© 2012 THOSS Media & **DE GRUYTER** DOI 10.1515/aot-2012-0028 — Adv. Opt. Techn. 2012; 1(4): 237–248

Tutorial

Donis G. Flagello and Daniel G. Smith*

Calculation and uses of the lithographic aerial image

Abstract: Beginning with the seminal Dill papers of 1975, the aerial image has been essential for understanding the process of microlithography. From the aerial image, we can predict the performance of a given lithographic process in terms of depth of focus, exposure latitude, etc. As lithographic technologies improved, reaching smaller and smaller printed features, the sophistication of aerial image calculations has had to increase from simple incoherent imaging theory, to partial coherence, polarization effects, thin film effects at the resist, thick mask effects, and so on. This tutorial provides an overview and semi-historical development of the aerial image calculation and then provides a review of some of the various ways in which the aerial image is typically used to estimate the performance of the lithographic process.

Keywords: aerial image; image simulation.

***Corresponding author: Daniel G. Smith,** Nikon Research Corp. of America, 12490 N Rancho Vistoso Blvd. Ste 130, Tucson, AZ 85737, USA, e-mail: dsmith@nikonrca.com
Donis G. Flagello: Nikon Research Corp. of America, 12490 N Rancho Vistoso Blvd. Ste 130, Tucson, AZ 85737, USA

1 Introduction

The early days of microlithography developed more as an art than a science. Lithographers made do with rudimentary approximations of imaging and an incomplete understanding of resist processes. This changed in 1975, when Dill's seminal papers [1–5] ushered in an evolution of microlithography as a science [6, 7]. In these papers, a complete theory of resist behavior was developed, new instruments were used to characterize the fundamental properties of resist, and the beginnings of a more complete imaging theory were created using the modulation transfer function (MTF), which would give us what came to be known as the aerial image – in film, we would call this the film image or, in water, the aquarial image.

Although the MTF was useful as a first approximation in some cases, the problems with the MTF were severalfold. The MTF treats sinusoidal objects, whereas lithographers need to image an array of shapes such as edges, lines, spaces, elbows, contact holes, and isolated features. Of course, any object can be decomposed into its Fourier components, to which the MTF applies [8, 9]. However, the MTF assumes incoherent imaging, whereas lithographers never actually use totally incoherent imaging; even a pupil filling ratio of $\sigma=1$ is not truly incoherent because light from outside the aperture is not there to contribute to the dark-field image. Finally, the MTF provides only modulation, whereas lithographers are more concerned with the image slope and the resultant size of the pattern in photoresist.

Two basic formulations of partially coherent lithographic image simulation emerged: the Abbe formulation and the Hopkins formulation. Following the Dill papers, in the 1970s and early 1980s, authors such as Kintner [10], Subramanian [11], Matsumoto [12], Cole [13], and others adapted the Hopkins formulation [14], with computers of that day, to the problem of efficiently calculating partially coherent aerial images. The central component to this formulation is the transmission cross coefficient, which we will see could be computed in advance for a specific illumination condition and used over and over for different mask patterns. The Abbe formulation is also included in the popular formulation. While this method has its origins in Abbe's imaging theory [15], it is really a collection of many separate contributions, including those of Debye [16] and Hopkins. This formulation is conceptually more tractable and more easily programmed; however, the application of the theory to computational algorithms is often computationally demanding.

2 Computing the aerial image

The basic picture of a simple lithographic imaging system is illustrated in Figure 1, where the imaging is projected to infinity by a condenser lens, and the reticle image is relayed to the wafer telecentrically with an afocal projection lens having a magnification equal to the ratio of focal lengths $\frac{f_2}{f_1}$. As is nearly always done, the illumination is

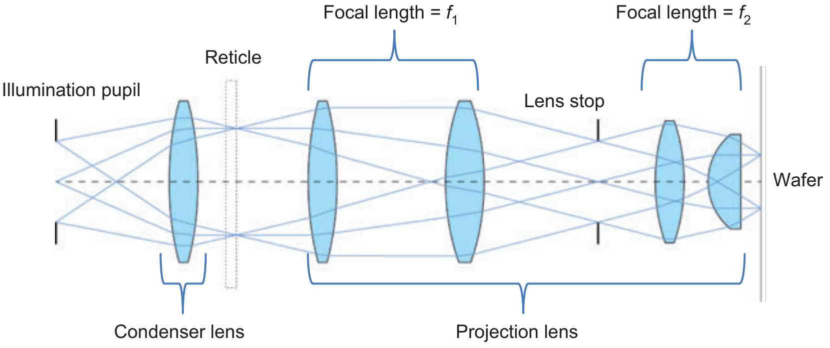


Figure 1 The lithographic imaging system is generally assumed to consist of a Kohler illumination and a double telecentric projection lens. In this way, each point in the illumination pupil can be represented as a plane wave in reticle space, and each plane wave transmitted by the reticle can be treated as a point source in the imaging pupil.

considered to be Koehler so that each point in the illumination pupil is manifested as a plane wave arriving at the reticle.

While Figure 1 shows a description of a simple optical system used for lithography, the derivations for aerial image formulation usually use a 'black-box'-type of optical system. In both cases, the system is assumed to have Koehler illumination with a double telecentric imaging system. However, in the 'black-box' case, there is no lens detail. Instead, the imaging model concentrates on conjugate planes given by the reticle/wafer plane and entrance/exit pupils combinations. Figure 2 illustrates this, where lenses are just given as block components without exact ray tracing. The entrance and exit pupils are replaced by focal spheres with radii of f_1 and f_2, respectively. The illumination source is also shown imaged into the projection lens.

Because the Hopkins and Abbe image formulations are equivalent, it is possible to obtain one from the other by a change of variables and an order of integration. We will start with the more conceptually tractable Abbe formulation and then derive the Hopkins formulation to show how they are related. The basic concept behind the Abbe formulation is that the image intensity can be thought of as the sum of intensities generated by light from independent coherent source points. The derivation given here is based on previous works [17, 18] with scalar imaging concepts laid out by Goodman [19] combined with a direction cosine spectrum description of the propagation of the scalar field given by Harvey [20] and Hopkins. Previously, Hopkins showed that the entrance pupil surface exists as the far-field distribution of the reticle (the object), while the image plane exists as the far-filed distribution of the image. Therefore, we can use the techniques of Huygens-Fresnel diffraction [21] and Fourier relationships to model the diffraction from the reticle to the entrance pupil and the diffraction from the exit pupil to the image plane. The propagation from the entrance to the exit pupil is further given by a conjugate mapping that represents the action of the lens with the appropriate obliquity terms due to the conservation of power requirements.

We will start with a generalized propagation of any point source from the illumination pupil to the reticle and then to the entrance pupil of the imaging lens. Let each point on the entrance pupil of the projection lens be represented by a vector in the direction cosine notation, i.e., $\bar{\rho}=\alpha\hat{i}+\beta\hat{j}$. For a circular pupil, we can define the object space numerical aperture by

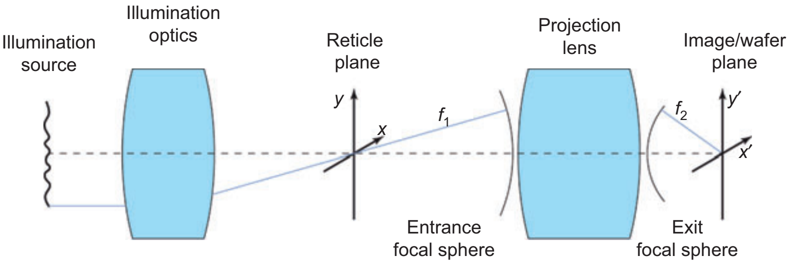


Figure 2 'Black-box' model of the lithographic imaging system, showing a Koehler illumination of the reticle and the illumination source imaged into the entrance pupil of the projection lens. The imaging action of the lens is shown by the entrance and exit pupils.

$$NA=n\sin\theta_{\max}=n\sqrt{\alpha_{\max}^2+\beta_{\max}^2}, \tag{1}$$

where $\theta_{\max}$ is the maximum marginal ray angle defined in the object or reticle space, and n is the index of refraction, which is assumed to be 1.

As we have Koehler illumination, where the source is imaged into the entrance pupil, and as the source can be considered as a distribution of point sources, we define an effective source distribution $J(\vec\rho_s)$ in the entrance pupil using a vector notation to denote each source point as $\vec\rho_S=\alpha_S\hat{i}+\beta_S\hat{j}$. Using the Huygens-Fresnel principle, the electric field from the reticle to the entrance pupil can be given as,

$$A(\vec\rho;\vec\rho_S)=A_0\gamma\iint O(\vec r)e^{i2\pi(\vec r\cdot\vec\rho)}e^{-i2\pi(\vec r\cdot\vec\rho_S)}d\vec r=A_0\gamma\tilde{O}(\vec\rho-\vec\rho_S). \tag{2}$$

For simplicity, all spatial coordinates and distances are normalized to the vacuum wavelength, λ. Hence, the vector $\vec r$ describes the normalized spatial coordinate position on the reticle given as $\vec r=(x\hat{i}+y\hat{j})/\lambda$. γ is a direction cosine defined by $\gamma=\sqrt{1-(\alpha^2+\beta^2)}=\cos\theta$ and can be considered as an obliquity term. $O(\vec r)$ is the reticle pattern of interest. Fourier transform is indicated by a tilde so that $\tilde{O}(\vec\rho-\vec\rho_s)$ is the angular spectrum of the object, which is shifted by $\vec\rho_S$, the angular coordinate of the illuminating point source. We note that implicit in this description is that the imaging system is shift invariant, and the implications of this will be discussed later in this section.

The field at the image is given by the propagation of the field distribution at the exit pupil. Modifying a previous work [17] to include a general source point and an image side index of refraction results in,

$$U'(\vec r';\vec\rho'_S,z')=U_0{}'F^{-1}\left\{\frac{A'(\vec\rho';\vec\rho'_S)e^{-i2\pi n'\gamma'z'}}{n'\gamma'}\right\}, \tag{3}$$

where the term in the brackets is an inverse Fourier transform. The primed notation denotes either the image (wafer) region or the exit pupil region. The phase term in *n′z′* has been added as a wafer focus term, with the convergence of the amplitude at *z′*=0.

From the entrance pupil, the field is mapped to the exit pupil via the operation of the lens, which truncates, scales, and modifies the field (such as in the case of aberrations and/or apodization). The image side numerical aperture is defined by,

$$NA'=n'\sin\theta'_{\max}=n'\sqrt{\alpha'^2_{\max}+\beta'^2_{\max}}, \tag{4}$$

Typically, in lithography, *n′* is either 1 or approximately 1.44 for water at a vacuum wavelength of 193 nm. The system magnification can be defined, using the Abbe sine condition, as

$$m=\frac{h'}{h}=\frac{NA}{NA'}, \tag{5}$$

Where *h* and *h′* are the object and image heights, respectively, and with the sine condition being defined by,

$$nh\sin\theta=n'h'\sin\theta'. \tag{6}$$

Rosenbluth [22] has detailed a complete discussion of the radiometric terms and subsequent obliquity factors that are necessary in a treatment of image simulation. Radiometric consistency must be insured when mapping from the entrance to the exit pupils, i.e., power conservation must be ensured between the differential areas from the entrance to the exit pupil. Therefore, the following condition must be met,

$$\left|A'(\vec\rho';\vec\rho'_S)\right|^2 da'=\left|A(\vec\rho-\vec\rho_S)\right|^2 da, \tag{7}$$

and the differential areas are given as,

$$da=f_1^2\frac{d\alpha d\beta}{\gamma}\text{ and }da'=f_2^2n'\frac{d\alpha'd\beta'}{\gamma'}, \tag{8}$$

where we have assumed that the object side refractive index is 1. The complete exit pupil distribution is obtained by using a lens pupil function to represent the action of the lens. For perfect circularly symmetric lenses, this is a circular top-hat function with a maximum diameter of 2*NA′*. Hence, when magnification scaling is also considered in conjunction with the power conservation condition, the exit pupil distribution can be shown to be,

$$\begin{aligned}A'(\vec\rho';\vec\rho'_S)&=\sqrt{n'\gamma'\gamma}\cdot A(m\vec\rho'-m\vec\rho'_S)\cdot P(\vec\rho')\\&=A'_0\sqrt{n'\gamma'\gamma}\tilde{O}(m\vec\rho'-m\vec\rho'_S)\cdot P(\vec\rho').\end{aligned} \tag{9}$$

Substituting this into Eq. (3) gives the final field at the wafer for one general source point,

$$U'(\vec r';\vec\rho_S,z')=U'_0F^{-1}\left\{\sqrt{\frac{\gamma}{n'\gamma'}}\cdot\tilde{O}(m\vec\rho'-m\vec\rho'_S)\cdot P(\vec\rho')e^{-i2\pi n'\gamma'z'}\right\}. \tag{10}$$

The partially coherent aerial image at a given focus is then found by summing all the intensities across the effective source and is given by,

$$I(\vec r';z')=\iint J(\vec\rho'_S)\left\|U'(\vec r';\vec\rho'_S,z')\right\|^2 d\vec\rho'_S. \tag{11}$$

The Hopkins formulation is found by expanding this equation and rearranging the order of integration. We simplify the procedure by dropping the leading constant complex factor U_0' in Eq. (10), which has no consequence in the intensity calculation. We also let m=1. The magnification (or reduction) and the subsequent obliquity factors are easily brought in as scaling and pupil apodization factors during the computational implementations of either the Abbe or Hopkins formalisms. Hence, Eq. (11) becomes,

$$\begin{aligned} I(\vec{r}') &= \iint J(\vec{\rho}_S')\left\|U'(\vec{r}';\vec{\rho}_S')\right\|^2 d\vec{\rho}_S' \\ &= \iint J(\vec{\rho}_S')\left\|F^{-1}\left\{\tilde{O}(\vec{\rho}'-\vec{\rho}_S')\cdot P(\vec{\rho}')\right\}\right\|^2 d\vec{\rho}_S' \\ &= \iint J(\vec{\rho}_S')\left\{\iint \tilde{O}(\vec{\rho}_1'-\vec{\rho}_S')\cdot P(\vec{\rho}_1')e^{-i2\pi(\vec{r}'\cdot\vec{\rho}_1')}d\vec{\rho}_1'\right\} \\ &\quad \times\left\{\iint \tilde{O}^*(\vec{\rho}_2'-\vec{\rho}_S')\cdot P^*(\vec{\rho}_2')e^{i2\pi(\vec{r}'\cdot\vec{\rho}_2')}d\vec{\rho}_2'\right\}d\vec{\rho}_S'. \end{aligned} \tag{12}$$

By a change of variables, the form of the integral terms in curly brackets can be written as,

$$\begin{aligned} &\iint \tilde{O}(\vec{\rho}_n'-\vec{\rho}_S')\cdot P(\vec{\rho}_n')e^{-i2\pi(\vec{r}'\cdot\vec{\rho}_1')}d\vec{\rho}_n' \\ &\quad = \iint \tilde{O}(\vec{\rho}_n')\cdot P(\vec{\rho}_n'+\vec{\rho}_S')e^{-i2\pi(\vec{r}'\cdot\vec{\rho}_1')}d\vec{\rho}_n'. \end{aligned} \tag{13}$$

Using the form of Eq. (13) into Eq. (12) gives the Hopkins formulation,

$$\begin{aligned} I(\vec{r}') &= \iint J(\vec{\rho}_S')\left\{\iint \tilde{O}(\vec{\rho}_1')\cdot P(\vec{\rho}_1'+\vec{\rho}_S')e^{-i2\pi(\vec{r}'\cdot\vec{\rho}_1')}d\vec{\rho}_1'\right\} \\ &\quad \times\left\{\iint \tilde{O}^*(\vec{\rho}_2')\cdot P^*(\vec{\rho}_2'+\vec{\rho}_S')e^{i2\pi(\vec{r}'\cdot\vec{\rho}_2')}d\vec{\rho}_2'\right\}d\vec{\rho}_S' \\ &= \iint\iint\iint \tilde{O}(\vec{\rho}_1')\tilde{O}^*(\vec{\rho}_2')e^{i2\pi\vec{r}'\cdot(\vec{\rho}_2'-\vec{\rho}_1')}P(\vec{\rho}_1'+\vec{\rho}_S') \\ &\quad P^*(\vec{\rho}_2'+\vec{\rho}_S')J(\vec{\rho}_S')d\vec{\rho}_1'd\vec{\rho}_2'd\vec{\rho}_S' \\ &= \iint\iint \tilde{O}(\vec{\rho}_1')\tilde{O}^*(\vec{\rho}_2')e^{i2\pi\vec{r}'\cdot(\vec{\rho}_2'-\vec{\rho}_1')}\tilde{C}(\vec{\rho}_1',\vec{\rho}_2')d\vec{\rho}_1'd\vec{\rho}_2'. \end{aligned} \tag{14}$$

The last term is called the *C-function* [23] or the *transmission cross coefficient* and is given as,

$$\tilde{C}(\vec{\rho}_1',\vec{\rho}_2') = \iint P(\vec{\rho}_1'+\vec{\rho}_S')P^*(\vec{\rho}_2'+\vec{\rho}_S')J(\vec{\rho}_S')d\vec{\rho}_S'. \tag{15}$$

The transmission cross coefficient, or just TCC, can be considered the hallmark of the Hopkins formulation. The TCC is a cross correlation between the imaging pupil P, its complex conjugate P^*, and the effective source J, and as such, it is a four-dimensional function of two positions within the pupil. The integrand of the TCC is commonly described with a diagram such as that shown in Figure 3 where the integrand is nonzero only in the region of overlap between the three functions. In some simple cases, the TCC can have a closed form that can be derived geometrically. In other cases, it can be computed numerically and stored in memory for latter calculations of the aerial image.

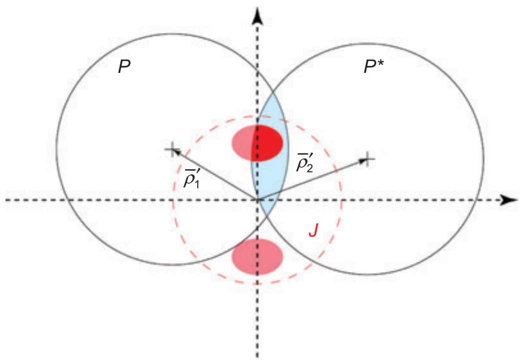


Figure 3 This is a graphical representation of the TCC in the plane of $\vec{\rho}$. The effective source J is represented in red as a dipole illumination pattern bordered by a dashed circle. The pupil function P and its complex conjugate P^* are represented by two closed circles in black offset by $\vec{\rho}_1'$ and $\vec{\rho}_2'$, respectively. The integral is then performed over the entire plane, but the integrand in nonzero only in the region of overlap between all three functions.

The key difference between the Hopkins and Abbe formulations is the order in which the integrals are performed. In the Hopkins formulation, the integral over the source and the pupil are performed first through the TCC. The Abbe formulation, also known as the 'sum of sources' approach, takes the integral over the source last. Another viewpoint is that the Abbe formulation presents a lens pupil-centered approach, while the Hopkins mathematics derives an illumination source-centered approach. Because of these differences, the formulations are well suited to different applications.

Whenever imaging is desired with fixed source and fixed pupil constraints, the Hopkins formulation can be more efficient. This occurs because the TCC can be precalculated. This is commonly done in optical proximity correction (OPC) and mask optimization today. However, when the source constraints need to be variable, as in a source optimization (SO) procedure or a simultaneous source mask optimization (SMO), the Abbe formulation is often used. Also, many consider the Abbe formulation easier to understand and implement. This formulation is usually used when a general purpose algorithm is needed for aerial image simulation.

Until the numerical aperture (NA) of the imaging system reached the range of 0.5–0.6, the basic scalar formulations of Hopkins or Abbe imaging outlined above worked very well and made excellent predictions of aerial images. The propagation of the aerial image field into the photoresist and subsequent films was often approximated using a simple Beer's law [24],

$$I(\vec{\rho}\,';z)=I(\vec{\rho}\,')T_{resist}e^{-\alpha z}, \tag{16}$$

where α is the photoresist absorption and T_{resist} is the transmission of the photoresist/air interface. The equation here only assumes one photoresist film with no reflection; however, more elaborate approximations, such as Mack [25], were able to predict film images with effects such as standing waves.

In the drive to smaller feature sizes, larger NAs forced us to include the effect of polarization and the vector nature of light. Flagello [17, 18] showed how to combine the Hopkins and Abbe theories with the thin film theory [26] and polarization decomposition [27] to obtain a vector theory of image formation in thin films such as photoresist. The basic idea is that the field in the image space is decomposed into plane wave components when the illumination source is polarized. These field components can be treated separately using the classic thin film theory to determine the fields at any position within the resist by deriving a film term as a function of $\vec{\rho}\,'$ and z. The field components are added vectorally to produce the irradiance within the film or the so-called *film image*. Unfortunately, the term *aerial image* is still being used today for reference to the *aquarial image* or the *film image*.

Another assumption implicit in the above development is that the mask is a so-called thin mask or that the Kirchhoff approximation holds [19]. This assumption came into a serious question with the advent of phase shifting masks. For the features at the wafer plane with dimensions $<\lambda/2$, the Kirchhoff approximation must be treated carefully to avoid substantial errors. To include the effects of the thick mask, electromagnetic magnetic field (EMF) treatments, such as finite difference time domain (FDTD) [28], finite element methods [29], waveguide methods [30], or rigorous coupled wave analysis (RCWA) [31] have been used to estimate the transmitted near field, which can then be decomposed into a spectrum of plane waves and propagated to the wafer by the Abbe formulation. Unfortunately, the EMF treatments can be very costly computationally, so care must be taken with respect to how many times the thick-mask transmitted field is computed. For cases where the shift invariance can be invoked, we can compute the object spectrum using rigorous EMF methods with only one angle of illumination and obtain the others by a simple shift of its spectrum around in the pupil. However, when the transmitted field varies significantly over the range of illumination, and there is shift-invariance, then, the rigorous mask transmittance must be computed for a larger subset and then interpolated to a finer pupil grid.

The Abbe formulation shown here assumes a generalized form for the index of refraction of the imaging medium. Specifically, the direction cosine vector in the image/exit pupil space must be defined by,

$$\vec{\rho}\,'=n'\alpha'\hat{i}+n'\beta'\hat{i} \quad \text{and} \quad n'\gamma'=n'\sqrt{1\text{-}\left(\alpha'^2+\beta'^2\right)}. \tag{17}$$

Historically, the aerial image formulations for lithographic applications were derived assuming both image and object spaces had an index of refraction of unity. As such, many of the available aerial image simulators did not use a definition such as Eq. (17) nor could they handle indices other than 1 for the medium between the lens and the wafer. This seriously impacted the ability to obtain a valid film image. Also, water immersion lithography required that the aerial image algorithms had to be modified because the incident medium had changed. The procedure was outlined by Mulkens [32] in 2004, where three variable substitutions in a pure aerial image algorithm allowed for accurate aquarial and film image formation, where in this case the ‘hat’ indicates a normalized quantity rather than the usual unit vector as was used above:

- Change the numerical aperture to ensure that the values are greater than 1 using $N\hat{A}={}^{NA}\!/_{n'}$.
- Create an effective wavelength that scales with the index of refraction, i.e., $\hat{\lambda}={}^{\lambda}\!/_{n'}$.
- The refractive index of each film, n_j, on the wafer, including the photoresist film, is normalized to the index of refraction of the immersion medium. Hence, $\hat{n}_j={}^{\hat{n}_j}\!/_{n'}$.

While current commercial simulators have corrected their algorithms, these scaling substitutions give some insight into some of the advantages of immersion. From (1) we see that the increase in *NA* is directly proportional to the index. From (2), we notice that the effective wavelength for a 193-nm lithographic immersion exposure tool will have $\hat{\lambda}=193\ \text{nm}/1.44=134\ \text{nm}$. Therefore, the use of immersion is similar to a decrease in wavelength. This is significant, as decreasing the wavelength below 193 nm requires substantial development costs, while the use of water with existing systems has proven a cost-effective solution for resolution improvement. Finally, from (3), we see that the ratio of indices will get smaller with immersion. The implication of this is that the top surface of the photoresist/water interface will have a much lower reflectivity. This can result in more power transmitted into the photoresist resulting in a faster exposure throughput and lower, probably, for scattered radiation to affect the image.

3 Making use of the aerial image

Aerial images must have a medium to interact with to be utilized. By themselves, they are only a mathematical construct, as the physical act of measuring the aerial image, whether it is a sensor, a photosensitive medium, or even the human eye, will interact with its properties. However, the creation of aerial images into a photoresist is necessary to create a pattern or stencil after a chemical development. Optical lithography uses two basic components for imaging: a lens system and the photoresist. Although, it is the final film image that is used by the photoresist, we can gain some understanding of the many interacted imaging mechanisms if we consider the lens system as 'presenting' the aerial image (or aquarial image for immersion) to the photoresist. This is then modified to form a film image that is subsequently used for further processing such as post exposure baking and chemical development. In this section, we attempt to examine some of these interactions with the photoresist, which requires the creation of models for the photoresist response to the image and to understand the metrics necessary to predict the performance of the developed image in a photoresist film.

3.1 Normalization and bulk photoresist response

One of the critical aspects in understanding how the aerial image interacts with the photoresist is to understand the initial exposure normalization and response. Exposure (often called dose or exposure dose) into a photoresist is defined as,

$$E = I \cdot t, \tag{18}$$

where t is the exposure time in seconds, and I is the irradiance in mW/cm^2 resulting in the exposure, E, in units of mJ/cm^2. Typically, the irradiance of the aerial image is normalized such that an open frame pattern, i.e., a 100% transmission clear area produces an aerial image irradiance of $\hat{I}(x,y)=1.0$, where we have used a carrot to denote normalization, and we have dropped the prime notation for the image plane for simplification. This normalization requires the angular spectrum of the object or reticle patterns, shown in Eq. (2), to be scaled by the transmission. In practice, to achieve a specific distribution of exposure due to the aerial image, we would multiply the normalized irradiance by a nominal exposure, E_n, which has an implicit time dependence. Hence, we can write the aerial image exposure as,

$$E(x,y) = \hat{I}(x,y) \cdot E_n \ . \tag{19}$$

The outcome of this is that the representation of the exposure mechanism is a multiplicative scaling of the aerial image. This becomes advantageous when formulating models of photoresist behavior.

Figure 4 shows an example of the bulk response of the positive photoresist to an open (100% transmission reticle) exposure to actinic radiation. The photoresist has been coated on a substrate with an absolute thickness, T_0. For a substantial amount of exposure, the normalized resist thickness changes relatively slowly. However, at an inflection exposure, E_i, the photoresist starts to lose film thickness at a more rapid rate, with a slope of γ_{resist}, until the point where no photoresist is remaining on the substrate. This exposure is termed the 'clearing dose' or E_0. An ideal photoresist would have $\gamma_{\text{resist}}=\infty$ and $T=T_0$ up to the point E_i with $E_i=E_0$.

3.2 Combining photoresist response and aerial images

A simple model for the aerial images uses the bulk response of the photoresist as the primary mechanism of development. While only a first-order approximation, its simplicity gives insight into the role of the aerial image in lithographic performance. The model assumes that if we have an aerial image exposure into the photoresist, as given by Eq. (19), then for all positions where $E(x,y) \geq E_0$, the photoresist will be totally developed. For a perfect resist, for all exposures where $E(x,y)<E_0$, the photoresist will have a thickness comparable to the initial thickness. The width of the patterned features is, therefore, determined

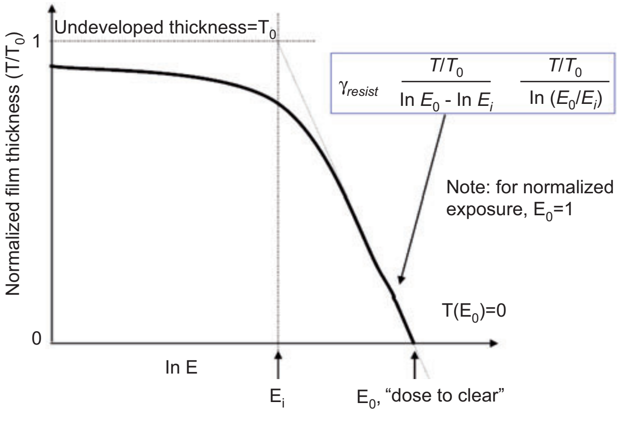


Figure 4 Photoresist response to exposure.

at the E_0 positions, and the photoresist structure will have vertical sidewalls. This is called the *threshold* model of photoresists and does not consider the finite thickness of the resist, but only the positions where $E(x,y) \geq E_0$.

Figure 5A examines the threshold model in further detail, where we assume that $E_0=1.0$ and $E(x,y)=\hat{I}(x,y)$. The red curve is the outcome of an aerial image simulation of a 300-nm isolated space. If this distribution is multiplied by 3.125, as given by the blue curve, the width of the photoresist space will exactly be 300 nm. This corresponds to an irradiance threshold of 0.32 on the original curve. Figure 5B shows the predicted resist structure with straight sidewalls. We can define a critical dimension (CD) as the measure size of the critical pattern. In this case, it is our target space size of 300 nm. The effective nominal exposure for the threshold model is just the reciprocal of the irradiance threshold. We can write this as,

$$E_n = \frac{1}{I_{\text{threshold}}} \cdot E_0. \qquad (20)$$

With a less than perfect resist, $\gamma_{\text{resist}} << \infty$, the sidewall structures can be sloped. A model of this was developed in the 1980s and is called the 'tone reproduction model' of the photoresist [33]. It is a slight modification of the threshold model to include more effects of the aerial image into the finite thickness of the resist. Figure 6 is an example of the model using experimental photoresist data. The aerial image is of a 1-μm isolated line model at $\lambda=365$ nm with NA=0.42. The exposed image is essentially mapped into the photoresist pattern using the bulk response function of the photoresist. The size of the photoresist line is clearly defined by the E_0 positions and Eq. (20).

More advanced photoresist models usually use the full film image. The assumption is that the higher-order effects such as the higher NA image interaction with the photoresist films, diffusion of acids and solvents, and a full treatment of the photoresist development requires a full dimensional approach to modeling. The limited dimensional approach of the simple threshold model cannot accurately explain the effects that occur in the depth of the photoresist especially when the image formation is due to the larger angles of incidence into the photoresist. Figure 7 shows a comparison of the aerial image, the film image, and a fully developed image of a 5-bar 0.25-mm pattern using $\lambda=248$ nm exposure and an NA=0.57. A simple threshold model would only output the widths of the 5 bars when only examining the aerial image. This may be a sufficient image model depending on the application. However, if a higher degree of accuracy is needed, the film image is necessary to model the depth structure. When this is used in a more rigorous photoresist model, we see that the prediction of the photoresist images show some rounding at the top and some slight asymmetry between the center and outer bars.

3.3 Applications of aerial image models with photoresist

The aerial image, in conjunction with a model of photoresist, can be used to explore and study a variety of

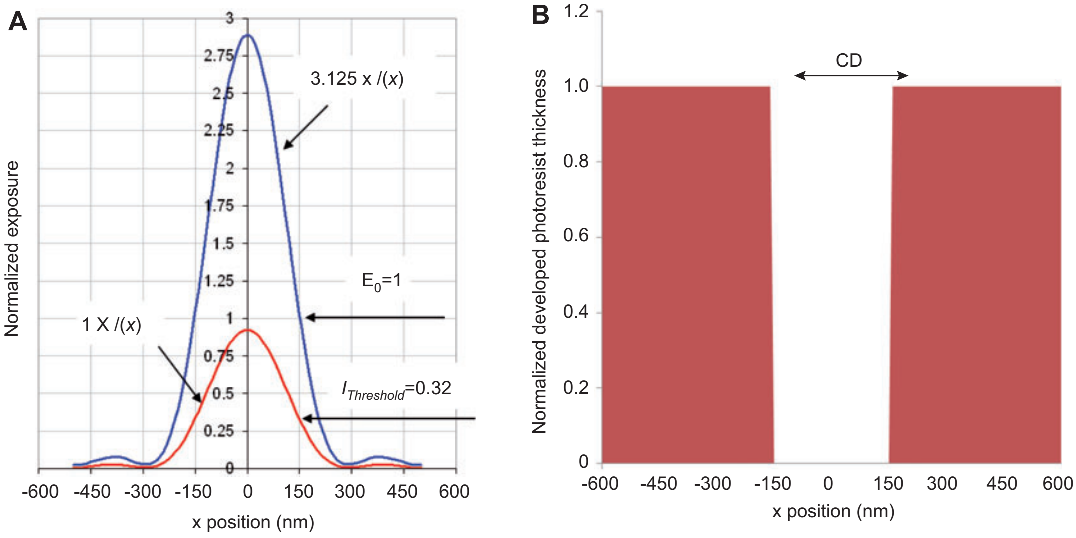


Figure 5 (A) Exposed aerial images of a 300-nm isolated space pattern, showing two different exposures and the location of the threshold size of the space. (B) The predicted photoresist structure using a threshold model, where the CD defines the width of the space.

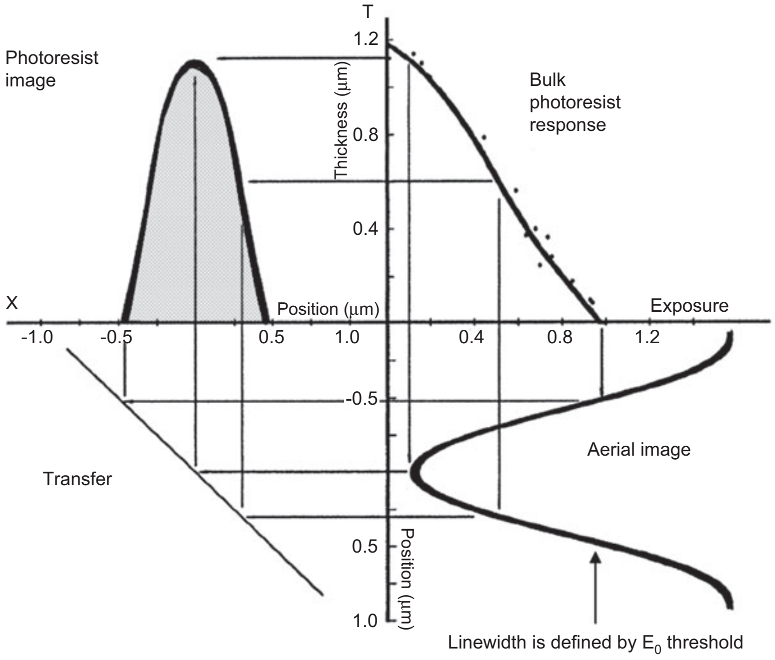


Figure 6 The ‘tone reproduction model’ of the photoresist showing the mapping of the aerial image exposure to the photoresist image via the bulk photoresist response function.

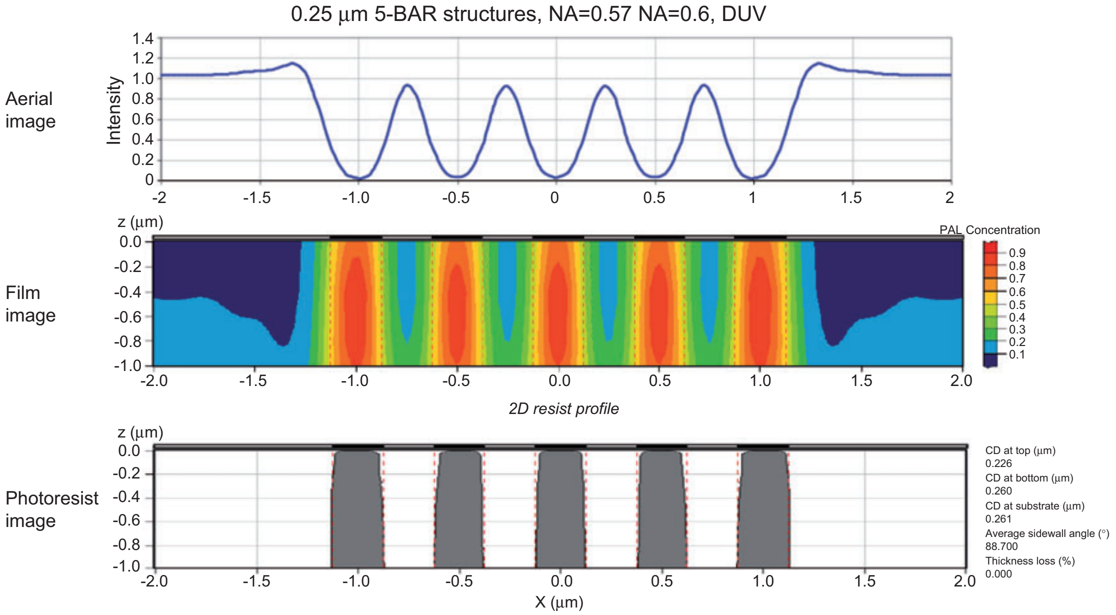


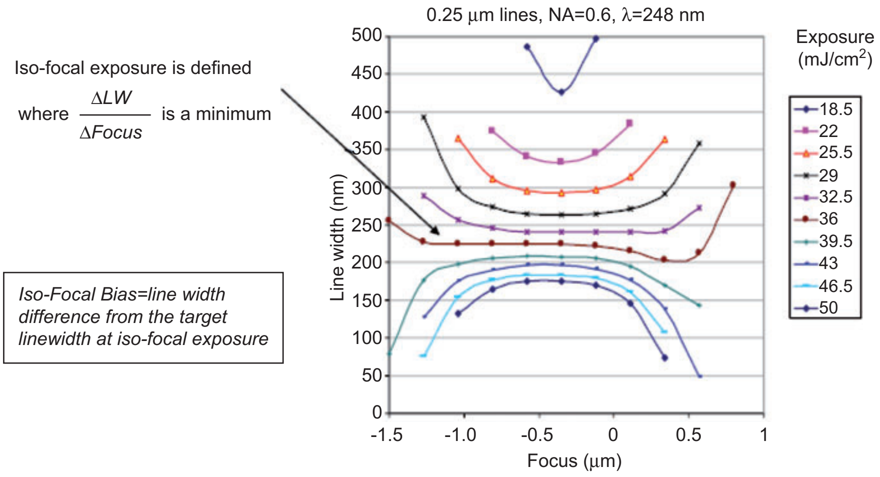
Figure 7 Construction of the photoresist image from the aerial image, to the film image, to the printed photoresist image.

parameters and conditions of either the optical system in question or the photoresist process. The behaviors of the developed photoresist to various amounts of defocus, exposure, lens aberrations, illumination source coherence, photoresist thickness, etc., are just a few of the parameters that have been explored. Once the photoresist image has been constructed from the aerial image, the CDs are usually calculated for a given pattern and then plotted against various other parameters. A historic example of this is the Bossung curve (named after John Bossung of Perkin

Elmer in the 1970s). This is a permutation where the line width is plotted against focus, for incremental amounts of exposure. Figure 8 illustrates a typical Bossung curve and demonstrates the optimal, isofocal exposure, where the line width varies minimally with changing focus.

Another useful way to evaluate the imaging performance from the aerial image is to plot iso-CD contours against dose and focus, the so-called E-D window [34], where the contours of the target CD, plus and minus some tolerance (typically ±10%), are plotted, leaving a gap in between. The red rectangle (sometimes an ellipse is used instead) that fits entirely within this gap has a height equal to the exposure latitude (EL) and a width equal to the depth of focus (DOF). There is a family of such rectangles (or ellipses) that fit within this gap, and so the family of the EL and DOF can be plotted against one another to provide what is known as the process window – exposure latitude vs. depth of focus. The examples of the E-D window and the process window are illustrated in Figure 9.

Another use of the aerial image is to examine the differences in the critical dimension that can result from the proximity of other features or the type of feature. This 'optical proximity effect,' can be caused by illumination coherence, aberrations, scattered light, photoresist effects, etch, and other processes. It is primarily thought of as having most to do with diffraction and coherence of the source. A simple way of characterizing this is to plot the width of specific features with some nominal width or CD, at a given focus and exposure, for a range of pitches. The plot of this optical proximity effect (OPE) illustrates how the photoresist width will vary in the transition from equal lines and spaces, to isolated (or nearly isolated) structures. An example of the OPE curve, or sometimes the 'coherence curve' [35], is illustrated in Figure 10, where a 100-nm line



Iso-focal exposure is defined where $\frac{\Delta LW}{\Delta Focus}$ is a minimum

Iso-Focal Bias=line width difference from the target linewidth at iso-focal exposure

Figure 8 This figure illustrates the concept of the Bossung curve and the isofocal exposure.

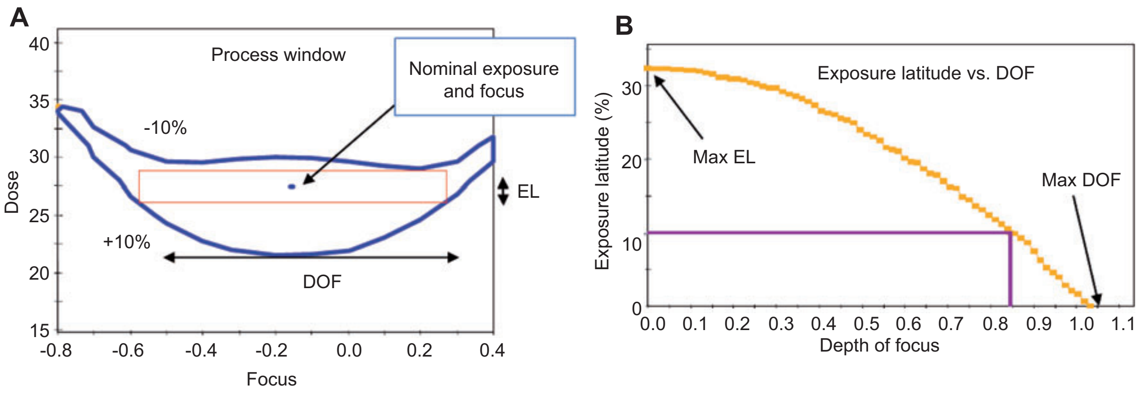


Figure 9 The left plot shows an example of an E-D window, where the contours of the target CD +10% and -10% are plotted against dose and focus. The gap in between the two contours can be inscribed by a family of rectangles (or ellipses) whose height is the EL and whose width is the DOF. The right plot, known as the process window, is a plot of EL vs. DOF.

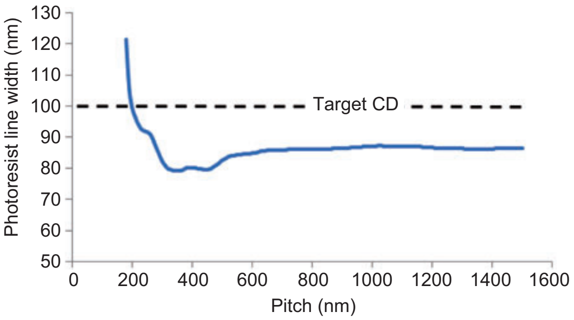


Figure 10 This figure is an example of the OPE curve, or coherence curve, where the printed line width is plotted against the pitch for a given focus and exposure, with the dense lines on the left and the increasing line isolation to the right.

has been simulated using λ=193 nm with annular illumination and a simple threshold model. The effect in this example is mainly driven by the annular illumination, which tends to be only optimized within a finite range of pitches. Outside of this range, the aerial image requires a different nominal exposure-to-size to the target CD.

Finally, we can use aerial images to examine the effect of reticle or mask errors. As an illustration, if we assume an error in a mask feature by, say, 4-nm, then with an optical system having a reduction factor (m=1/4) of 4, the error in the resist should be 1 nm. However, because of diffraction, coherence, and all the other subtle effects of lithographic imaging, there can be a strong nonlinearity between the object to image mapping. The effect is described by a factor called the *mask error enhancement factor* (MEEF), and we can define it by,

$$MEEF=\frac{\partial CD_{resist}}{\partial CD_{mask}}. \tag{21}$$

The MEEF can be seen as creating an effective optical magnification and reduction such that

$$m_{\text{effective}}=m\cdot MEEF \text{ and } R_{\text{effective}}=R^{-1}\cdot MEEF. \tag{22}$$

The aerial image with a simple photoresist model can be used to directly estimate the MEEF and to understand some of the mechanisms that cause it. Figure 11A shows the result of a simulation of dense lines (50% duty cycle) as a function mask CD (in wafer scale units). The NA=0.7 and λ=258 nm using a circular source with a lens pupil filling ratio of σ=0.5. We also plot ±10% CD lines for reference. The photoresist exposure has been targeted for the 150-nm lines. Ideally, we want to target the exposure at one CD setting and have all the other feature sizes fall within the ±10% zone. However, the strong nonlinearity is immediately noticed. This is especially evident for line sizes below 150 nm. The main cause is that most of the diffracted electric field from the reticle is not being captured by the lens due to an insufficient NA. We expect the MEEF to very high in this line size region. Figure 11B illustrates this where the MEEF has been calculated for an NA of 0.7 and 0.8. At 130 nm, the MEEF is approximately 15 for NA=0.7 compared to a MEEF of 2.6 for NA=0.8.

We can also gain an understanding of the fundamental mechanism of how MEEF arises by directly looking at the aerial image. Figure 12 examines the aerial image for a mask CD of 130 nm using the parameters for Figure 11. It is clearly evident that the high NA has a greater dynamic

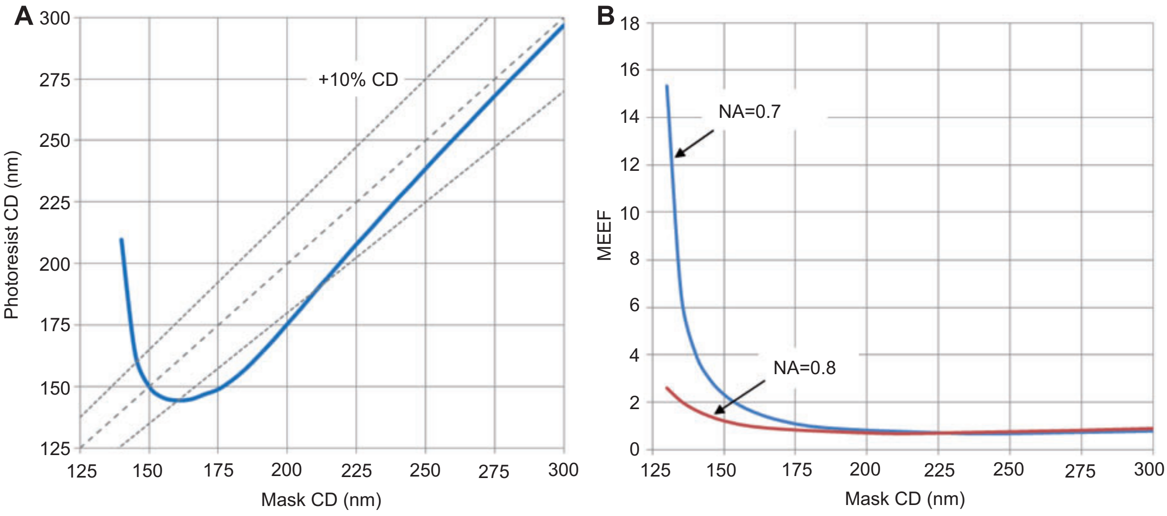


Figure 11 (A) The linearity curve for the dense lines. The mask CDs are in the wafer scale units. The photoresist exposure has been targeted for a mask CD=150 nm. (B) MEEF as a function of dense lines for NA=0.7 and NA=0.8.

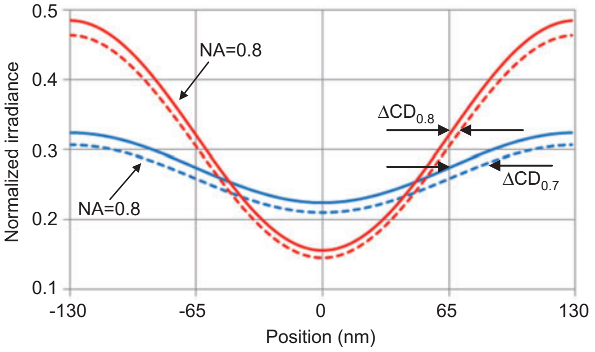


Figure 12 Aerial images of the 130-nm line (solid) and the 134-nm line (dashed). The higher NA has less line width change.

range, i.e., contrast, as well as a steeper image slope at the threshold levels at the CD positions, which is ±65 nm. Therefore, any small change in the size of the aerial image would affect NA=0.7 with a much greater sensitivity. This translates into the higher MEEF that is seen.

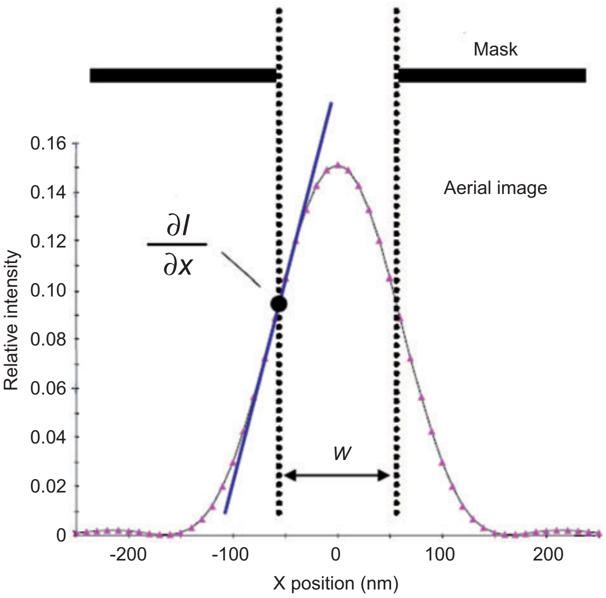


Figure 13 Illustration of the image slope.

3.4 Image log slope

In many fields outside of lithography, contrast is a common metric for image sharpness. However, contrast only takes into account the maximum and minimum intensities, which generally fall above and below the threshold, and are positioned some distance from the transition that determines the sharpness of a feature edge. For this reason, the lithographers often think in terms of the slope of the intensity in the vicinity of the transition from dark to light as illustrated in Figure 13 because the greater the image slope, the sharper the transition and the sharper the feature edge.

Because the magnitude of the image slope scales with intensity, it is important to normalize the image slope by the target exposure value. Furthermore, it is also useful to normalize by the target line size so that changes in the edge position are expressed as a fraction of the line size. The resulting unitless metric is known as the normalized image log-slope or NILS [36], as expressed in Eq. (23).

$$NILS=\frac{w}{I}\frac{\partial I}{\partial x}=w\frac{\partial \ln I}{\partial x}. \tag{23}$$

For a given process, there is a minimum NILS required to print a feature at the target CD. As small changes in the exposure latitude have a linear effect on the NILS, Mack [36] has shown that we can usually apply Eq. (23) to extrapolate the exposure latitude from a few experimental sample points. The constants *A* and *B* are determined from the experimental data.

$$\%EL=A(NILS\text{-}B). \tag{24}$$

Defocus impacts the NILS by blurring the imaging and reducing the slope. However, there is a minimum NILS required to print a feature at the target CD; below that value, nothing will print. Today's processes require the NILS to be above approximately 1.5–2.0, so that although the NILS can be estimated from the aerial image below the cutoff, the NILS less than the minimum correspond to the images outside the depth of focus.

4 Summary

The need for rigorous methods in calculating the lithographic aerial image became apparent with the advent of the Dill papers in 1975. Since then, there has been a steady increase in the complexity and precision of the mathematical methods keeping pace with the increased resolution requirements demanded by Moore's law. The aerial image calculations now include the effects of the custom partially coherent illumination, thick masks, thin films on the wafer, polarization effects, and the vector nature of light. With these aerial image calculations, lithographers have a variety of ways to estimate the imaging performance of their process, by automatic computation of Bossung curves, process windows, OPE curves, MEEF, and more.

Received June 19, 2012; accepted August 13, 2012

References

[1] F. H. Dill, SPIE 5754, 377–382 (2005).
[2] F. H. Dill, IEEE Trans. Electron Devices, ED-22, 440–444 (1975).
[3] F. H. Dill, W. Hornberger, P. Hauge and J. Shaw, IEEE Trans. Electron Devices, ED-22, 445–453 (1975).
[4] K. L. Konnerth and F. H. Dill, IEEE Trans. Electron Devices, ED-22, 453–456 (1975).
[5] F. H. Dill, A. R. Neureuther, J. A. Tuttle and E. J. Walker, IEEE Trans. Electron Devices, ED-22, 456–464 (1975).
[6] C. A. Mack, Proc. SPIE 5754 (2004).
[7] A. R. Neureuther, Proc. SPIE 6924 (2008).
[8] R. Barakat and S. Lerman, App. Optics 6, 545–548 (1967).
[9] P. K. Modal and S. Slansky, Optica Acta 17, 779 (1970).
[10] E. C. Kintner, App. Opt. 17, 2747–2753 (1978).
[11] S. Subramanian, App. Opt. 20, 1854–1857 (1981).
[12] K. Matsumoto, K. Konno and K. Ushida, in 'Development and application of photolithography simulation program for step-and-repeat projection system', Proc. Kodak Microelectronics Seminar, 74–79 (Nov. 14–15, 1983).
[13] D. C. Cole, E. Barouch, U. Hollerbach and S. A. Orszag, Jpn. J. Appl. Phys. 31, 4110–4119 (1992).
[14] H. H. Hopkins, App. Opt. 20, 1854–1857 (1981).
[15] M. Born and E. Wolf, in 'Principles of Optics', 7th edition, (Pergamon, New York, 1999) pp. 476–472.
[16] P. Debye, Ann. Phys. Lpz. 30, 755–776 (1909).
[17] D. G. Flagello and A. E. Rosenbluth, J. Vac. Sci. Technol. B 10, 2997–3003 (1992).
[18] D. G. Flagello, T. Milster and A. E. Rosenbluth, J. Opt. Soc. Am. A 13, 53–64 (1996).
[19] J. W. Goodman, in 'Introduction to Fourier Optics' (McGraw-Hill, San Francisco, 1968).
[20] J. E. Harvey, Am. J. Phys. 47, 974–980 (1979).
[21] Max Born and Emil Wolf, in 'Principles of Optics', 7th edition, Chapter 8 (Pergamon, New York, 1999).
[22] A. E. Rosenbluth, J. T. Azpiroz, K. Lai, K. Tian, D. O. S. Melville, M. et al., Proc. SPIE 6924, 69240V (2008).
[23] D. S. Goodman, 'Stationary Optical Projectors', PhD Dissertation (University of Arizona, 1979).
[24] Max Born and Emil Wolf, in 'Principles of Optics', 7th edition (Pergamon, New York, 1999) p. 219.
[25] C. A. Mack, Proc. SPIE 538, 207–220 (1985).
[26] H. A. Macleod, in 'Thin Film Optical Filters' (McGraw-Hill, New York, 1989).
[27] M. Mansuripur, J. Opt. Soc. Am. A3, 2086–2093 (1986).
[28] A. Wong and A. R. Neureuther, IEEE Trans. Electron Devices 41, 895 (1994).
[29] G. Wojcik, J. Mould, R. Fergusson, R. Martino and K. K. Low, Proc. SPIE 2197, 455–465 (1994).
[30] K. D. Lukas, H. Tanabe and A. J. Strojwas, J. Opt. Soc. Am. A 13, 2187 (1996).
[31] B. H. Kleemann, J. Bischoff and A. K. Wong, Proc. SPIE 2726, 334 (1996).
[32] J. Mulkens, D. Flagello, B. Streefkerk and P. Graeupner, J. Microlith. Microfab. Microsyst. 3, 104–114 (January 2004).
[33] D. G. Flagello and A. T. S. Pomerene, Proc. SPIE 772, 6–20 (1987).
[34] B. J. Lin, in 'Optical Lithography: Here Is Why'. (SPIE, Bellingham, Wash. (1000 20th St. Bellingham, WA 98225-6705, USA), 2010).
[35] Y. Borodovsky, Proc. SPIE 2440, 750–770 (1995).
[36] C. A. Mack, in 'Fundamental Principles of Optical Lithography: The Science of Microfabrication', Chapter 9, (John Wiley & Sons, London, 2007).

Donis G. Flagello has worked within the photolithography field for 30 years. He has numerous publications and patents specializing in many aspects of image formation, including high NA imaging systems, polarization, immersion lithography, aberration and lens metrology, optical system modeling and simulation.

In 1981, Donis joined IBM where he worked in multiple areas of optical lithography that covered manufacturing, development and research fields. In 1994, he joined ASML and concentrated on optimizing and understanding optical exposure tools, with specific attention to aberrations and their relationships to lithographic features that are used in high volume chip production. He is currently working at Nikon Research Corporation of America in Tucson, Arizona where he holds the title of NRCA Fellow. His department concentrates on imaging issues and solutions associated with advanced lithography for future generations. Donis has authored over 50 papers and two dozen patents. He is a member of OSA and a Fellow of SPIE. He holds a Ph.D. from the University of Arizona in Optical Science.

Daniel G. Smith received his PhD from the Optical Sciences Center and has been working as an optical designer in the field of Microlithography at Nikon Research Corporation of America since 2004.