Using multiple focal planes to enhance depth of focus

C. A. Spence, D. C. Cole and B. B. Peck

IBM Corporation,
Essex Junction, VT 05452

ABSTRACT

The technique of exposing chips several times, changing the focal plane of the stepper between exposures, has been proposed by others as a way of improving the depth of focus when printing contact holes. We have investigated this technique with aerial image simulation, using criteria based on exposure and defocus latitudes. In addition, we have performed experiments using an i-line stepper to establish the correlation between simulation and what is achievable when printing images in resist. In this paper we investigate the following questions: (i) what is the optimum number of focal planes; (ii) what is the best separation of these planes; (iii) what is the effect on exposure dose; and (iv) what is the net process improvement?

1. INTRODUCTION

As we use optical lithography to print smaller features, we find that the depth-of-focus (DOF) budget decreases rapidly. For a lens of a given numerical aperture (NA) and an exposure wavelength (λ) the available depth of focus may be approximated as $\pm 0.5\ \lambda/NA^2$. In practice, however, the depth of focus is reduced by: wafer flatness, chuck flatness, lens field curvature and autofocus instabilities (lens heating, atmospheric pressure variations and repeatability errors). The small usable depth of focus restricts the options of circuit designers by limiting the topography (vertical height difference) allowed on the wafer. This restriction is serious for stacked-capacitor DRAM designs where the capacitor structures create large height differences between the array and pitch-limited circuits. This problem is particularly acute at the contact mask level. Consequently, several techniques have been proposed to increase the depth of focus of optical systems when printing contact holes, including: phase shift masks,[1,2] pupil apodization,[3] and exposure of multiple focal planes at a single chip location.[4,5] This last approach is attractive since it does not require that new masks or lenses be built. In addition, the spacing of contact holes on the mask is not constrained by oversized rim or subresolution structures required in some phase shift mask schemes. This paper presents our study of this multiple focal plane (MFP) technique. Fukuda et al.[4,5] (especially reference 5) have already explored many of the benefits and disadvantages of this technique, which they have called FLEX. We have used different simulation and experimental techniques; despite this, our results are in good agreement with the work of Fukuda et al.

For the simulations we have concentrated on what we will call the exposure-defocus (E-D) tree format[6]. Although this approach only considers the aerial image (as opposed to developed resist profiles), it does contain information about the slope of the aerial image profile at the nominal feature size, as we pass through focus, and can reveal more about the process window than a plot of image contrast or peak intensity. We note that our simulation work is based entirely on scalar diffraction theory.

Predicting the developed resist profile from an aerial image is difficult since even the best models do not account for the effects of developer and reactant diffusion which is an important effect in

the development of high-aspect-ratio contact holes.[7] We have therefore performed experiments to see how much of the theoretical benefits could be realized in practice. Electrical test structures[8] have been used to measure sizes of contact holes.

2. SIMULATION

2.1 The Exposure-Defocus (E-D) Tree

The bulk of the simulations have been performed using the E-D format developed by Burn Lin[6]. To give a brief explanation, consider the case of the idealized aerial image shown in Figure 1. The intensity of the aerial image at the mask edge is I_{nom}. The ratio of I_{nom} to the open frame intensity (i.e. $I_{open\ frame}/I_{nom}$) gives us the relative dose required to print the feature at the desired size ("on-size"), assuming a simple threshold development model (the feature develops where the intensity is greater than a certain threshold). If we wish to print the feature 10% larger than nominal, we see that the intensity of the aerial image ($I_{+10\%}$) is lower than I_{nom}, and so consequently the dose is required to develop out the feature is higher.

To generate an E-D tree we plot the dose required to print on-size as well as over- and under-sized features. This data is collected for many focal positions. When we join the doses required at different focus positions, to print features of the same size we produce the 'branches' of the E-D tree (Figure 2). Having generated the E-D tree, we calculate the process window as follows. A range of feature sizes that the device can tolerate, e.g. ±10%, and a range of exposure doses that the resist is likely to receive due to process variations (e.g. ±10%) are selected. We then maximize the depth of focus while satisfying these constraints.

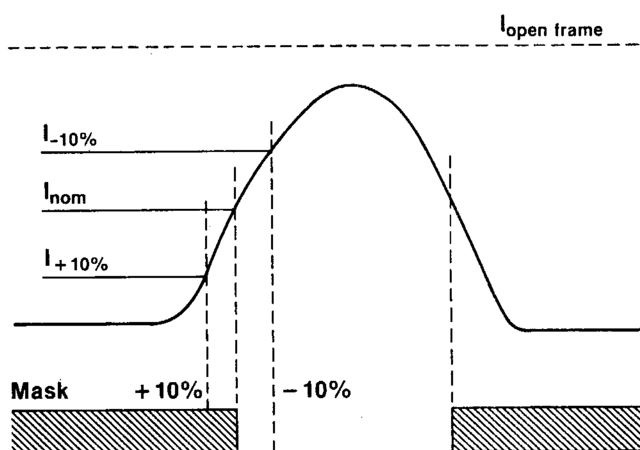


Figure 1. Horizontal cross section of idealized aerial image.

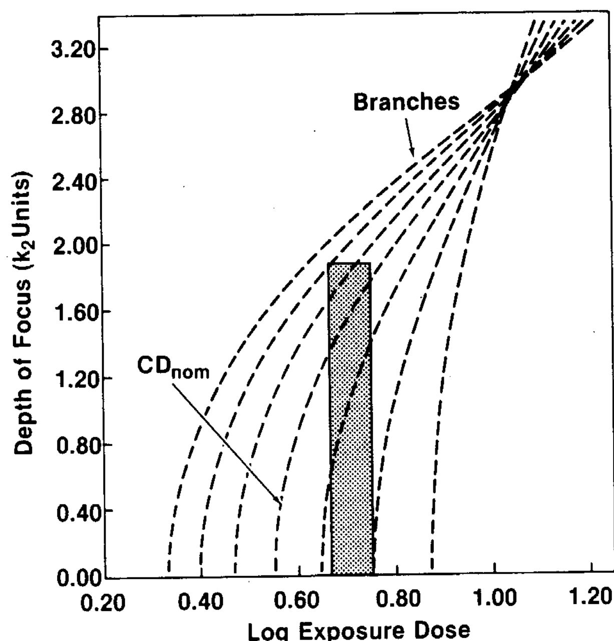


Figure 2. E-D tree for 0.7-μm contact hole printed on a 0.35NA I-line stepper, σ = 0.65 (k_1 = 0.67). The branches of the E-D tree are spaced by 10% CD. The shaded area shows the process window for ±20% dimension control.

The choice of parameters to define the process window is not a simple matter since they depend on the device design, the resist system and the underlying substrate. Some realistic values may be derived as follows. From device performance considerations, a ±20% variation in the linewidth of a feature due to photolithography might be acceptable. However, we must also allow for mask making errors and lens aberrations (e.g. astigmatism and coma, these effects are not accounted for in our E-D trees). A more realistic tolerance for variation in resist image size (solely caused by changes in the aerial image of a perfect lens) might therefore be ±10%.

Exposure dose variations are caused by stepper dosimetry errors and illumination non-uniformity over the exposed field. Variations in the thickness of underlying dielectric films and substrate reflectivity changes also change the amount of energy coupled into the resist film during exposure. As an example, when the thickness of underlying dielectric layers changes, it can result in changes in the coupled energy of as much as ±23%.[9] Using an antireflective coating and a broad-band source, this variation in coupled energy can be reduced to ±4%.[9] Thus the dose variation, accounting for all sources including uniformity and repeatability, can be in the range ±5% to ±30%. A variation of ±10% is considered typical for a good i-line or deep-UV process.

2.2 Benefits of multiple focal planes (MFP) for contact holes

To generate an E-D tree using multiple focal planes, we sum the aerial images (suitably weighted) from the different focal planes and proceed as described in section 2.1. Again, the same threshold model will be assumed for the photoresist development. The result for a contact hole, $k_1 = 0.67$ (with k_1 = Feature Size/(λ/NA)) and partial coherence (σ) = 0.65, using two focal planes (one above the nominal focal plane by $0.75\lambda/NA^2$ and one below the nominal focal plane by $0.75\lambda/NA^2$) is shown in figure 3. We notice that the branches of the tree are closer together

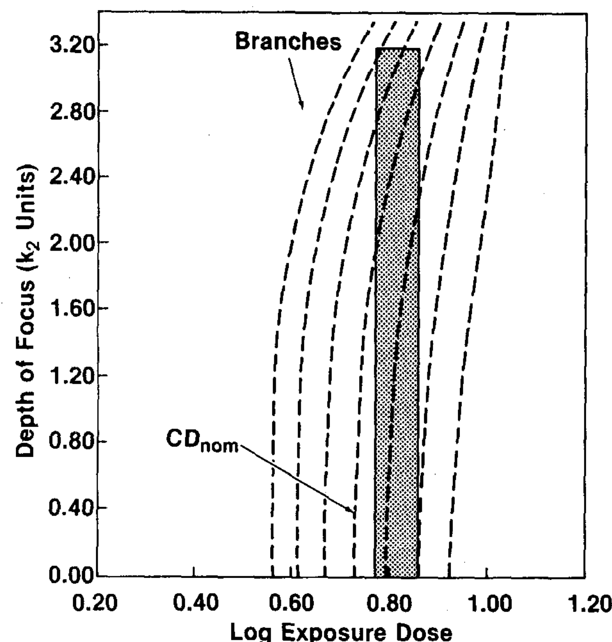


Figure 3. E-D tree for 0.7-μm contact hole printed on a 0.35NA i-line stepper, σ = 0.65 (k1 = 0.67) exposed with two focal planes ±2.5 μm (±0.84λ/NA2) from the nominal focus position. The branches of the E-D tree are spaced by 10% CD. The shaded area shows the process window for ±10% dimension control.

because of the reduction in image slope; however, the branches are also straighter. If we maximize the depth of focus for a ±10% critical dimension (CD) and ±10% exposure dose variation, illustrated in Figure 3 by the shaded box, a large improvement in the depth of focus is clearly predicted when using multiple focal planes. (Note: to convert from k_2 units to microns, multiply by $0.5\lambda/NA^2$)

Varying the separation of the focal planes changes the E-D trees. The effect of this change on the maximum depth of focus for several exposure dose and feature size tolerances has been calculated. The results are shown in figures 4a-d for the case of contact holes, $k_1=0.7$ and 0.5 ($\sigma=0.5$), with two or three equally weighted focal planes. We notice spectacular improvements in the maximum depth of focus (except in the case where a ±20% exposure tolerance is requested for a ±10% feature-size tolerance). The best separation of focal planes, when using three planes, appears to be between 3 and 4 k_2 units (i.e. 1.5 λ/NA^2 and 2.0 λ/NA^2) and around 4 k_2 units for two focal planes ($2\lambda/NA^2$) -- a surprisingly large separation. The relative improvement depends on the exposure and CD tolerances, as illustrated in figures 4a-d. The results are given in tables 1a and 1b.

We summarize the results for contact holes as follows:

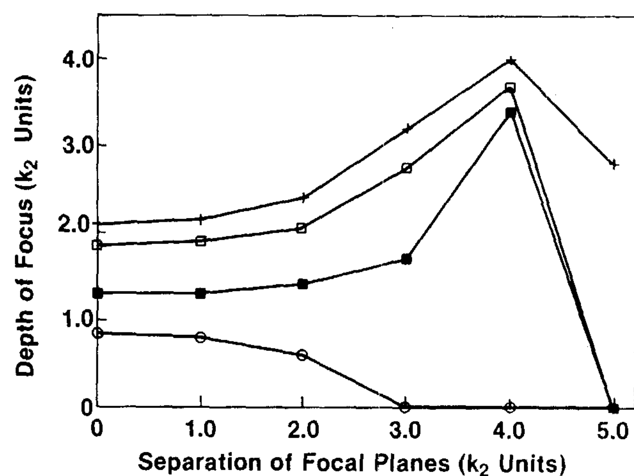
- Apart from the case of ±10% feature-size and ±20% exposure-dose tolerances, the multiple focal plane technique yields an increase in depth of focus.
- The improvement is greater for contact holes at a k_1 factor of 0.7 than for a k_1 of 0.5.
- The optimum focal plane separation is between 1.5 and 2.0 λ/NA^2 when using two focal planes and about $1.5\lambda/NA^2$ when using three focal planes. (This result is true for k_1 factors of both 0.7 and 0.5.)
- When using three focal planes, instead of two, we see additional improvement in depth of focus. When using only two planes, the optimization is more sharply peaked than with three focal planes.
- The exposure dose (when using the best multiple focal plane conditions) increases by about 45% for $k_1 = 0.7$ contacts (Figs. 2 and 3) and by about 100% for $k_1 = 0.5$ contacts.

2.3 Application of multiple focal planes to other patterns

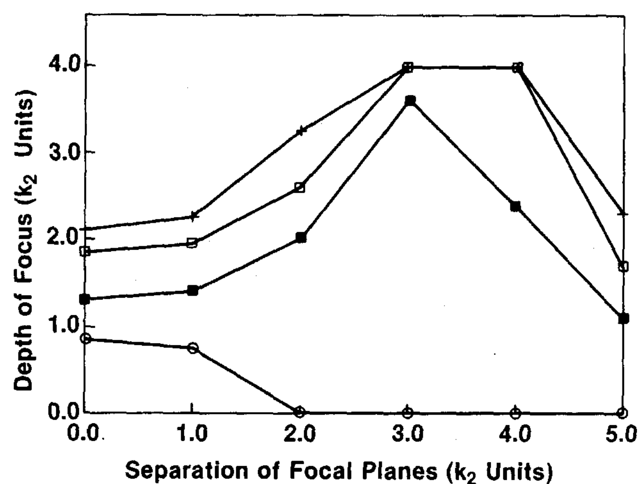
The basis of the multiple focal plane technique is the trade-off between increasing depth of focus at the cost of decreasing image contrast. In the case of contact holes, we have seen that this trade-off appears to be favorable. We have also simulated the effects of a multiple focal plane approach on isolated spaces and gratings for a stepper with $\sigma=0.5$. The results are shown in tables 2 and 3.

The key points arising from the study of isolated spaces and gratings are:

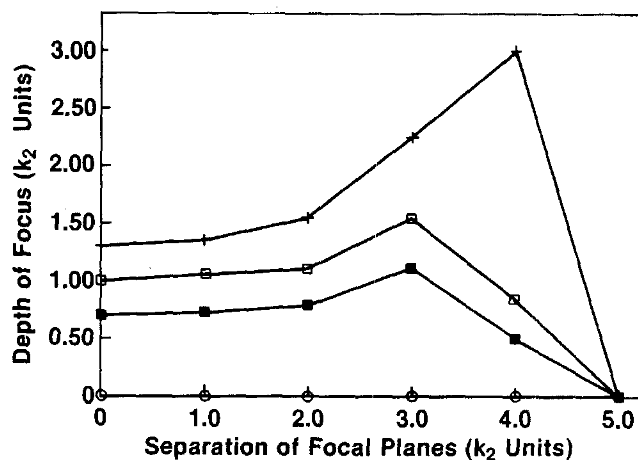
- For isolated spaces, at $k_1 = 0.7$, there is a small improvement in the depth of focus.
- For the isolated space, at $k_1 = 0.5$, the use of multiple focal planes appears to make the process worse; except for the case of a feature size tolerance of ±20% and exposure dose tolerance of ±10%.



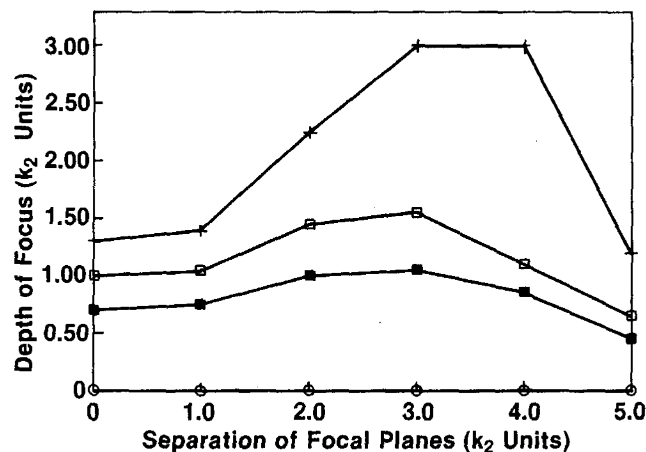
(a)



(b)



(c)



(d)

+10% Exp. Dose; 20% CD
□ 20% Exp. Dose; 20% CD
■ 10% Exp. Dose; 10% CD
○ 20% Exp. Dose; 10% CD

Figure 4. (a) Depth of focus as determined from the E-D trees as a function of focal-plane separation for a contact hole ($k_1 = 0.7$) using two focal planes ($\sigma = 0.5$). (b) Depth of focus as determined from the E-D trees as a function of focal-plane separation for a contact hole ($k_1 = 0.7$) using three focal planes ($\sigma = 0.5$). (c) Depth of focus as determined from the E-D trees as a function of focal-plane separation for a contact hole ($k_1 = 0.5$) using two focal planes ($\sigma = 0.5$). (d) Depth of focus as determined from the E-D trees as a function of focal-plane separation for a contact hole ($k_1 = 0.5$) using three focal planes ($\sigma = 0.5$).

- For gratings, we see improvements for $k_1 = 0.7$, feature size tolerance $\pm20\%$, exposure dose tolerance $\pm10\%$, but for all other cases the multiple focal plane technique degrades the performance.
- The optimum focal plane separation is $1.5\lambda/NA^2$ when using three focal planes. (This result is true for k_1 factors of both 0.7 and 0.5.)

Table 1a Improvement in DOF using MFP for contact holes with $k_1 = 0.5$.

	Exposure Dose ± 10%		Exposure Dose ± 20%	
	CD ± 10%	CD ± 20%	CD ± 10%	CD ± 20%
2 Planes	> 50%	> 100%	No Process	> 50%
3 Planes	> 50%	>100%	No Process	> 50%

Table 1b Improvement in DOF using MFP for contact holes with $k_1 = 0.7$.

	Exposure Dose ± 10%		Exposure Dose ± 20%	
	CD ± 10%	CD ± 20%	CD ± 10%	CD ± 20%
2 Planes	> 100%	> 100%	Worse	> 100%
3 Planes	> 100%	>100%	Worse	> 100%

Table 2a Improvement in DOF using MFP for isolated spaces with $k_1 = 0.5$.

	Exposure Dose ± 10%		Exposure Dose ± 20%	
	CD ± 10%	CD ± 20%	CD ± 10%	CD ± 20%
2 Planes	Worse	> 60%	No Process	Worse
3 Planes	Worse	> 60%	No Process	Worse

Table 2b Improvement in DOF using MFP for isolated spaces with $k_1 = 0.7$.

	Exposure Dose ± 10%		Exposure Dose ± 20%	
	CD ± 10%	CD ± 20%	CD ± 10%	CD ± 20%
3 Planes	Same	> 80%	Worse	> 30%

Table 3a Improvement in DOF using MFP for gratings with $k_1 = 0.5$.

	Exposure Dose ± 10%		Exposure Dose ± 20%	
	CD ± 10%	CD ± 20%	CD ± 10%	CD ± 20%
2 Planes	Worse	Worse	No Process	Worse
3 Planes	Worse	Worse	No Process	Worse

Table 3b Improvement in DOF using MFP for gratings with $k_1 = 0.7$.

	Exposure Dose ± 10%		Exposure Dose ± 20%	
	CD ± 10%	CD ± 20%	CD ± 10%	CD ± 20%
3 Planes	Worse	>100%	Worse	Worse

The reason the multiple focal planes technique works well for contacts, but poorly in the cases of isolated spaces and gratings, as was also observed in Ref. 5, may be understood as follows. Defocusing causes light from the bright central region to be 'pushed' into the adjacent dark areas. The contact hole is a two-dimensional structure; defocusing pushes light from the center disc into an outside ring. The area of the ring surrounding the central disc is much larger than the area of the central disc. Thus, while the intensity of the central disc is reduced, the intensity in the surrounding disc remains low and so the contrast is not degraded too much. In comparison, an isolated space is a one-dimensional structure. The area of the adjacent strips relative to the central bright line is therefore greater and the degradation of contrast by defocusing is worse. For a grating there are two bright lines adjacent to each dark space, so the effect of defocusing (in terms of pushing light into dark areas) is even worse than for an isolated line.

2.4 Summary of the simulation of multiple focal plane (MFP) technique

- MFP works well for contacts
- MFP can give some improvements for isolated lines (with relaxed CD control)
- MFP is not useful for gratings ($k_1 < 0.7$)
- Additional improvement when using three versus two focal planes
- The improvement in the depth of focus is a sharply peaked function of the focal plane separation.
- Optimum separation of focal planes is between 1.5 and 2.0 λ/NA^2 for two planes and three planes
- Dose increases by 45% (contact, $k_1 = 0.7$), > 100% (contact, $k_1 = 0.5$)

3. EXPERIMENTAL

One of the key results of the simulation study is that the optimum plane separation is $1.5\lambda/NA^2$, i.e. $0.75\lambda/NA^2$ above and below nominal focus. Since the Rayleigh depth of focus is usually quoted as being $\pm 0.5\lambda/NA^2$, the question arises as to whether this theoretical improvement can actually be translated into a usable improvement on the production line, or the research laboratory.

We used a GCA 8500 stepper with a 0.35 NA lens, $\sigma = 0.65$ and exposure wavelength of 365 nm in our experimental study. The focus position of this tool can be adjusted over a 10-μm range. Only 2 focal planes were used since the simulation indicated that the 10-μm focal range of the tool would be insufficient to show large improvements when using three planes.

To implement the multiple focal planes technique, a separate pass was made for each exposure, i.e. two per site, for a total of 162 passes to yield a 9 by 9 focus-exposure matrix. Each chip received two exposures; each dose was half the specified dose for the location. The focus positions were set at half the focal plane separation desired, both above and below the nominal focus position.

The photoresist used in the study was IBM Spectralith[10] (Hoechst AZ-7400). The resist was 1.1-μm thick. Wafers received a 90 sec, 95°C post-exposure bake. The wafers were spray developed for 120 secs using 0.24N TMAH. The experimental study concentrated on contact holes, because these are the features that benefit most from the MFP technique.

3.1 Measurement Techniques

To obtain large amounts of data on the printing of contact holes over a range of defocus and exposure conditions, an electrically testable structure was used (Figure 5). The design and calibration of this structure has been described elsewhere.[8] The conducting substrate used was a 500 Å TiSi film on 2000 Å oxide (ρ_s = 30 - 50 Ω/cm^2). After exposure and development the TiSi was etched using a Cl_2/HCl etch in an AME 5000 RIE etch tool. The resist was stripped and the wafers were probed using a Prometrix EM-1 electrical tester. Figure 6 shows the correlation between electrical and SEM measurements for the TiSi film. The contact hole sizes were calculated from electrical resistance measurements using equation 1 in reference 8.

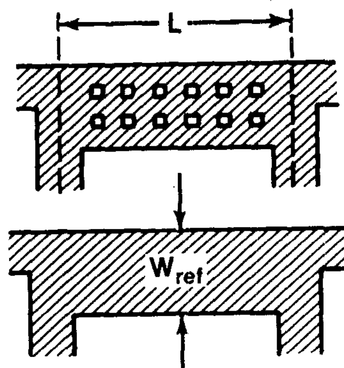


Figure 5. Electrical structure used to measure contact hole size.

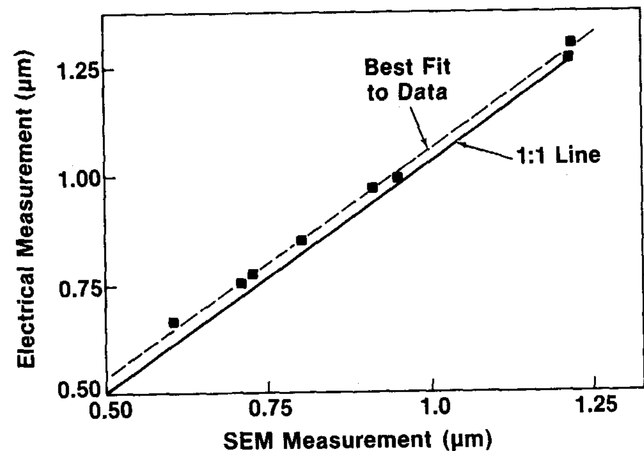


Figure 6. Calibration of electrical test structure for contact holes.

Where B is the diameter of the contact hole, W_{ref} is the width of the reference bar (Figure 5), W is the width (measured electrically) of the bar containing the contact holes, L is the distance between the voltage taps, and N is the number of contact holes in the bar. Over the range studied, we see good linearity and a constant 0.03-μm offset.

3.2 Experimental Results/Discussion

Figures 7a-c show the exposure/defocus behavior of 0.7-μm (k_1 = 0.67) contact holes for focal plane spacings of 0, 3 and 5 μm. We measured the dose to print on size features at best focus as a function of focal plane separation, and compared it with simulation (Figure 8). The simulation predicts a higher dose than found experimentally, but the trend is confirmed by experiment, where the dose changes slowly at first and then more rapidly as the focal plane separation increases.

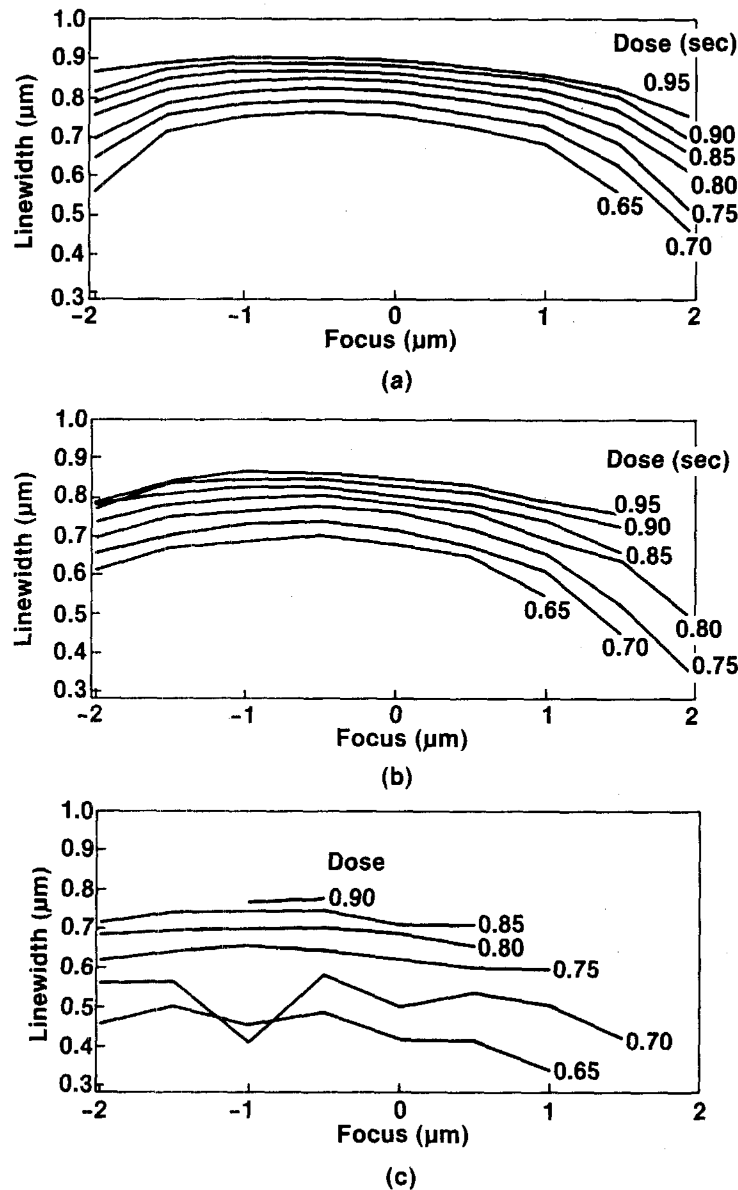


Figure 7. Exposure defocus behavior for 0.7-µm contact holes exposed on 0.35NA I-line stepper, σ = 0.65. (a) 0-µm focal-plane separation, (b) 3-µm focal-plane separation, and (c) 5-µm focal plane separation.

We see that at a focal plane spacing of 3 μm (1 λ/NA^2; see Figure 7b) there is virtually no difference in process window compared with normal exposure (Figure 7a), as predicted by simulation. However, the experimental data is inadequate to show whether there is the large improvement predicted by simulation at a focal plane spacing of 5 μm (= 1.68 λ/NA^2; see Figure 7c). The data presented in Figure 7c is somewhat noisy and there is no data at higher exposure doses. This is because the higher dose required to print contacts, together with the reduction in contrast when using the multiple focal planes technique, results in the probe lines for the electrical structure (1-μm wide) not printing, even though the contact holes still print. The electrical test structure is being redesigned to overcome this problem. In addition, because the stepper has a total focus travel of 10 μm, we can only look at a focal range of $\pm$2.5 μm, when printing with a focal-plane separation of 5 μm. Hence even at doses when the probe lines were intact, we could only study defocusing behavior over a limited range. Despite these problems, SEM cross sections show that although the probe lines were removed, the resist profiles of the contact holes were fairly steep even at a focal-plane separation of 5 μm (Figures 9a-b).

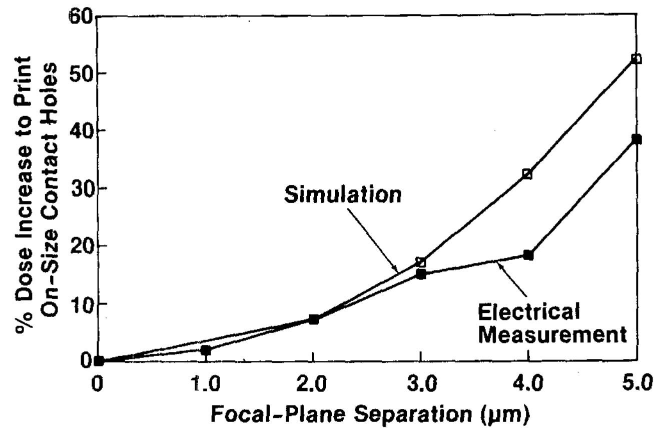


Figure 8. **Dose to print on-size as a function of focal-plane separation for 0.7-μm contacts printed on a 0.35NA, i-line stepper (σ = 0.65).**

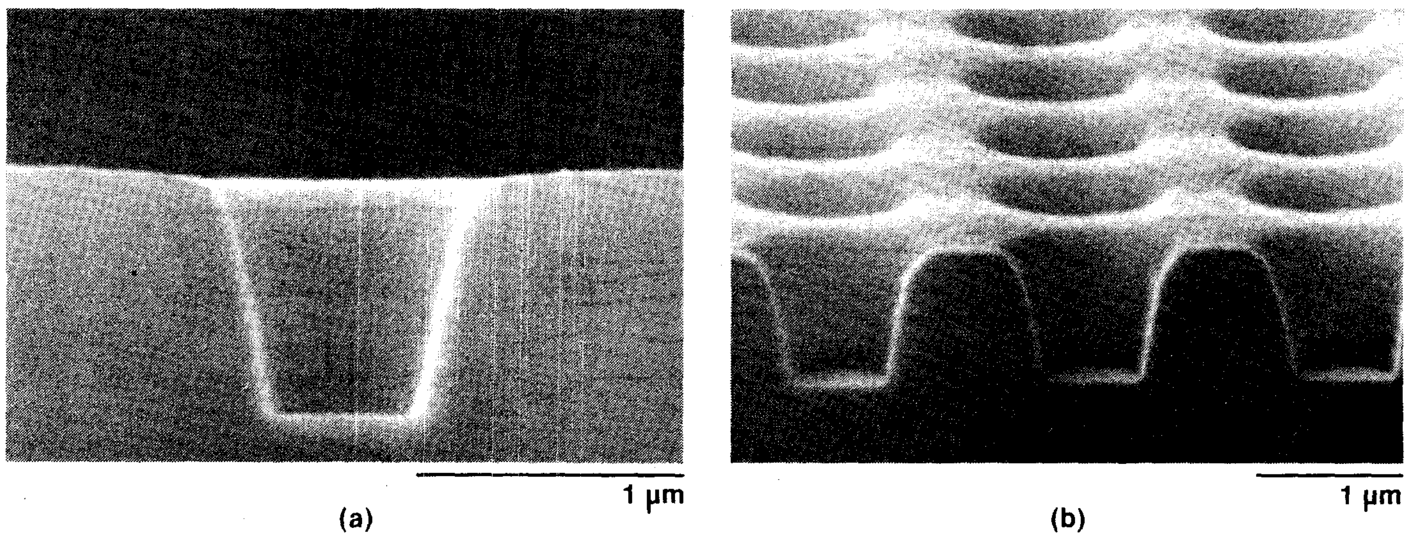


Figure 9. **Cross section SEM of 0.75-μm contact holes with 0.5-μm focal plane separation; (a) isolated (3.25-μm pitch), and (b) dense (1.75-μm pitch).**

3.3 Experimental Conclusions

- Electrical test structures are very useful for gathering large amounts of data to determine the contact process window.
- The increase in dose as focal-plane separation increases was predicted by simulation and verified by experiment.
- Profiles of contact holes are reasonably steep even at a focal-plane separation of 5 μm (1.68 λ/NA^2).
- The need for careful design of test structures is highlighted.

ACKNOWLEDGEMENTS

We thank Jim Peterson for providing the TiSi substrate process; James Bruce for his program to analyze the electrical measurements of contact holes; Nancy Kenyon for etching the wafers; Leon Moszkowicz for performing the SEM measurements used to calibrate the electrical measurement procedure; and Don Samuels for designing the test mask used.

REFERENCES

1. T. Terasawa, N. Hasegawa, N. Kurosaki and T. Tanaka, Proc. SPIE Vol. **1088** p. 25 (1989).
2. J. Garofalo, R. Kostelak and T. Yang, Proc. SPIE Vol. **1463** p. 151 (1991).
3. H. Fukuda, T. Terasawa and S. Okazaki, J. Vac. Sci. B **9(6)** p. 3113 (1991).
4. H. Fukuda, N. Hasegawa, T. Tanaka and T. Hayashida, Electron Device Letters **8** p. 179 (1987).
5. H. Fukuda, N. Hasegawa and S. Okazaki, J. Vac. Sci. Tech. B **7(4)** p. 667 (1989).
6. B. Lin, IEEE Electron Devices **27** p. 931 (1980).
7. J. Bruce, R. Leidy and D. Cole, Proc. KTI Microelectronics Seminar p. 205 (1991).
8. B. Lin, J. Underhill, D. Sundling and B. Peck, Proc. SPIE Vol.**921** p. 164 (1988).
9. D. Dunn, J. Bruce and M. Hibbs, Proc. SPIE Vol. **1463** p. 8 (1991).
10. M. Cagan, D. Kyser, C. Lyons, G. Hefferon and S. Muira, Proc. KTI Microelectronics Seminar p. 177 (1990).