ENHANCED, QUANTITATIVE ANALYSIS OF RESIST IMAGE CONTRAST UPON LINE EDGE ROUGHNESS (LER)

by

Michael Victor Williamson

ENHANCED, QUANTITATIVE ANALYSIS OF RESIST IMAGE CONTRAST UPON LINE EDGE ROUGHNESS (LER)

by

Michael Victor Williamson

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

Enhanced, Quantitative Analysis of Resist Image Contrast upon Line Edge Roughness (LER)

by

Michael Victor Williamson

B.A. (University of Illinois, Urbana) 1995
B.S. (University of Illinois, Urbana) 1995
M.S. (University of California, Berkeley) 2000

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy
in

Engineering-Electrical Engineering and
Computer Sciences

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, BERKELEY

Committee in charge:
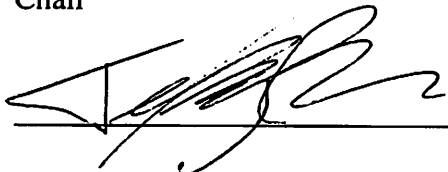
Professor Andrew R. Neureuther, Chair
Professor Jeffrey Bokor
Professor Eugene E. Haller

Fall 2003

The dissertation of Michael Victor Williamson is approved:

_____ 11/6/03
Chair   Date

_____ 11/5/03
                                                  Date

_____ 11/6/03
                                                  Date

University of California, Berkeley

Fall 2003

Enhanced, Quantitative Analysis of Resist Image Contrast upon Line Edge Roughness (LER)

Copyright 2003

by

Michael Victor Williamson

Abstract

Enhanced, Quantitative Analysis of Resist Image Contrast upon Line Edge Roughness (LER)

by

Michael Victor Williamson

Doctor of Philosophy in Electrical Engineering

University of California, Berkeley
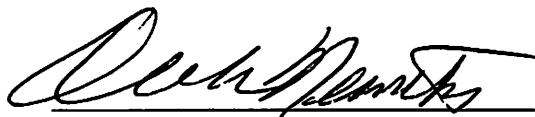
Professor Andrew R. Neureuther, Chair

This thesis examines the dependence of photoresist line edge roughness (LER) on litho-graphic aerial image contrast. A wide range of image contrasts are created using a double exposure programmed image technique where the total exposure dose and the ratio of fore-ground to background exposure dose are varied. As contrast is altered, one-dimensional LER and two-dimensional side edge roughness (SER) are measured on UVII-HS, UV210, SEPR-463, and Apex-E.

Two-dimensional SER studies of around 20 samples focus upon the shape of photoresist roughness. Roughness peak to peak spacing has a bimodal distribution, with one mode dis-tributed around roughly 30 nm and the other mode centered around 100 nm. Roughness cur-vature is on the order of 1°, much wider than the atomic force microscope (AFM) probe tip used to accurately measure roughness.

LER is first examined with multiple photoresists at 9 different contrast levels each. Some unexpected results of this experiment lead to a new experiment and to the creation of a new form of defining contrast based solely upon the background exposure, the background con-trast. This contrast is orthogonalized as best as possible to the standard (max - min) contrast

1

and the image slope at clear dose, and approximately 200 LER measurements are collected at 8 different contrast levels. Rigorous statistical analysis conclusively proves that LER tracked background contrast best, with LER $\propto$ background contrast^{-1} approximately. The statistical F-ratio fitting LER measurements to a predicted LER model using background contrast was 163.3, whereas the F-ratio fitting data to a predicted LER model using the next best form of contrast was 56.1.

Finally, an image deblurring algorithm often used in astronomy was applied to scanning electron microscope (SEM) images using 3 different SEM machines in order to improve image quality in the high frequency domain and obtain better LER measurements on these SEMs. The algorithm was rigorously tested. It was able to deblur SEM images in most any cases and it altered LER measurements roughly 11-15%, however statistical variation in LER among the samples was so large than any potential improvements in LER measurement accuracy were undetectable. Nonetheless, some unexpected benefits of the deblurring algorithm are that it can help troubleshoot SEM machines and create a larger operating specification window for the SEM tool.

Professor A.R. Neureuther

Committee Chairman

2

Acknowledgements

No experience occurs in a vacuum. The completion of a dissertation is perhaps the best example of this truism. Professor Andy Neureuther, my advisor, provided incalculable guidance, support, and trust in my capabilities. He is a brilliant and extremely well-rounded individual and I enjoyed being able to work under his tutelage. Most of all, I enjoyed watching how his mind worked. He understands optics on an instinctive level, envisioning solutions in his head during breakfast at least as often as with a pen and paper. I would also like to thank Professor Jeffrey Bokor. He is an amazing instructor, explaining extremely complicated concepts and mathematics in optics with ease. I only hope I can keep all of that knowledge in my head. I had some wonderful discussions with Professor Grant Willson at the University of Texas- Austin. He provided a great deal of insight into the chemistry of photoresists, and the reaction and diffusion mechanisms after lithographic exposure. I had many wonderful instructors and classes during my graduate studies. I want to quickly thank Professors Chen-ming Hu, Eugene Haller, Costas Spanos, Nathan Cheung, Tsu-Jae King, Michael Lieberman, and David Presti, as well as the professors mentioned earlier, for providing excellent classes that I thoroughly enjoyed.

Many thanks go to Professor William Oldham, SRC, and DARPA for providing a research platform that spanned UC- Berkeley, Stanford, UT- Austin, Georgia Tech, and several corporations. I would specifically like to thank Dr. Houle and Dr. Sanchez at IBM Almaden for enlightening discussions and ASM Lithography for the stepper tool. This collaboration provided an opportunity to interact with experts throughout the entire country on a fairly regular

i

I would also like to thank all my fellow colleagues in the research group for all of their help. Please don't be offended by the barrage: Tom Pistor, Kostas Adam, Yunfei Deng, Mike Lam, Garth Robins, Yashesh Shroff, Mike Shumway, Jacob Poppe, Dan Ceperley, Frank Gennari, Greg McIntyre, Hideaki Oshima, Sachan Madahar, Yijian Chen, and Scott Hafeman. Yash, thanks for providing me a daily distraction from work (or was it the other way around?). I am especially grateful towards Lei Yuan, Ebo Croffie, and Mosong Cheng for the numerous discussions we had regarding photoresist acid diffusion, reaction and development. Big thanks go to Jason Cain for helping me through the statistics, making sure I properly knew what things like "independently, identically, normally distributed" meant. Vivian Kim and Ruth Gjerde, you helped me muddle through the red tape of academia; I am very grateful. I also want to send thanks to Kim Chan for *all* her help with experiments and Xiaofan Meng for help with the atomic force microscope. I owe Evan Stateler my gratitude for ensuring that the stepper and the electron microscope were in such wonderful condition.

Lastly, I would like to thank all those who helped keep me sane through all the tough periods. Mom, Dad, I obviously owe you everything. Sorry I'll be leaving you once again as I head up to Portland. Brian, Rich, I promise not to send any more Fourier-transformed family photos. Herr Doktor Lau, I already miss our adventures, ranging from Beauty Bar to the American River. Jeff, Jerry, Tim, Andy, Mark, and Vonnie, thanks for sharing my obsession with games. Jolese, I'll miss our meandering talks. Most of all, I want to thank my wife Diana. You saw me through thick and thin, most of the time even with a smile on your face. I am so happy and lucky to have you in my life. I owe my eternal gratitude to Awmawni;

thanks for accepting me into your family. Now if only I can figure out how to spell that with Latin characters. Finally, I would like to thank San Pedro for spiritual guidance and awareness.

"No puedo mas"

-- El ingenioso hidalgo Don Quijote de la Mancha, part II, Miguel de Cervantes

Table of Contents

List of Figures

List of Tables

1 Introduction

A fundamental problem facing the semiconductor industry at the 65 nm technology node and below is local line edge roughness (LER). At the 65 nm node, which is expected by the end of 2003, the budget for the 3σ linewidth variation of critical feature dimensions is 3.7 nm [1]. This budget goal has no known manufacturable solutions. Roughness is defined as any deviation from a perfectly straight edge. After lithographic processing, roughness on the photoresist is passed on to the underlying gate during etching, damaging device performance, notably increasing leakage currents [2 - 3]. Figure 1-1 contains scanning electron micrographs (SEMs) of two photoresist trenches created by exposing the photoresist to light and developing the exposed regions. Image (a) has little roughness while (b) has a high level of roughness. The goal of this thesis is twofold: to better understand the root causes of LER, specifically how the exposure aerial image affects LER, and to more accurately and reproducibly measure LER.
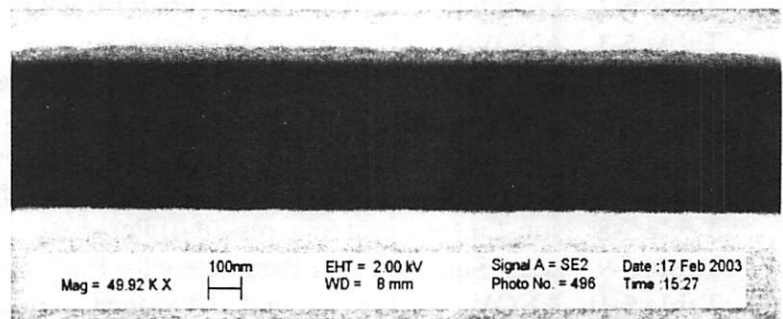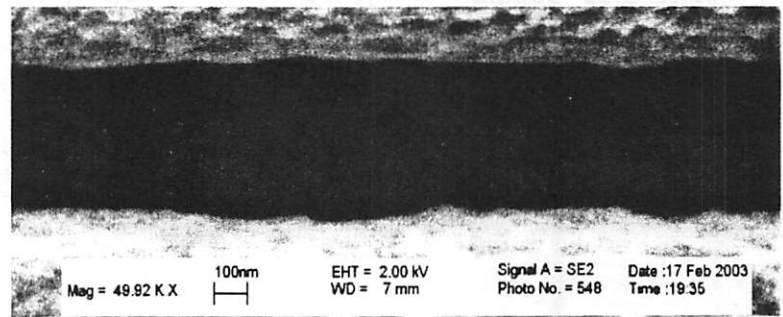
Figure 1-1: Top image (a) is a scanning electron micrograph of a trench in photoresist with little roughness. Bottom image (b) has large roughness.

Lithographic photoresists designed for 248 nm wavelength Krypton-Fluoride (KrF) lasers, which are now well established chemically amplified photoresists (CARs), typically have a line edge roughness nearly consuming all the critical dimension control tolerance [3]. Photoresists designed for 193 nm wavelength (Argon Fluoride) illumination, which are now being used in the most aggressive fabrication processes, demonstrate even poorer line edge roughness [4]. The 157 nm wavelength (Fluorine) photoresists lag further behind both of the others [5]. The International Technology Roadmap for Semiconductors (ITRS), created by the Semiconductor Industry Association (SIA), sets a pace the semiconductor industry intends to follow and recognizes the difficulty in reaching certain requirements such as the control of critical features. As critical feature sizes shrink, the tolerances around those feature sizes shrink as well, creating increasingly difficult engineering problems. Yet in order to continue the exponential growth predicted by Moore's Law [6] and ensure the economic viability of the semiconductor industry, the ITRS must be followed.

Understanding line edge roughness requires understanding the many aspects of the lithographic process. The lithographic imaging process will be discussed in some detail in chapter two and has been discussed in greater detail in many handbooks and textbooks [18-21]. To briefly summarize, one must consider the photoresist exposure, bake, and develop processes as an amplifier of the original aerial image. This is because, during the lithographic imaging process, the deprotection of positive tone photoresist (or the cross-linking of negative tone photoresist) acts to some extent as an amplifier. After lithographic exposure, photoresist-coated wafers are heated in a post exposure bake (PEB). At this step the essence of chemically amplified photoresist processing occurs. Photoacid generator (PAG) first releases a pro-

2

tonic acid (H+) during exposure, then that protonic acid diffuses significant lengths and catalytically deprotect the photoresist resin during the PEB.

Perhaps more important is the development process. After the PEB, photoresist that is sufficiently deprotected is removed when the wafer is sprayed with or dipped into a developing solution. The amplifying effect mentioned above lies within this process. A photoresist resin polymer molecule with a sufficient level of deprotected sites is soluble in the developer solution. However, a molecule with slightly less deprotection is fully insoluble [33]. It is this sharp contrast, from fully soluble to fully insoluble with little change in polymer composition, which accounts for most of the gain amplifier-like behavior.

There are several factors that contribute to photoresist LER during lithographic processing. LER has been attributed to phase separation, photoresist resin size and structure, acid diffusion, and lithographic imaging. Lithographic imaging includes imaging using EUV, DUV, and electron beam technologies. In order to be able to reduce LER and better control gate CD, the effect these factors have upon roughness needs to be limited. This thesis focuses upon how lithographic imaging affects LER. The next paragraph will expound upon the other causes, however all of them will be covered in further detail in chapter two.

Various studies have shown that phase separation [7-8, 10, 12, 70], photoresist resin size and structure [7, 9, 13, 34], photogenerated acid diffusion [11, 30-32, 48, 73+], and photoresist thickness [22] all play a role in LER formation. Phase separation is a natural process whereby immiscible substances, those with differing polarities, divide into distinct parts. A classic example is a mixture of water and oil. The lithographic process is based upon converting a fairly non-polar polymeric species into a polar species which can be dissolved in an aqueous developer. Therefore, phase separation may occur during lithography and larger col-

loidal structures of like species may form, causing LER. The photoresist resin size and struc-

ture are also culprits suspected of causing LER. Each polymer resin molecule is quite large

with respect to LER, approximately 5 nm in size. However, the resin cannot arbitrarily be

shrunk; there are many critical physical parameters of the photoresist such as the glass transi-

tion temperature that depend upon the size of the polymeric resin. Photogenerated acid diffu-

sion during the post-exposure bake process, another suspected cause of LER, has perhaps

received the most study. The chemically amplified lithographic process requires acid diffu-

sion. Diffusion is a statistical process: some acids may diffuse further than others. This vari-

ation in acid diffusion lengths leads to a variation in the location of the line edge, or LER.

Finally, the lithographic aerial image obviously plays a role in how a pattern is formed in

the photoresist. Therefore, it must also have some influence upon the roughness of that pat-

tern edge, or line edge roughness. Whether the lithography is extreme ultra-violet (EUV),

DUV, electron beam, or ion beam lithography, the basic principle is the same. An image is

transferred to a resist which is sensitive to that type of lithographic exposure. In the case of

optical lithography (DUV and EUV), light is shone either through or onto a photomask. This

mask allows some of the light to pass through it (or reflect off of it). By allowing some light

to pass and rejecting the remainder, an image is transferred into the photoresist. Advanced

lithography prints features that are roughly the size of the wavelength of light printing them,

and in the last few years the features are even smaller. This creates aerial images in the photo-

resist with grey regions-- where some light is received, but not the full level. These non-ideal

images create LER in the photoresist. The quality of the aerial image, or the similarity

between it and an ideal image, is referred to as the aerial image contrast, or simply the con-

trast.

As mentioned earlier, the resist development process acts as an amplifier. The photoresist slope generally appears as an enhanced version of the original image intensity slope. This also means it enhances any statistical variation due to acid diffusion or PAG distribution. There is a region near the profile edge where the lithographic exposure intensity slowly varies from sub-threshold to a level sufficient to allow photoresist deprotection. As the intensity varies more slowly, the effects of the aforementioned statistical variations are enhanced. Several studies, the first of which were performed by Sanchez et al. at IBM Almaden, have been conducted to determine the affects of aerial image profile upon LER, using optical and electron beam lithography platforms [14 - 17, 74]. Sanchez altered the aerial image via laser interferometry. In this thesis, the aerial image is altered by placing a background clear field image upon the foreground image, the proper image. Figure 1-2 displays a cutplane of 8 different
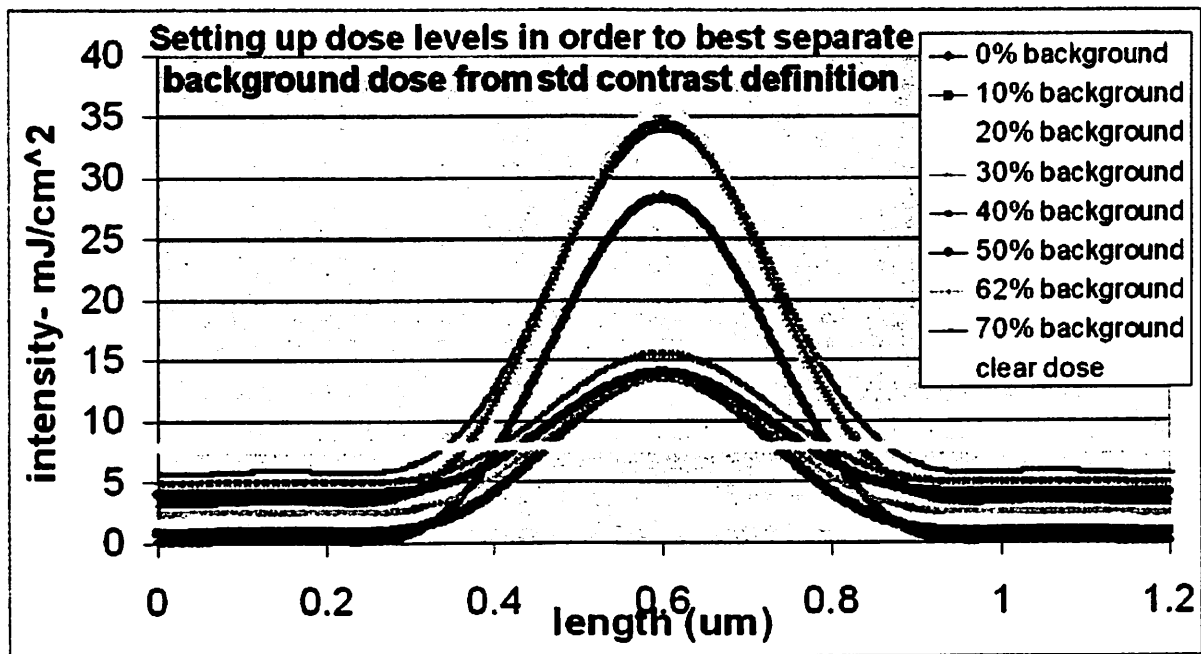
Figure 1-2: Cutplane of the 8 different aerial image profiles used to create an experimental design to be used in chapter 5. Profiles such as the 62% background profile are very poor, while images such as the 0% background and 20% background are very good.

aerial image profiles created using this technique. The bulk of this thesis addresses how lithographic imaging and aerial image contrast impact line edge roughness.

The remainder of the thesis addresses photoresist roughness metrology. Photoresist roughness is often defined as either a 1-dimensional phenomenon, commonly referred to as line edge roughness (LER), or a 2-dimensional phenomenon, referred to in this paper as side edge roughness (SER). LER is a top-down view of a line or trench; the measurement captures roughness information along the line, but not throughout the height of the line. No information regarding the roughness perpendicular to the substrate surface can be collected. A top-down SEM image, such as figure 1-1, can be used effectively to collect LER data. Figure 1-3 shows a SER measurement taken using an atomic force microscope (AFM). SER measure-
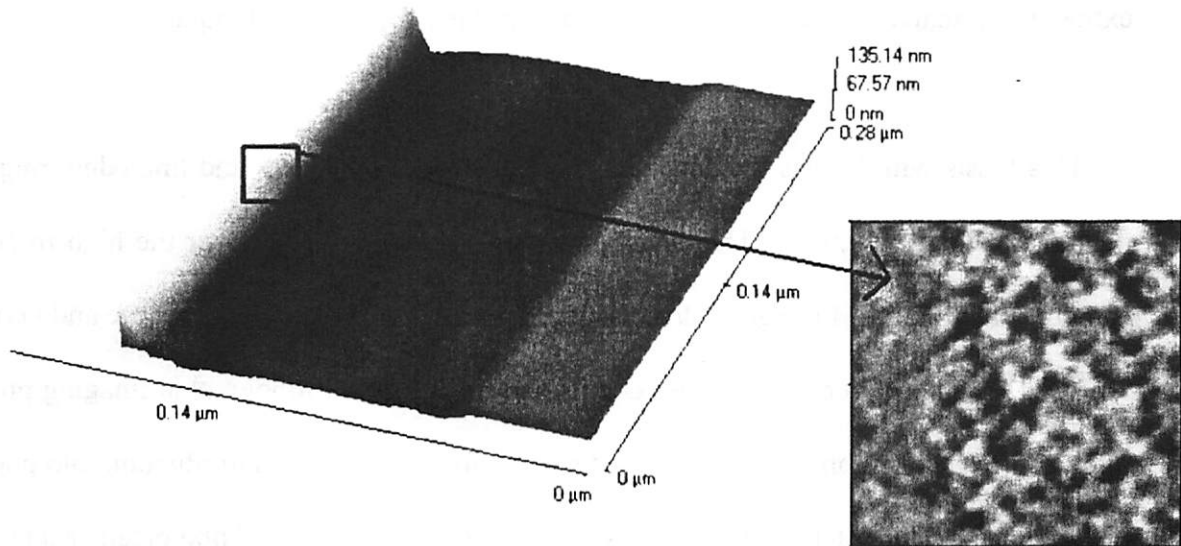
Figure 1-3: AFM image of trench in photoresist. SER is found by flattening the edge highlighted in the box onto a plane, as shown. This flattened image yields a roughness measurement along the line and normal to the substrate.

ments yield roughness information along the printed line pattern and also normal to the substrate.

Non-idealities in the SEM or AFM blur the image so that some of the high frequency data in the image are lost. A technique is presented here that recovers the high frequency component of an SEM image by removing some of the introduced blur. (The technique could be applied to AFM images as well.) The deblurring technique uses a maximum likelihood algorithm based upon the Richardson-Lucy Blind Deconvolution algorithm, originally developed for use in image processing fields, especially in astronomy [27-29]. This algorithm has been used for other imaging sciences as well, such as fluorescence microscopy [69]. Details of how to adapt this technology for use in measuring LER are given, the ability of this technique to indeed improve LER measurement accuracy and repeatability is shown, and the validity of the technique and algorithm in deblurring an SEM image is demonstrated. Further information regarding the condition of an SEM machine used to capture deblurred image can be extracted by analyzing the images recovered from the deblurring technique.

This thesis will discuss the dependence of side edge roughness and line edge roughness upon aerial image contrast. The development of a technique to recover the high frequency components of an SEM image is also discussed, along with tests to demonstrate and verify its efficacy. Chapter two covers further details and theory of the lithographic imaging process. The photoresist reaction will be examined more fully, including an introduction into photoresist parameters, and interesting questions and strategies to understand photoresist characteristics. It also briefly details the parameters of the photoresists used throughout this study. Chapter three discusses the side edge roughness (SER) results obtained, examining the SER

data in frequency space to get an idea of the spatial frequency spectrums of various photoresists. Chapter four details the experimental design and results of the first in-depth aerial image contrast versus LER study conducted. The correlation between contrast and LER is studied using three commercial DUV photoresists, Shipley's UV210 and UVII-HS and Shin Etsu's SEPR-463, with some unexpected results. Chapter five builds upon the results of chapter four, creating a new definition for aerial image contrast, the background contrast. After designing a new experiment using UVII-HS, rigorous statistical results are extracted concluding that this new background contrast is critical to controlling LER. This chapter presents the most compelling evidence that reducing background flare is critically needed to reduce LER. Chapter six discusses the difficulty in extracting accurate LER measurements. It relays some of the basic theory involved in the capture of real images and the concept behind image deblurring. It also touches upon why the high frequency components of the SEM image are vital to extract accurate LER data. Chapter seven provides experimental evidence validating the efficacy of the maximum likelihood algorithm in deblurring images. Furthermore, it shows that LER measurements are altered significantly before and after deblurring, and demonstrates how the deblurring algorithm can help assess the conditions of the SEM tool. Finally, chapter 8 summarizes this thesis and provides insight into valuable directions to explore for future photoresist roughness studies.

2 Lithographic Imaging and Pattern

Transfer

This chapter summarizes the photolithography process, focusing upon the aspects that most influence LER. It gives a perspective on current work in the literature regarding lithographic principles that influence LER, and provides an overview of the experimental methods applied in this thesis to study LER. The first few sections describe the aerial image, how the photoresist reacts to this image, and how this study alters the image. The latter sections discuss photoresist deprotection and development, and the other phenomena briefly mentioned in chapter one such as phase separation.

2.1 The Aerial Image Profile

The shape of the aerial image used to print lithographic patterns is critical. The entire industry of optical lithography, and in fact semiconductor processing, has been so successful because it has been able to scale down the lithographic wavelength, thereby scaling the minimum printable feature. Changing technologies to shorter and shorter lithographic wavelengths allows ever-smaller features to be printed. As Lord Rayleigh observed roughly 100 years ago, the minimum resolvable distance, δ, between two objects is dependent upon the wavelength of light [38]:

Equation 2-1:

$$\delta = 0.61 \frac{\lambda}{NA}$$

where λ is the wavelength of light used to observe the objects, and NA is the numerical aperture of the optical system used to observe the objects. In the case of lithography, the objects are defined rather than observed. The equation holds true regardless. Transferring to lower wavelengths, or larger numerical apertures, is exceedingly difficult and costly; photolithography is the most expensive processing step in the manufacture of integrated circuits. In order to avoid escalating costs and engineering difficulties, many creative alternatives to decreasing the wavelength have been made, such as phase shift masks [21]. Nonetheless, the smallest feature size is always near the resolution limit of the lithographic tool. This creates aerial image profiles with poor image intensity slopes.

Equation 2-1 is based upon the fact that at shorter wavelengths and larger numerical apertures, the slope of the aerial image profile becomes much steeper. As an
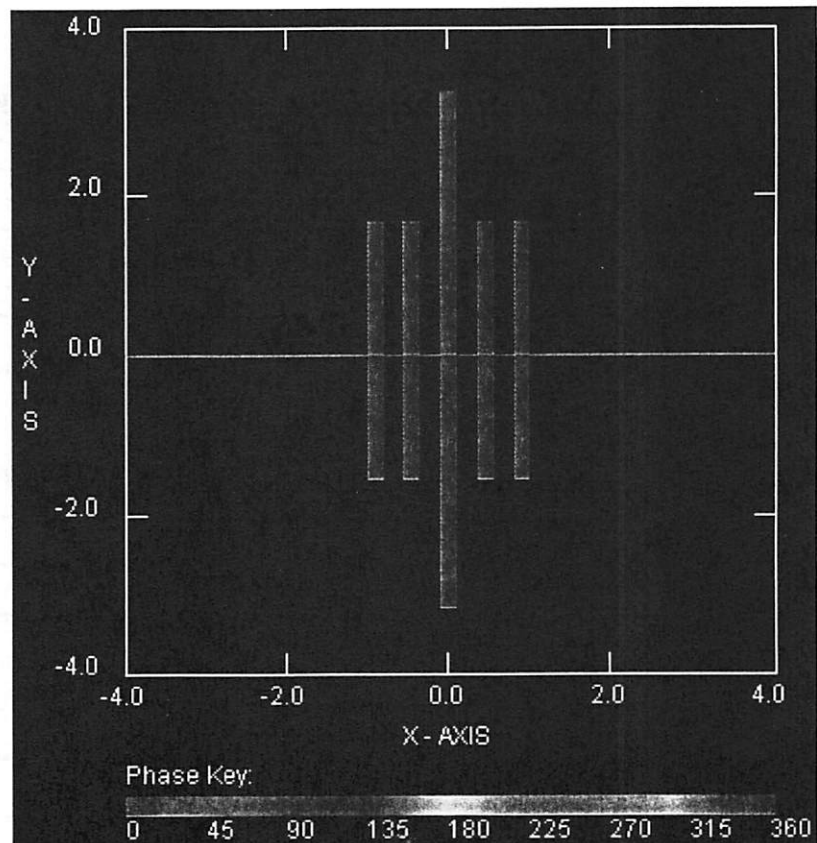
Figure 2-1: Mask pattern of 5 rectangular trenches (clear areas) in chrome photomask. Units are in microns. Black is chrome, blue is 0 phase clear area. The red line represents the cut plane used for figure 2-2.

10

example, figures 2-1 shows a mask pattern with some open areas, forming trenches on positive tone photoresist. Figure 2-2 is a cutplane of the aerial image profile, or the shape of the aerial image intensity, along the center of the mask pattern on figure 2-1 at various illumination conditions. The open areas of the mask are 0.2 μm wide with 0.25 μm space in between the open areas. They are 3.2 μm long, except for the center line which is 6.4 μm long. SPLAT, an aerial image simulator first developed by Toh at the University of California- Berkeley [40], is used to obtain the results shown in figure 2-2. A simple binary mask is used with illumination conditions as given in the figure. The partial coherence, or σ, remains constant at 0.5. As the wavelength drops or the numerical aperture increases, finer features can be resolved, as predicted in equation 2-1.
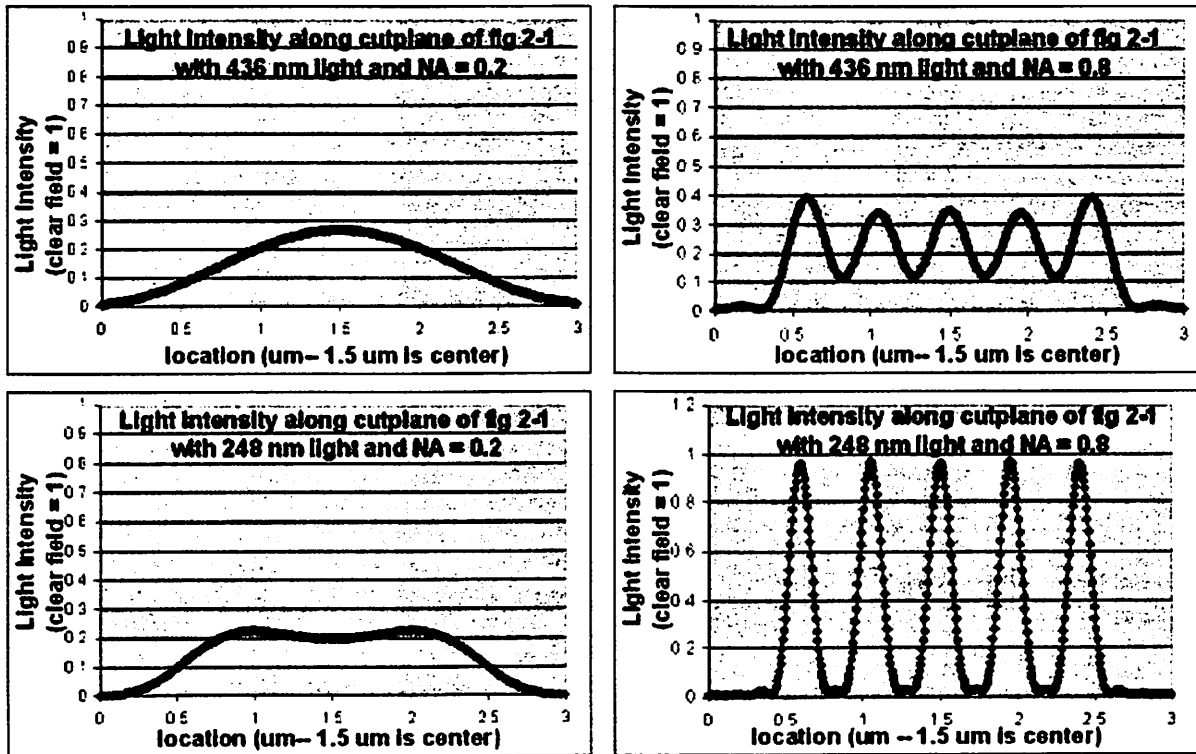
Figure 2-2: SPLAT simulations of the aerial image intensity across the cutplane marked in red on Figure 2-1. Throughout the 4 simulations only the numerical aperture and illumination wavelength of the system are altered. Only the case with 248 nm light and 0.8 NA yields a quality image.

However, another important feature to notice is that the *slope* of the aerial image intensity also becomes larger as the wavelength drops or the numerical aperture increases. Upon consideration, this seems obvious. To resolve two neighboring luminescent objects, the space between them must appear dark. In order for that region to appear dark, the aerial image slope in changing from high to low intensity must be large, like figure 2-2 (d). It might seem just as obvious that poor aerial image intensity slope will increase line edge roughness: the aerial image at an edge is greyer, or has a shallower image slope, therefore there will be more areas of photoresist along the edge which will partially develop. This may lead to LER. The standard contrast definition that has been used most often is:

Equation 2-2:

$$\text{Standard Contrast} = \frac{(\text{maximum energy dose} - \text{minimum energy dose})}{(\text{maximum energy dose} + \text{minimum energy dose})}$$

This contrast definition does not take the aerial image intensity slope into consideration. Describing the entire aerial image profile by simply comparing standard contrast definitions, as in equation 2-2, does not capture the full picture. An aerial image with a poor slope near the resist edge might still have a low minimum energy dose and a high maximum energy dose, yielding a standard contrast near unity (good contrast). An aerial image with superior slope near where the resist edge is defined-- referred to as the clear dose, the development threshold, or the dose to clear-- but maximum and minimum energy doses comparable to those in the poor slope case will likely yield better feature resolution and lower aerial-image associated LER, yet the standard measure of contrast defined in equation 2-2 does not capture the difference between these two aerial image profiles. Many researchers, notably Levenson and

Mack, have begun to use the intensity slope, defining a new form of aerial image contrast called the normalized image log-slope (NILS). NILS is defined as below:

Equation 2-3:

$$NILS = w\frac{d\ln(I)}{dx}$$

where I is image intensity, x is the position along the pattern edge, and w is the nominal line-width of the pattern [41-42].

2.2 Photoresist Response to Aerial Image Profile

Several different photoresists are used throughout this study to determine the correlation between aerial image contrast and line edge roughness. Regarding LER, different photoresists respond differently to aerial image profile changes [15]. LER increases as image contrast degrades for all photoresists, but with some photoresists LER increases nearly monotonically with decreasing contrast while with others LER increases almost as a step function.

In order to further highlight the differences between the various photoresists studied, dis-solution contrast curves, also called Hurter-Driffield (HD) plots, are established for all of the photoresists. The dissolution contrast curve is a plot of the relative thickness of photoresist versus the natural logarithm of the exposure dose [43-44]. A highly non-linear, step function response to image intensity is desirable for a photoresist, in order to provide desirable gain. In this fashion, the photoresist profile can be better than the aerial image profile and a mask pattern can be transferred into the photoresist with high fidelity. Generally, the photoresist is unchanged when it is exposed to light with an intensity significantly below the threshold dose,

E_0. As the intensity approaches the threshold dose, the photoresist is developed away until at E_0, where it is completely removed. The resist contrast or response can be described as:

Equation 2-4:

$$\gamma \equiv \lim_{E \to E_0} \frac{-dT}{d[\ln(E)]}$$

where T is photoresist thickness normalized to full thickness equal to 1, E is the exposure dose, and E_0 is the threshold exposure dose. Figure 2-3 is an HD plot for Shin Etsu's SEPR-463 resist. Chapter three fully describes the lithographic processing conditions used for the side edge roughness experiments, which are the same conditions used in this portion to collect the photoresist exposure dose response curves for determining the photoresist contrast parameter, γ. In figure 2-3, a line is drawn tangent to the slope of the response curve as the relative resist thickness approaches 0. From the slope of this line, using equation 2-4, γ is calculated as 3.45. This was done for all four resists studied: UV210 (Shipley), SEPR-463 (Shin Etsu), UVII-HS (Shipley), and Apex-E (IBM). In some cases the experiment is repeated to ensure that an accurate contrast parameter is extracted. The results of these calculations are summarized in table 2-1. Apex-
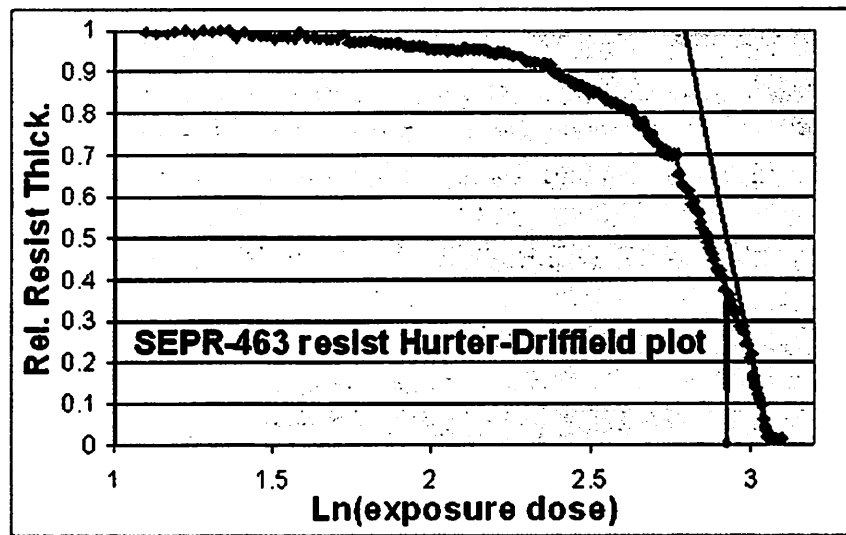
Figure 2-3: A Hurter-Driffield plot for Shin Etsu's SEPR-463 photoresist. In this case, the contrast parameter, γ, is 3.45.

E and SEPR-463 have shal-
lower resist responses, or
resist contrast. In the case of
Apex-E, the resist contrast

	SEPR-463	UV210	Apex-E	UVII-HS
gamma:	3.45	16.7	1.27	5.8

Table 2-1: Summary of resist responses, or resist contrasts, of the four different photoresists studied. Higher γ means sharper photoresist response to exposure.

does not provide sufficient non-linearity, or sufficient gain, to make it a suitable candidate for some of the experiments designed in chapters four and five. This is discussed further in those chapters.

2.3 Creating Various Aerial Image Profiles

In order to study how the aerial image profile affects LER, a mechanism for effectively shaping the aerial image is necessary. The work presented in chapters three, four, and five all use the technique described here to alter aerial image contrast. In this section, three separate aspects of aerial image profile degradation are examined, and specific contrast definitions are used to help quantify these forms of profile degradation. The contrast definitions are standard contrast, as described in equation 2-2; peak intensity contrast, which is simply the maximum image intensity; and slope contrast, which is the light intensity slope at the dose needed to clear the photoresist. Figure 2-4 highlights the different type of aerial image profiles which these three different forms of contrast discern.

A brief discussion of the phenomena which each form of contrast is expected to emphasize is presented. The image slope emphasizes a localized phenomenon dominating LER. Therefore a strong correlation between LER and image slope contrast suggests that forces which blur the image, such as acid diffusion, are not dominant in that resist. Or, more accurately, those forces are not influencing LER significantly and LER is a very localized phenom-
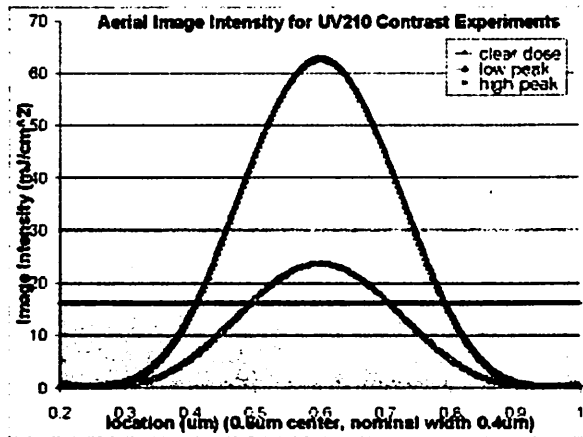
(a)
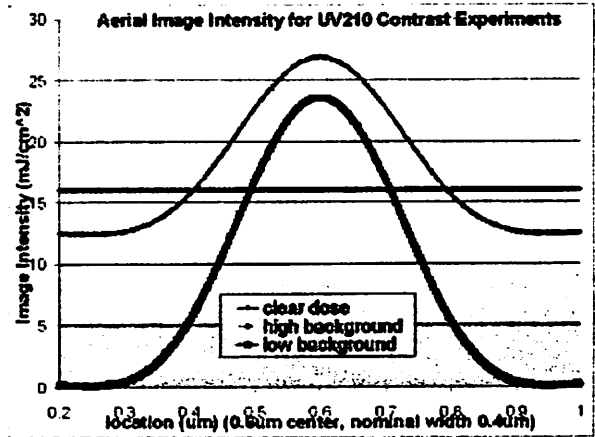
(b)
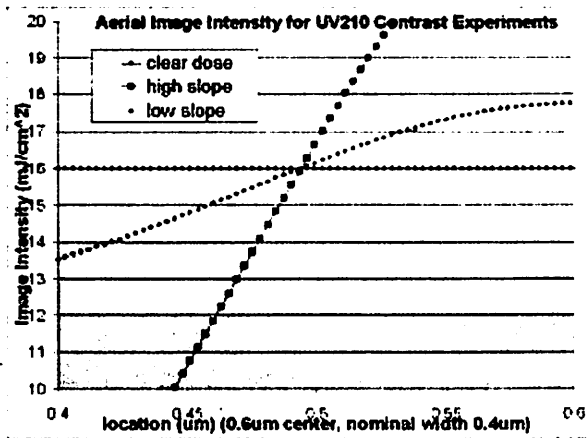
(c)

Figure 2-4: These graphs are aerial image simulations of an isolated trench, or clear area on the mask, with intentionally degraded images (see figure 2-5). They demonstrate the three different phenomena examined with respect to aerial image contrast induced LER. The black line represents the expected low LER case and the grey line the high LER case. (a) compares two peak intensities, (b) compares different background intensities, and (c) compares two image slopes at clear dose.

enon. The standard resist contrast is the only form that includes the background, or minimum, exposure dose. This indirectly suggests that if the correlation between LER and standard contrast is strongest, then background flare is the chief image profile-related cause for LER. Since standard contrast includes both the minimum exposure energy (background dose) and the maximum exposure energy (foreground peak dose), one cannot conclusively prove that a correlation with standard contrast is solely due to background flare. This issue will be discussed further in chapters four and five. The peak image intensity relates to longer range, high acid concentration areas driving LER.

The aerial image profile variations required to emphasize those different forms of contrast are created by two relatively simple methods. The first is to simply over- or underexpose the photoresist, thereby making the aerial image intensity smaller or larger relative to the photoresist clear dose threshold, a physical parameter which cannot be altered without changing the photoresist chemistry, the wavelength of exposure, or some other physical attribute of the system. The typical target for good lithographic exposures is to have the photoresist threshold clear dose equal to 30% of the clear field exposure dose. For example, UVII-HS, one of the photoresists used throughout this work, clears at around 8.1 mJ/cm^2. It's target exposure dose is around 27 mJ/cm^2. This dose leads to the highest NILS for many lithographic patterns [44].
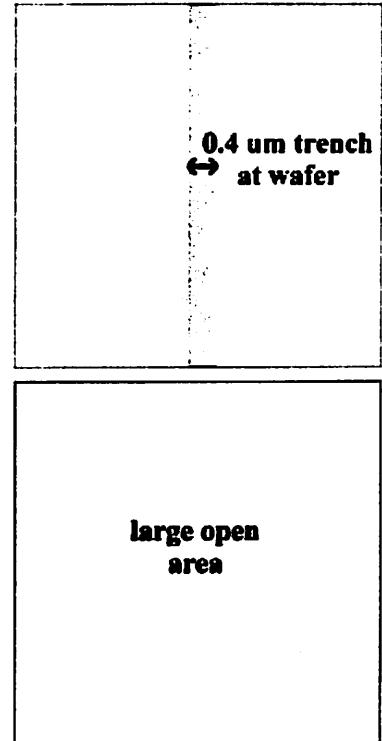
Figure 2-5: mask patterns used. (a) is the "foreground" image, as described in the text. (b) is the "background" pattern.

Overexposing or underexposing will diminish the contrast. Figure 2-4 (a) is an example of altering the aerial image profile by simply over- or underexposing. It mostly affects the peak intensity contrast, but also affects the clear dose slope contrast. The second method of altering the aerial image profile is to apply a background dose, a blanket exposure dose which is seen throughout all areas of the photoresist, both the unpatterned and the patterned regions. Figure 2-4 (b) is an example of this. Applying a background dose affects mostly the standard contrast, but can also affect either the slope contrast, the peak intensity contrast, or both, depending on how the foreground dose is altered. A background dose can be applied by exposing the photoresist with the proper foreground dose and then exposing the same area of the photore-

sist with a blanket clear field dose. Figure 2-5 shows the mask pattern used throughout this study to form lithographic images with varying levels of background dose. The background dose mimics flare which may be present in the lithography tool.

Adding a background dose to the image is especially critical because light scattering, or flare, in lithography tools is expected to be a major concern in extreme ultraviolet (EUV) lithography [53, 59]. As shown by Krautschik et al., flare in an exposure tool roughly scales inversely to wavelength squared. This is because nanoscale bumps in the focusing mirrors (or lenses, in the case of DUV) are two-dimensional and the wavelength of light is one-dimensional. As the wavelength halves, for instance, the number of wavelengths that can "fit" into a mirror deformity increases fourfold. Or, a jump in exposure wavelength from 193 nm to 13.4 nm would lead to a roughly 200 fold increase in flare, all other considerations being equal. All things are not kept equal, however. Mirror smoothness has improved greatly and flare at EUV has already been lowered down to roughly 20%, but at a considerable cost to mirror manufacturing. (Flare with 193 nm tools is on the order of 1% - 2%.) The effects that flare will have upon aerial image profile and LER will therefore become more critical as EUV lithography comes to fruition.

The background exposure method used here mimics flare quite well. There are two significant differences between the background exposure method used here and flare; they both concern the fact that exposure flare happens simultaneously with the proper foreground exposure while the background exposure method used here happens after the foreground has been exposed. (The background dose could also be applied before the foreground; the key issue is that they cannot happen simultaneously in this experimental set-up.) First, two separate light sources that are shone onto an object simultaneously will add their electric fields together.

18

Two light sources that are shone onto an object at different times will add in the intensity domain. When adding intensities, the phase can be discounted and there can be no destructive interference, or subtraction of light intensities. In the electric field domain, phase needs to be considered and destructive interference can occur. However, flare is completely incoherent with itself and with the foreground exposure dose. The light path of flare is such that one can assume the light reaching the wafer from flare is completely incoherent. This is especially true since the ASML stepper, and nearly all other lithography tools, has a diffuser to help create nearly full temporal incoherence. The second issue is the photoresist itself: does the photoresist have "memory"? Are two exposures at 10 mJ/cm^2, for instance, equivalent to one exposure at 20 mJ/cm^2? Yes, two separate exposures are equivalent to one with a dose equal to the two. This is clear if one thinks of how exposure doses are measured. The exposure *power* is never a concern, only the exposure *dose*. For example, a 10 mJ/cm^2 exposure dose that requires 100 ms to expose (0.1 W/cm^2 power) is equivalent to a 10 mJ/cm^2 exposure dose requiring 10 ms to expose (1 W/cm^2 power). Since this is true, the photoresist does have memory, only the overall exposure dose is important, and two separate exposures are equivalent to one exposure equaling the two. This was confirmed with several tests where the clear dose was found with a single exposure and with a foreground and background double exposure. Throughout the dozens of tests run, the single exposure clear dose was equal to the combined double exposure clear doses to within 0.1 mJ/cm^2, which is within experimental error.

2.4 Deprotection Reactions and Kinetics, and Photoresist Development

The actual mechanism in which irradiated photoresist is developed has not been discussed yet. This section will discuss the post-exposure lithographic processes: post-exposure bake

19

and development. As mentioned in chapter one, photogenerated acid diffusion and image amplification due to development create a photoresist profile that is often different from the original aerial image profile. Molecular model simulations confirm this statement [71]. Since the photoresist profile does not match the aerial image profile, it is essential to understand the post exposure processes that create the photoresist profile in order to understand how the image profile affects LER.

Before moving to the post-exposure processes and the models behind them, a brief introduction to typical DUV chemically amplified lithography is given. Ito, Willson, and

Figure 2-6: A typical photoresist resin: polyhydroxystyrene with t-butoxycarbonyl protecting groups.

Frechet first devised chemically amplified resists in the early 1980's as a means to provide resists that were highly sensitive to lithographic exposure in order to maintain the high process throughput the industry had come to expect [37]. Positive tone chemically amplified photoresists generally consist of 3 key components and possibly one or more additional components: the resin, the photoacid generator (PAG), and the solvent. The additional components typically include a quenching base, an emulsifier, and possibly other additives. The resin gives structure, etch resistance, and provides the protected groups that will allow pattern transfer. Figure 2-6 is an example of a typi-

Figure 2-7: A typical PAG: trisulfonium arsenic hexafluoride salt. Upon irradiation, the trisulfonium ion breaks apart, releasing H^+.

cal photoresist resin: poly(4-hydroxystyrene) (PHOST) with tert-butoxycarbonyl (tBOC) protecting groups. It is a polymer with a fairly non-polar portion, the tBOC, and a fairly polar portion, the hydroxy group. The left-hand monomer is that containing the tBOC. The photoacid generator, upon irradiation at the lithographic wavelength of light, generates the initial protonic acid that provides catalytic deprotection. Figure 2-7 is the chemical structure for a typical PAG, a trisulfonium salt. The solvent provides low viscosity initially to allow the resist to be spun onto the wafer. It is evaporated during pre-exposure bake so that the photoresist can become solid once on the wafer. Propylene glycol methyl ether acetate (PGMEA) is a common solvent. The quenching base limits the potential diffusion length of the catalytic acids by neutralizing them upon contact. Emulsifiers, if present, help to keep the entire mixture well dispersed.

A potential source of LER within the photoresist can already be identified: the photoresist resin. The photoresist resin consists of linear polymer molecules. The size of such a molecule is quite large, relative to the desired LER of photoresist. A linear polymer molecule is a long chain of monomer units. In practice, the size is typically measured in a dilute solution. In solution the polymer is free to assume the lowest free energy state, one which resembles a wrapped ball of twine, or more accurately a wrapped ellipsoid of twine. In this state, linear polymer molecules of the size used for photoresist resins, approximately 2000 g/mole number averaged molecule weight, are around 5 nm in diameter. Photoresist solutions ready for application, however, are concentrated. The polymer resin is further concentrated during pre-exposure bake when the solvent is almost entirely evaporated. At this point it is impossible to measure the diameter of the photoresist resin polymer; it has been stretched and distorted into oblong shapes. More importantly, many polymer molecules intertwine within each other to

21

form very large, complex aggregates of polymers. As the size of the photoresist increases, the likelihood of larger aggregates increases. Lin addresses the effects of increasing polymer size upon LER in reference [7], showing that resists with smaller polymer molecular weights demonstrate less LER.
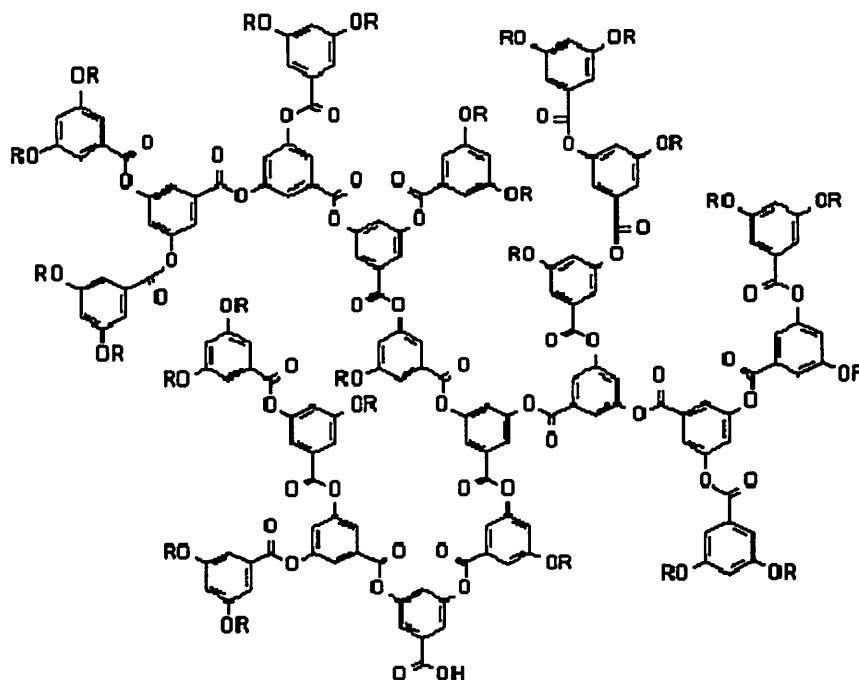
Figure 2-8: Chemical structure of a portion of a dendritic polymer with 3 active growth sites.

Tully takes another approach to limit photoresist resin size, creating photoresist resins that are dendritic in shape [35]. Dendritic polymers are derived from monomers that have three or more active growth sites instead of two; see figure 2-8. The extra growth site yields polymers that are ever-branching, eventually forming spheres with the highest atomic density at the periphery of the polymer, as shown in figure 2-9. Due to this structure, the spherical shape is maintained and polymer entanglement is minimized. LER studies comparing a dendritic polymer resist to standard linear polymer resists demonstrate that the LER of the dendritic resist is significantly smoother than that of standard linear resists [9, 34]. Unfor-
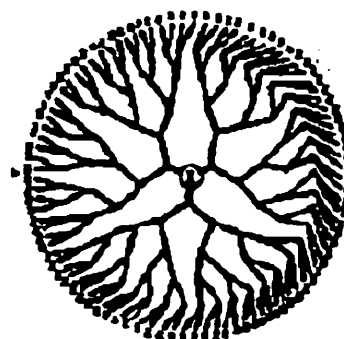
Dendrimer molecule

Figure 2-9: sketch of a dendrimer molecule.

tunately, other problems such as exposure sensitivity and manufacturing difficulties reduce the potential for dendritic resists to enter the marketplace.

Moving on to the deprotection and development processes, these occur during the post-exposure bake (PEB). The photoresist profile is define during these steps. During PEB the

$$CO_{2(g)} + H_2C = C(CH_3)_{2(g)} +$$

$$H^+ +$$
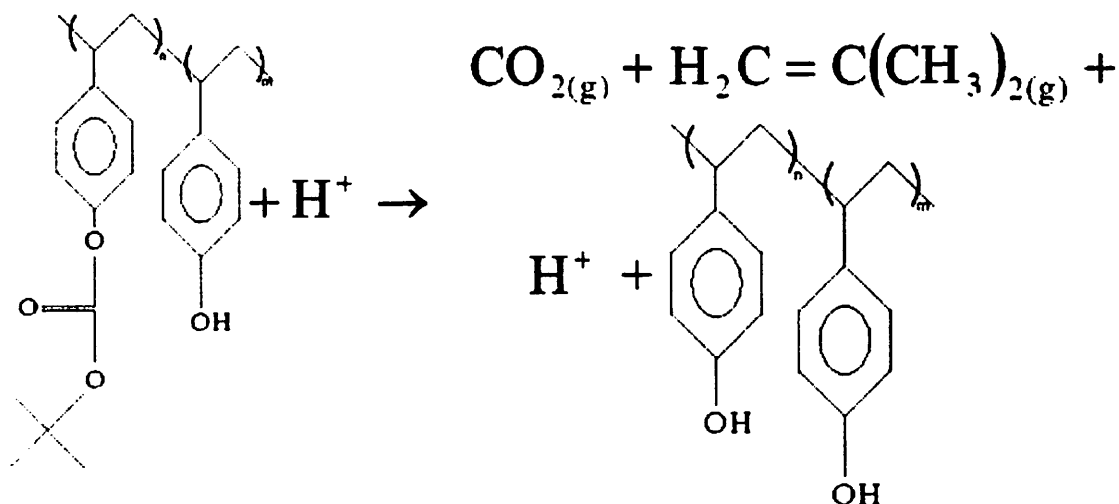
Figure 2-10: The tBOC deprotection reaction. The protonic acid is recovered and the polymer product is significantly more polar than the polymer reactant.

photogenerated acid diffuses to protected groups, such as the tBOC shown on the left hand side of figure 2-6. The acid reacts with the tBOC to produce a hydroxyl group-- now a deprotected group-- and gaseous products that are removed. Figure 2-10 shows the simplified chemical reaction. There are two key features of the reaction: it is catalytic, and the polymer product is significantly more polar than the polymer reactant. The former allows significantly greater exposure sensitivity, allowing much quicker exposure times and faster throughput. The latter allows the deprotected product polymer to dissolve in the developer while the protected polymer remains insoluble.

Photogenerated acid diffusion during the post-exposure bake process has received a great deal of study. The success of optical lithography at deep ultra-violet (DUV) wavelengths

owes itself to the chemical amplification (CA) process [37], and at the heart of CA is photogenerated acid diffusion. Chemical amplification is a multi-step, catalytic reaction chain. In the first step, light shines upon a photo-acid generator (PAG) which then emits a protonic acid (H^+), referred to as a photogenerated acid. Next the H^+ diffuses to polar protected groups on the photoresist resin. The acid then reacts with the protected group as shown in figure 2-10, deprotecting it. This reaction is catalytic: another H^+ is emitted and can react with another protected group. This catalytic process allows much greater sensitivity to the original lithographic exposure. Relatively few PAGs need to be exposed in order to allow sufficient deprotection. However, the lithographic pattern is largely defined by the ability of the H^+ to diffuse towards protected sites. Since protonic acid is the smallest atom possible, it has a fairly high diffusivity. Postnikov, Croffie, Cheng, and Yuan have all shown that photogenerated acid diffusion needs to be taken into consideration in order to properly predict the true location of the line edge [30-32, 36].

Acid diffusion and reaction can be mathematically modelled quite readily. Equations 2-5 and 2-6 represent the deprotection reaction rate and the photogenerated acid diffusion rate, respectively.

Equation 2-5:

$$\frac{\partial A}{\partial t} = k_r (P_0 - A) H$$

Equation 2-6:

$$\frac{\partial H}{\partial t} = \nabla (D \nabla H) - k_l H$$

where A is the concentration of deprotected sites, H is the concentration of acids, D is the diffusivity of acids, P_0 is the initial concentration of protected sites, k_r is reaction rate, and k_l is the acid loss rate. Equation 2-5 is coupled with equation 2-6 since the rate of deprotection is dependent upon the concentration of photoacids. Equation 2-6 represents the diffusion of photoacids assuming Fickean diffusion. In the case of Fickean diffusion, the acids are not affected by neighbors. The acids will follow a random walk pattern in all directions. The change in acid concentration over time is simply due to concentration gradients: acids in a high acid concentration area will "move" over to a low acid concentration area simply because statistically there are more acids following a random walk in the high acid concentration area than in the low acid concentration area. Fickean diffusion is of course well understood; the author mentions this merely to emphasize that concentration **gradient** is the only key parameter in this case.

Photogenerated acids do not move via simple Fickean diffusion, however. Postnikov et al. use an elegant experimental design to show that photo-generated acids diffuse much more slowly through a polar species than through a more non-polar species, in their case poly(4-hydroxystyrene) and poly(t-butyloxycarbonyloxystyrene), respectively [30]. As the polar species concentration rises, the diffusivity of the H^+ decreases. The protonic acid is thermodynamically more attracted to a polar environment, hence it cannot diffuse away from that environment as easily. This type of diffusion is called reduced, non-Fickean diffusion. Yuan and Croffie, using predictive simulations and fits to experimental data, demonstrate that acids experience reduced non-Fickean diffusion [31-32]. The diffusivity, D, used in equation 2-5 is not constant in this case. It is described as follows:

Equation 2-7:

$$D = D_0{}^{\omega A}$$

where D is the diffusivity, D_0 is the pre-exponential diffusion constant, A is the concentration of activated or deprotected sites, and ω is the non-linearity parameter. D_0 is now constant. In the case of reduced non-Fickean diffusion, ω is negative so that as the concentration of deprotected sites increases, the diffusivity of acid decreases. All these studies come to the same conclusion: diffusion of acid is limited when the medium becomes more polar. Since acid diffusion helps to determine the line edge, it obviously affects LER as well. Yuan is currently creating models and simulations to predict how acid diffusion affects LER [75].

After post-exposure bake, the wafer is dipped into developer, or perhaps the developer is sprayed onto the wafer. Due to the high solubility of deprotected polymer and the low solubility of protected polymer resin, one generally assumes the development process perfectly dissolves all the sufficiently deprotected resin and leaves all the sufficiently protected polymer intact. The only concern is whether the developer is sufficiently able to reach all the regions of the photoresist film where deprotected polymer is present, and whether phase separation is induced during development.

Phase separation can occur during post-exposure processing because the protected non-polar polymer species and the polar polymer species are in proximity in an aqueous developer. The non-polar species is referred to as protected-- it is protected from dissolution by the developer-- and the polar species is deprotected. The thermodynamics of the protected/deprotected mixture is similar to that of the oil and water mixture example given in chapter one:

protected polymers are thermodynamically driven to separate from deprotected polymers. Research by Lin and Yasin has shown that phase separation does indeed occur during post-exposure processing and that it leads to LER [7-8, 12, 39, 70]. Immediately before lithographic exposure, photoresist resin molecules are intertwined within each other. Little solvent is present, the photoresist consists mostly of polymer molecules intertwined and folded onto themselves. After exposure and post-exposure processing, some of the polymer molecules have been sufficiently deprotected so that they are now miscible in the aqueous developer and immiscible with protected polymers. At the line edge, neighboring protected and deprotected polymers begin to phase separate, with each of them forming little clumps of like polymers. When the deprotected polymer is dissolved into the aqueous developer, clumps of protected polymer along the line edge remain, creating LER. Lin et al. examined the effects of changing the photoresist solvent, photo-acid generator (PAG), and base additive upon phase separation and therefore LER [12]. However, the extent to which phase separation occurs during lithographic processing is minimized due to kinetics. Phase separation is driven by diffusion, and the diffusion of a polymer molecule within a polymer lattice is quite slow.

How does the aerial image profile tie into all of this; how does it affect LER? A poor lithographic image profile, such as the upper left graph of Figure 2-2, creates a photogenerated acid distribution that is slowly varying. This acid distribution in turn creates a deprotected polymer distribution that slowly varies, as predicted by the reaction and diffusion equations above. However, those equations do not tell the whole story. Those equations assume a continuous distribution and an initial PAG and photogenerated acid concentration that is large enough to be considered continuous. This is not true. There is a finite number of

27

PAGs yielding a finite number of photogenerated acids. These acids then diffuse following the models above, but with a quantized distribution. Diffusion and kinetics are statistical processes. The models will predict the average deprotected site concentration, but chance will provide the opportunity for some regions to have larger or smaller deprotected site concentrations than predicted. These differences are negligible when the aerial image profile is good: the average deprotected site concentration quickly rises or falls as the distance from the line edge increases. This is not true when the aerial image profile is poor. The statistical nature of the lithographic process becomes increasingly important, and its effects can be seen in a line with an edge that deviates significantly from its intended position.

2.5 Summary

A brief synopsis of the lithographic imaging and pattern transfer process was presented. It demonstrated that supplying a simple aerial image contrast value is not sufficient to fully describe the entire aerial image profile. Further knowledge of the profile may be needed when attempting to understand how the aerial image profile affects LER.

The contrast parameters for the four different photoresists used throughout this thesis were calculated. The contrast parameter is the response of the photoresist to exposure dose. An ideal photoresist has a very non-linear response so that it may act as an image amplifier. The response of Apex-E was insufficiently non-linear to allow it to be used for study in chapters four and five.

The mechanisms for creating the various aerial image profiles used throughout this study were discussed next. They involved either over- or underexposing the photoresist, and altering the ratio of foreground to background dose. The background dose mimics flare in the

lithography tool, expected to be a large concern at EUV lithography. Three different metrics for quantifying the aerial image profile were discussed: standard contrast, clear dose slope contrast, and peak intensity contrast.

The chemistry of the photoresist, the deprotection reaction, and development were discussed next. The components of a photoresist were enumerated, providing an example of each and explaining their utility. Due to its potential affect upon LER, the photoresist resin was discussed in further detail. The deprotection reaction was discussed, noting the key features: a catalytic process that converts a non-polar molecule into a polar molecule. The diffusion and reaction equations governing deprotection were presented, noting that protonic acid diffusion follows reduced, non-Fickean type II diffusion. Photoresist development was mentioned, focusing upon phase separation which may occur during development. Finally, the statistical nature of deprotection was discussed, which elucidates the connection between aerial image contrast and LER.

3 Side Edge Roughness and Analysis in the Frequency Space Domain

Side edge roughness (SER), as discussed in this chapter, is measured with an atomic force microscope (AFM). SER is measured throughout the photoresist edge in two dimensions, as shown in Figure 1-3. It is called SER to distinguish it from the top-down one dimensional roughness data. SER data and LER data correlate with each other quite well [16, 23], however some elements of roughness can be observed more easily with an AFM. Most of the effort here is to determine the height and size of side edge roughness and understand the physical nature of roughness. An initial study of how image contrast affects LER is presented, along with some discussion of how SER and LER compare with each other.

Roughness data tend to have a large level of noise and many samples need to be measured over many locations in order to obtain data that is statistically significant enough to be able to extract confident conclusions. Due to the relative difficulty in obtaining AFM measurements and the strong correlation between LER and SER, the pool of SER data collected are not of a sufficient size to be able to reach any solid conclusions regarding how image contrast affects LER. This is studied in further detail in chapters four and five.

3.1 Side Edge Roughness Experimental Set-Up

An ASM Lithography Pas 5500/90 248 nm wavelength stepper with 5 x demagnification is used for all exposures. The numerical aperture (NA) is 0.52; the partial coherence, σ, equals 0.62. The mask pattern transferred is shown in Figure 3-1. The pattern labeled "AFM

analysis" is a square block 40 μm by 40 μm and an 80 μm by 80 μm block as the background. After exposing the wafer once, the mask is shifted so that the background and the fore-ground exposure areas on the wafer are placed atop each other, then the wafer is exposed again. Figure 3-2 shows an SEM of a wafer exposed with the pattern just once, in order to view the pattern. After exposing the wafer

Figure 3-1: Mask pattern used for double exposure method. White is an open area, black is chrome. The smaller square is 40 μm x 40 μm on the wafer, the larger square is 80 μm x 80 μm.

twice, it is then post-exposure processed according to resist type as discussed below. More details of the pattern design are described below.

Image contrast levels are determined using SPLAT [40, 45], a lithographic imaging simulator. This simulator uses the theory of partial coherence imaging in its calculation. The numerical aperture and sigma used for all simulator calculations match those of the stepper tool, 0.52 and 0.62, respectively. Exposure dose ratios range from 100% foreground : 0% background (1:0) to 67% foreground:33% background (2:1) to 50% foreground : 50% background (1:1). Contrasts of these exposure doses range from 1.0 to 0.54 to 0.37, respec-

Figure 3-2: SEM image of exposed resist showing the "AFM analysis" mask pattern. The upper image is the background and the lower image is the foreground

tively. The exposure doses are chosen such that they would all cross at the same intensity level, 0.74, at the same location along the cutline.

The four commercial deep ultra-violet (DUV) 248 nm CARs studied are: UV210 (Shipley), SEPR-463 (Shin Etsu), UVII-HS (Shipley), and Apex-E (IBM). The two older resists, UVII-HS and Apex-E, are spun to around 8500 Å thickness. The two newer resists, UV210 and SEPR-463, are spun to around 4000 Å thickness. UVII-HS is pre-baked at 140 °C for 60 seconds. It is post-baked at 140 °C for 90 seconds, then developed in 1:1 solution of water:MF-312. UV210 is pre-baked at 130 °C for 60 seconds. It is post-baked at 130 °C for 60 seconds, then developed in LDD-26W for 45 seconds. Apex-E is pre-baked at 100 °C for 60 seconds. It is post-baked at 90 °C for 60 seconds and developed in 3:2 solution of water:MF-312. SEPR-463 is pre-baked at 110 °C for 90 seconds. It is post-baked at 100 °C for 90 seconds, then developed in SSFD-238. All wafers are coated with AR3-600 (Shipley) bottom anti-reflective coating (BARC) before applying the resist. Apex-E, due to its envi-
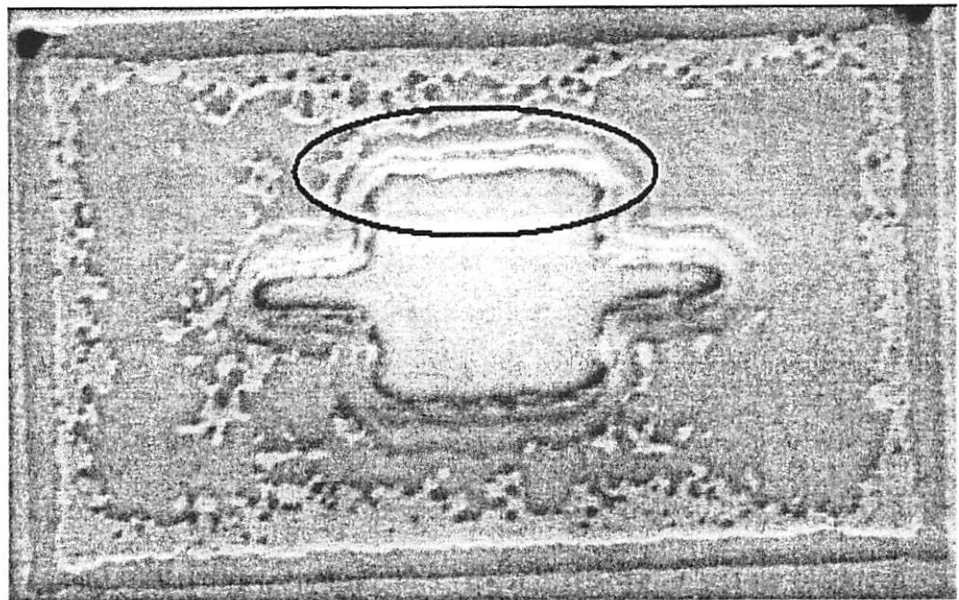
Figure 3-3: SEM image of UV210 photoresist exposed without anti-reflective coating. This sample was double exposed and overexposed, explaining the problems in the outer regions. At the edge of the patterned region, like the region circled, it is quite clear that there is a step-wise "tier" of photoresist. This is due to regions of higher and lower exposure orthogonal to the surface, characteristic of standing waves within the photoresist.

ronmental sensitivity, is coated with IBM's RTC top coat after coating the resist. The BARC coating, aside from preventing reflection and standing waves, acts as an adhesion promoter similar to hexamethyl disilazane (HMDS). Therefore the wafers were not coated with HMDS, nor were any of the wafers used throughout this study.

Initial studies not included in this dissertation did not use the BARC coating and instead simply used HMDS as the adhesion promoter. With the ASML 5500/90 248 nm stepper tool mentioned earlier, severe standing wave effects were clearly visible. Figure 3-3 is an SEM image of exposed UV210 photoresist that did not use the BARC coating. Since the same exposure tool was used, and always under similar lithographic conditions, Shipley's AR3-600 was adopted into the resist coating procedures early in the studies to prevent standing waves.

The AFM used is a TopoMetrix Explor 2000, slightly modified to help reduce external noise. A ThermoMicroscopes Supertip probe tip (with less than 20 nm tip radius of curvature) is tilted 12° in order to allow measurement of the pattern edge
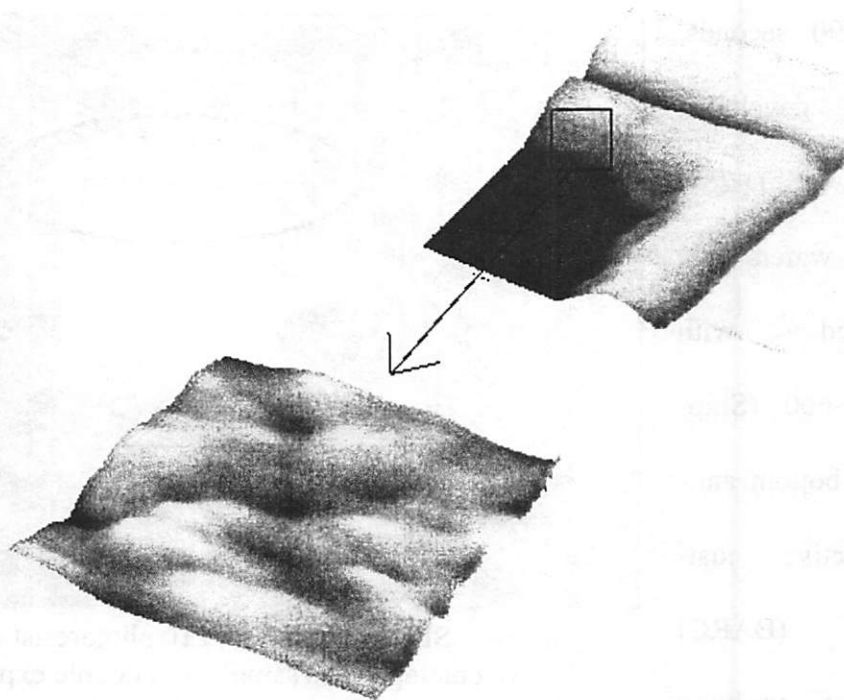
Figure 3-4: AFM SER measurement technique. A tilted probe tip measures data to create the image on the right. The highlighted box is then flattened to give an image analogous to that of top surface roughness (TSR) imaging.

without any interference due to the shape of the probe tip itself. The tip is always tilted at the same angle, in order to avoid systematic noise introduced by trigonometric effects [26]. Figure 3-4 demonstrates how the sidewall profile measurement data is then flattened, to give a set of data similar to that of top surface roughness (TSR) imaging.

3.2 Side Edge Roughness Results

The correlation between root mean square (RMS) SER and image contrast via AFM metrology is shown in figure 3-5 and table 3-1. The pattern used is the "AFM analysis" pattern shown in figures 3-1 and 3-2. Contrast is varied from 0.37 to 1.0. RMS roughness data is the 3σ deviation, calculated as follows:

Equation 3-1

$$RMS = 3 \times \sqrt{\frac{\sum_{i=1}^{n} |z_i - z_{ave}|^2}{n}}$$

where z_{ave} is the calculated point to which the perfectly straight line or flat plane is fit, z_i is the actual point where the edge is found, and n is the number of points being measured. Since the sampling size is generally so large, it is insignificant whether n or (n-1) is used in the denominator to determine RMS.

All AFM images are first flattened, such that deviations from z_{ave} arise solely due to roughness. Figure 3-5 and table 3-1 both contain the same volume of data. Dif-
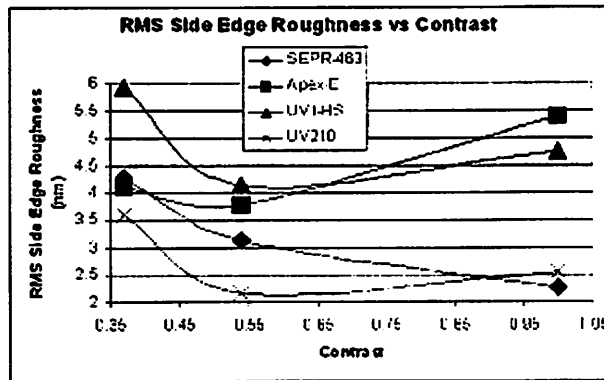
Figure 3-5: Measured correlation between SER and contrast for varying resists.

34

ferent formats are presented simply to different modes of view. In all four cases, the roughness does show a tendency to increase at the 0.37 contrast level, but the results are inconclusive. The overall standard deviation given in table

Contrast	SEPR-463 SER (nm)	Apex-E SER (nm)	UVII-HS SER (nm)	UV210 SER (nm)
1	2.25	5.4	4.74	2.55
0.54	3.12	3.78	4.14	2.16
0.37	4.26	4.11	5.94	3.6
overall standard deviation:			0.41	

Table 3-1: Measured correlation between SER and contrast for varying resists. The data herein are identical to figure 3-4.

3-1 is determined by collecting TSR measurements on unexposed areas from 28 runs total, with all four resist types. Due to the cost and difficulty of obtaining SER measurements, these repeat measurements are obtained using TSR data instead. In general, TSR correlates well with SER, however TSR is somewhat smoother [26]. Therefore, the overall standard deviation may be slightly larger than 0.41 nm.

3.3 Side Edge Roughness Shape

A benefit to measuring two dimensional SER instead of LER is that it allows a better opportunity to examine the shape of the roughness. The shape of the edge roughness is consistently transferred to the underlying layer. In the case of a relatively "soft" chemical etch with low ion bombardment energy, the underlying layer being etched will have some two dimensional topography. A "hard" physical etch with higher ion bombardment energy will create an underlying layer with one dimensional roughness that resembles tall "walls" of roughness [46]. Figure 3-6 shows a cartoon of the roughness of an underlying layer after the layer
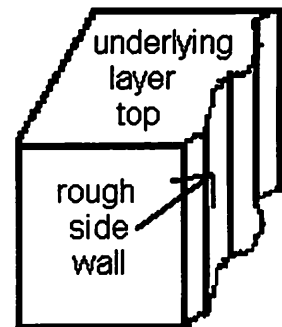
Figure 3-6: cartoon of pattern transfer during hard physical etching process.

35

Whole Image
Area Ra: 1.5156 nm
Area RMS: 1.9790 nm
Avg. Height: 9.9389 nm
Max. Range: 25.2643 nm

Partial Image
Area Ra: 1.3058 nm
Area RMS: 1.5652 nm
Avg. Height: 11.1112 nm
Height. Max 8.5803 nm
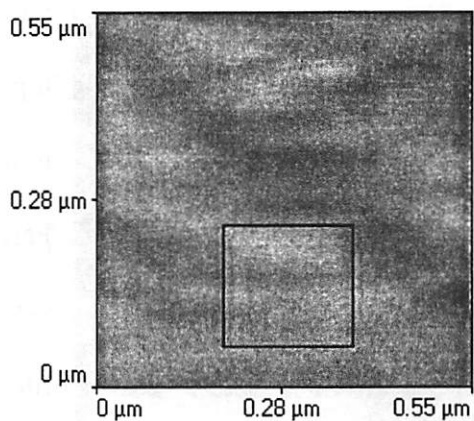
Figure 3-7: A typical AFM measurement for side edge roughness analysis.

was coated with photoresist and patterned with some level of roughness, then that roughness was transferred one dimensionally via a physical etch.

A typical AFM side edge roughness measurement is shown in figure 3-7. In this example the photoresist is UVII-HS. Twelve such roughness measurements were collected for figure 3-5 and table 3-1. Since the measurement can only be collected along the sidewall, the measurement area is relatively small compared to the distance between peaks and valleys of roughness. This prevents accurate analysis of the photoresist roughness in the frequency plane or Fourier plane. In order to look in the frequency domain, a

Whole Image
Area Ra: 2.1457 nm
Area RMS: 2.7404 nm
Avg. Height: 10.6041 nm
Max. Range: 25.9793 nm

Partial Image
Area Ra: 1.8900 nm
Area RMS: 2.4143 nm
Avg. Height: 10.5882 nm
Height. Max 17.1606 nm
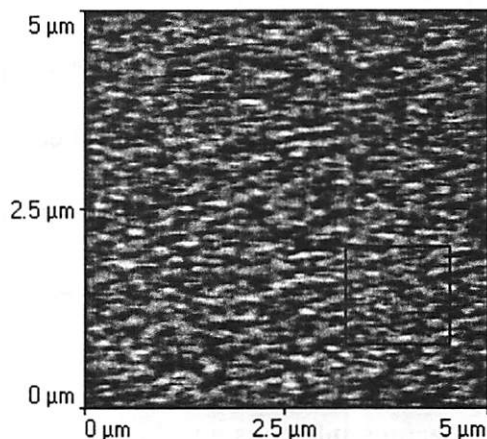
Figure 3-8: Typical AFM measurement used for TSR measurement.
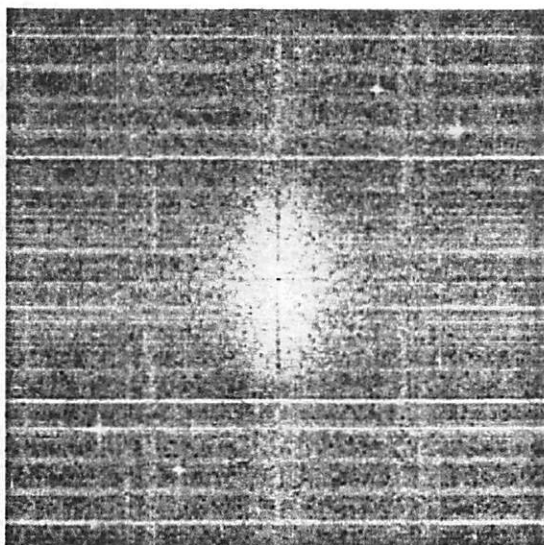
Figure 3-9: Typical AFM TSR measurement in frequency plane. White is high density of frequencies, black is low density.

larger measurement area is needed. Therefore, a top surface roughness (TSR) AFM measurement is made after partial exposure of the photoresist. Partial exposure of the photoresist yields a top surface that mimics the photoresist sidewall profile [47]. Figure 3-8 is an example of the partially exposed TSR measurement, with UV210 photoresist. This image is next analyzed in the Fourier plane, or frequency domain, as shown in figure 3-9. This plot shows the density of surface roughness peak to peak distances, where the center of the plot is peak to peak distance equal to 0. As can be seen, there is no peak to peak distance equal to 0, but the majority of peak to peak spacings are randomly distributed close to the center. This plot can be more easily examined in a bar chart separating peak to peak spacings into bins and looking at the average roughness peak to
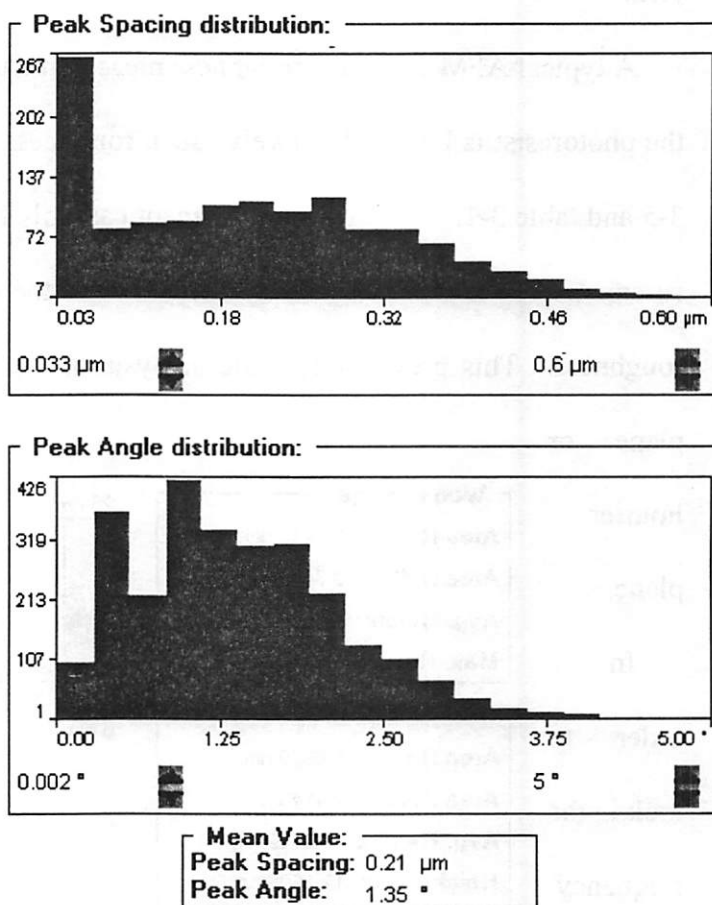
Figure 3-10: Bar chart listing peak to peak spacing and peak angle on a typical AFM TSR measurement.

peak distance. An example of such a chart with SEPR-463 photoresist is shown in figure 3-10. This figure is typical of the results gathered: the peak to peak spacing appeared bimodal, with many of

TSR Average Peak:Peak Spacing

Apex-E (nm)	SEPR-463 (nm)	UV210 (nm)	UVII-HS (nm)
143.3	160	130	93.3

Table 3-2: Average peak to peak roughness spacing.

the spacings less than 30 - 50 nm, and the other mode distributed around values 5 or 6 times

larger. Five TSR measurements were taken with each of the four photoresists. A summary of

the average peak to peak roughness spacing is shown in table 3-2.

The peak to peak spacing is on the order of 130 nm, the peak angle is typically around 1°,

and the average peak to valley height is around 3 - 4 nm. The average roughness height is

somewhat less than the RMS height given above. The partially developed photoresist top sur-

face roughness has a shape that resembles low, constantly rolling hills with a very mild slope.

There are very few sharp "spikes" in the roughness, nor are there any exceedingly flat areas.

As mentioned above, the partially developed TSR shape can be assumed similar to the SER.

3.4 Conclusions

The photoresist roughness peak to peak spacing ranged from 93 nm - 160 nm, however the

distribution was bimodal. One mode occurred below 50 nm and the other around 250 nm.

The spatial frequency spectrum of roughness showed random distribution. This information,

combined with very shallow peak angles, confirms low, rolling hills of roughness. There was

insufficient data collected comparing aerial image contrast to SER in order to make any strong

conclusions, but it was fairly clear that at very poor contrasts SER began to increase rapidly.

Further studies of aerial image contrast versus roughness were held to chapters four and five, which utilize LER data collected from SEMs.

4 Line Edge Roughness versus Aerial Image Contrast: Multiple Resists

The work presented in chapters four and five focuses on understanding *how* aerial image contrast affects line edge roughness. Previous studies have already shown that poor aerial image contrast can increase LER, and that increased LER at the gate layer can increase leakage current [2-3]. In this chapter, three separate aspects of aerial image profile degradation are examined, and contrast definitions are created to help quantify these forms of profile degradation. The contrast definitions are standard contrast, as described in equation 2-2; peak intensity contrast, which is simply the maximum light intensity; and slope contrast, which is the light intensity slope at the dose needed to clear the photoresist. Figure 2-4 demonstrates the different type of aerial image profiles that these three different forms of contrast highlight.

In this section three resists, UVII-HS, SEPR-463, and UV210, are studied in more detail with 9 contrast levels on each resist and 4 measurements at each treatment. IBM's Apex-E is removed from the work because it exhibits a high level of diffusion [32] and the contrast parameter, γ, is quite low, as calculated in chapter two. It impossible to examine all the degraded levels of contrast that can be studied for the other photoresists. Unexpected results lead to some interesting questions regarding the impact of background flare upon LER. The experiment described in chapter five addresses these questions.

4.1 LER versus Contrast (Multiple Resists) Experimental Set-Up

The mask patterns used to create the varying levels of contrast in this chapter are shown in figure 2-5. Pattern (a) in the figure is the foreground pattern. Grey corresponds to an open area in the mask and white is chrome. The pattern creates a 2.0 μm wide and several mm long trench on the mask, creating a trench nominally 0.4 μm wide at the wafer level. Pattern (b) is the background pattern. It is merely a large open area to simulate background flare, 100 μm wide at the wafer. The wafer is exposed to pattern (a), then immediately afterward exposed to pattern (b). Then the wafer is post-exposure processed as described earlier. By varying the overall dose and the ratio of foreground to background dose, the aerial image profile is changed to highlight varying levels of the three contrast types examined: standard contrast, slope contrast, and peak intensity contrast.

Image contrast levels are determined using SPLAT [40], a lithographic imaging simulator, with all optical parameters such as wavelength and partial coherence matching those of the lithography tool as given in section 3.1. Simulation of the trench with flare pattern shown in figure 2-5 is performed. Looking back upon figure 2-4, the examples generated show image intensities along a cutline of the trench as contrast is varied. Figures (a) and (b) show the entire aerial image along the trench and (c) shows a zoom near the photoresist clear dose, or the expected trench
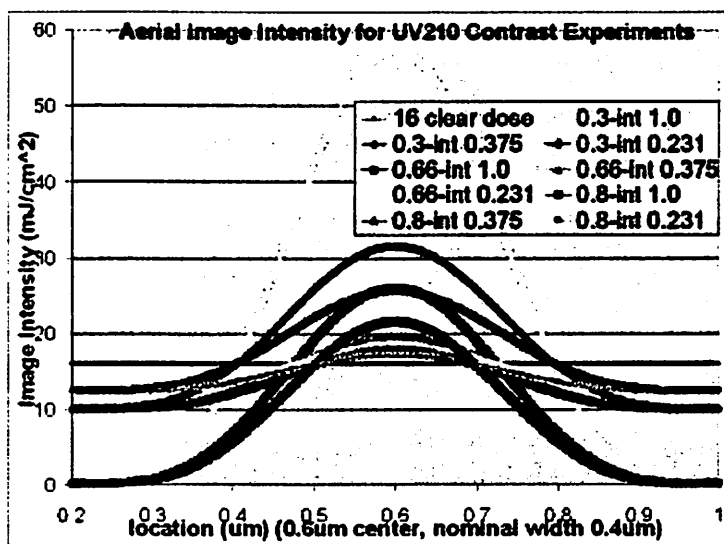
Figure 4-1: The 9 different aerial image profiles examined with UV210. 27 profiles total were examined, 9 with each resist.

41

edge. Figure 4-1 is a SPLAT simulation of the different image intensities used for the UV210 photoresist experiment correlating aerial image contrast with LER.

UV210, SEPR-463, and UVII-HS are studied. All three resists are spun to 6200 Å. Aside from the change in resist thickness, all resist coating and resist developing procedures are the same as those described in chapter three. Exactly as in chapter three, all wafers are coated with AR3-600 bottom anti-reflective coating (BARC) to prevent standing waves forming within the resist.

Images used to determine LER data are collected using a Leo 1550 SEM, at 2 kV electron acceleration voltage with a 30 μm aperture and pixel noise averaging techniques to reduce image noise. All SEM images, in this chapter and throughout the study, are collected at 50,000 x magnification unless otherwise noted. This magnification yields a 2.22 nm pixel size. Therefore, a perfectly straight line caught at a pixel edge can theoretically yield a 3 sigma root mean square (RMS) roughness as large as 3.33 nm. The LER of the trenches in these images is measured using 3 sigma RMS LER calculations. First, the contrast of the image is enhanced to determine the most exact line edge. Then a perfect line is fit to the trench wall and deviations from that perfect line are tabulated as the roughness. To clarify, LER data is presented here, not line width roughness (LWR). The former measures each wall of the trench separately, yielding deviations from the idealized perfect wall edge. The latter measures the width of the trench, or the space between each wall, and looks at deviations from the perfectly smooth line width. The calculations needed to enhance the contrast and extract LER are accomplished via MatLab code (version 6.1, The MathWorks, Inc.) created by the author.

This method to extract LER from SEM images is the same one used throughout all of this paper, including the work in chapter seven examining the LER of deblurred SEM images. An SEM image is captured, like figure 4-2 (a). A threshold contrast limit is set where all pixels brighter than that level are set to white and all pixels dimmer are set to black. Image 4-2 (b) is the result of this contrast enhancement. The coordinates at which the contrast switches from white to black are recorded. Figure 4-2 (c) represents the coordinates. A least squares linear fit to these coordinates is made. The line edge roughness is then determined by equation 3-1, where z_{ave} is the coordinate calculated via the least squares linear fit. The trenches examined

throughout this entire study are extremely long. The trenches are 0.4 μm wide and 6 mm long. These patterns allow simple linear fit-ting to the line edge. A perfectly printed trench with no LER will print a perfectly straight trench edge, or line edge, since the trench is so long rel-ative to the wavelength of exposure light. Thus, a simple linear fit to a
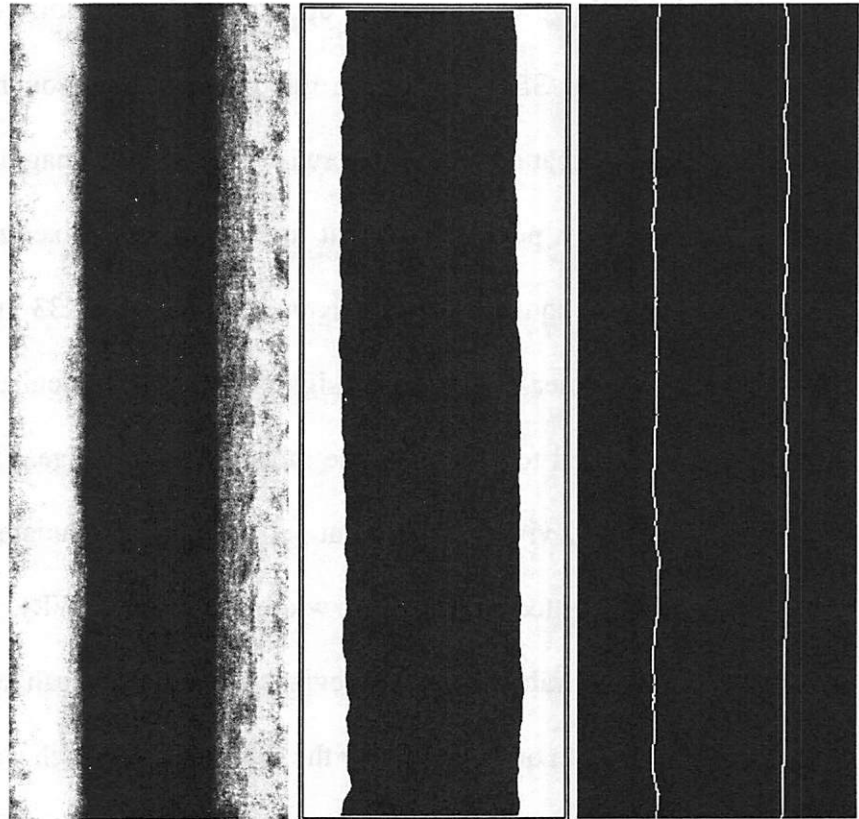
(a) (b) (c)

Figure 4-2: Method to calculated LER in this study. The SEM image, figure (a), is converted to image (b) by setting a thresh-old contrast level. Image (c) contains the coordinates of image (b) that transition from white to black. LER is calculated from image (c).

perfectly straight line is sufficient and LER is measured as a deviation from this perfectly straight line. All the measurements made for the aerial image contrast studies are conducted in this fashion. *Appendix B* contains the Matlab code created to enhance the contrast and calculate the LER.

4.2 LER versus Contrast (Multiple Resists) Results

The correlation between LER and 27 different aerial image profiles is examined. This leads to 27 different image slopes at clear dose energy level, 27 different peak image intensities, and 7 different standard resist contrast levels. Figure 4-2 shows all of the different aerial image profiles examined for UV210 resist. Table 4-1 lists all of the different standard contrasts, peak intensities, and image slopes examined for all three resists.

With the wide range of aerial image profiles examined, a wide variation in LER is seen between the vastly different profiles and a statistically significant correlation between LER and aerial image profile is detected. In the case of UVII-HS, for instance, the worst 3 sigma RMS LER is 24.8 nm and the best is 3.6 nm. The former occurs at 0.168 standard contrast, 12.6 $(mJ/cm^2)/\mu m$ slope at clear dose, and 8.7

Clear Dose Slope $(mJ/cm^2)/um$	Peak Intensity (mJ/cm^2)	'contrast' (max-min)/ (max+min)
12.6	9	0.181
15.2	9.5	0.21
20.5	9.6	0.306
24.7	10.5	0.347
24.9	11.9	0.37
30	13.6	0.54
33.7	14.4	0.997
34.6	17	
40.5	17.7	
41.7	18.8	
48.8	18.9	
54.7	20.6	
54.7	23.6	
56.2	24.6	
65.9	26.1	
66.5	26.2	
67.7	26.8	
92.3	28.4	
108	28.6	
108	31.8	
130.1	32.7	
145.8	33.6	
149.8	37.2	
149.8	39.4	
180.5	46.6	
287.9	62.8	
399.5	87.2	

Table 4-1: A summary of all of the different contrasts examined. They are listed as clear dose slope, peak intensity, and standard contrast formats.

44

mJ/cm^2 peak dose. The latter occurs at 0.995 standard contrast, 65.9 (mJ/cm^2)/μm slope at

clear dose, and 13.2 mJ/cm^2 peak dose. These contrast values are the poorest studied (similar

to the poorest profile in figure 4-1) and nearly the best studied (similar to the second tallest

low background profile in figure 4-1), respectively. Looking back at figure 1-1, it contains

SEM images of these two experimental samples. For each aerial image profile, at least four

LER measurements are made; the top and bottom edges of two different locations in the

trench are measured.

Correlation between LER and the above-mentioned forms of measuring contrast-- stan-

dard contrast, image slope, and peak intensity-- are extracted. Correlation is measured by fit-

ting curves of LER data versus contrast and comparing the R^2 fitting parameter. More

rigorous statistics are used in chapter five; the reason for limiting statistical analysis here is

explained at the end of this later in this chapter. R^2 fitting is found as follows in equation 4-1.

Equation 4-1:

$$R^2 = \frac{\frac{1}{n}\sum_{i=1}^{n}(\hat{Y}_i - M)^2}{\frac{1}{n}\sum_{i=1}^{n}(Y_i - M)^2}$$

$\hat{Y}_i$ is the expected value of the fitting equation, Y_i is each specific data point, M is the

overall mean of the data, and n is the total number of data points. R^2 equal to unity is a perfect

correlation; zero is no correlation. The sets of data from each resist are treated separately,

since it is possible that different resists have different mechanisms that lead to LER. In all

three different resists, the strongest correlation between LER and image contrast is found

between LER and standard contrast, as defined in equation 2-2. Table 4-2 contains a summary

	R^2 value UVII-HS	R^2 value UV210	R^2 value SEPR-463
LER vs clear dose slope (linear relationship)	0.454	0.1564	0.3739
LER vs <best fit> clear dose slope (non-linear relation)	0.721	0.2335	0.6783
-- (best fit equation):	$y=94.052\ x^{-0.6574}$	$y = 0.0001x^2 - 0.0462x + 9.4753$	$y = 285.86\ x^{-0.7636}$
LER vs peak intensity (linear relationship)	0.2643	0.0763	0.2274
LER vs <best fit> peak intensity (non-linear relation)	0.3805	0.0939	0.3851
-- (best fit equation):	$y=107.4\ x^{-0.9894}$	$y = 0.0026x^2 - 0.2485x + 11.141$	$y = 438.4\ x^{-1.1539}$
LER vs (max-min)/(max+min) cont. (linear relationship)	0.6726	0.2781	0.553
LER vs <best fit> cont. (non-linear relation)	0.8344	0.3131	0.734
-- (best fit equation):	$y=4.6173\ x^{-0.7951}$	$y=8.56\ e^{-0.554x}$	$y = 3.6954\ x^{-0.8942}$

Table 4-2: Summary of correlations between line edge roughness and the different forms of measuring aerial image contrast. With all three resists, LER versus standard contrast gave the strongest correlation, although UV210 did not yield a strong correlation with any definition of contrast.

of all the correlations. Correlations between LER and any form of measuring of contrast are all weak for UV210. Figures 4-3 through 4-5 show typical correlation plots. Figure 4-3 is the correlation between LER and the standard definition of contrast--

$$(max - min) / (max + min) -$$

- for SEPR-463. Figure 4-4 plots LER versus standard contrast for UVII-HS. Figure 4-5 is the same for UV210. A linear correlation of LER versus the specific form of contrast is included in table 4-2, however a linear fit is never optimal.
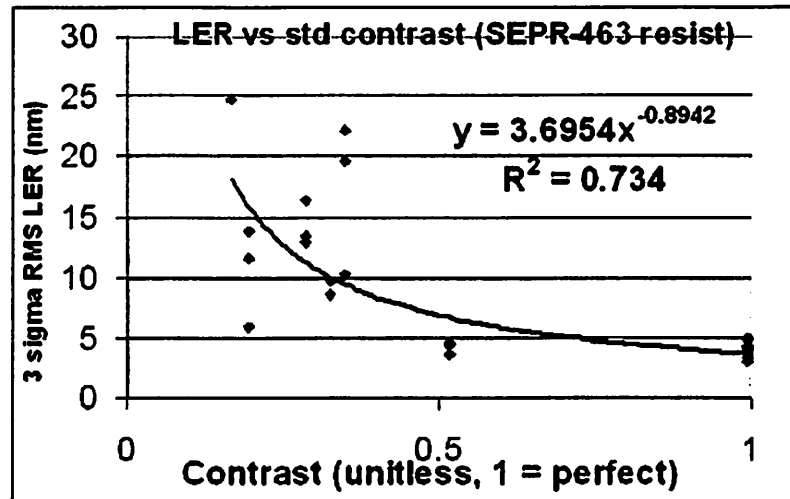
Figure 4-3: Plot of LER versus standard contrast for SEPR-463 resist. Correlation, R^2, is 0.734.

Some similarities between the strongest correlations are seen in this set of data. Logarithmic, exponential, power, inverse, binomial, and linear relationships between LER and the various forms of contrast are all examined. With all three resists, a power law relationship of

LER versus standard contrast emerges as the best fit. In the cases of UVII-HS and SEPR-463, the two resists which show a significant correlation between LER and standard contrast, the following approximate relationship is found:

Figure 4-4: Plot of LER vs standard contrast for UVII-HS resist. Correlation is 0.834.

Equation 4-2:

$$\text{Line Edge Roughness} \propto (\text{Standard Contrast})^{-0.85}$$

It is expected that as contrast degrades, LER increases. Therefore, the negative exponential relationship is anticipated. Furthermore, LER is not quite inversely related to aerial image contrast; the relationship is somewhat milder than that.

In the case of UV210, there are many data points where the LER is quite low even when contrast is poor. However, there are others where the LER is significant when contrast degraded.

Figure 4-5: Plot of LER vs standard contrast for UV210 resist. Correlation is very poor, at 0.313.

No strong correlation between LER and contrast can be found in the case of UV210.

A summary of all the raw data collected for this LER versus aerial image contrast experiment is presented in *Appendix A*.

4.3 LER versus Contrast (Multiple Resists) Conclusion and Discussion

The correlation between various aerial image contrast models and LER was studied with three different photoresists. LER versus the standard definition of contrast showed the strongest correlation. In the two resists that showed a significant correlation, UVII-HS and SEPR-463, there was a nearly inverse relationship: LER α (contrast)$^{-0.85}$, approximately.

The standard contrast definition depends upon the maximum dose and the minimum, or background, dose. The best fit between standard contrast and LER was an unexpected result. The experiment was designed either with the expectation that the slope contrast would be most significant, suggesting that LER was a localized phenomenon; or with the expectation that peak intensity contrast would be most significant, suggesting that long range effects such as acids diffusing over relatively large lengths dominate LER. In fact, the author was expecting the former to be true for photoresists with shorter diffusion lengths, such as UV210; and the latter to be true for older resists with larger diffusion lengths. Standard contrast was included in the study to provide a control, a method of measuring contrast that has been used for several years.

One explanation for why standard contrast fits best with LER is that the background dose had a significant effect upon LER. Standard contrast is the only form examined which defines itself, at least partially, using the minimum (or background) dose. This possibility seems counter-intuitive. One explanation how the presence of photo-generated acids in areas that received little exposure, a level of exposure below that needed to create enough acids to

deprotect the photoresist enough to be developed, could affect LER is given: The presence of photo-generated acids in the sub-threshold exposure area, or the background, change the over-all diffusion of acids into that area. For instance, a higher level of acids in the "unexposed" region (but still not enough acids to allow it to become exposed) prevents many acids in the exposed region to diffuse into that unexposed region. In the case of Fickean diffusion, whether the concentration in the background exposure area changes or whether the concentration in the peak intensity changes, the only key parameter is the concentration gradient. In this case, peak intensity contrast or a newly defined minimum intensity contrast should both fit LER data equally poorly or well. Standard contrast would maybe fit better than either of these contrasts, and the slope contrast should fit best of all. The results presented above, with peak intensity contrast not fitting well in any case and standard contrast fitting well in all cases, do not fit with any of the above expectations. As mentioned in chapter two, however, photogenerated acids do not move via simple Fickean diffusion.

A new experiment was designed in order to better understand this unexpected result. This experiment is detailed in chapter five.

5 Line Edge Roughness versus Aerial Image Contrast: In-Depth Study

This chapter describes a more in-depth study using only UVII-HS resist in order to highlight and carefully quantify the image metrics that correlate best with LER, particularly flare. The background dose seems to be contributing to LER in some form, otherwise the peak intensity contrast should have correlated at least as well as the standard contrast in chapter four. Work up to this point presented no form of contrast that measured solely background dose, so it is difficult to make any conclusive statements. In this chapter, a new experiment is run and a new form of contrast, the **background** contrast, is created.

Equation 5-1:

$$\text{Background Contrast} = \frac{(\text{clear dose} - \text{background dose})}{\text{clear dose}}$$

where "clear dose" is the minimum exposure dose needed to clear for that photoresist and "background dose" is the background (or flare) exposure dose level that the photoresist experienced. The clear dose is a fixed value for each photoresist and lithography system. It will depend only upon exposure wavelength, post-exposure bake, and other processing conditions. Similar to the standard contrast, a background contrast of 1 is perfect-- no flare experienced-- and with a contrast of 0 there is no patterning-- the entire area has been fully exposed, including sections meant to be unexposed. Of course, since only information regarding the background dose is given, it is entirely possible for an image with the background contrast equal to

1 to have poor image quality. For example, a photoresist coated wafer that receives no lithographic exposure whatsoever experiences background contrast equal to 1. Such a wafer will, of course, have poor imaging. Therefore, the background contrast cannot be the sole form of measurement used to define quality imaging. Instead, it is used as a tool to determine if indeed the background exposure level is as important as it seemed in chapter four.

5.1 Designing the In-Depth LER versus Contrast Study

In chapter four, several different aerial image profiles were examined. Throughout the entire experiment, though, only 3 different background exposure doses were studied. Examining three different levels of the background contrast would not be sufficient to order to find a statistically significant correlation. Therefore, a new experiment is designed. Eight different background contrast levels are studied, and 30 different LER measurements are taken at each contrast level, 15 measurements per trench side. With a pool of data this large, the number of photoresists studied has to be reduced. The author examines UVII-HS because it had the strongest correlation with standard contrast. As demonstrated in chapter four and by Sanchez et al., the LER of each photoresist has a different response as aerial image profile is altered. This study is not attempting to conclude that all photoresists will act similar to UVII-HS. However, UVII-HS is a model photoresist that still shares many attributes with the most modern 248 nm and 193 nm photoresists. The polymer side chains end in t-butyl esters which, after deprotection, form carboxylic acids [18, 19, 54]. This is still a dominant deprotection reaction in modern photoresist chemistry. Moreover, this work is attempting to uncover a previously unstudied phenomenon, a relation between flare and LER. Discovering

51

this in numerous photoresists would be ideal, but discovering it in one model photoresist is still quite significant.

Roughly 200 data points are collected, since some of the measurements had to be discarded. It is necessary to collect so many repeated measurements in order to have statistical confidence in the results. LER measurements are quite noisy, as already mentioned. In this experiment, as seen in appendix A2, the range between LER measurements at the worst contrast level is larger than the difference between the average of the LER measurements at that worst contrast level and the best contrast level. The range between individual 1 σ LER measurements at the worst level is 6.2 nm, yet the difference between the average of the worst level and the best level is only 5.3 nm. In order to make confident conclusions and find strong correlations, a large sample size and rigorous statistical analysis are needed. By taking 15 measurements per trench side per contrast level, the noise in the data is reduced by $\sqrt{15}$, or nearly 4.

Three different forms of measuring aerial image contrast are compared, once again. This time they are slope contrast, standard contrast, and the newly defined background contrast. The peak intensity contrast was discarded because chapter four demonstrated that this form of measuring contrast never correlated well with line edge roughness.

In order to extract the most useful data possible, the different forms of aerial image contrast must be made as orthogonal to each other as possible. It is impossible to make them fully orthogonal since they are all measuring portions of the same thing, aerial image profile. The standard contrast could be relatively high, for instance, while the background contrast is low by overexposing the photoresist in the exposed regions and allowing a fairly significant level of exposure in the background. Looking back at figure 1-2, it is a SPLAT simulation of the

aerial image profiles of the eight different contrast levels chosen in this experiment with maximized orthogonality. The ideal orthogonality is found by plotting the three different forms of contrast simultaneously. Figure 5-1 is a plot of all three forms of contrast, with the background contrast plotted nearly linearly. A high level of orthogonality
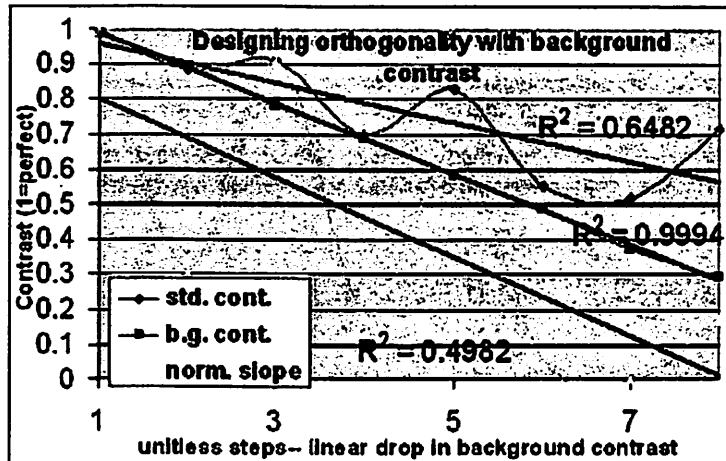
Figure 5-1: Plotting contrast levels from figure 1-4 in terms of the three different forms of measuring contrast. In this figure, it was plotted to make the background contrast fit as linear as possible. The other fits were relatively poor, providing a relatively high degree of orthogonality.

occurs when the R^2 fitting parameter-- fit to a linear equation-- for the linearly plotted contrast is close to 1 and the others are as close to 0 as possible. Figure 5-2 is the same plot except with the normalized slope plotted linearly. Aerial image profiles similar to those in figure 1-2 are adjusted until all three forms of measuring contrast were as orthogonal as possible. Orthogonality is determined via plots similar to figures 5-1 and 5-2. In figure 1-2, the aerial image profiles for the 0% back-
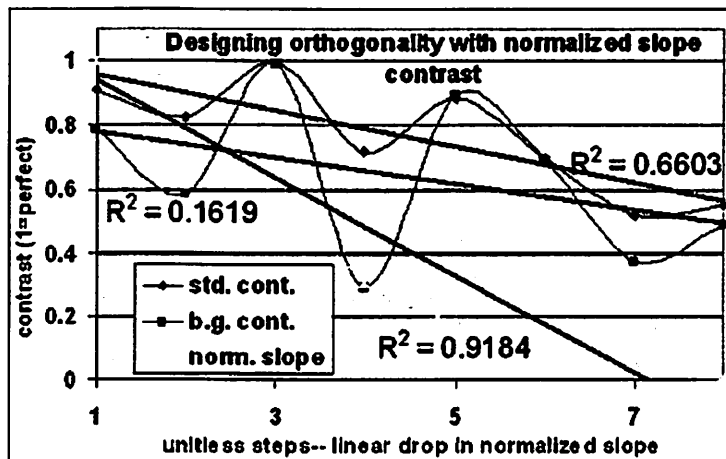
Figure 5-2: Plotting contrast levels from figure 1-4 in terms of the three different forms of measuring contrast. In this figure, it was plotted to make the normalized slope contrast fit as linear as possible. The other fits were relatively poor, providing a relatively high degree of orthogonality.

53

ground (0.99 background contrast) and the 62% background (0.38 background contrast) were fixed. These are two different background levels used in chapter four, therefore their aerial image profiles were already decided. In this way, the same wafers exposed in that chapter could be reused for this one. This saves time, but more importantly keeps some level of

	standard contrast:	background contrast:	norm. slope contrast:	slope contrast:	b.g. exposure level (% of clear)
1	0.99	0.99	1.00	14.30	0%
2	0.88	0.89	0.31	8.12	10%
3	0.91	0.79	0.91	13.50	20%
4	0.70	0.69	0.19	7.03	30%
5	0.83	0.58	0.64	11.09	40%
6	0.55	0.49	0.05	5.79	50%
7	0.52	0.38	0.00	5.38	62%
8	0.72	0.29	0.17	6.93	70%

Table 5-1: Summary of the contrast levels that yielded maximum possible orthogonality between the different forms of measuring contrast. The final column, background exposure level, is simply the percent of clear dose targeted for the background exposure, the second exposure. Slope contrast units are $(mJ/cm^2 \cdot \mu m)$, all other measures of contrast are unitless.

continuity between the two studies. Figure 1-2, figure 5-1, and figure 5-2 represent the best level of orthogonality between contrast forms. Table 5-1 is a summary of the different contrast levels in this most orthogonal case. For example, the samples measured from the 7th level of the experiment have standard contrast equal to 0.52, background contrast equal to 0.38, slope contrast equal to 5.38 $mJ/cm^2 \cdot \mu m$, normalized slope contrast equal to 0.

The normalized slope contrast is a simply a normalized version of the slope contrast used earlier. It is created to make it easier to examine the orthogonality of the contrasts compared to each other. The normalized slope contrast is defined as

Equation 5-1:

$$\text{Normalized Slope Contrast} = \frac{\text{slope contrast} - \text{minimum slope contrast}}{\text{max slope contrast} - \text{min slope contrast}}$$

where the slope contrast is the same as the slope contrast value used until this point, the minimum slope contrast is the smallest slope contrast examined in this study, and the maximum slope contrast is the largest slope contrast examined here. This is done to allow the slope contrast to be normalized between 0 and 1.

5.2 SEM Measurement Issues

The SEM tool used to collect data throughout this dissertation, the Leo 1550, has an asymmetric secondary electron detector, or secondary electron collector. The asymmetry in the detector leads to LER measurements that are dependent upon the orientation of the sample with respect to the detector. Figure 5-3 is an image of the inside of the Leo 1550 chamber. The secondary electron detector lies on one side of the electron gun and sample. For proper symmetry, the detector should be a ring circumscribing the gun so that the gun is at the center of the ring, instead of the detector lying just to one side of the gun as shown in the figure. In order to minimize any error due to the detector asymmetry, for all the measurements done in this section the trench is oriented so that the top and bottom edges of the trench always receive exactly the same orientation with respect to the secondary electron collector. The lines are oriented perpendicular to the direction of the secondary electron collector to create the largest difference between detection environments that the top and bottom edge experience. In this manner, any differences due to SEM measurement inaccuracies should become apparent during data analysis. Moreover, this will allow one of the two LER measurements to have the strongest secondary electron signal possible.

The significance of this effect was not considered in chapter four, so the samples were not necessarily placed in exactly the same position with respect to the secondary electron detector.

However, the effect is evident once it is noticed. In figure 1-1 (a), the top edge seems to have slightly better resolution and some level of information
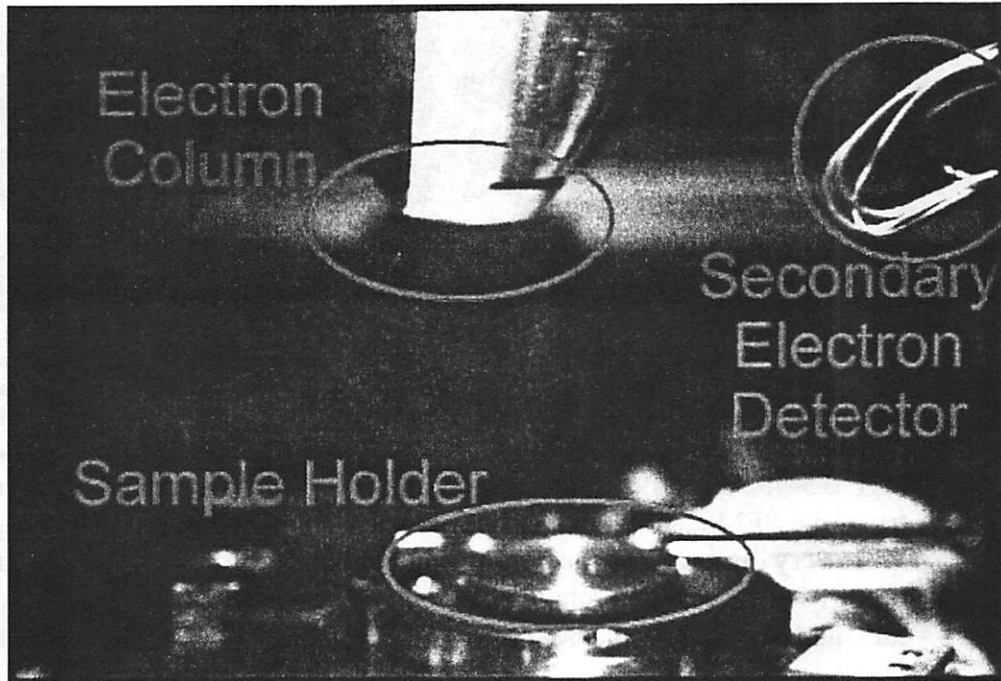
Figure 5-3: Image of the Leo 1550 sample chamber used. The secondary electron detector is not placed symmetrically around the electron column. Therefore signal detection is dependent upon the position of the sample relative to the detector.

in the z-direction can be seen on the top edge. None of this can be seen with the bottom edge. Yet, the sample was not tilted nor was the profile asymmetric. All samples in this section are measured with the orientation similar to figure 1-1 (a). Figure 5-4 is a sample image taken in this new run with this orientation. The SEM images are then artificially rotated counterclockwise after capturing the image so that the top side is the left side. This is done to make the image easier to read with the MatLab code created. The reader should keep in mind that "top" or "left" side LER values are those captured along the trench edge that looks similar to the top of figures 1-1 (a) and 5-4, and that the "bottom" or "right" side LER values are those captured along the trench edge similar to the bottom of figures 1-1 (a) and 5-4.

Mag = 49.96 K X	100nm ⊢──┤	EHT = 2.00 kV WD = 6 mm	Signal A = SE2 Photo No. = 490	Date :7 May 2003 Time :20:22
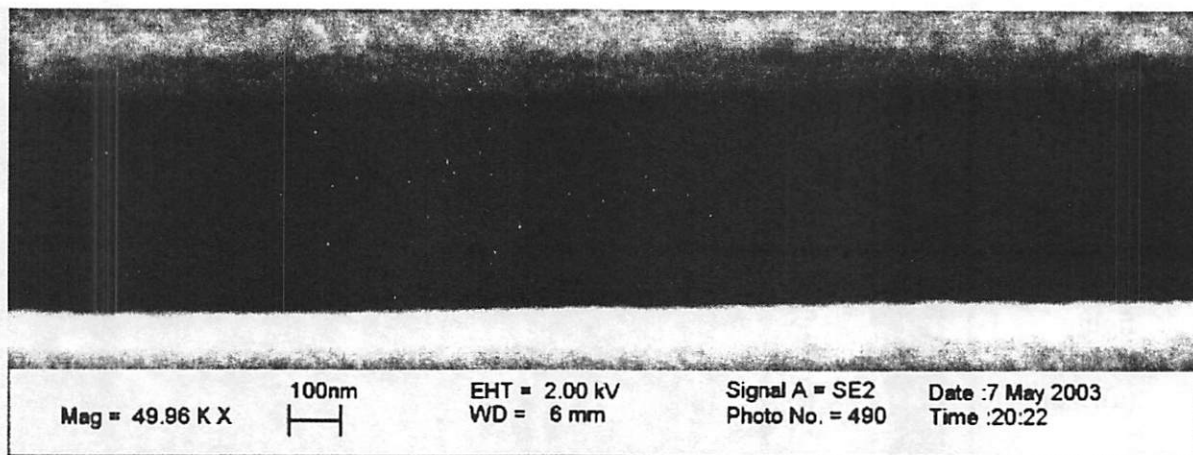
Figure 5-4: A sample SEM image from the 0.89 background contrast level. The asymmetry of the SEM electron detector is evident; the image along the top edge of the trench looks very dissimilar to the bottom edge of the trench. For this study, all SEM images are taken with the same orientation of the trench with respect to the secondary electron detector, and the top edge and bottom edge LER values are examined separately.

5.3 In-Depth LER versus Contrast Results using Treatment Averages

Data is first analyzed using R^2 fitting of the treatment averages similar to chapter four. Figures 5-5 through 5-7 are correlations between LER and the three different forms of contrast studied. The best fit is with the newly defined background contrast. R^2 is 0.76 fitting a power relation of background contrast versus LER on the left, or top, side of the trench. R^2 fitting, again for the power relation, is 0.79 for the background contrast versus LER on the
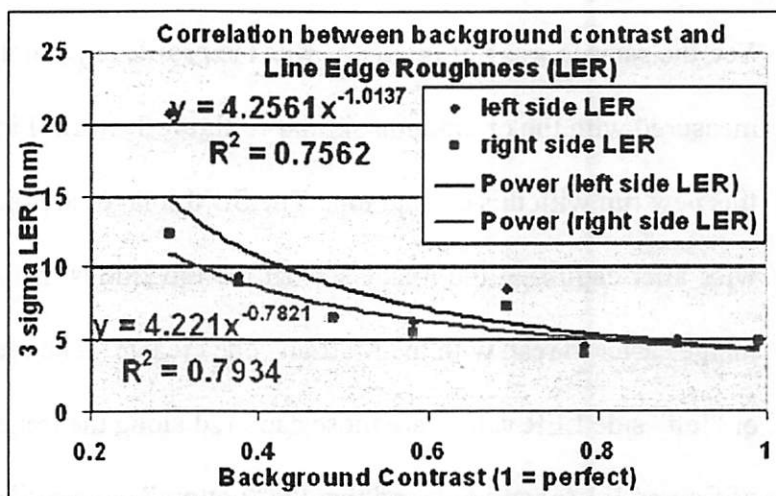
Figure 5-5: Correlation between background contrast and LER. Each data point represents 15 measurements, except in the case of 0.38 and 0.99 background contrast, which are 4 and 9, respectively.

right, or bottom, side of the trench. For both the left side and right side of the trench, a power relationship fit best, with a nearly inverse relationship.

Two data points in figure 5-5 significantly deviate from the fit, while all others fit quite well. LER values

Figure 5-6: Correlation between standard contrast and LER. Each data point represents 15 measurements, except in the case of 0.52 and 0.99 background contrast, which are 4 and 9, respectively.

for the 0.69 background contrast are considerably higher than the fit predicts. The following paragraph proposes a reason why this may have occurred. Also, the LER values for the poorest background contrast are very high. This may happen because, as the aerial image profile degrades, there is a point where lithographic imaging is no longer possible. The resist profile becomes worse than simply poor, but unrecognizable. Measuring LER is no longer possible Figure 5-8 is an example of a wafer exposed with such a poor aerial image that the trench

Figure 5-7: Correlation between clear dose slope contrast and LER. Each data point represents 15 measurements, except in the case of 5.38 and 14.3 mJ/cm^2·μm background contrast, which are 4 and 9, respectively.

58

did not fully develop. This aerial image profile was initially going to be another level used to evaluate all the photoresists in chapter four, however measuring LER was impossible and non-sensical. There is clearly some point at which the aerial image profile is so degraded that the LER increases very quickly and it deviates from any prior, comparatively mild, relationship between LER and contrast. Sanchez et al. also observed this region with several of the photo-resists they studied [15]. With several of the photoresists they studied, there was nearly a step function where the LER quickly degraded and it was soon impossible to measure sensical LER data.

As mentioned in the preceding paragraph, the 0.69 background contrast level LER data also did not fit the trend line as well as the other levels. The slope contrast at this level is exceedingly low, in order to better orthogonalize the experiment as mentioned earlier. The
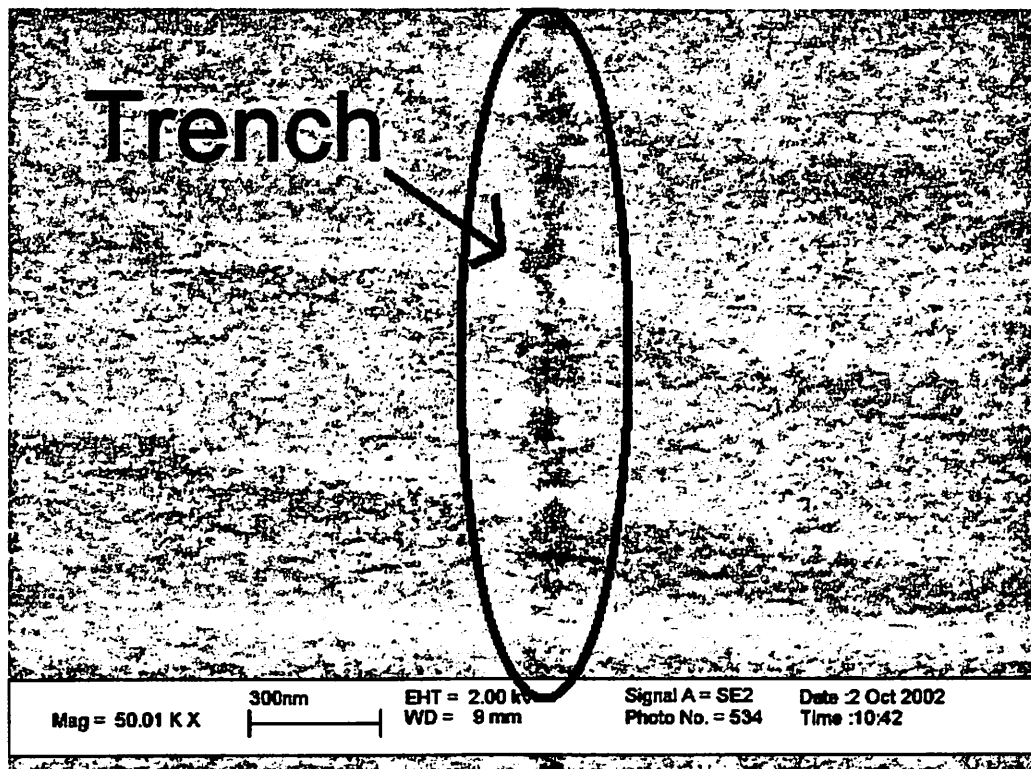
Figure 5-8: SEM image of trench receiving very poor exposure aerial image profile. It is impossible to measure LER in this case because the trench is not fully cleared; a large portion of the photoresist still remains in the trench area.

slope contrast at this level is 7.03 mJ/cm^2·μm. The neighboring background contrast values, 0.58 and 0.79, had slope contrasts equal to 11.09 mJ/cm^2·μm and 13.50 mJ/cm^2·μm, respectively. If the slope contrast is indeed also affecting LER, this would explain the relatively high LER at this point. However, the 0.38 background contrast level has even poorer slope contrast, 5.38 mJ/cm^2·μm, with only slightly worse LER data. It is therefore possible that the discrepancies in LER values are simply due to statistical error or some other phenomenon.

5.4 LER versus Contrast Study-- Rigorous Statistical Analysis

In order to state with confidence that the background contrast fits the experimental data best, analysis of variance (ANOVA) is needed. ANOVA is a statistically rigorous method for assessing whether data is fitting a model or following random noise. Further statistical analysis is performed using JMP software version 3.2 (SAS Institute, Inc.) [55-56]. First the best fitting equations are found. The fits are limited to a linear fit, logarithmic fit, exponential fit, and power fit. Polynomial fits are discounted because they are physically non-sensical. There is no physical reason why there should be either a maximum LER or minimum LER that does not occur at the maximum or minimum contrast explored. No local maximum or minimum should exist. This would mean the LER gets worse as contrast improves or gets better as contrast degrades. Both of these conclusions are physically illogical. Fits that involved a combination of certain equations, for instance the sum of a linear equation and a power equation, were also ignored. While these fits do make physical sense, especially when considering the argument made two paragraphs earlier that there might be a region of rapid LER increase, they are more difficult to model. The models become so complex that it is more difficult for the equations to converge upon an equation and even more difficult to find reasonable control

limits. The most important criteria throughout this statistical exercise are to find the form of contrast that best correlates with LER and to find which forms of contrast correlate enough to conclude that they contribute significantly to LER.

Figure 5-9 is the output to an exponential fitting model. The general form of the fitting equation must be defined by the user. In this case, the author chose the following general power law equation:

Equation 5-2:

$$\text{Left Side of Trench LER} = c_1 e^{(c_2 \cdot \text{Background Contrast})}$$

c_1 and c_2 are constants that are fit using the JMP statistical software, LER and background contrast are as defined early. LER values used throughout this section are 1 σ values instead of 3 σ values. This means that equation 1-2 is used as is, instead of multiplying it by 3. At some points throughout the analysis, it is necessary to find the variance within the set of LER data. For these calculations, the 1 σ RMS roughness values are needed.

Non-linear model fitting, such as that performed in figure 5-9 and equation 5-2, does not allow ANOVA. However, it needs to be performed upon all of the fitting equations in order to determine which best correlates with LER. Detailed descriptions of ANOVA can be found in several statistics textbooks that focus upon manufacturing or laboratory environments [57]. In order to use ANOVA, the relationship between the observation, in this case LER, and the effect, contrast, must be linear. The fitting equations mentioned above must therefore all be linearized so that analysis of variance can be applied. The following is an example to linearize a general power model fitting the left side of the trench LER to background contrast:
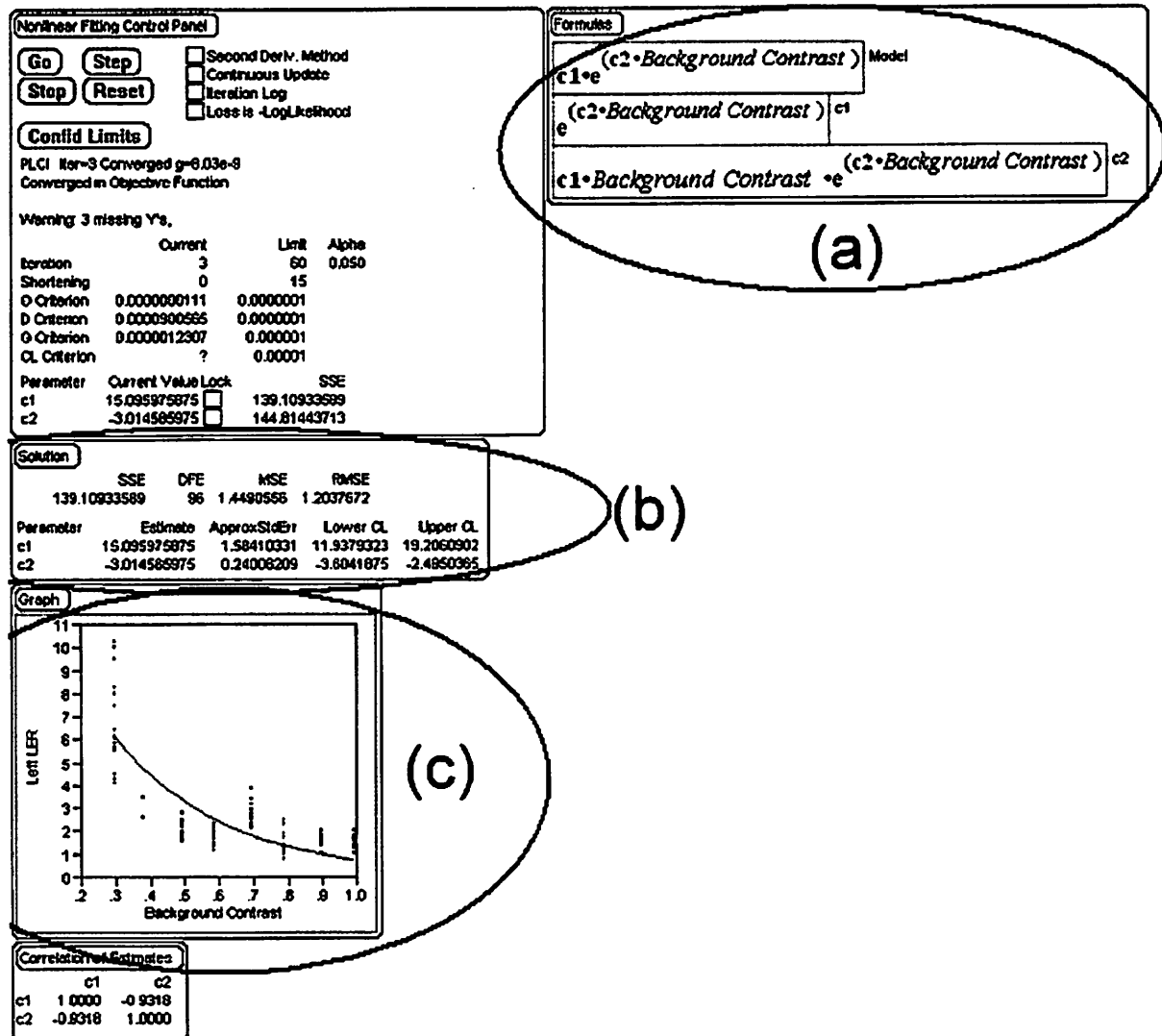
Figure 5-9: A typical output of statistical data from JMP software. The first line of portion (a) is the model formula for fitting the data, which is set by the user. In this case, it is an exponential form with two constants, c1 and c2, that are solved. The subsequent lines are the derivatives with respect to c1 and c2. Highlighted area (b) contains the solution for c1 and c2 parameters, as well as the upper and lower confidence limits, the sum of square errors (SSE), and other statistical information. Area (c) is a plot of the data and the model fit to the data.

Equation 5-3:

$$\text{Left Side of Trench LER} \ = \ c_1 (\text{Background Contrast})^{c_2}$$

Simply taking the logarithm of both sides yields:

Equation 5-4:

$$\ln(\text{LER}) = \ln(c_1) + c_2 \cdot \ln(\text{Background Contrast})$$

To properly apply ANOVA, the LER data all need to be converted to ln(LER) and the background contrast values need to be converted to ln(bg cont). Then the data is fit using the standard least squares fit like that described in equation 4-1. Table 5-2 shows fits of the parameters $\ln(c_1)$ and c_2 for the model in equation 5-4 using left side of trench LER versus background contrast. The intercept of the fitting equation is $\ln(c_1)$ and the slope is c_2. There-

Term	Estimate	Std Error	Lower 95%	Upper 95%
Intercept	0.2865402	0.054987	0.1773912	0.3956892
ln(BG Contrast)	-1.113174	0.087112	-1.28609	-0.940257

Table 5-2: Least squares fit for the parameters of the linearized equation 5-4, the dependence of left LER upon background contrast. "Intercept" is the term $\ln(c_1)$ and ln(BG Contrast) is the term c_2 as specified in equation 5-4.

fore, in the case of left LER versus background contrast, the optimal fitting equation is:

Equation 5-5:

$$\ln(\text{LER}) = (0.2865 \pm 0.0550) - (1.1132 \pm 0.0871) \cdot \ln(\text{Background Cont.})$$

Or, converting it back to the original power equation form:

Equation 5-6:

$$\text{LER} = (1.3318 \pm \sim 0.0732) \cdot (\text{Background Contrast})^{-(1.1132 \pm 0.0871)}$$

Studying the confidence limits provided in table 5-2 can further explain the model fit. Table 5-2 contains the 1 σ standard error values and the upper and lower 95% confidence intervals. Since the 95% confidence intervals for the exponential term do not include zero,

there is already at least 95% confidence that background contrast does influence LER. The errors given in equations 5-5 and 5-6 are 1 σ values. The errors on c_1 in equation 5-6 are not quite symmetric, however, since the symmetric errors were determined for $\ln(c_1)$. The errors written in equation 5-6 are approximate. The upper bound of c_1 is 1.4071 and the lower bound is 1.2605.

Further analysis of linearized equation 5-5 using ANOVA is shown in table 5-3. There were 97 data points, yielding 97 total degrees of freedom. There is only 1 level, or degree of freedom of the model, which is the background contrast. In the model, no other variable accounts for LER varia-
tions. The DF for error are found by subtracting the DF of the model from the total DF, yield-ing 96, since all other

Analysis of Variance				
Source	DF	Sum of Squares	Mean Square	F Ratio
Model	1	17.954209	17.9542	163.2946
Error	96	10.555183	0.1099	Prob>F
C Total	97	28.509392		<.0001

Table 5-3: ANOVA for the linearized power relationship model fitting background contrast to left side LER, equation 5-5. DF are the degrees of freedom.

variations in LER must be accounted for by random error in this case. The sum of squares (SS) for the model are found as follows:

Equation 5-7:

$$SS = \sum_{i=1}^{n} (x_i - \bar{x})$$

where x_i is the model predicted value at that data point. For instance, if the background contrast were 0.99 (the highest b.g. contrast tested), then the model predicted value for $\ln(\text{Left LER})$, x_i, would be 0.298 since $0.2865 - 1.1132 \cdot \ln(0.99) = 0.30$. $\bar{x}$ is the mean

of all the model predicted values. Next the mean square, or mean sum of squares (MSS), is simply the SS divided by the degrees of freedom:

Equation 5-8:

$$MSS = \frac{SS}{\text{degrees of freedom}}$$

The terms for the error are calculated almost identically, except the sum of squares due to error (SSE) is slightly different.

Equation 5-9:

$$SSE = \sum_{i=1}^{n} (y_i - x_i)$$

where x_i is the same as described for equation 5-7 and y_i is the individual data point. The mean sum of square errors (MSSE) is calculated by dividing the SSE by the DF of the error, analogous to equation 5-8. Finally, the ratio of the MSS over MSSE follows an F-statistical distribution.

Equation 5-10:

$$\frac{MSS}{MSSE} \sim F_{DF_{model}, DF_{error}}$$

The ratio of MSS over MSSE is then compared to the applicable F-distribution and the probability that the model is not actually fitting the data and that the data are simply random noise, at least with respect to the model, is found. The larger the F ratio, in this case it is 163.3, the lower the probability that the model is not fitting the data. The probability and the F ratio are shown in table 5-3 above. In this case, the probability that the model is **not** fitting the data is

less than 0.0001. Or, there is a greater than 99.99% probability that the model is fitting the data.

Similar calculations are made with the other forms of contrast, standard contrast and slope contrast, versus both left LER and right LER. These calculations are done with power, exponential, and logarithmic fitting functions. In all cases, the background contrast has the strongest F Ratio. Tables 5-4 and 5-5 are a summary of the left and right LER power function

	Left side of trench LER: Power function fitting:							
	C1, low	C1, estim.	C1, high	C2, low	C2, estim.	C2, high	Equation	F Ratio
B.G. Cont.	0.1774	0.2865	0.3957	-1.2861	-1.1132	-0.9403	LER=1.332 (b.g. cont)^1.1132	163.3
Std. Cont.	0.3621	0.536	0.7099	-1.6727	-1.1441	-0.6155	LER=1.709 (std. cont)^1.1441	18.46
Slope Cont	2.0876	2.718	3.3484	-1.1579	-0.8688	-0.5798	LER=15.15 (slope cont)^0.8688	35.61

Table 5-4: Summary of the equations best fitting LER on left side of trench to contrast. Background contrast has the highest F Ratio, next slope contrast, and finally standard contrast. All F Ratios are high enough that all forms of contrast show a significant effect upon LER. The C_1 estimates are those of $\ln(C_1)$. For example, C_1 for b.g. contrast is actually $e^{0.2865}$, or 1.3318.

	Right side of trench LER: Power function fitting:							
	C1, low	C1, estim.	C1, high	C2, low	C2, estim.	C2, high	Equation	F Ratio
B.G. Cont.	0.2244	0.3052	0.3861	-0.9423	-0.8045	-0.6667	LER = 1.357 (b.g. cont)^0.8045	134.3
Std. Cont.	0.2883	0.3977	0.5071	-1.409	-1.0719	-0.7349	LER = 1.4884 (std. cont)^1.0719	39.86
Slope Cont	1.8024	2.2139	2.6254	-0.8938	-0.7065	-0.5192	LER = 9.151 (slope cont)^0.7065	56.07

Table 5-5: Summary of the equations best fitting LER on right side of trench to contrast. Again, background contrast has the highest F Ratio, next slope contrast, and finally standard contrast. All F Ratios are high enough that all forms of contrast show a significant effect upon LER. The C_1 estimates are those of $\ln(C_1)$.

model fits, respectively. The exponential and logarithmic fits are left out of the tables because their fits were never as strong.

Tables 5-6 and 5-7 contain an example fit and ANOVA for an exponential function, which was typically the second best fit. These two tables are analogous to tables 5-2 and 5-3 above.

Term	Estimate	Std Error	Lower 95%	Upper 95%
Intercept	1.2039339	0.132355	0.9405884	1.4672794
Background Cont	-0.771421	0.180685	-1.130927	-0.411914

Table 5-6: These are the parameter estimates for the linearized exponential function described in equation 5-12. "Intercept" is the estimate for $\ln(c_1)$ and "Background Cont" is the estimate for c_1. This table is analogous to table 5-2, but with an exponential fit.

In this case, they present the data to fit the left side LER to background contrast. The exponential function first has to be linearized, so that the general exponential function

Equation 5-11:

$$\text{Left LER} = c_1 e^{c_2(\text{Background Contrast})}$$

becomes

Equation 5-12:

$$\ln(\text{Left LER}) = \ln(c_1) + c_2 \cdot \text{Background Contrast}$$

Parameters are found for the general exponential function 5-12, and it is then fit and analyzed

using ANOVA exactly as described above for the power function fit. As is evident from table 5-7, the F Ratio is nearly one order of

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	1	1.4503777	1.45038	18.2281
Error	81	6.4450421	0.07957	Prob>F
C Total	82	7.8954197		<.0001

Table 5-7: anova exponential left vs bg ANOVA for the linearized exponential model fitting background contrast to left side LER, equation 5-12. This is analogous to table 5-3.

magnitude lower than it was

for the power model fit. There is still a strong fit to the model; the model fits the data well

enough that the variations in LER cannot be considered mere random noise. However, the

power model gives a significantly stronger fit in this case, and all that were examined.

The statistical F-test is the most statistically rigorous method to prove whether a hypothesis is true or false, however it is difficult to intuitively understand. The next few paragraphs provide alternate ways to view how much better the background contrast fits compared to the other two forms of contrast. Only background contrast and standard contrast are explicitly compared here, although similar conclusions can be seen comparing background contrast to the slope contrast. First, the confidence ranges of the fitting equations are examined. This example will compare the best fit equations of standard contrast versus LER and background contrast versus LER, looking at only the LER of the left side of the trench. The linearized fitting equation for the background contrast, with the 95% confidence limits, is given in equation 5-5. Equation 5-13 is the linearized fitting equation for the standard contrast, with 95% confidence limits.

Equation 5-13:

$$\ln(\text{LER}) = (0.536 \pm 0.1739) - (1.1441 \pm 0.5286) \cdot \ln(\text{Standard Contrast})$$

Calculating LER values predicted from the standard contrast model, equation 5-13, yields 95% confidence interval LER values ranging from 1.45 nm to 2.07 nm for the best standard contrast level and LER values ranging from 2.15 nm to 6.07 nm for the poorest standard contrast level. The background contrast model, equation 5-5, yields a 95% confidence interval ranging from 1.21 nm to 1.50 nm at the best background contrast level and 3.82 nm to 7.3 nm at the worst contrast level. The confidence interval of the background contrast is 48% of the interval of the standard contrast at the best contrast level, and the interval of the background contrast is 89% of the standard contrast at the worst contrast level. And at every point in between, the confidence interval of the background contrast model is tighter than that of the standard contrast model.

Figures 5-10 and 5-11 also show that the confidence interval for the background contrast fit is much tighter than that for the standard contrast fit. Figure 5-10 is a plot of the ln(LER) predicted by the standard contrast model versus the ln(LER) of the actual data. The solid red line is the fitting equation, the dashed red lines are the 95% confidence limits, and the black dots are the individual 1 σ RMS LER measurements. The ln(LER) is plotted instead of LER so that the linearized fitting equation can be shown. Figure 5-11 is a similar plot, but using the background contrast model instead of the standard contrast model. It is clear to see that the confidence limits for the background contrast are much tighter than those for the standard contrast, and that the data fit the model better. As a side point, these figures also help to see how noisy LER measurements are and why so many measurements are needed in order to draw strong conclusions.

Differences between actual measurements and values predicted by a model are entirely due to random noise when the model is perfect. This noise may be due to measurement error or
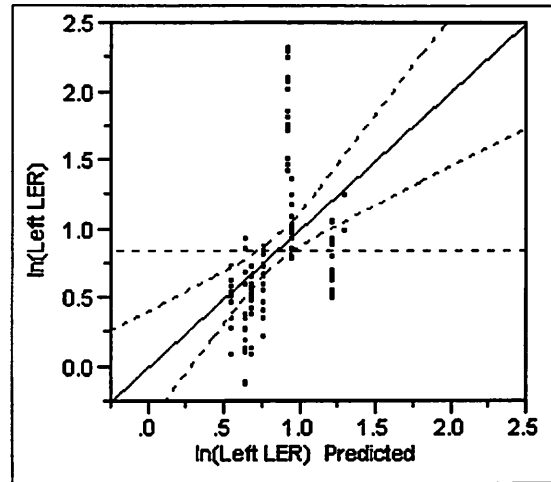
Figure 5-10: Plot of ln(LER) predicted by the standard contrast fitting model versus actual ln(LER) data. The solid red line is the fitting equation, the red dashed lines are the confidence limits, and the black dots are actual data.
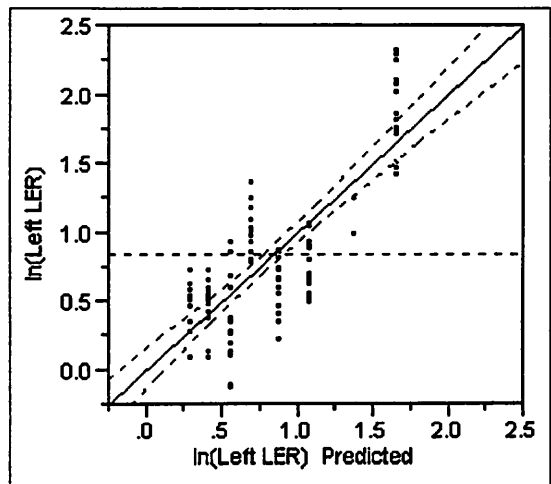
Figure 5-11: Plot similar to figure 5-10, but x-axis is ln(LER) predicted by the background contrast model instead of the standard contrast model. Again, the solid red line is the fitting equation, the dashed lines are the 95% confidence limits, and the dots are the actual data.

real variation in the data, but it will be independently and normally distributed, or in other words fully random. As the model fit degrades, the differences between actual measurements and predicted values become less random. Looking at residuals is a common practice to detect this lack of randomness. A residual is the difference between the actual measurement and the value predicted by the model. For instance, at 0.49 background contrast one of the 1

σ LER measurements is 2.27 nm and the model predicts 1 σ LER equal to 2.95 nm. The residual is (2.27 nm − 2.95 nm), or -0.68 nm. Figure 5-12 is a plot of all the residuals between the actual data-- LER on the left side of the trench-- and the background contrast model. In this plot the residuals are actually ln(LER) measured minus ln(LER) predicted; ln(LER) must be used since the linearized equation is examined. A perfect model would have the data normally and independently distributed around 0 at each predicted ln(LER) level. Studying figure 5-12, at the predicted ln(LER) equal to 0.70, which is 2.01 nm predicted LER and occurs at 0.69 background contrast, all of the residuals are above 0. The model does not fit the data perfectly. Comparing figure 5-12 to 5-13, however, it is immediately obvious that
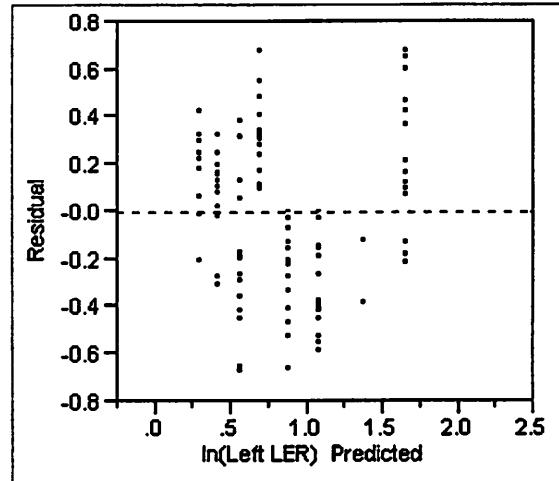
Figure 5-12: Plot of ln(LER) predicted by the background contrast fitting model versus residuals using that model. A perfect fit would have residuals that are independently and normally distributed around 0.
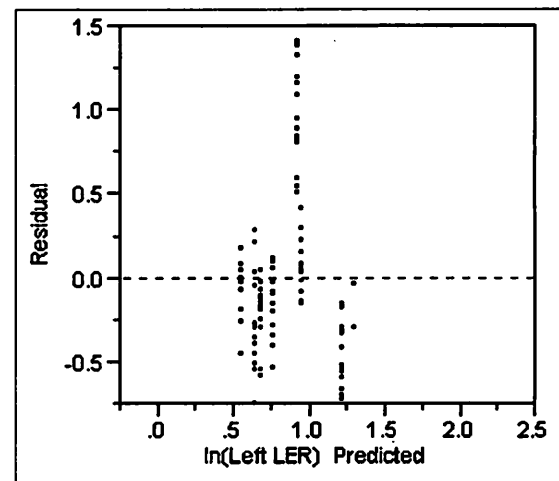
Figure 5-13: Plot of ln(LER) predicted by the standard contrast fitting model versus residuals using that model.

70

the background contrast model fits the data better than the standard contrast model. The largest residual for the standard contrast model is nearly 1.5, over double that of the background contrast model. Also, there are two predicted ln(LER) levels where all of the residuals are either above 0 or below 0. The level with residuals farthest off center, at 0.91 predicted ln(LER) and 0.72 standard contrast, is the same level with the poorest background contrast. This is further proof that the background contrast model fits LER better than the standard contrast model.

5.5 LER versus Contrast Study-- Models Based upon Multiple Forms of Contrast

With both the left side LER and the right side LER, the background contrast model provides the best fit, then the slope contrast model, then finally the standard contrast model. However all three contrasts showed a strong fit. This is due to two reasons mentioned earlier. First, these three definitions of contrast are not orthogonal to each other. As background contrast degrades, for instance, it is difficult if not impossible to keep the standard contrast from degrading as well. This can be prevented to some degree by increasing the maximum exposure dose, but the standard contrast cannot be kept entirely constant as background contrast changes. The second explanation is that several of the components of the contrast are important. Stepping back a little and thinking of the overall picture, it is obvious that background contrast alone cannot explain the correlation between line edge roughness and aerial image contrast. The background contrast does not consider the actual exposure dose, only the unwanted background dose, or flare. Therefore, it would be desirable to find an equation with the combination of all or two of the different forms of contrast that best describes the relationship between the aerial image profile and LER.

The simplest combination of forms of contrast would be a linear combination of two forms, say background contrast and slope contrast. It is clear from the data and the analysis up to this point that the dependence of LER upon contrast will not be linear. Assuming the trends presented above hold true for linear combinations-- and the power model still fits best for each individual contrast-- the general fitting equation would look like:

Equation 5-13:

$$\text{LER} = c_1(\text{b.g. cont.})^{c_2} + c_3(\text{slope cont.})^{c_4}$$

The fitting parameters, c_1 - c_4, are not the same fitting parameters as those found in tables 5-4 and 5-5. In fact, they cannot be the same or else the LER will be significantly overestimated. They are also not necessarily some linear combination of the previous equations. That is, c_2 and c_4 are not necessarily identical to their previous values. Equation 5-13 cannot be linearized in any fashion similar to equations 5-4 and 5-12, though. This would be possible if some relationship between the background contrast and slope contrast existed, then one form of contrast could be described in terms of the other. The only relationship the two have with each other is the relationship forced by the experiment. For example, when the background contrast is 0.58, the slope contrast is 11.09 mJ/cm^2·μm. Using this forced, or false, relationship will not lead to clear conclusions. There is no reason why the slope contrast has to be 11.09 mJ/cm^2·μm when the background contrast is 0.58. This is simply how the experiment was run. Therefore a non-linear fit needs to be made in order to solve for the fitting parameters.

The non-linear least squares fit used to find the optimal constants, c_1 - c_4, is significantly more complex than for the linear least squares fit. The JMP software uses the Gauss-Newton

iterative technique with step-halving. This iterative technique is described in detail in statistical textbooks, such as reference [58]. The results of this more complex method must be interpreted more carefully, as well. Iterations of the parameters to be solved continue until a least squares minimum is found. This can lead to a local minimum, or a false minimum. It is also possible that this method will not be able to find the proper global minimum after testing several different starting points. Critically important, there is no well defined R^2 statistic, ANOVA cannot be performed, and errors of the estimates are approximate, unlike in the linear case above.

First the accuracy of the non-linear least
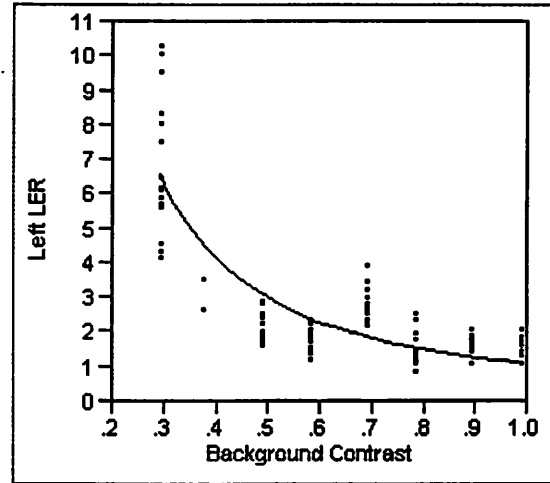
Figure 5-14: Fitting left side LER versus background contrast directly with a power model, equation 5-3, using non-linear least squares fitting. Table 5-8 contains the fitting parameters.

squares fitting is tested by solving equation 5-3 directly instead of linearizing it to equation 5-4 as done above. Figure 5-14 is a plot of the fit and table 5-8 contains the fitting parameters, c_1 and c_2, which are now calculated directly. The parameter estimations are significantly different from those calculated in the more rigorous linear least squares fit. In the linear least squares fit, c_1 is

	SSE	DFE	MSE	RMSE
	109.44960245	96	1.1401	1.0677547
Parameter	Estimate	ApproxStdErr	Lower CL	Upper CL
c1	1.0906678735	0.10598966	0.87635157	1.3211745
c2	-1.464723156	0.09363554	-1.6717188	-1.2734569

Table 5-8: Fitting parameters, c_1 and c_2, solved directly using non-linear least squares fitting. SSE, DFE, MSE, and RMSE are the same as defined earlier: sum of squares error, degrees of freedom for error, means square error, and root mean square error. The parameter estimations do not match those of the linear least squares fitting.

73

1.332 and c_2 is -1.113. With the non-linear fit, c_1 is 1.091 and c_2 is -1.465. The non-linear parameters do not even fall within the 95% confidence interval of the linear fits, and vice-versa.

Although the non-linear least squares fit does not seem as accurate, this model was attempted nonetheless. Examining linear combinations of background and slope contrast, of background and standard contrast, of slope and standard contrast, and of all three yielded no successful results. As mentioned above, the parameter estimates found in table 5-8 were calculated iteratively. After many iterations and testing several starting points for the initial iteration, the final parameter estimates, c_1 and c_2, were found. In all the cases using combinations of multiple contrasts, either the estimates were not converging onto a single value, or the value of the parameter estimate depended upon the initial starting value for the iteration. Regardless, the non-linear least squares model fitting method would not work for this task, and even if it did work, the results would be questionable.

As mentioned above, the parameters c_2 and c_4 used in equation 5-13 are not necessarily the same as those found individually as the exponential terms for the background contrast and slope contrast, respectively. However, one could logically hypothesize that the values should not change much. In other words, the shape of the curve representing the influence of each respective contrast upon LER should be similar. The author prefers to avoid making such an assumption, but must make it to allow for linear least squares fitting and analysis of variance. In this case, the equation fitting left side LER to background contrast and slope contrast is now:

74

Equation 5-14:

$$\text{Left LER} = A_1 + A_2 c_1 (\text{b.g. cont.})^{c_2} + A_3 c_3 (\text{slope cont.})^{c_4}$$

where, c_1 - c_4, are the same as those calculated when looking at solely background or slope contrast: 1.332, -1.1132, 15.15, and -0.8688, respectively. Parameters A_2 and A_3 are solved to find a ratio of which contrast dominates the LER estimation. Note that c_1 and c_3 are not actually held constant since $A_2 c_1$ is a new variable to replace c_1 from earlier, as is $A_3 c_3$. Any linear equation contains an intercept, which is A_1. A_1 should ideally be 0 since the two models combined should entirely predict LER, as they did in the previous linearized power models. Since c_2 and c_4 are kept constant, the values $c_1 (\text{b.g. cont.})^{c_2}$ and $c_3 (\text{slope cont.})^{c_4}$ can be calculated for each contrast level, thereby making equation 5-14 linear. Tables 5-9 and 5-10 contain the estimates for parameters A_1 - A_3 and the analysis of variance, respectively. Parameter A_3, the contribution due to slope contrast, has a negative estimate. This suggests that as the slope contrast gets worse, the LER improves. However, the 95% confidence interval includes 0. A better interpretation is that the effect of slope contrast is relatively insignificant compared to background contrast. The same is true of the intercept: while it is not 0, it is quite close to zero and the 95% confidence interval includes 0. Therefore the background contrast is almost

Term	Estimate	Std Error	Lower 95%	Upper 95%
Intercept	-0.256248	0.436046	-1.121914	0.6094173
BG Eqn.	1.3971907	0.104438	1.1898534	1.6045279
Slope Eqn.	-0.235607	0.204969	-0.642524	0.1713108

Table 5-9: Least squares fit for parameters A_1 - A_3 of equation 5-14, the dependence of left LER upon background contrast and slope contrast combined. "Intercept" is parameter A_1, "BG Eqn." is A_2, and "Slope Eqn." is A_3.

entirely accounting for

LER. Table 5-10 summa-

rizes the analysis of vari-

ance. The F Ratio is

exceedingly high and the

model is most certainly fit-

Analysis of Variance				
Source	DF	Sum of Squares	Mean Square	F Ratio
Model	2	275.17447	137.587	113.2060
Error	95	115.46019	1.215	Prob>F
C Total	97	390.63467		<.0001

Table 5-10: ANOVA for the model fitting background contrast and slope contrast to left side LER, equation 2-25. There are now 2 degrees of freedom in the model, corresponding to the two forms of contrast.

ting data. One must keep in mind that the assumption made earlier that c_2 and c_4 are the same as they were in the linear case is not necessarily true. The main conclusion that the background contrast accounts for almost all of the LER when compared to slope contrast can be stated with confidence, however the exact relationship is probably not that estimated in table 5-9 above.

All the possible linear combinations of LER versus contrast power are fit similar to equation 5-14. Table 5-11 (a) contains the fitting parameters and F ratio for left side LER versus

Left side of trench LER:
Linear Combination of contrasts (power function fitting):

A2 term	A3 term	Intercept estimate	A2 A2 estim.	A2 F Ratio	A3 A3 estim.	A3 F Ratio	overall F Ratio
background + slope cont.		-0.2562	1.3972	179	-0.2356	1.321	113.2
background + std. cont.		0.8219	1.4989	248.7	-0.8011	13.48	133.5
slope + std. cont.		0.9822	2.7686	19.69	-2.059	7.962	12.89

(a)

	Intercept estimate	B.G. estim.	B.G. F Ratio	Slope estim.	Slope F Ratio	Std estim.	Std F Ratio	overall F Ratio
All 3 contrasts	0.91	1.4039	235.6	1.3861	15.97	-2.1408	29.85	108.3

(b)

Table 5-11: Least squares fitting parameters and ANOVA to fit left side of trench LER with all linear combinations of contrast models. The intercept (A_1), A_2, and A_3 terms are those described in equation 5-14. Figure (a) shows all linear combinations of two forms of contrast, figure (b) has the linear combination of all three forms of contrast.

76

linear combinations of two forms of contrast; table 5-11 (b) contains the same information for a linear combination of all three forms of contrast. A_2 and A_3 terms are those described in equation 5-14. For example, the full equation for the linear combination of background and standard contrast, using the parameters estimated in table 5-11 (and the parameters estimated for the power fitting in table 5-4), the following relationship is found:

Equation 5-15:

$$\text{Left LER} = 0.82 + 1.50 \cdot 0.29 \cdot (\text{bg})^{-1.11} - 0.80 \cdot 0.54 \cdot (\text{std})^{-1.14}$$

or:

Equation 5-16:

$$\text{Left side LER} = 0.82 + 0.43 \cdot (\text{bg cont.})^{-1.11} - 0.43 \cdot (\text{std cont.})^{-1.14}$$

In order to simplify the presentation of data somewhat, an F ratio is given with each parameter instead of a $1\,\sigma$ range of parameter values. This ratio is to test whether the effect is zero. A high F ratio means that the effect is significant and that parameter is not just fitting random noise. The parameter F ratio, or effect F ratio, is the ratio of the mean square for the effect divided by the mean square for error, similar to how the other F ratios are calculated. The F ratios for the intercepts are not given because they are not effects. Table 5-12 contains the parameters and F ratios fitting combinations of contrast with the right side of the trench LER.

It is more difficult to draw definitive conclusions from the fits made with the linear combination of contrast terms than it is with one contrast definition. Since the individual power relationships were fixed, the parameters given in tables 5-11 and 5-12 are not fully optimized. Nevertheless, some powerful deductions can be made. First the intercepts, the A_1 parameters, are examined to give some degree of confidence that the fitting equations are reasonable. The

Right side of trench LER:

Linear Combination of contrasts (power function fitting):

A2 term	A3 term	Intercept estimate	A2 A2 estim. F Ratio		A3 A3 estim. F Ratio		overall F Ratio
background +	slope cont.	-0.4731	1.0145	129.8	0.2504	3.559	110.5
background +	std. cont.	-0.0453	1.1354	156.4	-0.0846	0.374	105.4
slope +	std. cont.	0.0969	1.4888	16.68	-0.4879	1.715	20.36

(a)

	Intercept estimate	B.G. B.G. estim. F Ratio		Slope Slope estim. F Ratio		Std Std estim. F Ratio		overall F Ratio
All 3 contrasts	-0.2366	1.0756	168.8	1.0541	22.76	-0.9866	18.98	94.09

(b)

Table 5-12: Least squares fitting parameters and ANOVA to fit right side of trench LER with all linear combinations of contrast models. The intercept (A_1), A_2, and A_3 terms are those described in equation 5-14. Figure (a) shows all linear combinations of two forms of contrast, figure (b) has the linear combination of all three forms of contrast.

1 σ range for the estimates includes zero in nearly all cases. These ranges are not given in tables 5-11 and 5-12. The only fits where the intercepts do not include zero within 1 σ are: (1) Fitting left side LER versus the combination of all three contrast definitions, and (2) fitting right side LER with the combination of background contrast and slope contrast. In the former case, the zero is at 1.05 σ below the estimate and in the latter case the zero is 1.02 1 σ above the estimate. It is expected for the intercept to be zero in all the linear combination fits since each fit on its own, with a pre-factor equal to one, fully accounts for the LER. If the intercepts do not approximate zero, one can concluded that the linear combination fitting is unsuccessful. However, having all the intercepts approximately equal to zero does not prove that the model is successful.

Next, the linear combination parameters A_2, A_3, and sometimes A_4 are examined. The parameter F ratios are large enough that the probability of type I error-- where the null hypothesis is mistakenly rejected-- is less than 0.01%. There are a few parameters in some of the

equations with lower F ratios. The F ratio for the slope contrast parameter in the equation fitting background and slope contrast versus left side LER is only 1.32, yielding a 25% chance of type I error. The ratio for the standard contrast in the fit of background and standard contrast versus right side LER is 0.3736, yielding a 54% chance of error. For the standard contrast in the fit of slope and standard contrast versus right side LER the F ratio is 1.72, a 19% chance of error. For the slope contrast parameter in the fit of background and slope contrast versus right side LER it is 3.56, yielding a 6.2% chance of error. In all cases, the fit for the background contrast parameter is strong enough that the F ratio never drops below 129.8, the F ratio for the standard contrast parameter is never above 29.8, and the best F ratio for the slope contrast parameter is 22.8.

Finally, the entire equations as a whole are examined statistically. The overall F ratio fitting slope and standard contrast to left side LER is 12.9 and the overall F ratio fitting slope and standard contrast to the right side LER is 20.4. The lowest overall F ratio that fit background contrast and any other combination of contrast is 94.1. All of the overall F ratios are high enough that the probability of type I error for the entire model overall is far less than 0.01%.

There are some significant deductions that can be made from the information summarized above. Foremost is that all the fits above are statistically significant. None of the overall F ratios are low enough to question this. Next is that the background contrast parameter is by far the most statistically significant of the three forms of contrast analyzed. This is seen on two levels: (a) the overall F ratios are always drastically higher when background contrast is included in the fit, and (b) the background contrast parameter F ratio is always appreciably higher than the F ratios of either slope or standard contrast. Finally, the background contrast

fits the data best alone, next best with one other parameter added, and worst with all three parameters. As mentioned initially, background contrast cannot solely be used to analyze the effects of a poor aerial image profile upon LER because it does not take the foreground contrast into consideration at all. This analysis suggests that perhaps neither the aerial image slope nor the standard contrast are the best methods to measure the contribution due to the foreground image. Another possibility is that the linear combination of fixed power law fits is not sophisticated enough to capture the effect of the slope contrast or standard contrast upon LER while also capturing the effect of the background contrast upon LER. The data does not allow for any further insight in this respect, however the author believes the latter to be a more accurate assessment.

It is difficult to determine whether the slope contrast or standard contrast provide the second best fit after background contrast. Some form of measuring the foreground image profile is necessary, but the data and analysis above do not provide any further insight.

A summary of all the raw data collected for this LER versus aerial image contrast experiment is presented in *Appendix A2*.

5.6 LER versus Contrast (UVII-HS Resist) Summary

The correlation between aerial image contrast models and LER are further studied, specifically focusing upon the correlation with a newly defined contrast definition, background contrast. In order to allow larger statistical sample sizing within the resist study, the experiment is limited to one photoresist, UVII-HS. The three forms of contrast examined-- standard contrast, slope contrast, and the new background contrast-- are orthogonalized as best as possible. Because of asymmetry in the SEM electron collector, the LER measured on the left (top) side

of the trench is analyzed separately from the LER measured on the right (bottom) side of the trench. Rigorous statistical testing, including analysis of variance, is performed upon the correlations found. Correlations of each contrast definition versus LER are executed, as well as correlations of combinations of contrasts versus LER.

The background contrast fits best with LER, in both the left side and right side LER cases. The relationship is approximately inverse, with left side LER $\propto$ background contrast$^{-1.11}$ and right side LER $\propto$ background contrast$^{-0.80}$. There are some difficulties with the statistical analysis of correlations involving combinations of contrast definitions. The analysis is not as rigorous as that for one form of contrast alone, however it shows that correlations which involved background contrast as one of the forms of contrast fit the data significantly better than combinations with only slope and standard contrast. Moreover, the correlation with solely background contrast fits better than any combination of contrasts.

The significance of these results in semiconductor processing is obvious, however the physical insight is much more elusive. Previous studies regarding LER, such as those mentioned in chapter 1 and earlier in this chapter, did not consider background flare to be a source of LER in and of itself. Background flare was only a concern with regards to LER in that it degraded the normalized image log slope or the standard contrast, or whatever other definition of contrast that was used which incorporated the foreground contrast as well.

The possible physical mechanism driving this strong correlation with the background dose was briefly touched upon in chapter 4. The mechanism is not much better understood after this analysis. Since the background contrast fits with LER far better than slope contrast or standard contrast, it seems possible that another phenomenon beyond diffusion accounts for the superb fit with background contrast. The author cannot confidently provide any physical

insight explaining this previously unobserved phenomenon. As mentioned earlier, Postnikov claimed that photogenerated acid diffusion is essentially halted when that acid is in a polar region, and Yuan claimed that diffusion in DUV photoresists is reduced, non-fickean diffusion. This explains how acids can easily move through unexposed regions of the photoresist, or the background. However, this still does not explain why levels of acid concentration beyond those needed to cause the photoresist to develop (unexposed background) would affect the roughness of the trench or line edge.

5.7 LER Dependence upon Aerial Image Contrast-- Overall Summary

A new definition for aerial image contrast was created, background contrast. The background contrast is a measure solely of the intensity of light reaching areas meant to be unexposed. In order to allow further statistical study, several measurements were repeated at each contrast level and 8 separate contrast levels were set. These levels were designed such that the three forms of contrast examined were as orthogonal to each other as possible. The three forms were background contrast, standard contrast, and image slope contrast. The peak intensity contrast was discarded since it was not fitting any of the data well in the previous experiment. This more rigorous experiment was designed with only UVII-HS photoresist.

Background contrast correlated with LER far better than either of the other forms of contrast. Linear combinations of the different forms of contrast, including combinations of background contrast and another form of contrast, did not correlate with LER as well as background contrast alone. This is an unexpected and somewhat counter-intuitive results. One would expect the foreground image to be most important in determining LER, however this experiment quite conclusively proved that it is the background dose, or flare, that most

strongly affects LER, at least for UVII-HS photoresist. As mentioned earlier and in several of the references, such as [15], photoresists all react differently with regards to how degraded aerial image contrast affects LER. However, they all show similar trends. It is therefore at least feasible, if not expected, that the background contrast will greatly affect LER for other photoresists as well.

This discovery is also significant because flare is becoming an increasingly critical issue as the exposure wavelength shrinks. The mechanics are discussed briefly above and in further detail in [53, 59]. The essence is that flare scales approximately inversely with the wavelength squared. Flare will become a much greater issue as the step towards EUV lithography is made. Many of the poorer contrast levels examined above that were made artificially might become actual levels of background contrast at EUV. For instance, the best reported EUV flare levels in 2001 were 20% [59]. Assuming the level can be brought down to roughly 10% and the exposure dose is still at 3.3 times the minimum dose to clear, the background contrast would be 0.67. At this level of flare, the expected contribution to LER from the poor aerial image would be 1.89 nm 3 σ LER, yielding total LER 5.93 nm 3 σ LER. (This calculation took the best fits to LER, those using just background contrast, and used the average of the left and right LER values.) Thus, the background flare contributes to 32% of the photoresist LER in this case. These numbers will be vastly different once EUV lithography is possible for semiconductor processing, but it is clear that background flare will have a strong impact upon LER.

6 SEM Image Capture and the Deblurring Algorithm

This chapter first discusses scanning electron microscope imaging, particularly as it applies to LER measurement. Then theory behind imaging and image capture is discussed, with special application to SEM images. Next, deconvolution algorithms used to deblur imperfect images are described. The Richardson-Lucy algorithm is first described, a precursor to the maximum likelihood algorithm used throughout this study. Then the maximum likelihood algorithm is discussed. Some time is devoted to explaining why this algorithm, which is able to recover image data in the high frequency domain, is particularly useful for measuring the LER seen on SEMs (or any other tool imaging LER, for that matter).

6.1 Advantages of LER Measurement via SEM

Line edge roughness is a relatively new concern in the semiconductor industry. Measurement techniques are still being developed, and no formal, standardized definition of LER has been determined [50]. Generally, LER from SEMs is most common method of roughness determination, however researchers also use SER, line width roughness (LWR), spatial frequency domain analysis and correlation lengths [72], and other methods. Other metrology tools such as optical scatterometry [24] and small-angle neutron scattering (SANS) [25] have been used to measure photoresist roughness, however SEMs have been the workhorse in the field.

With regards to LER measured from SEMs, there has been no formal decision throughout the industry as to which method of threshold detection is preferred to determine the line edge. Studies have shown that while different threshold detection algorithms yield different LER values, it is difficult to determine which algorithm is best [17, 23, 26]. Nelson compared several different methods of line edge detection, all of which computed somewhat different values for LER. None of them were found to be significantly superior to others using standard metrics, such as improved measurement repeatability. Yet the most difficult task in measuring photoresist roughness may be capturing the data. In order to measure roughness on the order of 2 nm RMS roughness accurately, the measurement capabilities need to be significantly better than that, on the order of 1 nm, using measurement tools such as the AFM and SEM.

LER measured by an SEM became the dominant means to measure photoresist roughness for several reasons. First, the scanning electron microscope is a relatively cost-effective tool already present in most semiconductor research laboratories and semiconductor manufacturing cleanrooms. The SEM also has a long and successful history, giving it excellent credibility within the metrology community. Max Knoll collected the first scanning electron image in 1935, and the first scanning electron microscope recognizable in today's form was designed by Sir Charles Oatley in 1952 [60]. Next, images can be collected very quickly, usually within a few seconds, depending upon the length of time averaging used to collect the image. Finally, the ultimate resolution of the SEM is quite fine. With the advancements that have been made over the years, the resolution achieved in practice has improved to around 1 nm or better. Finally, the operation of a scanning electron microscope, at least in principle, is relatively straightforward since it is quite similar to an optical microscope. This conceptual similarity between photon and electron beam microscopes has allowed relatively rapid advances

in SEM engineering and has simplified SEM use for operators already familiar with standard optical microscopes.

There are noteworthy disadvantages to the SEM as well. Most significantly, the tool can charge the photoresist during image capture. Without a means to discharge the insulating photoresist, it can heat up and become damaged [51]. In fact, energy deposition into the photoresist from bombarding electrons is so significant that electron beam lithography uses the same basic principles as SEM imaging. Ocola discusses how various electron acceleration voltages and energies affect energy deposition into the photoresist [74]. Proper precautions such as lower electron acceleration voltages can avoid such damage, however this is becoming more difficult more fragile resists designed for ArF and F_2 lasers.

6.2 Image Capture-- Image Blur and Noise

Since scanning electron microscopes function similar to optical microscopes, the methods used to manipulate and improve images created using optical tools should also improve images created using the electron microscope. Image manipulation is a growing field with a great deal of research being conducted. This chapter utilizes an image processing algorithm originally developed for use in astronomy and other optical image processing fields in order to improve the quality of the SEM images collected in the high frequency domain. This algorithm is the maximum likelihood algorithm. It recovers some of the high resolution detail, or the high frequency components, of the image by deblurring it. Before discussing the deblurring algorithm, SEM image capture and image blur must be understood.

SEM images are captured by focusing stream of electrons impinge upon a particular location in the sample. The location is raster scanned across the sample analogous to a television

set. The secondary electron collector measures the signal intensity of electrons that are ejected from the sample by electrons from the incident beam. It can pinpoint the location of the ejected secondary or backscattered electrons measured by comparing the time that the secondary electron arrived on the sample with the location of the raster scanning electron beam at that moment in time. The electron collector has no means to measure the location of origin of the electrons collected except by correlating the time to the location of the raster scanning electron gun. Typically the signal to noise ratio of SEM tools is quite low, since this process of emitting electrons from a sample is not very efficient. Some form of time averaging is calculated when the image is captured to improve the signal to noise ratio. Further discussions on the nature of secondary electron emission and backscattering electrons can be found in references [64-66].

The captured image, or any image, can be modeled as follows:

Equation 6.1:

$$g = h \otimes f + n$$

where f is a perfect image of the object, g is the blurred and noisy image, h is the point-spread function (PSF) of the system, and n is the (Poisson) noise. This model of an image is ideal for SEM image capture. h represents the blur caused by the SEM stage or optics. Yoshizawa [49] and Pfeiffer [52] discuss some of the sources of blur occurring in electron optics, most especially blur due to the dispersion of similarly charged electrons travelling in close proximity to one another. Image noise is represented by n. As mentioned above, SEM images are typically quite noisy. Figure 6.1 (a) is an SEM image with a significant level of image blur, and figure 6.1 (b) contains a high level of image noise. With blur, it is difficult to determine the exact shape of a sharp feature like a line edge. In the case of image noise, there is a large variation
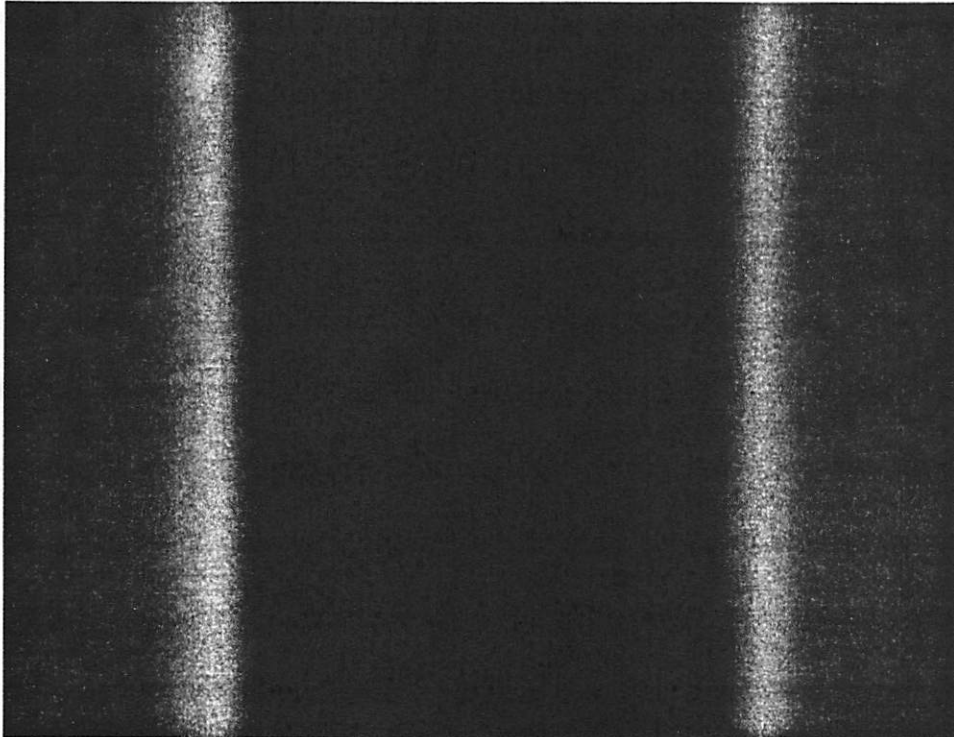
(a)
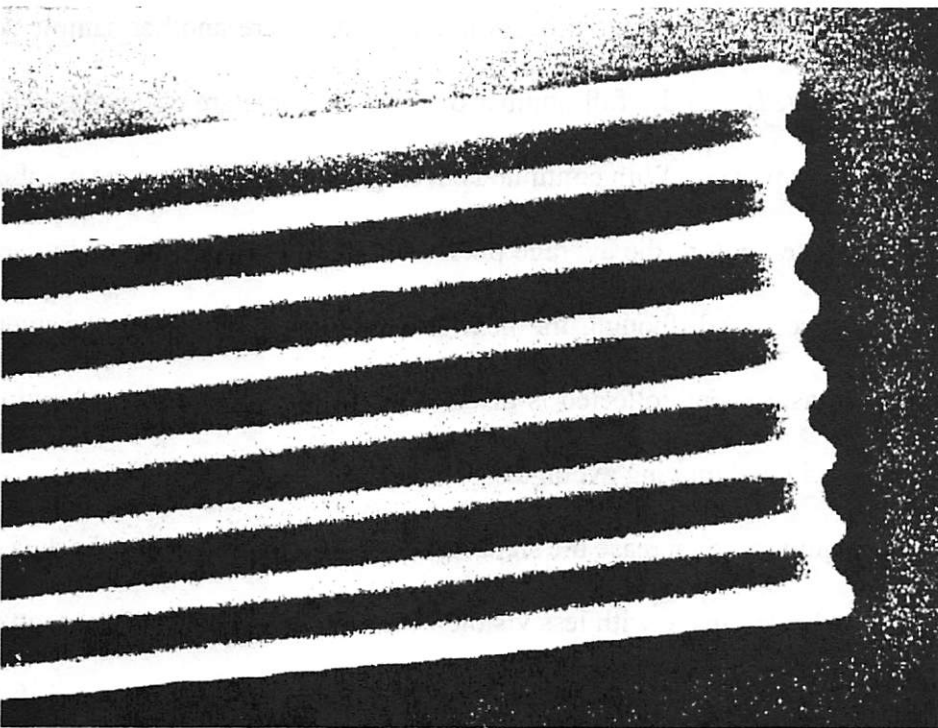
(b)

Figure 6.1: Examples of SEM image blur (a) and SEM image noise (b). Image (a) is a single trench. Image (b) is a tight pitch line:space pattern.

in signal intensity throughout an area that should yield a uniform signal. In figure 6.1 (b), for example, the area outside the line:space pattern should be a uniform grey or black color. Instead there are many white specks of noise in the image. Both noise and blur will adversely affect the ability to measure LER from the image. The level to which each

affects the measurement depends upon the type of measurement algorithm used. The maximum likelihood algorithm can compensate for image blur when noise is present, however it cannot compensate for the noise itself. SEM image noise is most often compensated by some sort of time averaging technique.

Typically two types of time averaging techniques are used to limit noise: pixel averaging and continuous averaging. Both techniques rely upon the same basic principle. The collection of secondary or backscattered electrons is a statistical phenomenon. As the sample is exposed for longer periods of time, the average signal recorded by the collector approaches the true mean of the signal. Anomalies affect the measured average less as the sample size increases. In the case of pixel averaging, a set number of measurements are collected for an individual pixel and, once collected, the average of the measurements is reported as a grey-scale pixel value on the SEM image. This process is repeated, where another sample set is collected and it is not reported until the full number of measurements are collected. This is done for all the pixels being imaged. With continuous averaging, there is also a set number of measurements being collected before the average pixel value is reported. Once a full set of measurements have been collected, though, the next measurement is added to the average value while the oldest measurement collected is discarded. In this manner the averaging is continuous, similar to a weighted moving average.

Both forms of averaging try to increase the signal to noise ratio by averaging the data collected over time. This yields an image with less visible Poisson noise, however it accentuates any errors introduced over time. Some examples of such error are electron beam drift and stage drift. In the former case, the electron beam does not return to exactly the same location when it retraces over pixels during the raster scan. In the latter case, either the sample or the

sample stage mechanically move over time. These time-dependent errors appear as blur on the image. In other words, these errors can be described as a convolution of the ideal image, f in equation 6.1, and the time-dependent variation h. Therefore, to a certain extent, Poisson noise can be lowered by increasing blur, or vice-versa.

Blur can be introduced by several mechanisms other than the drift mentioned above. Any optical aberrations, such as defocus, astigmatism, or aperture alignment, occur in electron optics as well. All non-idealities during image capture that affect all pixels equally can be lumped together as blur and represented as a convolution of the non-ideality's point-spread function (PSF) and the image. Moreover, these blurring mechanisms will generally take information stored in one pixel and spread that information through several neighboring pixels. For example, defocus aberrations will affect all pixels throughout the image equally, and the information stored in each pixel of the idealized image will instead be represented by that pixel and several of its neighbors. It is the former-- this uniform effect throughout the image, this linear phenomenon-- that the Richardson-Lucy (RL) algorithm and the maximum likelihood algorithm utilize to remove blur and recover the image. It is the latter-- the representation of one pixel in the ideal image by several pixels in the blurred image-- which causes the actual blurred image to lose the high frequency components of the image. The quality of the image is degraded greatly when the high frequency components of the image are lost. This is especially true when extracting LER values from SEM images. In most of the images collected throughout chapter seven, images taken at 50,000 x magnification, a pixel represents slightly more than 2 nm. If a line edge that should occur at one pixel location is instead blurred such that the transition through the line edge consumes 5 pixels, the line edge calculated from the blurred image could be off by 9 nm with respect to the idealized image. These

are the essential concepts that drive the desire for deblurring: (a) the ability to recover the image by utilizing the linear effects of blurring, and (b) the need for the high frequency components of the spectrum to create quality images that can be used for quantitative analysis.

6.3 The Blind Deconvolution Algorithm

An actual image may be represented as g in equation 6-1, with blur and noise reducing the image quality. As mentioned above, blur is uniform throughout the image, or linear. It can be represented as a single function, or in this study a single matrix, that is not dependent upon pixel location. This same function is applied to each pixel, unlike image noise, which is not uniform throughout the image. The iterative Richardson-Lucy algorithm removes some of the blur and recovers the image in the following manner [63, 27]:

Equation 6-2:

$$\hat{f}_{k+1} = \hat{f}_k\left(h * \frac{g}{h \otimes \hat{f}_k}\right) \equiv \psi(\hat{f}_k)$$

where $\hat{f}_k$ is the estimate of f after k iterations, * is the correlation operator, and Ψ is defined as the Richardson-Lucy function. The quantity $h \otimes \hat{f}_k$ is the reblurred image. As the estimate of the deblurred image, $\hat{f}_k$, approaches the perfect or ideal image of the object, f, the quantity $g/h \otimes \hat{f}_k$ approaches unity. This is because $h \otimes f$ is in fact g.

The algorithm used in this study is the maximum likelihood algorithm, an alteration of the Richardson-Lucy algorithm created by Biggs and Andrews [62, 63, 27], with the alterations made to help accelerate it. It is a blind deconvolution algorithm. The blurring PSF, h, need not be known. It is solved simultaneously with the deblurred image.

The maximum likelihood algorithm accelerates the estimation of the next iteration by magnifying the amplitude of the vector used to determine the value of the next iteration, as well as calculating a new direction for that vector, taking into account the magnified amplitude. The following examines the iteration of a single pixel. Let x_k be the iterated point, y_k be the predicted point, h_k be the direction vector, and α_k be the acceleration parameter. Then:

Equation 6-3:

$$y_k = x_k + \alpha_k h_k$$

where

Equation 6-4:

$$h_k \equiv x_k - x_{k-1}$$

Equation 6-5:

$$x_{k+1} = y_k + g_k$$

Equation 6-6:

$$g_k \equiv \psi(y_k) - y_k$$

This method uses vector extrapolation, adjusting the direction of the vector by taking into account the vector magnification, or the acceleration parameter, α_k. Biggs and Andrews delve into further detail in reference [27-29]. The Matlab script used to perform the maximum likelihood deconvolution is copied in *appendix B1*.

The essence of the blind deconvolution algorithm can be described in the following manner. Any level of blur in the image forces some loss of high frequency data. In other words, the blur makes neighboring pixels correlate more closely with one another. Therefore, in order to deblur the image and restore the high frequency components of the image, neighbor-

ing pixels in the deblurred image should be as uncorrelated with each other as possible. When examining a pixel location, the next iteration of the PSF should be calculated such that the neighboring pixels look as dissimilar to the specified pixel as possible. This is iteratively calculated throughout all pixel locations in the image. The final "deconvolving" PSF is therefore the closest estimate to the PSF which created the blur in the original image. The final "deconvolved" image is the closest estimate to the perfect image. The assumption that there should not be any correlation between neighboring pixels can be false in some instances, but this will be addressed in chapter seven.

6.4 Deblurring an SEM Image

By recovering the higher frequency components of the image, more accurate LER measurements are expected. LER of deblurred images is measured identically to that of chapters four and five, except that the original SEM image is first deblurred using the maximum likelihood algorithm described above. Before applying the algorithm, several SEM images were first magnified and studied to determine the average size of the blur seen in most cases. The maximum likelihood algorithm can blindly calculate the PSF shape and the deblurred image, but it must first be given a PSF size. Portions of SEM images with high frequency data were analyzed in order to
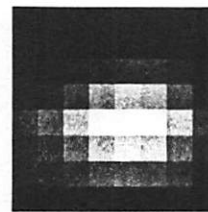
(a)
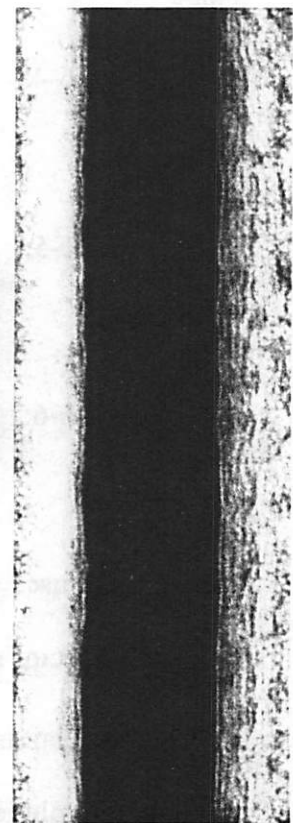
(b)

Figure 6-2: (a) is the PSF and (b) is the deblurred image from the SEM image in figure 3-2 (a). The deblurring algorithm was iterated nine times to obtain the PSF and image shown here.

93

determine the typical size of the image blur. In all cases, the blur was never larger than 8 pixels, and typically on the order of around 4-5 pixels. At 50,000 x magnification; each pixel is 2.22 nm for the Leo SEM tool used. An 8 pixel PSF size is chosen in order to capture all of the blur seen, thus the PSF is 17.8 nm x 17.8 nm. After the SEM image is captured, the image is deblurred off-line using Matlab with the maximum likelihood blind deconvolution algorithm. An SEM image similar to figure 4-3 (a) is read into Matlab. The initial PSF shape estimate is not important since it is solved; an 8 x 8 pixel matrix of ones is set as the initial PSF. The blind deconvolution is solved iteratively. The blurring PSF is solved simultaneously with the deblurred image. After several iterations, the shape of the PSF converges upon a value and the deblurred image is stable. Figure 6-2 (a) is the PSF solution and figure 6-2 (b) is the deblurred image solution after 9 iterations. Figure 6-2 is the deblurred solution to figure 4-3 (a). The deblurred image is then fed into the Matlab code to extract the LER data and images similar to figure 4-3 (b) and 4-3 (c) are extracted, as always.

Simply comparing the trench edges of figure 6-2 (b) to those of figure 4-3 (a) by visual examination, there are a few noteworthy changes. First, it is evident that the deblurred image has recovered the higher frequency components of the image to some degree. Next, the deblurred image has obtained some level of undesirable lower frequency error introduced. In some locations there are discontinuities on the order of the PSF, spots roughly 18 nm in size. These spots should not appear in the ideal image, but they are inevitably created via the reblurring process. A low level, systematic error is introduced in the low frequency spectrum of an image in order to recover data in the high frequency spectrum [62, 63, 67]. Both of these issues, as well as several others, will be covered in more detail in chapter seven.

6.5 Image Capture and Deblurring Summary

Several reasons were listed explaining why LER measurement via SEM image analysis is an excellent choice for measurement photoresist roughness. The fundamentals of SEM image capture were then described. SEM images require some form of time averaging, which can increase image blur. The Richardson-Lucy and the maximum likelihood algorithms were illustrated; the maximum likelihood algorithm is used in chapter seven. They provide a means to recapture the high frequency components of an image by removing the image blur, thereby providing enhanced LER values. Finally, an SEM image deblurring example was presented, enumerating the simple steps required to deblur an image.

7 Utilizing the Deblurring Algorithm for LER Collection, and Verifying the Efficacy and Validity of the Algorithm

The maximum likelihood deconvolution algorithm described in chapter six is used to deblur SEM images. LER is measured on these deblurred images to determine the effect the deblurring algorithm has upon LER. First the LER values of the deblurred images are compared to those of the original images, and any trend is observed. Next the LER values of several measurements of a sample on several different SEMs are collected to determine whether the LER variation amongst individual measurements and machines is tighter after deblurring.

The validity of the deblurring technique is verified. Several verification methods are applied. First, images with little information are deblurred, to verify whether the technique can deblur such images. Then images with repeating patterns are deblurred, for similar reasons. Next, the PSFs of artificially rotated, deblurred images are compared to normal deblurred images. Finally, extreme cases where the algorithm will fail are demonstrated, highlighting how one can tell when the algorithm fails.

A few other useful applications for the deblurring algorithm finish this chapter.

7.1 Effect of the Deconvolution Algorithm upon LER

The maximum likelihood algorithm has not been used previously to extract finer LER measurements from SEM images; the effect upon LER is unknown. The utility of the maxi-

mum likelihood algorithm has been proven in the fields of image processing and astronomy by the authors such as Boden [61]. In order for this algorithm to prove useful in LER metrology, however, a few requirements must be met. First, the algorithm must in fact change the LER measurements. Secondly, the new LER measurements must somehow be proven more accurate. Finally, the algorithm itself must be robust and trustworthy. This section will address the first issue, showing that LER values are in fact changed significantly.

Figure 7-1 shows the variation of LER for a set of measurements of 12 different samples over several deblurring iterations. LER values in the original sample are all normalized to one, so that variations in LER due to deblurring can be measured as deviations from unity. Deblurring iteration 0 is the original SEM image, hence the normalized LER equal to 1 at iteration 0 for all cases. In this case, there were 10 deblurring iterations. As mentioned above, the PSF for the initial iteration is simple an 8 x 8 pixel matrix of ones-- a grey plane. Therefore, there were only 9 solved PSF iterations and 10 solved deblurred image iterations. (A calculation is made using the grey plane PSF to produce a deblurred image.) Throughout this study, when the author quotes a number of iterations, he is referring to the number of
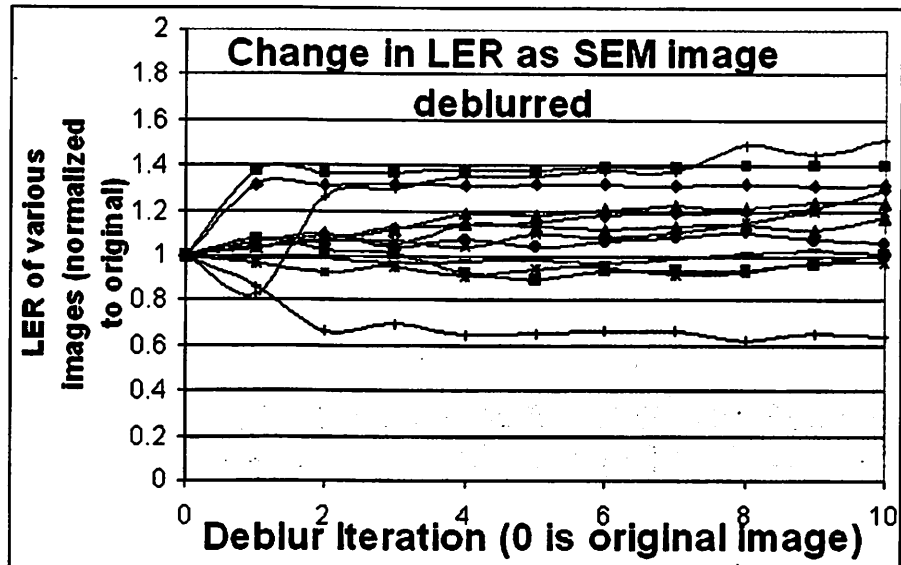
Figure 7-1: Variation of LER measurement as image is deblurred. 12 different samples shown. Normalized such that 1 is the original image LER.

solved PSF iterations. In this case, there were 9 solved PSF iterations (iteration 0 on figure 7-1 is the original image and iteration 1 is the iteration with the grey plane PSF). Two key results can be seen from this test. First, the LER changes drastically in the first two deblurring iterations, then holds roughly constant throughout the remaining iterations. Second, the change in LER can be quite drastic. The average normalized LER value for all samples for the last iteration is 1.152, or 15.2% average increase in LER. The maximum LER increase for the last iteration is 51.6% increase, and the minimum LER for the last iteration is 0.643, or a 35.7% decrease. Taking the average of iterations 2 - 9 (3 - 10 on figure 3-4), the average normalized LER for all samples is 1.112, or 11.2% average increase in LER. The maximum LER increase for iterations 2 - 9 is 40.2%, and the largest LER decrease for these iterations is 34.5%. Typically, larger changes in LER are seen when the measured LER is low in the original image. This is intuitive; a trench with large LER does not need a high resolution SEM image to detect it. However, when LER is already so smooth that it is barely detectable under ideal SEM conditions, any further blurring of the image due to SEM imaging non-idealities more easily distorts the calculated LER. Assuming that this technique does indeed improve LER measurement accuracy, it can provide significantly improved LER measurement data.

7.2 Efficacy of the Deconvolution Algorithm with respect to LER Measurement

Assuming the effect of image deblur upon LER is significant, the distribution of several LER measurements of a single sample should be tighter after deblurring than before. A blurry image adds more noise to LER measurements than an image that is nearly perfect. An experiment is designed to determine whether repeated LER values measured on the same sample using various SEM machines are more tightly distributed after deblurring the images than

those same measurements before deblurring the images. Three different field-emission SEM

(a)

(b)

99
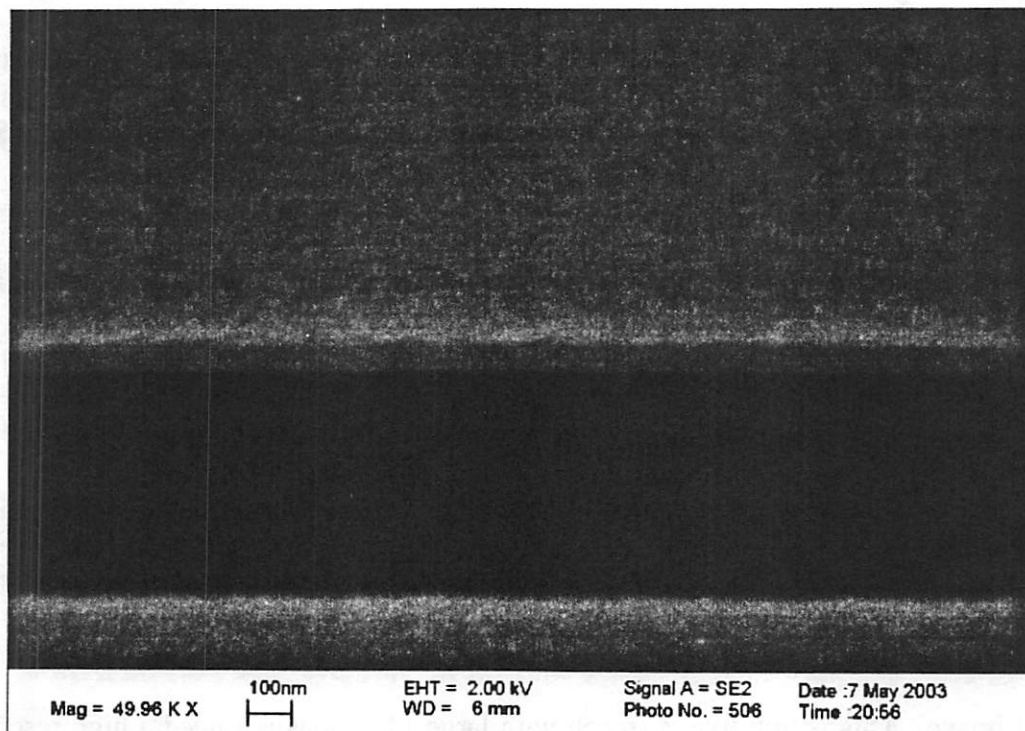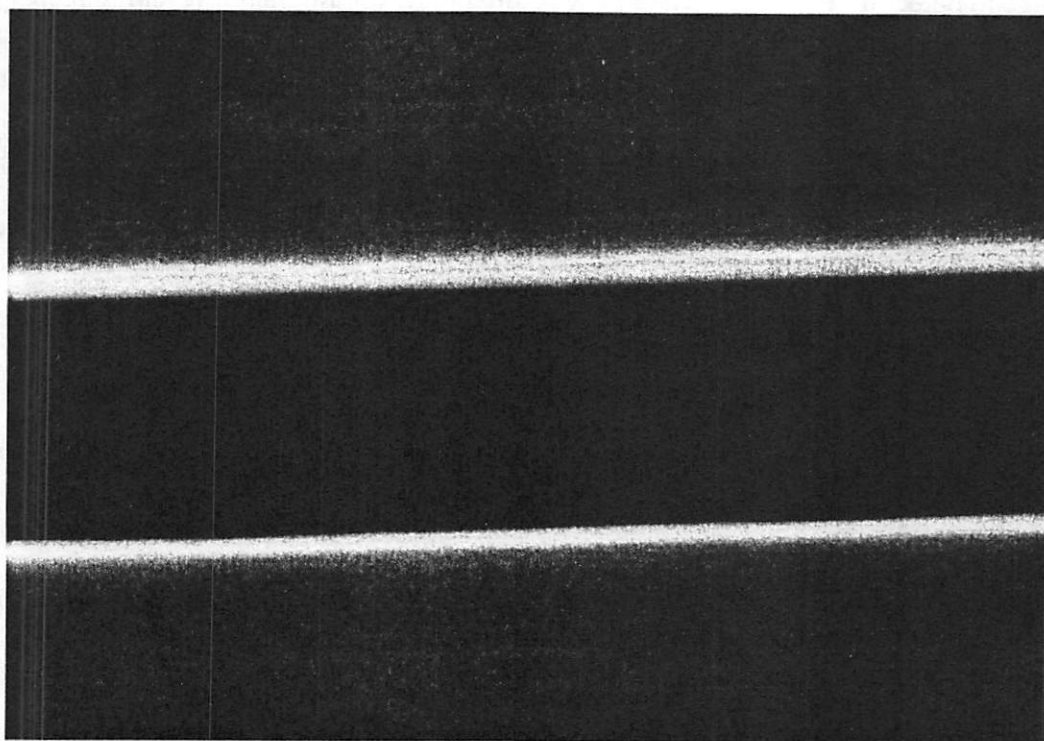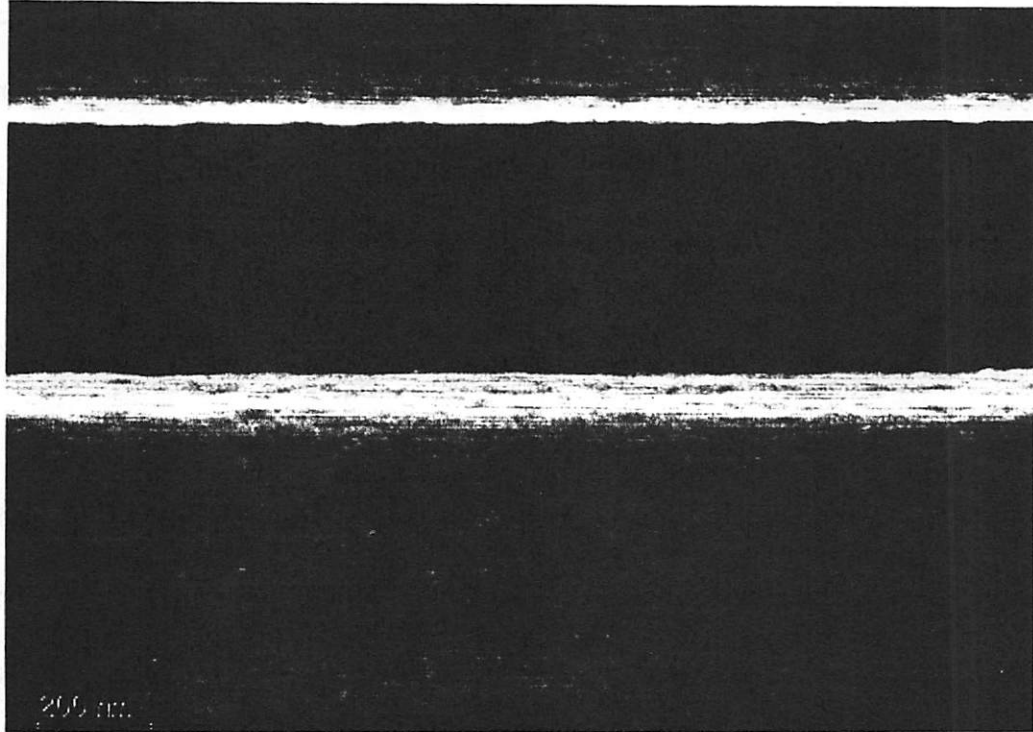
Figure 7-2: SEM images of the same sample trench, taken on different machines. Image (a) was captured on the Leo 1550 that has been used throughout this study, (b) was captured on the JEOL 6340 and (c) on the Hitachi S5000.

machines are used: The Leo 1550 that has been used throughout this study, a JEOL JSM-6340F at Lawrence Berkeley Laboratories (LBL), and a Hitachi S5000 run by the biology department of the University of California- Berkeley. Figure 7-2 contains SEM images of the same trench line taken using the three different SEM tools. Images (a) - (c) correspond to the Leo, the JEOL, and the Hitachi, respectively. Two samples coated with UVII-HS and exposed using the ASM Lithography stepper are studied. One received the best exposure level from the study in chapter four and the other received the worst exposure level from that study. Many LER measurements are taken with each sample and each SEM machine to allow statistical study.

LER measurements follow a chi-squared distribution, similar to the distribution of variances of normally distributed random variables. The variance between different LER mea-

surements taken does not quite follow the chi-squared distribution, but can be approximated as such. An often used statistical hypothesis test to compare whether two variances-- following a chi-squared distribution-- significantly unequal is the following:

Equation 7-1:

$$F_{\alpha,\,n_1-1,\,n_2-1} = \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}$$

where S_1^2 is the estimate of the variance and σ_1^2 is the true variance. If the two variances are equal to each other, they follow the F distribution. If not, they will not follow the F distribution. Targeting 95% confidence that the ratio of the variances is **not** unity, equation 7-1 can be rearranged as follows:

Equation 7-2:

$$\frac{S_2^2}{S_1^2} F_{0.025,\,n_1-1,\,n_2-1} \leq \frac{\sigma_2^2}{\sigma_1^2} \leq \frac{S_2^2}{S_1^2} F_{0.975,\,n_1-1,\,n_2-1}$$

and for the variances to indeed be different and the ratio not be equal to one, the range given in equation 7-2 cannot be equal to one. That is:

Equation 7-3 (a):

$$\frac{S_2^2}{S_1^2} F_{0.025,\,n_1-1,\,n_2-1} \geq 1$$

or:

101

Equation 7-3 (b):

$$\frac{S_2^2}{S_1^2} F_{0.975, n_1 - 1, n_2 - 1} \leq 1$$

in order for the two variances to be statistically different.

Four different pairs of variances are compared: the LER measurements of the left side of the trench of the sample with good aerial image contrast before and after deblurring the image, those before and after deblurring the right side on the sample with good contrast, those before and after deblurring the left side on the sample with poor contrast, and those before and after deblurring the right side on the on the poor contrast sample. Approximately 25 - 30 SEM images are captured on each machine. A total of 82 different SEM images are captured, creating 164 different images (original and deblurred) and 164 LER measurements. The deblurred images all pass through 9 iterations of deblurring before the final deblurred image is extracted. In none of these four cases are the variances significantly different. The strongest level of significance is the left side of the trench of the good image. In this case, there is a 39% confidence that the variances are different. This is an unacceptably low confidence in statistical testing. The poor level of confidence that the variances are different does not disclude the possibility that image deblurring improves the repeatability of LER measurements. Instead, this suggests that the variation between LER measurements along a trench or line is so large that any potential improvement (or deterioration) of LER accuracy and repeatability is overshadowed by the statistical variation in LER measurements.

Returning to the focus of section 7.1, the LER measurements before and after deblurring the sample with good contrast are compared. There are 46 SEM images taken with the good contrast sample; the remaining 36 images are the poor contrast sample. The average LER of

102

the good contrast sample before deblurring is 1.91 nm, 1 σ. The average LER of the same sample after deblurring is 2.05 nm, 1 σ. This is a 14.0% increase in LER measurements after deblurring. Earlier in this section an 11.2% increase in LER measurements after deblurring was reported.

7.3 Validity of Deconvolution Algorithm-- Low Information Image Recovery

Section 7.1 mentions that the algorithm must be robust and trustworthy when applied to SEM images. The maximum likelihood deconvolution algorithm requires an image that has a significant amount of information and contains no intentionally repeating pattern on the order of the size of the PSF. With regards to the former, the algorithm could not, for example, recover the blurring PSF of an entirely grey image. There is no variation in intensity between pixels. A grey image that is perfectly captured is identical to a grey image with blur. Such an image is meaningless. Yet, many SEM images used to calculate LER contain very little information aside from the line edge itself. The remainder of the image is grey, or nearly so. With regards to intentionally repeating patterns, the line edge does contain an intentionally repeating pattern on the order of the PSF. Assuming an image similar to any of those in figure 7-2, any point at the top edge of the trench has neighbors above it that are brighter, and neighbors below it that are darker. This is true all along the top edge. For these reasons, it is possible that the deconvolution algorithm may have difficulties properly deblurring SEM images.

To address the concern that SEM images of LER may not contain enough information, an image that does not contain much information, and certainly contains less than an image of a line or trench, is analyzed. An SEM image of an unpatterned area of photoresist is used as a low information image. First the image is taken with the best focus, astigmation, aperture

103

alignment, etc. Then the SEM electron optics are purposely unoptimized, to create a blurred image at exactly the same location on the sample. This blurred image is then deblurred off line using the deconvolution algorithm. Figure 7-3 is an example. Figure 7-3 (a) is an SEM image of a sample of unexposed SEPR-463 photoresist with proper focus, astigmatism, etc. Figure 7-3 (b) is an SEM image captured in exactly the same location on the sample, but with the electron gun purposely astigmated. This SEM image is deblurred off line. Figure 7-3 (c) is the deblurred image. Figures 7-3 (a) and 7-3 (b) are separate images, whereas figure 7-3 (c) is a reconstruction of figure 7-3 (b). It is evident from examining the images that the deblurred image looks significantly closer to the image with no astigmatism than the astigmated image does. The focus, aperture
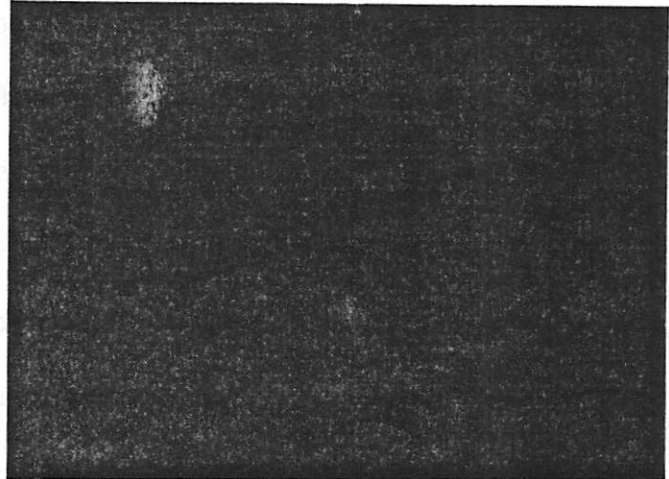
(a)
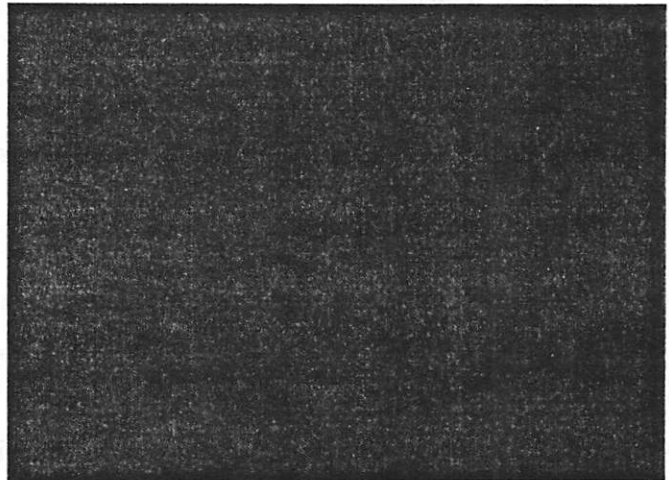
(b)
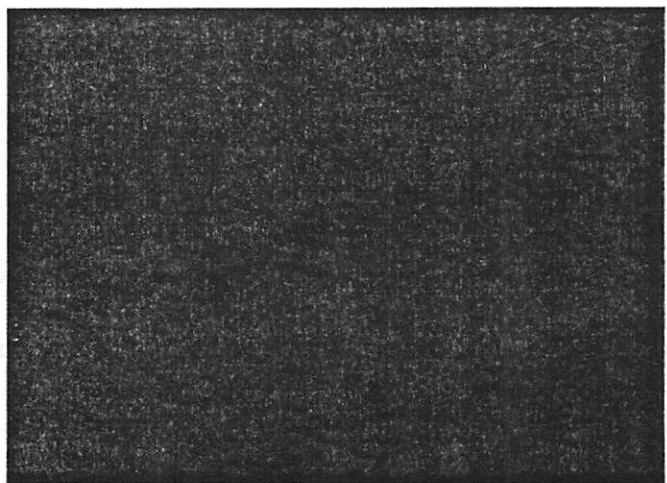
(c)

Figure 7-3: Unpatterned photoresist. (a) optimal SEM image, (b) blurred SEM image, (c) SEM image (b) deblurred using algorithm.

alignment, and astigmation are varied in several directions. The magnitude of the induced defocus, aperture misalignment, and astigmatism is also varied. As the magnitude of the aberration increases, the size of the PSF used to deblur has to be enlarged as well. With unpatterned photoresist like figure 7-3, the original image cannot be sufficiently recovered when the blur is large enough to require PSFs larger than 16 x 16 pixels (35.6 nm x 35.6 nm). Beyond that level, there is insufficient information within the image to allow the deblurring algorithm to function. As mentioned earlier, however, the magnitude of blur seen with images that are properly focused is typically on the order of 15 nm or less.

7.4 Validity of Algorithm-- Intentional, Repeating Pattern on SEM Image

The next hurdle to prove the efficacy of this algorithm with SEM images is to examine how intentional, regularly repeating patterns on the image affect the deblur. The PSF of an SEM image with a trench is compared to that of an image with no pattern. An intentional, regular pattern detectable within the size of the PSF, such as a line edge, may falsely be interpreted as blur. Figure 7-4 is an SEM image of a trench in photoresist, with 4 different PSFs calculated from that image. Figure 7-4 (a) is the PSF calculated from the entire SEM image as shown in the figure. PSF (b) is also calculated from the entire SEM image,
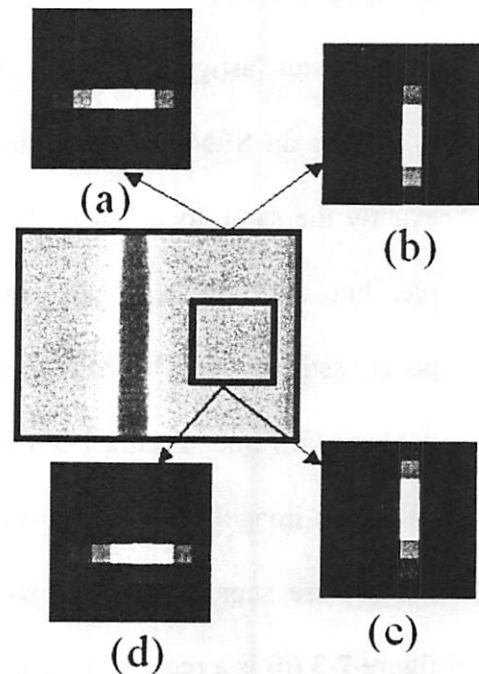
Figure 7-4: (a) and (b) are PSFs calculated using the entire SEM image. (a) is calculated from the image as shown, (b) is calculated from the image rotated clockwise 90°. (c) and (d) are PSFs calculated from the boxed area where no pattern is present. (c) is calculated from the rotated image, (d) is found from the image as shown.

but the image is first rotated 90° clockwise before deblurring. PSF (d) is calculated from just a portion of the SEM image, a portion with no pattern, only photoresist. PSF (c) is calculated from the same portion, rotated 90°. PSFs (b) and (c) are rotated 90° counterclockwise and compared to PSFs (a) and (d). The greyscale level of each of the pixels of PSFs (a) and rotated (b) agree with each other almost exactly, to greater than 99% overall. The same is true of PSF (d) and rotated PSF (c). Comparing PSF (a) with (d) and PSF (b) with (c), these PSFs agree with each other to greater than 95%. All the PSFs are calculated from the same image, simply different portions or rotations of the image. Therefore, all the PSF should agree identically with each other. With the image rotation, the difference between PSFs is negligible. When analyzing different portions of the image, the calculated PSF does change slightly. For cases where the only regular pattern is the line edge, as in figure 7-4, the change in the calculated PSF introduced by the regular pattern is small, but noticeable.

7.5 Validity of Algorithm-- In-Depth PSF Comparison

The PSF of the deblurring algorithm captures the blur within the imaging system, the SEM. PSFs of images that look significantly different, but were taken on the same SEM at the same time, should all have similar final deblurring PSF solution. A test is designed to allow statistical analysis. The PSFs of various SEM images of the same sample at 4 different stage rotations are compared. Figure 7-5 contains four SEM images captured at the four different sample rotations. The sample is first centered on the SEM stage, then it is rotated but not translated. Best focus, stigmation, etc. are found prior to image capture. At each rotation, 3 - 6 images are captured to allow for statistical significance. The images are captured in random order to avoid any time-dependent effects such as a slow gun misalignment over time. After
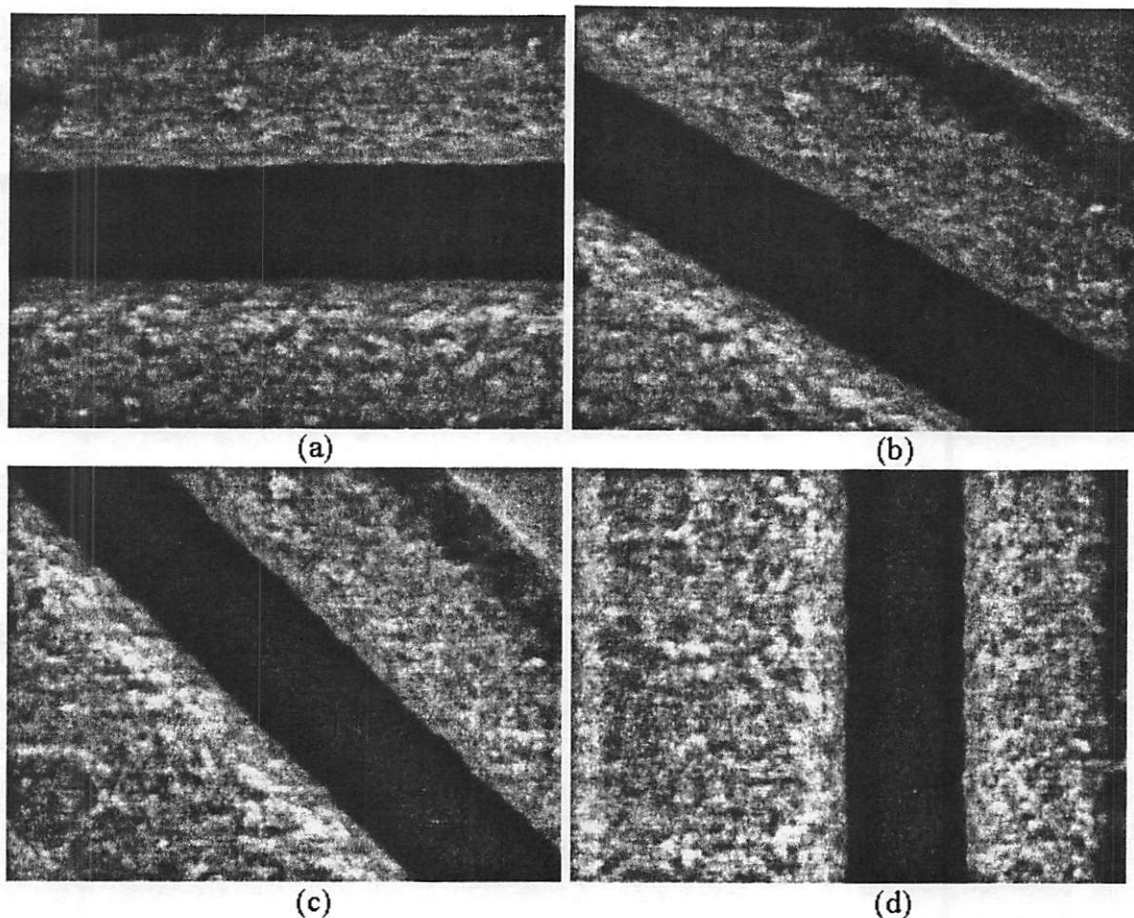
Figure 7-5: (a) - (d) are SEM images of the same trench area, taken with the trench horizontal, at 30°, at 45°, and vertical, respectively. In all cases, the sample is rotated while inside the SEM, then an image is taken. The sample is rotated, not the image.

image collection, the samples that are not horizontal are artificially rotated back to horizontal with Paint Shop Pro (by Jasc Software [68]). The images as collected and the rotated images are then deblurred. Figure 7-6 gives an example of this process. Figure 7-6 (a) and (b) are different images whereas (c) is simply (b) rotated 90°. Using analysis of variance statistics, a confidence level is found to determine how different the PSFs of the artificially rotated images are from those of the original images, and how similar rotated images are to other rotated images or original images to other original images. The similarity is compared quantitatively; the grey level of each pixel is examined numerically. When comparing PSFs, the pixel coor-
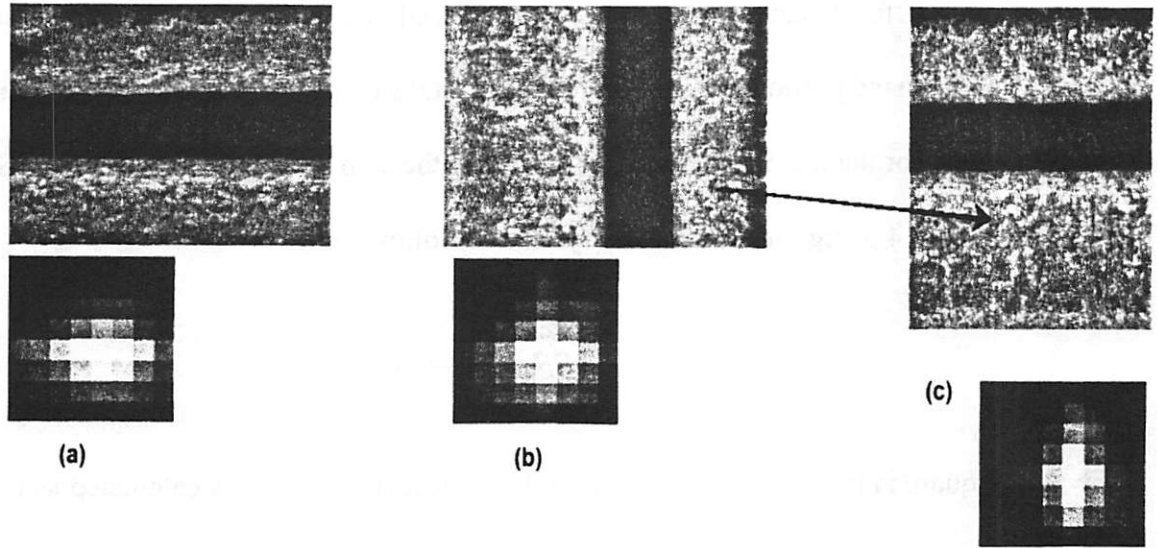
Figure 7-6: (a) and (b) are deblurred images and deblurring PSFs for the trench image taken at 0° and at 90°, respectively. (c) is the deblurred image and PSF taken after rotating the image from (b) back to a horizontal orientation.

dinate of one image is compared to the same pixel coordinate of the other image. Equation 7-4 summarizes the statistical comparison.

Equation 7-4:

$$\sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij}-\bar{y}_{..})^2 = n\sum_{i=1}^{a}(\bar{y}_{i.}-\bar{y}_{..})^2 + \sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij}-\bar{y}_{i.})^2$$

where y_{ij} is each individual measurement, $\bar{y}_{i.}$ is the average measurement within a group, $\bar{y}_{..}$ is the overall average of the measurements, a is the number of different groups, and n is the number of observations at each group. In this case there are 7 different group, or factors: one at 0°, 30°, 45°, and 90°, as well as one at 30°, 45°, and 90° after rotating back to 0°, horizontal. All the PSFs in the groups of original images are expected to be similar to each other, the PSFs of the images rotated back are expected to be similar to each other, and the two are expected to be different from each other. Therefore, these 7 factors can actually be thought of

108

as simply two different factors: the images as captured and the images rotated back to horizontal. The left hand portion of equation 7-4 is the total sum of squares, the next portion is the sum of squares for the factors, and the last portion is the sum of squares for the error. Simplifying equation 7-4 using this terminology yields the following.

Equation 7-5:

$$SS_{Total} = SS_{Factor} + SS_{Error}$$

The mean square is found by normalizing, and the f test statistic, F_0, is calculated as the ratio of the mean square of the factor divided by the mean square of error. F_0, equals 7.89.

Equation 7-6:

$$F_0 = \frac{MS_{Factor}}{MS_{Error}} = 7.89$$

In this case, with 2 factor levels and at least 48 measurements within each factor, the factor has 1 degree of freedom and the error has 94 degrees of freedom. With these degrees of freedom, the probability of falsely rejecting the null hypothesis is 0.6%. There is indeed a significant difference between the PSFs that are rotated back to horizontal and the PSFs from the original images.

Similar analysis of variance testing is conducted upon all 7 levels to confirm that there is no significant difference between the PSFs of all the original images, and to confirm that there is none between the PSFs of all the images rotated back to horizontal. The probability of rejecting the null hypothesis is far greater than 50% in all these cases. Therefore, there is no measurable or significant difference between these PSFs. All the original images that have not been rotated are similar to each other, and all the images that have been rotated back to horizontal are similar to each other.

7.6 Cases When the Algorithm Fails

In what cases, if any, will the maximum likelihood deconvolution algorithm fail? Can one detect a failed deblurred image? The deblurring technique is able to deblur successfully for cases where the image is mostly grey, cases where there is some level of intentionally repeating pattern on the order of the PSF, and cases where there are both. An entirely grey image is meaningless, as mentioned, but what happens when there is a significant amount of repeating patterns on the order of the PSF? Figure 7-7 (a) is an SEM image of a dense line:space pat-
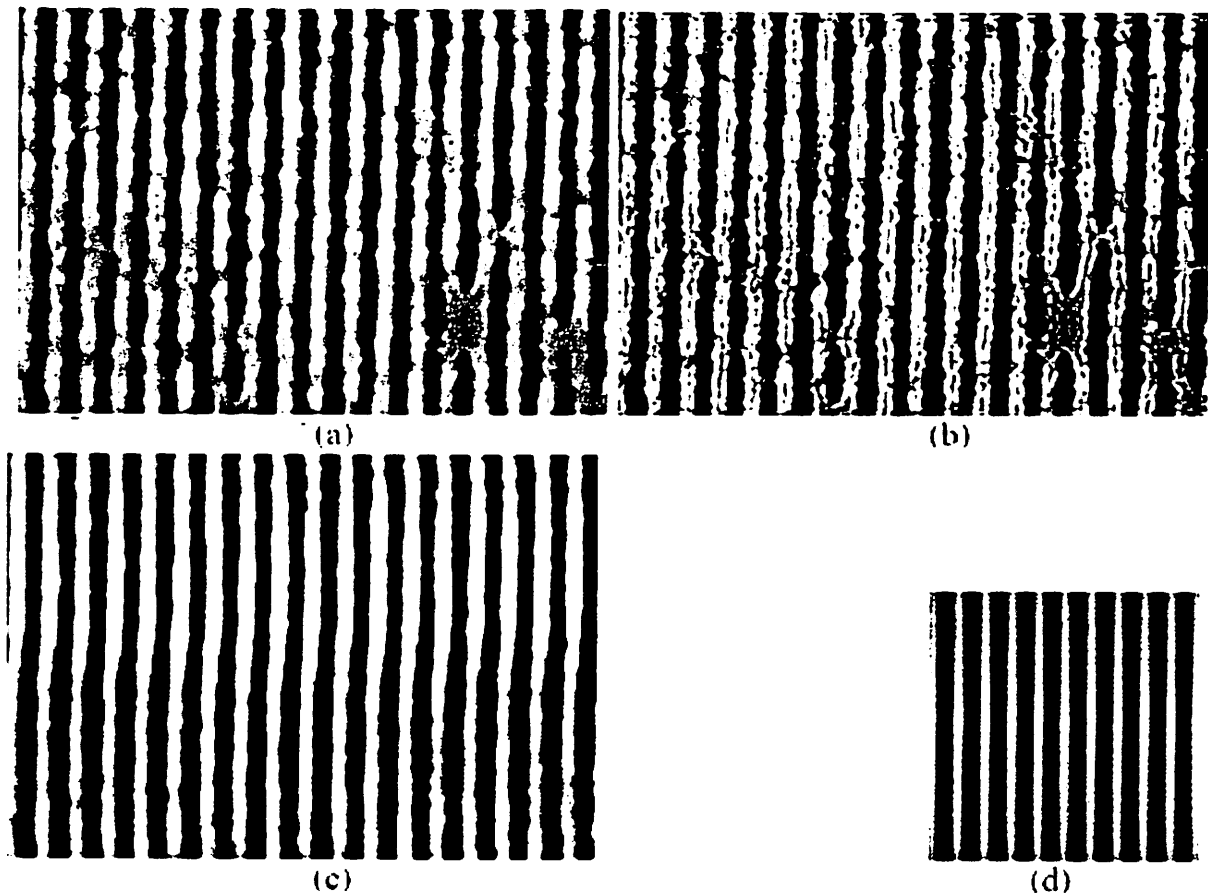
Figure 7-7: SEM image deblurring limitations. (a) is an SEM image without deblurring. (b) is the image deblurred fairly successfully with an 8 x 8 pixel PSF. (c) is the image deblurred with a 350 x 350 pixel PSF. (d) is the solved 350 x 350 pixel PSF. The PSF is shown to scale with the other images.

110

tern. Figure 7-7 (b) is a deblurred image using the typical 8 x 8 pixel PSF. Dimples are seen, but image fidelity is retained. Also, the PSF (not shown) has a roughly gaussian shape. Figure 7-7 (c) is a deblurred image using a 350 x 350 pixel PSF size. At this size, the PSF certainly encompasses several of the lines and spaces. During the iterative calculation of the PSF and deblurred image, the regular line:space pattern is mistakenly interpreted as some sort of image blur and the algorithm attempts to compensate for it. The final deblurred image is greatly distorted and the deblurring PSF, figure 7-7 (d), clearly contains the regular line:space pattern seen in the original image.

There are cases where the image deblurring algorithm cannot be applied. The final PSF after deblurring is completed readily provides information that can resolve whether the deblurring algorithm is mistakenly incorporating deliberate, repeating patterns as blur. Simple analysis of this PSF can determine the cases where the deblurring algorithm cannot be used. These cases are rare, however, since the PSF has to encompass a fairly large degree of repeatability. When the PSF is not large enough to capture both trench edges, the deblurring algorithm functions properly.

7.7 Utilizing the Deconvolution Algorithm for SEM Troubleshooting

The deblurring PSF contains information that can be used to help troubleshoot any problems with the SEM and it can be used to set process control limits upon the machine and recognize when it is operating out of spec. Also, the deblurring technique can provide wider tolerances for machine operation.

The maximum likelihood deblurring algorithm iteratively solves for the blurring point-spread function experienced in the electron optics hardware. Analysis of the PSF can yield

111

some information regarding any significant optical problems the SEM machine is experiencing. This portion of the study provides some initial examples empirically determined through the extensive use of this algorithm for SEM deblurring.
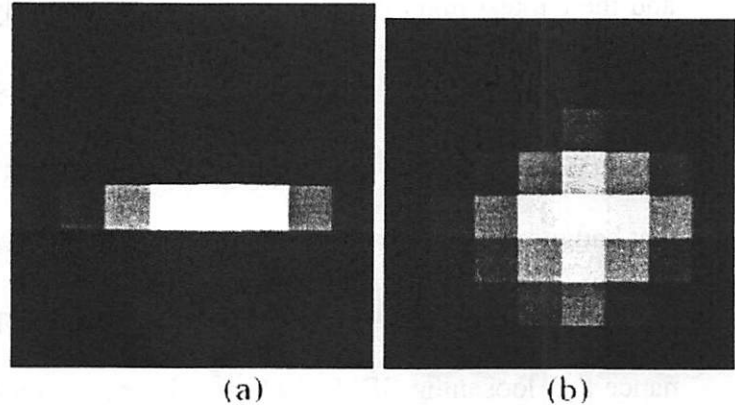
Figure 7-8: 8 x 8 pixel deblurring PSFs. Both PSFs were captured off of the Leo 1550, but during different days. (a) was captured on a day when the stage was drifting only in the x-plane, (b) was captured when it was drifting in both the x- and y-plane.

Figures 7-8 (a) and (b) are deblurring PSFs solved from SEM images that were collected using the same SEM machine, the Leo 1550, on different days. The author noticed, during SEM image capture of the original image, that the image corresponding to the PSF in figure 7-8 (a) was collected on a day when the SEM machine stage was drifting significantly in the x-plane. In figure 7-8 (b) the stage was also drifting, but it was drifting in both the x- and y-plane.

Qualitative examination of the correlation between the PSF and the stage drift continued. The shape of the PSF always matched the direction of stage drift. It was clear that stage drift was dominating the blur experienced. This was true of the other SEM machines studied as well, except the Hitachi. The stage of the Hitachi SEM was fixed after placing the sample and could not move. On days when little or no stage drift was seen, the shape of the deblurring PSF was more symmetric, suggesting that a radially symmetric blurring function, such as defocus, was typically the next highest cause of image blur.

All the images of PSFs shown in this paper have been scaled so that the shape of the PSF can be more easily seen. In most cases, the variation in intensity between the brightest pixel

and the darkest pixel is only a few percent, typically around 1% or 2%. If there were absolutely no blur in the original image, the deblurring PSF would be absolutely grey; all pixels in the solved PSF would have the exactly same value (1 / 64 since there are 64 pixels). Images that had more severe blur had a larger variation between the darkest and brightest pixels.

This information can be used for two critical purposes: improving SEM machine maintenance and loosening SEM machine tolerances. Quantitative analysis of the deblurring PSF can determine the state of the SEM machine blur and flag the machine when it needs servicing. Specific process control (SPC) limits can be established regarding the shape of the PSF. For instance, when the average of the innermost 16 pixels (4 x 4) varies appreciably from the average of the outermost 48 pixels, a flag can be sent that the machine is not in control.

One can loosen SEM tolerances since the deblurring algorithm is able to compensate for a significant level of the blur. For instance, pixel noise can be reduced by integrating the image over a longer period of time. As stated above, stage drift is already typically the largest component of blur detected from the deblurring algorithm. The tolerance for the SEM signal to noise ratio can be loosened somewhat since the deblurring algorithm can allow image integration over longer periods of time. The same could apply to whatever mechanism for image blur the SEM is experiencing, as long as the level of blur seen is not so high that it is impossible to recapture the image. As experienced throughout this study, that level has typically been around 12 - 16 pixels. Allowing looser tolerances on the SEM tool will make tool operation significantly cheaper. The tool will not be down for maintenance as often, and it will not require new parts and servicing as often. Both of these benefits result in cheaper LER data collection, and cheaper SEM operation in general.

113

7.8 Summary

The maximum likelihood deconvolution algorithm described in chapter six was used to deblur SEM images, measure LER, and verify the algorithm. LER values changed after deblurring. The amount of change depended in part upon the smoothness of the original roughness, yet LER values were around 11% to 14% larger on average after deblurring. There was no firm conclusion that the distribution of LER values after deblurring was tighter than before deblurring. However, this could be explained by the large deviation in LER measurements due to the statistical nature of LER.

The algorithm passed several tests to determine its ability to deblur SEM images. It functioned well when attempting to deblur images with little information or with a moderate level of repeating patterns on the order of the deblurring point-spread function. Also, the PSFs of several different artificially rotated deblurred images were compared to those of several unrotated deblurred images. The two were statistically different beyond 99% confidence, while simultaneously showing beyond 99% confidence that the PSFs of all the like images were alike. Some images with highly repeating patterns were deblurred with very large PSFs in order to force the algorithm to fail. It is clear, by examining the failed PSF, that the deblur failed. This is desirable, so that one can ensure that the deblurring algorithm is working properly.

Finally, other useful applications for the deblurring algorithm were listed. Analysis of the deblurring PSF can help troubleshoot an SEM machine. Also, a PSF that is far from its original grey matrix means that the image is experiencing a high degree of blur. Control limits placed upon the variation of the PSF can help control the stability of the SEM. Finally,

114

because the deblurring algorithm can recover some level of blur, original SEM images can tolerate further blur. This allows an SEM to function with looser tolerances, making LER data collection significantly cheaper and easier.

8 Conclusions

The previous chapters highlighted the critical need to further understand and minimize line edge roughness (LER). Improved critical dimension (CD) control is essential for the next generation of semiconductor devices, and LER degrades CD control. Earlier chapters proved that sub-threshold lithographic exposure in the background regions of the pattern have a strong influence upon LER. The last two chapters discussed LER metrology in further detail, focusing upon methods to improve the quality of an SEM image by deblurring the image. This deblurring technique was rigorously tested to ensure that it was properly and successfully enhancing the image quality.

Using the results provided in this work, some opportunities for future studies in LER and SEM image enhancement can be seen. This chapter will outline suggestions of some of the most promising directions such studies should take, as well as summarize the key data presented throughout the previous chapters.

8.1 Summary of LER versus Aerial Image Contrast

The most significant results and most in-depth study provided in this work were in chapter five and, to a lesser extent, chapter four. In these chapters the notion that background flare was greatly responsible, albeit not solely responsible, for aerial image-induced LER was first brought to light. Background flare has never before directly been considered a source of LER. Flare leads to a degraded aerial image profile, therefore degrading those aspects of the aerial image previously considered most responsible for LER such as the image slope. Others considered background flare as a concern only in this fashion, as a degradation of aerial image

116

quality. However this study proved that sub-threshold exposure in unpatterned areas, or background exposure, detrimentally affects LER in its own right. Moreover, the background exposure is probably more responsible for lithographic image-induced LER as those aspects of image contrast previously considered most important, such as image slope or standard contrast.

A model of the background contrast, defined in chapter five and dependent solely upon the exposure dose in areas meant to remain unexposed, versus LER fit better than any other form of contrast tested. The F ratio of the background contrast versus the left side of the trench LER was 163, compared to an F ratio equal to 18 for the standard contrast versus LER and 36 for the slope contrast versus LER. A power equation, with background contrast roughly inversely proportional to LER, fit best:

Equation 5-6:

$$1 \; \sigma \; \text{LER} \; = \; 1.332 \cdot (\text{background contrast})^{-1.1132}$$

Fitting background contrast to the right side of the trench LER, the F ratio was 134. The F ratios for standard contrast and slope contrast were 40 and 56, respectively.

Equation 8-1:

$$1 \; \sigma \; \text{LER} \; = \; 1.357 \cdot (\text{background contrast})^{-0.8045}$$

In all cases, the F ratio is certainly large enough to prove without a doubt that there is a correlation between all form of contrast and LER. It is impossible to fully orthogonalize each form of contrast to the others, so the strong fits to standard contrast and slope contrast might partially be due to their similarity to background contrast. Nonetheless, these strong fits suggested that perhaps a combination of the forms of contrast might fit best. Attempts to find fits

117

with many combinations of these forms of contrast versus LER did not find any fits superior to the fits versus background contrast alone. However, it was difficult to model combinations of resist contrast forms because the equations could not be properly linearized. Statistical F numbers can only be calculated from linear equations, therefore these non-linear equations had to be somewhat falsely linearized in order to allow full statistical analysis.

8.2 Summary of SEM Deblurring Algorithm

The maximum likelihood image deblurring algorithm, based on the Richardson-Lucy algorithm and chiefly used in the capture of astronomical images, was used to remove blur from images captured from SEMs and recover the high frequency components of those images. LER values were then calculated from the deblurred images. LER measurements after deblurring were around 11% - 15% higher than they were before deblurring. However one measurement was as large as 40.2% higher than before deblurring, while another was as low as 34.5%. This variation was greatly dependent upon the condition of the original image. If the original image already contained little blur, or if the LER of the original image was quite large, then the percent change before and after deblurring was less. Regardless, the change in LER measurement before and after deblurring was notable.

SEM images of photoresist lines and trenches contain little information, compared to a typical photograph or other image. They also often contain many repeating patterns. The maximum likelihood algorithm requires non-repeating information in the ideal image in order to be able to deblur properly. Therefore, it was possible that the algorithm may not deblur SEM images properly or sufficiently. Many tests were performed to ensure that enough information is present in SEM images, and that this information is not overly repeating. With

regards to the former, an image without any pattern was deblurred successfully. With regards to the latter, as long as reasonable limits were chosen for the PSF size-- it must be smaller than the smallest intentionally repeating pattern in the image-- the image was deblurred properly.

Finally, the accuracy of the SEM deblurring technique was tested. Are SEM measurements taken from deblurred images more accurate than SEM measurements from the original images? This was tested by collecting SEM measurements along a long trench of a sample using multiple SEM machines. Statistical testing of the measurements before and after deblurring showed little difference. It was impossible to conclude that the SEM measurements after deblurring were more approaching a more consistent or accurate "true" value. The suspected cause for this unexpected result is the high level of LER measurement variability. Taking several measurements of the trench at several different locations yields significantly different LER measurements. LER is a statistical phenomenon. The variability of LER measurements is larger than additional variability incurred by blurred images. In short, it is still possible, if not likely, that the deblurring technique still improves SEM image quality and provides more accurate LER measurements. However, the improved accuracy is slight compared to the inherent variability of LER measurements.

Some additional, unexpected benefits of the deblurring algorithm were noted; they are related to maintaining proper SEM operating conditions. First, blur experienced throughout the study was mostly due to SEM stage drift since some form of time averaging was needed during image capture. Therefore, the deblurring algorithm could be used in tandem with time averaging to lower image pixel noise while maintaining a relatively blur-free image. This would allow looser SEM tolerances without sacrificing image quality. Next, the symmetry of blur could be monitored, and one could easily determine whether blur is mostly in the x- or y-

direction, or whether it was fairly symmetric. The degree of blur could also be monitored, since the extent of blur in the original image was easily determined upon examination of the deblurring PSF. This information could be collected and monitored to keep track of how far the SEM machine is from optimal operating conditions. This is discussed further in section 8.3.

8.3 Opportunities for Future Research

With the newfound knowledge that background flare critically affects LER and EUV lithography struggling to control flare [53, 59], further research studying background contrast versus LER would provide obvious benefit to the field. This study proved conclusively that undesirable, sub-threshold exposure in unpatterned areas affects LER. Yet, it was unable to determine the physical theory behind why this background exposure affects LER so significantly, nor was it able to determine how other forms of contrast play a combined role with the background contrast. It is intuitively obvious that background contrast alone cannot determine LER, so it must be a combination of background contrast and some form of defining foreground contrast. Fuller understanding of this interdependence would highlight which aspects of lithographic tool design are most critical to improve CD control.

In order to determine how the different forms of contrast interact and better understand the physical processes that drive background contrast to be so important, one must attempt to orthogonalize the different forms of contrast-- and the different exposure and post-exposure mechanisms-- as best as possible. One possible method, which is unfortunately quite time- and material-consuming, is to hold one form of contrast as still as possible while adjusting the others. For instance, the clear dose image slope can be held constant while the background

120

exposure dose is altered, or vice-versa. Another approach that is more risky but may provide more insight is to hold one suspected physical phenomenon constant while altering others. For instance, if background contrast is important because it affects the acid distribution profile, and therefore acid diffusion pathways, it would be best to alter the background exposure dose (altering background contrast) while maintaining an acid distribution gradient that is roughly constant. This is risky because there may be several competing physical mechanisms to explain the same experimental results.

Platforms for future research with the deblurring algorithm are more numerous and easily applicable, however they may not be as useful to the industry overall. Two seen as most useful are presented here. The most useful research, perhaps, would involve experiments designed to use the deblurring PSF as an SEM tool process control. After deblurring, the further the PSF is from its original grey plane, the more the deblurring algorithm is compensating for blur within the machine. Specific process control (SPC) limits could be determined for SEMs where an upper limit for the deviations of the PSF from a grey plane could be set. Then one could track what elements of the machine typically account for the blur. For example, when the PSF deviation exceeds the SPC limits, the machine could be fully inspected to determine what caused the high level of blur. Eventually, analysis of the deblurring PSF could provide a simple and non-invasive means of SEM tool control.

Another useful study could be designed to determine how much larger the SEM operating window is when using the deblurring algorithm. Time averaging removes pixel noise but increases image blur. Therefore, when using the deblurring algorithm, SEM signal-to-noise ratio (SNR) does not need to be as high as it normally would. Time averaging could continue

for longer times, thus minimizing noise and increasing blur. The experiment could, for instance, purposely decrease the SNR of SEM tools and measure the image pixel noise after deblurring. This pixel noise could be compared to that of an SEM tool with a good SNR and no deblurring.

Appendix A

RAW DATA CORRELATING AERIAL IMAGE PROFILE TO L.E.R.

A.1 Raw Data used for Multiple Resist Study presented in Section 2.3.

UVII-HS Data:

| | | | | | overall average LER: | |
| | | | | | 10.50434 | |
norm. clear dose slope	norm. peak intensity	(max -min)/ (max+min) contrast	clear dose slope (mJ/cm^2)/um	peak intensity (mJ/cm^2)	LER 3 sigma	LER 1 sigma
1.78665863	1.108952213	0.18079014	14.4719349	8.982512921	22.99608	7.66536
2.152600759	1.179441702	0.20983996	17.43606615	9.553477785	11.92392	3.97464
2.900419854	1.177201563	0.30572874	23.49340081	9.535332656	10.75464	3.58488
3.494481751	1.290287726	0.346578	28.30530219	10.45133058	12.22533	4.07511
4.764423013	1.6770523	0.36983176	38.59182641	13.58412363	10.47027	3.49009
7.734452943	1.4723375	0.99690099	62.64906884	11.92593375	6.89367	2.29789
7.734452943	2.0974125	0.53951388	62.64906884	16.98904125	10.63407	3.54469
9.318618004	1.773900602	0.99690099	75.48080583	14.36859488	4.71102	1.57034
20.62520785	3.926233333	0.99690099	167.0641836	31.80249	3.98253	1.32751
1.78665863	1.108952213	0.18079014	14.4719349	8.982512921	24.76071	8.25357
2.152600759	1.179441702	0.20983996	17.43606615	9.553477785	16.20474	5.40158
2.900419854	1.177201563	0.30572874	23.49340081	9.535332656	12.96045	4.32015
3.494481751	1.290287726	0.346578	28.30530219	10.45133058	11.97138	3.99046
4.764423013	1.6770523	0.36983176	38.59182641	13.58412363	14.96586	4.98862
7.734452943	1.4723375	0.99690099	62.64906884	11.92593375	4.28838	1.42946
7.734452943	2.0974125	0.53951388	62.64906884	16.98904125	8.36088	2.78696
9.318618004	1.773900602	0.99690099	75.48080583	14.36859488	3.72039	1.24013
20.62520785	3.926233333	0.99690099	167.0641836	31.80249	4.38399	1.46133
1.78665863	1.108952213	0.18079014	14.4719349	8.982512921	23.59041	7.86347
2.152600759	1.179441702	0.20983996	17.43606615	9.553477785	13.14258	4.38086
2.900419854	1.177201563	0.30572874	23.49340081	9.535332656	6.35556	2.11852
3.494481751	1.290287726	0.346578	28.30530219	10.45133058	10.88958	3.62986
4.764423013	1.6770523	0.36983176	38.59182641	13.58412363	15.54162	5.18054
7.734452943	1.4723375	0.99690099	62.64906884	11.92593375	5.68356	1.89452
7.734452943	2.0974125	0.53951388	62.64906884	16.98904125	8.20905	2.73635
9.318618004	1.773900602	0.99690099	75.48080583	14.36859488	4.96395	1.65465
20.62520785	3.926233333	0.99690099	167.0641836	31.80249	4.14894	1.38298
1.78665863	1.108952213	0.18079014	14.4719349	8.982512921	18.13116	6.04372
2.152600759	1.179441702	0.20983996	17.43606615	9.553477785	15.66993	5.22331
2.900419854	1.177201563	0.30572874	23.49340081	9.535332656	8.79849	2.93283
3.494481751	1.290287726	0.346578	28.30530219	10.45133058	9.43734	3.14578
4.764423013	1.6770523	0.36983176	38.59182641	13.58412363	15.86307	5.28769
7.734452943	1.4723375	0.99690099	62.64906884	11.92593375	4.70964	1.56988
7.734452943	2.0974125	0.53951388	62.64906884	16.98904125	9.63204	3.21068
9.318618004	1.773900602	0.99690099	75.48080583	14.36859488	3.58503	1.19501
20.62520785	3.926233333	0.99690099	167.0641836	31.80249	3.59598	1.19866

UV210 Data:

norm. clear dose slope	norm. peak intensity	(max -min)/ (max+min) contrast	clear dose slope (mJ/cm^2)/um	peak intensity (mJ/cm^2)	LER 3 sigma	LER 1 sigma
				overall average LER: 6.746983		
1.78665863	1.108952213	0.180790136	28.58653808	17.7432354	9.3378	3.1126
2.152600759	1.179441702	0.209839964	34.44161214	18.87106723	**7.01271**	2.33757
2.900419854	1.177201563	0.305728742	46.40671766	18.835225	10.03191	3.34397
3.494481751	1.290287726	0.346578001	55.91170802	20.64460361	4.83528	1.61176
4.764423013	1.6770523	0.36983176	76.23076821	26.8328368	7.52559	2.50853
7.734452943	1.4723375	0.996900986	123.7512471	23.5574	4.17483	1.39161
7.734452943	2.0974125	0.539513885	123.7512471	33.5586	5.71947	1.90649
9.318618004	1.773900602	0.996900986	149.0978881	28.38240964	6.47214	2.15738
20.62520785	3.926233333	0.996900986	330.0033256	62.81973333	6.23586	2.07862
1.78665863	1.108952213	0.180790136	28.58653808	17.7432354	6.80625	2.26875
2.152600759	1.179441702	0.209839964	34.44161214	18.87106723	**12.78555**	4.26185
2.900419854	1.177201563	0.305728742	46.40671766	18.835225	11.9079	3.9693
3.494481751	1.290287726	0.346578001	55.91170802	20.64460361	3.85917	1.28639
4.764423013	1.6770523	0.36983176	76.23076821	26.8328368	13.1661	4.3887
7.734452943	1.4723375	0.996900986	123.7512471	23.5574	4.86225	1.62075
7.734452943	2.0974125	0.539513885	123.7512471	33.5586	5.9934	1.9978
9.318618004	1.773900602	0.996900986	149.0978881	28.38240964	3.87171	1.29057
20.62520785	3.926233333	0.996900986	330.0033256	62.81973333	4.60545	1.53515
1.78665863	1.108952213	0.180790136	28.58653808	17.7432354	7.73361	2.57787
2.152600759	1.179441702	0.209839964	34.44161214	18.87106723	**5.31249**	1.77083
2.900419854	1.177201563	0.305728742	46.40671766	18.835225	10.62489	3.54163
3.494481751	1.290287726	0.346578001	55.91170802	20.64460361	4.92606	1.64202
4.764423013	1.6770523	0.36983176	76.23076821	26.8328368	7.80795	2.60265
7.734452943	1.4723375	0.996900986	123.7512471	23.5574	4.4379	1.4793
7.734452943	2.0974125	0.539513885	123.7512471	33.5586	6.62625	2.20875
9.318618004	1.773900602	0.996900986	149.0978881	28.38240964	5.01216	1.67072
20.62520785	3.926233333	0.996900986	330.0033256	62.81973333	5.71647	1.90549
1.78665863	1.108952213	0.180790136	28.58653808	17.7432354	6.28686	2.09562
2.152600759	1.179441702	0.209839964	34.44161214	18.87106723	**6.54039**	2.18013
2.900419854	1.177201563	0.305728742	46.40671766	18.835225	8.10948	2.70316
3.494481751	1.290287726	0.346578001	55.91170802	20.64460361	4.57707	1.52569
4.764423013	1.6770523	0.36983176	76.23076821	26.8328368	9.96849	3.32283
7.734452943	1.4723375	0.996900986	123.7512471	23.5574	5.39217	1.79739
7.734452943	2.0974125	0.539513885	123.7512471	33.5586	5.08659	1.69553
9.318618004	1.773900602	0.996900986	149.0978881	28.38240964	5.09322	1.69774
20.62520785	3.926233333	0.996900986	330.0033256	62.81973333	4.43595	1.47865

124

SEPR-463 Data:

norm. clear dose slope	norm. peak intensity	(max -min)/ (max+min) contrast	clear dose slope (mJ/cm^2)/um	peak intensity (mJ/cm^2)	LER 3 sigma	LER 1 sigma
1.78665863	1.108952213	0.180790136	39.66382158	24.61873912		
2.152600759	1.179441702	0.209839964	47.78773685	26.18360578	11.6098	3.86992
2.900419854	1.177201563	0.305728742	64.38932075	26.13387469	13.4959	4.49864
3.494481751	1.290287726	0.346578001	77.57749488	28.64438752	8.61345	2.87115
4.764423013	1.6770523	0.36983176	105.7701909	37.23056106	22.1319	7.3773
7.734452943	1.4723375	0.996900986	171.7048553	32.6858925	4.93533	1.64511
7.734452943	2.0974125	0.539513885	171.7048553	46.5625575	3.66978	1.22326
9.318618004	1.773900602	0.996900986	206.8733197	39.38059337	4.0248	1.3416
20.62520785	3.926233333	0.996900986	457.8796142	87.16238	3.74061	1.24687
1.78665863	1.108952213	0.180790136	39.66382158	24.61873912		
2.152600759	1.179441702	0.209839964	47.78773685	26.18360578	5.90298	1.96766
2.900419854	1.177201563	0.305728742	64.38932075	26.13387469	16.4697	5.48991
3.494481751	1.290287726	0.346578001	77.57749488	28.64438752	9.67572	3.22524
4.764423013	1.6770523	0.36983176	105.7701909	37.23056106	19.6059	6.5353
7.734452943	1.4723375	0.996900986	171.7048553	32.6858925	4.89084	1.63028
7.734452943	2.0974125	0.539513885	171.7048553	46.5625575	4.46079	1.48693
9.318618004	1.773900602	0.996900986	206.8733197	39.38059337	3.39405	1.13135
20.62520785	3.926233333	0.996900986	457.8796142	87.16238	4.17162	1.39054
1.78665863	1.108952213	0.180790136	39.66382158	24.61873912	24.7607	8.25357
2.152600759	1.179441702	0.209839964	47.78773685	26.18360578	13.8609	4.6203
2.900419854	1.177201563	0.305728742	64.38932075	26.13387469	12.9605	4.32015
3.494481751	1.290287726	0.346578001	77.57749488	28.64438752		
4.764423013	1.6770523	0.36983176	105.7701909	37.23056106	10.239	3.41301
7.734452943	1.4723375	0.996900986	171.7048553	32.6858925	4.28838	1.42946
7.734452943	2.0974125	0.539513885	171.7048553	46.5625575	4.58259	1.52753
9.318618004	1.773900602	0.996900986	206.8733197	39.38059337	3.07782	1.02594
20.62520785	3.926233333	0.996900986	457.8796142	87.16238	3.29637	1.09879
1.78665863	1.108952213	0.180790136	39.66382158	24.61873912	24.7607	8.25357
2.152600759	1.179441702	0.209839964	47.78773685	26.18360578	13.8609	4.6203
2.900419854	1.177201563	0.305728742	64.38932075	26.13387469	12.9605	4.32015
3.494481751	1.290287726	0.346578001	77.57749488	28.64438752		
4.764423013	1.6770523	0.36983176	105.7701909	37.23056106	10.239	3.41301
7.734452943	1.4723375	0.996900986	171.7048553	32.6858925	4.28838	1.42946
7.734452943	2.0974125	0.539513885	171.7048553	46.5625575	4.58259	1.52753
9.318618004	1.773900602	0.996900986	206.8733197	39.38059337	3.07782	1.02594
20.62520785	3.926233333	0.996900986	457.8796142	87.16238	3.29637	1.09879

A.2 Raw Data used for In-Depth UVII-HS Study presented in Section 2.4.

All LER values are given as 1 σ RMS values.

std cont:	b.g. contrast:	norm. slope	slope cont.
0.994696188	0.990454831	1	14.3031952

measurement	leftLER	rightLER
1	1.10753	1.85214
2	2.08428	1.59396
3	1.74596	1.55236
4	1.88466	2.12244
5	1.44999	2.17732
6	1.7004	1.14589
7	1.34185	1.36812
8	1.82687	1.66565
9	1.62215	1.10288
average:	1.64041	1.62008444
std dev.:	0.299140006	0.38404338

std cont:	b.g. contrast:	norm. slope	slope cont.
0.884666959	0.894403109	0.52721764	8.12085049

measurement	leftLER	rightLER
1	1.68418	1.82626
2	2.10017	2.09833
3	1.15523	1.22392
4	1.11675	1.48271
5	1.77154	2.168
6	1.84769	1.42649
7	1.94198	1.91853
8	1.85321	1.28201
9	1.64675	1.16177
10	1.77138	1.64202
11	1.55424	1.46067
12	1.78951	1.53494
13	1.48716	1.41468
14	1.73896	1.78616
15		1.26654
average:	1.675625	1.57953533
std dev.:	0.274841019	0.31638558

std cont:	b.g. contrast:	norm. slope	slope cont.
0.909499879	0.785975331	0.93884374	13.5034846

measurement	leftLER	rightLER
1	2.57191	0.98952
2	1.15469	1.39024
3	2.40476	1.16501
4	1.47892	1.12073
5	1.99647	1.01201
6	1.34645	1.13732
7	0.91157	1.26308
8	1.31227	1.33507
9	1.12149	1.80221
10	0.90008	1.1653
11		1.06468
12	1.46345	1.44531
13	1.84885	1.56676
14	1.44085	1.67202
15	1.22948	1.7297
average:	1.512945714	1.32393067
std dev.:	0.51435973	0.26717548

std cont:	b.g. contrast:	norm. slope	slope cont.
0.699077204	0.692355267	0.44415226	7.03464497

measurement	leftLER	rightLER
1	3.03233	2.02369
2	2.85195	2.48539
3	3.97199	2.6418
4	2.74979	1.96897
5	2.68999	2.3363
6	2.57696	2.54521
7	2.79457	2.65197
8	3.52472	1.77683
9	2.74283	2.4663
10	3.27875	2.74787
11	2.84343	3.51422
12	2.25859	2.24526
13	2.24343	3.03046
14	2.40129	1.87299
15	2.22603	2.26322
average:	2.812443333	2.438032
std dev.:	0.488671304	0.45772728

std cont: 0.826060634	b.g. contrast: 0.58436301	norm. slope 0.754190999	slope cont. 11.08887056
measurement	leftLER	rightLER	
1	1.95118	1.84332	
2	2.37826	1.60374	
3	2.09706	1.56015	
4	4.21037	1.7127	
5	2.1388	1.42063	
6	2.14115	2.02259	
7	1.2629	1.93064	
8	1.85172	2.07567	
9	1.98266	1.42672	
10	2.27962	1.77196	
11	2.42697	1.80464	
12	1.44423	2.35257	
13	1.52865	1.79118	
14	1.61952	1.70777	
15	1.75438	2.26432	
average:	2.071164667	1.81924	
std dev.:	0.684840071	0.274665574	
std cont: 0.551269762	b.g. contrast: 0.490271132	norm. slope 0.348913344	slope cont. 5.789251988
measurement	leftLER	rightLER	
1	1.65926	1.8903	
2	2.03699	2.1684	
3	2.56098	1.73696	
4	2.9484	2.31119	
5	2.44966	1.78623	
6	2.87311	2.16344	
7	1.88888	2.95603	
8	1.71148	1.97471	
9	1.76138	1.85624	
10	2.00021	2.55682	
11	1.97789	2.09868	
12	2.27271	1.95823	
13	2.02189	2.23449	
14	2.53059	2.77011	
15	1.95999	2.49459	
average:	2.176894667	2.197094667	
std dev.:	0.408636127	0.36308624	

std cont:	b.g. contrast:	norm. slope	slope cont.
0.519237852	0.376079357	0.317929587	5.384092482
measurement	leftLER	rightLER	
1	3.54469	2.7827	
2	2.71441	3.16607	
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			
15			
average:	3.12955	2.974385	
std dev.:	0.587096618	0.271083527	

std cont:	b.g. contrast:	norm. slope	slope cont.
0.717211316	0.294168755	0.43621249	6.930820414
measurement	leftLER	rightLER	
1	4.59164		
2	4.39309		
3	9.55582	4.51653	
4	10.07664	4.72795	
5	6.52077	3.94778	
6	7.56734		
7	5.65531	4.49964	
8	5.78282	4.02245	
9	8.05214	3.42541	
10	5.90971	4.49881	
11	8.34105	3.32087	
12	4.21954		
13	10.35647	4.70929	
14	6.19517	4.21356	
15	6.18602	3.39398	
average:	6.893568667	4.116024545	
std dev.:	2.011871137	0.534569913	

Appendix B

MATLAB SCRIPTS AND CODE USED THROUGHOUT THE STUDY

B.1 Matlab script for the Blind Deconvolution Maximum Likelihood Algorithm.

This script is packaged with the Matlab software, the imaging package. This was not created by the author.

```
------------------------------------------------------------------------

function [J,P] = deconvblind(varargin),
%DECONVBLIND Image restoration using blind deconvolution algorithm.
%   [J,PSF] = DECONVBLIND(I,INITPSF) deconvolves image I using maximum
%   likelihood algorithm, returning both deblurred image J and a restored
%   point-spread function PSF. The resulting PSF is a positive array of
%   the same size as the INITPSF, normalized so its sum adds to 1. The
%   PSF restoration is affected strongly by the size of its initial
%   guess, INITPSF, and less by its values (an array of ones is a safer
%   guess).
%
%   To improve the restoration, additional parameters can be passed in
%   (use [] as a place holder if an intermediate parameter is unknown):
%   [J,PSF] = DECONVBLIND(I,INITPSF,NUMIT)
%   [J,PSF] = DECONVBLIND(I,INITPSF,NUMIT,DAMPAR)
%   [J,PSF] = DECONVBLIND(I,INITPSF,NUMIT,DAMPAR,WEIGHT)
%   [J,PSF] = DECONVBLIND(I,INITPSF,NUMIT,DAMPAR,WEIGHT,READOUT).
%
%   Additional constraints on PSF can be provided via a user supplied
%   function:
%   [J,PSF] = DECONVBLIND(...,FUN,P1,P2,...,PN)
%
%   FUN (optional) is a function describing additional constraints on the
%   PSF. There are four ways to specify FUN: as a function-handle, @, as
%   an inline object, or as a string containing either a function name or
%   a MATLAB expression. FUN is called at the end of each iteration. FUN
%   must accept the PSF as its first argument and may accept additional
%   parameters, P1, P2, ..., PN. FUN should return one argument, PSF,
%   that is the same size as the INITPSF, and satisfies the positivity
```

130

```
%   and normalization constraints.
%
%   NUMIT   (optional) is the number of iterations (default is 10).
%
%   DAMPAR  (optional) is an array that specifies the threshold deviation
%   of the resulting image from the image I (in terms of the standard
%   deviation of Poisson noise) below which the damping occurs. The
%   iterations are suppressed for the pixels that deviate within the
%   DAMPAR value from their original value. This suppresses the noise
%   generation in such pixels, preserving necessary image details
%   elsewhere. Default is 0 (no damping).
%
%   WEIGHT  (optional) is assigned to each pixel to reflect its recording
%   quality in the camera. A bad pixel is excluded from the solution by
%   assigning it zero weight value. Instead of giving a weight of one for
%   good pixels, you can adjust their weight according to the amount of
%   flat-field correction. Default is a unit array of the same size as
%   input image I.
%
%   READOUT (optional) is an array (or a value) corresponding to the
%   additive noise (e.g., background, foreground noise) and the variance
%   of the read-out camera noise. READOUT has to be in the units of the
%   image. Default is 0.
%
%
%   Note that the output image J could exhibit ringing introduced by the
%   discrete Fourier transform used in the algorithm. To reduce the
%   ringing use I = EDGETAPER(I,PSF) prior to calling DECONVBLIND.
%
%   Note also that DECONVBLIND allows you to resume deconvolution
%   starting from the results of an earlier DECONVBLIND run. To initiate
%   this syntax, the input I and INITPSF have to be passed in as cell
%   arrays, {I} and {INITPSF}. Then the output J and PSF become cell
%   arrays and can be passed as the input arrays into the next
%   DECONVBLIND call. The input cell array can contain one numeric array
%   (on initial call), or four numeric arrays (when it is the output from
%   a previous run of DECONVBLIND). The output J contains four elements,
%   where J{1}=I, J{2} is the image resulted from the last iteration,
%   J{3} is the image from one before last iteration, J{4} is an array
%   used internally by the iterative algorithm.
%
%   Class Support
%   -------------
%   I and INITPSF can be of class uint8, uint16, or double. DAMPAR and
%   READOUT have to be of the same class as the input image. Other inputs
%   have to be of class double. Output image J (or the first array of the
```

131

```
%   output cell) is of the same class as the input image I. Output PSF is
%   of class double.
%
%   Example
%   -------
%
%      I = checkerboard(8);
%      PSF = fspecial('gaussian',7,10);
%      V = .0001;
%      BlurredNoisy = imnoise(imfilter(I,PSF),'gaussian',0,V);
%      WT = zeros(size(I));WT(5:end-4,5:end-4) = 1;
%      INITPSF = ones(size(PSF));
%      FUN = inline('PSF + P1','PSF','P1');
%      [J P] = deconvblind(BlurredNoisy,INITPSF,20,10*sqrt(V),WT,FUN,0);
%      subplot(221);imshow(BlurredNoisy);
%            title('A = Blurred and Noisy');
%      subplot(222);imshow(PSF,[]);
%            title('True PSF');
%      subplot(223);imshow(J);
%            title('Deblured Image');
%      subplot(224);imshow(P,[]);
%            title('Recovered PSF');
%
%   See also DECONVWNR, DECONVREG, DECONVLUCY, EDGETAPER, PADARRAY,
%   PSF2OTF, OTF2PSF.

%   Copyright 1993-2002 The MathWorks, Inc.
%   $Revision: 1.4 $
%


%   References
%   ----------
%   "Acceleration of iterative image restoration algorithms, by D.S.C. Biggs
%   and M. Andrews, Applied Optics, Vol. 36, No. 8, 1997.
%   "Deconvolutions of Hubble Space Telescope Images and Spectra",
%   R.J. Hanisch, R.L. White, and R.L. Gilliland. in "Deconvolution of Images
%   and Spectra", Ed. P.A. Jansson, 2nd ed., Academic Press, CA, 1997.
%   "Light Microscopic Images Reconstructed by Maximum Likelihood
%   Deconvolution", Timothy J. Holmes et al. in "Handbook of
%   Biological Confocal Microscopy", Ed. James B. Pawley, Plenum
%   Press, New York, 1995

% Parse inputs to verify valid function calling syntaxes and arguments
[J,P,NUMIT,DAMPAR,READOUT,WEIGHT,sizeI,classI,sizePSF,FunFcn,FunArg] = ...
    parse_inputs(varargin{:});
```

```
% 1. Prepare parameters for iterations
%
% Create indexes for image according to the sampling rate
idx = repmat({':'},[1 length(sizeI)]);

wI = max(WEIGHT.*(READOUT + J{1}),0);% at this point  - positivity constraint
fw = fftn(WEIGHT);
DAMPAR22 = (DAMPAR.^2)/2;


% 2. L_R Iterations
%


lambda = 2*any(J{4}(:)~=0);
for k = lambda + [1:NUMIT],

  % 2.a Make an image and PSF predictions for the next iteration
  if k > 2,% image
    lambda = (J{4}(:,1).'*J{4}(:,2))/(J{4}(:,2).'*J{4}(:,2) + eps);
    lambda = max(min(lambda,1),0);% stability enforcement
  end
  Y = max(J{2} + lambda*(J{2} - J{3}),0);% image positivity constraint

  if k > 2,% PSF
    lambda = (P{4}(:,1).'*P{4}(:,2))/(P{4}(:,2).'*P{4}(:,2) + eps);
    lambda = max(min(lambda,1),0);% stability enforcement
  end
  B = max(P{2} + lambda*(P{2} - P{3}),0);% PSF positivity constraint
  sumPSF = sum(B(:));
  B = B/(sum(B(:)) + (sumPSF==0)*eps);% normalization is a necessary constraint,
  % because given only input image, the algorithm cannot even know how much
  % power is in the image vs PSF. Therefore, we force PSF to satisfy this
  % type of normalization: sum to one.

  % 2.b  Make core for the LR estimation
  CC = corelucy(Y,psf2otf(B,sizeI),DAMPAR22,wI,READOUT,1,idx,[],[]);

  % 2.c Determine next iteration image & apply positivity constraint
  J{3} = J{2};H = psf2otf(P{2},sizeI);
  scale = real(ifftn(conj(H).*fw)) + sqrt(eps);
  J{2} = max(Y.*real(ifftn(conj(H).*CC))./scale,0);
  J{4} = [J{2}(:)-Y(:) J{4}(:,1)];

  % 2.d Determine next iteration PSF & apply positivity constraint + normalization
  P{3} = P{2};H = fftn(J{3});
  scale = otf2psf(conj(H).*fw,sizePSF) + sqrt(eps);
  P{2} = max(B.*otf2psf(conj(H).*CC,sizePSF)./scale,0);
```

```
sumPSF = sum(P{2}(:));
P{2} = P{2}/(sumPSF + (sumPSF==0)*eps);

if ~isempty(FunFcn),
  FunArg{1} = P{2};
  P{2} = feval(FunFcn,FunArg{:});
end;
 P{4} = [P{2}(:)-B(:) P{4}(:,1)];
end

% 3. Convert the right array (for cell it is first array, for notcell it is
% second array) to the original image class & output the whole thing
num = 1 + strcmp(classI{1},'notcell');
if ~strcmp(classI{2},'double'),
  J{num} = changeclass(classI{2},J{num});
end

if num == 2,% the input & output is NOT a cell
  P = P{2};
  J = J{2};
end;


%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%
% Function: parse_inputs
function [J,P,NUMIT,DAMPAR,READOUT,WEIGHT,sizeI,classI,sizePSF,FunFcn,FunArg]
= ...
    parse_inputs(varargin),
%
% Outputs:
% I=J{1}   the input array (could be any numeric class, 2D, 3D)
% P=P{1}   the operator that distorts the ideal image
%
% Defaults:
%
NUMIT = [];NUMIT_d = 10;% Number of  iterations, usually produces good
                % result by 10.
DAMPAR =[];DAMPAR_d = 0;% No damping is default
WEIGHT =[];WEIGHT_d = 1;% All pixels are of equal quality, flat-field is one
READOUT=[];READOUT_d= 0;% Zero readout noise or any other
        % back/fore/ground noise associated with CCD camera.
        % Or the Image is corrected already for this noise by user.
FunFcn = '';FunFcn_d = '';
FunArg = {};FunArg_d = {};
funnum = [];funnum_d = nargin+1;
```

134

```
checknargin(2,inf,nargin,mfilename);% no constraint on max number because of FUN args


% First, assign the inputs starting with the cell/not cell image & PSF
%
switch iscell(varargin{1}) + iscell(varargin{2}),
case 0, % no-cell array is used to do a single set of iterations
  classI{1} = 'notcell';
  J{1} = varargin{1};% create a cell array in order to do the iterations
  P{1} = varargin{2};
case 1,
  msg = sprintf('In function %s, I and INITPSF must either both be cell arrays or both be
numeric arrays.',...
              mfilename);
  eid = sprintf('Images:%s:IandInitpsfMustBeOfSameType',mfilename);
  error(eid,msg);
case 2,% input cell is used to resume the interrupted iterations or
  classI{1} = 'cell';% to interrupt the iteration to resume them later
  J = varargin{1};
  P = varargin{2};
  if length(J)~=length(P),
    msg = sprintf('In function %s, I and INITPSF cell arrays must be of the same size.',mfile-
name);
    eid = sprintf('Images:%s:IandInitpsfMustBeOfSameSize',mfilename);
    error(eid,msg);
  end
end;


% check the Image, which is the first array of the J-cell
[sizeI, sizePSF] = padlength(size(J{1}), size(P{1}));
classI{2} = class(J{1});

checkinput(J{1},{'uint8' 'uint16' 'double'},{'real' 'nonempty' 'finite'},mfilename,'I',1);

if prod(sizeI)<2,
  msg = sprintf('In function %s, input image must have at least two elements.',mfilename);
  eid = sprintf('Images:%s:inputImageMustHaveAtLeast2Elements',mfilename);
  error(eid,msg);
elseif ~isa(J{1},'double'),
  J{1} = im2double(J{1});
end

% check the PSF, which is the first array of the P-cell
checkinput(P{1},{'uint8' 'uint16' 'double'},{'real' 'nonempty' 'finite' 'nonzero'},mfile-
name,'INITPSF',1);
```

135

```matlab
if prod(sizePSF)<2,
   msg = sprintf('In function %s, initial PSF must have at least two elements.',mfilename);
   eid = sprintf('Images:%s:initPSFMustHaveAtLeast2Elements',mfilename);
   error(eid,msg);
elseif ~isa(P{1},'double'),
   P{1} = double(P{1});
end


% now since the image&PSF are OK&double, we assign the rest of the J & P cells
len = length(J);
if len == 1,% J = {I} will be reassigned to J = {I,I,0,0}
   J{2} = J{1};
   J{3} = 0;
   J{4}(prod(sizeI),2) = 0;
   P{2} = P{1};
   P{3} = 0;
   P{4}(prod(sizePSF),2) = 0;
elseif len ~= 4,% J = {I,J,Jm1,gk} has to have 4 or 1 arrays
   msg = sprintf('In function %s, the input cells must consist of 1 or 4 numerical arrays.',mfile-
name);
   eid = sprintf('Images:%s:inputCellsMustBe1or4ElementNumArrays',mfilename);
   error(eid,msg);
else % check if J,Jm1,gk are double in the input cell
   if ~all([isa(J{2},'double'),isa(J{3},'double'),isa(J{4},'double')]),
      msg = sprintf(['In function %s, second, third, and forth array of the input image cell' ...
         ' have to be of class double.'],mfilename);
      eid = sprintf('Images:%s:ImageCellElementsMustBeDouble',mfilename);
      error(eid,msg);
   elseif ~all([isa(P{2},'double'),isa(P{3},'double'),isa(P{4},'double')]),
      msg = sprintf(['In function %s, second, third, and forth array of the input PSF cell' ...
               ' have to be of class double.'],mfilename);
      eid = sprintf('Images:%s:psfCellElementsMustBeDouble',mfilename);
      error(eid,msg);
   end
end;


% Second, Find out if we have a function to put additional constraints on the PSF
%

function_classes = {'inline','function_handle','char'};
idx = [];
for n = 3:nargin,
   idx = strmatch(class(varargin{n}),function_classes);
   if ~isempty(idx),
      [FunFcn,msg] = fcnchk(varargin{n}); %only works on char, making it inline
      if ~isempty(msg),
```

136

```
        eid = sprintf('Images:%s:fcnchkError', mfilename);
        error(eid,msg);
    end
    FunArg = [{0},varargin(n+1:nargin)];
    try % how this function works, just in case.
      feval(FunFcn,FunArg{:});
    catch
      Ftype = {'inline object','function_handle','expression ==>'};
      Ffcnstr = {' ',' ',varargin{n}};
      msg = sprintf(['DECONVBLIND cannot continue because user supplied' ...
              ' %s %s\n failed with the error below.\n\n%s '], ...
            Ftype{idx},Ffcnstr{idx},lasterr);
      eid=sprintf('Images:%s:userSuppliedFcnFailed',mfilename);
      error(eid,msg)
    end
    funnum = n;
    break
  end
end

if isempty(idx),
  FunFcn = FunFcn_d;
  FunArg = FunArg_d;
  funnum = funnum_d;
end

%
% Third, Assign the inputs for general deconvolution:
%
checknargin(3,7,funnum,mfilename);
switch funnum,
case 4,%                deconvblind(I,PSF,NUMIT,fun,...)
  NUMIT = varargin{3};
case 5,%              deconvblind(I,PSF,NUMIT,DAMPAR,fun,...)
  NUMIT = varargin{3};
  DAMPAR = varargin{4};
case 6,%              deconvblind(I,PSF,NUMIT,DAMPAR,WEIGHT,fun,...)
  NUMIT = varargin{3};
  DAMPAR = varargin{4};
  WEIGHT = varargin{5};
case 7,%              deconvblind(I,PSF,NUMIT,DAMPAR,WEIGHT,READOUT,fun,...)
  NUMIT = varargin{3};
  DAMPAR = varargin{4};
  WEIGHT = varargin{5};
  READOUT = varargin{6};
end
```

```
% Forth, Check validity of the gen.conv. input parameters:
%
% NUMIT check number of iterations
if isempty(NUMIT),
  NUMIT = NUMIT_d;
else %verify validity
    checkinput(NUMIT,{'double'},{'scalar' 'positive' 'integer' 'finite'},mfilename,'NUMIT',3);
end

% DAMPAR check damping parameter
if isempty(DAMPAR),
  DAMPAR = DAMPAR_d;
elseif (prod(size(DAMPAR))~=1)&~isequal(size(DAMPAR),sizeI),
  eid=sprintf('Images:%s:damparMustBeSizeOfInputImage',mfilename);
  error(eid,'If not a scalar, DAMPAR has to be size of the input image.');
elseif ~isa(DAMPAR,classI{2}),
  eid=sprintf('Images:%s:damparMustBeSameClassAsInputImage',mfilename);
  error(eid,'DAMPAR has to be of the same class as the input image.');
elseif ~strcmp(classI{2},'double'),
  DAMPAR = im2double(DAMPAR);
end

if ~isfinite(DAMPAR),
  eid=sprintf('Images:%s:damparMustBeFinite',mfilename);
  error(eid,'DAMPAR has to be finite.');        .
end        .

% WEIGHT check weighting
if isempty(WEIGHT),
    WEIGHT = repmat(WEIGHT_d,sizeI);
else
    checkinput(WEIGHT,{'double'},{'finite'},mfilename,'WEIGHT',5);
    if (prod(size(WEIGHT))~=1)&~isequal(size(WEIGHT),sizeI),
        eid=sprintf('Images:%s:weightMustBeSizeOfInputImage',mfilename);
        error(eid,'If not a scalar, WEIGHT has to have size of the input image.');
    elseif prod(size(WEIGHT))== 1,
        WEIGHT = repmat(WEIGHT,sizeI);
    end;
end

% READOUT check read-out noise
if isempty(READOUT),
  READOUT = READOUT_d;
elseif (prod(size(READOUT))~=1)&~isequal(size(READOUT),sizeI),
  eid=sprintf('Images:%s:readoutMustBeSizeOfInputImage',mfilename);
```

```
    error(eid,'If not a scalar, READOUT has to be size of the input image.');
elseif ~isa(READOUT,classI{2}),
    eid=sprintf('Images:%s:readoutMustBeSameClassAsInputImage',mfilename);
    error(eid,'READOUT has to be of the same class as the input image.');
elseif ~strcmp(classI{2},'double'),
    READOUT = im2double(READOUT);
end
if ~isfinite(READOUT),
    eid=sprintf('Images:%s:readoutMustBeFinite',mfilename);
    error(eid,'READOUT has to be finite.');
end;
```

B.2 Matlab Code Written to Calculate LER from an SEM Image.

The Matlab code reproduced below was written by the author and Lei Yuan jointly.

```
% This function trys to detect LER from SEM pictures

pack;
stringbase = strtok(stringfile, '.');

if (edge_detect == 'deblur')
    % this if statement below is used for the 9th deblurred image:
    stringfilenew = strcat(stringbase, '_deblur8_9.jpg');
    stringfull = strcat(stringdir, stringfilenew);
else
    % this if statement below is used for the original (not deblurred) image:
    stringfilenew = stringfile;
    stringfull = strcat(stringdir, stringfilenew);
end

% PP below is the image as read in to matlab:
PP = imread(stringfull);

figure(1)
imshow(PP);

stringfullshort = strtok(stringfull,'.');

image_size = size (PP);
```

```
d = image_size (1);
c = image_size (2);
% c is the x-coordinate "length", d is the y-coordinate "length", *after* flipping the image 90
degrees

for i=1:c
   for j=1:d

        % loc1 is the y-coordinate beginning, loc2 is the x-coordinate beginning, *after* flipping
the image 90 degrees
      loc1 = 0;
      loc2 = 0;
      MM(i,j) = PP(loc2+j,loc1+i);
      SS(i,j) = double(MM(i,j));
   end
end


% MM above is the image once shifted so the lines are vertical.  SS is the same thing
% but with "operatable" numbers-- you cannot do mathematical functions on MM:
figure(2);
imshow(MM);


%
% The following is to enhance the contrast.
%
WW = zeros(c,d);

for i=2:c-1
   for j=2:d-1
      WW(i,j)= ((SS(i,j)+SS(i-1,j)+SS(i-1,j+1)+SS(i-1,j-1)+SS(i,j-1)...
         +SS(i,j+1)+SS(i+1,j-1)+SS(i+1,j)+SS(i+1,j+1))/9.0);
   end
end

% This loop is to find the best threshold value for the contrast
% ickstart is the first contrast value and ickend is the last
ickstart = 1;
ickend = 81;
coffset = 0.09;
for ick = ickstart:ickend
   % used to find the optimum threshold contrast value, th

   MM = uint8(WW);

   JJ = MM;
   %
```

```
% The following converts the image into black and white.
% th is adjusted semi-manually.
%
A = min(MM);
B = max(MM);
min_MM = double(min(A));
max_MM = double(max(B));

% this program, "edge_find_h_exact.m", has the contrast improve by 0.01 steps
% with an offset given by coffset
th = coffset + (ick/100);

for i = 1:c
   for j = 1:d
      TT = double(MM(i,j));
      if TT > (min_MM+(max_MM-min_MM)*th);
         MM(i,j) = 255;
      else
         MM(i,j) = 0;
      end
   end
end

%
% Extract the line ---- matrix: edge
%
edge = zeros(c,d);
for i=1:c
   j=1;
   while j<d
      if MM(i,j+1) > 100
         if MM(i,j+2) >100
            if MM(i,j+3) < 100
               edge(i,j) = 255;
               break;
            end
         end
      end
      j=j+1;
   end

   j = d;
   while j>1
      if MM(i,j-1) > 100
         if MM(i,j-2) >100
            if MM(i,j-3) < 100
```

141

```
            edge(i,j) = 255;
                break;
            end
          end
        end
      j=j-1;
    end
end

% the following few lines of code output the data file

if (SEMmachine == 'dav')
    lx = 2.124689826;
end
if (SEMmachine == 'gia')
    lx = 1.769911504;
end
if (SEMmachine == 'lsi')
    lx = 3.06122449;
end
if (SEMmachine == 'lbl')
    lx = 3;
end
if (SEMmachine == 'leo')
    lx = 2.22222;
end
ly = lx;
% lx/ly are the nm/pixel (this is 2.2222nm for 50 kX magnification and 1024 x 768 image
size)

inc = 1;
right = 0;

iternum = (th*100);
iter = num2str(iternum);

imwrite(edge, strcat(stringfullshort,'_extracted_line_',iter,'.jpg'));

file = fopen(strcat(stringfullshort, '_edge_',iter,'.txt'),'w');
for i = 1:c
    for j = 1:d
        if edge(i,j)==255
            if right == 1
                edge3(3,inc) = j*ly;
                % third row in edge3 is x-location of right line
                fprintf(file,'%12.5e \n',edge3(3,inc));
```
142

```
            inc = inc + 1;
            right = 0;
        else
            edge3(1,inc) = i*lx;
            % first row in edge3 is y-location of both lines
            edge3(2,inc) = j*ly;
            % second row in edge3 is x-location of left line
            % note that the whole file gets "rotated 90 degrees" compared to the original SEM
            fprintf(file,'%12.5e %12.5e ',edge3(1,inc),edge3(2,inc));
            right = 1;
        end
    end
  end
end


fclose(file);


inc = inc - 1;


% calculating regression line fitting the LER
% I am removing the top 45 & bottom 45 pixels (100nm) because the top & bottom of the
image
% are not as "quality" as the middle portion
leftfit = polyfit(edge3(1,46:(inc-45)), edge3(2,46:(inc-45)), 1);
leftfittrue = polyfit(edge3(1,:), edge3(2,:), 1);
% leftfit(1,1) is the slope of the left line fit, leftfit(1,2) is the intercept
rightfittrue = polyfit(edge3(1,:), edge3(3,:), 1);
rightfit = polyfit(edge3(1,46:(inc-45)), edge3(3,46:(inc-45)), 1);


% calculating RMS roughness with only middle portion
leftsum = 0;
rightsum = 0;
for i = 46:(inc-45)
    leftsum = leftsum + (leftfit(1,1)*edge3(1,i) + leftfit(1,2) - edge3(2,i))^2;
    rightsum = rightsum + (rightfit(1,1)*edge3(1,i) + rightfit(1,2) - edge3(3,i))^2;
end
leftRM = leftsum / (inc - 91);
leftRMS(ick) = leftRM^0.5;
rightRM = rightsum / (inc - 91);
rightRMS(ick) = rightRM^0.5;


% the following few lines of code are to ensure that improper minima (like when the
% contrast is so high that the whole thing is seen to be white) are ignored. A "real"
% LER value cannot realistically be below around 1 since the detection limit is: 1.112 nm
if leftRMS(ick) < 0.7
    leftRMS(ick) = 20;
```

```
        end
        if rightRMS(ick) < 0.7
            rightRMS(ick) = 20;
        end

        LERsum(ick) = leftRMS(ick) + rightRMS(ick);

        % calculating RMS roughness
        leftsum = 0;
        rightsum = 0;
        for i = 1:inc
            leftsum = leftsum + (leftfittrue(1,1)*edge3(1,i) + leftfittrue(1,2) - edge3(2,i))^2;
            rightsum = rightsum + (rightfittrue(1,1)*edge3(1,i) + rightfittrue(1,2) - edge3(3,i))^2;
        end
        leftRMtrue = leftsum / (inc - 1);
        leftRMStrue(ick) = leftRMtrue^0.5;
        rightRMtrue = rightsum / (inc - 1);
        rightRMStrue(ick) = rightRMtrue^0.5;

    end

    stringdirnew = 'final_results\';
    stringfullnew = strcat(stringdir, stringdirnew, stringfilenew);
    stringfullnewshort = strtok(stringfullnew,'.');

    minLERsum = min(LERsum);
    minick = find(LERsum == minLERsum);
    mincont = (coffset*100) + minick;

    leftRMSmin = min(leftRMS);
    minickleft = find(leftRMS == leftRMSmin);
    mincontleft = (coffset*100) + minickleft;

    rightRMSmin = min(rightRMS);
    minickright = find(rightRMS == rightRMSmin);
    mincontright = (coffset*100) + minickright;

    file = fopen(strcat(stringfullnewshort, '_edge_summary.txt'),'w');
    fprintf(file,'%s\n',strcat('The SEM file named-- ',stringfilenew, '-- has the following LER val-
    ues using ',SEMmachine,' SEM machine:'));
    fprintf(file,'%s\n',' ');
    fprintf(file,'%s\n','Below is the contrast value which gave the minimum LER of left side of
    trench');
    fprintf(file,'%2.0f \n',mincontleft);
    fprintf(file,'%s\n','(And the minimum contrast using the alternative method-- which should not
    be saved-- would have been:)');
```

144

```
fprintf(file,'%2.0f \n',mincont);
fprintf(file,'%s\n',' ');
fprintf(file,'%s\n','Below is the minimum RMS LER for the left side of the trench');
fprintf(file,'%6.5f \n',leftRMSmin);
fprintf(file,'%s\n',' ');
fprintf(file,'%s\n','Below is the alternative LER for the left side of the trench (do not use)');
fprintf(file,'%6.5f \n',leftRMS(minick));
fprintf(file,'%s\n',' ');
fprintf(file,'%s\n',' ');
fprintf(file,'%s\n','Below is the contrast value which gave the minimum LER of right side of
trench');
fprintf(file,'%2.0f \n',mincontright);
fprintf(file,'%s\n','Below is the minimum RMS LER for the right side of the trench');
fprintf(file,'%6.5f \n',rightRMSmin);
fprintf(file,'%s\n',' ');
fprintf(file,'%s\n','Below is the alternative LER for the right side of the trench (do not use)');
fprintf(file,'%6.5f \n',rightRMS(minick));
fprintf(file,'%s\n',' ');
fprintf(file,'%s\n','Below is the minimum "true" RMS LER for the left side of the trench');
fprintf(file,'%s\n','(This value does not disclude the 100nm on each edge which are assumed to
be poor)');
fprintf(file,'%s\n','This is not the value to be used.  It is just for completeness" sake');
fprintf(file,'%6.5f \n',leftRMStrue(minick));
fprintf(file,'%s\n',' ');
fprintf(file,'%s\n','Below is the minimum "true" RMS LER for the right side of the trench');
fprintf(file,'%s\n','(This value does not disclude the 100nm on each edge which are assumed to
be poor)');
fprintf(file,'%s\n','This is not the value to be used.  It is just for completeness" sake');
fprintf(file,'%6.5f \n',rightRMStrue(minick));
fprintf(file,'%s\n',' ');
fprintf(file,'%s\n','The contrast values checked ranged from:');
fprintf(file,'%2.0f \n',((coffset*100) + ickstart));
fprintf(file,'%s\n','To:');
fprintf(file,'%2.0f \n',((coffset*100) + ickend));
fclose(file);
copyfile(strcat(stringfullshort, '_edge_',num2str(min(mincontleft)),'.txt'),strcat(stringfullnew-
short,'_left_data.txt'));
copyfile(strcat(stringfullshort, '_edge_',num2str(min(mincontright)),'.txt'),strcat(stringfull-
newshort, '_right_data.txt'));
copyfile(strcat(stringfullshort,'_extracted_line_',num2str(min(mincontleft)),'.jpg'), str-
cat(stringfullnewshort, '_extracted_left_line.jpg'));
copyfile(strcat(stringfullshort,'_extracted_line_',num2str(min(mincontright)),'.jpg'), str-
cat(stringfullnewshort, '_extracted_right_line.jpg'));

for ick = ickstart:ickend
    th = coffset + (ick/100);
```

```
    iternum = (th*100);
    iter = num2str(iternum);
    delete(strcat(stringfullshort, '_edge_',iter,'.txt'));
    delete(strcat(stringfullshort,'_extracted_line_',iter,'.jpg'));
end


pack;
```

Bibliography

[1]   International Technology Roadmap for Semiconductors, "2002 Update", http://public.itrs.net/Files/2002Update/Home.pdf

[2]   P. Oldiges, et al., "Modeling Line Edge Roughness Effects in sub 100 Nanometer Gate Length Devices", *International Conference on Simulation Semiconductor Processes and Devices IEEE*, p 131 (2000)

[3]   H. Namatsu, M. Nagase, T. Yamaguchi, K. Yamazaki, K. Kurihara, "Influence of Edge Roughness in Resist Patterns on Etched Patterns", *J. Vac. Sci. Technology B*, 16, p 3315 (1998)

[4]   C. W. Koh, K. H. Baik, "Lithographic Performances of Non-Chemically Amplified Resist and Chemically Amplified Resist for 193 nm Top Surface Imaging Process", *Journal of Photopolymer Science and Technology* 13 (4), p 539 (2000)

[5]   R. Kavanagh, et al., "Investigation of Lithographic Performance for 120 nm and sub 120 nm Gate Applications of Advanced ArF Resists based on VEMA Co-Polymers", *Proceedings of the SPIE*, 4690, p 141 (2002)

[6]   G.E. Moore, "Lithography and the Future of Moore's Law", Proceedings of the SPIE, vol. 2440, p 2 (1995)

[7]   Q. Lin, R. Sooriyakumaran, W. S. Huang, "Toward Controlled Resist Line Edge Roughness: Material Origin of Line Edge Roughness in Chemically Amplified Positive-Tone Resists", *Proceedings of the SPIE*, 3999, p 230 (2000)

[8]   S. Yasin, et al., "Influence of Polymer Phase Separation on Surface Roughness of Resist Features", *EIPBN 2003 Conference and Abstracts*, p 381 (2003); not yet fully published.

[9]   M. Williamson, A. Neureuther, "Lithographic Performance of Thin Dendritic Polymer Resists", *J. Vac. Sci. Technology B*, 18 (6), p 3345 (2000)

[10]   H. Fukuda, A. Yamaguchi, "LER as Structural Fluctuation of Resist Reaction and Developer Percolation", *Proceedings of the SPIE*, 4691, p 158 (2002)

[11]   J. Shin, Y. Ma, F. Cerrina, "Depth Dependence of Resist Line-Edge Roughness: Relation to Photoacid Diffusion Length", *J. Vac. Sci. Technology B*, 20(6), p 2927 (2002)

[12]   Q. Lin, D.L. Goldfarb, M. Angelopoulos, S.R. Sriram, J.S. Moore, "Line Edge Roughness in Positive-Tone Chemically Amplified Resists: Effect of Additives and Processing Conditions", *Proceedings of the SPIE*, 4345, p 78 (2001)

[13]   T. Yoshimura, H. Shiraishi, S. Okazaki, "Suppression of Acid Diffusion in Chemical Amplification Resists by Molecular Control of Base Matrix Polymers", *J. Appl. Physics.*, Part I 34, p 6786 (1995)

[14]   M. Ryoo, S. Shirayone, H. Oizumi, N. Matsuzawa, S. Irie, E. Yano, S. Okazaki, "Control of Line Edge Roughness of Ultrathin Resist Films Subjected to EUV Exposure", *Proceedings of the SPIE*, 4345, p 903 (2001)

[15]   Sanchez, M.I.; Hinsberg, W.D.; Houle, F.A.; Hoffnagle, J.A.; Ito, H.; Nguyen, C. "Aerial image contrast using interferometric lithography: effect on line-edge roughness", *Proceedings of the SPIE*, 3678, p160 (1999)

[16]   M Williamson, X. Meng, A. Neureuther, "A Study of the Effects of Image Contrast and Resist Types upon Line Edge Roughness (LER)", *Proceedings of the SPIE*, 4690, p 357 (2002)

[17] M. Yoshizawa, S. Moriya, "Comparative Study of Resolution Limiting Factors in Electron Beam Lithography using the Edge Roughness Evaluation Method", *J. Vac. Sci. Technology B*, 19(6), p 2488 (2001)

[18] S. MacDonald, H. Ito, and C. G. Willson, "Advances in the Design of Organic Resist Materials", *Microelectronic Engineering 1*, p 269 (1983)

[19] R.D. Allen, W.E. Conley, R.R. Kunz, "Deep-UV Resist Technology", *Handbook of Microlithography, Micromachining, and Microfabrication*, Vol. 1 (1997)

[20] J.D. Plummer, M.D. Deal, P.B. Griffin, *Silicon VLSI Technology: Fundamentals, Practice and Modeling*, Upper Saddle River, NJ: Prentice Hall (2000)

[21] K. Adam, *Domain Decomposition Methods for the Electromagnetic Simulation of Scattering from Three-Dimensional Structures with Applications in Lithography*, Doctoral Thesis, University of California- Berkeley (2001)

[22] A.P. Mahorowala, et al., "Impact of Thin Resist Processes on post-etch 'LER'", *SPIE Technical Summary Digest- Microlithography*, p 160 (2003); not yet fully published

[23] C. Nelson, et al., "Comparison of Metrology Methods for Quantifying the Line Edge Roughness of Patterned Features", *J. Vac. Sci. Technology B*, 17, p 2488 (1999)

[24] C. J. Raymond, et al., "Metrology of Subwavelength Photoresist Gratings using Optical Scatterometry", *J. Vac. Sci. Technology B*, 13, p 1484 (1995)

[25] E. K. Lin, "Feature-Shape and Line-edge Roughness Measurement of Deep Sub-micron Lithographic Structures using Small-angle Neutron Scattering", *Proceedings of the SPIE*, 4344, p 414 (2001)

[26] G. Reynolds, J Taylor, "Correlation of Atomic Force Microscopy Sidewall Roughness Measurements with Scanning Electron Microscopy Line-Edge Roughness Measurements

on Chemically Amplified Resists Exposed by X-Ray Lithography", *J. Vac. Sci. Technol B*, 17 (6) p 2723 (1999)

[27]   D.S.C. Biggs, M. Andrews, "Acceleration of iterative image restoration algorithms", *Applied Optics*, 36 (8), p 1766 - 75 (1997)

[28]   R.J. Hanisch, R.L. White, and R.L. Gilliland, "Deconvolutions of Hubble Space Telescope Images and Spectra", *Deconvolution of Images and Spectra,* Ed. P.A. Jansson, 2nd ed., Academic Press, CA, (1997)

[29]   T.J. Holmes, et al., "Light Microscopic Images Reconstructed by Maximum Likelihood Deconvolution", *Handbook of Biological Confocal Microscopy,* Ed. James B. Pawley, Plenum Press, New York, 1995

[30]   S. Postnikov, et al., "Study of Resolution Limits due to Intrinsic Bias in Chemically Amplified Photoresists", *J. Vac. Sci. Technol. B*, 17 (6), p 3335 (1999)

[31]   E. Croffie, *Simulation Tools for Optical Resist Models STORM*, Doctoral Thesis, University of California- Berkeley (2001)

[32]   L. Yuan, A. Neureuther, "Improving chemically-amplified resist modeling for 2D layout patterns", *Proceedings. of the SPIE*, 5039, (2003)

[33]   S. Burns, et al., "Advancements to the Critical Ionization Dissolution Model", *J. Vac. Sci. Technol. B*, 20 (2), p 537 (2002)

[34]   M. Williamson, *Lithographic Performance of Thin Dendritic Polymer Resists*, Master's Thesis, University of California- Berkeley (2000)

[35]   D.C. Tully, et al., "Dendrimer-based chemically amplified resists for sub-100-nm lithography", *Proceedings. of the SPIE*, 3999, p 1202, (2000)

[36] M. Cheng, *Comprehensive Model for Projection Photolithography: Rigorous, Fast and Novel Processing*, Doctoral Thesis, University of California- Berkeley (2002)

[37] H. Ito, et al., "New UV resists with negative or positive tone", *IEEE Symposium on VLSI Technology*, Digest of Papers, p 86 (1982)

[38] Lord Rayleigh, "On the Theory of Optical Images, with Special References to the Microscope", *Phil. Mag* (5), 42:167 (1896)

[39] S. Yasin, et al., "Influence of Polymer Phase Separation on Roughness of Resist Features in UVIII", *Micro and Nano Engineering 2003 Conference*, Abstracts, p 490 (2003)

[40] K. Toh, *Two-Dimensional Images with Effects of Lens Aberrations in Optical Lithography*, M.S. Thesis, University of California- Berkeley (1988)

[41] M.D. Levenson, et al., *IEEE Trans. Electron Devices*, Vol. ED-31 (6), p 753-763 (1984)

[42] C.A. Mack, "Using the Normalized Image Log-Slope", *Microlithography World*, p 23 Feb 2001

[43] C.A. Mack, S. Jug, D.A. Legband, *Proceedings of the SPIE*, 3677, p 415 (1999)

[44] T. Brunner, "Relationship Between the Slope of the HD Curve and the Fundamental Resist Process Contrast", *J. Vac. Sci. Technology B*, 17 (6), p 3362 (1999)

[45] K. Toh, A. Neureuther, "Identifying and Monitoring Effects of Lens Aberrations in Projection Printing", *Proceedings of the SPIE*, 772, p 202 (1987)

[46] L. Leunissen, R. Jonckheere, K. Ronse, "Influence of Gate Patterning on Line Edge Roughness", Conference Proceedings-- *Electron, Ion, and Photon Beam Technology and Nanofabrication* (2003); to be published in J. Vac. Sci. Technology B in 2003

[47] C. Nelson, S. Palmateer, T. Lyszczarz, "Metrology Methods for the Quantification of Edge-Roughness", *Proceedings of the SPIE*, 3332, p 19 (1998)

[48]   M. Yoshizawa, S. Moriya, "A Study of the Acid-Diffusion Effect on Line Edge Roughness

using the Ec ze Roughness Evaluation Method", *J. Vac. Sci. Technology B*, (2002)

[49]   M. Yoshizawa, "Edge Roughness Evaluation Method for Quantifying at-Size Beam Blur

in Electron Beam Lithography", *Proceedings of the SPIE*, 3997, p 301 (2000)

[50]   M. Sanchez, et al., "Line Edge Roughness: Status and Prospects for Measurement

Improvements", Conference Proceedings-- *Electron, Ion, and Photon Beam Technology

and Nanofabrication* (2003); to be published in J. Vac. Sci. Technology B in 2003

[51]   T. Azuma, et al., "Mechanism of ArF Resist Pattern Shrinkage in Critical Dimension

Scanning Electron Microscopy Measurement", Conference Proceedings-- *Electron, Ion,

and Photon Beam Technology and Nanofabrication* (2003); to be published in J. Vac. Sci.

Technology B in 2003

[52]   H.C. Pfeiffer, "PREVAIL-IBM's E-Beam Technology for Next Generation Lithography",

Proceedings of the SPIE, 3997, p 206 (2000)

[53]   C. Krautschik et al., "Impact of EUV Light Scatter on CD Control as a Result of Mask

Density Changes", *Proceedings of the SPIE*, 4688 p 289 (2002)

[54]   E. Reichmanis, et al., "Chemical Amplification Mechanisms for Microlithography",

Chem. Mater., 3 (3) p 394 (1991)

[55]   Information regarding JMP statistical software can be found at the home page: http://

www.jmp.com/

[56]   SAS Institute Inc. software manual: SAS Institute Inc., *JMP Statistics and Graphics

Guide, Version 3*, Cary, NC: SAS Institute Inc. (1995)

[57]   D.C. Montgomery, *Introduction to Statistical Quality Control*, 3rd edition, John Wiley &

Sons, Inc. (1997)

[58]   A. Sen, M. Srivastava, *Regression Analysis: Theory, Methods, and Applications*, New

       York, NY: Springer-Verlag (1990)

[59]   C. Krautschik, "Impact of Flare on the Imaging Performance of an EUV Exposure Tool",

       NGL Workshop, Pasadena, CA (2001)

[60]   D. McMullan, "Scanning Electron Microscopy 1928 - 1965", 51st Annual Meeting of the

       Microscopy Society of America, Cincinnati (1993)

[61]   A.F. Boden, et al., "Massively Parallel Spatially Variant Maximum-Likelihood Restora-

       tion of Hubble Space Telescope Imagery", *Journal of the Optical Society of America A*, 13

       (7), p 1537 (1996)

[62]   W.H. Richardson, "Bayesian-Based Iterative Method of Image Restoration", *Journal of

       the Optical Society of America*, 62 (1), p 55 (1972)

[63]   L.B. Lucy, "An Iterative Technique for the Rectification of Observed Images", *Astronom-

       ical Journal*, 79, p 745 (1974)

[64]   M.M. Mkrtchyan, R.C. Farrow, "Modeling of Electron Backscattering from Topographi-

       cal Marks", *Journal of Applied Physics*, 80 (12), p 7108 (1996)

[65]   H. Niedrig, "Electron Backscattering from Thin Films", *Journal of Applied Physics*, 53

       (4), p R15 (1982)

[66]   G. Messina, et al., "Physical Approximants to Electron Scattering", *Microelectronic Engi-

       neering*, 34 (2), p 147 (1997)

[67]   Personal communication with Avideh Zakhor, Electrical Engineering Professor at UC-

       Berkeley Focusing upon Video and Image Processing Techniques.

[68]   Information regarding Paint Shop Pro or Jasc Software can be found at their website:

       http://www.jasc.com/

[69]   T.J. Holmes, Y-H Liu, "Richardson-Lucy/Maximum Likelihood Image Restoration Algo-

rithm for Fluorescence Microscopy: Further Testing", *Applied Optics*, 28 (22), p 4930

(1989)

[70]   S. Yasin, et al., "Influence of Polymer Phase Separation on Roughness of Resist Features

in UVIII", Micro and Nano Engineering Conference, September 22nd - 25th 2003, Cam-

bridge, UK (2003)

[71]   G. Han, et al., "Modulation and Line Edge Roughness (LER) in Nano-Lithography",

Micro and Nano Engineering Conference, September 22nd - 25th 2003, Cambridge, UK

(2003)

[72]   L.H.A. Leunissen, et al., "Line Edge Roughness: Experimental Results using a Two-

Parameter Model for the Description", Micro and Nano Engineering Conference, Septem-

ber 22nd - 25th 2003, Cambridge, UK (2003)

[73]   S. Eder-Kapl, et al., "Line Edge Roughness (LER) Investigation on Chemically Amplified

Resist (CAR) Materials with Masked Helium Ion Beam Lithography", Micro and Nano

Engineering Conference, September 22nd - 25th 2003, Cambridge, UK (2003)

[74]   L.E. Ocola, "Modeling Electron Beam Lithography Resist Exposure for Nanopatterning

Applications", Micro and Nano Engineering Conference, September 22nd - 25th 2003,

Cambridge, UK (2003)

[75]   Lei Yuan, no published papers on this topic yet. Expect SPIE 2004 publication detailing

LER simulation model.