# Fast Inverse Lithography Technology

Daniel S. Abrams, Linyong Pang
Luminescent Technologies, Inc., 650 Castro Street, Suite 220, Mountain View, CA 94041, U.S.A.

Email: leo@luminescent.com

## ABSTRACT

Many RET technologies, such as rule and model based OPC, the use of sub-resolution assist features, and various PSM methodologies, can be thought of as heuristics employed in an attempt to design improved photomasks. Unfortunately, these traditional approaches are running into severe difficulties at advanced technology nodes (90nm and beyond). We discuss how one can find the optimal photomask by rigorously solving the lithography inverse problem. The design of the optimal mask takes into consideration not only pattern fidelity under nominal conditions, but also the size of the process window and the constraints and costs of mask manufacturing. By formulating the problem in a rigorous mathematical framework, we find highly optimal solutions which do not arise from traditional ad hoc approaches. The resulting masks often provide substantially improved depth-of-focus and exposure latitude, enabling geometries that may be otherwise unattainable.

**Keywords**: Inverse lithography, Inverse problem, lithography simulation, optimization, yield, process window, ILT, OPC, PSM

## INTRODUCTION

Increasingly, for semiconductor manufacturers moving to advanced nodes – 90nm, 65, 45, and below – the greatest challenge is lithography. This is because lithography is fundamentally constrained by basic principles of optical physics. At 65 nm, a line is less than a third of the effective wavelength; as the industry moves forward, optical diffraction and interference are becoming fundamental obstacles, not just second order effects.

It has long been known that the best lithography that is theoretically possible can be achieved by considering the design of photomasks as an inverse problem -- and then solving the inverse problem to find the optimal photomask for a given process, using a rigorous mathematical approach. Inverse Lithography Technology (ILT) has been explored for many years, starting with the pioneering work of B. E. A. Saleh and his students in the 1980s [1,2]. Although these early approaches to ILT often resulted in superb lithography, they were generally impractical in a production environment. Run-times were many orders of magnitude too slow, and the resulting masks were often too complex to manufacture.

In this paper, we discuss the first ILT approach that can rapidly solve for the optimal manufacturable photomask design and is suitable for use in a production environment.

Stated simply, the goal of inverse lithography is to find the optimal mask pattern to print a given target pattern, given the known transformation from mask to wafer. It is a rigorous mathematical approach to the design of photomasks for optimal lithography. By its nature, ILT is a methodology that is not limited to simple heuristic modifications of the target pattern; in other words, it explores regions of solution space that are very different from the original pattern. Perhaps most importantly, it is inherently pattern independent.

There are also some common myths associated with ILT. One of these is that ILT always results in a unique global optimum. Although this may be true for some algorithms, many inverse lithography approaches are based on local search heuristics which find a good -- but not necessarily global -- optimum. Moreover, because the inverse problem is mathematically ill-posed, there are many situations in which the global optimum is not unique. It is then up to the algorithm (or algorithm designer) to determine which of many equally good solutions is desired.

Another ILT myth is that such methods must be image or pixel based. The algorithm developed at Luminescent and discussed below is pixel based, but not every approach to inverse lithography necessarily is.

Finally, there is a common perception that inverse lithography algorithms must represent a solution in closed form, or are in some sense direct solutions to the problem. In point of fact, however, most ILT methods that have been developed are iterative at some level. This should not be surprising, because many (if not most) modern mathematical solvers involve iterations, and, done properly, such methods can be very efficient.

The key distinctive feature of ILT is the lack of pattern-dependent heuristics, and the ability to broadly explore wide areas of solution space. ILT algorithms routinely lead to mask patterns which are unanticipated. Consider the problem of placement of subresolution assist features (SRAFs). These are mask features that do not print on the wafer, are detached from the edges of the semiconductor pattern, and yet manipulate the light reaching the wafer so as to accentuate the wafer image. In the past, these were placed empirically, with great care, and frozen in place during the computation of the rest of the mask. ILT, on the other hand, can determine optimal SRAFs simultaneously with the rest of the mask.

The benefits of such an approach are many. The absence of segmentation scripts is a significant advantage. In addition, ILT is generally not susceptible to errors resulting from unanticipated patterns (because it is inherently pattern independent). By finding optimal mask patterns, superior pattern fidelity, larger process windows, and improved yields can all be achieved.

## HISTORICAL BACKGROUND

ILT was first developed by B. E. A. Saleh and others at the University of Wisconsin-Madison. For example, in 1981 Saleh and Sayegh [1] found optimized photomasks by a variation on simulated annealing, or “pixel flipping”— that is, they started with an initial guess, randomly flipped individual pixels, accepted changes that improved the quality of the solution (and rejected changes that degraded the solution), and repeated this process until the system converged on an optimal photomask. A few years later, Saleh and Nashold [2] described an algorithm using a sequence of projection operators in order to find a band limited function (corresponding to a continuous-tone or gray scale mask) which would optimally result in the desired image. Later, the same authors used a similar approach to find complex valued functions that corresponded to continuous tone phase masks.

In the early ‘90s, Yong Liu and Avideh Zakhor (at Berkeley) wrote a series of papers [3,4] describing various approaches to ILT. In one case, they used branch and bound and the simplex method. In another, they used what they called a “bacteria” algorithm in order to satisfy mask constraints.

In 2001, Rosenbluth et. al. (at IBM) described an ILT algorithm that analyzed diffraction orders in order to jointly optimize the photomask and the stepper illumination [5]. This approach solved first for an optimal wavefront and then in a second step tried to find the optimal photomask to generate the same diffraction pattern.

Although the researchers described above made significant contributions to the development of ILT, there are many others who have also made important contributions: for example, the work done by Wang. et. al. [6] (at Stanford, and later Numerical Technologies), and the OPERA program, by Oh et. al. (at Wonkwang University in South Korea) [7]. Most recently, Fuhner and Erdmann of the Fraunhofer Institute developed ILT using genetic algorithms [8]. The above summary is merely intended as a survey, and is certainly not one hundred percent inclusive.

As mentioned previously, these early approaches to ILT usually resulted in superb lithography. The patterns found often resulted in superior accuracy, improved process windows, and better pattern fidelity. However, they were generally impractical in a production environment, due to intractable run-times and/or unmanufacturable masks. For example, finding the optimal continuous tone or grey-scale mask is an easier mathematical problem than finding an optimal binary mask. However, only a binary mask is practical for current production.

## THEORETICAL FRAMEWORK

In order to formalize the problem, we require some definitions:

Mask function: $\psi$
Target pattern: $\Phi$
Forward operator: $f$
Wafer pattern: $\omega$

The forward operator may take into account all of the elements of the transformation from mask to wafer: for example, the electromagnetics of the 3D mask, the optics of illumination and the lens, the behavior of the photoresist, the dose and focus conditions, aberrations, etc. Thus

$$\omega = f(\psi) \qquad (1)$$

and we seek to find

$$\psi^* = f^{-1}(\Phi) \qquad (2)$$

where $\psi^*$ is the optimal mask function.

However, the problem thus stated is ill-posed; because the forward operator $f$ is many-to-one (that is, many different masks will yield identical on-wafer results), the function has no well-defined inverse. Moreover, for typical target patterns $\Phi$ (e.g., a drawn layout with Manhattan geometry and sharp corners), there does not exist any mask function $\psi$ for which $\Phi = f(\psi)$.

These issues are addressed by recasting the inverse problem as an optimization problem. We define a merit function, also called a cost function, energy function, or Hamiltonian (by analogy to quantum mechanics), and label it $H(\psi)$. This function is indicative of the quality of the solution, or the “goodness” of the mask. A simple example would be

$$H = \iint \left| f(\psi) - \Phi \right|$$

In other words, this Hamiltonian is the absolute value of the difference between the wafer image and the target pattern, integrated over the area of the mask. In practice, a number of additional elements may be included in the Hamiltonian. For example, the wafer pattern at various conditions throughout the process window (i.e., over or under exposed and/or plus/minus focus), the NILS of the image, the robustness against MEEF, or other factors as deemed appropriate. The actual functional form may be different from the form as described above as well. Elements that are not directly related to lithography may be included; for example, simple masks may be preferred over complex masks, and terms to this effect may be included in the Hamiltonian as well. What is essential is that the Hamiltonian is a functional of the mask function, and that minimizing said Hamiltonian allows us to find the optimal mask, according to the criteria we have chosen.

Another important aspect of the minimization problem comes in the form of constraints. A variety of constraints are imposed by the realities of mask manufacturing; for example, two disjoint chrome regions must be separated by a minimum distance, and a chrome line must have a minimum thickness. We address these constraints by defining a sub-space of the full Hilbert space of mask functions, and restricting our solution to this sub-space.

Using modern numerical methods and the latest processors, it is now possible to quickly solve the resulting minimization problem. Our implementation divides a large photomask into small regions called "work units". These are distributed to a cluster of compute nodes, which can then process many work units in parallel. The solutions are then stitched together to form a complete mask. A large number of real full-chip designs have been processed this way through our software. The results can thereby be obtained quite quickly. In one recent example, a large 4cm$^2$ die (wafer scale) was processed overnight.

## RESULTS AND EXAMPLES

The following example shows a mask pattern for a contact hole. The CD is 110nm, the numerical aperture is 0.78, and the mask is a continuous tone attenuated phase mask.

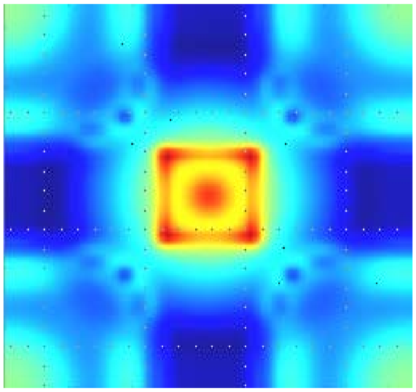

**Figure 1. Continuous tone mask. This "dark field"Deep blue is high density and red is low density.**

The same contact can be printed with an ordinary binary attenuated mask:

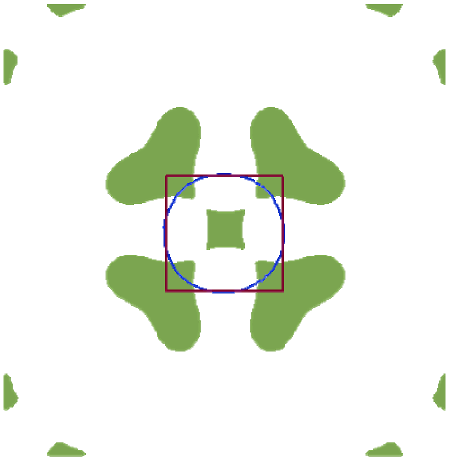

**Figure 2. Binary mask. The square is the target pattern, the circle is the forward image, and the remain figures are on mask.**

In Figure 2, we see the outline of the contours for the mask pattern in green; the square which represents the original target pattern, and a circle which represents the resulting wafer image. It is quite remarkable that the combination of a single small central region, with four distinct surrounding lobes, should optimally print the contact pattern. Such remarkable patterns illustrate the power of ILT, which finds solutions that are often unexpected. Notice also the assist features found further from the contact, around the perimeter of the image.
If the mask is constrained to Manhattan geometries, to assist in manufacturing, we find the following design:

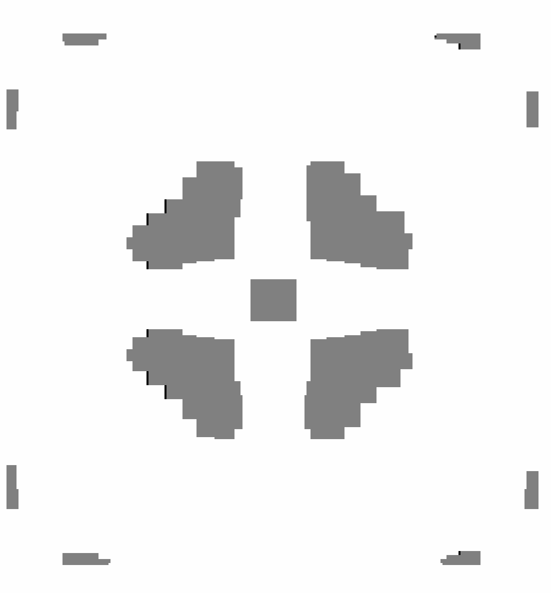

**Figure 3. Manhattan mask.**

We also found the optimal photomask for the following pattern:

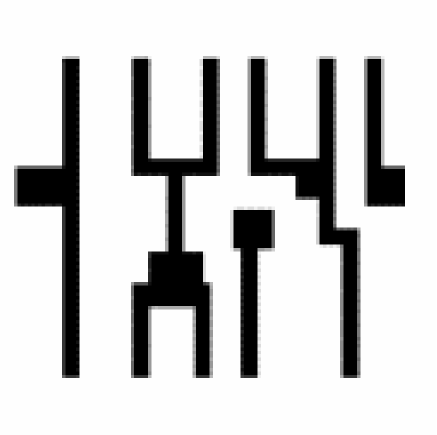

**Figure 4. XOR Pattern.**

The linewidth is 45nm, the numerical aperture is 1.2 (immersion), and we used annular illumination (.65::.95)

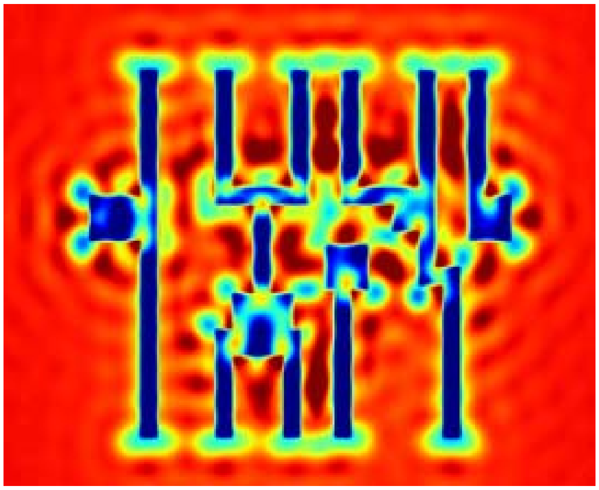

**Figure 5. Continuous tone XOR mask pattern.**

The continuous tone mask in figure 5 above can be used to print the XOR pattern; if we constrain the optimization to a binary mask, we find the following solutions:

**Figure 6a,b,c. Binary XOR mask patterns.**

In Figure 6a, on the left, we see the optimal continuous tone mask. In 6b, the middle pattern, we have the optimal mask pattern with Manhattan constraints. In 6c, we see the solution found without assist features. The left most mask will provide the best lithography, but is the hardest to manufacture. By contrast, the rightmost mask is the easiest to manufacture, but will not provide as good on wafer results as 6a and 6b. Finally, we show a simulated image of the wafer as printed by the Manhattan mask 6b:

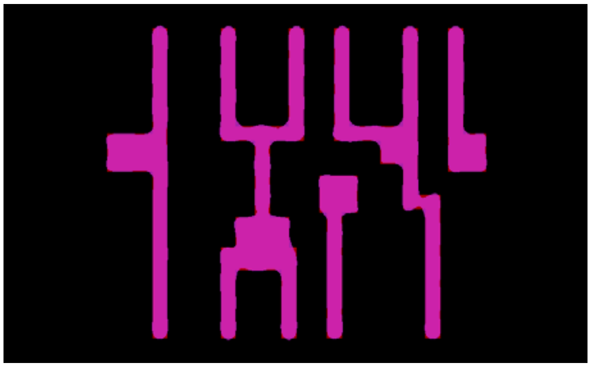

**Figure 7. XOR Simulated Wafer Pattern.**

In our last example, we consider the following SRAF design:

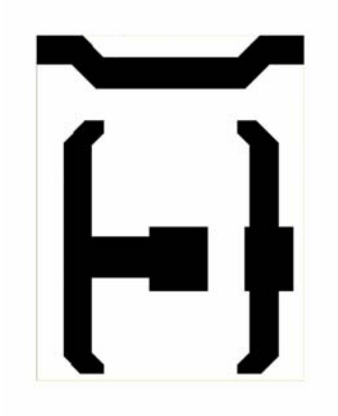

**Figure 8. SRAM Target Pattern.**

The continuous tone mask for this pattern appears as follows:

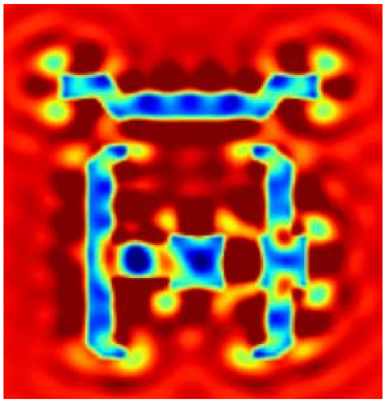

**Figure 9. Continuous tone SRAM mask.**

We can also find binary masks for the SRAM pattern, as follows:

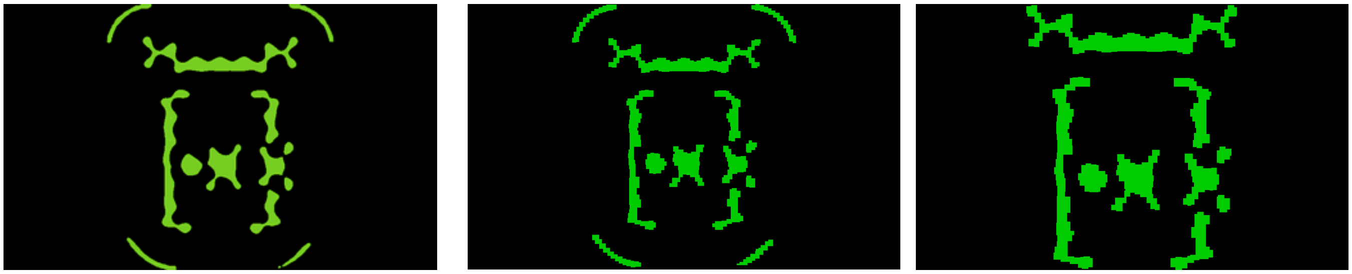

**Figure 10a,b,c. Binary masks for SRAM pattern.**

As with the XOR example, we can find masks with or without SRAFs, and with or without Manhattan constraints. The wafer pattern, as printed by the binary Manhattan mask, is as follows:

**Figure 11. SRAM Simulated wafer pattern.**

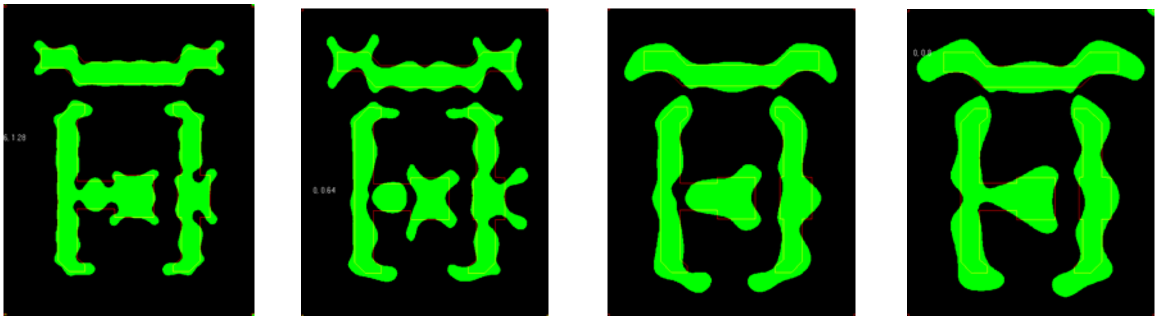

**Figure 12. SRAM mask pattern for several critical dimensions.**

It is also quite interesting to see how the optimal binary SRAM pattern changes with critical dimension. Figure 12 shows masks found with ILT with 130, 90, 65, and 55nm critical dimensions. As can be observed from the figure, each pattern is quite different from the others. However, without any prior knowledge of the optimal shape, or even of the topology or number of distinct chrome regions, ILT is capable of finding optimal solutions at each node.

Lastly, we show one example, as printed in silicon. Notice that, although the left part of the target pattern is separated into three distinct pieces on the photomask, it still prints the original target pattern quite accurately.

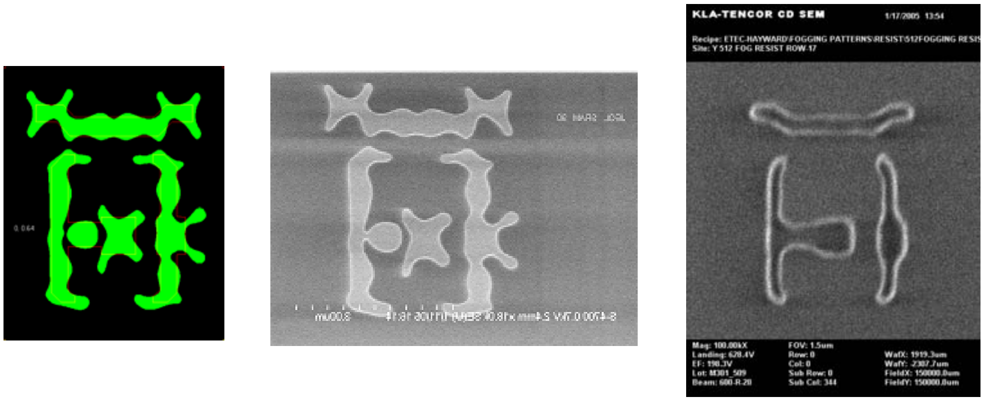

**Figure 13. SRAM pattern as GDS, on mask, and on wafer.**

## SUMMARY AND CONCLUSIONS

In summary, we have shown how ILT can find optimal photomask patterns in regions of solution space that may be very different from the original pattern. We have demonstrated a variety of results using ILT and shown mask manufacturing constraints incorporated into the solution. We have also described how modern ILT can scale to full chip and run fast. Our conclusion is that ILT is a viable technology for production manufacturing.

## ACKNOWLEDGEMENT

The authors would like to thank the whole Luminescent team for creating the ILT implementation that made this work possible. The authors would especially like to thank Grace Dai and Gordon Russell for their help in preparing ILT mask patterns for this presentation.

## REFERENCES

1. B.E.A. Saleh and S.I. Sayegh, “Reductions of errors of microphotographic reproductions by optical corrections of original masks”, Optical Eng. Vol. 20 pp 781-784 (1981)
2. K.M. Nashold and B.E.A. Saleh, “Image construction through diffraction-limited high-contrast imaging systems: an iterative approach”, J. Opt. Soc. Am.A, vol. 2 p. 635 (1985)
3. Y. Liu and A. Zachor, “Optimal binary image design for optical lithography”, Proc. SPIE Vol. 1264 pp 410-412 (1990)
4. Y. Liu and A. Zachor, “Binary and phase-shifting image design for optical lithography”, Proc. SPIE Vol. 1463 pp 382-399 (1991)
5. A. Rosenbluth et. al, “Optimum mask and source patterns to print a given shape”, JM3 vol. 1 pp 13-30 (2002)
6. Y-T Wang, Y.C. Pati, H. Watanabe and T. Kailath, “Automated design of halftoned double-exposure phase-shifting masks”, Proc. SPIE Vol. 2440 pp 290-301 (1995)
7. Y.H. Oh, and J-C Lee, Resolution enhancement through optical proximity correction and stepper parameter optimization for 0.12-um mask pattern, Proc. SPIE Vol. 3679 pp 607-613 (1999)
8. T. Fuhner and A. Erdmann, “Improved mask and source representations for automatic optimization of lithographic process conditions using a genetic algorithm”, Proc. SPIE Vol. 5754 pp 415-426 (2005)