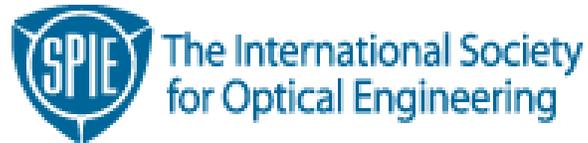


Copyright 2004 by the Society of Photo-Optical Instrumentation Engineers.



This paper was published in the proceedings of the 24th Annual BACUS Symposium on Photomask Technology, SPIE Vol. 5567-45 and is made available as an electronic reprint with permission of SPIE.

One print or electronic copy may be made for personal use only. Systematic or multiple reproduction, distribution to multiple locations via electronic or other means, duplication of any material in this paper for a fee or for commercial purposes, or modification of the content of the paper are prohibited.

The impact of mask topography on binary reticles at the 65nm node

Mark D. Smith^{*}, Jeffrey D. Byers, Chris A. Mack
KLA-Tencor Corp.

ABSTRACT

Most lithography simulation software, such as OPC decoration engines, employ the thin-mask approximation for imaging calculations. While it is well known that a more rigorous approach is often needed for alternating phase-shift reticles, a solution to the full Maxwell equations is rarely used for binary masks. In the past, both the patterns and the thickness of the patterns for COG and attenuated PSM were relatively small compared with the illuminating wavelength. For the future technology nodes, this will not be true. For example, scattering bars are typically a quarter to a third of the size of the main feature. This means that a 65nm isolated line with typically have 65nm to 85nm assist features (mask scale). We have found a non-constant bias through pitch for low k1 imaging that is not found with the thin-mask approach. Results are presented for varying feature sizes, chrome thickness, and for both lines and spaces.

Keywords: Kirchhoff approximation, thin-mask approximation, mask simulation, Maxwell equations, PROLITH

1 INTRODUCTION

One of the most important inputs to any optical lithography simulator is the model for the interaction between the illuminating electric field and the mask. Most lithography simulation software, such as OPC decoration engines, this model is the thin-mask approximation. The thin-mask approximation (or Kirchhoff boundary condition) assumes that the mask transmits light in an ideal way – different regions on the mask transmit the electric field with the ideal transmittance and phase, and the transition region between different types of features is a step function. Of course, the most accurate approach is to solve the Maxwell equations for the electric and magnetic fields, and when we do this, we find that the actual electric field does not match the predictions of the Kirchhoff approximation. The differences are especially dramatic for alternating phase shift masks, where there is an intensity imbalance between the shifted and unshifted clear regions on the mask.

The Kirchhoff boundary condition has such widespread use for several reasons. First, the solution of the full Maxwell equations requires much more computational time than the Kirchhoff approximation. This difference is especially important when performing large scale calculations, such as full-chip OPC decoration. Second, the key assumptions in the Kirchhoff boundary condition are that the thickness of the features on the mask is very small compared with the wavelength, and that the widths of the features on the mask are very large compared with the wavelength.

Until recently, both of these assumptions have been true for binary reticles. However, the shift to 193nm wavelength makes the chrome thickness (or MoSi thickness) relatively larger. Also, the widths of the features on the mask for the 65nm node are relatively smaller as well. While the main features would be 260nm on the 4x reticle, a scattering bar is typically one-quarter to one-third of the size of the main feature, so this feature would be 65nm to 85nm on a 4x reticle, which is less than one-half of the wavelength. The purpose of this paper is to investigate the impact of mask topography on binary reticles at the 65 nm node by performing a detailed comparison of the results predicted by the Kirchhoff approximation with the results calculated by solution to the Maxwell equations.

2 METHODOLOGY

We performed simulations of chrome on glass reticles with simple line-space patterns. PROLITH v8.1.1 with the Mask Topography option was used for all calculations. We simulated an ArF imaging tool with a partial coherence of 0.5 and a numerical aperture of 0.8. For calculations with the full Maxwell equations, we made the following assumptions about the topography and optical properties of the mask: The chrome features on the mask are a homogeneous material with

^{*} mark.d.smith@kla-tencor.com; phone 1-512-231-4212; 8834 North Capital of Texas Highway, Suite 301, Austin, TX 78759

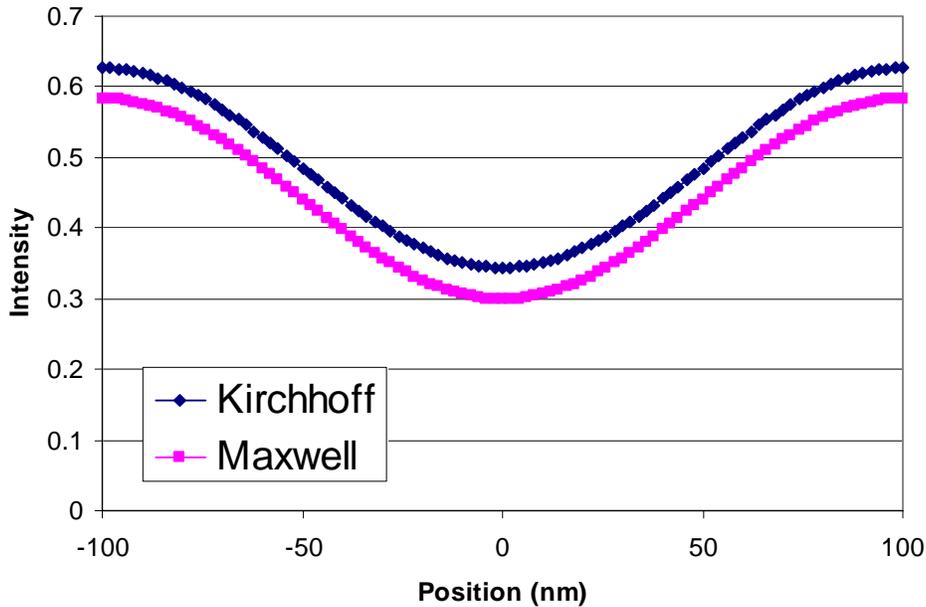


Figure 1: Comparison of the aerial image calculated with the Kirchhoff approximation and the full Maxwell equations for a 70nm line on a 200nm pitch.

optical properties of $\mathbf{n} = 1.40 + 3.26i$, with a nominal thickness of 100nm, and with straight, vertical sidewalls. The quartz substrate has optical properties of $\mathbf{n} = 1.474$. (Note that the actual chrome absorber used in modern masks is much more complex than a simple homogeneous layer, but making this assumption is convenient and should yield representative results.)

Calculated aerial images for 70nm lines on a 200nm pitch are shown in Figure 1 for both the Kirchhoff approximation and the solution of the full Maxwell equations. As shown in the figure, the difference between the two images is an offset in the intensity that causes the aerial image CD for the Maxwell calculations to be larger than the CD calculated with the Kirchhoff approximation. Shown in Figure 2 is the aerial image CD versus pitch for 70nm lines calculated with the Kirchhoff approximation and the full Maxwell equations – the difference between the CDs is more than 20nm! The reason for such a large difference is shown in Figure 3, where we plot the MEEF for this feature as a function of pitch. The MEEF is greater than 3 for all pitch values, so an error at the mask will be greatly amplified. If we view the error in the Kirchhoff approximation as an error in the effective mask CD, then we can convert the aerial image CD differences to a difference at the mask by dividing by the MEEF. This result is also shown in Figure 3, and we see that the effective error at the mask is around 7 to 8nm through pitch.

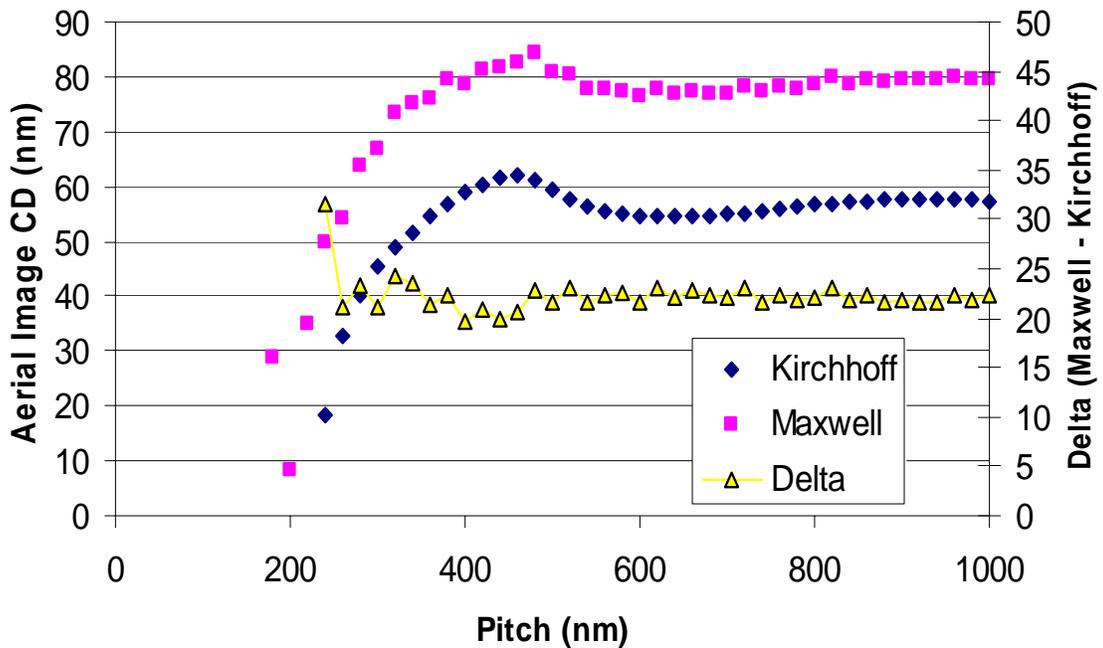


Figure 2: Aerial image CDs for 70nm lines through pitch for calculations with the Kirchhoff approximation and with the full Maxwell equations. The difference between the Kirchhoff and Maxwell CDs is plotted as “Delta”.

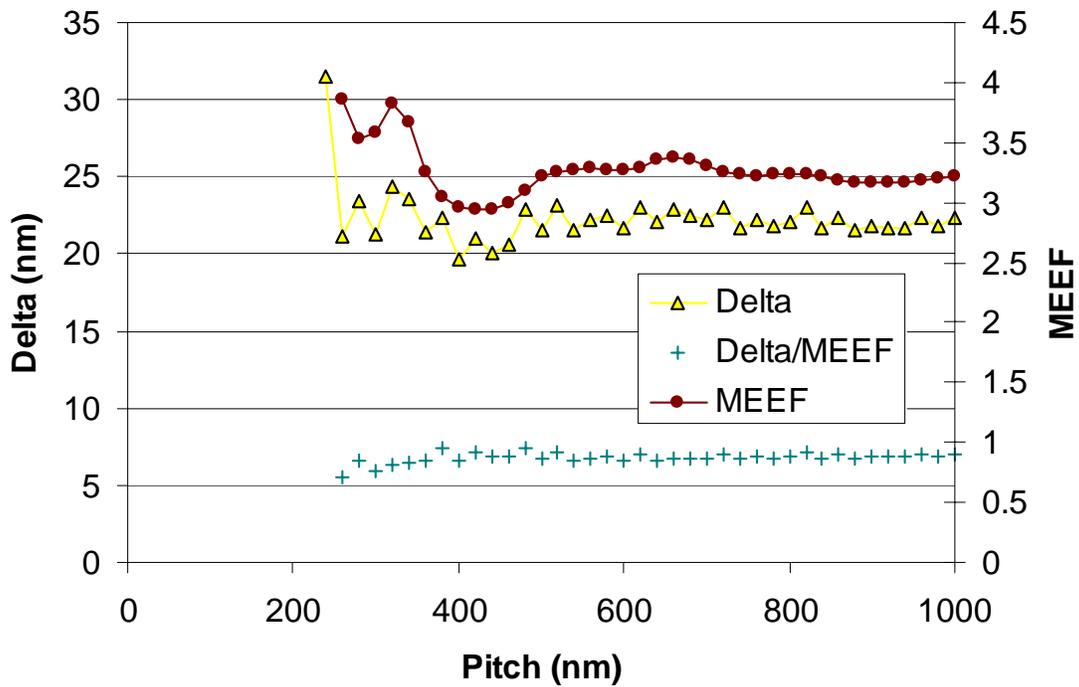


Figure 3: MEEF for 70nm lines through pitch as calculated with the Kirchhoff approximation. Also shown is the Delta between the Kirchhoff and Maxwell CDs shown in Figure 2, as well as the Delta/MEEF, which is about 7 to 8nm through pitch.

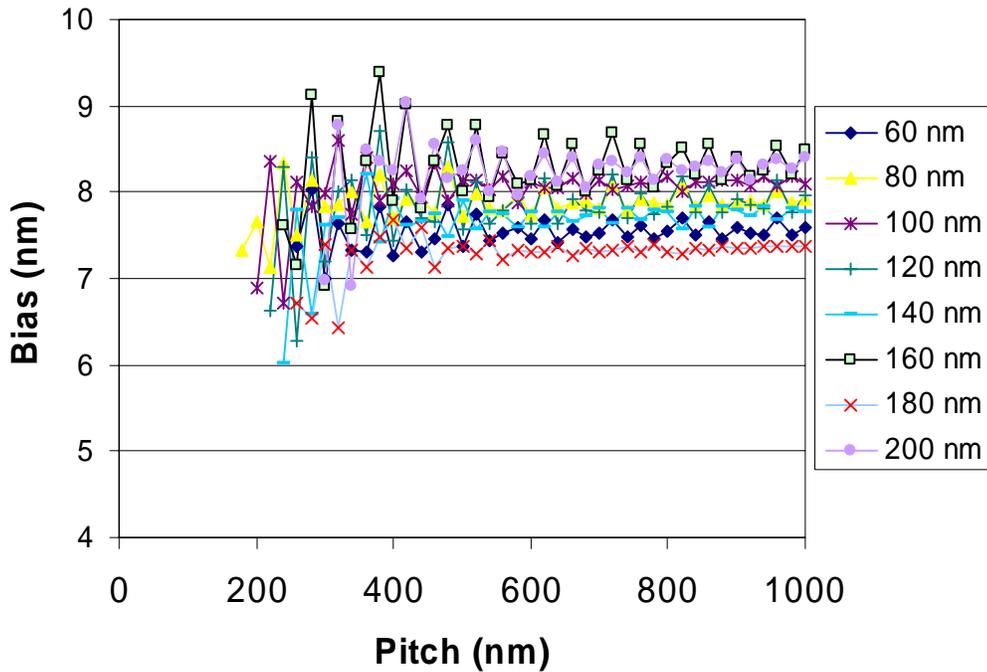


Figure 4: The bias applied to the Kirchhoff mask to match the prediction from the full Maxwell equations (aerial image CD) for a variety of feature widths through pitch.

3 MASK BIAS DUE TO THE KIRCHHOFF APPROXIMATION

The example in the previous section demonstrates the pitfalls of directly comparing aerial images calculated with the Kirchhoff approximation with those from a full Maxwell equations calculation. The difference between the predicted aerial image CDs may depend strongly on the optical settings for the imaging tool – a large MEEF will not only amplify mask manufacturing errors, but also amplify errors due to the Kirchhoff approximation! For this reason, we take a different approach in the remaining sections of this paper: we calculate the bias at the mask that is required for the Kirchhoff result to match the Maxwell result. The bias will be reported in wafer dimensions. As shown in Figure 3, we expect this bias to be around 7 to 8nm – the effective chrome “edge” appears wider for the Maxwell result compared with Kirchhoff. This approach has two main benefits. First, applying the bias at the mask allows us to sidestep the influence of the MEEF. Second, if we can quantify the bias between the Maxwell and Kirchhoff results, then it may be possible to apply the bias to “fix” Kirchhoff calculations in applications where the Maxwell approach is too computationally intensive, such as full chip OPC decoration.

The needed bias for a variety of feature widths is shown in Figure 4 as a function of pitch. On average, the bias is around 7 to 8nm, with a range of 3.3nm. As shown in the figure, the bias is both a function of pitch and feature width. We can use this information to attempt to improve the Kirchhoff results by applying a uniform bias of 8nm (4nm of extra chrome per edge). Shown in Figure 5 is an example of this approach for isolated lines and isolated spaces. The difference between the biased mask with the Kirchhoff approximation and the full Maxwell result is less than a few nanometers for all features except for small isolated spaces, where the differences become larger. One possible reason for this deviation is that for small spaces, the two chrome edges that form the space become close enough that they interact. At this point, the bias becomes dependent on the distance between adjacent edges. In general, we find that the bias approach is a useful correction, but if we need to predict the CD within a few nanometers, a global bias is not accurate enough. Solution to the full Maxwell equations is necessary because the required bias can vary with feature size and feature type. For small spaces, the error in the global bias approach is larger than the CD tolerances at the 65nm node.

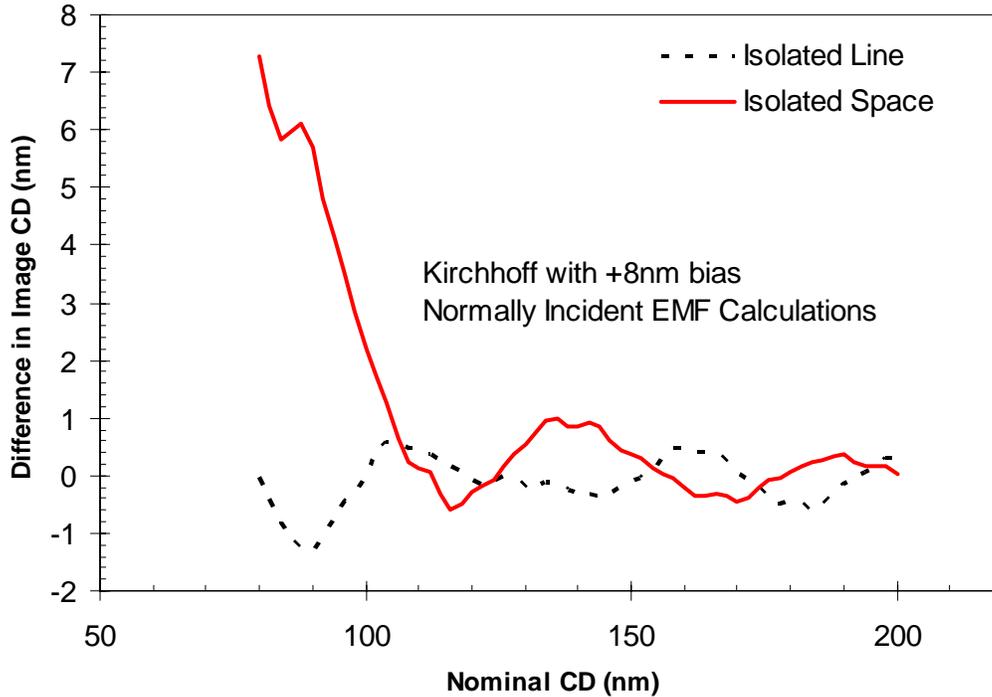


Figure 5: Comparison of the aerial image CDs calculated with the Kirchhoff approximation with an 8nm mask bias and the CDs calculated with solution to the full Maxwell equations. Results are shown for isolated lines and isolated spaces.

We also tested the accuracy of the mask bias for predicting process windows. For these calculations, we used the Lumped Parameter Model in PROLITH with typical parameters. The required bias was extracted from the data in Figure 4. As shown in Figure 6, the agreement between the Kirchhoff with bias and the Maxwell is quite good. We also tried changing the illuminator type from partially coherent to Quasar in order to determine if the required bias would depend on illuminator type. As shown in Figure 7, the agreement is quite good with the bias calculated with the partially coherent source shape, so it appears that the bias is not illuminator-dependent, but more work is required.

Next, we performed calculations where we varied the chrome thickness. The bias for a chrome thickness of 70nm is 5 to 7nm, as shown in Figure 8. Comparison of Figures 4 and 8 demonstrates that the bias decreases as the chrome thickness decreases. This trend is what is expected, as the Kirchhoff approximation requires an “infinitely thin” absorber, and this idealization is better met by the 70nm chrome thickness.

Finally, we investigated the impact of polarization on the calculated bias. All of the calculations up to this point have been for unpolarized illumination. Shown in Figure 9 is the calculated bias for unpolarized, x-polarized, and y-polarized illumination on a y-oriented, isolated, 70nm line. As shown in the figure, the bias for x-polarized light is around 6nm, and the bias for y-polarized light is around 10nm. The results for unpolarized illumination are halfway between the two polarized results, with a bias of around 8nm.

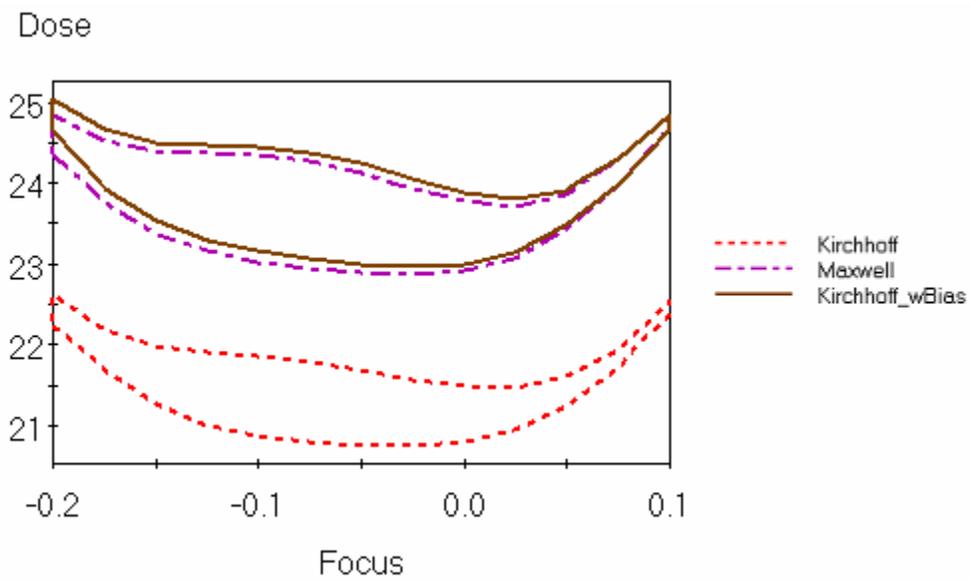


Figure 6: Process windows for dense lines and spaces calculated with the Kirchhoff approximation, Kirchhoff approximation with bias, and the full Maxwell equations.

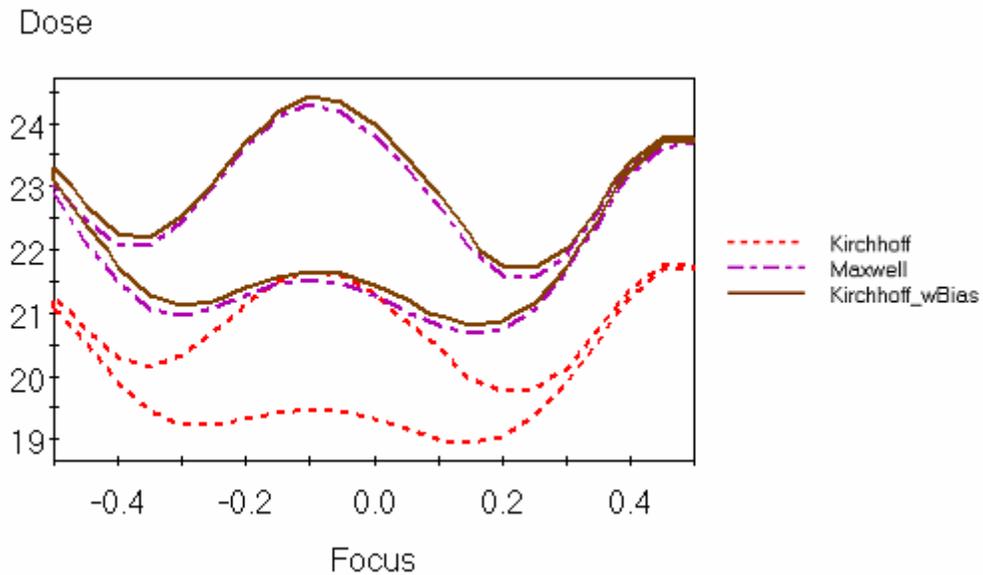


Figure 7: Process windows with Quasar illumination for dense lines and spaces calculated with the Kirchhoff approximation, Kirchhoff approximation with bias, and the full Maxwell equations. The mask bias is the same value used for Figure 6, and was calculated with a partially coherent source shape with $\sigma = 0.5$.

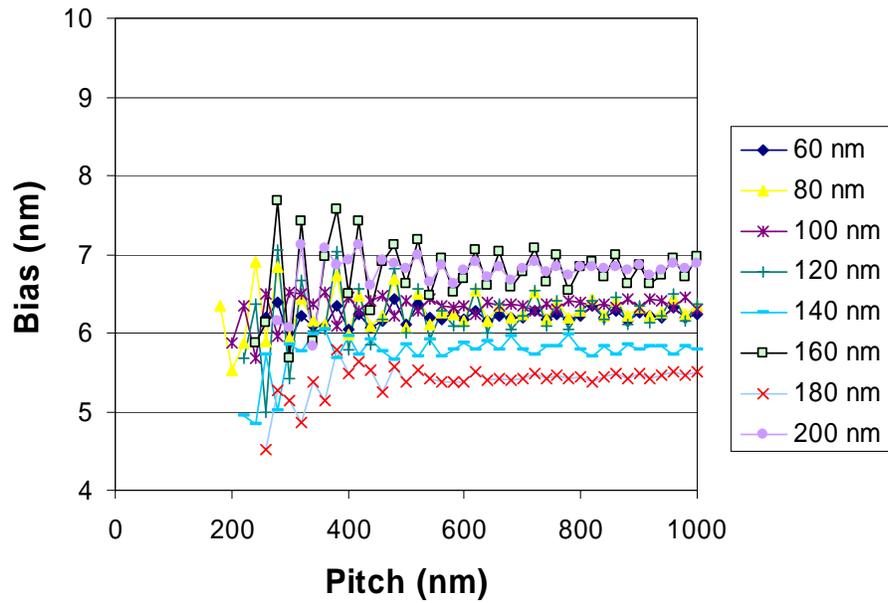


Figure 8: The bias applied to the Kirchhoff mask to match the prediction from the full Maxwell equations for a variety of feature widths through pitch for a chrome thickness of 70nm.

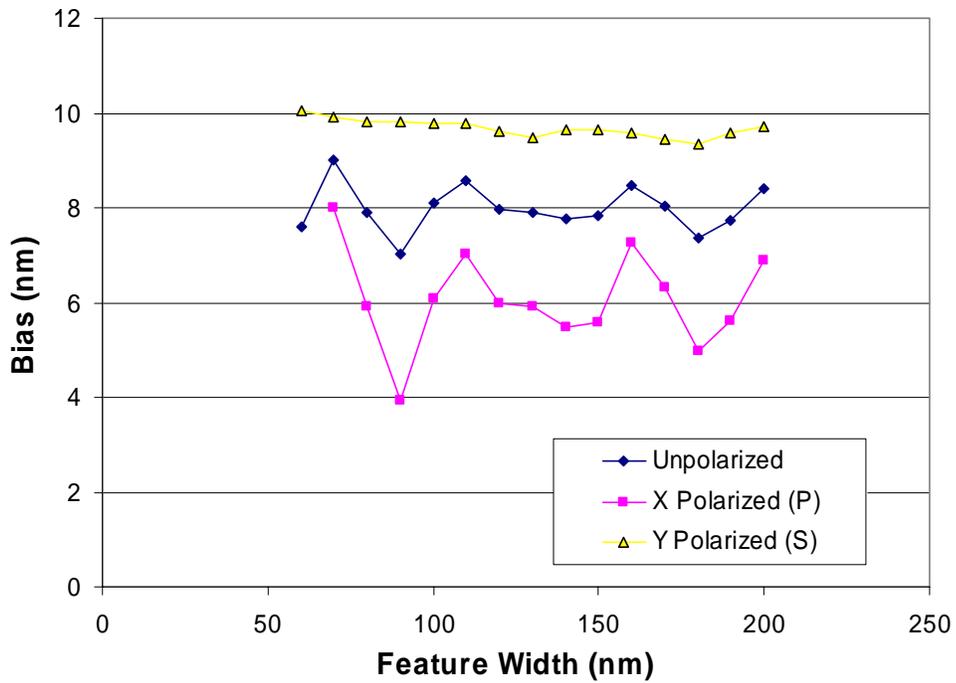


Figure 9: The bias applied to the Kirchhoff mask to match the prediction from the full Maxwell equations for a y-oriented isolated line for unpolarized, x-polarized, and y-polarized illumination.

4 NORMALLY INCIDENT APPROXIMATION

One additional approximation that is commonly used when calculating the electromagnetic transmittance of the mask is to assume that the results are independent of the angle of the incident illumination. This is called the normally incident approximation because the diffraction pattern calculated for normally incident light is used for calculations with off-axis illumination. This approximation is almost always used with the Kirchhoff approximation because the mask is assumed to be infinitely thin, so any “shadowing effects” due to the interaction between the off-axis illumination and the finite height of the topography on the mask can be ignored. In this section, we solve the Maxwell equations for a variety of different incident illumination angles at the mask. Each of these calculations will generate a separate diffraction pattern for the mask, and then when we calculate the aerial image, the appropriate diffraction pattern will be used for each point on the illuminator. In PROLITH, this is called the “range of angles” mode of calculation. For the partially coherent illuminator used in this study ($\sigma = 0.5$), the range of incident angles at the 4x mask is around $\pm 5.7^\circ$ from the normal. A comparison of the normally incident approximation with the “range of angles” approach is shown in Figure 10, where results are presented for the isolated space calculations from the previous section. As shown in the figure, the “range of angles” approach does not have a dramatic impact on the results, other than to dampen some of the oscillations in the results for larger space widths.

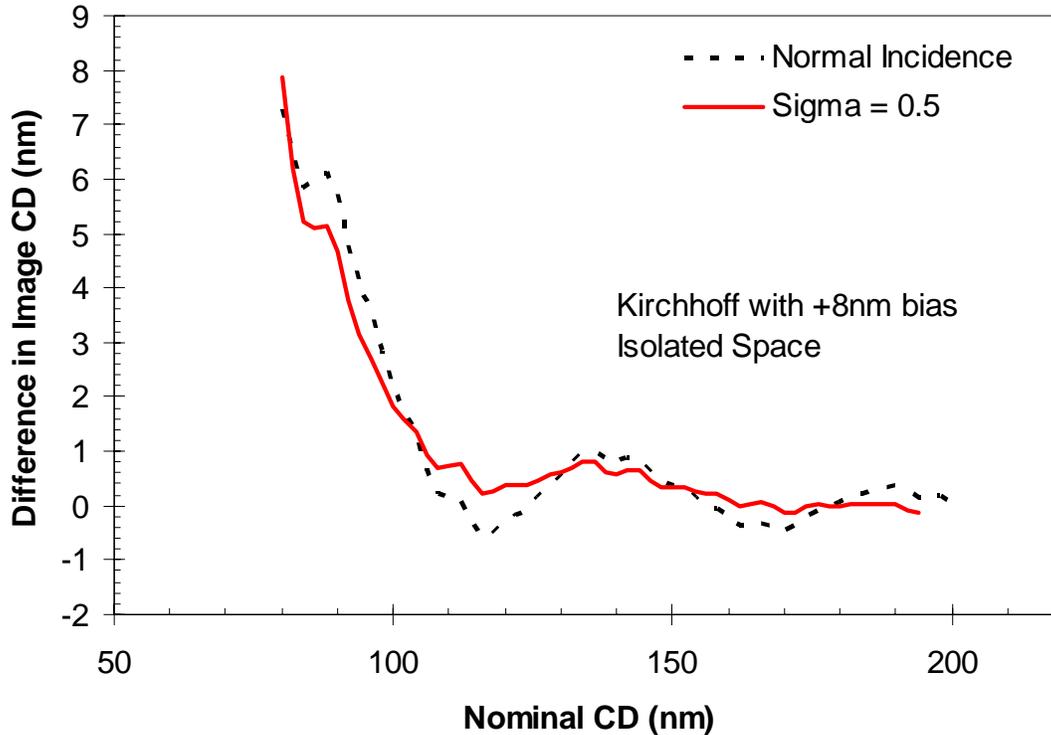


Figure 10: Comparison of the normally incident approximation with the “range of angles” calculation approach for the isolated space (again, showing the difference between the Maxwell and Kirchhoff aerial images).

5 SUMMARY AND CONCLUSIONS

The difference between the aerial images calculated with the Kirchhoff approximation and the full Maxwell equations can be approximated by a bias for chrome on glass reticles. We found that the bias was about 7 to 8nm for 100nm thick chrome, with a range of bias values of about 3.3nm for different line widths and pitches. For 70nm thick chrome, the bias was reduced to 5 to 7nm. The calculated bias could be used to generate a more accurate Kirchhoff model for the mask where the average bias was globally applied to the chrome features. Good results with the “Kirchhoff plus global bias” model were obtained for isolated lines and process window calculations, but significant deviations were found for small, isolated spaces. The bias was found to be polarization-dependent, with a 10nm bias for p-polarization and a 6nm average bias for s-polarization.

In general, the bias-type of Kirchhoff model may work well in applications where solution to the full Maxwell equations is intractable. The accuracy of this approach can be estimated by the range of bias values that we calculated (3.3nm), so this approach may be insufficient in applications where we want to calculate the desired mask feature size to a tolerance lower than 3nm (wafer scale). In addition, the approach seems to be especially limited for the case of small isolated spaces. Further work is needed in this area in order to better understand the interactions between darkfield features that are comparable to the wavelength of light. One area that will likely require special attention is the contact layer, where four edges are in close proximity instead of the two edges that form an isolated space. Further work is also required on the impact of chrome thickness, more realistic chrome film stack optical properties, other optical scanner settings, and attenuated phase-shift materials.