22 nm Lithography Using Near Field X-rays

Antony J. Bourdillon\textsuperscript{a}, Gwyn P Williams\textsuperscript{b}, Yuli Vladimirsky\textsuperscript{a} and Chris B Boothroyd\textsuperscript{c}
\textsuperscript{a}UhrlMasc Inc., P.O. Box 700001, San Jose, CA 95170; \textsuperscript{b}Jefferson National Accelerator Facility, 12000 Jefferson Avenue, Newport News, VA 23606; \textsuperscript{c}IMRE, 3 Research Link, Singapore 117602

\section*{ABSTRACT}

By using the Near Field in Proximity X-ray Lithography (PXL), the technique is demonstrated that extends beyond a resolution of 25 nm print feature size with 2:1 pitch to line width. "Demagnification by bias" of clear mask features is positively used in Fresnel diffraction together with multiple exposures of sharp peaks. Exposures are performed without lenses or mirrors between mask and wafer, and the "demagnification" is achieved in the selectable range 1X to 9X. Pitch is kept small by multiple, stepped exposures of sharp, intense, image peaks followed by single development. Low pitch nested lines are demonstrated. The optical field is kept compact at the mask. Since the mask-wafer gap scales as the square of the mask feature size, mask feature sizes and mask-wafer gaps are comparatively large. Because the features are themselves larger, the masks are more easily manufactured. Meanwhile Exposure times, for development levels high on sharp peaks, are short, and there are further benefits including defect reduction, virtual elimination of sidebands etc. A Critical Condition has been identified which is typically used for the highest resolution. Many devices, including batches of microprocessors, have been demonstrated previously by traditional 1X PXL which is the most mature of the Next Generation Lithographies and which is now further extended. For two-dimensional Near field patterning, temporal and spatial incoherence at the Critical Condition are used to show, not only that peculiarities in the aerial pattern, such as "ripple" and "bright spots", can be virtually eliminated, but also that there is an optimum demagnification, around 3X, in the Fresnel diffraction, where the contrast is highest. At this demagnification, patterns of various dimensions can be printed using various and appropriate demagnifications.

\textbf{Keywords:} PXL, Near field, Demagnification by Bias, Ultra High Resolution Lithography

\section*{1. INTRODUCTION}

Proximity X-ray Lithography (PXL) is on the International Technology Roadmap for Semiconductors, a contender for application in Next Generation Lithography. PXL is extensible to 15 nm print resolution using demagnification by bias, without lenses or mirrors\textsuperscript{[1]}. Prints with feature size down to 25 nm have been demonstrated\textsuperscript{[1][2][3][4]}. Typical 1-2kV broad band incident beam energies (1.2-0.6 nm wavelengths) were used, and demagnifications down to 6X were obtained. The mask feature sizes and mask-wafer gaps were large, about 20 \mu m. The technique, which has many novel features and is sometimes called Ultra-High Resolution Lithography\textsuperscript{[5]} (UHRL), employs Fresnel diffraction positively near a "Critical Condition" (see below and ref. [1]) and results in demagnification by bias. While the print is smaller than the corresponding clear mask feature, the demagnification is not generally uniform because the bias is more or less constant around the edge of the image so that the bias is subtracted from the size of the mask aperture when the print is developed. Meanwhile, the optical field is kept compact as in traditional 1X PXL. The demagnification has considerable importance in a wide range of applications including integrated circuit manufacture\textsuperscript{[6]} and manufacture of micro electromechanical systems. The importance is due additionally to the relative physical simplicity of proximity methods over competing projection systems, as also to relatively high throughputs and simple equipment requirements.
Thus PXL, which is the only one of next generation lithographies\cite{7,8,9,10,11} that is properly demonstrated, is extendible beyond previously supposed limits. Typically synchrotron radiation is used as the radiation source and this is naturally collimated by relativity so that the penumbra is controlled to within one nanometer. To provide uniform illumination, the X-ray beam is typically scanned, off an oscillating, grazing-incidence mirror, across the mask-wafer system. Exposures of full fields, typically $50 \times 30$ mm, are made in a time about 1 s. The mask is held stationary while the wafer is stepped and aligned between exposures on different fields.

Previously, we have reported the results of simulations\cite{2,12} which have shown the effects of (a) varying the mask-wafer gap about the critical condition; of (b) the residual effects of small absorber transmission; (c) the distortion to be observed in non-symmetric, two-dimensional features, of (d) high frequency “ripple” and “bright spots” parallel to longer dimensions, of (e) mask shaping to virtually eliminate the “ripple” and “bright spots”, and of (f) combined doses due to multiple exposures used to reduce pitch:line width\cite{3}. The simulations have been performed for both monochromatic incident radiation and for the wide band of wavelengths typically used in PXL. The simulations generally correspond to earlier experimental demonstrations\cite{1,2,3,4}.

The choice of magnification has previously been left open to compromises involving mask feature size, gap width, printed feature size etc. However there are indeed preferential regimes determined by Fresnel diffraction. Further theoretical understanding is needed of the effects of temporal and spatial coherence on the aerial images produced by the Fresnel diffraction. The analysis can be used to identify demagnification regimes that provide enhanced Critical Dimension (CD) control. The analysis can be further used to illustrate limitations in traditional 1X PXL and to optimise mask patterns for ideal print shapes.

We have previously listed other attempts at fine printing by methods equivalent to 1X contact printing\cite{3}. There have been further incremental attempts to enhance PXL some of which can, in principle, be employed in Near Field to extend further to 15 nm printed features. Some attempts have been uncompetitive for high resolution\cite{13} and involve unnecessarily complicated procedures in 1X mask making and exposure for phase shifting. More significant is the attempt to use shorter wavelength X-rays, about 0.4 nm, with diamond-like mask substrates\cite{14,15} and modified resists. It is clear that such incremental resolution enhancements can be improved by a large factor of 3 when adapted to Near Field. This is partly due to the larger mask-wafer gaps that can be employed, since the gap scales as the square of clear mask feature size. We review the status of Near field and proceed to apply theoretical models for understanding temporal and spatial coherence in PXL. This will affect the optimum demagnification that is conducive to high resolution, high contrast and good CD control.

Figure 1. Simulation of a Fresnel diffracted current, with wavelength $\lambda=0.8$ nm, passing through a slit of width 150 nm. The picture width is 1536 nm and the height 40 µm. The Critical Condition lies at a gap of 10 µm. Notice here the sharp peak and adjacent shoulders as in figure 7. (Courtesy Institute of Physics Publishing, ref 2).
2. SWEET SPOT

2.1 Currents from a clear mask feature

The most striking graphic demonstration of the necessity for demagnifying clear mask features, when high resolution is needed in PXL, comes from simulations of currents transmitted by clear mask features. Further details showing the effects of residual absorber transmission are given elsewhere[2] and are supplemented by the analysis of temporal coherence described below.

Fig. 1 is a universal dimensionless image showing the distribution of current below a clear mask feature[2]. If the clear mask feature has a width of 150 nm and the wavelength is 0.8 nm, then the vertical range shown is about 40 µm. With 0.4 nm wavelength, the range is 80 µm. It can be seen that there is a long "sweet spot" about one quarter down where the current is bright and narrow. This corresponds to the region near the Critical Condition (CC) defined below. Here, by controlled development, the highest resolution can be obtained since the single peak in the aerial image is most narrow. Further theoretical description of this optimum is given below.

Figure 2. Schematic diagram shows dense lines developed from multiple exposures of individually demagnified clear mask features[3] (courtesy Institute of Physics Publishing [3]).

Figure 3a. 61 nm lines printed in SAL 605 negative resist using a mask with 180 nm nominal line/space features (360 nm period). Left mask (152 nm); right print (61 nm). Courtesy Institute of Physics Publishing [2].
2.2 Schematic for demagnification by bias with multiple exposures
The peak is illustrated in the schematic figure 2 where the sharp peak at CC allows rapid exposure above a broad background. The ratio of pitch to line-width is reduced to 2:1, i.e. half pitch, by the method of multiple exposures with single development. The process is rapid because the stepping is typically blind and peaks are intense.

3. EXPERIMENTAL DEMONSTRATIONS

The following demonstrations were performed on a modified Suss XRS-200/2M X-ray stepper, owned by the Center for Nanotechnology and located on beamline ES-4 on the Alladin storage ring at the University of Wisconsin-Madison Synchrotron Radiation Center. A combination of optimized lithographic process and exposure conditions allowed the formation of lines down to 25 nm at 15-30 μm gaps for both negative (SAL605) and positive (UV-3, APEX-E) resists.[16]

3.1 Line prints using demagnification by bias
To demonstrate the fact of demagnification by bias, various line prints have been recorded using a synchrotron radiation source with a typical broad band radiation dose centered about 0.8 nm wavelength.[2, 3]. Some examples are shown in figure 3. SAL 605 resist was used to form the figures 3a and 3b and
2.2 Nested features
Figure 4a is a simulation of low pitch 45 nm lines for printing after double exposure. This simulation corresponds to the print in figure 4b and follows the schematic diagram in figure 2, except that the latter is for triple exposures closer to the Critical Condition.

4. TEMPORAL AND SPATIAL COHERENCE

4.1 The Critical Condition in one dimension
Consider firstly the Critical Condition as it applies in the imaging and printing of one dimensional features such as parallel lines. Figure 5 shows a schematic exposure system. Parallel rays of radiation pass through a clear mask feature and form a Fresnel pattern, or demagnified image, at a distance \( G \) below the mask feature. One ray is axial. A second ray suffers a phase lag which depends on the distance, \( s \), from the center line of the clear mask feature, i.e. the phase lag suffered by this ray at the resist depends on \( 2\pi s/\lambda \), where the wavelength is \( \lambda \). The amplitude at the wafer depends on the vectorial sum of the amplitudes of all rays passing through the clear mask feature.

Consider the dimensionless spatial coordinate, \( v \), defined:

\[
v = s \sqrt{\frac{2}{G \lambda}} = \sqrt{2 \hat{N}_P}
\]
where $s$ is a distance measured from the axis of the clear mask feature/clear mask feature in its plane (fig. 5); $G$ is the width of the mask/wafer gap; $\lambda$ is the wavelength of the radiation used, and $N$ is the number of Fresnel half zones across the clear mask feature.

If $\Delta s$ is the clear mask feature width and $\Delta v$ is the dimensionless spatial co-ordinate corresponding to $\Delta s$ at a given $G$ and $\lambda$, then:

$$\Delta v = \Delta s \sqrt{\frac{2}{G\lambda}}$$

(2)

$\Delta v$ can be called the dimensionless slit width. The vectorial addition of the amplitudes and phases of rays passing through the clear mask feature, and interfering constructively at the plane of the wafer, can be summed over all transmitted rays. The amplitudes are represented mathematically with well-known Fresnel integrals or can be summed graphically with Cornu's spiral\textsuperscript{17}, i.e. the vibration curve, shown in figure 6. The amplitude of the Fresnel pattern at a point on the wafer can be found by summing amplitudes and phases of corresponding rays\textsuperscript{12}. The vectorial summation is found graphically by connecting two points on Cornu's spiral. The Critical Condition occurs when the width of a transmitting mask feature, $\Delta s$, is related to the mask/wafer gap $G$ and X-ray wavelength $\lambda$ by the equation:

$$\frac{\Delta s}{\sqrt{\lambda G}} = 1.7$$

(3)

At the Critical Condition, summing over rays for which $\Delta s/2 > s > -\Delta s/2$, the amplitude at the wafer, on axis, is the longest vector which joins the extremities of the Cornu spiral. This vector is represented by $A$ in figure 6. The square on this vector is the maximum intensity. Off-axis at the wafer, asymmetric vectors on the spiral are used to represent vectorial sums of ray amplitudes.

The Critical Condition has a clear theoretical meaning with practical implications. However there is wide latitude in setting the conditions - through the selection of wavelength, mask feature size and gap - which are not practically critical. Cornu's spiral applies to the imaging of a long slit with monochromatic radiation; but we extend it, as described below, to applications using broad band illumination having temporal incoherence.

Notice meanwhile that, by holding the dimensionless $\Delta v$ in equation 2 constant, the gap, $G$, changes with the square of the slit width, $\Delta s$, and that $G$ depends inversely on $\lambda$. This is an important reason for demagnifying and a new reason for using shorter wavelengths\textsuperscript{18}.

4.2 The Critical Condition with Broad Band Illumination

Secondly, with broad band illumination, simulations require independent integration for both temporal and spatial coherence. When the wavelength, $\lambda$, is not monochromatic but is spread over a range $\delta \lambda$, it follows from equation 1 that $\delta \lambda/\lambda = -\delta v/2\lambda$. The Fresnel integrals, represented in Cornu's spiral, can be averaged as in the greyscale
Figure 6. Cornu's spiral [1], (full line) representing relative phases of rays transmitted by a slit onto a wafer. Vector A represents the sum of phases of transmitted rays striking the wafer, at the Critical Condition, opposite the slit axis. Other spirals are formed by averaging over spreads of dimensionless spatial co-ordinates: typically when $\delta v = \pm 0.2$ (dashed curve), and also when $\delta v = \pm 0.4$ (dotted curve) and $\delta v = \pm 0.6$ (dash-dot curve). These spreads correspond to broadband illumination and are used to show the effect of temporal incoherence.

Aerial images at the Critical Condition with various beam spreads: $0, \pm 20\%, \pm 40\%, \pm 60\%$ having temporal incoherence derived from Cornu's spiral.

Figure 7. Aerial images at the wafer corresponding to monochromatic $\delta v = 0$, typical $\delta v = \pm 0.2$, and also $\delta v = \pm 0.4$ and $\delta v = \pm 0.6$. Notice the inflection points at $v = 0.36$ where the slope is steeper, and at $v = 0.88$ where the slope is shallower. The slit edges lie geometrically opposite the dimensionless points $v = +1.2$ and $v = -1.2$ on the wafer.
curves in fig 6. The averaging procedure, by including vectorial additions of rays, accounts for the temporal incoherence. Specifically, with broad band illumination, such that the bandwidth corresponds to \( \delta v = \pm 0.2 \), the ray phases can be represented by the tangents on the corresponding curve in fig. 6. Corresponding phases at a wafer, for bandwidth ranges \( \delta v = \pm 0.4 \), and, for bandwidth ranges \( \delta v = \pm 0.6 \), are represented on corresponding curves. Summing rays over the dimensionless slit width provides corresponding aerial profiles in figure 7.

The new significance of these plots lies in the inflection points where the graphs at various \( \delta v \) cross over. The inflection points closest to the axis (\( v = 0 \)) occur at \( v = \pm 0.36 \) where the slope is steeper, and at \( v = \pm 0.88 \) where the slope is shallower. A steep slope is valuable for critical dimension (CD) control in printing. It is interesting, furthermore, that the inflection points occur on the same ordinate scale as the incident intensity (level 2) at the clear mask feature.

When \( v = 0.88 \) and \( \delta v = 0 \), the printing definition is extremely poor. This is close to the case used in traditional 1X proximity printing [8] where \( v = 1.2 = \Delta v/2 \) at the Critical Condition. Contrast is then further degraded with broad band as can be seen in figure 7 and additionally degraded by sidebands. Compared with the aerial image shown in the fig 7, sidebands at \( v > \Delta v/2 \), are increased by residual transmission from the mask absorber[11] and further increased in masks containing periodic structures. The fuzzy contrast, sometimes difficult to predict, that was traditionally used in 1X PXL is completely avoided in Near Field where rapid exposures are made, high on the aerial image instead of at the base.

However, the inflection points demonstrate a further optimisation of demagnification in Near Field. Previously, we have proposed[1, 2, 3, 4, 11] that demagnification by bias is selectable, depending on chosen development level and that optimisation depends on various factors including fabricability of masks and multiple exposure systems. It now appears that an additional feature needs to be taken into account and this may often dominate: where CD control is critical, contrast is highest at the inflection point, i.e. at a demagnification of 3X or 1/0.36. The inflection points occur at the dose level of incident radiation at the mask.

4.3 The Critical Condition for Two Dimensional Features

Consider thirdly the critical condition for two dimensional, asymmetric, clear mask features. Since the CC depends on clear mask feature size, \( \Delta s \), it cannot be maintained for two different dimensions at one time. CC then refers, by our convention[11], to the smaller dimension where the print resolution is the more critical.

Consider in consequence, features produced in the less-critical, longitudinal direction of a rectangular mask. To understand the independent effects of mask feature shape, wavelength, and gap for the 2D images, idealized simulations were performed for the intensities below the mask. A multislice method written in the Semper[18] image processing program was used. The program allows the Fresnel diffraction from arbitrarily shaped masks in one or two dimensions to be calculated at any distance from the mask. Examples of imaging with monochromatic radiation and with broad band illumination, both at CC and away from CC, are shown in figure 8[1, 11]. This shows the leveling of both "Ripple" and "Bright spots" in the pattern when broad beam is used. The leveling is a consequence of temporal incoherence.

To reduce undesirable effects of Ripple and Bright spots further, the spatial incoherence in the aerial pattern can be increased by an inverted variant of optical proximity correction. An example is shown in figure 9. The blurred 45 degree indent, at the ends of the rectangular pattern, results in a reduction of Ripple and Bright spots to insignificance when broad beam is additionally used, typically, as before.

Near Field can be used equally to print patterns of arbitrary shape, though optimisation occurs through a correct use of temporal and spatial incoherence. As an example, figure 10 shows a flag shaped pattern and the corresponding aerial image when the gap is set critical (\( \Delta v = 2.4 \)) for the narrow part of the flag pole. Blurred indents are employed at the ends of the pole and corners of the flag are also blurred and the pole can be conveniently printed 3X. Notice that the
width of the flag then has $\Delta \nu > 2.4$ and it is printed, with a smaller proportionate bias, close to 1X. The effects of various indents are shown in the figure. The results show how, by employing temporal and spatial incoherence, any shape can be printed with high resolution using Near Field X-rays. Notice that in the printing of complex shapes, the requirement for half pitch is relaxed.

Figure 8. At top is a two-dimensional rectangular mask slit of size 150 x 600 nm$^2$ with intensity scale white=1; at centre a simulated image due to 0.8 nm X-rays transmitted at a distance 9.8 $\mu$m through the clear mask feature behind the otherwise opaque mask, i.e. for critical $\Delta \nu = 2.4$ and intensity scale white =2.7; at center bottom corresponding simulation with broadband 0.62<\lambda<1.28 nm; at right corresponding simulations at distance 30 $\mu$m behind the opaque mask when $\Delta \nu = 1.4$; at left corresponding central profiles of adjacent images. Notice that the “Ripple” and “Bright Spots” most evident in the central image, are reduced with broadbeam. (Courtesy Institute of Physics Publishing [17])

Figure 9. Simulation of two-dimensional rectangular mask as in figure 8 but with indents: V-shaped 30°, 45°, 60° and blurred 45° and corresponding simulated images at the Critical Condition with broadbeam. Profiles are also shown. (Courtesy Institute of Physics [17])
4.4 Other Effects of Coherence

A further consequence of the diffraction is a reduction in printing errors due to mask defects when compared with 1X PXL. Inhomogeneities in mask patterns are smoothed by Fresnel diffraction at the print surface. Examples of such reduction is evident in figure 3a, and elsewhere [1,2,11].

Figure 10 Flag pole pattern with pole 150 x 600 nm², as in figure 8 plus flag 450 x 300 nm² and simulated images at the Critical Condition and a broadband of wavelengths 0.62<λ<1.24 nm. Effects of spatial incoherence are simulated by the use of various indents. At right are shown profiles of adjacent (fourth row) simulated image both vertically down the flag pole and horizontally across the flag. A development level can be chosen to print the pole 3X with the broader flag close to 1X.

4.5 Extensibility to 15 nm and further

Various facts lead us to revise our previous estimates [2] of the extensibility of Near Field methods. Recent studies that include treatments of blur [14,15] when combined with a careful study of present results, especially the 25 nm print in figure 3, together with equivalent earlier, but overlooked, reports [19][20] leads to a conclusion that previous estimates of photoelectric blur are exaggerated. It was previously supposed that the minimum feature size that can be printed by PXL results from the minimum due to competing effects of diffraction and photoelectric blur and that these are both dependent on the energy of photons used. The results reported would have been impossible if the previous estimates of blur are valid. In particular, the blur in PMMA produces negligible broadening at 20 nm print resolution when used with 0.8 nm X-rays incident at the mask and the printed feature size is dominated by diffraction. This is especially the case when demagnification by bias is deliberately used as in Near Field. Though primary photoelectric blurring is broad, the printed contrast in fact depends on the Auger electrons, and their range is not only much smaller than that of the primaries, but moreover does not change with increasing incident photon energy. This range only changes significantly with resist composition which can be carefully selected for optimization.

Blurring due to the range of Auger electrons is small down to 15 nm. The diffraction limits, shown in figure 11, depend on the experimental k value (k = smallest print feature size(λG)¹/² ~0.15) demonstrated using demagnification by bias on isolated lines. At dimensions for 15 nm prints, the gap approaches an experimental limit about 5 μm, and the printing is facilitated if a restriction to half pitch is relaxed. Since PMMA (polymethyl methacrylate) is less sensitive than chemically amplified resists such as SAL605, optimization will benefit from a selection of new resists beyond those currently used in both PXL and 248 nm optical lithography. Some of the requirements in resist development are common to all next generation lithographies and mutual transference is anticipated for extension to 15 nm prints. For print features down to 35 nm, the normal resists are sufficient and readily available.
5. CONCLUSION

Spatial and temporal incoherence in Near Field X-ray Lithography can, in principle, be used to identify demagnification regimes that provide enhanced Critical Dimension (CD) control. A demagnification about 3X allows optimum discrimination in the printing of edge features. This is due to an inflexion point in Fresnel diffraction when the wafer is placed at the Critical Condition. By contrast, the analysis shows that exposures with traditional 1X edges often lie near to plateaux - especially when broadband illumination is used - where intensity discrimination, employed critically in the printing, is minimal. Moreover, Near Field X-ray Lithography has further important advantages in both the control of printing and in increased wafer throughput. These are in addition to many other advantages that Near Field X-ray Lithography accrues, such as increased mask feature size and gap width, in addition to the outstanding feature of extensibility beyond 20 nm. The method opens the way to the manufacture of micromachines and integrated circuits of such small dimensions, whether using modern compact synchrotron light sources \cite{21} or, at a slower rate, with point sources \cite{22}.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{resolution_diagram.png}
\caption{Revised resolution limits showing competition between photon energy dependent Fresnel diffraction, which is a function of gap in UHRL, and typical range of Auger electrons in photoresist. The primaries cause background exposure that is not significant for print contrast.}
\end{figure}

\begin{center}
REFERENCES
\end{center}

\begin{enumerate}
\item Kong J R, Vladimirsky Y and Quinn L, \textit{Proc. MNE 2000}, Jena Sep 18-21, Germany
\item Y Vladimirsky and AJ Bourdillon \textit{US patent Nr. 6,383,698} (2002)
\item \textit{Solid State Technology}, February 2000, News Item, pp18-23
\end{enumerate}


