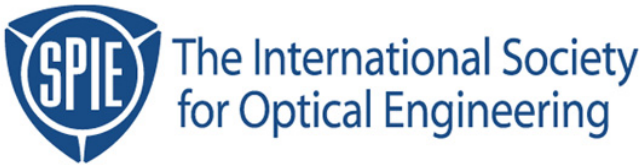


This paper was published in the proceedings of
Photomask and X-Ray Mask Technology VI
SPIE Vol. 3748, pp. 27-40.
It is made available as an electronic reprint with permission of SPIE.

Electron-beam lithography simulation for maskmaking, part IV: Effect of resist contrast on isofocal dose

Charles Sauer
Etec Systems, Inc., 26460 Corporate Avenue, Hayward, California 94545 USA

Chris A. Mack
FINLE Technologies, Austin, Texas

ABSTRACT

There is a certain exposure dose at which variations in electron-beam (e-beam) spot size have virtually no impact on the resulting feature width. In optical lithography, this phenomenon is well characterized and is called the *isofocal point*. The exposure that produces a flat response of linewidth versus spot size is called the *isofocal dose*, and the resulting feature width is called the *isofocal critical dimension (CD)*. It is intuitive that operating in the flat portion of the curve will have advantages from a process latitude perspective. Also, it is significant to note that the isofocal CD occurs at widths that are overexposed with respect to the target spacewidth. Typically, this difference is resolved by sizing data so that the dose to size approaches the amount needed to reach the isofocal point. As linewidths continue to shrink, sizing will become a point of contention, because resolution can be limited by the magnitude of data bias.

In this paper, we examine the effect of resist contrast on the difference between dose to size and dose to isofocal. ZEP 7000, a resist from Nippon Zeon, is examined and compared to resists with other different dissolution rates. Included in the resist selection is a high gamma resist that is modeled on chemically amplified resists (CARs).

Keywords: MEBES, ProBEAM/3D, ProDRM, modeling, e-beam lithography, spot size, isofocal dose

1. INTRODUCTION

While the e-beam theory of exposure has been extensive,[1,2,3] integrating the theory of exposure with e-beam resist development has not progressed to the same level as it has for optical lithography. As a result, the effect of writing and process strategies used in e-beam lithography on lithographic performance is not as well understood. An example of this is the use of isofocal exposure in e-beam mask generation. While the industry commonly employs data biasing, the theory as to why lithographic results improve is not well understood. This paper investigates the role of data biasing in maskmaking and how resist contrast and data biasing interact with each other.

2. ISOFOCAL EXPERIMENTS

Four resists were chosen and examined for their response to isofocal: PBS, Nippon Zeon's ZEP 7000, Shipley's SPR 700, and a "high gamma" resist. The first three were chosen because they are used at Etec, are commercially available, and have been tested for their dissolution rates at Etec using the ProDRM software package developed by FINLE Technologies. The high gamma resist is a synthesis of the parameters thought to make up a near infinite contrast material. This theoretical resist is included to determine the effect a "perfect" resist has on the lithographic response of an e-beam system.

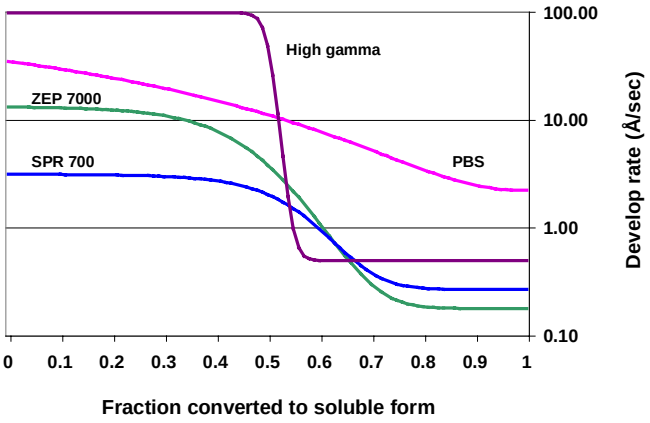
Measurements of isofocal conditions as shown in this section were made by modeling resist responses at various conditions of dose and spot size. All models used 10 kV energy, multipass gray (MPG) writing strategy, a 25 nm input address, a 300 nm resist thickness, and a 700 nm isolated (clear and dark tone) feature. The spot size was varied from 100 to 300 nm at 25 nm increments. The resist critical dimension (CD) was simulated under these conditions using ProBEAM/3D version 5.1n. Table 1 lists the dissolution rate parameters used in the experiments. A 2-D slice was made in each trial and measurements of the CD were made by ProBEAM/3D using the weighted threshold model at a 10% level (resist height from the bottom of the resist). Definition of the dissolution parameters in this paper follows the Mack model and has been described.[4]

Table 1. Dissolution parameters.

Resist	R_{max}	R_{min}	m_{th}	n
PBS	32.9	2.23	–1.0	2.22
ZEP 7000	13.1	0.18	0.45	7.27
SPR 700	2.88	0.27	0.55	7.15
High gamma	100	0.50	0.50	50.0

Figure 1 is a plot of the dissolution parameters for the four resists. The fraction of resist converted to the soluble form is plotted versus the develop rate in angstroms/second. As both Table 1 and Figure 1 show, there is a wide range of resist responses. The high gamma resist is close to infinite contrast, transmitting the aerial image of the exposure into the resist with little or no loss of fidelity. PBS, on the other hand, shows a transfer with a great deal of blurring of the aerial image into the developed resist. As e-beam resists, SPR 700 and ZEP 7000 contrasts are relatively high, but are intermediate between PBS and the high gamma resist.

Figure 1. Semilog plot of fraction converted versus the develop rate for four resists.



Figures 2 through 5 are plots of normalized dose versus linewidth or spacewidth for the four resists at varying spot sizes. A constant develop time was used for each of the four resists. Dose was normalized by comparing that dose to the dose to size for a 700 nm clear feature. This was done for each resist. Isofocal dose for a resist can be estimated at the point where changes in spot size will minimize CD changes.

As expected, the isofocal bias in each case is positive. Table 2 summarizes the results for each resist and feature tone. Results are different than expected — instead of a dependency of bias on resist contrast (with higher contrast resists exhibiting less bias), there is no obvious relationship. This is likely due to the fact that at isofocal, all resists tested exhibited retrograde (i.e., >90°) profiles. The retrograde profiles followed to a large degree the energy deposited in the resist. This is due to a substantial amount of forward scatter at 10 kV.

Results are specific to a 10 kV exposure model and cannot be generally applied to other exposure energies. As with all results in this paper, etch bias that occurs when manufacturing a mask has not been accounted for. The amount of etch bias will influence the optimal amount of lithographic bias that is selected. The difference between the isofocal dose of dark and clear features for the same resist can be explained by proximity. Dark features in this experiment have ~2 µm of surrounding exposure, which provides a higher net effective dose than that received by isolated clear features.

Figure 2. Dose versus (a) spacewidth and (b) linewidth as a function of spot size for PBS at a 55 sec develop time.

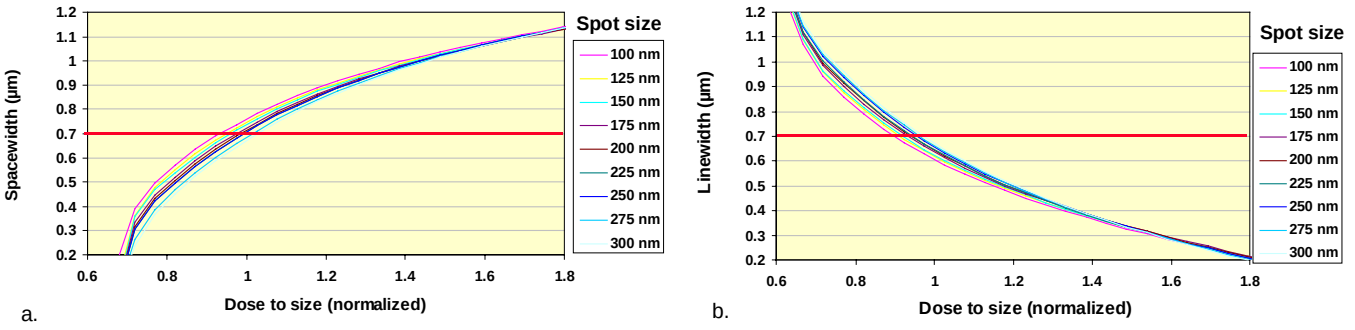


Figure 3. Dose versus (a) spacewidth and (b) linewidth as a function of spot size for ZEP 7000 at 310 sec develop time.

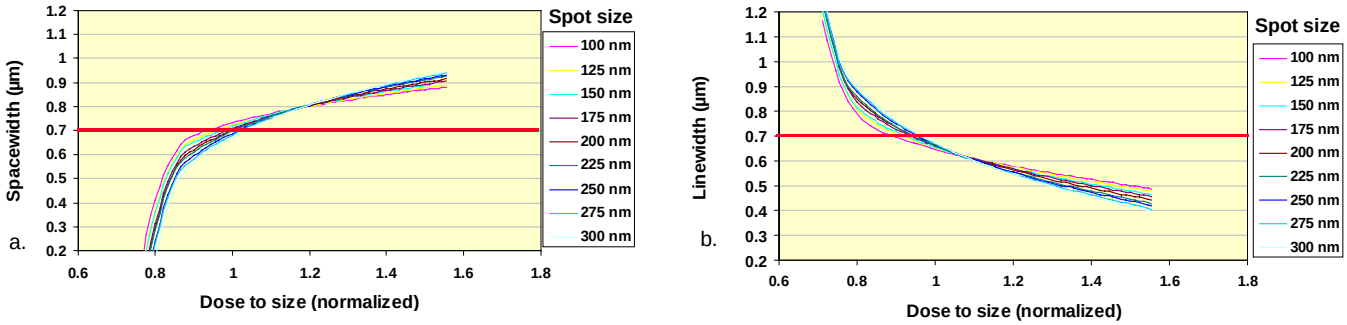


Figure 4. Dose versus (a) spacewidth and (b) linewidth as a function of spot size for SPR 700 at 650 sec develop time.

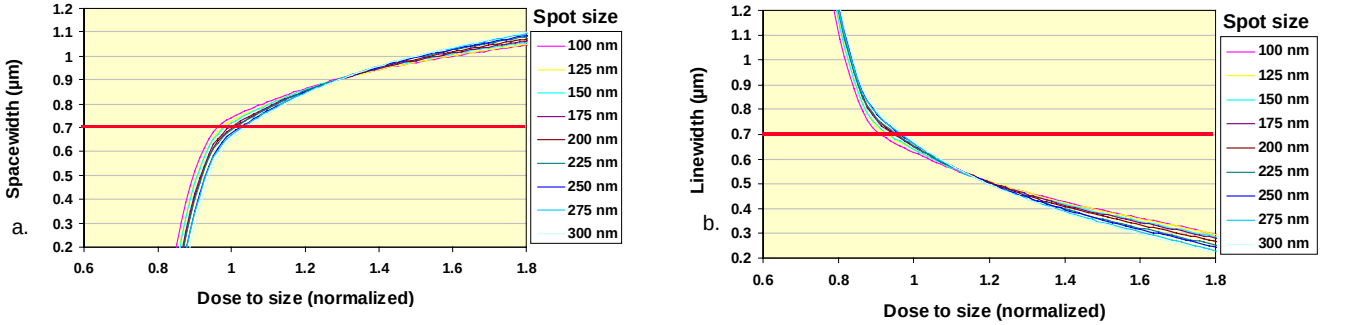


Figure 5. Dose versus (a) spacewidth and (b) linewidth as a function of spot size for a high gamma resist at a 180 sec develop time.

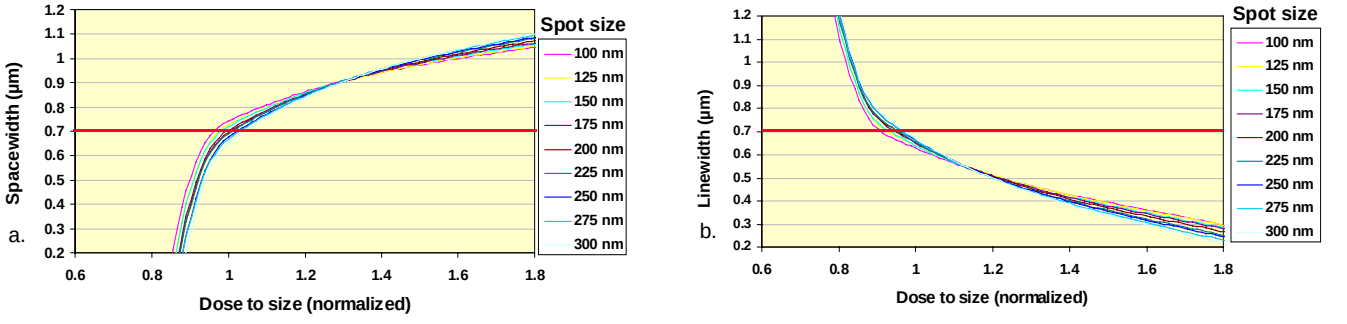


Table 2. Isofocal results.

Resist	Feature	Dose to Size (μC/cm^2)	Isofocal Dose (μC/cm^2)	Isofocal Bias (nm)
PBS	Clear	1.95	3.3	+450
	Dark	1.85	2.9	+380
ZEP 7000	Clear	9.0	10.4	+90
	Dark	8.5	9.6	+70
SPR 700	Clear	8.2	10.8	+230
	Dark	7.8	9.4	+160
High gamma	Clear	2.1	2.7	+190
	Dark	1.9	2.3	+140

Section 4 explores the need to operate at isofocal to get best lithographic results.

3. THEORY OF RASTER SCAN IMAGING

To examine the need for data bias, it is instructive to explore the physics of Gaussian, raster scan imaging. Raster scan imaging can be thought of as the formation of an image by the summation of many spatially distinct base images of (possibly) differing energies. In practice, this summation is always "incoherent," meaning the energy deposited into the photoresist for each base image is added together to form the total energy of the total image. A base image, for example, could simply be the intensity of the spot of a laser beam as it is projected onto the surface of a substrate (or more accurately, deposited into a photoresist film). For the simplest raster scan imaging scheme, only one base image is used, which is called a *spot* or *pixel* image, and each base image has equal energy. Ignoring, for the moment, the distribution of energy through the thickness of the resist film, the raster scan image *I(x,y)* can be described as the summation of many spatially shifted base images:

$$I(x,y) = \sum_{i=1}^{N} P(x-x_i, y-y_i) \tag{1}$$

where P(x,y) represents the image of a single pixel or spot, and the set of points (xi, yi) represent the centers of each pixel that are being summed to form the final image.

For both laser beam and e-beam raster scan writing tools, the pixel image can be fairly approximated as a Gaussian beam. As a simple starting place, let us assume that the only base image being used is a symmetric Gaussian pixel. Normalizing the pixel to have a peak intensity of 1,

$$P(x,y) = \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) = \exp\left(-\frac{x^2}{2\sigma^2}\right)\exp\left(-\frac{y^2}{2\sigma^2}\right) \tag{2}$$

where the width of the Gaussian is defined by σ. Assume that the raster scan strategy being employed uses a fixed address grid with grid spacings Δx and Δy. The image being produced is then given by Equation (1) with address grid points turned either on or off.

3.1 Image of a Single Spot

Two very simple examples illustrate the formation of a raster scan image. First, consider the simplest image: a single spot. In this case, only one address grid is "on" so the total image is just

$$I_{spot}(x,y) = \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) \tag{3}$$

Note that this is also the general result of the image in the vicinity of $x = y = 0$ whenever Δx and Δy are much greater than σ.

Using this simple result, we can derive some of the basic lithographic properties of the image. First, where is the edge of the image? In other words, when this image is printed in resist, where will the edges of the resist lie? Although this is a rather complicated question, it can be simplified by assuming the response of the resist to a Gaussian-shaped image is such that the resist edge will occur at an intensity threshold of I_{th}. For example, picking $I_{th} = 0.5$ would produce edges corresponding to the full width half maximum (FWHM) point of the image. In optical lithography, it has been shown empirically that most resists have near optimum performance when $I_{th} \approx 0.3$. For an arbitrary threshold intensity, the position of the left edge (at $y = 0$) is given by

$$x_{edge} = -\sigma\sqrt{2\ln(1/I_{th})} \tag{4}$$

where the center of the spot is at x = 0. As an example, the FWHM ($I_{th} = 0.5$) width ($= 2\,|x_{edge}|$) is equal to 2.355σ.

For the image of a spot, what is the quality of the image? One of the most useful metrics of image quality is the image log-slope: the slope of the logarithm of the image at the nominal photoresist edge. For the Gaussian spot image, the image log slope is just

$$log-slope = \left.\frac{d\ln I}{dx}\right|_{x_{edge}} = \frac{\sqrt{2\ln(1/I_{th})}}{\sigma} \tag{5}$$

It is quite apparent that Equation (5) predicts improved image quality (improved log-slope) for a smaller spot beam (i.e., smaller σ). The image log-slope can be normalized by multiplying by the desired feature width to give the dimensionless normalized image log-slope (NILS). For this case,

$$NILS = 4\ln(1/I_{th}) \tag{6}$$

If $I_{th} = 0.5$, the NILS will have a value of ~2.8. For $I_{th} = 0.3$, the NILS will become ~4.8. This increasing NILS with lower threshold intensity means that resist images printed at these lower threshold intensities will have much greater linewidth control.

3.2 Image of an Edge

A second, somewhat less trivial, example is the image of an edge where all pixels such that $x_i \geq 0$ are turned on.

$$I_{edge}(x,y) = A(y)\sum_{n=0}^{\infty}\exp\left(-\frac{(x-n\Delta x)^2}{2\sigma^2}\right) \tag{7}$$

where *A(y)* is the result of summing all pixels in the y-direction. An interesting special case occurs when the address grid becomes infinitely small. For such a case, the summation in Equation (7) becomes an integral, *A(y)* becomes a constant, and an "ideal edge" image can be formed:

$$I_{ideal-edge}(x,y) = \frac{1}{2}\mathrm{erfc}\left(-\frac{x}{\sqrt{2}\sigma}\right) \tag{8}$$

where the image has been normalized to have a peak value of 1 and erfc(z) is the complimentary error function. Note that this image is ideal only in the sense that it is a limiting behavior and, in fact, may not be the best image for a particular application.

Where is the edge for this image of an edge? For two cases, $\Delta x \gg \sigma$ and $\Delta x \ll \sigma$, the result can be determined analytically. When the address grid is much bigger than the Gaussian spot size, the edge position is just given by Equation (4), because the edge is influenced almost exclusively by the nearest spot. When the address grid is much smaller than the spot size, a simple result is also possible. Consider first the limiting case of an infinitely small address grid. Picking a specific threshold intensity, Equation (8) can be solved for the position of the edge.

$$\begin{aligned} For\ I_{th} = 0.3, \quad x_{edge} &\approx -0.523\sigma \\ For\ I_{th} = 0.5, \quad x_{edge} &= 0 \end{aligned} \tag{9}$$

When the address grid size Δx is >0 but still much smaller than the spot size, the edge position becomes

$$\begin{aligned} For\ I_{th} = 0.3, \quad x_{edge} &\approx -0.523\sigma - \Delta x/2 \\ For\ I_{th} = 0.5, \quad x_{edge} &= -\Delta x/2 \end{aligned} \tag{10}$$

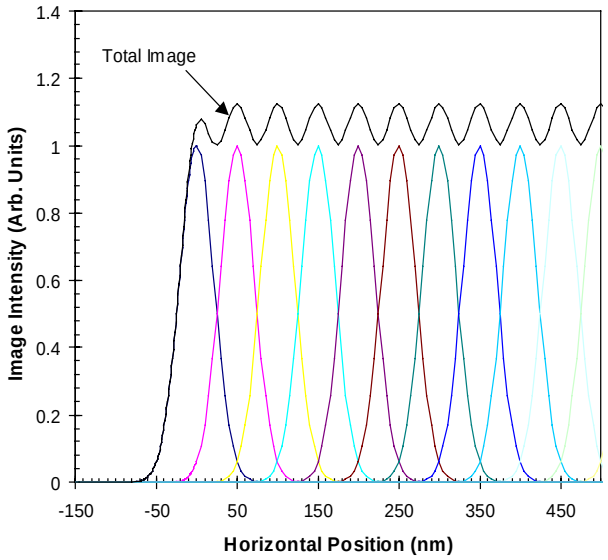
The image quality of the raster scan printed edge can again be described by the image log-slope. Defining the log-slope at the edge position given by Equation (10), the two cases are again $\Delta x \gg \sigma$ and $\Delta x \ll \sigma$. If the address grid is much larger than the spot size, the image log-slope is given by Equation (5), the log-slope of a single spot. When the address grid is much smaller than the spot size, Equation (8) can be used to derive the log-slope.

$$\begin{aligned} For\ I_{th} = 0.3, \quad log-slope &\approx \frac{1.16}{\sigma} \\ For\ I_{th} = 0.5, \quad log-slope &\approx \frac{0.798}{\sigma} \end{aligned} \tag{11}$$

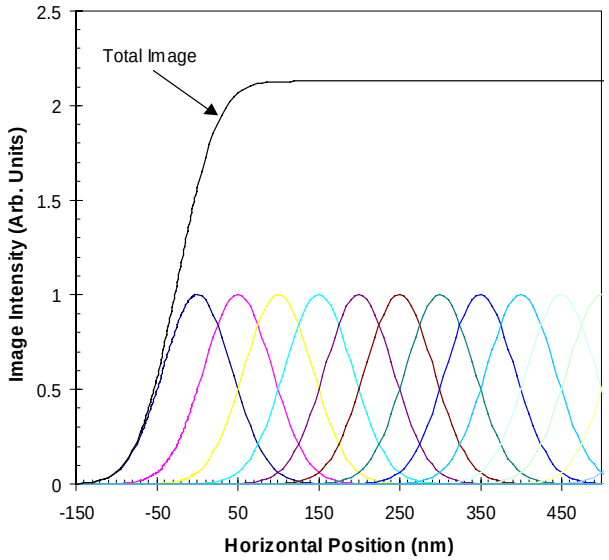
As in the case of a single spot, both a smaller spot size and a lower I_{th} results in a better quality image of the edge.

The limiting cases discussed in this section provide analytic solutions to the position of the edge (relative to the position of the center of the first spot) and the log-slope of the image at the edge. Numerical evaluation of Equation (7) can provide more general solutions. Figure 6 shows two examples of calculating the image of an edge for different spot sizes. From these calculated images, the edge positions and image log-slopes can also be calculated. Figures 7 and 8 show the results for different address grids and spot sizes assuming both $I_{th} = 0.3$ and $I_{th} = 0.5$. It is apparent that the behavior of a single spot and the behavior of the ideal edge (address grid = 0) provide an envelope for the imaging behavior.

Figure 6. The image of an edge calculated as the sum of many Gaussian spots for an address grid of 50 nm and spot sizes of (a) 50 nm and (b) 100 nm.

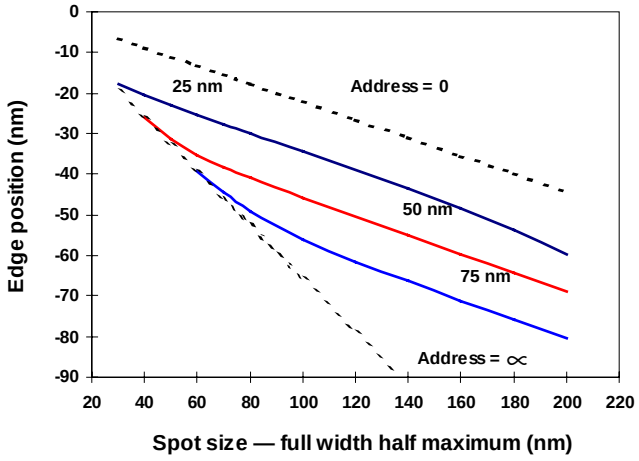


a.

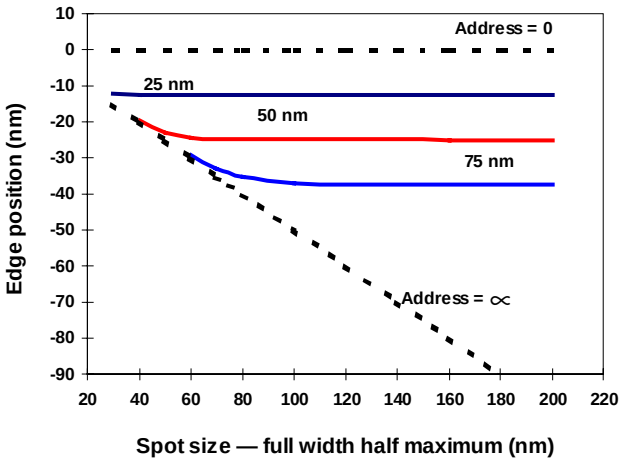


b.

Figure 7. The effects of spot size and address grid size on the position of the edge of an edge image assuming the edge occurs at (a) I_{th} = 0.3, and (b) I_{th} = 0.5. The limiting cases of a single spot and the case where the address grid is zero are shown as dotted lines.

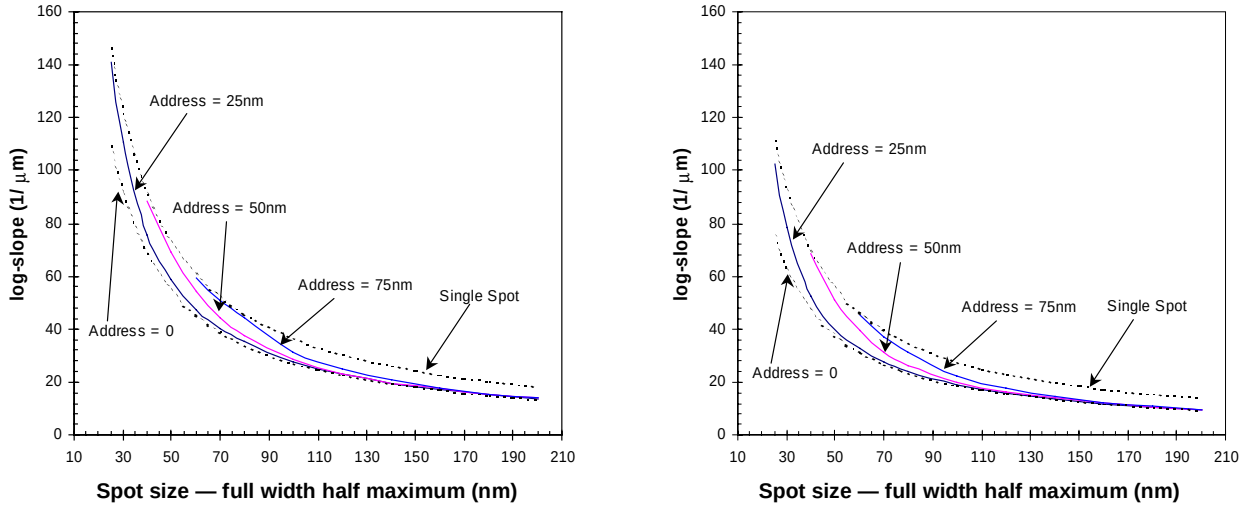


a.



b.

Figure 8. The effects of spot size and address grid size on the image log-slope of an edge image assuming the edge occurs at (a) I_{th} = 0.3, and (b) I_{th} = 0.5. The limiting cases of a single spot and the case where the address grid is zero are shown as dotted lines.



3.3 Summary of Raster Theory

The cases shown in this section are aerial images only and do not include forward scatter. The amount of biasing that is necessary to correct for a series of Gaussian spots is dependent on the writing grid and the spot size. More importantly, the assumption of the location of the threshold is a critical component of how much sizing is needed. It is clear that imaging submicron features with a series of Gaussian spots results in an image in which the feature drawn is larger than desired.

4. OPTIMIZATION OF LITHOGRAPHY — DESIGN OF EXPERIMENTS (DOE) OF SIMULATION DATA

For the optimization effort, the same four resists were used, PBS, ZEP 7000, SPR 700, and a high gamma resist. A four-variable, three-level, full-factorial design was chosen. The variables included exposure dose (±20%), development time (±25%), spot size (100–300 nm), and data bias (0–200 nm). The number of trials for each resist tested was 81. For this work, only 700 nm isolated clear features were examined.

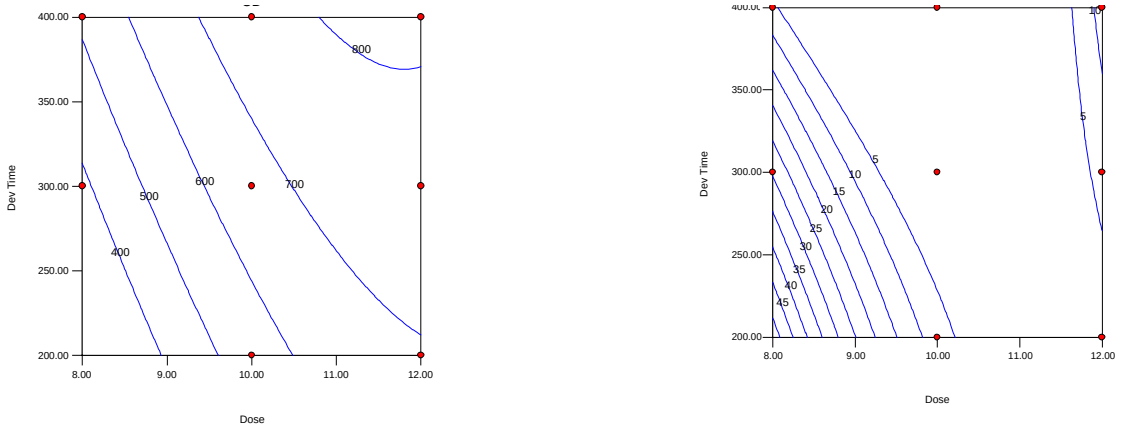
Table 3 lists the variables tested. The responses examined were CD and ΔCD/Δ% dose. The response ΔCD/Δ% dose is a parameter that is easy to simulate and is a key parameter in designing lithography tools. ProBEAM/3D version 5.1n was used for all simulations. A 2–D slice was made in each trial, and measurements of the CD were made by ProBEAM/3D using the weighted threshold model at a 10% level (from the bottom of the resist).

Table 3. Conditions for lithography optimization

Resist	Dose (μC/cm^2)	Develop Time (sec)	Spot Size (nm)	Data Bias (nm)
PBS	2, 2.5, 3	40, 55, 70	100, 200, 300	0, 100, 200
ZEP 7000	8, 10, 12	200, 300, 400	100, 200, 300	0, 100, 200
SPR 700	8, 10, 12	485, 650, 815	100, 200, 300	0, 100, 200
High gamma	2, 2.5, 3	135, 180, 225	100, 200, 300	0, 100, 200

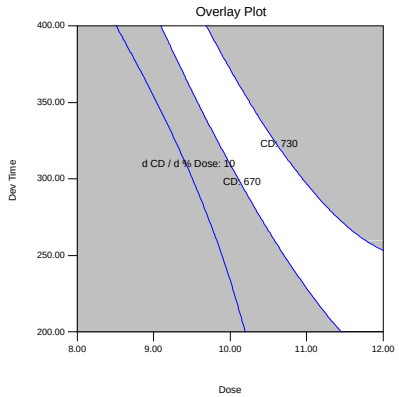
The data from each resist was used to fit a quadratic solution using the DOE software package Stat Expert version 5. Separate solutions to both CD and ΔCD/Δ% dose was derived. Figures 9a and 9b are examples of some of the plots generated. To find an optimum for the process, it is instructive to overlay the plots for CD and ΔCD/Δ% dose and find conditions of simultaneous solutions for both responses. Figure 10 is an overlay plot of the two responses, using acceptable conditions of 670–730 nm CD and a ΔCD/Δ% dose of 0–10nm/%. The white area on the graph in Figure 10 represents where both conditions are simultaneously satisfied. The CD constraints may be considered as the boundary conditions for a test of acceptable ΔCD/Δ% dose. This is a robust way to define process latitude for a series of exposures, develop times, spot size, and data bias conditions.[5]

Figure 9. Contour plots of ZEP 7000. Dose versus develop time at 200 nm spot and –100 nm bias for (a) CD and (b) ΔCD/Δ% dose.



a. b.

Figure 10. Simultaneous optimization of ZEP 7000 for CD and ΔCD/Δ% dose.



Appendixes A–D provide plots of simultaneous optimization analysis for each of four resists. The plots include three spot sizes and three data sizing conditions.

Appendix A contains graphs for PBS. For this series of plots, the target CD is 200 nm smaller than the other resists (470–530 nm). This reflects the need to wet etch PBS. There is no satisfactory solution for plots with a data bias of 0 and –100 nm. At a –200 nm data bias (current industry practice), solutions are shown that exhibit a reasonable process window. The rather low contrast of PBS is evident by the area of acceptable solutions for these nine plots.

Appendix B contains plots for ZEP 7000. The effect of increasing data bias is to shift the acceptable process window towards a higher range of doses. As the data bias increases, the acceptable process window is increased. At –200 nm data bias, the acceptable solution is out of the experimental conditions tested. Responses for ZEP 7000 indicate that the resist response to these conditions is robust.

Appendix C contains plots for SPR 700. As with the ZEP 7000 plots, data bias also shifts the acceptable window towards a higher range of doses, and the effect of increased data bias is to increase the process window.

Appendix D contains plots for the high gamma resist. The dissolution rate parameters have been set up so that little or no blurring of the image transfer is transferred into the resist. These plots describe the limit for 10 kV lithography and this combination of writing strategy, spot sizes, and input addresses. Results show clearly that spot size and bias are two key parameters for optimizing CD performance. Data bias has the most impact on process latitude, with higher data biases improving process latitude. Spot size also improves process latitude with smaller spots showing a small but significant improvement in process latitude.

5. CONCLUSIONS

All four resists tested exhibited an isofocal point that was larger than the designed CD. There was no obvious correlation between resist contrast and isofocal bias. This is likely because of the retrograde (i.e., larger than) 90-degree profiles that occur at isofocal when exposing at 10 kV. The results then may be specific only for the 10 kV case and cannot be extrapolated to higher voltages.

Modeling of raster scan writing shows that the use of a series of Gaussian spots to compose a submicron image results in features that are larger than the desired feature size. The amount of ideal data bias needed to compensate depends on a number of parameters, including writing grid, spot size, and the assumption of the location of the development threshold. Etch bias is also a key parameter that has not been examined in this paper.

Optimization of lithography performance was also examined for these four resists. Overlaying two responses, $\Delta CD/\Delta\%$ dose with CD control is one good technique to estimate process robustness. Results showed that the use of data bias is a key to improving process robustness. The effect of spot size on the process window showed that in most cases, reducing the spot size improves process performance.

6. ACKNOWLEDGMENTS

Thanks to David Alexander for his work on dissolution rate parameters.

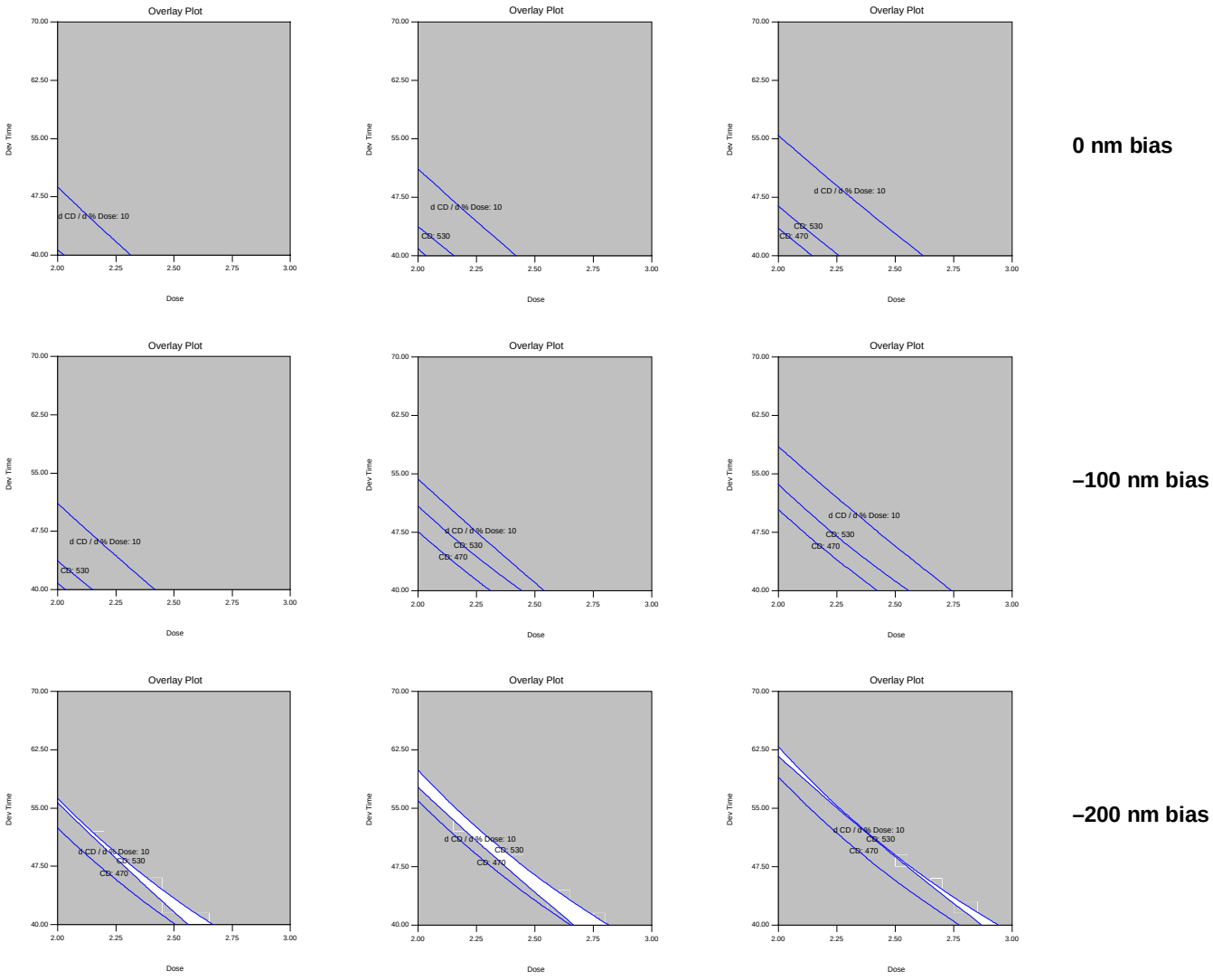
7. TRADEMARKS

8. REFERENCES

1. N. Eib, D. Kyser, and R. Pyle, *Electron Resist Process Modeling*, Chapter 4, *Lithography for VLSI, VLSI Electronics - Microstructure Science*, Academic Press, **vol. 16**, pp. 103–145, New York, 1987.
2. G. R. Brewer, *Electron-Beam Technology in Microelectronic Fabrication*, Academic Press , New York, 1980.
3. K. Valiev, *The Physics of Submicron Lithography*, Plenum Press New York, 1992.
4. C. Sauer, D. Alexander, and C. Mack, "Electron-beam lithography simulation for mask making: II. Comparison of the lithographic performance of PBS and EBR 900-MI," *17th Annual BACUS Symposium on Photomask Technology and Management*, SPIE, **vol. 3236**, pp. 413–423, 1997.
5. C. Mack, "Electron-beam lithography simulation for mask making: III. Effect of spot size, address grid, and raster writing strategies on lithographic performance with PBS and ZEP 7000," *18th Annual BACUS Symposium on Photomask Technology and Management*, SPIE, **vol. 3546**, pp. 32–44, 1998.

APPENDIX A. OPERATING CHARACTERISTIC OVERLAY PLOT — PBS

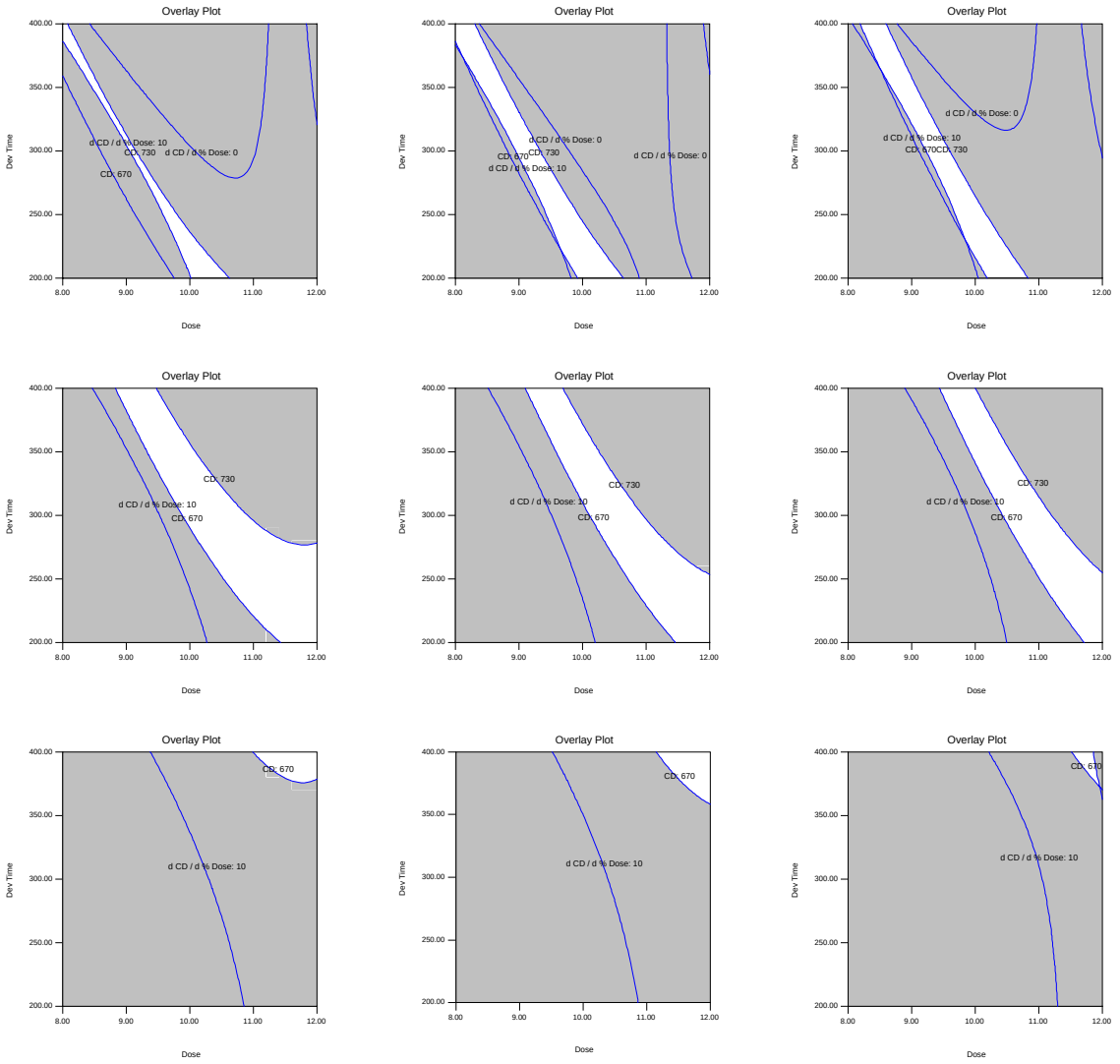
(700 nm clear features, MPG, 25 nm input address)



Operating Conditions (Wet Etch)
CD operating range: 470–530 nm
ΔCD/Δ% dose range: 0–10 nm/%

APPENDIX B. OPERATING CHARACTERISTIC OVERLAY PLOT — ZEP 7000

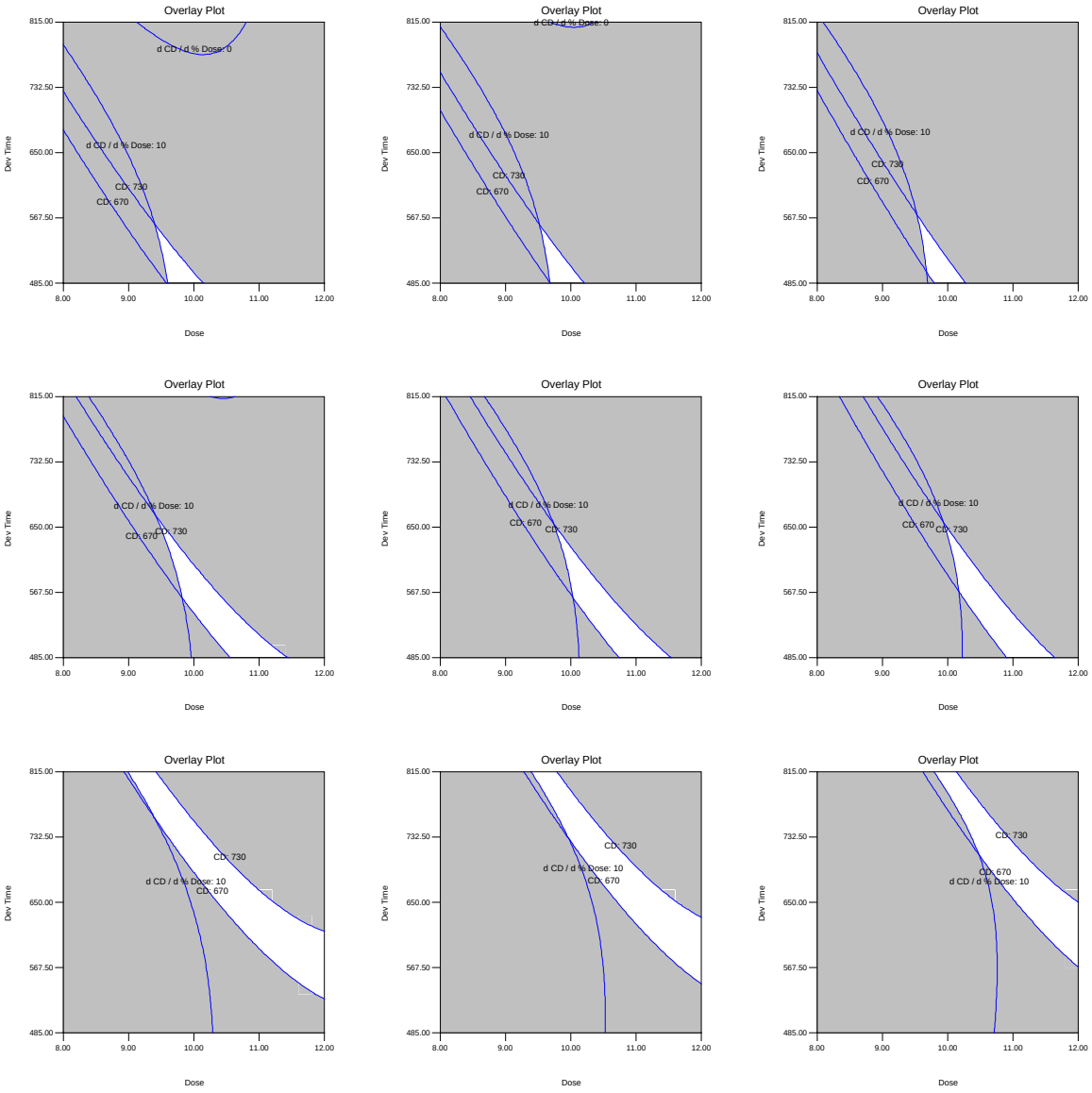
(700 nm clear features, MPG, 25 nm input address)



Operating Conditions (Dry Etch)
CD operating range: 670–730 nm
ΔCD/Δ% dose range: 0–10 nm/%

APPENDIX C. OPERATING CHARACTERISTIC OVERLAY PLOT — SPR 700

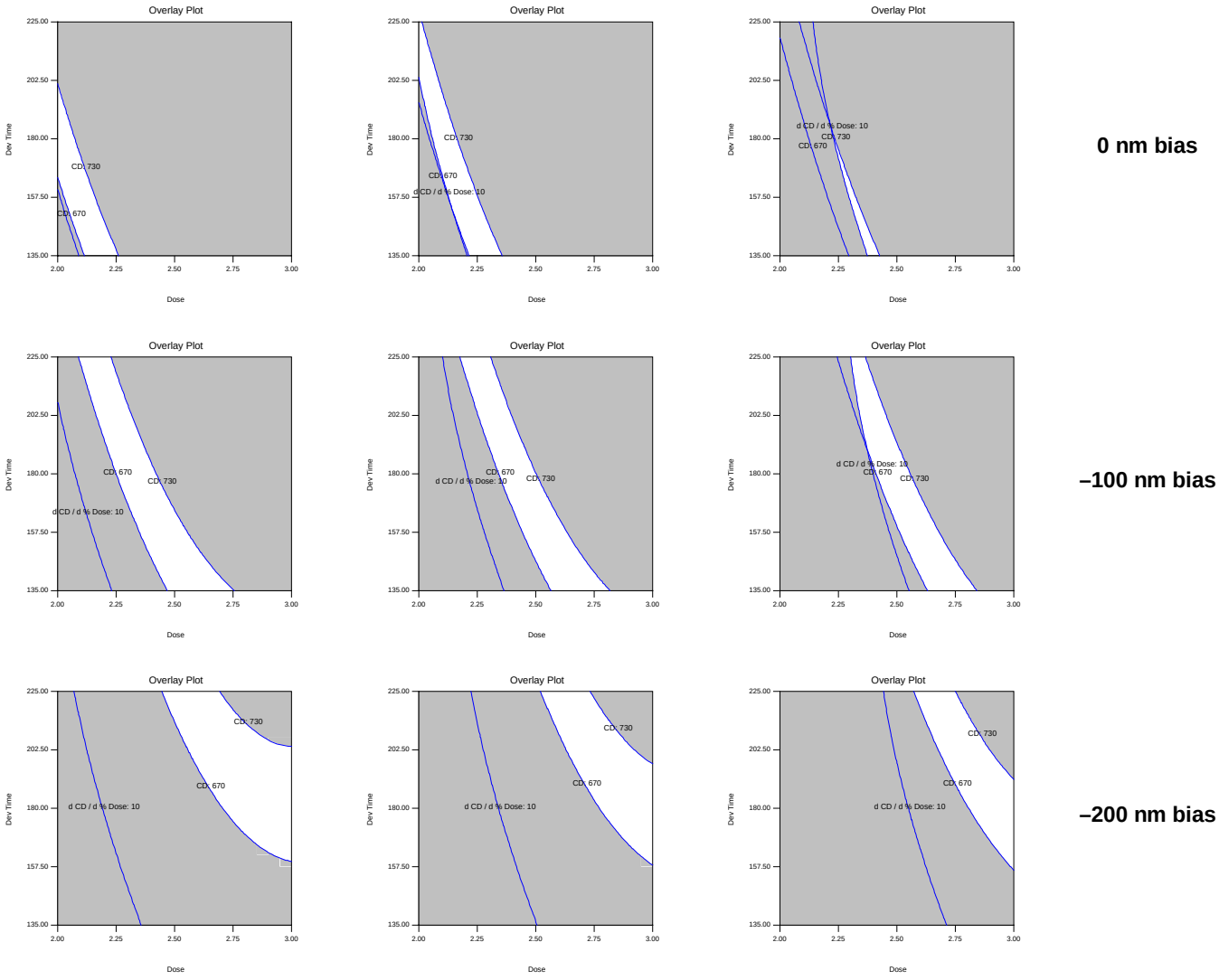
(700 nm clear features, MPG, 25 nm input address)



Operating Conditions (Dry Etch)
CD operating range: 670–730 nm
ΔCD/Δ% dose range: 0–10 nm/%

APPENDIX D. OPERATING CHARACTERISTIC OVERLAY PLOT — HIGH GAMMA RESIST

(700 nm clear features, MPG, 25 nm input address)



Operating Conditions (Dry Etch)
CD operating range: 670–730 nm
ΔCD/Δ% dose range: 0–10 nm/%