1

Introduction

1.1 OPTICAL LITHOGRAPHY

Complex circuitries of modern microelectronic devices are created by building and wiring millions of transistors together. At the heart of this technology is optical lithography. Optical lithography technology is similar in concept to printing, which was invented more than 3000 years ago [92]. In optical lithography systems, a mask is used as the template, on which the target circuit patterns are carved. A light-sensitive polymer (photoresist) coated on the semiconductor wafer is used as the recording medium, on which the circuit patterns are projected. Light is used as the writing material, which is transmitted through the mask, thus optically projecting the circuit patterns from the mask to the wafer. The lithography steps are typically repeated 20–30 times to make up a circuit, where each underprinting pattern must be aligned to the previously formed patterns. After a lengthy lithography process, a complex integrated circuit (IC) structure is built from the interconnection of basic transistors. Moore's law, first addressed by Intel cofounder G. E. Moore in 1965, describes a long-term trend in the history of computing hardware. Moore's law predicted that the critical dimension (CD) of the IC would shrink by 30% every 2 years. This trend has continued for almost half a century and is not expected to stop for another decade at least. As the dimension of IC reduces following Moore's law, optical lithography has become a critical driving force behind microelectronics technology. During the past few decades, our contemporary society has been transformed by the dramatic increases in electronic functionality and lithography technology. Two main factors of optical lithography attract the attention of scientists and engineers. First, since lithography is the cardinal part of the IC fabrication process, around 30% of the cost of IC manufacturing is attributed to the lithography steps. Second, the advance and ultimate performance of lithography determine further advances of the critical size reduction in IC and thus transistor speed and silicon area. Both of the above aspects drive optical lithography into one of the most challenging places in current IC manufacturing technology. Current commercial optical lithography systems are able to image features smaller than 100 nm (about one-thousandth the thickness of human hair) of the IC pattern. As the dimension of features printed on the wafer continuously

COPYRIGHTED MATERIAL

Computational Lithography By Xu Ma and Gonzalo R. Arce
Copyright © 2010 John Wiley & Sons, Inc.

shrinks, the diffraction and interference effects of the light become very pronounced resulting in distortion and blurring of the circuit patterns projected on the wafer. The resolution limit of the optical lithography system is related to the wavelength of light and the structure of the entire imaging system. Due to the resolution limits of optical lithography systems, the electronics industry has relied on *resolution enhancement techniques* (RETs) to compensate and minimize mask distortions as they are projected onto semiconductor wafers. There are three RET techniques: optical proximity correction (OPC), phase-shifting masks (PSMs), and off-axis illumination (OAI). OPC methods add assisting subresolution features on the mask pattern to correct the distortion of the optical projection systems. PSM methods modify both the amplitude and phase of the mask patterns. OAI methods exploit various illumination configurations to enhance the resolution.

1.1.1 Optical Lithography and Integrated Circuits

Optical lithography is at the heart of integrated circuit manufacturing. Generally, three stages are involved in the IC creation process: design, fabrication, and testing [92]. The flow chart of the IC creation process is illustrated in Fig. 1.1.

First, the IC products are defined and designed. In this stage, the abstract functional units such as amplifiers, inverters, adders, flip-flops, and multiplexers are translated into physically connected elements such as metal-oxide-silicon (MOS) transistors. Subsequently, the design results of the physically connected elements are exploited in the second stage of fabrication, where the desired circuit patterns are carved on the masks, which are to be replicated onto the wafers through an optical lithography process. After a series of development processes applied to the exposed wafer such as etching, adding impurities, and so on, the ICs are packaged and tested for functional

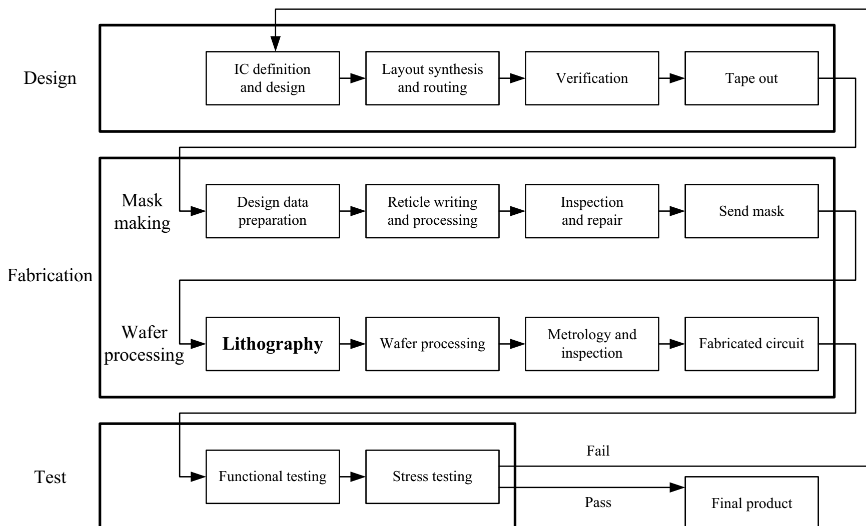

Figure 1.1 The flow chart of the IC creation process.

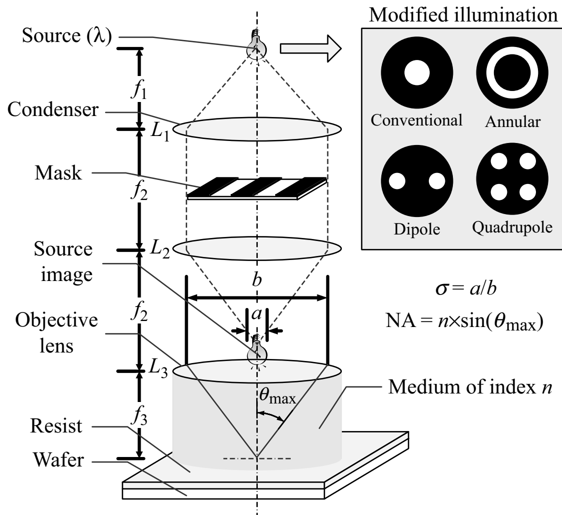

Figure 1.2 The scheme of a typical optical lithography system.

correctness and durability. During the entire IC creation process, optical lithography plays a significant role and is mainly responsible for the miniaturization of IC sizes.

Similar to printing, optical lithography uses light to print circuit patterns carried by the mask onto the wafer. The optical lithography system comprises four basic parts: an illumination system, a mask, an exposure system, and a wafer [92]. A typical optical lithography processing system is shown in Fig. 1.2. In Fig. 1.2, n is the diffraction index of the medium surrounding the lens. $\theta_{\max}$ is the maximum acceptable incident angle of the light exposed onto the wafer. The numerical aperture of the optical lithography system is defined as $\mathrm{NA} = n \sin \theta_{\max}$. The partial coherence factor $\sigma = \frac{a}{b}$ is defined as the ratio between the size of the source image and that of the pupil. Partial coherence factor measures the physical extent of the illumination. Larger partial coherence factor represents larger illumination and lower degree of coherence of the light source [92].

In the optical lithography process, the output pattern sought on the wafer is carved on the mask. Light emitted from the illumination system is transmitted through the mask, where the electric field is modulated by the transparent clear quartz areas and opaque chrome areas on the mask. Subsequently, the modulated electric field propagates through the exposure system and is finally projected onto the light-sensitive photoresist layer coated on the wafer, which is then partially dissolved by the solvents. The details of the photoresist processes and characteristics are discussed in Section 1.3.

1.1.2 Brief History of Optical Lithography Systems

Early optical lithography systems used contact lithography methods, where the mask is pressed against the photoresist-coated wafer during the exposure [11]. Since neither

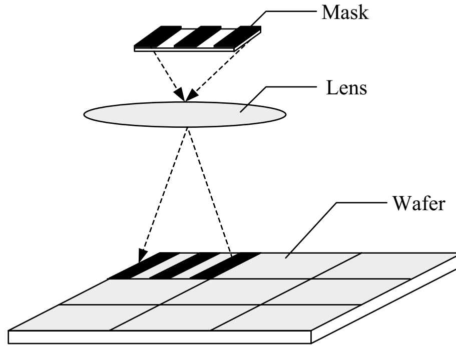

Figure 1.3 The configuration of wafer stepper.

the mask nor the wafer is perfectly flat, the hard contact method was used to push the mask into the wafer by applying a pressure ranging from 0.05 to 0.3 atm [11]. The advantage of contact lithography is that small features can be imaged using relatively cheap equipment. However, defects were generated on both wafer and mask due to the hard contact. In order to avoid defects, proximity lithography was introduced [11]. In proximity lithography, a gap, typically ranging from 10 to 50 μm, was maintained between mask and wafer. The primary disadvantage of proximity lithography is the resolution reduction due to the divergent light. Subsequently, projection lithography was developed so as to obtain high resolution without the defects associated with contact lithography. The projection optical lithography system is illustrated in Fig. 1.2. As shown in Fig. 1.3, in order to replicate a pattern onto a wafer of large scale, a wafer stepper is applied to repeat the lithography process for each small portion of the wafer. There are two configurations of wafer stepper: "step and repeat" and "step and scan" [4]. In the "step and repeat" configuration, the wafer is moved after each exposure until the total wafer has been exposed. Thus, the image size is limited by the largest size of lens field of sufficient imaging quality. In the "step and scan" configuration, the size of the lens field just covers a portion of the mask. The mask and wafer are scanned by the light source until the entire mask pattern is imaged on the wafer. In the past decade, optical lithography systems developed from deep ultraviolet lithography (DUVL) systems employing radiation with a wavelength of 300 nm to the ArF laser lithography system employing radiation with a wavelength of 193 nm.

As the critical dimension of the IC continuously reduces, lithography technology will be pushed into the next-generation lithography systems, including extreme ultraviolet lithography (EUVL), e-beam lithography (EBL), X-ray lithography, and ion-beam projection lithography [6]. Current research mainly concentrates on EUVL and EBL. In EUVL, radiation with a wavelength in the range from 10 to 14 nm is used to carry out projection imaging. In this wavelength range, the radiation is strongly absorbed in almost all materials. Thus, the EUVL systems must operate in near-vacuum environments. In addition, EUVL systems are entirely reflective, including masks [6]. On the other hand, EBL uses electron beam to directly write patterns on the wafer.

The primary advantage of the EBL is that it overcomes the diffraction limit of light. However, the major disadvantage is the low throughput.

1.2 RAYLEIGH'S RESOLUTION

The International Technology Roadmap for Semiconductors (ITRS) (2007 Edition), driven by Moore's law, shows the trend as depicted in Table 1.1 [1]. The critical dimension, which is the minimum feature size to be printed on the wafer, is limited by the Rayleigh's resolution [92].

According to the Fourier optics and the properties of lenses, the light energy passing through a mask forms a distribution in the pupil plane, which is proportional to the mask spectrum [25, 92]. The set of discrete spatial frequency of the mask pattern is referred to as the diffraction orders. Observing from the center of the mask, rays of low spatial frequency components travel with small angles, while those of high spatial frequency components travel with large angles. Therefore, lower frequency components pass closer to the center of the pupil. Higher frequency components are out of scope of the pupil and cannot be collected by the lens. Thus, the effect of the lens in the optical lithography system is equivalent to a low-pass filter, cutting off some high spatial frequency components of the mask pattern.

Consider the optical lithography system under coherent illumination. The spectrum of an isolated opening in the spatial domain is continuous such that some components of the signal spectrum always pass through the low-pass filter and image the pattern [4]. Figure 1.4 shows the imaging process of the isolated opening with a width of d.

Figure 1.5 shows the imaging process of a periodic pattern. The periodic patterns has discrete spatial frequency spectrum at intervals $\Delta k = \frac{2\pi}{p}$, where p is the period of the periodic pattern, referred to as the pitch. The periodic pattern shown in Fig. 1.5 depicts a pitch of p and a width of d. Let the diffraction index $n = 1$, the mth diffraction order diverges from the mask at an angle of θ_m, where

$$\sin\theta_m = m\frac{\lambda}{p}, \quad m = 0, \pm 1, \pm 2, \ldots, \tag{1.1}$$

Table 1.1 The International Technology Roadmap for Semiconductors

Technology Node	2009	2010	2011	2012	2013	2014
DRAM						
1/2 pitch (nm)	50	45	40	36	32	28
MPU						
1/2 pitch (nm)	52	45	40	36	32	28
Gate in resist (nm)	34	30	27	24	21	19
Physical gate length (nm)	20	18	16	14	13	11
Mask minimum features						
Nominal image size (nm)	135	120	107	95	85	76
Minimum primary feature size (nm)	94	84	75	67	59	53
Subresolution feature size (nm), opaque	67	60	54	48	42	38

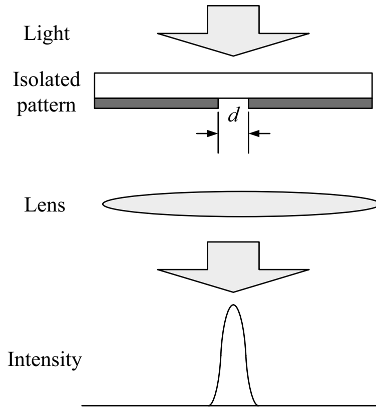

Figure 1.4 The imaging process of the isolated opening with a width of d.

where λ is the wavelength. It has been proven that at least two diffraction orders are needed to image the periodic pattern with distinguishable intensity variation [92].

According to Eq. (1.1), to allow the ± 1 diffraction orders to pass the low-pass filter,

$$\sin\theta_{\max} \geq \frac{\lambda}{p}. \tag{1.2}$$

Thus, the minimum distinguishable pitch is

$$p_{\min} = \frac{\lambda}{\sin\theta_{\max}} = \frac{\lambda}{\mathrm{NA}}. \tag{1.3}$$

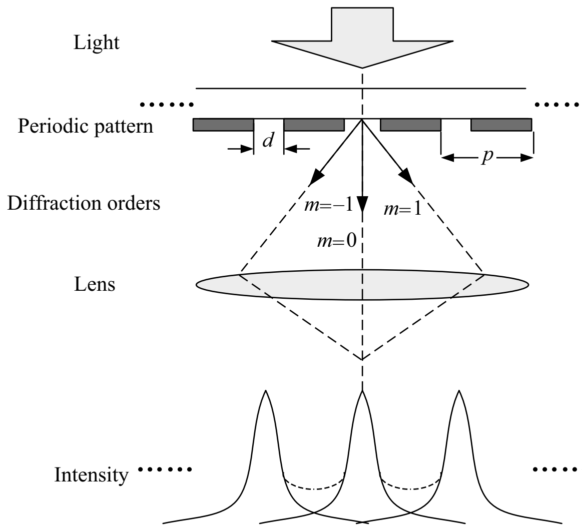

Figure 1.5 The imaging process of a periodic pattern with a pitch of p and a width of d.

In addition, the CD is defined as

$$\mathrm{CD} = \frac{p_{\min}}{2} = \frac{\lambda}{2\mathrm{NA}}. \tag{1.4}$$

In a partially coherent illumination system, the pitch and CD limits change as expected. Partially coherent illumination introduces the partial coherence factor $0 < \sigma < 1$ and thus enhances the resolution. Intuitively, the partially coherent illumination has nonzero line width. Thus, the partially coherent illumination system allows the ± 1 diffraction orders to pass through the low-pass filter; this is more difficult with a coherent illumination system. Therefore, the minimum distinguishable pitch in a partially coherent illumination system can be smaller than that in a coherent illumination system. In partially coherent illumination systems, the minimum distinguishable pitch is [92]

$$p_{\min} = \frac{1}{1+\sigma}\frac{\lambda}{\mathrm{NA}} \tag{1.5}$$

and

$$\mathrm{CD} = \frac{1}{1+\sigma}\frac{\lambda}{2\mathrm{NA}}. \tag{1.6}$$

When the dimension of the partially coherent illumination continuously extends, σ may become larger than 1. However, $\sigma > 1$ does not contribute to the enhancement of the resolution. Therefore, when $\sigma > 1$,

$$p_{\min} = \frac{\lambda}{2\mathrm{NA}} \tag{1.7}$$

and

$$\mathrm{CD} = \frac{\lambda}{4\mathrm{NA}}. \tag{1.8}$$

The above discussion of the Rayleigh's resolution takes only the diffraction effect into account. In order to incorporate the photoresist effect, resolution enhancement techniques, and so on, a process constant k is introduced to describe the comprehensive resolution limit:

$$R = \mathrm{CD} = k\frac{\lambda}{\mathrm{NA}}. \tag{1.9}$$

1.3 RESIST PROCESSES AND CHARACTERISTICS

Photoresist, or simply resist, is a photosensitive compound coated on the wafer, whose properties are changed by the impinging light radiation transmitted through the mask. The photoresist is first exposed in the optical lithography system. Subsequently, the wafer is immersed in a developer solution, then removed from the solution, rinsed off, and dried [11]. After the photoresist process and development, the mask pattern is replicated on the surface of the wafer. Photoresist used in IC fabrication normally consists of three components: a resin or base material, a photoactive compound (PAC),

and a solvent, which is used to control the mechanical properties, such as the viscosity of the base or keeping it in a liquid state [11].

There are two types of photoresist processes. These differ in the polarity. The thickness of the remaining photoresist after development is nonlinearly related to the exposure dose (referred to as the aerial image) exceeding a given threshold intensity. In a positive photoresist process, the PAC acts as an inhibitor before the exposure to reduce the dissolving rate of the photoresist when it is developed. Under exposure, chemical reaction occurs in the photoresist and changes the inhibitor to a sensitizer. Thus, the dissolving rate of the photoresist is increased. Because of these properties, almost all the photoresist material remains in the low-exposure areas on the wafer and is removed in the high-exposure areas. Between these two extremes is the transition region. The negative photoresist responds in the opposite manner. Positive photoresist tends to have the best resolution and is therefore much more popular for IC fabrication [11]. On the other hand, compared to positive photoresist, negative photoresist tends to exhibit better adhesion to various substrates such as Si, GaAs, InP, and glass, as well as metals, including Au, Cu, and Al. In addition, the current generation of G-, H-, and I-line negative photoresists exhibit higher temperature resistance over positive photoresists. The steps involved in a typical lithography process are shown in Fig. 1.6.

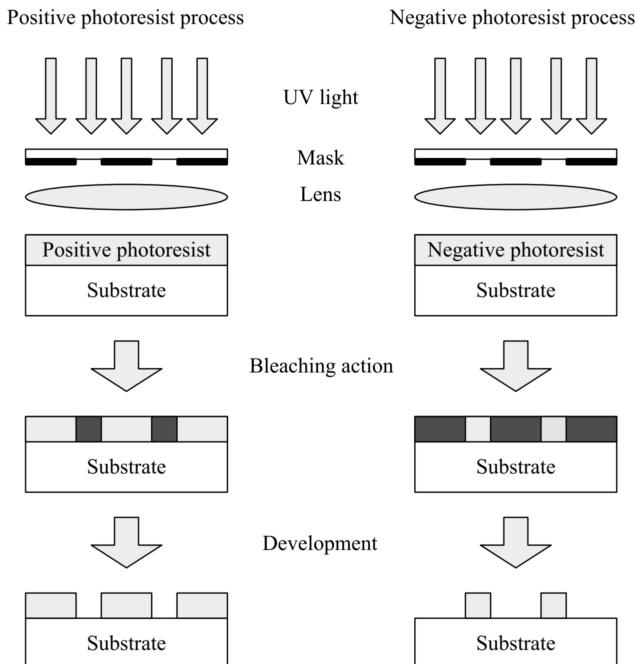

Figure 1.6 The steps involved in a typical optical lithography with positive photoresist process (left) and negative photoresist process (right).

Because of the photoresist absorption, the radiation intensity is decreased with increasing depth into the photoresist. The relationship between the radiation intensity I and the depth z is described by the logarithmic function as follows [11]:

$$I(z) = I_0 \mathrm{e}^{-\alpha z}, \tag{1.10}$$

where α is the optical absorption coefficient of the photoresist with unit of inverse of length. I_0 is the intensity on the top of the photoresist. The absorbance A is defined as

$$A = \frac{\int_0^T I_0 - I(z)\mathrm{d}z}{I_0 T} = 1 - \frac{1 - \mathrm{e}^{-\alpha T}}{\alpha T}, \tag{1.11}$$

where T is the thickness of the photoresist.

The positive photoresist is taken as an example to discuss the characteristics of photoresist in detail. For positive photoresist, the resist process can be characterized by the relationship between the thickness of the developed resist and the exposure dose for a fixed time. The exposure dose is defined as the light intensity multiplied by the exposure time. The plot of the normalized remaining thickness of resist versus the exposure dose is illustrated in Fig. 1.7. In Fig. 1.7, the x-axis represents exposure dose with unit mJ/cm^2 in logarithmic scale. The y-axis represents the normalized remaining thickness of resist in linear scale. The actual relationship is shown in solid line, which may be approximated by a piecewise linear curve shown in dashed line. There are three regions in the piecewise linear curve. Below the dose D_l is the low-exposure region. Beyond the dose D_h is the high-exposure region. Between the doses D_l and D_h is the transition region. Because of the logarithmic dependence on the exposure dose, the resist development is a nonlinear function of the dose. It is this nonlinearity that transforms sloped aerial images into relatively vertical photoresist profiles [92].

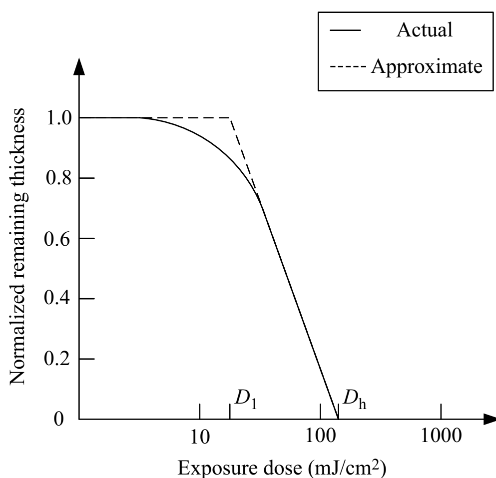

Figure 1.7 The plot of the normalized remaining thickness of resist versus the exposure dose.

The two extensively used metrics to measure the performance of the photoresist are the contrast and the critical modulation transfer function. The contrast is defined as

$$\gamma = \frac{1}{\log_{10}(D_h/D_l)}. \tag{1.12}$$

Thus, the contrast is just the slope of the piecewise curve shown in Fig. 1.7. The contrast depends on the resist material, the development process, the post-exposure bake processes, the wavelength of the light, the surface reflectivity of the wafer, and several other factors [11]. The lithographic image quality in general improves as the contrast is increased. In the ideal case, $D_l = D_h$; thus, $\gamma = \infty$. Therefore, the piecewise curve in Fig. 1.7 will reduce to a hard threshold function. Although this ideal limit cannot be reached in practice, the hard threshold approximation is used in the following chapters to simplify the photoresist model.

On the other hand, the modulation transfer function of an image is defined as

$$\mathrm{MTF} = \frac{I_{max} - I_{min}}{I_{max} + I_{min}}, \tag{1.13}$$

where I_{max} and I_{min} are the maximum and minimum intensities of the aerial image, respectively. The critical modulation transfer function (CMTF) is defined as

$$\mathrm{CMTF} = \frac{D_h - D_l}{D_h + D_l}, \tag{1.14}$$

where D_h and D_l are shown in Fig. 1.7. The CMTF can be explained as the approximate minimum MTF necessary to obtain a pattern [11]. In order to print an aerial image on the photoresist, the MTF of the aerial image must be larger than the CMTF of the resist.

1.4 TECHNIQUES IN COMPUTATIONAL LITHOGRAPHY

Due to the resolution limits of optical lithography systems, the electronics and photonics industry has relied on 2D and 3D resolution enhancement techniques to compensate and minimize mask distortions as they are projected onto semiconductor wafers [78, 92, 95]. According to Eq. (1.9), resolution in optical lithography obeys the Rayleigh's criterion: resolution$(R) = k\frac{\lambda}{\mathrm{NA}}$, where λ is the wavelength, NA is the numerical aperture taking on values around 0.9 for most lithography systems used today, and k is the process constant. In order to extend the limits of the resolution, the wavelength of the employed illumination is continuously reduced to enhance the resolution. On the other hand, NA is also increased by immersion lithography, where a liquid medium is filled in the space between the front lens and the photoresist. Different from these methods, RETs are applied to minimize the process constant k [11, 37, 76, 77]. RET methods manipulate the local amplitude and phase features of the optical wavefront to precompensate for image distortions. There are mainly three traditional kinds of techniques included in RETs such as optical proximity correction, phase-shifting masks, and off-axis illumination. In addition, second-generation RETs

have been proposed as the supplements of the traditional RETs. Second-generation RETs include a variety of techniques such as multiple mask exposure lithography, simultaneous source and mask optimization, photoresist tone reversing methods, and so on. The introduction of these approaches is described next.

1.4.1 Optical Proximity Correction

For low-k imaging, a sizable fraction of the transmitted light energy is concentrated in the high spatial frequency components of the mask spectrum. However, the low-pass filtering properties of the lens in the exposure systems cut off the high-frequency components and lead to image distortion. In general, there are four types of image distortion [92]. The first is the variation of the printed image under different environments with the same nominal critical dimension. The second type of distortion occurs when the changes of the nominal CD are not reflected linearly in the printed image. The third is the line shortening, and the final one is the corner rounding. In order to compensate for these image distortions, OPC methods modify the mask amplitude by the addition of subresolution features to the mask pattern such that the output patterns are as close to the desired pattern as possible. The typical scheme of optical proximity correction is shown in Fig. 1.8. The two types of OPC are rule-based approaches and model-based approaches. Rule-based approaches, which are simple to implement, can just compensate for the warping in local features. On the other hand, the model-based approaches used in this book rely on mathematical models to represent the image formation process of the optical lithography system and seek the global minimization of the cost function to improve the output pattern fidelity on the wafer. Model-based approaches include inverse and forward methods. In inverse methods, the optimization algorithms start from the desired output pattern and iteratively obtain the optimized layout. In the forward methods, the original layouts are continuously modified until the output patterns and the mask manufacturability properties are acceptable.

1.4.2 Phase-Shifting Masks

The application of the OPC methods encounters energy diffusing problems, where unwanted energy in the opaque (chrome) regions appears due to the close proximity of the neighboring transparent (quartz) features [69]. Figure 1.9 illustrates the imaging process of the OPC methods. In Fig. 1.9, the binary mask just includes clear area (quartz) and opaque area (chrome). Note that the energy diverges into the gaps between exposed areas on the wafer and reduces the resolution. Phase-shifting masks, commonly attributed to Levenson [35], induce phase shifts in the transmitted field that have a favorable constructive or destructive interference effect to remove the unwanted energy in the opaque regions. Three types of PSMs are extensively used in the IC fabrication industry: alternating phase-shifting masks, attenuated phase-shifting masks, and chromeless phase-shifting masks.

Alternating PSMs modulate the phases of the adjacent features on the mask by 180° out of phase with each other. The phase adjustment is implemented by the

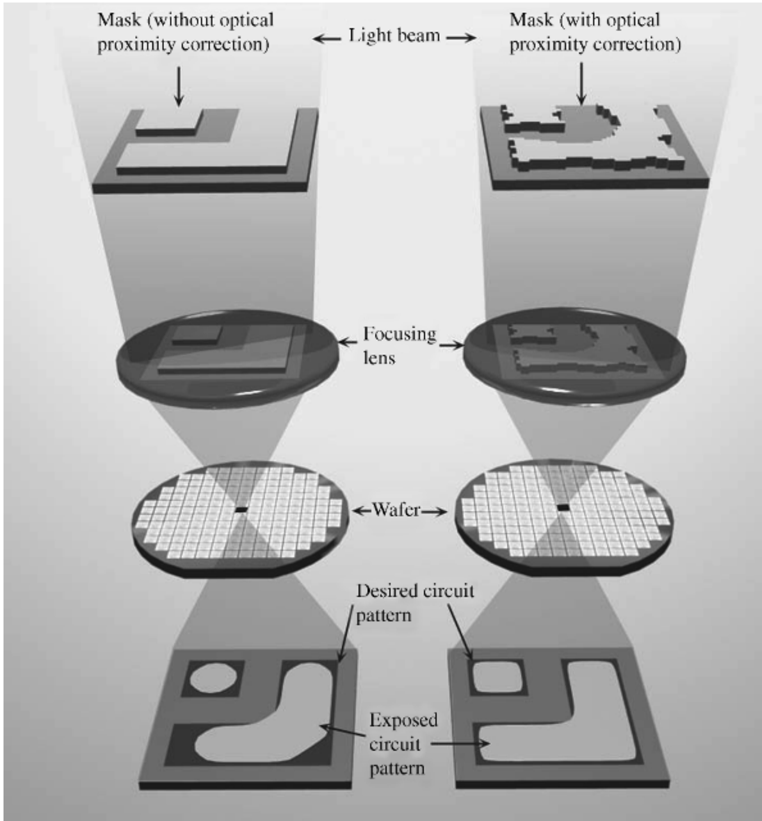

Figure 1.8 The typical scheme of optical proximity correction.

quartz etching. The features with 0° phase are referred to as clear areas, while shifting areas have 180° phase. Phase difference on the alternating PSM leads to destructive interference, removing the diffused energy in the opaque areas and resulting in better contrast and resolution of the printed image compared to binary masks. The imaging process of the alternating PSM is shown in Fig. 1.10.

The attenuated PSMs replace the chrome (opaque area) on the binary masks with molybdenum silicide (MoSi), through which the light partially transmits. The thickness of the MoSi layer introduces a phase shift of 180°. Unlike alternating PSMs where all transparent regions image onto the wafer, the background due to the partially transparent MoSi layer is not printed on the wafer [92]. Phase difference between the transparent regions and partially transparent regions leads to destructive interference. Compared to the alternating PSMs, attenuated PSMs achieve the process latitude improvement of sparse spaces such as an isolated contact without phase-shifting assisting features. However, the alternating PSMs outperform attenuated PSMs when printing narrow dark images [71, 92]. The imaging process of the attenuated PSM is shown in Fig. 1.11.

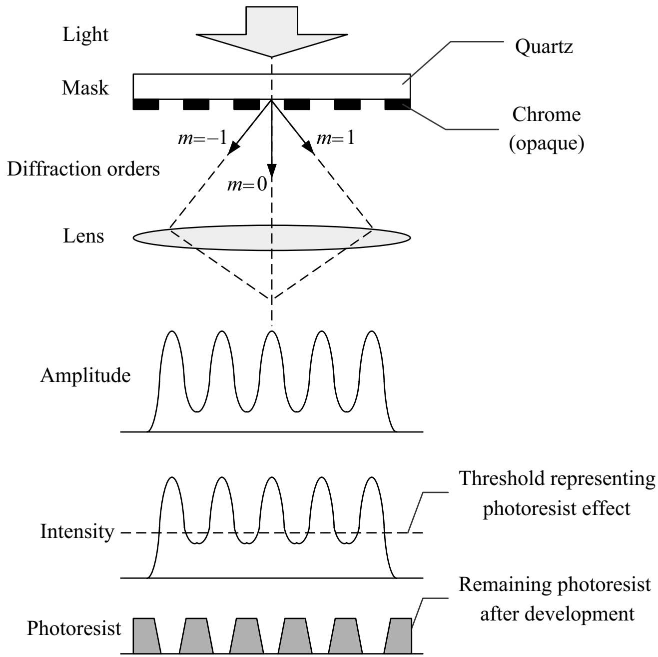

Figure 1.9 The imaging process of the OPC methods.

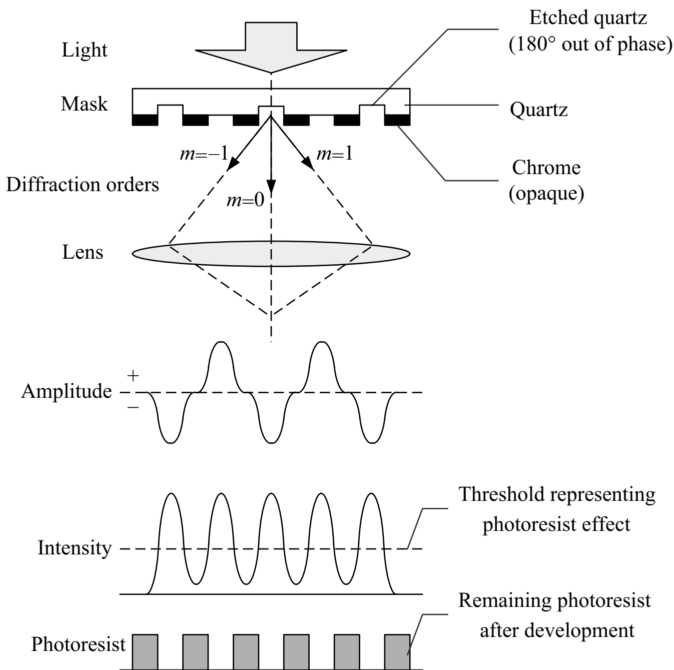

Figure 1.10 The imaging process of alternating PSM.

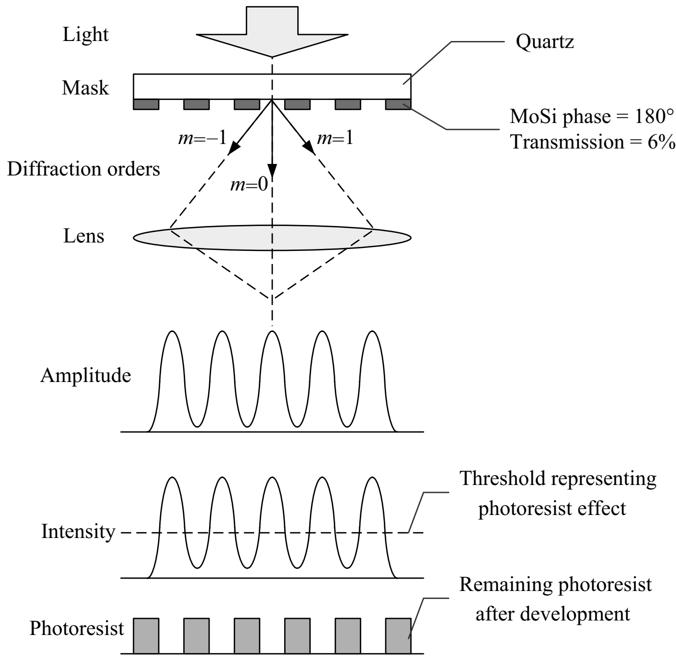

Figure 1.11 The imaging process of attenuated PSM.

Chromeless PSMs replace the MoSi layers of the attenuated PSMs with the etched transparent quartz layers, which preserve 180° phase. The interaction between quartz layers with 0° and 180° phases introduces destructive interference. The imaging process of the chromeless PSM is shown in Fig. 1.12.

1.4.3 Off-Axis Illumination

While OPC methods and PSMs modulate the amplitudes and phases of the features on the mask, off-axis illuminations modify the direction of the impinging light onto the wafer, thus influencing the diffraction orders captured by the lens. In the on-axis illumination system, all the 0 and ± 1 diffraction orders are collected by the lens. These collected light sources carry a lot of information of the background, rather than contributing to image formation. However, in OAI systems, just 0 and either one of ± 1 diffraction orders are collected. The diffraction between the two collected diffraction orders improves the imaging resolution. The OAI is implemented by designing the geometric pattern of the illumination. Common OAI configurations include dipole, quadrupole, and annular illuminations, among others [92]. The conventional illumination and a variety of OAIs are shown in Fig. 1.13. The conventional OAI methods derive the illumination pattern geometries by modifying the accepted diffraction orders by the lens to enhance the resolution limit and contrast. This book formulates the design of OAI pattern as an optimization problem, where both illumination and mask patterns are divided into pixels, each of which is optimized by gradient-based algorithms.

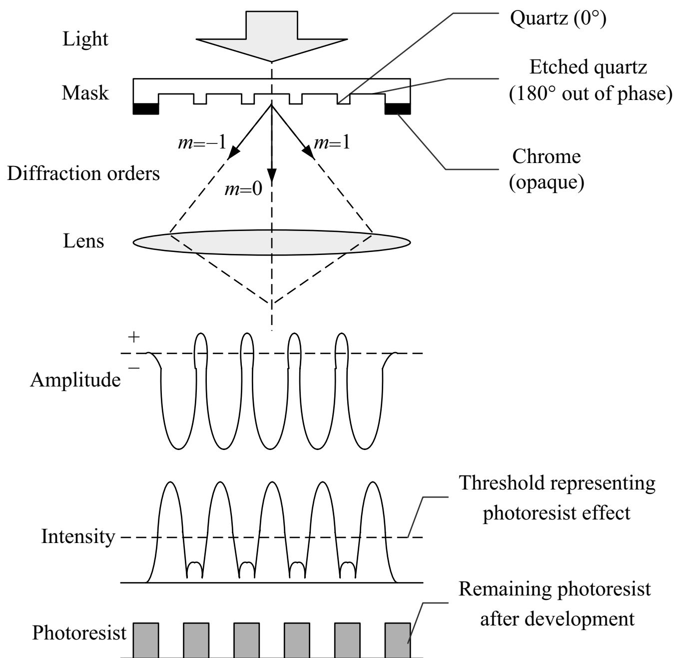

Figure 1.12 The imaging process of chromeless PSM.

1.4.4 Second-Generation RETs

During the past two decades, the continuous decrease in the critical dimension has motivated the development of second-generation RETs, which are not constrained in the range of the traditional RETs discussed above [8, 73, 84, 92]. Multiple mask exposure methods expose the coated photoresist layer several times with different mask

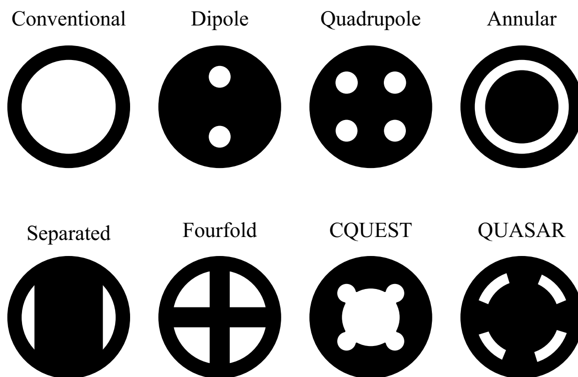

Figure 1.13 The conventional illumination and a variety of off-axis illuminations. *Top row* (from left to right): Conventional, dipole, quadrupole, and annular illuminations. *Bottom row* (from left to right): Separated, fourfold, CQUEST, and QUASAR illuminations.

patterns, capable of printing images even impossible by single exposure methods. In the multiple mask exposure methods, a dense circuit pattern is split into several relative sparse patterns. Masks are fabricated according to each of the sparse patterns and separately exposed on the wafer. There are two typical models of multiple mask exposure methods: double patterning lithography (DPL) and double exposure lithography (DEL). In DPL, two masks are exploited and each exposure is followed by its own etching process of the photoresist. On the other hand, DEL uses just one etching process after two exposures. Although multiple mask exposure methods are suitable for printing dense patterns, the drawback of these methods is the reduction of throughput.

Traditional RETs fix the illumination shape, thus limiting the degrees of freedom during the optimization of the mask pattern. In order to overcome this restriction, the simultaneous source and mask optimization methods have been developed recently where the illumination configuration and the mask pattern are designed simultaneously. The resulting source and mask patterns fall well outside the realm of known design forms. Usually, the optimized illumination is far from the OAIs discussed in Section 1.4.3.

Another new RET method is that of photoresist tone reversing, which exploits both positive and negative photoresist materials on the wafer and improves the lithography performance of small spaces in the output pattern. In addition to these three approaches, there are numerous other potential second-generation RETs, which may be found in relative literature.

1.5 OUTLINE

The organization of the book is as follows.

In Chapter 2, Abbe's formulation of the partially coherent imaging system is first summarized. As the simplified version of Abbe's formulation, the Hopkins diffraction model is used to represent the optical lithography system. Subsequently, three kinds of decompositions of the Hopkins diffraction model are discussed. First is the Fourier series expansion model, where the partially coherent imaging system is represented as the sum of several coherent systems. The accuracy of the Fourier series expansion model is the same as the direct discretization of the Hopkins diffraction model. Subsequently, two kinds of approximation models, referred to as the average coherent approximation model and the singular value decomposition (SVD) model, are summarized and used to reduce the computational complexity of the Fourier series expansion model. As the two limits of the partially coherent imaging system, coherent and incoherent imaging systems are discussed at the end of Chapter 2.

In Chapter 3, the rule-based RETs are described. First, the RET approaches are classified into rule-based, model-based, and hybrid RET approaches. Subsequently, the chapter focuses on the rule-based RETs, where rule-based OPC, PSM, and OAI approaches are described in detail. The rule-based OPC includes catastrophic OPC, one-dimensional OPC, line-shortening reduction OPC, and two-dimensional OPC. The rule-based PSM includes dark-field application and light-field application. For

all these rule-based RET methods, the rules to modify the masks and illuminations are summarized.

In Chapter 4, the fundamentals of optimization are discussed. First, the definition and the classification of different optimization problems are summarized. According to the classification, the inverse lithography optimization can be transformed to a continuous, unconstrained, nonlinear, and deterministic problem. Subsequently, the unconstrained optimization problems are discussed in detail. The methods to recognize minimizers are presented as several theorems. Two strategies, such as line search strategy and trust region strategy, are discussed to solve the unconstrained optimization problem. Particularly, the line search strategy includes the steepest descent method, the Newton method, the quasi-Newton method, and the conjugate gradient method. The trust region strategy includes the dogleg method, the two-dimensional subspace minimization method, and so on. In the following chapters, the steepest descent algorithm is applied to the gradient-based inverse lithography optimization.

In Chapter 5, the OPC and PSM optimizations for inverse lithography is developed under coherent imaging systems. The forward imaging process of the optical lithography systems is approximated as the Hopkins diffraction model followed by a sigmoid function. The Hopkins diffraction model represents the formation process of the aerial image. The sigmoid function represents the photoresist effect. The MSE between the desired pattern and the output pattern after the photoresist process is used as the cost function for the mask optimization. Based on this model, OPC and two-phase PSM optimization algorithms are developed. Next, generalized gradient-based PSM optimization methods are developed. These generalized algorithms provide highly effective four-phase PSMs capable of generating mask patterns with arbitrary Manhattan geometries.

In Chapter 6, a set of regularization frameworks for the OPC and PSM optimizations is discussed. The pole penalty is developed to reduce the pattern errors resulting from the discretization of the amplitude and phase of the optimized complex-valued mask. In order to influence the solution patterns to have more desirable manufacturability properties, a wavelet penalty is introduced. The wavelet penalty offers more localized flexibility than total variation penalty, which is traditionally employed in inverse problems. Furthermore, the comparison between wavelet penalty and total variation penalty is discussed.

In Chapter 7, OPC optimization approaches are developed for the partially coherent imaging systems based on the Fourier series expansion model. In order to reduce the computational complexity, the average coherent approximation model is applied to develop effective and more computationally efficient OPC optimization algorithms for inverse lithography. The advantages and the disadvantages of both algorithms are discussed and analyzed. Subsequently, the SVD model is used to develop computationally efficient PSM optimization algorithms under partially coherent illuminations for inverse lithography. These PSM optimization algorithms are most effective with small to medium partial coherence factors.

In Chapter 8, a variety of techniques to improve the performance of OPC and PSM optimizations are described. A double patterning optimization method for general inverse lithography is described where each exposure uses an optimized two-phase

mask. Furthermore, a novel DCT post-processing is derived to reduce the mask complexity and the output pattern error by cutting off the high-frequency components of the optimized masks in DCT domain. Finally, a photoresist tone reversing technique is exploited to improve the resolution limit.

In Chapter 9, simultaneous source and mask optimization (SMO) algorithms are described for both OPC and PSM designs. In this chapter, the SOCS model was first applied to decompose the partially coherent imaging systems. Then, the simultaneous source and mask design was formulated as an optimization problem, where the cost function was the square of the l^2-norm of the difference between the desired output pattern and the aerial image. Cost sensitivity was calculated and applied to drive the cost function in the descent direction during the optimization process. In order to influence the solution patterns to have more desirable manufacturability properties, topological constraints are added to the optimization framework.

In Chapter 10, the thick-mask effects are taken into account, as the CD printed on the wafer shrinks into the subwavelength regime, and the mask topography is considered as a 3D object. The OPC and PSM optimization methods are developed based on the boundary layer (BL) model to compensate for the thick-mask effects. In these algorithms, the model-based lithography methods are exploited to obtain the desired binary and phase-shifting masks.

In Chapter 11, the contributions of this book are concluded and the new future directions of RETs are outlined. A software guide for the accompanying Matlab codes is included in Appendix H.