Advances in Inverse Lithography

Thomas Cecil*[a] , Danping Peng[b] , Daniel Abrams[c] , Stanley J. Osher[d] , Eli Yablonovitch[e]

[a]Synopsys, 690 East Middlefield Road, Mountain View, CA 94043, United States

[b]TSMC, 2851 Junction Ave, San Jose, CA 95134, United States

[c]344 Addison Ave, Palo Alto, CA, 94301, United States

[d]UCLA Mathematics Dept. 520 Portola Plaza, Box 951555, Los Angeles, CA 90095, United States

[e]UC Berkeley EECS Dept., 253 Cory Hall, Berkeley, CA 94720-1770, United States

*Email: tcecil@synopsys.com

Abstract:

As Moore's law has marched forward, progressively shrinking chip designs at a consistent pace, the manufacturing of those chips has raced to keep up. Through advances in lithography hardware and software, we now are in the realm of the single-digit-nanometer design nodes at leading chip foundries. This paper will review the method of Inverse Lithography Technology, (ILT) which is the preferred computational method for photo-lithography. Indeed photo-lithographic masks were the first meta-surfaces, and still the most important meta-surface, economically, used in memory chips, storage, and microprocessors. Moreover, photolithographic mask designers were the early adopters of the mathematical optimization in optics, creating ILT, specifying intended wafer patterns and metrics, and using of the mathematical machinery of level sets, functional derivatives, and adjoints to drive the mask design process.

Keywords:

inverse problems, level set methods, lithography, mask synthesis, computational lithography, resolution enhancement technologies

Section 1: Introduction and Background on Mask Synthesis

Section 1.1: Historical Background

The past decades have seen a revolution in electromagnetics with a continuing search for interesting and important meta-materials [1,2] and photonic crystals[3,4] . The design of these electromagnetic structures has largely been guided by deep individual intuition, or by trial-and-error, like the original photonic bandgap materials[5].

Such an intuitive design approach is now being superseded by systematic optimization of the shapes of the interesting electromagnetic structures that achieve a desired goal. For example, the widest possible bandgap for the least index contrast could be a goal in photonic bandgap design[6]. In silicon photonics, insertion loss is often paramount[7]. Nowadays, these figures-of-merit can be plugged into a

computational optimization engine, to produce designs that appear very imaginative[8], even to the most skilled human designers.

Electromagnetics and especially optics are very late in adopting the new optimization mathematics. Other fields, such as linear programming, control theory, mechanical design, and artificial intelligence, adopted modern optimization techniques which were not yet playing a role in optics. These inverse methods can co-optimize thousands or millions of variables simultaneously making them ideal for shape optimization. The mathematics includes adjoints, duality, the level set methodology, etc. The key point is that the optimization mathematics represents a merger of calculus and linear algebra that could and should be taught to undergraduates.

Formal mathematical optimization is now being adopted in the electromagnetic meta-material field. But few of those practitioners are aware that the original meta-surfaces were actually the photolithographic masks that are essential in semiconductor manufacturing. Photolithographic masks are usually thin-film patterned Chromium metal, on glass, being in effect a meta-surface. Lithographic masks are the largest and most important practical application of meta-surfaces, propping up the giant electronics industry. Indeed photo-lithography has the earliest adoption of sophisticated mathematical shape optimization in optics, having been demonstrated for industry in 2000.

The name given to this form of sophisticated optimization varies by field, sometimes called adjoint optimization, but in artificial intelligence the same mathematics is called back-propagation. We have the name Inverse Electromagnetic Design, but it is also called Inverse Lithography Technology (ILT) in photomasks.

ILT was first explored by B. E. A. Saleh and others at the University of Wisconsin-Madison. For example, in 1981 Saleh and Sayegh[9] found optimized photomasks by "pixel flipping", a variation of steepest descent optimization. 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[10] 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[11] 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[12], allowing a randomly chosen mask object to change its position, or to expand/contract on a certain edge, 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[13]. 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.[14] at Stanford, and later Numerical Technologies, and the OPERA program, by Oh et. al. at

Wonkwang University in South Korea[15]. Most recently, Fuhner and Erdmann of the Fraunhofer Institute developed ILT using genetic algorithms[16]. 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.

Section 1.2: Introduction to Lithography

We begin with a short description of the lithography process used today for advanced chip fabrication[17]. Figure 1 shows a simplified schematic of the flow from wafer coating to light exposure to resist development and etching. The algorithmic design of the photomask, also known as the mask synthesis, is a critical step where geometric operations and optimizations take place. For commonly used binary masks made from chrome and glass, for example, the mask synthesis problem will output a design consisting of polygons representing either the chrome or glass regions of the mask. While the various coating, resist, and etch steps are highly tuned physical processes, they are not changing in terms of their response to a new design being synthesized, it is mainly the mask design which is optimized dynamically given a new chip.

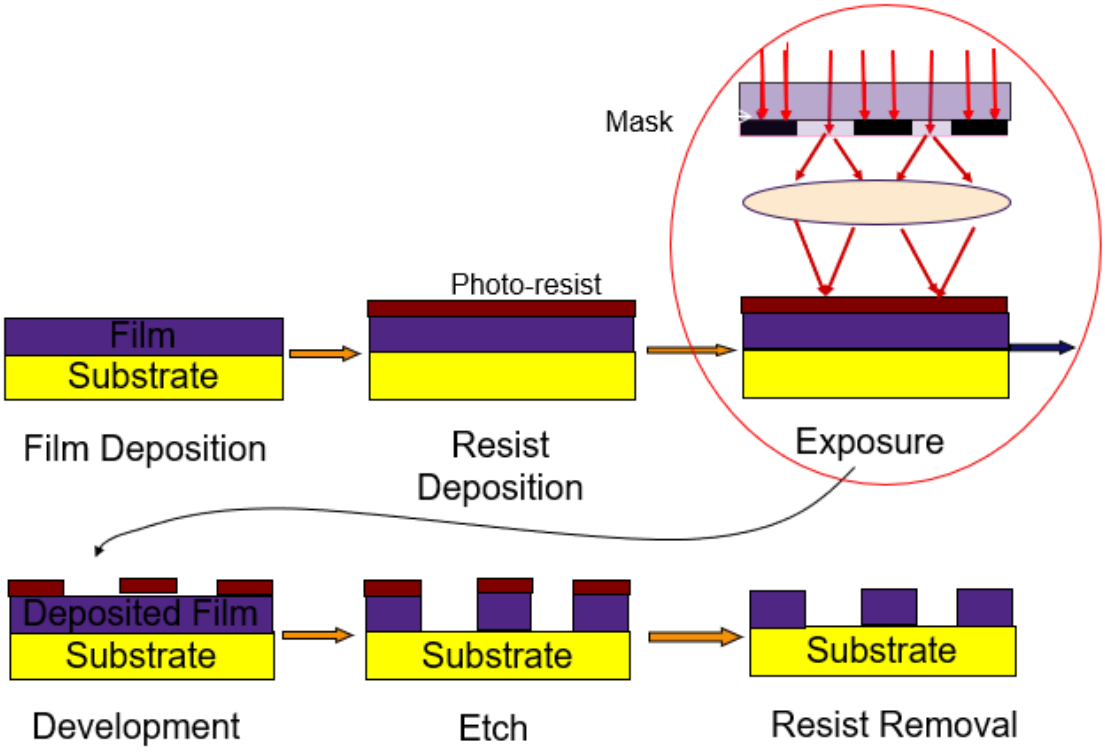


Figure 1: The lithography process, highlighting in the red circle the role of the mask as a key geometrical patterning step. Light sensitive photoresist is applied to the wafer and the mask patterns and imaging system determine which parts of the resist are removed.

The manufacturing challenge using lithography has a critical measure known as the k_1 factor, which is defined by the Rayleigh's equation,

$$R = k_1 \frac{\lambda}{NA},$$

(1)

where R is the half-pitch of the tightest one-dimensional line-and-space pattern which can be resolved by an imaging system, and where λ is the light wavelength and NA is the numerical aperture. To improve resolution, we have three knobs at our disposal: decrease wavelength λ, increase numerical aperture NA, or shrink the k_1 factor. The first two knobs involve expensive hardware improvement. During the 1990s-2000s, the wavelength of the imaging tools (called scanner) has progressed into the DUV regime, from λ=365nm to 248nm to 193nm, while the NA of a scanner has been getting larger, until it reached 1.35 with 193-immersion lithography at about 2005. Further reductions in wavelength were met with considerable technical and financial difficulties. After many false hopes and delays, EUV lithography using light with a wavelength of 13.5nm was introduced into high volume manufacturing in 2017, but the at the cost of decreasing the NA to 0.33. Active work is ongoing to increase the NA to 0.55 in the high-NA EUV scanner.

The third knob, the k_1 factor, has a theoretical limit of 0.25, but image contrast gets progressively worse as k_1 approaches this limit. Practical considerations limit k_1 to a higher value. As the k_1 factor was being reduced closer to its theoretical limit, more aggressive resolution-enhancement-techniques (RETs)[18] are demanded. In addition, as the k1 factor was decreased, it became ever more important to be able to control the process window (PW) to tighter and tighter tolerances, as any deviation from the optimally tuned configuration of the system could quickly reduce yields.

On the physical system side there were various RETs introduced to squeeze out as much improvement as possible such as: source mask optimization (SMO) to enhance patterning of dense arrays by exploiting off axis illumination, use of water immersion lithography to increase the NA, and enhancements to chemical resists on the wafer. The combined use of off axis illumination which induces stronger, longer range distortions in the imaging systems and the increased focus on PW improvements drove mask designs to start taking unintuitive shapes. Optical proximity correction (OPC), which was limited to perturbing the edges of the design pattern to produce a mask design could not correct for these diffraction effects thoroughly enough, as the correction demanded sub-resolution assist features (SRAFs) to produce the desired wafer patterns. SRAFs are mask patterns which have no direct corresponding wafer pattern to be printed by a particular SRAF polygon, but rather assist other main feature polygons in enhancing the PW of printed design patterns by increasing the exposure and focus ranges, for example, for which the patterns print within specified accuracy tolerances.

Inverse lithography technology (ILT) is a method of solving the mask synthesis problem by setting the problem up as an inverse problem, where the observed metrics on the wafer of imaging precision, robustness, and PW are measured, and drive an inverse solver algorithm to produce a mask which optimizes the metrics on the wafer. Figure 2 shows the progression in mask synthesis from design perturbations done in OPC, to a curvilinear ILT solution with separate main and assist feature polygons.

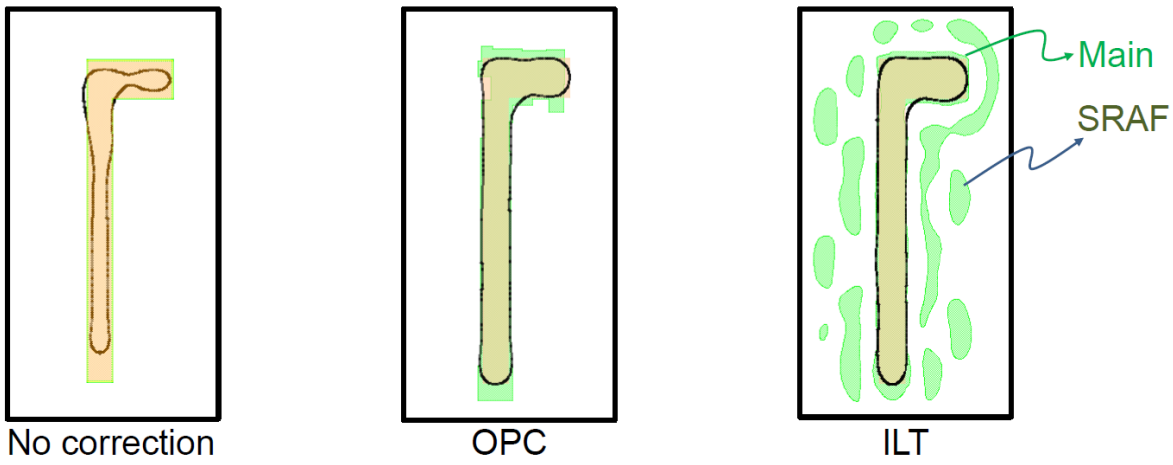


Figure 2: The effects of OPC, SRAFs, and ILT on lithography printability. Design: tan, mask: green, image contour: black.

In this paper we will discuss the inverse problem framework in Section 2: Inverse Problem Framework. Then in Section 3: Practical Considerations for manufacturing, we discuss manufacturing implications on

runtime and mask complexity. In Section 4: Newer computational tools, we explore some of the more recent software and hardware updates which can be applied to mask synthesis.

Section 2: Inverse Problem Framework for mask synthesis

Section 2.1 Level-Set Methods

As Moore's law was steadily shrinking chips, enabling more power for personal computers, one beneficiary of this was computational mathematics. In the 1990s this meant that physical simulation methods could be readily developed on off the shelf computational platforms, allowing for researchers to make rapid progress in developing new methods for real problems. One class of ideas which has been very successful in the realm of interface tracking are known as level-set methods[19,20]. Popularized beginning in the late 1980s, level-set methods treat the interface tracking problem by embedding the interface in a higher dimensional surface. This embedding enabled traditionally difficult geometric topology change operations such as splitting, merging, creation, and deletion to be handled without user algorithmic or emotional involvement, requiring no extra coding techniques.

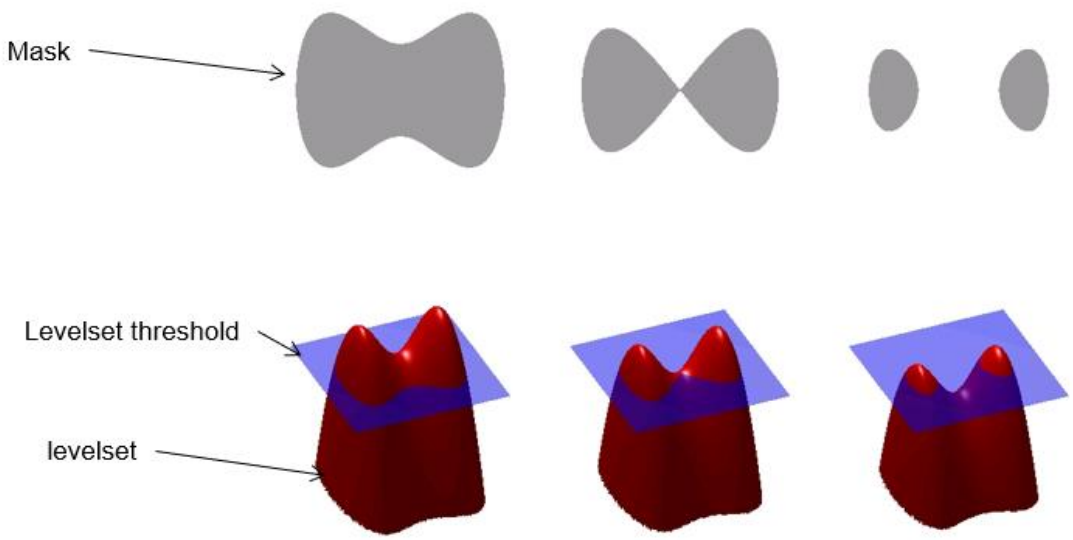


Figure 3: Level-set representation of a mask showing topology change.

Figure 3 shows an example of a mask polygon interface in multiple topological states including a critical singular formation where - two polygons are connected by a single point. The evolution of the interface for level set methods is changed from a set of polygon vertices being moved in (x,y) space to a higher dimensional surface being moved up and down in z space at each (x,y) location. The final interface is extracted by thresholding the level set.

One approach to model the motion of an interface is track the positions x(t), y(t) of particles on the interface as a function of time t. Mathematically, we can write this as

$$u(t) = \dot{x}(t), v(t) = \dot{y}(t)$$

where u(t) and v(t) are the instantaneous velocities of the interface for a given particle, and $\dot{x}(t)$, and $\dot{y}(t)$ denote time derivatives of x(t) and y(t).

In the level set approach, we instead track the interface with the use of a level set function $\varphi(x, y, t)$. Here, $\varphi > 0$ corresponds to one material, $\varphi < 0$ corresponds to the second material, and $\varphi = 0$ denotes the interface between the two materials. The evolution of φ over time is governed by the PDE

$$\frac{\partial \varphi}{\partial t} + u\frac{\partial \varphi}{\partial x} + v\frac{\partial \varphi}{\partial y} = 0.$$

(2)

where (u(t), v(t)) is the interface velocity as before. Thus, rather than tracking a large set of ODEs for each particle on the interface, we reduce the problem to a single PDE that governs the the level-set function phi.

If we want to study the motion in the normal direction to the interface we can obtain the Hamilton-Jacobi (HJ) PDE equation for the motion of φ,

$$\frac{\partial \varphi}{\partial t} + u_n|\nabla\varphi| = 0, where\ u_n = \left(\dot{x}(t), \dot{y}(t)\right) \cdot \nabla\varphi/|\nabla\varphi|.$$

(3)

This nonlinear HJ PDE can be solved adapting numerical methods originally developed for conservation laws in physics, as it has been shown that there is a close connection between these two types of equations[21,22].

Section 2.2 Inverse Problem Formulation

Pioneering work was done for the mask synthesis inverse problem over several years starting in the early 1980s[9.10.11]. This was extended by industry practitioners using various methods including pixelated mask representations[23,24,16,13,25,26] . At Luminescent Technologies, and later Synopsys, the method described below utilizing the level-set framework and cost function gradients derived using functional analysis and optimization was patented and pursued for production level deployment[27,28,29,30] . For a historical overview of the ILT field see[31]. These methods have had successful application beyond the lithography space in other related optical areas[32,33].

Given the level-set framework for interface tracking, we can now discuss the inverse problem formulation of the mask synthesis problem. The main goal will be to formulate a cost function, C, which measures the quality of printing on the wafer given a specific mask, represented as a level-set function, φ. Once this is done the method will be to find the functional gradient of C with respect to φ. Then we can use a wide range of gradient based optimization solvers to optimize C.

$$C = \iint (F(\varphi)(x, y) - T(x, y))^2 dxdy,$$

(4)

In equation (4) we describe an example cost function mathematically as the integral over the wafer plane in x,y of the squared difference between the image signal, F, of a given mask level-set, φ, and a target design function, T. The construction of the target, T, will depend on the user's knowledge of critical features on the wafer, as well as the level of processing we include in the forward simulation function, F. An example function F can be defined as

$$F = |\Pi\mathcal{M}(\varphi)|^2 = |\eta|^2 ,$$

(5)

where $\Pi = \mathcal{F}^{-1}\wp\mathcal{F}$ is the projection operator, $\mathcal{F}$ is a Fourier transform, $\wp$ is a pupil function, and $\mathcal{M}(\varphi)$ is a masking function. In practice we will use more accurate approximations to the physical phenomena such as resist development, R, where R=S(F-α), where S is a sigmoid function and α is a development threshold level.

The pupil cutoff function in Fourier space can be defined for this example as

$$\wp(k_x, k_y) = \begin{cases} 0, k_x^2 + k_y^2 \geq k_{max}^2 \\ 1, k_x^2 + k_y^2 \leq k_{max}^2 \end{cases}.$$

(6)

Meanwhile the masking function for a chrome and glass mask is a Heaviside function defined by

$$\mathcal{M}(\varphi) = \begin{cases} 1, \varphi \geq 0 \\ 0, \varphi < 0 \end{cases}.$$

(7)

In equation (5) we refer to Π as the projection operator, which can be thought of as a convolutional operator, and $\eta = \Pi\mathcal{M}(\varphi)$ as the electric field distribution on the wafer.

The functional derivative of C with respect to φ is found by using variational calculus[34], exploiting the adjoint of the projection operator, which we write as Π^*.

$$\frac{\delta C}{\delta \varphi} = \Pi^* \left[2\left(F(\varphi)(x,y) - T(x,y)\right)\delta(\varphi)\eta\right],$$

(8)

where $\delta(\varphi)$ is a Dirac delta function. We can then use the gradient, $\frac{\delta C}{\delta \varphi}$, in various gradient based optimization methods to minimize C, resulting in a mask interface that can be extracted from φ by contouring. The entire optimization procedure including some parts of the forward imaging is described in Figure 4.

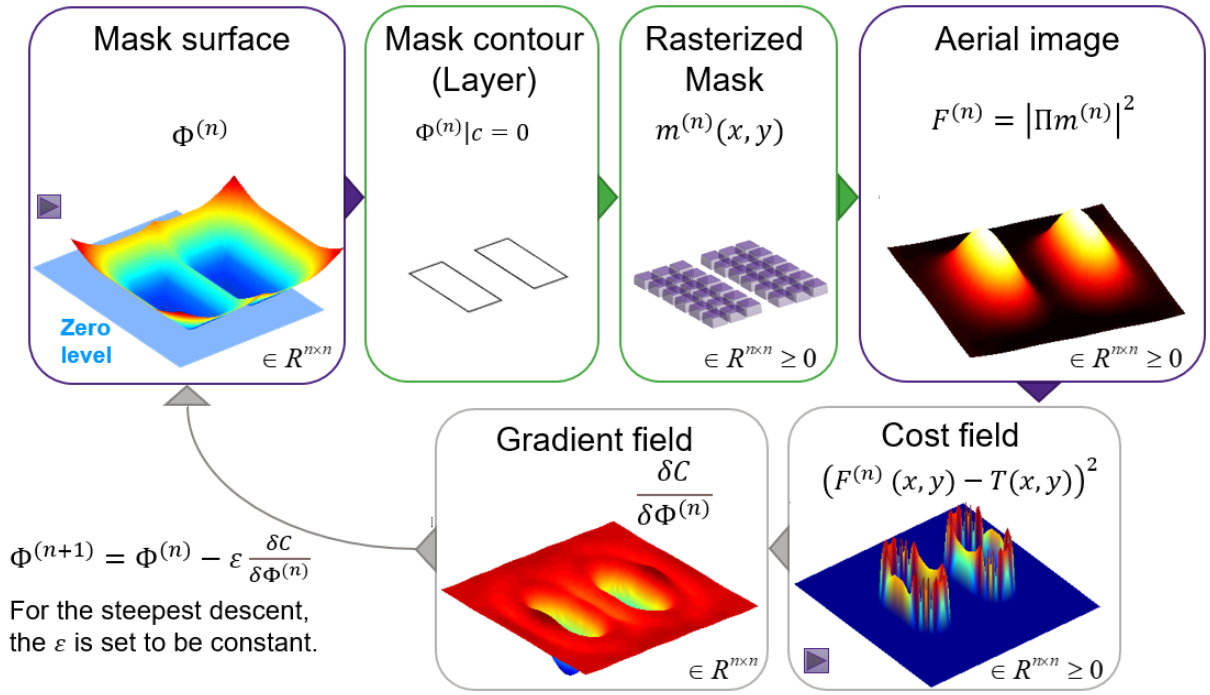


Figure 4: Forward imaging, cost function, and gradient computations in an iterative gradient descent solver for ILT.

We should note here that the approximations of the imaging system described in equations (4) -(7) are in practice more complicated, but the main method of deriving the gradient is similar. As with the methods used by machine learning (ML) neural network backpropagation algorithms, if the forward operator, F, is composed of differentiable functions, then exploiting their adjoints allows the calculus of variations to automatically derive the functional derivative of the C with respect to φ. Thus, the inverse problem formulation is well suited for other optical design problems and can be extended in a straightforward way to non-binary designs where φ can be a transmission map or other design space free variables. For example, the same method has been used for the SMO problem described above where the source map is computed using adjoint and gradient descent optimization methods as shown in Figure 5[35].

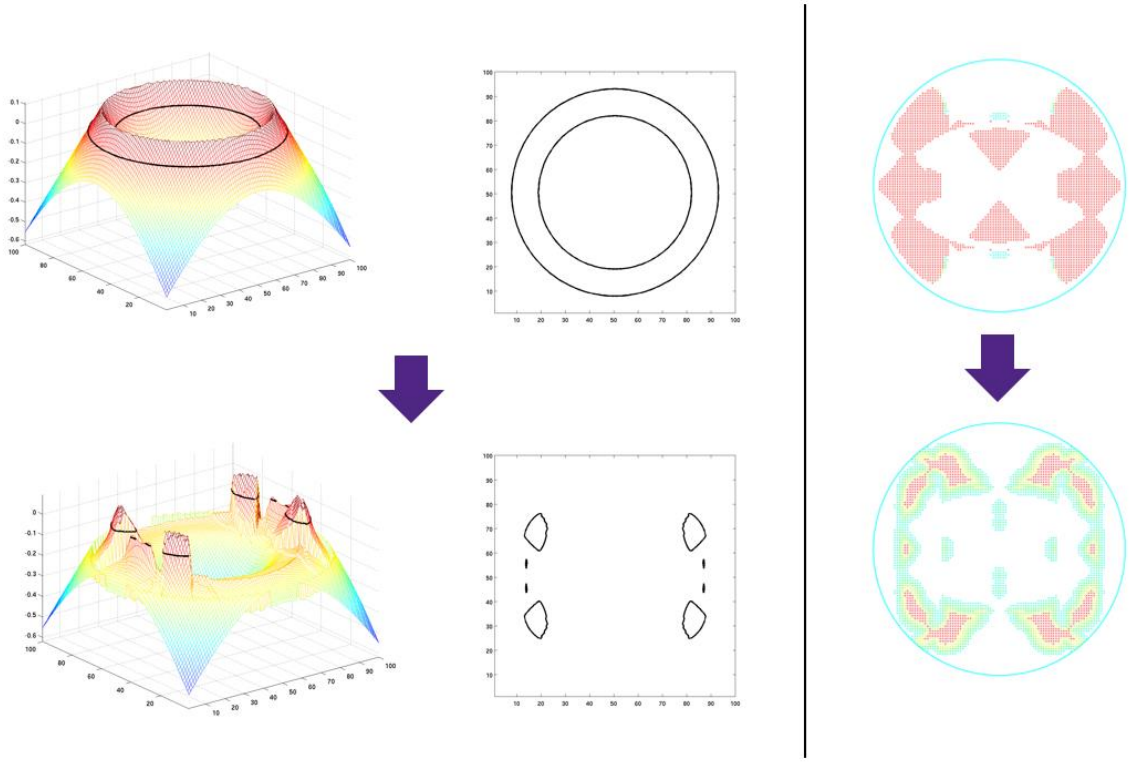


Figure 5: SMO source map results. Left: Level-set representation of annular initial source (top) and optimized binary source (bottom). Right: initial continuous transmission source map (top) and optimized source map (bottom).

Section 3: Practical Considerations for manufacturing

There are several practical considerations to be to be considered in the mask synthesis problem space. In this section we will cover some of the main areas including runtime constraints and manufacturability constraints.

Section 3.1 Runtime Constraints

A number of increasingly sophisticated methods for lithographic mask synthesis have been developed over the past few decades. Listed in order of increasing sophistication, they are:

1. Masks being the same as the wafer design, tolerating any lithographic process distortions
2. Rule based OPC: Masks being adjusted by geometric rules, relying on the user to write rule tables for geometric perturbations of the design to make the mask
3. Model based OPC: Masks being adjusted by model-based algorithms, perturbing input design geometry through simulation feedback
4. ILT: Masks being created by model-based algorithms, with input design geometry mainly providing wafer metrics but not mask degrees of freedom

Advancing from each step of this sequence to the next was somewhat of a paradigm shift and required significant qualification of new software algorithms. We discuss further extensions to this table in

Section 4. One of the main progressions in the table is the increased reliance on computational models, moving away from rules and controls that had proven their worth on previous chip tape outs. As computational complexity increased, so did runtime. This was mitigated somewhat by increased CPU processing speeds, but smaller designs demanded higher accuracy models, which grew larger as they attempted to capture all physical processes that could be measured. As each level in the table was realized in production, it became harder for the previous level to compete. Here are a few reasons:

1. Increased computational reliance replaced manual work such as rule table creation and design perturbation initialization/setup, thus not only was expertise in previous steps more difficult as device sizes decreased, but it was also not cultivated as widely.
2. The degrees of freedom for the mask increased as the time progressed, giving later methods a larger solution space.

While the runtime level for each subsequent step became somewhat normalized as the solution quality was recognized, there was still pressure to keep the speed as similar as possible to the previous level. The result of this pressure was to introduce ILT in hybrid solutions, mixing old and new technologies. Given the limited computational resources of a single processing unit, it is necessary to parallelize the processing of a full chip design during mask synthesis. Current parallelization can range over tens of thousands of CPU cores, working on millions of individual clips (known as templates) together representing a full design. ILT was introduced in localized regions, referred to as hot spots, which had the most difficult patterns[36,37]. These constituted a small percentage of the entire design, thus limiting the impact on throughput. There were also critical periodic memory arrays which were solved with ILT as the repeated nature of these patterns lent itself to high ratios of mask area versus correction runtime.

To smooth the transition from OPC to ILT, at the single template solver level there were several modifications to the ILT solutions as they were originally formulated. Here is a list of some of the critical migrations:

1. Modification of ILT cost functions to resemble OPC metrics of image edge-placement-error (EPE)
2. Manhattanization of ILT curvilinear masks to meet mask writer requirements
3. Improved algorithmic consistency to handle pixel-based aliasing that was less pronounced in OPC
4. Mask rule checking and correcting for level-set and curvilinear masks

All these steps are difficult problems which differentiate production level implementations from toy solutions. The semiconductor fabs have very little error budget allocated for any mask inconsistencies and deviations introduced by the mask synthesis software tools, as the noise in the physical fab processes consume most of the total error budget in manufacturing. These constraints lead us to the discussion in the next subsection on manufacturability.

Section 3.2 Manufacturability Constraints

As demonstrated in Figure 6, process windows from ILT and curvilinear masks were superior to those produced by OPC with Manhattan masks. However, there were numerous manufacturing issues to be confronted for the acceptance of ILT as a mainstream production mask synthesis solution[38,39].

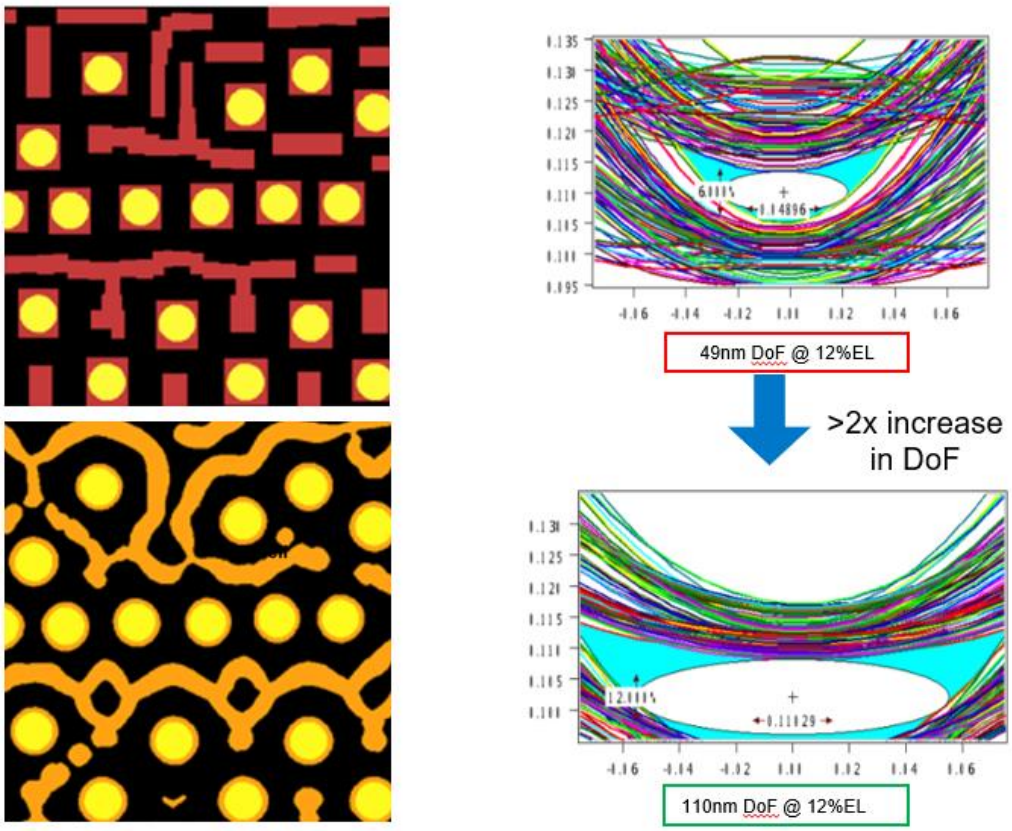


Figure 6: Manhattan mask example pattern (top left) and corresponding common process window (top right). Process window is the largest ellipse which can fit in the common acceptably printing area of a set of designs with varying process conditions in depth of focus (x-axis) and exposure latitude (y-axis). Curvilinear mask example pattern (bottom left) and corresponding process window (bottom right) showing increase in depth of focus for a given minimum exposure latitude.

The first issue was mask complexity. Existing mask writing tools were based on Manhattan geometry and fractured the mask polygons into rectangles so that the mask writing tools with shaped beams restrictions could expose them. As shown in Figure 7, even with manhattanization of curvilinear shapes, the fracture/shot count was complex, as fracturing tools were designed for OPC masks which were perturbations of Manhattan designs, not ILT masks which had less correspondence to the design geometry.

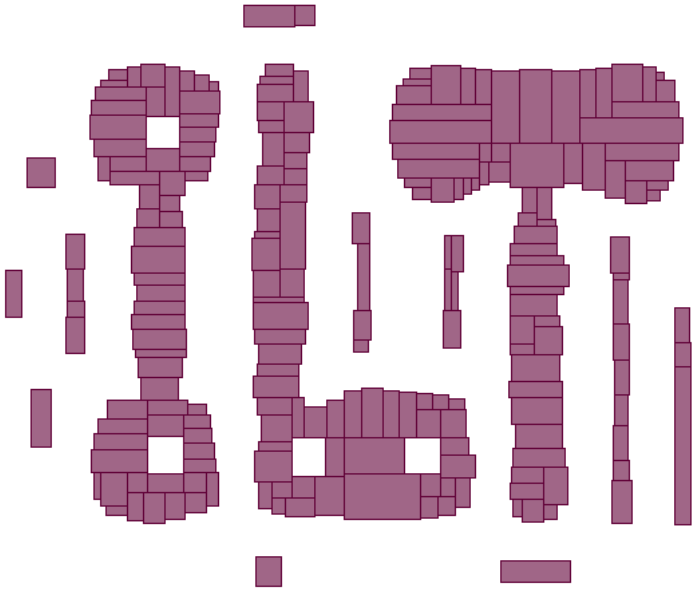


Figure 7: Fracturing of a manhattanized ILT mask.

As time progressed the mask writing industry made progress to the infrastructure to support curvilinear masks, including the introduction of multi-beam mask writing (MBMW) tools[40] which work in a more pixelated fashion that the traditional variable shaped beam (VSB) tools.

Within the mask tape out tool chain flow and within mask synthesis tools themselves, the data complexity is burdensome, and thus new mask representations were developed for ILT. Not only are the pixel transmission and level-set interface tracking methods used, but also lower dimensional representations of the mask, namely: curvilinear polygons and splines in 2d, and skeletons in 1d as shown in Figure 8 [41]. The multi-dimensional representations aid in both data volume as well as algorithmic complexity, and mirror the progress made in other level-set simulation areas by exploiting lower dimensional objects such as particles.

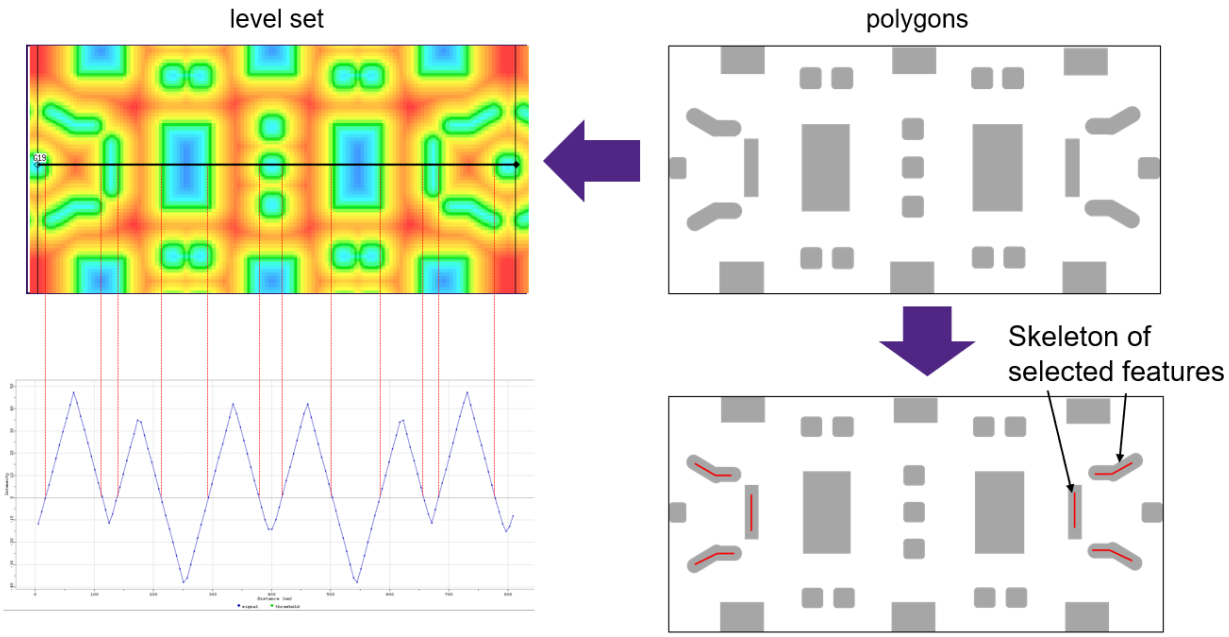


Figure 8: Various mask representations in 3d: level-sets, 2d: curvilinear polygon and spline, and 1d: skeletons.

One open question for the curvilinear mask tape out flow is: what are the specifications for mask manufacturability rule checks (MRC)? Traditional Manhattan masks had well defined rule table specifying width, space, corner to corner, aspect ratio, etc. checks which produced well defined algorithms to compute them. For curvilinear masks there are questions of: what is a corner, or where to measure a width[42,43]? It is possible that as these check definitions become more mature and accepted across the industry there will be more opportunity to define them based on physical manufacturing limitations which can potentially be modeled and penalized by geometric measurements which are well defined by various level-set differential operators. The optimization framework of ILT is in a good position to incorporate these checks, and is already doing so in practice, even though the commonly agreed upon definitions are still being debated at large.

A final area to be discussed is the consistency of solutions. As mentioned previously, a full design can be parallelized to millions of templates. Although the goal of the parallelization is to process any unique design geometry clip only once, the reality is that some small regions will be repeated in multiple unique templates, and thus solved numerous times. A critical challenge for ILT is to maintain consistency of the mask for all the solutions solved in parallel. This is an open area of study and often ignored in small scale image processing and optimization research as it rears its head only when many templates are solved independently and stitched back together, revealing boundary mismatches. The periodic array solving approach[44] is able to recognize repeated designs and guarantee consistency of them, and other pattern matching solutions have been developed to attempt to identify all unique placements of all localized design variations within a full chip, but a robust, parallel-consistent, gradient based optimization algorithm is something that would be welcome in the mask synthesis community.

Section 4: Newer computational tools

At the beginning of section 3.1 we introduced a list of steps in the evolution of lithographic mask synthesis. A fifth step could have been listed as: 5. Machine learning: Masks being created by ML models trained on ILT masks. This next shift in mask synthesis is underway today, both in incremental steps: as ML sub-models of portions of the simulation system trained on more expensive rigorous physical solvers; and as full replacements of ILT as deep ML models replicating the entire mask synthesis optimized solution[45,46,47,48]. Figure 9 shows a representation of the MLILT model flow. The acceptance of and migration to these ML models in addition to optimization methods is underway and there are still many open questions. For example, many ML training methods and model architectures have been designed for image classification problems, where probabilistic outputs are acceptable and errors in the low single digits are state of the art. For the mask synthesis problem, models must be much more robust, as even one error in a billion can cause a mask to fail, and masks have billions of polygons to correct. A clearer understanding of the following areas would be helpful to enhance user migration to MLILT solutions:

1. How can we know the extrapolation capability of ML models? When will they fail and how badly?
2. What are the most robust model architectures for ML mask synthesis problems?
3. How can we design training pattern sets to produce the best ML models?
4. Can ML models capture or retain physical information in a meaningful way?

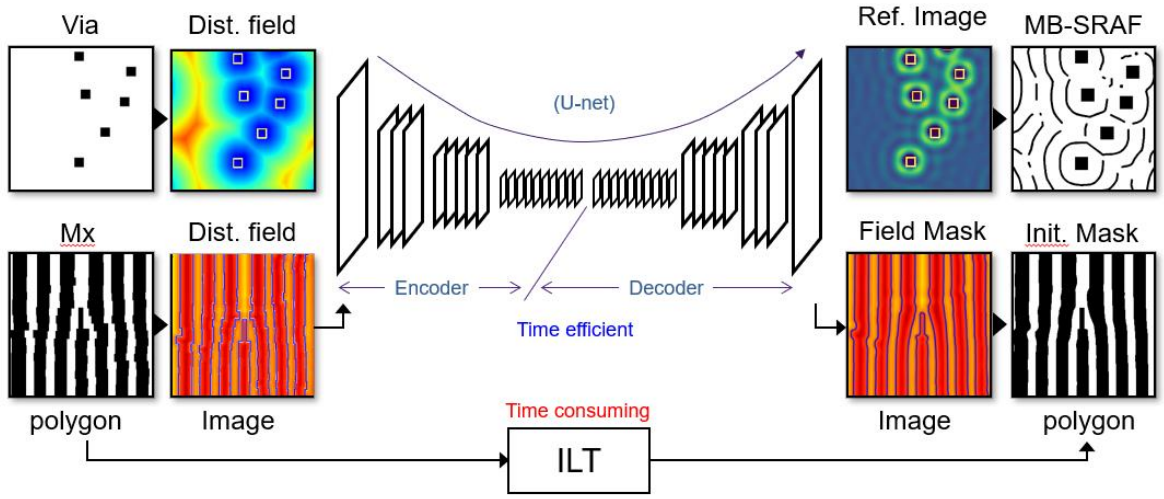


Figure 9: Representation of ML-ILT. A ML deep neural net model (U-Net) is trained on ILT design/mask pairs and can replicate the ILT mask synthesis optimized mask for examples of Via and metal (MX) layers of a chip.

One of the main drivers of ML deployment in the world is the continued advancement of GPU processing speed. The tensor operations in deep neural net models are dramatically improved on GPUs relative to CPUs. However, as noted above, the adjoint methods of the functional gradients for ILT are identical in many ways to the backpropagation methods used in ML training, and so there are many uses

within the non-ML ILT solvers for GPU acceleration. The optical system and approximations to other physical components in the simulation flow are often able to be modeled by convolutional type operators which can be accelerated on GPUs. The pixel-based nature of level-set method, including many localized geometric operations using it are well suited to GPU acceleration. Runtimes for level set based implementation of various geometric operations as compared to polygon operations performing the same task can be orders of magnitude faster when customized for GPU acceleration[49,50].

Section 5: Conclusion

In this paper we have attempted to outline the mask synthesis problem evolution, and the increasingly difficult manufacturing progression which led to the need for computationally intensive inverse solution methods. The introduction of mathematical level set techniques for interface tracking, and functional analysis tools for computing gradients of inverse problems led to the inverse lithography solution space becoming practical. These techniques have been introduced to other optical optimization areas over time as well. While there have been challenges over time to adopt the unintuitive and complex solutions produced by ILT, various hardware and software tools have been introduced which have increased manufacturing quality and decreased its run time. With the further introduction of machine learning and GPU processors, the level set inverse problem framework has gained even more advantages as it is readily adaptable to benefit from these acceleration tools. It will be interesting to follow the future trend of inverse problem solution methods in the optical field to see if their success can mirror that of ILT.

References

[1] J. Pendry and et al, "Controlling electromagnetic fields," *Science,* vol. 312, pp. 1780-1782, 2006.

[2] V. Galdi and et al, "Genetically Optimized Metasurface Pairs for Wideband Out-of-Phase," *IEEE ANTENNAS AND WIRELESS PROPAGATION LETTERS,* vol. 7, pp. 788-791 , 2008.

[3] E. Yablonovitch, "Inhibited Spontaneous Emission In Solid-State Physics And Electronics," *Physical Review Letters,* vol. 58, pp. 2059-2062, 1987.

[4] S. John, "Strong Localization Of Photons In Certain Disordered Dielectric," *Physical Review Letters,* vol. 58, pp. 2486-2489, 1987.

[5] E. Yablonovitch and et al, "Photonic Band-Structure - The Face-Centered-Cubic Case Employing," *Physical Review Letters,* vol. 67, pp. 2295-2298, 1991.

[6] C. Kao and et al, "Maximizing band gaps in two-dimensional photonic crystals by using level set methods," *APPLIED PHYSICS B-LASERS AND OPTICS,* vol. 81, pp. 235-244, 2005.

[7] A. Piggott and et al, "Inverse design and demonstration of a compact and broadband on-chip wavelength demultiplexer," *NATURE PHOTONICS,* vol. 9, p. 374, 2015.

[8] K. Yang and et al, "Inverse-designed non-reciprocal pulse router for chip-based LiDAR," *NATURE PHOTONICS,* vol. 14, p. 369, 2020.

[9] B. Saleh and S. Sayegh, "Reductions of errors of microphotographic reproductions by optical corrections of original masks," *Opt. Eng.,* vol. 20 (5), p. 781 –784, 1981.

[10] K. Nashold and B. Saleh, "Image construction through diffraction-limited high-contrast imaging systems: an iterative approach," *J. Opt. Soc. Am. A ,* vol. 2, 1985.

[11] Y. Liu and A. Zachor, "Optimal binary image design for optical lithography," *Proc. SPIE ,* vol. 1264, p. 410 –412, 1990.

[12] Y. Liu, A. Zakhor and M. Zuniga, "Computer Aided Phase Shift Mask Design with Reduced Complexity," *Trans. Semi. Manuf.,* vol. 9, no. 2, 1996.

[13] A. Rosenbluth and et al, "Optimum mask and source patterns to print a given shape," *Proc. SPIE,* vol. 4346 , p. 13 –30, 2002.

[14] Y. Wang and et al, "Automated design of halftoned double-exposure phase-shifting," *Proc. SPIE,* vol. 2440, pp. 290-301, 1995.

[15] Y. Oh and et al, "Resolution enhancement through optical proximity correction and stepper parameter," *Proc. SPIE,* vol. 3679, pp. 607-613, 1999.

[16] 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 , p. 415 –426 , 2005.

[17] F. Schellenberg, "Resolution enhancement technology: the past, the present, and extensions for the future," *Proc. SPIE 5377,* vol. Optical Microlithography XVII, 2004.

[18] A. Wong, Resolution Enhancement Techniques in Optical Lithography, Society of Photo Optical, 2001.

[19] S. Osher and J. Sethian, "Fronts propagating with curvature-dependent speed: algorithms based on Hamilton–Jacobi formulations," *J. Comput. Phys.,* vol. 79, pp. 12-49, 1988.

[20] S. Osher and R. Fedkiw, Level set methods and dynamic implicit surfaces, Springer Applied Mathematical Sciences 153, 2003.

[21] S. Osher and C. Shu, "High order essentially non-oscillatory schemes for Hamiton-Jacobi equations," *SIAM J. Numer. Anal. ,* vol. 28, pp. 902-921 , 1991.

[22] H. Zhao and et al, "A variational level set approach to multiphase motion," *J. Comput. Phys.,* vol. 127, pp. 179-195, 1996.

[23] Y. Borodovsky, "Pixelated phase mask as novel lithography RET," *Proc. SPIE,* vol. 6924 , 2008.

[24] Y. Granik and et al, "Fast pixel-based mask optimization for inverse lithography," *J J. Micro/Nanolithogr. MEMS MOEMS ,* vol. 5, no. 4, 2006.

[25] A. Poonawala and P. Milanfar, "OPC and PSM design using inverse lithography: a nonlinear optimization approach," *Proc. SPIE,* vol. 6154 , 2006.

[26] S. Jang and et al, "Manufacturability evaluation of model-based OPC masks," *Proc. SPIE,* vol. 4889, 2002.

[27] S. Abrams, S. Osher and D. Peng, "Method for time-evolving rectilinear contours representing photo masks". USA Patent 7124394, 2006.

[28] D. Abrams and L. Pang, "Fast inverse lithography technology," *Proc. SPIE ,* vol. 6154 , 2006.

[29] C. Hung and et al, "First 65 nm tape-out using inverse lithography technology (ILT)," *Proc. SPIE,* vol. 5992 , 2005.

[30] L. Pang and et al, "Laser and e-beam mask-to-silicon with inverse lithography technology (ILT)," *Proc. SPIE,* vol. 5992, 2005.

[31] L. Pang, "Inverse lithography technology: 30 years from concept to practical, full-chip reality.," *J. Micro/Nanopattern. Mats. Metro.,* vol. 20(3), 2021.

[32] M. Burger and et al, "Inverse problem techniques for the design of photonic crystals," *IEICE transactions on electronics,* vol. 87, no. 3, pp. 258-265, 2004.

[33] M. Burger and S. Osher, "A survey on level set methods for inverse problems and optimal design," *European journal of applied mathematics,* vol. 16, no. 2, pp. 263-301, 2005.

[34] I. Gelfand and S. Fomin, Calculus of Variations, Dover Publications, 2000.

[35] V. Tolani and et al, "Source-mask co-optimization (SMO) using level set method," *Proc. SPIE,* vol. 7488 , 2009.

[36] W. Sim and et al, "Hotspot fixing using ILT," *Proc. SPIE ,* vol. 7973, 2011.

[37] Y. Liu and et al, "Technique for correcting hotspots in mask patterns and write patterns". USA Patent 8082525, 2011.

[38] P. Martin and et al, "Manufacturability study of masks created by inverse lithography technology," *Proc. SPIE ,* vol. 5992, 2005.

[39] B. Kim and et al, "Trade-off between inverse lithography mask complexity and lithographic performance," *Proc. SPIE 7379, Photomask and Next-Generation Lithography Mask Technology XVI,* 2009.

[40] C. Klein and E. Platzgummer, "MBMW-101: world's 1st high-throughput multi-beam mask writer," *Proc. SPIE ,* vol. 9985, 2016.

[41] T. Cecil, "Rule based assist feature placement using skeletons". USA Patent 10262100, 2019.

[42] I. Bork and et al, "MRC for curvilinear mask shapes," *Proc. SPIE 11518,* vol. 11518, 2020.

[43] T. Cecil, "Mask rule checking for curvilinear masks for electronic circuits". USA Patent 11222160, 2022.

[44] G. Xiao and et al, "Hybrid inverse lithography techniques for advanced hierarchical memories," *Proc. SPIE 9052, Optical Microlithography XXVII,* vol. 9052, 2014.

[45] N. Jia and E. Lam, "Machine learning for inverse lithography: using stochastic gradient descent for robust photomask synthesis," *J. Opt.,* vol. 12, no. 4, 2010.

[46] S. Wang and et al, "Machine learning assisted SRAF placement for full chip," *Proc. SPIE,* vol. 10451, 2017.

[47] P. Liu, "Mask synthesis using machine learning software and hardware platforms," *Proc. SPIE 11327, Optical Microlithography XXXIII,* 2020.

[48] T. Cecil and et al, "Establishing fast, practical, full-chip ILT flows using machine learning," *Proc. SPIE 11327, Optical Microlithography XXXIII,* 2020.

[49] I. Torunoglu and et al, "A GPU-based full-chip inverse lithography solution for random patterns," *Proc. SPIE 7641, Design for Manufacturability through Design-Process Integration IV,* 2010.

[50] J. Zhang and et al, "GPU-accelerated inverse lithography technique," *Proc. SPIE 7379, Photomask and Next-Generation Lithography Mask Technology XVI,* 2009.

For Table of Contents Use Only

Advances in Inverse Lithography

for ACS Photonics

Authors:

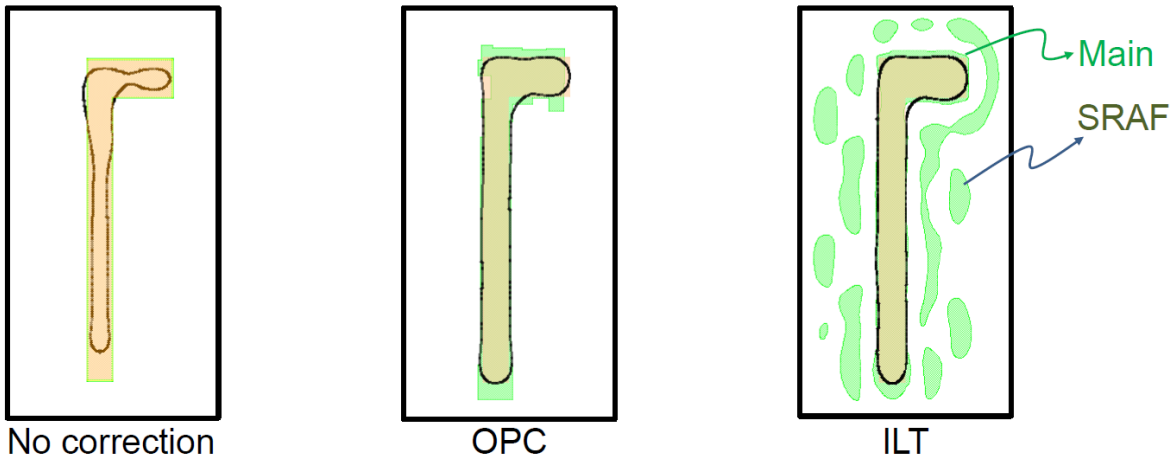
Thomas Cecil, Synopsys, tcecil@synopsys.com

Danping Peng, TSMC, DPENG@TSMC.COM

Daniel Abrams, daniel.xyzzy+123.ilt@gmail.com

Stanley Osher, UCLA Mathematics Dept., sjo@math.ucla.edu

Eli Yablonovitch, UC Berkeley EECS Dept., eliy@eecs.berkeley.edu



Inverse lithography mask synthesis.