}{s_t}\int_0^{s_t} m_{\frac{1}{R}}(p, D)dp$ (see Eq. 3.2), i.e., the average time required to develop a unit length.

The relationship between the dose and developing rate is non-linear and, therefore, $\bar{m}_{\frac{1}{R}}(D)$ is a non-linear function of D. However, a first-order approximation may be employed to evaluate Eq. 3.22. That is,

$$s_t \cdot \left(\bar{m}_{\frac{1}{R}}(D_{opt}) - \bar{m}_{\frac{1}{R}}(D_t)\right) \approx s_t \cdot \bar{m}'_{\frac{1}{R}}(D_t) \cdot (D_{opt} - D_t) \tag{3.23}$$

where $\bar{m}'_{\frac{1}{R}}(\ )$ is the first-order derivative of $\bar{m}_{\frac{1}{R}}(\ )$ with respect to D.

As an approximation, ΔT may be computed with $D = D_t$ instead of $D = D_{opt}$ in Eq. 3.22, i.e., $\int_{x_t}^{x_{opt}} m_{\frac{1}{R}}(x, D_{opt})dx \approx \int_{x_t}^{x_{opt}} m_{\frac{1}{R}}(x, D_t)dx$. Then,

$$\int_{x_t}^{x_{opt}} m_{\frac{1}{R}}(x, D_t)dx = s_t \cdot \bar{m}'_{\frac{1}{R}}(D_t) \cdot (D_{opt} - D_t). \tag{3.24}$$

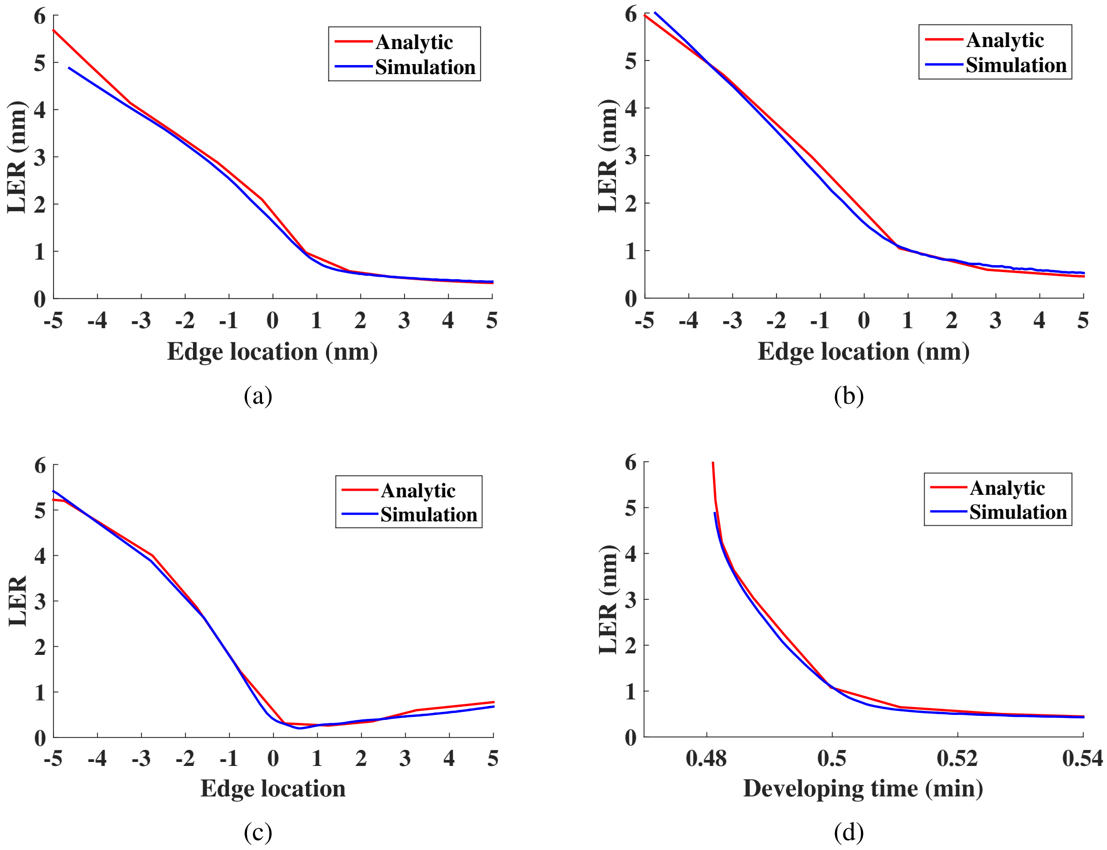
From Eq. 3.24, the optimal dose is derived as

$$\begin{aligned} D_{opt} &= \frac{\int_{x_t}^{x_{opt}} m_{\frac{1}{R}}(x, D_t)dx}{x_t \cdot \bar{m}'_{\frac{1}{R}}(D_t)} + D_t \qquad (3.25) \\ &= \frac{\displaystyle\int_{x_t}^{(m'_R)^{-1}\left(\frac{2}{3(\rho_T(s_t)+1)}\left(\sqrt{\int_0^{s_t}\sigma^2_{\frac{1}{R}}(s')ds' + \int_0^{s_t}\int_0^{s_t} Cov_{\frac{1}{R}}(p_1,p_2)dp_2dp_1}\right)^{-1}\right)} m_{\frac{1}{R}}(x, D_t)dx}{x_t \cdot \bar{m}'_{\frac{1}{R}}(D_t)} + D_t. \end{aligned}$$

Note that D_{opt} can be computed given the distribution of developing rate in the resist.

3.3 Results and Discussion

In order to verify the accuracy of the proposed analytic method, the estimation and minimization results are compared with those obtained through the simulation. A typical substrate system is employed where a resist layer of PMMA with a certain thickness is on the top of the Si substrate. A single feature of size 90 nm × 400 nm ($W \times L$) is exposed with a uniform dose. The LER computed from the analytic expression is compared with that obtained by the simulation for varying edge location (x) where $x = 0$ corresponds to the target edge location on the right edge, i.e., $x < 0$ and $x > 0$ correspond to the inside and outside of a feature. The edge location is controlled by changing the developing time with the same exposure (developing rate) distribution. In the simulation, the development process continues until the boundary of developed feature reaches each edge location.



(a) (b) (c) (d)

Figure 3.6: The LER estimated by the analytic method and simulation: (a) 300 nm PMMA on Si, beam energy of 50 keV, dose of 640 μC/cm^2, (b) 300 nm PMMA on Si, beam energy of 30 keV, dose of 640 μC/cm^2, (c) 100 nm PMMA on Si, beam energy of 50 keV, dose of 600 μC/cm^2, and (d)300 nm PMMA on Si, beam energy of 50 keV, dose of 640 μC/cm^2.

While the analytic method can be applied to any layer of resist, the results for the bottom layer are provided since the LER is usually largest at the bottom layer. In Fig. 3.6(a), the results for the resist thickness of 300 nm, beam energy of 50 keV, and dose level of 640 μC/cm^2 are provided. It can be seen that the LER estimated by the analytic method is well matched with the LER by the simulation. In Fig. 3.6(b), the results for a lower beam energy of 30 keV with the same dose are shown. Since the exposure is higher for 30 keV than 50 keV, for the same edge location, the developing time is shorter for 30 keV than for 50 keV. In Fig. 3.6(c), the results for a thinner resist (100 nm) are given, with a lower dose level of 600 μC/cm^2. A similar observation, i.e., a close match between the analytic and simulation results, can be made.

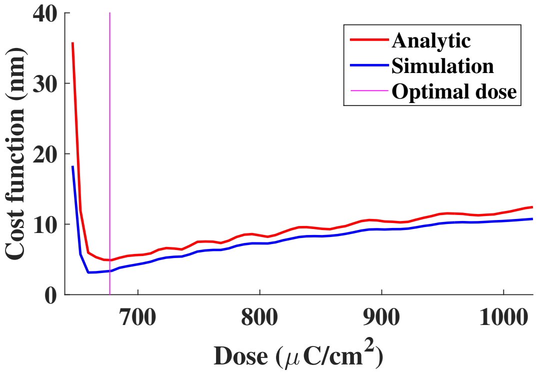
One observation to make is that the LER outside the feature ($x > 0$) in Fig. 3.6(c) increases while it continues to decrease in Fig. 3.6(a) and Fig. 3.6(b). Given a stochastic exposure distribution, the LER generally depends on two factors, the exposure fluctuation and contrast. A higher exposure contrast makes the LER smaller. On the other hand, ignoring the effect of the exposure contrast, the LER would be larger at a more positive edge location (i.e., further outside or equivalently a longer development path) since the exposure fluctuation is accumulated over a longer development path. These effects compete with each other. The LER may increase or decrease outside a feature depending which effect becomes more dominant. That is, the exposure contrast outside the feature in Fig. 3.6 (a) and (b) is relatively higher than that in Fig. 3.6(c).

For an edge location, the developing time is the same for both the analytic method and simulation. In Fig. 3.6(d), the LER is plotted as a function of developing time for both the analytic method and simulation (for the case of Fig. 3.6(a)). However, the shape of LER with the developing time does not have a turning point even when it is developed to the outside. A possible reason is that it normally takes a much longer time to develop further outside. That is, the turning point has not been reached in Fig. 3.6(d).

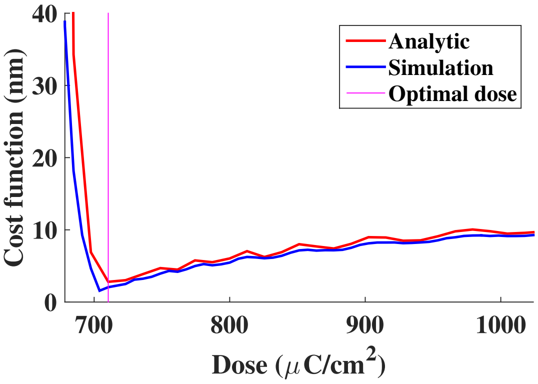
The cost function is evaluated by both the analytic method and simulation and the optimal dose, which minimizes the cost function, is computed by the analytic method, i.e., evaluating Eq. 3.26. The analytic minimization of LER and CD error by controlling the dose is compared with the minimization by the simulation with a fixed developing time. In Fig. 3.7(a), the results for the resist thickness of 300 nm and beam energy of 50 keV are provided. The cost function evaluated by the analytic method is well matched with that by the simulation, and the optimal dose obtained by the analytic method is very close to the optimal dose by the simulation. In Fig. 3.7(b) and Fig. 3.7(c), the results for a lower beam energy of 30 keV and the results for a thinner resist (100 nm) are shown, respectively. It is noted in Fig. 3.7(c) that the cost function shows a flat portion. When the exposure contrast over the feature edge is high (in general, higher for an upper layer of resist, a thinner resist, a higher beam energy, etc.), the exposure outside (even right outside) a feature is very low. In such a case, the width of developed feature increases very slowly as the dose (or developing time) increases, leading to the flat portion of the cost function (see Fig. 3.7(c)). One may still use the optimum dose or the lowest dose in the flat portion in practice since the cost function does not vary substantially in that region.

3.4 Summary

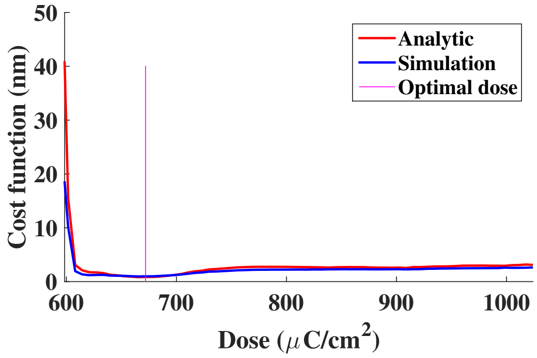
In this chapter, an analytic method of estimating and minimizing the LER is described. It has been shown that the results obtained by this analytic method of estimating and minimizing the LER are closely matched with the simulation results. Although this method focuses on the stochastic variation of developing rate, the analytic method described in this chapter is still applicable even when other factors contributing to the LER are taken into account, e.g., the randomness in the development process may be converted into an extra fluctuation in developing rate and reflected in the LER estimation.



(a)



(b)



(c)

Figure 3.7: The minimization results (minimization of the cost function with respect to the dose) by the analytic method and simulation: (a) 300 nm PMMA on Si, beam energy of 50 keV, developing time of 50 s, (b) 300 nm PMMA on Si, beam energy of 30 keV, developing time of 20 s, and (c) 100 nm PMMA on Si, beam energy of 50 keV, developing time of 10 s. The optimal dose computed by the analytic method is indicated by the vertical line in the graphs.

Chapter 4

A Large-scale Uniform Pattern with a Uniform Dose

In Chapter 3, an analytic method of estimating and minimizing the LER at any layer of resist is described for the case where a long single line is exposed with a uniform dose, denoted by D_0. In this chapter, an extension of the analytic method to a large-scale uniform pattern is described. In general, it is not straightforward to apply the single-line result of LER expression to a large circuit pattern since the exposure level varies with the location within the pattern (see Fig. 4.1).

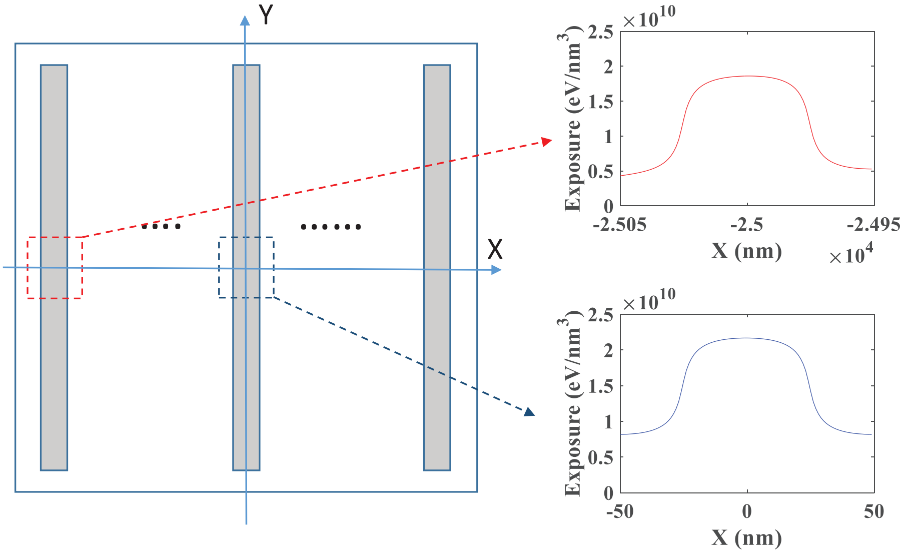


Figure 4.1: The exposure at the edge (red curve) and center (blue curve) of a line-space pattern generated by the Monte Carlo simulation: 300 nm PMMA on Si, dose of 640 μC/cm^2, beam energy of 50 keV, and pattern size of 50 μm $\times$ 50 μm where both the line-width and space are 50 nm.

However, for a large uniform pattern such as a line-space (L/S) pattern where the same feature is replicated uniformly, the exposure level varies gradually in space. Therefore, given the LER expression for a single line (Eq. 3.15), it is possible to derive the LER at a location in a large uniform pattern analytically by adjusting this expression depending on the location in the pattern. To simplify the derivation procedure and the final expression of LER, the analytic expression of LER is derived at three critical locations, i.e., the center, edge and corner (see Fig. 4.2), by modeling the differences of the stochastic exposure distribution among the locations and incorporating them into the single-line result. The LER at other locations may be obtained through an interpolation using the LER's at the critical locations.

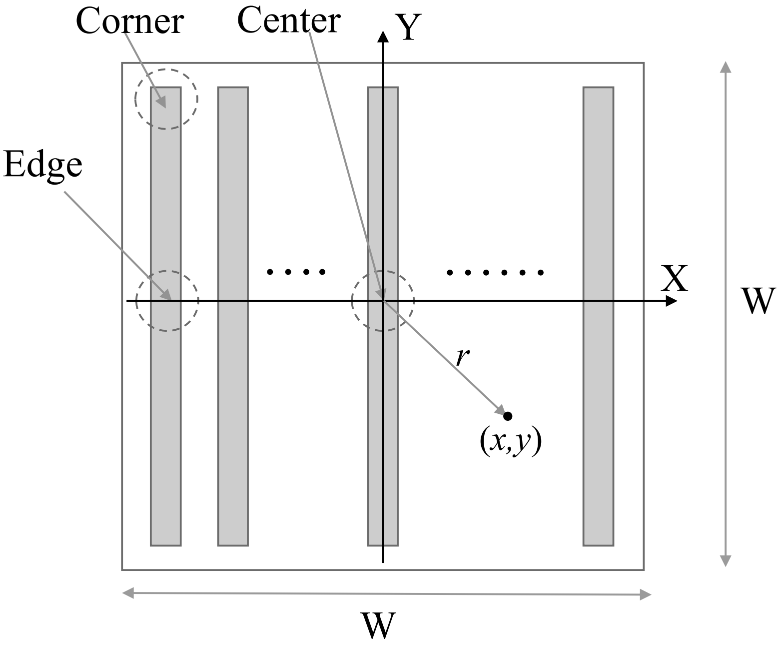


Figure 4.2: A large-scale uniform pattern with line features where three critical regions (corner, edge and center) are shown. The size of the pattern is $W \times W$.

The rest of this chapter is organized as follows. The analytic method of estimating the LER in a large uniform pattern is described in Sec. 4.1. The analytic method of minimizing the LER in a large uniform pattern is described in Sec. 4.2. The results of estimating and minimizing the LER in a large-scale uniform pattern are provided and discussed in Sec. 4.3, followed by a summary in Sec. 4.4.

4.1 Estimation of LER

The exposure level in a large pattern significantly varies with the location (see Fig. 4.1) and, therefore, the LER is also location-dependent. Given a location in such a pattern, it is possible to use the analytic expression of LER (Eq. 3.15) derived for a single line, i.e., compute the mean and variance of exposure along the development path reaching the location through a feature-by-feature convolution (Eq. 2.2 and Eq. 2.3), derive the mean and variance of the developing rate from the exposure and conversion formula (Eq. 2.4), and apply the analytic method for a single line, i.e., derive the LER from the distribution of developing rate. However, this approach requires the repetitive efforts of computing the exposure distribution at each location.

In this study, the LER expression for a single line is adjusted by considering the location dependency of exposure. In a large pattern, the exposure at a location may be decomposed into the "local" and "global" exposures [10] where the former refers to the exposure contribution from the forward-scattering of electrons and the latter to that from the back-scattering. The exposure distribution in the case of a single line is mainly of local exposure. Therefore, the difference in the exposure distribution between a single line and a large pattern comes from the global exposure which varies significantly with the location, as shown in Fig. 4.3. The adjustment of the analytic expression of LER for a single line is done according to the distribution of global exposure.

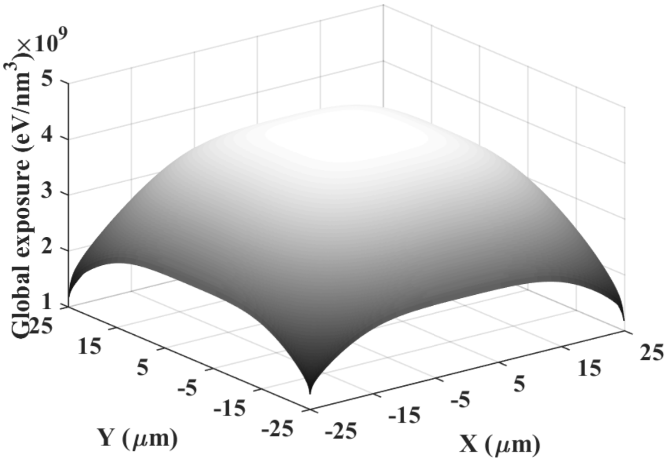


Figure 4.3: The global exposure distribution in a large-scale uniform pattern: 300 nm PMMA on Si, dose of 640 μC/cm^2, beam energy of 50 keV, and pattern size of 50 μm $\times$ 50 μm.

4.1.1 Modeling of Global Exposure

To model the global exposure, a PSF may be decomposed into two components [23], the local component, which has a relatively large magnitude and is sharp, describes the energy deposited by the forward-scattering of electrons, and the global component, which has a relatively low magnitude and is flat, describes that by the back-scattering (see Fig. 2.2). One common way to model PSF's is to employ a double-Gaussian function, which was proposed by Chang in 1975 [31], where one Gaussian function models the local component and the other models the global component. Therefore, the PSF at a layer (z) in the resist can be expressed as

$$\begin{aligned} psf(r,z) &= psf_F(r,z) + psf_B(r,z) \\ &= a_F(z)\exp\left(-\frac{r^2}{b_F(z)}\right) + a_B(z)\exp\left(-\frac{r^2}{b_B(z)}\right) \end{aligned} \tag{4.1}$$

where $r = \sqrt{x^2+y^2}$, $psf_F(r,z)$ is to model the local component with parameters of $a_F(z)$ and $b_F(z)$, and $psf_B(r,z)$ the global component with parameters of $a_B(z)$ and $b_B(z)$.

As in Eq. 2.1, the global exposure can be calculated by the convolution of the global component of a PSF, i.e., psf_B, and the dose distribution as

$$e_G(x,y,z) = \int_{-\frac{W}{2}}^{\frac{W}{2}} \int_{-\frac{W}{2}}^{\frac{W}{2}} D(x,y) \cdot psf_B(x-x', y-y', z)dx'dy' \tag{4.2}$$

where $D(x,y)$ depicts the dose distribution in a L/S pattern shown in Fig. 4.2, i.e., $D(x,y) = D_0$ when the point (x,y) is inside a line, and $D(x,y) = 0$ otherwise.

In the case of a large uniform pattern in Fig. 4.2, since the psf_B spatially varies slow (see Fig. 2.2), the global exposure may be computed without performing the feature-by-feature convolution. Instead, the entire pattern may be considered to be a single feature with the size of $W \times W$ where the effective dose is the actual dose scaled by the feature density d (e.g., $d = 0.5$ when the line-width and space are equal in a L/S pattern as shown in Fig. 4.2). From Eq. 4.1 and Eq. 4.2, the global exposure e_G can be computed by

$$e_G(x,y,z) = \int_{-\frac{W}{2}}^{\frac{W}{2}} \int_{-\frac{W}{2}}^{\frac{W}{2}} D_0 d \cdot psf_B(x-x', y-y', z)dx'dy'$$

$$= \int_{-\frac{W}{2}}^{\frac{W}{2}} \int_{-\frac{W}{2}}^{\frac{W}{2}} D_0 d \cdot a_B(z) \exp\left(-\frac{(x-x')^2+(y-y')^2}{b_B(z)}\right) dx'dy'. \quad (4.3)$$

Note that the integral of Eq. 4.3 does not lead to an explicit expression. Since the psf_B changes slowly over the space, the mean and variance of psf_B are approximated as follows (see Fig. 4.4):

$$\begin{aligned} m_{psf_B}(r) &= \begin{cases} M_B & \text{when} \quad r \leq r_0 \\ 0 & \text{when} \quad r > r_0 \end{cases}, \\ \sigma^2_{psf_B}(r) &= \begin{cases} {V_B}^2 & \text{when} \quad r \leq r_0 \\ 0 & \text{when} \quad r > r_0 \end{cases} \end{aligned} \quad (4.4)$$

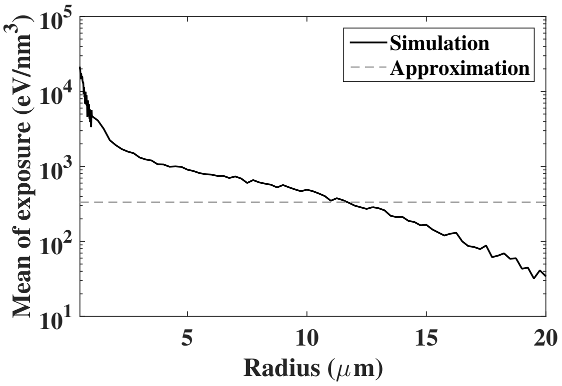
where M_B and ${V_B}^2$ are the average mean and variance of the global component of PSF within the PSF domain, respectively, and r_0 is the radius of the PSF domain. An example of this approximation is shown in Fig. 4.4 where r_0 is 20 μm.

Then, computing the global exposure at a point is equivalent to finding the overlapped area between the pattern and the domain of a PSF, which is a circle with the radius of r_0 centered at (x, y). Therefore, from Eq. 4.3 and Eq. 4.4, the mean of global exposure m_{e_G} can be derived as

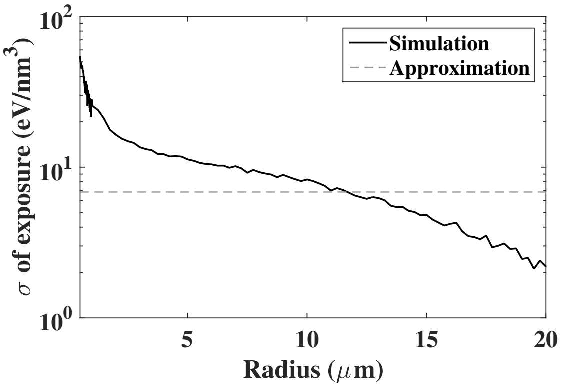
$$m_{e_G}(x,y) = \begin{cases} M_B \pi r_0^2 d & \text{when} \quad r < \frac{W}{2} - r_0 \\ \left(\begin{array}{c} \frac{1}{2}\sin\theta_1 \cdot \left(\frac{W}{2} - |x|\right) r_0 + \frac{1}{2}\sin\theta_2 \cdot \left(\frac{W}{2} - |y|\right) r_0 + \\ \left(\frac{W}{2} - |x|\right)\left(\frac{W}{2} - |y|\right) + \frac{1}{2}\left(\frac{3\pi}{2} - \theta_1 - \theta_2\right) \end{array} \right) M_B d & \text{when} \quad r \geq \frac{W}{2} - r_0 \end{cases} \quad (4.5)$$

Similarly, the fluctuation of global exposure $\sigma^2_{e_G}$, which affects the LER, can be derived as

$$\sigma^2_{e_G}(x,y) = \begin{cases} {V_B}^2 \pi r_0^2 d & \text{when} \quad r < \frac{W}{2} - r_0 \\ \left(\begin{array}{c} \frac{1}{2}\sin\theta_1 \cdot \left(\frac{W}{2} - |x|\right) r_0 + \frac{1}{2}\sin\theta_2 \cdot \left(\frac{W}{2} - |y|\right) r_0 + \\ \left(\frac{W}{2} - |x|\right)\left(\frac{W}{2} - |y|\right) + \frac{1}{2}\left(\frac{3\pi}{2} - \theta_1 - \theta_2\right) \end{array} \right) {V_B}^2 d & \text{when} \quad r \geq \frac{W}{2} - r_0 \end{cases} \quad (4.6)$$



(a)



(b)

Figure 4.4: The approximated (a) mean, M_B, and (b) standard deviation, V_B, of the global component of a PSF obtained from the Monte Carlo simulation: 300 nm PMMA on Si, dose of 640 μC/cm^2, and beam energy of 50 keV. M_B and V_B are indicated by the dashed lines.

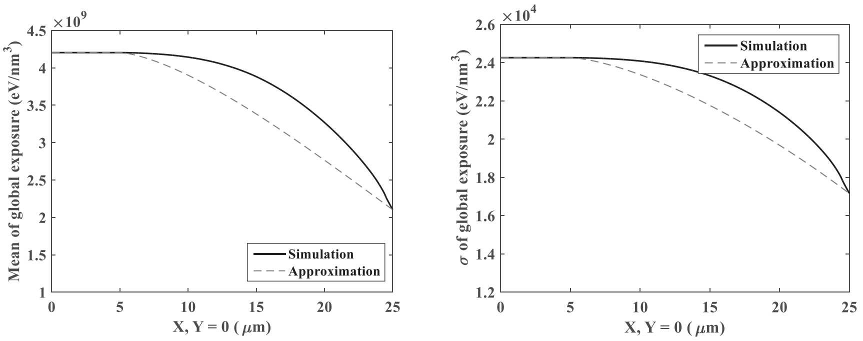
where

$$\begin{aligned} \theta_1 &= \cos^{-1}\left(\frac{\frac{W}{2}-|x|}{r_0}\right), \\ \theta_2 &= \cos^{-1}\left(\frac{\frac{W}{2}-|y|}{r_0}\right). \end{aligned} \tag{4.7}$$

The mean and variance of global exposure generated by the Monte Carlo simulation are compared with those by the approximate models (Eq. 4.5 and Eq. 4.6) in Fig. 4.5 where the pattern size ($W = 50$ μm) is larger than the PSF domain ($r_0 = 20$ μm). As shown in the figure, the mean and variance of global exposure derived by the approximate models are closely matched with those from the simulation at the critical locations where $(X = Y = 0)$,

$(X = 25, Y = 0)$, and $(X = Y = 25)$ correspond to the center, edge, and corner regions in Fig. 4.2.

Along X axis:



Along X=Y diagonal:

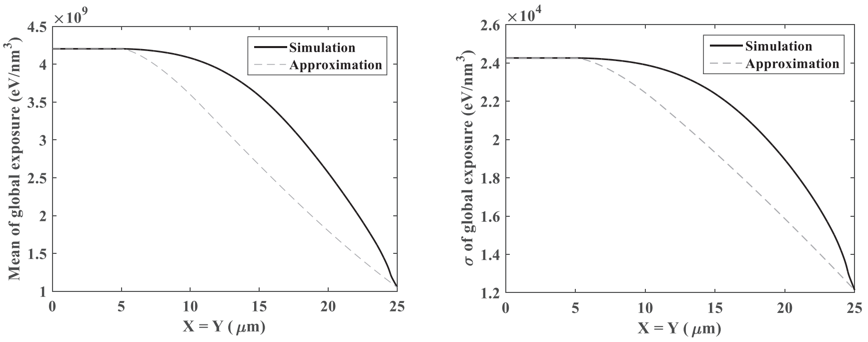


Figure 4.5: The analytic estimates of the mean and variance of global exposure are compared with the corresponding simulation results: 300 nm PMMA on Si, dose of 640 μC/cm^2 and beam energy of 50 keV.

4.1.2 Relationship Between Global Exposure and LER

The LER at a location may be modeled by two parts: the contribution from the local exposure, which is independent of the location, and the contribution from the global exposure. The exposure distribution in the case of a single line is mainly of the local exposure and, therefore, the LER derived for a single line (Eq. 3.15) may approximate the LER contribution from the local exposure in a large pattern.

Given a location at which the LER is to be estimated, the location is referred to by (x, y) for computing the global exposure as in Section 4.1.1 (see Fig. 4.2) and by x_l for expressing the edge location in the local coordinate (see Fig. 4.6), to be distinguished from the global variable x. The x_l represents the relative location of edge with respect to the (right) target edge location of a feature. That is, $x_l = 0$, $x_l < 0$ and $x_l > 0$ correspond to the cases where the actual edge is on the target location, inside the feature, and outside the feature, respectively. In the case of a single line, x_l is the same as x in Eq. 3.15. It should be mentioned that x_l is only an auxiliary variable, not independent of x.

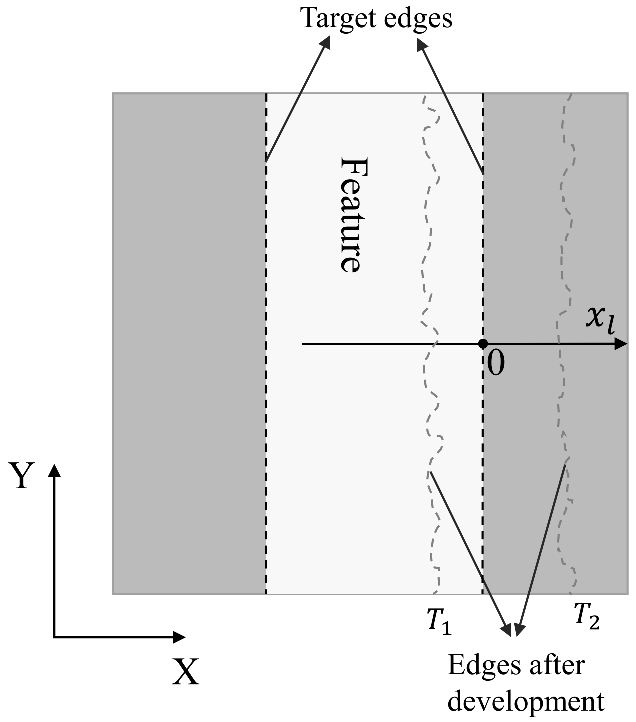


Figure 4.6: x_l is defined with respect to the target edge of a feature. After the development process, the boundary (edge) of remaining resist profile can be inside ($x_l < 0$) or outside ($x_l > 0$) the feature. For example, the two dashed curves represent the two possible boundaries for two different developing times, T_1 and T_2 where $T_1 < T_2$.

Let $\Delta\sigma(x_l|T,(x,y))$ denote any change of the LER due to the global exposure, compared to the LER for a single line. The fluctuation in the global exposure is relatively small compared to that in the (total) exposure and does not vary substantially within a small region. Also, the LER tends to increase as the exposure fluctuation increases. Therefore, $\Delta\sigma(x_l|T,(x,y))$ may be approximated to be linearly proportional to the added exposure fluctuation $\sigma_{e_G}(x,y)$.

$$\Delta\sigma(x_l|T,(x,y)) = A \cdot \sigma_{e_G}(x,y) \tag{4.8}$$

where A is is a positive proportional constant (change rate).

The proportional constant A may be found using the LER for a single line, denoted by $\sigma_{single}(x_l|T)$, where $\sigma_{e_G} = 0$ and the LER at the center in a large pattern, $\sigma_{center}(x_l|T)$, where $\sigma_{e_G}(0,0) = V_B r_0 \sqrt{\pi d}$ (see Eq. 4.6). The LER at the center can be derived by the analytic method for a single line, but with the distribution of developing rate at the center region. Then, A is expressed as

$$A = \frac{\sigma_{center}(x_l|T) - \sigma_{single}(x_l|T)}{V_B r_0 \sqrt{\pi d}}. \tag{4.9}$$

And the LER at (x, y) can be derived as

$$\begin{aligned} \sigma(x_l|T,(x,y)) &= \Delta\sigma(x_l|T,(x,y)) + \sigma_{single}(x_l|T) \\ &= \frac{(\sigma_{center}(x_l|T) - \sigma_{single}(x_l|T))\sigma_{e_G}(x,y)}{V_B r_0 \sqrt{\pi d}} + \sigma_{single}(x_l|T). \end{aligned} \tag{4.10}$$

The exposure contrast over a feature edge also affects the LER. A larger exposure contrast tends to lead to a smaller LER [30], since the room for variation among neighboring edge locations becomes smaller when the exposure decreases faster (over the edge). Therefore, this inverse relationship between the LER and exposure contrast is approximated to be linear. Then, the LER expression may be further adjusted according to the location-dependent exposure contrast. The exposure contrast $Contrast(x, y)$ is quantified as the exposure difference between the center of a feature and the middle point between features. An example of the location dependency of exposure contrast is provided in Fig. 4.7.

Finally, the LER is expressed as follows:

$$\begin{aligned} &\sigma(x_l|T,(x,y)) \\ &= \left(\frac{(\sigma_{center}(x_l|T) - \sigma_{single}(x_l|T))\sigma_{e_G}(x,y)}{V_B r_0 \sqrt{\pi d}} + \sigma_{single}(x_l|T) \right) \frac{Contrast_{single}}{Contrast(x,y)} \end{aligned} \tag{4.11}$$

where $Contrast_{single}$ is the exposure contrast in the case of a single line.

As can be seen from Eq. 4.11, the LER at a location (x, y) in a large pattern is determined by the global-exposure fluctuation and exposure contrast.

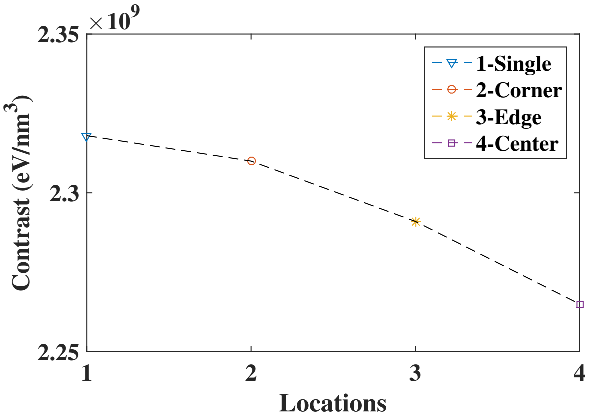


Figure 4.7: The location dependency of exposure contrast in a large pattern with 300 nm PMMA on Si, beam energy of 50 keV, dose of 640 μC/cm^2, and pattern size of 50 μm $\times$ 50 μm where both the line-width and space are 100 nm. The exposure contrast for the corresponding single-line pattern is also provided.

Given the LER expressions at the critical locations (Eq. 4.11), the LER at another location is computed through an interpolation of the LER's at the critical locations. Since the exposure distribution does not linearly vary with the location, the global exposure distribution (Fig. 4.3) is used as a reference for the interpolation. For example, if the value of global exposure at a location is between those at the center and edge, the LER at that location is derived according to the ratio of global exposure at those three locations, i.e., the ratio of LER at the three locations is same as that of global exposure.

4.2 Minimization of LER

As the exposure level varies with the location in a large pattern (Fig. 4.3), after the resist development process, the CD at a location can be significantly different from those at other locations, and so is the LER. Therefore, to minimize the LER and CD error in a large unform pattern, each location (x, y) requires a different dose level denoted by $D_{opt}(x, y)$. A two-step procedure is employed to derive the optimal dose distribution for a large uniform pattern.

4.2.1 Deconvolution Surface

Since given a developing time, the x_l at a location can be significantly different from others, a deconvolution surface is used to compensate the differences of exposure level among the locations in the first step (Fig. 4.3). The deconvolution surface is derived as [5]

$$D_s(x, y) = e_t \otimes^{-1} psf(x, y) \tag{4.12}$$

where e_t is the target (average) exposure within the feature, i.e., the exposure needed to reach the target location ($x_l = 0$) after the resist development, which can be calculated from the D_t for a single-line pattern (see Sec. 3.2) and $psf(x, y, z)$ through Eq. 2.1.

An example of the deconvolution surface is shown in Fig. 4.8.

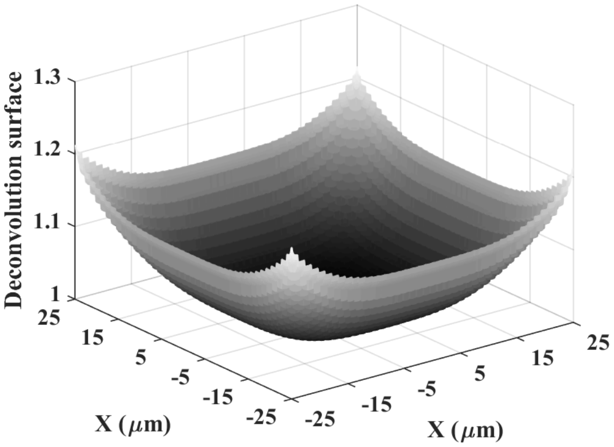


Figure 4.8: An example of the deconvolution surface generated by a PSF obtained from the Monte Carlo simulation: 300 nm PMMA on Si, dose of 640 μC/cm^2, and beam energy of 50 keV. And the pattern size is 50 μm × 50 μm.

According to Eq. 2.1, with the dose distribution of $D_s(x, y)$, all the locations in a large pattern have the same average exposure level of e_t inside the line. Since x_l is mostly determined by the exposure inside the line, especially when $x_l \leq 0$, the edge location x_l at all locations should be approximately the same, i.e., $x_l \approx 0$.

4.2.2 Optimal Dose

As explained in Sec. 3.2, since the LER and CD error may not be minimized simultaneously, the same cost function is employed, i.e.,

$$C(x_l) = |x_l - 0| + 3 \cdot LER. \tag{4.13}$$

To avoid the the feature-by-feature optimization, the optimal doses for the critical locations are derived first, and the optimal dose distribution $D_{opt}(x, y)$ is obtained through an interpolation using the optimal doses at the critical locations. For each critical location, since the LER is expressed as a function of x_l (Eq. 4.11), the same method described in Sec. 3.2 is used to derive the optimal x_l and, eventually, the optimal dose D_{opt}. Define ΔD_{opt} as $D_{opt} - D_s$, and given the ΔD_{opt}'s at the critical locations, the $\Delta D_{opt}(x, y)$ for a large pattern can be approximated through a 2-D interpolation as

$$\Delta D_{opt}(x, y) \approx a_0 + a_1 x + a_2 y + a_3 xy \tag{4.14}$$

where the coefficients are found by solving the following equation:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & \frac{W}{2} & 0 \\ 1 & -\frac{W}{2} & 0 & 0 \\ 1 & -\frac{W}{2} & \frac{W}{2} & -\frac{W^2}{4} \end{pmatrix} \cdot \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} \Delta D_{opt}(0, 0) \\ \Delta D_{opt}\left(0, \frac{W}{2}\right) \\ \Delta D_{opt}\left(-\frac{W}{2}, 0\right) \\ \Delta D_{opt}\left(-\frac{W}{2}, \frac{W}{2}\right) \end{pmatrix}. \tag{4.15}$$

With respect to Fig. 4.2, $\Delta D_{opt}(0, 0)$, $\Delta D_{opt}\left(-\frac{W}{2}, \frac{W}{2}\right)$, and $\Delta D_{opt}\left(0, \frac{W}{2}\right)$ and $\Delta D_{opt}\left(-\frac{W}{2}, 0\right)$ are the ΔD_{opt} for the center, corner, and two edge regions, respectively. Therefore, the optimal dose distribution $D_{opt}(x, y)$ can be expressed as

$$\begin{aligned} D_{opt}(x, y) &= D_s(x, y) + \begin{pmatrix} 1 & x & y & xy \end{pmatrix} \cdot \\ & \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & \frac{W}{2} & 0 \\ 1 & -\frac{W}{2} & 0 & 0 \\ 1 & -\frac{W}{2} & \frac{W}{2} & -\frac{W^2}{4} \end{pmatrix}^{-1} \cdot \begin{pmatrix} \Delta D_{opt}(0, 0) \\ \Delta D_{opt}\left(0, \frac{W}{2}\right) \\ \Delta D_{opt}\left(-\frac{W}{2}, 0\right) \\ \Delta D_{opt}\left(-\frac{W}{2}, \frac{W}{2}\right) \end{pmatrix}. \end{aligned} \tag{4.16}$$

4.3 Results and Discussion

In order to verify the accuracy of the proposed analytic method, its results are compared with those obtained through the simulation. A typical substrate system is employed where a resist layer of PMMA with a certain thickness is on the top of the Si substrate. A large L/S pattern with a size of 50 μm $\times$ 50 μm ($W \times W$) is exposed with a uniform dose.

While the analytic method can be applied to any layer of the resist, the results for the bottom layer are provided since the LER is usually largest at the bottom layer. The LER estimated by the analytic method is compared with that obtained through the simulation for varying edge location x_l where $x_l = 0$ corresponds to the target edge location (see Fig. 4.6). The edge location is controlled by changing the developing time with the same exposure (developing rate) distribution (see Fig. 4.6).

In Fig. 4.9, the results for a L/S pattern with the line-width of 100 nm are provided at the three critical locations, the corner, edge and center. The beam energy is 50 keV, the resist thickness is 300 nm, and the dose level is 64 μC/cm^2. Note that the developing time is different for the same edge location in a different region since the exposure distribution varies with location. It can be seen that the LER's estimated by the analytic method are well matched with those by the simulation. In Fig. 4.10, the results for a L/S pattern with a narrower line-width of 50 nm and a thicker resist (500 nm) are provided where the beam energy is 50 keV, and the dose level is 64 μC/cm^2.

One observation to be made is that the behavior of LER varying with the edge location varied (i.e., the shape of LER curve) does not change substantially with the location within a pattern. This is due to the fact that the spatial distribution of exposure cross a line feature has a similar shape that is independent of the location though the exposure level varies with the location. Note that the shape of exposure distribution mainly depends on the local exposure.

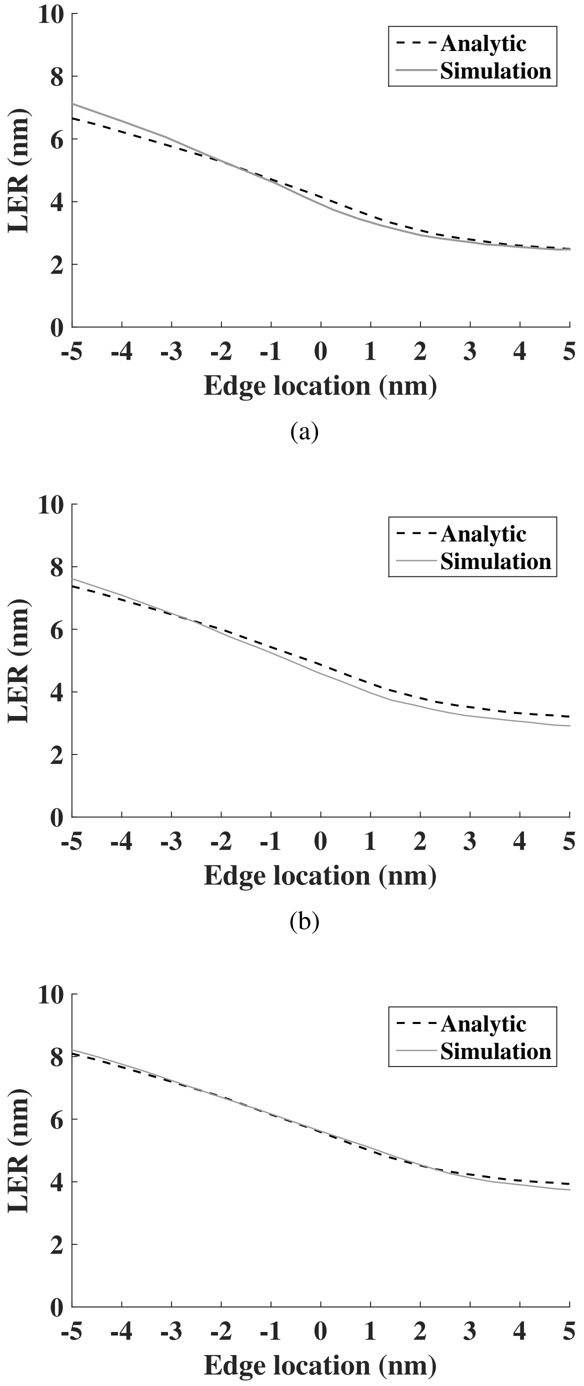
The exposure fluctuation at the center is larger than that at the edge or corner. Therefore, at the same edge location x_l, the LER at the center is also larger than that at the edge or corner, as a larger exposure fluctuation leads to a larger LER in these cases. Comparing Fig. 4.9 with Fig. 4.10, the LER's in Fig. 4.10 are significantly larger than those in Fig. 4.9. It is because the exposure for 500 nm resist usually has a smaller contrast than the exposure for 300 nm resist with the same beam energy of 50 keV, and the LER is inversely proportional to the contrast. Also, since the development process is cumulative from the top surface of resist to the bottom layer (see Fig. 2.5), the fluctuation $\sigma(T|s)$ for the resist thickness of 500 nm can be larger than that for the resist thickness of 300 nm, which also leads to a larger LER.

The LER's at other locations than the three critical locations are computed through an interpolation of the LER's analytically estimated at the critical locations. In Fig. 4.12, the LER's obtained through the interpolation are compared with the simulation results at three test locations shown in Fig. 4.11. The results in the figure verify that the LER's estimated at the three critical locations are sufficient to accurately compute the LER at any other location in a large uniform pattern through the interpolation.

The results of analytic minimization of the LER and CD error for a large uniform pattern are compared with those obtained from the simulation where the developing time is fixed. As can be seen in Fig. 4.13 where the beam energy is 50 keV, the resist thickness is 300 nm, for the critical locations, the optimal doses derived by the analytic method of a single line are closely matched with those by the simulation. And the optimal dose distribution of a large pattern derived by the interpolation also minimizes the LER and CD error at the test locations.

4.4 Summary

In this chapter, an analytic method of estimating the LER in a large-scale uniform pattern is described based on the LER expression for a single line. The analytic expression of LER derived for a single-line pattern is adjusted for the critical locations according to the stochastic information on the global exposure at those locations. Then, the LER at another location is derived through an interpolation of the LER's at the critical locations. Due to the global exposure, the CD varies significantly with the location in a large pattern. Therefore, a two-step minimization procedure is employed to derive the optimal dose distribution for a large pattern, which can minimize the CD error and LER at all locations. It has been shown that the results obtained by the analytic method are closely matched with those from the simulation.

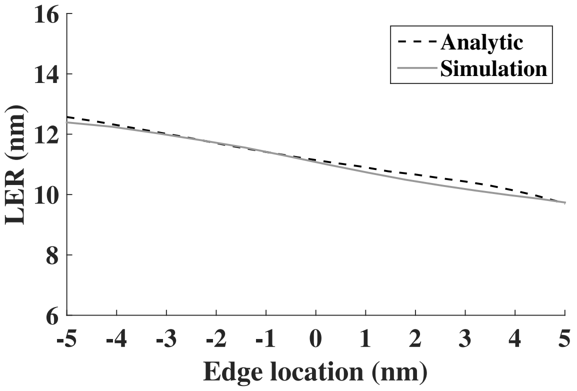


(a)

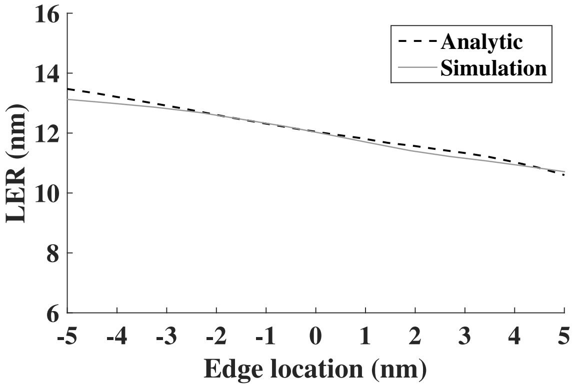
(b)

(c)

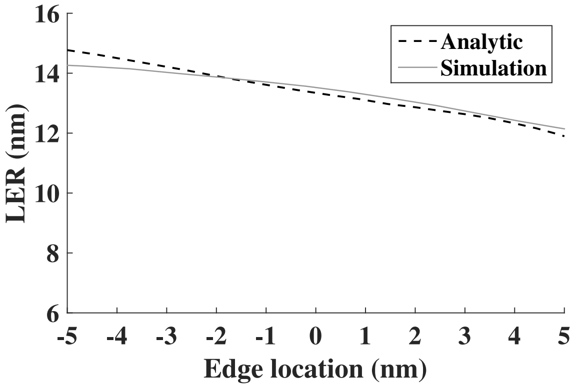
Figure 4.9: The LER estimated by the analytic method and simulation at the (a) corner, (b) edge, and (c) center: line-width of 100 nm, 300 nm PMMA on Si, beam energy of 50 keV, and dose of 64 μC/cm^2.



(a)



(b)



(c)

Figure 4.10: The LER estimated by the analytic method and simulation at the (a) corner, (b) edge, and (c) center: line-width of 50 nm, 500 nm PMMA on Si, beam energy of 50 keV, and dose of 64 μC/cm^2.

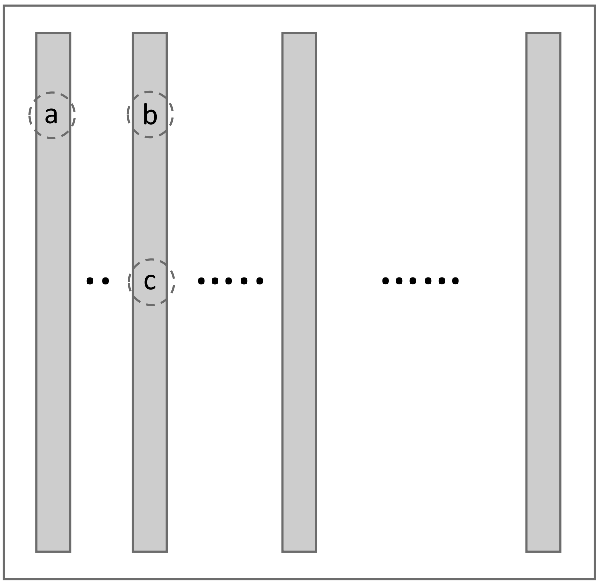
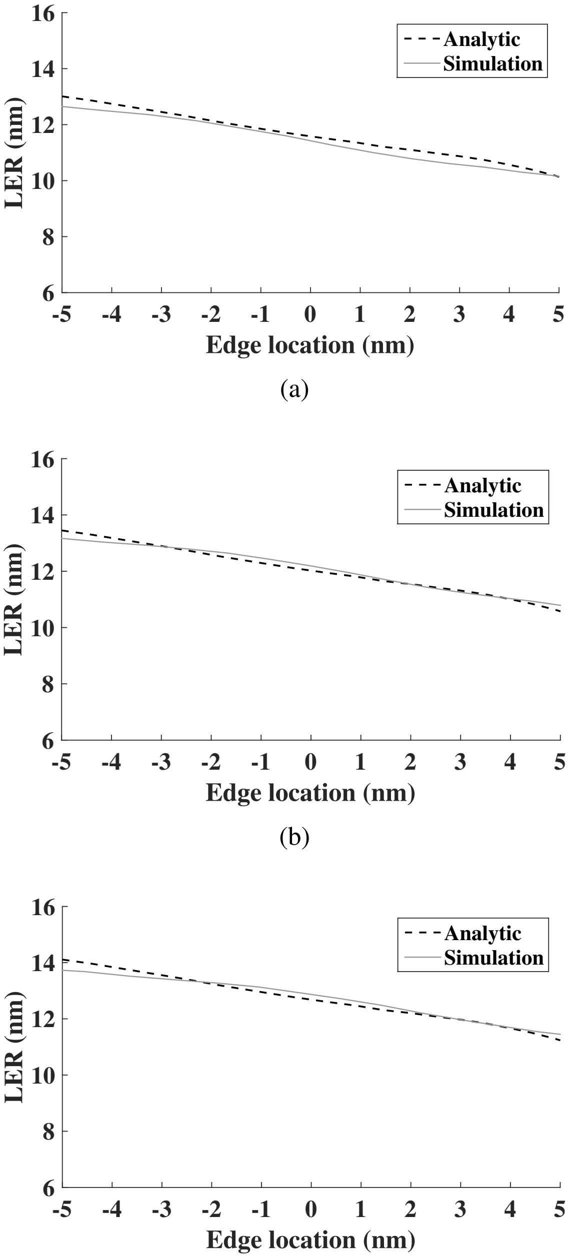


Figure 4.11: Three test regions (a, b, and c) marked by circles in a large uniform pattern.



(a)

(b)

(c)

Figure 4.12: The LER estimated by the analytic method and simulation at at the test locations (in Fig. 10): line-width of 50 nm, 500 nm PMMA on Si, beam energy of 50 keV, and dose of 64 μC/cm^2.

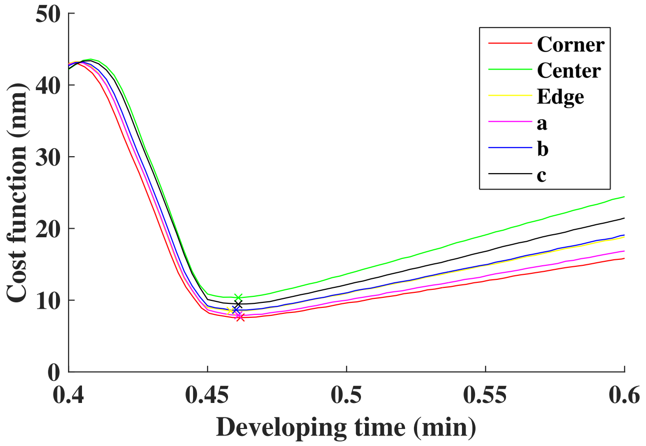


Figure 4.13: The minimization results (minimization of the cost function with respect to the dose) by the analytic method and simulation: 300 nm PMMA on Si, beam energy of 50 keV.

Chapter 5

Modeling Lithographic Process

In the previous chapters, the analytic method is verified using the results of the simulation described in Chapter 2. It is not uncommon that the results derived by the simulation are substantially different from the actual experimental results, since the simulation may not capture completely all the characteristics of e-beam lithographic process. Therefore, the experimental results can be used to verify the analytic method. However, unlike the simulation, the stochastic parameters, i.e., the mean and variance of exposure distribution in the resist, etc., required by the analytic method cannot be directly measured from the experimental results.

Therefore, in this chapter, a method to extract the stochastic information from the experimental results, i.e., SEM images, is described. In the first step, the line spread function (LSF) and conversion formula are modeled by matching the CD's measured from the SEM images, i.e., through iterations, the LSF and conversion formula are adjusted such that the difference between the modeled (through simulation) and measured (from SEM images) CD's is minimized. In the second step, the exposure fluctuation is modeled by matching the LER through iterations. Since the proposed method utilizes the experimental results of SEM images, its results are likely to be realistic. Another important aspect of the proposed method is that the LSF, conversion formula, and exposure fluctuation are modeled considering their interaction rather than individually. Therefore, a minor inaccuracy in one of them may be compensated by the others such that the modeled CD and LER can achieve a high accuracy compared with those measured from SEM images.

The rest of the chapter is organized as follows. The preprocessing of SEM images and the edge detection are described in Sec. 5.1. The modeling of e-beam lithographic process is described in Sec. 5.2. The results of modeling SEM images are discussed in Sec. 5.3, followed by a summary in Sec. 5.4.

5.1 Analysis of SEM Images

In the experiment, a typical substrate system is employed where a resist layer with a certain thickness is on top of the substrate. A pattern of multiple long lines with the fixed line-width and space is exposed with different levels of a uniform dose, while all the other e-beam lithographic parameters, i.e., beam energy, developing time, etc., are fixed. After the resist development process, SEM images are taken in a top-down view of the remaining resist profile. An example of SEM image is shown in Fig. 5.1, which is to be used in illustrating the steps of analyzing SEM images. Using image processing techniques, the feature boundaries (edges) are detected to measure the CD (line-width) and LER.

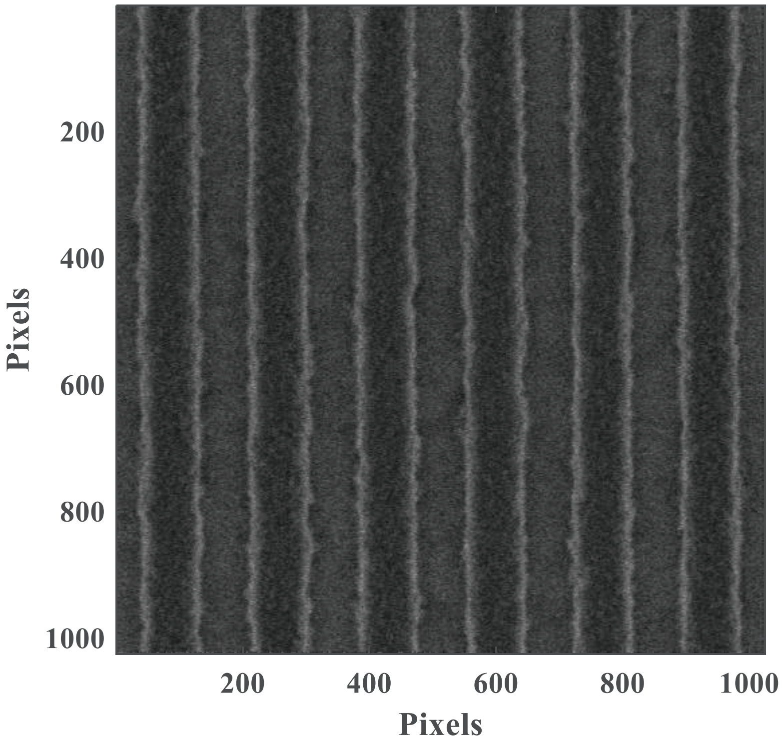
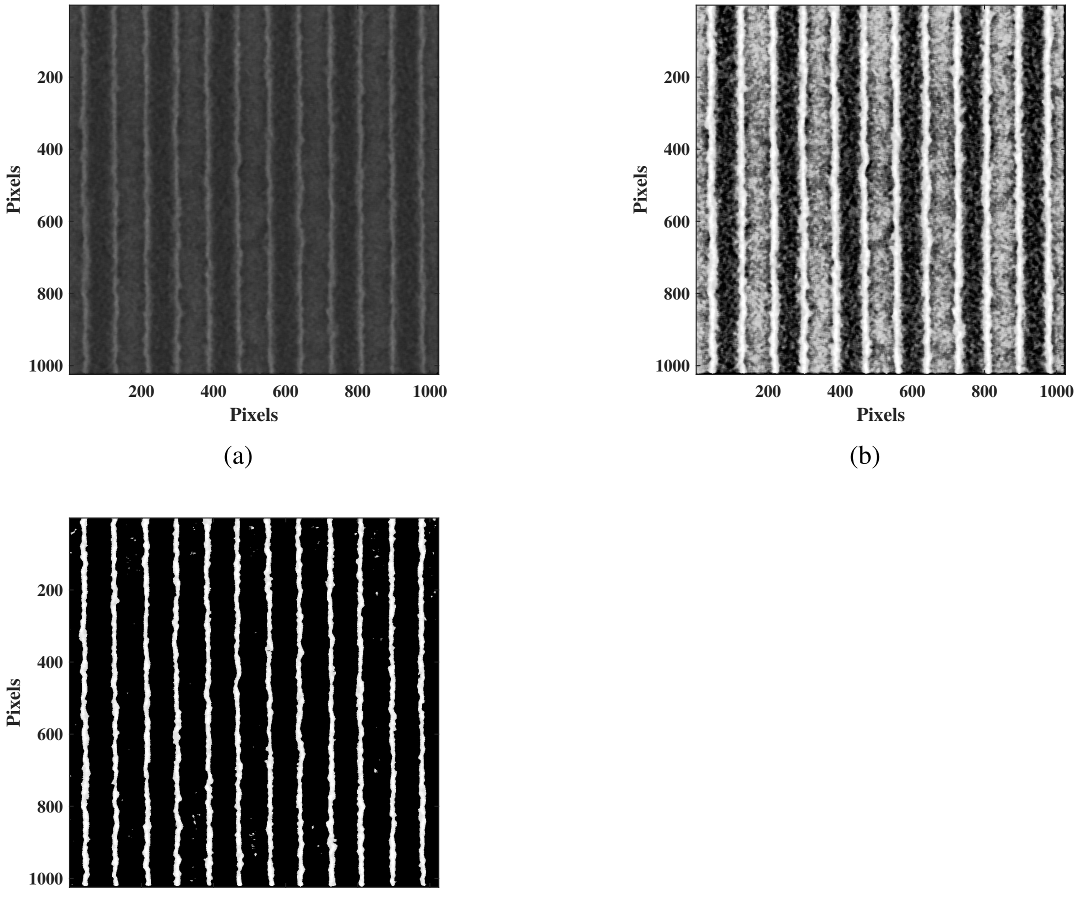


Figure 5.1: An example of SEM image taken in a top-down view of the remaining resist profile for a large L/S pattern where the line-width and space are 120 nm. The pixel interval is 1.4 nm.

5.1.1 Preprocessing

SEM images typically include a significant level of "salt and pepper" noise (see Fig. 5.1) and their quality, e.g., brightness or contrast, may vary from image to image or spatially within an image. Since the proposed method utilizes the information (the CD and LER) extracted from SEM images, it is important to reduce the noise level and enhance the image quality.



(a) (b)

(c)

Figure 5.2: An illustration of preprocessing the SEM image in Fig. 5.1: (a) noise reduction, (b) contrast enhancement, and (c) background removal.

A spatial-averaging filter such as a Gaussian filter reduces the noise level, but at the same time tends to smooth out the image detail such as the roughness of feature boundary. Also, such a filter is not very effective on the salt-and-pepper noise. On the other hand, the median filter is effective in reducing the salt-and-pepper noise without destroying the image detail significantly. Therefore, a median filter is employed in this study. The larger the size of median filter is, the more image detail can be lost. The smallest size of the median filter that removes most of the noise is found to be 9×9 pixels (12.6 nm $\times$ 12.6 nm). The SEM image (Fig. 5.1) after the noise reduction is shown in Fig. 5.2(a).

Since the boundary detection is based on the change of brightness, an SEM image with a low contrast is prone to a detection error. The contrast of such an image is enhanced by the histogram equalization for the better detection. The result of contrast enhancement is shown in Fig. 5.2(b). The contrast enhancement is carried out after the noise reduction not to increase the noise level.

5.1.2 Detection of Feature Boundaries

In SEM images, boundary (edge) regions are brighter than the background. To facilitate detecting feature boundaries, the background is removed by using a threshold, which is set to be the half of the average brightness in the edge regions assuming most of the background is set to zero after the contrast enhancement, as shown in Fig. 5.2(c).

However, as the noise in SEM images is not completely removed, using the threshold, i.e., half of the average brightness in the edge regions, may not be sufficient to remove all the background, e.g., some small regions may have a brightness higher than the threshold and survive the thresholding (see Fig. 5.3). Therefore, a window centered at each peak of brightness is employed to define the edge region and any area outside the edge regions is set to be zero.

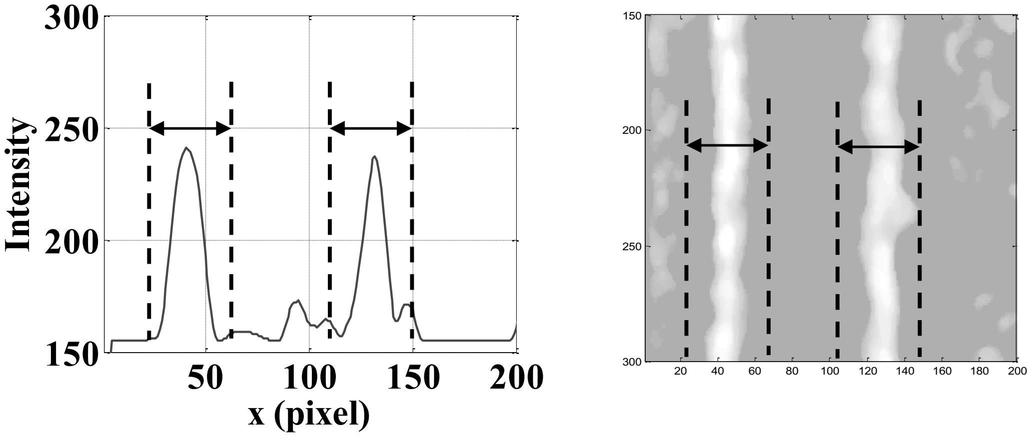


Figure 5.3: Edge region defined by a window centered at the peak brightness.

A 3×3 differential edge detector is employed (laterally) in edge regions, to obtain the gradient of the preprocessed SEM images. The maximum positive and negative gradients are searched to find the inner and outer edge locations (see Fig. 5.4(b)). The feature boundaries are normally continuous in space. Isolated edges such as a single pixel of edge are most probably due to the noise and thus are removed. In Fig. 5.4, the detected edges are overlaid with the original image where a high-fidelity result can be observed.

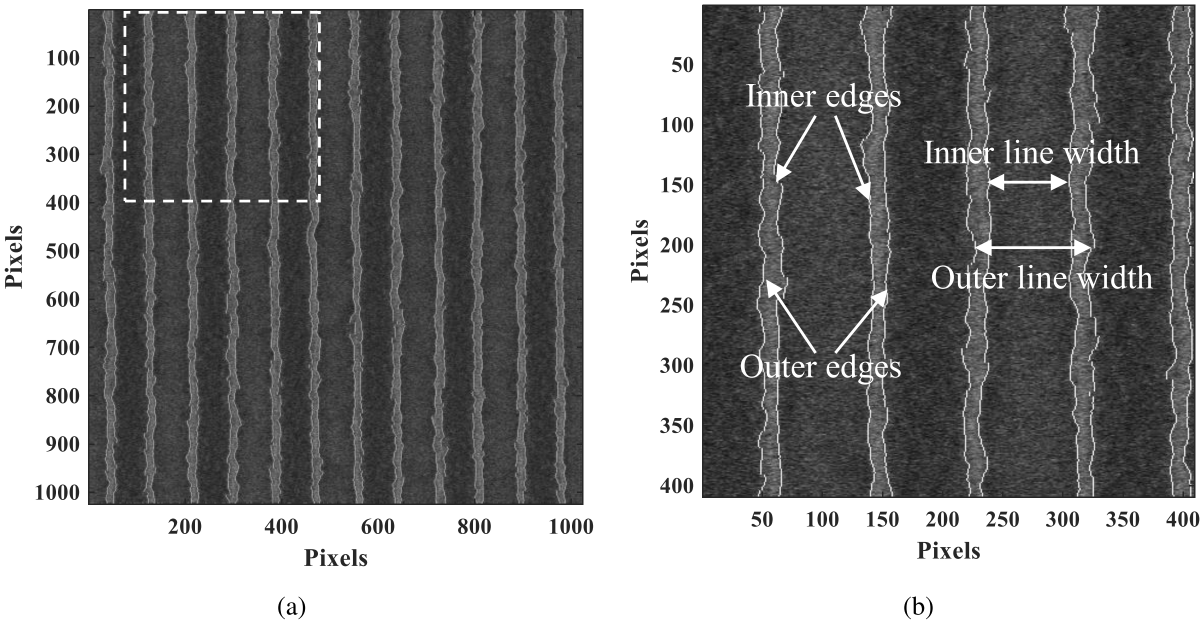
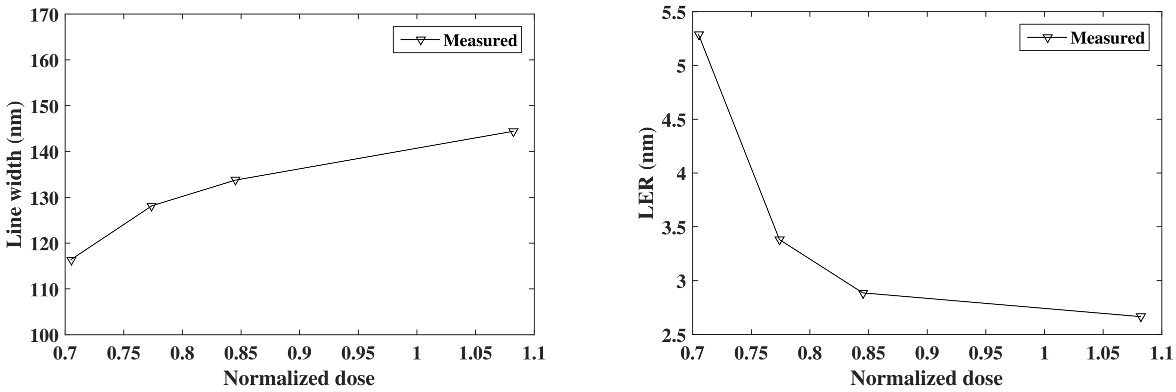


Figure 5.4: (a) Detected edges are overlaid with the original SEM image and (b) a zoomed-in region (the white box shown in (a)) where inner and outer edges are illustrated.

The outer and inner edges shown in Fig. 5.4(b) may correspond to the edges at the top and bottom layers of resist, respectively, in the case of an overcut resist profile. To utilize as much information extracted from SEM images as possible, both the inner and outer edges are considered, i.e., the CD is evaluated as the average of the inner and outer line-widths, and the LER is evaluated as the standard deviation of the inner and outer edge locations. The CD's and LER's measured from the SEM images with 4 different normalized dose levels (from 0.705 to 1.082) are plotted in Fig. 5.5. The proposed method does not require the absolute dose levels to be known, and the normalized dose is used only to distinguish different doses, i.e., relative dose levels, and the normalized dose of 1.0 does not have any special meaning.



(a) (b)

Figure 5.5: (a) The average line-widths and (b) the LER's measured from a set of SEM images with different normalized dose levels.

The detected edge locations are indicated by the pixel index, i.e., (i, j), of SEM image, e.g., for a 1024 by 1024 image, both i and j are from 1 to 1024. The physical edge location (x_i, y_j) are derived by (i, j) multiplied by the pixel size, e.g., if the pixel size is 1.4 nm, $(x_i, y_i) = (1.4i, 1.4j)$. To measure the LER accurately, the tilt angel η of the lines in the SEM images is estimated using the image moments [35]. The corrected edge locations are obtained by rotating the detected edge locations by the angle η as

$$\begin{aligned} x'_i &= x_i \cos\eta - y_j \sin\eta, \\ y'_j &= x_i \sin\eta + y_j \cos\eta \end{aligned} \tag{5.1}$$

where (x'_i, y'_j) is the corrected edge location of (x_i, y_j).

5.2 Modeling of E-beam Lithographic Process

In this section, the procedures of modeling the e-beam lithographic process, i.e., estimating the LSF, conversion formula, and exposure fluctuation based on the CD and LER measured from SEM images, are described. In the proposed method, the LSF, conversion formula, and exposure fluctuation are modeled by considering their interaction rather than individually. This approach allows a modeling error in one to be compensated in the others as long as the modeled CD and LER closely match with the measured ones. Also, the approach makes it unnecessary to know certain experimental details such as the absolute dose and developing time. For example, through the modeling process, the LSF and conversion formula adapt themselves to an assumed developing time.

5.2.1 Models

In the SEM images (see Fig. 5.1), only the line-space patterns are considered and, therefore, the exposure distribution in the resist can be computed using the LSF, denoted by $lsf(x, z)$, which describes the exposure distribution at a layer (z) when an infinitely long line with the single-pixel width is exposed. The 3-D spatial distribution of exposure $e(x, y, z)$ in the resist can be computed by the following convolution:

$$e(x, y, z) = \int D(x - x', y) lsf(x', z) dx'. \tag{5.2}$$

The conversion formula may be modeled by a Gaussian function as in Eq. 2.4 where only the left portion of the curve before its peak is used, i.e.,

$$R(x, y, z) = k_1 \cdot \exp\left(-\frac{(e(x, y, z) - k_2)^2}{k_3}\right) + k_4. \tag{5.3}$$

In Eq. 5.3, $k_1 + k_4$ is the maximum developing rate, k_2 is the lowest exposure leading to the maximum developing rate, k_3 is the rate contrast with respect to the exposure, and k_4 is the minimum developing rate for a non-zero exposure. Through the modeling process, the parameters $\{k_i\}$ are determined.

In the modeling process, given a distribution of developing rate, $R(x, y, z)$, the path-based simulation method [22] is used to derive the remaining resist profile from which the CD and LER are calculated and compared with the measured ones from SEM images.

5.2.2 Matching the Line-Width

Since the average line-width is most affected by the average developing-rate distribution which can be calculated from the LSF and conversion formula, the LSF and conversion formula can be modeled first by matching the average line-widths. To reduce the computation required by the modeling process, a two-step procedure is employed.

Without the Exposure Fluctuation

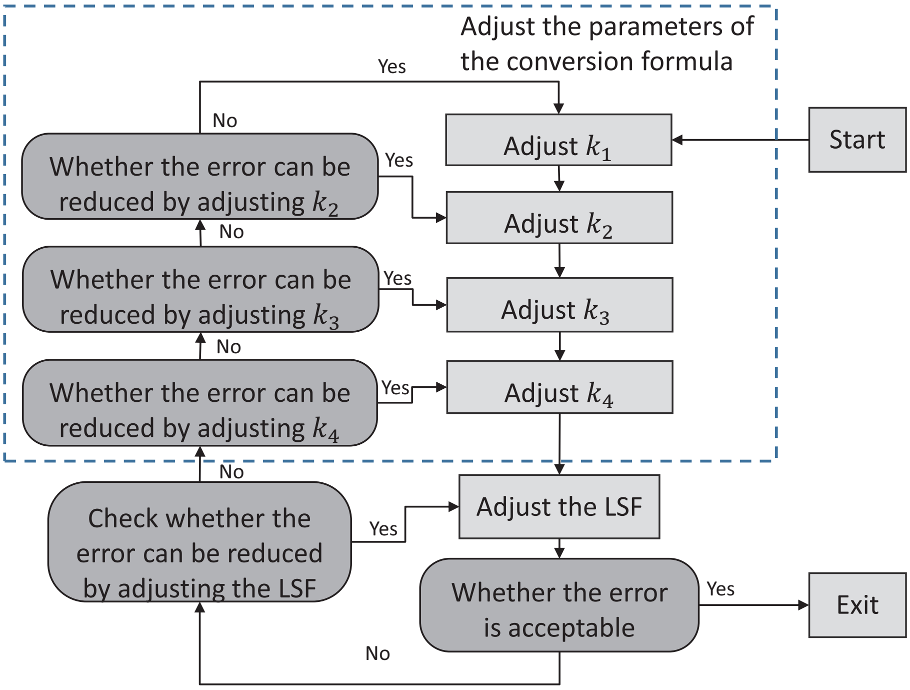


Figure 5.6: The flowchart of the iterative procedure for determining the conversion formula (and the LSF).

In the first step, without considering the effects of exposure fluctuation, an iterative procedure is employed, i.e., the LSF and conversion formula are adjusted through iterations such that the total difference (error) between the modeled and measured average line-widths is minimized. The total difference is the sum of the differences for all SEM images. The iterative procedure is described below and also in Fig. 5.6. Note that since the exposure fluctuation is not included in this step, only the 2-D (the X-Z plane) development simulation is needed, which makes practical the optimization of all the points of LSF along with the parameters in the conversion formula.

Step 1: Adjust the parameter k_1 in the conversion formula.

Step 2: Adjust the parameter k_2 in the conversion formula.

Step 3: Adjust the parameter k_3 in the conversion formula.

Step 4: Adjust the parameter k_4 in the conversion formula.

Step 5: Adjust the LSF.

Step 6: Check whether the total error is acceptable (minimized). If yes, stop the iteration. Otherwise go to step 7.

Step 7: Check whether the total error can be reduced by further adjusting the LSF. If yes, go to step 5. Otherwise, go to step 8.

Step 8: Check whether the total error can be reduced by further adjusting k_4. If yes, go to step 4. Otherwise, go to step 9.

Step 9: Check whether the total error can be reduced by further adjusting k_3. If yes, go to step 3. Otherwise, go to step 10.

Step 10: Check whether the total error can be reduced by further adjusting k_2. If yes, go to step 2. Otherwise, go to step 1.

Since both the LSF and conversion formula affect the line-width, after the adjustment of one parameter (k_i) in the conversion formula, each point in the LSF needs to be adjusted to derive the current "optimal" LSF, i.e., given a set of parameters ($\{k_i\}$) in the conversion formula, there only exists one optimal LSF that minimizes the total error. The flowchart of adjusting the LSF is shown in Fig. 5.7.

The overall procedure (see Fig. 5.6) terminates when the total error is acceptable (minimized), which indicates that both the LSF and conversion formula are optimized. To make the results realistic, in adjusting the conversion formula, the following constraints are imposed: the k_i's are positive and the conversion formula is a monotonically increasing function. Also, the following constraints are imposed on adjusting the LSF (see Fig. 5.7):

- The LSF is adjusted point-wise from the center to the outside iteratively until the local optimal value is found for each point.

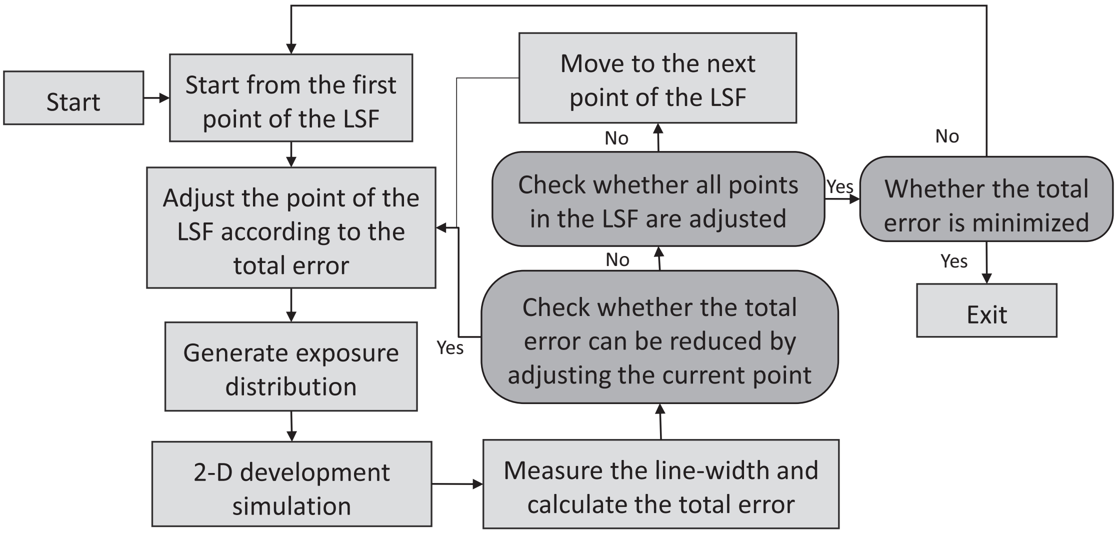
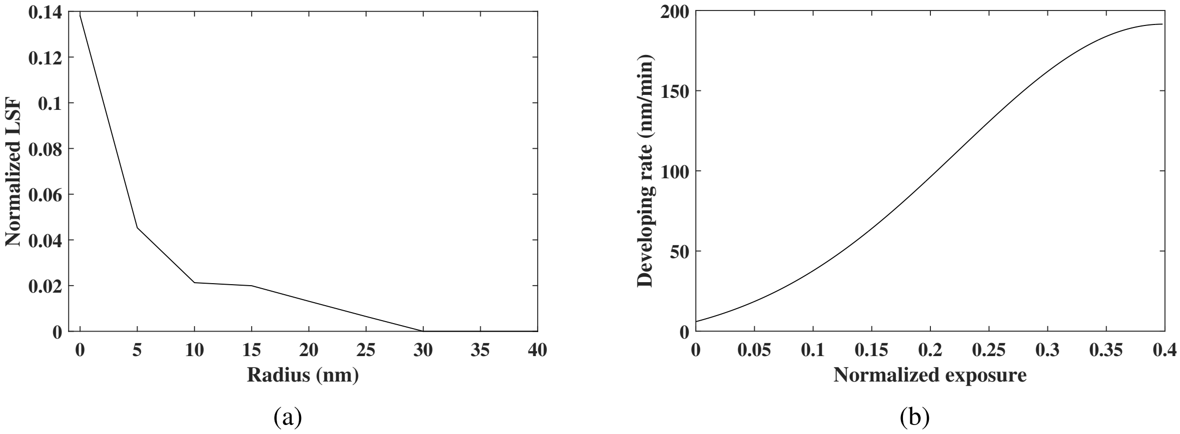


Figure 5.7: The flowchart of the iterative procedure for adjusting each point of the LSF.

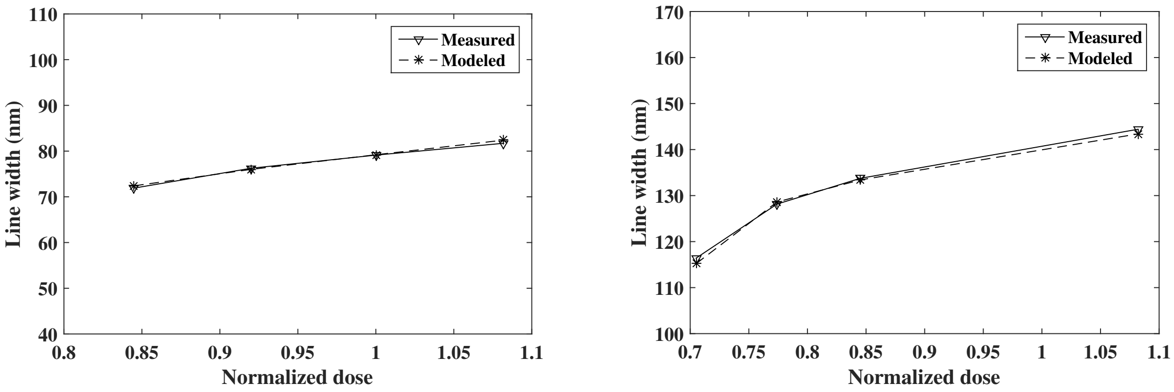
- All the previous points are set slightly lower than their optimal values so that the following points have more room to be adjusted.
- Each adjustment is linearly proportional to the current total error, since usually a larger error requires more adjustment.
- The current point of LSF is always set smaller than the previous points, since the LSF is a monotonously decreasing function.



(a) (b)

Figure 5.8: (a) The LSF and (b) conversion formula estimated without the exposure fluctuation.

The results of the modeled LSF and conversion formula are shown in Fig. 5.8, and the modeled line-widths at different doses and patterns are compared with their corresponding



(a)

(b)

Figure 5.9: The line-widths modeled through the simulation without the exposure fluctuation are compared with those measured from the SEM images : (a) L = S = 60 nm and (b) L = S = 120 nm.

measured ones from SEM images in Fig. 5.9. It can be seen that with the modeled LSF and conversion formula, the line-widths from the simulation are well matched with those measured from the SEM images.

With Exposure Fluctuation

Another factor that affects the line-width is the exposure fluctuation. For the same average developing-rate distribution, the exposure fluctuation can accelerate the development process and enlarge the line-width. This is due to the interaction between paths, i.e., the fast paths (paths with positive fluctuations) help the slow paths (paths with negative fluctuations) to achieve a larger average line-width, with the average developing-rate distribution unchanged. Therefore, in the second step, the results from the first step are further adjusted in an iterative procedure shown in Fig. 5.10 to compensate the effects of the exposure fluctuation.

In the second step, only the points in the LSF are adjusted assuming that the conversion formula derived in the first step is accurate enough. As the LSF and conversion formula together determine the line-width, a small error in the conversion formula may be compensated by adjusting the LSF. Also, to evaluate the effects of the exposure fluctuation, i.e., the interaction between development paths, the 3-D development simulation of the stochastic exposure needs to be used, which is significantly slower than the 2-D development simulation used in the first step. Therefore, adjusting each parameter in the conversion formula while optimizing the LSF

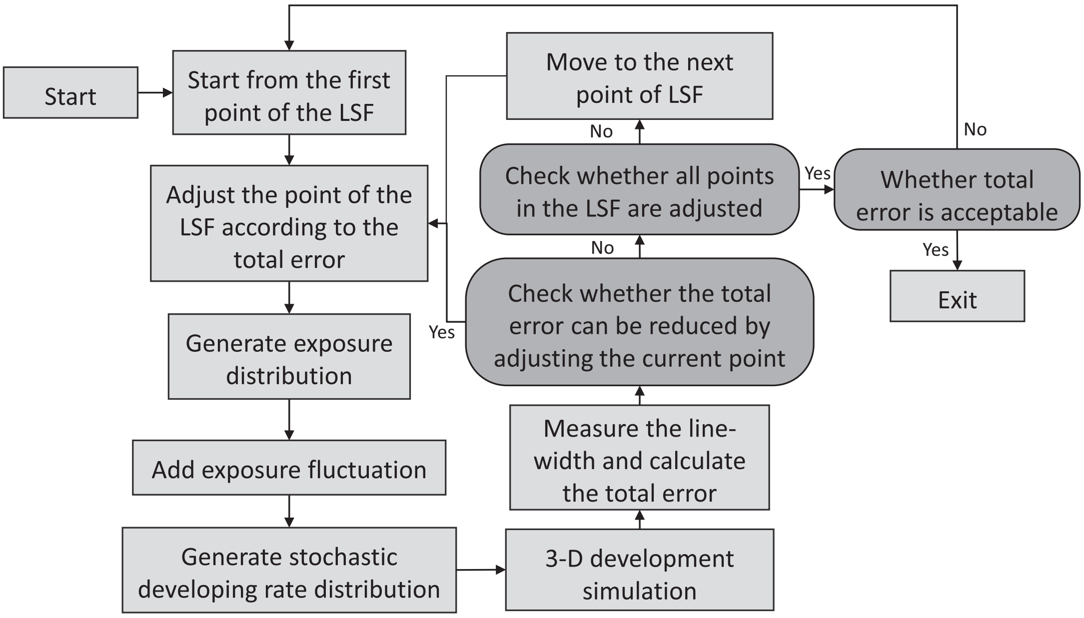
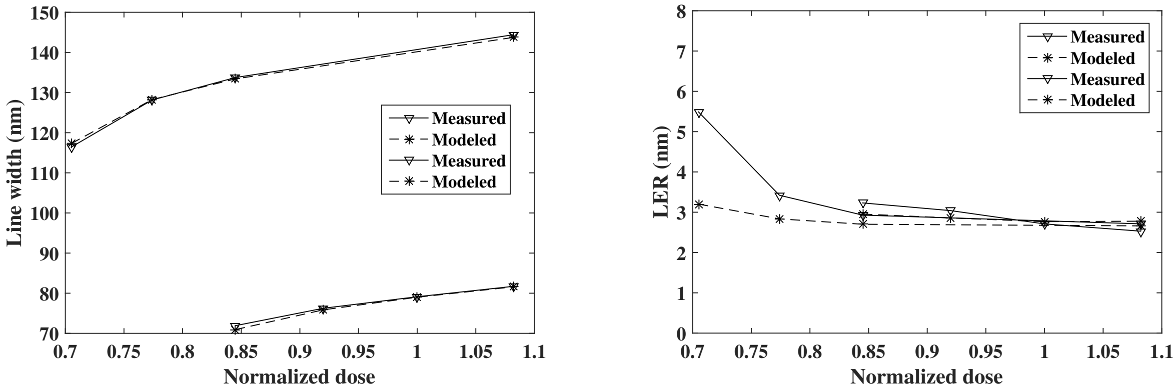


Figure 5.10: The flowchart of the iterative procedure for adjusting the LSF to compensate the effects of the exposure fluctuation on the line-width.

for the stochastic exposure is not practical. Without knowing the actual exposure fluctuation (in the experiments), a simple uniformly-distributed exposure fluctuation is added. To approximate the effects of the exposure fluctuation on the line-widths in the experiments, different levels of the exposure fluctuation are tested, and the one that can generate the same or similar LER level in the experiments is used in the optimization procedure (Fig. 5.10). The results are shown in Fig. 5.11.



(a) (b)

Figure 5.11: (a) The line-width and (b) LER modeled through the simulation with the exposure fluctuation are compared with those measured from the SEM images.

As can be seen from Fig. 5.11, after the adjustment of the LSF, the modeled line-widths are well matched with the measured ones, i.e., the effects of the exposure fluctuation on the line-widths are compensated. Note that the LER is not modeled in this step and, therefore, the difference between the LER's from the simulation and experiment can be significant, e.g., the LER's for the low doses in Fig. 5.11(b). Only the overall LER level is considered in this step to evaluate and compensate the effects of the exposure fluctuation on the line-widths.

5.2.3 Matching the LER

The (stochastic) fluctuation of exposure is modeled by matching the LER values obtained from the modeled exposure fluctuation to those measured in the SEM images. A Gaussian randomness is used to model the fluctuation of exposure, i.e., a zero-mean Gaussian noise, $e_n(x, y, z)$, is added to the (deterministic) exposure, $e_d(x, y, z)$, as follows:

$$\begin{aligned} e_s(x, y, z) &= e_d(x, y, z) + e_n(x, y, z), \\ e_n(x, y, z) &= G(x, y, z) \cdot \sigma_n(x, y, z) \end{aligned} \tag{5.4}$$

where $G(x, y, z)$ is generated by a Gaussian random number generator with (μ, σ) of $(0, 1)$, $e_s(x, y, z)$ represents the stochastic exposure, and $\sigma_n(x, y, z)$ determines the level of exposure fluctuation.

In general, a higher exposure level leads to a higher absolute fluctuation but a lower relative fluctuation. This property is incorporated into the implementation of exposure fluctuation as in Eq. 5.5.

$$\sigma_n(x, y, z) = \sigma_0 \left(\frac{e_d(x, y, z)}{\max\limits_{\forall x,y,z}(e_d(x, y, z))} \right)^{\alpha} \tag{5.5}$$

where $0 < \alpha < 1$.

It can be seen in Eq. 5.5 that as the exposure level ($e_d(x, y, z)$) increases, the absolute fluctuation $\sigma_n(x, y, z)$ becomes larger but the relative fluctuation, defined as $\frac{\sigma_n(x,y,z)}{e_d(x,y,z)}$, becomes smaller. Since one of the main sources of the exposure fluctuation is the random process of electron-scattering, there may exist a strong correlation of the exposure fluctuations between pixels. Let Cr_X and Cr_Y denote the correlations along the X and Y dimensions. To simplify the model, when $Cr_X = Cr_Y = 1$, the $e_s(x, y, z)$ is independent of the location (x, y, z). When $Cr_X = k$, every k pixels along the X dimension share the same level of fluctuation, and the same definition is applied to $Cr_Y = k$ but along the Y dimension.

In the optimization, four parameters, i.e., σ_0, α, Cr_X and Cr_Y, are determined by matching the LER. Since the exposure fluctuation also affects the line-width, the first point of LSF ($LSF(0)$) is adjusted to compensate the effects. Through a binary search, all of the five parameters, i.e., σ_0, α, Cr_X, Cr_Y, and $LSF(0)$, are adjusted such that the total error of the line-width (CD) and LER is minimized. The total error is defined as

$$TotalError = \sum_{dose\ level} (|Error_{CD}| + 10 \cdot |Error_{LER}|) \tag{5.6}$$

where the $Error_{CD}$ is the difference between the modeled and measure line-widths, and the $Error_{LER}$ is the difference between the modeled and measure LER's.

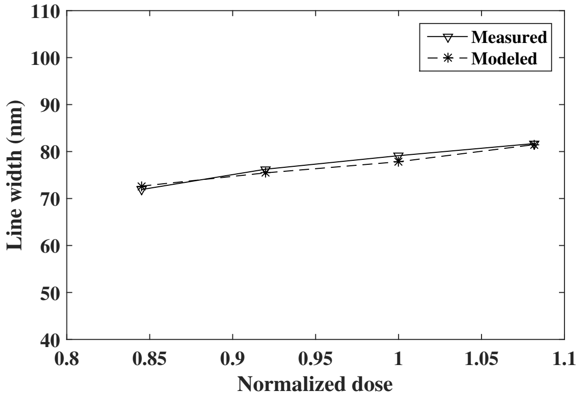
In most cases, the LER is much smaller than the CD. In order to achieve the percent LER error comparable to the percent CD error, a weight of 10 is multiplied to the LER error.

5.3 Results and Discussion

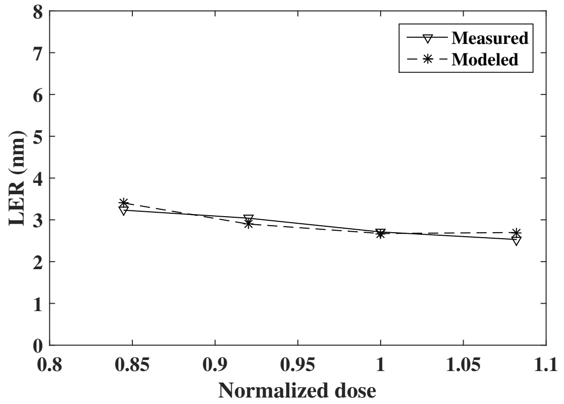
Two sets of SEM images obtained under the same experimental set-up (except the dose) are used in the modeling. One set includes 4 SEM images corresponding to the 4 normalized dose levels of 0.845, 0.920, 1.000, and 1.082 where both the line-width and space are 60 nm. The other set includes 4 SEM images corresponding to the 4 normalized dose levels of 0.705, 0.774, 0.845, and 1.082 where both the line-width and space are 120 nm. The beam energy is 50keV. Other details of the experiment are unknown.

In order to minimize the statistical uncertainty of the results, all of the lines in each SEM image, i.e., 12 lines in the pattern of 60 nm lines and 6 lines in the pattern of 120 nm lines, are utilized in measuring the line-width and LER. The line-width and LER obtained by the proposed modeling method are compared with those measured from the SEM images for the

60 nm line pattern in Fig. 5.12. It can be seen that the modeled line-width and LER are closely matched with the measured results. A similar close match can be observed for the pattern of 120 nm lines in Fig. 5.13.



(a)

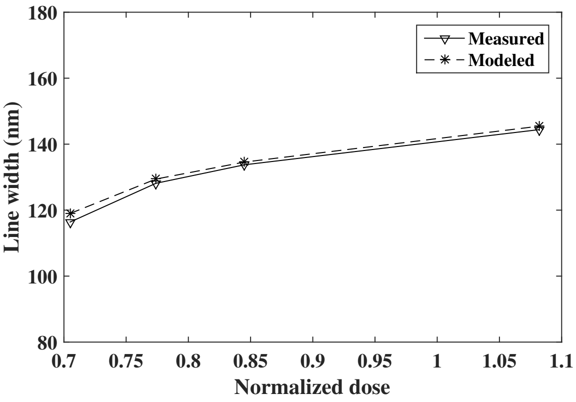


(b)

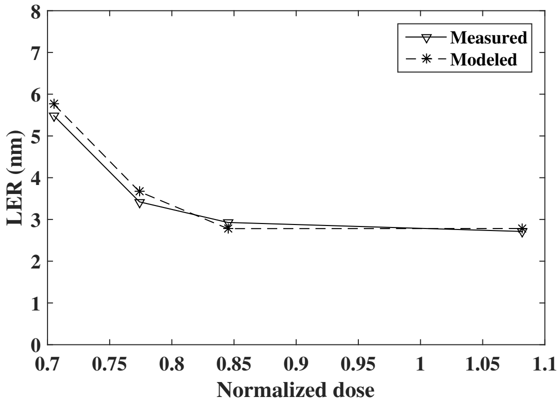
Figure 5.12: The line-width and LER modeled through simulation are compared with those measured from the SEM images of a L/S pattern with L = S = 60 nm: (a) average line-width and (b) LER.

5.4 Summary

Since the simulation may not be completely realistic, the experimental results may be used to verify the analytic method. However, the stochastic information, i.e., the mean and variance of the developing rate, required by the analytic method, cannot be directly measured from the experiments. In this chapter, a practical and accurate method of extracting the stochastic information of e-beam lithographic process, i.e., the LSF, conversion formula, and exposure



(a)



(b)

Figure 5.13: The line-width and LER modeled through the simulation are compared with those measured from the SEM images of a L/S pattern with L = S = 120 nm: (a) average line-width and (b) LER.

fluctuation, is described. One of the main advantages is that the method utilizes only the information extracted from SEM images with the normalized dose levels without having to know the complete set-up of e-beam lithographic process. Since the proposed method utilizes the experimental results of SEM images, its results are likely to be realistic. Another important aspect of the proposed method is that the LSF, conversion formula, and exposure fluctuation are modeled by considering their interaction rather than individually. Therefore, a minor inaccuracy in one of them may be compensated by the others such that the modeled CD and LER can achieve a high accuracy compared with the those measured from experimental results.

Chapter 6

Verification

In this chapter, the analytic method is further verified through a more thorough comparison. In the first section (Sec. 6.1), the experimental results are used to verify the analytic method, i.e., the stochastic information (the mean and variance of exposure distribution) extracted from SEM images (see Chapter 5), is used by the analytic method to derive the CD and LER, which are compared with those measured from the experimental results. In the second section (Sec. 6.2), all the factors that may affect the LER are controlled individually in generating PSF's to be employed in the simulation, and the corresponding results of the analytic method are compared with those from the simulation.

6.1 Verification from Experimental Results

In the previous charter, a method of extracting the stochastic information from the experimental results by modeling the e-beam lithographic process, i.e., modeling (estimating) the line spread function (LSF), conversion formula, and exposure fluctuation, through analyzing SEM images, is described. With the information extracted, the mean of exposure distribution $m_e(x, y, z)$ in the resist can be computed through the convolution of LSF and dose distribution as

$$m_e(x, y, z) = \int D(x - x', y) lsf(x', z) dx'. \tag{6.1}$$

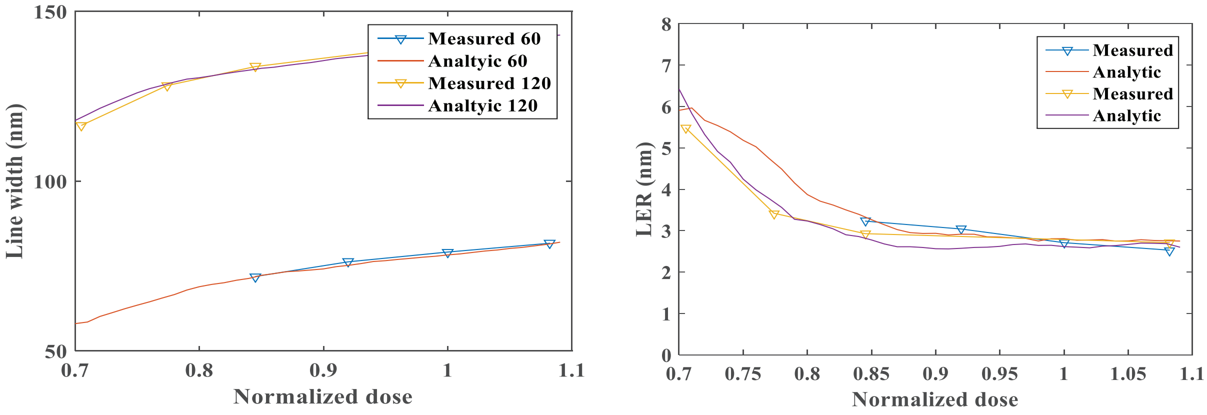
In addition to the mean of exposure distribution, the variance of exposure distribution is also required by the analytic method. Since the exposure fluctuation is modeled by a zero-mean Gaussian noise defined in Eq. 5.4 and Eq. 5.5, the variance of exposure distribution can be derived by the variance of the added Gaussian noise as

$$\sigma_e^2(x,y,z) = \sigma_n^2(x,y,z) = \sigma_0^2 \left(\frac{m_e(x,y,z)}{\max\limits_{\forall x,y,z}(m_e(x,y,z))} \right)^{2\alpha} \tag{6.2}$$

where α and σ_0 are the parameters derived by the modeling method.

Note that the deterministic exposure distribution $e_d(x,y,z)$ in Eq. 5.5 is equal to the mean of exposure distribution $m_e(x,y,z)$, when a zero-mean Gaussian noise is added.

Through the conversion formula (Eq. 5.3) derived in the modeling method, the mean and variance of developing-rate distribution can be derived from those of exposure. By applying the analytic method described in Chapter 3, the analytic results (CD and LER) are compared with the ones measured from SEM images in Fig. 6.1.



(a) (b)

Figure 6.1: The line-width and LER derived by analytic method are compared with those measured from SEM images of L/S patterns: (a) average line-width and (b) LER.

As can be seen from Fig. 6.1, the results obtained by the analytic method are closely matched with those measured from SEM images. However, since the sample size of the experimental results is relatively small (4 points for the pattern of 120 nm lines and 4 points for the pattern of 60 nm lines), the overall tendency (shape) of the LER derived by the analytic method appears to be different from that of the experimental results. In addition, some difference may be caused by the errors in the stochastic information, i.e., the mean and variance of exposure, derived by the modeling method (in Chapter 5), which are included in the analytic results.

6.2 Verification from Generated PSF's

In this section, more PSF's are generated from the Monte Carlo simulation to test the analytic method. It is shown [15, 32] that the LER is mainly affected by the following parameters: the fluctuation, contrast, and shape of exposure distribution. Therefore, each of these parameters is controlled individually in generating different sets of PSF's, and the corresponding results by the analytic method are compared with those from the simulation.

6.2.1 The Fluctuation of Exposure

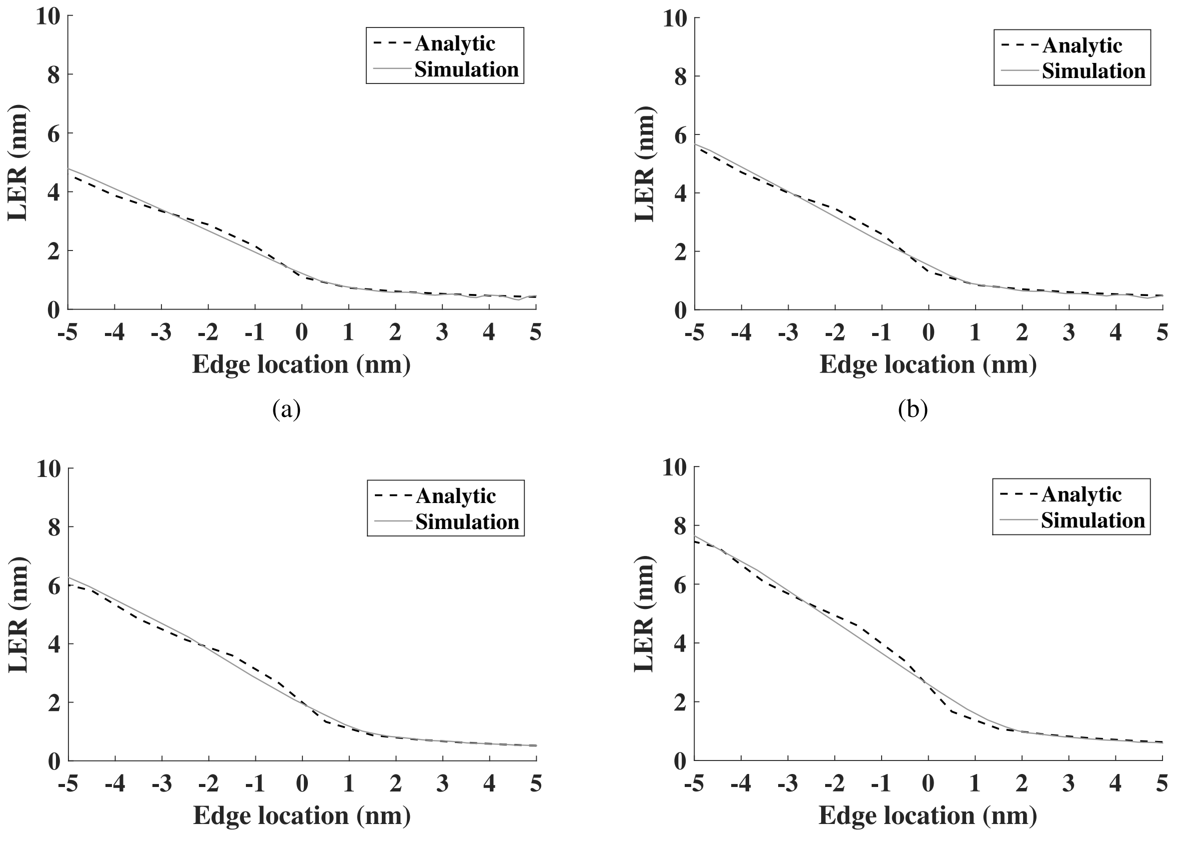
The LER is caused by a number of stochastically fluctuating effects such as shot noise, distributions of chemical species in the resist such as photoacid generator (PAG), resist development process, etc [15, 32]. All those effects can be described by or equivalently mapped to the fluctuation of exposure. The exposure fluctuation can be quantified in two ways, i.e., the absolution exposure fluctuation, which is defined as the standard deviation of exposure $\sigma_e(x, y, z)$, and the relative exposure fluctuation, which is defined as

$$\overline{\sigma}_e(x, y, z) = \frac{\sigma_e(x, y, z)}{m_e(x, y, z)}. \tag{6.3}$$

The Absolute Exposure Fluctuation

To fix the relative exposure fluctuation and change the absolute exposure fluctuation only, $\sigma_e(x, y, z)$ is set to a fixed proportion, e.g., from 5% to 20%, of the $m_e(x, y, z)$, such that the relative exposure fluctuation $\overline{\sigma}_e(x, y, z)$ is spatially invariant. In the first set of results, a typical substrate system is employed with a resist layer of 300 nm PMMA on Si. A single-line feature with a size of 100 nm × 400 nm ($W \times L$) is exposed with a uniform dose. The corresponding analytic and simulation results are compared in Fig. 6.2.

As expected, with a larger absolution exposure fluctuation, the LER is larger, especially inside the feature, i.e., the edge location is less than 0. The results from the analytic are closely matched with the simulation results. However, it is observed that the analytic method may generate a larger error for the case of a larger absolute exposure fluctuation. This is due to the



(a)

(b)

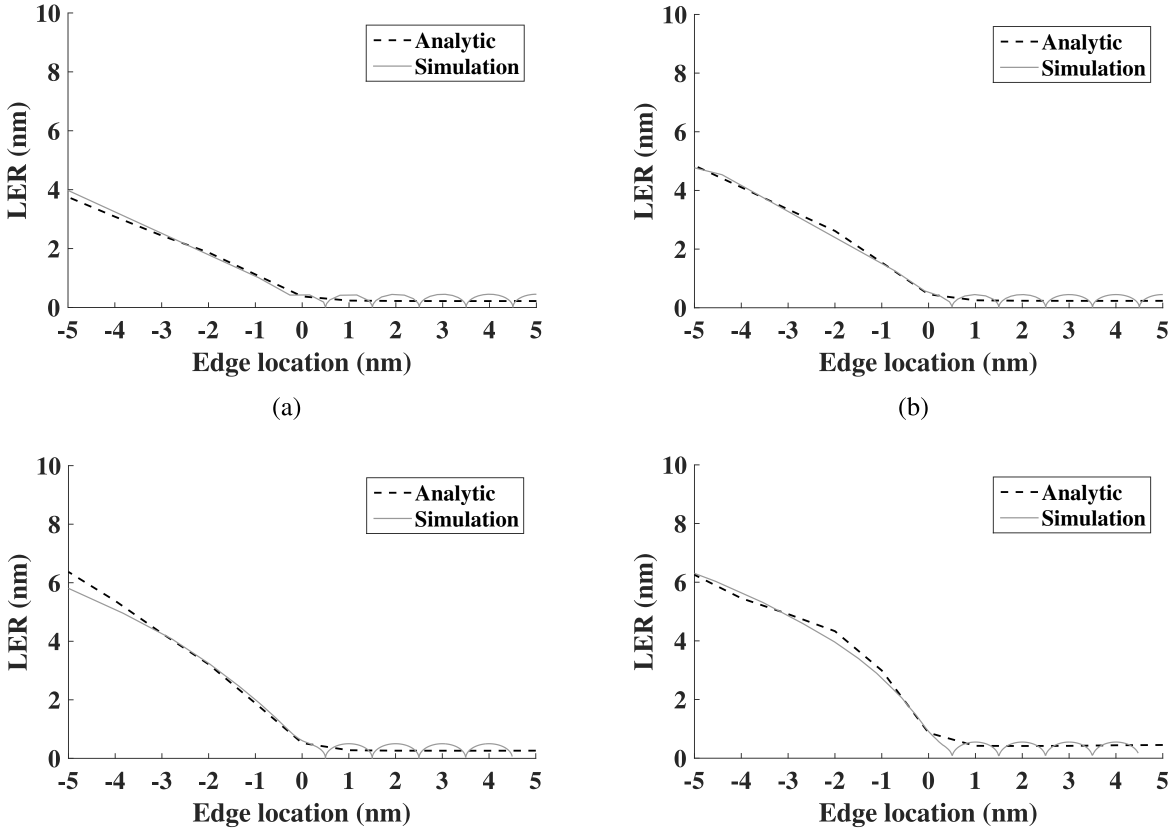
(c)

(d)

Figure 6.2: The LER estimated by the analytic method and simulation for 300 nm PMMA on Si, beam energy of 50 keV, dose of 640 μC/cm^2 with different absolute exposure fluctuations of (a) 5%, (b) 10%, (c) 15%, and (d) 20%.

interaction between development paths. For example, when the absolute exposure fluctuation is large, during the resist development process, the interaction is not only from the adjacent paths but also from the paths of several pixels away (in the Y dimension), which is not considered in the analytic method. It is also shown that in these cases, with different absolute exposure fluctuations, the shape of LER (as a function of edge location) does not change significantly, e.g., the overall shape is approximately scaled by a constant. Another set of PSF's for a thinner resist of 100 nm PMMA on Si is generated where the PSF is shaper. With the same absolute exposure fluctuations added, i.e., from 5% to 20% of $m_e(x, y, z)$, the analytic and simulation results are compared in Fig. 6.3.

A similar tendency is observed that with a larger absolute fluctuation, the LER is larger. For the simulation results in Fig. 6.3, some ripples can be observed where the LER level is low, i.e., the edge location is outside a feature. The ripples are due to an artificial effect introduced



(a) (b) (c) (d)

Figure 6.3: The LER estimated by the analytic method and simulation for 100 nm PMMA on Si, beam energy of 50 keV, dose of 640 μC/cm^2 with different absolute exposure fluctuations of (a) 5%, (b) 10%, (c) 15%, and (d) 20%.

by the simulation, i.e., the developing-rate distribution is discrete rather than continuous. When the LER level is low, most development paths are located in the same pixel and behave similarly. Therefore, the variation of the development paths, i.e., LER, can increase within the pixel. This effect is reduced when some development paths are moving into a different pixel. As a result, the LER level oscillates with the pixel size of 1 nm in the simulation results. When the LER level is high, i.e., the edge location is inside a feature, development paths are likely to be located in the different pixels and, therefore, no ripple can be observed.

By comparing Fig. 6.2 with Fig. 6.3, it can be seen that the LER for 100 nm PMMA is smaller than that for 300 nm PMMA. One of the reasons is that given the same beam energy of 50 keV, the contrast of exposure for 100 nm PMMA is much larger than that for 300 nm PMMA, and the LER is inversely proportional to the exposure contrast. Also, it is noticeable that in the case of a larger absolute fluctuation, the shape of LER is also changed in addition to its level.

The Relative Exposure Fluctuation

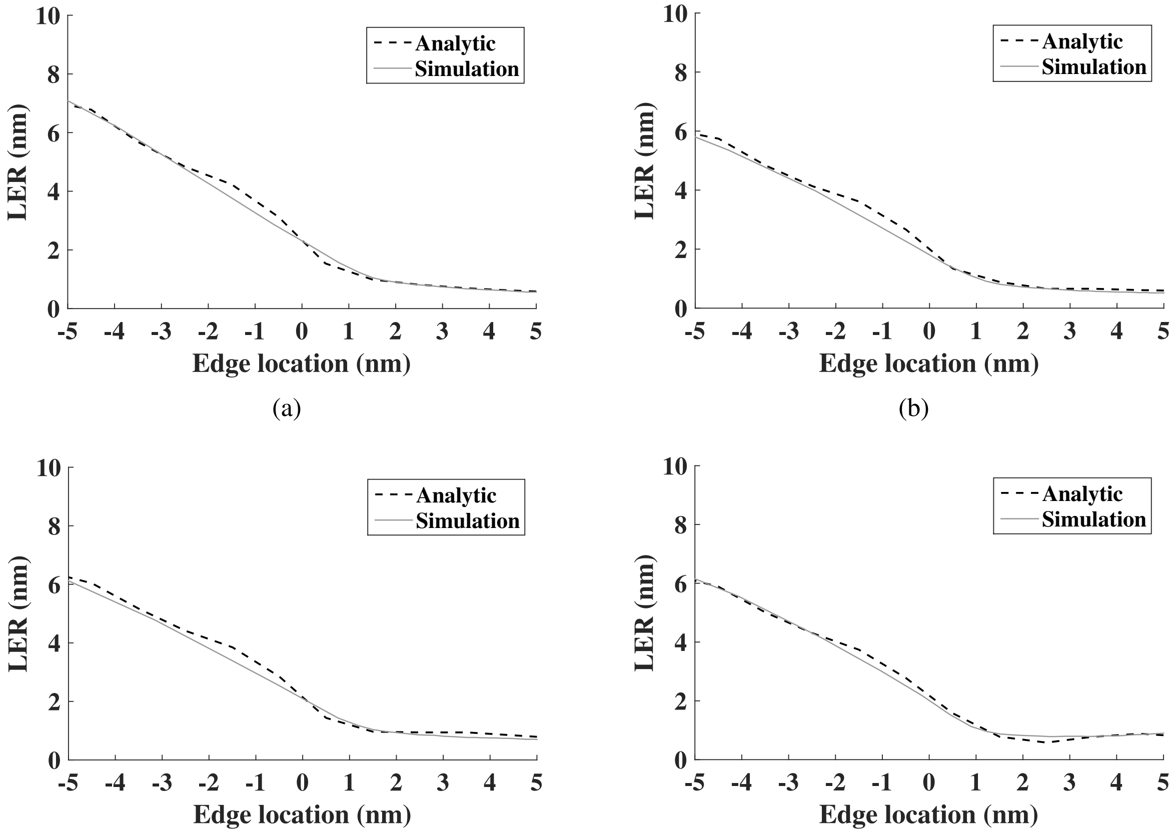
In the writing process of e-beam lithography, the energy deposited inside the feature, which is mainly contributed from the forward-scattering of electrons, can be significantly different from the energy deposited outside the feature, which is mainly from the back-scattering of electrons. As a result, the absolute fluctuation of the exposure inside the feature is much higher than that outside the feature, but the relative exposure fluctuation behaves in the other way. To model such a behavior, the following equation is used to derive the exposure fluctuation:

$$\begin{aligned}
\sigma_e(x,y,z) &= \sigma_0\left(\frac{m_e(x,y,z)}{\max\limits_{\forall x,y,z}(m_e(x,y,z))}\right)^{\alpha}, \\
\overline{\sigma}_e(x,y,z) &= \frac{\sigma_e(x,y,z)}{m_e(x,y,z)} = \sigma_0\left(\frac{1}{\max\limits_{\forall x,y,z}(m_e(x,y,z))}\right)^{\alpha}\frac{1}{m_e(x,y,z)^{1-\alpha}}. \qquad (6.4)
\end{aligned}$$

The relative fluctuation is modeled by α where $0 < \alpha \leq 1$, i.e., a smaller α leads to a larger relative fluctuation outside the feature. The absolute fluctuation is controlled by σ_0, which is set to 17% of $\max\limits_{\forall x,y,z}(m_e(x,y,z))$ to match the fluctuations of the stochastic exposure generated by the Monte Carlo simulation with the resist thickness of 300 nm PMMA, beam energy of 50 keV, and beam diameter of 3 nm. A single-line pattern with a size of 100 nm × 400 nm ($W \times L$) is exposed with a uniform dose of 640 μC/cm^2. The corresponding analytic and simulation results are compared in Fig. 6.4.

As can be seen from Fig. 6.4, the relative exposure fluctuation affects the shape of LER, especially outside the feature. Since the development process is cumulative, with a larger exposure fluctuation outside the feature, the LER may decrease less and even increase. Such effects are more significant for a shaper PSF, e.g., the PSF for a thinner resist of 100 nm PMMA with beam energy of 50 keV. The analytic and simulation results are compared in Fig. 6.5.

Comparing Fig. 6.5 and Fig. 6.4, with the same α (the same level of relative fluctuation), the LER outside the feature is affected more significantly by the relative fluctuation when the



(a) (b)

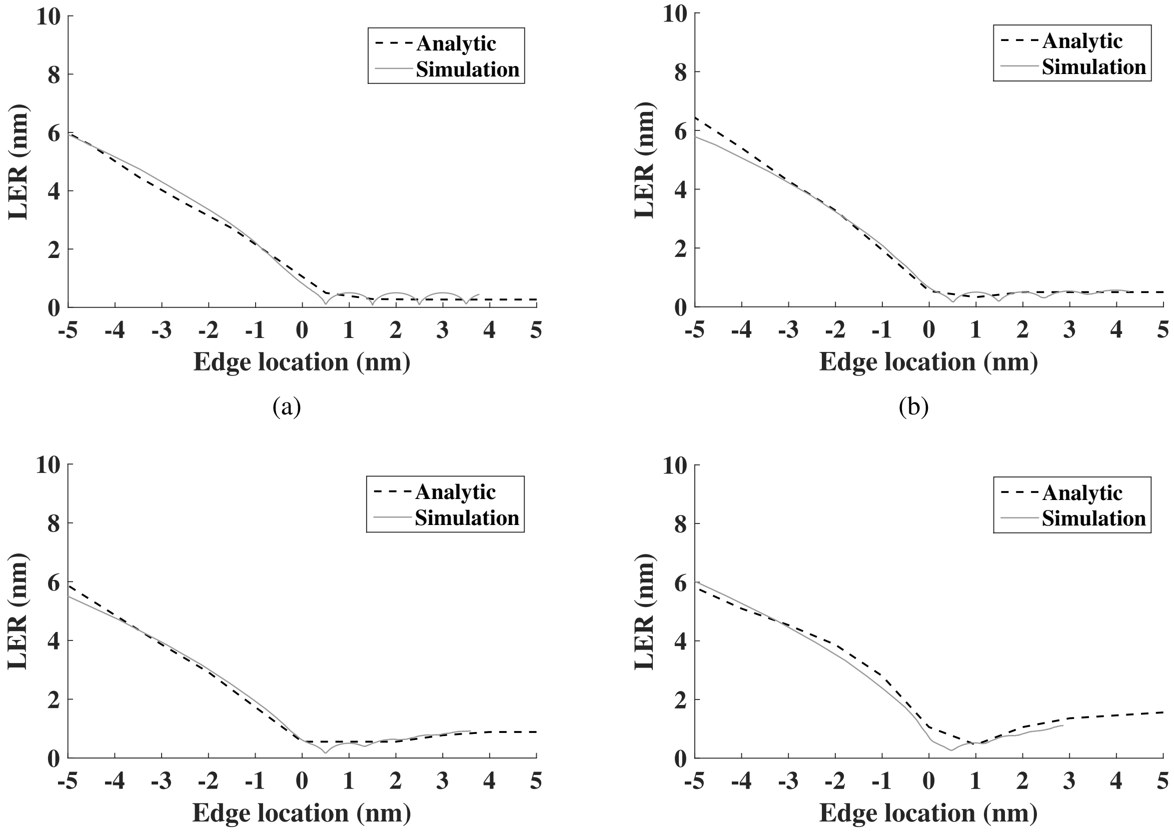
(c) (d)

Figure 6.4: The LER estimated by the analytic method and simulation for 300 nm PMMA on Si, beam energy of 50 keV, dose of 640 μC/cm^2 with different values of α: (a) 1, (b) 0.8, (c) 0.6, and (d) 0.4.

PSF is shaper. For a shaper PSF, the exposure decreases faster over the edge and stays flat outside the feature. Similarly, without considering the cumulative effect, the LER outside the feature also stays flat. At the same time, when there is more exposure fluctuation outside (a larger relative fluctuation), the LER tends to increase more as it is cumulative.

6.2.2 The Contrast of Exposure

Another factor, which affects the LER is the contrast of exposure. The contrast of exposure can be defined as the difference of exposure levels between the inside and the outside of a feature. Since this study focuses on the LER around the edge, i.e., edge location from -5 nm to 5 nm, the contrast of exposure is quantified as the difference between the peak of the exposure (at the middle point of the line feature) and the exposure at the edge, i.e., the edge location is 0 nm. To



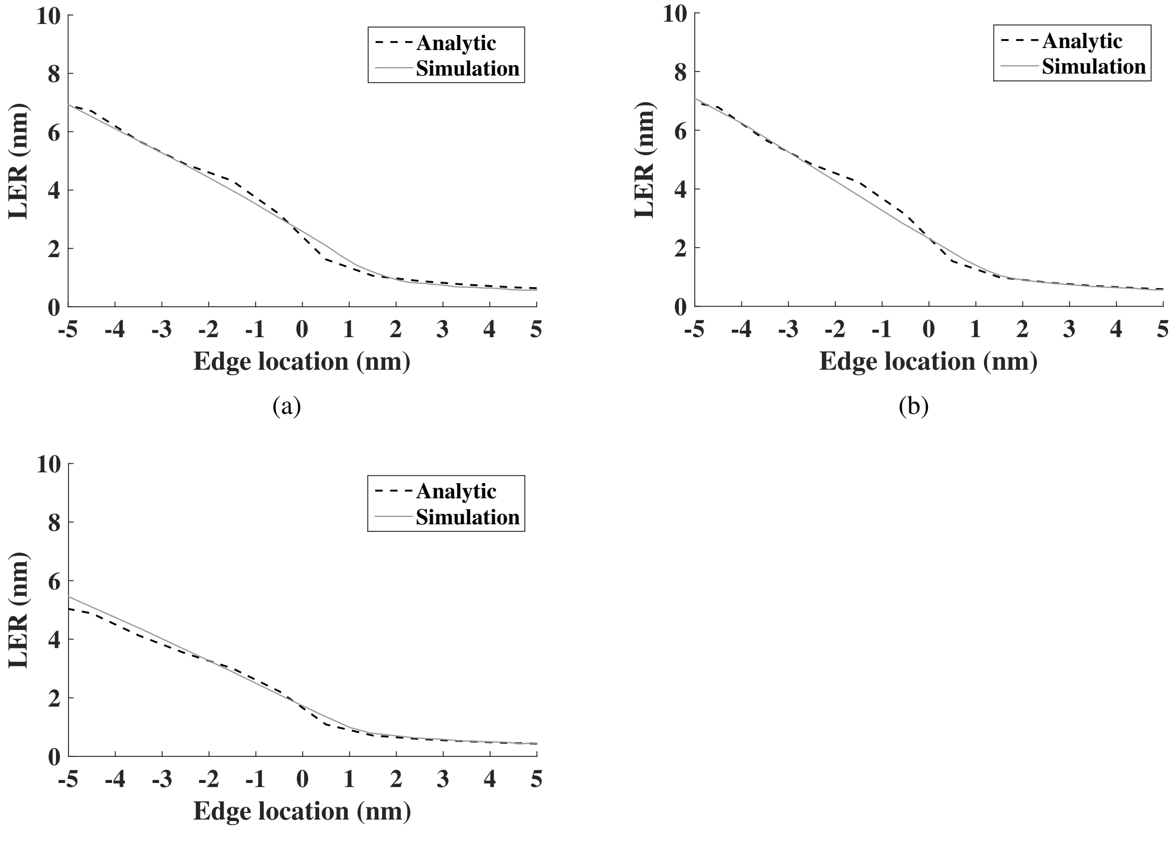
(a) (b) (c) (d)

Figure 6.5: The LER estimated by the analytic method and simulation for 100 nm PMMA on Si, beam energy of 50 keV, dose of 640 μC/cm^2 with different values of α: (a) 1, (b) 0.8, (c) 0.6, and (d) 0.4.

analyze the contrast of exposure without changing other parameters, the stochastic exposure is generated by the following procedure:

1. Compute the mean and variance of exposure from the mean and variance of PSF.
2. Scale the mean of exposure by a certain factor to change the contrast.
3. Add the same level of exposure fluctuation according to the variance of exposure.

In this procedure, the absolute exposure fluctuation is maintained while the contrast is changed. The relative exposure fluctuation is spatially invariant. For a typical substrate system with a resist layer of 300 nm PMMA on Si, define the contrast of exposure, generated by the Monte Carlo simulation when a single-line feature is exposed, to be 1. The cases of the exposure contrast changing from 0.8 to 1.2 are tested, and the analytic and simulation results are provided in Fig. 6.6.



(a)

(b)

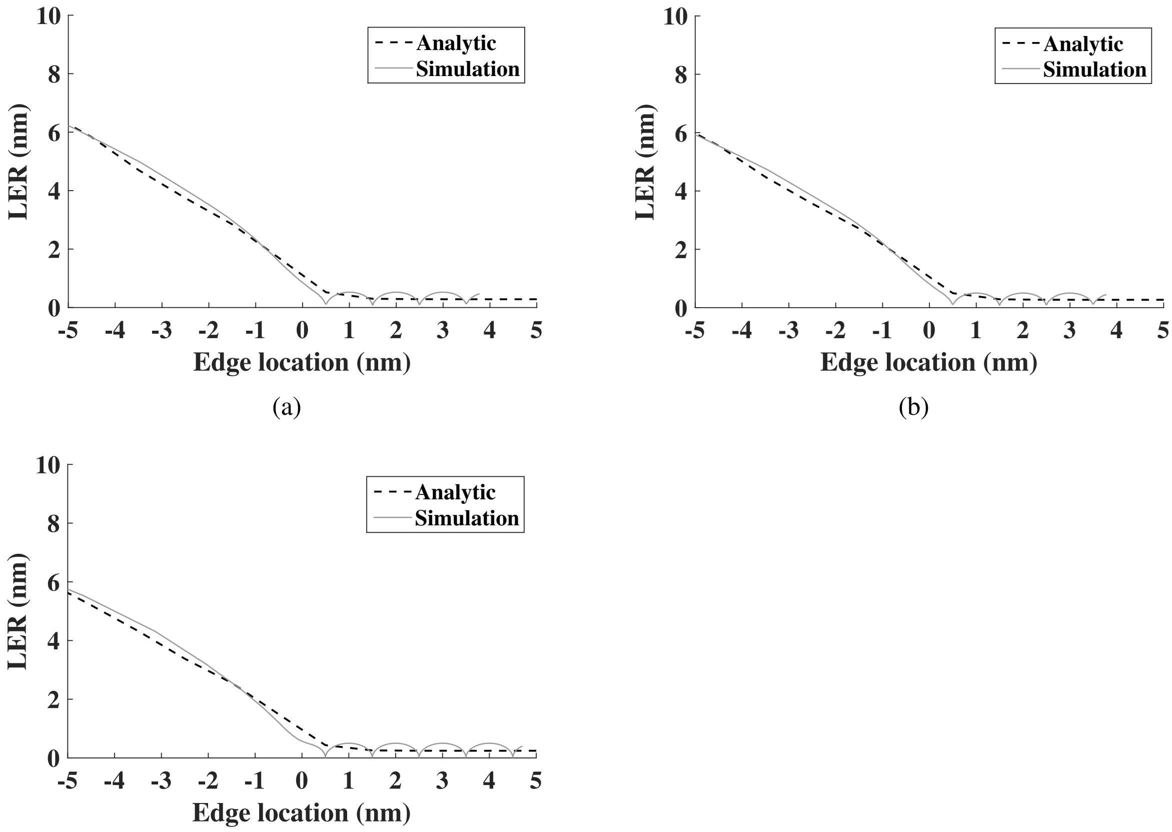
(c)

Figure 6.6: The LER estimated by the analytic method and simulation for 300 nm PMMA on Si, beam energy of 50 keV, dose of 640 μC/cm^2 with different values of exposure contrast: (a) 0.8, (b) 1, and (c) 1.2.

As can be seen from the results, a larger contrast of exposure leads to a smaller LER, as the LER is inversely proportional to the exposure contrast. Also, the cases of the exposure contrast for a sharper PSF, i.e., the PSF for the resist thickness of 100 nm PMMA on Si with the bean energy of 50 keV, are tested. The simulation and analytic results are compared in Fig. 6.7, and a similar tendency can be observed.

6.2.3 The Shape of Exposure

The shape of exposure distribution over the feature edge, which is determined by the shape of PSF, is one of the main factors that affect the LER. To test the analytic method thoroughly, a factor k is used to control the shape of PSF, i.e., when $k = 0$, the exposure distributions of all layers (z) share the same shape, and a larger difference among the layers is introduced as k increases. More specifically, when k increases, the exposure distribution at the top layer



(a) (b) (c)

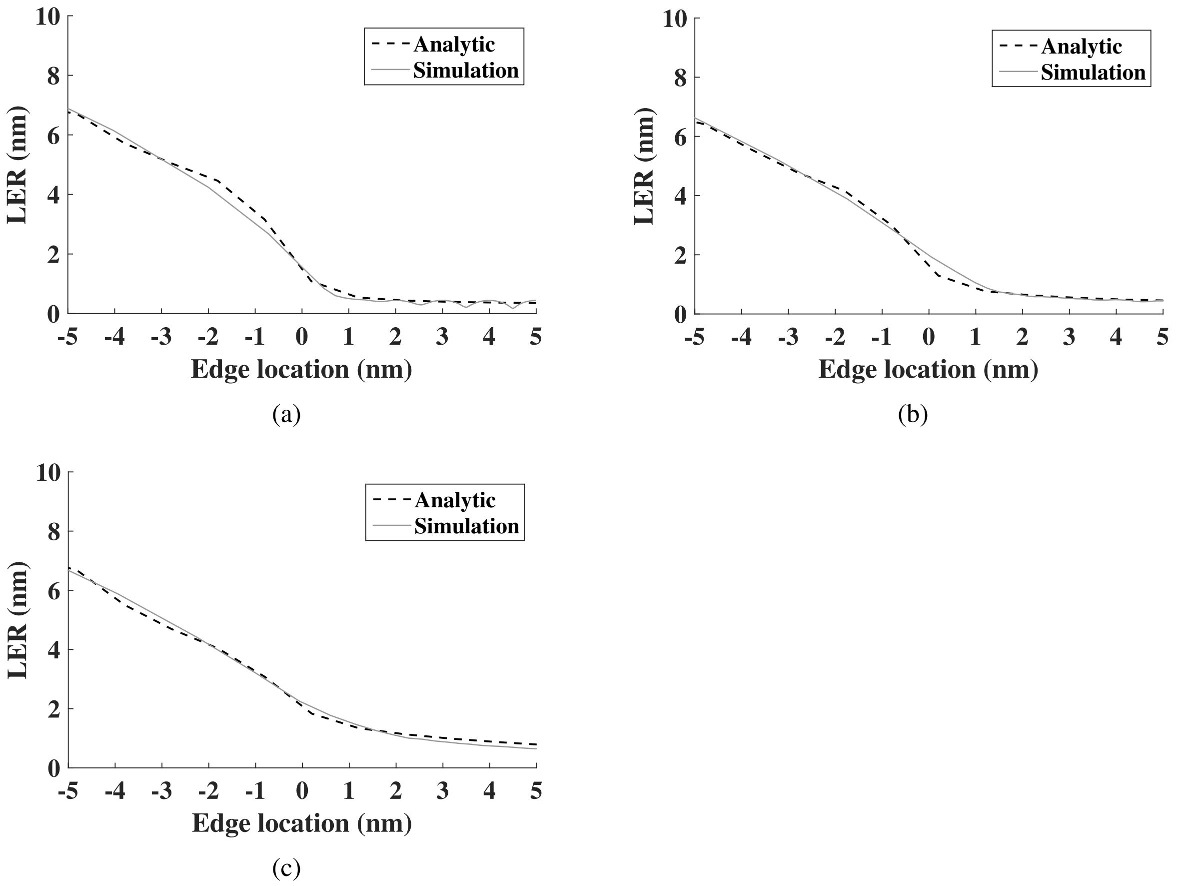
Figure 6.7: The LER estimated by the analytic method and simulation for 100 nm PMMA on Si, beam energy of 50 keV, dose of 640 μC/cm^2 with different values of exposure contrast: (a) 0.8, (b) 1, and (c) 1.2.

becomes sharper, and the exposure distribution at the bottom layer becomes broader. Given the exposure distribution at the middle layer ($e_{mid}(x, y)$) fixed, the exposure distributions at other layers, e.g., $e(x, y, z)$ at a layer z, are calculated by the following equation:

$$e(x,y,z) = \left(\left(\frac{e_0(x,y,z)}{\max\limits_{\forall x,y}(e_i(x,y))} \cdot \frac{\max\limits_{\forall x,y}(e_{mid}(x,y))}{e_{mid}(x,y)} - 1 \right) k + 1 \right) \frac{e_{mid}(x,y) \max\limits_{\forall x,y}(e_i(x,y))}{\max\limits_{\forall x,y}(e_{mid}(x,y))} \tag{6.5}$$

where $e_0(x, y, z)$ is generated by the Monte Carlo simulation.

For a typical substrate system of a resist layer of 300 nm PMMA on Si, different values of k are tested, and the analytic and simulation results of LER at the bottom layer are compared in Fig. 6.8.



(a)

(b)

(c)

Figure 6.8: The LER estimated by the analytic method and simulation for 300 nm PMMA on Si, beam energy of 50 keV, dose of 640 μC/cm^2 with different shapes of exposure defined by k: (a) 0, (b) 1, and (c) 2.

As can be seen from the results, the shape of LER, i.e., the LER varying with the edge location, has a strong correlation with the shape of exposure. With a sharp exposure shape, i.e, $k = 0$, the LER decreases rapidly as the edge location increases. And with a flatter shape of exposure, i.e., $k = 2$, the LER decreases slower from the inside to the outside of a feature.

Chapter 7

Concluding Remarks and Future Work

7.1 Conclusions

As the feature size decreases, the LER will eventually become a resolution-limiting factor in the e-beam lithography since the LER does not scale with the feature size. Therefore, it is essential to minimize the LER in order to achieve the highest resolution possible by the e-beam lithographic process. An simulation based approach may be employed. However, since the quantitative relationship between the LER and e-beam lithographic parameters is unknown, it is not easy to minimize the LER by optimizing those parameters.

This dissertation describes the efforts of developing an analytic method of estimating and minimizing the LER in e-beam lithography. First, a complete analytic model is briefly described including the models of PSF, exposure, conversion formula from exposure to developing rate, and development path, which are used in the analytic method. The simulation procedure whose results are used to verify the analytic method is also depicted. Then, an analytic method of estimating and minimizing the LER for a single-line pattern is developed based on the mean and variance of developing-rate distribution in the resist. It has been shown that the results obtained by the analytic method are closely matched with the simulation results.

Then, the analytic method is extended to a large-scale uniform pattern based on the LER expression for a single line, by relating the location dependency of LER to the spatially-varying global exposure. The analytic LER expressions at the critical locations are derived first, and the LER's at other locations are calculated through an interpolation of the LER's at the critical locations. It has been shown that the results obtained by the analytic method of estimating the

LER are closely matched with the simulation results. Then, the deconvolution surface is used to derive the optimal dose distribution of a large pattern, which can minimize the LER and CD error at all the locations in a large pattern.

Since the simulation may not be completely realistic, the experimental results (SEM images) can be used to verify the analytic method. However, since the stochastic information required by the analytic method cannot be directly measured from SEM images, a practical and accurate method of modeling the LSF, conversion formula, and exposure fluctuation directly from SEM images has been developed. One of the advantages is that the developed method utilizes only the information extracted from SEM images with the normalized dose levels without having to know the complete set-up of e-beam lithographic process. The results (CD and LER) obtained from the modeling method are shown to be closely matched with those measured from SEM images.

To verify the results of the analytic method, a set of experimental results is first utilized, i.e., by using the stochastic information extracted from SEM images, the LER estimated by the analytic method are compared with those measured from SEM images. It has been shown that the analytic estimates of LER are closely matched with the measured ones. Then, more PSF's are generated by the Monte Carlo simulation to control the factors that affect the LER. The fluctuation, contrast, and shape of exposure are controlled individually, and the corresponding analytic results are compared with the simulation results where a close match can be observed.

In conclusion, an analytic method of estimating and minimizing the LER for a single line and a large-scale uniform pattern exposed with a uniform dose has been developed. This method provides an alternative to the time-consuming and costly experiment and simulation in a LER study. Though the method is not completely analytic and requires numerical computation in some steps, the good accuracy achieved indicates that it has a potential to be further developed into a practical and useful tool in the area of e-beam lithography.

7.2 Future work

A possible future work is to make the method (described in this dissertation) more analytic and the results more explicit in terms of parameters such as the dose and developing time. In some steps of the analytic method, a numerical computation or an iterative procedure is required. It would be necessary to remove or minimize such computation and procedure in order to have a more analytic method. Also, it might be possible to improve the accuracy and applicability of the method such that the assumptions made in developing the method are not necessary.

Another possibility is to extend the analytic method for the cases where each feature is exposed with a non-uniform dose. A straightforward approach would be to compute the exposure distribution through the convolution given a spatial distribution of dose, convert the exposure into the developing rate, and then follow the procedures developed for a single line. However, the final expression of LER is likely to be too complicated to be practical and useful. An approach which can utilize the result of a single line for each region of a uniform dose within a feature might lead to a useful method.

References

[1] R. Rau, J. McClellan, and T. Drabik, "Proximity effect correction for nanolithography," J. Vac. Sci. Technol. B **14**, 2445 (1996).

[2] G. P. Watson, L. A. Fetter, and J. A. Liddle, "Dose modification proximity effect correction scheme with inherent forward scattering corrections," J. Vac. Sci. Technol. B **15**, 2309 (1997).

[3] G. P. Watson, S. D. Berger, and J. A. Liddle, "Characterizing GHOST proximity effect correction effectiveness by determing the worst-case error," J. Vac. Sci. Technol. B **16**, 3256 (1998).

[4] R. Murali, D. Brown, K. Martin, and J. Meindl, "Proximity effect correction using blur map in electron projection lithography," J. Vac. Sci. Technol. B **24**, 2936 (2006).

[5] Q. Dai, S.-Y. Lee, S.-H. Lee, B.-G. Kim, and H.-K. Cho, "Three-dimensional proximity effect correction for large-scale uniform patterns," J. Vac. Sci. Technol. B **29**, 06F314 (2011).

[6] H. Eisenmann, T. Waas, H. Hartmann, "PROXECCO-Proximity effect correction by convolution," J. Vac. Sci. Technol. B **11**, 2741 (1993).

[7] S.-Y. Lee and B. D. Cook, "PYRAMID-a hierarchical, rule-based approach toward proximity effect correction. I. Exposure estimation," IEEE Trans. Semicond. Manuf. **11**, 108 (1998).

[8] S.-Y. Lee and K. Anbumony, "Accurate control of remaining resist depth for nanoscale three-dimensional structures in electron-beam grayscale lithography," J. Vac. Sci. Technol. B **25**, 2008 (2007).

[9] M. Osawa, K. Takahashi, M. Sato, and H. Arimoto, K. Ogino, H. Hoshino, and Y. Machida, "Proximity effect correction using pattern shape modification and area density map for electron-beam projection lithography," J. Vac. Sci. Technol. B **19**, 2483 (2001).

[10] Q. Dai, S.-Y. Lee, S.-H. Lee, B.-G. Kim, and H.-K. Cho,"New types of dose distributions for vertical sidewall minimizing total dose in 3-D electron-beam proximity effect correction of nanoscale features," J. Vac. Sci. Technol. B **30**, 06F307 (2012).

[11] Q. Dai, S.-Y. Lee, S.-H. Lee, B.-G. Kim, and H.-K. Cho, "Estimation of resist profile for line/space patterns using layer-based exposure modeling in electron-beam lithography," Microelectron. Eng. **88**, 902 (2011).

[12] M. Nagase, H. Namatsu, K. Kurihara, K. Iwadate, K. Murase and T. Makino, "Critical Dimension Measurement in Nanometer Scale by Using Scanning Probe Microscopy," Jpn. J. Appl. Phys. **35**, 4166 (1996).

[13] J. Bolten, T. Wahlbrink, M. Schmidt, H. Gottlob, and H. Kurz, "Implementation of electron beam grey scale lithography and proximity effect correction for silicon nanowire device fabrication," Microelectron. Eng. **88**, 1910 (2011).

[14] I. W. Cho, J.-M. Park, H. Kim, J.-Y. Hong, S.-S. Kim, H.-K. Cho, and H.-K. Oh, "Reduction of line width and edge roughness by resist reflow process for extreme ultra-violet lithography," Advances in Resist Materials and Processing Technology, 7273 (2009).

[15] X. Zhao, S.-Y. Lee, J. Choi, S.-H. Lee, I.-K. Shin, and C.-K. Jeon, "Minimization of line edge roughness and critical dimension error in electron-beam lithography," J. Vac. Sci. Technol. B **32**, 06F505 (2014).

[16] J. A. Croon, G. Storms, S. Winkelmeier, I. Pollentier, M. Ercken, S. Decoutere, W. Sansen, and H. E. Maes, “Line edge roughness: characterization, modeling and impact on device behavior,” in IEDM02 International Electron Devices Meeting, 307 (2002).

[17] C.H. Diaz, H. Tao, Y. Ku, A. Yen, K. Young, “An experimentally validated analytical model for gate line-edge roughness (LER) effects on technology scaling,” Electron Device Letters, IEEE , **22**, 287 (2001).

[18] C. A. Mack, “Stochastic approach to modeling photoresist development,” J. Vac. Sci. Technol. B **27**, 1122 (2009).

[19] C. A. Mack, “Stochastic modeling of photoresist development in two and three dimensions,” J. Micro/Nanolith. MEMS MOEMS **9**(4), 041202 (2010).

[20] C. A. Mack, A Simple Model of Line-Edge Roughness,” Future Fab International, **34**, 64 (2010).

[21] R. Guo, S.-Y. Lee, J. Choi, S.-H. Lee, I.-K. Shin, C.-U. Jeon, B.-G. Kim, and H.-K. Cho, J. Vac. Sci. Technol. B **31**, 06F408 (2013).

[22] Q. Dai, R. Guo, S.-Y. Lee, J. Choi, S.-H. Lee, I.-K. Shin, C.-U. Jeon, B.-G. Kim, and H.-K. Cho, “A fast path-based method for 3-D resist development simulation,” Microelectron. Eng. **127**, 86 (2014).

[23] S.-Y. Lee, Q. Dai, S.-H. Lee, B.-G. Kim, and H.-K. Cho, “Enhancement of spatial resolution in generating point spread functions by Monte Carlo simulation in electron-beam lithography,” J. Vac. Sci. Technol. B **29**, 06F902 (2011).

[24] K.-S. Chen, I-K. Lin, and F.-H. Ko, “Fabrication of 3D polymer microstructures using electron beam lithography and nanoimprinting technologies,” J. Micromech. Microeng. **15**, 1894 (2005).

[25] H.-M. Yeh, and K.-S. Chen, “Development of a Digital-Convolution-Based Process Emulator for Three-Dimensional Microstructure Fabrication Using Electron-Beam Lithography,” IEEE Trans. Industrial electronics. **56**, 926 (2009).

[26] W. Guo, and H. H. Sawin, “Review of profile and roughening simulation in microelectronics plasma etching,” Journal of Physics D: Applied Physics, **40**, 194014 (2009).

[27] L. Lallement, A. Rhallabi, C. Cardinaud, and M. C. P. Fernandez, “Modelling of fluorine based high density plasma for the etching of silica glasses,” J. Vac. Sci. Technol. A **29**, 051304 (2011).

[28] M. Kotera, K. Yagura, and H. Niu, “Dependence of linewidth and its edge roughness on electron beam exposure dose,” J. Vac. Sci. Technol. B **23**, 2775 (2005).

[29] J.A. Sethian, “A fast marching level set method for monotonically advancing fronts,” Proc. Natl. Acad. Sci. **96**, 1591 (1996).

[30] X. Zhao, S.-Y. Lee, J. Choi, S.-H. Lee, I.-K. Shin, C.-U. Jeon, B.-G. Kim, and H.-K. Cho, “Dependency analysis of line edge roughness in electron-beam lithography,” Microelectron. Eng. **133**, 78 (2015).

[31] T. H. P. Chang, Proximity effect in electron-beam lithography, J. Vac. Sci. Technol., **12**, 1271 (1975).

[32] K. Suzuki, J. Shears, and B. Smith, “Microlithography: Science and Technology,” (Marcel Dekker, New York, 1998).

[33] D. Drouin, A. R. Couture, D. Joly, X. Tastet, V. Aimez, and R. Gauvin, CASINO V2.42 - A fast and easy-to-use modeling tool for scanning electron microscopy and microanalysis users, Scanning **29**, 92 (2007).

[34] P. Tsiartas, L. Flanagin, C. Henderson, W. Hinsberg, I. Sanchez, R.Bonnecaze, and C. G. Willson, “The Mechanism of Phenolic Polymer Dissolution: A New Perspective,” Macromolecules **30**, 4656 (1997).

[35] M. K. Hu, ”Visual Pattern Recognition by Moment Invariants”, IRE Trans. Info. Theory **8**, 179 (1962).