Modeling Image Formation In Layered Structures : Application To X-ray Lithography

Srinivas B. Bollepalli, M. Khan and F. Cerrina

University of Wisconsin, Madison, WI, USA
sbollep@xraylith.wisc.edu, khan@xraylith.wisc.edu, cerrina@xraylith.wisc.edu

ABSTRACT

In the fabrication of semiconductor devices using lithography, the modeling of the exposure process is very often needed. The elements of a typical exposure system from a modeling perspective comprise of a radiation source, a patterned mask and a wafer coated with a photo-resist. The diffracted image of the mask pattern exposes the photo-resist after propagating through a distance termed *mask-to-wafer gap*. The exposed wafer is later on chemically developed to form a semiconductor device. In short, the electric field propagates through a series of regions (*layers*) with various materials and topographies (*structures*) before forming an image on the wafer.

In this paper we describe a computational algorithm based on *angular spectrum propagation* approach to model image formation in layered structures with particular emphasis to X-ray Lithography (XRL). Several illustrative examples will be presented.

Keywords: Angular Spectrum, Diffraction, Defects, Fluorescence, X-ray Lithography.

Introduction

The modeling of X-ray Lithography (XRL) image formation [1,2] is based on the Fresnel approximation of the diffracted electric field. The complication of image formation in XRL stems from the system itself: mask topography, absorber side-wall slope variations, non-uniformities, materials, roughness in the mask and so on. To solve all these problems, a simulator, *CXrl Toolset*, based on a "layering" approach was created. In this method, the whole system - mask, gap, resist, wafer - is modeled as a sequence of non-uniform layers. With this, arbitrary mask and resist structures can be modeled as long as the in-plane distribution of the materials is known. The mask presents the most challenging problem - topography. In order to simulate this, an "ideal" mask pattern is subjected to a sequence of transformations. Some of these transformations include rounding of corners of a polygon, expansion or shrinking of features etc. This structure is then modeled as a series of thin transmission layers used in image calculations. The image is propagated from layer to layer by using classical scattering approach. A typical layered system is shown in above figure.

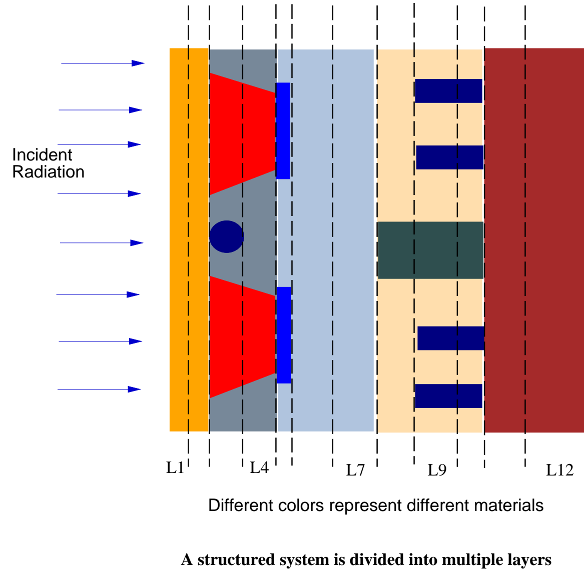

Figure 1: Illustration of layered system

In the application of modeling to harder X-rays (like those used in LIGA) it is necessary to take into account other effects. In particular, X-ray fluorescence must be included. The modeling approach and algorithm will be discussed in detail in the paper. An excellent agreement has been obtained with other detailed approaches. Several applications, such as modeling defects in masks, focusing of Fresnel zone plates etc. will be shown to illustrate the modeling for both MEMS and ULSI problems.

Angular Spectrum of Plane Waves

Let a monochromatic wave field $E(x, y, 0)$ produced by a collection of monochromatic waves be incident on the $z = 0$ plane. The Fourier transform of $E(x, y, 0)$ can be written as

$$FE_0(f_x, f_y) = \int\int_{-\infty}^{\infty} E(x, y, 0)e^{-j2\pi(f_x x + f_y y)}dxdy \quad (1)$$

$E(x, y, 0)$ can be represented in terms of its Fourier

transform as

$$E(x,y,0) = \int\int_{-\infty}^{\infty} FE_0(f_x,f_y)e^{j2\pi(f_x x+f_y y)}df_x df_y \quad (2)$$

The exponential part under the integral, can be regarded as an unit amplitude plane wave propagating in the direction

$$\hat{k} = (\lambda f_x)\hat{x} + (\lambda f_y)\hat{y} + \sqrt{1-(\lambda f_x)^2-(\lambda f_y)^2}\hat{z} \quad (3)$$

This decomposition of $E(x,y,0)$ into plane waves propagating in various directions, with complex amplitudes $FE_0(f_x,f_y)$ is known as *angular spectrum* representation.

Propagation of the angular spectrum

Given $E(x,y,0)$ we are interested in obtaining the field at a distance z, i.e., $E(x,y,z)$. Since $E(x,y,z)$ is assumed to be a scalar wave field, it must satisfy Helmholtz's equation at all source-free points,

$$\Delta^2 E + k_z^2 E = 0 \quad (4)$$

Let the angular spectrum representation of $E(x,y,z)$ be written as

$$E(x,y,z) = \int\int_{-\infty}^{\infty} FE_z(f_x,f_y)e^{j2\pi(f_x x+f_y y)}df_x df_y \quad (5)$$

Direct substitution of equation 5 in equation 4, results in

$$\Delta^2 FE_z(f_x,f_y) + k_z^2 FE_z(f_x,f_y) = 0 \quad (6)$$

An elementary solution to the above equation is

$$FE_z(f_x,f_y) = FE_0(f_x,f_y)exp[j(\frac{2\pi}{\lambda})\sqrt{(1-(\lambda f_x)^2-(\lambda f_y)^2)z}] \quad (7)$$

Therefore we can write

$$E(x,y,z) = \int\int_{-\infty}^{\infty} FE_0(f_x,f_y)exp[j(\frac{2\pi}{\lambda})\sqrt{(1-(\lambda f_x)^2-(\lambda f_y)^2)z}] \quad (8)$$

which describes the propagation of the angular spectrum. There are two important cases here. In the first case, when

$$(\lambda f_x)^2 + (\lambda f_y)^2) \leq 1 \quad (9)$$

we have homogeneous plane waves propagating in the z direction. When,

$$(\lambda f_x)^2 + (\lambda f_y)^2) > 1 \quad (10)$$

we have inhomogeneous or evanescent plane waves, that exponentially decay in the $-z$ direction. Both types of waves should be taken into account in the computations for accurate results. The validity of *angular spectrum* propagation has been studied by several authors and conditions for validity have been given by Lalor [5].

Transmission function of a layer

Consider a planar inhomogeneous layer of thickness d extending from $z = 0$ to $z = d$. Assuming that a monochromatic plane wave of wavelength λ and wavenumber $\frac{2\pi}{\lambda}$ is incident on this layer, we are interested in computing the electric field at the exit of the layer, i.e at $z = d$. Without any loss of generality, we can assume that inhomogeneity is only lateral; for we can treat even inhomogeneities in the z direction by treating the layer as a stack of several thinner layers of thicknesses $< d$. With this, each layer can be characterized by a spatially varying refractive index $\eta(x,y)$. At X-ray wavelengths (1 nm), the refractive indices of any material is complex, with real part being close to unity and a very small imaginary part.

When the medium the index of refraction is not anymore 1, but rather complex, we write

$$\tilde{n} = n + j\kappa \quad (11)$$

and thus in explicit form the effect of the propagation through a space z filled with uniform material can be written:

$$exp[j\ kz] = exp[-\kappa\ k_0 z]exp[j\ nk_0 z] \quad (12)$$

The first exponential represents the attenuation due to the propagation through a distance z, and the second the corresponding phase shift. Notice that $n = Re(\tilde{n})$ appears explicitly in the exponential, effectively increasing the optical thickness of the medium to $n\ z$. Finally, Eq. 12 represents the transmission function, which is position dependent and yields:

$$t(x,y) = exp[-\kappa(x,y)\ k_0 z]exp[j\ n(x,y)k_0 z] \quad (13)$$

In this equation, both κ and n are position dependent, and can be supplied in a bitmap form where each pixel has a transmission given by Eq. 13. Notice further that all the optical properties *of the layer* have been collapsed into the transmission function.

Propagation through a layer

The propagation of a scalar wave field through a layer can be obtained in two steps.

- By multiplication of the incoming wave-field with the transmission function of the layer
- Propagating the angular spectrum of the modulated wave field through a distance equal to the thickness of the layer.

They are described below. *Assuming that the thickness of the object is negligible*, the field *right after* the object will be:

$$E_t(x,y) = E_{in}(x,y)t(x,y) \quad (14)$$

where E_{in} represents the incoming plane wave. Here, we are implicitly applying Kirchhoff boundary conditions and assuming that the incoming wave field is not disturbed and that it is only modulated. For sufficiently thin layers, this is a good approximation [7].

In general, we can write the transmission of the layer using its Fourier representation:

$$t(x,y) = \int d\nu_x d\nu_y T(\nu_x,\nu_y) exp[-j2\pi(x\nu_x + y\nu_y)] \quad (15)$$

This is equivalent to decomposing the transmission object in a sum of infinite diffraction gratings, specified by the frequencies (ν_x, ν_y). Each will contribute to scatter the plane wave, with a weight proportional to the component T of the transmission function at that frequency.

After the wave has been transmitted through the layer, it will propagate through free-space until it reaches the image plane, located at a distance z. Essentially nothing will happen to the plane waves as they propagate; after moving a distance z, the wave is described by Eq. 15. Thus, the central idea is that *a plane wave is modulated by an object of transmission function* $T(x,y)$, *filtered into a new set of plane waves which are then propagated through free space.*

As stated, this represents the *Fresnel-Kirchhoff Approximation* of the diffraction theory. The main problem of this approach is the existence of sudden jumps in the electric field E_t after multiplication of E_{in} by t, at the edges of the features defined by t. We will discuss first the implementation of the FK case.

We notice that the Fresnel Approximation is obtained if we replace the k_z with its series expansion:

$$k_z = \sqrt{k^2 - k_x^2 - k_y^2} = k(1 - \frac{k_x^2}{2k^2} - \frac{k_y^2}{2k^2}) \quad (16)$$

While it is useful for simplifying formulae, for a computer implementation there is no need to use the Fresnel Approximation. We will use instead the exact formula.

General case

This case is more complicated, since we cannot disregard the variations in $E(x,y)$. The analysis remains however quite the same, if we replace the definition of $t(x,y)$ by:

$$E_t(x,y) = \int d\nu_x d\nu_y E_t(\nu_x,\nu_y) exp[-j2\pi(x\nu_x + y\nu_y)] \quad (17)$$

and

$$E_t(\nu_x,\nu_y) = \int dx\, dy E_t(x,y) t(x,y) exp[j2\pi(x\nu_x + y\nu_y)] \quad (18)$$

thus leading to

$$E_t(x,y,z) = \int d\nu_x d\nu_y E_t(\nu_x,\nu_y) exp[-j2\pi(x\nu_x + y\nu_y)] exp[-jk_z z] \quad (19)$$

where, once more, k_z is dependent explicitly on (ν_x, ν_y) It may be interesting to look at equation 21 for another point of view, i.e., that of a convolution kernel (a propagator)

$$K(\nu_x,\nu_y;z) = exp[-jk_z z] \quad (20)$$

k_z represents the propagation in the medium.

Algorithmic Implementation: The Diffraction

The algorithm must reproduce the following physical system:

1. An incoming field of complex amplitude $E(x,y)$ is incident on an object of complex transmittance $t(x,y)$.

2. A transmitted field $E_t(x,y)$ is formed by multiplying, in the real space, t and E:

   $$E_t(x,y) = t(x,y)E(x,y) \quad (21)$$

   where t is the complex transmission function.

3. The field E_t is transfered to the frequency domain by Fourier transform:

   $$E_t(\nu_x,\nu_y) = \int dx\, dy\, E_t(x,y) exp[j2\pi(x\nu_x + y\nu_y)] \quad (22)$$

4. The transmitted field E_t is considered to be the input to the propagation sub-system. The field is propagated through z by writing:

   $$E_{image}(\nu_x,\nu_y) = E_t(\nu_x,\nu_y)\ exp[-jk_z z] \quad (23)$$

   Remember that k_z is a function of (ν_x, ν_y).

5. The field is written again in the real domain by inverse Fourier Transform, yielding E_{out}.

6. If the propagation is completed, exit. Else, set $E = E_{out}$ and repeat from (1).

This applies to the case of:

- The object has zero thickness
- The propagation is in vacuum
- The lateral variation of $t(x,y)$ is "slow" in comparison to the wavelength.

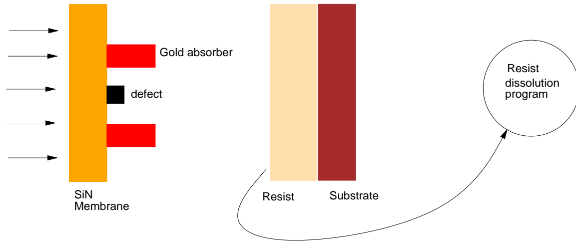

Figure 2: An in between mask defect

Dose absorption in resist and development

The absorbed dose in the resist is defined as the amount of energy absorbed per unit volume. Given a resist of thickness Z_r, the change in intensity inside the resist between depths, z and $z+dz$ as a function of depth z can be written as

$$dI(z) = I(z)\alpha dz \quad (24)$$

where α is the absorption coefficient and $I(z)$ is magnitude squared of the wave field at Z. $I(z)$ can be expressed as a function of the incoming intensity at the surface of the resist as

$$I(z) = I(0)e^{-\alpha z} \quad (25)$$

When integrated over time we get the delivered dose

$$D(z) = D(0)e^{-\alpha z} \quad (26)$$

Once the absorbed dose in the resist is obtained, it can be mapped into a dissolution rate specific to the resist which can then be used by a *resist dissolution* model to obtain the final developed pattern. In this paper, a *resist dissolution* model based on a *level set* approach is used. For more details see [8].

Illustrative Examples

Study of Defects in X-ray Masks

Defects in X-ray masks are caused by errors during mask fabrication and by contamination while using the mask for printing [10]. Since, XRL is a 1:1 printing process, a defect present in the mask can alter the printed pattern under suitable conditions. It is important to know the critical sizes of the defects that need to be removed before the mask is used for printing. This creates certain topographical problems from the aspect of modeling. With our layer-based approach, a study of defects can be made quite easily. This is illustrated with two examples.

First, consider an X-ray mask which has been patterned with 1:2 lines and spaces. Suppose a small stainless steel defect is placed in between a pair of lines in the mask. The thickness of the defect is set to be half

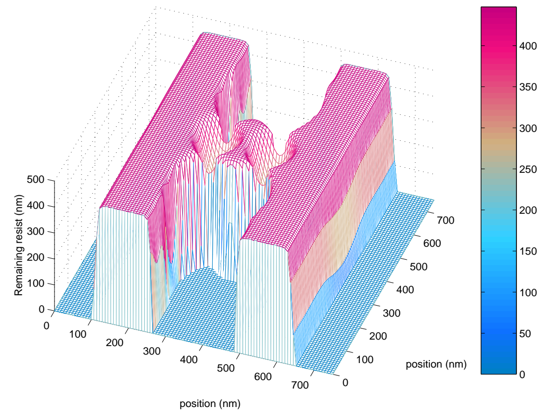

Figure 3: Effect of a mask defect on developed resist

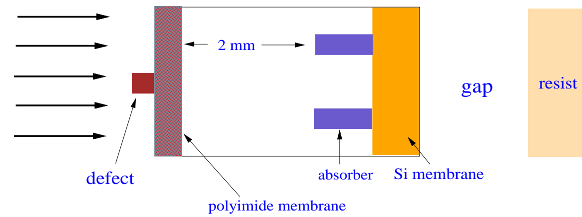

Figure 4: Schematic of a mask pellicle

that of the thickness of the gold absorber. The modeling of this scenario is illustrated in figure . APEX-E was used as the resist on the wafer. The exposed resist was developed using a program *FastPhoto* based on *level set* methods. The resulting developed pattern is shown in figure .

It has been suggested that X-ray masks be covered by a protective sheath or *pellicle* for avoiding contamination during exposures and for ease of cleaning. Here, as the second example, we would like to know the effect of a defect placed on the front of the *pellicle*. Again, we considered the same mask pattern as in the previous example. This setup is illustrated as in figure . The developed resist is shown in figure 4. Notice the bending of the lines because of distortion of incoming plane waves by the stainless steel defect.

Study of Image Formation from Point Sources

It is well known that point sources provide an alternate to synchrotron radiation for X-ray lithography. In spite of being compact and economical, there are some inherent disadvantages associated with point sources [6]. Point sources are characterized by a source divergence angle θ_s and a global divergence angle θ_g. As a result,

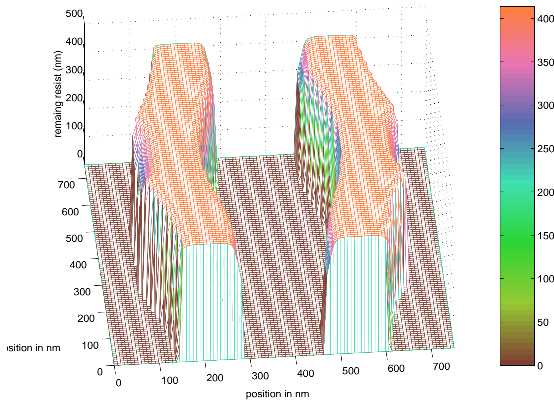

Figure 5: Effect of a pellicle defect

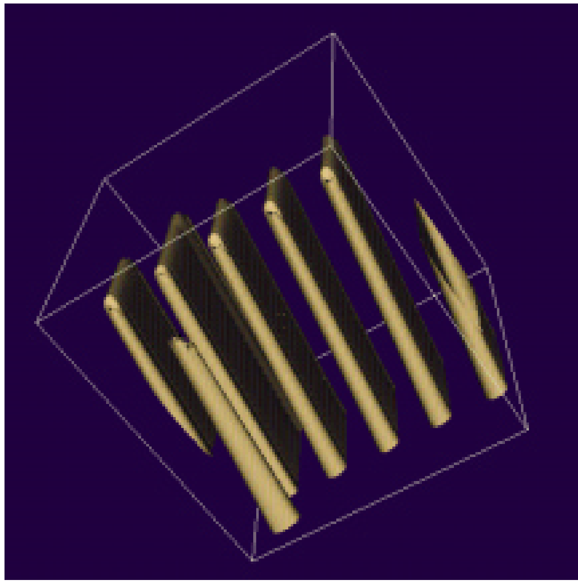

Figure 6: Effect of oblique incidence on developed resist

the fields at the corners and edges of the mask receive an oblique illumination if a collimator is not used. The result of developing a complex mask pattern to such an oblique illumination (20 mrads) is shown in figure . Notice that the resist side-walls are slanted.

Fluorescence effects

Studies such as those used in LIGA, use hard X-rays for illumination of the mask. A typical spectrum is shown in figure . At these short wavelengths, fluorescence effects should be considered in the model. When an inner shell atomic electron is ejected from an atom as a result of a collision process involving a photon or other incoming projectile, a vacancy is thus created in that electron's pre-collision sub-shell. In one mode of de-excitation called *fluorescence* a fluorescence or characteristic X-ray is emitted from the atom, with photon energy equal to the difference between the vacancy site

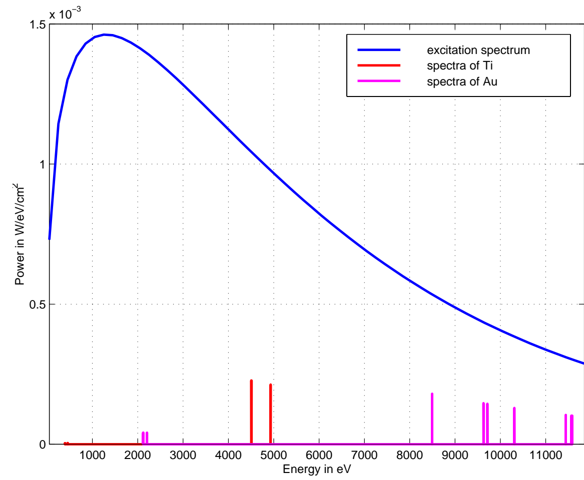

Figure 7: A example of fluorescence spectrum

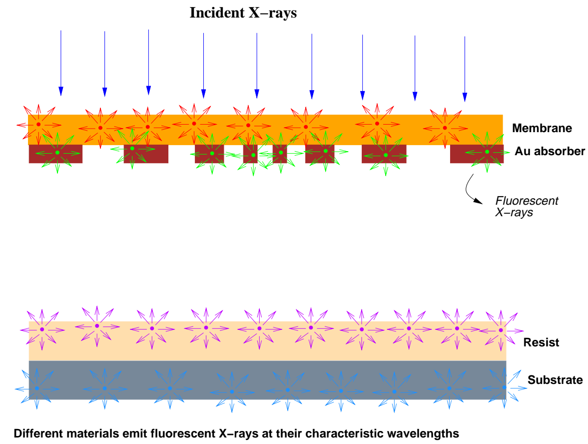

Figure 8: Illustration of fluorescent radiation

inner-shell energy level and the energy level of the particular outer shell which happens to supply the electron to fill the vacancy. From an XRL perspective, when X-rays strike the mask, there can be characteristic emissions from the *membrane*, *absorber*, *substrate* and the *resist* itself. The intensity of the emission depends upon the *fluorescence yield* which is defined as

$$\omega_i(Z) = \frac{f_i(Z)}{\nu_i(Z)} \tag{27}$$

where $f_i(Z)$ is the average number of fluorescence X-rays emitted and $\nu_i(Z)$ is the number of vacancies in the ith shell or sub-shell and Z is the atomic number of the element. In figure 7, the emission spectra from a *titanium membrane* and *gold absorber* are shown while taking into account the corresponding *fluorescence yields*. For an excellent review of tables of *fluorescence yields* see Hubble[9]. From an image formation perspective, the fluorescence radiation emitted by *membrane*, *absorber*,

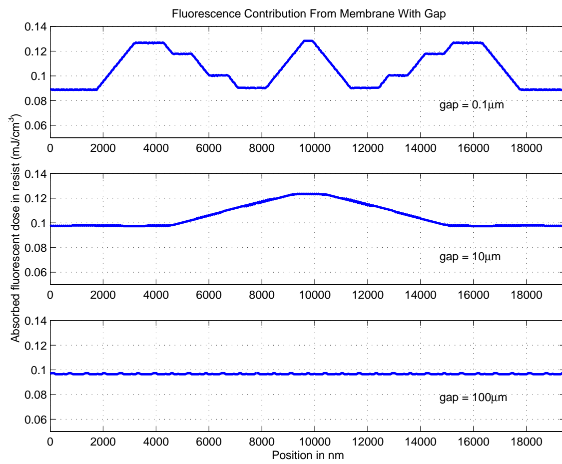

Figure 9: Membrane contributions

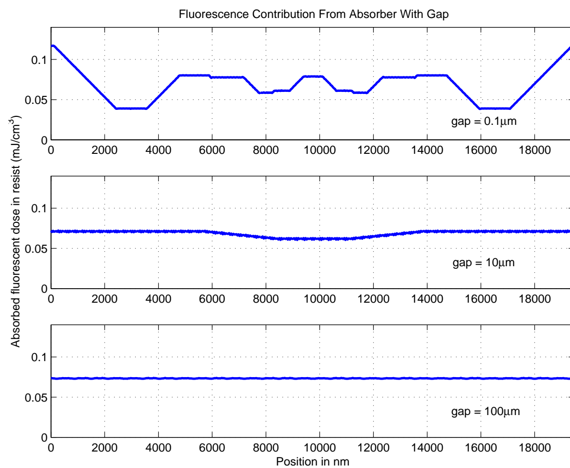

Figure 10: Absorber contributions

resist and *wafer* is incoherent. This causes computational problems, because we need to consider an infinite number of point sources and add their collective intensity contributions at the resist. In order to avoid this, we used the *phase-space* approach of geometrical optics for modeling. Figure 8, shows the fluorescent radiation emitted in various parts of the exposure system. The contribution from the *resist* and *substrate (Si)* are negligible compared to that of the *membrane* and *absorber* and hence have been ignored in our computations. The *membrane* contributions are shown in figurei 9 and the *absorber* contributions are shown in figure 10 for various mask-to-wafer gaps. It can been seen that at large gaps, the fluorescence contribution becomes uniform and is quite low. For the system illustrated in figure , we observed that the net increase in absorbed dose by the resist due to fluorescence effects is of the order of 0.6%. This becomes important only when the contrast of the structure is of prime importance.

Conclusions

In this paper we have described the modeling of image formation in layered structures using the angular spectrum propagation approach. Several examples have been presented as application to X-ray lithography. This modeling approach can be used to solve a variety of problems in an efficient manner. The use of the angular spectrum approach is computationally very efficient, as the the image propagation is carried out using Fourier transforms from layer to layer. However, for very large grid sizes, this may cause requirements for large memory. Some additional effects such as fluorescence, have been described.

REFERENCES

[1] B.E.A. Saleh and M.C. Teich, *Photonics*, p.111-121, Wiley 1991

[2] J.W. Goodman, *Introduction to Fourier Optics*, McGraw-Hill, 1968

[3] A. Sommerfeld, Optics, p. 237, Academic 1964

[4] W.H. Southwell, *Validity of the Fresnel Approximation in the near field*, JOSA, **71**, 7-14 (1981), Eqs. 7-9.

[5] E. Lalor, *Conditions for the Validity of the Angular Spectrum of Plane Waves*, JOSA, **58**, 1235 (1968)

[6] Cerrina F., *X-ray Lithography*, Chapter 3, Handbook of Micro-lithography, Micro-machining and Micro-Fabrication, Volume 1, 1997.

[7] Guo Z. Y., Cerrina F., *Modeling Proximity Lithography* ,IBM Journ. Res. Develop., **37** 331, (1993).

[8] Mumit Khan, B. S. Bollepalli, F. Cerrina., *A Semi-Empirical Resist Dissolution Model for Sub-micron lithographies*, MSM, 1998.

[9] J. H. Hubbell, et. al., *A Review, bibliography, and Tabulation of K, L and Higher Atomic Shell X-ray Fluorescence Yields*, J.Phys.Chem.Ref.Data, Vol 23, No.2, 1994.

[10] B. S. Bollepalli, S. D. Hector, J. Maldonado ,J. A. Leavey , F. Cerrina, M. Khan, *Simulation of X-ray mask defect printability*, Proc. SPIE, Vol 3048, pp 155-166, Emerging Lithographic Technolgies;David E. Seeger, Ed;